

武浩文, 赵艳玲, 韩桂军, 等. 集合调整卡尔曼滤波方法在 M_2 分潮数值模拟中的水深估计研究[J]. 海洋学报, 2022, 44(6): 10–21, doi:10.12284/hyxb2022057

Wu Haowen, Zhao Yanling, Han Guijun, et al. Bathymetry estimation using ensemble adjustment Kalman filter in the numerical simulation of M_2 constituent[J]. Haiyang Xuebao, 2022, 44(6): 10–21, doi:10.12284/hyxb2022057

集合调整卡尔曼滤波方法在 M_2 分潮 数值模拟中的水深估计研究

武浩文¹, 赵艳玲², 韩桂军^{1*}, 李威^{1*}, 曹力戈¹, 武晓博¹,
李超亮¹, 李云东¹, 周功赋¹

(1. 天津大学海洋科学与技术学院, 天津 300072; 2. 中国人民解放军 31010 部队, 北京 100081)

摘要: 数据同化利用观测信息对模型状态场调整的同时也可以对数值模型中的不确定参数进行估计, 从而改进数值模型, 提高数值模拟的精度。本文基于集合调整卡尔曼滤波方法, 采用广义坐标系统的美国普林斯顿大学海洋模式的外模式开展了渤海和部分黄海海域 M_2 分潮模拟中的水深估计研究。理想数据同化试验结果表明, 集合调整卡尔曼滤波方法能很好地降低模式模拟的水位误差并反演出“真实”的水深参数。而在 NAO.99Jb 和验潮站数据的实际数据同化试验中, 与验潮站数据相比较, 水深参数估计后, 模式模拟的 M_2 分潮振幅与迟角误差分别降低了 40.27% 和 49.19%。

关键词: 数据同化; 集合调整卡尔曼滤波; 数值模拟; 渤海; 黄海; M_2 分潮; 水深估计

中图分类号: P731.23

文献标志码: A

文章编号: 0253-4193(2022)06-0010-12

1 引言

随着计算机技术的不断发展, 海洋数值模型越来越成为人们研究和预测海洋的重要工具。对于所有的海洋数值模式而言, 参数化方案以及精确的参数值的给定对数值模拟结果有着重要的影响。在海洋数值模式调试的过程中, 参数值的给定通常采用试错的方法, 以使模拟结果接近观测。由于海洋数值模型的复杂性, 这是一个极其耗费计算资源与人力资源的过程。因此, 替代这种主观调试方法的途径之一是采用海洋数据同化技术, 通过将观测数据同化到海洋数值模型中, 在模型状态场调整的同时, 对参数进行估计, 从而得到一个合理的参数值。这样做的同时, 也可以在一定程度上得以缓解试错法在海洋数值模型调试中所带来的繁重任务。利用基于伴随模型的四维变

分同化方法与集合卡尔曼滤波方法均可以开展海洋数值模型中参数的最优估计。

以潮汐潮流的数值模拟为例, 模型中不确定的参数包括水深、低摩擦系数和开边界条件等。此前, 研究者们大多采用伴随方法来进行潮汐潮流模拟中这些参数的优化, 从早期基于线性浅水方程模型的理想数据同化试验^[1-4], 到后来基于非线性二维和三维潮汐数值模型的实际数据同化试验^[5-12], 均取得了很好的研究成果。相比较而言, 基于集合卡尔曼滤波方法开展潮汐潮流模拟中的参数估计研究尚较少。Ngodock 等^[13]利用状态增广集合卡尔曼滤波(Augmented State Ensemble Kalman Filter, ASEnKF)对一个全球三维海洋环流模型的潮汐外强迫进行修正。Wilson 和 Özkan-Haller^[14]、Wilson 等^[15-16]、Landon 等^[17]和 Moghimi 等^[18]利用基于集合的数据同化方法对近岸、河流和

收稿日期: 2021-06-22; 修订日期: 2021-09-03。

基金项目: 国家自然科学基金(41876014)。

作者简介: 武浩文(1997—), 男, 天津市人, 主要从事海洋数据同化研究。E-mail: 2955178840@qq.com

* 通信作者: 韩桂军(1970—), 女, 辽宁省新民市人, 教授, 主要从事海洋分析与预报研究。E-mail: guijun_han@tju.edu.cn; 李威(1978—), 男, 天津市人, 教授, 主要从事海洋数据同化、海洋数值分析和预报研究。E-mail: liwei1978@tju.edu.cn

潮汐汉道的水深参数进行最优估计。类似本文所开展的基于集合卡尔曼滤波方法对渤海和黄海这样的陆架浅海M₂分潮数值模拟中的水深参数进行估计研究尚鲜有报道。

本文所关注的渤海和黄海,由于其独特的地理位置和海底地形,使其潮波运动独具特征。前人的研究表明,黄海的潮能消耗约占本海区总能量的80%以上^[19-20],潮汐、潮流成为本海区海水运动最重要的过程。因此,在上述海域建立高精度的潮汐潮流模拟数值模型有着重要意义。本研究拟采用集合卡尔曼滤波方法,以集中反映该海域潮波运动特征的优势分潮——M₂分潮为例,对数值模型构建过程中所涉及的不确定参数——水深进行最优估计,以期获得更好的M₂分潮模拟结果。

2 数据和方法

2.1 数据

本研究的海区范围为35°~41°N, 117°~127°E,包括渤海和黄海的大部分海域(图1)。数值模型所使用的水深数据来自美国国家地球物理数据中心(U.S.National Geophysical Data Center, NGDC)的ETOPO2(<https://www.ngdc.noaa.gov/mgg/global/etopo2.html>),其分辨率为(1/30)°×(1/30)°。研究海区内M₂分潮网格化的潮汐调和常数来自日本国家天文台的NAO.99Jb数据^[21](https://www.miz.nao.ac.jp/staffs/nao99/index_En.html),其分辨率为(1/12)°×(1/12)°;南部开边界处M₂分潮的潮流调和常数来自美国俄勒冈州立大学的

TPXO9数据^[22](<https://tpxows.azurewebsites.net/>),其分辨率为(1/12)°×(1/12)°。由于NAO.99Jb数据在中国沿岸海域有较高的精度^[23],因此本研究将其作为数据同化试验中的观测数据来使用。此外,本研究还使用了研究海域内34个验潮站的M₂分潮调和常数^[6],其位置和编号如图1所示。

2.2 海洋数值模式

本研究用于关注海域潮汐潮流模拟的数值模型为广义坐标系统的美国普林斯顿大学海洋模式(Princeton Ocean Model with Generalized Coordinate System, POMgcs)^[24]的外模式,其动力学基本方程组为

$$\frac{\partial \zeta}{\partial t} + \frac{\partial UD}{\partial x} + \frac{\partial VD}{\partial y} = 0, \quad (1)$$

$$\begin{aligned} \frac{\partial UD}{\partial t} + \frac{\partial U^2 D}{\partial x} + \frac{\partial UV D}{\partial y} - fVD = -gD \frac{\partial \zeta}{\partial x} + \\ \frac{\rho_a C_d W_x \sqrt{W_x^2 + W_y^2}}{\rho_w} - C_b U \sqrt{U^2 + V^2} + \\ \frac{\partial}{\partial x} \left(H \times 2A_M \frac{\partial U}{\partial x} \right) + \frac{\partial}{\partial y} \left[HA_M \left(\frac{\partial U}{\partial y} + \frac{\partial V}{\partial x} \right) \right], \end{aligned} \quad (2)$$

$$\begin{aligned} \frac{\partial VD}{\partial t} + \frac{\partial UV D}{\partial x} + \frac{\partial V^2 D}{\partial y} + fUD = -gD \frac{\partial \zeta}{\partial y} + \\ \frac{\rho_a C_d W_y \sqrt{W_x^2 + W_y^2}}{\rho_w} - C_b V \sqrt{U^2 + V^2} + \\ \frac{\partial}{\partial y} \left(H^2 \times A_M \frac{\partial V}{\partial y} \right) + \frac{\partial}{\partial x} \left[HA_M \left(\frac{\partial U}{\partial y} + \frac{\partial V}{\partial x} \right) \right], \end{aligned} \quad (3)$$

