Using machine learning to predict fire-ignition occurrences from lightning forecasts

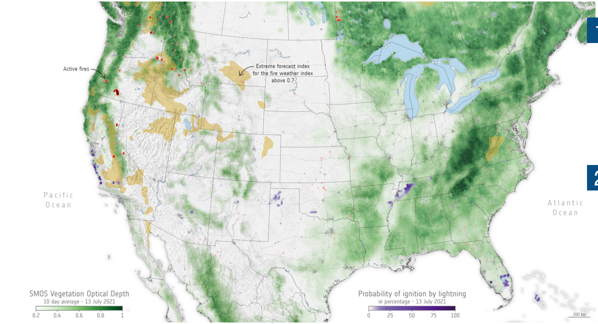
Francesca Di Giuseppe

Francesca.DiGiuseppe@ecmwf.int

European Centre for Medium-Range Weather Forecasts

Use of ESA data in aid of Copernicus Emergency Management Service-fire

Two new products have been prototyped using SMOS observations :



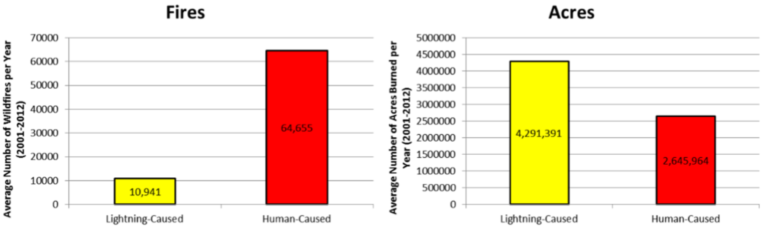
1. **probability of ignition by lighting.**
   Mostly dependent on the moisture content of fuel $\rightarrow$ a crucial predictor is observed soil moisture [1]

2. **vegetation available for burning**
   Strongly dependent on the available biomass $\rightarrow$ crucial predictor is Vegetation Optical depth[2]

[1] Coughlan et al 2021: Using machine learning to predict fire ignition occurrences from lightning forecasts. Meteorol Appl.

[2] Di Giuseppe, F., et al (2021). A global bottom-up approach to estimate fuel consumed by fires using above ground biomass observations. Geophysical Research Letters

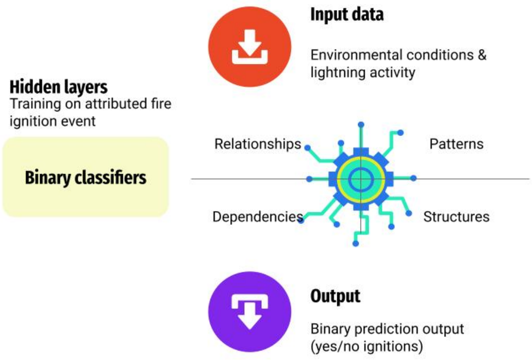
Lightning vs human ignited fires



- On average, there were more human-caused wildfires (64,655) per year than lightning-caused wildfires (10,941) in the U.S. from 2001-2012. However, on average there were more acres burned by lightning-caused wildfires (4,291,391) than by human-caused wildfires (2,645,964).
- There are around 800 lightning caused fires a year in the US

Data source : National Interagency Fire Centerhttps://www.nifc.gov/fire-information/statistics/human-caused

A machine learning framework for lightning ignited fires

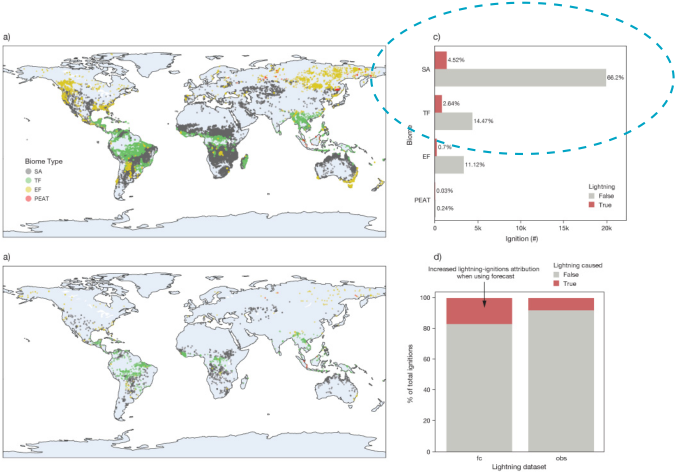


- Dry lightning are more likely to ignite a fire then lightning associated with heavy precipitation
- Key predictors of natural ignitions are fuel availability and flammability related to fuel moisture content.

A dataset of lightning ignited fires

Any newly active fires with no fire spread and not being agriculture related are deemed Ignition points.
Depending on the Cloud-to-Ground lightning density , points identified as IPs (6) are then split into caused by lightning or not.

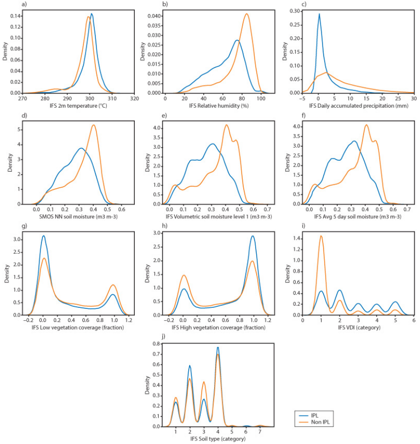
Global distribution of ignition points (2020)



- 60 % of fire ignitions happen in Savanna, but in proportion more lightning caused fires in forested areas
- Using lightning observations for attribution leads to 8% contribution rom lightning this would double (16 %) if attribution made with ECMWF forecast

