

A statistical global ~~burned~~burnt area model ~~tailored for seamless~~ integration into Dynamic Global Vegetation Models

Blessing Kavhu^{1,3}, Matthew Forrest¹, Thomas Hickler^{1,2}

¹Senckenberg Biodiversity and Climate Research Centre (SBIK-F), Frankfurt am Main, Germany.

²Institute of Physical Geography, Goethe University Frankfurt am Main, Frankfurt am Main, Germany

³Environmental Studies, University of California, Santa Cruz, 1156 High 5th, Santa Cruz, 95064, California, United States.

Correspondence to: Blessing Kavhu (kavhublessing@gmail.com)(kavhublessing@gmail.com)

Abstract. Fire-enabled Dynamic Global Vegetation Models (DGVMs) play an essential role in predicting vegetation dynamics and biogeochemical cycles amid climate change. ~~Modeling, but modelling~~ wildfires has been challenging in process-based biophysics-oriented DGVMs, ~~as human behaviour plays a crucial regarding the role. This of socioeconomic drivers. In this study aims, we aimed to reveal~~build a ~~simple~~ global statistical model ~~for the relationships between biophysical and that incorporates~~ socioeconomic drivers of wildfire dynamics ~~and, together with biophysical drivers, tailored for integration within DGVMs. Using monthly burned~~burnt area (BA) ~~that can be integrated into DGVMs. We data form the latest global burned area product from GFED5 as our response variable, we~~ developed ~~Generalised~~Generalized Linear Models (GLMs) to capture the relationships between potential predictor variables (~~biophysical and socio-economic~~) that are simulated by DGVMs and/or available in future scenarios ~~and the latest global burned area product from GFED5. Predictor variables were chosen to, We used predictors that~~ represent aspects of fire weather, vegetation structure and activity, human land use and ~~behaviour~~behavior and topography. ~~The final~~Based on an iterative process of choosing various variable combinations that represent potential key drivers of wildfires, we chose a model ~~was chosen by minimizing with minimum~~ collinearity and ~~by maximizing~~maximum model performance in terms of reproducing observations. ~~The final~~Our results show that the best performing (deviance explained 56.8%) and yet parsimonious model ~~included~~includes eight ~~socio-economic and biophysical~~ predictor variables encompassing the Fire Weather Index (FWI), a ~~novel~~Gross PrimaryMonthly Ecosystem Productivity Index (GPPIMEPI), Human Development Index (HDI), Population Density (PPN), Percentage Tree Cover (PTC), Percentage Non-Tree Cover (PNTC), Number of Dry Days (NDD), and Topographic Positioning Index (TPI). ~~Given its simplicity, our model demonstrated a remarkable capability, explaining 56.8% of~~When keeping the burnt area variability, comparable to other state-of-the-art global fire models: variables constant (partial residual plots). FWI, PTC, TPI and PNTC were positively related to BA, while GPPIMEPI, HDI, PPN, and NDD were negatively related to ~~wildfire~~BA. While the model effectively predicted the spatial distribution of ~~burned areas~~BA (Normalized Mean Error [NME] = 0.72), its standout performance lay in capturing the seasonal variability, especially in regions often characterized by distinct wet and dry seasons, notably southern Africa. ($R^2 =$

Formatted: Line spacing: single

Formatted: Font: 10 pt, Not Bold, Font color: Auto

Formatted: Underline, Font color: Blue

0.72 to 0.99), Australia ($R^2 = 0.68$) and parts of South America ($R^2 = 0.5975$ to 0.90). Our model reveals the robust predictive power of fire weather and vegetation dynamics emerging as reliable predictors of these seasonal global fire patterns. Finally, simulations with and without dynamically changing HDI revealed HDI as an important driver of the observed global decline in BA. The model presented model should be compatible with most, if not all, DGVMs used to develop future scenarios.

1-Introduction

Globally, the impacts of climate change continue to manifest through extreme weather events and changes in weather patterns (Clarke et al., 2019). Notably, climate change has led to more severe fire weather in large parts of the world and record fires have recently occurred in Australia and Canada, each burning more than 15 million ha (Barnes et al., 2023; Jain et al., 2024). Even though the effects of fires may be positive through contributing to selected natural ecosystem processes, large and frequent fires are often destructive and have far-reaching effects through loss of life, biodiversity, landscape aesthetic value, and increase in forest fragmentation and soil erosion (Bowman et al., 2017; Knorr et al., 2016; Nolan et al., 2022). The negative role of climate change in driving large and frequent burning has been well documented (Brown et al., 2023). However, climate change by itself does not fully account for the recent changes in global wildfire patterns as human activities are crucial drivers as well (Pausas and Keeley, 2021). For instance, recent empirical investigations have highlighted a notable 25% reduction in burnt area extent over the past two decades, explicitly attributing this decline to human activities (Andela et al., 2017). Wu et al. (2021) argue that future demographic and climate patterns will cause an increase in burned areas, particularly in high latitude warming and tropical regions. However, Knorr et al. (2016) concludes that under a moderate emissions scenario, global burned areas will continue to decline, but they will begin to rise again around mid-century with high greenhouse gas emissions. Cunningham et al. (2024), on the other hand reported that although total burned area is declining globally, extreme fire events are increasing as consequence of climate change especially in boreal and temperate conifer biomes. Future global fire dynamics are clearly driven by the overarching interaction between human activities (altered ignition patterns, surveillance and management) and climate (Krawchuk et al., 2009). Accurately evaluating these factors through modeling could guide prescribing solutions that will ensure reliable fire management and attainment of Sustainable Development Goals (SDGs) (Koubi, 2019; Robinne et al., 2018).

Modeling continues to be an essential tool for comprehending and forecasting wildfire dynamics, founded on the intricate interplay among fire weather, vegetation, and human activities (Bistinas et al., 2014; Hantson et al., 2016). Models for wildfire can be process-based or statistical. While process-based models delve into the physics and dynamics of wildfires and

vegetation, statistical models, on the other hand, tend to focus on analyzing historical data and identifying correlations to predict future wildfire events (Morvan, 2011; Xi et al., 2019). Process-based models such as fire-enabled DGVMs stand out in understanding interactions between climate, vegetation, and human activities in a mechanistic manner (Hantson et al., 2016; Rabin et al., 2017). However, their predictive skill is often not yet satisfactory (Hantson et al., 2020). One of their greatest limitations lies in representing the often-dominating effects of humans on fire ignitions, fire spread, and fire suppression in a mechanistic process-based way as this might be elusive (Archibald, 2016; DeWilde and Chapin, 2006; Hantson et al., 2022). Hence statistical approaches have often been used to evaluate human impacts on wildfires, in combination with weather and vegetation drivers (Haas et al., 2022; Kuhn-Régnier et al., 2021). Besides, some authors reported that the application of statistical models for ecosystems other than the ones used in their derivation is often not reliable (e.g. Perry, 1998). Integration of mechanistic process-based techniques and statistical methods remains a lasting solution to advance our understanding of fire dynamics.

Global wildfire modeling offers a macroscopic perspective, allowing researchers to analyze large-scale patterns across diverse ecosystems (Doerr and Santin, 2016; Flannigan et al., 2009). The strength of modeling fires at a global scale lies in its ability to capture overarching patterns (spatial, seasonal and inter-annual) that might provide valuable insights for strategic wildfire control. While one can argue about the potential oversimplification of local factors and the challenges in representing fine-scale heterogeneity, global models do, on the other hand, excel in capturing and understanding the effect of climate change, partly because they capture large climatic gradients (Robinne et al., 2018). The ability to capture the interconnectedness of ecosystems and fire regimes on a planetary scale contributes to a more holistic approach to understand global vegetation dynamics and carbon cycling (Bowman et al., 2013; Kelly et al., 2023). As such, studies on evaluating drivers of burnt areas at a global scale in the face of ongoing climatic shifts are crucial in ensuring sustainable management of vulnerable ecosystems.

There is a growing recognition of the significance of exploring both inter-annual and seasonal variations to comprehensively understand the dynamics of fire across diverse ecosystems (Dwyer et al., 2000).

1 Introduction

Globally, the impacts of climate change continue to manifest through extreme weather events and changes in weather patterns (Clarke et al., 2019). Notably, climate change has led to more severe fire weather in large parts of the world and record fires have recently occurred in Australia and Canada, burning more than 15 million and 7 million ha (Jain et al., 2024; MacCarthy et al., 2024). Even though the effects of fires may be positive through contributing to selected natural ecosystem processes, large and frequent fires are often destructive and have far-reaching effects through loss of life, biodiversity, landscape aesthetic

value, and increase in forest fragmentation and soil erosion (Bowman et al., 2017; Knorr et al., 2016; Nolan et al., 2022). The negative role of climate change in driving large and frequent burning has been well documented (Brown et al., 2023). However, climate change by itself does not fully account for the recent changes in global wildfire patterns as human activities are crucial drivers as well (Pausas and Keeley, 2021). For instance, recent empirical investigations have highlighted a notable 25% reduction in burnt area extent over the past two decades, explicitly attributing this decline to human activities (Andela et al., 2017). Wu et al. (2021) argue that future demographic and climate patterns will cause an increase in burnt areas, particularly in high latitude warming and tropical regions. However, Knorr et al. (2016) concluded that, under a moderate emissions scenario, global burnt areas will continue to decline, but they will begin to rise again around mid-century with high greenhouse gas emissions. Cunningham et al. (2024), on the other hand reported that although total burnt area is declining globally, extreme fire events are increasing as consequence of climate change especially in boreal and temperate conifer biomes. Future global fire dynamics are clearly driven by the overarching interaction between human activities (altered ignition patterns, surveillance and management) and climate (Krawchuk et al., 2009). Accurately evaluating these factors through modelling could guide prescribing solutions that will ensure reliable fire management and attainment of Sustainable Development Goals (SDGs) (Koubi, 2019; Robinne et al., 2018).

Modelling continues to be an essential tool for comprehending and forecasting wildfire dynamics, founded on the intricate interplay among fire weather, vegetation, and human activities (Bistinas et al., 2014; Hantson et al., 2016). Models for wildfire can be process-based or statistical. While process-based models delve into the physics and dynamics of wildfires and vegetation, statistical models, on the other hand, tend to focus on analyzing historical data and identifying correlations to predict future wildfire events (Morvan, 2011; Xi et al., 2019). Process-based models such as fire-enabled DGVMs stand out in understanding interactions between climate, vegetation, and human activities in a mechanistic manner (Hantson et al., 2016; Rabin et al., 2017). However, their predictive skill is often not yet satisfactory (Hantson et al., 2020). The predictive skill of process-based models is often limited due to incomplete representation of fire drivers, uncertainty in parameterization, and difficulties in accurately simulating human-fire interactions. One of their greatest limitations lies in representing the often-dominating effects of humans on fire ignitions, fire spread, and fire suppression in a mechanistic process-based way as this might be elusive (Archibald, 2016; DeWilde and Chapin, 2006; Hantson et al., 2020). Hence statistical approaches have often been used to evaluate human impacts on wildfires, in combination with weather and vegetation drivers (Haas et al., 2022; Kuhn-Régnier et al., 2021). Statistical approaches can effectively quantify and evaluate empirical relationships between fire occurrences and diverse predictors, providing flexibility in handling diverse data from multiple spatial and temporal scales. However, some authors reported that the application of statistical models for ecosystems other than the ones used in their derivation is often not reliable (e.g Perry, 1998). This is mainly because statistical models assume that the relationship between predictors and responses is stationery and context dependent, which is not typical of fires that are stochastic in nature.

127 Integration of mechanistic process-based techniques and statistical methods remains one common way forward to advance our
128 understanding of fire dynamics.

129

130 The integration of DGVMs and statistical models increasingly benefits from remote sensing data (Dantas De Paula et al.,
131 2020). Remote sensing provides spatially explicit observations, such as vegetation cover, leaf area index (LAI), and biomass,
132 which are used to initialize, calibrate, and validate DGVM simulations (Yang et al., 2020). Meanwhile, statistical models help
133 correct biases in DGVM outputs and enhance predictions by combining empirical relationships with mechanistic model results.
134 This integration enables more reliable modelling of global wildfires, offering a macroscopic perspective, and allowing
135 researchers to analyze large-scale patterns across diverse ecosystems (Doerr and Santín, 2016; Flannigan et al., 2009). The
136 strength of modelling fires at a global scale lies in its ability to capture overarching patterns (spatial, seasonal and inter-annual)
137 that might provide valuable insights for strategic wildfire control. While one can argue about the potential oversimplification
138 of local factors and the challenges in representing fine-scale heterogeneity, global models do, on the other hand, excel in
139 capturing and understanding the effect of climate change, partly because they capture large climatic gradients (Robinne et al.,
140 2018). The ability to capture the interconnectedness of ecosystems and fire regimes on a planetary scale contributes to a more
141 holistic approach to understanding global vegetation dynamics and carbon cycling (Bowman et al., 2013; Kelly et al., 2023).
142 As such, studies on evaluating drivers of burnt areas at a global scale in the face of ongoing climatic shifts are crucial in
143 ensuring sustainable management of vulnerable ecosystems.

144

145 There is a growing recognition of the significance of exploring both inter-annual and seasonal variations to comprehensively
146 understand the dynamics of fire across diverse ecosystems (Dwyer et al., 2000), partly because of the strong seasonal dynamics
147 of vegetation. Also, understanding seasonal cycles of fires helps to identify peak fire seasons, regions prone to seasonal
148 outbreaks, potential shifts in fire regimes over time and facilitating adaptive management strategies (Carmona-Moreno et al.,
149 2005). Incorporating monthly data in global fire modeling (Carmona-Moreno et al., 2005). Incorporating monthly data in global
150 fire modelling helps researchers to accurately capture seasonal variations in fire activity. Hence, global models developed
151 using monthly data are necessary.

152

153 Recent efforts have seen global burned area models based on diverse datasets and statistical approaches such as Convolution
154 Neural Network (CNN) (Bergado et al., 2021), Random Forest (RF) and generalized additive models (GAM) (Chuvieco et al.,
155 2021). However, the integration of these techniques into DGVM is yet to be realized. Haas et al. (2022)) developed a statistical
156 global model for burned area using a GLM, however, their model does only simulate annual dynamics, not seasonal patterns.
157 Generally, most earlier fire modules in DGVMs were informally parameterised models and do not consider the fuller range of
158 predictors available in a more rigorous statistical framework (Fosberg et al., 1999; Pfeiffer et al., 2013). This left an opportunity
159 to improve burned area models in DGVMs to accurately represent the detailed seasonal dynamics. To our knowledge, there

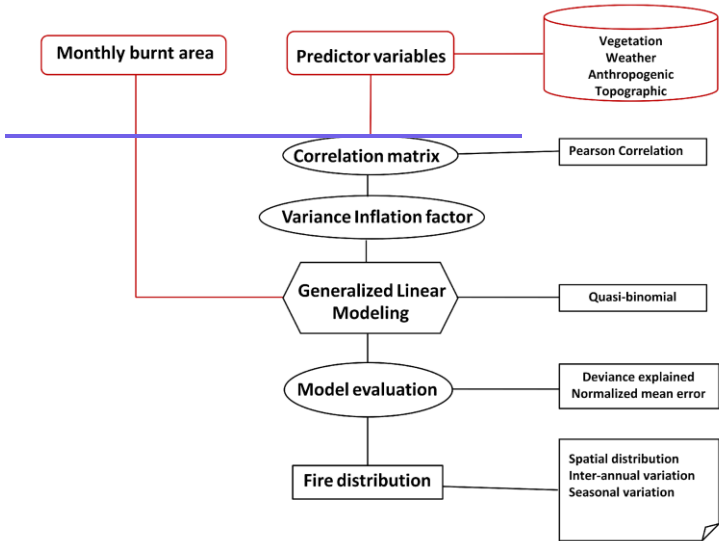
haven't been any reports on a simpler and more efficient statistical model specifically crafted to capture the seasonal cycles of global burnt areas, while also being easily integrated into DGVMs. Closing this gap can best be facilitated by using up-to-date remote sensing datasets pertinent to fire modeling. This integration can efficiently enhance our comprehension of inadequately understood factors while leveraging the potential of finely detailed temporal resolution burnt area datasets using a DGVM-integrable statistical model.

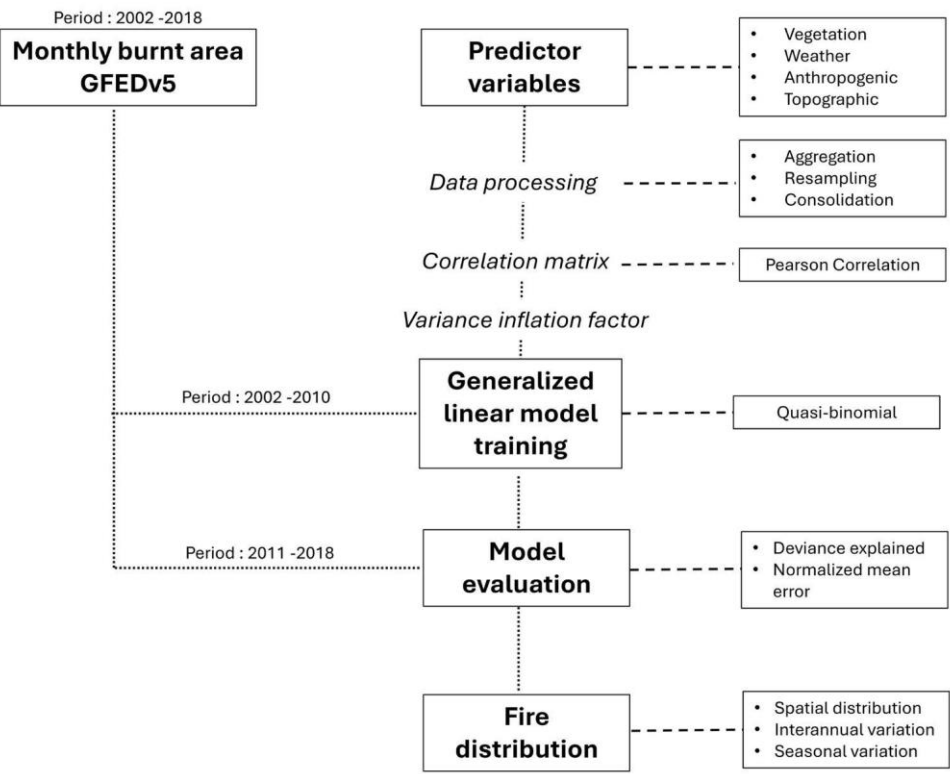
Recent efforts have seen global burnt area models based on Convolution Neural Network (CNN)(Bergado et al., 2021), Random Forest (RF) and generalized additive models (GAM)(Chuvieco et al., 2021) which are currently not easily integrated into DGVMs, although we note that recent work from Son et al. (2024) is an important step towards integration of an advanced recursive neural network in a DGVM. GLMs are easier to implement into DGVMs, and their partial residual plots show the relationships of each predictor with the response variable when all other drivers are kept constant, which facilitates a discussion of potential underlying mechanisms. The risk of overfitting can be minimized by only choosing potential driver variables that are mechanistically expected to play an important role and by choosing only a limited number of uncorrelated driver variables. Accordingly, Haas et al. (2022) developed a GLM for global burnt area with good model skill but without accounting for seasonal dynamics and without a focus on driver variables that can be predicted with DGVMs. Generally, most earlier fire modules in DGVMs such as the LPJ-LMfire(v1) were informally parameterized to predict seasonal fire cyclesmodels and do not consider the fuller range of predictors available in a more rigorous statistical framework (Fosberg et al., 1999; Pfeiffer et al., 2013). Nurrohman, et al., (2024) produced monthly fire predictions from downscaling of annual model outputs without building a statistical approach that is calibrated based on monthly inputs. This left an opportunity to improve burnt area models in DGVMs to accurately represent the detailed seasonal dynamics. To our knowledge, there haven't been any reports on a simpler and more efficient statistical model specifically crafted to capture the seasonal cycles of global burnt areas, while also being easily integrated into DGVMs. Closing this gap can best be facilitated by developing a statistical model based on variables pertinent to fire modelling, with the goal to later integrate it into DGVMs. This integration can efficiently enhance our comprehension of inadequately understood factors while leveraging the potential of finely detailed temporal resolution burnt area datasets.

The main aim of this research is to build a parsimonious statistical model for global seasonal ~~burned~~burnt areas that can be integrated into a DGVM. The specific objectives are to 1) to improve our understanding of major drivers of global ~~burned~~burnt area dynamics, 2) to leverage a GLM for predicting global burnt areas using DGVM-integrable predictors and 3) to evaluate the interannual and seasonal cycles of burnt area extent, both globally and regionally.

191 2 Data and Methods

192 In this study, we used GLM to assess the drivers and distribution of global wildfires based on a combination of vegetation,
193 weather, anthropogenic and topographic predictors. The spatial and temporal variability (interannual and seasonality) was also
194 evaluated. Fig. 1 provides an overview of the steps that were followed during modelingmodelling.
195





Formatted: Line spacing: single

Figure 1: Study workflow showing an overview of steps followed in model calibration and evaluation together with the outputs.

