Localized Forest Fire Risk Prediction: A Department-Aware Approach for Operational Decision Support

Paper #3087

1 Clustering target

This section present figure 1. This figure shows more detailed on the target clustering applied in this study.

2 Class imbalanced

This section present figure 2. This figure shows more detailed on the imbalanced data problem faced in this study.

3 Features list

Categorized Features

This section shows the final list of features that have been used for predicting burned area and fire occurrence.
Meteorological:

- angstroem_max
- angstroem_mean
- angstroem_min
- bui_max
- bui_min
- dailySeverityRating_min
- days_since_rain_max
- days_since_rain_mean
- days_since_rain_min
- dwpt16_max
- dwpt16_min
- dwpt_max
- dwpt_min
- fwi_max
- fwi_min
- ffmc_max
- ffmc_mean
- ffmc_min
- isi_mean
- kbdi_max
- kbdi_min
- munger_max
- munger_mean
- munger_min
- nesterov_max
- nesterov_mean
- nesterov_min
- path_mean
- prec24h16_max
- prec24h16_min
- prec24h_max
- prec24h_min
- prcp16_max
- prcp16_min
- prcp_max
- prcp_min
- precipitationIndexN3_max
- precipitationIndexN3_min
- precipitationIndexN5_max
- precipitationIndexN5_mean
- precipitationIndexN5_min
- rhum16_max
- rhum16_mean
- rhum16_min
- rhum_max
- rhum_mean
- rhum_min
- snow24h16_max
- snow24h16_min
- snow24h_max
- snow24h_min
- sum_consecutive_rainfall_max
- sum_consecutive_rainfall_min
- sum_rain_last_7_days_max
- sum_rain_last_7_days_mean
- sum_rain_last_7_days_min
- sum_snow_last_7_days_max
- sum_snow_last_7_days_min
- temp16_max
- temp16_min
- temp_max
- temp_min
- wdir16_max
- wdir16_mean
- wdir16_min
- wdir_max
- wdir_mean
- wdir_min
- wspd16_max
- wspd16_mean
- wspd16_min
- wspd_max

