# Geophysical and anthropogenic drivers for global and regional fire emission trends from 2001 to 2019

**Sifeng Wu** ( sifeng001@e.ntu.edu.sg )
Nanyang Technological University
**Stuart William Smith**
Stockholm University
**Janice Ser Huay Lee**
Nanyang Technological University

**Article**

**Keywords:**

**Posted Date:** April 14th, 2022

**DOI:** https://doi.org/10.21203/rs.3.rs-1537229/v1

# Geophysical and anthropogenic drivers for global and regional fire emission trends from 2001 to 2019

**Sifeng Wu**[1,2,*], **Stuart William Smith**[1,3], **and Janice Ser Huay Lee**[1,*]

[1]Asian School of the Environment and Earth Observatory of Singapore, Nanyang Technological University, Singapore
[2]Complexity Institute, Interdisciplinary Graduate Programme, Nanyang Technological University, Singapore
[3]Department of Physical Geography, Stockholm University, Sweden
[*]sifeng001@e.ntu.edu.sg
[*]janicelee@ntu.edu.sg

## ABSTRACT

Fire plays an important role in the earth system. Carbon emissions from fires affect the global carbon budget and consequently climate change. Biome-specific qualities such as vegetation, climate and human land-use change are altering long-term trends in carbon emissions from fires. Whilst some biomes and continents seem to witness increasing fire-derived carbon emissions, others see decreasing trends. Yet, currently a biome-wide perspective on fire-derived carbon emission trends is lacking. We used total carbon emissions from the Global Fire Emissions Database (version 4 with small fires) to investigate the spatial and temporal dynamics of fire emissions from 2001 to 2019 across the world. While the global trend for fire emissions stayed relatively unchanged (non-significant trend), this was due to contrasting trends in two geographical biomes. Specifically, a decreasing trend in tropical savannas and grasslands, notably across the African continent and an increasing trend in boreal forests, especially a dense cluster in the east Siberian taiga in Boreal Asia. The decrease in emissions in African savannas and grasslands was driven by a decline in vegetation, and thus fuel, and an increase in anthropogenic intervention, especially agricultural expansion. Increase in fire emissions from boreal forests in Boreal Asia was driven by agricultural activities and climate change towards a drier climate (e.g., lower humidity). For many biomes, fire-derived carbon emissions are driven by several anthropogenic activities, vegetation and climate drivers, likely due to the complex feedback governing emissions. This global and biome-wide study highlights that anthropogenic activities in relatively small regions can shape global fire-derived carbon emission trends. Monitoring and management interventions are needed to address increasing fire-derived carbon emission areas, particularly the east Siberian taiga, and other forested biomes where deforestation contributes to rising carbon emission trends.

# 1 Introduction

Carbon emissions from fires contribute around 2.2 gigatones of carbon per year globally during 1997–2016[1]. Part of this carbon released by fires is reabsorbed by ecosystems through biochemical processes in the vegetation and soils[2,3]. Long-term effects of fires and the emitted carbon vary depending on the region and biome. For example, in African savannas and grasslands, regular fires are part of the historical natural cycle of the ecosystem's regeneration process[4]. The emitted carbon from fires has been documented to be reabsorbed in roughly equivalent amounts through this regeneration process over a period of years to decades and such carbon emissions are considered as "fast respirations"[1,2]. However, for ecosystems such as boreal forests and forested peatlands in Equatorial Asia, fires can alter the carbon mass balance and turn the ecosystem into a net carbon source, because the carbon stored in the soil and vegetation are emitted to the atmosphere. This alteration can last for decades to centuries because of the impaired ecosystem functions to support vegetation regrowth and accumulation of soil carbon. Thus, carbon emissions from fires can have long-term effects on the global carbon cycle[2,3,5].

Over the past decades, fire regimes have shifted due to climate change and anthropogenic intervention across the world[6–9]. The decline in global burnt area has been extensively studied and shown to be mainly driven by agricultural expansion and human suppression of fires[10–12]. However, global carbon emissions from fires are relatively less studied and understood, especially in terms of the temporal trends and their distribution among regions and biomes, and the drivers of these spatially differentiated trends. It is important to investigate the temporal trends of carbon emissions from fires and the spatial distribution due to potentially contrasting effects on the global fire emission carbon budget. Modelling approaches and satellite derived data for global fire emissions covering the past two decades have been established and are constantly evolving with technological developments[1,3,13]. These datasets have been used in regional and global studies of fire emissions at a coarse spatial scale, often at the regional or ecosystem level[14,15], but have not been used to assess the spatial-temporal dynamics of global carbon emissions from fires at a higher resolution nor the drivers of these trends.

Fires and fire-derived carbon emissions are influenced by a wide range of geophysical and anthropogenic factors. Climate is a widely studied controlling factor for fires, and among all the climatic factors, rainfall and temperature are generally the most important factors for controlling fire regimes[6]. In addition, extreme climatic condition, such as extended periods of drought, increases the likelihood of extreme fire events in many regions[14,16,17]. For fire emissions, the availability of fuel and completeness of combustion have a large effect on the amount of carbon released from fires[1,18]. Fuel is typically measured by vegetation cover or primary plant productivity, and in some ecosystems such as peatlands, soil carbon. Combustion completeness is a measure of the percentage of burnt material against all available fuel, and is also impacted by geophysical factors such as wind and humidity[1,12,19]. Besides geophysical factors, anthropogenic changes to ecosystem landscape can determine trends in fire emissions. Human induced fires could be used for clearing the forests to generate grazing pasture or cropland; and suppression can be done by removing vegetation and preventing accumulation of fuel[9,10,12,20,21]. The impact of anthropogenic interventions on fires can be complex and differ depending on the intervention, ecosystem, and the interaction between ecosystem and human activity[22].

Here, we conduct a comprehensive spatial-temporal analysis for global carbon emissions from fires and evaluate the geophysical and anthropogenic drivers of these trends specific to world regions and biomes. We ask the following questions: (1) What is the overall trend in fire emissions at the global scale? (2) How do these trends vary across subregions defined by biome and geographical region? (3) What are the anthropogenic and geophysical drivers associated with fire emission trends in each subregion? (4) What are the most important drivers of fire emissions for each subregion? We first identified the overall trend of

fire emissions at the global scale, which we decomposed into smaller subregions. We further observed areas that contribute the most to fire emission trends using $2^\circ \times 2^\circ$ pixels. Next, we selected fire emission drivers for each of the subregions from the 18 potential drivers that cover aspects of climate, fuel, and anthropogenic interventions with a causal model using multispatial cross convergent mapping (multispatial CCM). Lastly, we ranked the importance of the selected drivers by their impact on fire emissions with machine learning models. We derived an understanding of the spatial-temporal dynamics of fire emissions and their underlying drivers across regions and biomes.

# 2 Data and methods

## 2.1 Overview

We used the total carbon emission data from the fourth version of Global Fire Emissions Database. (GFED4s). Therefore, the term "fire emission" in this paper specifically refers to fire-derived total carbon emissions, and we used this term along this paper for simplicity. We collected 18 candidate variables as potential drivers for driver selection. Our analysis includes three steps: 1) Identifying fire emission trends for the time period 2001-2019 at the global, subregional and pixel scales. We defined subregions as biomes within geographical regions; and set the resolution of the pixel at $2^\circ \times 2^\circ$. 2) Selecting drivers for fire emission trends from the 18 potential drivers with the causal model. The pixels that showed significant fire emission trends were used for driver selection for each subregion. The drivers we identified for fire emission trends were thereby specific to regions and biomes. 3) Using machine learning models to rank the driver importance by estimating their impacts on fire emissions at the subregional level.

## 2.2 Fire emission data

GFED4s is an open source database that provides a variety of fire-related indicators.[1,18,23]. Fire emission data in GFED4s were based on burnt area derived from satellite image, and was boosted by algorithms to account for small fires[1,23]. GFED4s is the most widely-used dataset in existing modelling studies for burnt area and fire emissions[12,13,24–26]. We used total carbon emissions from fire in GFED4s. Total carbon emission is separated by carbon species, namely $CO_2$, CO and $CH_4$ in GFED4s. The fractions of these species are calculated by converting carbon to dry matter and then multiplying their respective emission factors estimated for each gaseous species and fire type. Estimation for these carbon species $CO_2$, CO and $CH_4$ are subject to higher uncertainty due to the emission factors and were not included in our study[1,13,27].

