Acta Oceanol. Sin., 2024, Vol. 43, No. 3, P. 30–39
https://doi.org/10.1007/s13131-023-2229-z
http://www.aosocean.com
E-mail: ocean2@hyxb.org.cn

Application of the finite analytic numerical method to a flow-dependent variational data assimilation

Yan Hu[1], Wei Li[1], Xuefeng Zhang[2]*, Guimei Liu[3], Liang Zhang[2]

[1] Tianjin Key Laboratory for Marine Environmental Research and Service, School of Marine Science and Technology, Tianjin University, Tianjin 300072, China

[2] School of Marine Science and Technology, Tianjin University, Tianjin 300072, China

[3] National Marine Environmental Forecasting Center, Ministry of Natural Resources, Beijing 100081, China

Received 11 March 2023; accepted 6 June 2023

Abstract

An anisotropic diffusion filter can be used to model a flow-dependent background error covariance matrix, which can be achieved by solving the advection-diffusion equation. Because of the directionality of the advection term, the discrete method needs to be chosen very carefully. The finite analytic method is an alternative scheme to solve the advection-diffusion equation. As a combination of analytical and numerical methods, it not only has high calculation accuracy but also holds the characteristic of the auto upwind. To demonstrate its ability, the one-dimensional steady and unsteady advection-diffusion equation numerical examples are respectively solved by the finite analytic method. The more widely used upwind difference method is used as a control approach. The result indicates that the finite analytic method has higher accuracy than the upwind difference method. For the two-dimensional case, the finite analytic method still has a better performance. In the three-dimensional variational assimilation experiment, the finite analytic method can effectively improve analysis field accuracy, and its effect is significantly better than the upwind difference and the central difference method. Moreover, it is still a more effective solution method in the strong flow region where the advective-diffusion filter performs most prominently.

Key words: finite analytic method, advection-diffusion equation, data assimilation, flow-dependent

Citation: Hu Yan, Li Wei, Zhang Xuefeng, Liu Guimei, Zhang Liang. 2024. Application of the finite analytic numerical method to a flow-dependent variational data assimilation. Acta Oceanologica Sinica, 43(3): 30–39, doi: 10.1007/s13131-023-2229-z

1 Introduction

Advection-diffusion equation can be used to simplify the numerical simulation of many hydromechanical problems. Recently, it is extended to the variational data assimilation domain to estimate the flow-dependent background error covariance matrix (Hu et al., 2023). This new method is called the advection-diffusion filter (ADF), which is developed based on the diffusion filter. It is well known that the background error covariance matrix plays an essential role in variational data assimilation. However, it is difficult to be obtained accurately, because of insufficient statistical information on the state variables and limited computing resources. True background error covariance matrix should be anisotropic and varied with the flow field, that is, flow-dependent, but it is generally assumed to be isotropic and static for simplicity. Some scholars try to construct the flow-dependent background error covariance matrix. One of the popular ways is the diffusion filter method (Weaver and Courtier, 2001; Weaver and Mirouze, 2013; Weaver et al., 2016, 2021; Li et al., 2011; Zhang et al., 2015), it can simulate an anisotropic correlation operator by adjusting the diffusion tensors and avoids computing complicated background error covariance matrix directly. The ADF is similar to the diffusion filter, the difference is the ADF develops the flow-dependent background error covariance matrix by considering some fundamental dynamical process rather than adjusting diffusion tensors, where the advection-diffusion equation is the key of it. A large number of experiments have proved that ADF can propagate observational information anisotropically even if with a fixed diffusion coefficient. In addition, it achieved satisfactory results in the numerical forecast (Hu et al., 2023).

It is worth noting that the core of the ADF method is solving the advection-diffusion equation. In this process, it is important to choose a suitable discrete scheme because of the advection term with direction. Commonly, the finite difference method (Yan, 2006), the finite volume method (Eymard et al., 2000), and the finite element methods (Nassehi and King, 1991) are available. However, these methods are not all suitable for the ADF involving the cases in which the advection term dominates the dynamical processes. Under that scenario, the solution to the equation has a large gradient transition layer near the boundary once the central difference method (CDM) or finite element method is used, and then the numerical diffusion or non-physical numerical oscillation will occur (Rigal, 1989; Douglas and Russell, 1982; Du et al., 2000); upwind difference method (UDM) (Courant et al., 1952) was developed in the 1950s, in which the effect of the flow direction is considered to ensure that coefficients in the al-

Foundation item: The National Key Research and Development Program of China under contract Nos 2022YFC3104804, 2021YFC3101501, and 2017YFC1404103; the National Programme on Global Change and Air-Sea Interaction of China under contract No. GASI-IPOVAI-04; the National Natural Science Foundation of China under contract Nos 41876014, 41606039, and 11801402.
*Corresponding author, E-mail: xuefeng.zhang@tju.edu.cn

gebraic equations are always positive, thus the upwind scheme avoids numerical oscillation. However, the weighted coefficient of the upwind point is 1 and the downwind point is 0, the value of which cannot change with the peclet number (Feng and Zheng, 2006). In other words, this traditional upwind method lacks the "auto upwind" characteristics and over-reflects upwind conditions. It is worth noting that the upwind method of low order accuracy (Tao, 2001) can lead to the occurrence of false diffusion to further affect the accuracy of the numerical solution. To improve the calculation efficiency of the advection-diffusion equation with advection dominant, the finite analytic method (FAM) (Chen et al., 1989) based on the local analytic solution was proposed in the 1970s. In the construction of the scheme, the Taylor expansion is avoided around the singular grid nodes (Wang et al., 2014). This method has an auto-upwind mechanism, thus can provide more accurate analysis results. In this paper, the FAM is used to solve the advection-diffusion equation to improve the performance of the ADF. In the previous paper (Hu et al., 2023), the UDM is used to test the ADF, thus a comparison of the results of the ADF derived from the UDM and the FAM respectively is presented.

In the following two sections, the ADF and FAM theories are introduced, and three numerical examples are used to investigate the FAM. In Section 4, an ideal sea surface temperature (SST) assimilation experiment is conducted to evaluate the effects of FAM in flow-dependent variational assimilation. The conclusions are summarized in Section 5.

2 ADF and FAM

2.1 *ADF*

To simplify, the 1D advection-diffusion equation is introduced as follows:

$$\begin{cases} \dfrac{\partial \phi}{\partial t} + u\dfrac{\partial \phi}{\partial x} = a\dfrac{\partial^2 \phi}{\partial x^2} & 0 < x < D, 0 < t \leqslant T, \\ \phi = \omega(x) & 0 < x < D, t = 0, \\ \dfrac{\partial \phi}{\partial n} = 0 & x = 0, x = D, \end{cases} \tag{1}$$

where ϕ is the control variable, u is the flow velocity in the x direction, t is the time, n is the outer normal direction of the boundary, ω is the initial value of ϕ, a is the diffusion coefficient, $[0, D]$ is the calculation domain, T is the integral termination time. The first term on the left of the equality sign is the time-varying, the second term is the advection term, and the diffusion term is on the right of the equation. The relative proportion of advection and diffusion in a grid point can be described by Grid Peclet Number (GPN) defined as $\mathrm{pe} = \dfrac{u\Delta x}{a}$, where Δx is the spatial step. The smaller pe is, the weaker effect of the advection is. In contrast, the diffusion effect is weaker than the advection when the pe is larger.