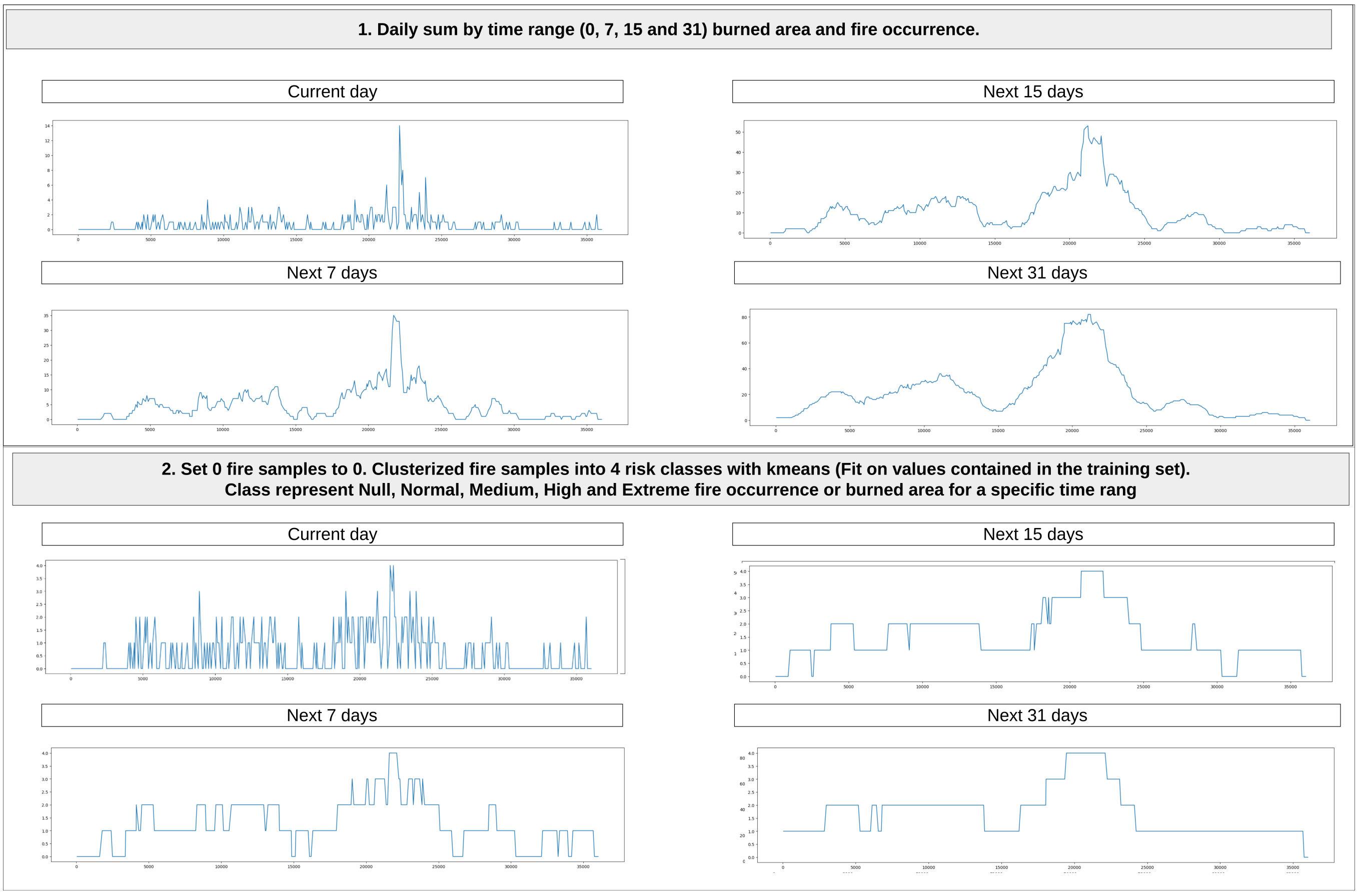


Figure 1: Process applied on each targets to obtain 5 classes risk. We use fire occurrence plots from Bouches du Rhone but same process has been applied at each department independently.

- wspd_mean
- wspd_min

Topographic:

- Bare soil_max
- Bare soil_mean
- Bare soil_min
- Chênes décidus_max
- Chênes décidus_mean
- Chênes sempervirents_max
- Conifères_max
- Conifères_mean
- Conifer_max
- Conifer_mean
- Crop_max
- Crop_mean
- Deciduous_max
- Deciduous_mean
- Douglas_max
- Douglas_mean
- Feuillus_max
- Feuillus_mean
- Feuillus_min
- Hêtre_max
- Hêtre_mean
- Lawn_max
- Lawn_mean
- Lawn_min
- Mélèze_max
- Mélèze_mean
- Mixtes_max
- Mixtes_mean
- NDMI_mean
- NDBI_max
- NDBI_min
- NDVI_max
- NDVI_mean
- NDSI_max
- NDSI_mean
- NDSI_min
- NDWI_max
- NC_max
- NC_mean
- NDVI_mean
- NDMI_mean
- NR_max
- PasDeRoute_max
- PasDeRoute_mean

