This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

An Explicit Data Assimilation Scheme for a Nonlinear Wave Prediction Model Based on a Pseudo-Spectral Method

Seongjin Yoon, Jinwhan Kim, *Member, IEEE*, and Wooyoung Choi

***Abstract*—A robust data assimilation scheme is presented for a wave model to predict evolving nonlinear ocean waves. The Fourier coefficients of the surface elevation and the free surface velocity potential are chosen for state variables and are propagated in time by solving numerically a set of nonlinear evolution equations using a pseudo-spectral method. The numerical solutions are then updated with noise corrupted measurements of the surface elevation with the aid of an explicit Kalman filter for which the time evolution of the error covariance matrix is found explicitly by solving analytically the linearized wave prediction model. After presenting an error analysis for this explicit data assimilation scheme, numerical simulations of the integrated nonlinear wave prediction model for long-crested waves of varying wave steepness are performed by using synthetic data with different noise characteristics. It is shown that the estimated surface wave fields agree well with the true states, and the present data assimilation scheme based on the explicit Kalman filter improves considerably the computational efficiency and stability, in comparison with a standard Kalman filter for which the error covariance matrix is found numerically.**

***Index Terms*—Data assimilation, Kalman filter, pseudo-spectral wave model, wave prediction.**

I. Introduction

ACCURATE prediction of evolving ocean waves is important for the design and safe operation of ships and offshore structures. Recently, there has been a growing concern regarding the occurrence of so-called "rogue waves," whose wave heights are typically greater than twice the significant wave height of a given wave spectrum [1], [2]. As the number of offshore structures in deep oceans increases, more structures are susceptible to damage by these extreme waves. Thus, the real-time prediction of ocean wave fields resolving not only wave amplitude, but also wave phase is required for offshore structures to avoid being caught in extreme wave conditions.

Unfortunately, these extreme waves appear spontaneously, without any warning, in spatially isolated locations. Therefore, their accurate prediction is nontrivial and observed field data often need to be assimilated into wave models to improve their prediction. Up until now, ocean wave data assimilation has been adopted mostly by spectral wave models for statistical descriptions of wave spectral changes [3]–[5]. However, the phase-averaged spectral wave models have limitation in that they fail to predict the exact location and time of extreme wave events; furthermore, it is troublesome to describe correctly the evolution of nonlinear waves and their interactions.

An alternative is to choose a deterministic approach using, for example, a phase-resolving nonlinear wave model, but it has been considered almost impractical for many decades, due to its high computational cost and the lack of sensors measuring accurately wave fields to initialize such a model. However, with the development of high performance computers and more reliable measurement sensors such as shipborne radars, it seems to be feasible to predict the occurrence of extreme waves once an efficient and robust numerical model is available.

Various numerical models have been developed to describe the evolution of nonlinear surface waves. However, even under the idealized potential flow assumption, most phase-resolving wave models have been found computationally too expensive to be adopted in practice to simulate three-dimensional ocean waves in a large computational domain. In recent years, a formulation based on asymptotic expansion in wave steepness proposed originally by West *et al.* [6] attracts much attention since it is found numerically efficient for broadband short-crested waves when it is combined with a pseudo-spectral method based on the fast Fourier transform [7]–[9]. Although the original formulation of West *et al.* [6] is based on the small wave steepness assumption, comparisons with laboratory experiments [10], [11] have shown that the pseudo-spectral wave model describes accurately, up to wave breaking, the nonlinear evolution of focusing and modulationally unstable wave groups as well as broad-band irregular wave fields. Furthermore, to simulate postbreaking waves, an eddy viscosity model describing energy dissipation due to wave breaking along with a wind forcing model was incorporated into the pseudo-spectral wave model [12], [13], which has been validated successfully with recent laboratory measurements [14]–[16]. Thus, the

Manuscript received August 22, 2014; revised November 28, 2014; accepted February 16, 2015. . This work was supported by the National Research Foundation of Korea funded by the Ministry of Education, Science, and Technology through the World Class University (WCU) program under Grant R31-2008-000-10045-0.

Associate Editor: K. Takagi.

S. Yoon is with the Division of Ocean Systems Engineering, KAIST, Daejeon 305-701, Republic of Korea, and also with the Department of Naval Architecture and Marine Engineering, University of Michigan, Ann Arbor, MI 48109-2145 USA (e-mail: seongjin@umich.edu).

J. Kim is with the Division of Ocean Systems Engineering, Korea Advanced Institute of Science and Technology, Daejeon 305-701, South Korea (e-mail: jinwhan@kaist.ac.kr).

W. Choi is with the Department of Mathematical Sciences, Center for Applied Mathematics and Statistics, New Jersey Institute of Technology, Newark, NJ 07102-1982 USA (e-mail: wychoi@njit.edu).

Digital Object Identifier 10.1109/JOE.2015.2406471

0364-9059 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

pseudo-spectral wave model seems to be a good candidate in predicting nonlinear ocean wave fields, with resolving a wide range of wavelength scales.

While the pseudo-spectral wave model is found effective for numerical simulations, its prediction capability relies heavily on accurate wave field measurements to initialize the model. Although improved greatly, measurement sensors could provide wave field data that contain both measurement errors and noise, leading to inaccurate wave predictions. Therefore, a technique to minimize the effects of such errors and noise is required for better performance of the wave prediction model.

Here we aim to develop a real-time data assimilation system by incorporating into the pseudo-spectral wave model a Kalman filter developed to produce optimal estimated wave fields over time from noise corrupted measurements with assuming that the system is linear and the measurement noise is additive, independent, and identically distributed Gaussian random process.

Although Kalman filtering is used widely for various applications, it is found that integrating the wave prediction model in time with a naive implementation of a Kalman filtering algorithm is problematic mainly for two reasons. One is its extra computational cost to find numerically the error covariance matrices and the other is numerical instability associated with time periodic applications of the data assimilation scheme since observed wave field data are available, for example, only after one full rotation of a shipborne radar antenna, as described later. An attempt to assimilate field measurements has been previously made for the pseudo-spectral wave model using a conjugate gradient method [17], but its computational cost is also high due to numerical evaluations of the gradient of the cost function. It is therefore useful to develop a data assimilation scheme resolving these two issues.

The paper is organized as follows. The pseudo-spectral wave model truncated at the third order in wave steepness is adopted in Section II for the prediction of evolving nonlinear surface wave fields while the linearized wave model written in Fourier space is used in Section III to develop a Kalman filtering system for data assimilation. As shown in Section IV, this enables one to evaluate the evolution of the error covariance matrix analytically, which not only reduces the computational cost, but also improves the stability of the system. To demonstrate the performance of the integrated wave prediction model, several numerical experiments are performed for various measurement noise characteristics and wave fields of different wave steepnesses. Detailed simulation conditions and results are discussed in Section V.

II. Nonlinear Wave Model

For inviscid, incompressible, and irrotational flows, the velocity potential $\phi(x, y, z, t)$ satisfies the Laplace equatio

$$\nabla^2\phi + \frac{\partial^2\phi}{\partial z^2} = 0 \quad \text{for} \quad -\infty < z \le \zeta \tag{2.1}$$

where ζ is the free surface elevation and ∇ is the two-dimensional horizontal gradient defined as

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right). \tag{2.2}$$

The free surface boundary conditions at $z = \zeta(x, y, t)$ are given by

$$\frac{\partial\zeta}{\partial t} + \nabla\phi\cdot\nabla\zeta = \frac{\partial\phi}{\partial z} \qquad \text{at} \quad z = \zeta(x, y, t), \tag{2.3}$$

$$\frac{\partial\phi}{\partial t} + \frac{1}{2}|\nabla\phi|^2 + \frac{1}{2}\left(\frac{\partial\phi}{\partial z}\right)^2 + \mathrm{g}\zeta = 0 \quad \text{at} \quad z = \zeta(x, y, t). \tag{2.4}$$

By introducing Φ, the velocity potential at the free surface, defined by

$$\Phi(x, y, t) = \phi(x, y, z = \zeta, t) \tag{2.5}$$

equations (2.3) and (2.4) can be written [18] as

$$\frac{\partial\zeta}{\partial t} + \nabla\Phi\cdot\nabla\zeta = \left[1 + (\nabla\zeta)^2\right] W, \tag{2.6}$$

$$\frac{\partial\Phi}{\partial t} + \frac{1}{2}[\nabla\Phi]^2 + \mathrm{g}\zeta = \frac{1}{2}\left[1 + (\nabla\zeta)^2\right] W^2 \tag{2.7}$$

where W is the vertical velocity evaluated at the free surface defined by

$$W = \left.\frac{\partial\phi}{\partial z}\right|_{z=\zeta}. \tag{2.8}$$

Under the small wave steepness assumption, when W is expanded as

$$W = \sum_{j=1}^{\infty} W_j[\zeta, \Phi] \tag{2.9}$$

we can obtain the nonlinear evolution equations for ζ and Φ, by substituting (2.9) into (2.6) and (2.7), as

$$\frac{\partial\zeta}{\partial t} = \sum_{j=1}^{\infty} Q_j[\zeta, \Phi], \qquad \frac{\partial\Phi}{\partial t} = \sum_{j=1}^{\infty} R_j[\zeta, \Phi] \tag{2.10}$$

where W_j, Q_j, and R_j have been assumed to be $O(\epsilon^j)$, with ϵ being the wave steepness, and can be found recursively, as shown, for example, in Choi *et al.* [9]. For numerical computations, depending on the desired accuracy and computational efficiency, the series on the right-hand sides should be truncated to a prescribed order of nonlinearity. This system has been further generalized to include the wave breaking effect [12], as mentioned previously, but, as our focus is the development of a data assimilation scheme for the nonlinear evolution model given by (2.10), we consider only nonbreaking waves in this paper.

