



HAL
open science

Calibration of tidal model parameters by data assimilation with Particle Swarm Optimisation

Anju Sebastian, Jérôme Thiebot, Emmanuel Poizot, Pierre-Henri Musiedlak,
Gwendoline Gregoire

► **To cite this version:**

Anju Sebastian, Jérôme Thiebot, Emmanuel Poizot, Pierre-Henri Musiedlak, Gwendoline Gregoire. Calibration of tidal model parameters by data assimilation with Particle Swarm Optimisation. 30th TELEMAT User Conference, Oct 2024, Chambéry (73), France. hal-04738099

HAL Id: hal-04738099

<https://hal.science/hal-04738099v1>

Submitted on 15 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Calibration of tidal model parameters by data assimilation with Particle Swarm Optimisation

A. Sebastian¹, J. Thiébot¹, E. Poizot¹, P. Musiedlak², G. Gregoire¹

anju.sebastian@unicaen.fr, Cherbourg, France

¹: LUSAC EA 4253, University of Caen Normandy, 60 rue Max-Pol Fouchet, France

²: HydroQuest SAS, 38240 Meylan, France

Abstract –In this study, we calibrate a Telemac2D model, which solves shallow water equations, to improve tidal predictions. Observational tide gauge data is assimilated using the Particle Swarm Optimization (PSO) algorithm, to calibrate and estimate two critical model parameters: bottom friction and boundary conditions. The boundary conditions include the amplitude and phase of the predominant tidal constituents, while bottom friction is estimated as both constant and spatially varying. Our results show that while spatially varying bottom friction provides a slight improvement over constant friction, the most significant reduction in model error is achieved by calibrating both the amplitude and phase of the tidal constituents. The study concludes that the phase correction of predominant tidal constituents is the most critical parameter for improving the model’s accuracy, emphasizing the importance of precise boundary condition calibration in tidal modeling.

Keywords: water elevation, PSO, calibration, boundary condition, bottom friction.

I. INTRODUCTION

Tidal models are essential tools in predicting coastal hazards and assessing environmental impacts. Tides in shallow water models are influenced by bottom friction, boundary conditions and bathymetry. The accuracy of tidal models is often compromised by uncertainties in input data. For instance, bottom friction is commonly treated as a constant despite its known variability across different spatial and temporal scales. Similarly, the number of tidal constituents and the resolution of boundary condition databases can significantly influence model performance. Traditional trial-and-error methods for calibrating these models are not only computationally expensive and time-consuming but also often fail to efficiently optimize model parameters.

Data assimilation (DA) methods can be efficient in calibrating coastal models and estimating optimal model parameters [1]. Traditional DA techniques include variational methods such as 3DVAR and 4DVAR with the adjoint method, as well as filtering methods like the Kalman filter. These techniques have been effectively employed to calibrate and estimate open boundary conditions (OBCs) and bottom friction in coastal ocean models. For instance, Taniguchi et al. [2] successfully calibrated OBCs in a tidal model using the ensemble Kalman filter and a model Green’s function via data assimilation. Similarly, Cao et al. [3] optimized the OBC in an internal tidal model using the adjoint method, assimilating satellite altimetry data. Chen et al. [4] and Jiang et al. [5]

further refined the spatially varying OBC using the adjoint method combined with an independent point scheme, also based on satellite altimeter data. Zhang et al. [6] implemented an adjoint-based 4DVAR data assimilation technique for OBC estimation, while Lardner [7] employed the adjoint variational data assimilation method to estimate boundary conditions using tide gauge data. Improved OBC estimation has also been achieved through the adjoint method with a trigonometric polynomial fitting scheme, assimilating multi-satellite data [8].

In addition to boundary condition calibration, data assimilation has been applied to the calibration of bottom friction. Wang et al. [9] estimated spatially and temporally varying bottom friction in a tidal model using the adjoint method, demonstrating its effectiveness. Moreover, Wang et al. [10] showed that data assimilation-based bottom friction parameter estimation can yield better results in global tidal models. The calibration of varying bottom friction has also been explored in tidal models by assimilating satellite altimetry data [11].

