We would like to thank R#2 for his comments that will help us to improve the paper.

[R#2]: *Section 2 (page 2765, lines 10-16). The Authors should discuss the implication of the choice of the inter-event time. In fact, some of their conclusions about the validity of the identified thresholds have to do with the role of the initial (pre-event) conditions on the triggering of the landslide. It is quite obvious that the choice of the inter-event time affects the "memory" of the previous event at the beginning of the new one. Such a memory depends also on the hydraulic properties of the soil cover under study (see another related comment below), so I expect that, for a given soil cover, the (arbitrary) choice of the inter-event time may hide the effects of the previously fallen rainfall. Indeed, the small number of events identified in the synthetic rainfall series used for the Monte Carlo simulation (less than 20 rainfall events a year) makes me think that the choice of a dry interval of at least 24 hours results in few long-lasting events separated by dry intervals long enough to allow the drainage of most of the previously infiltrated water from the soil cover (especially considering the high value of the hydraulic conductivity assumed for either the regolith or the fractured bedrock).*

[A]: We agree that the choice of the inter-event time is a delicate issue and that must be carried out very carefully. Several studies are dedicated to this issue (Restrepo-Posada and Eagleson, 1982; Bonta and Rao, 1988). From a general perspective, the choice of a minimum inter-event time for event separation affects two aspects of the methodology:

1. In the Monte Carlo simulation procedure, the (minimum) inter-event time is introduced to "switch" from the event-based transient response model (TRIGRS) to the initial condition computation model (drainage exponential equation) and viceversa. Being TRIGRS an event-based model, it is necessary to define what can be considered as an *event*. Adapting an approach similar to that proposed by Balistrocchi et al. (2009) and Balistrocchi et al. (2011), the inter-event time can be assumed as the minimum time needed to avoid the overlapping of the response produced by two subsequent rainfall events. To this end it is important to consider that the peak of pore pressure may be reached in general some hours after rainfall ceases. A criterion to define this switch interval may be hence that of choosing a value that is close to the maximum time interval, measured from the end of rainfall events, that is necessary to achieve the peak pressure head relatively to all the simulated rainfall-event responses. This means that in principle the switch time interval has to be chosen to be different for different sets of hydraulic/geotechnical parameters, but for sake of simplicity, and within the domain of parameter values for which the simulations were conducted, a single value may be accepted, causing negligible effects on the computation of the response. We have performed some preliminary tests, and we have seen that $\Delta t = 24$ h ensures the reach of the peak of transient response, within the range of simulated parameters. Hence the choice of this value.

2. In the *a posteriori* analysis of the simulations, the inter-event time choice is necessary to create the correspondence between landslide response and triggering rainfall, and in particular to define the rainfall intensity and duration pairs (I, D). It can be said that this stage is analogous to the beginning stage of someone that wants to analyse an historical series of rainfall and landslide data to derive empirical thresholds. In this case, an objective analysis requires the choice of a criterion to define the (I, D) triggering and non-triggering pairs. Several criteria have been applied in literature (see Berti et al, 2013 for a detailed discussion on this issue). Many of these criteria are based on the choice of a (minimum) inter-event time, as the criterion in our MS. In particular we have adopted the same criterion of Brunetti et al. (2010), which has in our case the advantage that it is based on a inter-event time of 24 h which is consistent with the "switch" time interval we have adopted.

To respond to the specific question by this reviewer, i.e. whether "the choice of a dry interval of at least 24 hours results in few long-lasting events separated by dry intervals long enough to allow the drainage of most of the previously infiltrated water from the soil cover", the results of Figure 4 and Table 4 of the MS show that this is not the case, since, referring to the $A/B = 10$ m case there are low intensity triggering events for all durations, and that there are $P_0 = 2$ events for which the initial condition is above the one that leads to FS = 1 (note that as explained in the MS, these P_0 values were excluded from the *a posteriori* analysis).

We will add some comments in the paper to better explain these issues.

