JSC370 Final Project: Global Natural Disaster Patterns from NASA EONET

Xucheng (Rolland) He, 1008905909

Table of contents

1 Introduction

Natural disasters such as wildfires, severe storms, floods, and volcanic eruptions are major sources of social and ecological risk. Understanding their temporal and spatial patterns can support hazard monitoring and policy planning. However, globally consistent event records are often fragmented across agencies.

This project uses NASA's Earth Observatory Natural Event Tracker (EONET) API to study global disaster observations in a unified framework. Compared with the midterm report, this final report keeps the same dataset and core topic but adds stronger statistical and modeling components, with clearer interpretation and limitations.

The final research questions are:

Main Research Question: How do global natural disaster observations recorded by NASA EONET vary across category, time, seasonality, and space, and how does predictive performance change when the target is defined at two scopes (full multi-category counts vs wildfire-only counts)?

Supporting Questions:

1. Which disaster categories dominate the observed global records, and how imbalanced is the dataset?
2. How have disaster observations changed over time, both overall and by category?
3. Is seasonality statistically different across categories?
4. Can simple benchmark models improve out-of-sample monthly prediction relative to a seasonal baseline?
5. Does model ranking or forecast quality change when inference is restricted to the wildfire-only subset?

Because the dataset is heavily dominated by wildfire observations, the primary inferential emphasis in this report is on wildfire-driven dynamics. Other categories are retained for descriptive context and imbalance diagnostics.

2 Methods

2.1 Data acquisition and wrangling

The data source is the NASA EONET v3 endpoint: `https://eonet.gsfc.nasa.gov/api/v3/events`. The API is paginated and does not require an API key. To ensure reproducibility and avoid repeated network calls during rendering, the flattened dataset is cached as `data/eonet_events.csv`.

2.2 Analytical workflow

The analysis has four layers:

1. **Descriptive analysis**: category shares, annual trends, monthly seasonality, and spatial concentration.
2. **Formal statistical testing**:
   - chi-square test of independence for `category × month`;
   - linear trend regression of annual counts on year.
3. **Predictive modeling** (new in final report): monthly count prediction by category.
4. **Evaluation**: out-of-sample metrics (MAE, RMSE, R2) on a strict time split.

This combination is appropriate for the project goals. Descriptive summaries answer the distribution questions directly and transparently. The chi-square test is suitable because both `category` and `month` are categorical variables and the question concerns dependence in a contingency table. The annual linear trend model is a parsimonious first diagnostic of long-run direction, while the train/test predictive comparison addresses whether observed temporal patterns carry usable forecasting signal rather than only retrospective structure.

Given severe class imbalance, this workflow is designed so that inferential claims are interpreted primarily through the dominant wildfire signal, while low-volume categories are used more cautiously for comparative pattern checks rather than strong standalone inference.

For modeling, observations are aggregated to monthly category counts:

$$Y_{t,c} = \#\{\text{observations in year-month } t \text{ and category } c\}.$$

Two models are compared:

- **Seasonal baseline**: predicts each category-month pair using its historical mean from training data.
- **Linear model (OLS)**: uses `year`, cyclic month features (`sin`, `cos`), and category one-hot indicators.

The OLS specification is:

$$Y_{t,c} = \beta_0 + \beta_1 \,\text{year}_t + \beta_2 \sin\left(\frac{2\pi m_t}{12}\right) + \beta_3 \cos\left(\frac{2\pi m_t}{12}\right) + \sum_{k=2}^{K} \gamma_k \,\mathbf{1}(c = k) + \varepsilon_{t,c},$$

where (m_t) is calendar month and category indicators use one reference category.

Training/testing protocol:

- Training: years `<= 2023`
- Testing: years `>= 2024` (strictly future data)

This split is appropriate because it enforces temporal ordering, so every test prediction uses only information that would have been available before 2024.

The seasonal baseline is intentionally simple and strong: if no model can consistently outperform historical category-month averages, then added model complexity is not yet justified. OLS is used as an interpretable extension of that baseline, allowing us to test whether linear time drift and smooth seasonal structure improve generalization.

As a nonlinear extension motivated by course material on smooth modeling, we additionally fit a GAM that replaces the linear year effect with a smooth term while preserving seasonal and category structure. We also add a lightweight XGBoost benchmark to test whether a flexible nonlinear learner materially improves out-of-sample prediction under the same temporal split.

For extended predictive experiments, we construct additional category-level time-series features, including month index, lagged counts, and short rolling means.

