Natural Hazards
and Earth System
Sciences

Open Access

Discussions

Interactive comment on "Quantitative comparison between two different methodologies to define rainfall thresholds for landslide forecasting" *by* D. Lagomarsino et al.

Anonymous Referee #1

Received and published: 13 April 2015

The paper compares two empirical models for rainfall threshold definition to assess the most effective one for landslide forecasting at regional scale. The two models have been already published in peer-reviewed international journals (Martelloni et al., 2012; Lagomarsino et al., 2013; Segoni et al., 2014a-b) and, according to authors, are used in local early warning systems in Italy. What is new in this paper is the quantitative comparison between these models, which are applied in two different areas and validated against available rainfall and landslide data. Although the paper is clearly written and properly illustrated, my feeling is that this goal has not been fully achieved. Here are my concerns.

1. The manuscript is unbalanced in the way that it contains a long Introduction (sec. 1), a quite detailed description of the two published models (sec. 2) but a very short discussion of models application (sec. 3-4) and test results (sect.5-6). In particular, the description of the two datasets used for model comparison is very poor. For instance, the ER dataset is described as follows (pg. 899 row 12): "The dataset consists in 71 landslides occurred in the period 2004-2007 and in the measurements of 9 automated rain gauges". What kind of landslides were triggered? In which materials? Were they all "deep-seated rotational slides" as previously mentioned? What about the critical rainfall events? The same apply to the Tuscany dataset. Please provide a detailed description of the two datasets so that results can be adequately interpreted.

2. While the aim of the paper is "to compare the forecasting effectiveness of different rainfall threshold models for landslide forecasting" (p.892, row 1), the two models are run in calibration, not in validation. In both cases (if I've well understood) only the calibration dataset is considered (2004-07 ER; 2000-07 Tuscany) and a validation analysis is not performed. A straight comparison should be made on the predictions of the two calibrated models. Therefore, the confusion matrixes and the comments in section 6 ("Results show that the both threshold models are characterized by satisfactory results: in all applications, the validation statistics are close to optimal values") are misleading. This is an important issue to address.

3. Why the overall number of days in Table 5 is only 177 against an analysis period of 7 years?

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 3, 891, 2015.