EGU2020-10025
https://doi.org/10.5194/egusphere-egu2020-10025
EGU General Assembly 2020
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

A data-mining approach to investigate El Niño damage in Peru

Fabio Brill and Heidi Kreibich
Section 4.4 Hydrology, GFZ German Research Centre for Geosciences, Potsdam, Germany (fbrill@gfz-potsdam.de)

Compound natural hazards, like El Niño events, which trigger torrential rain, mudslides, riverine and flash floods, cause high damage to society. An improved risk management based on reliable risk assessments are urgently needed. However, knowledge about the complex processes leading to El Niño damage is lacking, and so are loss models. We explore a large dataset of building damage from the coastal El Niño event 2017 in Peru. We use data-mining techniques to analyse data of damage grades of about 180.000 affected houses together with satellite observations and open geo-information. In a first step, we use unsupervised clustering (t-SNE + OPTICS) to separate regions of different dominant processes. Secondly, we train various supervised classification algorithms and create feature importance rankings per cluster, to identify drivers of observed damage for each of these regions. A comparison of different algorithms provides further insights about the potential and limitations of these methods and datasets. Results indicate that topographic wetness is the most important indicator, as selected by the algorithms, when using the entire dataset. Rainfall sum and maximum from TRMM satellite measurements are identified as damage driver despite the coarse spatial resolution. Also urbanity, based on a focal window around the global urban footprint, appears to play a role for the amount of damage. At least a coarse separation of processes is possible: the slope length and steepness, bare soil index, stream power index, and maximum rainfall are dominating the damage processes in lower mountain ranges and canyons, indicating rapid processes. Damage in upper mountain areas seem more influenced by the rainfall sum, local topographic position, and vegetation cover. In the lowlands, topographic wetness is very dominant, indicating ponding water or riverine floods. As opposed to previous work, this study constructs importance rankings based entirely on real observed damage to buildings. It is therefore a step towards data-driven damage assessments for El Niño events.