We would like to thank the reviewer for the comments and constructive suggestions relating to the underlying review of manuscript number *nhess-2017-166*. Please find below the authors' replies (in blue italics) to each of these comments:

L. Capra (Referee)

lcapra@geociencias.unam.mx

Received and published: 16 June 2017

The paper represents an original contribution aimed to defined lahar occurrence, that represents a very useful tool to be implemented in volcanoes where lahar monitoring systems are not available, or to anticipate the occurrence of an event respect to an early warning system. The model is based on two years records of lahars and their associated rainfalls of the Belham River Valley at Soufrière Hills Volcano, Montserrat. The 1-hour rainfall intensity is used to correlate lahar occurrence in dry and wet season, and lahar probability is defined considering also the 3-day antecedent rainfalls and the catchment evolution. The paper is well organized and nicely illustrated.

I have identified some points that need to be better discussed:

A more detailed description of how lahars were grouped in these three different categories is needed (small, medium and large) at least indicating which the main differences are: i.e. duration, magnitude (i.e. maximum amplitude from the seismic record?); runout, flow-depth?

- *Increased information can be included in the manuscript regarding the magnitude categories assigned to the lahars. These categories were assessed using visual inspection of the degree of channel inundation and flow depth (where possible); in addition to the assessment of the duration and amplitude of seismic signals. Lahar signals show continuous readings in the 2-5 hz and peak at approximately 30 hz. The highest recorded amplitudes are associated with discharge and solid load in the lahar (based on observations). Lahar signals were cross referenced to visual observations and carefully excluded from signals associated with primary activity and other seismic noise (such as construction vehicles).*

Can author also provide a simple description of these lahars, if they are debris flow or hyperconcentrated flow?

- *Detailed observations of lahars in the Belham River Valley have indicated that they are Newtonian and fully turbulent (Barclay et al., 2007; Susnik, 2009; Alexander et al., 2010; Froude et al., 2017) This interpretation is based on sampling of several small and large events and two detailed studies of flow deposits (2006-2009 and 2012-2015). Further details may be provided, however detailed observations of a flow and associated previous studies are fully referenced in Froude et al. (2017).*

In addition will be useful to have a table with rainfall characteristics (total accumulated rain, peak intensity) for some selected lahar events, some examples for each lahar category (small, medium, large) in dry and wet season.

- *The authors agree and feel that a multi-part figure illustrating the timeline of several rainfall events and the associated lahar activity (size, timing and duration) and rainfall characteristics (timing, cumulative rainfall and peak intensity) could be added to the manuscript and would be of significant benefit to the research.*

Why 1-hour rainfall intensity is here considered? Is a limitation due to the record? I don't know the weather conditions at Monserrat, but in other volcanoes (i.e. Merapi and Colima for example) especially for orographic rains (in the "dry" season), rainfall intensity is calculated over a 5 o 10 min. window, which is much more representative of these type of rains, of short duration (< 1 hours) and high intensity. Do shorter rainfalls (< 1 hrs) have triggered lahars at Montserrat? Is 1-hour peak intensity representative of different rainfall behaviors at Montserrat? Would you expect any difference in your model with a 10-min. peak rainfall intensity?

- *The reviewer is correct in identifying that 1-hour rainfall intensity was utilised in this study due to a limitation of the record (it was the maximum temporal resolution available). As noted by the reviewer, at other locations including Colima, Merapi and Tungurahua, 10-minute rainfall has been utilised and this has benefits in terms of assessing lahar triggering rainfall from short-duration high-intensity rainfall events which frequently occur in the tropics (e.g. Lavigne & Suwa, 2004; Capra et al. 2010; Jones et al. 2015). Short duration rainfall has resulted in lahars in the Belham Valley within the studied dataset and increased temporal rainfall data resolution would certainly be advantageous if available. However, the 1-hour approach has been demonstrated to be an effective basis for the methods developed in this study (Lavigne et al. 2000; Lavigne & Suwa, 2004; Jones et al. 2015). If incorporated alongside the current 1-hour peak rainfall intensity, 10-minute rainfall intensity could potentially be expected to further* *increase model performance by more appropriately capturing lahars triggered by short duration, high-intensity events. A discussion point relating to this concept could be added to the manuscript.*

Line 116. How the 1-hourPRIs threshold is defined?

- *In this study 1-hour peak rainfall intensity is defined as the maximum rainfall recorded in one hour during a single rainfall event. A single rainfall event is defined as a period of recorded rainfall in between two dry spells of six hours or longer. The 1-hour PRI thresholds referred to in the manuscript separate the dataset into those rainfall events which exceeded a given peak intensity threshold and those which did not, and examines the rate of lahar occurrence in each case. More detail regarding these definitions can be incorporated into the manuscript for clarity.*

Line 124-129. From figure 2 at least two large lahars occurred in the dry seson2, with accumulated rainfall less than 20 mm for at least one of them. There are any evidences of hydrophobicity? Which type of vegetation grows at Soufriere Hills volcano?

- *Prior to the onset of eruptive activity 62% of the Belham Catchment was densely vegetated with Dry Forest (29%), Mesic Forest (48%) and Wet Forest (13%), with dry forest subsequently identified as the dominant species found on re-vegetating pyroclastic deposits (Froude 2015). Previous studies in the Belham Valley have not identified evidence of hydrophobicity, such as previously identified at Colima by Capra et al. (2010). In the Belham Valley increased vegetation damage has been identified as increasing lahar occurrence (Barclay et al, 2007; Alexander et al, 2010) and increased lahar activity late in the wet season attributed to increased deposit saturation and decreased infiltration rates (Barclay et al, 2007). Figure 2 displays hourly rainfall and whilst it is correct that neither of the two large lahars in dry season two were triggered by rainfall events featuring 1-hour PRI values of >20 mmhr^{-1}, they were associated with rainfall events with significant total rainfall values of 39 mm (29/11/2011) and 22 mm (19/04/2012).*

In addition, small lahars are more common in the wet season. For example during dry seasons 1 and 2 only medium (and 2 large) lahars were recorded and small events are only observed in the wet season. Please add some consideration about this behaviour in the discussion section, at line 215-218.

- *Small events are indeed more common in the wet season, a factor attributed to "flash flood" responses to rainfall during periods of increased antecedent rainfall. Small magnitude pulses of lahar activity did occur due to rainfall during dry seasons 1 and 2, however these often occurred during rainfall events which also triggered larger magnitude pulses and as such the small pulses are superseded in Figure 2.*

Line 140-141. "This indicates that more intense rainfall is required to trigger lahars in the dry season than in the wet season." Can author please discuss this behaviour? Is this correlated with a higher permeability of the substratum in the dry season? How much rains accumulate during these high intensity events in the dry season?

- *The dataset indicated that lahars were statistically more likely to be triggered for a given peak rainfall intensity in the wet season compared to the dry season. This is thought to be a product of increased infiltration rates in the dry season associated with generally lower levels of antecedent rainfall. In terms of individual dry-season rainfall events that did not trigger lahars (of sufficient magnitude to be detected on the seismic records); 64 mm of rainfall was recorded on 4th/5th January 2011 and 73 mm on 4th/5th December 2011 without any recorded lahars. Recorded 3-Day antecedent rainfall was less than 3.1 mm at the onset of both rainfall events.*

Line 165: 3-day antecedent rainfall values is a common time interval also used in previous works, such as at Colima volcano, please add some references.

- *Absolutely, additional references including Capra et al. (2010) to the prior use of 3-day antecedent rainfall will be added. Information and references will also be included regarding the previous use of other timescales (including 24-hour and 7-day antecedent rainfall) and how 3-day rainfall was chosen as the optimal timescale within this study.*

Line 166. Can authors be more specific about the definition of the term "total cumulative rainfall since significant eruptive activity"? In their model will be the total rain since Phase 5? And, how this term reflect the catchment evolution?

- *The reviewer is correct, the term "total cumulative rainfall since significant activity" reflects the total rainfall since the end of Phase 5. This parameter is used as a proxy for catchment evolution within the model under the assumption that in the absence of further eruptive activity hydrogeomorphic drainage basin recovery will occur following the catchment disturbance associated with phase 5 (Pierson & Major, 2014).*

Line 215-218. This point needs a better discussion in light of Figure 2 (see previous comment at line 124-129).

- *As the reviewer identifies in their comment relating to line 124-129, large lahars are not exclusively triggered in the wet season and there are examples of large lahars in the dry season. However, the primary objective of the point in lines 215-218 is to emphasise that large lahars are frequently associated with the passage of large synoptic weather systems which produce large volumes of total rainfall. The increased frequency of rainfall events in the wet season (including such synoptic systems) results in an increase in the average antecedent rainfall, which is identified as contributing to the observed reduction in 1hr PRI based lahar initiation thresholds during the wet season.*

Line 225-227. This is questionable based on data here presented; see previous comment about figure 2.

- *As identified by the reviewer, the term "absence of large lahars in the dry season" should be replaced with "the reduction in the frequency of large lahars in the dry season" as there are a couple of examples of such flows within the studied dataset. However, this reduction is still attributed to a combination of the occurrence of fewer sustained catchment-wide synoptic weather systems and a reduction in average antecedent rainfall and thus saturation level of pyroclastic deposits.*

References:

- Alexander, J., Barclay, J., Susnik, J., Loughlin, S. C., Herd, R. A., Darnell, A., and Crosweller, S.: Sediment-charged flash floods on Montserrat: The influence of synchronous tephra fall and varying extent of vegetation damage, Journal of Volcanology and Geothermal Research, 194, 127-138, 10.1016/j.jvolgeores.2010.05.002, 2010.
- Barclay, J., Alexander, J., and Susnik, J.: Rainfall-induced lahars in the Belham Valley, Montserrat, West Indies, Journal of the Geological Society, 164, 815-827, 10.1144/0016-76492006-078, 2007.
- Capra, L., Borselli, L., Varley, N., Gavilanes-Ruiz, J. C., Norini, G., Sarocchi, D., Caballero, L., and Cortes, A.: Rainfall-triggered lahars at Volcán de Colima, Mexico: Surface hydro-repellency as initiation process, Journal of Volcanology and Geothermal Research, 189, 105-117, 10.1016/j.jvolgeores.2009.10.014, 2010.
- Froude, M. J.: Lahar Dynamics in the Belham River Valley, Montserrat: Application of Remote Camera-Based Monitoring for Improved Sedimentological Interpretation of Post-Event Deposits, PhD Thesis, School of Environmental Science, University of East Anglia, 2015.
- Froude, M.J., Alexander, A., Barclay, J., Cole, P. (2017) Interpreting flash flood paleoflow parameters from antidunes and gravel lenses: An example from Montserrat, West Indies, Sedimentology, DOI:10.1111/sed.12375
- Jones, R., Manville, V., and Andrade, D.: Probabilistic analysis of rain-triggered lahar initiation at Tungurahua volcano, Bulletin of Volcanology, 77, 10.1007/s00445-015-0946-7, 2015.
- Lavigne, F., Thouret, J. C., Voight, B., Young, K., LaHusen, R., Marso, J., Suwa, H., Sumaryono, A., Sayudi, D. S., and Dejean, M.: Instrumental lahar monitoring at Merapi Volcano, Central Java, Indonesia, Journal of Volcanology and Geothermal Research, 100, 457-478, 10.1016/S0377-0273(00)00151-7, 2000.
- Lavigne, F., and Suwa, H.: Contrasts between debris flows, hyperconcentrated flows and stream flows at a channel of Mount Semeru, East Java, Indonesia, Geomorphology, 61, 41-58, 10.1016/j.geomorph.2003.11.005, 2004.
- Pierson, T. C., and Major, J. J.: Hydrogeomorphic effects of explosive volcanic eruptions on drainage basins, Annual Review of Earth and Planetary Sciences, 42, 469-507, 10.1146/annurev-earth-060313-054913, 2014.

