WIDS COMPETITION 2026

Team name: RO- ALERT

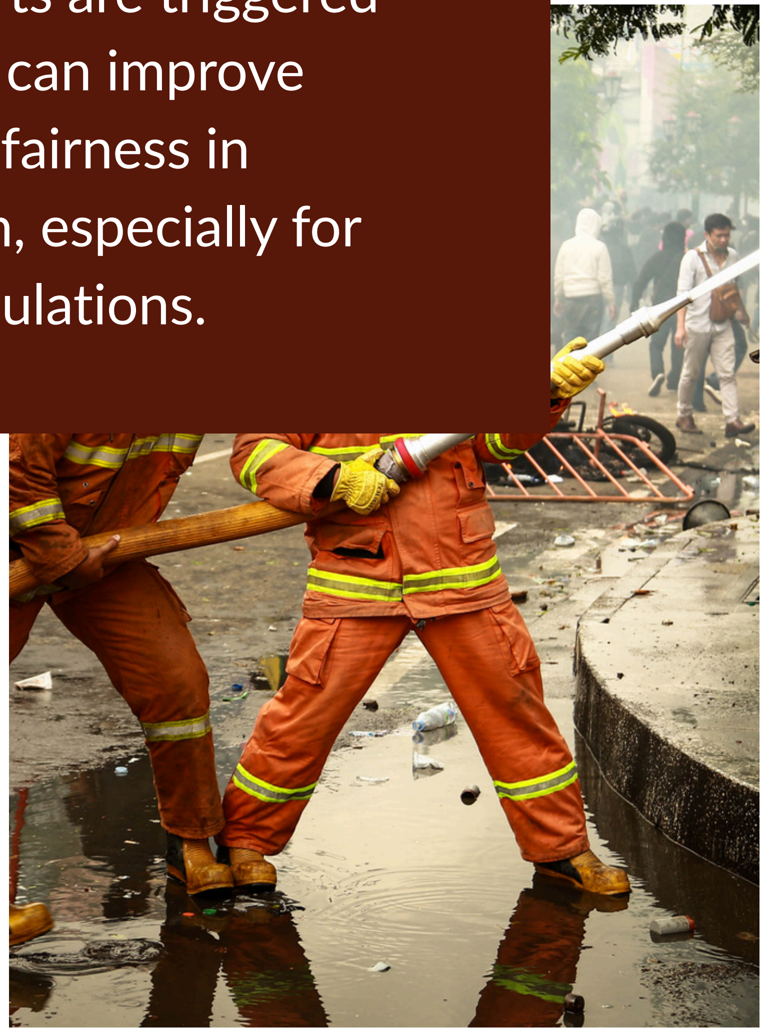
Team members: Sasu Sabrina
Sandu Bianca Antonia
Sandulescu Crina
Tilica Mihnea David
Zamfir Robert
Radu ALexandru

Bucharest University of Economic Studies

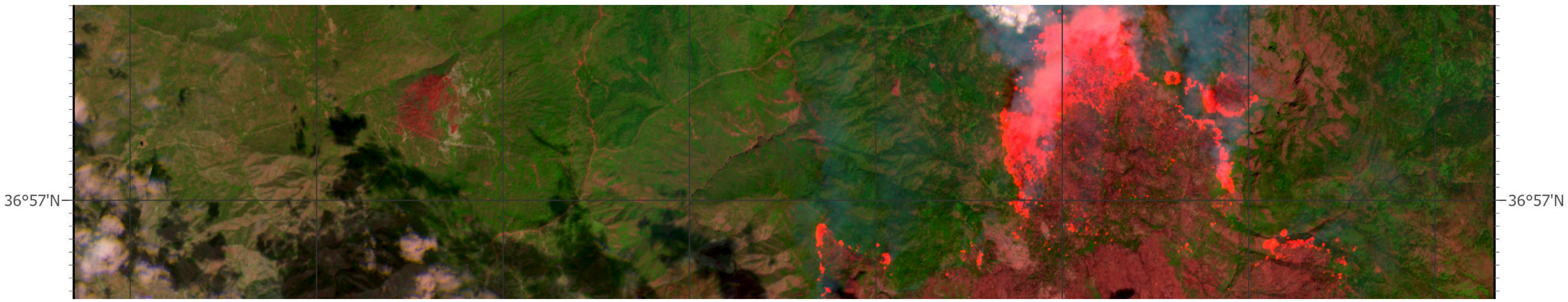
ROUTE 1

Core Question:
How can we reduce delays in evacuation alerts and improve response times for the communities that are most at risk?

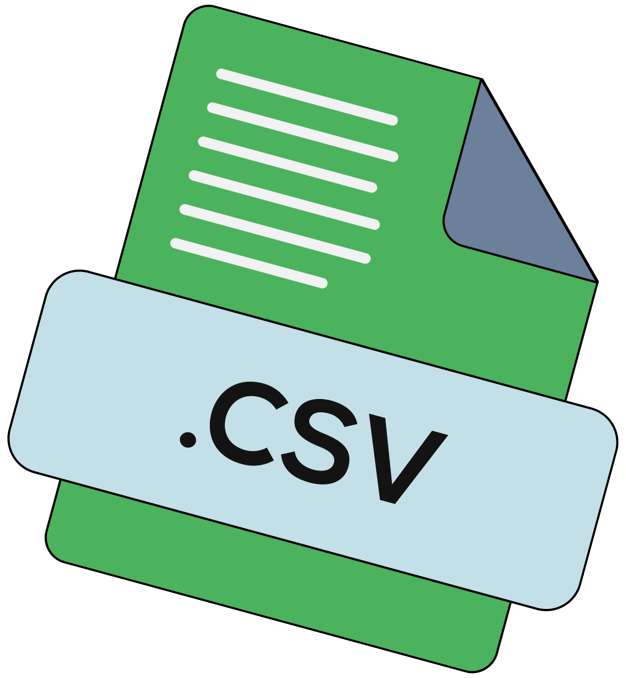
This route focuses on analyzing how and when evacuation alerts are triggered — and how we can improve timeliness and fairness in communication, especially for vulnerable populations.



DATASET OVERVIEW



DATA	DESCRIPTION
evac_zone_status_geo_event_map.csv	maps wildfire events to evacuation zones
evac_zones_gis_evaczone.csv	defines evacuation zones as spatial entities, including their identifiers, names, activity status
geo_events_geoevent.csv	records of geographic events, including wildfire incidents, with their location
geo_events_geoeventchangelog.csv	time-stamped updates to wildfire events, capturing changes in reported field



Data Cleaning

```
After cleaning:
  events    (61779, 12)
  changes   (178697, 5)
  evac_zones (37458, 16)
  evac_map  (4429, 3)

Extracted fields preview:
   id geo_event_type  containment  acreage is_fps  is_prescribed
0  76       wildfire        100.0     50.0  False          False
1  77       wildfire        100.0      0.0  False          False
2  78       wildfire          0.0      0.0  False          False
3  79       wildfire          0.0      0.0  False          False
4  80       wildfire        100.0      0.0  False          False
5  81       wildfire        100.0      0.0  False          False
6  82       wildfire          NaN      NaN  False          False
7  83       wildfire        100.0      1.0  False          False
8  84       wildfire          0.0      0.0  False          False
9  85       wildfire        100.0      0.0  False          False

Non-null counts for extracted fields:
containment      24977
acreage          50460
is_fps           26792
is_prescribed    61779
dtype: int64
```

