 doi:10.20944/preprints202108.0438.v1

Assessment on the Use of Meteorological and Social Media Information for Forest Fires Detection and Prediction in Riau, Indonesia

Anni Arumsari Fitriany[1,2], Piotr J. Flatau[3] , Khoirunurrofik Khoirunurrofik[1] and Nelly Florida Riama[2]

1 Faculty of Economics and Business, Universitas Indonesia; khoirunurrofik@ui.ac.id
2 Agency for Meteorology, Climatology, and Geophysics of the Republic of Indonesia; anni.arumsari@bmkg.go.id ; nelly.florida@bmkg.go.id
3 Scripps Institution of Oceanography, University of California San Diego; pflatau@ucsd.edu
* Correspondence: anni.arumsari@bmkg.go.id

Abstract: Early detection that results in early warning of forest fires occurrences in Indonesia, which are strongly related to land management practices (including peatlands), is necessary to mitigate land and forest fires in Indonesia. Riau has been chosen in this study because of its vulnerability to forest fires. The remoteness of this region is one reason for developing alternative warning tools using meteorological and social media information. This study identified tweets related to fires using carefully selected keywords, geoparsed to select messages relevant to fire occurrences, and binned within several Indonesian sub-regions in Riau Province. Content analysis was performed for 31 related online local newspapers. Assessment to study the correlation between meteorological and Twitter information with the forest fires was conducted. Existing approaches that the BMKG and other Indonesian agencies use to detect fire activities are reviewed, and a novel approach based on crowdsourcing of tweets is proposed. The results show a correlation between meteorological information and Twitter activity with satellites derived hotspot information. The policy implications of these results suggest that information should be included in the fire management system in Indonesia to support fire early detection as part of fire disaster mitigation efforts.

Keywords: forest fires, meteorology, newspaper, policy, Tweeter, Indonesia

1. Introduction

Indonesia has a very large area of forest. Based on the data from the Ministry of Environment and Forestry of the Republic of Indonesia in 2019, its total forest area is about 94,1 Million Ha, which is almost half of its total land-covered area. This condition not only could make Indonesia as the country which is rich in forestry resources, but also could create a problem for the country because of its forest fires disasters if they are not well managed.

The cost of fires could affect many aspects. Tacconi [1] mentioned that the economic cost of fires is not only caused by the fires themselves, but also by the smoke haze pollution it may cause. The Economic costs of fires would come from the loss of the timber and crops plantation, loss of indirect forest benefits such as flood protection as well as erosion and siltation, loss of biodiversity, fire fighting cost, loss of property, and even human life. While, the economic costs of smoke haze pollution would come from the loss in many important sectors in Indonesia, such as health, tourism, transport, industrial production, as well as fishing decline.

According to World Bank report [2] in 2019, the land and forest fires that raged in Indonesia is considered to interfere with economic growth if this problem persists. The World Bank has estimated a decrease of 0.09 and 0.05 points in the percentage of Indonesia's economic growth for 2019 and 2020 because of the forest fires. It estimated Indonesia's economic growth in 2019 at 5 percent and in 2020 at 5.1 percent.

It was also mentioned in the World Bank report that the most extensive forest and man-made land fires since the 2015 fire crisis in Indonesia was occurred in the 2019. On

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 August 2021 doi:10.20944/preprints202108.0438.v1

the 2019 fire events, a bulky haze was produced and the haze blanketed at least eight provinces and economic activities, both domestically and abroad, were hampered. By September 2019, respiratory health diseases had been reported by over 900,000 people, the operation of 12 national airports had been halted, and hundreds of schools in Indonesia, Malaysia, and Singapore had also been closed temporarily. Overall, it is also mentioned in the report that the estimation of the total damage and economic loss in eight affected provinces throughout June-October 20194 is USD 5.2 billion, or equivalent to half of Indonesia GDP, which was mainly contributed by the loss in the sectors of agriculture, transportation, trade, industry, and environmental (see Figure 1).

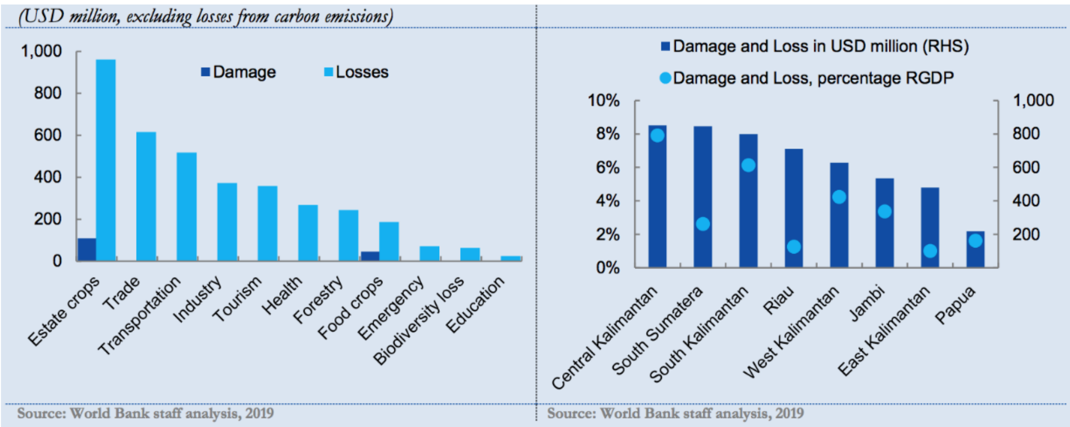

Figure 1. Damage and losses recorded from the 2019 Forest Fires in Indonesia (source: World Bank Report in 2019).

Based on the above mentioned World Bank publication on Indonesia Economic Quarterly: Investing in People released in December 2019, it was reported that the damage and losses recorded from the 2019 Forest Fires in Indonesia were more than USD 700 million or almost 2% of the Gross Regional Domestic Product (GRDP) of Indonesia. This means that if mitigation using available sources of information (such as meteorological and climatological phenomena as well as tweeters and newspapers related to forest fires) is carried out properly, the losses could be minimized. Assuming 50% of fire damage and losses can be reduced, it means that the Government can minimize losses up to USD 300-500 million per year, or about 1% of Indonesia's GRDP can be prevented from these losses.

Comparing the costs and benefits of improving Indonesia's forest fire early detection system by integrating meteorological and social media information (Annex 1). The data used in the calculation is qualitative data based on an assumption of the information derived from some sources. The actual losses of the forest fires events in Indonesia could be more than the assumption we took. According to data from World Bank, damage and losses of the 2019 forest fires showed USD 700 million while the cost and benefit calculation using the assumption that we did in this study showed about USD 234 Million for the cost. The cost and benefit analysis using that we did showed that the benefits of upgrading the system far outweigh the possible costs, which means upgrading the system is something which is worth it.

Indonesia’s land and forest fires are strongly related to peat soil and land management practices. This leads to high frequency of fire activity in these regions, usually during dry seasons. Because of the fuel type such fires are long-lasting but this depends on land cover management issues Miettinen, Shi, & Liew [3]. Kibanov et al. [3] write that fires in the peatland and their correlated haze events are the disasters that may create significant impacts. Several environmental conditions would generate the occurrence of fires, but in Indonesia, they are often generated by human activities, which mainly from the practices of agriculture which is illegal, for example, because of the burning activities using the slash-and-burn techniques in the preparation of the palm oil plantation from forests and peatlands Holmgren [4]. Kim et al. [5] indicate that environmental problem such as accelerated deforestation and economic losses had resulted in burning activities and further note that associated haze events seriously affect local and remote population health. The fires often produced the haze in the peatland in widely affected areas and in a relatively very short time. Several types of warning are issued depending on haze severity: from reducing outdoor activities to evacuation of the affected areas because of haze-related health issues (Frankenberg, McKee, & Thomas [6]).

According to Ulya et al. [7], forest fires would reduce institution (family, company, and Government) income of IDR 77.4 Million per Ha of burned area. In the social aspect, family is the most affected institution, with its reducing income is IDR 45,48 Million, which is higher than the company reducing income of IDR 20,42 Million and the government reducing income of 11,54 Million per Ha of burned area. The study also showed the strong relationship between forest activities sectors and the rural based agricultural sectors.

Considering the above mentioned economic costs from the loss in the community, company and government incomes by the forest fires, one of the urgent policy action required in Indonesia is the improvement of the forest fires early detection and warning system. Rogers and Tsirkunov [8] mentioned that the disaster risk reduction cost and benefit evidence consistently showed that the early warning system can save the lives and property. For European case study, several hundreds lives per year could be saved by hydro-meteorological information and early warning system (S. Hallegatte [9]).

Currently, Social Media, such as Twitter and newspaper, has been explored to be used in the policy design. Several researches on the how twitter and newspaper responds to the political issues (Biswas M. [10] and Kalogeropoulos A., et al. [11]) have been conducted. They showed how Twitter and Newspaper beneficial for politician through the identification of problems Tweets or mentioned in the media. Power, et al. [12] states that social media can be one of the valuable channels of communication. However, the adoption of its uses as a source of information to enhance public awareness was still rare in adoption (Anderson [13]). This is because it is not easy to frame the content of the social media properly, shift the information with a large volume, and trust the message in the social media (Lindsay [14]). It is also mentioned by Power et al. [12] that there has been a successful study on the use of Twitter for detecting disaster events notably the earthquake (Sakaki, Okazaki and Matsuo [15]; Robinson, Power, and Cameron [16]; Avvenuti, Cresci, Marchetti, Meletti, and Tesconi [17]). A new Twitter database to detect flood on a global scale in real-time with a detecting accuracy of approximately 90% were presented by De Brujin, et al. [18].

Early detection and prediction of forest fires are very important in order to prevent fires from spreading more intensively and for a quick locating of forest fires area which is needed to be immediately put into action. The Government of Indonesia has actually operated a tool developed from the Canadian Forest Service (CFS) by the Agency Assessment and Application of Technology (BPPT) in collaboration with the Agency for Meteorology, Climatology, and Geophysics (BMKG), and the Ministry of Environment and Forestry (KLHK), called the Fire Danger Rating System (FDRS).

Although this system has been operating in Indonesia since 1999, forest fires are still occured with the worst in 2015 (Figure 2). This indicates that other sources of information might be urgently needed to strengthen the efforts that the government has made. There are many information available in public that might correlate with the forest fires, one of which is the meteorological information, such as the El-Nino and Indian Ocean Dipole (IOD) Indices, and social media information, such as twitter and newspaper. Assessing the use of the above mentioned information in the forest fires early detection is important to study whether the information could be an option for the policy maker in Indonesia in their concerted effort in mitigating the forest fires.

