

Supplementary Materials for

“The season for large fires in Southern California is projected to lengthen in a changing climate”

Chunyu Dong^{1,2}, A. Park Williams^{2,3}, John T. Abatzoglou⁴, Kairong Lin¹, Gregory S. Okin², Thomas W. Gillespie², Di Long⁵, Yen-Heng Lin⁶, Alex Hall⁶, Glen M. MacDonald^{2*}

¹School of Civil Engineering, Sun Yat-sen University, and Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, China

²Department of Geography, University of California, Los Angeles, Los Angeles, CA, USA

³Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA

⁴Management of Complex Systems Department, University of California, Merced, Merced, CA, USA

⁵State Key Laboratory of Hydrosience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing, China

⁶Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, Los Angeles, CA, USA

*Corresponding author: Glen M. MacDonald

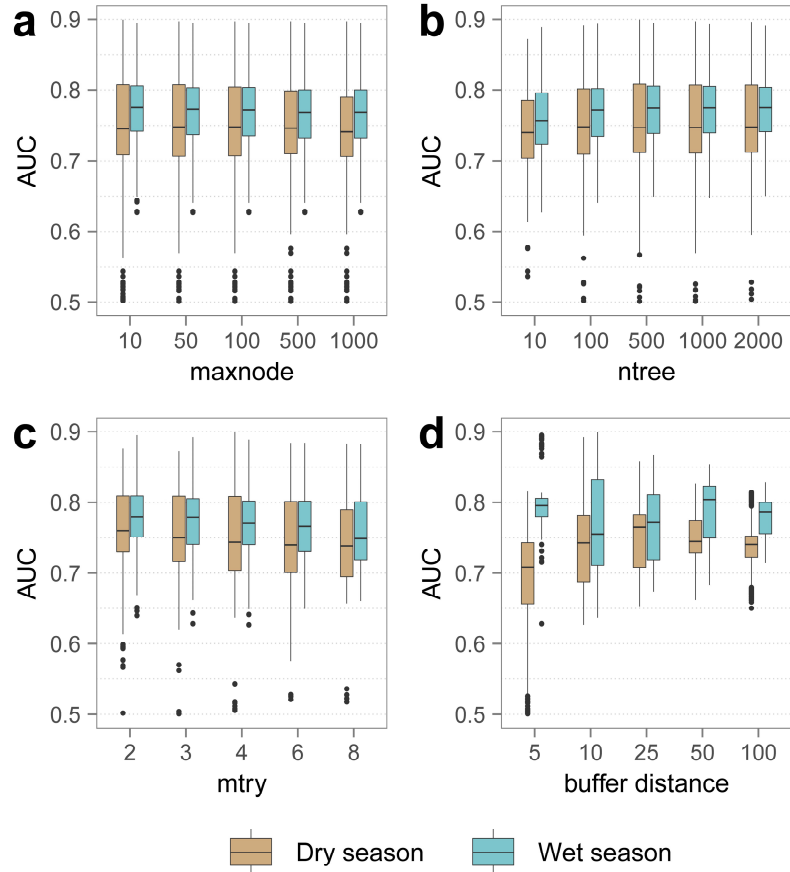
Email: glen@geog.ucla.edu

ORCID: 0000-0002-0267-9825

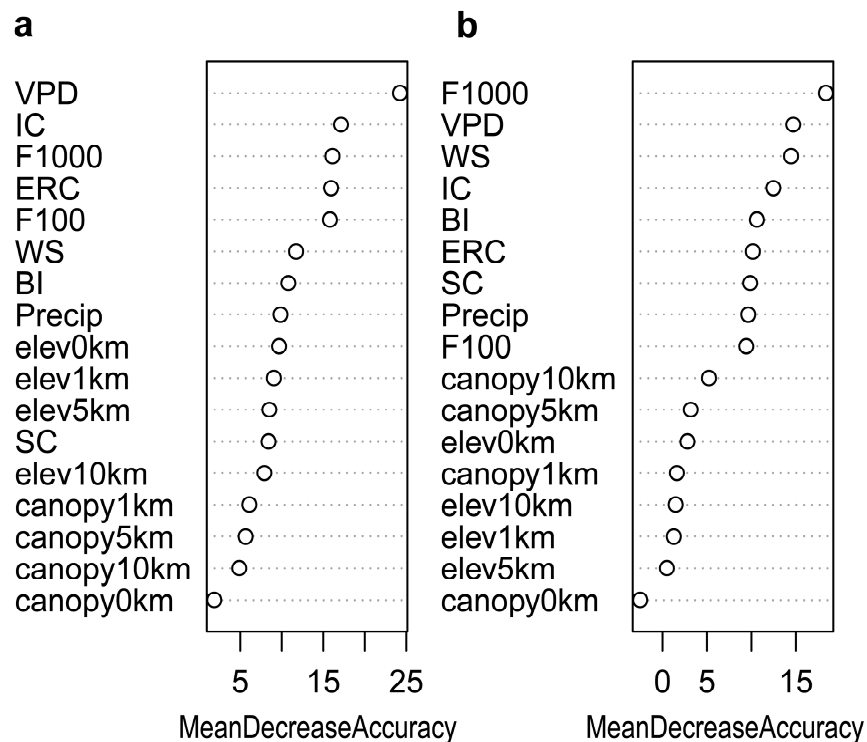
This PDF file includes:

