Rainfall analysis in the Indian Ocean by using 6-States Markov Chain Model

Miftahuddin1*, Maulidawani1, I Setiawan2, Y Ilhamsyah2, and Fadhli3

1Department of Statistics, Faculty of Mathematics and Sciences, Syiah Kuala University, Aceh Indonesia
2Department of Marine, Faculty of Marine and Fisheries, Syiah Kuala University, Aceh Indonesia
3Department of Physics, Faculty of Mathematics and Sciences, Syiah Kuala University, Aceh Indonesia

*e-mail: miftah@unsyiah.ac.id

Abstract. In this study, we used Markov chain approach to analyze rainfall dataset from one buoy (4N90E) in the eastern Equatorial Indian Ocean. This study aims to determine the opportunity for transition (displacement) of daily rainfall intensity, where there are six states or conditions of rainfall intensity, i.e. no rain, very weak, weak rain, moderate rain, heavy rain, and very heavy rain. The Markov Chain method used is the Chapman-Kolmogorov equation and the steady state equation. The investigation of the 6-states in Markov chain model show that dynamic probability of transition state for rainfall data is reflexive properties majority. By using the model, it is concluded that the transition rate matrix of the largest transition probability in the area of 4N90E occurs at the transition from state-1 to state-1 as much is 0.72 and state-2 to state-2 is 0.60. The transition probability value becomes 0.5184 and 0.36 for the same state dominant of two periods $P^2$. The use of the 1st order Markov chain is better than 2nd order.

1. Introduction
Rainfall analysis is important for many applications related to environment. In real-time, many natural phenomena have uncertainties or probabilistic characteristics, for example rainfall. Stochastic processes are mathematical tools that can be used to model time-dependent data phenomena [1]. The Markov chain is a stochastic process with discrete parameters that meet the nature of Markov, future events only depend on the circumstances of today. Rainfall data is time dependent data that shows the change of circumstances [2] [3]. Generally, if the probably process is known to be in a certain condition, then the chance of prediction the process in the future depends on the current events and does not depend on the previous event.

Markov chain models can have used to compute the probability, frequency, and rate (or rain intensity) of movement associated with each rain transition between rainfall states within a single observation cycle as well as the approximate amount of cycles rainfall spent in a particular state of rainfall. When rainfall observations are made at regular intervals in a period, the number of cycles can be interpreted as time in a state like rainfall. Time spent in all states prior to absorption can be summed to estimate time duration of rainfall event or as the total survival time; In application Markov chains model needs two basic assumptions: (i) transition probabilities are constant over time (time homogeneity) and (ii) the probability of the next
transition depends only on the current state (the first-order Markov property). These models are attractive for time-to-event analysis [2, 4]. Therefore, the Markov chain is a series of event processes where the chance of conditional future events depends on the present event. [1, 2, 3]. The annual rainfall in Indonesia is not always the same in variability. The magnitude variability of rainfall depends on time and region. But lately there has been a physical change in the atmosphere from the variability of normal weather components to extreme conditions over a long period of time and have a wide impact on various sectors of human life [5, 6]. There are several steps of rainfall events as follows, evaporation, uplift and cooling, condensation, aggregation, and precipitation. Factors affecting precipitation are for altitude (different areas of the Earth's surface) receive different amounts of precipitation, and based on latitude, it rains more in the areas near the equator than in the temperature zones and polar region. Furthermore, factors affecting precipitation is different level of humidity. Actually, the temperature is higher near the equator so there is more evaporation.

Information relating to future conditions cannot be determined with certainty but can only be predicted or forecasted. In forecasting rainfall, it can be done by recognizing patterns of rainfall in the past, of course with influential parameters. Rainfall has a similar pattern or relative different for a certain period of year. Although there are changes, but not so drastic. Rainfall forecast is used for agricultural planning and flood early warning or anything related to the influence of rain factors. In forecasting rainfall there are two approaches that can be taken, namely factors that cause rain and historical data. In this study the approach to historical data will be carried out through rainfall data [7]. Rainfall is a natural phenomenon that belongs to one of the climate variables and is observed at all times in every region. Rainfall data is time series process, which is random in nature. In this data have state transfer from one time to another which can be expressed as a state of low, medium or high intensity.

