Adjoint-Free Variational Data Assimilation into a Regional Wave Model

GLEB PANTELEEV

University of Alaska Fairbanks, Fairbanks, Alaska, and National Research Tomsk Polytechnic University, Tomsk, Russia

MAX YAREMCHUK AND W. ERICK ROGERS

Naval Research Laboratory, Stennis Space Center, Mississippi

(Manuscript received 12 September 2014, in final form 23 December 2014)

ABSTRACT

A variational data assimilation algorithm is developed for the ocean wave prediction model [Wave Model (WAM)]. The algorithm employs the adjoint-free technique and was tested in a series of data assimilation experiments with synthetic observations in the Chukchi Sea region from various platforms. The types of considered observations are directional spectra estimated from point measurements by stationary buoys, significant wave height (SWH) observations by coastal high-frequency radars (HFRs) within a geographic sector, and SWH from satellite altimeter along a geographic track. Numerical experiments demonstrate computational feasibility and robustness of the adjoint-free variational algorithm with the regional configuration of WAM. The largest improvement of the model forecast skill is provided by assimilating HFR data (the most numerous among the considered types). Assimilating observations of the wave spectrum from a moored platform provides only moderate improvement of the skill, which disappears after 3 h of running WAM in the forecast mode, whereas skill improvement provided by HFRs is shown to persist up to 9 h. Spaceborne observations, being the least numerous, do not have a significant impact on the forecast skill but appear to have a noticeable effect when assimilated in combination with other types of data. In particular, when spectral data from a single mooring are used, the satellite data are found to be the most beneficial as a supplemental data type, suggesting the importance of spatial coverage of the domain by observations.

1. Introduction

In the last decades, significant progress has been made in numerical modeling of oceanic waves (Cavaleri et al. 2007; Janssen 2008). With the advent of massive observations of sea surface roughness from satellites, the forecast skill of the wave models (WMs) acquired new prospects for further improvement through optimization of the poorly known model parameters with respect to the observations.

The early attempts to improve WM performance through data assimilation employed the nudging technique to reduce the misfit between the integral characteristics of the model spectra, such as significant wave height (SWH), and observations (Francis and Stratton 1990; Bauer et al. 1992; Breivik and Reistad 1994; Wittmann and Cummings 2004). More sophisticated assimilation schemes updated the model using optimal interpolation of several spectral characteristics (Lionello et al. 1992; Dunlap et al. 1998; Greenslade 2001). The approach was generalized to the spectral partitioning technique (Voorrips et al. 1997; Aouf and Lefevre 2006; Abdalla et al. 2006), which appeared to provide a reasonably good improvement of the forecast skill.

An interesting approach was proposed by Bauer et al. (1996), who explicitly inverted the linearized Wave Model (WAM; WAMDI Group 1988; Monbaliu et al. 2000) operator to obtain corrections to the wind fields induced by observations of the wave spectra. This method was constrained by a rather strong assumption that the response of the linearized model is localized both in space and time. Holthuijsen et al. (1997) explored a similar technique using a limited number of nonlocal perturbations that were derived from statistical analysis of the wind forcing error fields. The method demonstrated a 20%–25% reduction of the forecast errors at time scales less than a day.

Corresponding author address: Max Yaremchuk, Ocean Prediction, Naval Research Laboratory, Bldg. 1009, Stennis Space Center, MS 39529.
E-mail: max.yaremchuk@nrlssc.navy.mil

DOI: 10.1175/JTECH-D-14-00174.1

Voorrips et al. (1999) were among the first to explore the performance of the extended Kalman filter (KF) in application to WAM using simulations in simple geometries. The extended KF, its reduced-order approximation, and fixed-lag Kalman smoother were tested and they demonstrated better performance than optimal interpolation (Hasselmann et al. 1997), especially in estimating the wave field away from observations and in past wind corrections. More recently, Sannasiraj et al. (2006) proposed a hybrid assimilation scheme that combined a statistical error model at the observation points with spatial interpolation using the Kalman gain matrix. Their method was capable of reducing the forecast errors by 30%–60% over 12–24-h forecast periods. Operationally oriented sequential KF methods were also considered by Emmanouil et al. (2010, 2012), who employed WAM in combination with the second-order KF and optimal interpolation schemes to improve the model's performance in the North Atlantic.

Variational methods involving the adjoint models have been in use in WM data assimilation for more than two decades (Snyder et al. 1992; Barzel and Long 1994; De Las Heras et al. 1994, 1995; Hersbach 1997; Veeramony et al. 2010). One of the difficulties with this approach is in the numerical complexity of the WMs, which employ sophisticated nonlocal parameterizations of the source terms and operate in spatially variable spectral space. De Las Heras et al. (1994) used an approximation to the adjoint WAM, neglecting spatial derivatives in the evolution equation. Orzech et al. (2013) took an alternative approach of developing the adjoint for spectral advection and performed a series of experiments controlling steady-state open boundary conditions of the linearized version of spectral WMs for nearshore [Simulating Waves Nearshore (SWAN)] applications. Similar experiments were performed by Veeramony et al. (2010), who used the analytical adjoint developed by Walker (2006). Recently, Orzech et al. (2014) employed the adjoint of SWAN to optimize the offshore SWH observations.

In the present study we explore the feasibility of the new variational data assimilation technique (Yaremchuk et al. 2009; Trevisan et al. 2010) in application to WAM. This approach employs a sequence of low-dimensional subspaces that are iteratively updated in the process of finding a cost function minimum. A distinctive feature of the method is that it does not require development of the tangent linear and adjoint codes for implementation, which is important for WM applications. In addition, the adjoint-free four-dimensional variational data assimilation (a4DVAR) method is less vulnerable to the instabilities that develop in the adjoint versions of the numerical models due to the breakdown of the tangent linear approximation in strongly nonlinear regimes. It was also shown that the a4DVAR technique appears to be advantageous when observations are sparse and noisy.

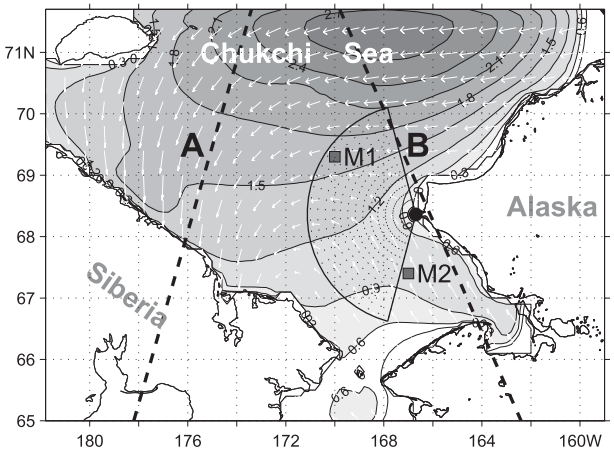

FIG. 1. Wind speed (white arrows) and SWH (contours, m) of the model solution at 0000 UTC 20 Sep 2011 ($t = 0$). Mooring positions are shown by black squares. HFR location and coverage area are given by the black circle with a sector. SWH data are acquired along the radar beams shown by dotted lines within the sector. Dashed lines are the tested tracks of *Envisat*.

We explore the performance of the a4DVAR technique with WAM configured in the 800 km × 800 km region of the Chukchi Sea (Fig. 1). In recent years this sector of the Arctic Ocean was subject to significant summer/fall storm activity due to the rapid loss of the ice cover, but its wave conditions remain largely unexplored. The choice of the region was partly motivated by the anticipated increase of shipping through the Bering Strait. This may require monitoring and forecasting of the wave conditions, which could be better done with data assimilation. Presented data assimilation (DA) experiments are performed with several types of observations that include SWH data from satellites, coastal HF radars, and spectral characteristics obtained from in situ observations at moored buoys.

In the following section, we present the details of experimental setting after a brief description of the a4DVAR assimilation technique. In section 3, the results of synthetic data experiments are presented and compared with the results of the traditional optimal interpolation (OI) method, observability of the wave field from different platforms is discussed, and the computational cost of the a4DVAR is considered. Additional discussion is offered in section 4, preceded by a brief summary and overall conclusions.

2. Methodology

a. a4DVAR technique

A variational DA algorithm performs a dynamically constrained search for a minimum of the cost function J, which quantifies the distance between observations and the numerical model solution used to assimilate the data. This minimization is performed with respect to a set of poorly known model parameters (e.g., the set of gridpoint values of the model fields at the start of time integration) that constitute the control vector $\mathbf{c}$. Assuming the differentiability of J with respect to c, in the vicinity of the minimum the cost function is represented by

$$J = J_o + \frac{1}{2}(\mathbf{c} - \mathbf{c}_o)^{\mathrm{T}}\mathbf{J}(\mathbf{c} - \mathbf{c}_o), \tag{1}$$

where J_o and $\mathbf{c}_o$ are the values of J and $\mathbf{c}$ at the minimum, respectively, and $\mathbf{J}$ is the Hessian matrix of the assimilation problem.