3 Results

3.1 Category dominance and imbalance

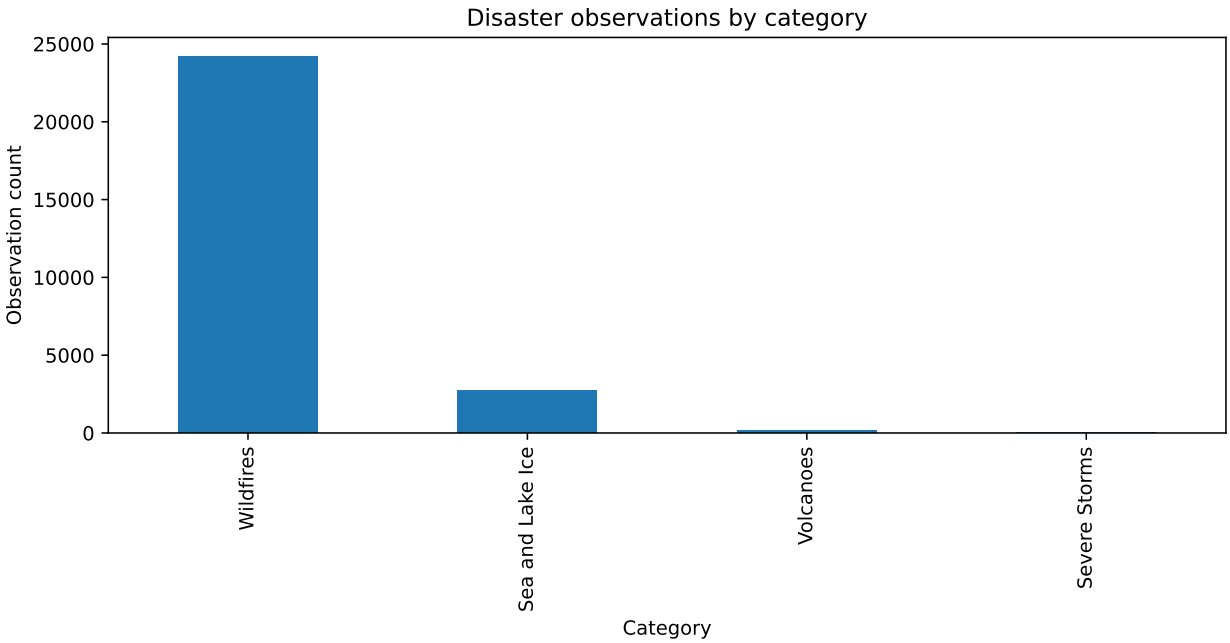

Figure 1: Distribution of EONET disaster observations by category.

The dataset contains 27,180 georeferenced observations. The largest class is Wildfires with 24,208 observations, representing 89.1% of all records. This concentration indicates that any aggregate result can be strongly influenced by wildfire dynamics.

Table 1: Category-level observation counts and shares.

	Category	Observations	Share
0	Wildfires	24208	89.1%
1	Sea and Lake Ice	2752	10.1%
2	Volcanoes	176	0.6%
3	Severe Storms	44	0.2%

3.2 Temporal trend and category-specific trajectories

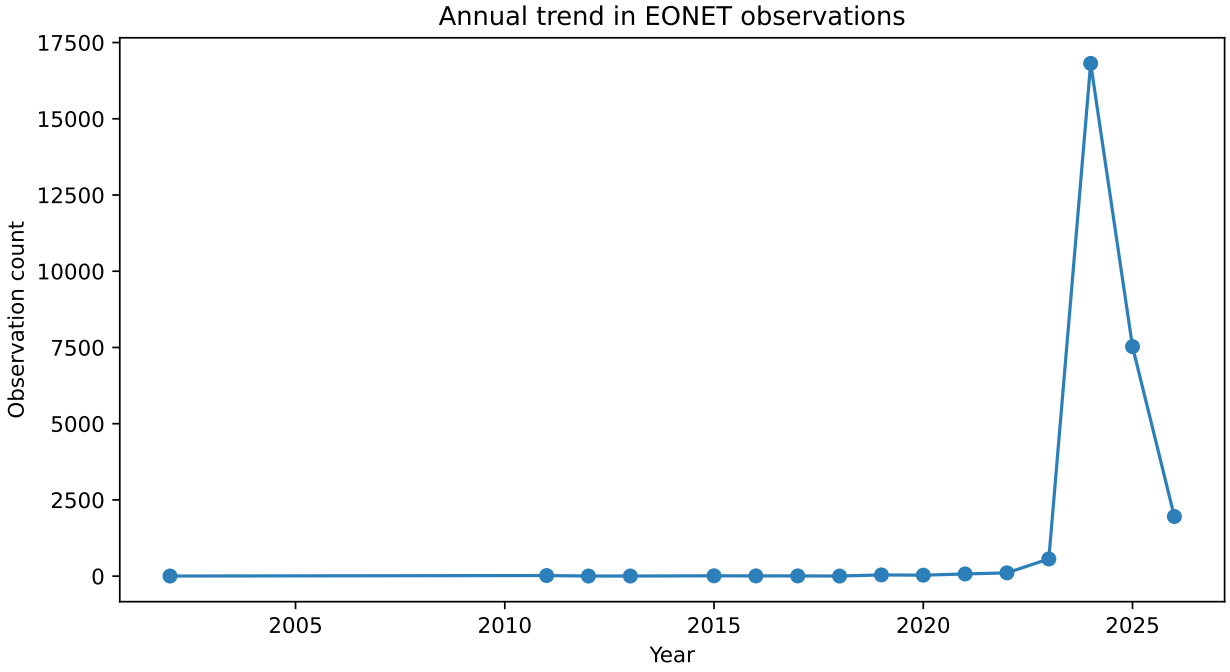

Figure 2: Annual total number of EONET disaster observations.

