



Wildfires and Subjective Well-being: Quantifying the Environmental and Health Impacts in Indonesia

Abstract

Wildfires are an increasingly frequent problem in Indonesia—a result of both climate change and intensive agriculture. In light of this problem, we use social media to study the impact that fires have on the people of Indonesia. We collected 12.9 million tweets posted in Indonesia over a 3 month period: from July 31, 2019 to November 10, 2019. For every tweet, we compute user sentiment using state-of-the-art Machine Learning and Natural Language Processing. We also extract conversation topics from the Twitter database and examine overall trends. Finally, we evaluate the impact of wildfires on subjective well-being in Indonesia, both on conversation—through topic modeling where we examine how often people discuss these wildfires—and on expressed sentiment—through an analysis of social media, to understand their impact on people’s subjective well-being. We find that both wildfires and general air pollution have negative impacts on subjective well-being. These results reveal an additional social cost of wildfires, and help justify their prevention and mitigation as a clear policy priority in Indonesia.

PRINCIPAL INVESTIGATOR: SIQI ZHENG

Samuel Tak Lee Professor of Urban and Real Estate sustainability, Department of Urban Studies and Planning, and Center for Real Estate, MIT; Faculty Director, MIT Sustainable Urbanization Lab

PROJECT INVESTIGATORS: AJKEL MINO NICOLAS GUETTA-JEANRENAUD JIANGHAO WANG



1 INTRODUCTION

In recent years, wildfires have become a major problem in Indonesia, as climate change extends drought periods and peatlands get replaced by large palm oil plantations (Dunne, 2019). These fires are the root cause for the Southeast Asian haze, a large-scale air pollution incident that occurs every year, to various degrees. It has been recorded since 1972 (Lee, 2015), and has affected multiple countries around Indonesia—namely Brunei, Malaysia, and Singapore—both in terms of health and economy (The Economist, 2013).

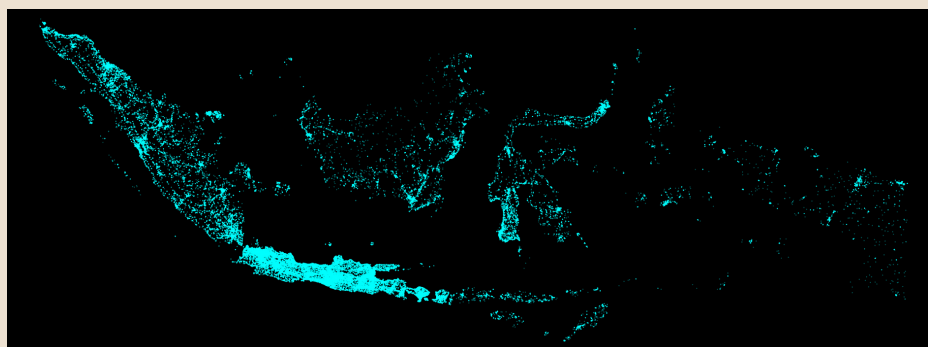
Existing literature documents the wide-ranging effects of wildfires, including on coral reef death (Abrams et al., 2003), carbon cycles (Shimel and Baker, 2002), and air pollution (Liu et al., 2016). Although the economic costs of these fires are well documented (Tacconi, 2003; Abriningrum et al., 2019), their negative impact on well-being is still largely ignored. Prior research has used social media data to show that environmental disasters can have significant emotional costs (Zheng et al., 2019; Baylis, 2020). However, none of these applications—to the best of our knowledge—have examined wildfires, nor focused on Indonesia specifically.

This project bridges the gap between these research questions, and produces a comprehensive analysis of the impact of Indonesian wildfires on subjective well-being. Drawing from the fields of environmental studies and computer science, and by leveraging public data drawn from Twitter, we evaluate the extent to which wildfires are discussed on social media, and their impact on communication topics and sentiment.

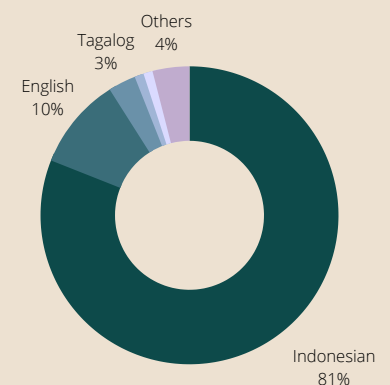
2 DATA AND METHODS

SOCIAL MEDIA DATA

Geotagged tweets posted in Indonesia between 2019 and 2020, have been collected. However, for the purpose of this project, we focus on a 3-month period from July 31, 2019 to November 10, 2019 when Indonesia declared an emergency alert on 6 provinces (IFRC, 2019). This dataset contains information regarding the exact location of the tweet, the tweet's text, as well as public information on the account doing the posting (number of followers, pages followed). Overall, our dataset covers the social media posts in the entire country. As is expected, the bulk of Twitter activity is concentrated in the most densely populated areas of the country, like Jakarta and Java (Fig. 1a).



