Supplementary Information

Quantitative assessment of fire and vegetation properties in historical simulations with fire-enabled vegetation models from the Fire Model Intercomparison Project

Stijn Hantson[1,2], Douglas I. Kelley[3], Almut Arneth[1], Sandy P. Harrison[4], Sally Archibald[5], Dominique Bachelet[6], Matthew Forrest[7], Thomas Hickler[7,8], Gitta Lasslop[7], Fang Li[9], Stephane Mangeon[10,a], Joe R. Melton[11], Lars Nieradzik[12], Sam S. Rabin[1], I. Colin Prentice[13], Tim Sheehan[6], Stephen Sitch[14], Lina Teckentrup[15,16], Apostolos Voulgarakis[10], Chao Yue[17].

[1] Karlsruhe Institute of Technology, Institute of Meteorology and Climate research, Atmospheric Environmental Research, Garmisch-Partenkirchen, Germany.
[2] Geospatial Data Solutions Center, University of California Irvine, CA 92697 Irvine, USA.
[3] Centre for Ecology and Hydrology, Wallingford OX10 8BB, UK.
[4] School of Archaeology, Geography and Environmental Sciences, University of Reading, Reading, UK.
[5] Centre for African Ecology, School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Private Bag X3, WITS, Johannesburg, 2050, South Africa.
[6] Biological and Ecological Engineering, Oregon State University, Corvallis, OR 97331, USA.
[7] Senckenberg Biodiversity and Climate Research Institute (BiK-F), Senckenberganlage 25, 60325 Frankfurt am Main, Germany.
[8] Institute of Physical Geography, Goethe-University, Altenhöferallee 1, 60438 Frankfurt am Main, Germany.
[9] International Center for Climate and Environmental Sciences, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China.
[10] Department of Physics, Imperial College London, London, UK.
[11] Climate Research Division, Environment and Climate Change Canada, Victoria, BC V8W 2Y2, Canada
[12] Institute for Physical Geography and Ecosystem Sciences, Lund University, 22362 Lund, Sweden.
[13] AXA Chair of Biosphere and Climate Impacts, Grand Challenges in Ecosystem and the Environment, Department of Life Sciences and Grantham Institute – Climate Change and the Environment, Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot SL5 7PY, UK.
[14] College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4RJ, UK.
[15] ARC Centre of Excellence for Climate Extremes, University of New South Wales, Sydney, NSW, Australia.
[16] Climate Change Research Center, University of New South Wales, Sydney, NSW 2052, Australia.
[17] Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, 91198 Gif-sur-Yvette, France.
[a] now at: Data 61, CSIRO, Brisbane, Australia

Correspondence to: Stijn Hantson (stijn.hantson@gmail.com)

Supplementary Information S1: Overview of the benchmarking reference datasets used

Burnt Area

Five global burnt fraction products were used in this study (Figure S1). We used the fourth version of the Global Fire Emissions Database (GFED4) for 1997-2013, which uses the MCD64 burnt area MODIS based product in combination with an empirical estimation of burnt area based on thermal anomalies when MODIS data was unavailable (Giglio et al., 2013). We also included a version where the MCD64 burnt area product was merged with the small fire detection approach developed by Randerson et al. (2012; GFED4s). The third dataset is the MODIS burnt area product MCD45 which is the only burnt area product not using MODIS thermal anomalies within its burnt area detection algorithm (2002-2013) (Roy et al., 2008). The fourth is the FireCCIv4.0 dataset based on MERIS satellite data (Alonso-Canas and Chuvieco, 2015), available for the period 2005-2011. The fifth is the FireCCI5.1 dataset based on MODIS 250m imagery (Chuvieco et al., 2018).

Fire emissions

Carbon emission by fires is estimated within the Global Fire Assimilation System (GFAS) based on satellite-retrieved fire radiative power (FRP) (Kaiser et al., 2012). Here we use the global GFAS data for the period 2000-2013.

Fire size and numbers

Estimates on mean size and number of fires can be produced using a floodfilling algorithm to extract individual fires (Archibald et al., 2013). Here we use the data as produced by Hantson et al. (2015) from the MCD45 global burnt area product (Roy et al., 2008). Only large fires ≥25ha (one MODIS pixel) are detected, with a considerable underestimation of fires < ~125ha. Therefore, a direct comparison with modelled fire numbers and size is meaningless, but evaluation of the spatial pattern in fire numbers and fire size can be performed.

