# CHAPTER 2 THEORIES

There are many methods for analyzing the sea level and sea surface height (SSH) in the Gulf of Thailand (GoT) and the South China Sea (SCS). In this chapter all methods used in this research are introduced. The trends of the observations are analyzed using two methods: least square linear regression and periodic cubic spline approximation. The Empirical Orthogonal Function (EOF) is used to reduce the spatial and temporal variability of the original data sets to new variables which cover most of the total variance. The wavelet transform and wavelet coherence are used to analyze the principal components of EOF. Furthermore, the descriptions of numerical ocean models: Ocean Circulation and Climate Advanced Modelling Project (OCCAM) and Princeton Ocean Model (POM) are given in this chapter.

## 2.1 Least Square Linear Regression

The best line fit between two paired variables is very useful in many applications. Linear least squares regression is a standard statistical analysis techniques. The basic idea is to find the line which minimizes the sum of the vertical distances squared between all data points and the least square line. Therefore, the line that fits best in this sense is called "least square fit" and the process of finding that line is called "least square linear regression". In this study, this method is used to find the trends of sea level for each tide gauge station.

The processes of least square linear regression are discussed. Firstly, considering $m$ pairs of data $(x_i, y_i), i = 1, \ldots, m$. Define $F(x) = \alpha x + \beta$ and the residual, $r_i$, for the data pair $(x_i, y_i)$ as

$$r_i = y_i - F(x_i) = y_i - (\alpha x_i + \beta),$$

where $\alpha$ and $\beta$ are the coefficients of $F(x)$.

The least squares fit is obtained by choosing the $\alpha$ and $\beta$ such that $\sum_{i=1}^{m} r_i^2$ is a minimum. To simplify the notation $\rho = \|r\|_2^2$ and find $\alpha$ and $\beta$ by minimizing $\rho = \rho(\alpha, \beta)$. The minimum requires

$$\frac{\partial \rho}{\partial \alpha}|_{\beta=\text{constant}} = 0 \quad \text{and} \quad \frac{\partial \rho}{\partial \beta}|_{\alpha=\text{constant}} = 0.$$

Fulfilling the differentiation leads to

$$\begin{aligned} S_{xx}\alpha + S_x\beta &= S_{xy}, & (2.1)\\ S_x\alpha + m\beta &= S_y, & (2.2) \end{aligned}$$

where $S_{xx} = \sum_{i=1}^{m} x_i x_i$, $S_x = \sum_{i=1}^{m} x_i$, $S_{xy} = \sum_{i=1}^{m} x_i y_i$ and $S_y = \sum_{i=1}^{m} y_i$.

Solving Eq. (2.1) and (2.2) for $\alpha$ and $\beta$ yields

$$\alpha = \frac{1}{d}(S_x S_y - m S_{xy}), \tag{2.3}$$
$$\beta = \frac{1}{d}(S_x S_{xy} - S_{xx} S_y), \tag{2.4}$$

where $d = S_x^2 - m S_{xx}$.

## 2.2 Approximation with Periodic Cubic Spline Function

In this section, the periodic cubic spline approximation is introduced for study the general trend of data sets and its application by Herrmann, 1996.

Given a data set $(x_k, y_k), k = 0, \ldots, n$ with $n \geq 3$ and $x_0 < x_1 < \ldots < x_n$. Let $y_n = y_0$ to get a periodic spline function and given control parameters $p_k, k = 0, \ldots, n$. To find a periodic cubic spline function $s(x)$ on $[x_0, x_n]$ with knots $x_0, \ldots, x_n$, that means we seek $4n$ coefficients $a_k, b_k, c_k$ and $d_k$ for $k = 0, \ldots, n-1$ such that

$$s(x) = s_k(x) = a_k + b_k(x - x_k) + c_k(x - x_k)^2 + d_k(x - x_k)^3, \tag{2.5}$$

in $[x_k, x_{k+1}]$ satisfying the following conditions:

**(i)** $s(x)$ and its first and second derivatives are continuous at the knots

$$\begin{aligned} s_{k-1}^{(j)}(x_k) &= s_k^{(j)}(x_k), \quad k = 1, \ldots, n-1, \quad j = 0, 1, 2 \\ s_{n-1}^{(j)}(x_n) &= s_1^{(j)}(x_1), \quad j = 0, 1, 2 \end{aligned}$$

**(ii)** $s(x)$ satisfies the minimal property

$$G(s) := \min \int_{x_0}^{x_n} [s''(x)]^2 dx, \tag{2.6}$$

**(iii)** $s(x)$ satisfies the following boundary conditions

$$H_k(\vec{a}) := [y_k - a_k]^2 \leq M_k, \tag{2.7}$$

where $M_k$ is slick variable for $k = 1, \ldots, n-1$.