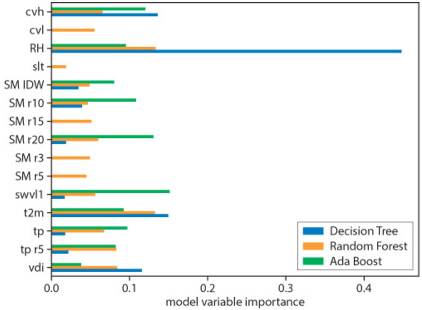
Environmental predictor selection

Discrimination potential based on "shift" in PDF for lightning ignition



Separation between PDF is a measure of discrimination.
→ Soil moisture statistically dryer when ignition occurs.
→ Temperature less relevant

Feature Importance for lightning

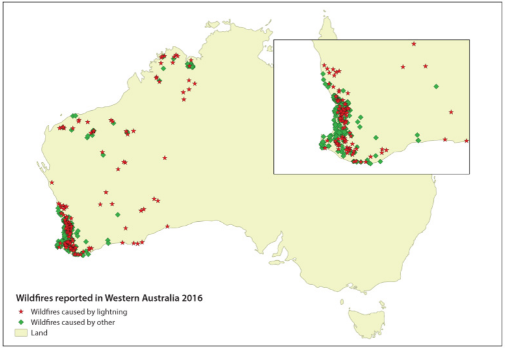


→ Automatic method to rank the importance of predictors for the accuracy of the final prediction
→ 78 % accuracy

Model	TN	FP	FN	TP	Specificity	Precision	Sensitivity	Accuracy	F1
Decision Tree	568217	161881	3854	4735	77.83	2.09	55.13	66.48	5.41
AdaBoost	564419	165679	2698	5891	77.31	2.72	68.59	72.95	6.54
Random Forest	576275	153823	3024	5565	78.93	2.67	64.79	71.86	6.63

Verification in real cases

Data provided by a Western Australia wildfire database, allowed a comprehensive verification on over 145 lightning ignited wildfires in 2016



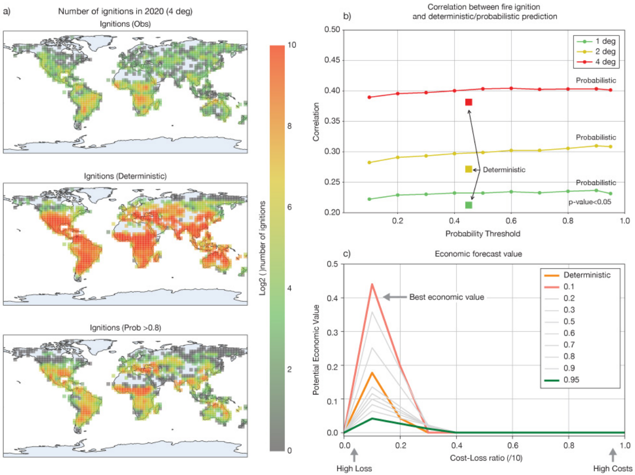
Model	TP	FN	Sensitivity
Decision Tree	110	37	0.75
Random Forest	111	36	0.76
AdaBoost	104	43	0.71

Sensitivity for the three models assuming a perfect lightning forecast on the discovery day.

This was reduced to around 50% when adding the uncertainties on the lightning forecast

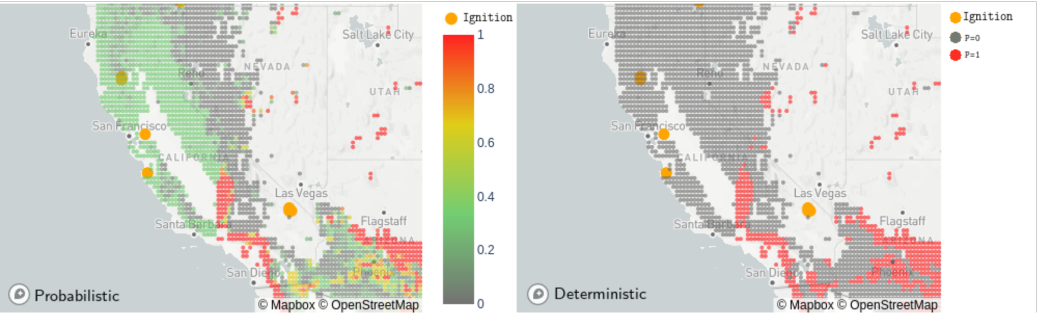
The benefit of a probabilistic approach

Comparing observed total fire ignitions with prediction



- → Fire ignition is a stochastic process and its prediction should rely on probabilistic approaches.
- → Using the probability of ignition rather than a binary outcome maximises the correlation between prediction and observed ignitions.
- → The probabilistic approach provides the best value-for-money when a potential economical value is attributed to the forecast.

What this means in a real case?



16 August 2020- California ignitions

A series of lightning ignition events happened in California in 2020. While in this region most of fires are connected to human activities, between August 16 and August 20 a siege of dry lightning from rare, intense summer thunder storms were the attributed to around 650 ignitions.

Conclusions

Value in remote observations to detect fire ignitions

Observed soil mosisture provides a key predictor for lightning ignited fire prediction

Opportunity for ML methods to predict inherently stochastic processes such as ignition

Very high out-of sample accuracy of 75% under certain conditions. However the use of lightning forecast could reduce this accuracy due to forecast errors

Optimal strategy is in the combination of ML and probabilistic methods

Using the probability of ignition rather than a binary outcome maximises the correlation between prediction and observed ignitions. The probabilistic approach provides the best value-for-money when a potential economical value is attributed to the forecast.