The three-dimensional variational (3D-Var) data assimilation can be attributed to minimizing the cost function as follows:

$$J = \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_\mathrm{b})^\mathrm{T}\boldsymbol{B}^{-1}(\boldsymbol{x} - \boldsymbol{x}_\mathrm{b}) + \frac{1}{2}(\boldsymbol{H}\boldsymbol{x} - \boldsymbol{y}_\mathrm{o})^\mathrm{T}\boldsymbol{R}^{-1}(\boldsymbol{H}\boldsymbol{x} - \boldsymbol{y}_\mathrm{o}), \tag{2}$$

where $\boldsymbol{x}$ represents the estimated state vector, $\boldsymbol{x}_\mathrm{b}$ is the background vector that provides a starting point for the variational estimation based on prior knowledge, $\boldsymbol{y}_\mathrm{o}$ is the observation vector, $\boldsymbol{H}$ is a bilinear interpolation operator from the model grid to the observation space, $\boldsymbol{R}$ is the observation error covariance matrix reflecting the uncertainty of the observation, and $\boldsymbol{B}$ is the background error covariance matrix carrying the prior statistical information of errors in $\boldsymbol{x}_\mathrm{b}$. To reduce the computational expense of directly manipulating the $\boldsymbol{B}^{-1}$ of Eq. (2) and speed up the convergence of the minimization procedure, Courtier (1997) and Derber and Rosati (1989) introduced a new control variable $\boldsymbol{\omega}$, which is defined as

$$\boldsymbol{\omega} = \boldsymbol{C}^{-1}(\boldsymbol{x} - \boldsymbol{x}_\mathrm{b}), \tag{3}$$

where $\boldsymbol{B} = \boldsymbol{C}\boldsymbol{C}^\mathrm{T}$. Thus, the cost function can be rewritten as:

$$J(\boldsymbol{\omega}) = \frac{1}{2}\boldsymbol{\omega}^\mathrm{T}\boldsymbol{\omega} + \frac{1}{2}(\boldsymbol{H}\boldsymbol{C}\boldsymbol{\omega} - \boldsymbol{d})^\mathrm{T}\boldsymbol{R}^{-1}(\boldsymbol{H}\boldsymbol{C}\,\boldsymbol{\omega} - \boldsymbol{d}), \tag{4}$$

where $\boldsymbol{d} = \boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}_\mathrm{b}$ is called the "innovation" vector. The 3D-Var is a minimization problem, the gradient of the cost function is necessary for the minimization algorithms. The gradient of Eq. (4) is as follows:

$$\nabla J = \boldsymbol{\omega} + \boldsymbol{C}^\mathrm{T}\boldsymbol{H}^\mathrm{T}\boldsymbol{R}^{-1}(\boldsymbol{H}\boldsymbol{C}\boldsymbol{\omega} - \boldsymbol{d}). \tag{5}$$

In the diffusion filter method, the solution of the diffusion equation is used to model $\boldsymbol{C}\boldsymbol{\omega}$ of Eq. (4) (Weaver and Courtier, 2001). Similarly, the anisotropic diffusion filter, that is ADF, can build flow-dependent correlation by integrating Eq. (1). In other words, the analysis is converted to optimize the initial value of Eq. (1). It is worth noting that the gradient of the cost function can be obtained by the adjoint method. Although the derivation of the adjoint model is complex, the auto-adjoint compiler (tapenade (Hascoët and Pascual, 2004) or Tangent and Adjoint Model Compiler (Giering and Kaminski, 1998)) is an alternative scheme. The ADF can perform more excellently in the strong flow area, but when the advection is stronger than diffusion, that is large GPN, the choice of the discrete scheme should be more rigorous. In the previous paper, the UDM is used for discrete the advection term. The detail of UDM is omitted because it is well-known. In a word, in the UDM, the value of the interface is approximated as that of the upstream node. Although the UDM considers the direction of flow more suitable for the ADF than the CDM, the loss of accuracy caused by false diffusion is very serious. In the next content, the FAM with auto-upwind mechanism is introduced.

2.2 *FAM*

The FAM is first proposed by Chen et al. (1989), it can overcome the disadvantage that the numerical solution of the CDM is easy to oscillate or diverge when the GPN is large. In addition, the auto-upwind mechanism is also a main highlight of this method. In general, the FAM is developed based on the analytical solution of the simple advection-diffusion equation. Then, the value of the compute nodes can be obtained by the analytical solution. Finally, simultaneous the analytic solutions of each grid to obtain the algebraic equations. The detail of the FAM is followed.

For the 1D stationary advection-diffusion equation as follows:

$$u\frac{\mathrm{d}\phi}{\mathrm{d}x} = a\frac{\mathrm{d}^2\phi}{\mathrm{d}x^2} + S, \tag{6}$$

where the definition of notations is the same as Eq. (1) and S is the source item, the analytic solution is $\phi = A_0 + B_0 \cdot \exp\left(\dfrac{u}{a}S\right) + \dfrac{S}{u}x$, in which A_0 and B_0 are the unclear constants and determined by

the boundary conditions. If u, a, and S are not the constants in the entire domain and the variations are very small at the interval (x_{i-1}, x_{i+1}), which are replaced by the mean values $\overline{u}$, $\overline{a}$, and $\overline{S}$ at this interval. Then the analytic solution is converted to $\phi = A_0 + B_0 \cdot \exp\left(\frac{\overline{u}}{\overline{a}}\overline{S}\right) + \frac{\overline{S}}{\overline{u}}x$, which is regarded as the approximate solution to Eq. (6) and is called the finite analytic solution. A dimensionless coordinate $\xi = \frac{x - x_i}{\Delta x}$ is available and its relation to the original coordinate is as follows:

$$\xi = \begin{cases} 1 & x_{i+1} = x_i + \Delta x, \\ 0 & x_i, \\ -1 & x_{i-1} = x_i - \Delta x, \end{cases} \tag{7}$$

where Δx is the grid interval, x_i is the calculated node, i is the grid index. With the dimensionless coordinate, the local solution of Eq. (6) can be rewritten as

$$\phi = A_0 + B_0 \cdot \exp(\overline{p}\xi) + \frac{\overline{S}\Delta x}{\overline{u}}\xi, \tag{8}$$

where $\overline{p} = \frac{\overline{u}\Delta x}{\overline{a}}$ (GPN). Therefore, in the interval (x_{i-1}, x_{i+1}),

$$\phi_{i+1} = \phi_{\xi=1} = A_0 + B_0\exp(\overline{p}) + \frac{\overline{S}(\Delta x)}{\overline{u}}, \tag{9}$$

$$\phi_{i-1} = \phi_{\xi=-1} = A_0 + B_0\exp(-\overline{p}) - \frac{\overline{S}(\Delta x)}{\overline{u}}, \tag{10}$$

$$\phi_i = \phi_{\xi=0} = A_0 + B_0. \tag{11}$$

Eliminating A_0 and B_0, then the finite analytic solution of Eq. (6) can be written as follows:

$$\exp\left(-\frac{1}{2}\overline{p}\right)\phi_{i+1} - 2\,\mathrm{ch}\left(\frac{1}{2}\overline{p}\right)\phi_i + \exp\left(\frac{1}{2}\overline{p}\right)\phi_{i-1} = -2\frac{\overline{S}\Delta x}{\overline{u}}\mathrm{sh}\left(\frac{1}{2}\overline{p}\right). \tag{12}$$

When $\overline{p} \ll 1$ the approximate expansions are as follows:

$$\exp\left(-\frac{1}{2}\overline{p}\right) \approx 1 - \frac{1}{2}\overline{p}, \tag{13}$$

$$\exp\left(\frac{1}{2}\overline{p}\right) \approx 1 + \frac{1}{2}\overline{p}, \tag{14}$$

$$\mathrm{ch}\left(\frac{1}{2}\overline{p}\right) \approx 1, \tag{15}$$

$$\mathrm{sh}\left(\frac{1}{2}\overline{p}\right) \approx \frac{1}{2}\overline{p}. \tag{16}$$

Substituting into Eq. (12), then

$$\frac{u_i}{2\Delta x}(\phi_{i+1} - \phi_{i-1}) = \frac{a_i}{(\Delta x)^2}(\phi_{i+1} - 2\phi_i + \phi_{i-1}) + S_i. \tag{17}$$

Obviously, Eq. (17) is the discrete form of the CDM. It shows that the CDM is a special case of the FAM, that is, the CDM is the approximation of the FAM when the effect in diffusion is far greater than that in advection.

