AttentionFire_v1.0: interpretable machine learning fire model for burned area predictions over tropics

Fa Li[1,2], Qing Zhu[1,*], William J. Riley[1], Lei Zhao[3], Li Xu[4], Kunxiaojia Yuan[1,2], Min Chen[5], Huayi Wu[2], Zhipeng Gui[6], Jianya Gong[6], James T. Randerson[4]

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

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

[5] Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, Madison, WI, USA

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

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

Abstract

African and South American (ASA) wildfires account for more than 70% of global burned areas and have strong connection to local climate for sub-seasonal to seasonal wildfire dynamics. However, representation of the wildfire-climate relationship remains challenging, due to spatiotemporally heterogenous responses of wildfires to climate variability and human influences. Here, we developed an interpretable Machine Learning (ML) fire model (AttentionFire_v1.0) to resolve the complex controls of climate and human activities on burned area and to better predict burned areas over ASA regions. Our ML fire model substantially improved predictability of burned area for both spatial and temporal dynamics compared with five commonly used machine learning models. More importantly, the model revealed strong time-lagged control from climate wetness on the burned areas. The model also predicted that under a high emission future climate scenario, the recently observed declines in burned area will reverse in South America in the near future due to climate changes. Our study provides reliable and interpretable fire model and highlights the importance of lagged wildfire-climate relationships in historical and future predictions.

1. Introduction

Wildfires modify land surface characteristics, such as vegetation composition, soil carbon, surface runoff, and albedo, with significant consequences for regional carbon, water, and energy cycles (Benavides-Solorio and Macdonald, 2001; Shvetsov et al., 2019; Randerson et al., 2006). Over African and South American (ASA) regions, where more than 70% of global burned area occurs, wildfires emit ~1.4 PgC yr^{-1} (~65% of global wildfire emissions (Werf et al., 2017a)) and dust and aerosols that can alter regional climate through radiative processes (Werf et al., 2017b; Etminan et al., 2016; Ramanathan et al., 2001). While greenhouse gas emissions contribute to climate change, other toxic species and airborne particulate matter from wildfires lead to substantial health hazards, including elevated premature mortality (Knorr et al., 2017; Lelieveld et al., 2015). In particular, wildfire particulate matter emissions across tropical regions have exceeded current anthropogenic sources and are predicted to dominate future regional emissions (Knorr et al., 2017).

Although total tropical wildfire burned area has declined over the past few decades due to climate change and human activities (Andela and Van Der Werf, 2014; Andela et al., 2017) (*e.g.*, from increases in population density, cropland fraction, and livestock density), wildfire still plays a significant role in mediating surface climate (Xu et al., 2020), biogeochemical cycles, and human health (Andela et al., 2017). Further, 21st century projections of increases in temperature, regional drought (Dai, 2013; Taufik et al., 2017), and precipitation variations may outweigh these direct human impacts and result in unprecedentedly fire-prone environments over a large fraction of Africa (Van Der Werf et al., 2008; Andela and Van Der Werf, 2014; Archibald et al., 2009) and South America (Pechony and Shindell, 2010; Malhi et al., 2008). These factors highlight the need for better understanding, prediction, and management of these critical fire regions to minimize economic losses, human health hazards, and natural ecosystem degradation. Therefore, improved understanding and accurate prediction of wildfire activity is increasingly important for effective fire management and sustainable decision-making.

Climate is acknowledged as one of the most dominant controllers on ASA wildfires (Chen et al., 2011; Andela et al., 2017). For example, precipitation variations contribute substantially to burned area patterns in southern and northern Africa (Andela and Van

Der Werf, 2014; Archibald et al., 2009), and are also closely linked to wildfire spatiotemporal dynamics in south America (Chen et al., 2011; Van Der Werf et al., 2008; Malhi et al., 2008). More importantly, the strong controls of climate on wildfires often show time-lags and the time-delay can be on the order of multiple months (Van Der Werf et al., 2008; Andela and Van Der Werf, 2014). Meanwhile, ocean dynamics (e.g., El Niño-Southern Oscillation, ENSO) may also exert considerable influences on ASA wildfires through influencing wet and wet-to-dry season climate and fuel conditions (Yu et al., 2020; Chen et al., 2016; Andela and Van Der Werf, 2014; Chen et al., 2011; Chen et al., 2017). The time-lags between ocean dynamics and wildfires can be even longer than that between climate and wildfires (Chen et al., 2020), which enable wildfire predictions ahead of fire season (Chen et al., 2011; Chen et al., 2016; Chen et al., 2020; Turco et al., 2018). The spatiotemporal responses of wildfires to climate changes are complicated by non-linear interactions among climate, vegetation, and human activities (Van Der Werf et al., 2008; Andela et al., 2017). In more xeric subtropical regions, increasing precipitation during the wet season can be the dominant controller on increasing wildfire during the following dry season (through regulation of fuel availability and fuel spatial structures) (Van Der Werf et al., 2008; Littell et al., 2009; Archibald et al., 2009). In contrast, increasing precipitation in more mesic regions results in excessive fuel moisture, thereby becoming the main limitation of dry-season wildfires (i.e., opposite fire trends are observed with increasing precipitation in northern and southern Africa) (Van Der Werf et al., 2008; Andela and Van Der Werf, 2014). In addition to natural processes, human activities are primary ignition sources and have shaped fire patterns in the ASA regions (Aragao et al., 2008; Archibald et al., 2009; Andela et al., 2017). Fire-use types driven by local socio-economic conditions and fire management policies may also affect the fire-climate relationships (Andela et al., 2017). Therefore, strong climate controls from wet season to dry season need to be considered along with fuel distributions and human activities for continental fire predictions under climate change.

Accurate predictive modeling of wildfire with skillful representation of how environmental and anthropogenic factors modulate the burned area is still challenging. State-of-the-art process-based fire models (*e.g.*, the Fire Model Intercomparison Project (Rabin et al., 2017)) have reasonably simulated the spatial distribution of burned areas. However, they generally do not accurately capture burned area seasonal variation and inter-annual trends and variability (Andela et al., 2017). Improving predictability and reducing uncertainties of process-based models require more sophisticated representation of fire processes and parameterization, which remain a long-term challenge (Bowman et al., 2009; Hantson et al., 2016; Teckentrup et al., 2019). In response to this challenge, data-driven statistical or Machine Learning (ML) approaches have been developed and demonstrated to effectively capture wildfire severity and burned area dynamics (Archibald et al., 2009; Chen et al., 2020; Chen et al., 2011; Zhou et al., 2020). However, the spatially heterogenous, non-linear, and time-lagged controls have been oversimplified (e.g., using linear models or only considering climate variables at specific time lags or seasons (Chen et al., 2011; Chen et al., 2016; Chen et al., 2020; Archibald et al., 2009; Gray et al., 2018a)) or have been black boxed.

For example, the commonly used neural network or deep learning models (Zhu et al., 2022; Joshi and Sukumar, 2021) themselves are complex and built upon hidden neural layers with non-linear activation functions and thus cannot directly identify the relative importance of different drivers for wildfires (Murdoch et al., 2019; Jain et al., 2020). A few ML models (e.g., decision tree and random forest) provide variable importance, however, such importance scores are constant across the entire dataset rather than spatiotemporally varied (Wang et al., 2021a; Yuan et al., 2022b). While post-hoc analyses could interpret ML models (Altmann et al., 2010; Lundberg and Lee, 2017), inconsistent and unstable explanations can be derived with different post-hoc methods or settings (Slack et al., 2021; Molnar et al., 2020). Such limitations impede an interpretable and reliable way to understand the critical spatiotemporal processes from wet season to dry season (Reichstein et al., 2019; Jain et al., 2020).

In this work, we developed a wildfire model (AttentionFire) leveraging on an interpretable Long-Short-Term-Memory (LSTM) framework to predict wildfire burned areas over Northern Hemisphere Africa (NHAF), Southern Hemisphere Africa (SHAF), and Southern Hemisphere South America (SHSA) (Giglio et al., 2013). We also focused on using the AttentionFire model to explore the dependency of simulated burned area on different drivers from wet season to dry season across different gridcells. We assessed model predictability with observed burned area from Global Fire Emission Database (GFED) and compared with five other machine learning based fire models.

2. Methods

2.1 AttentionFire model

The AttentionFire model is based on an interpretable attention-augmented LSTM (Liang et al., 2018; Qin et al., 2017; Guo et al., 2019; Li et al., 2020; Vaswani et al., 2017) framework. Like the traditional artificial neural network (ANN) models, the LSTM is also built upon neurons and the non-linear activation functions; specifically, the LSTM uses the gating mechanism (i.e., forget, input, and output gates) (Hochreiter and Schmidhuber, 1997; Wang and Yuan, 2019) to filter out useless information while keeping useful information underlying in the time series as hidden states (Fig. 1). Relative to traditional ANN, the LSTM has shown advantages in capturing short- and long-term dependencies in input time series (Hochreiter and Schmidhuber, 1997), such as the time-lagged controls from wet-to-dry season climate conditions on wildfires. However, LSTM cannot explicitly and dynamically select important drivers from multiple driving time series to make predictions (Qin et al., 2017; Liang et al., 2018; Guo et al., 2019; Li et al., 2020; Vaswani et al., 2017). Further, LSTM works as a black-box, lacking interpretability to identify the relative importance of each driver across different time steps (Guo et al., 2019; Li et al., 2020; Liang et al., 2018). Attention mechanisms overcome these challenges by adaptively assigning larger weights to more important drivers and time steps (Liang et al., 2018; Vaswani et al., 2017). Here we use attention mechanism to explicitly capture controlling factors of fire predictions with various time-lags (Fig. 1). Below are detailed descriptions of the fire model.

