ASM Sc. J., 20(2), 2025
https://doi.org/10.32802/asmscj.2025.1581

Rain, Rain, Go Away: Stochastic Model in Predicting the Future Rainfall

S. Samsuddin*, N. Rosdi, N.H. Zainal, N.A. Zailan, N.F. Zamri and R.M. Omar

School of Mathematical Sciences, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

This study investigates the use of the Markov Chain (MC) method to forecast future rainfall in Tanah Merah, Kelantan, by observing daily rainfall data and categorising it into five distinct states, denoted as S={1,2,3,4,5}, each representing different levels of rainfall intensity. A structured examination of rainfall transitions between states is made possible by the research's discrete definitions of the state space and the temporal set. The study reflects the dynamic character of weather patterns by capturing the possibility of changes in rainfall amounts through the creation of a Transition Probability Matrix (TPM) for each month. In addition to forecasting rainfall, the study computes the limiting distribution of the TPMs to create risk matrices for every state. These risk matrices, which are based on recent and past rainfall data, offer a probabilistic evaluation of future flood hazards. The monthly risk matrices provide important insights into flood prediction and disaster preparedness by showing how the likelihood of rainfall in each state can affect the chance of flooding in succeeding months. The study illustrates the potential of the MC technique in enhancing flood risk management in the region and improving rainfall forecasts by utilising the concept of long-run behaviour.

Keywords: Markov chain; rainfall; flood

I. INTRODUCTION

The development of flora and wildlife in the natural world depends on rainfall. It is necessary for people, plants, animals, and other living things (Abeywardena, 1955). In addition, rainfall, which is considered the most natural resource on earth, is a crucial resource for agricultural output. However, particularly in low-income areas, excessive rainfall can have a disastrous effect, especially on farmers (Meza-Pale & Yunez Naude, 2015). Then, receiving the right amount of rainfall to suit human demands and continue to use in daily life may be difficult due to changing global climate conditions. Therefore, in order to meet human needs and raise awareness of the potential for natural disasters that unexpectedly intense rainfall could cause, it is now crucial to evaluate rainfall patterns and attempt to predict rain.

According to Bopi *et al.* (2016), Sarawak and Sabah, as well as the Peninsular region, get 2500 mm and 3500 mm of rain annually, respectively. Because of several elements, including geology and winds, the frequency of rainfall differs according to each state in Malaysia. The Malaysian east coast has faced significant rainfall, especially during the North-East Monsoon, which has an average annual rainfall of 2700 mm. This monsoon typically lasts from November to March each year.

A day with more than 0.1 millimetres of rain may be referred to as a rainy day, according to Telipot (2000). There are a few states in Malaysia where it frequently rains heavily during specific times of the year, and among them is the state of Kelantan. For this study, the main focus would be the area of Tanah Merah. In addition, Tanah Merah experiences hot, humid and cloudy weather. Throughout the year, the temperature fluctuates between 25°C and 33°C, rarely dropping below 20°C or going over 34°C. Tanah Merah has a tropical monsoon climate with significant seasonal variations in monthly rainfall. Every year, Tanah Merah experiences

*Corresponding author's e-mail: shamshimah@tmsk.uitm.edu.my

248.38 wet days (68.05% of the total), with an average of 164.43 millimetres (6.47 inches) of precipitation. There are also several researchers that perform study in Tanah Merah on rainfall, but these studies concentrate more on various research goals like the development of a rainfall rate monitoring system, such as Bopi *et al.* (2016) and the analysis of annual maximum and partial duration rainfall series, in Ng *et al.* (2016).

Statistics can be used to predict rainfall, especially the amount, which becomes increasingly important as the year progresses. In addition to hydrological goals, it is used to model economic development, including typical corporate processes. For instance, land management systems, agricultural growth, the design of urban drainage systems, and other environmental initiatives are some of the industries that require a forecast of rainfall distribution each year, according to Parmar *et al.* (2017). According to Joseph and Ratheesh (2013), rainfall forecasting has become one of the world's most challenging scientific and technological conundrums.

Furthermore, one of the factors that might lead to several issues is incorrect or poor rainfall forecasting. The accuracy of Tanah Merah's different enterprises and the facilitation of local affairs were both aided by the accurate rainfall forecast. Therefore, this study employed the Markov Chain (MC) model to predict future rainfall in Tanah Merah, Kelantan. According to Karlin (2014), the MC has probability analysis for reducibility, periodicity, recurring and transitory states, and limiting distribution as a stochastic process. The Department of Irrigation and Drainage and World Weather Online are utilised to collect the three years of rainfall data in Tanah Merah, Kelantan, from 2018 to 2020.

II. MATERIALS AND METHOD

Rainfall data is gathered from Kusial station in Kelantan's Tanah Merah area, and it is recorded hourly, daily, and measured in millimetres. The Department of Irrigation and Drainage in Malaysia and World Weather Online (WWO) provided the daily rainfall data for three years in a row, starting on January 1st, 2018, and ending on December 31st, 2020 (1095 days number of observation).

The Discrete Time Markov Chain (DTMC) method is widely used in environmental research studies. For example, Holmes *et al.* (2021) used discrete-time Markov chains to study the stochastic behaviour of air pollution indices in Ontario, Canada; Yakasiri (2019) studied the stochastic approach for the state-wise forecast of wind speed using discrete-time Markov chains in Japan; and Nop *et al.* (2021) studied rainfall in ideal rainwater harvesting system operation in Japan. This study employed the DTMC model to replicate the rainfall data on day-to-day basis. From this basis, this study may observe 364 steps. Based on the Markov property in this DTMC model, regardless of the past state, the probability of transition to the next state only depends on the current state.

