Review of La Palma paper by Stéphane Abadie and co-workers

This paper is an elaborate study of the effect of three scenarios due to slope failure from La Palma Island on tsunami impact in French territories. It reports large amount of analysis and work, and there has been surely large efforts behind producing these outputs, in particular given that some of the tools employed such as the CFD THETHIS code are demanding to operate for such purposes. It is an important study because of the practical implications. On the other hand, the elaboration is also a drawback of the paper. There are many different models used, to illuminate different types of physics, merged with an attempt to make an impact assessment (the authors uses the term hazard). Moreover, the paper seems to have undergone several previous reviews with large changes, and would benefit from a better organisation.

Authors response: **Indeed, we were very surprised not to see our revised paper ending up with the initial reviewers. It is obviously more difficult to please a reviewer which did not ask for the specific changes made on the initial submission.**

Some related general comments are summarised briefly below, followed by a long list of line-by-line comments. These comments must be taken into account in a possible revision of the manuscript.

General comments:

It is not clear why mane different models are used for various purposes. I would have liked a simpler strategy where the authors choose a simple set of models. The physics is well known: the tsunamis are dispersive, and we need nonlinear shallow water models for the inundation. Right now, there is a patchwork of models, even including analytical solutions (which I suggest to remove), and it is hard to understand why a given model is used where. While I would suggest that this is much simplified, I would probably expect that the authors would like to keep as much as possible of these results. Hence, as a minimum, a much tightened up introduction is needed to much better explain the scope and how the different models are used, and why. I would also suggest to better distinguish impact studies and studies of physical effects (e.g. dispersion, model comparisons).

Authors response: **We thank the reviewer to open the door for a doable revision for this matter. Indeed, this work was performed in the framework of a national project gathering several French institutes developing or using different models. One of the underlying principle of this paper was to compare models and advertise the work performed within this project and therefore not to exclude anyone (also for political reasons). This principle makes the organization of the paper a bit difficult as maybe its reading. Considering the comment of this reviewer, we explained in the new version of the introduction this project aspect and defend its interest. [This response is labeled Response (1) for the next similar questions]**

See paragraphs added at the end of the introduction (p3, l3):

"Computations performed by Gisler et al. (2006) or Abadie et al. (2012) were both based on inviscid or quasi-inviscid slide flow. In the present paper, the computations carried out in Abadie et al. (2012) are redone, improving their accuracy by calibrating the slide fluid viscosity in order to approach a granular slide (Sections 2.1 and 3.1) with a Newtonian model. Then, the same filtering process as in Abadie et al. (2012) is applied with the new wave sources to produce a wave signal which can be propagated by depth-averaged models (Sections 2.4 and 3.2). The three wave sources are then propagated using FUNWAVE-TVD (Section 2.2.1) and the results in the Caribbean Sea, in Western Europe and in France (Section 3.3) analyzed.

One of the goal of the TANDEM program was also the comparison of the models developed or used by the different partners of the project namely: Calypso developed by CEA, Telemac2D developed by EDF, Funwave-TVD used by BRGM and SCHISM by Université des Antilles. Here we take the opportunity of this case study to compare models on a real case and analyze the differences. The interest is double. This project involves partners who are already in charge of tsunami hazard

assessment while others may play a role in this field at the national level in the future. The first interest is to provide an inter-comparison of the codes used at in the different institutes. This comparison will be valuable for future operational use. On the other hand, this comparison is made on a real case, therefore including all the inherent complexity and uncertainties (bathymetry, mesh, numerical parameters, physical parameters, etc.) usually associated to a practical case. Such a comparison is rarely attempted in usual benchmark exercises which focus more frequently on specific processes such as run-up, tsunami generation, etc..in order to make the interpretation easier. Nevertheless, even though the analysis is not straightforward because models are not based on the same assumptions, numerical methods, mesh types, a comparison including all the complexity may also be of interest as it allows to judge all the effect at once and potentially lead to practical recommendations valuable for future studies. Therefore the originality of this comparison on a real case is the second interest of this part of the study. Accordingly, the rest of the study is organized around a comparison of the different model results (see Sections 2.2.2 for description and 3.4 for the results comparison). Finally, tsunami impact is assessed in different areas in Section 3.5, and results interpreted and discussed in Section 4".

Another major issue, in particular when reading the introduction, is that you sense that the hazard study is attempting to make a best estimate of a landslide motion and wave generation based on laboratory glass bead experiments. However, nature will not behave this way, and there is a considerable uncertainty related to the process and the sliding material. Granted, one cannot perhaps expect that the computations can cover all these uncertainties, but as a minimum, the authors must make it crystal clear that there can be a much larger variability related to the tsunami generation and tsunamigenic strength. This is a limitation of the study.

