Analysis and Forecasting of Sea Ice Conditions with Three-Dimensional Variational Data Assimilation and a Coupled Ice–Ocean Model

ALAIN CAYA AND MARK BUEHNER
Meteorological Research Division, Atmospheric Science and Technology Directorate, Environment Canada, Dorval, Quebec, Canada

TOM CARRIERES
Marine and Ice Services Division, Meteorological Service of Canada, Environment Canada, Ottawa, Ontario, Canada

(Manuscript received 23 March 2009, in final form 13 July 2009)

ABSTRACT
A three-dimensional variational data assimilation (3DVAR) system has been developed to provide analyses of the ice–ocean state and to initialize a coupled ice–ocean numerical model for forecasting sea ice conditions. This study focuses on the estimation of the background-error statistics, including the spatial and multivariate covariances, and their impact on the quality of the resulting sea ice analyses and forecasts. The covariances are assumed to be horizontally homogeneous and fixed in time. The horizontal correlations are assumed to have a Gaussian shape and are modeled by integrating a diffusion equation. A relatively simple implementation of the ensemble Kalman filter is used to produce ensembles of the ice–ocean model state that are representative of background error and from which the 3DVAR covariance parameters are estimated.

Data assimilation experiments, using various configurations of 3DVAR and simpler assimilation approaches, are conducted over a 7-month period during the winter of 2006/07 for the Canadian east coast region. The only data assimilated are the gridded daily ice charts and RADARSAT image analyses produced by the Canadian Ice Service. All of the data assimilation experiments produce significantly improved short-term forecasts as compared with persistence. When assimilating the same data, the forecast quality from the experiments employing either the 3DVAR, direct insertion, or nudging is quite similar. However, assimilation of both the daily ice charts and RADARSAT image analyses in 3DVAR results in significant improvements to the sea ice concentration forecasts. This result supports the use of a data assimilation approach, such as 3DVAR, for combining multiple sources of observational data together with a sophisticated forecast model to provide analyses and forecasts of sea ice conditions.

1. Introduction
Accurate and timely sea ice information in the Northern Hemisphere is greatly needed because of the increasing focus on northern regions for reasons of both economic development and national sovereignty. The changing Arctic climate, as demonstrated by record low ice coverage in recent years, is leading to increased marine transportation and natural resource development in and around ice-infested waters. These changes to the climate also represent a challenge for traditional northern pop-
on a daily basis. Many users rely on these analyses for the planning of safe and efficient marine activities. The CIS ice analyses are also used to specify sea ice parameters required for operational NWP forecasts produced at CMC. The ice analyses are produced manually by analysts who subjectively combine information from numerous available data sources, including satellite imagery (e.g., synthetic aperture radar, passive microwave, visible, and infrared), ship and aircraft reports, and the analysis from the previous day. Because this subjective analysis process is highly labor intensive and must be completed within a limited time period, the analysts focus most of their effort on high-priority areas, such as those with ongoing shipping or ice-breaking activity. Consequently, the accuracy of the analyses is difficult to quantify and may not be consistent throughout the domain. Also, because the ice analysis is performed without considering the ocean and atmospheric state, it may not be optimally suited for initializing coupled numerical model forecasts.

Data assimilation techniques have been widely used to objectively combine numerous data sources with different spatial resolutions and accuracies for both atmospheric and ocean analysis/prediction applications (see, e.g., Kalnay 2003 for an overview of techniques). Recently, these techniques have also been applied to the objective analysis of sea ice data. Here, a few selected studies are highlighted, and more details are given in section 2 concerning the techniques explored in this study, three-dimensional variational data assimilation (3DVAR) and nudging. Van Woert et al. (2004) evaluated the quality of ice concentration forecasts from a coupled ice–ocean model covering much of the Northern Hemisphere that is initialized using the simple data assimilation technique direct insertion applied to ice concentration derived from Special Sensor Microwave Imager (SSM/I) data. They showed that the forecast accuracy relative to persistence varied throughout the year, with the lowest-quality forecasts being obtained during the winter freeze-up period. Several studies have focused on the assimilation of ice motion estimated from processing sequential satellite images. In the studies of Meier et al. (2000) and Meier and Maslanik (2003), optimal interpolation was used to assimilate ice motion estimated from SSM/I imagery to initialize a stand-alone ice model for the Arctic. The assimilation improved the agreement of the modeled ice motion with buoy observations and also substantially altered the average ice thickness in some regions. Meier and Maslanik (2003) also estimated ice motion error statistics as a function of environmental conditions from a free model simulation, from observations, and from analyses obtained by assimilating the observations. Similarly, using a coupled ice–ocean model of the Arctic Ocean, Zhang et al. (2003) assimilated ice motion from buoy motion and SSM/I imagery using optimal interpolation. They showed that the assimilation improved both ice motion and thickness.

Using the simpler approach of nudging and a similar model as Zhang et al. (2003), Lindsay and Zhang (2006) assimilated ice concentration where the weight given to the data depended on the model–observation misfit. In a separate step, they also assimilated ice motion using optimal interpolation as in Zhang et al. (2003) while adjusting the ice thickness to minimize changes in ice mass. They found that, relative to observations, the ice extent and the simulated ice draft were both significantly improved by this assimilation procedure. Stark et al. (2008) assimilated ice concentration and motion from satellite data in addition to a full set of ocean data using an optimal interpolation scheme and a global ice–ocean model. Although they adjusted ocean salinity in response to changes made to the ice mass, the ocean temperature was not directly modified by the assimilation of ice data. They obtained significant improvements for either ice concentration or ice motion from assimilating observations of only concentration or motion, respectively. However, little impact was seen on either variable from assimilating only observations of the other variable. Finally, Lysæter et al. (2003) used an ensemble Kalman filter (EnKF) applied to a coupled ice–ocean model of the Arctic Ocean to assimilate ice concentration derived from SSM/I. They demonstrated the complex spatial and multivariate covariances of the short-term forecast (background) error that could be estimated from the ensemble. These include a consistent negative covariance between ice concentration and sea surface temperature (SST) and more highly state-dependent relationships between ice concentration and salinity and between ice concentration and ice thickness.

The present study represents the first step toward developing an automated sea ice analysis system in Canada that will provide gridded objective estimates of the current sea ice conditions to satisfy the operational requirements of both CIS and CMC. The system uses a data assimilation technique (3DVAR) similar to techniques developed for NWP applications to combine multiple sources of remotely sensed data while employing a numerical forecast model to evolve information from past observations. Such advanced data assimilation techniques may be helpful regarding, for example, the problem of a lack of direct measurements of the three-dimensional ocean state under sea ice by providing a means for correcting the ocean model variables given only sea ice observations. The resulting estimates of the current sea ice conditions can then be used to initialize the sea ice component of a coupled numerical forecast model.

