# CHAPTER 2
# THEORIES

This chapter aims to present related theories for this study. They are divided into two subjects, which are climate features and statistical methods.

## 2.1 East Asian Winter Monsoon and Climate Modes

In general, the monsoon is an important climate feature, and there are the summer and the winter monsoons. Emphasizing the Asian monsoon, the EAWM is dominant climate feature during winter season, and there are less studies of the EAWM comparing to the Asian summer monsoon (He, 2007; Takahashi, 2006, 2010, 2011). The other important climate modes are the forcing from SST. The forcing from global SST anomalies play important role on the interannual variability of Asian winter monsoon (Ji, 1997). Thus, the modes of SST variabilities in the tropical Pacific Ocean and the Indian Ocean are ENSO and IOD are important for climate variability over the IDP region.

### 2.1.1 East Asian Winter Monsoon

The monsoon is generally to be identified by the changing of the seasonal winds. The EAWM is the one importance of monsoon system over the Asian continent during boreal winter season. Major surface characteristic is strong northeasterly wind over the coast of East Asia, which resulted from the pressure gradient due to thermal contrast between continent and ocean. The pathway of wind blowing splits into two branches. The first branch turns east ward moving to subtropical Pacific, whereas the second branch turns along coastline of East Asia moving to tropical region (Figure 2.1). To describe the intensity of EAWM, there is complexity by using differences of pressure or temperature because there are different trends over different areas. Alternatively, interannual variability of wind can be used to represent the variability of EAWM. Due to the characteristic of EAWM showing strong northeasterly wind, averaging of surface meridional wind over representative areas was selected to construct an index to describe the variability of EAWM (Chen, 2000). Not only the surface meridional wind, but the meridional wind at 1000 hPa can also be used to determine an index to describe the variability of EAWM (Ji, 1997).

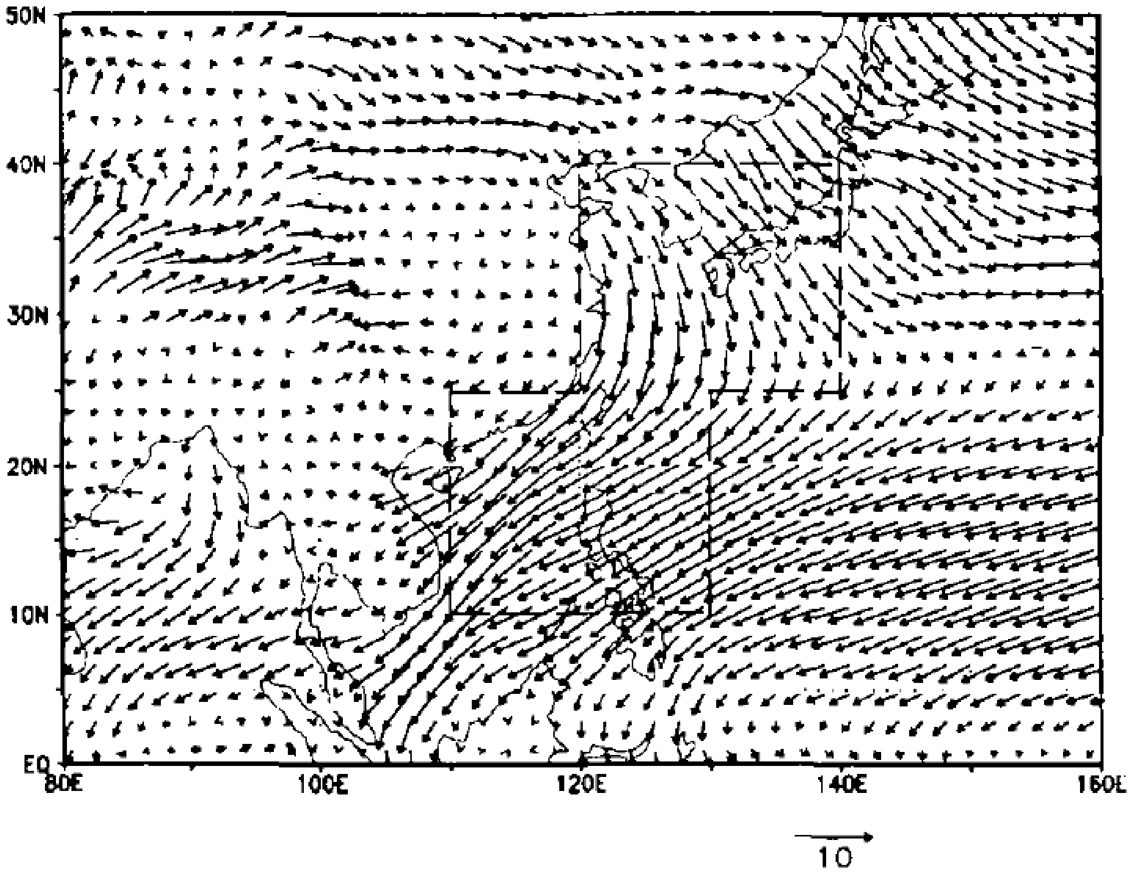



Figure 2.1: The average of surface wind (m/s) representing EAWM, and bounded area using to define the EAWM index (Chen, 2000).

One characteristic of the EAWM in terms of interannual annual variability of anomalous wind fields at 850 hPa was differentiated into two modes by utilizing the EOF technique to analyze the wind data of the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR), NCEP/NCAR reanalysis data. The leading mode consists of two distinct modes. The first distinct mode related to the intensity of the SH, and the corresponding time series show that the positive phase of the leading mode results to deepening of the East Asian trough at 500 hPa level. On the other hand, it is noteworthy that the PC time series of the second distinct mode does not show significant relationship with the SH and the East Asian trough, but it significantly correlated to ENSO indices than the first. It indicates that the second distinction is closely related to the atmospheric circulation anomalies in the Tropics and ENSO (as measured by equatorial eastern Pacific sea surface temperature). Therefore, the interannual variability of

EAWM is not simply reasoned from the pressure gradient between the SH and AL, and a simple index is not enough for the EAWM study (Wu, 2006).

