A data-driven model for Fennoscandian wildfire danger

Sigrid J. Bakke, Niko Wanders, Karin van der Wiel, and Lena M. Tallaksen

MOTIVATION

Wildfires are typically hard to predict, as their exact location and occurrence are driven by a variety of factors. Data-driven (machine-learning) models can identify dominant factors of complex and partly unknown processes, and can ultimately improve predictions and projections of wildfires in both the current and a future climate.

WHAT DID WE DO?

We developed a temporally and spatially explicit data-driven model for Fennoscandia to reach two main objectives:

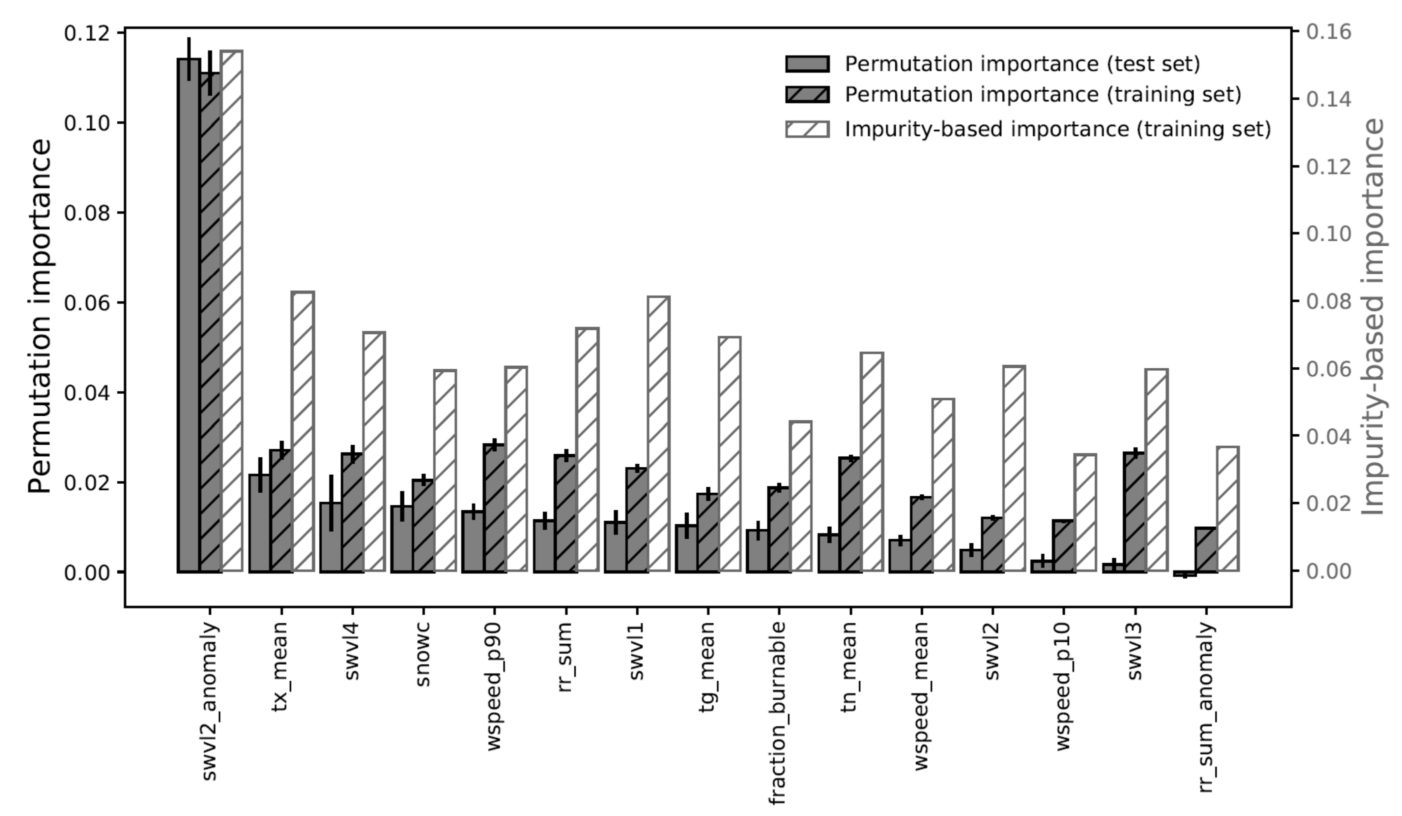
- **identify dominant predictors of wildfires**
- **construct monthly fire danger probability maps**

We evaluated our model by comparing its performance (test set ROC-AUC) with that of the established fire danger index FWI (Canadian forest fire weather index).