As part of the present study, a description of the basic data assimilation approach is given with emphasis on
the estimation and modeling of the background-error statistics. Numerical experiments are conducted in the context of a moderately high-resolution coupled ice–ocean model for the Canadian east coast region. The specific region and numerical model were chosen to facilitate the development of a prototype system; however, the software has been designed in such a way as to allow the data assimilation system to be relatively easily implemented over different geographical domains and with different numerical models (or even using persistence when a numerical model is not available). For this preliminary study, most of the experiments use only a single source of assimilated data. The CIS daily ice chart was chosen for assimilation, because it represents a synthesis of numerous sources of remotely sensed data (Carrieres et al. 1996). However, experiments are also performed in which two types of data are assimilated to demonstrate the ability of sophisticated data assimilation approaches to simultaneously extract ice information from multiple sources of observations. A 3DVAR approach is employed (e.g., Gauthier et al. 1999), with the background-error covariance parameters being estimated from an EnKF (e.g., Evensen 1994; Houtekamer and Mitchell 1998) run over a past period. The next section provides a description of the observations (used for both assimilation and verification of the analyses and forecasts), the coupled ice–ocean model, and the assimilation techniques used for the numerical experiments. Section 3 provides details on the parameterization and estimation of the background-error covariances. The results from the numerical experiments are presented in section 4. These numerical experiments include analyses and 48-h forecasts for the period between 5 December 2006 and 30 June 2007. Finally, conclusions are provided in section 5.

2. The assimilation system

a. Sea ice observational data

The observations assimilated for the purpose of this study are derived from the daily ice charts prepared in near–real time by CIS (Carrieres et al. 1996). Recent ice charts are available in graphical form on the CIS Web site (available online at http://ice-glaces.ec.gc.ca). These charts are subjective analyses of the instantaneous ice concentration and its stage of development. The daily ice charts may be considered as nowcasts, because ice information from observations is projected forward in time to 1800 UTC each day. Several types of remotely sensed data, ship and aircraft reports, and the analysis from the previous day all feed into the process. The stage of development is estimated by subjectively blending several types of information, including floe shape from satellite imagery, floe shape and color from helicopter-based visual observations, empirical relationships between freezing-degree days and the growth of level ice thickness, climatology, and the ice chart from the previous day. From the stage of development information, the observed distribution of ice thickness is inferred (using the ice thickness ranges in the World Meteorological Organization definitions for each stage of development) and then mapped onto the ice thickness categories used by the sea ice model. The partial concentrations of these ice thickness categories from the CIS daily ice charts are used for assimilation. This is convenient, because the data are available over the entire model domain (where there is sea ice), and the analysis time is chosen to coincide with the valid time for the ice charts (1800 UTC).

It was decided not to solely rely on the daily ice charts for evaluating the accuracy of analyses and forecasts produced by the data assimilation system and the ice–ocean model. This was done because of the heterogeneous accuracy of the ice charts and the tendency of the ice charts to resemble a persistence forecast in regions of low priority for the ice analysts. To avoid these problems, the subjective analyses of RADARSAT synthetic aperture radar images (which are an important input to the production of the daily ice charts) were used for verification in those regions and times where the image analyses were available. The image analyses contain the same type of information as the CIS daily ice charts from which an ice thickness distribution can be inferred. Though these data only cover a small portion of the model domain for any given time, the accuracy is expected to be relatively uniform over the image analysis. The total volume of data from image analyses is only approximately 15% of the volume from the daily ice charts. Also, generally only the accuracy of the resulting model forecasts and not the analyses are evaluated. Consequently, by using RADARSAT image analyses instead of the daily ice chart, the data used for verification are independent of the assimilated data involved in initializing the forecast.

Although most of the data assimilation experiments only use the CIS daily ice charts in the analysis procedure, two additional 3DVAR experiments were performed that also assimilate the CIS RADARSAT image analyses in combination with the daily ice charts. The daily ice charts are produced using the RADARSAT image analyses. Consequently, this is effectively equivalent to assimilating the same information twice. However, because the daily ice charts are assimilated assuming the same level of accuracy everywhere in the domain, the assimilation of the image analyses in combination with
the daily charts ensures that more weight is given to the observations (by assimilating twice as many observations) in the areas where the observations have a higher level of accuracy: that is, in the regions and at times when the image analyses are available. The daily ice charts and RADARSAT image analyses are assimilated using the same value for their observation-error variances.

b. Coupled ice–ocean model

The ice–ocean model used in this study is the Community Ice–Ocean Model (CIOM) as described by Yao et al. (2000). It is composed of the Princeton Ocean Model (Blumberg and Mellor 1987; Mellor 1996) and a multicategory sea ice model. The model has a grid resolution of $1/58$ longitude and $1/68$ latitude covering the horizontal domain in the northwest Atlantic Ocean shown by the light shaded area in Fig. 1. The prognostic ocean variables are the three-dimensional velocity, potential temperature, salinity, and turbulence quantities on 16 sigma levels unevenly spaced in the vertical. The ice model uses an Eulerian description of the ice thickness distribution and the prognostic ice variables are velocity and the partial ice concentrations of 24 thickness categories plus the open water fraction. The model configuration is essentially the same as described by Yao et al. (2000), except for the shortwave radiation; here, the clear-sky solar radiation of Shine et al. (1984) with the cloud factor of Laevastu (1960) is used. Atmospheric conditions used to force the CIOM are taken from the Canadian Global Environmental Multiscale (GEM) model (Côté et al. 1998) short-term regional forecasts refreshed every 3 h.

c. Variational assimilation method

The 3DVAR approach is employed to assimilate observations to correct a short-term forecast produced by the coupled ice–ocean model and used as the background state. The corrected model state (referred to as the analysis) is then used as the initial conditions to produce the next short-term forecast that serves as the background state for producing the subsequent analysis. In this study, the analyses are separated by 24 h; therefore, a 24-h forecast is used to produce each background state, starting from the previous analysis.

The data assimilation problem is cast in terms of minimizing a cost function that measures the fit of the analysis to the observations and to the background state. The standard 3DVAR cost function has the form

$$J(x) = \frac{1}{2}(x - x^b)^T B^{-1} (x - x^b)$$

$$+ \frac{1}{2}[y - H(x)]^T R^{-1} [y - H(x)],$$

where \(y\) is the observation vector, \(H\) is the (possibly nonlinear) observation operator that maps the model state into the space of the observations (including both spatial and possibly temporal interpolation), \(x^b\) is the background state, \(B\) is the background-error covariance matrix, and \(R\) is the observation-error covariance matrix. Under certain assumptions, the state \(x\) that minimizes (1) is also the maximum likelihood estimate (Lorenc 1986). For practical reasons, the cost function is transformed into its incremental form (Courtier et al. 1994) and also preconditioned with respect to the background term, resulting in the cost function

$$J(\xi) = \frac{1}{2} \xi^T \xi + \frac{1}{2}[y - H(x^b) - HB^{1/2} \xi]^T$$

$$\times R^{-1} [y - H(x^b) - HB^{1/2} \xi],$$

where \(\xi\) is referred to as the control vector and \(H\) is the tangent linear observation operator: that is, the operator \(H\) linearized about \(x^b\). Note that, because of the preconditioning, the inverse of \(B\) is not required, only the square root. In deriving (2), the approximation \(H(x^b + \Delta x) \approx H(x^b) + H \Delta x\) has been made. This approximation
results in \( J \) being a quadratic function with respect to \( \xi \); therefore, a unique global minimum always exists. The analysis \( x^a \) is related to the control vector by