Rainfall is the amount of water that falls to the surface of the earth in a certain time. The average rainfall in Indonesia every year potentially is different pattern and the rainfall that falls in the Indonesian region is influenced by several factors, including the shape of the terrain/topography, geography, environment, altitude from sea level, the direction of the slope, wind direction parallel to the direction of the coast and air 2003 pressure [8][9][10]. According Aldrian et al [8], there are three of rainfall regions, i.e., A, B, and C. Aceh includes climate region based on rainfall is in region B. Region B has two peaks, in October–November and in March to May. Those two peaks are associated with the southward and northward movement of the inter-tropical convergence zone (ITCZ). The annual cycle of Region B would be similar to that of Region A. Region B, there is a significant correlation between SST and rainfall variability, indicating a strong possibility for seasonal climate prediction. For March to May is the most difficult season to predict the variability of rainfall. Rainfall intensity is a size of the amount of rain per unit on certain time period during the rain frequently. Rainfall is generally divided into 5 levels according to the intensity as presented in Table 2.1 below.

| Level    | Intensity (mm/minute) |
|----------|-----------------------|
| Very Weak| < 0.02                |
| Weak     | 0.02 – 0.05           |
| Moderate | 0.05 – 0.25           |
| Heavy    | 0.25 – 1              |
| Very Heavy| > 1                  |

Table 1 shows that based on intensity level of rainfall, we can obtain six levels condition related to parameters in state, where no rain is other state. Beside variability related to spatially
(topography, geography), or physically, and position of the buoy, we also consider variability of the main object (i.e. rainfall) in six levels.

The Indian Ocean is the third largest ocean in earth with location from 30E to 120E and from 46S to 25N and Aceh region on 4° 41' 42.4860'' N and 96° 44' 57.8292'' E, where western and southern Aceh part of Indian Ocean [11]. The purpose of this study is to obtain a transition probability function of rainfall dataset, that will be used to obtain state transition probabilities at any time in the period November 2006 - December 2018 using the 6-states Markov chain estimation. The benefits of research are to construct a probability function that can be used to obtain state transition probabilities at any time in the period November 2006 - December 2018 using the 6-states Markov chain.

2. Materials and methods

2.1. Rainfall variability
Rainfall variability presents significant agricultural risks in Aceh province where several agriculture field areas are depending rainfall intensity. This factor as imply that majority fields were not yet obtain irrigation system for plantation activities adequately. In climate variability, one indicator to detect variability is by the rainfall observation. Rainfall is a phenomenon which an event of the fall of liquid from the earth atmosphere which is liquid and or frozen form to the surface of the earth such us land, ocean, river, lake, shore, and others surface. Rain usually requires the existence of a thick layer of air atmosphere can find above the melting point of ice above the surface of the earth. In earth, rain is a condensation process (changes in the form of objects into denser forms) of water vapor in the air atmosphere into a grain of water that is heavy enough to fall and usually arrives on surface of the earth. Two processes that may occur together can push air into saturation before rain occur, which is cooling condition the air or increasing water vapor into the air space. Rainfall variability occurs with various intensity, size, speed, and rain frequently, such as raindrops have sizes ranging from large grains to small grains or very tiny [12].

Rainfall gauge is the height of rainwater collected in a flat place, does not evaporate, does not seep, and does not flow. Rainfall magnitudes are expressed in millimeters (mm) or inches (inch) units. The rainfall in 1 mm has the meaning in an area of 1 m² in a place that is accommodated by 1 mm of water or accommodated by 1 liter of water. Rainfall intensity (RI) is the amount of rainfall in a given unit of time, the bias of which is expressed in mm/hour, mm/day, mm/year, and some of which are often called daily rain, monthly rain, annual rain, and so on.

2.2. Rainfall classification
Rainfall is differentiated according to the intensity and the speed of the fall. In this investigation, we tried to obtain classification detail probability of several rainfall condition in called a state. From this approach, our expected that probability of state transition of rainfall events can be estimated. According to [13] the classification of rain is based on rainfall intensity, as seen table 1 as follows:

(a) Rain is very low, if the RI is less than 0.02 mm/minute.
(b) Rain is low, if the RI ranges from 0.002 to 0.05 mm/minute.
(c) Normal rain, if the RI ranges from 0.05 to 0.25 mm/minute.
(d) Heavy rain, if the RI ranges from 0.25 to 1 mm/minute.
(e) The rain is very heavy, if the RI more than 1 mm/minute.

