AGU

PUBLICATIONS

Supporting Information template

"Fires of Unusual Size: Future of Extreme and Emerging Wildfires in a Warming United States (2020-2060)"

Jilmarie J. Stephenes[1]*, Maxwell Joseph[3], Virginia Iglesias[1], Ty Tuff[1,4], Adam Mahood[5], Imtiaz Rangwala[2], Jane Wolken[2], and Jennifer Balch[1,4,6]

[1] Earth Lab, Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado, Boulder, CO 80303.[2] North Central Climate Adaptation Science Center, University of Colorado, Boulder, CO 80303. [3] Natural Capital Exchange, 2443 Fillmore St. #380-1418, San Francisco, CA 94115.
[4] ESIIL, the Environmental Data Science Innovation & Inclusion Lab, University of Colorado, Boulder, CO 80303. [5] Water Resources, USDA-ARS, Fort Collins, CO.[6] Geography, University of Colorado, Boulder, CO 80303.

Contents of this file

Introduction

The following provides more details about the fire count and burn area models along with the results from the eight Global Climate Models that were averaged to get the ensemble results presented in the main text.

Text S1. Model Details

    a. Fire Occurrence Model

    The model represents counts as a zero-inflated negative binomial random variable. This approach allows us to simultaneously account for the zero-inflation and overdispersion observed in the fire count data. The model defines a probability mass function for fires over 405 ha (approximately 1000 acres) in each ecoregion *s* (spatial scale *s* = 1,....,*S*) and time step *t* (monthly scale t = 1,....,T). The location parameter $\mu_{s,t}$ and the structural zero inflation parameters $\pi_{s,t}$ were able to vary in space and time. A log link function ensured $\mu_{s,t} > 0$ while a logit link function ensured $\pi_{s,t} \in (0,1)$. Linking by spatial and temporal units so that $\pi = (\pi_{s=1,t=1}, \pi_{s=2,t=1}, \ldots, \pi_{s=S,t=1}, \pi_{s=S,t=2}, \ldots, \pi_{s=S,t=T})$ as well as for μ, the location and zero inflation parameters were modeled by

$$\log(\mu) = \alpha^{(\mu)} + X\beta^{(\mu)} + \varphi^{(\mu)} + \log(a)$$

$$\text{logit}(\pi) = \alpha^{(\pi)} + X\beta^{(\pi)} + \varphi^{(\pi)}$$

where $\alpha^{(\mu)}$ and $\alpha^{(\pi)}$ are scalar parameters of intercepts, X is a design matrix (S × T) × p where

p is the number of input features, β and φ are column vector parameters with $\beta^{(\mu)}$ and $\beta^{(\pi)}$ being length p and $\varphi^{(\mu)}$ and $\varphi^{(\pi)}$ of length S × T with spatiotemporal adjustments, and a is an areas offset vector for the spatial units s repeated for each time step t. A multivariate horseshoe was used sharing information between the zero inflated and negative binomial location parameters (Peltola et al., 2014).

b. Burned Area

The model response y_i is the number of hectares burned over 405 ha for the ith fire event occurring in each spatial unit s_i at time step t_i. The model included covariate dependence through the location parameter: $\mu_i = \alpha + X_{(s_i, t_i)}\beta + \phi_{(s_i, t_i)}$, where α is an intercept parameter, $X_{(s_i, t_i)}$ is a row vector from the design matrix X, β is a vector of coefficients of length p, and $\phi_{(s_i, t_i)}$ is an adjustment for the spatial unit s and time step t. For the lognormal burned area model a univariate horseshoe prior was used.

c. Accounting for nonlinear forcing

The design matrix was created to include the spatially nonlinear effects of meteorological variables and population density. To account for the nonlinearity and to allow the coefficients for each basis vector to vary spatially we used B-splines (Wood, 2017). The univariate B-splines for the meteorological drivers and population density each had five degrees of freedom, generating 30 basis vectors. Interaction effects were added between each basis vector and the ecoregions to account for the spatial variability in the nonlinear effects (Brezger and Lang, 2006; Kneib et al., 2009). The interactions between ecoregions is captured through the hierarchical nesting of the 3 ecoregions levels; such that coefficients in level 3 may be related to coefficients in the level 2 ecoregion that contains the level 3 ecoregion and so forth up to level 1 which contains the level 2 ecoregion and a global effect. The interactions effects for each of the 30 basis vectors for each ecoregion level were included to allow information sharing across ecoregion level and ecoregions of similar of ecology. An adjustment for the global intercept for each ecoregion level was included to account for any spatial variability that is not related to population density or climate. This leads to a matrix of 3,472 values, that includes many zero values, that will increase the efficiency of computing μ and π. The random effects in space and time were created "using a temporally autoregressive, spatially intrinsically autoregressive formulation (Besag and Kooperberg, 1995; Banerjee et al., 2014)" (Joseph et al., 2019).

d. Posterior predictive inference for finite sample maxima

Joseph et al. (2019) compared empirical maxima to the predicted distribution of maxima to ensure that models capture tail behavior. They also performed predictive checks for the proportion of zero counts and totals for count and burned area models. Posterior predictive inference for finite sample maxima obtains a "distribution over maxima by marginalizing over unknowns including the number of events, size of each event and the parameters of their distributions (Marani and Ignaccolo, 2015)" (Joseph et al., 2019).

