ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

Connecting the dots: tracing the origins of 2015 equatorial Southeast Asian fires

To cite this article: Adrian Dwiputra *et al* 2026 *Environ. Res. Lett.* **21** 134044

ENVIRONMENTAL RESEARCH
LETTERS

OPEN ACCESS

RECEIVED
30 January 2026

REVISED
10 June 2026

ACCEPTED FOR PUBLICATION
24 June 2026

PUBLISHED
13 July 2026

LETTER

Connecting the dots: tracing the origins of 2015 equatorial Southeast Asian fires

Adrian Dwiputra[1,2,3,*], Ping Yowargana[3], Janice S H Lee[4], Hoong Chen Teo[2], Zu Dienle Tan[4], Yiwen Zeng[4,5], Andrey Krasovskiy[3] and Lian Pin Koh[1,2,*]

1 Department of Biological Sciences, National University of Singapore, Singapore, Singapore
2 Centre for Nature-based Climate Solutions, National University of Singapore, Singapore, Singapore
3 Biodiversity and Natural Resources (BNR) Group, International Institute of Applied System Analysis, Laxenburg, Austria
4 Asian School of the Environment and Earth Observatory of Singapore, Nanyang Technological University, Singapore, Singapore
5 School of Social Sciences, Nanyang Technological University, Singapore, Singapore
* Authors to whom any correspondence should be addressed.

E-mail: a.dwiputra@u.nus.edu and lpkoh@nus.edu.sg

Keywords: fire origin, graph theory, Southeast Asia, tropical landscape

Supplementary material for this article is available online

Abstract

Tropical landscapes in Southeast Asia are susceptible to the detrimental consequences of fires. However, investigations on ignition sources in the region are limited despite its importance in formulating prevention strategies. We addressed this gap using a systems approach: a graph-theoretic method to identify the spatial origins of the devastating 2015 equatorial Southeast Asian fires. We then fitted a Random Forest model to explain the suitable conditions of fire origins using the identified origin points ($n = 74\,539$–$74\,950$) and predictors that comprised climatic, biophysical, and anthropogenic factors. We found that most fires in the region (∼84%) have multiple (>1) origin points. Overall, our modelling results conform to the anthropogenic origins of fires in the region, though geographical and climatic variables, such as ecoregion and long-term average vapour pressure deficit, act as major predictors of fire origins. The high importance of ecoregions in our model calls for ecoregion-based fire mitigation approaches. Incorporating ecoregion in our analyses has allowed for the identification of previously overlooked areas with high fire origin likelihood and areas potentially undergoing a major increase in fire origin likelihood following likely anthropogenic landscape modifications. Our identification of multi-origin fires in the region emphasises that preventing the expansion of areas with high fire origin likelihood is our best chance of preventing such resilient hazards.

1. Introduction

Wildland fires—henceforth referred to as fires unless specified otherwise—often have detrimental consequences for global society, especially in the tropics. A global assessment that considered both socio-economic and ecological factors found that tropical rainforest landscapes are the most vulnerable to fires (Chuvieco *et al* 2014). Recurring fires in tropical rainforests can trigger a shift to an alternate stable state characterised by lower tree diversity and fire-tolerant species domination (Goldammer 2007), or inhibit regeneration, preventing full forest recovery even after decades (Drüke *et al* 2023). Therefore, effective fire prevention measures are necessary to mitigate the societal and ecological costs associated with large, uncontrollable fires in tropical rainforest landscapes (Page and Hooijer 2016). However, despite having the highest occurrences of fires (Cochrane 2009, Chen *et al* 2026), tropical fires remain understudied compared to temperate or boreal regions (Juárez-Orozco *et al* 2017).

During the El Niño-Southern Oscillation (ENSO)-impacted year of 2015, equatorial Southeast Asia experienced the largest fire catastrophe since 1997 (Chisholm *et al* 2016). ENSO causes abnormally long droughts, low precipitation, and extreme increases in surface air temperatures in the region (Wooster *et al* 2012, Thirumalai *et al* 2017, Lam *et al* 2026), which turn organic matter accumulated in vegetation and peats—that are moist under natural conditions—into dry fuels that easily ignite (Krasovskii *et al* 2018, Widyastuti *et al* 2021). The smoke and haze produced by the fires, which set at least 2.6 million hectares of land ablaze (SiPongi 2015), caused major health, economic, and ecological damage in the region. For just two months in 2015, fires in equatorial Southeast Asia emitted 227 $\pm$ 67 $\times$ 10^3 kg of carbon (Huijnen *et al* 2016). Approximately 100 000 premature deaths in Indonesia, Malaysia, and Singapore were attributable to the transboundary haze (Koplitz *et al* 2016). The World Bank reported that the 2015 fires and haze incurred an estimated cost of 16.1 billion USD to Indonesia, where most of the fires in the region occurred, excluding the economic impacts of fires on the neighbouring countries (Glauber *et al* 2016).

Investigating the causes of the 2015 fires in the region can provide useful information that supports the prevention of such catastrophic events. To uncover the actual causes and ignition sources of fires, forensic investigations of fires are indispensable (Vayda 2006). However, given the imbalance between the limited availability of human power and high fire occurrences—which often occur simultaneously—large-scale post-fire forensic investigations are rarely feasible. Some (failed) attempts to shed light on the causes of fire in 2015 had ultimately become a 'fixed' blame game that led to the criminalisation of smallholder farmers instead of systematic and transparent research into the underlying causes (Lin 2021). Moreover, the politically sensitive nature of fire management topics in the region (Tacconi 2003, 2016, Rochmyaningsih 2020), further complicates efforts to gather useful information for the development of effective fire prevention in the region.

We attempted to address such problems using a systems approach and devised a graph-theoretic (König 1990) method to identify the spatial origins of fires. These spatial origins are the best available approximation of fire ignition sources, which can be analysed with climatic, anthropogenic, and biophysical predictors to provide some clues on the potential factors associated with the 2015 equatorial Southeast Asian fires (Trang *et al* 2022). Through this study we aim to (1) demonstrate the use of a graph-based approach to identify fire origins, (2) describe the spatial characteristics of the 2015 fire origins, and (3) identify areas to prioritise for fire prevention measures. To empirically demonstrate the robustness of our approach and complement our analyses with uncertainty estimates, we included sensitivity analyses using different temporal thresholds during fire graph construction, which we describe in the following section.

2. Materials and methods

2.1. Study region

We studied fires in equatorial Southeast Asia between 7.458675° N 152.438459° E and 11.077562° S 94.977294° E—top right and bottom left coordinates, respectively (figure 1). Considering the spatial resolution of the available data used in the analysis—most of which had a spatial resolution of 1 km^2—we excluded islands that were only partially covered within the study region and islands smaller than 1 km^2.

2.2. Fire graph analysis

We downloaded standard-quality hotspot data recorded by MODIS (Giglio and Justice 2021) and VIIRS (NASA VIIRS Land Science Team 2020) satellites in 2015–2016 from the Fire Information for Resource Management System platform's archive (https://firms.modaps.eosdis.nasa.gov/download/). Compared to the near-real-time data, the standard-quality data we used had gone through a more rigorous quality filtering by the data provider, making it more suitable for our analysis (Giglio *et al* 2021). We included 2016 in our download for possible fires that started in late 2015 and ended in early 2016.

The first step in implementing graph theory (König 1990) was the construction of fire graphs. Each graph consisted of vertices/nodes interconnected through a set of unidirected edges from nodes with earlier detection time to their neighbours. The nodes were the satellite-recorded hotspots, and each edge represented likely fire spreads between a pair of neighbouring nodes. We assume remotely sensed hotspots that are spatiotemporally close to be non-independent of each other (Artés *et al* 2019); instead, they are likely members of the same fire spatial graph (network). Based on the temporal sequence of the hotspots, we can identify the spatial origins of each graph.

The fire graph construction began with a near table generation that we ran in ArcGIS Pro version 3.4.0 (ESRI Inc 2024). Each row in the table contained a pair of 'from' and 'to' nodes (points) connected by a directed edge. The maximum distance in the near table generation was 1500 metres; this parameter was set after considering the lowest spatial resolution of the hotspot data source (MODIS data spatial resolution is 1 km^2) and the queen neighbourhood contiguity criterion we used. Such a criterion allowed all points surrounding the focal point, including the ones diagonally located, to be considered neighbours.