式中, t 为时间; x 和 y 为水平方向的笛卡尔坐标; U 和 V 为水平方向的流速; $D = H + \zeta$,其中 ζ 为水位, H 为水

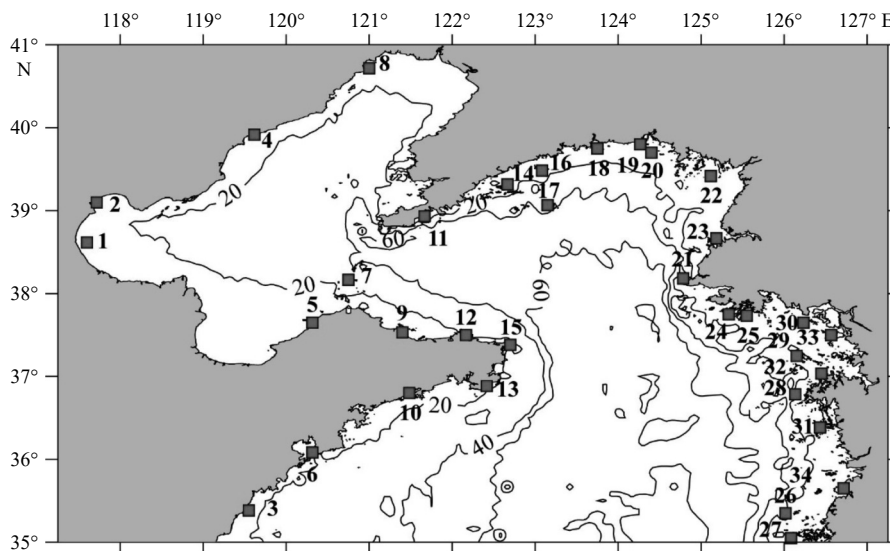


图1 研究海域范围及各验潮站位置和编号

Fig. 1 Study area and locations of tide gauges with numbers assigned

黑色实线分别为20 m、40 m和60 m等深线

The black solid lines indicate the isobaths of 20 m, 40 m, and 60 m

深; f 为科里奥利参数; g 为重力加速度; ρ_a 和 ρ_w 为空气和水的密度; C_d 为风应力拖曳系数; W_x 和 W_y 为水平方向的风速; C_b 为底摩擦系数; A_M 为水平涡动系数。

在本研究中, 数值模型的空间网格分辨率为 $(1/30)^\circ$, 时间步长为 12 min, 初始条件取为 0 值。底摩擦系数在整个海区取为常数 1.0×10^{-3} 。研究海域南部开边界处的水位和流速按下式给定:

$$Y = A \cos(\omega t - \varphi + V + u), \quad (4)$$

式中, Y 为水位或流速; A 为振幅; ω 为频率; φ 为迟角; V 和 u 为天文初相角。

2.3 集合卡尔曼滤波方法

本研究所采用的集合卡尔曼滤波方法是由 Anderson^[25] 提出的确定性方法, 即基于最小二乘框架的集合调整卡尔曼滤波 (Ensemble Adjustment Kalman Filter, EAKF)。对某一观测 y^o 而言, 该方法的实施包括如下两个步骤:

$$r_{x,y} = \Omega(a,b) = \begin{cases} -\frac{1}{4} \left(\frac{b}{a}\right)^5 + \frac{1}{2} \left(\frac{b}{a}\right)^4 + \frac{5}{8} \left(\frac{b}{a}\right)^3 - \frac{5}{3} \left(\frac{b}{a}\right)^2 + 1, & 0 \leq b \leq a \\ \frac{1}{12} \left(\frac{b}{a}\right)^5 - \frac{1}{2} \left(\frac{b}{a}\right)^4 + \frac{5}{8} \left(\frac{b}{a}\right)^3 + \frac{5}{3} \left(\frac{b}{a}\right)^2 - 5 \left(\frac{b}{a}\right) + 4 - \frac{2}{3} \left(\frac{b}{a}\right)^{-1}, & a < b \leq 2a \\ 0, & b > 2a \end{cases}, \quad (7)$$

式中, a 为经验给定的局地化半径, b 为观测点与状态变量格点之间的距离。

在数据同化过程中, 由于分析集合的不确定性 (标准差或集合离散度) 总是小于先验集合的不确定性, 随着同化的持续进行, 分析集合的离散度逐渐减小。但由于不可避免的模式偏差的存在, 使得所刻画先验场的不确定性小于实际, 高估了模式的准确性, 从而导致同化作用偏弱, 甚至引起滤波发散。因此, 对于状态估计, 本研究引入静态乘法膨胀方案, 即确定一个常数膨胀因子, 对各个集合成员相对于集合平均的扰动进行膨胀, 用以调整集合的离散度, 从而避免集合离散度降低, 导致滤波发散。而对于参数估计, 则引入条件静态膨胀方案^[27-28], 即通过判断当前时刻集合方差的大小, 决定是否对其进行参数膨胀。参数的条件静态乘法膨胀公式如下:

$$\tilde{\beta} = \bar{\beta} + \max\left(1, \frac{\alpha_0 \sigma_0}{\sigma_t}\right) (\beta - \bar{\beta}), \quad (8)$$

式中, β 、 $\tilde{\beta}$ 为膨胀前、后的参数集合; $\bar{\beta}$ 为参数的集合平均; α_0 为根据经验设定的参数膨胀系数; σ_0 和 σ_t 分别为初始集合和 t 时刻集合的标准差。

此外, 为了进一步提高 EAKF 参数估计的效果, 在后续实际数据试验中采用了强化的参数校正数据同化方法 (Data Assimilation Scheme for Enhance Para-

第 1 步, 计算观测增量:

$$\Delta y_k^o = \left(\sqrt{\frac{(\sigma^o)^2}{(\sigma^o)^2 + (\sigma_y^o)^2}} - 1 \right) (y_k^o - \bar{y}^o) + \frac{(\sigma_y^o)^2}{(\sigma^o)^2 + (\sigma_y^o)^2} (y^o - \bar{y}^o), \quad (5)$$

式中, Δy_k^o 为第 k 个集合成员在观测点 y_k 上的观测增量; σ^o 为观测误差的标准差; σ_y^o 为集合在该观测点上的先验 (同化前) 误差标准差; y_k^o 为第 k 个集合成员在该观测点上的先验值; \bar{y}^o 为投影到该观测点上的集合平均值。

第 2 步, 将观测增量投影到模式网格点上, 获得集合成员的更新值:

$$x_k^a = x_k^p + r_{x,y} \frac{cov_{x,y}^p}{(\sigma_y^p)^2} \Delta y_k^o, \quad (6)$$

式中, x_k^a 为第 i 个后验 (同化后) 集合成员; x_k^p 为第 i 个先验集合成员; $cov_{x,y}^p$ 为状态集合与观测之间的协方差; $r_{x,y}$ 为局地化因子, 采用 Gaspari-Cohn 函数形式^[26]:

meter Correction, DAEPC)^[29]。该方法的实现方式是在模型状态场的估计达到准平衡态后, 再启动对模型参数的估计。这是因为在数据同化的初期, 观测与状态变量之间的协方差由状态误差主导; 在进行一段时间的模型状态估计后, 状态误差降低, 参数误差则成为模式误差的主要成因, 即观测与状态变量之间的协方差反映的主要是由参数主导的信号, 因此, 采用 DAEPC 方法有助于强化参数调整过程中的信噪比, 从而提高参数估计的效果。