Many thanks to the reviewer for the comments and constructive suggestions relating to the underlying review of manuscript number *nhess-2017-166*. Please find below the authors' replies (in italics) to each of these comments:

T. Pierson (Referee)

tpierson@usgs.gov

Received and published: 4 July 2017

Attempts to use rainfall intensity/duration thresholds to effectively predict debris-flow occurrence in non-volcanic terrains and lahar occurrence in volcanic landscapes have been ongoing for decades. This paper, utilizing a rich data set from Montserrat and innovative statistical treatments of the data, makes an important contribution to the discussion. The paper is clearly and concisely written and the figures are quite good. Overall, I would like to see a bit more clarification of the methods used, more explanation (in plain English) of what the statistical treatments are attempting to show, and a broader discussion of the significance of the results in the context of other research. What makes this paper an important contribution is the authors' consideration of (1) catchment stability (measured as total cumulative rainfall since the last significant eruptive activity); and (2) the "false positives problem", i.e., when the occurrence of rainfall intensities above a threshold can, in some cases, trigger lahars but which in other cases do not. While the conclusions reached on both of these topics are a valuable contribution, more discussion of the significance of these findings in the context of previous studies would be extremely helpful.

There are several places in the paper where more attention is needed to clarify the research itself and its significance:

1) It would be helpful if there were a Methods section that summarized all of the approaches and assumptions used in the study. Explanations of these are currently scattered throughout the paper.

- *The authors agree that a restructure of the manuscript to include a consolidated methods section would be beneficial to the manuscript.*

2) The sentence in lines 52–56 is overly complex and confusing. In fact, a word seems to be missing.

- *Amendments to this sentence are required and would help to clarify this section. E.g.* "*Despite this geographic coincidence and the importance of climatic rainfall regimes on storm intensities, durations and antecedent conditions (all significant factors in lahar initiation: Pierson and Major (2014)), the impact of seasonal rainfall on rain-triggered lahar initiation has not previously been explicitly considered within the development of rain-triggered lahar hazard assessment tools.*"

3) In line 64 it would be good to say a bit more about what is meant by "temporal catchment development."

- *Absolutely, this is a key theme later in the manuscript and it would be beneficial to further develop the introduction to this topic at this point in the manuscript. Studies including but not limited to Major et al. (2000), Major & Yamakoshi (2005), Gran & Montgomery (2005) and Pierson & Major (2014) extensively cover this topic and could be used to provide key references when developing this concept within the manuscript.*

4) In lines 81 and 84 there is inconsistent capitalization of "Vulcanian."

- *This inconsistency will be rectified.*

5) At the beginning of section 4, please explain why data sets from different rain gauges are used for different time intervals. Different catch efficiencies can bias results between gauges, and local convective rainstorms can deliver different RF amounts to different gauges.

- *The different rain gauges were used for different time periods out of necessity, and it would indeed be advantageous to have both enhanced continuity of rain gauge location and increased spatial distribution of rainfall gauges across the catchment. As highlighted by the reviewer, the spatial variability in recorded rainfall from local convective rainstorms is certainly a consideration in the Belham Valley. However, the methods presented in this manuscript using the different rain gauges are shown to effectively forecast lahars, and this effectiveness could potentially be further enhanced at locations where networks of permanent gauges are present. Equipment failure is a common issue in monitoring volcanic environments and it of potential benefit that the method here is robust against this.*

6) More explanation is needed for how the peak rainfall intensity (PRI) of 1 hour was chosen for the analyses, and some discussion of PRIs used by other researchers is warranted.

- *One hour peak rainfall intensity was the highest temporal resolution available and as such was the selected resolution. Other studies have shown one-hour peak rainfall intensity to be an effective parameter in lahar initiation threshold assessment (e.g. Jones et al. 2015), although if higher temporal resolutions were available these would have the potential to enhance the performance of lahar forecasting tools, particularly with respect to more accurately capturing the intensities of local convective rainfall events. Previous studies have shown 10-minute rainfall (Arguden & Rodolfo, 1990; Tungol & Regalado, 1996; Lavigne et al. 2000; Lavigne & Suwa, 2004; Okano et al. 2012, Jones et al. 2015), 30-minute rainfall (Lavigne et al. 2000; Tungol & Regalado, 1996; Jones et al. 2015) and 1 hour rainfall (Lavigne et al. 2000; Lavigne & Suwa, 2004; Jones et al. 2015) to be useful parameters in the assessment of lahar hazard.*

7) What are the time lags between the PRIs and lahar initiations?

- *The authors agree that highlighting the lag time between recorded rainfall and lahar detection is important in portraying the potential benefits of the methods discussed in this manuscript. Examples of lag times will be displayed in a new figure displaying the timelines of individual lahar events and recorded rainfall data.*

8) Decline in lahar frequency and magnitude following catchment disturbance is a commonly reported phenomenon. Discussion is needed on how the results of this study specifically compare to the results of other studies.

- *A decline in lahar frequency following catchment disturbance is indeed a commonly reported phenomenon, although direct comparison of the results of this study to previous research is difficult due to the contrasting methods used. However, general comparisons of the conclusions of studies including Van Westen & Daag (2005), which identify increasing lahar initiation thresholds with time, would be beneficial to the manuscript.*

9) Sentence in lines 187–189 is unclear. Is there a word missing?

- *The authors agree that this sentence could be amended to improve its clarity. E.g. "ROC analysis plots the true positive rate against the false positive rate as a threshold (estimated lahar probability in this instance) is varied in order to assess how effectively the parameter discriminates between lahar and non-lahar producing rainfall events."*

10) In lines 193–194, the AUC produced by Eq. 2 is given for the analysis of all RF events. What is it for Eq. 3?

- *The AUC produced by Eq. 3 is 0.89 when all rainfall events are analysed, indicating that the AUC increases by a similar magnitude to that of Eq. 2 when all rainfall events (regardless of magnitude) are considered. This detail can be added to the manuscript.*

11) Discussion is needed for why the antecedent moisture index of 3-day previous rainfall was chosen. What indices have been used by other researchers?

- *A key point also raised by another reviewer, the discussion of the use of antecedent rainfall by other researchers will be expanded and specific mention will be given as to why 3-day rainfall was selected alongside other timescales for testing as an antecedent moisture index. When tested within this study, 3-day antecedent was the optimal timescale, as also utilised by Capra et al. (2010) at Colima, where the lower rainfall and higher evaporation rates made this shorter timescale more relevant than the 7-day timescale used in previous studies in Indonesia (Lavigne et al. 2000; Lavigne & Suwa 2004). As well as being heavily influenced by local climate (Capra et al. 2010), the optimal antecedent rainfall timescale is also influenced by the grain size of pyroclastic material in lahar source regions (Rodolfo & Arguden, 1991). 24-hour (Okano et al. 2012; Jones et al. 2015), 3-day (Capra et al. 2010; Jones et al. 2015) and 7-day (Lavigne et al. 2000; Lavigne & Suwa, 2004) antecedent rainfall have been used in previous research as a lahar initiation threshold assessment parameter.*

12) In lines 225–226, it would seem that the longer durations of the synoptic rainstorms are critical for providing the antecedent moisture during the wet season. It would be good to emphasize that here for the main reason that lahars are harder to trigger in the dry season.

- *An excellent point and a topic that needs to be further emphasised in the manuscript. The total volume of rainfall applied during the wet season during synoptic events is key to decreasing lahar initiation thresholds.*

13) In line 227, a reference for inefficient bulking in dry channels is in order.

- The authors agree, references to this process will be added to the manuscript, including Fagents & Baloga (2006), Doyle et. al (2011) and others.

14) Toward the end of the discussion section, a better explanation of the meaning and significance of the ROC analysis is needed. From what you have written, I assume (not being familiar with this analysis) that (1) AUC = 0.5 means the number of true positives equals the number of false positives, and that (2) AUC = 1.0 means the number of true positives is 100%. Is this the case?

- *This understanding of ROC analysis is correct, however further explanation of ROC analysis would be beneficial to the manuscript and could be implemented within the proposed updated methods section.*

15) How far above the PRI thresholds are the false-positive rainfall intensities? For example, if you set a PRI threshold of 25 mm/hr, how large a PRI can occur that does not trigger a lahar?

- *Taking the reviewer's example, if a strict threshold of 25 mm/hr was selected there would be 18 rainfall events in the study period above this threshold that would be expected to trigger lahars. Of these 18 rainfall events, there would be three false positives, with peak rainfall intensities of 26, 28 and 34 mm/hr respectively. All rainfall events exceeding 34 mm/hr that were analysed in this study triggered lahars. Consideration of this topic could be added to the manuscript as a discussion point.*

16) Figure 2 caption: Please explain the vertical dashed lines.

- *These dashed lines are periods where equipment failure occurred and resulted in a gap in the record. Further detail will be added to the caption to make this clearer.*

References:

- Alexander, J., Barclay, J., Susnik, J., Loughlin, S. C., Herd, R. A., Darnell, A., and Crosweller, S.: Sediment-charged flash floods on Montserrat: The influence of synchronous tephra fall and varying extent of vegetation damage, Journal of Volcanology and Geothermal Research, 194, 127-138, 10.1016/j.jvolgeores.2010.05.002, 2010.
- Arguden, A., and Rodolfo, K.: Sedimentologic and dynamic differences between hot and cold laharic debris flows of Mayon Volcano, Philippines, Geological Society of America Bulletin, 102, 865-876, 10.1130/0016-7606(1990)102<0865:saddbh>2.3.co;2, 1990.
- Barclay, J., Alexander, J., and Susnik, J.: Rainfall-induced lahars in the Belham Valley, Montserrat, West Indies, Journal of the Geological Society, 164, 815-827, 10.1144/0016-76492006-078, 2007.
- Capra, L., Borselli, L., Varley, N., Gavilanes-Ruiz, J. C., Norini, G., Sarocchi, D., Caballero, L., and Cortes, A.: Rainfall-triggered lahars at Volcán de Colima, Mexico: Surface hydro-repellency as initiation process, Journal of Volcanology and Geothermal Research, 189, 105-117, 10.1016/j.jvolgeores.2009.10.014, 2010.
- Doyle, E. E., Cronin, S. J., and Thouret, J. C.: Defining conditions for bulking and debulking in lahars, Geological Society of America Bulletin, 123, 1234-1246, 10.1130/B30227.1, 2011.
- Fagents, S. A., and Baloga, S. M.: Toward a model for the bulking and debulking of lahars, J Geophys Res-Sol Ea, 111, 10.1029/2005jb003986, 2006.
- Froude, M. J.: Lahar Dynamics in the Belham River Valley, Montserrat: Application of Remote Camera-Based Monitoring for Improved Sedimentological Interpretation of Post-Event Deposits, PhD Thesis, School of Environmental Science, University of East Anglia, 2015.
- Gran, K. B., and Montgomery, D. R.: Spatial and temporal patterns in fluvial recovery following volcanic eruptions: Channel response to basin-wide sediment loading at Mount Pinatubo, Philippines, Geological Society of America Bulletin, 117, 195-211, 10.1130/B25528.1, 2005.
- Jones, R., Manville, V., and Andrade, D.: Probabilistic analysis of rain-triggered lahar initiation at Tungurahua volcano, Bulletin of Volcanology, 77, 10.1007/s00445-015-0946-7, 2015.
- Lavigne, F., Thouret, J. C., Voight, B., Young, K., LaHusen, R., Marso, J., Suwa, H., Sumaryono, A., Sayudi, D. S., and Dejean, M.: Instrumental lahar monitoring at Merapi Volcano, Central Java, Indonesia, Journal of Volcanology and Geothermal Research, 100, 457-478, 10.1016/S0377-0273(00)00151-7, 2000.

