進化アルゴリズムを用いた海底地形データセットの最適化

稲津大祐，樋口知之，中村和幸
統計数理研究所／科学技術振興機構
〒106-8569 東京都港区南麻布 4-6-7
inazud@ism.ac.jp, higuchi@ism.ac.jp, nakakazu@ism.ac.jp

和文抄録

進化アルゴリズムを用い，領域海洋潮汐モデルの境界条件および物理パラメータを最適化した．水深は海洋モデルの境界条件であり，人工衛星による重力観測に基づいて推定された全球水深データから与えることが多い．しかし，人工衛星起源の水深データの精度は，海洋潮汐のシミュレーションにとってしばしば不十分である．我々は進化アルゴリズムを用い，アラスカ多島海の複数の特定の海域の平均水深を最適化した．最適化された平均水深は，船舶からの音波観測に基づく水深データに比較的忠実であった．一般的に，音波観測起源の水深データは，重力観測起源のデータに比べ正確であるが，データはまばらにしか存在しない．一方，最適化の結果，重力観測起源の水深データの誤差が非常に大きい場所を特定することができた．このように，進化アルゴリズムを用いる最適化手法は，海洋モデルの境界条件の修正を行う有効な手立てとなると期待される．

An evolutionary algorithm optimizes boundary bathymetric condition in an ocean tide model

Daisuke INAZU, Tomoyuki HIGUCHI and Kazuyuki NAKAMURA
The Institute of Statistical Mathematics/Japan Science and Technology Agency
4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, JAPAN
inazud@ism.ac.jp, higuchi@ism.ac.jp, nakakazu@ism.ac.jp

Abstract

An evolutionary algorithm (EA) is applied to a regional ocean tide model to optimize a boundary condition and a physical parameter in the model. The bathymetry is one of the boundary conditions in ocean models and is frequently given by global sea depth data based on estimation derived from the satellite gravity observation. Accuracy of the satellite-derived depth data is often insufficient for accurate ocean tide simulations. We apply the EA to modify the mean depth in specific regions in the Alaska Panhandle, U.S. The mean depth in those regions is optimized to be close to the multibeam-derived depth data which is generally accurate further more than the satellite-derived data but its spatial coverage is sparse. Also, a region with large errors of the satellite-derived data is successfully found through the optimization. The EA has a prospect of correction of erroneous boundary conditions for accurate ocean modeling.

1. Introduction

Predictions and forecasts should be accurate to represent real phenomena. One of familiar forecasts is the weather forecast. Recently ocean weather forecasts have been operationally developed in the U.S., Europe, and also in Japan (e.g., Sugiura et al., 2008). Accuracy of the forecasts is not always sufficient for our demands as found even in usual weather forecasts. This inaccuracy arises from various causes. Uncertainties of boundary conditions of the simulations strongly affect model accuracy. Bathymetry is a boundary condition and its accuracy is crucial to the coastal ocean modeling. Global bathymetry has been estimated based on gravity observation from the satellite (Smith and Sandwell, 1997). The satellite-derived depth data is useful and frequently sets the bathymetric condition in ocean models; however, the accuracy of the satellite-derived data is often insufficient especially for coastal ocean modeling (e.g., Sandwell and Smith, 2001). Some studies have addressed the problems with inversion modeling to modify inaccurate bathymetry (e.g., Losch and Wunsch, 2003; Hirose, 2005). On the other hand, approaches using evolutionary algorithms (EAs) are not used for those problems partly due to large numerical computation costs. Yet it is important to survey potential of the easy statistical method for parameter estimation and optimization in ocean models. In this paper, we apply the EA to a simulation model of the ocean tide.