The first and second derivatives of the spline function of Eq. (2.5) satisfy

$$\begin{aligned} s'(x) &= b_k + 2c_k(x - x_k) + 3d_k(x - x_k)^2 \quad \text{in } [x_k, x_{k+1}), \\ s''(x) &= 2c_k + 6d_k(x - x_k) \quad \text{in } [x_k, x_{k+1}). \end{aligned}$$

Therefore, the minimal property can be written as

$$G(s) := \min \int_{x_0}^{x_n} [2c_k + 6d_k(x - x_k)]^2 dx. \tag{2.8}$$

The solution of cubic spline interpolation can be written in a linear system as

$$\mathbf{S} \cdot \vec{c} = 3 \cdot \mathbf{Q} \cdot \vec{a}, \tag{2.9}$$

where $\vec{c} = (c_1, \ldots, c_{n-1})^T$, $\vec{a} = (a_1, \ldots, a_{n-1})^T$, $\mathbf{S}$ and $\mathbf{Q}$ are symmetric tridiagonal matrices whose entries satisfy

$$\begin{array}{lllllll} s_{k,k} & = & 2(h_{k-1}+h_k), & q_{k,k} & = & -(\frac{1}{h_{k-1}}+\frac{1}{h_k}), & k=0,\ldots,n-1 \\ s_{k,k+1} & = & s_{k+1,k} = h_k, & q_{k,k+1} & = & q_{k+1,k} = \frac{1}{h_k}, & k=0,\ldots,n-2 \\ s_{0,n-1} & = & s_{n-1,0} = h_{n-1}, & q_{0,n-1} & = & q_{n-1,0} = \frac{1}{h_{n-1}}, & \end{array}$$

with $h_k = x_{k+1} - x_k$, $k = 0, \ldots, n-1$ and $h_{-1} = h_{n-1}$.

Condition (iii) is the side condition at each knot. This allows us to control the approximation at each knot. For the minimal property condition (ii), to use the Lagrange multiplier method (see more details in Appendix) by introducing $n-1$ additional variables $z_0, \ldots, z_{n-1}$:

$$F(\vec{a}, z, \lambda) := \min \left[ G(\vec{a}) + \sum_{k=0}^{n-1} \lambda_k [H_k(\vec{a}) + z_k^2 - M_k] \right], \tag{2.10}$$

where $\lambda_0, \ldots, \lambda_{n-1}$ are the Lagrange parameters.

The necessary condition $\frac{\partial F}{\partial a_k}$ leads to

$$2 \cdot \mathbf{Q} \cdot \vec{c} + \mathbf{\Lambda} \cdot (\vec{a} - \vec{y}) = 0, \tag{2.11}$$

where $\mathbf{\Lambda} = diag(\lambda_0, \ldots, \lambda_{n-1})$.

Furthermore, the condition $\frac{\partial F}{\partial \lambda_k} = \frac{\partial F}{\partial z_k} = 0$ provides

$$\begin{aligned} H_k(\vec{a}) + z_k^2 - M_k &= 0, & (2.12) \\ 2 \cdot \lambda \cdot z_k &= 0, & (2.13) \end{aligned}$$

for $k = 0, \ldots, n-1$.

Pre-multiply Eq. (2.11) with $\mathbf{\Lambda}^{-1}$ and substitute $\vec{a}$ from Eq. (2.9) follow by pre-multiply it with $\mathbf{Q}$,

$$\begin{aligned} 2 \cdot \mathbf{\Lambda}^{-1} \cdot \mathbf{Q} \cdot \vec{c} + \vec{a} - \vec{y} &= 0, \\ 2 \cdot \mathbf{\Lambda}^{-1} \cdot \mathbf{Q} \cdot \vec{c} + \frac{1}{3} \cdot \mathbf{Q}^{-1} \cdot \mathbf{S} \cdot \vec{c} &= \vec{y}, \\ 6 \cdot \mathbf{Q} \cdot \mathbf{\Lambda}^{-1} \cdot \mathbf{Q} \cdot \vec{c} + \mathbf{S} \cdot \vec{c} &= 3 \cdot \mathbf{Q} \cdot \vec{y}. \end{aligned}$$

Thus,

$$(\mathbf{S} + 6 \cdot \mathbf{Q} \cdot \mathbf{\Lambda}^{-1} \cdot \mathbf{Q}) \cdot \vec{c} = 3 \cdot \mathbf{Q} \cdot \vec{y}. \tag{2.14}$$

**Definition 2.1** *Define*

$$\mathbf{A}(\mathbf{P}) := \mathbf{S} + 6 \cdot \mathbf{Q} \cdot \mathbf{\Lambda}^{-1} \cdot \mathbf{Q}, \tag{2.15}$$

*where* $\mathbf{P} := diag(p_0, \ldots, p_{n-1}) = \mathbf{\Lambda}$.

Eq. (2.14) can be written as the linear system

$$\mathbf{A}(\mathbf{P}) \cdot \vec{c} = 3 \cdot \mathbf{Q} \cdot \vec{y}. \tag{2.16}$$

