Supporting Information for

Benford's law as mass movement detector in seismic signals

Qi Zhou[1, 2†], Hui Tang[1], Jens M. Turowski[1], Jean Braun[1, 2], Michael Dietze[3,1], Fabian Walter[4], Ci-Jian Yang[1, 5], Sophie Lagarde[1, 2]

[1]Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Potsdam,

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

[3]Faculty of Geoscience and Geography, Georg-August-Universität Göttingen, Göttingen,

[4]Swiss Federal Institute for Forest, Snow and Landscape Research, Zürich,

[5]Department of Geography, National Taiwan University, Taipei,

Corresponding author: Qi Zhou (qi.zhou@gfz-potsdam.de, F427, Telegrafenberg, 14473 Potsdam, Germany)

Contents of this file

Introduction

This file includes supplementary text, figures, tables, and code for the manuscript titled *Benford's law as mass movement detector in seismic signals*.

Supplemental Text1 presents a comprehensive description of the methods employed to label debris flow events between the years 2013 and 2014. **Supplemental Text 2** illustrates the seismic data source and the catalog of debris flow events from 2017 to 2019. **Supplemental Text3** provides an elaborate explanation of the methods employed to fit the exponential curve as debris flow approaches the seismic station IGB02. Lastly, **Supplemental Text4** offers additional information regarding the definition of the training and validation classes, as well as details about the detector model elevation.

Figure S1 demonstrates the effects of different seismic data processing methods on Benford's Law (BL). **Figure S2** shows the input parameters of the debris flow detector model. **Figures S3** to **S4** and **Figure S5** display the BL results of training events between 2013 and 2014 and between 2017 and 2019, respectively, with a focus on partial event representations. **Figures S6** to **S9** present the BL results of different mass movements, and **Figures S10** to **S12** exhibit the BL results of various fluvial processes. **Figure S13** show the

kernel density for exponential fitting and power law exponent of noise and 38 BL-followed events. **Figure S14** display an example of exponential fitting curve results. **Figure S15** demonstrates the sensitivity of the debris flow detector model to different input parameters. **Figure S16** show the variation of power law exponent with different moving window sizes and detectors receiver operating characteristic. **Figures S17** to **S21** depict false-negative events encountered during the training process. **Figure S22** illustrates the relationship between BL and dataset range/signal frequency. **Figures S23** to **S25** highlight the potential of using a seismic array to improve detection performance, reduce the accuracy of false negative and improve the accuracy of true positive.

Table S1 provides parameters of seismometer stations for 2013-2014 and 2017-2019. **Table S2** to **S3** offer detailed information on start and end times, and whether BL was followed for the 24 training and 21 validation events. **Table S4** presents the seismic data source for different mass movements and fluvial processes to examine compliance with BL. **Table S5** displays the exponential fitting coefficients for all 45 debris flow events.

Text S1.

Methods to label debris flow training events (2013-2014)

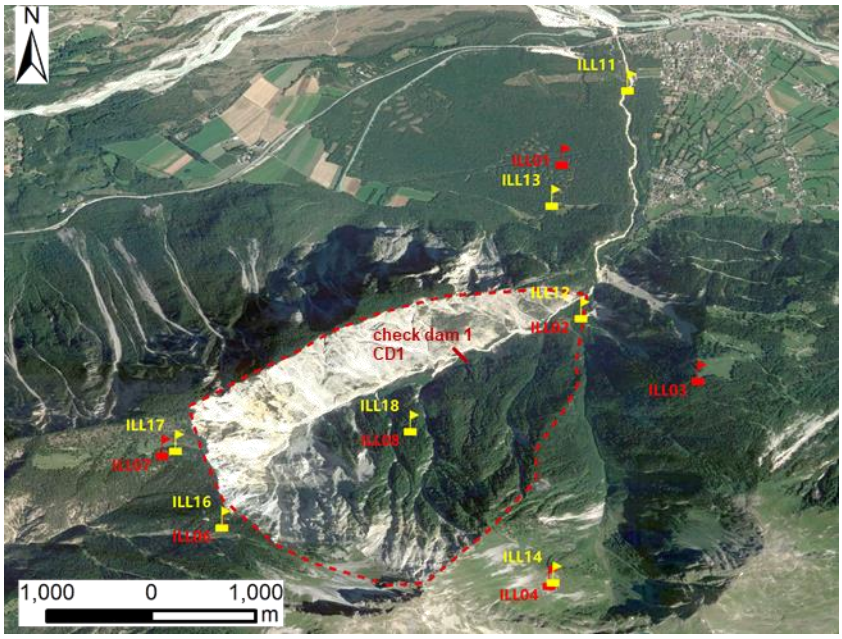
In order to build a training events catalog using data from 2013 to 2014, we utilized the data recorded by the seismic station IGB02 (located at coordinates 46.2863735252704, 7.62780978682399, elevation 933m, and 60m away from the trunk channel) to extract debris-flow events. The Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) warning systems provides traces for labelling ten debris flow events that occurred during the year 2013-2014, and we manually selected the event start and end time based on the seismic signals of station IGB02 (Table S2).

For debris flow events that may have occurred but were not recorded by the warning systems, we used the data from the seismic station IGB02 to extract events. The data were filtered between 1 and 90 Hz, and the spectrograms were plotted on a daily basis. Then the spectrograms were manually analyzed to retrieve the events. Here we refer to the methodology described in Burtin et al. (2014, 2016), Chmiel et al. (2021), Belli et al., (2022) and mainly focus on two main features that (1) the event duration of waveforms (debris flow should last more than 20 minutes), and (2) frequency features of the spectrogram (the characteristic frequency between 1 and 50 Hz, and peak frequency is around 7 Hz). For debris flows, we do not qualify an event corresponding to the surge, one debris flow could have different surges, and all these surges were merged as one event. Please refer to Table S2 for detailed catalog information. Because we do not have other data to validate the manually extracted events, it may introduce uncertainty.

Text S2.

Data source of debris flow validation events (2017-2019)

Between 2017 and 2019, there were 22 events collected by WSL (Chmiel et al., 2021), while the event 2017-05-19 is not available for ILL02. This event does not be included in the validation dataset. For event details, please refer to the supporting information for "*Machine Learning improves debris flow warning*". https://doi.org/10.1029/2020GL090874



Locations of seismometers at the Illgraben catchment from 2017 (red), 2018-2019 (yellow).

Text S3.

Methods to fit the exponential curve

For each BL-followed debris flow event, we first identify the one-minute time window t_i corresponding to the optimal goodness of fit within the manually labeled start and end times. Since the speed of debris flows can vary, the time required to record their approach by the seismometer may differ. To account for this variability, we select different lengths of raw waveforms (1 to 5 minutes before t_i, by minute) to fit an exponential curve. We then calculate the interquartile range (*iq*) of the raw waveforms using a 1-second sub-window. Next, we extract the dataset ***X*** during and before t_i. In cases where *iq* decreases at some points, we feed the data before the optimal of the ***X*** dataset to the *scipy.optimize.curve_fit* algorithm to obtain the exponential curve (Equation 1) and correlation coefficient R^2. Finally, we select the optimal fitting curve (optimal R^2) based on all fitting, and the results are listed in Table S5.

	$S(t) = a * e^{b*t} + c$	(1)

where *S* is the seismic signals (unit by counts), *t* is time (unit by second), and *a*, *b*, and *c* are the coefficient of the exponential function. If *b* is smaller than zero, then R^2 is zero.

The codes are available in content **Code S1**.

Text S4.

Define training and validation class

Our dataset has two labels: positive (debris flow) and negative (not debris flow). During the training process, the 24 events (Tables S2) between 2013 and 2014 were labeled positive based on data from station IGB02. An additional 1200 negative cases were randomly selected from outside the event at random start times and durations (between

20 minutes and 6 hours). The ratio between the positive and negative cases (P2N) is 0.02. During the validation process, the 21 debris-flow events between 2017 and 2019 (Tables S3) were labeled positive based on stations ILL02 and ILL12. A further 1050 negative cases were randomly selected using the same method as the training procedure with the same P2N ratio. The validation dataset of 1071 events was processed using the optimal detector determined during training, and the validation results with an existing random forest model trained with data from 2017 to 2019 recorded by the same seismic network using more than 70 seismic features (Chmiel et al., 2021).

Define TP, TN, FP, and FN class

We define the confusion matrix as follows: A true positive (TP) is an event that is classified as a debris flow by observation and our detector. A true negative (TN) is an event that is not classified as a debris flow by either observation or our detector. An event is labeled false positive (FP) when it is labeled as a debris flow by our detector but considered as a non-debris-flow event in observation. An event is labeled as false negative (FN) when an observed debris flow is not classified as debris flow by the detector.

Evaluate detector model

A good classification model should maximize the number of true-positive predictions and minimize the number of false-positive and false-negative predictions. The overall performance of the classifier is quantified as a true-positive rate (*TPR*, recall rate or sensitivity), false-positive rate (*FPR*, fall-out rate), false-negative rate (*FNR*, miss rate), and true-negative rate (*TNR*, specificity) which are calculated as Equation 2-5:

	$TPR = \frac{TP}{TP + FN}$	(2)
	$FPR = \frac{FP}{FP + TN}$	(3)
	$TNR = \frac{TN}{TN + FP}$	(4)
	$FNR = \frac{FN}{FN + TP}$	(5)

where TP, TN, FP, and FN are the number of true-positive, true-negative, false-positive and false-negative events.

We calculated the F-score or F1 score (F1) and the Threat Score (TS), which are commonly used to represent the predictive capability of classification models (Equation 6-7). The F1 measures the accuracy of a binary classification model and is a harmonic mean of precision and recall. When F1 gets the highest possible value (one), it means that the precision and recall are perfect, and F1 receives the lowest value (zero) when either the precision or the recall is zero. The TS is a measure of the overall performance of the classification model. In a perfect model, the threat value would be equal to one, with each

false prediction (false-negative or false-positive event) decreasing the value of the threat score.

