Variational Estimation of Wave-Affected Parameters in a Two-Equation Turbulence Model

XUEFENG ZHANG, GUIJUN HAN, DONG LI, XINRONG WU, AND WEI LI

Key Laboratory of Marine Environmental Information Technology, National Marine Data and Information Service, State Oceanic Administration, Tianjin, China

PETER C. CHU

Naval Ocean Analysis and Prediction Laboratory, Department of Oceanography, Naval Postgraduate School, Monterey, California

ABSTRACT

A variational method is used to estimate wave-affected parameters in a two-equation turbulence model with assimilation of temperature data into an ocean boundary layer model. Enhancement of turbulent kinetic energy dissipation due to breaking waves is considered. The Mellor–Yamada level 2.5 turbulence closure scheme (MY2.5) with the two uncertain wave-affected parameters (wave energy factor α and Charnock coefficient β) is selected as the two-equation turbulence model for this study. Two types of experiments are conducted. First, within an identical synthetic experiment framework, the upper-layer temperature "observations" in summer generated by a "truth" model are assimilated into a biased simulation model to investigate if (α, β) can be successfully estimated using the variational method. Second, real temperature profiles from Ocean Weather Station *Papa* are assimilated into the biased simulation model to obtain the optimal wave-affected parameters. With the optimally estimated parameters, the upper-layer temperature can be well predicted. Furthermore, the horizontal distribution of the wave-affected parameters employed in a high-order turbulence closure scheme can be estimated optimally by using the four-dimensional variational method that assimilates the upper-layer available temperature data into an ocean general circulation model.

1. Introduction

Observations ([Kitaigorodskii and Lumley 1983](#); [Thorpe 1984](#); [Anis and Moum 1992](#); [Terray et al. 1996](#); [Drennan et al. 1996](#); [Babanin 2006](#); [Kantha et al. 2010](#)) show that the dissipation of turbulent kinetic energy (TKE) is enhanced greatly near the sea surface by surface gravity waves under nonbreaking (including nonbreaking wave turbulence and Langmuir turbulence) and breaking waves. The breaking-wave-induced mixing has been broadly implemented into ocean circulation and mixing models (e.g., [Mellor and Blumberg 2004](#)). On the basis of the observational evidence of the surface wave breaking ([Osborn et al. 1992](#); [Agrawal et al. 1992](#)),

[Terray et al. (1996)](#) suggested a three-layer structure: The first layer (from the surface) is a wave-enhanced layer with the depth on the same order as the significant wave height, and the energy dissipation rate proportional to z^{-3} (z denotes the vertical distance from the sea surface), which is twice faster than the classical wall-layer dissipation. The second layer is the transition layer below the breaking zone (depth about $6z_0$; z_0 is the surface roughness length) ([Craig and Banner 1994](#)), with the energy dissipation rate proportional to z^{-2}. The third layer is the classic wall layer with the energy dissipation rate proportional to depth z^{-1}.

To model the wave-breaking-enhanced turbulence near the sea surface layer, [Craig and Banner (1994)](#) and [Craig (1996)](#) imposed a surface diffusion boundary condition on the turbulent kinetic energy equation (CB boundary condition) in the Mellor–Yamada (MY) turbulence closure model (1982). [Burchard (2001b)](#) simulated a wave-enhanced layer under breaking surface waves with a two-equation turbulence model including

Corresponding author address: Dr. Xuefeng Zhang, Key Laboratory of Marine Environmental Information Technology, National Marine Data and Information Service, State Oceanic Administration, 93 Liuwei Rd., Hedong District, Tianjin 300171, China.
E-mail: xfz_nmdis@126.com

the CB boundary condition. Mellor and Blumberg (2004) developed a wave-enhanced parameterization scheme with the CB boundary condition to overcome a weakness of the MY turbulence closure model that produces a shallower surface boundary layer and higher surface temperature during summertime warming in comparison to the observations (Martin 1985). Zhang et al. (2011a,b) identified the effect of breaking surface waves on upper-ocean boundary layer deepening in the Yellow Sea in summer utilizing the Princeton Ocean Model generalized coordinate system (POMgcs; Ezer and Mellor 2004). A well-mixed temperature surface layer in the Yellow Sea can be reconstructed successfully when the breaking-wave-enhanced turbulent mixing is considered.

In addition to the wave breaking, other wave-related processes are also important in modulating the upper mixed layer, such as the nonbreaking wave (Babanin and Haus 2009) and the Langmuir turbulence (Belcher et al. 2012). Some studies indicate that the effect of wave breaking on the upper-level turbulence is significant within the depth comparable to the wave height (Terray et al. 1996). However, for a deeper mixed layer, the impact of wave breaking would be small and the effect of Langmuir circulation and nonbreaking wave becomes important (Babanin 2006).

Uncertain wave-affected parameters exist in modeling wave-induced turbulence (nonbreaking or breaking waves), such as the critical value of the wave Reynolds number Re_{cr} in nonbreaking waves and the wave energy factor α and the Charnock coefficient β in breaking waves. These parameters are usually determined empirically or adjusted artificially. Studies have shown successful parameter estimation with a dynamical model using variational optimal control techniques (Derber 1987; Le Dimet and Talagrand 1986). For example, Yu and O'Brien (1991, 1992) used the variational method to assimilate meteorological and oceanographic observations into a one-dimensional oceanic Ekman layer model, to estimate the drag coefficient and the oceanic eddy viscosity profile and to investigate the effect of initial condition on the variational parameter estimation. Zhang et al. (2003) showed the capability of four-dimensional variational data assimilation (4D-VAR) in estimating uncertain parameters in numerical models. Peng and Xie (2006) developed a tangent linear model and an adjoint model of three-dimensional POM to construct a 4D-VAR algorithm for coastal ocean prediction. Effective error correction was found in initial conditions and wind stress in the storm surge simulation (Peng et al. 2007), and the drag coefficient was estimated in the storm surge prediction using the adjoint model of the three-dimensional POM (Peng et al. 2013). Peng and

Xie (2006) also pointed out that it is still an open issue as to whether it is meaningful to linearize the turbulence closure scheme in an atmospheric or oceanic model due to the high nonlinearity and discontinuity of the vertical turbulence. The nonphysical noise might be produced, and thus lead to numerical instability during the process of linearizing the turbulence closure scheme. They applied a simple but efficient way of avoiding the noise problem through neglecting the variation of the vertical diffusion coefficients in the linearization of the vertical turbulence scheme.

Despite earlier studies on the parameter estimation and model verification (e.g., Chu et al. 2001), the adjoint model of the turbulence closure scheme has not yet been thoroughly investigated with either nonwave breaking or wave breaking. Determination of wave-affected parameters in the turbulent mixing due to breaking waves using the variation method is selected as the major objective of this study. First, the upper-layer temperature "observations" are produced by a "perfect" model. Second, a biased assimilation is conducted to identify the capability of the variational method to optimally estimate the wave-affected parameters in the MY2.5 turbulence closure scheme. Third, the real temperature profiles at Ocean Weather Station *Papa* (OWS *Papa*) are assimilated into the ocean model to obtain the optimal wave-affected parameters.

2. Ocean boundary layer model

a. Mean equations and second-moment closure

Let (x, y) be the horizontal coordinates, z the vertical coordinate, and t the time. Following D'Alessio et al. (1998), equations governing the mean flow, temperature, and salinity in a horizontally homogeneous ocean boundary layer are given by

$$\frac{\partial u}{\partial t} - fv = \frac{\partial}{\partial z}\left(K_M \frac{\partial u}{\partial z}\right),$$

$$\frac{\partial v}{\partial t} + fu = \frac{\partial}{\partial z}\left(K_M \frac{\partial v}{\partial z}\right),$$

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial z}\left(K_H \frac{\partial T}{\partial z}\right) - \frac{\partial R}{\partial z},$$

$$\frac{\partial S}{\partial t} = \frac{\partial}{\partial z}\left(K_H \frac{\partial S}{\partial z}\right), \qquad (1)$$

where u, v are the velocity components in the x, y directions, respectively; T is the potential temperature; S is the salinity; f is the Coriolis parameter; and K_M and K_H are the vertical mixing coefficients for momentum and tracers, respectively.

The MY2.5 turbulence closure scheme, widely used in ocean models such as POM and the Regional Ocean Modeling System (ROMS), is a two-equation turbulence model,

$$\frac{\partial q^2}{\partial t} = 2\left\{ K_M\left[\left(\frac{\partial u}{\partial z}\right)^2 + \left(\frac{\partial v}{\partial z}\right)^2 \right] + \frac{g}{\rho_0} K_H \frac{\partial \rho}{\partial z} - \frac{q^3}{B_1 l} \right\}$$
$$+ \frac{\partial}{\partial z}\left(K_q \frac{\partial q^2}{\partial z} \right), \quad \text{and}$$

(2)

$$\frac{\partial q^2 l}{\partial t} = l E_1\left\{ K_M\left[\left(\frac{\partial u}{\partial z}\right)^2 + \left(\frac{\partial v}{\partial z}\right)^2 \right] + \frac{g}{\rho_0} K_H \frac{\partial \rho}{\partial z} - \frac{q^3}{B_1 l} \frac{W}{E_1} \right\}$$
$$+ \frac{\partial}{\partial z}\left(K_q \frac{\partial q^2 l}{\partial z} \right),$$

(3)

where q^2 is the turbulent kinetic energy times 2; l is the turbulent macroscale; K_q is the vertical mixing coefficient for turbulence; and ρ and ρ_0 are the density and reference density, respectively,

$$W = 1 + E_2(l/\kappa L)^2, \quad L^{-1} = (\eta - z)^{-1} + (H + z)^{-1},$$

where κ (=0.41) is the von Kármán constant, H is the water depth, η is the free-surface elevation, and E_1, E_2, and B_1 are empirical constants. The turbulent energy and macroscale equations are closed by

$$K_M = lq S_M, \quad K_H = lq S_H, \quad K_q = lq S_q, \quad (4)$$

where S_M and S_H are the stability functions.

b. Wave-affected parameters

Wave-affected parameters are included in the surface boundary conditions of the two-equation turbulence model. The first one is the CB boundary condition for q^2 (Craig and Banner 1994),

$$K_q \frac{\partial q^2}{\partial z} = 2\alpha u_\tau^3, \quad z = 0, \quad (5)$$

where u_τ is the water-side friction velocity and α is "wave energy factor." The second one is for the turbulent macroscale l (Terray et al. 1996, 1999),

$$l = \max(\kappa z_w, l_z), \quad (6)$$

where l_z is the "conventional" empirical length scale, which is calculated prognostically by the MY2.5 turbulence closure scheme; and z_w is the wave-related surface roughness length, which denotes the relevant scale of turbulence.

In the absence of surface waves, both α and z_w at the surface are set as zero in the MY2.5 turbulent closure scheme (Blumberg and Mellor 1987). However, when the effect of surface waves is considered, both α and z_w appear as constants or vary with the states of surface waves. Craig and Banner (1994) set α as 100 for wave ages embracing very young wind seas to fully developed situations. Terray et al. (1996) indicates that $\alpha = 150$ is an adapted value under breaking waves. Kraus and Turner (1967), Denman and Miyake (1973), and Gaspar (1988) also choose different values of α in their studies.