We specified maximum allowed time gaps between the connected 'from' and the 'to' points—henceforth referred to as temporal threshold—to +3 and +7 d. The latter threshold was chosen after considering the revisit interval of the satellites (from quasi-1 d for MODIS to quasi-8 d for VIIRS) and the

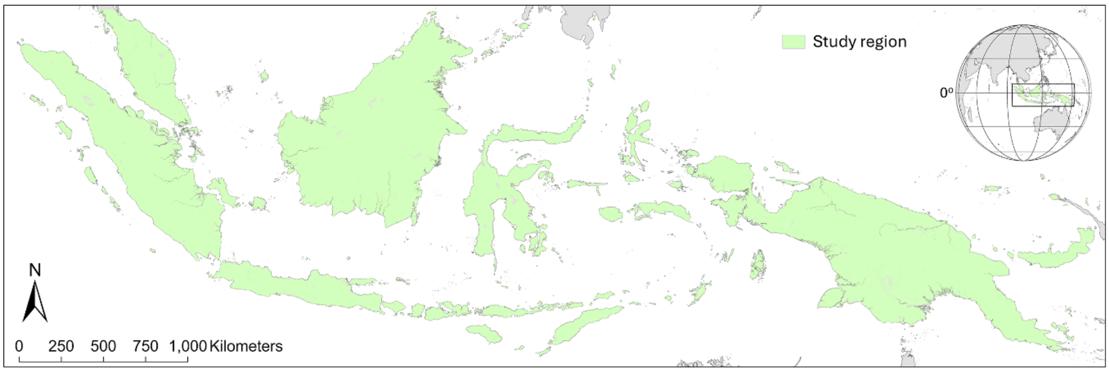


Figure 1. Study region map with an inset (top right) indicating our study's spatial extent and location, equatorial Southeast Asia.

smouldering or underground fires in peat swamps that could go undetected by the satellites for a few days before reigniting (Graham *et al* 2022). The former was added to provide uncertainty estimates and sensitivity analyses. Analysis steps that are described hereafter, unless otherwise described, were implemented in the same way between the independent graphs constructed with the two temporal thresholds. These table filtering steps were run in R version 4.3.3 (R Core Team 2024) using the package dplyr version 1.1.4 (Wickham *et al* 2023).

Following the table filtering steps, we applied a spatial filter to the remaining rows/edges in the near table. Edges that intersected waterbodies were removed because we assumed that waterbodies acted as fire breaks. We removed edges that intersected waterbodies using the *R* package terra version 1.8–29 (Hijmans 2025). We then used the remaining edges and their associated nodes for fire graph construction using *R* package igraph version 2.1.4 (Csárdi *et al* 2025).

2.2.1. Network filtering

Following the fire graphs construction, we retained only graphs with at least three connected 'from'—'to' nodes (points) to omit potential false positives. Then, for each graph, we created a minimum bounding polygon in the form of a convex hull to depict the spatial coverage of each network. Given the limited spatial resolution of the data used in our analysis, we omitted graphs with convex hull areas smaller than 1 km^2. We then excluded graphs that were located on waterbodies; to do so, we filtered graphs with convex hulls that covered four pixels or less of land area. To further increase the specificity of the fire graphs to wildland fires, we only included graphs with convex hulls covering less than 10% urban or barren lands out of the total land area covered by the respective graphs' convex hulls. Lastly, we omitted graphs with lifetime $\geqslant 5$ months to remove some false positives related to prolonged and consistent thermal anomalies, probably linked to volcanic activities (Tan *et al* 2020). We implemented these network filtering steps using *R* packages igraph (Csárdi *et al* 2025), terra (Hijmans 2025), and dplyr (Wickham *et al* 2023).

2.2.2. Origin points extraction

We extracted origin points from the filtered fire graphs in igraph 2.1.4 (Csárdi *et al* 2025). Origin points are nodes in fire graphs with no preceding neighbours, i.e. no neighbour within the same graph that occurred before them. In other words, origin points were characterised by the absence of any edges directed towards them and the presence of at least one edge directed from them. On the other hand, non-origin points in fire graphs, i.e. points with at least one edge directed towards them—from either origin or other non-origin points—were categorised as spread points. Given the chance that some fire spread events were faster than the revisit time of the satellites, the origin points comprise both the actual source(s) of fires and points that were spatiotemporally closest to the actual sources. The definition of origin and spread points can be mathematically expressed as:

$$G = (V,\ E) \text{ and } |V| \geqslant 3, v \in V$$
$$\deg^-(v_o) = 0 \text{ and } \deg^+(v_o) \geqslant 1$$
$$\deg^-(v_s) \geqslant 1$$

where

G : Graph
V : Nodes (Vertices)
E : Edges
$\deg^-$: Indegree, the number of unidirected edges coming towards a node
$\deg^+$: Outdegree, the number of unidirected edges coming from a node
v_o : Origin points
v_s : Spread points

Given these definitions, each graph contains at least three nodes that comprise at least one and at most $|V| - 1$ origin point(s). From this point on, the origin and spread points analysed belonged only to graphs whose origin points were detected in 2015.

2.3. Random forest with down-sampling (RFd)

We implemented RFd for two main purposes: to develop a fire origin likelihood model and to simulate the difference between models developed with fire origin input and models with fire spread input using Monte Carlo simulations. A recent comparative analysis (Valavi *et al* 2022) showed that RFd is among the best-performing methods in distribution modelling. Compared to the original random forest, RFd enforces balanced bootstrap sample sizes across groups analysed, even when the total number of samples differed, which substantially improves its performance compared to the original random forest (Valavi *et al* 2022). Before going into the modelling steps conducted in *R* package biomod2 version 4.2-6-2 (Thuiller *et al* 2024), we explained the steps taken in preparing the additional materials required for RFd implementation, namely absence datasets and spatial predictors.

2.3.1. Absence datasets

We followed the recommended practice from the biomod2 developer team (Thuiller *et al* 2024) of having 10 absence datasets to account for the model's uncertainty. We generated 10 absence datasets, each containing the same number of absence points as the origin points obtained from the graph analysis step, to enhance balance while retaining representativeness (Barbet-Massin *et al* 2012). These points were products of ten iterations of the balanced acceptable samples method implemented in *R* package spbal version 1.0.0 (Robertson *et al* 2025) across the study region outside 4 km buffers from the retained origin and spread points. We used 4 km buffers to avoid false negatives associated with the satellite data's spatial/temporal resolution.

2.3.2. Spatial predictors

We used eight spatial predictors in our analysis to model fire origin likelihood. We selected these predictors based on their relationship with fires documented in previous studies (Goldammer 2007, Tacconi 2016, Tan *et al* 2020, Wee *et al* 2024). In addition, the number of predictors included was designed to balance the number of anthropogenic and non-anthropogenic predictors (table 1) to minimise potential bias related to the anthropogenic origins of the fires in the study region.

Given the prevalence of human-ignited fires in the region (Cochrane 2009), we expected fire origin likelihood to respond to spatial predictors associated with anthropogenic activities. For instance, lowlands (Tan *et al* 2020), rural areas (Edwards *et al* 2020), agricultural lands (Varma 2003, Edwards *et al* 2020), and deforestation hotspots (Goldammer 2007, Busch and Ferretti-Gallon 2023, Wee *et al* 2024) are often associated with high fire likelihoods; we expected our predictors, such as Human Footprint, Land Cover, Accessibility, and Elevation, to be reflective of such tendencies. In terms of climate, we expected a positive relationship between vapour pressure deficit (VPD) and fire origin likelihood because higher VPD is indicative of drier and warmer conditions and found to be positively correlated with forest fire risks (Clarke *et al* 2022).

Following the preprocessing steps (supplementary material A), we calculated Pearson's correlation coefficient to assess potential multicollinearity in our predictors. The results (table 2) showed that the only pair with correlation coefficients larger than 0.5 was the elevation and VPD layers (−0.879). Nevertheless, we retained all predictors in our final model after considering the Random Forest's insensitivity to multicollinearity (Kaveh *et al* 2023) and to reduce a likely bias caused by the removal of correlated variables in Random Forest (Hanberry 2024).

While lightning has a potentially increasing significance as an ignition source in the region under climate change (Pérez-Invernón *et al* 2023), we did not incorporate it in our analysis for several reasons. Firstly, lightning is a rarer ignition source in equatorial Southeast Asia than in continental Southeast Asia (Eaton and Radojevic 2001, Cochrane 2009), and even there anthropogenic ignition sources are dominant to the lightning-ignited fires for—at least—the past few decades (Cochrane 2009). Moreover, most of the analysed fires occurred during an abnormally long dry season in 2015—due to the ENSO—which made lightning-ignited fires less likely as lightning in the region is closely coupled with rainstorms (Hamid *et al* 2001). Lastly, the coarse spatial resolutions of available lightning data, e.g. 9 km (Kaplan and Lau 2021) or 11 km (Cecil *et al* 2026), would have caused a substantial reduction in our mapping resolution. Related to this, during the early phase of our analysis, we included lightning frequency data from Kaplan and Lau (2021)—resampled into 1 km—and found it to be among the predictors with the lowest importance (supplementary material B).

2.3.3. Fire origin likelihood model

We input the identified origin points (as presence data), absence datasets, and the preprocessed predictors to RFd models. The model fitting was conducted in *R* package biomod2 version 4.2-6-2 (Thuiller *et al* 2024). We followed the parameter settings optimised by the package developer team, referred to as the *bigboss* setup (Thuiller *et al* 2024). To make a more informed decision on parameter settings, we compared another setup used by Valavi *et al* (2022) (table 3). We used cross-validation with 10 replications of random separation, with a 7:3—training: validation—ratio, for all models under comparison. The comparisons of true skill statistic (TSS) and area under the Receiver Operating Characteristic curve (ROC) values showed that the *bigboss* setup produced slightly better cross-validation accuracy (table 4) with shorter processing time (8 h 24 min vs 15 h 13 min for

Table 1. Spatial predictors for modelling of fire origin likelihood across equatorial Southeast Asia.

