Application of the stochastic transition matrix as a prediction pattern for the occurrence of disasters in Aceh province in 2017-2021

Siti Rusdiana¹, Asep Rusyana², T.Raja Ahmad Alfaruqi³

¹Department of Mathematics, Syiah Kuala University, Banda Aceh, Indonesia
²Department of Statistics, Syiah Kuala University, Banda Aceh, Indonesia
³Department of Mathematics, Syiah Kuala University, Banda Aceh, Indonesia
siti.rusdiana@unsyiah.ac.id

Abstract. Disaster is a series of events that threaten people's lives. The causes can be from natural factors and human factors that can cause casualties, environmental damage, property losses and psychological impacts. Historically, there have been many disasters in Indonesia, especially in the Aceh province, which has stimulated the government to produce policies of reducing the incidence of disasters, especially disasters caused by human error. The number of floods, fires and landslides occurring annually tends to be similar. Unlike the catastrophic tidal waves, tornadoes, earthquakes, and droughts, such disasters occur randomly and some never occur. Therefore, a method is needed to predict disaster event opportunities, one of which is the Markov chain method. To find out the predicted value of disaster opportunities in the Aceh province for 2017-2021. Based on the results of the study, it is known that the predicted probability of disasters in the province of Aceh in 2017-2021 does not increase and does not decrease in extreme ways from year to year, the accuracy of the predictive value of opportunities for disasters caused by humans (landslides, fires, and floods) in the Province of Aceh in 2017-2021 are classified as reasonable with error value of 26.93%.

1. Introduction
Disasters are divided into two, namely natural disasters and disasters caused by human activities. Natural disasters are disasters caused by events or a series of events caused by Landslides encompass a wide range of phenomena including slumps, rock falls, debris slides, and earth-, debris- and mud-flows. Disasters caused by human activities are disasters caused by human error such as flooding caused by the habit of throwing garbage in rivers, the fires caused by human error and landslides due to deforestation. (Forbes K 2011).

At this time disasters often occur in Aceh, especially disasters caused by human activity. By looking at the disasters that occurred in previous years, the province of Aceh can be said to be a disaster-prone province. Therefore, by looking at the value of the likelihood of future disasters, the government can find out how big the effect of the disaster management program that has been previously designed on each disaster in the future. One of methods on disaster predictions for the later years by using a transition probability matrix. Transition probability matrix is a matrix that contains information that regulates the movement of the system from one state to another (Langi, 2009).
2. Methods

Aceh Province is located between 01° 58' 37.2" - 06° 04' 33.6" north latitude and 94° 57' 57.6" - 98° 17' 13.2" east longitude with an average height of 125 meters above sea level. The disaster that occurred in the land of "Rencong" has been recapitulated into a book with the title Aceh in Numbers published by the Central Statistics Agency of Aceh Province. Following are the data on disasters caused by human activities, which was speculated by BPS Aceh from 2008-2017.

Table 1. Data on human-caused disasters that occurred in Aceh province from 2012-2017

| Year | Avalanche | Fire | Flood | Total |
|------|-----------|------|-------|-------|
| 2012 | 5         | 86   | 26    | 117   |
| 2013 | 67        | 441  | 192   | 700   |
| 2014 | 31        | 374  | 55    | 460   |
| 2015 | 12        | 20   | 65    | 97    |
| 2016 | 9         | 67   | 30    | 106   |
| 2017 | 2         | 50   | 32    | 84    |

(Source: Book of Aceh in Figures for 2012-2018 by BPS Aceh Provinci-Indonesia)

In this study, the data above will be tested whether there is a relationship between disaster and year variables using the chi-square method as follows:

\[ X^2 = \sum_{i=1}^{n} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \]  

(1)

where:

- \( X^2 \) = Distribution of Chi-square
- \( O_{ij} \) = The value of observation \((i-th \text{ observation})\)
- \( E_{ij} \) = \(i-th \) Expectation value.

\[ E_{ij} = \frac{\text{Total } i-th \text{ row value } \times \text{ total } i-th \text{ colom value}}{\text{total all values}} \]

To determine testing criterion, if \( X^2_{\text{count}} \leq X^2_{\text{table}} \), so \( H_0 \) accepted. It does the opposite, if \( X^2_{\text{count}} > X^2_{\text{table}} \), so reject \( H_0 \) (Ross, 2010).