Vegetation productivity

We use multiple datasets for vegetation productivity, both measurements from site locations and global upscaled estimates. The site-level GPP dataset is from Luyssaert et al (2007) and the site-level NPP combines these data with data from the Ecosystem Model/Data Intercomparison (EMDI; (Olson et al., 2001)) databases. Sites from managed or disturbed environments were not used. A recent compilation of NPP site-level estimates was compiled by Michaletz et al. (2014). The mean of observations was taken when more than 1 measurement was available within a 0.5º grid cell. We also use upscaled fluxnet GPP data (Jung et al., 2017; Tramontana et al., 2017). Kelley et al. (2013) showed that the spreading of data between fluxnet site observations in such upscaling artificially improved model performance, probably because it used similar input data and using methods which might emulate functional relationships used within DGVMs. Hence scores obtained by Jung should not be interpreted as true "benchmarking scores" but could help inform differences between models in relation to scores obtained from other comparisons like burnt area (See Figure S1).

Carbon in vegetation

A global dataset on aboveground vegetation biomass was recently produced by combining two existing datasets—Saatchi et al. (2011) and Bacchini et al. (2012)—using a reference dataset of field observations and estimates (Avitabile et al., 2016). However, this dataset only considers woody biomass and to be able to analyse vegetation carbon also for areas without tree cover we used the dataset generated by Carvalhais et al. (2003) whom combined the Saatchi et al. (2011) and Thurner et al. (2004) biomass datasets while providing a best estimate for herbaceous biomass.

Leaf Area Index (LAI)

We use the MODIS LAI product MCD15 which gives global LAI values each 8 days (Myneni et al., 2002) and the LAI dataset produced based on AVHRR (Claverie et al., 2016). The mean LAI over the period 2001-2013 is used for benchmarking.

Supplementary Information S2: Benchmarking metrics

Annual average burnt area is assessed using the Normalised Mean Error (NME) metric, which sums the difference between observations (obs) and simulation (sim) over all cells (i) weighted by cell area (A_i) and normalizes by the average distance from the mean of observations ($\overline{obs}$):

$$NME = \frac{\sum A_i|obs_i - sim_i|}{\sum A_i|obs_i - \overline{obs}|} \quad (1)$$

NME comparisons are conducted in three steps:

Step 1. As described above;

Step 2. obs_i and sim_i are the difference between observation or simulation and their respective means. ie $x_i \rightarrow x_i - \bar{x}$, removing systematic bias and describe the performance of the model around the mean.

Step 3. obs_i and sim_i from step 2 are divided by the mean deviation. i.e $x_i \rightarrow x_i/|x_i|$. This removes the influence of the variability and describes the models ability to reproduce the spatial pattern in burnt area.

Only results of Step 3 are used throughout the manuscript (see methods), but results from all steps are given in Supplementary table S1.

Seasonality comparisons are conducted in two parts: seasonal concentration (inverse of season length) and phase, or timing, of the season. Each month, m, can be represented by a vector in the complex plane whose direction (θ_m) corresponds to the time of year and length to the magnitude of the variable for that month:

$$\theta_m = 2 \cdot \pi \cdot (m-1)/12 \quad (4)$$

A mean vector L can be calculated by averaging the real (L_x) and imaginary (L_y) parts of the 12 vectors (x_m).

$$L_x = \Sigma_m x_m \cdot \cos(\theta_m) \quad (5)$$
$$L_y = \Sigma_m x_m \cdot \sin(\theta_m)$$

The mean vector length by the annual average describes the seasonal concentration (C) of burnt area, whilst it's direction (P) describes seasonal timing (phase):

$$C = \frac{\sqrt{L_x^2 + L_y^2}}{\Sigma_m x_m} \quad (6)$$

$$P = \arctan\left(\frac{L_x}{L_y}\right) \quad (7)$$

If burnt area in a given cell is concentrated all in one month, C is equal to 1 and P corresponds to that month. If burnt area is evenly spread throughout the year then concentration is zero and phase is undefined. If the phase of a cell is undefined in either observations or simulation, then it is not used in the comparison. Likewise, if a cell has zero annual average burnt area for either observations or simulation, then that cell is not included in the comparisons. Concentration is compared using NME step 1. Phase are compared using mean phase difference (MPD)

$$MPD = \frac{1}{\pi}\Sigma_i A_i \cdot arcos\left[cos\left(P_{sim,i} - P_{obs,i}\right)\right] / \Sigma_i A_i \quad (8)$$

MPD represents the average timing error, as a proportion of the maximum phase mismatch (6 months).

