Supplement of Nat. Hazards Earth Syst. Sci., 26, 1479–1513, 2026
https://doi.org/10.5194/nhess-26-1479-2026-supplement

Natural Hazards and Earth System Sciences
Open Access
EGU

Supplement of

What controls fire size in the South American Gran Chaco? Exploring atmospheric and landscape drivers through Remote Sensing

Rodrigo San Martín et al.

Correspondence to: Rodrigo San Martín (rodrigo.sanmartin@lsce.ipsl.fr)

S1. Supplementary Figures

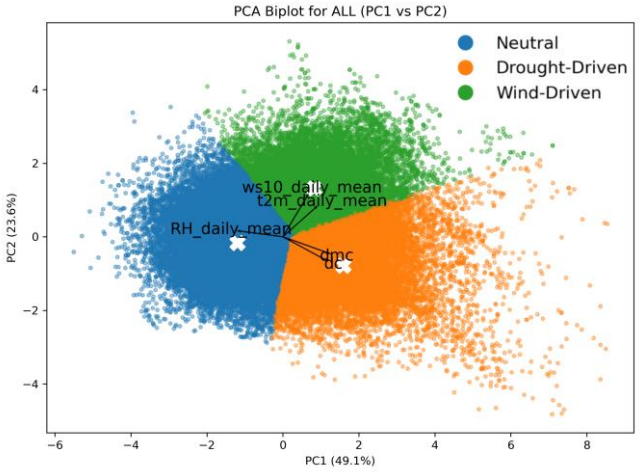

Fig. S1. Principal Component Analysis (PCA) biplot of pre-fire meteorological anomalies used for K-means clustering, showing the distribution of fire patches across the first two principal components (PC1 and PC2), which explain 49.1% and 23.6% of the total variance, respectively. The three clusters are color-coded and numbered as follows: Cluster 1 (blue) corresponds to Neutral conditions, Cluster 2 (orange) to Drought-Driven conditions (with high DC and DMC anomalies), and Cluster 3 (green) to Wind-Driven conditions (characterized by elevated wind speed and temperature anomalies). Arrows represent the contribution of the original variables to the PCA axes and the white crosses the cluster centroids. This ordination was used to guide the semantic naming of clusters.

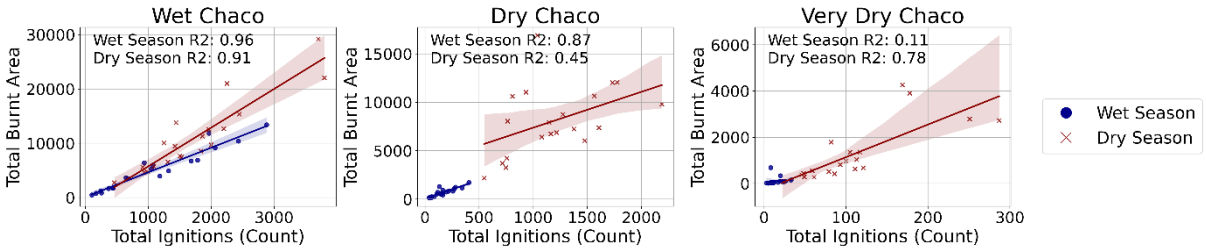

Fig. S2. Scatter plots and linear regressions between total annual BA and total annual ignitions between 2001 and 2022 in the Wet, Dry and Very Dry Chaco, divided into wet season fires (blue circles) and dry season fires (red crosses).

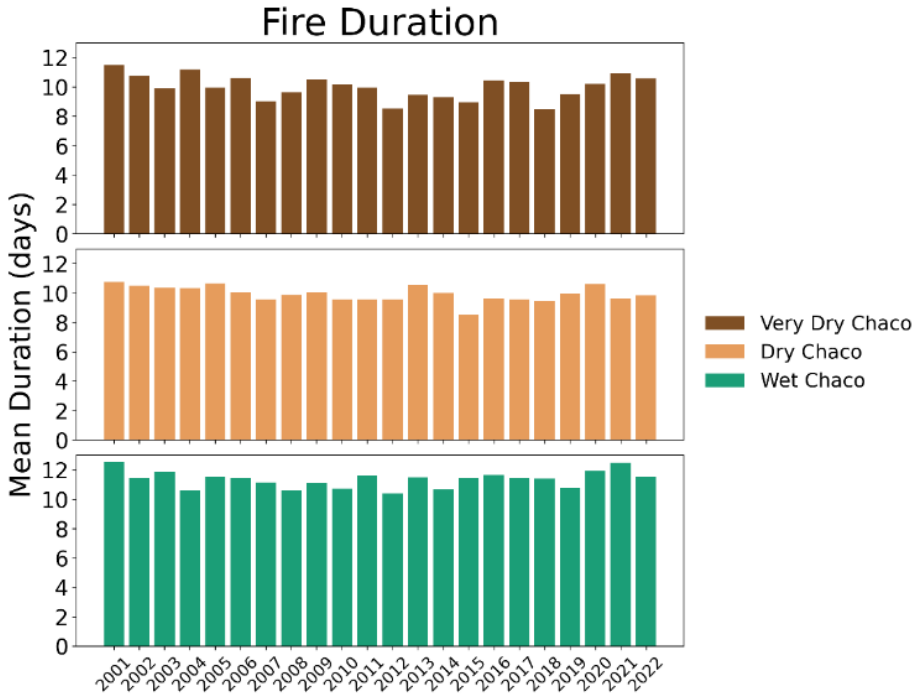

Fig. S3. Mean duration of individual fire patches between 2001 and 2022 in the Wet, Dry, and Very Dry Chaco, derived from FRYv2.0. Bars show annual averages for each subregion.

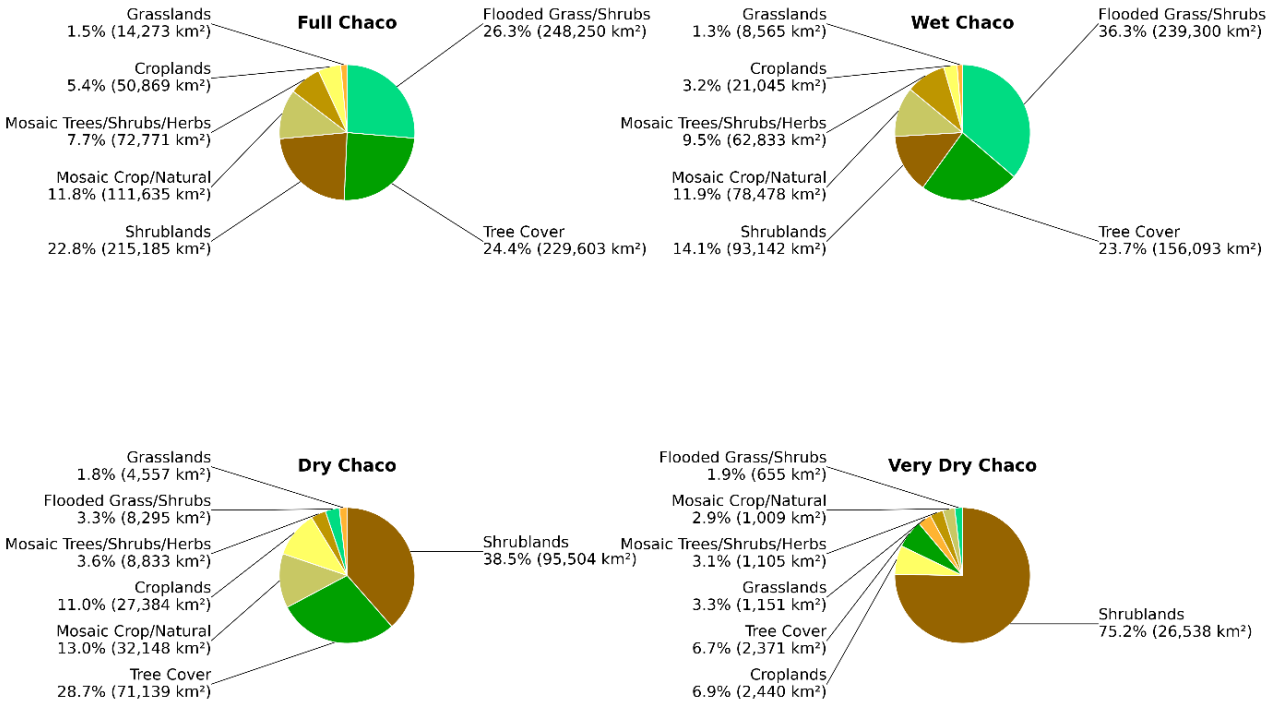

Fig. S4. Pie plots showing burnt land cover values within FRY fire polygons between 2001 and 2022 in the entire Gran Chaco and its subregions. Land cover derived from ESA CCI Medium Resolution Land Cover maps at 300 m.