The focus of this work is on the simultaneous estimation of various model parameters with minimum knowledge of their initial conditions. Traditional methods such as gradient based methods and filtering methods can be sensitive to the initial conditions and can be computationally expensive and time consuming as the number of calibration parameters increases [12,13,14]. In this case, we have tested the Particle Swarm Optimisation (PSO) algorithm for data assimilation and parameter estimation problem.

In this study, tidal elevation observations from tide gauges are assimilated into the model to calibrate both bottom friction and boundary conditions. The primary aim is to compare and distinguish the effects of bottom friction and boundary conditions on the calibration of the tidal model and to identify the most efficient approach for improving model performance.

In the following sections, Section 2 outlines the observational data, hydrodynamic model, calibration variables, and the data assimilation framework used in this study. Section 3 discusses the validation of the PSO algorithm and the calibration process, with a focus on the estimation of boundary conditions and bottom friction. Section 4 summarizes the findings of the study and outlines potential future directions.

II. METHODOLOGY

A. Study site and observations

The English Channel is a shallow epi-continental sea shared

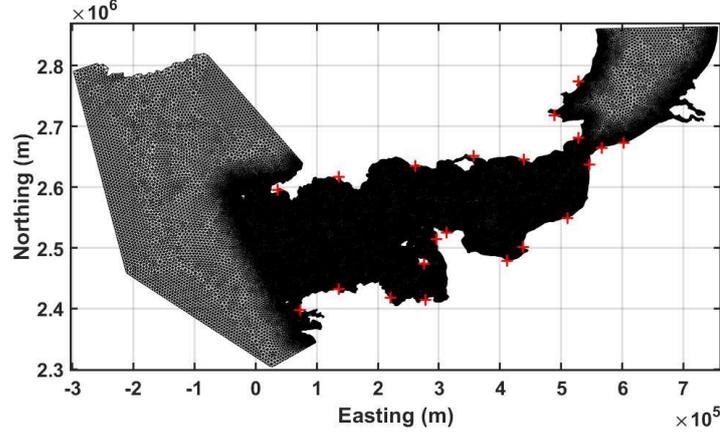


Figure 1. Study domain and the location of the tide gauges marked as red dots

between France to the south and the United Kingdom to the north. Most of the depth values fall between 50 to 80 m with a maximum depth of 174 m. The channel is characterized by semi-diurnal tides with a higher tidal range along the French coast.

Measurements of tidal elevation from 21 tide gauge locations (Figure 1) across the French and the English coasts are used in this study for data assimilation over a period of 10 days during October 2015. The particular period of the study is chosen based on the availability of the data (longer time periods had no effect on the methodology). The tide gauge data along the French coast are sourced from the SHOM (French navy) and along the English coast from the British Oceanographic Data Centre (BODC).

B. Hydrodynamic model setup

Telemac2D is the hydrodynamic model used to simulate tidal elevation. Telemac2D solves the second-order partial differential equations for depth-averaged free surface flow called shallow water equations derived from the full three-dimensional Navier Stokes equations. The vertically averaged equations are obtained by integrating from the bottom to the surface. This provides a system of equations with mass continuity and two momentum equations, considering a Newtonian fluid. The continuity equation and momentum equations are given by (1), (2) and (3) respectively.

$$\frac{\partial h}{\partial t} + \frac{\partial(hU)}{\partial x} + \frac{\partial(hV)}{\partial y} = 0 \quad (1)$$

$$\frac{\partial(hU)}{\partial t} + \frac{\partial(hUU)}{\partial x} + \frac{\partial(hUV)}{\partial y} = -gh \frac{\partial Z}{\partial x} + \frac{\partial}{\partial x} \left(h v_e \frac{\partial U}{\partial x} \right) + \frac{\partial}{\partial y} \left(h v_e \frac{\partial U}{\partial y} \right) + h F_x \quad (2)$$

$$\frac{\partial(hV)}{\partial t} + \frac{\partial(hUV)}{\partial x} + \frac{\partial(hVV)}{\partial y} = -gh \frac{\partial Z}{\partial y} + \frac{\partial}{\partial x} \left(h v_e \frac{\partial V}{\partial x} \right) + \frac{\partial}{\partial y} \left(h v_e \frac{\partial V}{\partial y} \right) + h F_y \quad (3)$$

where h is the depth, U and V are the velocity components (depth-averaged) in the x and y directions, t is the time, g is the gravitational acceleration, Z is the elevation of the free surface (m), v_e is the coefficient of diffusion (m^2/s^2), F_x and F_y are the source terms of the momentum equation, which include bottom friction and Coriolis force.