[R#2]: *Section 3 (page 2773, equation (9) and following discussion). As already pointed out by comments made by others, the chosen objective function is not novel. A possible improvement of such a commonly adopted approach, could be taking into account that in many real cases it is not obvious that a false alarm and a missing alarm have the same importance (the losses deriving from a missing alarm may be much more serious than the costs deriving from a false alarm), and it would be more effective for the decision-maker to define an objective function which accounts for such weights.*

[A]: For this comment see response to SC by S.L. Gariano. Regarding the improvement suggested here, in some of our previous works (see Peres and Cancelliere, 2012; Peres and Cancelliere, 2013) we have considered an objective function with different weights given to FN and FP. Nevertheless, the choice of the weight is crucial in the definition of the threshold and little information is usually available to do this choice. Hence we preferred to proceed in a "standard" way, in which FN and FP are weighted equally. The part on ROC analysis will be modified in order to account of these comments.

[R#2]: *Section 4 (page 2774, lines 13-20). More information should be provided about the calibration of the NRSP model, especially because the resulting synthetic series may affect the significance of the obtained I-D threshold (see the above comment about the effects of the choice of the minimum inter-event dry interval). Such concern about the NSRP calibration is motivated by the sentence at the beginning of page 2778, where the Authors say that in five years 190 events were recorded: nearly 40 events per year, which is around the double of the average yearly number of events of the synthetic generated series.*

[A]: We agree with the reviewer to provide more information on the calibration of the NSRP model. Specifically, it is true that the mean number of events per year considered from the simulated series is about 20. This is because rainfall in the months from April to August was neglected (See response to R#1). Moreover, the events having intensity always lower than the leakage flow rate have been removed from simulations as well, because it is sure that they produce no water table head rise. The inconsistency that results from the comparison of the results presented for the global validation test (where about 40 events per year are present) is due to the fact that in doing such test the months from April to August are not excluded and neither the cut of under-leakage events it is applied. This was done because the global validation test is aimed at testing every assumption done within the modeling framework, including the choice of neglecting some rainfall events.

If we exclude in the observed series the events in the period April-August, the mean number of events per year result of 19.18 which is statistically comparable with 19.83 of the simulated series (without the cut of the events with intensity less than the leakage flow rate, the mean number of events per year results to be 28.75 in the simulated series versus 28.91 in the observed series).

We will write this detail in future versions of the MS.

[R#2]: *Section 4 (page 2774, lines 22-26). Much more information about soil properties should be provided (it is not even written which kind of soil is studied). It seems that the Authors consider the obtained thresholds representative of an area as large as several tens of square kms. The variability of soil properties and slope morphology within such a large area could completely reduce the obtained results to a mere modeling exercise. The value of the critical wetness ratio corresponding to the assumed geotechnical soil properties and slope geometry should be given here.*

[A]: We agree that the paper will benefit of adding more information on the soil properties and on the study area. We have posted a specific supplement on this issue. (http://www.hydrol-earth-syst-sci-discuss.net/11/C2076/2014/hessd-11-C2076-2014-supplement.pdf). Critical wetness ratios will be specified as well.

[R#2]: *Section 5. The results of the sensitivity analysis to variations of geotechnical soil properties and soil cover thickness are quite interesting, but they should be completed also with the analysis of the effects of variations of the hydraulic properties (in particular the hydraulic conductivity) and, even more, of the ratio cd between the hydraulic conductivity of the fractured bedrock and that of the soil cover (as far as I understand, arbitrarily set to 0.1 in absence of experimental data). Indeed, as the failure is a-priori assumed to occur at the soil bedrock interface, I expect the results to be extremely sensitive to the variation of such parameters.*

[A]: We agree that it is worthed to complete the sensitivity analysis with the analysis of the effect of variations of the hydraulic properties. Hence we have conducted additional simulations, considering different values of the hydraulic conductivity K_S and the ratio c_d between the hydraulic conductivity of the fractured bedrock and that of the soil cover, in a interval centered on the values $K_s = 2 \times 10^{-5}$ and $c_d = 0.1$ considered in the MS. In particular we have conducted simulations to make similar plots of Figure 6 of MS for the following set of parameters: (a) $K_s = 10^{-5}$ m/s $= 36$ mm/h and $c_d = 0.1$ (b) $K_s = 2 \times 10^{-5}$ m/s $= 72$ mm/h and $c_d = 0.05$ (c) $K_s = 2 \times 10^{-5}$ m/s $= 72$ mm/h and $c_d = 0.05$ (d) $K_s = 3 \times 10^{-5}$ m/s $= 108$ mm/h and $c_d = 0.1$.