Solving the simple linear system Eq. (2.16) by choosing $p_0, \ldots, p_{n-1}$ arbitrarily. The symmetric band matrix $\mathbf{A}$ is a band matrix with 5 diagonal entries along diagonal and 3 entries in the upper right and the lower left corner,

$$\mathbf{A} = \begin{bmatrix} * & * & * & 0 & \ldots & 0 & * & * \\ * & \ddots & \ddots & \ddots & & & & * \\ * & \ddots & \ddots & \ddots & \ddots & & & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & & & \ddots & \ddots & \ddots & \ddots & * \\ * & & & & \ddots & \ddots & \ddots & * \\ * & * & 0 & \ldots & 0 & * & * & * \end{bmatrix},$$

where the elements are

$$\begin{aligned}
a_{k,k} &= 2(h_{k-1} + h_k) + 6\Big[\frac{1}{p_{k-1}h_{k-1}^2} + \frac{1}{p_k}\Big(\frac{1}{h_{k-1}} + \frac{1}{h_k}\Big)^2 + \frac{1}{p_{k+1}h_k^2}\Big], \\
&\qquad\qquad k = 0, \ldots, n-1 \\
a_{k,k+1} &= a_{k+1,k} = h_k - 6\Big[\frac{1}{p_k h_k}\Big(\frac{1}{h_{k-1}} + \frac{1}{h_k}\Big) + \frac{1}{p_{k+1}h_k}\Big(\frac{1}{h_k} + \frac{1}{h_{k+1}}\Big)\Big], \\
&\qquad\qquad k = 0, \ldots, n-2 \\
a_{k,k+2} &= a_{k+2,k} = \frac{6}{p_{k+1}h_k h_{k+1}}, \quad k = 0, \ldots, n-3 \\
a_{0,n-2} &= a_{n-2,0} = \frac{6}{p_{n-1}h_{n-2}h_{n-1}}, \\
a_{0,n-1} &= a_{n-1,0} = h_{n-1} - 6\Big[\frac{1}{p_{n-1}h_{n-1}}\Big(\frac{1}{h_{n-2}} + \frac{1}{h_{n-1}}\Big) + \frac{1}{p_0 h_{n-1}}\Big(\frac{1}{h_{n-1}} + \frac{1}{h_0}\Big)\Big], \\
a_{1,n-1} &= a_{n-1,1} = \frac{6}{p_0 h_1 h_{n-1}},
\end{aligned}$$

with $h_{-1} = h_{n-1}$, $p_{-1} = p_{n-1}$ and $p_n = p_0$.

## 2.3 Empirical Orthogonal Function Analysis

The Empirical Orthogonal Function (EOF) analysis is often used in meteorology and oceanography fields for analysis of spatial or temporal variability. This method is also known as Principal Component Analysis (PCA) and its procedure is equivalent to a data reduction method widely used in the social sciences known as factor analysis. The Empirical Orthogonal Functions (EOFs) are simply method for partitioning the variance of data series which are called "empirical" and defined by the covariance of data sets (Emery and Thomson, 2004). This technique can be applied in many applications such as dimensionality reduction, image processing and face recognition.

The main idea of EOFs is to separate the spatial and temporal variability of original data series into a smaller number of new variables that cover most of the total original variance. The new variables are called orthogonal functions or principal components (PCs). EOF analysis is used to separated the dominant modes of variations in oceanography (Nerem, et al., 1997; Chu, et al., 2003; Rong, et al., 2007; Rojsiraphisal, 2007). There are two approaches for computing EOFs. The first approach is obtained by constructing a covariance matrix of the data series then finding eigenvalues and eigenvectors of the covariance matrix. The second one uses the Singular Value Decomposition (SVD) of the data matrix to obtain all the components of the EOFs (eigenvalues, eigenvectors, and time-dependent amplitudes) without computing the covariance matrix. The EOFs determined by both methods are identical. In this thesis, the first approach is chosen to compute the EOF for data sets obtained from observations and the results of numerical ocean models.

**EOF Computation Using the Covariance Matrix Method**

The data sets which are measured at location $x = 1, \ldots, p$ for each time $t = 1, \ldots, n$ are arranged in matrix $\mathbf{F}$. Constructing a matrix $F(t, x)$ with size $n$ by $p$ as:

$$\mathbf{F} = \begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1p} \\ F_{21} & F_{22} & \cdots & F_{2p} \\ \vdots & & & \vdots \\ F_{n1} & F_{n2} & \cdots & F_{np} \end{bmatrix}, \tag{2.17}$$

where any entries $F_{tx} = F(t, x)$ for $t = 1, \ldots, n$ and $x = 1, \ldots, p$.

Next removing the mean from each of the $p$ time series in $\mathbf{F}$, to get the anomaly matrix $F'(t, x)$, that is

$$F'(t, x) = F(t, x) - \bar{F}(x), \tag{2.18}$$

where $\bar{F}(x) = \dfrac{1}{n} \sum_{t=1}^{n} F(t, x)$ for $x = 1, \ldots, p$.

The covariance matrix (Björnsson and Venegas, 1997; Limsakul, 2004) is obtained from

$$\mathbf{R} = \mathbf{F}'^T \mathbf{F}', \tag{2.19}$$

where $T$ denotes the transpose of matrix and the principal components are obtained by solving the eigenvalue problem

$$\mathbf{RC} = \mathbf{C\Lambda}, \tag{2.20}$$

where $\mathbf{\Lambda}$ denotes the diagonal matrix containing the eigenvalues $\lambda_i$ of $\mathbf{R}$ and the $i^{th}$ column of $\mathbf{C}$ is the $i^{th}$ eigenvector corresponding to the eigenvalues $\lambda_i$. Both matrices $\mathbf{\Lambda}$ and $\mathbf{C}$ have size $p$ by $p$.