As mentioned before, the variability of the EAWM has some relation to the ENSO (Chen, 2000;Wu, 2006). The study by modelling indicates that the forcing from global SST anomalies influences on the interannual variability of Asian winter monsoon, and tropical SST anomalies are important to play role on the Walker circulation (Ji, 1997). Focusing on the tropical Pacific Ocean, the EAWM is influenced by tropical SST anomalies as shown in Figure 2.2. This shows the EAWM becomes weak during El Niño event (SST anomaly in tropical eastern Pacific Ocean becomes positive), and the EAWM becomes strong during La Niña (SST anomaly in tropical eastern Pacific Ocean becomes negative) (Chen, 2000).

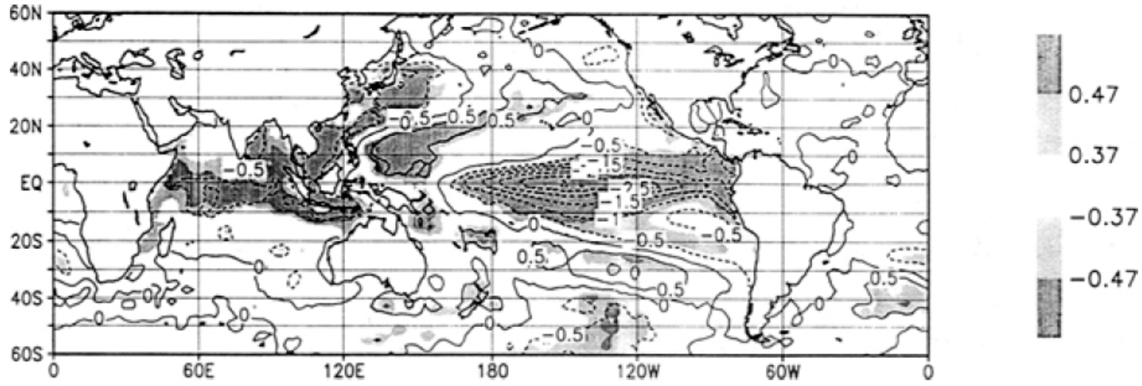



Figure 2.2: Composite differences of SST between strong and weak monsoons during the boreal winter (Chen, 2000).

The association between EAWM and ENSO related to changes in vertical circulation cells in terms of interannual variation. The proposed mechanism by Zeng (2011) describes the changes of the zonal Walker cell in Tropics, the midlatitude zonal cell, the western Pacific Hadley cell, and the eastern Pacific Hadley cell. The zonal Walker, the midlatitude zonal, and the western Pacific Hadley circulation cells are enhanced, but the eastern Pacific Hadley cell is weakened when the EAWM is strong. During weak EAWM, it shows the opposite. These resulted by SST anomalies in the tropical Pacific Ocean. Namely, the cold SSTA in the central tropical Pacific results in stronger zonal Walker, midlatitude zonal, and western Pacific Hadley circulation cells with a weakened eastern Pacific Hadley cell that leading to present a stronger EAWM, and vice versa for the weak monsoon.

There are scientific efforts on defining the intensity and revealing the variability of the EAWM using various kinds of indices. These indices were categorized into four categories as follows (Wang, 2010):

- Low-level wind indices,
- Upper zonal wind over the East Asia indices,
- Pressure gradient indices, and
- East Asian trough indices.

These indices were derived from a variety of atmospheric variables as shown in Table 2.1.

Table 2.1: General description of EAWM index categories (Wang, 2010).

| Category | Parameter for Index Calculation |
|---|---|
| Low-level wind indices | Meridional wind at 10 m<br>Meridional wind at1000 hPa<br>Zonal and meridional wind at 850 hPa |
| Upper zonal wind over the East Asia indices | Zonal wind at 300 hPa |
| Pressure gradient indices | Difference of sea level pressure between two areas |
| East Asian trough indices | Geopotential height at 500 hPa |

The indices of the first category were developed under a basis of the important characteristic of EAWM, which is a low-level wind blowing along the coast of East Asia (Wang, 2010). The low-level meridional wind index has significant signals over the tropical and eastern Pacific, and the indices more significantly correlated to the ENSO index than other categories that indicating a close relationship with ENSO (Wang, 2010).

The second and third categories indices were developed based on the concept of association between the features of the upper tropospheric East Asian jet stream, and the variations of the EAWM and thermal contrast between continents and oceans represented by differences in SLP, respectively. The second category index perhaps more suitable for describing variations of the EAWM in terms of the interdecadal variation than the interannual variation (Wang, 2010).

The fourth category indices used the East Asian trough characteristic to describe the EAWM. The indices under this category show higher significant correlation coefficients with AO than the Niño 3 index (Wang, 2010).

**2.1.2 Walker and Hadley Circulations**

The Walker and Hadley circulations are important vertical circulations in low latitudes. They are the strong driving force of large-scale circulation in the Tropics. Their variations affect climate variability, and produce strong convection in the western equatorial Pacific Ocean. The Walker circulation is the vertical circulation along the East-West axis near the equator, and driven by the contrast of SST in the Ocean as shown in Figure 2.3 for wintertime. Its intensity can be quantified by the Southern Oscillation index (SOI) (Tanaka, 2004), which related to ENSO phenomena.