(a)



(b)

Fig. 1: (a) Geographical coverage of Twitter data collected between July 31, 2019 and November 20, 2019. (b) Distribution of the languages of tweets posted in Indonesia.

SENTIMENT ANALYSIS AND TOPIC MODELING

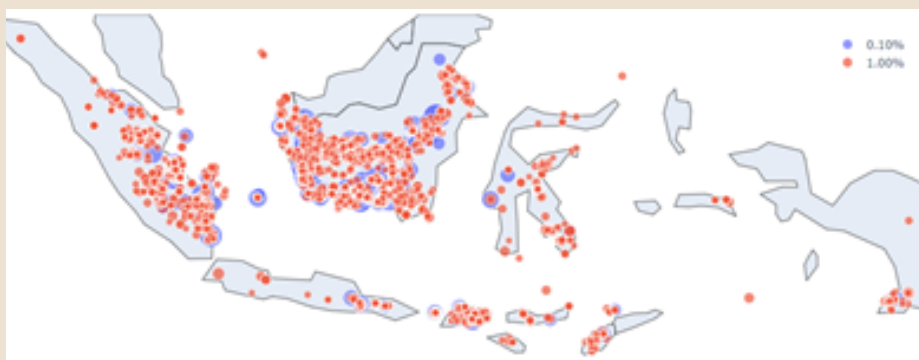
In order to extract relevant information from the Twitter data, we conduct sentiment analysis and topic modeling on the text of every tweet. These analyses are based on state-of-the-art methods of Natural Language Processing and Machine Learning. While most existing literature focuses on a single language of analysis (usually English), the methods we develop here are multilingual and tailored to Indonesia's social media landscape. Although a total of 61 languages are present in our Indonesian Twitter dataset, most of the posts are written in Indonesian. A significant share of the posts are in English, Tagalog, Spanish, and Turkish, while the rest of the languages are used less frequently (see Fig. 1b).

For the sentiment imputation, text entries were encoded into numerical representations using advanced techniques of Natural Language Processing called BERT. The Google AI-developed BERT (Devlin et al. 2018), is a pre-trained language model that has achieved exceptional performance levels on a diversity of natural language tasks. Here, we use the text representations—and hand-labeled training data—to build a sentiment classifier capable of tackling over 100 languages. This sets our approach apart from most dictionary-based approaches to sentiment imputation—which require as many dictionaries as there are languages in the data.

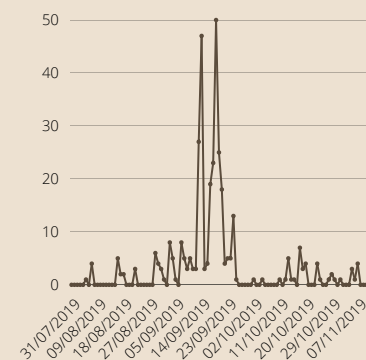
We also run preliminary models to identify the “Wildfires” topics in social media content. A key-word search methodology was used to understand how often users mention fires and smoke on twitter. Tweets mentioning the words “kebakaran” or “terbakar” were considered fire related, whereas those mentioning the words “asap” or “asbut” were considered smoke related.

WILDFIRE AND ENVIRONMENTAL DATA

Parallel to our analysis of social media data, we collected detailed wildfire data from the NASA Fire Information for Resource Management System (FIRMS) website. This data allows us to pinpoint the location of all the fires occurring in Indonesia. During our period of interest, 359,198 incidents took place in the country. Given the large number of fires, we choose to focus on the top 0.1% most severe ones (Fig. 2a). These were selected based on Fire Radiative Power (FRP), as calculated by MODIS (Giglio, 2000). The largest fires of 2019 took place in September (Fig. 2b).



(a)



(b)

Fig. 2: (a) Locations of the largest 1% (red) and 0.1% (blue) fires in Indonesia.
(b) Distribution of 0.1% largest wildfires over time.

We also gather various metrics that describe the air quality in Indonesia during this timeframe. Here, our data is collected from the NASA MERRA-2 model, and details the level of air pollutants available in each area at the time. Fine particulate matter with a diameter of 2.5 μm (PM-2.5) is a grouping of particles linked to damage in the human respiratory system (Xing, 2016). Hence we use this level as a measure of air pollution in our analysis, and construct a daily measure of PM-2.5 for every province in Indonesia.