Rainfall classification seen from the speed of fall of rainfall occurs on:

a) Drizzle, if the ration speed is around 0.5 mm/sec.
b) Smooth rain, if the speed of fall ranges from 2.1 mm/sec.
c) Normal rain, if the speed falls around 4-6.5 mm/sec.
d) Rain is very heavy, if the speed of fall ranges from 8.1 mm/sec.
2.3. Stochastic process

In our research, investigation towards rainfall dataset by using stochastic process. Stochastic process application is one technique to study the dynamic relationship of a series of events or processes whose events are uncertain. Suppose a stochastic process is a sequence or set as collection of indexed random variables $X_t$ or $X(t)$, where $t$ is the elements in set $T$ which correspond to the time value or set index. The stochastic process can be written as $\{X(t), t \in T\}$[14]. The set $T$ is the set of indices that construct a space, namely the parameter space. All possible values that might occur of the $X(t)$ random variable as a set of $S$ states are called state space. The set of conditions $S$ builds a space, that is the state of space. The sets $T$ and $S$ are finite and infinite sets. Furthermore, a given set is called a discrete set and a non-quantifiable set as a continuous set [15].

Let $X(t)$ is a rainfall sequence in $t$ time as parameter space and 6-states space of rainfall intensity. In this time parameter, the stochastic process can be divided into two forms that is

a). If $T = \{0, 1, 2, 3, \ldots\}$ then stochastic process has discrete parameter and can be written as $X_t$

b). If $T = \{t | t \geq 0\}$ then stochastic process has continuous parameter and can be written as $X(t)$.

Therefore, stochastic process classification is based on the state and its parameter spaces, the Markov chain can be grouped as in the table below:

| Parameter Space | State Space | Continuous |
|-----------------|-------------|------------|
| Discrete        | Markov chain| Markov process |
| Continuous      | Markov chain| Markov process |

If the future state of the process does not depend on the past and only depends on the present, this process is called the Markov Process. Knowledge of the current state of processes must be adequate. To predict the next Markov process in the future, state knowledge is needed currently. The Markov process has two time parameters, namely:

(a). Discrete parameter, Markov process with discrete parameter usually called with Markov chain, and

(b). Time continuous parameter called with Markov process.

2.3.1. Markov chain. Suppose the stochastic process $\{X(t), t \in T\}$ is called a Markov chain if this is known, then the probability of the state of a process in one step going forward is only affected by the present situation. Systematically can be written:

$$P(X(t_{n+1}) \in A | X(t) = x_t, t \leq t_n) = P(X(t_{n+1}) \in A | X(t_n) = x_{t_n})$$

(1)

of all events that are in the $A$ and from time when $t_n < t_{n+1}$. Where the probability of the process moves from the state $x_{t_n}$ where at time $t_n$, this event belongs to the state $\Lambda$ at time $t_{n+1}$ which does not depend on the process to reach $x_{t_{n+1}}$ from $x_{t_0}$, where $t_0$ is at the initial time (not depending on the path followed by the process) from $x_{t_0}$ to $x_{t_n}$. When $\{X_n, n = 0,1,\ldots\}$ is a discrete state time process, which says systematic from Markov

$$P(X_{n+1} = j | X_n = i_n, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P(X_{n+1} = j | X_n = i_n)$$

(2)

of all states $i_0, i_1, \ldots, i_{n-1}, i, j$ and for each time $n$, have an equation
2.3.3. Conformity test of markov chain model.

The Markov chain has a transition probability that depends on the latest information with the available processes. The discrete Markov chain process has Markov properties. The stochastic process \( \{X_n, n = 0, 1, \ldots \} \) if state \( S_{X_n} \) is limited or infinite is the stationary Markov chain (homogeneous time), from equation (3) become \( p_{i,j} \) which

\[
P(X_{n+1} = j | X_{n} = i) = p_{i,j} \tag{4}
\]

