Building a machine learning surrogate model for wildfire activities within a global earth system model

Qing Zhu[1,*], Fa Li[1,2], William J. Riley[1], Li Xu[3], Lei Zhao[4], Kunxiaojia Yuan[1,2], Huayi Wu[2], Jianya Gong[5], James Randerson[3]

[1] Climate and Ecosystem Sciences Division, Climate Sciences Department, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

[2] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China

[3] Department of Earth System Science, University of California Irvine, Irvine, CA, USA

[4] Department of Civil and Environmental Engineering, University of Illinois Urbana-Champaign, Champaign, IL, USA

[5] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China

*Correspondence to Qing Zhu (qzhu@lbl.gov)

Abstract

Wildfire is an important ecosystem process, influencing land biogeophysical and biogeochemical dynamics and atmospheric composition. Fire-driven loss of vegetation cover, for example, directly modifies the surface energy budget as a consequence of changing albedo, surface roughness, and partitioning of sensible and latent heat fluxes. Carbon dioxide and methane emitted by fires contribute to a positive atmospheric forcing, whereas emissions of carbonaceous aerosols may contribute to surface cooling. Process-based modeling of wildfires in earth system land models is challenging due to limited understanding of human, climate, and ecosystem controls on fire number, fire size, and burned area. Integration of mechanistic wildfire models within Earth system models requires careful parameter calibration, which is computationally expensive and subject to equifinality. To explore alternative approaches, we present a deep neural network (DNN) scheme that surrogates the process-based wildfire model with the Energy Exascale Earth System Model (E3SM) interface. The DNN wildfire model accurately simulates observed burned area with over 90% higher accuracy with a large reduction in parameterization time compared with the current process-based wildfire model. The surrogate wildfire model successfully captured the observed monthly regional burned area during validation period 2011 to 2015 (coefficient of determination, R^2 = 0.93). Since the DNN wildfire model has the same input and output requirements as the E3SM process-based wildfire model, our results demonstrate the applicability of machine learning for high accuracy and efficient large-scale land model development and predictions.

1. Introduction

Wildfires burn ~500 million hectares of vegetated land surface each year, which significantly modifies the physical properties and biogeochemical cycles of terrestrial ecosystems [*Andela et al.*, 2017; *Bond-Lamberty et al.*, 2007; *Pellegrini et al.*, 2018; *Randerson et al.*, 2006]. Living vegetation biomass, surface litter, and coarse woody debris are directly combusted and removed by wildfire [*Harden et al.*, 2006; *Walker et al.*, 2019]. It has been suggested that global forest would double if fire were eliminated [*Bond et al.*, 2005; *Lasslop et al.*, 2020]. Fire has multiple important consequences for the climate system, including directly releasing greenhouse gases (*e.g.*, CO_2, CH_4) [*Kasischke and Bruhwiler*, 2002; *Ross et al.*, 2013] and aerosols [*Jiang et al.*, 2020]; changing land surface albedo and energy budgets [*French et al.*, 2016; *Rother and De Sales*, 2020] and land-atmosphere exchanges of heat, mass, and momentum [*Chambers and Chapin*, 2002]; limiting plant transpiration and regional water recycling [*Brando et al.*, 2020; *Holden et al.*, 2018]; and reshaping forest composition [*Mekonnen et al.*, 2019]. In addition, biomass burning emits a large amount of fine particulate matter that contributes to about 30% of cloud condensation nuclei globally [*Day*, 2004]. Soil organic matter decomposition, nitrogen mineralization, and the richness and diversity of soil fungal communities [*Oliver et al.*, 2015] could also be influenced by wildfire through modifying litter substrate supply and degraded enzymatic activities [*Bowd et al.*, 2019; *Holden et al.*, 2018; *Pellegrini et al.*, 2018; *Pellegrini et al.*, 2020].

Climate change and land use activities have jointly affected fire spatial distribution, frequency, and intensity [*Andela et al.*, 2017; *Kelley et al.*, 2019; *Xu et al.*, 2020] since the pre-industrial era. For example, warmer and drier climate conditions enhance fuel aridity and favor fire occurrence in forest ecosystems where fuels have built up over a period of decades and centuries [*Abatzoglou and Williams*, 2016; *Williams et al.*, 2019]. Even if annual precipitation does not decline, redistribution of precipitation towards wet season extreme rainfall events could contribute to longer dry periods and thus more severe fire activity [*Xu et al.*, 2020]. Human activities often shape wildfire activity through regulating patterns of ignition and fire occurrence (*e.g.*, powerline ignition) [*Keeley and Syphard*, 2018] and suppressing wildfire activity by means of land fragmentation, fire management, and livestock grazing [*Andela et al.*, 2017]. In California, fire density is highly associated with population density and the distance to the wildland urban interface (WUI) [*Syphard et al.*, 2007]. At the global scale, along gradients of increasing population density, fire frequency initially increases by up to 20% and then gradually declines in more densely populated areas [*Knorr et al.*, 2014].

Although global wildfire burned area has declined over the recent two decades [*Andela et al.*, 2017], many vulnerable ecosystems and geographic regions have experienced significant increases in wildfire activity [*Abatzoglou and Williams*, 2016; *Walker et al.*, 2019] resulting in large losses of natural resources and economic assets [*Papakosta et al.*, 2017; *Stephenson et al.*, 2013]. Over western U.S. forests, wildfire has dramatically increased, costing billions of dollars each year and gaining wide public attention. This regional wildfire increase is mainly driven by concurrent increases of spring temperature and declining snowpack [*Westerling et al.*, 2006], mid-summer increases in vapor pressure deficit [*Williams et al.*, 2019], and increases in drought stress during fall [*Goss et al.*, 2020]. The enhancement of wet and dry oscillations favors initial vegetation growth and subsequent wildfire activity [*Heyerdahl et al.*, 2002; *Saha et al.*, 2019].

Wildfire models have played an important role in many aspects of wildfire research, including monitoring fire spread [*Finney*, 1998; *Radke et al.*, 2019], analyzing controllers of wildfire short-term and long-term variability [*Kelley et al.*, 2019], predicting severity of the upcoming fire seasons [*Preisler and Westerling*, 2007] and climate-scale fire variability [*Girardin and Mudelsee*, 2008; *Yue et al.*, 2013], and understanding the complex climate-wildfire-ecosystem feedbacks [*Clark et al.*, 2004; *Mekonnen et al.*, 2019; *Zou et al.*, 2020]. Two types of wildfire models are widely used: process-based models and data-driven statistical models. Process-based wildfire models consider detail processes related to natural fire ignition [*Prentice and Mackerras*, 1977], anthropogenic ignition [*Venevsky et al.*, 2002], fire spread and duration [*Thonicke et al.*, 2010], fire suppression [*Lenihan and Bachelet*, 2015], and fire mass and heat fluxes [*F Li et al.*, 2012]. Process-based wildfire models have been widely used in dynamic vegetation models and coupled earth system models (ESMs) with various complexities of parameterization [*Fang Li et al.*, 2019; *Rabin et al.*, 2017]. As more and more detailed fire processes are considered and parameterized, structural and parametric uncertainties may increase due to incomplete representation of individual processes and imperfect mathematical formulation [*Riley and Thompson*, 2017]. Historically, data-driven models were often used for fire behavior modeling and aim to track the ignition, spread, duration, and extinction of individual fires [*Finney*, 1998; *Radke et al.*, 2019] at fine spatial and temporal scales. This group of models are more relevant to operational fire research. While process-based wildfire models in the context of global vegetation models or earth system land models focuses on the gridcell aggregated fire burned area dynamics that are more relevant to researches on large scale patterns and climate scale predictions [*Fang Li et al.*, 2019; *Rabin et al.*, 2017]. This study particularly focuses on the second category of wildfire models.

Although explicit processes are simulated, the accuracy of process-based wildfire models are highly dependent on parameterization, which is computationally expensive [*Teckentrup et al.*, 2018; *Zhu and Zhuang*, 2014]. Data-driven models, however, directly link the driving variables (*e.g.*, climate factors) to the fire activity using simple statistical models or more sophisticated machine learning techniques, ignoring the explicit processes and feedbacks associated with wildfire [*Ganapathi Subramanian and Crowley*, 2018; *Radke et al.*, 2019; *Tonini et al.*, 2020]. Through training and validation, statistical representations of wildfire dynamics are learned by models using principles from machine learning. Data-driven wildfire models are diverse in terms of driving variables and model structure. For example, many current machine learning wildfire models rely on remote oceanic dynamics (*e.g.*, sea surface temperature variability) and atmospheric teleconnections to simulate land surface fire activities [*Chen et al.*, 2020; *Chen et al.*, 2011; *Yu et al.*, 2020]. Another group of data-driven wildfire models draws more heavily upon regional climate, plant functional type, and human infrastructure driver variables [*Coffield et al.*, 2019; *Sayad et al.*, 2019].

In this study, we develop a machine learning wildfire model using the process representation of wildfire in the Energy Exascale Earth System Model (E3SM) land model (ELMv1) [*Zhu et al.*, 2019], five observationally inferred burned area products [*Andela et al.*, 2019; *Giglio et al.*, 2018; *Joshua Lizundia-Loiola et al.*, 2020; *J Lizundia-Loiola et al.*, 2018; *Van Der Werf et al.*, 2017], and a deep neural network approach [*Goodfellow et al.*, 2016]. We implemented a deep learning model that can better capture the complex and non-linear interactions between controlling factors and wildfire activity. The objectives of this study are to surrogate the wildfire parameterization in ELMv1 with the deep neural network and improve the model simulated wildfire burned area across various fire regions [*Giglio et al.*, 2013].

2. Methodology

2.1 ELMv1 wildfire model

The process-based wildfire model in ELMv1 originates from the Community Land Model (CLM4.5) [*F Li et al.*, 2012]; we take this wildfire model as the baseline (hereafter refer to as BASE-Fire). BASE-Fire combines information regarding ignition, fuel conditions, surface climate, and anthropogenic suppression to simulate total burned area based on the fire counts and spread area of each fire (Figure 1). The fire count in BASE-Fire is modeled as the sum of anthropogenic ignition and natural ignition, where the latter is proportional to lightning density [*Prentice and Mackerras*, 1977] and the former is determined by population density [*Venevsky et al.*, 2002]. Human activity may also intentionally suppress wildfire occurrence if the fire is detected at early stage. For example, developed regions with high population density and gross domestic product are less likely to use fire to remove surface biomass. On the other hand, developed regions more likely suppress fire given more effective fire management policy and suppression capability. Fire count is also affected by surface fuel availability (aboveground biomass) and fuel combustibility (relative humidity, topsoil temperature and moisture). The fire spread area in BASE-Fire is modeled as an elliptical shaped region controlled by wind speed and fuel wetness (using topsoil (0 – 15 cm) moisture as a proxy). The fire duration is set to be one day based on a study that reported years 2001-2004 mean global fire persistence [*Giglio et al.*, 2006a]. BASE-Fire also does not explicitly consider roads, rivers, and firefighting activity [*Arora and Boer*, 2005].