Terray et al. (1996), Burchard (2001a), and Umlauf and Burchard (2003) suggest that z_w is the same order as the significant wave height (H_s). Further, Mellor and Blumberg (2004) summarized the work of Donelan (1990), Smith et al. (1992), and Janssen (2001), and obtained

$$z_w = (\beta \times 10^5)\frac{u_\tau^2}{g}, \quad (7)$$

where g is the gravitational acceleration, and β is the Charnock parameter (Chu and Cheng 2007), which varies from $\beta = 2$ (Stacey 1999) and $\beta = 0.32$ (Jones and Monismith 2008) to $\beta = 0.56$ (Carniel et al. 2009) to obtain the best performance in each numerical simulation. Mellor and Blumberg (2004) suggested that $\beta \sim O(1)$ is deemed correct under breaking waves. Stacey (1999) also indicated that $\beta \sim O(10)$ is too big a value to describe the surface boundary condition for the turbulent kinetic energy. It should be noted that β does not have to be so large if the other wave-induced mixing process is included in the models (Zhang et al. 2012).

c. Boundary conditions

The surface boundary conditions for q^2 and l are given by Eqs. (5) and (6). The bottom boundary conditions of q^2 and l are given by

$$q^2 = B_1^{2/3} u_{\tau b}^2, \quad \text{and} \quad (8)$$

$$l = \kappa z_0, \quad (9)$$

respectively, where $B_1 = 16.6$ (Blumberg and Mellor 1987) and $u_{\tau b}$ is the friction velocity associated with the bottom frictional stress. The surface and bottom boundary conditions of the mean flow and tracers are represented by

$$\left.\begin{aligned}\frac{\partial T}{\partial z} &= \frac{Q}{\rho_0 C_p} \\ S &= S_{\text{obs}} \\ K_M\left(\frac{\partial u}{\partial z}, \frac{\partial v}{\partial z}\right) &= \left(\frac{\tau_{wx}}{\rho_0}, \frac{\tau_{wy}}{\rho_0}\right), (\tau_{wx}, \tau_{wy}) = C_w \mathbf{u}_{10}(u_x, u_y) \\ C_w &= (0.75 + 0.067|\mathbf{u}_{10}|) \times 10^{-3}\end{aligned}\right\} \quad \text{at} \quad z = 0 \qquad (10)$$

and

$$\left.\begin{aligned}\frac{\partial T}{\partial z} &= 0 \\ \frac{\partial S}{\partial z} &= 0 \\ K_M\left(\frac{\partial u}{\partial z}, \frac{\partial v}{\partial z}\right) &= \left(\frac{\tau_{bx}}{\rho_0}, \frac{\tau_{by}}{\rho_0}\right), (\tau_{bx}, \tau_{by}) = C_d \mathbf{u}_b(u_{bx}, u_{by}) \\ C_d &= \max\left[\frac{\kappa^2}{\ln(z/z_0)^2}, 0.0025\right]\end{aligned}\right\} \quad \text{at} \quad z = -H(x), \qquad (11)$$

where Q is the surface net heat flux; C_p is the specific heat; S_{obs} is the observation of the sea surface salinity; τ_{wx} and τ_{wy} are the x and y components of the wind stress, respectively; $\mathbf{u}_{10}$ is the wind velocity at 10 m; u_x and u_y are x and y components of $\mathbf{u}_{10}$, respectively; τ_{bx} and τ_{by} are the x and y components of the bottom frictional stress, respectively; $\mathbf{u}_b$ is the bottom velocity; u_{bx} and u_{by} are the x and y components of $\mathbf{u}_b$, respectively; C_w and C_d are drag coefficients of the wind stress and the bottom stress, respectively; and z_0 is the bottom roughness parameter, taken as 0.01 m.

3. Variational approach

a. The variational theory within the least squares framework

The purpose of the variational analysis is to seek the optimal control variables by minimizing a well-defined cost function, in which a dynamical model including all the control variables is regarded as the strong constraints of the cost function. Within the least squares framework, a general form of the cost function can be defined as

$$J(p) = \frac{1}{2} \int_0^T \langle W(CX - X_{\text{obs}}), (CX - X_{\text{obs}}) \rangle, \quad (12)$$

where p is the vector of the control variables, X is the solution of the dynamical model

$$\frac{dX}{dt} = F(X),$$

and F is the differential operator. The symbol $\langle \ \rangle$ represents the inner product in the Euclidean space. Term W is the weight matrix. Term X_{obs} is the observation, and C is the projection operator from the model space to the observational space. Let

$$J(p^{\text{opt}}) = \min(p).$$

The optimal control variable p^{opt} is obtained from

$$\mathbf{\nabla} J(p^{\text{opt}}) = 0$$

with respect to all control variables. Here, $\mathbf{\nabla}$ is the gradient operator. The process for the variational analysis can be outlined as follows:

(i) Define a concrete cost function that reflects the misfit between the control variables and the available observations.
(ii) Calculate the value of the cost function $J(p)$ through integrating the dynamical model with a fixed time step.
(iii) Calculate the gradients of the cost function with respect to all control variables, $\mathbf{\nabla} J(p)$.
(iv) Minimize the cost function through a minimization algorithm according to the value of $J(p)$ and $\mathbf{\nabla} J(p)$.
(v) Estimate the optimal control variables p^{opt} according to the convergence criterion of the process of the minimization.

For executing the above-described process of the variational analysis (i)–(v), $\mathbf{\nabla} J(p)$ should be obtained in

advance, and in general, it is calculated by the adjoint model of the linearized dynamical model. To the first order the Taylor expansion of $J(p)$ is given by

$$J(p) = J(p_0) + \delta J(p), \qquad (13)$$

where $\delta J(p)$ is the variation of $J(p)$. On the one hand, $\delta J(p)$ is given by the definition of the variation:

$$\delta J(p) = \int_0^T \langle \nabla_X J(p_0), \delta X \rangle. \qquad (14)$$

On the other hand, $\delta J(p)$ can also be written according to Eq. (12),

$$\delta J(p) = \frac{1}{2} \int_0^T \left\langle W \frac{\partial C}{\partial X} \delta X, CX - X_{\mathrm{obs}} \right\rangle \\ + \frac{1}{2} \int_0^T \left\langle W(CX - X_{\mathrm{obs}}), \frac{\partial C}{\partial X} \delta X \right\rangle. \qquad (15)$$

With the symmetry of the inner product as well as a constant W matrix, Eq. (15) can be rewritten as

$$\delta J(p) = \int_0^T \left\langle W(CX - X_{\mathrm{obs}}), \frac{\partial C}{\partial X} \delta X \right\rangle. \qquad (16)$$

Let $\partial C / \partial X = A(X)$, thus Eq. (16) can be given as

$$\delta J(p) = \int_0^T \langle W(CX - X_{\mathrm{obs}}), A(X)\delta X \rangle, \qquad (17)$$

where $A(X)$ is the tangent linear operator. Equation (17) can be transposed according to the definition of the adjoint operator,

$$\delta J(p) = \int_0^T \langle WA^*(X)(CX - X_{\mathrm{obs}}), \delta X \rangle, \qquad (18)$$

where $A^*(X)$ is the adjoint operator of $A(X)$. Compared with Eq. (14), $\nabla_X J(p_0)$ can be described by

$$\nabla_X J(p_0) = WA^*(X)(CX - X_{\mathrm{obs}}). \qquad (19)$$

b. The adjoint model

According to Eq. (19), the gradient of the cost function with respect to the control variables can be

calculated using the adjoint model. The difference $CX - X_{\mathrm{obs}}$ is regarded as an external forcing of the adjoint model.

The general form of the adjoint model can be found in appendix A. The discretized adjoint model that computes the gradient of the cost function can be developed directly from the discretized dynamical model including Eqs. (1)–(11). In practical application, the source code of the adjoint model is constructed by combining the Tangent and Adjoint Model Compiler (TAMC) developed by Giering and Kaminski (1998) and a hand-coding correction. First, the adjoint code is generated by TAMC to avoid human errors and negligence, which are extremely easy to happen during the direct coding. Second, hand-coding correction is conducted to correct the AMC-generated code and to control the adjoint code structure. The errors in the adjoint code, which are induced from some irregular expressions of the forward numerical model, such as the partial array assignment and iterative use of intermediate arrays, are corrected through the hand coding. Finally, through the hand-coding correction, values of many intermediate results in the adjoint model are recorded into memory instead of recomputed to shorten the run time of the adjoint model, and some local variables and arrays are transferred to global attribute to improve the run efficiency of the adjoint model.

Once the cost function and its gradient are obtained from the dynamical model and its associated adjoint model, the minimization process is implemented to minimize the cost function through iterating the values of the control variables (T^n, T^{n-1}, α, and β) with the limited-memory Broyden–Fletcher–Glodfarb–Shanno (BFGS) quasi-Newton minimization algorithm (Liu and Nocedal 1989). During the minimization process, the maximum of α is set to 1000, and the maximum of β is set to 10 according to Mellor and Blumberg (2004) and Stacey (1999). The minima of the two wave-affected parameters are set to zero to keep realistic physical conditions. The minimization process is repeated until the convergence criterion of the gradient is reached. At that time, the optimal values of the control variables are obtained.

c. Cost function

In this study, the cost function is defined by

$$J(T^n, T^{n-1}, \alpha, \beta) = \frac{1}{2}(T^n - T_b^n)^{\mathrm{T}} B_1^{-1}(T^n - T_b^n) + \frac{1}{2}(T^{n-1} - T_b^{n-1})^{\mathrm{T}} B_2^{-1}(T^{n-1} - T_b^{n-1}) \\ + \frac{1}{2} \sum_{j=1}^M \sum_{i=1}^N [T_{j,i}(\alpha, \beta) - T_{\mathrm{obs}}]^{\mathrm{T}} R^{-1} [T_{j,i}(\alpha, \beta) - T_{\mathrm{obs}}], \qquad (20)$$

TABLE 1. All assimilation experiments and simulations within the identical synthetic experiment framework.

Name	Description	Control variables	Assimilation windows	Assimilation period	Assimilation depth
Truth model simulation	$\alpha = 200$ $\beta = 2$	—	—	—	—
Biased simulation	$\alpha = 100$ $\beta = 1$	—	—	—	—
PE	Parameter estimation	$T^n, T^{n-1}, \alpha, \beta$	1 day	1 day	30 m
PE_SST	Parameter estimation	$T^n, T^{n-1}, \alpha, \beta$	1 day	1 day	Sea surface
PE_β_TI	Parameter estimation with the perfect initial fields derived from the truth model simulation	β	1 day	1 day	30 m
PE_β_BI	Parameter estimation with the biased initial fields derived from the biased simulation	β	1 day	1 day	30 m

where the first two terms on the right side represent the background error terms that measure the misfit between the model's initial field and the background field. Terms T^n and T^{n-1} are the initial temperature values at the nth and $(n-1)$th time steps, respectively, which will be estimated optimally via the variational method. Terms T_b^n and T_b^{n-1} are the background temperature values at the nth and $(n-1)$th time steps, respectively, which can be derived from the model run. Both temperatures at the two consecutive time steps are considered as the control variables due to the utilization of the leapfrog time differencing scheme with the Asselin–Robert time filter (Robert 1966). Otherwise, initial shocks of the model states are likely to be produced during the variational estimation because of the inconsistence of the initial values at the two time steps. Terms B_1 and B_2 are the error covariance for T^n and T^{n-1}, respectively; for simplicity, both B_1 and B_2 use diagonal matrices, whose

values of the diagonal components are set to 10^{-4} in this study. The third term denotes the observation of the temperature at certain time intervals within the assimilation window, where $T_{j,i}$ and T_{obs} are the simulated and observed temperature at location i and time level j, respectively. Terms N and M are the number of grid points over the ocean and the number of time levels of observations, respectively. Term R is the error covariance for the observations, which also uses the same diagonal matrix as that of B_1.