2.1 Fire data

Monthly BA data for the period 2002 and 2018 were derived from monthly mean fractional BA from the GFEDv5. GFED-5 data are selected because of their improved ability to detect burnt area scars (Chen et al., 2023). Monthly BA data for the periods 2002 and 2018 were derived from monthly mean fractional BA from the GFED5. We selected this data because of

206 [their improved ability to detect burnt area scars \(Chen et al., 2023\)](#). GFED5 BA data are classified according to 17 major land
207 cover types using the MODIS classification scheme. We used this land cover information to remove burnt area in cropland
208 land cover type (type croplands and croplands/natural vegetation mosaic), to exclude the effect of cropland residue burning
209 which we suppose [is likely have](#) different drivers from burning in non-arable lands. [We used data for the period 2002 to](#)
210 [2010 for model training and data for 2011 to 2018 for model testing](#). BA data comes at a resolution of $0.25^{\circ} \times 0.25^{\circ}$, therefore
211 we aggregated it by a factor of 2 to a resolution of 0.5° . This was done for ease of processing at a global scale and at the same
212 time to ensure that our outputs are DGVM integrable since they are commonly applied at 0.5° globally.▲

Formatted: Line spacing: single

213 **2.2 Predictor variables**

214 [In this study, we only used variables which don't prohibit the use of the model for future projections](#). Whilst there are many
215 possible variables that could be tried as predictors of fire, especially in terms of socioeconomics predictors, we [here only use](#)
216 [variables which don't prohibit the use of the model for future projections](#). ~~These variables are~~restricted our selection to
217 [variables where](#): climate and vegetation variables typically available in a DGVM framework; socioeconomic variables with
218 future scenario projections; and time-invariant topographic variables. Previous studies used several variables that we couldn't
219 include due to lack of future scenario projections such as nighttime lights, cattle density ([Forkel et al., 2019a](#)), [Vegetation](#)
220 [optical depth \(Forkel et al., 2019b\)](#), [Lightning \(Rabin et al., 2017\)](#), [Soil moisture \(Mukunga et al., 2023\)](#), [soil fertility \(Aldersley](#)
221 [et al., 2011\)](#),([Forkel et al., 2019a](#)), [Vegetation optical depth \(Forkel et al., 2019b\)](#), [Lightning \(Rabin et al., 2017\)](#), [Soil moisture](#)
222 [\(Mukunga et al., 2023\)](#), [soil fertility \(Aldersley et al., 2011\)](#). Consequently, we considered predictor variables that are
223 compatible with DGVM integration to calibrate the model effectively. The chosen predictor variables were categorized based
224 on their representational nature and their roles in fire ~~modeling~~[modelling](#)([See Table 1](#)). ~~Table 1 provides a comprehensive~~
225 ~~overview of each variable category, including their sources and spatio-temporal resolutions.~~

Formatted: Font: Not Bold

Predictor	Abbreviations	Classification category (Climate, vegetation, landcover, landscape fragmentation, ignition, suppression topographic effect)	Original spatial resolution	Temporal resolution	Source
Percentage Grass cover	PGC	Vegetation	300m	Annual	ESA landcover CCI

Formatted: Line spacing: single

Percentage non-tree vegetation cover	PNTC	Vegetation	250m	Annual	MODIS - MOD44B (DiMiceli et al., 2011)
Topographic positioning index	TPI	Topography	90m	Static	digital elevation model products of global 250 m GMTED2010 and near-global 90 m SRTM4.1dev. Digital elevation model products of global 250 m GMTED2010 (GMTE data 2010) and near-global 90 m SRTM v4 (Jarvis et al., 2008)
Human Development Index	HDI	Ignition/suppression/fragmentation	subnational	Annual	Global data lab (Smits and Permanyer, 2019)
Road density	RD	Ignition/suppression/fragmentation	0.5° × 0.5°	Static	Global Roads Inventory Project (GRIP) database (Meijer et al., 2018)
Population density	PPN	Ignition/suppression/fragmentation	2.5 arc minutes	5-year intervals	Socioeconomic data and applications centre (SEDAC) Socioeconomic data and applications centre (SEDAC) (Klein Goldewijk et al., 2017)
Percentage crop cover	PCC	Fragmentation	5 arc minutes	Annual	HistorY Database of the Global Environment (HYDE 3.3HistorY Database of the Global

Formatted: Line spacing: single

					Environment (HYDE 3.3) (Klein Goldewijk et al., 2017)
Percentage pasture cover	PPS	Vegetation	5 arc minutes	Annual	History Database of the Global Environment (HYDE 3.3)History Database of the Global Environment (HYDE 3.3)(Klein Goldewijk et al., 2017)
Precipitation seasonality	PS	Climate	0.5° × 0.5°	Annual	Copernicus climate data store (Copernicus Climate Change Service 2018)
Fire weather index	FWI	Climate	0.5° × 0.5°	Monthly	Copernicus climate data store (Copernicus Climate Change Service 2018)
Precipitation of the driest quarter	PPNQ	Climate	0.5° × 0.5°	Annual	Copernicus climate data store (Copernicus Climate Change Service 2018)
Number of dry days	NDD	Climate	0.5° × 0.5°	Annual	Copernicus climate data store (Copernicus Climate Change Service 2018)
Percentage grazeland cover	PGZC	Vegetation	5 arc minutes	Annual	History Database of the Global Environment (HYDE 3.3)History Database of the Global Environment (HYDE 3.3) (Klein Goldewijk et al., 2017)

Percentage rangeland cover	PRC	Vegetation	5 arc minutes	Annual	History Database of the Global Environment (HYDE 3.3) History Database of the Global Environment (HYDE 3.3) (Klein Goldewijk et al., 2017)
Annual average precipitation	AAP	Climate	5 arc minutes	Annual	Copernicus climate data store (Copernicus Climate Change Service 2018)
Gross primary productivity	GPP	Vegetation	0.5° × 0.5°	Monthly	MOD17A1+MOD17A1 (Running and Zhao, 2019)
Aboveground biomass	AGB	Vegetation	0.5° × 0.5°	Longterm average	
Percentage Tree cover	PTC	Vegetation	250m	Annual	MODISMODIS - MOD44B (DiMiceli et al., 2011)
Fraction of Absorbed Photosynthetically Active Radiation	FAPAR	Vegetation	500m	Monthly (originally days)	8 MODISMODIS - MOD15A2H (Running and Zhao, 2019)

Table 1: List of predictor variables that were considered in this study including their classifications, resolution (spatial & temporal) and the respective data sources.

2.32.1 Vegetation-related predictors

We used eight vegetation predictor variables to comprehensively evaluate their role on global fire distribution. These variables encompass Percentage Grass Cover (PGC), Percentage Non-Tree Cover (PNTC), Percentage Crop Cover (PCC), Percentage Graze Cover (PGZC), Percentage Rangeland Cover (PRC), and Percentage Tree Cover (PTC). Previous work emphasizes the important role of vegetation on burnt area dynamics. For example, Thonicke et al. (2010), discussed the crucial role of vegetation structure in shaping fire occurrence, spread and intensity. PGC defines the land covered by grass, influencing fuel availability, while PNTC considers non-tree vegetation such as grass and shrubs, contributing to overall fuel dynamics. PCC reflects the presence of cultivated crops which have been found to suppress fire occurrence as they fragment the landscape acting and so act as a barrier to fire spread (Haas et al., 2022).

We used PGZC, PRC, PTNC and PTC to evaluate the relationship between landcover and burnt area distribution. Previous studies reported that landcover change has made a significant contribution to wildfire distribution (Gallardo et al., 2016; Villarreal and Vargas, 2021).

We used Gross Primary Productivity (GPP), Aboveground Biomass (AGB), and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) as proxies for vegetation health and productivity. Previous studies emphasized the varying effects of vegetation parameters on fire events (Bowman et al., 2020; Kuhn-Régnier et al., 2021).

Thonicke et al. (2010), Thonicke et al. (2010), for example, discussed the crucial role of vegetation structure in shaping fire occurrence, spread and intensity. Consequently, our study considered eight vegetation predictor variables to comprehensively evaluate their role on global fire distribution. These variables encompass Percentage Grass Cover (PGC), Percentage Non-Tree Cover (PNTC), Percentage Crop Cover (PCC), Percentage Graze Cover (PGZC), Percentage Rangeland Cover (PRC), and Percentage Tree Cover (PTC).

PGC defines the land covered by grass, influencing fuel availability, while PNTC considers non-tree vegetation such as grass and shrubs, contributing to overall fuel dynamics. PCC reflects the presence of cultivated crops which have been found to suppress fire occurrence as they fragment the landscape acting and so act as a barrier to fire spread (Haas et al., 2022). Previous studies reported that landcover change has a significant contribution to wildfire distribution (Gallardo et al., 2016; Vilar et al., 2021). (Haas et al., 2022). Previous studies reported that landcover change has made a significant contribution to wildfire distribution (Gallardo et al., 2016; Villarreal and Vargas, 2021). To understand the relationship between landcover and burnt area distribution, we incorporate PGZC, PRC, PTNC and PTC.

Numerous studies discussed the varying effects of vegetation parameters on fire events (Bowman et al., 2020; Kuhn-Régnier et al., 2021). Accordingly, Gross Primary Productivity (GPP), Aboveground Biomass (AGB), and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) were considered in this study as proxies for vegetation health and productivity. Numerous studies discussed the varying effects of vegetation parameters on fire events (Bowman et al., 2020; Kuhn-Régnier et al., 2021). Accordingly, Gross Primary Productivity (GPP), Aboveground Biomass (AGB), and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) were considered in this study proxies for vegetation health and productivity.

2.4 Vegetation2.2 Topographic-related predictors

We used topographic positioning index (TPI) to evaluate how topography can influence the occurrence and spread of fires. Topography has been reported to be more influential in regions with complex terrain and microclimatic conditions (Blouin et al., 2016; Fang et al., 2015; Oliveira et al., 2014). Some studies used slope (Cary et al., 2006) and surface area ratio (Parisien et al., 2011) in their models and reported topography to marginally contribute to wildfire dynamics. However, recent studies

273 reported some significant contributions of topography to global burnt area distribution when using the TPI (Haas et al., 2022).
274 TPI is designed to encompass and evaluate the complex influence of terrain features, such as elevation and slope, on the
275 distribution of burnt areas. Thus, TPI goes beyond simplistic representations of landscapes and offers a more nuanced
276 perspective on how terrain characteristics contribute to the occurrence and extent of wildfires. Given the role of terrain on fire
277 behavior and propagation patterns, the inclusion of TPI in this study allows for a comprehensive examination of wildfire
278 distribution.
279 Topography can influence the occurrence and spread of fires especially in regions with complex terrain and microclimatic
280 conditions (Blouin et al., 2016; Fang et al., 2015; Oliveira et al., 2014). (Blouin et al., 2016; Fang et al., 2015; Oliveira et al.,
281 2014). To capture the impact of topography, some studies used slope (Cary et al., 2006) (Cary et al., 2006) and surface area
282 ratio (Parisien et al., 2011) (Parisien et al., 2011) in their models and reported topography to marginally contribute to wildfire
283 dynamics. However, recent studies reported some significant contributions of topography to global burnt area distribution
284 when using the topographic positioning index (TPI) (Haas et al., 2022). (Haas et al., 2022). TPI is designed to encompass and
285 evaluate the complex influence of terrain features, such as elevation and slope, on the distribution of burned burnt areas. Thus,
286 TPI goes beyond simplistic representations of landscapes and offers a more nuanced perspective on how terrain characteristics
287 contribute to the occurrence and extent of wildfires. Given the role of terrain on fire behavior and propagation patterns, the
288 inclusion of TPI in this study allows for a comprehensive examination of wildfire distribution.

289 **2.52.3 Anthropogenic Influence Predictors**

290 We used the Human Development Index (HDI), Population Density (PPN), and Road Density (RD) to capture the impact of
291 anthropogenic factors on both fire ignition and suppression. The inclusion of HDI aims to encapsulate human influence on
292 ecological landscapes, thereby affecting the dynamics of both ignition and suppression processes. To comprehensively capture
293 the impact of anthropogenic factors on both fire ignition and suppression, our study integrates three key predictors: Human
294 Development Index (HDI), Population Density (PPN), and Road Density (RD). The inclusion of HDI aims to encapsulate
295 human influence on ecological landscapes, thereby affecting the dynamics of both ignition and suppression processes. HDI is
296 a composite index developed by the United Nations Development Program (UNDP) to assess long-term progress in three basic
297 dimensions of human development, including health (life expectancy at birth), education (mean years of schooling and
298 expected years of schooling), and standard of living (gross national income per capita) (Uddin, 2023). HDI values range from
299 0 to 1, with higher values indicating higher levels of human development. Although HDI itself may not directly relate to fire
300 occurrence, it stands as a valuable socio-economic indicator that significantly influences overall fire dynamics and
301 management, like how Gross Domestic Product (GDP) has been used in other fire models (Perkins et al., 2022). To address
302 the limitations of using GDP as a proxy for human development in predicting global fires, we opted for HDI. Previous research
303 has utilized GDP for this purpose (Zhang et al., 2023), (Zhang et al., 2023), however, GDP is an indicator of a country's
304 economic performance (Callen, 2008). In contrast, HDI is rather broad socioeconomic indicator, which we assume acts as a

proxy for factor such as investments and advancements in fire control methods, surveillance, technology, and outreach strategies increasing awareness. HDI data is much broader as it captures the country's social and economic development levels, making it a more suitable and consistent measure for our analysis. HDI evaluates a country or other administrative region's development status based on the critical factors of life expectancy, education, and income, providing a nuanced understanding of the socio-economic context shaping fire behavior (Teixeira et al., 2023). To evaluate model sensitivity to inclusion of HDI, we trained our model based on the three settings: including, excluding and holding HDI constant.

2.6.2.4 Weather-Related Predictors

To capture the impact of fire weather on the distribution of wildfires, we employed the Canadian Fire Weather Index (FWI), renowned for its comprehensive framework integrating diverse meteorological parameters to evaluate potential fire behavior and danger (de Jong et al., 2016). The FWI is widely adopted by fire management agencies facilitating informed decisions on fire prevention, preparedness, and suppression strategies, and global context has been shown to correlate well with burned area across the globe (Jones et al., 2022). We used the number of dry days (NDD) as a proxy for biomass production limitations. While it falls in the category of weather-related fire predictors, in this study it's an indirect indicator of how moisture availability can affect available combustible vegetation. We incorporated additional covariates capturing seasonal and annual weather dynamics that influence fires, including Precipitation Seasonality (PS), and Annual Average Precipitation (AAP). The selection of these predictors was informed by their significance in previous global fire modeling studies (Chuvieco et al., 2021; Joshi and Sukumar, 2021; Le Page et al., 2008; Mukunga et al., 2023; Saha et al., 2019), as well as insights from seminal works such as that by (Pechony and Shindell, 2010).

2. We employed the Canadian Fire Weather Index (FWI) to capture the impact of fire weather on the distribution of wildfires. FWI is renowned for its comprehensive framework integrating diverse meteorological parameters to evaluate potential fire behavior and danger. To capture the impact of fire weather on the distribution of wildfires, we employed the Canadian Fire Weather Index (FWI), renowned for its comprehensive framework integrating diverse meteorological parameters to evaluate potential fire behavior and danger (de Jong et al., 2016). The FWI is widely adopted by fire management agencies facilitating informed decisions on fire prevention, preparedness, and suppression strategies, and global context has been shown to correlate well with burnt areas across the globe (Jones et al., 2022). We used the number of dry days (NDD) as a proxy for biomass production limitations. While it falls in the category of weather-related fire predictors, in this study it's an indirect indicator of how moisture availability can affect available combustible vegetation. We incorporated additional covariates capturing seasonal and annual weather dynamics that influence fires, including Precipitation Seasonality (PS), and Annual Average Precipitation (AAP). The selection of these predictors was informed by their significance in previous global fire modelling studies (Chuvieco et al., 2021; Joshi and Sukumar, 2021; Le Page et al., 2015; Mukunga et al., 2023; Saha et al., 2019), as well as insights from seminal works such as that by Pechony and Shindell, (2010).

2.3 Data Processing

We harmonized the spatial and temporal resolution of the predictor dataset to conform to our analytical framework, which had a spatial resolution of 0.5° and a temporal resolution of one month. This involved employing techniques such as aggregation, resampling, and consolidation. For instance, while the native temporal resolution of FAPAR was 8 days, we transformed it into a monthly temporal resolution to align with our primary variable. Most predictors originally possessed an annual temporal resolution, except for FWI, GPP, and FAPAR, which were also available every month. For annual predictors, we replicated the same data for each month. Similarly, long-term variables like AGB, RD, and TPI were utilized every month to synchronize with the shorter-resolution predictors. PPN, which was available at a 5-year interval, was used monthly over the represented 5-year span. Kuhn-Régnier et al. (2021) Kuhn-Régnier et al. (2021) highlighted the important role of antecedent vegetation as key driver for global fires. To evaluate the role of fuel accumulation from the previous year on the burnt area, we derived the Gross Primary Productivity Index (GPPI) using monthly Gross Primary Productivity (GPP) data following Eq. (1). GPPI was originally defined as Monthly Ecosystem Productivity Index (MEPI) in the work by Forrest et al. (2024) Monthly Ecosystem Productivity Index (MEPI) using monthly Gross Primary Productivity (GPP) data following Eq. (1). MEPI was originally defined in the work by Forrest et al. (2024). This index allowed us to quantify the relationship between vegetation growth, fuel accumulation and subsequent fire activity, providing a more nuanced understanding of the factors influencing fire dynamics.

$$\text{GPP index MEPI} = \frac{GPP_m}{\max(GPP_m, GPP_{m-1}, \dots, GPP_{m-13})}$$

(1)

Where GPPm is the month’s GPP, and the denominator is the maximum GPP of the past 13 months. Furthermore, we calculated additional metrics including GPP12 (the mean gross primary productivity over the previous 12 months), (FAPAR12) (the mean fraction of absorbed photosynthetically active radiation over the past 12 months), and FAPAR6 (the mean FAPAR over the last 6 months). These metrics serve to capture average vegetation productivity, serving as refined indicators of fuel accumulation.

2.4 Statistical modeling and final predictor choice

To address variable collinearity, we conducted pairwise correlation analyses among predictor variables using the R statistical package (CoreTeam, 2014). Following established guidelines (Dormann et al., 2013), we applied the conventional threshold of R > 0.5 to enhance the model’s efficiency. This helped us identify and exclude correlated variables from the analysis. To address variable collinearity, we conducted pairwise correlation analyses among predictor variables using the R statistical

367 [package\(CoreTeam, 2014\)](#). Following established guidelines by Dormann et al. (2013) , we applied the conventional threshold
368 [of \$R > 0.5\$ to enhance the model's efficiency.](#)
369
370 ~~-Specifically, variables such as AGB, FAPAR12, FAPAR6, AAP, and RD were excluded due to their strong correlations with~~
371 ~~other variables (see Fig. 2). There were some correlated variables that were however returned in the model due to their~~
372 ~~significant contribution to fire modelling and model performance. For example NDD was strongly correlated to PTC (~~
373 ~~0.68), but keeping both increased the variance explained by the full model.~~
374

Formatted: Line spacing: single

Formatted: Line spacing: single

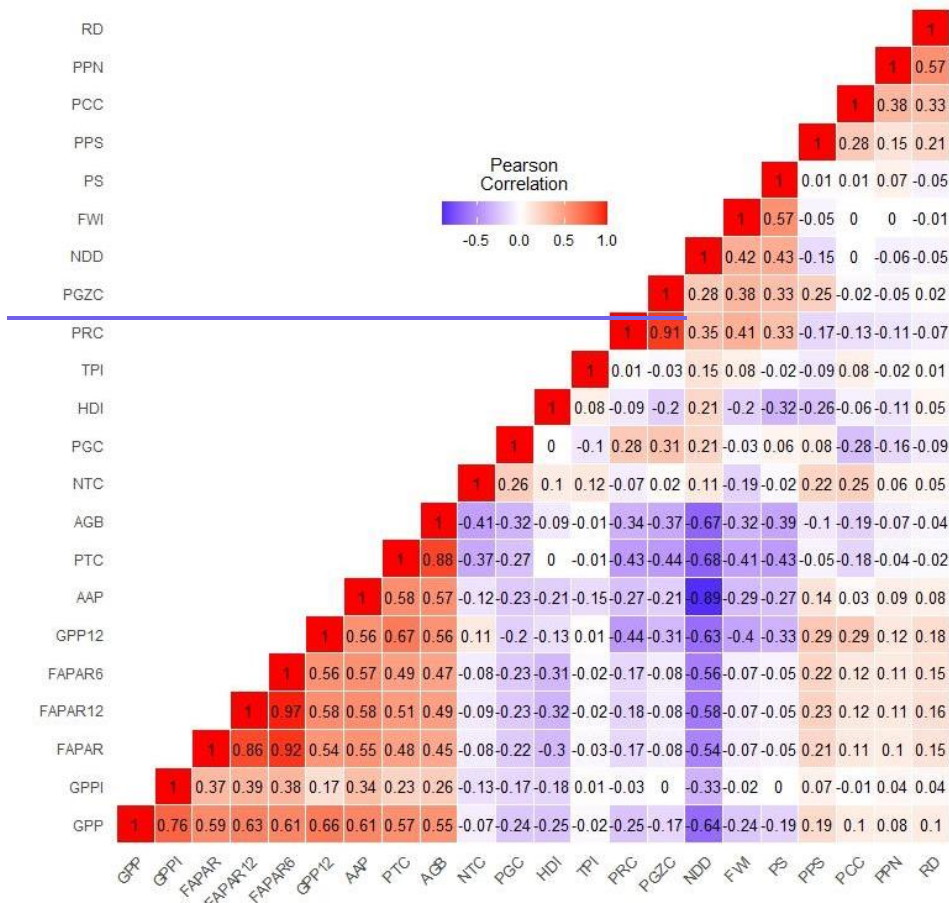


Figure 2: Correlation matrix of all the variables that were considered for modelling in this investigation

Moreover, we employed the Variance Inflation Factor (VIF) to evaluate collinearity among predictor variables, removing those with VIF values surpassing 5, as recommended by O'Brien. (2007)-O'Brien, (2007). Post collinearity tests, an additional 3 parameters were adopted to progressively select the best model, namely: 1) a simple (~ parsimonious) model which comprise

of a full suite of categories of covariate combinations (i.e vegetation, climate, topography, ignitions), 2) the deviance explained value and 3) the normalised mean square error value as illustrated in the making of Burnt Area Simulator For Europe (BASE) (Forrest et al., 2024). Fig A3 shows ten scatter plots of the final variable selection based on the optimum model, each depicting the relationship between the burnt area fraction and different environmental or socio-economic variables. The variables include the GPPi(Forrest et al., 2024). The variables include the MEPI, FWI, PNTC, HDI, PTC, TPI, NDD and PPN.