Overall counts rise sharply in the early 2020s, with a peak of 16,816 in 2024. The visual trend suggests a substantial recent increase, but interpretation should account for potential changes in monitoring intensity and reporting infrastructure.

Substantively, this figure is best interpreted as a change in *recorded observations* rather than a direct estimate of physical hazard frequency. In an operational monitoring system such as EONET, improved sensing coverage, reporting pipelines, and event tracking practices can increase counts even when underlying hazard processes are unchanged. This is why the trend is paired with formal inference and category-stratified diagnostics later in the report.

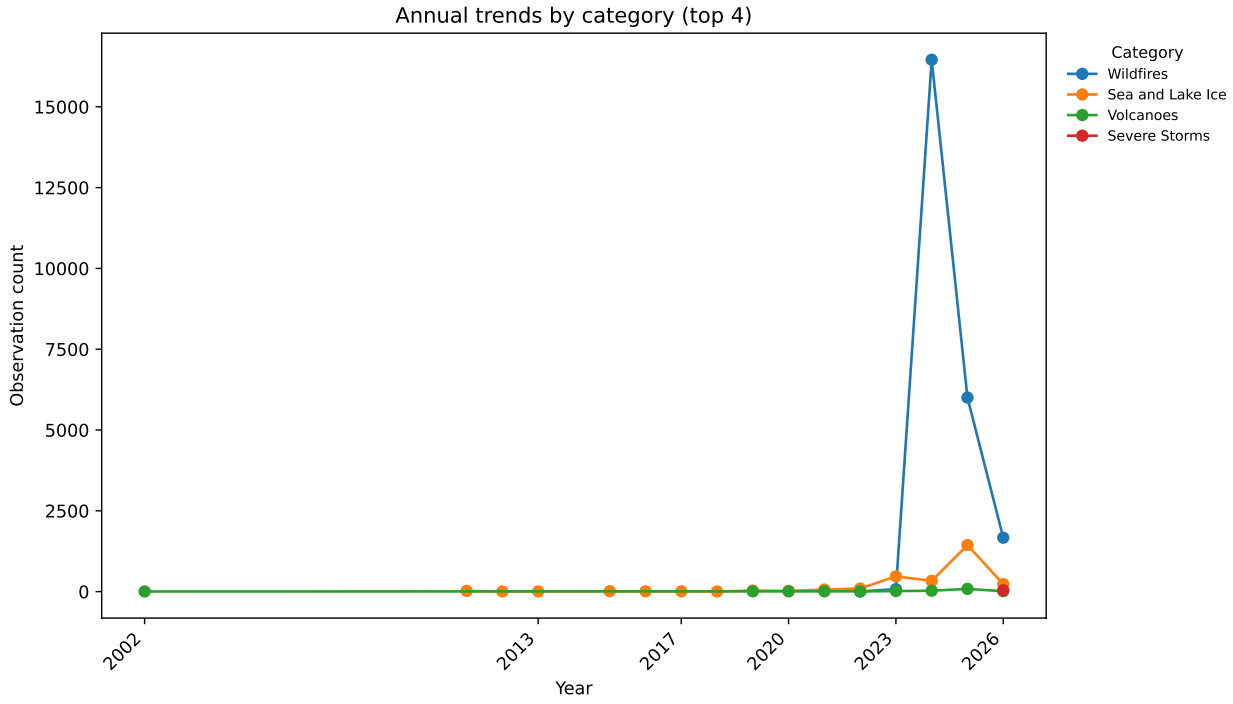

Figure 3: Annual trajectories of the most frequent disaster categories.

Category-stratified trajectories show that the recent surge is not equally shared across categories: wildfire-related observations account for most of the increase, while other hazard classes are flatter and lower-volume.

This decomposition materially changes the interpretation of the aggregate trend. Instead of concluding that all disaster classes are rising in parallel, the evidence points to composition effects: one dominant category contributes most of the total increase. As a result, pooled trend summaries should be read alongside category-level patterns to avoid over-generalization.