Telemac2D solves the above-mentioned equations using a finite element method on a triangular element mesh. The mesh resolution varies between 1 km to 5 km. The time step is set to 60 seconds and the model simulations are obtained for 10 days. The turbulence closure relies on the $k-\epsilon$ model. Bathymetry is from the digital elevation model of the SHOM with a resolution of 100 m. At the open boundaries, the model is forced with water elevation given by the Previmer database of 37 tidal constituents with a resolution of 2 km.

Amplitude and phase of two tidal constituents M_2 and S_2 are calibrated in the boundary condition. Amplitude is calibrated by multiplying a dimensionless factor a , where a is in the range 0.85 to 1.15. The phase is calibrated by adding a value θ , where θ is the phase shift that can vary between -10° and $+10^\circ$. It is to be noted that the a and θ for the tidal constituents M_2 and S_2 are considered separately in the calibration.

The bottom friction is estimated with the Nikuradse formula given by (4).

$$C_b = \left(\frac{\kappa}{\ln(h/ez_0)} \right)^2 \quad (4)$$

where C_b is the friction coefficient, κ is the von Karman constant and z_0 is the bottom roughness defined as the height above the seabed where the fluid velocity is zero. This variable (bottom roughness z_0) is used for the model calibration. z_0 is

considered either constant or varying in space for calibration tests (see Figure 2). When considering a spatially varying bottom friction, the calibration becomes more complex but considering a heterogeneous seabed can improve the model results [15]. In the English Channel, the seabed is highly heterogeneous as shown in Figure 2 where the spatial distribution of sediments (data from SHOM) is classified into four categories (hence four different roughness values). This classification is based on hypotheticalal mean based on the size and proportion of the sediments in each class given in the sediment map from SHOM. The sediment classes and the corresponding ranges of the bottom roughness values (in mm)

$$J(x) = J(x) = \frac{1}{2}(x - x^b)^T \cdot B^{-1} \cdot (x - x^b) + \frac{1}{2}(H_0(t) - H_T(x, t))^T \cdot R^{-1} \cdot (H_0(t) - H_T(x, t)) \quad (5)$$

where x is the calibration parameter values, x^b is the initial guess of the calibration parameters, R is the observation error covariance matrix (here, it is an identity matrix) and B is the background error covariance matrix. Here, weight is given to the observation considering greater variance for the background error such that first term in the cost function is neglected. As a result, $J(x)$ (m²) can be equated to the Root Mean Square Error (RMSE in m) by the following formula:

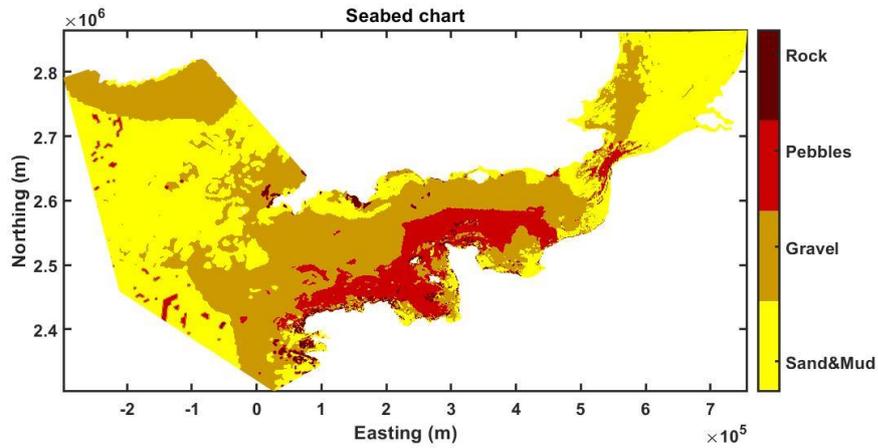


Figure 2. sediment map showing the 4 types of sediments, rock, pebbles, gravel and sand&mud

in this study are Sand&Mud [0.01-3], Gravel [1-7], Pebbles [4-25] and Rock [20-40]. The ranges of bottom roughness are taken from [16] and early studies in which models are calibrated with a comparable friction formula [17].

