Supplement of

A hydroclimatic model for the distribution of fire on Earth

Matthias M. Boer et al.

Correspondence to: Matthias M. Boer (m.boer@westernsydney.edu.au)

Contents

S1. Validation of the climate-fire response surface

The modelled $F_{0.99}$ response surface (Figure 1c) was validated against the 50% of the data that was not used for model fitting. The validation data was binned into 100 mm x 100 mm wide E, D bins and the corresponding values of $F_{0.99}$ identified for all bins with a minimum of 50 observations (N=191). Predicted values of $F_{0.99}$ were extracted from the modelled mean response surface using the same set of 191 E, D coordinates. To evaluate the fit of the model, a linear regression was fitted to the 191 pairs of observed (y_i) and predicted values ($\hat{y}_i$) of $F_{0.99}$ (Fig. S1.1):

$$\hat{y}_i = \alpha + \beta y_i \quad \text{(S1.1)}$$

where y_i and $\hat{y}_i$ are observed and predicted values of $F_{0.99}$ and α and β are fitted coefficients.

The linear regression model was highly significant (p<2.2e-16), with adjusted R^2=0.80, slope (β) of 0.83 (95% CI: 0.76-0.90) and intercept (α) of -0.004 and not significantly different from zero. As recommended by Piñeiro et al.(2008) and Valbuena et al. (2017), Theil's (1958) partial inequality coefficients were used to evaluate the proportions of the total model error due to the unexplained variance (U_{error}), model slope (U_{slope}), and bias (U_{bias}):

$$U_{error} = \sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \,/n \quad \text{(S1.2)}$$

$$U_{slope} = [(1-\beta)^2 \sum_{i=1}^{n}(\hat{y}_i - \overline{y_i})^2]/SS \quad \text{(S1.3)}$$

$$U_{bias} = [n.MD^2]/SS \quad \text{(S1.4)}$$

$$SS = \sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \quad \text{(S1.5)}$$

$$MD = \sum_{i=1}^{n}(\hat{y}_i - y_i)\,/n \quad \text{(S1.6)}$$

where $\hat{y}_i$, y_i are predicted and observed values of $F_{0.99}$, β is the slope of the linear regression of y_i against $\hat{y}_i$ (Fig. S1.1), SS is squared sum of squared residuals, MD the mean deviation between y_i and $\hat{y}_i$, and n the number of validation data pairs. The resulting U_{error}, U_{slope}, and U_{bias} were 0.73, 0.13, and 0.14, indicating that the largest proportion of predictive error was due to unexplained variance, while small proportions were associated to the slope and bias. Consistent with the slight overprediction of the $F_{0.99}$ model, the distribution of model residuals peaked at slightly negative values (Fig. S1.2). Relatively large residuals (e.g. absolute values > 0.2) were more common for validation points of PL-type fire (green dots, Fig. S1.1) than for validation points of DL-type fire (orange dots, Fig. S1.1); this was to be expected because observed $F_{0.99}$ of 0.4 and higher are more common in the domain of PL-type fire but very rare in the domain of DL-type fire.

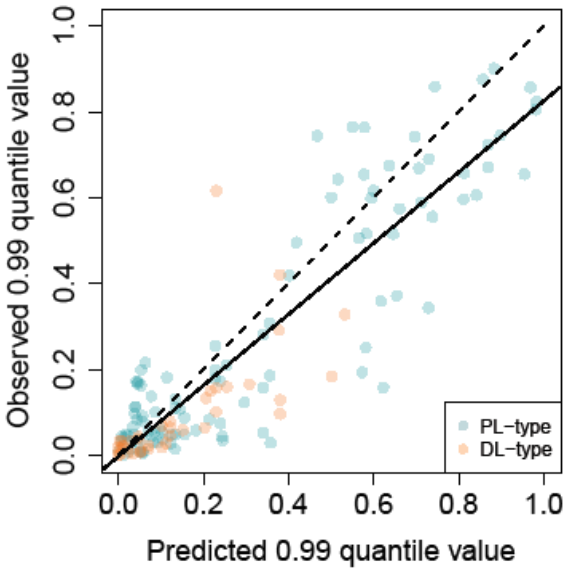


Figure S1.1. Observed versus predicted $F_{0.99}$ for 191 validation grid points evenly spread across E, D space. The continuous black curve is the linear regression line ($\hat{y}_i = -0.004 + 0.083 y_i$) and the dashed lines indicates the 1:1 relationship. Data points are colour coded according to the climate-fire domain being of PL-type (green) or DL-type (orange) fire.

