Graduate School of Frontier Sciences, The University of Tokyo

Department of Socio-Cultural Environmental Studies

2022

Master's Thesis

Exploring the influence of the boundary conditions in a coastal ocean model
沿岸海域モデルへの境界条件の影響

Submitted on July 15, 2022

Advisor: Professor Jun SASAKI

Yu Jingrui

ユ　ジンルイ

47-20680

Abstract

In the area of numerical simulation, the problem of open boundary conditions can be considered one of the most significant and challenging research topics of coastal modeling because it has a great impact on the solution within the model domain. Nowadays, in most studies applying coastal ocean models in Tokyo Bay, the open boundary conditions are either an artificial algorithm related to water depth and the period of a year, or a profile function deduced from a period of observation data or even a fixed constant or a rough time resolution dataset predicted by some ocean models. These methods are not accurate enough and cannot provide the real-time change of the connection exchange between the seawater inside the bay and the seawater outside the bay. Moreover, no study compares and combines the data measured at the observation station with the open boundary data provided by other ocean models.

In this paper, the water temperature data provided by the observation station at the mouth of Tokyo Bay and the water temperature and salinity data predicted by the ocean model HYCOM were combined and compared to explore the possibility of different sources of datasets and their combination to provide open boundary conditions for a coastal model TEEM in Tokyo Bay.

By replacing the artificial open boundary conditions of water temperature and

salinity set in TEEM with the high time resolution water temperature data measured by the observation station and the high time resolution salinity data predicted by HYCOM, the accuracy of the simulated water temperature and salinity inside Tokyo Bay was greatly improved, with the RRE of simulated water temperature decreased from 0.1207 to 0.0873, and the RRE of simulated salinity decreased from 0.2177 to 0.1943. Compared with only using open boundary datasets provided by HYCOM, the combination of the water temperature data measured by the observation station and the salinity data predicted by the ocean model can provide more accurate open boundary conditions. When the open boundary conditions provided by HYCOM are embedded with the model in the bay, attention should be paid to the way the two models with different water depth settings are coupled at the open boundary.

This study provides a way for the combination of the observation dataset and the dataset predicted by other ocean models to provide the open boundary conditions for the coastal model applied in Tokyo Bay, which can reflect the real situation of the seawater in the bay varied with the seawater outside the bay.

Acknowledgment

I am so grateful to Professor Jun Sasaki, the senior students, Wang Yulong, Wang Kangnian, Xu Rui, Endo sang, Li Peiran, and my lab mates who gave me a hand when I was very helpless and need the other's help. I have struggled in my research for two years. I have figured out something and learned much. I also experienced the process of exploring something by myself and finally, I found something is too difficult for me.

But I will applaud myself because I made it to the end. It is hard to live abroad alone. I haven't seen my family for three years because of the corona. I miss them very much and thanks for their support and understanding.

Tables of Contents

Figures and tables

Chapter 1 Introduction

In the area of numerical simulation, the problem of open boundary conditions can be considered one of the most significant and challenging branches of coastal modeling. One reason is that in coastal modeling, open boundary conditions greatly impact the solution of the simulated state variables in the inner domain of the study area. It is widely known that the open boundary conditions could lead to ill-posed problems (Orlanski, 1976), however, the prescribed open boundary conditions scheme doesn't exist.

As an ideal open boundary condition, on one hand, it should be transparent to those disturbances generated within the domain of modeled area allowing them to pass through the open boundary without significant distortion and cannot influence the solution within the domain of the ocean model. On the other hand, the solution inside the domain should reflect the disturbances of the external fields.

Tokyo Bay is a semi-enclosed bay, the freshwater discharge from rivers and offshore seawater from the bay mouth drive the estuarine circulation in the bay. With the rapid urbanization of cities around the bay, nutrient loading discharged into the bay from river runoff causes eutrophication. The bottom hypoxic water full of toxic hydrogen sulfide, upwell to the surface by wind, often happens in summer, kills fish and shellfish.

There have been many ocean models applied in Tokyo Bay to predict the

hydrodynamic process and water quality. Tabeta & Fujino (1996) used a fixed vertical profile of salinity estimated from the observed data inside Tokyo Bay because the observed data outside the bay were few. Sato et al., (2012) imposed the nutrient concentrations and dissolved oxygen based on the results of once-a-month field measurements at the open bay mouth. Sohma et al., (2018) specified the open boundary conditions of water temperature and salinity with the prescribed functions for the inflow from the outside and free-stream conditions for the outflow from the inside area. Liu et al., (2022) utilized the Blumberg and Khanta implicit condition as the open boundary conditions for storm surge simulation, tsunami simulation, and river flood simulation. This boundary condition is a modified version of the traditional radiation boundary condition. Aoki et al., (2022) applied lateral boundary conditions and the vertical profiles of water temperature and salinity were determined by the mean values observed for the bay mouth around ten years.

Nowadays, in most studies applying coastal ocean models in Tokyo Bay, the open boundary conditions are either an artificial algorithm related to water depth and the period of a year, or a profile function deduced from a period of observation data or even a fixed constant. These methods are not accurate enough and cannot provide the real-time change of the connection exchange between the seawater inside the bay and the seawater outside the bay. Some modelers used the ocean model (such as the JCOPE2 ocean model) for the Tokyo Bay coast model bay mouth open boundary condition, but the time resolution of the provided dataset is not high, a

day for one record (Masunaga et al., 2018). Moreover, no study compares and combines the data measured at the observation station with the open boundary data provided by other ocean models.

HYCOM is an eddy-resolving, real-time global and basin scale ocean prediction system (Chassignet et al., 2007). It takes the hybrid coordinate, which is isopycnic in the open ocean, but smoothly reverts to a terrain-following coordinate in the shallow coastal area and becomes z-level coordinates in the mixed layer. One of the most important capabilities of HYCOM is to provide boundary conditions for the higher-resolution regional and coastal models. Most of the existing studies using the HYCOM ocean model to provide open boundaries have focused on the Atlantic coast (Chassignet et al., 2003; Halliwell, 2004; Hyun-Sook Kim, 2013), while the open boundary conditions for coastal models along the North Pacific coast are very limited.

TEEM model is an integrated, layer-resolved, process-based, sediment-water coupled, long-term robust, and three-dimensional ecosystem model applied in Tokyo Bay (Amunugama & Sasaki, 2018). This model could be used to calculate the concentration of various state variables in seawater and sediment and reproduce their long-term biogeochemical processes. It is an important coastal ecosystem model in Tokyo Bay.

Water temperature and salinity are two of the most important and basic parameters in coastal modeling (Ji, 2017). Water temperature affects the stratification and

further influence the vertical mixing in the water system. Dissolved oxygen solubility is largely controlled by water temperature. Besides, many biochemical and physiological processes are also sensitive to water temperature. Salinity is one of the most important factors affecting estuarine circulation and often changes the stratification at estuaries more effectively than water temperature. In addition, the variation of salinity also has a great impact on the marine life that lives near estuaries. However, the open boundary conditions of water temperature and salinity set in TEEM model at the mouth of the bay are assumed artificial open boundary conditions, which cannot reflect the real-time influence of the seawater inside Tokyo Bay by the seawater outside the bay. This limits the authenticity and accuracy of the simulation results of the water temperature and salinity in the bay and further affects the accuracy of the simulation results of other variables in water and sediment.

In this thesis, the water temperature data provided by the observation station at the mouth of Tokyo Bay and the water temperature and salinity data predicted by the ocean model HYCOM were combined and compared to explore the possibility of different sources of datasets and their combination to provide open boundary conditions for the coastal model in Tokyo Bay. The functionality of coastal model TEEM was improved by replacing the artificial open boundary conditions with real-time changing water temperature and salinity datasets in high time resolution. In addition, the uncertainty caused by different embeddings of the coastal model and data obtained from the ocean model at the open boundary to the result accuracy

in Tokyo Bay was also figured out.

Chapter 2 Materials and Methods

2.1 The framework of the study

The present work includes three parts, the first part creates an interface for TEEM model to the open ocean on the bay mouth. The second part integrates the water temperature and salinity dataset predicted by HYCOM or measured at the observation station with TEEM model. In this part, the two methods for integrating datasets the rom ocean model with different settings in sigma layers were proposed and different groups of datasets were tested. The verification of the accuracy of simulation results were also calculated compared with the observation data. The third part includes the sensitivity of TEEM to changes in open boundary conditions and the visualization of the sensitivity analysis results. The schematic figure of the framework is shown in Figure 2.1.1.

2.2 Research area

Tokyo bay is a semi-enclosed bay with a mean depth of 19m. The freshwater discharge from rivers and offshore seawater from the bay mouth drives estuarine circulation in the bay. And the residence time of water in Tokyo Bay is 1-3 months. With the rapid urbanization of cities around the bay, nutrient loading discharged into the bay along with rivers causes eutrophication. That is the reason for algae bloom, which we often refer them as red tide. The species of algae blooming in

Tokyo Bay are mainly dinoflagellates and diatoms.

Figure 2.1. 1 The framework of the study

The bottom hypoxic water which is full of toxic hydrogen sulfide, upwell to the surface by wind, often happens in summer, and kills fish and shellfish. That is called blue tide.

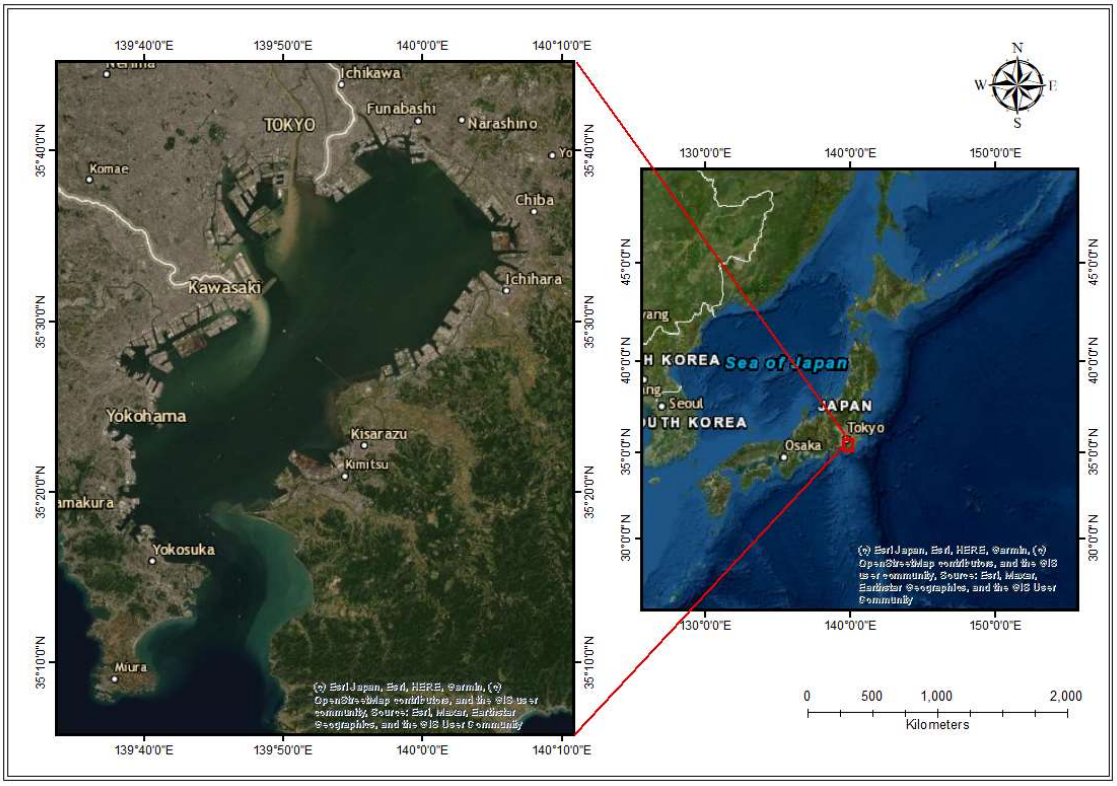

Figure 2.1. 2 The research area

2.2.1 The grid system of TEEM

To apply TEEM model in Tokyo Bay, the horizontal resolution is set as $2000\text{m} \times 2000\text{m}$ for every grid for both the pelagic model and the benthic model. In the vertical direction, the cartesian coordinate is transformed into a sigma coordinate for all state variables. The method for transformation is the same as equation1. There are 20 sigma layers in the pelagic model and 25 sigma layers in

the benthic model, respectively.

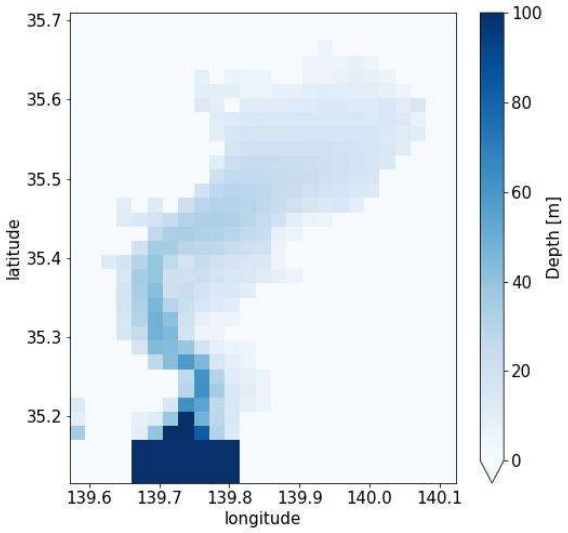

Figure 2.2. 1 The grid system set in TEEM

2.3 Model description and the driving data for TEEM model

2.3.1 Model description

TEEM model is a three-dimensional, layer-nested, water-sediment coupled, process-based numerical model (Amunugama & Sasaki, 2018),it integrates physical processes, biological processes, and chemical processes in seawater and sediment. All the processes are considered independently and coupled on their interface - the bottom of the water and the surface of the bed.

