RAINFALL VARIABILITY OVER THAILAND AND ITS POSSIBLE TELECONNECTION TO CLIMATE MODES

MS. SARINYA KIRTPHAIBOON
ID: 55300700402

A THESIS SUBMITTED AS A PART OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN ENVIRONMENTAL TECHNOLOGY

THE JOINT GRADUATE SCHOOL OF ENERGY AND ENVIRONMENT AT KING MONGKUT'S UNIVERSITY OF TECHNOLOGY THONBURI

2ND SEMESTER 2013

Rainfall Variability Over Thailand and Its Possible Teleconnection to Climate Modes

Ms. Sarinya Kirtphaiboon
ID: 55300700402

A Thesis Submitted as a Part of the Requirements
for the Degree of Master of Science
in Environmental Technology

The Joint Graduate School of Energy and Environment
at King Mongkut's University of Technology Thonburi

2nd Semester 2013

Thesis Committee

Prungchan Wongwises .... (Assoc. Prof. Dr. Prungchan Wongwises)	Advisor
U. Humpheries .... (Assoc. Prof. Dr. Usa Humpheries)	Co-advisor
Atsamon Limsakul .... (Dr. Atsamon Limsakul)	Member
Angkool Wangwongchai .... (Dr. Angkool Wangwongchai)	Member
Patama Singhruck .... (Dr. Patama Singhruck)	External Examiner

Thesis Title: Rainfall Variability Over Thailand and Its Possible Teleconnection to Climate Modes

Student's name, organization and telephone/fax numbers/email

Miss Sarinya Kirtphaiboon

The Joint Graduate School of Energy and Environment (JGSEE)

King Mongkut's University of Technology Thonburi (KMUTT)

126 Pracha Uthit Rd., Bangmod, Tungkru, Bangkok 10140 Thailand

Telephone: 0-8156-52501

Email: tukatyzaa@hotmail.co.th

Advisor's name, organization and telephone/fax numbers/email

Assoc. Prof. Dr. Prungchan Wongwises

The Joint Graduate School of Energy and Environment

King Mongkut's University of Technology Thonburi (KMUTT)

126 Pracha Uthit Rd., Bangmod, Tungkru, Bangkok 10140 Thailand

Telephone: 0-2872-6736

Email: prungchan.won@kmutt.ac.th

Co-advisor's name, organization and telephone/fax numbers/email

Assoc. Prof. Dr. Usa Humphries

Department of Mathematics, Faculty of Science

King Mongkut's University of Technology Thonburi (KMUTT)

126 Pracha Uthit Rd., Bangmod, Tungkru, Bangkok 10140 Thailand

Telephone: 0-2470-8831

Email: usa.wan@kmutt.ac.th

Topic: Rainfall Variability over Thailand and Its Possible Teleconnection to Climate Modes

Name of student: Ms. Sarinya Kirtphaiboon **Student ID:** 55300700402

Name of Advisor: Assoc. Prof. Dr. Prungchan Wongwises

Name of Co-Advisor: Assoc. Prof. Dr. Usa Humphries

ABSTRACT

This study analyzed the monthly rainfall data of the Global Precipitation Climatology Centre (GPCC) over Thailand, covering the period from 1971 to 2010 using the Empirical Orthogonal Function (EOF) technique. The most dominant mode accounts for 21.6% of the total variance, when interpreting the rainfall from the principal component time series and associated eigenvector. The next part was a study of the relationship with ENSO using smoothed anomalies with the low-pass filter between the Niño 3.4 index and the principal component time series It was found that the Niño 3.4 index leads the rainfall anomalies by 4 months. But it was of the relationship with Indian Ocean Dipole (IOD) using smoothed anomalies with the low-pass filter between Dipole Mode Index (DMI) and the principal component time series. It was found that the rainfall variability over Thailand has less correlation with IOD. This study used ENSO events divided into weak and strong intensity classes. It was based on composites of fourteen weak La Niña events, six strong La Niña events, twelve weak El Niño events and six strong El Niño events. It was found that there was high rainfall in La Niña events, whereas there was low rainfall in El Niño events. Also, we constructed the corresponding wind circulation and the sea level pressure in order to better understand the mechanisms associated with this events that have affected rainfall variability over Thailand.

Keywords: Rainfall over Thailand; EOF; GPCC; El Niño; La Niña; IOD

ACKNOWLEDGEMENTS

I would like to express my thanks to many people who encouraged and helped me throughout the course of this study: Assoc. Prof. Dr. Prungchan Wongwises, my thesis supervisor and Assoc. Prof. Dr. Usa Humphries, my thesis co-advisor, for their great elaborate guidance, friendly encouragement and tolerance with problems during the course of my study. I am also grateful to the members of the examination committee, Dr. Atsamon Limsakul from Environmental Research and Training Center, Department of Environmental Quality Promotion, Ministry of Natural Resources and Environment, Dr. Angkool Wangwongchai Department of Mathematics, Faculty of Sciences, King Mongkut's University of Technology Thonburi and Dr. Patama Singhruck from Department of Mathematics, Faculty of Sciences, Chulalongkorn University for their valuable comments and suggestions. And thanks to Mr. Sirapong Sooktawee, best brother in Asian Tropical Climate and Ocean-Atmosphere Modeling Laboratory, King Mongkut's University of Technology Thonburi for the suggestion and solving the problems during the course of study.

I would like to acknowledge the Global Precipitation Climatology Centre (GPCC) for the data set, the developer of the NCAR Command Language (NCL) for analysis and graphic visualization tool, the Joint Graduate School of Energy and Environment (JGSEE) and the Centre of Excellence on Energy Technology and Environment (CEE), Science and Technology Postgraduate Education, Research Development Office (PERDO) for its financial support of my Master degree studies, and the lecturers in JGSEE for their instructions, which were very useful for my thesis.

Last, but not least, I express my warm thanks to my family who has always been my inspiration and my guiding light. They have always supported me with all their love.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Cont')

CHAPTER 1
INTRODUCTION

1.1 Rationale

Thailand has an area of 513,120 square kilometers (198,120 square miles), and is situated at the southeastern part of the Indochinese peninsula, with latitude between $5°N$ to $20°N$ and longitude between $97°E$ to $105°E$. Thailand is under the influence of monsoon winds of seasonal character such as southwest monsoon and northeast monsoon. The southwest monsoon started in May and brings a stream of warm moist air from the Indian ocean towards Thailand, which causes abundant rain over the country, especially the wind ward side of the mountains. Rainfall during this period is not only caused by the southwest monsoon but also by the Inter Tropical Convergence Zone (ITCZ) and tropical cyclones, which produce a large amount of rainfall. May is the period of first arrival of the ITCZ to the southern part of Thailand. It moves northward rapidly and lies across southern China around June to early July causing a dry spell over upper Thailand. The ITCZ then moves south and lies over the northern and northeastern parts of Thailand in August and later over the central and southern part in September and October respectively. The northeast monsoon starts in October and brings the cold and dry air from the anticyclone in China over major parts of Thailand, especially the northern and northeastern parts. In the southern part, this monsoon causes mild weather and abundant rain along the eastern coast.

Agriculture and related sectors of Thailand are favoured by the Thai tropical monsoon climate. But in some abnormal years, different disasters, owing to the weather and climate, may be happened to cause the considerable damage and even the loss of life. The climate variability has been concerned about various aspects for example the characteristic of climate, rainfall and temperature. The impact of climate variability involves many problems such as drought and flood. Then in this research, the rainfall variability will be studied by using the accumulate of monthly rainfall over Thailand from the GPCC data [Schneider, 2011a] version V6 for the period 1901 to 2010 with a spatial resolution of $0.5° \times 0.5°$ latitude by longitude. The Empirical Orthogonal Function (EOF) Analysis will be used to analyze large data with linkage to relate in its possible teleconnection to the climate mode.

The purposes of this research are to study the rainfall variability over Thailand and to analyze the rainfall variability and its possible teleconnection to global or regional climate mode.

1.2 Literature Review

Pribadi et al. (2012) studied diurnal rainfall variations over three sub-domains (Java, Sumatra and Cerum islands) with distinct annual rainfall cycles within Indonesia using Empirical Orthogonal Function analysis. To use the resolution data is 3 hour of temporal and $0.25° \times 0.25°$ of the spatial from the Tropical Rainfall Measuring Mission during 2000-2009. It found that the eigenvalues of Boxes A, B and C are selected by mode 1 and 2 because they are represented a very large proportion of the variance that is 97%, 93% and 93%. To interpret the diurnal rainfall from the principle component (PCs) time series and associated eigenvectors (EOFs) found that box A shows amplitude over land duration 1500-1800 LST and over ocean duration 0900-1200 LST, box B shows amplitude over land duration 1800-2100 LST and over ocean duration 0600-1200 LST and Box C shows amplitude over land 1200-1800 LST and over ocean 0300-1200 LST.

Juneng and Tangang (2005) showed that in Southeast Asia, rainfall (SEAR) anomalies depend strongly on the phases of El Nino and La Nina using an extended empirical orthogonal function analysis. The dominant mode of SEAR anomalies evolves northeastward throughout a period from the summer when El Nino develops to spring the following year when the event weakens which is consistent with the northeastward migration of the ENSO-related anomalous outgoing radiation field. The evolution of the dominant mode of SEAR anomalies is in tandem with the evolution of ENSO related sea surface temperature anomalies. The strong ENSO-related anomaly tends to reside in regions south of the equator during boreal summer whereas the strong ENSO related anomaly tends to reside in regions north of the equator during boreal winter. The anomalous low-level circulation associated with ENSO-related SEAR anomaly indicates the strengthening and weakening of two off-equatorial anticyclones, one over the Southern Indian Ocean and the other over the western North Pacific.

Trenberth et al. (2002) studied the origins of delayed increases in global surface temperature accompanying El Nino events and explored the implications for the role of diabatic processes in ENSO using correlation and regression analysis of global mean surface temperatures, zonal means and fields of sea surface temperatures, land surface temperatures, precipitation, outgoing long wave radiation, vertically integrated diabatic heating and divergence of atmospheric energy transports, and ocean heat content in the Pacific. ENSO linearly accounts for 0.06°C of global surface temperature increase during 1950–1998. The warming events peak 3 months after SSTs in the Niño 3.4 region, somewhat less than is found in previous studies. The warming at the surface progressively extends to about ±30° latitude with lags of several months. A major part of the ocean heat loss to the atmosphere is through evaporation and thus is realized in the atmosphere as latent heating in precipitation, which drives teleconnections. Reduced precipitation and increased solar radiation in Australia, Southeast Asia, parts of Africa, and northern South America contribute to surface warming that peaks several months after the El Niño event.

1.3 Research Objectives

This study aims to:

1. Analyze the rainfall variability over Thailand,
2. Identify possible teleconnection between the dominant modes of rainfall variability over Thailand and climate modes.

1.4 Scope of the Research

The implementation of this work is as follows:

1. Analyze the dominant modes of rainfall variability over Thailand based on monthly data,
2. Analyze temporal variation in terms of frequency and period,
3. Identify possible teleconnection between the dominant modes of rainfall variability over Thailand and sea surface temperatures (SST) - related climate modes such as ENSO and IOD,
4. Interpretation of possible linking mechanisms between rainfall variability and climate modes.

CHAPTER 2

THEORIES

This chapter presents related theories, such as the statistical techniques and climate features.

2.1 Empirical Orthogonal Function Analysis

This study used the Empirical Orthogonal Function (EOF) analysis, which is a technique of descriptive multivariate statistics. It is a useful technique in the meteorology and oceanography. It is also known as Principle Component Analysis (PCA) which based on a linear transformation to extract information from the large data sets by decomposing to the orthogonal basis function while retaining as much as possible of the variations present in the data sets. EOF is considered from the covariance matrix of data and it decomposed a space-time data into spatial patterns and associated time indices. Then, it received the eigenvalues, eigenvectors and principle component time series. The eigenvectors are few data which is extracted from the large data sets and they are corresponding the eigenvalues of each mode where the number of EOF modes is equal to the number of stations. Because, it is without a certain criteria in selecting the number of modes, then it depends mainly on the purpose of the study (Hannachi et al., 2007; Compagnucci and Richer, 2008; William and Richard, 2001).

The process and methods in this study are as follows:

2.1.1 Formatting and preparing data for analysis: The EOF technique is essentially a matrix method by which the data can be analyzed when arranged into matrices. Linear transforms a continuous space-time field $\mathbf{Z} = \left[z_{tx}\right]_{n\times p}$ where $\mathbf{Z}$ is matrix of rainfall data and the element of z_{tx} is the rainfall amount collected at the time t ($t =$ 1,2,3,…,n) and station x ($x =$ 1,2,3,…,p). The formatting of data is shown as Figure 2.1.

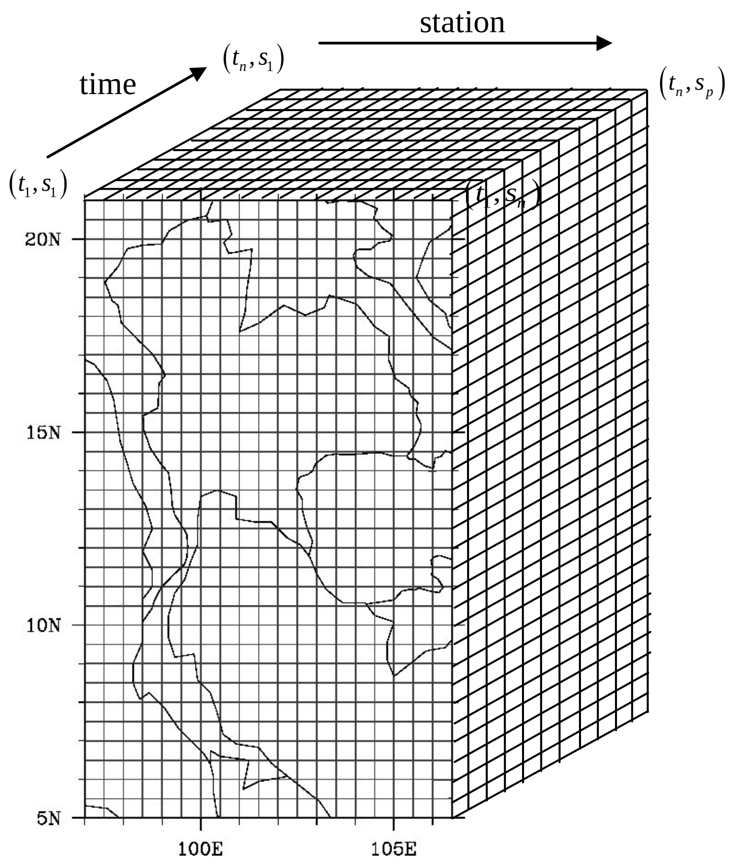


Figure 2.1: The data formatting

The matrix $\mathbf{Z}$ is arranged in a matrix as follows:

$$\mathbf{Z} \;=\; \left[z_{tx}\right]_{n\times p} \;=\; \begin{array}{c} \text{Station } (x_1, x_2, \ldots, x_p) \longrightarrow \\ \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1p} \\ z_{21} & z_{22} & \cdots & z_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{np} \end{bmatrix}_{n\times p} \end{array} \downarrow \text{Time } (t_1, t_2, \ldots, t_n) \tag{2.1}$$

Preparing data for analysis is used to compute the anomaly, which is different from the long-term average (climatology) in order to minimize errors arising from higher or lower than the normal data of each station or time (t). The time average mean $(\overline{\mathbf{Z}})$ is calculated from the average of column of matrix $\mathbf{Z}$ as follow:

$$\overline{\mathbf{Z}} = \begin{bmatrix} \overline{z}_1 & \overline{z}_2 & \ldots & \overline{z}_p \end{bmatrix} \tag{2.2}$$

where $\bar{z}_x = \frac{1}{n}\sum_{t=1}^{n}\mathbf{Z}_{tx}$

The anomalies matrix $(\mathbf{Z}')$ is computed from the time average mean, which is defined by

$$\mathbf{Z}' = \mathbf{Z} - \overline{\mathbf{Z}} \tag{2.3}$$

or, in matrix form is written as:

$$\mathbf{Z}' = \left[z'_{tx}\right]_{n\times p} = \begin{bmatrix} z_{11}-\bar{z}_1 & z_{12}-\bar{z}_2 & \cdots & z_{1p}-\bar{z}_p \\ z_{21}-\bar{z}_1 & z_{22}-\bar{z}_2 & \cdots & z_{2p}-\bar{z}_p \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1}-\bar{z}_1 & z_{n2}-\bar{z}_2 & \cdots & z_{np}-\bar{z}_p \end{bmatrix}_{n\times p}$$

$$= \begin{bmatrix} z'_{11} & z'_{12} & \cdots & z'_{1p} \\ z'_{21} & z'_{22} & \cdots & z'_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ z'_{n1} & z'_{n2} & \cdots & z'_{np} \end{bmatrix}_{n\times p} \tag{2.4}$$

where z'_{tx} is the anomalous data of matrix $\mathbf{Z}'$, which is collected at the time t (t = 1,2,3,…,n) and station x (x = 1,2,3,…,p).