Supplementary Fig. 1–6

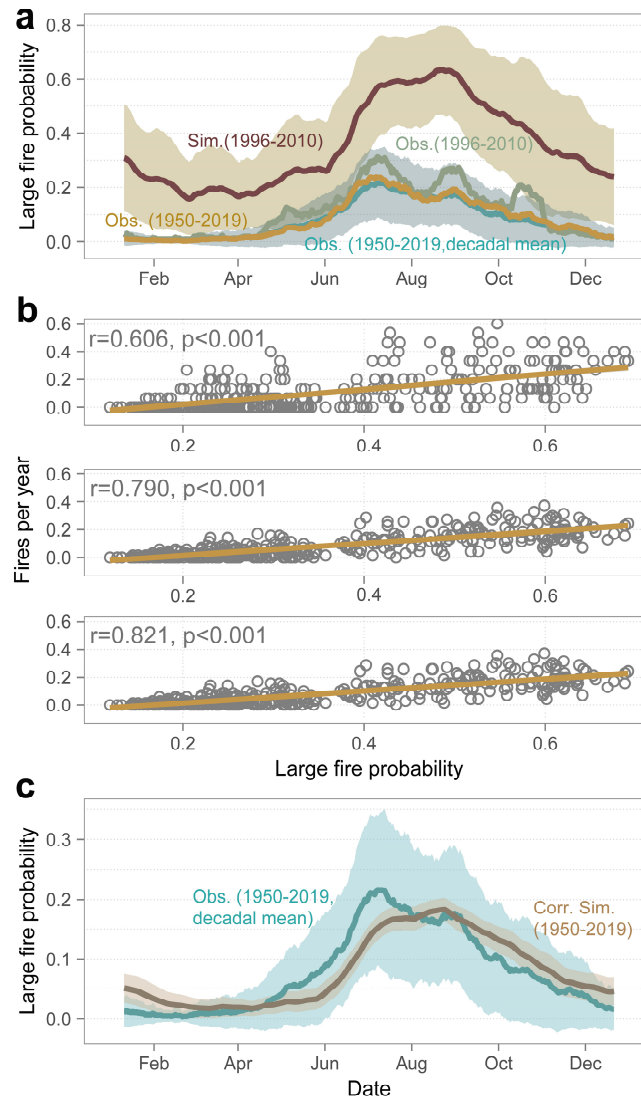
Supplementary Table 1–4



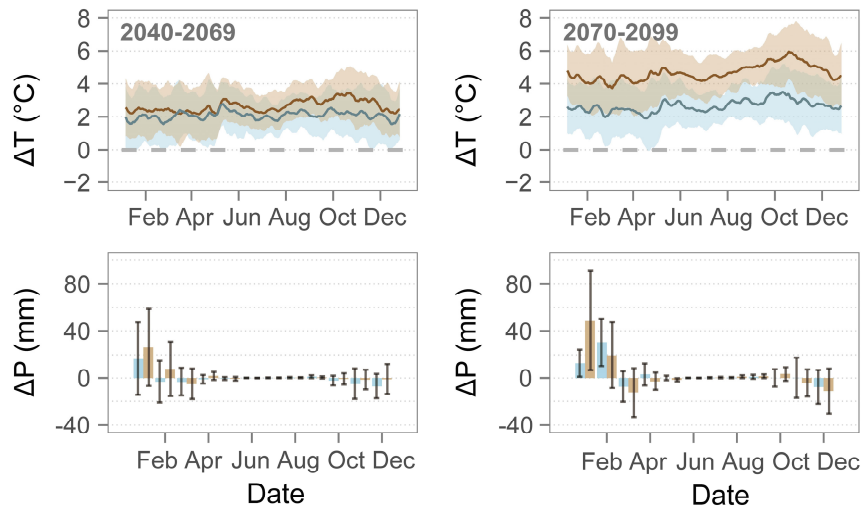
Supplementary Fig. 1 Variations in accuracy for the random forest modelling in large fire probability for different model parameters and station buffer distances. (a) *maxnode*, (b) *ntree*, (c) *mtry*, (d) station *buffer distances*. *maxnode* refers to the maximum number of terminal node trees that the forest can have. *ntree* indicates the number of trees to grow. *mtry* refers to the number of variables randomly sampled as candidates at each split. A higher AUC (area under the ROC curve) represents higher accuracy of the model, and an AUC of 1.0 indicates optimal modelling with 100% correct predictions. The detailed meaning of the *buffer distance* is provided in Methods. Meaning of boxplot elements: central line: median, box limits: upper and lower quartiles, upper whisker: $\min(\max(x), Q3 + 1.5 \times IQR)$, lower whisker: $\max(\min(x), Q1 - 1.5 \times IQR)$, black dots: outliers.



