



# Geophysical Research Letters®

## RESEARCH LETTER

10.1029/2022GL100832

## Real-Time Forecast of Catastrophic Landslides via Dragon-King Detection

Qinghua Lei<sup>1,2</sup> , Didier Sornette<sup>3</sup>, Haonan Yang<sup>1</sup>, and Simon Loew<sup>1</sup> 

<sup>1</sup>Department of Earth Sciences, ETH Zürich, Zürich, Switzerland, <sup>2</sup>Department of Earth Sciences, Uppsala University, Uppsala, Sweden, <sup>3</sup>Institute of Risk Analysis, Prediction and Management, Academy for Advanced Interdisciplinary Studies, Southern University of Science and Technology, Shenzhen, China

### Key Points:

- Tertiary creeps of catastrophic landslides accommodate dragon-kings showing statistically different properties than secondary slope creeps
- A predictive framework is developed to forecast catastrophic landslides by detecting signatures typical of the emergence of dragon-kings
- A phase diagram characterizes the detectability of dragon-kings against black-swans and discriminates catastrophic and slow landslides

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

Q. Lei,  
[qinghua.lei@erdw.ethz.ch](mailto:qinghua.lei@erdw.ethz.ch)

### Citation:

Lei, Q., Sornette, D., Yang, H., & Loew, S. (2023). Real-time forecast of catastrophic landslides via dragon-king detection. *Geophysical Research Letters*, 50, e2022GL100832. <https://doi.org/10.1029/2022GL100832>

Received 17 AUG 2022

Accepted 24 FEB 2023

### Author Contributions:

**Conceptualization:** Qinghua Lei, Didier Sornette, Simon Loew

**Data curation:** Qinghua Lei, Haonan Yang, Simon Loew

**Formal analysis:** Qinghua Lei

**Funding acquisition:** Qinghua Lei, Didier Sornette, Simon Loew

**Investigation:** Qinghua Lei, Didier Sornette, Simon Loew

**Methodology:** Qinghua Lei, Didier Sornette, Haonan Yang

**Project Administration:** Qinghua Lei

**Resources:** Didier Sornette, Simon Loew

**Software:** Qinghua Lei, Haonan Yang

© 2023. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs License](https://creativecommons.org/licenses/by/4.0/), which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

**Abstract** Catastrophic landslides characterized by runaway slope failures remain difficult to predict. Here, we develop a physics-based framework to prospectively assess slope failure potential. Our method builds upon the physics of extreme events in natural systems: the extremes so-called “dragon-kings” (e.g., slope tertiary creeps prior to failure) exhibit statistically different properties than other smaller-sized events (e.g., slope secondary creeps). We develop statistical tools to detect the emergence of dragon-kings during landslide evolution, with the secondary-to-tertiary creep transition quantitatively captured. We construct a phase diagram characterizing the detectability of dragon-kings against “black-swans” and informing on whether the slope evolves toward a catastrophic or slow landslide. We test our method on synthetic and real data sets, demonstrating how it might have been used to forecast three representative historical landslides. Our method can in principle considerably reduce the number of false alarms and identify with high confidence the presence of true hazards of catastrophic landslides.

**Plain Language Summary** Catastrophic slope failures that pose great threats to life and property remain difficult to predict due to the strong variability of slope behavior. As a result, only a limited number of large rock slope failures have been so far successfully forecasted with associated risks mitigated. Here, we propose a novel predictive framework to prospectively and quantitatively detect slope failure precursors with high confidence. Our research sheds light on one of the most challenging questions in landslide prediction: Would an active landslide slowly move or catastrophically fail in the future? Our method adds a new conceptual framework and operational methodology with a significant potential to support existing early warning systems and hence reduce landslide risks.

## 1. Introduction

With rapid urbanization and population growth, human habitats suffer from an increased exposure to landslide hazards worldwide (Froude & Petley, 2018). This situation tends to become more severe under climate change, where extreme rainfall, permafrost thaw, and glacier retreat may promote fatal landslides (Kirschbaum et al., 2020; Lacroix et al., 2022; Patton et al., 2021). One of the biggest unknowns in the real-time hazard assessment of an ongoing landslide is whether a catastrophic collapse (i.e., a runaway failure) would occur or not (Palmer, 2017). Forecasting catastrophic landslides is very challenging, because slope movements are highly variable in both space and time, depending on various external drivers and preconditioning factors (Glastonbury & Fell, 2008, 2010; Lacroix et al., 2020). Existing landslide hazard mitigation strategies relying on empirically defined alarm thresholds of rainfall intensity and/or landslide velocity (Crosta & Agliardi, 2002; Crosta et al., 2017; Krøgli et al., 2018) are plagued by large uncertainties. Although great advances have been made over the past decades to develop and deploy high-precision monitoring technologies to observe/detect unstable slope movements (Booth et al., 2013; Crosta et al., 2014; Hu et al., 2020; Manconi et al., 2016), only a limited number of catastrophic landslides have been successfully predicted. Thus, there is a fundamental need to rapidly and reliably forecast catastrophic landslides in real time, so that relevant warnings can be immediately issued to mitigate the associated risks.

We propose that a catastrophic landslide tends to manifest itself as being in the “dragon-king” regime—a double metaphor for an event of a predominant size/impact (like a “king”) and a unique origin (like a “dragon”) (Sornette & Ouillon, 2012). This classification can be justified in the context of classical creep theory (Courtney, 2000), where a creeping solid material prior to failure experiences a transition from a secondary or steady-state creep

**Supervision:** Qinghua Lei, Didier Sornette, Simon Loew  
**Validation:** Qinghua Lei, Haonan Yang  
**Visualization:** Qinghua Lei, Haonan Yang  
**Writing – original draft:** Qinghua Lei  
**Writing – review & editing:** Didier Sornette, Simon Loew

stage (having relatively constant strain rates of low magnitude) to a tertiary or accelerating creep stage (having continuously amplified strain rates of large magnitude). The strain rate bursts during the tertiary creep stage preceding the failure are dragon-kings, because they are much larger in size than those of the secondary creep stage (Main, 2000) and are born of a unique origin (i.e., positive feedback mechanism, where the growth rate of the strain rate increases with the strain rate itself) (Sammis & Sornette, 2002). These dragon-kings are expected to be “outliers” with statistically different properties compared to their smaller siblings (Sornette, 2002; Sornette & Ouillon, 2012), pointing to a possible avenue to predict material failure via dragon-king detection. If there is no positive feedback in the tertiary creep process, large strain rates as non-dragon-king outliers could still appear but do not lead to a catastrophic failure (Main, 1999).

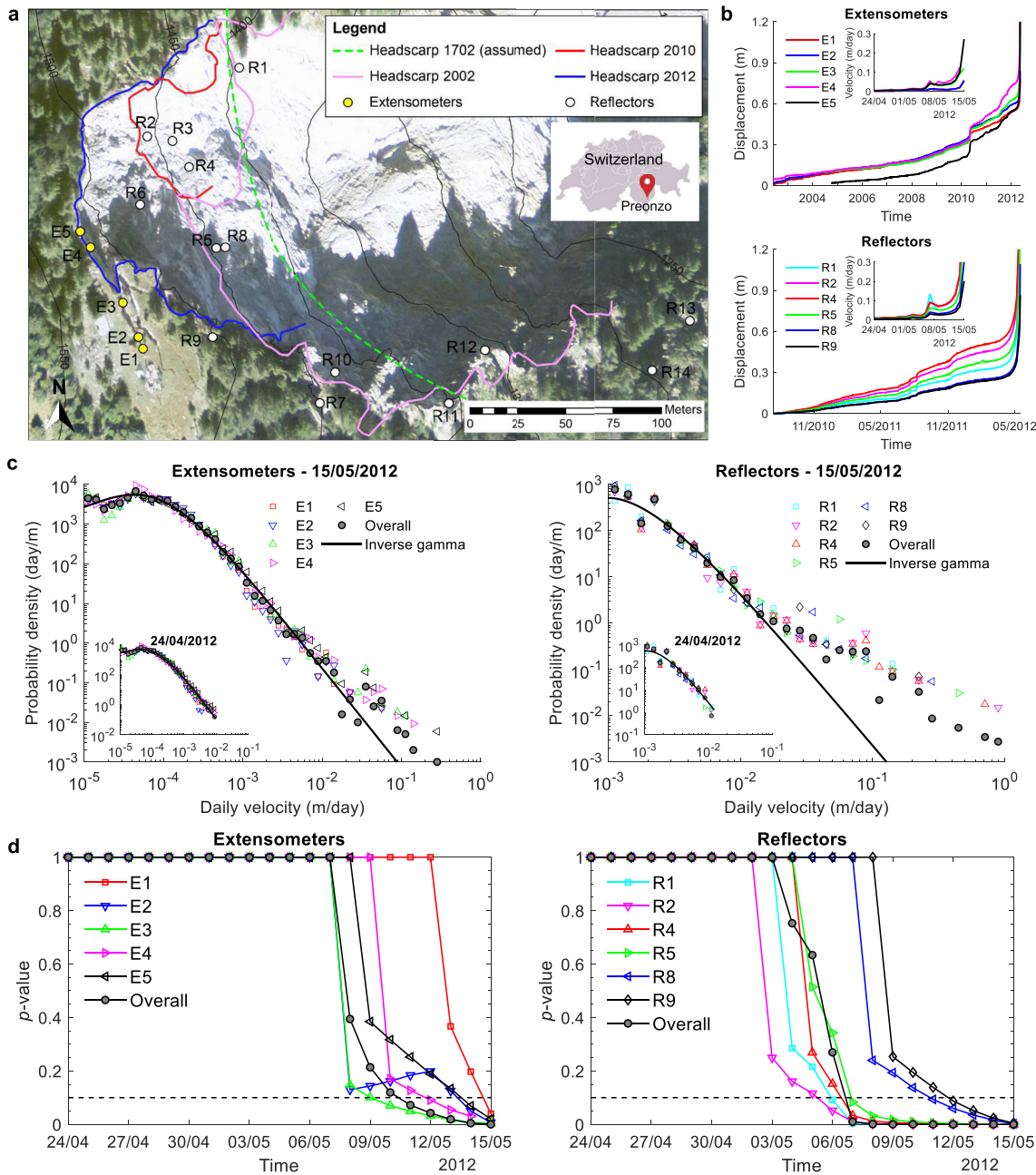
Building upon this paradigm, we develop a physics-based predictive framework for real-time forecast of catastrophic landslides via dragon-king detection. The inputs for our analysis are slope displacement time series that can be recorded by various slope monitoring techniques. We first compute slope velocities on a daily basis and describe their statistical properties using the inverse gamma distribution with parameters estimated based on the profile likelihood method (Text S1 in Supporting Information S1). We use the inverse gamma distribution as a parsimonious parameterization that is convenient to capture the coexistence of a power law (for medium/large velocities related to short-term intermittent creeps) and a rollover transition (for small velocities related to long-term background creeps) in the probability density function (PDF) of slope velocities during the secondary creep stage, as revealed by the actual observational data (see Section 2). Other distribution functions with a similar shape may also be used, for which the following outlier detection procedures are equally applicable. We then define a null hypothesis  $H_0$  that all the velocity data are drawn from the same population, whilst the alternative hypothesis  $H_1$  states that the largest velocities follow a distribution different from that of their smaller siblings. We identify outlier candidates based on the Anderson-Darling distance (Text S2 in Supporting Information S1) and further conduct an outward sequential test (Text S3 in Supporting Information S1) to exclude spurious outliers at a prescribed significance level  $\alpha$ . For the remaining outliers, we run a block test (Text S4 in Supporting Information S1) to calculate the  $p$ -value under  $H_0$ , and if it is less than  $\alpha$ ,  $H_0$  is rejected, indicating that velocity outliers have emerged and the slope has transitioned into a tertiary creep stage undergoing an acceleration crisis. Here,  $\alpha$  defines the probability of type I errors (i.e., false alarms). Considering the significant consequences of missing true hazards which occurs with probability  $\beta$ , where  $1 - \beta$  is the power of the method, we adopt  $\alpha = 0.1$  as a compromise that allows for a smaller  $\beta$ . This  $\alpha = 0.1$  is larger than the conventionally used level of 0.05 (a sensitivity analysis is given in Supporting Information S1), but this is justified to obtain more power. Finally, we use a phase diagram (Text S5 in Supporting Information S1) based on two dimensionless parameters, that is,  $\omega$  (a smaller  $\omega$  represents a system which is more heterogeneous/stochastic) and  $m$  (a larger  $m$  expresses that the slope acceleration is more nonlinear), to predict whether the slope during the acceleration crisis will evolve as a catastrophic landslide or, instead, as a slow landslide with no failure. Here, the discrimination of catastrophic/slow landslides is based on the presence (for  $m > 1$ ) or absence (for  $m \leq 1$ ) of positive feedbacks. In reality, slow landslides may exhibit a wide range of movement rates, from a few mm/yr to more than a hundred m/yr (IUGS, 1995; Lacroix et al., 2020). Catastrophic landslides tend to live in the dragon-king regime where the slope failure is potentially predictable, which is clearly distinct from a “black-swan” regime (Taleb, 2010), where the slope failure is unpredictable and could occur as a surprise (if such events have not yet occurred in the record). Below, we demonstrate our method by applying it to three representative historical landslide events at Preonzo (Switzerland), Veslemannen (Norway), and Moosfluh (Switzerland).

## 2. Landslide Data Sets and Forecast Analysis

### 2.1. Preonzo Landslide, Switzerland

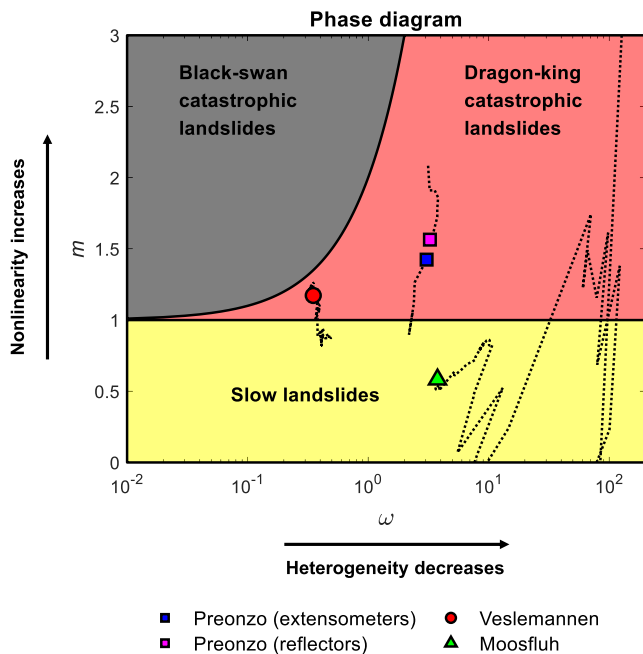
The Preonzo landslide is located at Alpe di Rosciro, above the village of Preonzo in the Riviera valley of southern Switzerland (Figure 1a). This instability complex, composed mainly of augen gneiss, has retrogressively experienced several failures since the 18th century (Gschwind et al., 2019). The collapse in February 1702 destroyed the church of the ancient village of Preonzo below the slope; in May 2002,  $\sim 120,000$  m<sup>3</sup> of rocks at the southern sector was released; in May 2010,  $\sim 20,000$  m<sup>3</sup> of rocks at the northern sector detached from the slope; in May 2012, a suspended large volume of  $\sim 210,000$  m<sup>3</sup> failed catastrophically. Our investigation is focused on the most recent 2012 event.