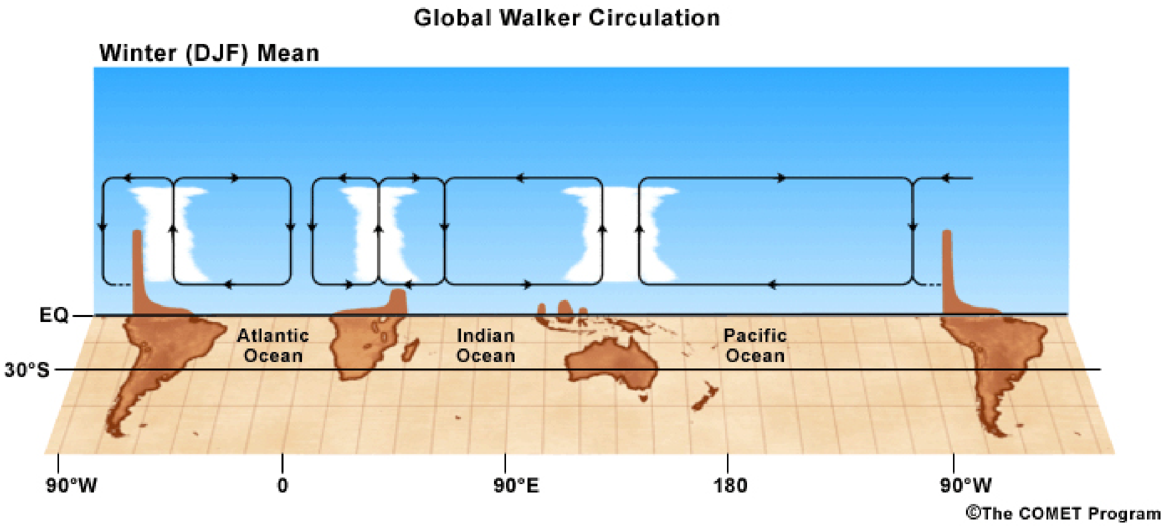



Figure 2.3: Global Walker circulation (source: http://28storms.com/longrange/wpcontent /uploads/2011/04/Global_Walker_Circulation.gif).

The Hadley circulation cell is a meridional atmospheric circulation cell in the tropical region. Its important feature is the rising of air mass motion over the Intertropical Convergence Zone (ITCZ), then moving polewards, and descending at the subtropical belts (Figure 2.4). The Hadley cell is driven by the meridional differences in heating by the global radiative process (Tanaka, 2004).

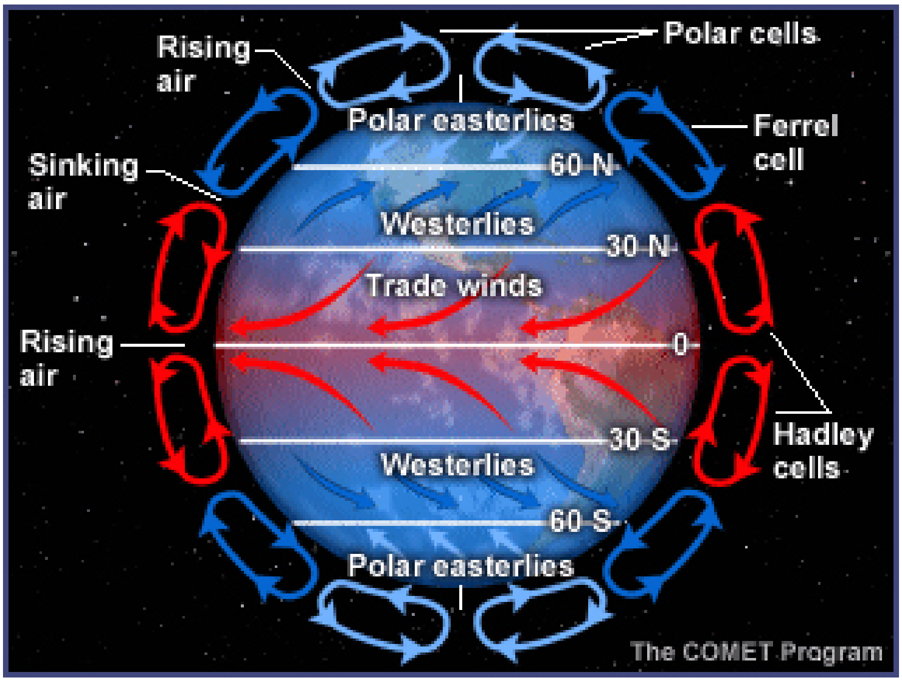



Figure 2.4: Polar, Ferrel, and Hadley circulation cells (source: http://minerva.union.edu).

### 2.1.3 El Niño-Southern Oscillation

Originally, the term El Niño (in Spanish, Niño means "the boy", the Christ-child) was related to the weak warm ocean current that flows along the coast of Peru and Ecuador around Christmastime, and associated with the unusually large warm pool in the Pacific Ocean that has the linkages with anomalous global climate patterns. Since there is collaboration between atmospheric and oceanic phenomena, the term El Niño has been tied with the southern oscillation phenomenon to be the El Niño-Southern Oscillation (ENSO). The close variation between El Niño and the Southern Oscillation is shown in Figure 2.5. The El Niño related to the warm phase of ENSO, whereas the cool phase of ENSO is "the girl" in Spanish that called La Niña phenomenon showing cooling of SST in the tropical Pacific Ocean (Trenberth, 1997). For mechanism during El Niño, the trade winds is weaken along the equator, whereas atmospheric pressure increases and decreases in the western Pacific and the eastern Pacific, respectively that related to anomalous warming of SST in the central and eastern Pacific Ocean with warm water in the western Pacific Ocean moves eastward resulting upwelling is decreased. Whereas, La Niña shows anomalous cooling of SST in the central and eastern Pacific Ocean with opposite characteristic to El Niño as shown in Figure 2.6 (McPhaden, 2006; Ashok, 2009). It shows linkages with various global climate patterns (Trenberth, 1997) such as the relation with EAWM (Chen, 2000; Wu, 2006; Zeng, 2011).

The temporal variation signals and spatial patterns of ENSO can be captured by the leading mode from the analysis of tropical SST anomalies in the Pacific Ocean (Diaz, 2001; Li, 2010). Singular value decomposition analysis can also detect ENSO (Trenberth, 2001). Therefore, an index using to characterize El Niño activity would strongly correlated to the time series representing the phenomenon given by the analysis as much as possible (Li, 2010). There are many indices using to indicate the ENSO events. The well known indices derived from SST behaviour suggested by Trenberth (2001) are Niño 1+2, Niño 3, Niño 4, and Niño 3.4, that represent the SST anomalies averages over corresponding areas, as shown in Figure 2.7. Among Niño indices, the Niño 3.4 describe well on the variability of the leading mode given by the singular value decomposition analysis (Trenberth, 2001), whereas Niño 3 and Niño 3.4 show strong correlation (greater than 0.9) with the leading mode given by the EOF analysis (Li, 2010).