The results are shown below, in Figures $1-5$.

In the captions we report the values of the time constant of the sub-horizontal drainage model $\tau = \frac{\frac{A}{B}(\theta_s - \theta_r)}{K_s \sin\delta}$ as well.

In fact we agree that the K_s and c_d may affect significantly the best performing ID power-law threshold (the a_1 parameter) and its performance (the maximum TSS).

First of all, for $K_s = 36$ mm/h memory is high and the $I-D$ model results inadequate (TSS< 0.8) for $A/B = 10$ m, though in the case of $d_{LZ} = 2$ m for high critical wetness ratios performance may be still acceptable (TSS>0.8)

In the $K_s = 72$ mm/h case, where the variation of c_d is considered, results indicate sensitivity to this parameter. In particular performances of the $I - D$ model increase with the increase of c_d and thresholds result higher.

In the $K_s = 108$ mm/h case performances get better, and the thresholds result higher.

The conducted sensitivity analysis shows that performances of the I-D model increase as τ decreases, the critical wetness ζ_{cr} increases, the soil depth d_{LZ} increases and the leakage c_d increases.

When we speak of memory, for the studied climate, results show that one may consider the specific upslope area to be "small" when it results $\tau \approx 3$ days.

Specifically, we have obtained that when $\tau < 3$ days ($A/B = 10$ m, $\theta_s - \theta_r = 0.3050$, $K_s = 2 \times 10^{-5}$, $\sin\delta = 0.643$) and the soil depth is greater than 1 m and $\zeta_{cr} > 0.5$, performances are acceptable (TSS> 0.8).

Conclusions will be revised taking into account of the results of these new simulations.

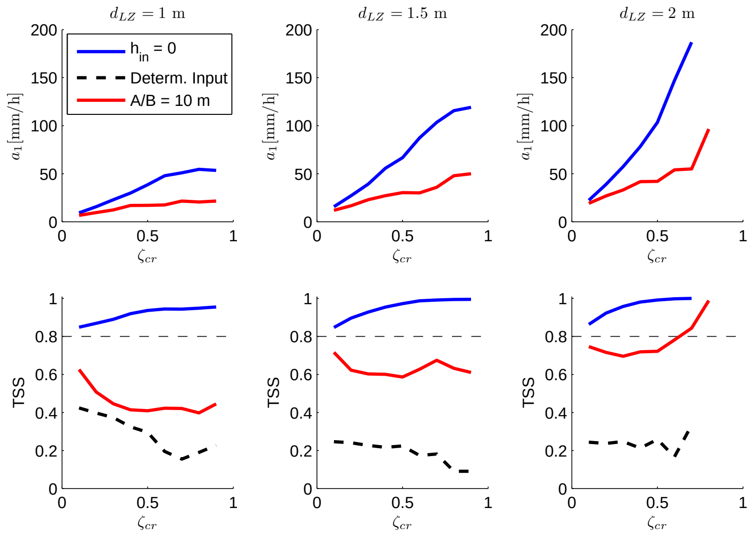

Fig. 1. Results for $K_s = 36$ mm/h and $c_d = 0.1$; $\tau = 5.5$ days.

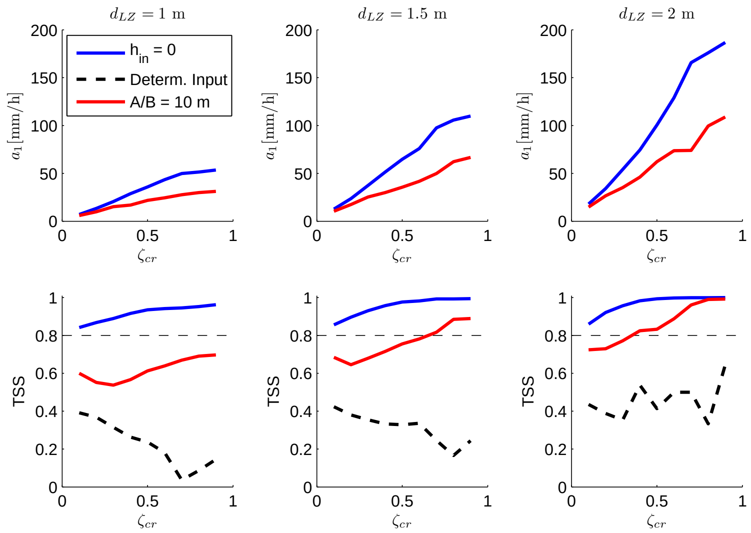

Fig. 2. Results for $K_s = 72$ mm/h and $c_d = 0.05$; $\tau = 2.7$ days.