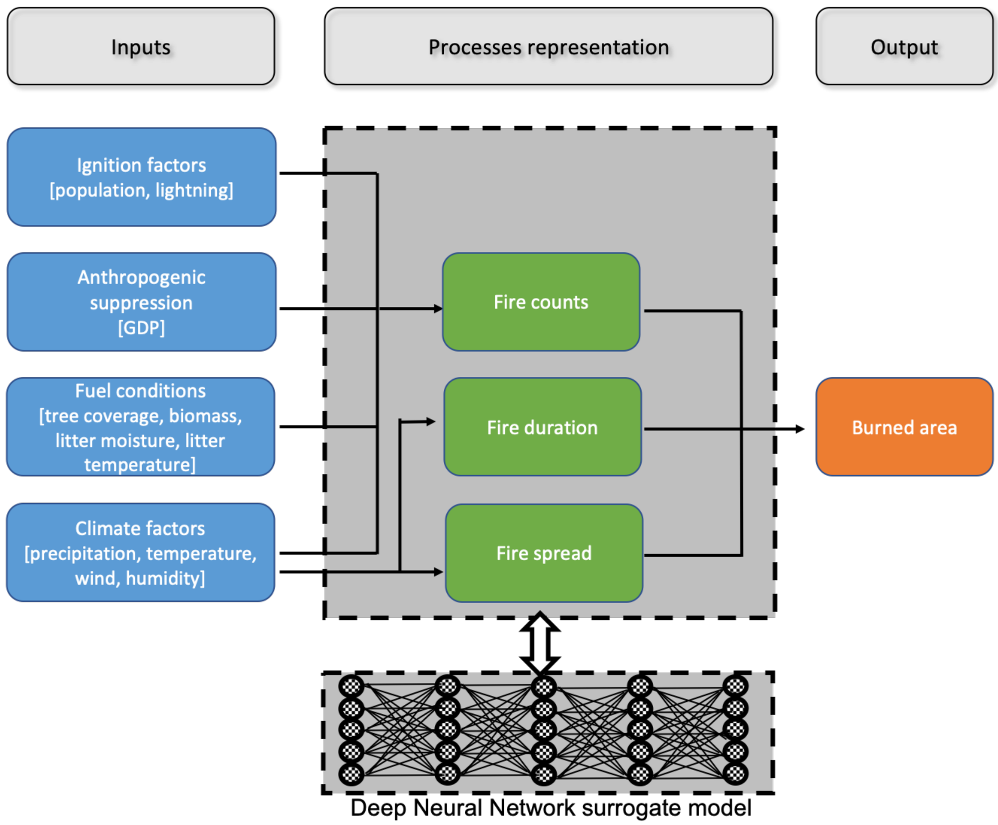

Figure 1. Schematic representation of the ELMv1 process-based BASE-Fire model and the components to be surrogated with the Deep Neural Network (DNN) model (dark grey).

2.2 Deep neural network wildfire surrogate model

We developed the new fire model in two steps: (1) surrogating BASE-Fire with a deep neural network (DNN) approach and (2) improving that surrogate model using five observationally inferred burned area products (Table S1). First, we surrogated BASE-Fire with a DNN approach (hereafter refer to as DNN-Fire) that uses the same input and output variables as BASE-Fire but treats the explicit intermediate processes (*e.g.*, ignition, fire spread) as latent variables coded by hidden layers in the DNN (Figure 1). DNN-Fire was developed with five hidden layers and five neurons in each layer for burned area simulation. The DNN approach uses a fully-connected feedforward neural network [*Schmidhuber*, 2015] that comprises input, hidden, and output layers:

$$h_1 = f_1(W_1 I + b_1) \quad (1)$$

$$h_2 = f_2(W_2 h_1 + b_2) \quad (2)$$

$$h_3 = f_3(W_3 h_2 + b_3) \quad (3)$$
$$h_4 = f_4(W_4 h_3 + b_4) \quad (4)$$
$$h_5 = f_5(W_5 h_4 + b_5) \quad (5)$$
$$O = f_6(W_6 h_5 + b_6) \quad (6)$$

where *I* denotes the input layer (*e.g.*, climate factors) with 11 neurons, each corresponding to an input variable listed in Table 1. h_1, h_2, h_3, h_4, and h_5 are five hidden vectors that are calculated with two steps. First is a linear combination of previous layers' input vector (h) and the trainable weight parameter matrix [$W_1, W_2, W_3, W_4, W_5, W_6$], considering biases b_1, b_2, b_3, b_4, b_5, and b_6. Then, nonlinear activation functions f_1, f_2, f_3, f_4, f_5, and f_6. are applied to the output from the previous step. In this study we used *softplus* as the activation function [*Zheng et al.*, 2015] that is a non-linear transformation of input signals. *O* denotes the output layer that summarize the latent variables from the last hidden layer (h_5) and calculate burned area.

Table 1. Input and output variables of ELMv1 BASE-Fire and surrogate DNN-Fire models

Variable category	Variable name	Data source and reference
	Input variables	
Fuel conditions	Tree coverage	LUH2 [*Hurtt et al.*, 2020]
	Fuel load	ELMv1 total biomass [*Zhu and Riley*, 2015; *Zhu et al.*, 2019]
	Fuel wetness	ELMv1 topsoil moisture [*Zhu and Riley*, 2015; *Zhu et al.*, 2019]
	Fuel temperature	ELMv1 topsoil temperature [*Zhu and Riley*, 2015; *Zhu et al.*, 2019]
Climate factors	Precipitation	GSWP3 [*Dirmeyer et al.*, 2006]
	Near surface temperature	GSWP3 [*Dirmeyer et al.*, 2006]
	Wind speed	GSWP3 [*Dirmeyer et al.*, 2006]
	Relative humidity	GSWP3 [*Dirmeyer et al.*, 2006]
Ignition	Population density	[*Dobson et al.*, 2000]
	Lightning frequency	NASA-LIS/OTD [*Cecil et al.*, 2014]
Anthropogenic suppression	GDP	[*van Vuuren et al.*, 2007]
	Population density	[*Dobson et al.*, 2000]

output variable	
Burned area	ELMv1 percentage burned area [*Zhu and Riley*, 2015; *Zhu et al.*, 2019]

Second, we improved the surrogate DNN-Fire by fine-tuning the weight parameters using observations (hereafter refer to DNN-Fire-OBS). Between 2001 and 2010, we initialized DNN-Fire-OBS's weight parameters (W_1, W_2, W_3, W_4, W_5, and W_6) using results from DNN-Fire, replaced the BASE-Fire burned area by the ensemble mean of five observationally inferred burned area products including GFEDv4s [*Van Der Werf et al.*, 2017], Fire_CCI51 [*Joshua Lizundia-Loiola et al.*, 2020], Fire_CCILT11 [*J Lizundia-Loiola et al.*, 2018], MODIS MCD64 [*Giglio et al.*, 2018], and Fire_Atlas [*Andela et al.*, 2019] (Table S1), and adjusted weight parameters until the model best reproduced the observed burned area. This two-step approach will also allow rapid parameterization of the Fire model as new fire data and baseline fire model results become available. DNN-Fire-OBS can be more easily generalized since BASE-Fire provides explicit physical guidance and a larger-than-observation input and output feature space for development of the machine learning fire model.

2.3 Model setup and simulation protocol

We ran ELMv1 with BASE-Fire at 1.9° by 2.5° spatial resolution [*Zhu et al.*, 2020; *Zhu et al.*, 2016] to generate training and testing datasets for the DNN wildfire model. BASE-Fire was first spun up for 600 years with accelerated soil decomposition followed by 200 years regular spinup with regular soil decomposition [*Koven et al.*, 2013]. The spinup simulations were forced with constant atmospheric CO_2 concentration (285 ppmv) and 1901-1920 repeated climate forcing from GSWP3 (Global Soil Wetness Project) [*Dirmeyer et al.*, 2006]. The purpose of the spinup was to initialize ecosystem carbon pools and stabilize plant and soil carbon and water fluxes. Restarting from the "spunup" conditions, a transient simulation was then conducted from 1901 to 2015 with GSWP3 transient climate forcing, atmospheric CO_2 concentrations, and nitrogen and phosphorus deposition [*Lamarque et al.*, 2005; *Mahowald et al.*, 2008]. Wildfire associated variables were selected for output with a monthly temporal resolution (Table 1).

BASE-Fire output from years 1981 to 2010 were used to train, test, and fine-tune DNN-Fire. We developed 14 region-specific models, corresponding to 14 widely used GFED regions. For each region, all land gridcells (comprising no fire history, infrequent fire, and repeated fire) were concatenated into one data matrix (where rows consist of the number of samples and columns of the number of variables). 80% of the data matrix was randomly sampled for the training dataset and the remaining 20% of the data were reserved for testing. Furthermore, the random sampling was stratified in order to reduce the risk of sampling, e.g., adjacent high fire grid cells. All grid cells were first divided into three "strata": low burn (0-33% percentile), median burn (33%-66% percentile), and high burn (67-100% percentile) grid cells based on the magnitude of the burn. The stratified random sample assured the sampled grid cells for training and testing had the same ratios of low/medium/high burn, thus eliminating the sampling bias from spatial autocorrelation [*Wang et al.*, 2012]. In addition to random sampling, we also investigated the impacts of data choice on the model performance, by sampling the testing datasets within specific years (*e.g.*, 2001-2002, 2003-2004, 2005-2006, 2007-2008, 2009-2010) and used the rest of the years for training. We found neglected differences among the models (Figure S1) indicating the choice of training/testing data years were not impactful. Therefore, we will discuss the results with stratified random sampling approach as the major results throughout the paper.

All training and testing datasets were normalized to the range [0, 1] with the following scaler:

$$X = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (7)$$

where X is the variable vector of interest and X_{min} and X_{max} are minimum and maximum values of X, respectively. During the training stage, we randomly initialized the weighting parameters (Eq. 1-6) and optimized them using the Adaptive Moment Estimation method [*Kingma and Ba*, 2014], which is a variant of the gradient descent optimization method but considers adaptive learning rate and momentum-like exponentially decaying gradients. The parameter optimization aimed to minimize a mean squared error cost function:

$$J = \frac{1}{n}\sum_{i=1}^{n}(y_i^{DNN} - y_i^{BASE})^2 \qquad (8)$$

where y_i^{DNN} and y_i^{BASE} are DNN-Fire and BASE-Fire generated burned area, respectively. Cost function J summarizes the overall magnitude of the error between the surrogate DNN-Fire and BASE-Fire. We then evaluated model performance using metrics of mean absolute error (Eqn. 9), Pearson correlation (Eqn. 10), and coefficient of determination (Eqn. 11).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i^{DNN} - y_i^{BASE}| \qquad (9)$$

$$p = \frac{covariance(y^{DNN}, y^{BASE})}{variance(y^{DNN}) variance(y^{BASE})} \quad (10)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i^{DNN} - y_i^{BASE})^2}{\sum_{i=1}^{n} (y_i^{BASE} - y_{mean}^{BASE})^2} \quad (11)$$

3. Results and discussion

3.1 Evaluation of wildfire surrogate model

BASE-Fire performed reasonably well for total global burned area (508 ± 53 Mha yr^{-1} (million hector per year) between years 2001 and 2010 compared with the observational long-term average of 424~484 Mha yr^{-1}; Figure 2, Table S1). BASE-Fire also captured the global declining trend of wildfire burned area over this time period, attributed to a decrease in tropical fires [*Andela et al.*, 2017]. At the regional scale, however, BASE-Fire underestimated tropical (S23.5° - N23.5°) burned area and overestimated temperate (N23.5° - N67.5°) and boreal (N67.5 above) burned area (Figure 2). Large spatial heterogeneity existed for BASE-Fire regional bias. For example, over tropical GFED regions, BASE-Fire overestimated wildfire burned area over Southern Hemisphere South America (SHSA), but underestimated wildfire burned area over both Southern and Northern Hemisphere Africa regions (SHAF and NHAF), despite an overall underestimation over the tropical region (Figure 3). In contrast, consistent overestimation occurred over all temperate GFED regions. For example, wildfire burned was overestimated by about a factor of 16 (~1 versus 16 Mha yr^{-1}) over the Europe GFED region (EURO) (Figure 3). Although there is room to improve BASE-Fire performance, the parameterization would involve large ensemble simulations and computational resources. Instead, we first used BASE-Fire generated data as training and testing datasets to parameterize DNN-Fire, then we fine-tuned the DNN-Fire model against observed burned area.