Additionally, some research on rainfall employs the MC model. A study from Setiawan and Ilhamsyah (2020) addresses rainfall analysis in the Indian Ocean using a 6-State MC model, Mahanta and Khosro (2018) discussed utilising an MC to analyse the rainfall conditions in Bangladesh, Da Silva (2019) used information from five weather stations spread throughout the state's mesoregions to examine the patterns of daily rainfall in the State of Paraíba through MC and Rohit *et al.* (2021) used MC in characterising the inherently stochastic nature of the dimensionless time distribution of extreme rainfall in India and United States.

Five different states are used in the study to represent rainfall classification. Each state has its intensity of rainfall., State 1 indicates that there will be no rain that day because the rainfall intensity is less than 1 mm/h. Next, the state will be 2 for rainfall intensities ranging from 1 to 10 mm/h, indicating that the rainfall on that day is light rain. Then, moderate rain, defined as 11 to 30 mm/h, will be classified as State 3.

Heavy rain will be classified as State 4 if the rainfall is between 31 and 60mm/h. Finally, State 5 denotes extreme rain with a rainfall rate of greater than 60 mm/h. The table below summarises all the rainfall states and rainfall intensity and their classification.

Table 1. Rainfall states, intensity, and its classification

State	Rainfall Intensity (mm/h)	Classification of rainfall
1	Less than 1	No Rain
2	1 to 10	Light Rain
3	11 to 30	Moderate Rain
4	31 to 60	Heavy Rain
5	More than 60	Extreme Rain

The rainfall is assumed to be an independent random variable, complying with time homogeneity, and therefore:

$$P_{ij} = \Pr(X_{k+1} = s_j \mid X_k = s_i)$$

where:

P_{ij}: The probability that is moving from state i to state j

k: The current time of the transition probability matrix, k=1,2,3,4 and 5

s_i: The state of the transition probability matrix at time i

s_j: The state of the transition probability matrix at time j.

The equation above is used to construct the transition probability matrix. However, since the daily rainfall data is not in probability form, some modification needs to be done to obtain the transition probability matrix. To obtain the monthly transition probability matrix, the rainfall sequences are created using daily rainfall data acquired from the Department of Irrigation and Drainage, as illustrated below:

$$Q = \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{bmatrix} n_{11} & n_{12} & n_{13} & n_{14} & n_{15} \\ n_{21} & n_{22} & n_{23} & n_{24} & n_{25} \\ n_{31} & n_{32} & n_{33} & n_{34} & n_{35} \\ n_{41} & n_{42} & n_{43} & n_{44} & n_{45} \\ n_{51} & n_{52} & n_{53} & n_{54} & n_{55} \end{bmatrix}$$

where n_{ij} denoted as frequency in state i followed by state j for i,j = {1, 2, 3, 4 and 5}.

The sequences of monthly rainfall transition probability matrix will be constructed as below, $P = \frac{n_{ij}}{\sum n_{ij}}$ for each row:

$$P = \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} \\ p_{21} & p_{22} & p_{23} & p_{24} & p_{25} \\ p_{31} & p_{32} & p_{33} & p_{34} & p_{35} \\ p_{41} & p_{42} & p_{43} & p_{44} & p_{45} \\ p_{51} & p_{52} & p_{53} & p_{54} & p_{55} \end{bmatrix}$$

The transition probability matrix analysis is conducted to confirm that the transition probability matrix is an ergodic MC. By classifying the state of the transition probability matrix, pij, this confirmation of an ergodic MC must be accomplished to identify the existence of a limiting distribution in this chain. The ergodic Markov chain, which is irreducible, positive recurrent, and aperiodic if and only if the MC, X_n, has an ergodic initial state, which also is of sequence k where (k = 1,2,3, 4,) for all n (Karlin, 2014). It has also been found that if the generated Markov chain model is irreducible, aperiodic, and recurrent, then an irreducible and aperiodic ergodic MC will be achieved if a finite-state MC indicates ergodic behaviour. Aside from that, this study's goal requires that the transition probability matrix not be an absorbing state to confirm a limiting distribution in this chain. The classification of this transition probability matrix can be divided into three sections: irreducible Markov chain, periodicity Markov chain, and recurrent and transient states.

a. Irreducible

According to Karlin (2014) and Ross (2014), the MC is irreducible when each state can return. It is also possible to conclude that if all states communicate with one another, the MC with finite states is irreducible [5].

b. Periodicity

According to Karlin (2014) and Ross (2014), if the period is less than 2, the state is aperiodic, whereas if the period is equivalent or greater than 2, the state is said to be periodic. According to [5], a state in a sequence is shown to be periodic if the chain can only return to it at multiples of a particular integer greater than 1.

c. Recurrent and Transient States

Recurrent and transient states can be distinguished if, for given state i, let fi be the probability that the process will ever return to state i. State i is recurrent if fi = 1 and transient if fi < 1. If j→i for every i and j, then it can also be said that the states are recurrent and that no transient state occurs.

d. Ergodic

Limiting distribution can be calculated if the MC is stationary, ergodic MC, and absorbing state. Long-run behaviour is forecasted using a limiting distribution with the steady-state probabilities independent from initial conditions. As a result, the states' probability will be stationary. The probability distribution $\pi = (\pi_1, \pi_2, \pi_3, \pi_4, \pi_5)$ is called the limiting distribution of the MC X_n if $\pi_j = \lim_{n\to\infty} P(X_n = j \mid X_1 = i)$ for $\sum \pi_j = 1$ for all i,j in the state space, where j = 1,2,3,4 and 5.

This convergence means that in the long run ($n \to \infty$), the probability of finding the MC in state j is approximate to π_j no matter in which state the chain began at time 0. Therefore, this study will obtain the values of π_j by using the following formula and steps:

$\pi_j = \sum_{n=1}^{5} \pi_n P_{nj}$ for j = 1,2,3,4 and 5. Thus, $\pi_j = 1$.

