AGU ADVANCING EARTH AND SPACE SCIENCES

Key Points:
- Deep neural networks (DNN) can accurately predict global burnt area fraction on a daily scale
- Integration of the DNN into a physics-based land model significantly improves the estimation of biomass burnt damage in vegetation dynamics
- The DNN accounts for regional fire variations by assigning varying degrees of importance to each predictor

Integration of a Deep-Learning-Based Fire Model Into a Global Land Surface Model

Rackhun Son[1,2] , Tobias Stacke[3] , Veronika Gayler[3] , Julia E. M. S. Nabel[1,3] , Reiner Schnur[3] , Lazaro Alonso[1], Christian Requena-Mesa[1], Alexander J. Winkler[1], Stijn Hantson[4], Sönke Zaehle[1] , Ulrich Weber[1], and Nuno Carvalhais[1,5,6]

[1]Max Planck Institute for Biogeochemistry, Jena, Germany, [2]Department of Environmental Atmospheric Sciences, Pukyong National University, Busan, South Korea, [3]Max Planck Institute for Meteorology, Hamburg, Germany, [4]Faculty of Natural Sciences, Universidad del Rosario, Bogotá, Colombia, [5]ELLIS Unit Jena, Jena, Germany, [6]Departamento de Ciências e Engenharia do Ambiente, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal

Abstract Fire is a crucial factor in terrestrial ecosystems playing a role in disturbance for vegetation dynamics. Process-based fire models quantify fire disturbance effects in stand-alone dynamic global vegetation models (DGVMs) and their advances have incorporated both descriptions of natural processes and anthropogenic drivers. Nevertheless, these models show limited skill in modeling fire events at the global scale, due to stochastic characteristics of fire occurrence and behavior as well as the limits in empirical parameterizations in process-based models. As an alternative, machine learning has shown the capability of providing robust diagnostics of fire regimes. Here, we develop a deep-learning-based fire model (DL-fire) to estimate daily burnt area fraction at the global scale and couple it within JSBACH4, the land surface model used in the ICON-ESM. The stand-alone DL-fire model forced with meteorological, terrestrial and socio-economic variables is able to simulate global total burnt area, showing 0.8 of monthly correlation (r_m) with GFED4 during the evaluation period (2011–2015). The performance remains similar with the hybrid modeling approach JSB4-DL-fire ($r_m = 0.79$) outperforming the currently used uncalibrated standard fire model in JSBACH4 ($r_m = -0.07$). We further quantify the importance of each predictor by applying layer-wise relevance propagation (LRP). Overall, land properties, such as fuel amount and water content in soil layers, stand out as the major factors determining burnt fraction in DL-fire, paralleled by meteorological conditions over tropical and high latitude regions. Our study demonstrates the potential of hybrid modeling in advancing fire prediction in ESMs by integrating deep learning approaches in physics-based dynamical models.

Plain Language Summary We develop a fire-vegetation model based on a hybrid approach integrating artificial intelligence (AI) techniques into physics-based models. Given the weather conditions, vegetation states, and human factors, our model estimates daily burned area fraction. The spatiotemporal variations in burned area are closely reproduced, especially over fire-prone regions, such as Africa, South America, and Australia. Our model is able to represent regional variations in the drivers of fire occurrence, showing different importance of input predictors for different regions. This approach shows the possibilities of using deep learning (DL) models to provide in-depth fire predictions in Earth system models.

1. Introduction

Fire is one of the main natural vegetation disturbance agents, and as such, a primary interactive component in the terrestrial ecosystem. Biomass burning affects the structure and dynamics of ecological processes (McLauchlan et al., 2020). Fire emissions alter atmospheric composition of trace gases and aerosol particles (Koppmann et al., 2005; Son, Kim, et al., 2022; Son, Ma, et al., 2022), with subsequent influences on land surface albedo (López-Saldaña et al., 2015), energy budgets (F. Li et al., 2017), climate (Liu et al., 2019; Voulgarakis & Field, 2015) and global biogeochemical cycles (Carcaillet et al., 2002; Crutzen & Andreae, 1990). Present-day global carbon emissions due to fire are approximately 1.5–3.0 PgC/yr (van der Werf et al., 2017). There is ample evidence that climate change has already resulted in increased fire risk and burned area in various areas around the world, and future increases are expected due to climate change (Seidl et al., 2017; Son et al., 2021). As fires are a significant source of greenhouse gases, there is the potential for positive (Harrison et al., 2018; Kurz et al., 1995) and negative feedbacks (Ward et al., 2012). Yet, important uncertainties remain to adequately

represent fires in Earth system models (ESMs), with uncertainties in the representation of fire disturbance still dominating the overall uncertainties in the estimation of carbon fluxes from land (Hardouin et al., 2022).

Global fire models have been developed based on empirical and physical understanding of the fire process, and these have been incorporated within dynamic global vegetation models (DGVMs) (Hantson et al., 2016). In the early stage of global fire modeling, burnt area was estimated based on the amount of dry fuel and the length of fire season (Thonicke et al., 2001). The representation of frequency of fire occurrence was advanced by considering weather-driven fire risk (Lenihan, 1998). Venevsky et al. (2002) added characteristics of fire spread by adopting the Rothermel's rate-of-spread (RoS) equations (Rothermel, 1972). Based on the RoS, more advanced fire related physical representations were introduced (Pfeiffer et al., 2013; Thonicke et al., 2010) and implemented in various DGVMs (Drüke et al., 2019; Lasslop et al., 2014; Yue et al., 2016). Human activity impacts are also considered as nonlinear functions for fire ignition and suppression based on population density, gross domestic product (GDP) and land-use changes (Kloster et al., 2010; le Page et al., 2015; F. Li et al., 2013).

Although there has been remarkable progress in global fire modeling, there are still many challenges remaining to represent the fire process and fire-vegetation interactions. For instance, fire characteristics, such as the completeness of combustion and plant mortality, are not robustly parameterized to reflect differences depending on vegetation types (Lasslop et al., 2014). Uncertainties in vegetation effects on fire remain as a main drawback in DGVMs (Forkel et al., 2019). Besides, while fire modeling has advanced with more sophisticated process based representations, there is still no agreement on the optimal level of complexity for a global fire model (Hantson et al., 2016).

Deep learning (DL), as a subset of machine learning (ML), has recently been incorporated in fire studies leading to significant advances within different aspects of fire science. For instance, convolutional neural network (CNN) is a class of deep learning algorithms that utilizes convolutional layers to extract spatial features from images or grid format data sets. The fundamental concept underlying CNN lies in its utilization of convolution layers. Convolution is a mathematical operation involving a small filter to detect and capture local patterns. By leveraging CNNs, the spatial behavior of fire has been successfully captured (Hodges & Lattimer, 2019; Radke et al., 2019). The long short-term memory modeling (LSTM, see Section 2) approaches have also demonstrated their capability in predicting fire damage and duration (Z. Li et al., 2021; Liang et al., 2019). As an extension of LSTM architecture, convolutional-LSTM (Shi et al., 2015) combines the advantages of both LSTM and CNN, rendering it particularly well-suited for tasks that necessitate the simultaneous understanding temporal patterns and spatial information. The model first processes input through convolutional operations to generate feature maps, similar to traditional CNNs. Subsequently, LSTM cells take these results as input and maintain their hidden states over time, facilitating the capture of temporal dependencies. Kondylatos et al. (2022) incorporated meteorological, environmental and anthropogenic drivers into a convolutional-LSTM to comprehensively address the spatiotemporal context for wildfire danger prediction. Other studies leveraged ML/DL methods to characterize various aspects of fire occurrence, such as fire weather (Son, Kim, et al., 2022; Son, Ma, et al., 2022), lightning ignition (Coughlan et al., 2021), fire susceptibility (Zhang et al., 2021) and fuel availability (D'Este et al., 2021).

The main objective of this study is to leverage the advantages of DL in enhancing a process-based model, with a specific focus on improving biomass burnt damage simulation. To achieve this, we first develop a DL-based global fire model, consisting of three independent modules representing weather-driven fire danger, land properties, and anthropogenic effects on burnt areas. Subsequently, we integrate this model into the land model of the Icosahedral Non-hydrostatic Earth System Model (ICON-ESM), serving as a surrogate for process-based fire representation. Lastly, we apply an interpretable method on the DL model to analyze regional characteristics and the limits of our hybrid modeling approach. Compared to a previous DL surrogate fire model (Zhu et al., 2022), our study has advances in two folds: (a) we incorporate LSTM based recurrent model architecture to consider time dependent memory effects from dynamic weather and vegetation processes; and (b) our model training was based on observational data sets, except for fuel load, allowing it to be coupled with any DGVM.

2. Methodology and Data

2.1. JSBACH4 and Its Simple Fire Scheme

JSBACH4 (Jena Scheme for Biosphere-Atmosphere Coupling in Hamburg version 4); (Nabel et al., 2020; Schneck et al., 2022), which is the land surface model used in the ICON-ESM, incorporates a simple fire
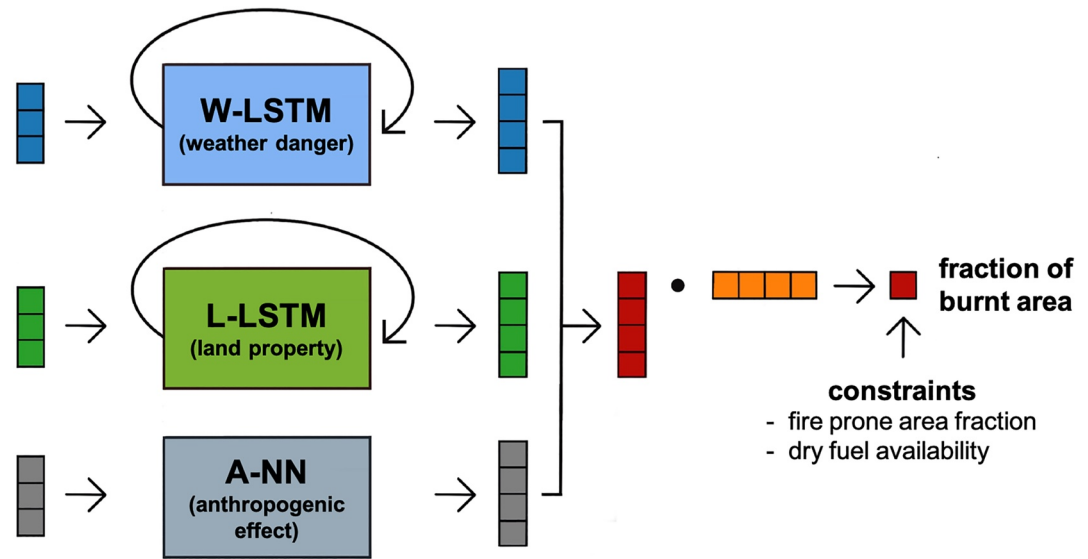
Figure 1. Flowchart of DL-fire model. An input vector with 50 predictors ([b,50], where b is the batch size) is divided into three sub-modules: 9 predictors for W-LSTM ([b,9], shown in blue vector), 23 predictors for L-LSTM ([b,23], green vector), and 18 predictors for A-NN ([b,18], gray vector). The output of each module is merged by element-wise multiplication, resulting in a dimension of 8 (red vector), which matches the number of PFTs, except for the bare land type, used in L-LSTM. The burnt fraction is finally calculated by summing up the merged output per PFT, which is obtained by taking the inner product between the output and the PFT vector (orange vector), and multiplying it with two physical constraint terms.