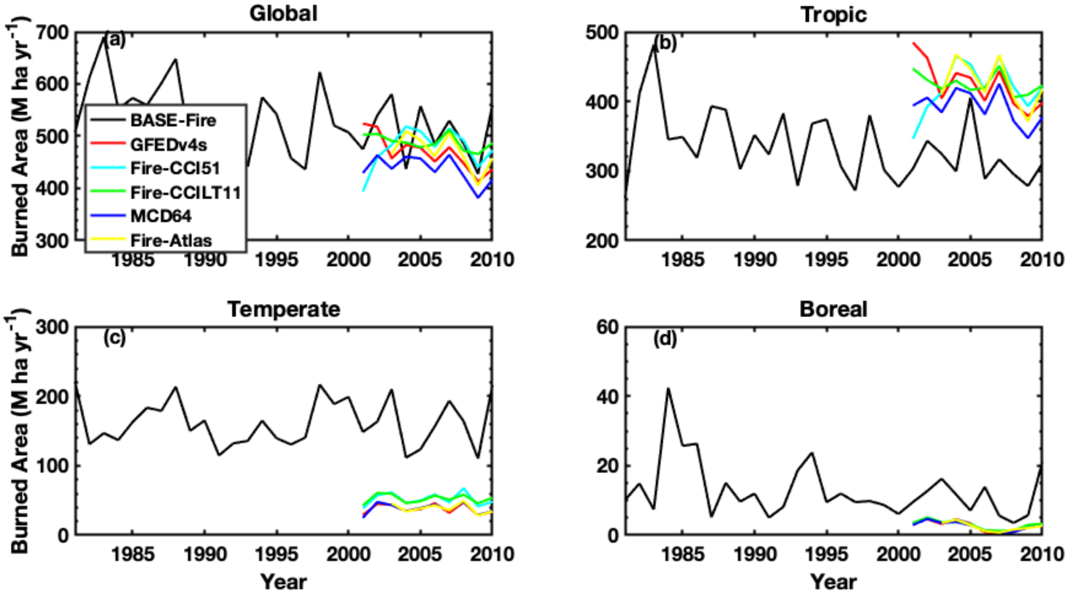

Figure 2. ELMv1 process-based model (BASE-Fire) simulated and five observationally inferred burned area products (Table S1) at (a) global scale; (b) Tropical (S23.5° -N23.5°); (c) Temperate (N23.5° - N 67.5°); and (d) Boreal (north of N 67.5°) regions.

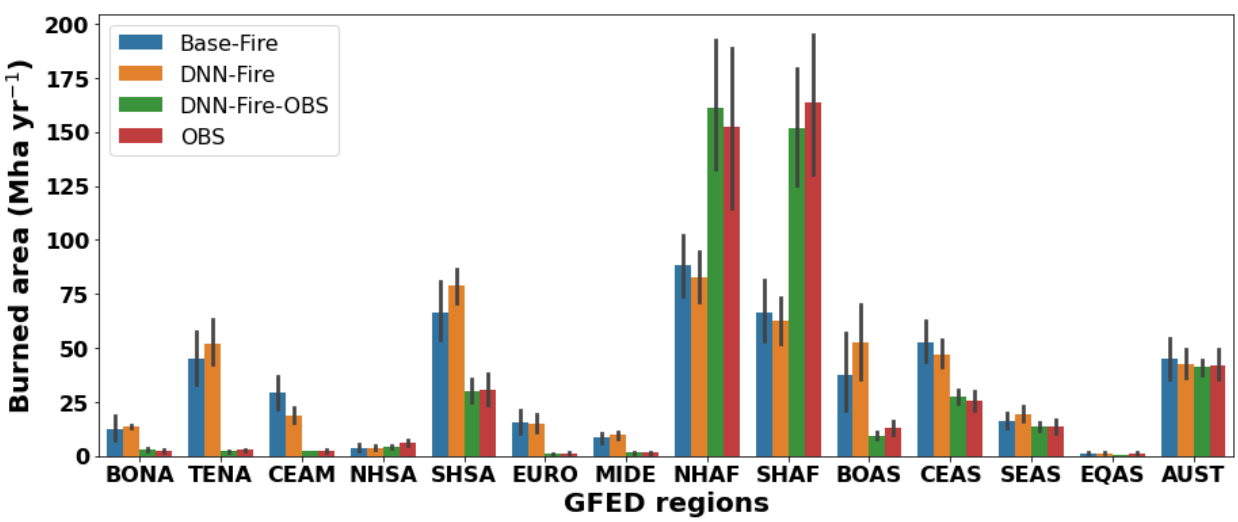

Figure 3. A comparison of wildfire burned area between estimates from the ELMv1 process-based model (BASE-Fire), Deep Neural Network wildfire model (DNN-Fire), Deep Neural Network wildfire model fine-tuned with observed burned area (DNN-Fire-OBS), and observations over 14 GFED fire regions.

Next we parameterized and compared DNN-Fire with BASE-Fire outputs. Using BASE-Fire generated 1.9° × 2.5° resolution datasets of surface fuel conditions (fuel load (vegetation biomass), fuel temperature (topsoil temperature), and fuel wetness (topsoil moisture)) with gridded climate forcing (GSWP3) [*Dirmeyer et al.*, 2006], land use (LUH2 dataset) [*Hurtt et al.*, 2020], and social economic [*Dobson et al.*, 2000; *van Vuuren et al.*, 2007] factors, DNN-Fire captured the spatial pattern of BASE-fire predicted wildfire activity (Figure 4, Figure S2). Across all GFED regions, mean absolute error of DNN-Fire was 4.4 Mha yr^{-1} (<1% of total burn area), with median and maximum errors of 1.8 and 13.0 Mha yr^{-1}, respectively (Figure 3). Equatorial Asia (EQAS), Northern Hemisphere South America (NHSA), Central America (CEAS), and Europe (EURO) regions had the lowest DNN-Fire errors (< 1.0 Mha yr^{-1}), while Southern Hemisphere Africa (SHAF), and Boreal Asia (BOAS) had the largest errors (10-13 Mha yr^{-1}). Overall, the correlation coefficient between BASE-Fire and DNN-Fire simulated burned area was 0.91 (p value < 0.01) and the coefficient of determination (R^2) was 0.79. Across seasons, DNN-Fire also reasonably captured the BASE-Fire peak fire months (June to October), which were dominated by Southern Hemisphere Africa and Southern Hemisphere South America (Figure 5).

By surrogating BASE-Fire, DNN-Fire is expected to have similar biases and uncertainties. The deficiency of BASE-Fire model will propagate to DNN-Fire. In our future work we will overcome such limitation by training multiple DNN-Fire models with ensemble simulations of BASE-Fire models that differ in critical parameters and vary in model structures.

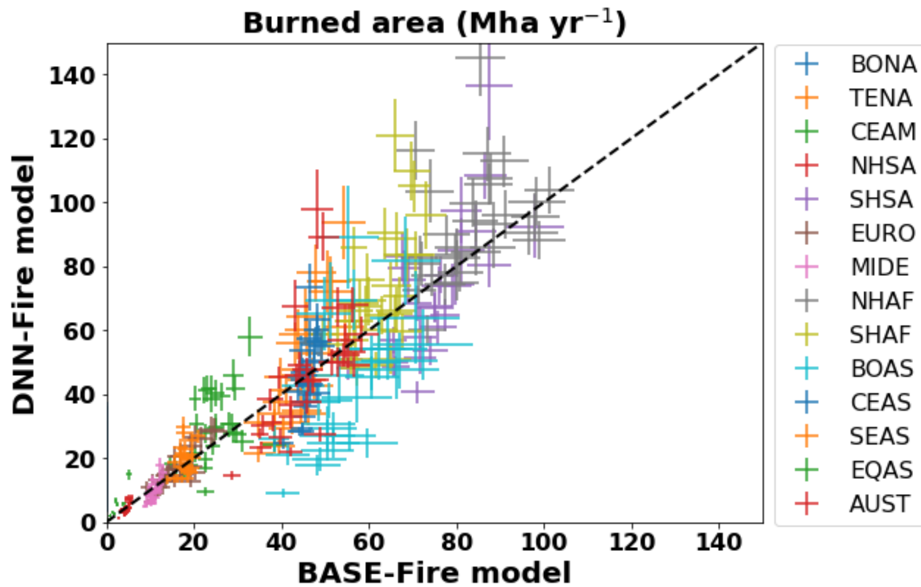

Figure 4. The performance of the Deep Neural Network wildfire model (DNN-Fire), compared with the original ELMv1 process-based wildfire model (BASE-Fire) over 14 GFED regions between years 2001 and 2010.

3.2 Calibrating the wildfire surrogate model using observations

Although the global pattern was reasonably captured, BASE-Fire had relatively large biases in several GFED regions, as discussed above. Since DNN-Fire was trained and validated only with BASE-Fire generated inputs (*e.g.*, fuel conditions) and outputs (burned area), we expect that, at best, DNN-Fire would have comparable biases as BASE-Fire. Starting from DNN-Fire, we further calibrated the model weighting parameters to create DNN-Fire-OBS and validated DNN-Fire-OBS performance using observed burned area from five existing burned area products (Table S1) between years 2001 and 2010. The calibration time cost several minutes with Intel Xeon Phi Processor 7250 processor.

Dramatic improvements were found in most of the 14 regions simulated by DNN-Fire-OBS (Figure 3). Overall, DNN-Fire-OBS simulated global long-term average burned area was 458 Mha yr^{-1} (compared with observational average 467 Mha yr^{-1}). Averaged across 14 regions, 73% reduction of mean absolute error was achieved by DNN-Fire-OBS, compared with the BASE-Fire model. Pearson correlation coefficient between the DNN-Fire-OBS simulated and observational burned area was 0.98 (p value < 0.001) with an R^2 of 0.97. Bias reduction was disproportionally distributed across the GFED regions (Figure 3). For example, severely burned regions, including Southern and Northern Hemisphere Africa (SHAF and NHAF) and Southern Hemisphere South America (SHSA) greatly benefited from the tuning and their regional biases were reduced by 88, 65, and 51 Mha yr^{-1} (or 88%, 89%, 98% reduction), respectively. Although Temperate Northern America (TENA) and Europe (EURO) wildfire burned area is relatively small (1-3 Mha yr^{-1}), the impacts of wildfire activity were significant due to their high population densities. DNN-Fire tended to overestimate the burned area in TENA and EURO by 47 and 13 Mha yr^{-1}, while DNN-Fire-OBS significantly reduced biases in both regions to less than 0.3 Mha yr^{-1} (a 97-98%% reduction).

BASE-Fire tended to overestimate inter-annual variability (IAV) and had opposite burned area anomalies between years 2001 and 2005. DNN-Fire dampened BASE-Fire's IAV, but had systematic overestimation of burned area. DNN-Fire-OBS agreed well with the observed IAV between years 2001 and 2010 (Figure 5a). The seasonal cycle was also improved in DNN-Fire-OBS in terms of reducing BASE-Fire's overestimation of burned area during peak fire seasons (Figure 5b, Figure S3), although we note that DNN-Fire-OBS is biased high during low fire seasons (March and April).