2. Ocean tides in the Alaska Panhandle

The target sea is the Alaska Panhandle in the U.S. which is characterized by complicated coastlines and bathymetry. Inazu et al. (accepted, henceforth I09) showed the following: In the case that the bathymetric condition in the region is configured by only the satellite-derived depth data called GINA (Lindquist et al., 2004), the simulated result of the ocean tide is cruel. However, they achieved to model the most accurate ocean tide in the region by setting the bathymetric condition with a combination of GINA and the multibeam-derived depth data from ship surveys. The multibeam-derived (MB) data is in general accurate more than the GINA bathymetry, but its coverage is sparse. In addition, the accuracy of the simulation is also shown to be sensitive to the bottom friction coefficient which is one of the physical parameters in the simulation model. Can we modify the inaccurate bathymetric condition in the case without the MB data? Is the bottom friction coefficient appropriately estimated? We use the EA to address the problem.

3. Data

Three kinds of data are used in the present study. One is the tide gauge (sea level) data observed at coastal sites (Fig. 1). Ocean tides are basically composed by specific, sinusoidal, diurnal and semidiurnal constituents such as M_2, S_2, K_1, and O_1 constituents. The subscripts of 2 and 1 denote semidiurnal and diurnal constituents, respectively. The information of respective tidal constituents at a coastal site is usually given by its amplitude and phase that are estimated by the harmonic analysis (Munk and Cartwright, 1966). The amplitude and phase of each tidal constituent is called the tidal harmonics which is obtained from time series of sea level at each coastal site. The observed tidal harmonics are used to validate the simulated tidal harmonics.

The other data is the two bathymetric data: GINA and the MB data. The GINA bathymetry is the global data but its accuracy seems insufficient at coastal regions. The original GINA bathymetry is to set the initial condition, and is sequentially updated in the process of the EA. The MB data is probably realistic more than GINA but its coverage is sparse over the Alaska Panhandle. The MB data is basically obtained from National Geographic Data Center. The finally optimized bathymetry is validated by the MB data.

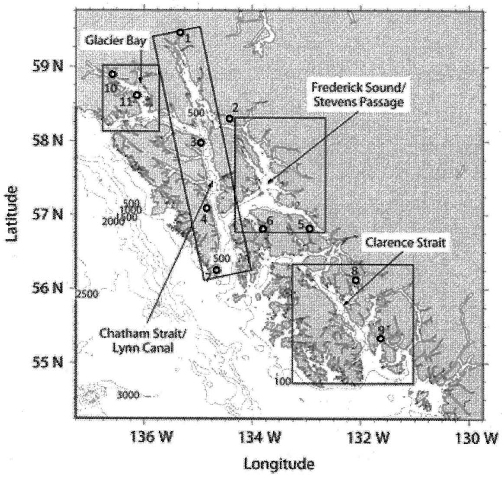

Figure 1. Computational domain of the simulation model. Contours indicate bathymetry in meters drawn from GINA. Open circles denote the coastal tide gauges used in the present study. Rectangles indicate specified regions in which the GINA bathymetry is perturbed to be optimized.

4. Simulation model

The simulation model of the ocean tide in the Alaska Panhandle is configured in this section. The simulation model is based on the following governing equations:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} + \mathbf{f} \times \mathbf{u} = -g\nabla\eta - \gamma_b \frac{\mathbf{u}|\mathbf{u}|}{H} + A_H \nabla^2 \mathbf{u} \, , \quad (1)$$

$$\frac{\partial \eta}{\partial t} + \nabla \cdot (\mathbf{u}H) = 0 \, \cdot \quad (2)$$

These partial differential equations are called the Navier-Stokes equation. A vector $\mathbf{u}$ is the horizontal depth-integrated velocity, η is the sea surface elevation, $\mathbf{f}$ is the Coriolis parameter which is a constant vector defined by latitude, g is the gravity acceleration (= 9.8 m sec^{-2}), and H is the total depth of water column. Dissipation terms are parameterized by two coefficients: γ_b for quadratic bottom friction, and A_H for horizontal viscosity.

The computational domain is shown in Fig. 1. The model grid spacing is set to $1' \times 1'$, and the number of the model grids is 329×449. The sea level elevation at the computational boundary is given by the accurate condition of the composite tides developed by I09 that forces the interior sea level elevation. The simulated variables are $\mathbf{u}$ and η. The GINA bathymetry gives H of the initial generation. A_H is given by an empirical value of 1 m^2 sec^{-1} based on the Richardson's low that is the same value used by I09. Since γ_b moderately affects the simulated results as mentioned in Section 2, H and γ_b are the target parameters to be optimized.