For all state \( i_0, \ldots, i_{n-1}, i_j \) and for \( n \geq 0 \). The conditional probability in equation (4) can be written as \( p_{i,j}(n) \), when there is no risk, it can be written \( p_{i,j} \). When state of Markov chain is defined as system using set integer \( N^0 = \{0, 1, \ldots \} \) as a state space, has a probability condition \( p_{i,j} = P(X_{n+1} = j | X_n = i) \) depending on \( n \) \[16, 17, 18\]. Definition of Markov chain order is a Markov chain said to be order or \( O(r) \) if the appearance of \( X_n \) only depends directly on the previous \( r \) state namely \( X_{n-1}, X_{n-2}, \ldots, X_{n-r} \) and is independent of other states \[16, 17\]. Suppose \( O(r) \) applies

\[
P(X_n = i_n | X_{n-1} = i_{n-1}, \ldots, X_{n-(r-1)} = i_{n-(r-1)}, \ldots X_1 = i_1) = P(X_n = i_n | X_{n-1} = i_{n-1}, \ldots, X_{n-(r-1)} = i_{n-(r-1)}, X_{n-r} = i_{n-r})
\]

2.3.2. Transition rate matrix. Transition rate matrix indicates that one-step transition probabilities do not change with increasing time which has the same probability between past, present and future which is assumed that the probability of transition is stationary. Probability provides information to describe the initial condition process of the probability mass function for \( X_0 \) in state \( i \) and the one-step transition probability in \( X_i \). The quadratic matrix is used for transitions probability as

\[
P_{ij} = P(X_1 = j | X_0 = i) \tag{5}
\]

The \( P \) matrix contains probabilities, where each element is not negative value and the number of each row is one. The fact that each non-negative matrix is equal to one is called a Markov chain \[15\]. If \( X_n = 1 \) is a process where state \( i \), for example \( i = 1, 2, 3, \ldots \) at time \( n \) with \( n = 1, 2, 3, \ldots, n-1 \) and the current state \( X_n = i \). The \( P \) matrix of \( n \)-state transition probability is denoted as

\[
P = \begin{bmatrix}
P_{11} & P_{12} & P_{13} & \cdots & P_{1n} \\
P_{21} & P_{22} & P_{23} & \cdots & P_{2n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
P_{n1} & P_{n2} & P_{n3} & \cdots & P_{nn}
\end{bmatrix} \tag{6}
\]

Suppose the matrix \( P = [P_{ij}] \) is named the transition probability matrix with \( P_{ij} \) is named transition probability from state \( i \) to state \( j \) \[10\], where:

\[
P_{ij} \geq 0 \ \text{dan} \ \sum_{j=1}^{n} P_{ij} = 1, \quad i = 1, 2, \ldots, n \tag{7}
\]

2.3.3. Conformity test of markov chain model. Suppose a Markov chain order is determined through selecting the most appropriate order from several possible alternative options. Then, it should be noted that it is desirable that the order be as low as possible. The first step in conducting a conformity test, is assumed first that the Markov chain is \( \sigma(r) \), where \( r \) denotes the maximum order desired by the researcher or the order
that can still be tolerated by the data. Next, compare it to what happens if the Markov chain is called \( \sigma(r-1) \). Thus, the hypothesis is:

\[ H_0: \text{the use of } \sigma(r) \text{ is not different with } \sigma(r-1) \text{ or } \sigma(r) = \sigma(r-1) \]

\[ H_1: \text{the use of } \sigma(r) \text{ is different with } \sigma(r-1) \text{ or } \sigma(r) \neq \sigma(r-1) \]

Therefore, based on considering of the definition of the Markov chain order, hypothesis can also be written as:

\[ H_0: p_{ij...kt} = p_{j...kt} \]

\[ H_1: p_{ij...kt} \neq p_{j...kt} \]

Test statistics are:

\[
\chi^2_{\text{calculation}} = \sum_{i,j}^{n} \sum_{k=1}^{n} \left( \sum_{l=1}^{n} (m_{ij...kt} - m_{ij...kl} p_{j...kl}) \right)^2
\]  

(8)

Rejection criteria is \( H_0 \) rejected if \( \chi^2_{\text{calculation}} > \chi^2_{1-a,df} \).