Data	Year	Source	Original resolution
Climatic			
Averaged vapour pressure deficit	1980-2014	Monthly data CHELSA-BIOCLIM+(Brun *et al* 2022)	1 km^2
Biophysical			
Elevation	—	NASADEM Merged DEM (NASA JPL 2020)	30 × 30 m^2
Ecoregion	—	Terrestrial ecoregions of the world (Dinerstein *et al* 2017)	Spatial polygons rasterised into 1 × 1 km^2
Distance to water bodies	2015	Global surface water (Pekel *et al* 2016)	30 × 30 m^2
		Southeast Asian peatland drainage canals (Dadap *et al* 2021)	5 × 5 m^2
Anthropogenic			
Dominant land cover	2014	ESA CCI Land Cover v. 2.0.7 (European Space Agency 2017)	300 × 300 m^2
Land cover diversity	2014	ESA CCI Land Cover v. 2.0.7 (European Space Agency 2017)	300 × 300 m^2
Human footprint	2014	An annual global terrestrial Human Footprint dataset (Mu *et al* 2022)	1 km^2
Accessibility	2015	Global travel time to cities (Weiss *et al* 2018)	1 km^2

Table 2. Pearson's correlation coefficients for all pairs of continuous predictors in the model.

Predictor	Human footprint	Land cover diversity	Elevation	Travel time to city	Mean VPD	Distance to water
Human footprint	1	0.391	−0.165	−0.490	0.326	−0.219
Land cover diversity	0.391	1	−0.251	−0.329	0.334	−0.262
Elevation	−0.165	−0.251	1	0.264	−0.880	0.351
Travel time	−0.490	−0.329	0.264	1	−0.323	0.205
Mean VPD	0.326	0.334	−0.880	−0.323	1	−0.397

Table 3. Summary of parameters used in RFd modelling with different settings under comparison.

Parameter name	Description	*Bigboss* (Thuiller *et al* 2024) settings	Valavi *et al* (Valavi *et al* 2022) settings
Ntree	Number of classification trees grown in the forest for each RFd	500	1000
Mtry	Number of variables randomly selected at each split	2	$\sqrt{\text{number of predictors}}$, rounded down = 2
Nodesize	The minimum number of observations in a terminal node. This parameter is often increased to reduce overfitting.	5	10

Table 4. Summary of cross-validation metrics used in comparing RFd models with different settings.

	bigboss (Thuiller *et al* 2024) settings		Valavi *et al* (2022) settings	
	3 d	7 d	3 d	7 d
Mean TSS	0.751	0.727	0.747	0.724
Mean ROC	0.951	0.942	0.950	0.942

models using the 7 d threshold). Therefore, our final results presented in this paper were the products of runs with the *bigboss* setup.

We fitted 10 RFd models following *bigboss* settings, each with a different absence dataset. After the model fitting concluded, we aggregated the model outputs, namely the permutation importance, response curves, and the model's prediction (Breiman 2001). The permutation importance calculation results were averaged across models and presented along with their associated standard deviation. For the response curves of continuous and categorical predictors, we plotted smoothed splines and the associated standard errors or their means and standard deviations, respectively.

2.3.4. Monte Carlo simulation

We quantified the difference between fire models developed using origin points and those developed with spread points by performing a Monte Carlo simulation. All model settings implemented in our Monte Carlo simulation, as well as the absence datasets, were identical to the fire origin likelihood modelling; the only difference was the presence data. Given more spread than origin points, our Monte Carlo simulation used presence data from 100 randomly sampled sets of spread points; each set comprised the exact number of points as the origin points used in the fire origin modelling. We ran 1000 iterations—all possible combinations of the presence (100) and the absence datasets (10)—of RFd in a high-performance computer with 320 GB of RAM across 2 nodes, each with 20 CPUs.

Following the Monte Carlo simulation, we statistically compared the outcomes of models developed with spread points against those developed using origin points. We fitted a linear mixed model (supplementary material C) to evaluate the differences between each of the models' predictor importance by restricted maximum-likelihood in *R* package lme4 version 1.1–36 (Bates *et al* 2015) coupled with lmertest version 3.1–3 (Kuznetsova *et al* 2017).

2.4. Fire origin likelihood map

We mapped fire origin likelihood at 1 km resolution by averaging the medians of fire origin likelihood model predictions (2.3.3) made with 3 d and 7 d thresholds. To complement the fire origin likelihood map, we mapped absolute differences between the 3 d and 7 d medians to visualise uncertainty estimates of the fire origin likelihood across the study area. Our fire origin likelihood map offers spatial highlights of where fire prevention in the region can be prioritised.

3. Results and discussion

3.1. Multiple origins of fires

We identified between 74 539 and 74 950 origin points from 14 964 to 15 424 fire graphs that occurred in 2015. The lower and higher values in the ranges were obtained from 7 d and 3 d thresholds, respectively. Most fire graphs ($\sim$84%) have multiple ($>$1) origin points (figure 2). Among them, a small proportion of graphs (6.8%–8.7%) have more origin points than spread points, while the remaining graphs have more spread points than origin points (figure 2(A)). The multiple origins in fire graphs indicate several spatiotemporally independent origin points merged at some point during the fire spreads that followed ignition events. While under ideal conditions origin points pinpoint the locations of ignition sources, in cases where the satellites' temporal resolution was lower than the speed of fire spread from their ignition source(s) or where cloud, haze, or canopy obscure hotspot detection (Giglio *et al* 2016, Liu *et al* 2020), fire origin points were the recorded points that were spatiotemporally closest to the actual ignition sources.

Fires that covered larger areas, as indicated by both the number of spread points and convex hull coverage, tend to have more origin points (figure 2). The largest graph identified in this analysis, with 50 906 points (nodes), was consistent with such general patterns (red outline, figure 3(B)). We observed that using the 3 d threshold would break such a graph into one large and twenty small graphs (yellow outline, figure 3(B)). Such results and the fact that several clusters of origin points were still detected as part of the large graph in the 3 d threshold seem to consistently suggest the multiple spatial origins of fires. From the systems perspective (Meadows and Wright 2011), fires with multiple origins are more resilient systems than those with fewer origins. Addressing such complex and resilient environmental hazards requires a better understanding of its ecological characteristics, which is covered by our modelling components.

Despite our study's focus on origin points, graph theory can be applied beyond the identification of origin and spread points. A graph-theoretical framework can contribute to fire behaviour studies that are often faced with issues related to data unavailability. For instance, we will be able to study landscape components that act as important spread points based on their degree of centrality, a graph metric that indicates how connected a node is (Aparício *et al* 2022).

Unfortunately, our current study's temporal scope was not ideal for covering spread behaviour as well. Limitations related to the number of hotspot-detecting satellites in 2015 produced data with a coarse temporal resolution, hence the temporal thresholds of 3 and 7 d. However, at the time this paper was written, the temporal resolution of hotspot detection has improved from quasi-8 d to sub-daily since 2024, following the launch of the newer satellites NOAA-20 in 2017 and NOAA-21 in 2022. Such improvements can facilitate further applications of graph theory to better understand fire behaviour. Regardless of our limitations, we hope this study can serve as an empirical basis to encourage wider applications of graph theory in fire analyses with data-limited contexts.

