Continental Shelf Research 31 (2011) 758–768

Contents lists available at ScienceDirect

Continental Shelf Research

journal homepage: www.elsevier.com/locate/csr

Research papers

A combined EOF/variational approach for mapping radar-derived sea surface currents

Max Yaremchuk [a,*], Alexei Sentchev [b]

[a] Naval Research Laboratory, Bldg. 1009, Stennis Space Center, MS 39529, USA
[b] Laboratoire d'Oceanographie et Geosciences (CNRS-UMR 8187), ULCO, 62930 Wimereux, France

ARTICLE INFO

Article history:
Received 11 July 2010
Received in revised form
13 January 2011
Accepted 24 January 2011
Available online 2 February 2011

Keywords:
High frequency radars
Surface currents
Variational interpolation

ABSTRACT

A technique for reconstruction of the 2d surface velocity field from radar observations is proposed. The method consecutively employs two processing techniques: At the first stage raw radial velocity data are subject to EOF analysis, which enables to fill gaps in observations and provides estimates of the noise level and integral parameters characterizing small-scale variability of the sea surface circulation. These parameters are utilized at the second stage, when the cost function for variational interpolation is constructed, and the updated radial velocities are interpolated on the regular grid.

Experiments with simulated and real data are used to assess the method's skill and compare it with the conventional 2d variational (2dVar) approach. It is shown that the proposed technique consistently improves performance of the 2dVar algorithm and becomes particularly effective when a radar stops operating for 1–2 days and/or a persistent gap emerges in spatial coverage of a basin by the HFR network.

Published by Elsevier Ltd.

1. Introduction

The technology of monitoring near-coastal currents by High Frequency Radars (HFRs) has been rapidly developing in the past decade. HFR observations are now extensively used to study near-shore circulation under a large variety of environmental conditions (e.g., Hisaki et al., 2001; Breivik and Sætra, 2001; Sentchev and Yaremchuk, 2007; Chavanne et al., 2007) helping to solve many applied problems in the coastal regions.

An obvious advantage of HFR observations is their availability in real time with nearly continuous temporal and spatial coverage of 10–15 min and 1–2 km, respectively. However, the back-scattered HFR signals suffer from to numerous distortions of artificial and natural origin. As a consequence, estimates of the along-beam sea surface velocities extracted from the Doppler shifts of the signals become unusable, resulting in numerous gaps in spatial coverage. These gaps may strongly degrade performance of the algorithms which extract the 2d sea surface velocity field from the HFR data (e.g., Kaplan and Lekien, 2007).

A natural way to fill these gaps is to take into account space-time correlations between the radial velocities. Assimilation of the HFR data into numerical models is the most straightforward and general approach, which has been under development in recent years (Lewis et al., 1998; Breivik and Sætra, 2001; Oke et al., 2002; Paduan and Shulman, 2004; Hoteit et al., 2009). The underlying idea is to combine dynamical constraints of a model with the history of dense spatio-temporal coverage by HFRs to produce the "best" estimate of the surface velocity at a given time. This approach, however, has a number of drawbacks hindering its implementation for real-time HFR data analysis: Beyond a relatively high computational cost, inverse numerical models have a large number of free parameters whose statistics is poorly known. The most problematic among those are the open boundary conditions, which are the major contributors to slow convergence of the HFR data assimilation schemes typically involving lengthy open boundaries.

An alternative more simple approach can be classified as the 2d objective analysis (OI) or spline interpolation (McIntosh, 1990): The corresponding least-squares algorithms differ from each other by specifying either the covariance function or its inverse. In application to HFR data the OI algorithms were implemented using empirical error covariances deduced from "normal modes" (Lipphardt et al., 2000), "open-boundary modes" (Kaplan and Lekien, 2007) and the data (Kim et al., 2007, 2008). A comparison of these methods was given by Yaremchuk and Sentchev (2009) who also proposed to augment the cost function with the terms penalizing grid-scale variability in the divergence and vorticity fields.

Although the 2d OI methods are computationally cheaper than the variational schemes involving dynamical information, they

* Corresponding author.
E-mail address: max.yaremchuk@nrlssc.navy.mil (M. Yaremchuk).

0278-4343/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.csr.2011.01.009

may perform poorly in the presence of large gaps in the data because information on the spatial structure of the velocity field within the gap is implicitly drawn from the idealized covariance function, which looses accuracy at large distances. A certain improvement of the covariance models can be obtained by considering their truncated expansions in the empirical orthogonal functions (EOFs), a technique successfully used in Kalman filtering (e.g., Tipett et al., 2003) and variational data assimilation (Fang et al., 2009; Yaremchuk et al., 2009).

The EOF-based estimates of the covariances rely upon time averaging and, therefore, may be successfully applied not only to model output but also to datasets with nearly continuous spatio-temporal coverage e.g., sea surface temperature (SST) or HFR data. Beckers and Rixen (2003) (hereinafter BR03) proposed an iterative EOF-based technique for filling gaps in the gridded SST images, which was successfully applied in Adriatic (Alvera-Azcarate et al., 2005). Kondrashov and Ghil (2006) developed the method further by including time correlations into the procedure under the assumption of statistical stationarity of the observed fields.

The goal of the present study is to design a HFR data interpolation algorithm capable of processing situations when *one or more radars are out of operation*. To do that we modify the BR03 method to make it suitable for processing HFR observations and combine it with the 2dVar technique. Large gaps in HFR data are filled using a truncated EOF decomposition of the radial velocity covariance matrix. Spatial correlations between the radial velocities are also used to estimate observational noise, assess its variance, and quantify the grid-scale variability of the velocity field. These parameters are inferred directly from observations and used to define the cost function weights for 2dVar mapping of the HFR data onto the regular grid.

The paper is organized as follows. In the next section we briefly describe the methodology of 2dVar interpolation and estimation of the error covariance via truncated EOF expansion. In the same section we also describe optimization of the truncation number and computation of the cost function weights. In Section 3 the method is verified using twin experiments with the HFR data simulated by a numerical model in a real domain (Monterrey Bay). Section 4 describes the results of experiments with the real observations off the Opal Coast of the Eastern English Channel. It is shown that the proposed algorithm significantly improves the accuracy of interpolation within the gaps typical for HFR observations, including the important case of a single operating radar. Conclusions and discussion of further development of the method complete the paper.

2. Methodology

The technique described here employs advanced statistical and variational methods to improve the accuracy of interpolation of HFR observations of surface currents. The underlying idea is to first retrieve spatial correlations between the radial velocities from the data, then use the obtained statistics to fill gaps in observations (interpolate in data space) and consistently define the cost function weights (inverse of the velocity error covariance) for 2dVar interpolation.

The proposed processing technique consists of four consecutive steps: (a) EOF analysis of the radial velocities; (b) Signal/noise separation and computation of the cost function weights; (c) Filling gaps in observations; (d) 2dVar interpolation of the preprocessed dataset.

2.1. Variational interpolation of the radial velocities

Consider an oceanic domain Ω partly bounded by the coastline $\partial\Omega_0$ where HFRs are located (Fig. 1). Projections v_k^n of the surface