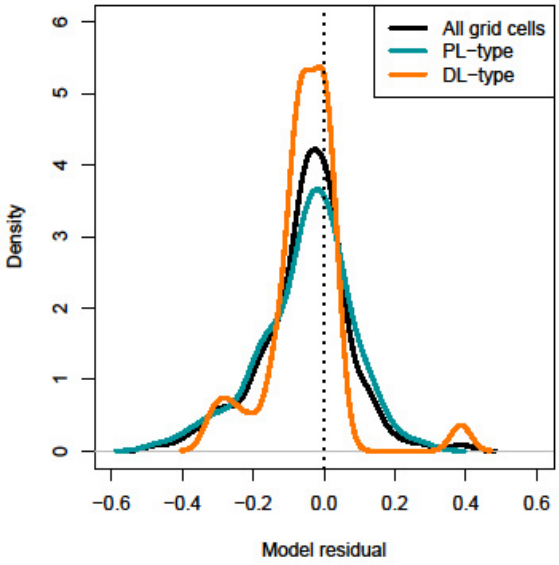


Figure S1.2. $F_{0.99}$ **model (eq. 4) residuals for 191 validation grid points evenly spread across** E, D **space. Green and orange curves show densities of residuals for validation points in the climate-fire domains of PL-type or DL-type fire, respectively.**

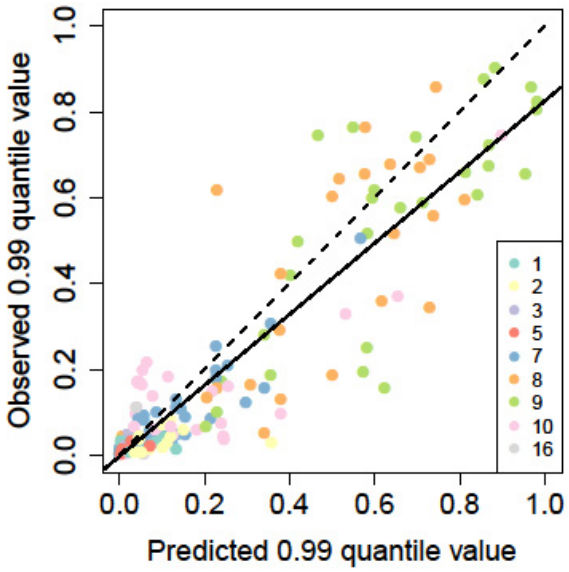


Figure S1.3. Observed versus predicted $F_{0.99}$ for 191 validation grid points evenly spread across E, D space. The continuous black curve is the linear regression line ($\hat{y}_i = -0.004 + 0.083 y_i$) and the dashed lines indicates the 1:1 relationship. Data points are colour coded according to the nine classes of the Moderate Resolution Imaging Spectroradiometer (MODIS) landcover classification (MCD12Q1, version 5.1) represented in the validation data points (N=191): 1) Evergreen needleleaf forest, 2) Evergreen broadleaf forest, 3) Deciduous needleleaf forest, 5) Mixed forest, 7) Open shrublands, 8) Woody savannas, 9) Savannas, 10) Grasslands, 16) Barren or sparsely vegetated.

S2. Summary statistics of observed and predicted fire potential by biome

Table S2.1. Quantiles of observed mean annual fractional burned area (F) and predicted potential mean annual fractional burned area ($F_{0.99}$) for the WWF Biomes included in this study. The global WWF Biome classification is described in Olson et al. (2001) and can be downloaded from: https://www.worldwildlife.org/publications/terrestrial-ecoregions-of-the-world