\[
x^a = x^b + B^{1/2} \xi^a,
\]

where \( \xi^a \) is the value of \( \xi \) that minimizes (2). The minimization of (2) is performed using a quasi-Newton descent algorithm. Because the innovation vector \( y = H(x^a) - H(x^b) \) does not depend on the control vector \( \xi \), it is only evaluated once at the beginning of the minimization, using the full observation operator. In the general case, \( H(x^a) \) is computed with the background state interpolated to the time of the observations using output from the nonlinear forecast model at several time steps throughout the 24-h assimilation window. Conversely, the linearized operator \( H \) is formulated under the assumption that all of the observations are valid at the center of the assimilation window. Consequently, the approach is still considered a three-dimensional assimilation scheme but is referred to as first guess at appropriate time (3D-FGAT).

For most of the experiments in the present study, however, only the CIS daily ice charts valid at 1800 UTC are assimilated and therefore time interpolation of the background state is not required. For two experiments, the RADARSAT image analyses produced by CIS are also assimilated. Because image analyses have a valid time equal to the valid time of the original RADARSAT image, which is generally not 1800 UTC, the FGAT approach is important when assimilating these data.

In this implementation, the error covariance matrices \( B \) and \( R \) are assumed to be time independent. To further simplify the problem, the observation errors are assumed uncorrelated, yielding a diagonal matrix \( R \), which is trivial to invert. These assumptions are also made for most implementations of variational data assimilation approaches for NWP. More detail regarding the background-error covariances is given in section 3.

d. Newtonian relaxation assimilation method

Newtonian relaxation, also known as nudging, is a simple data assimilation method that has been used to initialize operational ice–ocean model forecasts at CIS. In this approach a term proportional to the difference between the modeled and observed values is added to the corresponding model equation that governs the time evolution of the variable. In the present case, a term in the equation for the ice thickness distribution \( g(h) \) is added, resulting in the modified equation

\[
\frac{\partial}{\partial t} g + V \cdot (u g) = -\frac{\partial}{\partial h} (f g) + \psi + \text{diffusion} + \frac{2(g_o - g)}{(t_o - t)},
\]

where \( h \) is the ice thickness, \( f \) is the thermodynamic vertical growth rate of ice, \( u \) is the ice velocity vector, \( g_o \) is the observed ice thickness distribution, and \( t_o \) is the time of the observation. The mechanical redistribution function \( \psi \) represents the creation of open water and the ridging of ice during deformation. The last term in (4) is the nudging term that is active only during a 24-h period until one time step before the time of the observation and therefore remains finite. After each model time step, the ice thickness distribution \( g(h) \) is renormalized such that the total ice concentration plus open water fraction is kept equal to 1.0 at each model grid point. At the end of this nudging period, direct insertion is performed to replace the model values with the observed ice thickness distribution.

The nudging approach is relatively simple to implement, but it requires that only a single set of observations be assimilated and that they are of the same physical quantity and on the same grid as the model. Although this requirement is satisfied in the present case by assimilating only information from the gridded CIS daily ice chart (after horizontal interpolation to the model grid), it will not be the case when assimilating multiple sources of remotely sensed data. There is also no way for the observation of one physical variable (e.g., ice concentration) to directly impact on the analysis of another unobserved physical variable (e.g., ocean salinity and temperature). It is also not possible to use the nudging approach to assimilate observed variables that have only an indirect or nontrivial relationship with the variables to be analyzed. An example of this is the assimilation of the brightness temperatures from satellite-based passive microwave sensors to estimate sea ice and other surface and atmospheric properties.

3. Background-error covariance matrix

The background-error covariances are an essential aspect of any data assimilation system. They partially govern the level of influence the observations have on the analysis and how the information is spread both spatially and to other model variables that are not directly related to the observations. One method that can estimate background-error statistics is the EnKF (Evensen 1994; Houtekamer and Mitchell 1998). In the EnKF, the background-error statistics are estimated by evolving an ensemble of states, which are periodically updated using observations and statistical information regarding all known sources of error (e.g., from the observations, model, forcing data). However, the EnKF is computationally costly with a large ensemble and is prone to sampling errors for small ensembles. The idea here is to use a simpler, parameterized form of the
background-error covariance matrix for the 3DVAR system and to estimate the parameters using the EnKF background-error statistics obtained for a past period. Furthermore, for the first version of this system the background-error statistics used in the 3DVAR are assumed to be stationary. Consequently, the EnKF error statistics used to estimate the required covariance parameters are first averaged over time, thus further reducing the effect of sampling error. However, the background-error covariances may have significant seasonal variations, as shown by Lisæter et al. (2003), and this will be evaluated in a future study.

a. Parameterization of the background-error covariance matrix

The background-error covariance matrix and its square root can be expressed as

\[ \mathbf{B} = \mathbf{D} \mathbf{C} = \mathbf{D} \mathbf{C}^{1/2} (\mathbf{D} \mathbf{C}^{1/2})^T = \mathbf{B}^{1/2} (\mathbf{B}^{1/2})^T, \]

where \( \mathbf{D} \) is the diagonal matrix of background-error standard deviations and \( \mathbf{C} \) is the correlation matrix. By assuming separability between the horizontal and vertical correlations, the matrix \( \mathbf{C} \) can be decomposed into vertical and horizontal correlation matrices \( \mathbf{C} = \mathbf{C}_v \mathbf{C}_h \). In the present case, \( \mathbf{C}_v \) also contains the correlations between the partial ice concentrations for the 24 ice thickness categories. The horizontal correlations are incorporated in the 3DVAR data assimilation system using a diffusion operator (Weaver and Courtier 2001). A wide variety of correlation functions can be modeled using this approach. For simplicity, because the solution to the basic diffusion equation is a Gaussian function, a Gaussian correlation function can be employed efficiently by integrating

\[ \frac{d\mathbf{\eta}}{dt} = \mathbf{V}(\kappa \mathbf{V}\mathbf{\eta}), \]

where \( \kappa \) is the diffusion coefficient and \( \mathbf{\eta} \) is any scalar field that is a function of space and time. Integration of the spatially discretized version of Eq. (6) starting from a specified initial field of values is equivalent to multiplying this initial field by a correlation matrix with spatial correlations defined by the Gaussian function. The use of a diffusion equation to implement horizontal background-error correlations in a data assimilation system is both computationally efficient and has the benefit of being able to account for lateral boundaries. The latter feature is necessary to prevent observations on one side of an island or isthmus from directly affecting the analysis on the other side. Therefore, this approach may be especially advantageous in regions with many islands and narrow passages, such as the Canadian Arctic Archipelago.