Authors response: **We acknowledge this limitation. [This response is labeled Response (2) for the next similar questions]**

See new paragraph added in the discussion on that matter:

Second, we used a glass beads based experiment (Viroulet et al., 2013) to calibrate the Navier-Stokes simulation of the La Palma slide. If this is an improvement compared to the very coarse inviscid initial estimation (Abadie et al., 2012), which should be more considered as a worst case, such a laboratory experiment still is a huge simplification of the complexity expected in areal volcano collapse. An accurate description of such a complex process at real scale is still beyond the capabilities of current models. Therefore, there is here a very important source of uncertainty which the reader has to be aware of and this uncertainty propagates and affects the impact results. Furthermore, this work is not an hazard study which could have been performed for instance by considering different values of slide viscosity but at much higher computational cost. The position of this paper is rather to give an illustration of what could be expected from such an event by presenting results at least consistent with the current state of the art in terms of laboratory experiments and therefore propose an improvement compared to the previous published results on that case.

Finally, the title tsunami hazard is misleading, because the authors do not address return periods, in addition to lacking a proper treatment of the variability or sensitivity to landslide parameters as noted above. The title should hence be revised to take this into account.

Authors response: **We agree with the reviewer that the term hazard was used inappropriately in the initial version of the paper. It has been removed when possible from the manuscript (except at the beginning of the introduction). Moreover we have added a paragraph in the discussion section about this limitation and one of the reference suggested below (Grezio et al., 2017) [This response is labeled Response (3) for the next similar questions]**

See new paragraph added in the discussion on that matter:

Of course there are some limitations in this study which may provide the basis for future improvements. First, this study should not be considered as a hazard assessment stricto-sensu because the return period aspect is not considered and the sensitivity in the landslide parameters not covered extensively. For a review on Probabilistic Tsunami Hazard Analysis (PTHA) methods, the reader is referred to Grezio et al. (2017) for instance. Instead, the current study presents plausible particular scenarios based on state-of-the-art numeral models. Note that the Navier-Stokes model, which provides interesting information for this kind of processes, is still too heavy to be employed in PTHA computations.

Detailed comments:

Title: Probably the term "tsunamigenic strength from potential events" is better than hazard. After all, hazard refer to a temporal component, and should not really be used if return periods are not considered.
Authors response: **See response (3)**

New title: La Palma landslide tsunami: computation of the tsunami source with a calibrated multi-fluid Navier-Stokes model, impact assessment, and model intercomparison

Page 1 line 5: "for 5 minutes" --> "after 5 minutes". ***Authors response:*** **Done**

Page 2 line 8: "allow studying impact on France and Guadeloupe". Here you maybe emphasise more strongly that this is the scope? After all, the impact locally would be a more natural focus.
Page 1 - line 8: "Although the wave source seems to be reduced due to the rheology..." --> "Although the rheology applied in this study seemingly leads to smaller waves..."
Authors response: **Correction made (see point right after):**

Although <u>the slide modeling approach</u> applied in this study seemingly leads to smaller waves

Page 1 - line 9: add "mu(I)" ahead of rheology – ***Authors response:*** **Not appropriate. The approach used here is a calibration of a Newtonian model – the mu(I) is just used once in this paper to justify this approach, hence the correction made (point right before).**

Page 2 - line 7: It the term hazard is used properly, it would be useful to introduce a definition, and refer to at least one key paper. Use e.g. Grezio et al. (2017): Grezio, A., Babeyko, A., Baptista, M. A., Behrens, J., Costa, A., Davies, G., ... & Harbitz, C. B. (2017). Probabilistic tsunami hazard analysis: Multiple sources and global applications. Reviews of Geophysics, 55(4), 1158-1198.
Page 2 - line 13: On the complexity of these processes, please refer key review papers, Løvholt et al. (2015), Yavari-Ramshe and Ataie-Ashtiani (2016): Løvholt, F., Pedersen, G., Harbitz, C. B., Glimsdal, S., & Kim, J. (2015). On the characteristics of landslide tsunamis. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 373(2053), 20140376; Yavari-Ramshe, S., & Ataie-Ashtiani, B. (2016). Numerical modeling of subaerial and submarine landslide-generated tsunami waves—recent advances and future challenges. Landslides, 13(6), 1325-1368.
Authors response: **See response (3)**

