# The Application of Wavelet Transform to Analyze the Rainfall Data

**Nareemal Hilae**

**A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Mathematics and Statistics**

**Prince of Songkla University**

**2017**

**Title** The Application of Wavelet Transform to Analyze the Rainfall Data

**Author** Miss Nareemal Hilae

**Major Program** Mathematics and Statistics

---

**Advisor**

..............................................

(Dr. Kitipol Nualtong)

**Examination Committee:**

..............................................Chairperson

(Dr. Thammarat Panityakul)

...............................................Committee

(Dr. Kitipol Nualtong)

...............................................Committee

(Asst. Prof. Dr. Watcharin Klongdee)

The Graduate School, Prince of Songkla University, has approved this thesis as partial fulfillment of the requirements for the Master of Science Degree in Mathematics and Statistics.

..........................................................

(Assoc. Prof. Dr. Teerapol Srichana)

Dean of Graduate School

This is to certify that the work here submitted is the result of the candidate's own investigations. Due acknowledgement has been made of any assistance received.

.............................................. Signature
(Dr. Kitipol Nualtong)
Major Advisor

.............................................. Signature
(Nareemal Hilae)
Candidate

I hereby certify that this work has not been accepted in substance for any degree, and is not being currently submitted in candidature for any degree.

............................................. Signature

(Nareemal  Hilae)

Candidate

**ชื่อวิทยานิพนธ์** การประยุกต์การแปลงเวฟเลตกับปริมาณน้ำฝน
**ผู้เขียน** นางสาวนารีมาลย์ หิเล
**สาขาวิชา** คณิตศาสตร์และสถิติ
**ปีการศึกษา** 2559

## บทคัดย่อ

เราแปลงข้อมูลปริมาณน้ำฝนโดยใช้การแปลงฮาร์เวฟเลตและดูบีชีเวฟเลต ซึ่งการแปลงดังกล่าวใช้วิธีการดำเนินการเชิงสเกลและการเลื่อนขนาน เพื่อสร้างฐานหลักเชิงตั้งฉากในการประมาณค่าข้อมูลตั้งต้นและได้ผลลัพธ์ที่มีสมบัติที่ดีกว่าข้อมูลเดิม จากนั้นจะนำข้อมูลที่ได้เข้าสู่กระบวนการ ARIMA model เพื่อประมาณค่าและเปรียบเทียบค่าคลาดเคลื่อน ซึ่งพบว่าข้อมูลที่ผ่านการแปลงโดยใช้ฮาร์เวฟเลตและดูบีชี มีความคลาดเคลื่อนน้อยกว่าข้อมูลที่ไม่ผ่านการแปลงเวฟเลต และข้อมูลที่ผ่านการแปลงด้วยดูบีชีเวฟเลตยังมีประสิทธิภาพมากกว่าข้อมูลที่ผ่านการแปลงด้วยฮาร์เวฟเลตอีกด้วย

| | |
|---|---|
| **Title** | The Application of Wavelet Transform to Analyze the Rainfall Data |
| **Author** | Miss Nareemal Hilae |
| **Major Program** | Mathematics and Statistics |
| **Academic Year** | 2016 |

## ABSTRACT

We transformed the rainfall data by Haar and Daubechies wavelet function. Decompose Haar and Daubechies to scale and translate for constructing an orthogonal basis and also estimated a function to consist with the real value. Then, continue with ARIMA model to approximate and compare with the minimum value of mean absolute error (MAE)and root mean error (RMSE) to make a forecast in the future. We can see that the fitted ARIMA model of Haar and Daubechies discrete wavelet transformed data gives the smaller value of mean absolute error (MAE) and root mean error (RMSE) more than ARIMA of rainfall data. However, the model of Daubechies wavelet transformed data and gives better result more than Haar wavelet transformed data.

# ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisors, Dr. Kittipol Nualtong and Dr. Thammarat Panityakul for their continuous support of my graduate studies and research, for their inspiration, enthusiasm, patience, and immense knowledge. Their guidance helped me in all time for research.

My special appreciation is expressed to Dr. Kittipol Nualtong,
Asst. Prof. Dr. Wathcharin Klongde from Kornkan University and for many valuable comments and helpful suggestions.

I wish to thank all my teachers of the Department of Mathematics and Statistics, Prince of Songkla University for sharing their knowledge and support so that I can obtain this Master degree.

Nareemal Hilae

# CONTENTS

# LIST OF FIGURES

# CHAPTER 1

# Introduction

The wavelet analysis is synthesis of ideas that have intimate connections to several other parts of mathematics, notably phase-space analysis of signal processing. Generally speaking, the wavelet transform is a tool that divides up data, function, or operators into different frequency components and then studies each component with a resolution matched to its scale. It is a tool of analysing a function by correlating it with a two-parameter family of functions, obtained from translating and scaling of a single analysing function, called mother wavelet function. There are several types of wavelets families with different properties, with or without filters, orthogonal or biorthogonal such as Haar wavelet, Daubechies wavelet and Morlet wavelet ect. The focus of this thesis is to employ the orthogonal discrete wavelet transform, namely Haar wavelet transform and Daubechies wavelet transform.

Autoregressive integrated moving average (ARIMA)model is a forecasting a time series model that have been used in the forecasting literature. In order to provide estimates for the future, these models analyze the historical data, generally speaking, ARIMA model is a forecasting model that utilize historycal information to make predictions. ARIMA works best when data exhibits a outliers, it is usually superior to exponential smoothing techniques when the data is reasonably long and the correlation between past observation such as Fourier transform and wavelet transform to obtain a stationary time series.

We transformed the data by discrete wavelet transform with the based on linear algebra and processing with ARIMA model. We start to use a mother wavelet function, scaled and translated for construct an orthogonal basis and estimated function to consist with the real value. Then continue to approximate to make the forecasting and compare with the minimum value of MAE and RMSE in ARIMA model. In 2002 [6], Rumaih M.and Muhammad A used Saudi stock to represent that wavelet transform is good than other forecasting technique in forecasting non-stationary time series. In 2008 [5], Aggarwal et al, introduce that forecasting efficacy of the wavelet transforms based mixed model has been compare with the other three models. They will find the good evaluation for different wavelets were conducted, and notice that for make forecasting precision using haar wavelet transform to gave the best evaluations. In 2011 [4], S.Alwadi and Mohd Tahir Ismail use an amman stock market data from 1993 until 2009 to transform with the based on wavelet transform combined with ARIMA model, and then forecasting results will compare with the minimum value of MAE and RMSE considered to select the best ARIMA model of the daily return data. After we transform via Haar and Daubechies, we get the best result more than original return data.

This thesis will focus to transform the data by discrete wavelet transform with the based on linear algebra. We will start to construct an orthogonal basis by a mother wavelet which called Haar wavelet for estimated a function to consist with the real value. We used Haar wavelet and Daubechies wavelet to scale and translate in the discrete wavelet transform. The discrete Haar wavelet was transformed to construct an orthogonal basis by the theory of linear algebra, that was if $f \in L^2(\mathbb{R})$, $f$ can be written in the form

$$f = \sum_{j,k \in Z} < f, p_{j,k} > p_{j,k} + \sum_{j,k \in Z} < f, h_{j,k} > h_{j,k}$$

where $p_{j,k}(x) = 2^{j/2} p(2^{j/2}x - k)$ and $h_{j,k}(x) = 2^{j/2} h(2^{j/2}x - k)$.

For this study, we used the data of rainfall in Songkhla area from 2007 until 2014 which the Haar and Daubechies wavelet function to construct an orthogonal basis. We scaled, translated and used ARIMA model to approximate to make the forecasting and compare with the minimum value of the mean absolute error (MAE), the root mean square Error (RMSE).

In our work, we will be organize $5$ chapters. An introduction about application of wavelet transform to analyze the rainfall data is presented in chapter $1$. In chapter $2$, we summarize an introduction of wavelet transform. Summarise an introduction of ARIMA model in chapter $3$. Chapter $4$, we represented the experiment results in the forecasting method. The last chapter we summarize our mention preconcession.

First, we indicate state fundamental definition, example, theorem and some properties that will be used in the proceeding chapter.

In this thesis, we transform with the based on linear algebra which some properties of inner product space and now ready to state the definition and some properties as following :

## 1.1 Inner Product Space

**Definition 1.1.** Let $V$ be a vector space on $\mathbb{R}$. We call function $< \cdot, \cdot >$ : $V \times V \longrightarrow \mathbb{R}$ an inner product if for all $x, y$ and $z$ in $V$ and $c$ in $\mathbb{R}$, we have

1. $< x, y > = < x, y > + < z, y >$

2. $< cx + z, y > \; = \; c < x, y >$

3. $< x, y > \; = \; < y, x >$.

4. $< x, x > \; \geq \; 0$, and $< u \, , u > \; = \; 0$ iff $u \; = \; \overline{0}$.

**Definition 1.2.** Let $V$ be an inner product space and $u, \; v \in V$.

The norm of $u$ is written by $\|u\| =< u, u >^{\frac{1}{2}}$ . The distance of $u$ and $v$ denoted by

$$d(u, \; v) \; = \; \|u - v\|.$$

**Theorem 1.1.** *If $V$ be an inner product space and $x, \; y \in V$ then*

$$| < x, y > | \; \leq \; \|x\| \|y\|.$$

*Proof.* Details of the proof can be found in [2]. □

**Theorem 1.2.** *Let $V$ be an inner product space. If $x, \; y \in V$ and $c \in \mathbb{R}$,*

*1.* $\|x\| \; \geq \; 0$

*2.* $\|x\| \; \geq \; 0$ *if and only if* $x \; = \; 0$

*3.* $\|cx\| \; = \; |c| \|x\|$

*4.* $\|x + y\| \; \leq \; \|x\| \; + \; \|y\|$.

*Proof.* Details of the proof can be found in [2]. □

**Definition 1.3.** Let $V$ be an inner product space. Two vector $x, \; y \in V$ are orthogonal if $< x, y > \; = \; 0$.

**Definition 1.4.** Let $U$ and $V$ be subspace $S$ of a vector space $W$ such that $U \cap \; V \; = \{0\}$. The *direct sum* of $U$ and $V$ is denoted by

$$U \; \oplus \; V \; = \; \{x \; + \; y | \; x \; \in U \text{ and } y \in V\}.$$

**Definition 1.5.** Let $S$ be a subspace of the inner product space $V$. The orthogonal complement of $S$ is the set $S^{\perp} = \{x \in V \mid < x, s > = 0$ for all $s \in S\}$.

**Theorem 1.3.** *1. If $U$ and $V$ are subspace of $W$ with $U \cap V = \{0\}$, then $U \oplus V$ is also a subspace of $W$.*

*2. If $S$ is a subspace of the inner product space $V$, then $S^{\perp}$ is also a subspace of $V$.*

*Proof.* Details of the proof can be found in [2]. □

**Theorem 1.4.** *If $U$ and $V$ are subspace of $W$ with $U \cap V = \{0\}$, and $z \in U \oplus V$, then $w = u + v$ for unique $x \in U$ and $y \in V$.*

*Proof.* Details of the proof can be found in[2]. □

**Definition 1.6.** Let $V$ be an inner product space.

1. A subset $S \subseteq V$ is called **orthogonal** if $\forall x, y \in S, < x, y > = 0$.

2. $x \in V$ is a **unit vector** if $\|x\| = 1$.

3. A subset $S \subseteq V$ is **orthonormal** if $S$ is a orthogonal set and $\forall x \in S$, $\|x\| = 1$.

**Definition 1.7.** Let $V$ be a vector space and $S = \{v_1, v_2, ..., v_n\}$ be a subset of $V$. We say that $S$ **span** $V$ if every vector $v$ in $V$ can be written as a linear combination of vectors in $S$, that is $v = c_1v_1 + c_2v_2 + ... + c_nv_n$. for $v_1, v_2, ..., v_n \in S$ and $c_1, c_2, ..., c_n \in \mathbb{R}/\mathbb{C}$.

**Definition 1.8.** A **basis** of a vector space $V$ is defined as a subset $v_1, v_2, ..., v_n$ of vectors in that are linearly independent and span vector space $V$.

**Definition 1.9.** An **orthonormal basis** of V is an orthonormal list of vectors in V that is also a basis of V.

**Theorem 1.5.** *Let $V$ be an inner product space and $S = \{v_1, v_2, ..., v_k\}$ be an orthogonal subset of $V$ such that $\forall i,\ v_{i=\ 1,2,...,k}\ \neq 0$. If $y \in$ Span(S), then*

$$y \ = \ \sum_{i=1}^{k} \frac{< y, v_i >^2}{\parallel v_i \parallel} v_i.$$

*Proof.* Details of the proof can be found in [3]. □

**Theorem 1.6.** *Let $V$ be non-zero finite dimensional inner product space. Then $V$ has an orthonormal basis $B$. Furthermore, if $B \ = \ \{v_1, v_2, ..., v_n\}$ and $x \in V,\ then$*

$$x \ = \ \sum_{i\ =\ 1}^{k} < x, v_i > v_i.$$

*Proof.* Details of the proof can be found in [3]. □

## 1.2 Criteria of Statistics

Always some amount of error in every analysis, We examine and compare error to know reliability. In this thesis, we used Mean Error and Root Mean Square Error to examine error for compare to choose the best result.