We set our time period to be from 2001 to 2019. Data before 2001 were prolonged with additional active fire observations instead of MODIS 500m data[1]. As different satellite products can lead to diverging systemic bias, we did not include data before 2001[11,13,18,26]. The original spatial resolution of the fire emission data is $0.25^\circ \times 0.25^\circ$. We aggregated the spatial resolution to $2^\circ \times 2^\circ$ by summing up the emissions. It was not feasible to use the original resolution as our analysis using the causal model is highly computational intensive. At the original spatial resolution fire emissions occurred at least once in 158950 pixels. After aggregation, the number was reduced to 4117, about 38 times less than the original data. In addition, at the original resolution a large portion of the pixels were dominated by zero-valued data points (Supplementary Figure S1). The average percentage of zero-valued data points for all pixels observed fire emissions was 46.4% before aggregation, which was reduced to 17.4% after aggregating to $2^\circ \times 2^\circ$. Having too many zeros in the time series can be problematic for trend detection and causal analysis in terms of computational stability and model convergence. Instead of arbitrarily excluding low frequency pixels as adopted by some of the previous studies[11], we adopted spatial aggregation as it allowed for more complete spatial coverage.

## 2.3 Potential drivers for fire emission

There are four primary controls on fire activities: ignitions, fuel, meteorological conditions and anthropogenic interventions[1,9,18,25]. These primary controls are not independent but have complex interactions. Ignitions can start fires and affect the size of the subsequent burnt area and fire emissions. We did not include direct measures of ignition, but included indirect climatic and anthropogenic factors that could affect it[21,28]. After ignition, connectivity of fuel on the landscape can affect the spread of fire and consequently burnt area and fire emissions. Also, fuel availability directly determines the available carbon pool for combustion and thus carbon emissions[1,25,28]. Vegetation cover is often used as a proxy for fuel connectivity and fuel availability[11,12], and we included two vegetation indices *EVI* and *NDVI* to represent fuel (Table 1). Meteorological conditions affect fire emissions through spread of fire and the combustion completeness, namely how much of the available fuel is transformed into gaseous emissions[1,9]. Drought and temperature are the most often cited meteorological conditions that drive fire activities. Warmer and drier climates increase the likelihood of fires in many regions[16,29,30]. Both humidity and soil moisture are important factors that can affect fire emissions, because both aboveground living and dead vegetation biomass and soil organic matter are the sources of fire emissions[1,31,32]. Another meteorological variable, wind, also controls fire emissions because it can influence the spread of fires directly and the combustion completeness through the moisture of the fuel[33]. Therefore, we included several variables relating to water, humidity, heat and wind (Table 1). Anthropogenic intervention can take different forms and lead to varying consequences on fire emission. For example, land-use change and conversion to agricultural land can be the reason for decreasing fires by reducing fuel loads[11,12], whereas deforestation and logging can increase fire risks by opening the landscapes and allowing human initiated fires to access previously dense and closed wet-forests ecosystems[1,21]. The intensity of anthropogenic intervention is approximated by factors such as population density, economic development indicators, and percentage of crop land[11,21,34]. We included two anthropogenic factors to represent anthropogenic intervention, namely population density and percentage of agricultural land (Table 1).

We used 18 variables as potential drivers for fire emission (Table 1). We broadly categorized the 18 variables into 6 categories, namely (1) anthropogenic, (2) fuel, (3) heat, (4) humidity, (5) water, and (6) wind (Table 1). We retained some correlated variables that describe the same environmental aspect, for instance *EVI* and *NDVI*, both measuring the vegetation cover and can complement each other in global vegetation studies[35]. We did not eliminate them for following reasons: 1) to avoid making assumptions on which is the better indicator; 2) different measurements could add additional information in prediction; and 3) the subsequent machine learning models will not be undermined by correlated features[36]. Without making assumptions, the 18 variables were then selected and reduced by the causal model.

Details of data processing to obtain the 18 variables can be found in the Supplementary Material. Data for all 18 variables were spatially aggregated to $2° \times 2°$, if needed, in order to maintain a consistent spatial resolution with aggregated fire emission data. And all data were compiled on an annual basis consistent with fire emissions.

## 2.4 Statistical analysis

### *2.4.1 Robust regression for trend detection*

We applied robust linear regression to yield reliable estimations of the temporal trend of fire emissions[37]. Robust regression models operate by reducing the effect of points with high deviation from the mean, the regression results are less leveraged by the extremes. In our case, this means years with extreme fire emissions were given less weight in detecting a trend. This fits the purpose of this study as we were focused on the long-term consistent trends rather than the variations or extreme events. We explored different robust regression techniques, but these produced very similar results for temporal trends (Supplementary

**Table 1.** Potential drivers for fire emissions. The 18 drivers are divided into 6 categories. The relation to its category indicates the variable is positively (+) or negatively (−) related to the environmental condition of its category.

| Variable | Meaning | Category | Relation to its category |
|---|---|---|---|
| def | climatic water deficit (mm) | water | − |
| pdsi | Palmer Drought Severity Index | water | − |
| pr | precipitation (mm) | water | + |
| ro | runoff (mm) | water | + |
| swe | snow water equivalent (mm) | water | + |
| soil | soil moisture (mm) | water | + |
| aet | actual evapotranspiration (mm) | humidity | + |
| pet | reference evapotranspiration (mm) | humidity | − |
| vap | vapor pressure (kPa) | humidity | + |
| vpd | vapor pressure deficit(kPa) | humidity | − |
| srad | downward surface shortwave radiation ($W/m^2$) | heat | + |
| tmmn | min temperature (°C) | heat | + |
| tmmx | max temperature (°C) | heat | + |
| vs | wind speed at 10m (m/s) | wind | + |
| EVI | Enhanced Vegetation Index | fuel | + |
| NDVI | Normalized Difference Vegetation Index | fuel | + |
| population | population density | anthropogenic | + |
| agriculture | percentage of agricultural land | anthropogenic | + |

Figure S2). All results in the main manuscript are from the Huber regression, because it is a widely-used robust linear regression technique with applications in many fields[38,39].

The slope of the robust linear regression was used for determining trends. As the regression was for annual fire emission against year, the direction and magnitude of the slope coefficient represent the estimated changes in fire emission per year. Based on the direction of their slope, we classified the trend of fire emission into decreasing (negative slope) and increasing (positive slope) trends. The statistical significance of the trend, i.e., the slope of the regression model, was tested by the Wald Test, and we set the significance level at 0.05, meaning a significant trend requires $p \leq 0.05$.

### *2.4.2 Identifying drivers with a causal model*

The causal relationship between fire emission and all potential drivers was identified by a recently developed technique called multispatial convergent cross mapping (multispatial CCM)[40]. This technique is based on CCM, a method designed to detect causality in nonlinear systems from time series data. Original CCM method requires relatively long-term observations, which limited its application; whereas multispatial CCM resolved this challenge by leveraging spatial replications. Given time series data from multiple sites within the system, multispatial CCM can effectively detect casual relations between variables with much shorter time series compared to original CCM[40,41].

We defined subregions by biomes within geographical regions as the unit to apply multispatial CCM. Drivers for fire emission trends were identified for each subregion from the pixels that showed significant trend within it by multispatial CCM. The reasons for using subregions for multispatial CCM were: 1) to satisfy the assumption of multispatial CCM that the set of different sites as spatial replicates should come from a relatively homogeneous system[40]; and 2) to make the computational time feasible for our analysis. The geophysical division followed the basis regions defined in the GFED database. We reduced

the 14 basis regions to 10 for simplicity. The map of global biomes was provided by World Wildlife Fund[42], which classified global ecosystems into 14 different biomes based on climate and vegetation. We aggregated the original 14 biomes to 8 following a previous study[12]. Details of the geographical regions can be found in Supplementary Figure S3; aggregation of biomes in Supplementary Table S2 and Figure S4.

We obtained 44 subregions that contained at least one pixel that showed significant fire emission trend. We applied the causal model multispatial CCM to each subregion. Out of the total 44 subregions, the causal model successfully selected drivers for 31 subregions. The rest, such as tropical dry forests in Africa, contained too few pixels with significant fire emission trends (all $\leq 3$, Supplementary Figure S5) to produce a valid result. Our results on fire emission drivers represented 31 subregion, which accounted for around 93% of the global fire emissions from 2001 to 2019.

### *2.4.3 Quantifying driver importance*

After identifying the drivers for fire emission trends in each subregion, we quantified their relative importance with two steps: 1) building a machine learning model to simulate fire emissions for each subregion with the selected drivers as model input; and 2) interpreting the model to decouple contributions of drivers, which is the driver importance. Note that neither the machine learning model or its interpretation can answer the question of cause and effect, we need to use the causal model in the previous step to select drivers that caused changes in fire emissions as input for the machine learning model.