Page 3 line 9: Please update this sentence to say that you use mu(I). I think this is clearer than saying calibrated slide viscosity. ***Authors response:*** **As stated before, the mu(I) rheology is only used in one simulation in this paper. So mu(I) is not added in the sentence but in place, the sentence has been changed to be clearer on that point.**

Initial: In the present paper, the computations carried out in Abadie et al. (2012) are redone, improving their accuracy by calibrating the slide fluid viscosity in order to better represent a granular slide (Sections 2.1 and 3.1)

changed to: In the present paper, the computations carried out in Abadie et al. (2012) are redone, improving their accuracy by calibrating the slide fluid viscosity in order to approach a granular slide (Sections 2.1 and 3.1) with a Newtonian model.

Page 3 - line 29: "first instance of the motion" --> "initial motion". ***Authors response:*** **Done**

Page 3 - line 30: "solver code category" --> "type of solver". ***Authors response:*** **Done**

Page 3 - line 33: "close but not completely equivalent to models, also use to simulate landslide tsunami generation" --> "more sophisticated with respect to the slide motion than models such as" ***Authors response:*** **Done**

Page 4 - line 9: For a complete review, discuss also the model of Si et al 2018:
This model is more sophisticated material wise, but probably not able to tackle operational environments yet: Si, P., Shi, H., & Yu, X. (2018). Development of a mathematical model for submarine granular flows. Physics of Fluids, 30(8), 083302. ***Authors response:*** **Reference added and discussed as requested.**

Added:
Finally, Eulerian-Eulerian two-phase models such as the one described in Si et al. (2018b) and Si et al. (2018a) are very promising approaches able to describe the flow within the grains as well as the grain/grain interactions but their applicability to practical cases has not been demonstrated yet.

Page 4 line 22: Clarify where Newtonian and mu(I) rheologies are used, maybe reformulate: "Both Newtonian and mu(I) rheologies are used in the simulations". ***Authors response:*** **Done**

Page 4 - line 23: The experimental results cannot necessarily represent the real case realistically (glass beads are far from a realistic rock slope material). Hence, all the different viscosities may represent the reality, and should not be calibrated towards a single dataset. This is actually a misconception, the hazard analysis should ideally include this as an uncertainty. Hence please reformulate. ***Authors response:*** **done and see response (2)**

Page 5 - first paragraph. Please see above comment. I dont believe a single calibrated result represent the reality realistically. This does not mean that new simulations should be done, but the authors should make the reader aware of this uncertainty.
Authors response: **see response (2)**

Page 5 - line 15: Please explain that this is just possible value for the material parameter, and there is likely a rather large uncertainty that is not covered in our analysis. Otherwise, the reader gets the false impression that the wave generation is deterministic, which it is'nt.
Authors response: **Paragraph modified accordingly and see also response (2)**

To extrapolate these results for the La Palma computations, the following reasoning is adopted. First, it is assumed that the real slide is well represented by the granular medium used in the experiment. This approach is not deterministic as there is important differences between this experiment and the real case but at least it may be considered as a better assumption than the worst case scenario presented in Abadie et al. (2012)

Page 7 - line 3: Please clarify "can be upgraded"? Do you mean that it also contain dispersive features. In this case reformulate. It is BTW not clear why two types of dispersive models are used. Does this code have wetting and drying facilities?

Authors response: **Sentence reformulated:**

The user can choose to solve either the non-dispersive (NSW) or dispersive (Boussinesq model following Pedersen and Løvholt (2008)) non-linear long wave equations, written in spherical coordinates.

As explained in the text, the switch between non-dispersive and dispersive equations is realized between mother and daughter grids.

Yes, the code has wetting and drying facilities. It has been added in the text:

The wave impact assessment is realized using this mixed method for the French coasts and calculating run-up with wet and dry conditions.

Page 7 - line 25: Again, why is this model used? It is not clear why so many seemingly similar models are used, please elaborate.
Authors response: **see response (1)**

Page 7 - line 34: "In this work..." do you refer to Telemac? The meaning is not clear.
Authors response: **replaced by:**

In this work, the mesh <u>used in Telemac-2D</u> has 12.5 million of...