MISSING VALUES

Key Data Insights

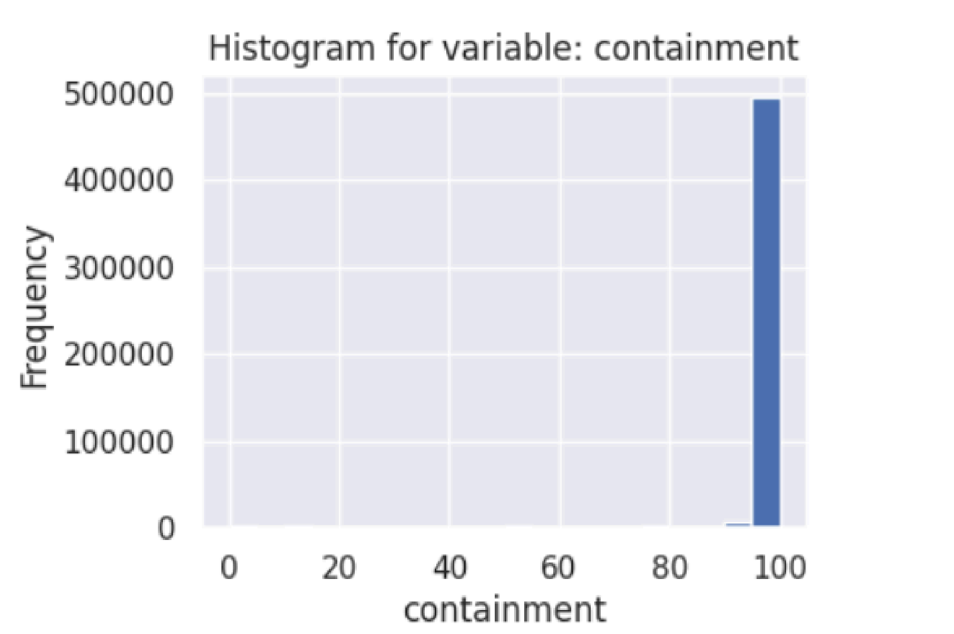
- Scale of Operations: 42,000+ unique incidents, providing a statistically significant dataset for wildfire modeling.
- Categorical Precision: Low cardinality in notification_type and is_prescribed (2 unique values) allows for perfect segmentation between planned burns and emergency wildfires.
- Priority Logic: While structure_threat is rare (99% missing), its presence represents a high-priority alert state that bypasses standard lag thresholds.

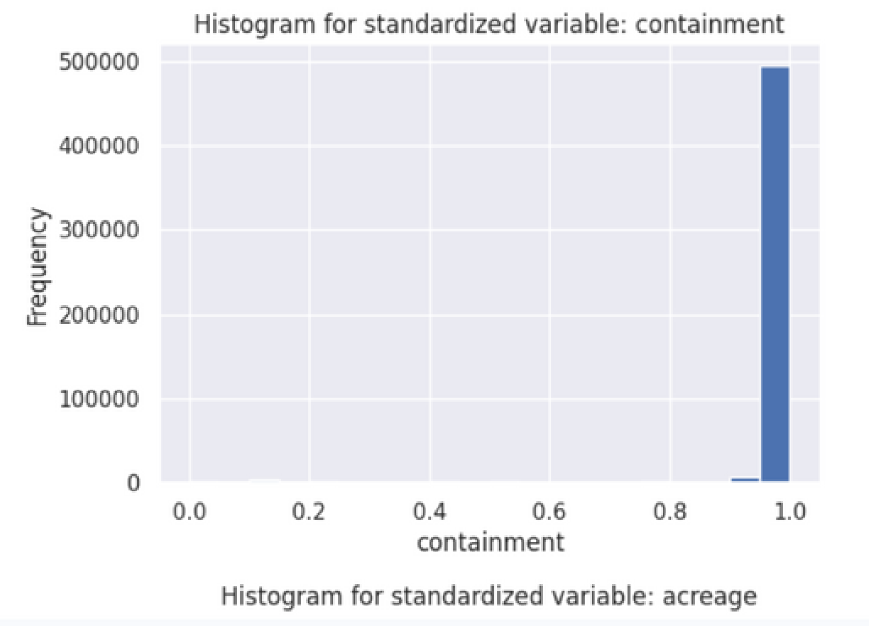
Missingness Category	Key Variables	Interpretation
High (>90%)	spotting, structure_threat, status, rate_of_spread	Critical Triggers: High sparsity is expected as these reflect extreme escalations. These act as system overrides rather than predictive features.
Moderate (10-60%)	external_status, display_name, uid_v2, containment	Incident Lifecycle: Missingness indicates rapid control. Many fires are contained (is_fps) before formal naming or zone assignment.
Low/No (0-3%)	acreage, date_created_log, alert_lag_min	Model Core: Our most reliable metrics. Complete timelines allow for robust Alert Lag analysis, the foundation of the predictive engine.

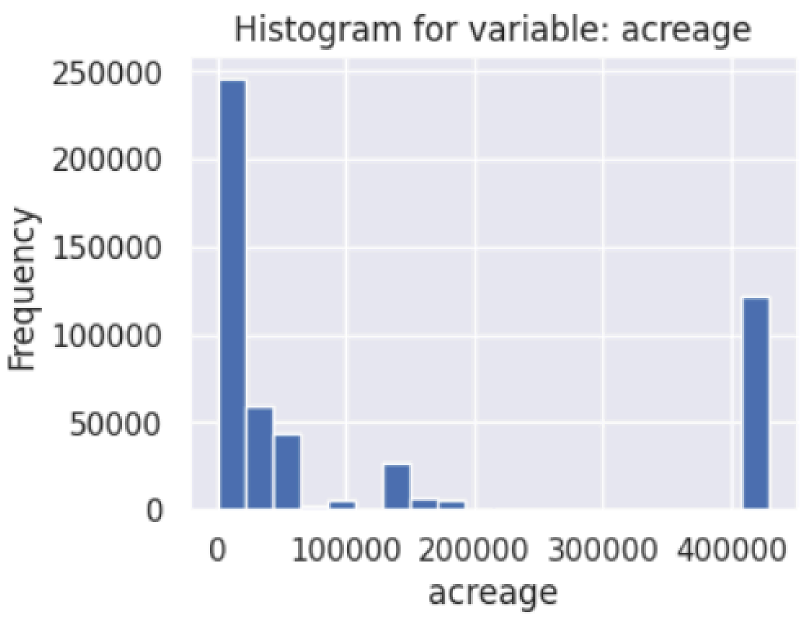
Missing Values

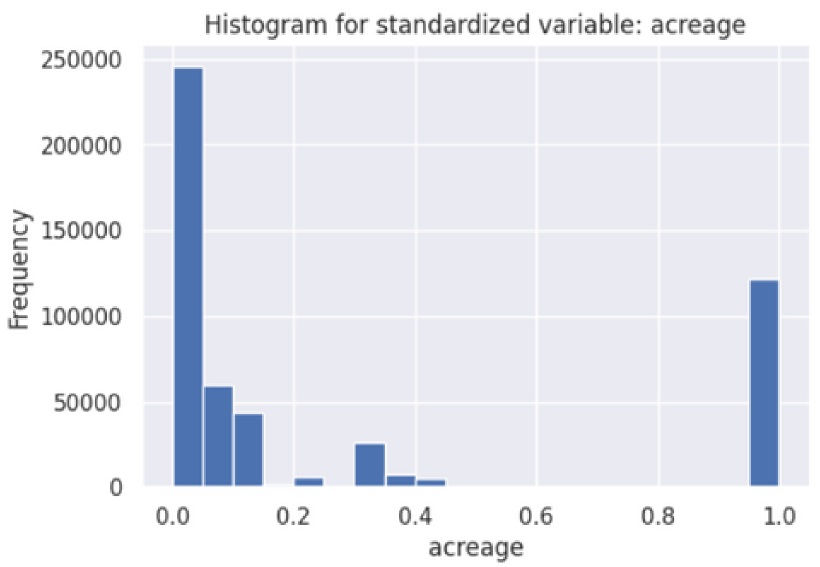
	Variable	Missing values count	Missing values %	Unique values count	Data type
33	pending_updates	592,538	100.00	0	float64
4	spotting	591,900	99.89	3	object
3	structure_threat	590,088	99.59	1	object
30	status	584,259	98.60	3	object
2	rate_of_spread	574,995	97.04	4	object
34	external_status	324,098	54.70	29	object
24	display_name	265,339	44.78	3,086	object
31	geom_label	264,898	44.71	3,077	object
22	date_modified_evac	264,898	44.71	3,120	object
28	source_extra_data	264,898	44.71	1	object
32	is_pending_review	264,898	44.71	1	object
21	date_created	264,898	44.71	3,120	object
23	is_active	264,898	44.71	2	object
20	id_evac	264,898	44.71	3,120	float64
25	region_id	264,898	44.71	82	float64
27	dataset_name	264,898	44.71	89	object
29	geom	264,898	44.71	3,078	object
26	source_attribution	264,898	44.71	[illegible]	object

