# RAINFALL-TRIGGERED LANDSLIDES: CONDITIONS, PREDICTION, AND WARNING

**LISA VICTORIA LUNA**

CUMULATIVE DISSERTATION
for the degree of
*doctor rerum naturalium (Dr. rer. nat.)*
in Georisks

submitted to the Faculty of Science
**University of Potsdam**

submitted on February 21, 2023 | defended on June 22, 2023

**First supervisor:**
Prof. Oliver Korup, PhD

**Second supervisor:**
Prof. Dr. Dr. h.c. mult. Jürgen Kurths

**Mentor:**
Prof. Taylor Schildgen, PhD

**Reviewers:**
Dr. rer. nat. habil. Wolfgang Schwanghart, University of Potsdam
Prof. Alison Duvall, PhD, University of Washington

Published online on the
Publication Server of the University of Potsdam:
https://doi.org/10.25932/publishup-60092
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-600927

## STATEMENT OF ORIGINALITY

I, Lisa Victoria Luna, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. Where studies were conducted jointly with co-authors, I state what their and my contributions were in the declaration of authorship.

All the work described in this thesis was carried out while in candidature for a doctoral degree at the Institute of Geosciences at the University of Potsdam. This dissertation is not substantially the same as any that I have submitted, or is being concurrently submitted, for a degree or diploma or other qualification at University of Potsdam or any other university or similar institution.

_____ _____

Place, Date Lisa Victoria Luna

# ACKNOWLEDGEMENTS

I am grateful to many people and institutions for support that has enabled me to complete this dissertation. My heartfelt thanks go to all of you.

- RTG NatRiskChange for providing not just funding but also a supportive network of curious minds.
- My advising team: My first advisor, Oliver Korup, for giving me freedom to explore, encouragement at the right moments, and honest feedback. My advisors at PIK, Jürgen Kurths and Norbert Marwan, for providing a productive working environment and a broader research network. My mentor Taylor Schildgen, for encouraging walks and much thoughtful advice.
- My collaborators: Georg Veh, for friendship and for always being excited to talk science with me, in concert hall lobbies and glacial valleys, Annette Patton, for many rewarding scientific Zoom exchanges and joint learning, Maria Isabel Arango Carmona, for helping with so much data, Sara Maria Vallejo Bernal, for sharing unabashed enthusiasm and ambition with me.
- Elisabeth Schönfeldt, for being the best PhD tandem buddy I could have wished for, through the highs and lows.
- The Australia wildfires task force: Matthias Kemter, Melanie Fischer, Elisabeth Schönfeldt, Kata Schmidt, and Joscha Vogel, for collaborative efforts that were a highlight of my PhD experience.
- Friends who went on many treasured long walks with me through pandemic lockdowns and PhD struggles, in Potsdam parks and across the Alps: Emma Loizeaux, Giorgia Di Capua, Alyna Lange, Stephanie Olen, Maria Schönen.
- The Potsdam Graduate School (POGS), particularly Kerstin Hille and Jasmin Blümlein, for providing indispensable learning experiences. The amazing women of my Mentoring Plus cohort.
- Laura Deming (International Counseling and Coaching) and Edda Wilde (Polyfon Coaching) for vital coaching and advice that helped me grow so much.
- Key teachers and mentors from my past who I feel helped me to reach this point. Although I'm surely forgetting some: Willy Amidon, Jeff Munroe, Doug Burbank, Burch Fisher, and Dylan Rood, for mentoring and treating me like a "real scientist" a decade before I submitted my doctoral dissertation. Hanna Fekete, Takeshi Kuramochi, Frauke Röser, Thomas Day, Niklas Höhne for teaching me professionalism. Jeff Howarth, Rhonda Spidell, Joan Alessi, Dale Huber, Agustin Kintanar.
- My Middlebury College GEOL0251 class, for giving me the opportunity to try on the persona of "Professor Lisa" and undoubtedly teaching me more than I taught you.

- Myriad people who I do not know but am indebted to, for the code, data, and learning resources they freely provide: in particular, Richard McElreath, Paul-Christian Bürkner, Matthew Kay. The contributors of StackOverflow. This document was formatted using a modified version of Oriol Pavón Aroca's thesis template (CC BY 4.0).
- Folks who helped me finalize this document: Melanie Fischer, for the German language summary, Elli Schönfeldt, Joscha Vogel, Matthias Kemter, Kata Schmidt, and Maria Arango Carmona, for dedicated proofreading.
- My parents, for always supporting me in whatever it was that I wanted to do. Leigh, for being there for me always.
- Myself. For grit, perseverance, self-reliance, and resilience. For keeping myself healthy and fed through a simultaneous global pandemic and PhD program.
- This work was funded by the DFG RTG "Natural Hazards and Risks in a Changing World" (NatRiskChange GRK 2043).

# ABSTRACT

Rainfall-triggered landslides are a globally occurring hazard that cause several thousand fatalities per year on average and lead to economic damages by destroying buildings and infrastructure and blocking transportation networks. For people living and governing in susceptible areas, knowing not only where, but also when landslides are most probable is key to inform strategies to reduce risk, requiring reliable assessments of weather-related landslide hazard and adequate warning. Taking proper action during high hazard periods, such as moving to higher levels of houses, closing roads and rail networks, and evacuating neighborhoods, can save lives. Nevertheless, many regions of the world with high landslide risk currently lack dedicated, operational landslide early warning systems.

The mounting availability of temporal landslide inventory data in some regions has increasingly enabled data-driven approaches to estimate landslide hazard on the basis of rainfall conditions. In other areas, however, such data remains scarce, calling for appropriate statistical methods to estimate hazard with limited data. The overarching motivation for this dissertation is to further our ability to predict rainfall-triggered landslides in time in order to expand and improve warning. To this end, I applied Bayesian inference to probabilistically quantify and predict landslide activity as a function of rainfall conditions at spatial scales ranging from a small coastal town, to metropolitan areas worldwide, to a multi-state region, and temporal scales from hourly to seasonal. This thesis is composed of three studies.

In the first study, I contributed to developing and validating statistical models for an online landslide warning dashboard for the small town of Sitka, Alaska, USA. We used logistic and Poisson regressions to estimate daily landslide probability and counts from an inventory of only five reported landslide events and 18 years of hourly precipitation measurements at the Sitka airport. Drawing on community input, we established two warning thresholds for implementation in the dashboard, which uses observed rainfall and US National Weather Service forecasts to provide real-time estimates of landslide hazard.

In the second study, I estimated rainfall intensity-duration thresholds for shallow landsliding for 26 cities worldwide and a global threshold for urban landslides. I found that landslides in urban areas occurred at rainfall intensities that were lower than previously reported global thresholds, and that 31% of urban landslides were triggered during moderate rainfall events. However, landslides in cities with widely varying climates and topographies were triggered above similar critical rainfall intensities: thresholds for 77% of cities were indistinguishable from the global threshold, suggesting that urbanization may harmonize thresholds between cities, overprinting natural variability. I provide a baseline threshold that could be considered for warning in cities with limited landslide inventory data.

In the third study, I investigated seasonal landslide response to annual precipitation patterns in the Pacific Northwest region, USA by using Bayesian multi-level models to combine data from five heterogeneous landslide inventories that cover different areas and time periods. I quantitatively confirmed a distinctly seasonal pattern of landsliding and found that peak landslide activity lags the annual precipitation peak. In February, at the height of the landslide season, landslide intensity for a given amount of monthly rainfall is up to ten times higher than at the season onset in November, underlining the importance of antecedent seasonal hillslope conditions.

Together, these studies contributed actionable, objective information for landslide early warning and examples for the application of Bayesian methods to probabilistically quantify landslide hazard from inventory and rainfall data.

# KURZFASSUNG

Durch Regenfälle ausgelöste Erdrutsche sind eine weltweit auftretende Gefahr, die im Durchschnitt mehrere tausend Todesopfer pro Jahr fordern und zu wirtschaftlichen Schäden führen, indem sie Gebäude und Infrastrukturen zerstören und Verkehrsnetze blockieren. Für Bewohner, sowie lokale Regierungen in potentiell gefährdeten Gebieten, ist es entscheidend zu wissen, nicht nur wo, sondern auch wann Erdrutsche am wahrscheinlichsten sind, um Strategien zur Verringerung des Risikos zu entwickeln. Dies erfordert zuverlässige Bewertungen der wetterbedingten Erdrutschgefahr und eine angemessene Warnung. Angemessene Maßnahmen während Hochrisikoperioden, wie der Umzug in höhere Etagen, die Sperrung von Straßen und Schienennetzen, sowie die Evakuierung von Wohngebieten, können Leben retten. In vielen Regionen mit hohem Erdrutschrisiko gibt es jedoch derzeit keine spezifischen, einsatzfähigen Frühwarnsysteme für Erdrutsche.

In einigen Regionen ermöglichte die zunehmende Verfügbarkeit von zeitlich-aufgelösten Erdrutschdaten datengestützte Ansätze zur Abschätzung der Erdrutschgefahr auf Grundlage von Niederschlagsbedingungen. In anderen Gebieten sind solche Daten jedoch nach wie vor spärlich, sodass geeignete statistische Methoden erforderlich sind, um die Gefährdung trotz einer begrenzten Datenmenge abzuschätzen. Die übergreifende Motivation für diese Dissertation besteht darin, unsere Fähigkeit zur rechtzeitigen Vorhersage von niederschlagsbedingten Erdrutschen zu verbessern, um Frühwarnsysteme zu erweitern und optimieren. Zu diesem Zweck habe ich Bayes'sche Inferenz angewandt, um die Erdrutschaktivität in Abhängigkeit von den Niederschlagsbedingungen probabilistisch zu quantifizieren und vorherzusagen. Meine Studien decken dabei sowohl eine breite räumliche Skala, welche von einer lokalen bis regionalen Betrachtung reicht, als auch eine von stündlich bis saisonal reichende zeitliche Skala ab. Diese Dissertation setzt sich aus drei Studien zusammen.

In der ersten Studie habe ich zur Entwicklung und Validierung statistischer Modelle für ein Online-Dashboard zur Erdrutschwarnung in der Kleinstadt Sitka, Alaska, USA, beigetragen. Wir verwendeten logistische und Poisson-Regressionen zur Einschätzung der täglichen Erdrutschwahrscheinlichkeit und der Anzahl der Erdrutsche auf Grundlage von nur fünf dokumentierten Erdrutschereignissen und 18 Jahren stündlicher Niederschlagsmessungen am Flughafen von Sitka. Basierend auf Hinweisen aus der Bevölkerung legten wir zwei Warnschwellenwerte für die Umsetzung des Dashboards fest, welches wiederum beobachtete Niederschläge und Vorhersagen des US-amerikanischen Wetterdienstes (US National Weather Service) nutzt, um Echtzeiteinschätzungen der Erdrutschgefahr zu liefern.

In der zweiten Studie habe ich Schwellenwerte für die Niederschlagsintensität und -dauer für Erdrutsche in 26 Städten weltweit, sowie einen globalen Schwellenwert für urbane Erdrutsche ermittelt. Dabei stellte ich fest, dass Erdrutsche in urbanen Gebieten

bei Niederschlagsintensitäten auftreten, die unter den zuvor gemeldeten globalen Schwellenwerten liegen, und dass 31 % der Erdrutsche in Städten durch moderate Niederschlagsereignisse ausgelöst wurden. Erdrutsche in Städten mit sehr unterschiedlichen klimatischen und topografischen Bedingungen wurden jedoch bei vergleichbaren kritischen Niederschlagsintensitäten ausgelöst: Für 77 % der Städte unterschieden sich die lokalen Schwellenwerte nicht von den globalen Schwellenwerten, was darauf hindeutet, dass eine zunehmende Urbanisierung die Schwellenwerte zwischen Städten angleicht und natürliche Schwankungen überlagern kann. Ich habe einen Basisschwellenwert festgelegt, der für die Warnung in Städten mit begrenzten Erdrutschdaten in Betracht gezogen werden könnte.

In der dritten Studie untersuchte ich saisonale Reaktionen von Erdrutschen auf jährliche Niederschlagsmuster im pazifischen Nordwesten der USA. Dafür verwendete ich Bayes'sche Mehrebenenmodelle, um Daten aus fünf heterogenen Erdrutschinventaren zu kombinieren, welche unterschiedliche Gebiete und Zeiträume abdecken. Ich fand heraus, dass Erdrutsche deutlich saisonabhängig sind und dass der Höhepunkt der Erdrutschaktivität mit einem zeitlichen Versatz auf den jährlichen Niederschlagsspitzenwert folgt. Im Februar, auf dem Höhepunkt der Erdrutschsaison, ist die Erdrutschintensität bei einer gegebenen monatlichen Niederschlagsmenge bis zu zehnmal höher als zu Beginn der Saison im November. Dies unterstreicht die Bedeutung von vorherigen saisonalen Hangbedingungen.

Zusammengefasst liefern die in dieser Dissertation vorgestellten Studien umsetzbare, objektive Informationen für die Frühwarnung vor Erdrutschen und Beispiele für die Anwendung von Bayes'schen Methoden zur probabilistischen Quantifizierung der Erdrutschgefahr mittels Bestands- und Niederschlagsdaten.

# STATEMENT OF CONTRIBUTIONS

This cumulative dissertation includes three manuscripts that have been published in or are intended for publication in international peer-reviewed journals. Chapter 2 is the result of a collaboration with Annette Patton, with input from our co-authors. Chapters 3 and 4 are primarily my work, with support from my co-authors. In detail, the contributions of individual co-authors to each manuscript were:

**Chapter 2**: A.I.P. and J.J.R. conceptualized the study and A.I.P. and L.V.L. co-led the study. L.V.L. and A.I.P. performed statistical analyses, produced figures, and co-wrote the text. J.J.R., A.J., O.K., and B.B.M. contributed to discussing results and reviewing the manuscript. L.V.L. and A.I.P. revised and edited the text.

**Chapter 3**: L.V.L. conceptualized and led the study, performed the statistical analysis, and drafted the manuscript. L.V.L., M.I.A.C, and E.L. processed the data. L.V.L. and G.V. designed the statistical models and created the figures. G.V., O.K., U.O., M.I.A.C, and E.L. contributed to discussing the results and reviewing the manuscript. L.V.L, G.V., and O.K. revised and edited the text.

**Chapter 4**: L.V.L and O.K. conceptualized the study. L.V.L. processed the data, conducted the statistical analysis, created the figures, and drafted the original manuscript. L.V.L. and O.K. revised and edited the text.

L.V.L. = Lisa V. Luna
M.I.A.C. = Maria Isabel Arango Carmona
E.L. = Elisabeth Lewis
G.V. = Georg Veh
O.K. = Oliver Korup
U.O. = Ugur Öztürk
A.I.P. = Annette I. Patton
J.J.R. = Josh J. Roering
A.J. = Aaron Jacobs
B.B.M. = Ben B. Mirus

## ADDITIONAL PUBLICATIONS

In addition to the chapters presented in this thesis, I contributed to the following peer-reviewed publications during my doctoral studies. Publications based substantially on work that I conducted before beginning my doctoral studies are indicated with a *.

Veh, G., Lützow, N., Tamm, J., **Luna, L. V.**, Hugonnet, R., Vogel, K., et al. (2023). Less extreme and earlier outbursts of ice-dammed lakes since 1900. *Nature*, 1–7. https://doi.org/10.1038/s41586-022-05642-9

Fisher, G. B., **Luna, L. V**., Amidon, W. H., Burbank, D. W., de Boer, B., Stap, L. B., et al. (2023). Milankovitch-paced erosion in the southern Central Andes. *Nature Communications*, 14(1), 424. https://doi.org/10.1038/s41467-023-36022-0 *

Fekete, H., Kuramochi, T., Roelfsema, M., Elzen, M. den, Forsell, N., Höhne, N., **Luna, L.V.**, et al. (2021). A review of successful climate change mitigation policies in major emitting economies and the potential of global replication. *Renewable and Sustainable Energy Reviews*, 137, 110602. https://doi.org/10.1016/j.rser.2020.110602 *

Kemter, M., Fischer, M., **Luna, L.V.**, Schönfeldt, E., Vogel, J., Banerjee, A., et al. (2021). Cascading hazards in the aftermath of Australia's 2019/2020 Black Summer Wildfires. *Earth's Future*, 9(3). https://doi.org/10.1029/2020EF001884

Stirling, M. W., Abbott, E. R., Rood, D. H., McVerry, G. H., Abrahamson, N. A., Barrell, D. J. A., Huso, R., Litchfield, N.J., **Luna, L.V.**, et al. (2021). First use of fragile geologic features to set the design motions for a major existing engineered structure. *Bulletin of the Seismological Society of America*, 111(5), 2673–2695. https://doi.org/10.1785/0120210026 *

# CONTENTS

# FIGURES

# TABLES

# 1 INTRODUCTION

## 1.1 RAINFALL-TRIGGERED LANDSLIDES

Rainfall-triggered landslides rank fourth after floods, storms, and earthquakes among the world's deadliest natural hazards (Haque et al., 2019), causing nearly 60,000 fatalities globally between 2004 and 2016 according to detailed compilations of news reports, press releases, and other sources (Froude & Petley, 2018) (Figure 1.1). Beyond fatalities, landslides routinely damage infrastructure and can lead to further economic losses by, for example, obstructing road networks (Postance et al., 2017). While global assessments of economic losses due to landslides are scarce, annual damage costs were estimated at 14.8 million USD (year 2000 $) for the San Francisco Bay Area, USA (Crovelli & Coe, 2009) and at 300 million USD for Germany (Klose et al., 2016). In 2022 alone, example rainfall-triggered landslide disasters included Petrópolis, Brazil, where landslides killed 231 people and destroyed 60 homes in February (Alcântara et al., 2022), Durban, South Africa, where landslides and flooding displaced over 40,000 people in April (ReliefWeb, 2022).

While geotechnical engineering measures may be able to stabilize slopes locally (Choi & Cheung, 2013), regularly updated assessments of landslide hazard are necessary to give officials and residents the information needed to make informed decisions to mitigate landslide impacts in susceptible areas (Kockelman, 1986; Larsen, 2008). A grand challenge for landslide researchers is therefore to predict reliably when landslides will occur and to disseminate this information in a timely manner, for example through landslide early warning systems (LEWS). Taking proper action during high hazard periods, such as moving to higher levels of houses, closing roads and rail networks, and evacuating neighborhoods, can save lives (Intrieri et al., 2013; Pollock & Wartman, 2020).

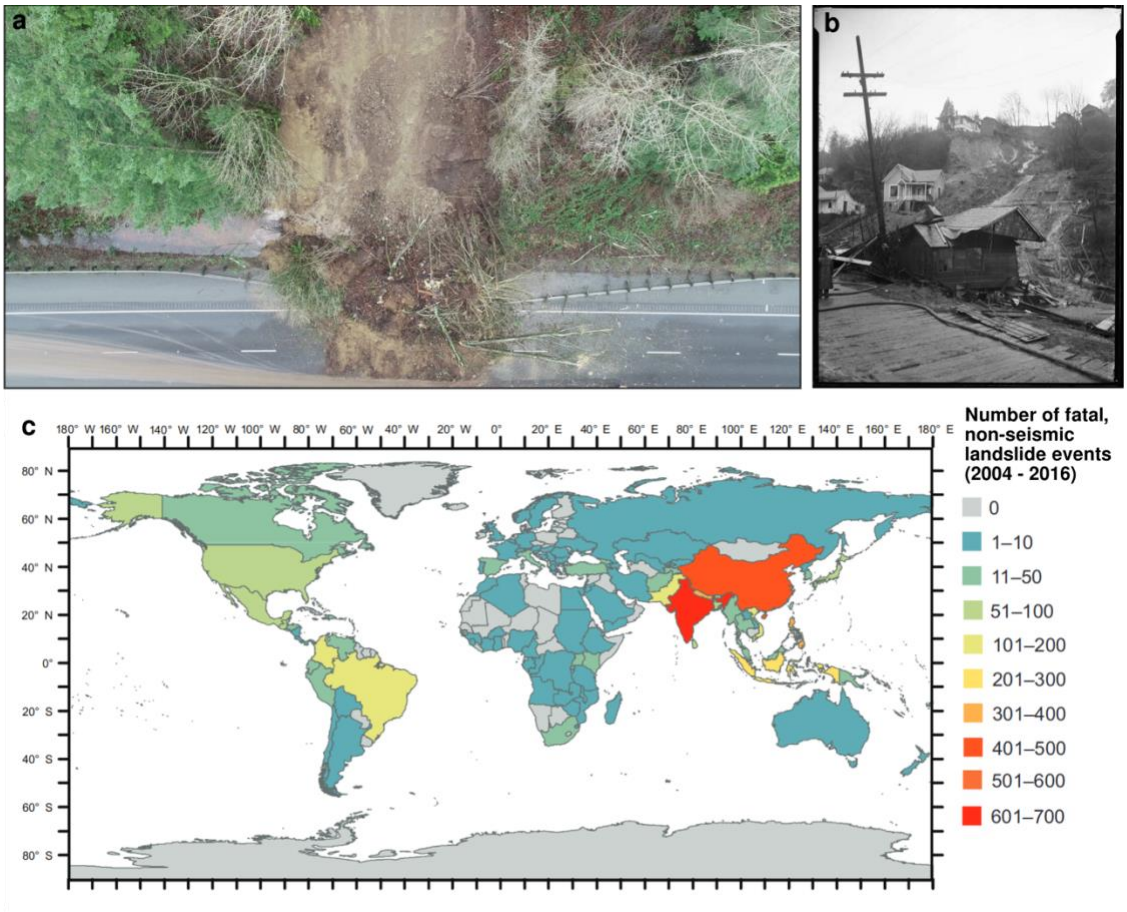



**Figure 1.1. Rainfall-triggered landslide impacts.** (a) Landslide damage to Interstate 5 in Washington, USA. Photo: Washington State Department of Transportation, CCBY NC-ND 2.0 https://creativecommons.org/licenses/by-nc-nd/2.0/). (b) A fatal landslide in Seattle, Washington, USA in 1921. Photo: PEMCO Webster & Stevens Collection, Museum of History & Industry, Seattle, Public Domain via Wikimedia Commons. (c) Globally reported fatal, non-seismically triggered landslide events between 2004 and 2016. Figure modified from (Froude & Petley, 2018).

In the past years, multi-temporal landslide inventory data has become increasingly available in some regions, enabling data-driven analyses of landslide timing, however, such data remains scarce in other regions (e.g. Kirschbaum et al., 2015; Mirus et al., 2020; Peres & Cancelliere, 2021; Stanley et al., 2021). Despite the pressing need for information, only five countries, 13 regions, and four metropolitan areas benefit from a dedicated, operating landslide early warning system, meaning that many areas with high landslide risk lack this information (Guzzetti et al., 2020). For example, the small, coastal town of Sitka, Alaska, USA, experienced a fatal landslide event in 2015 with no early warning system in place (Busch et al., 2021). As a result, the local community and technical experts determined the need for a LEWS; Chapter 2 of this thesis describes the development of such a system.

Most rainfall-triggered landslides happen when rainfall infiltrating into a potentially unstable hillslope causes an increase in pore water pressure in the hillslope material that reduces shear strength and causes the slope to fail (Iverson, 2000). A slope's stability at

the moment of failure is thus a function of environmental factors, like slope angle and material properties, antecedent soil moisture conditions, and the intensity and duration of the triggering rainfall (Bogaard & Greco, 2016). Therefore, information on antecedent hillslope conditions and rainfall forecasts have been used with statistical models to create forecasts of weather-related landslide hazard for early warning purposes (e.g. Peruccacci et al., 2017; Thomas et al., 2018). In this thesis, I use the term "landslide hazard" to mean the probability of at least one landslide occurring over a given time period and spatial extent, whereas other literature definitions consider landslide size as well (Malamud et al., 2004).

## 1.2 RAINFALL THRESHOLDS FOR LANDSLIDE EARLY WARNING

Landslide early warning systems can be divided into two general types, local LEWS and geographical or territorial LEWS, and are intended to warn people during periods of high hazard (Guzzetti et al., 2020; Pecoraro et al., 2019; Piciullo et al., 2018). Local landslide early warning systems focus on a single hillslope or small catchment and monitor very local conditions with high temporal resolution on scales of seconds to minutes (Pecoraro et al., 2019). These systems are designed to warn of an imminent failure or an event already in progress; for example, a system installed in the Illgraben catchment, Switzerland activates flashing lights and sirens at footpaths crossing a debris flow channel five to 15 minutes before the debris flow arrives (Badoux et al., 2009). Geographical (also called territorial) early warning systems, in contrast, operate on spatial scales ranging from city, regional, national, to global, and may incorporate timescales ranging from minutes to days (Chae et al., 2017; Guzzetti et al., 2020; Piciullo et al., 2018). These warning systems often rely on hillslope hydrological monitoring, precipitation measurements, and weather forecasts to assess hazard. For example, in Rio de Janeiro, Brazil the Alerta-Rio system incorporates 15 minute rainfall measurements, Doppler radar, and 6-hour meteorological forecasts to issue warnings (Calvello et al., 2015; Guzzetti et al., 2020). NASA operates a global landslide nowcast that relies on satellite precipitation measurements over the previous 7 days to assess current landslide susceptibility at 30 minute intervals (Kirschbaum & Stanley, 2018).

Most geographical LEWS rely on thresholds to issue warnings (Guzzetti et al., 2020; Piciullo et al., 2018). Generally, thresholds aim to separate conditions under which landslides are more likely from those when they are less likely; when the metrics being monitored (e.g. pore pressure, cumulative rainfall, or rainfall intensity and duration) cross the threshold, a warning is issued. Approaches to identifying thresholds typically combine landslide inventory data with rainfall measurements, either from gauges, radar, or satellite data, and/or with soil moisture or pore pressure measurements to identify the conditions under which landslides have, or have not, occurred in the past (e.g. Kirschbaum & Stanley, 2018; Saito et al., 2010; Scheevel et al., 2017; Thomas et al., 2018; Wicki et al., 2021).

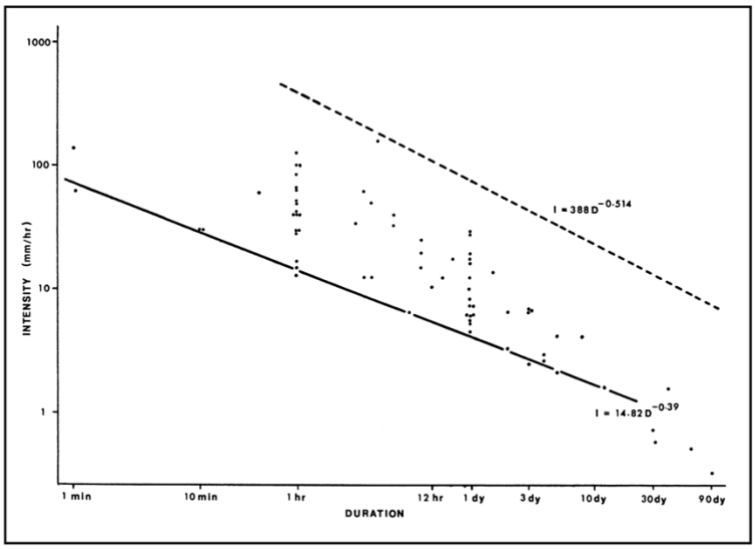



**Figure 1.2. The first global intensity-duration threshold for shallow landslides** (Caine, 1980). The solid line is the threshold; the dotted line approximates global maximum precipitation intensities.

However, previous studies have employed widely varying data types, metrics, and methods to identify such thresholds (Segoni et al., 2018). For example, national thresholds for Italy consider cumulative precipitation over a range of durations (Peruccacci et al., 2017), whereas thresholds for Seattle, Washington, USA rely on three day and 15 day precipitation totals (Scheevel et al., 2017). Since the first global rainfall intensity-duration (I-D) threshold for shallow landsliding was proposed in 1980 (Figure 1.2), thresholds have been widely investigated for different regions, but identified in differing ways (Caine, 1980; Guzzetti et al., 2008; Segoni et al., 2018). For instance, in Japan, quantile regression has been applied to learn thresholds from reported landslide triggering rainfall events from gauge corrected radar measurements (Saito et al., 2010), whereas in California, USA, the ratio of false alarms (rainfall above the threshold that did not trigger a landslide) to missed alarms (rainfall below the threshold that did trigger a landslide) was optimized to establish thresholds for post-wildfire debris flows from station-based observations (Staley et al., 2013).

This plethora of different thresholds and methods likely stems from the common observation that thresholds vary widely between regions, driven by differences in topography, lithology, vegetation type, soil properties, or other factors affecting slope stability, thus necessitating locally determined thresholds (Crosta, 1998; Guzzetti et al., 2008; Segoni et al., 2014). While it is clear that approaches vary widely, few investigations have objectively tested if and by how much thresholds vary between regions using consistent methods and datasets. For example, Guzzetti et al., 2008 compiled landslide triggering rainfall observations from different studies globally, finding that thresholds vary by climate zone, however, the original studies defined event rainfall in differing ways, complicating comparison. Moreover, very few studies estimate or report any type of threshold uncertainty. It therefore remains unclear if differences between regions are statistically significant. For instance, Segoni et al., 2014

estimated I-D thresholds for 25 neighboring sub-regions in Tuscany, Italy, reporting that thresholds varied between areas with differing mean annual precipitation, lithology, and topography and that such heavily localized thresholds outperformed thresholds that cover the whole region. However, with no threshold uncertainty having been documented, it is difficult to evaluate if these thresholds are credibly distinguishable, or whether improved performance is a product of overfitting.

This poses a challenge for towns, cities, or regions seeking to establish a LEWS, but with limited landslide inventory data, either because this data has not been collected, or because few landslides have occurred. It remains unclear to what extent such communities can rely on data or thresholds from neighboring or other regions for warning. Thus, a key research gap is to quantify threshold variability and uncertainty with a consistent methodology; Chapter 3 contributes to filling this gap by estimating variation between thresholds in 26 major metropolitan areas worldwide.

## 1.3 PROBABILISTIC MODELS FOR PREDICTING LANDSLIDE RESPONSE TO RAINFALL CONDITIONS WITH INVENTORY DATA

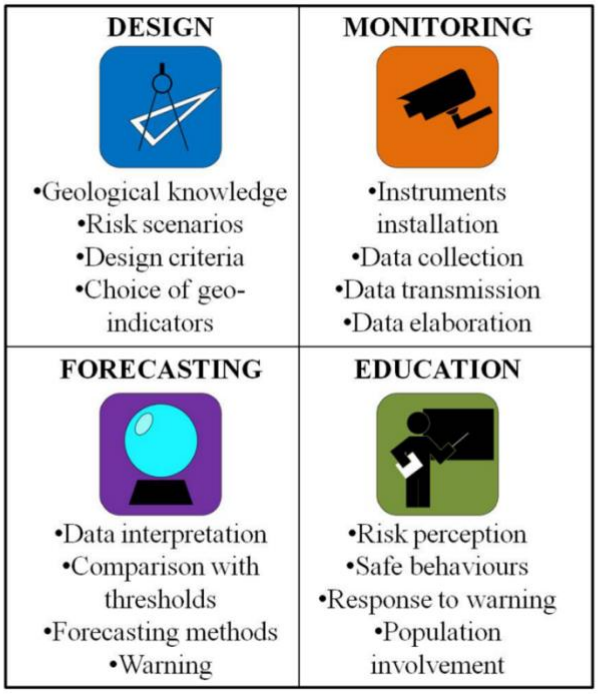



**Figure 1.3. Four fundamental components of a landslide early warning system.** Figure from (Intrieri et al., 2013).

One important element of a LEWS is the ability to forecast landslide activity, which requires models (Figure 1.3) (Intrieri et al., 2013). Thresholds are a simple model that express a binary representation of landslide occurrence. For a forecast rainfall event with an intensity and duration that exceeds an I-D threshold, for example, landslides

have occurred in the past, so under the assumption that the past is a sufficient representation of the future, we infer that landslides could occur during this event. Conversely, for a rainfall event below the threshold, no landslides have been documented in the past, so we infer that no landslides will occur now. However, in reality, landslides at times occur and do not occur under identical monitoring conditions in the same region, inevitably leading to some false alarms and some missed alarms (Conrad et al., 2021; Peres & Cancelliere, 2021). Probabilistic approaches to quantifying landslide hazard and establishing warning thresholds can be preferable, as they provide additional information beyond a binary output (Berti et al., 2012). While under-utilized in comparison to simpler threshold approaches, probabilistic statistical models can answer questions relevant for warning that binary thresholds are unable to address, such as:

1. How probable is a landslide given forecast rainfall conditions?

2. How many landslides can be expected during a given event?

3. How certain are our predictions, given the available data?

There is often confusion between the terms "prediction" and "forecast" in the landslides literature. Here, I use "prediction" to mean an estimate about an unobserved event, regardless of whether that event happens in the past, present, or future. Landslide forecasts typically refer to future predictions of landslide activity conditional on a rainfall forecast, whereas the term "nowcast" is also used to denote a prediction based on current or recent conditions (Stanley et al., 2021).

Landslide inventory data is a key resource to train statistical models for temporal landslide prediction. Recently, efforts to document landslide occurrences by, for example, collating news reports, mapping landslides after storms with LiDAR and aerial imagery, and collecting official records like highway department reports, have increased the availability of data with reported landslide timing in some areas (Froude & Petley, 2018; Kirschbaum et al., 2015) (Figure 1.4).

While valuable, these datasets are rarely complete, posing additional challenges (Steger et al., 2017). For example, inventories that are created by collecting news reports tend to have daily timestamps but are typically limited to road networks and urban areas where newsworthy landslides occur. These inventories may therefore under-report the number of landslides that occurred during a given time period. On the other hand, inventories created by mapping from satellite imagery, LiDAR, or aerial photos, tend to have more complete spatial coverage, but may only report landslide timing with a resolution of months to years, cover only single storm events, or report no temporal information whatsoever (Kirschbaum et al., 2015; Luna & Korup, 2022; Mirus et al., 2020).

In some well-documented regions, this results in a collection of heterogeneous datasets that all contain valuable temporal information about landslide activity in a given area, but which so far, have mostly been analyzed separately (Marc et al., 2018; Stanley et al.,

2020). In the Pacific Northwest region of the USA, for example, at least five different landslide inventories exist that cover different areas, time periods, and have been created in different ways. Landslides with temporal information in the Washington Landslide Compilation are linked to specific storm events, whereas the City of Seattle's inventory contains a time series of landslide reports dating back to the 1890s (City of Seattle, 2020; Washington Geological Survey, 2020). So far, most studies have analyzed these different types of temporal information separately: for example, some studies have investigated triggering conditions and spatial patterns of landsliding during individual storm events, but have not considered time series of reported landslides (Marc et al., 2018). However, the availability of data that captures these different aspects of landslide timing calls for modeling approaches that are able to combine these heterogeneous inventories to make best use of the available data. Chapter 4 explores a modeling approach that combines different inventories from the Pacific Northwest of the USA.

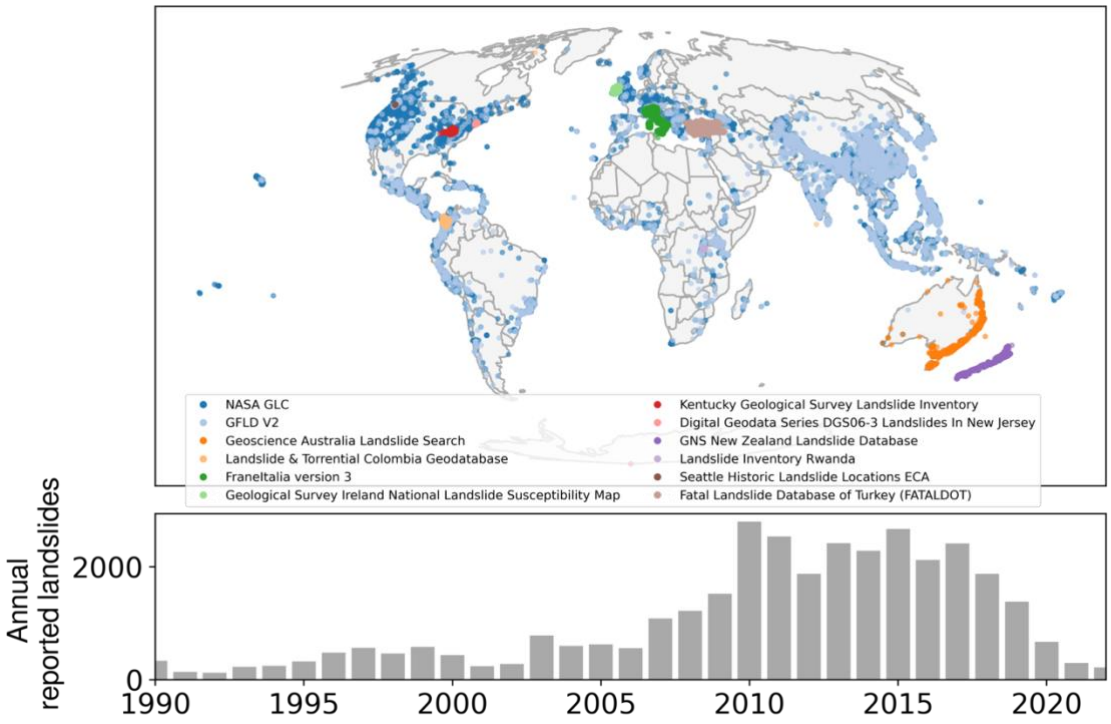



**Figure 1.4. Examples of openly available landslide inventories with daily timestamps.** Landslide data with temporal information has become increasingly available since ~2010.

In other regions, however, landslide inventory data is scarce or non-existent (Peres & Cancelliere, 2021). In Sitka, Alaska, for instance, only five landslide events were reported over the period from 2002 to 2020. By some methods, this would be far too few landslides to establish an early warning system; Peruccacci et al., 2017 reported that at least 75 landslides were needed to achieve stable parameter estimates with their models. Nevertheless, information on temporally varying landslide hazard is needed, requiring methods that are able to provide as much information as possible using the available data, while also quantifying the uncertainty that arises from having few data points. Chapter 2 applies probabilistic models to estimate landslide hazard with few reported landslide events.

In this context, Bayesian statistics offers an approach to analyzing data that has some advantages for predicting landslide activity. All Bayesian inference is based on Bayes' Rule, first described in a letter by Thomas Bayes in 1763 (Bayes, 1763), and subsequently formulated by Pierre Simon Laplace in 1825 (Laplace, 1825). Bayes' Rule is:

$$p(\theta \mid y) = \frac{p(y \mid \theta)p(\theta)}{p(y)}$$

*Equation 1.1*

where $p(\theta \mid y)$ is the probability of a set of model parameters $\theta$ conditional on the data y, and is called the posterior. $p(y \mid \theta)$ is the probability of the data conditional on the model parameters, or the likelihood, and $p(\theta)$ is the prior distribution of the parameters, which can be interpreted to represent our belief about the distribution of these parameters before seeing y. Bayesian inference thus updates prior knowledge about statistical model parameters with new data and can be used for statistical inference or for making predictions about unobserved events, past, current, or future (McElreath, 2020; van de Schoot et al., 2021). The posterior distribution provides not a single estimate of the "best" parameters for our model, but rather a probability distribution of all parameters that are compatible with the data and our prior knowledge. Using the full posterior distribution to make predictions of future events seamlessly propagates all uncertainty learned from the data into the predictions. Additionally, Bayesian inference allows us to incorporate prior knowledge, which can improve estimates for regions with limited inventory data. In western Canada, for example, (Nolde & Joe, 2013) used Bayesian inference to incorporate expert knowledge to reduce uncertainty of debris flow return level estimates based on a small inventory of only thirteen debris flows. Despite these advantages, and although Bayesian statistics have gained popularity in fields like ecology, epidemiology, economics, psychology, and hydrology (Renard et al., 2013; van de Schoot et al., 2021) in the past years, uptake in landslide research has been limited, with few exceptions (Berti et al., 2012; Lombardo et al., 2018, 2020; Nolde & Joe, 2013). In Chapters 2, 3, and 4, I apply Bayesian inference to estimate and predict landslide activity at a range of spatial and temporal scales.

## 1.4 Toward seasonal forecasts: expanding time frames for landslide warning

Most research on landslide early warning has focused on hourly to weekly timescales (e.g. Conrad et al., 2021; Peruccacci et al., 2017; Saito et al., 2010; Scheevel et al., 2017; Thomas et al., 2018). However, better understanding and predicting seasonal variations in landslide activity could help lead to seasonal landslide forecasts, potentially providing information to better plan and allocate resources for landslide response with several months lead time (Steger et al., 2022). So far, few studies globally have attempted to predict seasonal landslide activity as a function of monthly rainfall, although many regions of the world show a strong landslide seasonality. For example, in Central

America and the Caribbean, the number of reported landslides peaks during the fall Atlantic hurricane season (Sepúlveda & Petley, 2015), whereas the corresponding peak in Japan happens during the East Asian Summer Monsoon (Saito et al., 2010). Nevertheless, landslide studies mentioning seasonal patterns often simply plot histograms of monthly landslide counts, which are unable to predict, for example, the inter-annual variability in landslide activity, or landslide response to monthly rainfall patterns. Achieving seasonal forecasts requires steps to predict seasonal landslide activity based on monthly precipitation.

The Pacific Northwest region of the USA is among the best studied regions impacted by landsliding globally (e.g. Godt et al., 2006; LaHusen et al., 2020; Montgomery & Dietrich, 1994), and it is common knowledge among landslide researchers, emergency responders, and the public that landslides mostly occur during the wet winter season. This strong seasonality, combined with a multitude of available landslide inventory data, provides an ideal case to establish and test predictive models of seasonal landslide activity in Chapter 4.

## 1.5 OBJECTIVES AND RESEARCH QUESTIONS

The overarching motivation for this dissertation is to further our ability to predict rainfall-triggered landslides in time in order to improve warning. In pursuit of this goal, I apply Bayesian statistical models to landslide inventory and precipitation data to quantify landslide triggering rainfall conditions and predict landslide activity at spatial scales from a small town in Alaska, to large metropolitan areas globally, to a multi-state region, and at temporal scales from hourly to seasonal. I seek to fill existing research gaps by providing functional landslide warning for a small town with limited landslide inventory data, quantifying threshold uncertainty and variability between cities with a consistent methodology, exploring statistical methods to combine available, yet heterogeneous landslide inventories, and laying the groundwork for seasonal landslide forecasts.

This dissertation focuses on three related research questions, which I address in three chapters. These chapters are presented in order of increasing spatial area targeted: small town, major metropolitan area, and multi-state region (Figure 1.5).

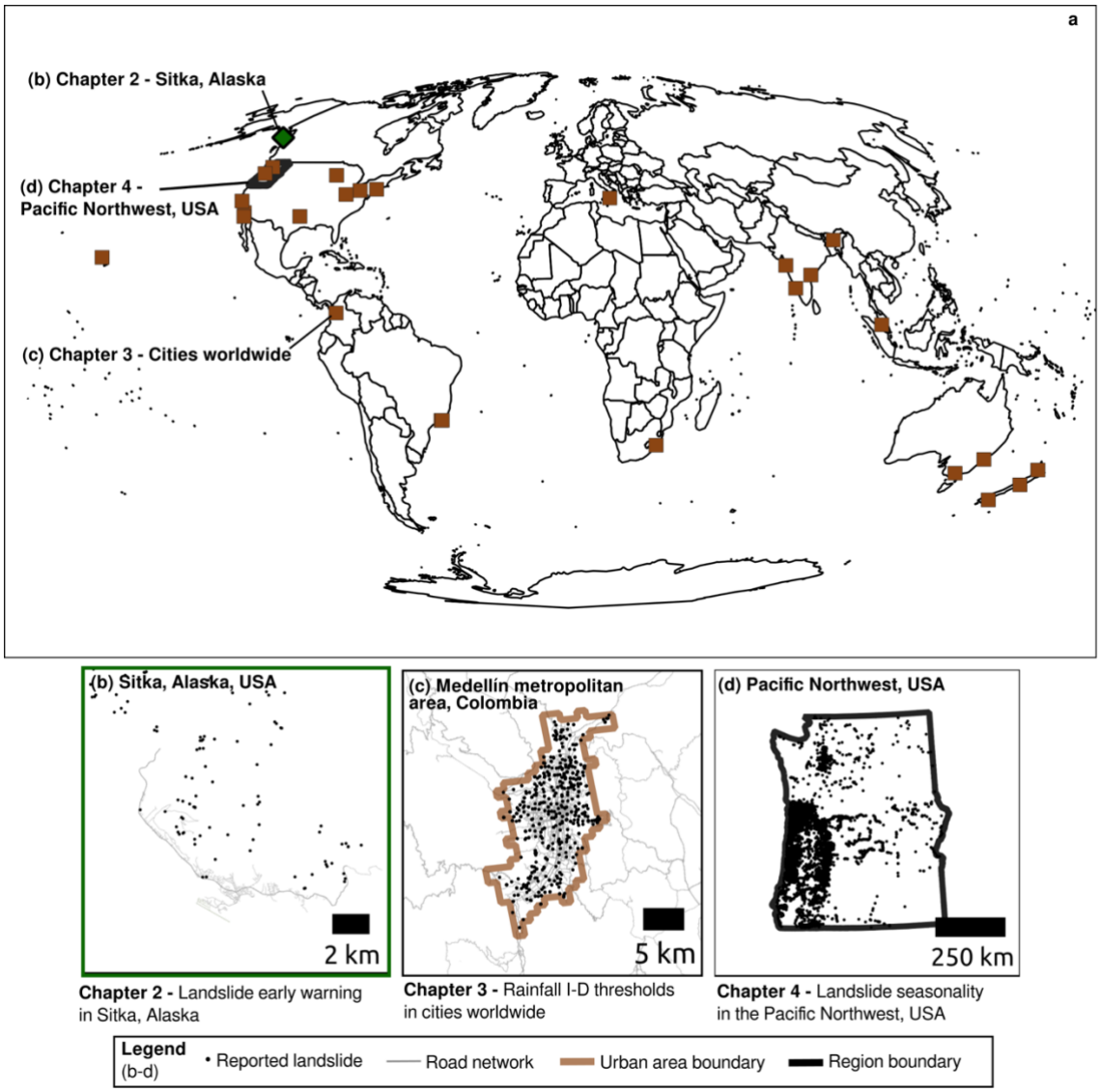



**Figure 1.5. Study areas and the structure of this thesis**. The chapters in this thesis are presented in order of increasing spatial area targeted (b – d). (a) Global overview of study areas. Brown squares are cities for which I estimate I-D thresholds in Chapter 3.

**1. Under which rainfall conditions should we warn for shallow landslides in Sitka, Alaska, USA?** *(Chapter 2)*

In Chapter 2, I contribute to developing and validating statistical models for an online landslide warning dashboard for the small town of Sitka, Alaska, USA. Our objectives are to estimate daily landslide probability and identify warning thresholds. We use an inventory of only five reported landslide events and 18 years of hourly precipitation measurements at the Sitka airport to train the models and establish two warning thresholds drawing on community input.

**2. How do intensity-duration thresholds for urban landslides vary between cities worldwide?** *(Chapter 3)*

In Chapter 3, I estimate rainfall intensity-duration thresholds for 26 cities worldwide and a global threshold for urban areas. My objectives are to update a prior global threshold with data from urban areas, quantify the variability between thresholds in different cities using consistent precipitation datasets and methodology, and assess potential drivers of threshold variation between cities.

**3. What is the relationship between reported landslide activity and seasonal rainfall patterns in the Pacific Northwest, USA?** *(Chapter 4)*

In Chapter 4, I investigate the seasonal pattern of landslide activity in the Pacific Northwest region, USA. My objectives are to test if statistical models objectively reveal seasonal variations in landslide activity and to test if landslide response to precipitation changes seasonally. I explore Bayesian multi-level models as a method to combine data from heterogeneous inventories to predict landslide activity at seasonal timescales.

## 1.6 REFERENCES


Alcântara, E., Marengo, J. A., Mantovani, J., Londe, L., San, R. L. Y., Park, E., et al. (2022). Deadly disasters in Southeastern South America: Flash floods and landslides of February 2022 in Petropolis, Rio de Janeiro. *Natural Hazards and Earth System Sciences Discussions*, 1–27. https://doi.org/10.5194/nhess-2022-163

Badoux, A., Graf, C., Rhyner, J., Kuntner, R., & McArdell, B. W. (2009). A debris-flow alarm system for the Alpine Illgraben catchment: design and performance. *Natural Hazards*, *49*(3), 517–539. https://doi.org/10.1007/s11069-008-9303-x

Bayes, T. (1763). LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F. R. S. communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S. *Philosophical Transactions of the Royal Society of London*, *53*, 370–418. https://doi.org/10.1098/rstl.1763.0053

Berti, M., Martina, M. L. V., Franceschini, S., Pignone, S., Simoni, A., & Pizziolo, M. (2012). Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach. *Journal of Geophysical Research: Earth Surface*, *117*(F4). https://doi.org/10.1029/2012JF002367

Bogaard, T., & Greco, R. (2016). Landslide hydrology: from hydrology to pore pressure. *WIREs Water*, *3*(3), 439–459. https://doi.org/10.1002/wat2.1126

Busch, L., Lempert, R., Izenberg, M., & Patton, A. (2021). Run Uphill for a Tsunami, Downhill for a Landslide. *Issues in Science and Technology*, *38*(1), 40–46.