The DA module ADAO [19] is coupled with Telemac2D. ADAO has various coded algorithms within it which enables the user to test different algorithms. Figure 3 represents the coupling of ADAO with Telemac2D that enables assimilating observational data and estimating the calibration parameters. The observational data $H_0(t)$ and the model data $H_T(x, t)$ are given to the DA framework along with the ranges of the variables to be calibrated. For the present purpose, $H_0(t)$ and $H_T(x, t)$ are time-series of water elevations at different tidal gauge locations. The model data $H_T(x, t)$ depends on x which is the parameters (z_{0i}, a_i, θ_i) to be calibrated. The DA framework updates the values of the variables in each iteration (i) of the algorithm. n is the total number of iterations. A cost function $J(x)$ is calculated with the DA framework which gives the error between the observation and the model output. The value of cost function is minimized as iterations progress. The equation for the cost function $J(x)$ is as follows:

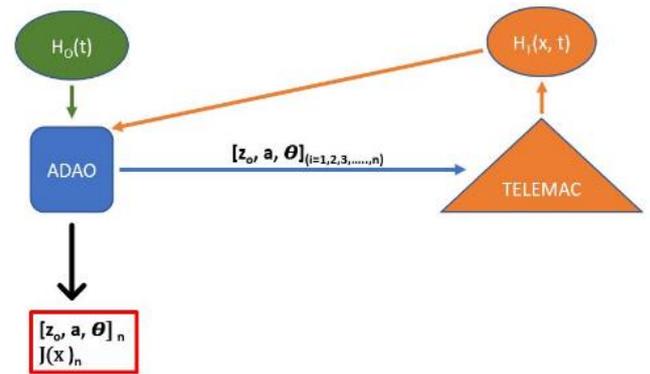


Figure 3. Data assimilation framework ADAO

$$RMSE = \sqrt{\frac{2J(x)}{n_{TG}n_t}}$$

where n_{TG} is the number of tide gauge locations and n_t is the length of the timeseries.

The DA algorithm tested in this study is the PSO (Particle Swarm Optimisation) which is a global optimisation algorithm [18]. PSO considers a swarm of particles that move towards

the best position or the optimal solution in the given bounds. Particles retain the memory of their previous best position during the search time both globally and individually. A set of iterations are employed in the process to reach the minimum error. At each iteration, a new velocity is assigned to the particle based on the global best position from which a new position is again given to the particle. The following two equations give the velocity vector and the position vector in PSO algorithm respectively:

$$v_{id} = wv_{id} + C_1r_1(P_{id} - X_{id}) + C_2r_2(P_{gd} - X_{id}) \quad (6)$$

$$X_{id} = X_{id} + v_{id} \quad (7)$$

where c_1 and c_2 are the acceleration constants. v_{id} is the particle's previous velocity. P_{id} is the previous best position of the particle. P_{gd} is the best particle position among the population. The second part on the left-hand side of the velocity equation (6) is the cognitive part which represents each particle's individual experience. The third part on the left-hand side of the equation (6) is called the social component where the particle gains information from its neighbouring particles. w is the inertia weight which controls the particle's search for the optimal value between the local and the global states.

Synthetic data are used to check the suitability of the algorithm in this study. In this case, the observation data (e.g., time-series of water elevations) is replaced by synthetic data, i.e., the results of a model simulation performed with randomly chosen (known) values of calibration parameters (e.g., values of bottom roughness). This model simulation and parameter values are considered to be the true state of the system. To check if the DA works well, the model results (e.g., time-series of water elevations) are assimilated and it is expected that the calibrated parameters (e.g., values of bottom roughness) will converge to the parameter values used to build the synthetic data.

