Global environmental controls on wildfire burnt area, size, and intensity

Olivia Haas1,2,∗, Iain Colin Prentice1,2 and Sandy P Harrison1,3

1 Leverhulme Centre for Wildfires, Environment and Society, Imperial College London, South Kensington, London SW7 2BW, United Kingdom
2 Georgina Mace Centre for the Living Planet, Department of Life Sciences, Imperial College London, Silwood Park Campus, Backhurst Road, Ascot SL5 7PY, United Kingdom
3 Geography and Environmental Science, University of Reading, Whiteknights, Reading RG6 6AH, United Kingdom

∗ Author to whom any correspondence should be addressed.
E-mail: o.haas20@imperial.ac.uk

Keywords: fire regimes, fuel loads, fire spread, landscape fragmentation, generalized linear model, vegetation models

Supplementary material for this article is available online

Abstract
Fire is an important influence on the global patterns of vegetation structure and composition. Wildfire is included as a distinct process in many dynamic global vegetation models but limited current understanding of fire regimes restricts these models’ ability to reproduce more than the broadest geographic patterns. Here we present a statistical analysis of the global controls of remotely sensed burnt area (BA), fire size (FS), and a derived metric related to fire intensity (FI). Separate generalized linear models were fitted to observed monthly fractional BA from the Global Fire Emissions Database (GFEDv4), median FS from the Global Fire Atlas, and median fire radiative power from the MCD14ML dataset normalized by the square root of median FS. The three models were initially constructed from a common set of 16 predictors; only the strongest predictors for each model were retained in the final models. It is shown that BA is primarily driven by fuel availability and dryness; FS by conditions promoting fire spread; and FI by fractional tree cover and road density. Both BA and FS are constrained by landscape fragmentation, whereas FI is constrained by fuel moisture. Ignition sources (lightning and human population) were positively related to BA (after accounting for road density), but negatively to FI. These findings imply that the different controls on BA, FS and FI need to be considered in process-based models. They highlight the need to include measures of landscape fragmentation as well as fuel load and dryness, and to pay close attention to the controls of fire spread.

1. Introduction
Fire is a necessary component of some ecosystems, but a destructive influence in others (Keeley et al 2011, Archibald et al 2013, 2018, Harrison et al 2021). In both cases, however, it represents a major influence on vegetation structure and composition. Recent changes in wildfire regimes, especially the incidence of exceptionally large fires in some regions, have raised serious concerns about how they will develop in response to projected future changes in climate and land use (Bond et al 2005, Pausas and Keeley 2009, Pausas and Ribeiro 2017).

Wildfire regimes are characterized as combinations of specific fire properties (Krebs et al 2010, Archibald et al 2013, 2018), of which the most cited are total burnt area (BA) and the sizes and intensities of individual fires. Although all these properties are understood to be jointly controlled by climatic, vegetation and land-cover conditions (Archibald et al 2013, 2018, Kelley et al 2019, Rogers et al 2020), the specifics remain unclear. This is manifested in the poor performance of process-based models designed to predict wildfire and its interactions with vegetation. Fire-enabled dynamic global vegetation models (DGVMs) have been used to simulate fire during
the historical period (e.g. Teckentrup et al 2019), and to predict how fire regimes might change in future (Kloster et al 2012, Sheehan et al 2015, Wu et al 2021). However, the Fire Model Intercomparison Project (FireMIP; Hantson et al 2016, Rabin et al 2017) showed that process-based models performed no better than empirical models in predicting BA (Forkel et al 2019a, Hantson et al 2020). Furthermore, no fire-enabled DGVM could predict global patterns in fire size (FS) better than a null model that assumed the same mean size everywhere (Hantson et al 2020). This failure has been attributed to inadequate understanding, and therefore incomplete or incorrect representation, of the key processes and their relationship to environmental factors (Forkel et al 2019a, Hantson et al 2020) in process-based models.

Several empirical studies have investigated the relationships between environmental factors and remotely sensed BA (e.g. Bistinas et al 2013, 2014, Knorr et al 2014, 2016, Andela et al 2017, Forkel et al 2017, 2019b, Kuhn-Régnier et al 2021). Global FS has been shown to vary along precipitation gradients, with greater aridity linked to larger fires (Hantson et al 2015). Human activity on the other hand strongly restricts FS through landscape fragmentation, which limits fire spread (Hantson et al 2015, Andela et al 2017, Laurent et al 2019). Intensity depends on fire type, with crown fires generally more intense than ground fires (Sugihara et al 2006, Keeley 2009, Archibald et al 2018). However, research into the controls of the geographic patterns in FS and fire intensity (FI) has been limited, mainly due to the lack of observations (Bowman 2018). Novel products from newly available satellite data sets (e.g. Laurent et al 2018, Andela et al 2019, Artés et al 2019) provide new opportunities to investigate these controls.

Four conditions must be met for a fire to start. There must be sufficient fuel; it must be dry, and therefore flammable; weather conditions must be suitable; and there must be an ignition source (Bradstock 2010). Once ignition occurs, however, other factors may influence whether the fire spreads, how rapidly it spreads, and how intense it is (Archibald et al 2013). In fire-enabled DGVMs, BA is sometimes derived as the product of the number of fires started and a mean FS. This concept intrinsically links ignition frequency to the total amount of burning, an assumption that has been challenged (Bistinas et al 2014, Harrison et al 2021). Globally, FI has been observed to increase with fire occurrence up to a threshold, also known as the intermediate fire occurrence-intensity (IFOI) hypothesis (Pausas and Ribeiro 2013, Luo et al 2017). In process-based models FS is often dependent on FI, defined as the amount of energy released by a wildfire, through rate-of-spread equations (Yue et al 2014, Hantson et al 2016). Indeed, as a key component of the Rothermel equation (Rothermel 1972), greater FI is assumed to increase the speed of fire spread, thus producing larger fires. However, this relationship has been shown to saturate in much of the world (Laurent et al 2019).