The main state variables in the water include labile organic carbon, refractory organic carbon, inert organic carbon, dissolved inorganic carbon, pCO2, and total alkalinity; dissolved oxygen; three communities of phytoplankton, each of them

blooms in different seasons; zooplankton; nutrients, such as Ammonium, Nitrate, Phosphate; sulfide; silica; and the suspended matter. The main state variables in the sediment are labile organic carbon, refractory organic carbon, inert organic carbon; dissolved oxygen; nutrients; sulfide; silica; and silt. The porosity changes with computing, thus the thickness of the sediment layer also changes with every time step. In the interface of the water column and the bed, the diffusion of dissolved nutrients and dissolved oxygen are undergoing computing, the particulate matter also continues to settle from water to bed or suspend from the bed to the above water column.

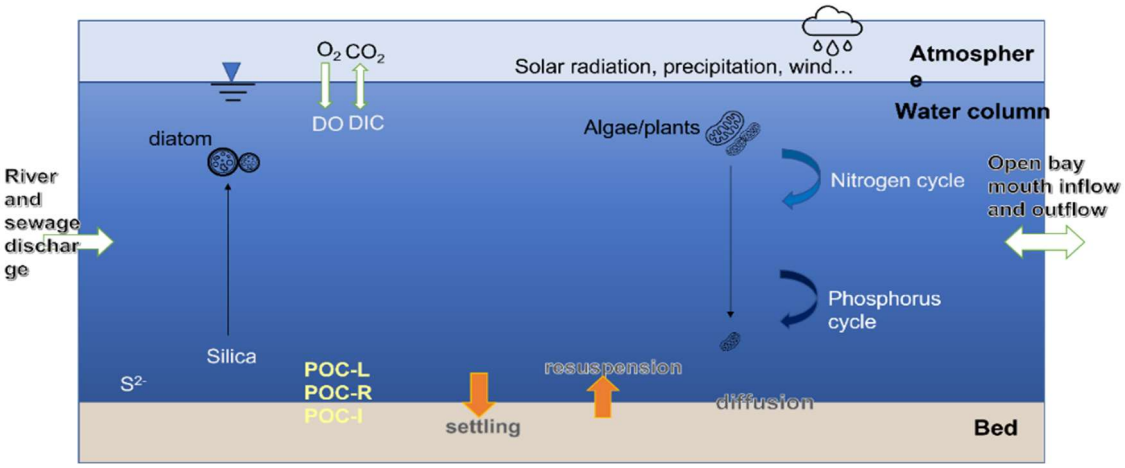

Figure 2.3. 1 The schematic of TEEM model

2.3.2 The driving data used as boundary conditions

The boundary conditions are set at the interface of surface seawater and atmosphere; the bottom of the bed; the river mouth and the sewage mouth on the bay head; and the open boundary conditions set at the bay mouth.

The surface boundary conditions are one of the main driving forces of the model, including air temperature, air pressure, vapor pressure, relative humidity, wind speed, wind direction, cloud cover, solar radiation, and precipitation. The dataset used as surface boundary conditions in the model is the GWO hourly dataset observed by Japan Meteorological Agency (JMA) which is the oldest and official observatory.

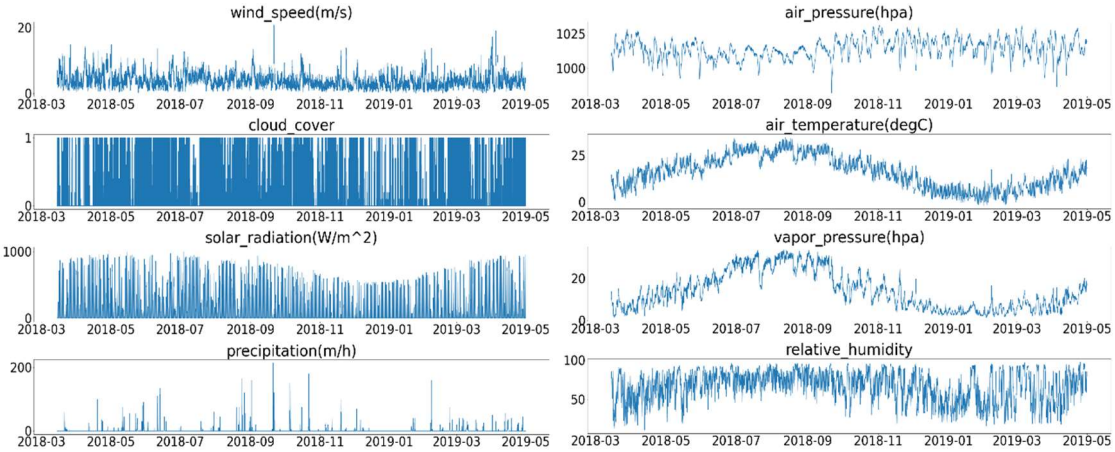

Figure 2.3. 2 Meteorological data set as surface boundary conditions

The river discharge of Arakawa is calculated with the empirical formula derived by (Suzuki, 2013), in which the discharge of main rivers flowing into Tokyo Bay is calculated by the water level observed at specific observations. The daily water level can be obtained from the water information system maintained by the Japanese government. The other rivers are obtained by the ratio of the average annual discharge of Arakawa and themselves.

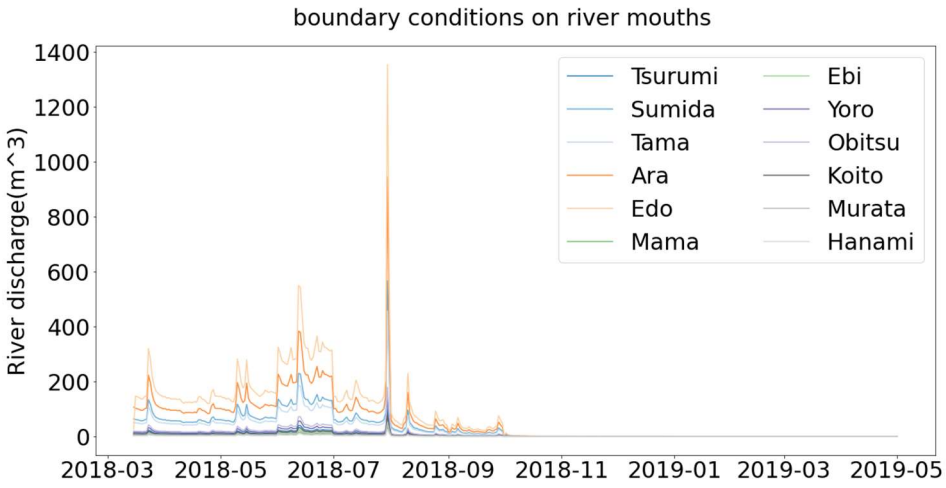

Figure 2.3. 3 The river discharge data adopted in the boundary condition system

The input data on the tide level was obtained at the Yokosuka monitoring station, which locates near the open mouth of Tokyo Bay. The data was downloaded from the website of the Japan Meteorological Agency. The original dataset has a time interval of 1 hour and then was interpolated into the time interval of 600 seconds. The part of the dataset shown in Figure 2.3.4.

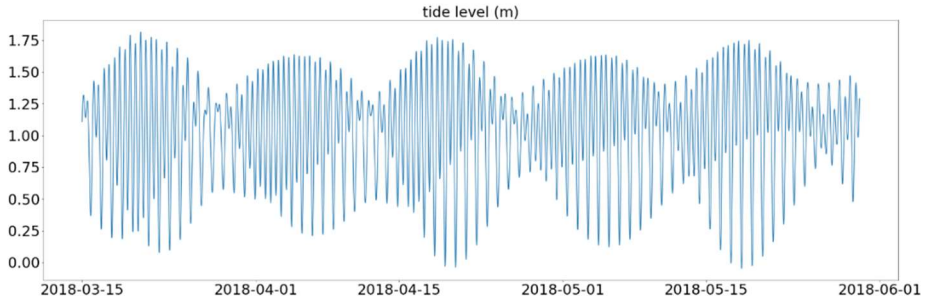

Figure 2.3. 4 The time-series tide level observed near the bay mouth.

2.4 The interface for open boundary conditions

The open boundary conditions of TEEM were set at the mouth of Tokyo Bay. The original open boundary conditions of the model are a temperature and salinity profile algorithm set artificially. In the case of OBC water temperature, the independent variables of the algorithm include the water depth, lowest temperature, which was set as 15 degrees centigrade, half of the range of temperature, one year period and the phase lag in hours which equals 0 when January 1st. In the case of salinity, the salinity at the open boundary depends only on the depth of the water with higher salinity in deeper layers. However, such artificial open boundary conditions cannot reflect the real-time changes of the seawater flowing from the open bay mouth, here, an interface was created at the open boundary to read the real-time observation data at the bay mouth.

As the water depth of the grids set on the mouth of Tokyo Bay are all set as 100m, the fluctuation of the free surface is negligible concerning this water depth in the sigma coordinate system. Thus, in the vertical profile, the water 'wall' on the bay mouth could be divided into 20 layers, which are consistent with the grids inside the bay, with each layer 5 meters.

$$\Delta depth = \frac{D}{k_{max}}$$

Where $\Delta depth$ is water depth on each layer, D is the total water depth set on the open boundary, k_{max} is the total number of layers. The time interval of the observation data set on the open boundary was interpolated to fit the marching

time of the model.

When coupled with the real-time observation dataset, the data should be assigned to the corresponding sigma layers from the surface to the bottom. Then distributed all grids in the same layer with the same value. The grids of seawater and the grids of the land were distinguished by using a parameter $mask(i,j)$.

$$mask(i,j) = \begin{cases} 1, & i \text{ is the grids of land} \\ 0, & i \text{ is the grids of sea} \end{cases}$$
$$OBC(i,j,k,t) = O(i,j,k,t) * mask(i,j)$$

Where i is the number on the x-coordinate for each grid on the horizontal plane, j is the number on the y-coordinate for each grid on the horizontal plane, j equals 1 on the open boundary, at the open boundary, $mask(i,j)$ can be denoted as $mask(i,1)$, where $j = 1$ means the grids on the south open boundary. k is the number of layers in the vertical direction, t is the marching time of TEEM model. The time interval of the observation data set on the open boundary was interpolated to fit the marching time of the model.

$OBC(i,j,k,t)$ is the data which will be assigned as open boundary condition, $O(i,j,k,t)$ is the real-time observation dataset read from the input file. $mask(i,j)$ is zero when the grid is on the land and one when grid is in the sea. The schematic of the data assigned on the open boundary marching with time is shown here. V is any variable for open boundary conditions, V(i,j,k,n) is the value of one grid for this

variable at the open boundary. i and j are grid numbers of x axis and y axis, respectively, j equals 1 at the open boundary, k is the grid number of the vertical sigma axis, and n is the timestep number. In the case of timestep is 100s and the initial time is 0s, the relationship between time (t) and timestep index (n) is when n is 0, t is 0s; When n is 1, t is 100s; When n is 2, t is 200s, and so on.

$V(i,j,k_{max},n_0)$	$V(i,j,k_{max},n_1)$	…	$V(i,j,k_{max},n_{n-1})$	$V(i,j,k_{max},n_n)$
$V(i,j,k_{max-1},n_0)$	$V(i,j,k_{max-1},n_1)$	…	$V(i,j,k_{max-1},n_{n-1})$	$V(i,j,k_{max-1},n_n)$
$V(i,j,k_{max-2},n_0)$	$V(i,j,k_{max-2},n_1)$	…	$V(i,j,k_{max-2},n_{n-1})$	$V(i,j,k_{max-2},n_n)$
…	…	…	…	…
$V(i,j,k_1,n_0)$	$V(i,j,k_1,n_1)$	…	$V(i,j,k_1,n_{n-1})$	$V(i,j,k_1,n_n)$

Figure 2.4. 1 The schematic of the data assigned on the open boundary marching with time

2.5 The preparation of dataset used as open boundary conditions

2.5.1 The water temperature and salinity data predicted by HYCOM

The water temperature and salinity data were extracted from HYCOM (Chassignet et al., 2007) from March 15th 2018 to April 30th 2019. The dataset can be downloaded on their website (https://www.hycom.org/hycom/overview). The data sets are not evenly distributed vertically, with more dense measurements in the upper ocean

and sparser measurements in the lower layers. The horizontal resolution is 0.08° for each grid near the mouth of Tokyo Bay. The time interval for each piece of the dataset is three hours.

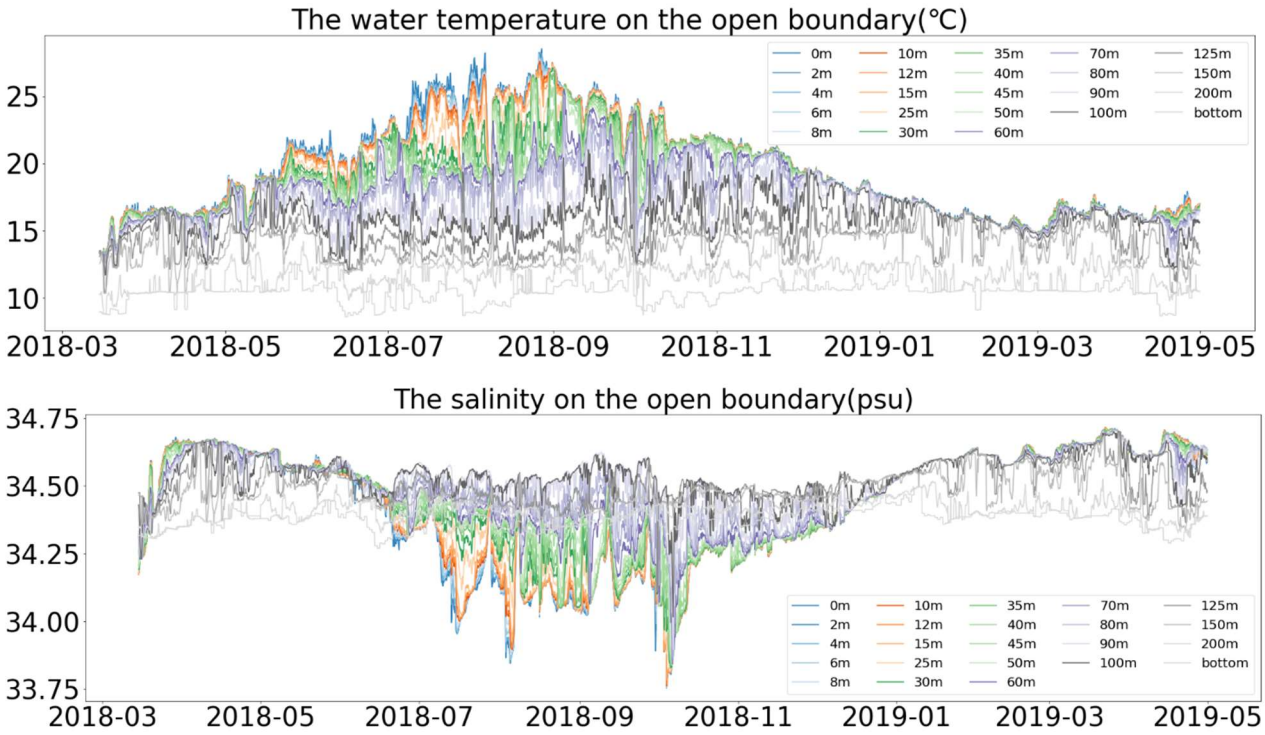

Figure 2.5. 1 The water temperature and salinity dataset predicted by HYCOM

2.5.2 The water temperature data measured by observation station

The temperature for the open boundary condition is set as the specified boundary condition. The time-series water temperature data, observed by Kaneda observation (35°09′31",139°41′42") which locates at the open mouth of Tokyo Bay, owns the water temperature observation records every twenty minutes. It has observations every five meters from the sea surface to a depth of 35 meters.