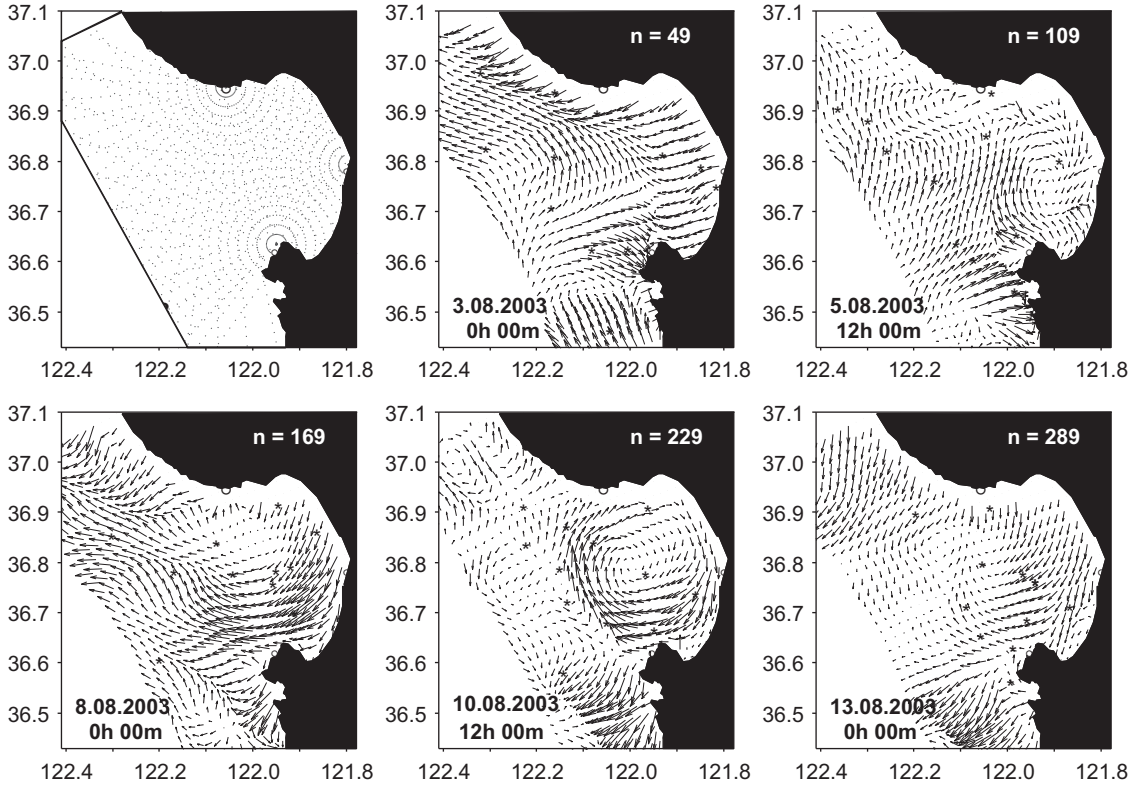

Fig. 1. Evolution of the reconstructed surface velocity field in the Monterrey Bay. Upper left panel shows distribution of the observation points. Solid line indicates the boundary of the interpolation domain Ω. Cross-validation points are shown by asterisks.

velocity field $\mathbf{v}^t(x,y,t^n)$ on the radar beam directions $\mathbf{r}_k$ are observed at times t^n at a discrete set of points $\mathbf{x}_k$, $k=1,...,K$ located along the beams (upper left panel in Fig. 1). Our goal is to obtain gridded estimates $\mathbf{v}^n$ of $\mathbf{v}^t(x,y,t^n)$ given v_k^n.

The number M of points of the regular grid defines the number of unknown values of the interpolated function to be determined from K observations at a given time t^n. In the particular case of the HFR data the interpolation grid step is often chosen such that $M\sim K/2$ because two velocity components have to be defined at a regular grid point, whereas only one component is measured at an observation point.

A standard approach to regularizing interpolation problems is penalizing small-scale variability (enforcing smoothness) of the interpolated fields (McIntosh, 1990). The latter is usually represented by the action of the Laplacian on the field subject to interpolation. In application to the HFR data Yaremchuk and Sentchev (2009) have shown that it is beneficial to enforce smoothness in the divergence $\mathrm{div}\mathbf{v}=\partial_x u+\partial_y v$ and vorticity $\mathrm{curl}\mathbf{v}=\partial_x v-\partial_y u$ patterns. This approach facilitates extraction large-scale components of these physically important features of circulation. In the present study we utilize similar technique. In addition to the terms proportional to $(\Delta\mathrm{div}\mathbf{v})^2$ and $(\Delta\mathrm{curl}\mathbf{v})^2$ we also enforce smoothness in the velocity field $\mathbf{v}$.

At a given time the velocity field $\mathbf{v}$ is obtained through the constrained minimization of the quadratic cost function:

$$J=\frac{1}{2K}\sum_{k=1}^{K}\sigma_k^{-2}[(\hat{P}_k\mathbf{v})\cdot\mathbf{r}_k-v_k]^2 + \frac{1}{2A}\int_{\Omega}\left[W^d(\Delta\mathrm{div}\mathbf{v})^2+W^c(\Delta\mathrm{curl}\mathbf{v})^2+W^u(\Delta\mathbf{v})^2\right]d\Omega \rightarrow \min_{\mathbf{v}|_{\mathbf{v}(\partial\Omega_0)=0}} \quad (1)$$

Here K is the number of observations, A is the area of the interpolation domain Ω and $\hat{P}_k$ is the local interpolation operator which projects the unknown velocity vectors onto the kth observation point from the apexes of the grid cell, enveloping that point. Factors σ_k^{-2},W^d,W^c and W^u are the inverse error variances of the corresponding squared quantities, so that J could be treated as the argument of the Gaussian pdf $\mathcal{P}(\mathbf{v})$ defined on the $2M$-dimensional space of the gridded velocity fields $\mathbf{v}$: $\mathcal{P}(\mathbf{v})\sim\exp[-J]$.

In contrast to the previous studies (Kaplan and Lekien, 2007; Yaremchuk and Sentchev, 2009), where regularization factors or their analogues were determined empirically, here we take the advantage of the rich temporal statistics provided by the HFRs and obtain the cost function weights from the statistical analysis of the data.

2.2. Signal/noise separation

Spectral decomposition provides the following representation of the covariance matrix:

$$\boldsymbol{C}=\boldsymbol{U}\Lambda\boldsymbol{U}^{\mathrm{T}}$$

where $\boldsymbol{U}$ is a rectangular matrix whose columns are the eigenvectors $\boldsymbol{e}^k$ (empirical orthogonal functions, EOFs) of $\boldsymbol{C}$ corresponding to the eigenvalues λ_k, and $\Lambda=\mathrm{diag}\{\lambda_k\}$. The eigenvalues quantify time variation of the spatial patterns in the radial velocity distributions described by the corresponding EOFs.

Having the EOF decomposition of the data at hand, the noise level could be estimated using the cross-validation (CV) technique (e.g., Beckers and Rixen, 2003). The technique provides a certain number of EOFs (modes) K_r, which describe the portion of variability of the radial velocities, that is well-resolved by the HFRs. The rest of the modes $\boldsymbol{e}_k$, $k>K_r$ are attributed to noise, whose spatial variability cannot be determined with statistical confidence.

Technically, K_r is computed as the number of modes which provide a minimum for the interpolation error at the randomly chosen set ω_c of CV points (Fig. 1). These points are temporarily removed from observations and constitute a small portion of the dataset to minimize their impact on the result of covariance estimate.

Note that locations of the CV points should change in time in order to keep the dimension of the covariance matrix $\boldsymbol{C}$ equal to K. This requirement may result in a certain distortion of the covariance estimate because, in the presence of the artificial gaps (introduced by CV points), the number of snapshots in the time-averaging operation depends on the pairs of points being correlated. In such case the covariance estimate may not be even positive-definite (e.g., von Storch and Zwiers, 2002) but with sufficiently small number of CV points one may hope this effect will be negligible. Technical details of minimizing the interpolation error with respect to K_r are given in Sections 2.3 and 3.2.

With the optimal cutoff number of modes K_r, the covariance matrix $\boldsymbol{C}$ can be decomposed into the well-resolved $\boldsymbol{C}_r$ and unresolved (noisy) $\boldsymbol{C}_n$ constituents:

$$\boldsymbol{C}=\boldsymbol{C}_r+\boldsymbol{C}_n\equiv\boldsymbol{U}_r\Lambda_r\boldsymbol{U}_r^{\mathrm{T}}+\boldsymbol{U}_n\Lambda_n\boldsymbol{U}_n^{\mathrm{T}} \quad (2)$$

where $\boldsymbol{U}_r$ is the $K_r\times K$ matrix, whose columns are the first (well-resolved) eigenvectors, $\Lambda_r=\mathrm{diag}\{\lambda_k\}$, $k=1,\ldots,K_r$; the eigenvectors in the columns of the $(K-K_r)\times K$ matrix $\boldsymbol{U}_n$ are attributed to noise, and $\Lambda_n=\mathrm{diag}\{\lambda_k\}$, $k=K_r+1,\ldots,K$.

The noise level v is estimated as

$$v=\left[\frac{\mathrm{Tr}\Lambda_n}{\mathrm{Tr}\Lambda}\right]^{\frac{1}{2}}\equiv\left[\sum_{i=K_r+1}^{K}\lambda_i/\sum_{i=1}^{K}\lambda_i\right]^{\frac{1}{2}} \quad (3)$$

whereas observation error variances σ_k^2, $k=1,...,K$ are represented by the diagonal elements of the matrix $\boldsymbol{C}_n=\boldsymbol{U}_n\Lambda_n\boldsymbol{U}_n^{\mathrm{T}}$.