Since this active landslide imposed a significant hazard to the important industrial facilities and transport routes located directly at the toe of the slope, a long-term monitoring campaign was implemented (Loew et al., 2017).



**Figure 1.** Measurement data and forecast analysis of the 2012 Preonzo landslide, Switzerland. (a) Overview of the slope with the monitoring system and headscarps indicated. (b) Slope displacement time series recorded by extensometers and reflectors (inset: daily velocities for the last 3 weeks before the failure). (c) Probability density function of daily velocities available until 15 May 2012 (inset: 24 April 2012), fitted to the inverse gamma distribution for the extensometer and reflector data, respectively (with outliers excluded). (d) Temporal evolution of  $p$ -values (the dashed line corresponds to a significance level of 0.1).

In 1999, five extensometers (with a measurement accuracy of at least 0.01 mm) were installed to measure the opening of tension cracks in the headscarp area (Figure 1a). In 2010, a robotic total station was installed at the valley floor to measure the slope surface displacement via 14 reflectors (with a measurement accuracy of  $\sim 0.6$  mm) (Figure 1a). Figure 1b shows the slope displacement time series recorded by the five extensometers and seven of the reflectors (other reflectors were either lost much earlier than the 2012 event or located outside the failure area with no relevant response). In summer 2010, an early warning system was implemented to mitigate the risk of slope failure, with velocity threshold-based intervention criteria defined for pre-alarm, public alarm, and evacuation (Loew et al., 2017). On 24 April 2012, a pre-alarm was released, and on 3 May, the industrial



**Figure 2.** Phase diagram of landslide evolution during the slope acceleration crisis. The  $\omega$ - $m$  parameter pair of each landslide is indicated by the corresponding marker with the dotted trajectory illustrating the  $\omega$ - $m$  footprint (see Movies S1, S2, and S3 for animation and Figures S6 and S7 in Supporting Information S1 for time periods). Each marker position represents the state of the landslide at the end of the time series.

at different stages, probably due to variations in slope topography/kinematics (Gschwind et al., 2019). In general, the instruments installed at the northern sector could detect outliers earlier than those at the southern sector, consistent with the landslide kinematic before failure (Gschwind et al., 2019; Loew et al., 2017). Reflectors identified outliers earlier than extensometers, possibly because reflector measurements represent the absolute displacements including internal deformations developed downslope of the headscarp. We consider the slope has entered a crisis (a tertiary creep stage) when the overall  $p$ -value drops below  $\alpha$ . At  $\alpha = 0.1$ , we could confirm the crisis on 7 and 11 May based on the overall  $p$ -values of extensometers and reflectors, respectively (Figure 1d and Movie S1). We then analyze the landslide evolution during the crisis by plotting the  $\omega$ - $m$  trajectories of this landslide in the phase diagram, which indicates that the outliers detected by reflectors were dragon-kings from the start, while those by extensometers also evolved into dragon-kings (Figure 2 and Movie S1). With this information, the resumption of the activities of the industrial facilities on 7–11 May may be considered to have subjected them to high risk. After 12 May, all individual  $p$ -values had considerably decreased, and on 14 May most of them (except that of E1) dropped below 0.1 (Figure 2 and Movie S1), suggesting that the dragon-kings became very mature and the slope was likely to collapse soon, consistent with the actual failure on 15 May. Being clearly in the dragon-king regime in the phase diagram (Figure 2), the  $\omega$ - $m$  footprint confirms the high predictability of this slope failure. Interestingly, the  $\omega$ - $m$  trajectories derived from extensometers and reflectors started from very different initial positions but converged with the approach to failure, implying that the movements of the sliding mass became more uniform prior to the collapse. Our results are compatible with the previous decisions (Loew et al., 2017) of closing the cantonal road on 12 May and the transnational highway on 14 May. Although the conventional velocity threshold-based approach performed almost equally well as our method for the Preonzo landslide (because large velocities only appeared close to failure), it could be problematic for other cases (see below).

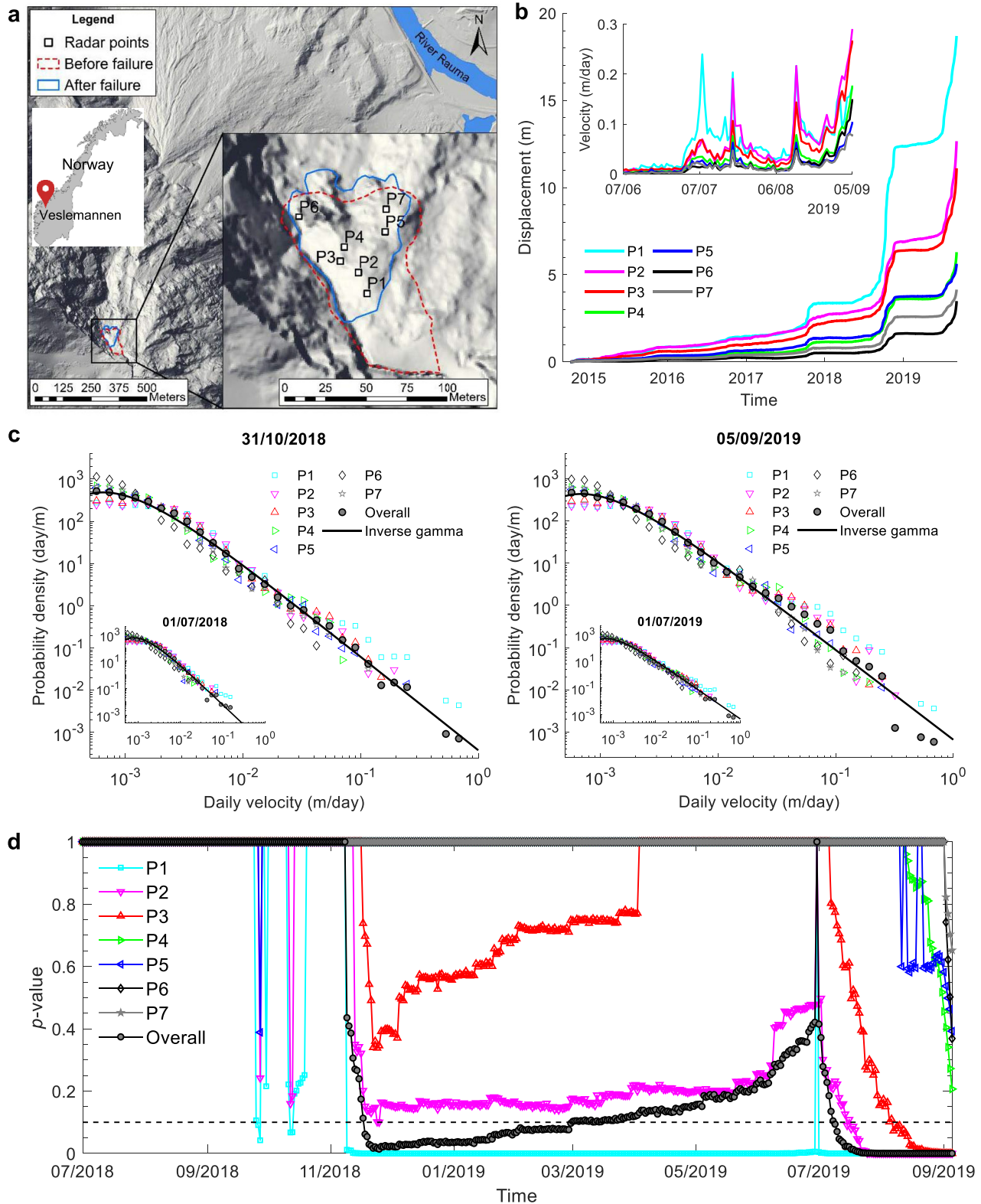
## 2.2. Veslemannen Landslide, Norway

The Veslemannen landslide is situated on a north-facing slope at Romsdalen, western Norway (Figure 3a) and is part of a dormant deep-seated gravitational mountain slope deformation. This currently active slope sector

facilities were evacuated. The landslide significantly accelerated after heavy rainfalls on 5–6 May (Figure 1b, insets). Then, the slope velocity reduced and workers resumed to work in the industrial facilities on 7–11 May. However, on 10–11 May, the slope started to accelerate again under no rainfall conditions. Finally, a public alarm was released with the cantonal road closed on 12 May and the transnational highway closed on 14 May, just before the slope failure on 15 May. This is one of the rare cases where early warning and intervention criteria worked successfully, but the decision-making authority was confronted with large uncertainties when interpreting available displacement data.

We revisit this event using our forecast method. Figure 1c gives the PDF of daily velocities derived from extensometer and reflector measurements until 15 May 2012 (inset shows the PDF of velocity data available until 24 April). We plot the individual PDF for each instrument as well as the overall PDF for all the extensometers/reflectors. The velocities below  $\sim 0.01$  m/day tend to obey an inverse gamma distribution, whereas those above this threshold seem to not belong to the same population. These large velocities only appeared in the last 3 weeks before the failure (Figure 1c and Figure S1 in Supporting Information S1). We perform statistical tests to rigorously determine whether velocity outliers exist and, if so, when they can be detected. We conduct pseudo-prospective predictions, assuming we are on any day before the failure and only the data recorded until that day are known. We define  $H_0$  that the velocity data follow an inverse gamma distribution. We determine the transition velocity, above which velocities are outlier candidates (Figure S2 in Supporting Information S1). We then exclude spurious outliers and calculate the individual/overall  $p$ -values under  $H_0$  for the period between 24 April and 15 May 2012 (Figure 1d). The systematic declines of all  $p$ -values suggest that outliers have emerged since  $H_0$  becomes less and less likely as  $p$ -values decrease. Different instruments sensed outliers





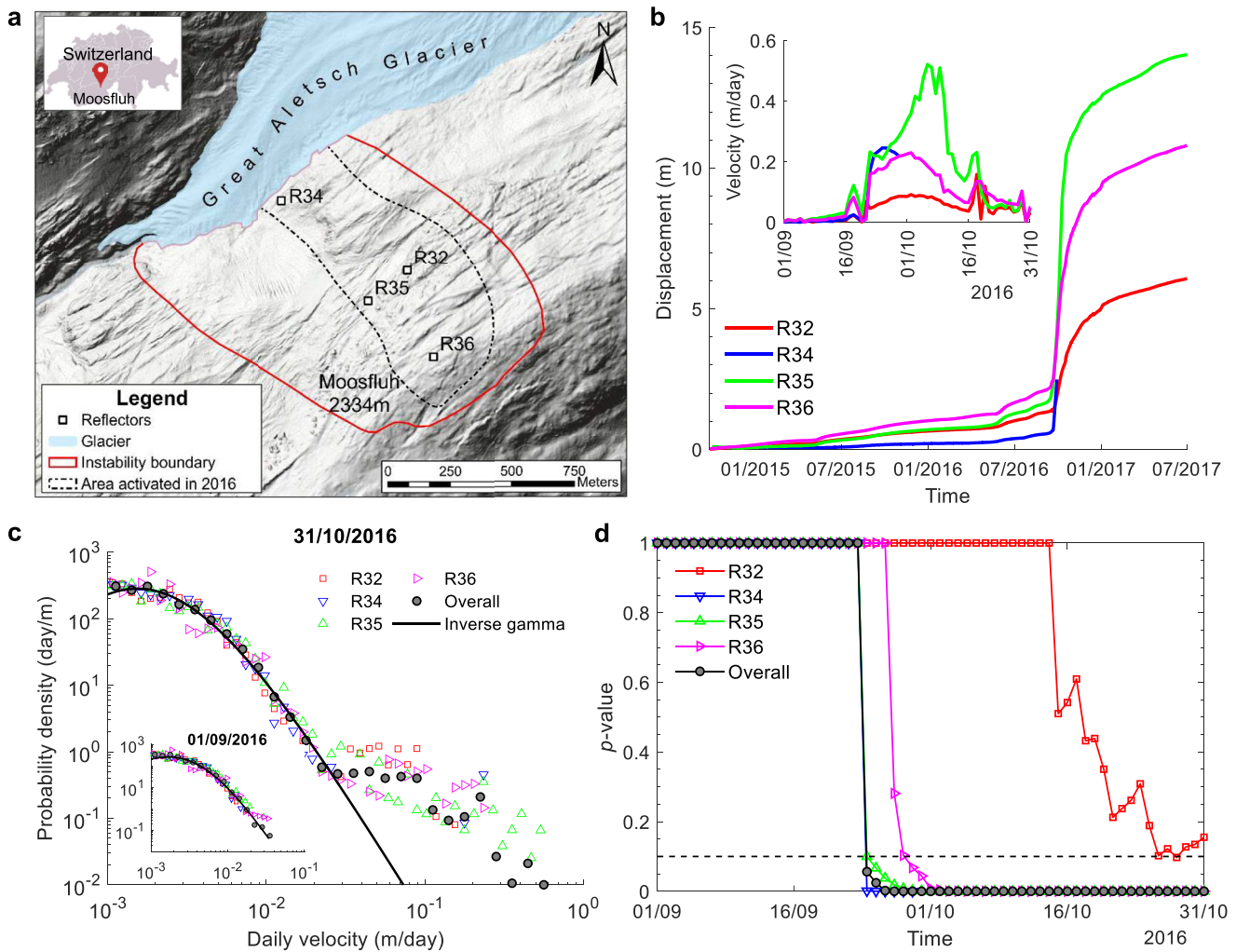
**Figure 3.** Measurement data and forecast analysis of the Veslemannen landslide, Norway. (a) Overview of the slope with the monitoring system and instability boundary indicated. (b) Slope displacement time series recorded by seven radar points (inset: daily velocities for the last 3 months before the failure). (c) Probability density functions of daily velocities available until 31 October 2018 (inset: 1 July 2018) and 5 September 2019 (inset: 1 June 2019), respectively, fitted to the inverse gamma distribution (with outliers excluded). (d) Temporal evolution of  $p$ -values (the dashed line corresponds to a significance level of 0.1).

consisting of high-grade metamorphic rocks is subject to sporadic mountain permafrost and deep seasonal frost (Magnin et al., 2019). Due to the high risk of slope collapse, a monitoring system with multiple surface and borehole-based instrumentations was initiated in 2009. Since October 2014, a ground-based interferometric synthetic-aperture radar system (with a measurement accuracy of 0.5 mm) has been installed in the valley floor as a key component of an early warning system to continuously monitor the entire unstable slope surface (Kristensen et al., 2021). Figure 3b shows the displacement time series recorded at seven radar points (see Figure 3a for their locations) over a 5-year period. The slope movements mainly occurred in summer/autumn seasons (possibly because the frozen layer thawed to permit rainwater infiltration deep into the slope) (Kristensen et al., 2021) and became increasingly large especially after 2017. On 5 September 2019, ~54,000 m<sup>3</sup> of suspended rocks collapsed catastrophically. Prior to the failure, the landslide had cumulatively displaced about 19 and 4 m in the upper and lower regions, respectively. The Norwegian Water Resources and Energy Directorate adopted a color-coded hazard assessment system (green, yellow, orange, and red) for risk management and mitigation (Krøgli et al., 2018). Hazard levels were determined according to annually adjusted velocity thresholds for different slope compartments (Kristensen et al., 2021). An evacuation of the risk zone was defined for red levels. In total, 16 red alarms were released: 1 in 2014, 1 in 2015, 1 in 2016, 2 in 2017, 6 in 2018, and 5 in 2019. Only after the last one, the slope failed, meaning that following the preceding 15 alarms, the population was evacuated but no collapse happened.