A quasi-binomial GLM was selected for modelingmodelling BA due to its capability to handle non-Gaussian error distributions, seamless integration into DGVMs and ability to generate partial residual plots (Bistinas et al., 2014; Haas et al., 2022; Lehsten et al., 2016), i.e. the effect of each predictor in the model while the others are held constant (Bistinas et al., 2014; Haas et al., 2022; Lehsten et al., 2016). Calibration of the model utilized data from 2002 to 2010 while testing utilized data from 2011 to 2018. Residual plots were utilized to examine the magnitude and nature of each predictor's relationship with wildfire burnt area distribution.

Model performance was assessed using the Normalized Mean Error (NME) following Kelley et al., (2013). Kelley et al. (2013). NME serves as a standardized metric for evaluating global fire model performance, facilitating direct comparison between predictions and observations. The NME was calculated following Eq. (2).

$$NME = \frac{\sum \frac{Ai \mid obs_i - sim_i \mid}{\sum Ai \mid obs_i - obs \mid}}{\sum} \quad (2)$$

The NME score was computed by summing the discrepancies between observations (obs) and simulations (sim) across all cells (i), weighted by the respective cell areas (Ai), and then normalized by the average distance from the mean of the observations. A lower NME value reflects superior model performance, with a value of 0 indicating a perfect alignment between observed and simulated values. BA fractions were treated as a probability ranging from 0 to 1, following a quasi-binomial distribution. We applied the logit link function based on the methodology outlined by Haas et al. (2022). We applied the logit link function based on the methodology outlined by Haas et al. (2022). After conducting a collinearity test, the models were systematically evaluated using various combinations of predictor variables. A total of 25 model runs were conducted, each incorporating different sets of variables while iteratively excluding some, to discern the extent to which each predictor explained variance when others were not included (see Table 2). A1). We followed the stepwise approach of variable inclusion, exclusion, interaction terms, log transformations, and polynomial transformations as described by Forrest et al. (2024). While their analysis focused on Europe, our objective was to replicate and apply the method at a global scale. To evaluate the reliability of the predicted interannual variability and seasonal cycles, we applied a regression function to determine the relationship (R²) between the observed and predicted trends. An R² of 1 shows good performance in our predictions and an R²

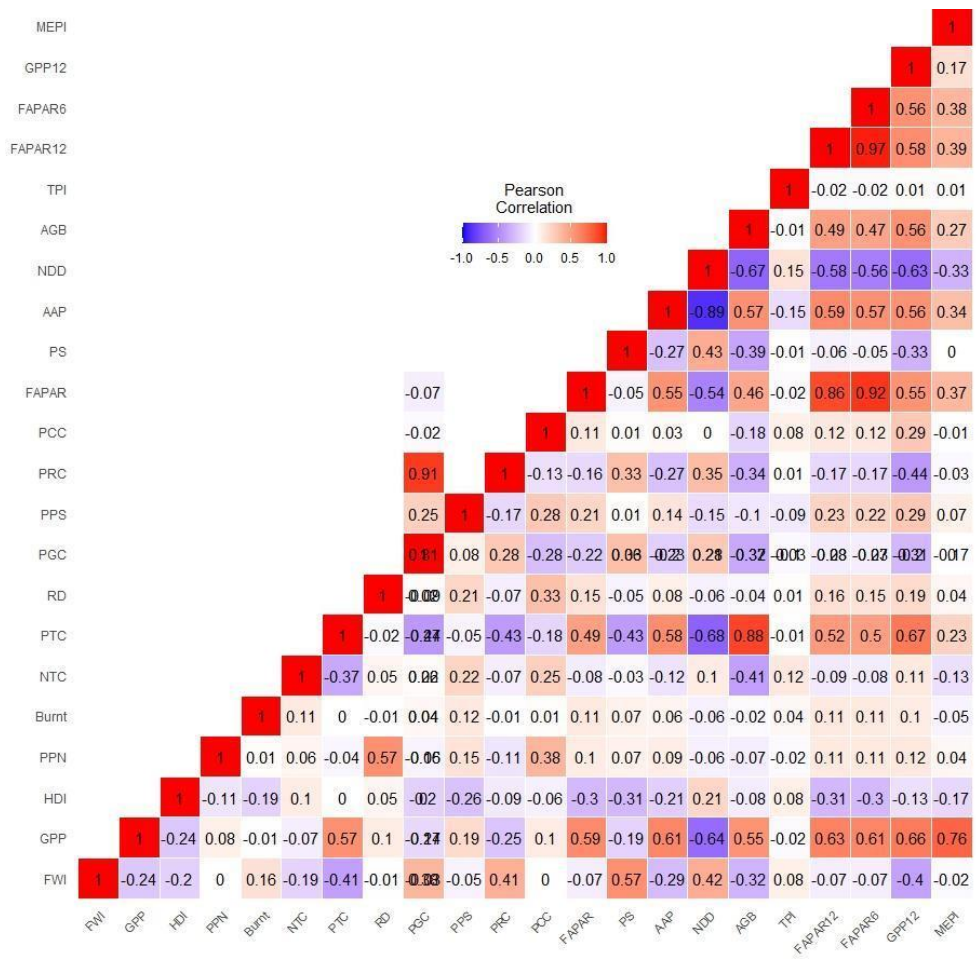
of 0 shows poor performance in our predictions. To assess the trend in predicted interannual variability, we used the Mann-Kendall test (Kendall, 1975; Mann, 1945). This widely used method detects monotonic trends in environmental data. Being non-parametric, it works for all distributions, does not require normality, but assumes no serial correlation. We extracted the trend test results and plotted a map of trend distributions across 14 different GFED regions to identify areas with significant predicted trends ($P<0.05$) from those with non-significant trends. The 14 GFED regions include, Boreal North America (BONA), Temperate North America (TENA), Central America (CEAM), Northern Africa (NHAF), Southern Africa (SHAF), Europe (EURO), the Middle East (MIDE), Equatorial Asia (EQAS), Southern Asia (SEAS), Boreal Asia (BOAS), Temperate Asia (TEAS), Australia and New Zealand (AUST), and Northern Hemisphere South America (NHSA), and Southern Hemisphere South America (SHSA). We extracted the trend test results and plotted a map of trend distributions across different GFED regions to identify areas with significant predicted trends ($P<0.05$) from those with non-significant trends.

3 Results

3.1 Correlation between variables

We found correlated variables that we had to exclude from the analysis. Specifically, variables such as AGB, FAPAR12, FAPAR6, AAP, and RD were excluded due to their strong correlations with other variables (see Fig. 2). There were however some variables that correlated but had to be returned to the model due to their significant contribution to fire modelling and model performance. For example, NDD was strongly correlated to PTC (~ -0.68), but both increased the variance explained by the full model.

432



433

434

435

Figure 2: Correlation matrix of all the variables that were considered for modelling in this investigation.

436

3.1 Optimal model selection and GLM results

437 Table 2 provides a list of results from our progressive inclusion of model covariates as we aimed to identify the optimum
438 model. The initial models (model 1 to model 3) progressively include more variables, however, a noticeable jump in deviance
439 explained when PNTC is added (Model 3: 0.5298). Models 4 to 8 involve adding vegetation (FAPAR) and various land use
440 types (PCC, PPS, PRC, PGC). This is accompanied by marginal improvement in deviance explained, indicating these factors
441 provide some additional predictive power but are not as impactful as existing vegetation covariates (such as GPP). Models 10
442 to 12 introduce polynomial terms for PTC. This results in an increase in deviance explained, peaking at around 0.558836 in
443 Model 12. Models 13 to 16 incorporate interactions between HDI and land use types (e.g., PCC and PRC), resulting in
444 marginal increase in deviance explained with the highest recorded in Model 15(~ 0.5664789). Models 19 to 25 fine-tune the
445 overall performance by incorporating various variables and their interactions. Model 24, which includes a comprehensive set
446 of climatic, vegetation, human, and topographic variables along with their interactions, achieves the highest deviance explained
447 (~0.5720048). The marginal improvements observed in subsequent models indicate that while additional variables contribute
448 to the model, the primary influencing factors were already identified by Model 19, however it was not the simplest model, (~
449 parsimonious), and comprised consisted of other variables that we don't have future projections for (e.g RD). We removed
450 some of the redundant variables till Model 24 (~11 variables), however, it was not as parsimonious as Model 25 (~8 variables).
451 Therefore, Model 25, which offers a balance of parsimony, simplicity, high deviance explained, and low NME, was selected
452 as the best model in this analysis.

453

454

Model	Formulae	Deviance explained	NME
model 1	-glm(burnt~ FWI + GPP + HDI + PTC + RD)	0.3548030	0.7472088
model 2	-glm(burnt~ FWI + GPP + HDI + PTC + RD + PGC)	0.3609393	0.7405652
model 3	-glm(burnt~ FWI + GPP + HDI + PTC + RD + PNTC)	0.5298061	0.7208771
model 4	-glm(burnt~ FWI + GPP + HDI + PTC + RD + PNTC + FAPAR)	0.5312036	0.7188448
model 5	-glm(burnt~ FWI + GPP + HDI + PTC + RD + PNTC + FAPAR + PCC)	0.5312697	0.7191269

Formatted: Line spacing: single

model 6	-glm(burnt ~ FWI + GPP + HDI + PTC + RD + PNTC + FAPAR + PCC + PPS)	0.5328183	0.7195616
model 7	-glm(burnt ~ FWI + GPP + HDI + PTC + RD + PNTC + FAPAR + PCC + PRC)	0.5313813	0.7193946
model 8	glm(burnt ~ FWI + GPP + HDI + PTC + RD + PNTC + FAPAR + PCC + PGC)	0.5349288	0.7190611
model 9	-glm(burnt ~ FWI + GPP + HDI + PTC + RD + PNTC + FAPAR + PCC + FAPAR12 + PGC)	0.5359802	0.7181930
model 10	-glm(burnt ~ FWI + GPP12 + HDI + poly(PTC, 2) + PNTC + FAPAR + PCC + FAPAR12 + PGC + PPN)	0.5295939	0.7172668

model 11	-glm(burnt ~ FWI + GPPI + HDI + poly(PTC, 2) + RD + PNTC + FAPAR12 + PGC + PS)	0.5579946	0.7193546
model 12	-glm(burnt ~ FWI + GPPI + HDI + poly(PTC, 2) + RD + PNTC + FAPAR12 + PS)	0.5571164	0.7192122
model 13	-glm(burnt ~ FWI + GPPI + HDI*PCC + PGC + RD + poly(PTC, 2) + PNTC + PS)	0.5569187	0.7214560
model 14	-glm(burnt ~ FWI + GPPI + HDI*PGC + RD + poly(PTC, 2) + PNTC + PS)	0.5570586	0.7222061
model 15	-glm(burnt ~ FWI + GPPI + HDI*PRC + RD + poly(PTC, 2) + PNTC + PS)	0.5664789	0.7154708
model 16	glm(burnt ~ FWI + GPPI*PNTC + HDI + RD + poly(PTC, 2) + PS)	0.5563012	0.7215202
model 17	-glm(burnt ~ FWI + GPPI + HDI + RD + poly(PTC, 2) + PNTC + PS + NDD)	0.5681926	0.7191069
model 18	-glm(burnt ~ FWI + GPPI + HDI + RD + poly(PTC, 2) + PNTC*PS + NDD + TPI)	0.5711503	0.7167015
model 19	-glm(burnt ~ FWI + GPPI + HDI + RD + poly(PTC, 2) + PNTC*PS + NDD + PGC + FAPAR12)	0.5708682	0.7175149

Model 20	glm(burnt ~ FWI + GPPI + HDI + poly(PTC, 2) + PNTC*PS + NDD + FAPAR12)	0.5677209	0.7182814
Model 21	glm(burnt ~ FWI + GPPI + HDI + poly(PTC, 2) + RD + PNTC*PS + NDD + TPI + PPV	0.5714474	0.7170576
Model 22	glm(burnt ~ FWI + GPPI + HDI + poly(PTC, 2) + PNTC*PS + NDD + TPI + PPV	0.5705348	0.7177887
Model 23	glm(burnt ~ FWI + GPPI + HDI + poly(PTC, 2) + RD + PNTC*PS + NDD + TPI + PPV	0.5714474	0.7170576
Model 24	glm(burnt ~ FWI + GPPI + HDI + poly(PTC, 2) + RD + PNTC*PS + NDD + TPI + PPV + AAP	0.5730048	0.7173093
Model 25	glm(burnt ~ FWI + GPPI + HDI + poly(PTC, 2) + PNTC + PPV + NDD + TPI	0.5682776	0.7186160
Model 26	glm(burnt ~ FWI + GPPI + HDI + poly(PTC, 2)* NDD + NTC + PPV + NTC + TPI	0.5687439	0.7194855

Table 2: Results of modeling attempts using different combinations of predictor variables using a progressive inclusion of covariates approach.

Our results reveal that each predictor variable incorporated in the final analysis significantly predicted the distribution of wildfires ($p < 0.05$), as outlined in Table 32.

-	Estimate	Std. Error	T value	Pr(> t)
(Intercept)	- 6.159e+00	2.349x10^-02	-262.17	<0.00001
FWI	9.296e-01	1.948x10^-03	477.28	<0.00001
MEPI	-2.270e+00	8.974x10^-03	-252.96	<0.00001
HDI	-1.680e+00	1.235x10^-02	-135.99	<0.00001
PNTC	5.170e-02	2.270x10^-04	227.78	<0.00001
poly(PTC,2)1	2.135e+03	1.114x10^01	191.55	<0.00001
poly(PTC,2)2	-9.783e+02	6.975	-140.27	<0.00001
TPI	2.225e-01	3.946x10^-03	56.39	<0.00001
NDD	-9.550e-03	4.757x10^-05	-200.78	<0.00001

463
464
465

466
467
468
469
470
471
472
473
474
475
476
477

PPN	-1.075e-03	1.808x10^-05	-59.48	<0.00001
-----	------------	--------------	--------	----------

Table 32. Summary of GLM coefficients for the final model, presenting t-values and p-values for predictors. The results indicate that all predictors in the final model were statistically significant aboutfor wildfire distribution (p < 0.05).

-	Estimate	Std.Error	T-value	Pr(> t)
(Intercept)	-6.159e+00	2.349e-02	-262.17	<0.00001
FWI	9.296e-01	1.948e-03	477.28	<0.00001
GPPI	-2.270e+00	8.974e-03	-252.96	<0.00001
HDI	-1.680e+00	1.235e-02	-135.99	<0.00001
PNTC	5.170e-02	2.270e-04	227.78	<0.00001
poly(PTC,2)1	2.135e+03	1.114e+01	191.55	<0.00001
poly(PTC,2)2	-9.783e+02	6.975e+00	-140.27	<0.00001
TPI	2.225e-01	3.946e-03	56.39	<0.00001
NDD	-9.550e-03	4.757e-05	-200.78	<0.00001
PPN	-1.075e-03	1.808e-05	-59.48	<0.00001

Our analysis results revealed the relationship between various predictors and BA distribution, as depicted in Fig. 3. Among the predictors studied, FWI, PNTC, PTC and TPI showed a positive relationship with BA distribution. Notably, FWI and PNTC showed particularly strong relationships, underscoring the substantial role of fire weather, fuel availability on the expansion of BA extent. Conversely, several predictors showed a negative relationship with BA distribution, including the ~~GPPI~~MEPI, HDI, PPN and NDD. A polynomial of PTC shows a slightly bell-shaped relationship with burnt area fraction.

Overall, our observations highlight the critical role of factors such as fire weather, fuel availability, vegetation cover, climate conditions, and landscape characteristics in shaping BA distribution patterns. Fig.3 visually represents the differential relationship of these predictors on BA distribution, offering a comprehensive overview of the underlying mechanisms driving wildfire dynamics.

Formatted: Line spacing: single

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Font: Bold

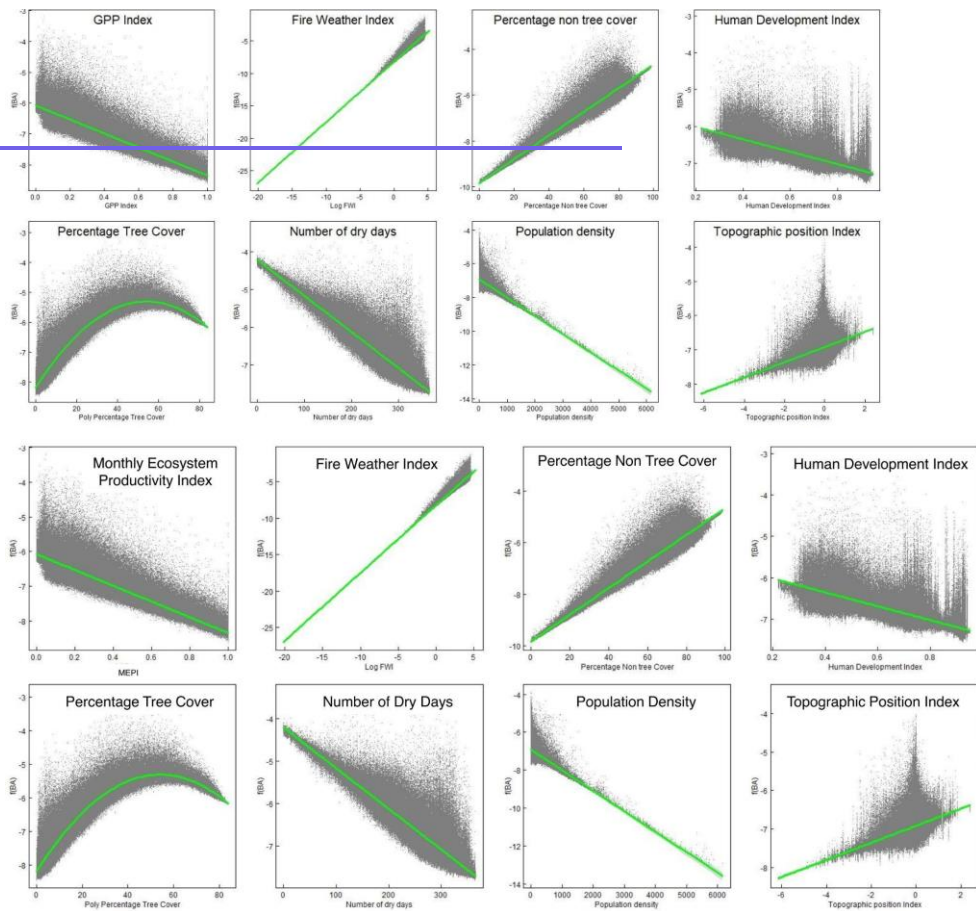


Figure 3: Partial Residual Plots illustrating the relationship between Burned Areas (BA) and the eight final predictor variables. These plots show the effect of each predictor while the others are held constant (Larson and McCleary 1972). These plots show the effect of each predictor while the others are held constant (Larsen and McCleary, 1972). Predictor variables were Gross Primary Production Index (GPP), Fire Weather Index (FWI), Percentage Non-

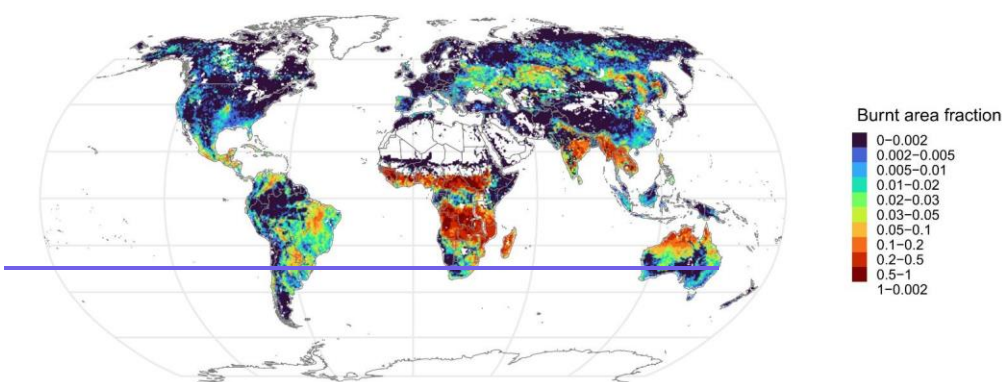
484 **Tree Cover (PNTC), Human Development Index (HDI), Percentage Tree Cover (PTC), Topographic Position Index**
485 **(TPI), Population Density (PPN) and Number of Dry Days (NDD).**

486
487 The model demonstrated strong performance in predicting BA, accounting for over 50% of the variability in burnt areas
488 (Deviance explained = 0.568). While our results slightly lagged those of a global fire distribution model by [Haas et al. \(2022\)](#),
489 ~~(2022)~~[Haas et al. \(2022\)](#), who found a deviance explained of 0.69, it's noteworthy that their model incorporated a broader
490 array of variables (16 predictors) and operated at a coarser temporal resolution (annual). Our model's performance, based on
491 eight predictors and operating at a finer temporal resolution (monthly), is considered satisfactory and parsimonious.