Increasing $\overline{p}$ until $\overline{p}^2 \ll 1$, then

$$\exp\left(-\frac{1}{2}\overline{p}\right) \approx 1 - \frac{1}{2}\overline{p} + \frac{1}{8}\overline{p}^2, \tag{18}$$

$$\exp\left(\frac{1}{2}\overline{p}\right) \approx 1 + \frac{1}{2}\overline{p} + \frac{1}{8}\overline{p}^2, \tag{19}$$

$$\mathrm{ch}\left(\frac{1}{2}\overline{p}\right) \approx 1 + \frac{1}{8}\overline{p}^2. \tag{20}$$

Substituting into Eq. (12), then

$$\overline{u}\left[\frac{1}{2}\left(1 + \frac{1}{4}\overline{p}\right)\frac{\phi_i - \phi_{i-1}}{\Delta x} + \frac{1}{2}\left(1 - \frac{1}{4}\overline{p}\right)\frac{\phi_{i+1} - \phi_i}{\Delta x}\right] = \frac{\overline{a}}{(\Delta x)^2}(\phi_{i+1} - 2\phi_i + \phi_{i-1}) + \overline{S}. \tag{21}$$

Considering $\alpha = \frac{1}{2}\left(1 + \frac{1}{4}\overline{p}\right), \beta = \frac{1}{2}\left(1 - \frac{1}{4}\overline{p}\right)$, where $\alpha + \beta = 1$, Eq. (21) is rewritten as

$$\overline{u}\left(\alpha\frac{\phi_i - \phi_{i-1}}{\Delta x} + \beta\frac{\phi_{i+1} - \phi_i}{\Delta x}\right) = \frac{\overline{a}}{(\Delta x)^2}(\phi_{i+1} - 2\phi_i + \phi_{i-1}) + \overline{S}. \tag{22}$$

When $\overline{u} > 0, \overline{p} > 0$, and $\alpha > \beta$; $\overline{u} < 0, \overline{p} < 0$, then $\alpha < \beta$, it can be found that $\overline{p}$ shows the characteristic of auto upwind by its size and sign.

For the 2D case with a uniform grid, the 1D finite analytic method will be implemented in the x and y direction, respectively, and then the coefficient matrix of the discretized algebraic equations becomes a five-diagonal matrix. Here the alternating direction implicit (ADI) format method is employed to resolve the algebraic equation system. In the ADI, the chasing method is used to solve the three-diagonal system in each coordinate direction. This method greatly reduces the computational workload of directly resolving the complete algebraic system (Tao, 2001).

With the FAM scheme, the 2D unsteady advection-diffusion equation can be discretized as follows:

$$\exp\left(-\frac{1}{2}\overline{p_1}\right)\phi_{i+1,j}^{n+\frac{1}{2}} - 2\,\mathrm{ch}\left(\frac{1}{2}\overline{p_1}\right)\phi_{i,j}^{n+\frac{1}{2}} + \exp\left(\frac{1}{2}\overline{p_1}\right)\phi_{i-1,j}^{n+\frac{1}{2}} + \exp\left(-\frac{1}{2}\overline{p_2}\right)\phi_{i,j+1}^{n} - 2\,\mathrm{ch}\left(\frac{1}{2}\overline{p_2}\right)\phi_{i,j}^{n} + \exp\left(\frac{1}{2}\overline{p_2}\right)\phi_{i,j-1}^{n} = 2\frac{\phi_{i,j}^{n+\frac{1}{2}} - \phi_{i,j}^{n}}{\Delta t/2}\frac{\Delta x}{\overline{u}}\mathrm{sh}\left(\frac{1}{2}\overline{p_1}\right), \tag{23}$$

$$\exp\left(-\frac{1}{2}\overline{p_2}\right)\phi_{i,j+1}^{n+1} - 2\,\mathrm{ch}\left(\frac{1}{2}\overline{p_2}\right)\phi_{i,j}^{n+1} + \exp\left(\frac{1}{2}\overline{p_2}\right)\phi_{i,j-1}^{n+1} + \exp\left(-\frac{1}{2}\overline{p_1}\right)\phi_{i+1,j}^{n+\frac{1}{2}} - 2\,\mathrm{ch}\left(\frac{1}{2}\overline{p_1}\right)\phi_{i,j}^{n+\frac{1}{2}} + \exp\left(\frac{1}{2}\overline{p_1}\right)\phi_{i-1,j}^{n+\frac{1}{2}} = 2\frac{\phi_{i,j}^{n+1} - \phi_{i,j}^{n+\frac{1}{2}}}{\Delta t/2}\frac{\Delta x}{\overline{u}}\mathrm{sh}\left(\frac{1}{2}\overline{p_1}\right), \tag{24}$$

where $n \in [0, N-1]$; $i \in [1, I-1]$; $j \in [1, J-1]$; $\overline{p_1} = \dfrac{u\Delta x}{a}$, $\overline{p_2} = \dfrac{v\Delta y}{b}$ and the solution domain is $[0, I] \times [0, J]$. i, j, and n are grid index and time index, respectively. I, J, and N are grid numbers and the total time step numbers, respectively. The other symbols are defined above.

3 Numerical examples

3.1 *1D steady advection-diffusion equation*

Consider the 1D steady advection-diffusion equation as follows:

$$\begin{cases} u\dfrac{\mathrm{d}\phi}{\mathrm{d}x} = a\dfrac{\mathrm{d}^2\phi}{\mathrm{d}x^2} + f(x) & 0 \leqslant x \leqslant 1, \\ \phi(0) = 0, \\ \phi(1) = 1, \end{cases} \tag{25}$$

the exact solution is $\phi(x) = \dfrac{f}{u}\left[x - \dfrac{\exp(ux/a) - 1}{\exp(u/a) - 1}\right]$. Where $a = 0.01$, $u = 1$, $f = 1$. The spatial step is 0.2, and the GPN is 20. From Fig. 1, one can see that the results of the FAM and the exact solution are nearly coincident, but the accuracy of UDM is lower than that of FAM. This is because the FAM is derived from the analytical solution of the steady equation, and the numerical solution of Eq. (25) that is obtained by FAM is approximate to the exact solution. However, the other solution derived from the UDM is only first-order precision, so the accuracy of UDM is lower. Therefore, in terms of calculation accuracy, the FAM is more suitable for the steady advection-diffusion equation than the UDM.

3.2 *1D unsteady advection-diffusion equation*

Consider the 1D unsteady advection-diffusion equation as follows:

$$\begin{cases} \dfrac{\partial\phi}{\partial t} + u\dfrac{\partial\phi}{\partial x} = a\dfrac{\partial^2\phi}{\partial x^2} & 0 \leqslant x \leqslant 2, \\ \phi(x, 0) = \sin(2\pi x), \\ \phi(0, t) = \exp(-4a\pi^2 t) \cdot \sin(-2\pi u t), \\ \phi(2, t) = \exp(-4a\pi^2 t) \cdot \sin[2\pi(2 - ut)], \end{cases} \tag{26}$$

the exact solution is $\phi(x, t) = \exp(-4a\pi^2 t) \cdot \sin[2\pi(x - ut)]$, where $u = 1$, $a = 0.05$, $t = 0.5$, $\Delta t = 0.025$. Figure 2 shows the exact solution and the numerical solutions derived from the UDM and the FAM, where the spatial step is 0.1 and 0.2 respectively. We can find that when $\Delta x = 0.1$, the numerical solution derived from the FAM almost coincides with the exact solution. The difference of 2-norm between them is 1.55×10^{-2}, where the computational formula of the 2-norm is $\|\phi\|_2 = \sqrt{\dfrac{\sum_{i=1}^{N} \phi_i^{\,2}}{N}}$. The other numerical solution shown in Fig. 2a is derived from the UDM, its error is larger than that of the FAM, and the difference of 2-norm between it and the exact solution is 0.16. Increasing the spatial step size, as shown in Fig. 2b, when $\Delta x = 0.2$, the effect of both numerical methods decreases, but the FAM is still better than the UDM, and the differences of 2-norm between the numerical solutions and the exact solution are 0.13 and 0.17, respectively. These results demonstrate that the FAM has more high accuracy than the UDM, and it is more suitable for solving the 1D unsteady advection-diffusion equation.