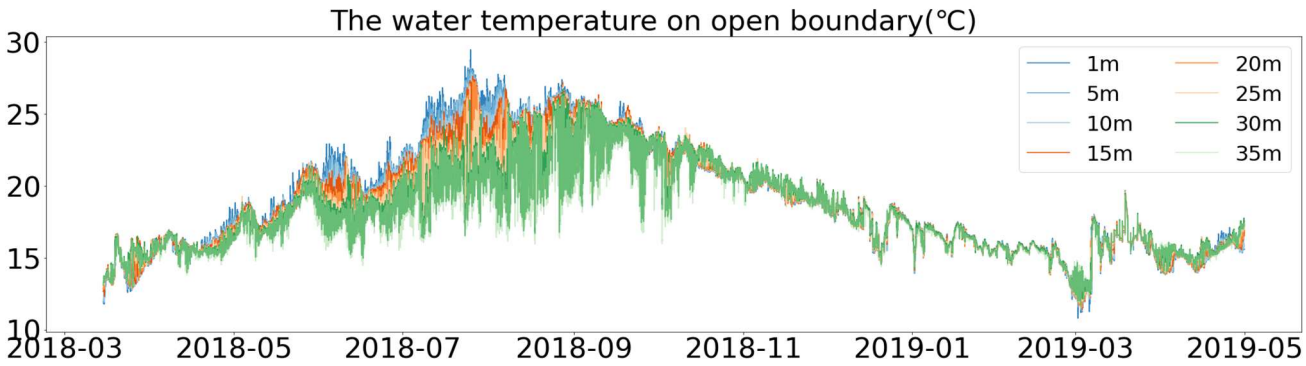

Figure 2.5. 2 The water temperature data measured at observation station

2.6 The method for coupling dataset to TEEM model

2.6.1 The setting method in vertical direction

In the vertical direction, the model was evenly divided into 20 layers in the vertical direction, using sigma coordinate system, while HYCOM uses a Cartesian coordinate system in the vertical direction near the mouth of Tokyo Bay, and the predicted data are dense at the surface of the depth, with one record every two meters, and sparse at the depth, with only one record every 10 or even 50 meters. In addition to the different coordinate system, HYCOM has a different water depth for each grid at the open boundary, the water depth at the center of the bay mouth is more than 200 meters, while the water depth for each grid is 100 meters here for TEEM. Thus, the two kinds of boundaries need to be aligned. There are two types of solution applied to solve this problem. One is to select only part of the data, and another is to compress the data in the corresponding layer.

For the first solution, only takes the first 100 meters of the dataset predicted by

HYCOM, and then converts the dataset to the sigma coordinate, and assigns the value to the corresponding layer on the open boundary, with 5 meters each layer. For the second solution, divide 200 meters into 20 layers, 10 meters each. The data of corresponding depth is distributed into corresponding layers. If there are multiple data in a layer, the value of this layer is the average value of these data. The data of the bottom layer is obtained by averaging the water temperature or salinity at 200 meters and the bottom water temperature or salinity value predicted by HYCOM.

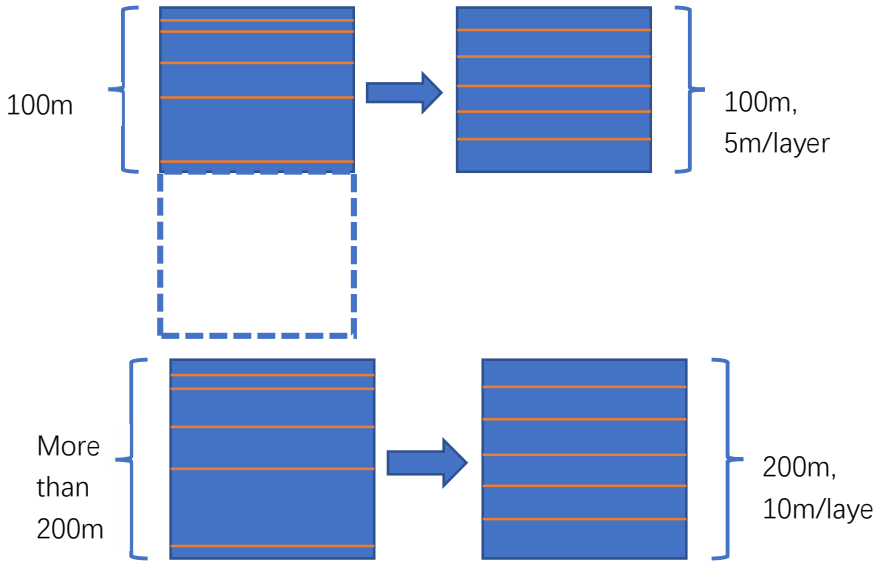

Figure 2.6. 1 The schematic for the setting of layers on the open boundary

2.6.2 The dataset groups used as open boundary conditions

The groups of datasets used as open boundary conditions were selected to simulate water temperature and salinity in TEEM. There are four groups only change the OBC water temperature: case 1, water temperature dataset measured by the observation station with original artificial salinity open boundary conditions; case 2, water

temperature dataset predicted by HYCOM above 100 meters with artificial salinity algorithm; case 3, water temperature dataset predicted by HYCOM compressed in all depths.

There are two groups only change the OBC salinity: in case 1, the original OBC water temperature was kept but changed the OBC salinity with the salinity predicted by HYCOM within 100m of the water surface; in case 2, the original OBC water temperature was kept and used the salinity dataset predicted by HYCOM but compressed in all depth.

There are six groups changing the water temperature and salinity open boundary conditions: case1, water temperature dataset predicted by observation station and salinity dataset predicted by HYCOM within 100m; case2, water temperature datasets measured at the observation station and salinity dataset predicted by HYCOM compressed in all depths; case3, water temperature dataset predicted by HYCOM compressed in all depths and salinity predicted by HYCOM within 100m; case4, water temperature dataset and salinity dataset predicted by HYCOM within 100m; case5, water temperature predicted by HYCOM within 100m and salinity dataset predicted by HYCOM compressed in all depths; case6, water temperature and salinity datasets predicted by HYCOM compressed in all depth.

2.7 Statistical parameters used to verify the accuracy of the simulation results

The statistical parameters used to verify the accuracy of the simulation results are

mean absolute error (MAE), root mean square error (RMSE), R^2, and RRE which is the ratio between the RMSE and the difference between the maximum and minimum values of the observed data. These parameters are calculated as:

$$MAE = \frac{1}{N}\sum_{n=1}^{N}|O^n - P^n|$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(O^n - P^n)^2}$$

$$RRE = \frac{RMSE}{O_{max} - O_{min}} \times 100$$

Where n is the number of data pair, O is the observation data, P is the simulated data, O_{max} is the maximum value of the observation data, O_{min} is the minimum value of the observation data.

R^2 was obtained from the regression model of the least square method for the observed and simulated data.

Chapter 3 Results

3.1 Error analysis of the results in the upper layer, middle layer, and bottom layer respectively

Error analysis of observed and modeled water temperature and salinity from 2018.03.15 to 2019.04.30 were conducted at Kawasaki station (35°29′25" N,139°50′02" E) which locates at the center of Tokyo Bay.

The statistical variables used to calculate the accuracy of the simulation results are the mean value of the observation data, the mean value of the modeled results, the mean absolute error, the root mean square error, RRE, and R square.

3.1.1 Water temperature datasets with original artificial salinity algorithm

There are three kinds of water temperature datasets: water temperature datasets measured by observation station, water temperature datasets predicted by HYCOM above 100 meters, and water temperature datasets predicted by HYCOM compressed in all depths.

The four cases are shown here: case 1, original artificial open boundary conditions for both water temperature and salinity; case 2, water temperature dataset measured by the observation station with original artificial salinity open boundary conditions; case 3, water temperature dataset predicted by HYCOM above 100 meters with artificial salinity algorithm; case 4, water temperature dataset predicted

by HYCOM compressed in all depths. The results of the four cases are shown here.

Table 3.1.1. 1 Error analysis of observed and modeled water temperature at Kawasaki station from 2018-03-15 to 2019-04-30 in case1.

layer	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max -ob mean)
up	18.58	16.6612	2.0132	2.456	21.25	0.1155
mid	17.69	15.3026	2.4364	2.9415	18.32	0.1605
bot	17.01	15.0115	2.1237	2.6767	16.45	0.1627

In the case of original open boundary settings, the mean value of the modeled water temperature is lower than the water temperature measured in the observation station for the three layers. The mean absolute errors of the three layers are more than 2. The RMSE of the middle layer is the largest and the RMSE of the upper layer is the smallest.

Table 3.1.1. 2 Error analysis of observed and modeled water temperature at Kawasaki station from 2018-03-15 to 2019-04-30 in case2.

layer	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max -ob mean)
up	18.58	17.6858	1.4651	1.8952	21.25	0.0891
mid	17.69	16.9918	1.1333	1.4611	18.32	0.0797
bot	17.01	17.0429	0.9085	1.1808	16.45	0.0717

When the open boundary condition of water temperature (OBC water temperature) was substituted with the water temperature dataset measured at the observation station, the mean values of the simulated water temperature in the upper and

middle layers are lower than the measured values, while the simulated water temperature in the bottom layer is slightly higher than the measured values. The mean absolute errors are around one for the three layers. The RMSE decreases from the upper layer to the lower layers.

Table 3.1.1. 3 Error analysis of observed and modeled water temperature at Kawasaki station from 2018-03-15 to 2019-04-30 in case3.

layer	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max -ob mean)
up	18.58	17.4818	1.4234	1.8979	21.25	0.0892
mid	17.69	16.6122	1.2367	1.6764	18.32	0.0915
bot	17.01	16.544	0.9907	1.351	16.45	0.0821

When the OBC water temperature dataset was used in HYCOM simulated water temperature data within 100m of the water surface and the artificial OBC salinity was kept, the mean values of the simulated water temperature of the three layers are lower than the observed values of the water temperature of the corresponding layers. The RMSE of the upper layer is larger than that of the middle layer and larger than that of the bottom layer.

Table 3.1.1. 4 Error analysis of observed and modeled water temperature at Kawasaki station from 2018-03-15 to 2019-04-30 in case4.

layer	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max -ob mean)
up	18.58	16.244	2.3946	2.8332	21.25	0.1332
mid	17.69	14.6881	3.0464	3.512	18.32	0.1917
bot	17.01	14.1068	2.9529	3.4132	16.45	0.2074

When the OBC water temperature dataset was used in compressed HYCOM water temperature data from the surface to bottom with the artificial OBC salinity, the simulated mean water temperature of the top layer is about 2 degrees lower than the measured value, and the middle and bottom layers are about 3 degrees lower than the measured value. The RMSE of the middle and bottom layers is higher than that of the surface layer.

3.1.2 Salinity datasets with original artificial water temperature algorithm

These are the comparisons of observed and modeled salinity at Kawasaki station from 2018-03-15 to 2019-04-30 when the artificial open boundary condition of water temperature was kept and substituted the artificial OBC salinity with the salinity datasets predicted by HYCOM. There are three cases: in case 1, the original

artificial OBC water temperature and salinity were kept; in case 2, the original OBC water temperature was kept but changed the OBC salinity with the salinity predicted by HYCOM within 100m of the water surface; in case 3, the original OBC water temperature was kept and used the salinity dataset predicted by HYCOM but compressed in all depth. The results of the three cases are shown in Table 3.1.2.1, Table 3.1.2.2, and Table 3.1.2.3.

Table 3.1.2. 1 Error analysis of observed and modeled salinity at Kawasaki station from 2018-03-15 to 2019-04-30 in case1.

layer	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max -ob mean)
up	30.59	27.9681	2.9211	3.8416	8.55	0.4493
mid	31.96	31.0211	1.2925	1.5711	6.55	0.2398
bot	33.25	32.3705	1.1564	1.3982	7.29	0.1918

In case 1, The mean value of simulated salinity data of each layer is lower than that of the measured data. In terms of RMSE, the salinity of the upper layer has the lowest simulation accuracy. The mean absolute error of the upper layer is about 3, while the maximum difference of salinity measured by the observation station is 8.55 in one year. The simulation accuracy of salinity in the surface layer is very low.