2.1.2 Computation

2.1.2.1 To find covariance matrix

The covariance matrix is determined by

$$\mathbf{R} = \frac{1}{n-1}\mathbf{Z}'^{\mathrm{T}}\mathbf{Z}' \tag{2.5}$$

where $\mathbf{Z}'^{T}$ is the transpose matrix of $\mathbf{Z}'$ and the matrix form is written as:

$$\mathbf{Z}'^{T} \;=\; \begin{bmatrix} z'_{11} & z'_{21} & \cdots & z'_{n1} \\ z'_{12} & z'_{22} & \cdots & z'_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ z'_{1p} & z'_{2p} & \cdots & z'_{np} \end{bmatrix}_{p\times n}$$

The covariance matrix can be calculated as:

$$\mathbf{R}_{p\times p} \;=\; \frac{1}{n-1}\begin{bmatrix} z'_{11} & z'_{21} & \cdots & z'_{n1} \\ z'_{12} & z'_{22} & \cdots & z'_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ z'_{1p} & z'_{2p} & \cdots & z'_{np} \end{bmatrix}_{p\times n}\begin{bmatrix} z'_{11} & z'_{12} & \cdots & z'_{1p} \\ z'_{21} & z'_{22} & \cdots & z'_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ z'_{n1} & z'_{n2} & \cdots & z'_{np} \end{bmatrix}_{n\times p}$$

$$\mathbf{R}_{p\times p} \;=\; \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & r_{pp} \end{bmatrix}_{p\times p} \tag{2.6}$$

where $\mathbf{R}$ is symmetric matrix and the element r_{ij} of the matrix $\mathbf{R}$ is the value of the covariance between the time series of the field at any pair of grid points that $i, j = 1, 2, 3, ..., p$

2.1.2.2 To compute the eigenvalues and their corresponding eigenvectors

By multiplying matrix $\mathbf{R}$ with some matrix $(\mathbf{E})$ is equal to the matrix $(\mathbf{E})$ multiplied by a scalar (λ) as follows:

$$\mathbf{RE} \;=\; \mathbf{E}\lambda$$

$$(\mathbf{RE}) - (\mathbf{E}\lambda) \;=\; \mathbf{0}$$

$$(\mathbf{R} - (\lambda\mathbf{I}))\mathbf{E} \;=\; \mathbf{0}$$

$$(\mathbf{R} - \mathbf{L})\mathbf{E} \;=\; \mathbf{0} \tag{2.7}$$

where $\mathbf{I}$ is the identity matrix with dimension $p \times p$. The elements of the matrix $\mathbf{I}$ on the diagonal are equal to 1 while other elements are equal to 0.

$\mathbf{L}$ is the diagonal matrix containing eigenvalues λ_k ($k = 1,2,3,\ldots,p$) of matrix $\mathbf{R}$, in which the other elements are equal to 0.

$$\mathbf{L} = \lambda\mathbf{I} = \begin{bmatrix} \lambda & 0 & \ldots & 0 \\ 0 & \lambda & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda \end{bmatrix}_{p\times p} \qquad (2.8)$$

From Equation (2.7), $\mathbf{R}-\mathbf{L}$ is the square matrix of homogeneous equations, which has a nontrivial solution $\lambda_1, \lambda_2, \lambda_3, ..., \lambda_p$, when $\det(\mathbf{R}-\mathbf{L}) = 0$, therefore

$$\begin{vmatrix} r_{11}-\lambda & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22}-\lambda & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & r_{pp}-\lambda \end{vmatrix}_{p\times p} = 0$$

It is easy to understand that the eigenvalues $\lambda_k (k = 1,2,3,..., p)$ of the matrix $\mathbf{R}$ can be arranged in the form of the proportional to the percentage of the variance as $\lambda_1 > \lambda_2 > \lambda_3 > ... > \lambda_p$ where each eigenvalue (λ_k) is accounted for mode k and shown as:

$$\%\text{ variance mode } k = \frac{\lambda_k}{\sum_{i=1}^{p}\lambda_i}\times 100 \qquad (2.9)$$

Substitute λ_k $(k = 1,2,3,..., p)$ in Equation (2.7), we obtain the eigenvectors $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3, ..., \mathbf{E}_p$,

where $\mathbf{E}_k = \begin{bmatrix} \mathbf{E}_1 & \mathbf{E}_2 & \ldots & \mathbf{E}_p \end{bmatrix}$; $k = 1,2,3,\ldots,p$

The eigenvectors $\mathbf{E}_{p\times p}$ from Equation (2.7) can be written in matrix form as:

$$\mathbf{E}_{p\times p} = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1p} \\ e_{21} & e_{22} & \cdots & e_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ e_{p1} & e_{p2} & \cdots & e_{pp} \end{bmatrix}_{p\times p} \qquad (2.10)$$

$$\begin{matrix} \downarrow & \downarrow & & \downarrow \\ \mathbf{E}_1 & \mathbf{E}_2 & & \mathbf{E}_p \end{matrix} \Longrightarrow \text{eigenvector } (\mathbf{E}_k)$$

where $\mathbf{E}_{p\times p}$ is matrix of eigenvectors corresponding to eigenvalues (λ_k)

$\mathbf{E}_k$ is column vectors which are instead of eigenvector of $\mathbf{R}$. The spatial variability of each mode is extracted from the original data sets $(k = 1, 2, 3, ..., p)$

Since the elements of the matrix $\mathbf{Z}'$ are real numbers and the elements of covariance matrix ($\mathbf{R}$) are greater than zero, then every eigenvalue (λ) of the matrix $\mathbf{E}_{p\times p}$ is greater than zero. Therefore, the matrix $\mathbf{E}_{p\times p}$ is orthogonal over space as follows:

$$(\mathbf{E}\mathbf{E}^T)_{p\times p} = (\mathbf{E}^T\mathbf{E})_{p\times p} = \mathbf{I}_{p\times p} \qquad (2.11)$$

2.1.2.3 To compute the principal component time series (time-dependent amplitudes) of each EOF mode.

It can derive by projecting the original data series $\mathbf{Z}'$ onto the eigenvectors $(\mathbf{E}_k)$ of matrix $\mathbf{E}_{p\times p}$ as follows:

$$\mathbf{A}_{n\times p} = \mathbf{Z}'_{n\times p}\mathbf{E}_{p\times p} \qquad (2.12)$$

or, it can be written in matrix form as

$$\mathbf{A}_{n\times p} = \begin{bmatrix} z'_{11} & z'_{12} & \cdots & z'_{1p} \\ z'_{21} & z'_{22} & \cdots & z'_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ z'_{n1} & z'_{n2} & \cdots & z'_{np} \end{bmatrix}_{n\times p} \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1p} \\ e_{21} & e_{22} & \cdots & e_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ e_{p1} & e_{p2} & \cdots & e_{pp} \end{bmatrix}_{p\times p}$$

$$\mathbf{A}_{n\times p} \quad = \quad \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}_{n\times p} \tag{2.13}$$

$$\begin{array}{ccc} \downarrow & \downarrow & \downarrow \\ \mathbf{A}_{t1} & \mathbf{A}_{t2} & \mathbf{A}_{tx} \end{array} \quad \Longrightarrow \quad \text{Principal component time series (PCs)}$$

where $\mathbf{A}_{tx} \ = \ \sum_{j=1}^{p} z'_{tj} e_{jx}$

The elements $\mathbf{A}_{tx}$ are the principal component time series (PCs) or expansion coefficient of EOF at time $(t = 1, 2, ..., n)$ and station $(x = 1, 2, 3, ..., p)$. They are orthogonal over time, which means that the time-averaged covariance of the amplitudes satisfies the expression:

$$\overline{a_{ti} a_{tj}} \ = \ \lambda_i \delta_{ij} \tag{2.14}$$

where $i, j \ = \ 1, 2, 3, ..., p$ and δ_{ij} is the Kronecker delta function:

$$\delta_{ij} = \begin{cases} 1, & j = i \\ 0, & j \neq i \end{cases}$$

The overbar in Equation (2.14) denotes the time-averaged value and the eigenvalues (λ) denotes the variance of each EOF mode calculated by:

$$\lambda_i \ = \ \overline{a_{ti} a_{ti}} \ = \ \frac{1}{n} \sum_{i=1}^{n} \left[a_{ti} \right]^2 \tag{2.15}$$

or, it can be written in matrix form as follows:

$$\mathbf{A}^T \mathbf{A} \ = \ \mathbf{L} \tag{2.16}$$

Since the eigenvectors and the principal component time series (PCs) are orthogonal over space and time, then the sum of variances in the eigenvalues is equal to the sum of variances in the original data as in the expression:

$$\sum_{x=1}^{p}\left\langle \frac{1}{n}\sum_{t=1}^{n}\left[z_{tx}\right]^{2}\right\rangle = \sum_{k=1}^{p}\lambda_{k} \tag{2.17}$$

Finally, the original data can be reconstructed by the eigenvectors and the principal component time series as:

$$\mathbf{Z}'_{tx} = \sum_{i=1}^{p} a_{ti}e^{T}{}_{xi} \tag{2.18}$$

where a_{ti} denotes the principal component time series $(t = 1, 2, 3, ..., n)$

e^{T}_{xi} denotes the transpose matrix of the eigenvectors elements $(x = 1, 2, 3, ..., p)$

In matrix notation:

$$\mathbf{Z}'_{n\times p} = \mathbf{A}_{n\times p}\mathbf{E}^{T}_{p\times p}\,. \tag{2.19}$$

Or,

$$\mathbf{Z}'_{n\times p} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}_{n\times p} \begin{bmatrix} e_{11} & e_{21} & \cdots & e_{p1} \\ e_{12} & e_{22} & \cdots & e_{p2} \\ \vdots & \vdots & \ddots & \vdots \\ e_{1p} & e_{2p} & \cdots & e_{pp} \end{bmatrix}_{p\times p} \tag{2.20}$$

2.2 Correlation Analysis

The correlation is considered in more detail when the time series analysis methods are examined. In statistics, dependence is any statistical relationship between two random variables or two sets of data. Correlation refers to any a board class of statistical relationships involving dependence (Hinkle et al., 1998; Chen and Popovich, 2002).

The most familiar measure of dependence between two variables is Pearson's correlation coefficient, commonly called simply the correlation coefficient (*r*) (Pearson, 1957; O'Brien, 1985; Ware and Benson, 1975). It is sensitive only to a linear relationship between two variables. It is obtained by dividing the covariance $(\mathrm{COV}(x, y))$ of two variables *x* and *y* by the product of their standard deviations (S_x and S_y) (Rodgers and Nicewander, 1988; Emery and Thomson, 2001). The correlation coefficient is defined as:

$$r_{xy} = \frac{\mathrm{COV}(x, y)}{\mathrm{S}_x \mathrm{S}_y} \tag{2.21}$$

If a series of *n* measurements of *x* and *y* written as x_i *and* y_i where i=1,2,…,N, then the correlation coefficient can be used to estimate the population Pearson correlation *r* between *x* and *y* as follows:

$$r_{xy} = \frac{1}{N-1} \sum_{i=1}^{N} \frac{(x_i - \bar{x})(y_i - \bar{y})}{\mathrm{S}_x \, \mathrm{S}_y}$$

Or,

$$r_{xy} = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N} (x_i - \bar{x})^2 \sum_{i=1}^{N} (y_i - \bar{y})^2}}$$

where $\bar{x}$ and $\bar{y}$ are means of *x* and *y*, and S_x and S_y are the standard deviations of *x* and *y*.

The coefficient value is bounded between -1 to 1. It equals -1 meaning that there is a negative correlation. On the contrary, it equals 1 meaning that there is a positive correlation. If it equals 0, it means that the variables are independent (Francis et al., 1999).

The correlation analysis is useful for comparing correlations among variables. For example, Juneng and Tangang (2005) presented the correlations between the Southeast Asia rainfall anomalies and Niño 3.4 region. Xu et al. (2012) presented the correlations between the variations of subsurface ocean temperature, and eastern and central Pacific ENSO. Hou and Yan (2011) presented the correlation between the total cloud amount anomalies with ENSO over the tropical Pacific.

On the other hand, the correlation coefficients are used to indicate the relationships 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 (Trenberth, 1984). The effective sample size for correlation coefficient of two time series data can be determined as follows:

$$N_{ef} = \frac{N}{(1 + r_{1x}r_{1y} + r_{2x}r_{2y} + ...)} \tag{2.22}$$

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.

2.3 Lanczos Filtering

Lanczos filtering is a Fourier method of filtering digital data (Duchon, 1979). Its principal feature is reduction of the amplitudes of Gibbs oscillation. The Fourier coefficients for the smoothed response function are determined by multiplying the original weight function by a function that Lanczos called the sigma factor. This method can be used to predict the main characteristics of the response function, to compare Lanczos response functions to those from other types of filters, and to extend the analysis to two dimensions.

Digital filtering involves transforming an input data sequence x_t into an output data sequence y_t where t is time, using the linear relationship as follows:

$$\mathrm{y_t} = \sum_{\mathrm{k}=-\infty}^{\infty} w_k x_{t-k} \tag{2.23}$$

in which w_k are suitably chosen weights, k is an entire number.

For an ideal filter, it is shown that

$$w_k = \frac{\sin 2\pi k f_c}{\pi k} \tag{2.24}$$

where f_c is the cut-off frequency, i.e. the frequency at which the response drops from one to zero. If there is a total of 2n-1 weights in the weight function, then in order to suppress the Gibbs oscillation, Lanczos suggested that the ideal response function is convolved for the following rectangular function as follows:

$$h(f)=\begin{cases} n/2f_N, & |f|\le f_N/n \\ 0, & |f|> f_N/n. \end{cases} \tag{2.25}$$

where f_N is the Nyquist frequency with value 0.5 cycle per data interval.

Then, the weight function of relation (2.24) becomes

$$\bar{w}_k=\frac{\sin 2\pi k f_c}{\pi k}\cdot\frac{\sin \pi k/n}{\pi k/n} \tag{2.26}$$

Therefore, it can be seen that the truncated weight function for the smoothed response is the product of that for the ideal filter and a $\sin X/X$ term denoted by sigma and called the sigma factor by Lanczos.

2.4 Testing of Significance

A statistically significant t-test result is one in which a difference between two groups is unlikely to have occurred because the sample happened to be atypical. Statistical significance is determined by the sizes of the differences between the group averages, the sample size, and the standard deviations of the groups. For practical purposes statistical significance suggests that two larger populations from which we sample are actually different. William Sealy Gosset introduced the t-statistic in 1908. The t-test is any statistical hypothesis test in which the test statistic follows a Student's t distribution if the null hypothesis is supported. It can be used to determine if two sets of data are significantly different from each other, and is most commonly applied when the test statistic would follow a normal distribution if the value of a scaling term in the test statistic were known. When the scaling term is unknown and is replaced by an estimate based on the data, the test statistic follows a Student's t distribution (Siegel, 1956).

Most t-test statistics have the form $t = \frac{Z}{s}$, where Z and s are functions of the data. Typically, Z is designed to be sensitive to the alternative hypothesis. s is a scaling parameter that allows the distribution of t to be determined.

As an example, in the one-sample t-test,

$$t = \frac{\overline{X}}{(\sigma/\sqrt{n})} \tag{2.27}$$

where $\overline{X}$ is the sample mean of the data, n is the sample size, and σ is the population standard deviation of the data, s is the sample standard deviation.

The assumptions underlying a t-test are that:

- Z follows a standard normal distribution under the null hypothesis,
- s^2 follows a χ^2 distribution with p degrees of freedom under the null hypothesis, where p is a positive constant.
- Z and s are independent.