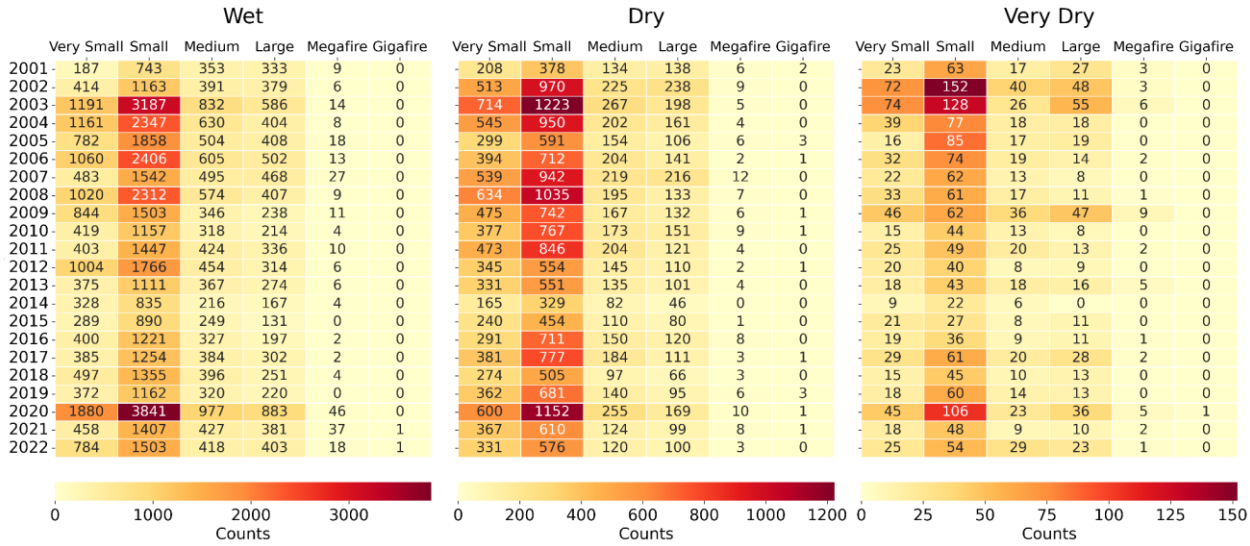

Fig. S5: Total counts of fire patches separated by size category between 2001 and 2022 in the Wet, Dry, and Very Dry Chaco.

Table S1. Number of fires detected by FRYv2.0 between 2001 and 2022 classified by fire size. WS: wet season; DS: dry season.

Region	**Very Small (0-1 km²)**		**Small (1-5 km²)**		**Medium (5-10 km²)**		**Large (10-100 km²)**		**Megafire (100-1000 km²)**		**Gigafire (> 1000 km²)**		**Total**
Season	**WS**	**DS**	**WS**	**DS**	**WS**	**DS**	**WS**	**DS**	**WS**	**DS**	**WS**	**DS**	
Wet	8414	6322	17,018	18,992	4340	5667	3264	4534	91	163	2	0	**68,807**
	14,736		**36,010**		**10,007**		**7,798**		**254**		**2**		
Dry	3526	5332	5754	10,302	1201	2485	841	1991	24	94	0	15	**31,565**
	8,858		**16,056**		**3,686**		**2,832**		**118**		**15**		
Very Dry	334	300	708	691	187	203	200	238	13	29	0	1	**2,904**
	634		**1,399**		**390**		**438**		**42**		**1**		
Total	**24,228**		**53,465**		**14,083**		**11,068**		**414**		**18**		**103,276**

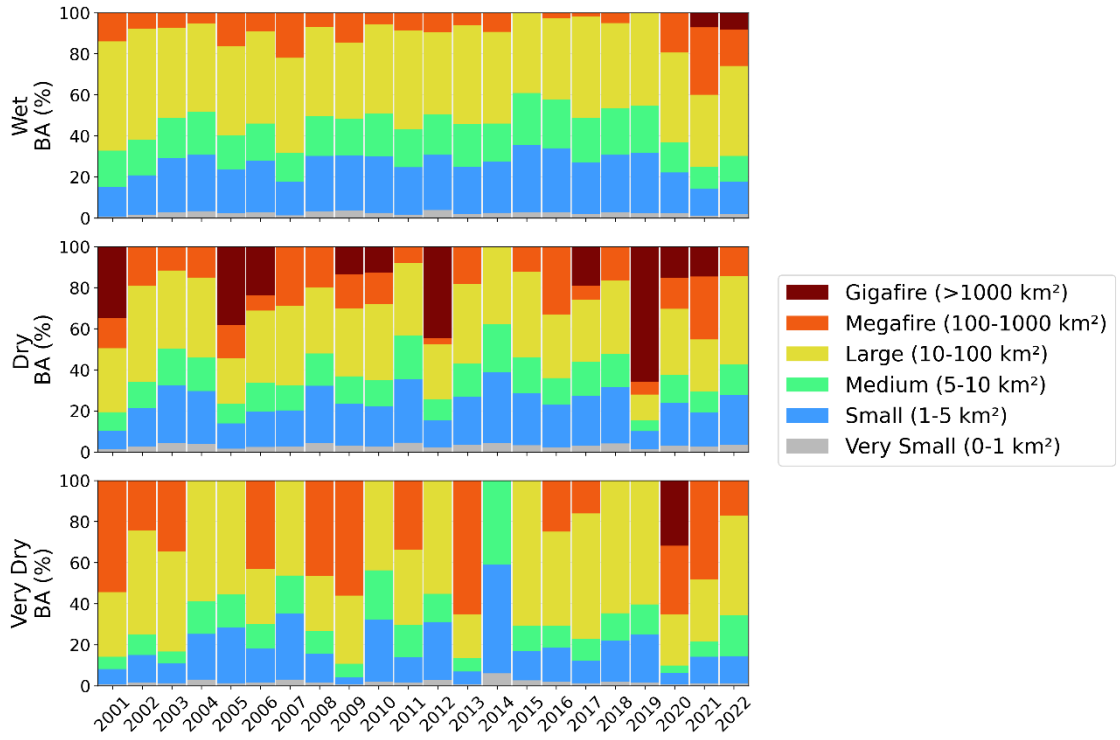

Fig. S6: Annual percentage distribution of burned areas across different size categories between 2001 and 2022 in the Wet, Dry, and Very Dry Chaco.

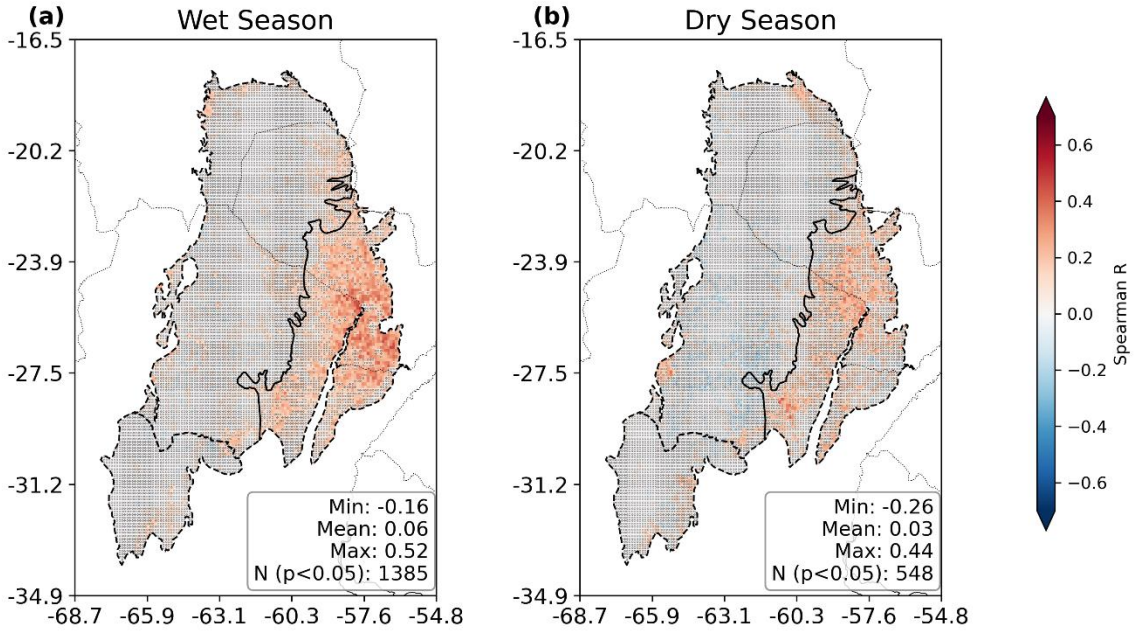

Fig. S7. Spatial distribution of pixel-wise Spearman correlation coefficients between monthly Fire Weather Index (FWI) anomalies and monthly burned area (BA) for the period 2001–2022: (a) Wet Season and (b) Dry Season. The color bar indicates the strength and direction of the correlation (from negative in blue to positive in red). Inset statistics summarize the distribution of coefficients (Min, Mean, Max). Pixels marked with small black circles represent non-significant correlations (p-value > 0.05), while unmarked pixels indicate significant correlations (p-value < 0.05). Only pixels with more than 3-time steps with burned area >0 were kept to avoid biased correlations related to very few or no fires.