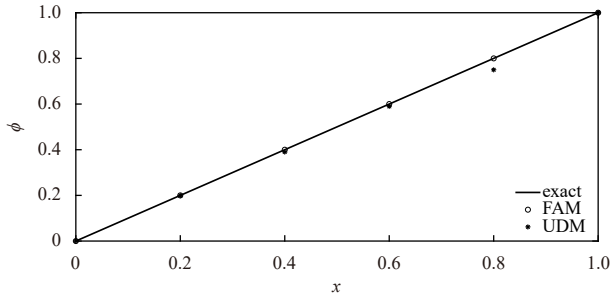

Fig. 1. The numerical solution of Eq. (25) was obtained by the upwind difference method (UDM) and the finite analytic method (FAM), and the exact solution.

3.3 *2D unsteady advection-diffusion equation*

In this paper, the FAM was originally developed to be applied to the ADF method, so the effect of the FAM used to solve the 2D unsteady advection-diffusion equation is investigated next. The numerical sample is as follows:

$$\begin{cases} \dfrac{\partial\phi}{\partial t} + u\dfrac{\partial\phi}{\partial x} + v\dfrac{\partial\phi}{\partial y} = a\dfrac{\partial^2\phi}{\partial x^2} + b\dfrac{\partial^2\phi}{\partial y^2} & 0 \leqslant x \text{ and } y \leqslant 2, \\ \phi(x, y, 0) = \dfrac{1}{4t+1}\exp\left[-\dfrac{(x-0.5)^2}{a} - \dfrac{(y-0.5)^2}{b}\right], \\ \phi(0, y, t) = \dfrac{1}{4t+1}\exp\left[-\dfrac{(-ut-0.5)^2}{a(4t+1)} - \dfrac{(y-vt-0.5)^2}{b(4t+1)}\right], \\ \phi(0, y, t) = \dfrac{1}{4t+1}\exp\left[-\dfrac{(x-ut-0.5)^2}{a(4t+1)} - \dfrac{(-vt-0.5)^2}{b(4t+1)}\right]. \end{cases} \tag{27}$$

The exact solution of the above equation is $\phi(x, y, t) = \dfrac{1}{4t+1} \times \exp\left[-\dfrac{(x-ut-0.5)^2}{a(4t+1)} - \dfrac{(y-vt-0.5)^2}{b(4t+1)}\right]$. Where $u = v = 0.8$, $a = b = 0.01$. The exact solution with $\Delta t = \dfrac{1}{160}$, $T = 1.25$ is shown in Fig. 3. The numerical solutions with the UDM and the FAM, and their error are shown in Figs 4–6, where the spatial step is 0.05×0.05, 0.02×0.02, and 0.01×0.01, respectively. Preliminary conclusions can be drawn by comparing the results of the three group experiments. When the spatial step is 0.05×0.05, the mean error of the UDM and the FAM is 0.013 and 0.003, respectively, and the accuracy of the latter is improved by 76.9% than that of

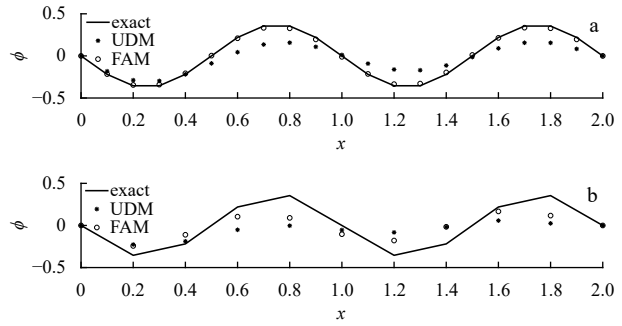

Fig. 2. The numerical solution of Eq. (26) derived from the upwind difference method (UDM) and the finite analytic method (FAM), and the exact solution, where $\Delta x = 0.1$ (a) and $\Delta x = 0.2$ (b), respectively.

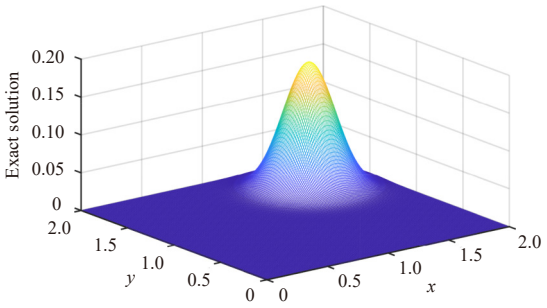

Fig. 3. The exact solution of Eq. (27).

the former. decreasing the spatial step to 0.02 × 0.02 and 0.01 × 0.01, the accuracy of the FAM improved by 67.17% and 79.41% than that of the UDM, respectively. In a word, the ability of the FAM to solve the 2D unsteady advection-diffusion equation is stronger than that of the UDM.

4 Application to SST assimilation experiment

For further accounting for the effects of the FAM in improving the quality of assimilation, a 2D ideal assimilation experiment with the ADF is carried out. In this section, the CDM and the UDM are used to solve the advection-diffusion equation involved in the ADF as contrast schemes for the FAM, respectively.

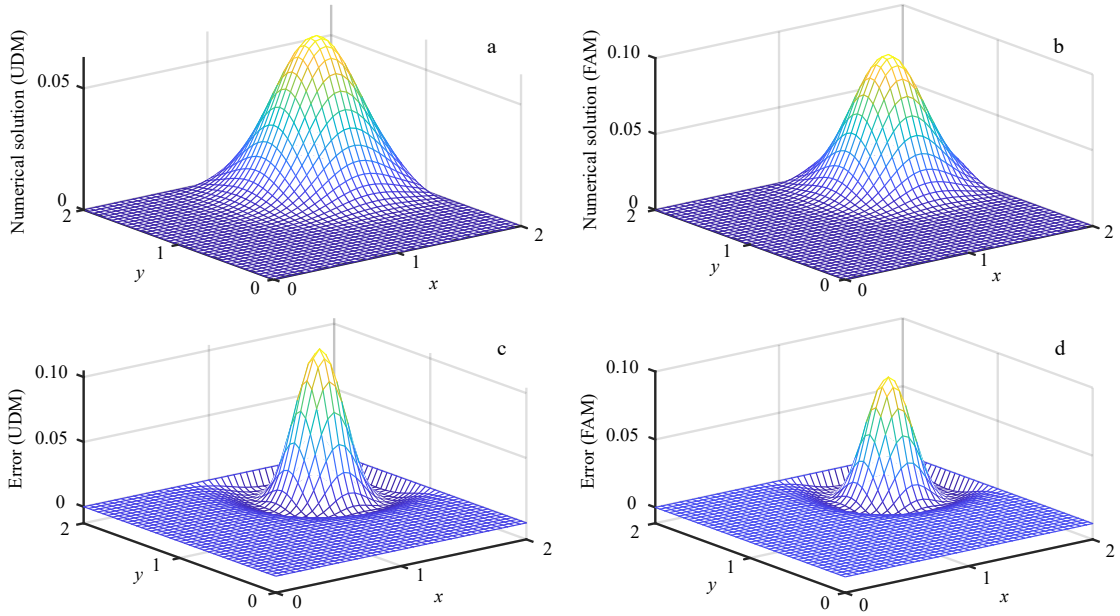

Fig. 4. The numerical solutionsof Eq. (27) obtained by the upwind difference method (UDM) (a) and the finite analytic method (FAM) (b), and their error distribution (c and d), where the spatial step is 0.05 × 0.05.