The diagonal elements of $\boldsymbol{C}$ are also used to estimate the variances σ_u^2, σ_c^2 and σ_d^2 of the respective fields $\Delta\mathbf{v}$, $\Delta\mathrm{curl}\mathbf{v}$ and $\Delta\mathrm{div}\mathbf{v}$. In the present study we assume that the corresponding inverse variances W^u, W^c and W^d do not vary in horizontal and compute them as the reciprocals of σ_v^2, σ_c^2 and σ_d^2. The values of σ_u^2, σ_c^2 and σ_d^2 are obtained through the following formulas

$$\sigma_u^2=\frac{2\mathrm{Tr}\boldsymbol{C}}{K\delta x^4},\quad \sigma_c^2=\frac{2\mathrm{Tr}\boldsymbol{C}}{K\delta x^4L_c^2},\quad \sigma_d^2=\frac{2\mathrm{Tr}\boldsymbol{C}}{K\delta x^4L_d^2} \quad (4)$$

where δx is the grid step of the interpolation grid and L_c, L_d are the scales of variability of the curl and divergence fields inferred from statistical analysis of the data.

2.3. Interpolation in the data space

The final step in preparing to interpolate is filling gaps in observations. Although the cost function (1) does not contain spatial correlations between the radial velocities, this information could be taken into account at the preliminary stage by the gap-filling technique proposed by Beckers and Rixen (2003) for the SST data. Below we give a brief description of the method with an emphasis on the difference in its application to HFR observations.

To fill a gap containing points $\mathbf{x}_i$ in a subdomain $\omega\subset\Omega$, the radial velocities $\mathrm{v}(\mathbf{x}_i)$ observed outside the gap ($\mathbf{x}_i\in\Omega\backslash\omega$) are expanded in K_r "resolved" eigenfunctions $\boldsymbol{e}^k$ of $\boldsymbol{C}$:

$$\text{find } \alpha_k:\ \sum_{\mathbf{x}_i\in\Omega\backslash\omega}\left[\boldsymbol{v}(\mathbf{x}_i)-\sum_{k=1}^{K_r}\alpha_k\boldsymbol{e}^k(\mathbf{x}_i)\right]^2\rightarrow\min_{\alpha_k} \quad (5)$$

and the expansion coefficients α_k are used to obtain radial velocities within the gap:

$$\boldsymbol{v}(\mathbf{x}_i \in \omega) = \sum_{k=1}^{K_r} \alpha_k \boldsymbol{e}^k(\mathbf{x}_i \in \omega) \quad (6)$$

Note that a gap may also include points from the CV set ω_c.

After this the EOF expansion is iteratively improved: a set of EOFs $\{\boldsymbol{e}^k_{(m)}\}$ on the mth iteration is computed using the covariance estimate $\boldsymbol{C}_{(m)}$ emerging from the "dataset" whose gaps are already filled with the help of the previous EOFs $\{\boldsymbol{e}^k_{(m-1)}\}$, then these new EOFs $\{\boldsymbol{e}^k_{(m)}\}$ are employed to fill the gaps again. The process terminates when the relative reduction of the mean interpolation error

$$\varepsilon^2 = \sum_{\mathbf{x}_i \in \omega_c} \left[\boldsymbol{v}_c(\mathbf{x}_i) - \sum_{k=1}^{K_r} \alpha_{(m)k} \boldsymbol{e}^k_{(m)}(\mathbf{x}_i) \right]^2 \quad (7)$$

computed over the CV set ω_c becomes smaller than the machine precision. The difference of our approach from BR03 is the following:

(a) at the first iteration we use the direct estimate of $\boldsymbol{C}$ (which may not necessarily have a positive definite spectrum) derived from the gappy dataset (BR03 algorithm fills the gaps with zeroes to obtain a positively definite estimate of the covariance matrix). The reason is that distortion of the spectrum by gaps in HFR data usually occurs at the high-wavenumber part of the spectrum, which is not used by the gap-filling process anyway, whereas the direct estimate of $\boldsymbol{C}$ gives somewhat better approximation of the leading eigenfunctions. Of course, a lot depends on the spatio-temporal structure of the gaps, but our experience with HFR data shows that this property is well kept for a typical set of HFR observations: we never encountered negative eigenvalues in neither simulated, nor real data;

(b) interpolation within the gaps is performed using (5) and (6), i.e. projection on the eigenvectors is performed *outside* the gaps: interpolated values within the gap areas are never used to compute projections of the data on the new eigenvectors. As a consequence, expansion coefficients α_i cannot be computed analytically via the inner product in the data space and the minimization problem (5) has to be solved numerically.

These modifications provide faster convergence of the iterative algorithm for the computation of the final set of EOFs (Section 3.2.1).

To summarize, the proposed method consists of four major steps:

1. EOF decomposition of the covariance matrix between the radial velocities.
2. Signal/noise separation and computation of the cost function weights.
3. Filling gaps in observations.
4. 2dVar interpolation of the preprocessed data set.

To assess the method's performance, we conducted twin-data experiments with simulated HFR data (Section 3) and real observations off the Opal Coast in northern France (Section 4).

3. Twin-data experiments

3.1. Setting

Setting of the twin-data experiments was chosen to mimic real observations conducted in the Monterrey Bay in summer of 2003 (Shulman and Paduan, 2009).

The "true" currents $\mathbf{v}^t$ were extracted from the 12.5-day run of the NCOM model forced by COAMPS (Hodur et al., 2002) winds. The model was configured on a curvilinear orthogonal grid (Paduan and Shulman, 2004) with a typical step of 1.8 km. Surface currents were sampled every hour along the beams of three radars which probed the radial components of the model surface velocity at 386, 407 and 349 points, respectively (Fig. 1, upper left panel). Therefore the dimension of the data space was $K=1142$. The total number of the gridpoints where velocity vectors were reconstructed was $M=560$, so the number of unknowns $2M=1120$ was approximately equal to the number of observations.

Radial velocities v_k "observed" at points $\mathbf{x}_k$ were defined by adding white noise w to projections of the model currents $\mathbf{v}^t$ on the beam directions $\mathbf{r}_k$:

$$v_k = (\hat{P}_k \mathbf{v}^t \cdot \mathbf{r}_k) + vVw. \quad (8)$$

Here $V=0.12$ m/s is the typical magnitude of $\hat{P}_k\mathbf{v}^t \cdot \mathbf{r}_k$ and v is the scalar parameter whose reciprocal has the meaning of signal/noise ratio. Three values of v (0, 0.1, and 0.3) were tested within each of six major series of twin-data experiments. Each series was characterized by specific structure of the artificial gaps introduced into the simulated dataset to assess the benefits of the gap-filling technique. These simulated data sets were the following:

(0) Without the gaps.
(a) With randomly distributed 1-point gaps (data loss $\gamma = 13.5\%$).
(b) With gaps, generated by obstacles, moving across the domain (Fig. 2): Each obstacle (ship) spoils data along three beams, whose intersection point coincides with the ship's position. We simulated back-and -forth motion of three ships, which effectively removed 6.9% of the data points from observations.
(c) With the gap created by discarding all observation points in the rectangular region (Fig. 2) for 1 day. This gap removed 28% of the data on August 10–11 and approximately $\gamma = 2\%$ of the data overall.
(d) With gaps generated by switching off for 12 h radars 2,3 on August 4, radar 1 on August 8 and radars 1,3 on August 12 (Fig. 2, $\gamma = 6.3\%$).
(e) With all the above mentioned gaps superimposed ($\gamma = 28.2\%$)

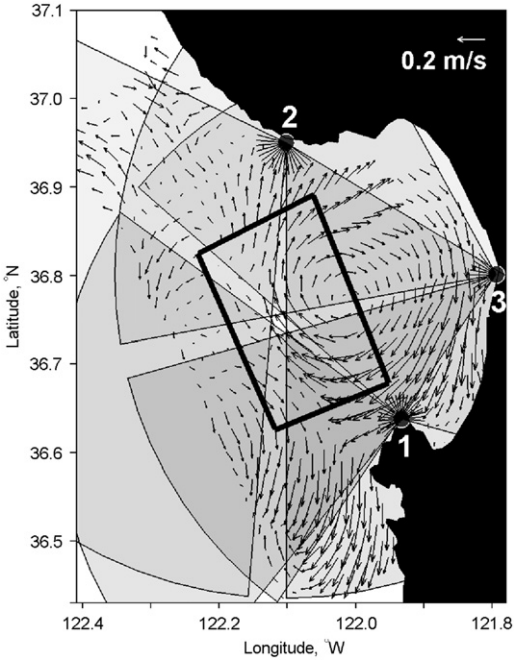

Fig. 2. Setting of the gap simulation experiments: Gray shading shows data acquisition areas of the radars in the presence of a simulated ship (case *b*) moving across the Monterrey Bay. Area within the rectangle shows the boundary of the data-free region in case *c*. Numbers enumerate radars switched off in case *d*.

To assess the effect of the preliminary interpolation in the data space (step 3, Section 2.3), we also compared the results of 2dVar interpolation of the raw data (with the gaps) and the results of 2dVar with the gaps filled.