5. Evolutionary algorithm

The optimization of the parameter using the EA is described in this section. The EA is used to maximize an objective function $L(\theta)$, where θ is the parameter to be optimized. The objective function is defined to be large when the difference between the observed and the simulated tidal harmonics is small (not shown).

It is expected to be difficult to optimize H and γ_b at all the model grids because of the large number of the model grids. The dimension of the parameter space should be limited to be reduced. Since the speed of surface gravity waves in straits and channels is roughly determined by the mean depth in those regions, we optimize the mean depth of the main straits and the bay by overall shallowing or deepening the depth in the respective straits and bay. The main straits and bay is roughly divided into four regions in the Alaska Panhandle: Glacier Bay, Lynn Canal/Chatham Strait, Frederick Sound/Stevens Passage, and Clarence Strait (Fig. 1). In the process of the optimization, the depth in the respective regions is sequentially given by perturbation terms multiplied by the original GINA depth. Since the coefficients of the dissipation including the bottom friction are usually assumed to be constant in ocean models, γ_b is uniformly optimized over the computational domain. Thus a degree of freedom of the optimization is limited to be strictly small and then the dimension of the parameter space is reduced to be five. The set of $\lfloor \alpha_j \mid j = 1, \cdots, 5 \rfloor$ is the parameter vector θ to be optimized. In the cases of $j = 1, \cdots, 4$, α_js are positive values which are multiplied by the original GINA depth in the respective four regions. If α_j is greater (smaller) than 1, the updated depth is deeper (shallower) than the original GINA bathymetry in the region. Since errors in the GINA bathymetry probably depends on locations, the set of $\lfloor \alpha_j \mid j = 1, \cdots, 4 \rfloor$ is given by the multipliers rather than differences in depth. A common logarithm of γ_b ($\log \gamma_b$) gives α_5 which is a negative value because γ_b basically takes 10^{-3}–10^{-1}. The modeled parameter is sequentially updated in the process of EA according to the objective function.

6. Results and discussions

The result of the optimization by the EA is shown in Fig. 2. The objective function and the parameter both seem to converge through iterations of the hundredth generation. The residual mostly converges after the 30th generation. The bathymetry in Clarence Strait is gradually modified and attains depth eight times as much as the original GINA bathymetry. In Glacier Bay, Chatham Strait/Lynn Canal, and Frederick Sound/Stevens Passage, the optimized depth is less than the depth twice as much as GINA. The result indicates that the bathymetric error of GINA is quite large in Clarence Strait compared in other areas. The optimized friction coefficient is around −2.5 in the log scale, that is, γ_b = 0.003, which is less than that obtained by I09.

The finally optimized bathymetry is compared to the MB data in terms of the mean depth in the respective regions (Table 1). Although the GINA bathymetry represents an unrealistically shallow Clarence Strait, the EA efficiently modifies the mean depth in the strait. The EA also modifies the mean depth to be realistic in Glacier Bay, Frederick Sound/Stevens Passage, and Chatham Strait/Lynn Canal.