Table 3.1.2. 2 Error analysis of observed and modeled salinity at Kawasaki station from 2018-03-15 to 2019-04-30 in case2.

layer	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max -ob mean)
up	30.59	28.344	2.6648	3.5581	8.55	0.4161
mid	31.96	31.426	1.1137	1.3733	6.55	0.2096
bot	33.25	32.7523	0.9152	1.2003	7.29	0.1646

When the OBC salinity was changed to the salinity datasets predicted by HYCOM within 100m of the water's surface and kept the original OBC water temperature, the simulated salinity values of the three layers are all lower than the measured values. The difference in the top layer is the largest (RMSE 3.55), and the difference in the bottom layer is relatively small (RMSE 1.2).

Table 3.1.2. 3 Error analysis of observed and modeled salinity at Kawasaki station from 2018-03-15 to 2019-04-30 in case3.

layer	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max -ob mean)
up	30.59	28.355	2.6508	3.5613	8.55	0.4165
mid	31.96	31.4342	1.1105	1.3702	6.55	0.2091
bot	33.25	32.7567	0.9072	1.1899	7.29	0.1632

In case 3, the OBC water temperature was the same as the original case but changed the OBC salinity with the compressed salinity dataset predicted by HYCOM. The errors of this case are almost the same as that of case2, and the distribution of RMSE in the vertical direction is also consistent, with the RMSE in the upper layer being

higher than that in the middle layer and higher than that in the bottom layer.

3.1.3 Water temperature datasets with salinity datasets

The cases of grouping the OBC water temperature dataset and OBC salinity dataset were also conducted. There are three kinds of OBC water temperature, water temperature measured at the observation station, water temperature predicted by HYCOM within 100m of the sea surface, and water temperature predicted by HYCOM compressed in the vertical direction. There are two kinds of OBC salinity, salinity dataset predicted by HYCOM within 100m of the sea surface, and salinity predicted by HYCOM compressed in all depths.

Free combination of any two data sets, plus an original base case, a total of seven cases were obtained: case1, original artificial open boundary conditions of water temperature and salinity; case2, water temperature dataset predicted by observation station and salinity dataset predicted by HYCOM within 100m; case3, water temperature datasets measured at the observation station and salinity dataset predicted by HYCOM compressed in all depths; case4, water temperature dataset predicted by HYCOM compressed in all depths and salinity predicted by HYCOM within 100m; case5, water temperature dataset and salinity dataset predicted by HYCOM within 100m; case6, water temperature predicted by HYCOM within 100m and salinity dataset predicted by HYCOM compressed in all depths; case7, water temperature and salinity datasets predicted by HYCOM compressed in all depth.

Table 3.1.3. 1 Error analysis of observed and modeled water temperature and salinity at Kawasaki station from 2018-03-15 to 2019-04-30 in case1.

variable	layer	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max - ob mean)
water temperature	up	18.58	16.6612	2.0132	2.456	21.25	0.1155
	mid	17.69	15.3026	2.4364	2.9415	18.32	0.1605
	bot	17.01	15.0115	2.1237	2.6767	16.45	0.1627
salinity	up	30.59	27.9681	2.9211	3.8416	8.55	0.4493
	mid	31.96	31.0211	1.2925	1.5711	6.55	0.2398
	bot	33.25	32.3705	1.1564	1.3982	7.29	0.1918

In case1, the annual average of the simulated water temperature is about 2 degrees lower than the predicted value for each layer. The RMSE of surface temperature was the lowest.

The largest difference between simulated and measured salinity was found in the surface layer, where the RMSE of the surface layer was three times that of the bottom layer.

Table 3.1.3. 2 Error analysis of observed and modeled water temperature and salinity at Kawasaki station from 2018-03-15 to 2019-04-30 in case2.

variable	layer	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max -ob mean)
water temperature	up	18.58	18.6204	1.5134	1.8992	21.25	0.0893
	mid	17.69	17.6652	1.7433	2.0666	18.32	0.1128
	bot	17.01	17.5066	1.2021	1.5992	16.45	0.0972
salinity	up	30.59	28.0171	2.8591	3.5807	8.55	0.4188
	mid	31.96	32.3504	0.7015	1.0437	6.55	0.1593
	bot	33.25	33.2119	0.6765	1.0085	7.29	0.1383

In case2, the annual mean of the simulated water temperature is very close to the annual mean of the measured water temperature in the upper, middle, and bottom layers, and the MAEs of the three layers are within 2 degrees. The mean value of simulated salinity exceeds the measured data in the middle layer, is similar to the measured data in the bottom layer, and is still low in the top layer. The largest error among the three layers is at the top, where the RMSE is more than three times that of the bottom layer.

Table 3.1.3. 3 Error analysis of observed and modeled water temperature and salinity at Kawasaki station from 2018-03-15 to 2019-04-30 in case3.

variable	layer	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max - ob mean)
water temperature	up	18.58	17.7292	1.4639	1.8664	21.25	0.0877
	mid	17.69	16.9879	1.1499	1.4834	18.32	0.0809
	bot	17.01	17.0249	0.8898	1.1716	16.45	0.0712
salinity	up	30.59	28.3774	2.5884	3.5675	8.55	0.4172
	mid	31.96	31.3552	1.0534	1.3578	6.55	0.2073
	bot	33.25	32.726	0.8847	1.1934	7.29	0.1637

In case3, from the perspective of MAE, the simulated water temperature of the three layers are about 1 degree different from the measured value, and the RMSE is the largest in the top layer and the smallest in the bottom layer. The simulated salinity of the three layers is all less than the measured data, with the largest difference at the top and the smallest difference at the bottom.

Table 3.1.3. 4 Error analysis of observed and modeled water temperature and salinity at Kawasaki station from 2018-03-15 to 2019-04-30 in case4.

variable	layer	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max - ob mean)
water temperature	up	18.58	16.2836	2.3577	2.8153	21.25	0.1324
	mid	17.69	14.7162	3.0178	3.5539	18.32	0.1939
	bot	17.01	14.1811	2.902	3.43	16.45	0.2085
salinity	up	30.59	27.917	2.9639	3.8416	8.55	0.4493
	mid	31.96	30.9565	1.342	1.6115	6.55	0.246
	bot	33.25	32.304	1.2224	1.4611	7.29	0.2004

In case4, the water temperature dataset predicted by HYCOM compressed in all depths and salinity predicted by HYCOM within 100m were used as the open boundary conditions.

The mean of the simulated water temperature is about 2 degrees lower than the mean of the measured water temperature in the upper layer, and about 3 degrees lower in the middle and lower layers. The maximum error of simulated water temperature occurs in the middle layer, with an RMSE of about 3.5. The mean values of simulated salinity were lower in each layer than the measured values, with the largest error occurring in the top layer, with an RMSE of about 3.8.

Table 3.1.3. 5 Error analysis of observed and modeled water temperature and salinity at Kawasaki station from 2018-03-15 to 2019-04-30 in case5.

variable	layer	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max - ob mean)
water temperature	up	18.58	17.5757	1.3767	1.8538	21.25	0.0871
	mid	17.69	16.7124	1.1815	1.6572	18.32	0.0904
	bot	17.01	16.695	0.9495	1.3222	16.45	0.0803
salinity	up	30.59	28.138	2.7699	3.7069	8.55	0.4335
	mid	31.96	31.1134	1.2189	1.4993	6.55	0.2289
	bot	33.25	32.509	1.0441	1.3261	7.29	0.1819

In case 5, the water temperature dataset and salinity dataset predicted by HYCOM within 100m were used as the open boundary conditions. The average simulated water temperature in each layer is about one degree lower than the measured value. The upper layer has the largest error (RMSE) of 1.8, while the lower layer has the smallest error (RMSE) of 1.3. The annual mean of the simulated salinity values was lower than the measured values in each layer, with the largest error occurring in the upper layer, where the RMSE is 3.7, about three times that of the lower layer.

Table 3.1.3. 6 Error analysis of observed and modeled water temperature and salinity at Kawasaki station from 2018-03-15 to 2019-04-30 in case6.

variable	layer	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max - ob mean)
water temperature	up	18.58	17.532	1.3941	1.8728	21.25	0.088
	mid	17.69	16.6525	1.2204	1.7025	18.32	0.0929
	bot	17.01	16.628	0.9874	1.3611	16.45	0.0827
salinity	up	30.59	28.1482	2.7612	3.7013	8.55	0.4329
	mid	31.96	31.1348	1.1981	1.4791	6.55	0.2258
	bot	33.25	32.5219	1.0273	1.306	7.29	0.1791

In case6, water temperature predicted by HYCOM within 100m and salinity dataset predicted by HYCOM compressed in all depths were set as the open boundary conditions for water temperature and salinity in TEEM. The average value of the simulated water temperature is about one degree lower than the measured value in each layer, and the error of the bottom layer is the smallest, with an RMSE of about 1.3. The simulated salinity was lower than observed in every layer, with the top layer having the largest error, with an RMSE of 3.7, three times that of the bottom layer.

Table 3.1.3. 7 Error analysis of observed and modeled water temperature and salinity at Kawasaki station from 2018-03-15 to 2019-04-30 in case7.

variable	layer	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max - ob mean)
water temperature	up	18.58	16.2663	2.374	2.8346	21.25	0.1333
	mid	17.69	14.688	3.046	3.6047	18.32	0.1967
	bot	17.01	14.1569	2.9298	3.4803	16.45	0.2115
salinity	up	30.59	27.9456	2.9381	3.8234	8.55	0.4471
	mid	31.96	30.9927	1.3112	1.5823	6.55	0.2415
	bot	33.25	32.3363	1.1926	1.4289	7.29	0.196

In case7, the open boundary conditions for water temperature and salinity were equipped with the water temperature and salinity datasets predicted by HYCOM compressed in all depths. The average simulated water temperature is about 2 degrees lower than the measured one at the top and about 3 degrees lower at both the middle and bottom layers. The error of the simulated water temperature in the middle layer is the largest, with an RMSE of 3.6. The average absolute error between the simulated salinity value and the measured value is 2.9 in the top layer, and the RMSE is 3.8, which is the largest error among the three layers.

To make the comparison of simulation results of each case more intuitive, the comparison diagrams of MAE, RMSE and RRE for water temperature and salinity on surface, middle and bottom layer of the seven cases are shown below. Where, the X-axis represents the case numbers from case1 to case7, and the Y-axis represents the statistical parameters MAE, RMSE or RRE. The blue line is the simulation results

of the upper layer, the orange line is the simulation results of the middle layer, and the gray line is the simulation results of the bottom layer.

