Applying machine learning to assess socio-economic wildfire risks

Carmen B. Steinmann[1,2], Jonathan Koh[3], Samuel Lüthi[1,2], Samuel Gübeli[1], Benoît P. Guillod[4], David N. Bresch[1,2]

[1]Institute for Environmental Decisions, ETH Zurich; [2]Federal Office of Meteorology and Climatology MeteoSwiss; [3]Institute of Mathematical Statistics and Actuarial Science, Oeschger Centre for Climate Change Research, University of Bern; [4]CelsiusPro AG, Zurich

Motivation

Wildfires cause extensive damage to physical assets exposed to them. So far, assessing the risk of these events remains an understudied area of global disaster risk assessment (Ward et al., 2020). Probabilistic risk estimates covering the range and likelihood of devastating events are crucial for various applications such as prioritizing adaptation measures and determining insurance pricing. In parallel, increasingly available data allows for the use of machine learning techniques to predict wildfire behaviour (Koh, 2023). In this context, a globally consistent, open-source wildfire risk model facilitates the accessibility of such analysis to stakeholder from both the public and private sector.

Method overview

In a first step we set up a statistical wildfire hazard model. Then, we build upon the open-source climate risk assessment platform CLIMADA (Aznar-Siguan and Bresch, 2019) to compute socio-economic impacts as the combination of the newly developed hazard, an exposure and its vulnerability (Fig 1).

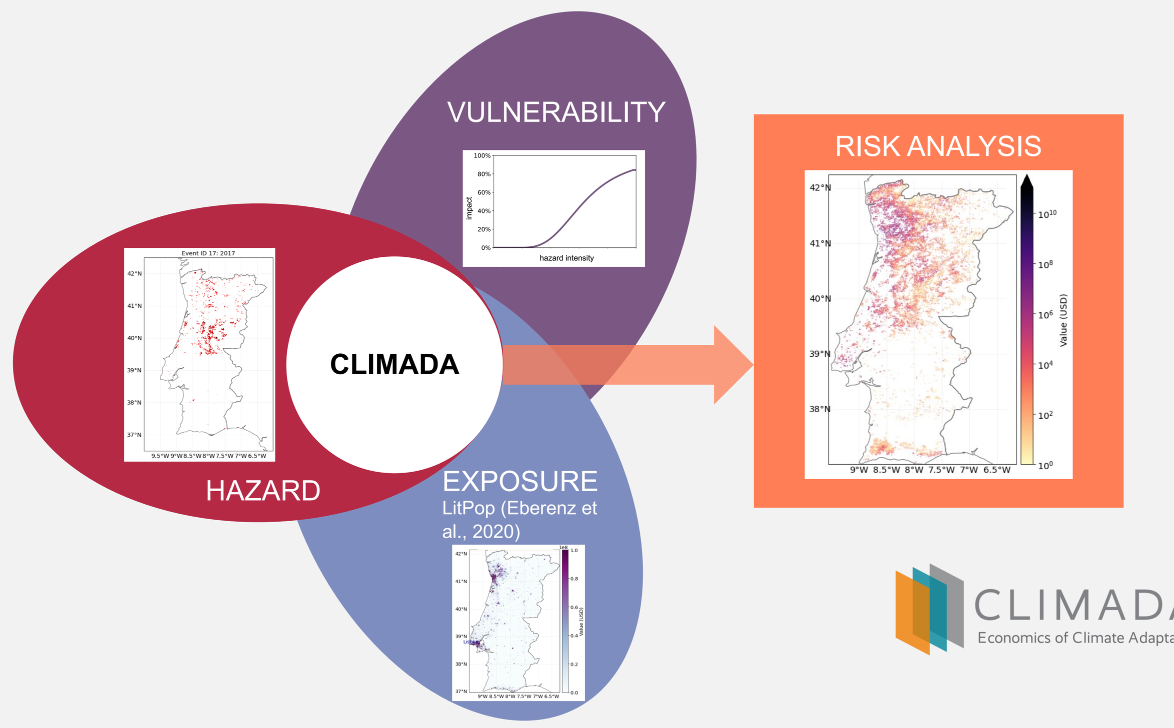

Fig. 1 The three components hazard, exposure and vulnerability as used in CLIMADA to compute wildfire risks

I. Statistical wildfire hazard model

We set up a country-specific wildfire model predicting the daily fraction of burnt area (0-1) per grid cell (4x4 km) based on covariates derived from open-source data.