Assume now for simplicity a linear framework and minimization of the following cost function with respect to the model state $\mathbf{x}$ at the initial time $t = 0$ $\mathbf{x}_0 \equiv \mathbf{c}$:

$$J = \frac{1}{2}(\mathbf{c} - \mathbf{c}^b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{c} - \mathbf{c}^b) + \frac{1}{2}\left[\sum_n (\mathbf{H}_n\mathbf{x}_n - \mathbf{d}_n^*)^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}_n\mathbf{x}_n - \mathbf{d}_n^*)\right]. \tag{2}$$

Here $\mathbf{c}^b$ is the first guess (background) model state at $t = 0$; $\mathbf{B}$ is its error covariance; summation is made over observation times t_n, when the data $\mathbf{d}_n^*$ are available; $\mathbf{R}$ is the observation error covariance; operators $\mathbf{H}_n$ project the respective model states $\mathbf{x}_n$ onto the vectors $\mathbf{d}_n^*$ of observed quantities; and the superscripted T stands for transposition.

The cost function (2) can be rewritten in terms of deviations $\tilde{\mathbf{x}} = \mathbf{x} - \mathbf{x}_b$ of the model states from their background values:

$$J = \frac{1}{2}\left[\tilde{\mathbf{x}}_0^{\mathrm{T}}\mathbf{B}^{-1}\tilde{\mathbf{x}}_0 + \sum_n (\tilde{\mathbf{H}}_n\tilde{\mathbf{x}}_n - \tilde{\mathbf{d}}_n)^2\right], \tag{3}$$

with $\tilde{\mathbf{H}}_n = \mathbf{R}^{-1/2}\mathbf{H}$ and $\tilde{\mathbf{d}}_n = \mathbf{R}^{-1/2}\mathbf{d}_n - \tilde{\mathbf{H}}_n\mathbf{x}_n^b$. Introducing the notation $\mathbf{M}^n$ for the model operator mapping $\mathbf{c}$ onto $\mathbf{x}_n$ and omitting tildes for convenience yields the explicit expression for $\mathbf{J}$ in terms of $\mathbf{c}$:

$$J = \frac{1}{2}\left[\mathbf{c}^{\mathrm{T}}\mathbf{B}^{-1}\mathbf{c} + \sum_n (\mathbf{H}_n\mathbf{M}^n\mathbf{c} - \mathbf{d}_n)^2\right]. \tag{4}$$

Equation (4) can be represented in the complete square form:

$$J = \frac{1}{2}(\mathbf{S}\mathbf{c} - \mathbf{b})^2, \tag{5}$$

where

$$\mathbf{S} = \begin{bmatrix} \mathbf{B}^{-1/2} \\ \mathbf{H}_1\mathbf{M}^1 \\ \vdots \\ \mathbf{H}_n\mathbf{M}^n \end{bmatrix}; \quad \mathbf{b} = \begin{bmatrix} 0 \\ \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_n \end{bmatrix}, \tag{6}$$

and $\mathbf{S}$ is the square root of the Hessian matrix $\mathbf{J} = \mathbf{S}^{\mathrm{T}}\mathbf{S}$. To find the minimum, one has to solve the normal equation

$$\mathbf{S}^{\mathrm{T}}\mathbf{S}\mathbf{c} = \mathbf{S}^{\mathrm{T}}\mathbf{b}. \tag{7}$$

The solution of (7) can be obtained by considering the equivalent left-preconditioned system

$$\mathbf{P}^{\mathrm{T}}\mathbf{S}\mathbf{c} = \mathbf{P}^{\mathrm{T}}\mathbf{b} \tag{8}$$

with the only requirement for the preconditioner $\mathbf{P}$ is to have the same range with $\mathbf{S}$ (e.g., Hayami et al. 2007). As seen in (8), the best preconditioner is $\mathbf{P}^* = \mathbf{S}\mathbf{J}^{-1}$, which delivers convergence in one iteration. In practice, however, (pseudo) inversion of the Hessian is computationally unfeasible, and various approximations to $\mathbf{P}^*$ are considered. The major principle of a preconditioned iterative solver is to implicitly build the solution of the form

$$\mathbf{c}_o = \mathbf{c}^b + \sum_m \alpha_m(\mathbf{P}^{\mathrm{T}}\mathbf{S})^m\mathbf{P}^{\mathrm{T}}\mathbf{r}_0; \quad \mathbf{r}_0 = \mathbf{S}\mathbf{c}^b - \mathbf{b}, \tag{9}$$

which can be performed by a variety of methods involving Krylov subspaces $\mathcal{K}$ (e.g., Vuik et al. 1996). In particular, the well-known conjugate gradient minimization algorithm can be viewed as a version of the Krylov subspace technique with $\mathbf{P} = \mathbf{S}$; that is, $\mathbf{J}^{-1}$ is replaced by the identity operator $\mathbf{I}$.

The adjoint-free variational methodology (Yaremchuk et al. 2009) avoids the use of the adjoint model $\mathbf{M}^{n\mathrm{T}}$ in the construction of the preconditioner but employs explicit computation of the columns of $\mathbf{S}$ to invert $\mathbf{J}$ in $\mathcal{K}$ and performs consecutive orthogonalization of the search subspaces with respect to the inner product induced by the Hessian matrix in the control space (see the appendix).

In principle, the availability of the algorithm for inversion of $\mathbf{J}$ in a sequence of $\mathbf{J}$-orthogonal subspaces is a necessary condition for the monotonic decrease of the cost function with iterations, but it is not sufficient to achieve the global minimum in the linear case due to a possibility of the breakdown of the process. The latter occurs when the new search directions become (almost) linearly dependent on the directions generated at the

previous iterations. In the realistic nonlinear applications, however, the number of iterations rarely exceeds a few hundred due to computational constraints. Therefore, the key practical requirement to the algorithm for the iterative update of the search directions is its ability to quickly generate subspaces spanned by the leading projections of the initial residual $\mathbf{r}_0$ on the eigenvectors of the inverse Hessian matrix, which defines the solution $\mathbf{c}_{\text{opt}} = \mathbf{J}^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{b}$ to (7). If the final residual is sufficiently small (i.e., the model–data misfit is smaller than the observation error), the solution is assumed to be acceptable. An ad hoc method is to employ the sequence $\{\mathbf{M}^n\mathbf{r}_m\}$, as it contains projections of the residual $\mathbf{r}_m$ at the mth iteration on the most persistent dynamical modes of the system and may therefore capture a significant part of $\mathbf{c}_{\text{opt}}$.

In the present study we employ the technique of building the search subspaces $\mathcal{K}$ on the sequences $\{\mathbf{M}^n\mathbf{c}_m\}$. This approach provided acceptable solutions after 100 iterations and therefore proved to be satisfactory in the experiments with WAM.

The a4DVAR algorithm was coded as the following sequence of operations:

1) Specify the background initial conditions $\mathbf{c}$ and run the model, writing down model states $\mathbf{x}_n$ and their deviations from the data $\mathbf{H}_n\mathbf{x}_n - \mathbf{d}_n^*$ at the observation sampling frequency (15 min). Using the model run output, compute the cost function value J_0 and the auxiliary vector $\mathbf{Sc}$.
2) Extract the m leading EOFs $\mathbf{e}_j$, $j = 1, \ldots, m$ from the sequence $\{\mathbf{x}_n\}$ to form the basis in the search subspace.
3) Perturb the initial condition $\mathbf{c}_j \rightarrow \mathbf{c} + \varepsilon\mathbf{e}_j$ ($\varepsilon = 0.001$) and perform (in parallel) the ensemble of the perturbed model runs, computing the respective perturbed values of J_j and $\mathbf{Sc}_j$.
4) $\mathbf{J}$-orthogonalize the search basis $\{\mathbf{e}_j\}$ with respect to at most k bases obtained on the previous iterations and compute the optimal correction $\delta\mathbf{c}$ to the initial condition by minimizing J in the subspace spanned by the perturbations (see the appendix).
5) Update the initial condition $\mathbf{c} \leftarrow \mathbf{c} + \delta\mathbf{c}$ and compute the updated values of $\mathbf{x}^n, J$, and $\mathbf{Sc}$ by running the model.
6) If the stopping criterion is satisfied (either the magnitude of the cost function gradient normalized by its initial value is less than $10^{-3}J_0$ or the number of iterations exceeds 100), exit. Otherwise, go to 1 to start the next iteration.

b. *Setting the experiments*

1) WAM

WAM performs a time integration of the balance equation describing spectral density $F(\mathbf{x}, \mathbf{k}, t)$ for the wave component with the wavenumber $\mathbf{k} = (k_x, k_y)$ at the location $\mathbf{x} = (x, y)$:

$$\frac{\partial F}{\partial t} + \nabla \cdot (\boldsymbol{\upsilon} F) = \mathcal{S}(F, \mathbf{x}, \mathbf{k}, t), \tag{10}$$

where $\mathcal{S}$ is the sum of source functions, primarily composed of wind-forced generation, dissipation, and redistribution of the wave spectrum by nonlinear wave–wave interactions (WAMDI Group 1988); $\nabla = \{\nabla_x, \nabla_k\}$ stands for the gradient in the horizontal and wavenumber coordinates; and υ is the four-component vector of the respective wave-propagation velocities depending on the ambient current and constrained by the dispersion relationship for linear surface waves:

$$\sigma^2 = g|\mathbf{k}| \tanh|\mathbf{k}|h, \tag{11}$$

where σ is the wave angular frequency and $h(\mathbf{x})$ is the water depth.