1**. Mean absolute error(MAE)**

Mean absolute error is a quantity used to measure how close forecasts or predictions are to the eventual outcomes. The mean absolute error is given by

$$MAE \ = \ \frac{\sum_{i=1}^{N} |\text{actual value - predicted value}|}{N}.$$

2**. Root mean squared error (RMSE)**

Root mean squared error is the square root of the mean/average of the square of all of the error. RMSE is very commonly used and makes for an excellent general purpose error metric for numerical predictions

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\text{actual value - predicted value})^2}{N}}.$$

# CHAPTER 2

# Wavelet Transform

The wavelet transform is similar to the Fourier transform (or much more to the windowed Fourier transform) with a completely different good function. The main difference is this fourier transform decomposes the signal into sine and cosines. Wavelet transform is very suitable with non-stationary data and were used to adopt a mother function, scaled and translated. We can used wavelet transform is transforming with the based function which a mother wavelet to scale and translate and used orthogonal wavelets for discrete wavelet transform improvement and non-orthogonal wavelets for continuous wavelet transform improvement. The wavelet transform is divided into two sort, that is continuous wavelet transform and discrete wavelet transform.

## 2.1 Continuous Wavelet Transform

Now we start to state the definition of continuous wavelet transform. Generally, continuous wavelet transform can be expressed by the following.

**Definition 2.1.** Fix $\varphi \in L^2(\mathbb{R}),\ \varphi \neq 0$, call it the mother wavelet function. Given $a > 0, b \in \mathbb{R}$,

$$\varphi_{a,b}(x) = \frac{1}{\sqrt{a}}\varphi(\frac{x-b}{a}).$$

The wavelet transform of $f \in L^2(\mathbb{R})$ is

$$Wf(a,b) = < f, \varphi_{(a,b)} > = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} f(x)\overline{\psi(\frac{x-b}{a})}dx.$$

**Theorem 2.1.** *Let $f \in L^2(\mathbb{R})$, an inverse wavelet transform is*

$$f(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} Wf(a,b)\psi(\frac{x-b}{a})dadb.$$

*Proof.* Details of the proof can be found in [7]. □

For convenience to transform the data, we will used the discrete wavelet transform which the same main idea of the continuous wavelet transform, begin to used the mother wavelet function, scaling and translating this functon to construct an orthogonal basis.

## 2.2 Discrete Wavelet Transform

The discrete wavelet transform is an implementation of the continuous wavelet transform using a discrete set of the wavelet scales and translations obeying some defined rules. In other words, this transform decomposes the function (data) into mutually orthogonal set of wavelet. The wavelet can be construct from a scaling function $\phi(x)$, for each $j \in \mathbb{Z}$, define the approximation space

$$V_j = \text{span}\{\phi_{j,k}(x)\}_{k\in\mathbb{Z}}.$$

and one can show that $\forall_j \in \mathbb{Z}$, $\{\phi_{j,k}\}_{k=-\infty}^{\infty}$ is an orthogonal basis of $V_j$. We also define a mother function $\psi(x)$, for each $j \in \mathbb{Z}$ define the approximation space

$$W_j = \text{span}\{\psi_{j,k}(x)\}_{k\in\mathbb{Z}}$$

and one can show that $\forall_j \in \mathbb{Z}$, $\{\psi_{j,k}\}_{k=-\infty}^{\infty}$ is an orthogonal basis for $W_j$. Moreover, we can show that $\forall f \in V_{j+1}$,

$$f = \sum_{j,k=-\infty}^{\infty} < f, \phi_{j,k} > \phi_{j,k} + \sum_{j,k=-\infty}^{\infty} < f, \psi_{j,k} > \psi_{j,k}.$$

Finally, $L^2(\mathbb{R})$ can be decomposed as infinite orthogonal direct sum $L^2(\mathbb{R}) = V_0 \oplus W_0 \oplus W_1 \oplus ....$ In particular, $\forall f \in L^2(\mathbb{R})$,

$$f = f_0 + \sum_{j=0}^{\infty} w_j$$

where $f_0 \in V_0$ and $w_j \in W_j$.

### 2.2.1 Haar Discrete Wavelet Transform

Now, we summarize the Haar wavelet transform and some properties of Haar wavelet transform. Firstly, we introduce the Haar scaling function and mother wavelet function as follow.

**Definition 2.2.** Let

$$p(x) = \begin{cases} 1, x \in [0,1) \\ 0, \text{otherwise.} \end{cases}$$

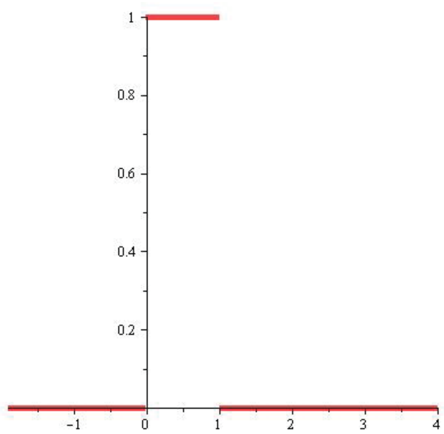


Figure 2.1: Haar scaling function $p(x)$.

For each $j, k \in Z$, define

$$p_{j,k}(x) = 2^{j/2} p(2^j x - k).$$

The collection $\{p_{j,k}(x) = 2^{j/2}p(2^jx - k)\}$ is referred to as the Haar scaling function. For each $j \in Z$, the collection $\{p_{j,k}(x) = 2^{j/2}p(2^jx - k)\ ; k \in \mathbb{Z}\}$ is referred to as the system of scale $j$ Haar scaling functions. Next, we state definition of Haar function as follow.

**Definition 2.3.**

$$h(x) = \begin{cases} 1, & x \in [0, \frac{1}{2}) \\ -1, & x \in (\frac{1}{2}, 1] \\ 0, & \text{otherwise.} \end{cases}$$

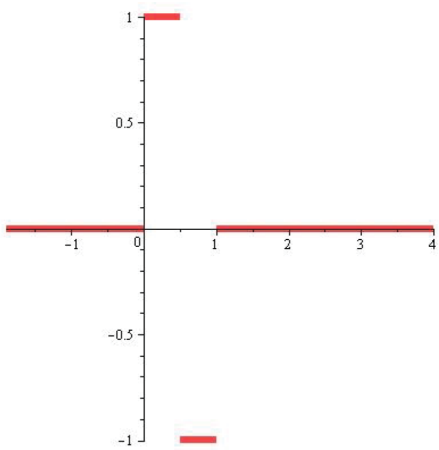


Figure 2.2: Haar wavelet function $h(x)$.

For each $j, k \in Z$, define

$$h_{j,k}(x) = 2^{j/2}h(2^jx - k).$$

The collection $\{h_{j,k}(x) = 2^{j/2}h(2^jx - k)\ ; j, k \in \mathbb{Z}\}$ is referred to as the Haar system on $\mathbb{R}$. For each $j \in Z$, the collection $\{h_{j,k}(x) = 2^{j/2}h(2^jx - k)\ ; k \in \mathbb{Z}\}$ is referred to as the system of scale $j$ Haar functions.

Now, we will introduce some definition and some properties for the understanding of proof and the construction of general wavelet bases.

**Definition 2.4.** For each pair of integer $j,\ k \in \mathbb{Z},$ define the interval $I_{j,k}$ by

$$I_{j,k} \ = \ [2^{-j}k,\ 2^{-j}(k+1)).$$

The collection of all such intervals is called the collection of dyadic subinterval of $\mathbb{R}$.

**Lemma 2.2.** *Let $j,\ j',\ k,\ k' \in \mathbb{Z}$ be such that $(j,\ k)\ \neq\ (j',\ k')$, then either*

*(i) $I_{j',k'}\ \cap I_{j,k}\ =\ \varnothing$, or*

*(ii) $I_{j',k'}\ \subset I_{j,k}$, or*

*(iii) $I_{j,k}\ \subset I_{j',k'}$.*

*Proof.* Details of the proof can be found in [7]. □

**Definition 2.5.** Given a dyadic interval at scale $j$, we write $I_{j,\ k}\ =\ I^{l}_{j,k}\ \cup\ I^{r}_{j,k}$, where $I^{l}_{j,k}$ and $I^{r}_{j,k}$, are scale $j+1$, to denote the left half and right half of the interval $I_{j,\ k}$. In fact, $I^{l}_{j,k}\ =\ I_{j+1,2k}$ and $I^{r}_{j,k}\ =\ I_{j+1,2k+1}$.

As we have a mother wavelet function, we scale and translate mother wavelet function for construct an orthogonal basis, so now we continue to summarize that some properties and theorem of the based on orthogonality of the Haar system.

**Theorem 2.3.** *The Haar system on $\mathbb{R}$ is an orthonormal system on $\mathbb{R}$.*

*Proof.* For $k \neq k',\ h_{j,k}(x)h_{j,k'}(x) = 0$ for all $x \in \mathbb{R}$.

Next, if $k = k^{'}$ , we have

$$< h_{j,k'}, h_{j,k} > \ = \ \int_{\mathbb{R}} |h_{j,k}|^2 dx = 1.$$

Assume that $j \neq j^{'}$. Without loss of generality it is sufficient to consider the case $j > j^{'}$.

Then, it follows that there 3 distinct possibility:

Case(i) $I_{j,k} \cap I_{j',k'} = \emptyset$ is elementary, as in this case $h_{j,k}(x)h_{j',k'}(x) = 0$ and so

$$< h_{j,k}, h_{j',k'} > \ = \ 0.$$

Case(ii) $I_{j,k} \subset I^l_{j',k'}$ implies that whenever $h_{j,k}$ is non-zero, $h_{j',k'}$ is constantly equal to $1$. Thus

$$< h_{j,k}, h_{j',k'} > \ = \ \int_{I_{j,k}} h_{j,k}(x)h_{j',k'}(x)dx \ = \ 0.$$

Similarly, the case(iii) $I_{j,k} \subset I^r_{j',k'}$ implies that whenever $h_{j,k}(x)$ is non-zero, $h_{j',k'}(x)$ is constantly equal to $-1$. Thus

$$< h_{j,k}(x), h_{j',k'}(x) > \ = \ \int_{I_{j,k}} h_{j,k}(x)h_{j',k'}(x)dx \ = \ 0.$$

□

Next, we estimate the function by discrete wavelet transform and then ready to study basic knowledge of some definition and conditions of an approximation operator.

**Definition 2.6.** For each $j \in Z$, define the approximation space $V_j$, by

$$V_j \ = \ \overline{\text{span}\{p_{j,k}\}}_{k\in\mathbb{Z}}.$$

**Theorem 2.4.** *Fix any $j \in \mathbb{Z}$. Then the set of functions $\{p_{j,k}\ , k \in \mathbb{Z}\}$ is an orthogonal basis for $V_j$.*

*Proof.* Details of the proof can be found in [7]. □

**Definition 2.7.** For each $j \in \mathbb{Z}$, we define the approximation space $W_j$, by

$$W_j \ = \ \overline{\text{span}\{h_{j,k}\}}_{k\in\mathbb{Z}}.$$

**Theorem 2.5.** *Fix any $j \in \mathbb{Z}$. Then the set of functions $\{h_{j,k}\ , k \in \mathbb{Z}\}$ is an orthogonal basis for $W_j$.*

*Proof.* Since $\{p_{j/2,k}\}_{k\in\mathbb{Z}}$ is an orthogonal basis of $V_{j+1}$ and $W_j \subset V_{j+1},$ every function in $W_j$ has an expansion of the type

$$f(x) = \sum_{k\in\mathbb{Z}} c_k p_{j/2,k}(x).$$

In dyadic interval of length $2^{-j}$ the average of $f$ is zero (since $f \in W_j$), so by combining two of the function $p_{j/2,k}$ we get,

$$f(x) = \sum_{k\in\mathbb{Z}} d_k h_{j,k}(x).$$

Conversely, every sum of this type is orthogonal to every function in $V_j$ since

$$p_{j,k} \perp h_{j,k}, \quad \forall j, k \in \mathbb{Z}.$$

□

**Theorem 2.6.** *(Splitting Theorem) For each $j, k \in \mathbb{Z}$*

*(i) $V_{j+1} = V_j \oplus Wj$*

*(ii) every $f \in V_{j+1}$ has expansion*

$$f = \sum_{k\in\mathbb{Z}} a_{j,k} p_{j,k} + \sum_{k\in\mathbb{Z}} b_{j,k} h_{j,k}.$$

*(iii) The set of function*

$$\{p_{j,k},\ k \in \mathbb{Z}\} \bigcup \{h_{j,k},\ k \in \mathbb{Z}\}$$

*is an orthogonal basis in $V_{j+1}$.*

*Proof.* (i) Let $r \in V_j$ , $q \in W_j$, we get $r + q \in V_{j+1}$.

Given $f = r + q$, and $V_j \perp W_j$. Hence $f \in V_{j+1}$.

(ii) Since $f \in V_{j+1}$ then $f$ has expansion

$$f = \sum_{k \in \mathbb{Z}} a_{j,k} p_{j,k} + \sum_{k \in \mathbb{Z}} b_{j,k} h_{j,k}.$$

(iii) Follow from (ii) and note that $h_{j,k} \perp p_{j,l} \quad \forall k, l \in \mathbb{Z}$, since $p_{j,l} \in V_j$ and $h_{j,k} \in W_j$, and $V_j \perp W_j$. Hence the set of function

$$\{p_{j,k},\ k \in \mathbb{Z}\} \bigcup \{h_{j,k},\ k \in \mathbb{Z}\}$$

is an orthogonal basis in $V_{j+1}$. □

**Lemma 2.7.** *$W_j \perp W_l$ for $j \neq l$.*

*Proof.* If $l < j$, then $W_l \perp V_j$, and $W_l \subset V_{l+1} \subset V_j$. By repeating this splitting over and over again we get,

$$V_j = W_{j-1} \oplus V_{j-1} = W_{j-1} \oplus W_{j-2} \oplus V_{j-2} = ....$$

□

**Theorem 2.8.** *The $L^2(\mathbb{R})$ can be decomposed as an infinite orthogonal direct sum*

$$L^2(\mathbb{R}) = V_0 \oplus W_0 \oplus W_1 \oplus ...$$

*In particular, each $f \in L^2(\mathbb{R})$ can be written uniquely as*

$$f = f_0 + \sum_{j=0}^{\infty} w_j$$

*where $f_0$ belongs to $V_0$ and $w_j$ belongs to $W_j$.*

*Proof.* If $f$ is an orthogonal to all of these function, since $V_j = W_{j-1} \oplus W_{j-2} \oplus ... \oplus W_0 \oplus V_0$ which $f_0 \in V_0,\ w_j \in W_j$, then we get $f_j$ is an orthogonal in $V_j$. So for each $f_j \in V_j$ can be decomposed uniquely as a sum

$$f_j = w_{j-1} + w_{j-2} + ... + w_0 + f_0.$$

See figure 2.3. □

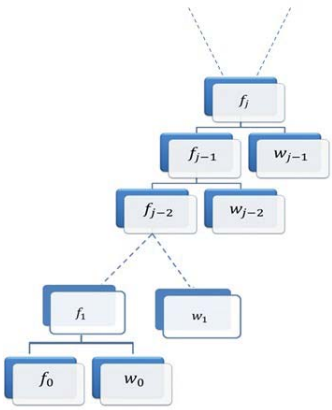


Figure 2.3: Haar decomposition.