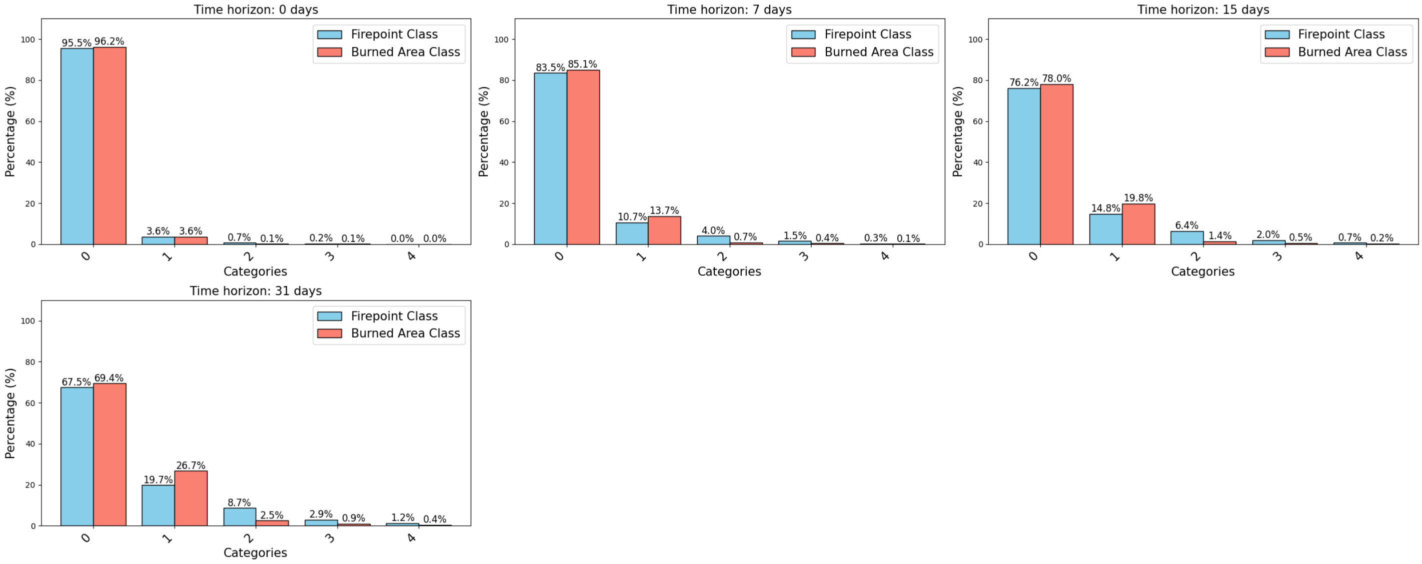


Figure 2: Distribution across time horizon in the test set for occurrence and burned area class. Classes 1, 2, 3, and 4 are outnumbered by class 0, especially classes 3 and 4, which make up barely 1.%.

- path_max
- path_mean
- Pin à crochets, pin cembro_max
- Pin d'Alep_max
- Pin laricio, pin noir_max
- Pin laricio, pin noir_mean
- Pin maritime_max
- Pin autre_max
- Pin sylvestre_max
- Pin sylvestre_mean
- Pins mélangés_max
- Pins mélangés_mean
- Peuplier_max
- Peuplier_mean
- Robinier_max
- Robinier_mean
- Sapin, épicéa_max
- Shrubland_max
- Shrubland_min
- argile_encoder_max
- argile_encoder_mean
- cosia_encoder_max
- cosia_encoder_mean
- cosia_encoder_min
- elevation_max
- elevation_mean
- elevation_min
- foret_encoder_max
- foret_encoder_mean
- foret_encoder_min

Socio-Economic:

- calendar_max
- calendar_sum
- confinement
- couvrefeux
- dayofweek
- holidays
- holidaysBorder
- isweekend
- month
- motorw...
- motorw...
- population_max
- population_mean
- population_min
- primary_max
- primary_mean
- secondary_max
- secondary_mean
- tertiary_max
- tertiary_mean
- ramadan

Historical:

- Past_burnedarea
- Past_risk
- cluster_encoder
- id_encoder

4 Length of time series

This section present the length of the time series selected for training and testing time models. Results are shows in 1.

5 Tree based configuration

Table 2 presents the configuration of tree based models.

6 Deep learning layers configuration

This section present the layers parameters for each deep learning model used in this study.

6.1 MLP

Table 3 shows the configuration of the MLP models.

6.2 GraphCast

Table 4 shows the configuration of the GraphCast model. Edges channels is set to 4 as in the original paper containing the x, y, z positions of the mesh node and the length of the edge. Similarly, the input mesh data is the 3D position of the node.

6.3 GRU

Table 5 shows the configuration of the GRU model.

6.4 LSTM

Table 6 shows the configuration of the LSTM models.

6.5 DilatedCNN

Table 7 shows the configuration of the DilatedCNN models.

6.6 GRUGNN

Table 8 shows the configuration of the GRUGNN models. 100 refers to the number of temporal features, 79 to the number of spatial features.

7 Comparis between binary models and multi classification

This section present the comparison between multi-classification and binary on Catboost, GRU and MLP models.