Given the appropriate initial/boundary conditions, ambient current, and wind forcing, (10) is integrated numerically to produce the evolution of the wave spectrum in the space–time domain of the model.

An important diagnostic formula available in the WAM package relates the squared SWH Q^2 with the spectral density through the following linear relationship:

$$Q^2(\mathbf{x}, t) = \hat{Q}N = 16\sum_{\mathbf{k}} F(\mathbf{x}, \mathbf{k}, t)\, d\mathbf{k}, \tag{12}$$

where $d\mathbf{k}$ denotes the gridcell area in the wavenumber space and summation is done over the entire grid.

The model was configured in the domain shown in Fig. 1 with the spatial resolution of $\delta x = 9$ km. There were $m_x = 4412$ active grid points in horizontal and $m_k = 600$ grid points (24 directions at 15° resolution and 25 logarithmically spaced frequencies between 0.0314 and 0.3091 Hz) in the wavenumber space. The total length of the state vector was $M = m_x \times m_k = 2\,647\,200$.

Distance between the model states was assessed in terms of the correlation coefficient C and the normalized rms difference S between the spectra:

$$C(F) = \frac{\langle F_1' F_2' \rangle}{\sqrt{\langle F_1'^2 \rangle \langle F_2' \rangle}}; \quad N' = N - \langle N \rangle \tag{13}$$

$$S(F) = \left[\frac{\langle (F_1 - F_2)^2 \rangle}{\sqrt{\langle F_1'^2 \rangle \langle F_2'^2 \rangle}} \right]^{1/2}, \tag{14}$$

where angular brackets denote averaging in space, time, and over the wavenumbers. Similar coefficients were

calculated to assess the differences between the scalar (SWH) and vector (wind speed) fields, with averaging performed just in space and time.

2) Cost function

The general form of the cost function used in the data assimilation experiments was identical to (4) with the M-dimensional vector $\mathbf{c} = N(t_0) - N_b(t_0)$ describing the difference between the gridded model state $F(\mathbf{x}, \mathbf{k}, t_0)$ and the background (first guess) state $F_b(\mathbf{x}, \mathbf{k}, t_0)$ at the start of model integration t_0. The first term in (4) was specified by

$$\mathbf{c}^{\mathrm{T}}\mathbf{B}^{-1}\mathbf{c} = W\sum_{\boldsymbol{x}}[(\mathbf{I} - a^2\nabla_x^2)\hat{Q}\mathbf{c}]^2 \qquad (15)$$

and was kept intact through all the experiments. In the aforementioned equation, W is the regularization weight; $\mathbf{I}$ is the $m_x \times m_x$ identity matrix; and the parameter a defines the high-frequency cutoff in horizontal coordinates, which enforces spatial smoothness in the deviations of the optimal state SWH from the background. By setting $a = 2\delta x$ throughout the experiments, SWH variability at spatial scales below a was heavily penalized. The regularization weight W was chosen to be inversely proportional to the squared mean of SWH in the background solution with the proportionality coefficient $\varepsilon_x = 0.01$.

As seen in (15), the parameter a can also be interpreted as the horizontal decorrelation scale of the SWH errors: The respective inverse error correlation is proportional to $(\mathbf{I} - a^2\nabla_x^2)^2$ and has the kernel that exponentially decays with the distance between the correlated points at the length scale of $4\sqrt{2/\pi}a \sim 60$ km (e.g., Yaremchuk and Smith 2011). Since $\mathbf{B}^{-1}$ is also proportional to $(\mathbf{I} - a^2\nabla_x^2)^2$, a similar spatial structure is induced in the spatial correlations between the spectral components of $\mathbf{c}$. Apparently, our choice of the decorrelation scale could be refined (see, e.g, Waters et al. 2013), but we consider it to be a reasonable approximation to reality given the objectives of our study. In the spectral subspace, (15) defines the inverse error covariance to have only one linearly independent column (specified by the components of $\hat{Q}$). As a consequence, spectral correlations at a given point are represented by an $m_k \times m_k$ correlation matrix whose elements are equal to 1 (thus implying 100% correlation between all the spectral components). This assumption has been routinely used in the sequential algorithms assimilating SWH (e.g., Wittmann and Cummings 2004).

The observation part of the cost function (4) is defined by the structure of the observation operators $\mathbf{H}_n$ and the respective error variances $\mathbf{R}_n$ described in section 2b(4).

To compare the results with the traditional OI method, we used the 2D OI approach (e.g., Wittmann and Cummings 2004; Waters et al. 2013) in application to the SWH data: at the observation times, the WAM state was sequentially updated by the OI analysis of the SWH field, which was projected onto the spectral components by multiplying the spectrum at a grid point by the ratio of the updated to predicted SWH values. The OI algorithm was configured with the same background error covariance $\mathbf{B}$, $\mathbf{R}_n$, $\mathbf{H}_n$ and using the same true and background solutions as the a4DVAR method.

3) True and background solutions

To perform the experiments, the true evolution of the wave field was generated by integrating WAM from the state of rest for 10 days under Oceanweather Inc. (OWI) wind forcing. The high-resolution OWI winds were developed by Oceanweather, Inc. using the methodology of Cardone et al. (1995, 1996). The winds were taken for the period 11–20 September 2011 and were updated every hour during the integration. The true state $N_t(\mathbf{x}, \mathbf{k}, t)$ shown in Fig. 2 was picked from the last 9 h of the model run (0000–0900 UTC 20 September 2011). Synthetic data (described in the next subsection) were picked from the true solution and then used for its reconstruction through minimization of the cost function (4).

Minimization started from the background run of the model, which was obtained as follows: The true model solution was averaged in time and space and the resulting spatially homogeneous spectrum was used as the initial condition for the background model run. The run was forced by the winds, which were different from those forcing the true solution. First, the true winds were horizontally smoothed to mimic the errors typical for reanalysis winds from meteorological centers that are usually available at a coarser (0.25°–1°) resolution and have to be interpolated on the fine resolution of grid of a regional ocean model. In the case considered, the smoothing was done by the isotropic Gaussian filter with the half-width of 25 km. After smoothing, the winds were rotated 35° counterclockwise to increase their distance from the truth to $S_{\mathrm{wind}} = 0.67$ [(14)]. The larger distance from the true forcing was needed for better assessment of the observation impact on the reconstruction of initial conditions, whose signature usually persists for 3–5 h in a typical wave model integration.

4) Synthetic observations

In this study two types of data are considered: moored observations of the wave spectra and SWH measurements from coastal HF radars and satellites. Each data