To develop a data assimilation scheme, we use the nonlinear evolution equations correct to the third order in wave steepness, which is the minimum order of nonlinearity to describe the resonant interaction among four gravity waves. In terms of ζ and Φ, the third-order system can be written [19] as

$$\frac{\partial\zeta}{\partial t} + \mathcal{L}[\Phi] + \nabla\cdot(\zeta\nabla\Phi) + \mathcal{L}\left[\zeta\mathcal{L}[\Phi]\right] + \nabla^2\left(\frac{1}{2}\zeta^2\mathcal{L}[\Phi]\right) + \mathcal{L}\left[\zeta\mathcal{L}\left[\zeta\mathcal{L}[\Phi]\right] + \frac{1}{2}\zeta^2\nabla^2\Phi\right] = 0, \tag{2.11}$$

$$\frac{\partial\Phi}{\partial t} + \mathrm{g}\zeta + \frac{1}{2}\nabla\Phi\cdot\nabla\Phi - \frac{1}{2}\left(\mathcal{L}[\Phi]\right)^2$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

$$-\mathcal{L}[\Phi]\Big(\zeta\nabla^2\Phi+\mathcal{L}\big[\zeta\mathcal{L}[\Phi]\big]\Big)=0. \tag{2.12}$$

In (2.11)–(2.12), the linear integral operator $\mathcal{L}$ is defined symbolically as $\mathcal{L}[\Phi]=-\mathcal{F}^{-1}[|k|\mathcal{F}[\Phi]]$ for the case of infinite depth water, where $\mathcal{F}$ and $\mathcal{F}^{-1}$ represent the Fourier and inverse Fourier transforms, respectively.

In the numerical wave prediction model, the free surface elevation $\zeta(x,t)$ and the free surface velocity potential $\Phi(x,t)$ are expressed in truncated Fourier series

$$\zeta(x,t)=\sum_{i=-N+1}^{N} Z_i(t)\,\mathrm{e}^{ik_ix}, \tag{2.13}$$

$$\Phi(x,t)=\sum_{i=-N+1}^{N} \Pi_i(t)\,\mathrm{e}^{ik_ix} \tag{2.14}$$

where N is the number of Fourier modes for positive wave numbers, and Z_i and Π_i are the complex Fourier coefficients of ζ and Φ, respectively. Since the surface elevation and the free surface velocity potential are real, $Z_{-i}=Z_i^*$ and $\Pi_{-i}=\Pi_i^*$, where the asterisks denote the complex conjugates. Then the system of nonlinear evolution equations given by (2.11)–(2.12) is integrated in time using a fourth-order Runge–Kutta scheme with a low-pass filter to eliminate aliasing errors introduced by the use of truncated Fourier series; see, for example, Choi *et al.* [9] for a detailed description.

Although the wave prediction model can be solved accurately with the numerical method described, the reliable deterministic prediction of wave fields depends on initial conditions for ζ and Φ, which are difficult to estimate correctly from field measurements by, for example, a shipborne marine radar. Even if available, such measurements are often noisy. Therefore, to improve its performance, the wave model should be assimilated with field measurements; otherwise, the prediction might be inaccurate. Here, using an explicit Kalman filter, we develop an efficient data assimilation scheme for the wave prediction model given by (2.10) and test its predictive capability.

III. Problem Formulation

A. State Space Representation

As the wave prediction model uses a pseudo-spectral method, the surface elevation and the free surface velocity potential are computed both in the spatial domain (or at given grid points) for nonlinear operations and in the spectral domain, in terms of their Fourier coefficients, for linear operations. Therefore, for Kalman filtering, the state-space representation can be made either in the spatial domain or in the spectral domain. Nonetheless, as discussed subsequently, it is found more convenient to use the Fourier coefficients of ζ and Φ in the spectral domain as state variables, from which the state vector $\mathbf{X}$ can be constructed as

$$\mathbf{X}=\begin{bmatrix} Z_1 \\ \vdots \\ Z_N \\ \hdashline \Pi_1 \\ \vdots \\ \Pi_N \end{bmatrix}. \tag{3.1}$$

Since sensors for wave field measurement such as a noncoherent marine radar, often detect only the free surface elevation, the measurement output vector $\mathbf{Y}$ is assumed to be given by

$$\begin{aligned}\mathbf{Y}&=\left[\begin{array}{c:c}\mathbf{I}_{(N\times N)} & \mathbf{O}_{(N\times N)}\end{array}\right]\mathbf{X}\\ &=\mathbf{H}\,\mathbf{X}\end{aligned} \tag{3.2}$$

where $\mathbf{H}$ is the $N\times 2N$ measurement sensitivity matrix that consists of the identity matrix $\mathbf{I}_{(N\times N)}$ for the free surface elevation and the zero matrix $\mathbf{O}_{(N\times N)}$ for the free surface velocity potential.

B. Linear Approximation of the Evolution Equations

Although the wave prediction model solves a system of nonlinear evolution equations given by (2.11)–(2.12), we develop a data assimilation scheme based on the linearized wave model, which improves considerably the efficiency and stability of the filter. More specifically, the resulting linear system is solved analytically, as shown in this section, and the error covariance matrix required for Kalman filtering is then evaluated analytically. Otherwise, its evaluation is computationally expensive, in particular, when the number of state variables is large in a large computational domain. Therefore, the use of the linear system improves significantly the computational efficiency. In addition, as discussed in Sections III-C and V-B, the analytic evaluation removes the numerical instability associated with a low-order time-integration scheme, which is often adopted for fast evaluation, but is conditionally unstable.

Although the linear system is adopted to improve numerical efficiency, its use for data assimilation can be justified from the fact that the nonlinear wave interaction for gravity waves is a slow process, whose characteristic time scale inversely proportional to the square of wave steepness for resonant four-wave interactions is much longer than a typical update period of Kalman filtering, which is the rotation period of a radar. Thus, this nonlinear aspect of gravity waves can be assumed to be insignificant during a relatively short filtering period. On the other hand, the effect of nonlinearity cannot be neglected in the long-term evolution of wave fields.

From (2.11)–(2.12), the linear evolution equations for the Fourier coefficients of ζ and Φ can be written as two coupled differential equations

$$\frac{\mathrm{d}Z_i}{\mathrm{dt}}=|k_i|\Pi_i, \tag{3.3}$$

$$\frac{\mathrm{d}\Pi_i}{\mathrm{dt}}=-\,\mathrm{g}Z_i \tag{3.4}$$

where k_i is the wavenumber of the ith mode. In the state vector representation, equations (3.3)–(3.4) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{X}=\left[\begin{array}{c:c}\mathbf{O}_{(N\times N)} & |\mathbf{k}|_{(N\times N)}\\ \hdashline -\mathrm{g}\mathbf{I}_{(N\times N)} & \mathbf{O}_{(N\times N)}\end{array}\right]\mathbf{X}\equiv\mathbf{F}\mathbf{X} \tag{3.5}$$

where $|\mathbf{k}|$ is an $N\times N$ diagonal matrix whose ith diagonal term is $|k_i|$. Notice that $\mathbf{F}$ is a $2N\times 2N$ block diagonal matrix. Then, the evolution equation for each Fourier component is completely decoupled from other Fourier components when

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

the system is written as a second-order differential equation for one variable. This is an advantage of the spectral state representation. If we use the spatial representation where the state variables are defined at grid points, the evolution equations for the state variables are coupled through the (nonlocal) integral operator $\mathcal{L}$.

C. Discrete-Time Kalman Filter

A detailed description of the discrete-time Kalman filter (hereafter referred to as the *standard* Kalman filter) can be found in, for example, Kalman [20] and Grewal [21]; therefore, only the final results are given here.

The discrete-time Kalman filter solves the state vector $\mathbf{X}$ and the error covariance matrix $\mathbf{P}$ in two distinct steps: *prediction* and *update*. The former step is to solve the nonlinear evolution equations for ζ and Φ while the data assimilation scheme is applied at every *update* step.

At every prediction time step, $\widehat{\mathbf{X}}$ and $\widehat{\mathbf{P}}$ are determined by

$$\widehat{\mathbf{X}}^n = \mathcal{N}[\widehat{\mathbf{X}}^{n-1}] \qquad \text{for every time step,} \tag{3.6}$$

$$\widehat{\mathbf{P}}^n = \mathbf{\Theta}^{n-1}\,\widehat{\mathbf{P}}^{n-1}\left(\mathbf{\Theta}^{\mathbf{T}}\right)^{n-1} + \mathbf{Q} \qquad \text{for every time step} \tag{3.7}$$

where $\mathcal{N}$ is the third-order nonlinear transition function (or the operator correct to the third order in wave steepness) of the wave model shown in (2.11)–(2.12); $\widehat{\mathbf{P}}$ is a $2N \times 2N$ matrix; the superscript n denotes the current time step such that $t_n = n\Delta t$ with Δt being a time step; and $\mathbf{Q}$ represents the process noise matrix that is assumed to be zero. In (3.7), the linearized state transition matrix $\mathbf{\Theta}$ is often computed numerically using, for example, the forward Euler time-integration scheme, which yields $\mathbf{\Theta} = \mathbf{I} + \mathbf{F}\Delta t$, where the linear operator $\mathbf{F}$ is defined in (3.5).