7.1 0 day horizon

Figure 3 shows the comparison between multi-classification and binary models at 0 day horizon.

7.2 7 day horizon

Figure 4 shows the comparison between multi-classification and binary models at 7 day horizon.

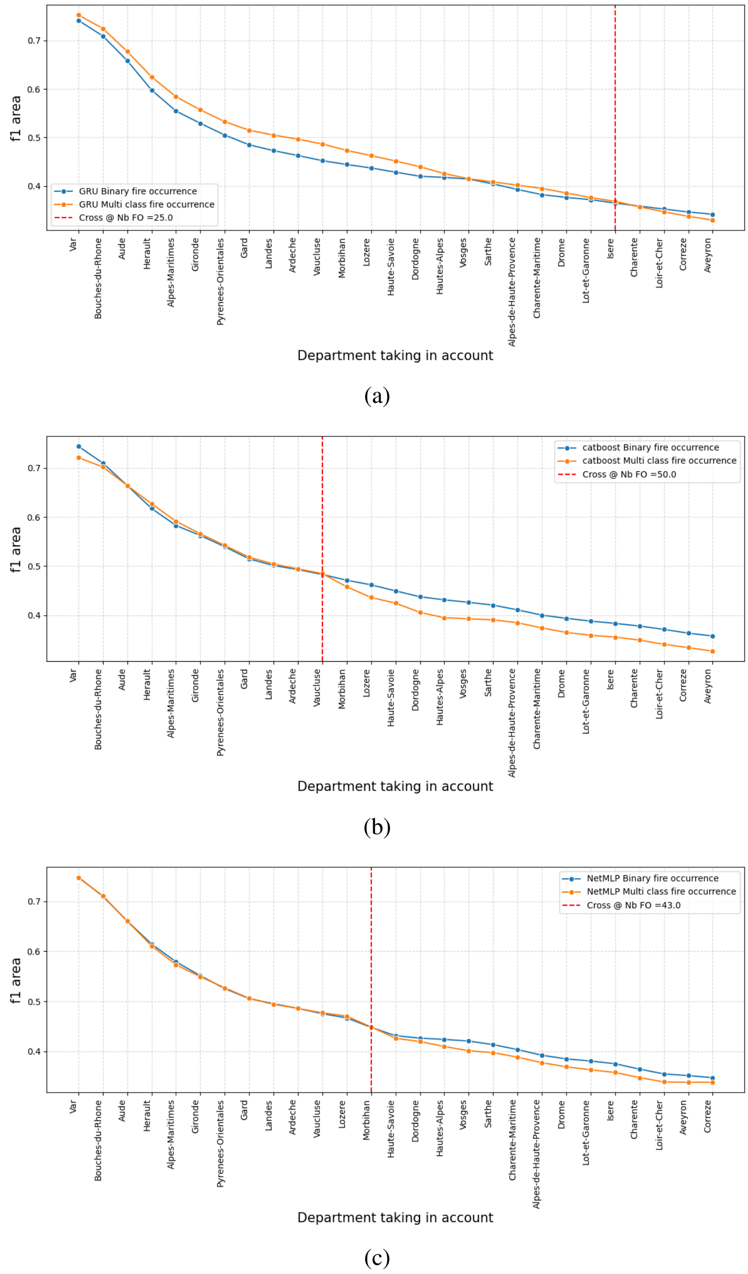


Figure 3: Comparison of multi-classification and binary F1 score area performance depending of the number of department for GRU (a), Catboost (b) and MLP (c) at 0 day horizon.

7.3 15 days horizon

Figure 5 shows the comparison between multi-classification and binary models at 15 day horizon.