Caine, N. (1980). The Rainfall Intensity - Duration Control of Shallow Landslides and Debris Flows. *Geografiska Annaler: Series A, Physical Geography*. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/04353676.1980.11879996

Calvello, M., d'Orsi, R. N., Piciullo, L., Paes, N., Magalhaes, M., & Lacerda, W. A. (2015). The Rio de Janeiro early warning system for rainfall-induced landslides: Analysis of performance for the years 2010–2013. *International Journal of Disaster Risk Reduction*, *12*, 3–15. https://doi.org/10.1016/j.ijdrr.2014.10.005

Chae, B.-G., Park, H.-J., Catani, F., Simoni, A., & Berti, M. (2017). Landslide prediction, monitoring and early warning: a concise review of state-of-the-art. *Geosciences Journal*, *21*(6), 1033–1070. https://doi.org/10.1007/s12303-017-0034-4

Choi, K. Y., & Cheung, R. W. M. (2013). Landslide disaster prevention and mitigation through works in Hong Kong. *Journal of Rock Mechanics and Geotechnical Engineering*, *5*(5), 354–365. https://doi.org/10.1016/j.jrmge.2013.07.007

City of Seattle. (2020). Historical Landslide Locations ECA. City of Seattle. Retrieved from https://data-seattlecitygis.opendata.arcgis.com/datasets/6ac72973a5784d90bda0a5f8a001d9f3_22

Conrad, J. L., Morphew, M. D., Baum, R. L., & Mirus, B. B. (2021). HydroMet: A New Code for Automated Objective Optimization of Hydrometeorological Thresholds for Landslide Initiation. *Water*, *13*(13), 1752. https://doi.org/10.3390/w13131752

Crosta, G. (1998). Regionalization of rainfall thresholds: an aid to landslide hazard evaluation. *Environmental Geology*, *35*(2–3), 131–145. https://doi.org/10.1007/s002540050300

Crovelli, R. A., & Coe, J. A. (2009). Probabilistic estimation of numbers and costs of future landslides in the San Francisco Bay region. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, *3*(4), 206–223. https://doi.org/10.1080/17499510802713123

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences*, *18*(8), 2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

Godt, J. W., Baum, R. L., & Chleborad, A. F. (2006). Rainfall characteristics for shallow landsliding in Seattle, Washington, USA. *Earth Surface Processes and Landforms*, *31*(1), 97–110. https://doi.org/10.1002/esp.1237

Guzzetti, F., Peruccacci, S., Rossi, M., & Stark, C. P. (2008). The rainfall intensity–duration control of shallow landslides and debris flows: an update. *Landslides*, *5*(1), 3–17. https://doi.org/10.1007/s10346-007-0112-1

Guzzetti, F., Gariano, S. L., Peruccacci, S., Brunetti, M. T., Marchesini, I., Rossi, M., & Melillo, M. (2020). Geographical landslide early warning systems. *Earth-Science Reviews*, *200*, 102973. https://doi.org/10.1016/j.earscirev.2019.102973

Haque, U., da Silva, P. F., Devoli, G., Pilz, J., Zhao, B., Khaloua, A., et al. (2019). The human cost of global warming: Deadly landslides and their triggers (1995–2014). *Science of The Total Environment*, *682*, 673–684. https://doi.org/10.1016/j.scitotenv.2019.03.415

Intrieri, E., Gigli, G., Casagli, N., & Nadim, F. (2013). Brief communication: Landslide Early Warning System: toolbox and general concepts. *Natural Hazards and Earth System Sciences*, *13*(1), 85–90. https://doi.org/10.5194/nhess-13-85-2013

Iverson, R. M. (2000). Landslide triggering by rain infiltration. *Water Resources Research*, *36*(7), 1897–1910. https://doi.org/10.1029/2000WR900090

Kirschbaum, D., & Stanley, T. (2018). Satellite-Based Assessment of Rainfall-Triggered Landslide Hazard for Situational Awareness. *Earth's Future*, *6*(3), 505–523. https://doi.org/10.1002/2017EF000715

Kirschbaum, D., Stanley, T., & Zhou, Y. (2015). Spatial and temporal analysis of a global landslide catalog. *Geomorphology*, *249*, 4–15. https://doi.org/10.1016/j.geomorph.2015.03.016

Klose, M., Maurischat, P., & Damm, B. (2016). Landslide impacts in Germany: A historical and socioeconomic perspective. *Landslides*, *13*(1), 183–199. https://doi.org/10.1007/s10346-015-0643-9

Kockelman, W. (1986). Some Techniques for Reducing Landslide Hazards. *Environmental & Engineering Geoscience*, *xxiii*(1), 29–52. https://doi.org/10.2113/gseegeosci.xxiii.1.29

LaHusen, S. R., Duvall, A. R., Booth, A. M., Grant, A., Mishkin, B. A., Montgomery, D. R., et al. (2020). Rainfall triggers more deep-seated landslides than Cascadia earthquakes in the Oregon Coast Range, USA. *Science Advances*, *6*(38), eaba6790. https://doi.org/10.1126/sciadv.aba6790

Laplace, P.-S. (1825). *Pierre-Simon Laplace Philosophical Essay on Probabilities: Translated from the fifth French edition of 1825 With Notes by the Translator*. Springer Science & Business Media.

Larsen, M. C. (2008). Rainfall-triggered landslides, anthropogenic hazards, and mitigation strategies. In *Advances in Geosciences* (Vol. 14, pp. 147–153). Copernicus GmbH. https://doi.org/10.5194/adgeo-14-147-2008

Lombardo, L., Opitz, T., & Huser, R. (2018). Point process-based modeling of multiple debris flow landslides using INLA: an application to the 2009 Messina disaster. *Stochastic Environmental Research and Risk Assessment*, *32*(7), 2179–2198. https://doi.org/10.1007/s00477-018-1518-0

Lombardo, L., Opitz, T., Ardizzone, F., Guzzetti, F., & Huser, R. (2020). Space-time landslide predictive modelling. *Earth-Science Reviews*, *209*, 103318. https://doi.org/10.1016/j.earscirev.2020.103318

Luna, L. V., & Korup, O. (2022). Seasonal Landslide Activity Lags Annual Precipitation Pattern in the Pacific Northwest. *Geophysical Research Letters*, *49*(18), e2022GL098506. https://doi.org/10.1029/2022GL098506

Malamud, B. D., Turcotte, D. L., Guzzetti, F., & Reichenbach, P. (2004). Landslide inventories and their statistical properties. *Earth Surface Processes and Landforms*, *29*(6), 687–711. https://doi.org/10.1002/esp.1064

Marc, O., Stumpf, A., Malet, J.-P., Gosset, M., Uchida, T., & Chiang, S.-H. (2018). Initial insights from a global database of rainfall-induced landslide inventories: the weak influence of slope and strong influence of total storm rainfall. *Earth Surface Dynamics*, *6*(4), 903–922. https://doi.org/10.5194/esurf-6-903-2018

McElreath, R. (2020). *Statistical Rethinking A Bayesian Course with Examples in R and STAN* (2nd Edition). Chapman and Hall/CRC Press. Retrieved from https://www.routledge.com/Statistical-Rethinking-A-Bayesian-Course-with-Examples-in-R-and-STAN/McElreath/p/book/9780367139919

Mirus, B. B., Jones, E. S., Baum, R. L., Godt, J. W., Slaughter, S., Crawford, M. M., et al. (2020). Landslides across the USA: occurrence, susceptibility, and data limitations. *Landslides*, *17*(10), 2271–2285. https://doi.org/10.1007/s10346-020-01424-4

Montgomery, D. R., & Dietrich, W. E. (1994). A physically based model for the topographic control on shallow landsliding. *Water Resources Research*, *30*(4), 1153–1171. https://doi.org/10.1029/93WR02979

Nolde, N., & Joe, H. (2013). A Bayesian extreme value analysis of debris flows. *Water Resources Research*, *49*(10), 7009–7022. https://doi.org/10.1002/wrcr.20494

Pecoraro, G., Calvello, M., & Piciullo, L. (2019). Monitoring strategies for local landslide early warning systems. *Landslides*, *16*(2), 213–231. https://doi.org/10.1007/s10346-018-1068-z

Peres, D. J., & Cancelliere, A. (2021). Comparing methods for determining landslide early warning thresholds: potential use of non-triggering rainfall for locations with scarce landslide data availability. *Landslides*, *18*(9), 3135–3147. https://doi.org/10.1007/s10346-021-01704-7

Peruccacci, S., Brunetti, M. T., Gariano, S. L., Melillo, M., Rossi, M., & Guzzetti, F. (2017). Rainfall thresholds for possible landslide occurrence in Italy. *Geomorphology*, *290*, 39–57. https://doi.org/10.1016/j.geomorph.2017.03.031

Piciullo, L., Calvello, M., & Cepeda, J. M. (2018). Territorial early warning systems for rainfall-induced landslides. *Earth-Science Reviews*, *179*, 228–247. https://doi.org/10.1016/j.earscirev.2018.02.013

Pollock, W., & Wartman, J. (2020). Human Vulnerability to Landslides. *GeoHealth*, *4*(10), e2020GH000287. https://doi.org/10.1029/2020GH000287

Postance, B., Hillier, J., Dijkstra, T., & Dixon, N. (2017). Extending natural hazard impacts: an assessment of landslide disruptions on a national road transportation network. *Environmental Research Letters*, *12*(1), 014010. https://doi.org/10.1088/1748-9326/aa5555

ReliefWeb. (2022). *South Africa: Floods and Landslides - Apr 2022 | ReliefWeb*. Retrieved from https://reliefweb.int/disaster/fl-2022-000201-zaf

Renard, B., Sun, X., & Lang, M. (2013). Bayesian Methods for Non-stationary Extreme Value Analysis. In A. AghaKouchak, D. Easterling, K. Hsu, S. Schubert, & S. Sorooshian (Eds.), *Extremes in a Changing Climate: Detection, Analysis and Uncertainty* (pp. 39–95). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-4479-0_3

Saito, H., Nakayama, D., & Matsuyama, H. (2010). Relationship between the initiation of a shallow landslide and rainfall intensity—duration thresholds in Japan. *Geomorphology*, *118*(1), 167–175. https://doi.org/10.1016/j.geomorph.2009.12.016

Scheevel, C. R., Baum, R. L., Mirus, B. B., & Smith, J. B. (2017). *Precipitation thresholds for landslide occurrence near Seattle, Mukilteo, and Everett, Washington* (USGS Numbered Series No. 2017–1039). *Precipitation thresholds for landslide occurrence near Seattle, Mukilteo, and Everett, Washington* (Vol. 2017–1039, p. 60). Reston, VA: U.S. Geological Survey. https://doi.org/10.3133/ofr20171039

van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., et al. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers*, *1*(1), 1–26. https://doi.org/10.1038/s43586-020-00001-2

Segoni, S., Rosi, A., Rossi, G., Catani, F., & Casagli, N. (2014). Analysing the relationship between rainfalls and landslides to define a mosaic of triggering thresholds for regional-scale warning systems. *Natural Hazards and Earth System Sciences*, *14*(9), 2637–2648. https://doi.org/10.5194/nhess-14-2637-2014

Segoni, S., Piciullo, L., & Gariano, S. L. (2018). A review of the recent literature on rainfall thresholds for landslide occurrence. *Landslides*, *15*(8), 1483–1501. https://doi.org/10.1007/s10346-018-0966-4

Sepúlveda, S. A., & Petley, D. N. (2015). Regional trends and controlling factors of fatal landslides in Latin America and the Caribbean. *Natural Hazards and Earth System Sciences*, *15*(8), 1821–1833. https://doi.org/10.5194/nhess-15-1821-2015

Staley, D. M., Kean, J. W., Cannon, S. H., Schmidt, K. M., & Laber, J. L. (2013). Objective definition of rainfall intensity–duration thresholds for the initiation of post-fire debris flows in southern California. *Landslides*, *10*(5), 547–562. https://doi.org/10.1007/s10346-012-0341-9

Stanley, T. A., Kirschbaum, D. B., Sobieszczyk, S., Jasinski, M. F., Borak, J. S., & Slaughter, S. L. (2020). Building a landslide hazard indicator with machine learning and land surface models. *Environmental Modelling & Software*, *129*, 104692. https://doi.org/10.1016/j.envsoft.2020.104692

Stanley, T. A., Kirschbaum, D. B., Benz, G., Emberson, R. A., Amatya, P. M., Medwedeff, W., & Clark, M. K. (2021). Data-Driven Landslide Nowcasting at the Global Scale. *Frontiers in Earth Science*, *9*. Retrieved from https://www.frontiersin.org/articles/10.3389/feart.2021.640043

Steger, S., Brenning, A., Bell, R., & Glade, T. (2017). The influence of systematically incomplete shallow landslide inventories on statistical susceptibility models and suggestions for improvements. *Landslides*, *14*(5), 1767–1781. https://doi.org/10.1007/s10346-017-0820-0

Steger, S., Moreno, M., Crespi, A., Zellner, P. J., Gariano, S. L., Brunetti, M. T., et al. (2022). Deciphering seasonal effects of triggering and preparatory precipitation for improved shallow landslide prediction using generalized additive mixed models. *Natural Hazards and Earth System Sciences Discussions*, 1–38. https://doi.org/10.5194/nhess-2022-271

Thomas, M. A., Mirus, B. B., & Collins, B. D. (2018). Identifying Physics-Based Thresholds for Rainfall-Induced Landsliding. *Geophysical Research Letters*, *45*(18), 9651–9661. https://doi.org/10.1029/2018GL079662

Washington Geological Survey. (2020). Landslides compilation--GIS data, February 2020: Washington Geological Survey Digital Data Series 12, version 5.2. Retrieved from https://fortress.wa.gov/dnr/geologydata/publications/data_download/ger_portal_landslide_compilation.zip

Wicki, A., Jansson, P.-E., Lehmann, P., Hauck, C., & Stähli, M. (2021). Simulated or measured soil moisture: which one is adding more value to regional landslide early warning? *Hydrology and Earth System Sciences*, *25*(8), 4585–4610. https://doi.org/10.5194/hess-25-4585-2021

# 2 Landslide initiation thresholds in data sparse regions: Application to landslide early warning criteria in Sitka, Alaska, USA[1]


## Abstract

Probabilistic models to inform landslide early warning systems often rely on rainfall totals observed during past events with landslides. However, these models are generally developed for broad regions using large catalogs, with dozens, hundreds, or even thousands of landslide occurrences. This study evaluates strategies for training landslide forecasting models with a scanty record of landslide-triggering events, which is a typical limitation in remote, sparsely populated regions. We train and evaluate 136 statistical models with a rainfall dataset with five landslide-triggering rainfall events recorded near Sitka, Alaska, USA, as well as >6,000 days of non-triggering rainfall (2002–2020). We use Akaike, Bayesian, and leave-one-out information criteria to compare models trained on cumulative precipitation at timescales ranging from 1 hour to 2 weeks, using both frequentist and Bayesian methods to estimate the daily probability and intensity of potential landslide occurrence (logistic regression and Poisson regression). We evaluate the best-fit models using leave-one-out validation as well as with testing a subset of the data. Despite this sparse landslide inventory, we find that probabilistic models can effectively distinguish days with landslides from days without. Although frequentist and Bayesian inference produce similar estimates of landslide hazard, they do have different implications for use and interpretation: frequentist models are familiar and easy to implement, but Bayesian models capture the rare-events problem more explicitly


[1] Patton, A. I., Luna, L. V., Roering, J. J., Jacobs, A., Korup, O., and Mirus, B. B.: Landslide initiation thresholds in data sparse regions: Application to landslide early warning criteria in Sitka, Alaska, USA, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2023-25, 2023. CC-BY-4.0. *Under consideration at NHESS.*

and allow for better understanding of parameter uncertainty given the available data. Three-hour precipitation totals are the best predictor of elevated landslide hazard, and adding antecedent precipitation (days to weeks) did not improve model performance. This relatively short timescale combined with the limited role of antecedent conditions reflects the rapid draining of porous colluvial soils on very steep hillslopes around Sitka. We use the resulting estimates of daily landslide probability to establish two decision boundaries for three levels of warning. With these decision boundaries, the frequentist logistic regression model incorporates National Weather Service quantitative precipitation forecasts into a real-time landslide early warning "dashboard" system (sitkalandslide.org). This dashboard provides accessible and data-driven situational awareness for community members and emergency managers.

## 2.1 INTRODUCTION

On August 18, 2015, an extreme rain event initiated more than 40 landslides on the islands near Sitka, Alaska, USA, including a debris flow that resulted in three fatalities (Busch et al., 2021). Over a six-hour period, the Sitka area received 2.5–3.25 inches of rain, and the three-hour storm totals had an estimated 45-year return period. Following this event, the community convened a GeoTask Force to identify priority questions related to landslide risk and hazard (Sitka Sound Science Center, 2016). Community leaders and technical experts determined the need for a landslide early warning system. This study results from the actions of the community to seek support to reduce landslide risks.

Landslide early warning has the potential to save lives by providing actionable information in advance of an imminent landslide event (e.g., Guzzetti et al., 2020). Landslide early warning systems consist of a prediction ("now-cast" or "forecast") of landslide occurrence, one or more thresholds for action, and a method for disseminating warning information. Decades of investigation around the world have demonstrated the value of using precipitation and hydrologic conditions to predict landslides (e.g., Chae et al., 2017; Guzzetti et al., 2020), but prediction strategies vary. Most studies determine decision thresholds that aim to separate periods when landslides are likely from periods when they are not. These thresholds may be based on precipitation intensity and duration, consider cumulative precipitation over different time periods (Guzzetti et al., 2008; Bogaard and Greco, 2018; Mirus et al., 2018b), and/or incorporate in situ hydrologic data or estimates of antecedent hillslope hydrological conditions (Mirus et al., 2018a; Thomas et al., 2018; Marino et al., 2020; Wicki et al., 2020). Thresholds may indicate the minimum accumulation of precipitation needed to initiate landslides or attempt to optimally separate triggering from non-triggering precipitation events (Segoni et al., 2018).

Accurately predicting rare events like landslides is challenging because the complex and spatially heterogenous processes that drive landslide initiation are difficult to characterize at sufficiently high resolution across broad regions. In this study, instead of

trying to predict the spatial occurrence, we focus on predicting the temporal occurrence of landslides (when and how many failures) within a given study area.

Both empirically and physically based hazard assessments and warning systems require sufficient in situ data to be developed, calibrated, and validated. For example, lack of high-resolution imagery and in situ measurements of parameters such as soil bulk density, thickness, and hydraulic properties hinders the development of physically based models. Detailed precipitation and hydrologic records with high temporal resolution (hourly or finer) rarely cover long timescales (years to decades). Additionally, remote, sparsely populated areas typically lack inventories of landslide occurrence. These limited datasets of landslide occurrence and associated triggering conditions make it challenging to develop empirical models for landslide initiation, which may have large uncertainties, are often difficult to validate, and cause detrimental false positives in early warning systems. Yet, vulnerability to landslides is often high in remote and data-sparse regions due to limited infrastructure and access to external aid (Cutter and Finch, 2008). Improving prediction of landslide hazards in remote regions is a critical step to supporting resilient communities.

In this study, we developed a landslide early warning system for the remote community of Sitka, Alaska, USA, (Fig. 2.1), which had a population of 8,407 in 2021 (U.S. Census Bureau, 2021). We trained statistical models with limited landslide inventory data to estimate landslide probability and the number of landslides in the study area based on observed and forecasted precipitation, and then used these models to establish thresholds for landslide early warning. We use the term "landslide prediction" to refer to estimates of elevated landslide hazard in the future based on forecasted precipitation.

### 2.1.1 Study area: Sitka, southeast Alaska

Landslides in southeast Alaska pose persistent hazards to the small, isolated communities that are on the flanks of hillslopes over-steepened by glaciers. The majority of failures are debris flows initiated by shallow landslides (Swanston and Marion, 1991; Johnson et al., 2000). Steep hillslopes with thin volcanic soils overlying till are especially susceptible to shallow-seated landslides (Swanston, 1970; Sidle and Swanston, 1981; Patton et al., 2022). Following the fatal debris flow event in Sitka on August 18, 2015 (Busch et al., 2021), community organizers identified the need to better understand both where and when landslides are likely to occur in Sitka.

The landscape surrounding Sitka (Fig. 2.1) is geomorphically complex (Patton et al., 2022), having been sculpted by tectonic activity (White et al., 2016; Elliott and Freymueller, 2020), Pleistocene glaciation (Hamilton, 1986; Mann, 1986), volcanic eruptions (Riehle et al., 1992b, 1992a), and a long history of human settlement (Sandberg, 2013; Lesnek et al., 2018).

The mid-latitude maritime climate in Sitka is characterized by high annual precipitation. During the last climatic normal (1981–2010), mean annual precipitation at sea level was

2205 mm (Wendler et al., 2016), but steep orographic gradients and complex topography result in spatially heterogenous climate and weather patterns. Mean monthly temperatures stay above freezing all year. Variable snowpacks accumulate in winter months, particularly at high elevations, but most precipitation occurs as non-freezing rain in coastal and low-elevation areas. Rainfall occurs year-round in southeast Alaska, but August-November are the wettest months.

In particular, atmospheric rivers (ARs) account for 90% of extreme precipitation in southeast Alaska, where "extreme" precipitation was statistically defined using the block maxima approach by identifying one extreme event per year and per season (Sharma and Déry, 2019; Sharma and Dery, 2020). The AR contribution to extreme precipitation is particularly high (>90%) from September to December. Across southeast Alaska, as well as much of western North America, ARs initiate the vast majority of shallow precipitation-related landslides, although a minority of those ARs actually trigger widespread landsliding (Jacobs et al., 2016; Oakley et al., 2017; Cordeira et al., 2019).

Given this geographic setting, the community of Sitka is exposed to persistent, although largely unquantified, landslide hazards (Miller, 2019; Busch et al., 2021; Patton et al., 2022). Although it is difficult or impossible to reduce landslide hazards across broad hillslopes, landslide early warning systems can greatly reduce landslide risk to life and safety in these areas. With sufficient warning, residents can voluntarily evacuate high-hazard neighborhoods.

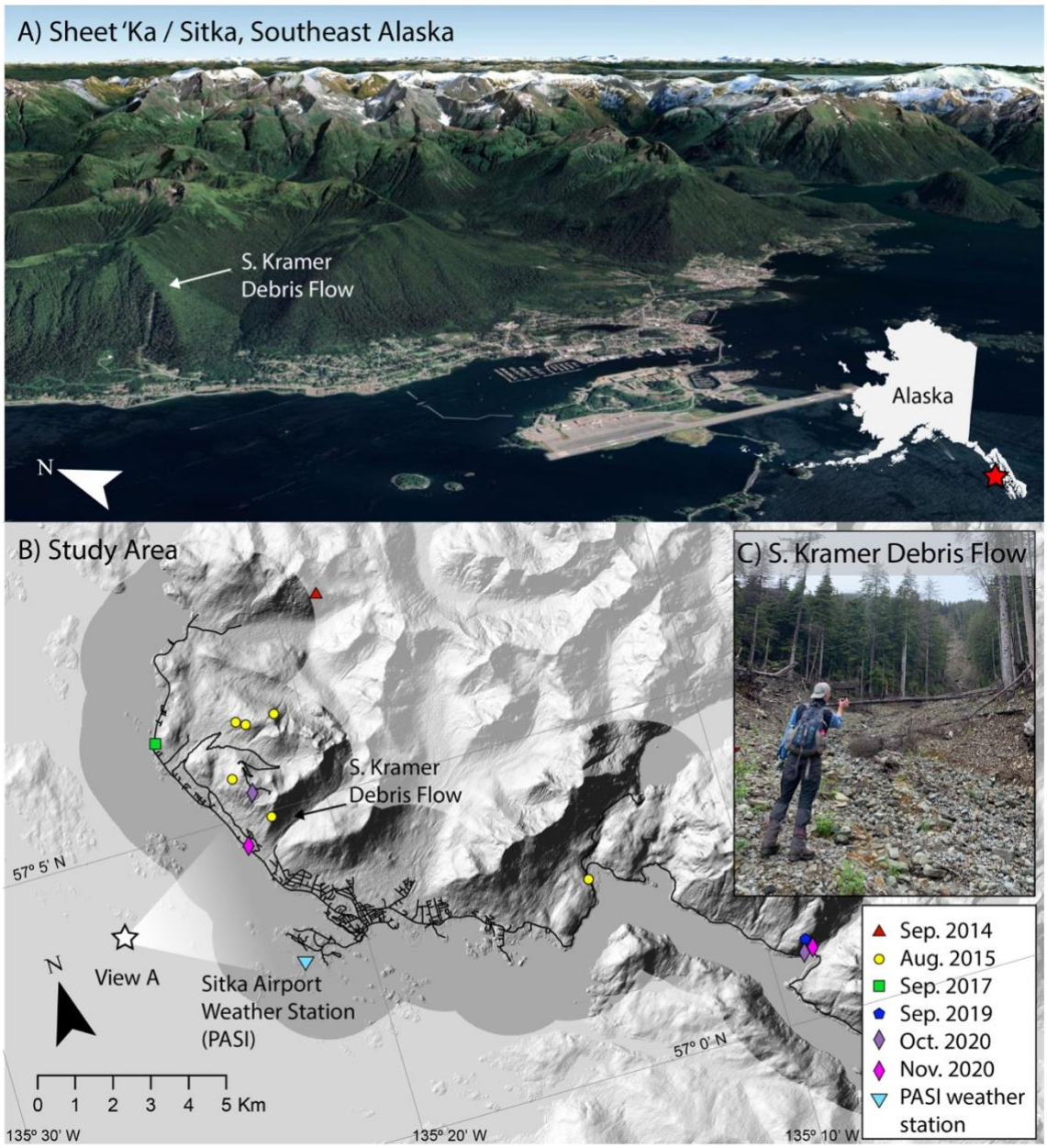



**Figure 2.1. Study area.** (A) Google Earth image of Sitka (Sheet'Ka), Alaska (©Landsat/Copernicus and Maxar Technologies, 2021). The town lies below over-steepened postglacial hillslopes susceptible to landslides. Some of the existing residential and municipal areas are built in landslide initiation or runout zones. (B) Map of recent rain-triggered landslides evaluated in this study shown on a shaded relief map from the U.S. Geological Survey 5-meter digital data, 2014. Higher contrast delineates areas within 2 km of the Sitka road network (dark lines). (C) Photo of the South Kramer Debris Flow, which initiated on August 18, 2015, and resulted in three fatalities. Photograph courtesy of the authors.

### 2.1.2 Developing precipitation thresholds for landslide warning

Landslide hazard estimates and precipitation thresholds exist at the global (Guzzetti et al., 2008; Kirschbaum and Stanley, 2018; Khan et al., 2021), regional (Piciullo et al., 2018; Segoni et al., 2018), and local scales (e.g., Mirus et al., 2018a), and even for individual landslides (Kristensen et al., 2021). Developing and applying new thresholds for landslide warning requires determining the most relevant variables and timescales to model landslide hazard in a particular region, considering data availability, and taking the risk tolerance of the targeted community into account. Almost every investigation of antecedent-triggering precipitation thresholds uses different observation timescales. These differences reflect different landslide types of interest (shallow versus deep-seated), hydrogeomorphic controls, climate, and the type and length of records available.

Hydrometeorological thresholds for landslide initiation have been proposed for nearby remote areas of British Columbia (Jakob et al., 2006) and suburban Vancouver (Jakob et al., 2012) in Canada, but no systematic landslide warning threshold currently exists at either local scales for towns within southeast Alaska or at the regional scale for southeast Alaska as a whole, despite its high susceptibility to slope failures (e.g., Darrow et al., 2022; Patton et al., 2022). Generations of knowledge in southeast Alaska and close observation of the natural environment provide rich understanding of landslides and other natural processes, but southeast Alaska lacks extensive written records of landslide occurrence with daily timestamps and sub-daily, spatially distributed precipitation records. This contrasts with many well-established landslide prediction models developed in the European Alps, Japan, and other data-rich regions that can draw on tens to thousands of observations of landslide-triggering precipitation and gridded precipitation datasets with high spatial and temporal resolution (e.g., Osanai et al., 2010; Saito et al., 2010; Berti et al., 2012; Lee et al., 2015; Leonarduzzi et al., 2017; Piciullo et al., 2017). Although previous estimates of rainfall thresholds have included only precipitation events that triggered landslides (Peruccacci et al., 2017), recent research has shown that including records of precipitation that did not trigger landslides can help sparse landslide datasets perform well (Peres and Cancelliere, 2021). Warning systems developed from hundreds to thousands of observed landslides are generally considered more trustworthy than those with few landslide-inducing events.

In southeast Alaska, the National Weather Service (NWS) forecasting products provide the best available warning information through weather and hydrologic watches, warnings, and advisories, but both communities and NWS forecasters have expressed a need for systematic analysis of landslide potential under different storm conditions (Busch et al., 2021). Recent investigations in Sitka, Alaska, (Booth et al., 2020; Chu et al., 2021; Vascik et al., 2021) and the community's desire for real-time landslide hazard assessments make this an ideal region to identify new precipitation thresholds and expand on established landslide prediction techniques for use in data-sparse regions. Our research objective is to provide the community of Sitka with a landslide early warning system that provides real-time and forecasted assessments of landslide hazard

to support individual and community-wide decision making. We estimate two metrics of daily landslide hazard in Sitka using statistical models trained with landslide inventory data and precipitation records. As described in detail in the methods section, our approach relies on models developed using hourly precipitation data from both landslide-triggering days (five) and all non-triggering days (>6000) within the period of record between 2002 and 2020.

## 2.2 METHODS AND DATA

To develop daily estimates of landslide hazard, we

- Compiled information about landslide occurrences with known timing near Sitka, Alaska, and weather-station precipitation data for a period of record with hourly precipitation data (2002–2020). Timing of each landslide was known within 12 hours or finer.

- Trained frequentist and Bayesian models (136 total) with historical records of precipitation and landslides to predict the daily probability of landslides (logistic regression) and the number of landslides (Poisson regression) that could occur based on cumulative precipitation. Logistic regression and Poisson regression are generalized linear models that can incorporate any number of predictor variables, including precipitation at different timescales and groundwater or hydrologic data.

- Compared the 136 models (Table 2.1) using a range of cumulative precipitation timescales to select the most appropriate model for the warning system. We considered models with a single predictor (cumulative "triggering" precipitation) and models with two predictors (cumulative "antecedent" precipitation and cumulative "triggering" precipitation during a specific time period on a given day).

- Checked the most appropriate model's sensitivity (and thus robustness) by removing individual landslide events (leave-one-out/jackknife validation).

- Using input from Sitka community members, established heuristic decision thresholds for multiple landslide warning levels based on the estimated probability of landsliding and expert judgement.

- Assessed how well a model trained on an earlier section of the time series was able to predict landslides in a later portion of the time series based on these thresholds, compared its predictive skill to a simpler alternative model based on historical landslide frequency, and evaluated how often landslide warnings would have been issued in the past.

The naming scheme we used for all models (frequentist versus Bayesian, logistic regression versus Poisson regression, precipitation timescales) is summarized in Table 2.1.

### 2.2.1 Data sources

To investigate landslide conditions in Sitka, we used existing hourly precipitation records from the nearby weather station (NWS station code PASI) operated by the Federal Aviation Administration (FAA) at Sitka Airport (NOAA NCEI, 2001). Climate records in Sitka go back to the early 19th century (Wendler et al., 2016), and hourly precipitation data (or sub-hourly) have been recorded at the airport weather station since 2002; we included all days with observations between November 12, 2002, and December 13, 2020. A nearby U.S. Climate Reference Network (USCRN) meteorological station (NWS station code SIKA2) also has documented sub-hourly precipitation since 2005 (Diamond et al., 2013). Some variation in precipitation is observed at these two locations, but for the purposes of simplicity we train the statistical models using a single precipitation gauge, the PASI gauge at the Sitka Airport.

Through a combination of air photo interpretation and field mapping, the U.S. Forest Service has curated an inventory of more than 12,000 landslides in southeast Alaska, with records dating back to the early 20th century, known as the Tongass National Forest Landslide Inventory (Tongass National Forest, 2017). To focus only on landslides likely to impact human safety and infrastructure, we subset the Tongass inventory to landslides within 2 km of the road network in Sitka, thus obtaining five days with recorded landslides out of 6,606 days with reported precipitation. We collected and synthesized information about the landslides near Sitka, including their timing and impacts.

**Table 2.1. Model naming system.** We evaluated 136 models fit to the complete landslide inventory data and precipitation records. We also evaluated one frequentist logistic regression model fit to a subset of these data, reserving some data for validation (training-test), and one simpler alternative model based on historical landslide frequency.

| Component | Type | Naming convention | Example |
|---|---|---|---|
| **Inference** | Frequentist | Model name begins with "F" | FL-1H1D |
| | Bayesian | Model name begins with "B" | BL-1H1D |
| **Model type** | Logistic regression | Second letter is "L" | FL-1H1D |
| | Poisson regression | Second letter is "P" | FP-1H1D |
| | Historical frequency | Labeled "simpler alternative" | SA |
| **Precipitation predictor 1** | Triggering precipitation (hours) | Number of hours is indicated as "#H" for 1, 3, 6, 12, or 24 hours | FL-3H |
| | Triggering precipitation (days) | Number of days is indicated as #D for 1,2,3,7, or 14 days; no hours indicated | FL-1D |
| **Precipitation predictor 2** | Antecedent precipitation | Number of days is indicated as "#D" for 1, 2, 3, 7, or 14 days | FL-1H2D |
| | No antecedent precipitation variable | No days indicated | FL-1H |
| **Training-test split** **(preferred model only)** | | | FL-TT-3H (frequentist logistic regression) |

### 2.2.2 Logistic and Poisson regression for estimating landslide hazard

Many previous works have used probabilistic techniques to predict landslide hazard (Brunetti et al., 2010; Berti et al., 2012; Tufano et al., 2019). In keeping with this practice, we used logistic regression to estimate the daily probability of landsliding as a function of precipitation. We also used Poisson regression to estimate intensity (number of landslides/day in the study region) as a proxy for the magnitude of the event. The outputs of logistic and Poisson regressions are useful because they provide a nuanced understanding of relative landslide hazard that allows practitioners to identify multiple working thresholds that lead to different levels of community response.

Logistic and Poisson regression are generalized linear models that can include any number of predictor variables. To determine the most effective precipitation timescale for estimating daily landslide hazard in Sitka, we tested a series of models with predictors at a range of timescales that include (a) triggering, or (b) triggering with antecedent precipitation. We considered two model set-ups: the first (trigger-only) estimates daily landslide hazard (probability or intensity on day $d$) as a function of cumulative precipitation during a specified time period $t$, which is either a sub-daily interval of day d or a time period leading up to and including day $d$. We investigated time periods t of 1, 3, 6, 12, and 24 hours and 2, 3, 7, and 14 days. The second model set-up (trigger-antecedent) introduces an additional predictor to describe cumulative precipitation during an antecedent time period a preceding day d, and uses only sub-daily time periods for $t$. We considered antecedent periods of 1, 2, 3, 7, and 14 days.

For each day of recorded precipitation between 2002 and 2020 $d$ (6,606 days), we used a series of moving windows to extract the maximum cumulative precipitation in each sub-daily time period $t$ on that day and cumulative precipitation for all other time periods $t$ and $a$ leading up to and including that day. This applies for both days with landslides and days without. For example, on a day with a landslide, 3-hour "triggering" precipitation ($P_t$) represents the highest cumulative three hours between midnight and 11:59 PM local time. A 1-day "antecedent" precipitation ($P_a$) period is the 24-hour period before midnight on the day of the landslide. The precipitation timescales we considered are designed to integrate with existing NWS precipitation forecasting products, which provide precipitation estimates for 3-hour intervals for the upcoming ~48 hours and 6-hour intervals for the following ~48 hours.

The logistic regression models have the form:

$$y_d \sim Bernoulli(p_d)$$
$$logit(p_d) = \beta_0 + \beta_1 P_t + \beta_2 P_a$$

*Equation 2.1*

where $y_d$ is a binary indicator of whether a landslide was observed on day $d$, $p_d$ is the probability of having a landslide on day $d$, and ~ indicates that $y_d$ is modeled as a Bernoulli distributed random variable. $\beta_0$ is the intercept of the generalized linear model;

$\beta_1$ is the coefficient of cumulative precipitation ($P_t$) during time period $t$; and $\beta_2$ is the coefficient of antecedent precipitation ($P_a$) during time period $a$, which is excluded in the single-predictor models. Logistic regression models are indicated with an L in their name (Table 2.1).

The Poisson regression models have a similar form:

$$z_d \sim Poisson(\lambda_d)$$
$$ln(\lambda_d) = \alpha_0 + \alpha_1 P_t + \alpha_2 P_a$$

*Equation 2.2*

where $z_d$ is the number of landslides observed on day $d$, $\lambda_d$ is the average intensity of landsliding (landslides/day/area) on day $d$, $\alpha_0$ is the intercept, $\alpha_1$ is the coefficient of cumulative precipitation, and $\alpha_2$ is the coefficient of antecedent precipitation, again excluded in the single-predictor models. Poisson regression models are indicated with a P in their name (Table 2.1).

Landslide days (five days) are rare compared to non-landslide days (> 6000), leading to an imbalance in the dataset that must be considered when setting decision thresholds.

We applied both frequentist (F, Table 2.1) and Bayesian (B, Table 2.1) approaches to fitting the logistic and Poisson regressions. Frequentist inference assumes that there is a true, unknown set of parameters and that the observed data result from an infinitely repeatable sampling experiment. Frequentist 95% confidence intervals around the point estimate for a parameter have a 95% probability of including the true parameter value, if the experiment were repeated a large number of times. Bayesian inference, in contrast, provides posterior parameter estimates, which are probability distributions of all possible parameter estimates that are compatible with the observed landslide data and our prior knowledge, which is specified in the form of a probability distribution. This is a useful property for estimating hazard from landslide inventories with few reported landslides because the posterior probability distribution quantifies how certain we can be of the parameter estimates, given few data points, and incorporates previous knowledge of landslide processes. A Bayesian 95% credibility interval contains 95% of the posterior probability, providing an arguably more intuitive estimate of uncertainty.

Both frequentist and Bayesian approaches have been applied in landslide research (e.g., Berti et al., 2012; Segoni et al., 2018). Frequentist approaches may be familiar to a wider range of users and are typically easy to apply out of the box in popular programming languages. Bayesian approaches offer potential advantages for small datasets, particularly because they quantify parameter uncertainty given the available data but are less commonly known and require sufficient expertise to select prior distributions. Here, we explore both inferences and compare their output and application.

For the Bayesian regressions, we chose weakly informative Student-*t* prior distributions that reflect our expectations about landslide activity in Sitka, specifically that (1) landslide probability and the number of landslides should increase with increasing

precipitation; (2) the probability of landsliding should be less than 50% at the mean precipitation in Sitka; (3) the average number of landslides should be fewer than 1 at the mean precipitation. Weakly informative Student-*t* priors are recommended defaults for Bayesian logistic regression (Gelman et al., 2008) that encode prior knowledge without overly influencing regression results. Additionally, when faced with an imbalanced dataset, as is the case here, such priors have been shown to produce stable regression coefficients, even in the case where there is near-perfect separation between landslide and non-landslide days (Gelman et al., 2008). Specifically, we chose the following:

$$\beta_0 \sim Student - t(3, -3, 2.5)$$
$$\beta_1, \beta_2 \sim Student - t(3, 3, 2.5)$$
$$\alpha_0 \sim Student - t(3, 0.5, 1)$$
$$\alpha_1, \alpha_2 \sim Student - t(3, -5, 1)$$

*Equation 2.3*

In the Bayesian regressions, precipitation values were standardized by subtracting the mean across all days and dividing by the standard deviation, also known as a z-score. These priors refer to standardized data, where the intercepts $\beta_0$ and $\alpha_0$ indicate the expected values for probability and intensity at the mean precipitation value.

We fit the frequentist models using the R *glm* software package, which relies on iterative weighted least squares to estimate parameters (R Core Team, 2019). We fit the Bayesian models using the R *brms* software package version 2.17.0 (Bürkner, 2017), which uses Hamiltonian Monte Carlo to estimate posterior parameter distributions as implemented in the Stan programming language (Stan Development Team, 2022). We ran four chains for 2000 iterations, discarding the first 500 draws as warm up. We checked the chains visually for convergence of parameter estimates; Gelman-Rubin convergence diagnostic (R-hat) values were in all cases 1, indicating convergence.

### 2.2.3 Model comparison and evaluation

We used several information criteria to compare models with different timescales of precipitation (1 hour to 14 days) to identify the best-performing model for use in a warning system. For the frequentist models, we calculated the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), which estimate out-of-sample prediction error. For the Bayesian models, we used approximate leave-one-out cross-validation as implemented in the R package loo version 2.5.1 (Vehtari et al., 2017), to estimate out-of-sample predictive accuracy (Leave-One-Out Information Criterion, LOOIC). We then chose the respective logistic regression and Poisson regression models with the lowest prediction error for further validation:

- FL-3H (Frequentist, logistic regression, 3-hour model),
- BL-3H (Bayesian, logistic regression, 3-hour model),

- FP-3H (Frequentist, Poisson, 3-hour model), and
- BP-3H (Bayesian, Poisson, 3-hour model).

Because the number of days with reported landslides is small compared to the number of days with no reported landslides, we evaluated the sensitivity of these four selected models (FL-3H, BL-3H, FP-3H, and BP-3H) to individual landslide events using leave-one-out cross validation. We removed each landslide event from the dataset and fit the models to the remaining data. We then evaluated the difference in parameter estimates between the complete dataset and the leave-one-out dataset.

In addition to the leave-one-out exercise for models trained on the entire data series (minus one landslide day), we also split the precipitation time series into a training and testing dataset to evaluate how well the model can be expected to predict landslides in the future. The training dataset is composed of the precipitation and landslide records from November 2002 – November 2019. The test dataset is from December 2019 – November 2020. We trained the logistic regression model on the training data and then predicted the probability of landslides for all days between December 2019 – November 2020 based on observed precipitation data. Although this "testing" window represents a relatively small portion of total days in the dataset, it does include 365 days and two of five (40%) landslide days. To test the sensitivity of our results of the length of the training and test periods, we also flipped the training and test periods (i.e., trained on the year 2020 and tested on the previous 17 years) and performed a similar validation.

One approach to understanding how well a model is able to predict landslides is to compare that model to alternative models (Table 2.1). We compared the skill of the best-fit frequentist and Bayesian logistic regression models (FL-3H, BL-3H) to a simpler alternative model (SA) that randomly guesses whether a landslide will occur based on the historical daily frequency of landslides (conceptually similar to tossing a weighted coin, where the sides are weighted according to how many landslide or non-landslide days have been recorded). This also provides a way to evaluate the added value of applying a more complex statistical model over a simpler model. The equation for this model is:

$$y_d \sim Bernoulli(p_d)$$
$$p_d = n_{ls}/n$$

*Equation 2.4*

where, as in the logistic regression models, $y_d$ indicates whether a landslide is observed, $p_d$ is the probability of landsliding, $n_{ls}$ is the number of days with reported landslides, and $n$ is the total number of days on record.

We compared the best-fit logistic regression models (FL-3H, BL-3H) to the simpler model using the Brier Skill Score (BSS). The Brier Score (BS) is the mean squared error of the predicted probabilities and is calculated as

$$BS = \frac{\sum_{d=1}^{n}(p_d - y_d)^2}{n}$$

*Equation 2.5*

and where lower scores indicate better skill (Brier, 1950).

The BSS then compares the logistic regression model ($BS_{logistic}$) to the simpler alternative model ($BS_{SA}$):

$$BSS = 1 - \frac{BS_{logistic}}{BS_{SA}}$$

*Equation 2.6*

where a BSS of 0 indicates that the models have the same skill, a BSS > 0 indicates that the logistic regression model outperforms the reference historical daily frequency model (i.e., weighted coin toss), and a BSS < 0 indicates that the logistic regression model performs worse than the simpler model.

### 2.2.4 Setting multiple decision thresholds for different hazard levels

Although landslide probability and intensity can be estimated for any precipitation over a specified period using the fitted logistic and Poisson regression models, decision thresholds must be chosen to specify when to issue warnings. Extensive conversations with emergency responders and community leaders revealed a variety of perspectives and priorities for Sitka's landslide warning system and different levels of risk tolerance (Busch et al., 2021). For example, emergency responders who are concerned about the considerable cost of false alarms preferred a higher hazard threshold in favor of fewer false alarms. Other citizens were comfortable with some false alarms, preferring to be alerted whenever there was an elevated chance of landslides. These previous findings informed our selection of multiple warning levels because each threshold must compromise between missed and false alarms, and dual thresholds can inform different types of decision-making with different alert levels (e.g., Mirus et al., 2018b).

The trade-off between missed alarms and false alarms is illustrated using a confusion matrix. A confusion matrix for a single threshold is a 2 × 2 matrix that shows the number of true alarms, false alarms, missed alarms, and true no-alarms by comparing the predicted outcome based on the probabilistic model and threshold with the true outcome. Metrics calculated from the confusion matrix can reveal optimal thresholds based on the user's tolerance for false alarms or missed alarms. In imbalanced datasets with few landslide days and many no landslide days, typically applied metrics for logistic regression thresholding like accuracy and the Receiver-Operating-Characteristic

(ROC) curve are less informative because they over-emphasize the importance of true no-alarms while masking the threshold's performance for true alarms and false alarms. For rare events, considering precision (ability to issue true alarms while avoiding false alarms) and recall (ability to issue true alarms while avoiding missed alarms) is preferable (Saito and Rehmsmeier, 2015).

Precision is defined as:

$$Precision = \frac{TrueAlarms}{TrueAlarms + FalseAlarms}$$

*Equation 2.7*

Recall is defined as:

$$Recall = \frac{TrueAlarms}{TrueAlarms + MissedAlarms}$$

*Equation 2.8*

To satisfy varying levels of risk tolerance within the community, we set two warning thresholds based on landslide probability estimated by best-performing frequentist logistic regression model (FL-3H). The lower threshold is set such that the system would have missed no landslide in the past (recall = 1), and the upper threshold is set such that every day with a landslide probability above the threshold has been associated with landslides in the past (precision = 1). Given the limited number of landslide days, a range of thresholds can achieve these results, calling for a heuristic approach in choosing final warning thresholds. We built a confusion matrix to illustrate how often warnings based on these thresholds would have been issued in the past and document the outcome of the event. We also used a confusion matrix to evaluate the number of warnings at each level that would have been issued between December 2019 and 2020 using the model trained on an earlier section of the time series as described in section 2.2.3.

Ideal practice would include models tested with historical precipitation forecasts (rather than observed precipitation), but archived forecast data are not readily available. Instead, we set preliminary thresholds using observed precipitation totals, and necessarily assume that forecasts are accurate. This introduces additional uncertainty in the landslide warning system, which relies on weather forecasts. Detailed analysis of the uncertainty in precipitation forecasts is beyond the scope of this paper, but validation and evaluation of the warning system could be used to refine warning thresholds.

## 2.3 RESULTS

### 2.3.1 Landslide events

Over the 18 years with hourly precipitation records, five rain events in Sitka have resulted in one or more landslides (Fig. 2.2; Table 2.2). In most cases, landslide timing is known within the hour or can be estimated based on eyewitness constraints and the precipitation record. All five landslide events were characterized by a few hours of intense precipitation (Fig. 2.2).

Other landslides have occurred near Sitka but are > 2 km from the road network and sensitive infrastructure. For example, the Starrigavan Landslide occurred several kilometers from town in 2014 and impacted a popular recreation area. Local accounts indicate that precipitation at the initiation site was much higher than precipitation observed at the Sitka Airport. Pronounced spatial heterogeneity in precipitation and weather is typical of southeast Alaska (Hennon et al. 2010; Shanley et al., 2015; Roth et al., 2018), which emphasizes the value of considering only very local (< 2 km from the road network) landslides for training prediction models using station-based precipitation data.

While antecedent precipitation conditions varied during these landslide events, the short-term (several hour) precipitation totals were high. For example, four of the events had peak 1-hour precipitation with >2-year return intervals as calculated by the NWS (Perica et al., 2012). Peak 3-hour precipitation during landslide events had between 2- and 25-year return intervals. Precise timing from eyewitness accounts and news records are available for the S. Kramer, Halibut Point, Medvejie, and Sand Dollar Dr. landslides, which all occurred within a few hours of peak precipitation recorded at the Sitka Airport.

When compared to the record of all precipitation that did not initiate landslides over the last 18 years, the events that produce landslides occur at the extreme high end of the distribution of cumulative precipitation (Fig. 2.3).