The application of synthetic data serves multiple purposes. Firstly, tuning the PSO parameters for this specific case study is crucial to ensure that the algorithmic parameters (e.g., number of particles, number of iterations) are set appropriately. Secondly, synthetic data helps demonstrate the effectiveness of PSO in calibrating multiple parameters (we test up to eight calibration parameters), with a focus on reducing the objective function $J(x)$. Thirdly, as the tide gauge data is only available at the boundaries of the study domain (near the coast), with no observations in the middle of the study domain, it is essential to verify that the available observational data is sufficient to achieve satisfactory performance in the data assimilation framework.

After validating the PSO algorithm with synthetic data, we assimilate real observations from tide gauges. Several tests will be performed with different combinations of the parameters given in Table 1. The aim is to achieve the minimum value of the $J(x)$ with the optimal values of the parameters within the given ranges of the values of the parameters. This will lead to the best combination of parameters to obtain the model predicted water elevation in good agreement with the observations of water elevation.

Table 1. Parameter combinations tested in calibration

Case	Parameter combination
1	z_0
2	$z_{01}, z_{02}, z_{03}, z_{04}$
3	$z_{01}, z_{02}, z_{03}, z_{04}, a_{M2}, \theta_{M2}, a_{S2}, \theta_{S2}$
4	$z_0, a_{M2}, \theta_{M2}, a_{S2}, \theta_{S2}$
5	$z_{01}, z_{02}, z_{03}, z_{04}, a_{M2}, a_{S2}$
6	$z_{01}, z_{02}, z_{03}, z_{04}, \theta_{M2}, \theta_{S2}$

III. RESULTS AND DISCUSSION

A. Validation using synthetic data

As a starting point, limited number of calibration parameters, say bottom friction have been considered for setting up the parameters of the DA algorithm. After several tests with different number of particles and number of iterations, the number of particles is set to 40 and the number of iterations to 12 since higher numbers didn't show any improvement in the cost function. Other parameters of the DA algorithm are set to the default values. To calculate the number of simulations, the number of particles is to be multiplied by the number of iterations giving 480 simulations. Later, more parameters to be calibrated are included in the DA process. The results have shown that the DA algorithm can handle numerous numbers of parameters without an increase in the computational time. This has the advantage of testing larger set of parameters, see the improvement in $J(x)$ and find the most influential parameters in the model performance.

The results from the above tests have given $J(x)$ nearly equal to zero indicating that the DA algorithm permitted to retrieve the true calibration parameters (the ones used to build the synthetic data). Further improvement of $J(x)$ is expected by adding more data (assimilated) points across the domain where no tide gauge data is available. However, this hasn't given a considerable improvement in $J(x)$. This suggests that the assimilated observation data and the number of data points from the tide gauges are sufficient for the DA framework to perform the best.

Table 2. True value of the parameters in the synthetic data vs the converged value after the DA

Parameters	True Value (Synthetic data)	Converged value
z_{01} (mm)	2.2	2.9
z_{02} (mm)	3.7	2.9
z_{03} (mm)	4	4.3
z_{04} (mm)	27	40
a_{M_2} (°)	1	0.99
θ_{M_2} (°)	4	4.1
a_{S_2} (°)	1.1	1.11
θ_{S_2} (°)	2	2.3

The use of synthetic data permitted to confirm that the DA is working properly. As an example, the synthetic data test with 21 data points (at all the tide gauge locations) and eight calibration variables has given satisfactory results. The convergence of the calibration parameters shown in Table 2 to the proposed true values in the reference case (synthetic data) with a minimum $J(x)$ of 1.46 m² [Figure 4], points out to the reliability of the algorithm. One of the bottom friction values (z_{04}) hasn't converged to the proposed true value. We expect that it is because the corresponding sediment class swaps lesser surface area when considering the total surface area of the study domain (Figure 2 - Rock).