	$F1 = \frac{2TP}{2TP + FN + FP}$	(6)
	$TS = \frac{TP}{TP + FN + FP}$	(7)

The results of F1 and TS are displayed in Figure S15, the detector model is considered the best when F1 is maximum.

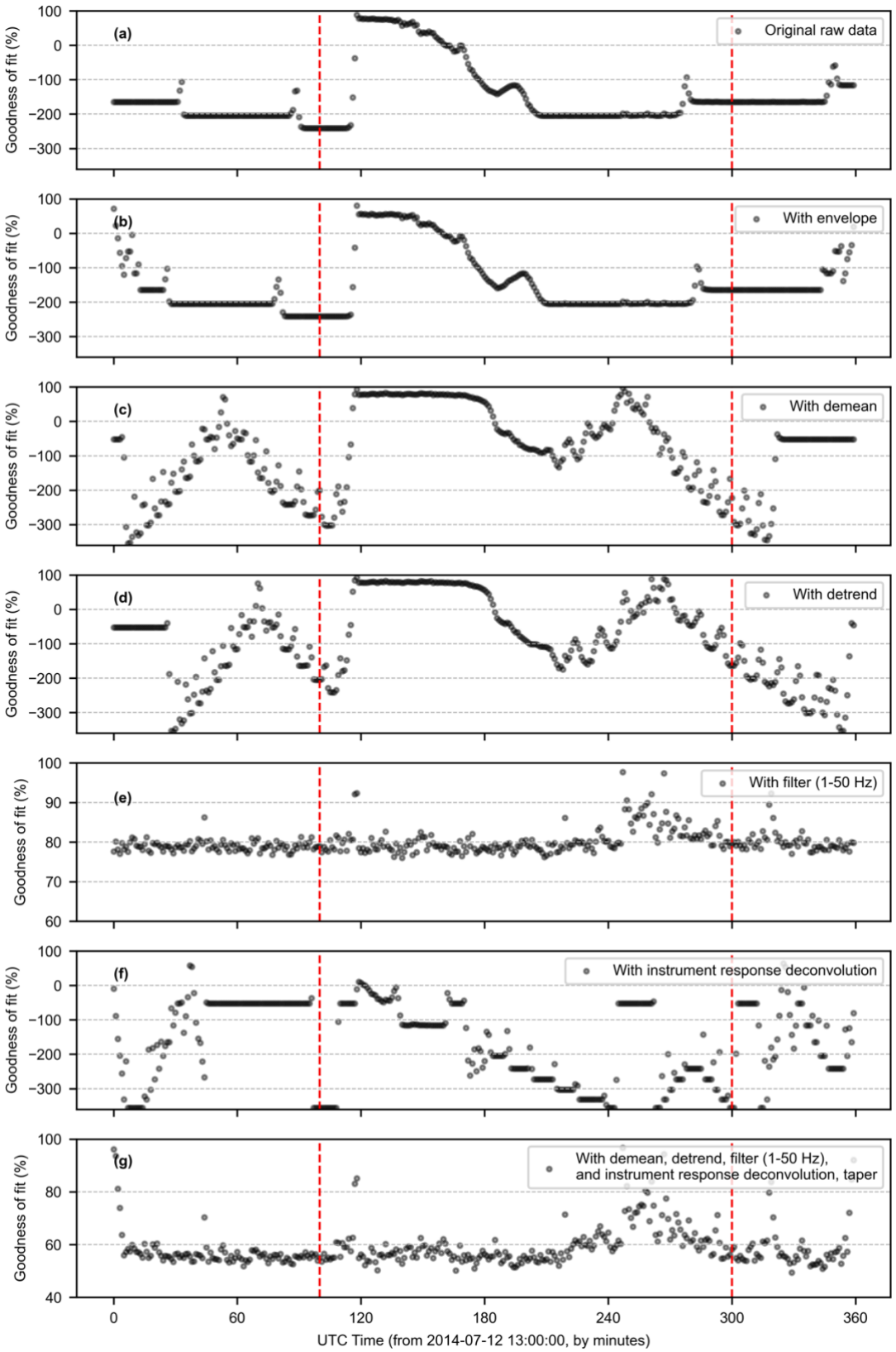


Figure S1. Effects of data processing methods on BL, coding by R ESEIS Dietze (2018) . The red dashed vertical lines are manually labeled start and end time. The codes are available in **4. Codes**.

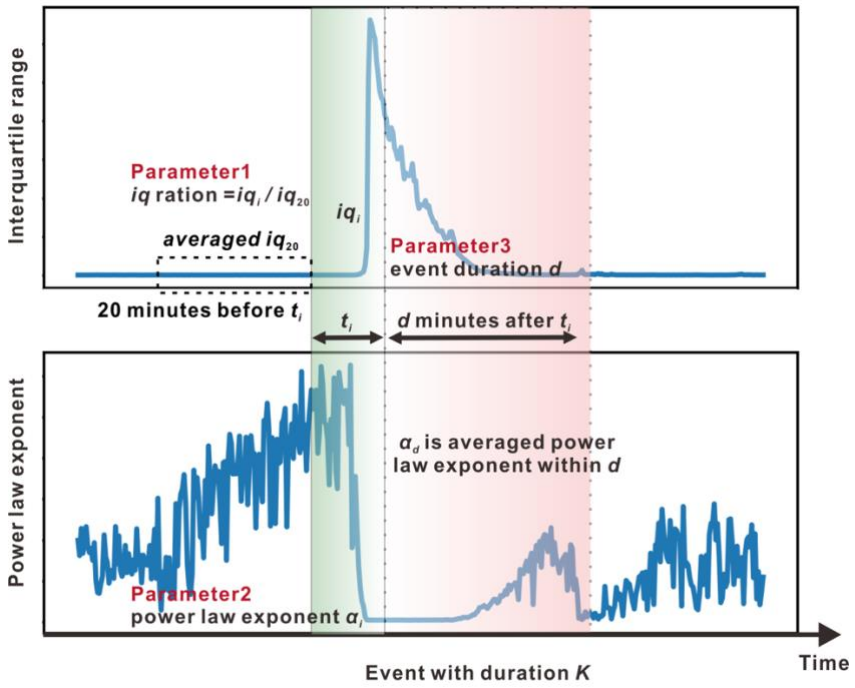


Figure S2. The input dimensionless parameters and range of event detector**.**

The input parameters of detector in Figure S2 are listed below:

	$R_{iq} = iq_i / iq_{i20}$,	(8)

where iq_i is the parameter 1, interquartile range ratio of time i and iq_{i20} is the average interquartile range of 20 minutes before the time i.

Range of input parameters		
Interquartile range (iq) ratio	Power law exponent α_i	Event duration d (unit: minute) or α_d
From 2 to 10	From 1.01 to 1.50	From 3 to 20
Optimal parameters from training data		
iq = 4	α_i = 1.25	d = 20

The validation process with optimal parameters could be completed in 113 seconds (do not count parameters calculating time, and codes were operated with 1 node and 48 G memory) via GFZ Cluster.

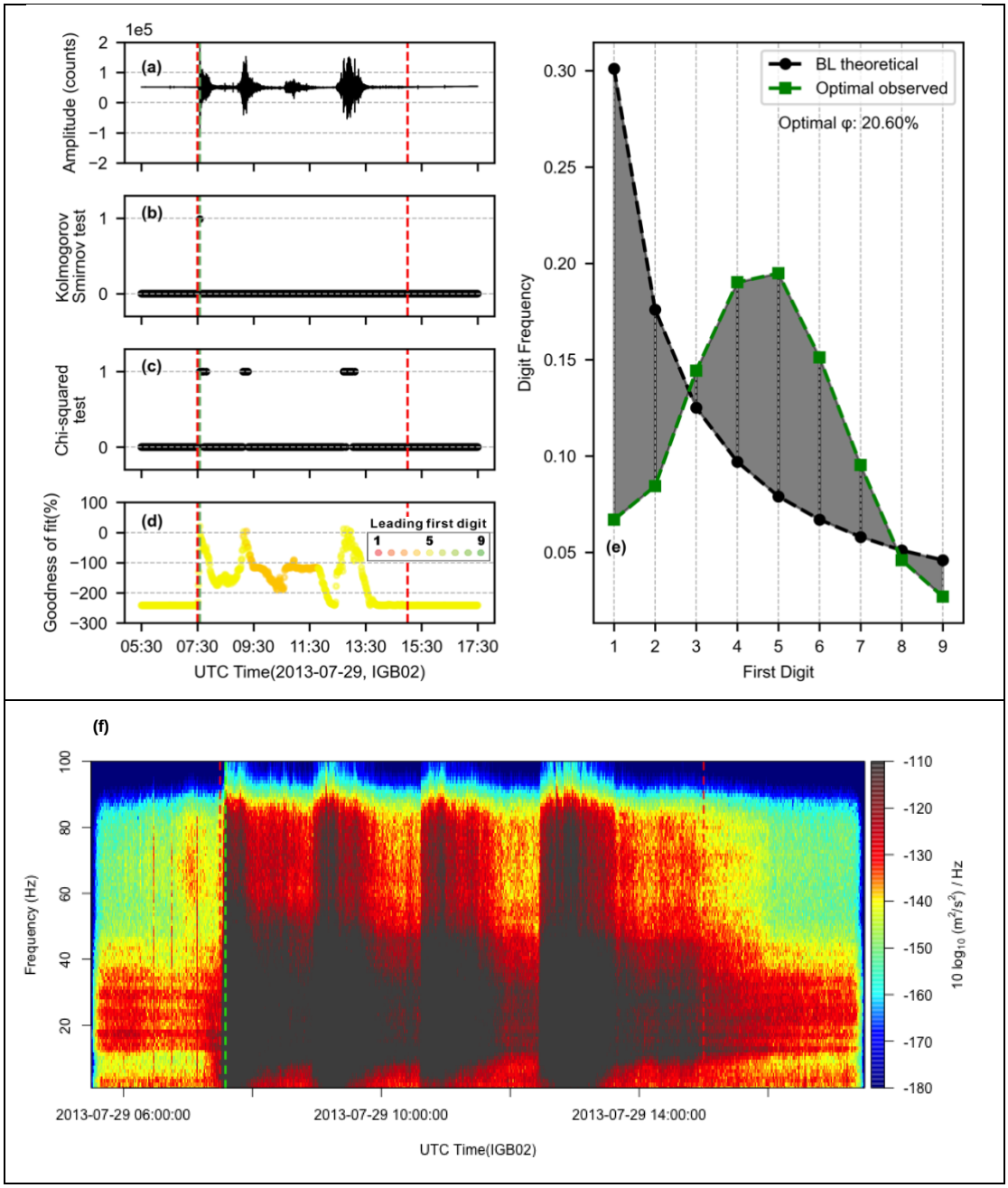


Figure S3. Raw waveform (a) and spectrogram (f) generated by a **flood** (2013-07-29, training events, WSL label, IGB02 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