- Lavigne, F., and Suwa, H.: Contrasts between debris flows, hyperconcentrated flows and stream flows at a channel of Mount Semeru, East Java, Indonesia, Geomorphology, 61, 41-58, 10.1016/j.geomorph.2003.11.005, 2004.
- Major, J. J., Pierson, T. C., Dinehart, R. L., and Costa, J. E.: Sediment yield following severe volcanic disturbance - A two-decade perspective from Mount St. Helens, Geology, 28, 819-822, 10.1130/0091-7613(2000)28<819:Syfsvd>2.0.Co;2, 2000.
- Major, J. J., and Yamakoshi, T.: Decadal-scale change of infiltration characteristics of a tephra-mantled hillslope at Mount St Helens, Washington, Hydrological Processes, 19, 3621-3630, 10.1002/Hyp.5863, 2005.
- Okano, K., Suwa, H., and Kanno, T.: Characterization of debris flows by rainstorm condition at a torrent on the Mount Yakedake volcano, Japan, Geomorphology, 136, 88-94, 10.1016/j.geomorph.2011.04.006, 2012.
- Pierson, T. C., and Major, J. J.: Hydrogeomorphic effects of explosive volcanic eruptions on drainage basins, Annual Review of Earth and Planetary Sciences, 42, 469-507, 10.1146/annurev-earth-060313-054913, 2014.
- Rodolfo, K., and Arguden, A.: Rain-lahar generation and sediment-delivery systems at Mayon Volcano, Philippines, in: Sedimentation in Volcanic Settings, edited by: Fisher, R., and Smith, G., SEPM, Special Publication 45, 71-87, 1991.
- Tungol, N., and Regalado, T.: Rainfall, acoustic flow monitor records, and observed lahars of the Sacobia River in 1992, in: Fire and Mud, Eruptions and Lahars of Mt Pinatubo, Philippines, edited by: Newhall, C., and Punongbayan, R., PHIVOLCS/University of Washington Press, Quezon City/Seattle, 1023-1033, 1996.
- Van Westen, C., and Daag, A.: Analysing the relation between rainfall characteristics and lahar activity at Mt Pinatubo, Philippines, Earth Surface Processes and Landforms, 30, 1663-1674, 2005.

Summary of Manuscript Edits

Please find below a summary of the changes made to the manuscript by the authors in response to the above reviews. A fully marked-up version of the manuscript can be found below this summary (references to line numbers refer to the marked-up manuscript below)

- Lines 53-57: This sentence has been reconstructed to address Pierson comment #2.
- Lines 65-67: Revision for clarity and to emphasise the point raised by Pierson in comment #3, which is discussed in the preceding lines.
- Line 82: Pierson comment #4 regarding the inconsistent capitalisation of the term 'Vulcanian' has been addressed.
- Lines 93-95: A brief description of the lahars has been added as requested by Capra in comment #2.
- Lines 103-107: Addresses Capra comment #6 by adding information regarding dominant vegetation types present in the location.
- Lines 115-154: This section has been redesigned to include a more consolidated methods section as advised by Pierson comment #1.
- Lines 119-122: Information added regarding why different rain gauges are used during the study period, an issue raised by Pierson comment #5.
- Lines 123-130: A description of the observation/detection methods used to identify lahars has been added in addition to information regarding how the lahars are categorised by magnitude. This is in response to Capra comment #1.
- Lines 131-143: Detail has been added regarding the use and definition of 1 hour peak rainfall intensity as discussed in Capra comments #4 and #5. This temporal resolution of rainfall data was the highest available in this case.
- Lines 133-137: A new figure has been created demonstrating timelines of rainfall data and lahar occurrence in response to Capra comment #3 and Pierson comment #17.
- Lines 144-154: Information regarding the methods used in the study (specifically analysis methods) has been transferred to this new consolidated methods section (Pierson comment #1). Some of this material has been moved to this section from later in the manuscript.
- Lines 173-175: A demonstration of the % of false positives present above an example threshold and details regarding the maximum non-lahar triggering rainfall intensity has been added~~.~~ (Pierson comment #15).
- Lines 207-208: Clarity regarding what is meant by the term "cumulative rainfall since significant eruptive activity" has been added to address Capra comment #10.
- Lines 210-214: Information regarding the antecedent rainfall timescales used in other studies and the reasons for the different timescales has been added to address Capra comment #9 and Pierson comment #11.
- Line 242: Results of ROC analysis added as requested in Pierson comment #10.
- Lines 252-254: Pierson comment #8 has been addressed by adding information regarding the difficultly in making direct comparisons to the results of previous studies. Lines 263-264 also address this point by referencing a previous study which highlights an increase in lahar initiation thresholds with time.
- Line 277: Adjustment made to the phrasing as identified by Capra comment #12.
- Lines 279-280: References added to support point as suggested in Pierson comment #13.
- Lines 296-302: Information regarding the rainfall timescales used in previous studies of lahar initiation thresholds has been added (Pierson comment #6, Capra comment #4).
- Line 335: The caption has been amended to add clarity to the figure as identified in Pierson comment #16
- Lines 337-339: New caption for new figure 3.
- Throughout the references section additional references have been added where appropriate.

- Line 474: New Figure 3 (Capra Comment #3, Pierson Comment #17).
- Line 485 Onwards: A new appendix has been created (including two tables and a figure) to describe ROC analysis more fully as identified by Pierson comment #14.

Real-time prediction of rain-triggered lahars: incorporating seasonality and catchment recovery

Formatted: Numbering: Restart each section

Robbie Jones[a*], Vern Manville[a], Jeff Peakall[a], Melanie Froude[bc], Henry Odbert[de]

[a]School of Earth and Environment, University of Leeds, Leeds. LS2 9JT, United Kingdom
[b]School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ, United Kingdom
[c]Department of Geography, University of Sheffield, 9 Northumberland Road, Sheffield, S10, UK
[d]School of Earth Sciences, University of Bristol, Wills Memorial Building, Queens Road, Bristol BS8 1RJ, United Kingdom
[e]Met Office, FitzRoy Road, Exeter, Devon, EX1 3PB, United Kingdom

[*]*Correspondence to:* Robbie Jones (~~eerj@leeds.ac.uk~~robbie_j_jones@outlook.com)

Abstract. Rain-triggered lahars are a significant secondary hydrological and geomorphic hazard at volcanoes where unconsolidated pyroclastic material produced by explosive eruptions is exposed to intense rainfall, often occurring for years to decades after the initial eruptive activity. Previous studies have shown that secondary lahar initiation is a function of rainfall parameters, source material characteristics and time since eruptive activity. In this study, probabilistic rain-triggered lahar forecasting models are developed using the lahar occurrence and rainfall record of the Belham River Valley at Soufrière Hills Volcano, Montserrat collected between April 2010 and April 2012. In addition to the use of peak rainfall intensity as a base forecasting parameter, considerations for the effects of rainfall seasonality and catchment evolution upon the initiation of rain-triggered lahars and the predictability of lahar generation are also incorporated into these models. Lahar probability increases with peak one-hour rainfall intensity throughout the two-year dataset, and is higher under given rainfall conditions in year one than year two. The probability of lahars is also enhanced during the wet season, when large-scale synoptic weather systems (including tropical cyclones) are more common and antecedent rainfall and thus levels of deposit saturation are typically increased. The incorporation of antecedent conditions and catchment evolution into logistic regression-based rain-triggered lahar probability estimation models is shown to enhance model performance and displays the potential for successful real-time prediction of lahars, even in areas featuring strongly seasonal climates and temporal catchment recovery.

1 Introduction

Lahars are rapidly flowing mixtures of rock debris and water (other than normal streamflow) from a volcano and represent a significant hazard due to their energetic nature and mobility (Smith and Fritz, 1989). Globally, 17% of historical volcano-related fatalities have occurred due to lahars (Auker et al., 2013); with decadal-scale hazards being created by some large eruptions (Major et al., 2000). Secondary, post-eruption lahars are dominantly the result of rainfall on unconsolidated pyroclastic deposits, which are typically remobilised by rilling due to Hortonian overland flow (Segerstrom, 1950; Waldron, 1967), undercutting and lateral bank collapse and headward erosion (Pierson, 1992); or by shallow landsliding of saturated tephra layers above basal décollement surfaces (Iverson, 2000; Manville et al., 2000).

At present, rain-triggered lahar hazard identification is predominantly based on observations as well as ground-based flow detection systems such as Acoustic Flow Monitors (AFMs) or trip-wires at locations where such resources are available (e.g. Marcial et al., 1996; Lavigne et al., 2000). Previous studies featuring post-lahar analysis of flow observations and rainfall records at a range of volcanoes have displayed a power-law relationship indicating that lahar initiation occurs along a continuum from short duration, high intensity rainfall events to long duration, low-intensity events (e.g. Rodolfo and Arguden, 1991; Capra et al., 2010; Jones et al., 2015). Enhancing the use of local telemetered rainfall gauge networks within lahar hazard monitoring and assessment has the potential to increase the number of available mitigation tools whilst avoiding the lag-time between flow initiation and flow detection inherent in ground-based detection and observation. Globally, such pre-emptive prediction and forecasting of rain-triggered lahars based on telemetered rainfall data is lacking, although initial application of real-time rainfall data for lahar prediction has demonstrated increased lahar warning times compared with ground-based flow detection (Jones et al., 2015).

The initiation of rain-triggered lahars is dependent on the characteristics of rainfall, pyroclastic deposits and topography, indicating that both the climatic regime of lahar-prone regions and the hydrogeomorphic response of drainage basins to eruptive activity are important considerations in rain-triggered lahar research (Pierson and Major, 2014). Regions of high rainfall seasonality are predominantly distributed in the tropics and sub-tropics either side of the equator (Wang et al., 2010); whilst approximately 46% of active volcanoes are identified as being located in the humid tropics (Rodolfo and Arguden, 1991). Despite this geographic coincidence and the importance of climatic rainfall regimes on storm intensities, durations and antecedent conditions (~~,~~ all significant factors in lahar initiation: ~~(~~Pierson and Major, (2014)), the impact of seasonal rainfall on rain-triggered lahar initiation has not previously been explicitly considered within the development of rain-triggered lahar hazard assessment tools.

Commented [A1]: Addresses Pierson comment #2

Following a discrete volcanic eruption, sediment yields in impacted fluvial systems are amongst the highest recorded globally, but decline exponentially (Major et al., 2000), which is consistent with other examples of disturbed earth systems (Graf, 1977). Mechanisms include a reduction in available particulate material, vegetation recovery, fragmentation of runoff-enhancing surface crusts, exposure of more permeable substrates and the stabilisation of rill networks (Leavesley et al., 1989; Schumm and Rea, 1995; Major et al., 2000; Major and Yamakoshi, 2005). Conversely, at locations featuring recurrent or persistent volcanic activity, the magnitude of the lahar hazard remains relatively constant with time due to the regular supply of new material (Thouret et al., 2014). As a result, temporal catchment development is another factor ~~which influences~~influencing lahar frequency and magnitude through time, ~~indicating that it is also an important consideration~~and should also be considered within the development of rain-triggered lahar hazard assessment tools.

