Numerical Analysis using Empirical Orthogonal Function Based on Multivariate Singular Value Decomposition on Indonesian Forest Fire Signal

P Septiawan¹, S Nurdiati¹ and A Sopaheluwakan²

¹Mathematics Department, Faculty of Mathematics and Natural Science, Bogor Agricultural University, Bogor, 16680, Indonesia
²Research and Development Center, Meteorology, Climatology and Geophysics Agency, Jakarta, 10720, Indonesia

Email: Septiawan.pandu@gmail.com

Abstract. Over the past years, fires have burn Indonesian forest annually. The strength of the fire and many of fire points is vary each year. In 1997 there was 9.75 million ha forest burned in one year with 206.6 million ton of carbon emissions. In 2015 there was 2.089 million ha forest and 805 million tons carbon emissions. Both years involved the biggest incidences of burned forest in Indonesia. The purpose of this research is to analyze the pattern of burned forest in Indonesia and to analyze linkages of variables that affect it. Datasets that used in this research is monthly data from 1997 until 2016 Global Fires Emissions Database (GFED) with variable burned area, GFEDs with variable carbon emissions, and Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS). Using Singular Value Decomposition and Empirical Orthogonal Function analysis, for the overlapping period detected that there is annual signal that dominate the incidences of fires in Indonesia, especially in Riau, Palembang, South Kalimantan and Central Kalimantan. There was 4 annual signal that can explain 72% of fires incidences in Indonesia in 1997-2016. In Riau there is another signal that have 6 months period caused by 6 months rainfall period.

Keywords : fire, GFEDs, empirical orthogonal function, singular value decomposition.

1. Introduction

Forest fires are one of the impacts of increasing levels of pressure on forest resources. The impacts associated with forest or land fires are environmental damage and pollution, such as damage to flora and fauna, soil and water. There are several parameters that affect forest fires, including carbon emissions and rainfall. In this research focused on seasonal rainfall in order to see seasonal patterns of forest fire patterns in Indonesia.

In a previous research entitled "Precipitation-fire linkages in Indonesia (1997-2015)" by Thierry Fanin and Guido R. van der Werf examined the relationship between rainfall and forest fires in Indonesia. This research used 2 datasets namely the Global Fire Emissions Database (GFED) and the Tropical Rainfall Measurement Mission (TRMM) 3B42 in 1997-2015 and focused on the Sumatra and Kalimantan regions. The dataset used in the research is a daily dataset that focuses on night time to prove its consistency in the daily period. In this research, 5 patterns were obtained in 4 regions with different spatial and temporal patterns.

Global Forest Watch Fires shows that more than half of Indonesia's forest fires occur on peatlands, which are found in Central Sumatra, South Sumatra, Central Kalimantan, South Kalimantan and Papua.
In 2015, there were 96,937 hotspots detected in Indonesia, with total emissions in Indonesia of 1,403 million tons of carbon. Based on the number of hotspots and carbon emissions, we can predict or look for patterns about forest fires [6].

The aim of this research is to analyze burned land, carbon emissions and rainfall data in Indonesia using combined Empirical Orthogonal Function method based on Singular Value Decomposition (SVD). The next aim is to find the seasonal joint patterns between variables of forest fires, carbon emissions, and rainfall. Then the aim is to analyze the linkages of these three variables to the shared patterns that are formed. This research is expected to provide a very dominant signal that affects forest fires in Indonesia related to rainfall, so that it can be used for future purposes in determining the prediction of forest fires in Indonesia.

2. Types and Data Sources
The first data used in this research is the Global Fire Emissions Database (GFED) by taking the burned area variable. The second data is a small Global Fire Emissions Database (GFEDs) with a small fire fraction to get emissions variable. Both of these data have a spatial resolution of 0.250 x 0.250 and a monthly temporal resolution with HDF format. The third data related to rainfall variable is Climate Hazards Group Infrared Precipitation with Station data (CHIRPS). CHIRPS data has a spatial resolution of 0.050 x 0.050 and a monthly temporal resolution with bil format. Data reduction and transformation in this research were carried out in accordance with the Indonesian domain.

3. Methods
The methods that will be used in this research are combined empirical orthogonal function (EOF) method based on singular value decomposition (SVD) and Fourier transformation.