We use our method to analyze the Veslemannen landslide. Figure 3c gives the PDF of daily velocities derived from the radar displacement data. This landslide had already accommodated some large velocities exceeding 0.5 m/day in 2018 (Figure S1 in Supporting Information S1). However, these large velocities tend to follow the same inverse gamma distribution as smaller ones. In the 2019 PDF, some deviation seems to have appeared. To rigorously detect velocity outliers, we first determine the transition velocity (Figure S3 in Supporting Information S1) and then calculate the  $p$ -values under  $H_0$  for the period between July 2018 and September 2019 (Figure 3d and Movie S2; the  $p$ -values before July 2018 are all unity, meaning that  $H_0$  is not rejected). In September and October 2018, only a few radar points (e.g., P1 and P2 in the upper region) temporarily exhibited drops in  $p$ -values, while most other points had no  $p$ -value response. In November 2018, the individual  $p$ -values of three radar points (P1-P3) and the overall  $p$ -value showed clear declines, but only P1 had its  $p$ -value below 0.1 (even if the slope has entered a crisis, the phase diagram analysis indicates that it is in the slow landslide regime; see Movie S2). It is only from mid-August 2019 that most radar points started to show  $p$ -value declines, and the overall  $p$ -value stayed close to zero. On 2 September 2019, all individual  $p$ -values began to drop substantially. This observation, in conjunction with the phase diagram analysis showing that this landslide has evolved into the dragon-king regime (Figure 2 and Movie S2), suggests that the slope was likely to collapse, in agreement with the catastrophic failure 3 days after. The  $\omega$ - $m$  footprint of the Veslemannen landslide close to the boundary between the black-swan and dragon-king regimes (Figure 2) explains the difficulty in predicting the failure of this landslide. However, it is noteworthy that, compared with the 15 false alarms by the system of the Norwegian Water Resources and Energy Directorate, our method would have correctly predicted the failure while giving at most 1 false alarm (if a majority rule is adopted when analyzing the  $p$ -value responses of the seven radar points), or none (if a strict rule is adopted requiring all the seven radar points to show a  $p$ -value decline), depending on the decision process associated with the  $p$ -value evolution and phase diagram analysis.

### 2.3. Moosfluh Landslide, Switzerland

The Moosfluh landslide, also consisting of high-grade metamorphic rocks, is located at the tongue of the Great Aletsch Glacier, southern Switzerland (Figure 4a). This landslide with a volume of ~75 million m<sup>3</sup> is characterized as a deep-seated gravitational mountain slope deformation (mainly controlled by deep toppling) (Glueer et al., 2020) and responds to an ongoing glacial retreat (Kos et al., 2016). It had a displacement rate less than 5 cm/yr before 2005, but started to accelerate after 2005 with a rate exceeding 0.5 m/yr in 2012 (Glueer et al., 2020; Kos et al., 2016). In 2013, a comprehensive monitoring system including 2 robotic total stations and about 30 reflectors was implemented to continuously monitor the slope surface displacement (Glueer et al., 2019). On 21 September 2016, the central sector of this landslide accelerated dramatically, starting from the glacier-covered toe and propagating toward the crest (Glueer et al., 2019). This dramatic acceleration could be explained by the formation of three secondary compound sliding surfaces within the toppling slope (Glueer et al., 2019). Four reflectors within the activated area (Figure 4a) recorded this acceleration (Figure 4b). Reflector R34 in the toe region witnessed a displacement of ~1.6 m within a week and then was lost on 29 September, while three



**Figure 4.** Measurement data and forecast analysis of the Moosfluh landslide, Switzerland. (a) Overview of the slope with the monitoring system and instability boundary indicated. (b) Slope displacement time series recorded by four reflectors (inset: daily velocities for the period of the 2016 crisis). (c) Probability density function of daily velocities available until 31 October 2016 (inset: 1 September 2016), fitted to the inverse gamma distribution (with outliers excluded). (d) Temporal evolution of  $p$ -values (the dashed line corresponds to a significance level of 0.1).

other reflectors (R32, R35, and R36) survived. The lowermost secondary slide continued to substantially displace until today, but the middle and upper ones strongly decelerated until summer 2017 and move today again with a velocity of  $\sim 1$  mm/day. This slope so far has experienced a maximum displacement of more than 100 m without failure, which may be attributed to the retreating but persistent buttress offered by the viscous glacier at the toe of the slope (Storni et al., 2020).

We analyze the 2016 Moosfluh landslide using our method. Figure 4c gives the velocity PDF derived from the displacement data by 31 October 2016. Small velocities obey an inverse gamma distribution, while large ones tend to follow a different distribution. Those large velocities only emerged in September and October 2016 (Figure 4c inset and Figure S1 in Supporting Information S1). To rigorously detect velocity outliers, we first determine the transition velocity (Figure S4 in Supporting Information S1) and then calculate the  $p$ -values under  $H_0$  for the period between 1 September and 31 October 2016 (Figure 4d and Movie S3). The systematic  $p$ -value declines qualify the emergence of outliers, whose significance is compatible with the  $\omega$ - $m$  path staying far away from the black-swan regime in the phase diagram (Figure 2). The phase diagram analysis further indicates that this slope had a strong potential of catastrophic failure at the beginning of the crisis (from 24 September to 5 October), but later evolved back into the slow landslide regime (Figure 2 and Movie S3), consistent with the observation that this slope did not collapse and continued to deform gradually.

### 3. Discussion

For the three historical landslides, we also show the results for a significance level  $\alpha$  of 0.05 (Figure S8 in Supporting Information S1) and those using different outlier test statistics (Texts S3 and S4 and Figure S9 in Supporting Information S1); no fundamental difference is observed. Furthermore, we have tested our method using realistic synthetic slope velocity data sets (Text S6 and Figures S10–S13 in Supporting Information S1). These results indicate that our method is robust and has a strong outlier detection capability (without producing type II errors in the tested cases, i.e., empirically  $\beta = 0$ , corresponding to the maximum possible power of the method, given the possible uncertainty of the prior null distribution). Our forecast method is superior to the conventional velocity threshold approach (Crosta & Agliardi, 2002), because it uses the information embedded in the evolution of the full velocity distribution rather than pointwise velocity magnitude changes to qualify the landslide crisis. It is worth emphasizing that it is important to detect the secondary-to-tertiary creep transition first (e.g., by using our outlier detection algorithm), before conducting the phase diagram analysis that is only applicable when the slope has entered the tertiary creep stage. Our method also provides a dynamic perspective to quantitatively diagnose the landslide state and adaptively predict its possible subsequent development, such that it could robustly eliminate many of the false alarms given by the conventional approach with a fixed viewpoint. Additionally, our method provides diagnostics of the limits of predictability by characterizing the detectability of dragon-kings against black-swans.

Our research sheds light on one of the biggest questions in landslide hazard assessments (Palmer, 2017): Would an active landslide slowly move or catastrophically fail in the future? Here, we have articulated this question using the dragon-king paradigm (Sornette, 2002; Sornette & Ouillon, 2012), where a catastrophic failure is characterized by dragon-kings mechanistically arising from positive feedbacks during the landslide process. Such positive feedbacks may be related to the velocity-weakening frictional behavior (Handwerker et al., 2016; Helmstetter et al., 2004) and/or compaction-induced pore pressure buildup (Iverson, 2005) of basal shear zones as well as the stress corrosion cracking (Sammis & Sornette, 2002) of rock bridges. To apply our method to real-time forecasts, great attention is needed when estimating some of the essential parameters like  $m$ , which may be subject to uncertainties due to the intrinsic fluctuations of real data. Different techniques (Cornelius & Voight, 1995) are available to reduce the uncertainty. We do not claim that our method provides predictive tools for all landslides under all circumstances, but the results presented here suggest that adding it to existing early warning systems of landslides will bring significant value.

### Data Availability Statement

The slope displacement monitoring data of the Preonzo landslide can be found in Data Set S1 or accessible from the ETH Zürich Research Collection (<https://doi.org/10.3929/ethz-b-000600495>). The slope displacement monitoring data of the Veslemannen landslide can be downloaded from the supplementary information of the paper by Kristensen et al. (2021) published in *Landslides* (<https://doi.org/10.1007/s10346-020-01609-x>). The slope displacement monitoring data of the Moosfluh landslide are accessible from the ETH Zürich Research Collection (<https://www.research-collection.ethz.ch/handle/20.500.11850/358779>).

### References

- Booth, A. M., Lamb, M. P., Avouac, J.-P., & Delacourt, C. (2013). Landslide velocity, thickness, and rheology from remote sensing: La Clapière landslide, France. *Geophysical Research Letters*, *40*(16), 4299–4304. <https://doi.org/10.1002/grl.50828>
- Cornelius, R. R., & Voight, B. (1995). Graphical and PC-software analysis of volcano eruption precursors according to the Materials Failure Forecast Method (FFM). *Journal of Volcanology and Geothermal Research*, *64*(3–4), 295–320. [https://doi.org/10.1016/0377-0273\(94\)00078-U](https://doi.org/10.1016/0377-0273(94)00078-U)
- Courtney, T. H. (2000). *Mechanical behavior of materials* (2nd ed.). Waveland Press, Inc.
- Crosta, G. B., & Agliardi, F. (2002). How to obtain alert velocity thresholds for large rockslides. *Physics and Chemistry of the Earth, Parts A/B/C*, *27*(36), 1557–1565. [https://doi.org/10.1016/S1474-7065\(02\)00177-8](https://doi.org/10.1016/S1474-7065(02)00177-8)
- Crosta, G. B., Agliardi, F., Rivolta, C., Alberti, S., & Dei Cas, L. (2017). Long-term evolution and early warning strategies for complex rockslides by real-time monitoring. *Landslides*, *14*(5), 1615–1632. <https://doi.org/10.1007/s10346-017-0817-8>
- Crosta, G. B., di Prisco, C., Frattini, P., Frigerio, G., Castellanza, R., & Agliardi, F. (2014). Chasing a complete understanding of the triggering mechanisms of a large rapidly evolving rockslide. *Landslides*, *11*(5), 747–764. <https://doi.org/10.1007/s10346-013-0433-1>
- 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>
- Glastonbury, J., & Fell, R. (2008). Geotechnical characteristics of large slow, very slow, and extremely slow landslides. *Canadian Geotechnical Journal*, *45*(7), 984–1005. <https://doi.org/10.1139/T08-021>
- Glastonbury, J., & Fell, R. (2010). Geotechnical characteristics of large rapid rock slides. *Canadian Geotechnical Journal*, *47*(1), 116–132. <https://doi.org/10.1139/t09-080>

### Acknowledgments

Q. L. is grateful for the support by the Swiss National Science Foundation (Grant 189882) and the National Natural Science Foundation of China (Grant 41961134032). D. S. acknowledges the support by the National Natural Science Foundation of China (Grant U2039202) and by the Shenzhen Science and Technology Innovation Commission (Grant GJHZ20210705141805017). S.L. thanks for the support by the Swiss National Science Foundation (Grant 172492). We acknowledge data and discussions with many colleagues from the Norwegian Water Resources and Energy Directorate, especially L. Kristensen and L.H. Blikra.



- Glueer, F., Loew, S., & Manconi, A. (2020). Paraglacial history and structure of the Moosfluh landslide (1850–2016), Switzerland. *Geomorphology*, 355, 106677. <https://doi.org/10.1016/j.geomorph.2019.02.021>
- Glueer, F., Loew, S., Manconi, A., & Aaron, J. (2019). From toppling to sliding: Progressive evolution of the Moosfluh landslide, Switzerland. *Journal of Geophysical Research: Earth Surface*, 124(12), 2899–2919. <https://doi.org/10.1029/2019JF005019>
- Gschwind, S., Loew, S., & Wolter, A. (2019). Multi-stage structural and kinematic analysis of a retrogressive rock slope instability complex (Preonzo, Switzerland). *Engineering Geology*, 252, 27–42. <https://doi.org/10.1016/j.enggeo.2019.02.018>
- Handwerger, A. L., Rempel, A. W., Skarbek, R. M., Roering, J. J., & Hilley, G. E. (2016). Rate-weakening friction characterizes both slow sliding and catastrophic failure of landslides. *Proceedings of the National Academy of Sciences of the United States of America*, 113(37), 10281–10286. <https://doi.org/10.1073/pnas.1607009113>
- Helmstetter, A., Sornette, D., Grasso, J.-R., Andersen, J. V., Gluzman, S., & Pisarenko, V. (2004). Slider block friction model for landslides: Application to Vaiont and La Clapière landslides. *Journal of Geophysical Research*, 109(B2), 1–15. <https://doi.org/10.1029/2002jb002160>
- Hu, X., Bürgmann, R., Schulz, W. H., & Fielding, E. J. (2020). Four-dimensional surface motions of the Slumgullion landslide and quantification of hydrometeorological forcing. *Nature Communications*, 11(1), 1–9. <https://doi.org/10.1038/s41467-020-16617-7>
- IUGS. (1995). A suggested method for describing the rate of movement of a landslide. *Bulletin of the International Association of Engineering Geology*, 52(1), 75–78. <https://doi.org/10.1007/BF02602683>
- Iverson, R. M. (2005). Regulation of landslide motion by dilatancy and pore pressure feedback. *Journal of Geophysical Research*, 110(F2), F02015. <https://doi.org/10.1029/2004JF000268>
- Kirschbaum, D., Kapnick, S. B., Stanley, T., & Pascale, S. (2020). Changes in extreme precipitation and landslides over High Mountain Asia. *Geophysical Research Letters*, 47(4), 1–9. <https://doi.org/10.1029/2019GL085347>
- Kos, A., Amann, F., Strozzi, T., Delaloye, R., Ruetz, J., & Springman, S. (2016). Contemporary glacier retreat triggers a rapid landslide response, Great Aletsch Glacier, Switzerland. *Geophysical Research Letters*, 43(24), 12466–12474. <https://doi.org/10.1002/2016GL071708>
- Kristensen, L., Czekirka, J., Penna, I., Eitzelmüller, B., Nicolet, P., Pullarello, J. S., et al. (2021). Movements, failure and climatic control of the Veslemannen rockslide, Western Norway. *Landslides*, 18(6), 1963–1980. <https://doi.org/10.1007/s10346-020-01609-x>
- Krøgli, I. K., Devoli, G., Colleuille, H., Boje, S., Sund, M., & Engen, I. K. (2018). The Norwegian forecasting and warning service for rainfall- and snowmelt-induced landslides. *Natural Hazards and Earth System Sciences*, 18(5), 1427–1450. <https://doi.org/10.5194/nhess-18-1427-2018>
- Lacroix, P., Belart, J. M. C., Berthier, E., Sæmundsson, Þ., & Jónsdóttir, K. (2022). Mechanisms of landslide destabilization induced by glacier-retreat on Tungnavíslarjökull area, Iceland. *Geophysical Research Letters*, 49(14), 1–11. <https://doi.org/10.1029/2022GL098302>
- Lacroix, P., Handwerger, A. L., & Bièvre, G. (2020). Life and death of slow-moving landslides. *Nature Reviews Earth & Environment*, 1(8), 404–419. <https://doi.org/10.1038/s43017-020-0072-8>
- Loew, S., Gschwind, S., Gischig, V., Keller-Signer, A., & Valenti, G. (2017). Monitoring and early warning of the 2012 Preonzo catastrophic rock slope failure. *Landslides*, 14(1), 141–154. <https://doi.org/10.1007/s10346-016-0701-y>
- Magnin, F., Eitzelmüller, B., Westermann, S., Isaksen, K., Hilger, P., & Hermanns, R. L. (2019). Permafrost distribution in steep rock slopes in Norway: Measurements, statistical modelling and implications for geomorphological processes. *Earth Surface Dynamics*, 7(4), 1019–1040. <https://doi.org/10.5194/esurf-7-1019-2019>
- Main, I. G. (1999). Applicability of time-to-failure analysis to accelerated strain before earthquakes and volcanic eruptions. *Geophysical Journal International*, 139(3), F1–F6. <https://doi.org/10.1046/j.1365-246x.1999.00004.x>
- Main, I. G. (2000). A damage mechanics model for power-law creep and earthquake aftershock and foreshock sequences. *Geophysical Journal International*, 142(1), 151–161. <https://doi.org/10.1046/j.1365-246x.2000.00136.x>
- Manconi, A., Picozzi, M., Coviello, V., De Santis, F., & Elia, L. (2016). Real-time detection, location, and characterization of rockslides using broadband regional seismic networks. *Geophysical Research Letters*, 43(13), 6960–6967. <https://doi.org/10.1002/2016GL069572>
- Palmer, J. (2017). Creeping Earth could hold secret to deadly landslides. *Nature*, 548(7668), 384–386. <https://doi.org/10.1038/548384a>
- Patton, A. I., Rathburn, S. L., Capps, D. M., McGrath, D., & Brown, R. A. (2021). Ongoing landslide deformation in thawing permafrost. *Geophysical Research Letters*, 48(16), 1–11. <https://doi.org/10.1029/2021GL092959>
- Sammis, C. G., & Sornette, D. (2002). Positive feedback, memory, and the predictability of earthquakes. *Proceedings of the National Academy of Sciences of the United States of America*, 99(suppl\_1), 2501–2508. <https://doi.org/10.1073/pnas.012580999>
- Sornette, D. (2002). Predictability of catastrophic events: Material rupture, earthquakes, turbulence, financial crashes, and human birth. *Proceedings of the National Academy of Sciences of the United States of America*, 99(suppl\_1), 2522–2529. <https://doi.org/10.1073/pnas.022581999>
- Sornette, D., & Ouillon, G. (2012). Dragon-kings: Mechanisms, statistical methods and empirical evidence. *The European Physical Journal - Special Topics*, 205(1), 1–26. <https://doi.org/10.1140/epjst/e2012-01559-5>
- Storni, E., Hugentobler, M., Manconi, A., & Loew, S. (2020). Monitoring and analysis of active rockslide-glacier interactions (Moosfluh, Switzerland). *Geomorphology*, 371, 107414. <https://doi.org/10.1016/j.geomorph.2020.107414>
- Taleb, N. N. (2010). *The black swan: The impact of the highly improbable* (2nd ed.). Penguin.