Through the 1-step transition probabilities \( P_{ij} \), then the transition probabilities for \( z \)-next step \( P_{ij}^z \) with the state \( i \) moving process to state \( j \) after \( z \) state additional transitions.

\[ P_{ij}^z = P(X_{z+m} = j|X_m = i), z \geq 1, \ i, j \geq 1 \]

By using the Chapman-Kolmogorov equation for the transition probability for \( z \) is as follows:

\[
P_{ij}^{z+m} = P(X_{z+m} = j|X_m = i) = \sum_{k=1}^{\infty} P(X_{z+m} = j|X_z = k, X_1 = i)P(X_z = k|X_1 = i)
\]  

(9)

If \( P_{ij}^z \) is a matrix of transition probability \( z \) periods, then: \( P^{(z+m)} = P(z) P^{(m)} \)

By using induction obtained,

\[
P^{(z)} = P \times P^{(z-1)} = P \times P \times P^{(z-2)} = \ldots = P^{(z)}
\]

(10)

The Markov chain is called to be irreducible when each state can reappear (vs: transient state). If the process is finite (limited number of states) and irreducible, all states are recurrent. An irreducible finite process must be containing minimum one recurrent state. In a process with an unlimited number of states (or infinite), it cannot be guaranteed that the Markov chain is irreducible with recurrent state. Whereas Markov chain known as absorbing state is a state that does not experience a move to the further state. Suppose the Markov chain has entered and is in one-state then it will be in that state. The Markov chain can be ergodic if it is irreducible, positive (non-null), recurrent, and aperiodic [2][18][19].

Suppose state \( j \) is called accessible condition from state \( i \) then there is a integer number \( n \geq 1 \) where \( P_{ij}^n > 0 \). Therefore, there is a positive probability that in finite state \( j \) transitions can be achieved from state \( i \). If state \( j \) can be reached from state \( i \) and state \( i \) can be achieved from state \( j \) (conversely), then state \( i \) and state \( j \) are said to mutually communicate with each other, and are notated as \( i \leftrightarrow j \). If the two states \( i \) and \( j \) uncommunicated with each other, then \( P_{ij}^n = 0 \) or every \( n \geq 1 \). The relation concept of communication satisfies as follows:

i. If itself communicate notated as \( i \leftrightarrow i \), \( i \) is state from any Markov chain (reflexive property)
ii. If two states \( i \leftrightarrow j \), then from communicate definition, \( j \leftrightarrow i \) due to \( i \) accessible from state \( j \) and state \( j \) can be achieved from state \( i \) (symmetry property).

iii. If three states \( i \leftrightarrow j \) and \( j \leftrightarrow k \), then \( i \leftrightarrow j \) (transitive property).

In this equivalent relation, the state space of a Markov chain can be grouped into equivalent classes. If a Markov chain only has one equivalent class (all states communicate with each other), then the Markov chain is called irreducible process [2, 3, 20].

![Buoy position (in circle) is used in this analysis.](image)

**Figure 1.** Buoy position (in circle) is used in this analysis.

This research is an applied on the Markov Chain model and the data used is obtained from the website https://www.pmel.noaa.gov/gtmba/. The data used is daily data in the period November 2006 to December 2018 with the number of data as many as 1972 observations at point 4N90E in the Indian Ocean as pilot model. For describe the observation area, as displayed in the Figure 1.

3. Results and discussion

3.1. Descriptive statistics of the rainfall data

In Table 3 shows that maximum and range values of the rainfall data is very close (similar). This shows the gap between the presence of rain and no rain is quite large in the 4N90E area. For more details, we can see as displayed in Table 3.

| Statistic   | Value | Statistic     | Value  |
|-------------|-------|---------------|--------|
| Minimum     | 0.0000| Range         | 9.730  |
| Maximum     | 9.4300| Variance      | 0.545  |
| Median      | 0.0200| Standard deviation | 0.738 |
| Mean        | 0.2615| Mode          | 1.830  |
| Quartile 1  | 0.0000| Interquartile | 5.933  |
| Quartile 3  | 0.2400| Skewness      | 51.558 |
The advantage of the histogram is to be able to look a data whether it is normally distributed or not and the skewness pattern, as in Figure 3 below. The distribution of rainfall shows tend to positively skewed distributions.

Based on Figure 2 rainfall data are not normally distributed because the data are tails sticking out to the right or more data is to tend the left. Means the mode value is smaller than the mean and median, where the mode frequency value is more than 1500 for rainfall data.