WWF Biomes		F – biome quantiles					$F_{0.99}$ – biome quantiles				
Code	**Description**	**0.025**	**0.250**	**0.500**	**0.750**	**0.975**	**0.025**	**0.250**	**0.500**	**0.750**	**0.975**
1	(Sub-)Tropical Moist Broadleaf Forests	0.00	0.00	0.00	0.01	0.14	0.02	0.11	0.29	0.44	0.78
2	(Sub-)Tropical Dry Broadleaf Forests	0.00	0.00	0.01	0.03	0.37	0.13	0.55	0.71	0.86	0.96
3	(Sub-)Tropical Coniferous Forests	0.00	0.00	0.01	0.01	0.05	0.06	0.30	0.49	0.63	0.88
4	Temperate Broadleaf Mixed Forests	0.00	0.00	0.00	0.00	0.03	0.00	0.03	0.12	0.26	0.63
5	Temperate Conifer Forest	0.00	0.00	0.00	0.00	0.03	0.01	0.02	0.05	0.14	0.41
6	Boreal Forests Taiga	0.00	0.00	0.00	0.01	0.04	0.00	0.00	0.01	0.01	0.03
7	(Sub-)Tropical Grasslands Savannas Shrublands	0.00	0.01	0.08	0.25	0.70	0.09	0.44	0.63	0.80	0.98
8	Temperate Grasslands Savannas Shrublands	0.00	0.00	0.00	0.03	0.14	0.02	0.05	0.08	0.21	0.56
9	Flooded Grasslands Savannas	0.00	0.03	0.10	0.26	0.75	0.06	0.56	0.77	0.84	0.94
10	Montane Grasslands Shrublands	0.00	0.00	0.00	0.01	0.18	0.01	0.04	0.16	0.41	0.66
11	Tundra	0.00	0.00	0.00	0.01	0.03	0.00	0.00	0.00	0.01	0.01
12	Mediterranean Forests Woodlands Scrub	0.00	0.00	0.00	0.02	0.05	0.06	0.13	0.17	0.25	0.43
13	Deserts Xeric Shrubland	0.00	0.00	0.01	0.05	0.18	0.02	0.05	0.10	0.19	0.77
14	Mangroves	0.00	0.00	0.00	0.02	0.12	0.03	0.04	0.07	0.38	0.76

S3. Classification of climate-fire domains using gradient analysis

The division between the PL-type and DL-type domains can be computed numerically and an analytical solution can also be attempted. As defined in the main text, the potential mean annual fractional burned area ($F_{0.99}$) depends on both mean annual actual evapotranspiration (E) and climatic water deficit (D). We can, therefore, see where in climate space $F_{0.99}$ is more dependent on E (productivity limited, PL-type) and where it is more dependent on D (dryness limited, DL-type). Since D is not independent of E (as in $D = E_0 - E$) we cannot directly examine how $F_{0.99}$ varies with either variable (where increases in E are larger we would observe PL-type conditions and vice versa). Instead, a different approach must be taken.

We take advantage of the fact that all three functions ($F_{0.99}$, D and E) are dependent entirely on observable variables (potential evapotranspiration, E_0, and precipitation, P) and perform a gradient analysis on these three functions. We compare the directional change (gradient; indicated by the symbol ∇) with respect to E_0 and P of $F_{0.99}$ and E and $F_{0.99}$ and D. We must consider that the units of E and D are different to those of $F_{0.99}$ and so is the order of magnitude of this function. Therefore, we only consider the directions of the gradients instead of the distances between the vectors. We do this by comparing the angles between ∇E and $\nabla F_{0.99}$, and ∇D and $\nabla F_{0.99}$ (Fig. S3.4). Where the first is smaller we assign the PL-type domain and vice versa. The angles between the gradients are computed through the simple trigonometric formula:

$$\theta_{EF_{0.99}} = \cos^{-1} \frac{\nabla E \nabla F_{0.99}}{|\nabla E||\nabla F_{0.99}|} \quad \text{(S3.7)}$$

$$\theta_{DF_{0.99}} = \cos^{-1} \frac{\nabla D \nabla F_{0.99}}{|\nabla D||\nabla F_{0.99}|} \quad \text{(S3.8)}$$

where $\theta_{EF_{0.99}}$ and $\theta_{DF_{0.99}}$ are the angles between $F_{0.99}$ and E and D respectively (Fig. S3.4).

Additionally, since we are only interested in which domain a grid point belongs to, we can omit the calculation of the angle and focus on the values of the cosines. We assign a domain association parameter *K*:

$$K = \frac{\nabla E \nabla F_{0.99}}{|\nabla E||\nabla F_{0.99}|} - \frac{\nabla D \nabla F_{0.99}}{|\nabla D||\nabla F_{0.99}|} \quad \text{(S3.9)}$$

K can then tell us which domain $F_{0.99}$ belongs to:

$$\begin{cases} K < 0 & \Rightarrow \quad \text{DL} - \text{type} \\ K = 0 & \Rightarrow \quad \text{undefined} \\ K > 0 & \Rightarrow \quad \text{PL} - \text{type} \end{cases}, \tag{S3.10}$$

within the limits $0 < E \leq E_2$ and $0 < D \leq D_2$ (see eq. 2-4, main text).