Wave-affected parameters α and β are expressed implicitly in Eq. (20), which are regarded as the independent variables of $T_{j,i}$. Therefore, the value of the cost function can be obtained when the model integrates for n time steps with the known initial values of T^n, T^{n-1}, α, and β. The cost function has the following form if the wave-affected parameters α and β have background values (α_b, β_b, respectively):

$$J(T^n, T^{n-1}, \alpha, \beta) = \frac{1}{2}(T^n - T_b^n)^{\mathrm{T}} B_1^{-1}(T^n - T_b^n) + \frac{1}{2}(T^{n-1} - T_b^{n-1})^{\mathrm{T}} B_2^{-1}(T^{n-1} - T_b^{n-1})$$
$$+ \frac{1}{2}\sum_{j}^{M}\sum_{i}^{n}[T_{j,i}(\alpha,\beta) - T_{obs}]^{\mathrm{T}} R^{-1}[T_{j,i}(\alpha,\beta) - T_{obs}] + \frac{1}{2}K_\alpha(\alpha - \alpha_b)^2 + \frac{1}{2}K_\beta(\beta - \beta_b)^2, \quad (21)$$

where K_α and K_β are coefficients controlling the best fits for data. In this study, we use the first form of the cost function [Eq. (20)] for avoiding the complexity of the cost function.

4. Synthetic experiments

a. Truth model simulation

Table 1 lists all the assimilation experiments and model simulations within an identical synthetic experiment framework. The truth model consists of Eqs. (1)–(3) with $\alpha = 200$ and $\beta = 2$. All six equations from Eqs.

(1)–(3) are discretized using the same implicit method as POM. The maximum depth is set to 250 m, with 60 vertical levels. The first 20 vertical levels are 0.0, 0.5, 1.0, 1.5, 2.0, 4.0, 6.0, 8.0, 10.0, 12.0, 14.0, 16.0, 18.0, 20.0, 22.0, 24.0, 26.0, 28.0, 30.0, and 35.0 m. The time step is 1 h. The model initial state is from 1 January 1961, including temperature and salinity, derived from the real observation at OWS *Papa*. The model is forced by the observational 10-min momentum and heat fluxes acquired online from (http://www.pmel.noaa.gov/OCS/Papa).

Starting from the initial conditions (1 January 1961), the truth model is run for 6 yr to generate a time series of
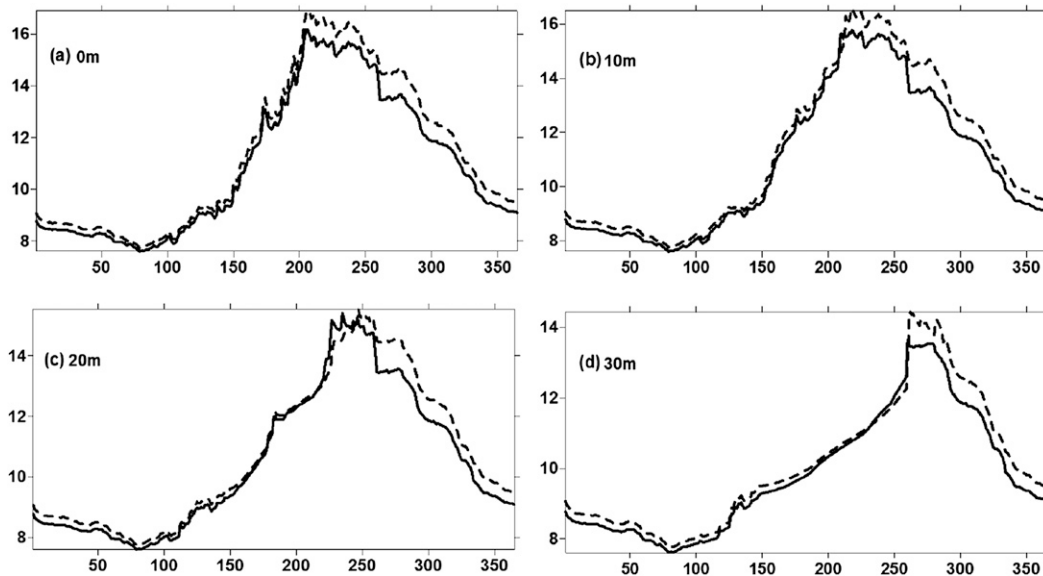
FIG. 1. Daily temperature in 1966 at (a) 0, (b) 10, (c) 20, and (d) 30 m at OWS *Papa* with the truth model simulation (solid curve) and the biased simulation (dashed curve).

the truth with the first 5 yr as the spinup period. The time of the observations of T is from 1 to 30 August 1966. The observations of T are produced through sampling the truth states at 1-h observational frequencies. The observation locations of T are consistent with those of the model vertical grids.

b. Biased simulation

The biased simulation uses the same truth model, but with different parameter settings. Therefore, the difference between the biased simulation and the truth model leads to the effect of the incorrect parameter settings. Figure 1 shows the simulated daily temperature at OWS *Papa* in 1966. The sea surface temperature (SST) from the biased simulation with $(\alpha, \beta) = (100, 1)$ is higher than that by the truth model simulation with $(\alpha, \beta) = (200, 2)$, and the maximum difference of the SST between the two simulations occurs in summer, namely, from the 200th day to the 240th day (solid line vs dashed line in Fig. 1a). An obvious difference of the temperature at 10-m depth in the two simulations also remains (Fig. 1b). The wave-affected parameters are half smaller in the biased simulation than in the truth model simulation, which suggests that the turbulent kinetic energy is too weak to mix the surface and subsurface water well in the biased simulation. After the 240th day (fall and winter), the temperature decreases gradually due to the convective mixing induced by the surface cooling. The temperatures at the surface and 10-m depth in the biased simulation remain higher than the counterpart in the truth model simulation due to the

insufficient wave-enhanced mixing in the biased simulation. Below 20 m, the effect of the wave-affected parameters on the temperature is not evident in summer (solid line vs dashed line in Figs. 1c and 1d), which indicates that the turbulent kinetic energy generated by the breaking surface gravity waves is dissipated only near the sea surface and does not penetrate into the deeper waters. The maximum difference in temperature at 30 m from the two simulations occurs in the fall (after the 250th day) with the temperature higher in the biased simulation than in the truth model simulation. Although the wave-affected parameters do not directly affect the temperature in the deeper layers in summer, it can affect the temperature indirectly by the SST due to the subsequent convective cooling in autumn and winter. Thus, the wave-affected parameters directly impact the temperature near the sea surface in summer, and indirectly impact the temperature in the deeper layers in autumn and winter.

We intend to investigate if the wave-affected parameters in a two-equation turbulence model can be estimated effectively through assimilating the temperature data into an ocean boundary layer model with the variational method. In addition, we want to understand how well the model state estimation/forecast can be improved through the estimated wave-affected parameters. In the next subsection, a series of synthetic experiments are carried out to address the issues.

c. Correctness test of the gradient

The code of the adjoint model is produced directly through the TAMC (of course, a hand-coding correction
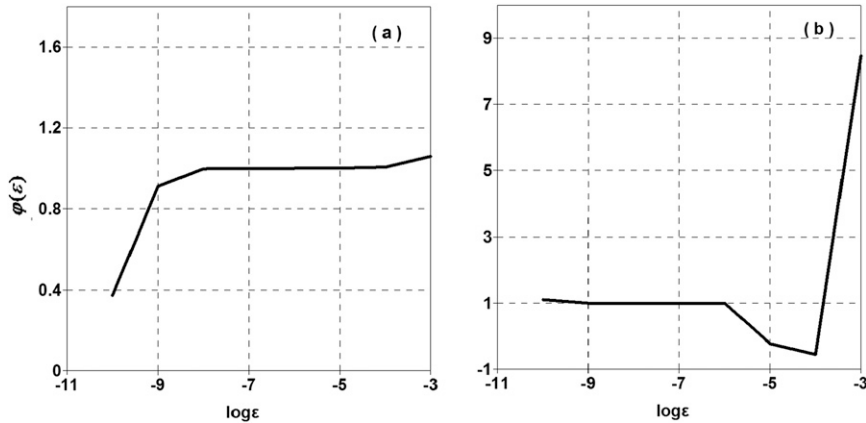
FIG. 2. The correctness test of the gradient with respect to (a) α and (b) β.

is necessary after that). According to the Taylor expression, one has

$$\lim_{\varepsilon \to 0} \varphi(\varepsilon) = \lim_{\varepsilon \to 0} \frac{J[x_0 - \varepsilon \nabla J(x_0)] - J(x_0)}{-\varepsilon \langle \nabla J(x_0), \nabla J(x_0) \rangle} \approx 1, \quad (22)$$

where x_0 is any control variable and the symbol $\langle \ \rangle$ represents the inner product. Figure 2 shows the correctness test of the gradient of the cost function with respect to α and β using Eq. (22). With respect to α, $\varphi(\varepsilon)$ converges to 1 as ε decreases from 10^{-3} to 10^{-8}, and decreases from 1 to 0.38 as ε decreases from 10^{-8} to 10^{-10}, which indicates the dominance of the computational errors in $\varphi(\varepsilon)$. With respect to β, $\varphi(\varepsilon)$ converges to 1 as ε decreases from 10^{-6}. Therefore, the adjoint coding is valid.

d. Parameter estimation

Figure 3 shows the time series of α and β during the parameter estimation (PE) described in Table 1, where both the assimilation window and assimilation period are set to 24 h and the assimilation depth is set to 30 m. Therefore, the processes (ii)–(vii) described by appendix C are executed 30 times to obtain time series of α and β from 1 to 30 August 1966. Figure 3b shows that β converges to its truth value (dashed line) after 9 days, while α converges to its truth value (dashed line in Fig. 3a) after about 15 days. Results show the wave-affected parameters in the high-order turbulent model can be estimated successfully using the upper-layer temperature observations through the variational control technique. For each cycle of the parameter estimation in the 30 days, the process of the minimization is iterated until the convergence criterion of the gradient is satisfied. Figure 4 shows the dependence of the cost function and the norm of the gradient on the number of iterations on 2 August 1966. The value of the cost function decreases rapidly from 4.3 to 0.8 within the first

five iterations, and it keeps the low value (0.8) steadily after the fifth iteration (Fig. 4a). However, the norm of the gradient oscillates dramatically to search the optimal declining direction of the gradients. The norm of the gradient becomes stable after the 130th iteration (Fig. 4b). The minimization process stops after 180 iterations, indicating the local minima of the wave-affected parameters for that day.