In this paper, we explored the causal relationships between environmental factors and geographic variation in fractional BA, FS and FI. These underlying relationships may differ from the emergent relationships with individual drivers. We develop generalized linear models (GLMs) for each fire property, using a common set of quantitative predictors representing vegetation, land cover, climate, and ignition sources. We explore the direction and power of the linear relationships fitted between each predictor and each fire property, when all other variables are held constant, to identify significant drivers and constraints for different aspects of the fire regime and assess their (differing) relative importance.

2. Data and methods

2.1. Fire data

BA data were derived from monthly mean fractional BA from the Global Fire Emissions Database (GFEDv4; Randerson et al 2018). This provides monthly BA at 0.25° × 0.25° resolution for the period from June 1995 to December 2016. FS was represented by median monthly FS from the Global Fire Atlas (Andela et al 2019). This provides global vector shapesfiles gridded at 500 m with the timing and location of each individual fire, along with its key properties, from 2003 to 2016. The median was used to account for the highly skewed distribution of FS. FI was approximated by a metric based on the median monthly fire radiative power (FRP) in the MCD14ML dataset (Giglio et al 2006). This provides geographic coordinates of active ~1 km × 1 km fire pixels detected by the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor onboard the Terra and Aqua satellites, from 2000 to 2017. FRP in megawatt (MW) is provided for each pixel. Only type 0 pixels (presumed vegetation fires) and with confidence >50% were retained. We divided the monthly median values by the square root of monthly median FS in the Global Fire Atlas. Although FRP has been used directly to indicate FI (Wooster and Zhang 2004, Wooster et al 2005), it is a measure of the instantaneous energy emission over the whole pixel. Normalizing by the square root of median FS provides a measure (in units of W m⁻²) more closely related to the fireline intensity of any given fire (see supplementary II available online at stacks.iop.org/ERL/17/065004/mmedia).

2.2. Predictor variables

A set of predictors representing vegetation, land cover, climate and ignition sources was identified from previous research on the drivers of wildfire regimes globally (Bistinas et al 2014, Forkel et al 2018, 2019b).
2.2.1. Vegetation, land cover and landscape fragmentation

Mean annual gross primary production (GPP) obtained from monthly outputs of the P-model (Stocker et al. 2020), a first-principles model of GPP that allows continuous acclimation of photosynthetic parameters to environmental variations in space and time (supplementary III), was used as an index of vegetation productivity. Land cover categories were used to represent vegetation types. Annual fractional shrubland, grassland and tree cover were obtained by converting the land cover types from the European Space Agency (ESA) Climate Change Initiative (CCI) Land Cover dataset, which provides annual maps from 1992 to 2015 (Li et al. 2018), to fractional cover following Poulter et al. (2015) using the conversion table in Forkel et al. (2017). Landscape fragmentation was represented by cropland cover, two landscape heterogeneity indexes and road density. This categorization is by no means exhaustive, however all three of these factors have been shown to influence fire activity and spread (Pfeiffer et al. 2013, Andela et al. 2017, Povak et al. 2018, Kelley et al. 2019, Harrison et al. 2021, Pausas and Keeley 2021). Vector ruggedness measure (VRM) and topographic position index (TPI) were used to represent topographic controls (Sappington et al. 2007). VRM captures variability in slope and aspect into a single measure, providing a measure for terrain heterogeneity, whilst TPI provides information on valleys (negative values) and ridges (positive values). We obtained 50 km aggregated information for these variables, derived from the digital elevation model products of global 250 m GMTED2010 and near-global 90 m SRTM4.1dev (Amatulli et al. 2018). Although the data layers only cover 2010, it was assumed that there would be no significant changes in VRM or TPI from 2010 to 2015 at this resolution. Fractional annual cropland cover was taken from version 3.2 of the HistorY Database of the global Environment (HYDE 3.2: Klein Goldewijk et al. 2017) database, which provides annual data from 2000 to 2017 at 0.25° × 0.25° resolution. Road density was obtained from the Global Roads Inventory Project (GRIP) database (Meijer et al. 2018), which provides average road density from 1979 to 2015 (although >50% of the data is from 2010 or later) at 0.5° × 0.5° resolution.

2.2.2. Climate

We calculated the total monthly number of dry days (DDs), monthly mean daily vapour pressure deficit (VPD), monthly mean daily diurnal temperature range (DTR) and monthly mean wind speed from the WFDE5 bias-adjusted ERA5 dataset (Cucchi et al. 2020) (supplementary 1). This provides hourly data from 1979 to 2018 at 0.5° × 0.5° resolution.

2.2.3. Ignition sources

Annual values of population density, representing the potential for human ignitions, were taken from HYDE3.2. Mean monthly lightning ground-strike density, representing potential natural ignitions, was obtained from the worldwide lightning location network (WWLLN) Global Lightning Climatology (WGLC) dataset (Kaplan and Lau 2021). This contains global monthly gridded data at 0.5° × 0.5° resolution from 2010 to 2020.