**Table 2.2. Summary statistics of recent landslide occurrences near Sitka.** Timing is listed as "precise" when the landslide timing is known within 30 minutes, and "approximate" when landslide timing is known to a broader window (~12 hours) within the date of initiation. *Four landslides that occurred in fall 2020 are attributed to either the October 26 or November 1 storms, but it is not known which storm. The landslides that initiated < 2 km of the road network occurred on known dates.

| Event name (local name for the most notable landslide) | Date, local time, precision | Total landslides | Landslides <2 km of road network | Failure type | Description of impacts and other notes |
|---|---|---|---|---|---|
| S. Kramer | Aug. 18, 2015, 9:30 am (precise) | 40+ | 6 | Debris flows initiated by shallow landslides | The debris flow near S. Kramer Avenue resulted in three fatalities. In addition to the six landslides counted here, multiple other landslides occurred in the region surrounding the study area. |
| Halibut Point Recreation Area | Sept. 4, 2017, Mid-day (approximate) | 1 | 1 | Landslide (type unspecified) | The landslide occurred in a recreation area, impacts unknown. |
| Medvejie | Sept. 20, 2019, 12:50 pm (precise) | 1 | 1 | Debris flow in an existing debris flow channel | Debris on the road to Medvejie Hatchery temporarily blocked access. |
| Harbor Mountain | Oct. 26, 2020 Early morning (approximate) | 6-10* | 2 | Debris flows | One debris flow temporarily blocked a highway. |
| Sand Dollar Drive | Nov. 1, 2020 6:00 pm (precise) *and* the night of Nov. 1 – Nov. 2 (approximate) | 2-6* | 2 | One debris flow; one fill-slope failure | The fill-slope failure occurred beneath a house on Sand Dollar Drive, impacting residential infrastructure. |

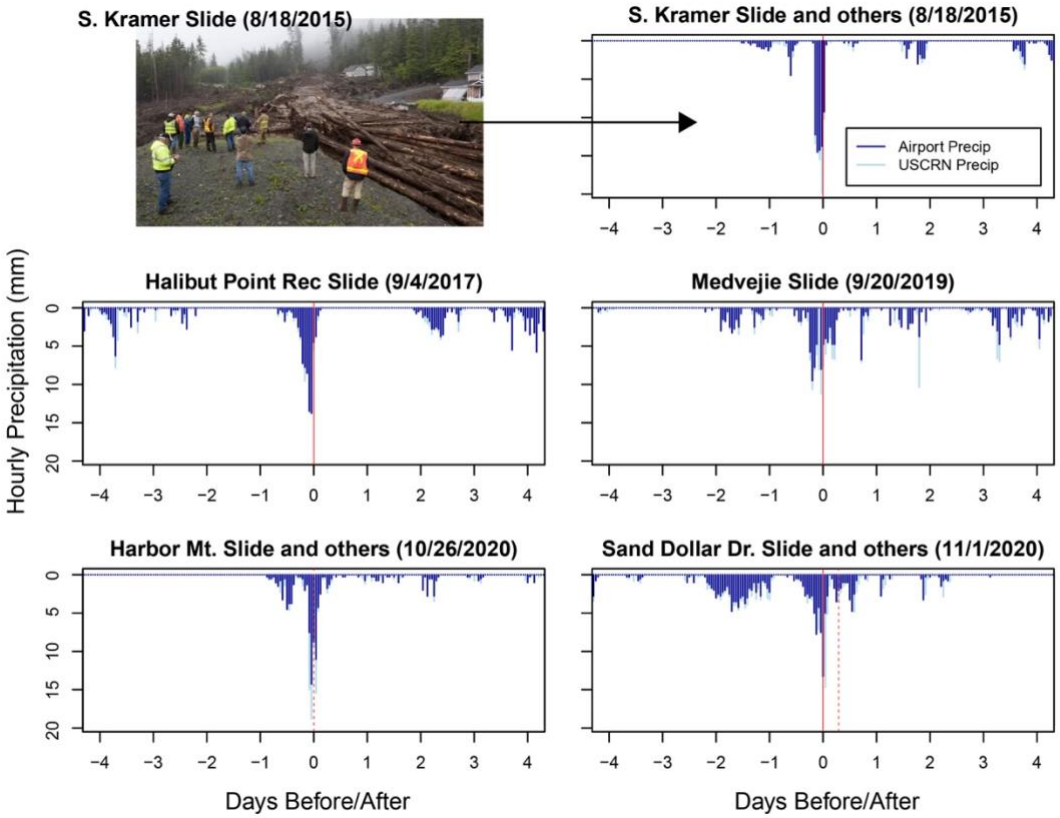



**Figure 2.2. Hourly precipitation before, during, and after landslide-initiating storms in Sitka.** Landslide prediction models in this paper were trained on the longer record of hourly precipitation at the weather station at the Sitka Airport (darker blue bars). Recorded precipitation from a nearby weather station at the U.S. Climate Reference Network (USCRN) site in Sitka is also displayed for comparison. Landslide timing is indicated by a solid red line for events where timing is constrained to ~30 minutes and a dashed line where timing is constrained within ~12 hours. Photo courtesy of James Poulson, Sitka Sentinel.

### 2.3.2 Landslide hazard prediction

We present results from both frequentist and Bayesian logistic regression and Poisson regression that predict landslide hazard (probability or intensity) based on precipitation totals over timescales ranging from 1 hour to 14 days (Fig. 2.3-2.6). Differences in model performance indicated by information criteria (section 2.3.3) show which timescales of precipitation provide the most useful prediction tools. For example, logistic regression based on 2-week precipitation totals is ineffective at separating landslide days from no-landslide days. Logistic regression based on 3-hour precipitation, however, does separate landslide days with some overlap (Fig. 2.3, 2.4).

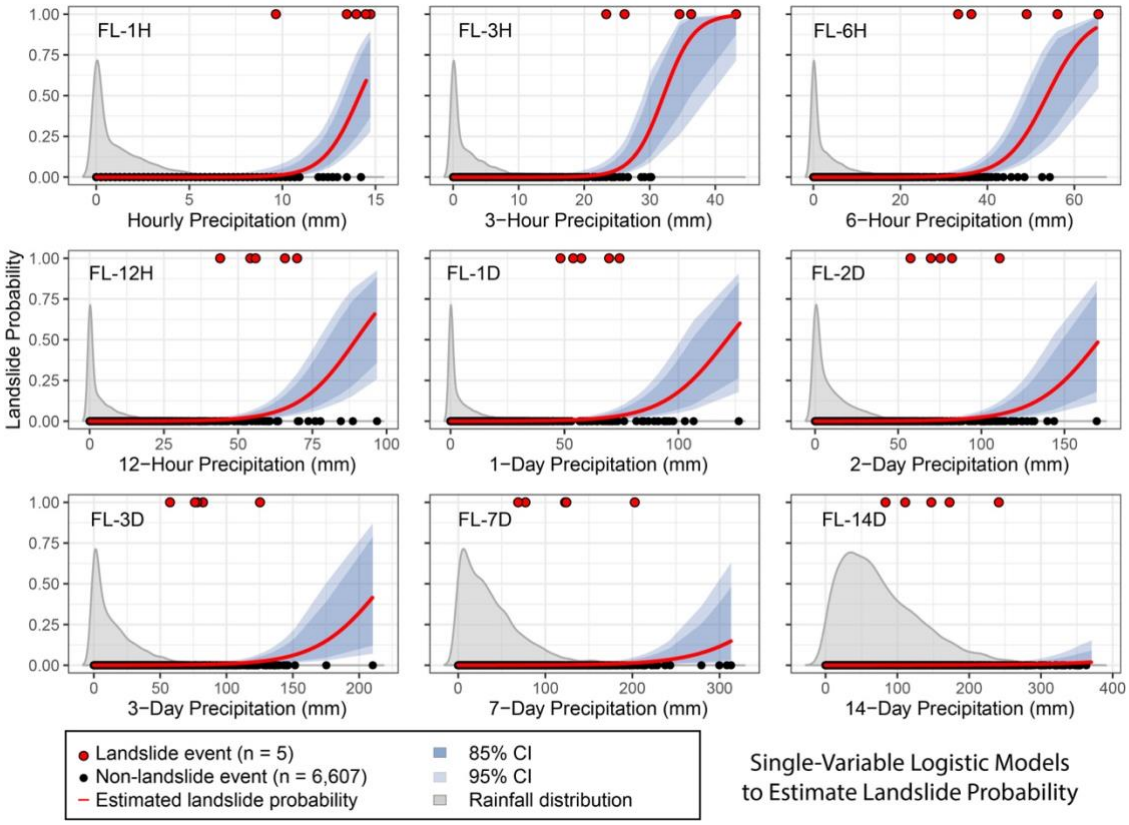



**Figure 2.3. Estimated daily landslide probability $p_d$ (red curve) from frequentist logistic regression based on different durations of precipitation from 1 hour to 2 weeks.** Confidence intervals (CIs) were estimated based on standard error. Event outcomes (red circles) of 1 indicate at least one landslide occurred, while 0 indicates no landslide occurred. Three-hour precipitation produces the model with the best fit (lowest estimated out-of-sample prediction error based on Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) fit parameters; Fig. 2.7). Kernel density distribution of all observed precipitation values are shown in gray, scaled for visual clarity.

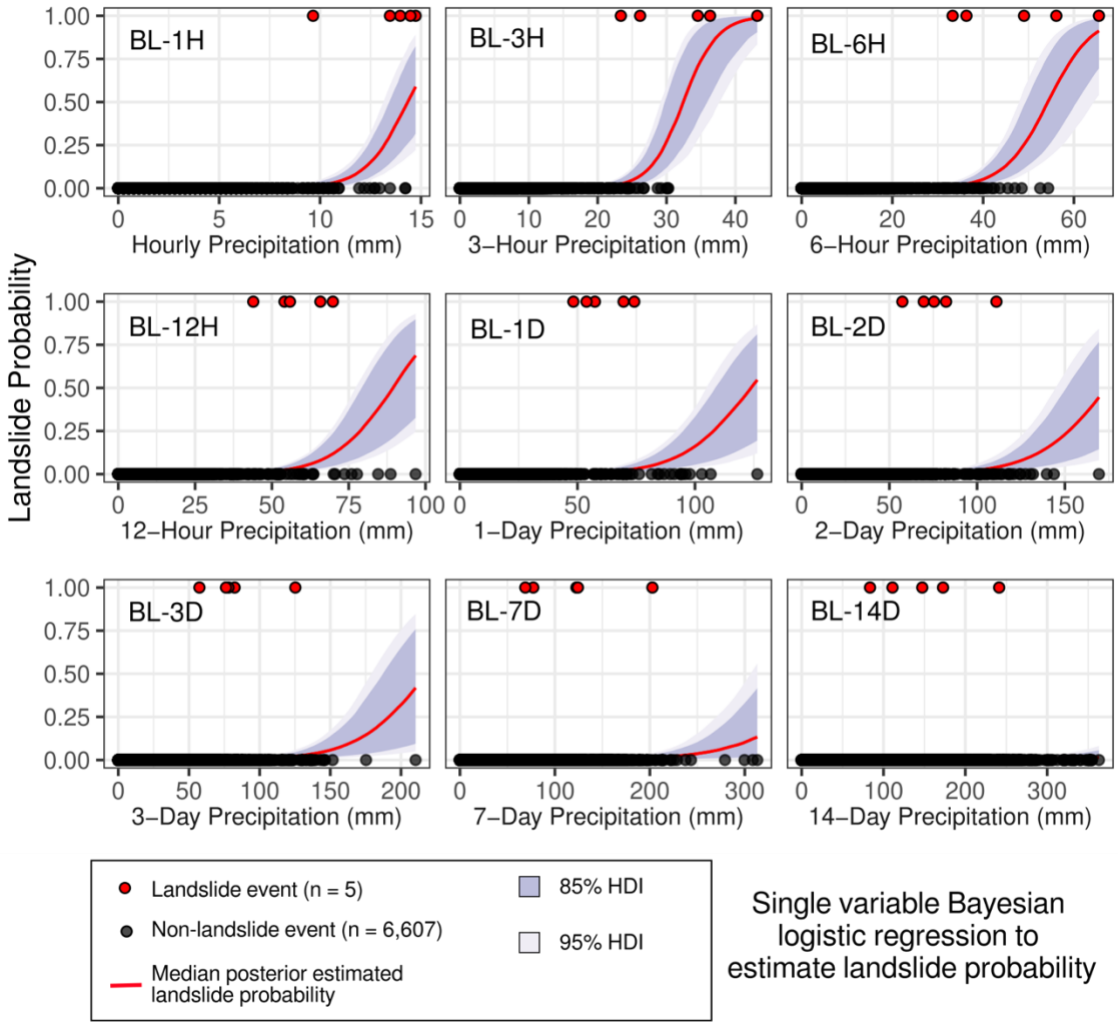



**Figure 2.4. Estimated daily landslide probability pd from Bayesian logistic regression**, showing the posterior median (red curve) with 85% (darker purple) and 95% (lighter purple) Highest Density Intervals (HDIs). The 95% HDI is the posterior estimate of parameter uncertainty and contains 95% of the distribution of all parameter values compatible with the data and our prior knowledge. At a single precipitation value, the 95% HDI contains the true landslide probability with 95% probability, conditional on the data, the model, and our prior knowledge. Three-hour precipitation gives the best out-of-sample predictive accuracy as measured by leave-one-out Information Criterion (LOOIC) (Fig. 2.8).

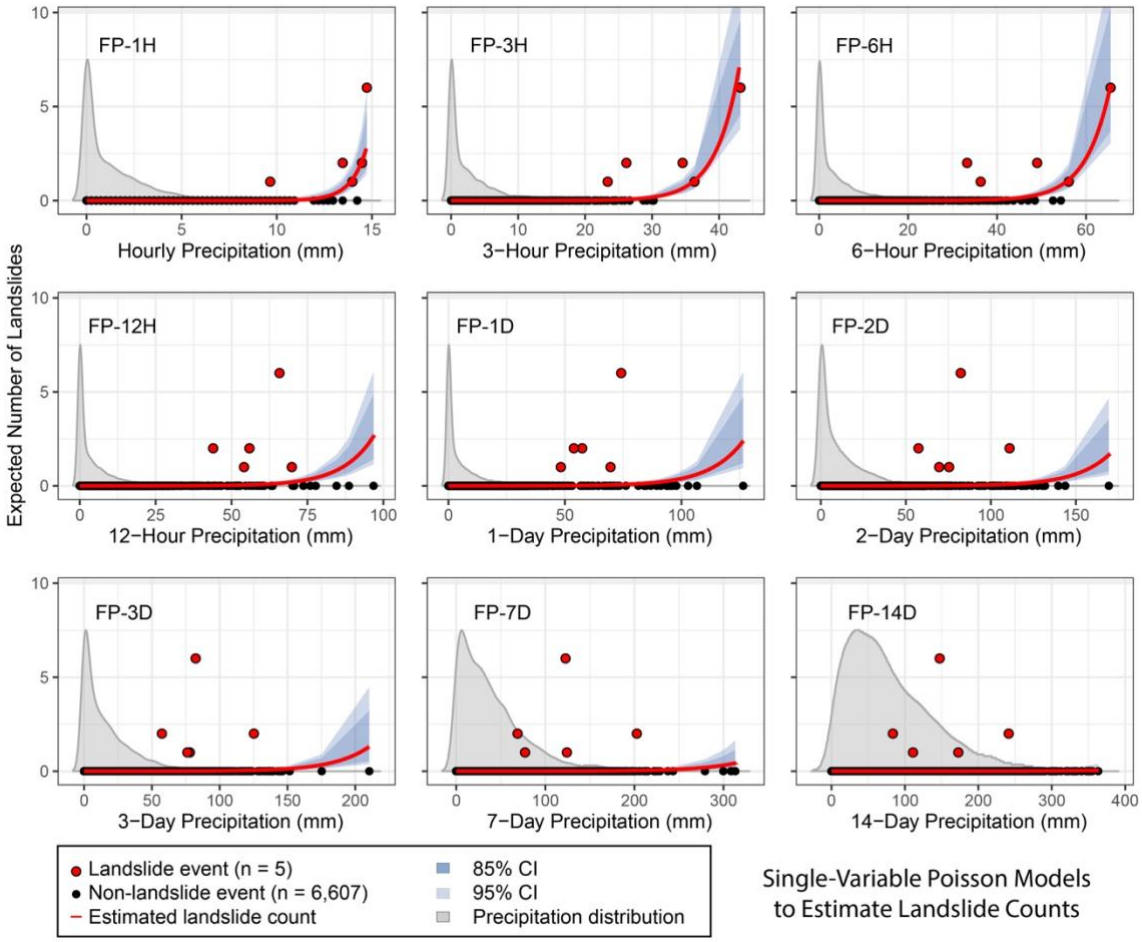



**Figure 2.5. Estimated daily average landslide count (red curve) from frequentist Poisson regression** ($\lambda_d$) based on different durations of precipitation from 1 hour to 2 weeks. Event outcomes (red circles) of ≥1 indicate the number of landslides that were reported, while 0 indicates no landslide reported. Confidence intervals (CIs) were estimated based on standard error. Three-hour precipitation produces the model with the best fit (lowest estimated out-of-sample prediction error based on Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) fit parameters (Fig. 2.7). Kernel density distribution of all observed precipitation values are shown in gray, scaled for visual clarity.

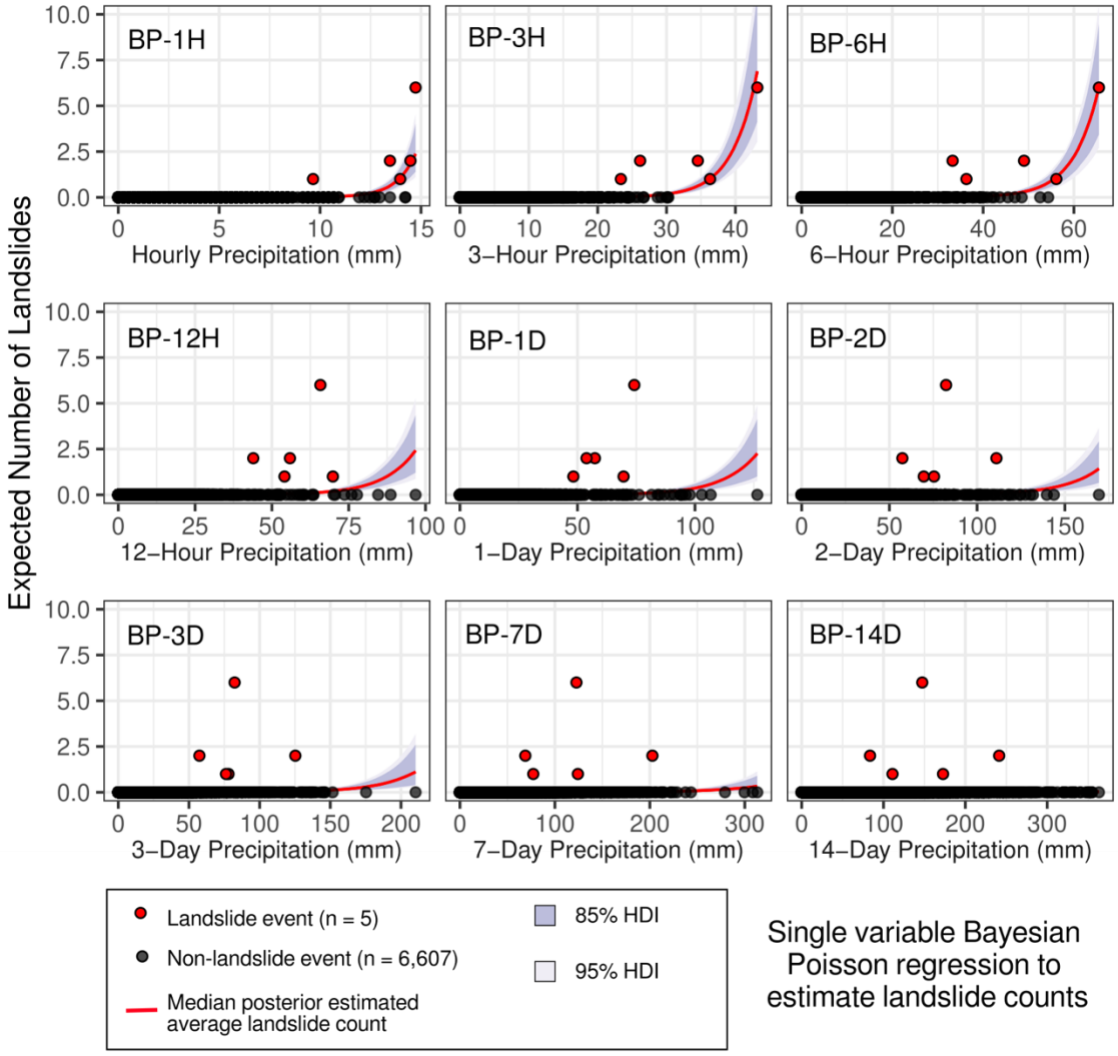



**Figure 2.6. Posterior Bayesian Poisson regression results**, showing the posterior median for the daily average estimated number of landslides $\lambda_d$ (red curve) with 85% (darker purple) and 95% (lighter purple) Highest Density Intervals (HDIs). Three-hour precipitation gives the best out-of-sample predictive accuracy as measured by leave-one-out Information Criterion (LOOIC) (Fig. 2.8).

### 2.3.3 Model comparison

Models that incorporated short-term precipitation (three hours) demonstrated best fit to the data and lowest estimated out-of-sample prediction errors (Figs. 2.7–2.8). All models that incorporate a metric of precipitation on the day of the landslide show lower AIC, BIC, and LOOIC values than models based on accumulated precipitation over multiple days.

We note that for many of the Bayesian models with longer precipitation timescales (>1 day), Pareto-k values for some of the landslide days were > 0.7, indicating that these models would be very unlikely to predict a landslide at that precipitation value. In contrast, Pareto-k values for all landslide days in the 3-hour logistic regression model are < 0.7, confirming that this model is not overly sensitive to the individual landslide points.

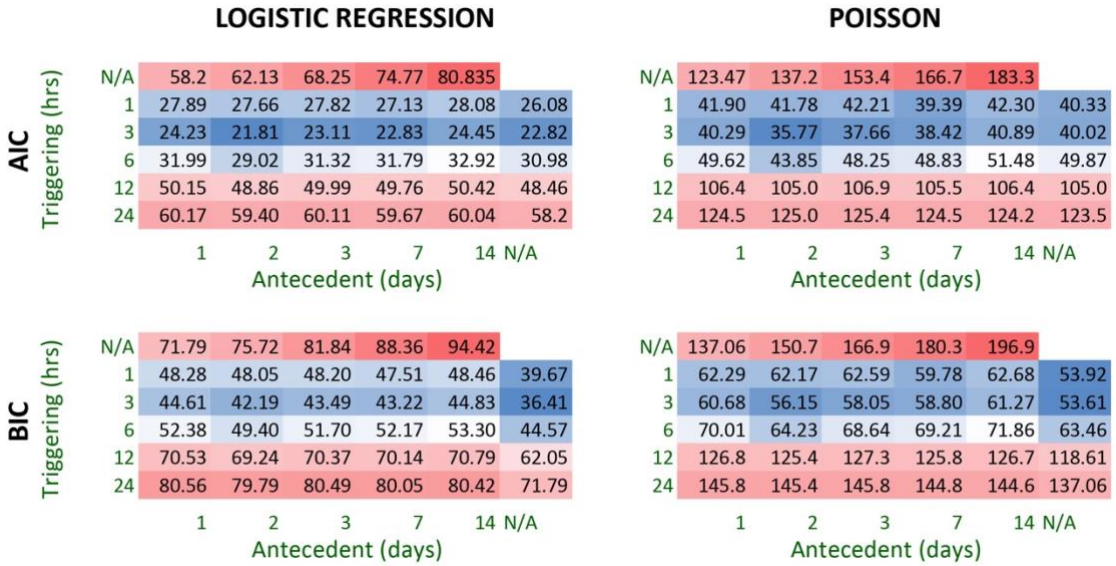



**Figure 2.7. Information criteria for a suite of frequentist models** that estimate landslide probability (logistic regression) or number of landslides (Poisson regression) for a given precipitation characteristic. Each cell corresponds to a model with one or two precipitation parameters, including daily maximum cumulative precipitation measured over 1–24 hours ("triggering") and antecedent precipitation measured over 1–14 days. Lower Akaike Information Criterion (AIC)/ Bayesian Information Criterion (BIC) values (blue) correspond to better model fit and higher AIC/BIC values (red) correspond to worse model fit, with BIC more heavily penalizing complex models with multiple predictor variables. AIC and BIC scores are specific to a regression type and should not be compared between the logistic regression (probability output) and the Poisson regression (count output).

We also qualitatively evaluated model fit by comparing the estimated landslide probability of the best-fit two-variable models that incorporate 3-hour triggering precipitation with 1-day (FL-3H1D) and 2-day (FL-3H2D) antecedent precipitation, respectively, and model FL-3H the best-fit one-variable model that considers only 3-hour triggering precipitation (Fig. 2.9). Using either the 24- or 48-hour antecedent precipitation as another predictor does modify the probability contours in Fig. 2.9B and 2.9D, but the tradeoff between false versus failed alarms is largely unchanged across all threshold values. Most observed landslides cluster at high to extreme triggering precipitation values and low or moderate antecedent precipitation totals. Increased model complexity does not significantly improve model fit for the available database of landslide occurrence (Figs. 2.7–2.8).

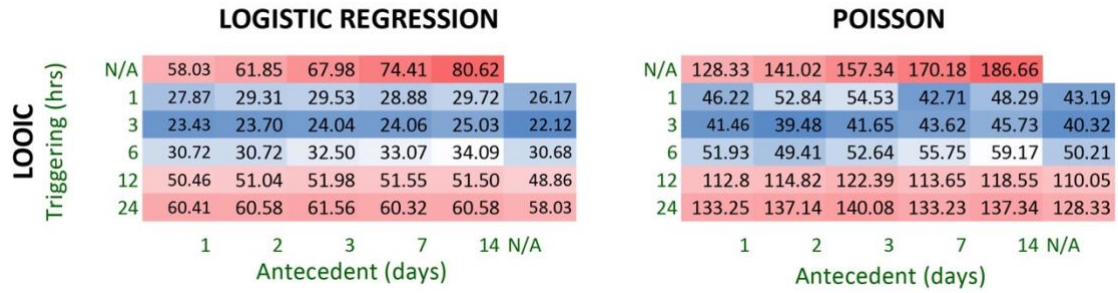



**Figure 2.8. Model comparison based on the Leave-One-Out-Information-Criterion (LOOIC) for a suite of Bayesian models** that estimate landslide probability and average count for a given precipitation characteristic. Each cell corresponds to a model with one or two precipitation parameters, including daily maximum cumulative precipitation measured over 1–24 hours ("triggering") and antecedent precipitation measured over 1–14 days. Lower LOOIC values (blue) correspond to higher out-of-sample predictive accuracy, and higher LOOIC values (red) correspond to lower accuracy.

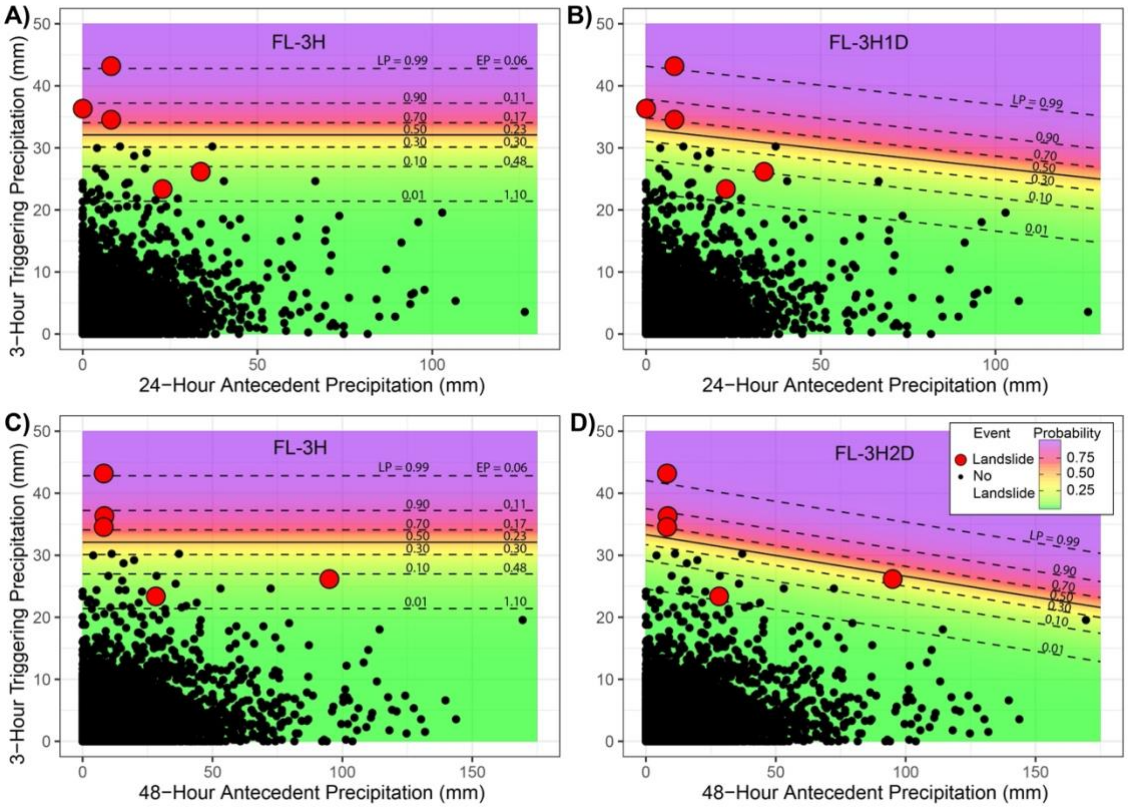



**Figure 2.9. Estimated landslide probability at varying precipitation values** from the preferred "trigger-only" frequentist logistic regression model (FL-3H, panels A and C) compared to models that include antecedent precipitation (panels B and D, models FL-3H1D and FL-3H2D, respectively). The color gradient and black contours show estimated landslide probability with reported data shown as black points (no landslide day) or red points (one or more landslides). Three-hour annual exceedance probabilities (EPs) reported by the NWS (Perica et al., 2012) also are plotted for the precipitation totals that generate the associated landslide probability (LP) contours from frequentist logistic regression in panels A and C.

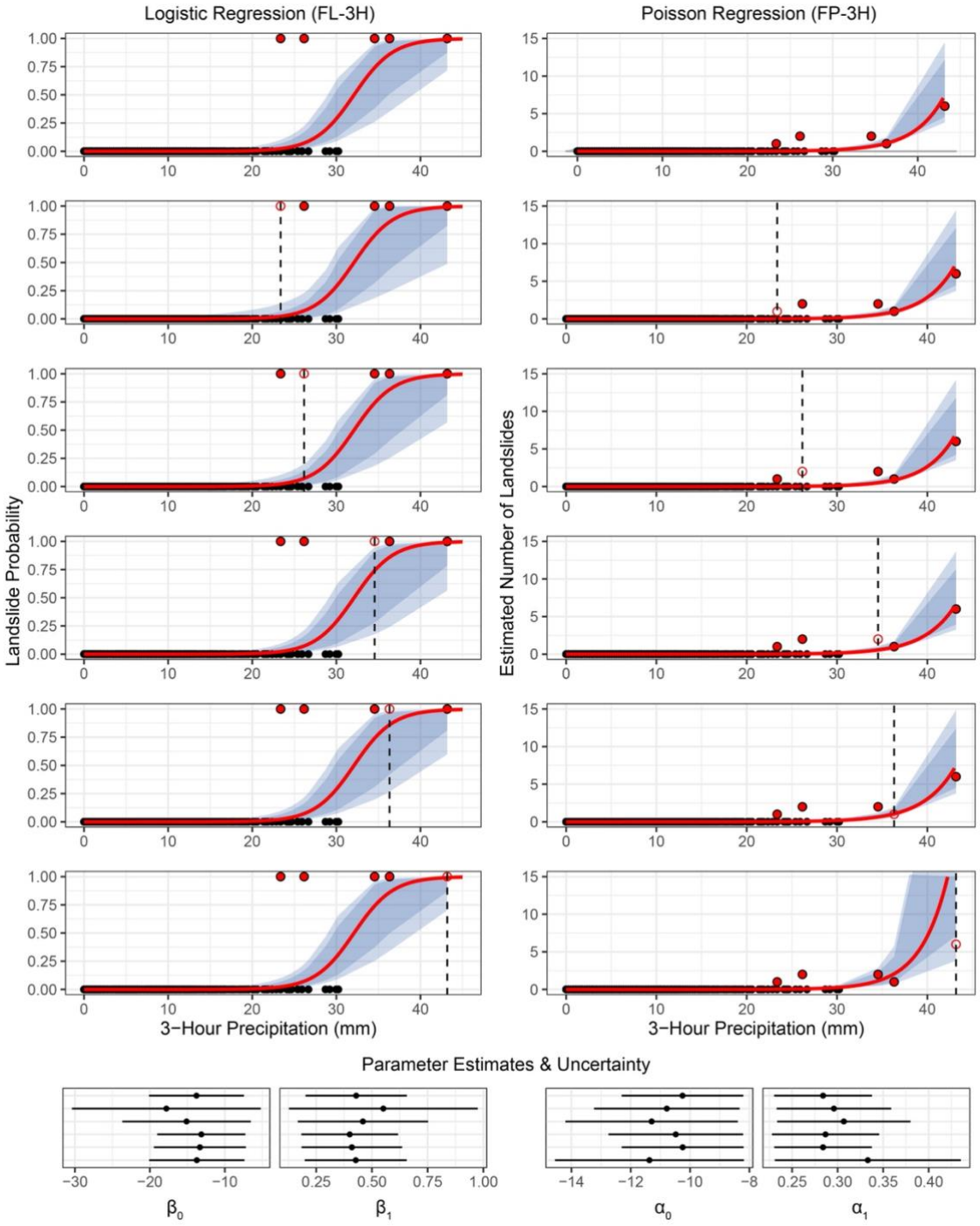



**Figure 2.10. Leave-one-out cross validation for the preferred frequentist logistic regression model** FL-3H (left column) and Poisson 3-hour model FP-3H (right column). Similar to Fig. 2.3 and Fig. 2.5, solid red points are landslide events, black points are non-landslide events, red lines show model estimates, and the dark and light shaded regions show 85% and 95% confidence intervals, respectively. Hollow red circles and dashed black lines show the landslide event that was omitted from each run. Model coefficient estimates are shown in the bottom panel with 95% confidence intervals based on the standard error. Model output and coefficient estimates remain largely unchanged when an individual landslide event is "missed" in the inventory, but the uncertainty bounds of the logistic regression and Poisson regression are sensitive to "missing" the landslide events with the lowest and highest precipitation, respectively.

Given the small number of observed landslide events in the dataset, we evaluated the sensitivity of the 3-hour models to individual landslide events using leave-one-out cross validation (Fig. 2.10 and Supplementary Fig. S2.1). Parameter estimates and their 95% confidence intervals for the leave-one-out models and the full dataset logistic regression models (FL-3H) overlap, indicating no relevant difference (Fig. 2.10). That we cannot distinguish these model parameters demonstrates that the model is not particularly sensitive to individual landslide points. The confidence intervals of the parameter estimates for the Poisson models also overlap with each other, with very high similarity in most cases, but we observe that the model is particularly sensitive to the single landslide day with six individual landslides. Further evaluations focus on the frequentist model (for ease of implementation) with lowest prediction error: frequentist model FL-3H.

### 2.3.4 Thresholds and predictive performance

Based on predicted daily landslide probability from the preferred logistic regression models (FL-3H, BL-3H), we established two decision thresholds for a landslide warning system (Fig. 2.11). A lower threshold was set at a probability of 0.01; in the past, this threshold would have resulted in no missed alarms (recall = 1). Any threshold below a probability of 0.023, based on model FL-3H estimates, results in a recall = 1; the threshold that maximizes precision at a recall = 1 would be 0.023, resulting in a precision of 0.22. We took a conservative approach and set the threshold lower, at 0.01. At 0.01, precision = 0.15, indicating that 28 false alarms would have occurred between 2002 and 2020 if this threshold had been used in the past.

Frequentist logistic regression indicates that a probability of 0.01 corresponds to a precipitation total of 21.3 mm in 3 hours (0.84 inch). Bayesian logistic regression indicates that a probability of 0.01 could correspond to precipitation values between 17.4 and 24.0 mm (95% HDI) (0.685 to 0.945 inch). An upper threshold that would have resulted in no false alarms (precision = 1) was set at 0.70. Based on model FL-3H estimates, any threshold above a probability of 0.31 results in a precision of 1, and thresholds that maximize recall for a precision of 1 range from 0.31 to 0.74 (Fig. 2.11). This wide range results from both few reported landslides and few reported precipitation values at the tail end of the precipitation distribution. At 0.70, recall = 0.6, indicating that two of the five reported landslide events occurred below this threshold. At this probability, frequentist logistic regression corresponds to precipitation of 34.0 mm in 3 hours (1.34 inches) and Bayesian logistic regression indicates precipitation between 31.0 and 39.2 mm (95% HDI) (1.22 to 1.54 inch).

At a precipitation total of 21.3 mm in 3 hours, the 3-hour Poisson models predict the occurrence of 0.015 landslides per day on average in the study area (FP-3H) or between 0.0034 and 0.031 landslides (95% HDI, BP-3H). At 34.0 mm of precipitation in 3 hours, Poisson regression predicts 0.56 (FP-3H) or between 0.25 and 0.86 landslides (95% HDI, BP-3H).

Decision thresholds for the landslide early warning system in Sitka were based on consideration of these ranges and our judgement. Using the frequentist logistic regression (FL-3H) model, probabilities < 0.01 were considered "Low" hazard; 0.01 < 0.70 were considered "Moderate" hazard; and > 0.70 were considered "High" hazard. The confusion matrix in Table 2.3 shows the warning levels that would have been generated by observed precipitation in 2002–2020 and the actual outcome. Probability of landslides for exceedance of our two alerts are consistent with other areas where dual- or multiple-thresholds are used (e.g., Chleborad et al., 2006).

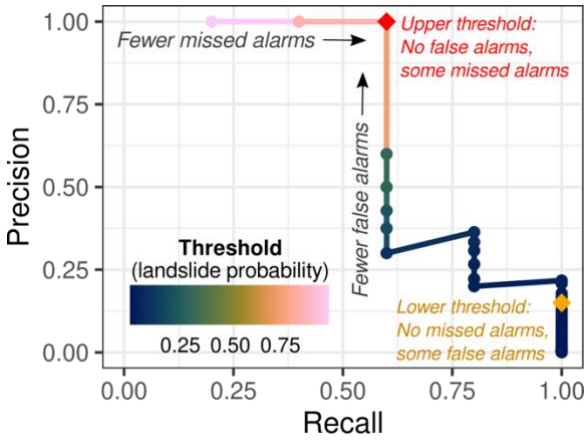



**Figure 2.11. Precision-Recall curve based the preferred frequentist logistic regression 3-hour model** (FL-3H). Recall is the true alarm rate and precision is the rate of true alarms/true alarms + false alarms. The upper threshold (red diamond) was heuristically set at a daily landslide probability of 0.7 to result in no false alarms (precision = 1). The lower threshold (orange diamond) was set at a probability of 0.1 to result in no missed alarms (recall = 1).

**Table 2.3: Warning levels that would have been generated between 2002 and 2020** by model FL-3H and the selected decision thresholds, showing the number of times each warning level would have been reached and the actual outcome. For example, a "High" warning would have been reached three times, and landslides occurred all three times; similarly, zero landslides occurred during times when "Low" probability of landslides would have been predicted by the model.

| | **"Low" Warning** | **"Moderate" Warning** | **"High" Warning** |
|---|---|---|---|
| **Landslide** | 0 | 2 | 3 |
| **No landslide** | 6573 | 28 | 0 |

At these thresholds, a moderate warning level would have been issued on 28 days between 2002 and 2020; two of those days actually resulted in landslides (Fig. 2.12A). A high warning level would have been issued three times, with all three days actually resulting in landslides. No landslide warning would have been issued on 6573 days, and no landslides would have occurred on a day without a landslide warning. This is useful for estimating the impact on a community based on the frequency of warnings. For example, the moderate warning level would have been issued 1-2 times per year (on average) in the historical record. However, while the confusion matrix summarizes how the model would have behaved in the past, it is not an indicator of how well the model can predict landslides because it uses the same dataset for validation as was used for training.

In the frequentist logistic regression model FL-TT-3H, we split the precipitation time series into training (November 12, 2002 – November 30, 2019) and test data (December 1, 2019 – November 30, 2020). Model FL-TT-3H is trained using only three landslide days and 6225 non-landslide days, and with thresholds at 0.01 and 0.70. This version of the model predicts elevated landslide probabilities on the days when landslides occurred in October and November of 2020 (Fig. 2.12B). Table 2.4 presents a confusion matrix for predicted warning levels for all days in the test dataset. A moderate warning would have been issued on two days, and both of those days did see landslides. No high warnings would have been issued. A low warning level would have been present on the remaining 363 days, with no landslides occurring. When we flip the training and testing periods and apply the same thresholds (probability = 0.01 and 0.7), the model would have issued warnings for all three testing landslide events, corresponding to a recall of 1 and no missed alarms (Supplementary Fig. S2.2, Table S2.1).

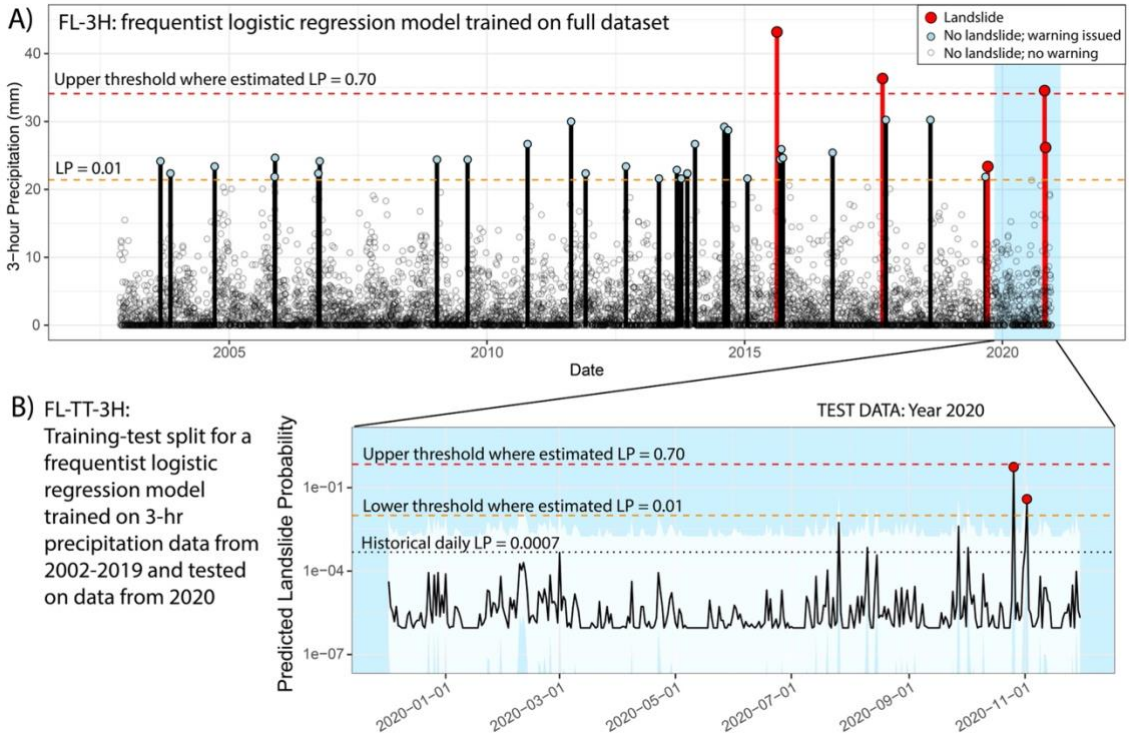



**Figure 2.12. Threshold exceedance between 2002 and 2020** (A) Timing of "moderate hazard" storm events (probability > 0.01 in the 3-hour frequentist logistic regression model, FL-3H) within the full period of record. All five landslide events in Sitka occurred within the last several years of record. Landslide probability is abbreviated as LP. (B) "Test" data for a model FL-TT-3H trained on the time series from 2002-2019. The black line indicates the point estimate and the gray field shows the 95% standard error. The model correctly predicts elevated hazard during the two landslide-initiating storms in 2020. The thresholds where probability = 0.01 and P = 0.70 are similar but not exactly the same in the two models (FL-3H and FL-TT-3H), such that one of the 2020 storms would have been predicted as "High" hazard (P > 0.70) in model FL-3H trained on the full dataset (A) but "Moderate" (0.01 < probability < 0.70) hazard in model FL-TT-3H trained on a subset of the dataset (B).

**Table 2.4: Confusion matrix for 2020 predictions**, based on model FL-TT-3H trained on 2002–2019 and with thresholds at probabilities of 0.01 and 0.7, showing the number of times each warning level would have been reached and the actual outcome. For example, a "Moderate" warning would have been reached twice, and landslides occurred both times.

| | **Low** | **Moderate** | **High** |
|---|---|---|---|
| **Landslide** | 0 | 2 | 0 |
| **No landslide** | 363 | 0 | 0 |

We also compared the performance of the logistic regression model FL-3H to a simpler model that randomly guesses if a landslide will occur based on historical frequency. The logistic regression model FL-TT-3H far outperforms random guessing (BSS = 0.44) for December 2019 – November 2020.

## 2.4 DISCUSSION

In this study, we applied logistic regression and Poisson regression to develop probabilistic daily estimates of landslide hazard in Sitka, Alaska, using limited landslide inventory data and nearly 20 years of hourly precipitation records (2002–2020). Based on these hazard estimates, we established two decision thresholds for landslide warning for implementation in a public-facing online dashboard that is driven with NWS forecast data and is automatically updated in real-time.

### 2.4.1 Probability based decision thresholds for landslide warning

Most commonly applied approaches to determining thresholds for landslide initiation seek to distinguish between precipitation and/or hydrological conditions that lead to landsliding from those that do not, or to determine a boundary below which landslides have not been previously observed. A disadvantage to such thresholds is that they (by design) provide only a binary outcome and no estimate of relative hazard. Probabilistic models, in contrast, estimate hazard and its uncertainty at every value of a predictor variable (e.g., maximum daily 3-hour precipitation), providing richer information than a binary threshold.

Identifying the most appropriate timescales for triggering and antecedent precipitation data influences the accuracy of landslide prediction tools (Gariano et al., 2020). By exploring the fit and predictive performance of selected precipitation timescales, including both triggering and antecedent precipitation, our models also provide insight into the physical processes that govern landslide initiation near Sitka.

We found that the 3-hour precipitation predictors best fit the data (e.g., models FL-3H, FP-3H, BL-3H, BP-3H), with negligible improvement in models that further incorporate 24- or 48-hour antecedent precipitation (e.g., models FL-3H1D, FL-3H2D, BL-3H1D). Including antecedent precipitation over timescales > 48 hours reduced model fit. Compared to some examples of cumulative precipitation thresholds in Seattle, Washington, which incorporate 3-day and 15-day antecedent precipitation totals (Chleborad et al., 2006; Scheevel et al., 2017), these timescales in Sitka are short. Similarly, intensity duration thresholds in Seattle rely on additional information on antecedent wetness for accurate performance (Godt et al., 2006).

These short timescales and lack of improvement with antecedent information for Sitka may result from multiple factors, including the steep topography, thin and locally permeable colluvial soils (Swanston and Marion, 1991), preferential flow and fracture-driven hydrology, unconstrained meso-scale storm patterns associated with landslide initiation in Sitka, and the small number of observed landslides in the dataset. We hypothesize that the importance of relatively short periods of intense precipitation in Sitka reflects the rapid hydrologic response of shallow, porous soils on fractured bedrock

that commonly are near saturation at critical failure depths. Antecedent information may be less predictive in this environment than in regions with thick or impermeable soils. Conversely, the lower performance of models using the 1-hour timescale indicates that shorter-duration bursts of intense rainfall are not necessarily sufficient to trigger landslides, and that some degree of sustained infiltration of rainfall is still needed.

Previous investigations on Chichagof Island, north of Sitka, demonstrate a rapid hydrologic response, with peaks in shallow pore pressure occurring within a few hours of observed precipitation and dissipating within several hours (Sidle, 1984); one investigation found that a shallow debris slide was most likely associated with maximum short-term intensity (2–6 hours) of precipitation, rather than storm totals (Sidle and Swanston, 1981).

Although the probabilistic outputs of logistic regression and Poisson regression are useful for understanding the relative magnitude of landslide hazard, it was necessary to establish decision boundaries for warning levels to communicate hazard to the public. As described in section 2.3.5, we selected two decision boundaries where frequentist logistic regression of maximum 3-hour precipitation (FL-3H) estimates a daily landslide probability of 0.01 and 0.70 for "Moderate" and "High" warning levels, respectively. Based on the historical record, moderate warnings would have been generated 31 times since 2002, and correctly predicted landslides only 3 times (Fig. 2.12). This means that there are many false alarms at the moderate warning level, but, by design, no missed alarms for this lower threshold. In comparison, the higher threshold was only crossed three times since 2002, all of which resulted in >1 landslide. These outcomes demonstrate the utility of having a tiered warning system, which provides more nuanced information about landslide hazard during a forecasted storm. As described below, these estimates do not account for unquantified error in precipitation forecasts or meso-scale atmospheric processes, which can generate above-threshold precipitation that may not be captured with traditional rain gauge networks (Collins et al., 2020).