Table 1: **Time series length selected for each model and prediction horizon**

Fire Occurrence				
Models	**0 days**	**7 days**	**15 days**	**31 days**
LSTM	10	5	15	5
GRU	10	15	15	5
GRU Binary	10	10	5	10
GRUGNN	10	5	5	5
DilatedCNN	15	10	15	15

Burned Area				
Models	**0 days**	**7 days**	**15 days**	**31 days**
LSTM	10	5	15	10
GRU	15	10	5	5
GRUGNN	10	10	5	5
DilatedCNN	15	10	15	5

Table 2: **Configuration of each tree based model**

Parameters	**XGBoost**	**Catboost**
Early Stopping Rounds	15	15
Learning Rate	0.001	0.001
Max Depth/Num Leaves	6	4
Min Child Weight	1.0	-
Max Delta Step	1.0	-
Subsample	0.5	0.7
Col Sample Bytree	0.7	0.6
Reg Lambda	1.7	1
Reg Alpha	0.7	0.27
N Estimators	10000	10000
Tree Method	hist	-

Table 3: MLP configuration

Linear 1	Linear 2	Linear 3	Linear 4
179, ReLU	179, ReLU	64, ReLU	5

Table 4: GraphCast configuration. PP refers to message passing

Grid2Mesh	Mesh MP	Mesh2Grid 3	Linear 1	Linear 2	Linear 3
179	179, 6 MLP MP layers	256	256	64	5

Table 5: GRU configuration

GRU	Norm 1	Dropout	Linear 1	Linear 2	Linear 3
2 layer, 0.03 dropout, 179, ReLU	BatchNorm	0.03	256, ReLU	64, ReLU	5

Table 6: LSTM configuration

LSTM	Norm 1	Dropout	Linear 1	Linear 2	Linear 3
1 layer, 64, ReLU	BatchNorm	0.03	64, ReLU	64, ReLU	5

Table 7: DilatedCNN configuration. Dil refers to dilation

Conv 1	Conv2	Conv 3	Norm, Dropout 1, 2, 3	Linear 1	Linear 2	Linear 3
179, Dil 1	179, Dil 3	Dil 3	Batchnorm, 0.03	64	64	5

Table 8: GRUGNN configuration.

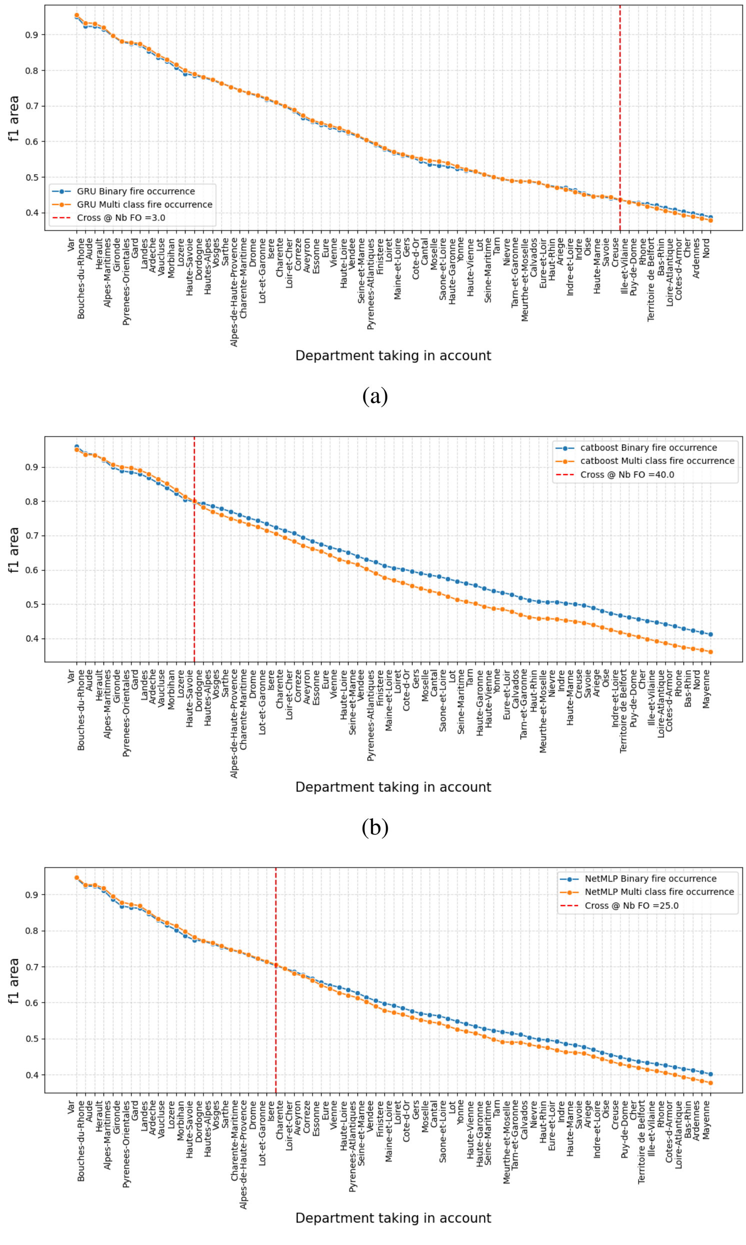
GRU	GCN	Norm, Dropout	Linear 1	Linear 2	Linear 3
2 layer, 0.03 dropout, 100	79	Batchnorm, 0.03	256	64	5