The quality of interpolation was monitored by three parameters: Velocity error e_v was defined as the mean absolute difference between the true $\mathbf{v}^t$ and interpolated $\mathbf{v}$ currents normalized by the typical magnitude of $\mathbf{v}^t$: $e_v = \langle|\mathbf{v}^t-\mathbf{v}|\rangle/\langle|\mathbf{v}^t|\rangle$, where angular brackets denote space-time averaging over the interpolation grid. Similar expressions were used to assess the interpolation qualities e_d, e_c of the divergence and vorticity fields:

$$e_d = \langle|\mathrm{div}(\mathbf{v}^t-\mathbf{v})|\rangle/\langle|\mathrm{div}\mathbf{v}^t|\rangle; \quad e_c = \langle|\mathrm{curl}(\mathbf{v}^t-\mathbf{v})|\rangle/\langle|\mathrm{curl}\mathbf{v}^t|\rangle$$

3.2. Results

3.2.1. Signal/noise separation

Without the gaps there are $N = Kn = 1142 \times 301 = 343{,}742$ observation points, where n is the number of hourly timesteps in the 12.5 day time window. The CV set was specified by randomly removing 10–13 points on each time layer (Fig. 1) with the total number of CV points $N_{cv}=3490$.

The cutoff number K_r was determined for the noise levels ($v=$ 0, 0.1 and 0.3) by minimizing the interpolation error (7). To examine sensitivity of K_r to changes in the CV set ω_c we varied ω_c by specifying different seed values for the random generator of the CV points' locations, keeping the ratio N_{cv}/N close to 1%. These experiments have shown very weak dependence of K_r on ω_c.

Left panel in Fig. 3 shows calculations of K_r for the experiments with $v=0.1$ and 0.3: for observations specified by (8) the S/N separation number is 38 for $v=0.1$ and 20 for $v=0.3$. The corresponding estimates of the noise level (Eq. (3)) are 0.093 and 0.29 in very good agreement with the true values. For the case of perfect observations ($v=0$) K_r appeared to be close to N as the dependence $\varepsilon(m)$ flattened out at large m and did not show any distinct minimum.

To speed up convergence of the BR03 iterative process we made two modifications discussed in Section 2.3. Fig. 3 (right panel) shows their effect for a particular case of $m=10$ modes and $v=0.3$: The gray curve was obtained when the first guess covariance was estimated without filling the CV gaps with zeroes but with α_i in (5) computed through summation over Ω with $v(\omega_c)=0$. The solid black curve in the same panel was obtained in a similar way except for summation in (5) was done over $\Omega\backslash\omega_c$ (i.e. filled CV points were not taken into account). Similar improvement in convergence was observed for other values of v and m with artificial gaps also included.

Fig. 4 demonstrates the impact of gaps on the covariance matrix spectrum and the performance of the gap-filling procedure for the "realistic" case e (28.2% loss of HFR data, all types of gaps are present). It is seen that the gaps have little impact on the ten leading eigenmodes of the spectrum. This could be explained by the fact that these modes are associated with the largest spatial scales, whereas the examined gaps tend to have stronger distorting effect on the small (case a) and intermediate (b,c,d) scales of variability.

The gap-filling procedure appears to be rather effective as it dramatically improves the spectral shapes in the region $10 < m < K_r$: spectra after gap-filling closely follow the observed ones for both values of noise level. Moreover, it appears that the gap-filling technique is able to extract useful signal from noisy observations as the gap-filled curves appear to follow the true spectrum more closely than the observed (without the gaps) curve in the vicinity of K_r ($v=0.3$).

3.2.2. Cost function weights

Spatial variability of the error variances σ_k retrieved from the gappy data was not smooth enough for adequate mapping. We examined dependence of σ_k on the distance from a radar. Fig. 5 shows a typical curve computed for the case $v=0.3$, $\gamma=28.2\%$. Errors increase from 2 to 4 cm/s in the ranges between 5 and 35 km and then more sharply to 6 cm/s between 35 and 45 km. This behavior can be explained by a certain loss of accuracy of the gap-filling technique at larger distances, where the density of observation points is getting smaller.

One may think of a possibility to take into account off-diagonal elements of the error covariance matrix $\boldsymbol{C}_n$ by specifying the data misfit weights in Eq. (1) in the full matrix form $\boldsymbol{C}_n^{-1} = \boldsymbol{U}_n\Lambda_n^{-1}\boldsymbol{U}_n^{\mathrm{T}}$ rather than in its diagonal approximation. Experiments with this formulation of the cost function have shown, however, that diagonal approximation $\boldsymbol{C}_n^{-1} \approx \mathrm{diag}\sigma_n^{-2}$ provides a better and more stable fit to the "truth". A possible reason is that the off-diagonal elements are obtained from the limited number of samples with less statistical confidence than the diagonal ones. Indeed, in a series of additional experiments the off-diagonal elements exhibited strong dependence on the structure of the gaps, whereas their spatial variation appeared too noisy even at small

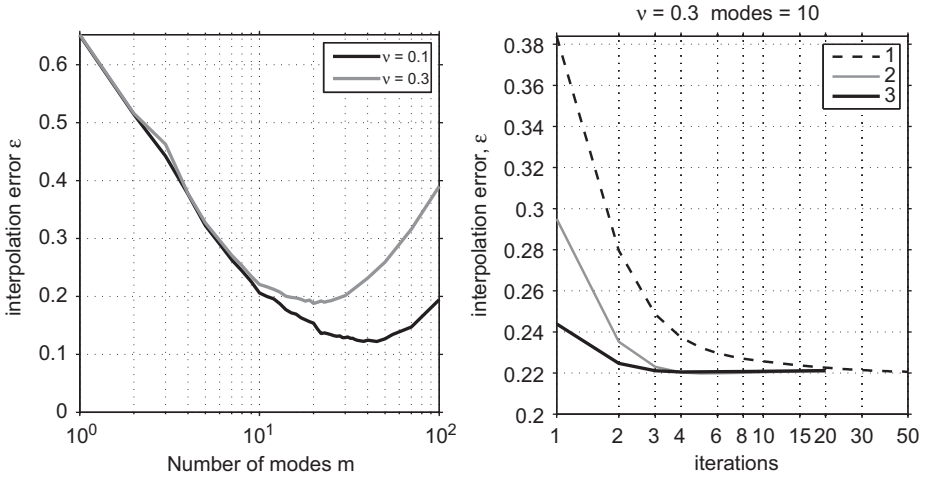

Fig. 3. Left: Interpolation error ε as a function of the number of modes m. Right: convergence rates in computation of this error ($v=0.3$,$m=10$) for the iterative process used in BR03 (1), for the same process but with the direct estimate of $\boldsymbol{C}$ at the zeroth iteration (2), and the process used in the present study (3).

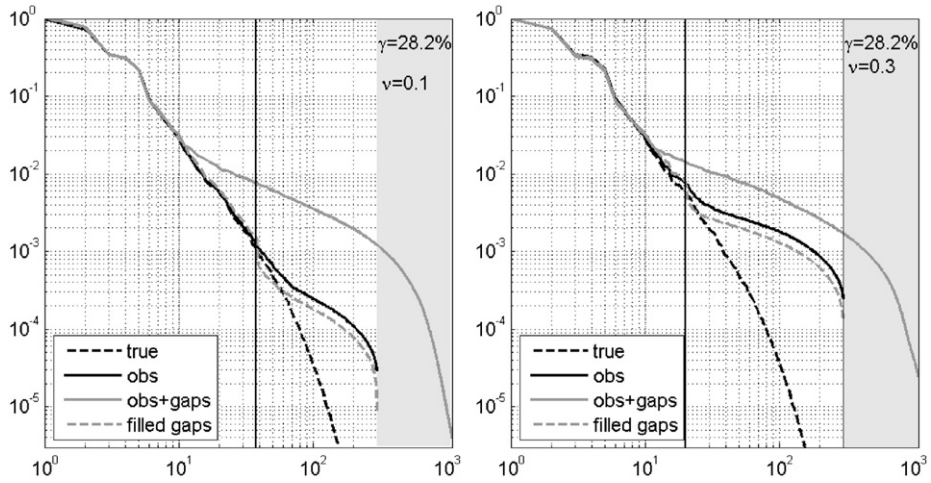

Fig. 4. Normalized spectra of the NCOM radial velocities taken directly from the model (black dashed curve), same velocities contaminated by noise (eq. (10), solid black curve), with gaps superimposed (solid gray) and the gaps removed (dashed gray). Shaded region on the right denotes the null space of the covariance matrix estimated with full ($N=300$) time series. Spectral components with $m>N$ emerge as a result of uneven lengths of the time series in the estimate of the covariance matrix elements—without gaps there are $K=1142$ time series with N elements each and spectral density is zero at $m>N$ for continuous observations. Note that presence of the gaps artificially elevates the relative spectral density at $m>10$.