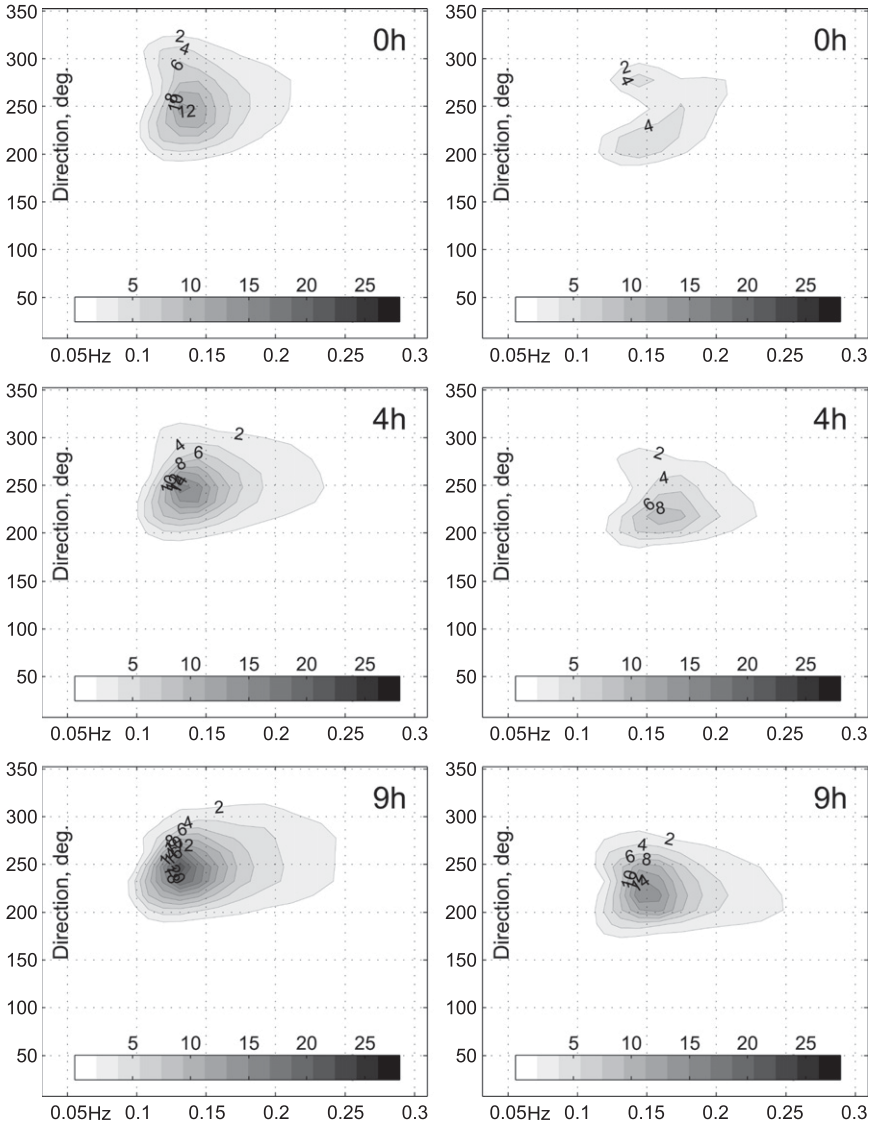

FIG. 2. Evolution of the horizontally averaged (left) true and (right) background spectra.

type had a specific structure of the observation operator, noise, and data generation procedure.

Two tested mooring sites are shown in Fig. 1: one was positioned in the open Chukchi Sea 150 km northwest of Cape Hope and the location of the second mooring corresponds to the real instrument maintained in the Bay of Kotzebu from July to December 2007 (Francis et al. 2011). Observations $\mathbf{d}^*$ by each mooring were generated by multiplying the true spectrum at any moment by the random factor $1 + \varepsilon\nu$, where ν is the white noise with unit variance and $\varepsilon_m = 0.01$. The observational error covariance matrices $\mathbf{R}$ for both moorings were diagonal with time-independent diagonal elements equal to $\langle \varepsilon N_t \rangle^2$. The respective observational operators $\mathbf{H}$ were time independent and picked the time-varying WAM spectra every 15 min from the grid point nearest to the buoy location, providing $4m_k = 2400$ observations per hour.

It should be noted that directional buoys provide only a few moments of the full spectrum and therefore constrain much smaller degrees of freedom of a model than m_k (usually, 2–3 times the number of frequencies, i.e., 50–80 in our case). These moments, however, are linear functions in the model state, and their calculation in terms of spectral components does not principally differ from Eq. 12. Assimilation of buoy moments would, of course, be a better approach in practice [like extending the method of Voorrips et al. (1997) to all the measured components of the directional spectrum], but in context of our study, this would fall somewhere in

between assimilating the full spectrum and its first moment (SWH) and would not be a useful addition to our analysis. We should also keep in mind that a user can assimilate postprocessed buoy data in the form of the full spectrum computed by the estimator available in the software provided with an instrument.

SWH observations were simulated by integrating the true spectrum field (Eq. 12) into the apexes of the grid cell containing a high-frequency radar (HFR) observation point followed by linear interpolation onto that point. After that, the SWH value was contaminated by random noise with the rms variance of 30 cm. HF radar observation points were located along the beams of the radar considered for placement at the Cape Hope site to monitor wave conditions north of the Bering Strait (Fig. 1). The above-described HFR observation operator computed SWH values along the 25 beams every 15 min, providing information to 535 model grid points within the sector shown in Fig. 1 (2140 observations per hour).

Synthetic satellite observations of sea surface roughness provided SWH data along the *Environmental Satellite* (*Envisat*) tracks shown in Fig. 1 with 9-km discretization (55 and 73 points for tracks A and B, respectively). These data were assumed instantaneous and satellite passage occurred for both tracks after 2 h of model integration. The respective observation operator was similar to the one used for HFR, except that it picked SWH values at the sequence of WAM grid points closest to the sampling points along the tracks (i.e., no spatial interpolation was used). Satellite SWH observations were contaminated similarly to HFRs with the rms error variance of 30 cm.

3. Results

Synthetic observations of SWH and wave spectra were assimilated into WAM using the adjoint-free variational technique. The model was constrained by data during the first 3 h of model integration and then integrated for 6 h to assess the improvement of the forecast skill. Performance of the method was quantified by calculating C (Eq. 13) and S (Eq. 14) between the optimized and true solutions. These quantities were computed with time averaging over three time intervals: 0–3 h (assimilation period), and two forecast periods of 3–6 h and 6–9 h. Five ensemble members were used for spanning the search subspace on every iteration. More details on building and orthogonalization of the ensemble members are given in the appendix. The 3-h ensemble model runs were executed in parallel and required 62 s of wall time per a4DVAR iteration on five processors of the 2.3-GHz cluster (1–1.5 h per experiment). The

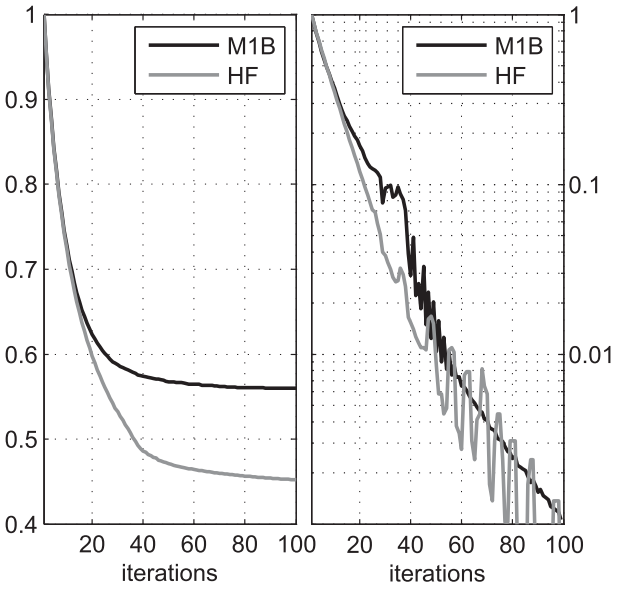

FIG. 3. Convergence of the a4DVAR algorithm in the assimilation experiments with M1B and HF data. (left) Cost function and (right) its gradient are normalized by their values at the start of assimilation.

sequential OI algorithm assimilating HFR data was executed in 74 s on a single processor.

A series of OI and a4DVAR experiments were conducted, involving assimilation of the data from five sources and their combinations: high-frequency radar at Point Hope (denoted by HF), two moorings (a4DVAR analyses only, locations shown in Fig. 1), and two *Envisat* tracks (A and B; Fig. 1). For comparison purposes, we conducted similar experiments with the OI method assimilating only SWH data from satellites and/or HF radar, which are abbreviated oHFA(B) and oHF, respectively, in the description below. With the exception of satellite tracks, all a4DVAR assimilation experiments demonstrated significant improvement of the model state in terms of its proximity to the true solution. The stopping criterion for optimization was the reduction of the cost function gradient 1000 times, which usually occurred after 80–100 iterations. By that time the value of J was typically reduced 2–3 times (Fig. 3). Maps of deviations from the truth of the spatially averaged background and optimized spectra at $t = 0$ are shown in Fig. 4. In most of the a4DVAR experiments, the initial error has been reduced to the values compatible with the wind forcing errors. The only exceptions were the results of optimal interpolation (Fig. 4b) and of the a4DVAR assimilation of SWH data from a single-satellite track (not shown): in these cases the optimized spectrum was only slightly different from the one produced by the background solution. For the OI case such a small

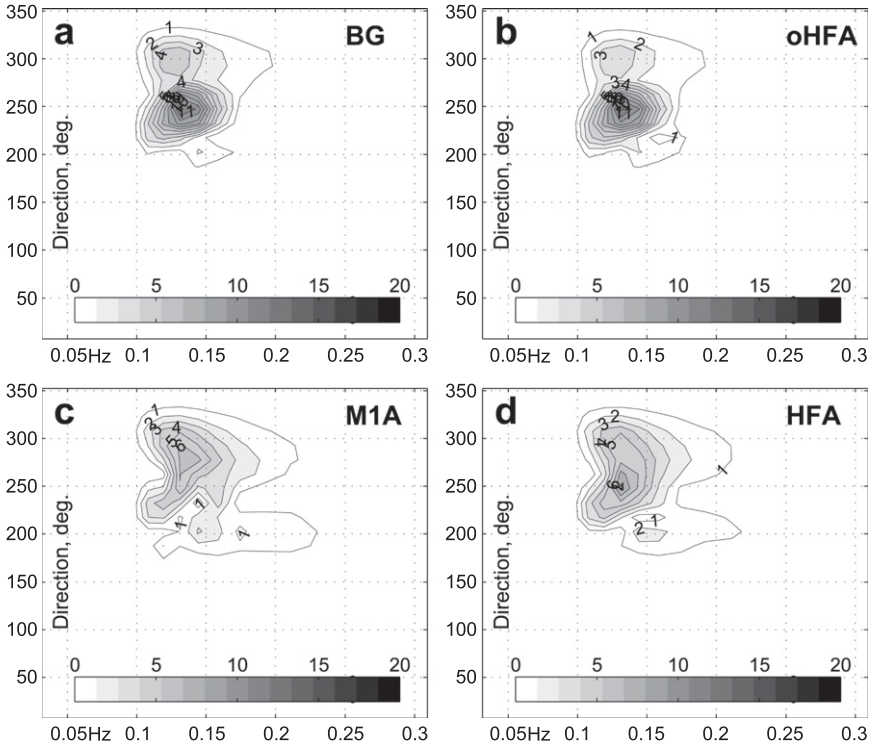

FIG. 4. Absolute difference between the horizontally averaged true spectrum at $t = 0$ and (a) background spectrum, (b) oHFA-optimized, (c) M1A-optimized, and (d) HFA-optimized spectra.