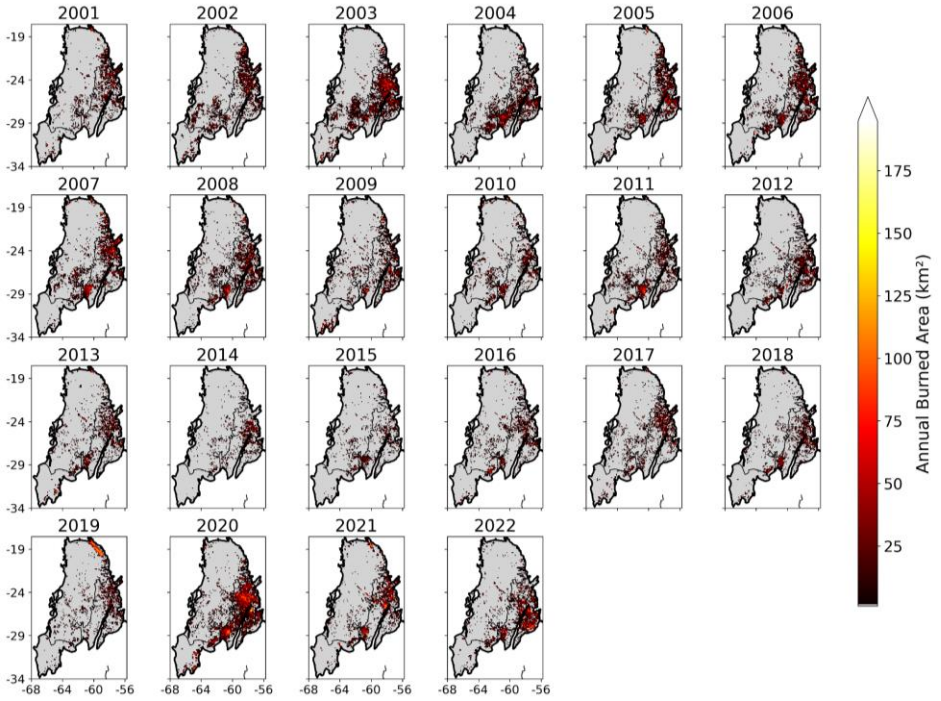

Fig. S8: Annual burned area maps of the Chaco region between 2001 and 2022. Burned areas extracted from FRYv2.0.

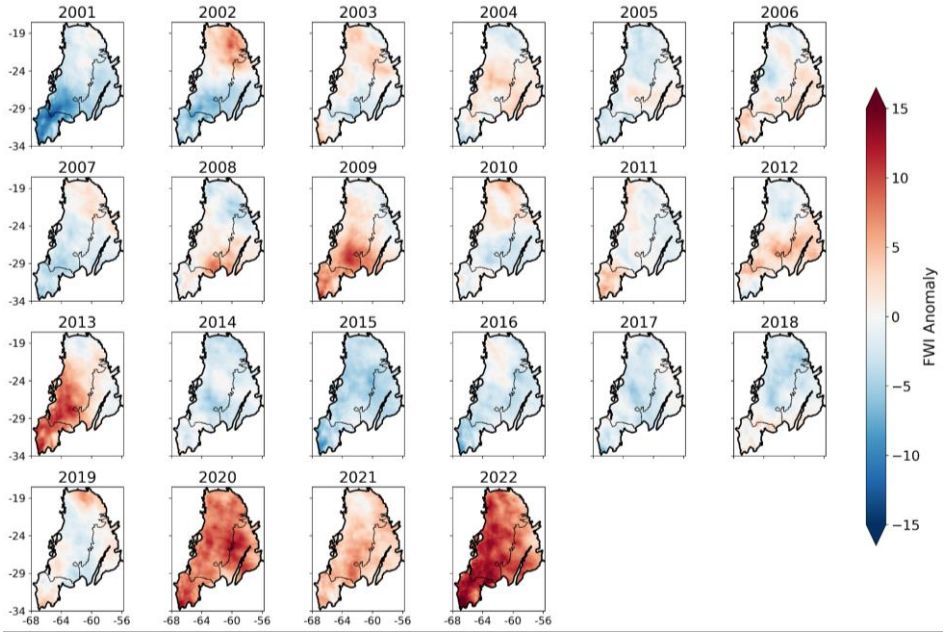

Fig. S9: Annual mean Fire Weather Index (FWI) anomalies with respect to the period 2001–2020, averaged for the Chaco region for each year between 2001 and 2022. FWI built from ERA5-Land.

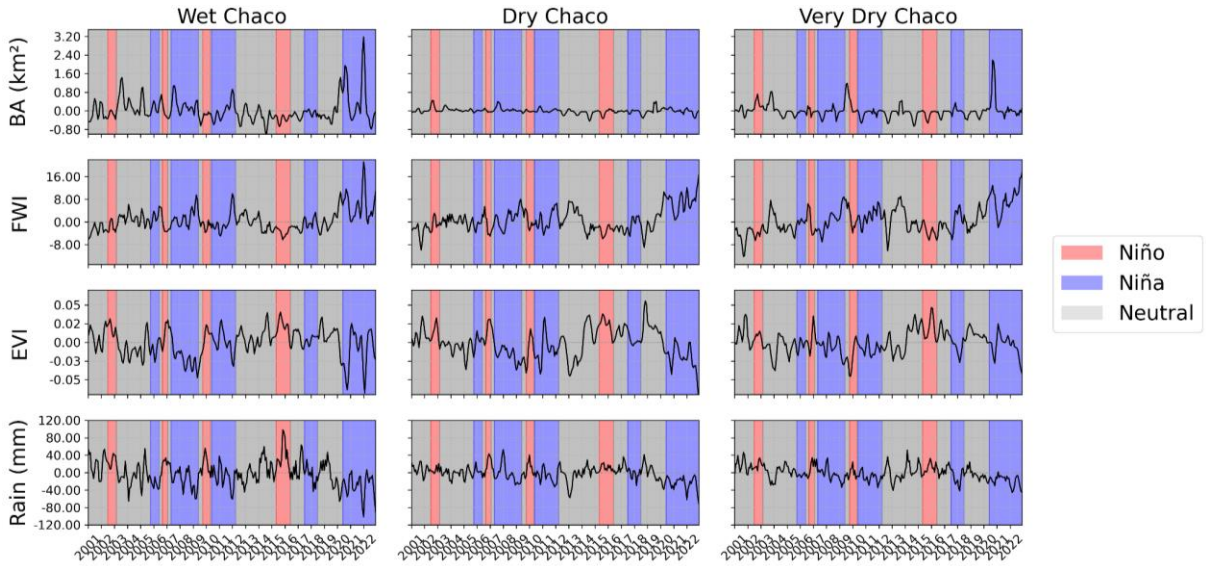

Fig. S10: Monthly anomalies of rainfall, vegetation (EVI), fuel dryness (FWI), and burned area in the Chaco subregions. Panels show 3-month running means of region-averaged anomalies for each variable, calculated from gridded (pixel-based) data and averaged over the Wet, Dry, and Very Dry Chaco subregions. Shaded backgrounds in the burned area panel indicate ENSO phases (red for El Niño, blue for La Niña), calculated with the Multivariate ENSO index (MEI).