Table S1: Full benchmarking table with all scores for each model, reference dataset and step considered.

	product	step	mean	random	CLM	CLASS-CTEM	JULES INFERNO	JSBACH SPITFIRE	LPJ-G GlobFIRM	LPJ-G SPITFIRE	SIMFIRE BLAZE	MC2	ORCHIDEE SPITFIRE
burnt area													
spatial	GFED4s	step1	1	1.07	0.60	0.81	0.63	0.67	0.73	0.84	0.83	0.69	0.71
	GFED4	step1	1	1.14	0.81	1.06	0.77	0.88	0.71	0.98	0.98	0.78	0.82
	MCD45	step1	1	1.07	0.68	0.94	0.68	0.74	0.72	0.93	1.00	0.71	0.81
	FireCCI4.0	step1	1	1.13	0.75	0.83	0.78	0.91	0.70	0.90	1.01	0.80	0.81
	FireCCI5.1	Step1	1	1.11	0.83	1.10	0.81	0.95	0.71	0.89	0.99	0.85	0.74
	GFED4s	step2	1	1.07	0.62	0.84	0.71	0.68	0.98	0.88	0.83	0.96	0.72
	GFED4	step2	1	1.14	0.84	1.12	0.76	0.89	0.98	1.00	1.02	0.99	0.91
	MCD45	step2	1	1.16	0.77	1.08	0.69	0.84	0.98	0.95	1.05	0.97	0.90
	FireCCI40	step2	1	1.13	0.77	0.91	0.78	0.92	0.99	0.91	1.05	1.02	0.88
	FireCCI51	Step2	1	1.11	0.85	1.14	0.84	0.96	0.98	0.90	1.01	1.02	0.77
	GFED4s	step3	1	1.07	0.63	0.79	0.72	0.70	1.06	0.94	0.88	1.00	0.72
	GFED4	step3	1	1.14	0.80	0.93	0.85	0.86	1.08	0.98	0.88	1.07	0.71
	MCD45	step3	1	1.16	0.65	0.81	0.72	0.69	1.12	0.93	0.92	1.02	0.70
	FireCCI40	step3	1	1.13	0.77	0.98	0.89	0.92	1.09	0.93	0.97	1.13	0.73
	FireCCI51	Step3	1	1.11	0.83	1.01	0.91	0.93	1.11	0.96	0.97	1.23	0.70
seasonal	GFED4s	step1	0.56	0.22	0.12	0.12	0.13	0.12		0.31			0.31
phase	GFED4	step1	0.49	0.47	0.34	0.35	0.41	0.42		0.33			0.31
	MCD45	step1	0.56	0.26	0.12	0.11	0.12	0.12		0.30			0.30
	FireCCI51	Step1	0.55	0.33	0.26	0.28	0.33	0.32		0.32			0.31
seasonal	GFED4s	step1	1	1.36	1.09	1.22	1.74	1.37		1.07			1.12