The statistical models presented here are designed to be adaptable as additional data and observation allow us to validate and refine the models. Bayesian reasoning in particular acknowledges such updates by evaluating how much has been learned in the revised posterior. Hydrologic monitoring has recently been implemented in Sitka, but the available data record is relatively short. Further evaluation of this hydrologic time series could improve understanding of the hydrologic conditions that trigger landslides. For example, the preferred statistical models could be updated for seamless integration of additional hydrologic data or other predictor variables into the models if they improve prediction accuracy.

### 2.4.2 Few landslide observations and many no-landslide observations: strengths

While the overall dataset of precipitation observations is large (> 6,600 days of hourly precipitation record), the number of landslide-inducing events in this highly localized dataset is small (five landslide events < 2 km from the road network). This imbalanced dataset results in large model uncertainty for extremely high precipitation values that have been rarely observed.

Our work confirms that useful precipitation thresholds may be defined without landslide events by including the distribution of non-triggering events, as demonstrated in recent investigations (Peres and Cancelliere, 2021). This is possible because high data availability for non-triggering events does constrain the relatively low probability of landslides at low precipitation values. In other words, prediction models built with non-triggering events provide larger datasets than those that consider only landslide-triggering events and can be robust when considered alone or in combination with known landslide occurrence. The value of low precipitation totals and non-triggering events are often overlooked in landslide prediction studies, but the large number of observations results in statistically robust models. This is reflected in our leave-one-out cross validation results, where we show that logistic regression parameter estimates are insensitive to individual landslide events, and our training-test thresholding results, where we show that a model trained with only three landslide events is able to issue warnings during two test events. Poisson regression results are more sensitive to the largest landslide event, but a model trained without the largest landslide event would have predicted the occurrence of multiple landslides. We find that our preferred logistic regression models are more skillful in estimating landslide hazard than a simpler alternative model in which daily landslide probability is estimated by historical daily frequency, analogous to randomly guessing whether a landslide will occur based on how often they have occurred in the past.

Based on previous surveys, conversations, and feedback with the community in Sitka (Busch et al., 2021; Izenberg et al., 2022), one particularly valuable result of our modeling is a well-constrained "low" probability of landsliding, which is possible due to the extensive non-triggering events in the precipitation record. Identifying the times when landslide occurrence is not likely (i.e., < 1% daily probability) allows residents to manage anxiety while living in a potentially hazardous landscape with frequent and sustained rainfall throughout the year.

### 2.4.3 Few landslide observations and many no-landslide observations: challenges

Although few landslide events combined with many non-landslide days resulted in robust statistical models, setting and validating decision thresholds based on only five landslide days presented additional challenges. For example, although Receiver-

Operating-Characteristic (ROC) curves are commonly used to select optimal decision thresholds based on logistic regression (Giannecchini et al., 2016), this approach is not as informative for highly imbalanced datasets as a Precision-Recall curve (Saito and Rehmsmeier, 2015) because a wide range of thresholds give a low false alarm rate. We therefore opted to consider the Precision-Recall curve (Fig. 2.11), which provides more information about how well a threshold can distinguish between true alarms and false alarms.

However, even when using the more appropriate Precision-Recall curve, three sources of uncertainty make the choice of threshold challenging: (1) because there are only five landslide events, a range of decision thresholds result in identical levels of precision and recall. For the upper threshold, for example, threshold values with the same precision and recall range from daily landslide probabilities of 0.31 to 0.74 based on the frequentist 3-hour logistic regression model (FL-3H), (2) at any of these potential threshold levels, the given probability could be associated with a range of precipitation values; for an upper threshold of 0.70, for example, these range from 31.0 to 39.2 mm with 95% posterior probability, based on the Bayesian 3-hour logistic regression model (BL-3H), and (3) without further correction, logistic regression based on imbalanced datasets can underestimate landslide probability (King and Zeng, 2003). A conservative yet easy-to-implement approach is to reduce the lower threshold below an optimal balance of precision and recall (threshold tuning), particularly if the primary goal of the (lower) decision threshold is to reduce risk of missed alarms.

Despite these uncertainties, two decision thresholds corresponding to specific precipitation values were required for implementation in the warning system to integrate with weather forecasts. We therefore chose a heuristic approach based on expert judgement to select precipitation values within the range of thresholds that lead to the desired precision and recall. We set the lower threshold to a probability of 0.01 in model FL-3H, which is generated by 21.3 mm in 3 hours (0.84 inch). We set the upper threshold at a probability of 0.70 in model FL-3H, which is generated by 34.0 mm in 3 hours (1.34 inches).

We also found challenges associated with the timing of landslides over the 18-year record: although the hourly precipitation record in Sitka starts in 2002, no landslides with well-constrained timing were reported until 2015. This presents an additional challenge when validating the selected thresholds for future predictive performance. An ideal approach would be to iteratively split the dataset into multiple training and test groups (k-fold cross-validation), each time fitting statistical models to the training set and testing performance with the test set. Because landslides in this dataset only occur in the final third of the dataset, k-fold cross-validation results in many training sets with no reported landslides, which are then unable to predict elevated landslide hazard and are not useful estimates of performance because the models applied in the warning system do include reported landslides in the training data. We emphasize that in the training-test split presented here (FL-TT-3H), a model trained on only three landslide points with our selected probability thresholds is able to issue moderate warnings during both testing landslide events. Although no false alarms occurred during the

testing period, some false alarms can be expected in the future, as the low warning threshold has been exceeded in the past without triggering landslides (Table 2.3). More reported landslides in the Sitka area in the future would allow for more extensive validation of the thresholds. We consider several potential explanations for this inconsistent frequency within the study period.

First, it is possible that small, isolated slope failures may have occurred prior to the major event in 2015 but were not well documented. We consider it unlikely, however, that large or extensive landslides occurred and were not observed between 2002 and 2015 in Sitka. A second potential explanation is that, although "moderate" hazard rain events have occurred throughout the period of record (Fig. 2.12), the only "high" hazard events on record occurred after 2015. Global warming is predicted to result in increased frequency of extreme precipitation in upcoming years and decades, but further study is needed to evaluate the links between changing precipitation patterns and landslide occurrence in southeast Alaska. A third potential explanation is increased human alteration of hillslopes. At least one recent landslide occurred in human-made fill material, which may have increased susceptibility to landslide failure compared to a natural slope by altering hydrologic characteristics and slope strength through vegetation removal, slope cutting, and addition of fill material (e.g., Beville et al., 2010; Bozzolan et al., 2020; Johnston et al., 2021). If intense precipitation events are becoming more frequent or if slope modification increases through expanded urban development (or both), landslides in the study area are likely to become more frequent in the future.

### 2.4.4 Experience with frequentist and Bayesian inference for estimating landslide hazard

We explored both frequentist and Bayesian approaches to fitting logistic regression and Poisson regression models for estimating landslide hazard. By design, both of these approaches produced similar results; however, they do have different implications for use and interpretation. Frequentist inference remains more commonly used in landslide research (Melillo et al., 2018; Segoni et al., 2018), indicating familiarity, and frequentist approaches tend to be straightforward to implement in commonly used statistical modeling software, like R glm applied here. However, when considering imbalanced datasets with rare events, frequentist logistic regression may underestimate landslide probability (King and Zeng, 2003), and parameter estimates may be unstable when near perfect separation between landslide and no landslide days occurs, as is the case in this dataset. We note that statistical strategies exist to correct for underestimation (King and Zeng, 2003) and to obtain stable parameter estimates (Kosmidis and Firth, 2021), which could be applied if Bayesian inference were unavailable as a cross-check. Additionally, frequentist confidence intervals must be estimated in an additional step, and their interpretation is arguably less intuitive than Bayesian credibility intervals. However, in this case we found frequentist inference to still be useful for defining heuristic decision thresholds.

Bayesian inference remains less common in landslide research, and although additional expertise is required to set prior distributions and interpret the results (e.g., McElreath, 2020; Bürkner, 2017) Bayesian inference allows for incorporation of prior knowledge, which is advantageous when few landslide events are reported. Here, we encoded our prior knowledge that landslide activity is likely to increase with increasing precipitation in a weakly informative prior, which by design has only a small influence on the posterior distribution. When few data are available, more informative priors based on other studies could be used to, for example, tell the model about a distribution of outcomes that are known to be possible from nearby areas, but were not observed in the small dataset at hand. Weakly informative priors have also been shown to lead to stable parameter estimates in the case of imbalanced datasets with rare events, which overcomes the problem of unstable parameter estimates that frequentist logistic regression can show without an additional correction, making these Bayesian models better suited to estimating hazard from imbalanced datasets (Gelman et al., 2008). Posterior distributions of parameter estimates provide intrinsic estimates of uncertainty learned from the data, which informed our understanding of the range of precipitation values that could be associated with a given decision threshold.

Overall, we conclude that frequentist models are familiar and easy to implement, but Bayesian models capture the rare-events problem more explicitly and allow for better understanding of uncertainty. Either model can be effectively used for probabilistic landslide models and to determine meaningful decision thresholds. Here we present thresholds for the easy-to-implement frequentist model, but consideration of the best-fit Bayesian model and parameter uncertainty improved our understanding of both models' strengths and weaknesses. Furthermore, either workflow is transferrable to other regions, but they would need to be trained on local data.

### 2.4.5 Landslide prediction and uncertainty based on weather forecasts

Accurate precipitation and landslide timing data facilitated the development of robust thresholds for low, moderate, and high landslide potential. Implementation of these thresholds into actionable information to provide advance warning of landslide potential hinges upon accurate precipitation forecasts. Uncertainty in the forecasted precipitation is added to uncertainty of the model and decision threshold. As storms approach and precipitation forecasts become more constrained, the precipitation uncertainty will be reduced (Fabry and Seed, 2009; Ashok and Pekkat, 2022). Thus, landslide predictions for the future (days in advance) become more accurate as the storm approaches (hours in advance). Effective warning education can encourage residents to stay alert for updated landslide predictions. Further studies would be useful to better quantify the magnitude of error expected in dynamic storm forecasting and the relationship between forecast uncertainty and time into the future.

Models trained on and applied to precipitation data from a single monitoring station (Sitka Airport) cannot account for spatial variability in precipitation totals. Although the geographically small study area described here is intended to minimize these impacts, mountainous areas (like Sitka) are characterized by spatially variable climate and weather patterns (Johnson and Hanson, 1995; Tullos et al., 2016; Napoli et al., 2019). Meso-scale atmospheric processes linked to spatial distribution of landslide initiation are difficult to model (Collins et al., 2020) and not typically incorporated in kilometers-scale precipitation forecasts.

Predicted landslide hazard could also be complemented by applying the model to recent precipitation observations as an estimate of current hazard. "Nowcasting" by looking at observed precipitation (e.g., Kirschbaum and Stanley, 2018) incorporates instrument and model error, but not weather forecast error, and can alert residents to hazardous conditions that exceeded previous predictions. This type of information provides indicators of immediate hazard but is less useful for developing emergency response plans or informing operational decisions, which require sufficient lead time to take suitable action. In Sitka, for example, the observed landslides with high resolution timing data occurred 1–3 hours following peak precipitation, which may still provide valuable time for emergency responders and risk-averse individuals to take actions that reduce risk if precipitation totals exceed forecasts.

### 2.4.6 Application to landslide early warning system in Sitka, Alaska

In Sitka, our best-fit frequentist model FL-3H (based on 3-hour precipitation) with the three warning levels (low, moderate, high) described in section 2.3.5 has been applied to a public-facing dashboard for situational awareness. This dashboard provides residents, emergency planners, and NWS forecasters with near-real-time updates of current and predicted landslide hazard (referred to as "risk" in the dashboard for ease of use by a non-technical audience) and suggests actions to mitigate risk. In Sitka, individual differences in risk tolerance create a need for contextualized risk information to be available to everyone in the community (Busch et al., 2021). To provide this service, project members worked with the community, web developers, and NWS forecasters to construct a series of warning levels that indicate the three levels of landslide hazard developed and tested in this work. The beta version of this dashboard is accessible at sitkalandslide.org, which at the time of writing, is functionally serving as a landslide early warning system used by the public to inform individual decision making and by NWS forecasters to guide special weather watch, warning, and advisory products.

## 2.5 CONCLUSIONS

In this study, we developed and evaluated probabilistic models for landslide hazard estimation built with a small landslide inventory. The best-fit models used 3-hour triggering precipitation only. Including antecedent precipitation in addition to triggering rainfall did not improve model fit for the available database of landslide occurrence relative to using only the triggering precipitation conditions.

Despite the small number of landslide events (five days with landslides), a large dataset of non-triggering events produces robust model results, albeit with higher uncertainty at high precipitation values. Validation through leave-one-out analysis demonstrates that the model is robust even if we assume that we missed a landslide event. Furthermore, training the model on only three of five landslide events and thousands of no-landslide events would still have resulted in a model that could correctly predict the subsequent two landslide events. This model outperforms a much simpler alternative model based on historical landslide frequency. Combined with probabilistic models, the small number of landslide events allowed for the development of usable decision thresholds for landslide warning.

Although frequentist and Bayesian inference produce similar estimates of landslide hazard by design, they do have different implications for use and interpretation: frequentist models are familiar and easy to implement, but Bayesian models capture the rare-events problem more explicitly and allow for better understanding of uncertainty.

Developing precipitation thresholds based on time intervals (e.g., 3 hours) that match NWS forecasting products allows for application to landslide predictions within the NWS operational framework. This landslide early warning system was developed in partnership with the community and prioritized community needs identified in previous studies (Busch et al., 2021). A publicly accessible web dashboard, sitkalandslide.org, uses our preferred frequentist logistic regression model (FL-3H) and precipitation thresholds to display current landslide hazard (based on recent precipitation) and “forecasted” landslide hazard (based on NWS forecasts) in real time.

## 2.6 ACKNOWLEDGEMENTS

This research was supported by National Science Foundation funding (Award #1831770). Lisa Luna was supported by the Deutsche Forschungsgemeinschaft (DFG) research training group “Natural Hazards and Risks in a Changing World" (NatRiskChange GRK 2043). We would also like to thank our collaborators including Robert Lempert, Ryan Brown, and Max Izenberg (RAND Corporation) and Lisa Busch (Sitka Sound Science Center) for leading conversations with Sitkans about risk perception; Jacquie Foss (U.S. Forest Service) for curating and clarifying entries in the Tongass Landslide Inventory and documenting landslide impacts; Jeff Frankl, Klaas Hoekema and Steph Wall (Azavea) for designing and building the public-facing

warning dashboard; and Cora Siebert and Jacyn Schmidt (Sitka Sound Science Center) for documenting storm impacts. We thank Corina Cerovski-Darriau (USGS) and Ugur Özturk (University of Potsdam) for their thoughtful reviews, which allowed us to improve the clarity and conclusions of this work. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

## 2.7 DATA AVAILABILITY

Data used in this manuscript are freely available. Raw weather data from the Sitka Airport are available through the University of Utah's MesoWest climate data portal for station PASI (https://mesowest.utah.edu/) with full records available through a Synoptic Data Service API (https://developers.synopticdata.com/mesonet/). The full Tongass National Forest landslide inventory (initiation points and landslide areas) is available from the U.S. Forest Service data portal (https://gis.data.alaska.gov). Processed data are available through GitHub (https://github.com/pattonai/sitka-lews) with a release on January 5, 2023 (doi:10.5281/zenodo.7508537).

## 2.8 AUTHOR CONTRIBUTIONS

A.I.P. and J.J.R. conceptualized the study and A.I.P. and L.V.L. co-led the study. A.I.P. and L.V.L. performed statistical analyses, produced figures, and co-wrote the text. J.J.R., A.J., O.K., and B.B.M. contributed to discussing results and reviewing the manuscript. L.V.L. and A.I.P. revised and edited the text. The authors have no competing interests.

## 2.9 REFERENCES


Ashok, S.P., and Pekkat, S., 2022, A systematic quantitative review on the performance of some of the recent short-term rainfall forecasting techniques: Journal of Water and Climate Change, v. 13, p. 3004–3029, doi:10.2166/wcc.2022.302.

Berti, M., Martina, M.L.V., Franceschini, S., Pignone, S., Simoni, A., and Pizziolo, M., 2012, Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach: Journal of Geophysical Research: Earth Surface, v. 117, p. F04006, doi:10.1029/2012JF002367.

Beville, S.H., Mirus, B.B., Ebel, B.A., Mader, G.G., and Loague, K., 2010, Using simulated hydrologic response to revisit the 1973 Lerida Court landslide: Environmental Earth Sciences, p. 1249–1257, doi:10.1007/s12665-010-0448-z.

Bogaard, T., and Greco, R., 2018, Invited perspectives: Hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: Proposing hydro-meteorological thresholds: Natural Hazards and Earth System Sciences, v. 18, p. 31–39, doi:10.5194/nhess-18-31-2018.

Booth, A.M., Sifford, C., Vascik, B., Siebert, C., and Buma, B., 2020, Large wood inhibits debris flow runout in forested southeast Alaska: Earth Surface Processes and Landforms, v. 45, p. 1555–1568 doi:10.1002/esp.4830.

Bozzolan, E., Holcombe, E., Pianosi, F., and Wagener, T., 2020, Including informal housing in slope stability analysis - an application to a data-scarce location in the humid tropics: Natural Hazards and Earth System Sciences, v. 20, p. 3161–3177, doi:10.5194/nhess-20-3161-2020.

Brier, G.W., 1950, Verification of forecasts expressed in terms of probability: Monthly Weather Review, v. 78, p. 1–3, doi:10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2.

Brunetti, M.T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D., and Guzzetti, F., 2010, Rainfall thresholds for the possible occurrence of landslides in Italy: Natural Hazards and Earth System Sciences, v. 10, p. 447–458, doi:10.1016/S1387-6473(03)00110-6.

Bürkner, P.C., 2017, brms: An R package for Bayesian multilevel models using Stan: Journal of Statistical Software, v. 80, doi:10.18637/jss.v080.i01.

Busch, L., Lempert, R., Izenberg, M., and Patton, A., 2021, Run uphill for a tsunami, downhill for a landslide: Issues in Science and Technology, v. 38, p. 40–46.

Chae, B.-G., Park, J.-J., Catani, F., Simoni, A., and Berti, M., 2017, Landslide prediction, monitoring and early warning: A concise review of state-of-the-art: Geosciences Journal, v. 21, p. 1033–1070, doi:10.1007/s12303-017-0034-4.

Chleborad, B.A.F., Baum, R.L., and Godt, J.W., 2006, Rainfall thresholds for forecasting landslides in the Seattle, Washington, area — Exceedance and probability: U.S. Geological Survey Open-File Report 2006-1064, 31 pp., doi:10.3133/ofr20061064.

Chu, M., Patton, A., Roering, J., Siebert, C., Selker, J., Walter, C., and Udell, C., 2021, SitkaNet: A low-cost, distributed sensor network for landslide monitoring and study: HardwareX, v. 9, p. e00191, doi:10.1016/j.ohx.2021.e00191.

Collins, B.D., Oakley, N.S., Perkins, J.P., East, A.E., Corbett, S.C., and Hatchett, B.J., 2020, Linking mesoscale meteorology with extreme landscape response: Effects of narrow cold frontal rainbands (NCFR): Journal of Geophysical Research: Earth Surface, v. 125, p. e2020JF005675, doi:10.1029/2020JF005675.

Cordeira, J.M., Stock, J., Dettinger, M.D., Young, A.M., Kalansky, J.F., and Ralph, F.M., 2019, A 142-year climatology of northern California landslides and atmospheric rivers: Bulletin of the American Meteorological Society, v. 100, p. 1499–1509, doi:10.1175/BAMS-D-18-0158.1.

Cutter, S.L., and Finch, C., 2008, Temporal and spatial changes in social vulnerability to natural hazards: Proceedings of the National Academy of Sciences of the United States of America, v. 105, p. 2301–2306, doi:10.1073/pnas.0710375105.

Darrow, M.M., Nelson, V.A., Grilliot, M., Wartman, J., Jacobs, A., Baichtal, J.F., and Buxton, C., 2022, Geomorphology and initiation mechanisms of the 2020 Haines, Alaska landslide: Landslides, v. 19, p. 2177–2188, doi:10.1007/s10346-022-01899-3.

Diamond, H.J., Karl, T.R., Palecki, M.A., Baker, C.B., Bell, J.E., Leeper, R.D., Easterling, D.R., Lawrimore, J.H., Meyers, T.P., Melfert, M.R., Goodge, G., and Thorne, P.W., 2013, U.S. Climate Reference Network after One Decade of Operations: Status and Assessment. Bulletin of the American Meteorological Society, v. 94, p. 485–498, doi:10.1175/BAMS-D-12-00170.1, Retrieved from https://www.ncei.noaa.gov/access/crn/qcdatasets.html.

Elliott, J., and Freymueller, J.T., 2020, A block model of present-day kinematics of Alaska and western Canada: Journal of Geophysical Research: Solid Earth, v. 125, p. e2019JB018378, doi:10.1029/2019JB018378.

Fabry, F., and Seed, A.W., 2009, Quantifying and predicting the accuracy of radar-based quantitative precipitation forecasts: Advances in Water Resources, v. 32, p. 1043–1049, doi:10.1016/j.advwatres.2008.10.001.

Gariano, S.L., Melillo, M., Peruccacci, S., and Brunetti, M.T., 2020, How much does the rainfall temporal resolution affect rainfall thresholds for landslide triggering?: Natural Hazards, v. 100, p. 655–670, doi:10.1007/s11069-019-03830-x.

Gelman, A., Jakulin, A., Grazia Pittau, M., and Su, Y.-S., 2008, A weakly informative default prior distribution for logistic and other regression models: The Annals of Applied Statistics, v. 2, p. 1360–1383, doi:10.1214/08-AOAS191.

Giannecchini, R., Galanti, Y., Amato, G.D., and Barsanti, M., 2016, Probabilistic rainfall thresholds for triggering debris flows in a human-modified landscape: Geomorphology, v. 257, p. 94–107, doi:10.1016/j.geomorph.2015.12.012.

Godt, J.W., Baum, R.L., and Chleborad, A.F., 2006, Rainfall characteristics for shallow landsliding in Seattle, Washington, USA: Earth Surface Processes and Landforms, v. 31, p. 97-110.

Guzzetti, F., Gariano, S.L., Peruccacci, S., Brunetti, M.T., Marchesini, I., Rossi, M., and Melillo, M., 2020, Geographical landslide early warning systems: Earth-Science Reviews, v. 200, p. 102973, doi:10.1016/j.earscirev.2019.102973.

Guzzetti, F., Peruccacci, S., Rossi, M., and Stark, C.P., 2008, The rainfall intensity-duration control of shallow landslides and debris flows: An update: Landslides, v. 5, p. 3–17, doi:10.1007/s10346-007-0112-1.

Hamilton, T.D., 1986, Correlation of quaternary glacial deposits in Alaska: Quaternary Science Reviews, v. 5, p. 171–180, doi:10.1016/0277-3791(86)90182-4.

Hennon, A., Paul, E., Amore, D., David, V., Dustin, T., and Melinda, B., 2010, Influence of forest canopy and snow on microclimate in a declining yellow-cedar forest of southeast Alaska: Northwest Science, v. 84, p. 73–87, doi:10.3955/046.084.0108.

Izenberg, M., Brown, R., Siebert, C., Heinz, R., Rahmattalabi, A., and Vayanos, P., 2022, A community-partnered approach to social network data collection for a large and partial network: Field Methods, v. 0, p. 1–7, doi:10.1177/1525822X221074769.

Jacobs, A., Holloway, E., and Dixon, A., 2016, Atmospheric rivers in Alaska – Yes they do exist, and are usually tied to the biggest and most damaging rain generated floods in Alaska: 2016 International Atmospheric Rivers Conference, p. 925.

Jakob, M., Holm, K., Lange, O., and Schwab, J.W., 2006, Hydrometeorological thresholds for landslide initiation and forest operation shutdowns on the north coast of British Columbia: Landslides, v. 3, p. 228–238, doi:10.1007/s10346-006-0044-1.

Jakob, M., Owen, T., and Simpson, T., 2012, A regional real-time debris-flow warning system for the District of North Vancouver, Canada: Landslides, v. 9, p. 165–178, doi:10.1007/s10346-011-0282-8.

Johnson, G.L., and Hanson, C.L., 1995, Topographic and atmospheric influences on precipitation variability over a mountainous watershed: Journal of Applied Meteorology, v. 34, p. 68–87, doi:10.1175/1520-0450-34.1.68.

Johnson, A.C., Swanston, D.N., and Mcgee, K.E., 2000, Landslide initiation, runout, and deposition within clearcuts and old-growth forests of Alaska: Journal of the American Water Resources Association, v. 36, p. 17–30, doi:10.1111/j.1752-1688.2000.tb04245.x.

Johnston, E.C., Davenport, F. V., Wang, L., Caers, J.K., Muthukrishnan, S., Burke, M., and Diffenbaugh, N.S., 2021, Quantifying the effect of precipitation on landslide hazard in urbanized and non-urbanized areas: Geophysical Research Letters, v. 48, p. e2021GL094038, doi:10.1029/2021GL094038.

Khan, S., Kirschbaum, D.B., and Stanley, T., 2021, Investigating the potential of a global precipitation forecast to inform landslide prediction: Weather and Climate Extremes, v. 33, p. 100364, doi:10.1016/j.wace.2021.100364.

King, G., and Zeng, L., 2003, Explaining rare events in international relations: International Organization, v. 55, p. 693–715, doi:10.1162/00208180152507597.

Kirschbaum, D., and Stanley, T., 2018, Satellite-based assessment of rainfall-triggered landslide hazard for situational awareness: Earth's Future, v. 6, p. 505–523, doi:10.1002/2017EF000715.

Kosmidis, B.I., and Firth, D., 2021, Jeffreys-prior penalty, finiteness and shrinkage in binomial-response generalized linear models: Biometrika, v. 108, p. 71–82, doi:10.1093/biomet/asaa052.

Kristensen, L., Czekirda, J., Penna, I., Etzelmüller, B., Nicolet, P., Pullarello, J.S., Blikra, L.H., Skrede, I., Oldani, S., and Abellan, A., 2021, Movements, failure and climatic control of the Veslemannen rockslide, Western Norway: Landslides, v. 18, p. 1963–1980, doi:10.1007/s10346-020-01609-x.

Lee, S., Won, J.S., Jeon, S.W., Park, I., and Lee, M.J., 2015, Spatial landslide hazard prediction using rainfall probability and a logistic regression model: Mathematical Geosciences, v. 47, p. 565–589, doi:10.1007/s11004-014-9560-z.

Leonarduzzi, E., Molnar, P., and McArdell, B., 2017, Predictive performance of rainfall thresholds for shallow landslides in Switzerland from gridded daily data: Journal of the American Water Resources Association, v. 5, p. 6612–6625, doi:10.1111/j.1752-1688.1969.tb04897.x.

Lesnek, A.J., Briner, J.P., Lindqvist, C., Baichtal, J.F., and Heaton, T.H., 2018, Deglaciation of the Pacific coastal corridor directly preceded the human colonization of the Americas: Science Advances, v. 4, p. eaar5040, doi:10.1126/sciadv.aar5040.

Mann, D.H., 1986, Wisconsin and Holocene glaciation of southeast Alaska: in Hamilton, T.D., Reed, K.M., and Thorson, R.M. (eds.) Glaciation in Alaska: The Geologic Record, Alaska Geological Society, p. 237–265.

Marino, P., Peres, D.J., Cancelliere, A., Greco, R., and Bogaard, T.A., 2020, Soil moisture information can improve shallow landslide forecasting using the hydrometeorological threshold approach: Landslides, v. 17, p. 2041–2054, doi:10.1007/s10346-020-01420-8.

McElreath, R., 2020, Statistical Rethinking: A Bayesian Course with Examples in R and STAN: Chapman and Hall/CRC Press, 612 pp.

Melillo, M., Brunetti, M.T., Peruccacci, S., Gariano, S.L., Roccati, A., and Guzzetti, F., 2018, A tool for the automatic calculation of rainfall thresholds for landslide occurrence: Environmental Modelling and Software, v. 105, p. 230–243, doi:10.1016/j.envsoft.2018.03.024.

Miller, D., 2019, Modeling Susceptibility to Landslides and Debris Flows in Southeast: TerrainWorks Report to the U.S. Forest Service, 43 pp.

Mirus, B.B., Becker, R.E., Baum, R.L., and Smith, J.B., 2018a, Integrating real-time subsurface hydrologic monitoring with empirical rainfall thresholds to improve landslide early warning: Landslides, v. 15, p. 1909–1919, doi:10.1007/s10346-018-0995-z.

Mirus, B.B., Morphew, M.D., and Smith, J.B., 2018b, Developing hydro-meteorological thresholds for shallow landslide initiation and early warning: Water, v. 10, p. 1274, doi:10.3390/w10091274.

Napoli, A., Crespi, A., Ragone, F., Maugeri, M., and Pasquero, C., 2019, Variability of orographic enhancement of precipitation in the Alpine region: Scientific Reports, v. 9, p. 13352, doi:10.1038/s41598-019-49974-5.

National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI), 2001, Global Surface Hourly Precipitation for Sitka Airport station PASI. NOAA National Centers for Environmental Information, Retrieved from https://www.ncei.noaa.gov/access/search/data-search/global-hourly.

Oakley, N.S., Lancaster, J.T., Kaplan, M.L., and Ralph, F.M., 2017, Synoptic conditions associated with cool season post-fire debris flows in the Transverse Ranges of southern California: Natural Hazards, v. 88, p. 327–354, doi:10.1007/s11069-017-2867-6.

Osanai, N., Shimizu, T., Kuramoto, K., Kojima, S., and Noro, T., 2010, Japanese early-warning for debris flows and slope failures using rainfall indices with Radial Basis Function Network: Landslides, v. 7, p. 325–338, doi:10.1007/s10346-010-0229-5.

Patton, A.I., Roering, J.J., and Orland, E., 2022, Debris flow initiation in postglacial terrain: Insights from shallow landslide initiation models and geomorphic mapping in southeast Alaska: Earth Surface Processes and Landforms, v. 47, p. 1583–1598, doi:10.1002/esp.5336.

Peres, D.J., and Cancelliere, A., 2021, Comparing methods for determining landslide early warning thresholds: potential use of non-triggering rainfall for locations with scarce landslide data availability: Landslides, v. 18, p. 3135–3147, doi:10.1007/s10346-021-01704-7.

Perica, S., Kane, D., Dietz, S., Maitaria, K., Martin, D., Pavlovic, S., Roy, I., Stuefer, S., Tidwell, A., Trypaluk, C., Unruh, D., Yekta, M., Betts, E., Bonnin, G., Heim, S., Hiner, L., Lilly, E., Narayanan, J., Yan, F., and Zhao, T., 2012, NOAA Atlas 14 Precipitation-Frequency Atlas of the United States Volume 7 Version 2.0: Alaska. U.S. Department of Commerce National Oceanic and Atmospheric Administration, 119 pp., Retrieved from https://www.weather.gov/media/owp/oh/hdsc/docs/Atlas14_Volume7.pdf, Data from https://hdsc.nws.noaa.gov/hdsc/pfds/pfds_map_ak.html.

Peruccacci, S., Brunetti, M.T., Gariano, S.L., Melillo, M., Rossi, M., and Guzzetti, F., 2017, Rainfall thresholds for possible landslide occurrence in Italy: Geomorphology, v. 290, p. 39–57, doi:10.1016/j.geomorph.2017.03.031.

Piciullo, L., Calvello, M., and Cepeda, J.M., 2018, Territorial early warning systems for rainfall-induced landslides: Earth-Science Reviews, v. 179, p. 228–247, doi:10.1016/j.earscirev.2018.02.013.

Piciullo, L., Gariano, S.L., Melillo, M., Brunetti, M.T., Peruccacci, S., Guzzetti, F., and Calvello, M., 2017, Definition and performance of a threshold-based regional early warning model for rainfall-induced landslides: Landslides, v. 14, p. 995–1008, doi:10.1007/s10346-016-0750-2.

R Core Team, 2019, R: A language and environment for statistical computing: R Foundation for Statistical Computing, Vienna, Austria, Retrieved from https://www.r-project.org/.

Riehle, J.R., Champion, D.E., Brew, D.A., and Lanphere, M.A., 1992a, Pyroclastic deposits of the Mount Edgecumbe volcanic field, southeast Alaska: Eruptions of a stratified magma

chamber: Journal of Volcanology and Geothermal Research, v. 53, p. 117–143, doi:10.1016/0377-0273(92)90078-R.

Riehle, J.R., Mann, D.H., Peteet, D.M., Engstrom, D.R., Brew, D.A., and Meyer, C.E., 1992b, The Mount Edgecumbe tephra deposits, a marker horizon in southeastern Alaska near the Pleistocene-Holocene boundary: Quaternary Research, v. 37, p. 183–202, doi:10.1016/0033-5894(92)90081-S.

Roth, A., Hock, R., Schuler, T. V, Bieniek, P.A., Pelto, M., and Aschwanden, A., 2018, Modeling winter precipitation over the Juneau Icefield, Alaska, using a linear model of orographic precipitation: Frontiers in Earth Science, v. 6, p. 20, doi:10.3389/feart.2018.00020.

Saito, H., Nakayama, D., and Matsuyama, H., 2010, Relationship between the initiation of a shallow landslide and rainfall intensity-duration thresholds in Japan: Geomorphology, v. 118, p. 167–175, doi:10.1016/j.geomorph.2009.12.016.

Saito, T., and Rehmsmeier, M., 2015, The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets: PLOS One, v. 10, p. e0118432, doi:10.1371/journal.pone.0118432.

Sandberg, E., 2013, A History of Alaska Population Settlement: Alaska Department of Labor and Workforce Development, 19 pp., Retrieved from https://live.laborstats.alaska.gov/pop/estimates/pub/pophistory.pdf.

Scheevel, C.R., Baum, R.L., Mirus, B.B., and Smith, J.B., 2017, Precipitation thresholds for landslide occurrence near Seattle, Mukilteo, and Everett, Washington: U.S. Geological Survey Open-File Report 2017-1039, 51 pp, doi:10.3133/ofr20171039.

Segoni, S., Piciullo, L., and Gariano, S.L., 2018, A review of the recent literature on rainfall thresholds for landslide occurrence: Landslides, v. 15, p. 1483–1501, doi:10.1007/s10346-018-0966-4.

Shanley, C.S., Pyare, S., Goldstein, M.I., Brinkman, T.J., Edwards, R.T., and Hood, E., 2015, Climate change implications in the northern coastal temperate rainforest of North America: Climatic Change, v.130, p. 155–170, doi:10.1007/s10584-015-1355-9.

Sharma, A.R., and Dery, S.J., 2020, Contribution of atmospheric rivers to annual, seasonal, and extreme precipitation across British Columbia and southeastern Alaska: Journal of Geophysical Research: Atmospheres, v. 125, p. e2019JD031823, doi:10.1029/2019jd031823.

Sharma, A.R., and Déry, S.J., 2019, Variability and trends of landfalling atmospheric rivers along the Pacific Coast of northwestern North America: International Journal of Climatology, v. 40, p. 544–558, doi:10.1002/joc.6227.

Sidle, R.C., 1984, Shallow groundwater fluctuations in unstable hillslopes of coastal Alaska: Zeitschrift für Gletscherkunde und Glazialgeologie, v. 20, p. 79–95.

Sidle, R.C., and Swanston, D.N., 1981, Analysis of a small debris slide in coastal Alaska: Canadian Geotechnical Journal, v. 19, p. 167–174, doi:10.1139/t82-018.

Sitka Sound Science Center, 2016, Sitka Geotask Force Summaries, Retrieved from https://sitkascience.org/research-projects/landslide-research/research/sitka-geotask-force/.

Stan Development Team, 2022, Stan User's Guide Version 2.31, Retreived from https://mc-stan.org.

Swanston, D.N., 1970, Mechanics of avalanching shallow till soils of southeast Alaska: USDA Forest Service Research Paper PNW-103, Pacific Northwest Forest and Range Experiment Station, Institute of Northern Forestry, Forest Service, U.S. Dept. of Agriculture, 17 pp.

Swanston, D.N., and Marion, D.A., 1991, Landslide response to timber harvest in southeast Alaska: 5th Federal Interagency Sedimentation Conference, v. 2, p. 49–56.

Thomas, M.A., Mirus, B.B., Collins, B.D., Lu, N., and Godt, J.W., 2018, Variability in soil-water retention properties and implications for physics-based simulation of landslide early warning criteria: Landslides, v. 15, p. 1265–1277, doi:10.1007/s10346-018-0950-z.

Tongass National Forest, 2017, Tongass National Forest Landslide Areas. U.S. Department of Agriculture Forest Service, Retrieved from https://gis.data.alaska.gov/datasets/usfs::tongass-national-forest-landslide-areas/.

Tufano, R., Cesarano, M., Fusco, F., and Vita, P. De, 2019, Probabilistic approaches for assessing rainfall thresholds triggering shallow landslides. The study case of the peri-Vesuvian area (Southern Italy): Italian Journal of Engineering Geology and Environment, v. 2019, p. 105–110, doi:10.4408/IJEGE.2019-01.S-17.

Tullos, D., Byron, E., Galloway, G., Obeysekera, J., Prakash, O., and Sun, Y.H., 2016, Review of challenges of and practices for sustainable management of mountain flood hazards: Natural Hazards, v. 83, p. 1763–1797, doi:10.1007/s11069-016-2400-3.

U.S. Census Bureau (2021). Population Estimates, July 1, 2011 (V2021) - Sitka city and borough, Alaska. Quick Facts, Retrieved from https://www.census.gov/quickfacts/fact/table/sitkacityandboroughalaska/PST045221

Vascik, B.A., Booth, A.M., Buma, B., and Berti, M., 2021, Estimated amounts and rates of carbon mobilized by landsliding in old-growth temperate forests of SE Alaska: Journal of Geophysical Research: Biogeosciences, v. 126, p. e2021JG006321, doi:10.1029/2021JG006321.

Vehtari, A., Gelman, A., and Gabry, J., 2017, Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC: Statistics and Computing, v. 27, p. 1413–1432, doi:10.1007/s11222-016-9696-4.

Wendler, G., Galloway, K., and Stuefer, M., 2016, On the climate and climate change of Sitka, southeast Alaska: Theoretical and Applied Climatology, v. 126, p. 27–34, doi:10.1007/s00704-015-1542-7.

White, C., Gehrels, G.E., Pecha, M., Giesler, D., Yokelson, I., McClelland, W.C., and Butler, R.F., 2016, U-Pb and Hf isotope analysis of detrital zircons from Paleozoic strata of the southern Alexander terrane (southeast Alaska): Lithosphere, v. 8, p. 83–96, doi:10.1130/L475.1.

Wicki, A., Lehmann, P., Hauck, C., Seneviratne, S.I., Waldner, P., and Stähli, M., 2020, Assessing the potential of soil moisture measurements for regional landslide early warning: Landslides, v. 17, p. 1881–1896, doi:10.1007/s10346-020-01400-y.

# 3 GLOBALLY SIMILAR RAINFALL THRESHOLDS FOR URBAN LANDSLIDES[2]

## SUMMARY

Today, more than half of people live in urban areas and 68% are projected to do so by 2050[1]. As growing cities expand into steeper, unstable terrain, residents are increasingly exposed to hazards such as rainfall-triggered landslides[2–8]. Despite the pressing need for operational forecasts[9,10], few cities have established landslide early warning systems[11], and the variability of triggering rainfall conditions between cities has yet to be systematically quantified at a global scale. Here, we estimate rainfall intensity-duration (I-D) thresholds for landsliding in 26 cities worldwide and their spread around a global mean threshold for urban landslides. We find that urban landslides occurred at lower critical rainfall intensities than previously reported global thresholds[12,13]. Yet, landslides were triggered above similar thresholds in cities despite widely varying climatic and topographic settings; median thresholds were indistinguishable from the global threshold in 77% of cities. Our results suggest that human hillslope modification and insufficiently maintained urban infrastructure promote failures on urban hillslopes, which is consistent with our observation that 31% of urban landslides occurred well below annual rainfall extremes. We infer that urbanization likely harmonizes rainfall thresholds between cities, overprinting natural variability, and provide a baseline threshold that could be considered for early warning in cities with limited landslide inventory data.

[2] Luna, L. V., Arango Carmona, M. I., Veh, G., Lewis, E., Ozturk, U., Korup, O. Globally similar rainfall thresholds for urban landslides. *Under consideration at Nature.*

## 3.1 INTRODUCTION

Rainfall-triggered landslides kill over 4,000 people annually on average, ranking with floods, storms, and earthquakes among the world's deadliest natural hazards[14,15]. Landslides are particularly destructive in densely populated urban areas[5,16]: for example, exceptional rainfall triggered numerous landslides in Petrópolis, Brazil in February, 2022, causing 231 fatalities and destroying 60 homes[16]. Shortly after, catastrophic landslides and flooding in the Durban, South Africa metropolitan region in April claimed 448 lives, devastated over 12,000 homes, and displaced over 40,000 people[17]. As cities grow to meet increasing population pressures, buildings and infrastructure often expand into steeper and potentially less stable terrain[4,8], driving projected future increases in landslide risk.[2] For example, the area exposed to landslides in Medellín, Colombia grew by 250% between 1994 and 2018, outpacing the total urban growth (33%) by nearly an order of magnitude.[7]

In cities, humans have modified most hillslopes from their natural states, and quantitative studies have found that urbanization can both heighten[3,6,8,18–21] and mitigate[22] rainfall-triggered landslide activity. Clearing vegetation, cutting and filling to create space for buildings, and loading hillslopes with infrastructure can decrease stability[6,8,19,20], whereas geotechnical engineering measures, such as grading slopes or installing retaining walls, aim to increase stability[23]. Most rainfall-triggered shallow landslides are set in motion by transient increases in pore water pressure that temporarily reduce hillslope strength[24]. Surface sealing or slope de-watering measures may lead to drier hillslope hydrological conditions, whereas leaky pipes, malfunctioning sewage systems, or over-watering lawns may elevate soil water content even during dry periods[19]. During rainstorms, concentrated runoff from paved surfaces, for example due to broken curbs or blocked culverts, can raise pore water pressure, whereas functioning storm sewer systems can route water away from unstable slopes.

For people living and governing in landslide prone areas, knowing when failures are most likely can save lives by informing strategies to mitigate risk[9,10,25]. Accordingly, a key research question over the past decades has been to quantify how much rain is needed to trigger landslides in pursuit of operational forecasts[12,26]; intensity-duration (I-D) thresholds[13] remain a popular empirical tool to identify critical rainfall conditions associated with reported landslides[26]. So far, few cities have established landslide early warning systems[11], and limited multi-temporal landslide inventory data has been one barrier to estimating thresholds for other cities. Rainfall thresholds have been found to vary widely between regions due to differences in environmental factors that influence slope stability like lithology, topography, and climate[12,27,28]. However, previous studies estimating thresholds for individual cities have applied diverging methodologies[26,29–31], complicating comparison, whereas those determining global thresholds have pooled data from rural and urban areas[12,13], such that we lack a systematic global appraisal for cities. Hence, cities with scarce or non-existent landslide inventory data have few options to learn from other areas with more available information. Whether and how

much thresholds vary between cities, as well as the validity of rural thresholds for urban areas, remain open questions that need reliable answers to inform early warning.

## 3.2 LEARNING RAINFALL THRESHOLDS

Here, we address these questions by systematically assessing rainfall thresholds for urban landslides in a global context. We compiled 4,251 urban, rainfall-triggered landslides with a reported date and location in 636 cities from 10 publicly available landslide inventories (Fig. 3.1, *Methods*). Where documented, most of these failures, which we broadly term "landslides," were shallow slides involving earth materials (89%)[32], along with debris flows (7%), rockfalls (3%), and complex movements (1%). We defined urban areas based on the Global Human Settlement Layer Urban Centre Database (GHS-UCDB), which mapped urban centers at 1 km resolution based on satellite imagery of built area and estimated population density[33]. We refer to these areas as "cities," although they may differ in places from administrative boundaries. For each landslide, we determined the triggering event rainfall with hourly station-based precipitation data from the nearest gauge within 25 km with available data from the Global Sub-Daily Rainfall Dataset (GSDR)[34,35], supplemented with observations for Medellín, Colombia and Durban, South Africa to increase coverage of cities in South America and Africa (*Methods*).

Rainfall thresholds are often expressed as approximate quantiles of intensity (mm $hr^{-1}$) conditioned on duration. We defined the triggering event duration $D$ as the time difference between the onset of rainfall and the peak hourly rainfall on the day of the landslide; the intensity $I$ is the cumulative rainfall over this period divided by the duration. We distinguished individual rainfall events using at least 48-hour dry intervals. We applied Bayesian multi-level quantile regression to estimate 10$^{th}$ percentile ($q_{10}$), median ($q_{50}$), and 90$^{th}$ ($q_{90}$) percentile I-D thresholds for all cities with at least five reported landslides and available rainfall data (26 cities with 1,216 landslides) (*Methods*). For each city, quantile regression estimates a threshold that, for a given rainfall duration, separates a specified fraction of landslides according to the recorded rainfall intensity. For example, an estimated 90% of reported landslides lie above the $q_{10}$ line for a given duration, which we define here as the least extreme landslide triggering events; similarly, 10% lie above the $q_{90}$ line, which delineates extreme landslide triggers. The multi-level model also estimates global mean threshold(s) across the 26 cities. We constructed one multi-level model for each quantile in a Bayesian implementation that returns full posterior distributions of thresholds (*Methods)*. We used the global threshold estimated by Guzzetti et al., 2008 (ref.[12]) from rural and urban data as an approximate prior for our models, and evaluated whether and by how much our posterior urban thresholds differed from this prior. We compared city-level thresholds at the global mean rainfall duration for urban landslides ($\bar{x}$), and defined thresholds to be credibly distinguishable when their posterior 95% Highest Density Intervals (HDI) do not overlap. We assessed whether and by how much city-level posterior thresholds differ from each other and compared these thresholds with respect to local mean annual

precipitation, topographic relief, and UN income group[36]. We also compared thresholds to annual rainfall maxima from nearby stations to evaluate threshold estimates in the context of the highest observed rainfall values in these cities.

## 3.3 RAINFALL-TRIGGERED URBAN LANDSLIDES WORLDWIDE

Our compilation shows that rainfall-triggered landslides have occurred in cities on every continent (Fig. 3.1). Compared to the 13,135 cities in the GHS-UCDB, the 636 cities with reported landslides tended to be located in higher relief areas in equatorial, temperate, and polar climate zones with at least 1500 mm of mean annual precipitation (Fig. 3.1c). Cities with landslides were less frequent in arid and snowy climate zones and low relief areas. The cities with the most landslides in our database were Seattle, USA, which had 993 reported landslides between 1897 and 2018 and Medellín, Colombia, which had 1092 reported landslides between 1880 and 2022.

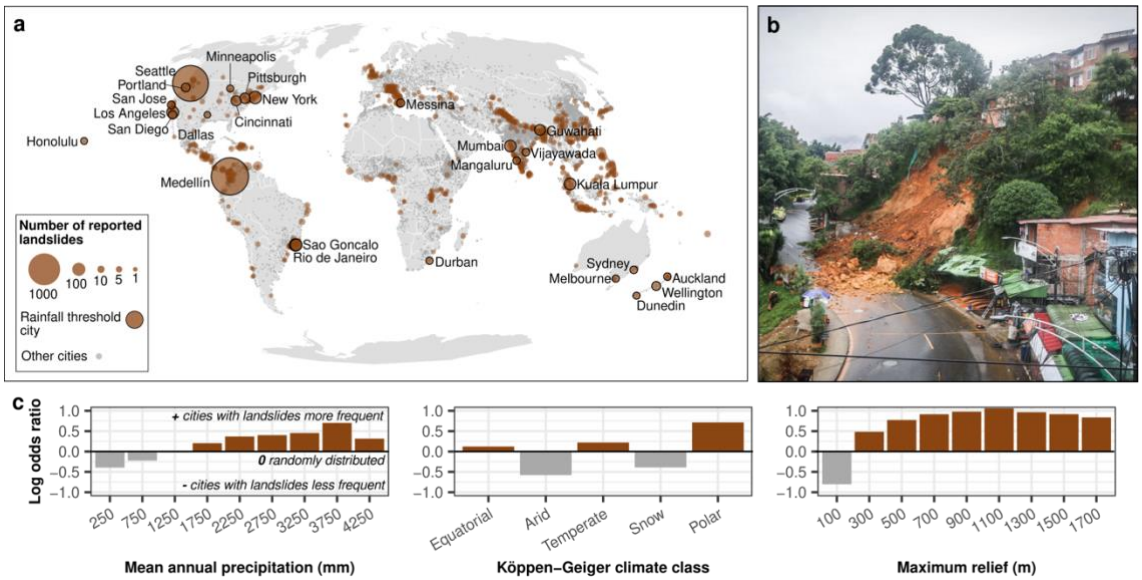



**Figure 3.1. Rainfall-triggered landslides in cities worldwide**. a, Locations of cities with reported rainfall-triggered landslides (brown circles) and all other cities in the GHS-UCDB database (gray points). Cities for which we estimated rainfall intensity-duration thresholds are labeled and outlined in black. b, An urban landslide in Medellín, Colombia in June, 2022. Photo by Carlos Augusto Restrepo. c, The distribution of cities with reported rainfall-triggered landslides compared to all cities in the GHS-UCDB database with respect to mean annual precipitation, climate class, and maximum topographic relief. Positive log odds ratios indicate that landslide cities were comparatively more frequent than all cities in a bin or category; negative values indicate that landslide cities were less frequent.