At every update period of T_u, or at $t = mT_u$ with m being a positive integer, $\widehat{\mathbf{X}}$ and $\widehat{\mathbf{P}}$ are updated to minimize the mean-square error between the measured value vector $\mathbf{Y}$ and the numerical solutions of the nonlinear wave prediction model $\widehat{\mathbf{X}}$, using the Kalman gain that is optimum in the linear sense. In other words, in the update step, $\widehat{\mathbf{X}}$ and $\widehat{\mathbf{P}}$ are improved to $\overline{\mathbf{X}}$ and $\overline{\mathbf{P}}$ as

$$\overline{\mathbf{X}}^n = \widehat{\mathbf{X}}^n + \mathbf{K}^n[\mathbf{Y} - \mathbf{H}\widehat{\mathbf{X}}^n], \tag{3.8}$$

$$\overline{\mathbf{P}}^n = [\mathbf{I} - \mathbf{K}^n\,\mathbf{H}]\,\widehat{\mathbf{P}}^n \tag{3.9}$$

where $\mathbf{H}$ is the $N \times 2N$ measurement sensitivity matrix given by (3.2) and the $2N \times N$ optimal Kalman gain matrix $\mathbf{K}^n$ is computed as

$$\mathbf{K}^n = \widehat{\mathbf{P}}^n\,\mathbf{H}^{\mathbf{T}}\left[\mathbf{H}\,\widehat{\mathbf{P}}^n\,\mathbf{H}^{\mathbf{T}} + \mathbf{R}\right]^{-\mathbf{1}}. \tag{3.10}$$

In (3.10), $\mathbf{R}$ is the $N \times N$ measurement noise matrix which, in this paper, is assumed to be an identity matrix.

Although the procedure described here is well-known, computing (3.7) for the error covariance matrix $\widehat{\mathbf{P}}^n$ in the prediction step could be problematic since its computational cost is high and a simple integration scheme such as the first-order Euler scheme is only conditionally stable with a relatively small region of numerical stability. Therefore, in this paper, we attempt to find a more efficient and more stable data assimilation scheme by finding analytically the error covariance matrix $\widehat{\mathbf{P}}^n$, as described in the following section.

IV. Explicit Data Assimilation

A. Analytic Solution for A Continuous-Time Kalman Filter

When the linear model is written in Fourier space as in (3.3)–(3.4), all Fourier modes are independent of each other, which implies that all cross-correlations between two different modes in the error covariance matrix are zeros. Therefore, we can decompose the total error covariance matrix $\mathbf{P}$ given by equation (4.1) at the bottom of the page into a number of lower-dimensional matrices

$$\mathbf{P}_i = \begin{bmatrix} P_{(i,i)} & P_{(i,N+i)} \\ P_{(N+i,i)} & P_{(N+i,N+i)} \end{bmatrix} \tag{4.2}$$

where $\mathbf{P}_i$ is the reduced error covariance matrix for the ith Fourier mode. Then, the equation for $\mathbf{P}_i$ can be found, by re-

$$\mathbf{P} = \left[\begin{array}{ccc:ccc} P_{(1,1)} & 0 & \cdots & P_{(1,N+1)} & 0 & \cdots \\ 0 & P_{(2,2)} & \cdots & 0 & P_{(2,N+2)} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ \hdashline P_{(N+1,1)} & 0 & \cdots & P_{(N+1,N+1)} & 0 & \cdots \\ 0 & P_{(N+2,2)} & \cdots & 0 & P_{(N+2,N+2)} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{array}\right], \tag{4.1}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

placing the discrete time formulation in (3.7) by the continuous time formulation [21], as

$$\dot{\mathbf{P}}_i = \mathbf{F}_i \mathbf{P}_i + \mathbf{P}_i \mathbf{F}_i^{\mathbf{T}} \tag{4.3}$$

where $\mathbf{F}_i$ is the reduced linear transition matrix corresponding to the ith Fourier mode

$$\mathbf{F}_i = \begin{bmatrix} 0 & |k_i| \\ -g & 0 \end{bmatrix}. \tag{4.4}$$

To find an analytic solution of the prediction equations for $\mathbf{P}_i$ given by (4.3), we first vectorize the reduced error covariance matrices $\mathbf{P}_i$. As $\mathbf{P}_i$ are symmetric matrices, *i.e.*, $P_{(i,N+i)} = P_{(N+i,i)}$, we can adopt half-vectorization, *i.e.*, the corresponding error covariance vector can be obtained from the lower triangular part of $\mathbf{P}_i$. Then, equation (4.3) is reduced to the following system of linear differential equations:

$$\begin{bmatrix} \dot{\alpha}_i \\ \dot{\beta}_i \\ \dot{\gamma}_i \end{bmatrix} = \begin{bmatrix} 0 & 2|k_i| & 0 \\ -\mathrm{g} & 0 & |k_i| \\ 0 & -2\mathrm{g} & 0 \end{bmatrix} \begin{bmatrix} \alpha_i \\ \beta_i \\ \gamma_i \end{bmatrix} \qquad \text{with}$$
$$\begin{bmatrix} \alpha_i \\ \beta_i \\ \gamma_i \end{bmatrix} = \begin{bmatrix} P_{(i,i)} \\ P_{(i,N+i)} \\ P_{(N+i,N+i)} \end{bmatrix}. \tag{4.5}$$

We should stress again that this simplification is possible since we develop a data assimilation scheme in Fourier space (in terms of Fourier coefficients).

The analytic solution of (4.5) can be obtained as

$$\begin{bmatrix} \alpha_i(t_0+T_u) \\ \beta_i(t_0+T_u) \\ \gamma_i(t_0+T_u) \end{bmatrix} = \mathbf{\Psi}_i(T_u) \begin{bmatrix} \alpha_i(t_0) \\ \beta_i(t_0) \\ \gamma_i(t_0) \end{bmatrix} \tag{4.6}$$

where $\mathbf{\Psi}_i$ are the transition matrices, which depend on the update period T_u, but are independent of the choice of initial time t_0: see equation (4.7) at the bottom of the page, where ω_i is the angular frequency for the ith Fourier mode given, from the linear dispersion relation, by ${\omega_i}^2 = \mathrm{g}|k_i|$.

There are a few special cases of interest that deserve some attention. Firstly, for $\omega_i T_u = l\pi$ ($l = 1, 2, 3, \ldots$), notice that $\mathbf{\Psi}_i$ become the identity matrix $\mathbf{I}$. In other words, if the update period T_u is an integer multiple of a half of the wave period, the error covariance matrix is periodic with a period of T_u. Even for this special case, it should be noticed that, although $P_{(i,j)}$ return to the original values through the prediction steps from $t = (m-1)T_u$ to $t = mT_u$, they are modified when the Kalman filter is applied at $t = mT_u$. Secondly, for $\omega_i = 0$, the transition matrix $\mathbf{\Psi}_0(T_u)$ is no longer the identity matrix, but becomes

$$\mathbf{\Psi}_0(T_u) = \begin{bmatrix} 1 & 0 & 0 \\ -gT_u & 1 & 0 \\ g^2T_u^2 & -2gT_u & 1 \end{bmatrix}. \tag{4.8}$$

The initial conditions for the reduced error covariance matrices $\mathbf{P}_i$ have often been chosen arbitrarily, but the same for all i. In this paper, $\mathbf{P}_i$ are set initially to $10\,\mathbf{I}$, where $\mathbf{I}$ is the 2×2 identity matrix; in other words, $\alpha_i(0) = \gamma_i(0) = 10$ and $\beta_i(0) = 0$. The choice of initial conditions for $\mathbf{P}_i$ will be further discussed in Section IV-C.

B. Explicit Update Procedures

Now the analytic solution for the time evolution of the error covariance matrix $\mathbf{P}$ is available from (4.6). Since the error covariance matrix needs to be evaluated only when numerical solutions are updated with measurements, it is no longer necessary to compute the error covariance matrix $\widehat{\mathbf{P}}$ using (3.7) at every prediction time step. Therefore, one can avoid a large number of multiplications between time-dependent matrices.

By letting $t_0 = (m-1)T_u$ with $m = 1, 2, \ldots$, the reduced error covariance vectors at the mth update step (or at $t = mT_u$) can be found from (4.6) and, from (4.1)–(4.2), can be reassembled to construct the corresponding error covariance matrix with $P_{(i,N+i)} = P_{(N+i,i)} = \beta_i$. Then, from (3.10), the reduced optimal-Kalman-gain matrix $\mathbf{K}_i$ for the ith Fourier component at $t = mT_u$ (or, equivalently, $n = mn_u$ with $n_u = T_u/\Delta t$) is found as

$$\mathbf{K}_i^{mn_u} = \begin{bmatrix} \frac{\widehat{\alpha}_i}{1+\widehat{\alpha}_i} \\ \frac{\widehat{\beta}_i}{1+\widehat{\alpha}_i} \end{bmatrix} \tag{4.9}$$

where $\widehat{f}_i \equiv f_i\big(mT_u\big)$ with $f_i = (\alpha_i, \beta_i, \gamma_i)$ represent the elements of the reduced error covariance vector at $t = mT_u$ right before they are updated.