Commented [A2]: Revision for clarity, Pierson comment #3 addressed by preceding existing paragraph

This study uses probabilistic and diagnostic methods, including binary logistic regression and Receiver Operating Characteristic (ROC) analysis, to develop real-time rainfall-based lahar forecasting tools which account for the impacts of seasonal rainfall and catchment recovery on lahar occurrence in the Belham Valley, Montserrat. Such hazard assessment tools have the potential to be utilised both as a stand-alone tool where ground-based detection equipment is unavailable, and in conjunction with instrumental monitoring techniques to increase lahar warning times.

2 Soufrière Hills Volcano, Montserrat

Soufrière Hills Volcano (SHV, Montserrat, Lesser Antilles, 16.72°N, 62.18°W) lies on the northern edge of the Inter-Tropical Convergence Zone in the eastern Caribbean and has a strongly seasonal climate. Rainfall-producing weather systems affecting the island fall into two broad categories; large-scale synoptic (>100 km across) systems and local mesoscale (<100 km across) systems (Froude, 2015). Both can produce high intensity precipitation, but large-scale events can potentially be forecast days in advance whereas this timescale reduces to hours for local weather systems (Barclay et al., 2006).

The andesitic dome-forming eruption of SHV began in July 1995 and has featured several phases of activity consisting of dome growth, dome collapse and ~~vulcanian~~ Vulcanian explosions as well as pauses in magma extrusion (Bonadonna et al., 2002; Komorowski et al., 2010; Stinton et al., 2014). Pyroclastic density currents (PDCs) have deposited fine-grained ash- and pumice-rich and coarser-grained blocky deposits around the volcano (Cole et al., 2002; Stinton et al., 2014), supplemented by tephra deposits from short-lived Vulcanian explosions and associated fountain-collapse flows and surges (Komorowski et al., 2010). Prevailing winds often distribute ash from weak plumes to the West, but larger plumes can also deposit to the North, East and South (Bonadonna et al., 2002). This intermittent eruptive activity has triggered a complex sedimentological response in drainages surrounding the volcano since 1995 (Barclay et al., 2006, 2007; Alexander et al., 2010; Froude, 2015).

Commented [A3]: Addresses Pierson comment #4

3 The Belham Catchment

Data from the Belham Valley, Montserrat (Fig. 1) were used to examine the influence of rainfall seasonality and catchment evolution on the occurrence of rain-triggered lahars between April 2010 and April 2012 (Fig. 2). Lahars have persisted in the valley since the onset of eruptive activity in 1995 and detailed observations of lahars in the Belham Valley have indicated that they are dominantly Newtonian and fully turbulent (Barclay et al., 2007; Alexander et al., 2010; Froude et al., 2017). Lahars have damaged infrastructure, including burying the Belham Bridge in 1998, resulting in the river bed being used as the primary transportation link between the "Safe Zone" and the "Daytime Entry Zone" (Barclay et al., 2007; Alexander et al., 2010).

Commented [A4]: Addresses Capra comment #2

The Belham Catchment had a pre-1995 surface area of c. 13.7 km^2, increasing to c. 14.8 km^2 early in the eruptive episode due to capture of a portion of Gage's fan (Froude, 2015). During eruptive episodes tephra fall and pyroclastic density current (PDC) deposits accumulate in the upper catchment. The destruction and burial of vegetation in the Belham Valley reduces the infiltration and interception of precipitation, and in combination with a reduction in surface roughness enhances run-off and erosion rates and promotes rain-triggered lahar generation (Barclay et al., 2007; Alexander et al., 2010; Froude, 2015). Prior to the onset of eruptive activity, 62% of the Belham Catchment was densely vegetated with Dry Forest (29%), Mesic Forest (48%) and Wet Forest (13%), with dry forest subsequently identified as the dominant species found on re-vegetating pyroclastic deposits (Froude, 2015). Previous studies in the Belham Valley have not identified evidence of hydrophobicity, such as previously identified at Colima by Capra et al. (2010). Aggradation and sedimentation in the upper catchment during periods of eruptive activity are counter-balanced during periods of quiescence by channel development and stabilisation, exposure of more permeable substrates, vegetation recovery and a reduction in available sediment (Froude, 2015). The data period used here coincides with a lack of substantial eruptive activity at SHV following the 11th of February 2010 dome collapse at the end of "Phase 5", which deposited stacked lobes of pumiceous PDC deposits up to 5.7 km from source in the Belham Valley (Stinton et al., 2014). This period of eruptive quiescence indicates that this study focuses on a time of channel development and stabilisation within the upper catchment of the Belham Valley.

Commented [A5]: Addresses Capra comment #6

4 ~~Rainfall and Lahar Record~~Methods

Commented [A6]: Movement towards a more consolidated Methods section as suggested by Pierson comment #1

The record used in this study (Fig. 2) comprises 0.1 mm resolution hourly precipitation data recorded at the MVO Helipad Gauge between February 2010 and February 2011, the St George's Hill gauge between March 2011 and May 2011, and the maximum of the St George's Hill and Windy Hill gauges (Fig. 1) between May 2011 and February 2012. While a continuous record from rain gauges with a better spatial distribution and density would be ideal to minimise differences in catch efficiencies and to capture local variations in convective and orographic rainfall, operating a fully functioning rain gauge network is technically challenging and generally a low priority during a volcanic crisis. The lahar database (Fig. 2) is compiled from inspection of seismic records and visual observations and lahars are categorised based on magnitude~~. Lahar size~~ (small, medium, large)~~ is estimated based on recorded seismic amplitude and occupied valley width alongside flow start time, end time and duration~~. These categories were assessed using visual inspection of the degree of channel inundation and flow depth (where possible); in addition to the assessment of the duration and amplitude of seismic signals. Seismic signals of lahars show continuous readings in the 2-5 Hz and peak at approximately 30 Hz. The highest recorded amplitudes are associated with the greatest discharges and sediment loads in observed lahars. Lahar signals were cross referenced to visual observations and carefully excluded from signals associated with primary volcanic activity and other seismic noise (such as construction vehicles).

Commented [A7]: Addresses Pierson comment #5

Commented [A8]: Addresses Capra comment #1

Within this study a designated minimum inter-event dry period of six hours is utilised, meaning that in common with several previous soil erosion studies a dry interval of six hours is needed to define the end of a single rainfall event (Wischmeier and Smith 1978; Todisco, 2014). Figure 3 ~~displays~~shows six examples of rainfall events (or series of consecutive rainfall events) which resulted in the observation or detection of lahars in the Belham River Valley. ~~Evident in Figure 3 is~~clearly displaying the lag time between the recording of rainfall (cumulative- and real-time progression of One Hour Peak Rainfall Intensity: 1hr PRI) and the observation/detection of lahars. ~~Alongside cumulative recorded rainfall, the real-time progression of the One Hour Peak Rainfall Intensity (1hrPRI: the highest temporal resolution available) of the rainfall event is displayed in Figure 3.~~ 1hrPRI has been identified as an effective parameter in lahar initiation threshold assessment during previous analysis (Jones et al., 2015). Division of the dataset into six-month moving windows, with staggered one-month start dates, facilitates the illustration of the seasonal variation in both the number of rainfall events exceeding 1hrPRI~~One Hour Peak Rainfall Intensity (1hrPRI: the highest temporal resolution available)~~ thresholds and the occurrence (and estimated magnitude) of lahars (~~Fig. 3~~Fig. 4).

Commented [A9]: Addresses Capra comment #4

Commented [A10]: Addresses Capra comment #4

This study uses binary logistic regression to develop lahar probability estimation models based on the 1hrPRI of a rainfall event, whilst also examining the impacts of incorporating considerations for seasonal and temporal effects within these models. Binary logistic regression is a statistical method ~~which~~that estimates the probability of a dichotomous outcome (the occurrence or non-occurrence of lahars in this case) using one or more independent variables (Hosmer Jr et al., 2013). Model performance is assessed using both the model chi-square test and Receiver Operating Characteristic (ROC) analysis (Fawcett, 2006). ROC analysis (Appendix 1) plots the true positive rate against the false positive rate as a threshold (estimated lahar probability in this instance) is varied in order to assess how effectively the parameter discriminates between lahar and non-lahar producing rainfall events. The area under the ROC curve (AUC) is a measure of the ability of a tool to distinguish between the two outcomes, and varies between 0.5 (no predictive ability, i.e. number of true positives equals number of false positives, or no better than guessing) and 1.0 (perfect predictive ability, i.e. 100% true positives and no false positives).

Commented [A11]: Addresses Pierson comment #14

5 Results

The six-month window between April and October is identified as the peak wet season in this study, with 1721 mm of recorded rainfall in the 2010 peak wet season (WS1) and 1455 mm in the 2011 peak wet season (WS2). The 2010/11 peak dry season (DS1) featured approximately 750 mm of rainfall, whilst 1076 mm of rainfall was recorded in the 2011/12 peak dry season (DS2). Mean WS1 and WS2 1hrPRIs are 5.2 mm hr^{-1} and 5.0 mm hr^{-1} respectively, whilst mean dry season 1hrPRIs are 2.2 mm hr^{-1} (DS1) and 3.3 mm hr^{-1} (DS2).

There is significant ($p < 0.01$) correlation between recorded rainfall on timescales of 1-168 hours and lahar occurrence. When lahars are categorised by estimated magnitude, large lahars are strongly correlated with longer-duration (>24 hours) rainfall events, produced by the passage of synoptic weather systems. Between April 2010 and April 2012 large flows were directly attributed to several named tropical cyclones (Fig. 2). In contrast, smaller lahars display increased correlation with the passage of short-duration (<24 hours) rainfall events, more commonly associated with mesoscale weather systems.