2.3. Data processing

A seasonal climatology over the common 6 years period from 2010 to 2015 was constructed for all variables with monthly data, eliminating inter-annual variability. This was the only common period for all the variables but, given the strong spatial patterns involved, we assume this is sufficiently long for our purpose. Annual variables were averaged over the 6 years. For monthly data, a single global layer was constructed from the seasonal climatology. For each grid cell, the value of the month with (on average) the maximum number of DD, the largest DTR, and the highest VPD were selected. Wind speed value was taken from the hottest month of the year (determined from the WFDE5 2 m air temperature (Cucchi et al. 2020)). For lightning, the mean value over the seasonal climatology was selected.

Additionally, we included two seasonality predictors to account for periods of high vs low productivity (seasonality of GPP) and wet vs dry seasons (seasonality of DD). These seasonality predictors were constructed from the GPP and DD climatology, respectively. To capture the seasonality of GPP, we divided the range of monthly values from the seasonal climatology by the mean value of all 12 months. The same was done to capture the seasonality of DD. For road density and the VRM, no additional processing was performed. This provided a set of 16 predictors (table 1). For the fire variables, the mean value over the seasonal climatology was selected. All variables were re-gridded to a 0.5° × 0.5° resolution grid in order to be compatible with the lowest resolution datasets and all analysis were conducted at this scale.

Predictor variables were transformed appropriately (table 1) to reduce differences in scale and improve interpretability (Schielzeth 2010). GPP was the only predictor with zero values. We applied logarithmic transformation to GPP, ignoring zero values in the model. Since we assume no fire occurs where GPP was zero, the missing values were re-coded as zero afterwards. Square root transformation was applied to predictors representing densities, for which zero values were assumed to be meaningful. To reduce distortion introduced by a few (extremely rare) very large values, FS and intensity were transformed to z-scores:
Table 1. Summary of predictor and fire response variables.

| Predictors | Data source | Abbreviation | Transformation |
|------------|-------------|--------------|---------------|
| Vegetation, land cover and landscape fragmentation | | | |
| Annual gross primary production (g C m\(^{-2}\) a\(^{-1}\)) | P-model (Stocker et al 2020) | GPP | Logarithmic |
| Gross primary production seasonality (unitless) | P-model (Stocker et al 2020) | GPP_seasonality | Logarithmic |
| Fractional shrubland cover | ESA CCI Landcover | Shrub | None |
| Fractional grassland cover | ESA CCI Landcover | Grass | None |
| Fractional tree cover | ESA CCI Landcover | Tree | None |
| Road density (km\(^{-2}\)) | GRIP (Meijer et al 2018) | Roads | Square root |
| Fractional cropland cover | HYDE 3.2 (Klein Goldewijk et al 2017) | Crop | None |
| Vector ruggedness measure | (Amatulli et al 2018) | VRM | None |
| Topographic position index | (Amatulli et al 2018) | TPI | None |
| Climate | | | |
| Maximum monthly number of dry days | WFDE5 (Cucchi et al 2020) | DD | Logarithmic |
| Seasonality of monthly number of dry days (unitless) | WFDE5 (Cucchi et al 2020) | DD_seasonality | Logarithmic |
| Maximum mean monthly vapour pressure deficit (Pa) | WFDE5 (Cucchi et al 2020) | VPD | Logarithmic |
| Maximum mean monthly diurnal temperature range (K) | WFDE5 (Cucchi et al 2020) | DTR | Logarithmic |
| Mean wind speed of the hottest month (m s\(^{-1}\)) | WFDE5 (Cucchi et al 2020) | Wind | Logarithmic |
| Ignition sources | | | |
| Population density (km\(^{-2}\)) | HYDE 3.2 (Klein Goldewijk et al 2017) | Popd | Square-root |
| Mean monthly lightning ground-strikes (km\(^{-2}\)) | WGLC WWLLN (Kaplan et al 2021) | Light | Square-root |
| Fire variables | | | |
| Monthly mean burnt area (fraction) | GFEDv4 (Randerson et al 2018) | BA | None (logit link function) |
| Monthly median fire size (km\(^2\)) | Global Fire Atlas (Andela et al 2019) | FS | Min–max normalized (log link function) |
| Monthly median fire intensity (W km\(^{-3}\)) | MCD14ML (Giglio et al 2006) | FI | Median FRP divided by square root of median FS, min–max normalized (log link function). |

\[ z_i = \frac{(x_i - \mu_i)}{\sigma_i} \]  \hspace{1cm} (1)

where \(\mu_i\) is the mean and \(\sigma_i\) is the standard deviation of the variable \(x_i\), and values of \(x_i\) yielding \(z\)-scores >3 or <-3 were excluded (Shiffler 1988). The remaining values were scaled using the min–max transformation:

\[ x_{ni} = \frac{[x_i - \min(x_i)]}{\max(x_i) - \min(x_i)} \]  \hspace{1cm} (2)

To confine the range of the transformed \(x_{ni}\) values to the interval (0, 1). This scaling helps interpretability but does not interfere with the statistical distribution of values (Justszczak et al 2002).