Finally, by substituting (4.9) into (3.8)–(3.9), the updated state variables $\overline{\mathbf{X}}_i^{mn_u}$ and the error covariance matrices $\overline{\mathbf{P}}_i^{mn_u}$ are obtained as

$$\overline{\mathbf{X}}_i^{mn_u} = \widehat{\mathbf{X}}_i^{mn_u} + \begin{bmatrix} \frac{\widehat{\alpha}_i}{1+\widehat{\alpha}_i} \\ \frac{\widehat{\beta}_i}{1+\widehat{\alpha}_i} \end{bmatrix} (Y_i - Z_i), \tag{4.10}$$
$$\overline{\mathbf{P}}_i^{mn_u} = \left[\mathbf{I} - \mathbf{K}^{mn_u}\,\mathbf{H}\right] \widehat{\mathbf{P}}^{mn_u}$$

$$\mathbf{\Psi}_i(T_u) = \begin{bmatrix} \frac{(\cos(2\omega_i T_u)+1)}{2} & \left(\frac{\omega_i}{\mathrm{g}}\right)\sin(2\omega_i T_u) & -\left(\frac{\omega_i^2}{2\mathrm{g}^2}\right)(\cos(2\omega_i T_u)-1) \\ -\left(\frac{\mathrm{g}}{2\omega_i}\right)\sin(2\omega_i T_u) & \cos(2\omega_i T_u) & \left(\frac{\omega_i}{2\mathrm{g}}\right)\sin(2\omega_i T_u) \\ -\left(\frac{\mathrm{g}^2}{2\omega_i^2}\right)(\cos(2\omega_i T_u)-1) & -\left(\frac{\mathrm{g}}{\omega_i}\right)\sin(2\omega_i T_u) & \frac{(\cos(2\omega_i T_u)+1)}{2} \end{bmatrix} \tag{4.7}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

$$= \begin{bmatrix} \frac{\widehat{\alpha}_i}{1+\widehat{\alpha}_i} & \frac{\widehat{\beta}_i}{1+\widehat{\alpha}_i} \\ \frac{\widehat{\beta}_i}{1+\widehat{\alpha}_i} & \widehat{\gamma}_i - \frac{\widehat{\beta}_i{}^2}{1+\widehat{\alpha}_i} \end{bmatrix} \tag{4.11}$$

where $\mathbf{X}_i = [Z_i,\ \Pi_i]^T$. Once the update procedures for $n = mn_u$ are completed, $\widehat{\mathbf{X}}_i^{mn_u}$ and $\widehat{\mathbf{P}}_i^{mn_u}$ need to be replaced by $\overline{\mathbf{X}}_i^{mn_u}$ and $\overline{\mathbf{P}}_i^{mn_u}$, respectively, before the next prediction time step starts. The nonlinear evolution equations given by (3.6) are then solved numerically to find $\widehat{\mathbf{X}}_i$ without updating them with measurements until the time step reaches the next update time step of $(m+1)n_u$. The update procedure described here will be referred to as the *explicit* Kalman filter.

As a special case, when the wave frequencies satisfy $\omega_i T_u = l\pi$ ($l = 1, 2, \ldots$), for which the error covariance matrix is periodic in time, it can be shown that, after the mth update, the updated state variables and error covariance matrix at $t = mT_u$ can be computed easily from their predicted values as

$$\overline{\mathbf{X}}_i^{mn_u} = \widehat{\mathbf{X}}_i^{mn_u} + \begin{bmatrix} \frac{\alpha_i^0}{1+m\alpha_i^0} \\ \frac{\beta_i^0}{1+m\alpha_i^0} \end{bmatrix} (Y_i - Z_i), \tag{4.12}$$

$$\overline{\mathbf{P}}_i^{mn_u} = \begin{bmatrix} \frac{\alpha_i^0}{1+m\alpha_i^0} & \frac{\beta_i^0}{1+m\alpha_i^0} \\ \frac{\beta_i^0}{1+m\alpha_i^0} & \gamma_i^0 - \frac{m(\beta_i^0)^2}{1+m\alpha_i^0} \end{bmatrix} \tag{4.13}$$

where $(\alpha_i^0, \beta_i^0, \gamma_i^0)$ denote the initial values of $(\alpha_i, \beta_i, \gamma_i)$ at $t = 0$. Notice that $(\alpha_i, \beta_i) \to 0$ as $t \to \infty$ or $m \to \infty$ while $\gamma_i \to \gamma_i^0 - (\beta_i^0)^2/\alpha_i^0$. This implies that, for these special frequencies of $\omega_i T_u = l\pi$, the error covariance matrix converge to a steady (constant) state as time increases. For other frequencies, it is nontrivial to find explicitly the behavior of the error covariance matrix as $t \to \infty$. However, since the filter system based on the linearized wave model is observable for the given $\mathbf{F}$ and $\mathbf{H}$ matrices, the error covariance matrix for frequencies different from $\omega_i T_u = l\pi$ is also expected to converge to a steady state [22].

On the other hand, for the zeroth Fourier component of $\omega_i T_u = 0$, it can be shown, from (4.6) with (4.8), that

$$\overline{\mathbf{X}}_i^{mn_u} = \widehat{\mathbf{X}}_i^{mn_u} + \begin{bmatrix} \frac{\alpha_i^0}{1+m\alpha_i^0} \\ \frac{\beta_i^0 - m\alpha_i^0 gT_u}{1+m\alpha_i^0} \end{bmatrix} (Y_i - Z_i), \tag{4.14}$$

$$\overline{\mathbf{P}}_i^{mn_u} = \begin{bmatrix} \frac{\alpha_i^0}{1+m\alpha_i^0} & \frac{\beta_i^0 - m\alpha_i^0 gT_u}{1+m\alpha_i^0} \\ \frac{(\beta_i^0 - m\alpha_i^0 gT_u)}{1+m\alpha_i^0} & \gamma_i^0 - \frac{\left(m\beta_i^{0\,2} + 2m\beta_i^0 gT_u - m^2\alpha_i^0 g^2 T_u{}^2\right)}{1+m\alpha_i^0} \end{bmatrix}. \tag{4.15}$$

It should be remarked that, for conservation of mass, the zeroth Fourier component for the predicted surface elevation Z_0 must remain unchanged after the update steps. From (4.14), one can see that this requires the zeroth Fourier component of the measured surface elevation must satisfy $Y_0 = 0$ if the initial value of Z_0 is 0. Therefore, before the update steps, this should be enforced to the measured data if their mean values are different from 0.

C. Error Analysis for the Explicit Data Assimilation Scheme

A data assimilation scheme using a standard Kalman filter could be numerically unstable due to the use of a forward Euler time integration scheme, which is only conditionally stable, as demonstrated through numerical simulations in Section V. It is therefore of interest to examine the stability characteristics of the explicit data assimilation scheme proposed in this paper.

For linear stability analysis, we consider the true state vectors $\mathbf{X}_i$ as

$$\mathbf{X}_i = \begin{bmatrix} Z_i \\ \Pi_i \end{bmatrix} \tag{4.16}$$

which are assumed to be governed by the linearized wave prediction model (3.5) so that

$$\dot{\mathbf{X}}_i = \mathbf{F}_i\, \mathbf{X}_i \tag{4.17}$$

where $\mathbf{F}_i$ are given by (4.4). When the error vectors $\mathbf{e}_i$ are defined by the difference between the true state vectors $\mathbf{X}_i$ and the estimated state vectors $\widehat{\mathbf{X}}_i$

$$\mathbf{e}_i = \mathbf{X}_i - \widehat{\mathbf{X}}_i \tag{4.18}$$

they should satisfy (4.17) since both $\mathbf{X}_i$ and $\widehat{\mathbf{X}}_i$ are governed by (4.17). Then, the analytic solution for $\mathbf{e}_i$ can be easily obtained, after elapsed time T_u from t_0, as

$$\mathbf{e}_i(t_0 + T_u) = \begin{bmatrix} C_i & \left(\frac{\omega_i}{\mathrm{g}}\right) S_i \\ -\left(\frac{\mathrm{g}}{\omega_i}\right) S_i & C_i \end{bmatrix} \mathbf{e}_i(t_0) \equiv \mathbf{\Omega}_i(T_u)\, \mathbf{e}_i(t_0) \tag{4.19}$$

where $C_i \equiv \cos(\omega_i T_u)$ and $S_i \equiv \sin(\omega_i T_u)$.