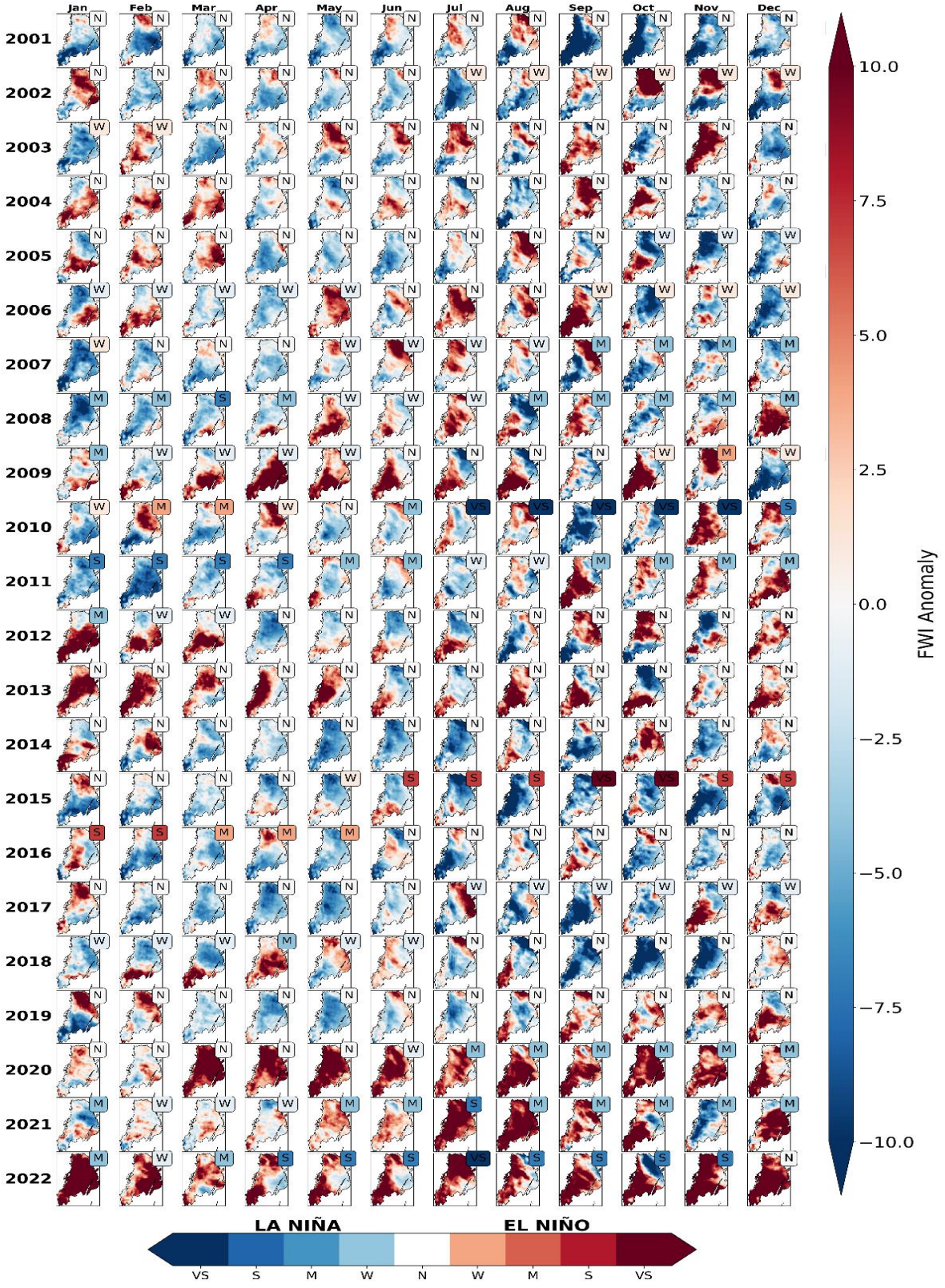

Fig. S11: The maps display the monthly anomalies (with 2001–2021 as the baseline) for the Chaco region for each year within the period. Additionally, each map counts with the Multivariate ENSO Index (MEI) showing the presence of an El Niño (EN; red) or La Niña (LN; blue) when during five consecutive three-month periods, MEI values are above +0.5 or below -0.5, respectively. Otherwise, the months are in a neutral (N) phase. The Niño/Niña events are classified by intensity based on the absolute MEI values. W: Weak (≥ 0.5); M: Moderate (≥ 1); S: Strong (≥ 1.5); VS: Very Strong (≥ 2).

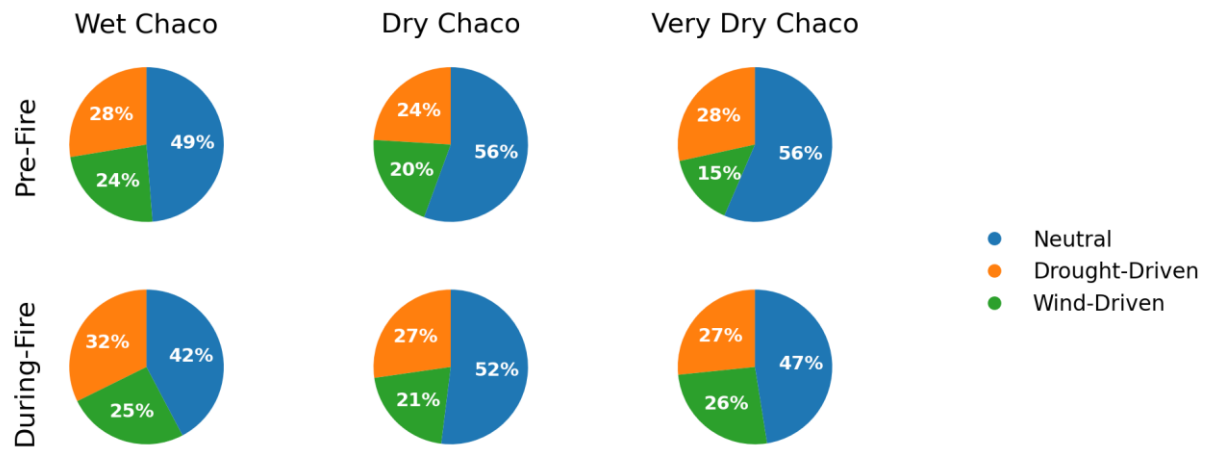

Fig. S12: Regional distribution of fire-weather types (FWTs) across the three Chaco subregions based on the Pre-Fire clustering (top row) and the During-Fire clustering (bottom row). Pie charts represent the proportion of fire patches assigned to each cluster—Drought-Driven (orange), Wind-Driven (green), and Neutral (blue)—based on pre-fire (0–3 days before ignition) and during-fire meteorological conditions.

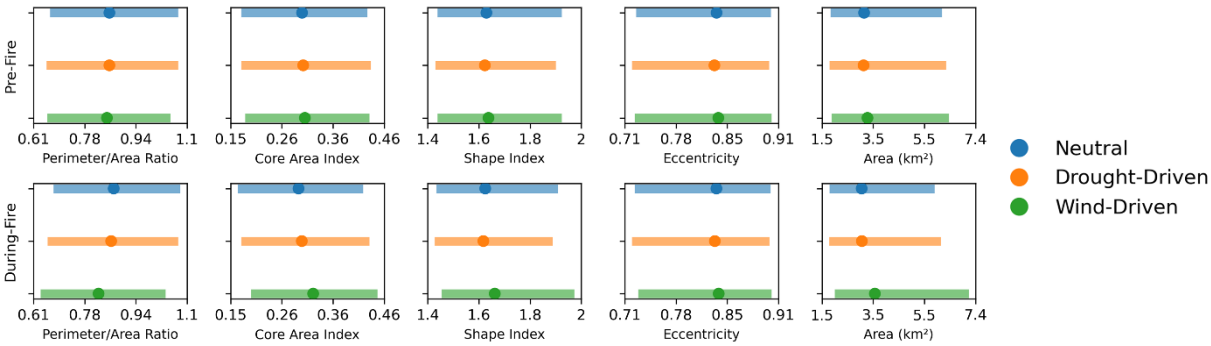

Fig. S13: Distribution of morphology variables by cluster (quartile–dot plots). For each morphology variable, the interquartile range (IQR; thick horizontal bar) and median (dot) are shown for each cluster, separately for Pre-Fire and During-Fire clustering (first and second rows, respectively). This visualizes the spread and central tendency of each variable within clusters, highlighting differences in fire patch morphology between cluster types and fire periods.

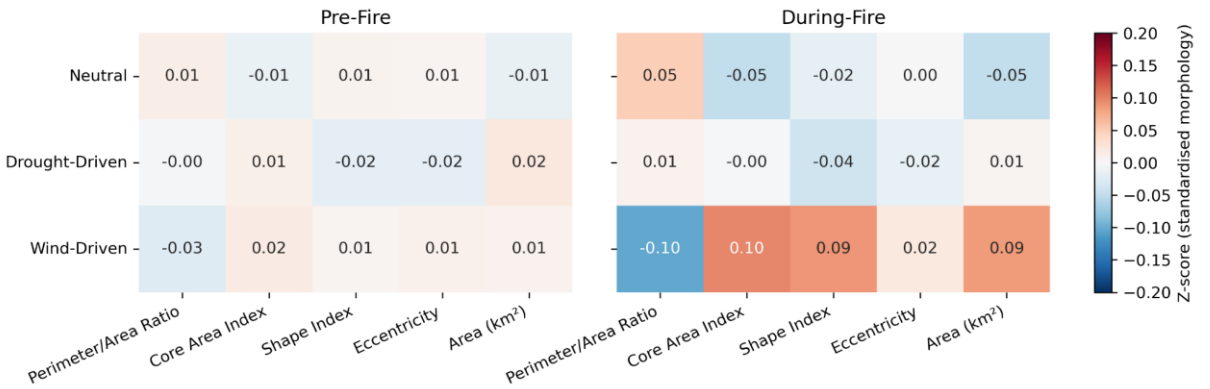

Fig. S14. Each heatmap shows the mean z-score (standardized value) of key fire patch morphology variables for each cluster, separately for Pre-Fire (left) and During-Fire (right) cluster assignments. Rows correspond to clusters (Neutral, Drought-Driven, Wind-Driven), and columns to morphology variables. The color scale indicates the relative position of each cluster's mean within the overall distribution, highlighting differences in fire patch shape and size between clusters and fire periods.