Figure 5 depicts the temporal variations of the natural logarithm of the cost function at OWS *Papa* from
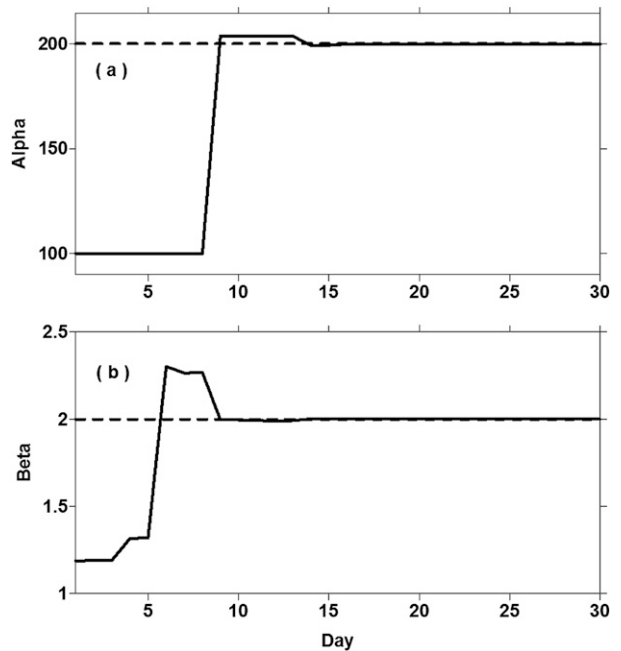

FIG. 3. Time series of the estimated wave-effected parameters (a) α and (b) β for PE from 1 to 30 Aug 1966 (solid curve), where both the assimilation window and the assimilation period are 1 day and the depth of the assimilation is 30 m. Here, the dashes curves show the truth (α, β) values.
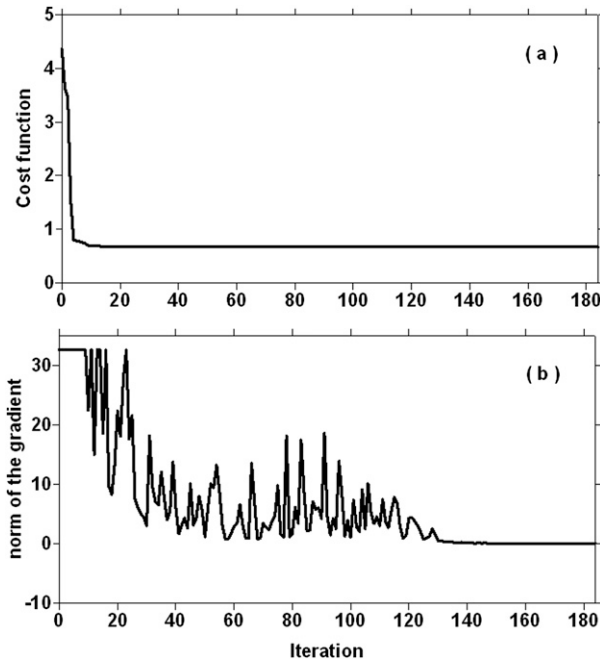
FIG. 4. Dependence of (a) the cost function and (b) the norm of the gradient on the number of iterations on 2 Aug 1966.

1 to 30 August 1966. The cost function (red line) decreases dramatically in the first 5 days and then decreases gently in the following 25 days. Both the background term (blue line in Fig. 5) and the observation term (black line in Fig. 5) of the cost function have a similar pattern with the total cost function. The two terms almost
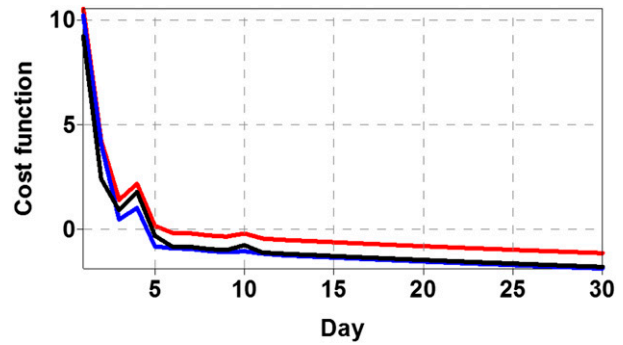

FIG. 5. Temporal variations of the natural logarithm of the cost function at OWS *Papa* from 1 to 30 Aug 1966. Here, the red, blue, and black curves are the total, background, and observation terms of the cost function.

converge to the same value after the 10th day, indicating the estimated initial temperature fields reach a balance between the background temperature and the observation.

The temporally varying wave-effected parameters (α, β) estimated from their different initial values on 1 August 1966 (Fig. 6) converge to their truth values within one month through the parameter optimization with the variational approach. It clearly shows that the variational assimilation approach is feasible for the wave-affected parameter optimization with different initial parameter values.

To evaluate the effect of the noise in the temperature observation on the wave-affected parameter estimation,
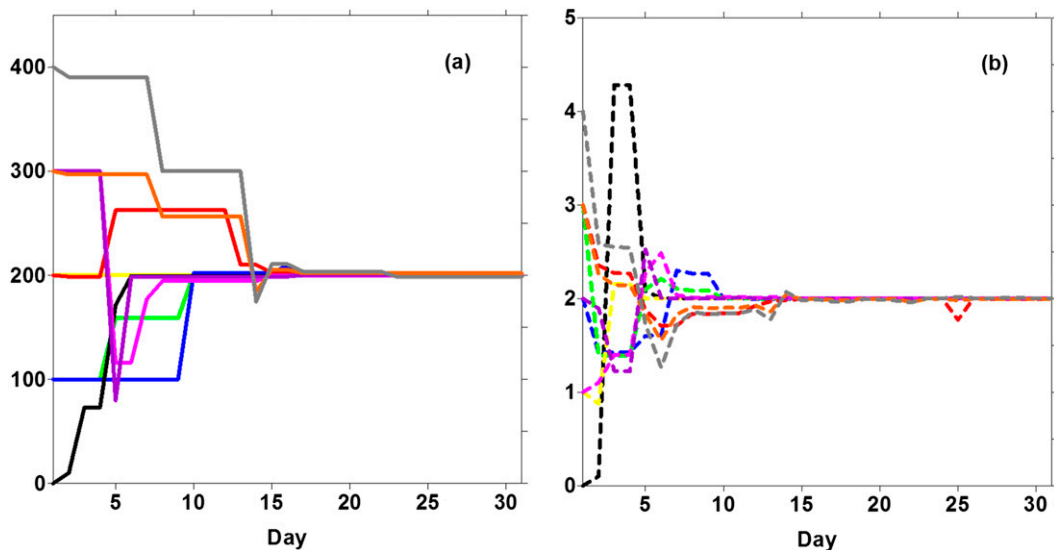

FIG. 6. Time series of the estimated wave-effected parameters (a) α and (b) β for different initial parameter values from 1 to 30 Aug 1966, where both the assimilation window and the assimilation period are 1 day and the depth of the assimilation is 30 m. Black, blue, green, yellow, red, pink, purple, orange, and gray solid lines in (a) and the corresponding dashed lines in (b) show values of $(\alpha, \beta) = (0, 0)$, $(100, 2)$, $(100, 3)$, $(200, 1)$, $(200, 3)$, $(300, 1)$, $(300, 2)$, $(300, 3)$, and $(400, 4)$ respectively.

TABLE 2. Dependence of the optimally estimated (α, β) on the standard deviation of temperature observation.

Std dev of temperature observation	Estimated value of α	Estimated value of β	Relative error of α (%)	Relative error of β (%)
10^{-3}	206.125	1.982	96.9	99.1
10^{-2}	167.343	2.098	83.6	95.1
0.05	120.033	2.114	60.0	94.3
0.1	100.096	1.889	50.0	94.5
0.5	100.068	0.866	50.0	43.3

based on the PE experiment, the white noises with different standard deviation are added to the temperature observation. Table 2 shows the dependence of the optimally estimated (α, β) on the error standard deviation of the temperature observation. The relative error of optimally estimated α decreases from 96.9% to 60%, and the relative error of the optimally estimated β decreases from 99.1% to 94.3% as the error standard deviation in the temperature observation increases from 0.001 to 0.05 K. It implies that the effect of observational noise on the estimation is more severe on α than on β, which means that it is more difficult to pick up the positive signal when the noise dominates the cost function and the corresponding gradients during the parameter estimation of α. When the standard deviation of the temperature observation increases to 0.5 K, both relative errors of the optimally estimated α and β are below 50%, which indicates that the level of the noise is not acceptable for assimilation purposes.

To determine if the wave-affected parameters can be estimated correctly using only the SST data, the second assimilation experiment, PE_SST, is conducted, in which only the SST observations are assimilated into the biased simulation model. Neither α (Fig. 7a) nor β (Fig. 7b) reaches their truth values (dashed curve) due to the poor constraint of the observation. When only the SST observations are assimilated, the subsurface temperature cannot be estimated accurately. Under this condition, the two parameters will be adjusted to the optimal values to fit the inaccurate temperature values to the greatest extent within a fixed time window, rather than converge to truth values. Therefore, the subsurface temperature observations are essential for estimating α and β reasonably well.

The dependence of the optimally estimated α (Fig. 8a) and β (Fig. 8b) on the assimilation window and assimilation period is investigated using different values from 1 to 30 August 1966 (Fig. 8). When the assimilation window and the assimilation period are 48 and 72 h, respectively, both parameters converge to their respective truth values (see black and blue lines in Figs. 8a and 8b). However, when the assimilation window and the assimilation period reach 96 and 120 h, respectively,

neither α nor β converges to their truth values within one month, which can be seen from the red and pink lines in Figs. 8a and 8b. It clearly shows that the parameter updating with the observation can improve the state estimation of the next cycle, and that the improved state estimation further enhances the quality of parameter estimation for the next cycle of parameter correction. When the assimilation window and the assimilation period are set to 120 h, the state-parameter optimization is performed only in six cycles within one month. Although the cost function decreases gradually, which can be seen from the dashed curve in Fig. 9, the control variables (the initial temperature T and the two parameters α, β) are not estimated reasonably well. In contrast, when the assimilation window and the assimilation period are set to 24 h, just as in the PE experiment, the state-parameter optimization can be performed in 30 cycles within one month, and the cost function can reach quasi equilibrium after 10 days (solid curve in Fig. 9).
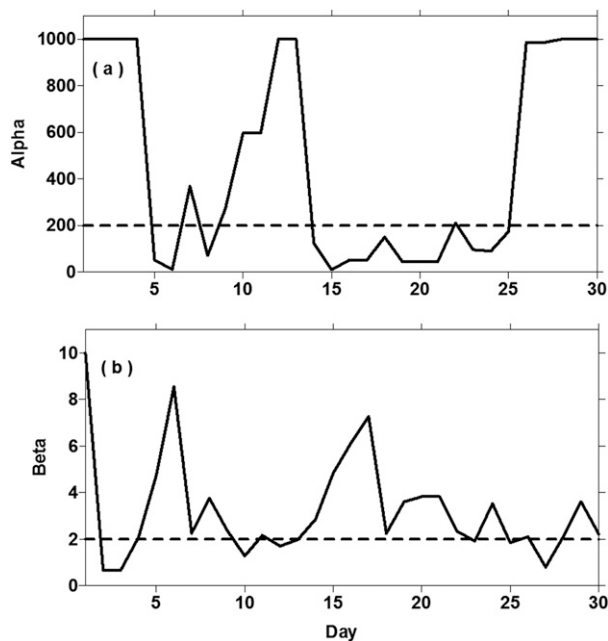