model implemented to estimate fire damage based on combustible fuel availability and fuel dryness (Jungclaus et al., 2022). As one of the most simple fire representations, it can be applied in any global land surface model. The primary objective of the fire scheme is more focused on the disturbance effect on natural land cover changes, rather than fire occurrence and interactions, limiting its role on vegetation dynamics and carbon cycling in ecosystems. Instead, the previous version of JSBACH (JSBACH3.2) used the SPITFIRE fire model (Thonicke et al., 2010) to simulate global fire regimes, but it has not yet been implemented in JSBACH4.

In the simple fire scheme, the fuel availability is represented by the total litter density (L) and is compared to the litter threshold (L_0). The fuel dryness is estimated from surface level air relative humidity ($\overline{rh_t}$) smoothed with a persistence factor (p) at each time step (Equation 1). When the humidity decreases lower than its threshold (rh_0), the fraction of burned area (FBA) is assumed to linearly increase as humidity decreases:

$$\overline{rh_t} = \overline{rh_{t-1}} \times p + \min(rh_t, 100) \times (1-p), \quad p = 0.95^{\frac{1}{48}} \tag{1}$$

$$FBA = FBA_{min} + \frac{1}{\tau} \times \frac{rh_0 - \overline{rh_t}}{rh_0} \quad \text{if } L > L_0 \text{ and } rh < rh_0 \quad \text{otherwise } 0 \tag{2}$$

where, τ denotes the frequency of fire occurrence: set as 6 years for woody and 2 years for grass type vegetation. The burnt fraction is computed on a daily time step and is utilized to update the relocation of carbon and nitrogen, assuming that all vegetation within the burned area perishes, notwithstanding the fact that only a part of it undergoes complete combustion. We take the simple fire model (hereafter referred to as JSB4-simple) as the baseline for model evaluation. The standalone version of JSBACH4 is used to run JSB4-simple with the default configurations as used in JSBACH3.2 and described in Reick et al. (2021).

2.2. Deep Learning (DL) Fire Model

The deep learning fire model (DL-fire) is composed of three modules: weather-driven fire danger, land properties and anthropogenic effects (Figure 1). The development of the modules for weather danger (W-LSTM) and land properties (L-LSTM) are based on the long short-term memory network approach (LSTM) (Hochreiter & Schmidhuber, 1997). LSTM is an advanced recursive neural network to handle temporal dynamic behaviors from

sequential data. The key aspect of the LSTM approach is its memory unit, called cell state that maintains information on states over timesteps, and its update is regulated by input and forget gates:

$$i_t = a_{\text{sigmoid}}(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

$$f_t = a_{\text{sigmoid}}(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{4}$$

$$o_t = a_{\text{sigmoid}}(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{6}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{7}$$

$$h_t = o_t \odot \tanh(c_t) \tag{8}$$

where i, f, o denote the input gate, forget gate, output gate and c, h denote cell and hidden state. The terms W and b refer to the weight matrices and bias vectors for each gate and the cell states (e.g., W_i is the matrix of weights for the input gate), a_{sigmoid} is the sigmoid function, $\tan h$ is the hyperbolic tangent function, and $\odot$ denotes the element-wise product of vectors. The output dimension of the LSTM is set to 8 to be equal with the number of the plant functional types (PFTs), except for the bare land type.

The anthropogenic effect module uses two layers of fully connected feed-forward network:

$$h_t = \text{act}(W_1 \cdot x_t + b_1) \tag{9}$$

$$o_t = W_2 \cdot h_t + b_2 \tag{10}$$

where x denotes the input vector for anthropogenic variables and h is hidden layer vectors. The W and b terms are weight matrices and bias vectors for the input and hidden vectors. The function act represents a nonlinear transformation using a *softplus* function (Dugas et al., 2000) in this study. The vector o is the output vector of the anthropogenic effect module that has the same dimension as the outputs of the W-LSTM and L-LSTM modules.

The final output, the fraction of burned area, is the computed sum of all PFTs, except for the bare land type, after multiplying results of the three modules and the fractions of PFTs (orange vector in Figure 1). Also, we use the fraction of bare land (f_bare) and snow (f_snow), fuel (above ground plant litter in JSBACH4) and relative humidity not only as LSTM input predictors, but also as constraints on fire occurrence and intensity:

$$\text{FBA} = \left(\sum o_w \times o_l \times o_a \times f\_\text{PFTs}\right) \times \text{fire prone area} \times \text{dry fuel availability} \tag{11}$$

$$\text{fire prone area} = 1 - f\_\text{bare} - f\_\text{snow} \tag{12}$$

$$\text{dry fuel availability} = \text{fuel}_\text{norm} \times S\left(1 - \frac{\text{rh}}{100}\right) \quad \text{if rh} > \text{rh}_0 \quad \text{otherwise } 0 \tag{13}$$

$$S(x) = \frac{1}{1 + e^{-20 \times (x - 0.5)}} \tag{14}$$

where o_w, o_l, o_a denote output vectors of W-LSTM, L-LSTM and anthropogenic effect modules and f_PFTs denotes the fractions of PFTs. We use sigmoidal curve function (S) to transform relative humidity into a non-linear space. rh_0 is the threshold of relative humidity for fire occurrence set as 60 (%), fuel_norm is normalized fuel using its maximum and minimum values during the training period (Equation 15).

2.3. Burnt Fraction

For model training and evaluation, we used daily burned area from the Global Fire Emissions Database (GFED4) (Randerson et al., 2015) and calculated the burnt fraction for each grid cell. The GFED4 burned area product is based on the Moderate Resolution Imaging Spectroradiometer (MODIS) Collection 5.1 (MCD64A1 v5.1), globally available at $0.25° \times 0.25°$ spatial resolution.

Extreme data imbalance between instances of fire and no-fire is observed over all regions (Table 1). If the data with a large proportion of no-fire instances are directly used for model training, it is highly likely to mislead

Table 1
Ratio Between Grid-Cell Level Fire/No-Fire Incidents Per Region

	Fire:no-fire	Step 1	Step 2	Ratio (step 2)
BONA	1:1313	1:301	1:1.0	300
TENA	1:412	1:61	1:1.23	50
CEAM	1:122	1:23	1:1.16	20
NHSA	1:85	1:20	1:1.02	20
SHSA	1:72	1:15	1:1.53	10
EURO	1:988	1:149	1:1.49	100
MIDE	1:1023	1:188	1:1.88	100
NHAF	1:27	1:8.4	1:1.69	5
SHAF	1:12	1:4.0	1:0.99	5
BOAS	1:721	1:128	1:1.27	100
CEAS	1:188	1:32	1:1.06	30
SEAS	1:104	1:24	1:1.19	20
EQAS	1:180	1:29	1:1.43	20
AUST	1:75	1:18	1:1.78	10

Note. The last column is for downsampling ratios used for step 2.

model outputs to converge into zero values. In order to reduce the risk of zero convergence, we adopt two strategies. We first used a Gaussian kernel with 30 days of window size to smooth the burned area (step 1 in Table 1). Subsequently, we downsample no-fire instances according to ratios in Table 1 (step 2 ratio), reducing the imbalanced ratios to be close to 1:1 for all regions.

2.4. Input Variables

The DL-fire uses 50 predictors which are divided into three sub-modules to predict burnt fraction illustrated in detail in Tables 2–4. The weather danger module (W-LSTM) uses 9 predictors, including anomalies of temperature, specific and relative air humidity. Weather variables, such as temperature, specific/relative air humidity, wind speed and precipitation, are obtained from ERA5 (Hersbach et al., 2020) and lightning climatology is based on a data set from the spaceborne Optical Transient Detector (OTD) and Lightning Imaging Sensor (LIS) on the Tropical Rainfall Measuring Mission (TRMM) satellite (Cecil et al., 2014). The anomalies are calculated by extracting daily climatology (mean values on a day of year basis) during the years 1950–2020.

The land property module (L-LSTM) takes 23 predictors including the water volumes in four soil layers which are obtained from ERA5-Land (Muñoz-Sabater et al., 2021) and the Leaf Area Index (LAI) which is derived from the collection-5 MODIS LAI product (Myneni et al., 2015). We also calculate daily anomalies for the water volumes and LAI using the above mentioned method during 1950–2020 and 2003–2020, respectively. The topographic factors, such as elevation, slope and roughness, are taken from Amatulli et al. (2018). The amount of fuel is simulated by JSB4-simple. The area distributions of plant functional types (PFTs) are obtained from Pongratz et al. (2008), given as inputs for running JSBACH4 and we remap PFTs to be nine types as outlined in Table 3.