For Rainfall Matrix (RM), was developed to show the different types of rainfall impact that Tanah Merah experiences throughout the upcoming month. In order to represent the likelihood of flood impact resulting from the long-run behaviour of rainfall, the Rainfall Matrix employs four different colours: green, yellow, orange and red. The level of flood hazard has also been observed in this Rainfall Matrix based on the Department of Irrigation and Drainage level of flood indicator that is used in Malaysia. The following is the explanation for the colours and likelihood of rainfall impact, the classification of flood hazard, and the indicator of flood impact.

Table 2. Likelihood of flood impact

Colour	Likelihood of Flood Impact
(green)	Very low occurrence of flood hazard
(yellow)	Low occurrence of flood hazard
(orange)	Moderate occurrence of flood hazard
(red)	High occurrence of flood hazard

Table 3. Classification of flood hazard

Flood Hazard	Explanations
Normal	The flood level is at a safe level.
Alert	The flood level is above the normal level.
Warning	The flood level increases to near flooding level and prepares for any evacuation action.
Danger	The flood level is causing considerable flooding, and evacuation is initiated.

Table 4. Indicator of flood impact

Limiting Distribution Probability	Flood Impact
< 0.20	Very low
0.20 to 0.40	Low
0.41 to 0.60	Moderate
> 0.60	High

Table 5 until Table 9 below is constructed based on the combination of Table 2 to Table 4 to indicate the events of flood hazard based on the likelihood of flood impact for each rainfall state.

Table 5. Justification for State 1 (no rain) flood hazard

Limiting Distribution Probability	Normal	Alert	Warning	Danger	Justification
Very low (<0.20)	Very low	Very low	High	High	The very low occurrence of no rain indicates that another state of rainfall has a high chance of flooding.
Low (0.20 to 0.40)	Moderate	Moderate	Moderate	High	The low occurrence of no rain indicates the moderate chances of rainfall intensity that will lead to flood, but the flood event might occur.
Moderate (0.41 to 0.60)	Low	Low	Very low	Very low	The moderate occurrence of no rain indicates that another state of rainfall has very low chances of flood occurrence.
High (>0.60)	High	Very low	Very low	Very low	The high occurrence of no rain indicates the very low chances of rainfall intensity that will lead to a flood.

Table 6. Justification for State 2 (light rain) flood hazard

Limiting Distribution Probability	Normal	Alert	Warning	Danger	Justification
Very low (<0.20)	Low	Low	Moderate	Moderate	The very low occurrence of light rain indicates the moderate chances of rainfall intensity that will lead to a flood.
Low (0.20 to 0.40)	Moderate	Low	Very low	Very low	The low occurrence of light rain indicates the very low chances of flood events with low alert and moderate normal flood impact.
Moderate (0.41 to 0.60)	Moderate	Very low	Very low	Very low	The moderate occurrence of light rain determines very low chances of rainfall intensity that will lead to a flood.
High (>0.60)	High	Very low	Very low	Very low	The high occurrence of light rain indicates that another state of rainfall has very low chances of rainfall intensity that will lead to a flood.

Table 7. Justification for State 3 (moderate rain) flood hazard

Limiting Distribution Probability	Normal	Alert	Warning	Danger	Justification
Very low (<0.20)	Moderate	Moderate	Low	Low	The very low occurrence of moderate rain indicates that the chances of a flood are low.
Low (0.20 to 0.40)	Low	Low	Moderate	Moderate	The low occurrence of moderate rain indicates that another state of rainfall has a reasonable chance of rainfall intensity leading to flooding.
Moderate (0.41 to 0.60)	Very low	Low	Moderate	Moderate	The moderate occurrence of moderate rain indicates that another state of rainfall has a very low chance of rainfall intensity that will lead to a flood. Still, the normal flood impact is very low as the chances for flood occurrence is moderate.
High (>0.60)	Very low	Low	High	High	The high occurrence of moderate rain indicates that another state of rainfall has a high chance of rainfall intensity leading to flooding.

Table 8. Justification for State 4 (heavy rain) flood hazard

Limiting Distribution Probability	Normal	Alert	Warning	Danger	Justification
Very low (<0.20)	High	Very low	Low	Low	The very low occurrence of heavy rain indicates that the chances of a flood are low.
Low (0.20 to 0.40)	Moderate	Low	Low	Low	The low occurrence of heavy rain indicates that another state of rainfall has low chances of rainfall intensity that will lead to a flood.
Moderate (0.41 to 0.60)	Very low	Low	High	High	The moderate occurrence of heavy rain indicates that another state of rainfall has high chances of rainfall intensity that will lead to flood, thus making the normal flood impact very low.
High (>0.60)	Very low	Very low	High	High	The high occurrence of heavy rain indicates high chances of rainfall intensity that will lead to a flood.

Table 9. Justification for State 5 (extreme rain) flood hazard

Limiting Distribution Probability	Normal	Alert	Warning	Danger	Justification
Very low (<0.20)	High	High	Very low	Very low	The very low occurrence of extreme rain indicates that the chances of rainfall intensity that will lead to a flood are very low.
Low (0.20 to 0.40)	High	Low	Low	Low	The low occurrence of extreme rain indicates that another state of rainfall has low chances of rainfall intensity that will lead to a flood.
Moderate (0.41 to 0.60)	Low	Low	High	Moderate	The moderate occurrence of extreme rain indicates that floods will occur in the upcoming days.
High (>0.60)	Very low	Very low	High	High	The high occurrence of extreme rain indicates that the flood level is considering the flood event that month.

III. RESULT AND DISCUSSION

Figure 1 shows the sequence of rain fall and Figure 2 show the matrix of transition probability, which was generated to obtain sequences of monthly matrix, and it also be plotted to observe the probability from one state to another in Figure 4. There are 265 sequences from no rain to no rain, showing the highest rainfall sequences in 2018 to 2020 and 145 series from extreme rain to extreme rain. Meanwhile, the lowest rainfall sequence between 2018 and 2020 is ten sequences, from heavy to no rain. Then, compared to other states, the no rain to no rain state has the most significant sequences from February to June and November. The highest sequences of light rain to light rain are found in July, August, and October, whereas the highest sequences of moderate rain to light rain are found in September. Finally, the highest sequences for December are the extreme rain to extreme rain state.