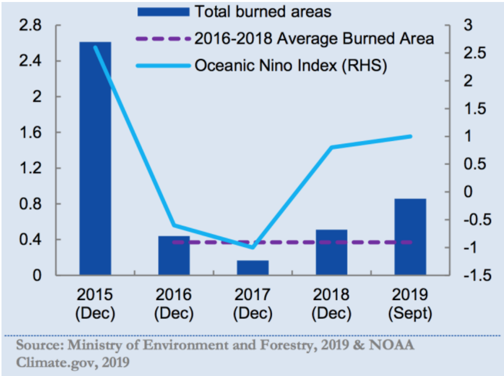

Figure 2. Total annual burned areas (in million hectares) and Ocean Nino Index, indicating likelihood of dry season that year (source: World Bank Report in 2019).

In this study, we will assess meteorological conditions leading to fires, such as El-Nino and IOD indices, and the community response to the forest fires in the social media, such as Twitter and newspaper in Riau Province in 2015-2019 timeframe. The question we would like to address is how meteorological information, such as the El-Nino and Indian Ocean Dipole (IOD) Indices, and the social media information, such as Twitter and newspaper, correlate with the fires occurrences in Riau and how the social media support the policy makers in taking a decision by a massively discussed forest fires issue in Twitter. The result of the study will provide better information whether meteorological information, such as the El-Nino and Indian Ocean Dipole (IOD) Indices, and the social media information, such as twitter and newspaper could be used for detecting and mitigating the forest fires events in Riau.

This study will explore more how meteorological information, such as the El-Nino and Indian Ocean Dipole (IOD) Indices, and the social media information, such as twitter and newspaper, could contribute to the early detection of the possibility or existence of fires to support early warning of forest fires in Indonesia. The degree to which the relationship between the dynamics of forest fires and the dynamics of information we can get from the meteorological information as well as from the social media can be used as early detection or early warning of forest fires.

This study is expected to contribute in complementing the research study or literature growth in this area, such as providing insight to policymakers in determining its policy choices for new practices in making a quick and more proper decision for mitigating forest fires from the weather and climate as well as social media information such as Twitter and newspaper.

This paper consists of five parts, beginning with an introduction and a summary of the research. The second part reviews the literature. The third part discusses research methods. The fourth part presents the results and discussions, and the final part offers the conclusions and limitations of the study.

2. Literature Review

The use of social media as the early detection and warnings for disaster events

The available data and information corresponding to the forest fires would be useful for the disaster management authority and policy makers to detect and mitigate the land and forest fires occurrences in advance. These data and information, including hotspot information, weather and climate information such as El-Nino and Indian Ocean Dipole, as well as information from the social media, such as newspaper and Twitter, could also be one of the informative sources for the land and forest fires policy in Indonesia.

Several studies have reported using the information from social media and newspapers to study public policy issues for many sectors, including disaster events. People could use Twitter as a media to share their concerns about the important issue, especially for those who are also active in providing comment on news outside Twitter and the political partisans (Kalogeropoulos, A., et al [11]). Study of Biswas, M. [10] showed how news media contributed to the process of policy making through the problems and priority area identification for the policy makers. The study tried to see how the mainstream and ethnic media response to one economic stimulus plan debate from the issue uploaded in the news media.

The use of social media in political communication was studied by Stefan Stieglitz and Linh Dang-Xuan [19]. The study tried to develop a framework that could be used as a support tool for the development of the toolset for collecting, storing, monitoring, analyzing, and summarizing user-generated content, which is politically relevant from social media to be used by the political institutions. Sinnenberg, L., et al. [20] has shown that Twitter has also been used in several health researches.

Power, et al. [12] states that social media can be one of the valuable channel of communication. However, the adoption of its uses as a source of information to enhance public awareness was still rare in adoption (Anderson [13]). This is because it is not easy to frame the social media content properly, shift the information with a large volume, and trust the message in the social media (Lindsay [14]). It is also mentioned by Power et al. (2015) that there has been a successful study on the use of Twitter for detecting disaster event particularly the earthquake (Sakaki, Okazaki and Matsuo [15]; Robinson, Power, and Cameron [16]; Avvenuti, Cresci, Marchetti, Meletti, and Tesconi [17]). A new Twitter database to detect flood on global scale in real time with detecting accuracy of approximately 90% were presented by De Brujin, et al. [18].

Power et al. [21] applied twitter to develop a system which can notify the identification of the near-real-time tweets related to event of fires using the keywords related to fires to identify messages of candidate. To focus on actual fire events and refine the results, a text classifier is then processed the tweets. New events were defines after a “quiet” inactive period tweets related to fires.

Several researches were conducted to study about the land and forest fires in Indonesia. Most of them concentrate on specific aspects such as meteorological phenomena, the social media, or the policy. Reid, et al. [22], Field, et al. [23] and Pan, et al. [24], studied how meteorological and climate phenomena relate to the Indoesian forest fires. Kibanov, et al. [3] studied on mining social media as the information source for peatland fire and haze disaster management. Forsyth, T. [25]

studied about public concerns in Newspaper about transboundary haze by comparing the case in Indonesia, Singapore, and Malaysia. Moreover, Panjaitan, et al. [26] studied the forest fires policy Indonesia, specifically on the role of central government and local government and the moderating effect of good governance on forest fire policy in Indonesia. In this study we try to assess if both information from the meteorological and social media information is still relevant to be used as the source of information for detecting forest fires to support policy maker in making the decision on the forest fires mitigation.

Indonesian Agencies – fire related warnings

The Government of Indonesia has tried many ways to prevent and mitigate the occurrence of land and forest fires. Indonesia's monitoring, prediction, analysis, and warnings of land and forest fires involve several Ministries and Agencies. The main ones are:

1. The Ministry for Environment and Forestry (KLHK);
2. The Agency for Meteorology, Climatology, and Geophysics (BMKG); and
3. The National Agency for Disaster Management (BNPB).

The KLHK offers a web dashboard for monitoring forest fires called SiPongi (http://sipongi.menlhk.go.id/home.main). Among several products they provide are hotspot distribution maps as well as information about the burnt area in Indonesia for 2014-2019 (http://sipongi.menlhk.go.id/pdf/luas_kebakaran). One can also access weekly hotspot and monthly hotspot satellite data and a static forest fires vulnerability map (see Fig. 3). In conjunction with the Geospatial Information Agency, the KLHK develops the forest fire vulnerability map using physical factors such as land cover, land topography,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 August 2021 doi:10.20944/preprints202108.0438.v1

weather climatology, and anthropogenic factors such as people behaviour and land permitted area location and local socio-culture.

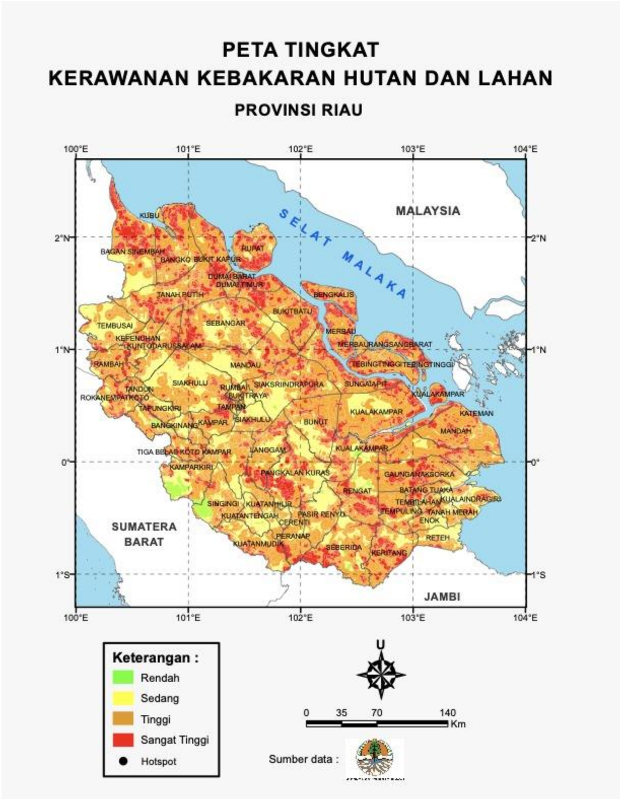

Figure 3. Forest fire vulnerability map developed by KLHK. Notes: Data classifies the vulnerability of the forest fires into 4 categorizes of fires vulnerability: very easy to be burnt (Sangat Tinggi), easy to be burnt (Tinggi), difficult to be burnt (Sedang), and very difficult to be burnt (Rendah).

The BNPB provides the geospatial dashboard, which provides information on forest fire mitigation and management (http://geospasial.bnpb.go.id/karhutla2019). It offers socio-economic vulnerability information about disasters as well.

The BMKG provides the internet web dashboard for forest fires information called Fire Danger Rating System (FDRS). The dashboard, which is developed by the Agency for Meteorology, Climatology, and Geophysics (BMKG) in collaboration with the Agency for the Assessment and Application of Technology (BPPT), and the Ministry of Environment and Forestry (KLHK) from the Canadian Forest Services (CFS), is publicly available at and provides information about the Fine Fuel Moisture Code and Fire Weather Index (see Figs 4 and 5), smoke distribution imagery, and hotspots distribution. The Fine Fuel Moisture Code and Fire Weather Index information are dynamic, derived from the weather parameter analysis, and based on the weather forecast.

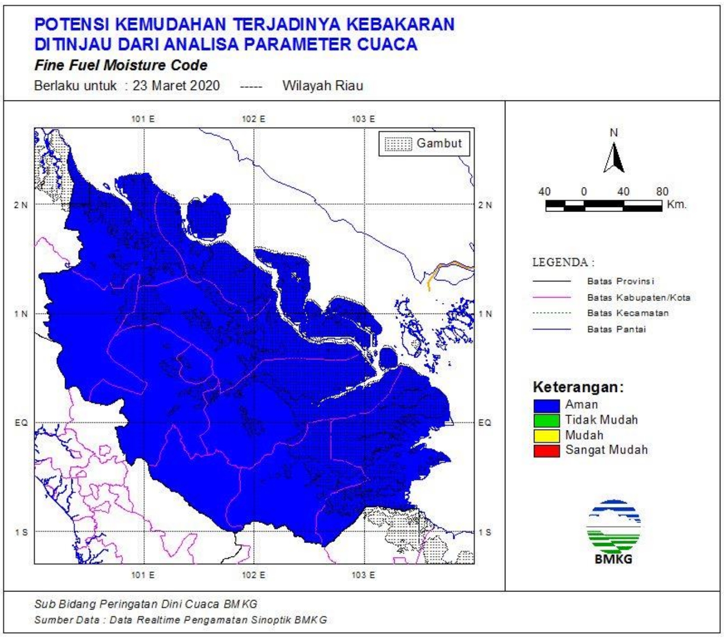

Figure 4. Fine Fuel Moisture Code Map developed by BMKG FDRS. Notes: Vulnerability to fires occurrence is based on Canadian Model (FFMC) and their Fine Fuel Mositure Code. Four categories are provided: very easy to be burnt (Sangat Mudah), easy to be burnt (Mudah), difficult to be burnt (Tidak Mudah), and very difficult to be burnt (Aman).