We used a machine learning model, Gradient Boosting Decision Tree (GBDT), to simulate fire emissions for each subregion with drivers selected by the causal model as GBDT model input. GBDT model is a nonlinear tree-based ensemble method that is efficient and robust and usually outperforms other machine learning methods in practice[36,39]. It includes a large family of tree-based models that are widely used in existing research in many fields including predicting fires. Previous studies found that GBDT models are capable of simulating fire emissions and performed well at the regional scale[15]. Optimization for GBDT models were performed with randomized search and evaluated by nested cross-validation (CV). We used variance explained as the score for GBDT model performance. It measures the ratio of the variance the model is able to explain against the total variance of the observations. Mean and variance of the performance scores were given by nested CV to avoid bias and overestimate of model performance[43]. High GBDT model performance scores with small variance means the model can consistently predict fire emissions accurately, which requires 1) that the causal model effectively selected the drivers that contained useful information to predict and explain fire emissions; and 2) the GBDT models captured some of the underlying mechanisms for fire emissions and thus can be used for interpretation[39].

While GBDT models are good at prediction, interpretation of the model is less straightforward compared to simple models such as linear regression due to its complex model structure. In order to gain insights on how changes in one input affect its output, additional model interpretation methods are required[44]. As such, we applied SHapley Additive exPlanations (SHAP) explainer to interpret the optimized GBDT models. SHAP explainer has been successfully applied in many fields of studies to identify the most relevant factors, for instance disease mortality risk factors based on ensemble tree-based models[44,45]. SHAP explainers use game theory to fairly distribute the “payout" (impact on output) among the inputs for the model being interpreted[46,47]. In our case, it can distribute the the impact of each driver on fire emissions based on the GBDT model and thus the importance of individual drivers. The impact of each driver on fire emissions were estimated over the range of driver values, which is the direct output of SHAP explainer (SHAP values). Positive values of impact indicate a positive impact of the driver on fire emissions, and vice versa. We summarized the impact of each driver by 1) the direction of its impact on fire emissions, which is calculated by the sign of the mean SHAP values and can be positive or negative;

and 2) its driver importance, which is calculated by the mean absolute SHAP values of the driver[15,46,47].

Statistical analysis was done in R[48] and Python[49]. Robust regression was implemented by `MASS`[50] package, and the significance test by `sfsmisc`[51]. Multispatial CCM was implemented by `multispatialCCM`[52]. GBDT was implemented by `LightGBM`[36], and its parameterization by `scikit-learn`[53]. SHAP explainer was implemented by `shap` using the `TreeExplainer`[45].

# 3 Results

## 3.1 Fire emission trend

At the global scale, fire emission did not show a significant trend ($p > 0.05$ for the regression slope, *Global*, Figure 1). This suggested the annual global fire emission did not change significantly between 2001 and 2019. The non-significant fire emission trend at the global scale was caused by the counteractive fire emission trends in tropical savannas and grasslands and boreal forests. Globally tropical savannas and grasslands showed a significant decreasing trend of fire emissions, while globally boreal forests showed a significant increasing trend (solid green and light green lines in *Global*, Figure 1). Tropical savannas and grasslands were the largest source of fire emissions, contributing 59% to global emissions, and boreal forests were the third largest source contributing 8% to global emissions (Supplementary Figure S6). Although emissions from tropical savannas and grasslands were a magnitude higher than from boreal forests, their increasing or decreasing rates of fire emissions were comparable. The former decreased at a rate of $-9.7 \pm 1.4 \times 10^{12} gC/year$ and the later increased at a rate of $7.4 \pm 2.2 \times 10^{12} gC/year$ (rates of fire emission trends estimated by the slope coefficients of Huber regression). Tropical savannas and grasslands and boreal forests were the only biomes that showed significant fire emission trends at the global scale, the other six biomes showed non-significant trends (Supplementary Figure S7).

More than 75% of fire emissions from tropical savannas and grasslands occurred in Africa, which also showed a significant decreasing trend (solid green lines in *Global* and *Africa*, Figure 1)). Emissions from tropical savannas and grasslands in other regions, e.g., South America, Australia, Equatorial Asia, and Middle East, had marginal contributions to global emissions and all showed a non-significant trend (Figure 1). Tropical wet forests were the second largest source for global fire emissions, but showed no significant trend in fire emissions globally and across regions (Figure 1). Boreal forests showed a significant increasing trend at the global scale, yet within subregions fire emissions for boreal forests in Boreal Asia, North America, and Europe were not significant (Figure 1). For the rest of the 5 biomes with lower fire emissions globally and in separate subregions, fire emission trends were mostly non-significant, except temperate forests and woodlands in Central Asia and North America, the former showed a significant decreasing trend and the latter increasing (Figure 1). Most regions exhibited strong interannual fluctuations in carbon emissions but non-significant trends, for example the significant spikes in annual fire emissions for tropical wet forests in South America, Central Asia and Equatorial Asia (Figure 1).

At the $2^\circ \times 2^\circ$ pixel scale, we detected 697 pixels that showed significant fire emission trends, which is 17% of all the pixels that experienced fire emissions. Of the pixels that showed significant fire emission trends, 361 pixels showed an increasing trend and 336 decreasing (Figure 2). Pixels that showed significant trend covered a total area of around $2.0 \times 10^7$ $km^2$, $0.9 \times 10^7$ $km^2$ for increasing and $1.1 \times 10^7$ $km^2$ for decreasing. A large number of pixels that showed an increasing trend were distributed in Central Asia, especially around India and north China. The dominant biomes for these pixels in Central Asia were temperate forests and woodlands, and shrublands. We also observed a large cluster of pixels ($6.0 \times 10^5$ $km^2$) with increasing fire emissions in boreal forests in the center of Boreal Asia, which is around the east of Siberia (Figure 2 (a)). This cluster alone contributed to increasing fire emissions at a rate of $2.8 \pm 0.7 \times 10^{12} gC/year$, slightly more than $1/3$ of the overall increasing rate of fire emissions from

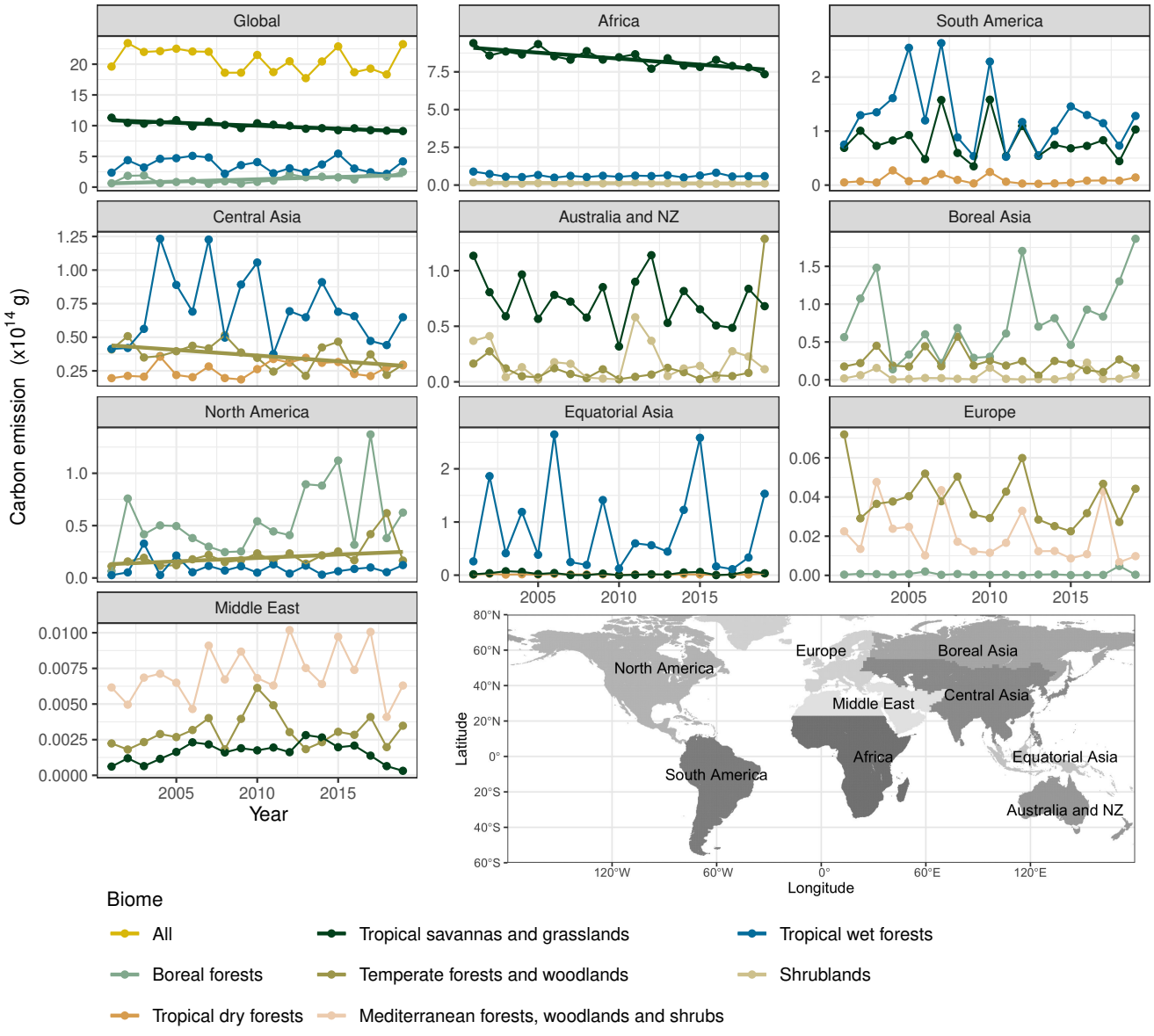


**Figure 1.** Global and regional fire-derived carbon emissions from different biomes from 2001 to 2019 and their temporal trends. For each region, only 3 biomes with the highest emissions were included, except that subplot *Global* also included total emissions from all biomes. Regions were arranged in a decreasing order in terms of their fire emissions from left to right and top to bottom, and their y-axis were in different scales. Overall temporal trends determined by Huber regression were denoted by straight lines and only significant trends ($p \leq 0.05$ for slope of the regression) were displayed in the figure. A map of the regions was included in the lower right corner for reference.

