

## Supplementary Information

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

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

<sup>1</sup>Sitka Sound Science Center, Sitka, Alaska, USA.

<sup>2</sup>University of Oregon, Eugene, Oregon, USA.

<sup>3</sup>Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany.

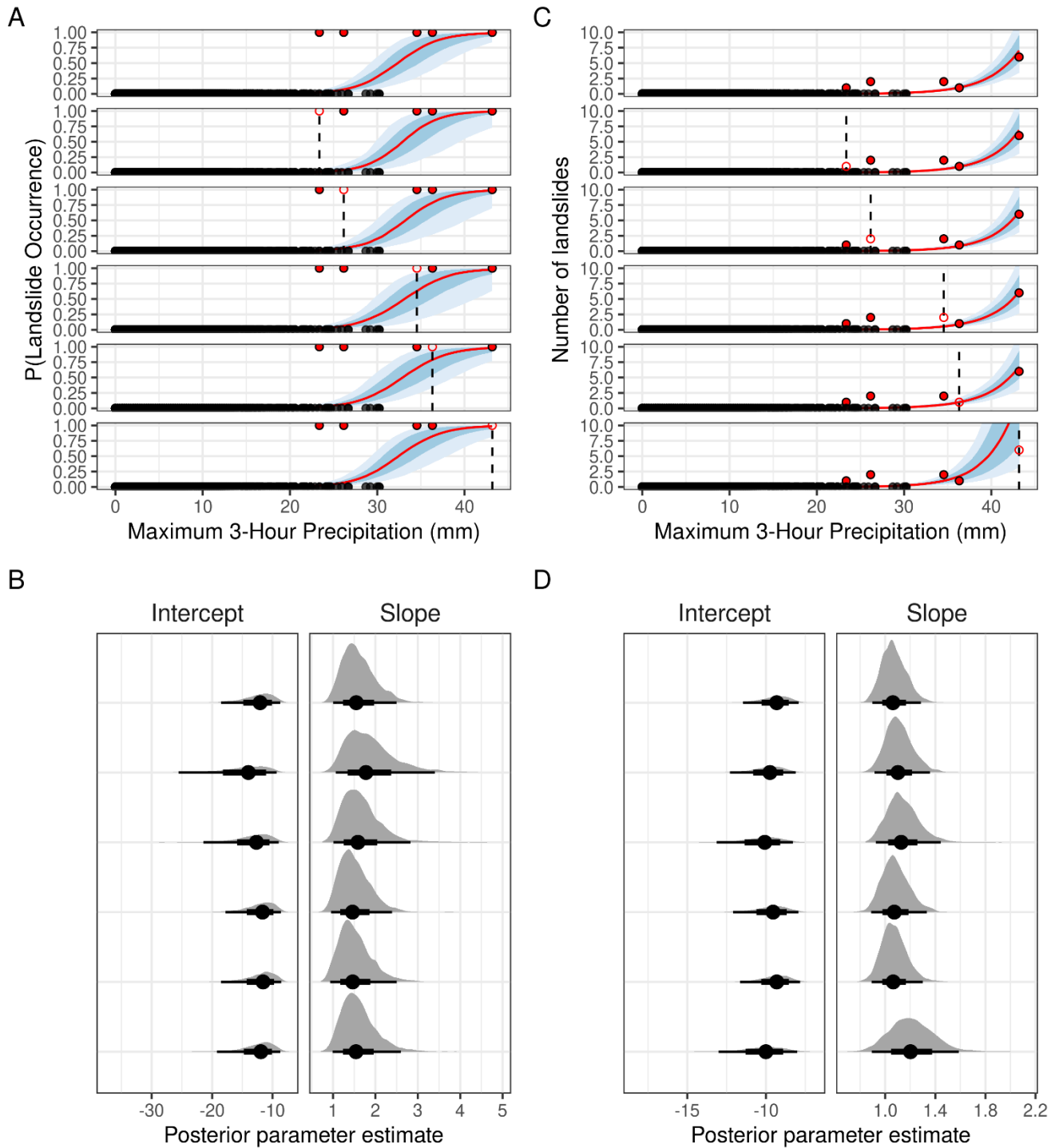
<sup>4</sup>Institute of Geosciences, University of Potsdam, Potsdam, Germany.

10 <sup>5</sup>Potsdam Institute for Climate Impact Research, Potsdam, Germany.

<sup>6</sup>NOAA National Weather Service Forecast Office Juneau, Alaska, USA.

<sup>7</sup>U.S. Geological Survey, Geologic Hazards Science Center, Golden, Colorado, USA.

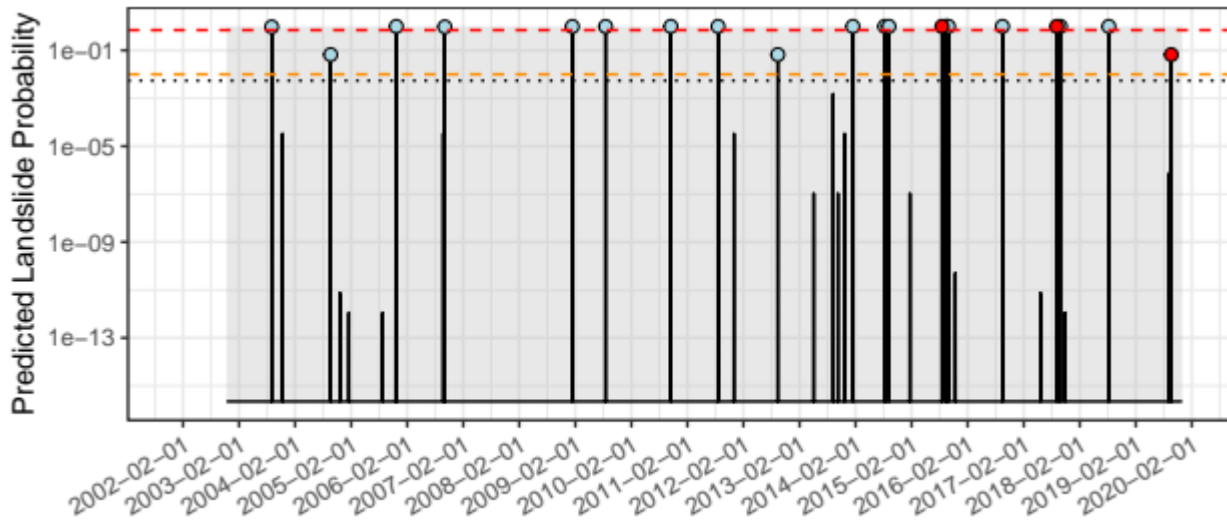
*Correspondence to:* Annette Patton ([apatton@sitkascience.org](mailto:apatton@sitkascience.org)) and Lisa Luna ([luna@uni-potsdam.de](mailto:luna@uni-potsdam.de))



15 **Fig. S1: 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

20

25 estimates in panels B and D refer to standardized data and are thus not directly comparable to the frequentist parameter estimates in Figure 10.



30 Fig. S2. Analog to Figure 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).

35 Table S1: 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