The $i^{th}$ pattern of EOF modes is described by the eigenvector corresponding to the $i^{th}$ eigenvalue. The EOF mode 1 is the eigenvector associated with the largest eigenvalue and EOF mode 2 is the eigenvector associated with the second largest eigenvalue, and so on. Since the matrix $\mathbf{C}$ satisfies $\mathbf{C}^T\mathbf{C} = \mathbf{C}\mathbf{C}^T = \mathbf{I}$ (where $\mathbf{I}$ is the identity matrix), this means that the EOFs are uncorrelated over space; i.e. each modes are orthogonal to each other. So, it is named that the "Empirical Orthogonal Function". Many authors refer to the patterns as the "EOFs", but some authors refer as the "principal component loading patterns" or just "spatial patterns". The time series are referred to as "EOF time series", "expansion coefficient time series", "expansion coefficients", "principal component time series" or just "principal components" (Björnsson and Venegas, 1997). In this research, the patterns and time series are referred as "EOFs" or "spatial patterns" and "principal components", respectively.

The principal components can be obtained from

$$\vec{a}_j = \mathbf{F}'\vec{c}_j, \tag{2.21}$$

where $\vec{a}_j$ are the projections of the map in $\mathbf{F}'$ on EOF mode $j$.

Reconstruct the data from EOFs and the principal components as

$$\mathbf{F}' = \sum_{j=1}^{p} \vec{a}_j(EOF_j). \tag{2.22}$$

The percentage of total variance the mode $k$ is in the form

$$\%\text{ of total variance mode } k = \frac{\lambda_k}{\sum_{i=1}^{p} \lambda_i} * 100. \tag{2.23}$$

In this research, the orthogonal constraints have been constructed in the EOF analysis so that

- The principal components (PCs) are orthogonal in time; i.e. there are no temporal correlation between any two principal components.
- The EOFs are orthogonal in space; i.e. there are no spatial correlation between any two EOFs.

In this study, the EOF analysis is used to analyze the grided data which includes land grid points from the observation and results from numerical ocean models. The EOF procedures have to compute by ignorance land using MATLAB as follow,

1. Assume the data is in a matrix, `anom`, with each row as one map and each column a time series for a given grided data.
2. Computing the covariance of matrix with land grid points or Not a Number (NaN) which this command already remove mean:

   ```
   Cov = nancov(anom,'pairwise');
   ```

3. Obtaining the eigenvalues and eigenvectors of the covariance matrix:

   ```
   [C, L] = eig(Cov);
   ```

   where `L` is a diagonal matrix of the eigenvalues corresponding to eigenvector matrix `C`.

4. Calculating the percentage of total variance:

```
k = diag(L)/trace(L) * 100;
```

5. Obtaining the principal components or amplitudes:

```
PCi = anom * C(:,i);
```

## 2.4 Wavelet Analysis

Wavelet analysis is a technique for time-frequency localization. The wavelet transform has been used in many fields such as mathematics, geophysics, image compression, electrical engineering, musical tones and de-noising noisy data. Comparison between sea level and other variations are analyzed using wavelet coherence (Hwang and Chen, 2000; Tiwari, et al., 2004; Rong, et al., 2007).

In this research, the wavelet transform based on wavelet analysis (Torrence and Compo, 1998) is used to investigate the principle components of sea level from difference sources which are obtained from EOFs analysis. Consider time series with equal time spacing $\delta t$, $x_n$, $n = 0, 1, \ldots, N-1$; $N$ denotes the number of times. Consider a mother wavelet or wavelet function, $\psi_0(\eta)$, which depends on non-dimensional for time parameter $\eta$. In this study, Derivation of a Gaussian (DOG) is used and given by

$$\psi_0(\eta) = \frac{(-1)^{m+1}}{\sqrt{\Gamma(m+\frac{1}{2})}} \frac{d^m}{d\eta^m}(e^{-\eta^2/2}), \tag{2.24}$$

where $m$ is the derivative. Note that DOG with $m = 2$ is called the Marr or Mexican hat function which is a real valued function and can be arrested both positive and negative oscillations of the time series as a separate peaks in wavelet power (Rojsiraphisal, 2007).

The continuous wavelet transform of a discrete time series $x_n$ is defined as the convolution of $x_n$,

$$W_n(s) = \sum_{k=0}^{N-1} \hat{x}_k \hat{\psi}^*(s\omega_k) e^{i\omega_k n \Delta t}, \tag{2.25}$$

where $s$ is a wavelet scale, $(\hat{\ })$ denotes Fourier transform, $(*)$ represents complex conjugate and $\omega_k$ represents an angular frequency which is defined as

$$\omega_k = \begin{cases} \dfrac{2\pi k}{N\Delta t} & : \ k \leq \frac{N}{2} \\[2ex] -\dfrac{2\pi k}{N\Delta t} & : \ k > \frac{N}{2}. \end{cases}$$

The Fourier transform of the DOG wavelet transform is in the form

$$\hat{\psi}_0(s\omega) = \frac{-i^m}{\sqrt{\Gamma(m+\frac{1}{2})}}(s\omega)^m e^{-(s\omega)^2/2}. \tag{2.26}$$

The wavelet power spectrum $W_n(s)$ is also magnitude of $W_n(s)$ because of the wavelet function $\psi(\eta)$ of the Mexican hat is a real valued function. In order to identify specific events from principal components achieved in EOFs section. The Mexican hat wavelet function and lag-1 autocorrelation for red noise background with 0.72 (Torrence and Compo, 1998) are used to calculate significant at the 5% level and global wavelet spectrum.