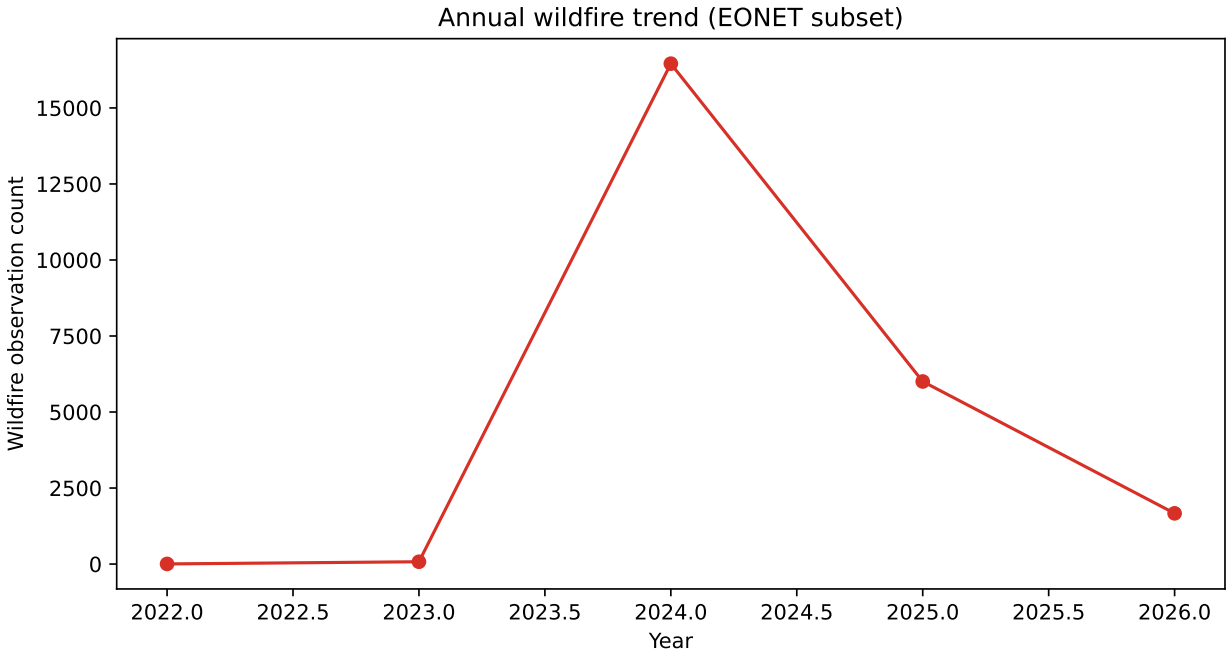

Figure 4: Wildfire-only annual observation trend.

The wildfire-only trend mirrors the aggregate surge, confirming that the overall temporal increase is primarily wildfire-driven. This dedicated subset analysis avoids dilution from sparse categories and provides a cleaner view of the dominant signal.

3.3 Seasonality and spatial concentration

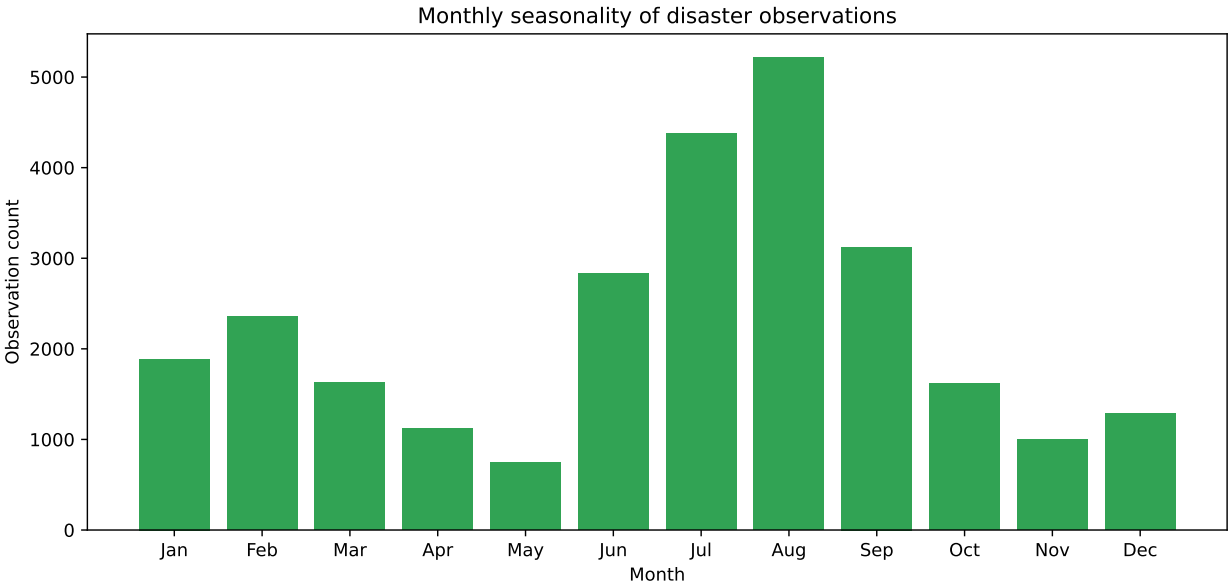

Figure 5: Monthly seasonality of total EONET observations aggregated across years.

