Variational Estimation of Wave-affected Parameters in a Two-equation Turbulence Model

Xuefeng Zhang[1,*] Guijun Han[1] Dong Li[1] Xinrong Wu[1] Wei Li[1] Peter C. Chu[2]

[1]Key Laboratory of Marine Environmental Information Technology, National Marine Data and Information Service, State Oceanic Administration, Tianjin, PRC

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

[1]

[1] *Corresponding author address*: Dr. Xuefeng Zhang.
E-mail: xfz.nmdis@hotmail.com

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2014		**00-00-2014 to 00-00-2014**

4. TITLE AND SUBTITLE	
Variational Estimation of Wave-affected Parameters in a Two-equation Turbulence Model	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Postgraduate School,Department of Oceanography,Naval Ocean Analysis and Prediction Laboratory,Monterey,CA,93943**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

Journal of Atmospheric and Oceanic Technology, In press.

14. ABSTRACT

A variational method is used to estimate wave-affected parameters in a two-equation turbulence model with assimilating the temperature data into an ocean boundary layer model. Enhancement of turbulent kinetic energy dissipation due to breaking waves is considered. The Mellor-Yamada 2.5 turbulence closure scheme (MY-2.5) 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 the 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.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**64**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Abstract

A variational method is used to estimate wave-affected parameters in a two-equation turbulence model with assimilating the temperature data into an ocean boundary layer model. Enhancement of turbulent kinetic energy dissipation due to breaking waves is considered. The Mellor-Yamada 2.5 turbulence closure scheme (MY-2.5) 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 the 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 et al. 1983; Thorpe 1984, 1992; Anis and Moum 1992; Terray el al. 1996; Drennan et al. 1996) show that the dissipation of turbulent kinetic energy (TKE) is enhanced greatly near the sea surface due to increasing shear by surface gravity waves under non-breaking (including Langmuir circulation) and breaking waves. The mixing induced by non-breaking waves directly affects or influences the upper-ocean mixing down to depths of the order of 100 m. With a wave amplitude-based Reynolds number (Re), an empirically determined critical value (Re_{cr}) is used to identify if the turbulence is generated by waves ($Re > Re_{cr}$) or not ($Re < Re_{cr}$); and to determine a depth of the upper ocean mixed layer from $Re = Re_{cr}$. Decrease of Re with depth confirms that within that depth the turbulence is generated by orbital movement of surface gravity waves; and below that depth there is no wave-induced turbulence (*Babanin* 2006; *Babanin and Haus* 2009).

The breaking wave induced mixing has been broadly implemented into ocean circulation and mixing models (e.g., Craig 1994). On the base of the observational evidences on the surface wave breaking (Osborn et al. 1992; Agrawal et al. 1992), Terray el 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 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, 1996) imposed a surface diffusion boundary condition on the turbulent kinetic energy equation (hereafter, 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 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 (2011, 2012) identified the effect of breaking surface waves on upper ocean boundary layer deepening in the Yellow Sea in summer utilizing the Princeton ocean model (POMgcs, Ezer et al., 2004). A well-mixed temperature surface layer in the Yellow Sea can be reconstructed successfully when the breaking wave enhanced turbulent mixing are considered.

In addition to the wave breaking, other wave-related processes are also important in modulating the upper mixed layer, such as the non-breaking wave (Babanin et al, 2009) and the Langmuir circulation (Stephen 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; Babanin et al, 2005). However, for a deeper mixed layer over 100 m depth, the impact of wave breaking would be small and the effect of Langmuir circulation and non-breaking wave becomes important (Babanin, 2005).

Uncertain wave-affected parameters exist in modeling wave-induced turbulence (non-breaking or breaking waves) such as critical value of the wave Renolds number (Re_{cr}) in non-breaking waves and wave energy factor (α) and 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) 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 4D-Variational method (4D-VAR) in estimating uncertain parameters in numerical models. Peng et al (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 (2012)]. Peng et al (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 non-wave 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 MY-2.5 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. Methodology

2.1 Ocean boundary layer model

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

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

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

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial z}(K_H \frac{\partial T}{\partial z}) - \frac{\partial R}{\partial z} \qquad (1)$$

$$\frac{\partial S}{\partial t} = \frac{\partial}{\partial z}(K_H \frac{\partial S}{\partial z})$$

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, 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 Regional Ocean Modeling System (ROMS), is a two-equation turbulence model,

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

$$\frac{\partial q^2 l}{\partial t} = lE_1(K_M((\frac{\partial u}{\partial z})^2 + (\frac{\partial v}{\partial z})^2) + \frac{g}{\rho_0}K_H\frac{\partial \rho}{\partial z} - \frac{q^3}{B_1 l}\frac{W}{E_1}) + \frac{\partial}{\partial z}(K_q \frac{\partial q^2 l}{\partial z}), \qquad (3)$$

where q^2 is the turbulent kinetic energy times two, l is the turbulent macroscale. K_q is the vertical mixing coefficient for turbulence, 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 Karman 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

$$K_M = lqS_M, \quad K_H = lqS_H, \quad K_q = lqS_q, \qquad (4)$$

where S_M and S_H are the stability functions.

2.2 Wave-affected parameters

Wave-affected parameters are included into 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),

$$K_q \frac{\partial q^2}{\partial z} = 2\alpha u_\tau^3, \qquad z = 0 \tag{5}$$

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),

$$l = \max(\kappa z_w, l_z) \tag{6}$$

where l_z is the "conventional" empirical length scale, which is calculated prognostically by the MY2.5 turbulence closure scheme; 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 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. In thepast, Kraus and Turner (1967), Denman and Miyake (1973), Gaspar (1988) also choose different values of α in their studies.

Terray et al. (1996), Burchard (2001a), 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:

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

where g is the gravitational acceleration, and β is the Charnock parameter (Chu and Cheng 2007), which varies from $\beta = 2$ (Stacey 1999), $\beta = 0.32$ (Jones and Monismith 2008) to β = 0.56 (Carniel et al. 2009) to obtain the best performance in each numerical simulation. Mellor and Blumberg (2004) suggested that β ~O(1) is deemed correct under breaking waves. Smith et al. (1997) also indicates that β ~O(10) is too big value to describe the surface boundary condition for the turbulent kinetic energy.

2.3 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

$$q^2 = B_1^{2/3} u_{\tau b}^2 \tag{8}$$

$$l = \kappa z_0 \tag{9}$$

where B_1 = 16.6 (Blumberg and Mellor 1987), $u_{\tau b}$ 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

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

and

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

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

2.4 The variational analysis

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-square framework, a general form of the cost function can be defined as

$$J(p)=\frac{1}{2}\int_0^T < W(CX-X_{obs}),(CX-X_{obs}) > \qquad (12)$$

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

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

F is the differential operator. < > is the inner product in the Euclidean space. W is the weight matrix. X_{obs} is the observation, and C is the projection operator from the model space to the observational space. Let

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

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

$$\nabla J(p^{obs}) = 0$$

with respect to all control variables. Here, ∇ is the gradient operator. The process for the variational analysis can be outlined as follows:

(a) Define a concrete cost function that reflects the misfit between the control variables and the available observations.

(b) Calculate the value of the cost function $J(p)$ through integrating the dynamical model with a fixed time step.

(c) Calculate the gradients of the cost function with respect to all control variables, $\nabla J(p)$.

(d) Minimize the cost function through a minimization algorithm according to the value of $J(p)$ and $\nabla J(p)$.

(e) Estimate the optimal control variables p^{obs} according to the convergence criterion of the process of the minimization.

For executing the above process of the variational analysis (a)-(e), $\nabla J(p)$ 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

$$J(p) = J(p_0) + \delta J(p) \tag{13}$$

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

$$\delta J(p)=\int_0^T<\nabla_X J(p_0),\delta X> \tag{14}$$

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

$$\delta J(p)=\frac{1}{2}\int_0^T< W\frac{\partial C}{\partial X}\delta X, CX-X_{obs}>+\frac{1}{2}\int_0^T< W(CX-X_{obs}),\frac{\partial C}{\partial X}\delta X> . \tag{15}$$

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

$$\delta J(p)=\int_0^T< W(CX-X_{obs}),\frac{\partial C}{\partial X}\delta X> \tag{16}$$

Let $\frac{\partial C}{\partial X}=A(\mathrm{X})$, thus Eq. (15) can be given as

$$\delta J(p)=\int_0^T< W(CX-X_{obs}), A(X)\delta X> \tag{17}$$

where $A(X)$ is called as the tangent linear operator. Eq.(17) can be transposed according to the definition of the adjoint operator:

$$\delta J(p)=\int_0^T< WA^*(X)(CX-X_{obs}),\delta X> \tag{18}$$

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

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

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 $CX-X_{obs}$ is regarded as an external forcing of the adjoint model.

2.5 The adjoint model

The dynamical model composed by Eq.(1)~(11) can be summarized in a general form as

$$\frac{\partial x}{\partial t} = F(x)$$
$$x\big|_{t_0} = x_0$$
$$x(t)\big|_{\Gamma} = y(t) \qquad (20)$$

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

The tangent linear model of Eq. (20) can be written by

$$\frac{\partial x'}{\partial t} = \frac{\partial F(x)}{\partial x} x'$$
$$x'\big|_{t_0} = x'_0 \qquad (21)$$
$$x'(t)\big|_{\Gamma} = y'(t)$$

where the prime is the perturbations of the state variables.