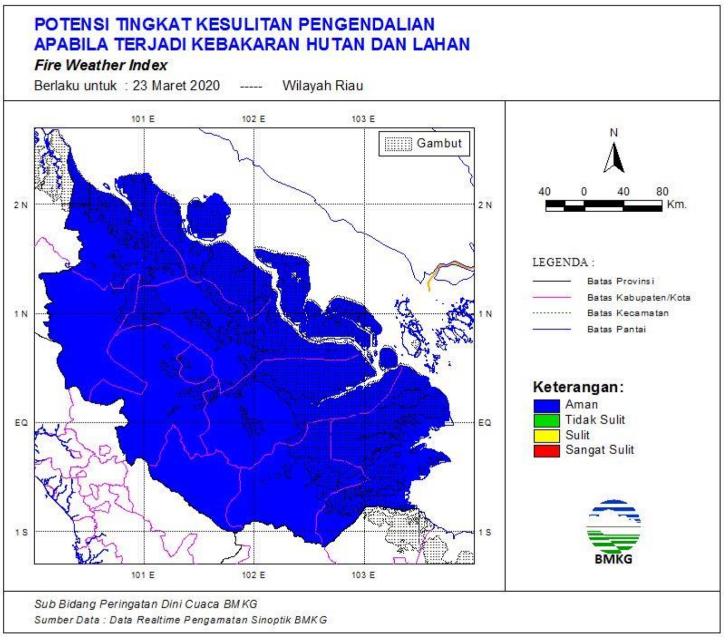

Figure 5. Fire Weather Index map example available from BMKG FDRS

The press release issued by the Ministry of Environment and Forestry (KLHK) on November 27, 2017 stated that the policy on forest and land fires emergency alert criteria, such as "standby status", are coordinated with several national institutions, such as National Agency for Disaster Management (BNPB), Agency for Climatology, and Geophysics (BMKG), National Institute of Aeronautics and Space (LAPAN). Such policies are also coordinated with the Local Disaster Management Agencies (BPBD), Natural Resources Conservation Centers, Center for Climate Change and Land and Forest Fires Control in Jawa-Bali-Nusa Tenggara, Kalimantan, Sumatera, Sulawesi, and Papua. It was also stated that the alert and emergency status on land and forest fires are determined on the basis of assessment of parameters, such as the ranking of forest and land hazards, temperature, number of days without rain, rainfall analysis, rainfall forecasts, hotspots, forest and land fires, smoke conditions, air quality conditions, visibility and the number of people suffering from health problems due to the land and forest fires.

For each level of alert and emergency conditions due to land and forest fires, the actions taken varies depending on its severity. Efforts to control land and forest fires will be more intense for higher alert and emergency levels. In the "Technical Criteria of Preparedness Status and Emergency as a Result of Forest and Land Fire" there are four severity levels: "Standby 3" (normal), "Standby 2" (alert), "Standby 1" (Emergency Standby), and "Emergency" (Emergency Response at Regency/City, Province or National Level).

The above-mentioned forest and land fire emergency technical criteria are regulated in the Minister of Environment and Forestry of the Republic of Indonesia number P.9/MENLHK/SETJEN/KUM.1/3/2018 dated 7 March 2018. The regulation defines that the Emergency or Emergency Response is a list of activities that should be taken at an immediate manner when a forest fires disaster occurred. The actions to be taken to deal with forest fires impacts include victims evacuation and rescue, securing properties, fulfilling basic necessities and shelter, managing refugee, and infrastructure restoration.

New methodology: Social-media based fire detection

Social-media-based fire detection can be important in Indonesia, where twitter are used by more than 75% of active users of the internet. It is recorded that Indonesia was the largest global tweet /user ratio in 2012, which was actively used by about 12% of the total population in Indonesia (Carley et al. [27]). More importantly, the development of a direct correspondence between the direct Twitter event observations and the larger-scale atmospheric phenomena from which it arise could be provided by the improved analysis of forest fires acquired through social media. This idea was recently exploited in a paper about use of Twitter data to predict floods in Indonesia (Baranowski et al. [28]). With this recognition, it is hoped that there will be a time increment for forest fires detection and warning, which might be useful for the decision makers, because Twitter observation in the field can identify this forest fires phenomena and because of anticipated improvements in the forest fire danger system warnings. The public attention conveyed and growth in twitter can become a source of information for the government and policy makers to quickly determine appropriate forest fire mitigation measures. Therefore, the three independent datasets are used in this study: hotspot satellite data, forest fires events from Twitter messages, and articles in local newspapers.

Social media can be an effective communication method but, at the same time, it is difficult to analyse large amount of social media content and decide its appropriateness to fire prediction problem. There are works related twitter use to disaster management such as earthquakes (Sakaki, Okazaki and Matsuo [15]; Robinson, Power and Cameron [16]). This is possible because of the nature of unexpectedly earthquake occurrences and because of the feature of people reacting to tweets when they experience them. The Tweets are processed by the Earthquake Alert and Report System (EARS) system after an earthquake event has been detected to get the information about impacts in the affected area in order to produce a damage assessment report. They also include measures to deter and minimize the effect of hoax (Power et al. [12]). Power, Robinson and Ratcliffe [21] have applied a similar methodology in detecting fires. In that study, the authors identified near-real-time tweets that describe the fire occurrences through a system of notification. The fires message candidate were identified by filtering the Tweets which are related to fire occurrences. After that, a text classifier is processing the Tweets to get the results refined for getting the actual fire occurrences. After an inactive "quiet" period of discussion of the fires related matters, new events are detected by their system (Power et al., [12]). Such system can be used by fire responders by defining a user's location and analysing tweets for its content using automatic tools such as natural Language Processing.

In aggregate tweets can be used for policy makers by helping in understanding localization and frequency and duration of fire events and long term determination of resources needed to its severity, long distance influence of smoke, or even determination of tweets contents for type of health problems experienced by population exposed to fires. Finally, policy maker can encourage use of social media for reporting purposes, develop twitter based web site providing "incident manager" interface for experts and general public which would contain such information as near real-time situational awareness, associated satellite and weather images, provide timely warnings, and showing fire locations. In a study conducted by Aditya et al. [29], the crowdsourcing information is used in a ground-based spatial validation of satellite data to get the more accurate location and fire severity. It was found that only some individuals could provide information, although it is expected that mobile reporting through crowdsensing could collect considerable data. They also mentioned that the availability of mobile network signals in either outer or almost inaccessible land was one of the main obstacles of getting the data using their geo-crowd app. They suggested locally store the reports on the users' mobile device and send them upon the availability of the network signal.

3. Materials and Methods

In order to develop initial understanding of the feasibility of using social media to augment fire monitoring in Indonesia and provide supplementary tools for the policy maker, we decided to perform several cases studies to understand availability of fire related tweets, newspapers reports and weather related (large scale) factors and their correlation with satellite derived product. To this end, we focus the study for Riau Province of the Sumatera Island, Indonesia.

We choose Riau Province in this study because Riau is one of the most vulnerable area of the land and forest fires location in Indonesia according to the BMKG data (Figure 6). Its location, which is very nearby with the neighboring countries such as Singapore and Malaysia, will create a problem in Indonesia itself and could create a transboundary haze problem for the mentioned countries.

Figure 6. The distribution of the prone area of land and forest fires in Indonesia (source: BMKG)

In this study, we use the hotspot data from MODIS satellite sensor on Terra and Aqua satellites for all levels of confidence from 2015 to 2019 for the occurrences of forest fires in Riau area. The meteorological and social media information that we used to be assessed in this study are the El-Nino and Indian Ocean Dipole Phase indices data for Indonesia

 doi:10.20944/preprints202108.0438.v1

for the period of 2015-2019, the newspaper data taken from online newspaper related to forest fires from 1st of January 2019 to 31st of December 2019, and the 5-years Twitter data from 2015 to 2019.

Satellite Data Collection, Construction, and Analysis

The satellite hotspot data from MODIS satellite sensor on Terra and Aqua satellites and SNPP satellite for all level of confidence were used to define the occurrences of land and forest fires from 2015 to 2019. The data was collected from an excellent LAPAN site: [16].

Lapan is managing the database which is obtained from NASA through their algorithm and quality control. The MODIS (Moderate-resolution Imaging Spectroradiometer) both on Terra and Aqua are based on channels 21 and 22 (3.92-3.98 micrometers) with spatial resolution of 1km. The S-NPP (Suomi National Polar Orbiting Partnership) satellite data is based on the VIIRS sensor and uses channel 14 between 3.55-3.93 micrometers with spatial resolution of 375m. These satellites are observing over the Indonesia twice daily on their polar orbit. The LAPAN system allows to retrieve data for a specific province or latitude and longitude box. One can also define a specific time period, confidence level, and satellite observing system.

Satellite database from the LAPAN site http://modis-catalog.lapan.go.id/monitoring/ was searched for Riau province only for all level of confidence. The data is in the form of a CSF files with latitude, longitude, relative humidity (%), name of a satellite system (Aqua, Terra, SNPP) and date.

Satellite data analysis is beyond the scope of this study. But, in short, satellite sensor and retrieve that the surface temperature is hotter than the surroundings by observing incoming radiance in the near-infrared spectrum region. There is a detailed description of the data algorithm in the following Guide http://103.51.131.166/guide/Panduan%20Website%20LAPAN%20Fire%20Hotspot%20v2.0%20(Juli%202020).pdf.

Meteorological Data Collection, Construction, and Analysis

The El-Nino and Indian Ocean Dipole Phase indices data for Indonesia were used to confirm whether there is a connection between the land and forest fires in Indonesia in 2019. The Niño 3,4 index data are taken from the Climate Prediction Center (CPC) of the National Weather Services, National Oceanic and Atmospheric Administration (NOAA). Data between 2015 to 2019 were used to investigate how the land and forest fires occurrences in Indonesia until 2019 were influenced by the the El-Nino and Indian Ocean Dipole. The datasets are available from BMKG data centre.