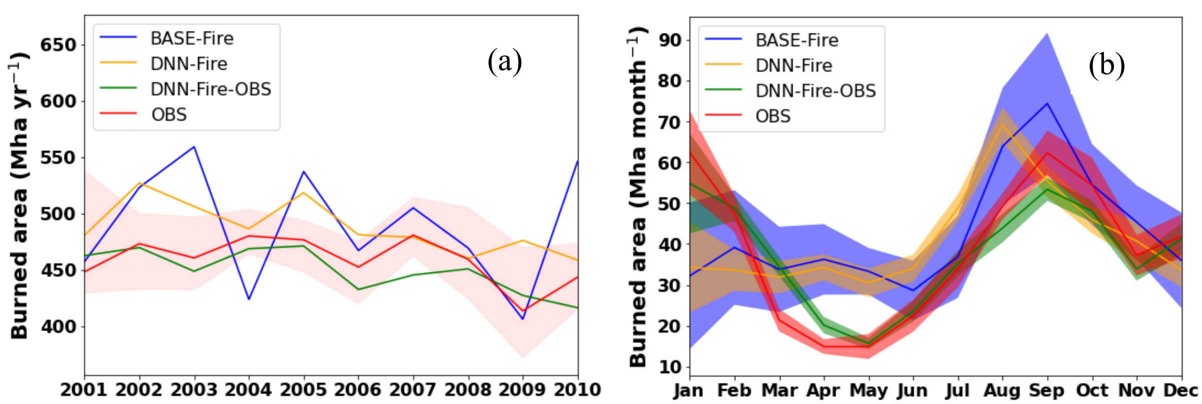

Figure 5. Inter-annual variation of burned area from years 2001 to 2010 (a) and the averaged seasonal cycle (b) of burned area estimated by the ELMv1 process-based wildfire model (BASE-Fire), Deep Neural Network wildfire model (DNN-Fire), Deep Neural Network wildfire model fine-tuned with observations (DNN-Fire-OBS), and observations.

3.3 Prognostic simulation and limitations

We next evaluated the DNN-Fire-OBS model against observations for the period 2011 to 2015, using data which were not used to train and validate the model. Overall, DNN-Fire-OBS simulated 469-514 Mha yr^{-1} global burned area, compared with observations 349-509 Mha yr^{-1}. Note that the large observational ranges were mainly due to the differences among the five burned area products rather than the inter-annual variability (Figure 6). Regionally, DNN-Fire-OBS overestimated NHAF, SHAF and SHSA annual burned area by 8, 6, 2 Mha yr^{-1}, respectively (Figure 6) compared with the observational mean. Averaged latitudinal distribution of simulated burned area during this period showed that global wildfire activity peaked around S10°- S15° and N5°-N10°, together accounting for burning 12-16% of the land surface (Figure 7). These two peaks were dominated by large burned area over Southern (SHAF) and Northern Hemisphere Africa (NHAF) fire regions. Compared with observations, DNN-Fire-OBS simulated reasonable burned area latitudinal distributions (Figure 7). We also compared the nine FireMIP models [*Rabin et al.*, 2017; *Teckentrup et al.*, 2018] and found diverse latitudinal distribution of burned area. The across model differences were much larger than the inter-annual variation simulated by each individual model, which indicated large model structural uncertainties. Validation was also conducted for the historical period 1981-2000, when most of the satellite based burned area data were not available. Compared with charcoal index inferred burned area during 1981-2000 (Figure S4), DNN-Fire-OBS model reasonably captured the declining of burned area from ~530 Mha yr^{-1} to 490 Mha yr^{-1}. In summary, DNN-Fire-OBS simulation is reasonably accurate and: (1) improved the simulated wildfire spatial and temporal distributions in ELMv1; (2) enabled effective and efficient parameterization of fires at regional scale.

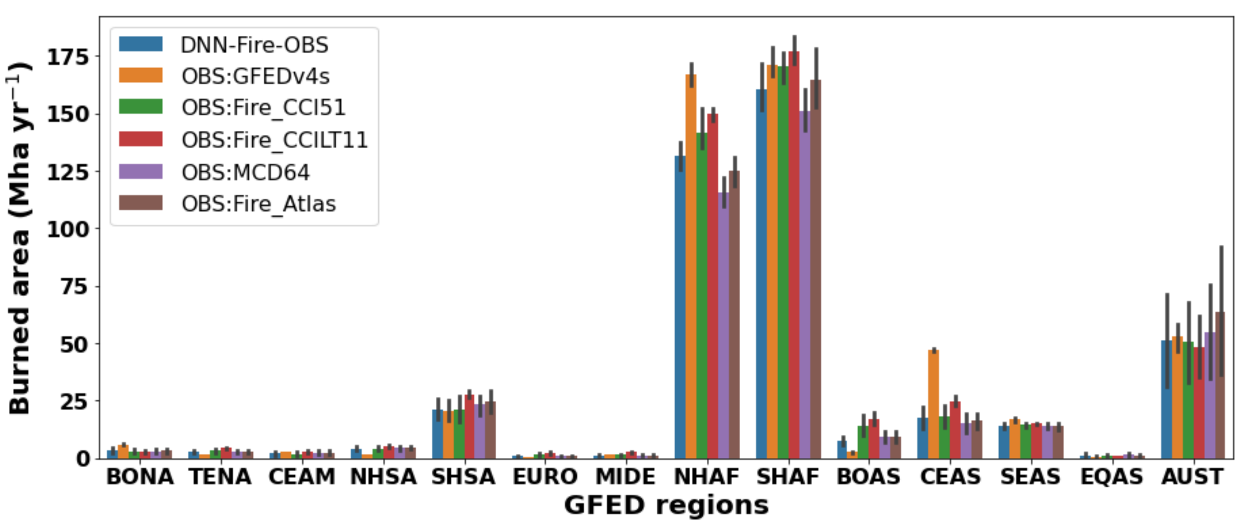

Figure 6. Prognostic simulation of annual wildfire burned area with the Deep Neural Network wildfire model fine-tuned with observations (DNN-Fire-OBS) compared with five burned area products (Table S1) over 2011-2015 for 14 GFED regions.

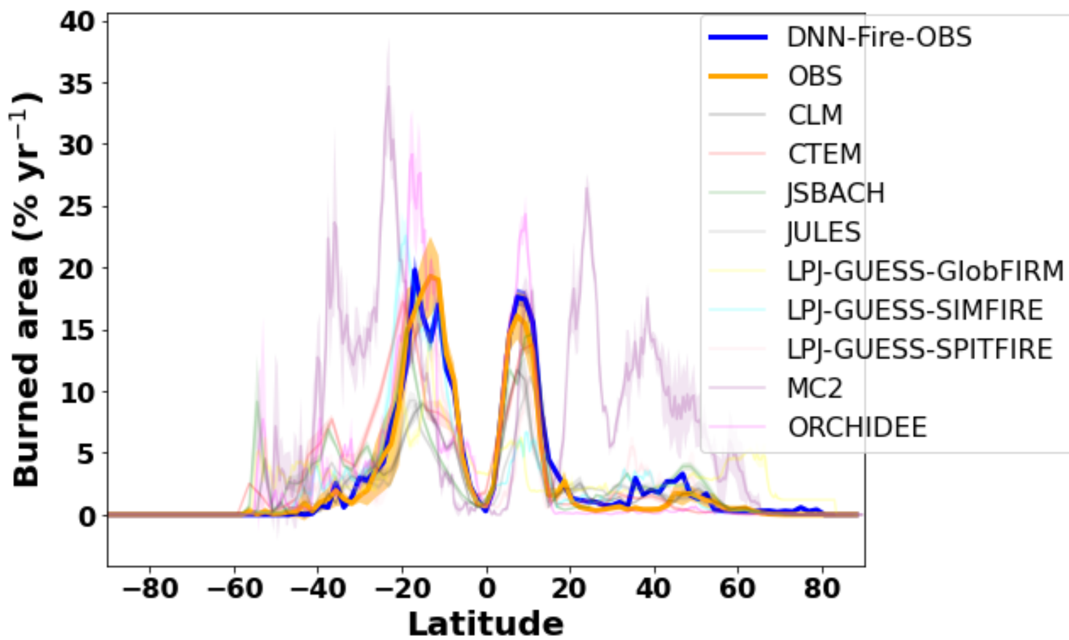

Figure 7. Prognostic simulation of wildfire burned area (2011-2015) with the Deep Neural Network wildfire model fine-tuned with observations (DNN-Fire-OBS) compared with observations and nine FireMIP models outputs.

This study focuses on design, development, and parameterization of the DNN fire model within the E3SM model interface. In this way the DNN model can be readily coupled in the future and iteratively simulate climate, ecosystem fuel conditions, and fire dynamics. This study is an important step towards fully coupling E3SM and the DNN-Fire models in the future. We acknowledge several challenges and limitations in our modeling framework. First, the DNN model uncertainty was subject to the accuracy of climate forcings as well as other physical driving variables simulated by the physical wildfire model (ELMv1). For example, in addition to the default GSWP3 climate forcings dataset used in the study, CRU-JRA [*Onogi et al.*, 2007] and NCEP-DOE2 [*Kanamitsu et al.*, 2002] reanalysis forcings were also widely used and potentially different from GSWP3 forcings. ELMv1 used climate forcing (*e.g.*, temperature, precipitation, wind speed, relative humidity) to simulate soil temperature, soil moisture, fuel load and so on. These simulated variables served as inputs for the DNN model and could also result in prediction uncertainty. It was challenging to eliminate the forcing uncertainties in this work, but we could at least evaluate the magnitude of these uncertainties. We ran the DNN-Fire-OBS model with alternative forcings of CRU-JRA, NCEP-DOE2, and CDAS soil moisture from 2001 to 2010 and compared the results with DNN-Fire-OBS driven by default inputs (GSWP3 climate and ELMv1 simulated soil moisture) (Figure S5). The results showed relatively larger uncertainties from climate forcing than that from soil moisture forcing particularly over the major fire regions (e.g., SHSA, SHAF, and NHAF). For fuel load, although no transient dataset of global living biomass existed yet, we directly compared the ELM model simulated biomass with the global estimate (GEOCARBON ~ 455 Pg C). We found that the modeled present-day biomass continuously increased from 425 to 470 Pg C and compared reasonably well with the global benchmark (Figure S6). Future work will focus on evaluating the uncertainties from dead fuel load and fuel temperature variables.

Second, the original ELMv1 wildfire model has a unified mathematical representation of how fuel, climate, and social-economic conditions control wildfire burned area [*F Li et al.*, 2012]. However, training one single DNN wildfire model across the globe will produce a model dominated by gridcells that have high burned area (*e.g.*, Africa). The performance of the trained DNN model, therefore, will likely have larger biases over the low fire gridcells although the globally aggregated burned area could be reasonable. We partly overcame this challenge by applying the widely used 14 GFED fire regions that assume unique and relatively uniform dynamics over each region [*Giglio et al.*, 2006b], and employed stratified random sampling method for training and testing datasets. Although the regionally specific wildfire model introduces additional complexity, it better represents distinct characteristics of wildfire activity over different climate regimes and biomes [*Zhu and Zhuang*, 2013; *Zou et al.*, 2019] and allows for future analyses of how the relevant controllers vary across the globe.