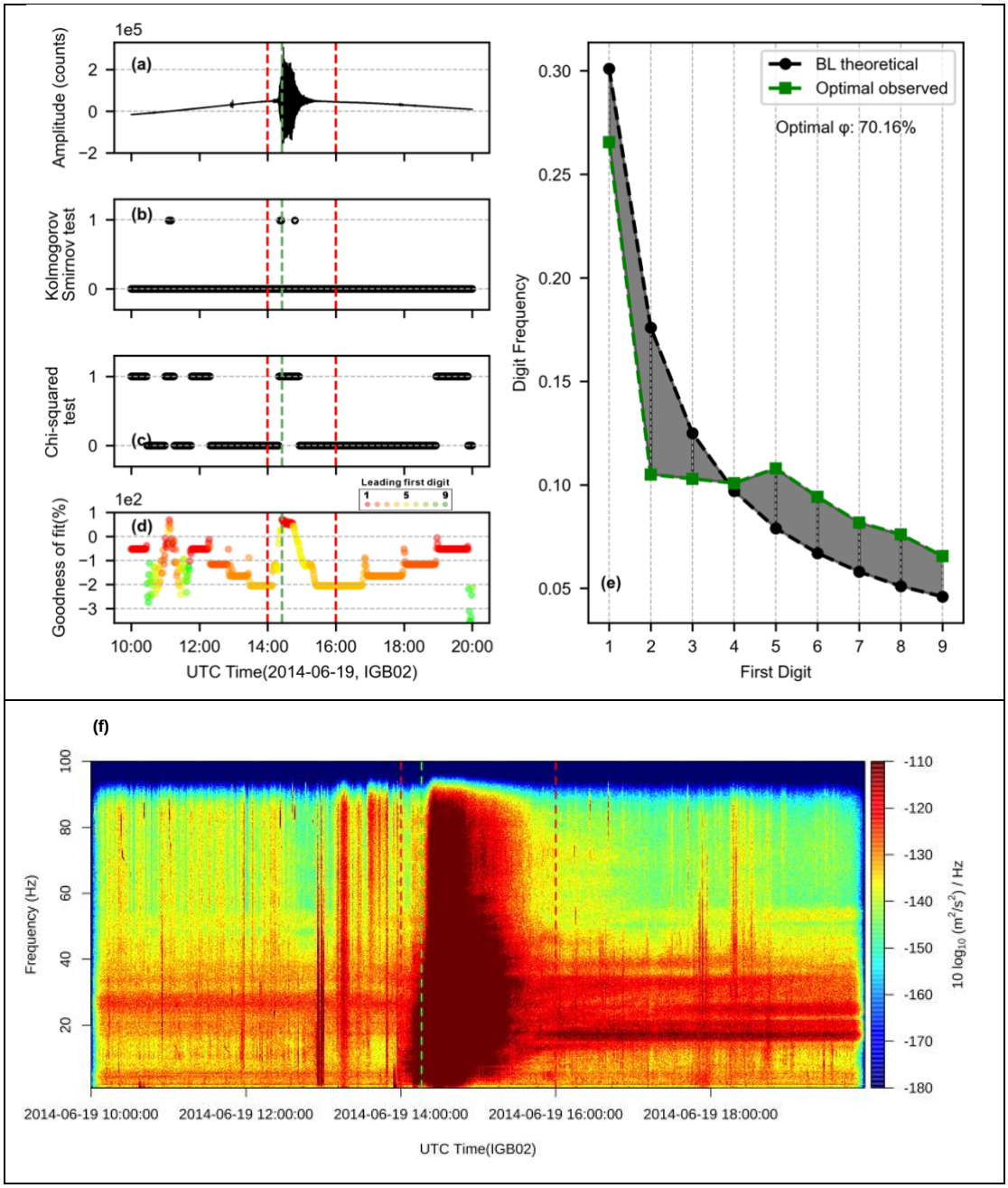


Figure S4. Raw waveform (a) and spectrogram (f) generated by a **debris flow** (2014-06-19, training events, GFZ label, IGB02 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

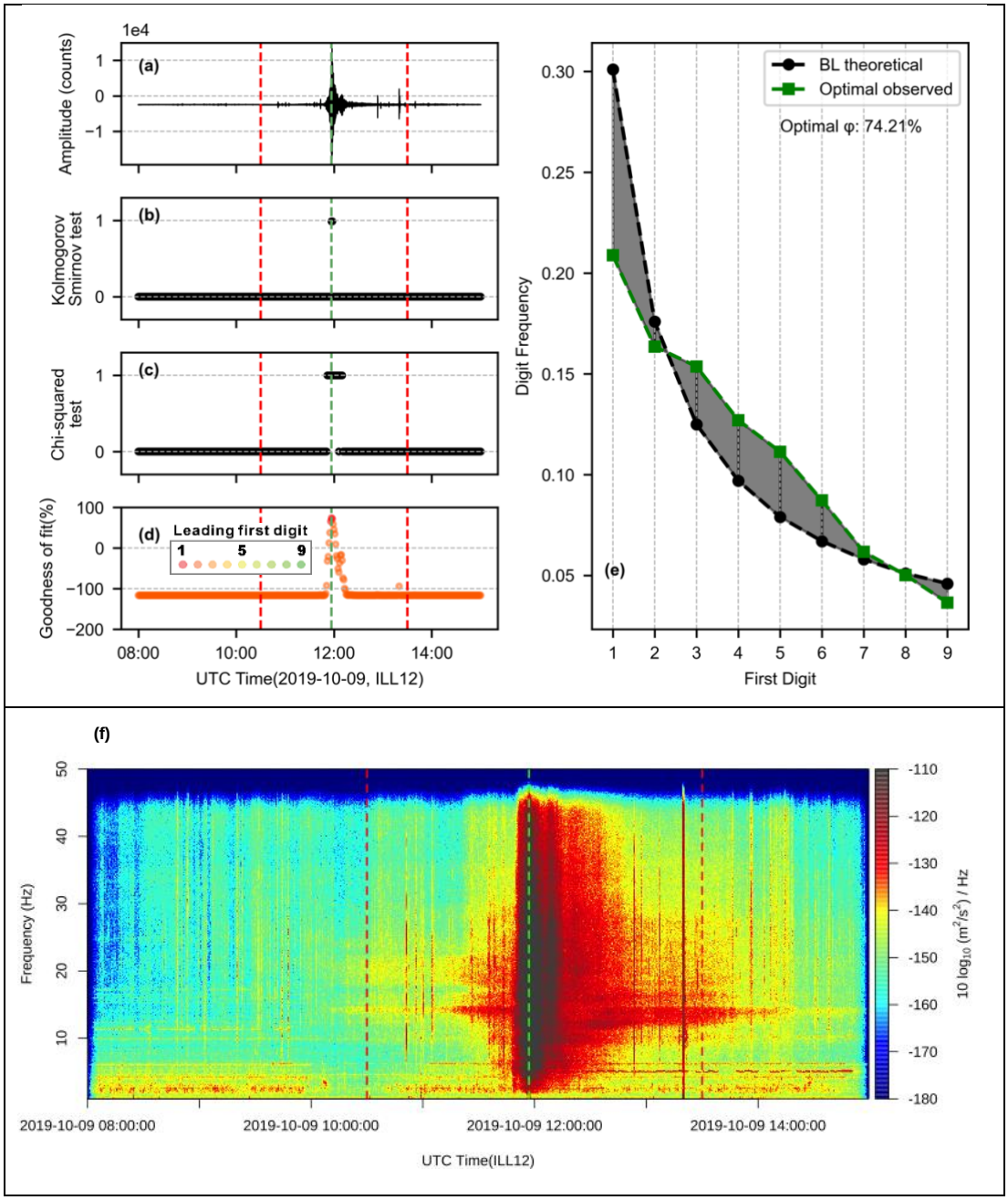


Figure S5. Raw waveform (a) and spectrogram (f) generated by a **debris flow** (2019-10-09, validation events, WSL label, ILL12 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