For the two vectors *w* and *z* in the Euclidean space, the adjoint operator L^* of the linear operator *L* can be defined as:

$$\langle z, Lw \rangle = \langle L^* z, w \rangle$$

In the Euclidean space, L^* is the transpose of *L*, namely $L^* = L^{\mathrm{T}}$. The adjoint model corresponding to (21) is given by

$$\frac{\partial \tilde{x}}{\partial t} = -\left(\frac{\partial F(x)}{\partial x}\right)^T \tilde{x}$$
$$\tilde{x}\big|_{t_E} = 0$$
$$\tilde{x}(t)\big|_{\Gamma} = 0 \qquad (22)$$

where $\tilde{x}$ represents the adjoint variables and t_E is the end time in the temporal integration of Eq. (20). The negative sign in the right side of the first equation in (22) indicates that the adjoint model integrates backward in time. When the adjoint model integrates backward to the initial time t_0, the corresponding $\tilde{x}|_{t=t_0}$ is the gradient of the cost function with respect to the state variables.

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 man-made 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 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 recomputation to shorten 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 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 Smith et al. (1997). 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.

2.6 Cost function

The cost function is defined by

$$\begin{aligned} J(T^n, T^{n-1}, \alpha, \beta) &= \frac{1}{2}(T^n - T_b^n)^T B_1^{-1}(T^n - T_b^n) + \frac{1}{2}(T^{n-1} - T_b^{n-1})^T B_2^{-1}(T^{n-1} - T_b^{n-1}) \\ &+ \frac{1}{2}\sum_{j=1}^{M}\sum_{i=1}^{N}(T_{j,i}(\alpha,\beta) - T_{obs})^T R^{-1}(T_{j,i}(\alpha,\beta) - T_{obs}) \end{aligned} \quad (23)$$

where the first two terms in the right side represent the background error terms that measure the misfit between the model's initial field and the background field. T^n and T^{n-1} are the initial temperature values at the n^{th} and $(n-1)^{th}$ time step respectively, which will be estimated optimally via the variational method. T_b^n and T_b^{n-1} are the background temperature values at the n^{th} 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. 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*. *N* and *M* are the number of grid points over the ocean and the number of time levels of observations. *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. (23), 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),

$$\begin{aligned} J(T^n, T^{n-1}, \alpha, \beta) &= \frac{1}{2}(T^n - T_b^n)^T B_1^{-1}(T^n - T_b^n) + \frac{1}{2}(T^{n-1} - T_b^{n-1})^T B_2^{-1}(T^{n-1} - T_b^{n-1}) \\ &+ \frac{1}{2}\sum_j^M \sum_i^n (T_{j,i}(\alpha,\beta) - T_{obs})^T R^{-1}(T_{j,i}(\alpha,\beta) - T_{obs}) + \frac{1}{2}K_\alpha(\alpha - \alpha_b)^2 + \frac{1}{2}K_\beta(\beta - \beta_b)^2 \end{aligned} \tag{24}$$

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 (23) for avoiding the complexity of the cost function.

3. Synthetic experiments

3.1. "Truth" model simulation

Table 1 lists all the assimilation experiments and model simulations within an identical synthetic experiment framework. The "truth" model consists of Eq.(1)-(3) with $\alpha = 200$ and $\beta = 2$. All the 6 equations from (1) to (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, 35.0 m. The time step is 1-hour. The model initial state is from Jan. 1, 1961, including temperature and salinity, derived from the real observation at OWS Papa. The model is forced by the observational 10-minutes momentum and heat fluxes from http://www.pmel.noaa.gov/stnp/data.html.

Starting from the initial conditions (Jan. 1, 1961), the "truth" model is run for 6-yr to generate time series of the "truth" with the first 5-yr as the spin-up period. The time of the "observations" of T is from Aug. 1, 1966 to Aug. 30, 1966. The "observations" of T are produced through sampling the "truth" states at 1-hour observational frequencies. The "observation" locations of T are consistent with those of the model vertical grids.

3.2. Biased simulation

The biased simulation uses the same "truth" model, but with different parameter settings. Therefore, the difference between the biased simulation and "truth" model leads to the effect of the "incorrect" parameter settings. Fig.1 shows the simulated daily temperature at OWS Papa in 1966. The sea surface temperature (SST) from the biased simulation with (α, β) = (100, 1) is higher than that by the "truth" model simulation with (α, β) = (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. dash line in Fig.1a). 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. dash 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 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 on the temperature near the sea surface in summer, and indirectly impact on 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.

3.3. Parameter estimation

Fig.2 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:

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

(b) 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. (23).

(c) 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.

(d) 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 initial upper layer temperature fields T^n , T^{n-1}.

(e) With the updated control variables from process (d), repeat processes (b) - (d) until the convergence criterion for the minimization is satisfied. The convergence criterion is defined as

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

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

(f) Integrate the model Eqs. (1) -(3) to the fixed time window ΔT_w using the optimal solution derived from process (e), and results are regarded as the new initial fields for the next integration.

(g) Use the new initial fields derived from the process (f) and the optimal wave-affected parameters derived from process (e), iterate the processes (b) to (f) to obtain time series of wave-affected parameters α and β.

Fig.3 shows the time series of α and β during the parameter estimation (PE) described in Table 1, where both assimilation window and frequency are set to 24 hours and the assimilation depth is set to 30 m. Therefore, the processes (b)~(g) are executed 30 times to obtain time series of α and β from Aug. 1, 1966 to Aug. 30, 1966. Fig.3b shows that β converges to its "truth" value (dash line) after 9 days, while α converges to its "truth" value (dash 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. Fig.4 shows the dependence of the cost function and the norm of the gradient on the number of iterations on Aug. 2, 1966. The value of the cost function decreases rapidly from 4.3 to 0.8 within first 5 iterations, and keeps the low value (0.8) steadily after the 5th 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 goes stable after the 130th iteration (Fig. 4b). The minimization process stops after 180th iterations, indicating the local minima of the wave-affected parameters for that day.

Fig.5 is the temporal variations of the natural logarithm of the cost function at OWS Papa from Aug. 1 to Aug. 30, 1966. The cost function (red line) decreases dramatically in the first 5 days, then decreases gently in the following 25 days. Both the background term (blue line in the Fig.5) and the observation term (black line in the Fig.5) of the cost function have a similar pattern with the total cost function. The two terms almost converge to the same value after the 10th days, 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 Aug.1, 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, 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 temperature observation. The relative error of optimally-estimated α decreases from 96.9% to 60%, and the relative error of optimally-estimated β decreases from 99.1% to 94.3% as the error standard deviation in temperature observation increases from 0.001 to 0.05K. It implies that the effect of observational noise on the estimation is more severe on α is than on β, which means that it is more difficult to pick up the useful signal when the noise dominates the cost function and corresponding gradients during the parameter estimation of α. When the standard deviation of temperature observation increases to 0.5K, 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 explore if the wave-affected parameters can be estimated correctly only using 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 to 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 frequency is investigated using different values from Aug.1 to Aug. 30, 1966 (Fig. 8). When the assimilation window and frequency are 48 hours and 72 hours, both parameters converge to their respective "truth" values (see black and blue lines in Figs. 8a and 8b). However, when the assimilation window and frequency reach 96 hours and 120 hours, 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 the improved state estimation further enhances the quality of parameter estimation for the next cycle of parameter correction. When the assimilation window and frequency are set to 120 hours, the state-parameter optimization is performed only in 6 cycles within one month. Although the cost function decreases gradually, which can be seen from the dash curve in Fig. 9, the control variables (the initial temperature *T*and two parameters α, β) are not estimated reasonably well. In contrast, when the assimilation window and frequency are set to 24 hours, just as 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).

The incorrect convergence of (α, β) suggests that the initial temperature field 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. For better understanding 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, norm of the projected gradient and value of β with respect to the number of iterations in PE_β_TI. The parameter β reaches its "truth" value at the 3rd iteration. The convergence criterion of the gradient is satisfied at the 4th 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 6th iteration, β reaches 3.302335, which is different from the "truth" value 2.0. Although β from PE_β_BI cannot converge to its "truth" value, it reaches its optimal value to compensate the error derived from the "biased" initial filed 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 both numerical and physical deficiencies of numerical model and enhance the performance of the model simulation to certain degree. Under this situation, parameters can only converge to their optimal value, instead of the "truth" values. In the next section, real temperature profiles from OWS Station Papa will be assimilated into the assimilation model to obtain the optimal wave-affected parameters (α, β).

4. **Real experiment**

The Papa Station locates in the North Pacific at (145°W, 50°N), where the currents are relatively weak and the local mixing modulates mainly the dynamical process in the upper ocean in summer. The observed temperature profiles from Aug 1 to Aug 31, 1966 at the site have 3h time interval and a coarser vertical resolution

(5 m) than the model grid points. There are 7 observational layers totally in the upper 30 m, namely 0,5,10,15,20, 25,30 m. Linear interpolation is used to fill the spatial gap between the modeled 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 00:00 GMT on Jan 1, 1961, and linearly interpolated to model grids. The high resolution ($1/6^o$) surface observed data (momentum and net heat fluxes) at the site are used to force the model. Fig.10a shows the daily observed (red curve) and simulated sea surface temperature from CTRL (black dashed curve) at the OWS Papa on Aug., 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 Aug. 1, 1966 simulated by the "truth" model in the above synthetic experiments. The initial values of (α, β) are also consistent with those in the "truth" model simulation. The length of the assimilation window is set to 3 days (8 real observational temperature profiles in each day, totally 24 profiles within 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 Aug., 1966. Table 6 shows the evolution of the cost function, α and β with respect to the number of iteration for RPE. After the 8th 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 exist 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 in the simple ocean boundary layer model. The optimal values of the parameters can only compensate 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 Aug 31, 1966, which are derived from RPE. Fig. 11 shows the observed (red curve) and simulated SST from RSE_Po (black solid curve) and

RSE_Pd (black dash curve) at OWS Papa from Sept. 1 to Sept 30, 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 vertical mixing coefficient for temperature K_H (Fig. 12b) at OWS Papa in Sept 1966 are much larger for all depths in RSE_Po (solid curve) than in RSE_Pd (dashed curve). The enhanced K_H in the upper 30m 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 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 repeatedly at the certain time intervals with more available observations.

