Response to the comments of reviewer 1 with regard to the discussion paper:

Mergili, M., Marchesini, I., Alvioli, M., Metz, M., Schneider-Muntau, B., Rossi, M., Guzzetti, F., 2014. A strategy for GIS-based 3-D slope stability modelling over large areas. Geoscientific Model Development Discussions 7, 5407–5445. doi:10.5194/gmdd-7-5407-2014.

We would first like to thank the reviewer for the constructive comments on our paper. It is good to hear that, according to the reviewer, the topic is of interest in principle and suitable for the journal. Below we address each comment in detail. Our responses are given in bold blue letters.

Changes in the manuscript, compared to the initial submission, are highlighted in yellow colour.

General Comments

The paper presents a strategy to improve the computational efficiency of a slope stability assessment model, r.rotstab, through multi-core processing and strategies for geometrical and geomechanical parameters sampling. This model allows the assessment of the susceptibility of slopes to shallow landslides through the computation of factor of safety on potential truncated ellipsoid surfaces of rupture (3D). The paper discussed parallel processing and several strategies for geotechnical parameters samplings. For large areas, a sequential approach would request huge computational times; therefore, the proposed approach is very valuable. The issues raised in the paper are valid for every physically-based model, and in particular for landslide susceptibility assessment.

Therefore, the problematic is valuable for the GMD community.

However, the description of the background and the model (part 1 and 2.1) are very similar to the paper from Mergili et al. (2014). Hence, the described background and the state-of-the-art are more oriented towards the problematic addressed in the former paper (needs for finite slope, physically-based models) [..]

We agree with the reviewer that much of the general framework is similar to the one presented in Mergili et al. (2014). We chose to describe in depth this part to present the reader with a self-contained paper. We have reduced this part in the revised manuscript, referring the interested reader to Mergili et al. (2014). In particular Chapter 1 is now focused on the existing last two paragraphs whilst the remaining part was condensed to the absolute minimum necessary. In contrast, it was hard for us to substantially shorten the first three paragraphs of Chapter 2.1 because (i) they are absolutely necessary to make the paper self-contained and (ii) the reviewer asks for the more detailed explanation of a number of aspects exactly concerning this part of the manuscript.

[..] while the state of practice in computational efficiency and sampling strategies is not really discussed. There is no information on why these sampling strategies have been selected and there is an offset between the abstract, background and state-of-the-art and the results and discussions.

The methods used (models, parallel processing and sampling) are already existing ones. Therefore, in my opinion, both the subject (computational efficiency) and the model are of interest to be published in GMD, but the paper should be totally restructured to reflect the reflections in sampling strategy and parallel computing before publication. Bibliography should also correspond to computational efficiency.

We further agree that more background and discussion on computing-specific topics and sampling strategies could add value to the article. We have modified the manuscript (and the reference list) in the way to cover these aspects in sufficient detail and to bring abstract, background and state of the art in line with the results and the discussion. Moreover the use of the sampling strategies chosen are better justified in the revised manuscript.

Specific comments

It is not clear enough what are the differences between r.rotstab and r.slope.stability, except the fact that r.slope.stability can be used with multi-cores computers.

Indeed, as explained in section 2.1 (P5413 L18 – P5414 L3), the main differences between the old r.rotstab model and r.slope.stability are (i) the capacity of r.slope.stability for parallel computing and (ii) the definition of a slope failure probability.

No information on computational times is provided. We cannot judge how necessary are parallel processing and parameters sampling. What are the capacities of the computers used?