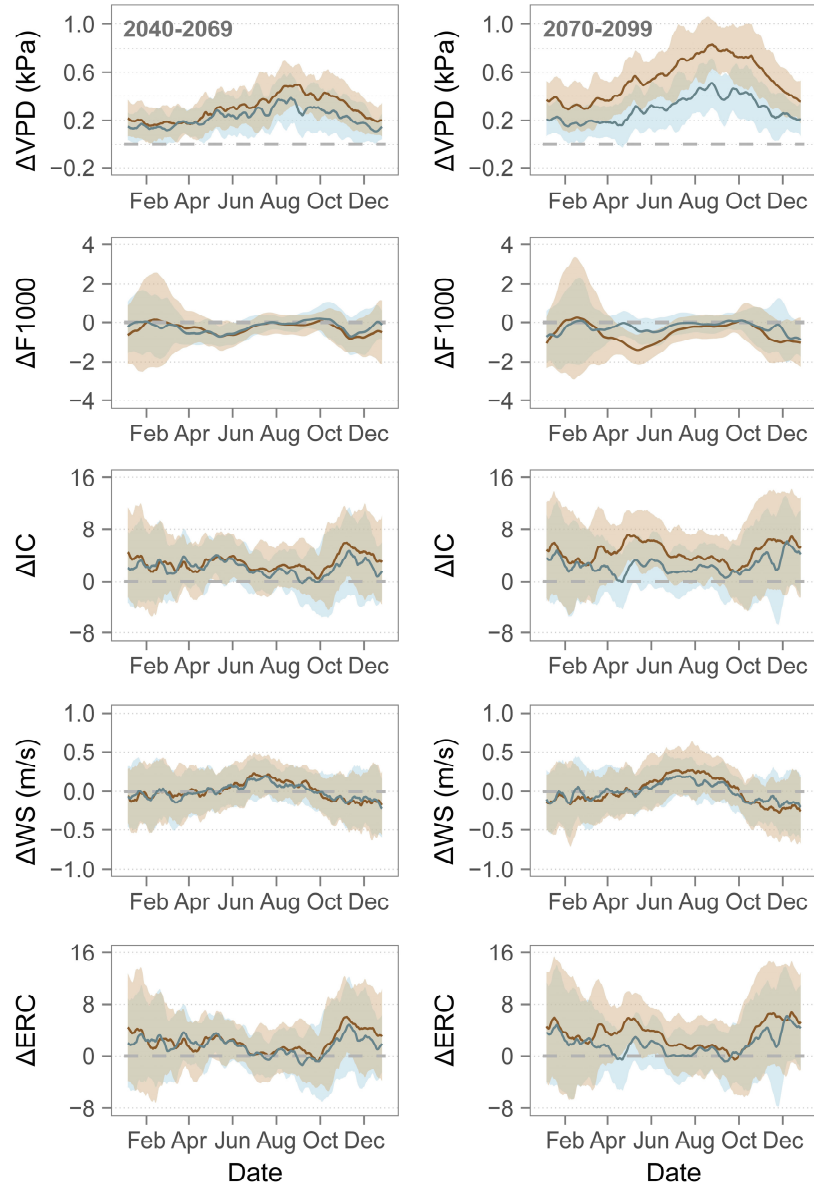
Supplementary Fig. 2 Relative importance of the random forest predictors for the (left) dry and (right) wet seasons based on the mean decrease in accuracy. A higher decrease in model accuracy indicates greater importance of the variable in the random forest model.



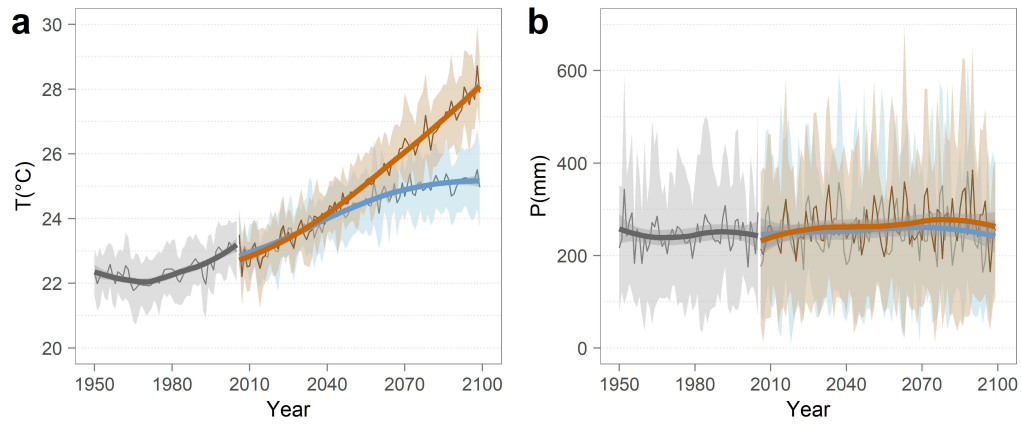
Supplementary Fig. 3 (a) Simulated mean daily LFP (dark brown) for the CSCA based on RAWs data between 1996–2010 and observed mean fire frequency for 1996–2010 (celadon), 1950–2019 (golden), and decadal mean during 1950–2019 (cyan). (b) Linear relations between simulated LFP and observed fire frequency for 1996–2010 (1st row), 1950–2019 (2nd row), and decadal mean of 1950–2019 (3rd row). (c) Bias-corrected simulations of daily LFP (dark brown) based on CMIP5 data (historical period: 1950–2005, RCP4.5 scenario: 2006–2019) and the observed seasonal changes in fire frequency during 1950–2019 (cyan). Shadings indicate intervals of mean ± 1 standard deviation.