In this study, the wavelet analysis is applied to principal components of sea surface height anomaly (SSHA) from TOPEX/ERS and principal components of sea surface height (SSH) from OCCAM and POM. The wavelet software is provided by Torrence and Compo (2005).

## 2.5 Wavelet Coherence

The wavelet coherence has been used to identify frequency of two time series, defined as the square of the cross spectrum normalized by the individual power spectra and gives quantity between 0 and 1 (Torrence and Compo, 1998). The wavelet transforms of two time series $X$ and $Y$ ($x_n, n = 0, \ldots, N-1$ and $y_n, n = 0, \ldots, N-1$) is defined as the convolutions of $x_n$ and $y_n$ with a scaled and translated version of $\psi_0(\eta)$:



$$W_n^X(s) = \sum_{n'=0}^{N-1} x_{n'}\psi^*\left[\frac{(n'-n)\delta t}{s}\right], \tag{2.27}$$

$$W_n^Y(s) = \sum_{n'=0}^{N-1} y_{n'}\psi^*\left[\frac{(n'-n)\delta t}{s}\right], \tag{2.28}$$

where the (*) indicates the conjugate and $\psi_0(\eta) = \pi^{-1/4}e^{i\omega_0\eta}e^{-\eta^2/2}$.

The cross wavelet spectrum is defined as

$$W_n^{XY}(s) = W_n^X(s)W_n^{Y*}(s), \tag{2.29}$$

where $(*)$ denotes complex conjugate.

The wavelet coherence of two time series is defined by the following (Torrence and Webster, 1999)

$$R_n^2(s) = \frac{|< s^{-1}W_n^{XY}(s) >|^2}{< s^{-1}|W_n^X(s)|^2 >< s^{-1}|W_n^Y(s)|^2 >}, \tag{2.30}$$

where $< \cdot >$ indicates smoothing in both time and scale. The factor $s^{-1}$ is used to convert to an energy density.

The wavelet coherence phase difference is defined as

$$\phi_n(s) = \tan^{-1}\left(\frac{\mathrm{Im}\{< s^{-1}W_n^{XY}(s) >\}}{\mathrm{Re}\{< s^{-1}W_n^{XY}(s) >\}}\right) \tag{2.31}$$

where Im{} and Re{} represent imaginary and real part.

In this research, the correlations of SSHA from TOPEX/ERS altimeter and SSHs from OCCAM (with low and high resolutions) and POM are analyzed by wavelet coherence. The wavelet coherence package are obtained from Grinsted, et al. (2005). The principal components obtained from EOFs analysis are used to find the relationship between SSHA and SSHs. The first part, principal components of SSHA from TOPEX/ERS and SSHs from OCCAM with resolutions of 1/4° and 1/12° in the GoT are investigated. Later, the principal components of SSHA from TOPEX/ERS and SSH from POM are analyzed in the GoT. Moreover, the principal components of SSHA from TOPEX/ERS and SSH from OCCAM with a resolutions of 1/4° are investigated in the SCS.

## 2.6 The Ocean Circulation and Climate Advanced Modelling Project (OCCAM)

The OCCAM is a primitive equation numerical global ocean model which is based on the Bryan-Cox-Semtner ocean general circulation model, but includes a free surface and improved advection schemes. This model uses an Arakawa B-grid (see Figure 2.1) in the horizontal and z-coordinate in the vertical. The OCCAM model is split into two parts: Model 1 and Model 2. Model 1 uses a standard latitude-longitude grid and covers the Pacific, Indian and South Atlantic Oceans. Model 2 uses a rotated grid and covers the North Atlantic and Arctic Oceans (Webb, et al., 1998). The model depth is calculated from the US Navy DBDB5 data set which uses grid of 1/4° and 1/12°. The depth is set to zero at the land grid points. For the resolution of 1/4°, the model has 36 levels with thickness from 20 meters near the surface up to 255 meters at the depth of 5500 meters. The model has 66 levels in the vertical while 14 levels in the upper 100 meters for the resolution of 1/12° (Coward and De Cuevas, 2005). The model is started from the Levitus annual mean temperature and salinity fields. The surface forcing uses European Centre for Medium-Range Weather Forecasts (ECMWF) monthly mean winds and is relaxed to the Levitus seasonal surface temperature and salinity fields.

