Supplement of

Technical note: Low meteorological influence found in 2019 Amazonia fires

Douglas I. Kelley et al.

Correspondence to: Douglas I. Kelley (douglas.i.kelley@gmail.com) and Chantelle Burton (chantelle.burton@metoffice.gov.uk)

The copyright of individual parts of the supplement might differ from the CC BY 4.0 License.
Model Evaluation

We used the same set of evaluation as Kelley et al. (2019). As we are also looking at annual variations for seasonal and sub-seasonal model simulation of burnt area, we also included additional comparisons for seasonal and interannual variability. Comparisons were made against MCD64A1 for July 2002 - June 2018. Monthly gridded simulations for all 1000 posterior samples were compared using Normalised Mean Squared Error (NMSE) (Kelley et al., 2013). Annual average and June to September average comparisons used Normalised Mean Error (NME) (Kelley et al., 2013) as recommended by the Fire Model Intercomparison Project (FireMIP, (Rabin et al., 2017)). NMSE and NME sum the squared (for NMSE) or absolute (for NME) distance between observations (obs) and reconstructed burnt area from a parameter set (sim(β)) over all cells (i) weighted by cell area (Ai) and normalised by mean variation in obs:

\[
NME = \frac{\sum_i A_i \times \left| \text{sim}(\beta_i) - \text{obs}_i \right|}{\sum_i A_i \times \left| \text{obs}_i - \text{obs}_i \right|}
\quad \text{and} \quad
NMSE = \frac{\sum_i A_i \times (\text{sim}(\beta_i) - \text{obs}_i)^2}{\sum_i A_i \times (\text{obs}_i - \text{obs}_i)^2}
\] (S1)

NME and NMSE comparisons were conducted in three steps:

1. As described above;
2. Comparing obs, and sim, after taking the difference between their respective means, thereby removing systematic bias
3. obs, and sim, are additionally divided by the mean deviation, which therefore describes the models’ ability to reproduce the spatial pattern in burnt area.

Seasonality comparisons were conducted in three parts: 1) modality - i.e how many “seasons” there are in a given year; 2) phase, or timing, of the season; and 3) seasonal concentration (inverse of season length).

To determine the modality of a given cell, we first calculated the monthly climatology (v0). The month of the minimum burnt area from this climatology was defined as the start of the “fire year”. We then found the position (P) of each maxima turning point (pi) throughout the year:

\[
v(1) = \min(v_0)
\]

\[
P = \left\{ p_i \mid \frac{dv(p_i)}{dt} = 0 \land \frac{d^2v(p_i)}{dt^2} < 0 \land v(p_i) < v(p_{i+1}) \right\}
\] (S2)
The modality (MOD) was then the prominence of each of these turning points (i.e, the minimum drop required to the next turning point), weighted by the phase distance (θ) to the next turning point. This is normalised by the height of the month of maximum burnt area

\[
\theta_p = 2 \times \pi \times \left( \frac{p - 1}{12} \right)
\]

\[
MOD = 1 + \frac{\sum_{i=1}^{p} \left( v(p_i) - \min\{v(p_i), v(p_{i+1})\} \right) \times \left( 1 - \cos(\theta_{p_i} - \theta_{p_{i+1}}) \right)}{2 \times \left( \max(v) - v(1) \right)}
\]

If there was no fire, then MOD is undefined and no comparison was made for that grid cell. If there were no turning points, then modality was set to 0. If there was one turning point, MOD was set to 1. The higher the number beyond that, the higher the modality the more “seasons” within a year. Two equally prominent peaks 6 months apart have a modality of 2. Observational and simulated MOD was then compared using NME.