5. Discussion and conclusion

The upper layer temperature data is assimilated into an ocean surface boundary layer model to 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 frequency, and the assimilation depth are set appropriately. The observational temperature profiles at the OWS Station 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 parameters estimation using the variational method can compensate in part the numerical and physical deficiencies of the model in the upper ocean. However, It should also be noted that the optimal values of the wave-affected parameters are not the so called "truth" values. The "optimal" values of the wave-affected parameters in real applications are only applicable to the specific time period, location, and model. The "optimal" values should vary temporally and spatially rather than being constants, which can be obtained by using the variational methods repeatedly at the certain time intervals and available observations (Peng et al, 2012). Although the optimal values of the wave-affected parameters are model-dependent (initial fields, time window of assimilation, model configuration, etc.), 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.

In general, the complex turbulent closure models are empirical and full of uncertainty in an ocean circulation model. Due to the high nonlinearity and discontinuity of the vertical turbulence, it is more difficult to linearize the complicated turbulence closure scheme than to linearize the momentum and tracer equations. Wave-affected parameters in high order turbulence closure schemes can modulate distinctly the vertical structure in the upper ocean. Therefore, it is essential to estimate their optimal values using observations deployed in the upper ocean through some robust data assimilation methods such as the variational method or the ensemble Kalman filter. Now, 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. Sensitivity of simulated temperature to parameters

It is essential to investigate model sensitivities with respect to parameters being estimated before parameter estimation. Fig. A1 shows the dependence of the cost function on α and β. It increases with the 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. A1b). 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 remain unperturbed. All the 100 ensemble members are started from the same initial conditions (Jan. 1, 1961). The biased-simulation model is integrated up to 6 years. Sensitivities are calculated with the model output from Aug. 1, 1966 to Aug. 31, 1966. This process is looped for the two wave-affected parameters. Fig. A2 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 30m. 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.

The sensitivities with respect to the wave-affected parameters are also investigated through calculating the gradients of the cost function with the parameters, namely $\frac{\partial J}{\partial \alpha}$ and $\frac{\partial J}{\partial \beta}$. Table A1 shows the dependence of the sensitivity on the initial values of the parameter 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 be disturbed by the noises arisen from the observational errors and the biased initial state fields during the parameter estimation.

Appendix B. Correctness test of the gradient

The code of the adjoint model is produced directly through the Adjoint Model Compiler (AMC) developed by Giering and Kaminski (1998) (Of course, a hand-coding correction is necessary after that), which means that it is not essential to produce the code of the tangent linear model explicitly. Therefore, only the correctness of the adjoint model is tested here.

According to the Taylor expression, one has

$$\lim_{\varepsilon\to 0}\varphi(\varepsilon)=\lim_{\varepsilon\to 0}\frac{J(x_0-\varepsilon\nabla J(x_0))-J(x_0)}{-\varepsilon<\nabla J(x_0),\nabla J(x_0)>}\approx 1 \qquad \text{(A1)}$$

where x_0 is any control variable, the symbol $<\ \ >$ represents the inner product. Fig. A3 shows the correctness test of the gradient of the cost function with respect to α and β using the Eq. (A1). 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.

Acknowledgement

This research was jointly supported by grants from the National Basic Research Program of China (2013CB430304), and the National Natural Science Foundation of China (under grants 41030854, 41106005, 41176003 and 41206178). Peter C. Chu was supported by the Naval Oceanographic Office.

Reference

Austin J A, Lentz S J, 2002. The inner shelf response to wind-driven upwelling and downwelling. Journal of Physical Oceanography, 32: 2171-2193.

Agrawal, Y.C., Terray, E.A., Donelan, M.A., Hwang, P.A., Williams, A. J., Drennan, W., Kahm, K., Kitaigorodskii, S., 1992. Enhanced dissipation of kinetic energy beneath breaking waves. Nature, 359, 219-220.

Anis, A., Moum, J.N., 1992. The superadiabatic surface layer of the ocean during convection. [J]. Journal of Physical Oceanography, 22, 1221-1227.

Babanin, A.V., Young, I.R., Mirfenderesk, H., 2005. Field and laboratory measurements of wave–bottom interaction. In: Townsend, M., Walker, D. (Eds.), Proc. 17th Australasian Coastal and Ocean Engineering Conf. and 10th Australasian Port and Harbour Conf., 20–23 September 2005, Adelaide, South Australia. The Institution of Engineers, Canberra, Australia, pp. 293–298.

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.

Babanin, A. V., and B. K. Haus (2009), On the existence of water turbulence induced by nonbreaking surface waves, J. Phys. Oceanogr., 39, 2675–2679, doi:10.1175/2009JP04202.1.

Blumberg, A.F., Mellor, G.L., 1987. A description of a three-dimensional coastal ocean circulation model. In: Heaps, N. (Ed.), Three Dimensional Coastal Ocean Models, vol. 4. American Geophysical Union, Washington, DC, p. 208.

Burchard, H., 2001a. On the q2l equation by Mellor and Yamada. Journal of Physical Oceanography, 31, 1377-1387.

Burchard, H., 2001b. Simulating the Wave-Enhanced Layer under Breaking Surface Waves with Two-Equation Turbulence Models. J. Phys. Oceanogr., 31, 3133–3145.

Carniel, S., Warner, J.C., Chiggiato, J., Sclavo, M., 2009. Investigating the impact of surface wave breaking on modeling the trajectories of drifters in the northern Adriatic Sea during a wind event. Ocean Modelling, 30, 225-239.

Chu, P.C., 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.

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. Marine Syst., 66, 122-129

Craig, P.D., Banner, M.L., 1994. Modeling wave-enhanced turbulence in the ocean surface layer. J. Phys. Oceanogr.,24, 2546-2559.

Craig ,P.D., 1996. Velocity profiles and surface roughness under wave breaking. J. Geophys.Res., 101, 1265-1277.

D'Alessio, S.J.D., K. Abdella, N.A. Mcfarlane, 1998. A new second-order turbulence closure scheme for modeling the oceanic mixed layer. Journal of Physical Oceanography, 28, 1624-1641.

Derber, John C., 1987. Variational four-dimensional analysis using quasi-geostrophic constraints. Mon. Weather Rev. 115, 998–1008.

Denman K L, Miyake M, 1973. Upper layer modification at ocean station Papa: observations and simulation. J. Phys. Oceanogr., 3, 185–196.

Donelan,M.A., 1990. Air-sea interaction.The Sea, B.LeNehaute and D.M.Hanes,Eds.,Ocean Engineering Science, 9, 239-292.

Drennan, W.M., Donelan, M.A., Terray, E.A., Katsaros, K.B., 1996. Oceanic turbulence dissipation measurements in SWADE. Journal of Physical Oceanography, 26, 808-815.

Ezer, T., Mellor,G.L., 2004. A generalized coordinate ocean model and a comparison of the bottom boundary layer dynamics in terrain-following and in z-level grids. Ocean Modelling, 6, 379-403.J Geophysical Res, 105, 16843-16855.

Gaspar, P., 1988. Modelling the seasonal cycle of the upper ocean. J. Phys. Oceanogr., 18, 161–180.

Giering, R., Kaminski, T., 1998. Recipes for adjoint code construction. ACM Trans. Math. Software 24, 437–474.

Janssen, P. A. E. M., 2001. Reply. J Phys Oceanogr., 31, 2532-2544.

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

Kitaigorodskii, S.A., Lumley, J.L.,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. Journal of Physical Oceanography, 13, 1977-1987.

Kraus, E.B, Turner, J.S., 1967. A one-dimensional model of the seasonal thermocline II. The general theory and its consequences.Tellus, 19:98-105

LeDimet, F.X., Talagrand, O., 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. Tellus 38A, 97–110.

Liu, D.C., Nocedal, J., 1989. On the Limited memory BFGS method for large scale optimization. Math. Program. 45, 503–528.

Martin, P.J., 1985. Simulation of the mixed layer at OWS November and Papa with several models.J.Geophys.Res., 90, 581-597.

Mellor, G.L., Yamada,T., 1982. Development of a turbulence closure models for geophysical fluid problems. Rev, Geophys., 20, 851-875.

Mellor. G.L, Blumberg, A.F., Wave breaking and ocean surface layer thermal response. J. Phys. Oceanogr., 2004, 34, 693-698.

Terray, E.A., Donelan, M.A., Agarwal, Y., Drennan, W.M., Kahma, K., Williams III, A. J., Hwang, P., Kitaigorodskii, S. A., 1996. Estimates of kinetic energy dissipation under breaking waves. J. Phys. Oceanogr., 26, 972–987.

Terray, E.A., Drennan, W.M., Donelan, M.A., 1999. The vertical structure of shear and dissipation in the ocean surface layer, in the Wind-driven Air-Sea Interface. Electromagnetic and Acoustic Sensing, Wave Dynamics and Turbulent Fluxes, Sydney,Australia,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. Journal of Fluid Mechanics, 142,151-170.

Osborn, T., Farmer, D.M., Vagle, S., Thorpe, S.A., Cure, M., 1992. Measurements of bubble plums and turbulence from a submarine. Atmos-Ocean, 30, 419-440.

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.

——, ——, 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.

Peng, S. Q., Y. Li, L. Xie,2012. Adjusting the wind stress drag coefficient in storm surge forecasting using an adjoint technique. Journal of atmospheric and oceanic technology, 30, 590-608.