3.1. Combined Empirical Orthogonal Function based on Singular Value Decomposition
Combined EOF is a generalization of EOF which can be considered as an EOF with combined parameter. Combined EOF is a method that allows simultaneous consideration of modes of different variable variations. One way is to combine several different data by equating the variables owned by both data. Defined a new data set by combining variables in data. The data matrix is then arranged in such a way that the bounded fields at the same time are arranged into the same column [8].

For example:

\[ Z = [Z_1, Z_2, Z_3, ..., Z_n] \]
\[ S = [S_1, S_2, S_3, ..., S_n] \]

Then:

\[ Y_n = \begin{bmatrix} Z_1 & Z_2 & Z_3 & ... & Z_n \\ S_1 & S_2 & S_3 & ... & S_n \end{bmatrix} \]

The SVD-based EOF method is used to reduce the data matrix into computer programs using software. Suppose X is any matrix measuring \( n \times p \) with rank \( (X) = r \). Singular value decomposition (SVD) of \( X \) is a factorization in the form of:

\[ X = U \Sigma V^T \]

with \( U = [u_1, u_2, ..., u_n] \) and \( V = [v_1, v_2, ..., v_n] \) are orthogonal matrix. \( U \) is \( n \times n \) matrix, \( V \) is \( p \times p \) matrix, and \( \Sigma \) is diagonal matrix \( n \times p \) with \( \sigma = \text{diag}(\sigma_1, \sigma_2, ..., \sigma_n) \) dan \( \sigma_1 \geq \sigma_2 \geq ... \geq \sigma_n \geq 0 \). Where \( \sigma_i \) is singular value matrix \( X \). \( U \) and \( V \) are left singular vector and right singular vector \( X \). The variant of \( i \) main component \( i = 1, 2, ..., r \) can be calculated by:

\[ \mu_i = \frac{\sigma_i^2}{\sum_{i=1}^{r} \sigma_i^2} \]
with $\mu_i, i = 1, 2, ..., r$ is the singular value of matrix $X$. In practice, $k$ mode EOF1 or the first major component with $k \ll r$ explains the proportion of the largest variant. EOF2 mode is a linear combination of all observed variables that are orthogonal to EOF1 mode and have the second largest variant and so on. Therefore, the $K$ EOF mode has a maximum variance of $k$ and does not correlate with the previous EOF mode [5].

3.2. Fourier Transformation

Fourier transformation method is used to convert the signal from the time domain to a signal with a frequency domain. Fourier transform is defined as

$$F(v) = \int_{-\infty}^{\infty} f(x)e^{-i2\pi vx} dx$$

$$f(x) = \text{signal with the time domain}$$

$$e^{-i2\pi vx} = \text{kernel function},$$

$$F(v) = \text{signal with a frequency domain},$$

$$v = \text{frequency},$$

$$x = \text{time}.$$

The functions $f(x)$ and $F(v)$ are called the Fourier transform pairs. The first equation is called an advanced transformation (analysis) to transform time to frequency and second equation is called an inverse (synthesis) transformation to transform function from frequency to time domain [11].

4. Research Steps

This stage begins with a literature study related to the method that will be used. SVD-based combined EOF method is used to reduce data matrix into computer programs using software. This method can reduce large data without removing the main information from the data. Reductions were carried out on GFED, GFEDs and CHIRPS data in the Indonesian coverage area. This method aims to determine the dominant patterns in bulk data both spatially and temporally. After that, the patterns that are formed will be analyzed and find the linkages between the three variables used.

CHIRPS data has a smaller grid resolution of $0.05^\circ \times 0.05^\circ$, so before being reduced, grid resolution is equalized by changing the resolution of CHIRPS data to $0.25^\circ \times 0.25^\circ$. The resolution is equalized by taking the average of each 5x5 CHIRPS data in a resolution of $0.05^\circ \times 0.05^\circ$, so that new data will be obtained that reflects the average of rainfall in a resolution of $0.25^\circ \times 0.25^\circ$. After that, the data is removed for data located in the ocean. This is done because the data used is not measured in the sea area, so that the pattern formed can focus on the land area.