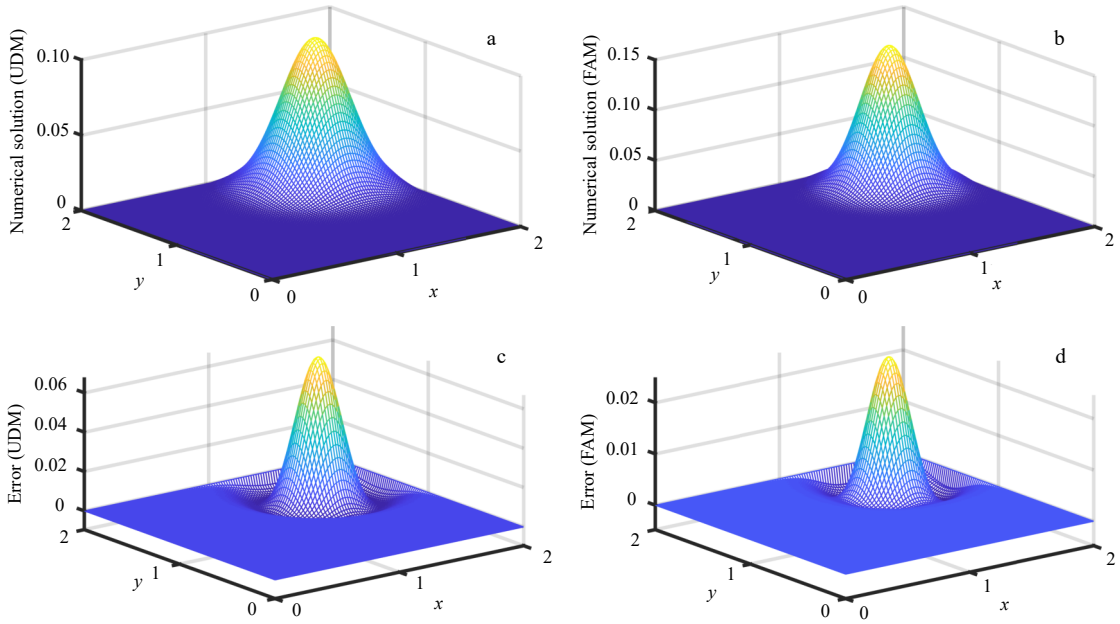

Fig. 5. The numerical solutions of Eq. (27) obtained by the upwind difference method (UDM) (a) and the finite analytic method (FAM) (b), and their error distribution (c and d), where the spatial step is 0.02 × 0.02.

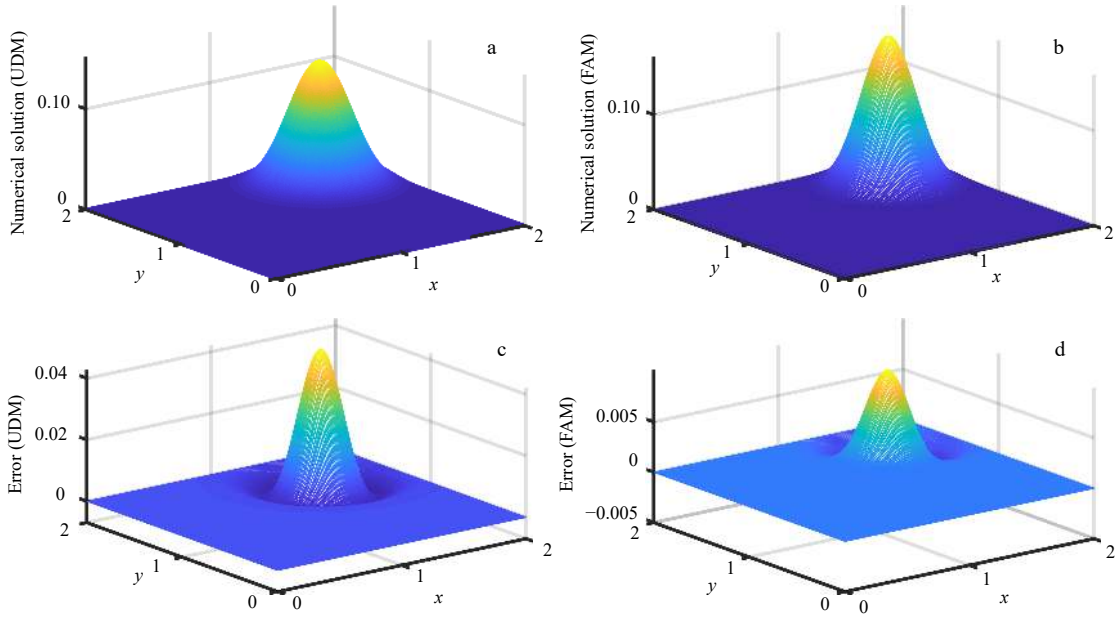

Fig. 6. The numerical solutions of Eq. (27) obtained by the upwind difference method (UDM) (a) and the finite analytic method (FAM) (b), and their error distribution (c and d), where the spatial step is 0.01 × 0.01.

4.1 *Introduction of the assimilation experiment*

The SST observation and flow velocity used in the ADF are derived from a 3D baroclinic primitive equation ocean model with a generalized coordinate system based on the Princeton Ocean Model (POMgcs) (Mellor et al., 2002; Ezer and Mellor, 2004). The main characteristics of the POMgcs include the MY-2.5 turbulent closure scheme (Blumberg and Mellor, 1987) for vertical mixing, the generalized sigma coordinate for vertical levels, curvilinear orthogonal coordinates, and an "Arakawa C" difference scheme in the horizontal direction. The POMgcs have three kinds of vertical layering ways that can be chosen: a z-coordinate, a σ-coordinate, and a σ-z hybrid coordinate (Ezer and Mellor, 2004). In this study, the hybrid method was used. More details about the POM can be found in the user guide (Mellor, 2002). The analysis area is 37.530°–37.609°N, 122.041°–122.140°E, which is Chudao Island of the Yellow Sea and an adjacent sea area. The grid horizontal resolution is 0.001° × 0.001° and there are 6 vertical levels. The maximum and minimum depths are 60 m and 5 m, respectively. The bottom topography is derived from the ETOPO1 dataset (https://www.ngdc.noaa.gov/mgg/global/relief/ETOPO1/), but the data are found to be highly questionable in the shallow regions, so it is modified by the local nautical charts to obtain a more realistic coastline. In addition, the meteorological force field of the model is from the National Centers for Environmental Prediction (NCEP), and the air-sea momentum and heat fluxes are obtained by the bulk formulas. The initial state field and lateral boundary conditions include the sea surface height, the ocean current, the temperature field, and the salt field, which are obtained from the daily mean fields and monthly fields of the China Ocean Reanalysis (CORA) (Han et al., 2013), respectively. The model spun up 10 days from March 1, 2005, and the output of the temperature on March 10 serves as the "true" state, as shown in Fig. 7a. Given that the spatial resolution of the analysis field is usually different from that of observation, selected one observation from each of the four analysis grid points. As a consequence, 600 observations were used to restore the "true" field, as shown in Fig. 7b. Figure 7c is the distribution of the flow field on March 10, 2005, which is applied to the advection-diffusion equation. The observation errors were assumed to be uncorrelated, so the matrix $\boldsymbol{R}$ was a diagonal matrix and all diagonal elements were normalized to 1.0. The background value was set as 0 and the background item was omitted temporarily. In the setting men-

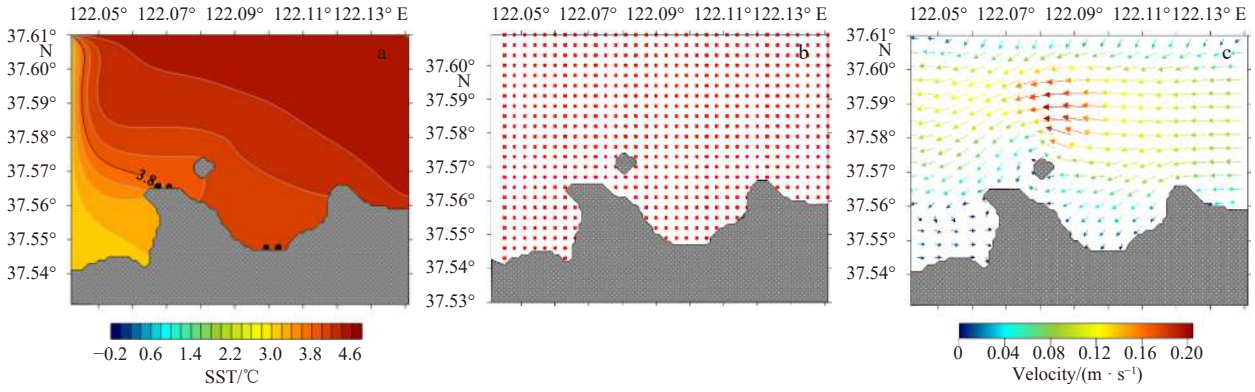

Fig. 7. The distribution of the true sea surface temperature (SST) field (a), the observations (b), and the flow field (c) used in the assimilation experiment.

tioned above, the analysis field on March 10 was formed by assimilating the pseudo-observations using the ADF. In the ADF, the diffusion coefficient needed to be set empirically. Different diffusion coefficient indicates the different correlation scales. When the same flow velocity is used, the larger the diffusion coefficient, the larger the correlation scale, and the observation information can be propagated to a further location; otherwise, the smaller the correlation scale, the closer the propagation distance. In other words, the fixed diffusion coefficient can only extract single-scale observation information. To compensate for this shortcoming, some multi-scale filter schemes (Li et al., 2011; Xie et al., 2011; He et al., 2008) using different diffusion coefficients were proposed. However, this was not the focus of this study, so the static diffusion coefficient was adopted. According to previous papers (Li et al., 2011; Xie et al., 2011; He et al., 2008), the diffusion coefficient was set within 1, so three different coefficients 0.8, 0.6, and 0.2 were used for the next experiment.