2.4. Statistical modelling

GLMs have previously been used to model BA (Lehsten et al 2010, Bistinas et al 2014) because they provide highly interpretable results (Nelder and Wedderburn 1972, McCullagh and Nelder 1989). GLMs are useful because they (a) handle response variables with highly non-Gaussian error distributions without the problems introduced e.g. by log-transformation of response variables, (b) are embedded in a well-established (multiple regression) framework, which allows quantification of the independent effects of multiple predictors even if they are partially correlated with one another, and (c) generate partial residual plots showing the effect of each predictor while the others are held constant (Larsen and McCleary 1972). The relative importance of each predictor was assessed using absolute \(t\)-values (the fitted regression coefficient of each predictor divided by its standard error) calculated with the package caret in R. These \(t\)-values are unitless and scale-invariant,
and in addition to providing a simple measure of the importance of each predictor (Grömping 2015), are directly related to the coefficient of determination ($R^2$): the squared $t$-value of each predictor $x_i$ represents the reduction in $R^2$ that occurs when that predictor is removed from the model (Darlington 1968, Bring 1996). Thus, the larger the $t$-value of a predictor, the more variance it explains. Variance inflation factors (VIFs) were examined to evaluate multicollinearity among predictors. A VIF of 1 indicates no collinearity; we eliminated predictors with VIF values >5 (O’Brien 2007). We used normalized mean error (NME) to evaluate the performance of the models, following Kelley et al (2013). NME is a standard metric to assess global fire model performances, allowing direct comparison with these other models (Kelley et al 2019, Hantson et al 2020). NME is defined as:

$$\text{NME} = \frac{\sum A_i |\text{obs}_i - \text{sim}_i|}{\sum A_i |\text{obs}_i - \text{obs}|}$$

where the difference between observations (obs) and simulation (sim) are summed over all cells ($i$) weighted by cell area ($A_i$) and normalized by the average distance from the mean of the observations (obs). An NME score has no upper limit, but smaller values signify better performance with a zero-value meaning perfect fit between observed and simulated values. All analysis were conducted in R using the stats, car, benchmarkMetrics and visreg packages.

Fractional BA was equated with the probability of burning, ranging between 0 and 1. We assumed this probability to follow a quasi-binomial distribution, and applied the logit link function. The distribution of FS and intensity is hypothesized to follow a power law (Kumar et al 2011, Hantson et al 2016). We assumed a quasi-Poisson distribution for both variables (due to their large overdispersion, with many very small values) and applied the log link function.

All three models were initially constructed using all 16 predictors (see supplementary IV). Population density has been hypothesized to indicate landscape fragmentation (Bistinas et al 2014, Knorr et al 2014) but is highly correlated with road density ($r = 0.31$). Therefore, we made two additional model runs, with all predictor variables excluding either population density or road density, to assess how much of the variance was explained by these predictors when the other is not present (see supplementary V). VPD, the absolute difference in water vapour content of the air and the water holding capacity of the atmosphere, influences long-term plant growth and photosynthesis (Park Williams et al 2013, Abatzoglou et al 2016, Grossiord et al 2020) but GPP seasonality is correlated with VPD ($r = 0.71$). We therefore made two separate model runs, using all predictors but excluding either GPP seasonality or VPD (see supplementary VI). From these runs, we present the best model for each fire property. Only predictors with $p < 10^{-5}$ were retained, resulting in three final models that includes only predictors with substantial explanatory power.

### 3. Results

All 16 predictors showed significant relationships with BA (table S2). GPP ($t = 63.00$), DD ($t = 70.23$) and DD seasonality ($t = 59.26$), grassland cover ($t = 52.91$), and VPD ($t = 39.11$) showed strong positive relationships with BA (figure 1). Shrubland cover ($t = 26.35$), DTR ($t = 19.82$) and TPI ($t = 18.86$) also showed positive relationships with BA. Lightning ($t = 12.35$) and population density ($t = 10.64$) showed positive relationships with BA (the relationship with population density however became negative when road density was excluded). Road density ($t = -37.32$), VRM ($t = -21.39$), tree cover ($t = -18.74$), cropland ($t = -10.05$) and wind ($t = -6.80$) showed negative relationships with BA (figure 1).

Eleven of the original 16 predictors were retained in the final FS model (table 2). GPP, grassland cover, TPI and population density were statistically insignificant, and seasonality of DD did not meet the $t$-value threshold (table S2). DTR ($t = 14.46$), and wind speed ($t = 14.41$), followed in importance by DD ($t = 11.16$), shrubland cover ($t = 7.61$), lightning ($t = 5.50$) GPP seasonality ($t = 5.18$) and VPD ($t = 5.26$) showed positive relationships with FS (figure 1). Cropland cover ($t = -22.42$), road density ($t = -16.47$), VRM ($t = 5.78$) and tree cover ($t = -5.25$) showed negative relationships with FS (figure 1).

Only nine of the original 16 predictors were retained in the final model for FI (table 2). DTR, wind speed cropland cover and TPI were statistically insignificant and VRM and seasonality of GPP did not meet the $t$-value threshold (table S2). Tree cover ($t = 9.08$) and road density ($t = 8.58$) showed positive relationships with FI, whereas VPD ($t = -47.55$), GPP ($t = -18.47$), DD ($t = -16.18$) and DD seasonality ($t = -14.27$) population density ($t = -13.49$) and lightning ($t = -7.15$) all showed negative relationships (figure 1).

Excluding population density did not change the significance or sign of any relationship with BA, FS or FI. Excluding road density led to a weak ($p = 0.05$) negative relationship between population density and BA, a highly significant negative relationship with FS, but no change for FI (table S3). Excluding VPD lead the seasonality of GPP to become insignificant with BA and to exhibit a weak ($p = 0.06$) positive relationship with FS. For FI, the exclusion of VPD lead to a large increase in the strength of the seasonality of GPP (from $t = 3.32$ to $t = 27.67$) (table S4).