correction can be explained by the fact that SWH data are weakly constrained by dynamics and can barely affect the shape and location of the spectra because they provide information only on their mean magnitude at a geographical position. Small spectral improvement of the a4DVAR experiments with a single-satellite track could be attributed to the small amount of data (55 SWH observations). As a consequence, the cost function is dominated by the regularization term, which implies 100% correlations in spectral space and is therefore capable of adjusting only the spectral magnitude.

These properties of the abovementioned assimilated solutions translate into their lower spectral forecast skill shown in Table 1, which also includes spectral errors from the other assimilation experiments. Abbreviations in the left column of Table 1 correspond to the types of data used in the experiment (e.g., HFA corresponds to assimilation of the HF data and SWH data from the *Envisat* track A). As seen, 6420 observations by the HF radar provide the largest impact on the improvement of the wave field during the assimilation period: the correlation coefficient C_{03} increases from 0.47 to 0.77, whereas S_{03} drops from 0.89 to 0.65. The OI method provides a relatively low skill improvement similar to assimilation of a single-satellite track (cf. row 1 and rows 5, 13, 14).

Direct measurement of the spectra by a single mooring (7200 observation points, rows 6, 9) also provide only a moderate increase of C_{03} to 0.52 and a decrease of S_{03} to 0.86. This can be partly explained by the fact that assimilated spectra occupy a small part of spectral domain (at most 15%–20%; Fig. 2). As a consequence, the effective number of observations with useful (nonzero) information on the state of the wave field should be reduced 5–7 times down to ~1500 data points on the total, which is compatible, by the way, to assimilating 3–7 spectral moments. Besides, mooring data do not provide any information on the spatial variability of the spectra, which appears to be crucial for the successful recovery of the true state.

In that respect, it is remarkable that adding much less numerous satellite data to moored spectral observations improves the performance of the assimilation system considerably. Combining moored and satellite data provides 30%–40% growth of the correlation coefficients and a 20%–25% drop of the normalized standard deviations from the true spectrum (cf. rows 6 and 9 with rows 7–8 and 10–11, respectively). At the same time, satellite SWH data do not add much new information to that containing in HFR observations (cf. rows 2 and 3–4), which monitor the same integral

TABLE 1. Normalized S and C between the optimized and true solutions for the experiments with various types of data. Subscripts 03, 36, and 69 correspond to time-averaging ranges of 0–3, 3–6, and 6–9 h of model integration.

No.	Expt	C_{03}	S_{03}	C_{36}	S_{36}	C_{69}	S_{69}
1	BG	0.47	0.89	0.48	0.87	0.59	0.80
2	HF	0.77	0.65	0.72	0.69	0.71	0.70
3	HFA	0.75	0.66	0.72	0.68	0.70	0.70
4	HFB	0.76	0.64	0.76	0.65	0.75	0.67
5	oHFA	0.48	0.87	0.49	0.86	0.59	0.79
6	M1	0.53	0.85	0.47	0.87	0.50	0.86
7	M1A	0.70	0.71	0.70	0.71	0.68	0.73
8	M1B	0.71	0.70	0.70	0.71	0.69	0.72
9	M2	0.52	0.87	0.50	0.87	0.57	0.86
10	M2A	0.65	0.76	0.69	0.72	0.67	0.73
11	M2B	0.69	0.75	0.69	0.72	0.68	0.73
12	M1M2	0.71	0.73	0.67	0.75	0.68	0.73
13	A	0.47	0.89	0.48	0.88	0.57	0.79
14	B	0.48	0.88	0.48	0.87	0.58	0.78

quantity for the whole assimilation period (3 h) and cover a significant part of the model domain (Fig. 1).

The importance of the spatial coverage by observations is confirmed by the result of the experiment with assimilation of the spectra from two moorings: The values of C and S in this case demonstrate a considerable improvement and become compatible (row 12 in Table 1) with those achieved with the joint assimilation of spectra from the single mooring and satellite SWH data (rows 7–8 and 10–11).

Normalized rms deviation term S from the true solution as a function of time is shown in Fig. 5. It is noteworthy that the moderate reduction of S acquired after assimilating M_1 data is completely lost after 2 h of integrating the model without data constraints: after $t = 5$ hours, deviation from the truth becomes even larger than that of the OI solution, which is very close to the respective deviation of the background. The situation is much improved with satellite data being taken into account (dashed lines in Fig. 5): the value of S significantly drops and becomes compatible with the values achieved by assimilating HFR observations. Complementing HFR data with satellite observations does not affect the mean value of S during the assimilation period, but it has a noticeable impact on the forecast skill of the model (quantified by S) up to the end of integration. Apparent convergence of all the curves by hour 9 is the consequence of the wind forcing, which tends to attract model trajectories to a state whose error is compatible with that of the wind ($S_{\text{wind}} = 0.67$).

The time dependence of the spatially averaged correlation coefficients $C(t)$ between the true and optimized spectra (Fig. 6) demonstrates similar behavior: assimilation of the M_1 spectra provides a slight improvement of

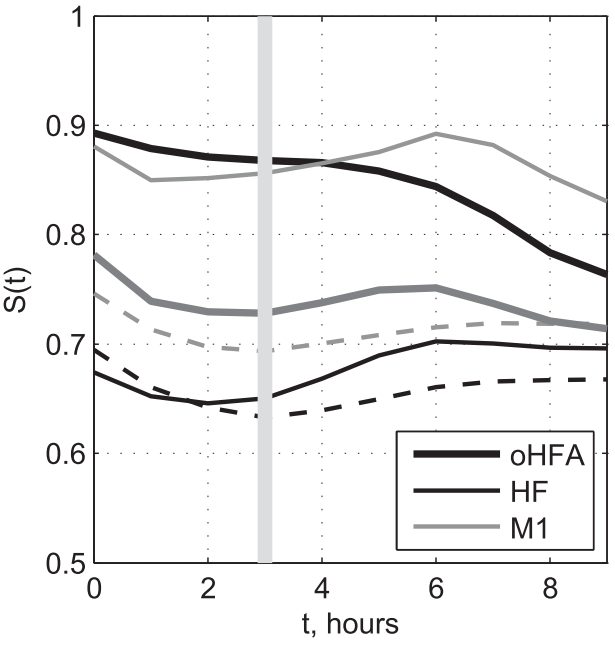

FIG. 5. Normalized rms deviations from the truth of the oHFA, HF-optimized, and M1-optimized spectra. Dashed lines correspond to augmenting the a4DVAR-assimilated data with satellite SWH observations from satellite passing through track B (experiments HFA and HFB). Thick gray line corresponds to the M1–M2-optimized a4DVAR solution. Vertical line marks the end of assimilation period.

the correlation (cf. gray and thick black curve in Fig. 6), which persists for 2 h of model integration. Adding satellite SWH data from track B instantly provides information on the higher spectral densities in the northern part of the Chukchi Sea (Fig. 1) and rises the correlation to 0.65–0.7 (dashed gray line in Fig. 6).

Inspection of Table 1 also shows that information from track A also increases the efficiency of assimilating spectra from moorings but to somewhat lesser extent. This phenomenon can be partly explained by the fact that track A does not cover the region of the highest SWH and therefore provides less information on the magnitude of spatial variability of the wave field. Similarly, assimilation of the M_2 data appears to be slightly less efficient than M_1, which can be partly attributed to the M_2 position at the periphery of the domain.

Table 1 also indicates that on a regional scale, instantaneous *Envisat* observations cannot provide a significant improvement to the background state if they are not accompanied by in situ measurements. At the same time, satellite data become quite valuable in complementing observations if the wave conditions are measured by a single mooring.