The OCCAM based on the Bryan-Cox-Semtner ocean general circulation model can be defined by the potential temperature, salinity, three components of velocity and others variables. These variables can be specified using a momentum equation for the time changes in velocity and an advection-diffusion equation for the changes of temperature and salinity. The system also needs a continuity equation, an equation of state and boundary condition to be specified. The horizontal momentum equation and the three dimensional advection/diffusion equations are

$$
\begin{aligned}
\frac{\partial \vec{U}}{\partial t} + (\vec{U}\cdot\nabla)\vec{U} + w\frac{\partial \vec{U}}{\partial z} + f\times\vec{U} &= -(\frac{1}{\rho_0})\nabla p + D_U + D_U, && (2.32)\\
\frac{\partial S}{\partial t} + (\vec{U}\cdot\nabla)S + w\frac{\partial S}{\partial z} &= D_S + F_{S'}, && (2.33)\\
\frac{\partial T}{\partial t} + (\vec{U}\cdot\nabla)T + w\frac{\partial T}{\partial z} &= D_T + F_{T'}, && (2.34)
\end{aligned}
$$

and the pressure (or vertical momentum), incompressibility and density equations are

$$\rho g = -\frac{\partial p}{\partial z}, \quad (2.35)$$

$$\nabla \cdot \vec{U} + \frac{\partial w}{\partial z} = 0, \quad (2.36)$$

$$\rho = \rho(T, S, p), \quad (2.37)$$

where $\vec{U}$ is the horizontal velocity $(u, v)$, $T$ is the potential temperature, $S$ is the salinity, $p$ is the pressure, $w$ is the vertical velocity, $\rho$ is the density, $t$ denotes time, $f$ is the Coriolis parameter (equals to $2\Omega \sin(\theta)$, $\Omega$ is the Earth's rotation rate and $\theta$ represents the latitude), $D$ is the diffusion, $F$ is the forcing and $g$ is the gravitational acceleration.

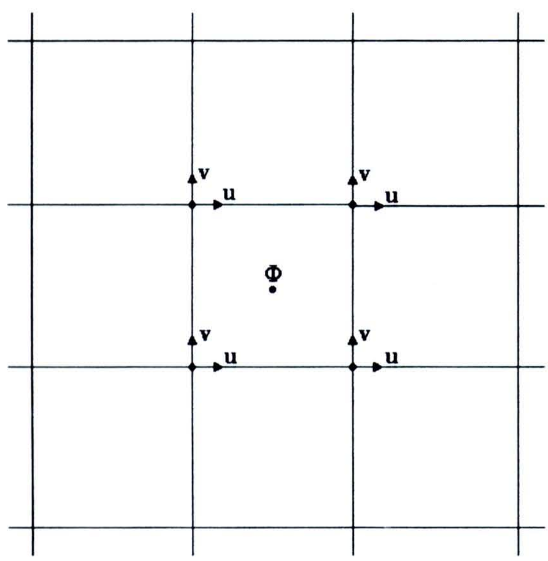


**Figure 2.1** The Arakawa staggered B-grid. The horizontal velocity components $(u, v)$ are defined at the corner of grid cell and the variables ($\Phi$ denotes $T, S$ and $p$) are defined at the center of grid cell.

In numerical ocean model, three important approximations are often made to reduce the computational load. The first assumption, the ocean is incompressible in the continuity equation. The second assumption, the vertical velocity is small and the terms involving can be neglected in the vertical momentum equation. The third assumption, the small changes in density can be neglected except where they affect the horizontal pressure gradient in the horizontal momentum equation. The resulting equations are called the primitive equations.

In this study, the monthly mean SSHs obtained from OCCAM which cover the GoT (98°E to 105°E and 6°N to 14°N) and the SCS (105°E to 123°E and 0°N to 25°N) are investigated. The SSHs from OCCAM model with resolution of 1/4° and 1/12° are taken during year 1985 - 2004 and 1988 - 2004, respectively.

## 2.7 The Princeton Ocean Model (POM)

Another numerical ocean model used in the study is a Princeton Ocean Model (POM) (Blumberg and Mellor, 1987). In this section, a brief description of the model applied in the GoT is given. The governing equations of the POM include equations of mass, momentum, temperature and salinity conservations. The sigma coordinate or pressure coordinate is used in this model (Aschariyaphotha, 2006). The density in each contour of vertical layers is quite constant in sigma coordinate system. This is attained by transformation of the governing equations from $z$ coordinate $(x, y, z, t)$ to the vertical sigma coordinate $(x^*, y^*, \sigma, t^*)$.

In this research, a two dimensional oceanic model is used for calculating the sea surface height and is derived by vertically integrated equations. This technique is known as mode splitting (Simons, 1974; Madala and Piacsek, 1977) which separates the vertically integrated equations (external mode) from the vertical structure equations (internal mode).

The barotropic (external mode) equations are obtained by integrating the vertical structure of equations over the depth. After integration, the continuity equation is

$$\frac{\partial \bar{u} D}{\partial x} + \frac{\partial \bar{v} D}{\partial y} + \frac{\partial \eta}{\partial t} = 0.$$