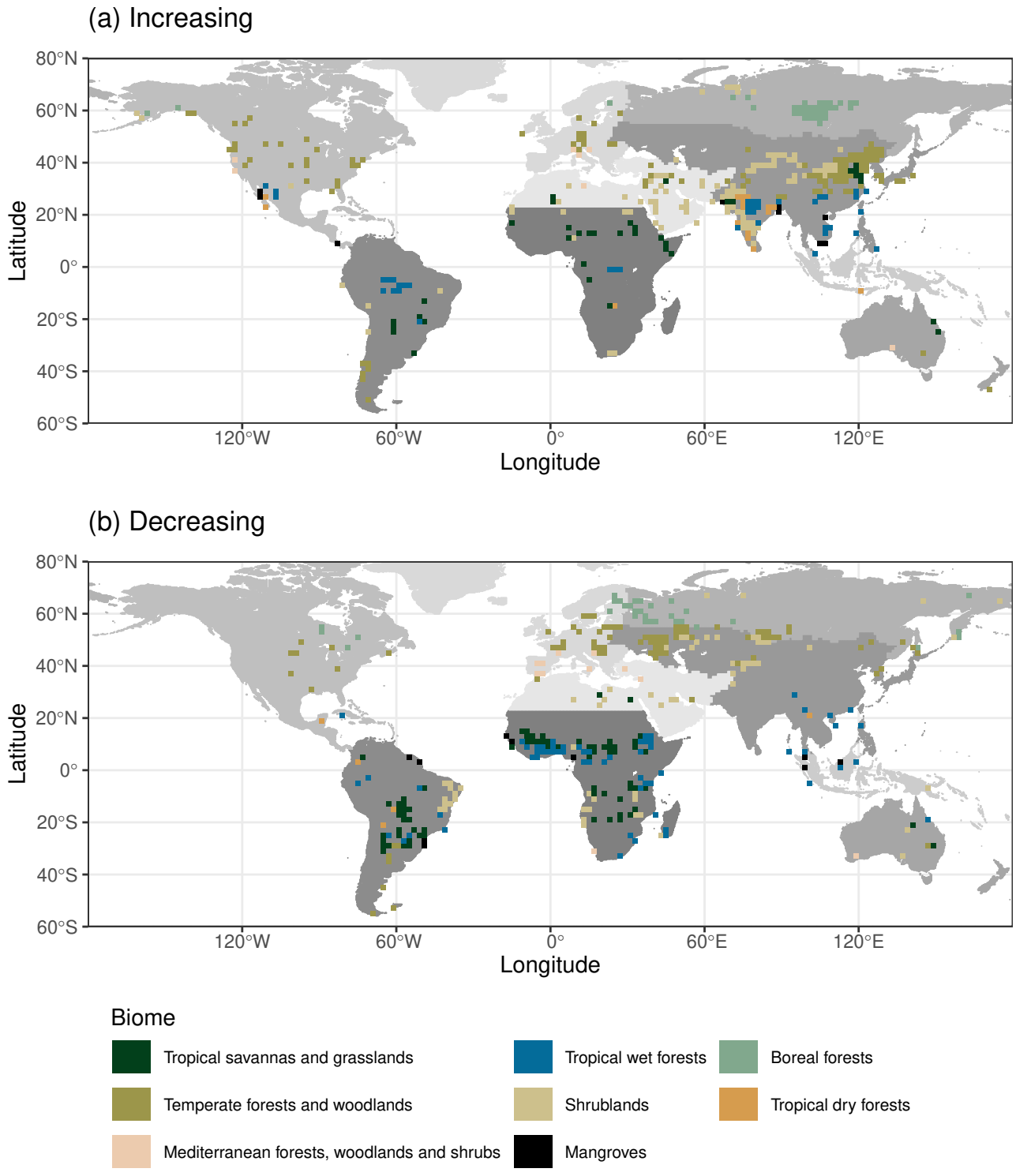


**Figure 2.** Pixels that showed significant (a) increasing fire emission trends and (b) decreasing fire emission trends from 2001 to 2019 ($p < 0.05$ for regression slopes). Color of the pixels indicated different biomes, and the different shades of grey in the base map indicated the geographical regions. Fire emission trends were detected by Huber regression.

boreal forests globally, which is $7.4 \pm 2.2 \times 10^{12} gC/year$. The large clusters of pixels that showed significant increasing trends in Boreal Asia and Central Asia accounted for 56% of that globally in terms of the number of the pixels. Three smaller clusters of pixels that showed significant increasing fire emission trend were observed in tropical wet forests in South America and Central Asia, and temperate forests and woodlands in Europe. The rest of pixels with increasing fire emissions in other regions did not form large clusters.

For pixels that showed decreasing fire emission trends, about 47% were located in Africa and South America. For Africa, most of the pixels that showed decreasing trends were in west and central Africa and the south of the equator; and were dominated by tropical savannas and grasslands, and tropical wet forests. Pixels that showed decreasing fire emission trends in South America were dominated by tropical savannas and grasslands. A considerable amount of pixels with decreasing fire emissions were observed in temperate forests and woodlands in Europe and west of Central Asia, and boreal forests closer to Europe. In total they accounted for 40% of the number of pixels that showed significant decreasing trends. The remaining pixels with decreasing fire emissions (13%) were scattered in other regions. Australia and Equatorial Asia, although hotspots for fire events, only contained a few pixels that showed significant long-term fire emission trends.

### 3.2 Drivers selected by causal model and their impact on fire emissions

For all 31 subregions, the causal model reduced 18 potential drivers to between 1 and 14 drivers for fire emission trends specific to the subregion. The number of drivers selected for each subregion suggested fire emission trends in some subregions had been driven by a more complex combination of drivers, while others by a small set of drivers. With the selected drivers as GBDT model input, 17 subregions achieved GBDT model performance scores higher than 50% variance explained (Supplementary Figure S8). Subregions that achieved good model performance did not necessarily have large numbers of drivers. For instance, the fire emissions from temperate forests and woodlands in South America were driven solely by wind speed (*vs*), a climatic driver, and the GBDT model scored around 50% of the variance explained; while the fire emission drivers for the same biome in Central Asia were more complex and had 13 in total, ranging from climatic to vegetation, and also achieved a performance score close to 50%. This suggested that the causal model effectively selected the drivers that have causal relations with fire emissions for most subregions.

We focused on the top 10 subregions that accounted for more than 80% of the global fire emissions from 2001 to 2019 (Figure 3). From these 10 subregions, 8 subregions achieved a GBDT model performance of more than 50% variance explained. Tropical wet forests in Equatorial Asia had the lowest GBDT model performance score, and boreal forests in Boreal Asia also had lower scores compared to the other subregions. Fire emission drivers for the same biome across different geographical regions can be different. For example, tropical savannas and grasslands in Africa, South America, and Australia, had different drivers associated with fire emissions. Vegetation was identified as an important driver for the same biome in these three regions, but climatic drivers were different, and anthropogenic drivers were only identified for African savannas and grasslands. Likewise, fire emission drivers for different biomes within the same geographical regions can be different as well. For example, the two biomes in Africa, tropical savannas and grasslands and tropical wet forests, had very different climatic drivers. The former only had one associated with wind, while the latter had multiple climatic drivers associated with heat, humidity and water.