In a specific type of t-test, these conditions are consequences of the population being studied, and of the way in which the data are sampled. For example, in the t-test comparing the means of two independent samples, the following assumptions should be met:

- Each of the two populations being compared should follow a normal distribution. This can be tested using a normality test, such as the Shapiro-Wilk or Kolmogorov-Smirnov test, or it can be assessed graphically using a normal quantile plot.
- If using Student's original definition of the t-test, the two populations being compared should have the same variance (testable using F-test, Levene's test, Bartlett's test, or the Brown-Forsythe test; or assessable graphically using a Q-Q plot). If the sample sizes in the two groups being compared are equal, Studeny's original t-test is highly robust to the presence of unequal variances (Markowski and Markoski, 1990).
- The data used to carry out the test should be sampled independently from the two populations being compared. This is generally not testable from the data, but if the

data are known to be dependently sampled, then the classical *t*-test discussed here may give misleading results.

Independent samples *t*-test is used when two separate sets of independent and identically distributed samples are obtained, one from each of two populations being compared. Explicit expressions that can be used to carry out various *t*-test are given below. In each case, the formula for a test statistic that either exactly follows or closely approximates a *t*-distribution under the null hypothesis is given. Also, the appropriate degrees of freedom are given in each case. Each of these statistics can be used to carry out either a one-tailed test or a two-tailed test. Once a *t* value is determined, a p-value can be found using a table of values from Student's t-distribution. If the calculated p-value is below the threshold chosen for statistical significance (usually the 0.10, 0.05, or 0.01 level), then the null hypothesis is rejected in favor of the alternative hypothesis (Elliott and Woodward, 2007).

2.4.1 One-sample *t*-test

In testing the null hypothesis that the population mean is equal to a specified value, one uses the statistic as follows:

$$t = \frac{\overline{x} - \mu_0}{s/\sqrt{n}} \tag{2.28}$$

where $\overline{x}$ is the sample mean, *s* is the sample standard deviation of the sample and *n* is the sample size. The degrees of freedom used in this test are *n*-1. Although the parent population does not need to be normally distributed, the distribution of the population of sample means, $\overline{x}$ is assumed to be normal. By the central limit theorem, if the sampling of the parent population is random, then the sample means will be approximately normal, The degree of approximation will depend on how close the parent population is to a normal distribution and the sample size, *n*.

2.4.2 Two-sample test

2.4.2.1 Equal sample size, equal variance

This test is only used when two sample sizes (that is the number, *n*, of participants of each group) are equal and it can be assumed that two distributions have the same variance. The *t* statistic to test whether the means are different can be calculated as follows:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{S_{x_1x_2} \bullet \sqrt{\frac{2}{n}}} \tag{2.29}$$

where $S_{x_1x_2} = \sqrt{\frac{1}{2}(S_{x_1}^2 + S_{x_2}^2)}$.

Here $S_{x_1x_2}$ is the grand standard deviation (1=group one and 2=group two). $S_{x_1}^2$ and $S_{x_2}^2$ are the unbiased estimators of the variances of the two samples. The denominator of *it* is the standard error of the difference between two means. For significance testing, the degrees of freedom for this test is $2n - 2$, where *n* is the number of participants in each group.

2.4.2.2 Unequal sample sizes, equal variance

This test is used only when it can be assumed that the two distributions have the same variance. The *t* statistic to test whether the means are different can be calculated as follows:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{S_{x_1x_2} \bullet \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \tag{2.30}$$

where $S_{x_1x_2} = \sqrt{\frac{(n_1 - 1)S_{x_1}^2 + (n_2 - 1)S_{x_2}^2}{n_1 + n_2 - 2}}$.

Here $S_{x_1x_2}$ is an estimator of the common standard deviation of the two samples. It is defined in this way so that its square is an unbiased estimator of the common variance whether or not the population means are the same. In these formulate, n is the number of participants (1=group one and 2=group two). $n-1$ is the number of degrees of freedom for either group, and the total sample size minus two (that is n_1+n_2-2) is the total number of degrees of freedom, which is used in significance testing.

2.4.2.3 Equal or Unequal sample sizes, unequal variances

This test, also known as Welch's t-test, is used only when the two population variances are not assumed to be equal and hence must be estimated separately. The t-statistic to test whether the population means are different is calculated as follows:

$$t=\frac{\overline{x}_1-\overline{x}_2}{S_{\overline{x}_1-\overline{x}_2}} \tag{2.31}$$

where $S_{\overline{x}_1-\overline{x}_2}=\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$.

Here S^2 is the unbiased estimator of the variance of the two samples, n is the number of participants (1=group one and 2=group two). For use in significance testing, the distribution of the test statistic is approximated as an ordinary Student's t distribution with the degrees of freedom calculated using $d.f.=\frac{(S_1^2/n_1+S_2^2/n_2)^2}{(S_1^2/n_1)^2/(n_1-1)+(S_2^2/n_2)^2/(n_2-1)}$. This is known as the Welch-Satterthwaite Equation. The true distribution of the test statistic actually depends on the two unknown population variances.

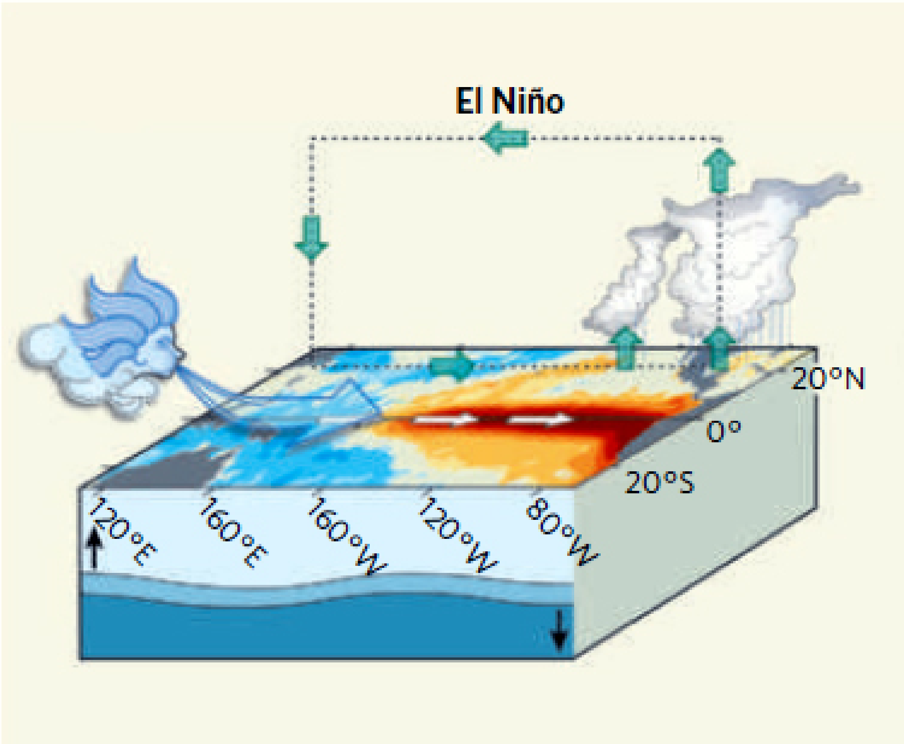
2.5 El Niño-Southern Oscillation

The term 'El Niño' is related to the weak warm ocean current that flows along the coast of Peru and Ecuador about Christmas-time, and associated with the unusually large warm pool in the Pacific Ocean that has linkages with anomalous global climate patterns (Niño means the boy in Spanish). The term El Niño has been tied with the southern oscillation event to be the El Niño-Southern Oscillation (ENSO) which is the interannual interaction of ocean-atmosphere in the tropical Pacific. The reverse event, the cooling of the eastern Pacific waters, was at first called Anti-El Niño, until it was realized that this literally meant the Anti-Christ. To avoid this unfortunate connotation, it was renamed La Niña which means the girl in Spanish (Trenberth, 1997).

The ENSO is the result of a cyclic warming and cooling of the surface ocean of the central and eastern Pacific. This region of the ocean is normally colder than it is as the equatorial location would suggest, mainly due to the influence of northeasterly trade wind, a cold ocean current flowing up the coast of Chile, and to the upwelling of cold deep water off the coast of Peru. At times, the influence of these cold water sources wane, causing the surface of the eastern and central Pacific to warm up under the tropical sun that is El Niño event. This results in heavy rainfall in South America, but severe droughts in eastern Australia. The more intense the El Niño, the more intense and extensive the Australian droughts. At other times, the injection of cold water becomes more intense than usual, causing the surface of the eastern Pacific to cool that is La Niña event. This results in droughts in South America and heavy rainfall, even floods, in eastern Australia. The timing of the cycle is irregular spanning anywhere from 2 years to over 7 years (Rasmusson and Carpenter, 1982; Chang et al., 2000).

The mechanism of El Niño is that the trade winds are weak along the equator, and atmospheric pressure increases in the western Pacific, whereas atmospheric pressure decreases in the eastern Pacific, which is related to anomalous warming of SST in the central and eastern Pacific Ocean with warm water in the western Pacific Ocean moving eastward and the 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 2.2a and Figure 2.2b respectively (McPhaden, 2006; Ashok, 2009).

(a)



(b)

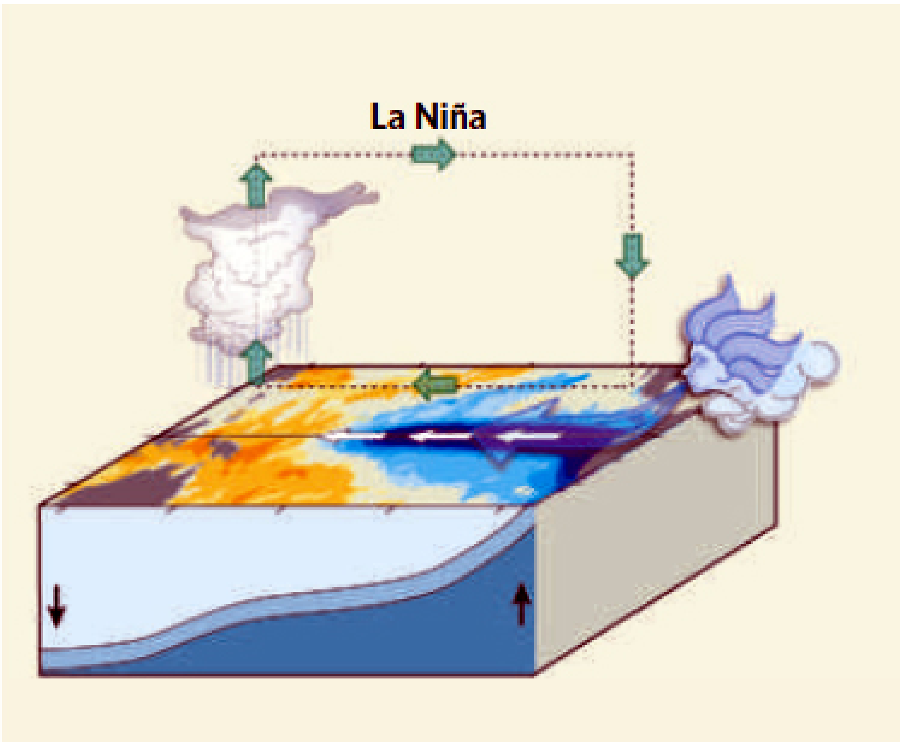


Figure 2.2: Schematic of (a) El Niño and (b) La Niña events (Ashok, 2009).

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 used to characterize El Niño activity would strongly correlate to the time series representing the event given by the analysis as much as possible (Li, 2010). There are many indices using to indicate the ENSO events. Trenberth (2001) suggested the well known indices derived from SST behavior such as Niño 1+2, Niño 3, Niño 4, and Niño 3.4, that represent the SST anomalies averages over corresponding areas, as shown in Figure 2.3. 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).

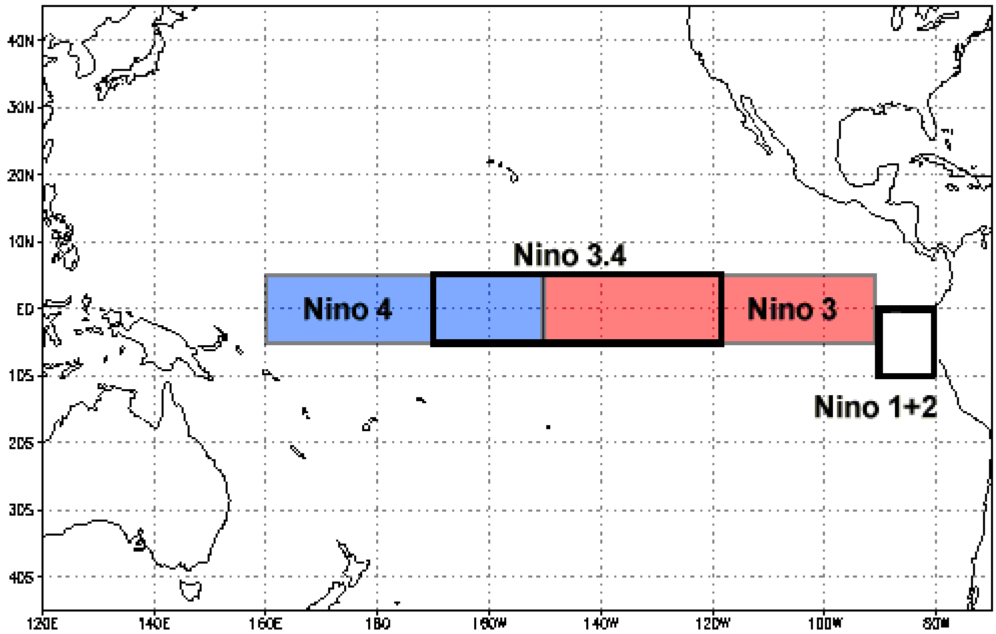


Figure 2.3: Area used for Niño indices.

(Source: https://www.ncdc.noaa.gov/teleconnections/enso/indicators/sst.php)

2.6 Indian Ocean Dipole

The name of the Indian Ocean Dipole or IOD was coined by Prof. Yamagata, Dr. Saji and other researchers of the Climate Variations Research Program (CVRP) of Frontier Research Center for Global Change (FRCGC) to represent the zonal dipole structure of the various coupled ocean-atmosphere parameters, such as SST, Outgoing Longwave

Radiation (OLR) and Sea Surface Height (SSH) anomalies. It is similar to ENSO, the change in temperature gradients across the Indian Ocean results in changes in the preferred regions of rising and descending moisture and air (Behera et al., 1999). Saji et al. (1999) studied the internal modes of variability of the Indian Ocean using EOF analysis on SSTA in the tropical Indian Ocean basin. It found that the first mode is monopole, known as an Indian Ocean Basin (IOB) mode which accounted for 30% of the total variance and had a high correlation with the Niño 3 index. The dipole is apparent for the second mode which accounted for 12% of the total variance. So, this pattern is called the IOD. Furthermore, many researches studied some relation between IOD and ENSO events but the conclusion is still controversial. Some studies found that 30% of IOD event occurred with ENSO (Rao at al., 2002; Yamagata et al., 2004). Some studies have found that the IOD event is an individual event that is unrelated with ENSO (Hastenrath, 2002; Baquero-Bernal et al., 2002; Dommenget and Latif, 2002; Ashok et al., 2003; Tozuka et al., 2006; Behera et al., 2006).

It is normally characterized by negative SST anomalies in the southeastern equatorial Indian Ocean (10°S-0°S, 90°E-110°E) and positive SST in the western equatorial Indian Ocean (10°S-10°N, 50°E-70°E). These conditions are indicated by the Dipole Mode Index (DMI) (Saji et al., 1999; Behera et al., 2006). A positive phase of IOD is characterised by cooler than normal water in the tropical eastern Indian Ocean and warmer than normal water in the tropical western Indian Ocean as shown in Figure 2.4a. Conversely, a negative phase of IOD is characterised by warmer than normal water in the tropical eastern Indian Ocean and cooler than normal water in the tropical western Indian Ocean as shown in Figure 2.4b which SST anomalies are shaded (red color is for warm anomalies and blue is for cold), white patches indicate increased convective activities and arrows indicate anomalous wind directions during IOD events. The direct impact of IOD events is rainfall variability in the tropical countries which are located around the Indian Ocean including eastern of Africa, South of Asia and the northern of Indochina peninsula (Ashok et al., 2001).