The K formula is further computed using the symbolics package in Matlab R2018a. However, the resulting model for K is complex and non-linear. Nevertheless, within the above limits of $F_{0.99}$ K is approximately linear (Fig. S3.5). We therefore used linear regression through the origin point on the numerically derived dataset for $K = 0$. The resulting line $E_0 = (1.94 \pm 0.02)P$ is significant (p-value < 0.001 and $R^2 = 0.9955$).

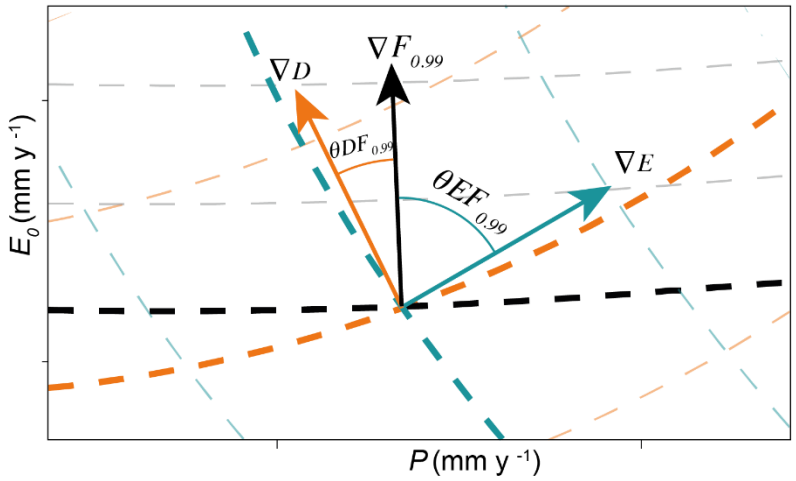
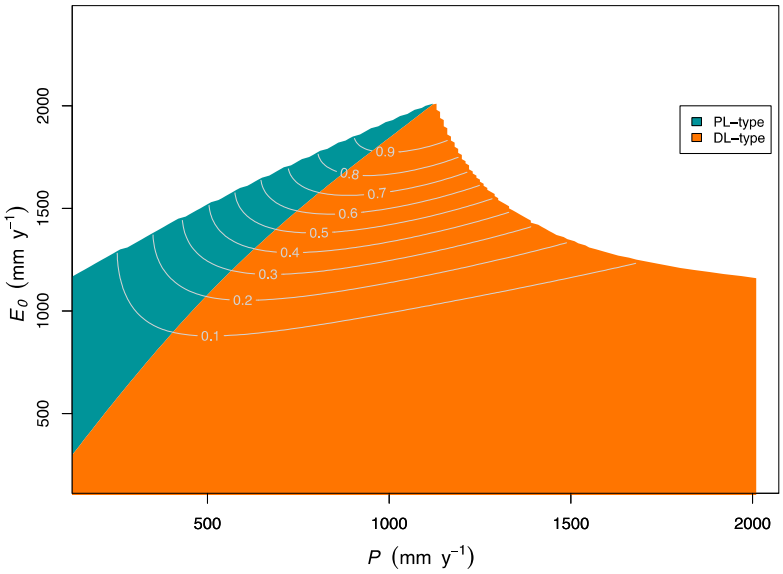


Figure S3.4. An illustration of how gradients of the three model components are compared. Contours of the response surfaces of $F_{0.99}$, E and D are shown in dashed black, green and orange curves, respectively, while the vectors $\nabla F_{0.99}$, ∇E, and ∇D are shown in solid black, green and orange arrows. Within the orthogonal P, E_0 space, where the angle between the gradients of $F_{0.99}$ ($\nabla F_{0.99}$) and D (∇D), $\theta_{DF_{0.99}}$, is smaller than the angle between the gradients of $F_{0.99}$ ($\nabla F_{0.99}$) and E (∇E), $\theta_{EF_{0.99}}$, we see the predominance of fuel dryness driving mean annual fractional burned area (DL-type fire) and where the angle $\theta_{DF_{0.99}}$ is larger than the angle $\theta_{EF_{0.99}}$ we see a predominance of fuel productivity being the dominant component (PL-type fire).