FIG. 7. As in Fig. 3, but for PE_SST, where only the SST observations are assimilated.
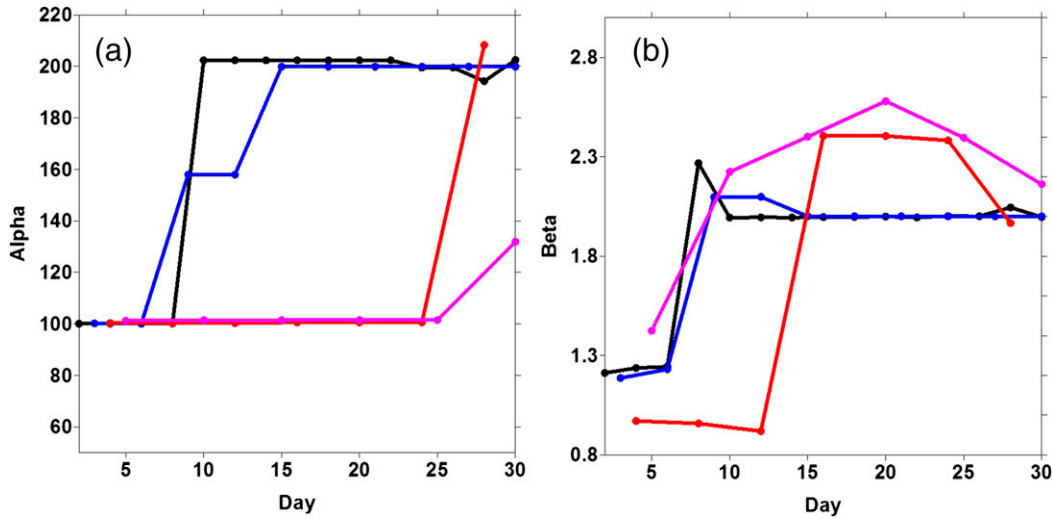
FIG. 8. Time series of the estimated wave-effected parameters (a) α and (b) β for a different assimilation window and assimilation period from 1 to 30 Aug 1966, where the depth of the assimilation is 30 m. Black, blue, red, and pink solid lines in (a) and (b) show the assimilation period and are 48, 72, 96, and 120 h.

It should be noted that the initial temperature values are regarded as not only the control variables being estimated but also the background temperature values of the cost function [see Eq. (20)] in the current minimization cycle. The incorrect convergence of (α, β) suggests that the initial temperature field (the background temperature values) in the current cycle is not adjusted well enough, which is regarded as the source of noise during parameter estimation using the variational method. Therefore, it is hard to obtain the accurate values of (α, β) before the state variables (T^n and T^{n-1}) attain the adequate accuracy. To better understand the issue, two other experiments are carried out, in which β is regarded as the only control variable. The experiment PE_β_TI described in Table 1 uses the "perfect" initial field that is generated by the truth model with the truth values of α and β; the other experiment, PE_β_BI, uses the "biased" initial field that is generated by the biased simulation with the biased values of α and β. Table 3 shows the evolution of the cost function, the norm of the projected gradient, and the value of β with respect to the number of iterations in PE_β_TI. The parameter β reaches its truth value at the third iteration. The convergence criterion of the gradient is satisfied at the fourth iteration. However, β estimated from PE_β_BI cannot converge to its truth value (Table 4). After the convergence criterion of the gradient is satisfied at the sixth iteration, β reaches 3.3023 35, which is different from the truth value of 2.0. Although β from PE_β_BI cannot converge to its truth value, it reaches its optimal value to compensate for the error derived from the biased initial field during minimizing the model-observation misfit.

In fact, in a 3D ocean circulation model, model biases arise from the imperfect dynamical core and empirical physical schemes even if the initial field is perfect. With a biased initial field alone, one expects that the parameter optimization can compensate for both the numerical and physical deficiencies of the numerical model and enhance the performance of the model simulation to a certain degree. In this situation, parameters can only converge to their optimal value, instead of the truth values. In the next section, real temperature profiles from OWS *Papa* will be assimilated into the assimilation model to obtain the optimal wave-affected parameters (α, β).

5. Real experiment

OWS *Papa* is located in the North Pacific at 50°N, 145°W, where the currents are relatively weak and the
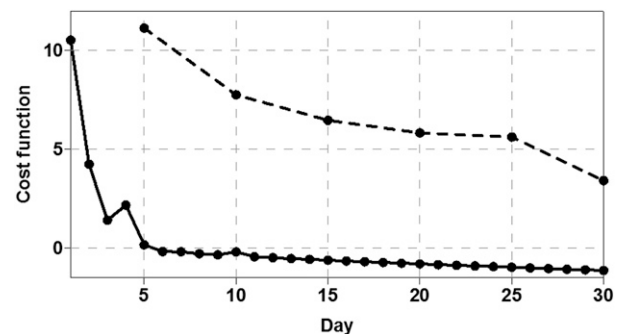

FIG. 9. Temporal variations of the natural logarithm of the cost function at OWS *Papa* from 1 to 30 Aug 1966. The solid and dashed curves represent PE and PE_5d, respectively, with black dots denoting the time that observations are assimilated.

TABLE 3. Evolution of the cost function, the norm of the projected gradient, and the value of β with respect to the number of iterations for the direct perturbed method with the perfect initial field.

Iteration step	Cost function	Norm of the projected gradient	Value of β
0	5.881	2.097	1.0
1	1.354×10^{-5}	7.406×10^{-2}	2.000 365
2	2.631×10^{-9}	1.032×10^{-3}	2.000 005
3	7.441×10^{-17}	3.042×10^{-7}	2.000 000
4	5.056×10^{-17}	5.693×10^{-9}	1.999 999

TABLE 4. As in Table 3, but with the biased initial field.

Iteration step	Cost function	Norm of the projected gradient	Value of β
0	2.319×10^{-2}	4.819	1.0
1	1.072×10^{-2}	3.351	3.350 811
2	1.071×10^{-2}	1.234	3.317 216
3	1.071×10^{-2}	9.003×10^{-3}	3.301 275
4	1.071×10^{-2}	1.982×10^{-3}	3.302 359
5	1.071×10^{-2}	6.043×10^{-6}	3.302 335
6	1.071×10^{-2}	3.627×10^{-6}	3.302 335

local mixing modulates mainly the dynamical process in the upper ocean in summer. The observed temperature profiles from 1 to 31 August 1966 at the site have a 3-h interval and a coarser vertical resolution (5 m) than the model grid points. There are seven observational layers totally in the upper 30 m, namely, 0, 5, 10, 15, 20, 25, and 30 m. Linear interpolation is used to fill the spatial gap between the modeled data and the observational data.

Table 5 lists all the assimilation experiments and model simulations within the real experiment framework. First, a control run without assimilating any observational data is called control (CTRL) to serve as the reference for the evaluation of assimilation experiments. The initial temperature and salinity are taken from those at 0000 UTC 1 January 1961 and linearly interpolated to model grids. The high-resolution (10 min) surface-observed data (momentum and net heat fluxes) at the site are used to force the model. Figure 10a shows the daily observed (red curve) and simulated sea surface temperature from CTRL (black dashed curve) at OWS *Papa* on August 1966. The simulated SST is higher than the observed SST by about 3°C (black dashed curve vs red curve). At the same time, the simulated mixed layer depth from CTRL is shallower than the observation by more than 10 m (black dashed curve vs red curve in Fig. 10b). The optimal values of (α, β) are estimated with the variational method to mitigate the bias between the model and the observation using the real summer temperature data.

The real parameter estimation (RPE) is described in the second row of Table 5. The initial field is generated from the results on 1 August 1966 simulated by the truth model in the above-mentioned synthetic experiments. The initial values of (α, β) are also consistent with those in the truth model simulation. The length of both the assimilation window and the assimilation period are set to 3 days (8 real observational temperature profiles in each day, for a total of 24 profiles for 3 days) and the assimilation depth is 30 m. The process of PE is similar to the process described in section 3, but with the real temperature observations at OWS *Papa* in August 1966. Table 6 shows the evolution of the cost function, α, and β with respect to the number of iterations for RPE. After the eighth iteration, the normalized cost function decreases to 5% of its initial value. The optimal values of α and β reach 107.48 and 3.98, respectively. The SST from RPE has a significant improvement compared to the simulated SST from CTRL (black solid curve vs black dashed curve in Fig. 10a), whose values are basically consistent with those of the observations (black solid curve vs red curve in Fig. 10a). The mixed layer depth is also more accurate from RPE than from CTRL (Fig. 10b). However, some discrepancy in the mixed layer depth still exists between RPE and the observation. This is because too many factors modulate the complicated thermodynamic processes of the upper mixed layer besides the surface gravity waves, such as horizontal advection, internal waves, upwelling, and entrainment. Many physical processes are not enclosed

TABLE 5. All the assimilation experiments and model simulations within the real experiment framework.

| Name	Description	Control variables	Assimilation windows	Assimilation period	Assimilation depth	Initial fields
CTRL	Simulation with $\alpha = 200$ $\beta = 2$	—	—	—	—	1 Aug 1966 from the truth model simulation
RPE	Real parameter estimation	T^n, T^{n-1}, α, β	3 days	3 days	30 m	Same as CTRL
RSE_Po	Simulation using the parameters estimated by RPE	—	—	—	—	31 Aug 1966, derived from RPE
RSE_Pd	Simulation using the same parameters as in CTRL	—	—	—	—	Same as RSE_Po
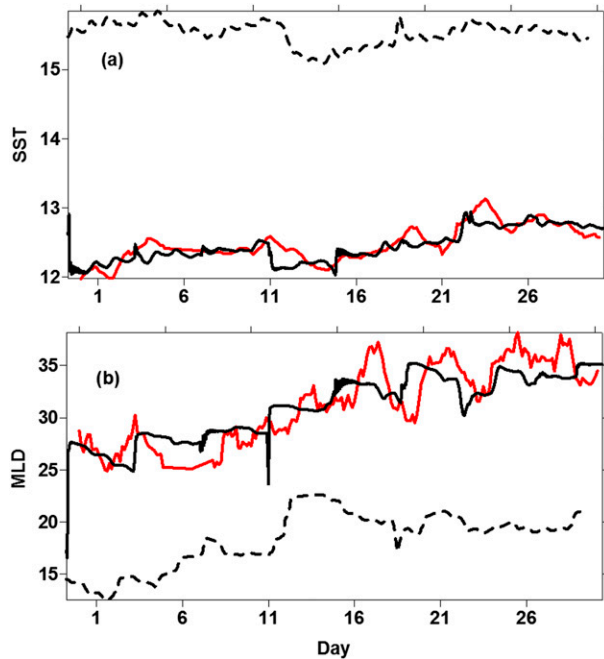
FIG. 10. (a) SST and (b) mixed layer depth from CTRL (black dashed curve) and RPE (black solid curve), and observations (red solid curve) at OWS *Papa* from 1 to 30 Aug 1966. Horizontal axis represents the day relative to 1 Aug 1966.

