# Wildfires Impact on $PM_{2.5}$ Concentration in Galicia Spain

## Supplementary material

### 1. Evaluation: Criteria and Benchmarks for Validation of Datasets

In this study, the analysis dataset was rigorously evaluated using several statistical scores: Rho (Pearson's Rank Correlation Coefficient), IOA (Index of Agreement), RMSE (Root Mean Square Error), MB (Mean Bias), FAC2 (Factor of Two), NMB (Normalized Mean Bias), HIT (Hit Rate), FAR (False Alarm Rate), and POC (Proportion of Correct Responses). Each of these scores provides a quantitative measure to assess the accuracy and reliability of the CAMS data.

Table SM1 outlines the benchmarks and thresholds established for each score. These benchmarks serve as criteria against which the computed values of all statistical scores are compared. If the computed value meets or exceeds the benchmark, it indicates that the CAMS dataset accurately represents the observed data. On the other hand, if the computed value falls below the benchmark, it suggests that there may be biases or discrepancies in the reanalysis dataset.

The evaluation of the model's performance involves several metrics. The Root Mean Square Error (RMSE) determines the standard deviation of the differences between the forecasted and observed values, helping to assess the variability in prediction errors. The Normalized Mean Bias (NMB) represents the average normalized difference between the model's forecasts and observations, providing insights into the bias and accuracy of the model, while the Mean Bias (MB) offers similar insights without normalization. The Factor of 2 (FAC2) metric indicates the fraction of data where predictions are within a factor of two of the observations, reflecting the model's reliability within a specific range. The Index of Agreement (IOA), which is the ratio of the RMSE to the potential correlation error, provides an overall measure of the model's predictive error. The spearman correlation (Rho) value represents the proportion of variance in the observations explained by model predictions, signifying the model's explanatory power. Additionally, the model's capability to correctly predict the occurrence of an event, including both exceedances and non-exceedances of a selected threshold, is measured by the proportion of correctness (POC), crucial for practical applications. The False Alarm Ratio (FAR) represents the frequency of false forecasts of extreme events, while the Hit Rate (HIT) measures the model's capability to correctly forecast an extreme event, both critical for understanding the model's reliability in predicting rare but significant events. For a more complete explanation, please consult Celis et al. (2021) supplementary material.

Table SM1. Validation statistical results

| **Station_name Meteogalicia** | **In this paper** | **Rho** | **IOA** | **RMSE** | **MB** | **FAC2** | **NMB** | **HIT** | **FAR** | **POC** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Benchmarck $PM_{2.5}$** | | >0.4 | >0.4 | <70 | +/- 25% | >50% | +/- 25% | >40% | >15 μg/m3 | >50% |
| Xubia | Xubia | 0.3 | 0.5 | 8.7 | -3 | 59.3 | -27.6 | 44.6 | 33.9 | 53 |
| Sur | Sur | 0.4 | 0.6 | 6.7 | -1.4 | 76.5 | -14 | 58.2 | 35.2 | 53.5 |
| Santiago-Caetano | Santiago-C | 0.6 | 0.7 | 6.9 | -2.3 | 77.5 | -19.6 | 79.9 | 18.5 | 59.2 |
| Coruna-Torre | La Coruña | 0.4 | 0.5 | 11.7 | -6.3 | 64.8 | -42.7 | 87.8 | 9.8 | 62.2 |
| Pontevedra-Campolongo | Pontevedra | 0.4 | 0.5 | 9.4 | -2.4 | 67.2 | -22.9 | 44.6 | 31.9 | 53.2 |
| Vigo-Coia | Vigo | 0.6 | 0.5 | 10.8 | -5.1 | 52.5 | -45.7 | 63 | 22.6 | 52.4 |
| Ferrol-Reina | Ferrol | 0.5 | 0.7 | 5.6 | -0.5 | 68.4 | -6.3 | 18.1 | 41.7 | 51.4 |
| Laza | Laza | 0.4 | 0.5 | 9.2 | 1.2 | 65.1 | 18.7 | 21 | 66.2 | 48.1 |
| Lugo-Fingoy | Lugo | 0.5 | 0.6 | 8.1 | -2.8 | 66.1 | -27.2 | 41.1 | 25 | 54 |
| Ourense-Gmez | Ourense | 0.4 | 0.5 | 8.9 | -2.2 | 62.3 | -22.5 | 41.1 | 33.5 | 52.6 |
| Centro-civico | Centro-civico | 0.5 | 0.6 | 8.1 | -1 | 72.3 | -10.5 | 48.1 | 38.1 | 52.5 |
| Grela | Grela | 0.5 | 0.7 | 7.1 | -3.6 | 74.9 | -30.2 | 67.6 | 15.3 | 58 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Fraga-Redonda | Fraga | 0.5 | 0.7 | 4.3 | -0.2 | 85 | -3 | 25.5 | 51.9 | 49.7 |
| Burela | Burela | 0.4 | 0.6 | 5.3 | -2.6 | 69.5 | -27.1 | 44.1 | 42.1 | 51.1 |
| San-Vicente_vigo | San-Vicente_vigo | 0.5 | 0.6 | 6.6 | -3.2 | 73.9 | -27.8 | 64.2 | 19.6 | 56.7 |
| Magdalena | Magdalena | 0.5 | 0.7 | 4.5 | -0.2 | 65.2 | -22.7 | 32 | 49.2 | 50.1 |