January

$$\begin{bmatrix} 10 & 2 & 3 & 3 & 2 \\ 8 & 9 & 1 & 0 & 4 \\ 2 & 6 & 2 & 2 & 2 \\ 0 & 1 & 2 & 0 & 7 \\ 0 & 5 & 6 & 5 & 10 \end{bmatrix}$$

February

$$\begin{bmatrix} 33 & 2 & 2 & 1 & 4 \\ 5 & 9 & 1 & 0 & 2 \\ 1 & 5 & 1 & 2 & 0 \\ 2 & 1 & 1 & 0 & 1 \\ 2 & 0 & 3 & 2 & 4 \end{bmatrix}$$

March

$$\begin{bmatrix} 54 & 6 & 3 & 1 & 1 \\ 7 & 2 & 2 & 0 & 2 \\ 3 & 3 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 \end{bmatrix}$$

April

$$\begin{bmatrix} 48 & 2 & 4 & 1 & 2 \\ 5 & 0 & 0 & 2 & 1 \\ 1 & 3 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 3 \\ 1 & 3 & 1 & 2 & 5 \end{bmatrix}$$

May

$$\begin{bmatrix} 53 & 3 & 4 & 2 & 1 \\ 7 & 1 & 2 & 0 & 2 \\ 3 & 4 & 2 & 0 & 0 \\ 0 & 2 & 1 & 1 & 0 \\ 0 & 2 & 0 & 1 & 1 \end{bmatrix}$$

June

$$\begin{bmatrix} 26 & 1 & 2 & 1 & 3 \\ 2 & 4 & 3 & 2 & 4 \\ 2 & 4 & 3 & 2 & 0 \\ 1 & 2 & 1 & 3 & 6 \\ 2 & 3 & 3 & 5 & 4 \end{bmatrix}$$

July

$$\begin{bmatrix} 7 & 3 & 1 & 1 & 0 \\ 3 & 22 & 9 & 3 & 2 \\ 2 & 11 & 12 & 1 & 1 \\ 0 & 3 & 1 & 1 & 2 \\ 0 & 0 & 4 & 1 & 2 \end{bmatrix}$$

August

$$\begin{bmatrix} 6 & 8 & 0 & 0 & 4 \\ 9 & 20 & 9 & 0 & 4 \\ 2 & 8 & 4 & 0 & 1 \\ 1 & 2 & 0 & 1 & 0 \\ 0 & 3 & 3 & 3 & 4 \end{bmatrix}$$

September

$$\begin{bmatrix} 0 & 3 & 4 & 0 & 1 \\ 6 & 9 & 10 & 3 & 2 \\ 1 & 12 & 6 & 3 & 3 \\ 1 & 6 & 4 & 2 & 1 \\ 0 & 0 & 1 & 6 & 5 \end{bmatrix}$$

October

$$\begin{bmatrix} 6 & 3 & 1 & 0 & 6 \\ 7 & 15 & 10 & 0 & 2 \\ 2 & 13 & 4 & 0 & 2 \\ 0 & 1 & 2 & 0 & 0 \\ 1 & 1 & 5 & 3 & 8 \end{bmatrix}$$

November

$$\begin{bmatrix} 22 & 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 2 \\ 2 & 3 & 2 & 0 & 3 \\ 3 & 0 & 0 & 7 & 36 \end{bmatrix}$$

December

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 3 \\ 2 & 2 & 2 & 6 & 0 \\ 1 & 0 & 1 & 5 & 65 \end{bmatrix}$$

Figure 1. Sequence of Monthly Rainfall in Tanah Merah, Kelantan for the years 2018 to 2020

January

$$\begin{bmatrix} 0.5000 & 0.1000 & 0.1500 & 0.1500 & 0.1000 \\ 0.3636 & 0.4091 & 0.0455 & 0.0000 & 0.1818 \\ 0.1429 & 0.4286 & 0.1429 & 0.1429 & 0.1429 \\ 0.0000 & 0.1000 & 0.2000 & 0.0000 & 0.7000 \\ 0.0000 & 0.1923 & 0.2308 & 0.1923 & 0.3846 \end{bmatrix}$$

February

$$\begin{bmatrix} 0.7857 & 0.0476 & 0.0476 & 0.0238 & 0.0952 \\ 0.2941 & 0.5294 & 0.0588 & 0.0000 & 0.1176 \\ 0.1111 & 0.5556 & 0.1111 & 0.2222 & 0.0000 \\ 0.4000 & 0.2000 & 0.2000 & 0.0000 & 0.2000 \\ 0.1818 & 0.0000 & 0.2727 & 0.1818 & 0.3636 \end{bmatrix}$$

March

$$\begin{bmatrix} 0.8308 & 0.0923 & 0.0462 & 0.0154 & 0.0154 \\ 0.5385 & 0.1538 & 0.1538 & 0.0000 & 0.1538 \\ 0.3750 & 0.3750 & 0.1250 & 0.0000 & 0.1250 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.2000 & 0.2000 & 0.4000 & 0.0000 & 0.2000 \end{bmatrix}$$

April

$$\begin{bmatrix} 0.8421 & 0.0351 & 0.0702 & 0.0175 & 0.0351 \\ 0.6250 & 0.0000 & 0.0000 & 0.2500 & 0.1250 \\ 0.1667 & 0.5000 & 0.0000 & 0.1667 & 0.1667 \\ 0.1667 & 0.1667 & 0.1667 & 0.0000 & 0.5000 \\ 0.0833 & 0.2500 & 0.0833 & 0.1667 & 0.4167 \end{bmatrix}$$