According to CPC climate guide Niño 3.4 index is for the (5N-5S, 170W-120W) region. The Niño 3.4 anomalies may represent the average equatorial SSTs across the Pacific from about the dateline to the South American coast. The Niño 3.4 index typically uses a 5-month running mean, and El Niño or La Niña events are defined when the Niño 3.4 SSTs exceed +/- 0.4C for a period of six months or more.The meteorological data sets are quality controlled by Climate Prediction Center (CPC) of the National Weather Services. No further quality control of the data were performed.

Newspaper Data Collection, Construction, and Analysis

The data was collected by searching fire related keywords and reading online newspaper articles. We tried to estimate the duration, location, latitude and longitude (based on city mentioned in an article), extend of the fire in kilometres. We also collected URL of the online newspaper and its name. Example of news organizations is Antara News, Tribun News, Ministry of Health, Sindonews, Republika, Detiknews, and other online accessible sources of information. The qualitative content analysis of the land and forest fires information from the newspaper was conducted. Information from the online newspapers were collected from the 1st of January 2019 to 31st of December 2019. The newspaper analysis gave us information about the date and duration of fire occurrences, the burn area and location, the reason why they occurred, and other information such as impact on

 doi:10.20944/preprints202108.0438.v1

people health, animals mortality related to ecosystem disruption, low visibility, decrease in air quality, as well as the limitations related to Government response of fire extinguishing. The purpose of newspaper analysis was not quantitative but was needed to better understand policy making related issues which would have been otherwise difficult to grasp from numerical information about number of hotspots and number of tweets.

Twitter Data Collection, Construction, and Analysis

Twitter messages from a mini blog with a single tweet containing no more than 280 characters were also used to confirm whether there is a connection between the Twitter and the land and forest fires events from 2015 to 2019 in Riau, Indonesia. The limitation of characters to be used in twitter forces messages to be short and precise. Twitter messages provides some metadata with every tweet, such metadata may include location of a tweet, author's name, time when a tweet was posted. However, gathering and analysing Twitter messages may be challenging because of data access and metadata uncertainties. There are several ways of getting historical tweets. For example, one can use Twitter API which is official (https://developer.twitter.com/en/docs.html); however for standard users there is only access to past 7 past days. There are also some paid datasets, such as GNIP (https://developer.twitter.com/en/enterprise). There are some tweeter datasets available, but none addresed fires in Indonesia. Therefore, we used scrapping script to get the Twitter data. The particular script which we used can be found on the github site: https://github.com/Jefferson-Henrique/GetOldTweets-python.

Accordingly, to this site when one enters a Twitter page a scroll loader starts. When one scrolls down one gets more and more tweets, all through calls to a JSON provider. Using tweeter search the deepest (oldest) tweets can be gathered. One can set TwitterCriteria which are collections of search parameters to be used. These can be setSince which allows to set a lower bound date to restrict search; setUntil which is an upper bound date to restrict a search, setQuerySearch which is a query text to be matched, setTopTweets which if "true" will retrive only the Top Tweets; setNear which defines a reference location area from where tweets were generated; setWithin which defines a distance radius from "near" location (e.g. 15mi); setMaxTweets which defined the maximum number of tweets to be retrieved. This technique allowed us to collect historical tweets for the whole Indonesia region.

Using the GetOldTweets-python we were able to search twitter for the 5-year period between 2015-2019 for the whole Indonesia using (separately) the following keywords:

1. forest fire: kebakaran hutan
2. land fire: kebakaran lahan
3. smoke: asap
4. no rain: tidak ada hujan
5. fire: api
6. drought: Kekeringan

In the 5-year period we located about 1,250,000 tweets which satisfied our search criteria for fires for the whole region of Riau, Indonesia. In contrast, only 50,000 of them contained keyword "kekeringan" (draught). For fire events, we did not include "no-rain" and "drought" keywords; these were used to understand number of tweets related to drought. The "smoke" keyword was used separately to understand possible "smoke" hazard related to fire events. Therefore "fire" kewords dominated by far the total amount of retrieved tweets.

To analyse tweets we need to know date and place as well as number of relevant tweets. Dates are provided by individual tweets. Number of relevant tweets can be deduced directly from the final selections. But the location is the most difficult. Due to privacy concerns, tweets do not include GPS latitude and longitude position (even though users can provide such data) which would have been the simplest way of obtaining location. Only less than 1% of all tweets contain such information. This significantly lowers the number of tweets available for analysis. To answer this limitation one can analyse

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 August 2021 doi:10.20944/preprints202108.0438.v1

tweet text and its metadata to provide an estimated event location. That is what the TAGGS algorithm does. The geoparsing algorithm of de Bruijn [18] was used to locate the geographical locations. In this work, the TAGGS was extended to find tweets related to fires in Indonesia.

The TAGGS algorithm allows geoparsing based on the user's nationality (for example Bahasa language) and their hometown. One can also discover direct referrals to specific locations. For example, it can be name of the province or name of a city in the message itself. Such estimation method may result in certain number of messages assigned to an inaccurate location, but we rely on a large number of tweets to identify fire events. Using such a large number of tweets, we get a large signal that minimizes noise due to potentially inaccurate geoparsing location.

4. Results and Discussions

4.1. The analysis of the meteorological and climatological conditions and the drought effects

The Nino 3.4 and the indices of the IOD from the NOAA's CPC in the 2015-2019 timeframe are shown in Figure 7. We can see that positive Niño 3.4 and IOD occurred between April - December 2015, March - July 2017, July - December 2018 and May - December 2019 thus providing insights into climatological forcings in Indonesia for that period.

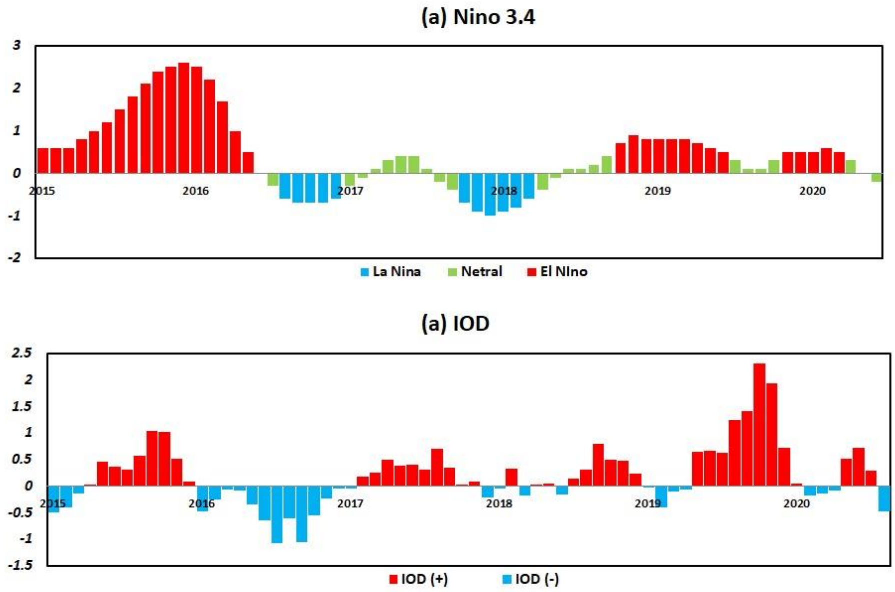

Figure 7. The Nino 3.4 and IOD indices during periods of droughts for 2015 to 2019.

The severe drought condition induced by the dry phase of El Niño event could worsen land and forest fires. A study by Wooster et al. [30] indicates that the phenomena of El Nino, by its influence on the precipitation, is a majorly wide-ranging or large-scale, interannual factor of climate that has a great influence on the fire activity magnitude caused by the large changes of the land cover, practices of the agricultural arrangement and fire ignitions caused by the human which are occurred in Kalimantan. Their results also showed that the forecast of El Nino provides ways of estimating fire events extent and magnitude some months earlier.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 August 2021 doi:10.20944/preprints202108.0438.v1

Examination of long-term data on land and forest fires based on SiPongi site shows that one of the worst Indonesian land and forest fires events occurred in the year of 2015 during one of the strongest ever recorded El Nino event (Figs 8 and 9).

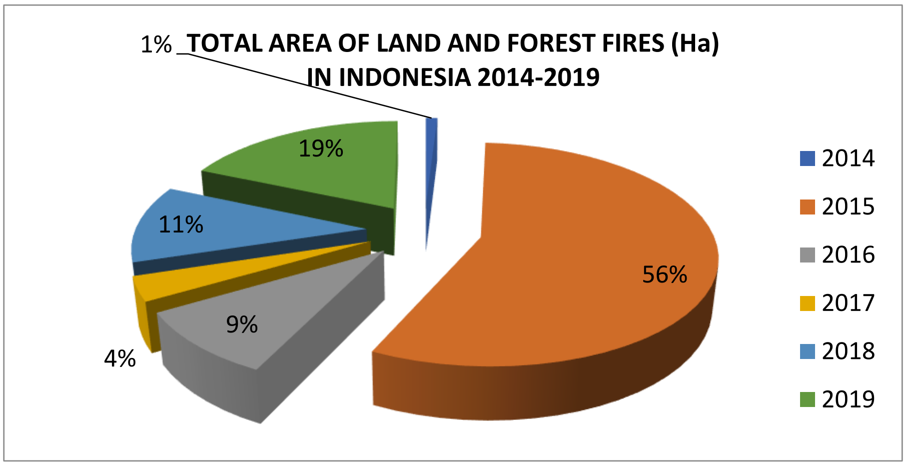

Figure 8. Summary of land and forest fires coverage area in Indonesia for 2014-2019 timeframe.