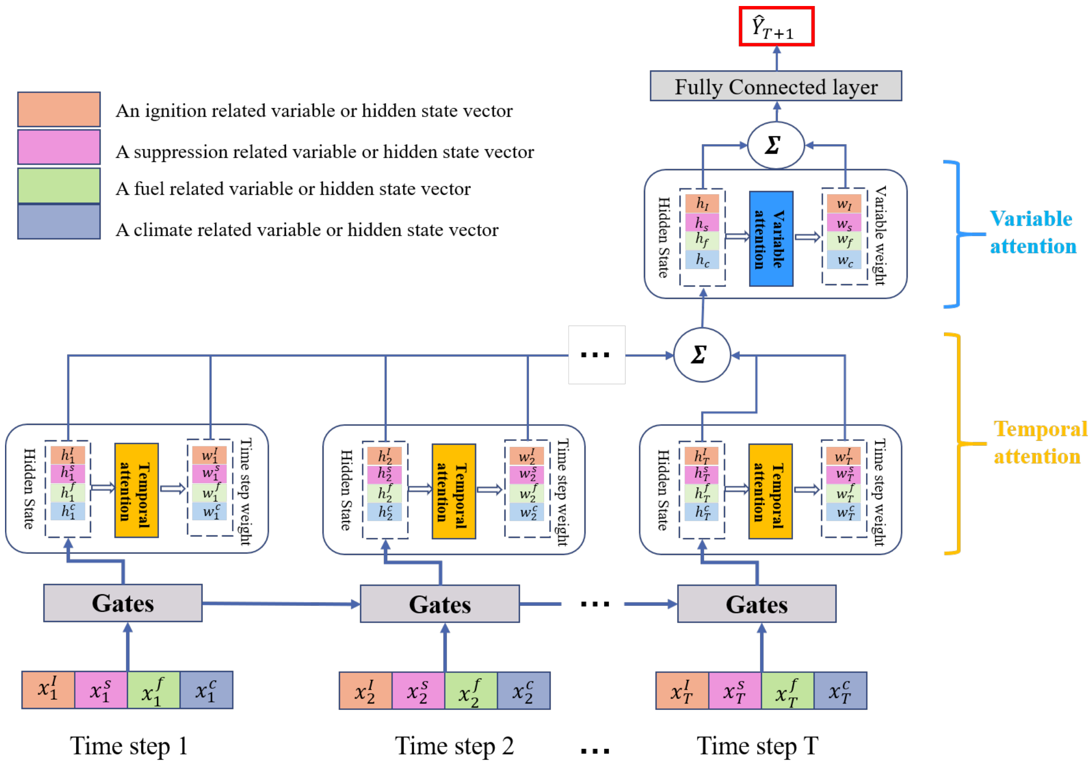

Fig. 1: An illustrative workflow for AttentionFire_v1.0 model prediction. Four kinds of drivers are considered: ignition related, suppression related, fuel, and climate. The temporal attention is used to identify important time steps for each kind of driver, while the variable attention is used to identify important drivers for final burned area prediction.

Given four categories of time series, $X = (X^{\mathrm{l}}, X^{\mathrm{s}}, X^{\mathrm{f}}, X^{\mathrm{c}})^T$, where T is the length of time series, we use $X^i = (x_1^i, x_2^i, \ldots, x_T^i)^T \in R^T$, where $1 \leq i \leq n$, to denote the i-th time series, and use $X_t = (x_t^1, x_t^2, \ldots, x_t^n)^T \in R^n$, where $1 \leq t \leq \mathrm{T}$, to represent the vector at time step t. x_{t}^l, x_{t}^s, x_{t}^f, and x_{t}^c represent the variables of ignition (*e.g.*, population density), suppression (*e.g.*, road network density), fuel availability (*e.g.*, living biomass), and climate (*e.g.*, precipitation) at time step t. The AttentionFire model aims to learn a nonlinear function F to map the n time series to the observed burned area Y_{T+1} at time step $\mathrm{T}+1$:

$$\hat{Y}_{T+1} = F(X^{\mathrm{l}}, X^{\mathrm{s}}, X^{\mathrm{f}}, X^{\mathrm{c}})^{T} \tag{1}$$

Where $\hat{Y}_{T+1}$is the predicted burned area at time step $\mathrm{T}+1$.

First, the model iteratively transforms the i-th driving variable at time step t to a hidden state vector h_t^i, where $1 \leq t \leq \mathrm{T}$ and $1 \leq i \leq n$ through LSTM gate mechanisms (please refer to Li et al. (2020) for the details of the Gates in Fig. 1). Second, as the importance of each time step varies, temporal attention is applied to h_t^i to calculate its corresponding weight or importance w_t^i. Third, the weighted summation h_{sum}^i of h_t^i is obtained to represent the summarized information for the i-th driving variable:

$$\begin{gathered} w_t^i = f_{attn}(h_t^i) \\ h_{sum}^i = \sum_{t=1}^{T} w_t^i h_t^i \end{gathered} \tag{2}$$

Where $h_t^i \in R^m$ is the hidden state vector of the i-th driving series at time step t, that stores the summary of the past input sequence (Hochreiter and Schmidhuber, 1997). w_t^i is the calculated weight for the i-th driver at time step t through attention function f_{attn}:

$$\begin{gathered} w_t^{i'} = \tanh\left(W_p h_t^i\right) \\ w_t^i = \frac{e^{w_t^{i'}}}{\sum_{j=1}^{T} e^{w_t^{j'}}} \end{gathered} \tag{3}$$

where $W_p \in R^{1\times m}$, is a parameter matrix that needs to be learned.

To furtherly capture the relative importance of the i-th driving variable compared to other driving variables, variable attention is used for the summarized information h_{sum}^{i} and h_T^i. Note that h_T^i is also a kind of summarized information derived by the LSTM (Hochreiter and Schmidhuber, 1997; Guo et al., 2019). The weight or importance of the i-th driving variable w_i is calculated as:

$$w_i' = \tanh\,(W_a[h_{sum}^{i}, h_T^i])$$
$$w_i = \frac{e^{w_i'}}{\sum_{j=1}^{n} e^{w_j'}} \tag{4}$$

Finally, using the weighted sum of all driving variables, the model generates the prediction $\hat{Y}_{T+1}$:

$$o_i = W_o[h_{sum}^{i}, h_T^i] + b_o$$
$$\hat{Y}_{T+1} = \sum_{i=1}^{n} o_i w_i \tag{5}$$

where $W_a \in R^{1\times 2m}$, is a learnable parameter matrix and the linear function with weight $W_o \in R^m$ and bias $b_o \in R$, along with attention calculated weight w_i, produce the final prediction result. The parameters of attention-based LSTM are learned via a back-propagation algorithm by minimizing the mean-squared error between predictions and observations (Guo et al., 2019; Leung and Haykin, 1991).

The AttentionFire model is implemented with python under Python 3 environment. The model is open-access at https://zenodo.org/record/6903284#.YvH8F-zMJmP under Creative Commons Attribution 4.0 International license. Detailed code and descriptions are included in the repository including loading datasets, model initialization, training, predicting, saving parameters, and loading the trained model (see more details in code availability section).

2.2 Baseline models and model settings

Five other widely used Machine learning (ML) models are used as baseline models to compare with AttentionFire model: ANN (Joshi and Sukumar, 2021; Zhu et al., 2021), decision tree (DT) (Amatulli et al., 2006; Coffield et al., 2019), random forest (RF) (Gray et al., 2018b; Yu et al., 2020; Li et al., 2018), gradient boosting decision tree (GBDT) (Coffield et al., 2019; Jain et al., 2020), and naive LSTM (Liang et al., 2019; Natekar et al., 2021; Gui et al., 2021; Mei and Li, 2019). The details of baseline models selected, including strengths, potential limitations, and their applications in wildfire studies and references are listed in Table 1. The ANN and LSTM have shown good performance on multiple earth science problems (Yuan et al., 2022a; Reichstein et al., 2019) including wildfires (Joshi and Sukumar, 2021; Zhu et al., 2021; Liang et al., 2019), however, the black-box nature of such models makes them lack interpretability. The DT method provides variable importance and is easily interpretable with its single tree structure, but prone to overfitting compared to RF and GBDT. The RF alleviates the overfitting through feature selection and ensemble learning (Breiman, 2001) while the GBDT avoids overfitting by constructing multiple trees with shallow depth (Ke et al., 2017). DT, RF, and GBDT provide variable importance scores for dominant driver inference, however, such importance scores are constant across the entire dataset and thus impede detailed interpretation of the variable importance like over space and time. The aforementioned ML models have been commonly used in wildfire science (Jain et al., 2020).

The inputs of climate and fuel-related variables for the first four models (non-sequence models) are variables of the latest three month available for prediction (Yu et al., 2020) while the corresponding inputs of naive LSTM and AttentionFire models are whole-year historical time sequences which cover dynamics from wet to dry seasons to capture short- and long-term dependencies underlying the input sequence(Qin et al., 2017; Vaswani et al., 2017; Guo et al., 2019; Li et al., 2020). The socioeconomic predictors (i.e., population, road density, livestock) consider only the more recent and available statistics typically reported at a year scale. For each model, we iteratively leave one-year dataset out (i.e., a holdout dataset that model has never seen) for testing, one year data for validation (to avoid overfitting during training (Yuan et al., 2022b; Jabbar and Khan, 2015)), and use the remaining dataset for model training (i.e., tunning model parameters). Details of the settings for used models in experiments are listed in Table S1.