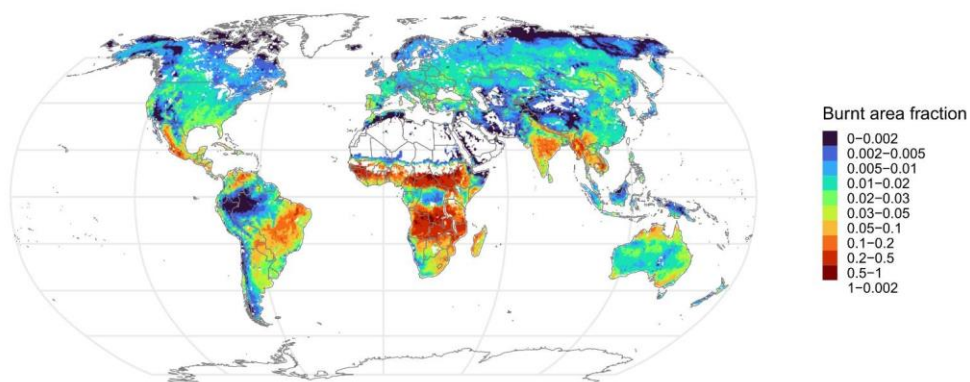
492
493 Assessment of model accuracy yielded an NME of 0.718, indicating a generally close correspondence between observed and
494 predicted burnt area patterns (see Fig. 4). ~~This level of accuracy is comparable to that reported by previous global fire models,~~
495 ~~such as (Haas et al., 2022) and (Hantson et al., 2016);~~[This level of accuracy is comparable to that reported by previous global](#)
496 [fire models, such as Haas et al. \(2022\) and Hantson et al. \(2016\)](#) which reported NMEs ranging from 0.60 to 1.10.

497
498 Spatially, our model effectively captured the distribution of BA in the tropics and the southern hemisphere, demonstrating
499 notable similarities between observed and predicted burnt area fractions on an annual basis (see Fig. 4). However, in
500 extratropical regions, particularly in the northern hemisphere, instances of over-prediction were observed. This discrepancy is
501 evident in the inconsistencies between observed annual distribution patterns and those predicted by the model.

Observed burnt area (GFED5)

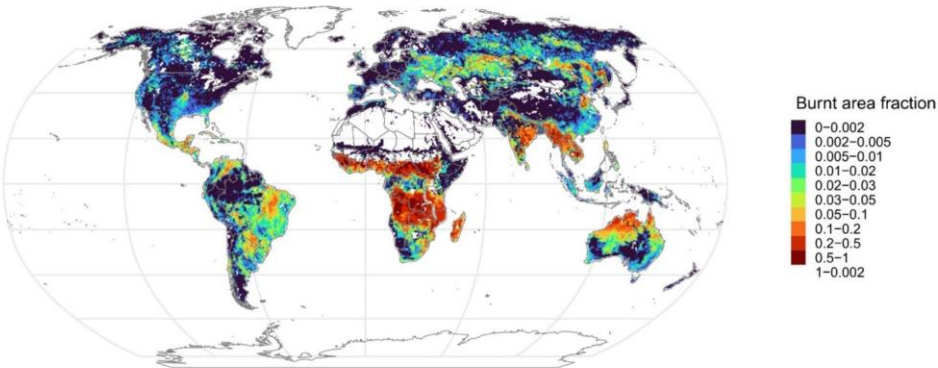


Predicted burnt area



502
503

Observed burnt area (GFED5): 2011 - 2018



Predicted burnt area: 2011 - 2018

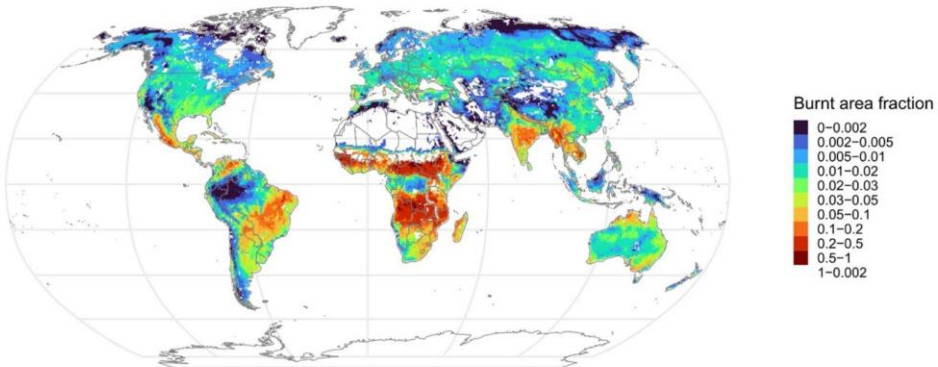
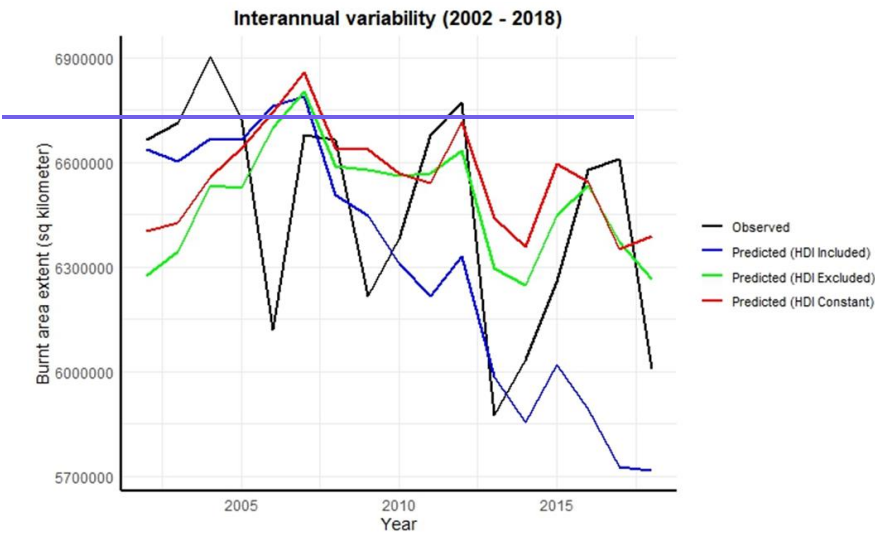


Figure 4: Annual burnt area fraction distribution map with the observed burnt area (top) and predicted burnt area (bottom).

3.2 Interannual variability

Our analysis results revealed a substantial global decrease in burnt areas exceeding 1 million square kilometers from 2002 to 2018, with the peak decline observed in 2004 (see Fig. 5). This downtrend was consistently observed in both the actual and predicted extent. Notably, the projected trend exhibited a steeper decline compared to the observed trend, indicating a potential underestimation of inter-annual variability by the model. However, it aligns with the decreasing patterns reported in earlier studies (Andela et al., 2017; Jones et al., 2022). Excluding and holding HDI constant in the model made the projected trend remain steady, suggesting the role of anthropogenic developments driving a downward trend in wildfire distribution.



3.2 Interannual variability

Our analysis results revealed a substantial global decrease in burnt areas exceeding 1 million square kilometers from 2002 to 2018, with the peak decline observed in 2004 (see Fig. 5). This downtrend was reproduced by the model, but the model underestimated the interannual variability and the model-predicted decline was stronger than observed. However, it aligns with the decreasing patterns reported in earlier studies (Andela et al., 2017; Jones et al., 2022). Excluding and holding HDI constant in the model resulted in smaller reduction of predicted BA over time.

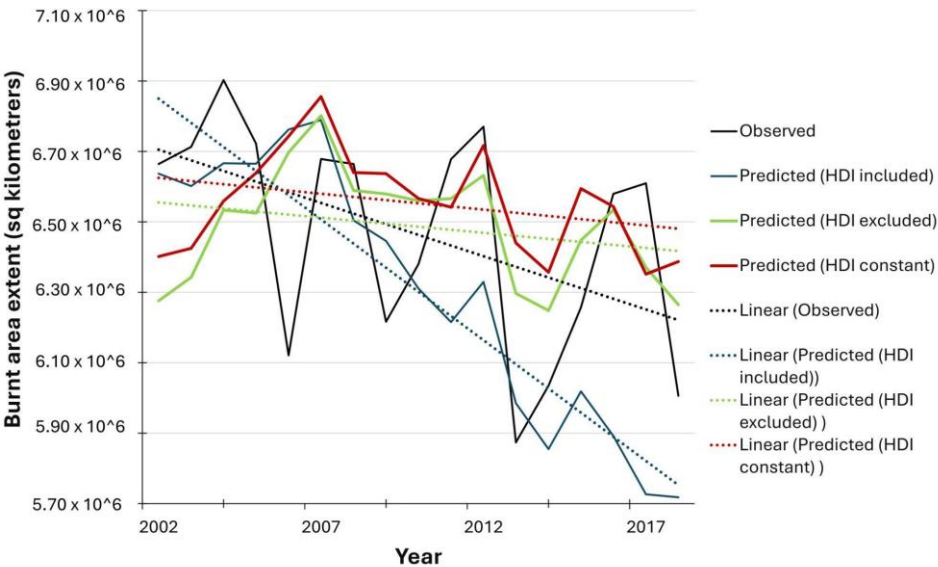
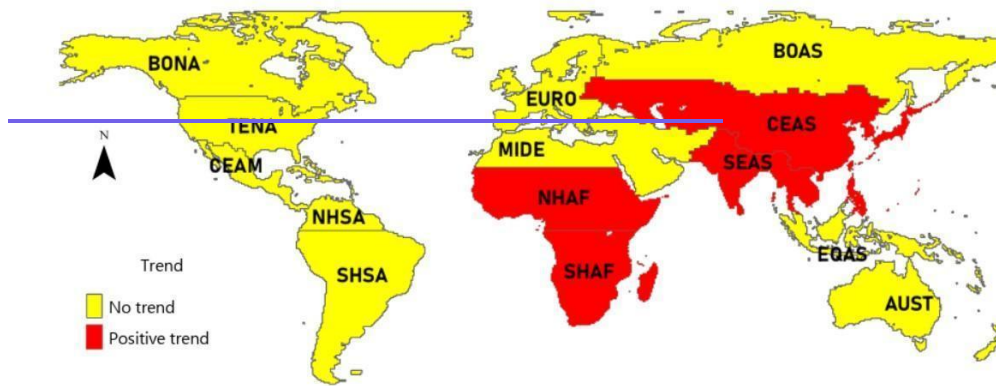


Figure 5: Interannual variability in burnt area extent showing the observed trend (based on GFED-detections) and predicted trends. Included are interannual trends GFED5 burnt estimates detection and model projections under different HDI treatments: when HDI was excluded, included and held constant from the value of the first year in the model.

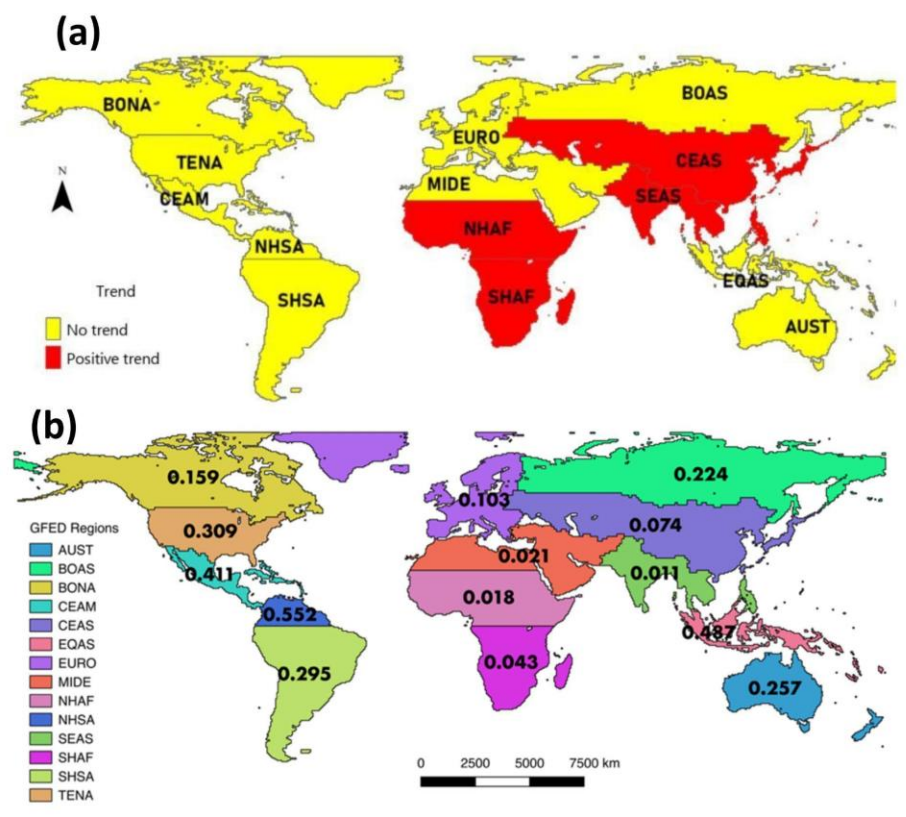
The Mann Kendall trend analysis further shows significant variation in the magnitude and direction of predicted burnt area extent across the 14 GFED regions (refer to Fig. 66a and Table A+A2). Five regions (SHAF, SHSA, NHAF, CEAS) predicted a significantly positive trend ($p < 0.05$) in burnt area extent, while the other regions predicted no significant trends (NHSA, SHSA, MIDE, TENA, AUST, EURO, EQAS, CEAM, BONA, BOAS). Overall, the projected positive trend predominated in GFED regions situated in central and southern Africa, and central and southern Asia. In contrast, the Americas, Australia, and Europe demonstrated no significant trend, as illustrated in Fig. 66a.



Formatted: Line spacing: single

538

Formatted: Line spacing: single



539

540

541 **Figure 6: Variation in the trend of interannual variability for burnt areas across different GFED regions. [Where \(a\)](#)**
542 **[shows the direction of the trend and \(b\) shows the spatial distribution of the strength of relationship \(r-square values\)](#)**
543 **[between observed and predicted interannual variability per GFED region.](#)**

544

545 Our predictive model performed poorly in predicting interannual variability as exhibited by a poor strength of relationship
546 between the predicted trend when compared to the observed ($R^2= 0.24$) (See Fig 76b and Fig A1). This poor relationship was

exhibited across most of the GFED regions ($R^2 < 0.50$), except for the NHSA which showed strong similarities between the predicted trend and observed trend ($R^2 = 0.55$). This observation suggests that the combination of covariates that we incorporated in this model has limited strength in capturing global interannual variability in burned-area. However, the predicted global trend is in sync with previously reported global trends (Jones et al., 2022).burnt area. However, the predicted global trend is in sync with previously reported global trends (Jones et al., 2022).

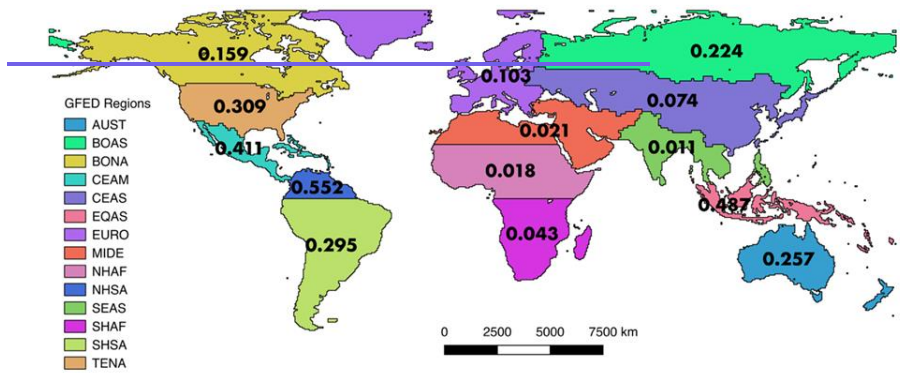
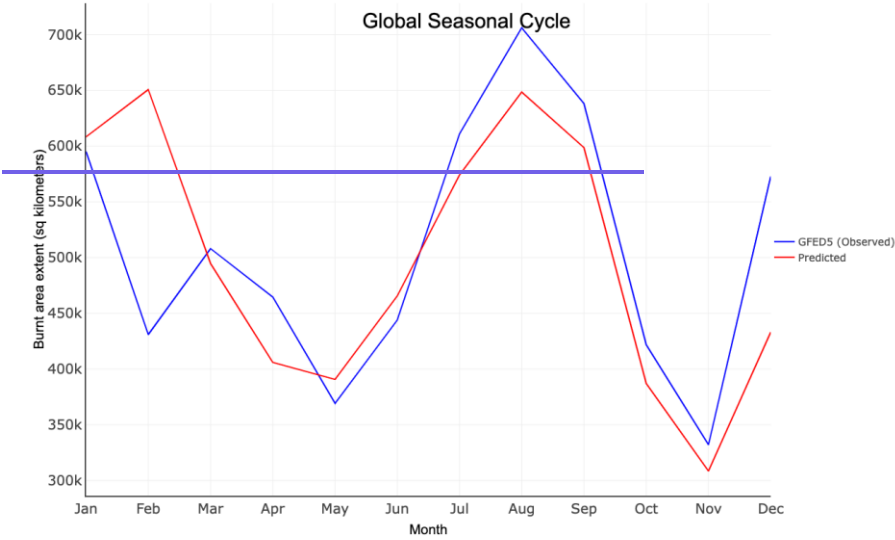


Figure 7: Spatial distribution of r-square values for the relationship between observed and predicted interannual variability per GFED region.

3.3 Seasonal Cycle

Our analysis results show that the global extent of BA shows an alternating seasonal cycle with strong peaks in February and August (see Fig. 87). The predicted pattern slightly underestimates the burnt area, however, appears to be closely knit with the observed trend ($R^2 = 0.54$). Like the global interannual trends, the strength of similarity between observed and predicted seasonal cycles varies according to the GFED region with R^2 ranging between 0.06 to 0.99 (refer to Fig. 98). The model predicted better in GFED regions that are situated in Southern Africa, South America, Australia and Asia ($R^2 > 0.50$) (see Fig. 98 and Fig A2). In contrast, poor seasonal predictions were recorded in GFED regions situated in North America, North Africa and Europe as indicated by a poor relationship between observed burnt area and predicted burnt area ($R^2 < 0.50$).

565



Formatted: Line spacing: single

Formatted: Line spacing: single

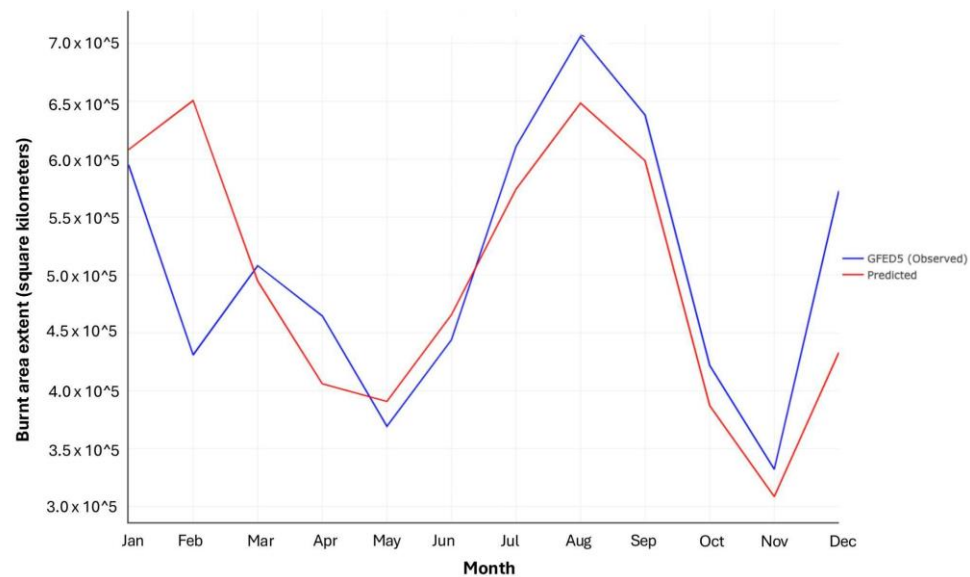
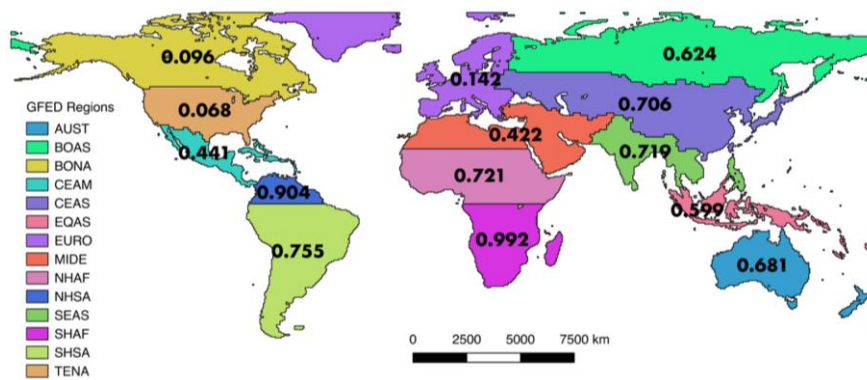


Figure 87: Global seasonal burnt area patterns showing the observed (GFED 5) and predicted burnt area extent.