115 **Figure S3.5. Distribution of the domains of PL- and DL-type fire in P, E_0 space (within the limits $0 < E \leq E_2$ and $0 < D \leq D_2$) estimated from an analysis of the direction of gradients $F_{0.99}$, E and D. Contours of the $F_{0.99}$ model are in grey.**

S4. Frequency of dry fuel moisture days against climatic water deficit

The mean annual frequency of days (1980-2017) with predicted fine dead fuel moisture content (DFMC) under 10% was computed using global MERRA2 reanalysis daily vapour pressure deficit data (VPD). Daily DFMC was predicted for global forest regions from daily VPD using the model proposed by Resco de Dios et al. (2015). The 10% DFMC threshold was shown to be a critical threshold in the level of fire activity in temperate forests of SE Australia (Nolan, Boer, Resco de Dios, Caccamo, & Bradstock, 2016), and similar thresholds were found for forests in Portugal (Boer et al., 2017). Mean annual climatic water deficit was computed from mean annual precipitation and potential evapotranspiration as described in the Methods section of the main text.

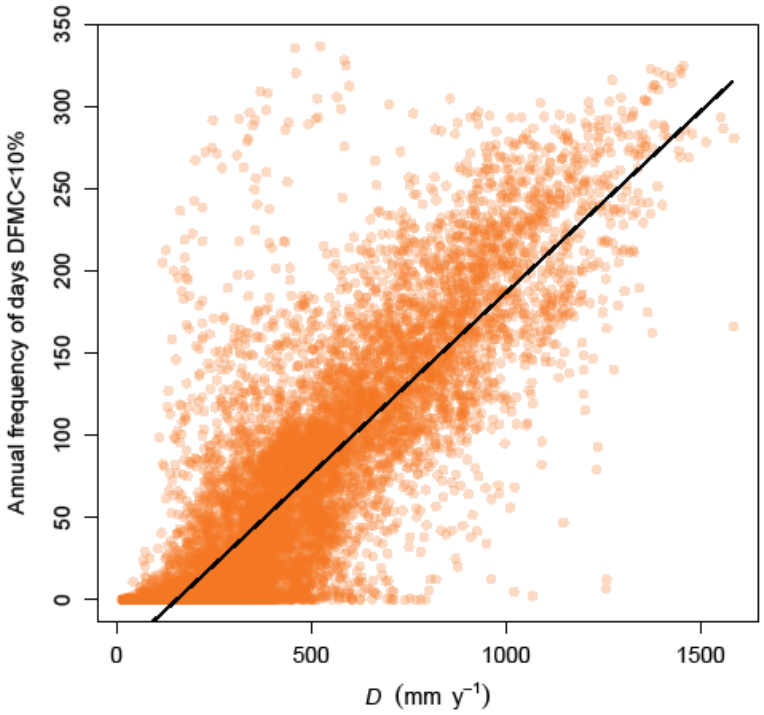


Figure S4.6. The mean annual frequency of days with predicted fine dead fuel moisture content (DFMC) below 10% in global forest regions against mean annual climatic water deficit, D. The continuous black curve is the linear regression model, y=-34.15+0.22x, which has an adjusted R^2=0.76 and p<2e-16)

References

Boer, M. M., Nolan, R. H., Resco De Dios, V., Clarke, H., Price, O. F., & Bradstock, R. A. (2017). Changing Weather Extremes Call for Early Warning of Potential for Catastrophic Fire. *Earth's Future, 5*(12), 1196-1202. doi:10.1002/2017EF000657

Nolan, R. H., Boer, M. M., Resco de Dios, V., Caccamo, G., & Bradstock, R. A. (2016). Large-scale, dynamic transformations in fuel moisture drive wildfire activity across southeastern Australia. *Geophysical Research Letters, 43*(9), 4229-4238. doi:10.1002/2016GL068614

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., . . . Kassem, K. R. (2001). Terrestrial ecoregions of the world: a new map of life on Earth. *Bioscience, 51*(11), 933-938.

Piñeiro, G., Perelman, S., Guerschman, J. P., & Paruelo, J. M. (2008). How to evaluate models: Observed vs. predicted or predicted vs. observed? *Ecological Modelling, 216*(3-4), 316-322.

Resco de Dios, V., Fellows, A. W., Nolan, R. H., Boer, M. M., Bradstock, R. A., Domingo, F., & Goulden, M. L. (2015). A semi-mechanistic model for predicting the moisture content of fine litter. *Agricultural and Forest Meteorology, 203*(0), 64-73. doi:http://dx.doi.org/10.1016/j.agrformet.2015.01.002

Valbuena, R., Hernando, A., Manzanera, J. A., Görgens, E. B., Almeida, D. R. A., Mauro, F., . . . Coomes, D. A. (2017). Enhancing of accuracy assessment for forest above-ground biomass estimates obtained from remote sensing via hypothesis testing and overfitting evaluation. *Ecological Modelling, 366*, 15-26. doi:https://doi.org/10.1016/j.ecolmodel.2017.10.009