3 FINDINGS AND CONCLUSIONS

The link between wildfires and air pollution has been well established in the scientific literature (Liu et al., 2016, Bo et al., 2020). The data we use here present the same intrinsic connection. On a province-day level, the number of large wildfires (top 1%) and the level of air pollution (PM-2.5) is highly correlated, with a Pearson coefficient of 0.47 (Fig. 3). Regressing the former on the latter – with environmental variable controls and province fixed-effects – yields a significant, positive coefficient. For the rest of our report, we consider the impact of both air pollution and wildfires on social media activity.

Topic analysis provides us with a first idea of the impact of wildfires on social media conversation in Indonesia. During the 3 month period of our analysis, there is not a lot of variation in the total number of tweets. However, the number of tweets on the topic of “wildfires” and “smoke” varies enormously (Fig. 4). A big increase can be observed between the 11th and 25th of September for both fire and smoke related tweets. At its peak, on the 17th of September, 0.9% of all tweets relate to smoke, and 0.4% relate to fire. This is consistent with the number of extreme wildfires taking place in the country: Fig. 2a highlighted that the peak in wildfire activity took place in September.

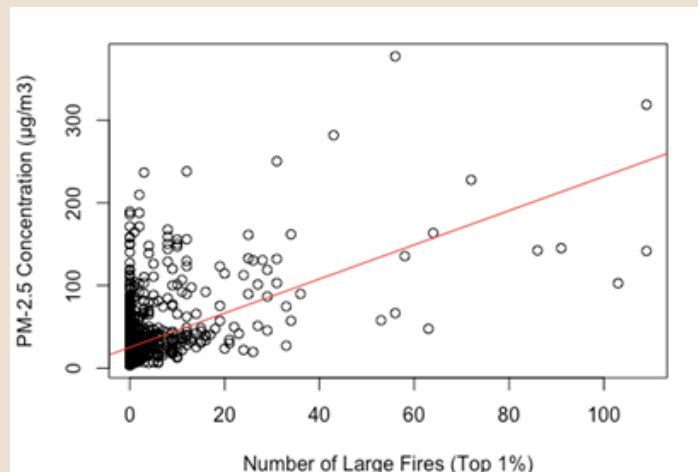


Fig. 3: Number of Large Fires by PM-2.5 Concentration, at province-day level.

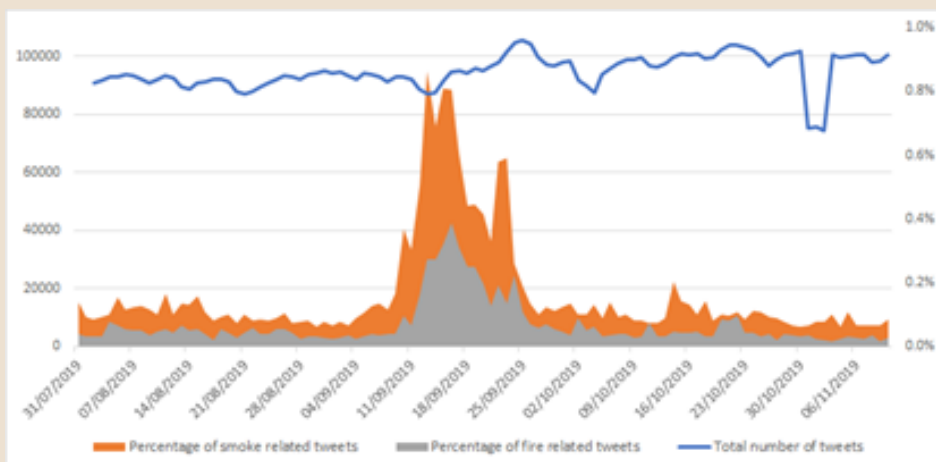


Fig. 4: Total number of tweets (3-day moving average), and share of wildfire and smog-related tweets in Indonesia.

In order to understand the impact of wildfires and resulting air pollution on sentiment in Indonesia, we aggregate our sentiment imputations to province-level and conduct a series of fixed-effect regressions. We control for other environmental variables (such as precipitation, temperature, and cloud coverage), as well as for day-of-week and province fixed effects. Our results show that both air pollution and wildfire proximity induce significant declines in expressed subjective well-being on social media. For air pollution, as PM-2.5 concentration increases, sentiment drops in a linear fashion, reaching a 3% drop for extreme pollution days (Fig. 5). The occurrence of a large wildfire in a province also accounts for a significant drop in wellbeing (0.5% drop in average sentiment).