After analyzing the entire territory of Indonesia, the analysis continued on the Sumatra and Kalimantan regions. This is done to find out patterns more specifically in areas where frequent forest fires occur annually. Then, the temporal signal formed from each mode will be converted into a periodic signal using the Fourier transform, so that the period of each pattern can be determined from each mode. This period will be used to determine the number of years the forest fires occur in Indonesia.

5. Result and Analysis

Indonesia is a country that has varying periods of rain in its territory. This is because Indonesia is located between two oceans, and has a land topology area that has many mountains. Most parts of Indonesia have a rainy period of twelve months, with the dry season occurring after June. In addition to that, in some regions in Indonesia which are located around the Earth's equator, there is a rainy period with a shorter period of time with six months of rain. For example, at Riau in one year there are two rainy seasons and two dry seasons, from February to March, and July to September.

GFED and GFEDs data are global data with a spatial resolution of $0.25^\circ \times 0.25^\circ$, while the CHIRPS extraction is a $0.05^\circ \times 0.05^\circ$, so the resolution of CHIRPS data needs to be changed to fit GFED and GFEDs data. The data that has been changed in resolution will be used in the overall pattern analysis of the territory of Indonesia, and data that has an initial resolution will be used to observe patterns in the regions of Kalimantan and Sumatra.
Figure 1. Indonesian rainfall distribution map for September 1997 with (a) data resolution of $0.05^\circ \times 0.05^\circ$ and (b) data resolution of $0.25^\circ \times 0.25^\circ$.

In this research the EOF mode obtained from each data matrix analysis is 240 modes. However, the fourth to 240th mode has a pattern similar to previous modes. The only difference lies in the strength of the signal. Therefore, this research only focused on three dominant modes in the analysis of each region. The first mode shows that there are patterns causing forest fires in Central Sumatra, South Sumatra, and South Kalimantan. When viewed from the spatial pattern of carbon emissions in Figure 2, many regions also have large carbon emissions but there are no forest fires. This is caused by the rainfall in the region is also high.

Figure 2. spatial pattern of variables (a) burned area, (b) carbon emission, (c) rainfall, and (d) period of the first temporal mode in Indonesia

Figure 3. spatial pattern and second temporal mode of the Indonesia

Figure 4. spatial pattern and third temporal mode pattern in Indonesia
In 1997-2017, this pattern is always occurred after July which was the time of the dry season in the territory of Indonesia which had a rainy period of twelve months. Based on this information it can be seen that the temporal pattern of forest patterns dominates variable was burned area. From Figure 2 it can be seen that the pattern has a period of six months and twelve months. This is influenced by the Indonesian rainy season which generally has a 12-month period, with 6 months of the rainy season alternating with 6 months of the dry season. However, for the area around the equator the rainy period of the region is 6 months, with 3 months of the rainy season alternating with 3 months of the dry season, resulting in 2 rainy periods in one year.

In the second and third mode observations that is shown in Figure 3 and 4, there were two quite different and contradictory modes in Sumatra and Kalimantan. The second mode of carbon emissions shows that most parts of Indonesia, especially in Sumatra and Kalimantan, have large emissions. However, large forest fires only occur in southern Sumatra and southern Kalimantan. This is because in Sumatra, Kalimantan and the other parts have high rainfall, which can prevent forest fires. The temporal pattern is the opposite of the first mode where in the second mode the pattern always occurs at the beginning of the year between January and mid-year. This pattern is a twelve-month period rainfall pattern in the area, so in the second mode the rainfall pattern dominates the third temporal pattern. From the third picture in Figure 3 it can be seen that the pattern has a twelve-month period, so this pattern only affects regions that have a period of twelve months of rainfall.

The carbon emission variable in the third mode shows the same as the second mode, where there is a large carbon emission in most parts of Sumatra and Kalimantan. However, only forest fires occurred in central Sumatra. This is because in Figure 4 almost all of Sumatra and Kalimantan have high rainfall except in Central Sumatra that shown in the third picture. Based on Figure 4 this pattern occurs within six-month period, so that only the area with six-month rainfall period that is affected by pattern in the third mode. This rainfall signal causes different time of dry season that only happened in equatorial area while twelve-month rainfall area has rain season. Consequently, this signal cause different pattern of carbon emission that does not occur in other area that only have twelve-month rainfall period. Based on the analysis of all regions of Indonesia, forest fires only occur in Sumatra and Kalimantan and separate in 2 patterns which has sixth and twelve period, therefore analysis will be carried out each of the two regions.