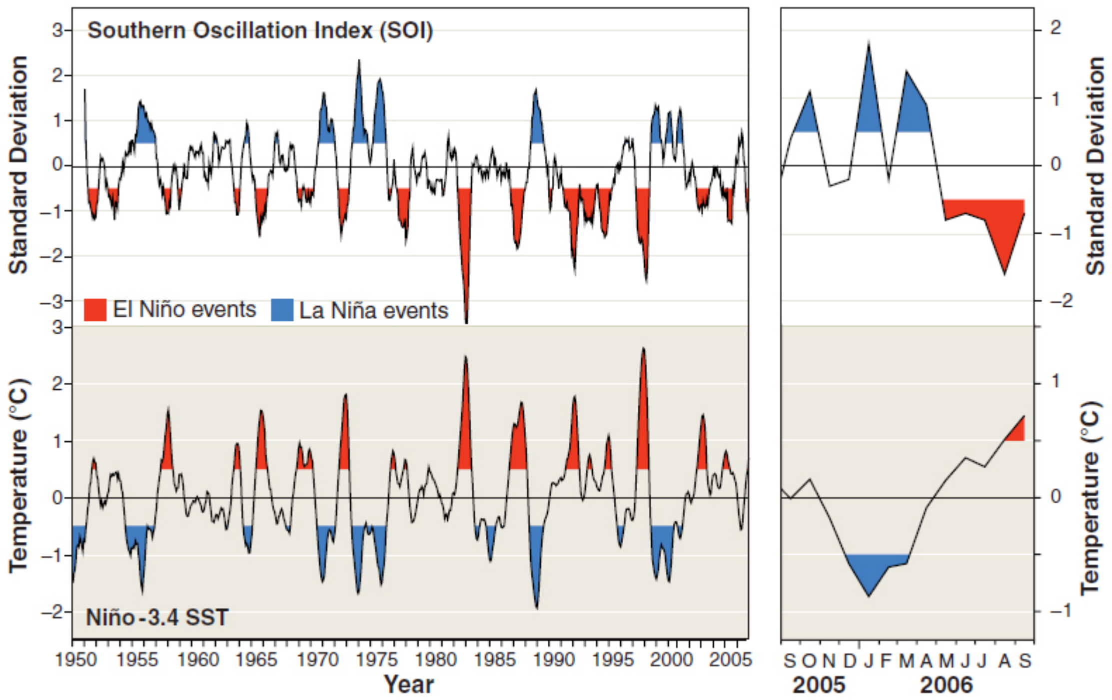
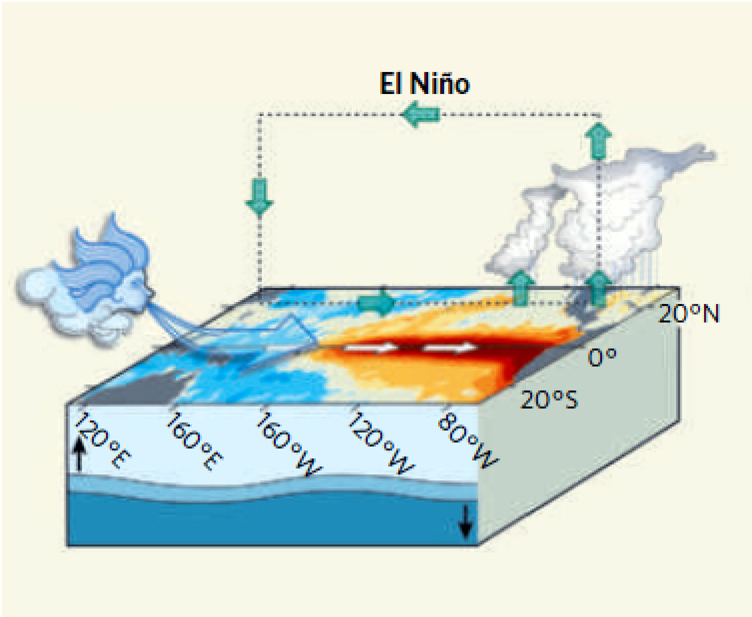



Figure 2.5: Time series of the Southern Oscillation index and Niño 3.4 index. (McPhaden, 2006).

(a)




(b)

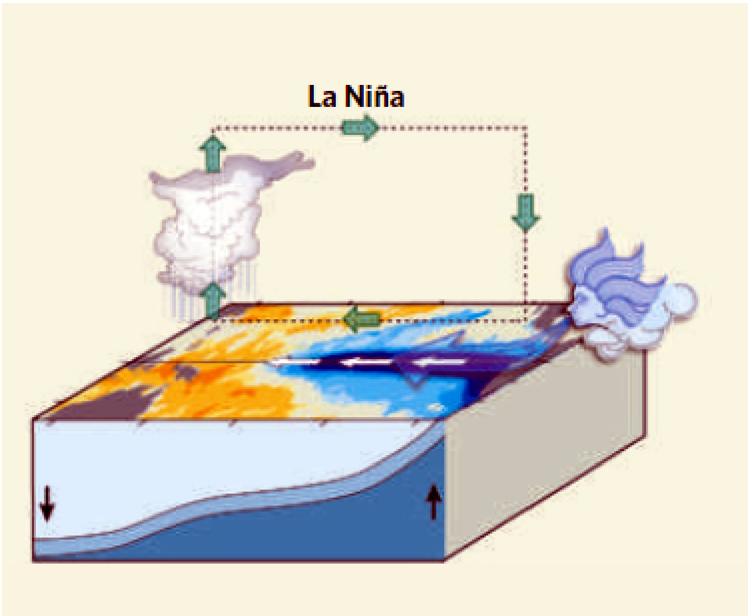



Figure 2.6: (a) El Niño and (b) La Niña events (Ashok, 2009).