Then the data is processed into a transition probability matrix. The model built from the disaster data summarized above is as follows:

\[ P = \sum_{j=1}^{3} p_{ij}, i = 1,2,3 \]

\[ P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \]  

(2)
In this case $P$ is a transition probability matrix with the conditions for each element of the matrix must be positive and because the matrix is a probability matrix, the sum of each row is $I$. The $P_{ij}$ matrix if notated into mathematical form is as follows:

$$P = [p_{ij}] \ \forall \ i,j = 1,2,3$$

where:

- $i = 1 = \text{in 2012}$ and $j = 1 = \text{probability of landslide disaster}$
- $i = 2 = \text{in 2013}$ and $j = 2 = \text{probability of fire disaster}$
- $i = 3 = \text{in 2014}$ and $j = 3 = \text{probability of flood disaster}$

So that $p_{ij}$ is probability for a disaster to occur in year $i$. In this research, a vector state is needed as an initial event matrix, where the initial vector state used contains the opportunities for a disaster event in 2015 which are modeled as follows:

$$x_0 = [x_1 \ x_2 \ x_3]$$

where:

- $x_0$ = the probability of disaster in 2015
- $x_1$ = Landslide disaster
- $x_2$ = Fire disaster
- $x_3$ = Flood disaster

To get the chance of each type of disaster in 2016 then:

$$x_1 = x_0 P^1$$

where:

- $x_1$ = the probability of disaster in 2016
- $x_0$ = the probability of disaster in 2015
- $P^1$ = transition probability matrix.

Further to obtain the occurrence of each type of disaster in 2017-2021 then:

$$x_n = x_{n-1} P^n ; n = 1,2,3,...,6$$

where:

- $x_n$ = the probability matrix of the type of disaster of the year $n$
- $P^n$ = the transition probability matrix

with $n = 1,2,3,...,6$ where

- $n = 1 = 2016$
- $n = 2 = 2017$
- $n = 3 = 2018$
- $n = 4 = 2019$
- $n = 5 = 2020$
- $n = 6 = 2021$
By multiplying line vectors and the transition probability matrix we get a new row vector. This method is called the Markov Chain method. To compute the accuracy of this method, it is necessary to compare using the MAPE method as follows:

\[ MAPE = \frac{1}{7} \sum \left| \frac{A_t - F_t}{A_t} \times 100\% \right| \]  

(6)

where:
- \( A_t \) = the actual value of the probability of a disaster occurrence \( t \)-th years in 2016
- \( F_t \) = the predictive value of 2016 disaster relief opportunities

with \( t = 1, 2, 3 \)
- \( t = 1 \) = Landslide disaster
- \( t = 2 \) = Fire disaster
- \( t = 3 \) = Flood disaster

3. Results and Discussion

3.1. Testing the Disaster Data Hypothesis in Aceh Province 2015-2016 Using the Chi-Square Method.
To see the effect / relation of the disaster to the year, the statements for \( H_0 \) and \( H_1 \) are as follows:
- \( H_0 \) = There is no correlation between the disaster variable and the year variable
- \( H_1 \) = There is a correlation between the disaster variable and the year variable

The following are the results of testing the hypothesis:

**Table 2** The Observation Value Of Disaster In Aceh Province In 2015-2016

| YEARS | Category of Disasters | Total |
|-------|-----------------------|-------|
|       | Landslides | Fires | Floods |       |
| 2015  | 12         | 20    | 65     | 97    |
| 2016  | 9          | 67    | 30     | 106   |
| Total | 21         | 87    | 95     | 203   |

**Table 3** The Expected Value Of Disaster In Aceh Province In 2015-2016

| YEARS | Category of Disasters | Total |
|-------|-----------------------|-------|
|       | Landslides | Fires | Floods |       |
| 2015  | 10.03       | 41.57 | 45.39  | 97    |
| 2016  | 10.97       | 45.43 | 49.61  | 106   |
| Total | 21          | 87    | 95     | 203   |

Using equation (1), the calculation is displayed in tabular form as follows:

\[
Total X^2_{value} = X^2_{value \ landslides} + X^2_{value \ fires} + X^2_{value \ floods}
\]

\[ = 0.74 + 21.44 + 16.22 \]

\[ = 38.4 \]

With the foregoing that \( X^2_{table} = 5.59 \) then \( X^2_{value} > X^2_{table} \) which means there is a link between the disaster variable and the year variable (reject \( H_0 \)).
3.2 Building A Transition Probability Matrix and Initial Vector State

To create an initial transition and vector state probability matrix, the disaster data table from 2012-2016 is needed, which is converted into the probability table as follows:

| Year | Landslides | Fires | Floods | Total |
|------|------------|-------|--------|-------|
| 2012 | 0.0427     | 0.7350| 0.2222 | 1     |
| 2013 | 0.0957     | 0.6300| 0.2743 | 1     |
| 2014 | 0.0674     | 0.8130| 0.1196 | 1     |
| 2015 | 0.1237     | 0.2062| 0.6701 | 1     |
| 2016 | 0.0238     | 0.5952| 0.3809 | 1     |