concentration	GFED4	step1	1	1.35	1.37	1.56	2.49	1.83		1.23		1.31
	MCD45	step1	1	1.36	1.26	1.40	2.40	1.70		1.15		1.23
	FireCCI51	Step1	1	1.36	1.48	1.72	2.75	2.06		1.26		1.45
	GFED4s	step2	1	1.36	1.07	1.20	1.14	1.28		1.06		1.13
	GFED4	step2	1	1.35	1.20	1.39	1.28	1.48		1.20		1.23
	MCD45	step2	1	1.36	1.18	1.31	1.29	1.49		1.15		1.21
	FireCCI51	Step2	1	1.36	1.27	1.53	1.38	1.64		1.22		1.34
	GFED4s	step3	1	1.36	1.16	1.15	1.24	1.15		1.13		1.22
	GFED4	step3	1	1.35	1.19	1.12	1.25	1.11		1.18		1.19
	MCD45	step3	1	1.36	1.14	1.08	1.26	1.13		1.12		1.20
	FireCCI51	Step3	1	1.36	1.25	1.22	1.33	1.21		1.20		1.27
IAV	GFED4s	step3	1	1.46	1.17	0.65	1.18	1.09	0.66	1.36	0.76	1.66	1.44
	GFED4	step3	1	1.27	0.98	1.62	1.23	0.89	1.04	1.08	1.00	1.41	1.25
	MCD45	step3	1	1.32	0.93	1.34	1.11	0.84	0.73	0.97	1.27	1.67	1.22
	FireCCI51	step3	1	1.42	1.18	1.53	1.24	1.27	1.73	1.27	1.23	1.87	1.12
fire emission													
spatial	GFAS	step1	1	1.08	0.72	1.12	0.74	0.81	0.75	0.94	0.95	0.64	1.11
		step2	1	1.08	0.74	1.33	0.79	0.91	0.97	0.98	0.98	0.99	1.22
		step3	1	1.08	0.78	0.85	0.73	0.74	1.13	1.03	0.91	1.06	0.86
seasonal phase	GFAS	step1	0.78	0.18	0.16	0.20	0.17	0.15		0.37			0.34
seasonal concentration	GFAS	step1	1	1.36	1.14	1.70	1.71	1.39		1.21			1.15
		step2	1	1.36	1.11	1.29	1.21	1.28		1.16			1.15
		step3	1	1.36	1.20	1.22	1.30	1.17		1.27			1.25

IAV	GFAS	step3	1	1.36	0.77	1.70	1.28	1.09	1.42	1.42	1.11	1.41	1.49
fire number													
spatial	Hantson	step3	1	1.19				0.96		0.83			0.76
fire size													
spatial	Hantson	step3	1	1.31				1.03		1.22			1.12
GPP													
spatial	Luyssaert	step1	1	1.39	2.71	1.28	1.39	1.48	2.45	1.83	2.17		1.80
		step2	1	1.39	1.32	1.23	1.33	1.15	1.13	1.01	1.07		0.95
		step3	1	1.39	1.49	1.41	1.46	1.39	1.41	1.24	1.37		1.09
spatial	Jung	step1	1	1.30	0.61	0.48	0.49	0.72	0.47	0.61	0.45		0.46
		step2	1	1.30	0.61	0.50	0.51	0.62	0.46	0.61	0.43		0.47
		step3	1	1.30	0.64	0.46	0.39	0.42	0.46	0.67	0.43		0.49
seasonal	Jung	phase	0.42	0.65	0.18	0.23	0.19	0.23		0.22			0.22
		step1	1	1.65	0.60	0.63	0.59	0.65		0.68			0.59
		step2	1	1.65	0.69	0.74	0.70	0.76		0.77			0.69
		step3	1	1.65	1.08	1.19	1.14	1.21		1.19			1.09
NPP													
spatial	Michaletz	step1	1	1.39	1.27	0.99	1.16	1.12	1.48	1.60	1.55	1.03	1.28
		step2	1	1.39	0.79	0.77	0.77	0.78	0.86	0.85	0.84	0.84	0.87
		step3	1	1.39	0.82	0.79	0.77	0.75	0.96	0.86	0.89	0.88	0.99
spatial	Luyssaert	step1	1	1.33	0.88	1.34	0.62	1.90	0.77	0.81	0.74	1.11	0.78
		step2	1	1.33	0.89	1.26	0.61	1.00	0.78	0.77	0.75	1.16	0.77

		step3	1	1.33	0.90	1.01	0.53	0.76	0.82	0.87	0.79	0.68	0.84
spatial	EMDI	step1	1	1.30	0.77	0.94	0.58	1.05	0.74	0.77	0.76	0.97	0.79
		step2	1	1.30	0.80	0.96	0.59	0.89	0.75	0.72	0.76	1.00	0.77
		step3	1	1.30	0.91	0.87	0.58	0.66	0.79	0.83	0.81	0.65	0.80
LAI													
spatial	MCD15	step1	1	1.29	0.96	0.76	0.78	0.88	0.69	1.85	0.66	1.89	
		step2	1	1.29	1.07	0.83	0.73	0.74	0.67	1.73	0.58	1.63	
		step3	1	1.29	0.60	0.53	0.44	0.78	0.70	0.61	0.57	0.63	
spatial	AVHRR	step1	1	1.34	1.38	1.43	1.03	0.74	0.79	2.57	0.81	2.58	
		step2	1	1.34	1.54	0.90	0.80	0.74	0.69	1.77	0.60	1.67	
		step3	1	1.34	0.81	0.71	0.49	0.65	0.74	0.62	0.61	0.64	
Carbon in vegetation													
spatial	Avitabile	step1	1	1.32	0.73	0.89	0.89	1.35	1.30	1.27	1.30	1.27	0.73
		step2	1	1.32	0.67	0.86	0.78	1.00	0.96	0.97	0.86	0.96	0.66
		step3	1	1.32	0.69	0.88	0.76	0.78	0.76	0.76	0.74	0.80	0.70
spatial	Carvalhais	step1	1	1.32	0.64	0.68	0.61	0.65	0.68	0.65	0.63	0.66	0.51
		step2	1	1.32	0.65	0.70	0.61	0.65	0.63	0.66	0.61	0.66	0.53
		step3	1	1.32	0.66	0.66	0.58	0.64	0.62	0.66	0.58	0.67	0.54