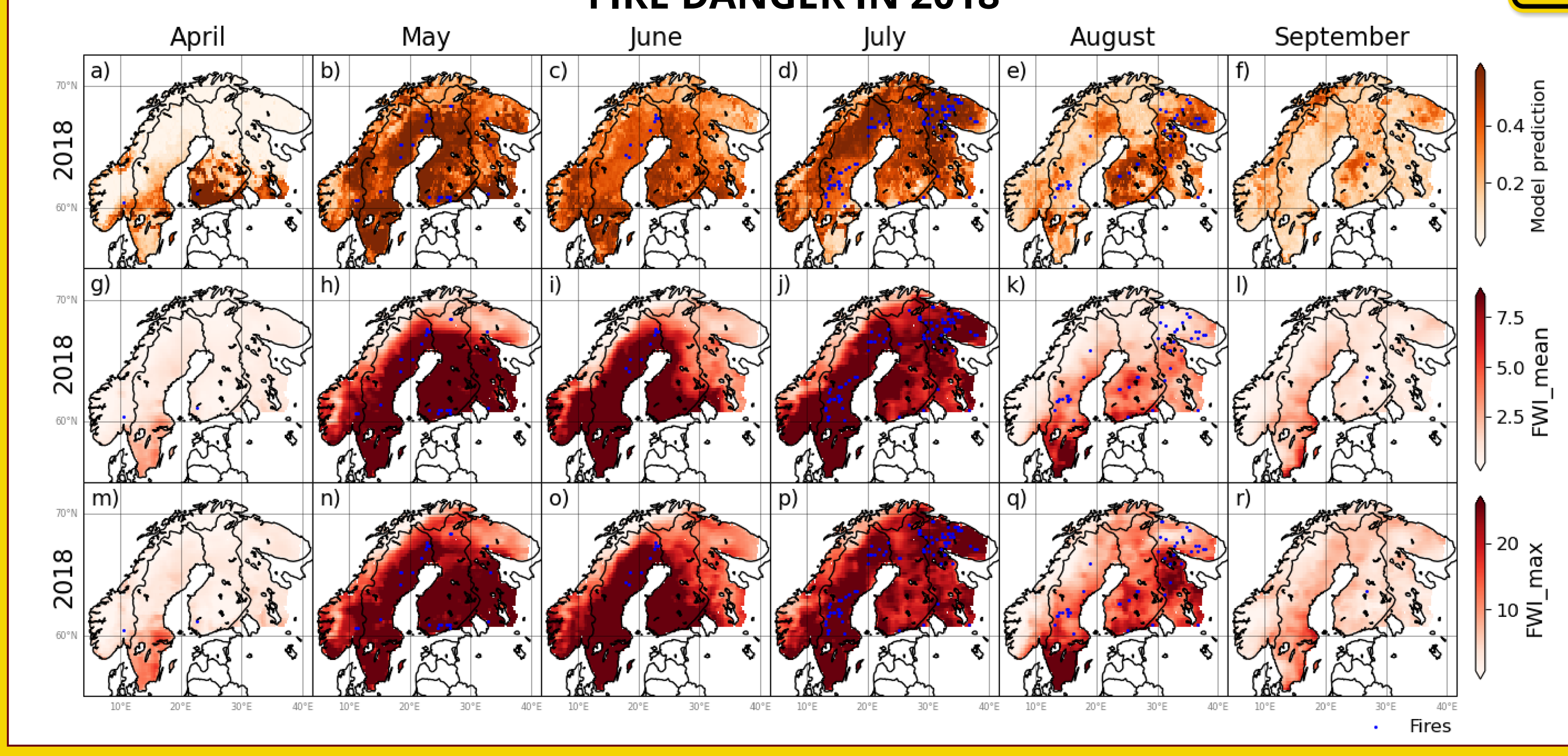
KEY FINDINGS

- The dominant predictor of wildfire is shallow soil moisture anomaly
- The predictors emphasise the importance of other predictors than weather alone, as has traditionally been used for fire danger indices
- The model produced somewhat different monthly fire danger maps as compared to FWI
- The model slightly outperformed FWI with an ROC-AUC of 0.79 vs 0.78 for FWI.