5.1 Probabilistic rain-triggered lahar analysis

The correlation between recorded peak rainfall intensity and the subsequent occurrence of lahars ~~occurrence~~ (Fig. 3) provides the platform for probabilistic analysis of lahar occurrence based on the 1hrPRI of a rainfall event. ~~Within this study a designated minimum inter-event dry period of six hours is utilised, meaning that in common with several previous soil erosion studies a dry interval of six hours is needed to define the end of a single rainfall event (Wischmeier and Smith 1978; Todisco, 2014).~~ Results show that lahar probability increases with greater 1hrPRI throughout the two-year study period. For example, of the 18 rainfall events which exceeded a 1hrPRI of 25 mm hr^{-1}, 15 were associated with the triggering of lahars, and all the rainfall events exceeding a 1hrPRI of 34 mm hr^{-1} triggered lahars. Additionally, higher ~~in both years of the Belham Valley dataset, with higher l~~lahar probabilities are observed in year 1 than year 2 for a specified 1hrPRI (~~Fig. 4~~Fig. 5)~~.~~, and e~~E~~mpirically-derived lahar probabilities for rainfall events featuring a given minimum 1hrPRI also fluctuate seasonally during the study period (~~Fig. 5~~Fig. 6). These 1hrPRI exceedance-based lahar probabilities (~~Fig. 5~~Fig. 6) are initially stable during the 6-month windows focused on WS1 before decreasing during DS1, increasing during WS2 and once again decreasing into DS2. This indicates that more intense rainfall is required to trigger lahars in the dry season than in the wet season. Throughout the two-year study period increased 1hrPRI correlates with increased lahar probability, displaying its effectiveness as a potential first-order lahar forecasting parameter. In addition to seasonal fluctuations in relative lahar probability, there is an overall decline in relative lahar probabilities across the two-year study period (Figs. 5~~4~~ & 6~~5~~). The relationship between 1hrPRI and lahar occurrence as well as the combination of seasonal fluctuation and temporal decline in lahar probability displayed in ~~Figure 5~~Figure 6 are~~is~~ examined further using binary logistic regression.~~, a statistical method which estimates the probability of a dichotomous outcome using one or more independent variables (Hosmer Jr et al., 2013)~~. In this instance the occurrence or non-occurrence of lahars (of any magnitude) is used as the dichotomous dependent variable and initially the 1hrPRI of a rainfall event is the singular independent variable. ~~Figure 6~~Figure ~~7A displays~~7 displays logistic regression-based lahar probability estimation models generated by this single-variable approach using four sub-datasets; *Year 1*, *Year 2*, *Wet Seasons* and *Dry Seasons*. Within each of these four models the model chi-square test indicated statistically significant lahar prediction ability ($p < 0.01$). ~~Figure 6~~Figure ~~7A~~

Commented [A12]: Addresses Pierson comment #15

Formatted: Not Superscript/ Subscript

~~displays~~7 displays higher estimated lahar probabilities at identical 1hrPRI values for Year 1 relative to Year 2 and Wet Seasons relative to Dry Seasons.

The potential benefit of incorporating considerations for seasonal and temporal effects within lahar forecasting models was investigated using further binary logistic regression. This approach selected alternate chronological rainfall events (minimum total rainfall ≥8 mm) from the two-year dataset, creating a model formulation dataset consisting of 74 rainfall events, of which 25 produced lahars. Lahar forecasting models were created from this model formulation dataset using binary logistic regression, and the remaining 73 rainfall events, of which 20 produced lahars, were retained for the assessment of the performance of the lahar forecasting models. Proxies for seasonal effects (antecedent rainfall on timescales of 1-90 days) and catchment recovery (long-term cumulative rainfall and days since significant eruptive activity) were tested in combination with 1hrPRI. The minimum event rainfall threshold of 8 mm (under which only two lahars occurred during the two-year dataset) was implemented for logistic regression and subsequent forecasting assessment in order to increase the balance between lahar and non-lahar outcomes and thus reduce skewed predicted probability.

Three-day antecedent rainfall displayed the biggest influence of the tested antecedent rainfall timescales upon the effectiveness of lahar forecasts, while total cumulative rainfall since significant eruptive activity (i.e. the end of Phase 5) best captured temporal catchment development effects. Therefore, the optimal lahar forecasting model developed from the model formulation dataset utilises 3-day antecedent rainfall and long-term cumulative rainfall alongside the first-order lahar forecasting parameter of 1hrPRI. A 3-day antecedent period was also used by Capra et al. (2010) at Colima, whereas a 7-day period was used in Indonesia (Lavigne et al., 2000; Lavigne and Suwa, 2004) where rainfall is higher and evaporation rates lower, and a 24-hour period was used at Mount Yakedake (Okano et al., 2012). The optimal antecedent rainfall timescale is a function of local climate (Capra et al., 2010) and the grain-size distribution of the pyroclastic deposits (Rodolfo and Arguden, 1991).

The reverse stepwise logistic regression method (Hosmer Jr et al., 2013), which involves the deletion of variables whose removal from the model results in a statistically insignificant deterioration of model performance, retained these three independent variables (1hrPRI, 3-day antecedent rainfall and total cumulative rainfall since significant eruptive activity). This model composition increased correct classification of rainfall event outcomes in the model formulation dataset from a null model value of 66% (when all events in the database are predicted to not trigger lahars) to 80% when using our explanatory variables, with model chi-square tests again indicating significant prediction ability ($p<0.01$). Model variables (X_i) and output regression coefficients (β_i) are used to construct lahar probability estimation equations by conversion of the logistic regression logit model (Eq. 1) in terms of probability.

$$logit(p) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \quad (1)$$

Eq. 2 displays the application of this to the multi-variable model, featuring the probability of lahar occurrence *(p)*, 1hrPRI (R_i), three-day antecedent rainfall (A_3) and cumulative rainfall since significant eruptive activity (C).

$$p = \frac{1}{1 + e^{-(-2.10 + 0.133R_i + 0.018A_3 - 0.215C)}} \quad (2)$$

Eq.3 displays the lahar probability estimation model produced by the same dataset using only 1hrPRI as an independent variable.

$$p = \frac{1}{1 + e^{-(-2.33 + 0.133R_i)}} \quad (3)$$

Commented [A13]: This sentence already addresses Pierson comment #11

Commented [A14]: Addresses Capra comment #10

Commented [A15]: Addresses Pierson comment #11, Capra comment #9

Application of Eqs. 2 & 3 to the 73 rainfall events in the forecasting assessment dataset produced two sets of model-derived lahar probability estimates. The lahar forecasting performance of the two models was then assessed relative to the actual outcomes (lahar or no lahar) of the rainfall events using ~~Receiver Operating Characteristic (~~ROC~~)~~ analysis. ~~(Fawcett, 2006). ROC analysis plots the true positive rate against the false positive rate as a threshold (estimated lahar probability) is varied in order to assess its ability to discriminate between lahar and non-lahar producing rainfall events. The area under the ROC curve (AUC) is a measure of the ability of a tool to distinguish between the two outcomes, and varies between 0.5 (no predictive ability, i.e. number of true positives equals number of false positives, or no better than guessing) and 1.0 (perfect predictive ability, i.e. 100% true positives and no false positives).~~ The multiple-variable lahar probability estimation model shown in Eq. 2 produced an AUC of 0.83 ($p<0.01$), whilst the single variable model shown in Eq. 3 produced an AUC of 0.79 ($p<0.01$) (~~Fig. 6~~Fig. 7B). The AUC produced by Eq. 2 increases to 0.93 if the 8 mm event threshold is removed and the multi-variable model is applied to all 508 rainfall events that were not used in model formulation (AUC given by Eq. 3 increases to 0.89 for equivalent parameters).

Commented [A16]: Addresses Pierson comment #10

6 Discussion

Analysis of the Belham Valley lahar occurrence and rainfall record over a two-year period indicates that lahar probability and magnitude is a function of: (i) temporal catchment evolution towards more stable conditions – lahars are harder to trigger with time; and (ii) seasonal variations in rainfall – lahars are more common in the wet season both in terms of frequency and probability relative to 1hrPRI.

The multi-year temporal trend is attributed to a declining supply of easily erodible pyroclastic material in the upper catchment, coupled with stabilisation of channel networks, vegetation re-growth, and increased infiltration as identified in several previous studies of lahar-prone regions following eruptive activity (e.g. Leavesley et al., 1989; Schumm and Rea, 1995; Major et al., 2000; Major and Yamakoshi, 2005). However, direct comparisons with other lahar-prone settings is not possible as differences in methodologies mean that common metrics such as sediment yield were not determined. The occurrence of several large rainfall events following Phase 5 of the eruption (Fig. 2) triggered a number of high-magnitude lahars within the Belham Valley, enhancing temporal channel development within the catchment and resulting in the widespread erosion and downstream transportation of pyroclastic material (Froude, 2015). Rapid re-vegetation during periods of eruptive quiescence has also been identified in the catchment (Froude, 2015), a process which increases infiltration, interception, evapotranspiration and surface roughness; reducing post-eruption runoff rates (Yamakoshi and Suwa, 2000; Ogawa et al., 2007; Alexander et al., 2010). Temporal increase in infiltration rates in the Belham Valley is also attributed to the exposure of more permeable substrates following the erosion of fine-grained surface tephra layers (Froude, 2015), a factor identified previously in studies of the landscape response to the 1980 eruption of Mt St Helens (Collins and Dunne, 1986; Leavesley et al., 1989). Collectively these processes would result in increasing lahar initiation thresholds with time (Van Westen and Daag, 2005).

Commented [A17]: Response to Pierson comment #8

Commented [A18]: Response to Pierson comment #8

Probabilistic analysis shows that throughout the two-year dataset utilised in this study, increased 1hrPRI results in increased lahar occurrence probability. Additionally, an increase in the absolute numbers of lahars and a reduction in rain-triggered lahar initiation thresholds are identified in the wet seasons. Seasonality in the nature and frequency of rainfall-generating weather systems controls this pattern. Large lahars are often associated with the passage of synoptic weather systems, which typically produce long-duration catchment-wide rainfall. This is demonstrated by the triggering of large lahars by several named storms during the study dataset including Hurricane Earl in August 2010, Tropical Storm Otto in October 2010 and Tropical Storm Maria in September 2011. Increased rainfall in the wet season also influences ~~the dominant~~ antecedent conditions within the catchment, resulting in reduced infiltration rates due to deposit saturation (Barclay et al., 2007). Increased antecedent rainfall can also produce runoff-enhancing surface seals (Segerstrom, 1950; Fohrer et al., 1999) and result in increased bulking efficiency during lahar transit due to high water contents in channel floor deposits (Iverson et al., 2011). These effects increase the overall probability of lahars in the wet season under given rainfall conditions due to flash-flood type responses to rainfall. The ~~absence~~ reduced frequency of large lahars in the dry season is attributed to the occurrence of fewer sustained catchment-wide synoptic weather systems as well as antecedent effects (low antecedent rainfall inhibits bulking efficiency in the dry season (Fagents and Baloga, 2006; Doyle et al., 2011; Iverson et al., 2011)). The development of lahar magnitude assessment methods, from the subjective classification used in this study, towards quantitative initial flow volume estimates has the potential to enhance probabilistic lahar forecasting by creating probabilistic hazard footprints (Mead et al., 2016). However, such quantitative assessment methods are highly data intensive relative to those developed ~~with~~in this study, requiring pre- and post-eruption digital elevation models, location specific rainfall intensity-frequency-duration thresholds and physical deposit characteristics as input data (Mead et al., 2016). These input data requirements prohibit practical implementation of fully-quantitative magnitude estimates within probabilistic rain-triggered lahar assessment at all but the most thoroughly monitored volcanoes.

The incorporation of considerations for temporal catchment development and seasonality of prevalent antecedent conditions into logistic regression-based lahar probability estimation models increases rain-triggered lahar forecasting performance. The addition of these considerations modulates purely 1hrPRI-based probability estimates to account for initial deposit moisture content and the degree of catchment recovery during a period of eruptive quiescence. ROC analysis indicates an excellent ability to differentiate between lahar and non-lahar outcomes (AUC = 0.83) when only larger rainfall events resulting in ≥8 mm of total rainfall are considered, and this ability improves even further (AUC = 0.93) when the 8 mm threshold is removed. The readily available model inputs of 1hrPRI, three-day antecedent rainfall and cumulative rainfall since significant eruptive activity can be easily assimilated into functional real-time lahar probability estimation models and produces real benefits. Rainfall gauge networks in volcanic areas are seldom designed with the intention of optimising their usefulness for detection and characterisation of rain-triggered lahar initiation: the 1hrPRI used in this study is based on the minimum temporal resolution of the data recorded. Previous studies have shown the utility of 10-minute (Arguden and Rodolfo, 1990; Tungol and Regalado, 1996; Lavigne et al., 2000; Lavigne and Suwa, 2004; Okano et al., 2012; Jones et al., 2015), 30-minute (Tungol and Regalado, 1996; Lavigne et al., 2000; Jones et al., 2015) and 60 minute (Lavigne et al., 2000; Lavigne and Suwa, 2004; Jones et al., 2015) rainfall data. Lahar forecasting using real-time telemetered rainfall data and these techniques has the potential to effectively predict secondary lahars and increase lahar warning times, even in areas where AFMs, proximal seismometers and trip wires are unavailable. Used in conjunction with ground-based detectors in instrumented catchments lahar warning times can be doubled (Jones et al., 2015).