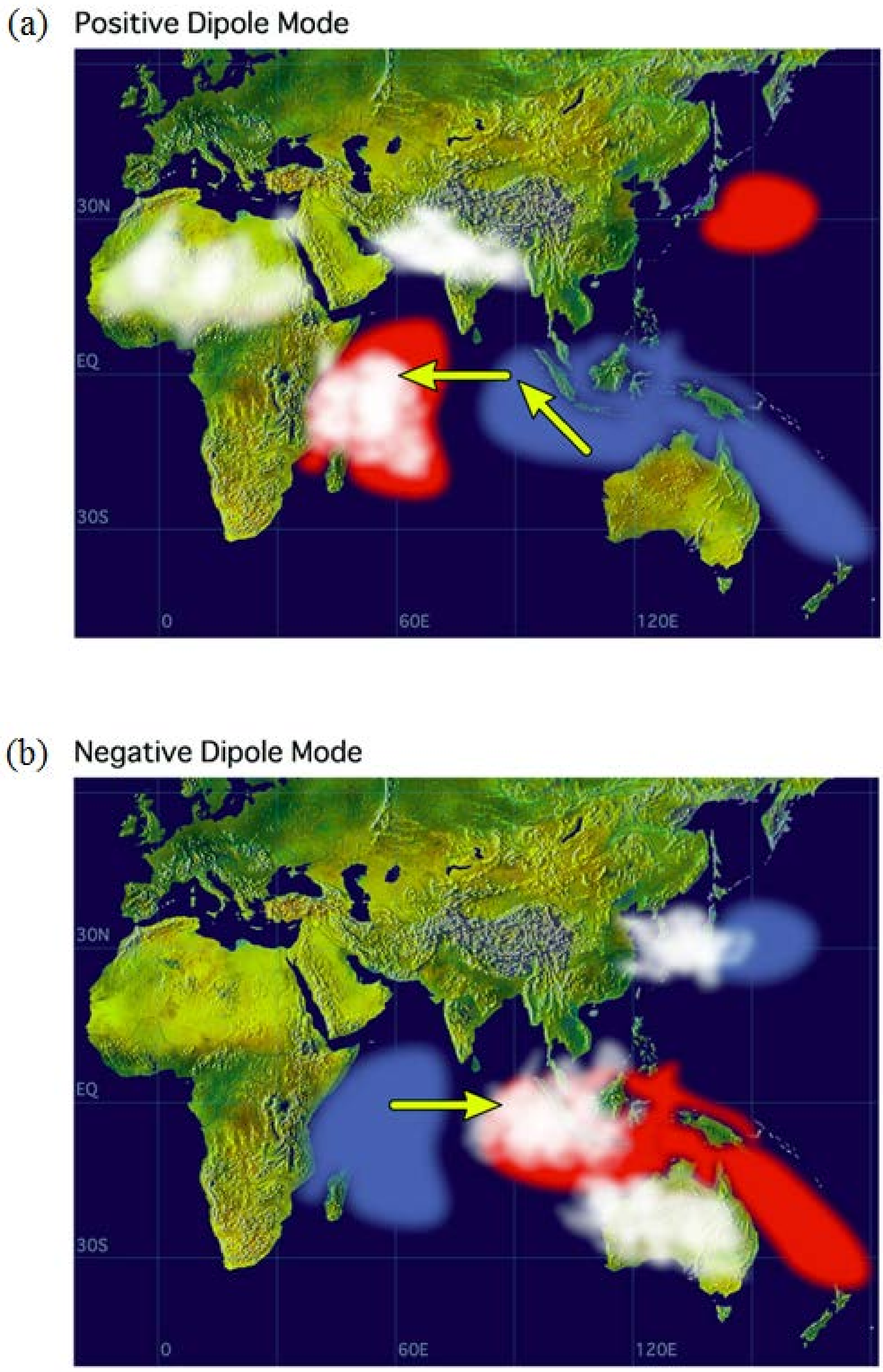


Figure 2.4: Schematic of (a) positive and (b) negative IOD events
(Source: http://www.jamstec.go.jp/frcgc/research/d1/iod/e/iod/about_iod.html).

CHAPTER 3

ANALYSIS OF RAINFALL VARIABILITY OVER THAILAND

In this study, the rainfall data was selected and the quality of rainfall data was controlled before analysis the variability by using the Empirical Orthogonal Function (EOF) method.

3.1 Data and Quality Control

The monthly gridded rainfall of the Global Precipitation Climatology Centre (GPCC) data was set with 0.5°×0.5° horizontal resolution (Schneider et al., 2011a) for the period of 1971 to 2010. The rainfall data were extracted for Thailand (5.5°N - 21°N, 97.5°E - 106°E) to analyze the dominant spatio-temporal modes by the EOF method.

3.1.1 GPCC Data

GPCC was established at Deutcher Wetterdienst (DWD, German Weather Service) in 1989 by invitation from the World Meteorology Organization (WMO) as a German contribution to the World Climate Research Programme (WCRP). It is the global analysis of monthly precipitation for the earth's land surface based on in situ rain gauge data. Since its start, the centre is the in situ component of the WCRP Global Precipitation Climatology Project (GPCP) (WMO, 1990). In 1994, the long-term operation of the GPCC has been requested by WMO in order to contribute to the climate monitoring activities of the Global Climate Observing System (GCOS). Since 1999, GPCC is one of the two global GCOS Surface Network Monitoring Centers (GSNMC) with special emphasis on precipitation. In mid December 2006, GPCC started its newest function as the WMO Commission for Basic Systems (CBS) Lead Center for GCOS data for Europe. The aim of GPCC is to serve user requirements regarding accuracy of the gridded precipitation analyses and timeliness of the product availability. The WCRP Global Energy and Water Experiment (GEWEX) for instance requests high spatial resolution and accuracy for the last two decades, while the priority of GCOS and IPCC is focused on long-term homogeneous time-series. Timeliness of products is ensured by cut-off dates for data processing and analysis. All GPCC analysis products result from the same quasi-operational data management and analysis system. However, depending on the required timeliness they differ with regard to the number of stations included and the level of data quality control being performed.

GPCC's database contains precipitation data on a mainly monthly basis from a variety of sources. The data distributed by the national meteorological and/or hydrological services (NMHSs) via the WMO Global Telecommunication System (GTS) to fulfil the needs of near real-time weather analysis and prediction and climate monitoring are available near real-time such as synoptic weather reports (SYNOP), from which monthly precipitation totals can be accumulated (Schneider et al, 1992) and monthly climate reports (CLIMAT). In addition to the GTS data, GPCC has required precipitation data from NMHSs from 190 countries which the spatial distribution of stations is shown as Figure 3.1 that meanwhile form the backbone of its data base but become available only with a larger delay (non real-time data). The near real-time products of the GPCC based on the GTS data are the First Guess Product (Ziese et al., 2011) which is based on synoptic weather reports received at DWD interpolated precipitation anomalies from more than 6,000 stations whereby an automatic-only QC is applied, and the Monitoring Product Schneider et al., 2011a) which is based on SYNOP and monthly CLIMAT reports received via GTS from 7,000-8,000 stations (after automatic and manual quality control). The non real-time products is Full Data Reanalysis which is based on all stations in GPCC data base supplying data for the individual month, near real-time, and non real-time. The full Data Reanalysis Product Version 6 covering the period from 1901 to 2010. The grid resolutions are 0.5°×0.5°, 1.0°×1.0° and 2.5°×2.5° geographical longitude by latitude (Schneider et al., 2011a, b, c). The data coverage per month varies from 10,800 stations at the beginning to more than 47,000 stations in 1986/1987 this jump is almost leveled, and the aforementioned increase during the past decade also gives an indication of how long it takes for worldwide collected rain gauge data to arrive at GPCC and to pass the rigorous QC before entering the GPCC data base, which is causing the decrease to 37,500 stations in 2000 and less thereafter as shown in Figure 3.2. It is being updated at irregular times subsequent to significant database enlargements and improvements.

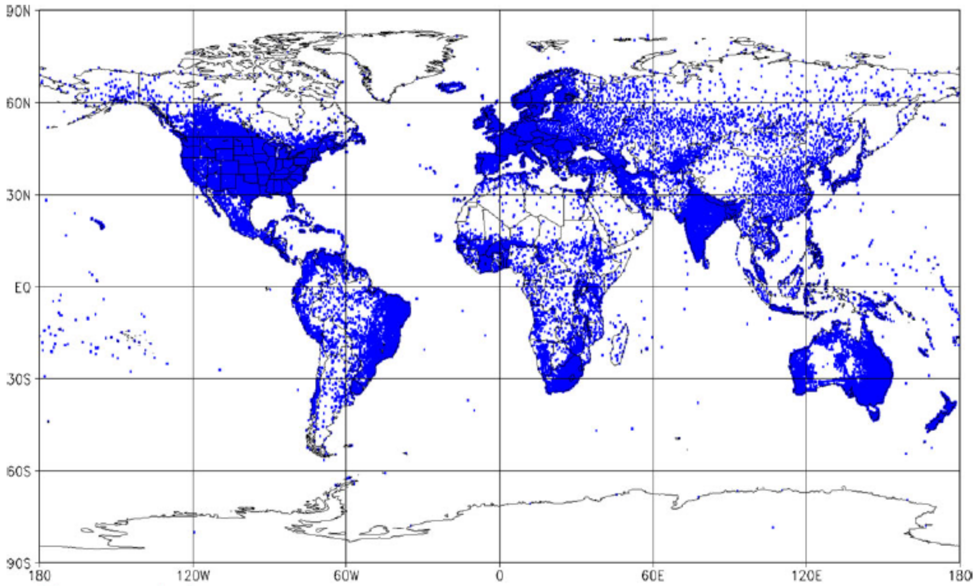


Figure 3.1: Spatial distribution of stations with normal climatology precipitation (Schneider et al., 2013)

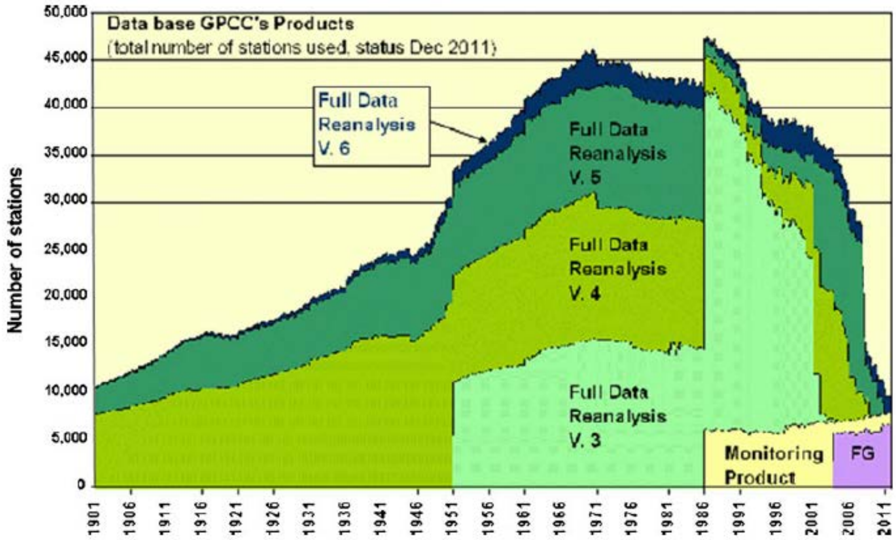


Figure 3.2: Total number of stations used for the GPCC products (near real-time First-Guess Product FG, Monitoring Product; non-real-time Full Data Reanalysis Product (Versions 3 to 6)) (Schneider et al., 2013)

3.1.2 Quality Control

The gridded products are generated by an operational analysis system with components for (a) integration of data from different sources, (b) quality-control, and (c) calculation of area-averaged precipitation on the grid cells.

The processing of observational precipitation data sets at the GPCC indicates that almost any larger input data set contains more or less all kinds of errors in the row station data. Raw data itself, as well as station meta information, can be affected by typing or coding errors and other modifications occurring on the way from the measurement at the station to the data archive. Therefore, a through quality control (QC) is necessary to detect and correct/eliminate such errors which otherwise would have a significant impact on the analysis results.

Toward the large variability of precipitation and the skewness of its frequency distribution, a fully automatic quality control would eliminate all data being classified as outliers including real extremes. However, these are very important to describe the variability of precipitation.

Therefore, QC processing at GPCC is semi-automatic in the way that the data classified as questionable by the automatic QC procedures undergo additional visual checks. The QC system of successive automatic and visual checks has been optimized with respect to the features of the different data sources and the specific meta information being available. Figure 3.3 shows a simplified scheme for the main steps of processing, quality control, archival, and analysis of precipitation data at the GPCC and distribution of its gridded precipitation products for the near real-time as well as non-real-time data.

3.1.2.1. Station identification and quality control of station meta information

The data sets received at GPCC are first checked for readability and then reformatted as given in Figure 3.3. To avoid a spatial misallocation of climatic data in the analysis, for the national/regional data sets supplied to the GPCC, the station locations are displayed by a climate data visualization software, and it is checked if all stations are located within the boundaries of the country. For stations located outside of the boundary, the geographical coordinates are checked with geographical information available via the Internet through geographical atlases or regional maps.

Subsequently, the uniform information data sets are loaded into GPCC’s relational data base management system, whereby the station meta data received from the different sources are checked against the meta-information archived in GPCC’s data bank. If the station meta information in the data set is identical with that of a station in the data base, the data are assigned to the station. If no similar station is existing in the data base, a new station is created therein. In case of discrepancies in station meta information between data set and GPCC's data base, the data supplier is contacted, if possible in a timely manner. Otherwise, the geographical coordinates of the station are checked with other sources of geographical information such as Google-Earth, geographical atlases, or regional maps. This rechecking of geographical information during each loading process is resulting in a continuous improvement of GPCC’s station data base and has led to a very high degree of reliability of its station meta information.

Observed discrepancies can be attributed in part to different spellings of station names, errors in the geographical coordinates or elevations. With regard to the geographical coordinates, typical errors on the order of sometimes 1°, 2°, or even up to 10° latitude or longitude are detected in many of the input data sets. In the elevator information, there are sometimes errors in the conversion of meters and feet, zero instead of missing elevator.

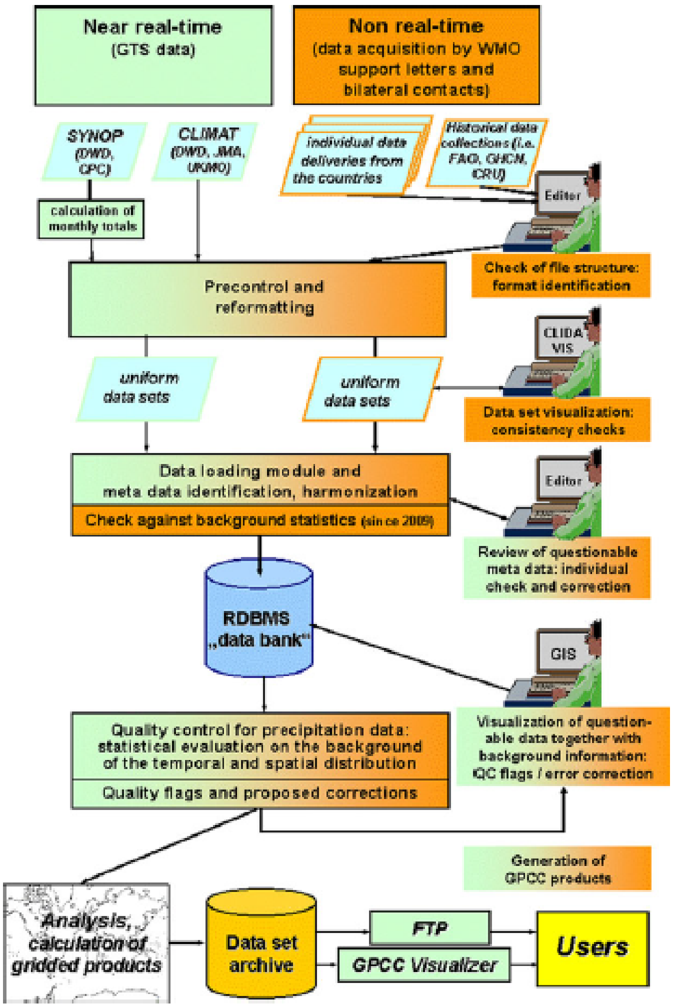


Figure 3.3: Simplified scheme for the main steps of processing, quality control, archival, and analysis of precipitation data at the GPCC and distribution of the gridded precipitation products

3.1.2.2. Quality control of the monthly precipitation data

In order to avoid mismatched or overall erroneous data sets into the data bank, all national and GTS precipitation data sets have already been pre-controlled separately using different techniques fitting the respective data sources as shown in Figure 3.3. Storing the data from the different data sources in parallel in the data bank together with the quality flags indicating the results of data processing is helpful in the QC processing and enables detection of errors by cross-checks of the data from the different sources.