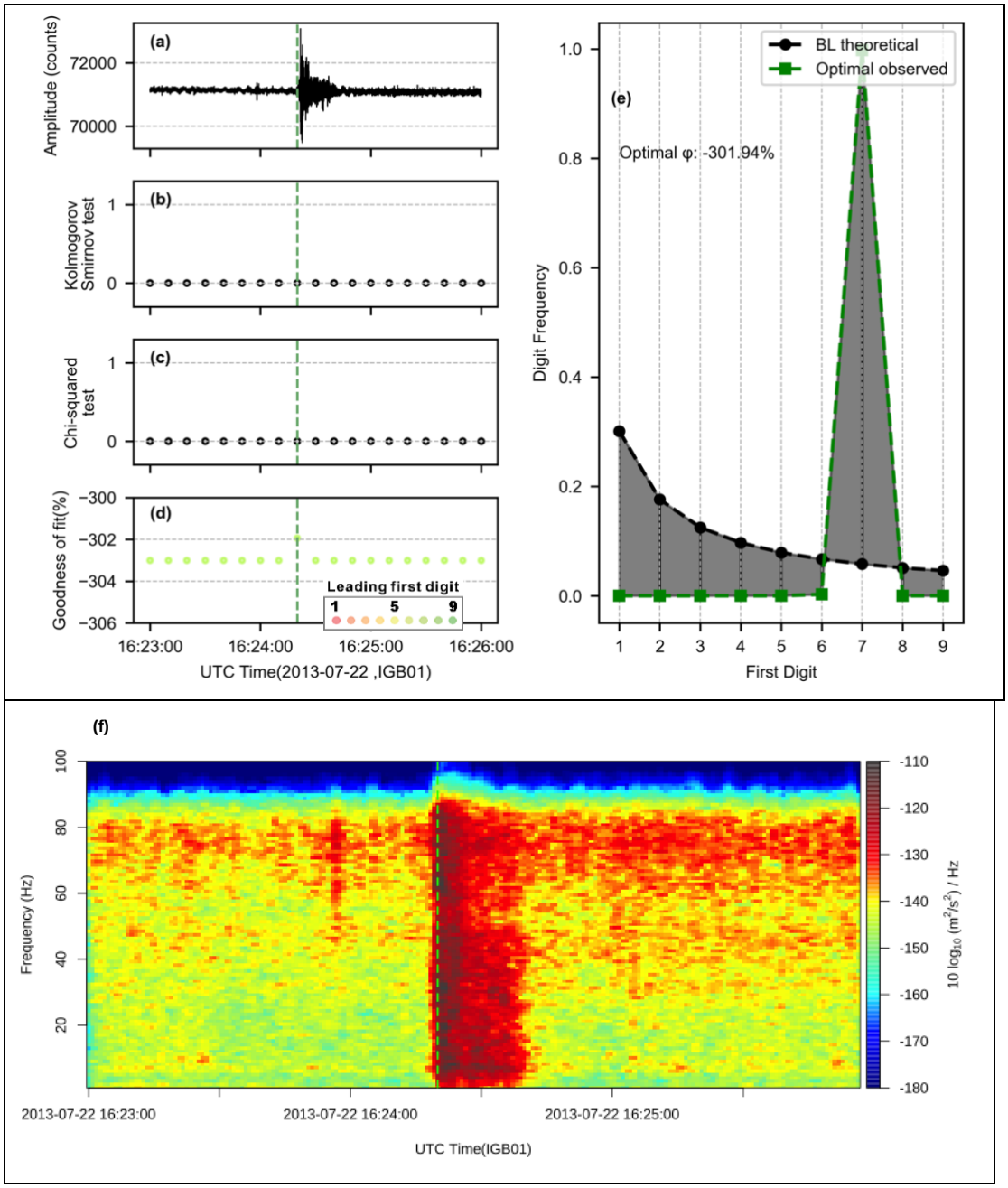


Figure S6. Raw waveform (a) and spectrogram (f) generated by the **Rockfall 1** (2013-07-22 IGB01 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The green dashed vertical line is the time for an optimal goodness of fit (e).

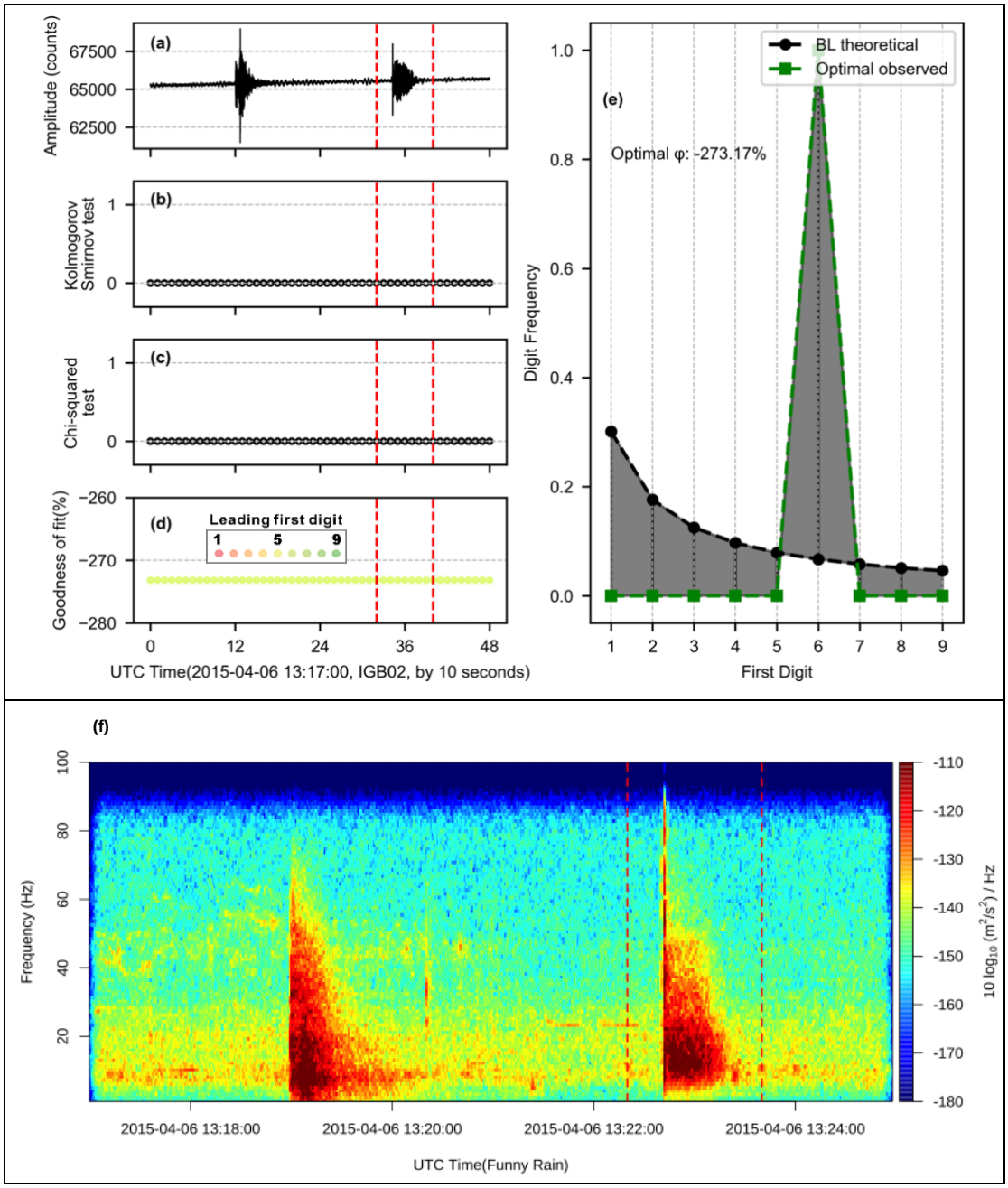


Figure S7. Raw waveform (a) and spectrogram (f) generated by the **Rockfall 2** (2015-04-06, Funny Rain station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The first event was a small long-distance earthquake. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

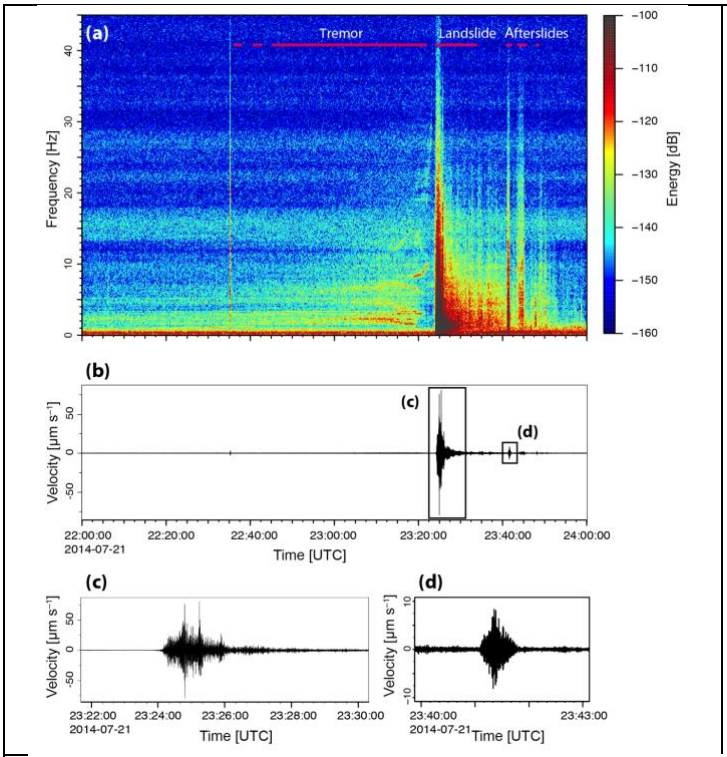


Figure S8-1. Spectrogram (a, unfiltered) and waveform (b, filtered between 1 and 45 Hz show precursory tremor and the after-slides at the Askja caldera on 21 July 2014. (c) Close-up of the landslide and (d) of one after-slide waveform. MOFO station, east component (Schöpa et al., 2018).