570
571
572

573

574
575
576
577
578

579

580
581
582
583
584
585
586
587
588

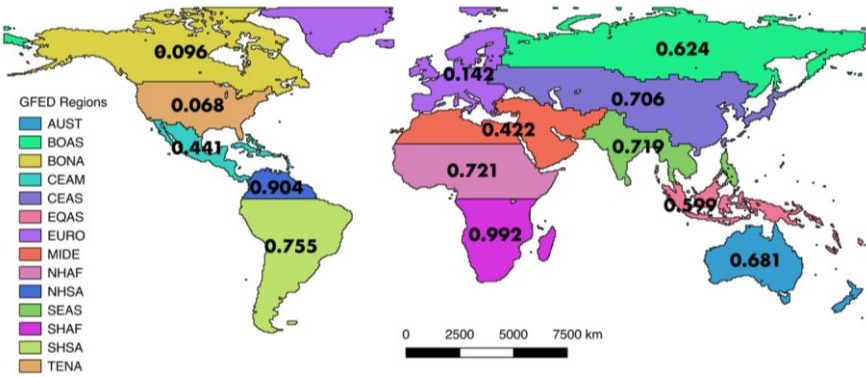


Figure 98: Spatial distribution of r-square values for the relationship between observed and predicted seasonal variability per GFED region.

4 Discussion

Wildfires are common phenomena whose dynamics may pose relevant impacts on the ecology of species and humans across different biomes. The evolving dynamics of wildfires are anticipated to undergo significant changes in the future due to global environmental shifts. In this study, we sought to tease apart statistical relationships between biophysical and socioeconomic drivers of wildfire dynamics and burned areas to facilitate DGVM integration and reliable prediction of future wildfire dynamics.

4.1 Main drivers of global burned area

We found a DGVM compatible parsimonious global statistical model made of FWI, PNTC, PTC, TPI, [GPPIMEPI](#), HDI, PPN and NDD. Of all the key variables, FWI and PNTC exhibited a strong positive relationship with fire occurrence, underscoring the importance of conducive fire-weather conditions and combustible fuel in driving wildfire occurrence and spread. High PNTC is most likely related to high amounts of flammable vegetation, such as grasses and shrubs. Our findings [show](#) that fire weather (~FWI) and fuel availability (~PNTC) influence [burnedburnt](#) area extent align with previous studies ([Andela et al., 2017](#); [Bistinas et al., 2014](#); [Forkel et al., 2019b](#); [Kuhn-Régnier et al., 2021](#));([Andela et al., 2017](#); [Bistinas et al., 2014](#); [Forkel et al., 2019b](#); [Kuhn-Régnier et al., 2021](#)). The other studies, however, did focus on [the annual burnedburnt](#) area, not the seasonal cycle, which is also crucial to adapt to changes in fire risk.

Our results show that higher PNTC leads to higher burnt area fractions. In contrast, areas with lower non-tree cover show lower burnt area fractions. Areas with high PNTC typically consist of grasses and shrubs (~ height < 2 m), while areas with low PNTC are often characterised by trees. Grasses and shrubs often encourage frequent burning much more than trees (Juli et al., 2017; Wragg et al., 2018).characterized by trees. Grass and shrubs often encourage frequent burning much more than trees (Juli et al., 2017; Wragg et al., 2018). Conversely, low PNTC indicates high tree cover, which is often less flammable, leading to fewer fires. Though our findings support previous literature indicating that regions with abundant combustible vegetation and favorable fire-weather conditions are prone to frequent burning (Kraaij et al., 2018; Thonicke et al., 2010),(Kraaij et al., 2018; Thonicke et al., 2010), we observed a surprising negative relationship between NDD and burnt area. Previous studies found a positive relationship between NDD and burnt area fractions (Haas et al., 2022), similar to(Haas et al., 2022), like our single factor plots of NDD and burnt area in Fig A3. This result most probably shows that relationships derived with annual data, as in the other studies mentioned here, cannot simply be transferred to seasonal fire predictions. Studies have shown that the effect of dryness on fire varies depending on vegetation communities in Mediterranean ecosystems (Cardil et al., 2019). Stott (2000)(Cardil et al., 2019). Stott (2000) echoed similar sentiments for tropical environments, indicating the complex relationship between vegetation, dryness and fire. Our efforts to investigate this complex relationship through an interaction term did not significantly improve our model accuracy (~ model 26). Hence, future studies may benefit from further exploring the complex relationship between dryness, and vegetation at a global scale, particularly the effect of incorporating polynomial terms on correlated predictors in a linear model.

Our findings revealed that HDI, GPPIMEPI and PPN are negatively associated with trends in global fire extent. For HDI, our findings implySpecifically, the negative relationship between HDI and burnt area implies that technological advancements, improved surveillance systems, and effective mitigation efforts play a significant role in limiting the extent of burnedburnt areas. Contrary to expectations based on Haas et al.(2022),Haas et al. (2022), PPN, which should correlate with more ignitions, does not appear to increase the burnt area extent (see Fig. 3). In fact, we observed that lower PPN corresponded to larger burnt areas, likely due to the impact of human activities on landscape fragmentation through road construction, and measures to suppress fires in human inhabited spaces to protect properties (Kloster et al., 2010). Saunders et al.(1991)(Kloster et al., 2010). Saunders et al., (1991) observed that the response of fire to changes in PPN is governed by two opposing processes, an increase in population leads to more ignition sources, while simultaneously prompting greater fire management efforts to suppress fires. They further highlighted that fire suppression rates are highest in densely populated areas. This suggests that the scale (both spatial and temporal) of analysis may influence the nature and extent to which PPN affects burnt area extent. Our results for the effect of PPN have important implications for DGVMs and land surface models. These models differ widely in the assumed effect of PPN, often using a unimodal response (Teckentrup et al., 2019). However, many DGVMs simulatesimulating BA annually, in some cases distributing the wildfires across seasons in a second step, using rather simplified assumptions (Teckentrup et al., 2019).Teckentrup et al., 2019). Similarly, we anticipated a positive relationship between GPPIMEPI and

Formatted: Line spacing: single

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

622 burnt areas, as [GPPIMEPI](#) is indicative of ecosystem dryness and flammability. However, our findings revealed a negative
623 relationship, indicating that other factors may be influencing the connection between [GPPIMEPI](#) and the extent of burnt areas.
624 Our findings are ~~inline with that of Forrest et al. (2024)~~in line with those of Forrest et al. (2024) who initially investigated the
625 effect of this index on burnt areas in Europe. Unlike previous global studies that utilized annual GPP, our research employed
626 a more refined measure, [GPPIMEPI](#). Future research could benefit from evaluating the relationships between [GPPIMEPI](#) and
627 burnt areas in other GFED regions and temporal scales.

628

629 4.2.1 Spatial variation in model performance

630 Our model exhibits stronger performance in predicting the spatial distribution of fires in southern Africa, Australia, and South
631 America than in other world regions, ~~potentially due to the seasonal patterns of key predictors (See Fig. 4). The stronger~~
632 ~~performance in these areas is likely due to the well-defined and predictable fire regimes~~ in these regions, ~~which can be. Since~~
633 ~~fire activity here is strongly governed by distinct wet-dry seasonal cycles, which align closely with climate variables such as~~
634 ~~precipitation, temperature, and vegetation productivity, factors that our model capture~~ effectively ~~captured~~ using linear
635 functions. ~~(Archibald, 2016; Van Der Werf et al., 2017). These regions typically exhibit lower interannual variability in fire~~
636 ~~occurrence, facilitating better model generalization.~~

637

638 Conversely, our model tends to overpredict fires in the northern hemisphere, particularly in North and Central America, as
639 well as Asia. ~~Annual burned area variability is relatively high in Asia, Europe, and Central America, which might~~Performance
640 ~~here declines as fire regimes are more heterogeneous and driven by a combination of biophysical and anthropogenic factors~~
641 ~~(Chuvieco et al., 2021; Forkel et al., 2019b). High interannual variability in burnt areas in these regions is due to irregular~~
642 ~~droughts, land use change, and fire suppression policies that make it more difficult to predict it (Chuvieco et al.,~~
643 ~~2024).~~prediction more challenging for linear models. Additionally, the influence of snow cover, freeze-thaw cycles, and varied
644 ignition sources in temperate and boreal regions further complicates seasonal pattern detection (Flannigan et al., 2009).
645 Chuvieco et al., (2021) reported about this challenge when building global models. Thus, our findings build upon existing
646 models on global ~~burned~~burnt area distribution. What sets our model apart from previous models is its ability to reliably
647 identify global seasonal fire distribution patterns. This simplicity offers a notable advantage, as it facilitates more nuanced
648 interpretation and implementation of DGVMs compared to annual models.

649 4.3 Attribution of global trends

650 Our model has contributed novel insights to the existing understanding of the factors influencing global fire trends (~~Joshi and~~
651 ~~Sukumar, 2021; Kraaij et al., 2018; Mukunga et al., 2023). Previous research, such as the work by Andela et al. (2017), (Joshi~~
652 ~~and Sukumar, 2021; Kraaij et al., 2018; Mukunga et al., 2023). Previous research, such as the work by Andela et al. (2017).~~

Formatted: Line spacing: single

Formatted: Font color: Auto

Formatted: Font color: Auto

primarily attributed the decline in global burnt areas to agricultural expansion and intensification. [Earl and Simmonds \(2018\)](#) supported this view, adding that increased net primary productivity in Northern Africa also played a significant role. However, our results suggest that human development is a more important driver than agricultural expansion alone. Despite the conventional emphasis on agricultural factors, our attempt to incorporate cropland and rangeland fractions as predictor variables did not substantially enhance our understanding of this trend (model 5 -10, Table 2). Interestingly, our analysis revealed that excluding the HDI from our model and holding it constant to the value of the first year predicted a steady trend that deviates from the observed negative trend in global fire extent and including HDI follows a decreasing trend that aligns with the observed trend (Fig. 5). This highlights the significant influence of HDI in projecting the purported negative global fire trend. HDI is rather broad socioeconomic indicator, which we assume acts as a proxy for factor such as investments and advancements in fire control methods, surveillance, technology, and outreach strategies increasing awarenessThe HDI is related to factors like advancements in fire control methods, surveillance, technology, and outreach strategies increasing awareness, particularly in response to the growing human technological developments. [\(Teixeira et al., 2023\)](#). Although these strategies are often implemented independently and on a smaller scale, their cumulative impact on global fire trends is substantial. Therefore, our model underscores the necessity for global initiatives aimed at enhancing fire control measures through comprehensive awareness campaigns, capacity-building efforts, resource mobilization, and the development and deployment of reliable surveillance technologies. By addressing these factors collectively, we can effectively mitigate the extent and severity of global wildfires, thereby safeguarding ecosystems and human livelihoods.

4.4 Interannual variability

Despite demonstrating the significant role of the HDI in predicting global fire trends, our model struggled to achieve high precision in forecasting interannual variability both globally (see Fig. 5) and within specific GFED regions (see Fig. S1). Recognizing that this limitation might stem from an inadequate representation of vegetation (fuel) dynamics, we incorporated FAPAR12 in models 9 to 12 (Table [2A1](#)) and GPP4MEP1 in models 11 to 26 (Table [2A1](#)). Unfortunately, these adjustments did not enhance our ability to predict the interannual variability of wildfires. Studies have found a relationship between increased precipitation in the years preceding the fire season and fire activity in the drier savanna regions of Southern Africa [\(Shekede et al., 2024\)](#)[\(Shekede et al., 2024\)](#). Hence, we also explored the role of previous fuel accumulation on subsequent fire seasons using GPP12 in model 10, respectively. While this approach did not improve global interannual predictions, it showed a slight enhancement in deviance explained (from 0.5357 to 0.5461). This improvement might have been confounded by the effects of the fire-aerosol positive feedback mechanism in Africa [\(Zhang et al., 2023\)](#)[\(Zhang et al., 2023\)](#) and periodic El Niño conditions, which can affect rainfall patterns and lead to drier vegetation conditions, reducing the predictability of fire occurrence, especially with linear models [\(Shikwambana et al., 2022\)](#). Other attempts at simulating global annual burned[\(Shikwambana et al., 2022\)](#). We note that in the recent comparison of fire-enabled DGVMs in the Fire Model Intercomparison Project (FireMIP) project [\(Hantson et al. 2020\)](#), all models did a poorer job of matching the interannual

variability than the spatial patterns by a considerable margin. The seven acceptably-performing models achieved a mean spatial NME (across all data and model comparisons) of 0.84 with respect to spatial patterns, but an NME of 1.15 for interannual variability. Other attempts at simulating global annual burnt area changes, for example with fire-enabled DGVMs (Fire Model Intercomparison Project (FireMIP)), LPJ-GUESS-GlobFIRM yielded similarly poor model performance concerning interannual variability (Hantson et al., 2020). Our modeling, and its successor, SPITFIRE, yielded similarly poor model performance concerning interannual variability (Hantson et al., 2020). Our modelling efforts highlight the complexity of accurately predicting wildfire trends and underscore the need for future research to identify covariates that more effectively capture the interannual variability of fires at a global scale.

4.5 Fire seasonality

The findings of this study exhibit robustness in capturing seasonal cycles ($R^2=0.536$), facilitated by the inclusion of monthly variables such as the GPPI and the logarithm of FWI, which are pivotal in delineating seasonal fire patterns. While the seasonal predictions demonstrated reliability across most GFED regions globally, notable exceptions were observed in North America, North Africa, and Europe ($R^2<0.50$). This discrepancy could be attributed to the intricate climatic conditions inherent to these regions, which influence fire weather in a manner that eludes simple linear modeling. Given the parsimonious design of our model, with only eight predictors, we think that the model performance is acceptable. For certain regions, it might be possible to increase model performance by implementing further region-specific predictors and relationships. Accurate predictions regarding the seasonal dynamics of diverse GFED regions can facilitate the identification of temporal windows when fires are prevalent, thereby furnishing valuable insights for simulating carbon emissions in DGVMs.

Globally, our model predicts a notable peak in burned areas during February and August. The February peak corresponds to dry conditions and fuel accumulation in regions such as NHSA, NHAF, and MIDE. In contrast, the August peak primarily emanates from tropical regions characterized by distinct seasonal patterns, particularly in SHSA, SHAF, and AUST. Here, the dry season augments the combustibility of accumulated fuel from the preceding wet season, facilitating fire spread. This observation corroborates earlier studies in the southern hemisphere, which underscore the prevalence of wildfires during prolonged dry spells (Magadzire, 2013; Shekede et al., 2024; Strydom and Savage, 2017). (Magadzire, 2013; Shekede et al., 2024; Strydom and Savage, 2017). Increased temperatures and desiccated vegetation substantially enhance the likelihood and severity of wildfires during the dry season. Conversely, the onset of the rainy season precipitates a marked reduction in the occurrence of wildfires in these regions. This underscores the enduring influence of fire weather and vegetation dynamics as principal drivers of seasonal burnt area cycles, with factors such as moisture content in vegetation and soil, as well as humidity, playing pivotal roles in modulating ignition and fire extent within ecosystems. The seasonal global forecasts generated by our model hold significant implications for guiding adaptive strategies, fire management and prevention at both regional and global scale.

4.6 Excluded drivers of burned area

The findings of this study exhibit robustness in capturing seasonal cycles ($R^2 = 0.536$), facilitated by the inclusion of monthly variables such as the MEPI and the logarithm of FWI, which are pivotal in delineating seasonal fire patterns. While the seasonal predictions demonstrate reliability across most GFED regions globally, notable exceptions were observed in North America, North Africa, and Europe ($R^2 < 0.50$) (see Fig 8). This discrepancy could be attributed to the intricate climatic conditions inherent to these regions, which influence fires in a manner that eludes simple linear modelling. For instance, regions with clear-cut wet and dry seasons tend to exhibit more regular fire cycles, largely governed by seasonal shifts in precipitation, temperature, and vegetation growth. These predictable patterns make them well-suited to linear modelling approaches (Van Der Werf et al., 2017). In contrast, areas in the northern hemisphere experience more irregular and less seasonally driven fire activity. Here, the interaction of drought events, land management, and socio-economic drivers introduces variability that weakens model performance (Chuvieco et al., 2021; Forkel et al., 2019b). Additionally, varied ignition sources in temperate and boreal zones disrupt consistent seasonal fire patterns (Flannigan et al., 2009). Given the parsimonious design of our model, with only eight predictors, we think that the model's performance is acceptable. For certain regions, it might be possible to increase model performance by implementing further region-specific predictors and relationships. Accurate predictions regarding the seasonal dynamics of diverse GFED regions can facilitate the identification of temporal windows when fires are prevalent, thereby furnishing valuable insights for simulating carbon emissions in DGVMs.

4.6 Model limitations and excluding drivers of burnt area

Several covariates initially considered, such as landcover variables (~PCC, PPS, PRC, PGC), vegetation (~FAPAR) and socioeconomic (~RD), did not make it to the final model (See Table 2) despite their potential relevance identified in previous studies (Forkel et al., 2019b; Hantson et al., 2015; Knorr et al., 2014; Pausas and Keeley, 2021; Perkins et al., 2022). The differences to A1 despite their potential relevance identified in previous studies (Forkel et al., 2019b; Hantson et al., 2015; Knorr et al., 2014; Pausas and Keeley, 2021; Perkins et al., 2022). The differences in our findings are related to differences in the statistical or modelling approach and the fact that most of these studies addressed annual BA patterns, not seasonal variations. Nevertheless, these other factors can clearly also be important for understanding fire dynamics, e.g. influencing fuel availability, landscape structure, and ignition sources. For instance, grazing lands can significantly impact fire behavior by altering fuel types and continuity, with areas used for grazing potentially reducing fuel loads (Davies et al., 2010; Strand et al., 2014) (Davies et al., 2010; Strand et al., 2014). Similarly, FAPAR indicates vegetation health and productivity, affecting fuel moisture content and thus fire risk (Pausas and Ribeiro, 2013) (Pausas and Ribeiro, 2013). However, these factors are apparently indirectly represented by the final model, as they are correlated to the driver variables in the final model. FAPAR, for example, is generally highly correlated with GPP. Furthermore, RD is associated with human-caused ignitions and fire

suppression capabilities (Forkel et al., 2019b). However, it was excluded here because its contributions were already effectively represented by HDI and PPN, which capture broader socioeconomic conditions and infrastructure impacts. Apart from that, Haas et al. (2022) observed a shift in the direction of contribution for covariates when PPN and RD are used together. Considering that we may not have future projections for RD unlike PPN, including the issue of collinearity, we decided to retain only PPN in our model. Furthermore, our attempt to include RD in our models 21, 23 and 24 (Table 2A1) yielded marginal improvements, which were not different from when we excluded it in model 25. Overall, the decision to exclude most of these covariates was aimed at reducing redundancy and multicollinearity, ensuring a balance between model complexity and predictive power. By focusing on more comprehensive variables with high explanatory power, the final model achieves robust explanatory power. However, the often-small differences in the deviance explained and the NME between different models imply that vegetation-fire modellers might also pick a slightly different set of variables for DGVM integration without using much predictive power.

4.7 Shortcomings and Recommendations

The findings of this study offer valuable insights into the underlying drivers and patterns shaping global fire dynamics. While our research represents relevant efforts in developing a streamlined model capable of accurately capturing seasonal variations in global fire distribution, it's important to acknowledge certain limitations. The selection of covariates and the statistical model was constrained by the necessity for integration within DGVMs applied to predict future dynamics, potentially omitting some previously identified key predictors (~lightning frequency, gridded livestock densities) and modeling techniques (~ Random Forest, Neural networks, XgBoost, CatBoost) for global fires (Forkel et al., 2019b; Joshi and Sukumar, 2021; Mukunga et al., 2023; Zhang et al., 2023). This might contribute to observed shortcomings in our model's ability to predict spatial fire distribution in certain regions and to capture interannual variability across many parts of the world. Future investigations should aim to explore the inclusion of other established predictors and methodologies in global fire modeling once they become easily compatible with DGVM integration. Despite these challenges, our study possesses intrinsic value, and the developed model stands as a relatively simple tool for informing global seasonal fire predictions.

4.7 Next steps for DGVM integration, future directions and model improvements

To integrate the model presented here into a DGVM and enable future predictions, the remotely sensed variables of vegetation state (PTC, PNTC and MEPI) must be replaced with equivalent variables from the DGVM. DGVMs include GPP and the cover fractions of vegetation types required to calculate PTC, PNTC and MEPI, so these variables provide a robust and universal coupling strategy to capture the effect of vegetation on burnt area. However, all model results are imperfect and biased to some degree, so the DGVM variables will not correspond perfectly with the remotely sensed ones used for model

training. This error will propagate to the burnt area calculation and so this discrepancy should be investigated. In the likely event that this discrepancy is not small, the GLM should be refitted using the model-calculated variables to implicitly account for biases in the DGVM simulation, although the burnt area estimated will still be dependent on the DGVM's skill to capture certain dynamics and states. However, we note that our comparatively restricted variable set and simple GLM approach will be more straightforward to integrate and less sensitive to errors in the DGVM simulated state than machine learning approaches with larger suites of predictor variables. For example-, Son et al., (2024) achieved excellent correspondence with observed data using an advanced recursive neural network which was partially integrated into the JSBACH DGVM. However, only the fuel predictor was taken from the prognostically simulated JSBACH model state, other high importance dynamic predictors (including PFT cover fractions and both absolute values and anomalies of LAI and water content of four soil layers) are all determined from fixed input data - remotely sensed of climate reanalysis. So, in this case, the quality of the results from hypothetical full integration will be dependent on the ability of JSBACH to simulate many more variables correctly. The model presented here is tailored for integration into a DGVM by using only a few variables which can be robustly predicted, and, as a simple GLM in contrast to more complex machine learning methods, is less prone to overfitting and relying on correlations in the data which may not hold in the DGVM predicted state.