3 数据同化方案设计

3.1 敏感性分析

Han 等^[8] 研究表明, 相较于底摩擦系数和开边界条件, 水深对潮汐潮流模拟结果有着较大的影响。在本研究进行水深参数估计前, 首先开展水深参数的敏感性分析试验。按照 Han 等^[8] 的方法, 将本文研究海域按照水深进行划分。经大量测试, 将区域按照水深划分为 0~20 m、20~40 m、40~60 m 和 60 m 以上 4 个区域 (如图 1 中的黑色等深线所示)。根据水深数据的精度, 本研究在上述 4 个区域内原水深数据的基础上, 分别叠加 0.5 m、2 m、4 m 和 6 m, 生成相应的有偏水深数据, 并保持模式其他参数不变, 进行约 10 d (以 M_2 分潮周期 12.42 h 计算, 共计 20 个周期) 自由

积分。在模式积分的最后一个M₂分潮周期(后文简称周期)内,将水位计算结果与原模式结果进行对比,计算各网格点水位的时间平均均方根误差(Root Mean Square Error, RMSE),结果如图2所示。由图可以看出,大部分区域的水位误差在0.1 m以上,且靠

近陆地的海域,水位误差更大,尤其在黄海的北部和东部,最大水位误差达到了0.6 m。这一水深参数的敏感性分析试验结果表明,在本研究所采用的水深参数区域划分方案下,M₂分潮的模拟结果能够产生明显的差异,为开展后续的数据同化试验奠定了基础。

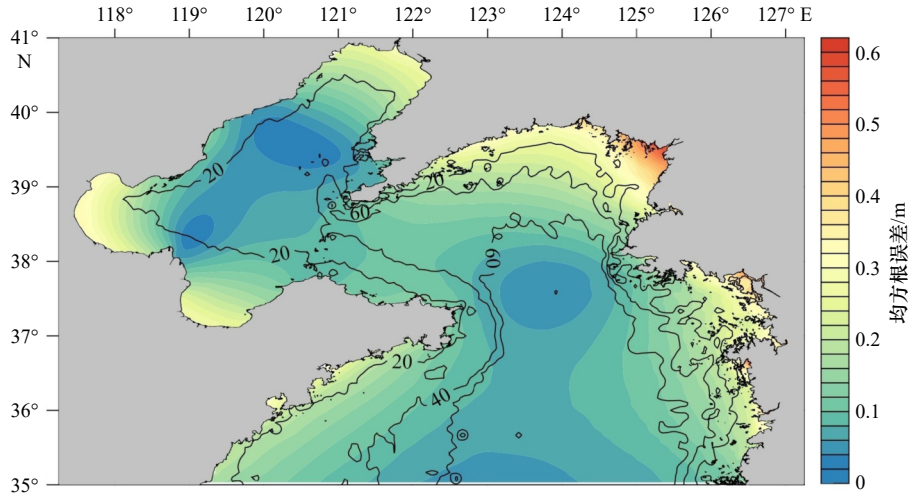


图2 基于有偏水深数据获得的水位数值模拟结果的时间平均均方根误差

Fig. 2 Time mean root mean square error of water level simulated with biased water depths

黑色实线分别为20 m、40 m和60 m等深线

The black solid lines indicate the isobaths of 20 m, 40 m, and 60 m

3.2 数据同化试验

本研究拟同时开展理想和实际水深参数估计试验。其中理想数据同化试验又称为孪生试验。在这两种数据同化试验中,均将水深增量作为待估计的参数,后文统一简称为水深参数。待估计的水深参数的设置采用与敏感性分析试验相同的分区方案,即划分为0~20 m、20~40 m、40~60 m和60 m这4个区域,从而有4个与其相对应的水深参数待估计。以某一个模式网格点(i, j)为例,估计后的水深按下式计算:

$$H_{i,j}^a = H_{i,j} + \Delta H_m^a, \quad (9)$$

式中, $H_{i,j}^a$ 为模式网格点的后验水深; $H_{i,j}$ 为由ETOPO2数据获得的水深; ΔH_m^a ($m = 1, 2, 3, 4$)为后验水深参数。

在后续的理想和实际数据同化试验中,集合数均设定为30,每个时间步均进行同化,同化时长设定为3个周期。对于状态估计,由于POMgcs模式采用蛙跳格式的时间差分方案,因此在同化试验中,当前的观测同时用于调整当前和上一个时间步的水位和潮流状态变量^[30]。同化后,利用更新后的水深继续积分约3 d以达到稳定。利用最后一个周期的水位数据进行调和和分析。需要注意的是,同化过程中,调整后的水深有可能会使得 $D < 0$ 。为此,在同化过程中为水深的调整设定一个阈值,以保证这种情况不会发生。

3.2.1 理想数据同化试验

在理想数据同化试验中,将ETOPO2的水深数据视为真实水深,利用该水深自由积分得到的水位数据作为真实场。在此基础上,将真实场叠加上标准差为0.1 m的高斯白噪声生成与模式网格点和时间积分步一致的水位观测场。

水深参数集合的形成与敏感性分析试验相同,即在0~20 m、20~40 m、40~60 m和60 m这4个区域内,分别以0.5 m、2 m、4 m和6 m为先验水深参数 ΔH_m^p ($m = 1, 2, 3, 4$),将其叠加高斯白噪声形成相应的先验水深参数集合,标准差分别取为各先验水深参数值的5%。

数值模型基于上述有偏的水深参数集合积分约7 d以达到稳定。需要说明的是,在测试中发现:同时进行状态和水深参数估计已经能够获得很好的试验结果,因此在理想数据同化试验中并未采用DAEPC方法。

3.2.2 实际数据同化试验

在实际数据同化试验中,利用NAO.99Jb的调和常数计算出模式积分的所有时间步、每隔3个模式水平网格点的水位值,连同验潮站处所计算得到的水位值一并作为本研究的观测数据。令4个先验水深参

数均为 0, 即 $\Delta H_m^p = 0$ ($m = 1, 2, 3, 4$), 然后采用叠加高斯白噪声的方式各自形成先验水深参数集合, 其标准差与理想试验相同。模型积分约 7 d 达到稳定。之后引入 DAEPIC 方法进行水深参数估计。在测试中发现, 进行模型状态场估计的 1 个周期后, 状态场估计的误差已达到准平衡态。因此, 进行实际数据同化试验时, 在第 1 个周期仅进行状态场估计, 之后同时进行状态场和水深参数估计。在实际数据同化试验中, 进行如下两组试验: (1) 仅利用来自 NAO.99Jb 的观测数据, 设置观测误差为 0.2 m。 (2) 同时利用来自 NAO.99Jb 和验潮站的观测数据, 设置 NAO.99Jb 数据的观测误差为 0.2 m, 验潮站数据的观测误差为 0.1 m。这两组试验中所给定的观测误差是通过多次数据同化试验后获得的最优结果。

4 结果

4.1 理想数据同化试验

本节主要讨论当数值模式的误差仅来自水深参数时, EAKF 方法的参数估计效果。通过多组试验结果对比, 发现局地化半径为 40 个网格点, 不进行状态场膨胀, 而参数膨胀系数为 1.3 时, 水深参数估计的效果最好。因此, 这里基于该组试验结果, 分析 EAKF 方法对前述 4 个参数的估计效果。

图 3a 和图 3b 分别为同化后空间平均的水位 RMSE 时间序列和 4 个水深参数估计值与真值偏差的时间序列。为方便进行比较, 将图 3b 中水深参数估计偏差进行标准化处理后形成图 3c, 即将 4 个水深参数估计偏差的时间序列分别除以初始时刻设置的水深参数值(0.5 m、2 m、4 m 和 6 m)。

由图 3a 和图 3b 可见, 在数据同化试验的初期, 模式水位误差由浅水区域主导, 因此 ΔH_1^p 很快达到稳定。与此同时, EAKF 的状态估计很快将水位误差调整到观测误差(0.1 m)以下。一段时间后, 状态误差已经很低, 此时模式误差主要来自 ΔH_2^p 和 ΔH_3^p 。在同化 3 个周期后, 两个参数在真实值附近稳定, 水位误差也进一步降低。而 ΔH_4^p 虽然很快达到稳定, 但稳定后呈周期性震荡, 与真实值始终维持一定的距离。同化结束后, 水位误差保持在 0.01 m 以下, 说明模式状态场误差已经大幅降低, 且同化后模式的自由积分(3 个周期后)结果与真实场也十分接近(图形未在此给出)。