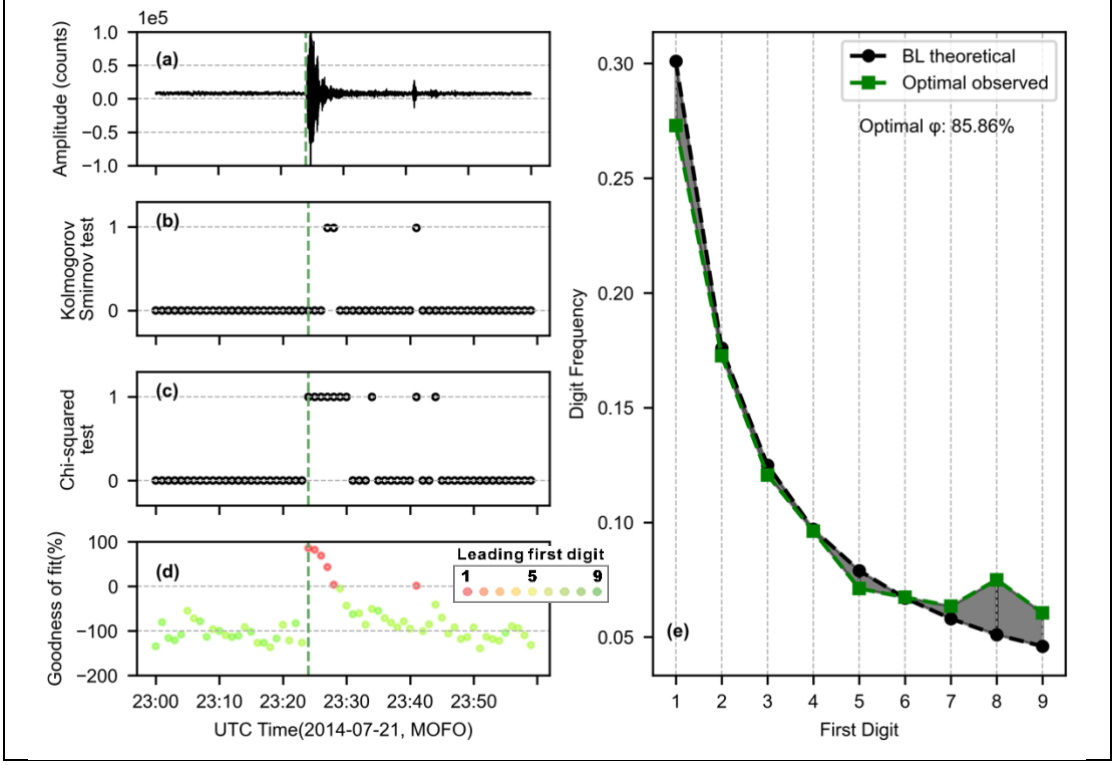


Figure S8-2. Raw waveform (a) and spectrogram (f) generated by the **landslide** (2014-07-21, MOFO station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The green dashed vertical line is the time for an optimal goodness of fit (e).

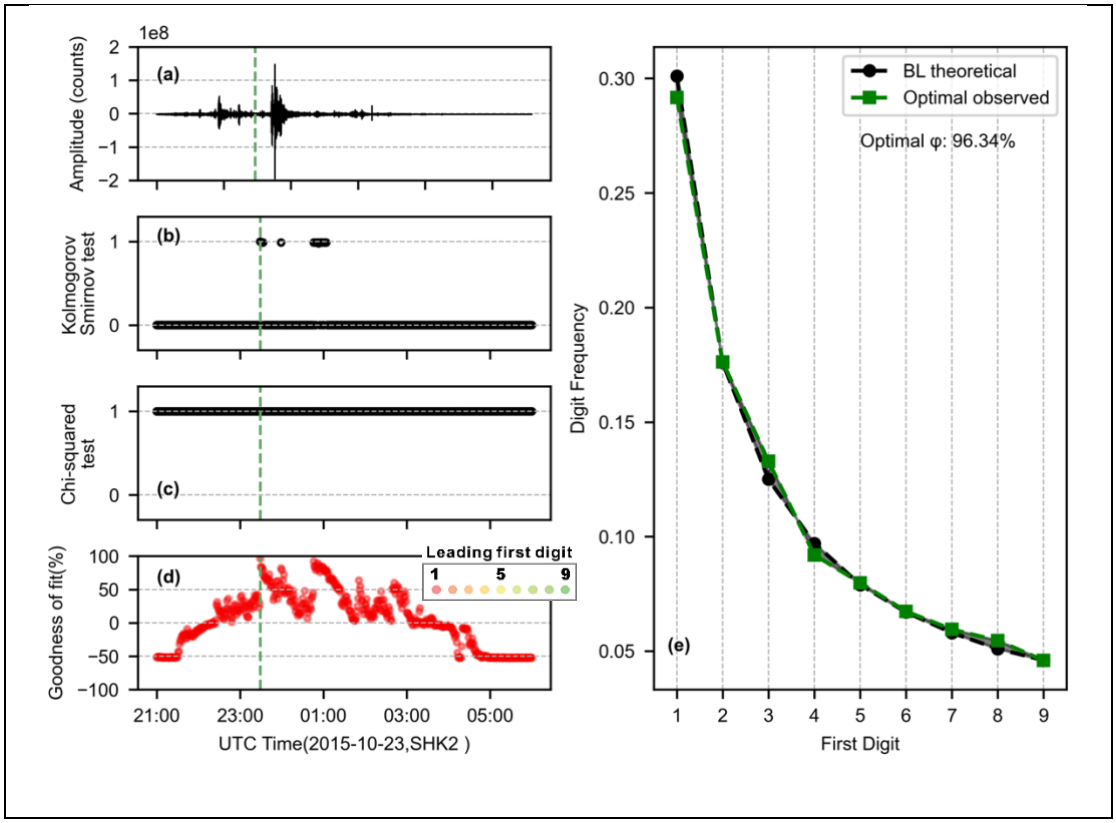


Figure S9. Raw waveform (a) generated by the a **hurricane-induced lahars** (2015-10-24 01:00, (Capra et al., 2018)). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The green dashed vertical line is the time for an optimal goodness of fit (e). *No instrument response information is available for spectrogram*.

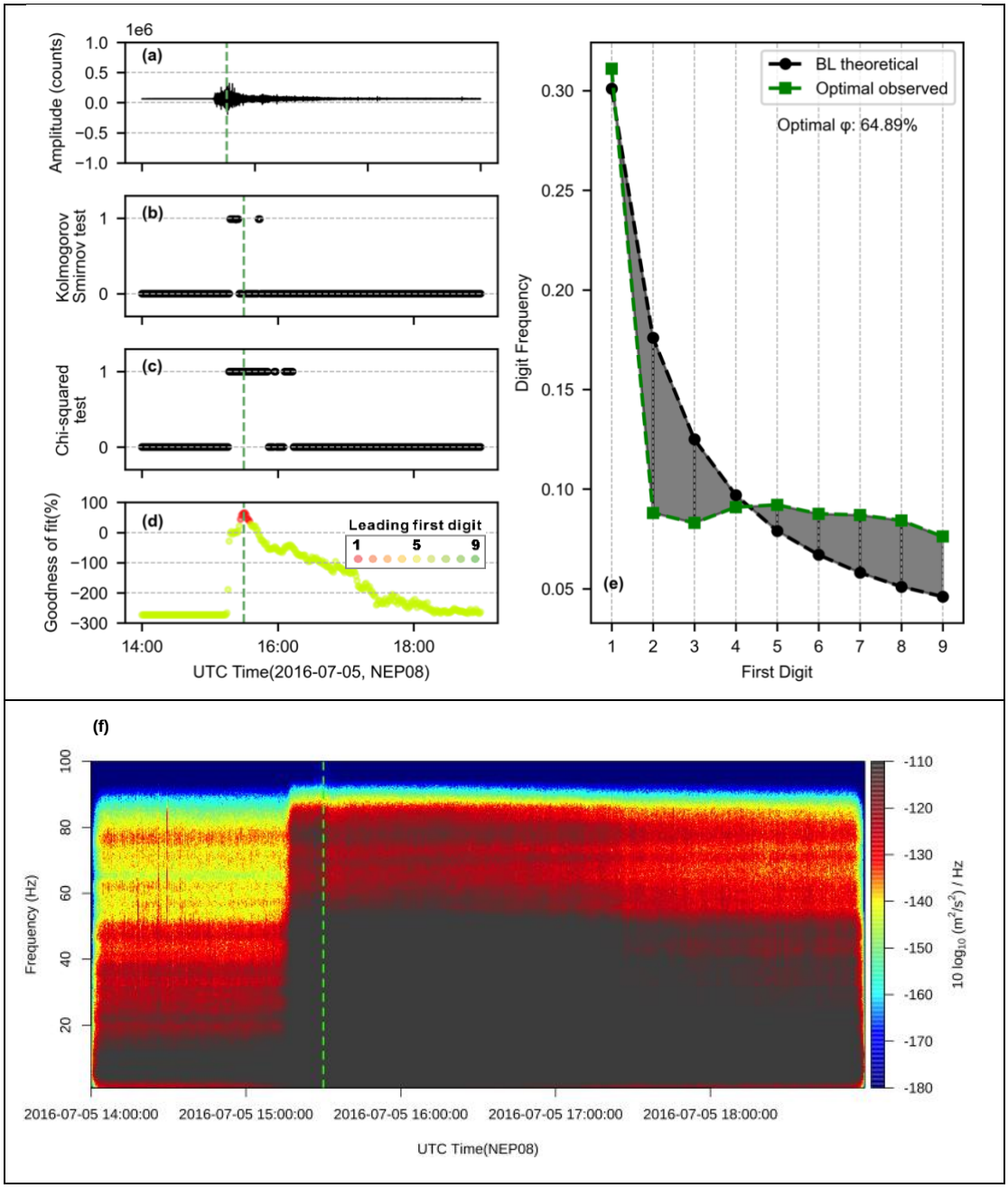


Figure S10. Raw waveform (a) and spectrogram (f) generated by a **glacial lake outburst flood** (2016-07-05, Hindi (NEP08) station, Cook et al., 2018). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The green dashed vertical line is the time for an optimal goodness of fit (e).

