THE IMPORTANCE OF VEGETATION BUILD UP FOR BURNED AREA SEASONALITY

ALEXANDER KUHN-REGNIER
APOSTOLOS VOULGARAKIS
COLIN PRENTICE
SANDY HARRISON
THE NEED FOR WILDFIRE MODELS

- Recent extreme weather events have caused wildfires to be scrutinised more closely (e.g. Herring et al., 2015, Yoon et al. 2015)
- Increasing wildfires as a result of climate change have been suggested on local scales (e.g. Westerling, 2006)
- Global changes in the proportion of burned area relative to wildfire emissions have also been observed (e.g. due to cropland changes (Niels Andela and van der Werf 2014; N. Andela et al. 2017))
- However, the short satellite record impedes the reliable estimation of changes in global burned area from observations (Forkel et al. 2019b)
- Wildfire models are a way to estimate future consequences of climate change scenarios using both data as well as an understanding of underlying processes
- Thus, they are key to understanding potential future variability of fire regimes
VEGETATION BUILD UP

- When constructing wildfire models, different effects must be considered
- Short term effects are dominated by weather, while vegetation properties linked to fuel build up are important for seasonal or longer timescales
- Thus, we need a better understanding of the timescales of fuel build up to understand its effects on fire regimes
- Past studies (e.g. Forkel et al. 2019a) have already demonstrated the importance of pre-season vegetation properties for burned area prediction
- Additionally, the relationships between burned area seasonality and vegetational drivers is still poorly understood
- This could be responsible for the shortcomings identified by Hantson et al., 2020 regarding current wildfire models' prediction of fire seasonality
METHOD

- A range of biophysical variables were selected as input factors
- These were then fed into a random forest machine learning model to predict global burned area
- To improve our understanding of the importance of pre-seasonal effects we shifted selected variables backwards in time
- Thus, we can evaluate the importance of antecedent effects by evaluating the feature importances of the shifted features
- Since the effects are coupled, an analysis of the vegetation drivers requires an analysis of biophysical variables in general
## METHOD - DATASETS

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Start</th>
<th>End</th>
<th>Frequency</th>
<th>Latitude (°)</th>
<th>Longitude (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avitabile Thurner AGB</td>
<td>static</td>
<td>static</td>
<td>static</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>Copernicus SWI</td>
<td>01/2007</td>
<td>11/2018</td>
<td>monthly</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>ERA5 Dry Day Period</td>
<td>01/1990</td>
<td>12/2018</td>
<td>monthly</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>ERA5 Temperature</td>
<td>01/1990</td>
<td>12/2018</td>
<td>monthly</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>ESA Landcover</td>
<td>01/1992</td>
<td>01/2015</td>
<td>yearly</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>GFED4</td>
<td>06/1995</td>
<td>12/2016</td>
<td>monthly</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>HYDE</td>
<td>01/2000</td>
<td>01/2017</td>
<td>yearly</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>MOD15A2H</td>
<td>02/2000</td>
<td>11/2018</td>
<td>monthly</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>VODCA</td>
<td>12/1997</td>
<td>07/2017</td>
<td>monthly</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>WWLLN Lightning</td>
<td>01/2010</td>
<td>12/2018</td>
<td>monthly</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Overall</td>
<td>01/2010</td>
<td>01/2015</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Input datasets are shown, along with their combined overall time period.
- All datasets were regridded to a common 0.25° grid at a monthly, climatological resolution (2010-2015).
DATASETS - PROCESSING

- GFED4: Global burned area
  - This dataset was chosen due to our focus on seasonality instead of small fires
- ERA5 precipitation data: contiguous number of dry days (‘Dry Day Period’, 0.1 mm / day threshold)
- ERA5 (2m) temperatures: diurnal temperature range and maximum temperature
- WWLLN Lightning dataset: lightning ground strikes
- HYDE: population density
  - Other indices like the night light development index (NLDI) have been shown to yield similar results at best
- Multiple vegetation properties: MOD15A2H – (fAPAR, LAI), VODCA – Ku-band VOD, ESA CCI Landcover – PFTs
  - Fractional PFT coverage from Poulter et al. 2015 (M. Forkel, personal communication)

---

**Dataset**

- **Avitabile Thurner AGB** (Avitabile et al. 2016)
- **Copernicus SWI** (Bauer-Marschallinger et al. 2018)
- **ERA5 Dry Day Period** (Copernicus Climate Change Service (C3S) (2017) 2017)
- **ERA5 Temperature**
- **ESA Landcover** (Li et al. 2018)
- **GFED4** (Giglio, Randerson, and Werf 2013)
- **HYDE** (Goldewijk, Beusen, and Janssen 2010)
- **MOD15A2H** (Myneni, Knyazikhin, and Park 2015)
- **VODCA** (Moesinger et al. 2019)
- **WWLLN Lightning** (Kaplan and Lau 2019)
## Datasets - Processing