May

$$\begin{bmatrix} 0.8413 & 0.0476 & 0.0635 & 0.0317 & 0.0159 \\ 0.5833 & 0.0833 & 0.1667 & 0.0000 & 0.1667 \\ 0.3333 & 0.4444 & 0.2222 & 0.0000 & 0.0000 \\ 0.0000 & 0.5000 & 0.2500 & 0.2500 & 0.0000 \\ 0.0000 & 0.5000 & 0.0000 & 0.2500 & 0.2500 \end{bmatrix}$$

June

$$\begin{bmatrix} 0.7879 & 0.0303 & 0.0606 & 0.0303 & 0.0909 \\ 0.1333 & 0.2667 & 0.2000 & 0.1333 & 0.2667 \\ 0.1818 & 0.3636 & 0.2727 & 0.1818 & 0.0000 \\ 0.0769 & 0.1538 & 0.0769 & 0.2308 & 0.4615 \\ 0.1176 & 0.1765 & 0.1765 & 0.2941 & 0.2353 \end{bmatrix}$$

July

$$\begin{bmatrix} 0.5833 & 0.2500 & 0.0833 & 0.0833 & 0.0000 \\ 0.0769 & 0.5641 & 0.2308 & 0.0769 & 0.0513 \\ 0.0741 & 0.4074 & 0.4444 & 0.0370 & 0.0370 \\ 0.0000 & 0.4286 & 0.1429 & 0.1429 & 0.2857 \\ 0.0000 & 0.0000 & 0.5714 & 0.1429 & 0.2857 \end{bmatrix}$$

August

$$\begin{bmatrix} 0.3333 & 0.4444 & 0.0000 & 0.0000 & 0.2222 \\ 0.2143 & 0.4762 & 0.2143 & 0.0000 & 0.0952 \\ 0.1333 & 0.5333 & 0.2667 & 0.0000 & 0.0667 \\ 0.2500 & 0.5000 & 0.0000 & 0.2500 & 0.0000 \\ 0.0000 & 0.2308 & 0.2308 & 0.2308 & 0.3077 \end{bmatrix}$$

September

$$\begin{bmatrix} 0.0000 & 0.3750 & 0.5000 & 0.0000 & 0.1250 \\ 0.2000 & 0.3000 & 0.3333 & 0.1000 & 0.0667 \\ 0.0400 & 0.4800 & 0.2400 & 0.1200 & 0.1200 \\ 0.0714 & 0.4286 & 0.2857 & 0.1429 & 0.0714 \\ 0.0000 & 0.0000 & 0.0833 & 0.5000 & 0.4167 \end{bmatrix}$$

October

$$\begin{bmatrix} 0.3750 & 0.1875 & 0.0625 & 0.0000 & 0.3750 \\ 0.2059 & 0.4412 & 0.2941 & 0.0000 & 0.0588 \\ 0.0952 & 0.6190 & 0.1905 & 0.0000 & 0.0952 \\ 0.0000 & 0.3333 & 0.6667 & 0.0000 & 0.0000 \\ 0.0556 & 0.0556 & 0.2778 & 0.1667 & 0.4444 \end{bmatrix}$$

November

$$\begin{bmatrix} 0.7857 & 0.0000 & 0.0000 & 0.0714 & 0.1429 \\ 0.0000 & 0.0000 & 0.0000 & 0.3333 & 0.6667 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \\ 0.2000 & 0.3000 & 0.2000 & 0.0000 & 0.3000 \\ 0.0652 & 0.0000 & 0.0000 & 0.1522 & 0.7826 \end{bmatrix}$$

December

$$\begin{bmatrix} 0.0000 & 0.3333 & 0.0000 & 0.0000 & 0.6667 \\ 0.0000 & 0.0000 & 0.0000 & 0.6667 & 0.3333 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.1667 & 0.1667 & 0.1667 & 0.5000 & 0.0000 \\ 0.0141 & 0.0000 & 0.0141 & 0.0563 & 0.9155 \end{bmatrix}$$

Figure 2. Transition Probability Matrix in Tanah Merah, Kelantan for the years 2018 to 2020

As a result, the periodicity of the MC is aperiodic since the period value is 1 for each month. If the return value for the character vector of the recurrent state is equivalent to 1, it is concluded that the MC has recurrent states. Since the communicating class value for each month are equal to 1, it revealed a list for each slot that has the names of the states returned in the communicating class. Thus, it is confirmed that the MC communicates with each other and is irreducible because the chain follows all the properties of communication, which are reflexivity, symmetry, and transitivity. In conclusion, the MC is ergodic as the MC is aperiodic, irreducible and has recurrent states for every month. It is proven that the transition probability matrix is stationary, and the limiting distribution exists.

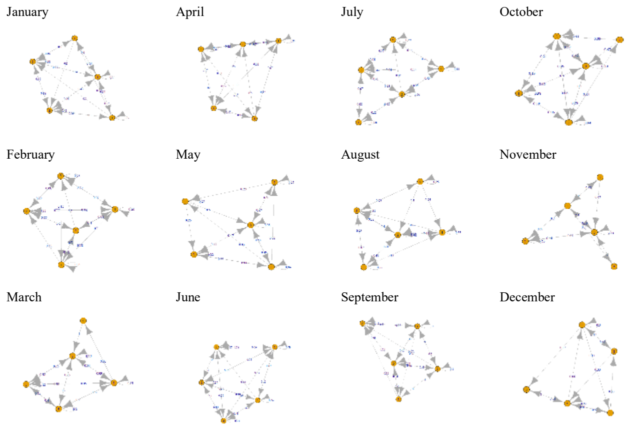

Figure 3. Transition Probability Diagram in Tanah Merah, Kelantan for the years 2018 to 2020

Table 10. Classification of classes of data for each month in 2018 to 2020

States	Classes	Return value
Irreducible	Aperiodic	Period = 1, Absorbing states = (0)
Periodicity	Aperiodic	Period = 1
Recurrent States	Has recurrent states	Recurrent {1,2,3,4,5}, *fi*=1
Transient States	Has no transient states	Transient = (0)
Communication	The class is communicated with each other	Communicating Class = 1
Conclusion	The Markov chain is ergodic	

The probability of long-run behaviour for each month beginning in January and ending in December for 2018 through 2020 is displayed in the Table 11. Five states classify rainfall, with State 1 indicating no rain and State 2 meaning light rain. States 3, 4, and 5 denote moderate, heavy, and intense rain, respectively.