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., Coauthors., 1992. Sea surface wind stress and drag coefficients:The HEXOS results. Bound-layer Meteor., 60, 109-142.

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.

Umlauf, L. and Burchard, H., 2003. A generic length-scale equation for geophysical turbulence models. Journal of Marine Research, 61, 235-265.

Yu, L., O'Brien, J.J., 1991. Variational estimation of the wind stress drag coefficient and the oceanic eddy viscosity profile. J. Phys.Oceanogr. 21, 709–719.

Yu, L., O'Brien, J.J., 1992. On the initial condition in parameter estimation. J. Phys. Oceanogr. 22, 1361–1364.

Zhang, A., Wei, E., Parker, B.B., 2003. Optimal estimation of tidal open boundary conditions using predicted tides and adjoint data assimilation technique. Continental Shelf Res. 23, 1055–1070.

Zhang X, Han G,Wang D, et al, 2011. Effect of surface wave breaking on the surface boundary layer of temperature in the Yellow Sea in summer. Ocean Modelling, 38: 267-279.

Zhang X, Han G, Wang D, et al, 2012. Summer surface layer thermal response to surface gravity waves in the Yellow Sea. Ocean Dynamics, 62: 983-1000.

Table captions

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

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

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

Table 4. Same as Table 2 but with biased initial field.

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

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

Table A1. Dependence of the sensitivity on the initial values of the parameter α and β.

Figure captions

Figure 1. Daily temperature in 1966 at (a) 0 m, (b) 10 m, (c) 20 m and (d) 30 m at the OWS Papa with the "truth" model simulation (solid curve) and the biased simulation (dashed curve).

Figure 2. Flowchart of the wave-affected parameter estimation using the variational method.

Figure 3. Time series of the estimated wave-effected parameters (a) α, and (b) β for PE from Aug.1 to Aug. 30, 1966 (solid curve), where both assimilation window and frequency are 1 day and the depth of the assimilation is 30 m. Here, the dashes curves show the "truth" (α, β) values.

Figure 4. Dependence of (a) the cost function and (b) the norm of the gradient on the number of iterations on Aug, 2, 1966.

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

Figure 6. Time series of the estimated wave-effected parameters (a) α, and (b) β for different initial parameter values from Aug.1 to Aug. 30, 1966. Here, both assimilation window and frequency are 1 day and the depth of the assimilation is 30 m. The black, blue, green, yellow, red, pink, purple, orange and gray solid curves in panel (a) and the corresponding dashed curves in panel (b) show values of (α, β) with the initial guess values of (0,0), (100,2), (100,3), (200,1), (200,3), (300,1), (300,2), (300,3) and (400,4), respectively.

Figure 7. Same as Figure 3 but for PE_SST with only the SST observations being assimilated.

Figure 8. Time series of the estimated wave-effected parameters (a) α, and (b) β for different assimilation window and frequency from Aug.1 to Aug. 30, 1966 with 30 m as the depth of the assimilation. Here, the black, blue, red, pink solid curves in panel (a) and panel (b) show the assimilation frequency are 48, 72, 96, and 120 hours.

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

Figure 10. (a) Sea surface temperature and (b) mixed layer depth from CTRL (black dashed curve) and RPE (black solid curve), and observations (red solid curve) at OWS Papa from Aug. 1 to Aug. 30, 1966. The horizontal axis represents the day relative to Aug. 1, 1966.

Figure 11. Sea surface temperature from observations (red solid curve), RSE_Po (black solid curve), and RSE_Pd (black dashed curve) at OWS Papa from Aug. 1 to Sept 30, 1966. The horizontal axis represents the day relative to Aug. 1, 1966.

Figure 12. Vertical profiles of the simulated monthly-averaged (a) two times turbulent kinetic energy q^2 (m^2s^{-2}), and (b) and vertical mixing coefficient for temperature K_H ($10^{-3}m^2s^{-1}$) from RSE_Po (solid curve) and RSE_Pd (dashed curve) at OWS Papa in Sep., 1966.

Figure A1. Dependence of the cost function on α and β for (a) $10 \geq \beta \geq 0$, and (b) $3 \geq \beta \geq 0$.

Figure A2. Depth-dependence of ensemble spread of temperature with respect to the wave-effected parameters α (dashed curve) and β (solid curve).

Figure A3. The correctness test of the gradient with respect to (a) α, and (b) β.

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

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

	model simulation				
PE_ β _BI	Parameter estimation with the “biased” initial fields derived from the biased simulation	β	1 day	1 day	30 m

Table 2. dependence of the optimally-estimated (α, β) on the standard deviation of temperature observation

Standard deviation 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%

Table 3. Evolution of the cost function, norm of the projected gradient and 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	1.354e-5	7.406e-2	2.000365
2	2.631e-9	1.032e-3	2.000005
3	7.441e-17	3.042e-7	2.000000
4	5.056e-17	5.693e-9	1.999999

Table 4. Same as Table 3 but with the biased initial field.

Iteration step	Cost function	Norm of the projected gradient	Value of β
0	2.319e2	4.819	1.0
1	1.072e2	3.351	3.350811
2	1.071e2	1.234	3.317216
3	1.071e2	9.003e-2	3.301275
4	1.071e2	1.982e-3	3.302359
5	1.071e2	6.043e-6	3.302335
6	1.071e2	3.627e-6	3.302335

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

Name	Description	Control variables	Assimilation windows	Assimilation frequency	Assimilation depth	Initial fields
CTRL	Simulation with $\alpha = 200$ β=2	–	–	–	–	1 Aug. 1966 from the "truth" model simulation
RPE	Real parameter estimation	T^n, T^{n-1} α, β	3 day	3 day	30 m	Same as CTRL
RSE_Po	Simulation using the parameters estimated by RPE	–	–	–	–	Aug. 31,1966, derived from RPE
RSE_Pd	Simulation using the same parameters as in CTRL	–	–	–	–	Same as RSE_Po

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

Table A1 Dependence of the sensitivity on the initial values of the parameter 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}

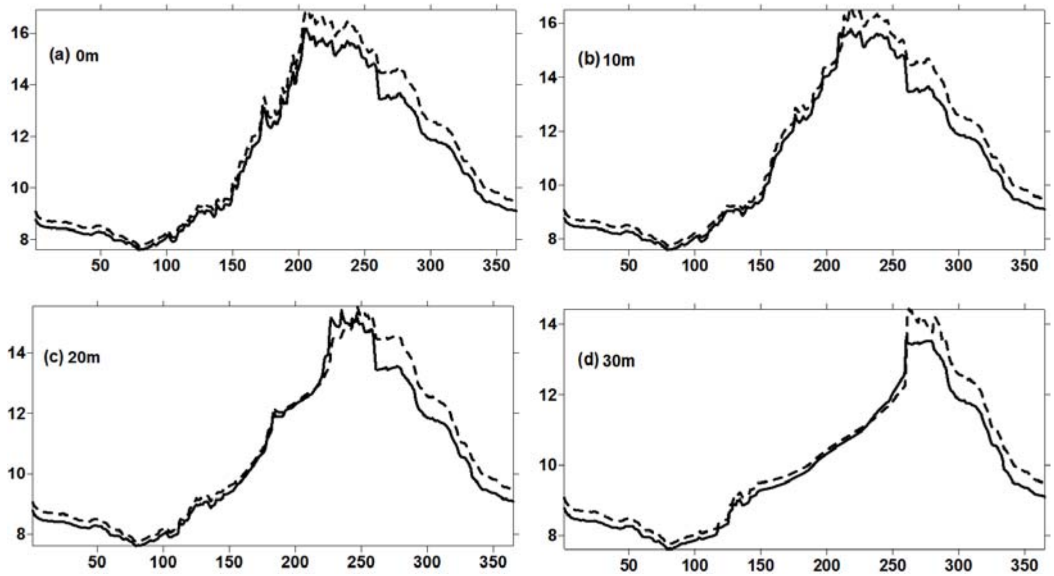

Figure1. Daily temperature in 1966 at (a) 0m, (b) 10m, (c) 20m and (d) 30m at the OWS Papa with the “truth” model simulation (solid curve) and the biased simulation (dashed curve).

