Lava flow hazard map of Piton de la Fournaise volcano
Chevrel et al.

In this manuscript, Chevrel et al. present an eruptive history database and methodology to produce numerical lava flow simulations and a probabilistic assessment of lava flow inundation hazard at Piton de la Fournaise. Their method builds on previous work done by these authors at other volcanoes to incorporate the complex spatiotemporal eruptive history of Piton de la Fournaise and, as they importantly highlight, the impact of topographic change at a frequently eruptive volcano on evolution in lava flow hazard. Overall, the manuscript is very good and will be of interest to a broad audience. The data, methods, and analysis are all appropriate, but some elements of the methodology and key discussion points are missing or wanting of more details. Below I summarize some major comments, questions, and suggestions. Pardon the length, these comments are largely minor, this is just a topic I am very interested in. There are also a number of typos and word order errors that just require a close reading with fresh eyes by the authors to fix.

Major comments

Using recent eruptions to assess methods and implications
A unique and powerful element of this work is the assessment of lava flow hazard through time using DEMs and eruptive history data that evolve over recent decades. Although these are used to generally state that changing terrain impacts hazard mapping and that recent flows generally occurred in areas that were previously deemed more likely, a missed opportunity is to use these maps and data to do a full hindcasting assessment and validate the method being presented. How well does the map based on data up to 1997 work for eruptions after 1997? Can this be quantified (e.g., Bevilacqua et al. 2017)? Is it better initially and then gets worse as topography changes more and more? What about up to 2009? This study has the distinction of allowing this critical discussion of how well we expect maps like this to do, what timeframe they're useful for, and where they fail given the availability of data over time and very frequent eruptions. These are topics that are all touched on but they are not framed in this fundamental way, where a hindcasting assessment would organize and strengthen the key conclusions. Showing that the method works well for recent decades would demonstrate validity and support further applications. I recommend adding a section to the discussion on this ("Validating hazard mapping with recent eruptions") and accompanying Fig. 7 with quantification of how well the map did for later flows through time. It is also regularly mentioned, but never demonstrated how topographic changes impact hazard assessment. Since this is a major conclusion of this work, and testable with the time series of hazard maps and eruptions, this is a good opportunity to show the changes caused by topographic evolution with visuals.

Treatment of spatiotemporal variability in eruptive behavior
As mentioned and assessed throughout the paper, there is significant variability in the frequency and style of eruptions both spatially and temporally at Piton de la Fournaise, including rare rift zone events, frequent (but episodic?) summit activity, and cyclic patterns in

eruption location and magnitude in recent decades. Currently, these are integrated empirically into a conditional hazard map (probability of lava inundation in the next eruption) that incorporates historical data or geologic mapping at various timescales depending on location. However, in the discussion it is then stated that this approach does not integrate all magnitudes or timescales depending on what the “next” eruption is. Certainly, eruptive history that is left out of the input data for hazard mapping will not be represented (e.g., lava lake activity that is not well preserved in the geologic or historical records), but it is unclear why the authors also do not feel that the map represents events that are included (e.g., recent large eruptions that end each eruptive cycle). If the input data are not representative, it would be helpful to offer suggestions for how they could be represented moving forward, or utilize a different methodology for the treatment of these spatiotemporal patterns. Similarly, if the results have limited utility to specific time periods or scenarios, incorporating those explicitly, such as producing different conditional probability maps for specific conditions or regimes (e.g., just the large events that occur only after smaller ones) or a map of probability of any lava inundation over a given time period (1 year, a decade, given that the recurrence rates are well characterized), would help in the discussion. Overall, it seems like the map in Fig. 6 is perhaps more representative of the current state of the volcano, and applicable on the scale of years to decades, than the text implies, because of the relatively complete recent record.

In terms of discussing the results, adding/reorganizing the discussion around the "Impacts of cyclicity (or spatiotemporal variability in eruptions more generally) on hazard assessment" instead of methods-based section headings (e.g., 8.1.2), would help focus discussion on a key topic that repeatedly is mentioned throughout the other sections. These cycles are introduced in the background section and talked about a lot in the discussion as changing in location and magnitude through time, but the reader is not shown what these look like or exactly how their properties evolve. This discussion should therefore be accompanied by a figure that shows this well. Maybe marking vent locations and flow extents by cycle position (in different colors) in map view and comparing this to hazard mapping results.

Methods questions

Input data for hazard mapping includes vent locations, as well as statistics on the number of lava flows, their lengths, etc. However, these are not clear in the methods given the potential for a given eruption to produce vents that are actually lengthy fissures, multiple vents, and multiple lava flows. Some clarification on how vents, flows, and locations were defined, any implications of these choices on the results, and references to literature on these challenges (e.g., Cappello et al. 2012; Runge et al. 2014), are needed. Given some uncertainty in how these may be defined, the high precision and number of significant figures in the tables seems overestimated and should be discussed.

For the application of DOWNFLOW, were these simulations run from a single vent or from all vents and/or a fissure geometry (lots of point sources along a line, say)? How were ocean-limited flows treated in defining the lava flow length distribution (line 305)? The DEM resolution is very important to the performance of DOWNFLOW, and separate calibrations are appropriate, so please provide the resolutions when describing the methods (e.g., line 287) and in the discussion of the results (line 453). In Fig. 4a, it is not clear that the calibration was completed though, given the best-fit seems unconstrained in the parameter-space (Δh >5?). How do multiple vents affect DOWNFLOW modeling/misfit? For long flows, the fit seems much worse (Fig. 4a) - greatly overpredicting the inundated area. Would a different model work better for these? Were other models considered? Can uncertainty from this type of misfit be propagated in this methodology?