At $t = mT_u$ when the mth update is to be made, the prediction error $\widehat{\mathbf{e}}_i^n$ propagated from the previously updated error at $t = (m-1)T_u$ can be found analytically as

$$\widehat{\mathbf{e}}_i^{mn_u} = \mathbf{\Omega_i}(T_u)\, \overline{\mathbf{e}}_i^{(m-1)n_u}. \tag{4.20}$$

Then, with assuming that the measured value vector $\mathbf{Y}$ satisfies $\mathbf{Y} = \mathbf{H}\ \mathbf{X}$ and subtracting $\mathbf{X}$ from both sides of (3.8), the error after the update step $\overline{\mathbf{e}}_i^{mn_u}$ can be determined by

$$\overline{\mathbf{e}}_i^{mn_u} = (\mathbf{I} - \mathbf{K}_i^{mn_u}\mathbf{H})\, \widehat{\mathbf{e}}_i^{mn_u} = \begin{bmatrix} \frac{1}{+\widehat{\alpha}_i} & 0 \\ -\frac{\widehat{\beta}_i}{1+\widehat{\alpha}_i} & 1 \end{bmatrix} \widehat{\mathbf{e}}_i^{mn_u} \tag{4.21}$$

where $\mathbf{K}_i^{mn_u}$ are given by (4.9). Notice that, for this error analysis, the measured state vector $\mathbf{Y}$ has been assumed to be free from measurement noise so that the error is introduced only through initial conditions for simplicity. When noise $\mathbf{r}$ is superimposed to $\mathbf{Y}$ such that $\mathbf{Y} = \mathbf{H}\ \mathbf{X} + \mathbf{r}$, it adds a constant vector to the right-hand side of (4.21).

Finally, from (4.20)–(4.21), the error propagated from the $(m-1)$th update to the mth update can be expressed as

$$\overline{\mathbf{e}}_i^{mn_u} = \left[\mathbf{I} - \mathbf{K}_i^{mn_u}\mathbf{H}\right] \mathbf{\Omega}_i(T_u)\, \overline{\mathbf{e}}_i^{(m-1)n_u} \equiv \mathbf{A}_m\, \widehat{\mathbf{e}}_i^{(m-1)n_u} \tag{4.22}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

where $\mathbf{A}_m$ is the transition matrix, describing the time evolution of $\mathbf{e}_i$ from $t = (m-1)T_u$ to $t = mT_u$ through both the prediction and update steps, defined by

$$\mathbf{A}_m = \begin{bmatrix} \frac{C_i}{1+\widehat{\alpha}_i} & \frac{\frac{\omega_i}{g} S_i}{1+\widehat{\alpha}_i} \\ -\frac{g}{\omega_i} S_i - C_i \frac{\widehat{\beta}_i}{1+\widehat{\alpha}_i} & C_i - \frac{\frac{\omega_i}{g} S_i \widehat{\beta}_i}{1+\widehat{\alpha}_i} \end{bmatrix} \tag{4.23}$$

where $\widehat{\alpha}_i$ and $\widehat{\beta}_i$ are the values of α_i and β_i at $t = mT_u$ prior to the update step. Then, after the mth update (with $n = mn_u$), the error evolved from the initial error $\mathbf{e}_i^0$ can be found as $\overline{\mathbf{e}}_i^{mn_u} = \mathbf{A}_{m,1}\, \mathbf{e}_i^0$, where $\mathbf{A}_{m,1} = \mathbf{A}_m \mathbf{A}_{m-1} \dots \mathbf{A}_1$.

Stability of the explicit data assimilation scheme can therefore be determined by the eigenvalues λ_i of $\mathbf{A}_m$, which are the roots of the following quadratic equation:

$$\lambda_i{}^2 - \frac{C_i(2+\widehat{\alpha}_i) - \left(\frac{\omega_i}{g}\right) S_i \widehat{\beta}_i}{1+\widehat{\alpha}_i} \lambda_i + \frac{1}{1+\widehat{\alpha}_i} = 0. \tag{4.24}$$

Notice (4.6)–(4.7) that $\widehat{\alpha}_i$ and $\widehat{\beta}_i$ depend not only on $\omega_i T_u$, but also on their initial values and m. Therefore, for stability of our data assimilation scheme, the eigenvalues λ_i must satisfy the condition of $\max_i |\lambda_i| < 1$ for all m.

(i) For $\omega_i T_u = l\pi$ with $l = 0, 1, 2, \dots$
Before examining the eigenvalues for all values of ω_i, we consider a special case of $\omega_i T_u = l\pi$ ($l = 0, 1, 2, \dots$), which simplifies (4.24) to

$$\lambda_i{}^2 - (-1)^l \frac{2+\widehat{\alpha}_i}{1+\widehat{\alpha}_i} \lambda_i + \frac{1}{1+\widehat{\alpha}_i} = 0 \tag{4.25}$$

whose roots are given by

$$\lambda_i = (-1)^l, \qquad \text{or} \qquad \lambda_i = \frac{(-1)^l}{1+\widehat{\alpha}_i}. \tag{4.26}$$

This implies that, as long as $\widehat{\alpha}_i > 0$, the data assimilation scheme is always neutrally stable (*i.e.*, $\max |\lambda_i| = 1$) for the waves with $\omega_i T_u = l\pi$, or, equivalently, $T_u/T_i = l$ with T_i being the wave period given by $T_i = 2\pi/\omega_i$. Notice that the condition of $\widehat{\alpha}_i > 0$ is fulfilled when α_i is chosen initially a positive real value, irrespective of m and initial conditions for β_i and γ_i.

(ii) For $\omega_i T_u \neq l\pi$
For frequencies different from $\omega_i T_u = l\pi$ with $l = 1, 2, \dots$, the eigenvalues depending on time dependent $\widehat{\alpha}_i$ and $\widehat{\beta}_i$ should be found numerically. To understand the effect of multiple updates, the eigenvalues of $\mathbf{A}_{m,1} \equiv \mathbf{A}_m \mathbf{A}_{m-1} \dots \mathbf{A}_1$ are computed after a large number of update steps, or, in this paper, $m = 100$ are applied.

As shown in Fig. 1, the absolute values of the eigenvalues of $\mathbf{A}_{100,1}$ are equal to 1, as expected, when $\omega_i T_u = l\pi$ with l being a positive integer; otherwise, they are less than 1 when $(\alpha_i,\, \beta_i,\, \gamma_i) = c\,(1, 0, 1)$ at $t = 0$, where $c = 10$. In fact, as long as c is a positive real value, the absolute values of the eigenvalues are found numerically to be always less than 1. Therefore, it can be concluded that the explicit data assimilation scheme with this choice of $(\alpha_i, \beta_i, \gamma_i)$ at $t = 0$ is neutrally stable and enables one to avoid the instability that might occur when the prediction equation for the error covariance is solved numerically with a conditionally stable time integration scheme.

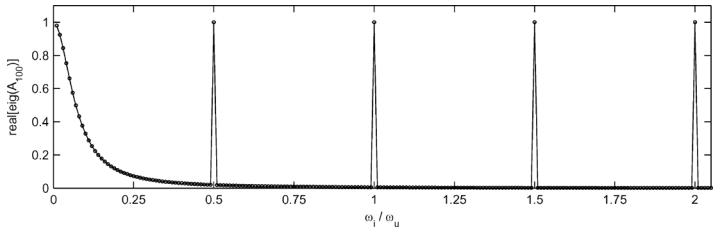


Fig. 1. Absolute values of the maximum eigenvalues of $A_{100,1}$.

When β_i are chosen initially to be nonzero and, in particular, greater than the values of α_i and γ_i, the eigenvalues are found numerically greater than 1. Therefore, for stability, this choice for β_i has to be avoided.

V. Numerical Simulations

A. Numerical Method and Parameters

To test the explicit data assimilation scheme proposed in this paper, a series of numerical experiments are conducted for long-crested (or unidirectional) irregular waves. We first assume that the true wave fields follow the system of third-order nonlinear evolution equations given by (2.11)–(2.12), and obtain ζ_{true} and Φ_{true} for the true wave fields by solving numerically the third-order system. Secondly, with assuming only surface elevation measurements are available from a noncoherent radar, the measured surface elevation $\zeta_{\text{measured}}(x,t)$ is constructed by adding a random noise field $r(x,t)$ to $\zeta_{\text{true}}(x,t)$ so that $\zeta_{\text{measured}} = \zeta_{\text{true}} + r$. Unless otherwise mentioned, the mean and standard deviation of the normal random noise field are chosen $\mu = 0$ and $\sigma = 0.15 H_s$, respectively, for our simulations. No information about the free surface velocity potential Φ is assumed available. The measured surface elevations are then used to update once at every T_u-interval numerical solutions of the third-order system for $\widehat{\zeta}$ and $\widehat{\Phi}$. These estimated (*i.e.,* predicted plus updated) wave fields are in turn compared with the true wave fields to quantify the difference between the two wave fields.