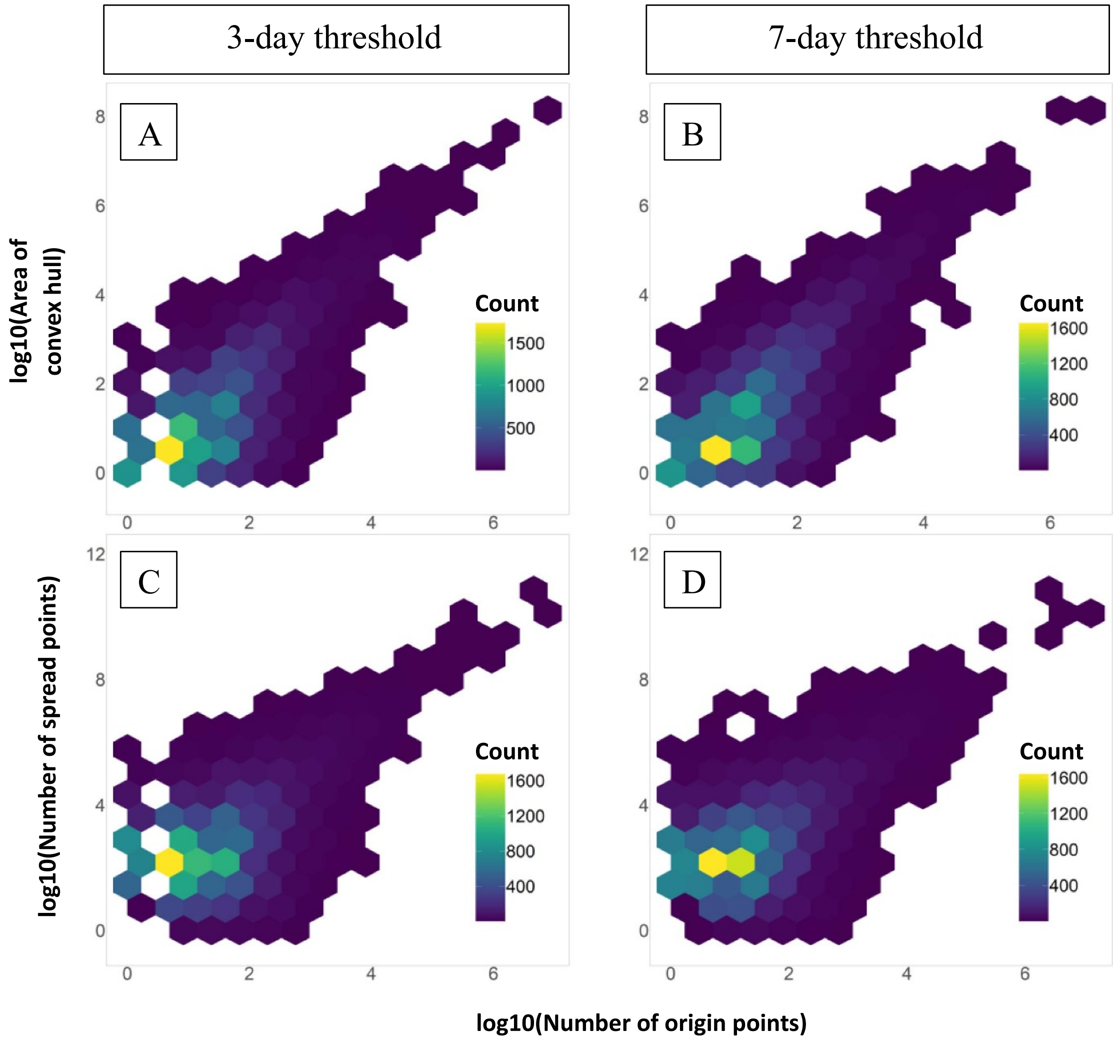


Figure 2. Hexagonal heatmap of bin counts (bins = 15) displaying the number of origin points against (A), (B) convex hull area and (C), (D) spread points. Consistent overall trends with small differences were observed between analyses using different temporal thresholds, the 3 d (A), (C) and 7 d (B), (D).

3.2. The combination of climate and people in shaping fire origins

The highest permutation importance of the ecoregion layer (figure 4) shows diverse ecological contexts associated with fire origins within the study region. Similarly, geographic variations in fire characteristics have also been recorded in continental Southeast Asia, at the provincial level in North Vietnam (Trang *et al* 2022). The ecoregion response plot (figure 5) shows the relative likelihood of ecoregion types in harbouring conditions suitable for fire origins, given all other predictors set to be equal to their respective average values. Since the modelled fire origin likelihood (section 3.4) is a product of the combined influences of all predictors, high values in the ecoregion's response curve do not necessarily imply high origin likelihood across the pixels within the ecoregion coverage, and vice versa. Our results suggest the need to incorporate ecoregion types in identifying areas to be prioritised for fire prevention.

VPD is the predictor with the second highest importance (figure 4). VPD represents long term climatic measurement of atmospheric dryness that affects vegetation (Novick *et al* 2024) and fires (Seager *et al* 2015, Clarke *et al* 2022). One of our response curves indicates an optimal range of VPD for fire origins in the region between 8 and 12 hPa (figure 6). A similar-shaped response to VPD has been documented in Sierra Nevada fires (Chen *et al* 2021), despite the lack of similarities in ecological contexts and focus on fire origin in such a study.

As climate change is expected to increase VPD in the tropics by 5.24 hPa in 2100 (Fang *et al* 2022), areas with high fire origin likelihood in the region will likely increase due to the increased fuel dryness (Novick *et al* 2024). Furthermore, VPD increase might also indirectly contribute to increased slash-and-burn activities as land demands increase due to crop yield reduction (Novick *et al* 2024); cropland expansions as a global response to drought associated with rainfall anomalies have been reported (Zaveri

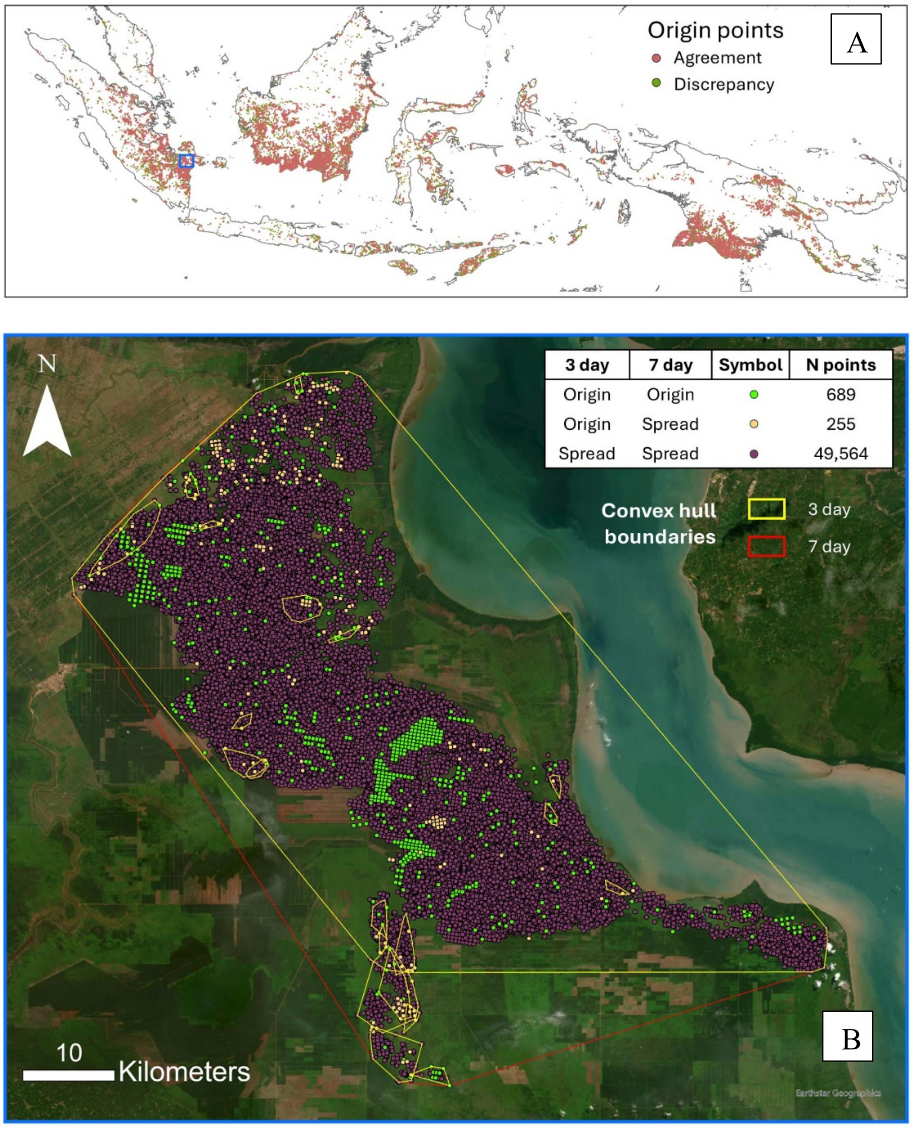


Figure 3. (A) The spatial distribution of the identified origin points across equatorial Southeast Asia in 2015. 'Agreement' and 'discrepancy' denote origin points identified consistently regardless of the temporal threshold and the ones that were not, respectively. (B) The largest identified fire graph (fire graph 13616) in Ogan Komering Ilir, with the origin and spread points mapped. This complex graph was first identified using the 7 d threshold. The total number of points on the table inset is less than 50 906 points because 398 points with duplicated coordinates were removed from the calculation.

et al 2020). These results suggest that incorporating VPD-fire feedback in future research may alter earlier findings on the invulnerability of Southeast Asian rainforests as tipping elements (Staal *et al* 2020, Armstrong Mckay *et al* 2022).

In general, our models capture the relationship of fire with the spatial predictors in similar ways to the hypothesised outcomes. The anthropogenic predictor with the highest importance is the dominant land cover. Needle-leaf deciduous forests, sparse vegetation, and broad-leaved evergreen forests had a higher likelihood of fire origin points (figure 5). High fire origin likelihood associated with deciduous forests was likely due to the shedding of leaves during the dry season, creating a buildup of fuel material (Stott *et al* 1990). In contrast, areas of sparse vegetation and grasslands can be associated with degradation and are often dominated by herbaceous vegetation that can be easily ignited (Goldammer 2007).

We noticed that broad-leaved evergreen forests had a lower likelihood than grasslands in models using a 7 d threshold (figure 5), suggesting possible mixing of spread points with origin points when the 3 d threshold was used. At face value, the high fire origin likelihood associated with broad-leaved evergreen forests seems to contradict previous findings, such as a study in Kalimantan peat swamps that shows most fires (68%–71%) originated in non-forest lands (Cattau *et al* 2016). However, the high fire origin likelihood (figure 6) associated with forests may have stemmed from rapid deforestations that occurred shortly before the ignitions (Goldammer 2007) and, therefore, not depicted in the land cover map used.