由标准化后的水深参数偏差(图 3c)可见, 在参数估计的起始阶段, ΔH_1^p 很快回到了真实值附近, 说明 ΔH_1^p 有很高的敏感性。在参数估计的第 1 和第 2 个周期, ΔH_1^p 产生了较明显的震荡, 这是由于 ΔH_2^p 、 ΔH_3^p 和 ΔH_4^p 仍未稳定, ΔH_1^p 在参数估计的过程中不断被调

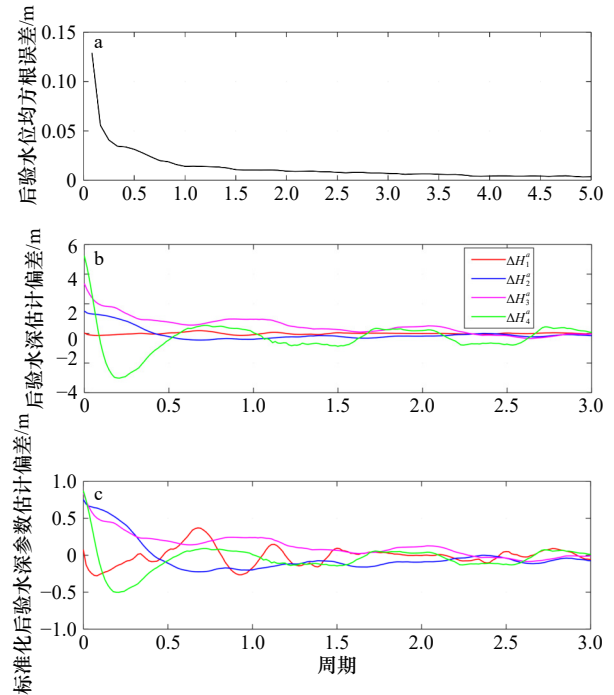


图 3 后验水位均方根误差(a)、后验水深参数估计偏差(b)时间序列和标准化的后验水深参数估计偏差时间序列(c)

Fig. 3 Time series of the root mean square error for the posterior water level (a), bias (b) and standardized bias (c) for the posterior water depth parameters

整。而到了第 3 个周期, ΔH_2^p 、 ΔH_3^p 和 ΔH_4^p 基本稳定后, ΔH_1^p 也保持在真值附近。 ΔH_2^p 和 ΔH_3^p 在整个参数估计过程中趋势比较平稳, 偏差减小的速度较慢, 说明这两处位置的敏感性相对较低, 需要较长时间的参数估计才能达到稳定。而 ΔH_4^p 在稳定后一直与真值保持着一定的偏差, 说明 ΔH_4^p 并不敏感, 即使没有回到真值, 也不能产生很大的状态误差。

图 4 为基于真实场和理想数据同化试验前后数据绘制的 M_2 分潮同潮图。其中, 迟角以格林威治时间(对应迟角为 0°)为基准绘制。由图 4b 可见, 同化前, 海区边缘处的振幅和无潮点附近的迟角均出现一定误差, 且渤海北部的无潮点位置相较于真实场(图 4a)出现偏移。同化后, 振幅和迟角(图 4c)相比于真实场均已无明显差异, 渤海北部无潮点的位置也得到改善。

表 1 给出了先验和后验振幅与迟角的空间平均误差。同化前, 振幅的误差约为 7.6 cm, 迟角误差约为 $12^\circ 54'$; 同化后, 振幅误差仅为 0.2 cm, 迟角误差仅为 $18'$ 。表明同化后, 模式误差整体上得以大幅降低。

综上所述, 在理想数据同化试验中, EAKF 方法通过对水深参数进行最优估计, 有效地降低了模式的偏差。这一结果也表明 EAKF 方法可以很好地应用

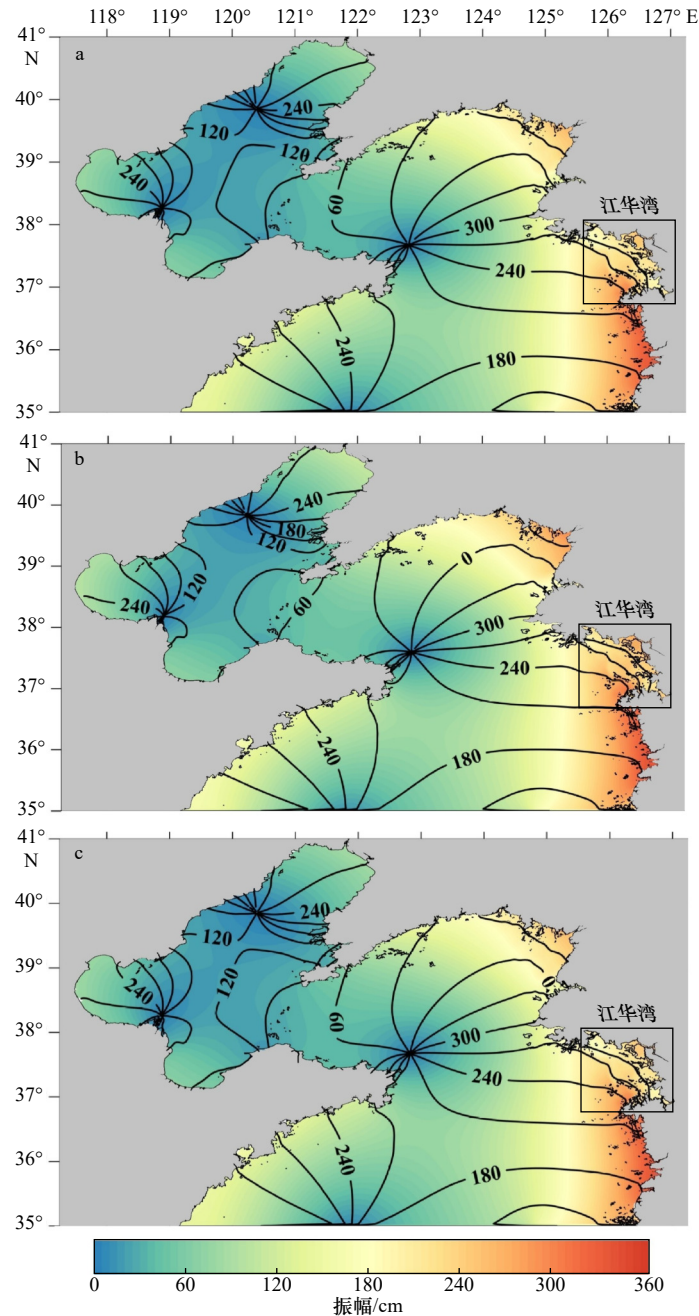


图4 利用真实场数据(a)、理想试验中的先验结果(b)和后验结果(c)绘制的M₂分潮同潮图

Fig. 4 Cotidal chart of M₂ constituent from the truth (a), prior (b) and posterior (c)

黑线为迟角,单位:(°)

The black line is phase lag, unit: (°)

于潮汐潮流数值模拟的水深参数估计中。

4.2 实际数据同化试验

4.2.1 基于NAO.99Jb数据的水深估计

在本试验中,将NAO.99Jb数据作为观测进行水深估计。经过多组试验,选取局地化半径为15个网格点、状态膨胀系数为1.03、参数膨胀系数为1.10的数据同化试验结果在此讨论。

图5为后验水位均方根误差(将数据同化试验得到的水位与NAO.99Jb数据对比)的时间序列分布图。

由图可见,在同化开始后,后验水位均方根误差在周期性震荡的过程中下降。这是由于模式计算结果与NAO.99Jb数据之间存在较大的相位误差,导致两者水位变化方向有时不同,从而使得在这段时间内,模式模拟的水位均方根误差不断增大。但总体上来看,水位均方根误差保持在观测误差(0.2 m)附近波动。而模式状态的调整无法改正这一相位误差。第1个周期后,开始同时进行状态和参数估计,这时由于水深参数的调整,使得水位模拟均方根误差不断减小,