Seasonality is strong: August has the highest count and May has the lowest. The peak month has about 7.0 times as many observations as the trough month, consistent with seasonal wildfire dynamics.

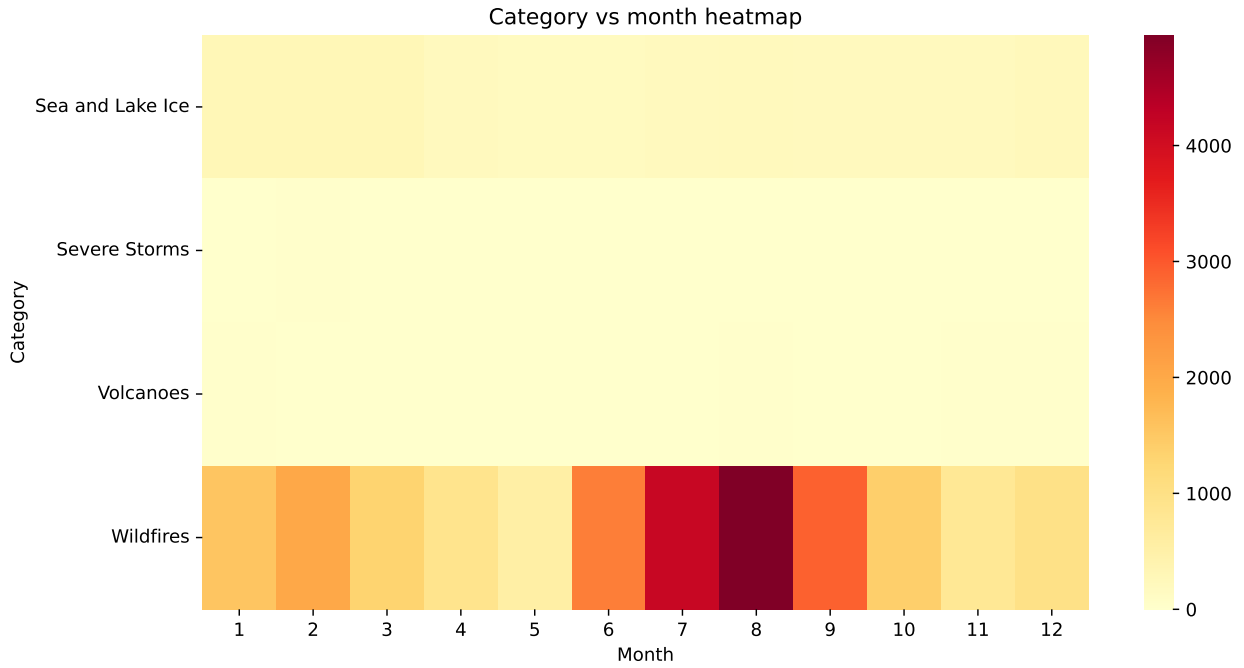

Figure 6: Heatmap of category-by-month observation counts in EONET.

The heatmap confirms that seasonality differs by category, motivating formal hypothesis testing rather than visual inspection alone.

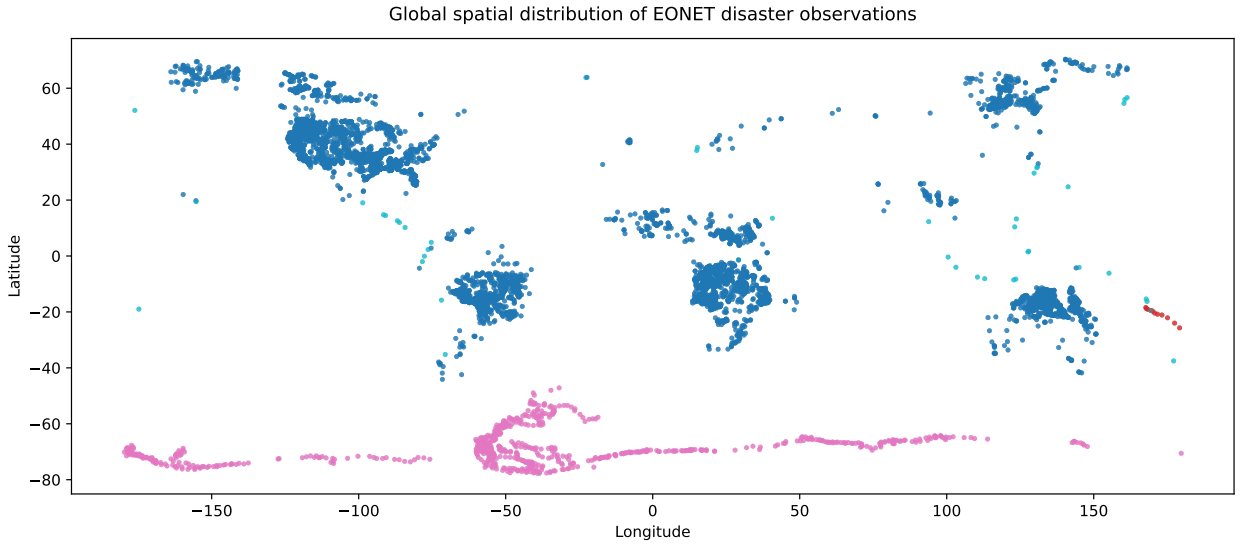

Figure 7: Global spatial distribution of georeferenced EONET observations.