The length scale \( L_c \) of the correlation function (as defined by Daley 1991) is given by

\[ L_c = \sqrt{2\kappa T_d}, \]

where \( T_d \) is the time over which Eq. (6) is integrated. In practice, \( \kappa \) can arbitrarily be set to \( L_c^2 \), resulting in an integration time of 0.5 from (7). Then, the only remaining parameter to be specified is the time step length, which is set small enough to ensure numerical stability. The square root of the horizontal correlation matrix, required to specify the square root of \( \mathbf{B} \), is given by

\[ \mathbf{C}_h^{1/2} = \mathbf{L}^{1/2} \mathbf{W}^{-1/2}, \]

where \( \mathbf{L} \) denotes the integration of the diffusion Eq. (6) and \( \mathbf{W} \) is a diagonal matrix defined in such a way as to ensure the equality

\[ \mathbf{L} = \mathbf{W}^{-1} \mathbf{L}^T \mathbf{W}. \]

In the present case, where \( \mathbf{V}^2 \) is the Laplacian operator applied to a latitude–longitude grid on the sphere, the \( i \)th diagonal element of \( \mathbf{W} \) is the cosine of the corresponding latitude,

\[ W_i = \cos(\lambda_i). \]

The diagonal matrix \( \mathbf{L} \) contains normalization factors computed numerically to ensure that the diagonal elements of \( \mathbf{C}_h \) are equal to unity. Two approaches for computing \( \mathbf{L} \) are described in the appendix.

b. Estimation of the background-error covariance parameters

As already mentioned, the EnKF is run independently over a past period to produce ensembles of model states (background ensembles) from which the 3DVAR background-error covariance parameters are estimated. The EnKF used here is the formulation of Whitaker and Hamill (2002) that does not require perturbed observations. The EnKF is run from 31 December 2002 to 28 April 2003, assimilating observations every 24 h at 1800 UTC. The assimilated quantity is the total ice concentration from CIS daily ice charts as described in section 2a, and the observation errors are assumed to be uncorrelated with a standard deviation of 0.1. The EnKF has 48 members. To include the effect of uncertainties in the atmospheric forcing on the coupled ice–ocean model variables, 16 different atmospheric states from the ensemble prediction system run at the Meteorological Service of Canada (Houtekamer et al. 1996) are used to
specify the upper boundary conditions for the CIOM (thus, each atmospheric forcing member is used for three members of the EnKF). It is noteworthy that model error is not taken into account in this EnKF; only uncertainties in the initial conditions, in the atmospheric forcing, and in the observations determine the estimated background-error covariances. Results from two separate EnKF experiments (not shown), one with and one without the atmospheric forcing uncertainty included, showed that this source of uncertainty is responsible for the majority of the resulting background-error variance for ice concentration.

In the EnKF, the analysis step involves all prognostic model variables, except the turbulence quantities. The background-error covariances used to assimilate observations during the analysis step are calculated each day from the ensemble of 48 background states,

$$B(t) = \frac{1}{N-1} XX^T,$$  

(11)

where the columns of the matrix $X$ are the deviations of the 48 background ensemble members from the ensemble mean state valid at time $t$ and $N$ is the ensemble size. To reduce the effect of spurious covariances between state variables that are separated by large distances within the EnKF analysis step, covariance localization is applied, following Houtekamer and Mitchell (2001), by taking the Schur (or element by element) product of a compactly supported correlation function, with the ensemble sample covariances. This compactly supported correlation function is the fifth-order piecewise rational function of Gaspari and Cohn (1999). It is found that a cutoff distance of 200 km gives the lowest root-mean-square difference between the 24-h ensemble mean forecast and the observations (this corresponds to a correlation length scale of 55 km, as defined by Daley 1991).

The error statistics used in the 3DVAR system are approximated to be static in time and horizontally homogeneous on depth levels. Consequently, the correlations and standard deviations estimated from the EnKF background ensemble are temporally and horizontally averaged over all grid points where sea ice is present in any ensemble member to give more robust estimates. Where there is no ice in all ensemble members, the ensemble spread is null in terms of ice concentration. This situation happens over a large portion of the model domain. Because we are mostly interested in the sea ice, we are not using these points in the horizontal averaging, because it would greatly reduce the averaged background-error standard deviation for ice concentration. Because of a large variation in the ocean depth over the model domain, it would not be appropriate to horizontally average the standard deviations and vertical correlations of the ocean variables on the 16 sigma levels used by the model. Instead, the EnKF ensemble members are first interpolated to 32 constant depth levels before estimating the error covariances. The background-error standard deviations diagonal matrix $D$ [in Eq. (5)] and the vertical background-error correlation matrix $C_v$ [which, in combination with $C_h$ from Eq. (8), corresponds with $C$ in Eq. (5)] for the model’s 24 ice thickness categories, ocean temperature, and salinity, including all multivariate relationships, are estimated. Presently, the ocean currents and ice velocities are not included in the background-error covariances and therefore are left unchanged from the values in the background state by the 3DVAR analysis.

Figure 2a shows the estimated background-error standard deviations of the partial ice concentrations for the 24 ice thickness categories, which generally decrease with increasing thickness. This corresponds with
a background-error standard deviation for total ice concentration of 0.093. Note that the variance for total ice concentration is the summation over all elements of the covariance matrix for partial ice concentrations, some of which are negative. Figure 2b shows the background-error correlation between the partial ice concentrations for five selected thickness categories and all of the other categories. All of the correlation functions have a similar shape with categories of similar thickness being positively correlated and categories with large differences in thickness having small negative correlations. Figures 3a,b show the estimated background-error standard deviations for temperature and salinity, respectively, as functions of ocean depth. The 32 depth levels are indicated by circles on the figures. The standard deviation for both of these variables monotonically decreases in a similar way with increasing depth.

It was mentioned earlier that the coupled ice–ocean model error is not taken into account in the implementation of the EnKF. This contributes to an underestimation of the resulting background-error variance estimated from the EnKF background ensembles. For this reason, the background-error standard deviations for all variables estimated from the EnKF (as shown above) have been multiplied by a factor of 2 to provide more realistic values for specifying the 3DVAR background-error covariances. The resulting background-error standard deviation for total ice concentration used in the 3DVAR experiments is 0.186. This value is thought to be more realistic, because it is comparable to the standard deviation of differences between CIOM forecasts and RADARSAT image analyses (as shown in section 4).