The anthropogenic effect module (A-NN) takes into account a total of 18 predictors from five different characteristics: population density (Klein Goldewijk et al., 2017), gross domestic product (GDP) and human development

Table 2
Input Predictors for Weather Danger Module (W-LSTM) and Their Brief Ecological Description

Temperature	(High)	ERA5 (Hersbach et al., 2020)
– Temperature at 2m	– Dry fuel moisture	
– Temperature anomaly	– Increase flammability	
Humidity	(High)	
– Specific humidity	– Preserve fuel moisture	
– Specific humidity anomaly	– Impede fire ignition & spread	
– Relative humidity	(Low)	
– Relative humidity anomaly	– Increase fuel availability	
Precipitation	(High)	
– Total precipitation	– Act as a natural fire break	
	– Support plant growth	
	(Low)	
	– Contribute drought conditions	
Wind speed	(High)	
– Wind speed	– Accelerate fire progression	
	– Intensify fire behavior	
Lightning	(High)	OTD/LIS (Cecil et al., 2014)
– Lightning climatology	– Ignite fire as a natural source	

Table 3
Input Predictors for Land Property Module (L-LSTM) and Their Brief Ecological Description

Water in soil layers	(High)	ERA5
– Volume of water in 4 soil layers lv1: 0–7 cm, lv2: 7–28 cm, lv3: 28–100 cm, lv4: 100–289 cm	– Promote vegetation growth (Low)	
– Anomalies in 4 soil layers	– Raise susceptibility to ignition	
	– Indicate potential fire severity	
LAI	(High)	MODIS (Myneni et al., 2015)
– LAI	– Increase fuel loads	
– LAI anomaly	– Enhance resilience against fire	
Topography	– Differ the types of vegetation	Amatulli et al. (2018)
– Elevation	– Affect the speed and direction of fire spread	
– Slope		
– Roughness		
Fuel	(High)	JSBACH4
– Above ground plant litter	– Increase fuel loads	
Plant functional types (PFTs)	– PFTs differ fuel characteristics	Pongratz et al. (2008)
– Snow (PFT_snow)		
– Tropical evergreen trees (PFT_tet)	(Grass)	
– Tropical deciduous trees (PFT_tdt)	– Easily ignitable fuels	
– Extra-tropical evergreen trees (PFT_eet)	– Facilitate fire spread	
– Extra-tropical deciduous trees. (PFT_edt)	– Lower biomass, reducing fire intensity	
– Raingreen shrubs (PFT_rs)	(Shrubs & trees)	
– Deciduous shrubs (PFT_ds)	– Higher biomass, increasing fuel loads	
– Grass (PFT_grass)	– Contribute to the potential for longer-lasting fires	
– Bare land (PFT_bare)		

index (HDI) (Kummu et al., 2018), total road density (Meijer et al., 2018) and 14 fractions representing the state of land use (Hurtt et al., 2020).

All the input variables are regridded and aggregated to a daily timestep and 0.25° spatial resolution to be consistent with the GFED4. Except for PFT fractions constrained in the range of [0,1], we normalized predictors using maximum and minimum values of each region based on the training period ($x_{r,\text{train\_max}}$ and $x_{r,\text{train\_min}}$, where r denotes a GFED region in Figure S1 of the Supporting Information S1), ideally to be in the range of [0,1]:

$$(x - x_{r,\text{train\_min}})/(x_{r,\text{train\_max}} - x_{r,\text{train\_min}}) \tag{15}$$

2.5. Model Setup for Training and Simulation With JSBACH4

We develop 14 regional models based on GFED reference regions (Figure S1 in Supporting Information S1). To train the models, we use 12 years (2004–2015) of data considering data availability for burnt fraction and all the input predictors. We randomly select 80% of the data set from the first 7 years (2004–2010) for training and the remaining 20% are for validation during the model training stage. We apply a stratified random sampling approach to preserve the same ratios between fire/no-fire incidents. The last 5 years (2011–2015) are used for performance evaluation.

The dimension of the hidden layer is set to be 64 for all the three module architectures and dropout regularization is implemented for the anthropogenic module layers with 10% of probability to randomly inactivate neural

Table 4
Input Predictors for Anthropogenic Effect Module (A-NN) and Their Brief Ecological Description

Population	(High)	HYDE3.2 (Klein Goldewijk et al., 2017)
– Population density	– More human-caused ignition	
	(Low)	
	– Less availability of firefighting resources	
Gross domestic product (GDP)	(High)	Kummu et al. (2018)
– GDP per capita	– Facilitate the implementation of environmental policies	
	(Low)	
	– Less infrastructure against fire	
Human development index (HDI)	(High)	
– HDI	– Higher education and awareness of fire prevention	
	– Responsible land management	
	– Effective emergency response	
Road density	(High)	GRIP4 (Meijer et al., 2018)
– Total road density	– More human-caused ignition	
	– Easier access for firefighting	
Human land use state fractions	– Land use changes can alter fire regimes	LUH2 (Hurtt et al., 2020)
– Forested primary land (LU_primf)		
– Non-forested primary land (LU_primn)	(Forested land)	
– Forested secondary land (LU_secdf)	– Larger amount of biomass	
– Non-forested secondary land (LU_secdn)	(Secondary land)	
– Urban land (LU_urban)	– Fuel accumulation depends on the stage of forest succession	
– C3 annual crops (LU_c3ann)	(Urban land)	
– C4 annual crops (LU_c4ann)	– More human ignition sources	
– C3 perennial crops (LU_c3per)	– Act as a barrier to fire spread	
– C4 perennial crops (LU_c4per)	(Crop land)	
– C3 nitrogen-fixing crops (LU_c2nfx)	– Agricultural activities introduce new ignition sources	
– Managed pasture (LU_pastr)	(Pasture & rangeland)	
– Rangeland (LU_range)	– Reduce fuel loads by grazing	
– Secondary mean biomass carbon density (kg/m^2, LU_secmb)	– Well-maintained areas can act as a barriers to fire spread	
– Secondary mean age (years, LU_secma)		

network nodes. For the LSTM modules, the sequence length of training data set is set to 14 days. We use the mean square error (MSE) loss function with ADAM optimizer (Kingma & Ba, 2014) by setting the learning rate to 0.001 and batch size to 1,024. To avoid overfitting on the training data set, we stop model training after a span of 30 epochs where no further improvement is observed in the validation data set.

The DL-fire is trained without coupling to the dynamics of JSBACH4, as an offline learning approach. When the DL-fire is integrated into JSBACH4, all the land properties are provided by physics-based dynamics processes, except for topography. The other predictors are set to be forced by data sets used for model training and it allows the evaluation of simulation results from the year 2001. We perform experiments on the R2B4 ICON-grid system with spin-up time of 51 years, starting from the year 1950, and evaluate simulation results from 2001 to 2015. During the spin-up period (before the year 2001), we set all anthropogenic variables to be static at the state of 1st January 2001.

2.6. Evaluation Metrics

To quantify the performance in simulating spatial variation, we apply the normalized mean error (NME) with area weights suggested by Hantson et al. (2020):

$$\mathrm{NME} = \sum_i A_i |o_i - m_i| \,/\, \sum_i A_i |o_i - \overline{o}| \tag{16}$$

where o_i denotes the observed value, m_i the simulated value and A_i cell area at grid cell i. $\overline{o}$ is the mean of the observed values. A smaller value of NME describes better agreement with observation and zero is for perfect match between observation and model simulation. If NME is larger than 1, model performance is worse than simple prediction with statistical mean value.

We calculate the Pearson correlation coefficient between daily (r_d), monthly (r_m) and interannual (r_i) variability in predicted burnt fraction and GFED4, and the mean phase difference (MPD) to evaluate seasonal variation (Kelley et al., 2013). To quantify a distance between two phases, time unit is first transformed as an angle vector:

$$\theta_m = 2\pi(m - 1)/12 \tag{17}$$

where m denotes month (January–December). Then real (L_x) and imaginary (L_y) component vectors are calculated by:

$$L_x = \sum_m x_m \, \cos(\theta_m) \tag{18}$$

$$L_y = \sum_m x_m \, \sin(\theta_m) \tag{19}$$

The phase (P) is described by direction of the vectors (Equation 20) and MPD quantifies the phase difference by Equation 21:

$$P = \arctan(L_x/L_y) \tag{20}$$

$$\mathrm{MPD} = \frac{1}{\pi} \sum_i A_i \times \arccos\left[\cos\left(\hat{P}_i - P_i\right)\right] \,/\, \sum_i A_i \tag{21}$$

where $\hat{P}_i$ is phase from model simulation and P_i from observation at grid cell i.

2.7. Layer-Wise Relevance Propagation

To interpret the decision making process of the DL-fire model, we apply the layer-wise relevance propagation (LRP) (Bach et al., 2015) to decompose contributions from the input space. LRP computes relevance scores for each individual input by propagating relevance from the model output back through the neural network layers. While the total amount of relevance scores in each layer is kept consistent, the relevance in a layer is redistributed to the previous layer considering weights and input values, and this process repeats until getting the scores for the input layer. Here, we normalized relevance scores for each timestep so that the absolute values sum up to 1. Then we composite the normalized scores during the evaluation period to compare relative attribution with a global aspect.

3. Results

3.1. DL-Fire Model Evaluation

Globally, the predicted burnt fraction shows a good overall accordance with the GFED4 estimates during the evaluation period (Figures 2a and 2b) with a NME of 0.64 (Table 5). The pattern of seasonal cycle is also accurately captured with 0.3 of MPD and 0.73 of r_d. Monthly aggregated predictions show a comparable correlation score ($r_m = 0.80$) to that of a previous DL model (0.76) (Joshi & Sukumar, 2021), noting that the evaluation period is different for both studies. However, high fractions, especially in the second half of the years 2011 and 2012, are underestimated (Figure 2c) indicating a degrading performance skill in interannual variability ($r_i = 0.35$).

