Response to REVIEWER N°1

Dear Reviewer

We thank you a lot for the punctual revision of our work. The paper represents a summary of my PhD thesis that I have adapted into a scientific publication. I am aware that the topics involved in CRHyME construction are very wide and touch different disciplines so I am grateful to have received useful hits for improving my work. In the next parts, you can find the reply to your suggestions and observations.

Review comment for "CRHyME (Climatic Rainfall Hydrogeological Model Experiment): a new model for geo-hydrological hazard assessment at the basin scale" by Andrea Abbate et al.

The authors present a new software that combines hydrological modeling with landslide and sediment transport modeling. With its distributed approach and physics-based modeling, CHRyME is a necessary complement to existing software.

This manuscript is an impressive paper that, because of its cross-disciplinary nature, addresses several very different topics. However, the sentence structure and the English do not always reach an adequate level, which, coupled with the vocabularies specific to the different themes, sometimes makes the text very difficult to understand. I would recommend checking the structure and meaning of each sentence using a translation tool. Furthermore, this article would benefit from additional references. I have made suggestions in this regard in the attached text.

We apologise for the English, and we have planned to do a language proofreading, splitting long sentences that may lead to difficulties in reading. About vocabulary we think that a list of acronymous or keywords should be made in one of the appendices in order to facilitate comprehension by the reader, also subdividing them into categories such as (i.e., climatology, meteorology, hydrology …)

The paper has been completely revised, pay more attention to sentence construction and vocabulary. Some parts were reduced and synthesized in order to make these paragraphs clearer to the reader. We are planning further language proofreading after this turn of revision. We added references that were required. We modified appendixes A and B: Appendix A contains more practical information about the code usage with an example of the .INI file for model initialization; Appendix B contain all the main acronymous used in the model and in this current paper. About the cases studied, a test case within a small catchment (30 km2) of Caldone River (Lecco, Lombardy, Italy) has been included for testing hydrological balance, sediment transport and dependence of the results with respect to the spatial resolution of the digital terrain model (Pars. 2.4, 3.1 and 4.1).

Here are the main points:

1) The connection between the different modules is not as clear as it should be in Figure 3. For example, the amount of sediment available for transport is not determined, as I would have expected, by the number and size of landslides triggered, but by the Gavrilovic equation. Similarly, the link between the hydrologic model and the water-dependent variables in the stability equations of the landslide module is also not clearly established. It is therefore difficult to assess the degree of novelty of the model as a whole.

This represents a critical part and we think that a paragraph should be dedicated for.

The amount of sediment is still determined using the Gavrilovich method since is quite well-established methodology for spatially distributed models. The mass wasting coming from landslides is not properly taken into account now since only shallow landslide “triggering” models are included (infinite slope) while erosion is not “perturbed” by local mass wasting. This is a feature we are planning to include in the future version of the model that I our opinion is very useful for now-casting prevision or singular case study analysis.

The link between the hydrological cycle and slope stability is made by soil moisture estimation, that is a variable needed to assess stability in partially saturated slope, avoiding the hypothesis of complete saturation (generally considered as conservative solution). Also this part should be better highlighted in the text.

Following these suggestions, we have rewritten paragraph n°2 in this way:

In 2.1 are included the novelties of the model

In 2.2 is described the model structure describing 2.2.1 model initialization (databases and required data series) 2.2.2 the hydrological equations of the model (adding further details and reworking on English to make it clearer) 2.2.3 the geo-hydrological equations about soil stability and erosion included in the model. A short paragraph 2.2.3.3 clarifies that erosion and sediment transport are fully integrated with the hydrological routine while for landslides only triggering condition is evaluated.

2) I am not very convinced by the 9-pixel buffer, nor by your validation using ROC curves (usually the area under the curve is calculated to quantify the quality of the indicator), for example. However, I think all these points would be much easier to accept if you developed a good "Model Limitations" section in your discussion.

