

Coupling a regional warning system to a semantic engine on online news for enhancing landslide prediction

Alessandro Battistini, Ascanio Rosi, Samuele Segoni, Filippo Catani, and Nicola Casagli
DST Department of Earth Science - University of Florence - Italy

Landslide inventories are basic data for large scale landslide modelling, e.g. they are needed to calibrate and validate rainfall thresholds, physically based models and early warning systems.

The setting up of landslide inventories with traditional methods (e.g. remote sensing, field surveys and manual retrieval of data from technical reports and local newspapers) is time consuming.

The objective of this work is to automatically set up a landslide inventory using a state-of-the art semantic engine based on data mining on online news (Battistini et al., 2013) and to evaluate if the automatically generated inventory can be used to validate a regional scale landslide warning system based on rainfall-thresholds.

The semantic engine scanned internet news in real time in a 50 months test period. At the end of the process, an inventory of approximately 900 landslides was set up for the Tuscany region (23,000 km², Italy). The inventory was compared with the outputs of the regional landslide early warning system based on rainfall thresholds, and a good correspondence was found: e.g. 84% of the events reported in the news is correctly identified by the model. In addition, the cases of not correspondence were forwarded to the rainfall threshold developers, which used these inputs to update some of the thresholds.

On the basis of the results obtained, we conclude that automatic validation of landslide models using geolocalized landslide events feedback is possible. The source of data for validation can be obtained directly from the internet channel using an appropriate semantic engine.

We also automated the validation procedure, which is based on a comparison between forecasts and reported events. We verified that our approach can be automatically used for a near real time validation of the warning system and for a semi-automatic update of the rainfall thresholds, which could lead to an improvement of the forecasting effectiveness of the warning system.

In the near future, the proposed procedure could operate in continuous time and could allow for a periodic update of landslide hazard models and landslide early warning systems with minimum human intervention.

References: Battistini, A., Segoni, S., Manzo, G., Catani, F., Casagli, N. (2013). Web data mining for automatic inventory of geohazards at national scale. *Applied Geography*, 43, 147–158.