## 2. $PM_{2.5}$ seasonality evolution

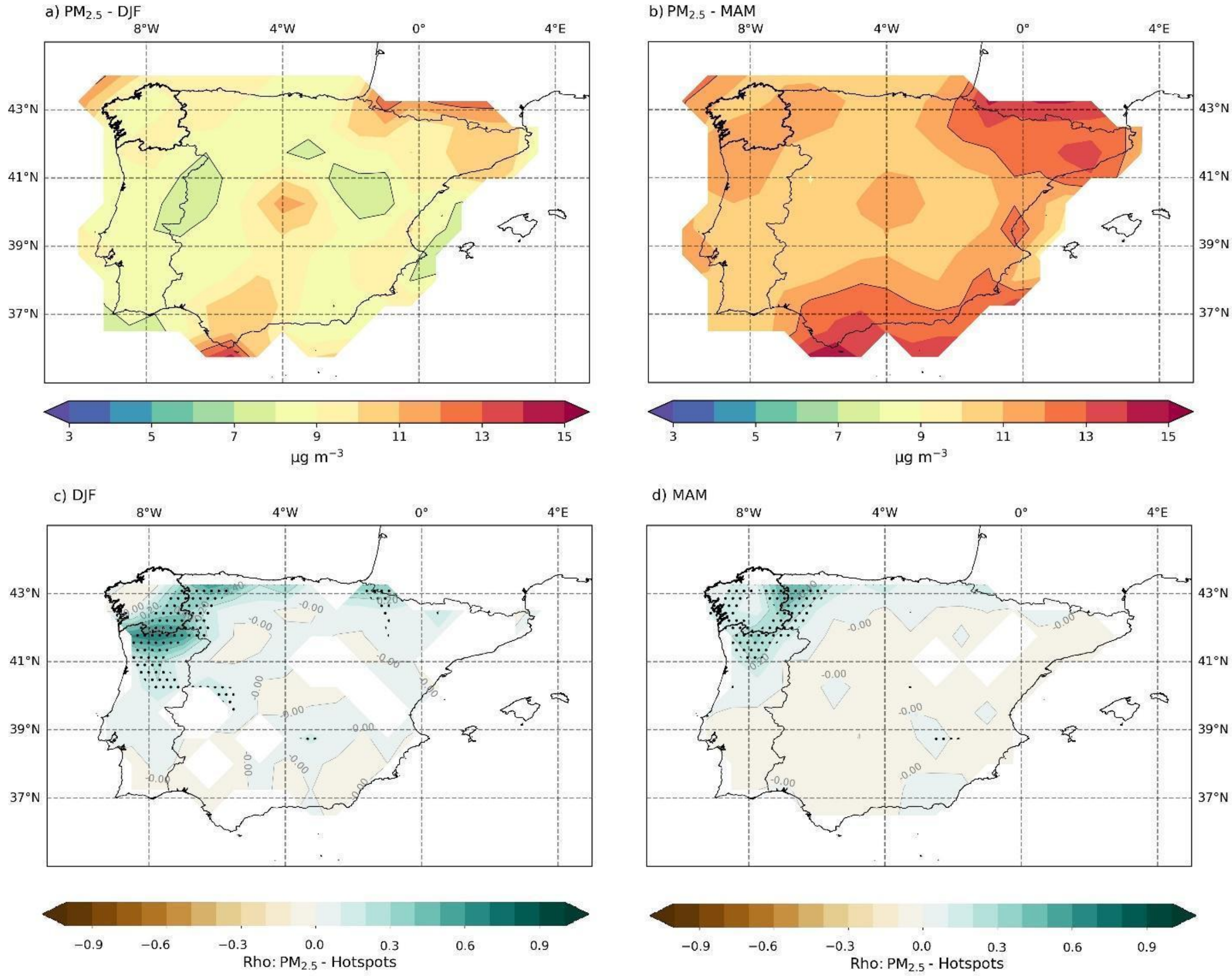


Figure SM1. Composites (seasonal average 2013-2022) of $PM_{2.5}$ pollution events captured by the Lagrangian tracker across the Iberian Peninsula for the DJF season (a) and MAM (b). Colors and contour lines represent the magnitude of each pollutant. Spearman correlation coefficient calculation (colors and contour lines) for DJF (c) and MAM (d) between seasonal averages of $PM_{2.5}$ and the number