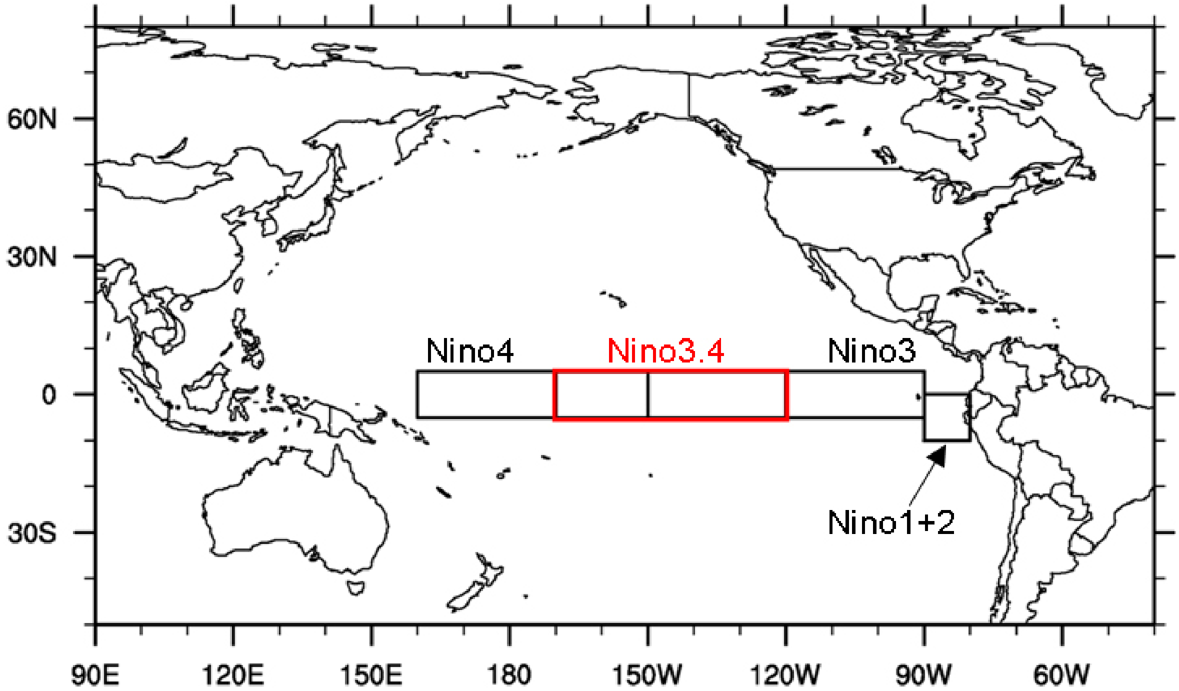


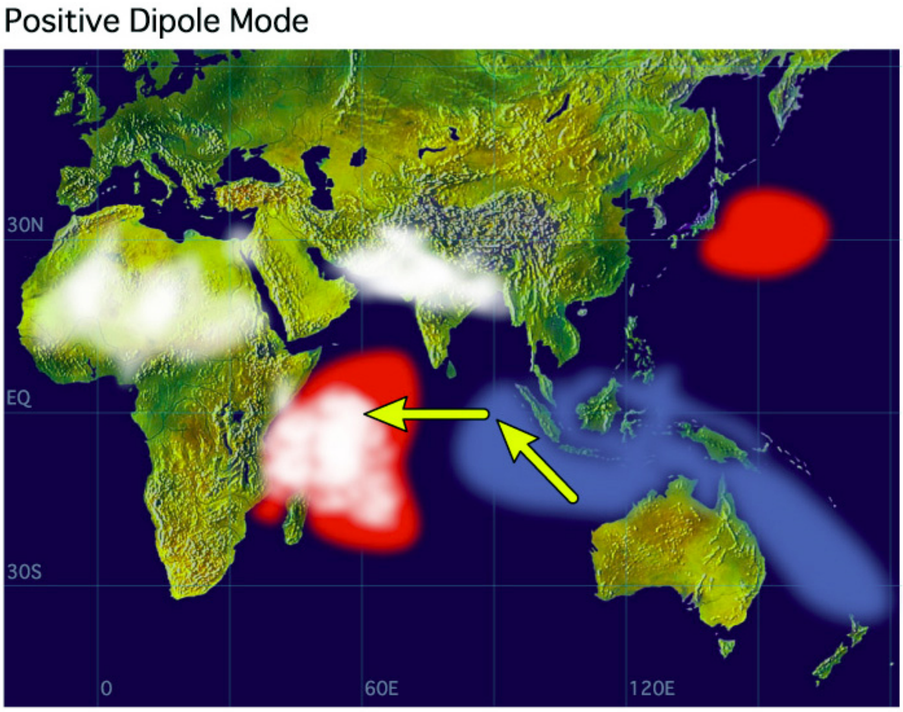

Figure 2.7: Areas used for Niño indices calculations.

### 2.1.4 Indian Ocean Dipole

Although the ENSO is an important phenomenon occurring in the Pacific Ocean, there is an important phenomenon in the Indian Ocean. Internal modes of variability of the Indian Ocean are recognized in 1999 by using EOF for analysis on SSTA in the tropical Indian Ocean basin. First mode of the SSTA is monopole, known as an Indian Ocean Basin mode (IOB) accounting for 30% of the total variance, and having a high correlation with the Niño 3 index. For the second mode, the dipole pattern is apparent. This mode explains about 12% of the total variance. The pattern has been called the “Indian Ocean Dipole (IOD)” as shown in Figure 2.8. The dipole mode are identified by an index time series which determined from the difference in SST anomaly between the two regions, the tropical western Indian Ocean (50°E-70°E, 10°S-10°N) and the tropical southeastern Indian Ocean (90°E-110°E, 10°S-Equator). A name of the index is the dipole mode index denoted as DMI. The positive phase of IOD indicates to the SSTA in the tropical western Indian Ocean than the tropical southeastern Indian Ocean (Figure 2.8a), whereas the negative phase exhibits the warm pool in the tropical southeastern Indian Ocean and the cool pool in the tropical western Indian Ocean (Figure 2.8b). Furthermore, the IOD correlated to changes in rainfall, leading to foods and drought in east Africa and Indonesia, respectively (Saji, 1999).