In the ADF, the discrete method of Eq. (1) was the FAM, the UDM, and the CDM, respectively. In the process of the experiment, the limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm (Liu and Nocedal, 1989) was used to minimize the cost function. The adjoint code was obtained by the tapenade compiler. The implementation detail of the ADF is shown in Fig. 8.

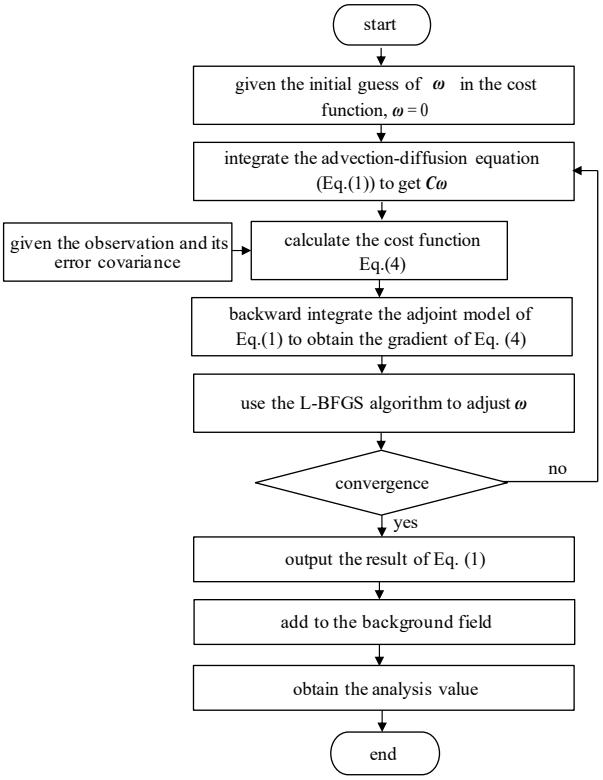

Fig. 8. The implementation flow chart of the advection-diffusion filter.

4.2 *Results analysis*

Figures 9–11 show the analysis field derived from the CDM, the UDM, and the FAM, where the diffusion coefficient of the advection-diffusion equation that is applied to the ADF is 0.8, 0.6,

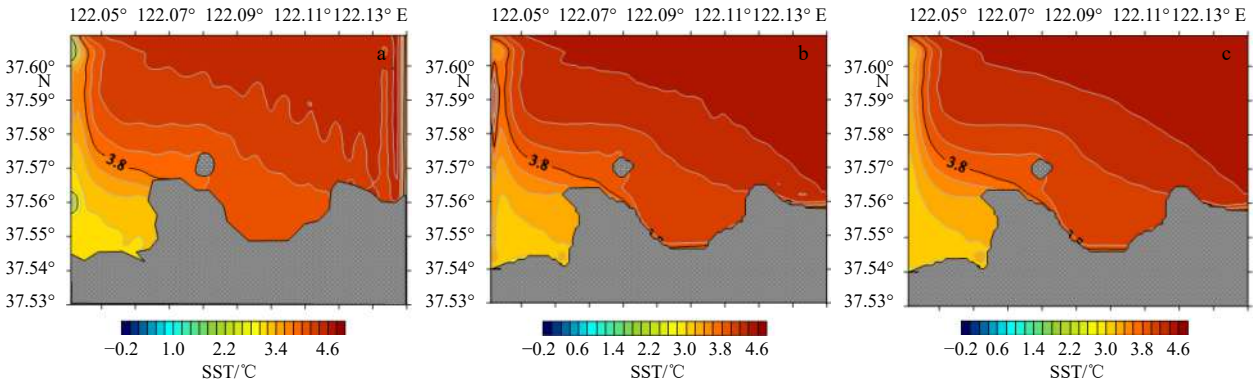

Fig. 9. The analyzed sea surface temperature (SST) field obtained by the upwind difference method (a), the central difference method (b), and the finite analytic method (c), where the diffusion coefficient is 0.8.

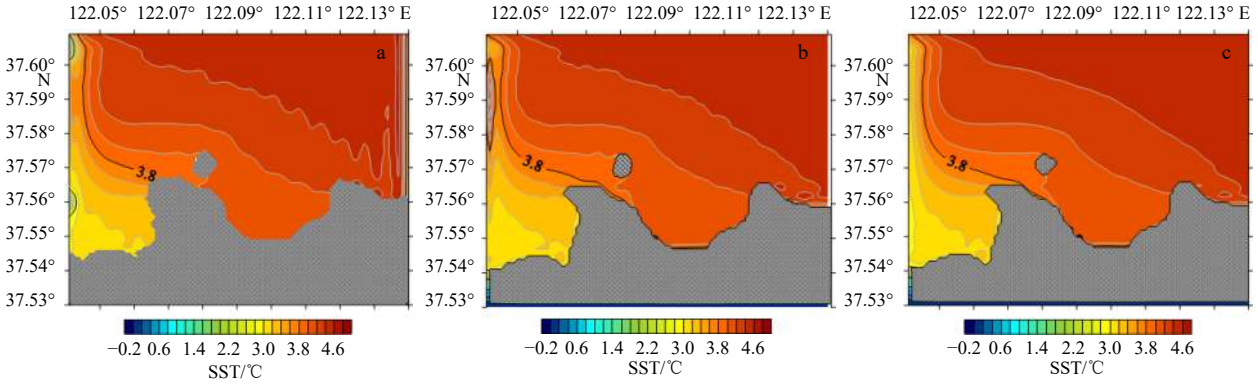

Fig. 10. The analyzed sea surface temperature (SST) field obtained by the upwind difference method (a), the central difference method (b), and the finite analytic method (c), where the diffusion coefficient is 0.6.

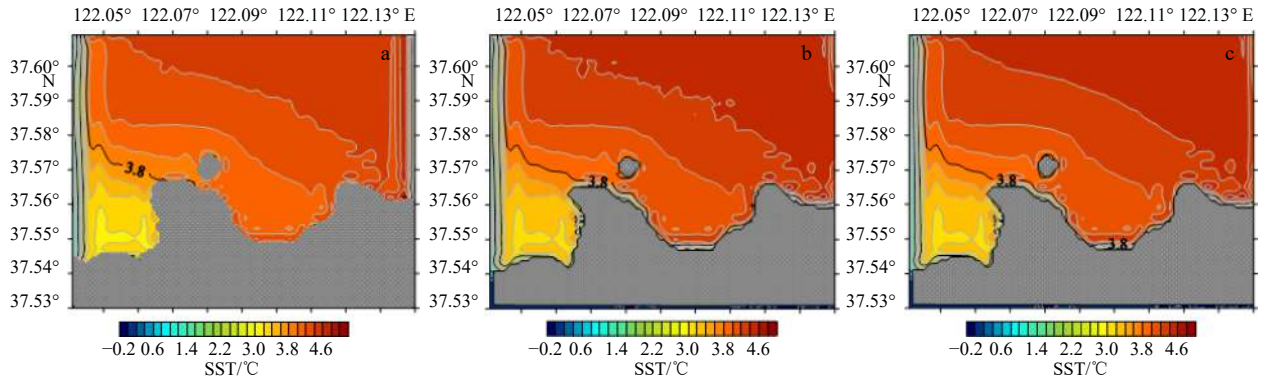

Fig. 11. The analyzed sea surface temperature (SST) field obtained by the upwind difference method (a), the central difference method (b), and the finite analytic method (c), where the diffusion coefficient is 0.2.