The probability represents the long-run behaviour for each month for the future year. The likelihood of limiting distribution for each state is obtained by calculating each of the limiting distributions of the MC, $\pi = (\pi 1, \pi 2, \pi 3, \pi 4, \pi 5)$ in each transition probability matrix for each month. The Limiting Distribution Matrix (LDM) table represents the likelihood of a fixed probabilities value. Each row in the LDM table reflects the limiting distribution probabilities of each state, commonly known as the long-run probabilities.

As shown in the Table 11, the limiting distribution probability matrix for January is [0.2256, 0.2515, 0.1495, 0.1065, 0.2668], indicating that the rainfall pattern will follow the probability pattern in the future. Thus, regardless of the weather on any given day in January, the likelihood of no rain is 0.2256, while the probabilities of light rain, moderate rain, heavy rain, and heavy rain are 0.2515, 0.1495, 0.1065, and 0.2668, respectively. For February, the limiting distribution probability matrix is [0.5274, 0.1890, 0.0942, 0.0575, 0.1319], showing that the average rainfall might follow the probability distribution in the foreseeable. Hence, the probability for no rain, light rain, moderate rain, heavy rain, and extreme rain is on any particular day in February is 0.5274, 0.1890, 0.0942, 0.0575, and 0.1319 correspondingly, while March's limiting distribution probability matrix is [0.7065, 0.1413, 0.0870, 0.0109, 0.0543], implying that future rainfall patterns will follow the probability pattern. Similarly, the chance of no rainfall on any given day in March is 0.7065. In contrast, the probabilities of light rain, moderate rain, and heavy rain are 0.1413, 0.0870, 0.0109, and 0.0543, respectively, regardless of the weather.

In addition, for the next three months, which is April, May and June, the limiting distribution probability matrix for these three months that implying in the future, the rainfall pattern will follow the likelihood pattern are [0.6216, 0.1018, 0.0670, 0.0707, 0.1389], [0.6848, 0.1304, 0.0978, 0.0435, 0.0435], and [0.3726, 0.1584, 0.1355, 0.1459, 0.1876]. Thus, the probability of no rain on any given day in these three months is 0.6216, 0.6848, and 0.3726, while the possibilities of light rain are 0.1018, 0.1304, and 0.1584. Afterwards, the probability of moderate rain in April, May, and June are 0.0670, 0.0978, and 0.1355. The likelihood of heavy rain for each month is 0.0707, 0.0435, and 0.1459. Lastly, the chances of extreme rain regardless of the weather for these three months are 0.1389, 0.0435 and 0.1876.

Then, for the months in July, August, and September, the limiting distribution probability matrix is [0.1304, 0.4239, 0.2935, 0.0761, 0.0761], [0.1947, 0.4465, 0.1748, 0.0433, 0.1407] and [0.0899, 0.3371, 0.2809, 0.1573, 0.1348]. It means that in the long run, the no rain chances for each month will be 0.1304, 0.1947, and 0.0899. The light rain probabilities will be 0.4239, 0.4465, and 0.3371 each month. Following that, the medium rain probabilities for each month will be 0.2935, 0.1748, and 0.2809. The heavy rain probability for July, August, and September will be 0.0761, 0.0433, and 0.1573, respectively, while the extreme rain probabilities for the same months will be 0.0761, 0.1407, and 0.1348.

Finally, for the months that are in October, November, and December, the limiting distribution probability matrix is [0.1726, 0.3609, 0.2384, 0.0326, 0.1956], [0.0566, 0.3611, 0.1194, 0.1279, 0.3350], and [0.2179, 0.4136, 0.0921, 0.0558, 0.2206]. It indicates that the long-term chances of no rain are 0.1726, 0.0566, and 0.2179 for each month. The probability of light rain will be 0.3609, 0.3611, and 0.4136. Then, the monthly medium rain probabilities will be 0.2384, 0.1194, and 0.0921. In the same future months for October, November, and December, heavy rain chances are 0.1956, 0.3350, and 0.2206, respectively. Meanwhile, heavy rain chances are 0.1956, 0.3350, and 0.2206.

Table 11. Probability of long run behaviour of rainfall for each month

Month	Probability
January	$\pi = [0.2256, 0.2515, 0.1495, 0.1065, 0.2668]$
February	$\pi = [0.5274, 0.1890, 0.0942, 0.0575, 0.1319]$
March	$\pi = [0.7065, 0.1413, 0.0870, 0.0109, 0.0543]$
April	$\pi = [0.6216, 0.1018, 0.0670, 0.0707, 0.1389]$
May	$\pi = [0.6848, 0.1304, 0.0978, 0.0435, 0.0435]$
June	$\pi = [0.3726, 0.1584, 0.1355, 0.1459, 0.1876]$
July	$\pi = [0.1304, 0.4239, 0.2935, 0.0761, 0.0761]$
August	$\pi = [0.1947, 0.4465, 0.1748, 0.0433, 0.1407]$
September	$\pi = [0.0899, 0.3371, 0.2809, 0.1573, 0.1348]$
October	$\pi = [0.1726, 0.3609, 0.2384, 0.0326, 0.1956]$
November	$\pi = [0.0566, 0.3611, 0.1194, 0.1279, 0.3350]$
December	$\pi = [0.2179, 0.4136, 0.0921, 0.0558, 0.2206]$

Based on the RM constructed below, for State 1, which has no rain, most of the month has a very low occurrence of no rain in July, August, September, October, and November, which will lead to high chances of flood. Thus, the red colour for warning and danger indicates the high occurrence of flood hazards as the flood level rises to near flooding levels. The flood level has reached critical levels, necessitating evacuation. January, June and December have a low occurrence of no rain. The colour for everyday alert and warning are orange because there is a possibility that flood hazard will occur, and the water level is above the safe level.