Page 9 - line 17: This is not a proper hazard assessment. Impact analysis or scenario analysis are better terms.
Authors response: **see response (3)**

Page 9 - line 22: Again, I miss the reasoning for choosing this model, and why other models are employed elsewhere. This is generally quite messy. You need a structured introduction upfront in the paper explaining these choices.
Authors response: **see response (1)**

Page 9 - line 31: Again, this is not hazard, probably something else but not hazard... Please revise sentence. ***Authors response:*** **"hazard" replaced by "impact" in the sentence**

Page 11 - line 4: Delete double "smaller" ***Authors response:*** **Done**

Page 11 - new paragraph marked red: Not clear what this paragraph add, it is confusing. We have repeatedly shown the effect of dispersion in previous studies. I dont see the need for doing this again, it disrupts the text.
Authors response: **It is very challenging to please successive reviewers who does not automatically always share the same point of view. Fourier transform analysis was explicitly requested by one of the former reviewer, hence this first revised version. We feel logical to keep the successive changes requested throughout the review process to respect this process.**

Page 11 line 31: This was analysed in more detail first by Løvholt et al. (2008), please notify and provide reference. ***Authors response:*** **Done**

Page 13 - line 27: Delete double punctuation. ***Authors response:*** **Done**

Page 13 - line 28: Again, this is not hazard assessment, but only an assessment of possible inundation or impact. Please revise title. ***Authors response:*** **done and see response (3)**

Page 13 - first three paragraphs of section 3.5: I find all this analytical analysis strange for a phenomena so strongly controlled by local phenomena. Why not limit the impact analysis to the local inundation study. I would suggest to skip this part, and only keep the part using NSW inundation analysis. The paper is overloaded with results, and this is for me a distraction. Moreover, such a rough analytical analysis could be worthwhile for assessing the hazard region, but not for a local analysis.

Authors response: **this aspect has been totally removed from the article.**

Page 15 - line 5: As said above, the authors does not seem to take into account that the dynamics and material behavior is uncertain, and that a simple glass bead experiment cannot be conveyed to real situation. The paragraph should be rewritten to better reflect this. Granted, the simulations fit better the experiments, but the authors have no guarantee that the slope failure will behave this way. Probably it will not.

Authors response: **see response (2)**

Page 15 - line 15: Again, please replace the term hazard assessment with something more appropriate, such as an impact assessment. The study is not broad enough and does not cover return periods, so cannot be coined a hazard study. ***Authors response:*** **replaced by "impact"** ***and*** **see response (3)**

Page 15 - line 31: This discussion of model effects is too long. I would suggest to shorten it dramatically, as results are shown above and the physics is well-known. Besides, the effects of dispersion have been investigated in previous studies. It can also be analysed with a dispersion number (e.g. Glimsdal et al., 2013)

Authors response: **We understand the point of view of the reviewer, but this discussion is justified in the context described at the end of the new introduction (p 3, l 9) (model comparisons and recommendations). It was also meant to answer the first reviews of the paper.**

Page 16 - line 23: Wynn and Masson found upward fining, which indicate long separations in time. This means that this was no real retrogression, but more likely separate events. On the other hand, I agree with the authors statement in the last part of this paragraph.

Authors response:

The present work did not explicitly take into account the possibility of a retrogressive scenario. Whether the flank collapse occurs en masse or in successive stages is obviously crucial in terms of wave generation. In this study, we proposed several slide volume scenarios which can be used for a crude assessment of the wave reduction in case the collapse occurred as several separate events with no interactions between the successive slides (e.g. the 20 km scenario may give an idea of what would happen if a 80 km slide were occurring progressively or in sequence). The interactions could be left for future research even though field evidences tend to show that these collapses may have occurred as separate events (Wynn and Masson, 2003) rather than in an actual retrogressive way.

Page 17 - line 13: See comment above several times on uncertainty, and reformulate accordingly.
Authors response: **Sentence modified**

Initial: The new wave source is reduced in half compared to previous estimations mainly due to the improved rheology calibration

changed to:
The new wave source is reduced in half compared to previous estimations mainly due to the larger value of slide viscosity used in this work

Page 17 - line 20: This sentence is not well formulated, I dont fully understand what you mean.
Authors response: **done**

Initial sentence: After 15 minutes of propagation in a Boussinesq model, the wave signal is still dispersive and therefore Boussinesq models should be recommended to use the source provided

modified as: The tsunami source calculated in this paper after 15 minutes of propagation in FUNWAVE-TVD and proposed to the community in the SEANOE repository is dispersive and therefore we recommend to use appropriate models (e.g., Boussinesq models) to propagate further this source in future studies.

Figure 8: Slide contours are very difficult to read. I suggest fewer and larger figures allowing the reader to see the details.

Authors response: **The Figure has been split in two figures (Figures 8 and 9) so as to respect the reviewer's wish.**