Table S2. Features ranges

Feature	Unit	Min	P5	Median	Mean	P95	Max	SD	% zeros
Fire size (target)	pixels	17	20	56	107.7	368	1874	160.8	0
Mean elevation	m	0	5	10	18.57	38	1921	61.27	0
Mean slope	%	0	0.48	0.77	1.19	1.91	63.41	2.83	0
LC evenness	–	0	0	0.72	0.61	0.98	1	0.33	16.4
LC entropy	–	0	0	0.68	0.66	1.33	1.84	0.43	16.4
Tree cover	%	0	0	0.1	0.23	0.88	1	0.29	37.9
Shrublands	%	0	0	0.08	0.24	0.93	1	0.31	41.1
Flooded areas	%	0	0	0	0.24	1	1	0.36	56.6
Trees/shrubs/herbs mosaics	%	0	0	0	0.08	0.43	1	0.16	61.2
Croplands/grasslands mosaics	%	0	0	0	0.12	0.64	1	0.22	58.1
Total precipitation during fire	mm	0	0	10.13	24.74	97.66	388.34	35.8	1.9
Maximum wind speed during fire	km h^{-1}	5.43	14.54	25.09	25.24	36.38	56.41	6.55	0
Extreme-wind days fraction	%	0	0	0.03	0.11	0.43	1	0.16	49.5
Extreme wind & direction index	–	0	0	0.01	0.07	0.26	1	0.11	49.5
LAI (previous growing season)	m^2/m^2	0.11	0.82	1.64	1.62	2.49	5.91	0.54	0
Population density	Persons/km^2	0	0.39	2.79	10.94	38.2	2388.6	39.72	0
Cattle density	Heads/km^2	0	437.5	2349.84	2515.59	5308.72	17657.53	1545.49	0
Road density (OSM)	km/km^2	0	0	0.18	0.3	0.94	18.4	0.51	19.9
Road density (MS)	km/km^2	0	0	0.21	0.34	1.04	12.73	0.47	9

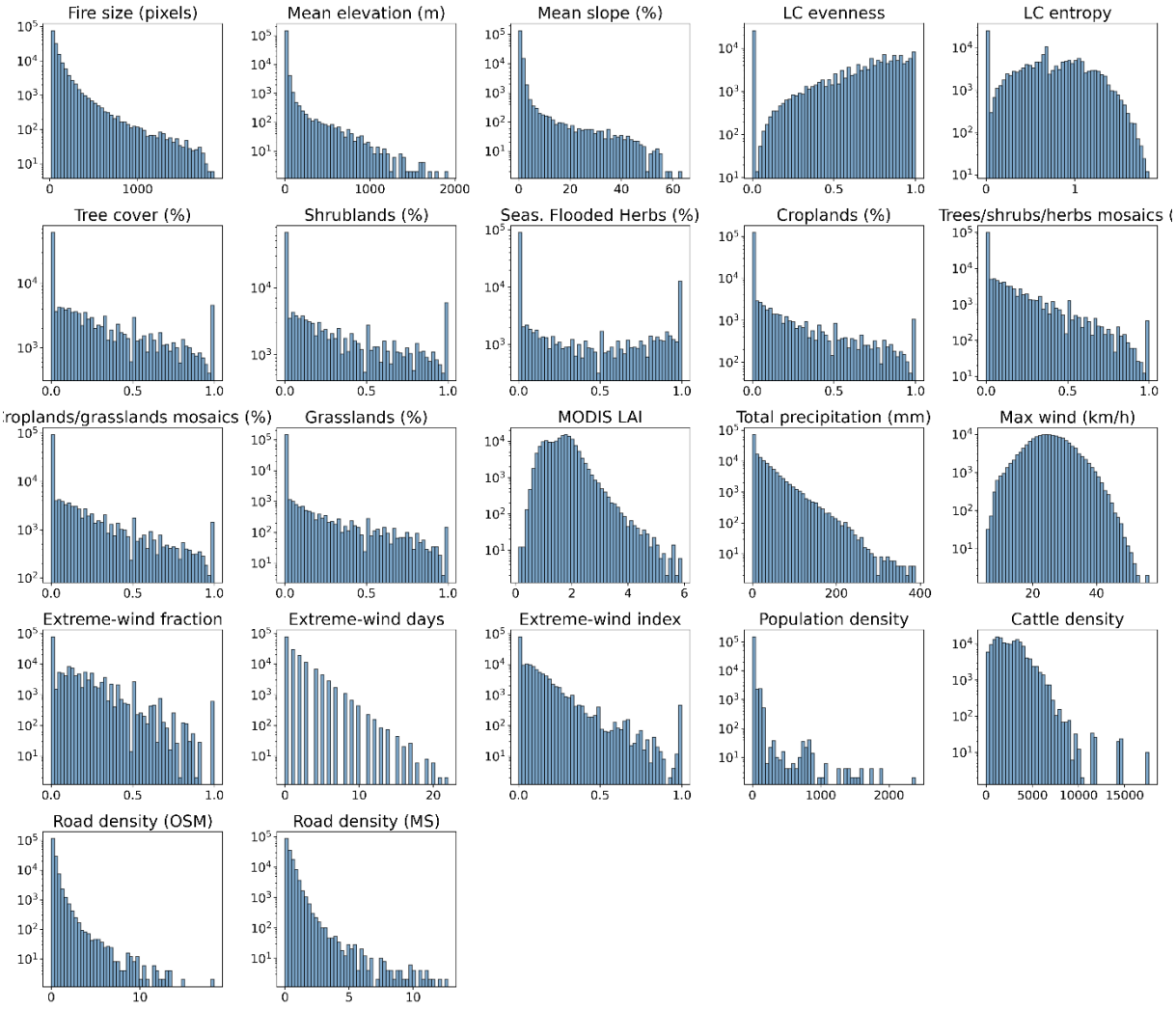

Fig. S15. Distributions of all features used in the Random Forest analysis, shown on a logarithmic y-axis to emphasize skewness and the prevalence of zero-inflated or heavy-tailed patterns. Each panel displays the empirical distribution of one predictor across all fire patches in the 1–100 km² range.

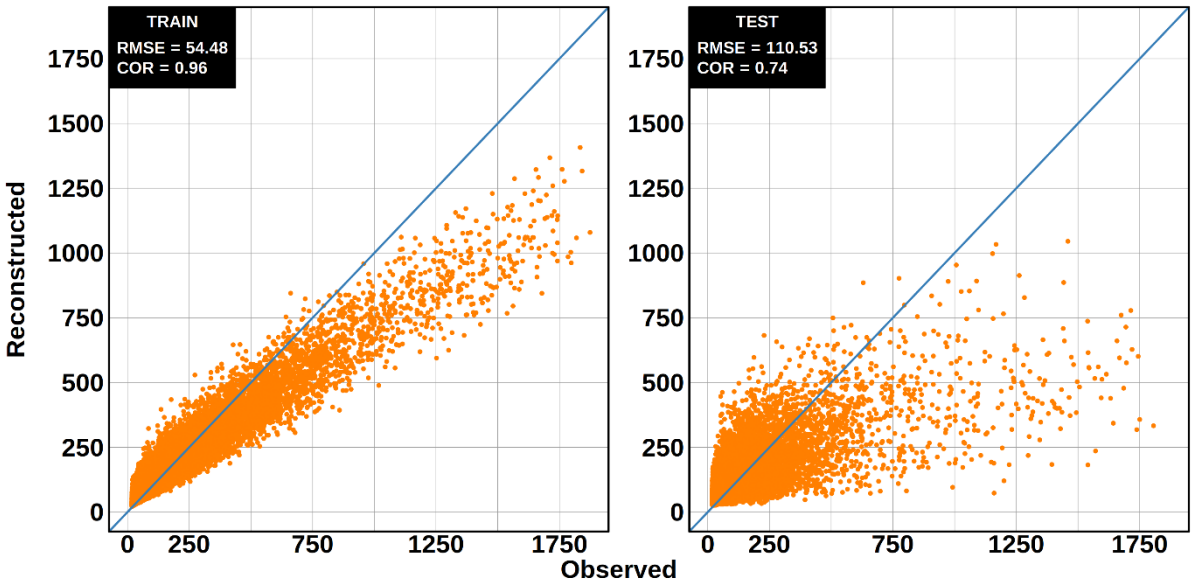

Fig. S16. Random Forest metrics for the Full Gran Chaco model.