We agree with you. The ROC methodology we applied state for identifying if the slope instabilities computed by the model CRHyME (preditor) were similar to those one recorded by the IFFI landslide catalogue (reference) after the rainfall events examined. This methodology is generally adopted for susceptibility mapping but here we have adapted it to measure the single-event model performance. This part will be revised and better explained.

In our opinion, addressing landslide failure to a single pixel instability is quite reductive since the uncertainties are related to local slope stability parameters and territorial morphology (we are using spatially distributed data coming from a worldwide database with different spatial resolutions that are ok for catchment simulation but still have high uncertainties on slope scale). This strategy (9-pixel buffer) was taken into account bearing in mind the possible risk in the surrounding areas (crowns and landslide boundary instabilities) according to regulations reported in the literature. As a result, we have extended the unstable areas a bit to be more conservative. Anyway, we will rethink this part to make the methodology description more straightforward and comment further on the dedicated section you have suggested.

This part (paragraph 2.3.2) was reworked and better explained. 9 pixel is a conservative choice that follows the guidelines proposed by the study conducted by (Harp et al., 2006). We have stressed the fact that all these choices are attempted to deal with the uncertainties that are large for this type of phenomenon. A "Model Limitation" paragraph (4.4) has been created in the discussion section where this and other encountered issues are better commented on with respect to the results obtained.

3) Also, you sometimes write assertions that are too strong or are not well supported by references. I have highlighted some of these in the text. Try to be more nuanced and explain more your modeling decisions.

Thank you very much for this observation. Further investigation in the literature will be done in this sense and a "smoothing" in some parts is advisable according to your suggestions.

Since the text has been profoundly revised we have erased assertions without references making the text more compact, fluid and consistent. The problems raised during the model tests have been commented on with respect to literature references highlighting warnings and giving possible solutions to try to solve them.

4) It is very easy to get lost in the names of different watersheds, subwatersheds, rivers, and stations. So I would recommend having a very clear map showing all the names and referring to it often. In addition, I

would suggest referring to stations with the river names in parentheses afterwards, and always specifying the type of thing you are referring to: "Nure rivers" not "Nure".

Thank you also for this piece of advice. We will provide a better representation of the sites we have investigated to make their identification through the text more clear.

We have also corrected this part both in Figures, tables and in the text

I would recommend this article to be accepted under Major revisions, since its content is very interesting, but the way it is presented does not put it to its advantage. However, I also think that Rejecting it would be a good option, notably because I think that the GMD (Geoscientific Model Development) journal could be a (more) appropriate fit for a resubmission. Indeed, this manuscript could be of interest for audiences in natural hazard, but also for more fluvial geomorphology or hydrology publics.

We thank you the reviewer for the comments and for the precious suggestions that will surely improve our work.

The Authors

Response to REVIEWER N°2

Dear Reviewer

We thank you a lot for the punctual revision of our work. The paper represents a summary of my PhD thesis that I have adapted into a scientific publication. I am aware that the topics involved in CRHyME construction are very wide and touch different disciplines so I am grateful to have received useful hits for improving my work. In the next parts, you can find the reply to your suggestions and observations.

Abbate et al. (2023) presented a modeling tool that simulates various surface processes, including landslides, debris flow, sediment transport, and erosion. The purpose of their paper is to present the structure of this model, along with some verifications conducted at multiple field sites across Italy. While the paper is interesting and offers an approach that contributes to closing the gap in our understanding and toolsets for landslide estimations, I have some concerns that must be addressed before its publication to improve the overall clarity of the manuscript. Although these are not fundamental modifications, addressing them could be time-consuming. Therefore, I recommend this paper for publication after a major revision or rejection at this time to provide the authors sufficient time for resubmission.

Here are the issues that need to be addressed:

The language in multiple places is unclear. I recommend proofreading to improve clarity.