### 2.2.2 Filter and Diagrams

After we study every conditions, we use all to writhed the diagrams by theorem as following for convenience to understanding.

**Lemma 2.9.** *The following relations hold* $\forall x \in \mathbb{R}$*:*

$$p_{j,k}(2^j x) = \frac{h_{j,k}(2^{j-1}x) + p_{j,k}(2^{j-1}x)}{2}$$

$$p_{j,k}(2^j x - 1) = \frac{p_{j,k}(2^{j-1}x) - h_{j,k}(2^{j-1}x)}{2}$$

*Proof.* Details of the proof can be found in [6]. □

**Theorem 2.10.** *Suppose*

$$f_j(x) \;=\; \sum_{k\in\mathbb{Z}} a_k^j p_{j,k}(x) \;\in V_j.$$

*Then* $f_i$ *can be decomposed as*

$$f_i \;=\; w_{j-1} + f_{j-1}$$

*where*

$$w_{j-1} \;=\; \sum_{k\in\mathbb{Z}} b_k^{j-1} h_{j,k}(2^{j-1}x-k) \;\; \in W_{j-1}$$
$$f_{j-1} \;=\; \sum_{k\in\mathbb{Z}} a_k^{j-1} p_{j,k}(2^{j-1}x-k) \;\; \in V_{j-1}$$

*with*

$$b_k^{j-1} \;=\; \frac{a_{2k}^j - a_{2k+1}^j}{2},\; a_k^{j-1} \;=\; \frac{a_{2k}^j + a_{2k+1}^j}{2}.$$

*Proof.* Divide the sum $f_j(x) = \sum_{k\in\mathbb{Z}} a_k p_{j,k}(2^j x - k)$ into even and odd terms :

$$f_j(x) \;=\; \sum_{k\in\mathbb{Z}} a_{2k} p_{j,k}(2^j x - 2k) \;+\; \sum_{k\in\mathbb{Z}} a_{2k} p_{j,2k}(2^j x - 2k - 1).$$

By Lemma 2.9, we get

$$f_j(x) = \sum_{k\in\mathbb{Z}} a_{2k}(h_{j,k}(2^{j-1}x-k) + p_{j,k}(2^{j-1}x-k))/2 \;+$$
$$\sum_{k\in\mathbb{Z}} a_{2k}(p_{j,k}(2^{j-1}x-k) - h_{j,k}(2^{j-1}x-k))/2.$$
$$= \sum_{k\in\mathbb{Z}} \frac{(a_{2k}-a_{2k+1})h_{j,k}(2^{j-1}x-k)}{2} + \sum_{k\in\mathbb{Z}} \frac{(a_{2k}+a_{2k+1})p_{j,k}(2^{j-1}x-k)}{2}$$
$$= w_{j-1} \;+\; f_{j-1}.$$

□

**Theorem 2.11.** *Suppose for* $0 \le j' < j$*. Then*

$$f(x) \;=\; \sum_{k\in\mathbb{Z}} a_k^{j-1} p_{j,k}(2^{j-1}x-k) \;\; \in V_j$$

*where the $a_l^{j'}$ are determined recursively for $j' = 1$, then $j' = 2$, and so on until $j = j'$, by algorithm*

$$a_l^{j'} = \begin{cases} a_k^{j'-1} + b_l^{j'-1};\ l = 2k \textit{ is even} \\ \\ a_k^{j'-1} - b_l^{j'-1};\ l = 2k+1 \textit{ is odd.} \end{cases}$$

*Proof.* $f(x) = f_0(x) + w_0(x) + ... + w_{j-1}(x)$ , $w_l \in W_l$ where

$f_0(x) = \sum_{k\in\mathbb{Z}} a_k^0 p_{j,k}(x-k) \in V_0$ and $w_l(x) = \sum_{k\in\mathbb{Z}} b_k^l p_{j,k}(2^l x - k) \in W_l$ for $0 \le l \le j-1$. Using lemma 2.9 with $x$ replaced by $x-k$, we have

$f_0(x) = \sum_{l\in\mathbb{Z}} (a_k^0 p_{j,k}(2x-2k) + a_k^0 p_{j,k}(2x-2k-1))$ and so

$f_0(x) = \sum_{l\in\mathbb{Z}} \widehat{a}_l^1 p_{j,k}(2x-l)$ where

$$\widehat{a}_l^1 = \begin{cases} a_k^0;\ l = 2k \text{ is even} \\ \\ a_k^0;\ l = 2k+1 \text{ is odd.} \end{cases}$$

Similarly, $w_0 = \sum_{l\in\mathbb{Z}} b_l^0 h_{j,k}(x-k)$ can be written as

$w_0(x) = \sum_{l\in\mathbb{Z}} \widehat{b}_l^1 h_{j,k}(2x-l)$ where

$$\widehat{b}_l^1 = \begin{cases} b_k^0;\ l = 2k \text{ is even} \\ \\ -b_k^0;\ l = 2k+1 \text{ is odd.} \end{cases}$$

$$f_0(x) + w_0(x) = \sum_{l\in\mathbb{Z}} a_l^1 p_{j,k}(2x-l) \text{ where}$$

$$a_l^1 = \widehat{a}_l^1 + \widehat{b}_l^1 \begin{cases} a_k^0 + b_k^0;\ l = 2k \text{ is even} \\ \\ a_k^0 - b_k^0;\ l = 2k+1 \text{ is odd.} \end{cases}$$

Adding $w_1(x) = \sum_{k\in\mathbb{Z}} b_k^1 h_{j,k}(2x-k)$ in to the both side we have

$$f_0(x) + w_0(x) + w_1(x) = \sum_{l\in\mathbb{Z}} a_l^2 p_{j,k}(2^2x - l) \text{ where}$$

$$a_l^2 = \widehat{a}_l^1 + \widehat{b}_l^1 \begin{cases} a_k^1 + b_k^1;\ l = 2k \text{ is even} \\ a_k^1 - b_k^1;\ l = 2k+1 \text{ is odd.} \end{cases}$$

$a_l^0,\ b_l^0-$ coefficients determine the $a_l^1-$ coefficients.

$a_l^1,\ b_l^1-$ coefficients determine the $a_l^2-$ coefficients, and so on in a recursive manner.

□

And now, we begin to define discrete filter $H$ and $L$ via there impulse responses, which are the sequence $h$ and $l$ :

$$h = (...0...\frac{-1}{2}\frac{1}{2}...0),\ l = (...0...\frac{1}{2}\frac{1}{2}...0).$$

If $\{x_k\} \in l^2$, then $H = h * x$ and $L = l * x$.

Thus we have

$$H(x)_k = (h * x)_k = \frac{1}{2}x_k - \frac{1}{2}x_{k+1},\ L(x)_k = (l * x)_k = \frac{1}{2}x_k + \frac{1}{2}x_{k+1}.$$

The operation of discarding the odd of coefficient of a sequence is called **downsampling**, denoted by $D$.

Go from level $j$ scaling coefficients $a_k^j$ to get the level $j-1$ scaling and wavelet coefficients.

Decomposition $f$ into $V_0$ and $W_{j'}$ component for $0 \leq j' \leq j$, $b_k^{j-1} = DH(a^j)_k$ and $a_k^{j-1} = DL(a^j)_k$.

Next, we define discrete filter $\widetilde{H}$ and $\widetilde{L}$ via there impulse responses, which are the sequence $\widetilde{h}$ and $\widetilde{l}$ :

$$\widetilde{h} = (...0... -1\ 1...0...),\ \widetilde{l} = (...0...1\ 1...0...).$$

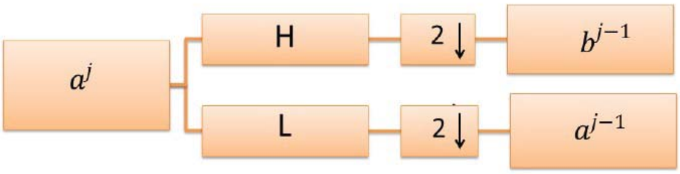


Figure 2.4: Haar decomposition diagram.

For a sequence $\{x_k\}$, we have

$$\widetilde{H} = (\widetilde{h} * x)_k = x_k - x_{k+1}, \; \widetilde{L} = (\widetilde{l} * x)_k = x_k + x_{k+1}.$$

If $x$ and $y$ are sequence in which the odd entries are all $0$, then

$$(\widetilde{h} * x)_l = \begin{cases} x_{2k};\; l = 2k \; is \; even \\ -x_{2k};\; l = 2k+1 \; is \; odd. \end{cases}$$

$$(\widetilde{l} * y)_l = \begin{cases} y_{2k};\; l = 2k \; is \; even \\ -y_{2k};\; l = 2k+1 \; is \; odd. \end{cases}$$

Assume that the $x_{2k+1}$ and $y_{2k+1}$ are $0$. Next, we set $x_{2k} = b_k^{j-1}$ and $y_{2k} = a_k^{j-1}$, that is

$$x = (...0 \; b_{-1}^{j-1} \; 0 \; b_0^{j-1} \; 0...)$$

and

$$y \;=\; (...0\; a_{-1}^{j-1}\; 0\; a_0^{j-1}\; 0...).$$

Adding the two sequence $\widetilde{h} * x$ and $\widetilde{l} * x$, we have

$$(\widetilde{h} * x)_l + (\widetilde{l} * y)_l = \begin{cases} x_{2k} + y_{2k};\; l = 2k \; is\; even \\ \\ x_{2k} - y_{2k};\; l = 2k+1 \; is\; odd. \end{cases}$$

We use $U$ to denote the **upsamping** operator, so

$$x \;=\; Ub^{j-1} \text{ and } y \;=\; Ua^{j-1}.$$

$$a_j \;=\; \widetilde{L}Ua^{j-1} \;+\; \widetilde{H}Ua^{j-1}$$

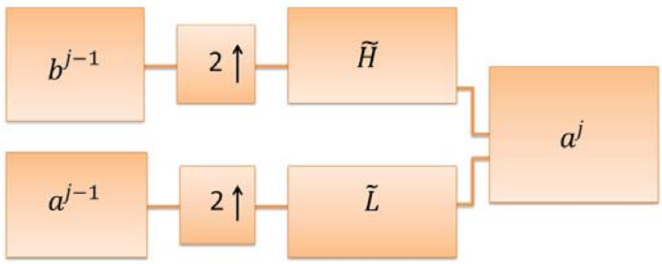


Figure 2.5: Haar reconstruction diagram.

### 2.2.3 Summary

After we study all of the above, can be summarized in the following steps as figure 2.6.

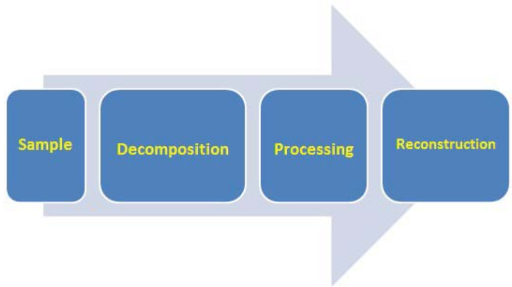


Figure 2.6: Step of Haar Wavelet transform.

Given function $y = f(t)$. Let $p$ and $h$ be the scaling Haar function and Haar(mother)wavelet function.

**STEP 1.** (Sample)

Choosing a positive integer. Set $a_k^J = f(\frac{k}{2^J})$.

Let us compute

$$f_J(x) = \sum_{k\in\mathbb{Z}} a_k^J p_{j,k}(2^J x - k).$$

**STEP 2.** (Decomposition)

$f_i$ can be decomposed as

$$f_j = w_{j-1} + f_{j-1}$$

where

$$w_{j-1} = \sum_{k\in\mathbb{Z}} b_k^{j-1} h_{j,k}(2^{j-1}x - k) \quad \in W_{j-1},$$

$$f_{j-1} \;=\; \sum_{k\in\mathbb{Z}} a_k^{j-1} p_{j,k}(2^{j-1}x-k) \quad \in V_{j-1}$$

The coefficients $b_{j-1}^{l},\ a_{j-1}^{l}$ are determined from the recursively by the algorithm

$$b_k^{j-1} \;=\; DH(a^j)_k \; and \; a_k^{j-1} \;=\; DL(a^j)_k,$$

where $H$ and $L$ are the high and low pass filters.

**STEP 3.** (Processing)

After we decompose, the function in the form

$$f_J(x) \;=\; \sum_{j=0}^{J-1} w_j + f_0$$

it can now be filtered by modifying the wavelet coefficients $b_j^k$.