Regionally developed models vary in their performance skills. All the regional models show a NME lower than 1.0 and the best score is achieved in the northern part of South America (NHSA, 0.48), whereas NME is relatively high in regions where it shows large burnt fractions, such as Boreal North America (BONA), the southern part of South America (SHSA), the southern part of Africa (SHAF) and Central Asia (CEAS). The model for
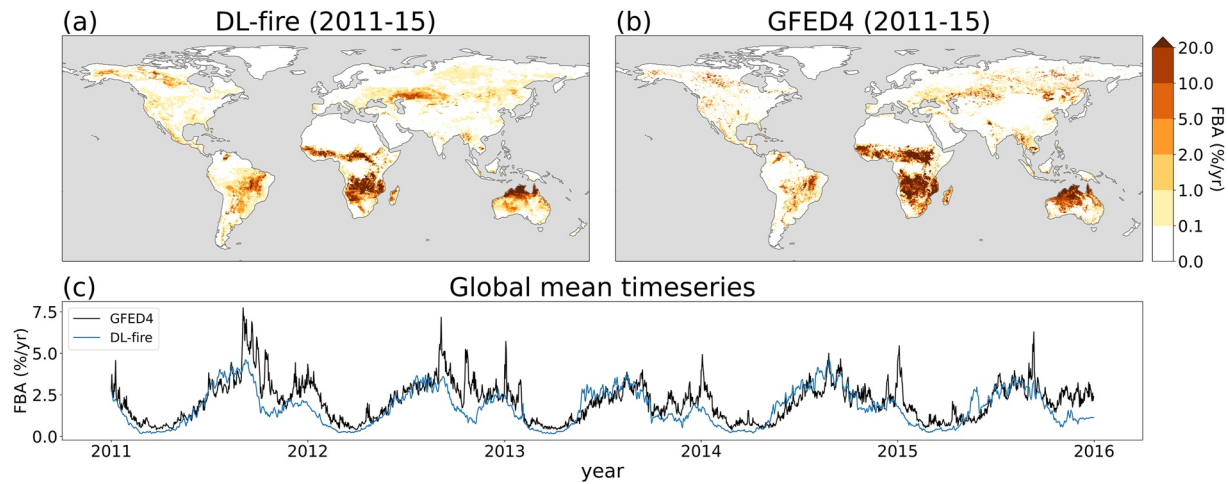
Figure 2. Spatial and temporal comparison between and GFED4 and DL-fire predictions. The maps of (a) DL-fire and (b) GFED4 visualize annual burnt fraction averaged over evaluation period (2011–2015). (c) Compares global mean of burnt fraction from GFED4 (black) and DL-fire (blue).

Central America (CEAM) shows high predictability in seasonal variation with 0.19 of MPD, and the BONA, SHSA, Africa and Equatorial Asia (EQAS) also perform well with a performance higher than 0.8 of r_d. The lowest daily correlations are obtained in the temperate North America (TENA, 0.47) and CEAS (0.41), showing underestimations in each of the fire seasons (Figures S2b and S2k in Supporting Information S1). 8 out of 14 regional models perform well on predicting interannual fire patterns with higher than 0.8 of r_i. The least interannual predictability is shown across Southeast Asia (SEAS) and SHAF ($r_i = -0.14, 0.08$) due to a lack in detecting high burnt fractions (Figures S2i and S2l in Supporting Information S1). These results, especially due to the SHAF region, cause a considerable drop in the interannual predictability at the global scale.

3.2. Coupling With JSBACH4

When the DL-fire model is coupled with JSBACH4 (JSB4-DL-fire), burnt fraction prediction skill is significantly enhanced in comparison to the simple fire model (JSB4-simple). JSB4-DL-fire improves NME score from 0.75 to 0.67 at the global scale, and NME decreases in 10 out of 14 regions (Table 6). Although burnt fractions in Africa and Siberia are underestimated, JSB4-DL-fire successfully captures the spatial variation of burnt fraction, especially across fire prone regions, such as Africa, South America, and Australia (Figure 3a).

Furthermore, burnt fractions in fuel-limited areas are improved to be close to zero in JSB4-DL-fire. JSB4-simple sets nonzero constant parameter for the minimum degree of fire damage (see Section 2), the results of JSB4-simple show higher than 0.1%/year of damage over almost all areas, including deserts and extremely cold regions (Figure 3c). Due to this over-simplified parameterization, arid areas and high latitudes, such as BONA, TENA, Europe (EURO), Middle East (MIDE) and Asia (BOAS and CEAS), show poor NME scores (2.34, 2.49, 2.06, 6.10, 1.40, and 1.39, respectively). These discrepancies are effectively addressed by JSB4-DL-fire with fuel and PFT constraints, improving NME to be lower than 1.0 across all the regions, except for MIDE.

The global spatial variation in fire seasonality is compared by visualizing the month with maximum fire damage per grid cell during the year 2001–2015 (Figure 3b). JSB4-DL-fire shows overall coincide fire season distribution with GFED4, and the best score of MPD is achieved over CEAM (0.19,

Table 5
Evaluation Metric Scores for DL-Fire

	NME	MPD	r_d	r_m	r_i
Global	0.64	0.30	0.73	0.80	0.35
BONA	0.90	0.36	0.81	0.95	0.92
TENA	0.77	0.35	0.47	0.64	0.92
CEAM	0.72	0.19	0.82	0.90	0.86
NHSA	0.48	0.31	0.74	0.85	0.85
SHSA	0.83	0.23	0.85	0.89	0.52
EURO	0.76	0.33	0.60	0.76	0.92
MIDE	0.49	0.31	0.62	0.72	0.30
NHAF	0.58	0.31	0.88	0.93	0.38
SHAF	0.96	0.33	0.90	0.94	0.08
BOAS	0.69	0.31	0.63	0.77	0.82
CEAS	0.86	0.39	0.41	0.55	0.97
SEAS	0.56	0.22	0.60	0.82	-0.14
EQAS	0.55	0.28	0.90	0.97	0.99
AUST	0.50	0.29	0.66	0.76	0.50

Table 6
Evaluation Metric Scores for JSB4-DL-Fire (JSB4-Simple)

	NME	MPD	r_d	r_m	r_i
Global	0.67 (0.75)	0.31 (0.30)	0.61 (−0.07)	0.79 (−0.07)	0.37 (0.17)
BONA	0.72 (2.34)	0.36 (0.34)	0.62 (0.45)	0.85 (0.56)	0.71 (0.44)
TENA	0.71 (2.49)	0.35 (0.28)	0.37 (0.32)	0.64 (0.48)	0.82 (0.82)
CEAM	1.53 (1.08)	0.19 (0.24)	0.70 (0.61)	0.82 (0.72)	0.62 (0.37)
NHSA	0.61 (0.68)	0.21 (0.21)	0.55 (0.61)	0.72 (0.71)	0.51 (0.53)
SHSA	0.83 (0.85)	0.21 (0.20)	0.81 (0.71)	0.89 (0.77)	0.78 (0.62)
EURO	0.70 (2.06)	0.38 (0.36)	0.29 (0.32)	0.55 (0.50)	0.34 (0.32)
MIDE	7.96 (6.10)	0.32 (0.31)	0.12 (0.61)	0.34 (0.75)	−0.12 (−0.18)
NHAF	0.58 (0.67)	0.37 (0.44)	0.75 (0.35)	0.87 (0.39)	0.80 (0.65)
SHAF	0.76 (0.82)	0.33 (0.28)	0.84 (0.80)	0.91 (0.86)	0.35 (0.14)
BOAS	0.68 (1.40)	0.35 (0.36)	0.60 (0.27)	0.78 (0.35)	0.76 (0.10)
CEAS	0.61 (1.39)	0.39 (0.32)	0.57 (−0.24)	0.67 (−0.32)	0.29 (−0.22)
SEAS	2.05 (0.88)	0.25 (0.19)	−0.02 (0.40)	−0.03 (0.54)	−0.04 (0.32)
EQAS	0.50 (0.81)	0.25 (0.26)	0.41 (0.63)	0.77 (0.74)	0.80 (0.90)
AUST	0.81 (0.72)	0.26 (0.33)	0.70 (0.48)	0.78 (0.55)	0.42 (0.62)

Table 6). Compared to JSB4-simple, the seasonal phase difference in AUST is also improved (MPD = 0.26), but JSB4-DL-fire achieves slightly increased scores in 8 out of 14 regions. Nevertheless, the most notable improvement in JSB4-DL-fire is found in temporal correlations. While the global mean of the JSB4-simple simulation has a statistically insignificant relationship with GFED4 (r_d, $r_m \approx 0$ and $r_i = 0.17$), the JSB4-DL-fire considerably increases the correlations ($r_d = 0.61$, $r_m = 0.79$, $r_i = 0.37$). We also compare their seasonality during 2011–2015 (DL evaluation period), showing that the month to month variability in JSB4-simple is highly underestimated, showing a limited range in monthly burned area values, whereas spatial and seasonal patterns of JSB4-DL-fire generally match well with GFED4 (Figure S3 in Supporting Information S1).

Regionally, the performance of JSB4-DL-fire is most marked in SHSA and SHAF (Figures 4e and 4i) with scores higher than 0.8 of r_d (Table 6). JSB4-DL-fire also effectively reduces underestimation in NHAF and AUST (Figures 4h and 4n) as well as the overestimation in BONA, BOAS and CEAS (Figures 4a, 4i, and 4k). Among 14 regions, JSB4-DL-fire enhances r_d in 9 and r_m in 12 of them. In terms of interannual variability, the biggest improvement is found in BOAS, increasing r_i from 0.1 to 0.76, whereas the variability in SEAS and MIDE are the least predictable (−0.04 and −0.12, respectively). Although JSB4-DL-fire outperforms JSB4-simple in general, in comparison to the model validation results forced by observation (Table 6), the predictability of DL-fire is degraded over almost all the regions by integrating with JSBACH4. These changes in predictability by being coupled with JSBACH4 will be further discussed in terms of JSBACH4 internal biases in the next section.