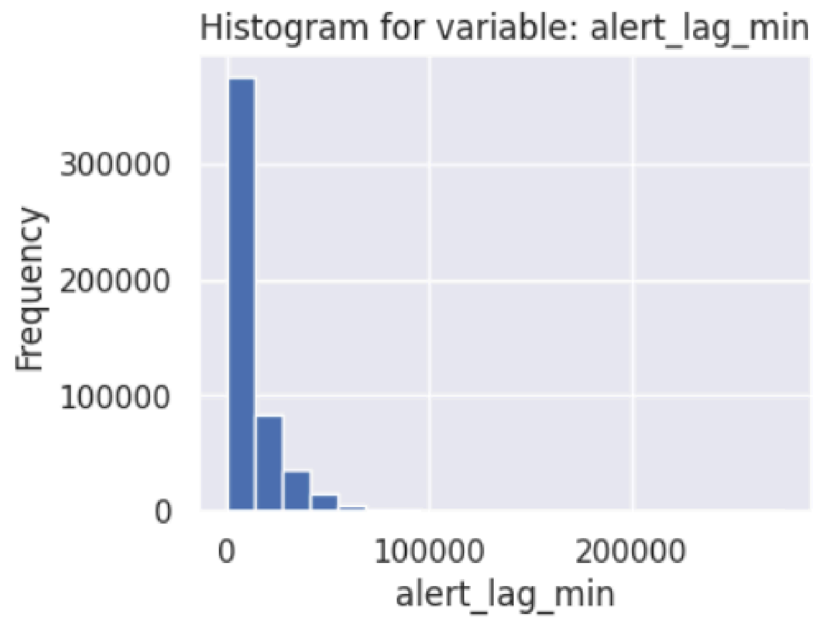
CLUSTERIZATION

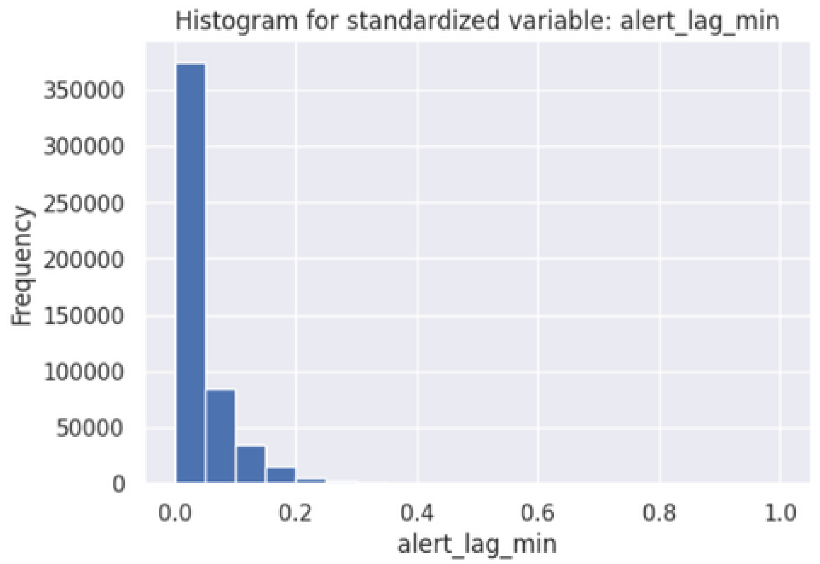








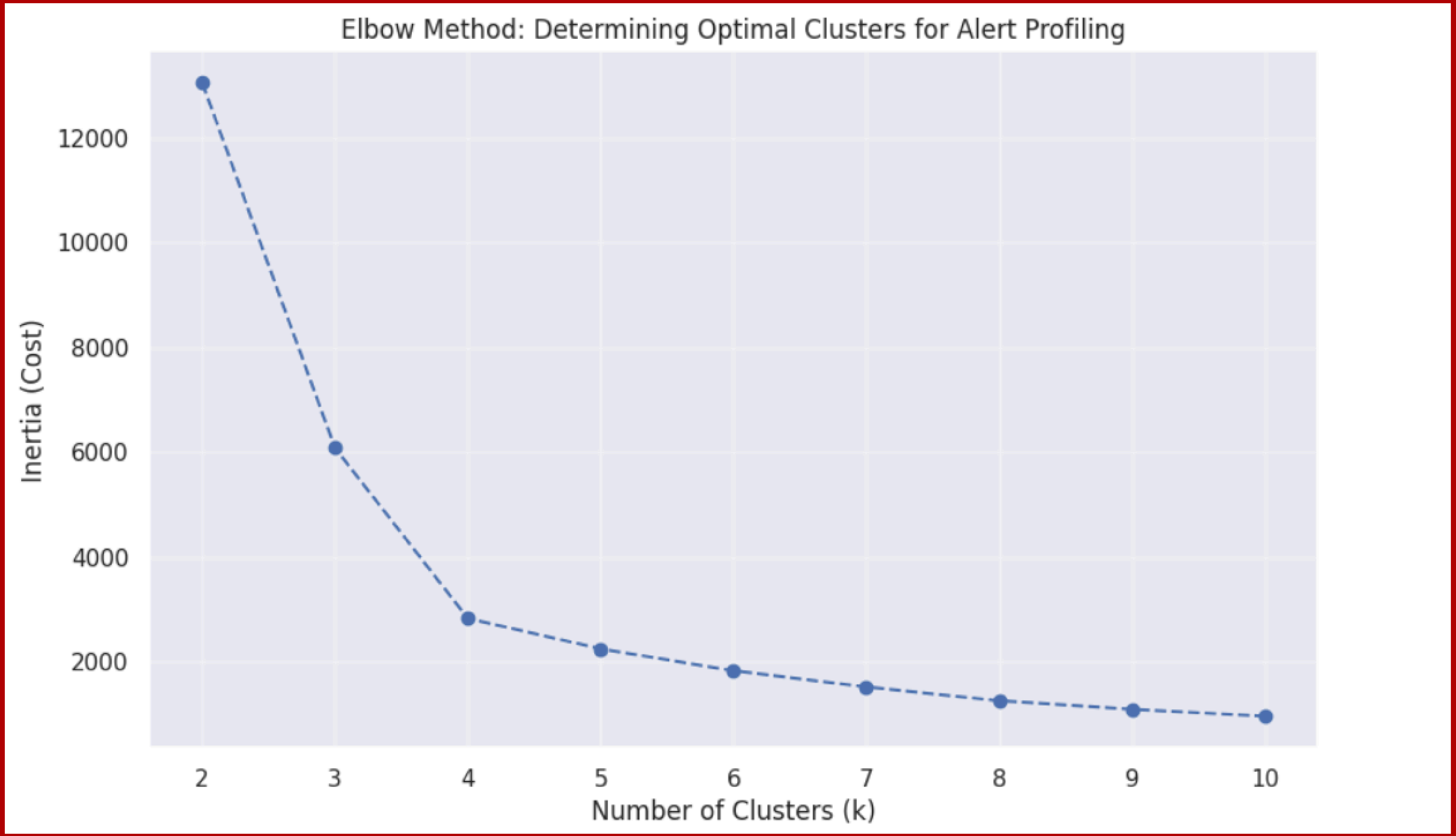




- **Containment**: The distribution is extremely concentrated near 100%, indicating that most recorded wildfire events are reported as nearly fully contained
- **Acreage**: The distribution is highly right-skewed, showing that the majority of fires affect relatively small areas, while a small number of extreme events account for very large burned acreages.
- **Alert lag (minutes)**: The distribution is strongly right-skewed, with most alerts issued within relatively short time spans, but with a long tail of cases experiencing very large delays.