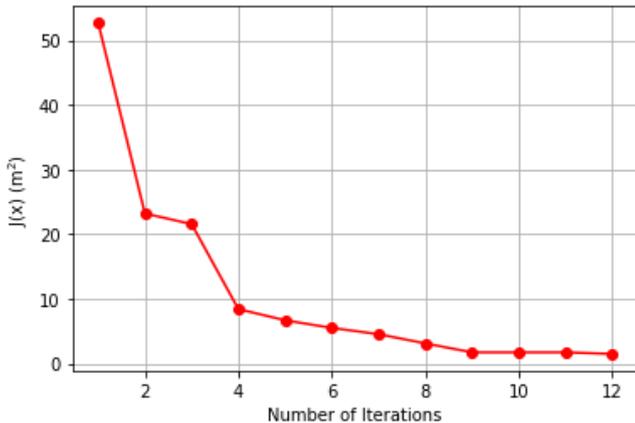


Figure 4. Reduction of $J(x)$ over iterations for the reference case using synthetic data

B. Assimilating real observation data

DA with the (real) observations from the 21 tide gauge locations is now done for different parameter combinations as listed in Table 1. Calibration using constant bottom friction and varying bottom friction have given $J(x)$ values over the iterations as 101.19 m² and 97.18 m² respectively. Even though constant and varying bottom friction give almost the

same minimised $J(x)$ values, varying bottom friction (4 types of seabed roughness) has given a $J(x)$ value by a small margin.



Figure 5. $J(x)$ values after calibration with different parameter combinations

Adding calibration variables for the boundary condition (modification of amplitude and phase) along with the bottom friction has resulted in a reduction of $J(x)$. The $J(x)$ values for the case 3 and case 4 are 78.6 m² and 80.13 m² respectively. When boundary condition variables are calibrated with both constant (case 3) and varying bottom friction (case 4), the errors show the same trend as seen in case 1 and case 2. Specifically, the cases that account for spatial variation in bottom friction consistently show lower $J(x)$ values compared to those with uniform bottom friction.

To study the influence of the amplitude and phase of M_2 and S_2 on $J(x)$, amplitude and phase are treated separately in case 5 and case 6. It is observed that exclusion of the phase or amplitude has a significant effect on $J(x)$. The variable set excluding phase has given $J(x)$ of 90.29 m² and excluding amplitude has given 77.24 m². Tuning of phase gives better improvement in $J(x)$ suggesting phase is the most important parameter to be calibrated in our model.

The RMSE equivalent of the $J(x)$ for case 1 and case 6 are 0.199 m and 0.174 m respectively. This depicts the significant improvement of the water elevation by a few centimetres from considering one calibration parameter, the constant bottom friction to multiple parameters including bottom friction and boundary condition.

IV. CONCLUSION

Coastal hydrodynamic models are inherently uncertain and require proper calibration to enhance their accuracy. This study presents the calibration of a tidal model using data assimilation combined with parameter estimation through the Particle Swarm Optimization (PSO) algorithm. The use of synthetic data generation has reinforced confidence in the methodology. The optimal estimation of bottom friction and

boundary conditions, followed by assimilation, has notably improved model performance. Incorporating the spatial variation of sediments has yielded better results, consistent with previous studies [15]. The study found that calibrating the model using both friction and boundary condition parameters is the most effective approach to improving results. Adjusting the amplitude and phase of tidal constituents in the boundary condition significantly impacts tidal elevation, with phase adjustments contributing more to error reduction than amplitude modifications. The influence of phase and amplitude adjustments on boundary conditions warrants further investigation to determine whether these effects stem from errors in the boundary condition database or the adjustments themselves. Additional experiments are needed to clarify the impact of boundary conditions, particularly the phase, on water elevation results.

ACKNOWLEDGEMENT

Anju Sebastian acknowledges the University of Caen for funding her PhD. The authors are grateful to the British Oceanographic Data Centre (BODC) for providing tidal gauge data, to the French navy SHOM (“Service Hydrographique et Océanographique de la Marine”) for providing access to bathymetric, sedimentological and tidal gauge data. We are grateful to IFREMER for providing the Previmer data (atlas of tidal constituents). We also thank the developers of Telemac and ADAO, a module for Data Assimilation and Optimization <http://www.salome-platform.org/>