TABLE 6. Evolution of the cost function (α, β) with respect to the number of iterations for the real assimilation.

Iteration step	Normalized cost function	Value of α	Value of β
1	1.0	107.12	4.40
2	0.52	107.45	4.44
3	0.26	107.61	4.37
4	0.29	107.97	4.23
5	0.26	107.39	3.79
6	0.14	107.48	3.86
7	0.17	107.52	3.96
8	0.05	107.54	3.98

vertically homogeneous. It indicates that the model performance can be effectively improved using the optimal parameters. However, more accurate model simulations are needed using the optimal values of parameters via the variational methods repeated at certain time intervals with more available observations.

6. Discussion and conclusions

Wave-affected parameters in high-order turbulence closure schemes can modulate distinctly the vertical structure in the upper ocean. For improving the performance of the model in simulating the upper-ocean mixed layer, it is essential to estimate the optimal values of the wave-affected parameters using available observations deployed in the upper ocean through some robust data assimilation methods. It is known that one of the advantages of the variational method is that it can seek a posterior maximum likelihood solution of the model parameters in terms of the best fitting of the modeling trajectory to the observations by minimizing a cost function that measures the distance between observations and model states within an appropriate minimization time window. Therefore, in this study, the

in the simple ocean boundary layer model. The optimal values of the parameters can only compensate for some model bias but not all. However, the result from RPE indicates that the variational estimation of wave-affected parameters can indeed reduce model biases and improve the model capability in the upper ocean.

To explore the impact of parameter estimation on model simulation, two validation experiments, RSE_Po and RSE_Pd, are conducted. The optimal parameters estimated from RPE are used in RSE_Po, and the default values of the parameters from CTRL are used in RSE_Pd. In addition, both experiments use the same initial fields on 31 August 1966, which are derived from RPE. Figure 11 shows the observed (red curve) and simulated SST from RSE_Po (black solid curve) and RSE_Pd (black dashed curve) at OWS *Papa* from 1 to 30 September 1966. The simulated SST is more consistent with the observations from RSE_Po than from RSE_Pd. The simulated twice-monthly averaged turbulent kinetic energy q^2 (Fig. 12a) and the vertical mixing coefficient for temperature K_H (Fig. 12b) at OWS *Papa* in September 1966 are much larger for all the depths in RSE_Po (solid curve) than in RSE_Pd (dashed curve). The enhanced K_H in the upper-30-m depth in RSE_Po, due to the improvement of the turbulent kinetic energy calculation, mixes the momentum from the winds downward through the water column and makes it more
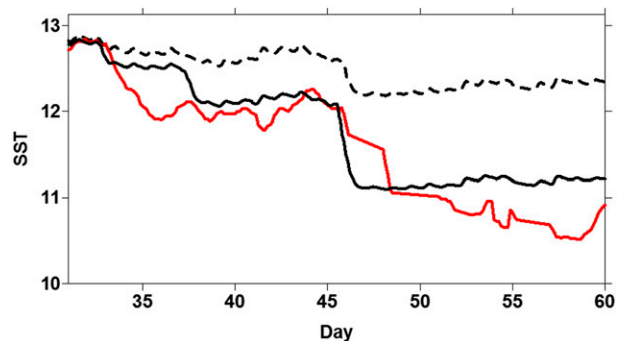

FIG. 11. SST from observations (red solid curve), RSE_Po (black solid curve), and RSE_Pd (black dashed curve) at OWS *Papa* from 31 Aug to 30 Sep 1966. Horizontal axis represents the day relative to 1 Aug 1966.
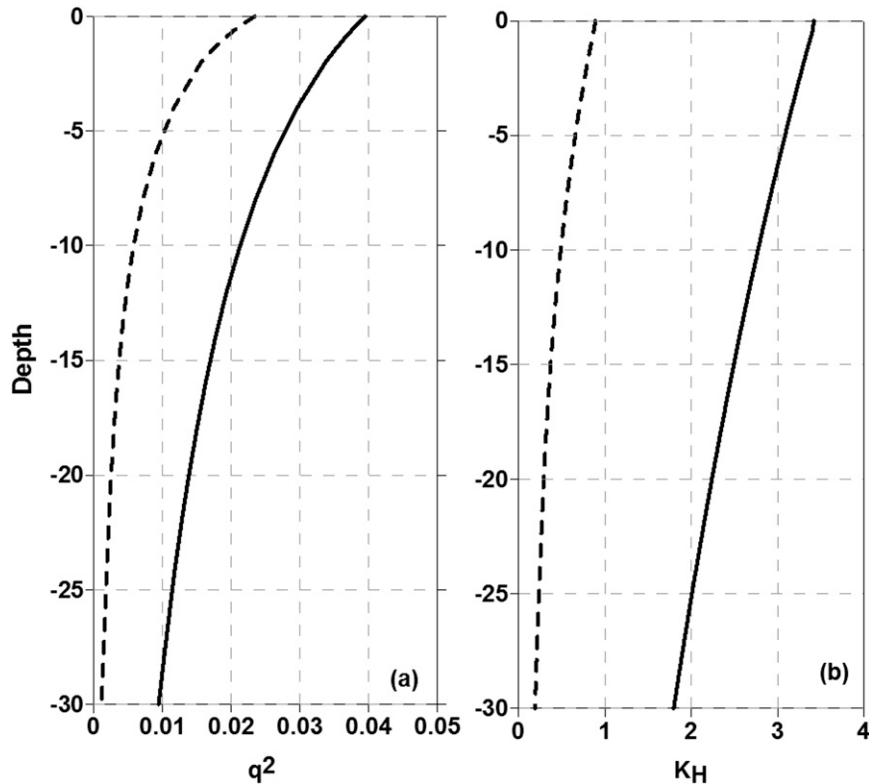
Fig. 12. Vertical profiles of the simulated monthly averaged (a) 2 times q^2 (m^2 s^{-2}) and (b) the vertical mixing coefficient for temperature K_H (10^{-3} m^2 s^{-1}) from RSE_Po (solid curve) and RSE_Pd (dashed curve) at OWS *Papa* in September 1966.

upper-layer temperature data are assimilated into an ocean surface boundary layer model to tentatively estimate the wave-affected parameters (α, β) employed in the MY2.5 two-equation turbulence model using the variational method. Within an identical synthetic experiment framework, the "truth" values of the wave-affected parameters in the high-order turbulence model can be retrieved successfully when the assimilation window, the assimilation period, and the assimilation depth are set appropriately. The observational temperature profiles at OWS *Papa* are also assimilated to correct the model bias arisen from multiple sources. By fitting the model results to the observations using the variational method, the optimal temperature field can be obtained in the upper 30 m through adjusting the wave-affected parameters to their optimal values.

Wave-affected parameter estimation using the variational method can compensate in part for the numerical and physical deficiencies of the model in the upper ocean. However, because of the existence of the deficiencies, the optimal values of the wave-affected parameters from the variational estimation might be not close to so-called truth values, even far away from the truth ones, which may induce that the upper-ocean

turbulent mixing is overestimated or underestimated. The optimal values of the wave-affected parameters in real applications are only applicable to the specific time period, location, and model. Further, the optimal values should vary temporally and spatially rather than being constants, which can be obtained by using the variational methods repeatedly at certain time intervals and the available observations (Peng et al. 2013). Although the optimal values of the wave-affected parameters are both model dependent (initial fields, time window of assimilation, model configuration, etc.) and observation dependent (sampling frequency, sampling errors, etc), as is indicated by this study, they can indeed mitigate the model biases from multiple sources, and obviously improve the performance of the model simulation.

Besides the wave-breaking parameters, other parameters in the wave-related processes can also be introduced into the model (which is compatible with those pertinent to the wave breaking) to estimate their optimal values. For instance, it is well known that Langmuir turbulence plays a key role in modulating the upper-layer mixing in the open sea. Recently, Harcourt (2013) introduces two more parameters into a second-moment closure turbulent model to describe the effect of the

Langmuir turbulence. One expects that the parameters associated with the Langmuir turbulence can be estimated together with the wave-breaking parameters using the variational method. Further, satellite remote-sensed SST data and in situ temperature data (such as the Argo floats) can provide a mass of temperature observations in upper oceans. Therefore, the optimal geographic-dependent distribution of the wave-affected parameters in a high-order turbulence closure scheme can be obtained using the 4DVar that assimilates the upper-layer available temperature data into ocean circulation models.

APPENDIX A

General Form of the Adjoint Model

The dynamical model composed of Eqs. (1)–(11) can be summarized in a general form as

$$\frac{\partial x}{\partial t} = F(x),$$
$$x|_{t_0} = x_0,$$
$$x(t)|_\Gamma = y(t), \tag{A1}$$

where x is the vector of model state variables, including u, v, T, S, q^2 and q^2l; x_0 is the model states at initial time t_0, and $y(t)$ is the boundary condition on Γ.

The tangent linear model of Eq. (A1) can be written as

$$\frac{\partial x'}{\partial t} = \frac{\partial F(x)}{\partial x} x',$$
$$x'|_{t_0} = x'_0,$$
$$x'(t)|_\Gamma = y'(t), \tag{A2}$$

where the prime is the perturbations of the state variables.

For the two vectors $\mathbf{w}$ and $\mathbf{z}$ in the Euclidean space, the adjoint operator L^* of the linear operator L can be defined as

$$\langle \mathbf{z}, L\mathbf{w} \rangle = \langle L^*\mathbf{z}, \mathbf{w} \rangle.$$

In the Euclidean space, L^* is the transpose of L, namely, $L^* = L^T$. The adjoint model corresponding to (A1) is given by

$$\frac{\partial \tilde{x}}{\partial t} = -\left[\frac{\partial F(x)}{\partial x}\right]^T \tilde{x},$$
$$\tilde{x}|_{t_E} = 0,$$
$$\tilde{x}(t)|_\Gamma = 0, \tag{A3}$$

where $\tilde{x}$ represents the adjoint variables and t_E is the end time in the temporal integration of Eq. (A1). The negative sign on the right side of the first equation in (A3) indicates that the adjoint model integrates backward in time. When the adjoint model integrates backward to the initial time t_0, the corresponding $\tilde{x}|_{t_0}$ is the gradient of the cost function with respect to the state variables (note that the difference in the state variables and observations should be regarded as the external forcing of the adjoint model in the practical applications).

APPENDIX B

Sensitivity of Simulated Temperature to Parameters

It is essential to investigate model sensitivities with respect to parameters being estimated before parameter estimation. Figure B1 shows the dependence of the cost function on α and β. It increases with increasing α and β in general. However, the local minimum of the cost function can be found near the region in which both α and β reach their default values (see Fig. B1b). The existence of the local minimum indicates that it is likely to estimate the optimal values of α and β if the values of the gradient with respect to the parameters can be calculated correctly in all the numerical iterations by the adjoint model.