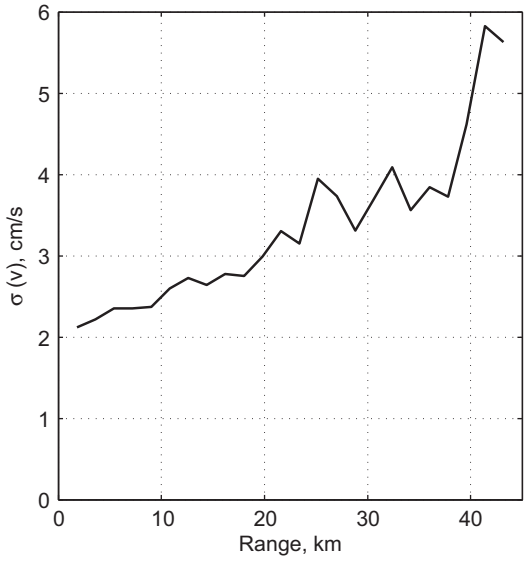

Fig. 5. Mean radial velocity error variance $\sigma(v)$ as a function of the distance from a radar.

separations between the data points, resulting in the overall loss of robustness of the estimate of $\boldsymbol{C}_n^{-1}$.

The efficiency of the approximation (4) of the regularization weights $W^u=\sigma_u^{-2}$, $W^c=\sigma_c^{-2}$ and $W^d=\sigma_d^{-2}$ was checked in a series of experiments, where the weights were varied to obtain the best fit to the "true" fields. These computations have shown that such "optimized" weights never departed more than 15% from the respective values obtained with Eq. (4), i.e. without the exact knowledge of the true velocity. At the same time interpolation errors e_v, e_c and e_d were only 5–8% larger than the "optimized" ones, showing feasibility of the estimate (4).

In the following sections we consider algorithm's performance in more detail, paying special attention to its skill in the presence of various types of gaps.

3.2.3. Random gaps

Table 1 compares the proposed interpolation algorithm with the 2dVar method (Yaremchuk and Sentchev, 2009). A robust 1–2% improvement of the interpolation error is observed in the velocity, divergence and vorticity fields for case *a* (randomly distributed 1-point gaps, Section 3.1). The improvement is significantly lower than the percentage of data loss (13.5%), primarily because filling random 1-point gaps affects information content on the grid scale which is poorly resolved anyway. Besides, data absence is largely compensated by observations in the points located in the immediate vicinity of the 1-point gaps at distances often smaller than the grid step. These neighboring points compensate missing data and provide the 2dVar interpolation with enough information on the larger-scale variability. Also note that the surface velocity field is recovered in most cases with a better accuracy e_v than the noise level $v=0.3$.

Table 1
Dependence of the interpolated field parameters e_v, e_c, and e_d on the structure of the gaps in HFR observations for $v=0.3$. The results of standard 2dVar (without filling the gaps) and 2dVar with the gaps filled are shown, respectively, in the left and right columns of the table cells.

case	γ (%)	e_v		e_d		e_c	
0	0.0	–	0.251	–	0.652	–	0.537
a	13.5	0.260	0.252	0.665	0.659	0.551	0.542
b	6.9	0.256	0.251	0.662	0.654	0.545	0.540
c	2.0	0.254	0.251	0.660	0.655	0.545	0.539
d	6.3	0.280	0.258	0.677	0.661	0.564	0.542
abcd	27.9	0.311	0.274	0.703	0.679	0.589	0.558

3.2.4. Moving ships

Moving ships (case *b*) spoil 6.9% of the entire set of 343,742 data points. Relative improvement of the 2dVar interpolation (line 2 in Table 1) is somewhat smaller than for the case of completely random gaps: Velocity field is better by 1.1% whereas vorticity and divergence fields show 0.9 and 1.0% improvements, respectively. Nevertheless, the proposed algorithm appears to be pretty robust with respect to this type of gaps as well.

3.2.5. Single gap

Much more difficulties emerge when a gap occupies a significant portion of the interpolation domain, as in case (*c*) (Fig. 2). To better illustrate the benefits of the gap-filling technique, we placed the gap at the location of an eddy-like structure seen in the

mouth of the Monterrey Bay around the 10th of August, 2003 (Fig. 6, left panel). Without interpolation, this eddy is not reproduced by the 2dVar technique (Fig. 6, right panel), simply because there is no information on the eddy in the velocity field, surrounding the gap. On the contrary, if the spatially inhomogeneous correlations between the radial velocities are taken into account, a certain portion of this eddy emerges from the prior statistical information, providing a much better skill to the 2dVar algorithm (Fig. 6, middle panel).

Table 1 does not give full impression of the improvement, because error data are averaged over the whole observation period (12.5 days), whereas the datasets in case (c) differ only on the 10–11th of August. If averaging is performed over the time period containing the gap (from 12PST 10.08.2003 to 12PST 11.08.2003) then the advantage is obvious: e_v is reduced from 0.49 to 0.32 (33% reduction). Similar error reductions are also observed for the vorticity and divergence fields (28 and 37%, respectively).

3.2.6. Switching off the radars

A common reason for low data return of a HFR system is malfunction of one or more radars. We simulated this kind of situation by switching off both northern radars for half a day on 4th of August, southern radar on 8th and two southern radars on 12th. The strongest reduction of the interpolation errors occurred on 12th of August, when the reconstructed (true) currents were generally perpendicular to the beams of the only operating radar. In that case the velocity error e_v reduced 64% (from 0.84 to 0.34) with 66% of the missing data being filled.

Velocity distributions show that 2dVar interpolation tends to align velocities along the beams of the only working radar (right panel in Fig. 7), producing rather unrealistic pattern (compare with the left panel in the same figure). After filling of the missing data at the southern radars, the skill of the 2dVar algorithm is significantly improved.

Advantage of the preliminary gap-filling is less visible in the case when only one of the three radars is switched off. This is because the major improvement occurs in the subregions covered by a single radar: When radar 1 in Fig. 2 was switched off on 8th of August, such regions emerged on the periphery of the domain and provided the major contribution to the 25% increase in e_v (computed by averaging over the 12 h period when the radar was switched off). Filling of the radar 1 data reduced e_v by 15% with similar reductions observed for the divergence and vorticity errors.

3.2.7. Combined gaps

Finally, all the gaps were combined together to obtain a "realistic" HFR record, characterized by 72% of the data return. Fig. 8 gives an overall comparison between the methods

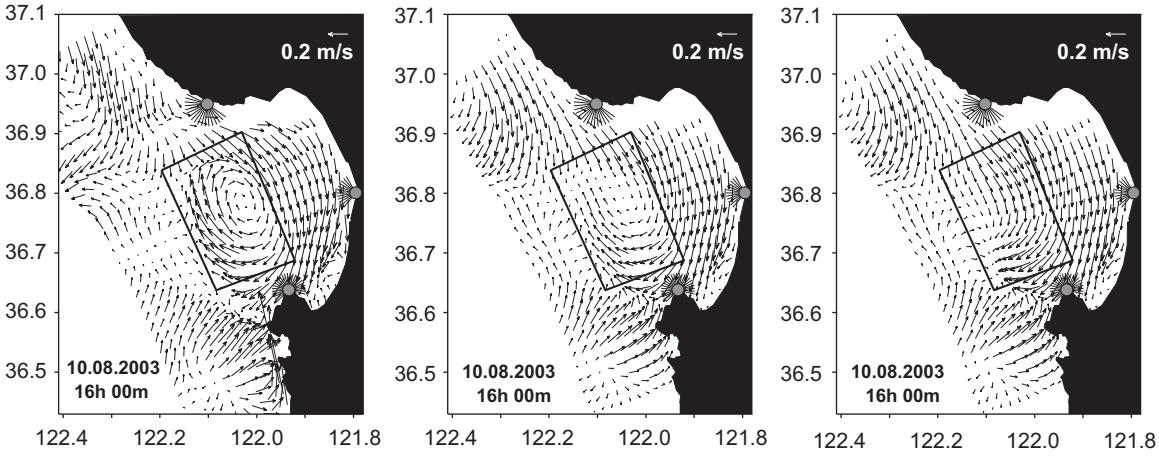

Fig. 6. The "true" velocity field (left panel) and velocity fields reconstructed by 2dVar with preliminary filling the rectangular gap (middle panel) and without (right panel).