In the 26 cities for which we estimated thresholds, landslide triggering rainfall events lasted, on average, 72 hours. At this duration, the global mean $q_{10}$ threshold was 0.3 mm $hr^{-1}$ (0.3 – 0.4, 95% Highest Density Interval (HDI)), $q_{50}$ was 1.0 (0.8 – 1.4) mm $hr^{-1}$ , and $q_{90}$ was 1.8 (1.4 – 2.3) mm $hr^{-1}$ (Fig. 3.2). Rio de Janeiro, Brazil had the highest $q_{50}$ threshold

(2.7 (2.2 – 3.3) mm $hr^{-1}$), indicating that the most rain was needed to trigger half of reported landslides and Medellín had the lowest $q_{50}$ threshold (0.5 (0.4 – 0.6) mm $hr^{-1}$), indicating that the least rain was needed (Extended Data Fig. 3.1 – 3.4).

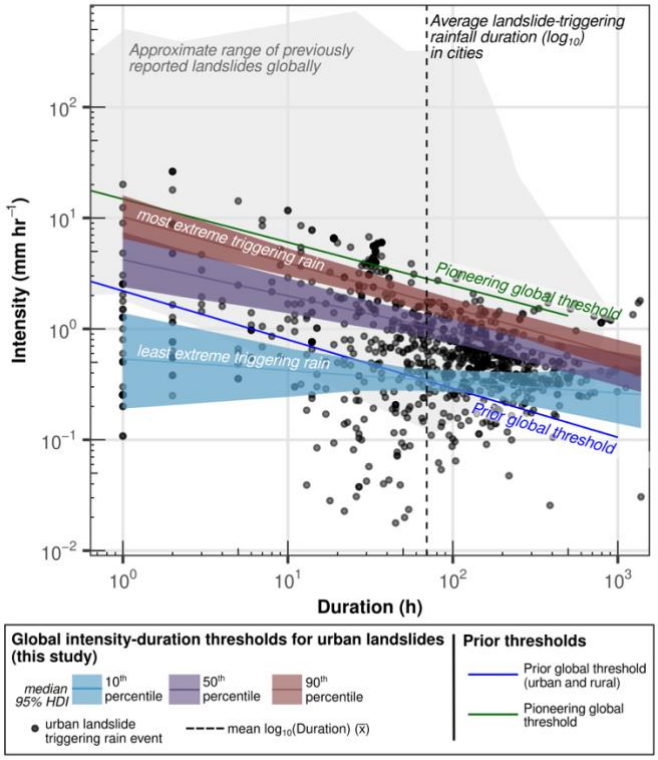



**Figure 3.2. Global intensity-duration thresholds for urban rainfall-triggered landslides**. Posterior global $q_{10}$, $q_{50}$, and $q_{90}$ rainfall intensity-duration thresholds for urban landslides from Bayesian multi-level quantile regression (shaded regions), compared to prior global thresholds from Guzzetti et al., 2008 (ref.) (blue line), and Caine, 1980 (ref.) (green line). Lines show the posterior median; shaded regions show the 95% highest density interval. Black points show the intensity and duration of landslide triggering rainfall events in cities (n = 1216). Gray shaded area is the approximate range of landslides reported globally (ref.).

I-D thresholds are commonly found to have negative slopes[12,13,29,37,38], reflecting the concept that at shorter durations, more intense rainfall is needed to trigger a landslide, whereas at longer durations, less intense rain that accumulates over time can cause failures. However, we found that in cities, the slope of the global $q_{10}$ threshold was indistinguishable from zero, and thus largely unrelated to duration (Fig. 3.2, Extended Data Fig. 3.5). Accordingly, intensities greater than ~0.3 mm $hr^{-1}$ were sufficient to trigger more than 10% of urban landslides, regardless of the rainfall duration. For rainfall lasting between one hour and three days, $q_{10}$ is below the corresponding prior global threshold[12], showing that less intense rain triggered landslides in cities than previously reported globally. Overall, 16% of urban landslides in our database occurred below the prior global threshold; these would have been missed alarms (false negatives) if that threshold had been used for warning. Both the $q_{50}$ and $q_{90}$ thresholds showed negative slopes; the latter is lower than the first established global threshold for shallow

landslides and debris flows[13], indicating that, on average across cities, 90% of landslides occurred below this threshold.

Previous global thresholds compiled data from studies that used differing methods to define event rainfall, noting some inconsistencies in documentation[12,13]. A strength of our approach is that we applied a consistent method to define event rainfall and estimate thresholds for all cities. We tested the sensitivity of our results to the length of the dry period that defines the beginning of the event. Choosing a three-hour dry period as opposed to the 48 hour results presented above lead to a shorter average event duration (6.5 hours vs 72 hours) but only marginally higher thresholds. For example, at a duration of 10 hours, the global $q_{10}$ estimate was 1.1 (0.9 – 1.3) mm $hr^{-1}$ as compared to 0.4 (0.3 – 0.7) mm $hr^{-1}$ (Extended Data Fig. 3.6). The slope of $q_{10}$ becomes positive, such that the threshold becomes higher at longer durations.

We argue that the capacity of comparatively low intensity rainfall events to trigger landslides in urban areas partly reflects anthropogenic processes that have modified slopes in cities and leads to the mismatch between urban rainfall thresholds and prior global thresholds[12,13]. For example, when leaks from under maintained municipal water systems saturate slopes before it begins raining, a lower intensity event can trigger a landslide. Additionally, sealed surfaces may concentrate flow that causes landslides during events that would not have triggered failures in rural areas. Physics-based modeling of fatal landslides in Campos do Jordão, Brazil in 2000, for example, showed that pipe leakage before the triggering rainfall event, combined with cut slopes and construction activity, lead to slope failures that would not have occurred in undisturbed areas.[19] Similarly, independent statistical modeling of landslide activity in the San Francisco Bay Area, USA, found that landslide activity was higher in urban areas than in rural areas for a given amount of cumulative rainfall[18]. We found that at the average duration of three days, similar intensities were needed to trigger landslides in urban areas as previously reported globally. At this duration, storms may deliver sufficient rain to saturate slopes and trigger landslides in both rural and urban areas.

## 3.4 DISTRIBUTION OF THRESHOLDS AMONG CITIES

While previous work has argued that thresholds cannot be exported from one region to another because of differences in environmental factors affecting slope stability,[27,28] we found that $q_{50}$ thresholds for 77% of cities are indistinguishable from the global threshold at the average duration (Fig. 3.3, Extended Data Fig. 3.3). We found no credible differences between thresholds in the cities with the highest (Kuala Lumpur, Malaysia, 2570 mm/yr) and lowest mean annual precipitation (Tijuana, Mexico, 180 mm/yr), and note the lack of trends in thresholds with increasing mean annual precipitation (Fig. 3.3b, Extended Data Fig. 7), climate zone (Fig. 3.3c), or topographic relief (Fig. 3.3d).

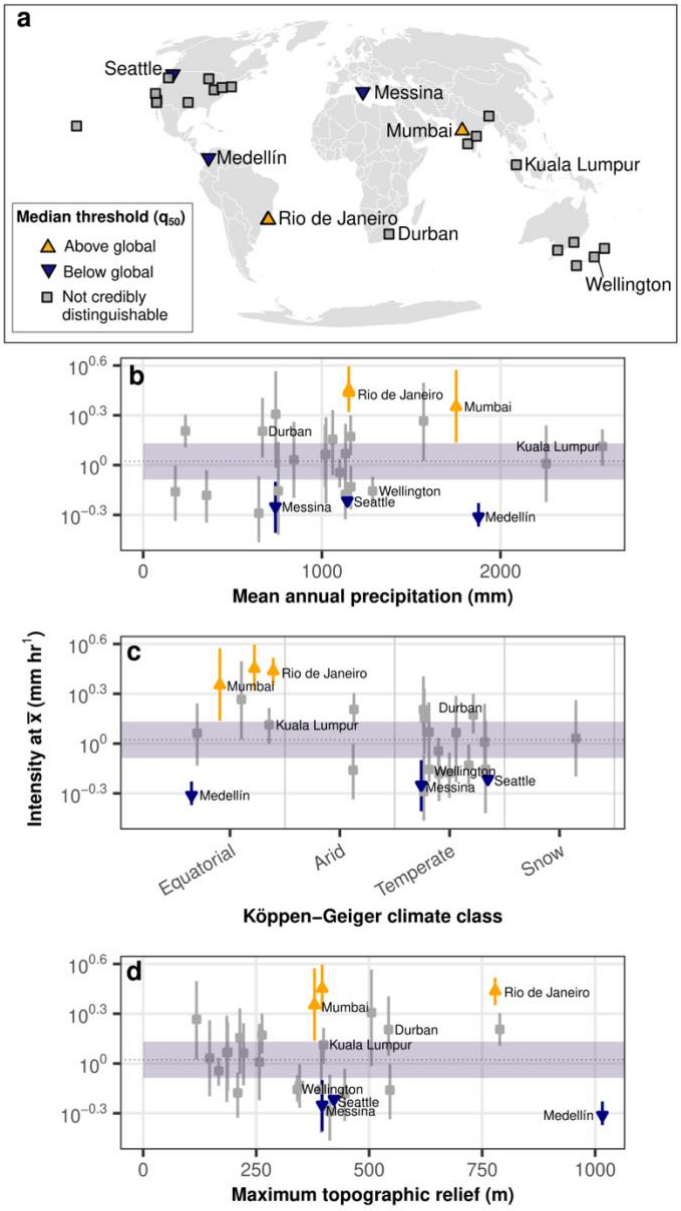



**Figure 3.3. City-level thresholds compared to global threshold and local environmental conditions**. a, Map of threshold cities, indicating whether the q50 threshold was above, below, or indistinguishable from the global mean threshold at average duration ($\bar{x}$). Labeled cities correspond to example cities from each continent in Figure 3.4. b, c, d Posterior parameter estimates for $q_{50}$ threshold intensity at the average landslide-triggering rainfall event duration (72 hours) compared to mean annual precipitation, climate class, and maximum topographic relief for each city (n = 26). The points are the posterior median and the bars are the 95% HDI. The purple dotted line and shaded area are the median and 95% HDI estimates for the global $q_{50}$ threshold.

In natural settings, hillslopes adjust to local climate conditions through preferential removal of material that is prone to failure during moderate, frequently storms, requiring more extreme storms to trigger further landslides[39,40]. This concept has been proposed to explain spatial patterns of landsliding during Typhoon Talas in Japan, for example, where rainfall anomaly was shown to predict landslide activity better than

accumulated precipitation during the storm[41]. Accordingly, intensity-duration thresholds have been found to be higher in regions with greater mean annual precipitation[12,37,38,42]. Our results, however, show that urban areas have rainfall thresholds that are inconsistent with these findings, and instead suggest that thresholds are similar among cities regardless of local climate and topography.

One explanation for the similarity in thresholds between cities is that urban hillslopes are not adjusted to local environmental conditions, but rather to urbanization activities. It is plausible that slopes that have been cut for housing development would fail at similar rainfall intensity and duration in any city because the similar steep slope angles and loads from buildings overprint any previously existing natural differences in environmental conditions. Fill slopes, in which material is added to a slope to create a locally level surface, could be expected to fail similarly in Kuala Lumpur, Malaysia or San Diego, USA, their stability determined not by geologic processes over millennial timescales[43,44] but rather by material strength and geotechnical engineering measures. In this way, human hillslope modification may exert a primary control on rainfall thresholds in built-up, densely populated areas, overshadowing natural environmental factors.

Although it is possible that cities with more extensive geotechnical engineering geared to reduce slope instability could have higher thresholds, we did not observe a systematic difference between cities in different UN income classes (Extended Data Fig. 3.7), which could broadly approximate available resources for hazard mitigation. In China, a national scale statistical analysis of landslide activity showed that urbanization reduced the number of landslides in cities compared to rural areas, attributed to more effective engineering measures in cities, but the impact of the amount of funding for landslide protection measures was ambiguous[22].

## 3.5 URBAN LANDSLIDE TRIGGERING RAIN IN CONTEXT

To evaluate our estimated thresholds in the context of rainfall extremes, we compared thresholds for each city to the maximum intensity rainfall events recorded at gauges in that city across a range of durations for each year on record (Fig. 3.4, Extended Data Fig. 3.1). We found that in most cities, $q_{10}$ and $q_{50}$ thresholds fell below the lowest yearly maxima at shorter durations. At 12-hour duration, for example, the median estimated $q_{10}$ thresholds were below the yearly maxima in 92% of cities; $q_{50}$ thresholds in 73% of cities. In Seattle, for instance, $q_{10}$ was estimated at 0.2 (0.2 – 0.2) mm $hr^{-1}$ and the $q_{50}$ threshold at 0.9 (0.8 – 1.0) mm $hr^{-1}$, whereas the maximum annual rainfall intensities ranged from 1.7 mm $hr^{-1}$ to 8.3 mm $hr^{-1}$. This indicates that the $q_{10}$ and $q_{50}$ thresholds have been exceeded at least annually in these cities over the period of record, and we infer that landslide-triggering rainfall events in cities are not exclusively tied to extremes. At longer durations, however, thresholds in fewer cities were below the yearly maxima, for example, at 100-hour duration, only 15% of cities had $q_{50}$ thresholds below this range. These results are consistent with previous observations that I-D thresholds display

shallower slopes than rainfall intensity-duration-frequency curves, resulting from hillslope hydrological processes[45]. Overall, 31% of reported landslides fell below the envelope defined by the yearly maxima, showing that moderate intensity events triggered a substantial share of urban landslides. In contrast, the $q_{90}$ thresholds fell among the annual maxima in most cities, showing that the most extreme landslide triggering events were often associated with the highest rainfall intensities on record. While we acknowledge that the spatial variability of intense rainfall remains a persistent challenge in measuring landslide-triggering rain[46], we did not find a systematic correlation between rainfall intensity and distance to the gauge (Extended Data Fig. 3.9).

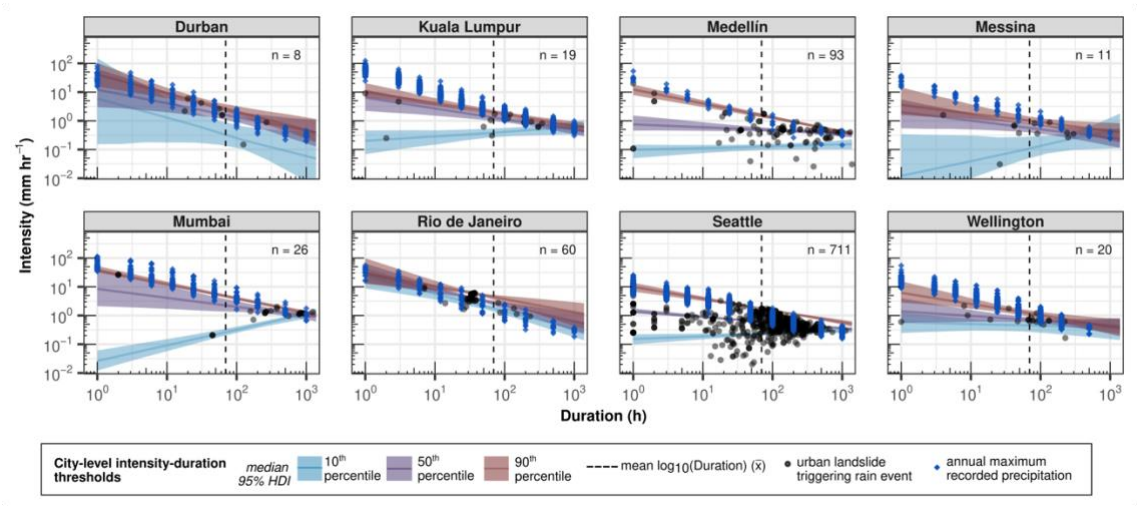



**Figure 3.4. City-level thresholds in context.** Estimated $q_{10}$, $q_{50}$, and $q_{90}$ thresholds for example cities from each continent, compared to the maximum recorded precipitation at nearby gauges for each year on record at a range of durations (blue diamonds). The black points show reported landslide triggering rain events, n refers to the number of reported landslides. Plots for all 26 cities are shown in Extended Data Fig. 3.1.

The capacity of frequent, moderate intensity rainfall events to trigger landslides supports our interpretation that at least some hillslopes in cities are not adjusted to local climate conditions. It also presents a challenge for landslide early warning. To be effective, a landslide early warning system must balance false alarms and missed alarms[47]. Although the $q_{10}$ thresholds estimated here by design would have missed an estimated 10% of landslides at a given duration if used for warning, they would have also been exceeded at least annually by rain events that triggered no landslides, leading to false alarms. Alternative hydro-meteorological thresholds that incorporate antecedent hillslope conditions as well as triggering rainfall have been proposed as a technique to reduce such false alarms in landslide early warning systems.[45,48,49] Another key challenge for effective warning in cities will be to quantify the impact that urban hydrology has in causing landslides during moderate rain events.

## 3.6 DISCUSSION AND IMPLICATIONS

Anthropogenic landscape modification has emerged as a major geomorphic force in the past century[50], and growing evidence suggests that urbanization alters rainfall-triggered landslide activity compared to rural areas by changing hillslope susceptibility and modulating the hydro-meteorological processes that trigger slope failures[3,6,18–20,22]. Our results indicate that urbanization may overprint local environmental factors, leading to similar thresholds among cities, and show that global thresholds for urban landslides are lower than previously reported thresholds that relied on data from urban and rural areas. Thresholds for rural areas may thus be too optimistic for warning in cities, leading to missed alarms, but so far, regional and global early warning efforts have yet to incorporate potential differences in landslide response to triggering rainfall between urban and rural areas[42,51,52]. For example, NASA's global, real-time Landslide Hazard Assessment for Situational Awareness (LHASA V2) takes topography and geology into account, but does not yet include land use[51]; future updates could assess whether introducing urbanization improves predictive performance.

Both urbanization and climate change have already intensified rainfall[53,54]. While our results show that the most extreme landslide triggering rain fell among the highest events on record in cities, 31% of reported landslides were triggered during moderate rainfall events. Such "everyday disasters"[55] thus make up a substantial share of urban landslides. I-D thresholds describe the distribution of rainfall conditions that have triggered landslides in the past, but they disclose little about the probability or frequency of landsliding. If, as projected, the frequency or magnitude of intense rainfall events rise in the future[56], rainfall thresholds may be crossed more often and more intense storms may increase landslide probability[57]. However, urbanization activities may play a more important role than climate change in setting landslide hazard in cities[58].

Although rainfall-triggered landslides pose a hazard in hundreds of cities worldwide (Fig. 3.1), only four metropolitan areas currently benefit from a dedicated, operating landslide early warning system: Hong Kong Special Administrative Region, China, Rio de Janeiro, Brazil, Seattle, USA, and Chittagong, Bangladesh, while few others are covered by regional or national systems[11]. We estimated I-D thresholds for 24 urban areas without a landslide early warning system. Additionally, our global thresholds for urban landslides could serve as a baseline threshold for warning in cities with limited landslide inventory data and the similarity between thresholds in cities suggests that urban areas with sparse local data can use information from other cities to inform warning efforts. Our multi-level modeling structure is able to estimate thresholds for cities with few, or even no, reported landslides; these estimates are informed by thresholds in other urban areas with more available data. Such thresholds can be evaluated and updated as additional local data becomes available; in particular, different stakeholder groups may have differing preferences for the balance of false alarms and missed alarms in an operational system[47]. Expanding early warning coverage to additional cities in the future can allow more people to stay safer during high landslide hazard periods, potentially saving lives[9,25]

## 3.7 MAIN REFERENCES


1. United Nations. *World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420)*. https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf (2019).

2. Ozturk, U. *et al.* How climate change and unplanned urban sprawl bring more landslides. *Nature* **608**, 262–265 (2022).

3. Dille, A. *et al.* Acceleration of a large deep-seated tropical landslide due to urbanization feedbacks. *Nat. Geosci.* **15**, 1048–1055 (2022).

4. Depicker, A. *et al.* Historical dynamics of landslide risk from population and forest-cover changes in the Kivu Rift. *Nat Sustain* **4**, 965–974 (2021).

5. Cui, Y. *et al.* The cost of rapid and haphazard urbanization: lessons learned from the Freetown landslide disaster. *Landslides* **16**, 1167–1176 (2019).

6. Holcombe, E. A., Beesley, M. E. W., Vardanega, P. J. & Sorbie, R. Urbanisation and landslides: hazard drivers and better practices. *Proceedings of the Institution of Civil Engineers - Civil Engineering* **169**, 137–144 (2016).

7. Kühnl, M., Sapena, M., Wurm, M., Geiß, C. & Taubenböck, H. Multitemporal landslide exposure and vulnerability assessment in Medellín, Colombia. *Nat Hazards* (2022) doi:10.1007/s11069-022-05679-z.

8. Smyth, C. G. & Royle, S. A. Urban landslide hazards: incidence and causative factors in Niterói, Rio de Janeiro State, Brazil. *Applied Geography* **20**, 95–118 (2000).

9. Schuster, R. L. & Highland, L. M. The Third Hans Cloos Lecture. Urban landslides: socioeconomic impacts and overview of mitigative strategies. *Bull Eng Geol Environ* **66**, 1–27 (2007).

10. Larsen, M. C. Rainfall-triggered landslides, anthropogenic hazards, and mitigation strategies. in *Advances in Geosciences* vol. 14 147–153 (Copernicus GmbH, 2008).

11. Guzzetti, F. *et al.* Geographical landslide early warning systems. *Earth-Science Reviews* **200**, 102973 (2020).

12. Guzzetti, F., Peruccacci, S., Rossi, M. & Stark, C. P. The rainfall intensity–duration control of shallow landslides and debris flows: an update. *Landslides* **5**, 3–17 (2008).

13. Caine, N. The Rainfall Intensity - Duration Control of Shallow Landslides and Debris Flows. *Geografiska Annaler: Series A, Physical Geography* (1980).

14. Froude, M. J. & Petley, D. N. Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences* **18**, 2161–2181 (2018).

15. Haque, U. *et al.* The human cost of global warming: Deadly landslides and their triggers (1995–2014). *Science of The Total Environment* **682**, 673–684 (2019).

16. Alcântara, E. *et al.* Deadly disasters in Southeastern South America: Flash floods and landslides of February 2022 in Petropolis, Rio de Janeiro. *Natural Hazards and Earth System Sciences Discussions* 1–27 (2022) doi:10.5194/nhess-2022-163.

17. ReliefWeb. *South Africa: Floods and Landslides - Apr 2022 | ReliefWeb*. https://reliefweb.int/disaster/fl-2022-000201-zaf (2022).

18. Johnston, E. C. *et al.* Quantifying the Effect of Precipitation on Landslide Hazard in Urbanized and Non-Urbanized Areas. *Geophysical Research Letters* **48**, e2021GL094038 (2021).

19. Mendes, R. M., de Andrade, M. R. M., Tomasella, J., de Moraes, M. A. E. & Scofield, G. B. Understanding shallow landslides in Campos do Jordão municipality Brazil: disentangling the anthropic effects from natural causes in the disaster of 2000. *Natural Hazards and Earth System Sciences* **18**, 15–30 (2018).

20. Bozzolan, E., Holcombe, E., Pianosi, F. & Wagener, T. Including informal housing in slope stability analysis – an application to a data-scarce location in the humid tropics. *Natural Hazards and Earth System Sciences* **20**, 3161–3177 (2020).

21. Wasowski, J. Understanding rainfall-landslide relationships in man-modified environments: a case-history from Caramanico Terme, Italy. *Environmental Geology* **35**, 197–209 (1998).

22. Li, G., Lei, Y., Yao, H., Wu, S. & Ge, J. The influence of land urbanization on landslides: An empirical estimation based on Chinese provincial panel data. *Science of The Total Environment* **595**, 681–690 (2017).

23. Choi, K. Y. & Cheung, R. W. M. Landslide disaster prevention and mitigation through works in Hong Kong. *Journal of Rock Mechanics and Geotechnical Engineering* **5**, 354–365 (2013).

24. Iverson, R. M. Landslide triggering by rain infiltration. *Water Resources Research* **36**, 1897–1910 (2000).

25. Pollock, W. & Wartman, J. Human Vulnerability to Landslides. *GeoHealth* **4**, e2020GH000287 (2020).

26. Segoni, S., Piciullo, L. & Gariano, S. L. A review of the recent literature on rainfall thresholds for landslide occurrence. *Landslides* **15**, 1483–1501 (2018).

27. Crosta, G. Regionalization of rainfall thresholds: an aid to landslide hazard evaluation. *Environmental Geology* **35**, 131–145 (1998).

28. Segoni, S., Rosi, A., Rossi, G., Catani, F. & Casagli, N. Analysing the relationship between rainfalls and landslides to define a mosaic of triggering thresholds for regional-scale warning systems. *Natural Hazards and Earth System Sciences* **14**, 2637–2648 (2014).

29. Godt, J. W., Baum, R. L. & Chleborad, A. F. Rainfall characteristics for shallow landsliding in Seattle, Washington, USA. *Earth Surface Processes and Landforms* **31**, 97–110 (2006).

30. Garcia-Urquia, E. & Axelsson, K. Rainfall thresholds for the occurrence of urban landslides in Tegucigalpa, Honduras: an application of the critical rainfall intensity. *Geografiska Annaler: Series A, Physical Geography* **97**, 61–83 (2015).

31. Althuwaynee, O. F., Pradhan, B. & Ahmad, N. Estimation of rainfall threshold and its use in landslide hazard mapping of Kuala Lumpur metropolitan and surrounding areas. *Landslides* **12**, 861–875 (2015).

32. Hungr, O., Leroueil, S. & Picarelli, L. The Varnes classification of landslide types, an update. *Landslides* **11**, 167–194 (2014).

33. Florczyk, A. *et al.* Description of the GHS Urban Centre Database 2015. (2019) doi:10.2760/037310 (online).

34. Lewis, E. *et al.* GSDR: A Global Sub-Daily Rainfall Dataset. *Journal of Climate* **32**, 4715–4729 (2019).

35. Lewis, E. *et al.* Quality control of a global hourly rainfall dataset. *Environmental Modelling & Software* **144**, 105169 (2021).

36. Florczyk, A. *et al. GHS Urban Centre Database 2015, multitemporal and multidimensional attributes, R2019A, v1.2.* https://data.jrc.ec.europa.eu/dataset/53473144-b88c-44bc-b4a3-4583ed1f547e (2019).

37. Saito, H., Nakayama, D. & Matsuyama, H. Relationship between the initiation of a shallow landslide and rainfall intensity—duration thresholds in Japan. *Geomorphology* **118**, 167–175 (2010).

38. Zêzere, J. L. *et al.* Rainfall thresholds for landslide activity in Portugal: a state of the art. *Environ Earth Sci* **73**, 2917–2936 (2015).

39. Iida, T. Theoretical research on the relationship between return period of rainfall and shallow landslides. *Hydrological Processes* **18**, 739–756 (2004).

40. Benda, L. & Dunne, T. Stochastic forcing of sediment supply to channel networks from landsliding and debris flow. *Water Resources Research* **33**, 2849–2863 (1997).

41. Marc, O., Gosset, M., Saito, H., Uchida, T. & Malet, J.-P. Spatial Patterns of Storm-Induced Landslides and Their Relation to Rainfall Anomaly Maps. *Geophysical Research Letters* **46**, 11167–11177 (2019).

42. Peruccacci, S. *et al.* Rainfall thresholds for possible landslide occurrence in Italy. *Geomorphology* **290**, 39–57 (2017).

43. Campforts, B., Shobe, C. M., Overeem, I. & Tucker, G. E. The Art of Landslides: How Stochastic Mass Wasting Shapes Topography and Influences Landscape Dynamics. *Journal of Geophysical Research: Earth Surface* **127**, e2022JF006745 (2022).

44. Montgomery, D. R. & Dietrich, W. E. A physically based model for the topographic control on shallow landsliding. *Water Resources Research* **30**, 1153–1171 (1994).

45. Bogaard, T. & Greco, R. Invited perspectives: Hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: proposing hydro-meteorological thresholds. *Natural Hazards and Earth System Sciences* **18**, 31–39 (2018).

46. Destro, E. *et al.* Spatial estimation of debris flows-triggering rainfall and its dependence on rainfall return period. *Geomorphology* **278**, 269–279 (2017).

47. Intrieri, E., Gigli, G., Casagli, N. & Nadim, F. Brief communication: Landslide Early Warning System: toolbox and general concepts. *Nat. Hazards Earth Syst. Sci.* **13**, 85–90 (2013).

48. Thomas, M. A., Mirus, B. B. & Collins, B. D. Identifying Physics-Based Thresholds for Rainfall-Induced Landsliding. *Geophysical Research Letters* **45**, 9651–9661 (2018).

49. Crozier, M. J. Prediction of rainfall-triggered landslides: a test of the Antecedent Water Status Model. *Earth Surface Processes and Landforms* **24**, 825–833 (1999).

50. Cendrero, A., Forte, L. M., Remondo, J. & Cuesta-Albertos, J. A. Anthropocene Geomorphic Change. Climate or Human Activities? *Earth's Future* **8**, e2019EF001305 (2020).

51. Stanley, T. A. *et al.* Data-Driven Landslide Nowcasting at the Global Scale. *Frontiers in Earth Science* **9**, (2021).

52. Krøgli, I. K. *et al.* The Norwegian forecasting and warning service for rainfall- and snowmelt-induced landslides. *Natural Hazards and Earth System Sciences* **18**, 1427–1450 (2018).

53. Fowler, H. J. *et al.* Anthropogenic intensification of short-duration rainfall extremes. *Nat Rev Earth Environ* **2**, 107–122 (2021).

54. Li, Y. *et al.* Strong Intensification of Hourly Rainfall Extremes by Urbanization. *Geophysical Research Letters* **47**, e2020GL088758 (2020).

55. Bull-Kamanga, L. *et al.* From everyday hazards to disasters: the accumulation of risk in urban areas. *Environment and Urbanization* **15**, 193–204 (2003).

56. Prein, A. F. *et al.* The future intensification of hourly precipitation extremes. *Nature Clim Change* **7**, 48–52 (2017).

57. Gariano, S. L. & Guzzetti, F. Landslides in a changing climate. *Earth-Science Reviews* **162**, 227–252 (2016).

58. Crozier, M. J. Deciphering the effect of climate change on landslide activity: A review. *Geomorphology* **124**, 260–267 (2010).

## 3.8 METHODS

### 3.8.1 Data

**Urban area data**

We defined urban areas using the Global Human Settlement Layer Urban Centre Database (GHS-UCDB), available at http://data.europa.eu/89h/53473144-b88c-44bc-b4a3-4583ed1f547e [36]. These polygons represent the 2015 extent of densely populated built-up areas with at least 50,000 inhabitants; the areas are mapped at 1-km resolution using satellite remote sensing and population data, and do not necessarily coincide with administrative boundaries[33]. We labeled urban centers, which we term "cities" for simplicity, by the main name of the urban center, but note that other cities or parts thereof may also fall into a given polygon. For example, the urban area we label as Seattle, USA, contains portions of Seattle, Tacoma, Bellevue, and Everett.

Information on mean annual precipitation, Köppen-Geiger climate class, and UN income class for each urban area were taken from the GHS-UCDB. We simplified the Köppen-Geiger climate classes following Kottek et al. 2006 (ref.[59]). We calculated maximum relief within an urban area by subtracting the minimum elevation from the maximum elevation value from the 90-m resolution Shuttle Radar Topography Mission (SRTM)[60] digital elevation model using Google Earth Engine.

**Landslide inventory data**

We compiled publicly available landslide inventory data at global, national, and regional scales (Table 3.1). Based on available attributes from each inventory, we assigned common classes for landslide type, material, trigger, and spatial uncertainty to each landslide. We then subset this dataset to landslides that occurred within the urban areas. We considered only landslides with a documented rainfall trigger. To be able to identify landslide-triggering precipitation, we included only landslides with a reported daily timestamp and a positional uncertainty of <25 km. For some inventories where spatial uncertainty was unknown, we estimated it based on the original source of the landslide information; for example, inventories provided as GIS data by a geological survey or similar entity were estimated to have a spatial uncertainty of <5 km. Overall, we compiled 12,248 urban landslides worldwide, 4251 (35%) of which have a documented rainfall trigger, daily timestamp, and spatial uncertainty of <25 km. 636 urban areas have at least one rainfall-triggered landslide meeting these criteria (Fig. 3.1).

**Table 3.1. Summary of landslide inventory data**

| Inventory | Scale | Number of urban, rainfall-triggered landslides with known date and location | Number of landslides in cities for which we estimated thresholds | Data availability |
|---|---|---|---|---|
| **NASA Global Landslide Catalog**[61] | Global | 1761 (41%) | 380 (31 %) | https://doi.org/10.1016/j.geomorph.2015.03.016 |
| **Global Fatal Landslide Database V2**[14] | Global | 244 (5.7%) | 32 (2.6%) | https://doi.org/10.5194/nhess-18-2161-2018 |
| **Geoscience Australia Landslide Search**[62] | National | 11 (0.26%) | 4 (0.3%) | http://pid.geoscience.gov.au/dataset/ga/74273 |
| **Landslide & Torrential Colombia Database**[63] | National | 1121 (26%) | 93 (7.6%) | https://geohazards.com.co/ |
| **FraneItalia V3**[64] | National | 161 (3.8%) | 8 (0.7%) | https://doi.org/10.17632/zygb8jygrw.3 |
| **GNS New Zealand Landslide Database**[65] | National | 26 (0.61%) | 22 (1.8%) | https://doi.org/10.1007/s10346-017-0843-6 |
| **Landslide Inventory Rwanda**[66] | National | 2 (0.047%) | 0 (0%) | https://doi.org/10.4121/15040446.v1 |
| **Kentucky Geological Survey Landslide Inventory**[67] | Regional | 3 (0.071%) | 2 (0.2%) | https://doi.org/10.13023/kgs.ic31.12 |
| **Digital Geodata Series DGS06-3 Landslides in New Jersey**[68] | Regional | 60 (1.4%) | 40 (3.3%) | https://www.state.nj.us/dep/njgs/geodata/dgs06-3.htm |
| **Seattle Historic Landslide Locations ECA**[69] | Regional | 862 (20%) | 635 (52%) | https://data-seattlecitygis.opendata.arcgis.com/datasets/SeattleCityGIS::historic-landslide-locations-eca/about |

## Precipitation data and event rainfall definition

We estimated landslide-triggering rainfall using hourly station-based observations from the Global Sub-Daily Rainfall Dataset (GSDR) of 26,297 gauges worldwide[34]. We used a quality-controlled version of the GSDR as described by Lewis et al. 2021 (ref.[35]). Hourly precipitation data has been shown to be necessary for determining rainfall thresholds for shallow landsliding, as daily data cannot sufficiently capture the short, intense periods of rain responsible for triggering some slope failures[70,71]. To increase coverage of cities in South America and Africa, we supplemented the GSDR with station data for Durban, South Africa and Medellín, Colombia. Data for Durban were obtained from the South African Weather Service (SAWS) and used under license for the current study. These data are available upon reasonable request and with permission of SAWS. Data for Medellín were obtained from the Colombian Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM), were used with permission for this study, and are available at http://dhime.ideam.gov.co/atencionciudadano/.

We defined landslide triggering rainfall events as the interval between the onset of rainfall after a 48-hour dry period with <0.1 mm of cumulative precipitation and the peak hourly rainfall on the day that the landslide occurred. The length of this period is the duration and the total accumulated amount of precipitation is the cumulative event rainfall. Dividing the cumulative event rainfall by the duration gives the intensity. We acknowledge a wide range of similar definitions for rainfall intensity in landslide studies, but stress the consistency of our method across the full dataset. There is no standard approach to define event rainfall in the landslide rainfall threshold literature. While authors agree that the definition should be objective and reproducible[72–74], a departure from early methods that relied on manual definition, they have applied a wide range of techniques[26]. The length of the dry period defining the beginning of the event rainfall is a key modeling choice. Generally, longer dry periods will lead to longer rainfall events with lower average intensities as more short periods of low-intensity rain are incorporated into the event rainfall[74]. This will result in overall lower thresholds, and a threshold must be considered in the context of these modeling choices. Previous studies have chosen dry periods ranging from 8 hours in California, USA to 144 hours in Liguria, Italy[37,72–74]. Here, we chose a 48-hour dry period to define the beginning of a triggering event, and tested the sensitivity of our results to this choice by repeating the analysis with a 3 hour dry period.

We calculated event rainfall from all stations within 25 km of the landslide location, assigning the nearest gauge with available data. We only considered gauges with >90% complete data in the event period and linearly interpolated across any remaining gaps in the time series. In the case that the nearest gauge measured <0.01 mm of precipitation during the event period before the landslide, we assigned the next nearest gauge. We considered all cities with at least five landslides with precipitation data coverage for I-D threshold analysis. This resulted in 1216 landslides in 26 cities with at least five landslides with coverage. 27% of landslides were within 5 km of the selected gauge, and 85% were within 15 km (Extended Data Fig. 3.8); 90% of these landslides had no missing data in the event rainfall period.

The choice of gauge may impact threshold estimates, as the spatial variability of rainfall remains a persistent challenge in identifying rainfall thresholds[46,75,76]. Particularly in areas with high topographic relief or in the case of storms with a small spatial footprint, even nearby rain gauges may capture the landslide triggering rain inaccurately[46,77]. Previous studies have considered gauges within 2 km to over 15 km[28,72,74], while others have proposed using gauge-corrected radar products to overcome this issue.[37,76] Yet such products are unavailable globally. From gauges within a chosen radius, many previous studies have selected the nearest gauge, as we do here, or have applied more complex methods based on distance, accumulated rainfall, and duration. Melillo et al., 2018 (ref.[72]), for example, assigned a weighting factor to each gauge that maximizes closer gauges with higher accumulations and shorter durations, while Segoni et al., 2014 (ref.[73]) selected the gauge with the highest return period rainfall. We chose a nearest neighbor approach for its simplicity and because it does not condition the threshold estimates on higher observed rainfall values.

For each gauge associated with a landslide, we extracted annual block maximum cumulative precipitation for each year on record at durations of 1, 3, 6, 12, 24, 48, 100, 200, 500, and 1000 hours (Fig. 3.4, Extended Data Fig. 3.1). We assigned blocks to the year in which they ended: for example, a 1000-hour period ending on January 1, 2020 is assigned to 2020. We used only years with at least 90% complete data and linearly interpolated any further no data values. We refrained from estimating return periods of landslide triggering rainfall from these data because of the widely varying record lengths (Extended Data Table 3.1) [34].

### 3.8.2 Bayesian multi-level quantile regression for identifying intensity-duration thresholds

**Quantile regression for identifying I-D thresholds**

Rainfall intensity-duration thresholds for shallow landsliding are typically represented as a power law, where the average rainfall intensity ($I$) during a landslide-triggering rainfall event is a function of the event's duration ($D$) [12,13].

$$I = AD^B$$

*Equation 3.1*

After log-transforming, this relationship becomes linear: $\log(I)$ scales linearly with $\log(D)$, $B$ is the (often negative) slope and $\log(A)$ is the intercept.

$$\log(I) = \log(A) + B\log(D)$$

*Equation 3.2*

Quantile regression estimates the conditional quantile of a response variable, in this case $\log(I)$, with respect to a predictor variable, in this case $\log(D)$. Whereas linear regression

estimates the conditional mean of the response variable, quantile regression estimates the conditional median or any other quantile of choice. In essence, quantile regression splits the landslide observations into two parts: for the 10th percentile, for example, an estimated 10% of observed landslides fall below the line and 90% above the line for a given duration. Here, we used quantile regression to estimate the 10th ($q_{10}$), 50th ($q_{50}$), and 90th ($q_{90}$) percentiles of landslide-triggering rainfall intensities with duration to characterize the distribution of rainfall intensities associated with reported landslides. While previous studies have employed a wide range of methods to estimate thresholds[12,13,26,37,74], we note that quantile regression is more robust against outliers than linear regression, which is a useful property in the case that there are errors in the reported landslide dates.[37] Importantly, quantile regression does not require the assumption that residuals are normally distributed, which is hardly the case for observations of landslide-triggering rainfall intensities; the results of quantile regression are also invariant with respect to the log-transformation of both *I* and *D*.

**Bayesian inference**

Bayesian inference updates prior knowledge in light of new information.[78,79] Here, we aimed to update previous knowledge on global I-D thresholds using landslide inventory data from urban areas. Bayes' Rule, which forms the foundation for all Bayesian inference, is stated as:

$$P(\Theta \mid r) = \frac{P(r \mid \Theta)P(\Theta)}{P(r)}$$

*Equation 3.3*

where Θ is the set of parameters that define the I-D threshold and $r$ are our observations of landslide-triggering rain. $P(\Theta|r)$ is the *posterior*, which is a probability distribution across all I-D thresholds that are compatible with the observed data and *prior* knowledge $P(\Theta)$. $P(r)$ is the *marginal or expected likelihood* and serves to normalize the posterior probability distribution. Posteriors provide an intuitive way of expressing parameter uncertainty: rather than obtaining an optimal I-D threshold, the model learns the whole range of plausible thresholds, weighted by their probability. Credibility intervals summarize this posterior distribution. We report the 95% Highest Density Interval throughout, and interpret parameter estimates to be credibly distinguishable when their 95% HDIs do not overlap. Bayesian inference therefore offers one solution to quantifying parameter uncertainty for rainfall thresholds, a task often neglected in landslide research[26]; very few thresholds reported in the literature include any estimate of parameter uncertainty. We used the global I-D threshold reported by Guzzetti et al.[12] as a prior. By comparing the posterior distribution with the prior distribution, we measured how much we have learned by considering new data from urban areas (Extended Data Fig. 3.5).

**Multi-level models for pooling information and quantifying variation between urban areas**

Multi-level models estimate parameters for different groups within one model. We chose city as a grouping variable to be able to learn the variation between cities. The motivation for this choice is that landslide early warning systems can be implemented at the city level and that an individual city has a broadly similar climatic and geologic

setting. As opposed to estimating I-D thresholds for each city individually, our multi-level model estimates I-D thresholds for all cities within one model. An I-D threshold is learned for each city, informed by the I-D thresholds of other cities, thus taking advantage of all data. Also, a global mean I-D threshold across cities is learned, as is the variation between different cities. Multi-level models are well suited to datasets with differing numbers of observations in each group, as is the case here, with hundreds of landslide observations in some cities and fewer than 10 in others. The structure prevents overfitting groups with many data points and generally improves inference for groups with few data points.[80] As the global mean threshold represents the distribution of thresholds across all cities, it is not overly influenced by cities with many data points.

A key advantage of Bayesian inference for estimating I-D thresholds is that the parameter estimates are conditional on the data: reported landslides and estimated rainfall intensities and durations (Equation 3.3). When a city has few reported landslides, the posterior distribution of the parameter estimates will be larger, and vice versa. Although our landslide database is undoubtedly incomplete, our threshold estimates are conditional on the available data and the posterior distributions report threshold uncertainty accordingly. With our model, an I-D threshold can be estimated for a city with few or even no reported landslides; this would then be the global mean threshold informed completely by the other cities. With additional landslide observations, the threshold for a given city will become more specific to that city, if the threshold deviates from the global mean. As the threshold uncertainty is transparently quantified in the posterior, a user can decide what level of uncertainty is necessary for their application (for example in an operational landslide early warning system) and assess if this condition is met with the available data, or if more data should be collected.

**Model set up**

We fit regressions for three quantiles ($q_{10}$, $q_{50}$, and $q_{90}$), resulting in three multi-level models. Within each model, we estimated thresholds for all cities with at least five reported landslides with available precipitation data: 26 cities, and a global mean threshold across these cities. We first log-transformed the intensity ($I$) and duration ($D$) values for each landslide-triggering rain event, and then standardized the data by subtracting the mean across all events and dividing by the standard deviation, such that the mean log-transformed intensity and duration are zero.

Bayesian quantile regression relies on an Asymmetric Laplace (AL) likelihood, in which the quantile of interest $\tau$ is specified and a location parameter ($\mu$) and positive scale parameter ($\kappa$) are learned from the data.

The parameters of the linear model that define the threshold for a given quantile are estimated as follows:

$$z_y(y_{ij}) \sim AL(\mu_{ij}, \kappa, \tau)$$

$$\mu_{ij} = \alpha_j{}' + \beta_j{}' z_x(x_{ij}$$

*Equation 3.4*

where $z_y$ *and* $z_x$ are functions that standardize $\log_{10}$ transformed observations of rainfall intensity ($y$) ($\log(I)$) and duration ($x$) ($\log(D)$). $i$ indicates individual landslides, and $j$ indicates the group: city. $\alpha_j{}'$ and $\beta_j{}'$ are the group-level intercept and slope of the standardized threshold.

In this multi-level model, the parameters for each city are drawn from a multi-variate normal distribution of group-level parameters, which serves to partially pool information across cities.

$$\begin{bmatrix} \alpha_j{}' \\ \beta_j{}' \end{bmatrix} \sim MVNormal\left[\left(\frac{\overline{(\alpha')}}{(\beta')}\right), S\right]$$

$$S = \begin{bmatrix} \sigma_{\alpha'}^2 & \sigma_{\alpha'}\sigma_{\beta'} \\ \sigma_{\alpha'}\sigma_{\beta'} & \sigma_{\alpha'}^2 \end{bmatrix} T$$

$$T = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

*Equation 3.5*

where $\overline{\alpha'}$ and $\overline{\beta'}$ are the population-level means that represent the standardized global mean threshold. The covariance matrix $S$ is composed of the group-level standard deviations $\sigma_{\alpha'}$ and $\sigma_{\beta'}$ and $T$, the correlation matrix with correlation coefficient $\rho$.

Prior distributions are required for $\kappa$, $\overline{\alpha'}$, $\overline{\beta'}$, $\sigma_{\alpha'}$, $\sigma_{\beta'}$, and $T$. We used the global I-D threshold determined by Guzzetti et al. 2008 (ref.[12]) as an approximate prior for the population-level mean parameters. We standardized their reported values to be compatible with our model set-up that used standardized data. Because no estimates of I-D threshold uncertainty were reported, we used the point estimates as the means of Student-$t$ distributions to achieve informative prior distributions for $\overline{\alpha'}$, $\overline{\beta'}$, and conducted a prior predictive check to ensure that these values were reasonable. For $\kappa$, $\sigma_{\alpha'}$, $\sigma_{\beta'}$, we chose weakly informative Student-$t$ priors that are implemented as defaults in the R package brms. For $T$, we chose a non-informative Cholesky Lewandowski-Kurowicka-Joe (LKJ) distribution, which makes all correlation matrices equally likely.

$$\begin{aligned} \kappa &\sim Studentt(3,0,2.5) \\ \bar{\alpha}' &\sim Studentt(3,-0.76,0.1) \\ \bar{\beta}' &\sim Studentt(3,-0.56,0.1) \\ \sigma_{\alpha'} &\sim Studentt(3,0,2.5) \\ \sigma_{\beta'} &\sim Studentt(3,0,2.5) \\ R &\sim LKJCholesky(1) \end{aligned}$$

*Equation 3.6*

We fit the models with the R package brms version 2.17.0[81]. Brms calls Stan, a probabilistic programming language that uses Hamiltonian Monte Carlo to numerically approximate the posterior parameter space[82]. We ran the sampler for 4,000 iterations for four independent chains and discarded the first 1,000 draws as warm up, resulting in 12,000 total draws. No numerical divergences occurred after warm up. In all cases, the Gelman-Rubin potential scale reduction factor R-hat was equal to 1, indicating chain convergence.

After fitting the models, we transformed parameter estimates for $\bar{\alpha}'$, $\bar{\beta}'$, $\alpha_j'$, $\beta_j'$ to the log-scale for easier and more familiar interpretation.

$$\alpha_j = \alpha_j' sd(y_i) + \overline{(y_i)}$$
$$\bar{\alpha} = \bar{\alpha_j} sd(y_i) + \overline{(y_i)}$$

*Equation 3.7*

where $\alpha_j$ and $\bar{\alpha}$ represent the estimated $\log_{10}(I)$ for each city at the average $\log_{10}(D)$ across all landslides ($\bar{x_i}$) and the global mean threshold, respectively. ($\bar{y_i}$) is mean $\log_{10}(I)$ across all landslides.

$$\log_{10}(A_j) = sd(y_i)(\alpha_j' - \beta_j'(\frac{\overline{(x_i)}}{sd(x_i)}) + \overline{(y_i)})$$
$$\log_{10}(\bar{A}) = sd(y_i)(\bar{\alpha}' - \bar{\beta}'(\frac{\overline{(x_i)}}{sd(x_i)}) + \overline{(y_i)})$$

*Equation 3.8*

where $\log_{10}(A_j)$ and $\log_{10}(\bar{A})$ are the estimated $\log_{10}$ intensity for each city and the global mean threshold, respectively, at a $\log_{10}(D)$ of zero, or a 1-hour rainfall duration.

$$B_j = \beta_j = \beta_j'(\frac{sd(y_i)}{sd(x_i)})$$
$$\bar{B} = \bar{\beta} = \bar{\beta}'(\frac{sd(y_i)}{sd(x_i)})$$

*Equation 3.9*

where $B_j = \beta_j$ and $\bar{B} = \bar{\beta}$ are the slope of the thresholds for individual cities and the global mean, respectively, in log-log space, or the estimated order of magnitude change in intensity with one order of magnitude change in duration.