The forecast errors provided by the HFA data assimilation using OI and a4DVAR techniques and

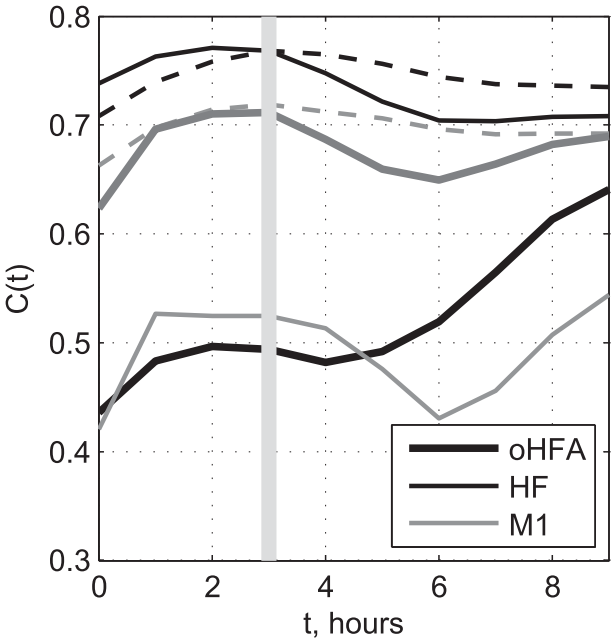

FIG. 6. As in Fig. 5, but for correlations.

averaged over the period of 3–9 h are compared in Fig. 7 in terms of the horizontal distributions of the SWH, peak period, and wave direction errors. that The a4DVAR technique provides 30%–50% better forecast skill in terms of the SWH (0.28 vs 0.37 m) and peak period (0.63 vs 0.91 s). Although discrepancies in the peak period near the southern and eastern boundaries are comparable in both solutions, the a4DVAR method demonstrates a significant advantage over OI in the northern Chukchi Sea and south of Cape Hope, resulting in approximately 10-cm smaller SWH errors throughout the entire domain. A local maximum in the a4DVAR peak period errors is also observed southwest of Cape Hope (Fig. 7b) that can be partly explained by a sharper gradient in the peak period field of the true solution (Fig. 7a).

The OI solution demonstrates a slightly better skill in forecasting the wave direction (the mean difference of 13.1° vs 16.9°). However, in the OI assimilation experiments with other types of SWH data this number varied within 13°–13.3° and was quite close to the respective characteristic (13.2°) of the background solution.

In general, our experiments have shown that the OI method tends to improve the amplitude of the spectrum and that it has only a slight impact on its shape and position in the frequency-direction coordinates. In contrast, the a4DVAR technique is capable of improving these characteristics as well, since it performs optimization along the most persistent dynamical modes of the governing (10). This important property of the a4DVAR algorithm provides a significantly better approximation of the true solution and an improved forecast skill.

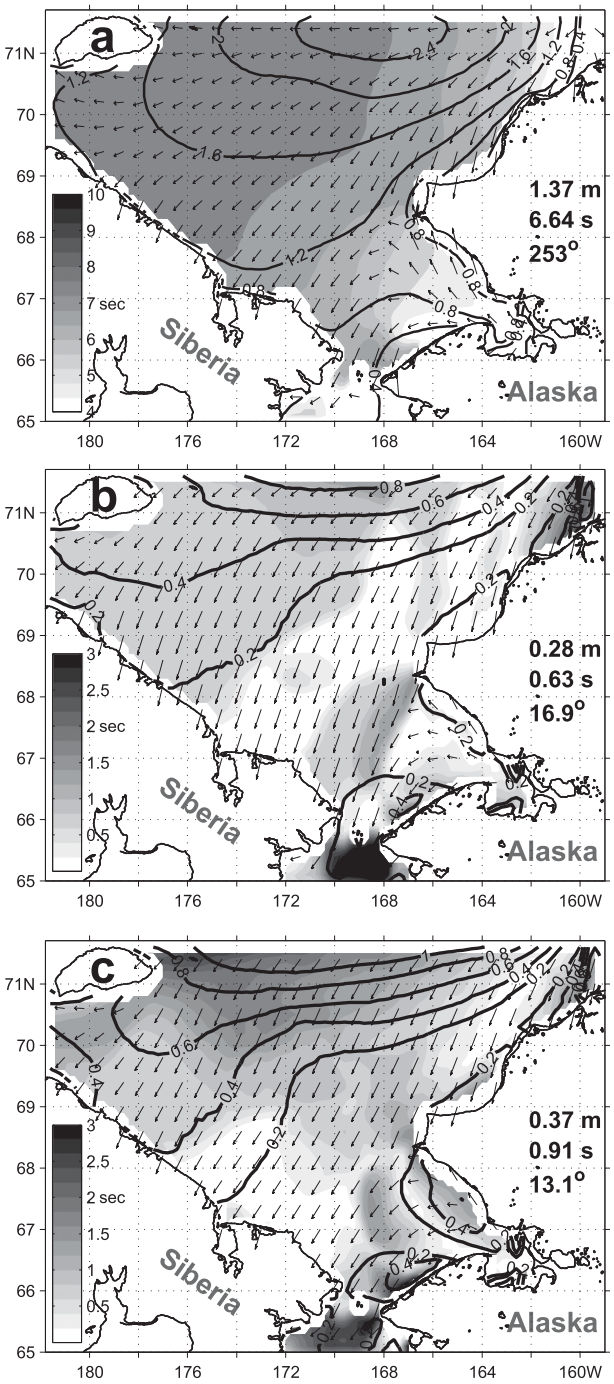

FIG. 7. Time-averaged (3–9 h) values of SWH (contours, m), peak period (shading, s), and wave direction (arrows) of (a) the true state and the respective absolute differences of the (b) HF-optimized and (c) oHF solutions from the true state. Domain-averaged values of the fields are shown on the right.

4. Conclusions and discussion

In this study, the adjoint-free variational data assimilation algorithm of Yaremchuk et al. (2009) has been applied to the fully nonlinear regional WAM configured in the Chukchi Sea. Performance of the method was tested in a series of twin-data experiments with an objective to assess the efficiency of various wave-observing platforms in reconstruction of the surface waves on a regional scale. Results indicate that an HF radar monitoring of the SWH field is more efficient than a mooring measuring the wave spectrum at a single point. However, augmenting spectral observations from the mooring with satellite SWH data boosts the system's efficiency to the level of an HF radar, which does not gain as much new information from satellite data.

Application of the adjoint-free variational method to optimization of the initial conditions is a new development in oceanic wave modeling that provides a pathway to bypass the burden of development and maintenance of the tangent linear and adjoint code. Besides, the adjoint-free approach is insusceptible to intrinsic instabilities arising in the adjoint model when the parent model solution is contains nonlinear instabilities. The a4DVAR method is based on probing the shape of the cost function by a sequence of ensembles updated in the iterative process. In this study we utilized the simplest possible technique of generating the ensemble members by the EOF analysis of model solutions. Although our experiments demonstrated a reasonably good performance of the method, more advanced schemes, such as building the search subspaces (spanned by the ensemble members) on the cost function gradients, could be developed to achieve faster convergence.

An extremely important issue of the a4DVAR development in application to wave modeling is optimization of the wind forcing, whose uncertainties provide the major contribution to model errors at time scales larger than a few hours. One possible solution is to employ wind error statistics and augment the search subspaces (ensembles) by new dimensions spanned by the leading EOFs of the wind error fields. However, extension of the a4DVAR technique to wind correction capability is a separate theoretical issue lying beyond the scope of the present paper.

An interesting finding of the experiments with synthetic data is the important role of satellite data (sporadic in the setting considered) in complementing spectral observations at a single or sparsely located moorings. At the same time, the revealed inadequacy of assimilating solely *Envisat* data (rows 13 and 14 in Table 1) is apparently an artifact of the space–time size of the domain and substantial wind forcing errors. In a series of experiments with $S_{\text{wind}} = 0.1$, satellite data provided much better improvement of the initial conditions, although not at the levels of HFR observations.

In the experiments we did not consider all types of data available for regional modeling. In particular, directional spectra (with a coarser resolution and larger uncertainties) can be obtained from high-frequency radars (Graber and Heron 1997; Hisaki 2005) and spaceborne (Tison et al. 2008; Ren et al. 2010) platforms. Simplifying assumptions have been also made in configuring WAM: the background currents ($\sim$0.2–0.4 m s^{-1}) were assumed to be negligible in comparison with the wave velocities ($\sim$5–10 m s^{-1}) of the major spectral peak and stationary conditions at the open boundaries were specified due to relatively short period of model integration. These assumptions can be relaxed in the experiments with real data spanning longer time periods in smaller domains.

The 4D variational (either adjoint or adjoint free) methods of data assimilation require multiple model runs and are certainly more expensive than sequential methods (e.g., based on optimal interpolation), which update the model state in the process of single integration. The 4DVAR methods do, however, deliver better forecast skill and are therefore competitive with sequential algorithms when the computational cost is not a primary issue. In our case, one assimilation run was performed at the expense of approximately 500 direct model runs, which is compatible to the cost of a typical adjoint 4DVAR run. The major advantage of the proposed adjoint-free technique is the absence of necessity to develop and maintain the adjoint and tangent linear codes for the assimilative model.