表 1 理想数据同化试验中 M_2 分潮的先验和后验振幅与迟角空间平均误差

Table 1 Spatial averaged errors of amplitude and phase lag of M_2 constituent from the prior and posterior in twin experiment

	振幅误差/cm	迟角误差
同化前	7.6	12°54'
同化后	0.2	18'

最后维持在 0.1 m 附近,且小于所给定的 NAO.99Jb 数据的观测误差(0.2 m),说明数据同化试验是成功的。

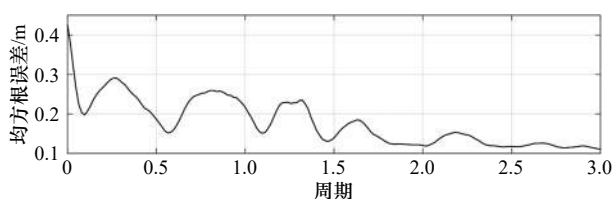


图 5 后验水位均方根误差的时间序列

Fig. 5 Time series of the root mean square error for the posterior water level

图 6 为基于 NAO.99Jb 和实际数据同化试验前后数据绘制的 M_2 分潮同潮图。由图 6c 可见,整体上,同化后模式的调和与分析结果更接近 NAO.99Jb 数据,尤其是在黄海的东北和东南部,同化后振幅明显减小,迟角的分布也更趋近 NAO.99Jb。同时,各无潮点的位置也有一定改善:渤海北部的无潮点向左偏移,渤海南部、黄海中部无潮点则向下移动。对比图 6a、图 6b 和图 6c 可见,江华湾部分误差在同化后误差较大。导致这一现象的原因可能是由于本研究中估计的水深参数个数较少,使得在当前这样的参数估计方案下,估计后的水深参数虽然使得总体的误差在减小,但这种减小的方式有可能是以某些区域的模拟误差减小、某些区域的模拟误差增大的方式而获得,而不是以整个研究区域的模拟误差同时减小的方式获得。因此,设计更好的水深参数估计方案是后续需要进一步开展的研究工作。

表 2 给出了实际数据同化试验中 M_2 分潮的调和常数在同化前后与 NAO.99Jb 相比的空间平均误差。从整体上来看,无论是振幅还是迟角,后验结果都优于先验结果。

4.2.2 基于 NAO.99Jb 和验潮站数据的水深估计

在 4.2.1 节的实际数据同化试验基础上,加入验潮站数据进行水深估计。在大量同化试验方案测试的基础上得到如下两组较优的数据同化试验:试验 1 为使用全部 NAO.99Jb 和验潮站数据,开展水深估计试验;试验 2 为模式网格点先验水深小于 20 m 且非渤

海区域内仅使用验潮站数据,其余部分则使用全部数据,开展水深参数估计试验。

在上述同化试验中,将验潮站数据的观测误差设置为 0.1 m, NAO.99Jb 数据的观测误差设置为 0.2 m。其余试验设置与 4.2.1 节相同。基于试验 1 和试验 2 的结果,将 4.2.1 节的 NAO.99Jb 数据同化试验作为对比试验,进行结果讨论分析。

表 3 给出了同化前、试验 1、试验 2 和对比试验与验潮站振幅,迟角的空间平均误差。由表 3 可以看出,各试验的整体误差相较于同化前均有所降低。其中,试验 2 的误差最小,相比于同化前,振幅误差减小了 40.27%,迟角误差降低了 49.19%;而试验 1 的振幅和迟角误差只分别降低了 28.52% 和 29.39%。

图 7 给出了试验 1、试验 2 和对比试验所得到的 M_2 分潮同潮图以及在验潮站位置处同化前后振幅和迟角误差的变化情况。图 8 为先验、试验 1、试验 2 和对比试验在各验潮站位置处的调和常数误差统计直方图。由图 7 和图 8 可见,相较于同化前,试验 1 中验潮站位置处的振幅和迟角误差减小的站点数分别为 21 个和 26 个,试验 2 中分别为 25 个和 30 个。这表明虽然试验 2 相较于试验 1 所使用的观测信息更少,但验潮站位置处振幅和迟角的整体误差以及误差分布都优于试验 1。这可能是由于在同化试验中的水深参数估计中出现了“少数服从多数”的现象,即 EAKF 方法根据大部分观测信息来调整水深,使观测信息较多的地方水位模拟结果得到改进,而观测信息较少的地方水位模拟结果反而变差。具体来说,在试验 1 中,黄海北部的模拟结果更好,渤海的模拟误差则较大,其余验潮站位置处接近对比试验。而试验 2 中江华湾区域改进明显,渤海和黄海北部的误差则接近对比试验。这表明在先验水深小于 20 m 且非渤海区域的 NAO.99Jb 数据使 EAKF 更倾向于优化黄海北部地区。这也说明去除这些观测信息后,其余观测信息的比重被调整,反而使更多区域的水深参数被优化。

在图 8a 和图 8b 中,试验 2 相较试验 1 得到优化的站点共 20 个。其中,渤海和江华湾部分的振幅得到改进,尤其是江华湾区域的改进尤为明显;而黄海北部和南部反而变差。在图 8c 和图 8d 中,试验 2 的迟角优于试验 1 的验潮站同样是 20 个。其中,黄海北部、渤海和黄海东南部的提升较为明显,江华湾、黄海西部和南部则有所下降。因此,试验 2 虽然在整体上振幅和迟角的误差更小,但误差的空间分布没有明显改进,说明在后续的研究中仍需要进一步探索如何改进参数估计的同化试验方案。