Supplementary Fig. 4 ESM ensemble-mean changes in simulated temperature (top) and precipitation (bottom) by the middle (top, 2040–2069) and end (bottom, 2070–2099) of the 21st century compared with the historical period (1970–1999) for the RCP4.5 (blue) and RCP8.5 (tan) scenarios. Temperature series refer to daily averages, and precipitation data are monthly totals. Whiskers and shaded areas represent ± 1 standard deviation. A low-pass filter was applied to remove the highest 20% frequencies to reduce noise in the temperature time series. Twenty-one-day moving averages for temperature are shown here.



Supplementary Fig. 5 ESM ensemble-mean changes in simulated VPD (top), F1000 (2nd row), IC (3rd row), wind speed (WS, 4th row), and ERC (bottom) by the middle (left, 2040–2069) and end (right, 2070–2099) of the 21st century compared with the historical period (1970–1999) for the RCP4.5 (blue) and RCP8.5 (tan) scenarios. Shaded areas represent ± 1 standard deviation. A low-pass filter was applied to remove the highest 20% frequencies to reduce data noise. The 21-day moving averages of the time series are shown here.



Supplementary Fig. 6 ESM ensemble means of simulated temperature (left) and precipitation (right) for the historical (1950–2005, grey) and future (2006–2099, blue: RCP4.5, tan: RCP8.5) periods. Annual averages of temperature and total precipitation for January–March are shown. Shaded areas represent ± 1 standard deviation. A low-pass filter was applied to remove the highest 20% frequencies to reduce noise in the time series. Bold lines (grey: historical, blue: RCP4.5, orange: RCP8.5) refer to locally weighted smooth series with the 95% confidence intervals indicated as shadings.

Supplementary Table 1 Stations that have RAWs observations for 1996–2010 and downscaled CMIP5 climate projections for both the historical period (1950–2005) and future scenarios (2006–2099, RCP4.5 and RCP8.5).

Name	NWSID	County	Lat. (N)	Lon. (W)	Elev. (m.s.l.)
Acton	45438	Los Angeles	34.45	118.20	792
Alpine	45701	San Diego	32.83	-116.74	622
Anza	45616	Riverside	33.56	-116.67	1195
Arroyo Grande	44915	San Luis Obispo	35.19	-120.43	183
Bell Canyon	45735	Orange	33.54	-117.59	213
Beverly Hills	45442	Los Angeles	34.13	-118.41	384
Branch Mtn	44901	San Luis Obispo	35.18	-120.08	1150
Burns Canyon	45125	San Bernardino	34.21	-116.63	1829
Cameron FS	45704	San Diego	32.72	-116.46	1049
Camp 9	45441	Los Angeles	34.36	-118.42	1219
Carrizo	44916	San Luis Obispo	35.10	-119.77	759
Case Springs	45731	San Diego	33.45	-117.42	707
Cheeseboro	45313	Ventura	34.18	-118.72	503
Chilao	45436	Los Angeles	34.33	-118.03	1661
Chuchupate	45302	Ventura	34.81	-119.01	1494
Claremont	45443	Los Angeles	34.14	-117.71	501
Converse	45105	San Bernardino	34.19	-116.91	1712
Corona	45618	Riverside	33.87	-117.55	195
Del Valle	45445	Los Angeles	34.43	-118.68	390
Descanso	45707	San Diego	32.86	-116.62	1061
El Cariso	45619	Riverside	33.65	-117.41	926
Fawnskin	45101	San Bernardino	34.27	-116.90	2103
Fremont Canyon	45736	Orange	33.81	-117.71	543
Goose Valley	45724	San Diego	33.07	-116.85	466
Julian	45708	San Diego	33.08	-116.59	1292
Keenwild	45604	Riverside	33.67	-116.77	1500
Leo Carrillo	45447	Los Angeles	34.05	-118.94	15
Los Prietos	45203	Santa Barbara	34.54	-119.78	311
Lost Horse	45614	Riverside	34.02	-116.19	1280
Las Flores	45733	Orange	33.29	-117.45	1756
Malibu Hills	45433	Los Angeles	34.06	-118.63	480
Mill Creek	45435	Los Angeles	34.38	-118.08	1530
Montecito	45218	Santa Barbara	34.46	-119.65	457
Mormon Rocks	45114	San Bernardino	34.32	-117.50	1006
Mt Laguna	45709	Mt Laguna	32.88	-116.42	1756
Oak Grove	45710	San Diego	33.39	-116.80	839
Ojai	45315	Ventura	34.45	-119.23	233
Pine Hills	45711	San Diego	33.02	-116.63	1097