3.3. Model Interpretation

To understand how the DL fire model makes its predictions, we implement LRP for evaluating the contribution of each predictor. Globally, the fraction of bare land shows the highest absolute attribution with more than 16.3% of relevance score (Figure 5a). Its role, as a key component in identifying no or low risk of fire, is highlighted across regions, where there are large portions of arid lands or deserts, such as SHSA, MIDE, SHAF, and AUST (Figures S4e, S4g, S4i, and S4n in Supporting Information S1). Fuel load also shows a high ratio of contribution (14.1%) based on its multiple roles as a constraint (7.4%) as well as an input of L-LSTM (6.7%). The volume of water in the 4th soil layer (SWL4) counts as the 3rd key factor associated with burned fraction in that it can be considered an extreme condition when dryness has reached deeper soil layers. Considering that the sum of soil dryness-related variable scores occupies 34.4% of the total relevance, the changes in soil dryness play as key drivers in the DL-fire.
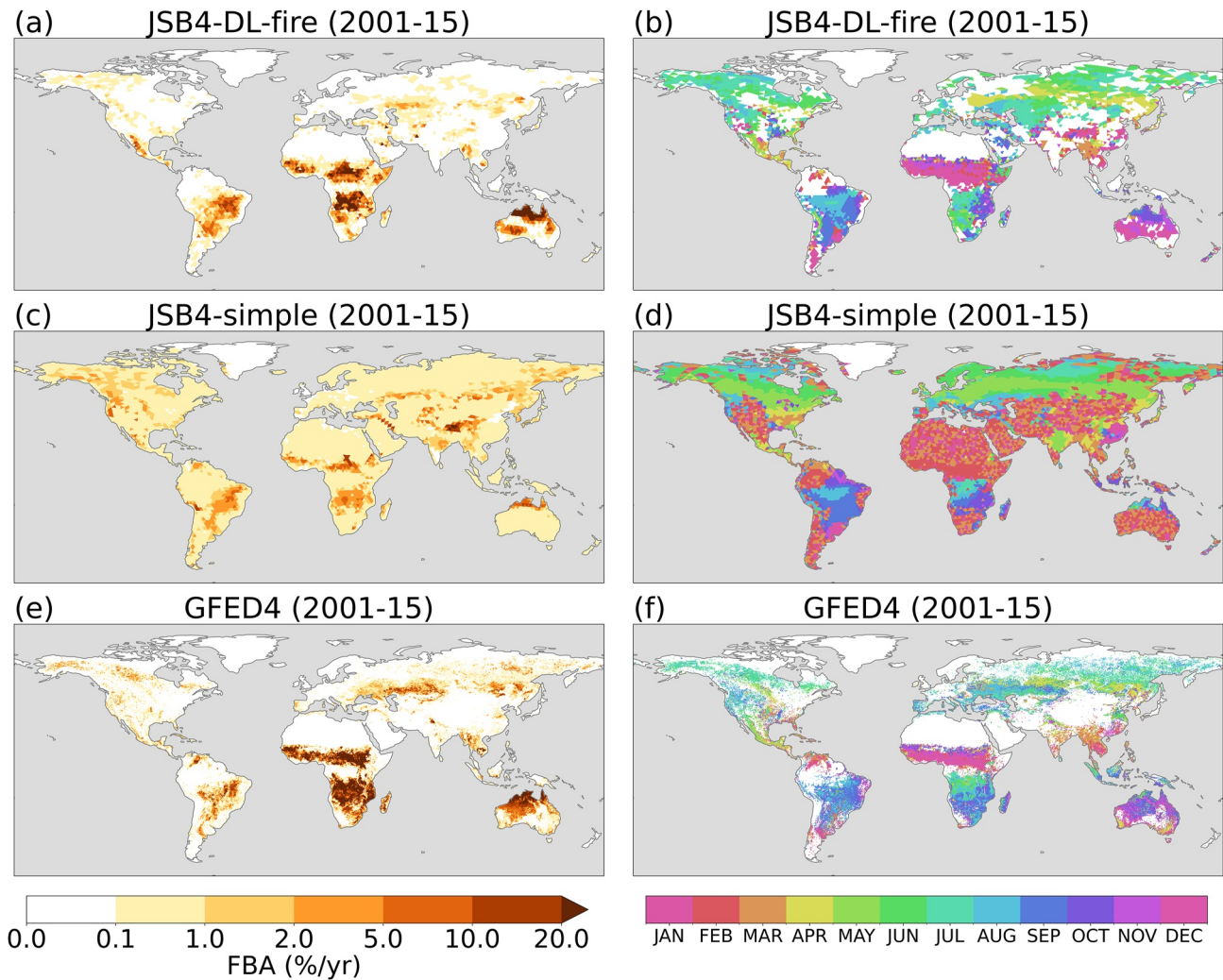
Figure 3. Spatial maps of burnt fraction and its seasonality. The maps on the left (a. JSB4-DL-fire, c. JSB4-simple and e. GFED4) show annual burnt fraction averaged over the years 2001–2015, and the right (b. JSB4-DL-fire, d. JSB4-simple, and f. GFED4) visualize the peak month of burnt fraction. All areas with annual burnt fraction less than 0.1%/yr are masked out (white).

Meteorological predictors, despite their small impacts in the global aspect (6.2%, Figure 5b), display significant importance in some tropical and high latitude regions. For instance, tropical rain forests are very fire-resistant during the wet season due to high humidity. Models trained over NHSA and EQAS show high relevance of relative humidity and temperature to capture the climatic characteristics and their distinct seasonality (Figures S4d and S4m in Supporting Information S1). The strong influence of meteorological predictors is also noticeable over BONA and BOAS, especially temperature contributes the most (12.3% and 16.4% respectively) (Figures S4a and S4j in Supporting Information S1). These results are associated with fire-climate interactions in boreal forests where fire frequency and extent are affected depending on temperature variation (Hu et al., 2015; Kim et al., 2020) and their positive feedbacks under climate change (Oris et al., 2014).

4. Discussion and Conclusion

In this study, we introduce a deep learning based fire model (DL-fire) and implement it within the physics-based land surface model JSBACH4. The DL-fire predicts burnt fraction based on weather conditions, land properties and anthropogenic effects, performing well in predicting spatial and seasonal variation. When the DL-fire operates as a coupled module within JSBACH4 (JSB4-DL-fire), the quality of fire damage simulation improves
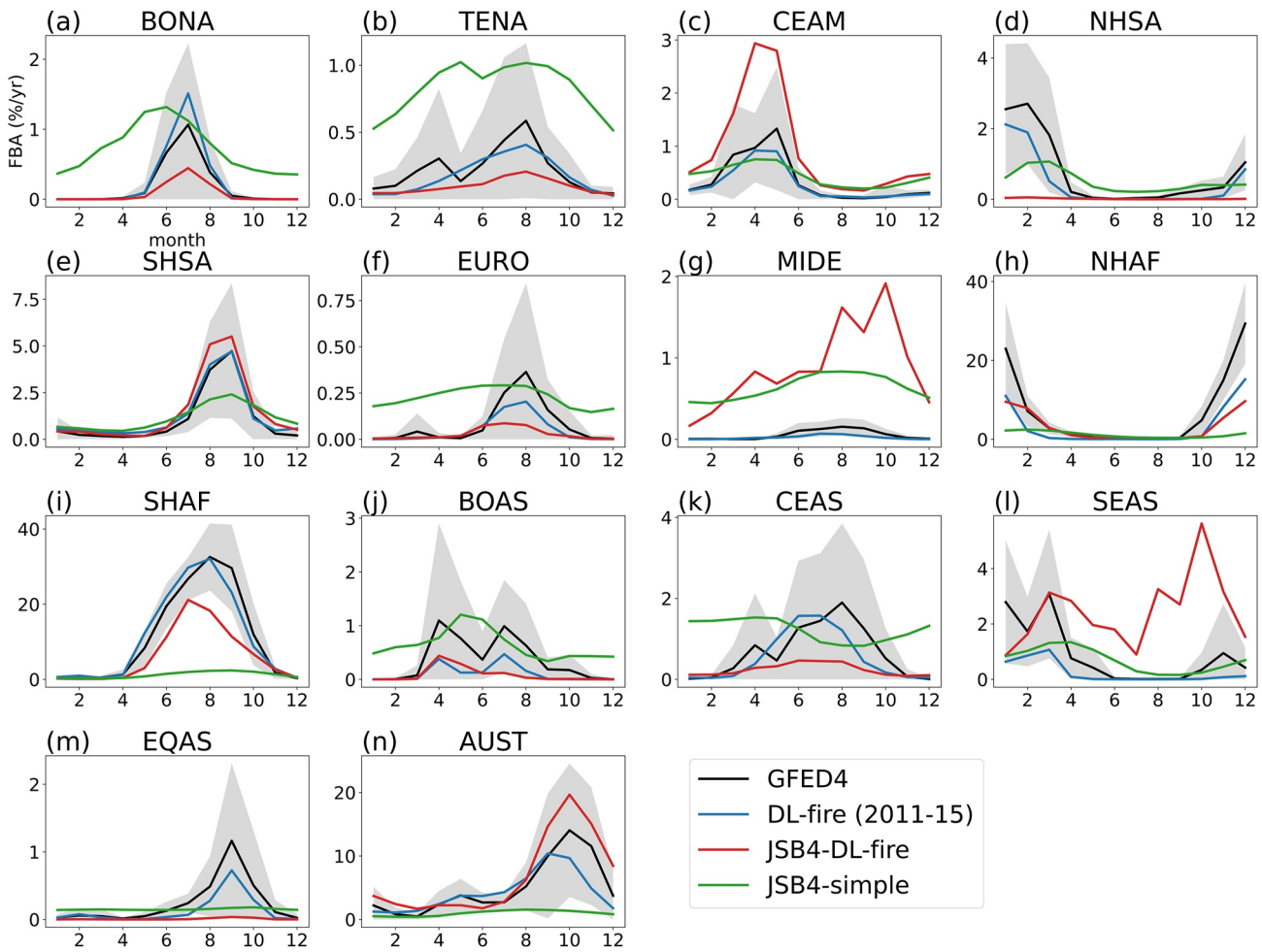
Figure 4. Comparison of monthly mean burnt fraction. Burnt fractions for GFED4 (black), JSB4-DL-fire (red), JSB4-simple (green) during 2001–2015 and DL-fire (blue) during 2011–2015 are averaged for each month and compared on each GFED region (Figure S1 in Supporting Information S1). Gray shadings indicate 1-sigma intervals of the GFED4.