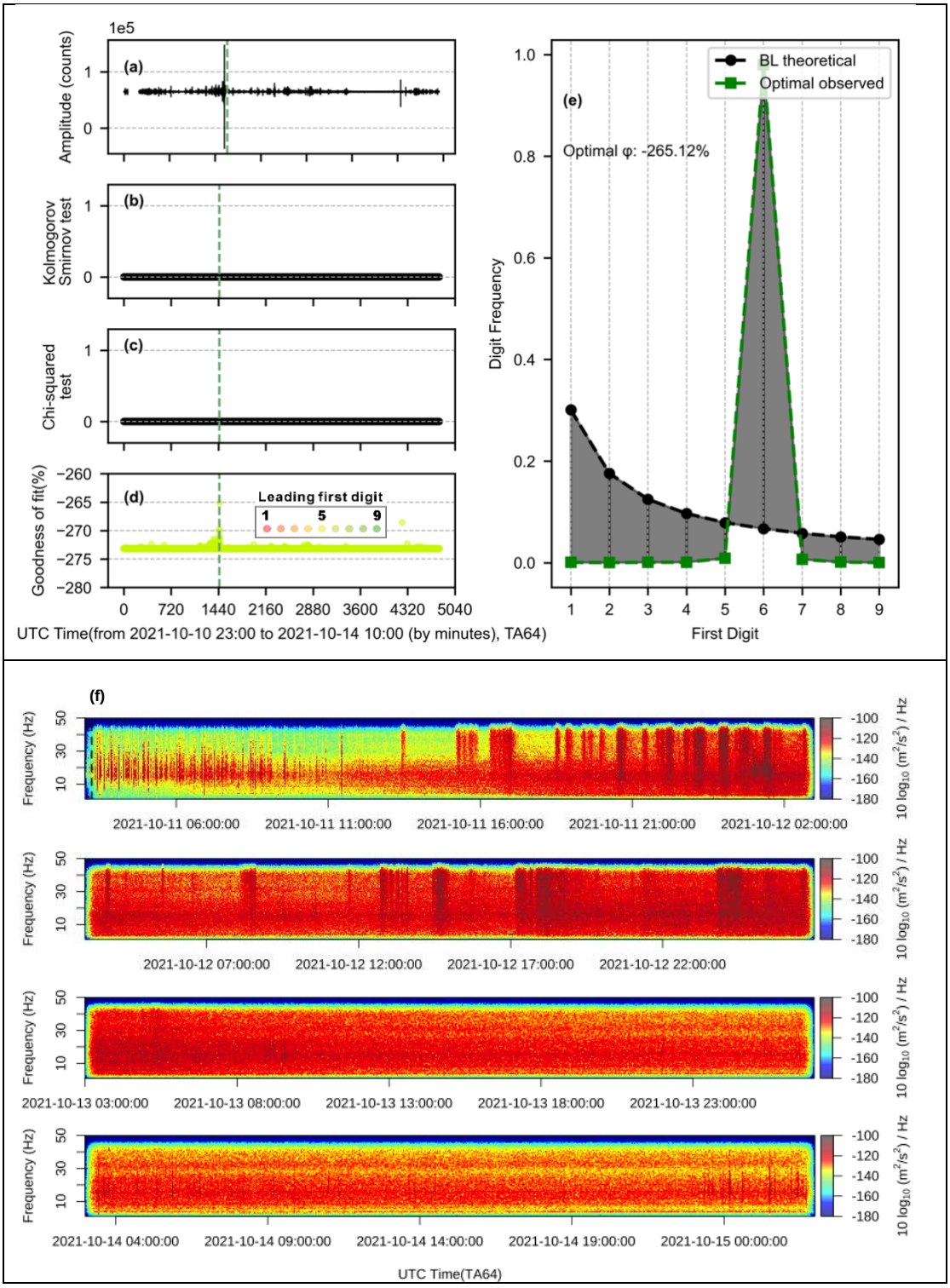


Figure S11. Raw waveform (a) and spectrogram (f) generated by a **bedload transport** (from 2021-10-10 to 2021-10-14, TA64 station). (b) and (c) results from Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First digit distribution (e) of BL theoretical and observed optimal periods. The black dashed vertical line is the time for an optimal of the goodness of fit (e).

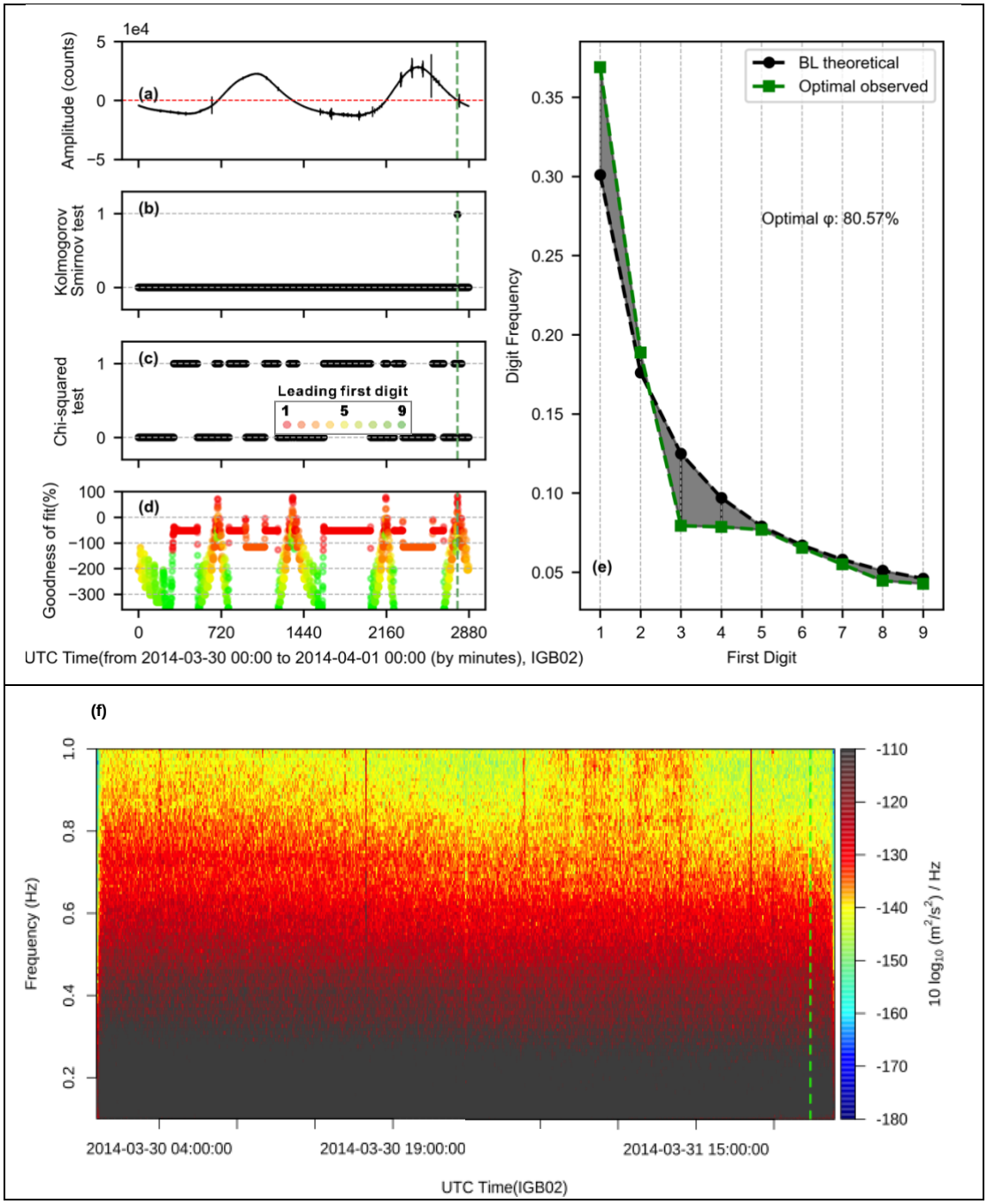


Figure S12. Raw waveform (a) and spectrogram (f) generated by **a long-period seismic signals (LP) and when LP is close to zero counts fluctuation** (from 2014-03-30 to 2014-04-01, IGB02 station, follow BL). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The green dashed vertical line is the time for an optimal goodness of fit (e). The max and min is -227 counts and 342 counts, the *iq* is 119 counts at 2014-03-31 22:22 to 22:23.

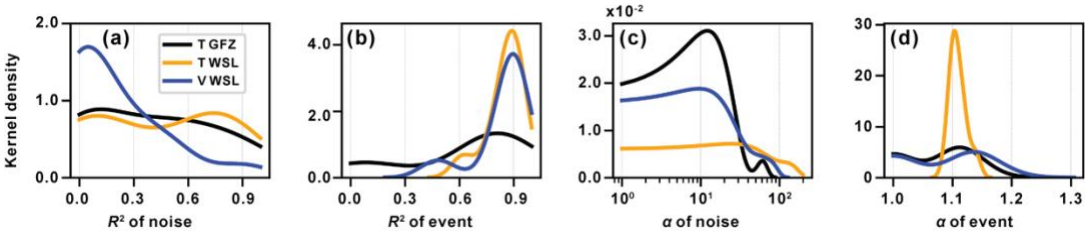


Figure S13. Kernel density for exponential fitting and power law exponent. (a) and (b) are the kernel density (bandwidth is 0.5) of R^2 of the exponential fitting. The coefficients of the exponential fitting for all events are listed in Table S5. (c) and (d) are kernel density (bandwidth is 0.5) of power law exponent α. The data displayed in (a) to (d) correspond to the optimal goodness of fit during 38 BL-followed debris flow events. T GFZ and T WSL represent training event datasets during 2013-2014 with GFZ and WSL labels, respectively. V WSL represents the validation event dataset during 2017-2019.