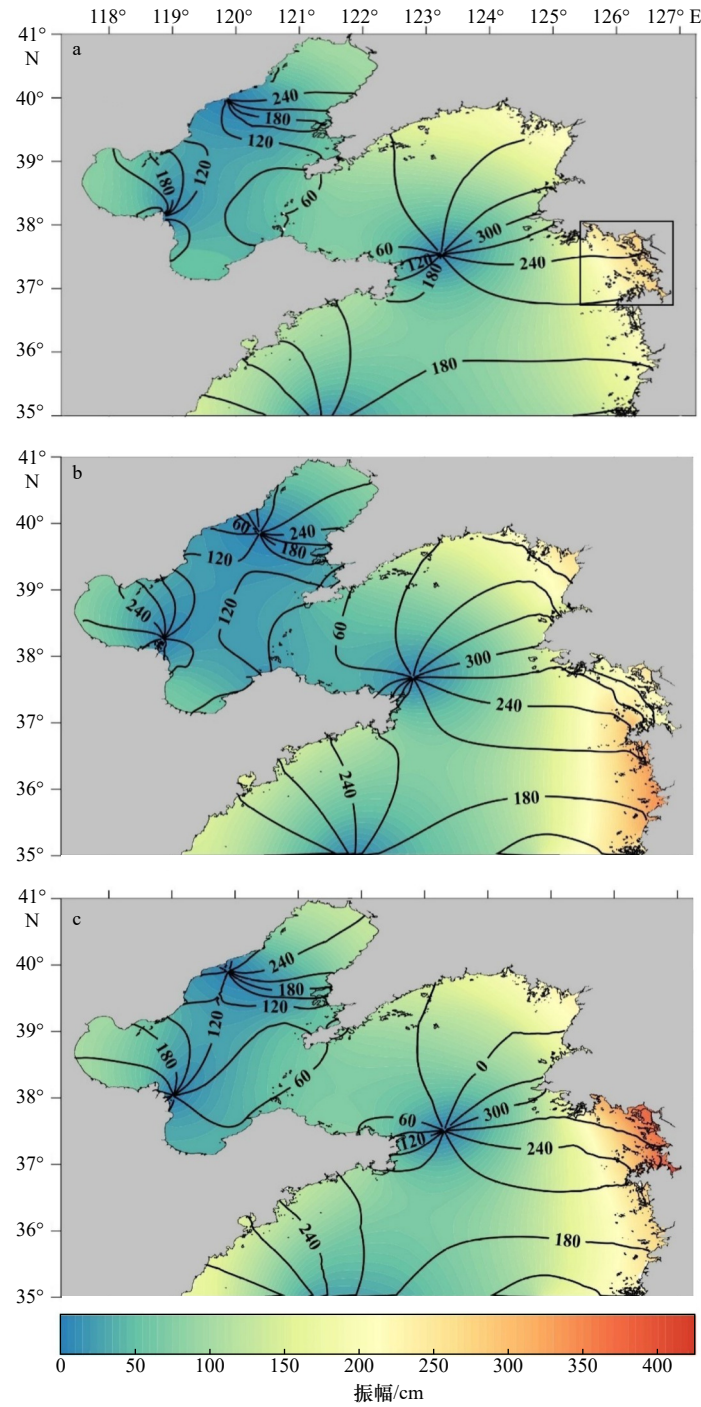


图6 利用NAO.99Jb数据(a)和NAO.99Jb实际数据同化试验中的先验(b)和后验(c)结果绘制的M₂分潮同潮图
 Fig. 6 Cotidal chart of M₂ constituent from the NAO.99Jb (a), prior (b) and posterior (c)

黑线为迟角,单位:(°)

The black line is phase lag, unit: (°)

此外,由图8c和图8d可见,试验1中葫芦岛站(图1)中4号的迟角存在较大误差,达到了175°10',几乎是反位相。对照其地理位置可以看出,该处距离渤海北部的无潮点较近(图7b)。由于无潮点模拟位置的偏移,在该点处产生了较大的误差,从而也导致了试验1模拟的迟角整体误差大幅增大(表3)。如果误差统

计中不考虑该验潮站的迟角误差,则试验1模拟的迟角误差为16°36'。相比较而言,对比试验的迟角误差则为16°56'。因此,从这个角度来说,试验1相较于对比试验,迟角的模拟仍然有所改进。这从对比图7b和图7f中验潮站误差变小的个数上也可以得到佐证。

在上述实际数据同化试验中,4个水深参数的调

表 2 NAO.99Jb 实际数据同化试验中 M₂ 分潮的先验和后验振幅与迟角空间平均误差

Table 2 Spatial averaged errors of amplitude and phase lag of M₂ constituent for the prior and posterior in the NAO.99Jb data assimilation experiment

	振幅误差/cm	迟角误差
同化前	26.1	24°46'
同化后	18.1	16°23'

整量分别约为 6 m、-4 m、15 m 和 10 m。需要说明的是, 不同于理想数据同化试验有“真实”的水深参数可供比较, 对于实际数据同化试验来说, 并没有这样的“真实”水深存在。而对于数值模式来说, 其模拟结

表 3 同化前、试验 1、试验 2 和对比试验与验潮站振幅、迟角的平均空间误差

Table 3 Spatial averaged errors of amplitude and phase lag from the model, experiment 1, experiment 2 and NAO.99Jb data assimilation experiment with respect to those from tide gauges

	振幅误差/cm	迟角误差
同化前	29.8	30°07'
试验1	21.3	21°16'
试验2	17.8	15°18'
对比试验	21.8	16°56'

果的误差主要来自模式的动力框架本身以及不确定的开边界条件、底摩擦系数和水深。因此, 在本研究

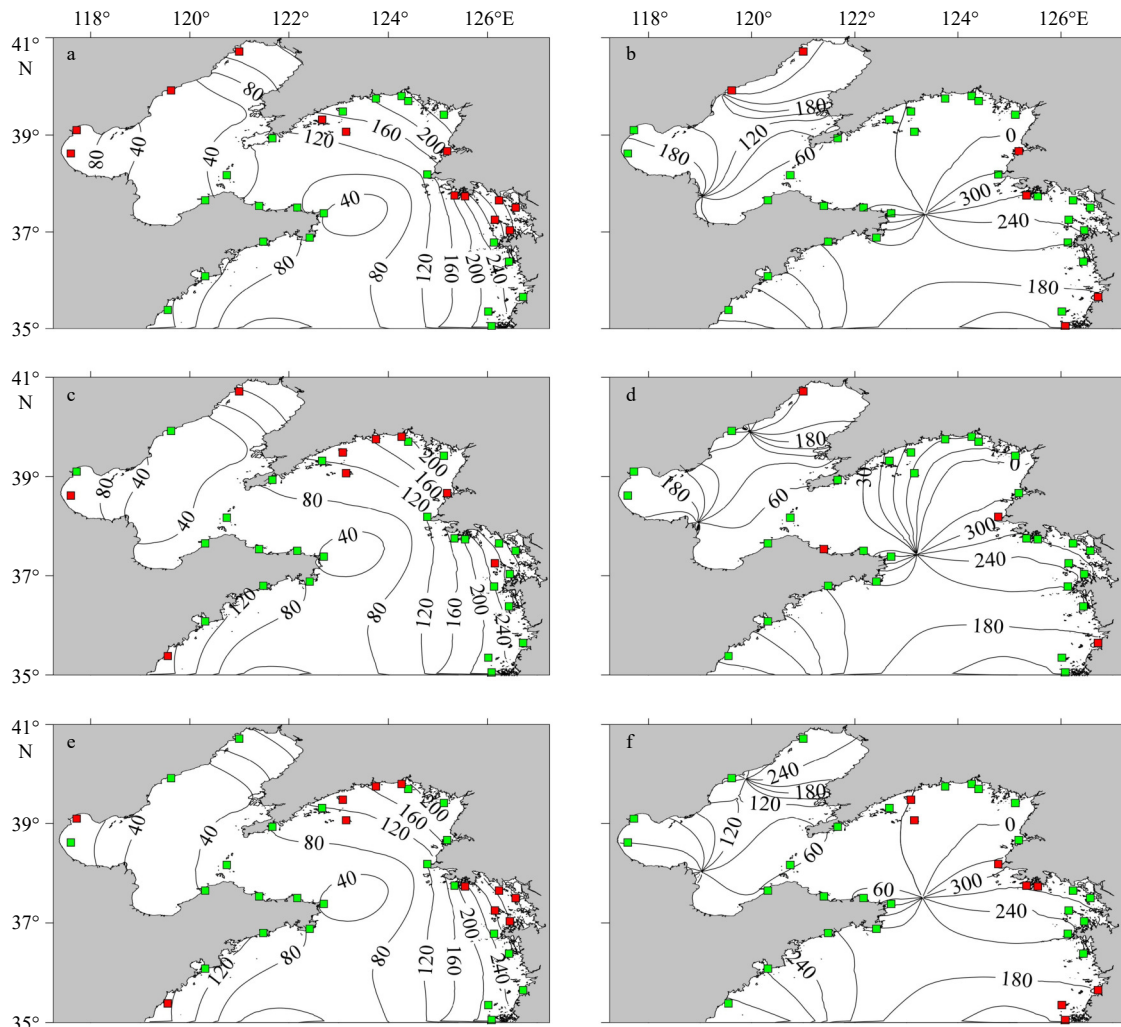


图 7 试验 1(a, b)、试验 2(c, d)和对比试验(e, f)的振幅(a, c, e; 单位: cm)和迟角(b, d, f; 单位:(°)及其在各个验潮站位置的数值相较于同化前的变化情况

Fig. 7 Amplitude (a, c, e; unit: cm) and phase lag (b, d, f; unit: (°)) from experiment 1 (a, b), experiment 2 (c, d) and NAO.99Jb data assimilation experiment (e, f), and change of errors at each tide gauge

绿色表示相较于同化前误差变小, 而红色表示误差变大

Green squares represent the errors are reduced and red squares represent the errors are increased with respect the prior