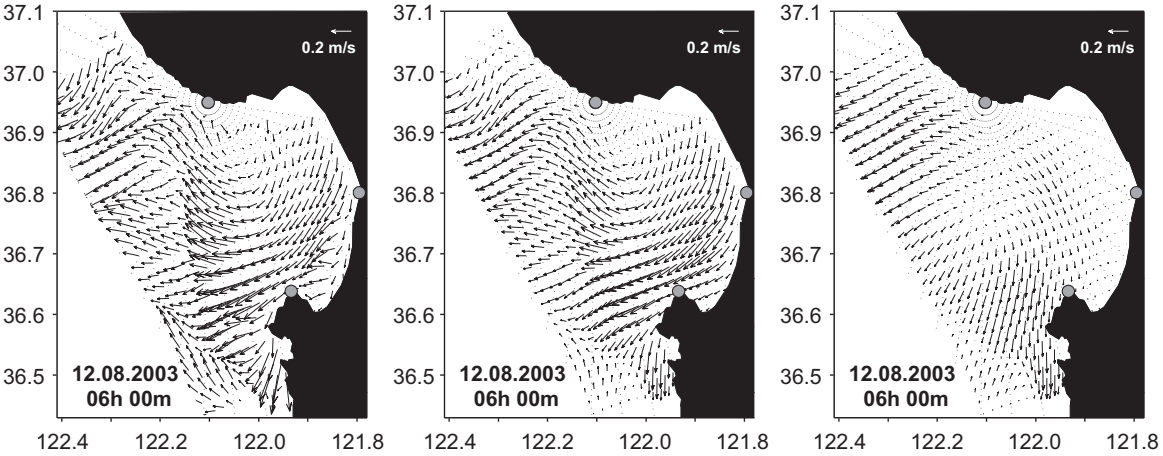

Fig. 7. The "true" velocity field (left panel) and velocity fields reconstructed by 2dVar with preliminary filling the gaps caused by switching off two southern radars (middle panel) and without filling (right panel).

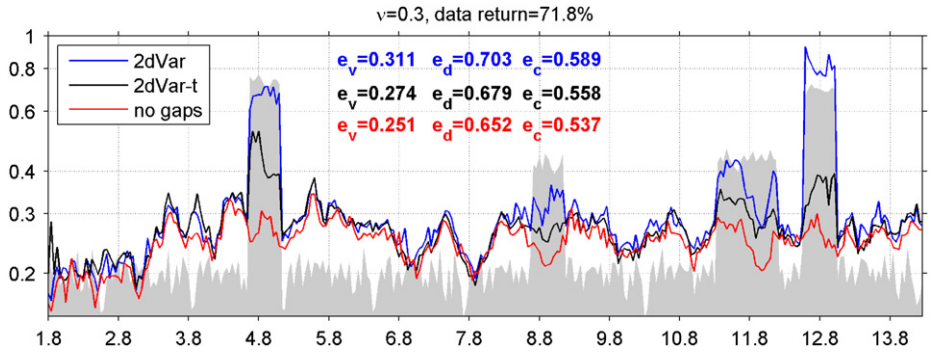

Fig. 8. Velocity interpolation error e_v for the "perfect" dataset (without gaps, red line), for the gappy data with (black) and without (blue) preliminary EOF-based interpolation of the radial data. Shaded area denotes the part of data occupied by the gaps. Particularly severe losses of data are observed during simulated radar malfunctions (4, 8 and 12 of August) and on August 11th, when "observations" were removed from a large region shown in Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

in terms of e_v, e_d and e_c. It is obvious that EOF-based gap-filling of the radial data is particularly advantageous during a "severe data loss" events caused by either malfunction of a radar (8.8) or two (4.8, 12.8); or by data loss in a region, whose size is considerably larger than the grid step (11.8).

Beyond these periods, when only 1-point and ship-generated gaps are present, the proposed algorithm still has some (1–3%) advantage over the 2dVar in terms of e_v, e_d and e_c (see Table 1 and numbers in Fig. 8).

It is also noteworthy that the proposed technique allows to retrieve the sea surface state with an accuracy $e_v=0.27$ better than the noise level $\nu=0.3$ (Fig. 8) even in the case of 28% loss of observations. The conventional 2dVar technique ($e_v=0.31$, Fig. 8) demonstrates somewhat lower skill, primarily because of much poorer performance during the heavy data loss periods.

4. Real data experiments

To test the algorithm with real data, we processed HFR observations obtained in the course of the ERMANO experiment off the Opal coast of the Pas de Calais in northern France.

4.1. The data

In May–June 2003, two HF radars were deployed to monitor surface currents: one radar was located on the Cape Gris Nez (CGN) and the other one was 12 km farther south, at Wimereux (WMX, Fig. 9). The entire 35-day record from 0.00 CET 01.05.2003 to 23.40 CET 04.06.2003 was used for testing. Surface currents were sampled every 20 min at 10° azimuthal resolution defined by the beam width. The radial velocity data were binned along the beams at 1.8 km resolution. Grid cells with less than 75% data returns were excluded from consideration, constraining the interpolation domain to the ranges less than 20 km (Sentchev and Yaremchuk, 2007) and the total number of "good" observation points to $K=203$. Overall, the analyzed records were characterized by 87% of data return (68,059 of 511,560 observations were discarded). Approximately 10% of the missing radial velocities were due to data acquisition problems at the Wimereux radar on May 3 (3 h) and May 21–22 (21 h). The instrumental accuracy of the measurements was 5 cm/s.

Regional velocity pattern (Fig. 9) is dominated by the M2 tidal constituent which contributes 77% to the total velocity variation in the area. The maximum velocities reach 1.8 m/s with the typical magnitude of the velocity vector of 0.52 m/s. Observed radial velocities had the rms amplitude of $\varepsilon_r=0.34$ m/s.

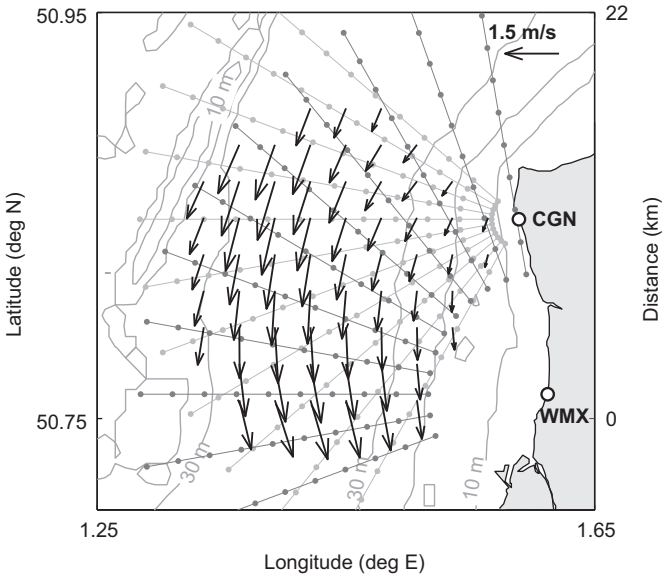

Fig. 9. Surface velocity at 12.20GMT 24.05 2007 in the ERMANO study area. Contours show the bathymetry in meters. Gray dots indicate locations of the surface velocity measurements by two radars (shown by circles).

To estimate observational noise level and the quality of interpolation, a set of CV points ω_c was removed from the data. Every 20 min 8–12 locations were randomly selected, and radial velocities observed at these points were extracted from the dataset. In total, 24,097 (4.7%) observations were removed.

The quality of interpolation was estimated as the mean absolute difference between the values of the interpolated velocity at the CV points and the radial velocities measured at these points:

$$e_v^* = \langle |v_k - \hat{P}_k \mathbf{v} \cdot \mathbf{r}_k| \rangle$$

Here index k enumerates the CV points and angular brackets denote averaging over ω_c. The total number of gaps in observations (including the CV points) was 92,156 (18%).

4.2. Results

Similar to twin-data experiments, the noise level was determined by minimizing the CV interpolation error (7). Dependencies of the normalized interpolation errors ε and e_v^* on the number of modes k demonstrated distinct minima at $k=33$–35 (Fig. 10). We selected $N_r=33$ as the noise cutoff number. The observational noise level computed through Eq. (3) was close to 0.15, or 5.1 cm/s,

in a good correspondence with the above estimate of the instrumental error.

In experiments with real data the true velocity field is unavailable, so the quality of interpolation of the curl and divergence fields e_c, e_d cannot be assessed. The interpolation quality was monitored by a single parameter e_v^* estimated at the CV points. Fig. 11 shows time evolution of e_v^* for two cases: with and without filling the gaps. It is seen that the preliminary gap-filling technique is able to reduce the time-averaged relative interpolation error to 0.16 (5.5 cm/s), a value very close to the observational noise level. On the contrary, the mean value of e_v^* without preliminary gap-filling appears more than two times higher (0.35) indicating a significant benefit of combining 2dVar interpolation with the EOF analysis.