Additionally, the impact of DEM resolution is worth testing with regard to varying Δh between time periods/DEMs, rather than attributing all differences to thickness changes (e.g., 455). Coarser DEMs (such as the 1997 DEM v. the >2008 DEMs) have built-in flow spreading from pixel size, potentially greater DEM uncertainty, and significant effective smoothing, which generally seems to change the best-fit Δh significantly. If you resample the later DEMs to 25 m instead of 5 m, does the best-fit Δh change? There may be changes related to flow thickness, too, but DEM resolution should be integrated into this discussion (also in terms of validation – does the 25 m data perform worse?).

Quantitative analysis of hazard probabilities
There are a number of places where results are described, but using inundation values for a given pixel, rather than integrating the results to answer the broader questions, such as "what is the probability of lava inundation in the Enclos during the next eruption?" (line 22) or "what is the probability that the next lava flow will intersect the coastal road?" (line 339, this is written as if integrated, but it's just the probability within a given pixel, not along the whole road). The actual probabilities are possible to calculate with these data though (the first is dictated empirically, but the second could be calculated based on the model results).

Line-by-line comments

33 – Many lava flow hazard maps also incorporate time (probability of inundation at a location over a given time interval), as opposed to just a conditional probability on the condition of an eruption occurring (e.g., Bebbington 2013; Cappello et al. 2015). Given the known recurrence rates in this study, this extension would potentially be possible here as well, and relevant for applications to land use planning.
64 – "complete" – that seems risky statement given how extensive burial is also described
86 – "Morphological" – replace with "Geological"
120 – Introducing these here as low frequency, high impact events will set up why they are so difficult to forecast in the discussion. They're in the data, but they are not the "most likely" event, and less spatially concentrated.
150 – Expand this to fully introduce the spatiotemporal patterns in eruptive history. Given that episodes on the order of centuries are also later invoked for the methods of database assembly and the short/long-term applicability of the analysis, these could also be introduced here in this background section.
192 – Using the longest flows only will overpredict flow length and yield higher probabilities overall. Perhaps add more explanation (as mentioned in the 'methods' questions above)

Tables 1 and 2 – Are all flows mapped or are some flows that are unmapped (or poorly mapped) but were reported included in these counts? I think since these are directly compared to each other and they have the same rows, it would be much easier for the reader to just combine them into one table.
236 - Is this every single cone, spatter rampart, and fissure? Or grouped/summarized somehow?
242 – What function is used for bandwidth, is it symmetric or asymmetric (or fissure geometries) and why?
Section 7 – The first sentence sets up the goal of 'the next effusive eruption", but this is also making retrospective maps for other timespans, so make sure the plan of making multiple hazard maps that represent different times to look at any changes over time is introduced at the start of section 7. (will make line 323 less confusing on the first read).
Sections 7.1 and 7.2 are interpretive and should be in the discussion. 7.1 can become 8.1 and describe the results, including the integrated example probabilities mentioned above. 7.2 can be modified/replaced into the core validation exercise, comparing the maps and subsequent flows more quantitatively/explicitly to show how well the method works. Overlaying flow outlines directly on the hazard maps may be especially helpful. Where it works/doesn't work will then set up the following existing/re-organized discussions of spatiotemporal cyclicity, short/medium/long-term applications, importance of topographic updates, etc.
334 – Clearly state that these are conditional probabilities
Section 8.1.1 - Overall, a 'long-term' hazard map using present-day topography in a frequently active area is most useful on the scale of the next few decades. A short-term forecast map (hours to a month, say) should be using short-term vent opening probability map integrating monitoring data, not the whole volcano over centuries (and up-to-date topography). A century-scale map needs to incorporate the potential of the volcano to change regimes entirely (it's been stated that eruption frequency can vary spatially and temporally over centuries at Piton de la Fournaise) and the present-day topography is expected to change over time (although it will likely be similar for a long time, large-scale subsidence and such can happen in addition to lava emplacement). So maps like this are most useful in that next years-decades period mostly useful for annualized to lifetime-scale planning and risk, but not shorter or longer term. Products for these other timeframes can also be produced with the same methodology, just with different input data, and that is worth saying here.
425 – This contradicts line 303 (although it's broadly true of course!)
429–445 – It seems like these events are represented accurately in the input data (infrequent, less spatially confined), and thus the hazard map. That they occur very sporadically at locations that erupt infrequently does not mean the map is wrong (low-likelihood, large-magnitude). Based on this work, are there recommendations for how these should be treated, or maybe presented differently (e.g., as "worst-case" scenarios or similar)? What work would be needed to include lava lake activity in a future hazard map?
471 – A good place to include some numbers of the integrated probability of trails being inundated in the next eruption, or visually showing the exposure on the map.
477 - Can you reconcile this? It is an aspect of drainage/integration by accumulation from many possible vents, but it's extremely nonintuitive to readers and the public. Would it be valuable to

make a "proximal" hazards layer, say, using a vent opening contour to define a proximal hazard region subject to tephra, ballistics, gasses and near-vent lava flows?
487 - Explicitly highlight that in a short-term or atypical scenario, this framework allows updating of the input data (DEMs, vent locations, flow properties) to quickly produce a probabilistic map or specific flow forecast scenario as needed.
Fig. 6 – I don't see green or white lines. Perhaps adjust line widths and colors?