The ensemble spread of T is used to evaluate the relevant sensitivities quantitatively. For α and β, 100 Gaussian random numbers are generated with the standard deviation being 5% of the default value and superimposed into the parameter being perturbed, while the other parameter remains unperturbed. All 100 ensemble members are started from the same initial conditions (1 January 1961). The biased simulation model is integrated up to 6 years. Sensitivities are calculated with the model output from 1 to 31 August 1966. This process is looped for the two wave-affected parameters. Fig. B2 shows the ensemble spread of T with respect to α and β at different depths. The ensemble spread of T near the sea surface is more than 0.09 with respect to β and less than 0.02 with respect to α. The sensitivity of T is obviously larger to β than to α for the whole depth, especially in the upper 30 m. Small sensitivity in the lower layer indicates that the noise may be stronger than the signal during the parameter estimation when the lower-layer temperature observations are assimilated into the bias simulation model.
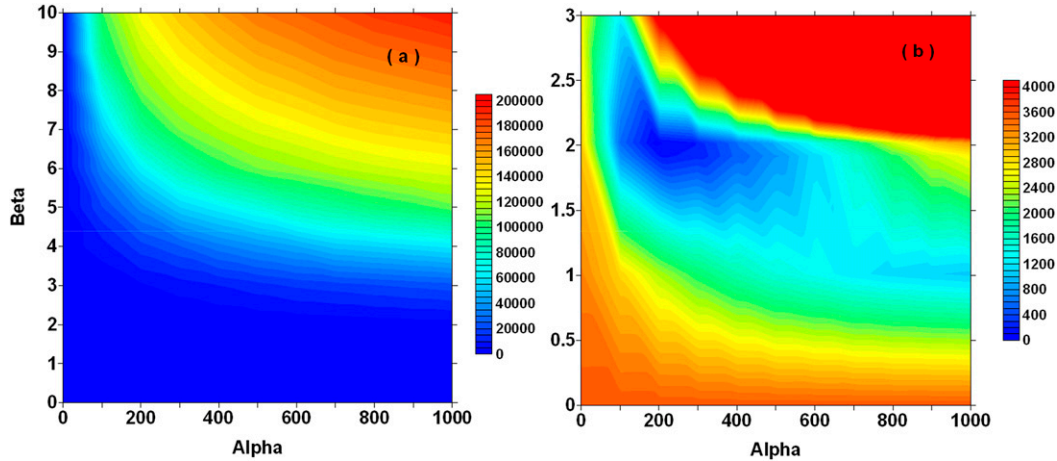
FIG. B1. Dependence of the cost function on α and β for (a) $10 \geq \beta \geq 0$ and (b) $3 \geq \beta \geq 0$.

The sensitivities with respect to the wave-affected parameters are also investigated through calculating the gradients of the cost function with the parameters, namely, $\partial J/\partial \alpha$ and $\partial J/\partial \beta$. Table B1 shows the dependence of the sensitivity on the initial values of the parameters α and β. When the initial parameter values (α, β) are set exactly to the truth values (200, 2), both sensitivities are very close to zero. In general, the sensitivity is several orders of magnitude greater on β than on α. It indicates that the parameter α is more vulnerable to being disturbed by the noises arisen from the observational errors and the biased initial state fields during the parameter estimation.

APPENDIX C

Process of the Wave-Affected Parameter Estimation

Figure C1 shows a flowchart of the wave-affected parameter estimation with the variational method. The

process for the wave-affected parameter estimation is outlined as follows:

(i) Begin with the initial field on 1 August 1966 and use the different values of wave-affected parameters from the truth for the biased simulation.

(ii) Integrate the model Eqs. (1)–(3) forward to a fixed time window ΔT_w and calculate the value of the cost function $J(T^n, T^{n-1}, \alpha, \beta)$ using Eq. (20).

(iii) Integrate the adjoint model backward in time and calculate the values of the gradient of the cost function with respect to the control variables ∇J.

(iv) With the values of the cost function $J(T^n, T^{n-1}, \alpha, \beta)$ and the gradient ∇J, use the BFGS algorithm to obtain the new values of the control variables, namely, the two wave-affected parameters α, β and the initial upper-layer temperature fields T^n, T^{n-1}.

(v) With the updated control variables from process (iv), repeat processes (ii)–(iv) until the convergence criterion for the minimization is satisfied. The convergence criterion is defined as
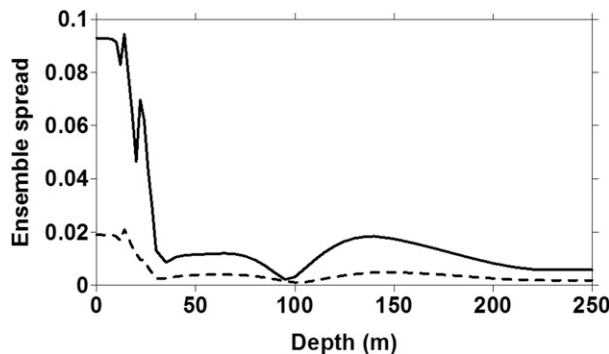

FIG. B2. Ensemble spread of temperature with respect to the wave-effected parameters α (dashed curve) and β (solid curve) at different depths.

TABLE B1. Dependence of the sensitivity on the initial values of the parameters α and β.

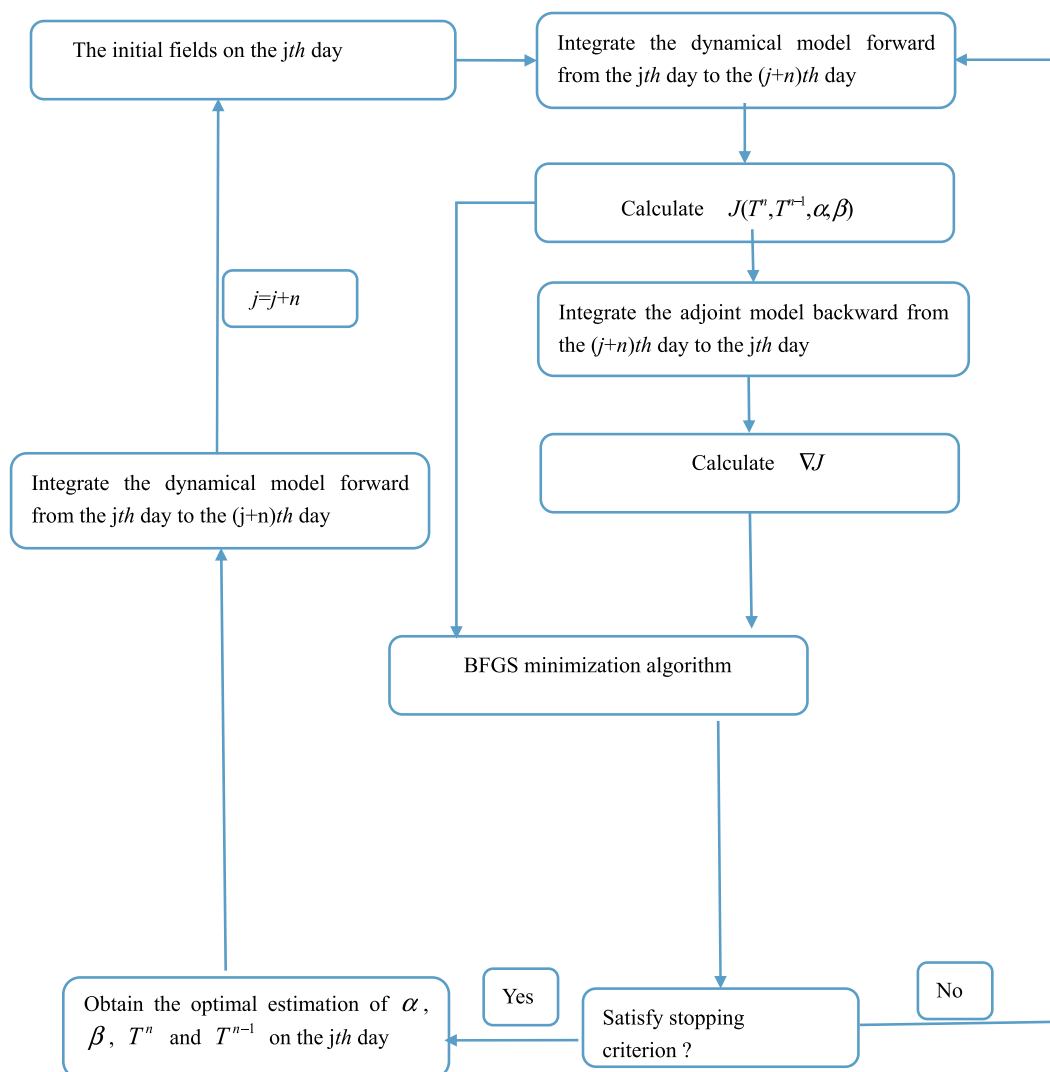| Initial values of (α, β)	Sensitivity of α	Sensitivity of β
(0, 0)	-7.5×10^{-5}	2.6×10^4
(100, 1)	-4.69	-574.96
(100, 2)	-42.63	-5325.69
(100, 3)	158.70	1.79×10^4
(200, 1)	-7.41	-4427.36
(200, 2)	4.0×10^{-11}	-3.3×10^{-10}
(200, 3)	172.36	2.73×10^4
(300, 1)	-23.82	-1.61×10^4
(300, 2)	7.76	2.50×10^3
(300, 3)	429.53	1.57×10^5
(400, 4)	471.18	1.80×10^5

FIG. C1. Flowchart of the wave-affected parameter estimation with the variational method.

$$|\nabla J|/|\nabla J_0| < 0.01.$$

The solution of the control variables that satisfies the convergence criterion is regarded as the optimal solution.

(vi) Integrate the model Eqs. (1)–(3) to ΔT_w using the optimal solution derived from process (v), and the results are regarded as the new initial fields for the next integration.

(vii) Use the new initial fields derived from the process (vi) and the optimal wave-affected parameters derived from process (v), iterate the processes (ii) to (vi) to obtain the time series of wave-affected parameters α and β [note that the background fields will also be updated by the new initial fields in process (ii)].

REFERENCES

Agrawal, Y. C., E. A. Terray, M. A. Donelan, P. A. Hwang, A. J. Williams, W. Drennan, K. Kahm, and S. Kitaigorodskii, 1992: Enhanced dissipation of kinetic energy beneath breaking waves. *Nature,* **359,** 219–220, doi:10.1038/359219a0.

Anis, A., and J. N. Moum, 1992: The superadiabatic surface layer of the ocean during convection. *J. Phys. Oceanogr.,* **22,** 1221–1227, doi:10.1175/1520-0485(1992)022<1221:TSSLOT>2.0.CO;2.

Babanin, A. V., 2006: On a wave-induced turbulence and a wave-mixed upper ocean layer. *Geophys. Res. Lett.,* **33,** L20605, doi:10.1029/2006GL027308.

——, and B. K. Haus, 2009: On the existence of water turbulence induced by nonbreaking surface waves. *J. Phys. Oceanogr.,* **39,** 2675–2679, doi:10.1175/2009JPO4202.1.