Spatially, events are globally distributed but highly clustered in specific regions, indicating uneven

hazard activity and/or uneven observation density.

Importantly, spatial concentration in this dataset should not be interpreted as a direct map of intrinsic hazard risk. EONET reflects where events are observed, reported, and tracked with higher continuity; therefore, the map primarily represents observed-event density generated by both hazard processes and monitoring/reporting-system coverage.

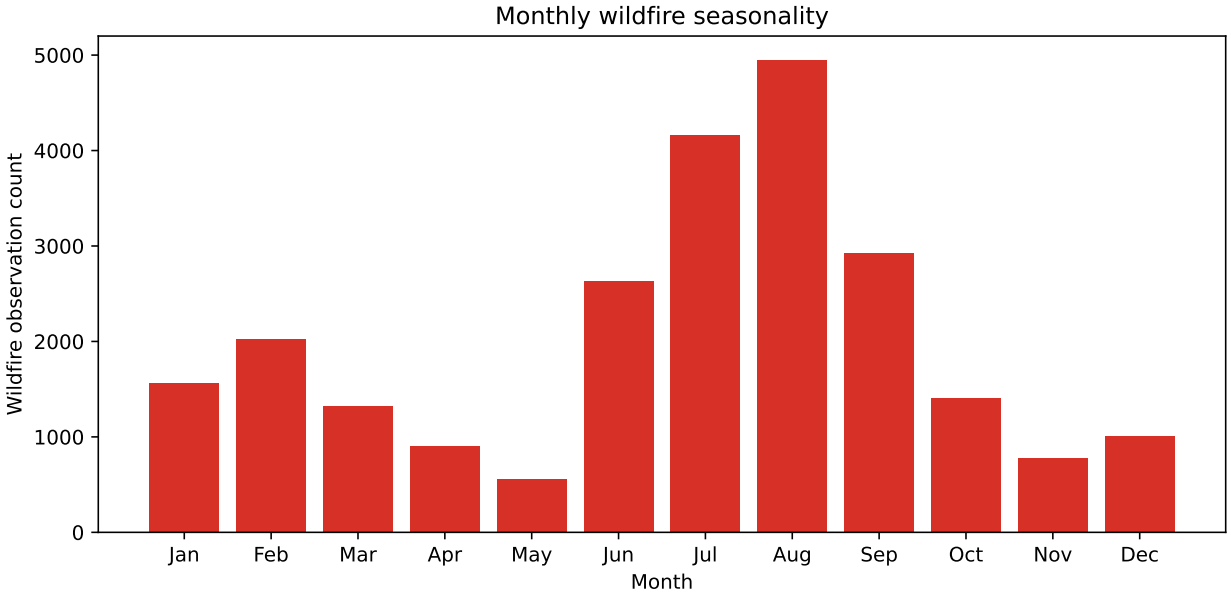

Figure 8: Wildfire-only monthly seasonality pattern.