Table 1. Strengths, potential limitations, and applications of selected baseline models in wildfire studies.

Model (acronym)	Strengths	Potential limitations	Applications
Random Forest (RF) (Breiman, 2001)	Provide variable importance; Alleviate overfitting through feature selection and ensemble learning;	Constant variable importance rather than varied; time-consuming when building large trees; may not perform well on time series with lags	(Gray et al., 2018b; Yu et al., 2020)
Decision Tree (DT) (Safavian and Landgrebe, 1991)	Provide variable importance; easy to interoperate the single tree	Prone to overfitting; constant variable importance rather than varied; time-consuming when building a large tree; may not perform well on time series with lags	(Amatulli et al., 2006; Coffield et al., 2019)

Gradient Boosting Decision Tree (GBDT) (Ke et al., 2017)	Alleviate overfitting by building multiple shallow trees; generally fast because of the shallowness of each tree built	Constant variable importance rather than varied; may not perform well on time series with lags	(Coffield et al., 2019; Jain et al., 2020)
Artificial Neural Network (ANN) (Ke et al., 2017)	Show good performance on complex and non-linear problems; alleviate overfitting through techniques like dropout and regularization	Lack of interpretability; hard to know the optimal neural network structures for different problems	(Joshi and Sukumar, 2021; Zhu et al., 2021)
Long-Short-Term-Memory (LSTM) (Hochreiter and Schmidhuber, 1997)	Show good performance on time series predictions; alleviate overfitting through techniques like dropout and regularization	Lack of interpretability; may not be suitable for non-time series problems; vanishing gradient problem when deployed to long time series (Li et al., 2020; Liang et al., 2018)	(Liang et al., 2019; Natekar et al., 2021)

2.3 Datasets and experiments

Satellite-based global burned area dataset (Global Fire Emissions Database (Giglio et al., 2013)) is used as prediction target, and datasets of various socio-environmental drivers are used as model inputs. Population density, livestock density, road-network density, and land use are considered as anthropogenic factors on fire ignition and spread. Fuel variables include fuel moisture, live and dead vegetation biomass. Seven meteorology variables from NCEP-DOE Reanalysis are considered, including air temperature, precipitation, surface pressure, wind speed, specific humidity, downward shortwave radiation, and vapor pressure deficit. Details of each dataset and corresponding references are listed in Table 2. The raw datasets were unified to the same spatial resolution (T62 resolution: ~210 km at the equator) at the monthly scale with a covering period from 1997 to 2015.

In addition to the local socio-environmental drivers, we also explored the impacts of ocean indices on burned area predictions. Chen et al. (2011) found that wildfires in South America were closely linked to the Oceanic Niño Index (ONI), and Atlantic multidecadal Oscillation (AMO) index. The ONI and AMO reflected the sea surface temperature (SST) anomalies in the tropical Pacific and north Atlantic. The SST anomalies directly affected ocean-atmosphere interactions and thus the wet, wet-to-dry, and onset of dry season climate in South America (Chen et al., 2011). The two indexes were significantly correlated with peak fire month wildfires 3 to 7 months later and could predict fire season wildfires in many regions of South America with lead times of 3 to 5 months (Chen et al., 2011). The controls of SST anomalies in tropical Pacific on climate and thus wildfires were also found in northern and southern Africa (Andela and Van Der Werf, 2014). In addition, SST anomalies in tropical northern and southern Atlantic could also affect wildfires in South America (Chen et al., 2016) and Africa (Yu et al., 2020; Chen et al., 2020). Therefore, we included ocean indices (Table 2) and investigated their impacts on wildfire predictions with the AttentionFire model (see section 3.4).

Table 2. Input and output variables and datasets of the AttentionFire model.

Variable category	**Variables (abbreviation, units)**	**Spatial (temporal) resolution**	**Dataset and reference**
Wildfire	Burned area (BA, hectares month^{-1})	0.25 degree (monthly)	Global Fire Emissions Database 4 (Giglio et al., 2013)
Climate	Precipitation (RAIN, mm s^{-1}), temperature (TA, K), surface air pressure (PA, Pa), specific humidity	~1.9 degree (monthly)	NCEP-DOE Reanalysis 2 (Kanamitsu et al.,

	(SH, kg kg^{-1}), downward short-wave radiation (SW, W m^{-2}), wind speed (WIND, m s^{-1}), vapor pressure deficit (VPD, hPa) (VPD calculated according to (Bolton, 1980))		2002)
Fuel conditions	Fuel moisture (FUELM, %), coarse wood debris (CWDC, gC m^{-2} s^{-1}), vegetation biomass (VegC, gC m^{-2} s^{-1}), litter biomass (LitterC, gC m^{-2} s^{-1})	~1.9 degree (monthly)	ELM prognostic simulations (Zhu et al., 2019)
Human activities	Population density (Popu, persons $grid^{-1}$)	~1km (yearly)	(Dobson et al., 2000)
	Road density (Road, km km^{-2})	0.5 degree (yearly)	(Meijer et al., 2018)
	Livestock density (LS, number of livestock $grid^{-1}$)	0.5 degree (yearly)	(Rothman-Ostrow et al., 2020)
Land cover	Bare soil (Bare, %), Forest (Forest, %), and Grass (Grass, %)	0.25 degree (yearly)	LUH2 (Hurtt et al., 2020a)
Oceanic indices	Ocean Niño Index (ONI), Atlantic multidecadal Oscillation (AMO) index, Tropical Northern Atlantic (TNA) Index, and Tropical Southern Atlantic (TSA) Index	monthly	NOAA Climate Indices (Noaa, 2021)

For future projection (2016-2055) of burned area with AttentionFire model, land use changes (Hurtt et al., 2020b), population growth, projected climate and fuel from fully coupled Earth System Model (ESM) simulations of CMIP6 (O'neill et al., 2016) under low (SSP126) and high (SSP585) emission scenarios were used as the ML model input, respectively. The reason to select 2016-2055 as the projected period was that during 2016-2055, the 99th percentiles of precipitation, temperature, and vapor pressure deficit were within the range of corresponding historical observations, which means that the trained model has covered the range of most projected drivers in the near-future and can alleviate extrapolation uncertainty caused by climate change. We also made longer projection till the end of 21st century and analyzed its longer-term trend (see section 3.4). All available ESMs with outputs of historical and future (SSP126 and

SSP585) fuel availability (i.e., biomass of coarse wood debris, vegetation, and litter) and climate variables (Table 2) were selected, including ACCESS-ESM1-5 (Ziehn et al., 2020), CESM2 (Danabasoglu et al., 2020), NorESM2-LM (Seland et al., 2020), NorESM2-MM (Seland et al., 2020), and TaiESM1(Wang et al., 2021b). For each ESM, the variable bias was corrected with the mostly used linear scaling method (Maraun, 2016; Dangol et al., 2022; Shrestha et al., 2017) which adjusted the bias in model simulations based on the ratio of modeled and observed variable mean value. Then the bias corrected variables of each ESM were used to drive AttentionFire model for future burned area projection. Finally, given the uncertainty of each ESM, the multi-model ensemble (MME) mean of projected burned area was calculated (Li et al., 2022) and analyzed. Details of the bias correction method can be found in Maraun (2016). For future projections, temporally constant road and livestock density were used due to the lack of future data in the two scenarios (i.e., SSP585 and SSP126), and the AttentionFire model was not coupled in the ESMs. Such limitation and uncertainty were discussed in section 3.5.

3 Results and Discussions

3.1 Model predictability on burned area spatial-temporal dynamics

The AttentionFire model accurately captured the spatial distribution and temporal variations (Fig. 2 and Fig. S1) of wildfire burned areas over NHAF, SHAF, and SHSA regions. The AttentionFire model had the lowest mean absolute errors (MAEs) between model predicted and observed (GFED) grided monthly burned areas among the six ML approaches. The gridded MAEs of burned area for AttentionFire were 110, 142, and 39 Kha yr^{-1} in NHAF, SHAF, and SHSA regions, which were respectively 6%~66%, 13%~65%, and 11%~42% lower than the other 5 ML approaches in the three regions. These results highlight the capability of the AttentionFire model to capture critical driving factors of burned area across time and space.

The fact that the AttentionFire model outperformed the other five models (Fig. 2g-i) indicates the benefit of skillfully integrating time-lagged and spatially heterogenous controls from critical drivers on wildfires. Compared to non-sequence models (i.e., RF, MLP, DT, and GBDT), the AttentionFire model adaptively captured historical dependencies of wildfires on climate conditions from wet to dry seasons (Van Der Werf et al., 2008; Archibald et al., 2009; Andela and Van Der Werf, 2014; Chen et al., 2011) (more detailed analysis is provided in next section). Compared to the naive LSTM models, the variable and temporal attention mechanisms integrated in AttentionFire has proven to be beneficial to model performance.

The spatial heterogeneity and temporal variation of wildfire responses to complex environmental and human factors have made wildfire predictions challenging, especially at large spatial scales (Chen et al., 2016; Littell et al., 2016; Andela and Van Der Werf, 2014; Chen et al., 2011; Zhou et al., 2020). The capability of the AttentionFire model to reasonably predict spatial and temporal distributions of burned area ahead of fire season allows more time to explore and implement management options, such as allocation of firefighting resources, fuel clearing or targeted burning restrictions (Chen et al., 2011).