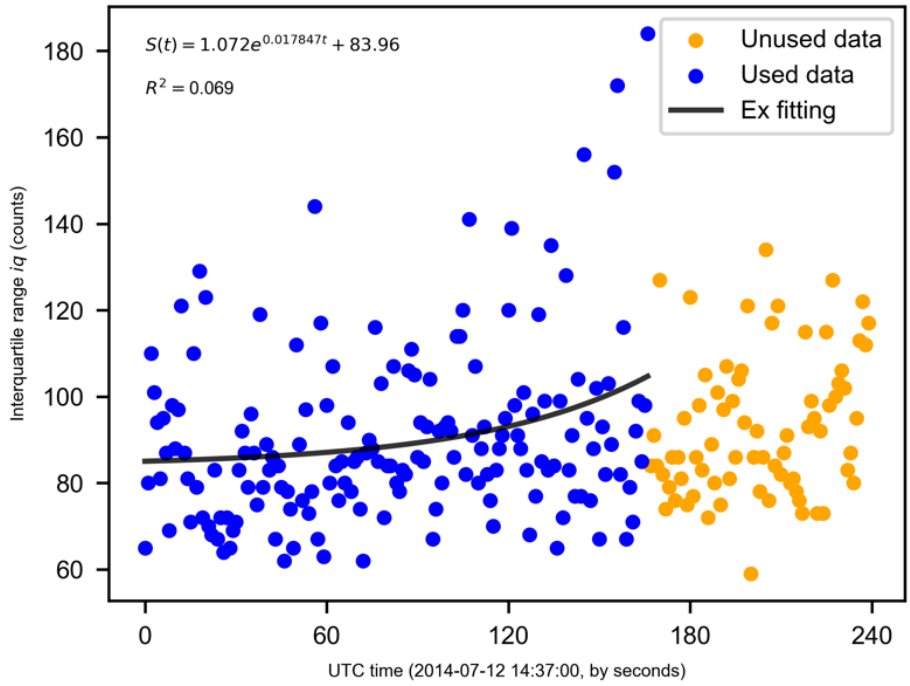


Figure S14. Exponential fitting of event 2014-07-12 debris flow noise.

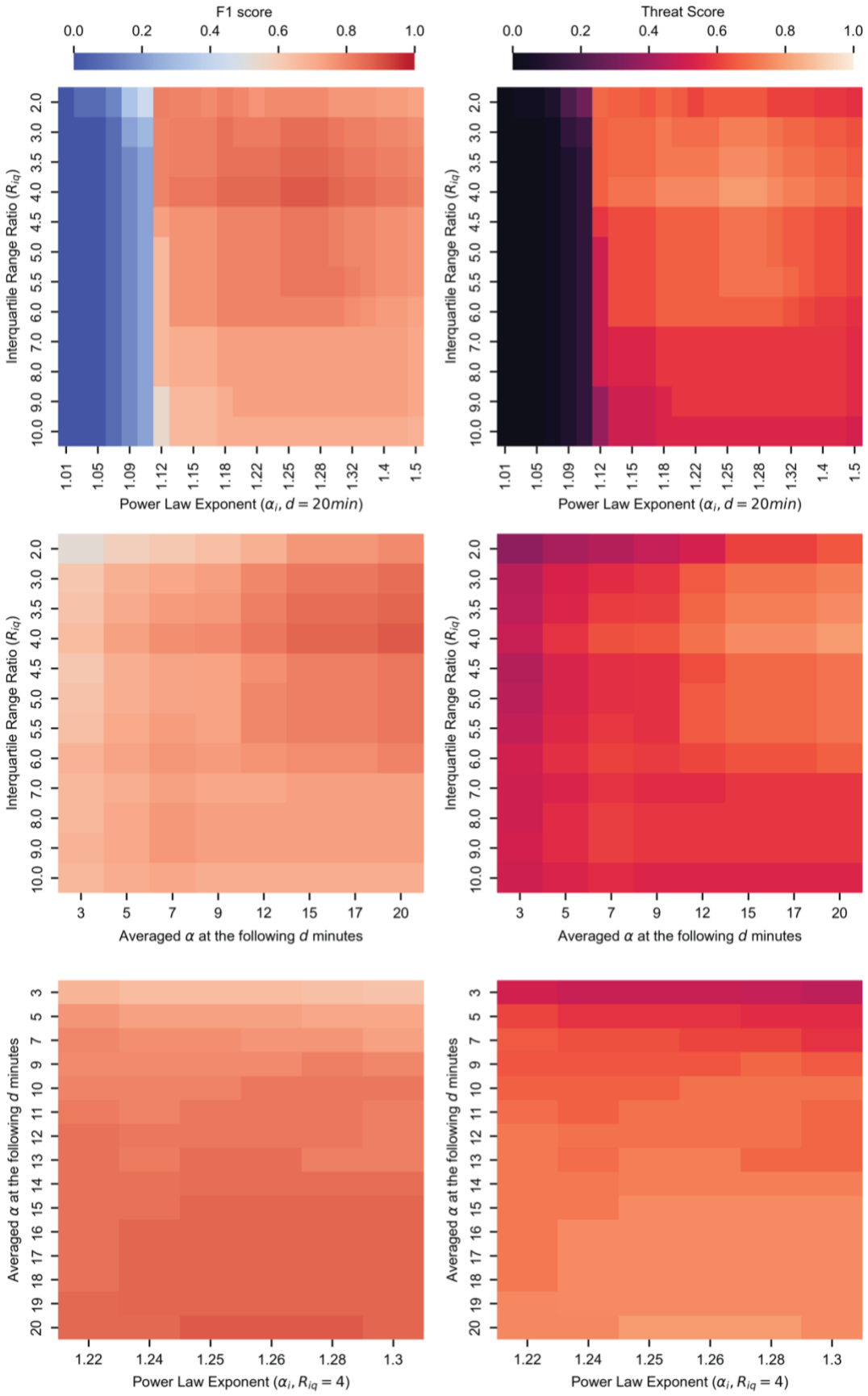


Figure S15. F1 score and Threat score for different debris-flow detector parameters.

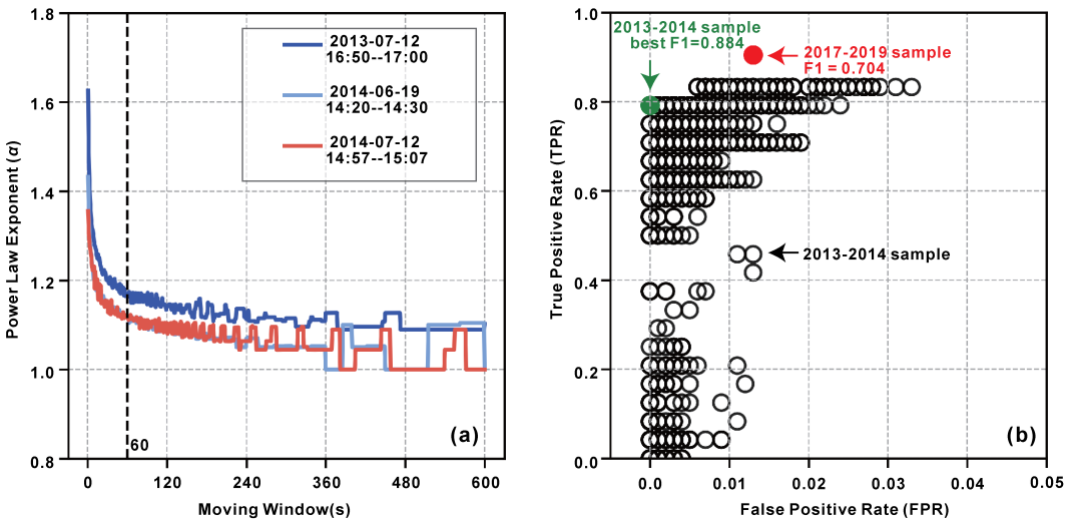


Figure S16. Variation of power law exponent with different moving window sizes and receiver operating characteristic of different detectors. (a) Averaged power law exponent α with different moving windows. The black dashed line indicates the 60 seconds window used in this work. (b) Receiver operating characteristic ROC of detectors with different parameters. The best detector (F1 0.884) obtained from the 2013-2014 training sample corresponds to FPR is 0, TPR is 0.792, and outputs of this detector for the sample from 2017-2019 FPR is 0.013, TPR is 0.905.