Initial attempts to directly estimate the covariances between the partial concentration of each ice thickness category and either ocean temperature or salinity produced complex multivariate relationships that were difficult to interpret and led to occasionally undesirable behavior in the 3DVAR analysis. For example, when assimilating data that led to a reduction in the ice thickness while maintaining the same total ice concentration, this was accompanied by a decrease in the ocean temperature. This change to the ocean temperature led to the ice again becoming thicker during the subsequent forecast as a result of enhanced freezing. To eliminate this problem, the relationship between the partial ice concentrations and ocean temperature was simplified by assuming that corrections to the ocean temperature are only related to corrections to the total ice concentration,

$$\Delta T^k_b = -\alpha^k \sum_{i=1}^n \Delta g_i,$$  \hspace{1cm} (12)

where $T^k_b$ is the “balanced” component of temperature, $\alpha$ is a positive parameter to be determined, the superscript $k$ is an index referring to the ocean depth level, and $g_i$ is the partial ice concentration for the $i$th ice thickness category. Therefore, changes to ice thickness that do not affect the total ice concentration will have no effect on ocean temperature. The parameter $\alpha$ is estimated by a linear regression analysis applied to the EnKF background ensembles. The covariances of the “unbalanced” component of ocean temperature, which is assumed to be uncorrelated with ice concentration, is also computed from the EnKF ensembles. The ocean temperature increment in the 3DVAR analysis is computed as the sum of the balanced and unbalanced components of temperature. The same approach is also used to model the covariances between ocean salinity and ice concentration.
concentration. The approach of using such a balance relation to describe multivariate covariances is similar to the approach used in variational data assimilation systems for NWP (e.g., Gauthier et al. 1998; Derber and Bouttier 1999).

Figure 4a shows the background-error covariance between total ice concentration and ocean temperature as a function of ocean depth. The covariance is negative near the surface and rapidly approaches zero with increasing depth to become essentially negligible at depths greater than 100 m. The negative sign means that, when the assimilation of observations related to ice concentration results in an increase to the ice concentration relative to the background state, the ocean temperature will be decreased. This change to the ocean temperature will be strongest near the surface and decrease to nearly zero below 100-m depth. Similarly, Fig. 4b shows the background-error covariance between total ice concentration and ocean salinity.

It is interesting to note that, consistent with the spatially and temporally averaged covariances just discussed, an examination of the raw ensemble correlation between total ice concentration and SST for any given day during the 4-month period showed negative values almost everywhere in the horizontal domain (not shown). This is consistent with the results of Lisæter et al. (2003). The raw ensemble correlation between total ice concentration and sea surface salinity (SSS) is also negative almost everywhere in the horizontal domain (not shown) during the same 4-month period. This is consistent with the explanation of Lisæter et al. (2003) for the case when advection dominates the effect of thermodynamic ice melting and growing, which is the case along the Labrador coast during the period when the ice edge is progressing southward.

The fact that the sign of the correlation of total ice concentration with SST and SSS does not change during the January–April period provides some justification for the approach of using static background-error statistics for assimilating data with the 3DVAR system.

The background-error horizontal correlation length scale, required to specify the coefficients used in the diffusion operator as discussed in section 3a, is also estimated from the ensemble of 24-h forecasts from the EnKF. The direct calculation of a local correlation length scale in gridpoint space by fitting a chosen analytical correlation function to the sample correlations is impractical because of the high computational cost involved. Instead, a local estimate based on the definition of length scale $L_c$ given by Daley (1991) is used:

$$L_c = \frac{1}{[\rho''(0)]^{1/2}},$$

where $\rho''$ is the second derivative of the correlation function. By using a centered-difference discretization of the second derivative, an estimate of the length scale for each analysis variable can be efficiently obtained from the EnKF background ensembles as a function of location and time. The approach will likely lead to a noisier estimate than with other approaches, but the noise can be reduced by spatial or temporal averaging of the estimated length scale.

Figure 5 shows the average of the zonal and meridional background-error correlation length scales for total ice concentration (Fig. 5a) and SST (Fig. 5b) calculated using (13) after temporally averaging over one month (March 2003). Note that the length scale for total ice concentration (Fig. 5a) is approximately 20 km over most of the area near the Newfoundland and Labrador coast where ice is present. Some localized areas have
a length scale up to 40 km. In areas where ice is not present, the length scale cannot be computed; therefore, these areas have no shading. The length scale for SST (Fig. 5b) is somewhat shorter, around 15 km, over most of the region shown, except for the southeast corner, where the length scale is much larger. Similar results for the subsurface levels show that the background-error horizontal correlation length scale for ocean temperature and salinity does not vary significantly with depth and the temporally and spatially averaged length scale for temperature, salinity, and ice concentration are all approximately 20 km. For this reason, the simplification is made to use the same length scale, equal to 20 km, for all analyzed variables at all depths.

4. Results

Several data assimilation experiments are conducted to evaluate the impact of using various configurations of the 3DVAR as compared with the nudging and other simpler approaches to produce analyses of the ice–ocean state. The experiments cover the period from 5 December 2006 to 30 June 2007. The partial concentrations of 24 ice thickness categories derived from CIS daily ice charts (Newfoundland and part of the eastern Arctic region covering the northern portion of the model domain) are assimilated every 24 h during the 7-month period. No ocean observations are assimilated. For two of the 3DVAR experiments, the same type of information from RADARSAT image analyses valid between 1200 and 2400 UTC each day is also assimilated. These two experiments demonstrate the ability of the 3DVAR approach to simultaneously assimilate multiple sources of information, which is not possible with nudging and other simpler approaches.

For each experiment, 48-h coupled ice–ocean model forecasts are produced from the analyses produced each day. In addition to using 3DVAR, three additional experiments are performed using simpler approaches. A persistence experiment consists of producing forecasts by persisting the ice thickness distribution from the CIS daily ice chart 48 h forward in time, and therefore this experiment does not make use of the ice–ocean model. A direct insertion experiment is performed by initializing the ice–ocean model forecasts each day by replacing the ice thickness distributions with those from the CIS daily ice chart. A nudging experiment is performed by assimilating the ice information from the daily ice charts following the approach described in section 2d. The different 3DVAR data assimilation experiments are performed to assess the sensitivity of the analysis and subsequent model forecast to the observation-error variance, to the background-error horizontal correlation length scale, and to the inclusion of corrections to the ocean variables in the analysis step. Two additional experiments also show the impact of additionally assimilating the RADARSAT image analyses. The results are presented in terms of the statistics of forecast-minus-observation differences for both total ice concentration and effective ice thickness against RADARSAT image analyses and CIS daily ice charts (Newfoundland region only) in Tables 1 to 4. The term “effective ice thickness” is defined as the ice thickness that would be obtained by evenly distributing the ice over the entire grid cell while maintaining the same ice volume. When comparing against RADARSAT image analyses, the 15- and 39-h forecast lead times were chosen because most of the image analyses are available around 0900 UTC each day within this model domain. The verification statistics only include points within the region of maximum ice extent.
computed from the 1970–2000 weekly ice climatology and also only such points located south of 57°N (corresponding to the northern limit of the domain included in the Newfoundland region CIS daily ice charts). As an example, the dark shaded region in Fig. 1 indicates the regions used for the verification during the week of 12 March.

a. Impact on analysis

To demonstrate the general behavior of a typical 3DVAR analysis step, the input and output fields of the analysis on 2 March 2007 are shown in Figs. 6 and 7 (from the fourth 3DVAR experiment, which is described later). Figures 6a,b show the CIS daily ice chart (i.e., the observations) and the 24-h model forecast (i.e., the background state) of total ice concentration, respectively. Figure 6c shows the difference between the total ice concentration from the observations and the background state. Large differences occur along the ice edge and near the coast of Newfoundland. The assimilation of the partial concentrations of the 24 thickness categories brings the analysis much closer to the total ice concentration from the ice chart, as shown in Fig. 6d. Figure 7a shows the corresponding total ice concentration analysis increment, which is the correction made to the background state by the 3DVAR to produce the analysis. The amplitude of the correction to the total ice concentration field can be large, up to 0.8 in this case. Similarly, the analysis increment of the effective ice thickness is shown in Fig. 7b. As a result of this particular 3DVAR analysis, the largest changes to ice thickness are positive (up to about 1 m) and occur along the Labrador coast. The corresponding SST and SSS analysis increments are shown in Figs. 7c,d, respectively. As expected from the negative correlation between ice concentration and both temperature and salinity discussed earlier, an increase in ice concentration corresponds to a decrease in temperature and salinity. The temperature analysis increment can be as large as 1°C at the ocean surface.