The advantage of the gap-filling technique is most vividly seen during the periods when the Wimereux radar was not working (see increase in the data loss to 45–50% on May 3 and 21–22 in Fig. 11). Fig. 12 shows tidal ellipses obtained by averaging of the interpolated currents over the 24-hour period from 12.00 CET 21.05 to 12.00 CET 22.05 for the cases with (a) and without (b) preliminary gap-filling. The pattern in Fig. 12a appears to be completely unrealistic as the major axes of the ellipses tend to align along the beams of the only operating radar in Cape Gris Nez. Fig. 12b is apparently more close to reality since the spatial distribution of the ellipses is much more similar to the one obtained by averaging the currents over the two 24-hour periods immediately before and immediately after the gap (from 12.00 20.05 to 12.00 21.05 and from 9.00 22.05 to 9.00 23.05): During these two periods both radars were in full operation with the average data return of approximately 9% (Figs. 11, 12c).

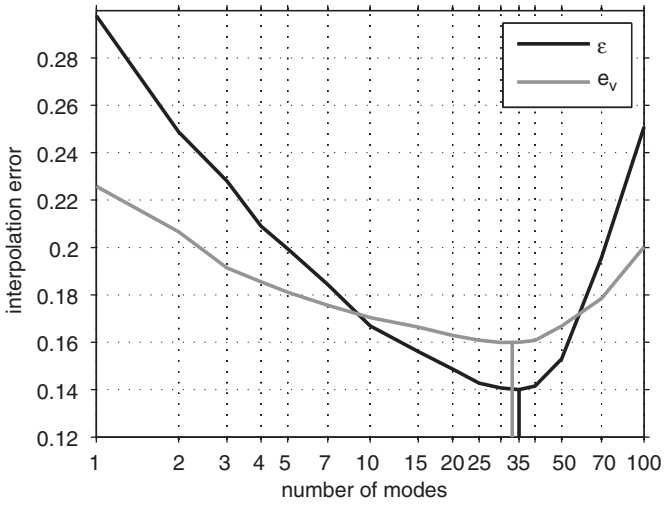

Fig. 10. Normalized interpolation errors ε and e_v as functions of the number of modes. The curves are normalized by the observed radial velocity variance ε_r.

Robustness of the noise separation and gap-filling algorithms was investigated in the similar way as in twin-data experiments: the CV subset ω_c was varied in size from 2 to 5% of the total number of observations by changing the parameters of the random generator of the CV points. The resulting values of K_r were found to vary between 32 and 36 ($\nu = 16-13\%$) while the respective values of e_v^* were even more stable varying in the range of 0.158–0.164.

We also studied the effect of the length of time averaging T on the quality of interpolation: The averaging interval was centered at 23.20 GMT 21.05.2003 (middle of the gap in Fig. 11) and varied between 3 and 20 days. Results of these experiments have shown that e_v^* comes close to the saturation (noise) level when $T > 6$ days, i.e. approximately 12 periods of the dominant wave (M2) were necessary to provide statistically robust cross-correlations between the radial velocities of two radars.

5. Discussion and conclusions

Observations of surface currents by high frequency radars are disrupted by environmental factors causing numerous gaps in the data. The number of gaps increases with distance and result in the loss of accuracy at far ranges. Moreover, since continuous operation of at least two radars is crucial for successful reconstruction of the velocity field, even a short-term radar malfunction may interrupt monitoring of the surface velocity and result in the dramatic loss of accuracy in prediction of particle trajectories that is important in many practical applications.

In the present study we combined the EOF analysis with the 2dVar interpolation technique to successfully process occasional single-radar coverage events and improve the overall quality of monitoring of sea surface currents by the HF radars. EOF analysis of the radial velocities provides (a) statistically rigorous estimation of the weights for the 2dVar algorithm, and (b) a set of spatial patterns (EOFs) capable to fill large gaps in the data caused by radar malfunctions. Our approach takes the advantage of the frequent time sampling by the HFRs and employs observation history to estimate the leading modes of variability of the radial velocities. Similar to SST analysis (Beckers and Rixen, 2003; Alvera-Azcarate et al., 2005), these modes are used to fill the gaps in HFR observations, which frequently occur in practice. Numerical experiments with simulated and real data have shown that preliminary gap-filling is extremely beneficial during occasional periods of heavy data loss associated with radar malfunctioning: With the proposed technique, the interpolation errors during these periods are typically reduced 1.5–2 times providing much more realistic velocity distributions (Figs. 6, 7, 12).

The interpolation method can be summarized as a four-step procedure: EOF analysis of the radial velocities; estimation of the noise and the cost function weights; filling gaps in observations, and finally, retrieving of the velocity vectors from the filled dataset.

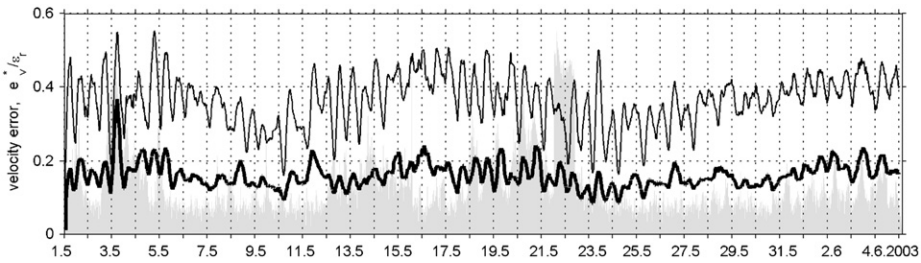

Fig. 11. Radial velocity errors at the cross-validation points for 2dVar interpolation with (bold curve) and without preliminary gap-filling. Errors are normalized by the observed radial velocity variance ε_r and smoothed with the 2-hour running mean. Shading indicates the relative amount of gaps in the data.

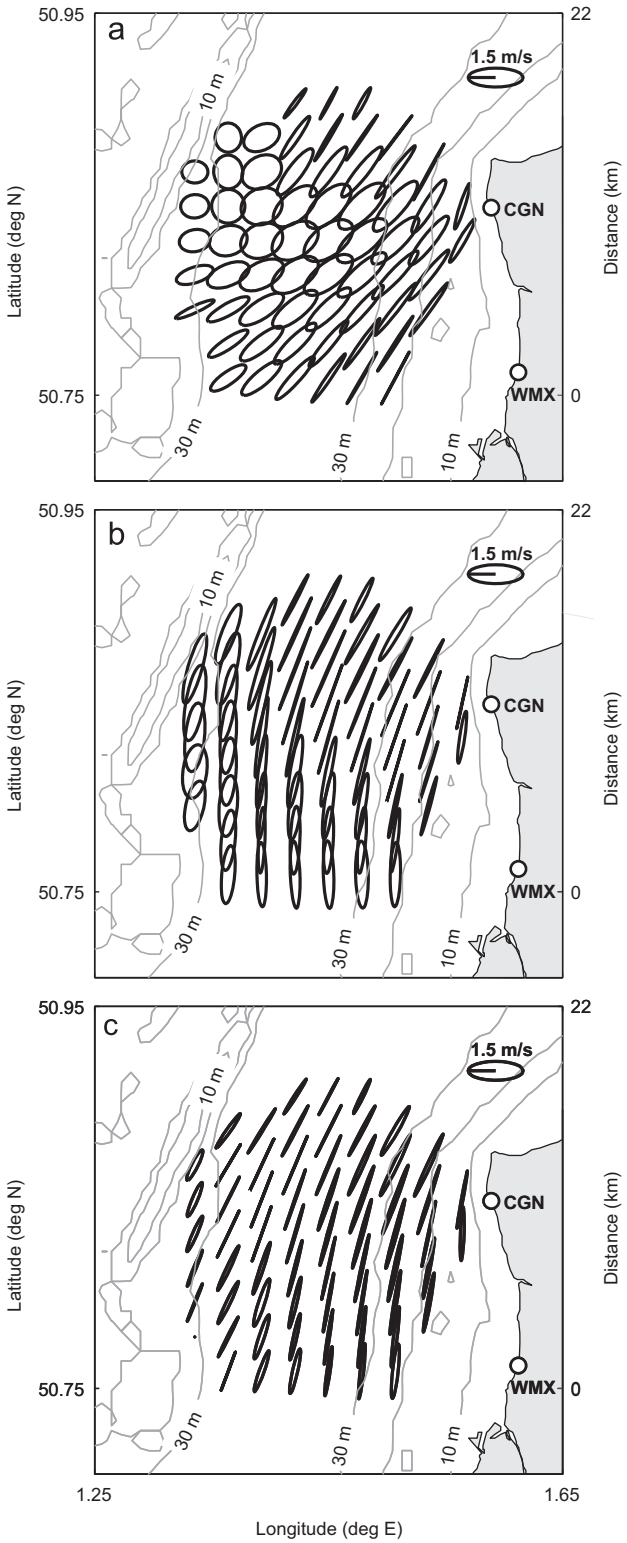

Fig. 12. Tidal ellipses computed by averaging of the surface velocities obtained by the standard 2dVar method (a); and the proposed EOF/2dVar technique (b,c). 24-hour time averaging is performed over the period of one operating radar (starting 12.40GMT 21.05) for (a) and (b) and over two 24-hour intervals preceding the gap (starting 12.40 20.05) and following immediately after (starting 12.00 22.05) for (c). Every second ellipse is shown.