Based on Figure 3, it shows that the rainfall data is not normally distributed and the data experienced significant fluctuations. In this time period show that peak season of rainfall on October-November as described [8]. However, the rainfall change fluctuation is still occurred until July.

3.2. State determination
In state determination, we use event as a set of parts from the sample space (see section 2.2), namely an event with certain conditions. The incident space is a collection of all events from a statistical experiment denoted by S. S = \{X_n, n \geq 0\} where X_n namely the condition of the rain (not rain, very low rain, low rain, normal rain, heavy rain, very heavy rain) and stated by criterion for example, \(X_n=0\) (not rain), \(X_n=1\) (very low) if rainfall intensity < 0.022 (mm/minute), \(X_n=2\) (low) if rainfall intensity 0.022 – 0.05 (mm/minute), \(X_n=3\) (normal) if rainfall intensity 0.05 - 0.25 (mm/minute), \(X_n=4\) (heavy rain) if rainfall intensity 0.25–1.00 (mm/minute), and \(X_n=5\) (very heavy rain) if rainfall intensity > 1 (mm/minute). In addition, we used parameter classification for rainfall in this investigation is as described in table 1 with Markov process. In this process, we can use time unit (daily) to observe distribution and pattern of rainfall based on trend, seasonal and annual patterns.
3.3. Determination of the order of the Markov chain

This test aims to determine the order used in forecasting rainfall in the next period. Testing between order 0 and order 1, with hypothesis:

\[ H_0: p_{0j} = p_{1j} = p_{2j} = p_{3j} = p_{4j} = p_{5j} = p_j \]

\[ H_1: \text{there are at least one different transition probability} \]

The frequency of conditions for the month at \( t-1 \), where \( t \) is the period November 2016 to December 2018 can be seen in Table 4.

### Table 4. Frequency of conditions for the month at the time \( t-1 \) in 4N90E.

| State       | Not rain | Very low rain | Low rain | Normal rain | Heavy rain | Very heavy rain | Total |
|-------------|----------|---------------|---------|-------------|------------|----------------|-------|
| Frequency   | 1186     | 106           | 94      | 25          | 270        | 163            | 1972  |

To obtain the transition probability value is used formulation \( P(A) = \frac{n(A)}{n(S)} \) so that it is obtained:

\[ p_{11} = \frac{n(1)}{n(S)} = \frac{1186}{1972} = 0.60 \]
\[ p_{12} = \frac{n(2)}{n(S)} = \frac{106}{1972} = 0.05 \]
\[ p_{13} = \frac{n(3)}{n(S)} = \frac{94}{1972} = 0.04 \]
\[ p_{14} = \frac{n(4)}{n(S)} = \frac{25}{1972} = 0.01 \]
\[ p_{15} = \frac{n(5)}{n(S)} = \frac{270}{1972} = 0.13 \]
\[ p_{16} = \frac{n(6)}{n(S)} = \frac{163}{1972} = 0.08 \]

Based on the above results is obtained a transition probability matrix:

\[ P_{ijklmno} = [p_{11}, p_{12}, p_{13}, p_{14}, p_{15}, p_{16}] \Rightarrow P_{ijklmno} = [0.60, 0.05, 0.04, 0.01, 0.13, 0.08] \]

The frequency of conditions for the month at time \( t-1 \), where \( t \) is the period of November 2016 to December 2018 can be seen in Table 5.

### Table 5. Pattern of state transfer at point 4N90E.

| State | 1   | 2   | 3   | 4   | 5   | 6   | Total |
|-------|-----|-----|-----|-----|-----|-----|-------|
| 1     | 557 | 11  | 11  | 102 | 53  | 30  | 764   |
| 2     | 5   | 33  | 4   | 7   | 6   | 0   | 55    |
| 3     | 9   | 5   | 31  | 13  | 5   | 8   | 71    |
| 4     | 81  | 15  | 11  | 173 | 55  | 20  | 355   |
| 5     | 60  | 7   | 11  | 49  | 128 | 26  | 281   |
| 6     | 19  | 3   | 3   | 15  | 18  | 42  | 100   |
| Total | 731 | 74  | 71  | 359 | 265 | 126 | 1626  |