The average analysis increments over the 7 months of the experiment for total ice concentration and SST are depicted in Figs. 8a,b, respectively, again from the fourth 3DVAR experiment, which is described later. There is a small net positive ice concentration analysis increment (up to 0.11) and negative SST analysis increment (as large as −0.12) in the northern part of the region shown, which is along the ice edge off the Labrador coast. This suggests that the 24-h model forecast tends

---

### Table 1. Standard deviation of the total ice concentration difference between model forecasts and CIS data for seven experiments where ice thickness distribution from the CIS daily ice charts are assimilated.

<table>
<thead>
<tr>
<th>Expt name</th>
<th>Obs-error std dev</th>
<th>Ocean update</th>
<th>Lc (km)</th>
<th>Image analysis</th>
<th>Daily ice chart</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>15-h forecast</td>
<td>39-h forecast</td>
<td>Analysis</td>
</tr>
<tr>
<td>Persistence</td>
<td>0</td>
<td>No</td>
<td>0</td>
<td>0.190 0.224</td>
<td>0.000 0.145</td>
</tr>
<tr>
<td>D.I.</td>
<td>0</td>
<td>No</td>
<td>0</td>
<td>0.174 0.195</td>
<td>0.000 0.142</td>
</tr>
<tr>
<td>Nudging</td>
<td>0</td>
<td>Yes**</td>
<td>0</td>
<td>0.173 0.195</td>
<td>0.000 0.141</td>
</tr>
<tr>
<td>3DVAR 1</td>
<td>0.0002</td>
<td>No</td>
<td>0</td>
<td>0.173 0.195</td>
<td>0.000 0.142</td>
</tr>
<tr>
<td>3DVAR 2</td>
<td>0.02</td>
<td>No</td>
<td>0</td>
<td>0.172* 0.194*</td>
<td>0.037 0.144</td>
</tr>
<tr>
<td>3DVAR 3</td>
<td>0.02</td>
<td>No</td>
<td>20</td>
<td>0.173 0.194*</td>
<td>0.050 0.143</td>
</tr>
<tr>
<td>3DVAR 4</td>
<td>0.02</td>
<td>Yes</td>
<td>20</td>
<td>0.176 0.197</td>
<td>0.050 0.139*</td>
</tr>
</tbody>
</table>

* The lowest standard deviation value for the verification score.
** Ocean state modified indirectly through coupled ice–ocean model physics during the 24-h nudging period.

---

### Table 2. Standard deviation of the effective ice thickness difference (in cm) between model forecasts and CIS data for seven experiments where ice thickness distribution from the CIS daily ice charts are assimilated.

<table>
<thead>
<tr>
<th>Expt name</th>
<th>Obs-error std dev</th>
<th>Ocean update</th>
<th>Lc (km)</th>
<th>Image analysis</th>
<th>Daily ice chart</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>15-h forecast</td>
<td>39-h forecast</td>
<td>Analysis</td>
</tr>
<tr>
<td>Persistence</td>
<td>0</td>
<td>No</td>
<td>0</td>
<td>23.5 27.2</td>
<td>0.3 16.4</td>
</tr>
<tr>
<td>D.I.</td>
<td>0</td>
<td>No</td>
<td>0</td>
<td>22.0 25.1*</td>
<td>0.3 15.8*</td>
</tr>
<tr>
<td>Nudging</td>
<td>0</td>
<td>Yes**</td>
<td>0</td>
<td>22.0 25.3</td>
<td>0.3 15.8*</td>
</tr>
<tr>
<td>3DVAR 1</td>
<td>0.0002</td>
<td>No</td>
<td>0</td>
<td>21.9* 25.1*</td>
<td>0.3 15.8*</td>
</tr>
<tr>
<td>3DVAR 2</td>
<td>0.02</td>
<td>No</td>
<td>0</td>
<td>22.0 25.4</td>
<td>5.8 16.2</td>
</tr>
<tr>
<td>3DVAR 3</td>
<td>0.02</td>
<td>No</td>
<td>20</td>
<td>22.0 25.2</td>
<td>7.0 16.1</td>
</tr>
<tr>
<td>3DVAR 4</td>
<td>0.02</td>
<td>Yes</td>
<td>20</td>
<td>22.2 25.3</td>
<td>7.0 16.2</td>
</tr>
</tbody>
</table>

* The lowest standard deviation value for the verification score.
** Ocean state modified indirectly through coupled ice–ocean model physics during the 24-h nudging period.
to underestimate the ice concentration in this area, on average. Farther south, close to Newfoundland, the opposite is true but with a smaller amplitude.

b. Impact on short-term forecasts: Total ice concentration

Results from evaluating the forecasts of total ice concentration in terms of the standard deviation of differences with CIS RADARSAT image analyses and CIS daily ice charts are shown in Table 1. The second column shows the observation-error standard deviation used to assimilate each of the partial concentrations of the 24 thickness categories from the CIS daily ice charts. As a point of reference, the value of 0.02 used for several 3DVAR experiments corresponds with a standard deviation for total ice concentration of 0.098, which is smaller than the value used for the background-error standard deviation of total ice concentration, 0.186. The third column indicates whether the ocean variables are modified during the 3DVAR analysis through the background-error covariances between ice concentration and the ocean temperature and salinity shown in section 3b. The remaining columns give the verification scores averaged over the entire 7-month period computed over all points where the weekly sea ice climatology indicates a nonzero probability of ice.

Comparing the standard deviations in Table 1 shows an improvement in the forecasts produced from all of the data assimilation experiments relative to the persistence experiment and that this improvement is larger when verifying against image analyses than the daily ice charts. This is an indication that the daily ice charts have a tendency to resemble a persistence forecast in certain areas and during certain periods, as already discussed. It also shows that the coupled ice–ocean model can produce forecasts of total ice concentration with a lead time of up to 48 h that are significantly better than persistence, when verified against the image analyses. The differences among all of the data assimilation experiments are much smaller than the difference between any of these experiments and the persistence forecasts when measured against image analyses. Specifically, the differences in standard deviation for the 15- and 39-h forecasts, as verified against image analyses, among all of the data assimilation experiments shown in Table 1 are generally not statistically significant (with 95% confidence level), except for the fourth 3DVAR experiment.