The momentum equations are

$$\begin{aligned} \frac{\partial \bar{u} D}{\partial t} + \frac{\partial \bar{u}^2 D}{\partial x} + \frac{\partial \bar{u}\bar{v} D}{\partial y} & -\tilde{F}^x - f\bar{v}D + gD\frac{\partial \eta}{\partial x} = - < wu(0) > + < wu(-1) > \\ & + G_x - \frac{gD}{\rho_0}\int_{-1}^{0}\int_{\sigma}^{0}(D\frac{\partial \rho'}{\partial x} - \sigma'\frac{\partial D}{\partial x}\frac{\partial \rho'}{\partial \sigma})d\sigma' d\sigma, \quad (2.38) \\ \frac{\partial \bar{v} D}{\partial t} + \frac{\partial \bar{u}\bar{v} D}{\partial x} + \frac{\partial \bar{v}^2 D}{\partial y} & -\tilde{F}^y - f\bar{u}D + gD\frac{\partial \eta}{\partial y} = - < wv(0) > + < wv(-1) > \\ & + G_y - \frac{gD}{\rho_0}\int_{-1}^{0}\int_{\sigma}^{0}(D\frac{\partial \rho'}{\partial y} - \sigma'\frac{\partial D}{\partial y}\frac{\partial \rho'}{\partial \sigma})d\sigma' d\sigma, \quad (2.39) \end{aligned}$$

where $u$, $v$ are the horizontal velocity components, $w$ denotes the vertical component of velocity, $g$ is the gravitational acceleration, $f = 2\Omega \sin\theta$ is the Coriolis parameter ($\Omega$ is the speed of angular rotation of the earth given by $\Omega = 7.2921 \times 10^{-5}\ rad\, s^{-1}$), $\rho_0$ is a reference density of sea water, $D = H + \eta$ ($H$ is the bottom topography and $\eta$ is the sea surface elevation), the overbars denote vertically integrated velocities such as

$$\bar{u} = \int_{-1}^{0} u d\sigma,$$

and the terms $(< wu(0), wv(0) >)$ and $(< wu(-1), wv(-1) >)$ are wind stress and bottom stress (the components are opposite in sign), respectively. The quantities of the horizontal viscosity $\tilde{F}^x$ and $\tilde{F}^y$ are defined as

$$\begin{aligned} \tilde{F}^x &= \frac{\partial}{\partial x}(2H\bar{A}_m\frac{\partial \bar{u}}{\partial x}) + \frac{\partial}{\partial y}[H\bar{A}_m(\frac{\partial \bar{u}}{\partial y} + \frac{\partial \bar{v}}{\partial x})], \quad (2.40) \\ \tilde{F}^y &= \frac{\partial}{\partial y}(2H\bar{A}_m\frac{\partial \bar{v}}{\partial x}) + \frac{\partial}{\partial x}[H\bar{A}_m(\frac{\partial \bar{u}}{\partial y} + \frac{\partial \bar{v}}{\partial x})], \quad (2.41) \end{aligned}$$

where $A_m$ is the horizontal turbulent diffusion coefficient (Mellor, 2004) and the dispersion terms are defined as

$$G_x = \frac{\partial \bar{u}^2 D}{\partial x} + \frac{\partial \bar{u}\bar{v} D}{\partial y} - \tilde{F}^x - \frac{\partial \bar{u^2} D}{\partial x} - \frac{\partial \overline{uv} D}{\partial y} + \bar{F}^x, \tag{2.42}$$

$$G_y = \frac{\partial \bar{u}\bar{v} D}{\partial x} + \frac{\partial \bar{v}^2 D}{\partial y} - \tilde{F}^y - \frac{\partial \overline{uv} D}{\partial x} - \frac{\partial \bar{v^2} D}{\partial y} + \bar{F}^y. \tag{2.43}$$

### 2.7.1 Initial conditions

Assume that the water velocity is zero when the model starts computation. The initial conditions for running model consist of bottom topography, temperature, salinity and wind stress from observations at each grid point. That is

$$\begin{aligned} \eta &= 0, \\ u &= v = 0, \\ T' &= S' = 0, \\ \tilde{\rho} &\equiv \rho(\tilde{T}, \tilde{S}, z), \end{aligned}$$

where $\tilde{T}$ and $\tilde{S}$ are taken from observations.

### 2.7.2 Boundary conditions

**Wind Stress**

The wind stress is described as

$$\rho_0 A_{mv}(\frac{\partial u}{\partial z}, \frac{\partial v}{\partial z})_{\sigma \to 0} = (\tau_{0x}, \tau_{0y}), \tag{2.44}$$

where $(\tau_{0x}, \tau_{0y})$ is wind stress in eastward and northward directions. The wind stress is calculated from monthly mean wind of the ECMWF. The wind stress scheme is specified by Mellor (2004),

$$\begin{aligned} (\tau_{0x}, \tau_{0y}) &= \rho_a C_D |\vec{V_a} - \vec{V_l}|(U_a - U_l, V_a - V_l), \\ &\approx \rho_a C_D |\vec{V_a}|(U_a, V_a), \end{aligned} \tag{2.45}$$

where $\rho_a$ is the sea level air density, $(U_a, V_a)$ is the wind velocity $(\vec{V_a})$ measured from 10 m above the sea surface, $(U_l, V_l)$ is the surface layer ocean current velocity $(\vec{V_l})$ and $C_D$ represents the drag coefficient from Matthias and Godfrey (1994) which is calculated by

$$C_D = \min\left\{0.001 + 0.00007|\vec{V_a}|, 0.0025\right\}.$$

**Bottom Boundary Forcing**

On the bottom of the gulf, the bottom friction condition is calculated from Mellor (2004) as

$$\rho_0 A_{mv}(\frac{\partial u}{\partial z}, \frac{\partial v}{\partial z}) = (\tau_{bx}, \tau_{by}), \tag{2.46}$$