5.1. Kalimantan

Kalimantan is an area that has the second largest peatland area in Indonesia. The area of peatland in Kalimantan reaches 4.78 million hectares (ha), the area of peatland constitutes 32% of the total peatland in Indonesia. In fact, almost every year there are forest fires in Kalimantan, especially in peatland areas.
The first mode of analysis in the Kalimantan region shows that carbon emissions tend to be high in areas around the coast of Kalimantan or areas close to the ocean, while rainfall in the first mode is high in the regions of Central Kalimantan to North Kalimantan. The pattern that appears in the first mode has a dominant period of six months and twelve months. This is influenced by different rainfall period in some parts of Kalimantan. The equatorial area of Kalimantan has two rainfall signals, there are six-month and twelve-month period, while in South Kalimantan which is the area that forest fire occurred has twelve-month rainfall only. Interesting patterns are found in the second and third modes. The second and third modes of burned area variable have a similar pattern to the first mode. However, the variables of carbon emissions and rainfall in all modes have significant differences.

In the second mode, the temporal pattern that formed has a high frequency in the beginning of the year to mid-year, then low in the middle of the year until the end of the year. Based on the formed pattern, the pattern is the rainfall of the Kalimantan region with a period of twelve months, so it can be concluded that the second mode is dominated by the twelve-month rainfall pattern. High emissions in second mode appear in regions that have twelve-month rainfall period, but because of peatland area, emissions also always appear in almost of Kalimantan.

The pattern formed in the third mode of the Kalimantan region has a high frequency of rainfall in the area around the equator and has a six-month temporal pattern period. Based on these two things, the third mode pattern is strongly influenced by rainfall patterns that have a six-month period. In this pattern the burned area was slightly compared to the previous mode, so it can be concluded that Kalimantan has a six-month forest fire pattern but the frequency is quite small and dominated by forest fire patterns with a period of twelve months. This also mean that forest fires in Kalimantan mostly occur on regions that has large of peatland area in South Kalimantan that produce high carbon emissions during dry seasons.

**Figure 6.** spatial patterns of variables (a) burned area (b) carbon emissions (c) rainfall in the second mode of Kalimantan

**Figure 7.** spatial patterns of variables (a) burned area (b) carbon emissions (c) rainfall in the third mode of Kalimantan

**Figure 8.** period of temporal patterns in the (a) second mode and (b) third mode of Kalimantan

The pattern formed in the third mode of the Kalimantan region has a high frequency of rainfall in the area around the equator and has a six-month temporal pattern period. Based on these two things, the third mode pattern is strongly influenced by rainfall patterns that have a six-month period. In this pattern the burned area was slightly compared to the previous mode, so it can be concluded that Kalimantan has a six-month forest fire pattern but the frequency is quite small and dominated by forest fire patterns with a period of twelve months. This also mean that forest fires in Kalimantan mostly occur on regions that has large of peatland area in South Kalimantan that produce high carbon emissions during dry seasons.
5.2. Sumatra

Further analysis was carried out on the Sumatra region. The first mode of burning area variable shows that fires occur in the area around Palembang and Riau. Carbon emission variables show that high carbon emissions occur in almost all parts of Sumatra but focuses on areas where there is a forest fire. Rainfall variables indicate that higher rainfall occur in areas close to the south coast or close to the Indian Ocean. Consequently, although some area has high emissions, there was no fires when rainfall on that regions was also high. The temporal pattern in the first mode shows that high frequency occurs two times in one year and has two dominant periods, six months and twelve months. This shows that the dominant variable in the temporal pattern is rainfall, with six-month rainfall signal in Riau and twelve-month rainfall period around Palembang. These two periods will be analyzed separately in the second and third mode.