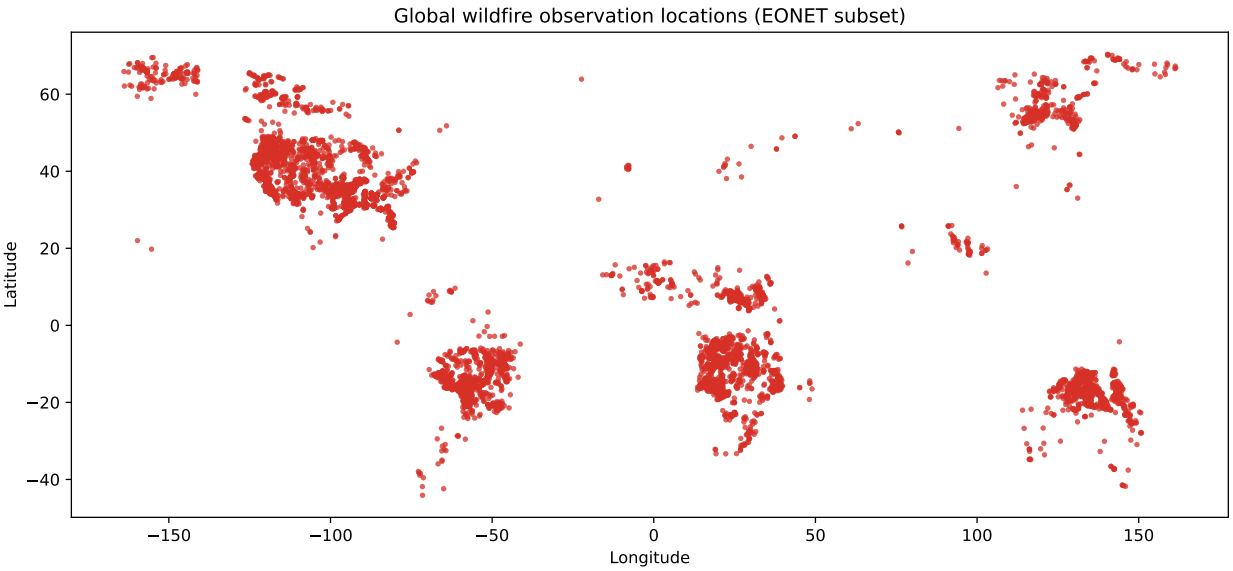

Figure 9: Wildfire-only spatial distribution.

These wildfire-only plots reinforce that the strongest temporal and seasonal signatures in the full dataset are driven by wildfire records, and that wildfire observations themselves remain spatially

concentrated rather than uniformly distributed.

3.4 Statistical tests

Table 2: Formal hypothesis-testing results for seasonality and long-run trend.

	Test	Statistic	p-value
0	Chi-square: category vs month	chi2=1628.32, dof=33	$p < 0.001$
1	Linear trend: annual count ~ year	slope=307.55, R2=0.185	$p = 0.096$

The chi-square test indicates strong evidence that category and month are associated ($p < 0.001$), meaning seasonal patterns vary across disaster types. For the annual linear trend, the estimated slope is positive, but the result is not statistically significant at the conventional 5% level ($p = 0.096$, R2=0.185). Therefore, the evidence for a long-run increase is weak to marginal rather than conclusive.

3.5 Predictive modeling (train/test evaluation)

Table 3: Out-of-sample predictive performance on test years (>= 2024).

	Model	MAE	RMSE	R2
0	Seasonal baseline	412.988	941.582	-0.228
1	OLS (linear year + month cycle + category)	410.456	942.813	-0.231
2	GAM (smooth year + month cycle + category)	393.350	928.331	-0.193
3	XGBoost (extended feature benchmark)	400.396	934.921	-0.210

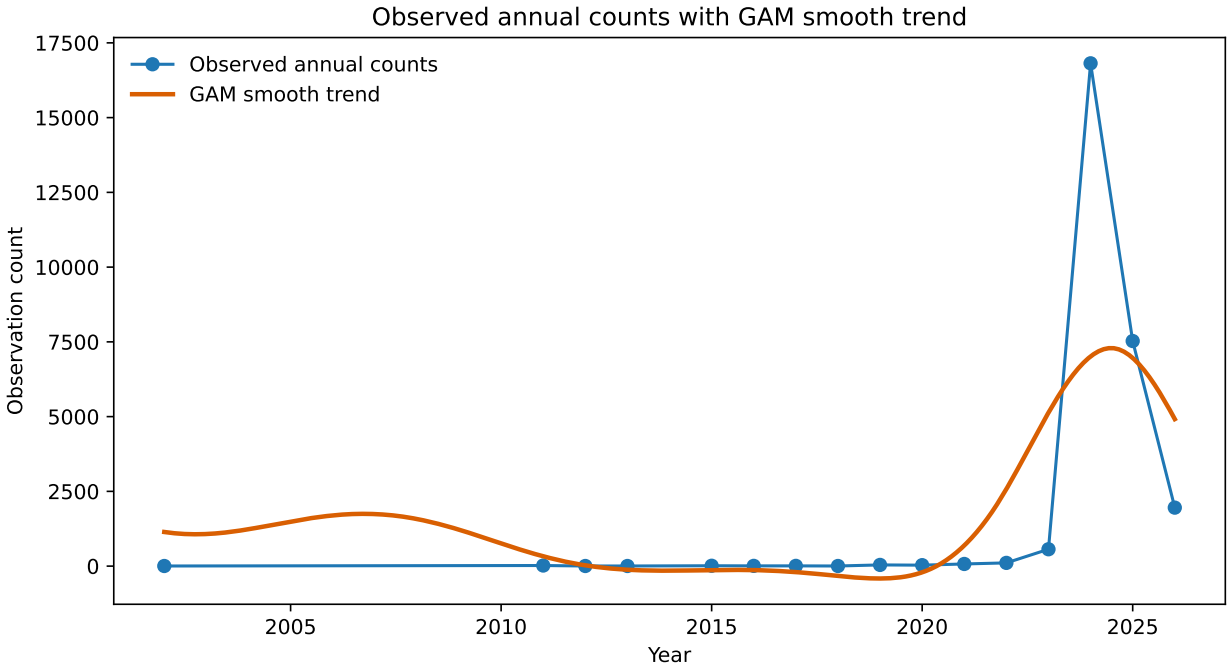

Figure 10: Smoothed annual trend from GAM overlaid on observed annual counts.

Table 4: Category-level RMSE comparison across baseline, OLS, GAM, and XGBoost.