of hot temperature points (VIIRS) (Figure 1). Black dots represent locations with statistical significance ($p$-value < 0.05). The plot was constructed using the Cartopy Python package (Met Office, 2015).

## 3. Characteristics of pollution plumes develop from wildfires

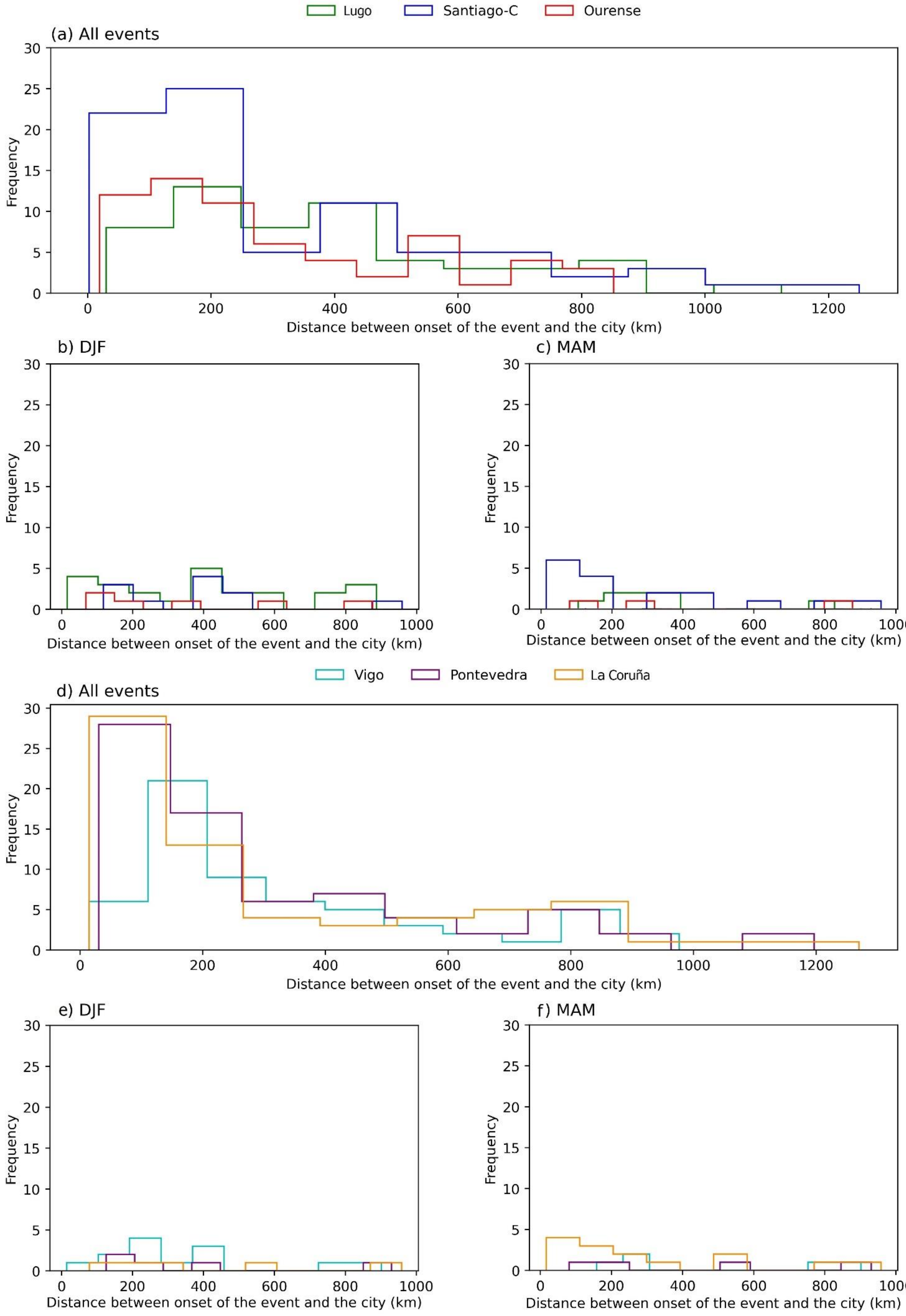
Lugo
Santiago-C
Ourense
(a) All events
Frequency
Distance between onset of the event and the city (km)
b) DJF
c) MAM
Vigo
Pontevedra
La Coruña
d) All events
e) DJF
f) MAM