Results of the present study can be used in the design of the wave monitoring systems supported by the variational data assimilation into state-of-the-art wave models (Booij et al. 1996; Tolman and the WAVEWATCH III Development Group 2014) lacking the comprehensive adjoint and tangent linear codes. Nowadays, development of such systems becomes particularly important in the Arctic seas, which experience a significant surge in SWH variability associated with ongoing ice retreat induced by climate change (e.g., Francis and Atkinson 2012; Thomson and Rogers 2014).

Acknowledgments. This study was supported by the International Arctic Research Center, NSF Grants 1107925 and 1203740. It was also supported by the Office of Naval Research (Program Element 0602435N, project ''Adjoint-free 4DVAR for ocean models'' by a grant from the Gulf of Mexico Research Initiative

through the Consortium for Advanced Research on Transport of Hydrocarbon in the Environment (CARTHE) and by Grant 2013-220-04-157 from the government of the Russian Federation.

APPENDIX

Hessian Projection and Orthogonalization

In the present study, each iteration of the a4DVAR minimization starts with the reference model run from the initial condition $\mathbf{c}$ obtained at the previous iteration. This run produces the a sequence of model states subject to principal orthogonal decomposition analysis. The leading modes $\mathbf{e}_n$, $n = 1, \ldots, N$ ($N = 5$) of the analysis are utilized to perturb the reference initial condition $\mathbf{c} \rightarrow \mathbf{c} + \varepsilon\mathbf{e}_n$, to execute N parallel model runs, and to compute the vectors $\mathbf{z}_n = \mathbf{S}\mathbf{e}_n$, whose mutual scalar products $\mathbf{z}_n^{\mathrm{T}}\mathbf{z}_k \equiv \mathbf{e}_n^{\mathrm{T}}\mathbf{J}\mathbf{e}_k$ can be used to perform Gram–Schmidt orthogonalization in the control space (spanned by $\mathbf{e}_k$) with respect to the inner product induced by the Hessian matrix $\mathbf{J}$.

The perturbation vectors $\delta\mathbf{c}_n = \varepsilon\mathbf{e}_n$ span the search subspace, which can be parameterized by the N-dimensional vector $\boldsymbol{\alpha}$ such that

$$\delta\mathbf{c} = \mathbf{E}\boldsymbol{\alpha}, \tag{A1}$$

where $\mathbf{e}_n$ constitute the columns of the $N \times M$ matrix $\mathbf{E}$ projecting the Hessian on the search subspace. Multiplying (7) by $\mathbf{E}^{\mathrm{T}}$, and taking (A1) into account, yields the projection of the normal equations on search subspace:

$$\mathbf{E}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{S}\mathbf{E}\boldsymbol{\alpha} = \mathbf{E}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{b}. \tag{A2}$$

To solve (A2) it is necessary to compute both the Hessian projection $\mathbf{E}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{S}\mathbf{E}$ and the result of action of its square root $\mathbf{E}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}$ on $\mathbf{b}$. These computations are straightforward.

The matrix $\mathbf{S}\mathbf{E}$ is computed explicitly columnwise by taking the differences between the perturbed and unperturbed residuals $\mathbf{S}(\mathbf{c} + \varepsilon\mathbf{e}_n) - \mathbf{S}\mathbf{c} = \varepsilon\mathbf{S}\mathbf{e}_n$. Dividing the resulting N vectors $\mathbf{z}_n = \varepsilon\mathbf{S}\mathbf{e}_n$ by ε, the columns of $\mathbf{S}\mathbf{E}$ are obtained. Explicit computation of $\mathbf{E}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{S}\mathbf{E} \equiv (\mathbf{S}\mathbf{E})^{\mathrm{T}}\mathbf{S}\mathbf{E}$ is straightforward: one has to compute $N(N+1)/2$ inner products of all the possible pairs of $\mathbf{z}_n$ (e.g., Zupanski 2005).

In the similar manner, the perturbation vectors $\mathbf{e}_n$ can be orthogonalized with respect to the inner product induced by $\mathbf{J}$ if the vectors $\mathbf{z}_n$ obtained on the previous iterations are kept in memory.

To assess $\mathbf{E}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{b}$, consider a small perturbation of J, induced by $\mathbf{e}_n$:

$$\begin{aligned}\delta J_n &= \frac{1}{2}(\delta\mathbf{c}_n^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{S}\delta\mathbf{c}_n - \delta\mathbf{c}_n^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{b} - \mathbf{b}^{\mathrm{T}}\mathbf{S}\delta\mathbf{c}_n + \mathbf{b}^{\mathrm{T}}\mathbf{b}) - \frac{\mathbf{b}^{\mathrm{T}}\mathbf{b}}{2} \\ &= \frac{1}{2}\delta\mathbf{c}_n^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{S}\delta\mathbf{c}_n - \delta\mathbf{c}_n^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{b}.\end{aligned} \tag{A3}$$

The first term on the rhs of (A3) is negligible because $\delta\mathbf{c}_n = \varepsilon\mathbf{e}_n$ and $\varepsilon \ll 1$. As a consequence,

$$\delta J_n = -\varepsilon\mathbf{e}_n^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{b}. \tag{A4}$$

Collecting δJ_n into a single vector $\delta\mathbf{J}$ and dividing by ε yields the rhs of (A2):

$$-\delta\mathbf{J}/\varepsilon = \mathbf{E}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{b}. \tag{A5}$$

As a result, the system of normal equations in the search subspace can be solved at a relatively low cost if the vectors $\mathbf{z}_n$ and the respective perturbations of the cost function δJ_n are available. In the a4DVAR algorithm, these quantities are computed independently during the parallel runs of the perturbed model.

REFERENCES

Abdalla, S., J.-R. Bidlot, and P. Janssen, 2006: Global validation and assimilation of ENVISAT ASAR wave mode spectra. *Proc. SEASAR Workshop,* Frascati, Italy, European Space Agency, 6 pp. [Available online at http://earth.esa.int/cgi-bin/confseasar19bd.html?abstract=176.]

Aouf, L., and J.-M. Lefevre, 2006: Assimilation of directional wave spectra in the wave model WAM: An impact study from synthetic observations in preparation for the SWIMSAT satellite mission. *J. Atmos. Oceanic Technol.,* **23,** 448–463, doi:10.1175/JTECH1861.1.

Barzel, G., and R. B. Long, 1994: Wave model fitting using the adjoint technique. *Dynamics and Modeling of Ocean Waves,* G. Komen et al., Eds., Cambridge University Press, 447–453.

Bauer, E., S. Hasselmann, and K. Hasselmann, 1992: Validation and assimilation of Seasat altimeter wave heights using the WAM wave model. *J. Geophys. Res.,* **97,** 12 671–12 682, doi:10.1029/92JC01056.

——, K. Hasselmann, I. R. Young, and S. Hasselmann, 1996: Assimilation of wave data into the wave model WAM using an impulse response function method. *J. Geophys. Res.,* **101,** 3801–3816, doi:10.1029/95JC03306.

Booij, N., L. H. Holthuijsen, and R. C. Ris, 1996: The "SWAN" wave model for shallow water. *Coastal Engineering 1996: Proceedings of the Twenty-Fifth International Conference,* B. L. Edge, Ed., Vol. 1, American Society of Civil Engineers, 668–676.

Breivik, L.-A., and M. Reistad, 1994: Assimilation of *ERS-1* altimeter wave heights in an operational numerical wave model. *Wea. Forecasting,* **9,** 440–451, doi:10.1175/1520-0434(1994)009<0440:AOAWHI>2.0.CO;2.

Cardone, V. J., H. C. Graber, R. E. Jensen, S. Hasselmann, and M. J. Caruso, 1995: In search of the true surface wind field in SWADE IOP-1: Ocean wave modeling perspective. *Global Atmos. Ocean Syst.,* **3,** 107–150.

——, R. E. Jensen, D. T. Resio, V. R. Swail, and A. T. Cox, 1996: Evaluation of contemporary ocean wave models in rare extreme events: The ''Halloween storm'' of October 1991 and the ''storm of the century'' of March 1993. *J. Atmos. Oceanic Technol.,* **13,** 198–230, doi:10.1175/1520-0426(1996)013<0198:EOCOWM>2.0.CO;2.

Cavaleri, L., and Coauthors, 2007: Wave modelling: The state of the art. *Prog. Oceanogr.,* **75,** 603–674, doi:10.1016/j.pocean.2007.05.005.

De Las Heras, M. M., G. Burgers, and P. A. E. M. Janssen, 1994: Variational wave data assimilation in a third-generation wave model. *J. Atmos. Oceanic Technol.,* **11,** 1350–1369, doi:10.1175/1520-0426(1994)011<1350:VWDAIA>2.0.CO;2.

——, ——, and ——, 1995: Wave data assimilation in the WAM wave model. *J. Mar. Syst.,* **6,** 77–85, doi:10.1016/0924-7963(94)00019-8.