To find ζ_{true}, the surface wave field is initialized with the JONSWAP spectrum characterized by two parameters, the significant wave height H_s and the peak wave frequency ω_p. From the relationship between the frequency spectrum $S_\omega(\omega)$ and the wavenumber spectrum $S_k(k)$ given by

$$S_k(k) \equiv S_\omega(\omega) \frac{\mathrm{d}\omega}{\mathrm{d}k} \tag{5.1}$$

the JONSWAP spectrum originally written in the frequency domain is given, in the wavenumber domain, by

$$S_k(k) = \frac{\alpha k_p^2 H_s^2}{2k^3} \exp\left[-\beta \left(\frac{k_p}{k}\right)^2\right] \gamma^{\exp\left[-(\sqrt{k}-\sqrt{k_p})^2/2\sigma^2 k_p\right]} \tag{5.2}$$

where $\alpha = 5/16$, $\beta = 5/4$, γ is the peak enhancement, and $k_p = \omega_p^2/\mathrm{g}$ is the peak wavenumber corresponding to the peak frequency.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Then, from the two independent definitions of the mean energy density E given by

$$E = \int_0^{k_{\max}} S_k(k)\,\mathrm{d}k \simeq \sum_{i=0}^{N/2} S_k(k_i)\,\Delta k,$$
$$E = \frac{1}{L}\int_0^L \zeta^2\,\mathrm{d}x \simeq \sum_{i=0}^{N/2} 2\,|Z_i|^2 \tag{5.3}$$

with L being the length of the computational domain, the Fourier coefficients of the free surface elevation at $t = 0$ are estimated as

$$Z_i = |Z_i|\,\mathrm{e}^{-2\pi i\delta_i}, \quad |Z_i| = \sqrt{\frac{1}{2}S_k(k_i)\Delta k} \quad \text{at } t = 0 \tag{5.4}$$

with normally distributed random phases δ_i over a range between 0 and 2π. On the other hand, the Fourier coefficients of the initial free surface velocity potential are estimated, using linear theory, by

$$\Pi_i = \mathrm{i}c_i Z_i \quad \text{at } t = 0 \tag{5.5}$$

where the phase velocities c_i for gravity waves in deep water are defined as $c_i = \mathrm{g}/k_i$.

For our simulations, the peak enhancement factor and the peak frequency are fixed to be $\gamma = 3.3$ and $\omega_p = 2\pi$ sec^{-1}, respectively, while three different significant wave heights for the JONSWAP spectrum are considered: $H_s = 0.05$, 0.075, and 0.1. The corresponding wave period T_p and wavenumber k_p are $T_p = 1$ sec and 4.03 m^{-1}, respectively, while the wave steepness defined by $\epsilon_p = k_p H_s/2$ is given by $\epsilon_p = 0.108$, 0.151, and 0.202 m for the three different significant wave heights.

The nonlinear evolution equations given by (2.11)–(2.12) are then solved to obtain a true wave field, ζ_{true} and Φ_{true}, using a pseudo-spectral method based on the fast Fourier transform with integrating in time using a fourth-order Runge-Kutta scheme with a time step of $\Delta t/T_p = 0.02$.

The length of the computational domain is $L = 128$ m, which corresponds approximately to 82 peak wavelengths. The total number of Fourier modes is chosen to be $N_F = 1024$. However, to remove aliasing errors due to the use of finite Fourier series, the highest one-half Fourier modes are filtered out for every time step for the third-order nonlinear computation [23]. Thus, the number of physically-effective Fourier modes including both positive and negative wavenumbers is 512.

To test our data assimilation scheme, we repeat our computations for $\widehat{\zeta}$ and $\widehat{\Phi}$ with the same parameters used to compute ζ_{true} and Φ_{true} except for initial conditions. Instead of initializing the surface wave fields using equations (5.4)–(5.5), we assume that there is initially no information about the true wave field and, therefore, the Fourier coefficients of initial wave fields are set to zero so that

$$Z_i = 0, \qquad \Pi_i = 0 \qquad \text{at } t = 0. \tag{5.6}$$

Then, we update the solution using (4.10) with an update period of $T_u = 0.4$ sec or the update frequency of $\omega_u = 15.71$ rad/sec. Notice that the dimension of the state vector (or the Fourier coefficients of the surface elevation) N is 256, a quarter of $N = 1024$, which is the number of positive wavenumber modes.

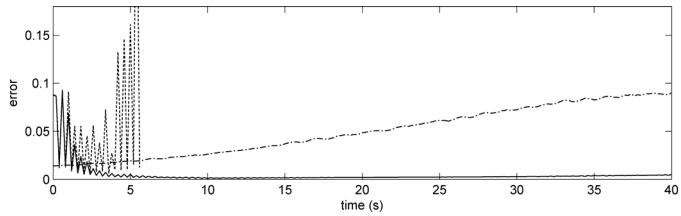


Fig. 2. Time evolution of the normalized error defined by (5.7) for the explicit data assimilation scheme with measurements generated by a true wave field of $H_s = 0.075$ m plus normal random noise with $\mu = 0$ and $\sigma = 0.15H_s$. Solid line: the explicit Kalman filter; dotted line: the classical Kalman filter; dashed-dotted line: no filtering.

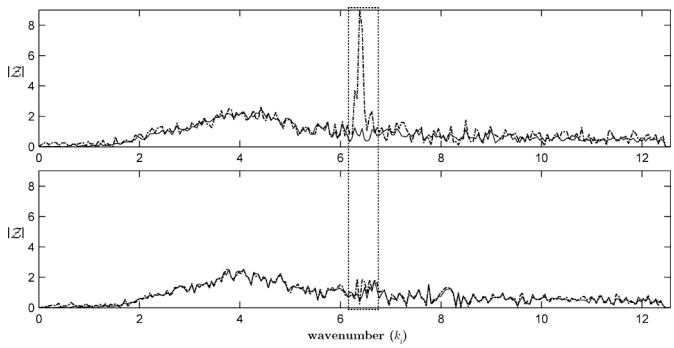


Fig. 3. Fourier coefficients of the free surface elevation at $t/T_p = 4.08$ for the numerical simulations presented in Fig. 2 with the standard (top) and explicit (bottom) Kalman filters. Notice that the numerical solution using the standard Kalman filter is about to become unbounded at this instant. In each panel, the true free surface elevation (solid line) is compared with the estimated free surface elevation (dotted line).

All filtered simulations presented here are performed using the either standard or explicit Kalman filter for the first 10 wave periods ($t/T_p = 10$). Then, the update step is turned off and no update is made afterwards. To monitor the filter performance, we measure the normalized error between the true and estimated wave elevations as

$$\text{error} = \frac{\int_0^L [\zeta_{\mathrm{true}}(x) - \widehat{\zeta}(x)]^2\,\mathrm{d}x}{{H_s}^2 L}. \tag{5.7}$$

B. Comparison Between Different Filters

For a given set of physical parameters, we perform numerical simulations with three different filters: (i) the standard Kalman filter based on a numerically-computed $\mathbf{P}$, as described in Section III-C, with zero initial conditions; (ii) the explicit Kalman filter based on $\mathbf{P}$ found analytically in Section IV, with zero initial conditions; (iii) no filter with initial conditions given by ζ_{true} and Φ_{true} superimposed with normal random noise whose mean and standard deviation are $\mu = 0$ and $\sigma = 0.15H_s$, respectively.

Errors from the true wave field for three different simulations are computed using (5.7) and their time evolutions are shown in Fig. 2. When no filter is applied, the error introduced through the imposed initial noise grows almost linearly in time and no reliable prediction can be made. This implies that noisy measurements cannot be used directly without any filtering. When

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

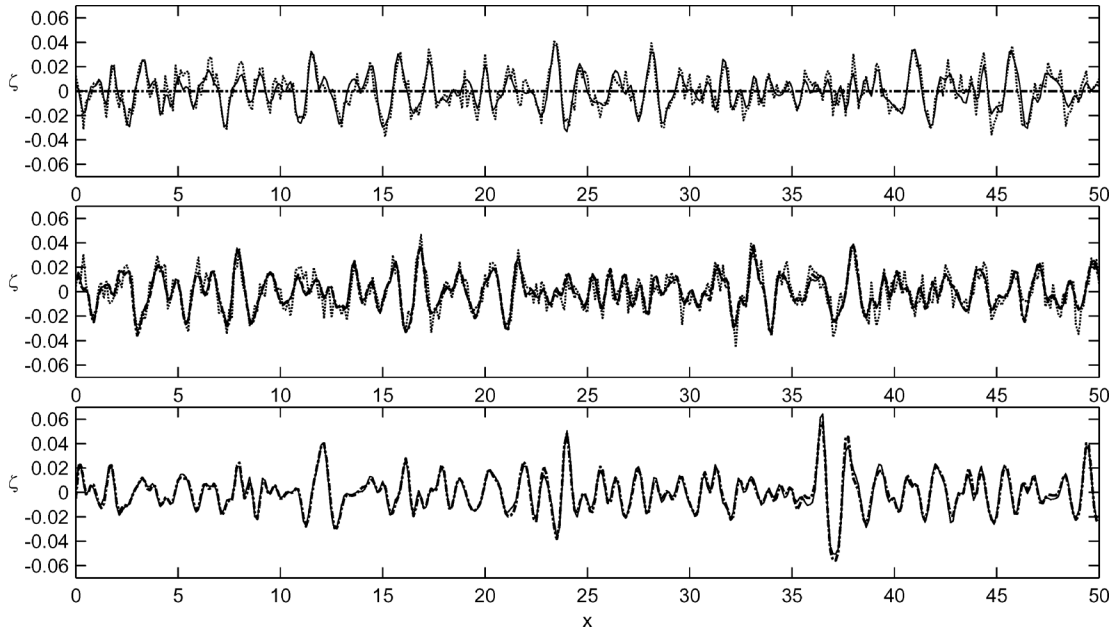


Fig. 4. Comparison of the true (solid line), measured (dashed line), and estimated (thick dash-dot line) free surface elevations at $t/T_p = 0$, $t/T_p = 10$, and $t/T_p = 60$ (from top to bottom). The explicit Kalman filter is used and the wave filed is initialized with $H_s = 0.05$ m and $k_p = 4.03\ \mathrm{m}^{-1}$ ($\epsilon_p = k_p H_s/2 = 0.108$).

the standard Kalman filter is used, the error oscillates with an amplitude much larger than the noise level and grows rapidly in time. Soon the wave prediction model becomes unstable and blows up in a finite time (less than $t/T_p = 6$) due to the conditionally stable forward Euler time-integration scheme in the prediction step along with the periodic update with noisy measurements, as discussed in Section III-C. This indicates that the standard Kalman filter should be used with great care with the pseudo-spectral wave prediction model. On the other hand, the simulation using the explicit Kalman filter shows a promising result where the error is maintained at the level of noise or even smaller. This observation supports our stability analysis presented in Section IV-C, where the linearized wave prediction system is shown neutrally stable.