7.4 31 days horizon

Figure 6 shows the comparison between multi-classification and binary models at 31 day horizon.

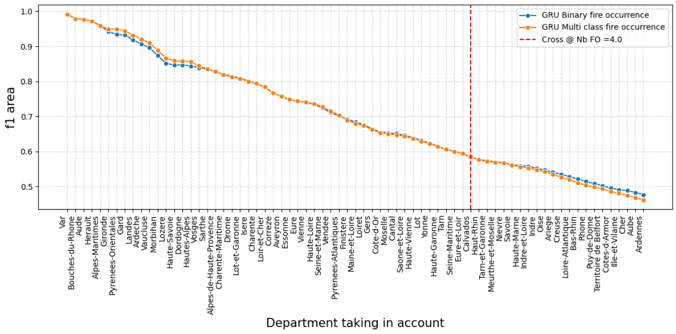
8 Features importance

This section features importance based on shap values for Catboost models. Figure 7 shows features importance of Burned area prediction of Catboost model.

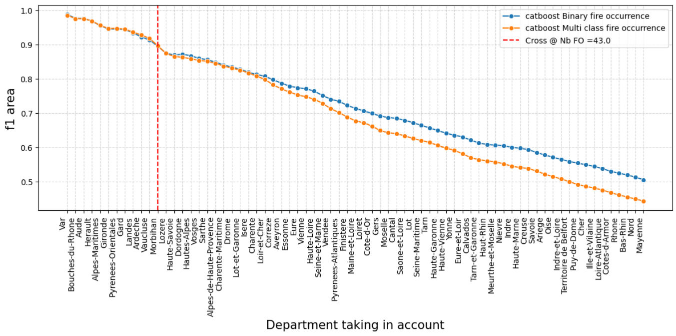


(a) (b) (c)

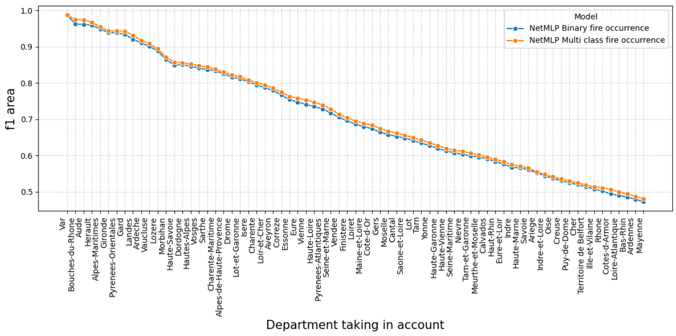
Figure 4: Comparison of multi-classification and binary F1 score area performance depending of the number of department for GRU (a), Catboost (b) and MLP (c) at 7 days horizon.



(a)

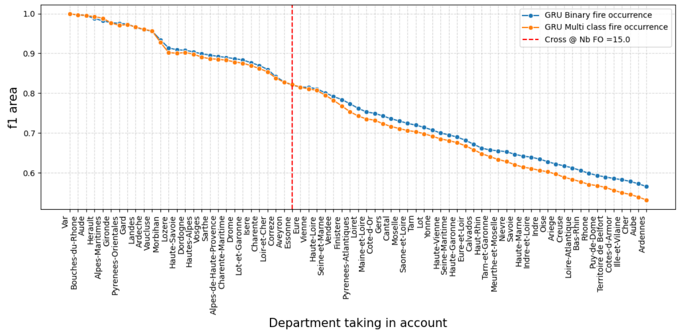


(b)