## References From the Supporting Information

- Barndorff-Nielsen, O. E., & Cox, D. R. (1994). *Inference and asymptotics* (1st ed.). Chapman & Hall/CRC.
- Bell, A. F. (2018). Predictability of landslide timing from quasi-periodic precursory earthquakes. *Geophysical Research Letters*, 45(4), 1860–1869. <https://doi.org/10.1002/2017GL076730>
- Bell, A. F., Naylor, M., Heap, M. J., & Main, I. G. (2011). Forecasting volcanic eruptions and other material failure phenomena: An evaluation of the failure forecast method. *Geophysical Research Letters*, 38(15), 1–5. <https://doi.org/10.1029/2011GL048155>
- Chikkagoudar, M. S., & Kunchur, S. H. (1983). Distributions of test statistics for multiple outliers in exponential samples. *Communications in Statistics - Theory and Methods*, 12(18), 2127–2142. <https://doi.org/10.1080/03610928308828596>
- Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4), 661–703. <https://doi.org/10.1137/070710111>
- Forbes, C., Evans, M., Hastings, N., & Peacock, B. (2011). *Statistical distributions* (4th ed.). John Wiley & Sons.
- Hawkins, D. (1980). *Identification of outliers*. Chapman & Hall.
- Kilburn, C. R. J., & Voight, B. (1998). Slow rock fracture as eruption precursor at Soufriere Hills Volcano, Montserrat. *Geophysical Research Letters*, 25(19), 3665–3668. <https://doi.org/10.1029/98GL01609>
- Kimber, A. C. (1982). Tests for many outliers in an exponential sample. *Applied Statistics*, 31(3), 263. <https://doi.org/10.2307/2348000>
- Lin, C., & Balakrishnan, N. (2014). Tests for multiple outliers in an exponential sample. *Communications in Statistics - Simulation and Computation*, 43(4), 706–722. <https://doi.org/10.1080/03610918.2012.714030>

- Malamud, B. D., Turcotte, D. L., Guzzetti, F., & Reichenbach, P. (2004). Landslides, earthquakes, and erosion. *Earth and Planetary Science Letters*, 229(1–2), 45–59. <https://doi.org/10.1016/j.epsl.2004.10.018>
- Malevergne, Y., & Sornette, D. (2006). *Extreme financial risks: From dependence to risk management*. Springer-Verlag.
- Süli, E., & Mayers, D. F. (2003). *An introduction to numerical analysis*. Cambridge University Press.
- Voight, B. (1989). A relation to describe rate-dependent material failure. *Science*, 243(4888), 200–203. <https://doi.org/10.1126/science.243.4888.200>
- Wheatley, S., & Sornette, D. (2015). Multiple outlier detection in samples with exponential and Pareto tails: Redeeming the inward approach and detecting Dragon Kings. *Swiss Finance Institute Research Paper Series*, (pp. 15–28). Retrieved from <http://www.ssrn.com/abstract=2645709>
- Wheatley, S., Sovacool, B., & Sornette, D. (2017). Of disasters and dragon kings: A statistical analysis of nuclear power incidents and accidents. *Risk Analysis*, 37(1), 99–115. <https://doi.org/10.1111/risa.12587>

**Real-Time Forecast of Catastrophic Landslides via Dragon-King Detection**

Qinghua Lei<sup>1,2</sup>, Didier Sornette<sup>3</sup>, Haonan Yang<sup>1</sup>, Simon Loew<sup>1</sup>

<sup>1</sup>Department of Earth Sciences, ETH Zürich, Zürich, Switzerland

<sup>2</sup>Department of Earth Sciences, Uppsala University, Uppsala, Sweden

<sup>3</sup>Institute of Risk Analysis, Prediction and Management, Academy for Advanced Interdisciplinary Studies, Southern University of Science and Technology, Shenzhen, China

**Contents of this file**

Text S1 to S6  
Figures S1 to S13

**Additional Supporting Information (Files uploaded separately)**

Captions for Dataset S1  
Captions for Movies S1 to S3

**Introduction**

This document provides supporting information to complement the theory, analysis, results, and discussions in the main Letter. Text S1 presents the mathematical formulation of the inverse gamma distribution and its parameter estimation based on the maximum likelihood estimation method. Text S2-S4 gives the details of our outlier detection procedures, including the determination of the transition velocity for identifying outlier candidates, an outward sequential test for excluding spurious candidates, and a block test of the significance of remaining multiple outliers. Text S5 describes the derivation of the phase diagram characterizing landslide evolution. Text S6 gives an examination of our forecast method based on synthetic slope velocity datasets. Figure S1 gives the time series of daily slope velocities for the three historical landslides (Preonzo, Veslemannen, and Moosfluh). Figures S2-S4 show the probability density functions of daily velocities of the three historical landslides on different selected dates as well as the results of transition velocity determination. Figure S5 illustrates the determination of parameter  $m$  for the three historical landslides via least square fitting on the bi-logarithmic plot of slope acceleration versus slope velocity. Figures S6 and S7 show the temporal evolution of parameters  $m$  and  $\omega$ , respectively, for the three historical landslides. Figure S8 shows the temporal evolutions of  $p$ -values calculated for the three historical landslides with a predefined significance level  $\alpha$  of 0.05, while Figure S9 gives those calculated using the classical test statistics. Figures S10-S13 shows the results of examining our forecast method based on synthetic slope velocity datasets. Dataset S1 gives the slope displacement monitoring data of the Preonzo landslide. Movies S1-S3 show the pseudo-prospective forecast analyses of the Preonzo, Veslemannen, and Moosfluh landslides.

### Text S1. Inverse gamma distribution and its parameter estimation.

The probability density function of the three-parameter inverse gamma distribution (Forbes et al., 2011; Malamud et al., 2004) is written as:

$$p(v; a, b, c) = \frac{a^b}{\Gamma(b)} \left( \frac{1}{v-c} \right)^{b+1} \exp\left( -\frac{a}{v-c} \right), \quad (\text{S1})$$

where  $v$  is the slope velocity,  $a$  is a scale parameter,  $b$  is a shape parameter,  $c$  is a constant, and  $\Gamma(\cdot)$  is the gamma function. The parameters need to meet the conditions of  $a > 0$ ,  $b > 0$ , and  $c < v$ . The inverse gamma distribution has an exponential rollover with an essential singularity for small  $v$  values and a power law decay (with an exponent of  $b+1$ ) for medium and large  $v$  values, which may be related to the so-called long-term background and short-term intermittent slope creeps (Crosta & Agliardi, 2002), respectively. The rollover may also be attributed to the resolution limit of in-situ measurements that cannot completely detect all the small velocities. In the Letter, we use the inverse gamma distribution as a parsimonious parameterization that is convenient to capture the coexistence of a power law (for medium/large velocities) and a rollover transition (for small velocities) in the probability density distribution of slope velocities during the pre-crisis stage (i.e., the slope secondary creep stage). Some other distribution functions with a similar shape may also be used to fit the data, for which our dragon-king detection algorithm is considered to be equally applicable.

For  $N$  given measurements of the observable,  $\mathbf{V} = \{v_1, v_2, \dots, v_N\}$ , we may rank them in an ascending order as  $\mathbf{V} = \{v_{(1)} \leq v_{(2)} \leq \dots \leq v_{(N)}\}$ . If we want to estimate  $a$ ,  $b$ , and  $c$  from a subset  $\mathbf{v} = \{v_{(1)} \leq v_{(2)} \leq \dots \leq v_{(n)}\} \subseteq \mathbf{V}$ ,  $n \leq N$ , we construct the likelihood function as:

$$L(a, b, c; \mathbf{v}) = \prod_{i=1}^n \frac{a^b}{\Gamma(b)} \left( \frac{1}{v_i - c} \right)^{b+1} \exp\left( -\frac{a}{v_i - c} \right), \quad (\text{S2})$$

and further write the log-likelihood function as:

$$l(a, b, c; \mathbf{v}) = \ln L(a, b, c; \mathbf{v}) = n b \ln a - n \ln \Gamma(b) - (b+1) \sum_{i=1}^n \ln(v_i - c) - a \sum_{i=1}^n \frac{1}{v_i - c}, \quad (\text{S3})$$

The goal is to find the parameter values that maximize this log-likelihood function, i.e.:

$$[\hat{a}, \hat{b}, \hat{c}] = \arg \max_{a, b, c} l(a, b, c; \mathbf{v}), \quad (\text{S4})$$

which is however nontrivial due to its strong nonlinearity. To circumvent this issue, we use the profile likelihood method (Barndorff-Nielsen & Cox, 1994) with  $a$  and  $b$  treated as nuisance parameters, such that the profile log-likelihood is defined as:

$$l_p(c; \mathbf{v}) = \max_{a, b, c} l(a, b, c; \mathbf{v}) = l(\hat{a}, \hat{b}, c; \mathbf{v}), \text{ with } [\hat{a}, \hat{b}] = \arg \max_{a, b, c} l(a, b, c; \mathbf{v}). \quad (\text{S5})$$

We conduct a grid search over a series of fixed values of  $c$  within a plausible range, e.g.,  $[-v_{\max}, v_{\min}]$ , where  $v_{\max}$  and  $v_{\min}$  are the maximum and minimum values in  $\mathbf{v}$ , respectively. To reduce the computational cost, the search starts with a coarse grid, which is adaptively refined locally around the maximum. For each fixed value of  $c$ , we estimate  $a$  and  $b$  by constraining  $\partial l / \partial a = \partial l / \partial b = 0$ , which leads to:

$$\frac{n\hat{b}}{\hat{a}} - \sum_{i=1}^n \frac{1}{v_i - c} = 0, \quad (\text{S6a})$$

$$n \ln \hat{a} - n\psi(\hat{b}) - \sum_{i=1}^n \ln(v_i - c) = 0, \quad (\text{S6b})$$

where  $\psi(\cdot)$  is the digamma function. The equations above can be further reduced to:

$$g(\hat{b}) = \ln n + \ln \hat{a} - \ln \sum_{i=1}^n \frac{1}{v_i - c} - \psi(\hat{b}) - \frac{1}{n} \sum_{i=1}^n \ln(v_i - c) = 0, \quad (\text{S7})$$



and solved using the Newton-Raphson method (Süli & Mayers, 2003) with an initial guess for  $b$  determined by the method of moments (Forbes et al., 2011) as  $[\mu(\mathbf{v}-c)/\sigma(\mathbf{v}-c)]^2+2$ , where  $\mu(\cdot)$  and  $\sigma(\cdot)$  denote the mean and standard deviation. Finally, we obtain the estimate for  $c$  from:

$$\hat{c} = \arg \max_c l_p(c; \mathbf{v}). \quad (\text{S8})$$

### Text S2. Determination of the transition velocity.

The transition velocity  $\zeta$  is defined as the threshold above which velocities tend to deviate from the inverse gamma distribution and are considered as outlier candidates. It is used to find the maximum number of possible outliers rather than firmly identify the true outliers, which is the task of the outward sequential test described in the next section. We use the Anderson-Darling distance, which emphasizes more the tail of a distribution compared to other distance metrics (Malevergne & Sornette, 2006), to characterize the deviation of velocity data from the inverse gamma distribution:

$$A^2(\xi; \mathbf{v}) = -n - 2 \sum_{i=1}^n \{w_i \ln[F(v_{(i)})] + (1-w_i) \ln[1-F(v_{(i)})]\},$$

with  $\mathbf{v} = \{v_{(1)} \leq v_{(2)} \leq \dots \leq v_{(n)}\} \subseteq \mathbf{V}$ ,  $v_{(n)} \leq \xi$ , and  $n \leq N$ , (S9)

where  $w_i = 2i/(2n+1)$  and  $F(\cdot)$  is the cumulative distribution function of the inverse gamma distribution. The transition velocity is therefore estimated from:

$$\hat{\xi} = \arg \min_{\xi} A^2(\xi; \mathbf{v}). \quad (\text{S10})$$

In our forecast analysis, we determine the transition velocity based on the velocity data  $\mathbf{V}$  that are only available until the day of making the forecast. In Figs. S1-S3, we show the determined transition velocity for the three historical landslides (Preonzo, Veslemannen, and Moosfluh) at different selected dates, with the insets plotting  $A^2$  as a function of  $\zeta$ .

### Text S3. Outward sequential test for outlier diagnosis.

An outward sequential test is used to diagnose the outlier candidates (identified as the velocity values beyond the transition velocity) in order to exclude spurious ones. The testing procedures are as follows. For the ranked  $N$  measurements  $\mathbf{V} = \{v_{(1)} \leq v_{(2)} \leq \dots \leq v_{(N)}\}$ , we specify the number of outlier candidates  $k$  according to  $v_{(N-k)} \leq \zeta < v_{(N-k+1)}$ , where  $\zeta$  is the transition velocity (defined in the previous section). Thus, the subset of  $\mathbf{V}$  denoted as  $\mathbf{v}^* = \{v_{(N-k+1)} \leq v_{(N-k+2)} \leq \dots \leq v_{(N)}\}$  contains all the outlier candidates. We start with testing if the  $k$ th largest candidate  $v_{(N-k+1)}$  is an outlier by deleting the other  $k-1$  largest ones. If the null is rejected,  $k$  outliers are declared; otherwise, the  $(k-1)$ th largest candidate  $v_{(N-k+2)}$  is tested and so on. This ‘‘outward’’ sequential procedure of testing increasingly larger outlier candidates,  $v_{(j)}$  with  $j = N-k+1, N-k+2, \dots, N$ , stops as soon as a null is rejected. If none of the nulls are rejected, no outlier is detected. The outward sequential test is effective in minimizing the masking and swamping errors (Kimber, 1982; Lin & Balakrishnan, 2014; Wheatley & Sornette, 2015).