**STEP 4.** Reconstruction

Take the modified function, $f_J$ by

$$f \;=\; f_0 + w_0 + ... + w_{j-1},$$

with

$$f_0 \;=\; \sum_{k\in\mathbb{Z}} a_k^0 \phi(x-k) \in V_0 \; and \; , w_{j'} \;=\; \sum_{k\in\mathbb{Z}} b_k^{j'} \phi(2^{j'}x-k) \in W_{j'}.$$

This is accomplished by the reconstruction algorithm

$$a_j \;=\; \widetilde{L}Ua^{j-1} \;+\; \widetilde{H}Ub^{j-1}.$$

### 2.2.4 Daubechies Discrete Wavelet Transform

Daubechies wavelet transforms is a discrete wavelet transform, it is defined in the same way as the Haar wavelet transform by computing the running averages and difference via scalar function and mother function the only difference between them consists in how these scaling signal and wavelets are defined. Daubechies wavelet are family of orthogonal wavelet set characterizing by a maximal of vanishing moments for same given support. In general the Daubechies wavelets are chosen to have the highest number $N$ of vanishing moments, for given support width $2N-1$. The name $D-N$ is the length or number of tape. By the same idea with Haar wave transform, The Daubechies wavelet of class $D-2N$ is a function $\psi \in L^2(\mathbb{R})$ defined by

$$\psi(x) : \; = \sqrt{2}\sum_{k=0}^{2N-1}(-1)^k h_{2N-1-k}\psi(2x-k),$$

where $h_0, ..., h_{2N-1} \in \mathbb{R}$ are the constant the *filter coefficient* satisfies the condition

$$\sum_{k=0}^{N-1} h_{2k} \; = \; \frac{1}{\sqrt{2}} \; = \; \sum_{k=0}^{N-1} h_{2k+1},$$

as well as, for $l = 0, 1, ..., N-1$,

$$\sum_{k=2l}^{2N-1+2l} h_k h_{k-2l} = \begin{cases} 1, \text{if } l = \; 0 \\ \\ 0, \text{if } l \neq \; 0, \end{cases}$$

and the $Daubechies\ scaling\ function\ \varphi$, given by

$$\varphi(x) \; = \; \sqrt{2}\sum_{k=0}^{2N-1} h_k\varphi(2x-k)$$

and

$$\varphi(x) \; = \; 0 \;\text{ for } x \in \mathbb{R} - \{0, 2N-1\}$$

as well as

$$\int_{\mathbb{R}} \varphi(2x-k)\varphi(2x-l)dx \; = \; 0 \text{ for } k \neq l.$$

In notation $\psi_{j,k}(x) = 2^{\frac{j}{2}}\psi(2^j x - k)$ and $\widetilde{\psi}_{j,k}(x) = 2^{\frac{j}{2}}\widetilde{\psi}(2^j x - k)$ for $j, k \in \mathbb{Z}$, $(\psi_{j,k})$ and $\widetilde{\psi}_{j,k}$ are Riesz bases dual to each other. In particular, a function $f \in L^2(\mathbb{R})$ can be represented both as

$$f(x) = \sum_{j,k\in\mathbb{Z}} < f, \widetilde{\psi}_{j,k} > \psi_{j,k}(x) \text{ and} \sum_{j,k\in\mathbb{Z}} < f, \psi_{j,k} > \widetilde{\psi}_{j,k}(x).$$

Similar with Haar wavelet transform we can compute the Daubechies wavelet transform by the simple algorithm as Haar wavelet transform.

# CHAPTER 3

# ARIMA Model

In this section we will introduce the basic knowledge of autoregressive integrated moving average (ARIMA) model, which is a best model to approximate to make the forecasting in a short times and get the minimum value of error less than another method. Generally, the initial selection of an ARIMA model is based on an examination of a plot of times series, so as before construct a model, we try to understanding of time series as follows.

## 3.1 Time Series

A time series is a sequence of numerical data points in successive order. A time series tracks the movement of the chosen data points, such as a securitys price, over a specified period of time with data points recorded at regular intervals. There is no minimum or maximum amount of time that must be included, allowing the data to be gathered in a way that provides the information being sought by the investor or analyst examining the activity. We can observe the time series a finite number of times, for example a sequence of random variables $(X_1,\ X_2, ..., X_n)$ is just an $n-$ dimensional random variable which called a $stochastic\ process$ [8].

**Definition 3.1.** A time series model for the observed data $\{x_t\}$ is a specification of the possible only the means and covariances of a sequence of random variables $\{X_t\}$ of which $\{x_t\}$ is hypothesized to be a realization.

A time series is a set of observation $x_t$, each one being record at a specific time $t$. For example, the time series of global mean temperature anomaly from $1850$ until $2010$. and Dow Jones Industrial Average $(DJI)$ inflation adjusted price from $1999$ until $2013$, see figure $3.1$ and $3.2$ as following.

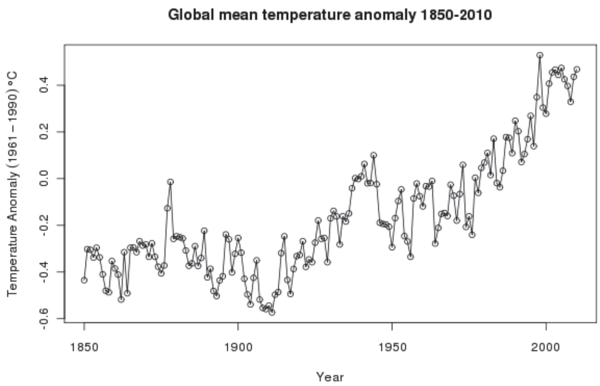


Figure 3.1: The data of global mean temperature anomaly $1850 - 2010$.

Site: http://www.fromthebottomoftheheap.net/2011/06/11/global-warming-since-1995-now-significant

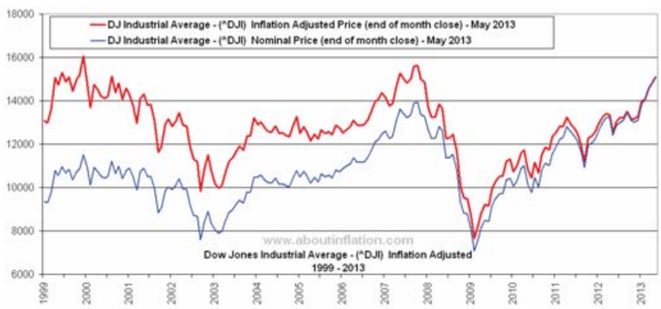


Figure 3.2: Dow Jones Industrial Average $(DJI)$ inflation adjusted price.

Site:http://www.aboutinflation.com/inflation-adjusted-charts/us-index-sectors-inflation-adjusted-charts/dow-jones-industrial-average-inflation-adjusted-chart

Moreover, the primary objective of time series analysis is to develop mathematical models that provide possible descriptions for sample data such as white noise, random walk with drift, autoregressive and others. Now we will shown some example of time series model in example $1$ and $2$ as following.

**Example 1.** (White noise process)

A sample kind of generated series might be a collection of uncorrelated random variables, $w_t$, with mean $0$ and finite variance $\sigma_w^2$. The time series generated from uncorrelated variables is used a model for noise engineering applications where it is called $white\ noise$, we shall some denoted the process as $w_t \sim w_n(0, \sigma_w^2)$.(See figure 3.3)

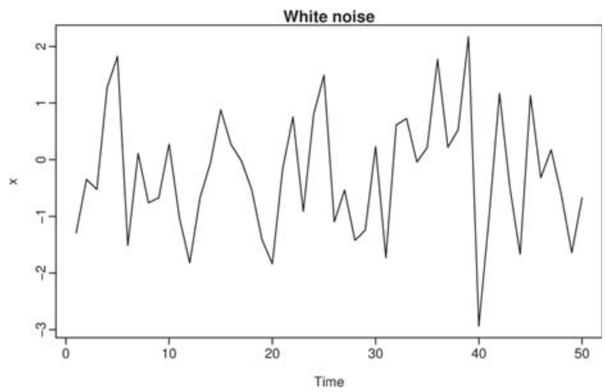


Figure 3.3: White niose.

Site:https://www.otexts.org/fpp/2/2

We often require stronger conditions and need the noise to be independent and identically distributed(iid) [9] random variables with mean $0$ and variance $\sigma_w^2$. We will discriminate this by saying white independent noise, or by writing $w_t \sim iid(0, \sigma_w^2)$.(See figure $3.4$.)

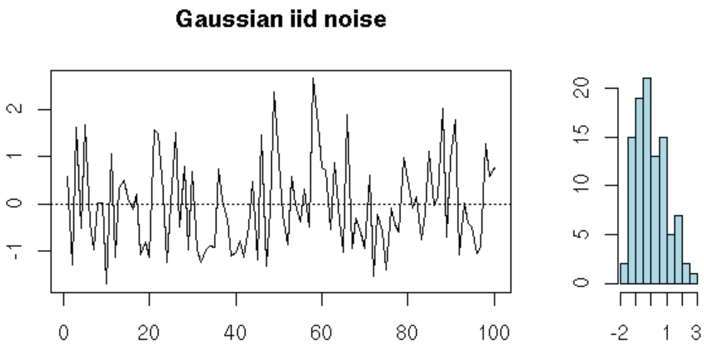


Figure 3.4: IID white niose.

Site:http://zoonek2.free.fr/UNIX/48

**Example 2.** (Random walk with drift process)

A model for analysing trend such as in the global temperature data in figure $3.5$, is the random walk with drift model given by

$$x_t = \delta + x_{t-1} + w_t \tag{3.1}$$

for

$t$ is equal $1,\ 2,\ 3, ...,$ with initial condition $x_0 = 0$,

$w_t$ is a white noise,

$\delta$ is the drift constant.


Figure 3.5: The model of a random walk with drift of USD-Euro exchange rate.

Site:https://people.duke.edu/ rnau/411rand.htm

In real life data are not stationary, we were varied the data for determined the exact model. So the assumptions of stationarity below apply after any trend or seasonal effect have been removed.

## 3.2 Stationary Time Series

We start to describe the general behavior of a process as the evolves over time, generally rely only on properties define by the means and covariance, we are led to the following definitions.

**Definition 3.2.** Let $\{X_t, \quad t \in T\}$ where $T$ is called the $index$ or parameter set be a stochastic process with Var($X_t \ < \infty$).

The mean function of $X_t$ is

$$\mu_x(t) \ = \ E(X_t), \text{ for } t \in T.$$

The covariance function of $X_t$ is

$$\gamma_x(r,s) \;=\; \mathrm{C}ov(X_r, X_s),\; r,s \in T.$$

**Definition 3.3.** The time series $\{X_t,\; t \in \mathbb{Z}\}$ is said to be $weakly$ stationary if

(i) $\mathrm{V}ar(X_t) \;<\; \infty$,

(ii) $\mu_x(t) \;=\; E(X_t),\; \forall t \in \mathbb{Z}$,

(iii) $\gamma_x(r,s) \;=\; \gamma_x(r+t,\; s+t)\;\; \forall r,s,t \in \mathbb{Z}$.

(iiii) implies that $\gamma_x(r,s)$ is a function of $r-s$, and it is convenient to define

$$\gamma_x(h) \;=\; \gamma_x(h,0).$$

The value "$h$" is referred to as the "$lag$".

**Definition 3.4.** Let $\{X_t,\; t \in \mathbb{Z}\}$ be a stationary time series. The autocovariance function (ACVF) of $\{X_t\}$ is

$$\gamma_x(h) \;=\; \mathrm{C}ov(X_{t+h}, X_t).$$

The autocorrelation function (ACF) is

$$\rho_x(h) \;=\; \frac{\gamma_x(h)}{\gamma_x(0)}.$$

## 3.3 Linear Process

We will now turn to an examination of a large class of useful time series model. These are almost all defined in term of the operator $L$. As the simplest example, consider the autoregressive model defined by

$$Y_t \;=\; \phi Y_{t-1} + \epsilon_t \tag{3.2}$$

where $\phi$ is a constant with $|\phi| < 1$ and $\epsilon_t$ is a sequence of uncorrelated random variables, each with mean $0$ and variance $\sigma^2$.

**Definition 3.5.** The time series $Y_t$ define by

$$Y_t \;=\; \sum_{u=-\infty}^{\infty} \psi_u \epsilon_{t-p},$$

where $\epsilon$ is a white noise series and

$$\sum_{u=-\infty}^{\infty} |\psi_u|^2 \;<\; \infty,$$

is call *general linear process*.

The general linear process depends on both past and future value of $\epsilon_t$ is said to be casual. Casual processes are perfected for forecasting because they reflect the way in which we believe the real world works.

Many time series can be represented as linear processes. This provides a unifying underpinning for time series theory, but may be of limited practical interest because of the potentially infinite number of parameters required.

## 3.4 Autoregressive Series

In an autoregressive model, we forecast the variable of interest using a linear combination of past value of variable. The term autoregressive indicated that it is a regression of the variable against itself.

**Definition 3.6.** A real value stochastic process $\{Y_t\}$ is said to be autoregressive of order $p$, denote by AR(p) if there exist $\phi_0, \phi_1, ..., \phi_p \in \mathbb{R}$ with $\phi_p \neq 0$ and a white noise $e_t$ such that

$$Y_t \;=\; \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2}, ..., +\phi_p Y_{t-p} + e_t,\; t \in \mathbb{Z}. \tag{3.3}$$

**Example 3.** (The AR$(1)$ Series)

The AR$(1)$ series is defined by

$$Y_t \;=\; \phi Y_{t-1} + \epsilon_t. \tag{3.4}$$

Because $Y_{t-1}$ and $\epsilon_t$ are uncorrelated, the variance of this series is

$$\text{Var}(Y_t) \;=\; \phi^2 \,\text{Var}(Y_{t-1}) + \sigma_\epsilon^2.$$

If $\{Y_t\}$ is stationary then $\text{Var}(Y_t) \;=\; \text{Var}(Y_{t-1}) \;=\; \sigma_Y^2$ and so

$$\sigma_Y^2 \;=\; \phi^2 \sigma_Y^2 + \sigma_\epsilon^2.$$

This implies that

$$\sigma_Y^2 \;>\; \phi^2 \sigma_Y^2$$

and hence

$$1 \;>\; \phi^2.$$

## 3.5 Moving Averages

**Definition 3.7.** Let $\theta_0, \theta_1, ..., \theta_p \in \mathbb{R}$ with $\theta_q \neq 0$ and white noise $\{\epsilon_t\}$.