For the top 10 subregions, fuel was the most common driver, a driver for 9 subregions with the exception of tropical wet forests in Equatorial Asia (Figure 3). The importance of the two drivers relating to fuel, *EVI* and *NDVI*, varied among subregions. *EVI* and *NDVI* were the most important drivers for

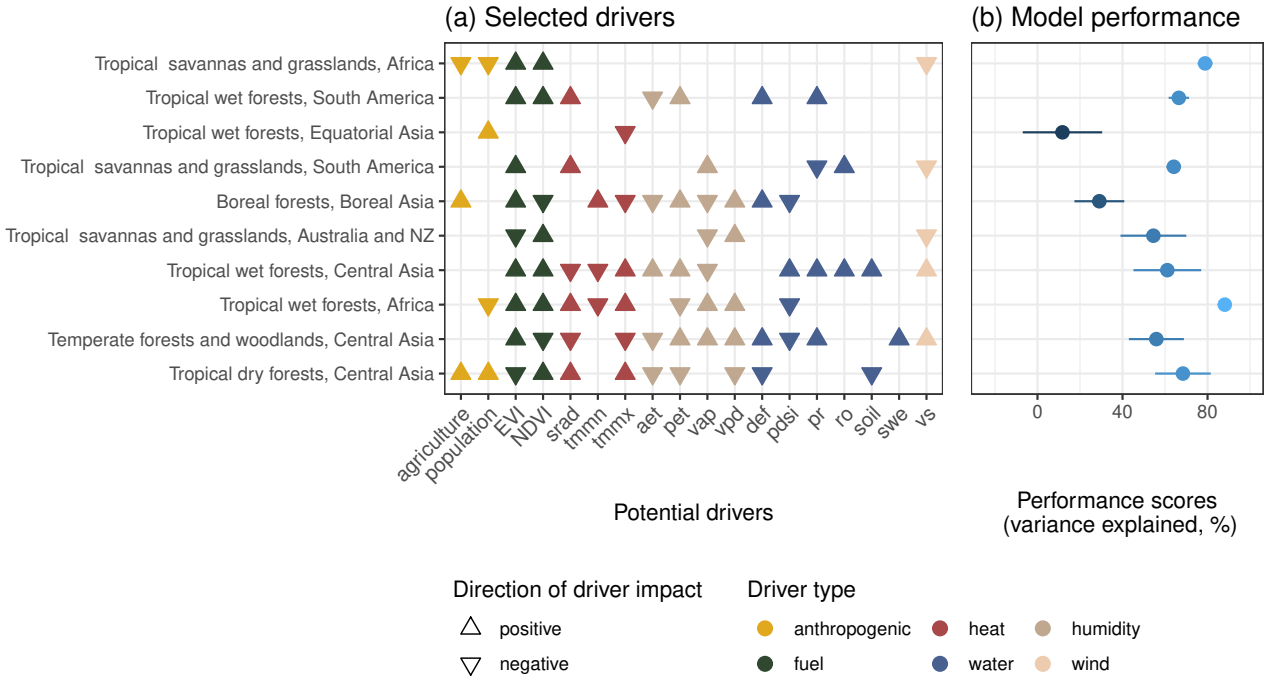


**Figure 3.** Drivers selected by causal model for subregions defined by region and biome (a) and their GBDT model performance with the selected drivers as model input (b). Only top 10 subregions with the highest fire emissions were displayed, and were arranged in a decreasing order by their fire emissions from top to bottom. Triangles denoted the selected drivers for each subregion. The orientation of the triangles represented the overall impact of the driver on fire emissions for the particular subregion, which were calculated by the sign of the average estimated impact on fire emissions (SHAP values). An upwards triangle means a positive impact, and downwards negative. All 18 potential drivers were divided into 6 categories, which were indicated by the color of the triangles. Model performance scores were measured by variance explained by the GBDT model with the selected drivers. The dots and horizontal lines represented the mean and standard deviation of performance scores from the 5-fold cross validation for each subregions. A high performance score means the model predicts fire emissions accurately. A negative score means the model predicts less accurately than having the mean fire emission value as a constant estimate.

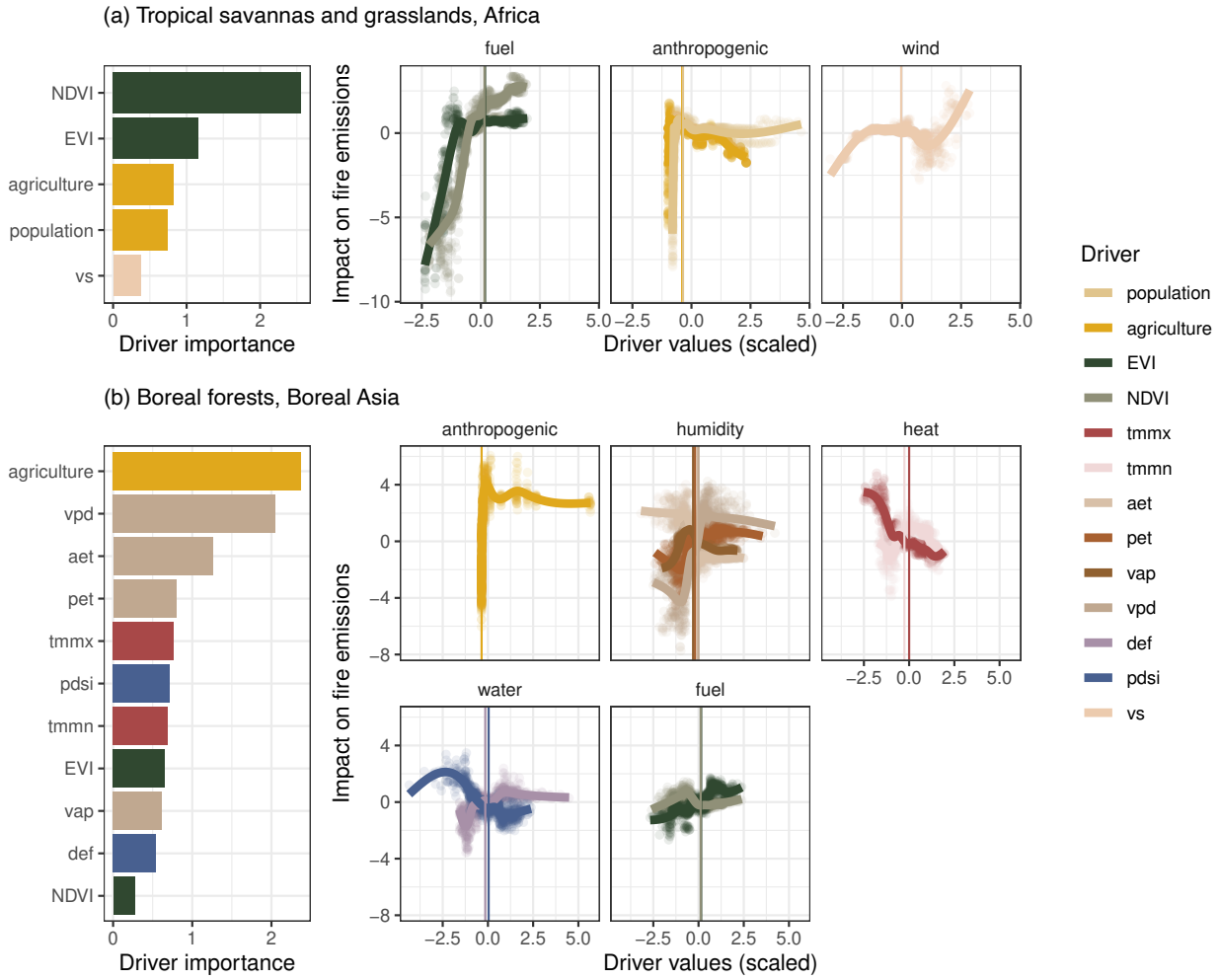


**Figure 4.** Driver importance and their estimated impact on fire emissions for tropical savannas and grasslands in Africa (a) and boreal forests in Boreal Asia (b). Drivers were arranged by their importance (overall impact on fire emissions) from top to bottom in a decreasing order in the bar plots. Importance of drivers were calculated by the mean absolute values of their estimated impact on fire emissions across the range of driver values. The estimated impact on fire emission against driver values were plotted by driver categories on the right hand side. Driver values were scaled by $(X - \bar{X})/\sigma(X)$, where $X$ represents a driver, and $\bar{X}$ and $\sigma(X)$ its mean and standard deviation. Estimated impact on fire emissions were SHAP values calculated by SHAP explainer. Negative values of estimated impact indicated negative impact on fire emissions, and vice versa. The vertical lines in each subplot indicated the medians of the respective drivers. Note that the medians of some drivers were very close and thus the vertical lines almost overlapped.

tropical savannas and grasslands in Africa, the biggest contributor to global fire emissions; while they were the least important drivers for boreal forests in Boreal Asia (Figure 4). Typically, low fuel loads decreased fire emissions, while higher fuel loads increased fire emissions. For example, in African savannas and grasslands, lower *EVI* and *NDVI* values had a negative impact on fire emissions but an increase in *EVI* and *NDVI* resulted in a positive impact on fire emissions, and this shift from negative impact to positive occurred around the medians (Figure 4 (a)). Fuel had an overall positive impact on fire emissions for

most subregions, highlighting the role of vegetation as having a positive relationship with fire emissions (Figure 3). The two drivers representing fuel, *EVI* and *NDVI*, could have opposite impacts on fire emissions for some subregions, but the overall impacts of fuel were nonetheless positive. For example, *EVI* and *NDVI* had opposite impacts on fire emissions for Asian boreal forests, *EVI* positive and *NDVI* negative (Figure 3). However, the positive impact of *EVI* outweighed the negative impact of *NDVI* because the former ranked higher in driver importance than the latter (Figure 4), resulting in an overall positive impact of fuel on fire emissions. This is also true for tropical savannas and grasslands in Australia, and temperate forests and woodlands and tropical dry forests in Central Asia. Regardless of the opposite directions of the two drivers associated with fuel, the overall impacts of fuel were still positive (Figure 3 and Supplementary Figure S9).