Dunlap, E. M., R. B. Olsen, L. Wilson, S. De Margerie, and R. Lalbeharry, 1998: The effect of assimilating ERS-1 fast delivery wave data into the North Atlantic WAM model. *J. Geophys. Res.,* **103,** 7901–7915, doi:10.1029/97JC02570.

Emmanouil, G., G. Galanis, and G. Kallos, 2010: A new methodology for using buoy measurements in sea wave data assimilation. *Ocean Dyn.,* **60,** 1205–1218, doi:10.1007/s10236-010-0328-9.

——, ——, and ——, 2012: Combination of statistical Kalman filters and data assimilation for improving ocean waves analysis and forecasting. *Ocean Modell.,* **59–60,** 11–23, doi:10.1016/j.ocemod.2012.09.004.

Francis, O. P., and D. E. Atkinson, 2012: Synoptic forcing of wave state in the southeast Chukchi Sea, Alaska, at an offshore location. *Nat. Hazards,* **62,** 1169–1189, doi:10.1007/s11069-012-0142-4.

——, G. G. Panteleev, and D. Atkinson, 2011: Ocean wave conditions in the Chukchi Sea from satellite and *in situ* observations. *Geophys. Res. Lett.,* **38,** L24610, doi:10.1029/2011GL049839.

Francis, P. E., and R. A. Stratton, 1990: Some experiments to investigate the assimilation of SEASAT altimeter wave height data into a global wave model. *Quart. J. Royal Meteor. Soc.,* **116,** 1225–1251, doi:10.1002/qj.49711649512.

Graber, H. C., and M. L. Heron, 1997: Wave height measurements from HF radar. *Oceanography,* **10,** 90–92.

Greenslade, D. J. M., 2001: The assimilation of ERS-2 significant wave height data in the Australian region. *J. Mar. Syst.,* **28,** 141–160, doi:10.1016/S0924-7963(01)00005-7.

Hasselmann, S., P. Lionello, and K. Hasselmann, 1997: An optimal interpolation scheme for the assimilation of spectral wave data. *J. Geophys. Res.,* **102,** 15 823–15 836, doi:10.1029/96JC03453.

Hayami, K., J.-F. Yin, and T. Ito, 2007: GMRES methods for least squares problems. National Institute of Informatics Tech. Rep. NII-2007-09E, 28 pp. [Available online at http://www.nii.ac.jp/TechReports/07-009E.pdf.]

Hersbach, H., 1997: The adjoint of the WAM model. KNMI Scientific Rep. WR 97-01, 28 pp.

Hisaki, Y., 2005: Ocean wave directional spectra estimation from an HF radar with a single antenna array: Observation. *J. Geophys. Res.,* **110,** C11004, doi:10.1029/2005JC002881.

Holthuijsen, L. H., N. Booij, M. van Endt, S. Caires, and C. Guedes Soares, 1997: Assimilation of buoy and satellite data in wave forecasts with integral control variables. *J. Mar. Syst.,* **13,** 21–31, doi:10.1016/S0924-7963(96)00121-2.

Janssen, P. A., 2008: Progress in ocean wave forecasting. *J. Comput. Phys.,* **227,** 3572–3594, doi:10.1016/j.jcp.2007.04.029.

Lionello, P., H. Günther, and P. A. E. M. Janssen, 1992: Assimilation of altimeter data in a global third-generation wave model. *J. Geophys. Res.,* **97,** 14 453–14 474, doi:10.1029/92JC01055.

Monbaliu, J., R. Padilla-Hernández, J. C. Hargreaves, J. C. Carretero Albiach, W. Luo, M. Sclavo, and H. Günther, 2000: The spectral wave model, WAM, adapted for applications with high spatial resolution. *Coastal Eng.,* **41,** 41–62, doi:10.1016/S0378-3839(00)00026-0.

Orzech, M. D., J. Veeramony, and H. Ngodock, 2013: A variational assimilation system for nearshore wave modeling. *J. Atmos. Oceanic Technol.,* **30,** 953–970, doi:10.1175/JTECH-D-12-00097.1.

——, ——, and S. Flampouris, 2014: Optimizing spectral wave estimates with adjoint-based sensitivity maps. *Ocean Dyn.,* **64,** 487–505, doi:10.1007/s10236-014-0700-2.

Ren, L., Z. Mao, H. Huang, and F. Gong, 2010: Satellite-based RAR performance simulation for measuring ocean wave spectrum based on SAR inversion spectrum. *Acta Oceanol. Sin.,* **29,** 13–20, doi:10.1007/s13131-010-0046-7.

Sannasiraj, S. A., V. Babovic, and E. S. Chan, 2006: Wave data assimilation using ensemble error covariances for operational wave forecast. *Ocean Modell.,* **14,** 102–121, doi:10.1016/j.ocemod.2006.04.001.

Snyder, R. L., L. M. Lawson, and R. B. Long, 1992: Inverse modeling of the action-balance equation. Part I: Source expansion and adjoint-model equations. *J. Phys. Oceanogr.,* **22,** 1540–1555, doi:10.1175/1520-0485(1992)022<1540:IMOTAB>2.0.CO;2.

Thomson, J., and E. Rogers, 2014: Swell and sea in the emerging Arctic Ocean. *Geophys. Res. Lett.,* **41,** 3136, doi:10.1002/2014GL059983.

Tison, C., G. Carayon, J. Lambin, P. Castillan, J. C. Souyris, and D. Hauser, 2008: A spaceborne radar for directional wave spectrum estimation: First performance simulations. *2008 IEEE International Geoscience and Remote Sensing Symposium: Proceedings,* Vol. 1, IEEE, I-347–I-350, doi:10.1109/IGARSS.2008.4778865.

Tolman and the WAVEWATCH III Development Group, 2014: User manual and system documentation of WAVEWATCH III version 4.18. NOAA/NWS/NCEP Tech. Note, 311 pp. [Available online at http://polar.ncep.noaa.gov/waves/wavewatch/manual.v4.18.pdf.]

Trevisan, A., M. D. Isidoro, and O. Talagrand, 2010: Four-dimensional variational assimilation in the unstable subspace and the optimal subspace dimension. *Quart. J. Roy. Meteor. Soc.,* **136,** 487–496, doi:10.1002/qj.571.

Veeramony, J., D. Walker, and L. Hsu, 2010: A variational data assimilation system for nearshore applications of SWAN. *Ocean Modell.,* **35,** 206–214, doi:10.1016/j.ocemod.2010.07.008.

Voorrips, A. C., V. K. Makin, and S. Hasselmann, 1997: Assimilation of wave spectra from pitch-and-roll buoys in a North Sea wave model. *J. Geophys. Res.,* **102,** 5829–5849, doi:10.1029/96JC03242.

——, A. W. Heemink, and G. J. Komen, 1999: Wave data assimilation with the Kalman filter. *J. Mar. Syst.,* **19,** 267–291, doi:10.1016/S0924-7963(98)00094-3.

Vuik, K., A. G. J. Sevink, and G. C. Herman, 1996: A preconditioned Krylov subspace method for the solution of least squares problems in inverse scattering. *J. Comput. Phys.,* **123,** 330–340, doi:10.1006/jcph.1996.0027.

Walker, D. T., 2006: Assimilation of SAR imagery in a nearshore spectral wave model. General Dynamics Advanced Information Systems Final Rep. 200236-F, 33 pp.

WAMDI Group, 1988: The WAM model—A third generation wave prediction model. *J. Phys. Oceanogr.,* **18,** 1775–1810, doi:10.1175/1520-0485(1988)018<1775:TWMTGO>2.0.CO;2.

Waters, J., L. R. Wyatt, J. Wolf, and A. Hines, 2013: Data assimilation of partitioned HF radar wave data into Wavewatch III. *Ocean Modell.,* **72,** 17–31, doi:10.1016/j.ocemod.2013.07.003.

Wittmann, P. A., and J. A. Cummings, 2004: Assimilation of altimeter wave measurements into WAVEWATCH III. *Eighth Int. Workshop on Wave Hindcasting and Forecasting,* North Shore, Oahu, Hawaii, U.S. Army Engineer Research and Development Center, Environment Canada, and JCOMM, F3. [Available online at http://www.waveworkshop.org/8thWaves/Papers/F3.pdf.]

Yaremchuk, M., and S. Smith, 2011: On the correlation functions associated with polynomials of the diffusion operator. *Quart. J. Roy. Meteor. Soc.,* **137,** 1927–1932, doi:10.1002/qj.893.

——, D. Nechaev, and G. G. Panteleev, 2009: A method of successive corrections of the control subspace in the reduced-order variational data assimilation. *Mon. Wea. Rev.,* **137,** 2966–2978, doi:10.1175/2009MWR2592.1.

Zupanski, M., 2005: Maximum likelihood ensemble filter: Theoretical aspects. *Mon. Wea. Rev.,* **133,** 1710–1726, doi:10.1175/MWR2946.1.