A more detailed comparison between three simulations can be found in Fig. 3, where we compare the Fourier coefficients of the surface elevation at $t/T_p = 4.08$, at which the numerical solution based on the standard Kalman filter becomes unbounded. Notice that the Fourier coefficient for $k_i = 6.29\ \mathrm{m}^{-1}$ has grown to a much higher value than those of any other wave number modes. The wave frequency ω_i corresponding to this dangerous wave number is approximately $\omega_i = 7.855$ rad/sec, which is one half of the update frequency $\omega_u = 15.708$ rad/sec so that $\omega_i/\omega_u \approx 1/2$. This is the smallest frequency ratio for neutral stability for which the maximum eigenvalue is one, as described in Section IV-C. Due to extra perturbations introduced by numerical approximation might produce this instability. The result in Fig. 3 demonstrates that the wave prediction system with the explicit Kalman filter is indeed neutrally stable, as predicted in Section IV-C, and shows no sign of instability at one half the update frequency.

It should be emphasized again that, in addition to being stable, the explicit Kalman filter is computationally much more efficient than the standard Kalman filter since it is no longer necessary to compute numerically the error covariance matrix at every prediction time step. For the two-dimensional wave evolution, the computational efficiency of the explicit Kalman filter will be much more pronounced.

C. Numerical Simulations With the Explicit Kalman Filter

Here we describe detailed numerical simulation results using the explicit Kalman filter. For each simulation presented in this section, the wave field is predicted for $t/T_p = 10$ with assimilating measurements at every update period T_u and, then, is predicted for the next $t/T_p = 50$ after the data assimilation scheme is turned off.

Figs. 4 and 5 show the time evolution of the estimated free surface elevation using the explicit Kalman filter compared with the true and measured free surface elevations at $t/T_p = 0$ (initial state), $t/T_p = 10$ (when the update step is turned off), and $t/T_p = 60$ for two different wave steepnesses: $\epsilon_p = 0.108$ and 0.202. As shown in Fig. 4, the estimated free surface elevation matches well with the true free surface elevation for the case of $\epsilon_p = 0.108$. However, it can be seen in Fig. 5 that, as the wave steepness increases, the difference between the true and estimated states grows after the filter is turned off. This can be seen more clearly in Fig. 6, where the time evolution of the normalized error defined by (5.7) is presented. This demonstrates that the noise introduced by assimilated measurements through the update steps for $t/T_p \leq 10$ makes numerical simulations deviate in time from the true states, in particular, when the wave steepness is no longer small. This behavior is due to the fact that the explicit Kalman filter is based on the wave model linearized about zero states. Thus the data assimilation scheme fails to distinguish (nonlinear) bound waves from free waves, particularly, in the estimation of the velocity potential. To be discussed later, this shows clearly a need of a more robust data assimilation

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

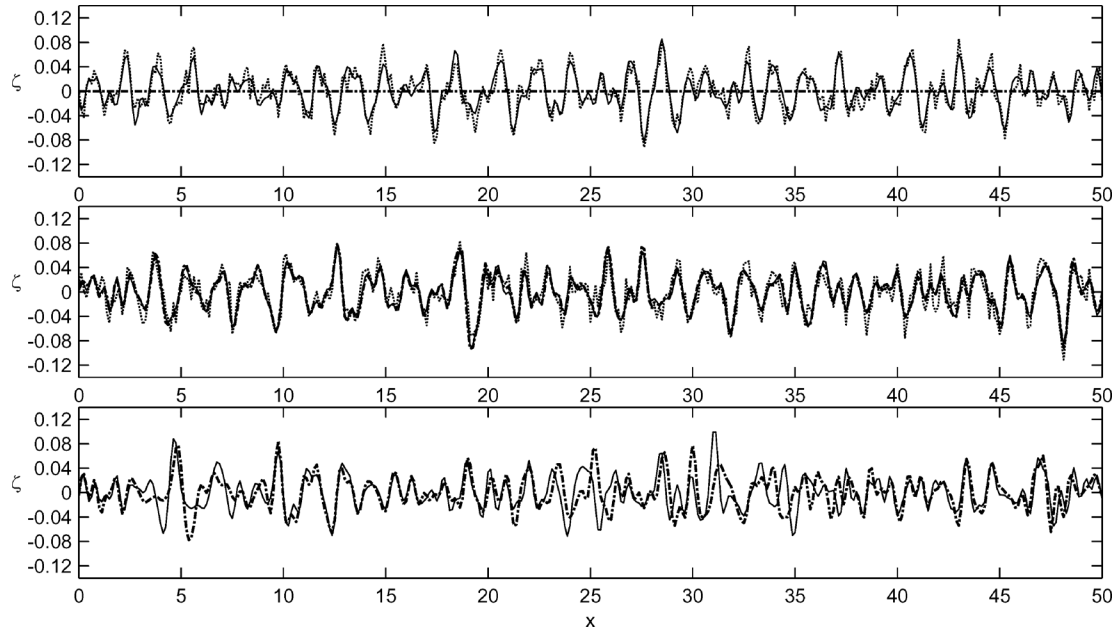


Fig. 5. Comparison of the true (solid line), measured (dashed line), and estimated (thick dash-dot line) free surface elevations at $t/T_p = 0$, $t/T_p = 10$, and $t/T_p = 60$ (from top to bottom). The explicit Kalman filter is used and the wave filed is initialized with $H_s = 0.1$ m and $k_p = 4.03\ \mathrm{m}^{-1}$ ($\epsilon_p = k_p H_s/2 = 0.202$).

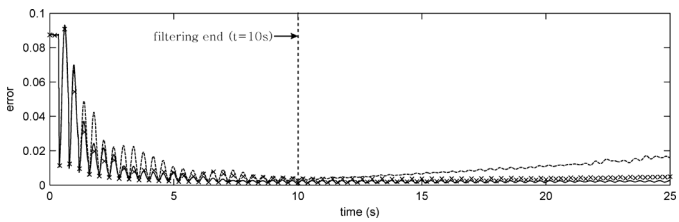


Fig. 6. Time evolution of the normalized error defined by (5.7) for the explicit data assimilation scheme. Solid line: $H_s = 0.05$ ($\epsilon_p = 0.108$); cross symbol: $H_s = 0.075$ ($\epsilon_p = 0.151$); dashed line: $H_s = 0.10$ ($\epsilon_p = 0.202$). The measurements for data assimilation are generated by true wave fields plus normal random noise with $\mu = 0$ and $\sigma = 0.15H_s$.

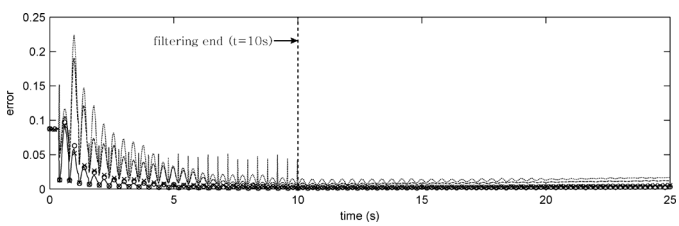


Fig. 7. Time evolution of the normalized error defined by (5.7) for a wave field initialized with $H_s = 0.075$ m ($\epsilon_p = 0.108$). Solid line: normal random noise with $\mu = 0$, $\sigma = 0.15H_s$; dashed line: normal random noise with $\mu = 0, \sigma = 0.30H_s$; cross symbol: Weibull random noise with $b = 1, \sigma = 0.15H_s$; dotted line: Weibull random noise with $b = 1$, $\sigma = 0.30H_s$; circle symbol: Weibull random noise with $b = 2, \sigma = 0.15H_s$.

scheme for the prediction of highly nonlinear wave fields when only noisy measurements are available.

To examine the performance of the explicit Kalman filter for different noise characteristics, a number of synthetic measurements are generated by adding to true states random noise with three different statistical characteristics: normal distribution, Weibull distribution with the shape parameter $b = 1$, and Weibull distribution with $b = 2$. In addition, two different standard deviations, $\sigma = 0.15H_s$ and $\sigma = 0.3H_s$ are used for the first two distributions. Fig. 7 compares the error between true and estimated wave elevations for five different statistical characteristics of random noise. There is a slight increase in error when the standard deviation becomes $0.3H_s$, but the error does not significantly affect the overall performance of the assimilation scheme, including the case of the biased Weibull random noise. In Fig. 7, we can observe a jump in the error at each update, but this error is filtered out just after the next time step.