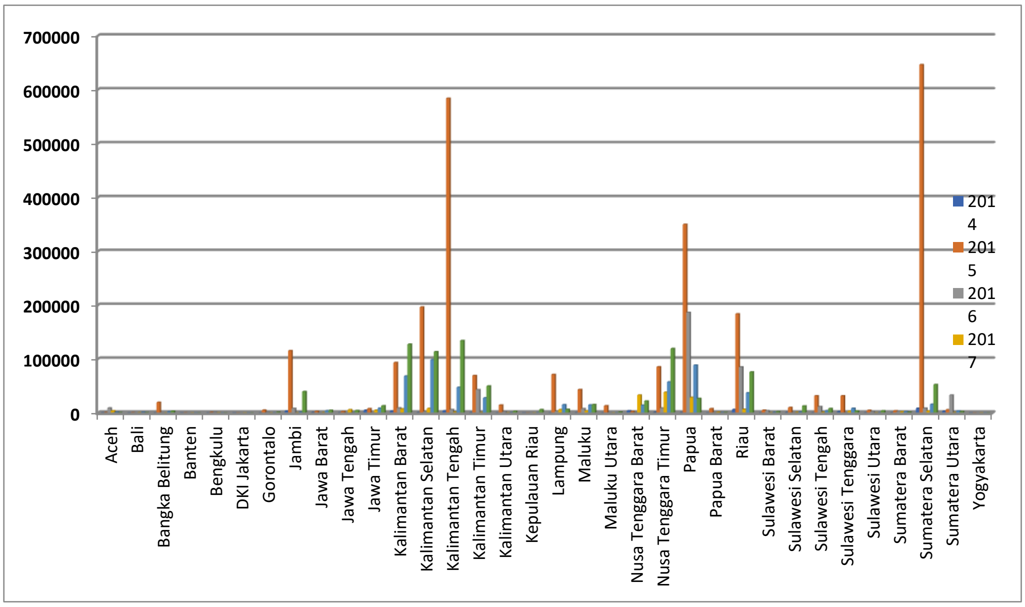

Figure 9. Summary of land forest and land fires areas (Ha) for each province in Indonesia in 2014-2019. Note: The areas of land and forest fires are calculated using the analysis of the OLI /TIRS satellite Landsat 8 images overlaid on hotspot distribution data as well as the hotspot ground check and blackout reports conducted by Manggala Agni Summary of land and forest fires coverage area in Indonesia for 2014-2019 timeframe.

Several studies indicate that forest fires in Indonesia are correlated with dynamical weather conditions and long-term climate indices. For example, Pan, et al [24] shows the connection between Indonesian fires and droughts with the El-Nino and the Indian Ocean Dipole phase from the year of 1979 to 2016 (Fig. 10).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 August 2021 doi:10.20944/preprints202108.0438.v1

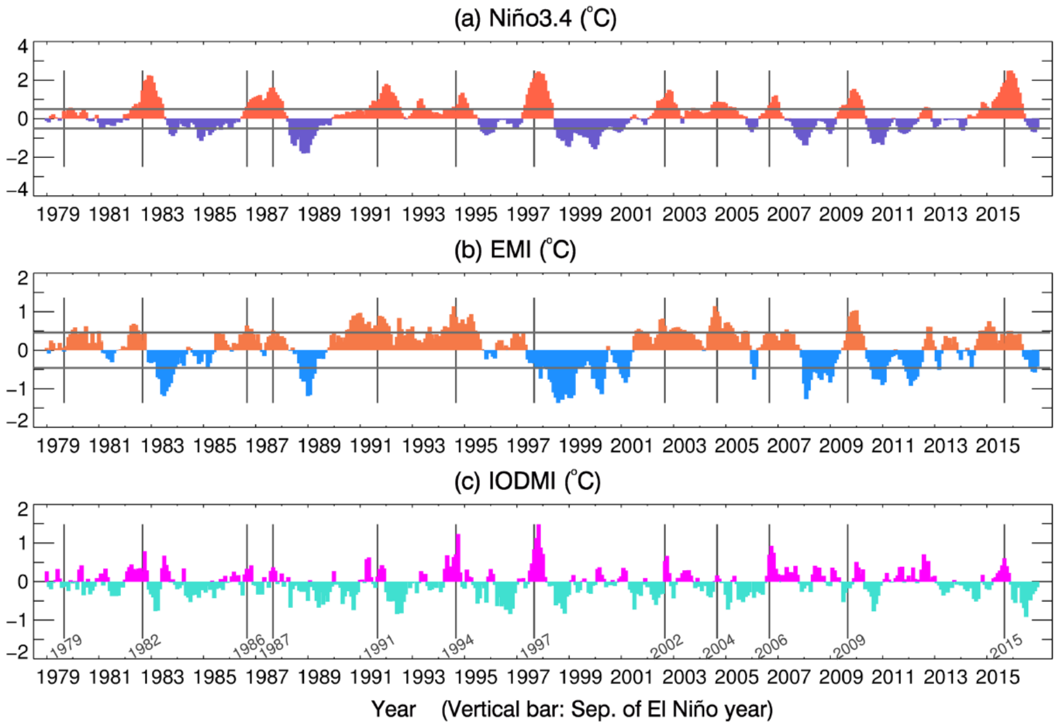

Figure 10. Time series of indices based on sea surface temperature anomalies from 1979 to 2016. Note: (a) Niño 3.4 (b) EMI, and (c) IODMI (red = positive IODMI and blue = negative IODMI). Month of September of El-Nino years, are marked by the vertical lines of grey colour in (a) – (c). horizontal lines of grey colour in (a) stand for the ±0.5oC thresholds (Nino 3.4 ≥ 0.5oC represent the El-Nino event), and horizontal lines of grey colour in (b) stand for the ±0.46oC thresholds (EMI ≥ 0.46oC represent the Central Pacific type event). EMI is the Index of El-Niño Modoki; IODMI is the Index for the Indian Ocean Dipole Model.

Considering the different El Niño regimes presence advanced the previous understanding of climate variability's role to the Indonesian fire activity. (ii) the interaction between El Niño and the Indian Ocean Dipole (IOD). The classify 12 El Niño events during 1979–2016 time period into Eastern Pacific (EP) and Central Pacific (CP) types (four and eight El Niño events, respectively) and analyse observational datasets of sea surface temperature, precipitation, drought code, carbon emission associated with biomass burning, optical depth of aerosol and the visibility. In the weather condition of EP type, Indonesian droughts and fires occurred more intensely and prolonged, where the carbon amount emitted is almost doubled if compared with the weather in CP type. When the El-Nino in CP type is separated for negative and positive phase of IOD, the intensity of less burning fire occurrences which has a shorter duration are mostly related to phenomena of the phase of IOD with weakly negative and positive phases. The intensity of fires also show the diversity of geographic. It is showed that the fires event in southern part of Kalimantan is always more fires are always more severe than the fires in the southern part of Sumatera during whole El-Nino events, in southern Kalimantan than in southern Sumatra in all El Niño events, even though Southern part of Kalimantan is less dry then in southern part of Sumatera.

Power, et al. [24] review that social media could act as an effective canal of communication. But, Anderson [13] informed that the use of social media has not been adopted. However, they state that its use as a source of situational awareness is not adopted commonly. That may be because of some mixture difficulties in framing the content of the social media, go through a large number of the available information and the trust issue in their content (Lindsay [14]). However, there has been a success result study on the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 August 2021 doi:10.20944/preprints202108.0438.v1

detection of earthquake disaster using social media (Sakaki, Okazaki and Matsuo [15]; Robinson, Power, and Cameron [16]; Avvenuti, Cresci, Marchetti, Meletti, and Tesconi [17]).

Study by Pan et al [24] shows the connection between fires occurrence in Indonesian and the droughts associated with the El-Nino and phase of Indian Ocean Dipole from 1979 to 2016 for the season period of fire in August to October (Figure 11).

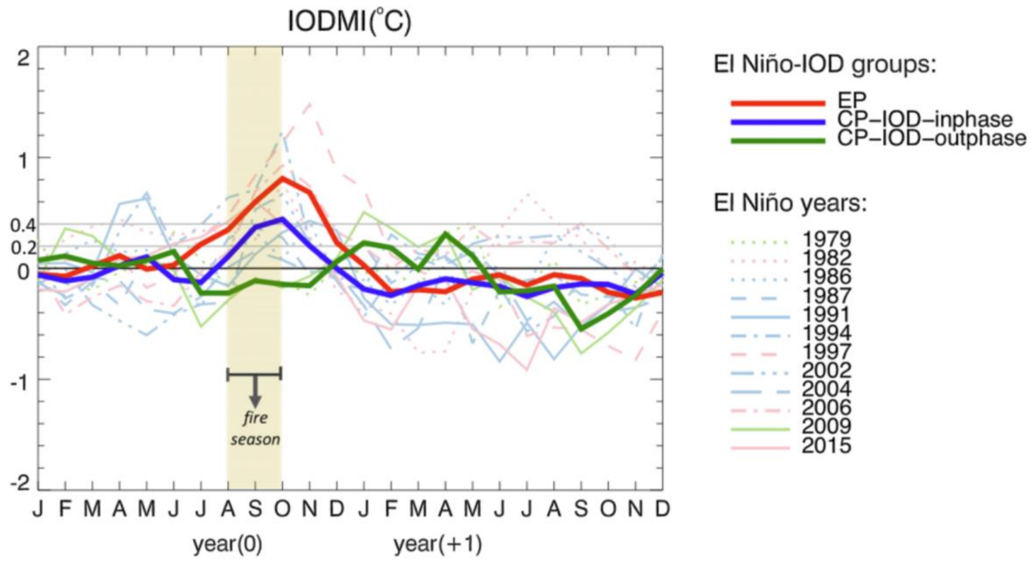

Figure 11. The 2-year monthly variation of the IODMI index in three El Niño-IOD groups and individual El Niño years (units:°C). Note: Year (0) represents the El Niño year as labelled in the legend, while the year (+1) represents the following year. The major fire season (August–September–October) in Indonesia is highlighted with a light brown shade. IODMI = Indian Ocean dipole model index; IOD = Indian Ocean Dipole; EP = Eastern Pacific; CP = Central Pacific.