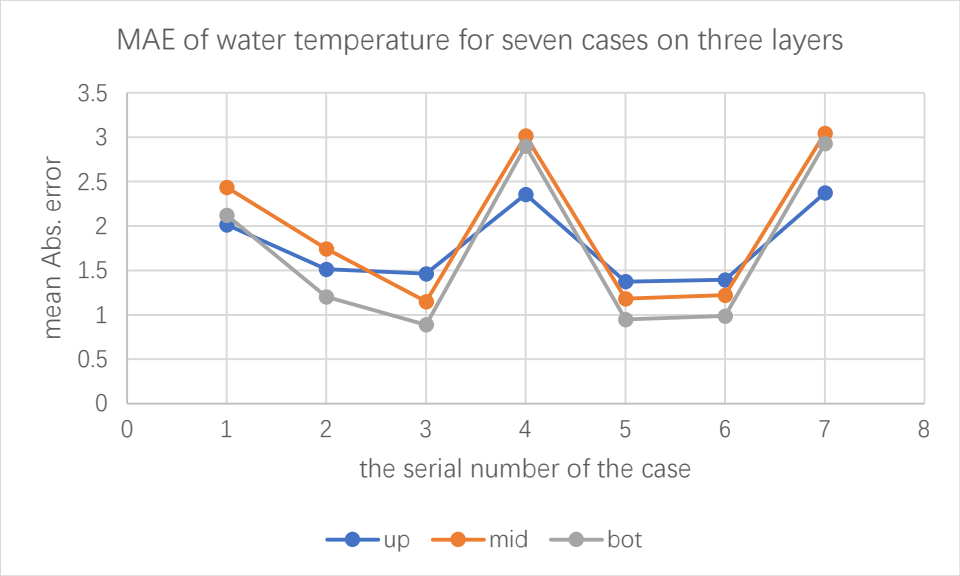

Figure 3.1.3. 1 MAE of water temperature for seven cases on three layers

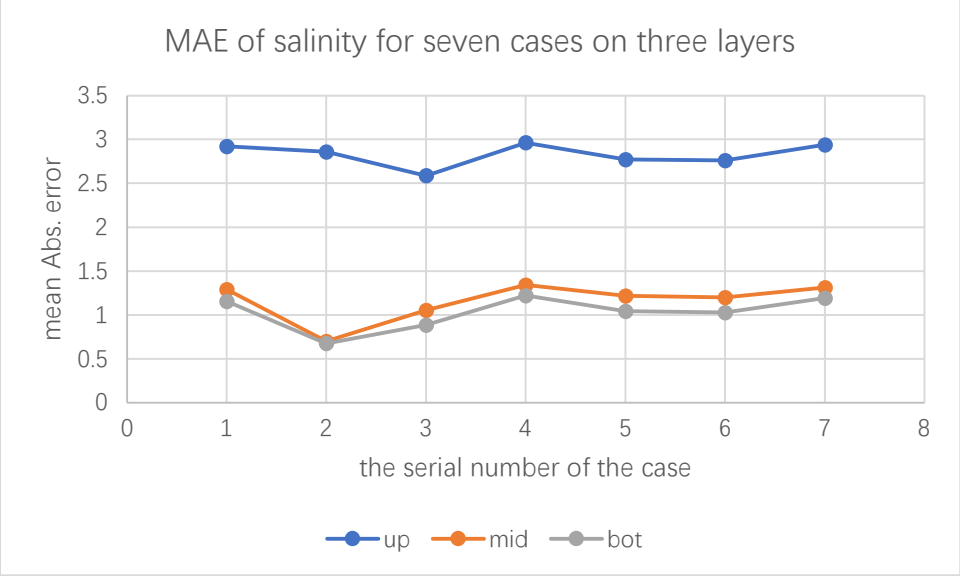

Figure 3.1.3. 2 MAE of salinity for seven cases on three layers

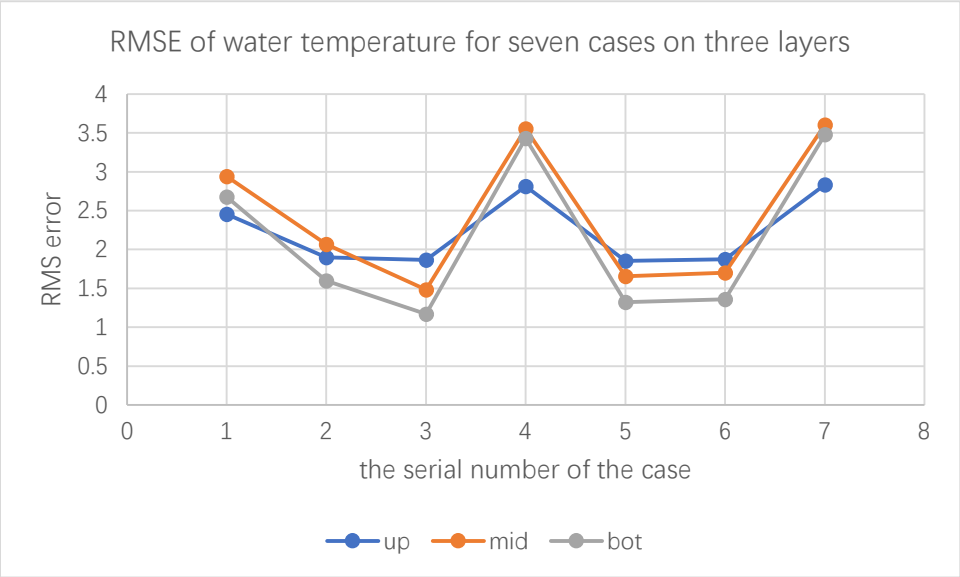

Figure 3.1.3. 3 RMSE of water temperature for seven cases on three layers

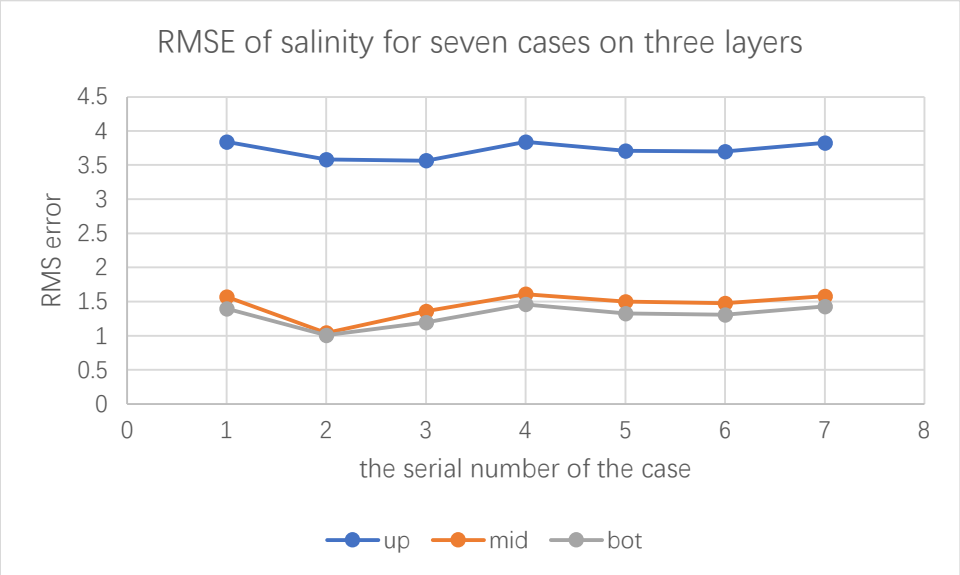

Figure 3.1.3. 4 RMSE of salinity for seven cases on three layers

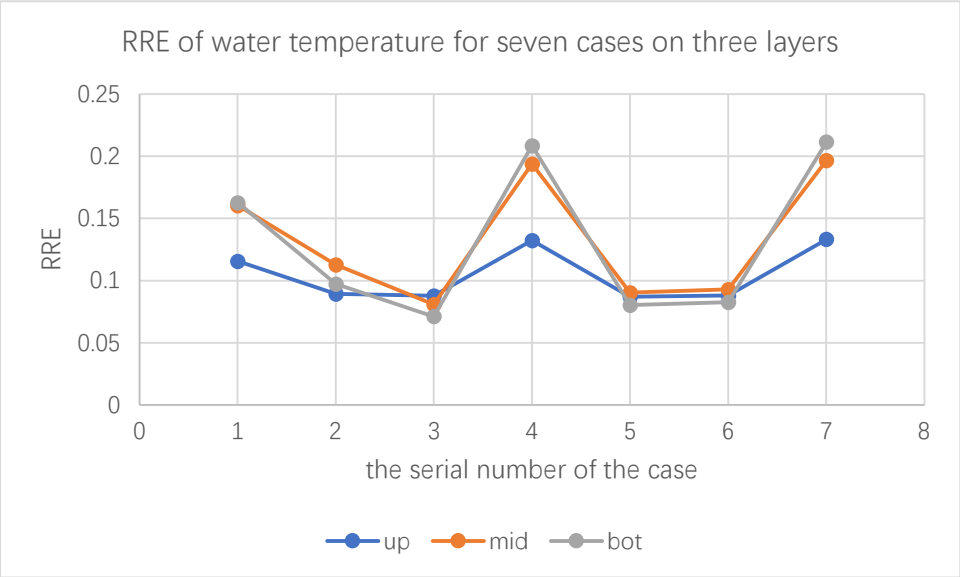

Figure 3.1.3. 5 RRE of water temperature for seven cases on three layers

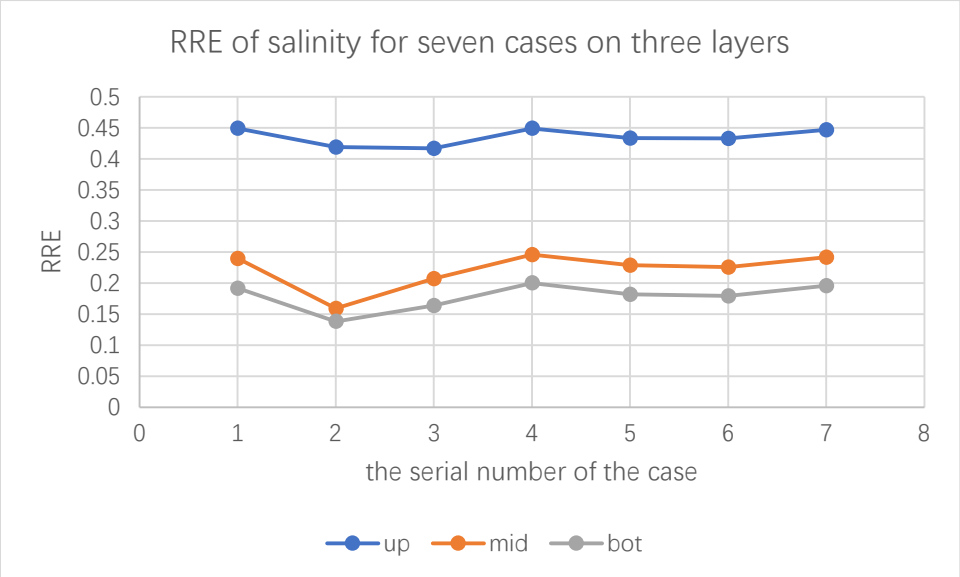

Figure 3.1.3. 6 RRE of salinity for seven cases on three layers

3.2 Error analysis of the results in three layers at two observation stations

Error analysis of observed and modeled water temperature and salinity from 2018.03.15 to 2019.04.30 were conducted at two stations- Kemegawa station

(35°29′25" N,139°50′02" E) which locates at the head of Tokyo Bay and Kawasaki station (35°29′25" N,139°50′02" E) which locates at the center of Tokyo Bay. Compared with Kawasaki station, Kemegawa station is further away from the bay mouth and influenced less by the changes in open boundary conditions set on the bay mouth.

The cases are the same as the part 3.1.3: case1, original artificial open boundary conditions of water temperature and salinity; case2, water temperature dataset predicted by observation station and salinity dataset predicted by HYCOM within 100m; case3, water temperature datasets measured at the observation station and salinity dataset predicted by HYCOM compressed in all depths; case4, water temperature dataset predicted by HYCOM compressed in all depths and salinity predicted by HYCOM within 100m; case5, water temperature dataset and salinity dataset predicted by HYCOM within 100m; case6, water temperature predicted by HYCOM within 100m and salinity dataset predicted by HYCOM compressed in all depths; case7, water temperature and salinity datasets predicted by HYCOM compressed in all depth.

The comparison results are shown here.

Table 3.2. 1 Error analysis of observed and modeled water temperature and salinity at Kemegawa and Kawasaki stations from 2018-03-15 to 2019-04-30 in case1.

station	variable	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max -ob mean)	R^2
Kemegawa	temperature	17.26	15.1854	2.1245	2.6871	22.25	0.1207	0.935
	Salinity	30.78	29.798	1.3521	1.9231	8.82	0.2177	0.343
kawasaki	Temperature	17.76	15.6584	2.1911	2.6987	21.25	0.1269	0.91
	salinity	31.93	30.4532	1.79	2.5286	10.51	0.2405	0.414

Table 3.2. 2 Error analysis of observed and modeled water temperature and salinity at Kemegawa and Kawasaki station from 2018-03-15 to 2019-04-30 in case2.

station	variable	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max -ob mean)	R^2
Kemegawa	temperature	17.26	17.3485	1.9864	2.4681	22.25	0.1109	0.83
	Salinity	30.78	30.6627	1.18	1.8279	8.82	0.207	0.267
kawasaki	Temperature	17.76	17.9307	1.4863	1.865	21.25	0.0877	0.872
	salinity	31.93	31.1931	1.4124	2.2307	10.51	0.2122	0.45

Table 3.2. 3 Error analysis of observed and modeled water temperature and salinity at Kemegawa and Kawasaki station from 2018-03-15 to 2019-04-30 in case3.

station	variable	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max -ob mean)	R^2
Kemegawa	temperature	17.26	16.3969	1.4488	1.9525	22.25	0.0877	0.912
	Salinity	30.78	30.1042	1.1465	1.7156	8.82	0.1943	0.367
kawasaki	Temperature	17.76	17.2473	1.1679	1.5337	21.25	0.0721	0.922
	salinity	31.93	30.8195	1.5088	2.309	10.51	0.2197	0.426

Table 3.2. 4 Error analysis of observed and modeled water temperature and salinity at Kemegawa and Kawasaki station from 2018-03-15 to 2019-04-30 in case4.

station	variable	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max -ob mean)	R^2
Kemegawa	temperature	17.26	14.8021	2.4896	3.08104	22.25	0.1384	0.926
	Salinity	30.78	29.7078	1.4205	1.9508	8.82	0.2209	0.348
kawasaki	Temperature	17.76	15.0603	2.7592	3.2823	21.25	0.1543	0.892
	salinity	31.93	30.3925	1.8428	2.5489	10.51	0.2425	0.416

Table 3.2. 5 Error analysis of observed and modeled water temperature and salinity at Kemegawa and Kawasaki station from 2018-03-15 to 2019-04-30 in case5.

station	variable	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max -ob mean)	R^2
Kemegawa	temperature	17.26	16.2522	1.3425	1.9435	22.25	0.0873	0.926
	Salinity	30.78	29.8448	1.3126	1.8513	8.82	0.2096	0.362
kawasaki	Temperature	17.76	16.9944	1.1692	1.6259	21.25	0.0764	0.928
	salinity	31.93	30.5868	1.6776	2.4323	10.51	0.2314	0.425

Table 3.2. 6 Error analysis of observed and modeled water temperature and salinity at Kemegawa and Kawasaki station from 2018-03-15 to 2019-04-30 in case6.

station	variable	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max -ob mean)	R^2
Kemegawa	temperature	17.26	16.2041	1.3597	1.9616	22.25	0.0881	0.928
	Salinity	30.78	29.8651	1.2964	1.8362	8.82	0.2079	0.363
kawasaki	Temperature	17.76	16.9375	1.2006	1.6592	21.25	0.078	0.928
	salinity	31.93	30.6016	1.6622	2.4216	10.51	0.2304	0.424

Table 3.2. 7 Error analysis of observed and modeled water temperature and salinity at Kemegawa and Kawasaki station from 2018-03-15 to 2019-04-30 in case7.

station	variable	Obs. Mean	modeled mean	mean Abs. error	RMS error	obs change	RRE(RMSE/ob max -ob mean)	R^2
Kemegawa	temperature	17.26	14.7824	2.5087	3.111	22.25	0.1398	0.924
	Salinity	30.78	29.737	1.3997	1.9336	8.82	0.2189	0.349
kawasaki	Temperature	17.76	15.0371	2.7832	3.3237	21.25	0.1563	0.888
	salinity	31.93	30.4249	1.814	2.5274	10.51	0.2404	0.416

Same as in Section 3.1.3, the comparison diagrams of MAE, RMSE and RRE for water temperature and salinity of the seven cases at the two stations are shown here. The X axis represents the number of the case, from case1 to case7, and the Y axis represents the statistical parameters, MAE, RMSE, RRE or R squared. The blue line is the simulation results of Kemegawa observation station and the orange line is the simulation results of Kawasaki observation station.