Our outward sequential test is based on the so-called max-robust-sum (MRS) statistic defined for the  $j$ th outlier candidate as (Wheatley & Sornette, 2015):

$$T_{j,k}^{\text{MRS}} = \frac{v_{(j)}}{\sum_{i=1}^{N-k} v_{(i)}}, \quad (\text{S11})$$

which has an improved performance (because no outlier candidate contaminates the denominator) compared to the classical max-sum (MS) statistic defined as (Kimber, 1982):

$$T_j^{\text{MS}} = \frac{v_{(j)}}{\sum_{i=1}^j v_{(i)}}. \quad (\text{S12})$$

The results based on the MRS statistic are shown in the Letter, while those based on the MS statistic are presented in Figure S8. For a prescribed significance level  $\alpha$ , we determine the corresponding critical value through Monte Carlo simulations (10,000 runs) of the null (i.e., inverse gamma distribution) with parameters estimated from the subset excluding outlier candidates, i.e.,  $\mathbf{V}-\mathbf{v}^* = \{v_{(1)} \leq v_{(2)} \leq \dots \leq v_{(N-k)}\}$ . If the test statistic derived from the actual measurement data exceeds the critical value, the null is rejected. The analysis in the Letter is for  $\alpha = 0.1$ , while an additional analysis for  $\alpha = 0.05$  is given in Figure S7.

#### **Text S4. Block test of multiple outliers.**

To quantify the significance of outliers, we further conduct a block test (Hawkins, 1980) on the  $r$  remaining outliers  $\mathbf{v}^{**} = \{v_{(N-r+1)} \leq v_{(N-r+2)} \leq \dots \leq v_{(N)}\}$ ,  $0 \leq r \leq k$  (with spurious ones excluded from the outlier candidate list) under the null hypothesis  $H_0$  that all the  $N$  measurements in  $\mathbf{V}$  are drawn from the same population. Our block test is based on the so-called sum-robust-sum test statistic (SRS) defined for the  $r$  candidate outliers as (Wheatley et al., 2017; Wheatley & Sornette, 2015):

$$T_r^{\text{SRS}} = \frac{\sum_{i=N-r+1}^N v_{(i)}}{\sum_{i=1}^{N-r} v_{(i)}}, \quad (\text{S13})$$

which has an improved performance (because the outliers do not contaminate the denominator) compared to the classical sum-sum (SS) statistic defined as (Chikkagoudar & Kunchur, 1983):

$$T_r^{\text{SS}} = \frac{\sum_{i=N-r+1}^N v_{(i)}}{\sum_{i=1}^N v_{(i)}}. \quad (\text{S14})$$

The results based on the SRS statistic are shown in the Letter, while those based on the SS statistic are given in Figure S8. To compute the  $p$ -value for the null hypothesis  $H_0$ , we first perform Monte Carlo simulation experiments to generate a large number of realizations drawn from the null (i.e., inverse gamma distribution) with parameters estimated from  $\mathbf{V}-\mathbf{v}^{**} = \{v_{(1)} \leq v_{(2)} \leq \dots \leq v_{(N-r)}\}$ . Then, the  $p$ -value is determined as the proportion of the realizations having a test statistic larger than that calculated from the actual measurement data. The number of Monte Carlo realizations is set as 10,000 in order to have the calculated  $p$ -value being accurate to at least 2 decimal digits (Clauset et al., 2009). If the calculated  $p$ -value is less than the significance level  $\alpha$ , the null is rejected; otherwise, we cannot reject the null. The analysis in the Letter is for  $\alpha = 0.1$ , while an additional analysis for  $\alpha = 0.05$  is given in Figure S7 (note that there are some subtle differences between the  $p$ -values for  $\alpha = 0.1$  and those for  $\alpha = 0.05$ , since different  $r$  numbers of outliers might be determined from the outward sequential test at the different  $\alpha$  levels, leading to different  $p$ -values calculated from the block test).

#### **Text S5. Phase diagram of landslide evolution.**

If the null is rejected in the block test (i.e., the presence of outliers is significant), we further analyze the landslide evolution and failure potential using a phase diagram constructed on the  $\omega$ - $m$  plane. The definitions and derivations of the dimensionless parameters  $\omega$  and  $m$  are elaborated below.

The slope acceleration behavior during a crisis (i.e., tertiary creep stage) can often be phenomenologically described by the following nonlinear dynamic equation (Bell, 2018; Bell et al., 2011; Crosta & Agliardi, 2002; Kilburn & Voight, 1998; Sammis & Sornette, 2002; Voight, 1989):

$$\frac{\dot{v}}{\dot{v}_0} = \left( \frac{v}{v_0} \right)^m, \quad (\text{S15})$$

where  $v$  is the slope velocity,  $\dot{v}$  is the slope acceleration,  $v_0$  and  $\dot{v}_0$  are the onset velocity and onset acceleration, respectively, at the starting time  $t_0$  of the slope acceleration crisis, and  $m$  is an exponent defining the degree of nonlinearity. The solution to the above equation is given by (Voight, 1989):

$$v = v_0 \exp \left[ \frac{\dot{v}_0}{v_0^m} (t - t_0) \right], \text{ for } m = 1, \quad (\text{S16a})$$

or

$$v = \left[ \frac{\dot{v}_0}{v_0^m} (1 - m)(t - t_0) + v_0^{1-m} \right]^{1/(1-m)}, \text{ for } m \neq 1, \quad (\text{S16b})$$

where  $t$  is the time (on a daily basis). For  $m > 1$ , a positive feedback is present, due to the fact that the growth rate of  $v$  defined as  $d(\ln v)/dt$  increases as a power law of  $v$  itself, characterizing a super-exponential dynamic with a finite-time singularity (Sammis & Sornette, 2002) and leading to a *catastrophic landslide*. For  $m < 1$  or  $m = 1$ , the growth rate of  $v$  respectively decreases or plateaus with the increase of  $v$ , such that the slope continuously deforms with no failure and is therefore classified as a *slow landslide*.

For a catastrophic landslide ( $m > 1$ ), the predictability of failure depends on whether the large velocities related to the crisis (i.e., generated by slope tertiary creeps) can be statistically discriminated from those generated before the crisis (i.e., generated by slope secondary creeps). These pre-crisis large velocities correspond to the tail of the inverse gamma distribution which may degenerate into a power law, given that  $v \geq v_0 \gg c$ :

$$p(v; a, b) = \frac{a^b}{\Gamma(b)} \left( \frac{1}{v} \right)^{b+1}. \quad (\text{S17})$$

Thus, we may estimate the number of pre-crisis large velocities as:

$$n_p = N \int_{v_0}^{\infty} p(v; a, b) dv = \frac{Na^b}{b\Gamma(b)} v_0^{-b}, \quad (\text{S18})$$

where  $N$  is the total number of observation days. On the other hand, we may estimate the number of crisis-related velocities by simply calculating the time to failure (in days):

$$n_c = t_f - t_0 = \frac{v_0}{\dot{v}_0(m-1)}, \quad (\text{S19})$$

where  $t_f$  is the time of failure (i.e., when  $v \rightarrow \infty$ , which means in practice the transition to the inertial sliding regime during which the landslide develops in matters of seconds to minutes). Theoretically speaking, the failure is potentially predictable if  $n_c$  is larger than  $n_p$  (i.e., the number of large velocities is greater than that expected from the null), corresponding to the ‘‘dragon-king’’ regime (Sornette & Ouillon, 2012). Otherwise, the failure tends to be unpredictable, corresponding to the ‘‘black-swan’’ regime (Taleb, 2010). This allows us to define the boundary between the black-swan and dragon-king regimes from  $n_p = n_c$ , which leads to:

$$m = \omega(a, b, v_0, \dot{v}_0, N) + 1, \quad (\text{S20})$$

where  $\omega = v_0^{b+1} b \Gamma(b) / (\dot{v}_0 N a^b)$  is a dimensionless parameter related to the statistical parameters of the null (i.e.,  $a$  and  $b$ ), the onset condition of the crisis (i.e.,  $v_0$  and  $\dot{v}_0$ ), and the observation period

(i.e.,  $N$ ). More intuitively,  $\omega$  is related to the heterogeneity/stochasticity of the system (the less heterogeneous/stochastic the system is, the higher  $\omega$  is), while  $m$  is related to the nonlinearity of the slope acceleration process (the more nonlinear the landslide behavior is, the higher  $m$  is). It is worth mentioning that the actual boundary between the black-swan regime and the dragon-king regime may be less sharp than predicted theoretically, because the detection of dragon-kings may require  $n_c$  to be significantly larger than  $n_p$  (in a statistical sense). It can be seen that a landslide is more prone to be in the black-swan regime if the system is more heterogeneous (e.g., the null regime accommodates large velocities with a large  $a$  and/or a small  $b$ ) or stochastic (e.g., the onset velocity  $v_0$  is small or the onset acceleration  $\dot{v}_0$  is large). These conditions would cause the dragon-kings to be hidden within the tail of the inverse gamma distribution and thus become undetectable by our  $p$ -value analysis, so that dragon-kings could grow secretly and be spuriously mistaken as black-swans.

We implement the following procedures to process the landslide datasets in order to derive their  $\omega$  and  $m$  values. We estimate  $a$  and  $b$  using the profile likelihood method (see Text S1). We estimate the onset velocity  $v_0$  as the transition velocity  $\zeta$ ; see equation (S10). We then estimate the  $m$  and  $\dot{v}_0$  values, by fitting equation (S15) to the plot of slope acceleration versus slope velocity (Figure S4). Here, only the true velocity outliers, i.e.,  $\mathbf{v}^{**}$ , are used with the acceleration data computed as the time derivative of the velocity data. In Figures S5 and S6, we show the derived  $\omega$  and  $m$  values as well as their temporal variations for the three historical landslides.

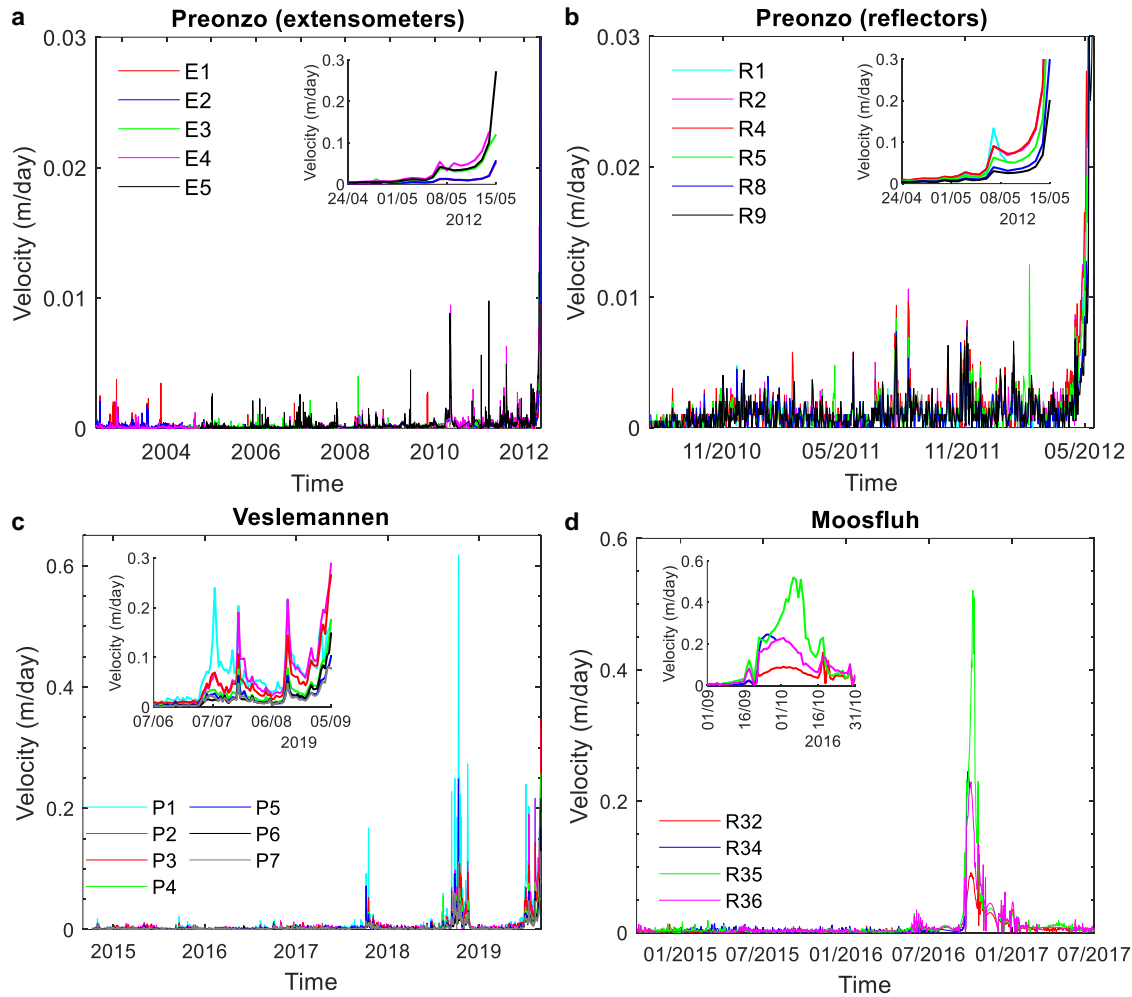
### **Text S6. Examination of the forecast method using synthetic slope velocity datasets.**

We test our forecast method using realistic synthetic slope velocity datasets generated from the probability density function of the inverse gamma distribution for the pre-crisis phase (i.e., the slope secondary creep stage), and the dynamic formulation of slope acceleration for the crisis phase (i.e., the slope tertiary creep stage).