In addition to the previously described QC for GPCC's Monitoring Products, the full data base has repeatedly been checked statistically for outliers over the last years for each new release of the gridded precipitation climatology and the Full Data Reanalysis. The statistical check of outliers: The course of the repeated QC processing for the releases V.4, V.5 and V.6 of the Full Data Reanalysis, the time series of overall about 10,000 stations had been checked visually since for each case of a station with suspicious data, generally two to four and sometimes even more neighbouring stations have been checked foe spatial consistency that can confirm as correct. The missing values which instead of "0" are the one of the biggest problems with the raw data. In the QC processing, they performed a systematic check for erroneous "0" values that revealed and eliminated automatically, after through pre-checks for data subsets. In the case of corrections, the original data are kept in GPCC's data base and the corrections are archived additionally as a higher quality level.

Visual check of spatial consistency in the following steps:

- Misplaced stations caused by erroneous geographical information
- Individual errors causing an erroneous climatological normal or maximum/ minimum
- Quasi-systematic errors such as conversion of units inch, feet, millimeters
- In some specific cases, data for some months/years have been found to be shifted by 1, 2, or more months, or even a whole year in some cases.

Check of temporal homogeneity: tested by applying a moving t-test which checks the homogeneity over time and allows deeper analysis of the data set in this regard, revealing significant inhomogeneities ($t \geq 6$) in some regions. On the basis with the QC of the station data from the different sources and statistical evaluation, the GPCC has set up a priority scheme according to which data are being selected for its analysis.

Since the GPCC data is through the process of testing and checking the quality of data, then, this chapter checked the missing values on the gridded points. Quality Control (QC) was undertaken prior to the data analysis. The QC procedure for data is checking the missing values of the yearly data which are 40 years. The daily rainfall data is from GPCC data over Thailand (5.5°N - 21°N, 97.5°E - 106°E) with resolution of 0.5°×0.5° latitude and longitude for the period from 1971 to 2010. The results of checking missing values of data still contain a few missing values as Figure 3.4.

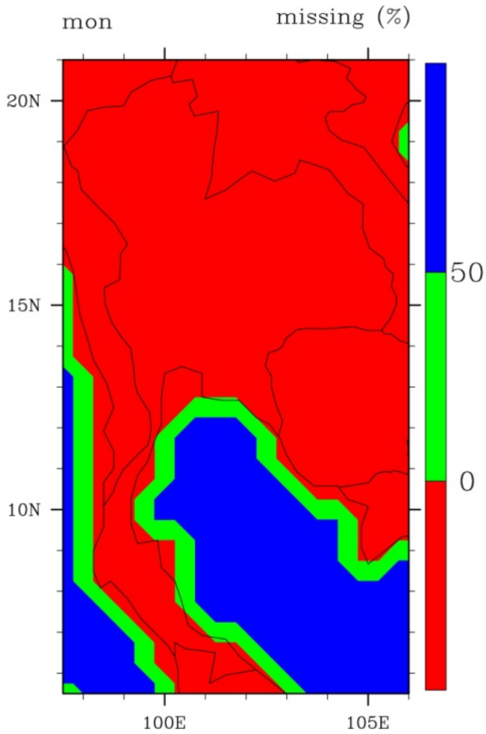


Figure 3.4: Checking the missing values of the GPCC data

3.2 Climatology Data

This study applied the monthly rainfall data of the full data reanalysis product version 6 from The Global Precipitation Climatology Centre (GPCC), in a horizontal resolution of $0.5° \times 0.5°$ latitude and longitude, for 1971 to 2010 (Schneider et al., 2011). The climatology is the long-term average of rainfall data over Thailand (5.5°N - 21°N, 97.5°E - 106°E) as shown in Figure 3.5.

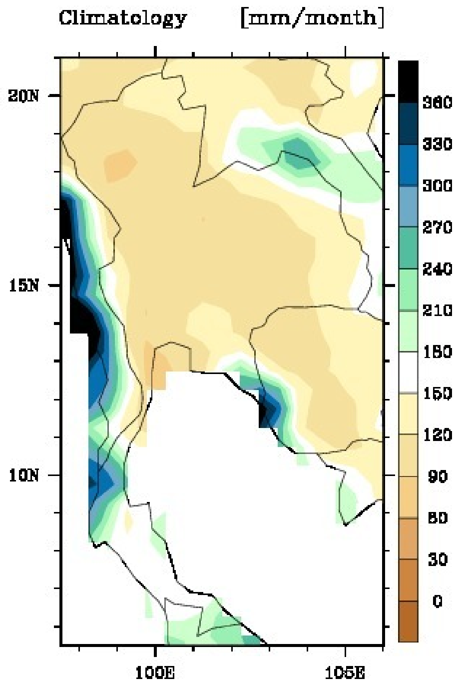
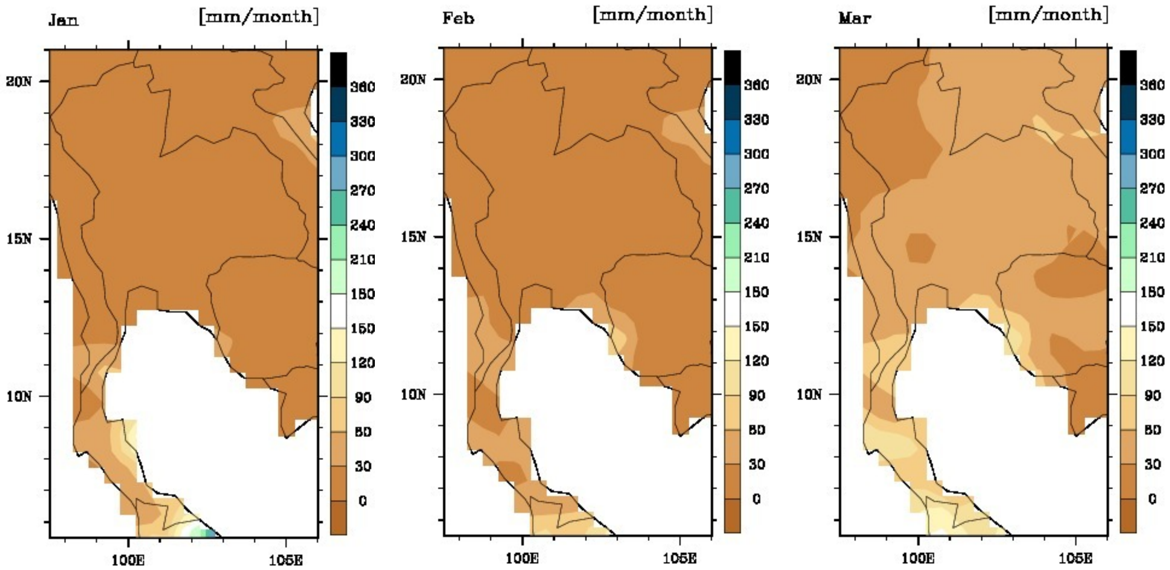


Figure 3.5: The climatology of rainfall data from GPCC data during 1971 to 2010

The average of total monthly rainfall data (January to December) from GPCC data over Thailand for the period from 1971 to 2010 as shown in Figure 3.6.



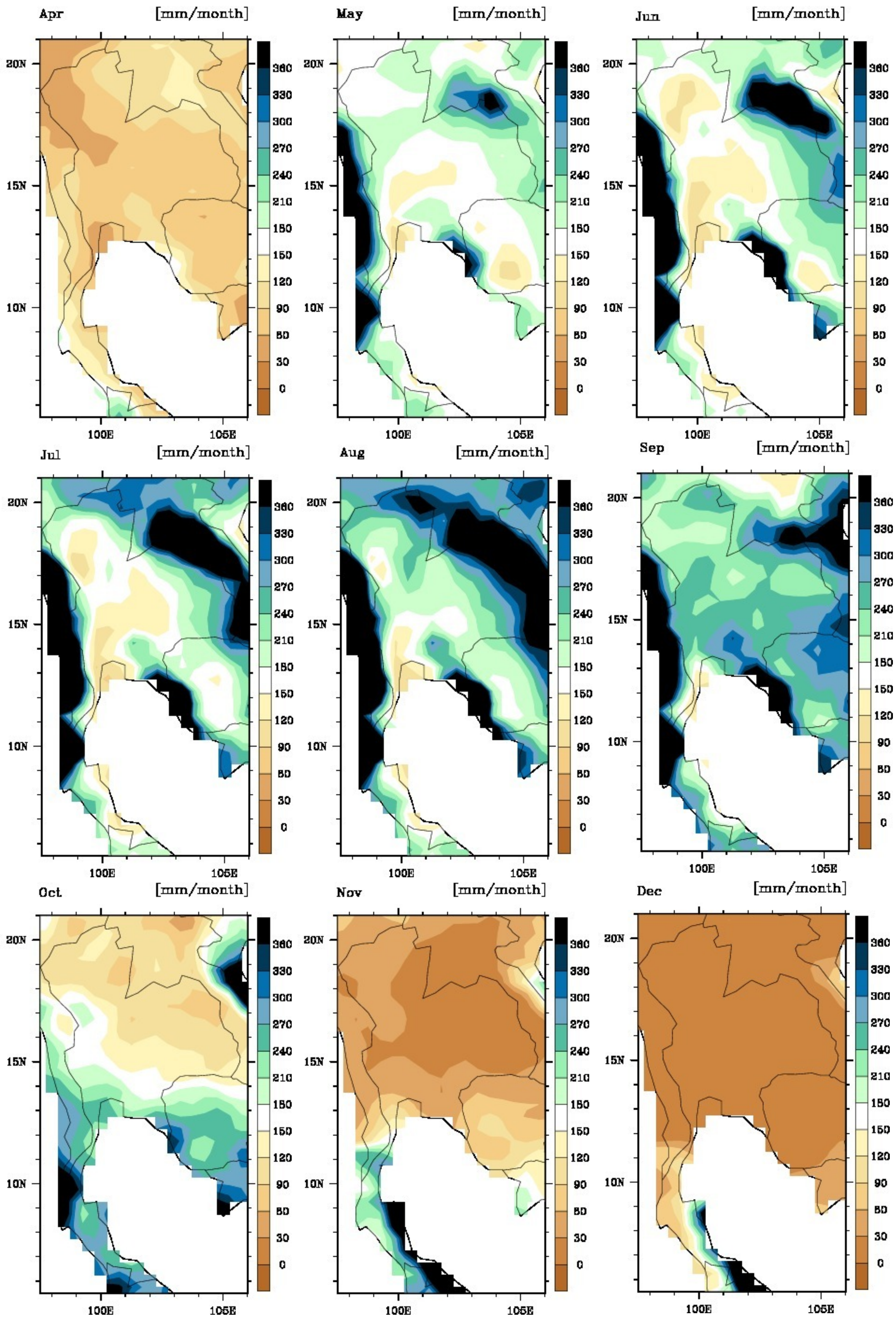


Figure 3.6: The averaged of total monthly rainfall from GPCC data during 1971 to 2010

After having studied the annual variation of rainfall over Thailand between of GPCC data averaged from 1971 to 2010, it was found that in January to April had less rainfall. Since May is start of the rainy season, there is high rainfall until October. November is the start of the dry winter season, so there is low rainfall, as shown in Figure 3.7.

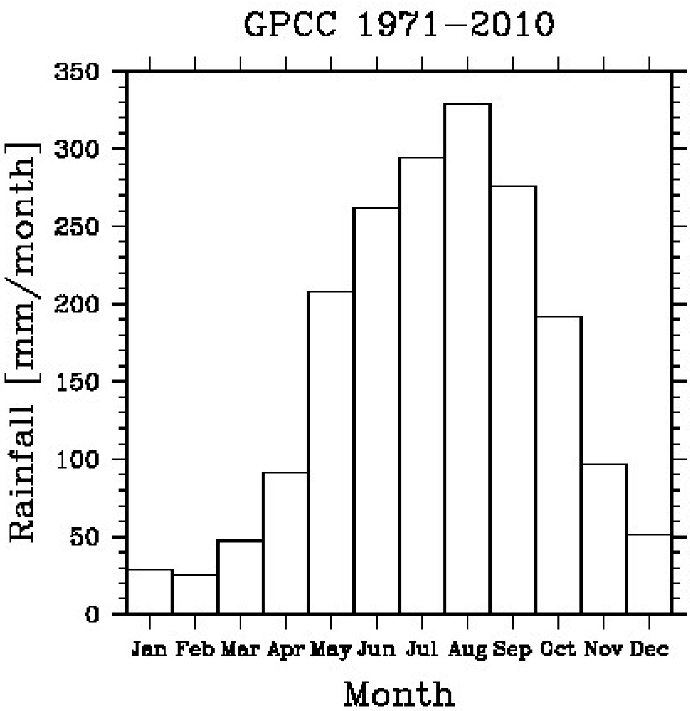


Figure 3.7: The annual variation of rainfall over Thailand

3.3 Rainfall Variability over Thailand

This section discusses the variability rainfall data and analysis of the spatial and temporal pattern by EOF. The EOF analysis is based on a linear transformation to extract information from the large data sets by decomposing to the orthogonal basis function while retaining as much as possible of the variations present in the data sets.

The GPCC data was analyzed by the EOF method for explaining the variance of rainfall over Thailand during the period of 1971 to 2010. Figure 3.8 shows the proportional percentage of the variability of rainfall data. The first mode represented the variability of the rainfall data over Thailand about 21.6% of the total variance as shown in Figure 3.9. The second mode represented the variability of the rainfall over Thailand about 9.1% of the total variance as shown in Figure 3.8.

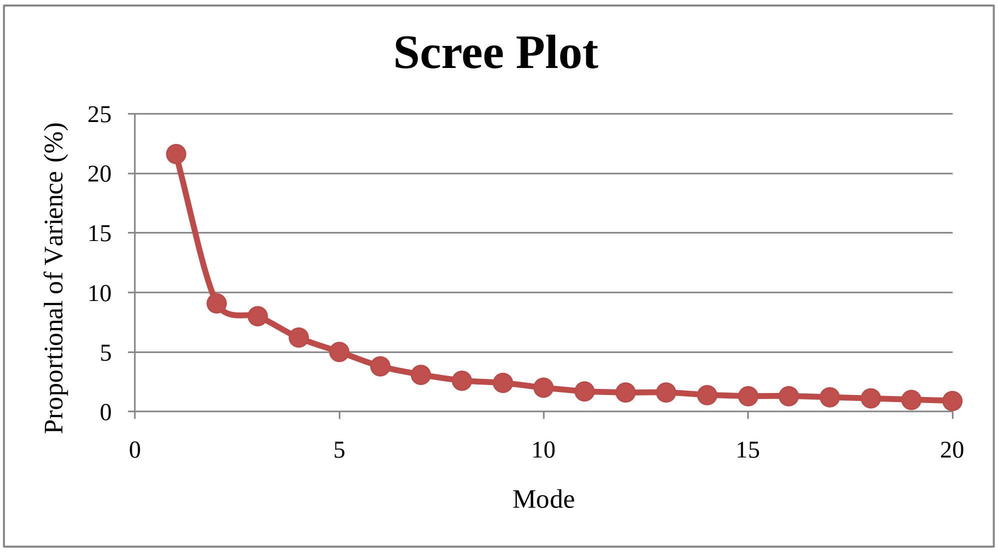


Figure 3.8: The proportional to the variability of each EOF mode of GPCC data

The result analysis of the eigenvectors and principal component time series in the first mode show that the relation or the variance of the eigenvectors is positive over most of Thailand as shown in Figure 3.9(a). Figure 3.9(b) shows the values of the standardized principal component time series of this mode. Since the values in the spatial pattern are strongly positive over most of Thailand, it follows that when the principal component (PC) in the time series is positive most of Thailand has higher than normal rainfall rainfall, and when the PC is negative most of Thailand has lower than normal rainfall. An exception to this result occurs in southern Thailand at latitudes below 9°N where the spatial pattern of the most dominant mode is only weakly positive.

When studying the relation between the eigenvector and the principal components, it was found that there was high rainfall (standardized PC greater than 2.0) in June 1975, August 1978, October 1983, June 1985, July 1994, September 1996, July 1997, April 1999, May 2001, July 2006 and October 2010. On the contrary, there was low rainfall (standardized PC less than −2.0) in September 1971, September 1974, August 1976, June 1977, October 1979, July 1983, July 1984, May 1987, August 1988, May 1992, October 1994, July 1998, August 1998 and October 2004 (see small circles in Figure 3.9(b)).