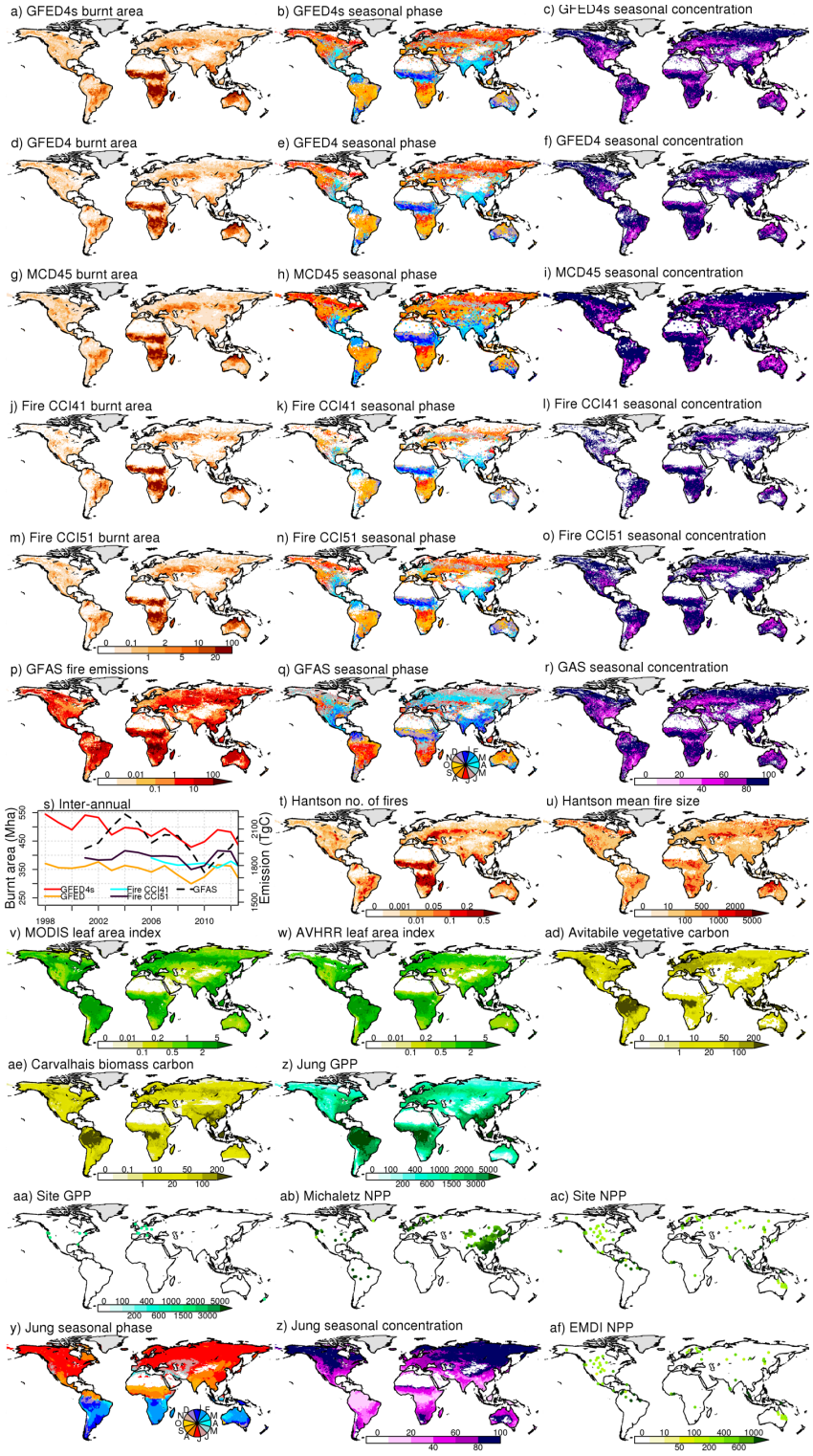

Figure S1: Overview of the reference datasets used.