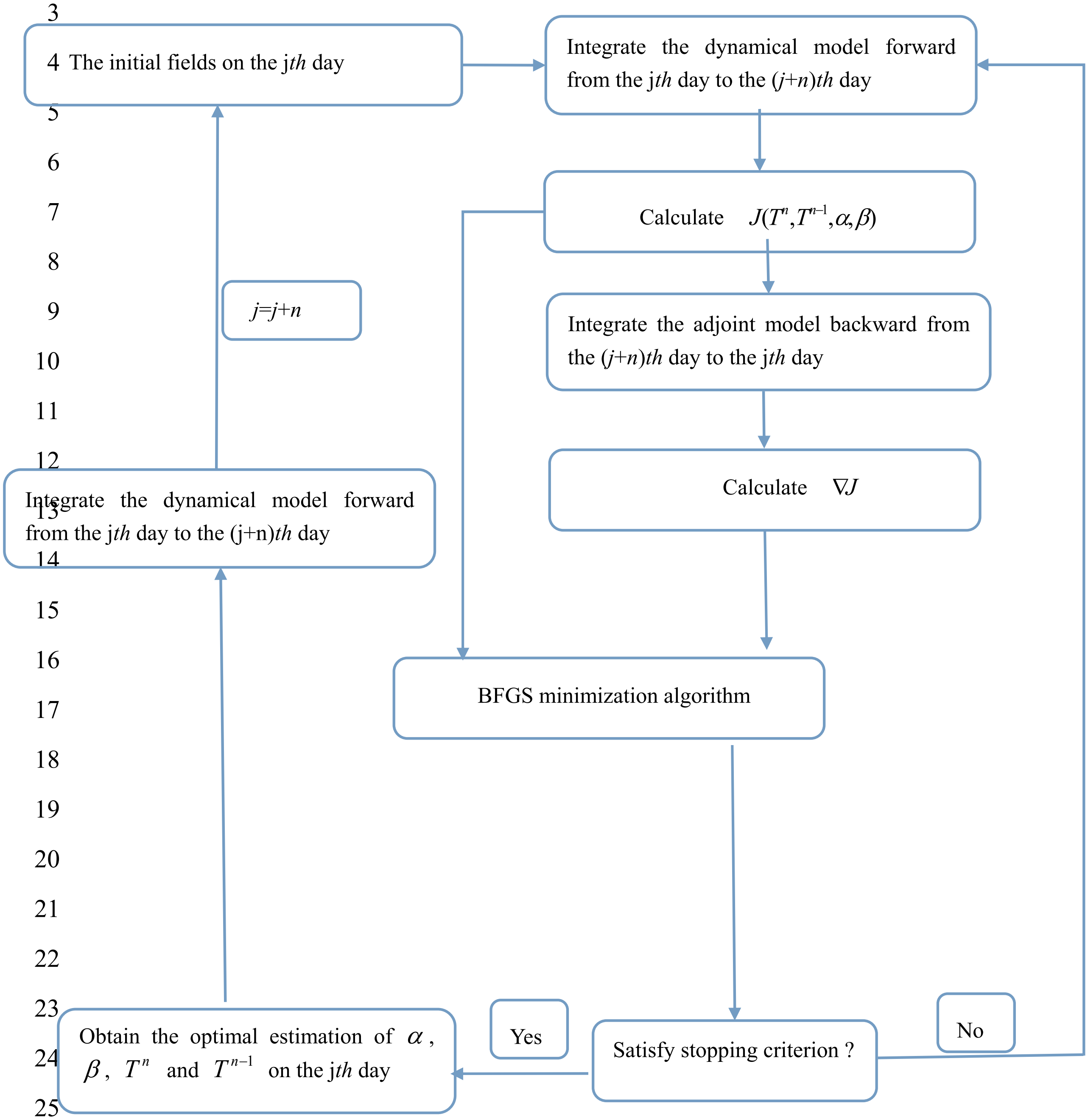

Figure 2. Flowchart of the wave-affected parameter estimation with the variational method.

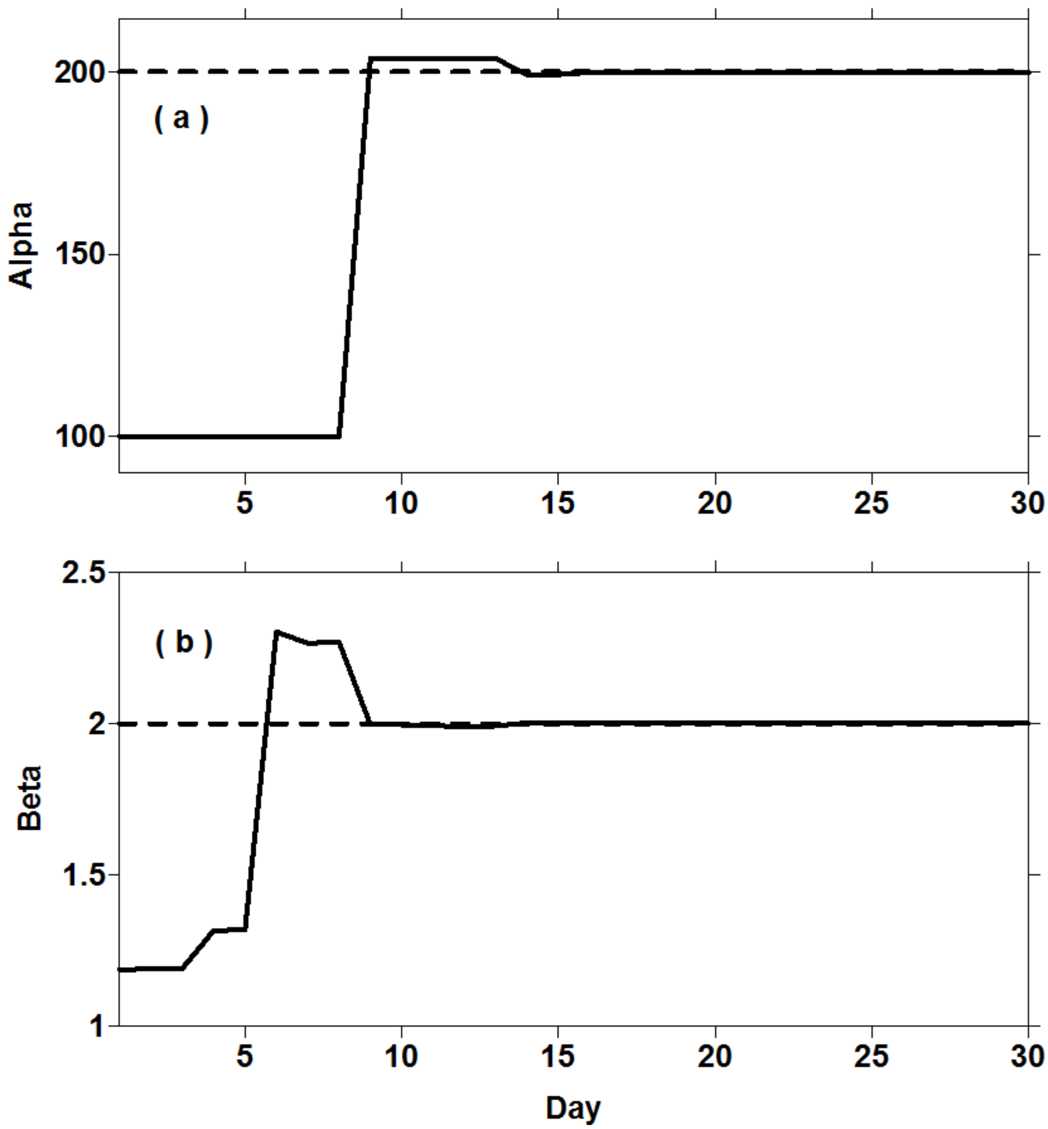

Figure 3. Time series of the estimated wave-effected parameters (a) α , and (b) β for PE from Aug.1 to Aug. 30, 1966 (solid curve), where both assimilation window and frequency are 1 day and the depth of the assimilation is 30m. The dashes curves show the "truth" (α, β) values.

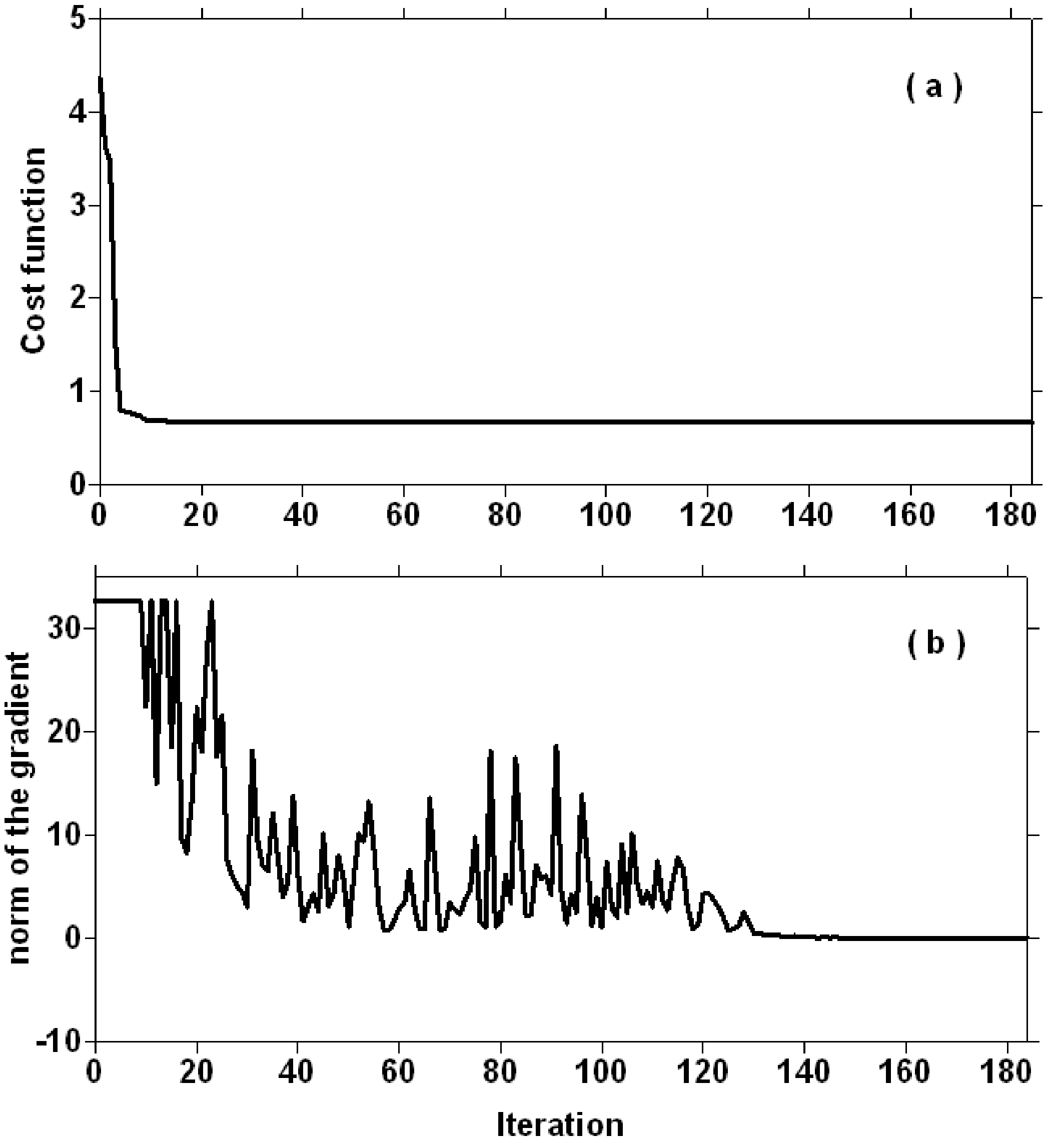

Figure 4. Dependence of (a) the cost function and (b) the norm of the gradient on the number of iterations on Aug, 2, 1966.