Finally, our GFED region-based parameterization strategy relied on the combination of climate and biome types, while an alternative parameterization strategy for DNN-Fire model could be based on plant functional type distributions. Based on our analysis, the PFT-based DNN-Fire model had similar performance compared with the GFED-based model (Figure S7, S8). Since the current version of the E3SM land model does not allow PFT changes driven by climate, both GFED-based and PFT-based models may not fully capture the changes of fire dynamics due to longer-time scale fire regimes changes.

4. Conclusions

In this study, we first surrogated the baseline ELMv1 wildfire model with a Deep Neural Network (DNN) approach (Pearson correlation coefficient = 0.91 (p value < 0.01), $R^2 = 0.79$). The development was based on inputs and outputs from the baseline ELMv1 wildfire simulation, which is process-based and reasonably simulates global burned area, although regional biases existed. We then calibrated the neural network weights using the years 2001-2010 observationally inferred burned area. The final calibrated DNN wildfire model (DNN-Fire-OBS) was shown to be more accurate over the 14 GFED regions. For example, reductions in absolute error over Africa, South America, and Europe were by ~90%. More importantly, the DNN-Fire-OBS model parameters could be calibrated within minutes, compared with traditional ELMv1 parameterization ensemble simulations that consume a large amount of computational time. The improved DNN-Fire-OBS model also accurately prognosed global and regional burned area in the five-year period following the training period from 2011 to 2015 (modeled 469-514 Mha yr^{-1}). We conclude that the improved surrogate wildfire model (DNN-Fire-OBS) developed in this study can serve as an effective alternative to the process-based fire model currently used in ELMv1. More broadly, we conclude that machine learning techniques can facilitate earth system model development, parameterization, and uncertainty reduction with high efficiency and accuracy.

Acknowledgements

This research was supported by Energy Exascale Earth System Modeling (E3SM, https://e3sm.org/) Project and the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computation (RUBISCO) Scientific Focus Area, Office of Biological and Environmental Research of the U.S. Department of Energy Office of Science. Lawrence Berkeley National Laboratory (LBNL) is managed by the University of California for the U.S. Department of Energy under contract DE-AC02-05CH11231.

Author contribution

Q.Z., W.J.R, designed the study, Q.Z., W.J.R, L.X.., and J.T.R designed model experiments, Q.Z. and F.L. wrote code and run experiments, L.Z, K.Y, H.W., J.G all contribute to the results interpretation, and writing.

Code availability

https://github.com/qzhu-lbl/ANN_wildfire

Data availability

GFEDv4s: https://daac.ornl.gov/VEGETATION/guides/fire_emissions_v4.html

Fire_CCI51: https://geogra.uah.es/fire_cci/firecci51.php

Fire_CCILT11: https://geogra.uah.es/fire_cci/fireccilt11.php

MCD64: https://modis-fire.umd.edu/files/MODIS_C6_Fire_User_Guide_C.pdf

Fire_Atlas: https://www.globalfiredata.org/fireatlas.html

FireMIP model outputs: https://zenodo.org/record/3555562/accessrequest

Reference

Abatzoglou, J. T., and A. P. Williams (2016), Impact of anthropogenic climate change on wildfire across western US forests, *Proceedings of the National Academy of Sciences*, *113*(42), 11770-11775.

Andela, N., D. Morton, L. Giglio, Y. Chen, G. Van Der Werf, P. Kasibhatla, R. DeFries, G. Collatz, S. Hantson, and S. Kloster (2017), A human-driven decline in global burned area, *Science*, *356*(6345), 1356-1362.

Andela, N., D. C. Morton, L. Giglio, R. Paugam, Y. Chen, S. Hantson, G. R. Van Der Werf, and J. T. Randerson (2019), The Global Fire Atlas of individual fire size, duration, speed and direction, *Earth System Science Data*, *11*(2), 529-552.

Arora, V. K., and G. J. Boer (2005), Fire as an interactive component of dynamic vegetation models, *Journal of Geophysical Research: Biogeosciences*, *110*(G2).

Bond, W. J., F. I. Woodward, and G. F. Midgley (2005), The global distribution of ecosystems in a world without fire, *New phytologist*, *165*(2), 525-538.

Bond-Lamberty, B., S. D. Peckham, D. E. Ahl, and S. T. Gower (2007), Fire as the dominant driver of central Canadian boreal forest carbon balance, *Nature*, *450*(7166), 89-92.

Bowd, E. J., S. C. Banks, C. L. Strong, and D. B. Lindenmayer (2019), Long-term impacts of wildfire and logging on forest soils, *Nature Geoscience*, *12*(2), 113-118.

Brando, P., B. Soares-Filho, L. Rodrigues, A. Assunção, D. Morton, D. Tuchschneider, E. Fernandes, M. Macedo, U. Oliveira, and M. Coe (2020), The gathering firestorm in southern Amazonia, *Science advances*, *6*(2), eaay1632.

Cecil, D. J., D. E. Buechler, and R. J. Blakeslee (2014), Gridded lightning climatology from TRMM-LIS and OTD: Dataset description, *Atmospheric Research*, *135*, 404-414.

Chambers, S., and F. Chapin (2002), Fire effects on surface-atmosphere energy exchange in Alaskan black spruce ecosystems: Implications for feedbacks to regional climate, *Journal of Geophysical Research: Atmospheres*, *107*(D1), FFR 1-1-FFR 1-17.

Chen, Y., J. T. Randerson, S. R. Coffield, E. Foufoula-Georgiou, P. Smyth, C. A. Graff, D. C. Morton, N. Andela, G. R. van der Werf, and L. Giglio (2020), Forecasting global fire emissions on subseasonal to seasonal (S2S) time scales, *Journal of advances in modeling earth systems*, *12*(9), e2019MS001955.

Chen, Y., J. T. Randerson, D. C. Morton, R. S. DeFries, G. J. Collatz, P. S. Kasibhatla, L. Giglio, Y. Jin, and M. E. Marlier (2011), Forecasting fire season severity in South America using sea surface temperature anomalies, *Science*, *334*(6057), 787-791.

Clark, T. L., J. Coen, and D. Latham (2004), Description of a coupled atmosphere–fire model, *International Journal of Wildland Fire*, *13*(1), 49-63.

Coffield, S. R., C. A. Graff, Y. Chen, P. Smyth, E. Foufoula-Georgiou, and J. T. Randerson (2019), Machine learning to predict final fire size at the time of ignition, *International journal of wildland fire*.

Day, C. (2004), Smoke from burning vegetation changes the coverage and behavior of clouds, *PhT*, *57*(5), 24-24.

Dirmeyer, P. A., X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki (2006), GSWP-2: Multimodel analysis and implications for our perception of the land surface, *Bulletin of the American Meteorological Society*, *87*(10), 1381-1398.

Dobson, J. E., E. A. Bright, P. R. Coleman, R. C. Durfee, and B. A. Worley (2000), LandScan: a global population database for estimating populations at risk, *Photogrammetric engineering and remote sensing*, *66*(7), 849-857.

Finney, M. A. (1998), *FARSITE, Fire Area Simulator--model development and evaluation*, US Department of Agriculture, Forest Service, Rocky Mountain Research Station.

French, N. H., M. A. Whitley, and L. K. Jenkins (2016), Fire disturbance effects on land surface albedo in Alaskan tundra, *Journal of Geophysical Research: Biogeosciences*, *121*(3), 841-854.

Ganapathi Subramanian, S., and M. Crowley (2018), Using spatial reinforcement learning to build forest wildfire dynamics models from satellite images, *Frontiers in ICT*, *5*, 6.

Giglio, L., L. Boschetti, D. P. Roy, M. L. Humber, and C. O. Justice (2018), The Collection 6 MODIS burned area mapping algorithm and product, *Remote sensing of environment*, *217*, 72-85.

Giglio, L., I. Csiszar, and C. O. Justice (2006a), Global distribution and seasonality of active fires as observed with the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) sensors, *Journal of geophysical research: Biogeosciences*, *111*(G2).

Giglio, L., J. T. Randerson, and G. R. Van Der Werf (2013), Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4), *Journal of Geophysical Research: Biogeosciences*, *118*(1), 317-328.

Giglio, L., G. Van der Werf, J. Randerson, G. Collatz, and P. Kasibhatla (2006b), Global estimation of burned area using MODIS active fire observations.

Girardin, M. P., and M. Mudelsee (2008), Past and future changes in Canadian boreal wildfire activity, *Ecological Applications*, *18*(2), 391-406.

Goodfellow, I., Y. Bengio, and A. Courville (2016), *Deep learning*, MIT press Cambridge.

Goss, M., D. L. Swain, J. T. Abatzoglou, A. Sarhadi, C. A. Kolden, A. P. Williams, and N. S. Diffenbaugh (2020), Climate change is increasing the likelihood of extreme autumn wildfire conditions across California, *Environmental Research Letters*, *15*(9), 094016.

Harden, J. W., K. L. Manies, M. R. Turetsky, and J. C. Neff (2006), Effects of wildfire and permafrost on soil organic matter and soil climate in interior Alaska, *Global Change Biology*, *12*(12), 2391-2403.

Heyerdahl, E. K., L. B. Brubaker, and J. K. Agee (2002), Annual and decadal climate forcing of historical fire regimes in the interior Pacific Northwest, USA, *The Holocene*, *12*(5), 597-604.

Holden, Z. A., A. Swanson, C. H. Luce, W. M. Jolly, M. Maneta, J. W. Oyler, D. A. Warren, R. Parsons, and D. Affleck (2018), Decreasing fire season precipitation increased recent western US forest wildfire activity, *Proceedings of the National Academy of Sciences*, *115*(36), E8349-E8357.

Hurtt, G. C., L. Chini, R. Sahajpal, S. Frolking, B. L. Bodirsky, K. Calvin, J. C. Doelman, J. Fisk, S. Fujimori, and K. K. Goldewijk (2020), Harmonization of global land-use change and management for the period 850–2100 (LUH2) for CMIP6, *Geoscientific Model Development Discussions*, 1-65.

Jiang, Y., X.-Q. Yang, X. Liu, Y. Qian, K. Zhang, M. Wang, F. Li, Y. Wang, and Z. Lu (2020), Impacts of wildfire aerosols on global energy budget and climate: The role of climate feedbacks, *Journal of Climate*, *33*(8), 3351-3366.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. Hnilo, M. Fiorino, and G. Potter (2002), Ncep–doe amip-ii reanalysis (r-2), *Bulletin of the American Meteorological Society*, *83*(11), 1631-1644.

Kasischke, E. S., and L. P. Bruhwiler (2002), Emissions of carbon dioxide, carbon monoxide, and methane from boreal forest fires in 1998, *Journal of Geophysical Research: Atmospheres*, *107*(D1), FFR 2-1-FFR 2-14.

Keeley, J. E., and A. D. Syphard (2018), Historical patterns of wildfire ignition sources in California ecosystems, *International journal of wildland fire*, *27*(12), 781-799.

Kelley, D. I., I. Bistinas, R. Whitley, C. Burton, T. R. Marthews, and N. Dong (2019), How contemporary bioclimatic and human controls change global fire regimes, *Nature Climate Change*, *9*(9), 690-696.

Kingma, D. P., and J. Ba (2014), Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980*.

Knorr, W., T. Kaminski, A. Arneth, and U. Weber (2014), Impact of human population density on fire frequency at the global scale, *Biogeosciences*, *11*(4), 1085-1102.