The result analysis of the eigenvectors and the principal component time series in the second mode shows that the relation or the variance of the eigenvectors is positive over the centre, eastern and southern regions of Thailand but it is negative over the northern and northeastern regions, as shown in Figure 3.10(a). Figure 3.10(b) shows the values of the standardized principal component time series, which have positive and negative values.

(a) (b)

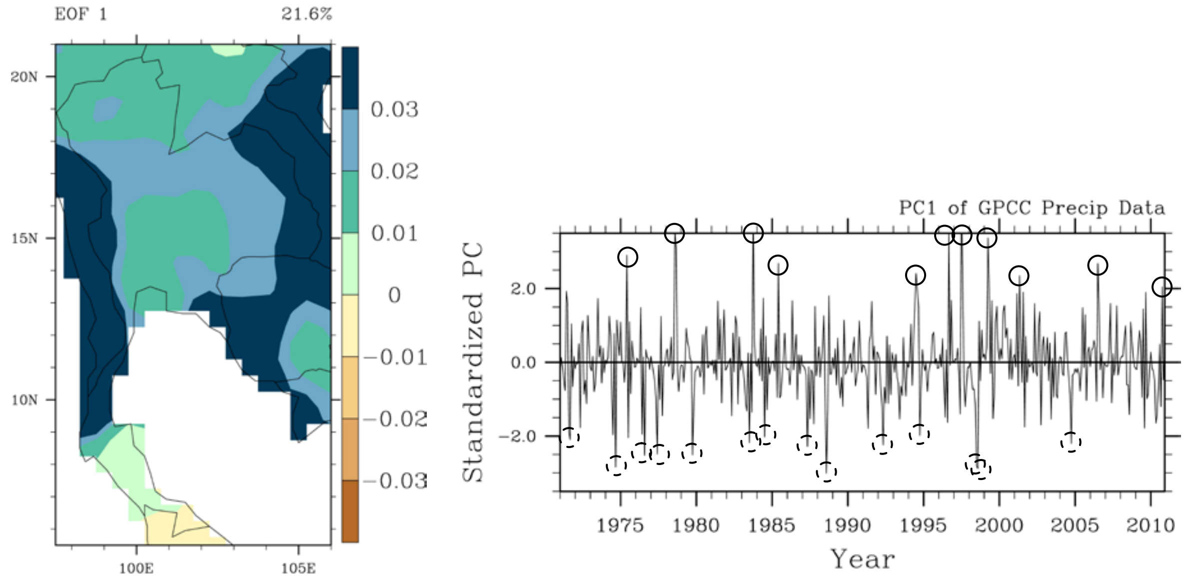


Figure 3.9: The EOF results of (a) the eigenvector and (b) the principal component time series of the first mode of GPCC data over Thailand over 1971 to 2010.

(a) (b)

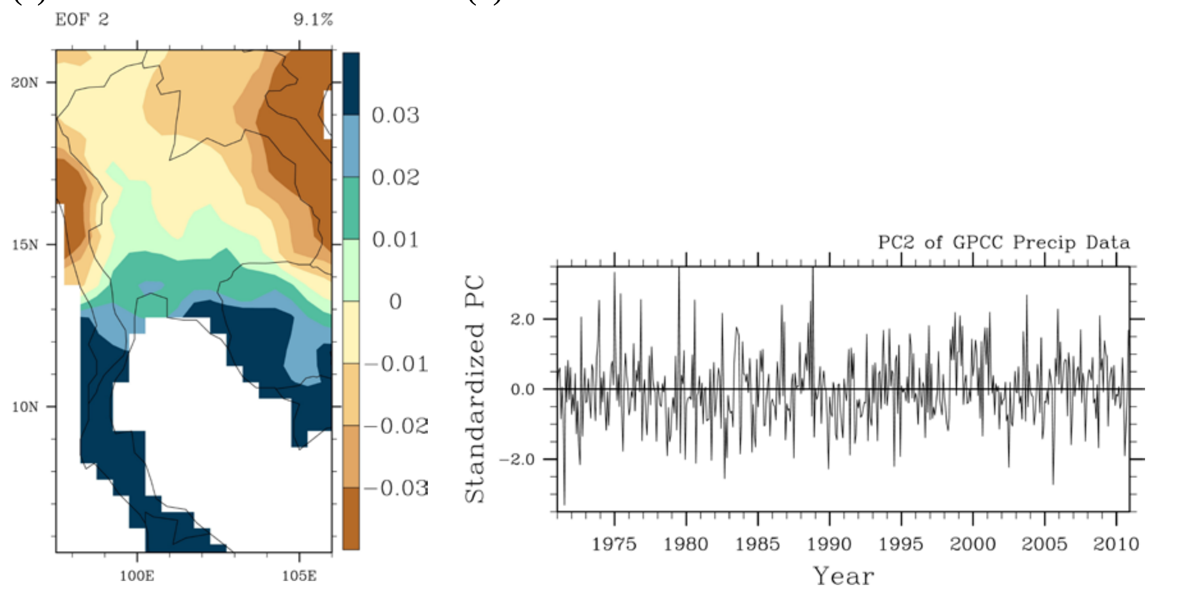


Figure 3.10: The EOF results of (a) the eigenvector and (b) the principal component time series of the second mode of GPCC data over Thailand from 1971 to 2010.

To ensure the rainfall analysis by using the EOF method, Figure 3.11(a) shows there is positive composite of rainfall over Thailand, and Figure 3.11(b) shows negative composite of rainfall over Thailand. It was found that there was greater than average rainfall almost everywhere in figure 3.11(a) and less than average rainfall almost everywhere in figure 3.11(b). Figure 3.11(c) shows the difference between the positive and negative composite of the rainfall over Thailand, with the contour interval 20 mm. The shaded region, covering most of the study area shows where the difference is significant at the 95% confidence level by Student's *t*-test.

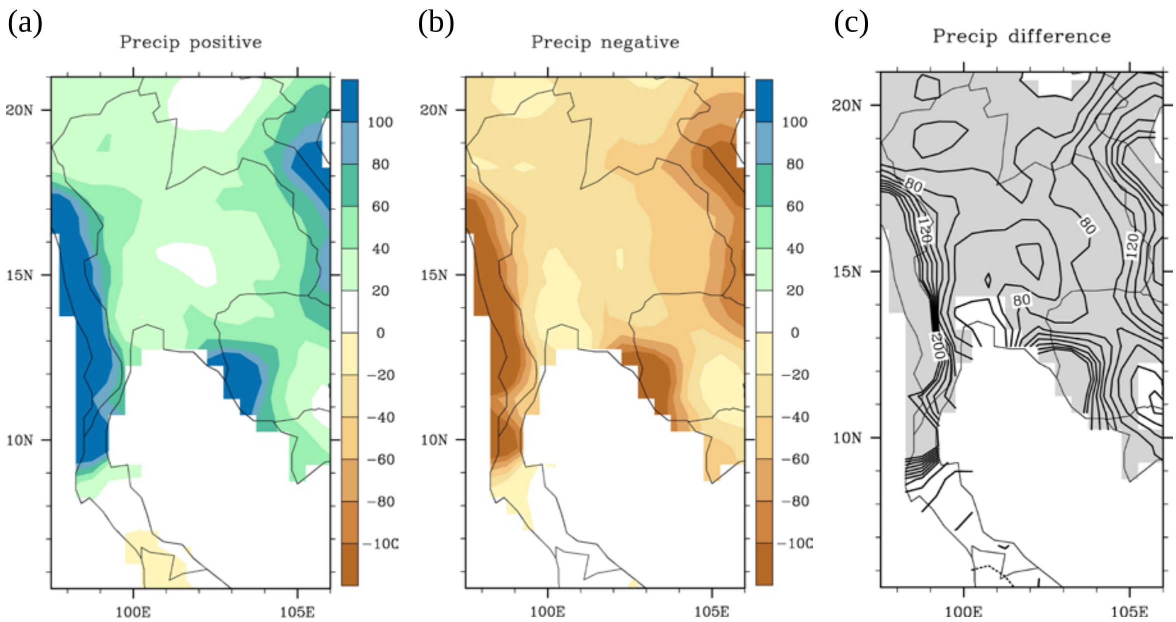


Figure 3.11: The composite of rainfall anomalies over Thailand between (a) positive values and (b) negative values of the first EOF mode, and (c) difference value from 1971 to 2010. The contour interval is 20 mm and the shading shows where the difference is significant at the 95% level as determined by Student's *t* test.

CHAPTER 4
RELATIONSHIP AND ASSOCIATION BETWEEN RAINFALL VARIABILITY AND CLIMATE MODES

Teleconnection is a natural phenomenon that occurs repeatedly in a period and has stability of the variance in several periods, which may be from one day to many centuries. The main causes are from the interaction and dynamics of the climate system. The teleconnection is a phenomenon on the planetary-scale which has the variability and impact across the oceans and continents. The format and style of teleconnection reflect the change in the large area of the circulation of atmosphere which influences for temperature, rainfall, frequenty and strength of storms (Wang et al, 2000; Diaz et al, 2001).

4.1 Teleconnection to ENSO

There has been much research analysis of the atmospheric responses to ENSO (Feng et al., 2010; Hoerling & Kumar, 2002; Wu et al., 2008). Teleconnections between the tropical Pacific and the remainder of the globe have been found in numerous observational analyses (Bjerknes, 1969; Mistra, 2004). While the rainfall variability over Thailand may be linked to global climate, it is important to understand the variations. Because Thailand lies between the equatorial Indo-Pacific basins, the rainfall variability over Thailand is linked to ENSO. This section reports our study of this relationship.

4.1.1 Relationship between Rainfall Variability over Thailand and ENSO

The ENSO event illustrated Sea Surface Temperature (SST) variability in the equatorial Pacific Ocean. There are many indices used to indicate the ENSO events. Trenberth (2001) suggested the well known indices derived from SST behavior such as Niño 1+2 (0°-10°S, 90°W-80°W), Niño 3 (5°N-5°S, 150°W-90°W), Niño 4 (5°N-5°S, 160°E-150°W), and Niño 3.4 (5°N-5°S, 170°W-120°W), that represent the SST anomalies averages over corresponding areas as shown in Figure 4.1. 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). The El Niño event occurs when SST anomalies exceed 0.4°C and the La Nina event occurs when SST anomalies below -0.4°C (Trenberth, 1997). It used normalized monthly rainfall anomalies (the anomalous monthly rainfall divided by the corresponding monthly standard deviations).

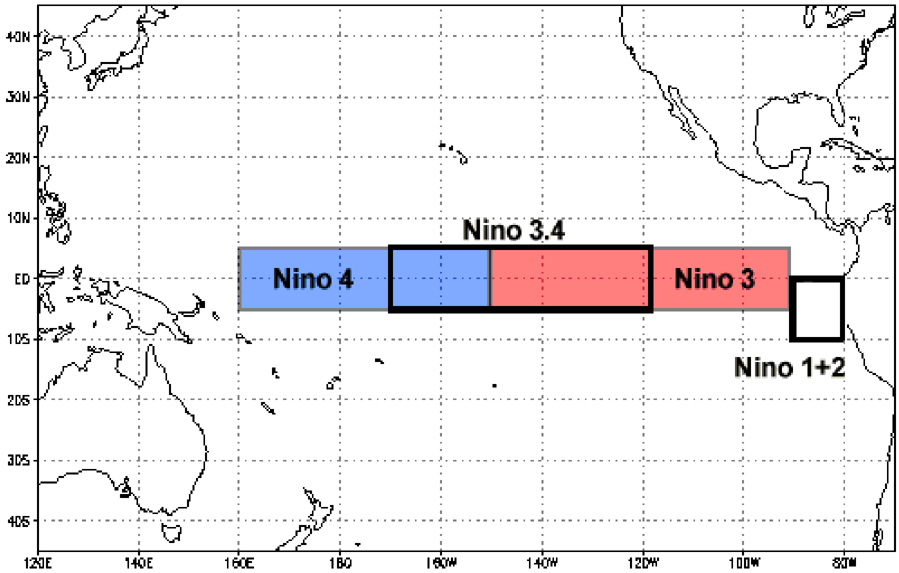


Figure 4.1: Niño 3.4 regions (5°N - 5°S, 170°W - 120°W)
(https://www.ncdc.noaa.gov/teleconnections/enso/indicators/sst.php)

This research studied the correlation between Nino 3.4 SST index and the principal component of rainfall variability over Thailand. Figure 4.2 shows the time series plots of these quantities. The correlation coefficient between them is -0.109876. It is negative, because a high SST anomaly gives a low rainfall anomaly, and a low SST anomaly gives a high rainfall anomaly. However, they are rarely independent. To study the autocorrelation analysis as shown in Table 4.1, the adjustment of sample size (N_{eff}) is 474. It found that the sample size decreased, then the correlation between Niño 3.4 Index and principal component increased.

Trenberth (1997) studied El Niño events during the 5-month running means of sea surface temperature anomalies in the Niño 3.4 region. Then, it used normalized monthly rainfall anomalies (the anomalous monthly rainfall divided by the corresponding monthly standard deviations) with the low-pass filter (Duchon, 1979).

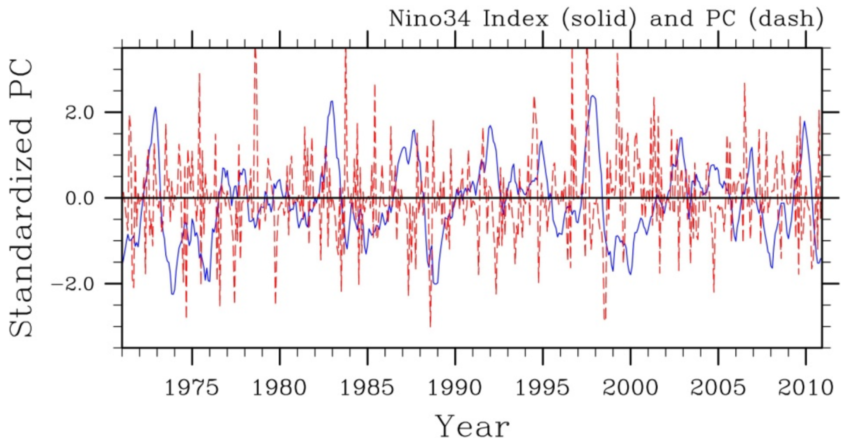


Figure 4.2: Time series plots of Niño region 3.4 SST index (solid line) and principal component of the first mode (dash line).

Table 4.1: The Autocorrelation of Niño 3.4 SST Index and the principal component.

Lag time	Niño 3.4 Index	Principal component
0	1.000000000	1.000000000
1	0.955358600	-0.027282480
2	0.868833700	0.092103740
3	0.765195900	-0.087159510
4	0.649258600	-0.011118800
5	0.526664000	0.047957360
6	0.403817500	-0.005514535
7	0.284384200	0.023777000
8	0.174237000	0.015913700
9	0.077395130	-0.101142400
10	-0.001963018	0.048489760
11	-0.063067830	-0.110442100

Figure 4.3 shows the correlation of the smoothed anomalies between Niño 3.4 SST index (blue line) and the principal component (red line) is -0.1651. It is a negative relationship because Niño 3.4 index indicates the Sea Surface Temperature (SST) anomalies and the principal component indicates the rainfall anomalies. Then, the high SST has low rainfall whereas the low SST has high rainfall. However, they are rarely independent. To study the autocorrelation analysis as shown in Table 4.2, the adjustment of

sample size (N_{eff}) is 241. It was found that the sample size decreased, then the correlation between Niño 3.4 Index and principal component increased.

