

A data mining approach to derive flood-related economic vulnerability of companies

Tobias Sieg (1,2), Heidi Kreibich (1), Kristin Vogel (2), Bruno Merz (1,2)

(1) GFZ German Research Centre for Geosciences, Section 5.4 Hydrology, Telegrafenberg, 14473 Potsdam, Germany, (2) University of Potsdam, Institute of Earth and Environmental Science, Karl-Liebknecht-Strasse 24-25, 14476 Potsdam, Germany

The assessment of vulnerability gained more and more attention in flood risk research during the recent years. However, there is still not much knowledge available about flood vulnerability of companies and its influencing factors. This study follows the natural sciences concept which defines vulnerability as the degree of loss to a given element at risk resulting from flooding of a given magnitude.

Machine learning algorithms like Random Forests (RFs) are promising approaches, since they consider many influencing variables and as such allow for a detailed assessment of flood vulnerability. Only these variables which are meaningful for the differentiation of a certain target variable are used by the derived models. This allows for an identification of relevant damage influencing variables and hence for a more detailed picture of flood vulnerability of companies.

This study aims to identify relevant damage influencing variables by means of the variable importance provided by Random Forests. The data sets used are taken from two surveys conducted after the floods in the Elbe and Danube catchments in the years 2002 and 2013 in Germany. Damage to buildings ($n = 430$), equipment ($n = 651$) as well as goods and stock ($n = 530$) are taken into account. The analysis is done for the entire data set as well as for four groups of different company sectors and the corresponding data subsets.

Relevant damage influencing variables separated by sector and assets are identified as, for example, the degree of contamination or precautionary measures undertaken before the flood event. The results provide insight into the damage processes and improve data-acquisition in future surveys by, for instance, asking specific questions for company sectors and assets.