ELBOW METHOD

k=3



CLUSTERS DESCRIPTION

Clus_km	count
0	384449
1	121676
2	12615

Weight Centers

Clus_km	containment	acreage	alert_lag_min
0	99.112004	29724.314312	8856.294203
1	99.000000	429500.670075	16077.549650
2	23.736187	8468.268438	13323.715908

Df_head

	containment	acreage	alert_lag_min	Clus_km
7	100.0	704.0	10294.900376	0
8	100.0	15563.0	4536.511185	0
9	100.0	15563.0	4536.511185	0
10	100.0	15563.0	4536.511185	0
11	100.0	15563.0	4536.511185	0

Cluster 0

contains the majority of observations (≈ 384,000 events), is characterized by very high containment (≈ 99.1%), moderate average fire size (≈ 29,700 acres), and a mean alert lag of about 8,856 minutes, indicating routine or controlled incidents

Cluster 1

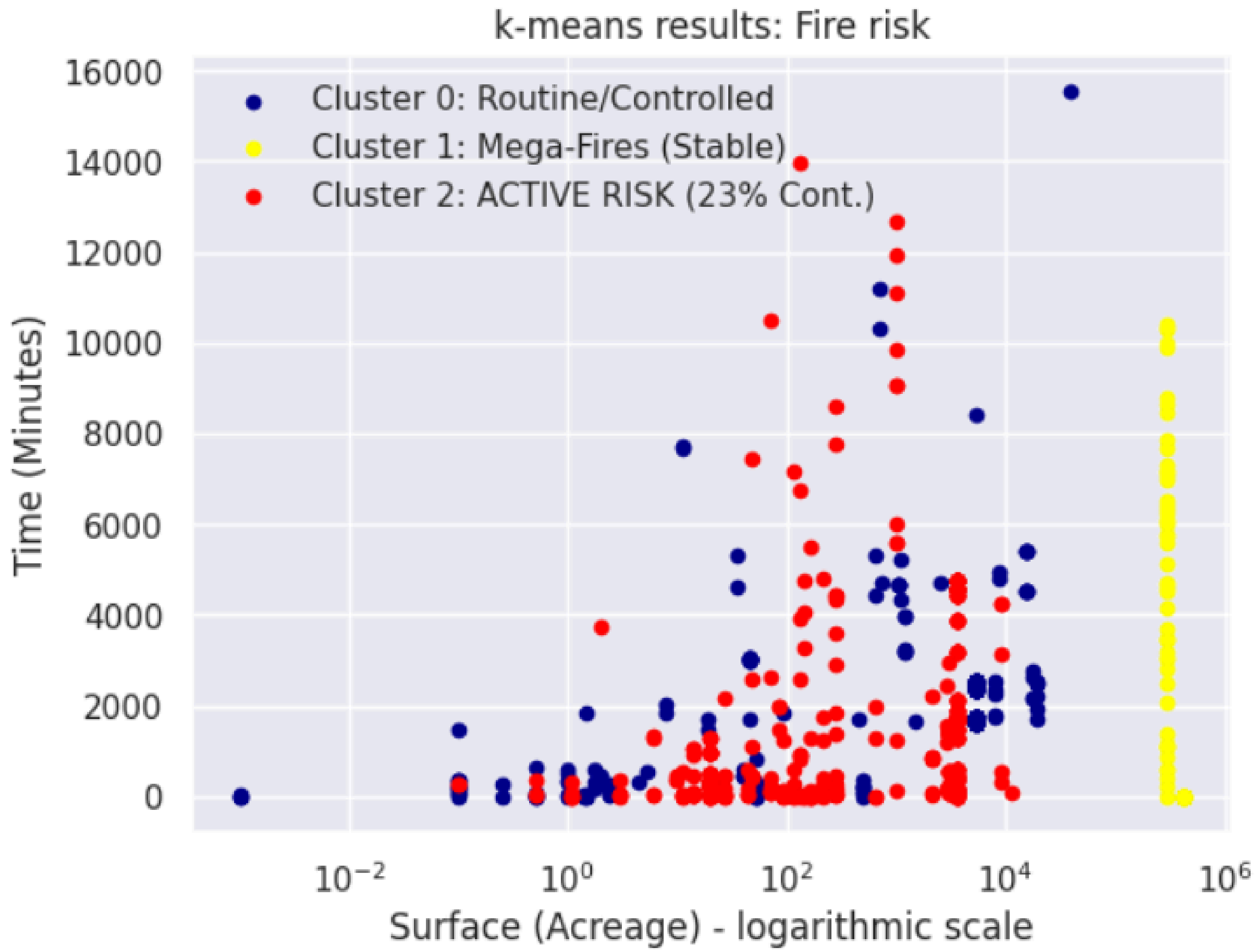
21,700 events, shows similarly high containment (≈ 99%) but an extremely large average acreage (≈ 429,500 acres) and the longest alert lag (≈ 16,078 minutes), reflecting large-scale, complex mega-fires

Cluster 2

the smallest group (≈ 12,600 events), stands out with very low containment (≈ 23.7%), smaller average fire size (≈ 8,468 acres), and a high alert lag (≈ 13,324 minutes), quantitatively confirming that low containment and active fire dynamics can lead to severe delays

K-MEANS METHOD

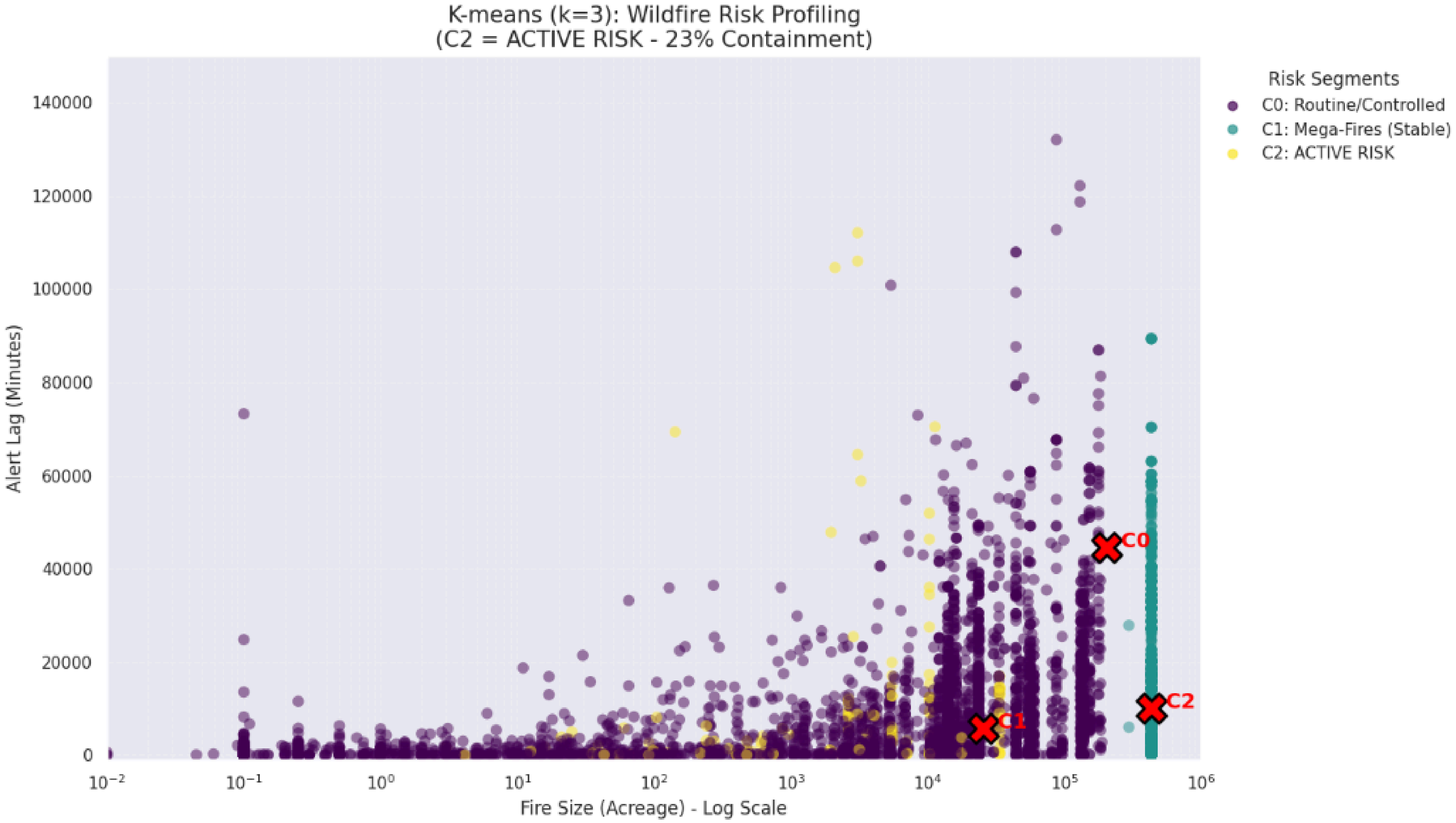
Cluster 0 (blue) groups routine or controlled fires, which generally have moderate acreage and shorter alert delays, although some variability remains due to operational complexity.
Cluster 1 (yellow) corresponds to mega-fires, characterized by extremely large burned areas but relatively stable and consistent alert timing, reflected by the vertical concentration at very high acreage values.
Cluster 2 (red) represents active-risk fires, where lower containment and ongoing fire dynamics are associated with longer and more variable alert delays, even at moderate acreage levels.