The process

$$Y_t \;=\; \mu - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2}, ..., -\theta_p \epsilon_{t-q} + \epsilon_t, \tag{3.5}$$

is said to be moving averages of order $q$, denoted by MA(q).

No additional condition are require to ensure stationary.

The autocovariance function for MA(q) process is

$$\gamma(u) = \begin{cases} 1+\theta_1^2+...+\theta_q^2; & u \ = \ 0 \\ \theta_u+\theta_1\theta_{u+1}...+\theta_{q-u}\theta_q^2; & u \ = \ 1 \\ 0 & \text{otherwise} \end{cases}$$

which say there is only a finite span of dependence on the series.

It is easy to distinguish MA and AR series by the behavior of their autocorrelation function. The acf for MA series *cuts off* sharply while that for an AR series dacays exponentially (with a possible sinusoidal ripple superimposed).

**Example 4.** (The M$A(1)$ Series)

The M$A(1)$ is defined by

$$Y_t \ = \ \epsilon_t + \theta\epsilon_{t-1}. \tag{3.6}$$

It has the autocovariance function

$$\gamma(u) = \begin{cases} (1+\theta^2)\sigma^2; & u \ = \ 0 \\ \theta\sigma^2; & u \ = \ 1 \\ 0 & \text{otherwise} \end{cases}$$

and the autocorrelation

$$\gamma(u) = \begin{cases} \frac{\theta}{1+\theta^2}; & u \ = \ 1 \\ 0 & \text{otherwise.} \end{cases}$$

## 3.6 Autoregressive Moving Averages

A model with autoregressive terms combined with a model having moving average term to get a mixed autoregressive-moving average model.

**Definition 3.8.** If a series satisfies

$$\phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2}, ..., +\phi_p Y_{t-p} - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2}, ..., -\theta_q \epsilon_{t-q} + \epsilon_t \tag{3.7}$$

it convenient to use the notation ARMA$(p, q)$, where $p$ is the order of the autoregressive part and $q$ is the order of moving average part, to represent these models and $\epsilon_t$ is a white noise.

**Example 5.** (The ARMA$(1, 1)$ Series)

The A$RMA(1, 1)$ series is defined by

$$Y_t \;=\; \phi Y_{t-1} + \epsilon_t + \theta \epsilon_{t-1}. \tag{3.8}$$

The derive the autocovariance function for $Y_t$, note that

$$E(\epsilon_t Y_t) \;=\; E[\epsilon_t(\phi Y_{t-1} + \epsilon_t + \theta\epsilon_{t-1})] = \sigma^2_\epsilon$$

and

$$E(\epsilon_t Y_{t-1}) \;=\; E[\epsilon_t(\phi Y_{t-1} + \epsilon_t + \theta\epsilon_{t-1})] = (\phi + \theta)\sigma^2_\epsilon.$$

Multiplying equation $3.8$ by $Y_{t-u}$ and taking expectation yield

$$\gamma(u) = \begin{cases} \phi\gamma(1) + (1 + \theta(\phi + \theta))\sigma^2_\epsilon \;; & u \;=\; 0 \\ \phi\gamma(0)\theta\sigma^2_\epsilon \;; & u \;=\; 1 \\ \phi\gamma(u-1) \;; & u \;\geq\; 2. \end{cases}$$

Solving the two equations produces

$$\gamma(0) = \frac{(1 + 2\theta\phi + \theta^2)}{1 - \phi^2}\gamma_\epsilon^2$$

and using the last recursive shows

$$\gamma(u) = \frac{(1 + \theta\phi)(\phi + \theta)}{1 - \phi^2}\phi^{u-1}\gamma_\epsilon^2 \qquad \text{for } u \geq 1.$$

The autocorrelation function can then be composed as

$$\rho(u) = \frac{(1 + \theta\phi)(\phi + \theta)}{1 - \phi^2}\phi^{u-1} \qquad \text{for } u \geq 1.$$

The pattern here is similar to that for AR$(1)$, except for the first term.

## 3.7 The Partial Autocorrelation Function

The autocorrelation function of an MA series exhibits different behavior from that of AR and general ARMA series. The acf of an MA series cuts of sharply whereas those for AR and ARMA series exhibit exponential decay (with possible sinusoidal behavior superimposed). This make it possible to identify an ARMA series as being a purely MA one just by plotting its autocorrelation function. The partial autocorrelation function provides a similar way of identifying a series as a purely AR one. Given a stretch of time series values

$$..., Y_{t-u}, Y_{t-u+1}, ..., Y_{t-u}, Y_t, ...$$

the partial correlation of $Y_t$ and $Y_{t-u}$ is the correlation between these random variables which is not conveyed through the intervening values. If the $Y$ values are normally distributed, the partial autocorrelation between $Y_t$ and $Y_{t-u}$ can be defined as

$$\phi(u) \;=\; \text{cor}(Y_t, Y_{t-u}|Y_{t-1}, ..., Y_{t-u+1}).$$

A more general approach is based on regression theory. Consider predicting $Y_t$ based on $Y_{t-1}, ..., Y_{t-u+1}$. The prediction is

$$\widehat{Y} \;=\; \beta_1 Y_{t-+1} + \beta_2 Y_{t-2} + ... + \beta_{u-1} Y_{t-u+1}$$

with the $\beta$ chosen to minimise

$$E(Y_t - \widehat{Y}_t)^2.$$

It is also possible to *think backwards in time* and consider predicting $Y_{t-u}$ with the same set of the predictor will be

$$\widehat{y}_{i-u} \;=\; \beta_1 Y_{t-u+1} + \beta_2 Y_{t-u+2} + ... + \beta_{u-1} Y_{t-1}.$$

(The coefficients are the same because the correlation structure is the same whether the series is run forwards or backwards in time). The partial correlation function at lag $u$ is the correlation between the prediction errors.

$$\phi(u) \;=\; \text{cor}(Y_t - \widehat{Y}_t, Y_{t-u} - \widehat{Y}_{t-u})$$

By convention we take $\phi(1) \;=\; \rho(1)$. It is quite straightforward to compute the value of $\phi(2)$. Using the results of of linear prediction [9], the best predictor of $Y_t$ based on $Y_{t-1}$ is just $\rho(1)Y_{t-1}$. Thus

$$\text{C}ov(Y_t - \rho(1)Y - t - 1, Y_{t-2} - \rho(1)Y_{t-1}) \;=\; \sigma_Y^2(\rho 2 - \rho(1)^2)$$

and

$$\text{V}ar(Y_t - \rho(1)Y_{t-1}) \;=\; \sigma_Y^2(1 - \rho(1))^2.$$

This mean that

$$\phi_2 \;=\; \frac{\rho(2) - \rho(1)^2}{1 - \rho(1)^2}. \tag{3.9}$$

**Example 6.** For the AR$(1)$ series, recall that

$$\rho(u) \;=\; \phi^u \qquad (u \;\geq\; 0).$$

Substituting this into equation $3.9$ we found that

$$\phi(2) \;=\; \frac{\phi^2 - \phi^2}{1 - \phi^2} \;=\; 0.$$

## 3.8 Integrated Models

In (example $3.2$), we saw that if $x_t$ is a random walk $x_t \;=\; x_{t-1} + w_t$, then by differencing $x_t$, we find that $\nabla x_t \;=\; w_t$ is stationary. In my situation, time series can be though of as being composed of two component, a nonstationary trend component. For example $3$, we considered the model

$$x_t \;=\; \mu \;+\; y_t \tag{3.10}$$

where $\mu_t \;=\; \beta_0 + \beta_1 t$ and $y_t$ is stationary. Differencing such a process will lead to a stationary process:

$$\nabla x_t \;=\; x_t - x_{t-1} \;=\; \beta_1 + y_1 - y_{t-1} \;=\; \beta_1 + \nabla y_t.$$

Another model that leads to first differencing is the cae in which $\mu_t$ in $(3.10)$ is a stochastic process and slowly varying according to a random walk. That is

$$\mu_t \;=\; \mu_{t-1} + \upsilon_t$$

where $\upsilon_t$ is stationary. In this case,

$$\nabla x_t \; = \; \nu_t + \nabla y_t,$$

is stationary. If $\mu_t$ in $(3.10)$ is quadratic, $\mu_t \; = \; \beta_0+\beta_1 t+\beta_2 t^2$, then the differenced series $\nabla^2 y_t$ is stationary. Stochastic trend model can also lead to higher order differencing. For example, suppose

$$\mu_t \; = \; \mu_{t-1} + \nu_t \;\; and \;\; \nu_t \; = \; \nu_{t-1} + e_t,$$

where $e_t$ is stationary. Then, $\nabla x_t \; = \; \nabla y_t$ is not stationary, but

$$\nabla^2 x_t \; = \; e_t + \nabla^2 y_t$$

is stationary.

The ARIMA model is broadening of the class of ARMA model to include differencing. The basic idea is that if differencing the data at some order $d$ process an ARMA process, then the original process is said to be ARIMA.

## 3.9 Autoregressive Integrated Moving Average

An autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model which stationary condition. These models are fitted to time series data either to better understand the data or to predict future points in the series (forecasting). They are applied in some cases where data show evidence of non-stationary, where an initial inferencing step (corresponding to the ”integrated” part of the model) can be applied to reduce the non-stationary.

**Definition 3.9.** Suppose that the time series $Y_t$ has a polynomial trend of degree $d$. Then we can eliminate this trend by considering the process ($\triangle^d Y_t$), obtain by $d$ times differencing as described in linear filtering of time series [2].If the filter process ($\triangle^d Y_t$) is an ARMA$(p, q)-$ process satisfying the stationary condition of unit circle that is

$$1 - a_1 z - a_2 z^2 - ... - a_p z^p \neq 0, \;\; for\, |z| \leq 1.$$

The original process $Y_t$ is said to be Autoregressive Integrated Moving average of order p, q, d, denoted by ARIMA$(p, d, q)$. In this case constant $\theta_0, \theta_1, ..., \theta_p \; and \; \phi_0, \phi_1, ..., \phi_p \in$
$\mathbb{R}$

exist such that

$$\triangle^d Y_t \; = \; \sum_{u=1}^{p} \theta_u \triangle^d Y_{t-u} \; + \; \sum_{w=0}^{p} \phi_w \epsilon_{t-w},$$

where $\{\epsilon_t\}$ is a white noise.

## 3.10 ARIMA model Analysis

In time series analysis, we can divide four steps to processed with ARIMA model.

### 3.10.1 Model identification

We first to checked stationary of time series data with criteria of statistics such as mean, variance and autocorrelation function are always stationary or we can checked stationary with looking at the time plot of the data, the ACF plot is also useful for identifying non-stationary time series. For a stationary time series, the ACF will drop to zero relatively quickly, while the ACF of non-stationary data decreases slowly.

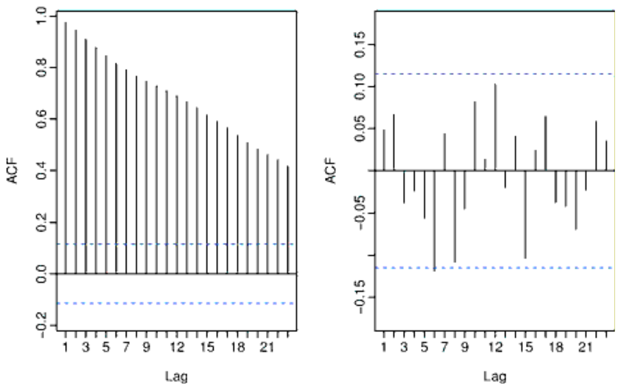


Figure 3.6: ACF stationary.

Site:https://www.otexts.org/fpp/8/1

The pattern of ACF and PACF are an exponentially decreases to zero in a quickly times. We discuss the classical ARIMA style analysis based on Autocorrelation function and partial autocorrelation function as well as model selection via information criteria.

ACF and PACF are determined the ordered of ARIMA model, if ACF that dies out gradually and PACF that cut off sharply after a few lags, we will get ordered of AR, and if PACF that dies out gradually and ACF that cut off sharply after a few lags, we will get ordered of MA.

### 3.10.2 Estimated parameters

After we selected a model, parameters for that must be estimated. The parameter in ARIMA models are estimated by minimizing the sum of the fitting errors.

In addition, the residual mean square error, an estimate of the variance of the error $\epsilon_t$ is computed.

The residual mean square error is defined as

$$s^2 = \frac{\sum_{t=1}^{n} \epsilon_t^2}{n-r} = \frac{\sum_{t=1}^{n}(Y_t - \widehat{Y}_t)^2}{n-r}$$

where $\epsilon_t$ is the residual at time $t$,

$n$ is the number of residuals,

$r$ is the total number of parameters estimated.

The residual mean square error is useful for assessing fit and comparing different models. It is also used to calculate forecast error limits.

### 3.10.3 Model Checking

Before using the model for forecasting, it must be checked. Normally, a model is enough if the residuals cannot be used to improve the forecasts, that is the residuals should be random.

1. Many of the same residual plots that are useful in regression analysis can be developed for the residuals from an ARIMA model. A histogram and a normal probability plot and time sequence plot are particularly helpful.