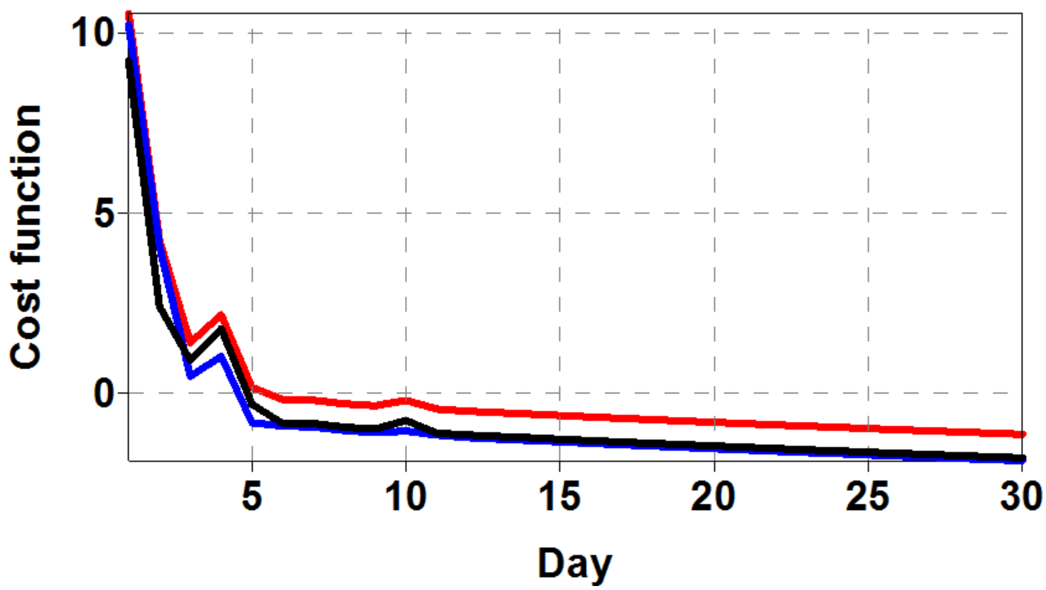

Figure 5. Temporal variations of the natural logarithm of the cost function at OWS Papa from Aug. 1 to Aug. 30, 1966. Red, blue and black lines are the total terms, the background term and the observation term of the cost function respectively.

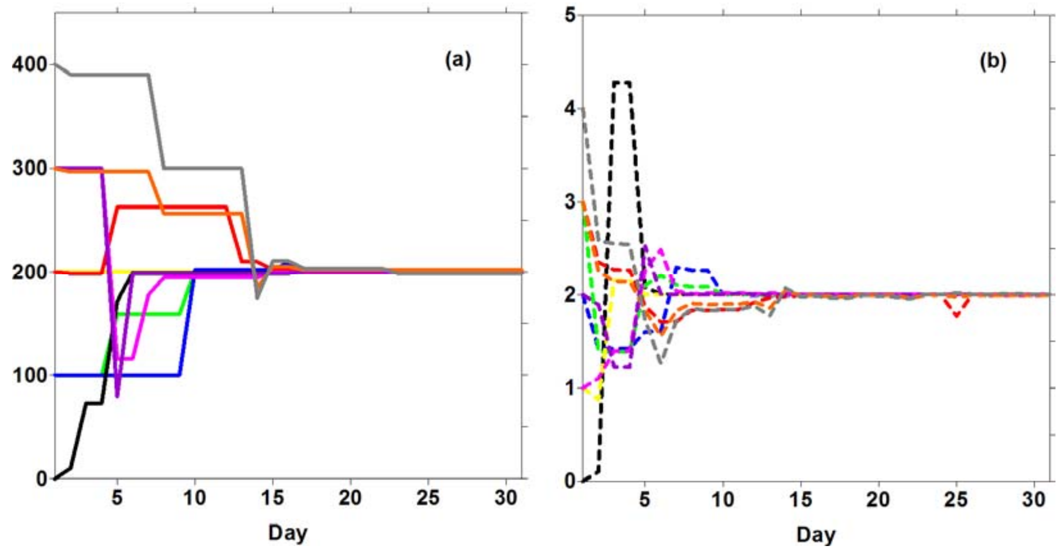

Figure 6. Time series of the estimated wave-effected parameters (a) α, and (b) β for different initial parameter values from Aug.1 to Aug. 30, 1966, where both assimilation window and frequency are 1 day and the depth of the assimilation is 30m. Black, blue, green, yellow, red, pink, purple, orange and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

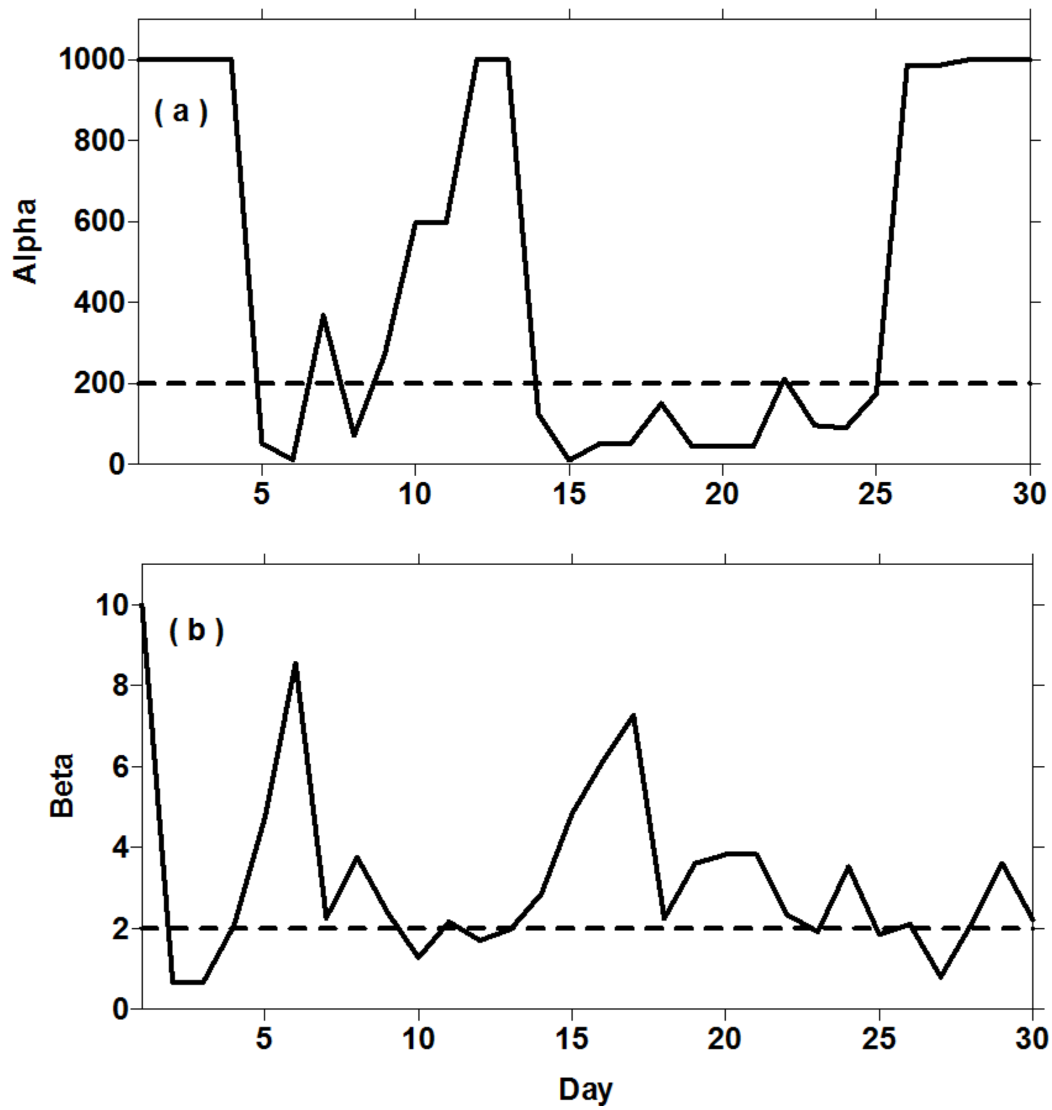

Figure 7. Same as Figure 3, but for PE_SST, where only the SST observations are assimilated.