## 3.9 METHODS REFERENCES


59. Kottek, M., Grieser, J., Beck, C., Rudolf, B. & Rubel, F. World Map of the Köppen-Geiger climate classification updated. *metz* **15**, 259–263 (2006).

60. Jarvis, A., Reuter, H. I., Nelson, A. & Guevara, E. *Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database*. https://bigdata.cgiar.org/srtm-90m-digital-elevation-database/.

61. NASA. Global Landslide Catalog. (2018).

62. Geoscience Australia. Landslide Search. (2012).

63. Geohazards Semillero de investigación. Inventario de eventos - Colombia. (2020).

64. Calvello, M. & Pecoraro, G. FraneItalia: a catalog of recent Italian landslides (version 3). (2021) doi:10.17632/ZYGB8JYGRW.3.

65. Rosser, B., Dellow, S., Haubrock, S. & Glassey, P. New Zealand's National Landslide Database. *Landslides* **14**, 1949–1959 (2017).

66. Uwihirwe, J. Data underlying the research of Integration of observed and model derived groundwater levels in landslide threshold models in Rwanda. (2021) doi:https://doi.org/10.4121/15040446.v1.

67. Crawford, M. M. Kentucky Geological Survey Landslide Inventory: From Design to Application. (2014) doi:https://doi.org/10.13023/kgs.ic31.12.

68. New Jersey Geological and Water Survey. Digital Geodata Series DGS06-3 Landslides In New Jersey. (2018).

69. City of Seattle. Historical Landslide Locations ECA. (2020).

70. Gariano, S. L., Melillo, M., Peruccacci, S. & Brunetti, M. T. How much does the rainfall temporal resolution affect rainfall thresholds for landslide triggering? *Nat Hazards* **100**, 655–670 (2020).

71. Marra, F. Rainfall thresholds for landslide occurrence: systematic underestimation using coarse temporal resolution data. *Nat Hazards* **95**, 883–890 (2019).

72. Melillo, M. *et al.* A tool for the automatic calculation of rainfall thresholds for landslide occurrence. *Environmental Modelling & Software* **105**, 230–243 (2018).

73. Segoni, S., Rossi, G., Rosi, A. & Catani, F. Landslides triggered by rainfall: A semi-automated procedure to define consistent intensity–duration thresholds. *Computers & Geosciences* **63**, 123–131 (2014).

74. Staley, D. M., Kean, J. W., Cannon, S. H., Schmidt, K. M. & Laber, J. L. Objective definition of rainfall intensity–duration thresholds for the initiation of post-fire debris flows in southern California. *Landslides* **10**, 547–562 (2013).

75. Iadanza, C., Trigila, A. & Napolitano, F. Identification and characterization of rainfall events responsible for triggering of debris flows and shallow landslides. *Journal of Hydrology* **541**, 230–245 (2016).

76. Nikolopoulos, E. I., Borga, M., Creutin, J. D. & Marra, F. Estimation of debris flow triggering rainfall: Influence of rain gauge density and interpolation methods. *Geomorphology* **243**, 40–50 (2015).

77. Lengfeld, K. *et al.* Use of radar data for characterizing extreme precipitation at fine scales and short durations. *Environ. Res. Lett.* **15**, 085003 (2020).

78. McElreath, R. *Statistical Rethinking A Bayesian Course with Examples in R and STAN*. (Chapman and Hall/CRC Press, 2020).

79. van de Schoot, R. *et al.* Bayesian statistics and modelling. *Nature Reviews Methods Primers* **1**, 1–26 (2021).

80. Gelman, A. & Hill, J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. (Cambridge University Press, 2007).

81. Bürkner, P.-C. brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* **80**, 1–28 (2017).

82. Stan Development Team. Stan Modeling Language Users Guide and Reference Manual, version 2.28. (2022).

## 3.10 ACKNOWLEDGMENTS

L.V.L. was funded by DFG GRK 2043 NatRiskChange. M.I.A.C was funded through BMBF project BB-KI Chips (grant 16DHBKI020). Precipitation data collection for the GSDR (E.L.) was supported by the European Research Council (INTENSE; grant: ERC-2013-CoG-617329) and the Global Energy and Water Exchanges (GEWEX) Hydroclimatology Panel Cross-Cutting project on sub-daily precipitation extremes. We thank Joaquin Ferrer for help with data collection. We thank Daissy Herrera Posada and Edier Aristizábal Giraldo for providing landslide inventory data for Colombia and the South African Weather Service for providing precipitation data for Durban. We especially thank authors and institutions who have made their landslide inventory and precipitation data public, making this work possible.

## 3.11 AUTHOR CONTRIBUTIONS

L.V.L. conceptualized and led the study, performed the statistical analysis, and drafted the original text. L.V.L., M.I.A.C, and E.L. processed the data. L.V.L. and G.V. designed the statistical models and created the figures. G.V., O.K., U.O., M.I.A.C, and E.L. contributed to discussing the results and reviewing the manuscript; L.V.L, G.V., and O.K revised the manuscript.

# 4 Seasonal landslide activity lags annual precipitation pattern in the Pacific Northwest[3]


## Abstract

Seasonal variations in landslide activity remain understudied compared to recent advances in landslide early warning at hourly to daily timescales. Here, we learn the seasonal pattern of monthly landslide activity in the Pacific Northwest from five heterogeneous landslide inventories with differing spatial and temporal coverage and reporting protocols combined in a Bayesian multi-level model. We find that landslide activity is distinctly seasonal, with credible increases in landslide intensity, inter-annual variability, and probability marking the onset of the landslide season in November. Peaks in landslide probability in January and intensity in February lag the annual peak in mean monthly precipitation, and landslide activity is more variable in winter than in summer, when landslides are rare. For a given monthly rainfall, landslide intensity at the season peak in February is up to ten times higher than at the onset in November, underlining the importance of antecedent seasonal hillslope conditions.


[3] Published as:
Luna, L. V., & Korup, O. (2022). Seasonal Landslide Activity Lags Annual Precipitation Pattern in the Pacific Northwest. **Geophysical Research Letters**, 49(18), e2022GL098506. https://doi.org/10.1029/2022GL098506. 

## 4.1 INTRODUCTION

Landslides regularly cause fatalities and damage infrastructure in many areas worldwide (Froude & Petley, 2018), and extensive research has focused on anticipating where landslides are likely to occur (Reichenbach et al., 2018). For people living in susceptible areas, however, better constraining when landslides are likely is key to reducing risk, as taking proper action can increase survival rates (Pollock & Wartman, 2020). Globally, advances in landslide early warning have mostly concentrated on hourly to daily timescales (Baum & Godt, 2010; Guzzetti et al., 2020; Mirus, Becker, et al., 2018; Stanley et al., 2021). Much less attention has been given to quantifying seasonal patterns of landslide activity, which would allow for improved planning and emergency preparedness to be able to quickly react to short-term warnings.

To characterize landslide seasonality, previous studies have, for example, modeled monthly rock fall frequency along a rail corridor in British Columbia (Pratt et al., 2019), investigated seasonal trends in modeled daily landslide hazard in the Pacific Northwest (PNW) of the United States (Stanley et al., 2020), or explored seasonal changes in intensity-duration thresholds for rainfall-triggered landslides in Italy (Napolitano et al., 2016; Nikolopoulos et al., 2015; Peruccacci et al., 2012). Other authors have connected increased landslide activity with seasonal changes in precipitation and hillslope hydrologic conditions, especially during the wetter winter season in the PNW (Godt et al., 2006; Mirus, Morphew, et al., 2018). Studies reporting a seasonal distribution of landslides often rely on two metrics: the number (or frequency) of landslides and the presence or absence of landslides (Saito et al., 2010; Schneuwly-Bollschweiler & Stoffel, 2012; Sepúlveda & Petley, 2015).

Here, our objectives are to (1) test with statistical models whether these two descriptive metrics objectively reveal seasonal variations in landslide activity in the Pacific Northwest, and (2) test if landslide response to precipitation changes seasonally. We use Bayesian inference to estimate monthly landslide intensity and probability from five different landslide inventories from the region (Figure 4.1, Table S4.1), employing one set of models that learn from landslide inventories alone and another that is conditioned on monthly precipitation over the inventory areas. Using Bayesian models in landslide research (Berti et al., 2012; Korup, 2021; Lombardo et al., 2020; Nolde & Joe, 2013) has the advantage of providing intrinsic and objective estimates of parameter uncertainty, a metric too often neglected in landslide studies (Segoni et al., 2018).

Yet, a challenge in modeling landslide activity is that different inventories report landslides in various ways, and with differing spatial and temporal coverage (Steger et al., 2017). In the PNW, a number of inventories capture quite diverse aspects of landsliding (Figure 4.1, Table S4.1). The NASA Global Landslide Catalog (NASA GLC), for example, largely relied on news and highway department reports, documenting landslides along roads and in urban areas (Kirschbaum et al., 2015; NASA, 2018). The Washington Landslide Compilation (WLC), on the other hand, mapped widespread

shallow landsliding events from LiDAR, aerial photographs, and field visits (Washington Geological Survey, 2020). Bayesian multi-level regression makes use of the combined diverse, if not seemingly incompatible, information on landslide timing contained in these heterogeneous inventories, while still providing estimates for each inventory individually. Our results provide quantitative expectations for monthly landslide probability, intensity, and inter-annual variability in the PNW, inclusive of all uncertainty learned from the data.

## 4.2 STUDY AREA AND DATA

The PNW has one of the highest concentrations of mapped landslides in the United States (Mirus et al., 2020). We focused on the states of Washington and Oregon (Figure 4.1d), which are topographically characterized by two mountain ranges, a lower coastal range and the higher Cascades Range, to the east of which relief is generally lower. Precipitation in the region has a strong west to east gradient controlled to first order by orographic effects. Mean annual precipitation to the west of the Cascades ranges from ~1000 to >4000 mm and drops to <500 mm to the east (PRISM Climate Group, 2021).

We analyzed five landslide inventories that cover parts or all of Washington and Oregon and include information on landslide timing (Figure 4.1, Table S4.1). Mapped landslide concentration is highest in and to the west of the Cascades, and the length of record varies from 6 to 122 years between inventories (Figure 4.1). We considered only those landslides with at least a known month of occurrence, which represent <1 – 100% of each inventory (Table S4.1). Where recorded, most landslides were categorized as rainfall-triggered shallow landslides or debris flows. Depending on the inventory, landslides were mapped through field, LiDAR, or imagery analysis, or collected from news articles, highway reports, consulting reports or other records (Table S4.1).

We subset the NASA GLC to Washington and Oregon (NASA, 2018). We treated the Statewide Landslide Information Database for Oregon (SLIDO) historical points dataset (SLIDOh) and the landslide polygons dataset (SLIDOd) separately, as they resulted from different reporting protocols (Franczyk et al., 2020). We also included all landslides with monthly time-stamps from the Seattle Historic Landslide Locations (Seattle) (City of Seattle, 2020) and WLC (Washington Geological Survey, 2020) inventories, with the exception of the 1980 eruption of Mt. St. Helens.

We used two 4-km resolution precipitation datasets over the inventory areas: the PRISM 30-Year Normals (Norm91m) describes average monthly precipitation (Ppt) between 1991 and 2020, while PRISM AN81m provides monthly estimates of precipitation (PRISM Climate Group, 2021). For consistency with the normals data, we analyzed all years between 1991 and 2020 in the AN81m data, computing the spatial mean of both datasets across each inventory footprint area (Figure 4.1).

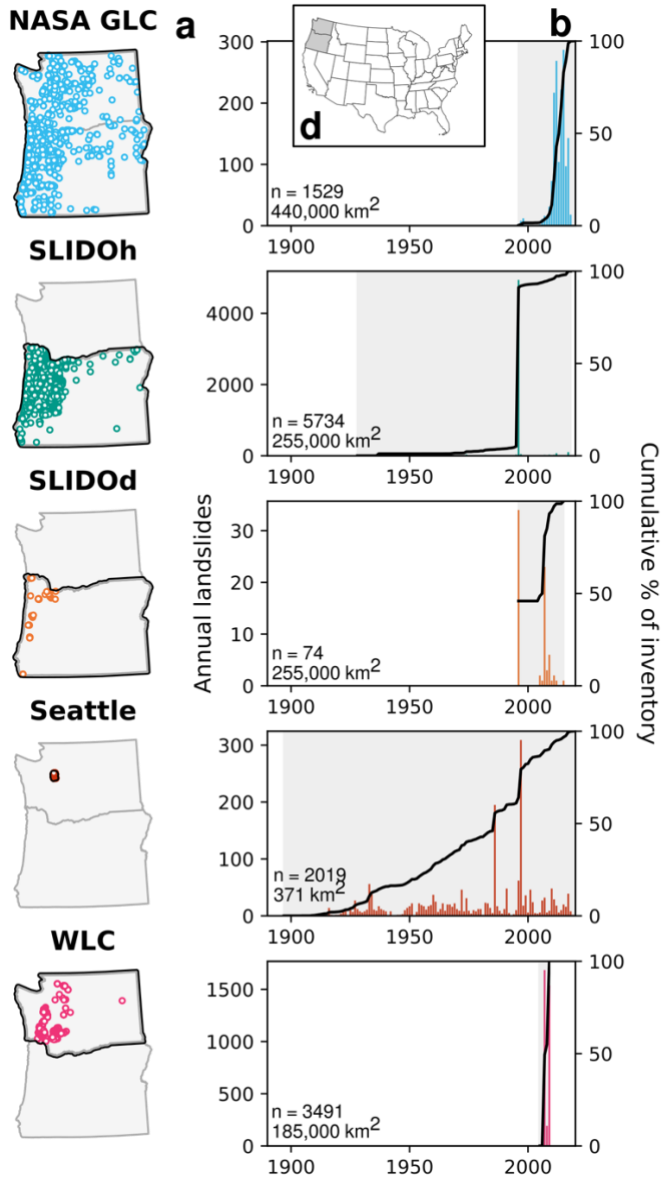



**Figure 4.1: Reported landslides in the Pacific Northwest from five inventories**. (a) Locations of landslides with monthly time-stamps (open circles) and footprint area covered by each inventory (black outline) in Washington and Oregon (b). (c) Annual time series of landslides (colored bars) over the period of record (gray shading). Black line indicates cumulative proportion of landslides recorded over time. (d) Box plots show the distribution of landslides in each month (box encloses the interquartile range, horizontal line is the median, whiskers cover 1.5 times the interquartile range); "+1" allows for display of months with no landslides on a log-scale.

## 4.3 METHODS

We trained two variants of Bayesian multi-level regression models to learn the seasonal pattern of landslide activity at monthly resolution. We used negative binomial regression to estimate the number of landslides reported in a given month (intensity) and logistic regression to estimate the presence or absence of reported landslides (probability). For each regression, we first fit models to inventory data alone to obtain the seasonal landslide pattern without any other predictors (landslide-only models). We then included spatially averaged monthly precipitation per inventory area (1991-2020) as a predictor in these generalized linear models to test for a seasonal landslide response to precipitation (landslide-precipitation models).

We chose Bayesian multi-level models because they are able to share information between landslide inventories, taking advantage of the diverse information contained in the different inventories (multi-inventory models). At the same time, the models provide parameter estimates for each inventory individually, thus respecting some of the important differences between inventories, like the area covered. In these models, the landslide data are grouped by month of occurrence and by inventory, hence one set of parameters is learned for each month and inventory. These parameter estimates are informed by a higher-level distribution of (hyper-)parameters that is also learned from the data and acts as an adaptive prior for each individual month and inventory. In this way, information is shared across inventories, which has a regularizing effect and generally improves estimates for groups with few observations while preventing overfitting to groups with many observations (Gelman & Hill, 2007; McElreath, 2020). For comparison, we also fit the landslide-only models to each inventory separately (single-inventory models). By binning our data into monthly intervals, we acknowledge the lack of any finer, for example daily, resolution that would require regression models with autocorrelation terms.

We fit the landslide-only negative binomial models to the number of landslides ($y_i$) that occurred in each month $m$ recorded in each inventory $v$. For each month and inventory, the model learned a mean ($\mu_{m,v}$) intensity or expected number of landslides per area, and a shape parameter ($\phi_{m,v}$). The negative binomial distribution's variance is a function of $\mu$ and $\phi$ and represents the inter-annual variability in the number of monthly landslides. In the five single-inventory models, the data were grouped by month and in the multi-inventory model, the data were grouped by month and inventory (Equation 1).

$$
\begin{gathered}
y_{i,m,v} \sim NegativeBinomial(\mu_{m,v}, \phi_{m,v}) \\
ln(\mu_{m,v}) = (\beta_0 + \beta_{0,m} + \beta_{0,v}) + ln(A_v) \\
ln(\phi_{m,v}) = \gamma_0 + \gamma_{0,m} + \gamma_{0,v} \\
\beta_{0,m} \sim Normal(0, \sigma_{0,m}) \\
\beta_{0,v} \sim Normal(0, \sigma_{0,v}) \\
\gamma_{0,m} \sim Normal(0, \psi_m) \\
\gamma_{0,v} \sim Normal(0, \psi_v)
\end{gathered}
$$

*Equation 4.1*

where $\beta_0$ and $\gamma_0$ are population-level intercepts, $\beta_{0,m}$ and $\gamma_{0,m}$ are group-level intercepts for month, and $\beta_{0,v}$ and $\gamma_{0,v}$ are group-level intercepts for inventory (which are excluded in the single-inventory models). $A_v$ is the footprint area of the inventory. The distribution of the group-level intercepts, which serve as adaptive priors for each individual group-level intercept, are specified as normal distributions with zero means and standard deviations $\sigma_{0,m}$ and $\psi_m$, and $\sigma_{0,v}$ and $\psi_v$. Prior distributions must be defined for $\beta_0$, $\gamma_0$, $\sigma_{0,m}$, $\psi_m$, $\sigma_{0,v}$, and $\psi_v$. For all models, we used weakly informative Student's $t$ priors that are implemented as widely applicable defaults in the R package brms, which we used to fit the models (Bürkner, 2017).

The landslide-precipitation negative binomial models follow the same structure, but include precipitation as a predictor.

$$
\begin{gathered}
y_{i,m,v} \sim NegativeBinomial(\mu_{i,m,v}, \phi_{m,v}) \\
ln(\mu_{m,v}) = (\beta_0 + \beta_{0,m} + \beta_{0,v}) + (\beta_1 + \beta_{1,m} + \beta_{1,v})P_{i,m,v} + ln(A_v) \\
ln(\phi_{m,v}) = \gamma_0 + \gamma_{0,m} + \gamma_{0,v} \\
\begin{bmatrix} \beta_{0,m} \\ \beta_{1,m} \end{bmatrix} \sim MVNormal\left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{0,m}^2 & \sigma_{0,m}\sigma_{1,m}\rho_m \\ \sigma_{0,m}\sigma_{1,m}\rho_m & \sigma_{1,m}^2 \end{bmatrix} \right) \\
\begin{bmatrix} \beta_{0,v} \\ \beta_{1,v} \end{bmatrix} \sim MVNormal\left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{0,v}^2 & \sigma_{0,v}\sigma_{1,v}\rho_v \\ \sigma_{0,v}\sigma_{1,v}\rho_v & \sigma_{1,v}^2 \end{bmatrix} \right) \\
\gamma_{0,m} \sim Normal(0, \psi_m) \\
\gamma_{0,v} \sim Normal(0, \psi_v)
\end{gathered}
$$

*Equation 4.2*

where $\beta_1$ is the population-level coefficient of precipitation, $\beta_{1,m}$ and $\beta_{1,v}$ are group-level coefficients with corresponding group-level standard deviations $\sigma_{1,m}$ and $\sigma_{1,v}$, and $P_{i,m,v}$ is the spatially averaged precipitation over the inventory area for a specific month between 1991 and 2020, standardized to the mean precipitation across all inventories and all months. The correlations $\varrho_m$ and $\varrho_v$ between group-level intercepts and coefficients of precipitation are also learned. We chose a uniform prior for $\beta_1$ and a non-informative Cholesky LKJ prior for the correlation matrix (Lewandowski et al., 2009; Stan Development Team, 2022).

Besides estimating the number of landslides, we also used logistic regression to model the presence or absence of reported landslides. Logistic regression relies on a Bernoulli

likelihood with parameter $p_{m,v}$ expressing the probability of having at least one landslide reported in a given month. We fit the logistic regression to data that indicate whether one or more landslides were recorded in an inventory in a given month ($z_i$). Again, we fit a model to data from each inventory and a model to data from all inventories for the landslide-only models (Equation 3).

$$
\begin{aligned}
z_{i,m,v} &\sim Bernoulli(p_{m,v}) \\
logit(p_{m,v}) &= \alpha_0 + \alpha_{0,m} + \alpha_{0,v} \\
\alpha_{0,m} &\sim Normal(0, \tau_{0,m}) \\
\alpha_{0,v} &\sim Normal(0, \tau_{0,v})
\end{aligned}
$$

*Equation 4.3*

where $p_{m,v}$ is a function of a population-level intercept $\alpha_0$ and group-level intercepts $\alpha_{0,m}$ for months and $\alpha_{0,v}$ for inventories (excluded from the single-inventory models), and $\tau_{0,m}$ and $\tau_{0,v}$ are the standard deviations of the group-level intercepts.

The landslide-precipitation logistic regression models again include standardized precipitation as a predictor.

$$
\begin{aligned}
z_{i,m,v} &\sim Bernoulli(p_{i,m,v}) \\
logit(p_{i,m,v}) &= (\alpha_0 + \alpha_{0,m} + \alpha_{0,v}) + (\alpha_1 + \alpha_{1,m} + \alpha_{1,v}) P_{i,m,v} \\
\begin{bmatrix} \alpha_{0,m} \\ \alpha_{1,m} \end{bmatrix} &\sim MVNormal\left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{0,m}^2 & \tau_{0,m}\tau_{1,m}\upsilon_m \\ \tau_{0,m}\tau_{1,m}\upsilon_m & \tau_{1,m}^2 \end{bmatrix} \right) \\
\begin{bmatrix} \alpha_{0,v} \\ \alpha_{1,v} \end{bmatrix} &\sim MVNormal\left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{0,v}^2 & \tau_{0,v}\tau_{1,v}\upsilon_v \\ \tau_{0,v}\tau_{1,v}\upsilon_v & \tau_{1,v}^2 \end{bmatrix} \right)
\end{aligned}
$$

*Equation 4.4*

where $\alpha_1$ is the population-level coefficient of precipitation, and $\alpha_{1,m}$ and $\alpha_{1,v}$ are group-level coefficients for months and inventories with corresponding standard deviations $\tau_{1,m}$ and $\tau_{1,v}$. $\upsilon_m$ and $\upsilon_v$ are the correlations between group-level intercepts and coefficients of precipitation.

We fit all models with the R package brms version 2.17.0 (Bürkner, 2017). Brms calls Stan, a probabilistic programming language that uses Hamiltonian Monte Carlo to approximate the posterior parameter space (Stan Development Team, 2022). We ran the sampler for 2000 iterations for four independent chains, discarding the first 500 draws as warm up, and checked that the chains converged. The single-inventory negative binomial model for WLC did not converge, likely because nearly all landslides recorded in this inventory happened in two months; we excluded this model from further analysis.

The model results are posterior parameter estimates conditional on the landslide inventory data, and, for the landslide-precipitation models, monthly precipitation (Figure 4.2, Figure 4.3, Figures S4.1-4.4). We defined parameter estimates to be credibly

distinguishable, for example between subsequent months, if their 95% highest-density-intervals (HDIs) did not overlap. Finally, we integrated over the posterior parameter estimates to characterize the distribution of landslide activity, thus naturally propagating all uncertainties in our summarized simulated landslide counts and presences/absences (Figure 4.2). These are called posterior predictive distributions in Bayesian statistics. We also report posterior expectations across all precipitation values for November and February for the landslide-precipitation models (Figure 4.3).

## 4.4 RESULTS

Both our landslide-only multi-inventory models confirm a distinct seasonal pattern of landslide activity in the PNW (Figure 4.2). Expected monthly landslide intensity is highest in January and February, decreases from March to May, and is lowest in June (dark green points in Figure 4.2A). Posterior estimates remain low from July to October, before increasing in November and December. This trend is consistent across all inventories, despite differences in reporting protocols and sample size. We can credibly distinguish (95% HDI) mean landslide intensity between subsequent months from February to April and from October to November in the NASA, Seattle, and SLIDOh inventories (Figure 4.2a). Landslide intensities per unit area are highest in the Seattle inventory and lowest in SLIDOd, with January landslide intensities in Seattle exceeding those in SLIDOd by more than three orders of magnitude. Monthly variance is also seasonal: highest from November to February and lowest from May through October (Figure 4.2b). The amplitude of the seasonal intensity pattern in a given inventory is smaller than between inventories (Figure 4.2a, Figure S4.1). Simulated posterior distributions of landslide counts are consistent with the data. Depending on the inventory, the model estimates up to tens to thousands of landslides in winter with 1% posterior probability, whereas median estimates range from zero to two landslides (Figure 4.2c).

Considering that median estimated counts are zero for all inventories between April and November, we also modeled the probability of any reported landslide in a given month. Our landslide-only multi-inventory logistic models show that landslides are most likely in January and become less so through June. The probability of landsliding is low, but non-zero, from June to October and increases again in November and December (Figure 4.2c). We observed credible differences between October and November and March and April in the Seattle and SLIDOh inventories. Landslide probabilities and their seasonal amplitude throughout the year are highest in the NASA inventory and lowest in WLC; again, we found that estimated variability between inventories is higher than between months (Figure S4.1).

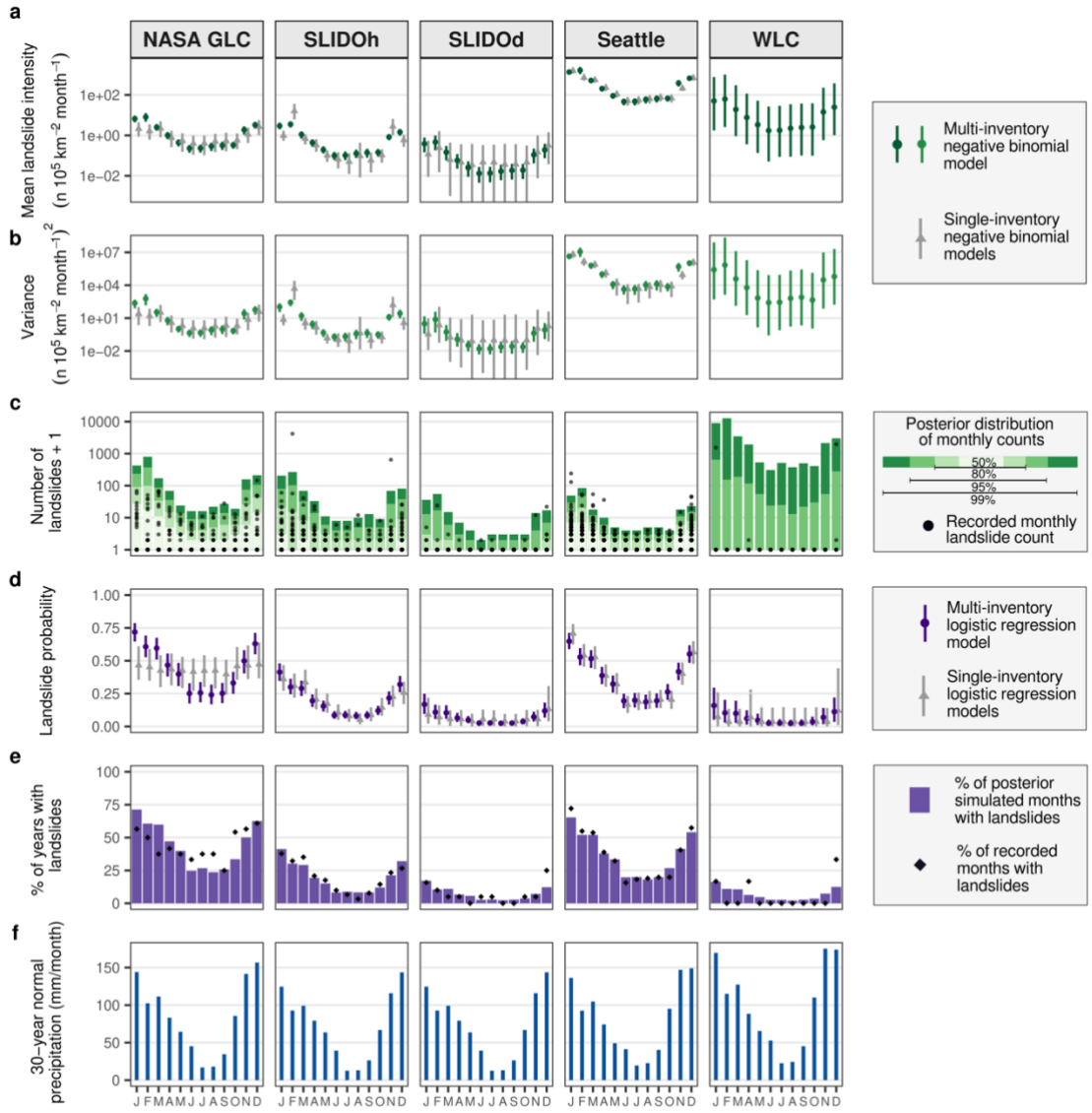



**Figure 4.2. Seasonal pattern of landslide activity from inventory-only models compared to 30-year normal monthly precipitation**. (a) Posterior parameter estimates for mean monthly landslide intensity, (b) variance, and (d) probability (median and 95% HDI). Two months are considered credibly different when their 95% HDIs do not overlap. (c) Distribution of posterior simulated counts and (e) months with landslides from the multi-inventory models. (f) Mean monthly precipitation over the inventory areas from 1991–2020 PRISM climate normals (PRISM Climate Group, 2021).

Both landslide-only multi-inventory models reveal a seasonality in landslide activity that some single-inventory models are unable to show (NASA, SLIDOd, WLC; gray symbols in Figure 4.2). The negative binomial multi-inventory model smooths the disjointed November and February landslide intensity peaks in the separate SLIDOh model, for example, and raises winter intensities for the NASA inventory, in each case informed by seasonal trends from the other inventories. Monthly uncertainties in the

multi-inventory models are also well below those in single-inventory models, especially for catalogs containing mostly months without reported landslides (SLIDOd).

Comparing these results to 30-year normal monthly precipitation (PRISM Climate Group, 2021) over the areas covered by these inventories shows an offset between the annual precipitation pattern and landslide activity (Figure 4.2f). Mean monthly precipitation increased markedly from September to October, but landslide activity did not credibly increase until November. Similarly, precipitation in these areas peaked in November or December, whereas our results show that landslide probability peaked in January and intensity in February, lagging the annual peak in precipitation by one to two months.

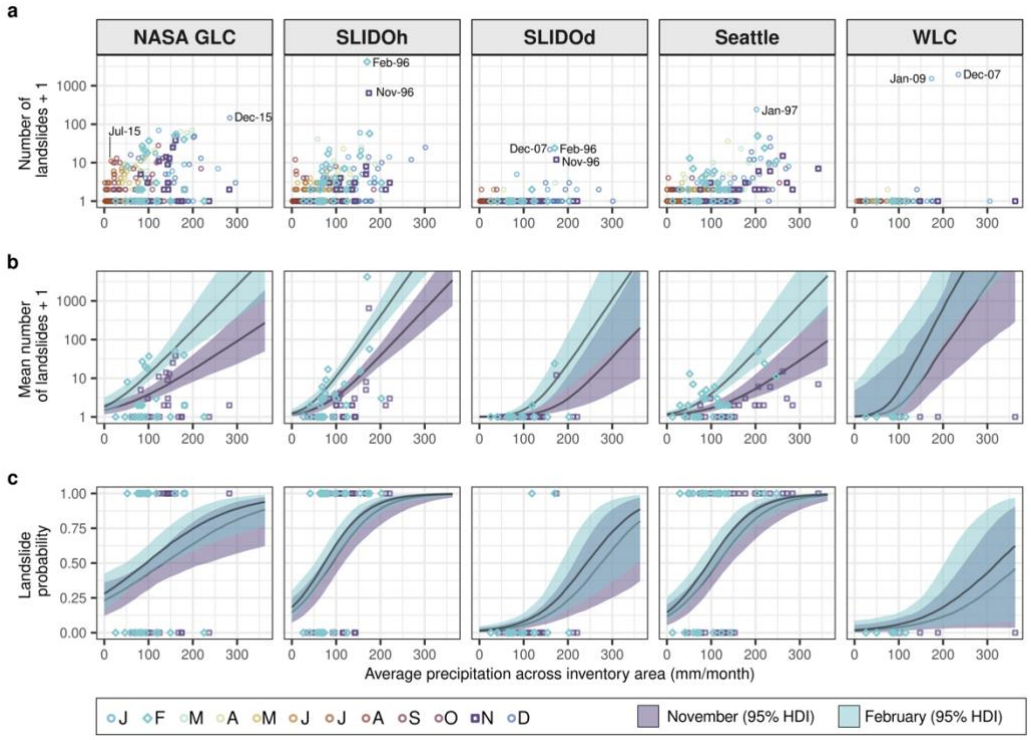



**Figure 4.3. Monthly landslide response to precipitation from Bayesian regression**. (a) Reported landslide count by month and inventory. Labeled points represent the highest counts reported in each inventory or months discussed in the text (month-year format). Posterior estimates of mean landslide intensity (b), and probability (c) with average precipitation across inventory area for November and February.

Considering average precipitation as an additional predictor of both mean landslide intensity and probability shows a positive relationship between precipitation and both intensity and probability, regardless of month (Figure 4.3, Figure S4.3, Figure S4.4). However, across all precipitation values, landslide intensity is substantially elevated in February compared to November in all inventories (Figure 3); these differences are credible at the mean precipitation value for NASA GLC, SLIDOh, and Seattle (Figure S4.4). For example, in the NASA GLC, our model estimates an order of magnitude more landslides for 200-mm average precipitation in February than for the same amount in November (Figure 4.3b). Landslide probability, on the other hand, remains

indistinguishable between months for a given precipitation average (Figure 4.3, Figure S4.4). In all inventory areas, an average of 200 mm of precipitation would produce a similar estimated probability in both November and February; at monthly rainfall means above 200 mm, our models estimate that reported landslides are >95% probable in the SLIDOh and Seattle inventories.

## 4.5 DISCUSSION AND CONCLUSIONS

We investigated patterns of monthly landslide activity in the Pacific Northwest with Bayesian multi-level models that unite data from five heterogeneous landslide inventories. Our multi-inventory models combine data from inventories with differing spatial and temporal coverage, data density, and reporting protocols to learn a regional seasonal pattern that some inventories show less distinctly and with higher uncertainties. Multi-inventory model results for monthly landslide intensity and probability both show a distinct seasonal pattern, with landslide activity peaking in winter, declining to a summer low, and increasing again in the fall. Credible increases in average monthly landslide intensity, inter-annual variability, and probability between October and November objectively mark the onset of the landslide season. Landslide intensity and probability increase with precipitation in all months, but landslide intensity is much higher for a given precipitation average at the peak of the landslide season in February compared to the onset in November (Figure 4.3).

This landslide seasonality in the PNW is largely linked to precipitation (Godt et al., 2006; Mirus, Morphew, et al., 2018), and most documented triggers or types in the inventories refer to rainfall-triggered shallow translational slides or flows. Heavy precipitation in the PNW mostly occurs between late October and mid-March, often linked to atmospheric rivers (Neiman et al., 2008, 2011) (Figure 4.2). Flooding in the PNW is also highly seasonal, with annual discharge peaks from November to January (Dougherty & Rasmussen, 2019; Neiman et al., 2011; Villarini, 2016). Our results expand previous conceptual models by showing that landslide seasonality lags the annual precipitation pattern by one to two months. The November landslide season onset follows the October precipitation increase; peak landslide probability in January and intensity in February come after November and December peaks in mean monthly precipitation. This lag indicates that landslide activity is highest when hillslopes have become sufficiently saturated, consistent with studies that have shown antecedent rainfall and resulting excess groundwater to be key predictors of landslide hazard on shorter, i.e. hourly to daily timescales in the PNW (Godt et al., 2006; Mirus, Becker, et al., 2018; Scheevel et al., 2017).

Besides pointing to monthly lags, our models also show that a given amount of rain is expected to initiate far more landslides in February than in November (Figure 4.3), likely because slopes are primed for failure after having accumulated moisture over the winter (Godt et al., 2006). For example, in both February and November 1996, large storms triggered widespread shallow landslides and debris flows that are reported in the SLIDOh inventory (Burns et al., 1998; Harp et al., 1997) (Figure 4.3). Despite similar mean precipitation totals in those months, the February storm triggered over 4000 reported landslides by delivering heavy rain to already wet, and in some areas snow-covered, hillslopes, whereas the November storm triggered a still disastrous 645 reported landslides. The February 1996 storm resulted in estimated direct damages of $100 million USD in Oregon alone (Wang et al., 2002). Unlike landslide intensity, landslide probability conditioned on precipitation remains comparable between all

months, suggesting that while antecedent conditions are important for initiating many landslides, sufficient rain in any month is likely to initiate at least one landslide (Figure 4.3).

Our models also highlight that landslide counts are more variable in winter than in summer, when landslides are less probable. Nevertheless, summer landslide probability exceeds 25% in the NASA inventory. In July, 2015, for example, 10 landslides were reported along roads or rivers in the NASA GLC; those with reported triggers were initiated by downpours, showing the effects of summer convective storms on the seasonal pattern of landsliding.

Overall, our multi-level modeling framework is able to learn more about regional landslide seasonality than models trained on individual inventories (Figure 4.2), but retains the data structure of all inventories. By design, a single seasonal pattern is learned from all of the inventories, but the amplitudes of monthly intensities or probabilities within that pattern differ by inventory. We find that the effects of inventory on landslide intensity exceed the effects of seasonality (Figure 4.2, Figure S4.1); the resulting estimated number of landslides per area in Seattle is much higher than in the region covered by WLC, for example. Similarly, estimated landslide probability in January is 8% in the WLC inventory and 69% in the NASA inventory, despite the spatial overlap between these inventories. This apparent mismatch could arise from the physics of landslide occurrence or from the details of the recording protocol. Landslide activity in urban areas and along roads may be higher than in rural areas (Johnston et al., 2021), spatial variations in landslide susceptibility may be captured by different inventory footprints, or average landslide intensities may have differed in the time periods covered by the inventories (Lombardo et al., 2020). A more likely explanation is reporting bias arising from more detailed observations in urban areas and along roads (Steger et al., 2017). The monthly time-stamped landslides included in our study are a subset of all reported landslides: nearly all landslides reported in Seattle have a time-stamp, whereas only 8% of landslides reported in WLC have a time-stamp (Table S1). 99% of dated landslides in WLC occurred on three unique days during major storms, whereas landslides in the Seattle inventory represent 823 days over 125 years (Table S1). Even if landslide activity in Seattle and the area covered by WLC were identical in space and time, a higher probability of recording landslides in Seattle would lead to higher apparent intensities and probabilities.

Given the different recording protocols between inventories, estimates from multi-inventory models (Figure 4.3) have additional advantages (Figure S4.2). Shrinkage refers to the difference between the single-inventory model parameter estimates and the multi-inventory model estimates and demonstrates how the inventories learn from each other (McElreath, 2020). The NASA inventory records landslides mostly along highways and in urban areas, and less so regional episodes that are recorded in WLC or SLIDOh. Parameter shrinkage pulls January and February intensity estimates up for the NASA inventory, as it learns about regional landslide episodes from the other inventories (green and gray points in Figure 4.2A). Conversely, the single-inventory

estimates for SLIDOh are dominated by the regional storms in November and February 1996 (Figure 4.2, Figure S4.2). Shrinkage pulls those estimates towards the global mean of the data, thus preventing overfitting. The same effect is observed when comparing the share of simulated months with landslides to the share of recorded months with landslides (Figure 4.3B, Figure S4.2). Single-inventory models tend to overfit the data, while the multi-inventory model generalizes better.

Altogether, our results show that Bayesian multi-level models are a useful and underexplored way to combine and reconcile information from multiple and seemingly incompatible landslide inventories. Potential further applications for combining inventory data or learning the variation between other groups are wide ranging: examples include intensity-duration thresholds for different seasons (Nikolopoulos et al., 2015; Peruccacci et al., 2012) or regions (Guzzetti et al., 2008), or combining multiple inventories in regression-based susceptibility models (Reichenbach et al., 2018).

Better understanding landslide activity at monthly to seasonal time-scales has the potential to improve emergency preparedness. Our results show that PNW landslide activity peaks in January to February, lagging mean monthly precipitation, and that similar rainfall leads to substantially higher intensities at the landslide season peak than at the onset. Because our models flexibly learn landslide activity by month from inventory data, they could be used to investigate landslide seasonality in regions with patterns as diverse as the winter-summer bimodal pattern observed in the western Himalaya and Karakoram (Hunt & Dimri, 2021), the East Asian Summer Monsoon peak in Japan (Saito et al., 2010), or the Atlantic hurricane season fall peak in Central America and the Caribbean (Sepúlveda & Petley, 2015). An important next step in seasonal landslide research will be to predict monthly intensity or probability for specific years, potentially considering global-interannual climate variability (Emberson et al., 2021), and eventually leading to operational forecasts. Such efforts will also pave the way for studies of whether and how climate change alters landslide seasonality.

## 4.6 ACKNOWLEDGMENTS

This research was funded by the DFG RTG “Natural Hazards and Risks in a Changing World" (NatRiskChange GRK 2043). We thank Natalie Lützow for help with data collection, Jürgen Kurths and Norbert Marwan for feedback, and Ben Mirus and an anonymous reviewer for comments that helped to improve the manuscript.

## 4.7 OPEN RESEARCH

The landslide inventory data are available from:

- NASA Global Landslide Catalog via https://data.nasa.gov/Earth-Science/Global-Landslide-Catalog/h9d8-neg4 (accessed 2021-04-13). Open Database License. (NASA, 2018)

- Statewide Landslide Information Database for Oregon, release 4.2 (SLIDO-4.2) via https://www.oregongeology.org/pubs/dds/p-slido4.htm (accessed 2021-01-15). Public. (Franczyk et al., 2020)

- Seattle Historic Landslide Locations ECA. https://data-seattlecitygis.opendata.arcgis.com/datasets/6ac72973a5784d90bda0a5f8a001d9f3_22/explore?location=47.616250%2C-122.328600%2C11.91 (accessed 2021-04-13). PDDL License. (City of Seattle, 2020)

- Washington Landslide Compilation. https://gis-qa.dnr.wa.gov/portal/home/item.html?id=da4255f770f84144b01c54010d533f4d (accessed 2021-04-13). Public. (Washington Geological Survey, 2020)

PRISM 30-year climate normals 1991-2020 (Norm91m, v M3, Ppt, November 2021) are available at https://prism.oregonstate.edu/normals/ (accessed 2022-02-14). Monthly PRISM AN81m (v M3, Ppt, July 2015) precipitation data is available at https://ftp.prism.oregonstate.edu/monthly/ppt/ (accessed 2022-08-18). (PRISM Climate Group, 2021)

Models were fit using brms version 2.17.0, available at https://cran.r-project.org/web/packages/brms/index.html (GPL-2) (Bürkner, 2017)

The Python Jupyter and R Markdown Notebooks used for this analysis are available from Luna (2022), and can be found at https://github.com/lvluna/landslide-seasonality.

# 4.8 REFERENCES

Baum, R. L., & Godt, J. W. (2010). Early warning of rainfall-induced shallow landslides and debris flows in the USA. Landslides, 7(3), 259–272. https://doi.org/10.1007/s10346-009-0177-0

Berti, M., Martina, M. L. V., Franceschini, S., Pignone, S., Simoni, A., & Pizziolo, M. (2012). Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach. Journal of Geophysical Research: Earth Surface, 117(F4). https://doi.org/10.1029/2012JF002367

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software, 80, 1–28. https://doi.org/10.18637/jss.v080.i01

Burns, S. F., Burns, W. J., James, D. H., & Hinkle, J. C. (1998). Landslides in the Portland, Oregon Metropolitan Area Resulting from the Storm of February 1996: Inventory Map, Database and Evaluation (p. 43). Portland State University. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.694.3602&rep=rep1&type=pdf

City of Seattle. (2020). Historical Landslide Locations ECA. City of Seattle. Retrieved from https://data-seattlecitygis.opendata.arcgis.com/datasets/6ac72973a5784d90bda0a5f8a001d9f3_22

Dougherty, E., & Rasmussen, K. L. (2019). Climatology of Flood-Producing Storms and Their Associated Rainfall Characteristics in the United States. Monthly Weather Review, 147(11), 3861–3877. https://doi.org/10.1175/MWR-D-19-0020.1

Emberson, R., Kirschbaum, D., & Stanley, T. (2021). Global connections between El Nino and landslide impacts. Nature Communications, 12(1), 2262. https://doi.org/10.1038/s41467-021-22398-4

Franczyk, J. J., Burns, W. J., & Calhoun, N. C. (2020). Statewide Landslide Information Database for Oregon, release 4 (SLIDO-4.2). Oregon Department of Geology and Mineral Industries. Retrieved from https://www.oregongeology.org/pubs/dds/p-slido4.htm

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. Natural Hazards and Earth System Sciences, 18(8), 2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

Gelman, A., & Hill, J. (2007). Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press.