2. The individual autocorrelations $r_k(\epsilon_t)$ should be small and generally within $\pm\frac{2}{\sqrt{n}}$ of zero.

3. The residual autocorrelations as a group should be consistent with those produce by random error.

An overall check of model enough is provided by chi-square $(\chi^2)$ test based on the Ljung-Box $Q$ statistics. This test tooks at the sizes of the residuals autocorrelations as a group. The test statistic $Q$ is

$$Q_m = n(n+2)\sum_{k=1}^{m}\frac{r_k^2(\epsilon)}{n-k}$$

where

$r_k(\epsilon)$ is the residual autocorrelation at lag $k$.

$n$ is the number of residuals.

$m$ is the number of time lags unclouded in the test.

## 3.11 Forecasting

1. Once an enough model has been found, forecast for one period or sevaral period into the future can be made.

2. As more data become available, the same ARIMA model can be used the generated revised forecasts from another time origin.

3. If the character of the series appear to be changing overtime, the new data can be used reestimate the model parameter. For easily to understanding, we will summarized all step of ARIMA model by diagram as following.

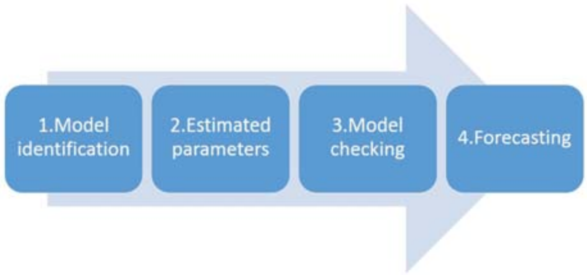


Figure 3.7: Step of ARIMA model.

# CHAPTER 4

## Results and Discussion

In this chapter, we present the results of all the data which transformed by discrete wavelet transform and ARIMA model. We first to transform the data with discrete wavelet transform via Haar and Daubechies. Next, we continue to approximate to make the forecasting and compare with the minimum value of MAE and RMSE. We determine the behavior of the data which transformed by discrete wavelet transform and then use R program to determined the behavior of the data which ARIMA model.

The last six years ago, Songkhla is destroyed by flooding. We interested to study and predict the rainfall in Songkhla, we can keep up with flooding in the future. Consider daily rainfall in Songkhla, $2007 - 2012$ (see figure 4.1-4.6).

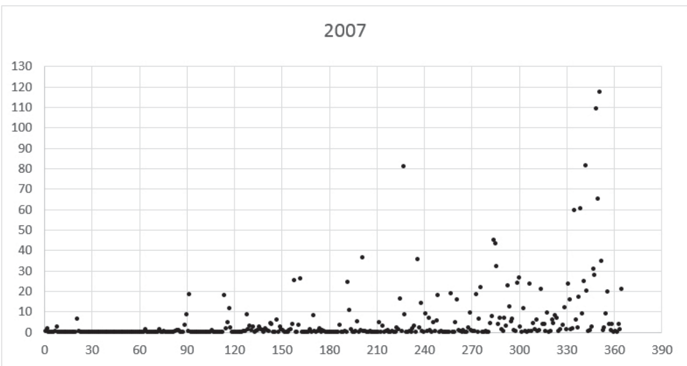


Figure 4.1: Daily rainfall in Songkhla, 2007.

For figure 4.1, we have found that the rain has increased and most rain at the end of year.

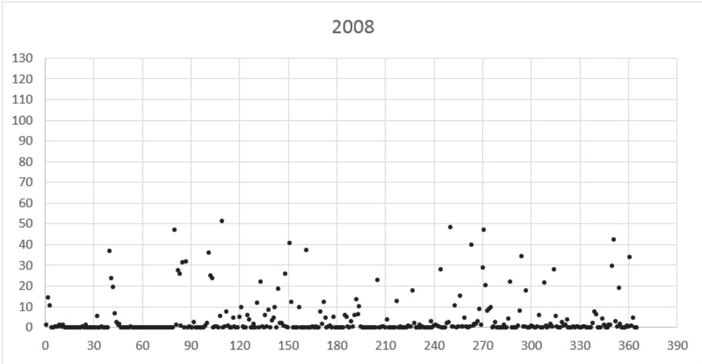


Figure 4.2: Daily rainfall in Songkhla, 2008.

For figure $4.2$, we can see that there's lots of rain all year and most rain on day $100^{th}$ to $180^{th}$.

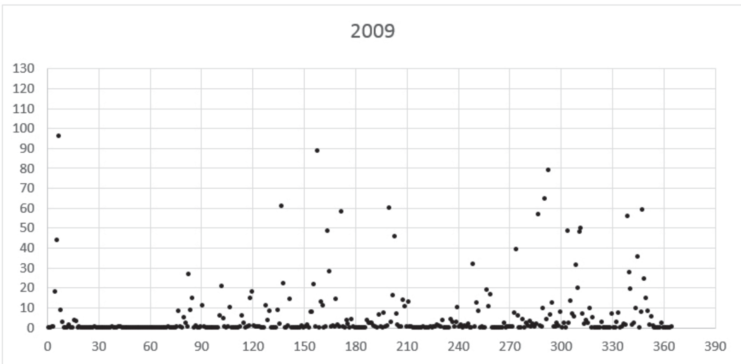


Figure 4.3: Daily rainfall in Songkhla, 2009.

For figure $4.3$, we can see that no rain at early year and most rain after day $80^{th}$ up to end of year.

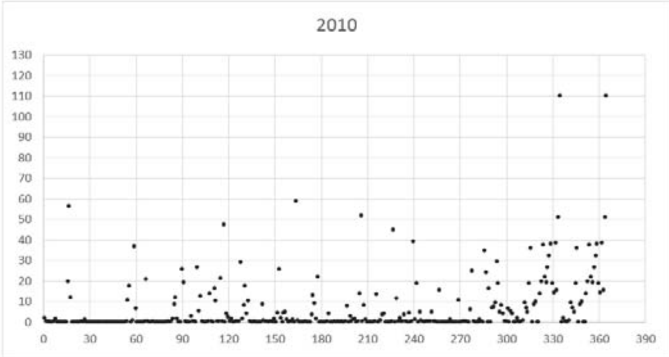


Figure 4.4: Daily rainfall in Songkhla, 2010.

For figure 4.4, we can found that almost rain everyday and most rain at the end of year.

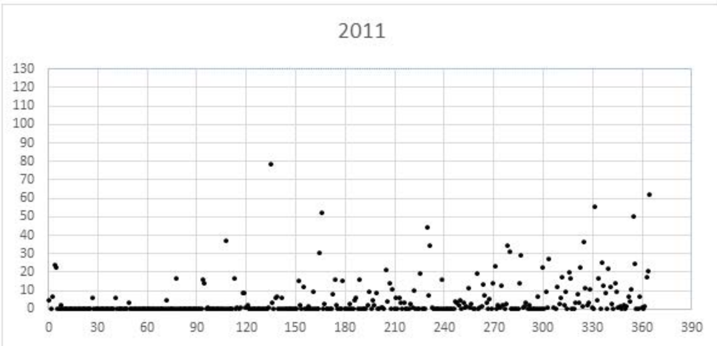


Figure 4.5: Daily rainfall in Songkhla, 2011.

For figure 4.5, we can found that there's a lot of rain from day $150^{th}$ until the end of year.

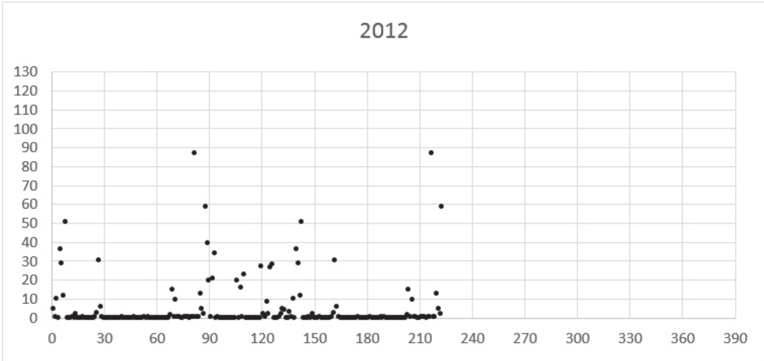


Figure 4.6: Daily rainfall in Songkhla, 2012.

For figure $4.6$, we can found that the rain occur sometimes.

On the figure $4.1 - 4.6$, we will describe the behavior of the data of rainfall in Songkhla area from $2007 - 2012$. Mostly, The first period of year we found that a few fall rain and the end of year we found that a lot of rain. Notice that an average increased so that the series may not stationary.

We consider three step to analyzed the data. Firstly, the daily rainfall data is transformed by Haar and Daubechies discrete wavelet and continue to approximate to make the forecasting, compare in ARIMA model and to show that result in the last. we can see step of analysis on the figure as below to understanding step as the following.

## 4.1 Step of Discrete Wavelet Transform Analysis

Figure $4.1 - 4.6$ show the daily rainfall in Songkhla, $2007 - 2012$, there are a few rainfall the first period, a lot of rainfall in the final period for each years. Therefore the average of rainfall is increasing so that the time series may not stationary.

Consequently, we transform daily rainfall data by the Haar wavelet trans-


Figure 4.7: Step of analyze the data of rainfall.

form and Daubechies wavelet transform as following scatter plot by the algorithm as shown in 2.1.1 (figure$4.8 - 4.19$)

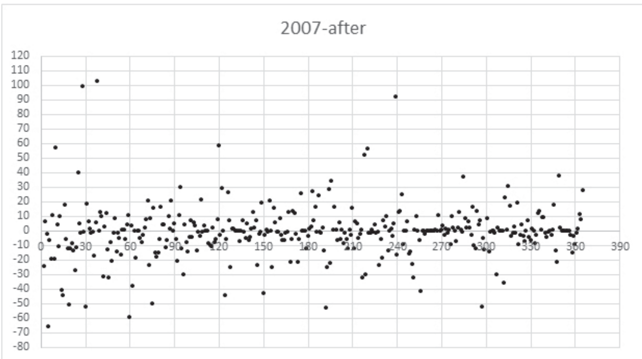


Figure 4.8: Haar wavelet transformed rainfall data, 2007.

For figure 4.8, we can found that the distribution of rainfall data is spreading almost difference period and an average do not change over time. Therefore, the data may be stationary.

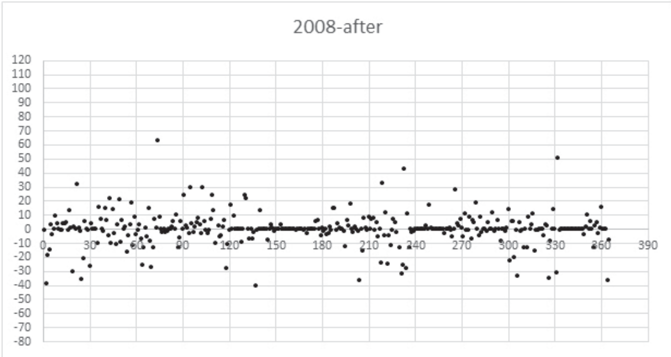


Figure 4.9: Haar wavelet transformed rainfall data, 2008.

For figure 4.9, we can found that the distribution of rainfall data is spreading and some period the data is equal zero. Hence mean equal zero. Therefore, the data may be stationary.

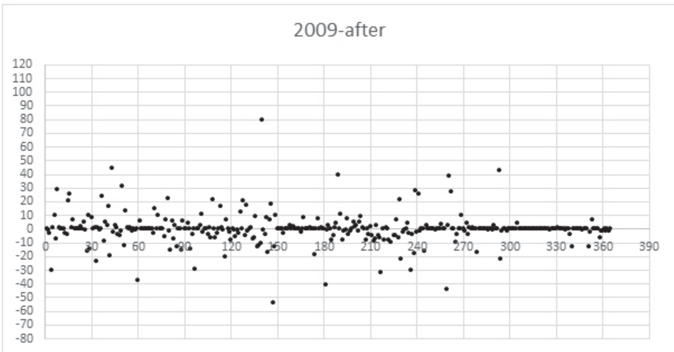


Figure 4.10: Haar wavelet transformed rainfall data, 2009.

For figure 4.10, we can found that the rainfall data is equal zero and mean equal zero. Hence, the data may stationary.

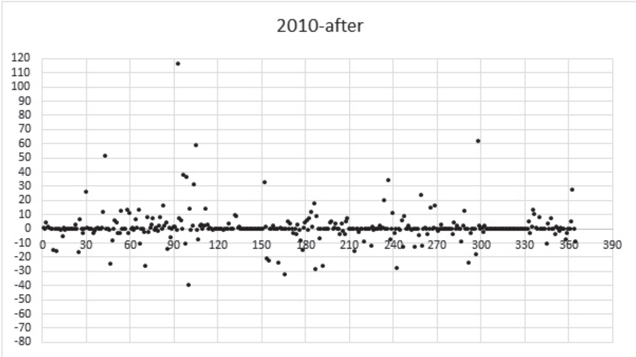


Figure 4.11: Haar wavelet transformed rainfall data, 2010.

For figure 4.11, we can found that some period the rainfall data is spreading. For some period the data is equal zero. Mean equal zero. Therefore, the data may be stationary.

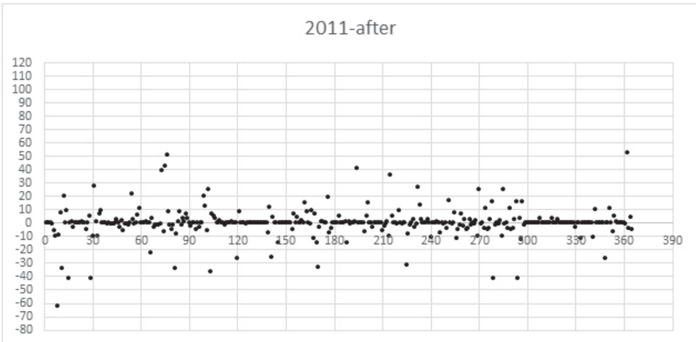


Figure 4.12: Haar wavelet transformed rainfall data, 2011.