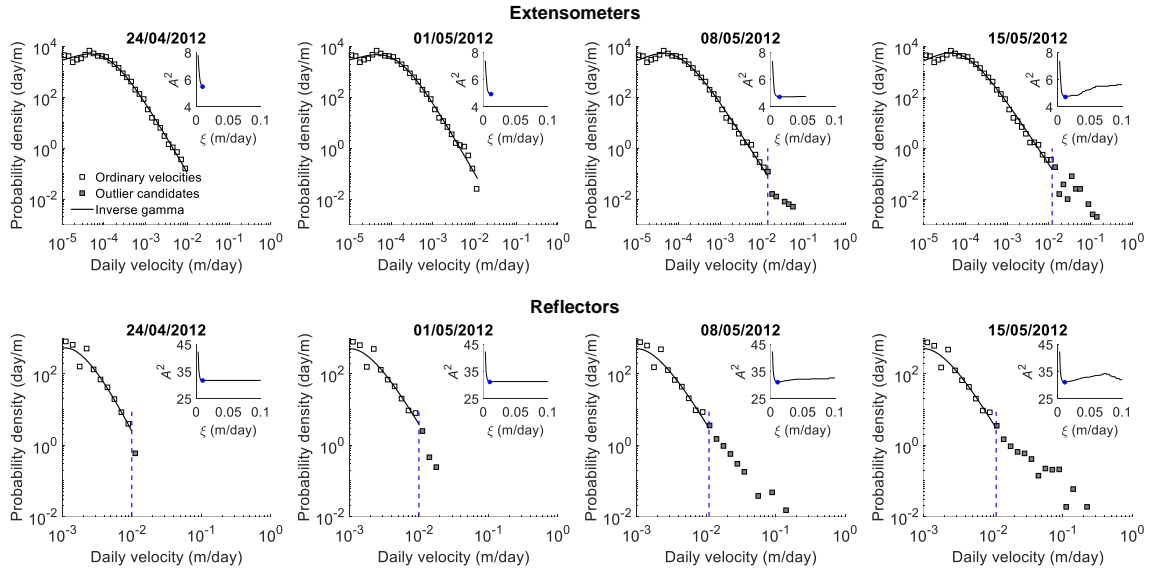
To cover typical observed values and regimes, we explore a broad parameter space of  $b = 0.2, 0.4, \dots, 3$  and  $m = 0, 0.2, \dots, 3$ , while assuming  $a = 2 \times 10^{-4}$  m/day,  $c = -3 \times 10^{-5}$  m/day,  $t_0 = 1001$  days,  $v_0 = 0.01$  m/day, and  $\dot{v}_0 = 1 \times 10^{-3}$  m<sup>2</sup>/day. We run 5 realizations (simulating the datasets of 5 measuring instruments D1-D5) with their  $m$  values assigned with a standard deviation of 0.1, reflecting the fact that different compartments of a single slope can have different  $m$  values. We use equations (S16) to generate velocities during the crisis that starts from  $t_0$  and ends 100 days after (if no failure happens) or at an earlier date (if a failure happens before reaching this maximum number of crisis days). For the pre-crisis phase before  $t_0$ , we sample  $t_0 - 1$  random velocities from the inverse gamma distribution. We set a maximum observation period of 1100 days; if a failure happens earlier, the actual observation period would be correspondingly shorter.

In Figures S9 to S11, we show three representative synthetic velocity datasets generated for different  $b$ - $m$  pairs, i.e., (i)  $b = 0.6$  and  $m = 1.6 \pm 0.1$ , (ii)  $b = 1.4$  and  $m = 1.6 \pm 0.1$ , and (iii)  $b = 1.4$  and  $m = 0.4 \pm 0.1$ , as well as the forecast analysis results using our method. It can be seen that, in general, the synthetic datasets realistically resemble the empirical ones shown in the Letter. The first synthetic case illustrates a black-swan catastrophic landslide, where many large velocities have already appeared before the crisis such that the dragon-kings are hidden and undetectable from the  $p$ -value analysis (see Figure S9). The second case represents a dragon-king catastrophic landslide and the crisis can be predicted well before the failure from the  $p$ -value analysis (see Figure S10). The last case is a slow landslide with no failure although all the  $p$ -values systematically drop below 0.1 (see Figure S11). In Figure S12, we summarize our prediction results for all the synthetic cases (240 combinations of the  $b$  and  $m$  parameters) in the  $\omega$ - $m$  phase diagram. In general, the results show a good consistency with the theoretically classified landslide regimes, except for a few cases near the boundaries, which is understandable considering the randomness in generated synthetic datasets and delicate calibration of parameters in finite noisy samples. The simulations and tests presented here using synthetic datasets generated with controlled parameters support the validity of our forecast method.

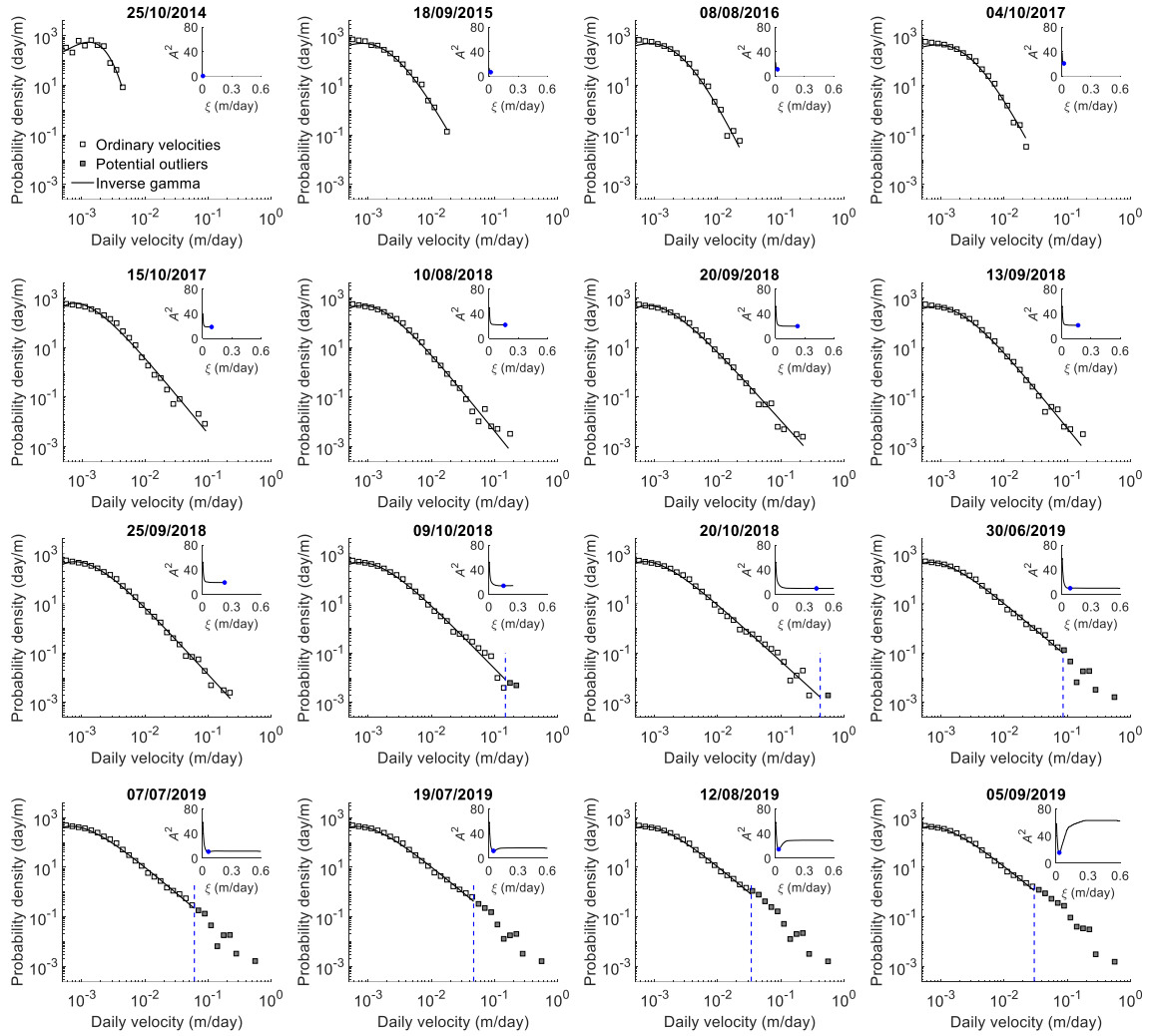




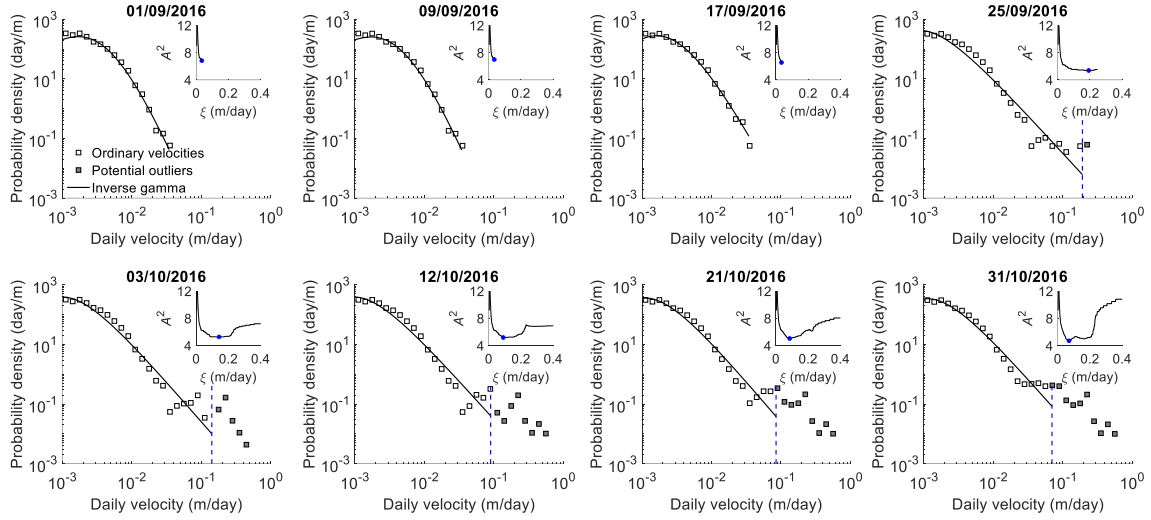
**Figure S1.** Time series of daily slope velocities for (a) the Preonzo landslide (based on extensometer measurements), (b) the Preonzo landslide (based on reflector measurements), (c) the Veslemannen landslide, and (d) the Moosfluh landslide. Insets show the velocity time series for the period of slope acceleration crisis of each landslide.



**Figure S2.** Probability density functions of daily velocities of the Preonzo landslide on different selected dates close to the 2012 failure event. For each specific date, the analysis is based on the velocity data derived from the slope displacement time series recorded until that day by the extensometers (upper panel) and reflectors (lower panel). Inset: determination of the transition velocity  $\zeta$ , beyond which velocities are considered as outlier candidates; the estimate of  $\zeta$  corresponds to the point (marked by a blue-colored circle) where the Anderson-Darling distance  $A^2$  reaches the global minimum.

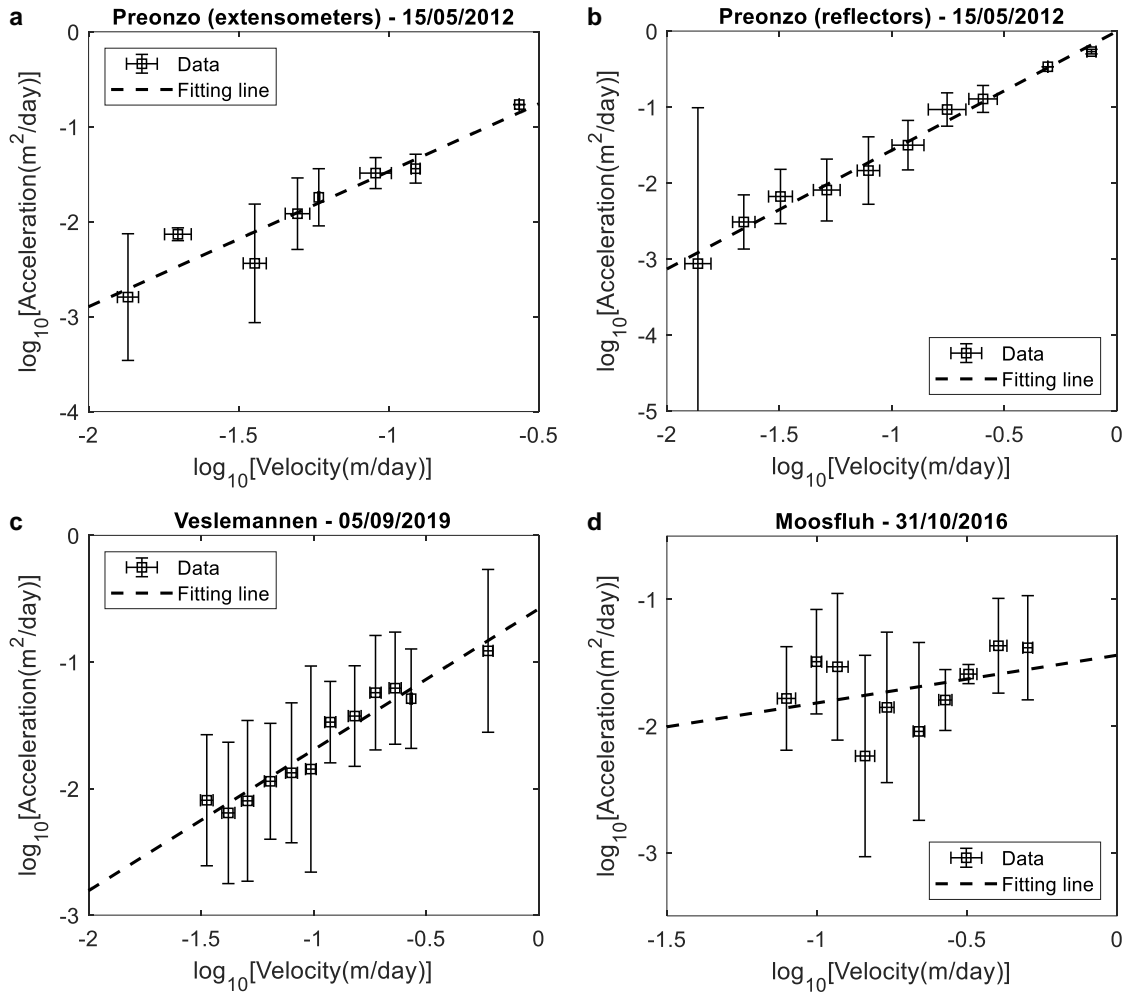


**Figure S3.** Probability density functions of daily velocities of the Veslemannen landslide on 16 different dates when red hazard alarms were previously released. For each specific date, the analysis is based on the velocity data derived from the slope displacement time series recorded until that day. Inset: determination of the transition velocity  $\zeta$ , beyond which velocities are considered as outlier candidates; the estimate of  $\zeta$  corresponds to the point (marked by a blue-colored circle) where the Anderson-Darling distance  $A^2$  reaches the minimum.