The CIS ice charts contain values of ice concentration that are only stored to the nearest tenth. Therefore, in addition to the standard statistical significance test, the impact on the precision of the verification statistics caused by the limited precision of the verifying observations was examined. This was done by randomly perturbing the verifying observations according to the uniform distribution \([-0.05, 0.05]\), which is below the resolution of the observations. As a result, the standard deviations shown in Table 1 with respect to the image analyses only changed by, at most, 0.0024. This is smaller than the differences between the persistence and any of the data assimilation experiments and also smaller than the difference between the fourth 3DVAR experiment and the other data assimilation experiments.

### Table 3. Standard deviation of the total ice concentration difference between model forecasts and CIS data for four 3DVAR experiments that demonstrate the impact of also assimilating RADARSAT image analyses.

<table>
<thead>
<tr>
<th>Expt name</th>
<th>Obs-error std dev</th>
<th>Ocean update</th>
<th>$L_c$ (km)</th>
<th>Image analysis</th>
<th>Daily ice chart</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15-h forecast</td>
<td>39-h forecast</td>
</tr>
<tr>
<td>3DVAR 3</td>
<td>0.02</td>
<td>No</td>
<td>20</td>
<td>0.173</td>
<td>0.194</td>
</tr>
<tr>
<td>3DVAR 3+</td>
<td>0.02</td>
<td>No</td>
<td>20</td>
<td>0.170*</td>
<td>0.193*</td>
</tr>
<tr>
<td>3DVAR 4</td>
<td>0.02</td>
<td>Yes</td>
<td>20</td>
<td>0.176</td>
<td>0.197</td>
</tr>
<tr>
<td>3DVAR 4+</td>
<td>0.02</td>
<td>Yes</td>
<td>20</td>
<td>0.172</td>
<td>0.196</td>
</tr>
</tbody>
</table>

* The lowest standard deviation value for the verification score.

### Table 4. Standard deviation of the effective ice thickness difference (cm) between model forecasts and CIS data for four 3DVAR experiments that demonstrate the impact of also assimilating RADARSAT image analyses.

<table>
<thead>
<tr>
<th>Expt name</th>
<th>Obs-error std dev</th>
<th>Ocean update</th>
<th>$L_c$ (km)</th>
<th>Image analysis</th>
<th>Daily ice chart</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15-h forecast</td>
<td>39-h forecast</td>
</tr>
<tr>
<td>3DVAR 3</td>
<td>0.02</td>
<td>No</td>
<td>20</td>
<td>22.0</td>
<td>25.2*</td>
</tr>
<tr>
<td>3DVAR 3+</td>
<td>0.02</td>
<td>No</td>
<td>20</td>
<td>21.9*</td>
<td>25.2*</td>
</tr>
<tr>
<td>3DVAR 4</td>
<td>0.02</td>
<td>Yes</td>
<td>20</td>
<td>22.2</td>
<td>25.3</td>
</tr>
<tr>
<td>3DVAR 4+</td>
<td>0.02</td>
<td>Yes</td>
<td>20</td>
<td>22.0</td>
<td>25.3</td>
</tr>
</tbody>
</table>

* The lowest standard deviation value for the verification score.
The first 3DVAR experiment is designed to produce analyses that are similar to those produced by direct insertion. This is attained by using a small observation-error standard deviation (100 times smaller than for the other 3DVAR experiments), no background-error correlations between the partial ice concentrations for the different ice thickness categories, no horizontal background-error correlations, and no update to the ocean variables. As a consequence, the analyses produced by this experiment fit much more closely to the assimilated data than the other 3DVAR experiments as seen in the standard deviation of the differences between the analysis and the daily ice chart.

The configuration of the second 3DVAR experiment is similar to the first, except that a larger, more realistic observation-error standard deviation is used and the background-error correlations between the partial ice concentrations are included. Even though this results in analyses that do not fit the assimilated observations as closely, the standard deviation of the differences between the forecasts and image analyses are slightly reduced. Compared with this experiment, the third 3DVAR experiment includes horizontal background-error correlations with a length scale of 20 km. This added constraint again decreases the fit to the assimilated observations while producing a very small difference with respect to both image analyses and the daily ice charts.

Finally, the fourth 3DVAR experiment includes the update to the ocean variables during the analysis step resulting from the background-error covariances between the ice and ocean variables. It is interesting to note that, when compared with the third experiment, for this experiment the standard deviations for total ice concentration are slightly higher with respect to the image analyses but lower when measured against the daily ice charts. Because the daily ice charts are assimilated and they have the tendency to resemble a persistence forecast in some regions at certain times, the
improved fit of the forecasts to the daily ice charts may indicate that the ocean update in the 3DVAR experiment helps preserve the assimilated information in the model forecast for a longer period.

The total ice concentration bias with respect to both image analyses and daily ice charts were also calculated but are not shown. For all experiments, the amplitude of the bias is at least an order of magnitude smaller than the comparable values for standard deviation. The bias for all experiments is positive (corresponding to too much ice in the forecasts) for the verification against RADARSAT image analyses, whereas the opposite is true for the verification against daily ice charts. Furthermore, as the forecast progresses, the bias becomes less positive (or more negative) for all experiments, except for persistence. This consistent temporal trend during the forecast suggests that, on average throughout the 7-month period, the model may be melting the ice too rapidly and/or not forming enough new ice. Alternatively, it may be due to an underestimation of ice concentration north of 57°N that is then advected into the verifying domain during the forecast. This is consistent with the relatively large mean positive analysis increment for ice concentration seen between 55° and 57°N in Fig. 8a.

c. Impact on short-term forecasts: Effective ice thickness

Evaluation of the forecasts of ice thickness is more difficult because direct measurements are not available. The ice thickness information used to evaluate the forecasts is inferred from the CIS RADARSAT image analyses and daily ice charts as described in section 2a. Table 2 shows the standard deviation of the differences with respect to image analyses and daily ice charts for all experiments in a similar format as used for the total ice concentration results shown previously. All approaches produce similar results, except for persistence, which results in a slightly higher standard deviation with respect to
the image analyses. Similar to the results for total ice concentration, the inclusion of the update to the ocean variables does not improve the forecasts. This may be due to the simplifying assumptions adopted in the parameterization of the background-error covariances: namely, assuming the covariances between the ice and ocean variables are spatially and temporally homogeneous.

The reason why the standard deviation is not exactly zero at 0 h for the persistence, direct insertion, and nudging experiments is due to a slight inconsistency in the treatment of points within a small region where the CIS daily ice charts for the Newfoundland and eastern Arctic regions overlap. Data from the two ice charts within this overlap region are averaged for the purpose of assimilation, but only the data from the Newfoundland ice charts are used for verification.