According to Page et al. [31], climate variability and practices of land use are the cause of the Indonesian fires. Goldammer [32], Goldammer & Seibert [33], and Simorangkir [34] mentioned that the key role for the occurrences of fires are human activities to instantly open the land to be used for plantation and agriculture. Nevertheless, Y. Chen et al. [35] and Hendon [36] mentioned that the equatorial Pacific Ocean anomalies of sea surface temperature (SSTA) processes which is called as the El-Nino are regulating the intensity of fires in Indonesia. This condition is triggered by the drier of air masses in Indonesia because of the weakening of the Walker circulation on the years of El-Nino occurrences (Lim et al. [37]), which triggered the increased droughts and more severe fires (Field et al. [23]). Numerous studies have shown that the El-Nino events triggering the intensive fires in Indonesia (C. C. Chen et al. [35]; Fanin & van der Werf [38]; Field et al. [23]; Wooster et al. [30]; Yin et al. [39]). Pan et al. [24] recently studied that El-Nino events could be classified into two main categories or type based on the maximum SSTA location. They are the Eastern Pacific (EP or canonical) and the Central Pacific (CP or Moldoki) types (Ashok et al. [40]; Kao & Yu [41]; Yeh et al. [42]). Maximum warm SSTA maximum warm in the Eastern part of Pacific Ocean indicated the EP type, and the SSTA maximum warm in the Central part of Pacific Ocean indicated the EP type. Lee and McPhaden [43] informed that the El-Nino in CP type in recent decade have been more frequently appeared. Ashok & Yamagata [40], Kao & Yu [41], Yu & Kim [44], Wang & Wang [45], and Zhang et al. [46] studies showed that The different types of the El-Nino control walker circulation different locations and intensities. These then modifies the global circulation of atmosphere as well as the pattern of the precipitation by their correlation. The earlier studies of fires in Indonesia (Fanin and van der Werf [38]; Field et al. [23]) showed that the phase of Indian Ocean Dipole (IOD) might also modulate fires, particularly when occurring in conjunction with appropriate El Nino phase. According to Saji et al. [47], the

IOD is the phenomena of an ocean-atmosphere that occurred in the tropical Indian ocean created in a summer that is boreal, reaches its maximum height in autumn, and quickly falls of in winter. Interactions of the IOD and El-Nino El Niño and IOD create variation on the pattern of the precipitation for the different types of El-Nino (Wang & Wang [45]; Zhang et al. [46]). During the EP types of the El-Nino, the Indonesian drought and fires occurred prolongedly and more intensely, where the Carmon emission amounts become almost doubled compared to the CP types of El-Nino. When the CP type of El-Nino further separated into the positive and negative phase of IOD, it can be shown that weakly positive or even negative phase of the phenomena of the IOD caused the less burning intensity of the fire seasons.

4.2. The analysis of the newspaper data

We have analysed the content of 31 news of 2019 fires as reported in the online Riau Antara Newspaper (https://riau.antaranews.com). From the content news, it was noted that 2019 forest fires season started on December 31, 2018. The areas burnt were mostly the peatland type with some areas covered by shrub vegetation, the palm oil trees and the rubber plantation vegetation. In early 2019 some of the newspapers reported that forest fires resulted because of high temperatures in the dry season without rainfall and with strong winds and from land clearing for plantations and grass burning by the people.

On January 11, 2019 newspapers reported that burning area has reached up to 267,5 hectares and that the local disaster management in Riau considered “standby” (ready for the action) status related to forest and land fires in Riau, but to issue a final decision they needed to coordinate with various institutions such as the Indonesian Agency for Meteorology, Climatology, and Geophysics (BMKG) and the Riau Provincial Government. Subsequently, on February 18, 2019 the “standby” status related to land and forest fires had been decided for Bengkalis Regency and Dumai City but not yet determined at the whole Riau province level. It was also mentioned that the “standby” status could be issued at the provincial level if two more regencies or cities determined that such status would be warranted. It took more than a month that a “standby” status for forest and land fires in the Province of Riau was reached up. By that time 841,71 hectares in Riau was burnt. It was also mentioned that once the Riau Provincial Government determined the “standby” status regarding land and forest fires, they could immediately coordinate with the National Disaster Management Agency (BNPB) to get support and assistance, including asking for additional information helicopters and others resources.

Some of the newspapers also reported difficulties in the field during efforts to extinguish fires. Locations are sometimes difficult to reach because of no access to roads or difficult terrain, where use of small boats through mangrove forest is required or when the access is by foot only. On occasion, winds are so strong that fires spread too quickly to react.

In summary, the constraints of using newspapers in this study are related to lack of information on the exact location of fires. Also, newspaper information is often delayed, and fire start times are not provided. However, newspapers do provide insights into problems that occur in the field when fire extinguishing is being conducted and how the Government responds facing many regulations, which often result in the down-top approach creating delayed actions.

4.3. The analysis of the Twitter data

Twitter data on land and forest fires in Riau from 2015 to 2019 was compared with the occurrences of fires from MODIS satellite images for the same period (Figure 12). These plots comparing number of tweets and the number of hotspots show some consistency between them. For example, when the number of hotspots in 2015, 2016 and 2019 is large, the number of tweets is also high.

 doi:10.20944/preprints202108.0438.v1

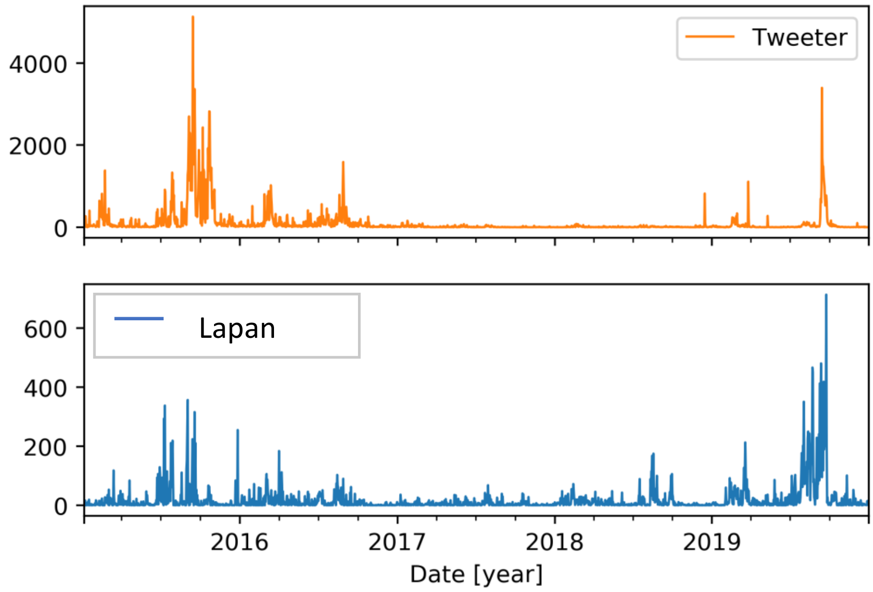

Figure 12. Number of tweets and hotspot occurrence from MODIS satellite images in Riau from 2015 to 2019.

4.4 The Empirical analysis on Social Media Response

The Annual trend of the scatterplot between hotspot and the twitter from 2015 to 2019 (figure 10) showed that the response from hotspot occurrences in twitter (R2) were very low, only about 4.4% in 2015, 22.5% in 2016, 0.1% in 2017, 12.3% in 2018, and 8.9% in 2019, and the correlation between hotspot and twitter data for the whole year ranges between 0,04 to 0,39. It may be because of insufficient data in some months where there are no tweet or no hotspot. However, for dry season between August to October when drought conditions occur in Riau there is better correlation, between 0,21 to 0,47 (Table 1). When hotspot occurrences are high as in 2015 and 2019, the correlation is lower comparing to years when the hotspot occurrences are smaller (Figure 13).

 doi:10.20944/preprints202108.0438.v1

Table 1. Yearly correlation between number of hotspot and number of tweets in Riau from 2015 to 2019

Year	Tweet-Hotspot Correlation (Jan-Dec)	Tweet-Hotspot Correlation (Aug-Oct)
2015	0,271357568	0,209109939
2016	0,384721617	**0,47389312**
2017	0,053482047	0,209109939
2018	0,042049973	**0,47389312**
2019	0,387262216	0,20602366

Source: Authors' calculation

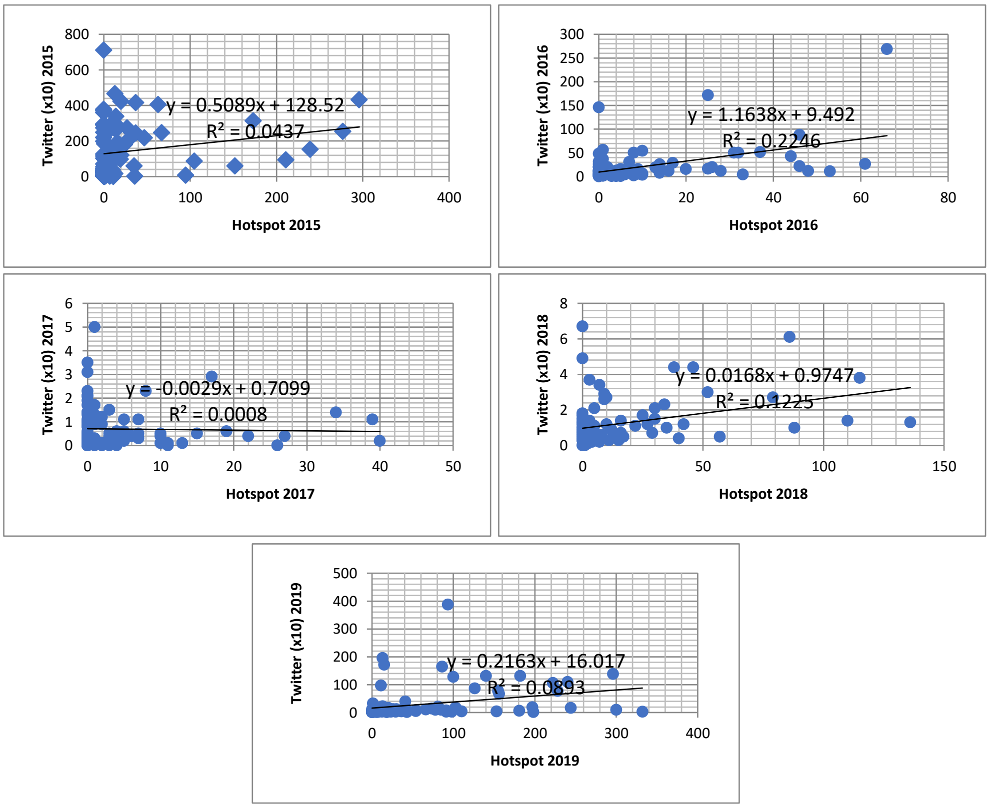

Figure 13. Annual trend of the scatterplot between hotspot and the twitter from 2015 to 2019

There is also a correlation between monthly hotspot and tweet numbers and during 2015 and 2019, such correlation are high, 71% and 84%, respectively (Fig. 14 and Table 2).