Meanwhile, February is the only month with moderate occurrence for no rain, which means there is a low possibility for the flood to occur. In the remaining month, which is March, April, and May, no rain is high. Therefore, there are minimal chances of a flood event because the colour for alert, warning, and danger are green, which indicates a very low occurrence of flood hazard, and the flood level is at a safe level.

Next, State 2 is light rain. This state also has the most significant number of the month with a very low occurrence, same as State 1, but State 2 for light rain. February, March, April, May, and June have a very low event of light rain, implying a moderate probability of flooding due to rainfall intensity. January, September, October, and November have the same light rain occurrence, which is low. The moderate light rain occurrence is in July, August, and December. The light rain for the occurrence of low and moderate both have very low chances a flood event to occur. There is no existence of a high occurrence of light rain.

Moreover, moderate rain is labelled as State 3. There are only two types of flood impact in this state: low and low. Nine over twelve months have very low flood impact: January, February, March, April, May, June, August, November, and December. The normal and alert have the likelihood of moderate occurrence of flood event, which is in orange colour. The warning and danger are in yellow as low flood hazards. July, September, and October have low flood impacts for moderate rain. The colour for normal, alert, warning and danger is the opposite colour for very low, which means that the colour for normal and alert are yellow, and the colour for warning and danger is orange.

Furthermore, State 4 is indicated as heavy rain. All the months, including January, February, March, April, May, June, July, August, September, October, November and December, have the same range of the limiting distribution probability, which is less than 0.20, indicating the very low occurrence of heavy rain. In all months, the likelihood of the flood level has already been reduced to a safe level is high, and the flood level is increasing to near flooding level is low. Hence, the flood event's chances to occur are low for all months.

Lastly, for State 5, which is extreme rain. The result for State 5 is very similar to State 3 because there is only low and very low for flood impact. Starting with February, March, April, May, June, July, August, and September, this month has a very low extreme rainfall occurrence. The colour for the warning and danger shows that there will be a very low flood event to happen. In other months, such as January, October, November and December, extreme rainfall is low. Thus, the possibility of flood will be low because the chances of the flood level in a safe level are high.

Table 12. Monthly Rainfall Matrix

Januari Rainfall Matrix

Flood Impact

Rainfall State	Normal	Alert	Warning	Danger
1 (0.23)	Moderate	Moderate	Moderate	High
2 (0.25)	Moderate	Low	Very Low	Very Low
3 (0.15)	Moderate	Moderate	Low	Low
4 (0.11)	High	Very Low	Low	Low
5 (0.27)	High	Low	Low	Low

February Rainfall Matrix

Flood Impact

Rainfall State	Normal	Alert	Warning	Danger
1 (0.53)	Low	Low	Very Low	Very Low
2 (0.19)	Low	Low	Moderate	Moderate
3 (0.09)	Moderate	Moderate	Low	Low
4 (0.06)	High	Very Low	Low	Low
5 (0.13)	High	High	Very Low	Very Low

March Rainfall Matrix

Flood Impact

Rainfall State	Normal	Alert	Warning	Danger
1 (0.71)	High	Very Low	Very Low	Very Low
2 (0.14)	Low	Low	Moderate	Moderate
3 (0.08)	Moderate	Moderate	Low	Low
4 (0.01)	High	Very Low	Low	Low
5 (0.05)	High	High	Very Low	Very Low

April Rainfall Matrix

Flood Impact

Rainfall State	Normal	Alert	Warning	Danger
1 (0.62)	High	Very Low	Very Low	Very Low
2 (0.10)	Low	Low	Moderate	Moderate
3 (0.07)	Moderate	Moderate	Low	Low
4 (0.07)	High	Very Low	Low	Low
5 (0.14)	High	High	Very Low	Very Low

May Rainfall Matrix

Flood Impact

Rainfall State	Normal	Alert	Warning	Danger
1 (0.67)	High	Very Low	Very Low	Very Low
2 (0.13)	Low	Low	Moderate	Moderate
3 (0.10)	Moderate	Moderate	Low	Low
4 (0.04)	High	Very Low	Low	Low
5 (0.04)	High	High	Very Low	Very Low

June Rainfall Matrix

Flood Impact

Rainfall State	Normal	Alert	Warning	Danger
1 (0.37)	Moderate	Moderate	Moderate	High
2 (0.16)	Low	Low	Moderate	Moderate
3 (0.14)	Moderate	Moderate	Low	Low
4 (0.15)	High	Very Low	Low	Low
5 (0.19)	High	High	Very Low	Very Low

July Rainfall Matrix

Flood Impact

Rainfall State	Normal	Alert	Warning	Danger
1 (0.13)	Very Low	Very Low	High	High
2 (0.42)	Moderate	Very Low	Very Low	Very Low
3 (0.30)	Low	Low	Moderate	Moderate
4 (0.08)	High	Very Low	Low	Low
5 (0.08)	High	High	Very Low	Very Low

August Rainfall Matrix

Flood Impact

Rainfall State	Normal	Alert	Warning	Danger
1 (0.19)	Very Low	Very Low	High	High
2 (0.45)	Moderate	Very Low	Very Low	Very Low
3 (0.17)	Moderate	Moderate	Low	Low
4 (0.04)	High	Very Low	Low	Low
5 (0.14)	High	High	Very Low	Very Low

September Rainfall Matrix

Flood Impact

Rainfall State	Normal	Alert	Warning	Danger
1 (0.09)	Low	Low	Very Low	Very Low
2 (0.34)	Low	Low	Moderate	Moderate
3 (0.28)	Moderate	Moderate	Low	Low
4 (0.16)	High	Very Low	Low	Low
5 (0.13)	High	High	Very Low	Very Low