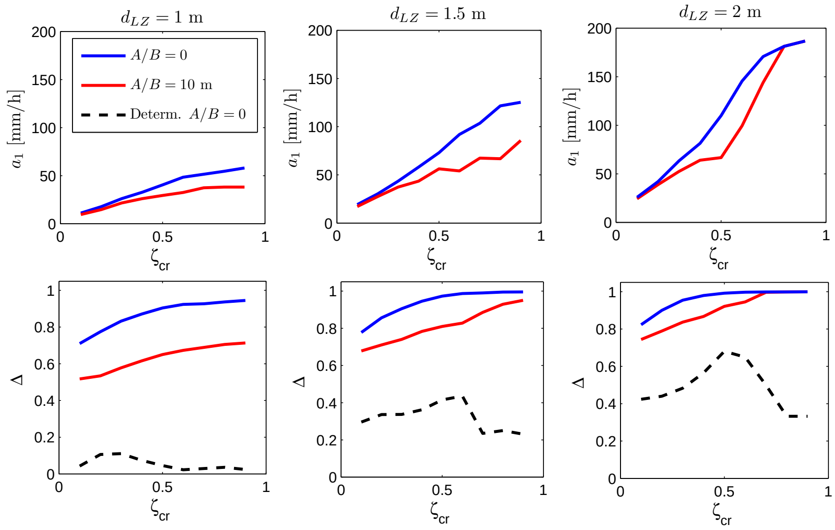

Fig. 3. Results for $K_s = 72$ mm/h and $c_d = 0.1$; $\tau = 2.7$ days (Figure 6 of MS, here shown for completeness).

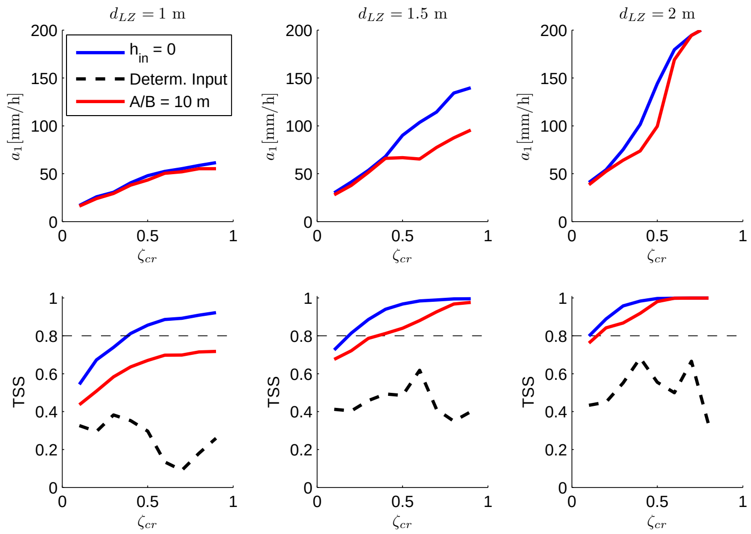

Fig. 4. Results for $K_s = 72$ mm/h and $c_d = 0.2$; $\tau = 2.7$ days.