noticeably compared to the simple fire scheme in JSBACH4. However, the predictability of JSB4-DL-fire is not as accurate as the validation results of DL-fire forced by observation. Since the only differences between the two are from land property predictors, either observed or simulated, its main reason is presumed to be internal biases of JSBACH4.
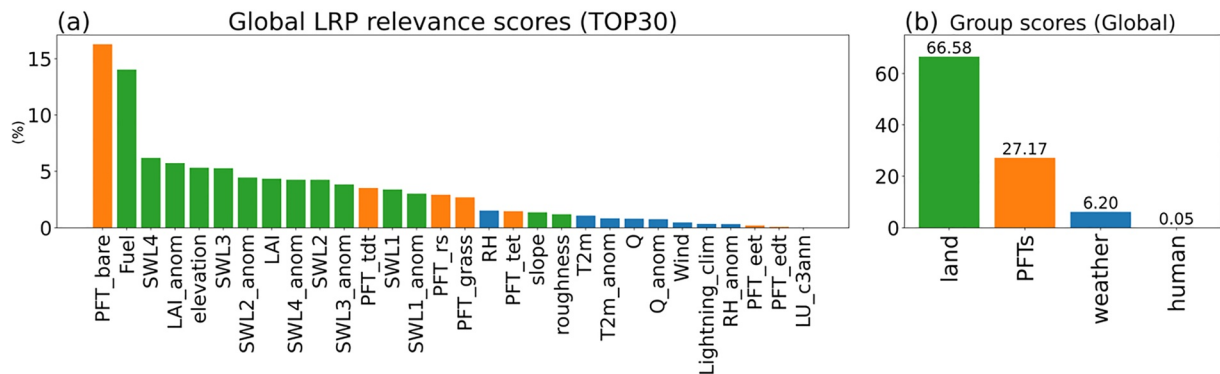

Figure 5. Global predictor importance assessment. (a) Shows predictors with the highest 30 LRP relevance scores and they are color-coded in four groups: weather conditions (blue), land properties (green), anthropogenic effects (gray) and PFTs (orange). Full names of PFTs and land use states (LU) are in Tables 3 and 4. (b) Compares the relevance between the groups and their scores are displayed on top of bars.

To investigate the impact of JSBACH4 internal biases on fire prediction, we compare the predictors from a validation data set and the simulated by JSBACH4. In terms of global perspective, the JSB4-DL-fire predictions overall underestimate burned fraction from May to September, and subsequently its rising and falling seasonal pattern is roughly a month lagged from September to February (Figure S5 in Supporting Information S1). These similar discrepancies are found in LAI over Africa. The simulated LAI in NHAF is overall underestimated with a month lagged peak in its seasonality (Figure S6h in Supporting Information S1). In SHAF, LAI shows opposite seasonal behavior from July to November (Figure S6i in Supporting Information S1), causing an underestimation of fire damage (Figure 4i).

Regionally, MIDE and SEAS show the most apparent discrepancies due to overestimation in JSB4-DL-fire. JSBACH4 shows a tendency to underestimate water contents in all the soil layers (Figures S7–S10 in Supporting Information S1), except for the content of the first layer (SWL1) in MIDE (Figure S7g in Supporting Information S1). Considering that water availability in the topmost layer plays a vital role for vegetation (Seneviratne et al., 2010) and agricultural productivity (Battista et al., 2016), the biases of SWL1 can mislead DL-fire to exaggerate combustible fuel amount or its conditions on the ground. Similarly, overestimated durations of burnt fraction and LAI in SEAS coincide with each other (Figure 4l; Figure S6l in Supporting Information S1). To effectively address internal biases of physics-based models, it was suggested to merge deep learning as an external post-processing method (Reichstein et al., 2019; Son, Kim, et al., 2022; Son, Ma, et al., 2022). However, this approach is not directly applicable in this study due to dynamical interactions between predictors and DGVMs. Instead, an online training approach, developing the deep learning model concurrently with running DGVMs will be our next step to advance the function of DL-fire in ESMs.

Representing interannual variability in global burnt area is yet a continuous effort for improvement in fire-enable DGVMs. Most of the DGVMs have not yet proven to successful in reproducing interannual variability (Hantson et al., 2020), and their limited skills cause uncertainties for the global carbon budget estimation (Bastos et al., 2020). Previous DL model showed ability to capture observed interannual patterns, but it still requires further verification due to its short evaluation period (Joshi & Sukumar, 2021). Although JSB4-DL-fire either performs well at a global scale, significant regional improvements are observed with higher than 0.7 of r_i over 6 out of 14 regions (Table 6). These results suggest that ML/DL based hybrid approach can be a solution for the interannual variability problems in DGVMs.

A critical limitation of our DL-fire model is its lack of mechanical understanding. In contrast to process-based fire models that incorporate scientific principles to account for physical processes and ecological interactions, our current model primarily focuses on estimating burned area statistics without explaining underlying fire processes, such as ignition, spread and extinction. Consequently, this limitation can potentially result in inconsistent or unrealistic simulations in terms of fire frequency and duration. Furthermore, another issue may arise regarding the fire-climate-carbon feedback when the land model is coupled as a part of an ESM, which has not been evaluated in this study.

To address the limitations of ML/DL, it has been suggested to develop independent models specifically tailored for fire ignition and spread (Tang et al., 2023). Also, there has been a proposal to integrate computational fluid dynamics and finite element methods with ML (Ye & Hsu, 2022). These suggestions highlight the potential for further advancements in fire modeling by leveraging the strengths of ML/DL in conjunction with established fire dynamics. By pursuing these avenues of research, our upcoming study will focus on developing DL processes that effectively harmonize fire dynamics.

Humans influence fire regimes in various ways that either promote or limit fire. Population growth and urban expansion generally result in the proliferation of human ignition sources, thereby increasing the likelihood of fires. In contrast, fire suppression efforts and changes in land-use practices contribute to a decline in fire activity (Andela et al., 2017; Bowman et al., 2011). Our model underrates roles of these factors showing conspicuously low global relevance (0.05%, Figure 5b). These consequences can be due to a coarse time resolution of anthropogenic data set. Since all the anthropogenic variables are interpolated from annual records or used as static values, they cannot provide any information associated to seasonal variation or anomalous daily events. Besides, some of the major man-made fire damages, particularly agricultural burnings, can be explained by weather seasonality and vegetation states (Korontzi et al., 2006). However, it should be pointed out that our model globally utilized C3 annual crops (c3ann) the most among anthropogenic drivers (Figure S11a in Supporting Information S1) to identify crop related activities, and regionally in NHAF, BOAS, SEAS, and EQAS (Figures S11i, S11k, S11m, and S11n in

Supporting Information S1). Population follows as the second influential anthropogenic factor and HDI also show relatively higher relevances in developed regions (0.02% in TENA and 0.07% in EURO), echoing their socioeconomic impacts on fire (F. Li et al., 2013; Teixeira et al., 2021). These results may suggest its potential of further improvement of human impacts on fire activities with more sophisticate data set and adapted model architecture.

Small fires play a crucial role in shaping the global-scale patterns of burned areas and carbon emissions (Van Der Werf et al., 2017). However, our current model lacks training data for sub-500m fires, as it is not readily available on a daily scale. To compensate, we additionally evaluate our model's performance by considering burned fractions, incorporating sub-500m fires, but on a monthly scale (Randerson et al., 2012). Overall, the metrics demonstrate similar scores (Table S1 in Supporting Information S1), with a slight increase in NME over regions where the burnt fraction is underestimated, such as NHSA, NHAF, and SEAS (Figures 4d, 4h, and 4i), due to the inclusion of small fire, leading to an augmented observed burnt fraction. EURO shows the most notable decline in performance with r_m of 0.61, although its biases are reduced by 0.66 in NME. We assume that the presence of a very small amount of burnt fractions resulting from Gaussian kernel smoothing prevents significant degradation in the model's performance, despite not explicitly considering small fires.

However, this approach does not comprehensively address the imbalance between small and big fires (Table S2 in Supporting Information S1), which could contribute to the suboptimal performance in extreme burnt fractions. The inclusion of small fires and their distinction from big fires should be considered another essential task to further enhance the model's predictive capabilities.

Regarding a global or local training approach, it can be argued which one in particular is a better option, either one single global model or multiple regional models. A global coverage model can be efficient in terms of model development and coupling with DGVMs, but for it not to lose regional characteristics, it may require more trainable parameters and higher complexity in architecture. When we tested to train a singular global model with the identical architecture as our local models (Figure S12 in Supporting Information S1), its global mean prediction accuracy notably decreased ($r_m = 0.1$). As previously discussed (Zhu et al., 2022), the global training approach only demonstrates relatively satisfactory results across Africa, where high burned areas prevail (NME = 0.73 and 0.72 for NHAF and SHAF, respectively), while substantial biases are observed in other regions with low-fire areas. Considering that we were unable to achieve the desired outcome even with a higher dimension of hidden layers (not shown), there is a need to explore more advanced model architectures that can effectively incorporate local fire characteristics (Lehmann et al., 2014). For the local approach, there are two major points to be considered: (a) the number of regions that should be considered and, (b) whether a unified or a specialized model design for each region should be developed. Exploration of these options would enable us to further upgrade prediction performances, however, this is not addressed in this study.

One of the main purposes of ESMs is to project climate changes based on future scenarios. However, in this study, we decide not to project future fire regime changes with DL-fire, although it is technically executable. This is because our model is currently composed of 14 regional models, and it cannot practically reflect global bioclimatic changes. Finally, we argue that further approaches should focus on developing and training one global DL model coupled with the host land surface model, and by that learning aspects of regional fire variability which would support conducting fully hybrid projection simulations.

Data Availability Statement

We utilized the burned area data set from the GFED4 archive within the Global Fire Emissions Database (Randerson et al., 2015) as the primary target for developing and evaluating our model. Furthermore, meteorological date were retrieved from ERA5 (Hersbach et al., 2020), lightning climatology data were provided through the NASA Earth Science Data and Information System (ESDIS) project and the Global Hydrology Resource Center (GHRC) Distributed Active Archive Center (DAAC) (Cecil et al., 2014). The Earth Science Data Systems (ESDS) facilitated access to LAI from the MCD15A3H version 6 of the Moderate Resolution Imaging Spectroradiometer (Myneni et al., 2015). Global topography data are credited to Amatulli et al. (2018), GDP and HDI data sets are attributed to Kummu et al. (2018). Population density data were obtained from the History database of the Global Environment (Klein Goldewijk et al., 2017), total road density data was sourced from the Global Roads Inventory Project data set (Meijer et al., 2018), and land use states were acquired from the Land-Use Harmonization project (Hurtt et al., 2020). Our model simulation results are openly available in Zenodo at Son et al. (2023).