In comparison to the vegetation variables, the inclusion of the other variables (FWI, HDI, PPN, NDD, TPI) is trivial as they can either be prescribed input variables or can be calculated from the climate input. Finally, to build a fully coupled vegetation-fire model, it is then necessary to include the effects of the simulated fire on the vegetation. For this step we can utilise the mortality and combustion components of fire models already available and integrated into DGVMs, for example the BLAZE model (Rabin et al., 2017) or the appropriate equations in SPITFIRE (Thonicke et al., 2010). These parameterizations may need to be adjusted to account for the different simulated burnt area.

5. Conclusions

Global fire patterns undergo constant changes influenced by fluctuations in vegetation, weather conditions, topography, and anthropogenic factors. Despite numerous attempts in previous studies We sought to describe global fire distribution, the development build a parsimonious statistical model for global seasonal burnt areas that can be integrated into a DGVM. The specific objectives were to 1) to improve our understanding of a concise model capable major drivers of explaining and global burnt area dynamics, 2) to leverage a GLM for predicting global fire patterns, particularly one that seamlessly integrates with DGVMs, is essential for a reliable assessment burnt areas using DGVM-integrable predictors and 3) to evaluate the interannual and seasonal cycles of the impacts of global change. In this study, we burnt area extent, both globally and regionally.

811 We present a parsimonious statistical model specifically tailored for global ~~seasonal-burned~~burnt areas, with the goal of
812 integration into DGVMs.

813

814 ~~FWI, PTC, TPI and PNTC were positively related to BA, while MEPI, HDI, PPN, and NDD were negatively related to BA.~~
815 Our findings highlight the significance of socio-economic advancements, particularly those improving fire management
816 strategies, as evidenced by the negative trend in global fire extent predicted through the inclusion of the HDI as a crucial socio-
817 economic predictor in our model. ~~Additionally, fire weather and vegetation dynamics, specifically the FWI and the novel GPP~~
818 ~~index, emerged as robust predictors of seasonal global fire patterns.~~ While our parsimonious model exhibited limitations in
819 predicting the interannual variability of global fires, it demonstrated commendable accuracy in forecasting the spatial ~~and~~
820 ~~seasonal distribution of wildfires.~~ (NME = 0.72). ~~The strength of similarity between observed and predicted seasonal cycles~~
821 ~~varied according to the GFED region with R² ranging between 0.06 to 0.99. Its standout performance laid in capturing the~~
822 ~~seasonal variability, especially in regions often characterized by distinct wet and dry seasons, notably southern Africa (R² =~~
823 ~~0.72 to 0.99), Australia (R² 68) and South America (R² = 0.75 to 0.90). Our predicted interannual variability exhibited poor~~
824 ~~strength of relationship between the predicted trend when compared to the observed (R²= 0.24)~~

825

826 We hope that our research outcomes will stimulate a more rigorous implementation of global fire models within DGVM
827 frameworks. This, in turn, will contribute to a deeper understanding of the intricate dynamics of global fire patterns and
828 enhance our capacity to effectively manage and mitigate the consequences of evolving fire regimes in the face of ongoing
829 global changes.

830

831 **Code and data availability**

832 The code used in this analysis model fitting, and plotting is available at <https://doi.org/10.5281/zenodo.14177016>. Data used
833 for model fitting are available at <https://doi.org/10.5281/zenodo.14110150>.

834

835

836 **Author contribution**

837 BK contributed to ~~conceptualisation~~conceptualization of the model and data analysis and model fitting. MF and TH supported
838 ~~in~~ developing the statistical framework and interpreting the results. BK drafted the manuscript, with input from MF and TH.

839

840 **Acknowledgements**

Formatted: Line spacing: single

Formatted: Line spacing: 1.5 lines

Formatted: Font: Bold

841 This project has received funding from German Research Foundation (DFG) grant “Fire in the Future: Interactions with
842 Ecosystems and Society (FURNACES) project” under grant number HI 1538/14-1. BK received a salary from FURNACES.
843

844 **Competing interests**

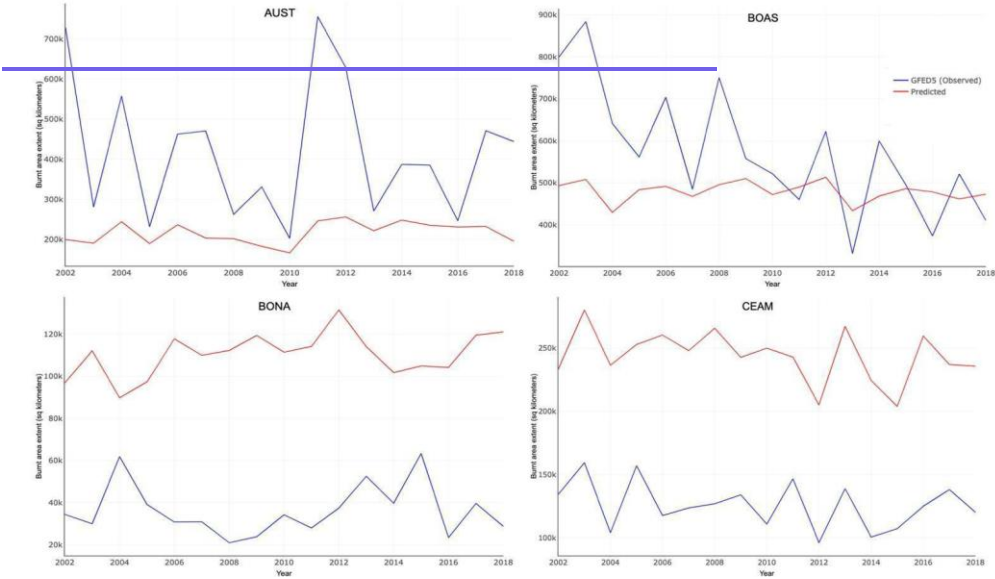
845 The authors declare that they have no known competing financial interests or personal relationships that could have influenced
846 the work reported in this paper.
847
848
849
850
851
852
853

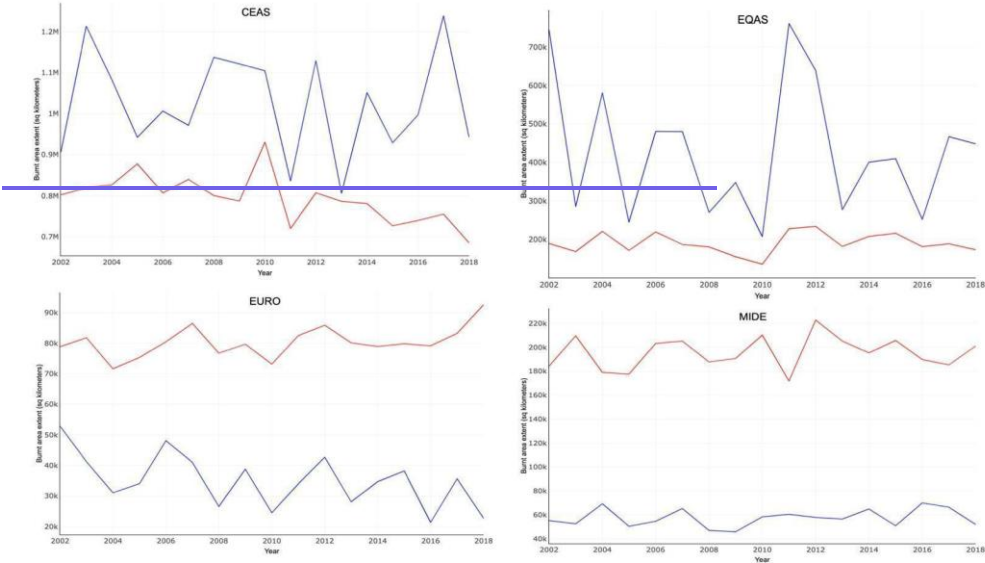
854 **Appendices**
855

856 Table A1: Results of modelling attempts using different combinations of predictor variables using a progressive
857 inclusion of covariates approach. For Normalized Mean Error (NME), higher values are represented by warmer
858 colours (with red indicating the highest error), while lower values appear in cooler colours (green indicating the lowest
859 error). In contrast, for Deviance Explained, higher values are shown in cooler colours (green indicating better
860 performance), and lower values in warmer colours (red indicating poorer performance). An optimal model is indicated
861 by a combination of cooler colours for both metrics, whereas a combination of warmer colours suggests poor model
862 performance.
863

<u>Model</u>	<u>Formulae</u>	<u>Deviance explained</u>	<u>NME</u>
<u>model 1</u>	<u>glm(burnt ~ FWI + GPP + HDI + PTC + RD)</u>	<u>0.3548030</u>	<u>0.7472088</u>
<u>model 2</u>	<u>glm(burnt ~ FWI + GPP + HDI + PTC + RD + PGC)</u>	<u>0.3699393</u>	<u>0.7495652</u>
<u>model 3</u>	<u>glm(burnt ~ FWI + GPP + HDI + PTC + RD + PNTC)</u>	<u>0.5298061</u>	<u>0.7208771</u>
<u>model 4</u>	<u>glm(burnt ~ FWI + GPP + HDI + PTC + RD + PNTC + FAPAR)</u>	<u>0.5312036</u>	<u>0.7188448</u>
<u>model 5</u>	<u>glm(burnt ~ FWI + GPP + HDI + PTC + RD + PNTC + FAPAR + PCC)</u>	<u>0.5312697</u>	<u>0.7191269</u>

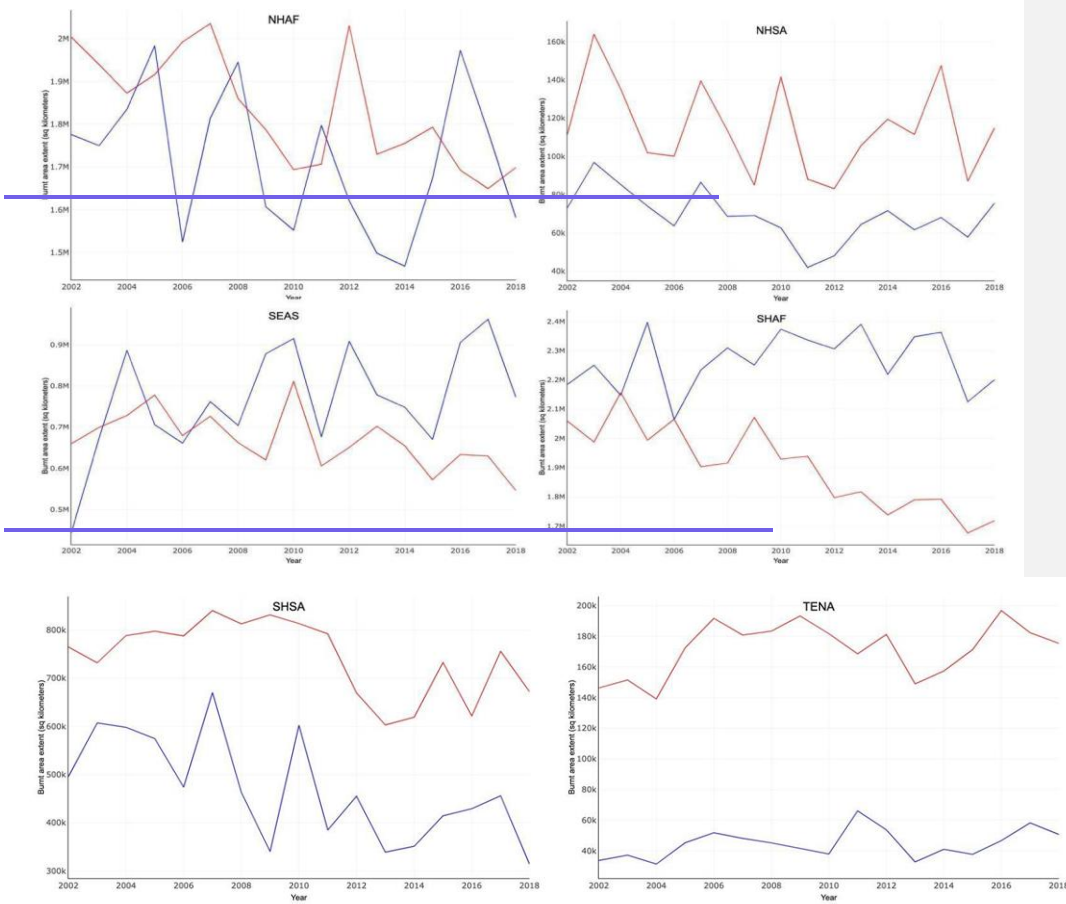
model 6	glm(burnt ~ FWI + GPP + HDI + PTC + RD + PNTC + FAPAR + PCC + PPS)	0.5328183	0.7195616
model 7	glm(burnt ~ FWI + GPP + HDI + PTC + RD + PNTC + FAPAR + PCC + PRC)	0.5313813	0.7193946
model 8	glm(burnt ~ FWI + GPP + HDI + PTC + RD + PNTC + FAPAR + PCC + PGC)	0.5349288	0.7190611
model 9	glm(burnt ~ FWI + GPP + HDI + PTC + RD + PNTC + FAPAR + PCC + FAPAR12 + PGC)	0.5359802	0.7181930
model 10	glm(burnt ~ FWI + GPP12 + HDI + poly(PTC, 2) + PNTC + FAPAR + PCC + FAPAR12 + PGC + PPN)	0.5295939	0.7172668





Formatted: Line spacing: single

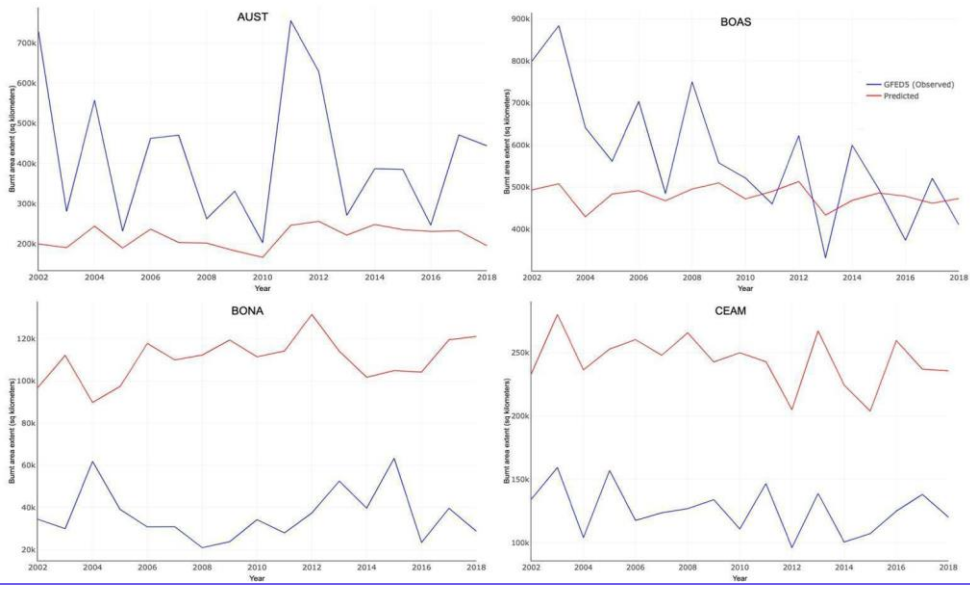
Formatted: Line spacing: single



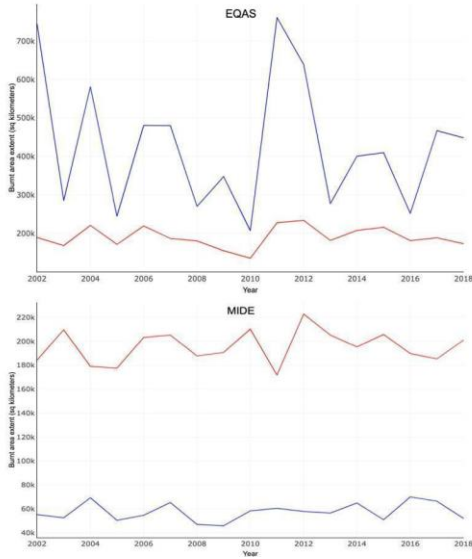
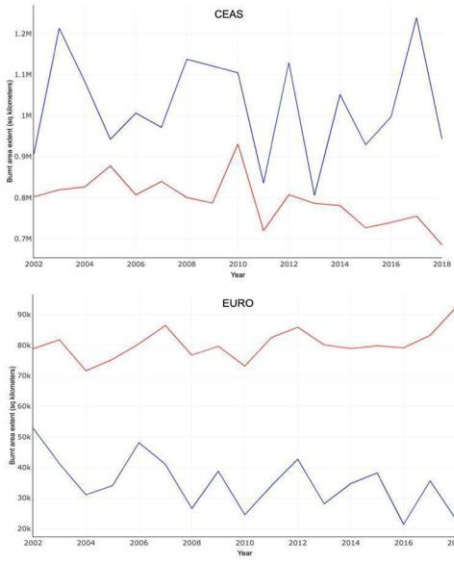
model 11	glm(burnt ~ FWI + MEPI + HDI + poly(PTC, 2) + RD + PNTC + FAPAR12 + PGC + PS)	0.5579946	0.7193546
model 12	glm(burnt ~ FWI + MEPI + HDI + poly(PTC, 2) + RD + PNTC + FAPAR12 + PS)	0.5571164	0.7192122
model 13	glm(burnt ~ FWI + MEPI + HDI*PCC + PGC + RD + poly(PTC, 2) + PNTC + PS)	0.5569187	0.7214560

Formatted: Line spacing: single

<u>model</u> <u>14</u>	<u>glm(burnt ~ FWI + MEPI + HDI*PGC + RD + poly(PTC, 2) + PNTC+ PS)</u>	<u>0.5570586</u>	<u>0.7222061</u>
<u>model</u> <u>15</u>	<u>glm(burnt ~ FWI + MEPI + HDI*PRC + RD + poly(PTC, 2) + PNTC + PS)</u>	<u>0.5664789</u>	<u>0.7154708</u>
<u>model</u> <u>16</u>	<u>glm(burnt ~ FWI + MEPI*PNTC + HDI + RD + poly(PTC, 2) + PS)</u>	<u>0.5563012</u>	<u>0.7215202</u>
<u>model</u> <u>17</u>	<u>glm(burnt ~ FWI + MEPI + HDI + RD + poly(PTC, 2) + PNTC + PS + NDD)</u>	<u>0.5681926</u>	<u>0.7191069</u>
<u>model</u> <u>18</u>	<u>glm(burnt ~ FWI + MEPI + HDI + RD + poly(PTC, 2) + PNTC* PS + NDD + TPI)</u>	<u>0.5711503</u>	<u>0.7167015</u>
<u>model</u> <u>19</u>	<u>glm(burnt ~ FWI + MEPI + HDI + RD + poly(PTC, 2) + PNTC* PS + NDD + PGC + FAPAR12)</u>	<u>0.5709692</u>	<u>0.7175149</u>
<u>Model</u> <u>20</u>	<u>glm(burnt ~ FWI + MEPI + HDI + poly(PTC, 2) + PNTC* PS + NDD + FAPAR12)</u>	<u>0.5677209</u>	<u>0.7182814</u>
<u>Model</u> <u>21</u>	<u>glm(burnt ~ FWI + MEPI + HDI + poly(PTC, 2) + RD + PNTC*PS + NDD + TPI + PPN</u>	<u>0.5714474</u>	<u>0.7170576</u>
<u>Model</u> <u>22</u>	<u>glm(burnt ~ FWI + MEPI + HDI + poly(PTC, 2)+ PNTC*PS + NDD + TPI + PPN</u>	<u>0.5705348</u>	<u>0.7177887</u>
<u>Model</u> <u>23</u>	<u>glm(burnt ~ FWI + MEPI+ HDI + poly(PTC, 2) + RD + PNTC*PS + NDD+ TPI+ PPN</u>	<u>0.5714474</u>	<u>0.7170576</u>
<u>Model</u> <u>24</u>	<u>glm(burnt ~ FWI + MEPI + HDI + poly(PTC, 2) + RD + PNTC*PS + NDD + TPI + PPN + AAP</u>	<u>0.5720048</u>	<u>0.7173093</u>
<u>Model</u> <u>25</u>	<u>glm(burnt ~ FWI + MEPI + HDI + poly(PTC, 2) + PNTC + PPN + NDD + TPI</u>	<u>0.5682776</u>	<u>0.7186160</u>
<u>Model</u> <u>26</u>	<u>glm(burnt ~ FWI + MEPI + HDI + poly(PTC, 2)* NDD + NTC + PPN + NTC + TPI</u>	<u>0.5687439</u>	<u>0.7194855</u>