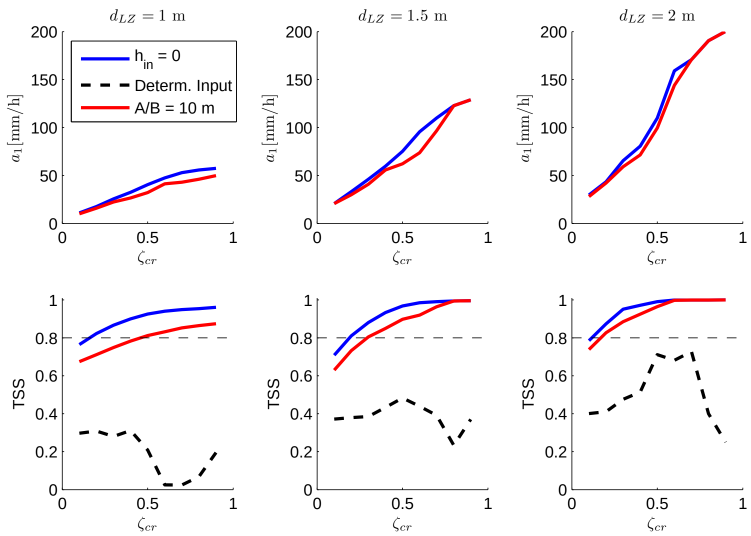

Fig. 5. Results for $K_s = 108$ mm/h and $c_d = 0.1$; $\tau = 1.8$ days.

Editorial issues

[R#2]: *Page 2769, last three lines, and page 2770, first line. The explanation of the meaning of A, B and their ratio is unclear and should be reformulated.*

[A]: We will explain better the meaning of A and B and their ratio, since we will provide a more detailed description of the assumptions by Rosso et al., 2006, on which the drainage model is based. See also response to R#1

[R#2]: *Page 2774, lines 9-10. Possibly the events of 25 October 2009 and 1 October 2009 are the same event (with wrong dates): somewhere else in the paper it is written that only four landslides occurred during the considered period.*

[A]: There are typing errors of the landslide dates. The correct ones are reported in the supplement regarding case study data (http://www.hydrol-earth-syst-sci-discuss.net/11/C2076/2014/hessd-11-C2076-2014-supplement.pdf). Only four of the events listed have been then considered in the global validation test, because the observed rainfall time series covers a period which includes only those 4 events.

[R#2]: *Page 2776, line 8. Replace "correspond" with "corresponds". Page 2778, line 19. Insert "as" between "soon" and "soil". Page 2779, line 2. Delete the word "for" after "Traditionally".*

[A]: We accept all these corrections.

References

Balistrocchi, M. and Bacchi, B.: Modelling the statistical dependence of rainfall event variables through copula functions, Hydrol. Earth Syst. Sci., 15, 1959-1977, doi:10.5194/hess-15-1959-2011, 2011.

Balistrocchi, M., Grossi, G., and Bacchi, B.: An analytical probabilistic model of the quality efficiency of a sewer tank, Water Resour. Res., 45, W12420, doi:10.1029/2009WR007822, 2009.

Baum, R. L., Savage, W. Z., and Godt, J. W.: TRIGRS – A FORTRAN program for transient rainfall infiltration and grid-based regional slope-stability analysis, version 2.0, US Geological Survey Open-File Report 2008–1159, Reston, Virginia, 2008. 2763, 2764, 2768, 2769, 2770, 2790

Berti, M., Martina, M., Franceschini, S., Pignone, S., Simoni, A., and Pizziolo, M.: Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach, J. Geophys. Res.-Earth, 117, F04006, doi:10.1029/2012JF002367, 2012. 2762, 2763

Bonta, J. V. and Rao, A. R.: Factors affecting the identification of independent storm events, J. Hydrol., 98(3–4), 275–293, doi:10.1016/0022-1694(88)90018-2, 1988.

Brunetti, M. T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D., and Guzzetti, F.: Rainfall thresholds for the possible occurrence of landslides in Italy, Nat. Hazards Earth Syst. Sci., 10, 447–458, doi:10.5194/nhess-10-447-2010, 2010. 2762, 2771

Peres, D. J., Cancelliere, A. Development of rainfall thresholds for landslide early warning in Sicily: A method based on the use of rainfall annual maxima series (2012) Rendiconti Online Societá Geologica Italiana, 21 (PART 1), pp. 580-582.

Peres, D. J., Cancelliere, A. Defining rainfall thresholds for early warning of rainfall-triggered landslides: The case of North-East Sicily (2013) Landslide Science and Practice: global Environmental Change, 4, pp. 257-263.

Restrepo-Posada, P. J. and Eagleson, P. S.: Identification of independent rainstorms, J. Hydrol., 55(1–4), 303–319, doi:10.1016/0022-1694(82)90136-6, 1982.