K-MEANS METHOD

Cluster C0 (Routine/Controlled) groups the majority of events, spanning a wide range of fire sizes but generally associated with lower to moderate alert delays, reflecting incidents that are operationally managed despite variability in scale. Cluster C1 (Mega-Fires, Stable) is concentrated at very large acreage values, showing that extremely large fires tend to exhibit more consistent alert timing, likely due to sustained monitoring and established response protocols.

Cluster C2 (Active Risk) combines very large fire sizes with relatively higher and more variable alert lags, indicating situations where low containment and ongoing fire dynamics increase operational uncertainty and delay alert escalation.



ANOVA

ANOVA Hypotheses

- **HO:** There is no difference between the averages of the tested groups.
- **H1:** There is a statistically significant difference in at least one group average.

```
ANOVA for variable: containment
F-statistic = 2404104.409
p-value = 0.0000000000
→ Differences of containment between the 3 profile risks are statistical significant (α = 0.05).

ANOVA for variable: acreage
F-statistic = 5371789.773
p-value = 0.0000000000
→ Differences of acreage between the 3 profile risks are statistical significant (α = 0.05).

ANOVA for variable: alert_lag_min
F-statistic = 11934.175
p-value = 0.0000000000
→ Differences of alert_lag_min between the 3 profile risks are statistical significant (α = 0.05).
```

Classification models

- Decision tree
- Random forest
- XGBoost

```
DecisionTreeClassifier
DecisionTreeClassifier(max_depth=6, random_state=42)
```

```
RandomForestClassifier
RandomForestClassifier(max_depth=10, n_estimators=300, n_jobs=-1,
                       random_state=42)
```

```
XGBoost trained.
```

```
test_size=0.30
```

The data will be split on train/test subsets based on a 70-30 split.

Target variable

	count
delayed_alert	
1	436279
0	34957

The target variable, delayed_alert, is a binary indicator derived from alert_lag_min. Records with missing or negative alert lag values are removed. A delay threshold of 60 minutes is applied: delayed_alert is set to 1 if alert_lag_min exceeds 60 minutes, and 0 otherwise.

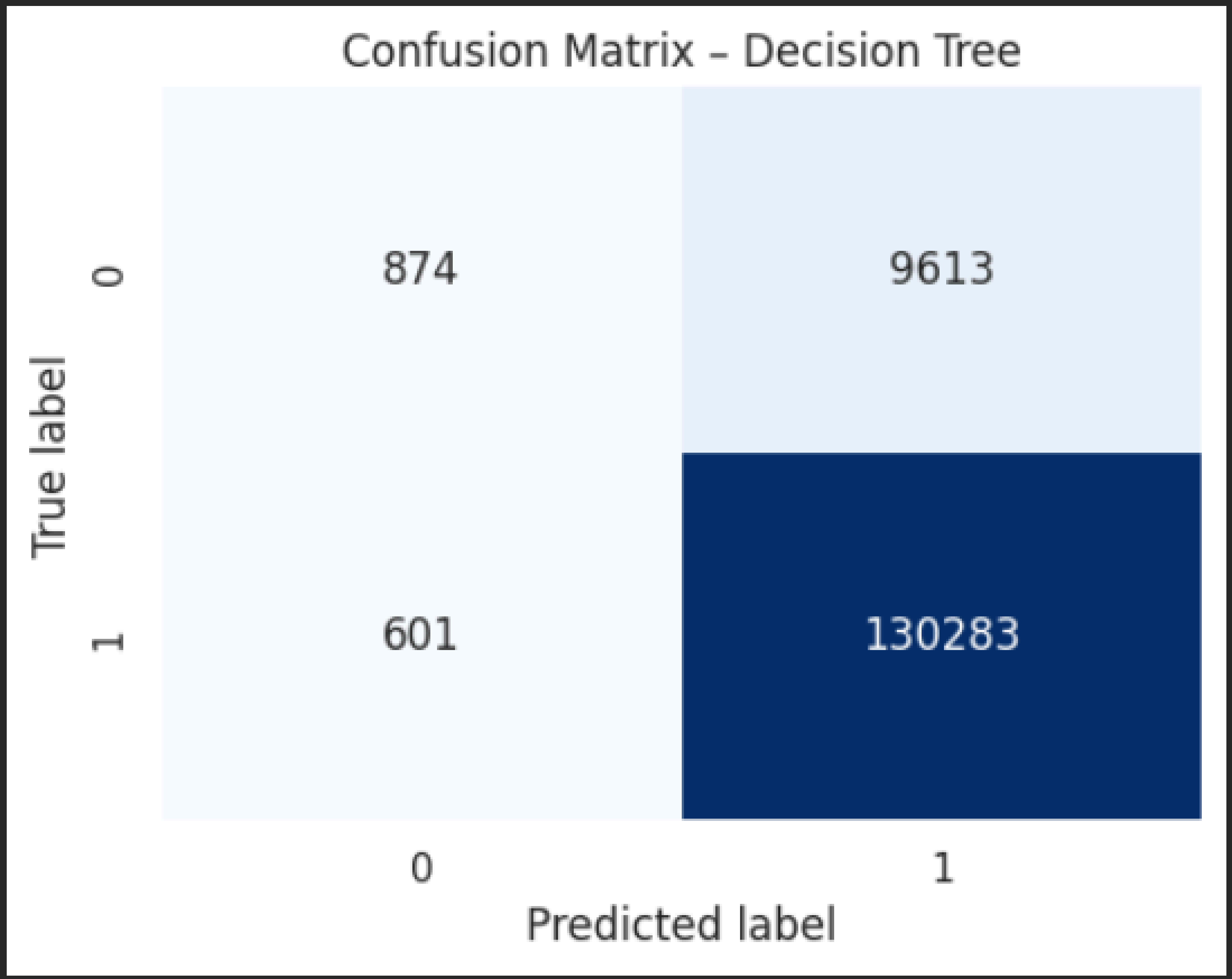
Features

Variable type	Variable name	Interpretation
Numerical	acreage	Represents the total burned area of the wildfire, serving as a proxy for the scale and operational complexity of the incident.
Numerical	containment	Indicates the percentage of the wildfire perimeter that has been contained, reflecting how much of the fire is under control at a given time.
categorical	is_fps	Model Core: Our most reliable metrics. Complete timelines allow for robust Alert Lag analysis, the foundation of the predictive engine.
categorical	is_prescribed	A binary indicator identifying whether the fire is a prescribed (planned) burn, which typically follows different alerting and response procedures than unplanned wildfires.