Since Saji (1999) discovered the IOD, many research studies have been undertaken to understand this phenomenon more (Behera, 2003; Luo, 2008, 2010; Krishnan, 2011). In 2006 to 2007, there were consecutive occurrences of IOD events. The simulation of coupled global climate model, SINTEX-F successfully predicted both evidences in the 2006 and 2007 for positive IOD events up to around four seasons ahead (Luo, 2008). There is linkage between IOD and ENSO such as an extreme IOD perhaps significantly improve El Niño and its onset forecast that give more understanding on prediction of climate in the Indian Ocean and the Pacific (Luo, 2010). For the relationship between SST in the tropical Indian Ocean and the Asian monsoon circulation, the IOB appears to have stronger relationship with the Asian summer monsoon than the IOD mode, but the IOD appears to shows a stronger relationship with the Asian winter monsoon than that of the Asian summer monsoon (Yang, 2010).

(a)




(b)

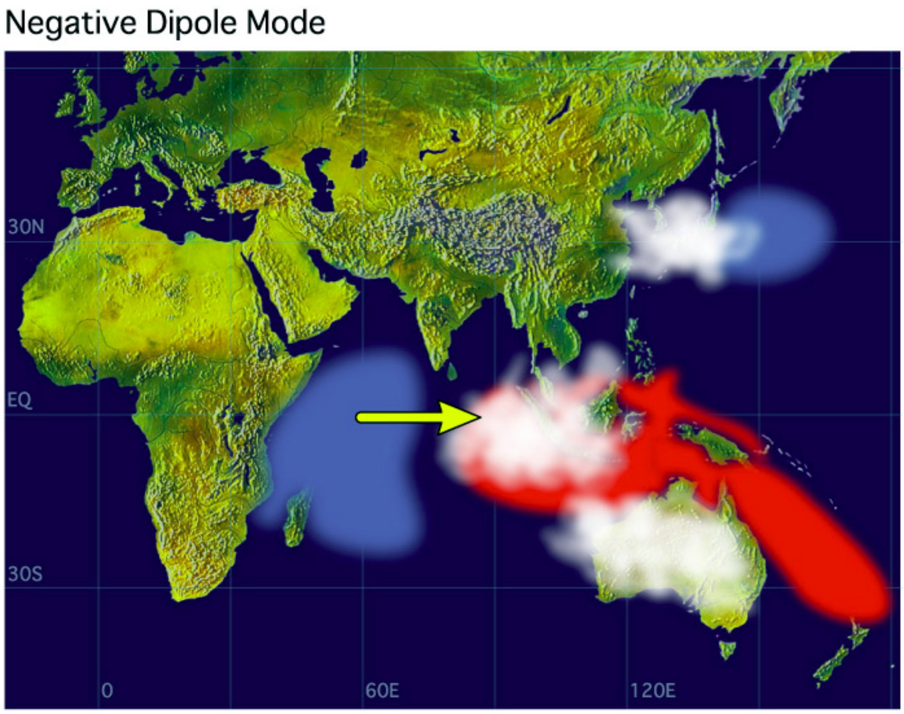



Figure 2.8: Schematic of (a) positive and (b) negative IOD events (source: the Japan Agency for Marine-Earth Science and Technology (JAMSTEC), http://www.jamstec.go.jp/frcgc/research/d1/iod/e/iod/about_iod.html).

## 2.2 Statistical Methods

There are various statistical methods that have been used in climate analyses and atmospheric sciences. In this study, the major methods used for analyses are the empirical orthogonal function (EOF), regression, and correlation analyses. The EOF method has been used to reveal variability of climate feature. The others are regression method using to summarize the relationship between two variables in terms of linear relationship by simple linear regression, and Pearson correlation using to investigate the correlation of paired data.

### 2.2.1 Empirical Orthogonal Function

The EOF analysis has been commonly used for the analysis of atmospheric data. It is similar to the principal component analysis denoted as PCA, and they are the same procedure. The EOF method decomposes large data set into small pieces of data that are linear combinations of the original data, and aims to determine uncorrelated linear combinations (Wilks, 2006). Originally, the EOF was used to decompose a continuous space-time data into uncorrelated combinations that explain maximum variance (Hannachi, 2007). The method has been often used for analyses of scalar variables such as precipitation, sea surface temperature, and temperature (Diaz, 2001; Limsakul, 2008; Li, 2010; Limsakul, 2010). The important procedure for the analysis is data forming for gridded data, which described by Hannachi (2007) as follows:

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{bmatrix}, \tag{2.1}$$

where, $p$ and $n$ denoted as total numbers of grid points of data and observed times, respectively. Next is the determination of time average of data for each grid point that is

$$\overline{x}_i = \frac{1}{n}\sum_{k=n}^{t}\overline{x}_{ki}\ , \tag{2.2}$$

where $i$ = 1, .., $p$ and $k$ = 1, …, $n$. The time average data were used to determine anomalies such as

$$x'_{t,i} = x_{t,i} - \overline{x}_i , \tag{2.3}$$

The anomalies ($x'$) were used as elements to form a matrix of anomalies ($\mathbf{X}'$) for further analysis. Next, it presents the procedure of EOF analysis. Firstly, the sample covariance matrix used for analysis is defined by

$$\mathbf{S} = \frac{1}{n}\mathbf{X}'^{T}\mathbf{X}' . \tag{2.4}$$

The EOF result is obtained to determine the solution of the eigenproblem, that is:

$$\mathbf{SE} = \lambda\mathbf{E} . \tag{2.5}$$

The result consists of eigenvalues ($\lambda_k$) and eigenvectors ($\mathbf{E}_k$), where $k$ = 1, …, $p$. The eigenvector is often used to present geographical information, whereas the eigenvalues is used to measure the variance of data accounted for corresponding eigenvector direction. To present the sequence of eigenvalues, they are sorted in decreasing direction. It is useful to account variance for each mode in percentage as

$$\frac{\lambda_k}{\sum_{k=1}^{p}\lambda_k}\times 100 . \tag{2.6}$$

Although the eigenvectors used to present spatial patterns, it is important to present the corresponding temporal variation. Projection of anomalies field onto the $k^{\text{th}}$ eigenvector as

$$a_k = \mathbf{XE}_k . \tag{2.7}$$

where $a_k$ is the $k^{\text{th}}$ principal component (PC), which is often used to describe in terms of time variation. So the $k^{\text{th}}$ eigenvalue represent variance of the corresponding mode, the $k^{\text{th}}$ eigenvector reveals its spatial pattern, and the $k^{\text{th}}$ PC presents temporal variation that are result of EOF analysis.