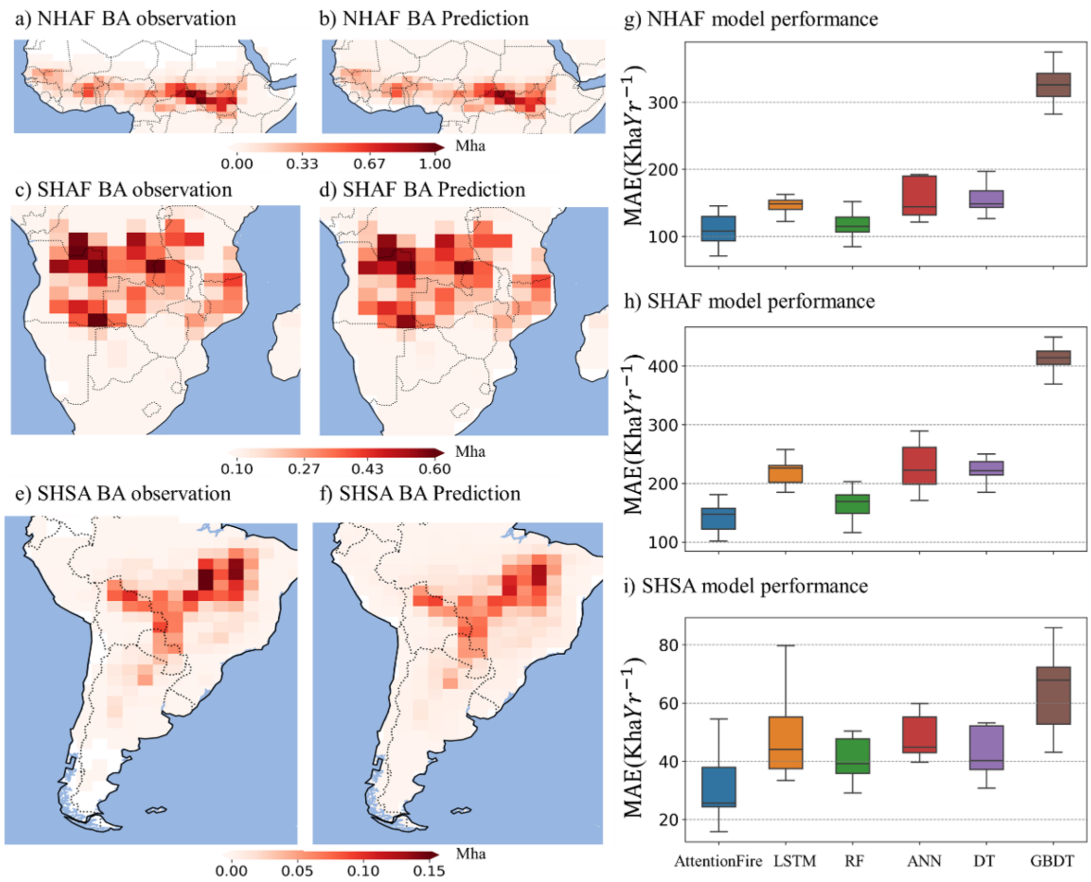

Fig. 2. The AttentionFire model accurately captured burned area spatial dynamics. Spatial distribution of observed and AttentionFire predicted fire season mean burned area (BA) with one-month lead time in Northern Hemisphere Africa (NHAF) (a-b), Southern Hemisphere Africa (SHAF) (c-d), and Southern Hemisphere South America (SHSA) (e-f) regions. (g-i) Performance (in terms of mean absolute error between predicted and observed burned area) of AttentionFire and other five baseline models, including Long-Short-Term-Memory (LSTM), random forest (RF), artificial neural network (ANN), decision tree (DT), and gradient boosting decision tree (GBDT).

3.2 Dominant drivers of tropical burned area dynamics

The AttentionFire model dynamically weights variable importance and highlights critical temporal windows (Qin et al., 2017; Vaswani et al., 2017; Liang et al., 2018;

Guo et al., 2019; Li et al., 2020) that maximize model predictability. Therefore, the variable weights could inform dominant physical processes, while the temporal weights reflect the temporal dependency structure, making it interpretable for spatial-temporal analysis. For the AttentionFire model predictions, the variable weights showed that climate wetness exerted strong and spatial heterogenous controls on burned areas. Specifically, precipitation (for SHAF and SHSA regions) and vapor pressure deficit (VPD; for NHAF region) played the most important roles (Fig. 3) in burned area prediction during fire seasons (defined as the four months with the largest burned areas, Fig. S2), and the control strengths from those climate wetness variables on fires were significantly (one-tailed t-test, p-value<0.05) stronger in regions with larger burned areas (gridcells with top 10% burned areas) than those with smaller burned areas (gridcells with last 90% burned areas) (Fig. 4 a-f).

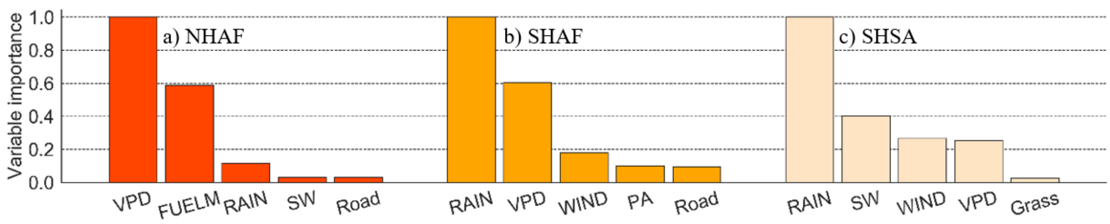

Fig. 3: Ranked top-five important variables for fire-season burned area in Northern Hemisphere Africa (NHAF) (a), Southern Hemisphere Africa (SHAF) (b), and Southern Hemisphere South America (SHSA) (c). For each gridcell within each study region, there is a mean variable weight, representing the importance of the variable for fire prediction in the gridcell. For each region, the variable weights are summed weighted by its corresponding mean burned areas, and normalized.

In AttentionFire model predictions, the precipitation and VPD explained ~66% - ~80% (Fig. S3) of the annual mean fire season wildfire burned areas. Variations of VPD and precipitation not only affect fire season ignition likelihood and fire spread (Sedano and Randerson, 2014; Holden et al., 2018) through fuel moisture, but also regulate vegetation growth, fuel structure (Gale et al., 2021) (e.g., fuel composition and spatial connectivity), and fuel availability (Mueller et al., 2020; Littell et al., 2009; Littell et al., 2016; Van Der Werf et al., 2008). The importance of these climate wetness variables confirms the dominant roles of local water balances and air dryness for wildfire prediction from sub-seasonal to seasonal scales (Littell et al., 2016; Archibald et al., 2009; Chen et al., 2011) , especially in regions with large burned areas.

Furthermore, we found that the emergent functional relationships between climate wetness and wildfire burned area were parabolic (Fig. S3): *i.e.*, enhancement of historical precipitation or decline of historical VPD (indicating wetter conditions) first increased burned area in more xeric conditions, then suppressed burned area under more mesic conditions, consistent with previous findings in subtropical regions (Andela and Van Der Werf, 2014; Van Der Werf et al., 2008). The transition points of these emergent functional relationships (thresholds at which the relationships reverse) were region specific, and these relationships may be useful for developing, tuning, and benchmarking wildfire models (Zhu et al., 2021).

For the time lags between those dominant climate wetness variables and fire-season burned areas, our results demonstrated that burned area over NHAF was more modulated by relatively short-term wetness (VPD during wet-to-dry and onset of dry season, from September to December), while SHAF and SHSA burned areas depended more on long-term wetness (precipitation during wet and wet-to-dry season, December to March in SHAF, and November to April in SHSA) (Fig. 4g-i). The short-term variations of climate wetness can directly affect near-surface temperature and moisture availability, which affect fuel flammability (Littell et al., 2016; Holden et al., 2018), while the long-term wetness (*e.g.*, during rainy season) can affect fuel availability, composition, and spatial connectivity, which can result in even stronger long time-lagged controls on dry-season burned areas (Abatzoglou and Kolden, 2013; Littell et al., 2016; Chen et al., 2011; Van Der Werf et al., 2008; Archibald et al., 2009; Andela and Van Der Werf, 2014).