Koven, C. D., W. J. Riley, Z. M. Subin, J. Y. Tang, M. S. Torn, W. D. Collins, G. B. Bonan, D. M. Lawrence, and S. C. Swenson (2013), The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4, *Biogeosciences*, *10*, 7109-7131.

Lamarque, J. F., J. T. Kiehl, G. P. Brasseur, T. Butler, P. Cameron-Smith, W. D. Collins, W. J. Collins, C. Granier, D. Hauglustaine, and P. G. Hess (2005), Assessing future nitrogen deposition and carbon cycle feedback using a multimodel approach: Analysis of nitrogen deposition, *Journal of Geophysical Research: Atmospheres*, *110*(D19).

Lasslop, G., S. Hantson, S. P. Harrison, D. Bachelet, C. Burton, M. Forkel, M. Forrest, F. Li, J. R. Melton, and C. Yue (2020), Global ecosystems and fire: Multi-model assessment of fire-induced tree-cover and carbon storage reduction, *Global change biology*, *26*(9), 5027-5041.

Lenihan, J. M., and D. Bachelet (2015), Historical climate and suppression effects on simulated fire and carbon dynamics in the conterminous United States, *Global Vegetation Dynamics: Concepts and Applications in the MC1 Model, edited by: Bachelet, D. and Turner, D., AGU Geophysical Monographs*, *214*, 17-30.

Li, F., M. Val Martin, M. O. Andreae, A. Arneth, S. Hantson, J. W. Kaiser, G. Lasslop, C. Yue, D. Bachelet, and M. Forrest (2019), Historical (1700–2012) global multi-model estimates of the fire emissions from the Fire Modeling Intercomparison Project (FireMIP), *Atmospheric Chemistry and Physics*, *19*(19), 12545-12567.

Li, F., X. Zeng, and S. Levis (2012), A process-based fire parameterization of intermediate complexity in a Dynamic Global Vegetation Model, *Biogeosciences*, *9*(7).

Lizundia-Loiola, J., G. Otón, R. Ramo, and E. Chuvieco (2020), A spatio-temporal active-fire clustering approach for global burned area mapping at 250 m from MODIS data, *Remote Sensing of Environment*, *236*, 111493.

Lizundia-Loiola, J., M. Pettinari, E. Chuvieco, T. Storm, and J. Gómez-Dans (2018), ESA CCI ECV Fire Disturbance: Algorithm Theoretical Basis Document-MODIS, version 2.0, edited, Fire_cci_D2.

Mahowald, N., T. D. Jickells, A. R. Baker, P. Artaxo, C. R. Benitez-Nelson, G. Bergametti, T. C. Bond, Y. Chen, D. D. Cohen, and B. Herut (2008), Global distribution of atmospheric phosphorus sources, concentrations and deposition rates, and anthropogenic impacts, *Global Biogeochemical Cycles*, *22*(4).

Mekonnen, Z. A., W. J. Riley, J. T. Randerson, R. F. Grant, and B. M. Rogers (2019), Expansion of high-latitude deciduous forests driven by interactions between climate warming and fire, *Nature plants*, *5*(9), 952-958.

Oliver, A. K., M. A. Callaham Jr, and A. Jumpponen (2015), Soil fungal communities respond compositionally to recurring frequent prescribed burning in a managed southeastern US forest ecosystem, *Forest Ecology and Management*, *345*, 1-9.

Onogi, K., J. Tsutsui, H. Koide, M. Sakamoto, S. Kobayashi, H. Hatsushika, T. Matsumoto, N. Yamazaki, H. Kamahori, and K. Takahashi (2007), The JRA-25 reanalysis, *Journal of the Meteorological Society of Japan. Ser. II*, *85*(3), 369-432.

Papakosta, P., G. Xanthopoulos, and D. Straub (2017), Probabilistic prediction of wildfire economic losses to housing in Cyprus using Bayesian network analysis, *International journal of wildland fire*, *26*(1), 10-23.

Pellegrini, A. F., A. Ahlström, S. E. Hobbie, P. B. Reich, L. P. Nieradzik, A. C. Staver, B. C. Scharenbroch, A. Jumpponen, W. R. Anderegg, and J. T. Randerson (2018), Fire frequency drives decadal changes in soil carbon and nitrogen and ecosystem productivity, *Nature*, *553*(7687), 194-198.

Pellegrini, A. F., S. E. Hobbie, P. B. Reich, A. Jumpponen, E. J. Brookshire, A. C. Caprio, C. Coetsee, and R. B. Jackson (2020), Repeated fire shifts carbon and nitrogen cycling by changing plant inputs and soil decomposition across ecosystems, *Ecological Monographs*, e01409.

Preisler, H. K., and A. L. Westerling (2007), Statistical model for forecasting monthly large wildfire events in western United States, *Journal of Applied Meteorology and Climatology*, *46*(7), 1020-1030.

Prentice, S., and D. Mackerras (1977), The ratio of cloud to cloud-ground lightning flashes in thunderstorms, *Journal of Applied Meteorology*, *16*(5), 545-550.

Rabin, S. S., J. R. Melton, G. Lasslop, D. Bachelet, M. Forrest, S. Hantson, J. O. Kaplan, F. Li, S. Mangeon, and D. S. Ward (2017), The Fire Modeling Intercomparison Project (FireMIP), phase 1: experimental and analytical protocols with detailed model descriptions, *Geoscientific Model Development*, *10*(3), 1175-1197.

Radke, D., A. Hessler, and D. Ellsworth (2019), FireCast: Leveraging Deep Learning to Predict Wildfire Spread, paper presented at IJCAI.

Randerson, J. T., H. Liu, M. G. Flanner, S. D. Chambers, Y. Jin, P. G. Hess, G. Pfister, M. Mack, K. Treseder, and L. Welp (2006), The impact of boreal forest fire on climate warming, *science*, *314*(5802), 1130-1132.

Riley, K., and M. Thompson (2017), An uncertainty analysis of wildfire modeling, *Natural hazard uncertainty assessment: modeling and decision support. Monograph*, *223*, 193-213.

Ross, A. N., M. J. Wooster, H. Boesch, and R. Parker (2013), First satellite measurements of carbon dioxide and methane emission ratios in wildfire plumes, *Geophysical Research Letters*, *40*(15), 4098-4102.

Rother, D., and F. De Sales (2020), Impact of Wildfire on the Surface Energy Balance in Six California Case Studies, *Boundary-Layer Meteorology*, 1-24.

Saha, M. V., T. M. Scanlon, and P. D'Odorico (2019), Climate seasonality as an essential predictor of global fire activity, *Global Ecology and Biogeography*, *28*(2), 198-210.

Sayad, Y. O., H. Mousannif, and H. Al Moatassime (2019), Predictive modeling of wildfires: A new dataset and machine learning approach, *Fire safety journal*, *104*, 130-146.

Schmidhuber, J. (2015), Deep learning in neural networks: An overview, *Neural networks*, *61*, 85-117.

Stephenson, C., J. Handmer, and R. Betts (2013), Estimating the economic, social and environmental impacts of wildfires in Australia, *Environmental Hazards*, *12*(2), 93-111.

Syphard, A. D., V. C. Radeloff, J. E. Keeley, T. J. Hawbaker, M. K. Clayton, S. I. Stewart, and R. B. Hammer (2007), Human influence on California fire regimes, *Ecological applications*, *17*(5), 1388-1402.

Teckentrup, L., G. Lasslop, D. Bachelet, M. Forrest, S. Hantson, F. Li, J. R. Melton, C. Yue, A. Arneth, and S. P. Harrison (2018), Simulations of historical burned area: A comparison of global fire models in FireMIP, *EGUGA*, 17537.

Thonicke, K., A. Spessa, I. Prentice, S. P. Harrison, L. Dong, and C. Carmona-Moreno (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.

Tonini, M., M. D'Andrea, G. Biondi, S. Degli Esposti, A. Trucchia, and P. Fiorucci (2020), A Machine Learning-Based Approach for Wildfire Susceptibility Mapping. The Case Study of the Liguria Region in Italy, *Geosciences*, *10*(3), 105.

Van Der Werf, G. R., J. T. Randerson, L. Giglio, T. T. Van Leeuwen, Y. Chen, B. M. Rogers, M. Mu, M. J. Van Marle, D. C. Morton, and G. J. Collatz (2017), Global fire emissions estimates during 1997-2016, *Earth System Science Data*, *9*(2), 697-720.

van Vuuren, D. P., P. L. Lucas, and H. Hilderink (2007), Downscaling drivers of global environmental change: Enabling use of global SRES scenarios at the national and grid levels, *Global environmental change*, *17*(1), 114-130.
Venevsky, S., K. Thonicke, S. Sitch, and W. Cramer (2002), Simulating fire regimes in human-dominated ecosystems: Iberian Peninsula case study, *Global Change Biology*, *8*(10), 984-998.
Walker, X. J., J. L. Baltzer, S. G. Cumming, N. J. Day, C. Ebert, S. Goetz, J. F. Johnstone, S. Potter, B. M. Rogers, and E. A. Schuur (2019), Increasing wildfires threaten historic carbon sink of boreal forest soils, *Nature*, *572*(7770), 520-523.
Wang, J.-F., A. Stein, B.-B. Gao, and Y. Ge (2012), A review of spatial sampling, *Spatial Statistics*, *2*, 1-14.
Westerling, A. L., H. G. Hidalgo, D. R. Cayan, and T. W. Swetnam (2006), Warming and earlier spring increase western US forest wildfire activity, *science*, *313*(5789), 940-943.
Williams, A. P., J. T. Abatzoglou, A. Gershunov, J. Guzman-Morales, D. A. Bishop, J. K. Balch, and D. P. Lettenmaier (2019), Observed impacts of anthropogenic climate change on wildfire in California, *Earth's Future*, *7*(8), 892-910.
Xu, X., G. Jia, X. Zhang, W. J. Riley, and Y. Xue (2020), Climate regime shift and forest loss amplify fire in Amazonian forests, *Global Change Biology*, *26*(10), 5874-5885.
Yu, Y., J. Mao, P. E. Thornton, M. Notaro, S. D. Wullschleger, X. Shi, F. M. Hoffman, and Y. Wang (2020), Quantifying the drivers and predictability of seasonal changes in African fire, *Nature Communications*, *11*(1), 1-8.
Yue, X., L. J. Mickley, J. A. Logan, and J. O. Kaplan (2013), Ensemble projections of wildfire activity and carbonaceous aerosol concentrations over the western United States in the mid-21st century, *Atmospheric Environment*, *77*, 767-780.
Zheng, H., Z. Yang, W. Liu, J. Liang, and Y. Li (2015), Improving deep neural networks using softplus units, paper presented at 2015 International Joint Conference on Neural Networks (IJCNN), IEEE.
Zhu, Q., and W. J. Riley (2015), Improved modelling of soil nitrogen losses, *Nature Climate Change*, *5*(8), 705-706.
Zhu, Q., W. J. Riley, C. M. Iversen, and J. Kattge (2020), Assessing impacts of plant stoichiometric traits on terrestrial ecosystem carbon accumulation using the E3SM land model, *Journal of Advances in Modeling Earth Systems*, *12*(4), e2019MS001841.
Zhu, Q., W. J. Riley, J. Tang, N. Collier, F. M. Hoffman, X. Yang, and G. Bisht (2019), Representing nitrogen, phosphorus, and carbon interactions in the E3SM Land Model: Development and global benchmarking, *Journal of Advances in Modeling Earth Systems*, doi: 10.1029/2018MS001571.
Zhu, Q., W. J. Riley, J. Tang, and C. D. Koven (2016), Multiple soil nutrient competition between plants, microbes, and mineral surfaces: model development, parameterization, and example applications in several tropical forests, *Biogeosciences*, *13*, 341-363, doi:10.5194/bgd-12-4057-2015.
Zhu, Q., and Q. Zhuang (2013), Improving the quantification of terrestrial ecosystem carbon dynamics over the United States using an adjoint method, *Ecosphere*, *4*(10), doi: dx.doi.org/10.1890/ES1813-00058.00051.