REFERENCES

- [1] Fringer OB, Dawson CN, He R, Ralston DK, Zhang YJ. The future of coastal and estuarine modeling: Findings from a workshop. *Ocean modelling*. 2019 Nov 1;143:101458.
- [2] Taniguchi N, Mutsuda H, Arai M, Sakuno Y, Hamada K, Huang CF, Guo J, Takahashi T, Yoshiki K, Yamamoto H. Application of coastal acoustic tomography: calibration of open boundary conditions on a numerical ocean model for tidal currents. *Frontiers in Marine Science*. 2024 Mar 19;11:1351390.
- [3] Cao A, Chen H, Zhang J, Lv X. Optimization of open boundary conditions in a 3D internal tidal model with the adjoint method around Hawaii. In *Abstract and Applied Analysis* 2013 (Vol. 2013, No. 1, p. 950926). Hindawi Publishing Corporation.
- [4] Chen H, Cao A, Zhang J, Miao C, Lv X. Estimation of spatially varying open boundary conditions for a numerical internal tidal model with adjoint method. *Mathematics and Computers in Simulation*. 2014 Mar 1;97:14-38.
- [5] Jiang D, Chen H, Jin G, Lv X. Estimating smoothly varying open boundary conditions for a 3D internal tidal model with an improved independent point scheme. *Journal of Atmospheric and Oceanic Technology*. 2018 Jun 1;35(6):1299-311.
- [6] Zhang J, Wang YP. A method for inversion of periodic open boundary conditions in two-dimensional tidal models. *Computer Methods in Applied Mechanics and Engineering*. 2014 Jun 15;275:20-38.
- [7] Lardner RW. Optimal control of open boundary conditions for a numerical tidal model. *Computer Methods in Applied Mechanics and Engineering*. 1993 Feb 1;102(3):367-87.
- [8] Jiao S, Zhang Y, Pan H, Lv X. Improved estimation of the open boundary conditions in tidal models using trigonometric polynomials fitting scheme. *Remote Sensing*. 2023 Jan 13;15(2):480.
- [9] Wang D, Zhang J, Wang YP. Estimation of bottom friction coefficient in multi - constituent tidal models using the adjoint method: Temporal variations and spatial distributions. *Journal of Geophysical Research: Oceans*. 2021 May;126(5):e2020JC016949.
- [10] Wang X, Verlaan M, Veenstra J, Lin HX. Data-assimilation-based parameter estimation of bathymetry and bottom friction coefficient to improve coastal accuracy in a global tide model. *Ocean Science*. 2022 Jun 9;18(3):881-904.
- [11] Lu X, Zhang J. Numerical study on spatially varying bottom friction coefficient of a 2D tidal model with adjoint method. *Continental shelf research*. 2006 Oct 1;26(16):1905-23.
- [12] Moore, A. M., Arango, H. G., Broquet, G., Powell, B. S., Weaver, A. T., & Zavala-Garay, J. (2011). The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part I–System overview and formulation. *Progress in Oceanography*, [91\(1\), 34-49](#).
- [13] Hamill, T. M., Whitaker, J. S., & Snyder, C. (2001). Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Monthly Weather Review*, [129\(11\), 2776-2790](#).
- [14] Kalnay, E. (2003). *Atmospheric modeling, data assimilation and predictability*. Cambridge university press.
- [15] Warder SC, Angeloudis A, Piggott MD. Sedimentological data-driven bottom friction parameter estimation in modelling Bristol Channel tidal dynamics. *Ocean Dynamics*. 2022 Jun;72(6):361-82.
- [16] Soulsby RL. Dynamics of marine sands: a manual for practical applications. *Oceanographic Literature Review*. 1997;9(44):947
- [17] Guillou N, Thiébot J. The impact of seabed rock roughness on tidal stream power extraction. *Energy*. 2016 Oct 1;112:762-73.
- [18] R.C. Eberhart, Y. Shi, Particle swarm optimization: developments, applications and resources, *Proceedings of the IEEE Congress of Evolutionary Computation 1* (2001) 27–30.
- [19] ADAO, a module for Data Assimilation and Optimization, <http://www.salome-platform.org/>
- [20] JM Hervouet, R. Ata (2017). User manual of open software TELEMAC-2D, Report, EDF-R&D, <http://www.opentelemac.org/>, ISBN:9780470035580.