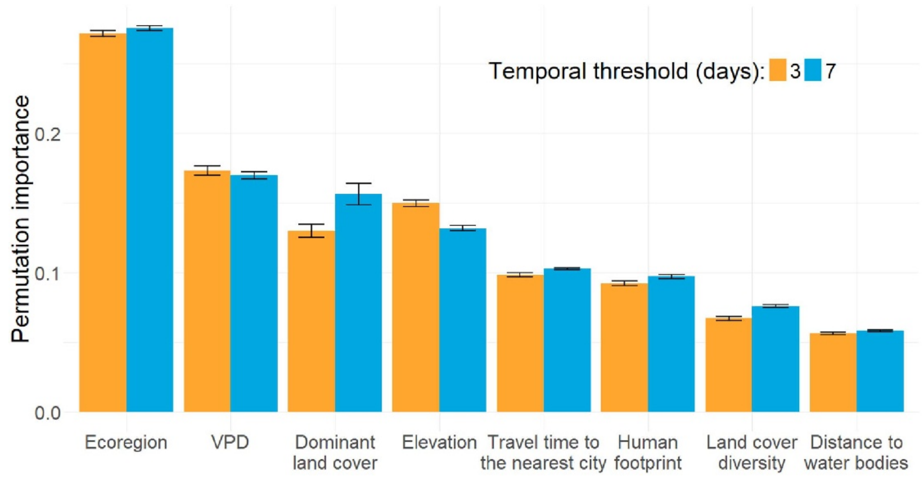


Figure 4. Permutation importance (mean and standard deviation) of predictors used in fire origin models. Each bar was calculated based on 3 permutations and 10 absence sets (n = 30).

The response curves of access to cities, human footprint, and elevation (figure 6) collectively seem to describe the spatial characteristics of land-use frontiers associated with fire origins (Busch and Ferretti-Gallon 2023). Land-use frontiers are parts of landscapes where land-use conversions, such as the conversion of natural ecosystems into production land, are concentrated (Meyfroidt *et al* 2024). The slight increase in fire origin likelihood at higher elevation (figure 6) may reflect the upward expansion of land-use frontiers in the region (Feng *et al* 2021). Future research equipped with higher spatiotemporal-resolution land cover data, ideally at sub-annual time steps, can expand the empirical basis for this hypothesis, while also identifying the rapid deforestation prior to fires that inflated the fire origin likelihood associated with broad-leaved evergreen forests.

Despite their relative rarity compared to human ignited fires (Cochrane 2009, Taylor 2010), natural ignitions may offer an alternative explanation for the observed patterns. For instance, other than the aforementioned upward expansion of land-use frontiers, volcanism might also contribute to the increase in fire origin likelihood at higher elevations (figure 6). Such a possibility still exists because we only excluded false positive graphs with lifetime $\geqslant$5 months, allowing fires associated with temporary increases in volcanic activity to be retained. Research into natural ignition mechanisms in the region, supported by better quality data, may offer estimates on the proportions of naturally ignited fires and their expected dynamics in climate change scenarios, as demonstrated by Pérez-Invernón *et al* (2023) for global lightning ignitions.

3.3. More apparent anthropogenic links with origin points than spread points

Our results showed statistically significant differences in models' variable importance between models fitted with inputs that comprised only origin and spread points (figure 7). With spread points as input, all anthropogenically altered predictors in the model—land cover, human footprint, travel time—and distance to water bodies, which acts as a proxy of accessibility in riparian contexts (Sumarga 2017), have lower importance values than with origin points as input. Correspondingly, we hypothesise that such effects will be even more pronounced as we use input points that are closer to the ignition sources. While insufficient to provide direct evidence of ignition mechanisms, these results—complementing the results from other parts of our study—conform with the literature that most fires in the region are human-ignited (Cochrane 2009, Edwards *et al* 2020).

Our results suggest that studies on fire drivers should ideally use origin points as inputs. Ignoring significant differences between origin and spread points, as found in our study, will produce biased estimations of the potential predictor's importance. For instance, studies on potential fire drivers in Kalimantan's peat swamp forests would have underestimated the significance of anthropogenic factors when forest boundaries were used as a subsetting layer because most fires there originated in non-forest lands (Cattau *et al* 2016). Consequently, the mixing of origin and spread points as inputs may dilute the importance of anthropogenic factors, given the large ratio between spread and origin points (figure 2(A)).

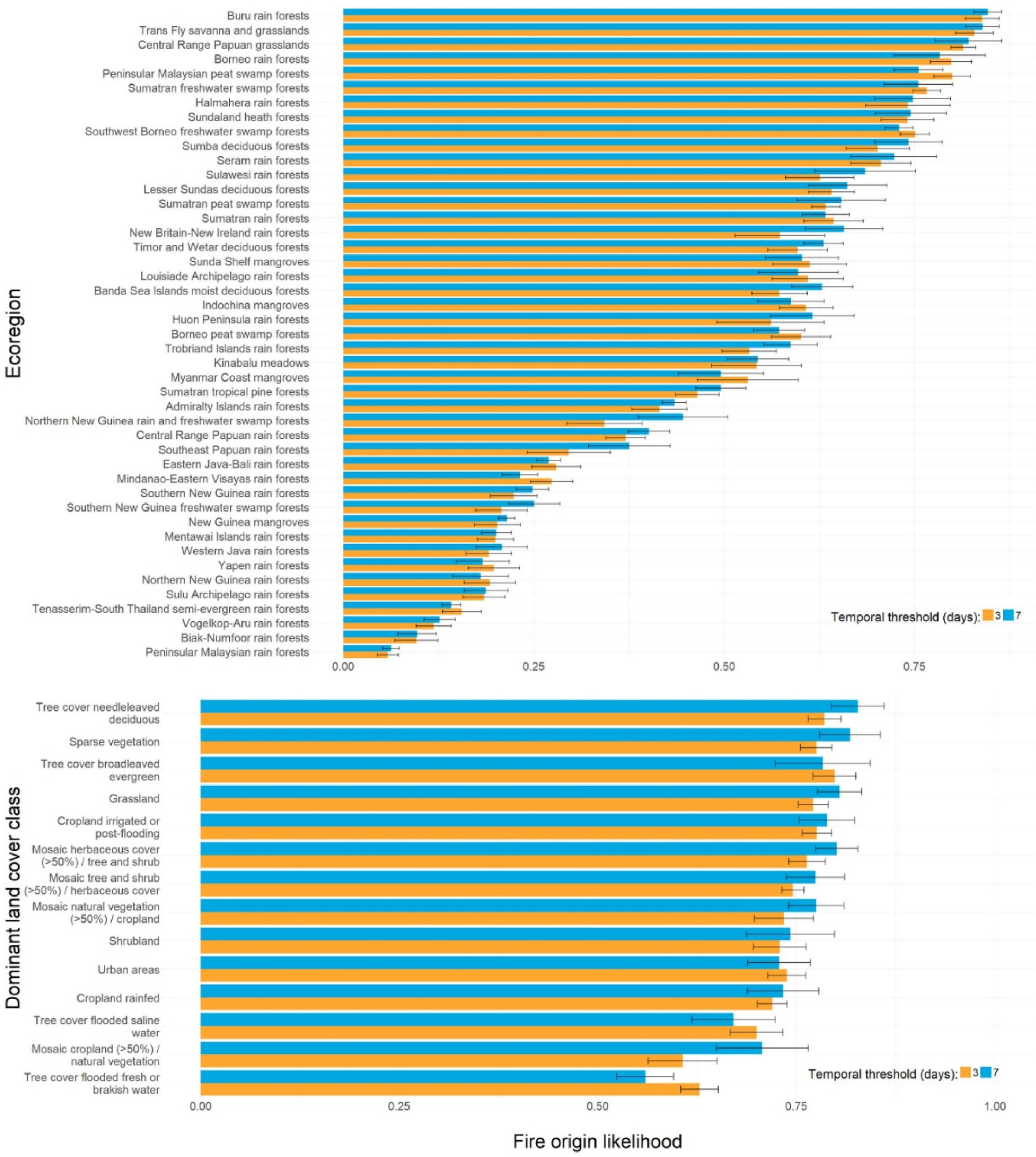


Figure 5. Response plots of the categorical model predictors: Ecoregion and dominant land cover. Error bars are standard deviations calculated based on 10 spatially balanced absence sets (n = 10).