Godt, J. W., Baum, R. L., & Chleborad, A. F. (2006). Rainfall characteristics for shallow landsliding in Seattle, Washington, USA. Earth Surface Processes and Landforms, 31(1), 97–110. https://doi.org/10.1002/esp.1237

Guzzetti, F., Peruccacci, S., Rossi, M., & Stark, C. P. (2008). The rainfall intensity–duration control of shallow landslides and debris flows: an update. Landslides, 5(1), 3–17. https://doi.org/10.1007/s10346-007-0112-1

Guzzetti, F., Gariano, S. L., Peruccacci, S., Brunetti, M. T., Marchesini, I., Rossi, M., & Melillo, M. (2020). Geographical landslide early warning systems. Earth-Science Reviews, 200, 102973. https://doi.org/10.1016/j.earscirev.2019.102973

Harp, E. L., Chleborad, A. F., Schuster, R. L., Cannon, S. H., Reid, M. E., & Wilson, R. C. (1997). Landslides and Landslide Hazards in Washington State Due to February 5-9, 1996 Storm (p. 33). U.S. Dept. of the Interior, U.S. Geological Survey. Retrieved from https://www.unisdr.org/preventionweb/files/1585_Washhrp.pdf

Hunt, K. M. R., & Dimri, A. P. (2021). Synoptic-scale precursors of landslides in the western Himalaya and Karakoram. Science of The Total Environment, 776, 145895. https://doi.org/10.1016/j.scitotenv.2021.145895

Johnston, E. C., Davenport, F. V., Wang, L., Caers, J. K., Muthukrishnan, S., Burke, M., & Diffenbaugh, N. S. (2021). Quantifying the Effect of Precipitation on Landslide Hazard in Urbanized and Non-Urbanized Areas. Geophysical Research Letters, 48(16), e2021GL094038. https://doi.org/10.1029/2021GL094038

Kirschbaum, D., Stanley, T., & Zhou, Y. (2015). Spatial and temporal analysis of a global landslide catalog. Geomorphology, 249, 4–15. https://doi.org/10.1016/j.geomorph.2015.03.016

Korup, O. (2021). Bayesian geomorphology. Earth Surface Processes and Landforms, 46(1), 151–172. https://doi.org/10.1002/esp.4995

Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. Journal of Multivariate Analysis, 100(9), 1989–2001. https://doi.org/10.1016/j.jmva.2009.04.008

Lombardo, L., Opitz, T., Ardizzone, F., Guzzetti, F., & Huser, R. (2020). Space-time landslide predictive modelling. Earth-Science Reviews, 209, 103318. https://doi.org/10.1016/j.earscirev.2020.103318

Luna, L. V. (2022). landslide-seasonality (Version 0.0.0). https://doi.org/10.5281/zenodo.6299024

Nolde, N., & Joe, H. (2013). A Bayesian extreme value analysis of debris flows. Water Resources Research, 49(10), 7009–7022. https://doi.org/10.1002/wrcr.20494

McElreath, R. (2020). Statistical Rethinking A Bayesian Course with Examples in R and STAN (2nd Edition). Chapman and Hall/CRC Press. Retrieved from https://www.routledge.com/Statistical-Rethinking-A-Bayesian-Course-with-Examples-in-R-and-STAN/McElreath/p/book/9780367139919

Mirus, B. B., Morphew, M. D., & Smith, J. B. (2018). Developing Hydro-Meteorological Thresholds for Shallow Landslide Initiation and Early Warning. Water, 10(9), 1274. https://doi.org/10.3390/w10091274

Mirus, B. B., Becker, R. E., Baum, R. L., & Smith, J. B. (2018). Integrating real-time subsurface hydrologic monitoring with empirical rainfall thresholds to improve landslide early warning. Landslides, 15(10), 1909–1919. https://doi.org/10.1007/s10346-018-0995-z

Mirus, B. B., Jones, E. S., Baum, R. L., Godt, J. W., Slaughter, S., Crawford, M. M., et al. (2020). Landslides across the USA: occurrence, susceptibility, and data limitations. Landslides, 17(10), 2271–2285. https://doi.org/10.1007/s10346-020-01424-4

NASA. (2018). Global Landslide Catalog. NASA. Retrieved from https://data.nasa.gov/Earth-Science/Global-Landslide-Catalog/h9d8-neg4

Napolitano, E., Fusco, F., Baum, R. L., Godt, J. W., & De Vita, P. (2016). Effect of antecedent-hydrological conditions on rainfall triggering of debris flows in ash-fall pyroclastic mantled slopes of Campania (southern Italy). Landslides, 13(5), 967–983. https://doi.org/10.1007/s10346-015-0647-5

Neiman, P. J., Ralph, F. M., Wick, G. A., Lundquist, J. D., & Dettinger, M. D. (2008). Meteorological Characteristics and Overland Precipitation Impacts of Atmospheric Rivers Affecting the West Coast of North America Based on Eight Years of SSM/I Satellite Observations. Journal of Hydrometeorology, 9(1), 22–47. https://doi.org/10.1175/2007JHM855.1

Neiman, P. J., Schick, L. J., Ralph, F. M., Hughes, M., & Wick, G. A. (2011). Flooding in Western Washington: The Connection to Atmospheric Rivers. Journal of Hydrometeorology, 12(6), 1337–1358. https://doi.org/10.1175/2011JHM1358.1

Nikolopoulos, E. I., Borga, M., Marra, F., Crema, S., & Marchi, L. (2015). Debris flows in the eastern Italian Alps: seasonality and atmospheric circulation patterns. Natural Hazards and Earth System Sciences, 15(3), 647–656. https://doi.org/10.5194/nhess-15-647-2015

Peruccacci, S., Brunetti, M. T., Luciani, S., Vennari, C., & Guzzetti, F. (2012). Lithological and seasonal control on rainfall thresholds for the possible initiation of landslides in central Italy. Geomorphology, 139–140, 79–90. https://doi.org/10.1016/j.geomorph.2011.10.005

Pollock, W., & Wartman, J. (2020). Human Vulnerability to Landslides. GeoHealth, 4(10), e2020GH000287. https://doi.org/10.1029/2020GH000287

Pratt, C., Macciotta, R., & Hendry, M. (2019). Quantitative relationship between weather seasonality and rock fall occurrences north of Hope, BC, Canada. Bulletin of Engineering Geology and the Environment, 78(5), 3239–3251. https://doi.org/10.1007/s10064-018-1358-7

PRISM Climate Group. (2021). PRISM Spatial Climate Datasets for the Coterminous United States. Oregon State University. Retrieved from https://prism.oregonstate.edu/

Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., & Guzzetti, F. (2018). A review of statistically-based landslide susceptibility models. Earth-Science Reviews, 180, 60–91. https://doi.org/10.1016/j.earscirev.2018.03.001

Saito, H., Nakayama, D., & Matsuyama, H. (2010). Relationship between the initiation of a shallow landslide and rainfall intensity—duration thresholds in Japan. Geomorphology, 118(1), 167–175. https://doi.org/10.1016/j.geomorph.2009.12.016

Scheevel, C. R., Baum, R. L., Mirus, B. B., & Smith, J. B. (2017). Precipitation thresholds for landslide occurrence near Seattle, Mukilteo, and Everett, Washington (USGS Numbered Series No. 2017–1039). Precipitation thresholds for landslide occurrence near Seattle, Mukilteo, and Everett, Washington (Vol. 2017–1039, p. 60). Reston, VA: U.S. Geological Survey. https://doi.org/10.3133/ofr20171039

Schneuwly-Bollschweiler, M., & Stoffel, M. (2012). Hydrometeorological triggers of periglacial debris flows in the Zermatt valley (Switzerland) since 1864. Journal of Geophysical Research: Earth Surface, 117(F2). https://doi.org/10.1029/2011JF002262

Segoni, S., Piciullo, L., & Gariano, S. L. (2018). A review of the recent literature on rainfall thresholds for landslide occurrence. Landslides, 15(8), 1483–1501. https://doi.org/10.1007/s10346-018-0966-4

Sepúlveda, S. A., & Petley, D. N. (2015). Regional trends and controlling factors of fatal landslides in Latin America and the Caribbean. Natural Hazards and Earth System Sciences, 15(8), 1821–1833. https://doi.org/10.5194/nhess-15-1821-2015

Stan Development Team. (2022). Stan Modeling Language Users Guide and Reference Manual, version 2.28. Retrieved from https://mc-stan.org

Stanley, T. A., Kirschbaum, D. B., Sobieszczyk, S., Jasinski, M. F., Borak, J. S., & Slaughter, S. L. (2020). Building a landslide hazard indicator with machine learning and land surface models. Environmental Modelling & Software, 129, 104692. https://doi.org/10.1016/j.envsoft.2020.104692

Stanley, T. A., Kirschbaum, D. B., Benz, G., Emberson, R. A., Amatya, P. M., Medwedeff, W., & Clark, M. K. (2021). Data-Driven Landslide Nowcasting at the Global Scale. Frontiers in Earth Science, 9. Retrieved from https://www.frontiersin.org/articles/10.3389/feart.2021.640043

Steger, S., Brenning, A., Bell, R., & Glade, T. (2017). The influence of systematically incomplete shallow landslide inventories on statistical susceptibility models and suggestions for improvements. Landslides, 14(5), 1767–1781. https://doi.org/10.1007/s10346-017-0820-0

Villarini, G. (2016). On the seasonality of flooding across the continental United States. Advances in Water Resources, 87, 80–91. https://doi.org/10.1016/j.advwatres.2015.11.009

Wang, Y., Summers, R. D., & Hofmeister, R. J. (2002). Landslide Loss Estimation Pilot Project in Oregon (Open-File Report No. O-02-05). State of Oregon Department of Geology and Mineral Industries. Retrieved from https://oregonexplorer.info/data_files/OE_location/northcoast/documents/NorthCoastPDFs/O0205.pdf

Washington Geological Survey. (2020). Landslides compilation--GIS data, February 2020: Washington Geological Survey Digital Data Series 12, version 5.2. Retrieved from https://fortress.wa.gov/dnr/geologydata/publications/data_download/ger_portal_landslide_compilation.zip

# 5 Discussion and synthesis

## 5.1 Discussion

In this section, I return to each of my primary research questions, discussing my results in the context of my objectives, and exploring emerging avenues for future work.

### 5.1.1 Under which rainfall conditions should we warn for shallow landslides in Sitka, Alaska, USA?

In 2015, the town of Sitka, Alaska was impacted by a fatal landslide event. Following this event, the community and technical experts determined the need for a landslide early warning system that would give local officials and residents real-time information on current landslide hazard (Busch et al., 2021). In Chapter 2, I contributed to developing the statistical models and warning thresholds behind an online landslide early warning dashboard that provides updated estimates of landslide hazard in Sitka based on weather forecasts. Our objectives were to estimate daily landslide probability based on rainfall conditions and to establish decision thresholds for warning. A key challenge for developing such a system for Sitka was that there were only five reported landslide events near infrastructure with known dates between 2002 and 2020, the period for which station-based precipitation records were available. Similar systems that cover larger regions in other parts of the world are often based on hundreds or even thousands of landslide reports (e.g. Peruccacci et al., 2017; Scheevel et al., 2017).

We estimated daily landslide probability using logistic regression and found that three-hourly precipitation totals are the best predictor of elevated landslide hazard in Sitka. Including antecedent precipitation on timescales from days to weeks in the models did not measurably improve performance, which we attributed to rapid drainage from porous colluvial soils on steep hillslopes. Some regional warning thresholds rely only on reported landslide triggering rainfall, but do not consider rainfall that did not trigger

landslides (Peruccacci et al., 2017). By including such non-triggering precipitation in our models, we were able to constrain conditions under which landslides are very improbable, consistent with observations that including non-triggering precipitation can improve threshold robustness (Peres & Cancelliere, 2021). We observed that posterior uncertainty of daily landslide probability is considerably higher at higher precipitation values, which that occurred infrequently in the record. For example, an estimated landslide probability of 0.7 could be associated with three-hourly cumulative rainfall ranging from 31 to 39 mm. In future work, it may be possible to reduce this uncertainty by including a more informative prior based on expert knowledge or available information from other regions. For instance, Nolde & Joe, 2013 used a prior based on expert knowledge to reduce uncertainty of debris flow return level estimates in western Canada. Alternatively, a multi-level modeling approach, like the one I applied in Chapters 3 and 4, could be used to incorporate information from nearby regions and thus include more observations.

Considering these uncertainties, we established two decision thresholds that determine the three warning levels in the online dashboard: low, moderate, and high. The upper threshold was set such that no false alarms (warning, but no landslide) would have occurred in the past; forecast rainfall above this threshold indicates a high hazard situation that may warrant measures like evacuations or road closures. The lower threshold was set such that no missed alarms (landslide, but no warning) would have occurred in the past; when forecast rainfall is below this threshold, residents can feel assured that landslides are very improbable. Although it was necessary to choose exact precipitation values for these thresholds for implementation in the dashboard, we acknowledge that a range of thresholds could have produced the same validation results. For example, for the upper threshold, estimated probabilities ranging from 0.31 to 0.74 produced the same precision and recall, requiring a heuristic approach to setting the final threshold at an estimated probability of 0.7, and a three-hourly cumulative precipitation of 34 mm. The three level warning system reflects some of this uncertainty, as probabilities between 0.3 and 0.7 are covered by the moderate warning level. We set the lower boundary at an estimated probability of 0.01 (21 mm in three hours), which is a more conservative probability than the theoretically optimal one if we consider precision and recall. Therefore, we accepted some additional false alarms in exchange for a lower chance of missed alarms.

By splitting the time series into training and test periods, we showed that our models can predict landslide probability from measured precipitation. However, in the dashboard, we rely on weather forecasts to predict future landslide hazard. Currently, we assume that measured precipitation and weather forecasts are interchangeable, and do not account for uncertainty in the forecast precipitation. Future updates could assess whether including such uncertainty improves performance, although this may prove difficult to assess with few landslide observations. Issues could arise if the weather forecasts underestimate three-hourly rainfall accumulation, leading to missed alarms, or regularly overestimate precipitation, leading to many false alarms.

We estimated hazard for areas within two kilometers of the road network based on a single meteorological station located at the airport, and in doing so, assumed that this station is sufficiently representative of this area for warning purposes. The spatial variability of intense rainfall remains a persistent challenge for landslide early warning, particularly in areas where topography enhances this variability (Destro et al., 2017; Iadanza et al., 2016). In Sitka, orographic effects likely do lead to spatially variable rainfall. However, landslides in southeast Alaska are thought to be mostly associated with atmospheric rivers, which deliver moisture from the Pacific and cause 90% of extreme precipitation in the region (Sharma & Déry, 2020). In this case, the airport station may be sufficiently representative of the broad mesoscale atmospheric conditions for effective warning, whereas in regions where landslides are primarily triggered by spatially limited convective thunderstorms, this is a greater problem (Iadanza et al., 2016). However, this assumption could be tested using data from radar or from a newly installed network of tipping bucket rain gauges in Sitka, both of which could also be explored to develop neighborhood scale hazard nowcasts.

### 5.1.2 How do intensity-duration thresholds for urban landslides vary between cities worldwide?

Rainfall-triggered landslides can be particularly deadly in densely populated urban areas (Alcântara et al., 2022; Cui et al., 2019), but many cities with high landslide risk lack dedicated landslide early warning systems (Guzzetti et al., 2020). Thresholds have been observed to vary widely by climate zone, lithology, and topographic setting (Crosta, 1998; Segoni et al., 2014), but a systematic comparison of thresholds between cities using a consistent methodology has been missing. It has therefore been unclear to what extent a city with limited landslide inventory data, like Sitka, can rely on data from other cities or regions to inform warning efforts. In Chapter 3, I estimated rainfall intensity-duration (I-D) thresholds for 26 cities worldwide and a global threshold for urban landslides. My objectives were to update a prior global threshold with data from urban areas, to quantify the variability between thresholds in different cities using a consistent methodology, and to assess potential drivers of threshold variation between cities.

I applied Bayesian inference to update a prior global threshold for shallow landslides that relied on combined data from urban and rural areas (Guzzetti et al., 2008). I found that my estimated 10$^{th}$ percentile ($q_{10}$) global threshold for urban areas is lower than the prior threshold at durations up to three days, and that overall, 16% of urban landslides in my database would have occurred below the prior global threshold. This result indicates that landslides in densely populated urban areas have been triggered at lower rainfall intensities than previously reported globally, which may result from urbanization activities.

When assessing variability between cities, I intriguingly found that in most cities (77%), median ($q_{50}$) thresholds are not credibly distinguishable from my estimated global $q_{50}$

threshold for urban areas. This finding stands in stark contrast to previous observations that thresholds vary by lithology, topographic setting, mean annual precipitation, or climate zone (Brunetti et al., 2010; Crosta, 1998; Guzzetti et al., 2008; Segoni et al., 2014; Zêzere et al., 2015). I interpret this contrast to result from a harmonizing effect of urbanization activities and suggest that anthropogenic landscape modification has overprinted any pre-existing natural variation in rainfall thresholds between cities. The Bayesian implementation that I applied provides estimates of threshold parameter uncertainty, which has often gone undocumented in previous threshold studies. I found that, when considering this uncertainty, and using a consistent methodology, most cities are not credibly distinguishable from the global threshold. While dozens of studies have sought to quantify rainfall thresholds in regions worldwide, the multitude of methods in use have made comparison between different regions difficult (Segoni et al., 2018). A new question that arises from my findings is to what extent thresholds that have been previously observed to vary widely would be found to be indistinguishable if parameter uncertainty were considered using consistent methods and data.

An important implication of my results is that cities with limited landslide inventory data may be able to borrow thresholds and data from other cities for their own warning efforts, or to use the global thresholds that I estimated here. For some users, the uncertainty of these global thresholds may be deemed too high and would need to be reduced for implementation in an operational warning system, for example, by collecting additional local landslide observations. Additionally, as in Sitka, users may be interested in validating and tuning thresholds to the number of false alarms or missed alarms that would have occurred if these had been used for warning, depending on particular needs. Nevertheless, these thresholds provide a baseline.

The 26 cities with at least five reported landslides and available rainfall data for which I estimated thresholds broadly cover the climatic and topographic distribution of all 636 cities with reported landslides in my database. However, I note that that these cites over-represent large-population cities in high income countries. This reflects the greater availability of both landslide inventory and hourly rainfall data in the largest cities and in countries like the United States, Italy, Australia, and New Zealand. Future efforts to collect landslide inventory data for global studies should attempt to expand coverage, particularly in Africa, and consider the potential of satellite data to fill gaps in rainfall station observations (Brunetti et al., 2018; Marc et al., 2022; Thomas et al., 2019). Incorporating satellite rainfall estimates from the Global Precipitation Measurement (GPM) Mission would have increased the number of cities with five landslides with rainfall coverage in my study from 26 to 88 (Huffman et al., 2020). However, preliminary analysis revealed substantial differences between satellite estimated and station measured landslide triggering rainfall, so I deferred using satellite based data to fill these gaps until further validation has been completed.

The I-D threshold has remained a popular empirical tool for characterizing landslide-triggering rain (Segoni et al., 2018), but has limited ability to capture the more detailed hydro-meteorological processes that lead to landslides (Bogaard & Greco, 2018). Positive pore pressure, which sets most rainfall triggered landslides in motion, is a

function of antecedent soil moisture content and the rates at which water infiltrates and drains from a potentially unstable slope, a process that has been termed filling-storing-draining (Bogaard & Greco, 2016). At short durations, intensity-duration pairs capture triggering rain, but hardly account for antecedent conditions in the hillslope. At longer durations, however, intensity is averaged over days or weeks. In this case, triggering rain is blended with antecedent rain, while draining and evapotransporation are not accounted for. Including in-situ hillslope hydrological measurements may help to define thresholds that can more deftly separate these physical processes in areas where such data is available (Thomas et al., 2018; Wicki et al., 2021).

### 5.1.3 What is the relationship between reported landslide activity and seasonal rainfall patterns in the Pacific Northwest, USA?

As in Chapters 2 and 3, most landslide early warning research has focused on hourly to weekly timescales, while few studies have sought to predict seasonal variations in landslide activity (Steger et al., 2022). In the Pacific Northwest (PNW) of the USA, it is common knowledge that landslides occur during the wetter winter season (Godt et al., 2006; Mirus et al., 2018), but up until now, such a seasonal pattern of landslide activity had not been quantitatively investigated at regional scale using landslide inventory data. In Chapter 4, my objectives were to test if statistical models objectively reveal seasonal variations in landslide activity in the PNW and to test if landslide response to precipitation changes seasonally. I estimated monthly landslide probability and intensity (number of landslides per area) by using Bayesian multi-level models to unite data from five landslide inventories with differing spatial areas covered, time periods, and collection methods.

My statistical models did objectively reveal seasonal variations in landslide activity, characterizing the seasonal pattern of landsliding in the PNW in far more detail than had previously been published. I found that credible increases in monthly landslide intensity, inter-annual variability, and probability between October and November mark the onset of the landslide season. Landslide probability peaks in January, whereas landslide intensity peaks in February, both lagging the annual precipitation peak by one to two months. I found that the inter-annual variability in the number of landslides is much higher in winter months than in summer months, when landslides are uniformly rare.

For a given amount of monthly precipitation, I found that estimated landslide intensity is up to ten times higher at the peak of the landslide season in February than in November at its onset, demonstrating that landslide response to precipitation does change seasonally, and explaining the observed lag between peak landslide activity and the annual precipitation peak. I attribute this differing response to precipitation to the importance of antecedent hillslope conditions: as hillslopes become saturated over the course of the winter, a given amount of monthly precipitation leads to more landslides

compared to early in the season. Additionally, rain-on-snow events may contribute to this pattern: when rainfall quickly melts accumulated snowpack, the additional water infiltrating into hillslopes has led to widespread landsliding in several notable historical landslide events in the PNW. With a monthly rainfall forecast, these models could be used to forecast the probability and intensity of landsliding in the PNW.

In a recent study of seasonal landsliding in South Tyrol, Italy, Steger et al., 2022 found that more precipitation was required to trigger a landslide in summer than in winter or spring, which they attributed to vegetation and temperature effects. While we observe similar effects, namely that more precipitation is needed to trigger landslides in drier months than in wetter months, we interpret this signal in different ways. According to Steger et al., 2022's interpretation, denser vegetation growing on hillslopes in summer leads to higher interception of potentially triggering precipitation and increased temperature-driven evapotranspiration, which dries out slopes between triggering rain events. Future work could test each of these mechanisms in both areas. Together, our studies indicate that the same amount of rainfall leads to seasonally differing landslide response, with the implication that seasonally variable warning thresholds may be warranted in some areas (Nikolopoulos et al., 2015).

I found that by combining five different landslide inventories that each reported differing aspects of landslide activity in a multi-level model, the model could learn a regional pattern of landslide activity that some individual inventories failed to show, and with reduced parameter uncertainty. Incomplete inventory data is a much discussed challenge in landslide prediction research, both in space and time (Kirschbaum et al., 2015; Steger et al., 2017). However, while landslide inventories may suffer from some limitations, they also represent the best data source we have in order to develop predictive models, requiring approaches that are able to make best use of available information. My results showed that multi-level models are one way forward for handling this challenge in areas where multiple inventories exist.

However, moving forward, the landslide research community needs a toolkit of options for making best use of available, yet incomplete inventory data. To develop this, we may be able to learn from other fields that have confronted similar challenges. For example, landslide inventory data is analogous to "presence-only" data in ecology (Hefley et al., 2015; Warton & Shepherd, 2010). In these types of datasets, only presences are reported, whether that be a landslide, a type of tree, or a bird sighting. However, at places or during time periods when no observation is reported, we cannot be sure that no event occurred, only that no event was reported. Ecologists have addressed this issue by concurrently modeling the observation process with the question of interest. For example, Hefley et al., 2015 incorporated expert knowledge of the probability of spotting whooping cranes, an endangered bird species, to estimate the spatial distribution of crane habitat from opportunistic sightings. Such methods could also be explored in the future to overcome limitations related to incomplete landslide inventory data.

## 5.2 SYNTHESIS

The overarching motivation of this dissertation was to improve our ability to predict rainfall-triggered landslides in time with the goal of improving warning. As a whole, the work enclosed in the three chapters in my dissertation has taken important steps toward that goal, which I summarize as the following key achievements:

1. **My thesis expanded the spatial coverage of information for landslide early warning.** In Chapter 2, I contributed to developing the statistical analysis behind a real-time landslide early warning dashboard for a small town in Alaska that has previously experienced a fatal landslide but had no warning system. This information is available for emergency responders and the public at all times. In Chapter 3, I estimated intensity-duration thresholds for 26 cities worldwide, many of which do not have an operating landslide early warning system, and a global threshold for urban landslides. I proposed a modeling structure that can be used to estimate thresholds for cities with limited landslide inventory data, informed by data from other cities. I found that many urban landslides were triggered at moderate rainfall intensities that were lower than previously reported global thresholds. However, median thresholds for most cities are indistinguishable from my estimated global mean threshold for urban landslides, suggesting that anthropogenic landscape modification in densely populated urban areas has overprinted any pre-existing natural variability in rainfall thresholds between cities.

2. **My thesis applied modeling approaches that are able to make use of available landslide inventory and rainfall data to predict landslide activity and estimate uncertainty.** In Chapters 2 and 4, I applied Bayesian statistical models to estimate landslide probability and counts based on landslide inventory and precipitation data. These methods are established in other fields, but are novel in landslide research. In Chapter 2, I contributed to estimating landslide activity in Sitka using models trained with only five reported landslide events over 18 years. By including observations of rainfall that did not trigger landslides, we were able to constrain conditions under which landslides are improbable, whereas posterior distributions showed higher uncertainties at higher precipitation totals that have been observed less frequently. Based on these estimates, we established two warning thresholds that consider three-hourly cumulative precipitation. In Chapter 4, I used multi-level models that combined five heterogeneous landslide inventories from the Pacific Northwest, USA that covered different spatial areas, timescales, and were collected in different ways. This approach revealed a distinct seasonal pattern of landslide activity that some inventories were unable to show in isolation, and reduced uncertainties compared to estimates from individual inventories. This method offers one solution to the challenge of predicting landslide activity with incomplete inventory data. In Chapter 2, I presented intensity-duration thresholds with uncertainty estimates, which have been largely neglected in I-D studies. These studies,

and the open access code that accompanies them, can serve as examples for other studies wishing to apply similar methods.

3. **My thesis made inroads toward seasonal landslide forecasts, broadening the timescales considered for warning.** In Chapter 3, I predicted seasonal landslide activity in the Pacific Northwest based on monthly rainfall data, which added objective, quantitative information about seasonal patterns of landslide probability and intensity in the region. I moved beyond a common understanding that landslides occur during the wetter winter season by showing that peak landslide activity lags the annual precipitation peak by one to two months, and that, for a given monthly rainfall amount, landslide intensity is nearly ten times higher in February than in November, highlighting the importance of antecedent hillslope conditions.

# 5.3 REFERENCES

Alcântara, E., Marengo, J. A., Mantovani, J., Londe, L., San, R. L. Y., Park, E., et al. (2022). Deadly disasters in Southeastern South America: Flash floods and landslides of February 2022 in Petropolis, Rio de Janeiro. *Natural Hazards and Earth System Sciences Discussions*, 1–27. https://doi.org/10.5194/nhess-2022-163

Bogaard, T., & Greco, R. (2016). Landslide hydrology: from hydrology to pore pressure. *WIREs Water*, *3*(3), 439–459. https://doi.org/10.1002/wat2.1126

Bogaard, T., & Greco, R. (2018). Invited perspectives: Hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: proposing hydro-meteorological thresholds. *Natural Hazards and Earth System Sciences*, *18*(1), 31–39. https://doi.org/10.5194/nhess-18-31-2018

Brunetti, M. T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D., & Guzzetti, F. (2010). Rainfall thresholds for the possible occurrence of landslides in Italy. *Natural Hazards and Earth System Sciences*, *10*(3), 447–458. https://doi.org/10.5194/nhess-10-447-2010

Brunetti, M. T., Melillo, M., Peruccacci, S., Ciabatta, L., & Brocca, L. (2018). How far are we from the use of satellite rainfall products in landslide forecasting? *Remote Sensing of Environment*, *210*, 65–75. https://doi.org/10.1016/j.rse.2018.03.016

Busch, L., Lempert, R., Izenberg, M., & Patton, A. (2021). Run Uphill for a Tsunami, Downhill for a Landslide. *Issues in Science and Technology*, *38*(1), 40–46.

Crosta, G. (1998). Regionalization of rainfall thresholds: an aid to landslide hazard evaluation. *Environmental Geology*, *35*(2–3), 131–145. https://doi.org/10.1007/s002540050300

Cui, Y., Cheng, D., Choi, C. E., Jin, W., Lei, Y., & Kargel, J. S. (2019). The cost of rapid and haphazard urbanization: lessons learned from the Freetown landslide disaster. *Landslides*, *16*(6), 1167–1176. https://doi.org/10.1007/s10346-019-01167-x

Destro, E., Marra, F., Nikolopoulos, E. I., Zoccatelli, D., Creutin, J. D., & Borga, M. (2017). Spatial estimation of debris flows-triggering rainfall and its dependence on rainfall return period. *Geomorphology*, *278*, 269–279. https://doi.org/10.1016/j.geomorph.2016.11.019

Godt, J. W., Baum, R. L., & Chleborad, A. F. (2006). Rainfall characteristics for shallow landsliding in Seattle, Washington, USA. *Earth Surface Processes and Landforms*, *31*(1), 97–110. https://doi.org/10.1002/esp.1237

Guzzetti, F., Peruccacci, S., Rossi, M., & Stark, C. P. (2008). The rainfall intensity–duration control of shallow landslides and debris flows: an update. *Landslides*, *5*(1), 3–17. https://doi.org/10.1007/s10346-007-0112-1

Guzzetti, F., Gariano, S. L., Peruccacci, S., Brunetti, M. T., Marchesini, I., Rossi, M., & Melillo, M. (2020). Geographical landslide early warning systems. *Earth-Science Reviews*, *200*, 102973. https://doi.org/10.1016/j.earscirev.2019.102973

Hefley, T. J., Baasch, D. M., Tyre, A. J., & Blankenship, E. E. (2015). Use of opportunistic sightings and expert knowledge to predict and compare Whooping Crane stopover habitat. *Conservation Biology*, *29*(5), 1337–1346. https://doi.org/10.1111/cobi.12515

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K.-L., Joyce, R. J., Kidd, C., et al. (2020). Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG). In V. Levizzani, C. Kidd, D. B. Kirschbaum, C. D. Kummerow, K. Nakamura, & F. J. Turk (Eds.), *Satellite Precipitation Measurement: Volume 1* (pp. 343–353). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-24568-9_19

Iadanza, C., Trigila, A., & Napolitano, F. (2016). Identification and characterization of rainfall events responsible for triggering of debris flows and shallow landslides. *Journal of Hydrology*, *541*, 230–245. https://doi.org/10.1016/j.jhydrol.2016.01.018

Kirschbaum, D., Stanley, T., & Zhou, Y. (2015). Spatial and temporal analysis of a global landslide catalog. *Geomorphology*, *249*, 4–15. https://doi.org/10.1016/j.geomorph.2015.03.016

Marc, O., Oliveira, R. A. J., Gosset, M., Emberson, R., & Malet, J.-P. (2022). Global assessment of the capability of satellite precipitation products to retrieve landslide-triggering extreme rainfall events. *Earth Interactions*, *1*(aop), 1–42. https://doi.org/10.1175/EI-D-21-0022.1

Mirus, B. B., Becker, R. E., Baum, R. L., & Smith, J. B. (2018). Integrating real-time subsurface hydrologic monitoring with empirical rainfall thresholds to improve landslide early warning. *Landslides*, *15*(10), 1909–1919. https://doi.org/10.1007/s10346-018-0995-z

Nikolopoulos, E. I., Borga, M., Marra, F., Crema, S., & Marchi, L. (2015). Debris flows in the eastern Italian Alps: seasonality and atmospheric circulation patterns. *Natural Hazards and Earth System Sciences*, *15*(3), 647–656. https://doi.org/10.5194/nhess-15-647-2015

Nolde, N., & Joe, H. (2013). A Bayesian extreme value analysis of debris flows. *Water Resources Research*, *49*(10), 7009–7022. https://doi.org/10.1002/wrcr.20494

Peres, D. J., & Cancelliere, A. (2021). Comparing methods for determining landslide early warning thresholds: potential use of non-triggering rainfall for locations with scarce landslide data availability. *Landslides*, *18*(9), 3135–3147. https://doi.org/10.1007/s10346-021-01704-7

Peruccacci, S., Brunetti, M. T., Gariano, S. L., Melillo, M., Rossi, M., & Guzzetti, F. (2017). Rainfall thresholds for possible landslide occurrence in Italy. *Geomorphology*, *290*, 39–57. https://doi.org/10.1016/j.geomorph.2017.03.031

Scheevel, C. R., Baum, R. L., Mirus, B. B., & Smith, J. B. (2017). *Precipitation thresholds for landslide occurrence near Seattle, Mukilteo, and Everett, Washington* (USGS Numbered Series No. 2017–1039). *Precipitation thresholds for landslide occurrence near Seattle, Mukilteo, and Everett, Washington* (Vol. 2017–1039, p. 60). Reston, VA: U.S. Geological Survey. https://doi.org/10.3133/ofr20171039

Segoni, S., Rosi, A., Rossi, G., Catani, F., & Casagli, N. (2014). Analysing the relationship between rainfalls and landslides to define a mosaic of triggering thresholds for regional-scale warning systems. *Natural Hazards and Earth System Sciences*, *14*(9), 2637–2648. https://doi.org/10.5194/nhess-14-2637-2014

Segoni, S., Piciullo, L., & Gariano, S. L. (2018). A review of the recent literature on rainfall thresholds for landslide occurrence. *Landslides*, *15*(8), 1483–1501. https://doi.org/10.1007/s10346-018-0966-4

Sharma, A. R., & Déry, S. J. (2020). Variability and trends of landfalling atmospheric rivers along the Pacific Coast of northwestern North America. *International Journal of Climatology*, *40*(1), 544–558. https://doi.org/10.1002/joc.6227

Steger, S., Brenning, A., Bell, R., & Glade, T. (2017). The influence of systematically incomplete shallow landslide inventories on statistical susceptibility models and suggestions for improvements. *Landslides*, *14*(5), 1767–1781. https://doi.org/10.1007/s10346-017-0820-0

Steger, S., Moreno, M., Crespi, A., Zellner, P. J., Gariano, S. L., Brunetti, M. T., et al. (2022). Deciphering seasonal effects of triggering and preparatory precipitation for improved shallow landslide prediction using generalized additive mixed models. *Natural Hazards and Earth System Sciences Discussions*, 1–38. https://doi.org/10.5194/nhess-2022-271

Thomas, M. A., Mirus, B. B., & Collins, B. D. (2018). Identifying Physics-Based Thresholds for Rainfall-Induced Landsliding. *Geophysical Research Letters*, *45*(18), 9651–9661. https://doi.org/10.1029/2018GL079662

Thomas, M. A., Collins, B. D., & Mirus, B. B. (2019). Assessing the Feasibility of Satellite-Based Thresholds for Hydrologically Driven Landsliding. *Water Resources Research*, *55*(11), 9006–9023. https://doi.org/10.1029/2019WR025577

Warton, D. I., & Shepherd, L. C. (2010). Poisson point process models solve the "pseudo-absence problem" for presence-only data in ecology. *Annals of Applied Statistics*, *4*(3), 1383–1402. https://doi.org/10.1214/10-AOAS331

Wicki, A., Jansson, P.-E., Lehmann, P., Hauck, C., & Stähli, M. (2021). Simulated or measured soil moisture: which one is adding more value to regional landslide early warning? *Hydrology and Earth System Sciences*, *25*(8), 4585–4610. https://doi.org/10.5194/hess-25-4585-2021

Zêzere, J. L., Vaz, T., Pereira, S., Oliveira, S. C., Marques, R., & Garcia, R. A. C. (2015). Rainfall thresholds for landslide activity in Portugal: a state of the art. *Environmental Earth Sciences*, *73*(6), 2917–2936. https://doi.org/10.1007/s12665-014-3672-0

# 6 SUPPLEMENTARY MATERIALS

The following sections contain supplementary materials for each of the three core chapters:

- Supplementary Information – Chapter 2
- Extended Data Figures and Extended Data Tables for Chapter 3
- Supplementary Information – Chapter 4

# Supplementary Information – Chapter 2

# Landslide initiation thresholds in data sparse regions: Application to landslide early warning criteria in Sitka, Alaska, USA

Patton, Annette I.[1,2], Luna, Lisa V. [3,4,5], Roering, Joshua J.[2], Jacobs, Aaron[6], Korup, Oliver[3,4], Mirus, Benjamin B.[7]

[1]Sitka Sound Science Center, Sitka, Alaska, USA.
[2]University of Oregon, Eugene, Oregon, USA.
[3]Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany.
[4]Institute of Geosciences, University of Potsdam, Potsdam, Germany.
[5]Potsdam Institute for Climate Impact Research, Potsdam, Germany.
[6]NOAA National Weather Service Forecast Office Juneau, Alaska, USA.
[7]U.S. Geological Survey, Geologic Hazards Science Center, Golden, Colorado, USA.

*Correspondence to*: Annette Patton (apatton@sitkascience.org) and Lisa Luna (luna@uni-potsdam.de)

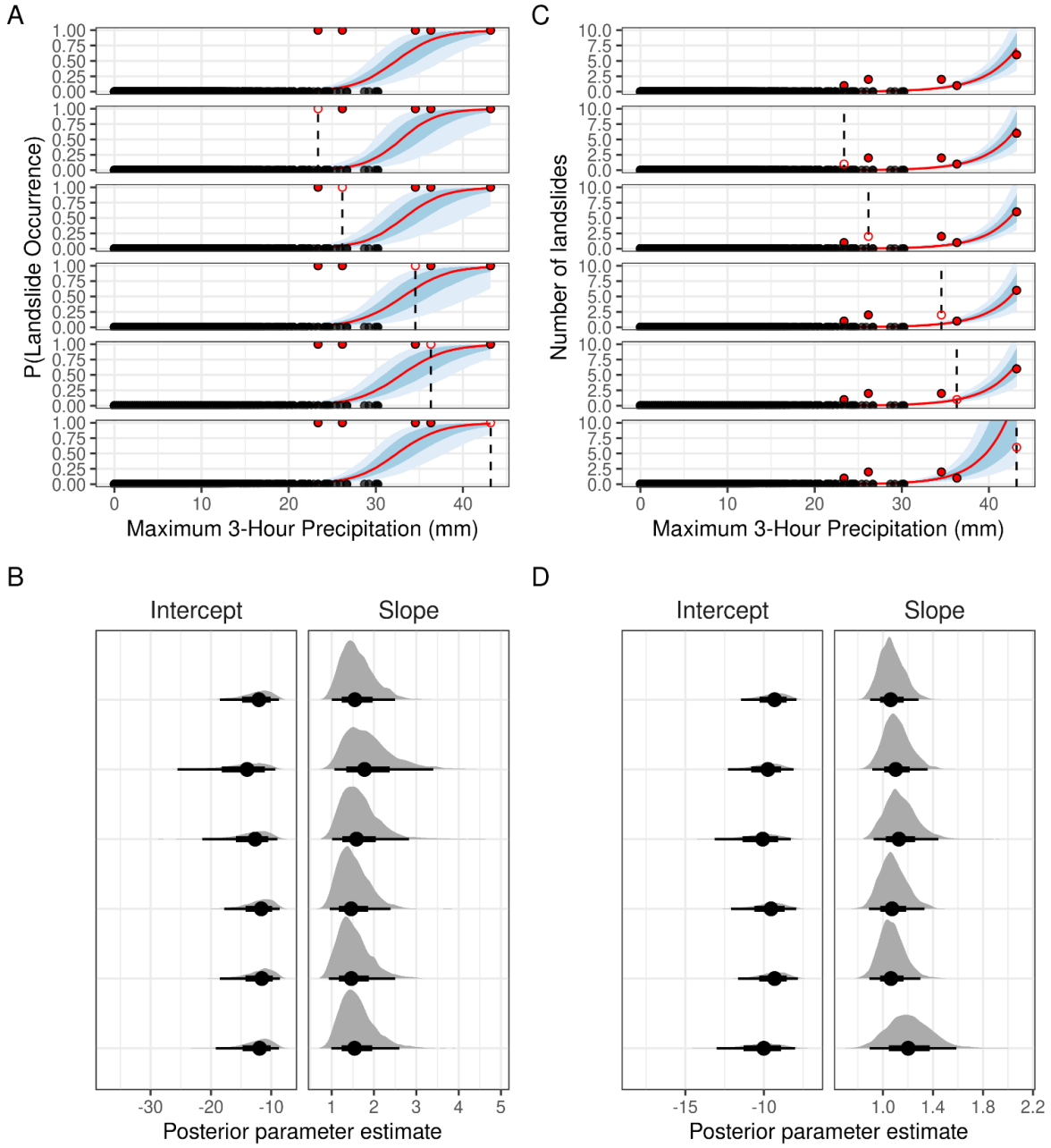



**Fig. S2.1: Leave-one-out cross validation of the preferred (A) Bayesian 3-hour model (BL-3H) and (C) Poisson 3-hour model (PL-3H). Solid red points are landslide events, black points are non-landslide events, red lines show median posterior model estimates, and the dark and light blue shaded regions show the 66% and 95% credibility intervals, respectively. Hollow red circles and dashed black line show the landslide event that was omitted from each run. The 95% High Density Interval (HDI) in panel C indicates the expected average number of landslides at each precipitation value. (B) Logistic regression posterior parameter estimates with one removed landslide event (second to sixth rows) are not credibly distinguishable from the model trained on all points (top row). (D) Poisson regression posterior parameter estimates are most sensitive to the event with six landslides, but are also not credibly distinguishable from the model trained on all points. The gray shaded area in panels B and D show the posterior parameter distributions; the point is the median parameter estimate and the thicker and thinner lines show the 66% and 95% credibility intervals, respectively. Note that the posterior parameter estimates in panels B and D refer to standardized data and are thus not directly comparable to the frequentist parameter estimates in Figure 10.**

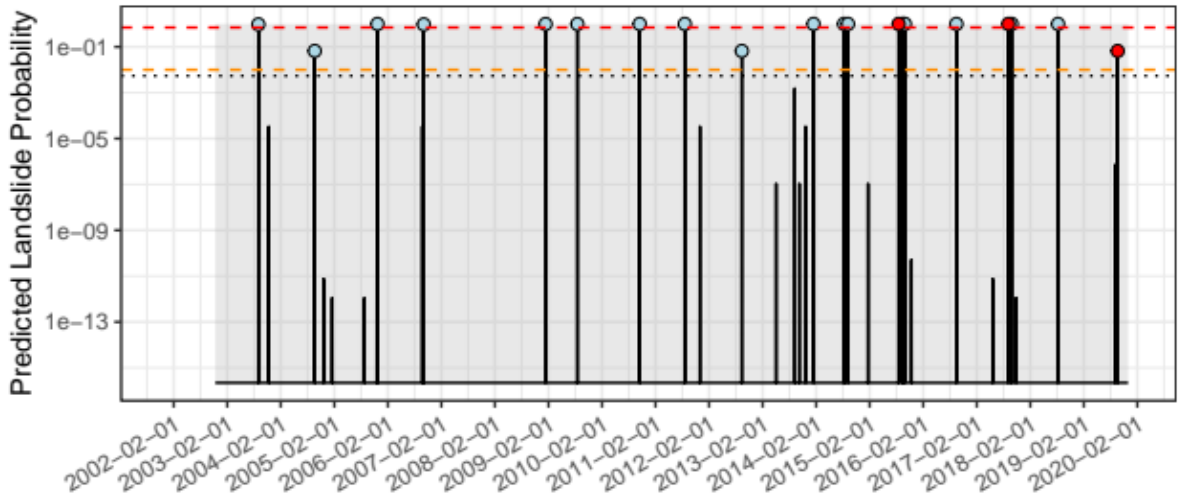



**Fig. S2.2. Analog to Figure 2.12b in the main text, with the training and testing sets reversed. This figure shows testing results for the period 2002 – November 2019, similar to model FL-TT-3H but trained on only one year of data: December 2019 – November 2020. Light blue points indicate false alarms; red points indicate true alarms. No missed alarms would have occurred, and the remaining days are true no alarms. Dashed lines show the upper threshold where estimated landslide probability = 0.70 (red); the lower threshold where landslide probability = 0.01 (orange); and the historical daily landslide probability = 0.0007 (black dotted line).**

**Table S2.1: Confusion matrix for 2002-November 2019 predictions, based on model FL-TT-3H trained on December 2019-November 2020 and with thresholds at probabilities of 0.01 and 0.7, showing the number of times each warning level would have been reached and the actual outcome.**

| | Low | Moderate | High |
|---|---|---|---|
| Landslide | 0 | 1 | 2 |
| No landslide | 6206 | 2 | 16 |

*Extended Data Figures and Extended Data Tables for Chapter 3*

# Globally similar rainfall thresholds for urban landslides

Lisa V. Luna[1,2,3], Maria Isabel Arango Carmona[1], Georg Veh[1], Elizabeth Lewis[4] , Ugur Ozturk[1,5], Oliver Korup[1,2]

[1]Institute of Environmental Science and Geography, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany.
[2]Institute of Geosciences, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany.
[3]Potsdam Institute for Climate Impact Research, Telegrafenberg, 14473 Potsdam, Germany.
[4]School of Engineering, Newcastle University, Newcastle Upon Tyne, UK, NE1 7RU
[5]Helmholtz-Centre Potsdam–GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany.

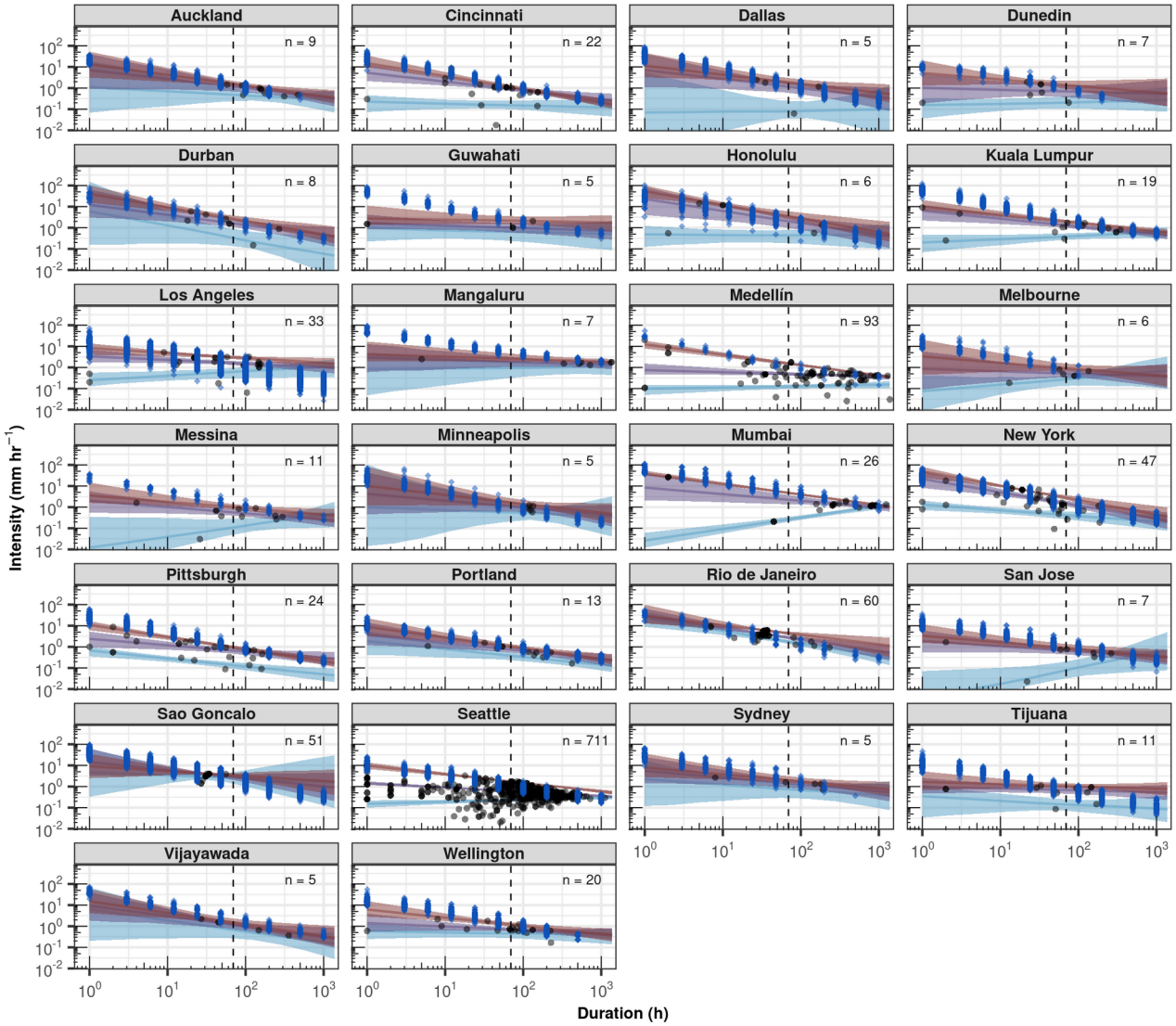
## Contents of this file



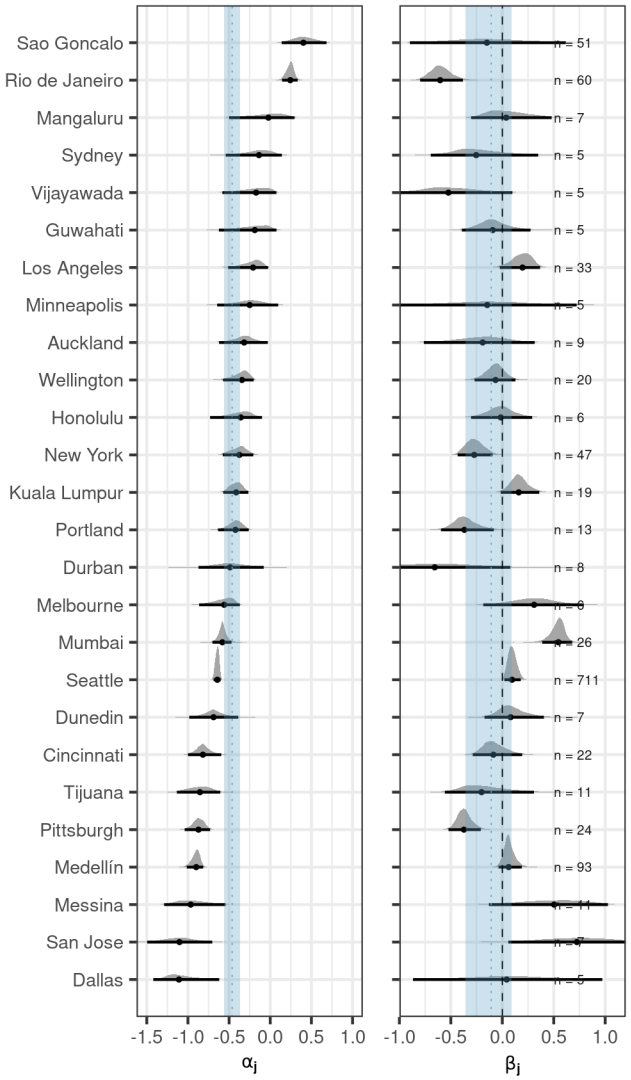
## List of figures:

**List of tables:**

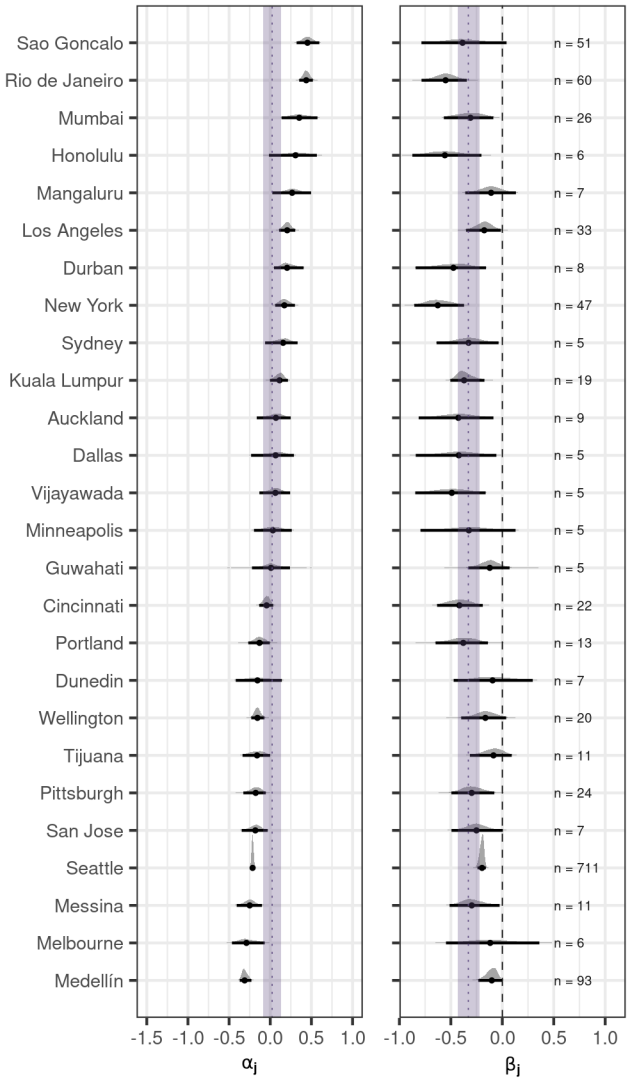
- EDT3.1 – Summary of number of landslides and gauges used for each city
- EDT3.2 – Posterior parameter estimates for global thresholds
- EDT3.3 – City-level posterior parameter estimates for $q_{10}$ thresholds
- EDT3.4 – City-level posterior parameter estimates for $q_{50}$ thresholds
- EDT3.5 – City-level posterior parameter estimates for $q_{90}$ thresholds



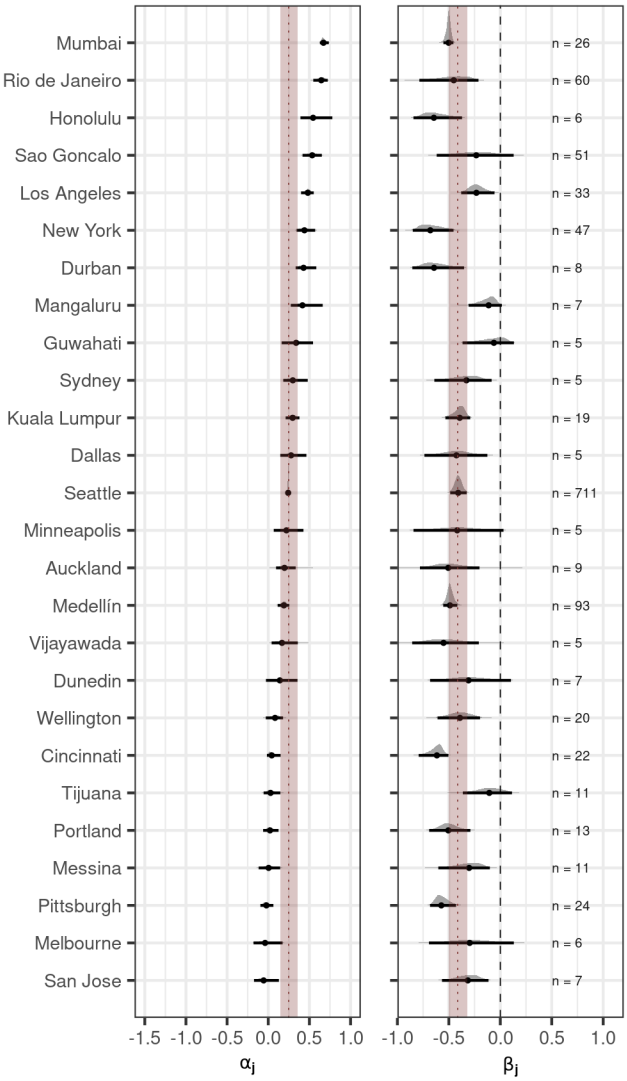

**Extended Data Figure 3.1. City-level intensity-duration thresholds for all cities compared to annual maximum rainfall intensities.** Estimated $q_{10}$ (light blue), $q_{50}$ (purple), and $q_{90}$ (red) thresholds for all cities. Lines are the posterior median; shaded regions show the 95% HDI. Blue diamonds are the maximum recorded precipitation intensities at nearby gauges for each year on record at a range of durations, and black points show reported landslide triggering rain events. The dashed line is the average duration of landslide triggering rain across cities, and n refers to the number of reported landslides.