S2. **Individual GCM predictions**

Out of all the models the IPSL and MRI predicted significantly fewer fires per year and burned area per year than the others, with the Had_ES model predicting approximately 2.5 times more fires and burned area than the IPSL model.

Model	Burned Area Per Year	Number of Fires Per Year
CanESM2	5.66M (min:3.47M; max:13.58M)	571 (min:393; max: 898)
CNRM_CM5	5.96M (min:3.79M; max:10.39M)	558 (min:419; max: 851)
CSIRO	6.23M (min:4.21M; max: 10.64M)	605 (min:432; max: 990)
Had_CC	6.53M (min: 4.39M; max: 9.82M)	661 (min:495; max: 952)
Had_ES	7.71M (min: 4.58M; max: 13.89M)	738 (min:527; max: 1116)
IPSL_MR	2.58M (min:1.64M; max: 7.25M)	300 (min:221; max: 749)
MIROC	5.91M (min:3.92M; max 9.25M)	608 (min:457; max: 849)
MRI	4.12M (min: 2.93M; max: 13.72M)	414 (min:314; max: 607)

Table S1. Median predicted burned area per year (acres) and number of fires per year from 2020-2060 for the Contiguous U.S. along with the minimum and maximum from the 2000 iterations run per model.
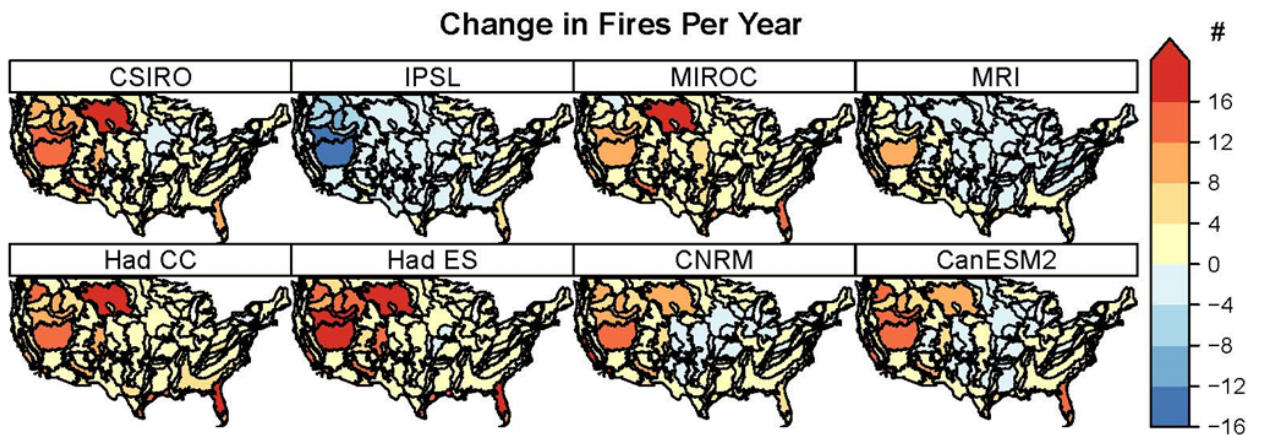
Figure S1. Change in the number of fires per year per ecoregion comparing predicted 2020-2060 vs. modeled 1990-2019 values for each of the eight GCMS.
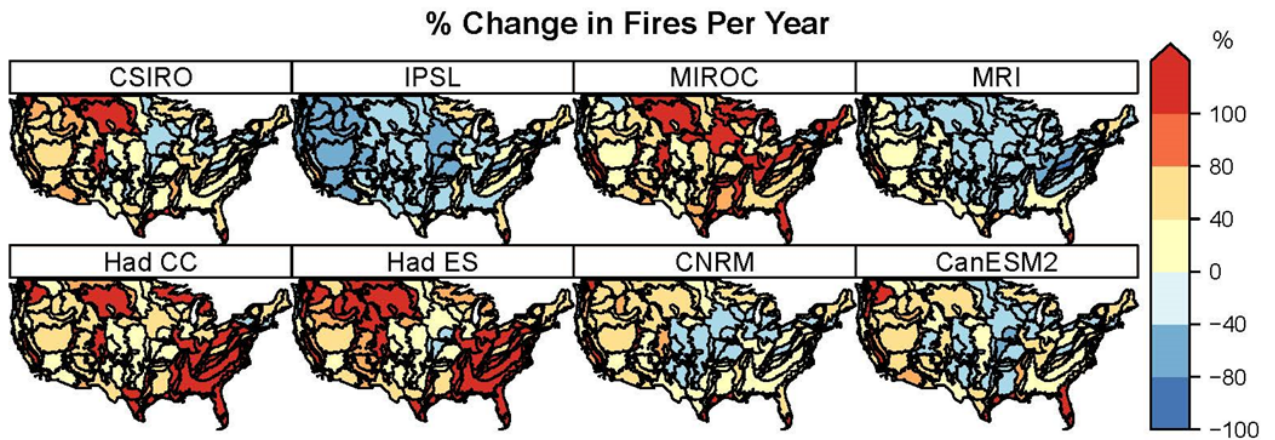
% Change in Fires Per Year