and 0.2, respectively. As can be seen from the three figures, the analysis field obtained by the FAM is very close to the true field, but the analysis results produced by the UDM and the CDM have a larger error than the FAM. When the diffusion coefficient is set as 0.8, the root mean square errors (RMSEs) of the FAM, the UDM, and the CDM are 0.065 4, 0.143 0, and 0.107 2, respectively. The error of the FAM is much smaller than that of the CDM and the UDM by almost 54.3% and 39.0%. Reducing the diffusion coefficient to 0.6, the RMSEs of them are 0.088 2, 0.165 9, and 0.126 3, respectively. The advantages of the FAM are still very significant, the improvement rate are 46.8% and 30.2%. When the diffusion coefficient is 0.2, the error of the analysis results is remarkable, especially in the boundary region where the observations are scarce. This is because the diffusion coefficient is too small and then observation information cannot be extracted adequately. In this case, the errors of the three schemes (UDM, CDM, and FAM) are 0.491 0, 0.454 6, and 0.468 6, respectively. The advantage of the FAM does not come through. However, such a small diffusion coefficient will not be adopted in practical application. In addition, Comparing the UDM and the CDM, one can see that the CDM outperformed the UDM, the reason is that the CDM has higher precision than the UDM. When the GPN is in the appropriate range, the result of the CDM is superior to that of the UDM. However, since the FAM is derived from the analytical solution, it is not only superior to the UDM but also to the CDM with second-order accuracy. In most cases, the ability of the FAM is more outstanding than the other two schemes.

The effect of the flow-dependent variational assimilation is especially prominent in the strong flow area, and the application results of the FAM in the strong flow area are particularly significant. Thus the analysis results over the strong flow region (37.58°–37.60°N, 122.07°–122.10°E) are shown separately, which further evaluates the FAM's capability of improving the assimilation quality, where the "true" field in this region is plotted in Fig. 12. The analysis fields with three solving methods and different diffusion coefficient are shown in Fig. 13, where the diffusion coefficient is 0.8, 0.6, and 0.2, respectively. The discrete method of the advection term from top to bottom is the UDM, the CDM, and the FAM. From the figures, we can find that the distribution of the temperature field derived from the FAM is closer to the true field than that from the other two schemes under the same diffusion coefficient, even though the diffusion coefficient is 0.2. In conclusion, the FAM is more suitable for solving the advection-diffusion equation than widely used the CDM and the UDM, thus it is an alternative method for the ADF.

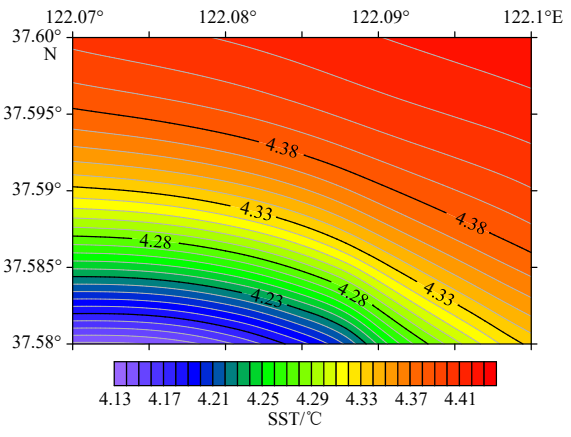

Fig. 12. The true sea surface temperature (SST) field over the strong flow area.

5 Summary

The ADF method is an extent of diffusion filter to build a flow-dependent background error covariance matrix by introducing the flow field to the diffusion equation model. In other words, the core of the ADF method is the advection-diffusion equation. However, the solution of the advection-diffusion equation is very complex, due to the strong directionality of the advection term. The CDM has two-order accuracy and is widely used in solving the diffusion term, but it is not suitable for the situation where advection is dominant, which will lead to numerical oscillation. Although the UDM avoids this deficiency, it over-reflects the upwind condition and leads to pseudo diffusion. In this paper, the FAM is used to discretize the advection-diffusion equation in ADF, which combines the numerical method and analytical method, thus the accuracy of it is very high. In addition, the FAM has the characteristic of auto upwind, which takes a more rational account of the effect of upstream and downstream nodes on the calculation nodes. In this paper, the 1D FAM is applied in the x and y direction respectively, and then ADI is used to solve the five-diagonal linear systems. Compared with traditional 2D FAM, this operation is more concise and avoids the complicated calculation process, and speeds up the calculation speed. It is characterized by steady and unsteady 1D numerical examples with different GPNs. Results indicate that the FAM has a great advantage over the UDM. The reason is that the FAM derived from the analytical solution of the steady advection-diffusion equation is more reasonable than the UDM, even though the GPN becomes large. Since the ADF involves a 2D problem, the 2D unsteady advection-diffusion equation numerical example is also evaluated, and

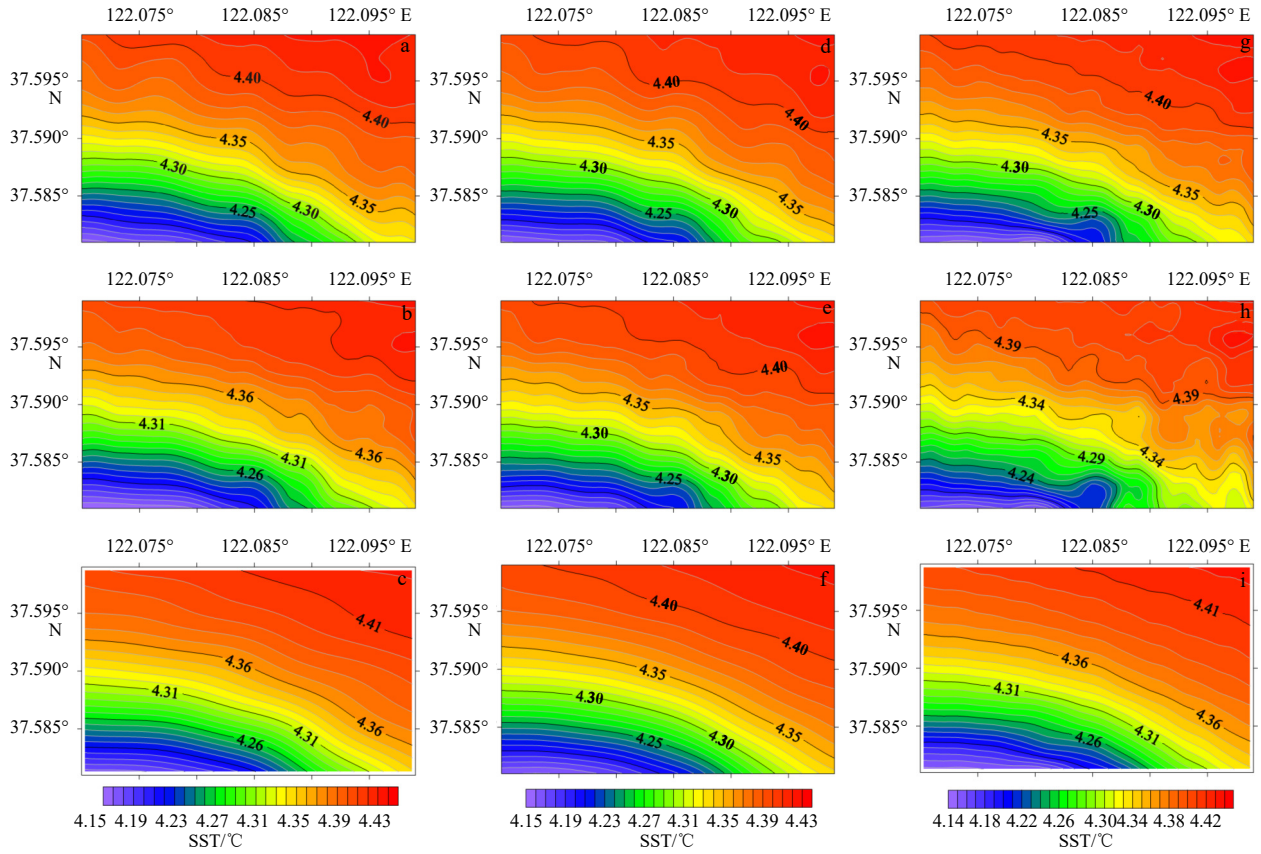

Fig. 13. The analyzed sea surface temperature (SST) fields with the upwind difference method (a, d, and g), the central difference method (b, e, and h), and the finite analytic method (c, f, and i), where the diffusion coefficient in the left is 0.8, that in the middle is 0.6, and that in the right is 0.2.

the FAM has a lower error than the UDM in the three kinds of spatial steps. After that, an assimilation experiment is designed to explore the effects of the FAM on the ADF, where the UDM and the CDM are used as control tests. The result indicates the RMSE of the analysis field derived from the FAM is the lowest among the three solving methods, even though the CDM has two-order precision and the UDM considers the direction of the flow. The effect of the ADF is more remarkable in the strong flow area, thus the performance of FAM in the strong flow area needs to be examined. The analysis field in the strong flow area is evaluated separately. The distribution of the SST field obtained by FAM is closer to the true field than the other two methods.