References

Amatulli, G., Domisch, S., Tuanmu, M.-N., Parmentier, B., Ranipeta, A., Malczyk, J., & Jetz, W. (2018). A suite of global, cross-scale topographic variables for environmental and biodiversity modeling [Dataset]. *Scientific Data*, *5*(1), 1–15. https://doi.org/10.1038/sdata.2018.40

Andela, N., Morton, D. C., Giglio, L., Chen, Y., van der Werf, G. R., Kasibhatla, P. S., et al. (2017). A human-driven decline in global burned area. *Science*, *356*(6345), 1356–1362. https://doi.org/10.1126/science.aal4108

Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K.-R., & Samek, W. (2015). On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLoS One*, *10*(7), e0130140. https://doi.org/10.1371/journal.pone.0130140

Bastos, A., O'Sullivan, M., Ciais, P., Makowski, D., Sitch, S., Friedlingstein, P., et al. (2020). Sources of uncertainty in regional and global terrestrial CO_2 exchange estimates. *Global Biogeochemical Cycles*, *34*(2), e2019GB006393. https://doi.org/10.1029/2019gb006393

Battista, P., Chiesi, M., Rapi, B., Romani, M., Cantini, C., Giovannelli, A., et al. (2016). Integration of ground and multi-resolution satellite data for predicting the water balance of a Mediterranean two-layer agro-ecosystem. *Remote Sensing*, *8*(9), 731. https://doi.org/10.3390/rs8090731

Bowman, D. M. J. S., Balch, J., Artaxo, P., Bond, W. J., Cochrane, M. A., D'antonio, C. M., et al. (2011). The human dimension of fire regimes on Earth. *Journal of Biogeography*, *38*(12), 2223–2236. https://doi.org/10.1111/j.1365-2699.2011.02595.x

Carcaillet, C., Almquist, H., Asnong, H., Bradshaw, R. H. W., Carrion, J. S., Gaillard, M.-J., et al. (2002). Holocene biomass burning and global dynamics of the carbon cycle. *Chemosphere*, *49*(8), 845–863. https://doi.org/10.1016/s0045-6535(02)00385-5

Cecil, D. J., Buechler, D. E., & Blakeslee, R. J. (2014). Gridded lightning climatology from TRMM-LIS and OTD: Dataset description [Dataset]. *Atmospheric Research*, *135*, 404–414. https://doi.org/10.1016/j.atmosres.2012.06.028

Coughlan, R., di Giuseppe, F., Vitolo, C., Barnard, C., Lopez, P., & Drusch, M. (2021). Using machine learning to predict fire-ignition occurrences from lightning forecasts. *Meteorological Applications*, *28*(1), e1973. https://doi.org/10.1002/met.1973

Crutzen, P. J., & Andreae, M. O. (1990). Biomass burning in the tropics: Impact on atmospheric chemistry and biogeochemical cycles. *Science*, *250*(4988), 1669–1678. https://doi.org/10.1126/science.250.4988.1669

D'Este, M., Elia, M., Giannico, V., Spano, G., Lafortezza, R., & Sanesi, G. (2021). Machine learning techniques for fine dead fuel load estimation using multi-source remote sensing data. *Remote Sensing*, *13*(9), 1658. https://doi.org/10.3390/rs13091658

Drüke, M., Forkel, M., von Bloh, W., Sakschewski, B., Cardoso, M., Bustamante, M., et al. (2019). Improving the LPJmL4-SPITFIRE vegetation–fire model for South America using satellite data. *Geoscientific Model Development*, *12*(12), 5029–5054. https://doi.org/10.5194/gmd-12-5029-2019

Dugas, C., Bengio, Y., Bélisle, F., Nadeau, C., & Garcia, R. (2000). Incorporating second-order functional knowledge for better option pricing. *Advances in Neural Information Processing Systems*, *13*.

Forkel, M., Andela, N., Harrison, S. P., Lasslop, G., van Marle, M., Chuvieco, E., et al. (2019). Emergent relationships with respect to burned area in global satellite observations and fire-enabled vegetation models. *Biogeosciences*, *16*(1), 57–76. https://doi.org/10.5194/bg-16-57-2019

Hantson, S., Arneth, A., Harrison, S. P., Kelley, D. I., Prentice, I. C., Rabin, S. S., et al. (2016). The status and challenge of global fire modelling. *Biogeosciences*, *13*(11), 3359–3375. https://doi.org/10.5194/bg-13-3359-2016

Hantson, S., Kelley, D. I., Arneth, A., Harrison, S. P., Archibald, S., Bachelet, D., et al. (2020). Quantitative assessment of fire and vegetation properties in simulations with fire-enabled vegetation models from the Fire Model Intercomparison Project. *Geoscientific Model Development*, *13*(7), 3299–3318. https://doi.org/10.5194/gmd-13-3299-2020

Hardouin, L., Delire, C., Decharme, B., Lawrence, D. M., Nabel, J. E. M. S., Brovkin, V., et al. (2022). Uncertainty in land carbon budget simulated by terrestrial biosphere models: The role of atmospheric forcing. *Environmental Research Letters*, *17*(9), 094033. https://doi.org/10.1088/1748-9326/ac888d

Harrison, S. P., Bartlein, P. J., Brovkin, V., Houweling, S., Kloster, S., & Prentice, I. C. (2018). The biomass burning contribution to climate–carbon-cycle feedback. *Earth System Dynamics*, *9*(2), 663–677. https://doi.org/10.5194/esd-9-663-2018

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis [Dataset]. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hodges, J. L., & Lattimer, B. Y. (2019). Wildland fire spread modeling using convolutional neural networks. *Fire Technology*, *55*(6), 2115–2142. https://doi.org/10.1007/s10694-019-00846-4

Hu, F. S., Higuera, P. E., Duffy, P., Chipman, M. L., Rocha, A. V., Young, A. M., et al. (2015). Arctic tundra fires: Natural variability and responses to climate change. *Frontiers in Ecology and the Environment*, *13*(7), 369–377. https://doi.org/10.1890/150063

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., et al. (2020). Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6 [Dataset]. *Geoscientific Model Development*, *13*(11), 5425–5464. https://doi.org/10.5194/gmd-13-5425-2020

Joshi, J., & Sukumar, R. (2021). Improving prediction and assessment of global fires using multilayer neural networks. *Scientific Reports*, *11*(1), 1–14. https://doi.org/10.1038/s41598-021-81233-4

Jungclaus, J. H., Lorenz, S. J., Schmidt, H., Brovkin, V., Brüggemann, N., Chegini, F., et al. (2022). The ICON Earth system model version 1.0. *Journal of Advances in Modeling Earth Systems*, *14*(4), e2021MS002813. https://doi.org/10.1029/2021ms002813

Kelley, D. I., Prentice, I. C., Harrison, S. P., Wang, H., Simard, M., Fisher, J. B., & Willis, K. O. (2013). A comprehensive benchmarking system for evaluating global vegetation models. *Biogeosciences*, *10*(5), 3313–3340. https://doi.org/10.5194/bg-10-3313-2013

Kim, J.-S., Kug, J.-S., Jeong, S.-J., Park, H., & Schaepman-Strub, G. (2020). Extensive fires in southeastern Siberian permafrost linked to preceding Arctic Oscillation. *Science Advances*, *6*(2), eaax3308. https://doi.org/10.1126/sciadv.aax3308

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. ArXiv Preprint ArXiv:1412.6980.

Klein Goldewijk, K., Beusen, A., Doelman, J., & Stehfest, E. (2017). Anthropogenic land use estimates for the Holocene–HYDE 3.2 [Dataset]. Earth System Science Data, *9*(2), 927–953. https://doi.org/10.5194/essd-9-927-2017

Kloster, S., Mahowald, N. M., Randerson, J. T., Thornton, P. E., Hoffman, F. M., Levis, S., et al. (2010). Fire dynamics during the 20th century simulated by the Community Land Model. *Biogeosciences*, *7*(6), 1877–1902. https://doi.org/10.5194/bg-7-1877-2010

Kondylatos, S., Prapas, I., Ronco, M., Papoutsis, I., Camps-Valls, G., Piles, M., et al. (2022). Wildfire danger prediction and understanding with Deep Learning. *Geophysical Research Letters*, *49*(17), e2022GL099368. https://doi.org/10.1029/2022gl099368

Koppmann, R., Von Czapiewski, K., & Reid, J. S. (2005). A review of biomass burning emissions, part I: Gaseous emissions of carbon monoxide, methane, volatile organic compounds, and nitrogen containing compounds. *Atmospheric Chemistry and Physics Discussions*, *5*(5), 10455–10516.

Korontzi, S., McCarty, J., Loboda, T., Kumar, S., & Justice, C. (2006). Global distribution of agricultural fires in croplands from 3 years of Moderate Resolution Imaging Spectroradiometer (MODIS) data. *Global Biogeochemical Cycles*, *20*(2). https://doi.org/10.1029/2005gb002529

Kummu, M., Taka, M., & Guillaume, J. H. A. (2018). Gridded global datasets for gross domestic product and Human Development Index over 1990–2015 [Dataset]. Scientific Data, 5(1), 1–15. https://doi.org/10.1038/sdata.2018.4

Kurz, W. A., Apps, M. J., Stocks, B. J., & Volney, W. J. A. (1995). Global climate change: Disturbance regimes and biospheric feedbacks of temperate and boreal forests. Biotic Feedbacks in the Global Climatic System: Will the Warming Feed the Warming, 119–133.