Poppy Park	45440	Los Angeles	34.73	-118.38	841
Ranchita	45729	San Diego	33.21	-116.51	1274
Rose Valley	45314	Ventura	34.54	-119.18	1015
Santa Fe Dam	45437	Los Angeles	34.12	-117.95	152
Santa Rosa Plateau	45623	Riverside	33.53	-117.23	604
Saugus	45412	Los Angeles	34.43	-118.53	442
Tanbark	45421	Los Angeles	34.17	-117.77	792
Target Range	45732	San Diego	33.37	-117.36	280
Valley Center	45734	San Diego	33.23	-116.99	418
Warm Springs	45426	Los Angeles	34.58	-118.55	1503
Whittier Hills	45446	Los Angeles	33.98	-118.01	290

Supplementary Table 2 Models from the Coupled Model Intercomparison Project (CMIP5) included in this study. All models were used for both the historical (1950–2005) and future analyses (2006–2099, RCP4.5 and RCP8.5).

MODEL	Modelling Group	Group Acroynm
bcc-csm1-1	Beijing Climate Center, China Meteorological Administration College of Global Change and Earth System Science, Beijing	BCC
BNU-ESM	Normal University	GCESS
CanESM2	Canadian Centre for Climate Modelling and Analysis Centre National de Recherches Météorologiques and Centre Européen de Recherche et Formation Avancée en Calcul	CCCMA
CNRM-CM5	Scientifique	CNRM- CERFACS
CSIRO- Mk3.6.0	Commonwealth Scientific and Industrial Research Organization and Queensland Climate Change Centre of Excellence	CSIRO- QCCCE
GFDL- ESM2G	NOAA Geophysical Fluid Dynamics Laboratory	NOAA GFDL
GFDL- ESM2M	NOAA Geophysical Fluid Dynamics Laboratory	NOAA GFDL
HadGEM2- CC	Met Office Hadley Centre	MOHC
HadGEM2-ES	Met Office Hadley Centre	MOHC
inmcm4	Institute for Numerical Mathematics Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute, and National Institute for Environmental Studies	INM
MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute, and National Institute for Environmental Studies	MIROC
MIROC-ESM- CHEM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute, and National Institute for Environmental Studies	MIROC
MIROC5	Meteorological Research Institute	MIROC
MRI-CGCM3		MRI

Supplementary Table 3 Number of fires captured by the RAWS station buffer zones for dry and wet seasons in CSCA between 1996–2010. There are 579 large fires (> 40 ha) in total.

Buffer distance (km)	Number of fires (Dry season)	Number of fires (Wet season)	Percent against all fires
5	51	35	14.9%
10	145	60	35.4%
25	368	110	82.6%
50	437	119	96.0%
100	455	124	100.0%

Supplementary Table 4 Best combinations of buffer distance and model parameters for the dry and wet seasons, and the respective evaluation metrics for the best models.

	Parameter/Index	Dry season	Wet season
Buffer distance		10 km	10 km
	maxnode	500	100
Model parameters	mtry	2	2
	ntree	5,00	500
	recall	0.96	0.90
	precision	0.03	0.003
Evaluation indices	FPrate	0.18	0.16
	accuracy	0.82	0.84
	AUC	0.89	0.87