The strongest relationships differed between models. BA was driven above all by variables related...
to fuel availability (GPP and grass cover: positive relationships) and dryness (DD and the seasonality of DD: positive relationships) (figure 2). Cropland cover and road density (negative relationships) were the two strongest drivers of FS (positive relationships), followed closely by DTR and wind speed (positive relationships) (figure 2). VPD and GPP showed the two strongest relationships with FI (figure 2), both being negative relationships.

The predictive power of the BA model ($R^2 = 0.69$) was substantially greater than that for FS ($R^2 = 0.29$) or FI ($R^2 = 0.27$), suggesting these last two properties are influenced by factors not included in the analysis. Despite modest $R^2$ values, however, all three models capture the broad geographic patterns of the observations (figure 3). We obtained NME scores of 0.51, 0.86 and 0.86 for the BA, FS and FI models, respectively. Our BA and FS models outperformed all models included in the Fire Intercomparison Project, which had NME scores ranging from 0.60 to 1.10 for BA and 0.96–0.98 for FS (Hantson et al 2020). Although no direct comparison with the FireMIP models is possible for FI, this model performed as well as the FS model. All three models perform better than a null mean model, which assumes the observational mean globally and produces an NME score of 1. All three GLMs tend to compress the reconstructed range of values, leading to apparent over- (under-) prediction at the low (high) extremes. However, the partial residual plots (figure 1) do not show systematic biases suggesting that the apparent compression may simply reflect the highly stochastic nature of wildfire. The fitted models represent an estimate of the most probable outcome, whereas the observations are what actually happened—including some exceptionally large and/or intense fires, but also many grid cells where no fires were detected. The observations also have a systematic low-end bias, as fires smaller than a certain size are not detected by MODIS (Rotera et al 2019, Ramo et al 2021).

4. Discussion

We have shown that the environmental controls on geographic patterns of BA, FS and FI are different. Here we discuss these differences, how they might be interpreted, and their implications for improving predictive models.

BA was shown to be primarily driven by fuel availability (represented by GPP and grass cover) and dryness (particularly DD, but also VPD and DTR) and constrained by road density. These results are consistent with previous research (Bistinas et al 2014, Andela et al 2017, Forkel et al 2019a, 2019b, Kelley et al 2019, Kuhn-Régnier et al 2021) and with the IFOI hypothesis that the main drivers of fire activity are fuel availability and fire weather, varying in importance globally along the productivity gradient.
Figure 1. Partial residual plots for BA, FS and FI as functions of (GPP) gross primary production, (GPP (s)) seasonality of GPP, (shrub) shrubland cover, (grass) grassland cover, (tree) tree cover, (roads) road density, (crop) cropland, (VRM) vector ruggedness measure, (TPI) topographic position index, (DD) dry days, (DD (s)) dry days seasonality, (VPD) vapour pressure deficit, (DTR) diurnal temperature range, (wind) wind speed, (popd) population density and (light) lightning ground-strikes. Green represents vegetation and land cover predictors; purple represents fragmentation predictors; red represents climate predictors; blue represents ignition source predictors.

Figure 2. T-values for each predictor included in the final BA (A), FS (B) and FI models (C) showing relative importance of each predictor. The larger the absolute t-values of a predictor within a model, the more variance that individual predictor explains in that model. Green represents vegetation and land cover predictors; purple represents fragmentation predictors; red represents climate predictors; blue represents ignition source predictors.
In addition, both lighting and population density positively influence BA. The apparent fire-suppressing effect of population density (noted by Bistinas et al. 2014, Knorr et al. 2014, among others) is represented in our analysis by road density. Roads provide barriers to fire spread, hence the negative relationship of BA to road density. When this effect is accounted for, population density relates positively to BA. The positive relationship with TPI suggests that fractional BA increases as we move from valleys to ridges, and that ridges do not constrain BA. This result is in line with previous research suggesting that ridges do not necessary provide fire breaks, and that valleys, characterized by higher soil moisture, lower insolation, and higher terrain shading exhibit stronger control on fire spread (Povak et al. 2018). Terrain heterogeneity (VRM) however, strongly constrains BA. Although wind speed promotes fire spread, wind speed was negatively related to BA, in line with previous research showing the limited effect of wind speeds on BA at a coarse global scale (Lasslop et al. 2015).

Factors promoting fire spread, namely wind speed and dryness (particularly DTR, but also VPD) were shown to be the primary drivers of FS while cropland cover, road density, and terrain heterogeneity (VRM), all agents of landscape fragmentation, strongly constrain FS. The effect of wind speed was significant for FS. TPI was not statistically significant, suggesting that the presence of steep slopes and ridges do not necessarily control fire spread (Povak et al. 2018). This is consistent with research suggesting fuel continuity is important for FS (Viedma et al. 2009, Hantson et al. 2015, Laurent et al. 2019). In contrast with BA, FS showed no significant relationship to GPP or grassland cover—indicating fuel availability is less important for FS than BA, and that large fires tend not to occur in highly productive environments.