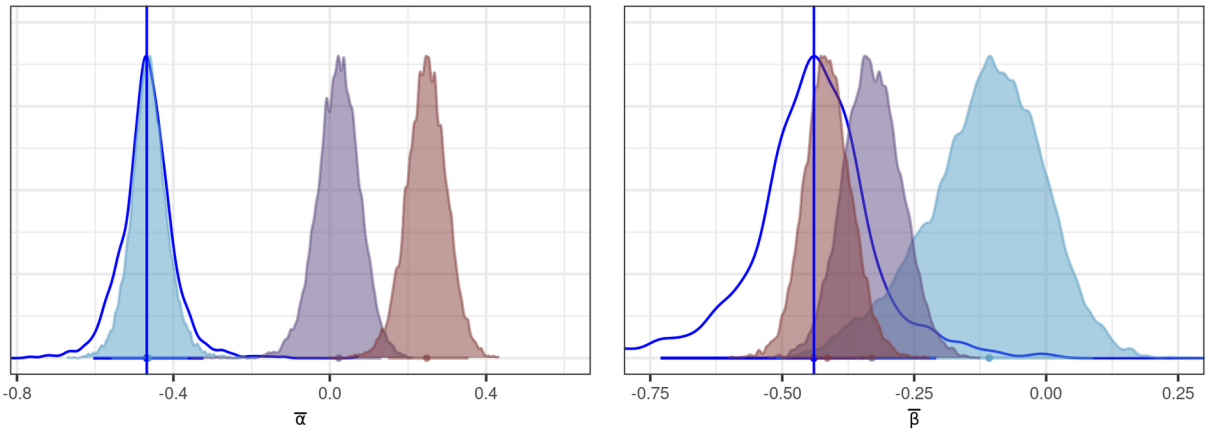

**Extended Data Figure 3.2. Posterior parameter estimates by city for $q_{10}$ thresholds.** $\alpha_j$ represents the posterior estimated $\log_{10}$ intensity at the average duration for each city, $\beta_j$ is the change in order of magnitude in $\log_{10}$ intensity with one change in order of magnitude of $\log_{10}$ duration. Black points are posterior medians, black bars are 95% HDIs. The blue line (median) and shaded region (95% HDI) are the global mean threshold parameter estimates ($\bar{\alpha}$ and $\bar{\beta}$). The dashed black line is a visual aid at 0 (no change in $\log_{10}$ intensity with $\log_{10}$ duration), and n indicates the number of reported landslides.



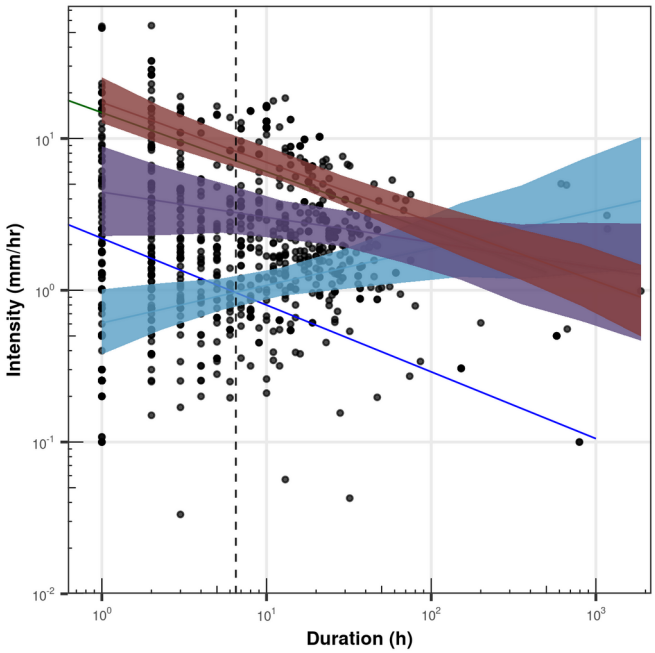

**Extended Data Figure 3.3. Posterior parameter estimates by city for $q_{50}$ thresholds.** $\alpha_j$ represents the posterior estimated $\log_{10}$ intensity at the average duration for each city, $\beta_j$ is the change in order of magnitude in $\log_{10}$ intensity with one change in order of magnitude of $\log_{10}$ duration. Black points are posterior medians, black bars are 95% HDIs. The purple line (median) and shaded region (95% HDI) are the global mean threshold parameter estimates ($\bar{\alpha}$ and $\bar{\beta}$). The dashed black line is a visual aid at 0 (no change in $\log_{10}$ intensity with $\log_{10}$ duration), and n indicates the number of reported landslides.



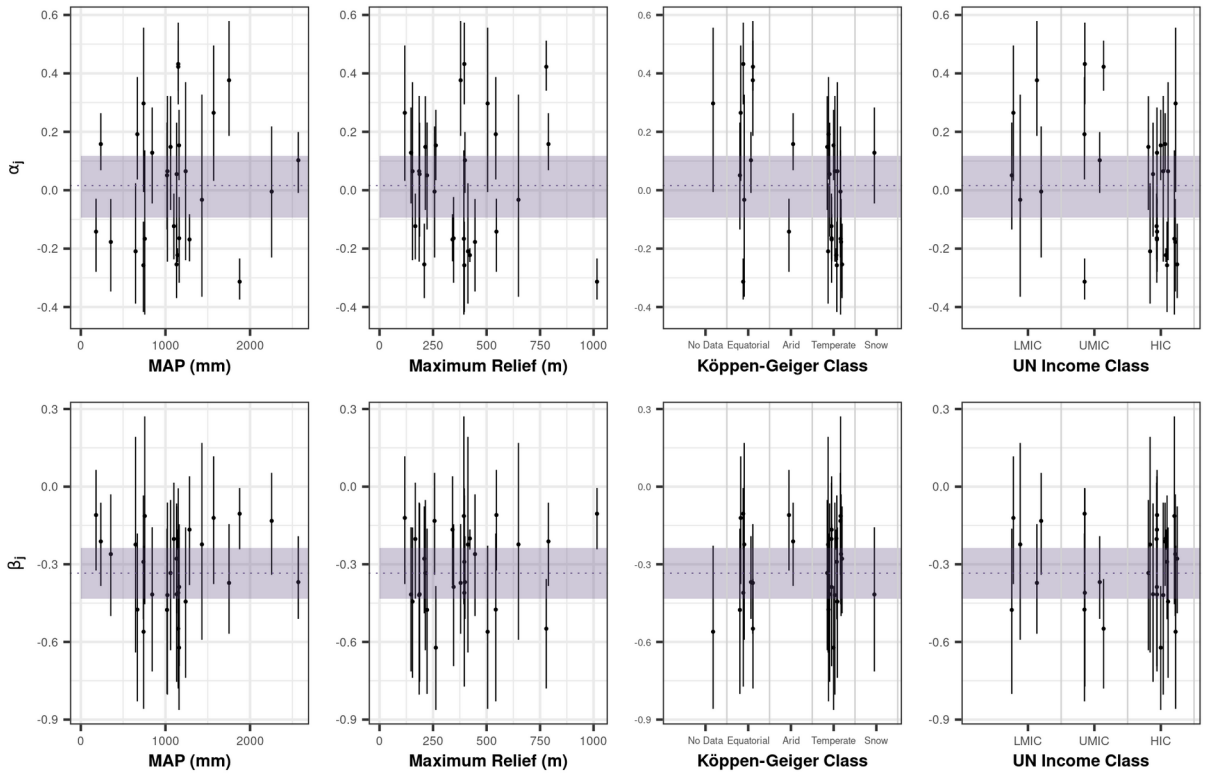

**Extended Data Figure 3.4. Posterior parameter estimates by city for $q_{90}$ thresholds.** $\alpha_j$ represents the posterior estimated $\log_{10}$ intensity at the average duration for each city, $\beta_j$ is the change in order of magnitude in $\log_{10}$ intensity with one change in order of magnitude of $\log_{10}$ duration. Black points are posterior medians, black bars are 95% HDIs. The purple line (median) and shaded region (95% HDI) are the global mean threshold parameter estimates ($\bar{\alpha}$ and $\bar{\beta}$). The dashed black line is a visual aid at 0 (no change in $\log_{10}$ intensity with $\log_{10}$ duration), and n indicates the number of reported landslides.



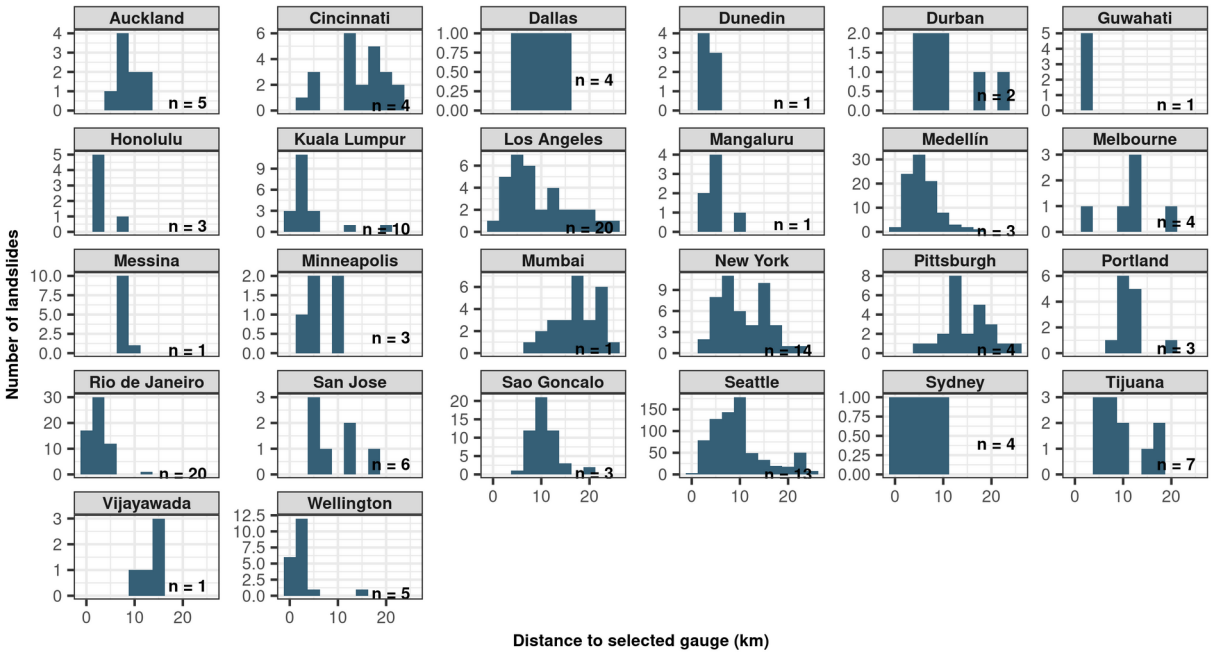

**Extended Data Figure 3.5. Comparison of posterior global thresholds for urban landslides to the prior global threshold.** $\bar{\alpha}$ is the global mean $\log_{10}$ intensity at the average $\log_{10}$ duration; $\bar{\beta}$ is the change in order of magnitude in $\log_{10}$ intensity with one order of magnitude change in $\log_{10}$ duration. The vertical blue line is the point estimate for the global threshold estimated by Guzzetti et al., 2008, as the mean of a Student-*t* informative prior distribution. The light blue, purple, and red shaded regions show the posterior global $q_{10}$, $q_{50}$, and $q_{90}$ threshold parameter estimates, respectively. Points are the posterior median, bars are the 95% HDI.



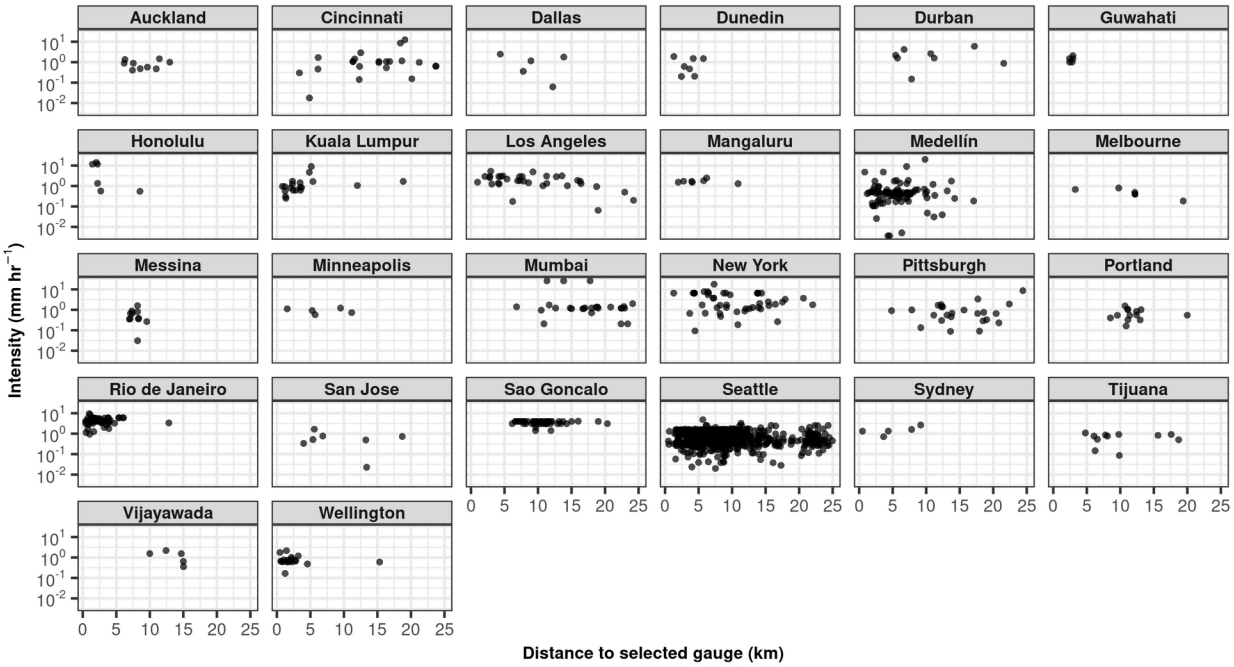

**Extended Data Figure 3.6. Global thresholds with alternative event rainfall definition.** Posterior global $q_{10}$ (light blue), $q_{50}$ (purple), and $q_{90}$ (red) rainfall intensity-duration thresholds when a dry period of at least 3 hours is used to define the onset of event rainfall. Shaded regions are the 95% HDI, and lines are the median. Black points show the intensity and duration of landslide triggering rainfall events. The blue line is the prior threshold from Guzzetti et al., 2008, the green line from Caine, 1980. The vertical dashed line is the average $\log_{10}$ duration across all urban landslides.




**Extended Data Figure 3.7. Posterior parameter estimates for city-level $q_{50}$ threshold compared to mean annual precipitation, maximum relief in urban area, Köppen-Geiger climate class, and UN income group.** For each city, black points are the posterior median, bars are the 95% HDI. The purple line (median) and shaded region (95% HDI) are the global mean threshold parameter estimates ($\bar{\alpha}$ and $\bar{\beta}$).




**Extended Data Figure 3.8. Distance to selected gauge for all landslides in each city.** n indicates the number of unique gauges associated with landslides.




**Extended Data Figure 3.9. Landslide-triggering rainfall intensity compared to distance to selected gauge.** Black points indicate reported landslides in each city.

**Extended Data Table 3.1. Summary of number of landslides and gauges used for each city**

| City | Number of reported landslides with available rainfall data | Year of earliest reported landslide | Year of latest reported landslide | Number of gauges associated with landslides | Year of earliest rainfall record | Year of latest rainfall record |
|---|---|---|---|---|---|---|
| **Auckland** | 9 | 1997 | 2017 | 5 | 1966 | 2018 |
| **Cincinnati** | 22 | 2008 | 2016 | 4 | 1951 | 2015 |
| **Dallas** | 5 | 2007 | 2016 | 4 | 1948 | 2015 |
| **Dunedin** | 7 | 1998 | 2018 | 1 | 1998 | 2018 |
| **Durban** | 8 | 2009 | 2016 | 2 | 1993 | 2017 |
| **Guwahati** | 5 | 2004 | 2005 | 1 | 1969 | 2009 |
| **Honolulu** | 6 | 2007 | 2016 | 3 | 1963 | 2015 |
| **Kuala Lumpur** | 19 | 2004 | 2012 | 10 | 1974 | 2011 |
| **Los Angeles** | 33 | 2005 | 2015 | 21 | 1949 | 2015 |
| **Mangaluru** | 7 | 2010 | 2017 | 1 | 1970 | 2016 |
| **Medellín** | 93 | 2017 | 2022 | 2 | 2017 | 2020 |
| **Melbourne** | 6 | 2011 | 2015 | 4 | 1992 | 2014 |
| **Messina** | 11 | 2009 | 2013 | 1 | 2003 | 2015 |
| **Minneapolis** | 5 | 2013 | 2014 | 3 | 1949 | 2015 |
| **Mumbai** | 26 | 2004 | 2013 | 1 | 1970 | 2013 |
| **New York** | 47 | 1952 | 2015 | 14 | 1949 | 2015 |
| **Pittsburgh** | 24 | 2007 | 2016 | 4 | 1900 | 2015 |
| **Portland** | 13 | 2007 | 2015 | 3 | 1949 | 2015 |
| **Rio de Janeiro** | 60 | 2009 | 2013 | 1 | 1997 | 2016 |
| **San Jose** | 7 | 2007 | 2014 | 6 | 1949 | 2015 |
| **Sao Goncalo** | 51 | 2010 | 2010 | 20 | 1997 | 2016 |
| **Seattle** | 711 | 1948 | 2015 | 13 | 1949 | 2015 |
| **Sydney** | 5 | 2011 | 2015 | 4 | 1985 | 2014 |
| **Tijuana** | 11 | 2008 | 2015 | 5 | 1949 | 2015 |
| **Vijayawada** | 5 | 2010 | 2010 | 1 | 1970 | 2009 |
| **Wellington** | 20 | 2002 | 2017 | 5 | 1961 | 2018 |

**Extended Data Table 3.2. Posterior parameter estimates for global thresholds.** To plot the posterior median thresholds: $\log_{10}(I)=\log_{10}(\bar{A})+(\bar{\beta})\log_{10}(D)$

| | $\log_{10}(\bar{A})$ | | | $\bar{\alpha}$ | | | $\bar{\beta}$ | | |
|---|---|---|---|---|---|---|---|---|---|
| **Quantile** | **Posterior 2.5%\| Median \| 97.5% of HDI** | | | **Posterior 2.5%\| Median \| 97.5% of HDI** | | | **Posterior 2.5%\| Median \| 97.5% of HDI** | | |
| **0.1** | -0.6 | -0.2 | 0.3 | -0.6 | -0.5 | -0.4 | -0.4 | -0.1 | 0.1 |
| **0.5** | 0.4 | 0.7 | 0.9 | -0.1 | 0.0 | 0.1 | -0.4 | -0.3 | -0.2 |
| **0.9** | 0.9 | 1.1 | 1.3 | 0.1 | 0.2 | 0.4 | -0.5 | -0.4 | -0.3 |

**Extended Data Table 3.3. City-level posterior parameter estimates for $q_{10}$ thresholds**

| | $\log_{10}(\bar{A})$ | | | $\bar{\alpha}$ | | | $\bar{\beta}$ | | |
|---|---|---|---|---|---|---|---|---|---|
| **City** | **Posterior 2.5%\| Median \| 97.5% of HDI** | | | **Posterior 2.5%\| Median \| 97.5% of HDI** | | | **Posterior 2.5%\| Median \| 97.5% of HDI** | | |
| **Auckland** | -1.2 | 0.0 | 1.3 | -0.6 | -0.3 | 0.0 | -0.8 | -0.2 | 0.3 |
| **Cincinnati** | -1.1 | -0.6 | -0.3 | -1.0 | -0.8 | -0.6 | -0.3 | -0.1 | 0.2 |
| **Dallas** | -2.9 | -1.2 | 0.8 | -1.4 | -1.1 | -0.6 | -0.9 | 0.0 | 1.0 |
| **Dunedin** | -1.4 | -0.8 | -0.5 | -1.0 | -0.7 | -0.4 | -0.2 | 0.1 | 0.4 |
| **Durban** | -0.8 | 0.8 | 2.2 | -0.9 | -0.5 | -0.1 | -1.4 | -0.7 | 0.1 |
| **Guwahati** | -0.7 | 0.0 | 0.3 | -0.6 | -0.2 | 0.1 | -0.4 | -0.1 | 0.3 |
| **Honolulu** | -0.9 | -0.3 | 0.1 | -0.7 | -0.4 | -0.1 | -0.3 | 0.0 | 0.3 |
| **Kuala Lumpur** | -1.2 | -0.7 | -0.3 | -0.6 | -0.4 | -0.3 | 0.0 | 0.2 | 0.4 |
| **Los Angeles** | -0.9 | -0.6 | -0.3 | -0.5 | -0.2 | 0.0 | 0.0 | 0.2 | 0.4 |
| **Mangaluru** | -1.3 | -0.1 | 0.7 | -0.5 | 0.0 | 0.3 | -0.3 | 0.0 | 0.5 |
| **Medellín** | -1.3 | -1.0 | -0.8 | -1.0 | -0.9 | -0.8 | 0.0 | 0.1 | 0.2 |
| **Melbourne** | -2.1 | -1.1 | -0.4 | -0.9 | -0.6 | -0.4 | -0.2 | 0.3 | 0.8 |
| **Messina** | -3.1 | -1.9 | -0.5 | -1.3 | -1.0 | -0.5 | -0.1 | 0.5 | 1.0 |
| **Minneapolis** | -1.8 | 0.0 | 2.0 | -0.6 | -0.2 | 0.1 | -1.1 | -0.1 | 0.7 |
| **Mumbai** | -1.9 | -1.6 | -1.2 | -0.7 | -0.6 | -0.5 | 0.4 | 0.5 | 0.7 |
| **New York** | -0.1 | 0.1 | 0.3 | -0.6 | -0.4 | -0.2 | -0.4 | -0.3 | -0.1 |
| **Pittsburgh** | -0.5 | -0.2 | 0.0 | -1.0 | -0.9 | -0.7 | -0.5 | -0.4 | -0.2 |
| **Portland** | -0.4 | 0.3 | 0.7 | -0.6 | -0.4 | -0.3 | -0.6 | -0.4 | -0.1 |
| **Rio de Janeiro** | 0.9 | 1.4 | 1.7 | 0.1 | 0.2 | 0.3 | -0.8 | -0.6 | -0.4 |
| **San Jose** | -3.5 | -2.5 | -1.1 | -1.5 | -1.1 | -0.7 | 0.1 | 0.7 | 1.3 |
| **Sao Goncalo** | -0.5 | 0.7 | 1.8 | 0.1 | 0.4 | 0.7 | -0.9 | -0.1 | 0.6 |
| **Seattle** | -1.0 | -0.8 | -0.7 | -0.7 | -0.6 | -0.6 | 0.0 | 0.1 | 0.2 |
| **Sydney** | -0.9 | 0.3 | 1.2 | -0.5 | -0.1 | 0.1 | -0.7 | -0.3 | 0.3 |
| **Tijuana** | -1.5 | -0.5 | 0.1 | -1.1 | -0.9 | -0.6 | -0.6 | -0.2 | 0.3 |
| **Vijayawada** | -0.7 | 0.8 | 1.8 | -0.6 | -0.2 | 0.1 | -1.1 | -0.5 | 0.1 |
| **Wellington** | -0.6 | -0.2 | 0.1 | -0.6 | -0.3 | -0.2 | -0.3 | -0.1 | 0.1 |

**Extended Data Table 3.4. City-level posterior parameter estimates for $q_{50}$ thresholds**

| | $\log_{10}(\bar{A})$ | | | $\bar{\alpha}$ | | | $\bar{\beta}$ | | |
|---|---|---|---|---|---|---|---|---|---|
| **City** | **Posterior 2.5%\| Median \| 97.5% of HDI** | | | **Posterior 2.5%\| Median \| 97.5% of HDI** | | | **Posterior 2.5%\| Median \| 97.5% of HDI** | | |
| **Auckland** | 0.0 | 0.9 | 1.6 | -0.2 | 0.1 | 0.2 | -0.8 | -0.4 | -0.1 |
| **Cincinnati** | 0.3 | 0.7 | 1.1 | -0.1 | 0.0 | 0.0 | -0.6 | -0.4 | -0.2 |
| **Dallas** | 0.0 | 0.9 | 1.6 | -0.2 | 0.1 | 0.3 | -0.8 | -0.4 | -0.1 |
| **Dunedin** | -0.7 | 0.0 | 0.7 | -0.4 | -0.2 | 0.1 | -0.5 | -0.1 | 0.3 |
| **Durban** | 0.5 | 1.1 | 1.8 | 0.0 | 0.2 | 0.4 | -0.8 | -0.5 | -0.2 |
| **Guwahati** | -0.1 | 0.2 | 0.6 | -0.2 | 0.0 | 0.2 | -0.3 | -0.1 | 0.1 |
| **Honolulu** | 0.6 | 1.4 | 1.9 | 0.0 | 0.3 | 0.6 | -0.9 | -0.6 | -0.2 |
| **Kuala Lumpur** | 0.3 | 0.8 | 1.1 | 0.0 | 0.1 | 0.2 | -0.5 | -0.4 | -0.2 |
| **Los Angeles** | 0.2 | 0.5 | 0.8 | 0.1 | 0.2 | 0.3 | -0.4 | -0.2 | 0.0 |
| **Mangaluru** | -0.2 | 0.5 | 1.1 | 0.0 | 0.3 | 0.5 | -0.4 | -0.1 | 0.1 |
| **Medellín** | -0.3 | -0.1 | 0.2 | -0.4 | -0.3 | -0.2 | -0.2 | -0.1 | 0.0 |
| **Melbourne** | -1.1 | -0.1 | 0.8 | -0.5 | -0.3 | -0.1 | -0.5 | -0.1 | 0.4 |
| **Messina** | -0.3 | 0.3 | 0.7 | -0.4 | -0.2 | -0.1 | -0.5 | -0.3 | 0.0 |
| **Minneapolis** | -0.3 | 0.6 | 1.6 | -0.2 | 0.0 | 0.3 | -0.8 | -0.3 | 0.1 |
| **Mumbai** | 0.3 | 0.9 | 1.6 | 0.1 | 0.4 | 0.6 | -0.6 | -0.3 | -0.1 |
| **New York** | 0.9 | 1.3 | 1.7 | 0.1 | 0.2 | 0.3 | -0.9 | -0.6 | -0.4 |
| **Pittsburgh** | 0.0 | 0.4 | 0.7 | -0.3 | -0.2 | -0.1 | -0.5 | -0.3 | -0.1 |
| **Portland** | 0.1 | 0.6 | 1.1 | -0.3 | -0.1 | 0.0 | -0.6 | -0.4 | -0.1 |
| **Rio de Janeiro** | 1.1 | 1.5 | 1.8 | 0.4 | 0.4 | 0.5 | -0.8 | -0.6 | -0.3 |
| **San Jose** | -0.3 | 0.3 | 0.8 | -0.3 | -0.2 | 0.0 | -0.5 | -0.3 | 0.0 |
| **Sao Goncalo** | 0.5 | 1.2 | 1.8 | 0.3 | 0.5 | 0.6 | -0.8 | -0.4 | 0.0 |
| **Seattle** | 0.1 | 0.1 | 0.2 | -0.2 | -0.2 | -0.2 | -0.2 | -0.2 | -0.2 |
| **Sydney** | 0.1 | 0.7 | 1.4 | -0.1 | 0.2 | 0.3 | -0.6 | -0.3 | 0.0 |
| **Tijuana** | -0.3 | 0.0 | 0.4 | -0.3 | -0.2 | 0.0 | -0.3 | -0.1 | 0.1 |
| **Vijayawada** | 0.3 | 1.0 | 1.6 | -0.1 | 0.1 | 0.2 | -0.8 | -0.5 | -0.2 |
| **Wellington** | -0.2 | 0.1 | 0.6 | -0.2 | -0.2 | -0.1 | -0.4 | -0.2 | 0.0 |

**Extended Data Table 3.5. City-level posterior parameter estimates for $q_{90}$ thresholds**

| | $\log_{10}(\bar{A})$ | | | $\bar{\alpha}$ | | | $\bar{\beta}$ | | |
|---|---|---|---|---|---|---|---|---|---|
| **City** | **Posterior 2.5%\| Median \| 97.5% of HDI** | | | **Posterior 2.5%\| Median \| 97.5% of HDI** | | | **Posterior 2.5%\| Median \| 97.5% of HDI** | | |
| **Auckland** | 0.5 | 1.1 | 1.7 | 0.1 | 0.2 | 0.3 | -0.8 | -0.5 | -0.2 |
| **Cincinnati** | 1.0 | 1.2 | 1.5 | 0.0 | 0.0 | 0.1 | -0.8 | -0.6 | -0.5 |
| **Dallas** | 0.5 | 1.1 | 1.7 | 0.1 | 0.3 | 0.5 | -0.7 | -0.4 | -0.1 |
| **Dunedin** | 0.1 | 0.7 | 1.3 | 0.0 | 0.1 | 0.4 | -0.7 | -0.3 | 0.1 |
| **Durban** | 1.2 | 1.6 | 2.0 | 0.3 | 0.4 | 0.6 | -0.9 | -0.6 | -0.4 |
| **Guwahati** | 0.1 | 0.4 | 1.0 | 0.2 | 0.3 | 0.5 | -0.4 | -0.1 | 0.1 |
| **Honolulu** | 1.4 | 1.7 | 2.0 | 0.4 | 0.5 | 0.8 | -0.8 | -0.6 | -0.4 |
| **Kuala Lumpur** | 0.8 | 1.0 | 1.3 | 0.2 | 0.3 | 0.4 | -0.5 | -0.4 | -0.3 |
| **Los Angeles** | 0.6 | 0.9 | 1.1 | 0.4 | 0.5 | 0.6 | -0.4 | -0.2 | -0.1 |
| **Mangaluru** | 0.3 | 0.6 | 1.2 | 0.3 | 0.4 | 0.7 | -0.3 | -0.1 | 0.0 |
| **Medellín** | 0.9 | 1.1 | 1.3 | 0.1 | 0.2 | 0.3 | -0.6 | -0.5 | -0.4 |
| **Melbourne** | -0.3 | 0.5 | 1.3 | -0.2 | 0.0 | 0.2 | -0.7 | -0.3 | 0.1 |
| **Messina** | 0.2 | 0.6 | 1.2 | -0.1 | 0.0 | 0.1 | -0.6 | -0.3 | -0.1 |
| **Minneapolis** | 0.1 | 1.0 | 1.9 | 0.1 | 0.2 | 0.4 | -0.8 | -0.4 | 0.0 |
| **Mumbai** | 1.5 | 1.6 | 1.7 | 0.6 | 0.7 | 0.7 | -0.6 | -0.5 | -0.5 |
| **New York** | 1.4 | 1.7 | 1.9 | 0.3 | 0.4 | 0.6 | -0.9 | -0.7 | -0.5 |
| **Pittsburgh** | 0.8 | 1.0 | 1.2 | -0.1 | 0.0 | 0.1 | -0.7 | -0.6 | -0.4 |
| **Portland** | 0.5 | 0.9 | 1.4 | -0.1 | 0.0 | 0.1 | -0.7 | -0.5 | -0.3 |
| **Rio de Janeiro** | 1.1 | 1.5 | 2.0 | 0.5 | 0.6 | 0.7 | -0.8 | -0.5 | -0.2 |
| **San Jose** | 0.2 | 0.5 | 1.1 | -0.2 | -0.1 | 0.1 | -0.6 | -0.3 | -0.1 |
| **Sao Goncalo** | 0.4 | 1.0 | 1.5 | 0.4 | 0.5 | 0.7 | -0.6 | -0.2 | 0.1 |
| **Seattle** | 0.8 | 1.0 | 1.1 | 0.2 | 0.2 | 0.3 | -0.5 | -0.4 | -0.3 |
| **Sydney** | 0.5 | 0.9 | 1.6 | 0.2 | 0.3 | 0.5 | -0.6 | -0.3 | -0.1 |
| **Tijuana** | -0.2 | 0.2 | 0.7 | -0.1 | 0.0 | 0.1 | -0.4 | -0.1 | 0.1 |
| **Vijayawada** | 0.6 | 1.2 | 1.8 | 0.0 | 0.2 | 0.4 | -0.9 | -0.6 | -0.2 |
| **Wellington** | 0.5 | 0.8 | 1.2 | 0.0 | 0.1 | 0.2 | -0.6 | -0.4 | -0.2 |

# Supplementary Information – Chapter 4

*Geophysical Research Letters*

Supporting Information for

**Seasonal landslide activity lags annual precipitation pattern in the Pacific Northwest**

L. V. Luna[1,2,3], O. Korup[1,2]

[1]Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany.

[2]Institute of Geosciences, University of Potsdam, Potsdam, Germany.

[3]Potsdam Institute for Climate Impact Research, Potsdam, Germany.

## Contents of this file



## Introduction

This supplementary information includes:

Table S4.1, which summarizes information about each landslide inventory

Figure S4.1, which shows posterior parameter estimates for group-level variations in the landslide-only multi-inventory models

Figure S4.2, which shows posterior predictive distributions from the landslide-only single-inventory models, similar to Figure 2 panels c and e

Figure S4.3, which shows posterior parameter estimates for the landslide-precipitation negative binomial models

Figure S4.4, which shows posterior parameter estimates for the landslide-precipitation logistic regression models

**Table S1. Landslide Inventory Summary**

| Inventory | Short title | Footprint area ($km^2$) | Reported landslides | Reported landslides with a monthly time-stamp | % of inventory with a monthly time-stamp | Earliest reported landslide | Latest reported landslide | Years covered | % of years with at least 1 reported landslide | Number of unique months with at least 1 reported landslide | % of reported landslides occurring in top 3 months | Source of reported landslides |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NASA Global Landslide Catalog (subset to WA and OR) | NASA GLC | 439795 | 1530 | 1529 | 100 | 1996 | 2019 | 23 | 83 | 125 | 18 | *News articles, highway department reports* |
| Statewide Landslide Information Database for Oregon (SLIDO), release 4.2, historical points dataset | SLIDOh | 255014 | 14139 | 5734 | 41 | 1928 | 2018 | 90 | 58 | 213 | 85 | *Aggregation including public lands and city and county records* |
| Statewide Landslide Information Database for Oregon (SLIDO), release 4.2, landslide polygons dataset | SLIDOd | 255014 | 53311 | 74 | 0.14 | 1996 | 2015 | 19 | 53 | 16 | 74 | *Field, LiDAR, and imagery based landslide mapping* |
| Seattle Historic Landslide Locations | Seattle | 371 | 2100 | 2019 | 96 | 1897 | 2019 | 122 | 89 | 537 | 21 | *Unknown; Geotechnical consulting reports* |
| Washington Landslide Compilation | WLC | 184781 | 45297 | 3491 | 7.7 | 2003 | 2009 | 6 | 50 | 4 | 100 | *Field, LiDAR, and imagery based landslide mapping* |

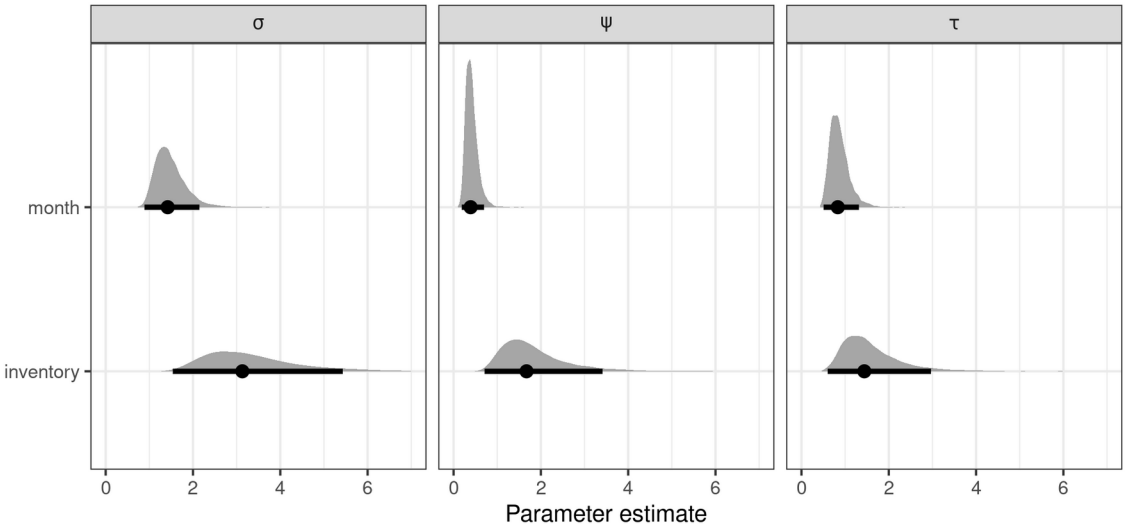



**Figure S4.1. Posterior parameter estimates for landslide-only group-level variations.** Estimates show the learned variation between groups for mean monthly intensity (σ), the negative binomial shape parameter (ψ), and the probability of landsliding (τ) in the multi-inventory models (Eq. 1, 3). In all cases, we observe that the variation between inventories are greater than the variation between months within an inventory.

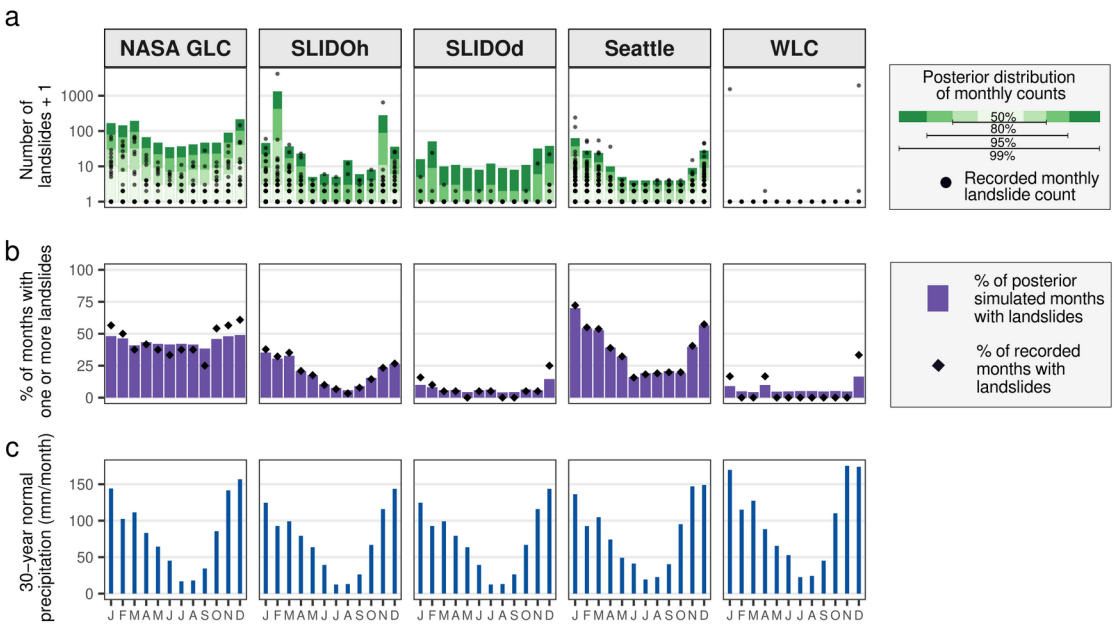



**Figure S4.2. Seasonal pattern of landslide activity from single-inventory models compared to mean monthly precipitation.** (a) Distribution of posterior simulated counts of landslides occurring in each month in the individual inventory areas from the single-inventory negative binomial models; black points represent observed data. The WLC single-inventory model is excluded because the model did not converge. (b) Summary of the posterior simulated distribution from the single-inventory

logistic regressions, reflecting the percentage of predicted months with one or more landslides. Black diamonds show the percentage of recorded months with one or more landslides in each inventory. (c) Mean monthly precipitation over the inventory areas from the 1991 – 2020 PRISM climate normals (PRISM Climate Group, 2021).

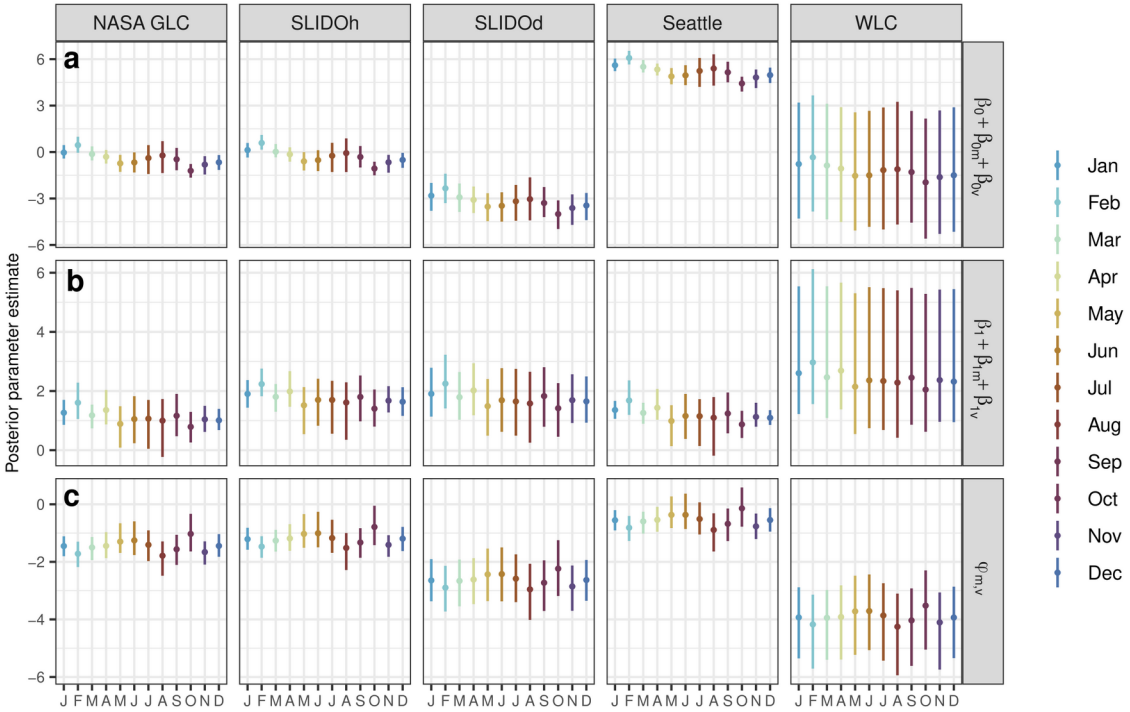



**Figure S4.3. Posterior parameter estimates for the landslide-precipitation negative binomial models.** Points represent the median estimate, bars represent the 95% HDI. We consider estimates credibly distinguishable when the 95% HDI does not overlap. Because these estimates refer to standardized precipitation (Eq. 2), the intercept for each month in each inventory (a) is interpreted as the expected intensity at the mean precipitation value across all months and inventories. We observed credible differences between November and February for the NASA GLC, SLIDOh, and Seattle inventories.

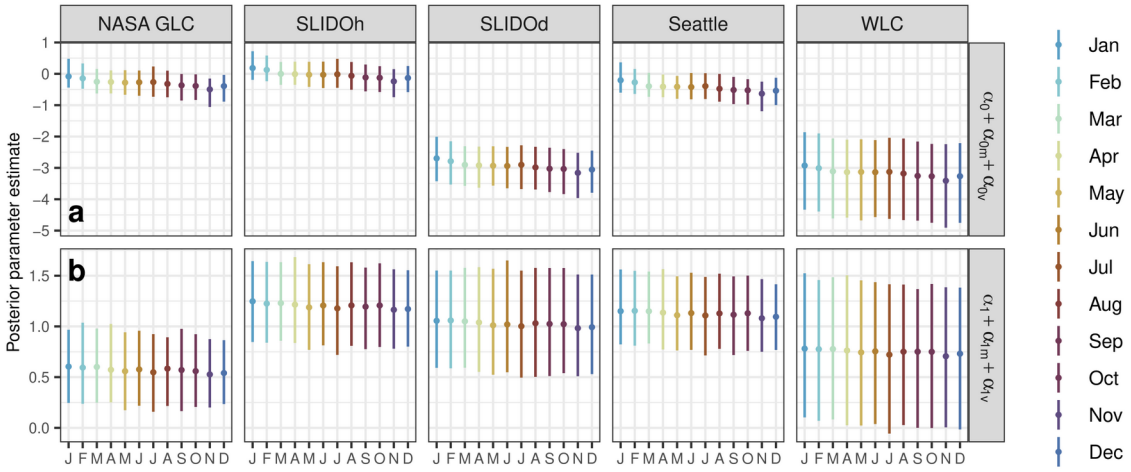



**Figure S4.4. Posterior parameter estimates for the landslide-precipitation logistic regression models.** Points represent the median estimate, bars represent the 95% HDI. Because these estimates refer to standardized precipitation (Eq. 4), the intercept for each month in each inventory (a) is interpreted as the expected probability at the mean precipitation value across all months and inventories. We observed no credible differences between any months in any of the inventories.