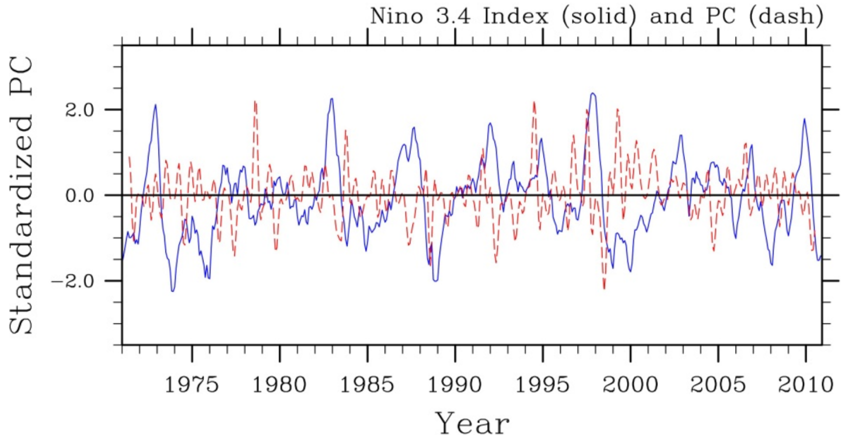


Figure 4.3: Time series plots of smoothed anomalies with the low-pass filter of Niño region 3.4 SST index (blue line) and principal component (red line), using the 5-term filter.

Table 4.2: The Autocorrelation of smoothed anomalies of Niño 3.4 SST Index and the principal component.

Lag time	Niño 3.4 Index	Principal component
0	1.000000000	1.000000000
1	0.971341900	0.787497500
2	0.894113000	0.330450200
3	0.786233400	-0.028466160
4	0.664772800	-0.119975800
5	0.539792900	-0.037906420
6	0.415972800	0.040788610
7	0.296881500	0.034320560
8	0.186797900	-0.028470230
9	0.089961910	-0.083024840
10	0.009384785	-0.089387330
11	-0.053768870	-0.053040520

The correlation of the lag time of smoothed anomalies between rainfall anomalies and Niño 3.4 SST index is shown in Table 4.3. The zero value of lag time represented the rainfall anomalies at Niño 3.4 SST index. The positive value of lag time represented Niño 3.4 SST index leads the rainfall anomalies by the number of months. Example from Table 4.3, lag 4 is the highest correlation about -0.3106 which means Niño 3.4 SST index leads the rainfall anomalies by 4 months.

Table 4.3: The correlation of lag time between Niño 3.4 SST index and the rainfall anomalies.

Lag time	Correlation
0	-0.1651
1	-0.2037
2	-0.2436
3	-0.2843
4	-0.3106
5	-0.3060
6	-0.2679
7	-0.2102
8	-0.1535
9	-0.1126
10	-0.0888
11	-0.0714
12	-0.0474

In this study, an index based on the area averaged SST anomaly in the Niño 3.4 region is employed and refers to it as the Niño 3.4 index as shown in Figure 4.4 (Trenberth, 1997). It is characterized by the warming (El Niño) and cooling (La Niña) sea surface temperature (SST) anomalies in the eastern and central equatorial Pacific (Rasmusson, 1982; Chang et al., 2000).

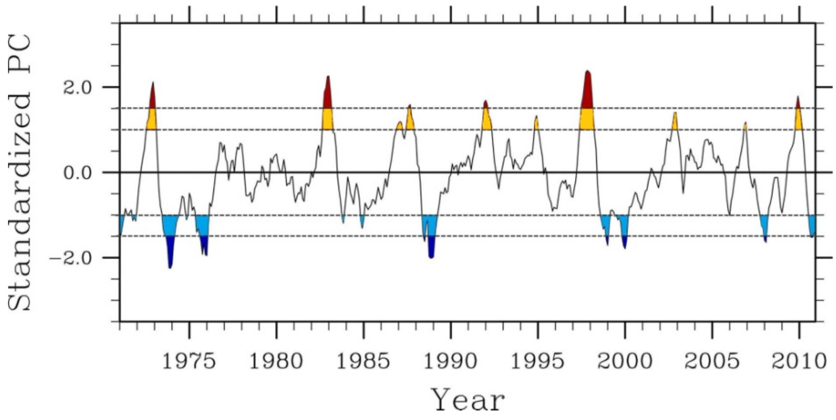


Figure 4.4: Niño 3.4 SST anomaly index for the period 1971-2010.

The Nino 3.4 index can classify ENSO events into the weak and strong intensity, the weak ENSO refers to the Niño 3.4 index larger than or equal to 1.0°C to 1.5°C, and the strong ENSO refers to the Niño 3.4 index larger than or equal to 1.5°C (Bulic and Brankovic, 2007). Then, our analysis is based on composites of fourteen weak La Niña events, six strong La Niña events, twelve weak El Niño events and six strong El Niño events as shown in Table 4.4.

Table 4.4: Years in the period 1971-2010 in the SST categories extending from strong cold (La Niña) to strong warm (El Niño) of ENSO events.

SST category	**No. of year**	**Years**
Strong La Niña	6	1973, 1975, 1988, 1999, 2000, 2008
Weak La Niña	11	1971, 1974, 1975, 1976, 1983, 1984, 1988,1989, 1998, 2007, 2010
Weak El Niña	8	1983, 1991, 1994, 1997, 1998, 2002, 2006, 2010
Strong El Niño	6	1972, 1982, 1987, 1992, 1997, 2009

The composite averages of rainfall anomalies over Thailand for weak-ENSO periods, strong-ENSO periods and their differences are presented in Figure 4.5. There was high rainfall in the cold phase (La Nina) (Limsakul et al., 2007) especially in the southern part of Thailand during the weak and strong periods, as shown in Figures 4.5(a) and 4.5(b). There was low rainfall in the warm phase (El Nino) (Limsakul et al., 2007) especially in the southern part of Thailand during the weak and strong periods, as shown in Figures 4.5(c) and 4.5(d). The differences between the La Nina and El Nino rainfall anomalies during the weak and strong periods are shown in figures 4.5(e) and 4.5(f). These differences were higher during the weak-ENSO periods than during the strong-ENSO periods.

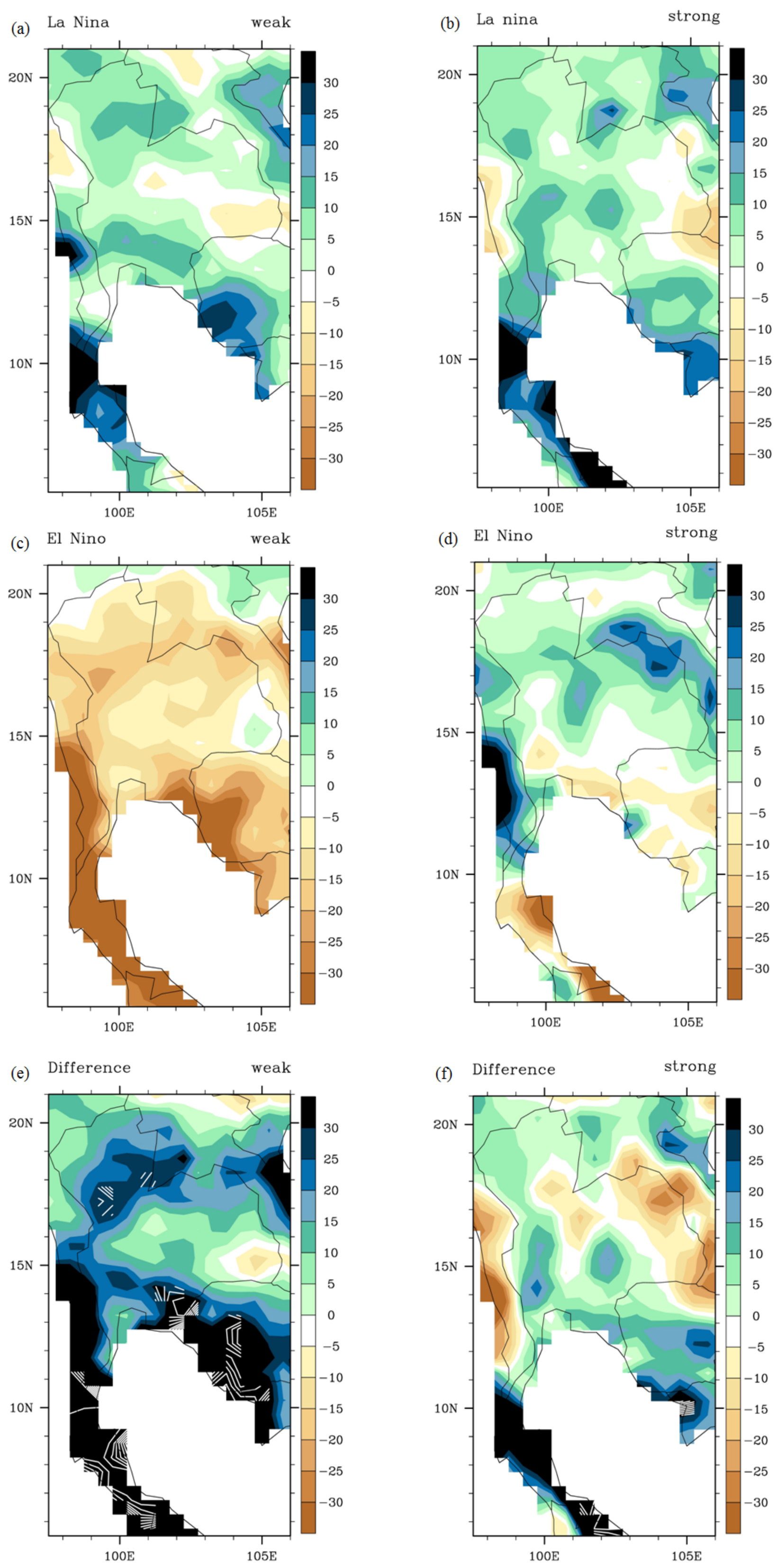


Figure 4.5: The monthly composite of the rainfall anomalies over Thailand for El Niño event and La Nina event and their differences during the weak-ENSO period (a, c and e,

respectively) and the strong-ENSO period (b, d and f, respectively) over the period from 1971 to 2010. The shaded interval is 5 mm and the contour interval indicates 95% significance level determined by Student's *t* test.

4.1.2 The mechanisms associated with ENSO

To understand the relationship between rainfall and ENSO, corresponding composite wind circulation and the sea level pressure maps are constructed. Figure 4.6 shows the composite maps of the differences between the La Nina and El Nino wind anomalies at the 850 hPa level and the differences between the La Nina and El Nino sea level pressure anomalies for the weak and strong ENSO periods. The region to the west of the Pacific Ocean, including Thailand, Burma, India, Laos, Cambodia, Vietnam and the southern part of China had negative differences between the La Nina and El Nino sea level pressure anomalies in both the weak and strong ENSO periods. This indicates lower pressures in Southeast Asia during the La Nina events than during the El Nino events. These figures also show that the northern Pacific high has a positive anomaly difference, so the high is intensified during the La Nina events and weakened during the El Nino events. The wind from the east into Southeast Asia (Wang et al., 2000, Juneng and Tangang, 2005) is then strengthened during the La Nina and weakened during the El Nino events, as shown by the wind anomaly differences in Figure 4.6. The interaction of these easterly winds from the Pacific with the south-westerly flow from the Indian Ocean then produce stronger than normal convergence and rainfall in Southeast Asia during the La Nina events, and reduced rainfall during the El Nino events. Kumar et al. (1999) identified the southeastward shift in Walker circulation anomalies which is similar to these patterns. The ENSO-related SST shifts and the associated shifts in Walker circulation and correlations between Thailand rainfall and Walker circulation all appear to be consistent (Singharattana et al., 2005). Clearly, these shifts will have implications to rainfall variability in terms of their relationship to ENSO.

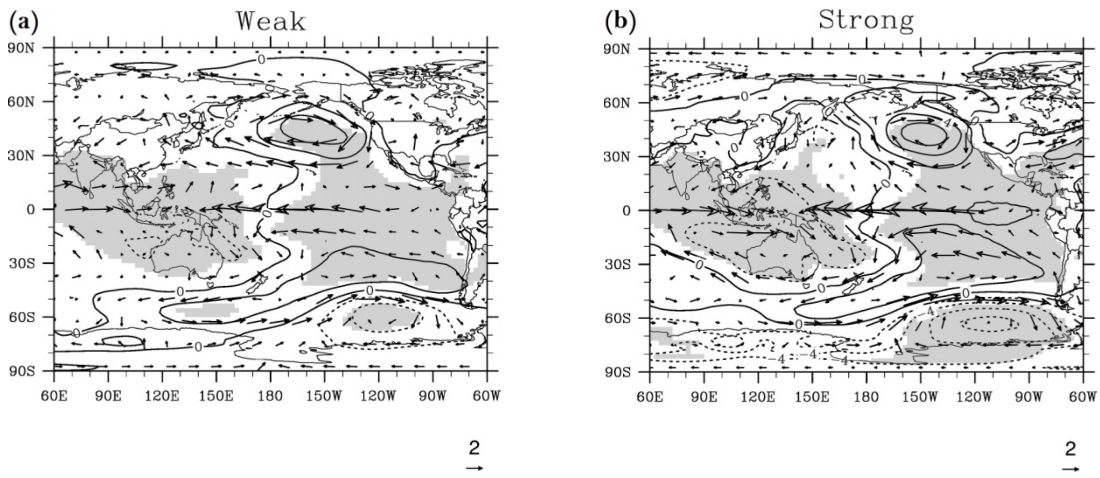


Figure 4.6: Composite maps of the differences between the La Nina and El Nino wind anomalies at the 850 hPa level and the sea level pressure anomalies for the (a) weak-ENSO periods and (b) strong-ENSO periods from 1971 to 2010. The contour interval for the sea level pressure differences is 2 hPa. Continuous contours are zero and positive; dashed contours are negative. The shaded region shows where the sea level pressure anomaly differences are significant at the 95% confidence level by the Student's *t*-test.

4.2 Teleconnection to IOD

There are many studied that analyze the atmospheric responses to IOD (Saji et al, 1999; Webster et al., 1999). The direct impact of IOD events is rainfall variability in the tropical countries which are located around the Indian Ocean including eastern of Africa, South of Asia and the northern of Indochina peninsula (Ashok et al., 2001). While the rainfall variability over Thailand may be linked to global climate, it is important to understand the variations. Then, this research studied the relationship with IOD because Thailand is close to the equatorial Indo-Pacific basins.

Relationship between Rainfall Variability over Thailand and IOD

IOD events can be detected using the Dipole Mode Index (DMI), which is defined as the difference in SST anomalies between the tropical western Indian Ocean (10°S-10°N, 50°E-70°E) and eastern Indian Ocean (10°S-0°S, 90°E-110°E) as shown in Figure 4.7. It has been normalized by standard deviations (0.3°C) (Saji et al., 1999). The correlation between DMI (solid line) and the principal component (dash line) of the first and the second EOF mode showed in Figure 4.8(a) and 4.8(b). It found that the rainfall variability over Thailand has less correlation with DMI about -0.00271 in the first mode and -0.00621 in the second mode. However, they are rarely independent. To study the autocorrelation analysis as shown in Table 4.5, the adjustment of sample size (N_{eff}) is 474. It found that the sample size decreased, then the correlation between DMI and principal component increased.