Figure 9. spatial patterns of variables (a) burned area (b) carbon emissions (c) rainfall, (d) period of temporal patterns in the first mode of Sumatra

The pattern in the second mode has similar burning area pattern to the first mode, but larger fires occur in the area around Palembang. While the carbon emission spreads in the southern Sumatra region and higher rainfall is in northern Sumatra. The temporal pattern that is formed has a high frequency around the beginning of the year and has twelve months period, so it can be concluded that the formed temporal pattern is strongly influenced by the South Sumatra rainfall pattern for a twelve-month period.

Figure 10. spatial patterns of variables (a) burned area (b) carbon emissions (c) rainfall in the second mode of Sumatra
The pattern formed in the third mode is the opposite of the pattern in the second mode. Carbon emissions in the second mode are high in the northern Sumatran region around Riau and high rainfall is in the southern Sumatran region. This causes the third mode of burning area to occur only in areas around Riau. The temporal pattern formed has a six-month period and has a high frequency two times in one year, so it can be concluded that the formed temporal pattern is strongly influenced by rainfall which has a six-month period. Although there were high emissions in almost of Sumatra, this pattern only caused fires in Riau that has also six-month period of Rainfall.

From the analysis of these three modes, it can be concluded that the first mode is a combination of the second and third modes, while the second and third modes of forest fire patterns are influenced by the period of rainfall that occurs in areas where forest fires occur. Therefore, further analysis is carried out to see separately the patterns formed in northern Sumatra and southern Sumatra. Separation is carried out by dividing into two parts of Sumatra region transversely, so that a two region with the same matrix resolution is formed.
The discussion will then focus on separating the six-month and twelve-month periods. The spatial pattern that appears in the first mode shown in Figure 13(a) is a pattern similar to the first mode of the Sumatra region shown in Figure 9 and Figure 11. The difference lies in the temporal pattern and period. The formed temporal pattern is a pattern that is strongly influenced by the burned area. The next difference is the period of the pattern which has high frequency at six-month period. This causes two forest fires in one year in Riau.

\[
\begin{array}{ll}
\text{Period (Month)} & \text{Frequency} \\
(a) & \\
(b) & \\
\end{array}
\]

**Figure 14.** period of temporal patterns in the (a) first and (b) second mode in North Sumatra

**Figure 15.** period of temporal patterns in the (a) first and (b) second mode in South Sumatra

In the spatial pattern of the second mode, formed burned area resembles the pattern in the first mode. While the temporal pattern in the second mode is a similar with the first mode. This pattern is a rainfall pattern that has a six-month period as shown in Figure 14. Significant differences occur in emission and rainfall where both of these variables provide a spatial pattern that is opposite to the first mode. This is because the second pattern is strongly influenced by rainfall, such as in the temporal pattern so that the spatial pattern is different with the pattern in the first mode which is dominated by burned area. Rainfall that dominated the third mode will make high frequency in temporal mode when there was high rainfall, different with burned area, second mode temporal signal frequency will high when there was large burned area.

Based on Figure 13(b) it can be seen that the spatial pattern in South Sumatra resembles the spatial pattern in the first mode of Sumatra which is shown in Figure 9 and Figure 10. In addition, the temporal pattern is also almost the same, the difference lies in the period of the temporal pattern, where in Figure 8 (d) there are two dominant periods of six and twelve months, in Figure 15 there is only one dominant period of twelve months. The formed temporal pattern is a signal that strongly influenced by the burned area. This can be seen from the high signal frequency around the end of the year when the dry season occurs in Indonesia which has a period of twelve months of rainfall.

The spatial pattern in the second mode shown in Figure 13(b) is a pattern similar to the pattern that appears in the first mode, however the frequency is different. Significant differences are seen in the temporal pattern. The temporal pattern in the second mode is strongly influenced by the rainfall pattern of the region, so the pattern has a high frequency when the beginning of the year to mid-year. This is the time of the rainy season in the twelve-month rain period in Sumatra. In addition, there are also periods that appear in the temporal pattern of six-month rainy periods. However, these signals tend to be small and are dominated by rainfall signals that have a twelve-month period.