Formatted: Line spacing: single



Formatted: Line spacing: single

Formatted: Line spacing: single

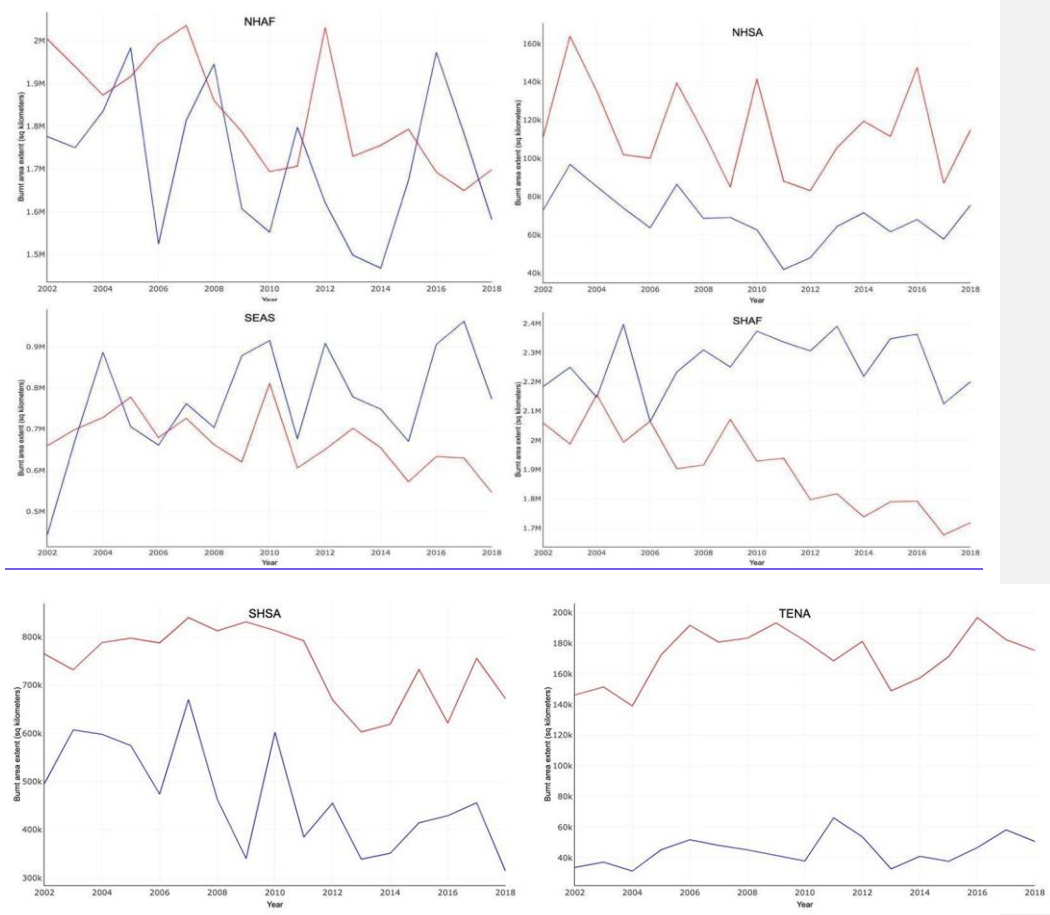
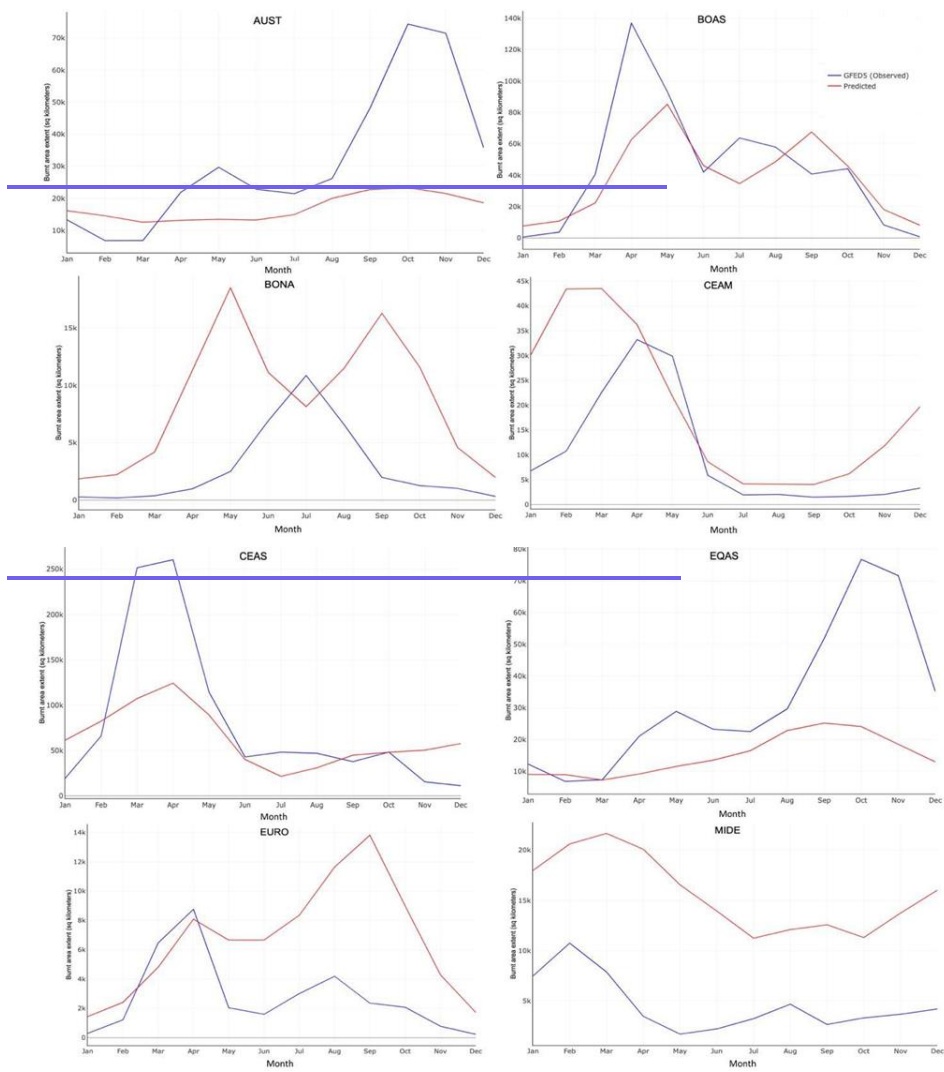
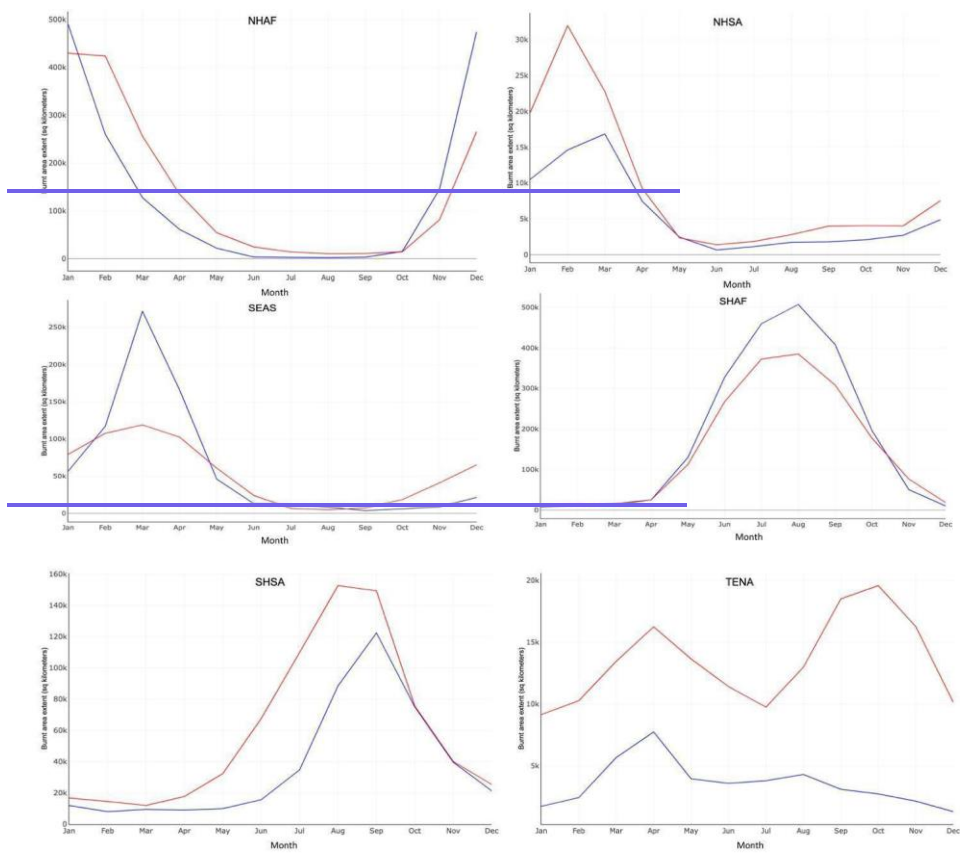


Figure A1: Shows the observed (in red) and predicted (in blue) interannual variability in burnt area fractions across different GFED regions.

Formatted: Line spacing: single

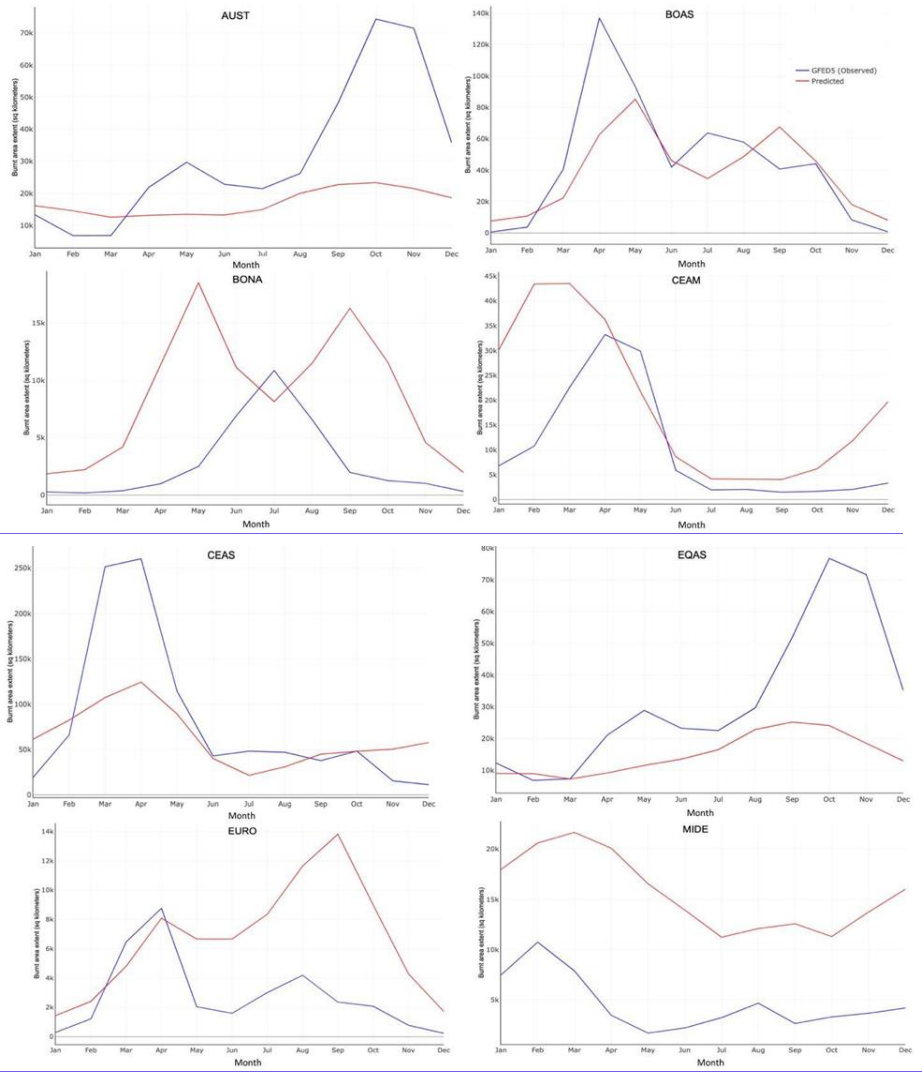


Formatted: Line spacing: single



883

884



Formatted: Line spacing: single

Formatted: Line spacing: single

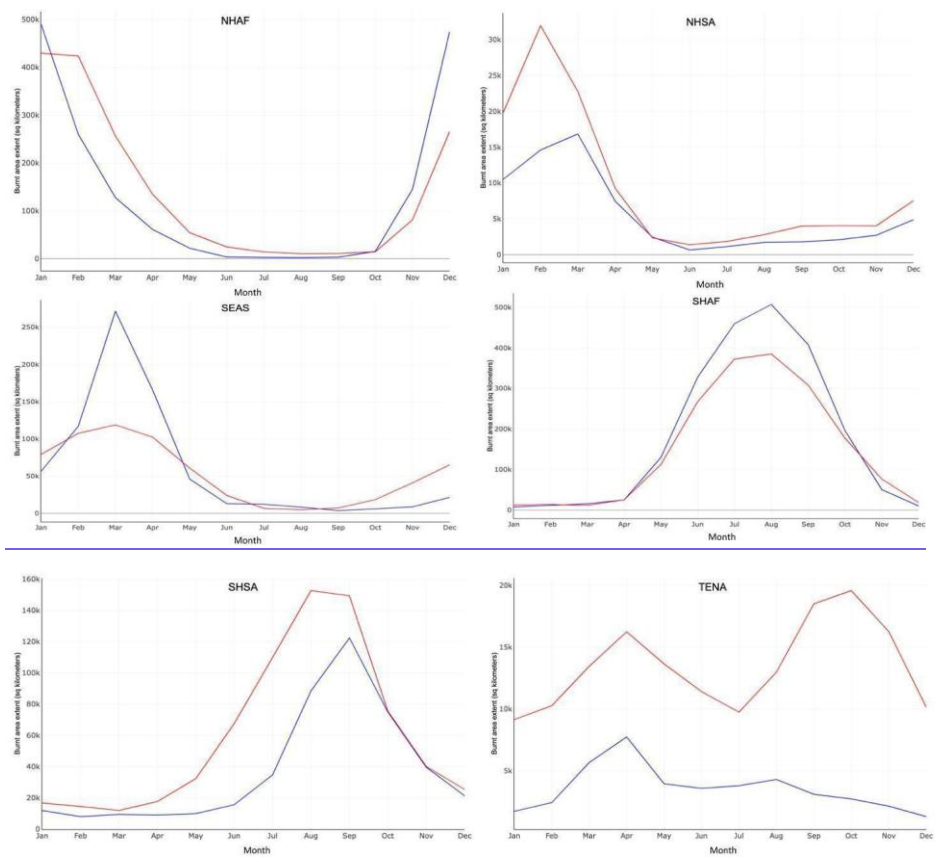
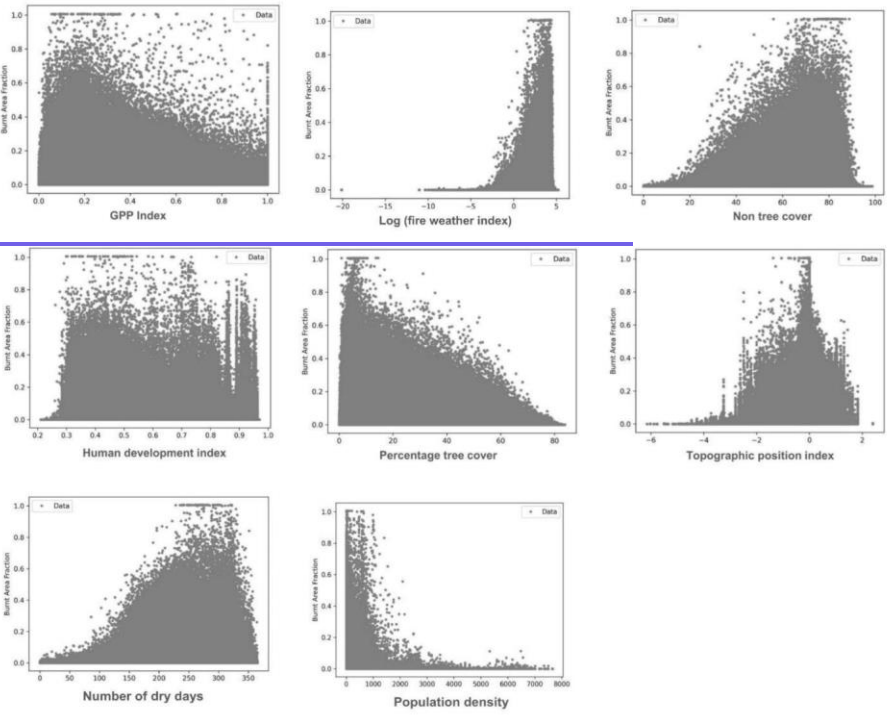


Figure A2: Shows the observed (in red) and predicted (in blue) seasonal variability in burnt area fractions across different GFED regions.



Formatted: Line spacing: single

Formatted: Line spacing: single

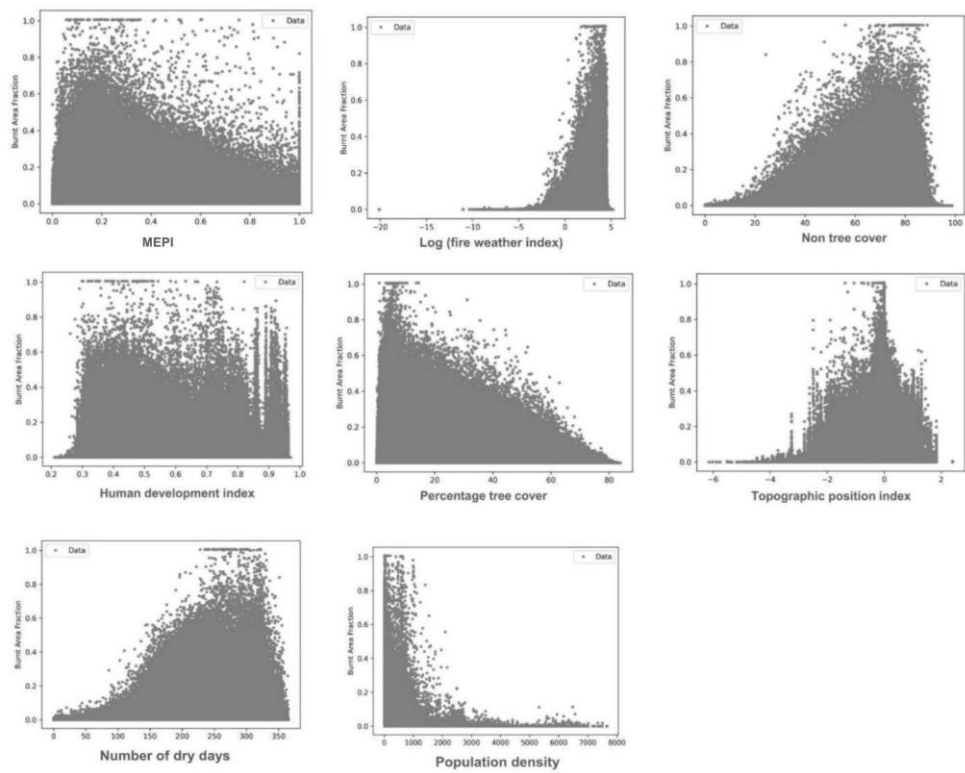


Figure A3: Scatter plots illustrating single-factor relationships between burnt area fraction and various environmental and socio-economic variables: GPP Index, Fire Weather Index, Percentage Non-Tree Cover, Human Development Index, Percentage Tree Cover, Topographic Position Index, Percentage Dry Days, Road Density, Precipitation Seasonality and Annual Precipitation Index. The plots highlight distinct patterns, such as the negative correlation between percentage tree cover and burnt area fraction, and the positive correlation between number of dry days and burnt area fraction.

Formatted: Line spacing: single

Region	Sen's slope	P-value
BONA	558.354	0.1082
TENA	895.8292	0.4338
CEAM	-1963.035	0.1494
NHSA	-1601.363	0.387
SHSA	-9119.019	0.0529
EURO	189.2956	0.387
MIDE	202.3893	0.9016
NHAF	-22329.83	0.0026
SHAF	-28205.43	0.0001
BOAS	-1560.25	0.1494
CEAS	-8342.713	0.0011
SEAS	-9671.238	0.0034
EQAS	69.04606	0.9671
AUST	1141.46	0.3434

Table A4A2: Mann-Kendall test results for trend analysis across GFED regions from 2002 to 2018. Regions with significant trends are in bold (NHAF, SHAF, CEAS, SEAS); the remaining ten regions show insignificant trends.

904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931

References

[Larsen, W.A. and McCleary, S.J., 1972. The use of partial residual plots in regression analysis. Technometrics, 14\(3\), pp.781-790.](#)

Aldersley, A., Murray, S.J., Cornell, S.E., 2011. Global and regional analysis of climate and human drivers of wildfire. *Sci. Total Environ.* 409, 3472–3481.

Andela, N., Morton, D.C., Giglio, L., Chen, Y., van der Werf, G.R., Kasibhatla, P.S., DeFries, R.S., Collatz, G.J., Hantson, S., Kloster, S., 2017. A human-driven decline in global ~~burned~~[burnt](#) area. *Science* 356, 1356–1362.