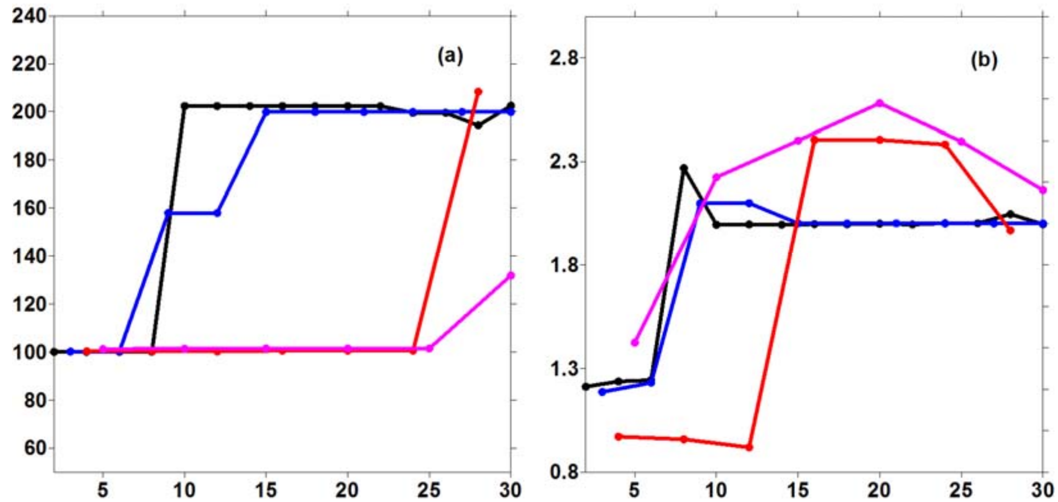

Figure 8. Time series of the estimated wave-effected parameters (a) α , and (b) β for different assimilation window and frequency from Aug.1 to Aug. 30, 1966, where the depth of the assimilation is 30m. Black, blue, red, pink solid lines in panel (a) and panel (b) show the assimilation frequency are 48, 72, 96 and 120 hours.

.

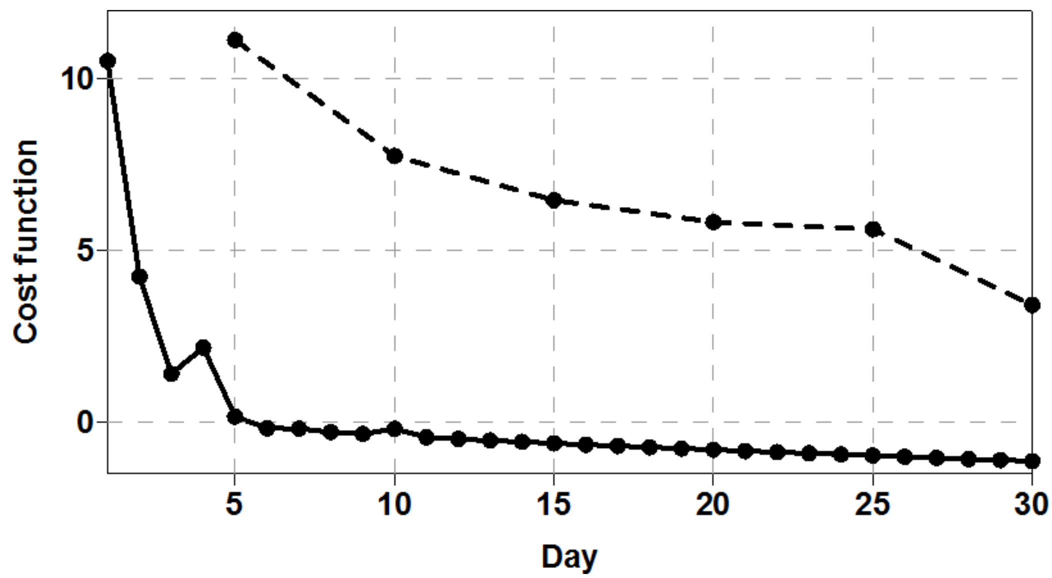

Figure 9. Temporal variations of the natural logarithm of the cost function at OWS Papa from Aug. 1 to Aug. 30, 1966. The solid and dashed curves represent PE and PE_5d with black dots denoting the time that observations are assimilated.

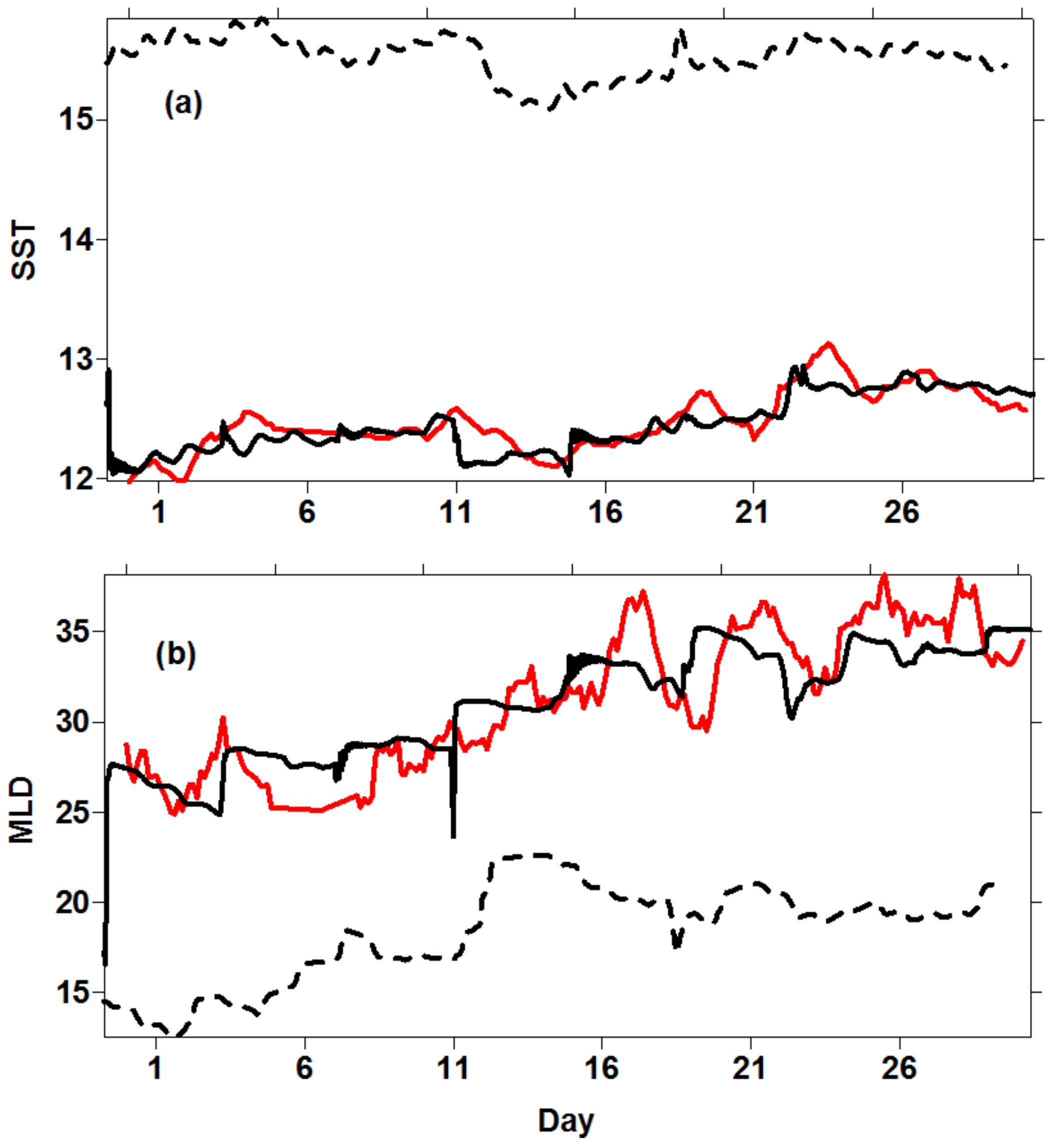

Figure 10. (a) Sea surface temperature and (b) mixed layer depth from CTRL (black dashed curve) and RPE (black solid curve), and observations (red solid curve) at OWS Papa from Aug. 1 to Aug. 30, 1966. The horizontal axis represents the day relative to Aug. 1, 1966.

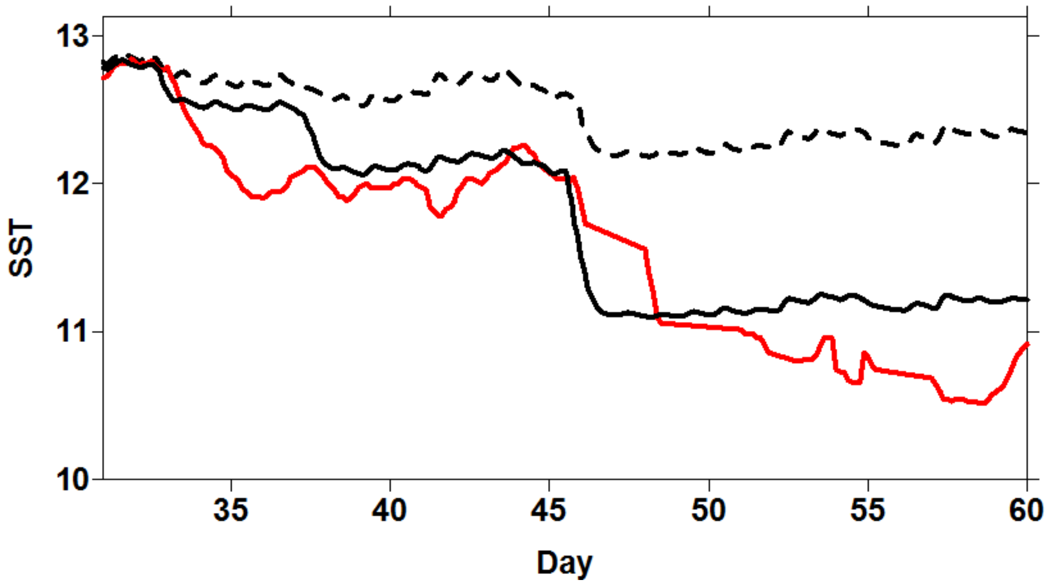

Figure 11. Sea surface temperature from observations (red solid curve), RSE_Po (black solid curve), and RSE_Pd (black dashed curve) at OWS Papa from Aug. 1 to Sept 30, 1966. The horizontal axis represents the day relative to Aug. 1, 1966.