We have added information on the capacities of the computer used as well as on the computational times. In particular we used a 48 cores (AMD Opteron, frequency of 2.2 GHz and cache of 512 KB) computer with 140 GB of RAM and running a 12.04 LTS Ubuntu GNU/Linux OS with the 3.5.0-26-generic kernel image. The computational times fluctuate depending on the settings of the experiments. Some values of Tp are specified in the revised manuscript (section 5.1): for d_e=2500, the computational time is approx. 110,000 seconds with 1 core and 1 tile, whilst it reduces to approx. 4,700 seconds with 42 cores and 182 tiles. Further, a new table (Table 5) was introduced, summarizing the evaluation results and computing times of test 2 (Section 5.2)

Global confusion between “surface” and “plane” of rupture. Not every failure occurs following a plane.

We agree with this comment and have replaced the term “plane” with “surface”.

In order to be clearer in the description of the model, it would be good to mention that no inter-column forces are considered in the r.slope.stability model.

We will mention this important aspect in the revised manuscript but this requires that we keep a brief introduction to the model from a geotechnical point of view.

The randomization process for W and L is not discussed. Do the authors also use a strategy to increase computational efficiency, or is a Monte-Carlo strategy used? What is the methodology used to ensure a proper repartition of ellipsoids over the whole area? Regular or random sampling?

The ellipsoid parameters are sampled randomly. The corresponding paragraph was reworked and the following sentence was added to the revised manuscript to improve clarity: "Simple pseudo-random numbers, generated separately for each parameter of each ellipsoid, are used to define the centre coordinates as well as the values of L, W, D and zb, constrained by user-defined minima and maxima for each parameter. Testing a sufficiently large number of ellipsoids ensures a proper repartition of the ellipsoids over the study area, and the consideration of a large variety of possible ellipsoid dimensions.".

The hypothesis of soil saturation is not discussed, even if it is a quite conservative hypothesis. It seems reasonable to make this hypothesis for the purpose of the paper, but the results should be analyzed accordingly. The final Pf maps correspond to "probability" of failure in the worst case scenario, and they do not correspond to current probability of failure.

This issue is a very important one. In fact, the computed probability of failure is valid for the worst-case assumption of fully saturated conditions, as correctly stated by the referee. Partial saturation is more difficult to treat from a geotechnical point of view and shall be the subject of future studies. This limitation is now clearly stated in the discussion of the revised manuscript.

Technical comments

P2 l22: Is this zone the same one as in Mergili et al. 2014? In this case, why the area is different?

The area given in Mergili et al. (2014) refers to another definition of the boundaries than the area given in the present paper.

P3 l4: "consisting" instead of "consiststing"

Thank You, this error was corrected!

P3 l6: not all the physically-based models assume the surface of rupture to be a plane. (i.e. circular assumptions with Jambu or Morgenstern-Price approaches).

The term "plane" (issue already raised above) was replaced by the term "surface" throughout the entire manuscript.

P3 l14: “forces” instead of “forcess”

Thank You, this will be corrected!

P3 l23: references to Baum et al. (2002) and (2010) are missing in the references section.

Thank You – these two references in the text are not necessary and were removed.

P4 l29: is the notion of large areas really commonly related to number of pixels, and not to sizes, or the existence of several objects (i.e. slopes)? A single slope can have ~10ˆ8 pixels, according to the resolution

From a computational point of view it is related to the number of pixels.

P6 l10-11: does the offset correspond to the offset mentioned l 15. In this case, it is better to mention zb l10.

Yes, it is z_b – in the revised manuscript, we use the symbol already in line 10.

P6 l.19 What could be considered as “relatively small pixels”? Could the ratio W/pixel size~3 be considered big enough?

The issue of pixel size was discussed in Mergili et al. (2014) – this reference is added at the corresponding place of the revised manuscript. Pixel size = W/3 is certainly at least at the verge of yielding a shape resembling an ellipsoid. However, we note that (i) this is the extreme and most ellipsoids are much larger and (ii) more importantly, the ellipsoid is an idealized shape which will not exactly occur in nature anyway – i.e., if the failure plane is no ellipsoid but some kind of polygon, this does not make the geotechnical computations wrong – it just means that the shape tested is slightly different.