Further research to expand the length of the current two-year study period would develop the understanding of the catchment recovery-driven temporal trends in lahar occurrence identified within this study. Likewise, the application of these techniques to additional volcanoes would facilitate both the further examination of the

Commented [A19]: This existing sentence already addresses Pierson comment #12

Commented [A20]: Addresses Capra comment #12

Commented [A21]: Addresses Pierson comment #13

Commented [A22]: Addresses second part of Pierson comment #6 and Capra comment #4 performance of the lahar forecasting models and the investigation of other important parameters contributing to the frequency and magnitude of rain-triggered lahar initiation.

7 Conclusions

This study demonstrates the development and enhancement of logistic regression-based rain-triggered lahar probability estimation models for real-time lahar forecasting using the lahar occurrence and rainfall record of the Belham Valley, Montserrat between April 2010 and April 2012. The incorporation of both antecedent rainfall and considerations for temporal catchment development into such models alongside the first-order lahar forecasting parameter of peak rainfall intensity is shown to improve lahar forecasting performance. Rainfall seasonality and catchment recovery are identified as important factors in the severity of the rain-triggered lahar hazard at Soufrière Hills Volcano, Montserrat, and by extension similar volcanoes worldwide. Seasonal influences increase both the absolute number of lahars and the probability of lahar occurrence under pre-defined rainfall conditions during the wet season due to antecedent effects. Lahar probability is also shown to decline with time under given antecedent and peak rainfall intensity conditions as a product of catchment evolution. Our results demonstrate the potential for successful real-time prediction of secondary lahars using readily available input data, even in areas featuring strongly seasonal climates and periods of eruptive quiescence.

Competing Interests

The authors declare that they have no conflict of interest.

Acknowledgements

This research was supported by STREVA (NERC/ESRC consortium NE/J02483X/1) and we are thankful to the Montserrat Volcano Observatory (MVO) for permission to use the lahar database and rain gauge dataset. We thank Thomas Pierson and Lucia Capra for their constructive reviews which helped improve the paper, and Editor Thomas Glade.

Figure Captions

Figure 1: Location map of Montserrat and Soufrière Hills Volcano.

Figure 2: Timeline illustrating hourly rainfall data (above) and rain-triggered lahar activity (below) in the Belham Valley, Montserrat between April 2010 and April 2012 (with minor gaps (~~shaded~~stippled ornament) due to equipment failure). S, M, and L on the vertical axis represent Small, Medium and Large lahars respectively, see text for details.

Figure 3: Timelines displaying examples of lahar triggering rainfall in the Belham Valley, Monserrat between April 2010 and April 2012. Alongside the timing of lahar observation and/or detection, the cumulative recorded rainfall (mm) and One Hour Peak Rainfall Intensity (1hrPRI – mm hr^{-1}) of the rainfall events are displayed.

~~Figure 3~~Figure 4: Illustration of the seasonal fluctuations in lahar occurrence displayed using 6-month data windows with 1-month staggered start dates. Vertical bars indicate the number of lahar events, categorised by magnitude, in each 6-month period. Background contours display the number of rainfall events exceeding specified ~~Peak~~ One Hour Peak Rainfall Intensity (1hrPRI) thresholds, in each 6-month period.

~~Figure~~ Figure 54: Lahar probability, classified by magnitude, as categorised One Hour Peak Rainfall Intensity (1hrPRI) increases. (a) April 2010-April 2012 (b) April 2010-April 2011 (c) April 2011-April 2012.

~~Figure~~ Figure 65: Seasonal and temporal effects on lahar probability. Contour graph of empirically-derived lahar probability relative to the exceedance of One Hour Peak Rainfall Intensity (1hrPRI) thresholds in 6-month moving data windows with 1-month staggered start dates. White numbers and dashed lines show temporal trends. Following the empirically-derived 4 mm hr^{-1} PRI contour, there is a 20% probability of a lahar if this threshold is exceeded at 1 (6-month start date of 13/10/2010). This probability increases to 38% at 2 (13/04/2011); and declines to 18% at 3 (13/10/2011). Alternatively, reading horizontally across the graph for a lahar probability of 38% the associated PRI threshold increases from 4 mm hr^{-1} at 2 (13/04/2011) to approximately 15 mm hr^{-1} at 4 (13/10/2011).

~~Figure~~ Figure 76: Assessment of binary logistic regression-based lahar probability estimation models in the Belham Valley, Montserrat. (a) Illustration of four binary logistic regression-based lahar probability estimation models created from *Year 1*, *Year 2*, *Wet Season* and *Dry Season* data. (b) ROC curves assessing the lahar forecasting performance of an exclusively One Hour Peak Rainfall Intensity (1hrPRI)-centric logistic regression-based lahar probability estimation model and a multi-variable (1hrPRI, antecedent rainfall and long-term cumulative rainfall) model.

Commented [A23]: Addresses Pierson comment #16

Formatted: Font: 9 pt, Bold

Formatted: Normal, Space After: 12 pt

Formatted: Font: 9 pt, Bold

Formatted: Font: 9 pt, Bold

Formatted: Font: 9 pt, Bold

References

Alexander, J., Barclay, J., Susnik, J., Loughlin, S. C., Herd, R. A., Darnell, A., and Crosweller, S.: Sediment-charged flash floods on Montserrat: The influence of synchronous tephra fall and varying extent of vegetation damage, Journal of Volcanology and Geothermal Research, 194, 127-138, 10.1016/j.jvolgeores.2010.05.002, 2010.

Arguden, A., and Rodolfo, K.: Sedimentologic and dynamic differences between hot and cold laharic debris flows of Mayon Volcano, Philippines, Geological Society of America Bulletin, 102, 865-876, 10.1130/0016-7606(1990)102<0865:saddbh>2.3.co;2, 1990.

Auker, M. R., Sparks, R. S. J., Siebert, L., Crosweller, H. S., and Ewert, J.: A statistical analysis of the global historical volcanic fatalities record, Journal of Applied Volcanology, 2, 10.1186/2191-5040-2-2, 2013.

Barclay, J., Johnstone, J. E., and Matthews, A. J.: Meteorological monitoring of an active volcano: Implications for eruption prediction, Journal of Volcanology and Geothermal Research, 150, 339-358, 10.1016/j.jvolgeores.2005.07.020, 2006.

Barclay, J., Alexander, J., and Susnik, J.: Rainfall-induced lahars in the Belham Valley, Montserrat, West Indies, Journal of the Geological Society, 164, 815-827, 10.1144/0016-76492006-078, 2007.

Bonadonna, C., Mayberry, G. C., Calder, E. S., Sparks, R. S. J., Choux, C., Jackson, P., Lejeune, A. M., Loughlin, S. C., Norton, G. E., Rose, W. I., Ryan, G., and Young, S. R.: Tephra fallout in the eruption of Soufriere Hills Volcano, Montserrat, Geological Society, London, Memoirs, 21, 483-516, 10.1144/gsl.mem.2002.021.01.22, 2002.

Capra, L., Borselli, L., Varley, N., Gavilanes-Ruiz, J. C., Norini, G., Sarocchi, D., Caballero, L., and Cortes, A.: Rainfall-triggered lahars at Volcán de Colima, Mexico: Surface hydro-repellency as initiation process, Journal of Volcanology and Geothermal Research, 189, 105-117, 10.1016/j.jvolgeores.2009.10.014, 2010.

Cole, P. D., Calder, E. S., Sparks, R. S. J., Clarke, A. B., Druitt, T. H., Young, S. R., Herd, R. A., Harford, C. L., and Norton, G. E.: Deposits from dome-collapse and fountain-collapse pyroclastic flows at Soufriere Hills Volcano, Montserrat, Geological Society, London, Memoirs, 21, 231-262, 10.1144/gsl.mem.2002.021.01.11, 2002.

Collins, B. D., and Dunne, T.: Erosion of tephra from the 1980 eruption of Mount St Helens, Geological Society of America Bulletin, 97, 896-905, 10.1130/0016-7606(1986)97<896:eotfte>2.0.co;2, 1986.

Doyle, E. E., Cronin, S. J., and Thouret, J. C.: Defining conditions for bulking and debulking in lahars, Geological Society of America Bulletin, 123, 1234-1246, 10.1130/B30227.1, 2011.

Fagents, S. A., and Baloga, S. M.: Toward a model for the bulking and debulking of lahars, Journal of Geophysical Research, 111, 10.1029/2005jb003986, 2006.

Fawcett, T.: An introduction to ROC analysis, Pattern Recognition Letters, 27, 861-874, 10.1016/j.patrec.2005.10.010, 2006.

Fohrer, N., Berkenhagen, J., Hecker, J. M., and Rudolph, A.: Changing soil and surface conditions during rainfall - Single rainstorm/subsequent rainstorms, CATENA, 37, 355-375, Doi 10.1016/S0341-8162(99)00026-0, 1999.

Froude, M. J.: Lahar Dynamics in the Belham River Valley, Montserrat: Application of Remote Camera-Based Monitoring for Improved Sedimentological Interpretation of Post-Event Deposits, PhD Thesis, School of Environmental Science, University of East Anglia, 2015.

Graf, W. L.: The rate law in fluvial geomorphology, American Journal of Science, 277, 178-191, 1977.

Formatted: Normal, Space After: 0 pt, No bullets or numbering, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Hosmer Jr, D. W., Lemeshow, S., and Sturdivant, R. X.: Applied logistic regression, John Wiley & Sons, 2013.

Iverson, R. M.: Landslide triggering by rain infiltration, Water Resources Research, 36, 1897-1910, 10.1029/2000wr900090, 2000.

Iverson, R. M., Reid, M. R., Logan, M., LaHusen, R. G., Godt, J. W., and Griswold, J. P.: Positive feedback and momentum growth during debris-flow entrainment of wet bed sediment, Nature Geoscience, 4, 116-121, 10.1038/NGEO1040, 2011.

Jones, R., Manville, V., and Andrade, D.: Probabilistic analysis of rain-triggered lahar initiation at Tungurahua volcano, Bulletin of Volcanology, 77, 10.1007/s00445-015-0946-7, 2015.

Komorowski, J. C., Legendre, Y., Christopher, T., Bernstein, M., Stewart, R., Joseph, E., Fournier, N., Chardot, L., Finizola, A., Wadge, G., Syers, R., Williams, C., and Bass, V.: Insights into processes and deposits of hazardous vulcanian explosions at Soufrière Hills Volcano during 2008 and 2009 (Montserrat, West Indies), Geophysical Research Letters, 37, 10.1029/2010gl042558, 2010.

Formatted: Justified, Indent: Left: 0 cm, First line: 0 cm, Line spacing: 1.5 lines, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Lavigne, F., and Suwa, H.: Contrasts between debris flows, hyperconcentrated flows and stream flows at a channel of Mount Semeru, East Java, Indonesia. Geomorphology, 61, 41-58, 2004.