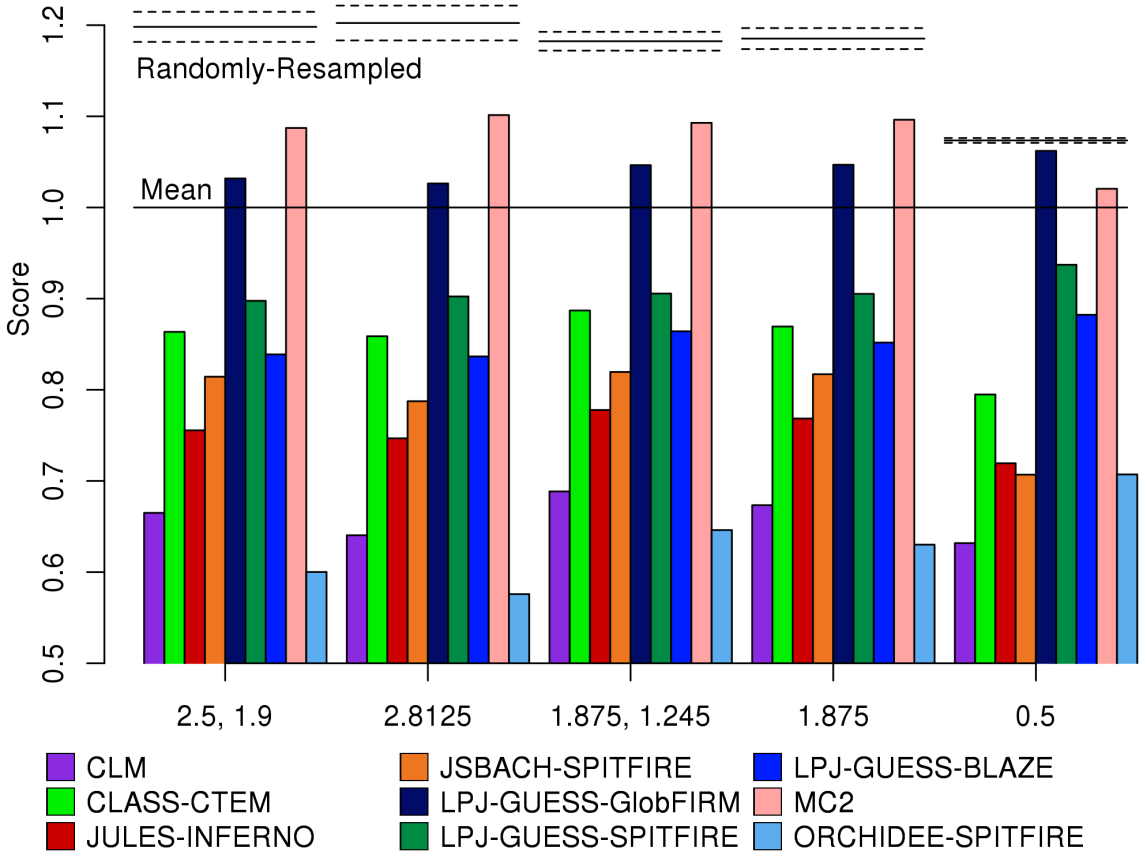

Figure S2: Benchmark scores for each model compared against GFED4s burnt area step 3 at different resolutions for CLM, CLASS-CTEM, JULES-inferno, JSBACH, LPJ-GUESS-GlobFIRM, LPJ-GUESS-SPITFIRE, LPJ-GUESS-BLAZE, MC2 and ORCHIDEE respectively. Each block (left to right) shows the comparison conducted by resampling model output and GFED4s to the 2.5 x 1.9° grid of CLM; 2.8125x2.8125° CLASS-CTEM grid; 1.875 x 1.25° JULES-INFERNO grid, 1.875x1.875° JSBACH-SPITFIRE grid; and the 0.5x0.5° grid used by all other models and for the benchmarking.