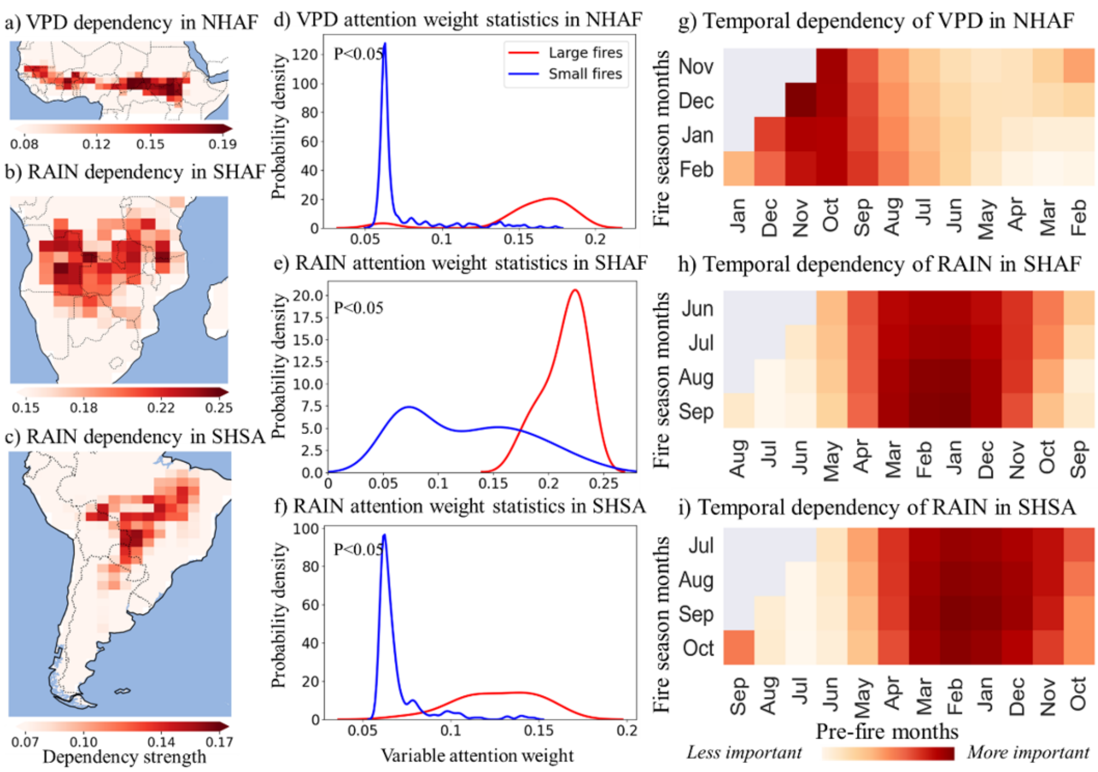

Fig. 4. Spatial-temporal importance of climate wetness variables for burned area dynamics. (a-c) Spatial importance of climate wetness variables for fire-season burned areas. (d-f) statistical comparison of the climate wetness variable importance over regions with large and small burned areas. (g-i) fire season burned area dependency on the history of the climate wetness driver over Northern Hemisphere Africa (NHAF), Southern Hemisphere Africa (SHAF), and Southern Hemisphere South America (SHSA) regions.

Previous work has shown that when and where fires occurred during dry season can be affected by precipitation induced fuel availability patterns during wet and during wet-to-dry transition seasons in savannah ecosystems (Van Der Werf et al., 2008; Archibald et al., 2009; Andela and Van Der Werf, 2014). Also, precipitation variations during wet and wet-to-dry transition seasons in the tropical forest ecosystem can affect soil recharge during wet season and further affect plant transpiration, local surface humidity, and precipitation during the following dry season (Chen et al., 2011; Ramos Da Silva et al., 2008; Malhi et al., 2008). The exact responses of fires to short-and-long term climate variations depend on both local wetness and fuel conditions (*e.g.*, fires in wetter ecosystems with enough fuel availability can be mainly limited by the length of dry season, while fires in drier ecosystems can be limited by fuel availability during wet season (Van Der Werf et al., 2008; Andela and Van Der Werf, 2014)). Therefore, an effective way of integrating the climate wetness history (*i.e.*, AttentionFire model) can lead to more accurate predictions of burned area spatial-temporal dynamics.

3.3 Possible usage of oceanic index for long-leading time predictions

In ASA regions, large-scale variations of oceanic dynamics can directly influence local climate (e.g., precipitation variations during wet seasons (Chen et al., 2011; Andela and Van Der Werf, 2014)) through time-lagged controls of teleconnections and indirectly influence fires during following dry seasons (Chen et al., 2016; Chen et al., 2011; Andela et al., 2017). Therefore, we hypothesized that ocean dynamics might benefit AttentionFire model predictions, especially for long leading time fire predictions through providing additional information that has not been reflected in local climate and land surface conditions (Chen et al., 2016; Chen et al., 2011; Andela et al., 2017; Chen et al., 2020; Ma et al., 2022).

We compared model performance for short term (1-4 month ahead), and long term (5-8 month ahead) fire predictions with and without considering the four oceanic indices (OIs). Relative to the MAE of short-term predictions, the mean MAE of long-term predictions without and with teleconnections increased by ~34% and ~14% in NHAF, ~34% and ~15% in SHAF, and ~17% and ~7% in SHSA, respectively, indicating the decline of system predictability with longer leading time (Figure 5). However, for long-term predictions, including OIs could decrease the mean MAE by ~20%, ~19%, and ~11% in NHAF, SHAF, and SHSA regions, respectively, compared with the case without oceanic indexes. While the mean variable importance of OIs was consistently lower that of local climate (Fig. S4) across the three regions, the OIs did provide additional information for long-term predictions with lower biases (Fig. 5). The results demonstrated the potential usage of teleconnections for long leading time burned area predictions (Chen et al., 2020; Chen et al., 2016; Chen et al., 2011).

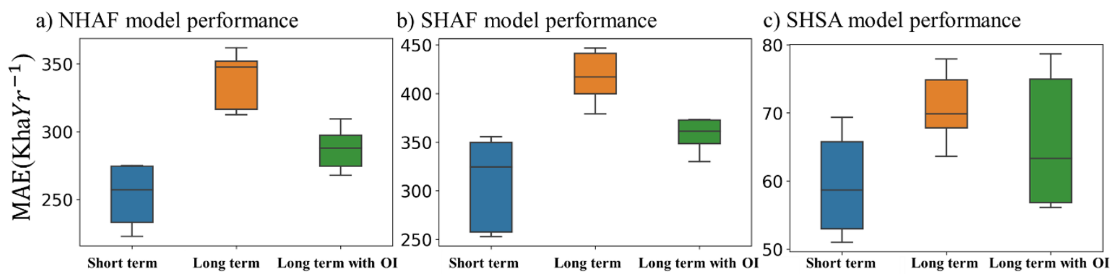

Fig. 5: Performance of AttentionFire burned area predictions with 1~4 month leading time (short-term) and with 5-8 month leading (long-term). MAE is mean absolute error. Four ocean indices which have been widely used for fire prediction over South American and African regions were considered for long-term forecast, including Oceanic Niño Index, Atlantic multidecadal Oscillation index, Tropical Northern Atlantic Index, and Tropical Southern Atlantic Index.

3.4 Future trends of burned area over Africa and South America

Due to climate change and human activities (Andela et al., 2017), strong but opposing trends of burned areas have been observed in Northern (decreasing) and Southern (increasing) Hemisphere Africa (Andela and Van Der Werf, 2014), and within different regions of Southern Hemisphere America (Andela et al., 2017) during the recent two decades, resulting in an overall declining burned area trend in Africa and South America. However, whether this decline will persist is under debate. On one hand, the projected increases in population, expansion of agriculture, mechanized (fire-free) management, and fire suppression policies will likely continue to decrease burned areas (Andela and Van Der Werf, 2014) (e.g., human activities were regarded as one of the main drivers for fire decline in NHAF region). On the other hand, future climate change (Dai, 2013; Taufik et al., 2017) could outweigh human impacts and result in unprecedented fire-prone environments in the tropics (Pechony and Shindell, 2010;

Malhi et al., 2008) (e.g., fires showed strong dependency with climate wetness in NHAF, SHAF (Andela and Van Der Werf, 2014; Archibald et al., 2009) and SHSA (Chen et al., 2011) regions).

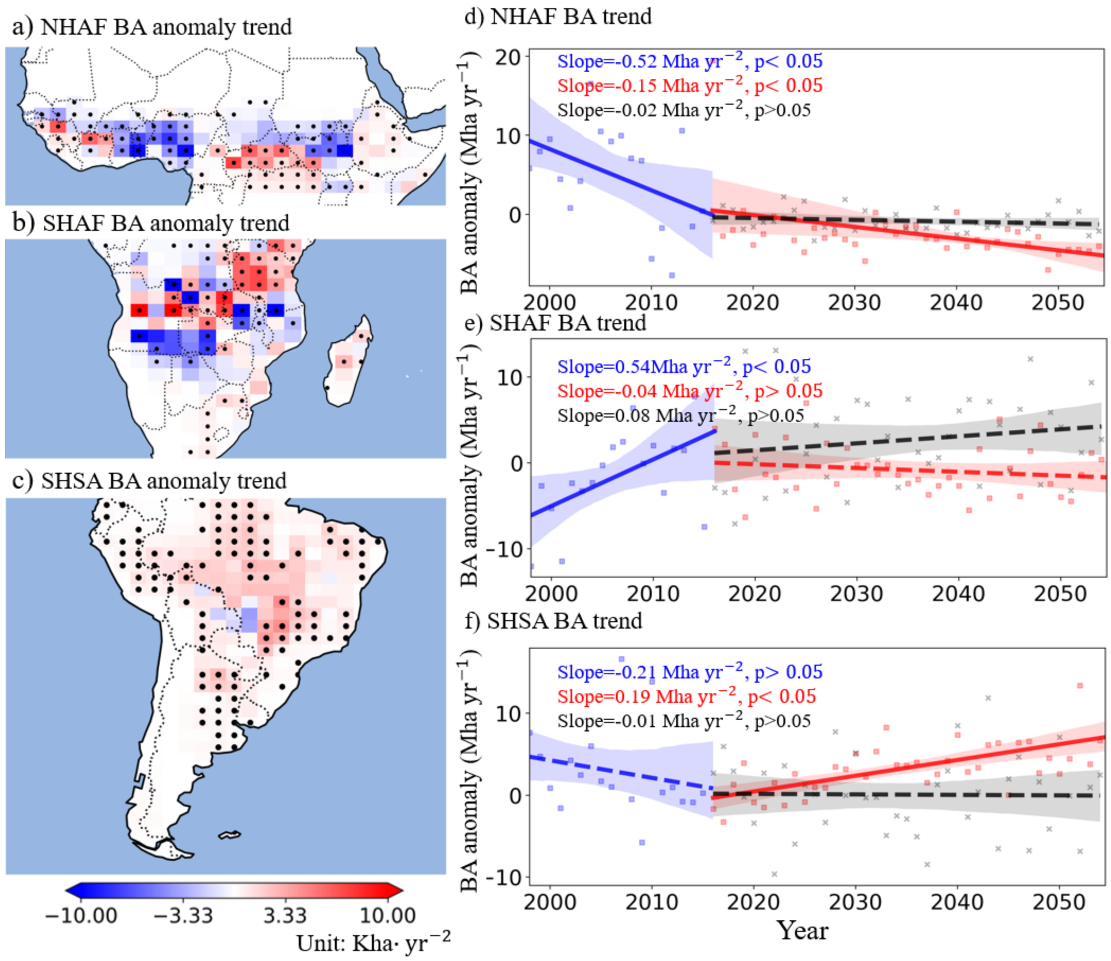

Fig. 6: Future burned area trends under the SSP585 high emission scenario. (a-c) spatial distribution of fire season burned area trends using drivers with interannual variations; dots in (a-c) indicate gridcells with statistically significant changes in the trend. (d-f) regionally aggregated burned area changes with historical mean subtracted. Blue and red lines respectively represent burned area anomaly in history and future; the black line represents future burned area trend while removing the interannual variations of the dominant variable. Solid lines represented significant BA trends (*p* value <0.05)

while dashed lines represented non-significant BA trends.