We apologise for the English, and we have planned to do a language proofreading, splitting long sentences that may lead to difficulties in reading. About vocabulary we think that a list of acronymous or keywords should be made in one of the appendices in order to facilitate comprehension by the reader, also subdividing them into categories such as (i.e., climatology, meteorology, hydrology …)

The paper has been completely revised, pay more attention to sentence construction and vocabulary. Some parts were reduced and synthesized in order to make these paragraphs clearer to the reader. We are planning further language proofreading after this turn of revision. We added references that were required. We modified appendixes A and B: Appendix A contains more practical information about the code usage with an example of the .INI file for model initialization; Appendix B contain all the main acronymous used in the model and in this current paper. About the cases studied, a test case within a small catchment (30 km2) of Caldone River (Lecco, Lombardy, Italy) has been included for testing hydrological balance, sediment transport and dependence of the results with respect to the spatial resolution of the digital terrain model (Pars. 2.4, 3.1 and 4.1).

The introduction should include a clear statement about the novelties of this newly proposed model to motivate readers. The authors should answer questions such as what CRHyME brings to the table that other models are incapable of doing and why landslide researchers should consider adopting this model instead of others.

We comprehend that probably our introduction is not so effective for the reader. We will try to fix it, summarizing concepts a bit and giving a clear statement of CRHyME capabilities:

- Versatility with respect to the modelled process (flood, landslide, erosion);
- Climate model outputs and meteorological reanalysis outputs compatibility;
- Fast modelling even though is a spatially distributed model;
- Fast coding, modules modifications, no constraints on spatial and temporal resolution;
- Availability of Opens Source dataset and Python libraries;

- Multihazard approach, case study analysis and future projection
- …

We have added 4 main points that sum up the novelties of this model. In addition, we have clearly stated why the CRHyME model should be preferred stressing the fact that try to cover the gaps on which existing models may fail (i.e no open source, no modular, slower, no flexible to model and initialization data, no geo-hydro integrated, no worldwide reproducibility due to lack of data etc …). Moreover, the introduction was clarified in some passages and adding references where required.

More elaboration on the interactions of different model layers and within-layer processes is needed. Although some of the model components that solve hydrological processes were inherited from another model, providing the full picture of such processes is essential within this paper. For example, the formulations of how groundwater is handled and how percolation/exfiltration processes achieve communications between the soil layer and groundwater layer are not available.

We agree with this suggestion. In the PhD thesis, all these parts have been extensively commented on and supported by literature references. Here we have summarized the main points but we are understanding that further clarifications are needed. However, groundwater is handled considering a hydrogeological basin coincident with the hydrological ones. This is a strong assumption that is constrained by a lack of maps and data describing worldwide the behaviour of local aquifers. The groundwater layer should be intended as another "bucket" under the shallow soil that communicates through percolation (when is not saturated) and exfiltration (when is completely saturated, i.e. groundwater table is close to the topsoil).

Further details were added in paragraph 2.2, in particular in 2.2.2 where the hydrological model is described. About groundwater, it is clarified how the depth of the aquifer was calculated and how was implemented with respect to the superficial hydrology (Percolation refill groundwater). How exfiltration terms are evaluated both in the upper layer and in the groundwater layer is clarified.

The equations are difficult to follow. Although the nomenclature of the terms is provided in the appendix, explanations of some terms are still missing (e.g., Fds in Eq. 8). Furthermore, some symbols explained in the text, are not found in the equations, such as the temperature coefficient, T, not being present in any of the equations (Eq. 12, 13, 14). I recommend that the authors review each symbol in the equations and explain them clearly right after the equation.

Thank you for these details. We have demanded all the documentation in the appendix but probably a "local" description is better. About the T parameter we have already spotted the mistake and we will fix it after the revision.

As for the hydrological module, the same improvements were included in paragraph 2.2.3, correcting equation terms and simplifying their explanation. We added further references where required.

In Figure 7, please use consistent units. For example, m3/s is used in the plots, but mm is used in the error metrics table, which makes it difficult to evaluate the errors. Moreover, the RMSE values for cumulative water volumes in the error metrics table do not offer any useful information. They might be removed.

Ok, we will revise these parts to make them more consistent with the text.

We have proofread all the maps, figures and tables pay particular attention to the description in the caption of all the required information to understand the data.