IMPORTANCE OF MODEL PREDICTORS



FIRE DANGER IN 2018



DATA AND METHODS

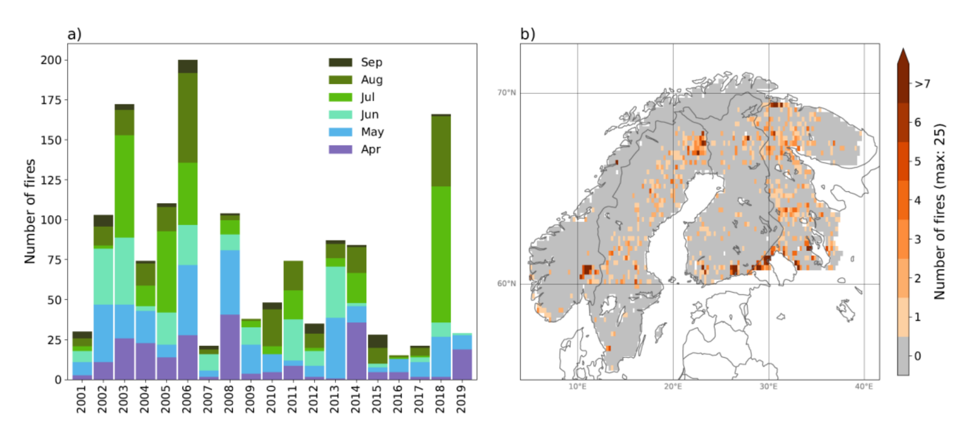
Target data:
Monthly fire occurrence from v5.1.1cds burned area product. The data is extremely imbalanced, with only 1439 of the 444 030 data points (0.3 %) classified as fire.

Potential predictors:
Hydrometeorological indices based on data from E-OBS, ERA5-Land and v5.1.1cds. We chose only predictors that are available in most climate models and transferable to different climate scenarios.

Machine learning algorithm:
The data-driven model was trained by using the Random Forest probability classifications, in which target data were weighted inversely proportional to the class frequencies.

Splitting of the data:
Whole years were assigned to the training set (14 years) and test set (5 years), as well as to each cross-validation fold (2 years), to reduce the dependencies between data points that could lead to too optimistic results.

TARGET DATA: FIRE OCCURRENCES



Model training:
We applied 7-fold cross-validation to tune the complexity parameter maximum tree depth and the number of predictors using backward elimination with updated permutation importances.

Model evaluation:
ROC-AUC (the area under the curve of the receiver operating characteristic). It tackles extreme imbalanced data and enable comparison of FWI and our model without a preset classification threshold. ROC-AUC calculates the area under the curve of true-positive rate vs. false-positive rate for different classification thresholds.

POTENTIAL PREDICTORS

Precipitation	rr_sum	Monthly precipitation sum
	rr_sum_anomaly	Anomalies of rr_sum
Temperature	tg_mean, tn_mean and tx_mean	Monthly mean of daily mean, daily minimum and daily maximum temperature
	tx_max	Monthly maximum of daily maximum temperature
	tg_mean_anomaly, tn_mean_anomaly and tx_mean_anomaly	Anomalies of tg_mean, tn_mean and tx_mean
Meteorological drought	SPI2, SPI3, SPI6 and SPI9	SPI [−3,3] over 2, 3, 6 and 9 months, calculated from rr_sum
	SPEI2, SPEI3, SPEI6 and SPEI9	SPEI [−3,3] over 2, 3, 6 and 9 months, calculated from rr_sum minus monthly potential evapotranspiration, calculated based on tg, tn and tx
Wind speed	wspeed_mean	Monthly mean 10 m wind speed
	wspeed_p10 and wspeed_p90	Monthly 10th and 90th percentile of daily 10 m wind speed
Snow	snowc	Monthly average fraction of grid cell occupied by snow
Soil moisture	swvl1, swvl2, swvl3 and swvl4	Monthly mean volumetric soil water in soil layer 1 (0–7 cm), layer 2 (7–28 cm), layer 3 (28–100 cm) and layer 4 (100–289 cm)
	swvl1_anomaly, swvl2_anomaly, swvl3_anomaly and swvl4_anomaly	Anomalies of swvl1, swvl2, swvl3 and swvl4
Land cover	fraction_burnable	Fraction of the cell corresponding to vegetated land covers that could burn

Natural Hazards and Earth System Sciences
Open Access

Bakke et al. (2023). DOI: 10.5194/nhess-23-65-2023