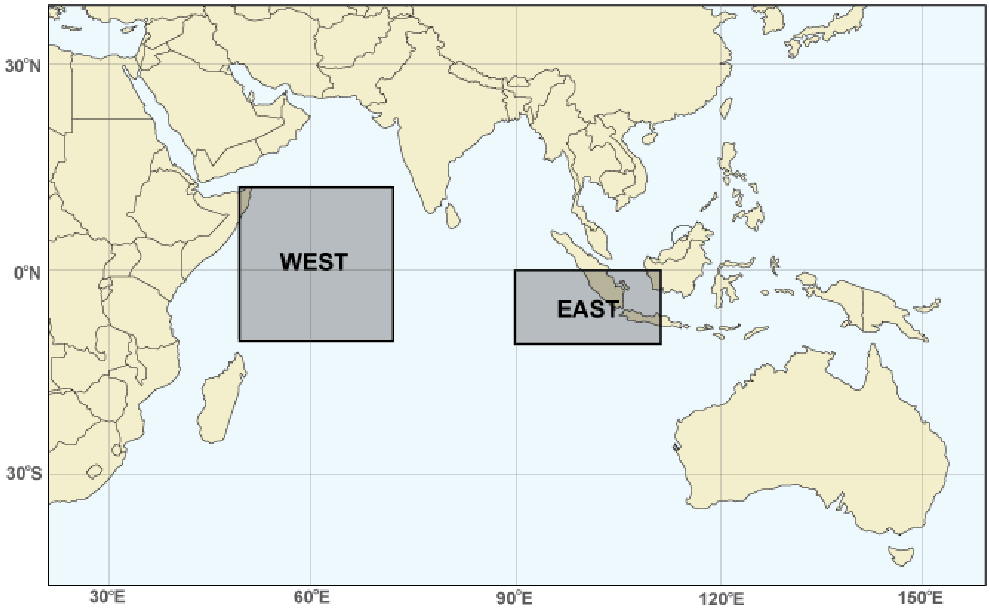


Figure 4.7: The region of SST anomalies between the tropical western Indian Ocean (10°S-10°N, 50°-70°E) and eastern Indian Ocean (10°S-0°, 90°-110°E)
(Source: http://www.bom.gov.au/climate/IOD/about_IOD.shtml)

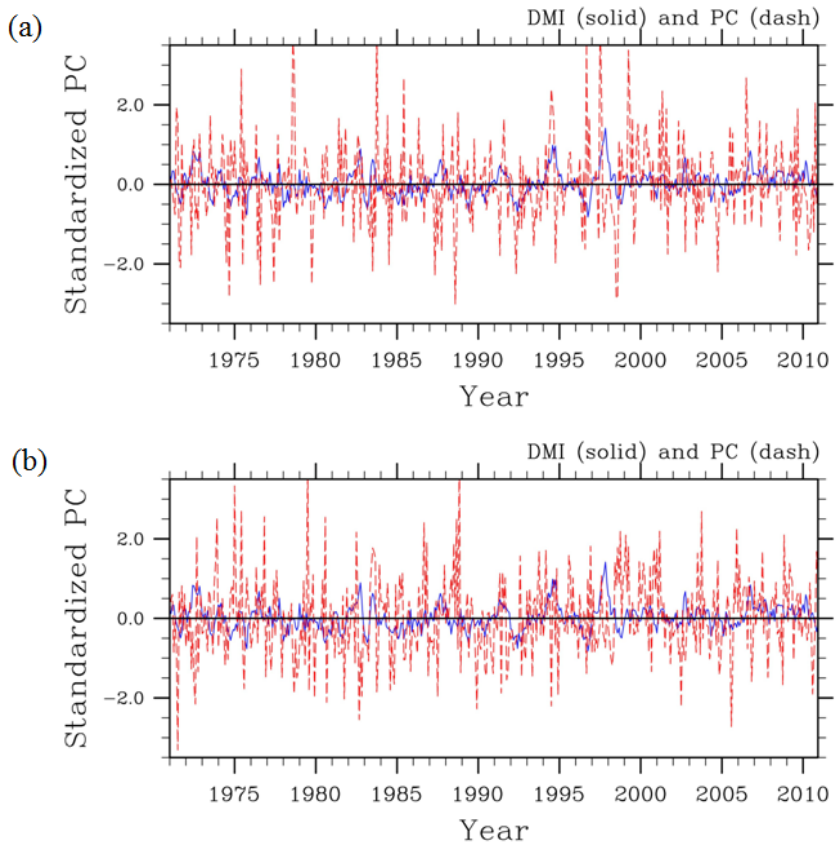


Figure 4.8: Time series plots of DMI (solid line) and principal component (dash line) of (a) the first mode and (b) the second mode.

Table 4.5: The Autocorrelation of DMI and the principal component.

Lag time	DMI	Principal component
0	1.000000000	1.000000000
1	0.7782927000	-0.027282480
2	0.5846329000	0.092103740
3	0.4097738000	-0.087159510
4	0.2808003000	-0.011118800
5	0.1791611000	0.047957360
6	0.1241788000	-0.005514535
7	0.0842163900	0.023777000
8	0.0327468100	0.015913700
9	-0.0137394100	-0.101142400
10	-0.0577931400	0.048489760
11	-0.0879918500	-0.110442100

CHAPTER 5
CONCLUSIONS

The main results found this study are summarized here. In the EOF analysis of the rainfall variability over Thailand during the period 1971 to 2010 the most dominant mode, accounting for 21.6% of the total variance, had a spatial pattern with the same sign (positive) over almost all of Thailand. Consequently, when the PC was positive there was more rainfall than normal and when the PC was negative there was less rainfall than normal almost everywhere. An exception to this general result occurs in the south at latitudes below 9°N.

The study of the teleconnection between rainfall over Thailand and ENSO, found that the differences between the La Niña and El Niño rainfall anomalies were higher during the weak-ENSO periods than during the strong-ENSO periods. Niño 3.4 SST index leads the rainfall anomalies by 4 months. But in the study of the teleconnection between rainfall over Thailand and IOD it was found that the rainfall variability over Thailand has less correlation with IOD.

Furthermore, the maps of the 850 hPa wind and sea level pressure anomalies showed that the wind from the east into Southeast Asia is strengthened during the La Nina events, producing more than normal rainfall in Southeast Asia, and is weakened during the El Nino events, producing less than normal rainfall in Southeast Asia and associated shifts in Walker circulation and correlations between Thailand rainfall and Walker circulation. The interaction of these easterly winds from the Pacific with the south-westerly flow from the Indian Ocean then produce stronger than normal convergence and rainfall in Southeast Asia during the La Nina events, and reduced rainfall during the El Nino events.

REFERENCES

Ashok, K. and Yamagata, T. (2009), Climate change: The El Niño with a difference, *Nature*, **461**, pp. 481-484.

Ashok, W., Guan, Z. and Yamagata T. (2003), A look at the relationship between the ENSO and the Indian Ocean dipole, *J. Meteorol. Soc. Jpn.*, **81**, pp. 41-56.

Baquero-Bernal, A., Latif, M. and Legutke, S. (2002), On dipolelike variability of sea surface temperature in the tropical Indian Ocean, *J. Clim.*, **15**, pp. 1358-1368.

Behera, S.K., Luo, J.J., Masson, S., Rao, S.A., Sakuma, H. and Yamagata, T. (2006), A CGCM study on the interaction between IOD and ENSO, *J. Clim.*, **19**, pp. 1608 – 1705.

Bjerknes, J. (1969), Atmospheric Teleconnections from the Equatorial Pacific, *Mon. Wea. Rev.* **97**(3), pp. 163-172.

Bulic, I.H. and Brankovic, C. (2007), ENSO forcing of the Northern Hemisphere climate in a large ensemble of model simulations based on a very long SST record, *Clim. Dyn.*, **28**, pp.231-254.

Chang, C., Zhang, Y. and Li, T. (2000), Interannual and interdecadal variations of the East Asian summer monsoon and tropical Pacific SSTs. Part I: Roles of the subtropical ridge. J. Clim. **13**, pp. 4310-4325.

Chen, P.Y. and Popovich, P.M. (2002), Correlation: Parametric and Nonparametric Measures. Newbury Park, CA: Sage Publications, Inc.

Compagnucci, R.H. and Richer, M.B. (2008), Can principle component analysis provide atmospheric circulation or teleconnection pattern. *Journal of climatology*, **28**, pp. 703-726.

Diaz, H. F., Hoerling, M. P. and Eischeid, J. K. (2001), ENSO variability, teleconnections and climate change, *International Journal of Climatology*, **21**, pp. 1845-1862.

Dommenget, D. and Latif, M. (2002), A cautionary note on the interpretation of EOFs. *J. Climate*, **15**, pp. 216-225.

Duchon, C. (1979), Lanczos Filtering in One and Two Dimensions, *Journal of Apply Meteorology Climatology*, **18**, pp. 1061-1022.

Elliott, A.C. and Woodward, W.A., (2007), Comparing one or two means using the t-test. Statistical Analysis Quick Reference Guidebook. SAGE.

Emery, J. W. and Thomson, E. R. (2001), Data Analysis Methods in Physical Oceanography. Second edition. Library of Congress Cataloging.

Feng, J., Chen, W., Tam, C.Y. and Zhou, W. (2010), Different impacts of El Nino and El Nino Modoki on China rainfall in the decaying phases, *International Journal of Climatology*, **31**(14), pp. 2091-2101, doi:10.1002/joc.2217.

Francis, D.P., Coats, A.J. and Gibson, D. (1999), How high can a correlation coefficient be?. *Int. J. Cardiol.*, **69**(2), pp. 185-199. doi:10.1016/SO167-5273(99)00028-5.

Hannachi, A., Jolliffe, I.T. and Stephenson, D.B. (2007), Empirical orthogonal functions and related techniques in Atmospheric Science: A review. *International Journal of climatology*, **27**, pp. 1119-1152.

Hastenrath, S. (2002), Dipoles, temperature gradients, and tropical climate anomalies. *Bull. Amer. Meteor. Soc.*, **83**, pp. 735–738.

Hinkle, D.E., Wiersma, W. and Jurs, S.G. (1998), Applied Statistics for the Behavioral Sciences, 4th ed. Boston, Houghton Mifflin Company.

Hoerling, M.P. and Kumar, A. (2002), Atmospheric response patterns associated with tropical forcing, *Journal of Climate*, **15**, pp. 2184-2203.

Hou, M. and Yan, X. (2011), A seasonal scale analysis correlating total cloud amount anomalies with ENSO over the tropical Pacific. *Theor. Appl. Climatol.*, **103**, pp. 471-478. doi:10.1007/s00704-010-0318-3.

Juneng, L. and Tangang, T.F. (2005), Evolution of ENSO-related rainfall anomalies in Southeast Asia region and its relationship with atmosphere-ocean variations in Indo-Pacific sector. *Climate Dynamics*, **25**, pp. 337-350. doi: 10.1007/s00382-005-0031-6.

Kumar, K.K., Rajagopalan, B. and Cane, A.M. (1999), On the weakening relationship between Indian Monsoon and ENSO. *Science*, **284**, pp. 2156–2159.

Li, G., Ren, B. et al. (2010), Indices of El Niño and El Niño Modoki: An improved El Niño Modoki index, *Advances in Atmospheric Sciences*, **27**, pp. 1210-1220.

Limsakul, A., Limjirakan, S. and Suthamanuswong, B. (2007), Spatio-temporal changes in total annual rainfall and the annual number of rainy days in Thailand. *Journal of Environmental Research*, **29**(2), pp. 1-21.

Markowski, C.A. and Markowski, E.P. (1990), Conditions for the Rffectiveness of a Preliminary Test of Variance. *The American Statistician*, **44**(4), pp. 322-326.

McPhaden, M. J., Zebiak, S. E. and Glantz, M. H. (2006), ENSO as an integrating concept in earth science, Science, 314, pp. 1740-1745. Onogi, K., Tsutsui, J. et al. (2007), The JRA-25 reanalysis, *Journal of the Meteorological Society of Japan*, **85**, pp. 369-432.

Mistra, V. (2004), The Teleconnection between the Western Indian and the Western Pacific Oceans, *Amer. Meteor. Soc.*, **132**, pp. 445-455.

O'Brien, R. (1985), The relationship between ordinal measures and their underlying values: Why all the disagreement? Quality & Quantity, **19**, pp. 265–277.

Pearson, K. (1957), The grammar of science. New York: The Meridian Library.

Pribadi, A., Wongwises, P., Humphries, U., Limsakul, A. and Wangwongchai, A. (2012), Diurnal rainfall variation over the three rainfall regions within Indonesia based on ten years of TRMM data. *Journal of Sustainable Energy & Environment*, **3**, pp. 81-86.

Rao, S. A., Behera, S. K., Masumoto, Y. and Yamagata, T. (2002), Interannual subsurface variability in the tropical Indian Ocean with a special emphasis on the Indian Ocean Dipole. *Deep-Sea Res. II*, **49**, pp. 1549-1572.

Rasmusson, E.M. and Carpenter, T.H. (1982), Variation in tropical sea surface temperature and surface wind fields associated with Southern Oscillation/El Niño. *Mon. Wea. Rev.*, **110**, pp. 354–384.

Rodgers, J.L. and Nicewander, W.A. (1998), Thirteen ways to look at the correlation coefficient. *The American Statistician*, **42**(1), pp. 59-66.

Saji, N. H., Goswami, B. N., Vinayachandran, P.N. and Yamagata, T. (1999), A dipole mode in the tropical Indian Ocean, *Nature*, **401**, pp. 360-363.

Schneider, U., Becker, A., Finger, P., Meyer, C.A., Rudolf, B. and Ziese, M. (2013), GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. Theoretical and Applied Climatology, http://dx.doi.org/10.1007/s00704-013-0860-x.

Schneider, U., Becker, A., Finger, P., Meyer, C.A., Rudolf, B. and Ziese, M. (2011a), GPCC Full Data Reanalysis Version 6.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data. DOI: 10.5676/DWD_GPCC/FD_M_V6_050.

Schneider, U., Becker, A., Finger, P., Meyer, C.A., Rudolf, B. and Ziese, M. (2011b): GPCC Full Data Reanalysis Version 6.0 at 1.0°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data. DOI: 10.5676/DWD_GPCC/FD_M_V6_100.

Schneider, U., Becker, A., Finger, P., Meyer, C.A., Rudolf, B. and Ziese, M. (2011c): GPCC Full Data Reanalysis Version 6.0 at 2.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data. DOI: 10.5676/DWD_GPCC/FD_M_V6_250.

Schneider, U., Henning, D., Hauschild, H., Reiss, M. and Rudolf, B. (1992), Zur Berechnung monatlicher Niederschlagshöhen aus synoptischen Meldungen. Meteorol. Zeitschrift N.F. 1/1, pp. 22-31 (English summary).

Siegel, S. (1956), Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill.

Singhrattna, N., Rajagopalan, B., Kumar, K.K. and Clark, M. (2005), Interannual and Interdecadal Variability of Thailand Summer Monsoon Season. *Journal of Climate*, **18**, pp. 1697–1708.

Tozuka, T., Miyasaka, T., Chakraborty, A., Mujumdar, M., Behera, S.K., Masumoto, Y., Nakamura, H. and Yamagata, T. (2006), Univ. of Tokyo coupled general circulation model (UTCM1.0), *Ocean Atmos. Res. Rep*., **44**, Univ. of Tokyo, Tokyo.

Trenberth, K.E., Caron, J.M., Stepaniak, D.P. and Worley S. (2002), Evolution of El Nino-Southern Oscillation and global atmospheric surface temperatures, *Journal of Geophysical Research*, **107**(8), pp. 1-14. doi: 10.1029/2000JD000298.

Trenberth, K. E. and Stepaniak, D. P. (2001), Indices of El Niño evolution, *Journal of Climate*, **14**, pp. 1697-1701.

Trenberth, K.E. (1997), The definition of El Niño, *Bulletin of the American Meteorological Society*, **78**, pp. 2771-2777.

Trenberth, K.E. (1984), Some effects of finite sample size and persistence on meteorological statistics. Part I: Autocorrelations. *Mon. Wea. Rev*., **112**, pp. 326-332.

Wang, B., Wu, R. and Fu, X. (2000), Pacific-East Asian teleconnection: How does ENSO affect East Asian climate?, *Journal of Climate*, **13**, pp. 1517-1536.

Ware W. and Benson J. (1975), Appropriate statistics and measurement scales. *Science Education*, **59**, pp. 575-582.

Webster, P. J., Moore, M.A., Loschnigg, P.J. and Leben, R.R. (1999), Coupled oceanic-atmospheric dynamics in the Indian Ocean during 1997-98. *Nature*, **401**, pp. 356-360.

William, J. E. and Richard, E. T. (2001), Data Analysis Methods in Physical Oceanography. Second edition. Library of Congress Cataloging.

WMO (1990), The Global Precipitation Climatology Project-Implementation and Data Management Plan. WMO/TD-No. 367, pp. 47 and 6 Appendices, Geneva.

Wu, L., Li, C. and Yang, C. (2008), Global Teleconnection in Response to a Shutdown of the Atlantic Meridional Overturning Circulation, *Journal of Climate*, **21**, pp. 3002-3019, doi: 10.1175/2007JCLI1858.1.

Xu, K., Zhu, C.W. and He, J.H. (2012), Linkage between the dominant modes in Pacific subsurface ocean temperature and the two type ENSO events. *Chinese Science Bulletin*, **57** (26), pp. 3491-3496. doi:10.1007/s11434-012-5173-4.

Yamagata, T., Behera, S.K., Luo, J.J., Masson, S., Jury, M.R. and Rao, S.A. (2004), Coupled ocean-atmosphere variability in the tropical Indian Ocean, Geoph. Monog. Series, **147**, pp. 189–211, doi:10.1029/147GM12, 2004.

Ziese, M., Becker, A., Finger, P., Meyer, C.A., Rudolf, B. and Schneider, U. (2011), GPCC First Guess Product at 1.0°: Near real-time first guess monthly land-surface precipitation from rain-gauges based on SYNOP data. doi: 10.5676/DWD_GPCC/FG_M_100.