EGU23-11434, updated on 19 Apr 2024
https://doi.org/10.5194/egusphere-egu23-11434
EGU General Assembly 2023

Evaluating the explainability and performance of an elementary versus a statistical impact-based forecasting model

Sahara Sedhain[1], Marc van den Homberg[2], Aklilu Teklesadik[2], Maarten van Aalst[1,3], and Norman Kerle[1]

[1]University of Twente, Faculty of Geo-information Science and Earth Observation (ITC), Netherlands
[2]Netherlands Red Cross, 510 humanitarian data team, Netherlands
[3]Red Cross Red Crescent Climate Centre, Netherlands

The disaster risk community has notably shifted from a response-driven approach to making informed anticipatory action choices through impact-based forecasting (IBF). Algorithms are being developed and improved to increase impact prediction abilities, and to allow automatic triggers to reduce the reliance on human judgement. However, as complexities in modelling algorithms increase, it becomes more difficult for decision makers to interpret and explain the results. This reduces the accountability and transparency, and can lead to lower adoption of the models. Therefore, humanitarian decision-makers can benefit from a mechanism to evaluate different IBF approaches, which has not yet been developed. Through a case study of anticipatory action for tropical cyclones in the Philippines, we evaluated two very different approaches to IBF: (1) a statistical trigger model that uses a machine learning algorithm with several predictor variables, and (2) an elementary trigger model that combines damage curves and weighted overlay of vulnerability indicators, to predict the impact and prioritize areas for intervention. The models were evaluated based on their performance for damage prediction and their sensitivity to different risk indicators for Typhoon Kammuri (2019) in the Philippines. The study also proposed a way of characterising the explainability specific to an IBF model, and that gives clarity on which elements, and why, influence the results, done via a model card. To facilitate this process a prototype interactive decision portal was built, which shows decision makers the sensitivity of the results to variations in input parameters. The results show that in relative terms the elementary model performed better and would have allowed to maximise impact reduction through early action, suggesting that, for this particular case, complex was not necessarily a better choice. However, the uncertainty in both models due to limitations in the initial hazard forecast indicates that multiple models need to be evaluated for practical cases that cover different characteristics of the hazard and socio-vulnerable situations. For this, the evaluation framework we developed can be expanded across operational IBF projects.