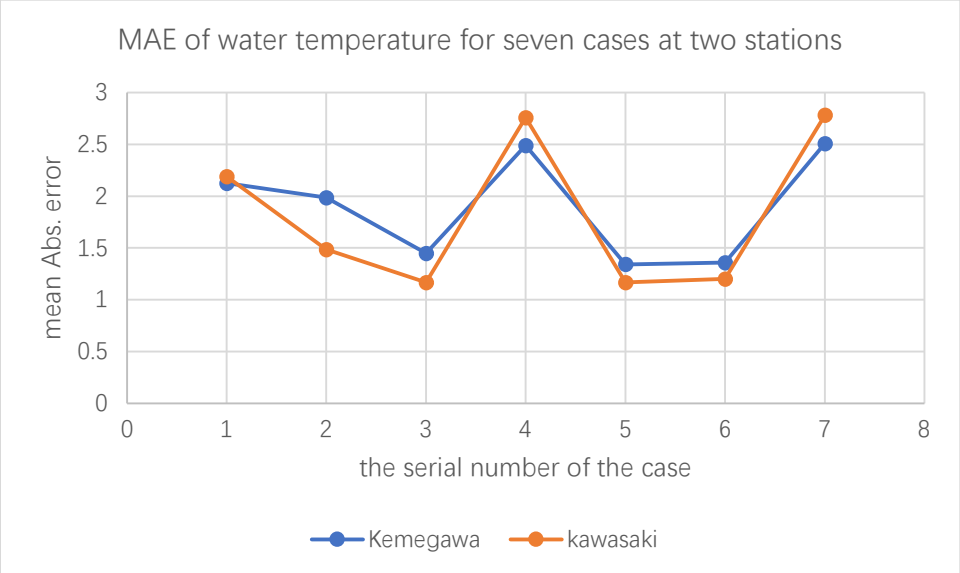

Figure 3.2. 1 MAE of water temperature for seven cases at two stations

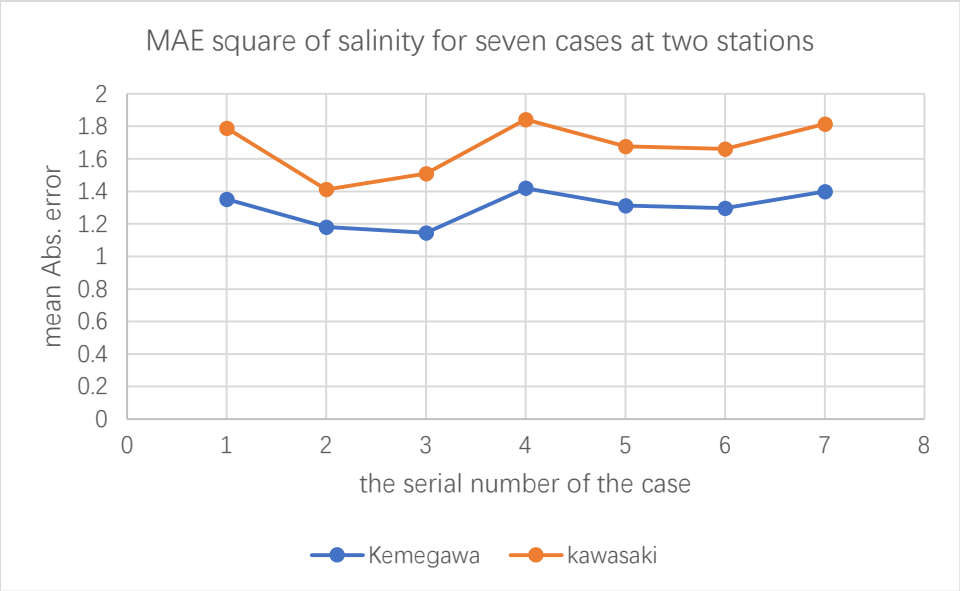

Figure 3.2. 2 MAE square of salinity for seven cases at two stations

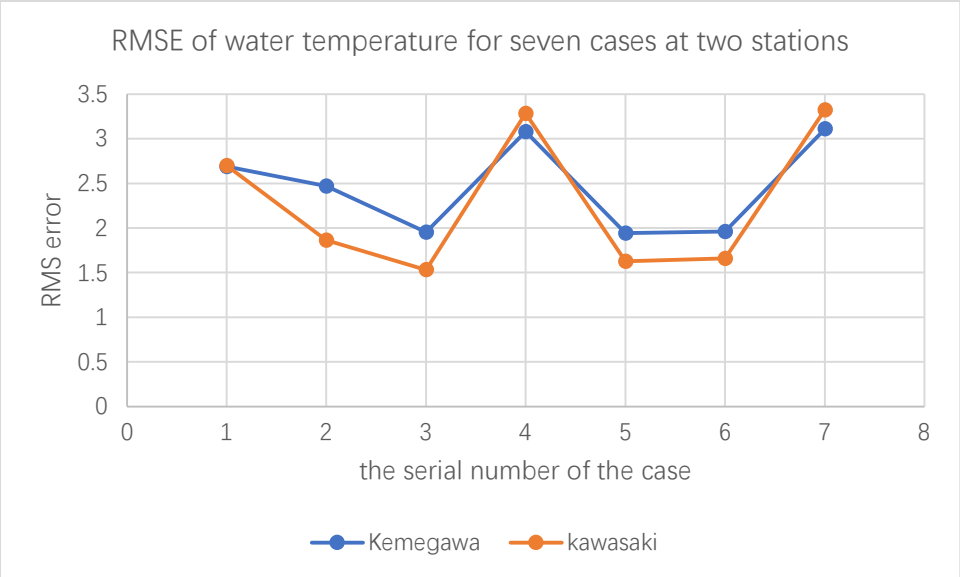

Figure 3.2. 3 RMSE of water temperature for seven cases at two stations

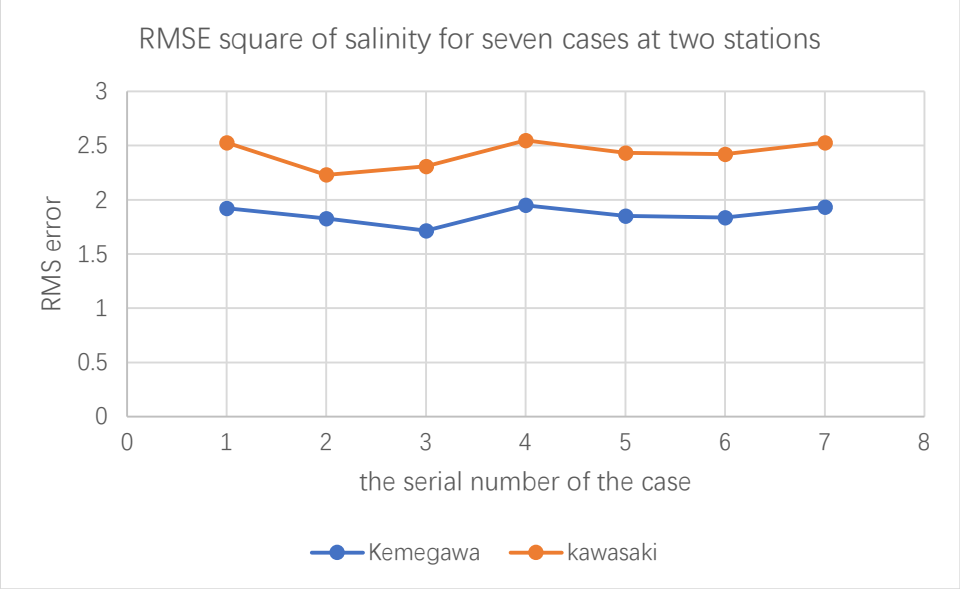

Figure 3.2. 4 RMSE square of salinity for seven cases at two stations

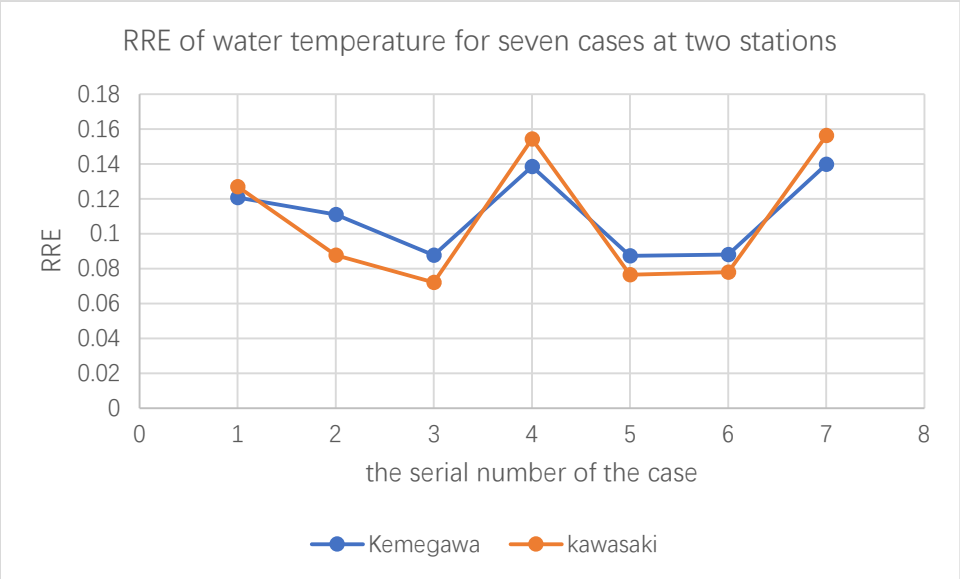

Figure 3.2. 5 RRE of water temperature for seven cases at two stations

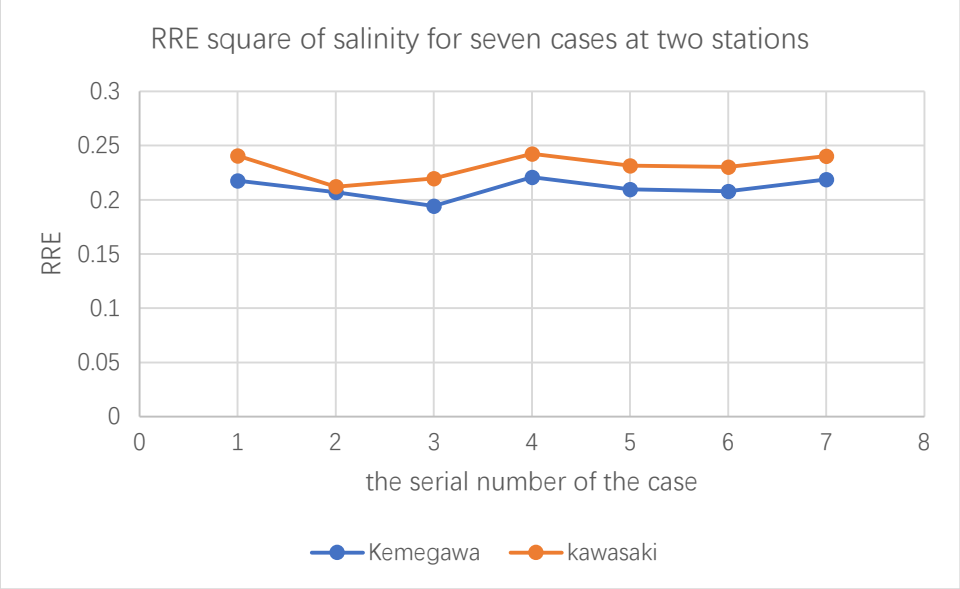

Figure 3.2. 6 RRE square of salinity for seven cases at two stations

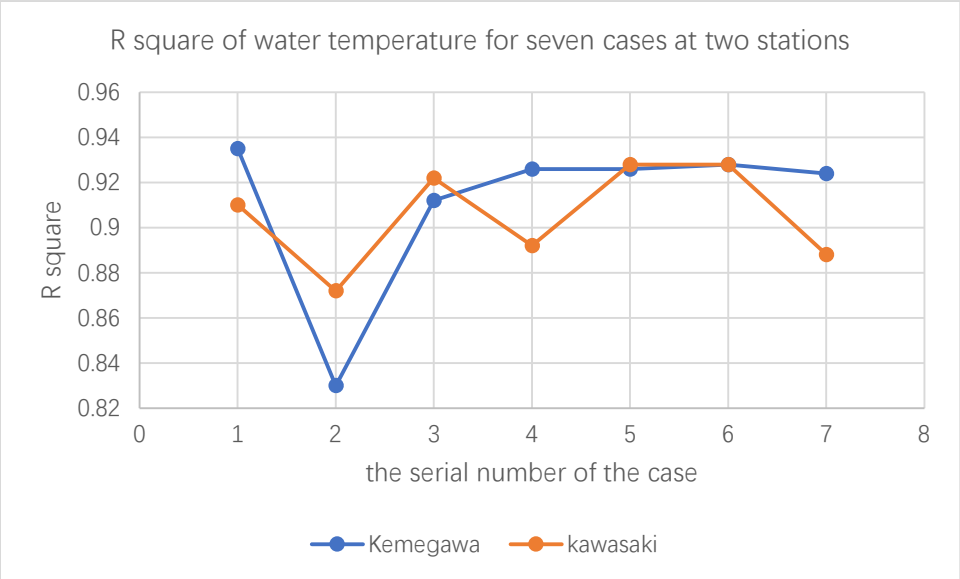

Figure 3.2. 7 R square of water temperature for seven cases at two stations

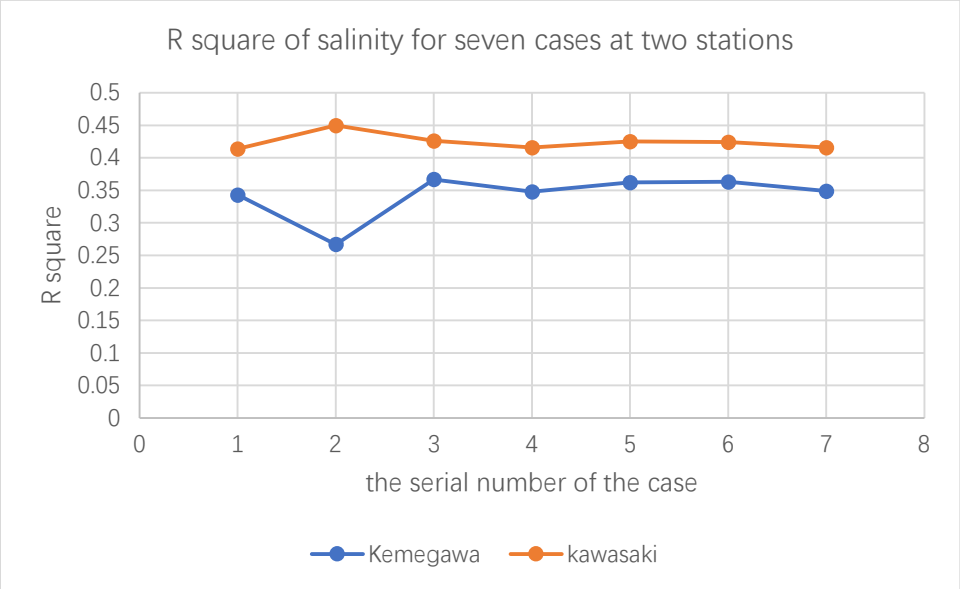

Figure 3.2. 8 R square of salinity for seven cases at two stations

For the Kemegawa observation station, the simulated water temperature of case5 is the closest to the observed value, RRE is only 0.0873, and the R square of the simulated water temperature and the observed value is 0.926, indicating a strong

correlation between the two data sets. The smallest error between simulated salinity and observed value is case5, with an RMSE of only 0.1943. The mean values of simulated temperatures are closest to the observed values for case6, and the mean values of simulated salinity are closest to the observed values for case3. The largest R square of simulated water temperature and observed value is case1, and the largest R square of simulated salinity and observed value is case3.