At least one driver belonging to the four climatic categories, namely heat, water, humidity and wind, were observed for all subregions (Supplementary Figure S8). Among the four climatic categories, heat, humidity and water were more commonly observed as drivers, while wind was identified as a driver for fewer subregions (Figure 3 and Supplementary Figure S8). For the top 10 subregions, 8 subregions had drivers relating to heat and humidity, 7 for water; while wind was only observed in 5 of them and usually ranked lower in terms of driver importance. Some subregions had more climatic drivers than others, suggesting that climate has more control over fire emissions for those regions. For example, African tropical savannas and grasslands had only one climatic driver *vs* which was also the least important driver. On the other hand, boreal forests in Boreal Asia had 8 climatic drivers covering 3 categories, namely humidity, heat, and water; which also ranked high in driver importance (Figure 4). Therefore, in our analysis climate played a much more important role for driving fire emissions in boreal forests in Boreal Asia than tropical savannas and grasslands in Africa.

Anthropogenic factors were identified to be fire emission drivers for 16 subregions in total, and 5 for the top 10 subregions with highest emissions. For the two anthropogenic factors, *population* was identified as a fire emission driver in more subregions than *agriculture*. This suggested anthropogenic interventions in general as reflected by *population* had a broader impact on fire emissions through different means other than agricultural activities. The impact of anthropogenic drivers on fire emissions were negative for some subregions and positive for others. For African tropical savannas and grasslands, both *population* and *agriculture* had a negative impact on fire emissions (Figure 3). When *population* and *agriculture* were low (below the medians), this negative impact was greater since their impacts were far below 0. With the increase of the two anthropogenic factors, the impact became positive around the median values, and then decreased and became negative again (Figure 4). For boreal forests in Boreal Asia, the anthropogenic driver *agriculture* had an overall positive impact on fire emissions (Figure 3), and the pattern of its impact were quite different from that in African savannas and grasslands (Figure 4). After around the median of *agriculture*, its impact stayed positive and the values were far higher than that in African savannas and grasslands.

The impact of drivers under the same driver category for a subregion did not necessarily follow the same pattern. For example, drivers under the categories fuel and anthropogenic for tropical savannas and grasslands in Africa aligned for their impacts on fire emissions (Figure 4 (a)). For boreal forests in Boreal Asia, drivers under the category humidity were consistent for their overall impact on fire emissions, which was a negative impact - high humidity led to lower fire emissions. Note that *vpd* and *pet* were inversely related to humidity (Table 1), therefore, the positive impacts of these two drivers on fire emissions represent a negative impact of humidity (Figure 3). However, the patterns of impact for the 4 drivers relating to humidity were all highly nonlinear and displayed different patterns (Figure 4 (b)). For Asian boreal forests, heat and water had different directions of impact, and their patterns of impact differed as well, highlighting the complexity of interacting climatic effects on fire emissions.

# 4 Discussion

Our study makes an important contribution towards understanding trends and drivers of fire-derived total carbon emissions from 2001 to 2019 at a global scale and at a scale specific to biome and region. Long-term trends associated with global fire emissions have been relatively stable, a result of two counteractive fire emission trends which we identify across the region of African savannas and grasslands and large cluster in Asian boreal forests. Trends across subregions vary widely, with some subregions showing consistent increasing or decreasing trends in fire emissions, while others showing large fluctuations in annual fire emissions. Vegetation and climatic variables were major drivers for fire emission trends across subregions, although anthropogenic variables were significant in driving the trends of fire emissions from two of the largest contributing sources, African savannas and grasslands and Asian boreal forests. These geophysical and anthropogenic drivers ranked differently in driver importance across our subregions and underlie the observed decreasing or increasing fire emission trends across time.

At a global scale, vegetation was the most common important driver in fire emission trends across subregions. This corroborates with global studies that identified fuel as the most important driver for changes in global fires[12]. Vegetation generally had a positive impact on fire emissions in our results, which aligns with our understanding that fire is constrained by the availability of vegetation and other forms of fuel[54,55]. From our results, climate drove fire emission trends in most subregions, aligning with the literature that climatic conditions, especially rainfall, temperature and humidity, are important controls for global fires[6,30]. Anthropogenic interventions were identified as drivers of fire emission trends for a smaller number of subregions compared to vegetation and climate. However, the influence of anthropogenic intervention can be stronger in some subregions such as African savannas and grasslands. Although previous studies in Africa savannas found climate conditions such as increasing moisture availability drives increases in fire using spatial data[56], our analysis showed climatic factors were less important for long-term trends. This concurs with previous studies that found land-use conversion to be a dominant driver for decreased burnt area in African savannas over the past two decades[10].

Although anthropogenic intervention had a negative impact in Africa, it can have a positive impact on fire emissions in other subregions such as tropical wet forests in Equatorial Asia and tropical dry forests in Central Asia. This corroborates with studies that showed using fires to clear and prepare the land for crops can be the reason for the large areas that showed increasing fire emissions in Central Asia[57]; and that deforestation have led to large-scale fires in Equatorial Asia[29]. For forested biomes, anthropogenic intervention through intensive deforestation can lead to a higher risk of fires in the intact forests and consequently increasing fire emissions, which is further amplified by climate change[58,59]. A small cluster of pixels in our study showed significant increasing fire emission trends in tropical wet forests in South America, likely a result of deforestation in this area[60]. This risk of increasing fire emissions from forests associated with anthropogenic interventions and climate change has strong implications for future fire emission trends because forests have much higher fuel loads compared to savannas and grasslands, resulting in higher fire emissions per unit burnt area[3,61].

Of the forested biomes that have higher risks of increasing fire emissions, we further identified the main contributor of increasing fire emissions, which is boreal forests, especially a dense cluster of pixels in Asia. Our finding corroborates with previous studies that identified forests and boreal forests as the reason for "stable" fire emissions in spite of the decreased burnt area globally[3,10,61]. The cluster that showed significant increasing fire emissions in Asian boreal forests belongs to the ecoregion *East Siberian Taiga*[42] (Supplementary Figure S10). The Siberian tundra has permafrost soils (frozen for more than 2 years) and agricultural activity tends to follow rivers and road networks in this region (Supplementary Figure S10). We find climatic conditions, especially lower humidity, as a driver of increased fire emissions in this

region, similar to regional studies that have connected significantly increased burned area to warmer and drier climates[62–64]. In addition to climate, we also found agricultural activity as an important driver for increasing fire emissions in this region, and the linkage between agricultural land-use and fire emissions has received far less research attention[63]. Moreover, increasing carbon emissions from burning of boreal forests are likely to result in a net long-term carbon source since boreal forests take decades to centuries to regenerate[3,65]. Fires over boreal forests can also lead to a positive feedback of further carbon emissions by accelerating thawing of underlying permafrost soils[64,66]. Our results highlight the importance of boreal forests in Asia in contributing towards global fire emissions, and the need to target *East Siberian Taiga* for detailed monitoring and management interventions to mitigate further fire emissions.

Our study on fire emissions focused on the long-term consistent trend in fire emissions which is limited by the time period for our observation (2001-2019), and did not investigate the interannual variations or extreme fire events. While several extreme fire events have taken place in some subregions such as the wet forests in South America[67] and Equatorial Asia[16], and the savannas and woodlands in Australia[14], our study did not capture any significant long-term increasing trends in fire emissions for these subregions, and the number of pixels with significant fire emission trends were not high. However, we should be cautious to extend our results to predict future changes in fire emissions, and it is possible that fire emissions in these regions experience significant increase under future climate change and/or anthropogenic intervention. We observed subregions where fire emission trends were heavily driven by climate, such as tropical wet forests and tropical savannas in South America, and Asian boreal forests. This aligned with existing studies that highlighted the risk of climate change on fires in these regions[58,64]. Since fire emission trends in these subregions have been heavily controlled by climate, it is likely that their future fire emissions are more vulnerable to climate change, and should receive more attention.