Formatted: Font: Not Italic

Lavigne, F., Thouret, J. C., Voight, B., Young, K., LaHusen, R., Marso, J., Suwa, H., Sumaryono, A., Sayudi, D. S., and Dejean, M.: Instrumental lahar monitoring at Merapi Volcano, Central Java, Indonesia, Journal of Volcanology and Geothermal Research, 100, 457-478, 10.1016/S0377-0273(00)00151-7, 2000.

Leavesley, G., Lusby, G., and Lichty, R.: Infiltration and erosion characteristics of selected tephra deposits from the 1980 eruption of Mt St Helens, Washington, USA, Hydrological Sciences, 34, 339-353, 1989.

Major, J. J., and Yamakoshi, T.: Decadal-scale change of infiltration characteristics of a tephra-mantled hillslope at Mount St Helens, Washington, Hydrological Processes, 19, 3621-3630, 10.1002/Hyp.5863, 2005.

Major, J. J., Pierson, T. C., Dinehart, R. L., and Costa, J. E.: Sediment yield following severe volcanic disturbance - A two-decade perspective from Mount St. Helens, Geology, 28, 819-822, 10.1130/0091-7613(2000)28<819:Syfsvd>2.0.Co;2, 2000.

Manville, V., Hodgson, K. A., Houghton, B. F., Keys, J. R. H., and White, J. D. L.: Tephra, snow and water: complex sedimentary responses at an active, snow-capped stratovolcano, Ruapehu, New Zealand, Bulletin of Volcanology, 62, 278-293, 2000.

Marcial, S., Melosantos, A., Hadley, K., LaHusen, R., and Marso, J.: Instrumental Lahar Monitoring at Mount Pinatubo, in: Fire and Mud, Eruptions and Lahars of Mt Pinatubo, Philippines, edited by: Newhall, C., and Punongbayan, R., PHIVOLCS/University of Washington Press, Quezon City/Seattle, 1015-1023, 1996.

Mead, S., Magill, C., and Hilton, J.: Rain-triggered lahar susceptibility using a shallow landslide and surface erosion model, Geomorphology, 273, 168-177, 10.1016/j.geomorph.2016.08.022, 2016.

Ogawa, Y., Daimaru, H., and Shimizu, A.: Experimental study of post-eruption overland flow and sediment load from slopes overlain by pyroclastic-flow deposits, Unzen volcano, Japan, Géomorphologie~~-~~: relief, processus, environnement, 13, 237-246, 10.4000/geomorphologie.3962, 2007.

Formatted: Justified, Indent: Left: 0 cm, First line: 0 cm, Line spacing: 1.5 lines, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Okano, K., Suwa, H., and Kanno, T.: Characterization of debris flows by rainstorm condition at a torrent on the Mount Yakedake volcano, Japan. Geomorphology, 136, 88-94, 2012.

Formatted: Font: Not Italic

Pierson, T. C., and Major, J. J.: Hydrogeomorphic effects of explosive volcanic eruptions on drainage basins, Annual Review of Earth and Planetary Sciences, 42, 469-507, 10.1146/annurev-earth-060313-054913, 2014.

Pierson, T. C., Janda, R. J., Umbal, J. V., and Daag, A. S.: Immediate and long-term hazards from lahars and excess sedimentation in rivers draining Mt. Pinatubo, Philippines. U.S. Geological Survey Water-Resources Investigations Report, 92-4039, 183-203, 1992

Rodolfo, K., and Arguden, A.: Rain-lahar generation and sediment-delivery systems at Mayon Volcano, Philippines, in: Sedimentation in Volcanic Settings, edited by: Fisher, R., and Smith, G., SEPM, Special Publication 45, 71-87, 1991.

Schumm, S. A., and Rea, D. K.: Sediment Yield from Disturbed Earth Systems, Geology, 23, 391-394, 10.1130/0091-7613(1995)023<0391:Syfdes>2.3.Co;2, 1995.

Segerstrom, K.: Erosion studies at Paricutin, State of Michoacan, Mexico, USGS Bulletin, 965-A, 164 pp, 1950.

Smith, G. A., and Fritz, W. J.: Volcanic influences on terrestrial sedimentation, Geology, 17, 375-376, 1989.

Stinton, A. J., Cole, P. D., Stewart, R. C., Odbert, H. M., and Smith, P.: The 11 February 2010 partial dome collapse at Soufriere Hills Volcano, Montserrat, Geological Society, London, Memoirs, 39, 133-152, 10.1144/m39.7, 2014.

Thouret, J. C., Oehler, J. F., Gupta, A., Solikhin, A., and Procter, J. N.: Erosion and aggradation on persistently active volcanoes-a case study from Semeru Volcano, Indonesia, Bulletin of Volcanology, 76, 10.1007/S00445-014-0857-Z, 2014.

Todisco, F.: The internal structure of erosive and non-erosive storm events for interpretation of erosive processes and rainfall simulation, Journal of Hydrology, 519, 3651-3663, 10.1016/j.jhydrol.2014.11.002, 2014.

Waldron, H. H.: Debris flow and erosion control problems caused by the ash eruptions of Irazu Volcano, Costa Rica, United States Geological Survey, Bulletin 1241-I, 37 p., 1967.

Tungol, N., and Regalado, T.: Rainfall, acoustic flow monitor records, and observed lahars of the Sacobia River in 1992, in: Fire and Mud, Eruptions and Lahars of Mt Pinatubo, Philippines, edited by: Newhall, C., and Punongbayan, R., PHIVOLCS/University of Washington Press, Quezon City/Seattle, 1023-1033, 1996.

Van Westen, C., and Daag, A.: Analysing the relation between rainfall characteristics and lahar activity at Mt Pinatubo, Philippines, Earth Surface Processes and Landforms, 30, 1663-1674, 2005.

Wang, B., Kim, H.-J., Kikuchi, K., and Kitoh, A.: Diagnostic metrics for evaluation of annual and diurnal cycles, Climate Dynamics, 37, 941-955, 10.1007/s00382-010-0877-0, 2010.

Wischmeier, W., and Smith , D.: Predicting rainfall erosion losses - A guide to conservation planning, Agricultural Handbooks (USA) No. 537, US Department of Agriculture, Washington DC, 1978.

Yamakoshi, T., and Suwa, H.: Post-eruption characteristics of surface runoff and sediment discharge on the slopes of pyroclastic-flow deposits, Mt Unzen, Japan, Transactions, Japanese Geomorphological Union, 21, 469-497, 2000.

Formatted: Justified, Indent: Left: 0 cm, First line: 0 cm, Line spacing: 1.5 lines, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Formatted: Font color: Blue

Fig.1

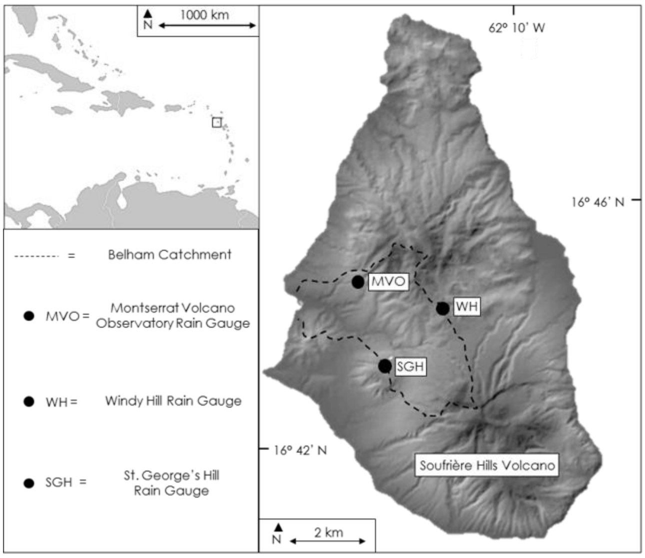
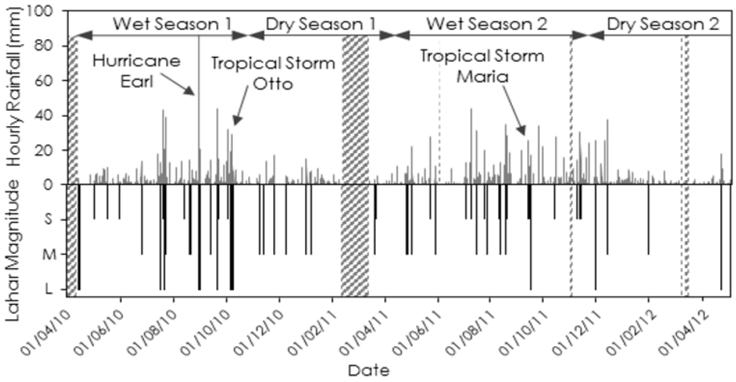
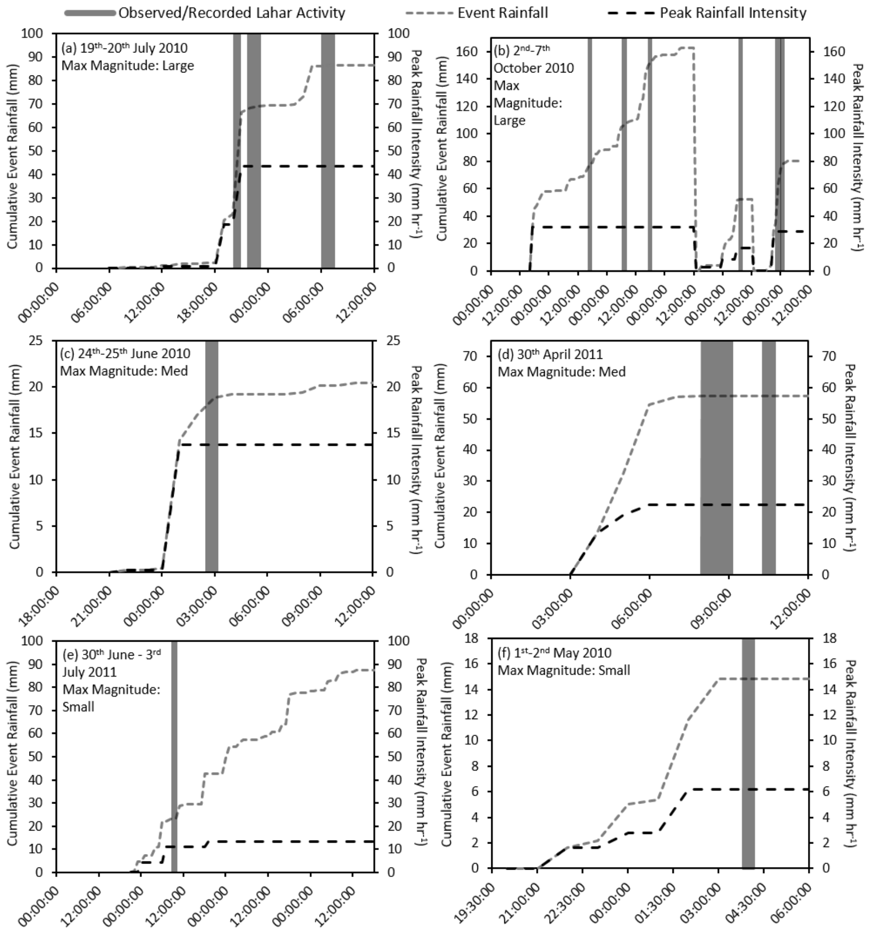