	Category	Baseline	OLS	GAM	XGBoost
0	Wildfires	1451.61	1453.76	1432.32	1442.08
1	Sea and Lake Ice	92.60	87.73	44.93	75.98
2	Severe Storms	16.36	14.48	71.12	29.40
3	Volcanoes	7.49	5.94	54.64	6.68

For Research Question 4, out-of-sample prediction remains challenging under the current setup, but the extended benchmarks add nuance. OLS does not improve RMSE over the seasonal baseline, while GAM and XGBoost reduce overall RMSE (941.582 baseline vs 928.331 GAM and 934.921 XGBoost). However, all models still have negative test R2 (-0.228, -0.231, -0.193, -0.210), so forecasting quality remains limited in absolute terms.

These extensions provide robustness depth rather than replacing the main narrative. The GAM result suggests that a smooth nonlinear year effect is more suitable than a constant linear drift for this dataset, but category-level RMSE remains heterogeneous across models. This mixed pattern indicates that performance is constrained not only by model class, but also by the available feature signal and the observation structure of the source data.

Table 5: Wildfire-only out-of-sample benchmark (monthly counts, test years >= 2024).

	Model	MAE	RMSE	R2
0	Seasonal baseline	915.026	1451.663	-0.645

	Model	MAE	RMSE	R2
1	OLS	855.292	1413.602	-0.560
2	GAM	921.004	1457.853	-0.659
3	XGBoost	913.428	1449.103	-0.639

The wildfire-only benchmark isolates predictive behavior in the dominant category and removes cross-category mixing effects. This provides a direct check of whether model gains persist when the target is restricted to wildfire monthly counts only.

Table 6: Prediction-scope comparison between full multi-category and wildfire-only targets.

	Target	Model	MAE	RMSE	R2
0	Full dataset	Seasonal baseline	412.988	941.582	-0.228
1	Full dataset	OLS	410.456	942.813	-0.231
2	Full dataset	GAM	393.350	928.331	-0.193
3	Full dataset	XGBoost	400.396	934.921	-0.210
4	Wildfire-only	Seasonal baseline	915.026	1451.663	-0.645
5	Wildfire-only	OLS	855.292	1413.602	-0.560
6	Wildfire-only	GAM	921.004	1457.853	-0.659
7	Wildfire-only	XGBoost	913.428	1449.103	-0.639

Table 6 shows that predictive conclusions are scope-dependent and numerically distinct. In the full multi-category setting, GAM has the lowest RMSE (928.331) compared with baseline (941.582) and OLS (942.813). In the wildfire-only setting, OLS has the lowest RMSE (1413.602) compared with baseline (1451.663) and GAM (1457.853). This rank reversal confirms that preferred model choice depends on target definition. At the same time, out-of-sample R2 remains negative in both scopes (full: -0.228 to -0.193; wildfire-only: -0.645 to -0.560), indicating limited absolute forecast skill despite partial RMSE gains.

4 Conclusions and Summary

This final report extends the midterm analysis into a more complete statistical workflow. Relative to the midterm, key changes include explicit handling of class imbalance, addition of nonlinear benchmarks (GAM and XGBoost), and a two-scope prediction design that compares full multi-category and wildfire-only targets. Four conclusions stand out:

1. The EONET dataset is strongly imbalanced, with Wildfires contributing 89.1% of observations.
2. Total recorded observations increase sharply in recent years, peaking at 16,816 in 2024.
3. Seasonality is substantial and category-specific, supported by both visual diagnostics and a significant chi-square test.
4. Prediction results depend on scope: in the full multi-category task, GAM achieves the best RMSE (928.331), while in the wildfire-only task, OLS achieves the best RMSE (1413.602). However, all out-of-sample R2 values remain negative (full scope best R2 = -0.193; wildfire-only best R2 = -0.560), so overall predictive fit is still limited.

The broader implication is that global disaster monitoring data should be interpreted with attention to category imbalance and observation-system effects, not only hazard dynamics. In this dataset, that means conclusions are most reliable for wildfire-dominant behavior, while inference for sparse categories remains provisional. Main limitations include potential reporting heterogeneity across years, the use of observation counts rather than physical severity, and simple linear model structure. Future work can incorporate richer covariates (climate indices, regional land cover, remote-sensing intensity metrics) and compare nonlinear or hierarchical models.

At a bigger-picture level, this project shows that disaster-monitoring datasets are dual-purpose artifacts: they describe natural events, but they also encode how institutions observe the Earth. For policy and research use, this means trend, seasonality, and spatial concentration should be interpreted as signals from both hazards and measurement systems. A publication-ready analysis therefore needs both statistical rigor and measurement-awareness, especially when conclusions may inform real-world preparedness decisions.

Interactive versions of the key maps and trend views are provided on the project GitHub website, while this PDF presents the corresponding publication-ready static figures and statistical interpretation.

For reproducibility, this report runs entirely from repository-relative paths and a cached dataset in `data/eonet_events.csv`.