The FI metric was shown to be positively influenced by tree cover and road density and constrained by dryness (particularly VPD, and also DD). Lack of precipitation has been shown to limit fuel build-up (Keeley and Syphard 2017, Kuhn-Régnier et al. 2021, Pausas and Keeley 2021) and high atmospheric dryness (VPD) limits plant growth (Fu et al. 2022). The negative relationship with dryness is therefore not surprising since we expect fuel load to be an important driver of FI. Although GPP is negatively related to the FI measure, when VPD was excluded, the seasonality of GPP became the strongest driver of the FI model, an effect that did not translate into the other two models. This would suggest that sufficient atmospheric moisture along with seasonal changes in productivity, provide the dense fuel loads necessary for intense fires. These dense fuel loads are mostly restricted to areas with high tree cover, explaining the negative relationship with grassland and shrubland cover (Archibald et al. 2013, Luo et al. 2017, Archibald et al. 2018). The overall negative relationship with GPP could therefore be explained by intense fires mainly occurring in regions with a seasonal variation in productivity, more characteristic of the high latitude
forests than the more productive, tropical ones. A plausible reason for the positive relationship between FI and road density is the correlation between intense deforestation fires and the access roads required in remote forest areas. The observed negative relationship between lightning and FI is expected since frequent fire limits fuel build-up (Pausas and Ribeiro 2013, Luo et al 2017). Similarly, the negative relationship between population density and FI can be interpreted as a trade-off between the frequency and intensity of fires, probably compounded by active fire suppression in densely populated regions. The FI metric shared few strong predictors with either of the other models. It was not limited by landscape fragmentation; wind speed and DTR were insignificant, and it showed an opposite relationship with dryness to the other models. This implies that the widely assumed positive relationship between FI and FS does not hold generally. Saturation of this relationship in fire-prone ecosystems (Laurent et al 2019) has been attributed to feedback between BA and fuel connectivity, with landscape fragmentation increasing through the fire season due to previous burns, thus limiting FS. Our results are consistent with this but suggest additional factors may influence the relationship between FS and intensity. The limiting effect of dryness on FI suggests fuel loads are important, consistent with previous research (Forkel et al 2019a, Kuhn-Régnier et al 2021). Antecedent conditions and current dryness may promote fine and flammable fuels that favour large but cooler fires, while limiting build-up of the larger, more dense fuels required for intense fires (Hantson et al 2015, Kuhn-Régnier et al 2021).

Some of the original 16 predictors proved unimportant for the FS and FI models. Although all predictors included in the final models were highly significant ($p < 10^{-3}$), the partial residual plots suggest that there may be cases where the relationship is not linear (e.g. GPP), although in no case did the inclusion of a non-linear (polynomial) term in the models improve predictability. Together with the relatively low level of predictability of these two models, this suggests that other factors need to be investigated. There is growing evidence of fire regimes differing in similar biomes, as well as different biomes showing similar fire regimes—phenomena that cannot be explained without consideration of fire-adaptive plant traits (Hollingsworth et al 2013, Pausas 2015, Pausas et al 2016, Pausas and Ribeiro 2017, Archibald et al 2018). The boreal forests in North America and Eurasia are a case in point, where crown and surface fires respectively are dominant (Wooster and Zhang 2004, Wirth 2005, Sitnov and Mokhov 2018). The distribution of fire-adaptive plant traits may be crucial to understand FS and intensity (Harrison et al 2021). Unfortunately, although there is information about fire-adaptive traits for some regions, including southern Europe and eastern Australia (Tavşanoğlu and Pausas 2018, Falster et al 2021), such information is limited elsewhere; this is a key data gap.

Our results have implications for fire modelling. Process-based models generally rely on the product of fire counts and FS to simulate BA. Our results suggest that the controls on BA, FS and FI should be modelled separately. Our results also add to the growing literature suggesting the use of population density to determine number of ignitions should be reviewed (Bistinas et al 2014, Knorr et al 2014, 2016, Andela et al 2017). Our results suggest it would be better to include landscape fragmentation explicitly as a control on fire spread. Population density apparently has a globally negative effect on BA unless an alternative measure of landscape fragmentation, such as road density, is included (table S3). Thus, population density as used in current models is a surrogate for human fire suppression and landscape fragmentation (Bistinas et al 2013, Knorr et al 2014, 2016, Harrison et al 2021); any positive impact of population on ignitions will only be captured when fragmentation is explicitly considered. Some process-based models account for the emergent relationship between population and fire properties (Harrison et al 2021), whereby increasing population density leads to increased ignitions up to a threshold and further increases lead to suppression (Rabin et al 2017). We have shown that some elements of landscape fragmentation exhibit strong spatial control on fire spread that can be modelled at a global scale. In addition to factors of landscape fragmentation considered here, additional features such as deforestation, non-flammable land cover and previous burns can all fragment the landscape and thus effect fuel continuity (Laurent et al 2019, Harrison et al 2021, Pausas and Keeley 2021). There is a pressing need to investigate how each of these features of landscape fragmentation influence different aspects of the fire regime and integrate them to global wildfire modelling explicitly. Although lightning showed a positive relationship with BA and FS, the relationship was not particularly strong. Lightning also constrained FI, possibly via a trade-off between intensity and frequency. Process-based models may therefore have overestimated the importance of ignition sources generally. Indeed, fire-enabled DGVMs struggle to simulate fire starts accurately (Forkel et al 2019a).

We have represented FS using data from the Global Fire Atlas and intensity with FRP values from MCD14ML dataset. The Global Fire Atlas relies on the MODIS sensor, so the minimum size of a recorded fire is one MODIS pixel (0.21 km$^2$) and thus our data does not include fires smaller than this. However, the environmental drivers of small fires are not expected to differ from those of larger fires. Although FRP has been used as a measure of FI, it represents the radiative energy emitted and does not account for convection and conduction. FI in the Rothermel equation represents the intensity of the flaming front.
of an individual fire, whereas retrieved FRP values from MCD14ML integrate radiative energy over a 1 km × 1 km pixel. The normalization introduced here provides a partial solution to this problem and accurately highlights the regions of high FI. However, work is required to derive fireline intensity more reliably from remotely sensed products.