Table S3. Random Forest metrics

Model Name	Domain	Configuration	Train		Test	
			COR	RMSE	COR	RMSE
Full Chaco	Full Gran Chaco	Full set	0.96	54.62	0.74	110.09
Wet Chaco	Wet Chaco	Regional	0.96	54.48	0.74	110.77
Dry Chaco	Dry Chaco	Regional	0.96	51.46	0.72	104.09
Very Dry Chaco	Very Dry Chaco	Regional	0.97	81.92	0.73	156.26
Wet Season	Full Gran Chaco (season)	Seasonal	0.96	55.35	0.71	108.57
Dry Season	Full Gran Chaco (season)	Seasonal	0.96	56.56	0.73	115.51
Pre-Wind-Driven	Full Gran Chaco (clusters)	FWT cluster (Pre)	0.96	56.79	0.71	110.44
Pre-Drought-Driven	Full Gran Chaco (clusters)	FWT cluster (Pre)	0.96	61.18	0.71	126.17
Pre-Neutral	Full Gran Chaco (clusters)	FWT cluster (Pre)	0.96	55.42	0.72	110.28
During-Wind-Driven	Full Gran Chaco (clusters)	FWT cluster (During)	0.96	60.93	0.73	123.67
During-Drought-Driven	Full Gran Chaco (clusters)	FWT cluster (During)	0.96	60.53	0.7	123.26
During-Neutral	Full Gran Chaco (clusters)	FWT cluster (During)	0.96	52.74	0.73	104.12
No Topography	Full Gran Chaco	No Topography	0.95	60.3	0.67	119.05
MS Roads	Full Gran Chaco	MS Roads	0.96	54.82	0.74	110.72

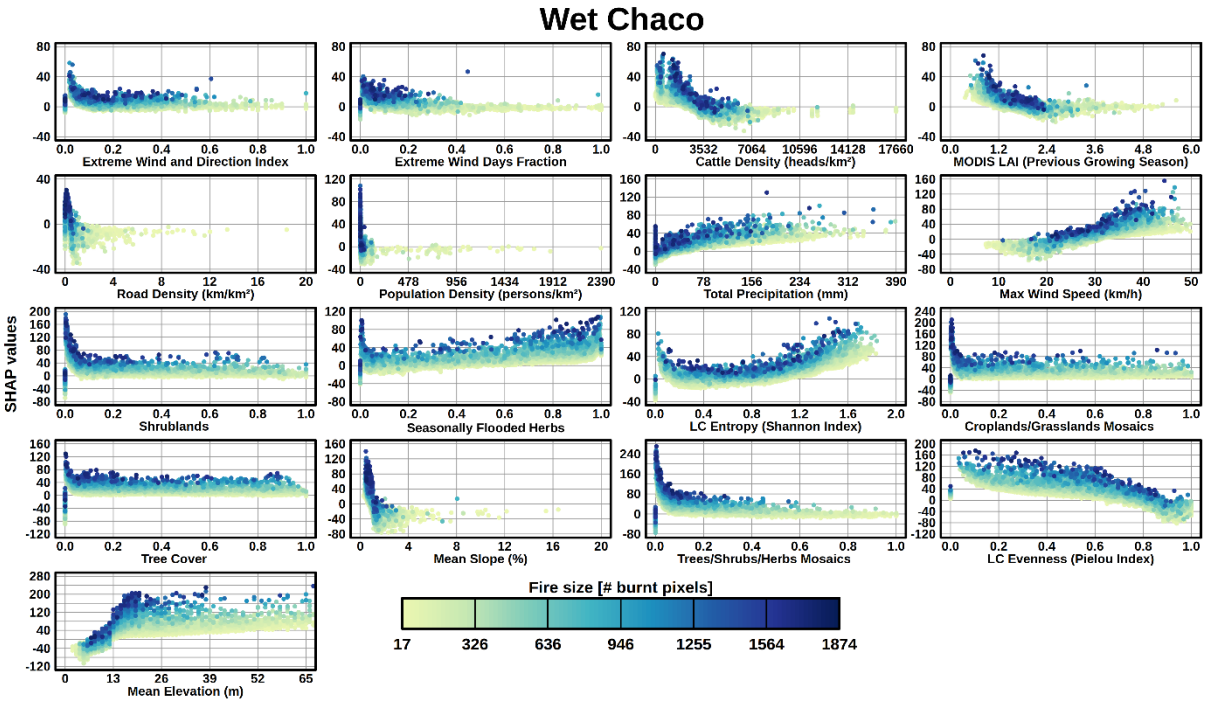

Fig. S17. SHAP dependence plots for the 17 explanatory variables used to predict fire patch size (n_cell) with the Random Forest model trained on fire patches between 1 km² and 100 km² in the Wet Chaco between 2001 and 2022. Each panel shows the SHAP value (y-axis) across the range between 0 and the 0.995 quantile of a given feature (x-axis), illustrating the marginal effect of that feature on the model's output. Dots are colored by fire size (number of burned pixels), with darker tones indicating larger fires. Land cover classes represented as fractions.

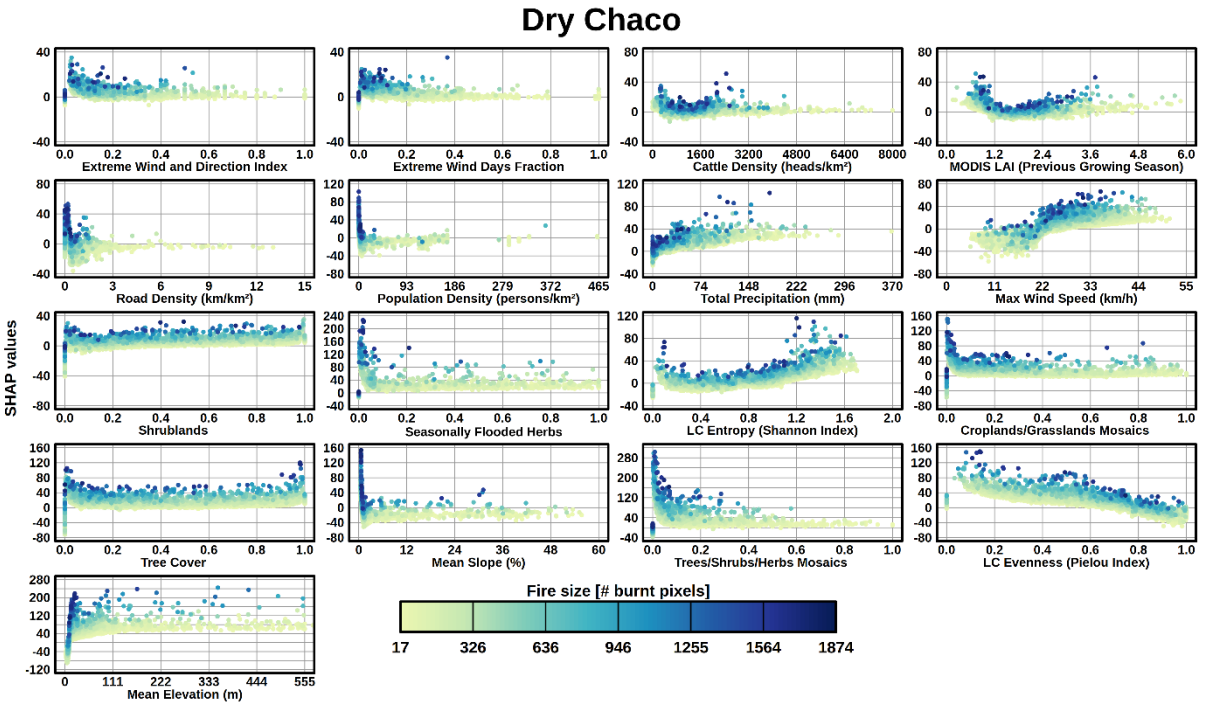

Fig. S18. SHAP dependence plots for the 17 explanatory variables used to predict fire patch size (n_cell) with the Random Forest model trained on fire patches between 1 km² and 100 km² in the Dry Chaco between 2001 and 2022. Each panel shows the SHAP value (y-axis) across the range between 0 and the 0.995 quantile of a given feature (x-axis), illustrating the marginal effect of that feature on the model's output. Dots are colored by fire size (number of burned pixels), with darker tones indicating larger fires. Land cover classes represented as fractions.