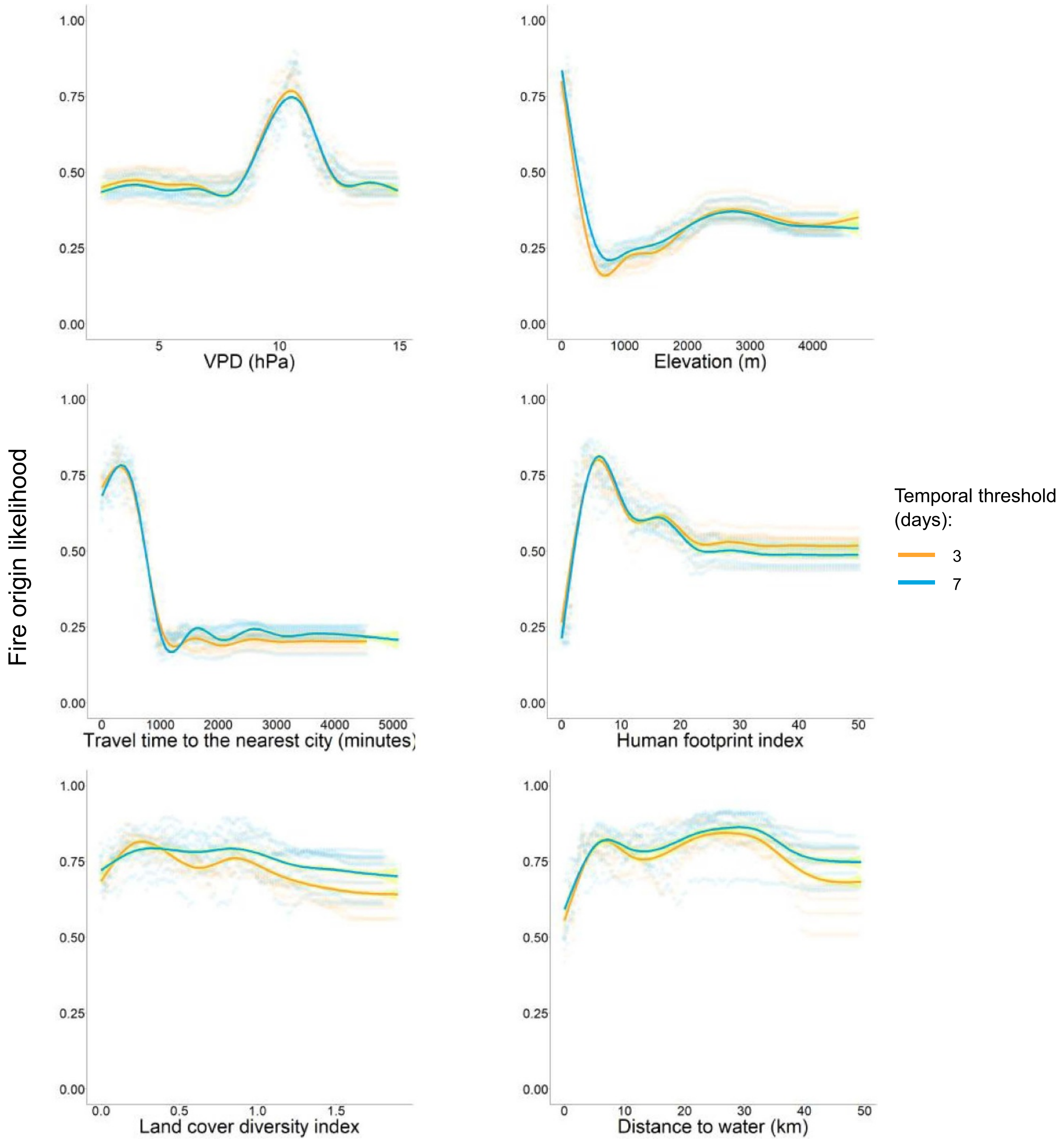
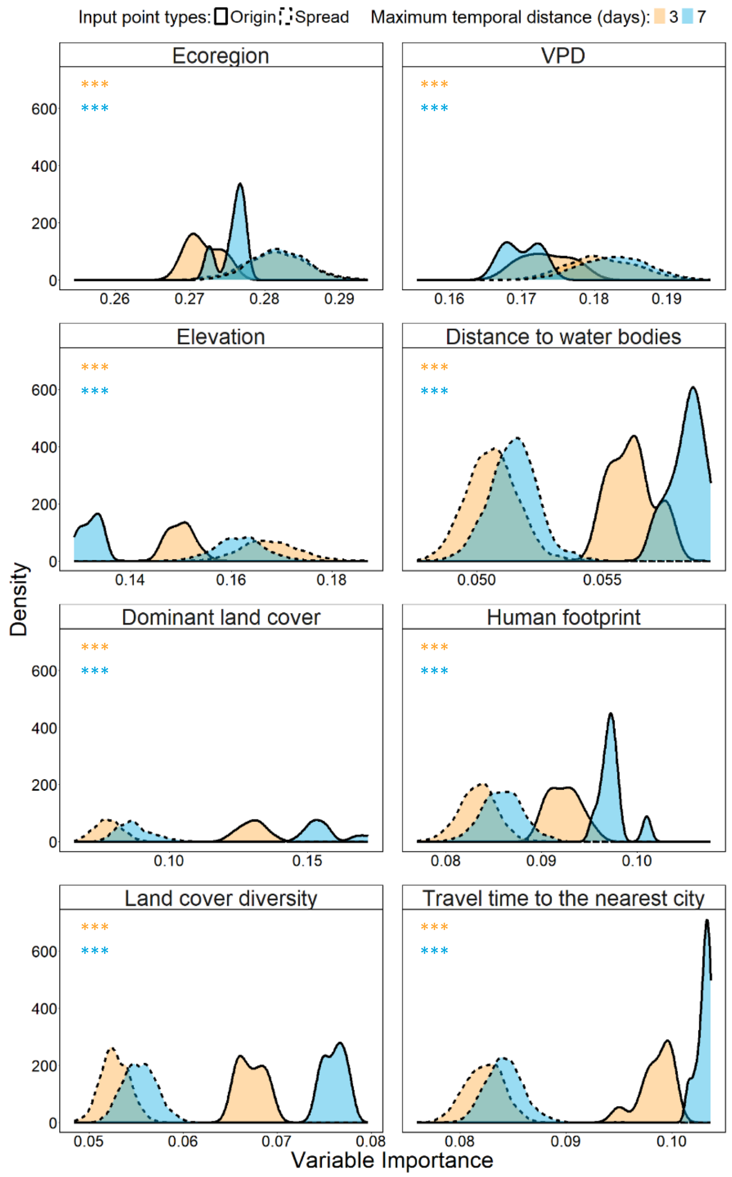


Figure 6. Response curves of the continuous model predictors, with smoothed spline (blue) and the associated standard error (yellow), highlight trends to indicate the complex relationship between different spatial predictors and fire origin likelihood. Variation intervals were calculated based on 10 spatially balanced absence sets (n = 10).

3.4. Regional priorities for fire preventions

We suggest prioritising fire prevention efforts in areas with high fire origin likelihood (figure 8(A)), especially ones that also belong to ecoregions associated with high likelihood (figure 5). While in most part the map depicted the known, relatively well-studied fire-prone areas such as the eastern coast of South Sumatra, Central Kalimantan, and Southern New Guinea (Tan *et al* 2020), we also identified an overlooked ecoregion, namely Buru rainforest (figure 5). Unfortunately, more than half of its area was already associated with high fire origin likelihood values in 2015 (figure 8(A) inset) without ever being noticed—to the best of our knowledge—in previous studies. Existing tools, such as the digital community-based fire management, can be prioritised here and in other areas with similar urgency.

Our findings also draw attention beyond areas already mapped with high fire origin likelihood. For instance, our results (figure 5) suggest extensive areas with low fire origin likelihoods in Borneo rainforest ecoregion (figure 8(A)) could shift into high fire origin likelihood areas following large-scale landscape changes that create suitable conditions for fire origins, such as prospective road networks traversing the Heart of Borneo following the establishment of the new Indonesian capital Nusantara (Teo *et al* 2020). Our findings on multi-origin fires in equatorial Southeast Asia emphasise that preventing the expansion of areas with high likelihood of fire origin is our best chance of mitigating such resilient hazards.



Star signs indicate statistically significant differences between two point types under comparison: origin and spread points. Accordingly, orange stars indicate significant differences between the origin and spread points analysed using a 3-day threshold, while the blue ones indicate the origin-spread differences in the 7-day threshold.

*** p-value < 0.001

Figure 7. Changes in variable importance indicate the significant influence of input points in unravelling or concealing the links between fire origins and anthropogenic factors in the region. Please note the different scales on the x-axes.

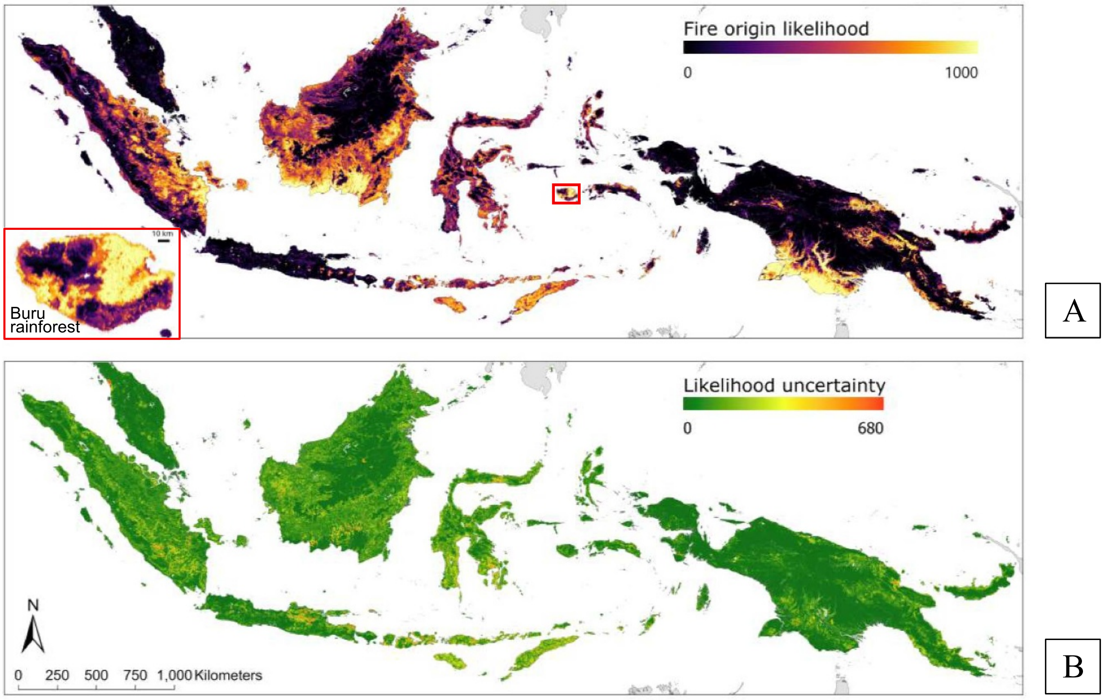


Figure 8. Fire origin likelihood across equatorial Southeast Asia in 2015 (A) and associated uncertainty (B). The likelihood values are averaged median composites of predictions obtained using 3 d and 7 d thresholds; each median composite was constructed from 10 model runs with different spatially balanced absence points. The likelihood values, which are scaled by 1000, depict the variability of fire origin likelihood among 1 × 1 km^2 pixels within the study area (relative values).