P8 l5-6: also variability of the geometry parameters (i.e. d)

Yes, thank You, this is mentioned in the revised manuscript.

P8 l16: Does the number “n” of samples of samples to be collected correspond to the samples from the ground, (in this case, not consistant with the “n” used in the rest of the paper, e.g. p9 l3)?

The number n does not refer to the samples collected from the ground, but to the statistical samples. We have collected field data in a much smaller count and used that data to build the probability density functions for each of the measured soil parameters. The values of the parameters used for each run are given by sampling the PDF of the corresponding parameter. We have made this aspect clear in the revised manuscript by referring to “statistic samples” which are “considered”.

P9 l4: “largest” instead of “lagest”,

Thank You, this error was corrected!

P9 l5: is it correct to consider the probability of failure for a pixel to be the largest Pf computed for the different parameters combinations? The value could be representative of the propability, but not the propability of failure per see.

From a strictly geotechnical point of view (which we follow here), it is the largest failure probability out of all intersecting ellipsoids which is interesting for us as, for each pixel, only the most critical ellipsoid is relevant. It is NOT the largest value of P_f out of all parameter combinations we consider – P_f is a result of combining the values of FOS yielded with all parameter combinations.

P9 l20-25: It is not clear how the sampling is performed: for a) and b), you select a different combinations (c', φ') for each ellipsoid, while for c), you pick one combination, and you consider the parameters homogenous over the whole area? Why don't you use the last option also for a) and b)?

We try explaining this issue more clearly in the revised manuscript. It is not one combination of parameters we use in c), it is rather one set of parameter combinations. As this set is not defined randomly, but deterministically, it is not necessary to define a new set for each ellipsoid. With a) and b), the set of parameters/combinations is sampled randomly for each ellipsoid. Our results seem to show that the regular sampling strategy is an efficient alternative to the classic Monte Carlo approach, at least in our geotechnical / geomorphological settings, and using the current implementation of r.slope.stability.

P9 l25: Please mention that in the paper, it is this solution (application to three parameters) which is applied. It makes confusion after in sect. 4

The application to three parameters is made clear in the revised manuscript.

P10 l14: According to Eq 4, the density is not in ellipsoids per pixel, but ellipsoid per unit of surface (here, meter).

The average number of ellipsoids per pixel is dimensionless. As long as the pixels are much smaller than the ellipsoids, we consider it a valid approximation with regard to pixels. Equally, it is a valid approximation with regard to any other square unit, as long as the square unit is much smaller than the ellipsoid size (here, square metres would be appropriate).

P10 l 16: "A" is not described here, and appears different from the parameter in Eq2

In Eq. 4, A is the study area size. This is now explained, and the variable was renamed to A_s in order to remove the confusion with A in Eq. 2.

P12 l15: does the inventory correspond only to scarps, to reflect areas of departure?

As most landslides in the Collazzone Area have a limited mobility, we have decided to use the entire landslides. Using only the scarps would mean that most of the landslide area is considered as observed negative as the deposit overlays most of the scarps which can therefore not be defined properly. There is certainly a slight overestimate of observed positives due to this. This aspect is made clear in the revised manuscript.

P13 l24: Please also add the standard deviation of c’

The standard deviation of c’ is given in Table 2. However, as we use an exponential pdf, it is meaningless for the modelling.

P16 l20: It is good to notice, but isn’t it normal? If you don’t have a new job to give to an available processor, this processor is somehow useless for the computation.

The referee is right, this is actually clear. The sentence is needed for introducing the following sentence, but it was reformulated to “There is, of course, no gain in terms of speedup at $p > t$ (not shown in Fig. 5).”.

P19 l22 (and Figure 8): which parameter sampling strategy has been selected?

It was equal density sampling (strategy c). This is now clearly mentioned in Section 5.2 and in the captions of Figs. 7b and 8.