Decision Tree

The confusion matrix shows that the model correctly identifies 130,283 delayed alerts (true positives) while missing 601 delayed cases (false negatives), resulting in a very high recall for the delayed class of 0.9954. However, performance on the non-delayed class is weak: only 874 non-delayed alerts are correctly classified, while 9,613 are incorrectly flagged as delayed, which explains the very low recall of 0.0833 for class 0. The overall accuracy is 0.9278, but this is driven mainly by the dominance of the delayed class. The ROC AUC of 0.8355 indicates limited discrimination ability, reflecting the model's difficulty in separating non-delayed alerts from delayed ones.

```
================ Decision Tree =================
```

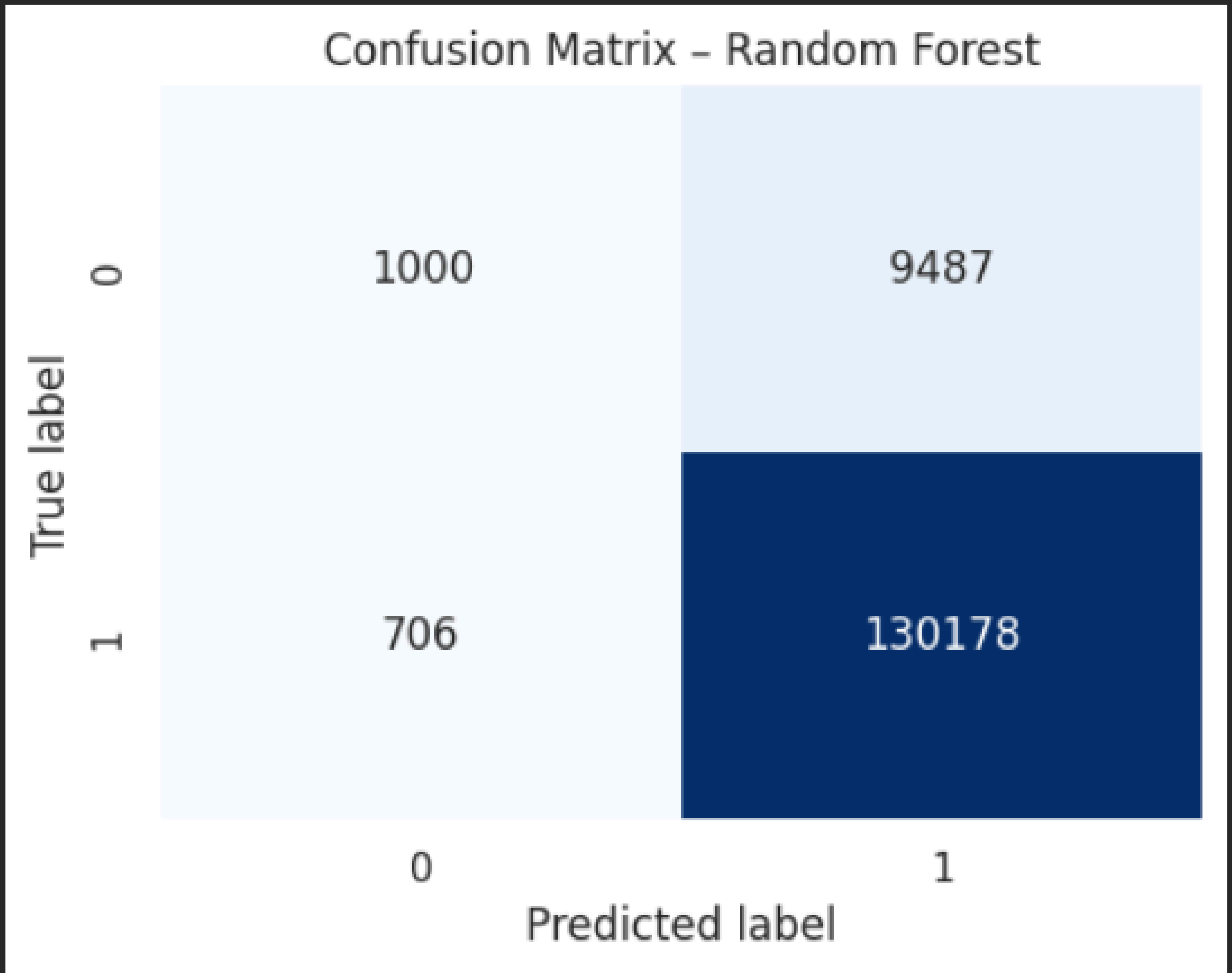



```
Classification report:
              precision    recall  f1-score   support

           0     0.5925    0.0833    0.1461     10487
           1     0.9313    0.9954    0.9623    130884

    accuracy                         0.9278    141371
   macro avg     0.7619    0.5394    0.5542    141371
weighted avg     0.9062    0.9278    0.9017    141371

ROC AUC: 0.8355
```

Random forest

The confusion matrix indicates 130,178 true positives and 706 false negatives, leading to a delayed-alert recall of 0.9946, meaning that almost all delayed alerts are detected. For the non-delayed class, the model correctly classifies 1,000 cases, but still misclassifies 9,487 non-delayed alerts as delayed, yielding a recall of 0.0954 for class 0. The overall accuracy is 0.9279, similar to the Decision Tree, but with slightly better identification of non-delayed cases. The ROC AUC of 0.8474 reflects improved class separation compared to the Decision Tree, although misclassification of non-delayed alerts remains substantial.

```
================ Random Forest ================
```

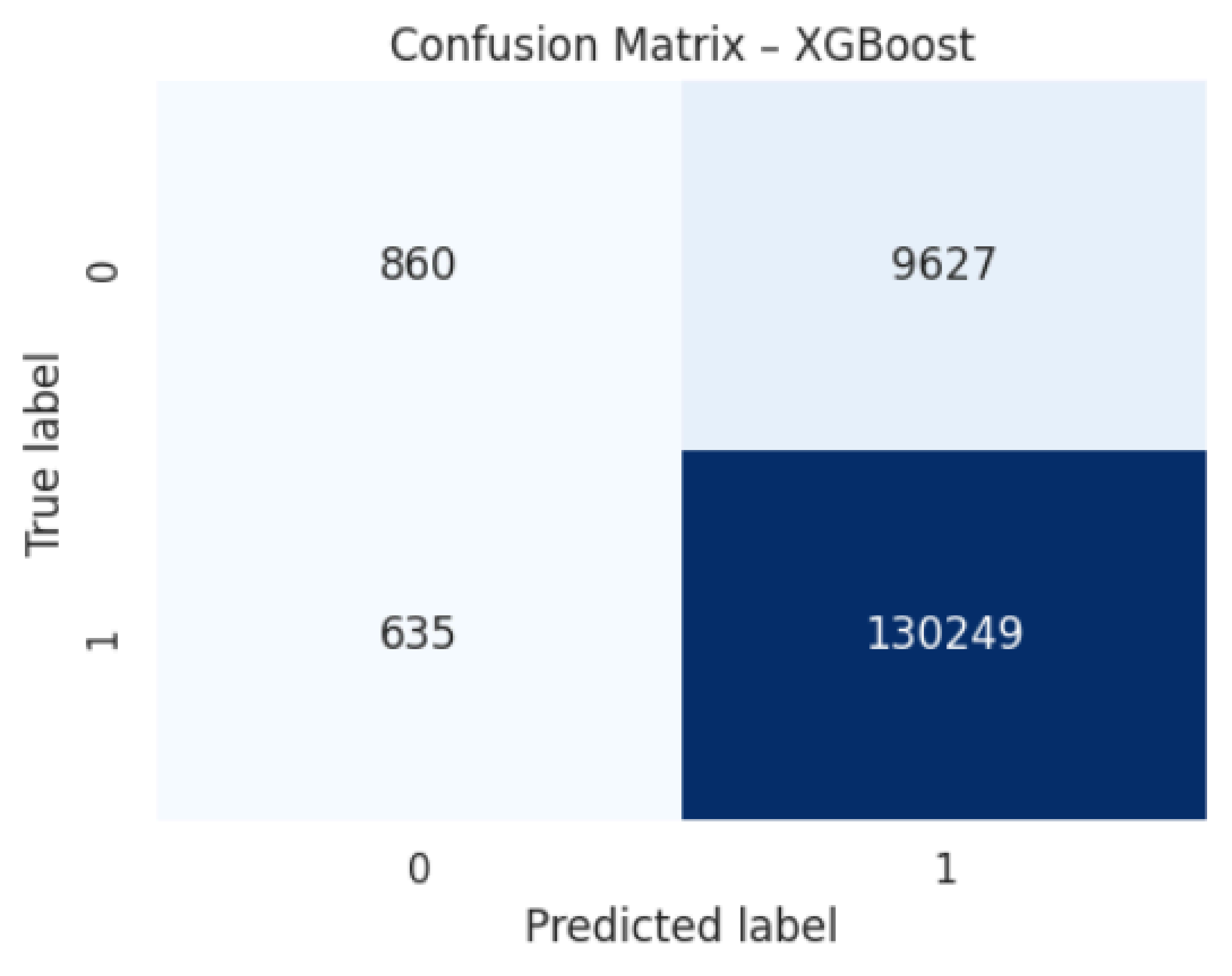



```
Classification report:
              precision    recall  f1-score   support

           0     0.5862    0.0954    0.1640     10487
           1     0.9321    0.9946    0.9623    130884

    accuracy                         0.9279    141371
   macro avg     0.7591    0.5450    0.5632    141371
weighted avg     0.9064    0.9279    0.9031    141371

ROC AUC: 0.8474
```