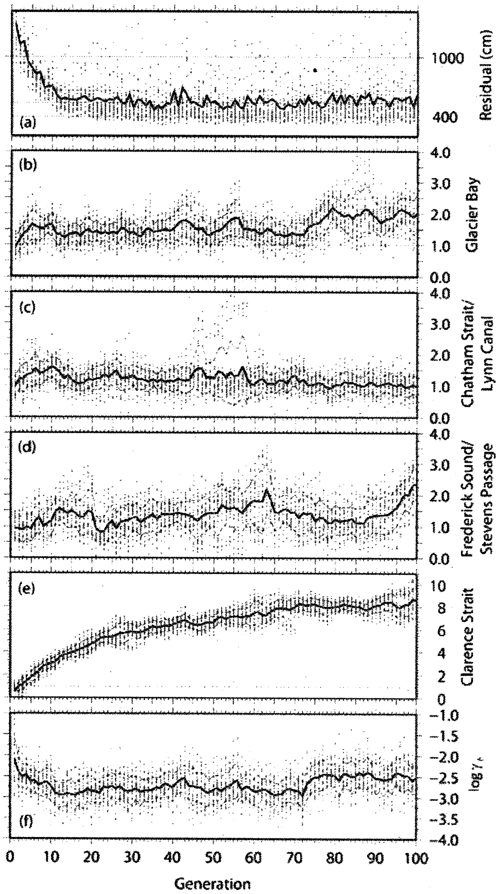

Figure 2. Evolution of the objective function and the parameter for each optimal parameter candidate through the optimization. (a) Residual denotes $L(\theta)^{-1}$ in centimeters instead of the objective function itself. The vector $\lfloor \alpha_j \mid j = 1, \cdots, 5 \rfloor$ is plotted in (b)–(f). Solid and dashed lines indicate mean and standard deviation for each generation.

Table 1. Mean depth of the MB data and GINA averaged over the respective rectangular regions in Fig. 1, and ratios of the depth of MB to GINA and of the optimized solution (Opt.) to GINA $\lfloor \alpha_j \mid j = 1, \cdots, 4 \rfloor$.

Region	MB (m)	GINA (m)	MB/GINA	Opt./GINA (EA)
Glacier Bay	112	68	1.6	2.0
Chatham Strait/Lynn Canal	218	212	1.0	1.0
Frederick Sound/Stevens Passage	170	141	1.2	1.5
Clarence Strait	218	50	4.4	8.0

7. Summary

We have used an EA to optimize boundary condition and physical parameter for an accurate ocean tide modeling. The optimized bathymetry becomes more realistic than the original satellite-derived bathymetry. Although the EA is very convenient for implementation, it is shown that the target parameter is reasonably optimized through the reduction of the degree of freedom of the parameter in Section 5. Similar procedure is expected to be available in other coastal and marginal seas such as Yellow/East China Seas, Okhotsk Sea, Indonesian Archipelago, and Arctic/Antarctic Oceans in which accuracy of the bathymetric data is suspicious (e.g., Tanaka et al., 2008).

Acknowledgements

The simulation code of the ocean model used in the present study was provided by Dr. Naoki Hirose in Kyushu University, Japan. The data of tide gauge observations in Glacier Bay was provided by Dr. Christopher F. Larsen in University of Alaska Fairbanks, U.S. The authors also wish to thank Dr. Tadahiro Sato in Tohoku University, Japan, for his numerical processing of the tidal data in Glacier Bay. Most calculations were performed on NEC SX-6 system in The Institute of Statistical Mathematics. Funding was provided by a CREST (Core Research for Evolutional Science and Technology) program of Japan Science and Technology Agency.

References

Hirose, N., Journal of Oceanography, 61, pp. 789–794. (2005)

Inazu, D., et al., Journal of Oceanography, accepted.

Lindquist, K. G., et al., EOS, Transactions American Geophysical Union, 85, pp. 186. (2004)

Losch, M. and Wunsch, C., Journal of Atmospheric and Oceanic Technology, 20, pp. 1685–1696. (2003)

Munk, W. H., and Cartwright, D. E., Philosophical Transactions of the Royal Society of London, A59, pp. 533–581. (1966)

Sandwell, D. T. and Smith, W. H. F., pp. 441–457. In Satellite Altimetry and Earth Sciences, Academic Press, San Diego. (2001)

Smith, W. H. F. and Sandwell, D. T., Science, 277, pp. 1956–1962. (1997)

Sugiura, N., et al., Journal of Geophysical Research, 113, C10017, doi:10.1029/2008JC004741. (2008)

Tanaka, Y., et al., Journal of Geophysical Research, 112, C10021, doi:10.1029/2007JC004172. (2007)