For figure 4.12, we can see that the rainfall data is equal zero, mean equal zero. Hence, the data may stationary.

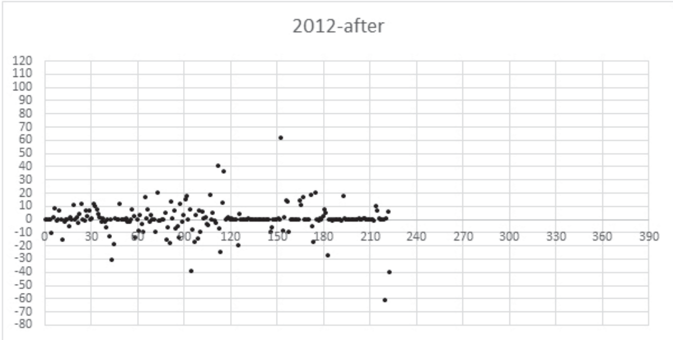


Figure 4.13: Haar wavelet transformed rainfall data, 2012.

For figure $4.13$, we can see that for some period the rainfall data is spreading. After day $120^{th}$ the data is zero. Hence, the data may be stationary.

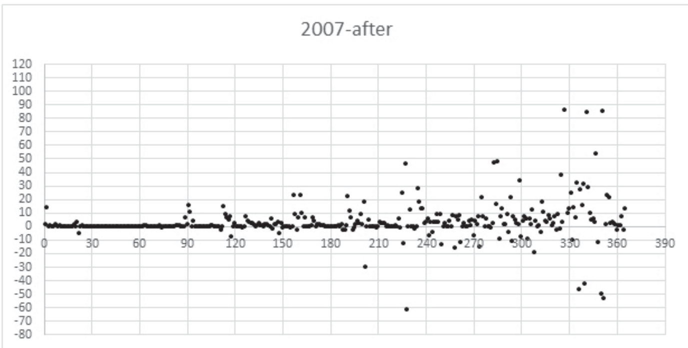


Figure 4.14: Daubechies wavelet transformed rainfall data, 2007.

For figure $4.14$, we can see that for early years the rainfall data in $2007$ is most equal zero. After day $120^{th}$ the data is almost spreading. Mean is equal zero. Hence, the data may be stationary.

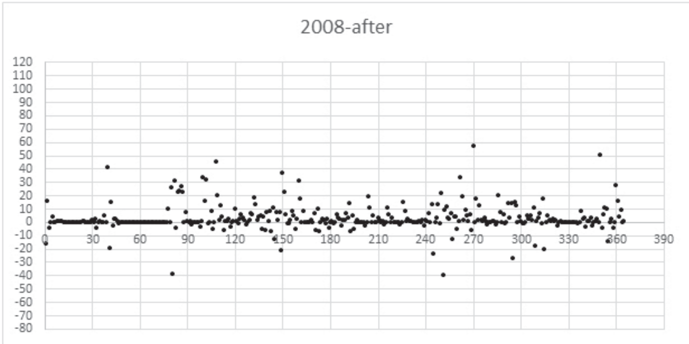


Figure 4.15: Daubechies wavelet transformed rainfall data, 2008.

For figure $4.15$, we can see that After day $120^{th}$ the data in $2008$ is most spreading. Mean is equal zero. So that the data may be stationary.

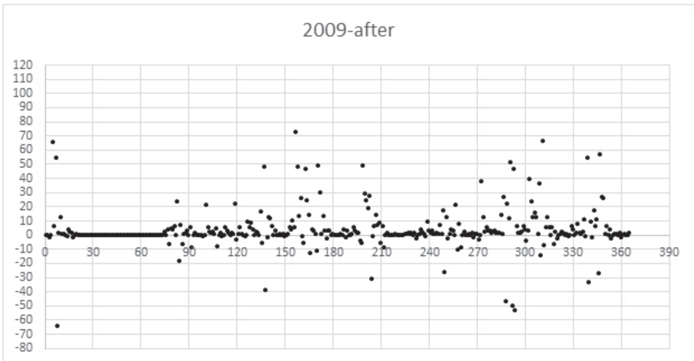


Figure 4.16: Daubechies wavelet transformed rainfall data, 2009.

For figure $4.16$, we can see that After day $80^{th}$ the distribution of data in $2009$ is alternating spread and equal zero until the end of the year. Mean is equal zero. Therefore, data may be stationary.

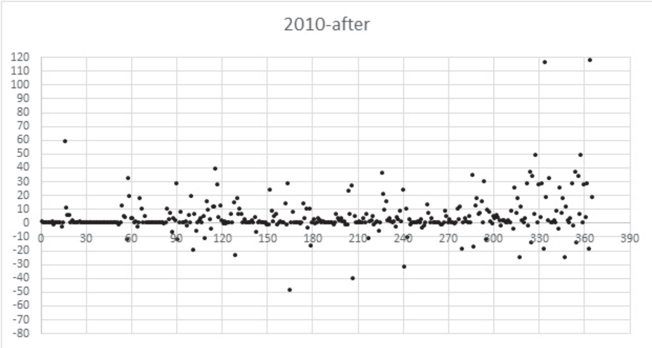


Figure 4.17: Daubechies wavelet transformed rainfall data, 2010.

For figure $4.17$, we can see that the distribution of data in $2010$ is almost near zero and most equal zero. Spreading after day $300^{th}$ until the end of the year.Mean is equal zero. Therefore, data may be stationary.

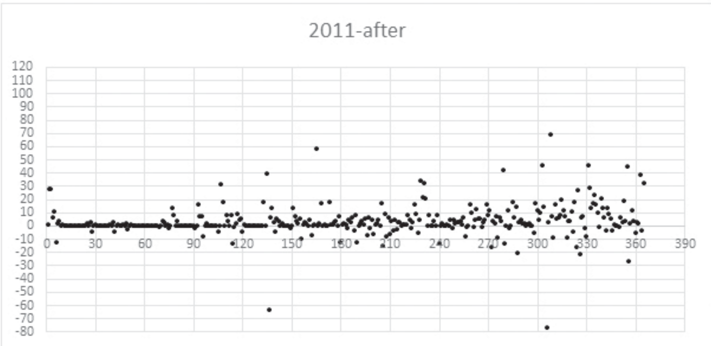


Figure 4.18: Daubechies wavelet transformed rainfall data, 2011.

For figure 4.18, We can see that the distribution of data in 2011 is most spreading after day $250^{th}$ until the end of the year. Mean is equal zero. Therefore, data may be stationary.

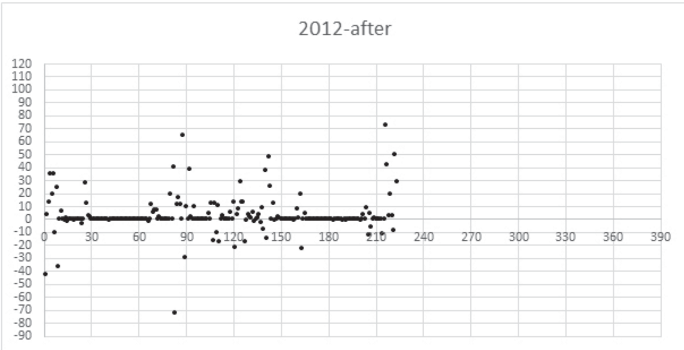


Figure 4.19: Daubechies wavelet transformed rainfall data, 2012.

For figure 4.19, We can see that the distribution of data in 2012 is spreading and equal zero at some period. Mean is equal zero. Therefore, data may stationary.

On the figure 4.8-4.19, we can see that, after we transform the natural of rainfall data has increased, decreased and refuted. The Haar and Daubechies wavelet transformed much different and their average nearly to equal zero. Therefore, the data may stationary.

## 4.2 Step of ARIMA Model Analysis

In this section, we try to find the fitted model to make the forecasting in the future. In ARIMA model we can divided three steps to find the best model.

### 4.2.1 Model Identification

In this section, we try to find the fitted ARIMA models. We known that if the graph of autocorrelation function of constancy averages time series is die out rapidly and cut off, then that time series is stationary. Autocorrelation function die out rapidly and cut off (see figure 4.20 and 4.22). Therefore the data after we transform are stationary.

Next step of fitted ARIMA model is determining the numbers of AR and/or MA terms by considered the autocorrelation function(ACF) and partial autocorrelation function(PACF) plots of different series.

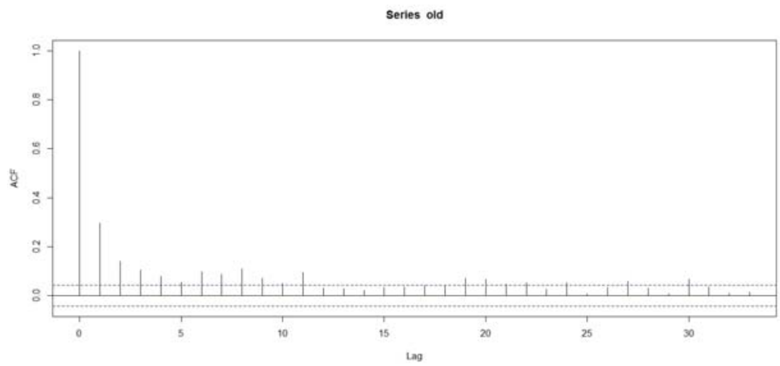


Figure 4.20: ACF of raw rainfall data.

For figure 4.20, the ACF die out. it is not clear what the attributes of model.

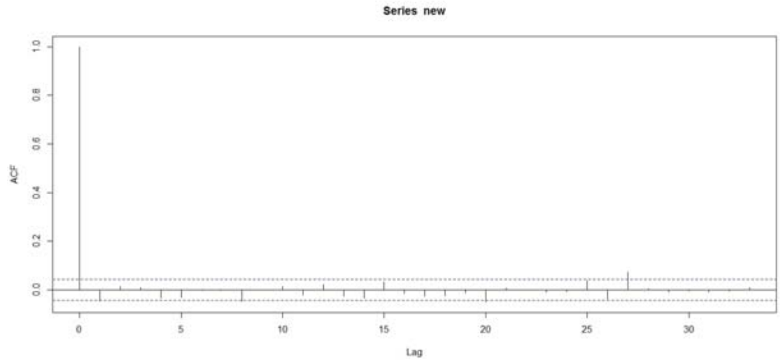


Figure 4.21: ACF of haar discrete wavelet data.

For figure $4.21$, the ACF cut off in $1^{st}$ lag. We get AR(1) model.

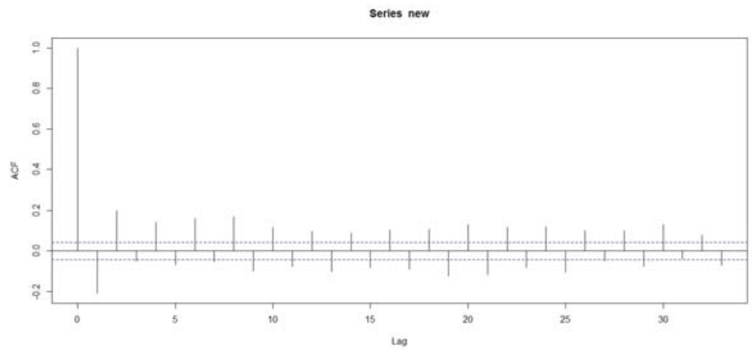


Figure 4.22: ACF of daubechies discrete wavelet data.

For figure $4.22$, the ACF cut off in $2^st$ lag. We get AR(2) model.

### 4.2.2 Model Estimation

After a time series has been stationeries, the next step in fitting an ARIMA model is to determine whether AR or MA terms are needed to correct any autocorrelation that remains in the differences series. By looking at the autocorrelation function (ACF) and partial autocorrelation (PACF) plots of the difference series, you can tentatively identify the numbers of AR and/or MA terms that are needed. You are already familiar with the ACF plot: it is merely a bar chart of the coefficients of correlation between a time series and lags of itself. The PACF plot is a plot of the partial correlation coefficients between the series and lags of itself.

After we plot ACF, PACF we get ARIMA$(1,1,1)$ for the rainfall data. And then for once a tentative model, the parameters for that model must be estimated. Now, we get the parameter of ARIMA$(1,1,1)$ as belows.

$$\text{ar1} \;=\; -\,0.3878,\;\; \text{ma1} \;=\; -\,1.0000 \text{ and a constant of the model} \;=\; 0.0000.$$

Therefore the fitted equation of ARIMA$(1,1,1)$ is

$$\widehat{Y}_t \;=\; -\,0.3878Y_{t-1} - 1.0000\epsilon_{t-1}.$$

After we plot ACF, PACF we get ARIMA$(1,0,1)$ for Haar discrete wavelet transform. For once a tentative model, the parameters for that model must be estimated. Now, we get the parameter of ARIMA$(1,0,1)$ as belows.

$$\text{ar1} = -0.1888, \text{ ma1} = 0.1355 \text{ and a constant of the model} = 0.1109$$

Therefore the fitted equation of ARIMA$(1, 0, 1)$ is

$$\widehat{Y_t} = 0.1128 - 0.1888Y_{t-1} + 0.1355\epsilon_{t-1}.$$

For Daubechies discrete wavelet transform we get we get ARIMA$(2, 0, 2)$ and the parameter as follows

$$\text{ar1} = -0.0250, \text{ ar2} = 0.9424, \text{ ma1} = -0.0041, \text{ ma2} = -0.8614 \text{ and}$$

a constant of the model $= 3.4994$. Therefore the fitted equation of ARIMA$(2, 0, 2)$ is

$$\widehat{Y_t} = 3.4478 - 0.025Y_{t-1} + 0.9420Y_{t-2} - 0.0041\epsilon_{t-1} - 0.8614\epsilon_{t-1}.$$

### 4.2.3 Model Checking

Finally, before using the model we will checked for adequacy that is the ACF plot of the residuals from the ARIMA(p,d,q) model shows all correlations within the threshold limits indicating that the residuals are behaving like white noise. A portmanteau test returns a large p-value, also suggesting the residuals are white noise. We will used Boxcox test to checked ACF and printed p-value of residuals. we can see figure as below.