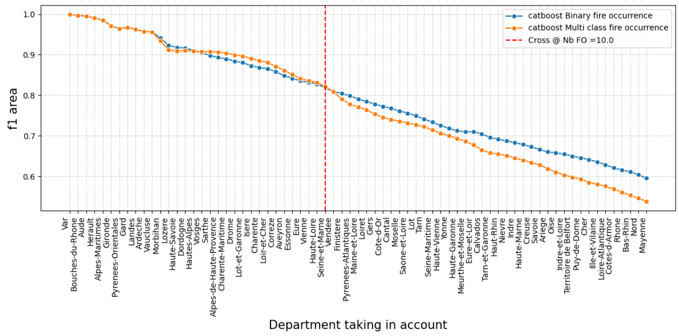


(c)

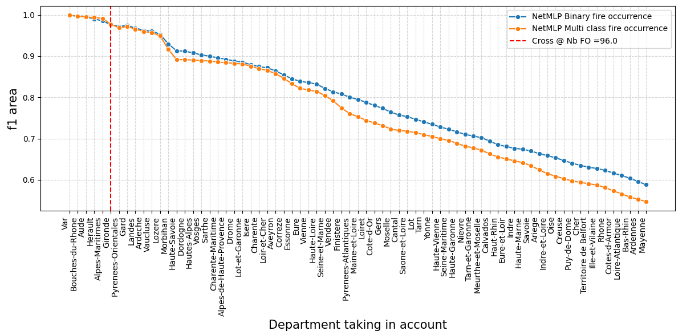
Figure 5: Comparison of multi-classification and binary F1 score area performance depending of the number of department for GRU (a), Catboost (b) and MLP (c) at 15 days horizon.



(a)



(b)



(c)

Figure 6: Comparison of multi-classification and binary F1 score area performance depending of the number of department for GRU (a), Catboost (b) and MLP (c) at 31 days horizons.

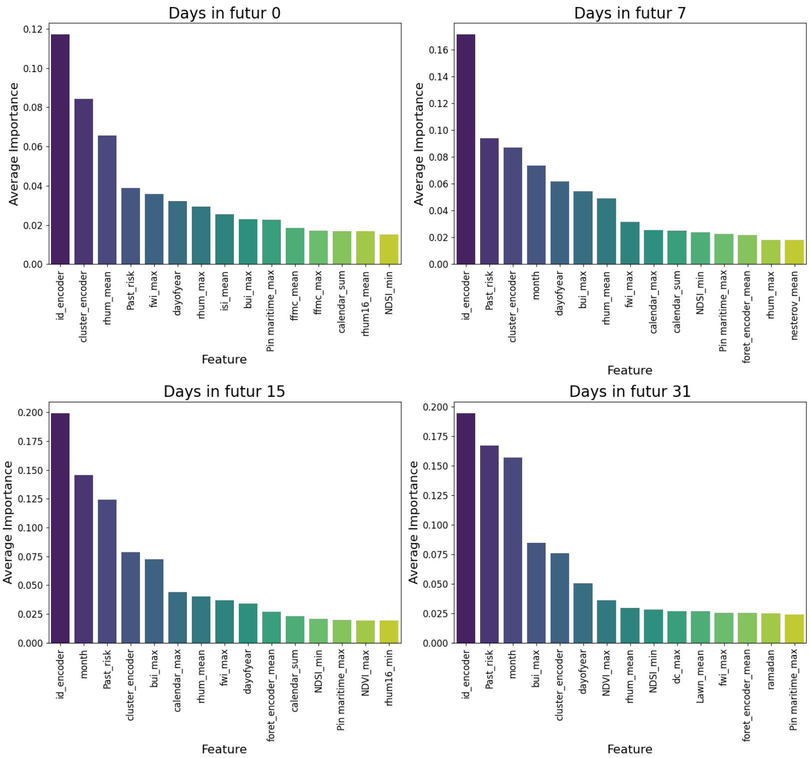


Figure 7: Top 15 shap values computed on multi-classification Catboost model on different time horizons for Burned area prediction. ID encoder correspond to Departmment ID