Figure S2. Percent change in the number of fires per year per ecoregion predicted 2020-2060 vs. modeled 1990-2019 for each of the eight GCMs.
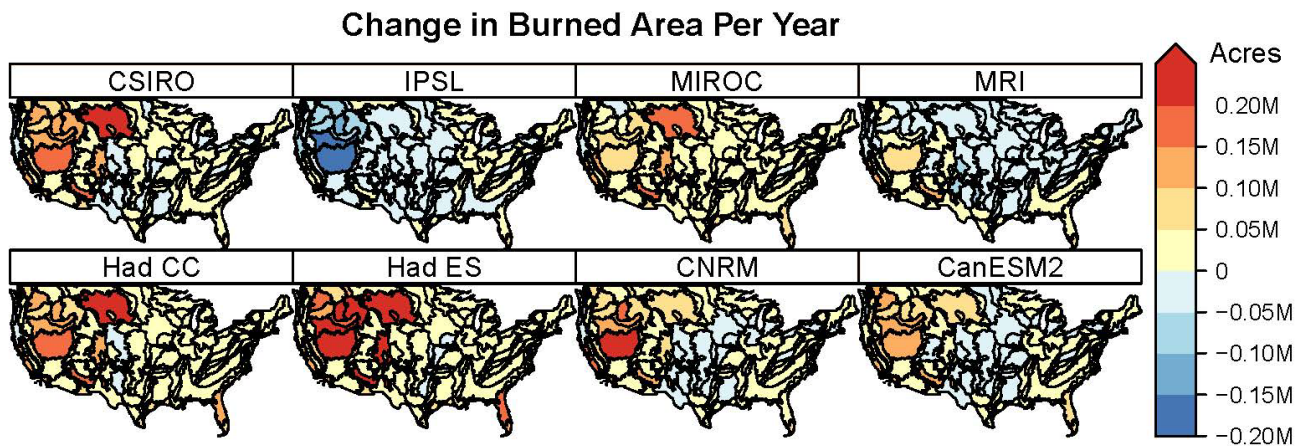
Change in Burned Area Per Year

Figure S3. Change in Burned Area per year per ecoregion comparing predicted 2020-2060 vs. modeled 1990-2019 values for each of the eight GCMS.
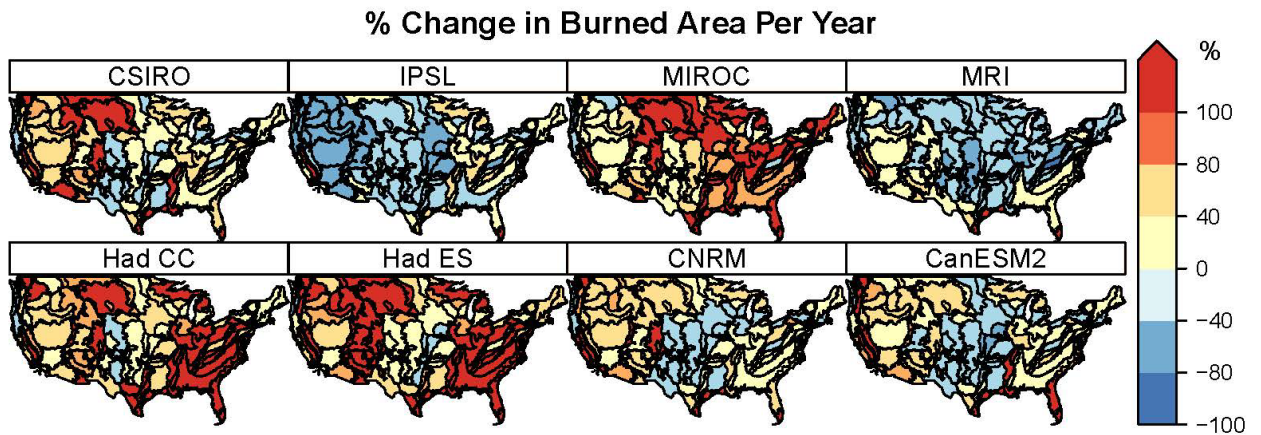
% Change in Burned Area Per Year