Figure SM2. Composition, frequency of the mean geodesic distances between the start of the event and the cities (Pontevedra, Lugo and Ourense) for the annual average (a), for the winter season DJF (b) and the spring season MAM (c). In a similar way but for the cities of Vigo, La Coruña, Santiago-C, (d-f).

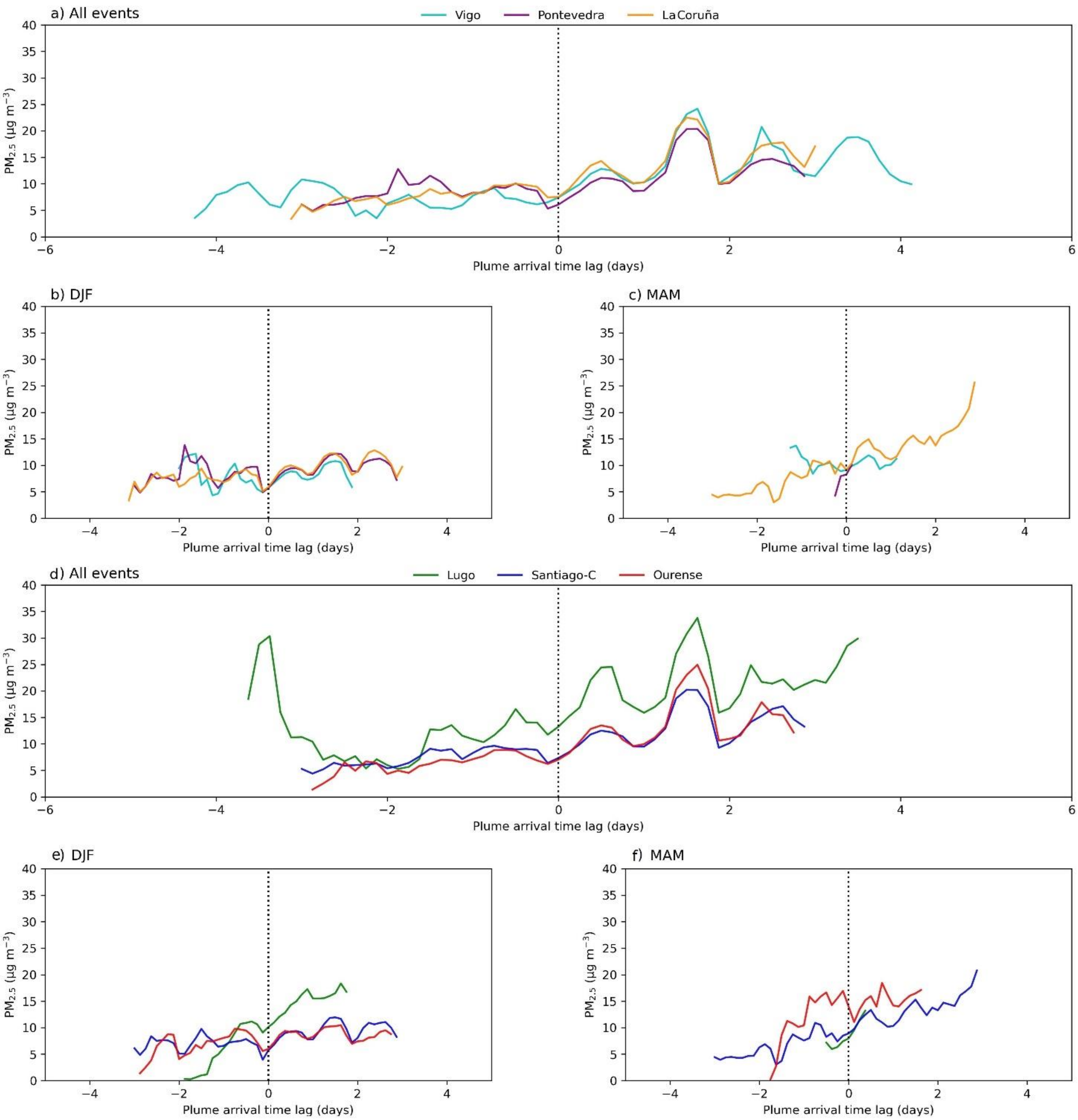


Figure SM3. Composition, $PM_{2.5}$ pollution columns (concentration) of the cities of Vigo, La Coruña, Santiago-C when the events occur during the annual average (a) the winter season (b) and the spring season (c). The same occurs with the cities of Pontevedra, Lugo and Ourense (d-f).

## 4. Characterization of land use types affected by wildfires in the NIP

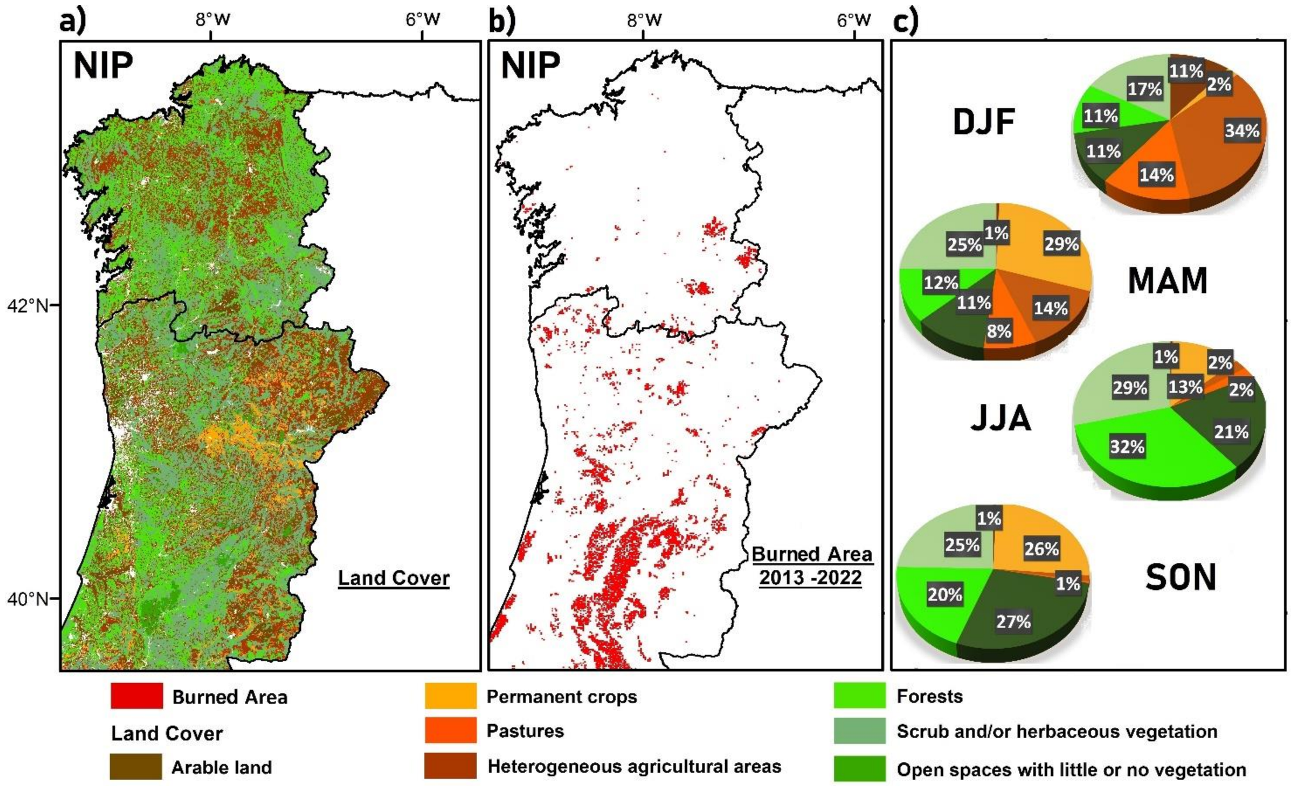


Figure SM4: (a) Map of CORINE land cover in the NIP region at the second level of land use type. (b) Burned areas in the NIP region for the years 2016-2022, showing the extent of the affected areas. (c) Percentage relationship of CORINE land cover types affected by forest fires across the seasons DJF, MAM, JJA, and SON.