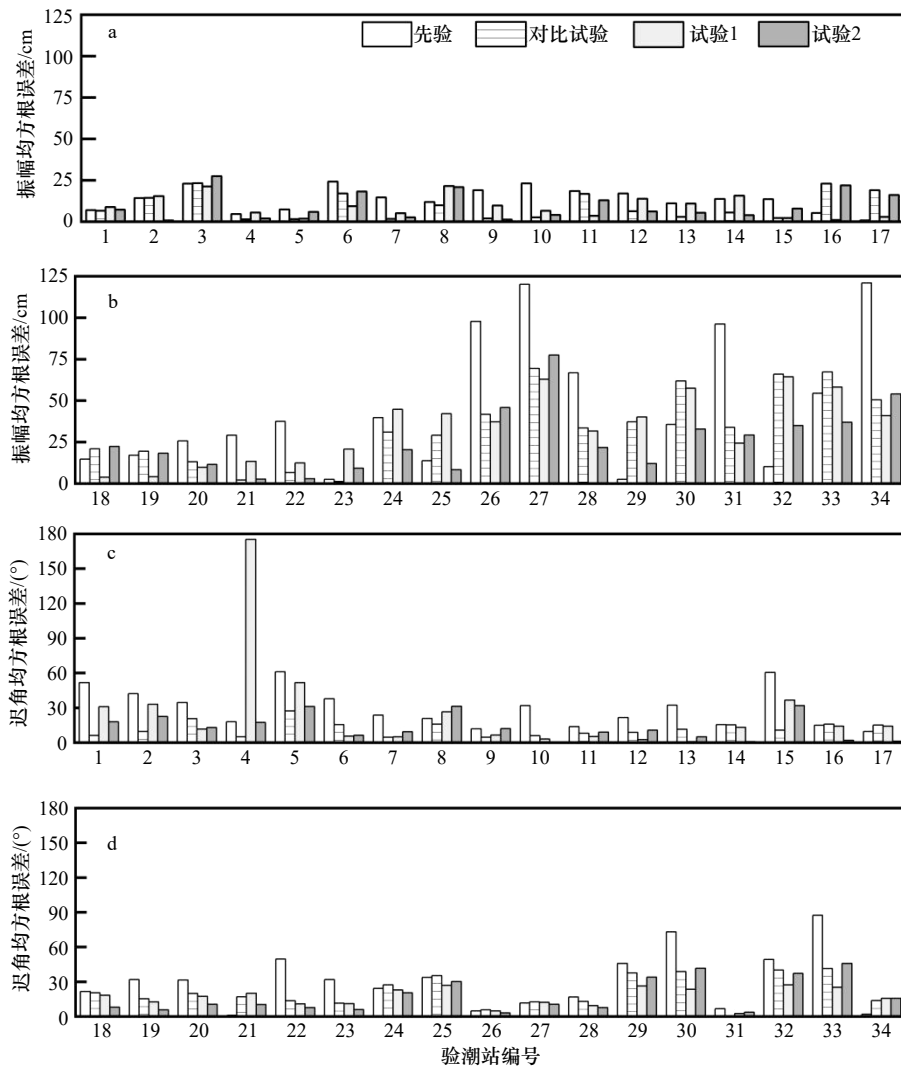


图8 模式先验、对比试验、试验1和试验2的振幅(a, b)和迟角(c, d)在验潮站位置处的均方根误差

Fig. 8 Amplitude (a, b) and phase lag (c, d) root mean square errors of the prior, NAO.99Jb data assimilation experiment, experiment 1, and experiment 2 with respect to the tide gauges

验潮站编号1~34与图1中对应

The tidal gauge station numbers correspond to those in Fig.1

中,由于只进行了水深的参数估计,上述模式的动力框架本身以及不确定的开边界条件和底摩擦系数所导致的模拟误差都涵盖在水深参数的调整中。因此,调整后的水深其实并不代表实际海域所观测的水深。

5 结论

本研究基于集合调整卡尔曼滤波(EAKF)方法,采用广义坐标系统的美国普林斯顿大学海洋模式(POMgcs)的外模式开展了渤海和部分黄海海域M₂分潮模拟中的水深估计研究。通过理想数据同化试验(又称孪生试验)和基于NAO.99Jb与验潮站数据,并采用强化的参数校正数据同化方法(DAEPIC)的实际数据同化试验结果的讨论分析,得出主要结论如下。

(1)EAKF方法在理想数据同化试验中可以实现对潮汐潮流模拟中的水深参数的最优估计,还原水深参数的“真实场”。这表明,EAKF方法可以很好地应用于潮汐潮流模拟中的状态场和参数估计。

(2)在实际数据同化试验中,利用DAEPIC方法开展了状态和水深参数估计。当仅使用NAO.99Jb数据作为观测时,EAKF方法可以通过实际的水深估计,提高数值模型对M₂分潮振幅和迟角的模拟精度。相较于NAO.99Jb数据,水深参数估计后的振幅误差减小了30.65%,迟角误差减小了33.86%。

(3)在NAO.99Jb数据基础上加入验潮站数据的水深估计试验中,通过研究改进观测系统,进一步提高了水深估计的效果。对比验潮站数据,模拟的M₂

分潮振幅与迟角误差相较于同化前分别减小了 40.27% 和 49.19%。在先验水深小于 20 m 且非渤海的区域仅使用验潮站的观测信息后,水深参数估计可以在整体上进一步提高 M_2 分潮的模拟效果,但误差的空间分布没有明显改进。

值得注意的是,在实际数据同化试验中,江华湾

区域的误差最大。这是由于该区域海底地形复杂,且水深较浅,数值模式本身模拟结果的准确度不高。即使进行了数据同化,对状态场和水深参数估计后,该区域的误差仍然较大。因此,需要在本研究的基础上,进一步开展数值模式调试的同时,设计更为适合的数据同化方案,以提高状态场和参数估计的效果。

参考文献:

- [1] Das S K, Lardner R W. Variational parameter estimation for a two-dimensional numerical tidal model[J]. *International Journal for Numerical Methods in Fluids*, 1992, 15(3): 313–327.
- [2] Lardner R W, Al-Rabeh A H, Gunay N. Optimal estimation of parameters for a two-dimensional hydrodynamical model of the Arabian Gulf[J]. *Journal of Geophysical Research: Oceans*, 1993, 98(C10): 18229–18242.
- [3] 朱江, 曾庆存, 郭冬建, 等. 利用伴随算子法从岸边潮位站资料估计近岸模式的开边界条件[J]. *中国科学(D辑: 地球科学)*, 1998, 41(3): 330–336.
Zhu Jiang, Zeng Qingcun, Guo Dongjian, et al. Estimating open boundary conditions from coastal tidal observations by adjoint approach[J]. *Science in China Series D: Earth Sciences*, 1998, 41(3): 330–336.
- [4] 吕咸青, 张杰. 如何利用水位资料反演开边界条件(一)[J]. *水动力学研究与进展(A辑)*, 1999, 14(4B): 92–102.
Lv Xianqing, Zhang Jie. How to invert open boundary condition from water level data (I)[J]. *Chinese Journal of Hydrodynamics*, 1999, 14(4B): 92–102.
- [5] 韩桂军, 何柏荣, 马继瑞, 等. 利用伴随法优化非线性潮汐模型的开边界条件 I: 伴随方程的建立及“孪生”数值试验[J]. *海洋学报*, 2000, 22(6): 27–33.
Han Guijun, He Bairong, Ma Jirui, et al. Optimizing open boundary conditions of nonlinear tidal model using adjoint method I: The establishment of adjoint model and twin-experiment[J]. *Haiyang Xuebao*, 2000, 22(6): 27–33.
- [6] 韩桂军, 方国洪, 马继瑞, 等. 利用伴随法优化非线性潮汐模型的开边界条件 II: 黄海、东海潮汐资料的同化试验[J]. *海洋学报*, 2001, 23(2): 25–31.
Han Guijun, Fang Guohong, Ma Jirui, et al. Optimizing open boundary conditions of nonlinear tidal model using adjoint method II: Assimilation experiment for tide in the Huanghai Sea and the East China Sea[J]. *Haiyang Xuebao*, 2001, 23(2): 25–31.
- [7] 吕咸青, 吴自库, 殷忠斌, 等. 渤、黄、东海潮汐开边界的1种反演方法[J]. *青岛海洋大学学报*, 2003, 33(2): 165–172.
Lv Xianqing, Wu Ziku, Yin Zhongbin, et al. Inversion of the tides with the open boundaries of the Bohai, Huanghai and Donghai Seas[J]. *Periodical of Ocean University of China*, 2003, 33(2): 165–172.
- [8] Han Guijun, Li Wei, He Zhongjie, et al. Assimilated tidal results of tide gauge and TOPEX/POSEIDON data over the China seas using a variational adjoint approach with a nonlinear numerical model[J]. *Advances in Atmospheric Sciences*, 2006, 23(3): 449–460.
- [9] Zhang Jicai, Lu Xianqing. Parameter estimation for a three-dimensional numerical barotropic tidal model with adjoint method[J]. *International Journal for Numerical Methods in Fluids*, 2008, 57(1): 47–92.
- [10] Qian Shouguo, Lv Xianqing, Sun Rencheng, et al. The estimation of open boundary conditions in the Bohai Sea with POD 4D VAR data assimilation[C]//*Proceedings of 2016 International Conference on Applied Mathematics, Simulation and Modelling*. Beijing: Atlantis Press, 2016: 9–11.
- [11] Qian Suhui, Wang Daosheng, Zhang Jicai, et al. Adjoint estimation and interpretation of spatially varying bottom friction coefficients of the M_2 tide for a tidal model in the Bohai, Yellow and East China Seas with multi-mission satellite observations[J]. *Ocean Modelling*, 2021, 161: 101783.
- [12] Wang Daosheng, Zhang Jicai, Wang Yaping. Estimation of bottom friction coefficient in multi-constituent tidal models using the adjoint method: temporal variations and spatial distributions[J]. *Journal of Geophysical Research: Oceans*, 2021, 126(5): e2020JC016949.
- [13] Ngodock H E, Souopgui I, Wallcraft A J, et al. On improving the accuracy of the M_2 barotropic tides embedded in a high-resolution global ocean circulation model[J]. *Ocean Modelling*, 2016, 97: 16–26.
- [14] Wilson G, Özkan-Haller H T. Ensemble-based data assimilation for estimation of river depths[J]. *Journal of Atmospheric and Oceanic Technology*, 2012, 29(10): 1558–1568.
- [15] Wilson G W, Özkan-Haller H T, Holman R A. Data assimilation and bathymetric inversion in a two-dimensional horizontal surf zone model[J]. *Journal of Geophysical Research: Oceans*, 2010, 115(C12): C12057.
- [16] Wilson G W, Özkan-Haller H T, Holman R A, et al. Surf zone bathymetry and circulation predictions via data assimilation of remote sensing observations[J]. *Journal of Geophysical Research: Oceans*, 2014, 119(3): 1993–2016.
- [17] Landon K C, Wilson G W, Özkan-Haller H T, et al. Bathymetry estimation using drifter-based velocity measurements on the Kootenai River, Idaho[J]. *Journal of Atmospheric and Oceanic Technology*, 2014, 31(2): 503–514.
- [18] Moghimi S, Özkan-Haller H T, Wilson G W, et al. Data assimilation for bathymetry estimation at a tidal inlet[J]. *Journal of Atmospheric and Oceanic Technology*, 2016, 33(10): 2145–2163.
- [19] An H S. A numerical experiment of the M_2 tide in the Yellow Sea[J]. *Journal of the Oceanographical Society of Japan*, 1977, 33(2):

- 103–110.
- [20] 方国洪. 黄海潮能的消耗[J]. 海洋与湖沼, 1979, 10(3): 200–213.
Fang Guohong. Dissipation of tidal energy in Yellow Sea[J]. *Oceanologia et Limnologia Sinica*, 1979, 10(3): 200–213.
- [21] Matsumoto K, Takanezawa T, Ooe M. Ocean tide models developed by assimilating Topex/Poseidon altimeter data into hydrodynamical model: a global model and a regional model around Japan[J]. *Journal of Oceanography*, 2000, 56(5): 567–581.
- [22] Egbert G D, Bennett A F, Foreman M G G. Topex/Poseidon tides estimated using a global inverse model[J]. *Journal of Geophysical Research: Oceans*, 1994, 99(C12): 24821–24852.
- [23] 李大炜, 李建成, 金涛勇, 等. 利用验潮站资料评估全球海潮模型的精度[J]. 大地测量与地球动力学, 2012, 32(4): 106–110.
Li Dawei, Li Jiancheng, Jin Taoyong, et al. Accuracy estimation of recent global ocean tide models using tide gauge data[J]. *Journal of Geodesy and Geodynamics*, 2012, 32(4): 106–110.
- [24] Mellor G L, Häkkinen S M, Ezer T, et al. A generalization of a sigma coordinate ocean model and an intercomparison of model vertical grids[M]//Pinaridi N, Woods J. *Ocean Forecasting*. Berlin, Heidelberg: Springer, 2002: 55–72.
- [25] Anderson J L. An ensemble adjustment Kalman filter for data assimilation[J]. *Monthly Weather Review*, 2001, 129(12): 2884–2903.
- [26] Gaspari G, Cohn S E. Construction of correlation functions in two and three dimensions[J]. *Quarterly Journal of the Royal Meteorological Society*, 1999, 125(554): 723–757.
- [27] Aksoy A, Zhang Fuqing, Nielsen-Gammon J W. Ensemble-based simultaneous state and parameter estimation with MM5[J]. *Geophysical Research Letters*, 2006, 33(12): L12801.
- [28] Tong Mingjing, Xue Ming. Simultaneous estimation of microphysical parameters and atmospheric state with simulated radar data and ensemble square root Kalman filter. Part II: parameter estimation experiments[J]. *Monthly Weather Review*, 2008, 136(5): 1649–1668.
- [29] Zhang Shaoqing, Liu Zhengyu, Rosati A, et al. A study of enhance parameter correction with coupled data assimilation for climate estimation and prediction using a simple coupled model[J]. *Tellus A: Dynamic Meteorology and Oceanography*, 2012, 64(1): 10963.
- [30] Zhang Shaoqing, Anderson J L, Rosati A, et al. Multiple time level adjustment for data assimilation[J]. *Tellus A*, 2004, 56(1): 2–15.

Bathymetry estimation using ensemble adjustment Kalman filter in the numerical simulation of M_2 constituent

Wu Haowen¹, Zhao Yanling², Han Guijun¹, Li Wei¹, Cao Lige¹, Wu Xiaobo¹,
Li Chaoliang¹, Li Yundong¹, Zhou Gongfu¹

(1. School of Marine Science and Technology, Tianjin University, Tianjin 300072, China; 2. The 31010 Army of PLA, Beijing 100081, China)

Abstract: Data assimilation can estimate the uncertain parameters in the numerical model while adjusting the state variables with observations to improve the simulation results through enhancing the numerical model. Based on the ensemble adjustment Kalman filter (EAKF) and the external mode of the Princeton ocean model with generalized coordinate system (POMgcs), a bathymetry estimate is performed in the M_2 constituent simulation of the Bohai Sea and part of the Yellow Sea. The results of the ideal data assimilation experiment or identical twin experiment show that the EAKF method can retrieve the “truth” bathymetry. In the practical data assimilation experiment of the NAO.99Jb and tide gauge data, by comparing with the 34 tide gauges, the model simulated amplitude and phase lag errors of M_2 constituent are reduced by 40.27% and 49.19%, respectively, by use of the posterior estimate of the bathymetry.

Key words: data assimilation; EAKF; numerical simulation; Bohai Sea; Yellow Sea; M_2 constituent; bathymetry estimation