Phase and concentration comparisons were conducted as per (Kelley et al., 2013), each month, \( p \), is represented by a vector whose direction corresponds to the time of year (equation S3) and length to the magnitude of the variable for that month. A mean vector \( L \) was calculated by averaging the \( x \) (\( L_x \)) and \( y \) (\( L_y \)) components of the 12 vectors (\( x_p \)).

\[
L_x = \sum p x_p \times \cos(\theta_p)
\]

\[
L_y = \sum p x_p \times \sin(\theta_p)
\]

(S5)
The mean vector length by the annual average described the seasonal concentration \((C)\) of burnt area, whilst it’s direction \((P)\) described season timing:

\[
C = \sqrt{L_x^2 + L_y^2} / \Sigma x \cdot y \\
P = \arctan(L_x / L_y)
\]  

(S6) (S7)

If burnt area in a given cell was concentrated all in one month, \(C\) was equal to 1 and \(P\) corresponded to that month. If burnt area was evenly spread throughout the year then concentration was set to zero and phase undefined and was not used in the comparison. Likewise, if a cell had zero annual average burnt area for either observations or simulation, then that cell was not included in the comparisons. Concentration was compared using NME step 1. Phases were compared using mean phase difference \((MPD)\)

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

(S8)

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

We assessed temporal variations using spearman's rank (Dodge, 2008; Lasslop et al., 2018). This was performed for each grid cell both monthly and, to assess inter-annual variability, annually-averaged burnt area. The score was then the area-weighted average comparison of all grid cells.

Smaller \(NME\), \(NMSE\) and \(MPD\) scores indicate a better agreement between simulation and observation, with a perfect score (i.e., simulation that perfectly matches observations) of 0. Greater spearman's rank scores indicate better performance, with a score of 1 occurring with perfect ordering in the simulation and -1 complete reverse ordering. We also used three null models to help interpret the score as per (Burton et al., 2019; Kelley et al., 2019). The mean null model compared the mean of all observations with the observations. For \(NME\) and \(NMSE\), the mean null model was always 1 as these metrics are normalised by the mean difference. As there is no ordering in a mean null model, the spearman's rank comparison gives a score of 0. The best “single value” model compared the median of observations to observations. For \(NME\) and \(NMSE\), the mean null model was always 1 as these metrics are normalised by the mean difference. As there is no ordering in a mean null model, the spearman's rank comparison gives a score of 0. The best “single value” model compared the median of observations to observations. By definition, it’s score is less than or equal to the mean model score for \(NME\) and, again, a score of 0 for spearman's rank. The mean and median null model scores for \(MPD\) depends on the observations. The “randomly resampled” null model compared randomly-resampled observations (without replacement) to the observations. The score depends on the resampling order, so we used 1000 bootstraps to determine the null models’ distribution.
Our monthly \textit{NMSE} step 1 scores of 0.896-0.931 (based on 5-95\% percentile, Fig. S1) are much better than all null models. \textit{NMSE} is proportional to our error term in equation 3, indicating good convergence of the model, though the Fig. 1 time series demonstrates that a lot of the models’ posterior spread is still within the error term rather than parameter uncertainty, suggesting that further development of the models’ process representation is still possible (see discussion).

Our reconstructed annual average burnt area \textit{NME} step 1 scores of 0.687-0.711 is comparable to the comparison between training observations and simulation in the global model, which obtained scores of 0.603-0.630 (Kelley et al., 2019). Our conversion to a coarser grid and pragmatic choice of variables have therefore not been detrimental for model performance, possibly helped by our restricted geographic range. Step 3 scores of 0.726-0.773, are 7-13\% better than our best null model which suggests good performance in the spatial pattern of burnt areas. The model spatial burnt area is even better for the fire season (Fig. 2 and 3), with step 1 scores of 0.605-0.626 and step 3 of 0.573-0.623. However, the median null model also improves, which suggests that capturing the fire season should be easier, though it should be noted that other fire models normally struggle to pick up high burnt areas during fire season peaks (Hantson et al., 2020).

The model largely identifies regions with bimodal fire seasons (Fig. S2), though with 33.6\% of the posteriors Step 3 scores being greater than the mean null model suggests that this is not easily captured by the model. Seasonal human fire manipulation is often a cause of bimodal fire systems (Archibald et al., 2009; Hall et al., 2016), which is not included in the modelling framework and again would be useful in future studies (see discussion). For single modal systems, the timing of the fire season is very well captured in the model (Fig. S3), with a score of 0.170-0.172, 45-56\% better than the closest null model. Most (93.6\%) seasonal concentrations of the model are better than the mean null model for step 3. However, the fire season is often too short in agricultural areas (Fig. S2). Again, this could be due to the lack of seasonal variation in human fire starts (Archibald et al., 2009; Hall et al., 2016).