Acknowledgments

L P K is supported by the Singapore National Research Foundation (NRF-RSS2019-007) and the Singapore National Research Foundation and Agency for Science, Technology and Research LCER Phase 2 funding programme (U2303D4101). J S H L is funded by the Climate Transformation Program, a Tier 3 Academic Research Fund from the Singapore Ministry of Education (MOE-MOET32022-0006). A D is supported by National University of Singapore (NUS) Research Scholarship (NUS RCA-Apr2021 2021-0588). This work receives support from the RESTORE+ project (www.restoreplus.org), which is part of the International Climate Initiative (IKI), supported by the Federal Ministry for the Environment, Climate Action, Nature Conservation and Nuclear Safety (BMUKN) based on a decision adopted by the German Bundestag. The high-performance computational facility involved in this research was supported by NUS IT's Research Computing group. The earlier version of our analysis was developed in collaboration with ICRAF Indonesia in 2016. We acknowledge the use of data from NASA's Fire Information for Resource Management System (FIRMS) part of NASA's Land, Atmosphere Near real-time Capability for Earth observations (LANCE) and NASA's Earth Science Data and Information System (ESDIS).

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Supplementary Materials available at: https://doi.org/10.1088/1748-9326/ae81ca/data1.

Data and code availability

All analyses in this paper were conducted using publicly available data accessible at the sources mentioned in Materials and Methods in detail. Codes used for the analyses are available at: https://github.com/adwiputra/fireModeling/releases/tag/v1.0.0.

Conflict of interest

All authors declare no conflict of interest.

Author contributions

A D, P Y, J S H L, Z Y and L P K designed the study. P Y and L P K provided funding acquisition and resources. A D, P Y, J S H L and H C T collected and analysed the data. A D, P Y, J S H L and L P K interpreted the data. A D, Z D T, P Y, J S H L, H C T, Z Y, A K, and L P K wrote the manuscript.

ORCID iDs

Adrian Dwiputra 0000-0002-9325-571X
Ping Yowargana 0000-0002-4854-0044
Janice S H Lee 0000-0001-6368-6212
Hoong Chen Teo 0000-0003-0127-978X
Zu Dienle Tan 0000-0001-5854-6987
Yiwen Zeng 0000-0001-7787-314X
Andrey Krasovskiy 0000-0003-0940-9366
Lian Pin Koh 0000-0001-8152-3871

References

Aparício B A, Pereira J M C, Santos F C, Bruni C and Sá A C L 2022 Combining wildfire behaviour simulations and network analysis to support wildfire management: a Mediterranean landscape case study *Ecol. Indic.* **137** 108726

Armstrong Mckay D I, Staal A, Abrams J F, Winkelmann R, Sakschewski B, Loriani S, Fetzer I, Cornell S E, Rockström J and Lenton T M 2022 Exceeding 1.5 °C global warming could trigger multiple climate tipping points *Science* **377** eabn7950

Artés T, Oom D, De Rigo D, Durrant T H, Maianti P, Libertà G and San-Miguel-Ayanz J 2019 A global wildfire dataset for the analysis of fire regimes and fire behaviour *Sci. Data* **6** 296

Barbet-Massin M, Jiguet F, Albert C H and Thuiller W 2012 Selecting pseudo-absences for species distribution models: how, where and how many? *Methods Ecol. Evol.* **3** 327–38

Bates D, Mächler M, Bolker B and Walker S 2015 Fitting linear mixed-effects models using lme4 *J. Stat. Soft* **67** 1–48

Breiman L 2001 Random forests *Mach. Learn.* **45** 5–32

Brun P, Zimmermann N E, Hari C, Pellissier L and Karger D N 2022 CHELSA-BIOCLIM+ A novel set of global climate-related predictors at kilometre-resolution (available at: www.envidat.ch/dataset/bioclim_plus)

Busch J and Ferretti-Gallon K 2023 What drives and stops deforestation, reforestation, and forest degradation? An updated meta-analysis *Rev. Environ. Econ. Policy* **17** 217–50

Cattau M E, Harrison M E, Shinyo I, Tungau S, Uriarte M and DeFries R 2016 Sources of anthropogenic fire ignitions on the peat-swamp landscape in Kalimantan, Indonesia *Glob. Environ. Change* **39** 205–19

Cecil D J, Buechler D E, Lang T J, Virts K S and Mach D M 2026 Lightning climatology datasets from TRMM LIS, ISS LIS, and OTD *J. Appl. Meteorol. Climatol.* **65** 51–72

Chen B, Jin Y, Scaduto E, Moritz M A, Goulden M L and Randerson J T 2021 Climate, fuel, and land use shaped the spatial pattern of wildfire in California's Sierra Nevada *JGR Biogeosci.* **126** e2020JG005786

Chen X-T, Kang S-C, Ji Z-M, Yang J-H, Hu Y-L, Xu M, Ma M-M and Zhou Y 2026 Divergent trajectories of global fires: overall decline, regional intensification *Adv. Clim. Change Res.* **17** 575–87

Chisholm R A, Wijedasa L S and Swinfield T 2016 The need for long-term remedies for Indonesia's forest fires *Conserv. Biol.* **30** 5–6

Chuvieco E, Martínez S, Román M V, Hantson S and Pettinari M L 2014 Integration of ecological and socio-economic factors to assess global vulnerability to wildfire *Glob. Ecol. Biogeogr.* **23** 245–58

Clarke H, Nolan R H, De Dios V R, Bradstock R, Griebel A, Khanal S and Boer M M 2022 Forest fire threatens global carbon sinks and population centres under rising atmospheric water demand *Nat. Commun.* **13** 7161

Cochrane M A 2009 *Tropical Fire Ecology: Climate Change, Land Use, and Ecosystem Dynamics* (Springer) Published in association with Praxis Pub

Csárdi G, Nepusz T, Müller K, Horvát S, Traag V, Zanini F and Noom D 2025 igraph for R: R interface of the igraph library for graph theory and network analysis (available at: https://zenodo.org/doi/10.5281/zenodo.7682609)

Dadap N C, Hoyt A M, Cobb A R, Oner D, Kozinski M, Fua P V, Rao K, Harvey C F and Konings A G 2021 Drainage Canals in Southeast Asian peatlands increase carbon emissions *AGU Adv.* **2** e2020AV000321

Dinerstein E *et al* 2017 An ecoregion-based approach to protecting half the terrestrial realm *BioScience* **67** 534–45

Drüke M, Sakschewski B, Von Bloh W, Billing M, Lucht W and Thonicke K 2023 Fire may prevent future Amazon forest recovery after large-scale deforestation *Commun. Earth Environ.* **4** 248

Eaton P and Radojevic M 2001 Forest fires and regional haze in Southeast Asia (Nova Science Publishers) (available at: https://books.google.at/books?id=pQrbAAAAMAAJ)

Edwards R B, Naylor R L, Higgins M M and Falcon W P 2020 Causes of Indonesia's forest fires *World Dev.* **127** 104717

ESRI Inc 2024 ArcGIS Pro version 3.4.0. (available at: https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview)

European Space Agency 2017 Land cover CCI product user guide version 2 (available at: maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf)

Fang Z, Zhang W, Brandt M, Abdi A M and Fensholt R 2022 Globally increasing atmospheric aridity over the 21st century *Earth's Future* **10** e2022EF003019

Feng Y, Ziegler A D, Elsen P R, Liu Y, He X, Spracklen D V, Holden J, Jiang X, Zheng C and Zeng Z 2021 Upward expansion and acceleration of forest clearance in the mountains of Southeast Asia *Nat. Sustain.* **4** 892–9

Giglio L and Justice C NASA MODIS Science Data Support Team 2021 MODIS Collection 6 Global Monthly Fire Location Product (MCD14ML) (available at: https://earthdata.nasa.gov/firms)

Giglio L, Schroeder W and Hall J V 2021 MODIS Collection 6 and Collection 6.1 Active Fire Product User's Guide

Giglio L, Schroeder W and Justice C O 2016 The collection 6 MODIS active fire detection algorithm and fire products Remote Sensing of Environment **178** 31—-41

Glauber A J, Moyer S, Adriani M and Gunawan I 2016 *The Cost of Fire: An Economic Analysis of Indonesia's 2015 Fire Crisis* the World Bank, Jakarta (https://doi.org/10.1596/23840)

