EGU21-12930
https://doi.org/10.5194/egusphere-egu21-12930
EGU General Assembly 2021

Near real-time identification of extreme events for weather index insurance using machine learning algorithms

Luigi Cesarini[1], Rui Figueiredo[2], Beatrice Monteleone[1], and Mario Martina[1]
[1]Scuola Universitaria Superiore IUSS, Pavia, Italy
[2]CONSTRUCT-LESE, Faculty of Engineering, University of Porto, 4200-465 Porto, Portugal

A steady increase in the frequency and severity of extreme climate events has been observed in recent years, causing losses amounting to billions of dollars. Floods and droughts are responsible for almost half of those losses, severely affecting people's livelihoods in the form of damaged property, goods and even loss of life. Weather index insurance is an innovative tool in risk transfer for disasters induced by natural hazards. In this type of insurance, payouts are triggered when an index calculated from one or multiple environmental variables exceeds a predefined threshold. Thus, contrary to traditional insurance, it does not require costly and time-consuming post-event loss assessments. Its ease of application makes it an ideal solution for developing countries, where fast payouts in light of a catastrophic event would guarantee the survival of an economic sector, for example, providing the monetary resources necessary for farmers to sustain a prolonged period of extreme temperatures. The main obstacle to a wider application of this type of insurance mechanism stems from the so-called basis risk, which arises when a loss event takes place but a payout is not issued, or vice-versa.

This study proposes and tests the application of machine learning algorithms for the identification of extreme flood and drought events in the context of weather index insurance, with the aim of reducing basis risk. Neural networks and support vector machines, widely adopted for classification problems, are employed exploring thousands of possible configurations based on the combination of different model parameters. The models were developed and tested in the Dominican Republic context, leveraging datasets from multiple sources with low latency, covering a time period between 2000 and 2019. Using rainfall (GSMaP, CMORPH, CHIRPS, CCS, PERSIANN and IMERG) and soil moisture (ERA5) data, the machine learning algorithms provided a strong improvement when compared to logistic regression models, used as a baseline for both hazards. Furthermore, increasing the number of information provided during model training proved to be beneficial to the performances, improving their classification accuracy and confirming the ability of these algorithms to exploit big data. Results highlight the potential of machine learning for application within index insurance products.