Response: MODIS burnt area

Covariates

- Land use fractions – based on Copernicus Global Land Service Land Cover (Buchhorn et al, 2019)
- Gridded population (Center for International Earth Science Information Network - CIESIN, 2018)
- Daily and monthly maximum vapor pressure deficit – based on ERA5 (Hersbach et al., 2023)

Feature Engineering

- Implicit propagation: Fraction of MODIS burnt area in neighbouring cells in preceding time steps
- Fragmentation: Average number of neighbouring cells sharing the same land use type – based on land cover maps with a resolution of 100m (Buchhorn et al, 2019)

Method

We make use of a machine learning model based on the efficient regression tree boosting system XGBoost, which also gives a measure of importance of each covariate (Koh, 2023).

II. Socio-economic impact and risk assessment

Impacts are computed by combining a hazard, exposure and vulnerability. The vulnerability describes the historic relationship between wildfires and caused impacts. It is derived by combining historic hazard intensities, an exposure and damage records. Lüthi et al. (2021) deduct the vulnerability for MODIS hotspots. In this study, we transfer this approach using MODIS burnt area as hazard intensity. The used exposure layer LitPop globally consistently disaggregates asset value data proportional to a combination of nightlight intensity and geographical population data (Eberenz et al., 2020). The damage records are extracted from the International Disaster Database EM-DAT (Guha-Sapir et al., 2021). The derived vulnerability can then be used to compute both historic and probabilistic impacts and thereon deduct common risk metrices such as impact return period curves and the average annual impact.

References

- Aznar-Siguan, G. and D.N. Bresch. 2019, 7. CLIMADA v1: a global weather and climate risk assessment platform. Geoscientific Model Devel- opment 12(7): 3085–3097.
- Buchhorn, Marcel, Bruno Smets, Luc Bertels, Bert De Roo, Myroslava Lesiv, Nandin-Erdene Tsend-bazar, Martin Herold, and Steffen Fritz (Sept. 2020). “Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe”.
- Center for International Earth Science Information Network - CIESIN - Columbia University (2018). “Gridded Population of the World, Version 4 (GPWv4): Population Count”. In: Revision 11. Pal-isades, New York: NASA Socioeconomic Data and Applications Center (SEDAC).
- Eberenz, S., D. Stocker, T. Röösli, and D. N. Bresch (2020): Asset exposure data for global physical risk assessment. Earth System Science Data 12 (2), pp. 817-833, 2020.
- Guha-Sapir, D., R. Below, and P. Hoyois. 2021. EM-DAT: The CRED/OFDA International Disaster Database - www.emdat.be – Universit ́e Catholique de Louvain – Brussels – Belgium
- Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2023): ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS)
- Koh, J. Gradient boosting with extreme-value theory for wildfire prediction. Extremes(2023).
- Lüthi, S., Aznar-Siguan, G., Fairless, C., and Bresch, D. N.: Globally consistent assessment of economic impacts of wildfires in CLIMADA v2.2, Geosci. Model Dev., 14, 7175–7187, 2021.
- Ward, P J, V Blauhut, N Bloemendaal, J E Daniell,MC de Ruiter,MJ Duncan, R Emberson, S F Jenkins, D Kirschbaum, M Kunz, S Mohr, S Muis, G A Riddell, A Sch.fer, T Stanley, T I E Veldkamp, and H C Winsemius (2020). “Review article: Natural hazard risk assessments at the global scale”. In: Natural Hazards and Earth System Sciences 20.4, pp. 1069–1096.

Partner/Sponsor: MeteoSwiss, Innosuisse, °CelsiusPro

Carmen B. Steinmann
Doctoral Student
Weather and Climate Risks
carmen.steinmann@usys.ethz.ch