The transition probability matrix will be formed by taking data from 2012 to 2014 and then changing it to the $P = [p_{ij}]$ matrix while 2015 data will be used as the initial vector state $x_0 = [x_i]$ as follows:

$$P = \begin{bmatrix} 0.0427 & 0.735 & 0.222 \\ 0.0957 & 0.630 & 0.2743 \\ 0.0674 & 0.813 & 0.1196 \end{bmatrix}$$

$$x_0 = [0.1237, 0.2062, 0.6701]$$

3.3. The Analysis of Predictions of Opportunities for Disasters Due to Human Action in 2016-2021

The predicted value of probability of disaster is searched by looking at the value of each vector state as in equation (5). Using equation (5), the following table predictions the results using the Markov Chain method:

| Vector State | Year | The Type of Disasters |
|--------------|------|-----------------------|
| $x_1$        | 2016 | 7.01% 76.57% 16.42% 100% |
| $x_2$        | 2017 | 8.73% 66.73% 24.52% 100% |
| $x_3$        | 2018 | 8.41% 68.39% 23.17% 100% |
| $x_4$        | 2019 | 8.46% 68.11% 23.39% 100% |
| $x_5$        | 2020 | 8.45% 68.15% 23.36% 100% |
| $x_6$        | 2021 | 8.46% 68.15% 23.37% 100% |

It shows that fire is the highest disaster prediction value which means that fire is the most disaster compared to other disasters for the following year. Landslides become disasters with the lowest predicted value of disasters each year. Also seen the predicted value of flood disasters rose by 8.07% in 2016-2017 and then settled at 23% in 2017 to 2022.
3.4. Measuring the Accuracy of Prediction Value of Chances of Disasters Caused by Human Action in 2016-2021

The results that have been obtained must be tested for accuracy using the MAPE method in equation (8). Field data for 2016 will be compared with prediction data for 2016 using MAPE as shown in the table as follows:

Table.6 Accuracy Measures of Predictions Value of Disasters Due to Human Action by MAPE Method in 2016

| The Type of Disasters | Actual Predictions for 2016 (A_t) | 2016 Markov Chain Disaster Prediction (F_t) | A_t - F_t | A_t - F_t / A_t × 100% |
|-----------------------|----------------------------------|---------------------------------------------|-----------|--------------------------|
| Landslides            | 0,0849                           | 0,0701                                      | 0,0148    | 0,1743                   | 17,43%                     |
| Fires                 | 0,6320                           | 0,7657                                      | -0,1337   | -0,2116                  | 21,16%                     |
| Floods               | 0,2830                           | 0,1642                                      | 0,1188    | 0,4198                   | 41,98%                     |
| Sum                  | 1                                | 1                                           | -0,0001   | 0,3826                   | 80,57%                     |

\[
MAPE = \frac{1}{3} \times 80,57\% = 26,86\%
\]

4. Conclusion

The number of possible disasters each year tends to decrease from 2016 to 2018 and then rises slightly in 2019 and reaches a steady state in 2020 to n. So the steady state value for landslides is 8,472%, fires are 68,171%, and floods are 23,356%. Fire is a disaster with a very high probability of occurrence every year which is around 68%. Landslide is a disaster with the lowest predictive value with a value of around 8%. If in one year it is certain that at least 1 time a disaster has occurred, the disaster that is more likely to occur is a fire.

5. References

[1] Indrianti, N., Ming, T., Toha., I.S. 2001. Model Perencanaan Kebutuhan Bahan dengan Mempertimbangkan Waktu Kadaluarsa Bahan. Media Teknik 2: 60-65.
[2] Amri MR, Yulianti G, Yunus R, Wiguna S, Ady AW, Ichwana AN, Randongkir RE, Septian RT, 2010 Risiko Bencana Indonesia BNPB Jakarta
[3] Bappeda Aceh 2012 Aceh Dalam Angka Edisi 2012 Bappeda Aceh, Aceh
[4] Bappeda Aceh 2013 Aceh Dalam Angka Edisi 2012 Bappeda Aceh, Aceh
[5] Bappeda Aceh 2014 Aceh Dalam Angka Edisi 2012 Bappeda Aceh, Aceh
[6] Bappeda Aceh 2015-2016 Aceh Dalam Angka Edisi 2012 Bappeda Aceh, Aceh
[7] Forbes K and Broadhead J 2011. The role of trees and forests in the prevention of landslides and rehabilitation in Asia, RAP PUBLICATION Vol 19 pg. 1–7.
[8] Langi, Yohanes. 2009. Penentuan Klasifikasi State pada Rantai Markov. Jurnal Ilmiah Sains, volume 9(1) hal.63-67.
[9] Ross, M. Sheldon 2010 Introduction to Probability Models Tenth Edition. Elsevier, USA
[10] Syarifuddin S, Irma S, dan Sukarna 2014 Aplikasi Analisis Rantai Markov Untuk Memprediksi Status Pasien Rumah Sakit Umum Daerah Kabupaten Barru Online Jurnal of Natural Science Vol.3 pg. 313 – 321.