For the Kawasaki observation station, case3 had the smallest error between simulated water temperature and observed data, and the RRE was only 0.0721. Case2 had the smallest error between simulated salinity and observed data, and the RRE was 0.2122. The mean of simulated water temperature is closest to the mean of observed water temperature in case3, and the mean absolute error of case3 is also the smallest, which is 1.1679. The annual mean of simulated salinity is closest to the observed value in case2, with an average absolute error of 1.4863. The R square of simulated water temperature and observed water temperature is the largest in case4 and case6, both of which are 0.928. The correlation between simulated salinity and measured salinity is strongest for case2, whose R square is 0.45.

Chapter 4 Discussion

4.1 Water temperature datasets with original artificial salinity algorithm

In all cases where only the OBC water temperature was replaced while the original OBC salinity was retained, compared with the base case (3.1.1 case1), when the OBC water temperature was replaced by the water temperature data measured by the observation station (3.1.1 case2), the average simulated water temperature increased by about one degree in the upper and middle layers but is still lower than the observed value, and the lower layer water temperature is slightly higher than the observed value. In addition, only in this case is the mean of the simulated water temperature at the bottom higher than the observed mean. It may be because the water temperature below 35 meters is assigned the same value as the water temperature at 35 meters, thus the simulated water temperature is slightly higher than the observed one.

When the OBC water temperature was changed to the water temperature dataset predicted by HYCOM within 100m of the sea surface (3.1.1 case3), the mean of the simulated temperature is lower than the observed values at all layers. The mean absolute error of each layer is larger than that of the OBC water temperature replaced by the dataset measured at the observation station.

When the open boundary water temperature was replaced by compressed HYCOM water temperature data in all depths (3.1.1 case4), the mean absolute error of the

middle and lower layers is about three degrees. The error of this case is the largest among these cases, which is even larger than the error of the base case without changing the boundary conditions, especially at the bottom level.

In summary, for the modeled water temperature, keeping the artificial salinity open boundary condition unchanged, and only changing the OBC water temperature with the three kinds of datasets, the case with the smallest error is replaced with the water temperature data of the observation station, the second is replaced with the water temperature dataset predicted by HYCOM within 100m, and the worst is replaced with the HYCOM water temperature of the full-layer compression. However, even for the case with the smallest error, the average absolute error of the upper layer is still 1.42, which is the largest among the three layers. This may be due to errors caused by other boundary conditions, such as solar radiation, air temperature, vapor pressure, and so on in the surface boundary conditions, which would also affect the simulated water temperature.

In terms of the accuracy of the simulated water temperature, the simulated water temperature obtained by HYCOM's full-layer compressed water temperature is lower than that of other cases, and this difference is most prominent in the bottom layer.

4.2 Salinity datasets with original artificial water temperature algorithm

When substituting the artificial OBC salinity with the salinity dataset predicted by HYCOM within 100m of water surface (3.1.2 case2), the average simulated salinity of each layer is also lower than the observed data. But compared with the base case (3.1.2 case1), the mean absolute error of each layer is improved and closer to the observed value. The mean absolute error of the top layer is reduced by 0.3 compared with the base. However, the RRE at the top is still high, at more than 40%. When the artificial open boundary condition for salinity was changed to the salinity dataset predicted by HYCOM compressed in all depths (3.1.2 case3), compared with the previous two examples, the mean of each layer of salinity is increased, and the mean absolute error is decreased. Among the cases where only the salinity on the open boundary was replaced, this case simulates the salinity best.

The best of the three cases is the one that uses the HYCOM salinity dataset with full-depth compression. In this case, the bottom layer has the smallest error and the top layer has the largest error.

Unlike the three statistics, the mean value, MAE and RMSE, all give absolute values of the simulation-observation discrepancies, RRE is often used to express the relative discrepancies to measure the model performance. It is defined as the ratio of RMSE to the observed change. However, even in the case with the smallest error of simulated salinity effect, the RRE of simulated salinity and measured data on the surface layer still exceeds 0.4, which may be caused by errors brought by other

boundary conditions. Such as errors in river runoff estimates, or uncertainties in wind speed and precipitation in surface boundary conditions.

4.3 Water temperature datasets with salinity datasets

Among the cases where OBC water temperature was equipped with water temperature measured at the observation station and the salinity datasets predicted by HYCOM, for error, the RMSE of simulated water temperature and observed water temperature is smaller in 3.1.3 case3 (water temperature datasets measured at the observation station and salinity dataset predicted by HYCOM compressed in all depths), especially the RMSEs of middle and bottom layers are much smaller. However, the average water temperature in 3.1.3 case2 (water temperature dataset predicted by observation station and salinity dataset predicted by HYCOM within 100m) is closer to the average of the measured water temperature, for each layer.

For the error between simulated salinity and observed values, the error in the surface layer of case3 is lower than that in the surface layer of case2, but the error in the middle and bottom layers of case3 is higher than that in the middle and bottom layers of case2. This is different from changing the OBC salinity without changing the OBC water temperature.

For these two cases, as far as the observation error is concerned, Kawasaki's RRE between simulated salinity and observed values in case3 and case2 is slightly larger than that in case2.

However, the mean values of the simulated salinity in case2 are closer to the observed values. Using RRE to compare the simulated salinity accuracy of the two observation stations, Kemegawa's simulated salinity accuracy in case3 is the highest. To compare which open boundary conditions simulate better water temperature and salinity, all layers need to be compared together. When the two models are set in different water depths or different numbers of sigma layers in the vertical direction, it is necessary to discuss how to dock the open boundary to achieve the best simulation effect of the model inside the bay. For example, the simulated salinity at the bottom has a smaller error than at the middle and top, while the average simulated salinity at the three layers is all lower than the observed values, thus the OBC salinity may need to be pushed up from the bottom layer. For HYCOM, the error of the simulated water temperature at the upper layer is smaller than that at the bottom and middle layers. When the open boundary water temperature is changed from the water temperature dataset HYCOM within 100m of the water surface to the water temperature predicted by HYCOM compressed in all depths, the error of the simulated water temperature at the bottom is further increased, and the modeled mean value is also reduced. Therefore, the temperature value of the HYCOM dataset at the upper layer needs to be expanded to the lower layer.

4.4 Time series comparison of simulated water temperature and simulated salinity of three cases with observed data at two observation stations

For the simulated water temperature, the time series plots for three cases were chosen: case0, original artificial open boundary conditions of water temperature and salinity; case1, water temperature datasets measured at the observation station and salinity dataset predicted by HYCOM compressed in all depths; case3, water temperature predicted by HYCOM within 100m and salinity dataset predicted by HYCOM compressed in all depths.

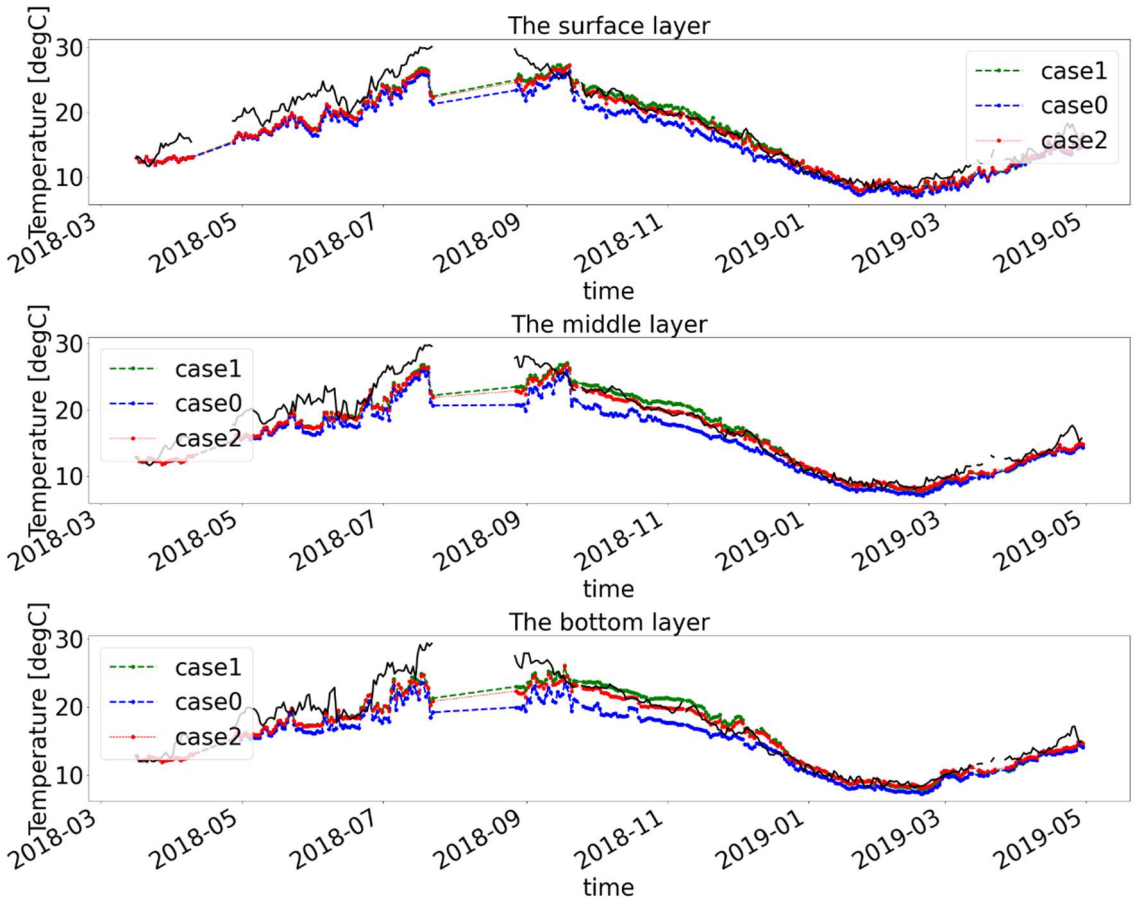

Figure 4.4. 1 The water temperature comparison at Kemegawa station from 2018-03-15 to 2019-04-30

Figure 4.4.1 shows the simulated water temperature of three layers for the three cases compared with the observation data. The line in black is the observation data measured at Kemegawa station, and the blue line is for case0, the green line is for case1 and the red line is for case2. We can see that the difference between the three cases is the largest from September 2018 to January 2019, among which the water temperature simulated by case0 is the lowest, the water temperature simulated by case1 is the highest, and the water temperature simulated by case2 is the closest to the measured value during this period.

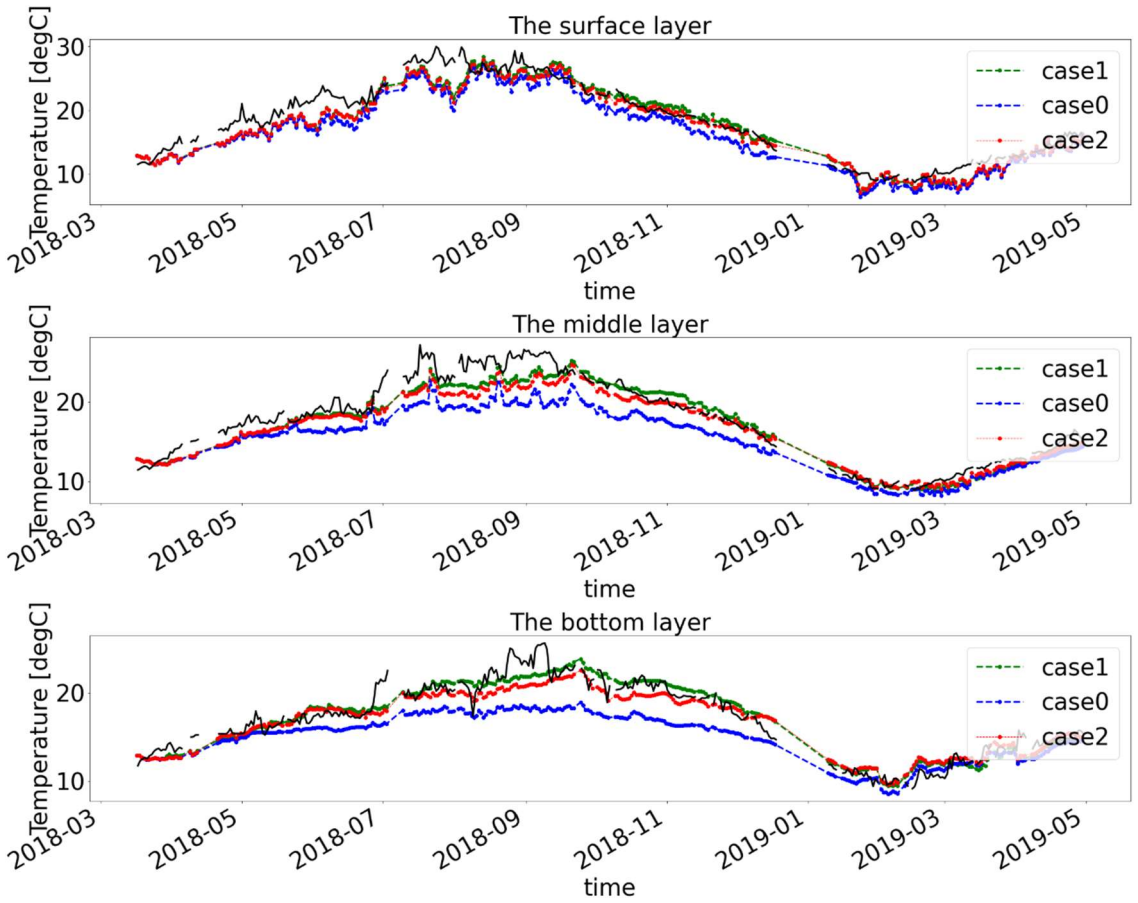

Figure 4.4. 2 The water temperature comparison at Kawasaki station from 2018-03-15 to 2019-04-30

Figure 4.4.2 shows the simulated water temperature of three layers for the three cases compared with the observation data. The time series plot of the simulated water temperature of the three cases shows that the difference in the simulated water temperature in the three cases is the smallest in the top layer and the largest in the bottom layer. The simulated water temperature of case1 is higher than that of case2, and the simulated water temperature of the case0 is the lowest.