Lasslop, G., Thonicke, K., & Kloster, S. (2014). SPITFIRE within the MPI Earth system model: Model development and evaluation. Journal of Advances in Modeling Earth Systems, 6(3), 740–755. https://doi.org/10.1002/2013ms000284

Lehmann, C. E. R., Anderson, T. M., Sankaran, M., Higgins, S. I., Archibald, S., Hoffmann, W. A., et al. (2014). Savanna vegetation-fire-climate relationships differ among continents. Science, 343(6170), 548–552. https://doi.org/10.1126/science.1247355

Lenihan, J. M. (1998). Simulating broad-scale fire severity in a dynamic global vegetation model. Northwest Science, 72(2), 91–103.

le Page, Y., Morton, D., Bond-Lamberty, B., Pereira, J. M. C., & Hurtt, G. (2015). HESFIRE: A global fire model to explore the role of anthropogenic and weather drivers. Biogeosciences, 12(3), 887–903. https://doi.org/10.5194/bg-12-887-2015

Li, F., Lawrence, D. M., & Bond-Lamberty, B. (2017). Impact of fire on global land surface air temperature and energy budget for the 20th century due to changes within ecosystems. Environmental Research Letters, 12(4), 44014. https://doi.org/10.1088/1748-9326/aa6685

Li, F., Levis, S., & Ward, D. S. (2013). Quantifying the role of fire in the Earth system–Part 1: Improved global fire modeling in the Community Earth System Model (CESM1). Biogeosciences, 10(4), 2293–2314. https://doi.org/10.5194/bg-10-2293-2013

Li, Z., Huang, Y., Li, X., & Xu, L. (2021). Wildland fire burned areas prediction using long short-term memory neural network with attention mechanism. Fire Technology, 57(6), 1–23. https://doi.org/10.1007/s10694-020-01028-3

Liang, H., Zhang, M., & Wang, H. (2019). A neural network for wildfire scale prediction using meteorological factors. IEEE Access, 7, 176746–176755. https://doi.org/10.1109/access.2019.2957837

Liu, Z., Ballantyne, A. P., & Cooper, L. A. (2019). Biophysical feedback of global forest fires on surface temperature. Nature Communications, 10(1), 1–9. https://doi.org/10.1038/s41467-018-08237-z

López-Saldaña, G., Bistinas, I., & Pereira, J. M. C. (2015). Global analysis of radiative forcing from fire-induced shortwave albedo change. Biogeosciences, 12(2), 557–565. https://doi.org/10.5194/bg-12-557-2015

McLauchlan, K. K., Higuera, P. E., Miesel, J., Rogers, B. M., Schweitzer, J., Shuman, J. K., et al. (2020). Fire as a fundamental ecological process: Research advances and frontiers. Journal of Ecology, 108(5), 2047–2069. https://doi.org/10.1111/1365-2745.13403

Meijer, J. R., Huijbregts, M. A. J., Schotten, K. C. G. J., & Schipper, A. M. (2018). Global patterns of current and future road infrastructure [Dataset]. Environmental Research Letters, 13(6), 064006. https://doi.org/10.1088/1748-9326/aabd42

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., et al. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications [Dataset]. Earth System Science Data, 13(9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Myneni, R., Knyazikhin, Y., & Park, T. (2015). MCD15A3H MODIS/Terra+Aqua Leaf Area Index/FPAR 4-day L4 Global 500m SIN Grid V006 [Dataset]. NASA EOSDIS Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MCD15A3H.006

Nabel, J. E. M. S., Naudts, K., & Pongratz, J. (2020). Accounting for forest age in the tile-based dynamic global vegetation model JSBACH4 (4.20 p7; git feature/forests)–a land surface model for the ICON-ESM. Geoscientific Model Development, 13(1), 185–200. https://doi.org/10.5194/gmd-13-185-2020

Oris, F., Asselin, H., Ali, A. A., Finsinger, W., & Bergeron, Y. (2014). Effect of increased fire activity on global warming in the boreal forest. Environmental Reviews, 22(3), 206–219. https://doi.org/10.1139/er-2013-0062

Pfeiffer, M., Spessa, A., & Kaplan, J. O. (2013). A model for global biomass burning in preindustrial time: LPJ-LMfire (v1. 0). Geoscientific Model Development, 6(3), 643–685. https://doi.org/10.5194/gmd-6-643-2013

Pongratz, J., Reick, C., Raddatz, T., & Claussen, M. (2008). A reconstruction of global agricultural areas and land cover for the last millennium [Dataset]. Global Biogeochemical Cycles, 22(3). https://doi.org/10.1029/2007gb003153

Radke, D., Hessler, A., & Ellsworth, D. (2019). FireCast: Leveraging deep learning to predict wildfire spread. IJCAI, 4575–4581.

Randerson, J. T., Chen, Y., Van Der Werf, G. R., Rogers, B. M., & Morton, D. C. (2012). Global burned area and biomass burning emissions from small fires. Journal of Geophysical Research, 117(G4), G04012. https://doi.org/10.1029/2012jg002128

Randerson, J. T., van der Werf, G. R., Giglio, L., Collatz, G. J., & Kasibhatla, P. S. (2015). Global fire emissions database, version 4.1 (GFEDv4) [Dataset]. ORNL DAAC. https://doi.org/10.3334/ORNLDAAC/1293

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat (2019). Deep learning and process understanding for data-driven Earth system science. Nature, 566(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1

Reick, C. H., Gayler, V., Goll, D., Hagemann, S., Heidkamp, M., Nabel, J. E. M. S., et al. (2021). JSBACH 3-the land component of the MPI Earth system model: Documentation of version 3.2.

Rothermel, R. C. (1972). A mathematical model for predicting fire spread in wildland fuels (Vol. 115). Intermountain Forest & Range Experiment Station, Forest Service.

Schneck, R., Gayler, V., Nabel, J. E. M. S., Raddatz, T., Reick, C. H., & Schnur, R. (2022). Assessment of JSBACHv4. 30 as a land component of ICON-ESM-V1 in comparison to its predecessor JSBACHv3. 2 of MPI-ESM1. 2. Geoscientific Model Development, 15(22), 8581–8611. https://doi.org/10.5194/gmd-15-8581-2022

Seidl, R., Thom, D., Kautz, M., Martin-Benito, D., Peltoniemi, M., Vacchiano, G., et al. (2017). Forest disturbances under climate change. Nature Climate Change, 7(6), 395–402. https://doi.org/10.1038/nclimate3303

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., et al. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. Earth-Science Reviews, 99(3–4), 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., & Woo, W. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. Advances in Neural Information Processing Systems, 28.

Son, R., Kim, H., Wang, S.-Y. S., Jeong, J.-H., Woo, S.-H., Jeong, J.-Y., et al. (2021). Changes in fire weather climatology under 1.5°C and 2.0°C warming. Environmental Research Letters, 16(3), 34058. https://doi.org/10.1088/1748-9326/abe675

Son, R., Kim, H. C., Yoon, J.-H., & Stratoulias, D. (2022). Estimation of surface PM$_{2.5}$ concentrations from atmospheric gas species retrieved from tropomi using deep learning: Impacts of fire on air pollution over Thailand. Available at SSRN 4255502.

Son, R., Ma, P., Wang, H., Rasch, P. J., Wang, S., Kim, H., et al. (2022). Deep Learning provides substantial improvements to county-level fire weather forecasting over the western United States. Journal of Advances in Modeling Earth Systems, 14(10), e2022MS002995. https://doi.org/10.1029/2022ms002995

Son, R., Stacke, T., Gayler, V., Nabel, J. E. M. S., Schnur, R., Silva, L. A., et al. (2023). Integration of a deep-learning-based fire model into a global land surface model [Software]. Zenodo. https://doi.org/10.5281/zenodo.7728155

Tang, R., Jin, M., Mao, J., Ricciuto, D. M., Chen, A., & Zhang, Y. (2023). TSECfire v1. 0: Quantifying wildfire drivers and predictability in boreal peatlands using a two-step error-correcting machine learning framework. Geoscientific Model Development Discussions, 1–25.

Teixeira, J. C., Folberth, G. A., O'Connor, F. M., Unger, N., & Voulgarakis, A. (2021). Coupling interactive fire with atmospheric composition and climate in the UK Earth System Model. *Geoscientific Model Development*, *14*(10), 6515–6539. https://doi.org/10.5194/gmd-14-6515-2021

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*(6), 1991–2011. https://doi.org/10.5194/bg-7-1991-2010

Thonicke, K., Venevsky, S., Sitch, S., & Cramer, W. (2001). The role of fire disturbance for global vegetation dynamics: Coupling fire into a dynamic global vegetation model. *Global Ecology and Biogeography*, *10*(6), 661–677. https://doi.org/10.1046/j.1466-822x.2001.00175.x

van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., et al. (2017). Global fire emissions estimates during 1997–2016. *Earth System Science Data*, *9*(2), 697–720. https://doi.org/10.5194/essd-9-697-2017

Venevsky, S., Thonicke, K., Sitch, S., & Cramer, W. (2002). Simulating fire regimes in human-dominated ecosystems: Iberian Peninsula case study. *Global Change Biology*, *8*(10), 984–998. https://doi.org/10.1046/j.1365-2486.2002.00528.x

Voulgarakis, A., & Field, R. D. (2015). Fire influences on atmospheric composition, air quality and climate. *Current Pollution Reports*, *1*(2), 70–81. https://doi.org/10.1007/s40726-015-0007-z

Ward, D. S., Kloster, S., Mahowald, N. M., Rogers, B. M., Randerson, J. T., & Hess, P. G. (2012). The changing radiative forcing of fires: Global model estimates for past, present and future. *Atmospheric Chemistry and Physics*, *12*(22), 10857–10886. https://doi.org/10.5194/acp-12-10857-2012

Ye, Z., & Hsu, S.-C. (2022). Predicting real-time deformation of structure in fire using machine learning with CFD and FEM. *Automation in Construction*, *143*, 104574. https://doi.org/10.1016/j.autcon.2022.104574

Yue, C., Ciais, P., Zhu, D., Wang, T., Peng, S. S., & Piao, S. L. (2016). How have past fire disturbances contributed to the current carbon balance of boreal ecosystems? *Biogeosciences*, *13*(3), 675–690. https://doi.org/10.5194/bg-13-675-2016

Zhang, G., Wang, M., & Liu, K. (2021). Deep neural networks for global wildfire susceptibility modelling. *Ecological Indicators*, *127*, 107735. https://doi.org/10.1016/j.ecolind.2021.107735

Zhu, Q., Li, F., Riley, W. J., Xu, L., Zhao, L., Yuan, K., et al. (2022). Building a machine learning surrogate model for wildfire activities within a global Earth system model. *Geoscientific Model Development*, *15*(5), 1899–1911. https://doi.org/10.5194/gmd-15-1899-2022