XGBoost

The confusion matrix shows 130,249 true positives and 635 false negatives, resulting in a delayed-alert recall of 0.9951, which means delayed alerts are almost never missed. For non-delayed alerts, 860 cases are correctly identified, while 9,627 are incorrectly labeled as delayed, corresponding to a recall of 0.0820 for class 0. The overall accuracy reaches 0.9274, again driven by strong performance on the delayed class. The ROC AUC value of 0.8465 indicates strong discrimination ability, slightly below Random Forest in this run, but still clearly higher than the Decision Tree.

```
================ XGBoost =================
```

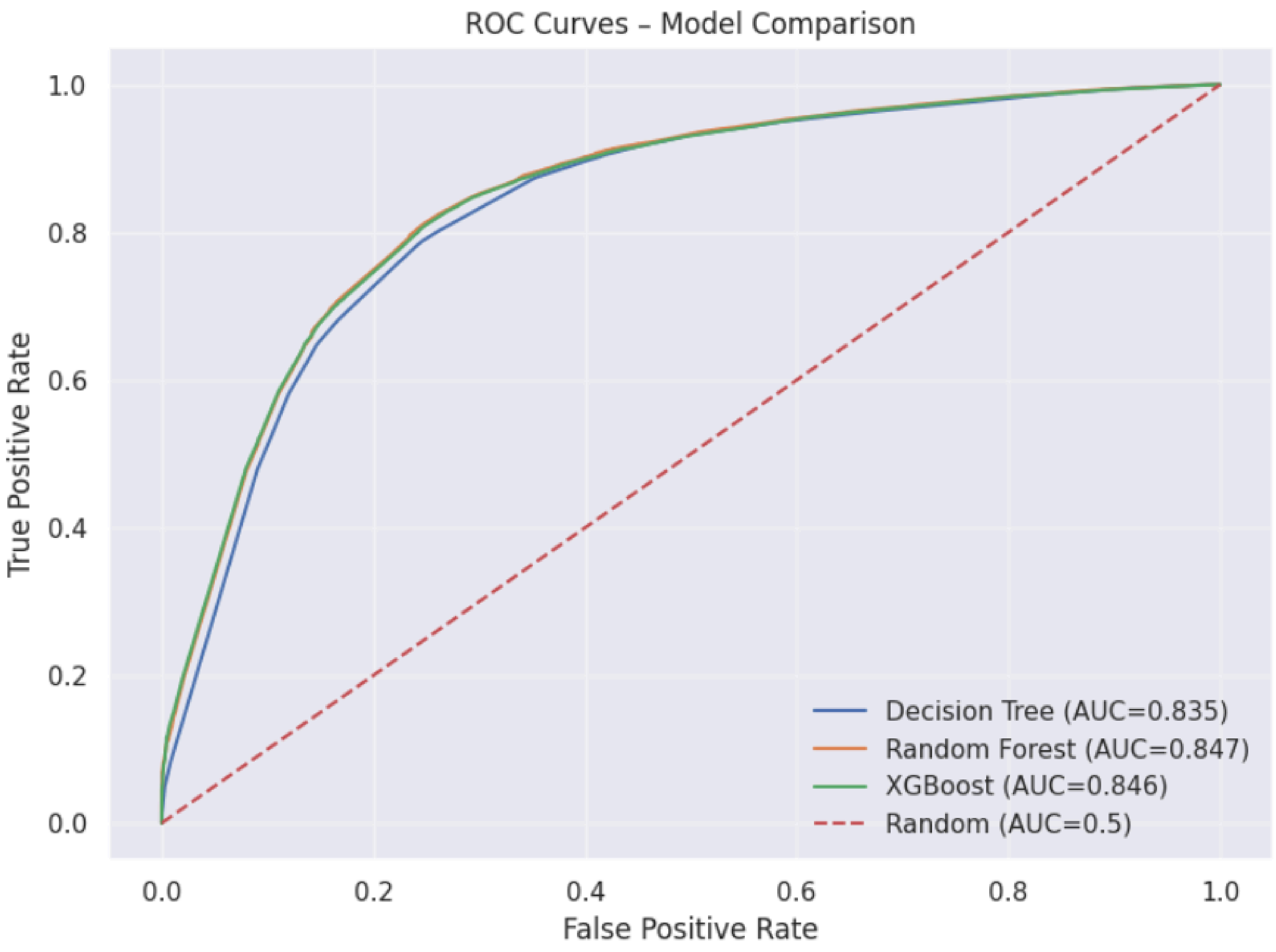



```
Classification report:
              precision    recall  f1-score   support

           0     0.5753    0.0820    0.1435     10487
           1     0.9312    0.9951    0.9621    130884

    accuracy                         0.9274    141371
   macro avg     0.7532    0.5386    0.5528    141371
weighted avg     0.9048    0.9274    0.9014    141371

ROC AUC: 0.8465
```

Model comparison

The results indicate that all three models perform at a consistently high level, with accuracy values of approximately 92.7–92.8%, which shows that the available information captures well the factors associated with alert delays. This suggests that the problem is well defined and that the selected features contain meaningful signals for distinguishing between delayed and non-delayed alerts.

The ROC curves further confirm this result, as the Random Forest and XGBoost curves consistently lie above the Decision Tree curve for most values of the false positive rate. This indicates that, for the same proportion of false alarms, these models are able to correctly identify a larger share of delayed alerts. In practical terms, both ensemble models provide more reliable probabilistic signals that can be adjusted to different operational thresholds, without a disproportionate increase in unnecessary alerts.



Model Comparison Table ===

Model	Accuracy	Precision	Recall	F1	ROC_AUC
Random Forest	0.9279	0.9321	0.9946	0.9623	0.8474
XGBoost	0.9274	0.9312	0.9951	0.9621	0.8465
Decision Tree	0.9278	0.9313	0.9954	0.9623	0.8355

Among the evaluated models, Random Forest achieves the highest ROC AUC value (≈ 0.847), followed very closely by XGBoost (≈ 0.846), while the Decision Tree records a lower value (≈ 0.836). Although the differences between Random Forest and XGBoost are small, the ROC AUC values indicate that the ensemble-based models provide superior discrimination between delayed and non-delayed alerts across different decision thresholds.