In real applications the HFR time series may exceed 1–2 years, and selection of the time interval for estimating the covariances becomes important. In the present study the sea surface velocity was dominated by tidal motions, and we have shown that averaging over more than 12 periods of the dominant wave is adequate. In the near-coastal areas with weak tidal currents (e.g., lakes, semi-enclosed seas) the time interval should be long enough to statistically capture the major events typical for regional sea surface dynamics. Our ongoing experience with multi-year HFR observations in the Gulf of Lyon (Forget et al., 2008) shows that the presented method also works well in a region whose dynamics is quite different from the tidally dominated regime considered here: circulation in the Gulf is characterized by sporadic mistral wind-driven events on the background of strong mesoscale activity and relatively weak tidal/inertial oscillations. It was found that 4-month moving average was adequate enough to apply the technique successfully.

The proposed method may not be applicable to short HFR records characterized by 1–2 brief and strong "events" occurring during the observation period. In that case the cross-validation technique may not properly work and the S/N separation model may be invalid because noise statistics becomes far from Gaussian.

From the computational point of view, the proposed algorithm is not expensive: the most time-consuming part is calculation of the S/N separation number K_r which requires multiple interpolation runs in the data space (Figs. 3, 10). In our case these computations consumed only several minutes of CPU time of a single 2.66 GHz processor. For larger grids ($K,M \sim 10^4$) the required CPU time may be close to an hour, but can be easily reduced using multiple processors because the time-consuming computation of the curve in the left panel of Fig. 3 is readily parallelized in the number of modes. Besides, since K_r is unlikely to change in time significantly, this lengthy computation have to be executed only once: a search for a minimum in Fig. 3, 10 could be done more effectively when a good guess for K_r is available from the first computation.

In view of the above, the technique can be applied in near-real time with only minor modification. In this case error statistics is estimated by averaging over the period preceding the moment of data acquisition and then updated by replacing the oldest data in the time series by the new readings. Similar updates are made to other quantities described in Section 2.2. However, since the number of samples used for statistical estimation is fairly large, the updates can be made once in a while, e.g., when contribution c_n of new data to the time series exceeds 1–2%. This "real-time" approach has been tested with the HFR observations in the Gulf of Lyon (Forget et al., 2008). Preliminary results show that time-averaged interpolation error ϵ_v^* is reduced significantly when the variational method is combined with the gap-filling technique, but this reduction is not sensitive to the frequency of updates when $c_n < 2.5\%$ (every 3 days).

Presented material and preliminary results with multi-year HFR observations do suggest that the proposed technique may be useful in processing a large variety of HFR datasets with significant loss of data.

Acknowledgements

This study was partly funded by the Office of Naval Research (Program element 0602435N) and completed during MY visit to the Université du Littoral with support from the French Ministry of Education and Research.

References

Alvera-Azcarate, A., Barth, A., Rixen, M., Beckers, J.M., 2005. Reconstruction of incomplete oceanographic data sets using empirical orthogonal functions: application to the Adriatic Sea surface temperature. Ocean Modell. 9, 325–346.

Beckers, J.M., Rixen, M., 2003. EOF calculations and data filling from incomplete oceanographic observations. J. Atm. Oceanic Tech. 20 (12), 1839–1856.

Breivik, O., Sætra, O., 2001. Real time assimilation of HF radar currents into a coastal ocean model. J. Marine Syst. 28, 161–182.

Chavanne, C., Janekovic, I., Flament, P., Poulain, P.-M., Kuzmic, M., Gurgel, K.-W., 2007. Tidal currents in the northwestern Adriatic: high-frequency radar observations and numerical model predictions. J. Geophys. Res. 112, C03S21. doi:10.1029/2006JC003523.

Fang, F., Pain, C.C., Navon, I.M., Piggott, M.D., Gorman, G.J., Farrell, P.E., Allison, P.A., Goddard, A.J.H., 2009. A POD reduced-order 4D-var adaptive mesh ocean modelling approach. Int. J. Num. Meth. Fluids 60, 709–732. doi:10.1002/fld.1911.

Forget, P., Barbin, Y., Andre, G., 2008. Monitoring of surface ocean circulation in the Gulf of Lions (North-West Mediterranian Sea) using WERA HF radars. In: Proceedings of IGARSS, CD ROM, Boston, USA (7–11 July).

Hisaki, Y., Tokeshi, T., Fujie, W., Sato, K., Fujii, S., 2001. Surface current variability east of Okinawa Island obtained from remotely sensed and in situ observational data. J. Geophys. Res. 106 (C12), 057–073.

Hodur, R.M., Pullen, J., Cummings, J., Hong, X., Doyle, J.D., Martin, P.J., Rennick, M.A., 2002. The coupled ocean-atmospheric mesoscale prediction system (COAMPS). Oceanography 15 (1), 88–98.

Hoteit, I., Cornuelle, B., Kim, S.Y., Forget, G., Khl, A., Terrill, E., 2009. Assessing 4D-VAR for dynamical mapping of coastal high-frequency radar in San Diego. Dyn. Atmosph. Oceans 48 (1–3), 175–197. doi:10.1016/j.dynatmoce.2008.11.005.

Kaplan, D., Lekien, F., 2007. Spatial interpolation of surface current data based on open-boundary modal analysis. J. Geophys. Res. 112, C12007. doi:10.1029/2006JC003984.

Kim, S.Y., Terrill, E., Cornuelle, B., 2007. Objectively mapping HF radar-derived surface current data using measured and idealized data covariance matrices. J. Geophys. Res. 112, C06021. doi:10.1029/2006JC003756.

Kim, S.Y., Terrill, E., Cornuelle, B., 2008. Mapping surface currents from HF radar radial velocity measurements using optimal interpolation. J. Geophys. Res. 113, C10023. doi:10.1029/2007JC004244.

Kondrashov, D., Ghil, M., 2006. Spatio-temporal filling of missing points in geophysical datasets. Nonlin. Process. Geophys. 13, 151–159.

Lewis, J.K., Shulman, I., Blumberg, A.F., 1998. Assimilation of the Doppler radar current data into numerical ocean models. Cont. Shelf Res. 18, 541–559.

Lipphardt, B.L., Kirwan, A.D., Grosch, C.E., Lewis, J.K., Paduan, J.D., 2000. Blending HF radar and model velocities in Monterey Bay through normal mode analysis. J. Geophys. Res. 105 (C2), 425–450.

McIntosh, P.C., 1990. Oceanographic data interpolation: objective analysis and splines. J. Geophys. Res. 95 (C8), 529–541.

Oke, P.R., Allen, J.S., Miller, R.N., Egbert, G.D., 2002. A modelling study of the three-dimensional continental shelf circulation off the Oregon. Part I. Dynamical analysis. J. Phys. Oceanogr. 32, 1383–1403 J. Geophys. Res. 109, C07S09..

Paduan, J.D., Shulman, I., 2004. HF radar data assimilation in the Monterey Bay area. J. Geophys. Res. 109, C07S09.

Sentchev, A., Yaremchuk, M., 2007. VHF radar observations of surface currents off the northern Opal coast in the eastern English Channel. Cont. Shelf Res. 27, 2449–2464.

Shulman, I., Paduan, J., 2009. Assimilation of HF radar-derived radials and total currents in the Monterrey Bay area. Deep-Sea Res. II 56 (3–5), 149–160.

von Storch, H., Zwiers, F.W., 2002. Statistical Analysis in Climate Research. Cambridge University Press, 484pp.

Tipett, M.K., Anderson, J., Bishop, C., Hamill, T.M., Whitaker, J.S., 2003. Ensemble square root filters. Mon. Wea. Rev. 131, 1485–1490.

Yaremchuk, M., Sentchev, A., 2009. Mapping radar-derived sea surface currents with a variational method. Cont. Shelf Res., 1711–1722.

Yaremchuk, M., Nechaev, D., Panteleev, G., 2009. A method of successive corrections of the control subspace in the reduced-order variational data assimilation. Mon. Wea. Rev. 137 (9), 2966–2978.