N
1000 km
62° 10' W
16° 46' N
16° 42' N
= Belham Catchment
MVO = Montserrat Volcano Observatory Rain Gauge
WH = Windy Hill Rain Gauge
SGH = St. George's Hill Rain Gauge
MVO
WH
SGH
Soufrière Hills Volcano
N
2 km

Fig.2


Wet Season 1
Dry Season 1
Wet Season 2
Dry Season 2
Hurricane Earl
Tropical Storm Otto
Tropical Storm Maria
Hourly Rainfall (mm)
Lahar Magnitude
100
80
60
40
20
0
S
M
L
01/04/10
01/06/10
01/08/10
01/10/10
01/12/10
01/02/11
01/04/11
01/06/11
01/08/11
01/10/11
01/12/11
01/02/12
01/04/12
Date

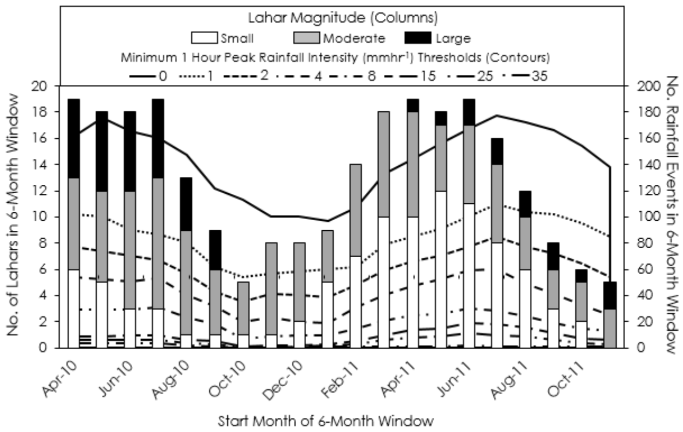
474 Fig. 3



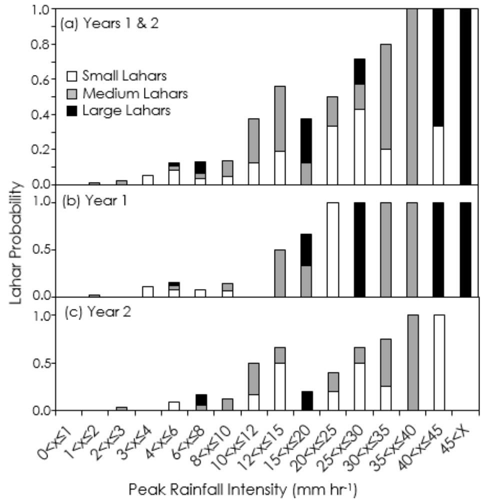
475

Commented [A24]: New figure as suggested by Capra comment #3 and Pierson comment #17

~~Fig. 3~~Fig. 4

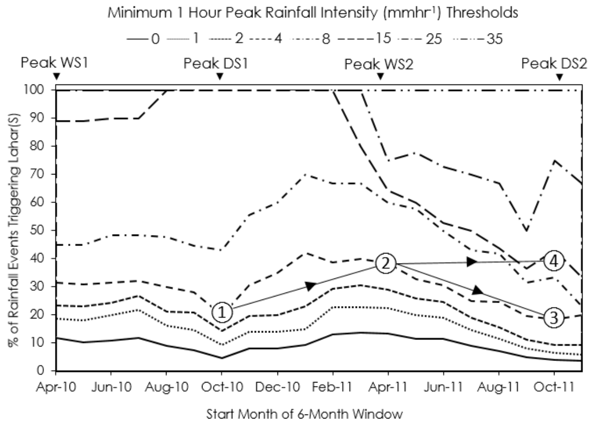

Lahar Magnitude (Columns)
Small
Moderate
Large
Minimum 1 Hour Peak Rainfall Intensity (mmhr⁻¹) Thresholds (Contours)
0
1
2
4
8
15
25
35
No. of Lahars in 6-Month Window
No. Rainfall Events in 6-Month Window
Start Month of 6-Month Window
Apr-10
Jun-10
Aug-10
Oct-10
Dec-10
Feb-11
Apr-11
Jun-11
Aug-11
Oct-11

~~Fig. 4~~Fig. 5

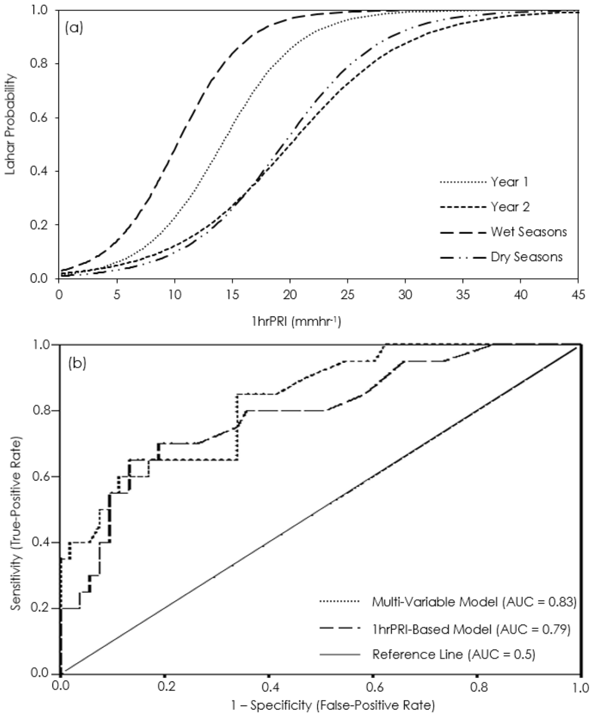


(a) Years 1 & 2
Small Lahars
Medium Lahars
Large Lahars
(b) Year 1
(c) Year 2
Lahar Probability
1.0
0.8
0.6
0.4
0.2
0.0
0.5
0<x≤1
1<x≤2
2<x≤3
3<x≤4
4<x≤6
6<x≤8
8<x≤10
10<x≤12
12<x≤15
15<x≤20
20<x≤25
25<x≤30
30<x≤35
35<x≤40
40<x≤45
45<X
Peak Rainfall Intensity (mm hr-1)

~~Fig. 5~~Fig. 6


Minimum 1 Hour Peak Rainfall Intensity (mmhr^{-1}) Thresholds
0
1
2
4
8
15
25
35
Peak WS1
Peak DS1
Peak WS2
Peak DS2
% of Rainfall Events Triggering Lahar(S)
0
10
20
30
40
50
60
70
80
90
100
1
2
3
4
Apr-10
Jun-10
Aug-10
Oct-10
Dec-10
Feb-11
Apr-11
Jun-11
Aug-11
Oct-11
Start Month of 6-Month Window

~~Fig. 6~~Fig. 7



(a)
Lahar Probability
1hrPRI (mmhr⁻¹)
Year 1
Year 2
Wet Seasons
Dry Seasons
(b)
Sensitivity (True-Positive Rate)
1 – Specificity (False-Positive Rate)
Multi-Variable Model (AUC = 0.83)
1hrPRI-Based Model (AUC = 0.79)
Reference Line (AUC = 0.5)

Appendix I

Commented [A25]: New appendix explaining ROC analysis in response to Pierson comment #14

Receiver Operating Characteristic (ROC) analysis is a statistical technique that is used to illustrate the diagnostic ability of a binary classifier system (i.e. a system that subdivides the elements of a given dataset into two groups, for example the presence or absence of a disease, a pass or a fail in a test etc.). The method was first developed by electrical and radar engineers during World War II, and has since been used in psychology, medicine, meteorology, and forecasting of natural hazards.

A graphical plot, or Receiver Operating Characteristics curve (ROC curve) is often used to illustrate the effect of varying the value of the classifying parameter (for example the number of cancer cells per microlitre of blood or the pass mark in the previous example). The ROC curve is generated by plotting the true positive rate (TPR) against the false positive rate (FPR) as the value of the classifying, or threshold parameter, is changed. There are four possible outcomes from a binary classifier (Table A1): (i) correct prediction of an event that really did occur = true positive; (ii) incorrect prediction of an event that did not occur = false positive; (iii) predicting no event when an event does happen = false negative; and (iv) correct prediction that no event occurs and no event really does occur = true negative.

Imagine a situation where there are 200 patients undergoing a medical test, where alpha is some diagnostic threshold for having a medical condition. At a given value of alpha, the contingency table could resemble Table A2.

Here, the TPR is the number of true positives divided by the total number of predicted positives (both true and false), or 70/(70+30) = 0.70

The FPR is the number of false positives divided by the total number of predicted negatives (both true and false), or 28/(28+72) = 0.28

Thus for this value of alpha, the corresponding point would plot at (0.63, 0.28) on Figure A1 (the white square). By systematically varying the value of the threshold parameter alpha, a whole series of 2x2 contingency tables would be generated, producing an array of points in ROC space and hence a curve (the dashed line).

A 100% rate of prediction (all true positives) would plot at (0, 1) on Figure A1 (the grey circle), whereas a 50% accurate rate of prediction (i.e. guessing the outcome of a coin toss) would plot at (0.5, 0.5). Random guesses thus plot along a diagonal line: points above the line represent predictions better than random, points below the line predictions worse than random.

Appendix I: Table Captions

Table A1: 2x2 contingency table showing the possible outcomes of a binary classifier system.

Table A2: 2x2 contingency table for 200 patients undergoing a medical test for the presence or absence of a condition.

Appendix I: Figure Captions

Fig. A1: ROC space and plots of the prediction examples discussed in the text.

Formatted: Font: Bold

Table A1

Total population	Event happens	Event does not happen
Predict it happens	True positive	False positive
Predict it does not happen	False negative	True negative

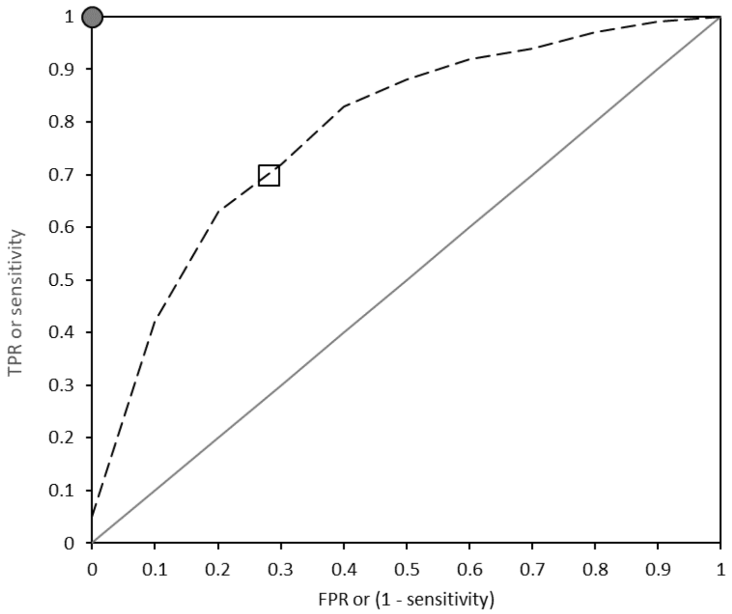
Formatted Table

Table A2

	Has condition	Has no condition
Predict has condition	70	30
Predict has no condition	28	72

Formatted Table

Fig. A1



TPR or sensitivity
FPR or (1 - sensitivity)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1