Overall, the model correctly ranks the months’ burnt area (Fig. S1 and S2) with a score of 0.497-0.543, demonstrating that the overall season is well represented. The model also correctly ranks the ordering of annual burnt areas (Fig. S3), scoring 0.227-0.236 - better than all null models, thereby showing that it is able to capture the interannual variability. For our AAD region, these improve further, with 0.703-0.724 for monthly ranking and 0.412-0.422 for annual ranking (Fig. S1).
Figure S1: Metric scores for the models posterior for steps 1-3 (red, orange and yellow) and median (grey dot-dashed line), mean (light blue dashed) and randomly-resampled (blue distribution) null models.
Figure S2: Seasonal comparisons for (top row) MCD64A1 observed and (2nd row) 50% percentile of the models posterior for (1st column) modality, (2nd column) phase of the timing of the fire season and (3rd column) concentration (inverse of season length) of the fire season. The bottom row shows the spatial pattern of spearman’s rank comparisons between MCD64A1 and the 50% percentile of the models posterior and a monthly and annual timestep.
Figure S3: 2002-2018 climatology for MCD64A1 observed burnt area on the 50% percentile model for each region. The thick line shows the median of all cells in the region, while shaded areas show full range.
Figure S4: Changes in tree cover and agriculture from 2002-2018. Top row shows (from left to right) percentage tree cover from VCF (Dimiceli et al., 2015), and cropland, pasture and total agricultural cover from HYDEv3.1 (Klein Goldewijk et al., 2010). The 2nd row shows the corresponding trends in tree cover and agriculture in percentage over 2002-2018. As variables are bounded, we use the trend analysis from (Kelley et al., 2019) using the “greenbrown” R package (Forkel et al., 2013, 2015). Dots signify areas of significant trends (p-value < 0.05). The bottom map shows areas of increases, decreases or no significant trend in tree cover and agriculture. Red areas indicate our “area of active deforestation” (F in Fig. 1).
Table S1: Controls, drivers and target variables and data sources used in the model

<table>
<thead>
<tr>
<th>Control</th>
<th>Variable</th>
<th>Calculated as</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuel continuity “Fuel” (%)</td>
<td>Total vegetation cover (%)</td>
<td>1 - bare cover</td>
<td>MODIS Vegetation Continuous Fields (VCF) (Dimiceli et al., 2015)</td>
</tr>
<tr>
<td></td>
<td>Maximum seasonal anomalies in water availability</td>
<td>( \frac{n}{n_{max}} - 1 ) (see row below)</td>
<td>NCEP/NCAR soil moisture (Kalnay et al., 1996)</td>
</tr>
<tr>
<td>Fuel moisture “Moisture” (%)</td>
<td>Soil moisture (%)</td>
<td>( \theta ) - 0</td>
<td>NCEP/NCAR relative humidity, temperature. (Kalnay et al., 1996)</td>
</tr>
<tr>
<td></td>
<td>Equilibrium fuel moisture content (%)</td>
<td>as per (Kelley et al., 2014)</td>
<td>GPCP precipitation (monthly) (Adler et al., 2003)</td>
</tr>
<tr>
<td></td>
<td>Tree Cover (%)</td>
<td></td>
<td>VCF (Dimiceli et al., 2015)</td>
</tr>
<tr>
<td>Potential ignitions “Ignitions” (no. km²)</td>
<td>Lightning strikes (strikes km⁻²)</td>
<td>Cloud-to-ground as per (Kelley et al., 2014)</td>
<td>LIS/OTD lightning flash counts(Cecil et al., 2014)</td>
</tr>
<tr>
<td></td>
<td>Population density (people km⁻²)</td>
<td></td>
<td>HYDEv3.1(Klein Goldewijk et al., 2010)</td>
</tr>
<tr>
<td></td>
<td>Pasture (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anthropogenic suppression “Suppression”</td>
<td>Cropland (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Population density (people/km²)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target</td>
<td>Burnt area</td>
<td></td>
<td>MCD64A1 burned area (Giglio et al., 2018)</td>
</tr>
</tbody>
</table>
Table S2: Driving data for optimization and simulation.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>density</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimization</td>
<td>Used</td>
<td>Note used</td>
<td></td>
</tr>
<tr>
<td>Simulation</td>
<td>Used</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
References