In the discussion part: the recommendation for n 9ˆ3 is valid for area assuming a unique parameterization over the whole area. Would it be the same with soil-type specific areas? In this case, where ranges of variations of parameters could be probably be reduced, would n smaller than 9ˆ3 suitable?

This is a very interesting question. If we could reduce the parameter variations, it is likely that also n could be reduced. However, with the data we have right now, also the variations within each soil type are quite large. The comment of the referee will be an interesting aspect to explore in a future paper, obtaining more data (these efforts are going on) and/or conducting theoretical experiments. A short note on this issue was added to the discussion of the revised manuscript.

Would it be possible to consider different soil water content conditions?

In principle it would be possible and interesting, but (see above) partial saturation is more difficult to treat from a geotechnical point of view and shall be the subject of future studies. This is clearly stated in the discussion of the revised manuscript.

In references, Iverson and Major (1986) is not referenced in the text.

Thank You – the reference has been removed from the bibliography.

Response to the comments of reviewer 2 with regard to the discussion paper:

Mergili, M., Marchesini, I., Alvioli, M., Metz, M., Schneider-Muntau, B., Rossi, M., Guzzetti, F., 2014. A strategy for GIS-based 3-D slope stability modelling over large areas. Geoscientific Model Development Discussions 7, 5407–5445. doi:10.5194/gmdd-7-5407-2014.

The proposed paper introduces a) an open-source, multi-core processing application (r.slope.stability) on landslide susceptibility mapping over large areas capable of computing both FoS (factor of safety) and the probability of slope failure (Pf) parameters; b) the efficiency and fastness of this application compared to the single-core version (r.rotstab); c) parameterization strategies on field-measured and heterogeneous geotechnical and soil depth datasets; d) and how it affects the landslide susceptibility map (FoS and Pf) for shallow landslides of Collazzone area in Umbria. Each of these works contain novelties and therefore very valuable for the GMD community.

We would first like to thank the reviewer for the constructive comments on our paper. It is good to hear that, according to the reviewer, all aspects considered contain novelties and are valuable for the GMD community. Below we address each comment in detail. Our responses are given in bold blue letters.

Changes in the manuscript, compared to the initial submission, are highlighted in yellow colour.

However this wide range of topics makes difficult to maintain the focus of the paper. This should be the efficiency, fastness and accuracy of the multi-core processing algorithm.

Therefore more technical details on the hardware and comparisons should be provided on the different runs (e.g. in tabular form).

We have added information on the details of the hardware: we use a 48 cores (AMD Opteron, frequency of 2.2 GHz and cache of 512 KB) computer with 140 GB of RAM and running a 12.04 LTS Ubuntu GNU/Linux OS with the 3.5.0-26-generic kernel image. A new Table (Table 5) was introduced, summarizing the evaluation outcomes and the computation times of test 2 (Section 5.2).

More detailed evaluation/validation of the results on the test site compared to the earlier landslide susceptibility maps and the landslide inventory might help the reader to put the results in a broader context.

Earlier work in the Collazzone area, using statistical methods for computing landslide susceptibility, yields higher values of A_{ROC} (0.71 – 0.75, depending on the method, Rossi et al., 2010), whilst the susceptibility index introduced by Mergili et al. (2014) yields comparable values (0.68 – 0.70). Even though these results are not fully comparable due to different inventories and reference units

used, they indicate that the geotechnical parameters have to be better understood in order to make physically-based models superior to statistical ones. We have added these aspects to the discussion.

The input parameters are perfectly summarized in tabular form, the results of each sampling strategies (Sect. 5.2) should also be presented similarly with shorter discussion, helping the easier comparison and maintaining the focus of the paper.

We have added a table (Table 5) summarizing the results obtained with each tested combination of parameter settings to Sect. 5.2.

We did not find a lot of potential to shorten the discussion as (i) it was already condensed to the most essential issues in the initial version and (ii) there were some requests from the other reviewer to add additional aspects.