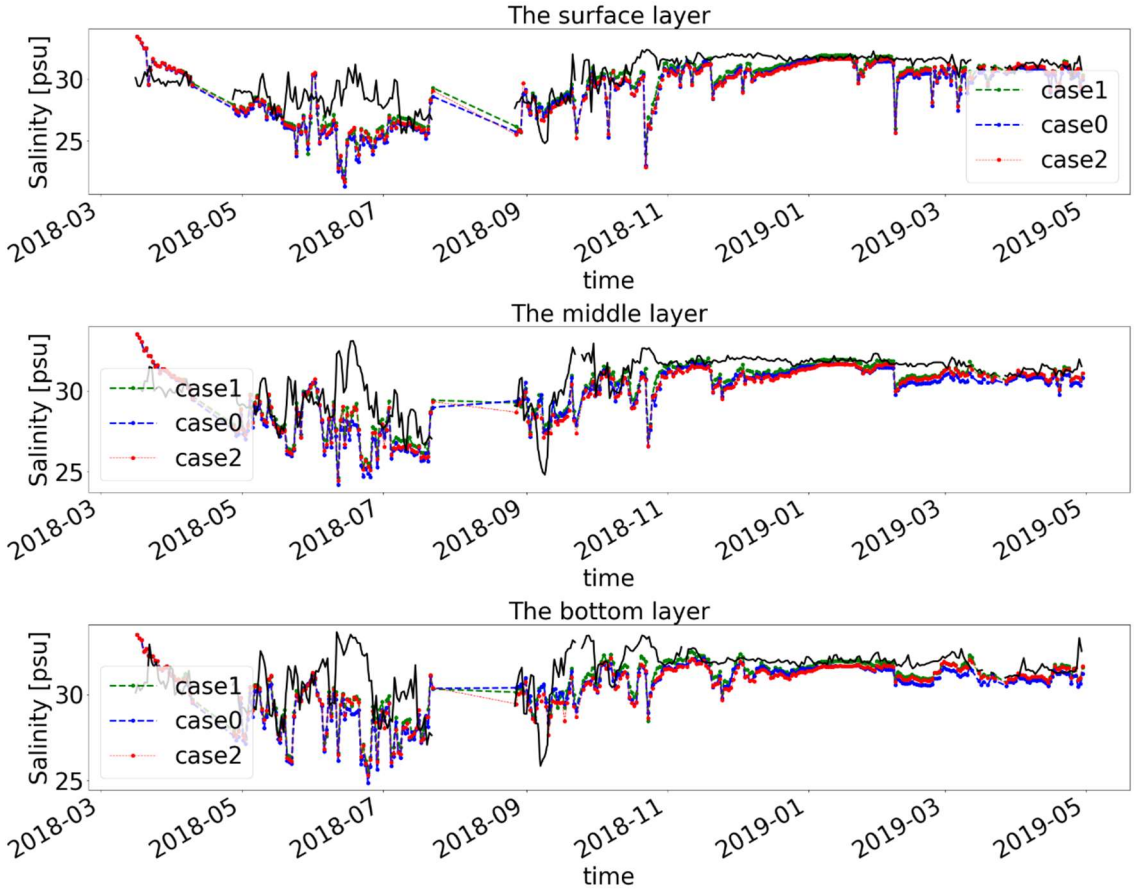

Figure 4.4. 3 The simulated salinity comparison at Kemegawa station from 2018-03-15 to 2019-04-30

The line in black is the observation data measured at Kemegawa station, and the blue line is for case0, the green line is for case1 and the red line is for case2. The difference in simulated salinity of the three cases in this observation station is not obvious. From November 2018 to January 2019, the simulated salinity value of the case2 was the smallest among the three cases, but it exceeded the other two cases after January 2019.

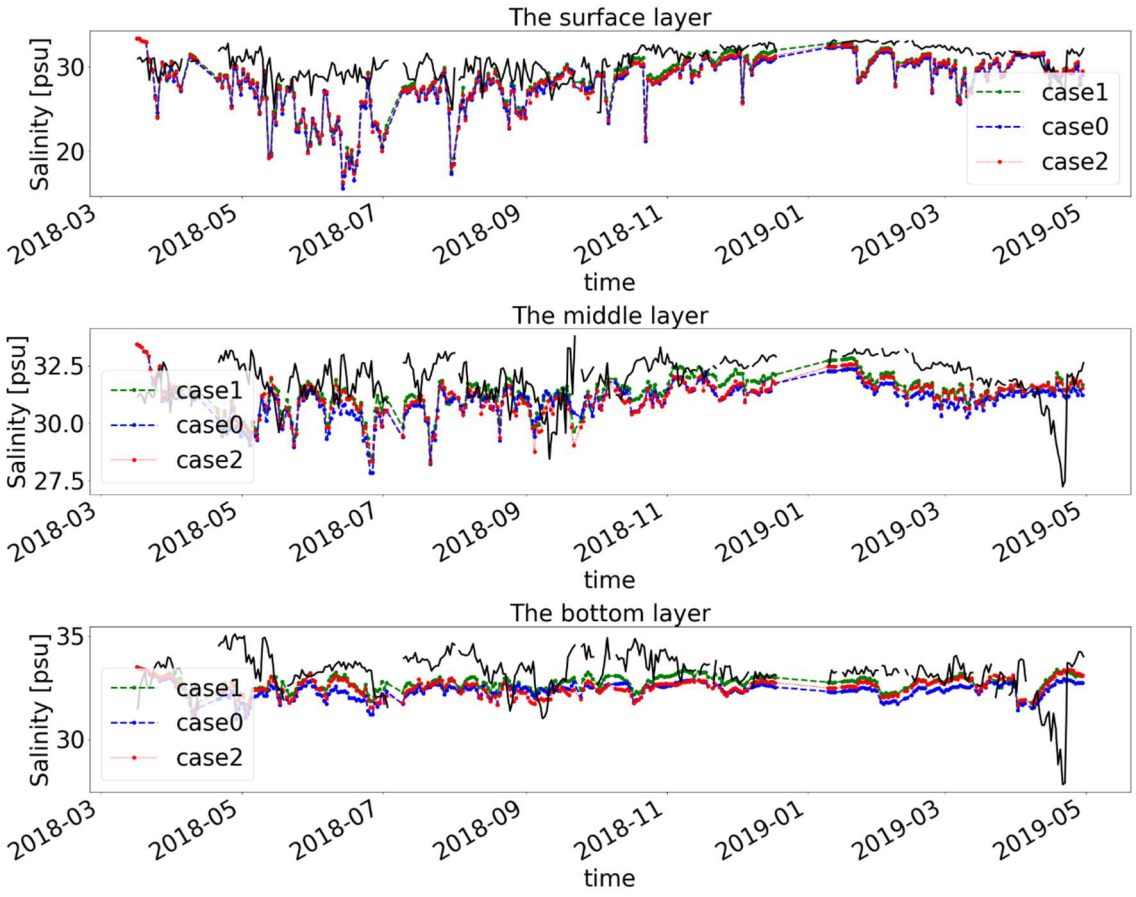

Figure 4.4. 4 The simulated salinity comparison at Kawasaki station from 2018-03-15 to 2019-04-30

The difference between the three cases at this observation station is even greater.

The simulated salt of case1 was higher than that of the other two cases throughout the year. The difference between the three cases is smallest at the top layer, and larger at the middle and bottom.

Chapter 5 Conclusion

By replacing the artificial open boundary conditions of water temperature and salinity originally set in TEEM with the high time resolution water temperature data measured by the observation station and the high time resolution salinity data predicted by HYCOM, the accuracy of the simulated water temperature and salinity inside Tokyo Bay was greatly improved. The functionality of TEEM was increased, and the modified model can reflect real changes in water temperature and salinity at the open boundary over time.

Compared with only using the water temperature and salinity datasets predicted by the ocean model to provide boundary conditions, the combination of the water temperature data measured by the observation station and the salinity data predicted by the ocean model can provide a more accurate open boundary conditions, and the accuracy of the simulated water temperature and salinity in the bay is higher.

Compared with the water temperature data of the observation station, the simulated water temperature obtained from using the OBC water temperature provided by HYCOM is low, regardless of how the data is inserted into TEEM at the bay mouth. This indicates that the water temperature provided by HYCOM at the mouth of Tokyo Bay may be lower.

When the open boundary conditions provided by HYCOM are embedded with the

model in the bay, attention should be paid to the way the two models with different water depth settings are coupled at the open boundary. When the HYCOM data is compressed in all depths and inserted into each layer of TEEM, compared with the water temperature and salinity datasets of HYCOM at 100 meters within the sea surface, the water temperature at the open boundary becomes colder, and the salinity increases.

Reference

Amunugama, M., & Sasaki, J. (2018). Numerical modeling of long-term biogeochemical processes and its application to sedimentary bed formation in Tokyo Bay. In *Water (Switzerland)* (Vol. 10, Issue 5). https://doi.org/10.3390/w10050572

Aoki, K., Shimizu, Y., Yamamoto, T., Yokouchi, K., Kishi, K., Akada, H., & Kurogi, H. (2022). Estimation of inward nutrient flux from offshore into semi-enclosed sea (Tokyo Bay, Japan) based on in-situ data. *Estuarine, Coastal and Shelf Science*, *274*(May), 107930. https://doi.org/10.1016/j.ecss.2022.107930

Chassignet, E. P., Hurlburt, H. E., Smedstad, O. M., Halliwell, G. R., Hogan, P. J., Wallcraft, A. J., Baraille, R., & Bleck, R. (2007). The HYCOM (HYbrid Coordinate Ocean Model) data assimilative system. *Journal of Marine Systems*, *65*(1-4 SPEC. ISS.), 60–83. https://doi.org/10.1016/j.jmarsys.2005.09.016

Chassignet, E. P., Smith, L. T., Halliwell, G. R., & Bleck, R. (2003). North Atlantic simulations with the Hybrid Coordinate Ocean Model (HYCOM): Impact of the vertical coordinate choice, reference pressure, and thermobaricity. *Journal of Physical Oceanography*, *33*(12), 2504–2526. https://doi.org/10.1175/1520-0485(2003)033<2504:NASWTH>2.0.CO;2

Halliwell, G. R. (2004). Evaluation of vertical coordinate and vertical mixing algorithms in the HYbrid-Coordinate Ocean Model (HYCOM). *Ocean*

Modelling, *7*(3–4), 285–322. https://doi.org/10.1016/j.ocemod.2003.10.002

Kim, H., Lozano, C.J., Tallapragada, V., Iredell, D., Sheinin, D., Tolman, H.L., Gerald, V.M., & Sims, J. (2014). Performance of Ocean Simulations in the Coupled HWRF-HYCOM Model. Journal of Atmospheric and Oceanic Technology, 31, 545-559.

Ji, Z. G. (2017). Hydrodynamics and water quality: Modeling rivers, lakes, and estuaries. *Hydrodynamics and Water Quality: Modeling Rivers, Lakes, and Estuaries*, *89*(39), 1–581. https://doi.org/10.1002/9781119371946

Liu, F., Sasaki, J., Chen, J. et al. (2022) Numerical assessment of coastal multihazard vulnerability in Tokyo Bay. *Nat Hazards*. https://doi.org/10.1007/s11069-022-05533-2

Masunaga, E., Uchiyama, Y., Suzue, Y., & Yamazaki, H. (2018). Dynamics of Internal Tides Over a Shallow Ridge Investigated With a High-Resolution Downscaling Regional Ocean Model. *Geophysical Research Letters*, *45*(8), 3550–3558. https://doi.org/10.1002/2017GL076916

Orlanski, I. (1976). A simple boundary condition for unbounded hyperbolic flows. *Journal of Computational Physics*, *21*(3), 251–269. https://doi.org/10.1016/0021-9991(76)90023-1

Sato, C., Nakayama, K., & Furukawa, K. (2012). Contributions of wind and river effects on DO concentration in Tokyo Bay. *Estuarine, Coastal and Shelf Science*, *109*, 91–97. https://doi.org/10.1016/j.ecss.2012.05.023

Sohma, A., Shibuki, H., Nakajima, F., Kubo, A., & Kuwae, T. (2018). Modeling a coastal ecosystem to estimate climate change mitigation and a model demonstration in Tokyo Bay. *Ecological Modelling*, *384*(July), 261–289. https://doi.org/10.1016/j.ecolmodel.2018.04.019

Suzuki, K. (2013). Effect of Wind and Fresh Water Discharge on Water Exchange and Hypoxia in Tokyo Bay. *Tech. Note Port and Airport Res. Inst.*

Tabeta, S., & Fujino, M. (1996). Comparison of simulation results and field data on currents and density in Tokyo Bay. *Journal of Marine Science and Technology*, *1*(2), 94–104. https://doi.org/10.1007/BF02391165