Considering land use changes, population growth, and projected climate and fuel conditions under the SSP585 high emission scenario, our model predicted that burned areas in the NHAF region will continue to decline; the currently increasing trend will be dampened in the SHAF region, and the currently decreasing trend will be reversed in SHSA region (Fig. 6). The increasing trend in SHSA, decreasing trend in NHAF, and dampened trend in SHAF under SSP585 were robust when projecting burned area till the end of 21st century (Fig. S5). Over NHAF and SHSA, burned area trends at the gridcell level were mostly robust (Fig. 6a, c; p <0.05) and of the same sign, thus resulting in a robust trend at regional scale. Under the low emission scenario (i.e., SSP126), the decreasing trend in NHAF disappeared (Fig. S5a) and the increasing trend in SHSA was reduced by ~69% (Fig. S5c), implying the big influences of climate changes and socioeconomic development pathways on future burn area changes in the two regions.

To investigate what drives future burned area changes under SSP585, we iteratively surrogated each driver with its climatology while keeping the other factors the same. Burned area changing trends in NHAF and SHSA were mostly affected by VPD changes because removing VPD inter-annual changes resulted in non-significant burned area trends at the whole NHAF and SHSA region (Fig. 6a, c). VPD was projected to continuously increase due to warming but had different implications over NHAF and SHSA. Over the relatively fuel abundant SHSA region, increased VPD will likely increase burned area (Pearson r= 0.64, p value <0.05, Fig. S6) through increasing fuel dryness and combustibility (Kelley et al., 2019; Chen et al., 2011; Malhi et al., 2008; Van Der Werf et al., 2008). In contrast, over the semi-arid savannah dominated NHAF region (less fuel, compared with SHSA), higher VPD could decrease burned area (Pearson r= -0.71, p value <0.05, Fig. S6) through limiting plant growth and fuel availability (Van Der Werf et al., 2008; Andela and Van Der Werf, 2014; Andela et al., 2017). For the SHAF, population growth and climate changes showed stronger influences on burned area changes (Andela and Van Der Werf, 2014) while the heterogeneity of wildfire responses finally led to a non-significant trend at the regional scale (Fig. 6). Our findings highlight the importance of climate changes on understanding future burned area dynamics, and motivate better representation of climate wetness effects on wildfire dynamics in process-based and machine learning-based wildfire prediction models.

3.5 Directions for future research

The time lagged controls of climate on ASA wildfires are critical for sub-seasonal to seasonal wildfire prediction (Chen et al., 2020; Andela and Van Der Werf, 2014; Chen et al., 2011) but remain less well represented due to the complex interactions among fire, climate, fuel, and human activities. Here we deployed the interpretable AttentionFire model to understand and predict fire dynamics in ASA region. We revealed the dominant, spatially heterogenous, and time lagged controls of climate wetness on ASA wildfires. Such climate wetness importance on ASA wildfires was consistent with previous findings (Andela and Van Der Werf, 2014; Chen et al., 2011) and also confirmed by the other three tree-based ML models (i.e., DT, RF, and GBDT) with variable importance (e.g., precipitation and VPD were regarded as the top-five most important variables in Fig. S7). However, differences existed across model identified most important drivers (Fig. 3 versus Fig. S7). The variable importance of AttentionFire model was spatiotemporally varied (Fig. 4) while tree-based model provided variable importance was constant over the entire dataset. We showed that the climate wetness was more (less) important in areas with large (small) burned areas and its importance also varied over time (Fig. 4), but the other MLs did not explicitly distinguish such differences. Albeit the higher accuracy and generally acceptable computation speed of AttentionFire (Table S2), its memory consumption and model training time could be up to 141% and 22 times higher than the other ML models. The implementation of LSTM in AttentionFire model is series instead of parallel, therefore, future work could improve the model efficiency by deploying some easy-for-parallel-computing time series prediction frameworks (e.g., temporal convolutional attention (Lin et al., 2021) and self-attention (Mohammadi Farsani and Pazouki, 2020; Vaswani et al., 2017)).

This study focused on wildfire prediction in ASA region and we showed the performance improvement of AttentionFire model by representing the time-lagged controls of climate on wildfires. Whether the AttentionFire model can also outperform other ML models in other regions may depend on the dependency strength and time lags between wildfires and climate variables. For example, in North American boreal forests, lightning was identified as the major driver of the interannual variability in burned area by influencing the number of ignitions in dry-season (Veraverbeke et al., 2017). In such region, AttentionFire model might not outperform other ML models due to the less dominance of time-lagged controls. In regions like western US and India where wildfires showed time-lagged dependencies with local climate (Littell et al., 2009; Kale et al., 2022) and some extreme wildfires were caused by persistent drought from wet to dry seasons with multi-month lags (Taufik et al., 2017; Littell et al., 2016), the AttentionFire model could be potentially useful.

With the fully coupled ESMs of CMIP6, we analyzed future burned area changes under high (SSP585) and low (SSP126) emission scenarios in the ASA region. While the MME mean was considered, substantial uncertainty has been found across different ESMs in history (Yuan et al., 2022a; Yuan et al., 2021; Wu et al., 2020) and future (Li et al., 2022; Lauer et al., 2020). Further work therefore is needed to narrow the projection uncertainty of ESMs (e.g., with constraints of causality (Nowack et al., 2020; Li et al., 2022) and observations (Tokarska et al., 2020; Lauer et al., 2020)). Meanwhile, for future projections, although land use and land cover changes, population growth, and climate and fuel changes were considered, constant livestock and road density were adopted due to lack of data. The impacts of livestock and road density therefore need further exploration with available data under different future scenarios. In addition, the AttentionFire model currently is not coupled with the ESM, therefore, the feedbacks among fires, climate, and biomass were ignored. To analyze such feedbacks, the AttentionFire model needs to surrogate the original fire module and be coupled within the ESM (Zhu et al., 2021).

4. Conclusions

This study developed an interpretable machine learning model (AttentionFire_v1.0) for burned areas predictions over African and south American regions. Compared with observations and other five widely used machine learning baseline models, we demonstrated the effectiveness of the AttentionFire model to capture the magnitude, spatial distribution, and temporal variation of burned areas. “Attention” mechanisms enabled the interpretation of complex but critical spatial-temporal patterns (Li et al., 2020; Guo et al., 2019; Liang et al., 2018; Vaswani et al., 2017; Qin et al., 2017), thus uncovering the black-boxed relationships in machine learning models for burned area predictions. We demonstrated the spatiotemporally heterogenous and strong time-lagged controls from local climate wetness on burned areas. Furthermore, under the SSP585 high emission scenario, our results suggested that the increasing trend in burned area over southern Africa will be dampened, and the declining trend in burned area over fuel abundant southern America will reverse. This study highlights the importance of skillful representation of spatiotemporally heterogenous and strong time-lagged climate wetness effects on understanding wildfire dynamics and developing advanced early fire warning models.

Acknowledgements:

This research was supported by the Director, Office of Science, Office of Biological and Environmental Research of the US Department of Energy under contract no. DEAC02-05CH11231 as part of their Regional and Global Climate Modeling program through the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computation Scientific Focus Area (RUBISCO SFA) project and as part of the Energy Exascale Earth System Model (E3SM) project.

Code availability

The source code of AttentionFire_v1.0 and all baseline machine learning models is archived at Zenodo repository: https://zenodo.org/record/7416437#.Y5JnBXbMK5c, under Creative Commons Attribution 4.0 International license, with four zip files: data, data_preparation, model, and example. The "data" file contains the links to all raw datasets used to drive the model (e.g., burned areas, climate forcing). The "data_preparation" file contains the code to preprocess the raw datasets and make them be ready for training and testing the AttentionFire model. The "model" file contains the python code of AttentionFire model. The "example" file gives a detailed example of how to use the AttentionFire model for burned area predictions.

There is also a tutorial file “Data_Model_Tutorial” that contain descriptions on (1) how to load the raw datasets; (2) how to prepare the input and output datasets for ML model; (3) how to initialize the ML model and run the model (4) how to train the ML model and use the trained ML model for predictions; (5) how to save and load the model parameters and save the predicted results.