Zhu, Q., and Q. Zhuang (2014), Parameterization and sensitivity analysis of a process-based terrestrial ecosystem model using adjoint method, *Journal of Advances in Modeling Earth Systems*, *6*(2), 315-331.
Zou, Y., Y. Wang, Z. Ke, H. Tian, J. Yang, and Y. Liu (2019), Development of a REgion-specific ecosystem feedback fire (RESFire) model in the Community Earth System Model, *Journal of Advances in Modeling Earth Systems*, *11*(2), 417-445.
Zou, Y., Y. Wang, Y. Qian, H. Tian, J. Yang, and E. Alvarado (2020), Using CESM-RESFire to understand climate-fire-ecosystem interactions and the implications for decadal climate variability, *Atmospheric Chemistry and Physics*, *20*(PNNL-SA-150222).

Supplementary Material

Table S1. Burned area datasets used in this study

Dataset name	Temporal range	Spatial resolution	Burned area, mean (std)	Citations
GFEDv4s	1997-2015	0.25 degree	455(39)	(Van Der Werf, Randerson et al. 2017)
Fire_CCI51	2001-2019	0.25 degree	476(26)	(Lizundia-Loiola, Otón et al. 2020)
Fire_CCILT11	1982-2018	0.25 degree	484(20)	(Lizundia-Loiola, Pettinari et al. 2018)
MCD64	2001-2019	0.25 degree	424(35)	(Giglio, Boschetti et al. 2018)
Fire_Atlas	2003-2016	0.25x0.25 degree	459(43)	(Andela, Morton et al. 2019)

Note: the long-term average global burned area was calculated using data with the same overlapping temporal range (2003-2015), unit Mha yr^{-1}

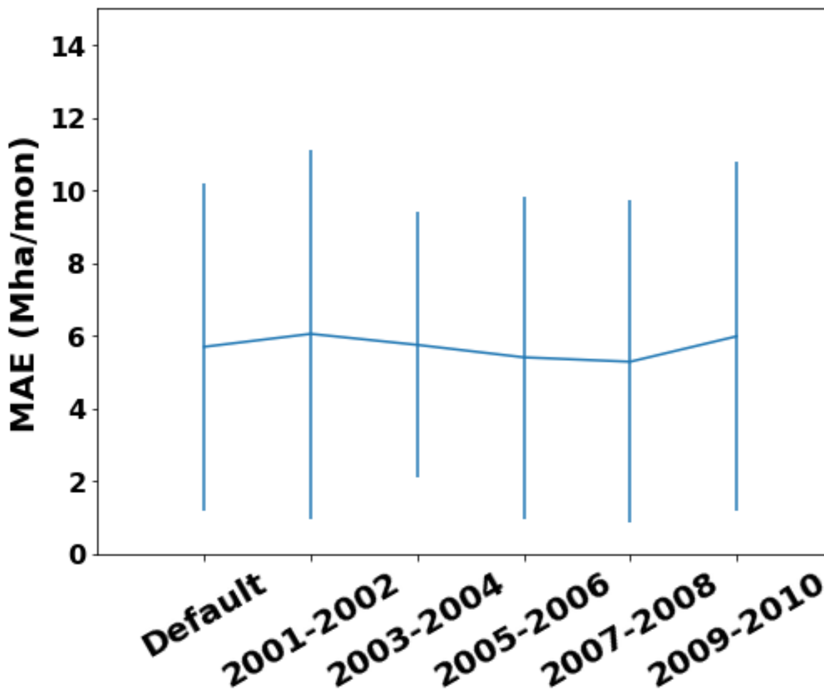

Figure S1. Model performance evaluated with testing datasets of default (20% randomly selected samples), or fixed to 2001-2002 period, 2003-2004 period, 2005-2006 period, 2007-2008 period, and 2009-2010 periods (the rest of the dataset was used as a training dataset.).

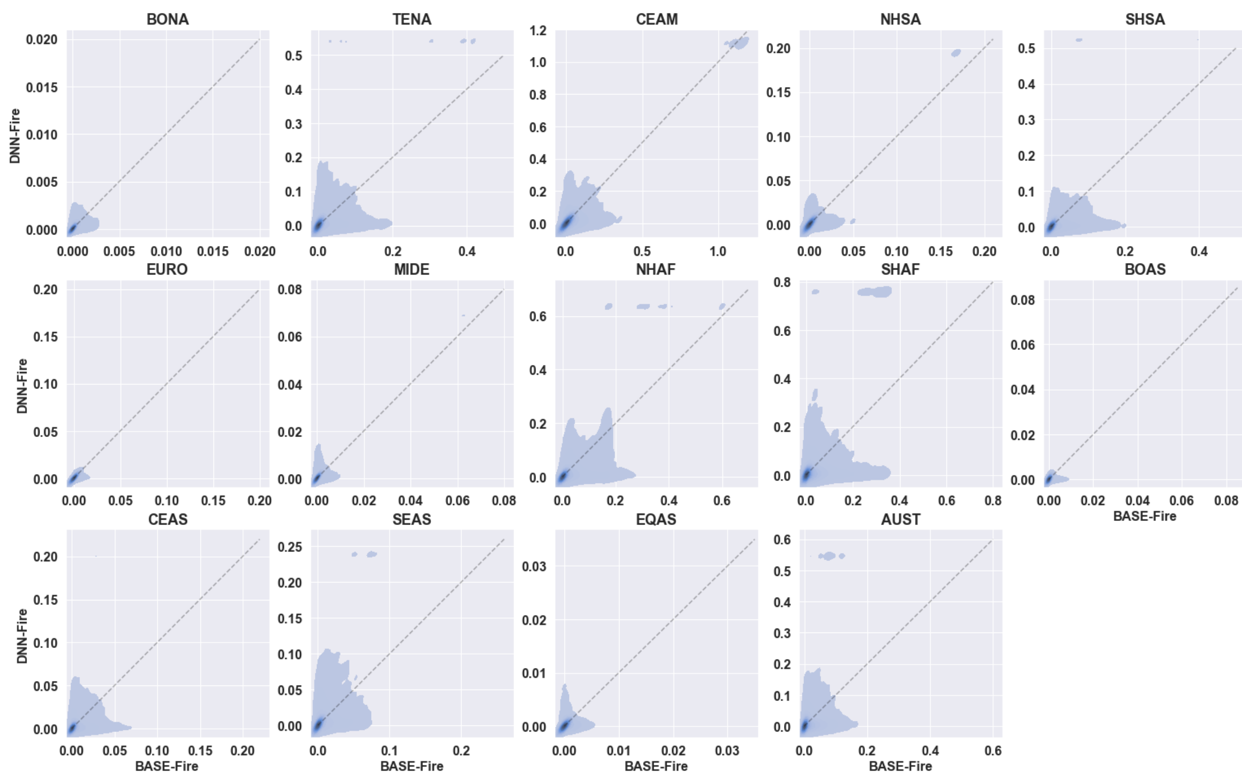

Figure S2. Performance of surrogate model (DNN-Fire) compared with ELMv1 process-based model (BASE-Fire).

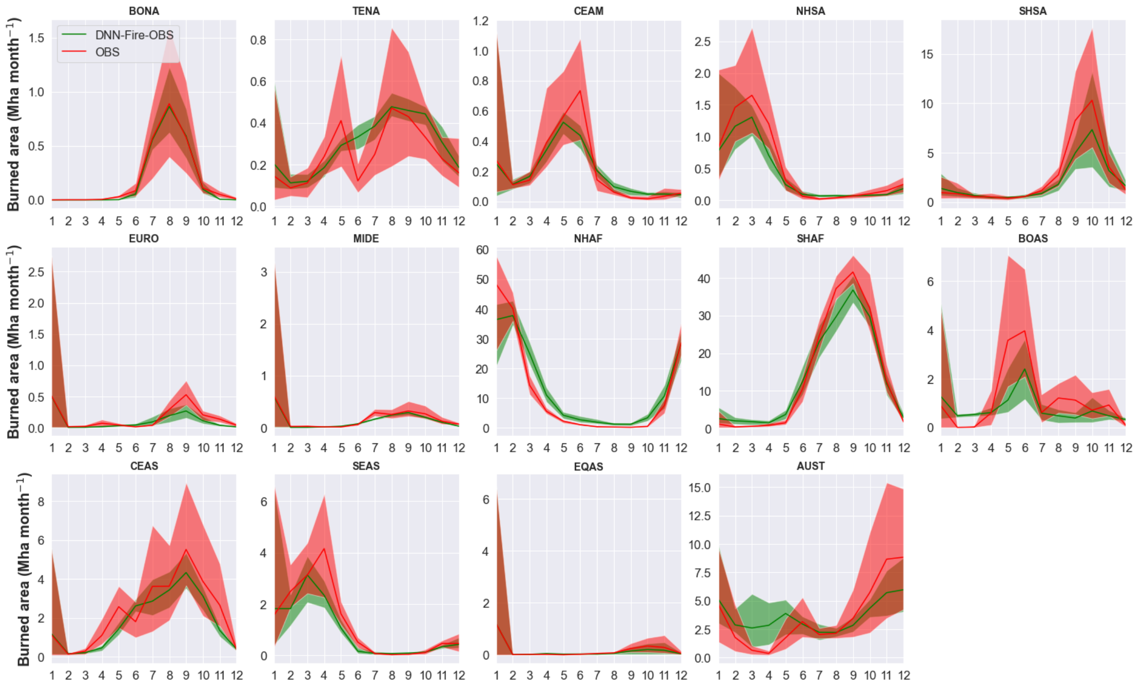

Figure S3. Seasonal cycles of fine-tuned Deep Neural Network wildfire model (DNN-Fire-OBS) and observations over 14 GFED fire regions.

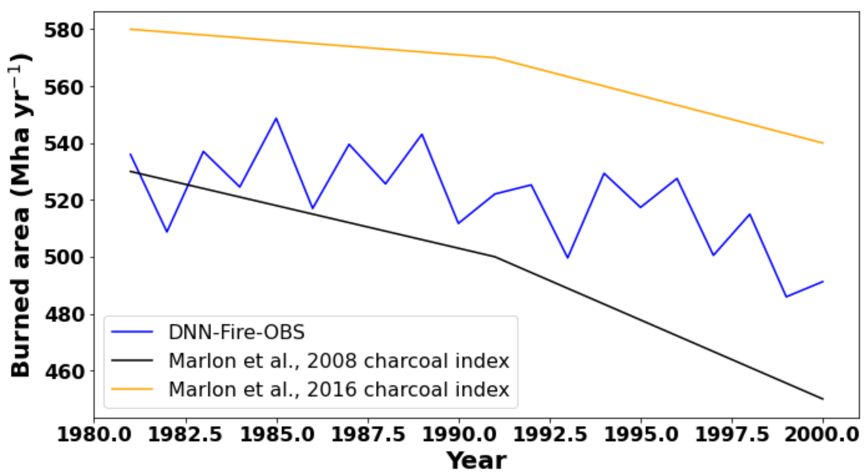

Figure S4. Comparison of DNN-Fire-OBS model simulated global burned area during 1981-1999 with two charcoal index inferred burned area.