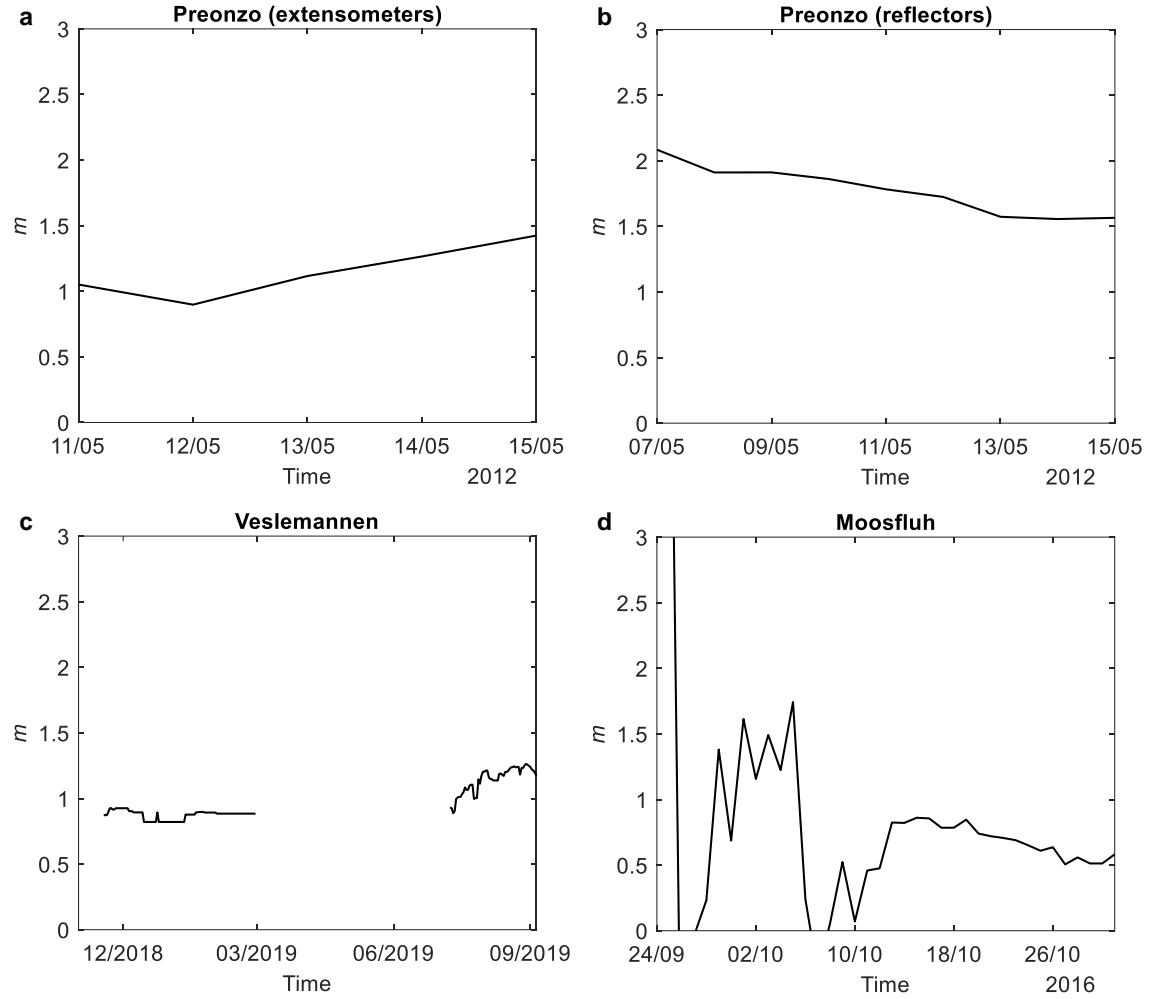


**Figure S4.** Probability density functions of daily velocities of the Moosfluh landslide on different selected dates during the 2016 crisis. For each specific date, the analysis is based on the velocity data derived from the slope displacement time series recorded until that day. Inset: determination of the transition velocity  $\zeta$ , beyond which velocities are considered as outlier candidates; the estimate of  $\zeta$  corresponds to the point (marked by a blue-colored circle) where the Anderson-Darling distance  $A^2$  reaches the global minimum.

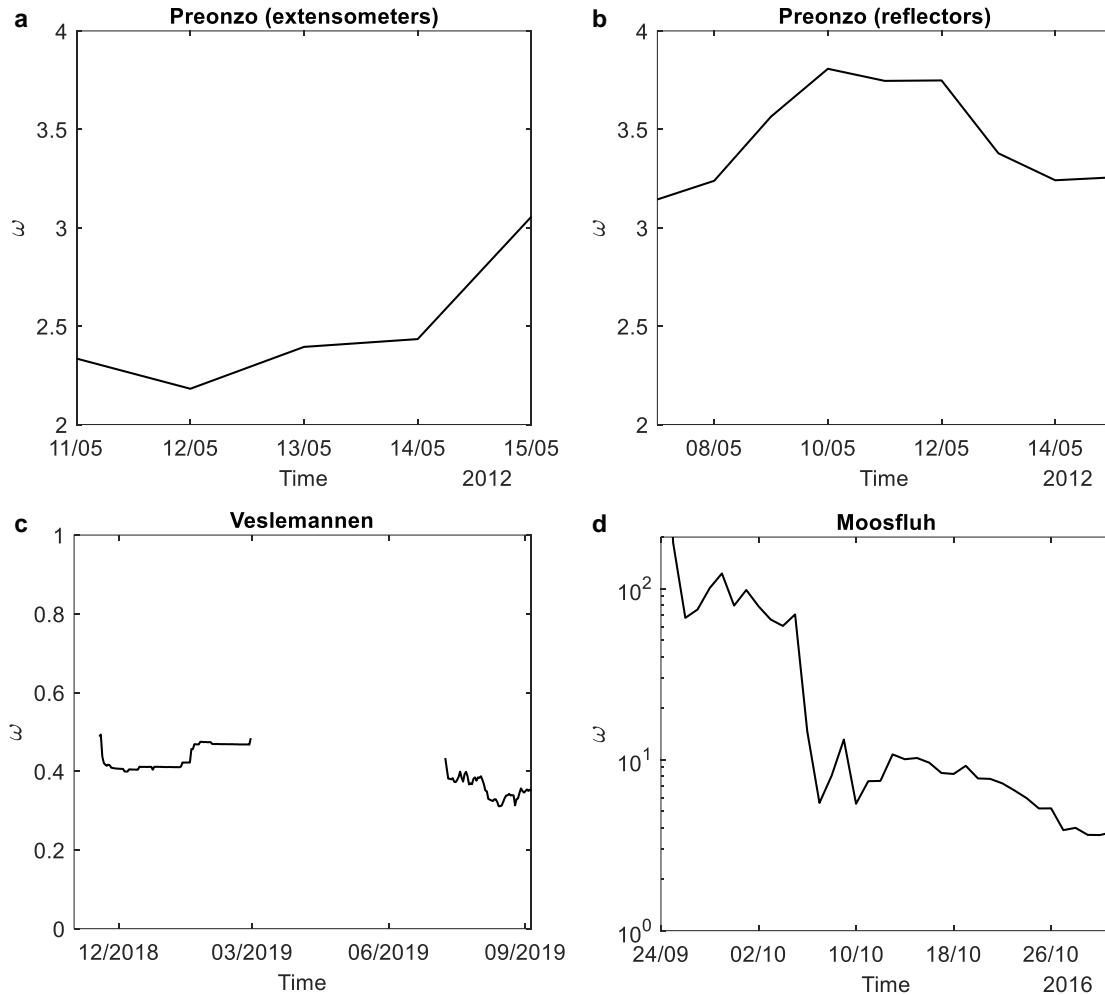




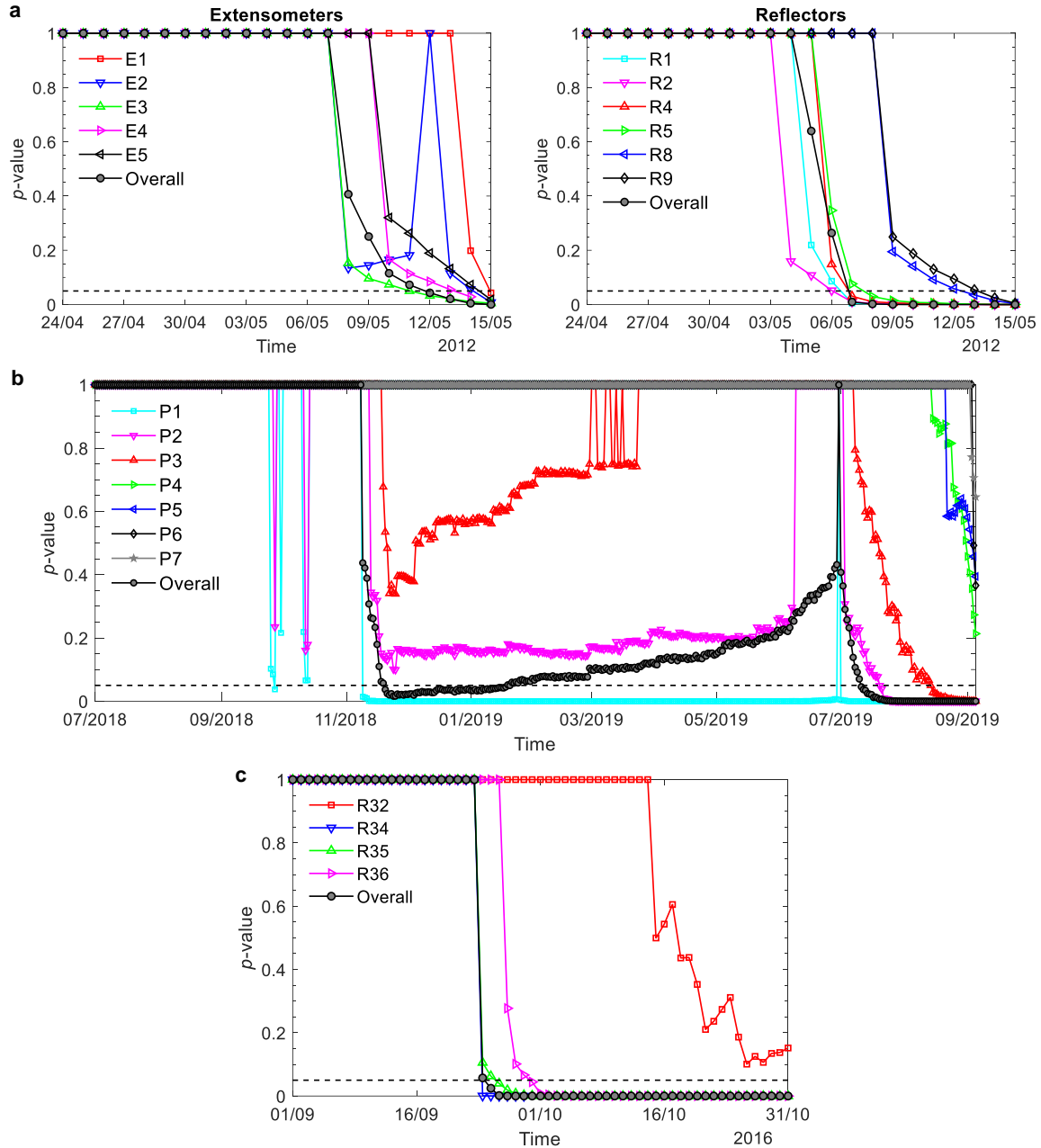
**Figure S5.** Estimation of parameter  $m$  for the three historical landslides. Derivation of the  $m$  value via a least square fitting on the bi-logarithmic plot of slope acceleration versus slope velocity based on (a) the Preonzo extensometer measurements, (b) the Preonzo reflector measurements, (c) the Veslemannen measurements, and (d) the Moosfluh measurements.



**Figure S6.** Temporal evolution of parameter  $m$  of the three historical landslides. It is estimated based on the available data until the given date (i.e., the horizontal axis) from (a) the Preonzo extensometer measurements, (b) the Preonzo reflector measurements, (c) the Veslemannen measurements, and (d) the Moosfluh measurements. Note that parameter  $m$  is only calculated for the period when the slope is subject to a crisis defined as when the overall  $p$ -value drops below the significance level  $\alpha$ .

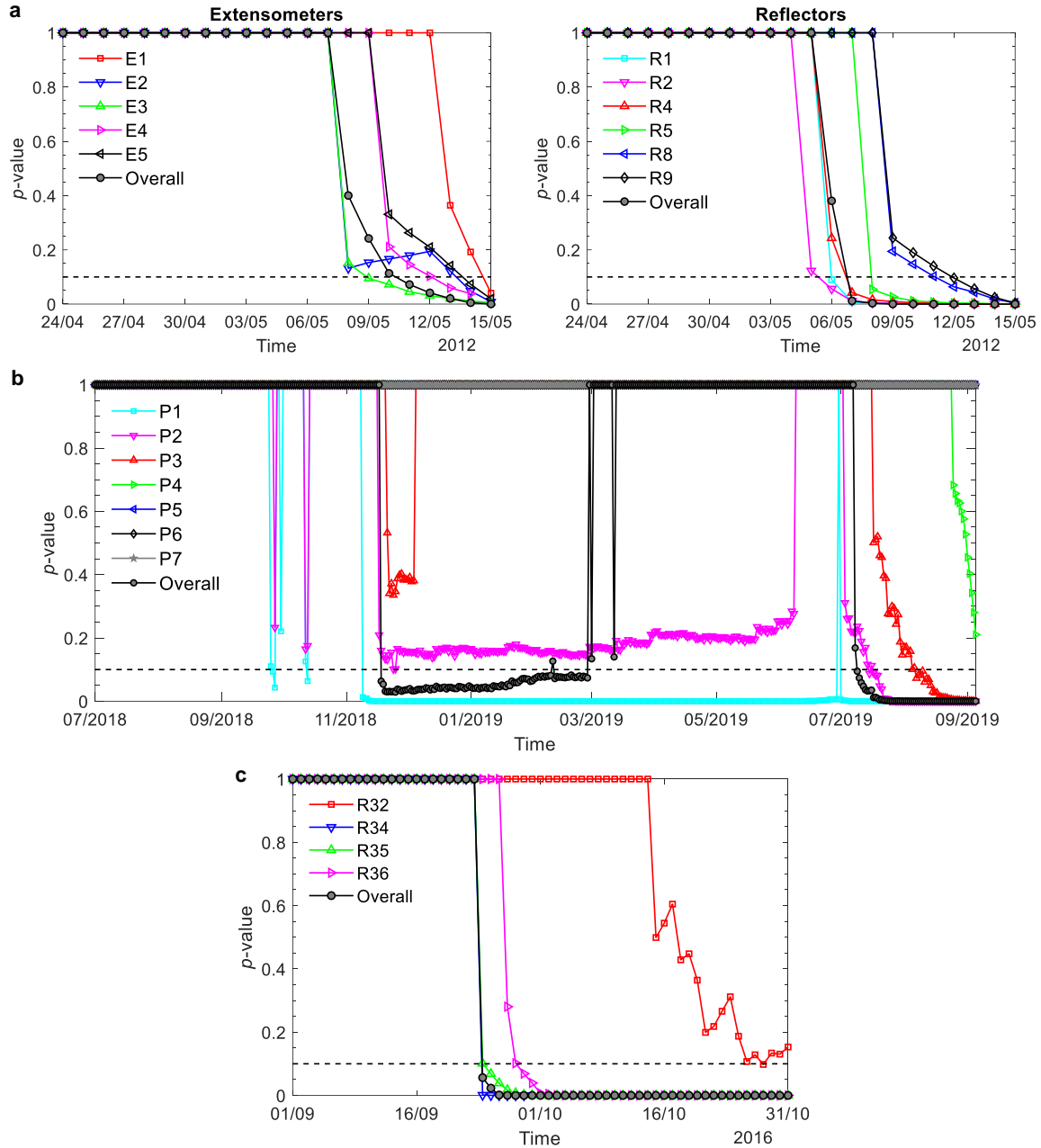


**Figure S7.** Temporal evolution of parameter  $\omega$  of the three historical landslides. It is estimated based on the available data until the given date (i.e., the horizontal axis) from (a) the Preonzo extensometer measurements, (b) the Preonzo reflector measurements, (c) the Veslemannen measurements, and (d) the Moosfluh measurements. Note that parameter  $\omega$  is only calculated for the period when the slope is subject to a crisis defined as when the overall  $p$ -value drops below the significance level  $\alpha$ .