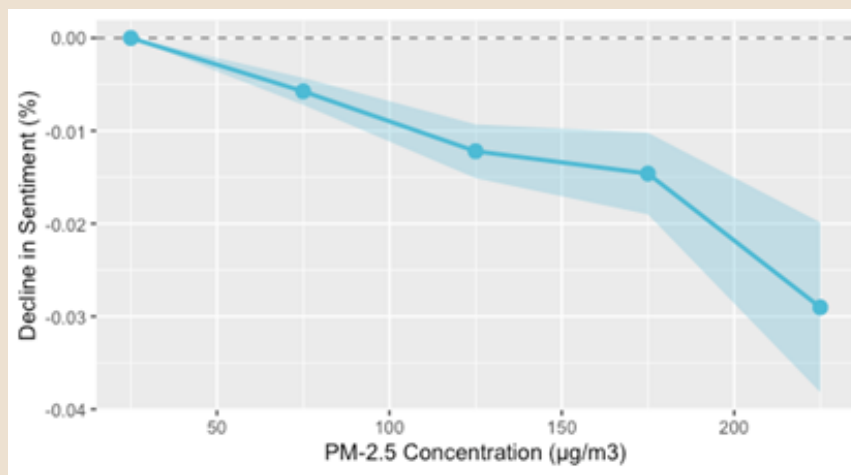


Fig. 5: Increased air pollution levels have a negative impact on expressed sentiment.

Overall, these results point to a significant emotional toll of wildfires on the Indonesian population. In part, wildfires are experienced directly by nearby inhabitants: wildfire season causes a significant increase in topical conversations on social media, and provinces with ongoing fires display consistently lower sentiment. However, the strongest effect on sentiment is observed through the indirect channel of air pollution. This supports the idea that even if people are not aware of ongoing natural disasters, their well-being is affected by long-lasting effects.

An understanding of these dynamics has important implications for governments in the fight against wildfires. These results open the door to a host of new costs—healthcare costs, well-being costs—when estimating the effects of wildfires, and justify additional investment in wildfire prevention and mitigation. A granular understanding of sentiment and keywords associated with these social media posts also provides valuable tools for targeted communication and effective environmental policy-making. They place the threat of wildfires front and center as a clear policy urgency, both for long-term benefit and short-term stability.

REFERENCES

- Abriningrum, D. E., Adriani M., Arsianti, Medina Y., Pinxten J., Rahmadanti, R.D., Rezza A. A., and Sari V.A.. (2019). *Indonesia Economic Quarterly, December 2019: Investing in People*, World Bank.
- Baylis, P. (2020). Temperature and temperament: Evidence from Twitter. *Journal of Public Economics*, 184, 104161.
- Bo, M., Mercalli, L., Pognant, F., Berro, D. C., & Clerico, M. (2020). Urban air pollution, climate change and wildfires: The case study of an extended forest fire episode in northern Italy favoured by drought and warm weather conditions. *Energy Reports*, 6, 781-786.
- Devlin, J., Chang, M. W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Dunne, D. (2019). The Carbon Brief Profile: Indonesia. *Carbon Brief*.
<https://www.carbonbrief.org/the-carbon-brief-profile-indonesia>
- E. H. (2013) "Why is South-East Asia's annual haze so hard to deal with?". *The Economist*.
- Giglio, Louis. (2000). MODIS Thermal Anomalies/Fire Products [Data set]. NASA Land Atmosphere Near real-time Capability for EOS Fire Information for Resource Management System. <https://doi.org/10.5067/FIRMS/MODIS/MCD14ML>
- IFRC (2019). Information Bulletin 01. Indonesia: Forest and Land Fires in 6 provinces in Borneo and Sumatra Island.
- Lee, Min Kok (2015). "Haze in Singapore: A problem dating back 40 years". *The Straits Times*. Archived from the original on 2 October 2015.
- Liu, J. C., Mickley, L. J., Sulprizio, M. P., Dominici, F., Yue, X., Ebisu, K., ... & Bell, M. L. (2016). Particulate air pollution from wildfires in the Western US under climate change. *Climatic change*, 138(3-4), 655-666.
- Tacconi, L. (2003). Fires in Indonesia: Causes. *Cost and Policy Implications* (Bogor: CIFOR).
- Xing, Y. F., Xu, Y. H., Shi, M. H., & Lian, Y. X. (2016). The impact of PM2. 5 on the human respiratory system. *Journal of thoracic disease*, 8(1), E69.
- Zheng, S., Wang, J., Sun, C., Zhang, X., & Kahn, M. E. (2019). Air pollution lowers Chinese urbanites' expressed happiness on social media. *Nature Human Behaviour*, 3(3), 237–243.