To obtain the transition probability value is used formulation \( P(A) = \frac{n(A)}{n(S)} \) so that it is obtained:

\[ p_{11} = \frac{n(1)}{n(S)} = \frac{557}{764} = 0.72 \]
\[ p_{12} = \frac{n(1)}{n(S)} = \frac{11}{764} = 0.01 \]

...
By using equation (6), obtained the transition probability matrix in 6-states is:

\[
\begin{bmatrix}
\frac{n(1)}{n(s)} & \frac{n(2)}{n(s)} & \frac{n(3)}{n(s)} & \frac{n(4)}{n(s)} & \frac{n(5)}{n(s)} & \frac{n(6)}{n(s)} \\
11 & 31 & 31 & 15 & 18 & 18 \\
76 & 71 & 71 & 100 & 100 & 100 \\
55 & 55 & 55 & 100 & 100 & 100 \\
335 & 335 & 335 & 100 & 100 & 100 \\
19 & 19 & 19 & 100 & 100 & 100 \\
335 & 335 & 335 & 100 & 100 & 100 \\
20 & 20 & 20 & 100 & 100 & 100
\end{bmatrix}
\]

\[
p_{13} = \frac{11}{76} = 0.14, \quad p_{14} = \frac{31}{71} = 0.43, \quad p_{15} = \frac{31}{71} = 0.43, \quad p_{16} = \frac{15}{100} = 0.15, \quad p_{23} = \frac{19}{335} = 0.05, \quad p_{25} = \frac{18}{100} = 0.18, \quad p_{26} = \frac{42}{100} = 0.42
\]

The results of the above probability can be visualized in the form of a transition diagram as follows:

Figure 4. Diagram of transition probability \( P \) with 6-states.

To test the hypotheses in this section, we used the equation (8) and test by using the chi-square distribution with degrees of freedom \( s^2 - (s - 1)^2 \) where \( H_0 \) rejected if \( x^2_{\text{calculation}} > x^2_{\text{table}}(\alpha) \).

\[
x^2_{\text{calculation}} = 2 \left[ \sum_{i=1}^{s} \sum_{j=1}^{s} \sum_{k=1}^{s} n_{ijklmn} (\ln p_{ijklmn} - \ln p_{ijklmn}) \right]
\]

\[
x^2_{\text{calculation}} = 2042.1972
\]
From the Chi-square table with a significant level $\alpha = 0.05$ and the degree of freedom obtained is $x_{\text{table}}^2(\alpha) = 21.026$. Because $x_{\text{calculation}}^2 = 2042,1972 > x_{\text{table}}^2(\alpha) = 21.026$ then $H_0$ is rejected at a significant level of 5%. This means that order 1 and order 2 differ from one another. In processing data for point 4N90E must be modeled using order 1 where the matrix measures 6 x 6.

For a matrix of transition probability 2 periods is

$$P^2 = \begin{bmatrix}
0.5184 & 0.0001 & 0.0001 & 0.0169 & 0.0036 & 0.0009 \\
0.0081 & 0.3600 & 0.0049 & 0.0144 & 0.0100 & 0.0000 \\
0.0144 & 0.0049 & 0.1849 & 0.0324 & 0.0049 & 0.0121 \\
0.0484 & 0.0016 & 0.0009 & 0.2304 & 0.0225 & 0.0025 \\
0.0441 & 0.0004 & 0.0009 & 0.0289 & 0.2025 & 0.0081 \\
0.0361 & 0.0009 & 0.0009 & 0.0225 & 0.0324 & 0.1764
\end{bmatrix}$$

![Figure 5. Diagram of transition probability $P^2$ for 2 periods with 6-states.](image)

4. Conclusion

The result of existing rainfall data and predicted rainfall model in state transition can be compared and calculated by using stochastic process. Characteristics of variability of rainfall in the sampling point (4N90E) is different value of transition probability of 6 states as stated in the matrix of transition probability and visualization in diagram of transition probability.

It can be concluded that the majority communication state with transition probabilities value of rainfall data occurs in the state transition from state-1 to state-1 by 0.72 and from state-2 to state-2 as much is 0.60. Both transitions are reflexive property more dominant compared to different state. The transition probability value becomes 0.5184 and 0.36 for the same state dominant of two periods $P^2$. 
Both transitions are reflexive property more dominant compare with different state. The transition probability value becomes 0.5184 and 0.36 for the same state dominant of two periods $P^2$. The use of the first order Markov chain with 6-states are better than order 2 for the 4N90E area in the Indian Ocean.