Since the coefficients of the algebraic equations of the FAM are exponential, the calculation speed using the finite analytic method is slightly slow. However, anisotropic diffusion filters show the potential for estimating a flow-dependent background error covariance in 3D-Var. Despite the promising results produced in the realistic application, much work is needed to explore the impact on multivariable data assimilation, longer-time forecast, and spatial multiscale variational analysis. However, the success of this research depends on the solution of accurate advection-diffusion equations.

Acknowledgement

The authors would like to express our gratitude to Wenxin Sun for his valuable suggestion.

References

Blumberg A F, Mellor G L. 1987. A description of a three-dimensional coastal ocean circulation model. In: Heaps N S, ed. Three-Dimensional Coastal Ocean Models, Volume 4. Washington, DC: American Geophysical Union, 1–16

Chen C J, Sheikholeslami M Z, Bhiladvala R B. 1989. Finite analytic numerical method for two-point boundary value problems of ordinary differential equations. Computer Methods in Applied Mechanics & Engineering, 75(1–3): 61–76

Courant R, Isaacson E, Rees M. 1952. On the solution of nonlinear hyperbolic differential equations by finite differences. Communications on Pure & Applied Mathematics, 5(3): 243–255, doi: 10.1002/cpa.3160050303

Courtier P. 1997. Variational methods. Journal of the Meteorological Society of Japan, 75(1B): 211–218

Derber J, Rosati A. 1989. A global oceanic data assimilation system. Journal of Physical Oceanography, 19(9): 1333–1347, doi: 10.1175/1520-0485(1989)019<1333:AGODAS>2.0.CO;2

Douglas J Jr, Russell T F. 1982. Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures. Siam Journal on Numerical Analysis, 19(5): 871–885, doi: 10.1137/0719063

Du Zhengping, Liu Xiaoyu, Lu Jinfu. 2000. Quadratic monotone interpolation characteristic difference method for convection-diffusion equation. Journal of Tsinghua University (Science and Technology) (in Chinese), 40(11): 1–4

Eymard R, Gallouët T, Herbin R. 2000. Finite volume methods. In: Ciarlet P G, Lions J L, eds. Handbook of Numerical Analysis. Amsterdam: Elsevier, 7: 713–1018

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(3/4): 379–403, doi: 10.1016/S1463-5003(03)00026-X

Feng Minquan, Zheng Bangmin. 2006. Auto up-wind and skew up-

wind numerical solution method of 2D convection-diffusion equation under high reynolds number. Journal of Sichuan University (Engineering Science Edition) (in Chinese), 38(6): 18–23

Giering R, Kaminski T. 1998. Recipes for adjoint code construction. ACM Transactions on Mathematical Software, 24(4): 437–474, doi: 10.1145/293686.293695

Han Guijun, Fu Hongli, Zhang Xuefeng, et al. 2013. A global ocean reanalysis product in the China Ocean Reanalysis (CORA) project. Advances in Atmospheric Sciences, 30(6): 1621–1631, doi: 10.1007/s00376-013-2198-9

Hascoët L, Pascual V. 2004. TAPENADE 2.1 User's Guide. France: National Institute for Research in Computer Science and Control

He Zhongjie, Xie Yuanfu, Li Wei, et al. 2008. Application of the sequential three-dimensional variational method to assimilating SST in a global ocean model. Journal of Atmospheric & Oceanic Technology, 25(6): 1018–1033, doi: 10.1175/2007JTECHO540.1

Hu Yan, Zhang Xuefeng, Li Dong, et al. 2023. Anisotropic diffusion filters for flow-dependent variational data assimilation of sea surface temperature. Ocean Modelling, 184: 102233, doi: 10.1016/j.ocemod.2023.102233

Li Dong, Wang Xidong, Zhang Xuefeng, et al. 2011. Multi-scale 3D-VAR based on diffusion filter. Marine Science Bulletin (in Chinese), 30(2): 164–171

Liu D C, Nocedal J. 1989. On the limited memory BFGS method for large scale optimization. Mathematical Programming, 45(1): 503–528, doi: 10.1007/BF01589116

Mellor G L. 2002. Users Guide for A Three-Dimensional, Primitive Equation, Numerical Ocean Model. Princeton: Princeton University

Mellor G L, Häkkinen S M, Ezer T, et al. 2002. A generalization of a sigma coordinate ocean model and an intercomparison of model vertical grids. In: Pinardi N, Woods J, eds. Ocean Forecasting: Conceptual Basis and Applications. Heidelberg: Springer, 55–72, doi: 10.1007/978-3-662-22648-3_4

Nassehi V, King S A. 1991. Finite element methods for the convection diffusion equation. International Journal of Engineering, 4(3): 93–100

Rigal A. 1989. Numerical analysis of two-level finite difference schemes for unsteady diffusion–convection problems. International Journal for Numerical Methods in Engineering, 28(5): 1001–1021, doi: 10.1002/nme.1620280503

Tao Wenquan. 2001. Numerical Heat Transfer (in Chinese). 2nd ed. Xi'an: Xi'an Jiaotong University Press, 135–176

Wang Yanfeng, Liu Zhifeng, Wang Xiaohong. 2014. Finite analytic numerical method for three-dimensional fluid flow in heterogeneous porous media. Journal of Computational Physics, 278: 169–181, doi: 10.1016/j.jcp.2014.08.026

Weaver A, Courtier P. 2001. Correlation modelling on the sphere using a generalized diffusion equation. Quarterly Journal of the Royal Meteorological Society, 127(575): 1815–1846, doi: 10.1002/qj.49712757518

Weaver A T, Chrust M, Ménétrier B, et al. 2021. An evaluation of methods for normalizing diffusion-based covariance operators in variational data assimilation. Quarterly Journal of the Royal Meteorological Society, 147(734): 289–320, doi: 10.1002/qj.3918

Weaver A T, Mirouze I. 2013. On the diffusion equation and its application to isotropic and anisotropic correlation modelling in variational assimilation. Quarterly Journal of the Royal Meteorological Society, 139(670): 242–260, doi: 10.1002/qj.1955

Weaver A T, Tshimanga J, Piacentini A. 2016. Correlation operators based on an implicitly formulated diffusion equation solved with the Chebyshev iteration. Quarterly Journal of the Royal Meteorological Society, 142(694): 455–471, doi: 10.1002/qj.2664

Xie Yuanfu, Koch S, Mcginley J, et al. 2011. A space-time multiscale analysis system: A sequential variational analysis approach. Monthly Weather Review, 139(4): 1224–1240, doi: 10.1175/2010MWR3338.1

Yan Chao. 2006. Computational Fluid Mechanics Methods and Applications (in Chinese). Beijing: Beijing University of Aeronautics and Astronautics Press, 1–266

Zhang Xuefeng, Li Dong, Chu P C, et al. 2015. Diffusion filters for variational data assimilation of sea surface temperature in an intermediate climate model. Advances in Meteorology, 2015: 751404, doi: 10.1155/2015/751404