We are aware that our results are constrained by data availability at the temporal and spatial scale of our study. Findings from this study are based on GFED4s emissions data, which are known to underestimate burnt area because of the inability to account for small fires. This constitutes a considerable underestimation for Africa where a large proportion of fires are small fires[27,68] amongst other regions[26]. Therefore, our findings may be underestimating a significant proportion of emissions from small sized fires[13]. In our analysis of drivers, our selection was limited by globally available datasets. For example, pasture landcover was found to be a good indicator for anthropogenic intervention in relation to burnt area in global studies[10,12], but harmonized and reliable global annual time series pasture landcover data from 2001 to 2019 is lacking[69]. Missing important drivers could be a reason for poor GBDT model performance for some subregions. For example, the relatively low GBDT model performance scores for tropical wet forests in Equatorial Asia can be a result of lacking a representation of soil carbon as a driver (fuel source) and its contribution to fire emission estimates. Large areas of tropical wet forests in Equatorial Asia are peatlands, where burning of organic carbon in soil contribute greatly to the emissions[1,32,70].

Changing climate and anthropogenic interventions are major concerns for increases in fires across the world, which can have a positive feedback on climate through fire-vegetation interactions and gaseous emissions. Global databases such as GFED provide spatially explicit time series data for fire emissions, allowing us to make valuable observations on temporal trends of fire emissions at different spatial scales. Our analysis at a global scale provides a first look at fire emission trends estimated at the global, subregional and pixel ($2^\circ \times 2^\circ$) level, and the drivers for increasing or decreasing trends specific to regions and biomes. Although we observed fire emissions to be relatively stable over the years 2001-2019, special attention should be paid to boreal forests, which showed significant increasing fire emissions. Based on the global distribution of fire emission trends, we can identify important regions to target for future monitoring and management, such as the ecoregion *East Siberian Taiga*. Our results on the drivers for fire emission trends highlight which geophysical and anthropogenic factors need to be targeted to mitigate global

fire emissions. For example, climatic factors, primarily humidity, and agricultural activity have been shown to drive increasing fire emissions in Asian boreal forests, predisposing this subregion to high risk of fires under climate warming. While mitigating climate change would require action at a global level, at a regional level, monitoring and managing agricultural activities in this subregion could have an out-sized contribution towards mitigating global fire emissions. A reduction in vegetation cover has driven decreasing trends in fire emissions in tropical savannas and grasslands in Africa and South America. However, a decline in vegetation cover in forests and woodlands biomes result in previously intact forests being made more susceptible to fires, subjecting these ecosystems to future risks of shifting fire regimes and higher fire emissions[58]. This risk is further amplified by climate change due to the complex feedback between vegetation, climate and fires[54,58]. Considering that increasing fire emissions are more controlled by the burning of forests than savannas and grasslands[3,61], we highlight the importance of mitigating deforestation and fires in forests and woodlands biomes to reduce global carbon emissions from fires.

## Data availability

All data in this paper are open source data that can be accessed online. Code to process the data and perform the analysis will be made available on Github after publication.

## References


1. van der Werf, G. R. *et al.* Global fire emissions estimates during 1997–2016. *Earth Syst. Sci. Data* **9**, 697–720 (2017).
2. Friedlingstein, P. *et al.* Global carbon budget 2020. *Earth Syst. Sci. Data* **12**, 3269–3340, DOI: 10.5194/essd-12-3269-2020 (2020).
3. Zheng, B. *et al.* Increasing forest fire emissions despite the decline in global burned area. *Sci. Adv.* **7**, eabh2646, DOI: 10.1126/sciadv.abh2646 (2021).
4. Hempson, G. P., Archibald, S. & Bond, W. J. The consequences of replacing wildlife with livestock in Africa. *Sci. Reports* **7**, 1–10 (2017).
5. Pellegrini, A. F. A. *et al.* Fire frequency drives decadal changes in soil carbon and nitrogen and ecosystem productivity. *Nature* **553**, 194–198, DOI: 10.1038/nature24668 (2018).
6. Archibald, S., Lehmann, C. E., Gómez-Dans, J. L. & Bradstock, R. A. Defining pyromes and global syndromes of fire regimes. *Proc. Natl. Acad. Sci.* **110**, 6442–6447 (2013).
7. Taylor, A. H., Trouet, V., Skinner, C. N. & Stephens, S. Socioecological transitions trigger fire regime shifts and modulate fire–climate interactions in the Sierra Nevada, USA, 1600–2015 CE. *Proc. Natl. Acad. Sci.* **113**, 13684–13689, DOI: 10.1073/pnas.1609775113 (2016).
8. Lindenmayer, D. B. & Taylor, C. New spatial analyses of Australian wildfires highlight the need for new fire, resource, and conservation policies. *Proc. Natl. Acad. Sci.* **117**, 12481–12485, DOI: 10.1073/pnas.2002269117 (2020).
9. McLauchlan, K. K. *et al.* Fire as a fundamental ecological process: research advances and frontiers. *J. Ecol.* **108**, 2047–2069, DOI: 10.1111/1365-2745.13403 (2020).
10. Andela, N. *et al.* A human-driven decline in global burned area. *Science* **356**, 1356–1362, DOI: 10.1126/science.aal4108 (2017).
11. Forkel, M. *et al.* Recent global and regional trends in burned area and their compensating environmental controls. *Environ. Res. Commun.* **1**, 051005, DOI: 10.1088/2515-7620/ab25d2 (2019).
12. Kelley, D. I. *et al.* How contemporary bioclimatic and human controls change global fire regimes. *Nat. Clim. Chang.* **9**, 690–696 (2019).
13. Pan, X. *et al.* Six global biomass burning emission datasets: intercomparison and application in one global aerosol model. *Atmospheric Chem. Phys.* **20**, 969–994, DOI: 10.5194/acp-20-969-2020 (2020).
14. van der Velde, I. R. *et al.* Vast CO2 release from Australian fires in 2019–2020 constrained by satellite. *Nature* **597**, 366–369 (2021).
15. Wang, S. S.-C., Qian, Y., Leung, L. R. & Zhang, Y. Interpretable machine learning prediction of fire emissions and comparison with FireMIP process-based models. preprint, Biosphere Interactions/Atmospheric Modelling/Troposphere/Physics (physical properties and processes) (2021). DOI: 10.5194/acp-2021-634.

**16.** Field, R. D. *et al.* Indonesian fire activity and smoke pollution in 2015 show persistent nonlinear sensitivity to El Niño-induced drought. *Proc. Natl. Acad. Sci.* **113**, 9204–9209, DOI: 10.1073/pnas.1524888113 (2016).

**17.** Walker, X. J. *et al.* Cross-scale controls on carbon emissions from boreal forest megafires. *Glob. Chang. Biol.* **24**, 4251–4265 (2018).

**18.** Giglio, L., Randerson, J. T. & van der Werf, G. R. Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). *J. Geophys. Res. Biogeosciences* **118**, 317–328, DOI: 10.1002/jgrg.20042 (2013).

**19.** Giglio, L. Global estimation of burned area using MODIS active fire observations. *Atmos. Chem. Phys.* 18 (2006).

**20.** Carlson, K. M. *et al.* Effect of oil palm sustainability certification on deforestation and fire in Indonesia. *Proc. Natl. Acad. Sci.* **115**, 121–126, DOI: 10.1073/pnas.1704728114 (2018).

**21.** Sze, J. S., Jefferson & Lee, J. S. H. Evaluating the social and environmental factors behind the 2015 extreme fire event in Sumatra, Indonesia. *Environ. Res. Lett.* **14**, 015001, DOI: 10.1088/1748-9326/aaee1d (2019).

**22.** Hamilton, M., Salerno, J. & Fischer, A. P. Cognition of complexity and trade-offs in a wildfire-prone social-ecological system. *Environ. Res. Lett.* **14**, 125017, DOI: 10.1088/1748-9326/ab59c1 (2019).

**23.** Randerson, J. T., Chen, Y., van der Werf, G. R., Rogers, B. M. & Morton, D. C. Global burned area and biomass burning emissions from small fires. *J. Geophys. Res. Biogeosciences* **117**, DOI: 10.1029/2012JG002128 (2012).

**24.** Arora, V. K. & Melton, J. R. Reduction in global area burned and wildfire emissions since 1930s enhances carbon uptake by land. *Nat. Commun.* **9**, 1–10 (2018).

**25.** Abatzoglou, J. T., Williams, A. P., Boschetti, L., Zubkova, M. & Kolden, C. A. Global patterns of interannual climate–fire relationships. *Glob. Chang. Biol.* **24**, 5164–5175 (2018).