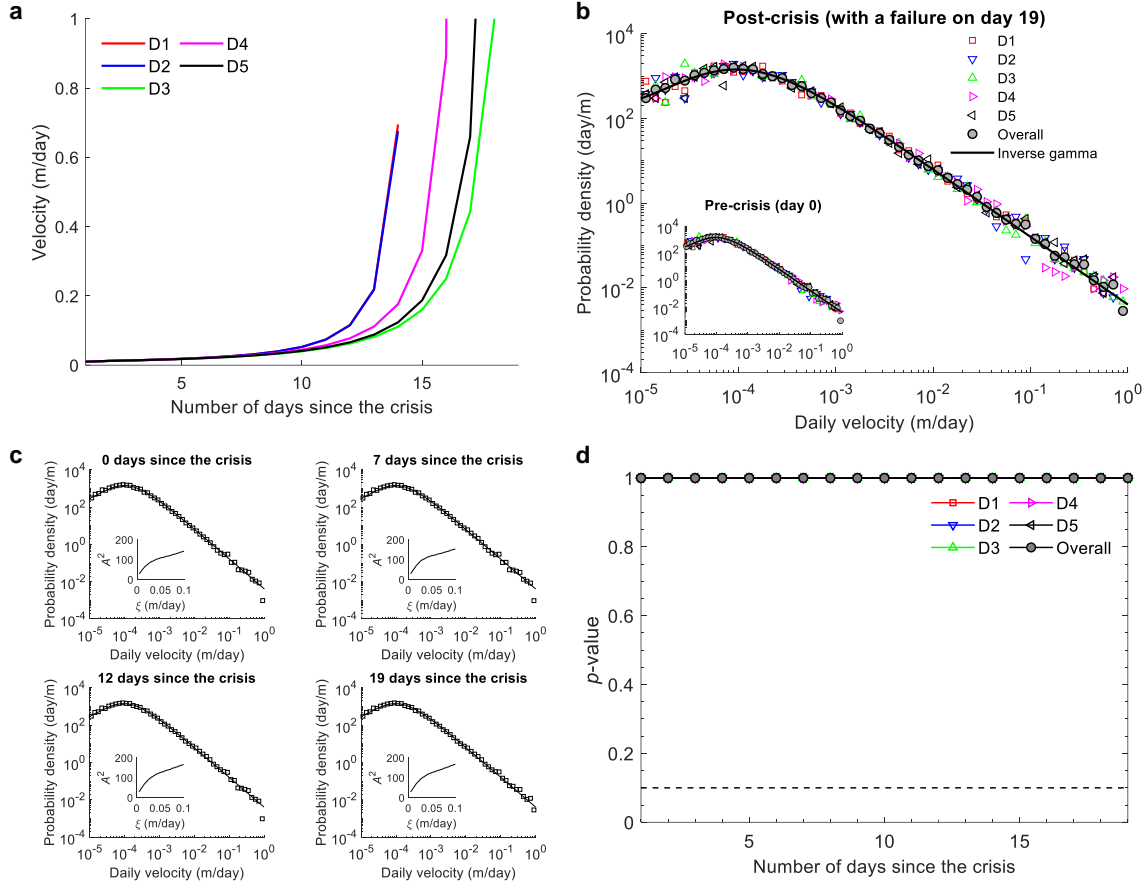


**Figure S8.** Temporal evolution of  $p$ -values calculated for the three historical landslides with a predefined significance level  $\alpha$  of 0.05 (indicated by the dashed line). (a) Preonzo landslide. (b) Veslemannen landslide. (c) Moosfluh landslide. Note: the test statistics adopted for null hypothesis testing here for  $\alpha = 0.05$  are the same as those used in the Letter for  $\alpha = 0.1$ ; there are some subtle differences between the  $p$ -values here and those presented in the Letter, since different numbers of outliers might be determined in the outward sequential test at the different significance levels, leading to different  $p$ -values calculated in the block test.

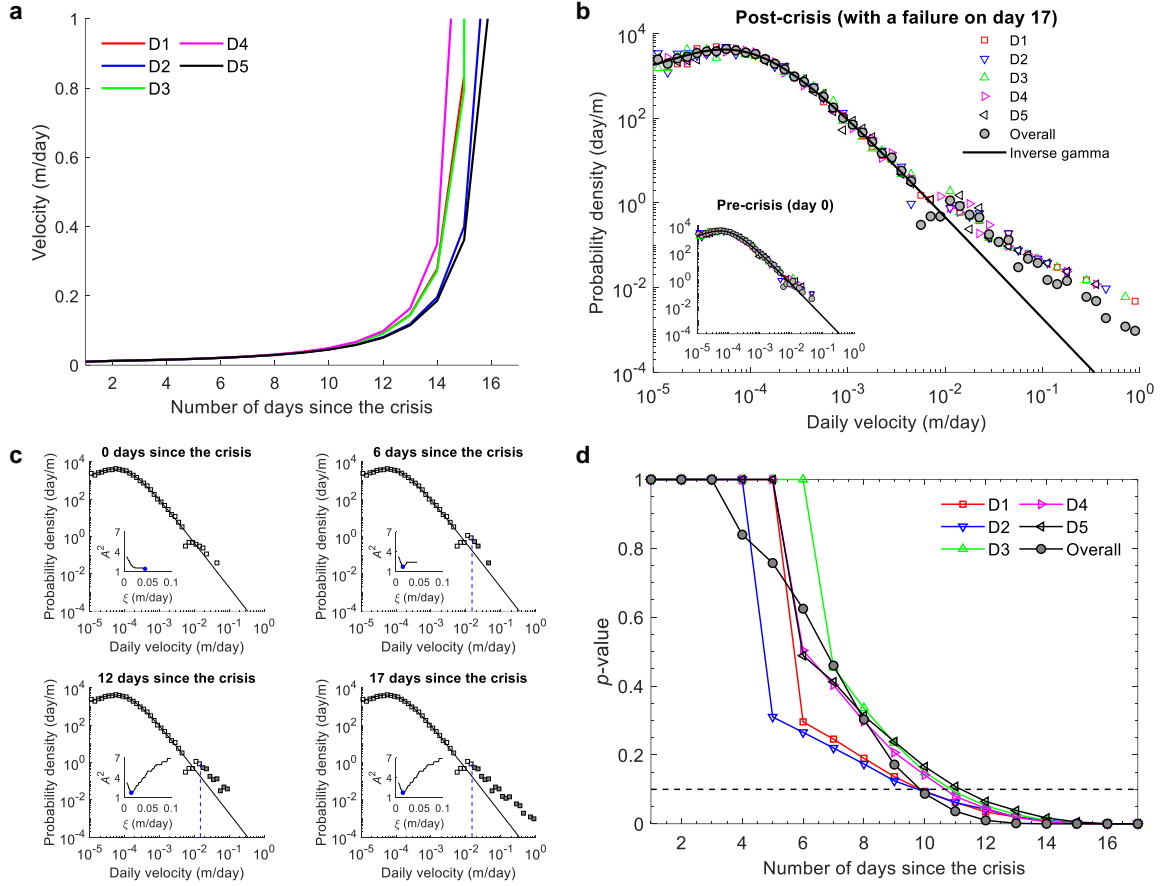




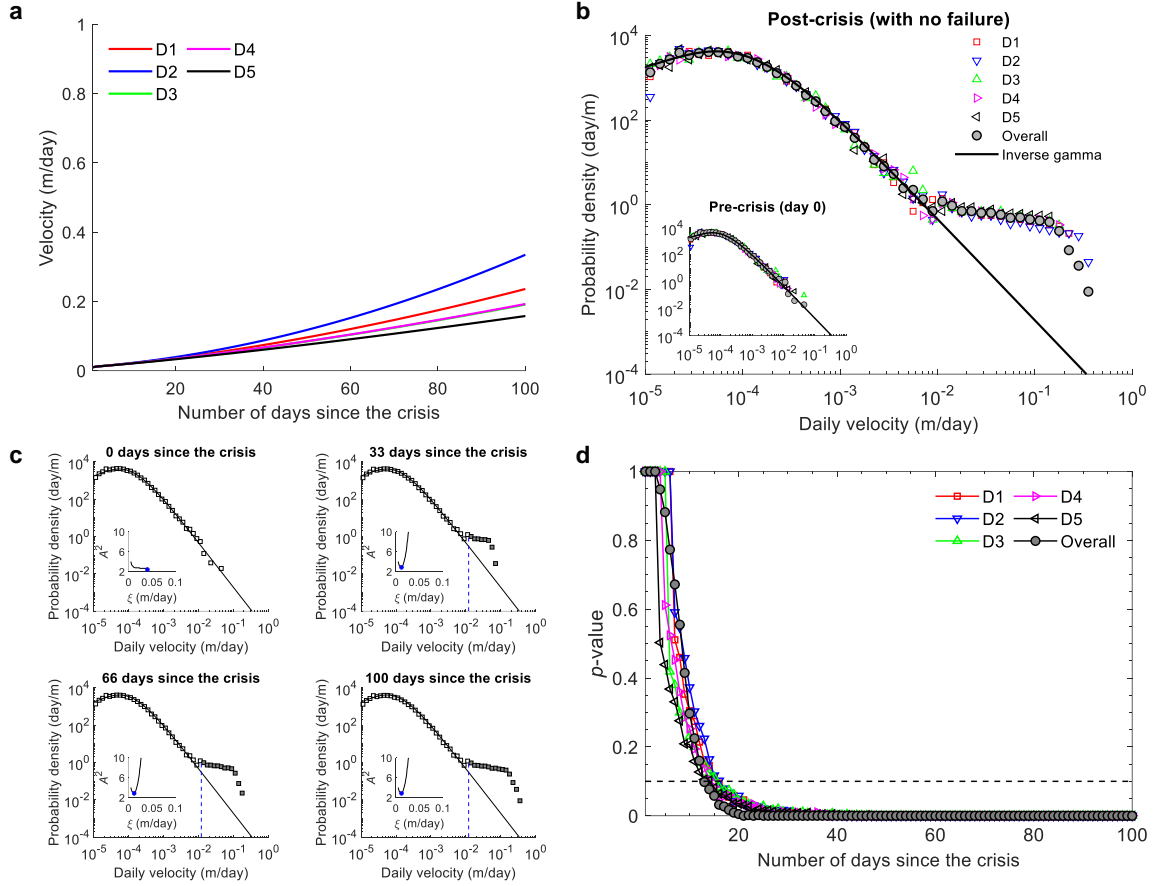
**Figure S9.** Temporal evolution of  $p$ -values calculated for the three historical landslides using the classical test statistics. (a) Preonzo landslide. (b) Veslemannen landslide. (c) Moosfluh landslide. The dashed line corresponds to a significance level  $\alpha$  of 0.1. The results are similar to those in the Letter, although the forecast performance seems to be less robust, especially for the Veslemannen landslide.



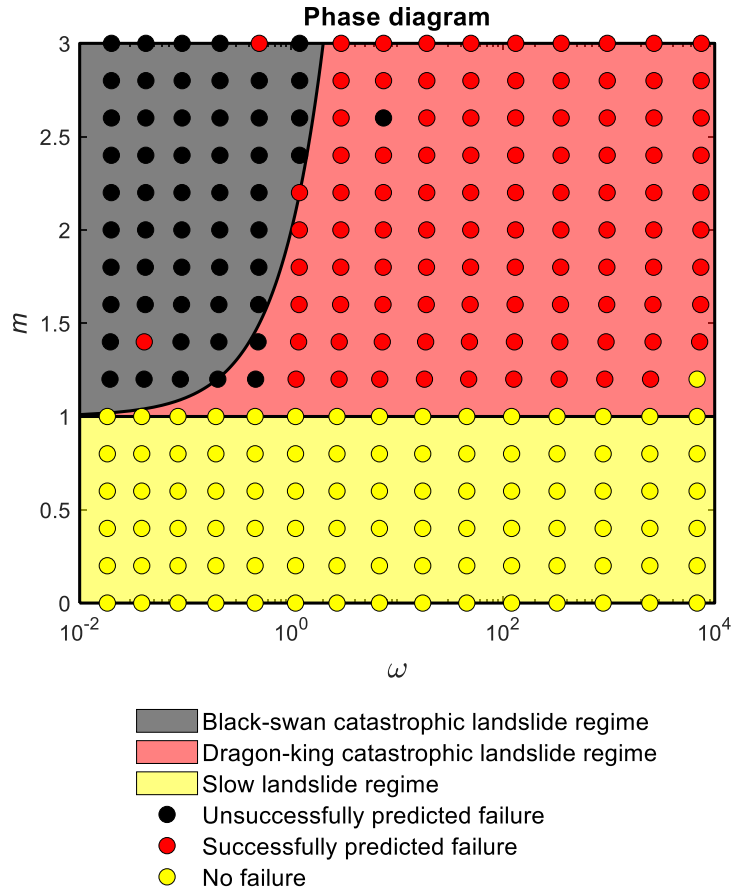
**Figure S10.** Examination of our forecast method using synthetic velocity datasets generated from an inverse gamma distribution with shape parameter  $b = 0.6$  in conjunction with the dynamic formulation, i.e., equation (S16), with parameter  $m = 1.6 \pm 0.1$ . (a) Generated synthetic datasets of daily slope velocities during the slope acceleration crisis. (b) Probability density functions of daily velocities available after the crisis with a failure (inset: before the crisis). (c) Determination of the transition velocity  $\zeta$  for different dates during the crisis (the estimate of  $\zeta$  corresponds to the point marked by a blue-colored circle in the inset; if no point is marked, the estimate of  $\zeta$  is infinity, which is the case here). (d) Temporal evolution of  $p$ -values (the dashed line corresponds to a significance level  $\alpha$  of 0.1).



**Figure S11.** Examination of our forecast method using synthetic velocity datasets generated from an inverse gamma distribution with shape parameter  $b = 1.4$  in conjunction with the dynamic formulation, i.e., equation (S16), with parameter  $m = 1.6 \pm 0.1$ . (a) Generated synthetic datasets of slope daily velocities during the slope acceleration crisis. (b) Probability density functions of daily velocities available after the crisis with a failure (inset: before the crisis). (c) Determination of the transition velocity  $\zeta$  for different dates during the crisis (the estimate of  $\zeta$  corresponds to the point marked by a blue-colored circle in the inset; if no point is marked, the estimate of  $\zeta$  is infinity). (d) Temporal evolution of  $p$ -values (the dashed line corresponds to a significance level  $\alpha$  of 0.1).



**Figure S12.** Examination of our forecast method using synthetic velocity datasets generated from an inverse gamma distribution with shape parameter  $b = 1.4$  in conjunction with the dynamic formulation, i.e., equation (S16), with parameter  $m = 0.4 \pm 0.1$ . (a) Generated synthetic datasets of slope daily velocities during the slope acceleration crisis. (b) Probability density functions of daily velocities available after the crisis with no failure (inset: before the crisis). (c) Determination of the transition velocity  $\zeta$  for different dates during the crisis (the estimate of  $\zeta$  corresponds to the point marked by a blue-colored circle in the inset; if no point is marked, the estimate of  $\zeta$  is infinity). (d) Temporal evolution of  $p$ -values (the dashed line corresponds to a significance level  $\alpha$  of 0.1).



**Figure S13.** Phase diagram of landslide evolution validated on synthetic velocity datasets. In the Letter, three different regimes are theoretically defined (indicated by the shaded areas): black-swan catastrophic landslides, dragon-king catastrophic landslides, and slow landslides. The forecast method is used to predict the failure of synthetic landslide cases generated from an inverse gamma distribution (for the pre-crisis phase) in conjunction with the dynamic formulation, i.e., equation (S16) (for the crisis phase), with a broad parameter space of  $\omega$  and  $m$  explored. Our criterion for a successful prediction of failure is that the individual  $p$ -values of the majority of instruments (i.e., at least 3 instruments over 5 in total) together with the overall  $p$ -value drop below the significance level  $\alpha$  of 0.1 before the failure.

**Additional Supporting Information (Files uploaded separately)**

**Dataset S1.** Slope displacement monitoring data of the Preonzo landslide.

**Movie S1.** Pseudo-prospective forecast of the Preonzo landslide.

**Movie S2.** Pseudo-prospective forecast of the Veslemannen landslide.

**Movie S3.** Pseudo-prospective forecast of the Moosfluh landslide.