Data availability

Burned area: Global Fire Emissions Database https://daac.ornl.gov/VEGETATION/guides/fire_emissions_v4.html

NCEP-DOE Reanalysis Climate forcings:

https://psl.noaa.gov/data/gridded/data.ncep.reanalysis2.html

Population: https://landscan.ornl.gov/

Road density: https://www.globio.info/download-grip-dataset

Livestock density: https://www.fao.org/dad-is/en/

Land cover change: https://luh.umd.edu/data.shtml

Oceanic index: https://psl.noaa.gov/data/climateindices/list/

Author contributions

QZ and FL designed the study. QZ, FL, and MC designed the model experiments. FL wrote the code and ran the experiments. LZ, WR, JR, LX, KY, HW, ZG, and JG all contributed to the interpretation of the results and writing of the paper.

References

Abatzoglou, J. T. and Kolden, C. A.: Relationships between climate and macroscale area burned in the western United States, International Journal of Wildland Fire, 22, 1003-1020, 2013.

Altmann, A., Toloşi, L., Sander, O., and Lengauer, T.: Permutation importance: a corrected feature importance measure, Bioinformatics, 26, 1340-1347, 2010.

Amatulli, G., Rodrigues, M. J., Trombetti, M., and Lovreglio, R.: Assessing long-term fire risk at local scale by means of decision tree technique, Journal of Geophysical Research: Biogeosciences, 111, 2006.

Andela, N. and Van Der Werf, G. R.: Recent trends in African fires driven by cropland expansion and El Niño to La Niña transition, Nature Climate Change, 4, 791-795, 2014.

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

Aragao, L. E. O., Malhi, Y., Barbier, N., Lima, A., Shimabukuro, Y., Anderson, L., and Saatchi, S.: Interactions between rainfall, deforestation and fires during recent years in the Brazilian Amazonia, Philosophical Transactions of the Royal Society B: Biological Sciences, 363, 1779-1785, 2008.

Archibald, S., Roy, D. P., van Wilgen, B. W., and Scholes, R. J.: What limits fire? An examination of drivers of burnt area in Southern Africa, Global Change Biology, 15, 613-630, 2009.

Benavides-Solorio, J. and MacDonald, L. H. J. H. P.: Post-fire runoff and erosion from simulated rainfall on small plots, Colorado Front Range, 15, 2931-2952, 2001.

Bolton, D.: The computation of equivalent potential temperature, Monthly weather review, 108, 1046-1053, 1980.

Bowman, D. M., Balch, J. K., Artaxo, P., Bond, W. J., Carlson, J. M., Cochrane, M. A., D'Antonio, C. M., DeFries, R. S., Doyle, J. C., and Harrison, S. P.: Fire in the Earth system, science, 324, 481-484, 2009.

Breiman, L.: Random forests, Machine learning, 45, 5-32, 2001.

Chen, Y., Morton, D. C., Andela, N., Giglio, L., and Randerson, J. T.: How much global burned area can be forecast on seasonal time scales using sea surface temperatures?, Environmental Research Letters, 11, 045001, 2016.

Chen, Y., Morton, D. C., Andela, N., Van Der Werf, G. R., Giglio, L., and Randerson, J. T.: A pan-tropical cascade of fire driven by El Niño/Southern Oscillation, Nature Climate Change, 7, 906-911, 2017.

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

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

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

Dai, A.: Increasing drought under global warming in observations and models, Nature climate change, 3, 52-58, 2013.

Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D., DuVivier, A., Edwards, J., Emmons, L., Fasullo, J., Garcia, R., and Gettelman, A.: The community earth system model version 2 (CESM2), Journal of Advances in Modeling Earth Systems, 12, 2020.

Dangol, S., Talchabhadel, R., and Pandey, V. P.: Performance evaluation and bias correction of gridded precipitation products over Arun River Basin in Nepal for hydrological applications, Theoretical and Applied Climatology, 148, 1353-1372, 2022.

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

Etminan, M., Myhre, G., Highwood, E., and Shine, K. J. G. R. L.: Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing, 43, 12,614-612,623, 2016.

Gale, M. G., Cary, G. J., Van Dijk, A. I., and Yebra, M.: Forest fire fuel through the lens of remote sensing: Review of approaches, challenges and future directions in the remote sensing of biotic determinants of fire behaviour, Remote Sensing of Environment, 255, 112282, 2021.

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

Gray, M. E., Zachmann, L. J., and Dickson, B. G.: A weekly, continually updated dataset of the probability of large wildfires across western US forests and woodlands, Earth System Science Data, 10, 1715-1727, 2018a.

Gray, M. E., Zachmann, L. J., and Dickson, B. G. J. E. S. S. D.: A weekly, continually updated dataset of the probability of large wildfires across western US forests and woodlands, 10, 1715-1727, 2018b.

Gui, Z., Sun, Y., Yang, L., Peng, D., Li, F., Wu, H., Guo, C., Guo, W., and Gong, J.: LSI-LSTM: An attention-aware LSTM for real-time driving destination prediction by considering location semantics and location importance of trajectory points, Neurocomputing, 440, 72-88, 2021.

Guo, T., Lin, T., and Antulov-Fantulin, N.: Exploring interpretable LSTM neural networks over multi-variable data, International Conference on Machine Learning, 2494-2504,

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

Hochreiter, S. and Schmidhuber, J.: Long short-term memory, Neural computation, 9, 1735-1780, 1997.

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

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

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., and Klein Goldewijk, K.: Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6, Geoscientific Model Development, 13, 5425-5464, 2020b.

Jabbar, H. and Khan, R. Z.: Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study), Computer Science, Communication and Instrumentation Devices, 70, 2015.

Jain, P., Coogan, S. C., Subramanian, S. G., Crowley, M., Taylor, S., and Flannigan, M. D.: A review of machine learning applications in wildfire science and management, Environmental Reviews, 28, 478-505, 2020.

Joshi, J. and Sukumar, R.: Improving prediction and assessment of global fires using multilayer neural networks, Scientific reports, 11, 1-14, 2021.

Kale, M. P., Mishra, A., Pardeshi, S., Ghosh, S., Pai, D., and Roy, P. S.: Forecasting wildfires in major forest types of India, Frontiers in Forests and Global Change, 5, 2022.

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

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y.: Lightgbm: A highly efficient gradient boosting decision tree, Advances in neural information processing systems, 30, 2017.

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

Knorr, W., Dentener, F., Lamarque, J.-F., Jiang, L., and Arneth, A.: Wildfire air pollution hazard during the 21st century, Atmospheric Chemistry and Physics, 17, 9223-9236, 2017.

Lauer, A., Eyring, V., Bellprat, O., Bock, L., Gier, B. K., Hunter, A., Lorenz, R., Pérez-Zanón, N., Righi, M., and Schlund, M.: Earth System Model Evaluation Tool (ESMValTool) v2. 0–diagnostics for emergent constraints and future projections from Earth system models in CMIP, Geoscientific Model Development, 13, 4205-4228, 2020.

Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., and Pozzer, A.: The contribution of outdoor air pollution sources to premature mortality on a global scale, Nature, 525, 367-371, 2015.

Leung, H. and Haykin, S.: The complex backpropagation algorithm, IEEE Transactions on signal processing, 39, 2101-2104, 1991.

Li, F., Zhu, Q., Riley, W. J., Yuan, K., Wu, H., and Gui, Z.: Wetter California projected by CMIP6 models with observational constraints under a high GHG emission scenario, Earth's Future, 10, e2022EF002694, 2022.

Li, F., Gui, Z., Wu, H., Gong, J., Wang, Y., Tian, S., and Zhang, J.: Big enterprise registration data imputation: Supporting spatiotemporal analysis of industries in China, Computers, Environment and Urban Systems, 70, 9-23, 2018.

Li, F., Gui, Z., Zhang, Z., Peng, D., Tian, S., Yuan, K., Sun, Y., Wu, H., Gong, J., and Lei, Y.: A hierarchical temporal attention-based LSTM encoder-decoder model for individual mobility prediction, Neurocomputing, 403, 153-166, 2020.

Liang, H., Zhang, M., and Wang, H.: A neural network model for wildfire scale prediction using meteorological factors, IEEE Access, 7, 176746-176755, 2019.

Liang, Y., Ke, S., Zhang, J., Yi, X., and Zheng, Y.: Geoman: Multi-level attention networks for geo-sensory time series prediction, IJCAI, 3428-3434,

Lin, Y., Koprinska, I., and Rana, M.: Temporal convolutional attention neural networks for time series forecasting, 2021 International Joint Conference on Neural Networks (IJCNN), 1-8,

Littell, J. S., McKenzie, D., Peterson, D. L., and Westerling, A. L.: Climate and wildfire area burned in western US ecoprovinces, 1916–2003, Ecological Applications, 19, 1003-1021, 2009.

Littell, J. S., Peterson, D. L., Riley, K. L., Liu, Y., and Luce, C. H.: A review of the relationships between drought and forest fire in the United States, Global change biology, 22, 2353-2369, 2016.

Lundberg, S. M. and Lee, S.-I.: A unified approach to interpreting model predictions, Advances in neural information processing systems, 30, 2017.

Ma, H., Yuan, K., Li, F., Leroy, C., and Bronevetsky, G.: Predicting climate conditions based on teleconnections, 2022.

Malhi, Y., Roberts, J. T., Betts, R. A., Killeen, T. J., Li, W., and Nobre, C. A.: Climate change, deforestation, and the fate of the Amazon, science, 319, 169-172, 2008.

Maraun, D.: Bias correcting climate change simulations-a critical review, Current Climate Change Reports, 2, 211-220, 2016.