October Rainfall Matrix

Flood Impact

Rainfall State	Normal	Alert	Warning	Danger
1 (0.17)	Low	Low	Very Low	Very Low
2 (0.36)	Low	Low	Moderate	Moderate
3 (0.24)	Moderate	Moderate	Low	Low
4 (0.03)	High	Very Low	Low	Low
5 (0.20)	High	High	Very Low	Very Low

November Rainfall Matrix

Flood Impact

Rainfall State	Normal	Alert	Warning	Danger
1 (0.06)	Low	Low	Very Low	Very Low
2 (0.36)	Low	Low	Moderate	Moderate
3 (0.12)	Moderate	Moderate	Low	Low
4 (0.13)	High	Very Low	Low	Low
5 (0.34)	High	High	Very Low	Very Low

December Rainfall Matrix

Flood Impact

Rainfall State	Normal	Alert	Warning	Danger
1 (0.22)	Low	Low	Very Low	Very Low
2 (0.41)	Low	Low	Moderate	Moderate
3 (0.09)	Moderate	Moderate	Low	Low
4 (0.06)	High	Very Low	Low	Low
5 (0.22)	High	High	Very Low	Very Low

IV. CONCLUSION

To sum up the preceding, the probability transition matrix is essential to indicate each state's probabilities. The transition probability matrix obtained is used to analyse the classes in the transition probability matrix. The limiting distribution probability can also be calculated using this transition probability matrix. Then, the transition probability matrix analysis is also crucial to obtain the ergodic MC as the ergodic MC can confirm the stationarity of the transition probability matrix to show the limiting distribution in the transition probability matrix exist.

Next, the long-run behaviour of the rainfall in Tanah Merah shows that heavy rain has the lowest probability of occurrence in that year. The limiting distribution method was executed to develop the future possibility for each state of the transition probability matrix that has been obtained. Other than heavy rain, the output established from the limiting distribution method demonstrates that the probability of no rain is the highest among the other states in those years, followed by moderate rain and extreme rain. However, the frequency of rainfall varies by state in accordance with the limited distribution probability.

Lastly, the RM can be used to interpret the long-run behaviour of rainfall probability in Tanah Merah into an understandable visualisation matrix. In other words, by using the likelihood of long-run behaviour, the RM for each month is generated to show the impact of rainfall probability for each state that can lead to the flood hazard in the upcoming month. Therefore, RM can be used by the government as one of their initial actions to forecast future floods.

V. ACKNOWLEDGEMENT

This paper is not supported by any external grant, and the authors are greatly appreciated reviewers for constructive comments.

VI. REFERENCES

Abeywardena, V 1955, Rainfall and Crops (Lecture to Trainees on Coconut Cultivation).

Bopi, NB, Ismail, AF, Zabidi, SA & Hasan, MK 2016, 'Development of a Rainfall Rate Monitoring System for Malaysia', Indian Journal of Science and Technology, vol. 9, p. 28.

da Silva Jale, J, Júnior, SFAX, Xavier, ÉFM, Stošić, T, Stošić, B & Ferreira, TAE 2019, 'Application of Markov chain on daily rainfall data in Paraíba-Brazil from 1995-2015', Acta Scientiarum, Technology, vol. 41, p. e37186.

Holmes, J & Hassini, S 2021, 'Discrete-time Markov chain modelling of the ontario air quality health index', Water, Air, & Soil Pollution, vol. 232, no. 4, p. 158.

Joseph, J & Ratheesh, TK 2013, 'Rainfall Prediction Using Data Mining Techniques', International Journal of Computer Applications, vol. 83, no. 8.

Karlin, S 2014, 'A first course in stochastic processes', Academic Press.

Mahanta, J, Dey, S & Khosro, P 2018, 'Analyzing Rainfall Condition of Bangladesh: An Application of Markov Chain', Thailand Statistician, vol. 16, no. 2, pp. 203-212.

Meza-Pale, P & Yunez-Naude, A 2015, The Effect of Rainfall Variation on Agricultural Households: Evidence from Mexico (No. 1008-2016-79827).

Ng, JL, Abd Aziz, S, Huang, YF, Wayayok, A & Rowshon, MK 2016, 'Analysis of annual maximum rainfall in Kelantan, Malaysia', in III International Conference on Agricultural and Food Engineering 1152, pp. 11-18.

Nop, C, Fadhil, RM & Unami, K 2021, 'A multi-state Markov chain model for rainfall to be used in optimal operation of rainwater harvesting systems', Journal of Cleaner Production, vol. 285, p. 124912.

Parmar, A, Mistree, K & Sompura, M 2017, 'Machine Learning Techniques for Rainfall Prediction: A Review', in International Conference on Innovations in Information Embedded and Communication Systems, vol. 3.

Rohith, AN, Gitau, MW, Chaubey, I & Sudheer, KP 2021, 'A multistate first-order Markov model for modeling time distribution of extreme rainfall events', Stochastic Environmental Research and Risk Assessment, vol. 35, pp. 1205-1221.

Ross, Sheldon M 2014, 'Introduction to Probability Models', Academic Press, 2014.

Setiawan, I & Ilhamsyah, Y 2020, 'Rainfall Analysis in the Indian Ocean by using 6-States Markov Chain Model', IOP Conference Series: Earth and Environmental Science.

Telipot, M 2000, 'Monthly and Annual Rainfall Variability in Peninsular Malaysia', Pertanika Journal, p. 73.

Yakasiri, M, Avrel, J, Sharma, S, Anuradha, M & Keshavan, BK 2019, 'A stochastic approach for the state-wise forecast of wind speed using discrete-time markov chain'. In TENCON 2019-2019 IEEE Region 10 Conference (TENCON) IEEE, pp. 575-580.