There are many applications of EOF analyses (Diaz, 2001; Hannachi, 2007; Limsakul, 2008; Li, 2010; Limsakul, 2010). Although the conventional EOF is usually used to decompose the scalar data field, such as temperature and precipitation (Diaz, 2001; Limsakul, 2008; Li, 2010; Limsakul, 2010) into space and time, there are vector parameters in an environment such as wind that are stored in the format of pairs of scalar fields. They consist of zonal and meridional wind components. To analyse a vector parameter, there is a method of EOF analysis similar to the conventional EOF for the paired data analysis, which is the EOF analysis for complex numbers (Hannachi, 2007). For example, Hardy (1978) applied the EOF analysis on observed wind data from meteorological monitoring stations. The data were formed in the exponential form of complex numbers (Hardy, 1978) as shown in Figure 2.9 to be the elements of the matrix for the EOF analysis. The analysis of complex numbers takes the advantage of the Hermitian matrix to give result that satisfies orthogonality (Hardy, 1978).

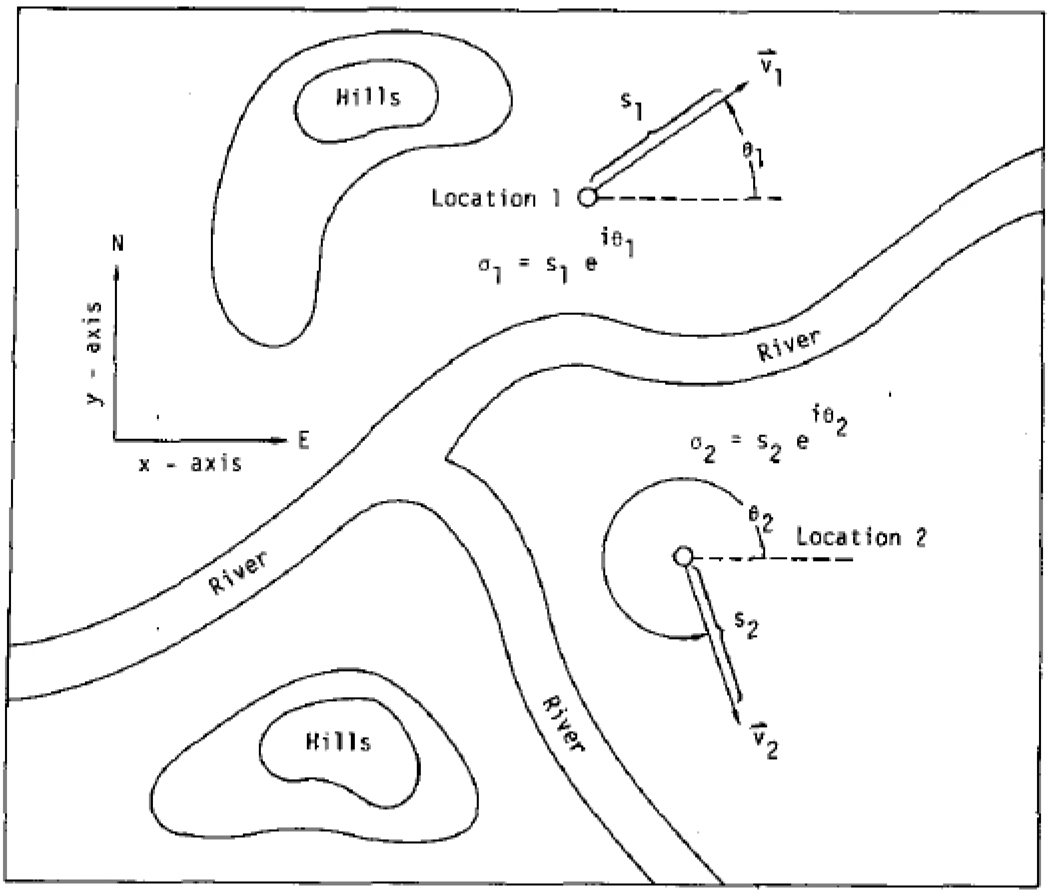



Figure 2.9: A schematic concept used to form a wind vector in exponential complex number form (Hardy, 1978).

There is a special matrix storing complex number elements, the Hermitian matrix, that gives real values for all eigenvalues from the solution of an eigenproblem. It is the answer for EOF analysis for wind data forming in complex numbers. However, the sample data used for the EOF analysis depend on difference of interest leading to has difference in size of the matrix. The difference in dimension of the matrix of complex numbers is required to treat as the square Hermitian matrix for EOF analysis.

Any matrix of complex numbers $\mathbf{H} = \mathbf{H}^*$, a Hermitian matrix, where $\mathbf{H}^*$ denotes the corresponding conjugate transpose of $\mathbf{H}$. Suppose that any sample matrices $\mathbf{S}$ and $\mathbf{A}$ of complex numbers,

$$(\mathbf{SA})^* = \mathbf{A}^*\mathbf{S}^* . \tag{2.8}$$

Let $\mathbf{A}$ equals $\mathbf{S}^*$, and $*$ denotes the conjugate transpose for the matrix of complex numbers. To form a covariance matrix of a matrix $\mathbf{S}$, thus

$$(\mathbf{SS}^*)^* = (\mathbf{S}^*)^*\mathbf{S}^* . \tag{2.9}$$

Because the property of the complex number matrix that is $(\mathbf{A}^*)^* = \mathbf{A}$ thus

$$(\mathbf{SS}^*)^* = \mathbf{SS}^* . \tag{2.10}$$

Let the matrix S has dimension of $N \times M$ , and $\mathbf{H} = \mathbf{SS}^*$, then Equation 3 becomes

$$\mathbf{H}^*_{N\times M} = \mathbf{H}_{N\times M} . \tag{2.11}$$

These show that the matrix $\mathbf{H}$ meets the criterion of the Hermitian matrix mentioned before. Therefore, it can form a vector parameter in the form of a complex number, construct the matrix of complex number elements, and make use of the Hermitian matrix for EOF analysis as described in study of Hardy (1977).