References
[1] Nicolas P 2013 Understanding Markov Chains, Examples and Applications Springer
[2] Sheldon M R 1996 Stochastic Processes 2nd Edition John Wiley & Sons, Inc
[3] Sheldon M R 2000 Introduction to Probability Models. Seventh Edition. Academic Press, California
[4] Salazar J C, Schmitt FA, Yu L, Mendiondo M S, Kryscio R J 2007 Shared random effects analysis of multi-state Markov models: application to a longitudinal study of transitions to dementia. Statist Med 26 568–580. http://www.ncbi.nlm.nih.gov/pubmed/16345024
[5] Nurdin 2011 Antisipasi Perubahan Iklim untuk Keberlanjutan Ketahanan Pangan (Gorontalo: Universitas Negeri Gorontalo)
[6] Paida 2011 Deteksi Perubahan Pola Curah Hujan Kota Makassar (Indonesia: Universitas Hasanuddin)
[7] Luk, K C, Ball, J E, and Sharma, A 2001 An Application of Artificial Neural Networks for Rainfall Forecasting. Mathematical and Computer Modelling 33 883-699
[8] Aldrion E and Susanto, RD 2003 Identification of three dominant rainfall regions within Indonesia and their relationship to sea surface temperature. Int. J. Climatol. 23 1435–1452
[9] Lee HS 2015 General Rainfall Patterns in Indonesia and the potential Impacts of Local Seas on Rainfall Intensity Journal of Water 7 1751-1768
[10] BMKG 2007 Bulletin Data Iklim 2006 Level of rainfall based on its intensity (Jakarta: BMKG)
[11] Miftahuddin and Y Ilhamsyah 2018 Modeling of sea surface temperature using linear models with autocorrelation Indian Ocean, IOP Conference Series: Earth and Environmental Science 176 1755-1315, 176-012038
[12] Triadmodjo 2008 Hidrologi Terapan (Yogyakarta: Beta Offset)
[13] Badan Meteorologi Klimatologi dan Geofisika 2006-2013 Prakiraan Curah Hujan (Jakarta: BMKG)
[14] Nisa, E F 2012 Analisis Model Stokastik Pada Proses Kelahiran dan Kematian Sel T Naif dengan Waktu Kontinu Rantai Markov Skripsi Universitas Islam Negeri Maulana Malik Ibrahim, Malang Indonesia
[15] Rahma Z and Miftahuddin 2010 Peramalan Pengaruh Unsur Iklim melalui Rantai Markov Waktu Kontinu Melalui 3 Keadaan: El Nino, La Nina dan Normal, Prosiding SEMIRATA BKS-PTN Barat Bidang MIPA ke 23, UNR1, Riau, Pusat Pengembangan Pendidikan Universitas Riau, p 152-160, ISBN: 978-979-1222-95-2
[16] Lefebvra, M 2007 Applied Stochastic Processes (New York: Springer Science Business Media LLC, 223 Spring Street)
[17] Masuku, F N, Langi, Y A R, and Mongi, C 2018 Analisis rantai Markov untuk memprediksi perpindahan konsumen maskapai penerbangan rute Manado-Jakarta. Jurnal Ilmiah Sains 18 2 (Indonesia)
[18] Feldman, R M and Valdez-Flores, C 2010 Applied Probability and Stochastic Processes (London, New York, Heidelberg, Dordrecht: Springer)
[19] Taylor, H M and Karlin, S 1993 An Introduction to Stochastic Modelling (California: Academic Press)
[20] Song C, Kuo L, Derby C A, Lipton R B, Hall C B. Multi-stage transitional models with random effects and their application to the Einstein Aging Study J. Biometrical 2011; 53:938–955. http://www.ncbi.nlm.nih.gov/pubmed/22020750. 10.1002/bimj.200900259

Acknowledges
The contribution of the Department of Statistics, Department of Physics, Faculty of Mathematics and Sciences, and the Department of Marine, Faculty of Marine and Fisheries, the Institute of Research and...
Public Services of Syiah Kuala University and DRPM Ristekdikti are large valuable in this study. We also would like to appreciate the reviewers of the second ICMS 2019 for their inputs, sharing and suggestions.