Machine learning models of forest vulnerability to fires, windthrows and insect outbreaks

Giovanni Forzieri (giovanni.forzieri@ec.europa.eu, gforzieri@gmail.com)
European Commission, Joint Research Centre, Ispra, Italy

Forest disturbance regimes are expected to intensify as Earth's climate changes. We investigated the vulnerability of European forests to fires, windthrows and insect outbreaks during the period 1979-2018 by integrating machine learning with disturbance data and satellite products. The proposed methodology is purely data-driven and therefore reproducible, applicable at large scales, and in line with the measurement/reporting/verification process of UNFCCC. Results of this study have been published in Forzieri et al. (2021).

Pre-processed data, codes and final vulnerability models developed in the afore-mentioned work are made publicly available here and briefly described to facilitate reproducibility and applicability.

Data and code organization

Pre-processed input data and computing codes for model development are stored in the folder ModelDevelopment. These routines allow for the reproducibility of the analyses and figures described in Forzieri et al. (2021). Vulnerability models are stored in the folder VulnerabilityModels. Models can be employed for user-specific applications; however, caution should be used in extrapolation outside the environmental ranges of disturbance records. All codes are written in MATLAB. Data and codes are organized as follows.

ModelDevelopment	disturbance[1]	code	Computing scripts
		input	Biomass losses and environmental features in disturbance records (disturbance_observed.xlsx[3]) Setup files
		output	Spatial maps of environmental features required to run vulnerability models in predictive mode
		disturbance_main.m	Main code to set/launch computing scripts
	multi[2]	code	Computing scripts
		input	Setup files
		output	Spatial maps used for binning in the climate space
		multi_main.m	Main code to set/launch computing scripts
	rasters	input	Spatial masks used for zonal statistics
VulnerabilityModels	disturbance[1]	disturbance_RFR_pfts[4]	Random Forest Model implemented as TreeBagger Matlab object (ensemble of bagged regression trees)

[1] natural disturbance: fire (wildfires), wind (windthrows), insect (insect outbreaks)

[2] multi: OVI (overall vulnerability index)

[3] data collection generated in R (version 4.0.1)/Python (version 3.8.5), contact: marco.girardello@ec.europa.eu. Environmental features are listed in Table 1. The biomass losses are retrieved from Google Earth Engine, contact: guido.ceccherini@ec.europa.eu.

[4] pfts: vulnerability models are derived for the whole mixture of multiple plant functional types (ALL) and for single plant functional types (BrDe broadleaved deciduous; BrEv broadleaved evergreen; NeDe needle leaf deciduous; NeEv needle leaf evergreen).

How to run the codes

Run the main code

Download the ModelDevelopment folder by preserving the paths of its subfolders. Launch the main code from the MATLAB command window. The main code will generate the output files and will save them automatically in dedicated subfolders. All analyses described in Forzieri et al. (2021) will be performed automatically.
Example: >> fire_main.m

Run single subroutines

Each single subroutine can be run separately as long as the preceding subroutines have been successfully completed. Global variables defined in the main code need to be recalled before running a given subroutine. The subroutine has to be launched from the path of the main code.

Run scripts for figure generation

Scripts for figure generation can be run as long as the required input data have been successfully generated from the main code or its subroutines. The variable named "pathOut" needs to be set manually as the main path where the ModelDevelopment folder is saved.

Run vulnerability models over new study domains

Example:
load fire_RFR_BrDe
[Yfit,stdevs] = predict(B,X);

- Output. It returns a vector of predicted vulnerabilities Yfit for the predictor data in the table or matrix X, based on the TreeBagger object B (ensemble of bagged regression trees). stdevs are standard deviations of the computed vulnerabilities over the ensemble of the grown tree.
- Input. B: TreeBagger Matlab object stored in the VulnerabilityModels folder. X: predictor data organized in table or matrix. Each column refers to a different predictor. The order of predictors (columns) in X has to be consistent with the order of predictors in the TreeBagger object (B.PredictorNames) and with its units of measure (table 1).

Category	Variable	Reference variable	Unit of measure of reference variable
Forest	Above ground biomass (Biomass)	bio	t ha^1
	Tree density	treedens	Number of trees per $1km^2$
	Tree age	age	years
	Leaf Area Index (LAI)	LAI	m^2 m^{-2}
	Tree height	height	m
Climate	Fire Weather Index (FWI)	FWI.le.1	days
	Moisture Index (MI)	AImin	mm $°C^{-1}$
	Cumulated precipitation (Pcum)	EvScale.ppt.1	mm
	Cumulated snow (Snow)	EvScale.sn.1	cm
	Short-term average anomaly in cumulated precipitation (avg aPcum)	anomean.EvScale.ppt.1	mm
	Maximum temperature (Tmax)	EvScale.tmax.1	°C
	Long-term average temperature (Long-term Tavg)	LongTerm.tmean.1981.2017.1/ NA.LongTerm.tmean.1981.2017.1	°C
	Short-term average anomaly in average temperature (avg aTavg)	anomean.EvScale.tmean.1	°C
	Maximum wind speed (Wind speed)	wind.le.1	m s^{-1}
	Short-term anomaly in Standardized Precipitation Evapotranspiration Index (avg SPEI)	mean.EvScale.SPEI12	-
Landscape	Slope	slope	-
	Elevation	altitude/dem	m.a.s.l.
	Homogeneity	homog.evi	-
	Coefficient of variation (CV)	cv.evi	-

Table 1. Environmental variables selected in the vulnerability models. Forest, climate and landscape features utilized to characterise the response functions of vulnerability to natural disturbances. Abbreviations used in text and figures of Forzieri et al. (2021) are in parentheses. Reference variable reports the corresponding variable name used in ModelDevelopment and VulnerabilityModels.

Running setup
All MATLAB codes have been ran under the following setup.

MATLAB Version: 9.3.0.713579 (R2017b)
MATLAB License Number: 702352
Operating System: Microsoft Windows 10 Enterprise Version 10.0 (Build 16299)
Java Version: Java 1.8.0_121-b13 with Oracle Corporation Java HotSpot(TM) 64-Bit Server VM mixed mode

MATLAB	Version 9.3	(R2017b)
Curve Fitting Toolbox	Version 3.5.6	(R2017b)
Image Processing Toolbox	Version 10.1	(R2017b)
Mapping Toolbox	Version 4.5.1	(R2017b)
Neural Network Toolbox	Version 11.0	(R2017b)
Parallel Computing Toolbox	Version 6.11	(R2017b)
Statistics and Machine Learning Toolbox	Version 11.2	(R2017b)

Data citation
Forzieri, G., Girardello, M., Ceccherini, G., Spinoni, J., Feyen, L., Hartmann, H., Beck, P.S.A., Camps-Valls, G., Chirici, G., Mauri, A., Cescatti, A., Emergent vulnerability to climate-driven disturbances in European forests. *Nature Communications*, 2021.