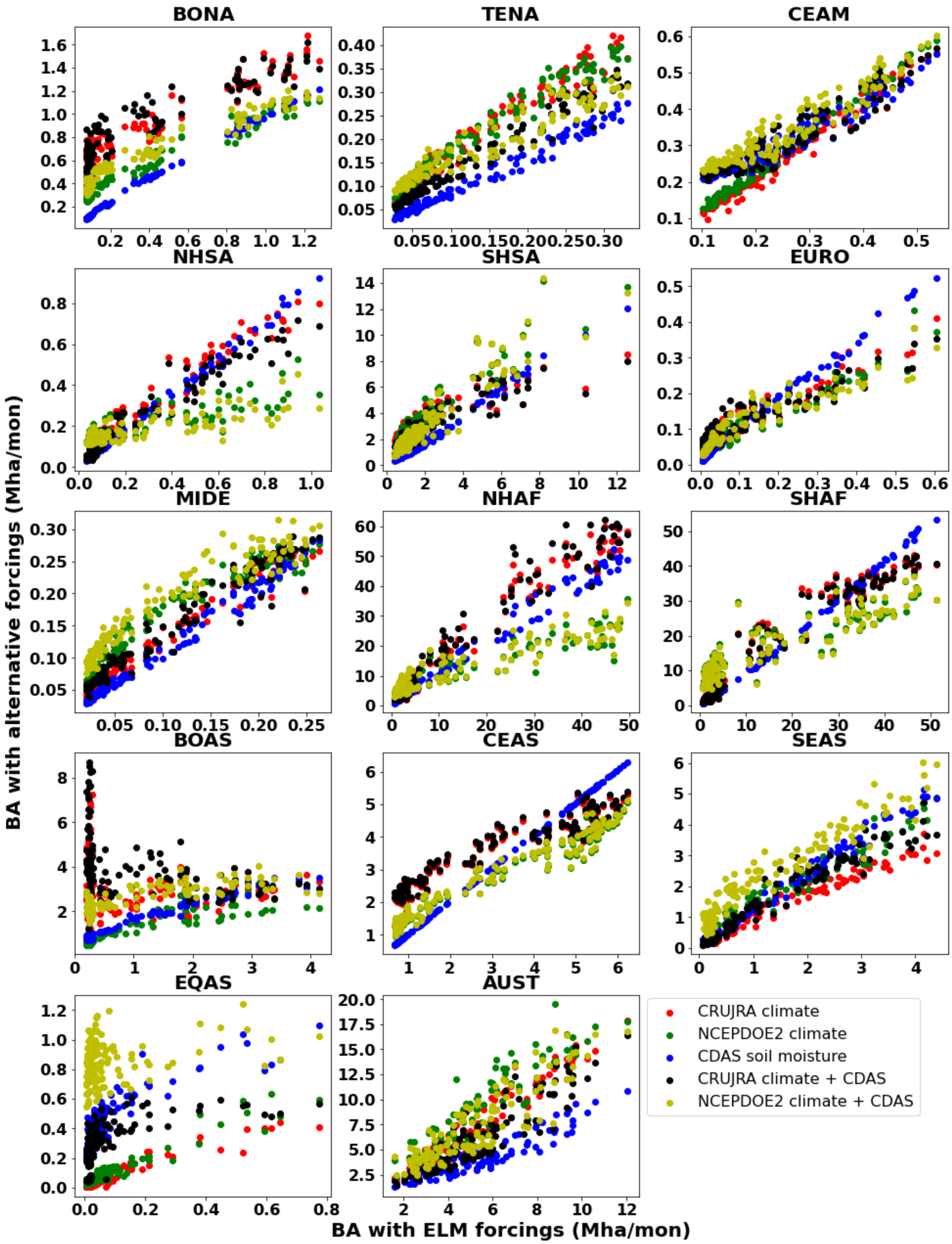

Figure S5. Sensitivity of modeled burned area (2001-2010 long-term averaged) to climate forcings (including temperature, precipitation, wind speed, relative humidity) and soil moisture.

X-axis was burned area simulated by the default model using GSWP3 climate forcing and ELMv1 simulated soil moisture. Y-axis were models with alternative climate forcing (CRUJRA, NCEPDOE2) and soil moisture product (NCEP CDAS soil moisture).

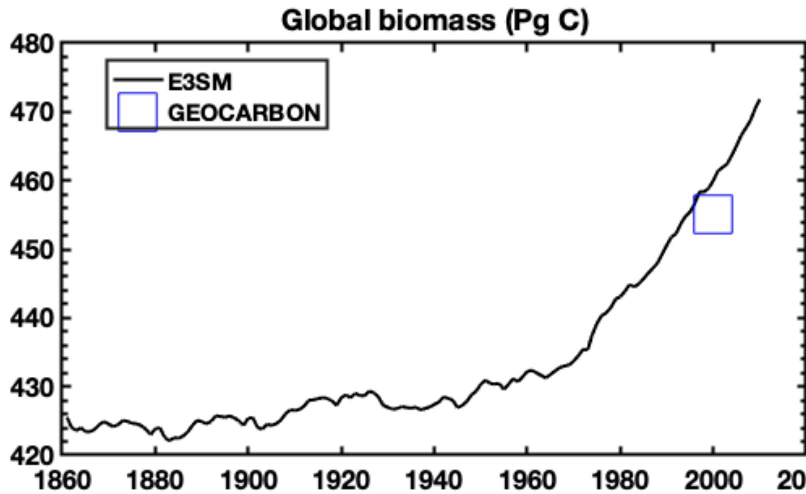

Figure S6. 3SM simulated global vegetation biomass [425-472 PgC] and observational based estimate of present-day living biomass (455 PgC GEOCARBON).

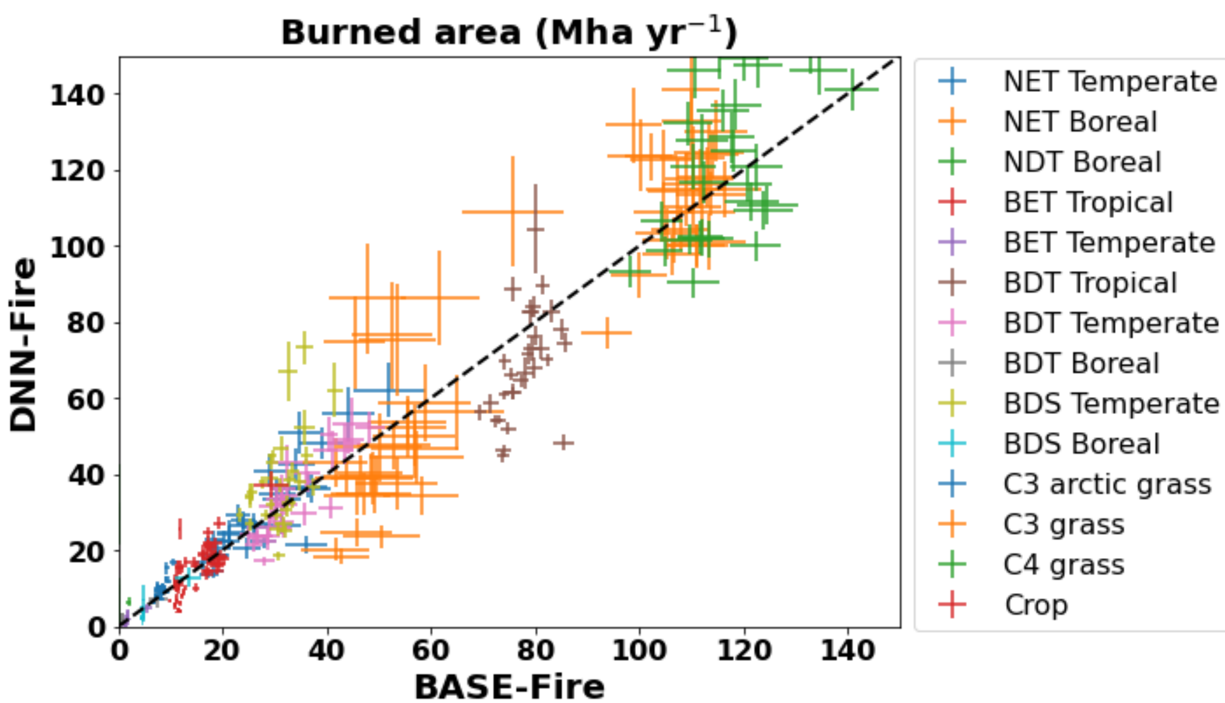

Figure S7. The performance of the Deep Neural Network wildfire model (DNN-Fire), compared with the original ELMv1 process-based wildfire model (BASE-Fire) aggregated over 14 plant functional types between years 2001 and 2010.

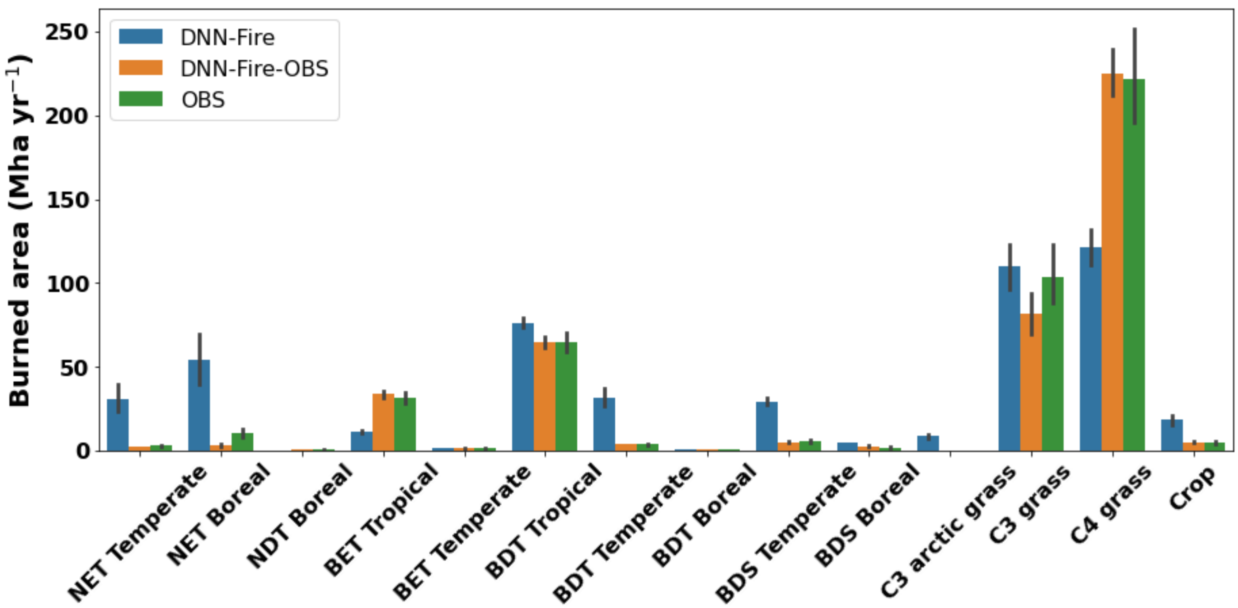

Figure S8. A comparison of wildfire burned area among Deep Neural Network wildfire model (DNN-Fire), Deep Neural Network wildfire model fine-tuned with observed burned area (DNN-Fire-OBS), and observations for 14 plant functional types.