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Source/Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Avitabile Thurner AGB</strong></td>
<td>(Avitabile et al. 2016)</td>
</tr>
<tr>
<td><strong>Copernicus SWI</strong></td>
<td>(Bauer-Marschallinger et al. 2018)</td>
</tr>
<tr>
<td><strong>ERA5 Dry Day Period</strong></td>
<td>(Copernicus Climate Change Service (C3S) (2017) 2017)</td>
</tr>
<tr>
<td><strong>ERA5 Temperature</strong></td>
<td></td>
</tr>
<tr>
<td><strong>ESA Landcover</strong></td>
<td>(Li et al. 2018)</td>
</tr>
<tr>
<td><strong>GFED4</strong></td>
<td>(Giglio, Randerson, and Werf 2013)</td>
</tr>
<tr>
<td><strong>HYDE</strong></td>
<td>(Goldewijk, Beusen, and Janssen 2010)</td>
</tr>
<tr>
<td><strong>MOD15A2H</strong></td>
<td>(Myneni, Knyazikhin, and Park 2015)</td>
</tr>
<tr>
<td><strong>VODCA</strong></td>
<td>(Moesinger et al. 2019)</td>
</tr>
<tr>
<td><strong>WWLLN Lightning</strong></td>
<td>(Kaplan and Lau 2019)</td>
</tr>
</tbody>
</table>

- Shifted variables (1, 3, 6, 9, 12, 18, and 24 months):
  - Vegetation optical depth (VOD)
  - Frequency of absorbed photosynthetically active radiation (fAPAR)
  - Leaf area index (LAI)
  - Dry day period

- For ease of interpretability (to extract anomalies) we further subtract the preceding seasonal cycle:
  - 12 → (12 − 0) Months
  - 18 → (18 − 6) Months
  - 24 → (24 − 0) Months
Input factors and their correlations are shown on the right. Vegetation indices are correlated with a Pearson's $r \sim 0.7$ across their monthly (climatological) temporal shifts. This large degree of correlation makes interpretation of feature importances difficult. Larger temporal shifts (not shown) generally have a very low correlation ($|r|<0.2$) as expected.
MODEL PERFORMANCE - CLIMATOLOGY

- $R^2$ (Train: 0.74, Test: 0.61) - decent performance with some overfitting

- Expected fitted features included a negative relationship for instantaneous vegetation proxies (fuel dryness) and a positive relationship with antecedent vegetation proxies (fuel accumulation), illustrated using the leaf area index (LAI) on the right (top and bottom, respectively)

- As seen on the next slide, other features are not always so clear-cut, e.g. dry day period:
  - A larger instantaneous dry day period increases burned area
  - The 6-month antecedent dry day period has the opposite relationship but the effect boundary is less distinct
FEATURE IMPORTANCE - SHAP

- SHAP values have the potential to provide robust local and global feature importance attributions (Lundberg and Lee 2017)
- The dry day period is ranked most important, with many samples responding very strongly to long dry periods (high feature value), as expected
- Large shifts (>12 months) are not amongst the 20 most important features, so they usually do not influence the burned area significantly
Gini importance and permutation feature importance (PFI) are (simpler) alternatives to explain feature importance.

However, all 3 methods agree that the following are important:
- Dry day period
- VOD (including its 1, 3, and 6-month shifts)
- fAPAR
- Maximum temperature

There is reasonable agreement with previous studies (e.g., Forkel et al. 2019a) but features more fine-grained temporal information.
INTERACTION EFFECTS

- Accumulated local effects (ALE) plots are a robust alternative to partial dependence plots (PDPs) without extrapolating to unlikely samples (amongst other benefits).

- First order effects shown here demonstrate the expected effects:
  - Increase in burned area with instantaneous dry day period (top)
  - Decrease in burned area with instantaneous fAPAR (bottom)
INTERACTION EFFECTS – AN EXAMPLE

- The interaction effect between fAPAR and dry day period (bottom) can be understood to **add to** the individual first-order effects shown before.

- The magnitude of the second-order interaction is **significant** compared to the first-order effects.

- So when fAPAR and dry day period are high, there is an additional **negative** effect, on top of the two aforementioned individual effects.

- Affected regions are shown in the map above (mean fAPAR across all times for regions where the shown conditions are met at any time).

- Affected biomes are usually adjacent to tropical forests, e.g. grass/tree savanna.

- Biome-specific effect?
FUTURE WORK

- Link more of the second-order interactions fitted by the model to global and/or regional effects
- Include solar induced fluorescence (SIF) as another proxy for plant productivity
- Perform a sensitivity analysis regarding the used burned area dataset
- Isolate the effects of antecedent meteorological conditions by keeping vegetational predictors fixed
- Account for antecedent burning explicitly