Figure S4. Percent change in Burned Area per year per ecoregion comparing predicted 2020-2060 vs. modeled 1990-2019 values for each of the eight GCMS.
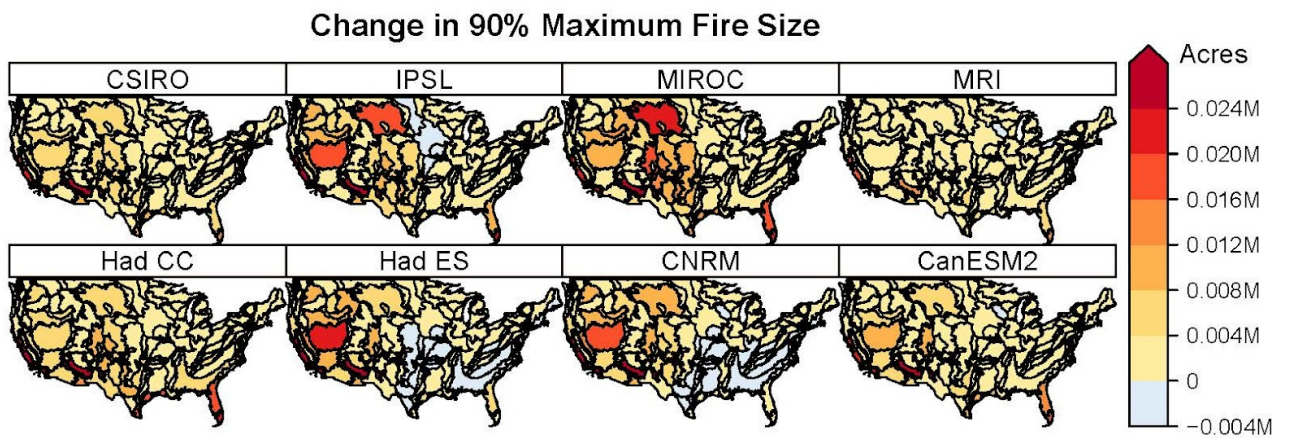
Change in 90% Maximum Fire Size

Figure S5. Change in 90% Maximum Fire Size per ecoregion comparing the predicted 2020-2060 vs. modeled 1990-2019 values for each of the eight GCMS.

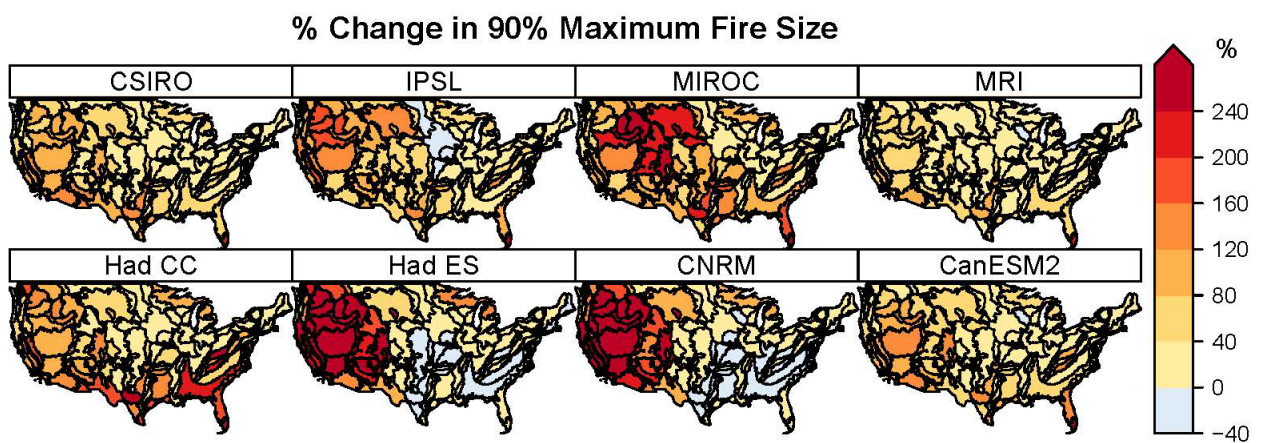## % Change in 90% Maximum Fire Size

Figure S6. Percent Change in 90% Maximum Fire Size per ecoregion comparing the predicted 2020-2060 vs. modeled 1990-2019 values for each of the eight GCMS.