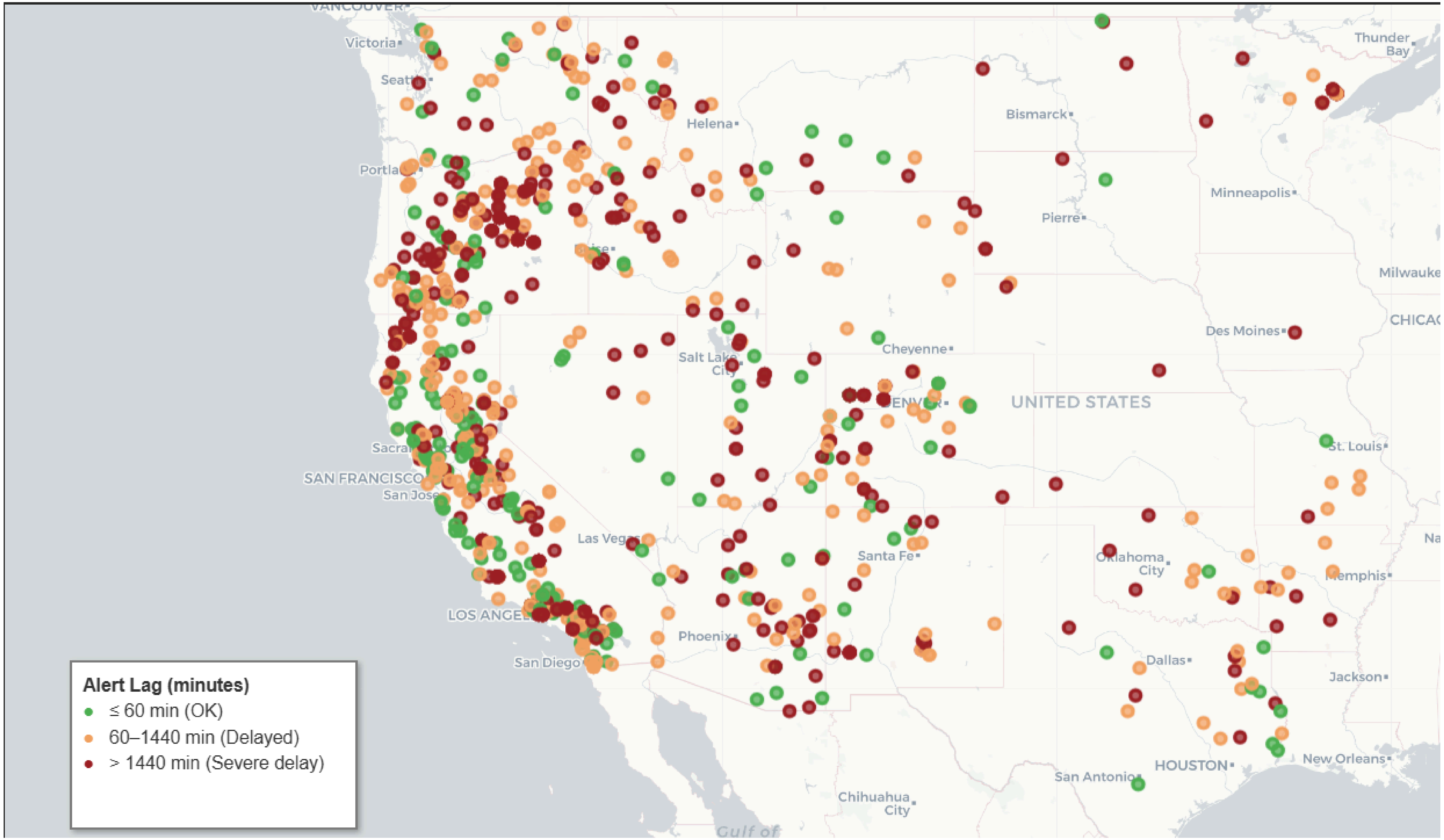
Alert messages in English/Spanish/Romanian

```
PLAYBOOK = {
    "LOW": {
        "EN": "Low risk. Monitor official updates and keep notifications enabled.",
        "RO": "Risc redus. Urmăriți actualizările oficiale și păstrați notificările active.",
        "ES": "Riesgo bajo. Siga las actualizaciones oficiales y mantenga las notificaciones activas."
    },
    "MEDIUM": {
        "EN": "Medium risk. Prepare for evacuation. Stay alert and review local guidance.",
        "RO": "Risc mediu. Pregătiți-vă pentru evacuare. Rămâneți în alertă și urmați indicațiile locale.",
        "ES": "Riesgo medio. Prepárese para evacuar. Manténgase alerta y siga las indicaciones locales."
    },
    "HIGH": {
        "EN": "High risk. If evacuation is ordered, leave immediately. Follow official instructions.",
        "RO": "Risc ridicat. Dacă există ordin de evacuare, plecați imediat. Urmați instrucțiunile oficiale.
        "ES": "Riesgo alto. Si hay orden de evacuación, salga inmediatamente. Siga las instrucciones oficial
    }
}
```

Small sample of the result

name	acreage	containment	rate_of_spread	structure_threat	spotting	alert_lag_min	p_delayed	alert_level	message_RO	message_EN	
Butler Fire	21058.00	100.0	None	None	None	4983.398080	0.999908	HIGH	Risc ridicat. Dacă există ordin de evacuare, p...	High risk. If evacuation is ordered, leave imm...	Riesgo alto. Si hay orden de evacuación, salga...
...	...	...	...	...	...	...	...	...	...	...	...
Hughes Fire	10425.00	100.0	None	None	None	167.492473	0.776546	HIGH	Risc ridicat. Dacă există ordin de evacuare, p...	High risk. If evacuation is ordered, leave imm...	Riesgo alto. Si hay orden de evacuación, salga...
Bahrman Fire	110.30	100.0	None	None	None	317.664085	0.683128	MEDIUM	Risc mediu. Pregătiți-vă pentru evacuare. Rămâ...	Medium risk. Prepare for evacuation. Stay aler...	Riesgo medio. Prepárese para evacuar. Manténga...
Vegetation Fire	0.50	100.0	None	None	None	10.062204	0.626235	MEDIUM	Risc mediu. Pregătiți-vă pentru evacuare. Rămâ...	Medium risk. Prepare for evacuation. Stay aler...	Riesgo medio. Prepárese para evacuar. Manténga...

Map of the alert lag



Conclusion

The supervised learning models demonstrate strong and consistent predictive performance in identifying delayed wildfire alerts. All evaluated models achieve an overall accuracy of approximately 92.7–92.8%, confirming that alert delays can be reliably predicted using the selected operational features. This indicates that the problem formulation is appropriate and that the chosen variables capture meaningful signals related to alert delays.

Among the evaluated models, Random Forest and XGBoost exhibit the strongest overall performance, with very similar results across all metrics. Both models achieve high precision (≈ 0.93) and exceptionally high recall (above 99%) for delayed alerts, indicating that nearly all truly delayed cases are correctly identified. This property is essential in an alerting context, where failing to detect delayed alerts represents the most costly type of error. In terms of discrimination power, Random Forest attains the highest ROC AUC (≈ 0.847), closely followed by XGBoost (≈ 0.846), while the Decision Tree records a lower value (≈ 0.836). Although the numerical differences are modest, they are consistent across ROC curves and evaluation metrics, highlighting the advantage of ensemble-based methods over a single-tree model.

The unsupervised clustering analysis identifies three statistically distinct wildfire risk profiles, each characterized by clearly different numerical patterns. One cluster corresponds to very large-scale incidents with near-total containment (≈ 99%) but the longest alert delays (over 16,000 minutes) and the highest delayed alert rate (above 98%), highlighting the operational complexity associated with mega-fires. A second cluster also exhibits high containment (≈ 99.1%) but significantly smaller average fire size (≈ 29,700 acres) and shorter alert delays (≈ 8,856 minutes), indicating faster alert handling when incident scale is reduced. The third cluster is defined by very low containment (≈ 23.7%), moderate fire size (≈ 8,468 acres), and long alert delays (over 13,300 minutes), demonstrating that low containment alone can drive substantial alert delays, independently of fire scale.

Population exposure and individual-level vulnerability are not directly modeled, and the system is designed to support decision-making rather than automate evacuation orders.

THANK YOU!