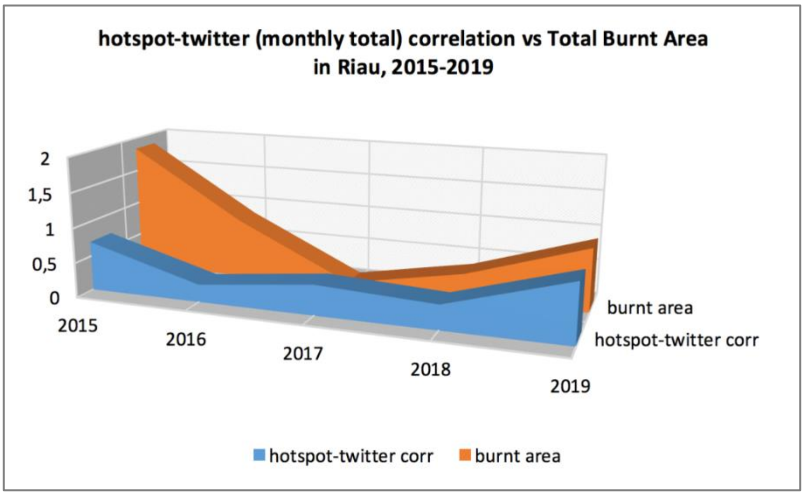

Figure 14. Total hotspot-twitter correlation and total burnt area in Riau, 2015-2019

Tabel 2. Hotspot-twitter (monthly total) correlation and total burnt area in Riau, 2015-2019

year	monthly hotspot-monthly twitter correlation	Total Burnt Area (in 100000)
2015	0,709382862	1,8380859
2016	0,251661499	0,8521951
2017	0,433118632	0,0686609
2018	0,35067847	0,3723627
2019	0,84346165	0,9055

Source: Authors' calculation

Dry season (August-September-October) fires in Indonesia correlate with Nino 3.4 which is positive, and the indices of IOD as seen in Fig. 8 for years 2015, 2018, and 2019. Positive Niño 3.4 and IOD indices in 2015, 2018, and 2019 are also correlated with the total burnt area as seen in Figure 2. This agrees with previous studies of Pan et al. [24] and others. We should note that fires that occurred in early 2016 may be related to the strong positive 2015 Niño 3.4 which lasted until May 2016.

Social media information from Newspapers and twitter data provide important information that can support land and forest fire management decisions. For example, one can derive such information as fire locations, time of fire occurrences, impact and risk faced by the communities, health and social issues related to smoke and fires, causes of fires and problems experienced in the field during fire extinguishing.

From the content analysis of the newspaper, requirements needed to establish the "standby status" for land and forest fires on the provincial level which results in support from the central government is often too long. This leads to worsening conditions of fire spreading and containment. The down-top approach would be fine for the safe conditions

 doi:10.20944/preprints202108.0438.v1

of weather and climate. But, top-down approach is strongly required during the extreme conditions of weather and climate. If the extreme weather and climate conditions could be predicted, a more flexible top-down approach could be applied.

The twitter analysis shows that the total monthly tweets and total monthly hotspots are well correlated, and this provides a beneficial information for usage of twitter data. Newspapers and twitter information associated with the forest fires show that they can provide additional near-real time information on forest fires detection that is beneficial for the forest fires management policy maker in deciding the fires prevention and mitigation. The evolving tweet might serve as an alert for decision makers to consider and determine the most appropriate steps needed to be taken to address problems on the ground.

Clearly there is no 1:1 correspondence between tweeter data and satellite hotspots. There can be several reasons for this: (1) Tweets can begin before satellite hotspots are observed particularly where fires are just originating (starting) and satellite detection algorithms are not sensitive enough. On the other hand, satellite hotspots may be observed in remote areas where mobile networks do not have sufficient coverage or where there are not many people. In such a case hotspots will occur first, before tweets, (2) Another reason is a tweet "memory" effect in which tweets grow with time and "explode" when the intensity of the event is large. There may be also a long decay of tweets after a particularly intense fire event. In other words, tweets are "social" phenomena". People tweet about health issues, displacement, or discuss the event. This is why we are observing long "tweeter memory", (3) There may be also discrepancy because some events lead to tweets which are not from a locality in which fire occurred. For example, relatives in Jakarta, may tweet about particularly large fires in Riau.

More research such as artificial intelligence Natural Language Processing (NLP) could be used to analyse fire related twitter messages and their discrepancies with hotspot data, For example, in context of tweeter and floods de Bruijin et al [48] adopted the cased multilingual version of BERT to categorize tweets into 2 groups, one is the group which is associated with the ongoing event ("relevant") and second is the group which is not associated with the ongoing event ("irrelevant"). BERT processes the correlations between the words and sub-words in a text using the Natural Language Processing (NLP) model which is based on the deep learning. However, such an analysis is beyond the scope of current work.

Even though tweets are "social media" our research illustrates of how objectively derived information from hotspots correlate with tweets, thus providing a glimpse of how they can be used in future events for proper policy making. For example, large events are well captured by tweets. There is strong, sharp rise in tweet amount and such gradient in tweet amounts can be used. We also observed "memory effect" in which tweets are more "sentiment" related, but nevertheless may be important for policy makers because even though the fires were extinguished but they are still discussed in villages and cities.

5. Conclusions

In this work, we conclude that social media fire information combined with the data on forest and land fires such as satellite derived hotspots, wind velocity and speed, rainfall and moisture content provides additional source of information which can be used for near-real time forecasting and polic decision making. The newspapers do provide insights into problems that occur in the field when fire extinguishing is being conducted and how the Government responds facing many regulations, which often result in the down-top approach creating delayed actions. We also conclude that large scale weather and oceanic state indices such as Indian Ocean Dipole and El Nino were well correlated with forest fires occurrences, therefore, providing additional estimation variable for forest fires detection. This work is relevant to the Riau province land and forest fires which are related to the existence of peat soil and land management practices. We notice that such practices lead to this region's high frequency of fire activities mostly during the dry seasons. It is observed that the remoteness of these areas is a factor in attempting to develop alternative detection tools for warning such as Twitter messages. The Government of

Indonesia has tried many ways to prevent and mitigate the occurrence of forest and land fires in Indonesia. In this study, existing approaches taken by the BMKG and other Indonesian agencies in estimating fire activities are reviewed, and novel approach based on crowdsourcing of tweeter is proposed.

We notice that social media together with satellite data and large-scale weather indices such as Indian Ocean Dipole and El Nino as well as mesoscale weather states such as passages of Convectively Coupled Equatorial Waves may be needed to increase fires and smoke timely detection and warnings. To this end, we suggest:

1. The top-down approach is required for the extreme weather and climate condition related to land and forest fires.
2. Since the dynamic condition of weather and climate phenomena is strongly correlated to the fire intensity in Indonesia, the Government needs to include them in its Fire Rating Danger System and forest fires vulnerability map through the identification of the El-Nino type and phase of IOD and including even shorter time scale events such as the Convectively Coupled Equatorial Waves and Madden-Julian Oscillations.
3. Newspaper and Twitter data could also be included in the fire management system in the country to support the early detection for warnings of the fires.

This study can be extended in several ways. The development of land and forest fire disaster threats index could be enhanced by combining the comprehensive information from the dynamical weather information about the related weather and climate conditions of the forest fires. In particular, research on forecasting of short terms weather events such as Convectively Coupled Kelvin Waves and short scale tropical disturbances such as Mesoscale Convective Systems and information of sudden onsets of droughts – so called "flash draughts"(c.f. Pendergrass et al. [49]) which may contribute to fires. Incentives to provide Twitter information could be developed, and enhanced mobile signal networking capabilities in remote areas would increase the viability of our proposed approach.

Author Contributions: Conceptualization, A.A.F. and P.J.F..; methodology, A.A.F, P.J.F. and K.; software, P.J.F.; validation, A.A.F. and P.J.F.; investigation, A.A.F.; writing—original draft preparation, A.A.F.; writing—review and editing, A.A.F., P.J.F., K. and N.F.R.; visualization, A.A.F. and P.J.F.; supervision, K.; project administration, K. and N.F.R. All authors have read and agreed to the published version of the manuscript.

Funding: P.J.F. has been supported in part by NSF grant 1724741 "Equatorial Line Observations (ELO) field campaign during the International Years of Maritime Continent" program managed by Eric T. DeWeaver. The Agency for Meteorology, Climatology, and Geophysics of the Republic of Indonesia (BMKG) contribution supported the Years of Maritime Continent ELO project.

Data Availability Statement: The satellite related to this paper can be downloaded from LAPAN site http://modis-catalog.lapan.go.id/monitoring/. El-Nino and Indian Ocean Dipole (IOD) Indices are available from https://psl.noaa.gov/gcos_wgsp/Timeseries/DMI/. All the newspaper data were collected by searching on-line open newspaper sources. The tweeter datasets analyzed and generated during the current study are not publicly available due to Twitter's privacy policy but are available from the corresponding author upon a reasonable request in line with the policy.

Code Availability The TAGGS code is publicly available on GitHub (https://github.com/jensdebruijn/TAGGS). The tweeter code can be found on the github site: https://github.com/Jefferson-Henrique/GetOldTweets-python.

Acknowledgments: The satellite hotspot data from MODIS satellite sensor on Terra and Aqua satellites and SNPP were obtained from LAPAN MODIS monitoring site. The Niño 3,4 index data are taken from the Climate Prediction Center (CPC) of the National