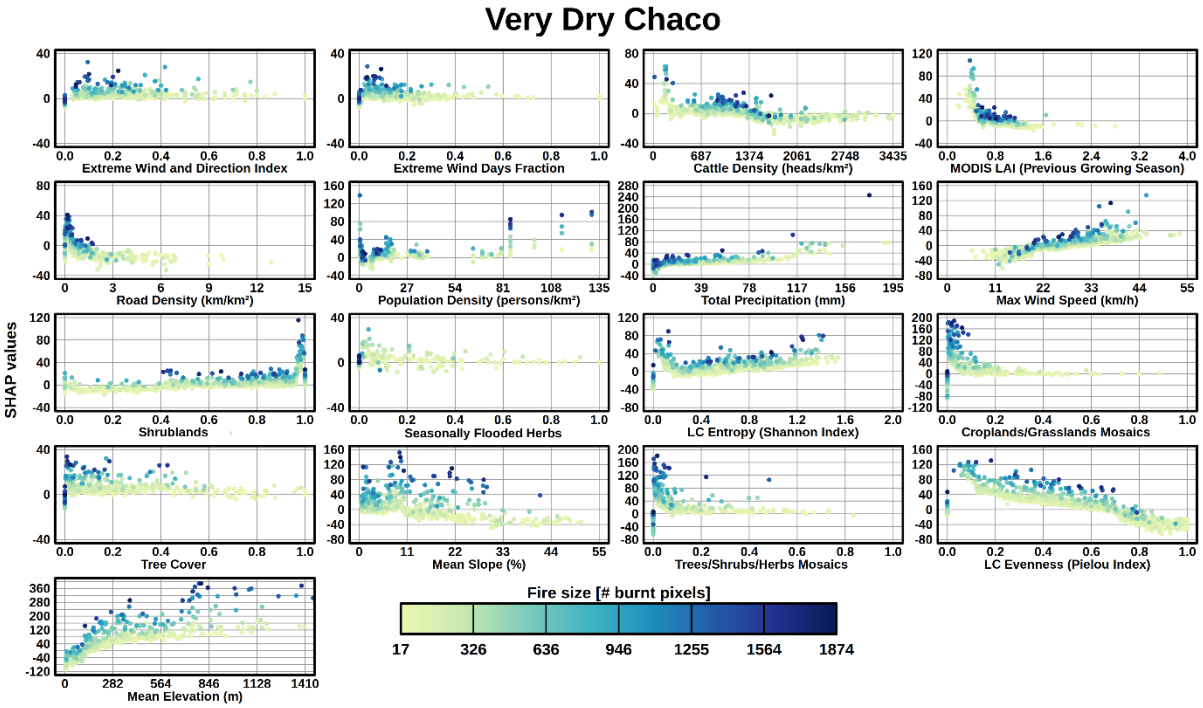

Fig. S19. SHAP dependence plots for the 17 explanatory variables used to predict fire patch size (n_cell) with the Random Forest model trained on fire patches between 1 km² and 100 km² in the Very Dry Chaco between 2001 and 2022. Each panel shows the SHAP value (y-axis) across the range between 0 and the 0.995 quantile of a given feature (x-axis), illustrating the marginal effect of that feature on the model's output. Dots are colored by fire size (number of burned pixels), with darker tones indicating larger fires. Land cover classes represented as fractions.

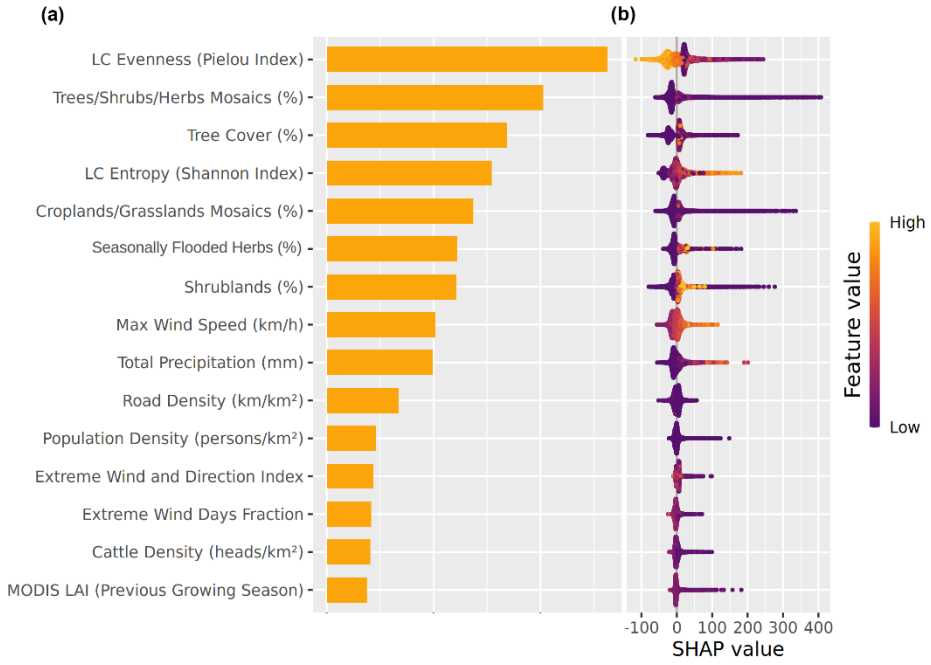

Fig. S20. Feature importance ranking for the sensitivity test Random Forest model predicting fire patch (FP) size across the entire Gran Chaco without topographic features (mean elevation and mean slope). (a) shows the average importance of each variable, expressed as the mean absolute SHAP value. (b) shows the SHAP values for all individual fire patches, indicating how low (purple) or high (yellow) feature values influence the prediction toward smaller or larger fires.

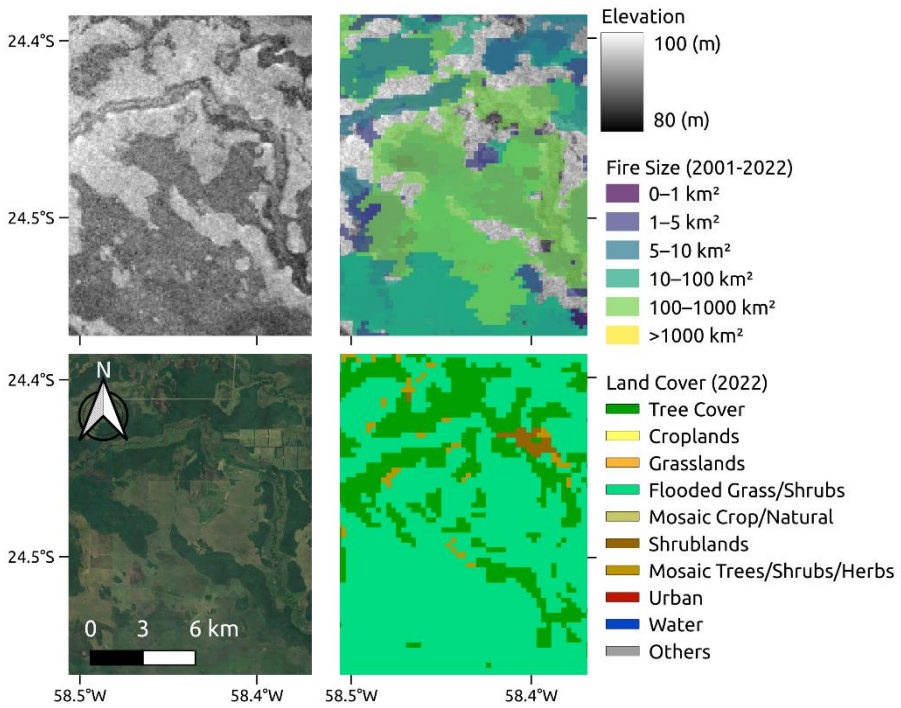

Fig. S21. Topographic and land-cover context of a representative sector of the Wet Chaco. The upper left panel shows the Shuttle Radar Topography Mission (SRTM) digital elevation model at 90 m resolution. The upper right panel displays the same elevation surface overlaid with FRY v2.0 fire polygons (2001–2022) colored by fire-size class. The lower left panel presents Google satellite imagery of the area (Map data © 2025 Google), and the lower right panel shows the ESA CCI Moderate-Resolution Land Cover (MRLC) map for 2022 (300 m).