Belcher, S. E., and Coauthors, 2012: A global perspective on Langmuir turbulence in the ocean surface boundary layer. *Geophys. Res. Lett.,* **39,** L18605, doi:10.1029/2012GL052932.

Blumberg, A. F., and G. L. Mellor, 1987: A description of a three-dimensional coastal ocean circulation model. *Three-Dimensional Coastal Ocean Models,* N. S. Heaps, Ed., Coastal and Estuarine Sciences, Vol. 4, Amer. Geophys. Union, 1–16.

Burchard, H., 2001a: On the q^2l equation by Mellor and Yamada. *J. Phys. Oceanogr.,* **31,** 1377–1387, doi:10.1175/1520-0485(2001)031<1377:OTQLEB>2.0.CO;2.

——, 2001b: Simulating the wave-enhanced layer under breaking surface waves with two-equation turbulence models. *J. Phys. Oceanogr.,* **31,** 3133–3145, doi:10.1175/1520-0485(2001)031<3133:STWELU>2.0.CO;2.

Carniel, S., J. C. Warner, J. Chiggiato, and M. Sclavo, 2009: Investigating the impact of surface wave breaking on modeling the trajectories of drifters in the northern Adriatic Sea during a windstorm event. *Ocean Modell.,* **30,** 225–239, doi:10.1016/j.ocemod.2009.07.001.

Chu, P. C., and K. F. Cheng, 2007: Effect of wave boundary layer on the sea-to-air dimethylsulfide transfer velocity during typhoon passage. *J. Mar. Syst.,* **66,** 122–129, doi:10.1016/j.jmarsys.2006.01.013.

——, S. H. Lu, and Y. C. Chen, 2001: Evaluation of the Princeton Ocean Model using the South China Sea Monsoon Experiment (SCSMEX) data. *J. Atmos. Oceanic Technol.,* **18,** 1521–1539, doi:10.1175/1520-0426(2001)018<1521:EOTPOM>2.0.CO;2.

Craig, P. D., 1996: Velocity profiles and surface roughness under wave breaking. *J. Geophys. Res.,* **101,** 1265–1277, doi:10.1029/95JC03220.

——, and M. L. Banner, 1994: Modeling wave-enhanced turbulence in the ocean surface layer. *J. Phys. Oceanogr.,* **24,** 2546–2559, doi:10.1175/1520-0485(1994)024<2546:MWETIT>2.0.CO;2.

D'Alessio, S. J. D., K. Abdella, and N. A. Mcfarlane, 1998: A new second-order turbulence closure scheme for modeling the oceanic mixed layer. *J. Phys. Oceanogr.,* **28,** 1624–1641, doi:10.1175/1520-0485(1998)028<1624:ANSOTC>2.0.CO;2.

Denman, K. L., and M. Miyake, 1973: Upper layer modification at ocean station *Papa*: Observations and simulation. *J. Phys. Oceanogr.,* **3,** 185–196, doi:10.1175/1520-0485(1973)003<0185:ULMAOS>2.0.CO;2.

Derber, J. C., 1987: Variational four-dimensional analysis using quasi-geostrophic constraints. *Mon. Wea. Rev.,* **115,** 998–1008, doi:10.1175/1520-0493(1987)115<0998:VFDAUQ>2.0.CO;2.

Donelan, M. A., 1990: Air-sea interaction. *Ocean Engineering Science, Parts A and B,* B. LeNehaute and D. M. Hanes, Eds., The Sea—Ideas and Observations on Progress in the Study of the Seas, Vol. 9, John Wiley and Sons, 239–292.

Drennan, W. M., M. A. Donelan, E. A. Terray, and K. B. Katsaros, 1996: Oceanic turbulence dissipation measurements in SWADE. *J. Phys. Oceanogr.,* **26,** 808–815, doi:10.1175/1520-0485(1996)026<0808:OTDMIS>2.0.CO;2.

Ezer, T., and G. L. Mellor, 2004: A generalized coordinate ocean model and a comparison of the bottom boundary layer dynamics in terrain-following and in *z*-level grids. *Ocean Modell.,* **6,** 379–403, doi:10.1016/S1463-5003(03)00026-X.

Gaspar, P., 1988: Modelling the seasonal cycle of the upper ocean. *J. Phys. Oceanogr.,* **18,** 161–180, doi:10.1175/1520-0485(1988)018<0161:MTSCOT>2.0.CO;2.

Giering, R., and T. Kaminski, 1998: Recipes for adjoint code construction. *ACM Trans. Math. Software,* **24,** 437–474, doi:10.1145/293686.293695.

Harcourt, R. R., 2013: A second-moment closure model of Langmuir turbulence. *J. Phys. Oceanogr.,* **43,** 673–697, doi:10.1175/JPO-D-12-0105.1.

Janssen, P. A. E. M., 2001: Reply. *J. Phys. Oceanogr.,* **31,** 2537–2544, doi:10.1175/1520-0485(2001)031<2537:R>2.0.CO;2.

Jones, N. L., and S. G. Monismith, 2008: Modeling the influence of wave-enhanced turbulence in a shallow tide- and wind-driven water column. *J. Geophys. Res.,* **113,** C03009, doi:10.1029/2007JC004246.

Kantha, L. H., U. Lass, and H. Prandke, 2010: A note on Stokes production of turbulence kinetic energy in the oceanic mixed layer: Observations in the Baltic Sea. *Ocean Dyn.,* **60,** 171–180, doi:10.1007/s10236-009-0257-7.

Kitaigorodskii, S. A., and J. L. Lumley, 1983: Wave turbulence interactions in the upper ocean. Part I: The energy balance of the interacting fields of surface wind waves and wind-induced three-dimensional turbulence. *J. Phys. Oceanogr.,* **13,** 1977–1987, doi:10.1175/1520-0485(1983)013<1977:WTIITU>2.0.CO;2.

Kraus, E. B., and J. S. Turner, 1967: A one-dimensional model of the seasonal thermocline II. The general theory and its consequences. *Tellus,* **19A,** 98–105, doi:10.1111/j.2153-3490.1967.tb01462.x.

Le Dimet, F.-X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus,* **38A,** 97–110, doi:10.1111/j.1600-0870.1986.tb00459.x.

Liu, D. C., and J. Nocedal, 1989: On the limited memory BFGS method for large scale optimization. *Math. Program.,* **45,** 503–528, doi:10.1007/BF01589116.

Martin, P. J., 1985: Simulation of the mixed layer at OWS November and Papa with several models. *J. Geophys. Res.,* **90,** 903–916, doi:10.1029/JC090iC01p00903.

Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure models for geophysical fluid problems. *Rev. Geophys.,* **20,** 851–875, doi:10.1029/RG020i004p00851.

——, and A. F. Blumberg, 2004: Wave breaking and ocean surface layer thermal response. *J. Phys. Oceanogr.,* **34,** 693–698, doi:10.1175/2517.1.

Osborn, T., D. M. Farmer, S. Vagle, S. A. Thorpe, and M. Cure, 1992: Measurements of bubble plums and turbulence from a submarine. *Atmos.–Ocean,* **30,** 419–440, doi:10.1080/07055900.1992.9649447.

Peng, S.-Q., and L. Xie, 2006: Effect of determining initial conditions by four-dimensional variational data assimilation on storm surge forecasting. *Ocean Modell.,* **14,** 1–18, doi:10.1016/j.ocemod.2006.03.005.

——, ——, and L. J. Pietrafesa, 2007: Correcting the errors in the initial conditions and wind stress in storm surge simulation using an adjoint optimal technique. *Ocean Modell.,* **18,** 175–193, doi:10.1016/j.ocemod.2007.04.002.

——, Y. Li, and L. Xie, 2013: Adjusting the wind stress drag coefficient in storm surge forecasting using an adjoint technique. *J. Atmos. Oceanic Technol.,* **30,** 590–608, doi:10.1175/JTECH-D-12-00034.1.

Robert, A. J., 1966: The integration of a low-order spectral form of the primitive meteorological equations. *J. Meteor. Soc. Japan,* **44,** 237–245.

Smith, S. D., and Coauthors, 1992: Sea surface wind stress and drag coefficients: The HEXOS results. *Bound.-Layer Meteor.,* **60,** 109–142, doi:10.1007/BF00122064.

Stacey, M. W., 1999: Simulations of the wind-forced near-surface circulation in Knight Inlet: A parameterization of the roughness length. *J. Phys. Oceanogr.,* **29,** 1363–1367, doi:10.1175/1520-0485(1999)029<1363:SOTWFN>2.0.CO;2.

Terray, E. A., M. A. Donelan, Y. Agarwal, W. M. Drennan, K. Kahma, A. J. Williams III, P. Hwang, and S. A. Kitaigorodskii, 1996: Estimates of kinetic energy dissipation under breaking waves. *J. Phys.*

Oceanogr., **26,** 792–807, doi:10.1175/1520-0485(1996)026<0792:EOKEDU>2.0.CO;2.

——, W. M. Drennan, and M. A. Donelan, 1999: The vertical structure of shear and dissipation in the ocean surface layer. *The Wind-Driven Air-Sea Interface: Electromagnetic and Acoustic Sensing, Wave Dynamics and Turbulent Fluxes,* M. L. Banner, Ed., University of New South Wales, 239–245.

Thorpe, S. A., 1984: The effect of Langmuir circulation on the distribution of submerged bubbles caused by breaking wind waves. *J. Fluid Mech.,* **142,** 151–170, doi:10.1017/S0022112084001038.

Umlauf, L., and H. Burchard, 2003: A generic length-scale equation for geophysical turbulence models. *J. Mar. Res.,* **61,** 235–265, doi:10.1357/002224003322005087.

Yu, L., and J. J. O'Brien, 1991: Variational estimation of the wind stress drag coefficient and the oceanic eddy viscosity profile. *J. Phys. Oceanogr.,* **21,** 709–719, doi:10.1175/1520-0485(1991)021<0709:VEOTWS>2.0.CO;2.

——, and ——, 1992: On the initial condition in parameter estimation. *J. Phys. Oceanogr.,* **22,** 1361–1364, doi:10.1175/1520-0485(1992)022<1361:OTICIP>2.0.CO;2.

Zhang, A., E. Wei, and B. B. Parker, 2003: Optimal estimation of tidal open boundary conditions using predicted tides and adjoint data assimilation technique. *Cont. Shelf Res.,* **23,** 1055–1070, doi:10.1016/S0278-4343(03)00105-5.

Zhang, X., G. Han, D. Wang, W. Li, and Z. He, 2011a: Effect of surface wave breaking on the surface boundary layer of temperature in the Yellow Sea in summer. *Ocean Modell.,* **38,** 267–279, doi:10.1016/j.ocemod.2011.04.006.

——, ——, X. Wu, W. Li, and D. Wang, 2011b: Effect of surface wave breaking on upper-ocean structure revealed by assimilating sea temperature data. *J. Trop. Oceanogr.,* **30** (5), 48–54.

——, ——, D. Wang, Z. Deng, and W. Li, 2012: Summer surface layer thermal response to surface gravity waves in the Yellow Sea. *Ocean Dyn.,* **62,** 983–1000, doi:10.1007/s10236-012-0547-3.