VI. Conclusion

In this paper, we have presented an explicit data assimilation scheme for a deterministic wave prediction model for nonlinear gravity waves. The nonlinear wave prediction model is based on asymptotic expansion originally proposed by West *et al.* [6] and is solved numerically using a pseudo-spectral method, where the Fourier coefficients of the surface elevation and the free surface velocity potential are computed at every time step. It is shown that the time evolution of the error covariance matrix for each Fourier mode is independent from other modes and can be found explicitly by solving analytically the linearized wave prediction model. Through both an error analysis and nonlinear numerical simulations, it is demonstrated that this explicit Kalman filter is not only numerically efficient, but also stable when noisy measurements are assimilated with the wave prediction model.

The explicit data assimilation scheme developed for the wave evolution in infinitely deep water can be easily applied to the case of finite-depth or shallow water when the integral operator $\mathcal{L}$ in (2.11)–(2.12) is redefined as $\mathcal{L}[\Phi] = -\mathcal{F}^{-1}[k \tanh(kh)\, \mathcal{F}[\Phi]]$, where h is the water depth. It is assumed that only surface elevation measurements are available from a noncoherent radar. When one uses a coherent radar from which radial velocity measurements are also available, the

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

measurement sensitivity matrix $\mathbf{H}$ can be modified without difficulty. A generalization to the two-dimensional wave evolution is also straightforward when the Fourier coefficients of state variables in the two-dimensional Fourier space are rearranged as a vector. One issue for two-dimensional wave fields might be the increase of dimension of the state vector, in particular, when the physical domain of interest is large. Since all Fourier modes are decoupled in the determination of the error covariance matrices, it should be of little concern. Nevertheless, if needed, the dimension of the state vector to be updated can be reduced by applying the explicit Kalman filter to a subset of Fourier modes (typically, of wave periods between 5 s and 20 s) that have relatively significant energy and could affect offshore structures.

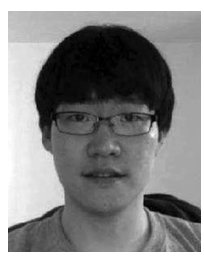
A couple of nontrivial improvements should be topics of future study. The explicit data assimilation scheme is shown reliable for most not-so-severe oceanic conditions and, when combined with the nonlinear wave model, would therefore serve as an efficient daily operational tool for the prediction of ocean waves. Nevertheless, the explicit Kalman filter considered here is based on linearization about zero states. For the prediction of more severe sea states, the scheme should be modified, without losing its efficiency, to take nonlinearity into account not only for prediction steps, but also for update steps. Another improvement should be made to handle temporal measurements of the surface elevation. Here it is assumed that spatial measurements of the surface elevation are available from a marine radar at every update period, but it is common to have only temporal measurements of the surface elevation from wave buoys only at a few spatial locations. Then, one should convert them to snapshots of the corresponding wave fields. This conversion can be made easily for linear wave fields using Fourier analysis, but must be done with care for nonlinear wave fields.

ACKNOWLEDGMENT

W. Choi would like to thank T. Waseda at the University of Tokyo for his helpful remarks on this work.

REFERENCES

[1] M. Onorato, A. R. Osborne, M. Serio, and S. Bertone, "Freak waves in random oceanic sea states," *Phys. Rev. Lett.*, vol. 86, p. 5831, 2001.
[2] P. K. Shukla, I. Kourakis, B. Eliasson, M. Marklund, and L. Stenflo, "Instability and evolution of nonlinearly interacting water waves," *Phys. Rev. Lett.*, vol. 97, p. 094501, 2006.
[3] A. C. Voorrips, A. W. Heemink, and G. J. Komen, "Wave data assimilation with the Kalman filter," *J. Mar. Syst.*, vol. 19, pp. 267–291, 1999.
[4] J. P. Pinto, M. C. Bernadino, and A. Pires Silva, "A Kalman filter application to a spectral wave model," *Nonlinear Process. Geophys.*, vol. 12, pp. 775–782, 2005.
[5] A. Zamani, A. Azimian, A. Heemink, and D. Solomatine, "Non-linear wave data assimilation with an ANN-type wind-wave model and Ensemble Kalman Filter (EnKF)," *Appl. Math. Modeling*, vol. 34, pp. 1984–1999, 2010.
[6] B. J. West, K. A. Brueckner, R. S. Janda, D. M. Milder, and R. L. Milton, "A new numerical method for surface hydrodynamics," *J. Geophys. Res.*, vol. 92, p. 11803, 1987.
[7] M. Tanaka, "A method of studying nonlinear random field of surface gravity waves by direct simulation," *Fluid Dyn. Res.*, vol. 28, pp. 41–60, 2001.
[8] M. Tanaka, "On the role of resonant interactions in the short-term evolution of deep-water ocean spectra," *J. Phys. Oceanogr.*, vol. 37, pp. 1022–1036, 2007.
[9] W. Choi, C. P. Kent, and C. Schillinger, "Numerical modeling of nonlinear surface waves and its validation," in *Advances in Engineering Mechanics*, A. T. Chwang, M. H. Teng, and D. T. Valentine, Eds. Singapore: World Scientific, 2005, pp. 94–110.
[10] Z. Tian, M. Perlin, and W. Choi, "Evaluation of a deep-water wave breaking criterion," *Phys. Fluids*, vol. 20, p. 066604, 2008.
[11] A. Goullet and W. Choi, "A numerical and experimental study on the nonlinear evolution of long-crested irregular waves," *Phys. Fluids*, vol. 23, p. 016601, 2011.
[12] Z. Tian, M. Perlin, and W. Choi, "Energy dissipation in two-dimensional unsteady plunging breakers and an eddy viscosity model," *J. Fluid Mech.*, vol. 655, p. 217, 2010.
[13] M. Perlin, W. Choi, and Z. Tian, "Breaking waves in deep and intermediate waters," *Ann. Rev. Fluid Mech.*, vol. 45, pp. 115–145, 2013.
[14] Z. Tian, M. Perlin, and W. Choi, "Frequency spectra evolution of two-dimensional focusing wave groups in finite depth water," *J. Fluid Mech.*, vol. 688, pp. 169–194, 2011.
[15] Z. Tian, M. Perlin, and W. Choi, "An eddy viscosity model for two-dimensional breaking waves and its validation with laboratory experiments," *Phys. Fluids*, vol. 24, p. 036601, 2012.
[16] Z. Tian and W. Choi, "Evolution of deep-water waves under wind forcing and wave breaking effects: Numerical simulations and experimental assessment," *Eur. J. Mech. B Fluids*, vol. 41, pp. 11–22, 2013.
[17] S. Aragh and O. Nwogu, "Assimilating synthetic radar data into a pseudo-spectral surface wave model," *J. Coastal Res.*, vol. 52, pp. 235–244, 2009.
[18] V. E. Zakharov, "Stability of periodic waves of finite amplitude on the surface of deep fluid," *J. Appl. Mech. Tech. Phys.*, vol. 9, pp. 190–194, 1968.
[19] W. Choi, "Nonlinear evolution equations for two-dimensional surface waves in a fluid of finite depth," *J. Fluid Mech.*, vol. 295, pp. 381–394, 1995.
[20] R. E. Kalman, "A new approach to linear filtering and Prediction Problems," *J. Basic Eng.*, vol. 82, pp. 35–45, 1960.
[21] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.
[22] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ, USA: Prentice Hall, 1980.
[23] S. A. Orszag, "On the elimination of aliasing in finite-difference schemes by filtering high-wavenumber components," *J. Atmos. Sci.*, vol. 28, p. 1074, 1971.



Seongjin Yoon received the B.S degree in naval architecture and ocean engineering from the Seoul National University, Seoul, Korea, and the M.S. degree in ocean systems engineering from the Korea Advanced Institute of Science and Technology, Daejeon, Korea. He is currently working toward the Ph.D. degree in naval architecture and marine engineering at the University of Michigan, Ann Arbor, MI, USA, working in the areas of X-ray CT and GPU computation.

Jinwhan Kim (S'05–M'10) received the B.S. and M.S. degrees in naval architecture and ocean engineering from the Seoul National University, Seoul, Korea. He received the M.S. and a Ph.D. degrees in aeronautics and astronautics from the Stanford University, Stanford, CA, USA.

He was with the Korea Research Institute of Ships and Ocean Engineering as a full-time Researcher. He was also with the Optimal Synthesis Inc. in California, as a Research Scientist. In 2010, he joined the Faculty of the Division of Ocean Systems Engineering at the Korea Advanced Institute of Science and Technology, Daejeon, Korea. He has been working in the areas of marine robotics and the guidance, control, and estimation of dynamical systems.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Wooyoung Choi received the B.S. and M.S. degrees in naval architecture and ocean engineering from the Seoul National University, Seoul, Korea, in 1980 and 1984, respectively, and the Ph.D. degree in engineering science from Caltech, Pasadena, CA, USA, in 1993.

He is currently a Professor in the Department of Mathematical Sciences at the New Jersey Institute of Technology, Newark, NJ, USA. His research interests include applied mathematics, fluid mechanics, and scientific computing with focusing on nonlinear wave phenomena in oceans.