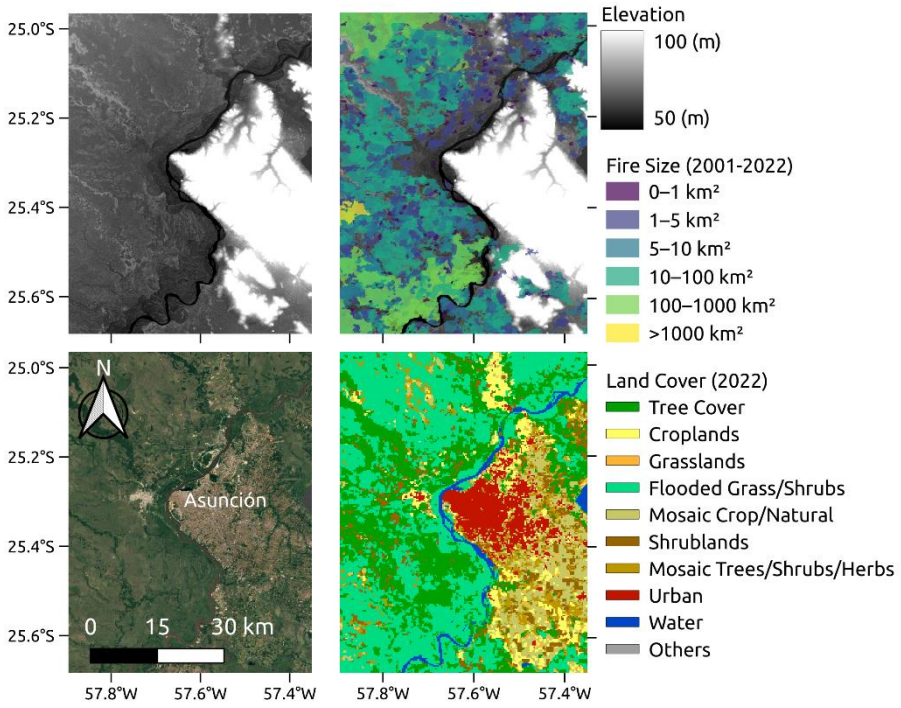

Fig. S22. Topographic and land-cover context over the city of Asunción (Paraguay's capital city). The upper left panel shows the Shuttle Radar Topography Mission (SRTM) digital elevation model at 90 m resolution. The upper right panel displays the same elevation surface overlaid with FRY v2.0 fire polygons (2001–2022) colored by fire-size class. The lower left panel presents Google satellite imagery of the area (Map data © 2025 Google), and the lower right panel shows the ESA CCI Moderate-Resolution Land Cover (MRLC) map for 2022 (300 m).

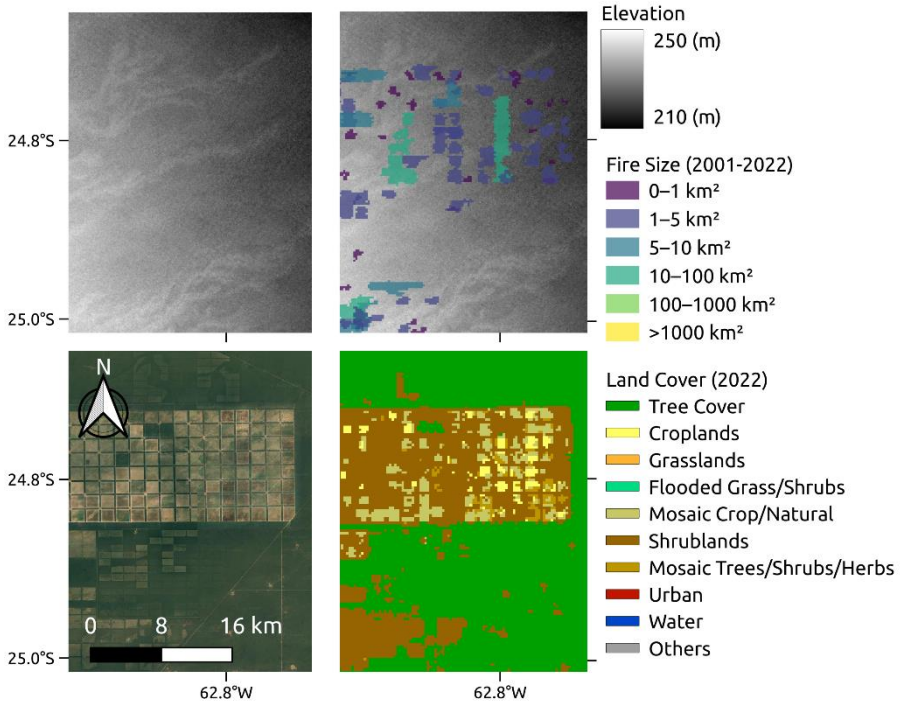

Fig. S23. Topographic and land-cover context over a deforested area within the Gran Chaco forest in Argentina. The upper left panel shows the Shuttle Radar Topography Mission (SRTM) digital elevation model at 90 m resolution. The upper right panel displays the same elevation surface overlaid with FRY v2.0 fire polygons (2001–2022) colored by fire-size class. The lower left panel presents Google satellite imagery of the area (Map data © 2025 Google), and the lower right panel shows the ESA CCI Moderate-Resolution Land Cover (MRLC) map for 2022 (300 m).

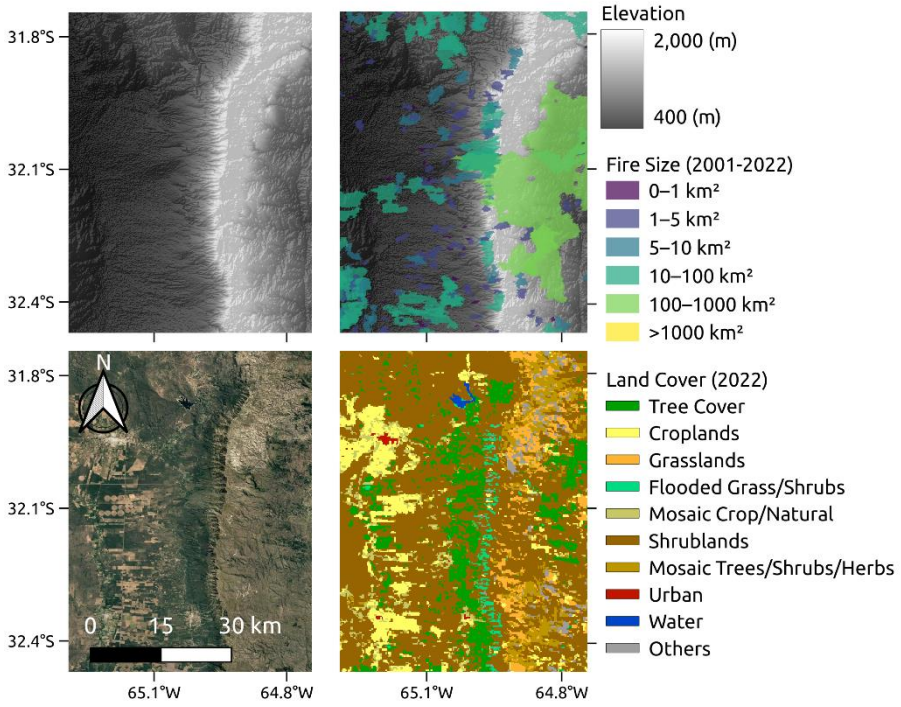

Fig. S24. Topographic and land-cover context over the “Sierras de Córdoba” in the Very Dry Chaco (Argentina). The upper left panel shows the Shuttle Radar Topography Mission (SRTM) digital elevation model at 90 m resolution. The upper right panel displays the same elevation surface overlaid with FRY v2.0 fire polygons (2001–2022) colored by fire-size class. The lower left panel presents Google satellite imagery of the area (Map data © 2025 Google), and the lower right panel shows the ESA CCI Moderate-Resolution Land Cover (MRLC) map for 2022 (300 m).

S2. Structural comparison of OSM and MS road density datasets

Before evaluating the effect of substituting OSM with MS road density in the RF models, we quantified how the two datasets differ across the landscape. OSM, being community-curated, captures paved and unpaved major roads reliably but tends to underrepresent informal, secondary, and seasonal tracks, particularly in remote ranching landscapes and sparsely populated areas of the western Dry and Very Dry Chaco. In our dataset, OSM identified 1.12 million km of linear features, whereas MS mapped 1.95 million km, an increase of roughly 74% in total detected road length.

The MS product, generated through automated detection from high-resolution imagery, identifies a much larger set of linear features: about 45–60% of MS segments have no corresponding OSM segment within 50 m, and a large fraction consist of faint dirt tracks, fencing lines, internal ranch access paths, and grid-aligned extraction routes. These additional features increase the apparent density of minor routes in areas where OSM shows limited coverage. As a consequence, MS expands the low-to-moderate density classes (0.1–2 km km^{-2}) and suppresses the long extreme tail observed in OSM.

These differences are spatially heterogeneous. In the Wet Chaco, correlations between datasets are high (0.85–0.90), whereas they decrease in the Dry (0.80–0.83) and Very Dry (0.65–0.72) Chaco. MS produces a smoother and more homogeneous density surface, while OSM remains highly skewed with extensive zero-density areas and occasional extreme values. Despite these differences, both datasets reproduce the broad east–west gradient in accessibility and exhibit comparable median densities.

Because road density interacts with landscape heterogeneity, we also examined differences across land-cover classes and structural indices. OSM density peaks in mosaic-rich and cropland–grassland transition landscapes, while MS spreads low-to-moderate values more broadly across shrublands and semi-open dry forests. These structural contrasts help explain why MS modifies the spatial distribution of road density without substantially altering model outcomes.