This analysis did not account for any inter-annual variability, such as changes in the length of the fire season or a departure the established fire regime within a given year. Though these factors have been shown to influence the different aspects of the fire regime (Pausas 2004, Pausas and Keeley 2021), we were interested in the underlying climatic, land cover and ignition spatial controls when the effect interannual variability was removed. Understanding these controls for different aspects of the fire regime is a first and crucial step to predicting how they will respond to different conditions between years and in the future. The timeframe of this analysis was relatively short, though similar to timeframes used in previous empirical analyses due to similar data availability issues. The strong spatial patterns displayed both by the fire variables and the predictors meant that key relationships could nonetheless be established. Each model reproduces the spatial patterns of regions of high BA, large FS and high FI, allowing useful conclusions to be drawn.

**Data availability statements**

All predictor variables are available using the references given in tables 1 and S1. The GFEDv4 data are available at https://daac.ornl.gov/VEGETATION/guides/fire_emissions_v4.html. The Global Fire Atlas data is available at www.globalfiredata.org/fireatlases.html. The MCD14ML dataset is available at https://earthdata.nasa.gov/earth-observation-data/near-real-time/frmrs/mcd14ml.

The data that support the findings of this study are openly available at the following URL/DOI: https://figshare.com/s/54af266b92e1d3b2316.

**Acknowledgments**

O H acknowledges support from the NERC Centre for Doctoral Training in Quantitative and Modelling skills in Ecology and Evolution (Grant No. NE/S007415/1) and from the Leverhulme Centre for Wildfires, Environment and Society (Grant No. RC-2018-023). S P H acknowledges support from the ERC-funded project GC2.0 (Global Change 2.0: Unlocking the past for a clearer future, Grant No. 694481) and from the Leverhulme Centre for Wildfires, Environments and Society. I C P acknowledges support from the ERC-funded project REALM (Re-inventing Ecosystem And Land-surface Models, Grant No. 787203) and from the Leverhulme Centre for Wildfires, Environments and Society. This work is a contribution to the LEMONTREE (Land Ecosystem Models based On New Theory, obseRvations and ExperimEnts) project, funded through the generosity of Eric and Wendy Schmidt by recommendation of the Schmidt Futures program (S P H, I C P). We thank Martin Wooster for discussion of FRP, Alexander Kuhn-Regnier and Michel Valette for sharing data and data processing scripts, and Keith Bloomfield, Ning Dong, David Orme and David Sandoval Calle for their assistance with data problems.

**Author contributions**

Concepts, strategy and interpretation were developed by O H, I C P and S P H jointly. O H performed the data curation, processing and analysis, and produced the graphics and Tables. O H wrote the original draft; S P H and I C P reviewed and edited it.

**Conflict of interest**

The authors declare that they have no conflict of interest.

**ORCID iD**

Sandy P Harrison  https://orcid.org/0000-0001-5687-1903

**References**

Abatzoglou J T, Kolden C A, Balch J K and Bradley B A 2016 Controls on interannual variability in lightning-caused fire activity in the western US Environ. Res. Lett. 11 045005

Amatulli G, Domisch S, Tuanmu M N, Parmentier B, Ranipeta A, Malczyk J and Jetz W 2018 A suite of global, cross-scale topographic variables for environmental and biodiversity modeling Sci. Data 5 1–15

Andela N, Morton D C, Giglio L, Chen Y, van der Werf G R, Kasibhatla P S, DeFries R S, Collatz G J, Hansson S and Kloster S 2017 A human-driven decline in global burned area Science 356 1356–62

Andela N, Morton D C, Giglio L, Paugam R, Chen Y, Hansson S, van der Werf G R and Randerson J T 2019 The Global Fire Atlas of individual fire size, duration, speed and direction Earth Syst. Sci. Data 11 529–52

Archibald S et al 2018 Biological and geophysical feedbacks with fire in the Earth system Environ. Res. Lett. 13 033003

Archibald S, Lehmann C E R, Gómez-Dans J L and Bradstock R A 2013 Defining pyromes and global syndromes of fire regimes Proc. Natl Acad. Sci. USA 110 6442–7

Artés T, Oon D, de Rigo D, Durrant T H, Maiani P, Libertà G and San-Miguel-Ayanz J 2019 A global wildfire dataset for the analysis of fire regimes and fire behaviour Sci. Data 6 1–11

Bistinas I, Harrison S P, Prentice I C and Pereira J M C 2014 Causal relationships versus emergent patterns in the global controls of fire frequency Biogeosciences 11 5087–101

Bistinas I, Oon D, Sá A C L, Harrison S P, Prentice I C and Pereira J M C 2013 Relationships between human population density and burned area at continental and global scales PLoS One 8 e61188