From all of these modes, the third modes from all analysis give different patterns with the other modes significantly. The different appears on both spatial and temporal pattern. All third modes have pattern occur in six-month period. This happened because six-month rainfall signal that is happened around equator. This signal lead to be different in dry seasons to the other area that has twelve-month rainfall signal. Therefore, all three modes will only be affected on equatorial area. Furthermore, the six-month rainfall also leads to dry season that happened at the same time with twelve-month period dry season. This phenomenon was showed by first mode in all region. All first modes showed forest fire
happened in areas where forest fires occur caused by both rainfall signal. However, the six-month period dry season was slightly weaker than the twelve-month dry season. Consequently, the frequency of six-month period was much lower than the twelve-month period.

6. Conclusion
Forest fires in Indonesia are phenomena that occur every year in Indonesia and are influenced by many factors. Most of carbon emissions variable in this research shows that most of region has high carbon emissions on all modes. It can be explained that in almost of region in Indonesia has a high temperature in afternoon and the rain will occur mostly after 2 pm. This fact make generally Indonesia will produce high emissions during afternoon, with different some of emissions depend on land type of each regions. The highest carbon emission was produced by peatland area. Patterns obtained from the analysis of burned area, carbon emissions, and rainfall the Indonesian region can be grouped into two. There are six-months period and twelve-months period patterns. This was same as rainfall in Indonesia that mostly can be grouped into sixth-month and twelve-month.

Forest fires in Indonesia occur annually in the regions of Central Sumatra, South Sumatra and South Kalimantan. In general, forest fires in Indonesia occur with a twelve-month period and occur when the dry season takes place in some of these areas. Besides that, number of peatland area in that regions produces high carbon emissions during dry season.

Besides having a twelve-month period, there are several areas that have a period of six months of forest fires. The area is a forest fire area whose rainfall is dominated by six months of rainfall such as the area around Riau. This is because when the rainfall is dominant for six months, there are two dry seasons which can cause forest fires twice a year. This six-month forest fire pattern does not appear in areas with little rainfall signal strength and is dominated by twelve months of rainfall. This can be seen in the analysis of the Kalimantan region that there are six-month forest fire patterns, but because of the dominant rainfall in Kalimantan twelve months, the small burning area is small.

Acknowledgments
This research was funded by the Directorate of Research and Community Service, Directorate General of Research and Development Strengthening, Ministry of Research, Technology, and Higher Education, in accordance with the Research Contract of the National Post-Graduate Team for Budget Year 2017 Number: 011 / SP2H / LT / DRPM / IV / 2017, April 20, 2017.

References
[1] Anton H, Rorres C. 2010. Elementary Linear Algebra. 10 Ed. New Jersey: J Wiley.
[2] Chen CC, Lin HW, Yu JY, Lo MH. 2016. The 2015 Borneo fires: What have we learned from the 1997 and 2006 El Niños? IOP Publishing Environ. Res. Lett. 11 104003
[3] Fernandes K, Verchot L, Baethgen W, Gutierrez-Velas V, Pinedo-Vasquez M, Martius C. 2017. Heightened fire probability in Indonesia in non-drought conditions: the effect of increasing temperatures. IOP Publishing Environ. Res. Lett. 12 054002
[4] Fanin T, Van der Werf GR. 2017. Precipitation–fire linkages in Indonesia (1997-2015). Biogeosciences, 14, 3995–4008
[5] Hannachi A. 2004. A primer for EOF analysis of climate data: Department of Meteorology, University of Reading. Reading RG6 6BB, UK
[6] Harris N, Mimmemeyer S, Stolle F, Aris OP. 2015. Indonesia’s Fire Outbreaks Producing More Daily Emissions than Entire US Economy. World Resources Institute.
[7] Leon SJ. 2010. Linear Algebra with Applications. 8 ed. New Jersey: Prentice Hall.
[8] Navarra A, Simoncini V. 2010. A Guide to Empirical Orthogonal Function for Climate Data Analysis: Springer.
[9] Nicholson WK. 2001. Elementary Linear Algebra. Singapore: McGraw-Hill.
[10] Van der Werf GR, Randerson JT, Giglio L, Collatz GJ, Mu M, Kasibhatla PS, Morton DC, DeFries RS, Jin Y, Van Leeuwen TT. 2010. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). Atmospheric Chemistry and Physics, 10 11707-11735
[11] Wiley E, Hansen. 2014. Fourier Transform: Principles and Applications. New Jersey: J Wiley.