The same figure indicates that the model predictions of river discharge are overestimated during the rainy seasons predominantly when using the MERIDA dataset, and all model predictions underestimate the discharge during the non-rainy periods (potentially due to the lack of baseflow). This result suggests that more water is partitioned as surface runoff in the model simulations, causing less infiltration. This may be an important issue causing underestimation of the landslides due to the reduced pore pressure in the subsurface. If the authors observed such model bias, please state it in the discussion section.

This represents a key point of our analysis since there we have investigated several things.

First of all, we have noticed that a correct rainfall field computation is essential. Reanalysis datasets are very useful since provide an already distributed field, avoiding spatial interpolation methods of the classical rain gauges. However, reanalysis is again a model product that may suffer from scale dependency. Therefore on a large basin, MERIDA can perform better since errors are less probable (more pixels are taken into account and the rainfall field is generally representative) while over a little catchment, errors may increase due to the lack of accurate resolution.

Secondly, the basins we have investigated are strongly regulated by hydropower plants that may delay and modify consistently the water discharges not only for a singular event but also at a seasonal scale. Therefore, we suspect these biases could be caused by hydropower management that is implicitly hidden in the hydrometer data series. Since the model is under development, we are interested into include in the future version of the model a rough hydropower management plan to improve the results.

Thirdly, we agree with your suggestion where probably infiltration processes and groundwater flows should be better calibrated. In this direction, we have spent several times testing different infiltration models (SCS CN, Horton and Green Ampt) adapting them to work continuously and not at a single event simulation. In this regard, we have planned further investigations.

This part is extensively discussed in paragraph 4.2. The problem has been disentangled highlighting the possible source of errors and also commenting on the sensibility of the model to the reconstruction of the rainfall field. The latter is a sort of "model limitation" (further described in par 4.4) that depends on the quality of the initial data that may perturb significantly the results and outcomes.

In Figure 10c, the model performed weakly during the initial year. The authors stated in the caption that this is due to initial condition uncertainties. I am curious why no spinup simulations were employed to minimize the impact of initial conditions on the model simulations.

In reality, we wanted to highlight that aspect: the CRHyME model needs at least a short period of 1-2 years of simulation in order to reach a "physically consistent" soil moisture redistribution especially if the basin is rather larger. For Trebbia (10a) and Parma (10c) rivers we have included in the series the "spin up" period while for Nure we have trimmed it for clarity. We will revise this part (the graphs) to explain better this topic and make it less ambiguous also in the main text.

This part has been added and better explained in the result section. SPIN UP simulation has been implemented in each of the model's runs we have carried out. Furthermore, the capability of CRHyME to make a restart, taking in memory the final states at a particular instant, has been highlighted in the discussion as a feature able to reduce simulation-time consumption.

In section 4.1, the hydrological performance of the model was discussed. While I agree that simulating the water cycle correctly in such basins is challenging due to the complexity of subsurface lithology and highly variable topography, the validity of the hydrological part of the proposed model may be more easily verified in a small, heavily monitored, less complex basin to ensure that the model performs satisfactorily. The authors should consider adding such model verification on top of existing test sites to verify the model's performance hydrologically.

A sensitivity analysis has been already done in two small catchments (less than 30 km2) located one in the Alps and the other in the Apennines. We have not included these analyses to reduce a bit the length of the paper but in principle, a short paragraph could be included. In these two catchments (well monitored) the performances of CRHyME have been tested also considering different spatial data resolution and different temporal data resolution. The performance of CRHyME was encouraging since hydrological and sediment transport assessment has been verified by on-site measurements (hydrometer and sediment bathymetries over a check dam).

We thank you the reviewer for the comments and for the precious suggestions that will surely improve our work.

The study about the Caldone catchment has been included in the paper. Caldone catchment is a well-monitored catchment near Lecco (Lombardy Italy) where several studies and field monitoring were conducted by Politecnico di Milano. In Caldone catchments the assessment of hydrological balance, spatial resolution data dependence and sediment transport calibration has been carried out. Further details are described in paragraphs 2.4, 3.1 and 4.1.

The Authors

[PS] We are planning the English proof reading sooner after your revision of the new corrected version of the paper.