Mei, Y. and Li, F.: Predictability comparison of three kinds of robbery crime events using LSTM, Proceedings of the 2019 2nd international conference on data storage and data engineering, 22-26,

Meijer, J. R., Huijbregts, M. A., Schotten, K. C., and Schipper, A. M.: Global patterns of current and future road infrastructure, Environmental Research Letters, 13, 064006, 2018.

Mohammadi Farsani, R. and Pazouki, E.: A transformer self-attention model for time series forecasting, Journal of Electrical and Computer Engineering Innovations (JECEI), 9, 1-10, 2020.

Molnar, C., Casalicchio, G., and Bischl, B.: Interpretable machine learning–a brief history, state-of-the-art and challenges, Joint European Conference on Machine Learning and Knowledge Discovery in Databases, 417-431,

Mueller, S. E., Thode, A. E., Margolis, E. Q., Yocom, L. L., Young, J. D., and Iniguez, J. M.: Climate relationships with increasing wildfire in the southwestern US from 1984 to 2015, Forest Ecology and Management, 460, 117861, 2020.

Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., and Yu, B.: Definitions, methods, and applications in interpretable machine learning, Proceedings of the National Academy of Sciences, 116, 22071-22080, 2019.

Natekar, S., Patil, S., Nair, A., and Roychowdhury, S.: Forest fire prediction using LSTM, 2021 2nd International Conference for Emerging Technology (INCET), 1-5,

NOAA: Climate indices, Retrieved 2021.05.04, from https://psl.noaa.gov/data/climateindices/list/, 2021.

Nowack, P., Runge, J., Eyring, V., and Haigh, J. D.: Causal networks for climate model evaluation and constrained projections, Nature communications, 11, 1-11, 2020.

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., and Lowe, J.: The scenario model intercomparison project (ScenarioMIP) for CMIP6, Geoscientific Model Development, 9, 3461-3482, 2016.

Pechony, O. and Shindell, D. T.: Driving forces of global wildfires over the past millennium and the forthcoming century, Proceedings of the National Academy of Sciences, 107, 19167-19170, 2010.

Qin, Y., Song, D., Chen, H., Cheng, W., Jiang, G., and Cottrell, G.: A dual-stage attention-based recurrent neural network for time series prediction, arXiv preprint arXiv:1704.02971, 2017.

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

Ramanathan, V., Crutzen, P., Kiehl, J., and Rosenfeld, D.: Aerosols, climate, and the hydrological cycle, science, 294, 2119-2124, 2001.

Ramos da Silva, R., Werth, D., and Avissar, R.: Regional impacts of future land-cover changes on the Amazon basin wet-season climate, Journal of climate, 21, 1153-1170, 2008.

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

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., and Carvalhais, N.: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195-204, 2019.

Rothman-Ostrow, P., Gilbert, W., and Rushton, J.: Tropical Livestock Units: Re-evaluating a Methodology, Frontiers in Veterinary Science, 7, 973, 2020.

Safavian, S. R. and Landgrebe, D.: A survey of decision tree classifier methodology, IEEE transactions on systems, man, and cybernetics, 21, 660-674, 1991.

Sedano, F. and Randerson, J.: Multi-scale influence of vapor pressure deficit on fire ignition and spread in boreal forest ecosystems, Biogeosciences, 11, 3739-3755, 2014.

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., and Kirkevåg, A.: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations, Geoscientific Model Development, 13, 6165-6200, 2020.

Shrestha, M., Acharya, S. C., and Shrestha, P. K.: Bias correction of climate models for hydrological modelling–are simple methods still useful?, Meteorological Applications, 24, 531-539, 2017.

Shvetsov, E. G., Kukavskaya, E. A., Buryak, L. V., and Barrett, K. J. E. R. L.: Assessment of post-fire vegetation recovery in Southern Siberia using remote sensing observations, 14, 055001, 2019.

Slack, D., Hilgard, A., Singh, S., and Lakkaraju, H.: Reliable post hoc explanations: Modeling uncertainty in explainability, Advances in Neural Information Processing Systems, 34, 9391-9404, 2021.

Taufik, M., Torfs, P. J., Uijlenhoet, R., Jones, P. D., Murdiyarso, D., and Van Lanen, H. A.: Amplification of wildfire area burnt by hydrological drought in the humid tropics, Nature Climate Change, 7, 428-431, 2017.

Teckentrup, L., Harrison, S. P., Hantson, S., Heil, A., Melton, J. R., Forrest, M., Li, F., Yue, C., Arneth, A., and Hickler, T.: Response of simulated burned area to historical changes in environmental and anthropogenic factors: a comparison of seven fire models, Biogeosciences, 16, 3883-3910, 2019.

Tokarska, K. B., Stolpe, M. B., Sippel, S., Fischer, E. M., Smith, C. J., Lehner, F., and Knutti, R.: Past warming trend constrains future warming in CMIP6 models, Science advances, 6, eaaz9549, 2020.

Turco, M., Jerez, S., Doblas-Reyes, F. J., AghaKouchak, A., Llasat, M. C., and Provenzale, A.: Skilful forecasting of global fire activity using seasonal climate predictions, Nature communications, 9, 1-9, 2018.

Van Der Werf, G. R., Randerson, J. T., Giglio, L., Gobron, N., and Dolman, A.: Climate controls on the variability of fires in the tropics and subtropics, Global Biogeochemical Cycles, 22, 2008.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I.: Attention is all you need, arXiv preprint arXiv:1706.03762, 2017.

Veraverbeke, S., Rogers, B. M., Goulden, M. L., Jandt, R. R., Miller, C. E., Wiggins, E. B., and Randerson, J. T.: Lightning as a major driver of recent large fire years in North American boreal forests, Nature Climate Change, 7, 529-534, 2017.

Wang, S. and Yuan, K.: Spatiotemporal analysis and prediction of crime events in atlanta using deep learning, 2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC), 346-350,

Wang, S. S. C., Qian, Y., Leung, L. R., and Zhang, Y.: Identifying key drivers of wildfires in the contiguous US using machine learning and game theory interpretation, Earth's future, 9, e2020EF001910, 2021a.

Wang, Y. C., Hsu, H. H., Chen, C. A., Tseng, W. L., Hsu, P. C., Lin, C. W., Chen, Y. L., Jiang, L. C., Lee, Y. C., and Liang, H. C.: Performance of the taiwan earth system model in simulating climate variability compared with observations and CMIP6 model simulations, Journal of Advances in Modeling Earth Systems, 13, e2020MS002353, 2021b.

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

Werf, G. R., Randerson, J. T., Giglio, L., Leeuwen, T. T. v., Chen, Y., Rogers, B. M., Mu, M., Van Marle, M. J., Morton, D. C., and Collatz, G. J. J. E. S. S. D.: Global fire emissions estimates during 1997-2016, 9, 697-720, 2017b.

Wu, G., Cai, X., Keenan, T. F., Li, S., Luo, X., Fisher, J. B., Cao, R., Li, F., Purdy, A. J., and Zhao, W.: Evaluating three evapotranspiration estimates from model of different complexity over China using the ILAMB benchmarking system, Journal of Hydrology, 590, 125553, 2020.

Xu, X., Jia, G., Zhang, X., Riley, W. J., and Xue, Y.: Climate regime shift and forest loss amplify fire in Amazonian forests, Global Change Biology, 26, 5874-5885, 2020.

Yu, Y., Mao, J., Thornton, P. E., Notaro, M., Wullschleger, S. D., Shi, X., Hoffman, F. M., and Wang, Y.: Quantifying the drivers and predictability of seasonal changes in African fire, Nature communications, 11, 1-8, 2020.

Yuan, K., Zhu, Q., Riley, W. J., Li, F., and Wu, H.: Understanding and reducing the uncertainties of land surface energy flux partitioning within CMIP6 land models, Agricultural and Forest Meteorology, 319, 108920, 2022a.

Yuan, K., Zhu, Q., Li, F., Riley, W. J., Torn, M., Chu, H., McNicol, G., Chen, M., Knox, S., and Delwiche, K.: Causality guided machine learning model on wetland CH4 emissions across global wetlands, Agricultural and Forest Meteorology, 324, 109115, 2022b.

Yuan, K., Zhu, Q., Zheng, S., Zhao, L., Chen, M., Riley, W. J., Cai, X., Ma, H., Li, F., and Wu, H.: Deforestation reshapes land-surface energy-flux partitioning, Environmental Research Letters, 16, 024014, 2021.

Zhou, W., Yang, D., Xie, S.-P., and Ma, J. J. N. C. C.: Amplified Madden–Julian oscillation impacts in the Pacific–North America region, Nature Climate Change, 10, 654-660, 2020.

Zhu, Q., Riley, W. J., Tang, J., Collier, N., Hoffman, F. M., Yang, X., and Bisht, G.: 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, 2019.

Zhu, Q., Li, F., Riley, W. J., Xu, L., Zhao, L., Yuan, K., Wu, H., Gong, J., and Randerson, J.: Building a machine learning surrogate model for wildfire activities within a global Earth system model, Geoscientific Model Development, 15, 1899-1911, 2022.

Zhu, Q., Li, F., Riley, W. J., Xu, L., Zhao, L., Yuan, K., Wu, H., Gong, J., and Randerson, J. T.: Building a machine learning surrogate model for wildfire activities within a global earth system model, Geoscientific Model Development Discussions, 1-22, 2021.

Ziehn, T., Chamberlain, M. A., Law, R. M., Lenton, A., Bodman, R. W., Dix, M., Stevens, L., Wang, Y.-P., and Srbinovsky, J.: The Australian Earth System Model: ACCESS-ESM1. 5, Journal of Southern Hemisphere Earth Systems Science, 70, 193-214, 2020.