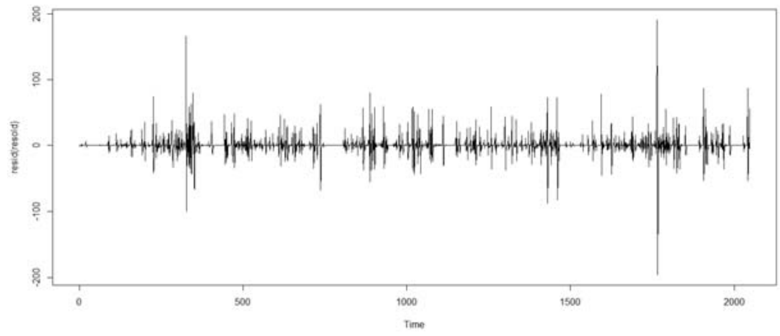


Figure 4.23: Residual of ARIMA$(1, 1, 1)$.

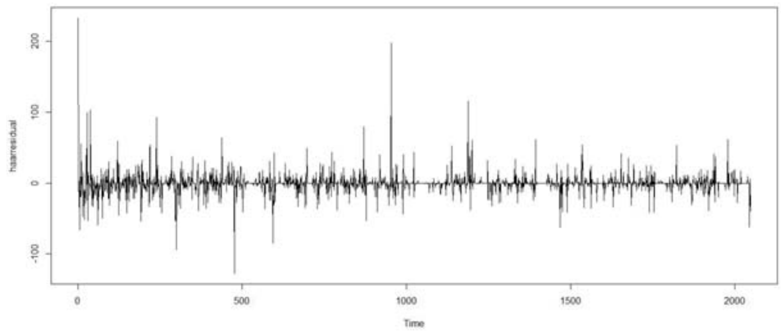


Figure 4.24: Residual of ARIMA$(1, 0, 1)$

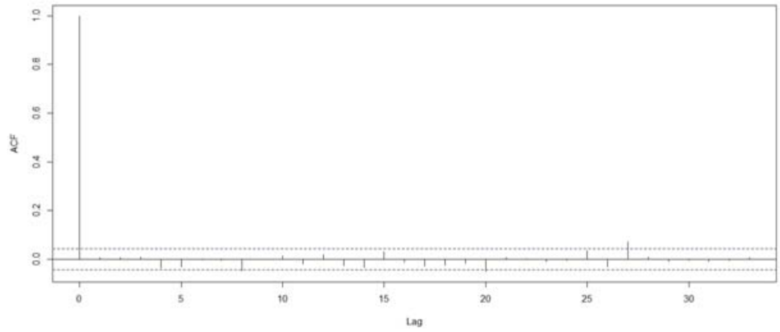


Figure 4.25: ACF of residual of ARIMA$(1, 0, 1)$

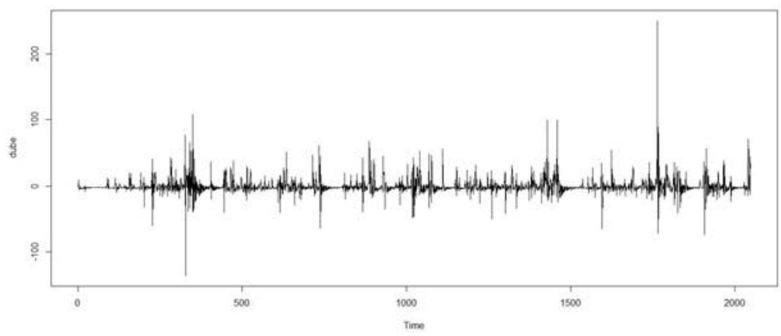


Figure 4.26: Residual of ARIMA(2,0,2)

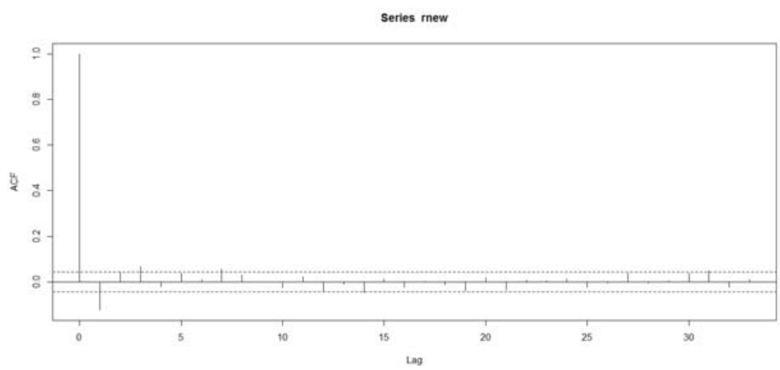


Figure 4.27: ACF of residual of ARIMA$(2, 0, 2)$

From figure 4.23, 4.24 and 4.26 are difficult to clearly describe feature. Now, we use the ACF to describe feature of residual of ARIMA(1,0,1) and ARIMA(2,0,2) models. Figure 4.25 and 4.27,the distribution of the rainfall data are random so that the residuals like white noise and the p-value greater than 0.05. Moreover, by the Shapiro wilk test, the residual of that models are normal so that the models is considered adequate. Hence this model adequacy. Now, we can use the fitted models of ARIMA(1,0,1) and ARIMA(2,0,2) to make a forecast in the next step.

### 4.2.4 Analysis result

In this section, we show the value of MAE and RMSE of ARIMA(1,0,1), ARIMA(2,0,2),which are fitted from the Haar wavelet transformed data,
the Daubechies wavelet transformed data and original ARIMA Songkla rainfall data respectively.

| Statistic fit | ARIMA(1,0,1) of Haar wavelet transformed data | ARIMA(2,0,2) of Daubechies wavelet transformed data | ARIMA(1,1,1) of raw rainfall data |
|---|---|---|---|
| Mean absolute error (MAE) | 6.743518 | 6.7349 | 6.843506 |
| Root mean square error (RMSE) | 14.91782 | 13.80871 | 15.28966 |

Figure 4.28: The statistical criteria for the ARIMA model

We can see that the values of RMSE for ARIMA(1,0,1), ARIMA(2,0,2), ARIMA(1,1,1) are $14.91782$, $13.80871$ and $15.28966$, respectively. The values of MAE for the three model are $0.009684946$, $0.008638414$ and $0.0193793$, respectively. This show that the ARIMA(1,0,1) which is the fitted model for raw Songkhla rainfall data has biggest values of both errors, while the ARIMA(2,0,2), the fitted model for the Daubechies wavelet transform data has the smallest value of both errors. Therefore we conclude that the fitted ARIMA models of both wavelet transformed data provides better fit than the fitted ARIMA model obtained from the raw data. Moreover, between the two fitted ARIMA models of the wavelet transform, the Daubechies wavelet transformed gives the better model.

# CHAPTER 5

# Conclusion

We transform the data with discrete wavelet transform via Haar and Daubichies. Then we continue to approximate with ARIMA model to make the forecasting and compare with the minimum value of MAE and RMSE. We determine the behavior of the data which transformed by discrete wavelet transform and then use R(3.1.3) program to determined the behavior of the data which ARIMA model.

As the last six years, in Songkhla which was also hard hit by flooding. We choose songkhla area is a sample in this study, to find the rainfall to keep up with flooding in the future.

We consider three step to analyzed the data. First, transformed by Haar and Daubechies discrete wavelet and continue to approximate to make the forecasting, compare in ARIMA model and to show that result in the last. we can see step of analysis on the figure as below to understanding step as following.


Figure 5.1: Step of analyze the data of rainfall.

## 5.1 Discrete wavelet transform analysis.

We transformed the data with discrete wavelet transform via Haar and Daubichies. From step of Haar and Daubechies wavelet transform in chapter $2$, we applied them with the all data of rainfall so we get the new result of Haar and Daubechies wavelet transform. We found that after transformed rainfall data via Haar and Daubechies we got the data is not much differences and an average may be constant which can make an observation, if an average is a constant we can bring about to stationary process. We notice every range of data we see no matter how much times we get an average is zero value according to stationary process.

## 5.2 ARIMA model analysis.

In this section, we try to find the fitted ARIMA model to make the forecasting in the future. We first to determine whether the series is stationary. From step$1$, an averages are constant. And autocorrelation function was die out rapidly and cut off. Therefore the data after we transform are stationary.

Now, the time series has been stationeries. For the next step we will fitting the model in ARIMA to determine whether AR or MA terms are needed to correct any autocorrelation that remains in the differences series. By looking at the autocorrelation function (ACF) and partial autocorrelation (PACF) plots of the differenced series, you can tentatively identify the numbers of AR and MA terms that are needed.

After we plot ACF, PACF and checked all properties of all ARIMA model we get the ARIMA$(1, 1, 1)$ for the rainfall data, ARIMA$(1, 0, 1)$ for Haar discrete wavelet transform and ARIMA$(2, 0, 2)$ for Daubechies wavelet transform which we see all model as following:

ARIMA$(1,1,1)$ :

$$\widehat{Y}_t \;=\; -0.3878Y_{t-1} - 1.0000\epsilon_{t-1}.$$

ARIMA$(1,0,1)$ :

$$\widehat{Y}_t \;=\; 0.1128 - 0.1888Y_{t-1} + 0.1355\epsilon_{t-1}.$$

ARIMA$(2,0,2)$ :

$$\widehat{Y}_t \;=\; 3.4478 - 0.025Y_{t-1} + 0.9420Y_{t-2} - 0.0041\epsilon_{t-1} - 0.8614\epsilon_{t-1}.$$

## 5.3 Analysis result.

The fit ARIMA model for the original return data is considered as ARIMA$(1,1,1)$ with root mean square error equal to $15.28966$, while the fit ARIMA model for the transform data by using Haar wavelet transform is selected as ARIMA$(1,0,1)$ with root mean square error equal to $14.91782$. Although the fit ARIMA model for the transform data using Daubechies wavelet transform is selected as ARIMA$(2,0,2)$ with root mean square error equal to $13.80871$. Therefore we can conclusion all of these criteria explain that the wavelet with ARIMA model is better than the ARIMA model. Moreover, the Daubechies wavelet transform gives more sufficient result and better than Haar wavelet transform in the forecasting. Results in chapter$4$ indicate that ARIMA model for the returns data after wavelet transforms produce smaller forecast error as compared to the ARIMA model for actual returns data. All of these criteria explain that the Daubechies wavelet transform gives more sufficient result and better than Haar wavelet transform in the forecasting for the next step.

Moreover, We can fitted ARIMA$(1,1,1)$, ARIMA$(1,0,1)$ and ARIMA$(2,0,2)$ to used with rainfall data in another area in thailand which we want to compare. And then we will applied and developed the model with another data to get maximum benefit for used in the next step in the future.

# Bibliography

[1] Akansu, Ali N.; Haddad and Richard A. (1992), *Multiresolution signal decomposition: transforms, subbands, and wavelets*, Boston, MA: Academic Press.

[2] A. Sheldon. (2010), *Linear Algebra Done Right*, Vol 3. Newyork: Springer Cham.

[3] Inner Product Spaceand Orthogonality, 2006.

[4] M. Alrumaih and A. Al- Fawzan, *Time series forecasting using wavelet denoising an application to Saudi stocks index*, J. king Saudi Univ, 14 (2002), 221-234.

[5] J. Contreras, R.Espinola, F. J. Nogales and A.b J.n Conejo, *ARIMA models to predict next-day electricity prices*, IEEE Transactions on Power Systems, 18(2005), 1014-20

[6] S. Alwadi and I.Mohd Tahir. *Selecting Wavelet Transform Model in Forecasting Financial Time series Data Based on ARIMA model*. Appied Mathematics Science, 5(2011), 316-324.

[7] F. Walnut, *An introduction to Wavelet analysis*, Springer Science and Business media, LLC, vol. 2, Birkhäuser, NewYork, 2004.

[8] Box, G. E. P., Jenskin, G , M., and Rensel, G. C. *Time series Analysis*, 3ed. Upper Saddle River, NJ :Prentice Hall, 1994.

[9] N. Aina A. R. and I. Mohd Tahir., *Denoising Malaysian Time Series Data*: A comparison Using Discrete and Stationary Wavelet Transforms. 2010 International Conference.on Science and Social Research, Malaysia 2010.

[10] Michael Falk., *A First Course on Time Series Analysis*. 2012 International Conference on Science and Social Research, Univesity of Wurzburg.

[11] Robert H. Shumway and David S. Stoffer.,*Time Series Analysis and Thier Applications*, J. king Saudi Univ, 14 (2002), 221-234.

# VITAE

**Name** Miss Nareemal Hilae

**Student ID** 5610220042

**Educational Attainment**

| Degree | Name of Institution | Year of Graduation |
|---|---|---|
| B.Sc. (Mathematics) | Prince of Songkla Univesity | 2012 |

**List of Publication and Proceeding**

H. Nareemal, N. Kitipol and P. Thammarat, The Application of Wavelet Transform to Analyze the Rainfall Data, *The* $22^{th}$ *Annual Meeting in Mathematics.,* Chiangmai University, Chiangmai, Thailand, June 02-04, 2017