where $(\tau_{bx}, \tau_{by})$ is the bottom stress parameterized as

$$(\tau_{bx}, \tau_{by}) = \rho_a C_z |\vec{V_b}| (U_{bx}, V_{by}), \tag{2.47}$$

where $(U_{bx}, V_{by})$ is the component of the nearest bottom velocity $(\vec{V_b})$ and $C_z$ is the coefficient of bottom friction which obtained from Mellor (2004) by

$$C_z = \max\left\{ \frac{\kappa^2}{\left[\ln\left\{\frac{(1+\sigma_{kb-1})H}{z_0}\right\}\right]^2}, 0.0025 \right\},$$

where $\kappa$ is the Von Kármán constant which is taken to be 0.4, $H$ is the bottom topography and $z_0$ is the bottom roughness which equals to 0.01 m.

### 2.7.3 The Numerical Scheme

#### External Mode Calculation

The external mode calculation results in updating the surface elevation and the vertically averaged velocities (Mellor, 2004; Aschariyaphotha, 2006). The time stepping process for external mode is shown in Figure 2.2. The external mode does a leap frog many times with a short time step, $\Delta t_E$, until $t = t^{n+1}$. The vertical and time averaged velocities, (UTF,VTF) and those from the previous time step, (UTB,VTB) are time averages of the external variables, (UA,VA).

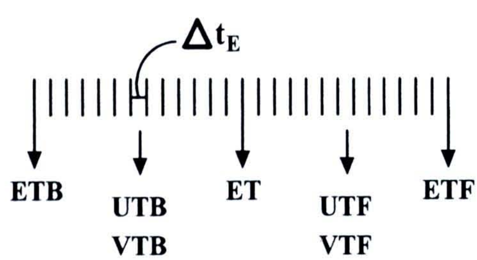


**Figure 2.2** The external mode calculation.

#### Grid Arrangement

The staggered grid arrangement for the external mode is an Arakawa C-grid (Mellor, 2004), as shown in Figure 2.3.

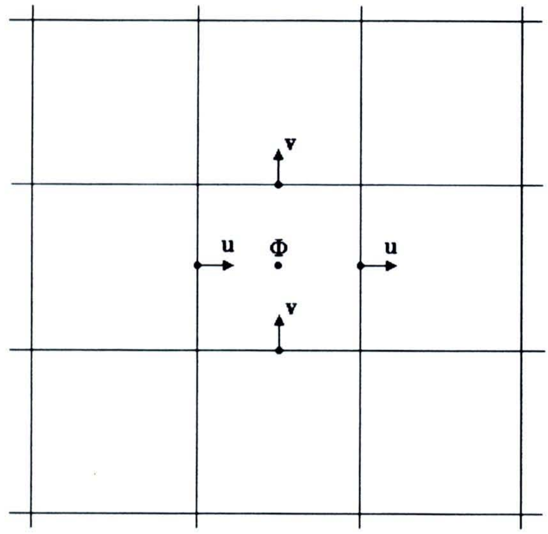


**Figure 2.3** The two dimensional external mode grid is an Arakawa C-grid. The horizontal components of velocity $(u, v)$ are defined along the middle borderline of grid cell and the variables ($\Phi$ denotes $T, S$ and $p$) are defined at the center of grid cell.

### Time Step Constraints

The Courant-Friedrichs-Lewy (CFL) computational stability conditions on the external mode is described as

$$\Delta t_E \leq \frac{1}{C_t}\left|\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2}\right|^{-1/2},$$

where $C_t = 2(gH)^{1/2} + U_{max}$; $U_{max}$ is the expected maximum velocity, $\Delta x$ and $\Delta y$ are the grid spacing in $x$ and $y$ directions.

### 2.7.4 Parameterizations of POM

Parameters used for driving the POM are shown in Table 2.1.

**Table 2.1** Parameters used in the POM

| Parameter | Symbol | Value | Unit |
|---|---|---|---|
| - External mode time step | $\Delta t_E$ | 30 | $s$ |
| - HORCON parameter | $C$ | 0.2 | – |
| - Turbulence Prandtl number | $P_{rt}$ | 0.9 | – |
| - Density of air | $\rho_a$ | 1.03 | $kg\ m^{-3}$ |
| - Density of sea water | $\rho_0$ | 1025 | $kg\ m^{-3}$ |
| - Gravity constant | $g$ | 9.806 | $m\ s^{-2}$ |
| - Angular speed of the Earth's rotation | $\Omega$ | $7.2921 \times 10^{-5}$ | $rad\ s^{-1}$ |
| - Coefficient of vertical eddy viscosity | $A_{mv}$ | $5 \times 10^{-5}$ | $m^2 s^{-1}$ |
| - Coefficient of vertical eddy diffusivity | $A_{hv}$ | $5 \times 10^{-5}$ | $m^2 s^{-1}$ |
| - Drag coefficient for the wind stress | $C_D$ | $[0.001, 0.0025]$ | – |
| - Drag coefficient for the bottom friction | $C_z$ | $[0.0025, 1]$ | – |
| - Von Kármán | $\kappa$ | 0.4 | – |
| - Bottom roughness parameter | $z_0$ | 0.01 | $m$ |