**2.2.2 Regression and Correlation Analyses**

To understand the linear relationship between two variables, a regression method is one of the methods often used. Generally, two variables, x and y used for the regression are called independent variable and dependent variable, respectively. Sometimes, the independent variable (*x*) has been called predictor, and the dependent variable (*y*) has been called the predictand. The representative linear relation between two variables is revealed by a simple linear equation, such as The EOF analysis has been commonly used

$$\hat{y} = a + bx, \tag{2.12}$$

where $\hat{y}$, *a*, and *b* are the predicted dependent variable, the constant, and the slope. The slope can be interpreted as rate of change for predicted dependent variable to predictor (Wilks, 2006). For climate analysis, it is useful to shows spatial pattern of the predicted value resulted from predictor. An example of using regression method is the study of Zhou (2010) that shows the regression map of winter (December to February) SST to normalized rainfall anomalies of period of January to March. This map shows meaning that is the change of SST respect to the normalized JFM rainfall anomalies in terms of spatial distribution as shown in Figure 2.10.

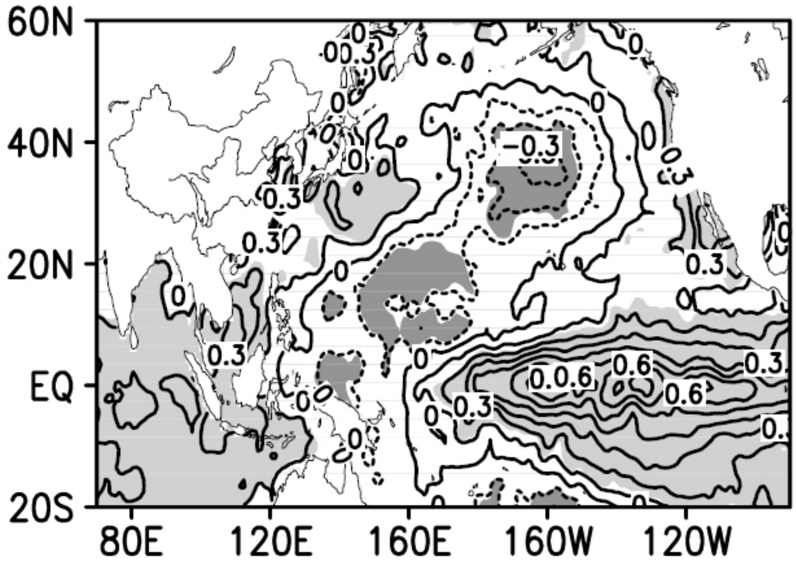



Figure 2.10: Regression map over South China of DJF SST with respect to normalized JFM rainfall anomalies (units: °C), whereas the significance at 0.05 level is denoted by shading (Zhou, 2010).

Another one often used to reveal the relation between climatic variables is correlation analysis. The ordinary (Pearson) correlation has been used to show linear relationship between two variables. The association between two variables can be measured by the correlation coefficient. The correlation has been determined by the ratio of covariance of two sample variables to the product of their standard deviations as follows:

$$r_{xy} = \frac{\mathrm{Cov}(x, y)}{\mathrm{S}_x \mathrm{S}_y}, \tag{2.13}$$

where $r_{xy}$, $\mathrm{Cov}(x, y)$, $\mathrm{S}_x$, and $\mathrm{S}_y$ are the Pearson correlation coefficient, covariance of *x* and *y* variables, and standard deviations of *x* and *y*, respectively. The coefficient value is bounded in a range $-1$ to 1. When the coefficient equals $-1$ meaning that there is negative perfect correlation, while it equals 1 meaning that there is positive perfect correlation (Wilks, 2006).

To display and compare correlations among many variables, the correlation matrix, or presenting as a table, is useful. For example, Wang (2010) presented the correlation coefficients as a correlation matrix of many EAWM indices to ENSO and AO indices leading to give a good interpretation and comparison of the EAWM indices. However, the correlation matrix is not suitable for a large number of data pairs or presenting in spatial pattern. Thus, it is better to geographically arrange correlation coefficients in map form to present useful information (Wilks, 2006). The correlation map has been used in fields of atmospheric, oceanic, and climatic analyses such as the study of Gong (2001) as shown in Figure 2.11.

On the other hand, correlation coefficients are used to indicate the relationship of two observed sample time series, which are assumed to be independent of such. However, they are rarely independent. Autocorrelation due to observation data sampling can be reflected by the first order (lag$-1$) of autocorrelation coefficient, and resulting to reduce the number of independent sample size (*N*). The adjustment of sample size called an effective sample size ($N_{ef}$) is required, and it can affect the confidence interval. The effective sample size for correlation coefficient of two time series data can be determined as follows (Trenberth, 1984):

$$N_{ef} = \frac{N}{(1 + r_{1x}r_{1y} + r_{2x}r_{2y} + ...)}, \tag{2.14}$$

where $r_{1x}$ and $r_{1y}$ are the first order autocorrelation coefficients for time series data of x and y, respectively, and $r_{2x}$ and $r_{2y}$ are the coefficients for the second order.

(a)


(b)

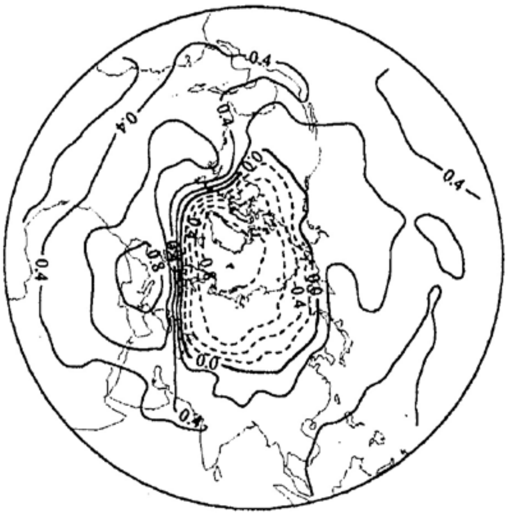



Figure 2.11: Correlation maps of SLP to (a) SH and (b) AO indices (Gong, 2001).