Goldammer J G 2007 History of equatorial vegetation fires and fire research in Southeast Asia before the 1997–98 episode: a reconstruction of creeping environmental changes *Mitigation Adapt. Strateg. Glob. Change* **12** 13–32

Graham L L B, Applegate G B, Thomas A, Ryan K C, Saharjo B H and Cochrane M A 2022 A field study of tropical peat fire behaviour and associated carbon emissions *Fire* **5** 62

Hamid E Y, Kawasaki Z and Mardiana R 2001 Impact of the 1997–98 El Niño event on lightning activity over Indonesia *Geophys. Res. Lett.* **28** 147–50

Hanberry B B 2024 Practical guide for retaining correlated climate variables and unthinned samples in species distribution modeling, using random forests *Ecol. Inf.* **79** 102406

Hijmans R J 2025 terra: spatial data analysis (available at: https://CRAN.R-project.org/package=terra)

Huijnen V, Wooster M J, Kaiser J W, Gaveau D L A, Flemming J, Parrington M, Inness A, Murdiyarso D, Main B and Van Weele M 2016 Fire carbon emissions over maritime southeast Asia in 2015 largest since 1997 *Sci. Rep.* **6** 26886

Juárez-Orozco S M, Siebe C, Fernández Y and Fernández D 2017 Causes and effects of forest fires in tropical rainforests: a bibliometric approach *Tropical Conserv. Sci.* **10** 1940082917737207

Kaplan J O and Lau K H-K 2021 The WGLC global gridded lightning climatology and time series *Earth Syst. Sci. Data* **13** 3219–37

Kaveh N, Ebrahimi A and Asadi E 2023 Comparative analysis of random forest, exploratory regression, and structural equation modeling for screening key environmental variables in evaluating rangeland above-ground biomass *Ecol. Inf.* **77** 102251

König D 1990 *Theory of Finite and Infinite Graphs* (Birkhäuser Boston) (https://doi.org/10.1007/978-1-4684-8971-2)

Koplitz S N *et al* 2016 Public health impacts of the severe haze in equatorial Asia in September–October 2015: demonstration of a new framework for informing fire management strategies to reduce downwind smoke exposure *Environ. Res. Lett.* **11** 094023

Krasovskii A, Khabarov N, Pirker J, Kraxner F, Yowargana P, Schepaschenko D and Obersteiner M 2018 Modeling burned areas in Indonesia: the FLAM approach *Forests* **9** 437

Kuznetsova A, Brockhoff P B and Christensen R H B 2017 **lmerTest** package: tests in linear mixed effects models *J. Stat. Soft* **82** 1–26

Lam T, Catto J L, Kay G, Dunstone N, Barciela R and Harper A B 2026 Major fires in Indonesian Borneo are possible under all ENSO phases *npj Nat. Hazards* **3** 46

Lin C K 2021 Afterlives of authority: an ethnography of fire prediction, social order, and technocracy in Indonesia (University of Michigan) (available at: http://deepblue.lib.umich.edu/handle/2027.42/169759)

Liu T, Mickley L J, Marlier M E, DeFries R S, Khan M F, Latif M T and Karambelas A 2020 Diagnosing spatial biases and uncertainties in global fire emissions inventories: Indonesia as regional case study *Remote Sens. Environ.* **237** 111557

Meadows H D and Wright D 2011 *Thinking in Systems: A Primer* (Chelsea Green Publishing)

Meyfroidt P *et al* 2024 Explaining the emergence of land-use frontiers *R. Soc. Open Sci.* **11** 240295

Mu H, Li X, Wen Y, Huang J, Du P, Su W, Miao S and Geng M 2022 A global record of annual terrestrial human footprint dataset from 2000 to 2018 *Sci. Data* **9** 176

NASA JPL 2020 NASADEM merged DEM global 1 arc second V001 (available at: https://lpdaac.usgs.gov/products/nasadem_hgtv001)

NASA VIIRS Land Science Team 2020 VIIRS (S-NPP) I band 375 m active fire product NRT (vector data) (available at: www.earthdata.nasa.gov/data/catalog/lancemodis-vnp14imgtdl-nrt-2)

Novick K A *et al* 2024 The impacts of rising vapour pressure deficit in natural and managed ecosystems *Plant Cell Environ.* **47** 3561–89

Page S E and Hooijer A 2016 In the line of fire : the peatlands of Southeast Asia *Phil. Trans. B* **371** 20150176

Pekel J-F, Cottam A, Gorelick N and Belward A S 2016 High-resolution mapping of global surface water and its long-term changes *Nature* **540** 418–22

Pérez-Invernón F J, Gordillo-Vázquez F J, Huntrieser H and Jöckel P 2023 Variation of lightning-ignited wildfire patterns under climate change *Nat. Commun.* **14** 739

R Core Team 2024 R: a language and environment for statistical computing (available at: www.r-project.org/)

Robertson B L, Davies P, Gansell O, Van Dam-Bates P and McDonald T 2025 spbal: an R package for spatially balanced master sampling *Aust New Zealand J. Stat.* **67** 320–36

Rochmyaningsih D 2020 Wildfire researcher deported amid growing rift between Indonesian government and scientists *Science* **367** 722–3

Seager R, Hooks A, Williams A P, Cook B, Nakamura J and Henderson N 2015 Climatology, variability, and trends in the U.S. vapour pressure deficit, an important fire-related meteorological quantity *J. Appl. Meteorol. Climatol.* **54** 1121–41

SiPongi M of E and F Republic of Indonesia 2015 Indikasi luas kebakaran-rekapitulasi luas kebakaran Hutan dan Lahan (Ha) per provinsi di Indonesia (available at: https://sipongi.menlhk.go.id/indikasi-luas-kebakaran)

Staal A, Fetzer I, Wang-Erlandsson L, Bosmans J H C, Dekker S C, Van Nes E H, Rockström J and Tuinenburg O A 2020 Hysteresis of tropical forests in the 21st century *Nat. Commun.* **11** 4978

Stott P A, Goldammer J G and Werner W L 1990 The role of fire in the tropical lowland deciduous forests of Asia *Fire in the Tropical Biota Ecological Studies* vol 84, ed J G Goldammer (Springer) pp 32–44

Sumarga E 2017 Spatial indicators for human activities may explain the 2015 fire hotspot distribution in central Kalimantan Indonesia *Tropical Conserv. Sci.* **10** 1940082917706168

Tacconi L 2003 *Fires in Indonesia: Causes, Costs and Policy Implications* (Center for International Forestry Research (CIFOR))

Tacconi L 2016 Preventing fires and haze in Southeast Asia *Nat. Clim. Change* **6** 640–3

Tan Z D, Carrasco L R and Taylor D 2020 Spatial correlates of forest and land fires in Indonesia *Int. J. Wildland Fire* **29** 1088

Taylor D 2010 Biomass burning, humans and climate change in Southeast Asia *Biodivers. Conserv.* **19** 1025–42

Teo H C, Lechner A M, Sagala S and Campos-Arceiz A 2020 Environmental impacts of planned capitals and lessons for Indonesia's new capital *Land* **9** 438

Thirumalai K, DiNezio P N, Okumura Y and Deser C 2017 Extreme temperatures in Southeast Asia caused by El Niño and worsened by global warming *Nat. Commun.* **8** 15531

Thuiller W, Georges D, Gueguen M, Engler R, Breiner F, Lafourcade B, Patin R and Blancheteau H 2024 biomod2: ensemble platform for species distribution modeling (available at: https://biomodhub.github.io/biomod2/)

Trang P T, Andrew M E, Chu T and Enright N J 2022 Forest fire and its key drivers in the tropical forests of northern Vietnam *Int. J. Wildland Fire* **31** 213–29

Valavi R, Guillera-Arroita G, Lahoz-Monfort J J and Elith J 2022 Predictive performance of presence-only species distribution models: a benchmark study with reproducible code *Ecol. Monogr.* **92** e01486

Varma A 2003 The economics of slash and burn: a case study of the 1997–1998 Indonesian forest fires *Ecol. Econ.* **46** 159–71

Vayda A P 2006 Causal explanation of indonesian forest fires: concepts, applications, and research priorities *Hum. Ecol.* **34** 615–35

Wee S J, Park E, Alcantara E and Lee J S H 2024 Exploring multi-driver influences on Indonesia's biomass fire patterns from 2002 to 2019 through geographically weighted regression *J. Geovis. Spat. Anal.* **8** 4

Weiss D J *et al* 2018 A global map of travel time to cities to assess inequalities in accessibility in 2015 *Nature* **553** 333–6

Wickham H, François R, Henry L, Müller K and Vaughan D 2023 dplyr: a GRAMMAR OF DATA MANIPULATION (available at: https://CRAN.R-project.org/package=dplyr)

Widyastuti K, Imron M A, Pradopo S T, Suryatmojo H, Sopha B M, Spessa A and Berger U 2021 PeatFire: an agent-based model to simulate fire ignition and spreading in a tropical peatland ecosystem *Int. J. Wildland Fire* **30** 71–89

Wooster M J, Perry G L W and Zoumas A 2012 Fire, drought and El Niño relationships on Borneo (Southeast Asia) in the pre-MODIS era (1980–2000) *Biogeosciences* **9** 317–40

Zaveri E, Russ J and Damania R 2020 Rainfall anomalies are a significant driver of cropland expansion *Proc. Natl Acad. Sci. USA* **117** 10225–33