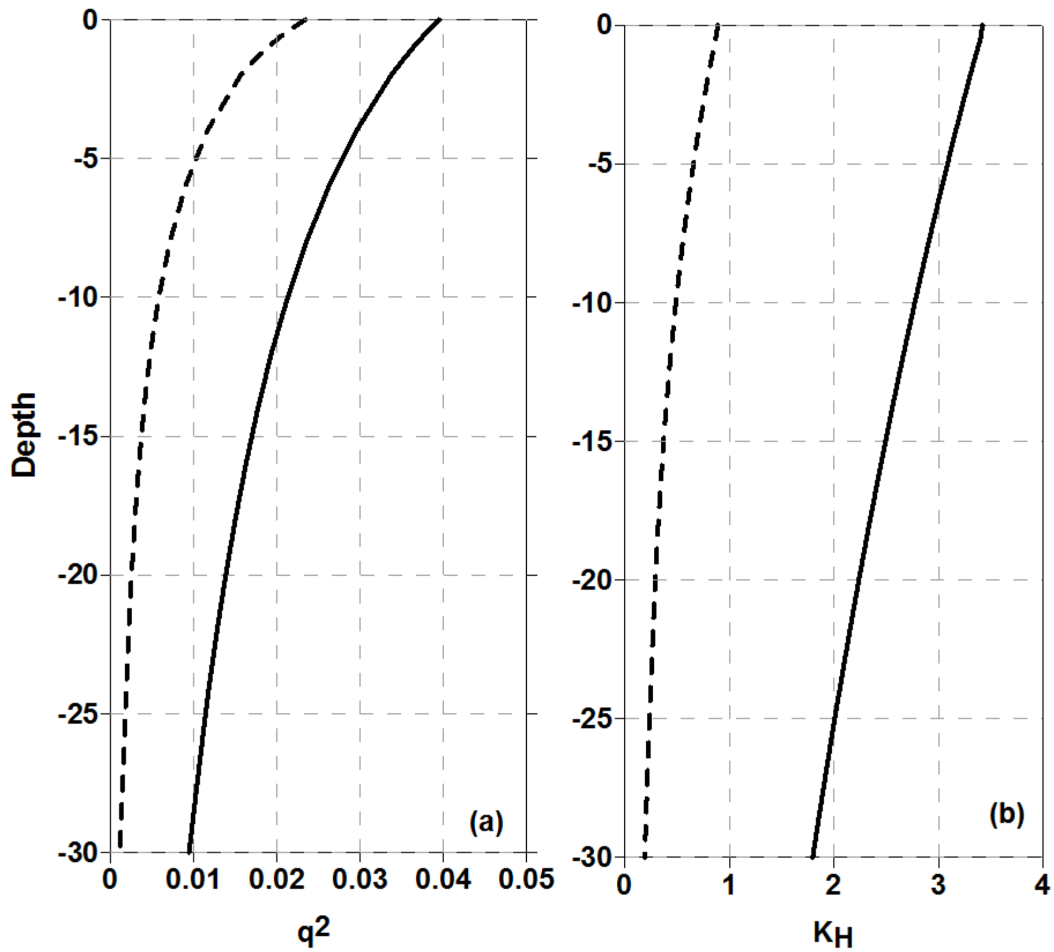

Figure 12. Vertical profiles of the simulated monthly-averaged (a) two times turbulent kinetic energy (m^2s^{-2}), and (b) and vertical mixing coefficient for temperature ($10^{-3}m^2s^{-1}$) from RSE_Po (solid curve) and RSE_Pd (dashed curve) at OWS Papa in Sept., 1966.

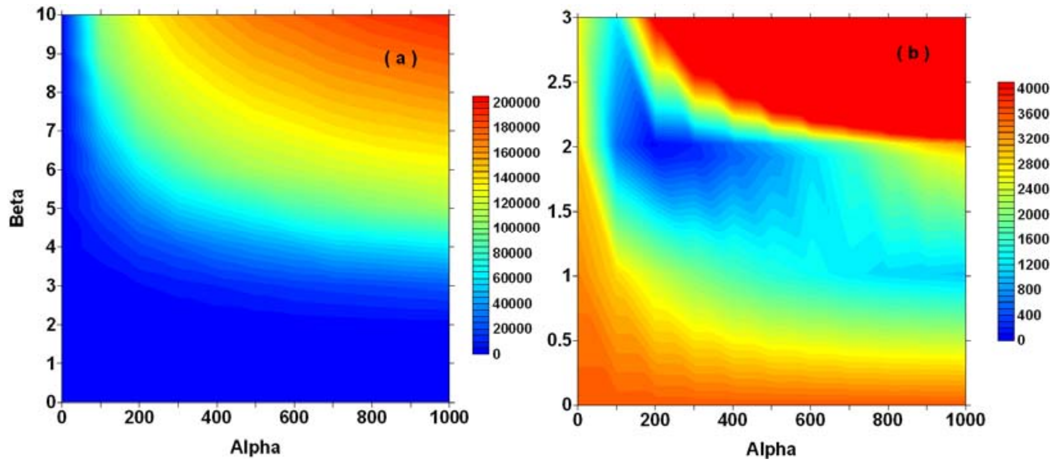

Figure A1. Dependence of the cost function on α and β with (a) $10 \geq \beta \geq 0$, and (b) $3 \geq \beta \geq 0$.

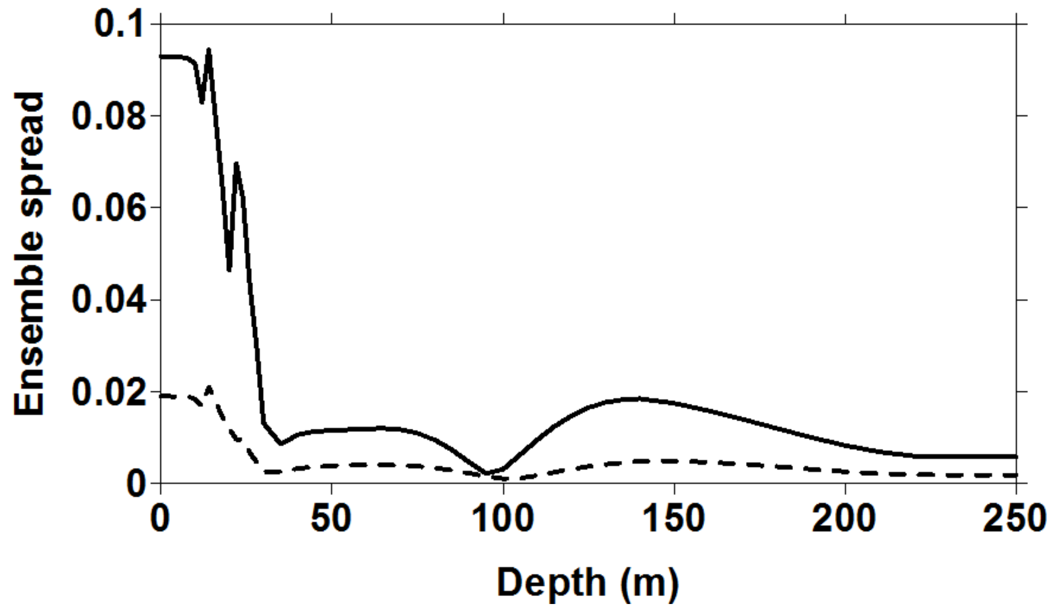

Figure A2. Ensemble spread of temperature with respect to the wave-effected parameters α (dashed curve) and β (solid curve) at different depths.

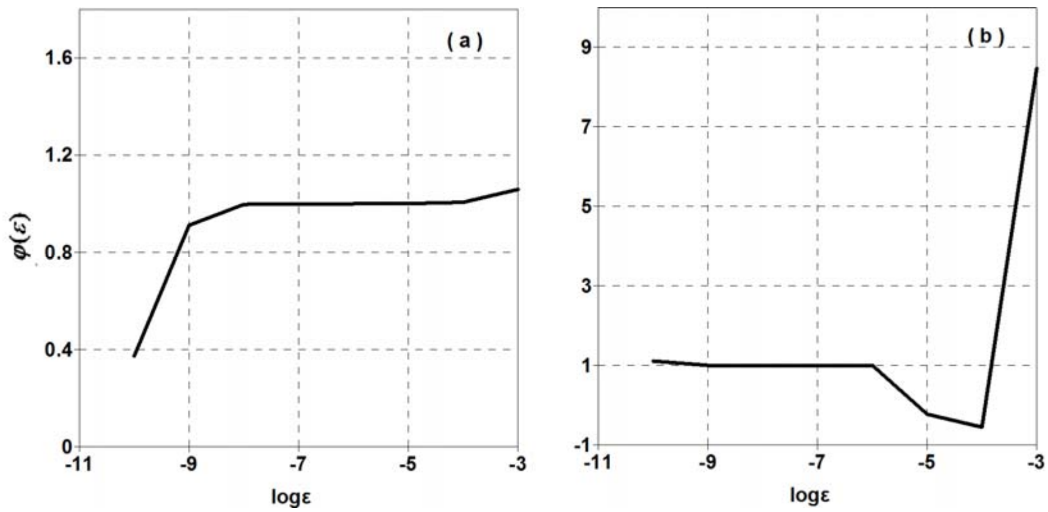

Figure A3. The correctness test of the gradient with respect to (a) α, and (b) β.