Bond W, Woodward F I and Midgley G F 2005 The global distribution of ecosystems in a world without fire New Phytol. 165 525–38
O'Brien R M 2007 A caution regarding rules of thumb for variance inflation factors Qual. Quant. 41 673–90
Park Williams A et al 2013 Temperature as a potent driver of regional forest drought stress and tree mortality Nat. Clim. Change 3 292–7
Pausas J G 2004 Changes in fire and climate in the eastern Iberian Peninsula (Mediterranean basin) Clim. Change 63 337–59
Pausas J G 2015 Bark thickness and fire regime Funct. Ecol. 29 315–27
Pausas J G and Keeley J E 2009 A burning story: the role of fire in the history of life Bioscience 59 593–601
Pausas J G and Keeley J E 2021 Wildfires and global change Front. Ecol. Environ. 19 387–95
Pausas J G, Pratt R B, Keeley J E, Jacobsen A L, Ramirez A R, Vilagrosa A, Paula S, Kaneakua-Pia I N and Davis S D 2016 Towards understanding respouting at the global scale New Phytot. 209 945–54
Pausas J G and Ribiero E 2013 The global fire–productivity relationship Glob. Ecol. Biogeogr. 22 728–36
Pausas J G and Ribiero E 2017 Fire and plant diversity at the global scale Glob. Ecol. Biogeogr. 26 889–97
Pfeiffer M, Spessa A and Kaplan J O 2013 A model for global biomass burning in preindustrial time: LPJ-LMfire (v1.0) Geosci. Model Dev. 6 643–85
Poulter B, MacLean N, Hartley A, Khlystova I, Arino O, Betts R, Bontemps S, Boettcher M, Brockmann C and Defourny P 2015 Plant functional type classification for earth system models: results from the European Space Agency’s Land Cover Climate Change Initiative Geosci. Model Dev. 8 2315–28
Povak N A, Hessburg P F and Salter R B 2018 Evidence for scale-dependent topographic controls on wildfire spread Ecosphere 9 e02443
Rabin S S et al 2017 The Fire Modeling Intercomparison Project (FireMIP), phase 1: experimental and analytical protocols with detailed model descriptions Geosci. Model Dev. 10 1175–97
Ramo R, Roteta E, Bistinas I, van Wees D, Bastarrrika A, Chuvieco E and van der Werf G R 2021 African burned area and fire carbon emissions are strongly impacted by small fires undetected by coarse resolution satellite data Proc. Natl. Acad. Sci. U.S.A. 118
Randerson J T, Chen Y, Wiggins E B, Hantson S, Andela N, Morton D C, Hall J, Giglio L and van der Werf G 2018 Development of the Global Fire Emissions Database (GFED): toward reconciliation of top-down and bottom-up constraints on fire contributions to variability and trends in carbeneous aerosol AGU Fall Meeting Abstracts 2018 pp A41E–01
Rogers B M, Balch J K, Goetz S J, Lehmann C E R and Turetsky M 2020 Focus on changing fire regimes: interactions with climate, ecosystems, and society Environ. Res. Lett. 15 030201
Roteta E, Bastarrrika A, Padilla M, Storm T and Chuvieco E 2019 Development of a Sentinel-2 burned area algorithm: Generation of a small fire database for sub-Saharan Africa Remote Sens. Environ. 222 1–17
Rothermel R C 1972 A Mathematical Model for Predicting Fire Spread in Wildland Fuels (United States: Intermountain Forest & Range Experiment Station, Forest Service)
Sappleton J M, Longshore K M and Thompson D B 2007 Quantifying landscape ruggedness for animal habitat analysis: a case study using bighorn sheep in the Mojave Desert J. Wildl. Manage. 71 1419–26
Schielzeth H 2010 Simple means to improve the interpretability of regression coefficients Methods Ecol. Evol. 1 103–13
Sheehan T, Bachelet D and Ferschweiler K 2015 Projected major fire and vegetation changes in the Pacific Northwest of the conterminous United States under selected CMIP5 climate futures Ecol. Modell. 317 16–29
Shiffer R E 1988 Maximum Z scores and outliers Am. Stat. 42 79–80
Sitnov S A and Mokhov I I 2018 A comparative analysis of the characteristics of active fires in the boreal forests of Eurasia and North America based on satellite data Izv. Atmos. Ocean. Phys. 54 966–78
Stocker B D, Wang H, Smith N G, Harrison S P, Keenan T F, Sandoval D, Davis T and Prentice I C 2020 P-model v1.0: an optimality-based light use efficiency model for simulating ecosystem gross primary production Geosci. Model Dev. 13 1545–81
Sugihara G, van Wagtendonk J W and Fites-Kaufman J 2006 Fire as an ecological process Fire in California’s Ecosystems (Berkeley, CA: University of California Press) pp 58–74
Tavangoli C and Pausas J G 2018 A functional trait database for Mediterranean Basin plants Sci. Data 5 180135
Teckentrup L, Harrison S P, Hantson S, Heil A, Melton J R, Forrest M, Li F, Yue C, Arneth A and Hickler T 2019 Response of simulated burned area to historical changes in environmental and anthropogenic factors: a comparison of seven fire models Biogeosciences 16 3883–910
Viedma O, Angeler D G and Moreno J M 2009 Landscape structural features control fire size in a Mediterranean forested area of central Spain Int. J. Wildland Fire 18 575–83
Wirth C 2005 Fire regime and tree diversity in boreal forests: implications for the carbon cycle Forest Diversity and Function (Berlin: Springer) pp 309–44
Wooster M J, Roberts G, Perry G L W and Kaufman Y J 2005 Retrieval of biomass consumption and totals from fire radiative energy release J. Geophys. Res. Atmos. 110 D24
Wooster M J and Zhang Y H 2004 Boreal forest fires burn less intensely in Russia than in North America Geophys. Res. Lett. 31 20
Wu C, Venevsky S, Sitch S, Mercado L M, Huntingford C and Staver A C 2021 Historical and future global burned area with changing climate and human demography One Earth 4 517–30
Yue C et al 2014 Modelling the role of fires in the terrestrial carbon balance by incorporating SPITFIRE into the global vegetation model ORCHIDEE—part 1: simulating historical global burned area and fire regimes Geosci. Model Dev. 7 2747–67