**26.** Humber, M. L., Boschetti, L., Giglio, L. & Justice, C. O. Spatial and temporal intercomparison of four global burned area products. *Int. J. Digit. Earth* **12**, 460–484 (2019).

**27.** Laris, P. On the problems and promises of savanna fire regime change. *Nat. Commun.* **12**, 4891, DOI: 10.1038/s41467-021-25141-1 (2021).

**28.** Veraverbeke, S. *et al.* Lightning as a major driver of recent large fire years in North American boreal forests. *Nat. Clim. Chang.* **7**, 529–534 (2017).

**29.** van der Werf, G. R. *et al.* Climate regulation of fire emissions and deforestation in equatorial Asia. *Proc. Natl. Acad. Sci.* **105**, 20350–20355, DOI: 10.1073/pnas.0803375105 (2008).

**30.** Balch, J. K. *et al.* Warming weakens the night-time barrier to global fire. *Nature* **602**, 442–448 (2022).

**31.** Konecny, K. *et al.* Variable carbon losses from recurrent fires in drained tropical peatlands. *Glob. Chang. Biol.* **22**, 1469–1480 (2016).

**32.** Dadap, N. C., Cobb, A. R., Hoyt, A. M., Harvey, C. F. & Konings, A. G. Satellite soil moisture observations predict burned area in Southeast Asian peatlands. *Environ. Res. Lett.* **14**, 094014, DOI: 10.1088/1748-9326/ab3891 (2019).

**33.** Touma, D., Stevenson, S., Lehner, F. & Coats, S. Human-driven greenhouse gas and aerosol emissions cause distinct regional impacts on extreme fire weather. *Nat. Commun.* **12**, 1–8 (2021).

**34.** Knorr, W., Arneth, A. & Jiang, L. Demographic controls of future global fire risk. *Nat. Clim. Chang.* **6**, 781–785 (2016).

**35.** Huete, A. *et al.* Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote. Sens. Environ.* **83**, 195–213, DOI: 10.1016/S0034-4257(02)00096-2 (2002).

**36.** Ke, G. *et al.* LightGBM: a highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems*, vol. 30 (Curran Associates, Inc., 2017).

**37.** Muhlbauer, A., Spichtinger, P. & Lohmann, U. Application and comparison of robust linear regression methods for trend estimation. *J. Appl. Meteorol. Climatol.* **48**, 1961–1970 (2009).

**38.** Huber, P. J. Robust estimation of a location parameter. In *Breakthroughs in statistics*, 492–518 (Springer, 1992).

**39.** Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning* (Springer, New York, 2009).

**40.** Clark, A. T. *et al.* Spatial convergent cross mapping to detect causal relationships from short time series. *Ecology* **96**, 1174–1181, DOI: 10.1890/14-1479.1 (2015).

**41.** Sugihara, G. *et al.* Detecting causality in complex ecosystems. *Science* **338**, 496–500, DOI: 10.1126/science.1227079 (2012).

**42.** Olson, D. M. *et al.* Terrestrial ecoregions of the world: a new map of life on earth. *BioScience* **51**, 933, DOI: 10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2 (2001).

**43.** Cawley, G. C. & Talbot, N. L. C. On over-fitting in model selection and subsequent selection bias in performance evaluation. *J. Mach. Learn. Res.* **11**, 2079–2107 (2010).

**44.** Molnar, C. *Interpretable Machine Learning* (lulu.com, 2020).

**45.** Lundberg, S. M. *et al.* From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2**, 56–67, DOI: 10.1038/s42256-019-0138-9 (2020).

**46.** Štrumbelj, E. & Kononenko, I. Explaining prediction models and individual predictions with feature contributions. *Knowl. Inf. Syst.* **41**, 647–665, DOI: 10.1007/s10115-013-0679-x (2014).

**47.** Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems*, vol. 30 (Curran Associates, Inc., 2017).

**48.** R Core Team. R: A Language and Environment for Statistical Computing (2021).

**49.** Guido, V. R. & L., D., Fred. *Python 3 Reference Manual* (CreateSpace, 2009).

**50.** Ziegel, E. R. Modern applied statistics with S. *Technometrics* **45**, 111 (2003).

**51.** Maechler, M. sfsmisc: Utilities from 'Seminar fuer Statistik' ETH Zurich (2020).

**52.** Clark, A. multispatialCCM: Multispatial Convergent Cross Mapping (2014).

**53.** Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

**54.** Hurteau, M. D., Liang, S., Westerling, A. L. & Wiedinmyer, C. Vegetation-fire feedback reduces projected area burned under climate change. *Sci. Reports* **9**, 1–6 (2019).

**55.** Marchal, J., Cumming, S. G. & McIntire, E. J. Turning down the heat: Vegetation feedbacks limit fire regime responses to global warming. *Ecosystems* **23**, 204–216 (2020).

**56.** Lehmann, C. E. *et al.* Savanna vegetation-fire-climate relationships differ among continents. *Science* **343**, 548–552 (2014).

**57.** Jethva, H. *et al.* Connecting crop productivity, residue fires, and air quality over northern India. *Sci. Reports* **9**, 1–11 (2019).

**58.** Nobre, C. A. *et al.* Land-use and climate change risks in the Amazon and the need of a novel sustainable development paradigm. *Proc. Natl. Acad. Sci.* **113**, 10759–10768 (2016).

**59.** Barni, P. E. *et al.* Logging Amazon forest increased the severity and spread of fires during the 2015–2016 El Nino. *For. Ecol. Manag.* **500**, 119652 (2021).

**60.** Libonati, R. *et al.* Twenty-first century droughts have not increasingly exacerbated fire season severity in the Brazilian Amazon. *Sci. Reports* **11**, 1–13 (2021).

**61.** van der Werf, G. R. *et al.* Interannual variability in global biomass burning emissions from 1997 to 2004. *Atmospheric Chem. Phys.* **6**, 3423–3441, DOI: 10.5194/acp-6-3423-2006 (2006).

**62.** Kim, J.-S., Kug, J.-S., Jeong, S.-J., Park, H. & Schaepman-Strub, G. Extensive fires in southeastern siberian permafrost linked to preceding arctic oscillation. *Sci. Adv.* **6**, eaax3308, DOI: 10.1126/sciadv.aax3308 (2020).

**63.** Kharuk, V. I. *et al.* Wildfires in the Siberian taiga. *Ambio* **50**, 1953–1974 (2021).

**64.** Talucci, A. C., Loranty, M. M. & Alexander, H. D. Siberian taiga and tundra fire regimes from 2001–2020. *Environ. Res. Lett.* **17**, 025001, DOI: 10.1088/1748-9326/ac3f07 (2022).

**65.** Chu, T., Guo, X. & Takeda, K. Remote sensing approach to detect post-fire vegetation regrowth in Siberian boreal larch forest. *Ecol. Indic.* **62**, 32–46 (2016).

**66.** Lin, S., Liu, Y. & Huang, X. Climate-induced Arctic-boreal peatland fire and carbon loss in the 21st century. *Sci. Total. Environ.* **796**, 148924, DOI: 10.1016/j.scitotenv.2021.148924 (2021).

**67.** Kumar, S. *et al.* Changes in land use enhance the sensitivity of tropical ecosystems to fire-climate extremes. *Sci. Reports* **12**, 1–11 (2022).

**68.** Ramo, R. *et al.* African burned area and fire carbon emissions are strongly impacted by small fires undetected by coarse resolution satellite data. *Proc. Natl. Acad. Sci.* **118**, e2011160118, DOI: 10.1073/pnas.2011160118 (2021).

**69.** Klein Goldewijk, K., Beusen, A., Van Drecht, G. & De Vos, M. The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years. *Glob. Ecol. Biogeogr.* **20**, 73–86 (2011).

**70.** Hu, Y., Fernandez-Anez, N., Smith, T. E. & Rein, G. Review of emissions from smouldering peat fires and their contribution to regional haze episodes. *Int. J. Wildland Fire* **27**, 293–312 (2018).


## Acknowledgements


We acknowledge Dr. Chew Lock Yue and Dr. Patrick Martin for their insightful suggestions that contributed greatly to this paper, and Dr. Liang Zhongyao for his early discussions on statistical analysis. We also acknowledge Dr. Amanda Schwantes, author of the Python package `geeDataExtract`, for her help with the code using her package to access Google Earth Engine data. S.W., S.W.S. and J.S.H.L. receive financial support from an Imperial-NTU Seed Grant (INCF-2021-006) for this study. S.W. is supported by NTU Research Scholarship (PhD).

## Author contributions statement

S.W. designed and conducted the analysis, interpreted the results, and wrote the paper. J.S.H.L advised on the analysis, interpreted the results, and revised the paper. S.W.S. interpreted the results and revised the paper. All authors reviewed the manuscript.

## Additional information

**Competing interests** The authors declare no competing interests.

# Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementary.pdf