Geosci. Model Dev. Discuss., 8, C2526–C2530, 2015
www.geosci-model-dev-discuss.net/8/C2526/2015/

Geoscientific Model Development Discussions Open Access

Interactive comment on "r.randomwalk v1.0, a multi-functional conceptual tool for mass movement routing" *by* M. Mergili et al.

J. Robl (Referee)

joerg.robl@sbg.ac.at

Received and published: 18 October 2015

Mergili et al., present a new open source tool to describe a variety of rapid mass movements by applying a random walk approach. Such codes have been and still are frequently used to assist natural hazard projects. In contrast to other codes, the presented tool offers a variety of additional options: the code allows different types of break criteria, account for the uncertainties, performs advanced statistics, allows the back-calculation of a set of observed mass movements and provides nice visualizations of the results. To achieve this functionality the code is integrated in the GRASS GIS framework and uses R.

Although not fully convinced by the random walk approach itself, the presented tool unifies the benefits of different random walk approaches presented by others in a free

and open GIS and add additional functionality. Hence, this code will be a valuable tool for the natural hazard community. The functions and methods are well described and the test areas differ significantly in terms of occurring rapid mass movements and model requirements, so that a wide range of model functions are successfully applied.

To be short, the test cases are very well chosen. The figures are informative and of high quality. The code and some datasets for testing are freely available from the website of the first author. The installation requires the compilation of GRASS GIS from source and additional hacking to get the r.randomwalk code correctly installed (at least on my rather old CentOS6.5 installation). This seems to be an obstacle and I wonder if it is possible to use the add-on system of GRASS GIS for releases in the future. The code itself worked flawlessly for the first test case and it is by far the most flexible and advanced random walk code for flow routing I have tested during the last ten years. Overall, I suggest publishing this study in GMD after minor revisions.

General Remarks: Every random walk approach requires defining several "unphysical" parameters like Rmax (user-defined maximum vertical run-up height) or Lseg, Lctrl (to avoid uneven flow paths) and several others. These parameters may lead to flow directions that are directed backwards or in a circular way if they are not set appropriately. Physical-based models that describe the motion of fluid by a depth averaged form of the Navier-Stokes-equations and an appropriate flow resistance law require less parameters and are in my opinion better constrained (although uncertainties are still large). The biggest advantage of a random walk approach over physical based models may be the computational performance of random walk codes that allow "Monte Carlo Simulations" and the exploration of the crucial parameter space. However, additional parameters to perform a random walk approach (see above) increase the parameter space dramatically. Sophisticated numerical methods (e.g. AMR: adaptive mesh refinement) in concert with modern and fast computers reduce to computational costs for physical based models and parameter studies can also be performed. Despite of that, these models provide us with flow height, velocity and momentum, travel time and other parameter that are important for mitigation strategies against natural hazards. Thus, I expect that random walk approaches describing natural hazards will vanish in the next decade(s) but in the meantime the code of Mergili et al. will probably the best choice.

After reading the manuscript several times I have the impression that the separation in chapters “3 Test sites and model parameterization” and “4 Results” is somehow confusing and follows the design layout of a classical natural hazard study. Maybe I’m wrong but I think it would be less confusing by merging chapter 3 and chapter 4, call it for example “3 Test cases” and stitch the model setup, the parametrization and the results of each test site together that will get the subheading 3.1, 3.2,3.3. However, this is just a suggestion.

There are many terms and abbreviations that require an exact definition. Some of them are defined in the context where they are used the first time; some are not defined at all (common knowledge?). Maybe the authors should implement one additional paragraph defining the most important terms (impact indicator score, release indicator score, impact indicator index, exposure indicator scores, impact probability raster map PI). All these terms and even worse their abbreviations together with a couple of dozen of variables are really confusing (at least for me).

In this context I encourage the authors to implement a new table 1 summarizing all abbreviations and variables with a short explanation (maybe instead of an additional paragraph?) Please implement one or two sentences on the computational performance of your model. How long will it take to perform each of the test cases on standard computer hardware?

Line by Line Remarks

P8194 - L10: parameter settings → parameter sets?

P8195 – L5-L13: I do not fully agree with your statements here (especially the first

sentence). Modelling rapid mass movements with numerical tools is always affected by large uncertainties related to initial parameters and fluid properties. However, I do not see why random walk models with empirical break criteria are “better” constrained than physically based models and should therefore applied in regions with sparse input data. Instead of calibrating parameters of flow resistance laws for physical based models, break criteria have to be calibrated (and a bunch of other parameters) for random walk models. In my opinion the main advantage of models based on a random walk approach may be that they cost less in terms of computational time and allow therefore (a) modelling many rapid mass movements on valley scale and (b) the exploration of a large variety of parameter sets without employing a computer cluster.

P8197-L2-4: “The parallelization procedure is implemented at the python level (analogous to the way described in Mergili et al., 2014) and serves for two purposes” First, I think that you should try to explain your parallelization approach in a few sentences. Model development is in the heart of this journal (although the approach has been already published in Mergili et al., 2014.). Second: The main reason for parallelization is the reduction of the computational time for Monte Carlo simulations, I assume.

P8197-L5-6: This statement is not clear to me. Maybe you can rephrase it.

P8200 – L15: IIS : was not explained before

P8200 – L16: parameter sensitivity tool AIMEC: Please explain the functionality of this tool and also the abbreviation.

P8200: L21.24: This is not clear to me. What do you mean with model development in this context?

P8200: L27 - : This is a cool feature!

P8201: L7 - : IF not defined before.

P8202: ROC – although this is common statistics, it should at least be explained by writing out the full term once.

P8202: L8: Maybe all these abbreviations (e.g. AUCROC) are frequently used by the natural hazard community. However, too many abbreviations hamper reading this paper and I wonder if all of them are really required for this study.

P8203-L16: using an ROC Plot → using a ROC Plot

P8203-L5-: "We assume that the values of nruns, Rmax, f_, fd , Lctrl, Lseg and the pixel size applied to the Acheron Rock Avalanche (see Sect. 3.1) are valid also for this study area." Well – I just wonder how you can assume that. Does this mean that we can assume this for all types of rapid mass movements worldwide ;-)? In this case I wonder why we need nruns, Rmax, f_, fd , Lctrl, Lseg as free parameters.

P8203-L14: How did you identify" the Gaussian distribution as the most suitable type of distribution for this purpose"? Please explain in one or two sentence(s).

P8208-L19-L22: I'm not sure if it is a good idea to use degrees for statistics. In the light of the non-linear relationship between percent and degrees - what will happen by averaging slope in degree and what can we learn from the STD in degree? This seems to be dangerous for rough flow paths with alternating high and low gradient flow segments?

8213 – L26: "Krenn et al., submitted" : Paper is missing in the list of references.

Table 2: What are the numbers in the brackets after the formulas?

Interactive comment on Geosci. Model Dev. Discuss., 8, 8193, 2015.