The effective ice thickness bias was also computed for all experiments. They are qualitatively very similar to the total ice concentration bias. Specifically, the bias is very small when compared with the standard deviations, and a negative temporal trend is seen as a function of forecast lead time. Thus, it is possible that the same physical process that is responsible for the decrease in total ice concentration as a function of forecast lead time is also responsible for the decrease in the effective ice thickness.

d. Results from assimilating both image analyses and daily ice charts

Two additional 3DVAR experiments were performed that assimilate all available RADARSAT image analyses valid between 1200 and 2400 UTC each day in addition to the same daily ice charts assimilated in the experiments already presented. The two additional experiments are based on the third and fourth 3DVAR experiments presented above. Table 3 shows the standard deviation of the differences in total ice concentration with respect to image analyses and daily ice charts for the third and fourth 3DVAR experiments (as also shown in Table 1) and from the new experiments (3+ and 4+). As can be expected, the fit of the analyses to the assimilated daily ice charts is decreased somewhat by assimilating the image analyses because small differences between the two types of assimilated data will draw the analyses away from the daily ice charts. However, a statistically significant improvement is seen in the quality of the ice concentration forecasts relative to the image analyses (for the 15-h forecasts) and daily ice charts (for the 24- and 48-h forecasts) from also assimilating the image analyses. Though there is also an improvement in the 39-h forecasts with respect to the image analyses for both experiments 3+ and 4+, it is not statistically significant. The results from 3DVAR experiment 3+ have the lowest standard deviations of all experiments when verifying against image analyses. Similarly, the results from 3DVAR experiment 4+ are the lowest when verifying against the daily ice charts.

The impact on the effective ice thickness from assimilating the RADARSAT image analyses, in addition to the daily ice charts, is shown in Table 4. A similar positive impact is seen as for ice concentration. That is, the improvements are statistically significant for the forecasts relative to the image analyses (for the 15-h forecasts) and daily ice charts (for the 24- and 48-h forecasts).

5. Conclusions

A 3DVAR system has been constructed to provide analyses of the ice–ocean state and to initialize forecasts produced with a coupled ice–ocean model. For the present study, configurations of the data assimilation
system and ice–ocean model for a domain over the Canadian east coast region were used. However, the software has been developed in such a way as to allow the data assimilation system to be relatively easily implemented over different geographical domains and with different numerical models (or even using persistence when a numerical model is not available). For this preliminary study, most of the experiments were performed by assimilating only a single type of data, the CIS daily ice charts. This was done to enable comparison between the results from the 3DVAR system, direct insertion, and nudging. However, for two additional experiments, the RADARSAT image analyses were also assimilated. Even though the 3DVAR analysis also modifies the ocean variables, the evaluation of the system focused exclusively on the quality of the ice analyses and subsequent forecasts of the total ice concentration and effective thickness.

The emphasis of the present study was on the parameterization and estimation of the background-error covariances. It was found that a simple implementation of the EnKF can provide useful information on background-error statistics, including multivariate relationships, for ice and ocean variables. The uncertainty in the atmospheric forcing explains a large fraction of the resulting background-error variance for ice concentration. However, the roles of imperfect ocean lateral boundary conditions and other sources of uncertainty in the coupled ice–ocean model, was not considered and should be evaluated in the future. A benefit of using the EnKF is that the direct examination of the raw ensemble covariances allowed the evaluation of the assumptions and simplifications in the parameterized 3DVAR background-error covariances. When only assimilating the CIS daily ice charts, all 3DVAR experiments resulted in forecasts of similar quality as compared with direct insertion and nudging, except for the experiment that included the update to the ocean variables. The latter produced forecasts that were significantly improved when verified against daily ice charts but degraded when verified against the RADARSAT image analyses.

The idealized context used for most of the experiments in this study, in which only the ice information from the gridded CIS daily ice charts was assimilated, likely lessens the benefits of advanced data assimilation approaches, such as 3DVAR, when compared with nudging or direct insertion. The results generally show very small changes in the quality of the forecasts when using direct insertion, nudging, or any of the 3DVAR configurations. A larger difference is seen between persistence and any of the experiments using model forecasts initialized with data assimilation, with the data assimilation experiments consistently producing better forecasts. The benefits of using 3DVAR versus direct insertion or nudging are more evident in the context of simultaneously assimilating multiple types of observations that differ from each other in terms of spatial resolution, accuracy, horizontal extent, and their relationship with the analysis variables. Moreover, the direct insertion and nudging approaches require the “observations” to be already combined into a single estimate and mapped onto the model grid. Therefore, these approaches are not suited to the assimilation of multiple sources of data or data with incomplete spatial coverage over the model domain. This was tested by simultaneously assimilating both CIS daily ice charts and RADARSAT image analyses in two additional 3DVAR experiments. In both of these experiments, significant improvements were obtained relative to the experiments that assimilated only the daily ice charts. These experiments also resulted in significantly better forecasts of ice concentration than for any of the other experiments.

The focus of our current research is to extend the system examined in the present study by developing and evaluating approaches to assimilate a diverse set of complementary satellite-based measurements with 3DVAR in a completely automated way. As in the application of such techniques to NWP, the 3DVAR approach to data assimilation is suited to take full advantage of diverse types of data by combining them in a physically and statistically sound way. Because variational data assimilation approaches are commonly used for initializing atmospheric models, adopting a similar approach for sea ice data also facilitates the eventual development of data assimilation strategies for fully coupled ice–ocean–atmosphere models.

Acknowledgments. The authors thank Dinh Hai Tran and Charles Tang for help in resolving issues related to the numerical model. The authors also acknowledge Anthony T. Weaver for help with implementing the diffusion operator. Comments by Mohammed Shokr helped improve an earlier version of the manuscript. This work was supported by the Canadian Program of Energy Research and Development (PERD).

APPENDIX

Calculation of Normalization Factors for the Horizontal Correlations Implemented with a Diffusion Equation

Defining $v_j = \delta_{ij}$ as the $j$th element of the vector $v$, it can be shown from (8) that the diagonal elements of $\Lambda$ are

$$\Lambda_i = (v_i^T v_i)^{-1/2}, \quad (A1)$$
where

\[ \hat{v}_i = \mathbf{W}^{-1/2} \mathbf{L}^{1/2} v_i, \]  

(A2)

The number of normalization factors corresponds to the number of grid points in the horizontal and the calculations using (A1) and (A2) can become too expensive. Alternatively, the normalization factors can be estimated by randomization. Let \( \mathbf{v} \) denote a vector of independent and normally distributed random numbers with the property that \( \mathbb{E}[\mathbf{v}] = 0 \) and \( \mathbb{E}[\mathbf{v} \mathbf{v}^T] = \mathbf{I} \). It can be shown that

\[ \Lambda_i \approx \left( \text{diag} \left\{ \frac{1}{R-1} \sum_{r=1}^{R} \hat{v}_r \hat{v}_r^T \right\} \right)^{-1/2}, \]  

(A3)

where \( R \) is the number of random vectors and

\[ \hat{v}_r = \mathbf{L}^{1/2} \mathbf{W}^{-1/2} v_r, \]  

(A4)

Although the normalization factors could be computed using randomization with a number of samples much smaller than the number of horizontal grid points found in typical applications, the exact calculation of the normalization factors is used in the present study.

REFERENCES