Weather Services, National Oceanic and Atmospheric Administration (NOAA). Authors acknowledge help of Maria Flatau with discussion of large scale indices. Michał Łabuz processed the TAGGS algorithm and collected twitter data. Katarzyna Barabasz helped with the analysis of the newspaper data and Dawid Gacek helped with data plotting. Ms. Tracy Stevens, which helped to improve the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tacconi, L. Fires in Indonesia: Causes. *Cost and Policy Implications (Bogor: CIFOR)* **2003**.
2. World Bank Indonesia Economic Quarterly, December 2019: Investing in People. **2019**.
3. Kibanov, M.; Stumme, G.; Amin, I.; Lee, J.G. Mining Social Media to Inform Peatland Fire and Haze Disaster Management. *Social Network Analysis and Mining* **2017**, *7*, 1–19.
4. Holmgren, P. Fire and Haze in Riau: Looking beyond the Hotspots. *Center for International Forestry and Research, Bogor, Indonesia* **2013**.
5. Kim, D.; Sexton, J.O.; Townshend, J.R. Accelerated Deforestation in the Humid Tropics from the 1990s to the 2000s. *Geophysical Research Letters* **2015**, *42*, 3495–3501.
6. Frankenberg, E.; McKee, D.; Thomas, D. Health Consequences of Forest Fires in Indonesia. *Demography* **2005**, *42*, 109–129.
7. Ulya, N.A.; Yunardy, S. Analisis Dampak Kebakaran Hutan Di Indonesia Terhadap Distribusi Pendapatan Masyarakat. *Jurnal penelitian sosial dan ekonomi kehutanan* **2006**, *3*, 133–146.
8. Rogers, D.; Tsirkunov, V. Costs and Benefits of Early Warning Systems. *Global assessment rep* **2011**.
9. Hallegatte, S. A Cost Effective Solution to Reduce Disaster Losses in Developing Countries: Hydro-Meteorological Services, Early Warning, and Evacuation. *World Bank policy research working paper* **2012**.
10. Biswas, M. An Exploratory Research: A Comparative Analysis of Mainstream and Ethnic Media Coverage of Social Policy Issues in the Economic Stimulus Plan Debate. *Journal of Comparative Social Welfare* **2010**, *26*, 13–26.
11. Kalogeropoulos, A.; Negredo, S.; Picone, I.; Nielsen, R.K. Who Shares and Comments on News?: A Cross-National Comparative Analysis of Online and Social Media Participation. *Social media+ society* **2017**, *3*, doi:10.1177/2056305117735754.
12. Power, R.; Robinson, B.; Colton, J.; Cameron, M.A. A Case Study for Monitoring Fires with Twitter. In Proceedings of the ISCRAM; 2015.
13. Anderson, M. Integrating Social Media into Traditional Emergency Management Command and Control Structures: The Square Peg into the Round Hole. In Proceedings of the Disaster and Emergency Management Conference Proceedings, 16-18 April 2012. Brisbane, Australia; 2012; pp. 18–34.
14. Lindsay, B.R. Social Media and Disasters: Current Uses, Future Options, and Policy Considerations. *Publisher: Congressional Research Service Washington, DC* **2011**.
15. Sakaki, T.; Okazaki, M.; Matsuo, Y. Tweet Analysis for Real-Time Event Detection and Earthquake Reporting System Development. *IEEE Transactions on Knowledge and Data Engineering* **2012**, *25*, 919–931.
16. Robinson, B.; Power, R.; Cameron, M. A Sensitive Twitter Earthquake Detector. In Proceedings of the Proceedings of the 22nd international conference on world wide web; 2013; pp. 999–1002.
17. Avvenuti, M.; Cresci, S.; Marchetti, A.; Meletti, C.; Tesconi, M. Ears (Earthquake Alert and Report System) a Real Time Decision Support System for Earthquake Crisis Management. In Proceedings of the Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining; 2014; pp. 1749–1758.

18. de Bruijn, J.A.; de Moel, H.; Jongman, B.; Wagemaker, J.; Aerts, J.C. TAGGS: Grouping Tweets to Improve Global Geoparsing for Disaster Response. *Journal of Geovisualization and Spatial Analysis* **2018**, *2*, 2.
19. Stieglitz, S.; Dang-Xuan, L. Social Media and Political Communication: A Social Media Analytics Framework. *Social network analysis and mining* **2013**, *3*, 1277–1291.
20. Sinnenberg, L.; Buttenheim, A.M.; Padrez, K.; Mancheno, C.; Ungar, L.; Merchant, R.M. Twitter as a Tool for Health Research: A Systematic Review. *American journal of public health* **2017**, *107*, e1–e8.
21. Power, R.; Robinson, B.; Ratcliffe, D. Finding Fires with Twitter. In Proceedings of the Proceedings of the Australasian Language Technology Association Workshop 2013 (ALTA 2013); 2013; pp. 80–89.
22. Reid, J.; Xian, P.; Hyer, E.; Flatau, M.; Ramirez, E.; Turk, F.; Sampson, C.; Zhang, C.; Fukada, E.; Maloney, E. Multi-Scale Meteorological Conceptual Analysis of Observed Active Fire Hotspot Activity and Smoke Optical Depth in the Maritime Continent. *Atmospheric Chemistry and Physics* **2012**, *12*, 2117–2147.
23. Field, R.; Shawki, D.; Tippett, M.; Saharjo, B.H.; Albar, I.; Atmoko, D.; Voulgarakis, A. Long-Lead Prediction of the 2015 Fire and Haze Episode in Indonesia.
24. Pan, X.; Chin, M.; Ichoku, C.M.; Field, R.D. Connecting Indonesian Fires and Drought with the Type of El Niño and Phase of the Indian Ocean Dipole during 1979–2016. *Journal of Geophysical Research: Atmospheres* **2018**, *123*, 7974–7988.
25. Forsyth, T. Public Concerns about Transboundary Haze: A Comparison of Indonesia, Singapore, and Malaysia. *Global Environmental Change* **2014**, *25*, 76–86.
26. Panjaitan, R.B.; Sumartono, S.; Sarwono, S.; Saleh, C. The Role of Central Government and Local Government and the Moderating Effect of Good Governance on Forest Fire Policy in Indonesia. *Benchmarking: An International Journal* **2019**, *26*, 147–159.
27. Carley, K.M.; Malik, M.M.; Kowalchuck, M.; Pfeffer, J.; Landwehr, P. Twitter Usage in Indonesia (Https://Ssrn.Com/Abstract=2720332). **2015**.
28. Baranowski, D.B.; Flatau, M.K.; Flatau, P.J.; Karnawati, D.; Barabasz, K.; Labuz, M.; Latos, B.; Schmidt, J.M.; Paski, J.A. Social-Media and Newspaper Reports Reveal Large-Scale Meteorological Drivers of Floods on Sumatra. *Nature communications* **2020**, *11*, 1–10.
29. Aditya, T.; Laksono, D.; Izzahuddin, N. Crowdsourced Hotspot Validation and Data Visualisation for Location-Based Haze Mitigation. *Journal of Location Based Services* **2019**, *13*, 239–269.
30. Wooster, M.; Perry, G.; Zoumas, A. Fire, Drought and El Niño Relationships on Borneo (Southeast Asia) in the Pre-MODIS Era (1980–2000). *Biogeosciences* **2012**, *9*, 317–340.
31. Page, S.; Hoscilo, A.; Langner, A.; Tansey, K.; Siegert, F.; Limin, S.; Rieley, J. Tropical peatland fires in Southeast Asia. In *Tropical fire ecology*; Springer, 2009; pp. 263–287.
32. Goldammer, J.G. History of Equatorial Vegetation Fires and Fire Research in Southeast Asia before the 1997–98 Episode: A Reconstruction of Creeping Environmental Changes. *Mitigation and Adaptation Strategies for Global Change* **2007**, *12*, 13–32.
33. Goldammer, J.G.; Seibert, B. The impact of droughts and forest fires on tropical lowland rain forest of East Kalimantan. In *Fire in the tropical biota*; Springer, 1990; pp. 11–31.
34. Simorangkir, D. Fire Use: Is It Really the Cheaper Land Preparation Method for Large-Scale Plantations? *Mitigation and adaptation strategies for global change* **2007**, *12*, 147–164.
35. Chen, Y.; Morton, D.C.; Andela, N.; Van Der Werf, G.R.; Giglio, L.; Randerson, J.T. A Pan-Tropical Cascade of Fire Driven by El Niño/Southern Oscillation. *Nature Climate Change* **2017**, *7*, 906–911.
36. Hendon, H.H. Indonesian Rainfall Variability: Impacts of ENSO and Local Air–Sea Interaction. *Journal of Climate* **2003**, *16*, 1775–1790.

 doi:10.20944/preprints202108.0438.v1

37. Lim, Y.-K.; Kovach, R.M.; Pawson, S.; Vernieres, G. The 2015/16 El Niño Event in Context of the MERRA-2 Reanalysis: A Comparison of the Tropical Pacific with 1982/83 and 1997/98. *Journal of climate* **2017**, *30*, 4819–4842.
38. Fanin, T.; Werf, G.R. Precipitation–Fire Linkages in Indonesia (1997–2015). *Biogeosciences* **2017**, *14*, 3995–4008.
39. Yin, Y.; Ciais, P.; Chevallier, F.; Van der Werf, G.R.; Fanin, T.; Broquet, G.; Boesch, H.; Cozic, A.; Hauglustaine, D.; Szopa, S. Variability of Fire Carbon Emissions in Equatorial Asia and Its Nonlinear Sensitivity to El Niño. *Geophysical Research Letters* **2016**, *43*, 10–472.
40. Ashok, K.; Behera, S.K.; Rao, S.A.; Weng, H.; Yamagata, T. El Niño Modoki and Its Possible Teleconnection. *Journal of Geophysical Research: Oceans* **2007**, *112*.
41. Yu, J.-Y.; Kao, H.-Y.; Lee, T.; Kim, S.T. Subsurface Ocean Temperature Indices for Central-Pacific and Eastern-Pacific Types of El Niño and La Niña Events. *Theoretical and Applied Climatology* **2011**, *103*, 337–344.
42. Yeh, S.-W.; Kug, J.-S.; Dewitte, B.; Kwon, M.-H.; Kirtman, B.P.; Jin, F.-F. El Niño in a Changing Climate. *Nature* **2009**, *461*, 511–514.
43. Lee, T.; McPhaden, M.J. Increasing Intensity of El Niño in the Central-equatorial Pacific. *Geophysical Research Letters* **2010**, *37*.
44. Yu, J.; Kim, S.T. Identifying the Types of Major El Niño Events since 1870. *International journal of climatology* **2013**, *33*, 2105–2112.
45. Wang, X.; Wang, C. Different Impacts of Various El Niño Events on the Indian Ocean Dipole. *Climate dynamics* **2014**, *42*, 991–1005.
46. Zhang, W.; Wang, Y.; Jin, F.; Stuecker, M.F.; Turner, A.G. Impact of Different El Niño Types on the El Niño/IOD Relationship. *Geophysical Research Letters* **2015**, *42*, 8570–8576.
47. Saji, N.; Goswami, B.N.; Vinayachandran, P.; Yamagata, T. A Dipole Mode in the Tropical Indian Ocean. *Nature* **1999**, *401*, 360–363.
48. de Bruijn, J.A.; de Moel, H.; Jongman, B.; de Ruiter, M.C.; Wagemaker, J.; Aerts, J.C. A Global Database of Historic and Real-Time Flood Events Based on Social Media. *Scientific data* **2019**, *6*, 1–12.
49. Pendergrass, A.G.; Meehl, G.A.; Pulwarty, R.; Hobbins, M.; Hoell, A.; AghaKouchak, A.; Bonfils, C.J.; Gallant, A.J.; Hoerling, M.; Hoffmann, D. Flash Droughts Present a New Challenge for Subseasonal-to-Seasonal Prediction. *Nature Climate Change* **2020**, *10*, 191–199.