Archibald, S., 2016. Managing the human component of fire regimes: lessons from Africa. *Philos. Trans. R. Soc. B Biol. Sci.* 371, 20150346. <https://doi.org/10.1098/rstb.2015.0346>

~~[Barnes, C., Boulanger, Y., Keeping, T., Gachon, P., Gillett, N., Haas, O., Wang, X., Roberge, F., Kew, S., Heinrich, D., 2023. Climate change more than doubled the likelihood of extreme fire weather conditions in eastern Canada-MacCarthy J, Tyukavina A, Weisse MJ, Harris N, Glen E. Extreme wildfires in Canada and their contribution to global loss in tree cover and carbon emissions in 2023. Global Change Biology. 2024;30\(6\):e17392.](#)~~

[Bergado, J.R., Persello, C., Reinke, K., Stein, A., 2021. Predicting wildfire burns from big geodata using deep learning. *Saf. Sci.* 140, 105276.](#)

Formatted: Line spacing: single

Formatted: Normal

Formatted: Font color: Auto

Bistinas, I., Harrison, S.P., Prentice, I.C., Pereira, J.M.C., 2014. Causal relationships versus emergent patterns in the global controls of fire frequency. *Biogeosciences* 11, 5087–5101.

Blouin, K.D., Flannigan, M.D., Wang, X., Kochtubajda, B., 2016. Ensemble lightning prediction models for the province of Alberta, Canada. *Int. J. Wildland Fire* 25, 421–432.

Bowman, D.M., Kolden, C.A., Abatzoglou, J.T., Johnston, F.H., van der Werf, G.R., Flannigan, M., 2020. Vegetation fires in the Anthropocene. *Nat. Rev. Earth Environ.* 1, 500–515.

Bowman, D.M., O'Brien, J.A., Goldammer, J.G., 2013. Pyrogeography and the global quest for sustainable fire management. *Annu. Rev. Environ. Resour.* 38, 57–80.

Bowman, D.M., Williamson, G.J., Abatzoglou, J.T., Kolden, C.A., Cochrane, M.A., Smith, A.M., 2017. Human exposure and sensitivity to globally extreme wildfire events. *Nat. Ecol. Evol.* 1, 0058.

Brown, P.T., Hanley, H., Mahesh, A., Reed, C., Strenfel, S.J., Davis, S.J., Kochanski, A.K., Clements, C.B., 2023. Climate warming increases extreme daily wildfire growth risk in California. *Nature* 621, 760–766.

Callen, T., 2008. What is gross domestic product. *Finance Dev.* 45, 48–49.

Cardil, A., Vega-García, C., Ascoli, D., Molina-Terrén, D.M., Silva, C.A., Rodrigues, M., 2019. How does drought impact burnedburnt area in Mediterranean vegetation communities? *Sci. Total Environ.* 693, 133603.

Carmona-Moreno, C., Belward, A., Malingreau, J.-P., Hartley, A., Garcia-Alegre, M., Antonovskiy, M., Buchshtaber, V., Pivovarov, V., 2005. Characterizing interannual variations in global fire calendar using data from Earth observing satellites. *Glob. Change Biol.* 11, 1537–1555.

Cary, G.J., Keane, R.E., Gardner, R.H., Lavorel, S., Flannigan, M.D., Davies, I.D., Li, C., Lenihan, J.M., Rupp, T.S., Mouillot, F., 2006. Comparison of the Sensitivity of Landscape-fire-succession Models to Variation in Terrain, Fuel Pattern, Climate and Weather. *Landsc. Ecol.* 21, 121–137. <https://doi.org/10.1007/s10980-005-7302-9>

Chen, Y., Hall, J., Van Wees, D., Andela, N., Hantson, S., Giglio, L., Van Der Werf, G.R., Morton, D.C., Randerson, J.T., 2023. Multi-decadal trends and variability in burnedburnt area from the fifth version of the Global Fire Emissions Database (GFED5). *Earth Syst. Sci. Data* 15, 5227–5259.

Chuvieco, E., Pettinari, M.L., Koutsias, N., Forkel, M., Hantson, S., Turco, M., 2021. Human and climate drivers of global biomass burning variability. *Sci. Total Environ.* 779, 146361.

Clarke, H., Tran, B., Boer, M.M., Price, O., Kenny, B., Bradstock, R., 2019. Climate change effects on the frequency, seasonality and interannual variability of suitable prescribed burning weather conditions in south-eastern Australia. *Agric. For. Meteorol.* 271, 148–157.

CoreTeam, Rd., 2014. Vienna: R foundation for statistical computing, 2014.

Cunningham, C.X., Williamson, G.J., Bowman, D.M.J.S., 2024. Increasing frequency and intensity of the most extreme wildfires on Earth. *Nat. Ecol. Evol.* 8, 1420–1425. <https://doi.org/10.1038/s41559-024-02452-2>

Davies, K.W., Bates, J.D., Svejcar, T.J., Boyd, C.S., 2010. Effects of long-term livestock grazing on fuel characteristics in rangelands: an example from the sagebrush steppe. *Rangel. Ecol. Manag.* 63, 662–669.

de Jong, M.C., Wooster, M.J., Kitchen, K., Manley, C., Gazzard, R., McCall, F.F., 2016. Calibration and evaluation of the Canadian Forest Fire Weather Index (FWI) System for improved wildland fire danger rating in the United Kingdom. *Nat. Hazards Earth Syst. Sci.* 16, 1217–1237.

DeWilde, L., Chapin, F.S., 2006. Human Impacts on the Fire Regime of Interior Alaska: Interactions among Fuels, Ignition Sources, and Fire Suppression. *Ecosystems* 9, 1342–1353. <https://doi.org/10.1007/s10021-006-0095-0>

Doerr, S.H., Santín, C., 2016. Global trends in wildfire and its impacts: perceptions versus realities in a changing world. *Philos. Trans. R. Soc. B Biol. Sci.* 371, 20150345.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leita, P.J., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36, 27–46.

Dwyer, E., Pinnock, S., Grégoire, J.-M., Pereira, J.M.C., 2000. Global spatial and temporal distribution of vegetation fire as determined from satellite observations. *Int. J. Remote Sens.* 21, 1289–1302.

Earl, N., Simmonds, I., 2018. Spatial and temporal variability and trends in 2001–2016 global fire activity. *J. Geophys. Res. Atmospheres* 123, 2524–2536.

Fang, L., Yang, J., Zu, J., Li, G., Zhang, J., 2015. Quantifying influences and relative importance of fire weather, topography, and vegetation on fire size and fire severity in a Chinese boreal forest landscape. *For. Ecol. Manag.* 356, 2–12.

Flannigan, M.D., Krawchuk, M.A., de Groot, W.J., Wotton, B.M., Gowman, L.M., 2009. Implications of changing climate for global wildland fire. *Int. J. Wildland Fire* 18, 483–507.

Forkel, M., Andela, N., Harrison, S.P., Lasslop, G., Van Marle, M., Chuvieco, E., Dorigo, W., Forrest, M., Hantson, S., Heil, A., 2019a. Emergent relationships with respect to burned area in global satellite observations and fire-enabled vegetation models. *Biogeosciences* 16, 57–76.

Forkel, M., Dorigo, W., Lasslop, G., Chuvieco, E., Hantson, S., Heil, A., Teubner, I., Thonicke, K., Harrison, S.P., 2019b. Recent global and regional trends in burned area and their compensating environmental controls. *Environ. Res. Commun.* 1, 051005.

Forrest, M., Hetzer, J., Billing, M., Bowring, S.P., Koszcor, E., Oberhagemann, L., Perkins, O., Warren, D., Arrogante-Funes, F., Thonicke, K., 2024. Understanding and simulating cropland and non-cropland burning in Europe using the BASE (Burnt Area Simulator for Europe) model. *EGUsphere* 2024, 1–55.

Fosberg, M.A., Cramer, W., Brovkin, V., Fleming, R., Gardner, R., Gill, A.M., Goldammer, J.G., Keane, R., Koehler, P., Lenihan, J., 1999. Strategy for a fire module in dynamic global vegetation models. *Int. J. Wildland Fire* 9, 79–84.

Gallardo, M., Gómez, I., Vilar, L., Martínez-Vega, J., Martín, M.P., 2016. Impacts of future land use/land cover on wildfire occurrence in the Madrid region (Spain). *Reg. Environ. Change* 16, 1047–1061.

997 Haas, O., Prentice, I.C., Harrison, S.P., 2022. Global environmental controls on wildfire burnt area, size, and intensity. *Environ.*
998 *Res. Lett.* 17, 065004.

999 Hantson, S., Andela, N., Goulden, M.L., Randerson, J.T., 2022. Human-ignited fires result in more extreme fire behavior and
1000 ecosystem impacts. *Nat. Commun.* 13, 2717.

1001 Hantson, S., Arneth, A., Harrison, S.P., Kelley, D.I., Prentice, I.C., Rabin, S.S., Archibald, S., Mouillot, F., Arnold, S.R.,
1002 Artaxo, P., 2016. The status and challenge of global fire modelling. *Biogeosciences* 13, 3359–3375.

1003 Hantson, S., Kelley, D.I., Arneth, A., Harrison, S.P., Archibald, S., Bachelet, D., Forrest, M., Hickler, T., Lasslop, G., Li, F.,
1004 Mangeon, S., Melton, J.R., Nieradzik, L., Rabin, S.S., Prentice, I.C., Sheehan, T., Sitch, S., Teckentrup, L., Voulgarakis, A.,
1005 Yue, C., 2020. Quantitative assessment of fire and vegetation properties in simulations with fire-enabled vegetation models
1006 from the Fire Model Intercomparison Project. *Geosci. Model Dev.* 13, 3299–3318. <https://doi.org/10.5194/gmd-13-3299-2020>

1007 Hantson, S., Lasslop, G., Kloster, S., Chuvieco, E., 2015. Anthropogenic effects on global mean fire size. *Int. J. Wildland Fire*
1008 24, 589–596.

1009 Jain, P., Barber, Q.E., Taylor, S., Whitman, E., Acuna, D.C., Boulanger, Y., Chavardès, R.D., Chen, J., Englefield, P.,
1010 Flannigan, M., 2024. Canada Under Fire—Drivers and Impacts of the Record-Breaking 2023 Wildfire Season. *Authorea Prepr.*

1011 Jones, M.W., Abatzoglou, J.T., Veraverbeke, S., Andela, N., Lasslop, G., Forkel, M., Smith, A.J., Burton, C., Betts, R.A., van
1012 der Werf, G.R., 2022. Global and regional trends and drivers of fire under climate change. *Rev. Geophys.* 60, e2020RG000726.

1013 Joshi, J., Sukumar, R., 2021. Improving prediction and assessment of global fires using multilayer neural networks. *Sci. Rep.*
1014 11, 3295.

1015 Juli, G., Jon, E., Dylan, W., 2017. Flammability as an ecological and evolutionary driver. *J. Ecol.* 105.

1016 Kelley, D.I., Prentice, I.C., Harrison, S.P., Wang, H., Simard, M., Fisher, J.B., Willis, K.O., 2013. A comprehensive
1017 benchmarking system for evaluating global vegetation models. *Biogeosciences* 10, 3313–3340.

1018 Kelly, L.T., Fletcher, M.-S., Menor, I.O., Pellegrini, A.F., Plumanns-Pouton, E.S., Pons, P., Williamson, G.J., Bowman, D.M.,
1019 2023. Understanding Fire Regimes for a Better Anthropocene. *Annu. Rev. Environ. Resour.* 48.

1020 Kendall, M., 1975. *Multivariate analysis.* Charles Griffin.

1021 Kloster, S., Mahowald, N.M., Randerson, J.T., Thornton, P.E., Hoffman, F.M., Levis, S., Lawrence, P.J., Feddema, J.J.,
1022 Oleson, K.W., Lawrence, D.M., 2010. Fire dynamics during the 20th century simulated by the Community Land Model.
1023 *Biogeosciences* 7, 1877–1902.

1024 Knorr, W., Arneth, A., Jiang, L., 2016. Demographic controls of future global fire risk. *Nat. Clim. Change* 6, 781–785.

1025 Knorr, W., Kaminski, T., Arneth, A., Weber, U., 2014. Impact of human population density on fire frequency at the global
1026 scale. *Biogeosciences* 11, 1085–1102.

1027 Koubi, V., 2019. Sustainable development impacts of climate change and natural disaster. *Backgr. Pap. Prep. Sustain. Dev.*
1028 *Outlook.*

1029 Kraaij, T., Baard, J.A., Arndt, J., Vhengani, L., Van Wilgen, B.W., 2018. An assessment of climate, weather, and fuel factors
1030 influencing a large, destructive wildfire in the Knysna region, South Africa. *Fire Ecol.* 14, 1–12.

1031 Krawchuk, M.A., Moritz, M.A., Parisien, M.-A., Van Dorn, J., Hayhoe, K., 2009. Global pyrogeography: the current and
1032 future distribution of wildfire. *PloS One* 4, e5102.

1033 Kuhn-Régnier, A., Voulgarakis, A., Nowack, P., Forkel, M., Prentice, I.C., Harrison, S.P., 2021. The importance of antecedent
1034 vegetation and drought conditions as global drivers of burnt area. *Biogeosciences* 18, 3861–3879. [https://doi.org/10.5194/bg-](https://doi.org/10.5194/bg-18-3861-2021)
1035 18-3861-2021

1036 Le Page, Y., Pereira, J.M.C., Trigo, R., Da Camara, C., Oom, D., Mota, B., 2008. Global fire activity patterns (1996–2006)
1037 and climatic influence: an analysis using the World Fire Atlas. *Atmospheric Chem. Phys.* 8, 1911–1924.

1038 Lehsten, V., Arneth, A., Spessa, A., Thonicke, K., Moustakas, A., 2016. The effect of fire on tree–grass coexistence in
1039 savannas: a simulation study. *Int. J. Wildland Fire* 25, 137–146.

1040 Magadzire, N., 2013. Reconstruction of a fire regime using MODIS [burnedburnt](#) area data: Charara Safari Area, Zimbabwe.
1041 Stellenbosch: Stellenbosch University.

1042 Mann, H.B., 1945. Nonparametric tests against trend. *Econom. J. Econom. Soc.* 245–259.

1043 Morvan, D., 2011. Physical phenomena and length scales governing the behaviour of wildfires: a case for physical modelling.
1044 *Fire Technol.* 47, 437–460.

1045 Mukunga, T., Forkel, M., Forrest, M., Zotta, R.-M., Pande, N., Schlaffer, S., Dorigo, W., 2023. Effect of Socioeconomic
1046 Variables in Predicting Global Fire Ignition Occurrence. *Fire* 6, 197.

1047 Nolan, R.H., Anderson, L.O., Poulter, B., Varner, J.M., 2022. Increasing threat of wildfires: the year 2020 in perspective: A
1048 *Global Ecology and Biogeography* special issue. *Glob. Ecol. Biogeogr.* 31, 1898–1905. <https://doi.org/10.1111/geb.13588>

1049 O’Brien, R.M., 2007. A caution regarding rules of thumb for variance inflation factors. *Qual. Quant.* 41, 673–690.

1050 Oliveira, S., Pereira, J.M., San-Miguel-Ayanz, J., Lourenço, L., 2014. Exploring the spatial patterns of fire density in Southern
1051 Europe using Geographically Weighted Regression. *Appl. Geogr.* 51, 143–157.

1052 Parisien, M.-A., Parks, S.A., Krawchuk, M.A., Flannigan, M.D., Bowman, L.M., Moritz, M.A., 2011. Scale-dependent
1053 controls on the area [burnedburnt](#) in the boreal forest of Canada, 1980–2005. *Ecol. Appl.* 21, 789–805.
1054 <https://doi.org/10.1890/10-0326.1>

1055 Pausas, J.G., Keeley, J.E., 2021. Wildfires and global change. *Front. Ecol. Environ.* 19, 387–395.

1056 Pausas, J.G., Ribeiro, E., 2013. The global fire–productivity relationship. *Glob. Ecol. Biogeogr.* 22, 728–736.

1057 Pechony, O., Shindell, D.T., 2010. Driving forces of global wildfires over the past millennium and the forthcoming century.
1058 *Proc. Natl. Acad. Sci.* 107, 19167–19170.

1059 Perkins, O., Matej, S., Erb, K., Millington, J., 2022. Towards a global behavioural model of anthropogenic fire: The
1060 spatiotemporal distribution of land-fire systems. *Socio-Environ. Syst. Model.* 4, 18130–18130.

1061 Perry, G.L.W., 1998. Current approaches to modelling the spread of wildland fire: a review. *Prog. Phys. Geogr.* 22, 222–245.

Pfeiffer, M., Spessa, A., Kaplan, J.O., 2013. A model for global biomass burning in preindustrial time: LPJ-LMfire (v1. 0). *Geosci. Model Dev.* 6, 643–685.

Rabin, S.S., Melton, J.R., Lasslop, G., Bachelet, D., Forrest, M., Hantson, S., Kaplan, J.O., Li, F., Mangeon, S., Ward, D.S., 2017. The Fire ~~Modeling~~Modelling Intercomparison Project (FireMIP), phase 1: Experimental and analytical protocols with detailed model descriptions. *Geosci. Model Dev.* 10, 1175–1197.

Robinne, F.-N., Burns, J., Kant, P., Flannigan, M., Kleine, M., de Groot, B., Wotton, D.M., 2018. Global fire challenges in a warming world. *IUFRO*.

Saha, M.V., Scanlon, T.M., D’Odorico, P., 2019. Climate seasonality as an essential predictor of global fire activity. *Glob. Ecol. Biogeogr.* 28, 198–210.

Saunders, D.A., Hobbs, R.J., Margules, C.R., 1991. Biological Consequences of Ecosystem Fragmentation: A Review. *Conserv. Biol.* 5, 18–32. <https://doi.org/10.1111/j.1523-1739.1991.tb00384.x>~~https://doi.org/10.1111/j.1523-1739.1991.tb00384.x~~

~~Son, R., Stacke, T., Gayler, V., Nabel, J. E. M. S., Schnur, R., Alonso, L., Requena-Mesa, C., Winkler, A. J., Hantson, S., Zaehle, S., Weber, U., and Carvalhais, N.: Integration of a Deep-Learning-Based Fire Model Into a Global Land Surface Model, *Journal of Advances in Modelling Earth Systems*, 16, e2023MS003710, <https://doi.org/10.1029/2023MS003710>, 2024.~~

Shekede, M.D., Kusangaya, S., Chavava, C.B., Gwitira, I., Chemura, A., 2024. A two-decade analysis of the spatial and temporal variations in ~~burned~~burnt areas across Zimbabwe. *PLOS Clim.* 3, e0000201.

Shikwambana, L., Kganyago, M., Xulu, S., 2022. Analysis of wildfires and associated emissions during the recent strong ENSO phases in Southern Africa using multi-source remotely-derived products. *Geocarto Int.* 37, 16654–16670.

Stott, P., 2000. Combustion in tropical biomass fires: a critical review. *Prog. Phys. Geogr. Earth Environ.* 24, 355–377. <https://doi.org/10.1177/030913330002400303>

Strand, E.K., Launchbaugh, K.L., Limb, R.F., Torell, L.A., 2014. Livestock grazing effects on fuel loads for wildland fire in sagebrush dominated ecosystems. *J. Rangel. Appl.* 1, 35–57.

Strydom, S., Savage, M.J., 2017. Potential impacts of climate change on wildfire dynamics in the midlands of KwaZulu-Natal, South Africa. *Clim. Change* 143, 385–397.

Teckentrup, L., Harrison, S.P., Hantson, S., Heil, A., Melton, J.R., Forrest, M., Li, F., Yue, C., Arneth, A., Hickler, T., Sitch, S., Lasslop, G., 2019. Response of simulated ~~burned~~burnt area to historical changes in environmental and anthropogenic factors: a comparison of seven fire models. *Biogeosciences* 16, 3883–3910. <https://doi.org/10.5194/bg-16-3883-2019>

Teixeira, J.C.M., Burton, C., Kelly, D.I., Folberth, G.A., O’Connor, F.M., Betts, R.A., Voulgarakis, A., 2023. Representing socio-economic factors in the INFERNO global fire model using the Human Development Index. *Biogeosciences Discuss.* 2023, 1–27.

Thonicke, K., Spessa, A., Prentice, I.C., Harrison, S.P., Dong, L., Carmona-Moreno, C., 2010. The influence of vegetation, fire spread and fire behaviour on biomass burning and trace gas emissions: results from a process-based model. *Biogeosciences* 7, 1991–2011.

Vilar, L., Herrera, S., Tafur-García, E., Yebra, M., Martínez-Vega, J., Echavarría, P., Martín, M.P., 2021. Modelling wildfire occurrence at regional scale from land use/cover and climate change scenarios. *Environ. Model. Softw.* 145, 105200.

Wragg, P.D., Mielke, T., Tilman, D., 2018. Forbs, grasses, and grassland fire behaviour. *J. Ecol.* 106, 1983–2001.

Wu, C., Venevsky, S., Sitch, S., Mercado, L.M., Huntingford, C., Staver, A.C., 2021. Historical and future global ~~burned~~[burnt](#) area with changing climate and human demography. *One Earth* 4, 517–530.

Xi, D.D., Taylor, S.W., Woolford, D.G., Dean, C.B., 2019. Statistical models of key components of wildfire risk. *Annu. Rev. Stat. Its Appl.* 6, 197–222.

Zhang, Y., Mao, J., Ricciuto, D.M., Jin, M., Yu, Y., Shi, X., Wulschleger, S., Tang, R., Liu, J., 2023. Global fire modelling and control attributions based on the ensemble machine learning and satellite observations. *Sci. Remote Sens.* 7, 100088.

Formatted: Line spacing: single