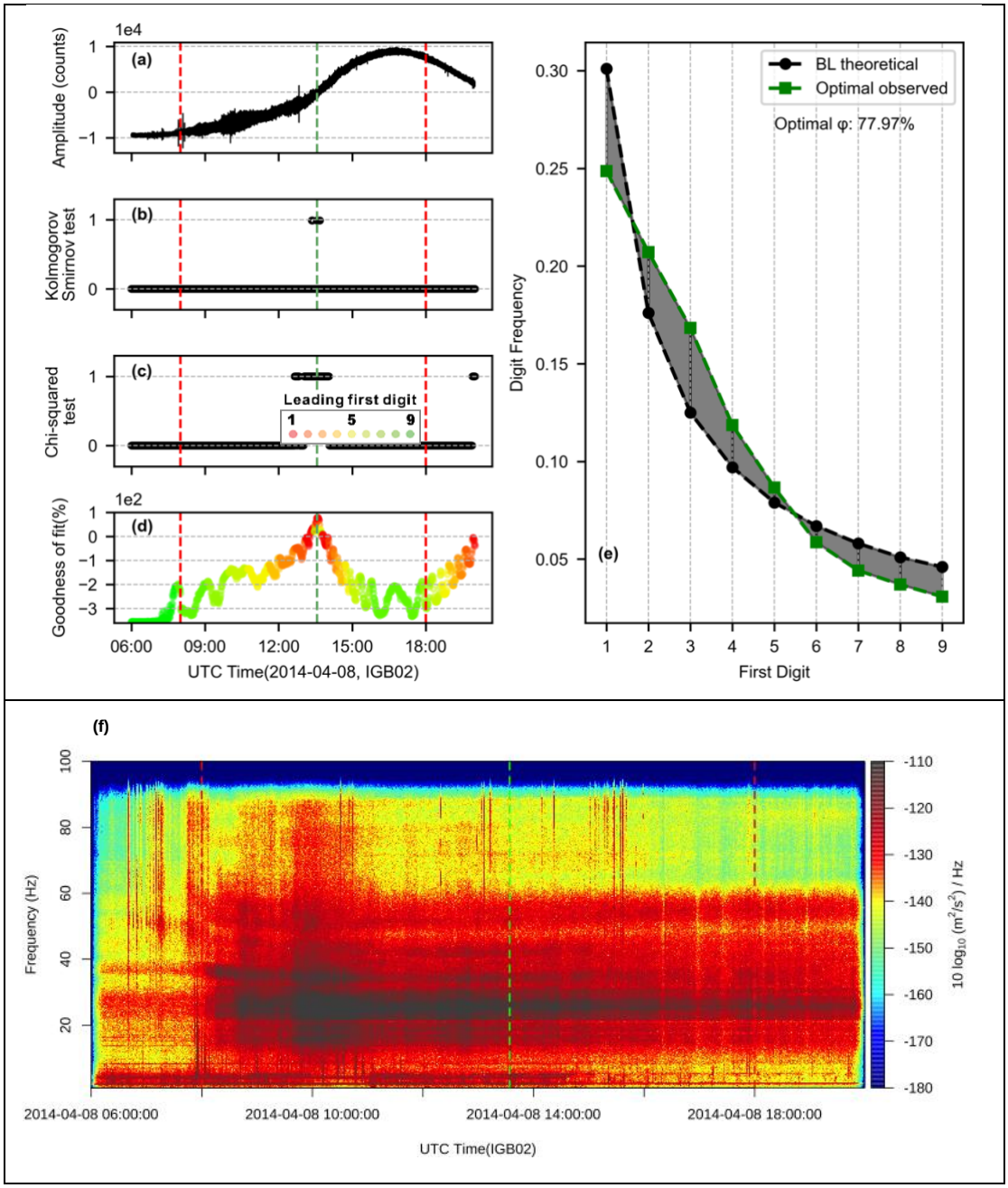


Figure S17. Raw waveform (a) and spectrogram (f) generated by a **debris flow** (**false-negative**, 2014-04-08, training events, GFZ label, IGB02 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

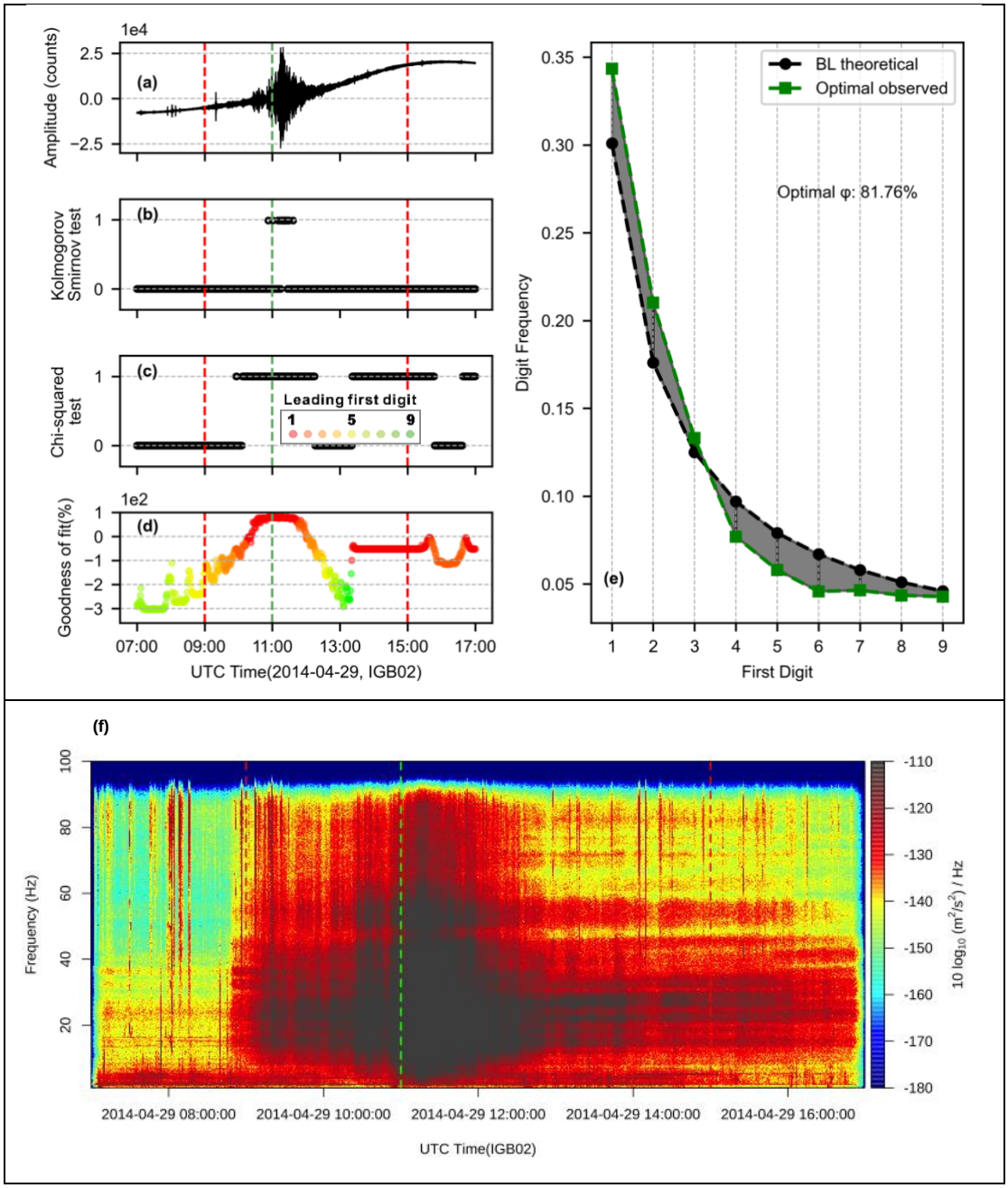


Figure S18. Raw waveform (a) and spectrogram (f) generated by a **debris flow** (**false-negative**, 2014-04-29, training events, GFZ label, IGB02 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

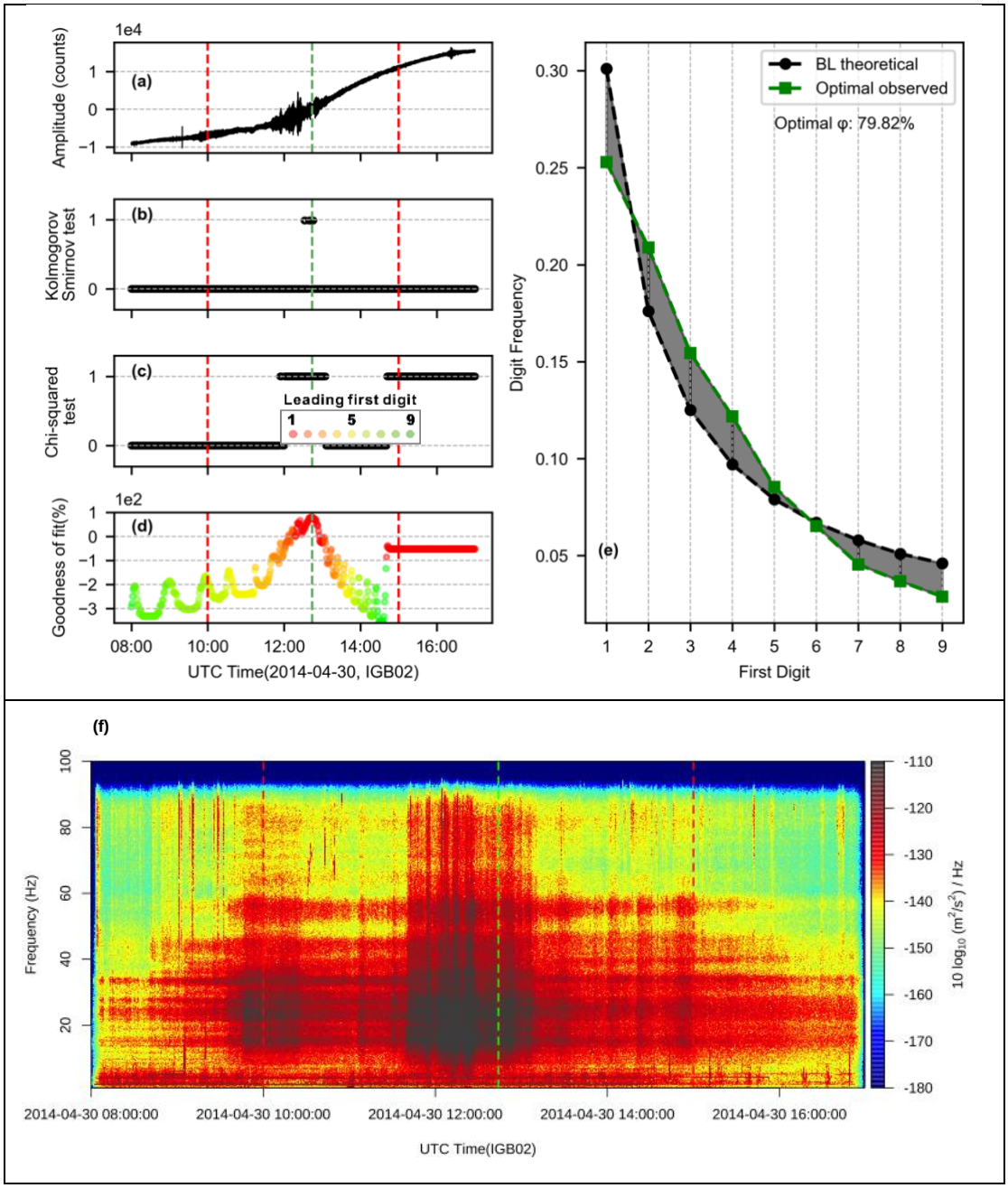


Figure S19. Raw waveform (a) and spectrogram (f) generated by a **debris flow** (**false-negative**, 2014-04-30, training events, GFZ label, IGB02 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

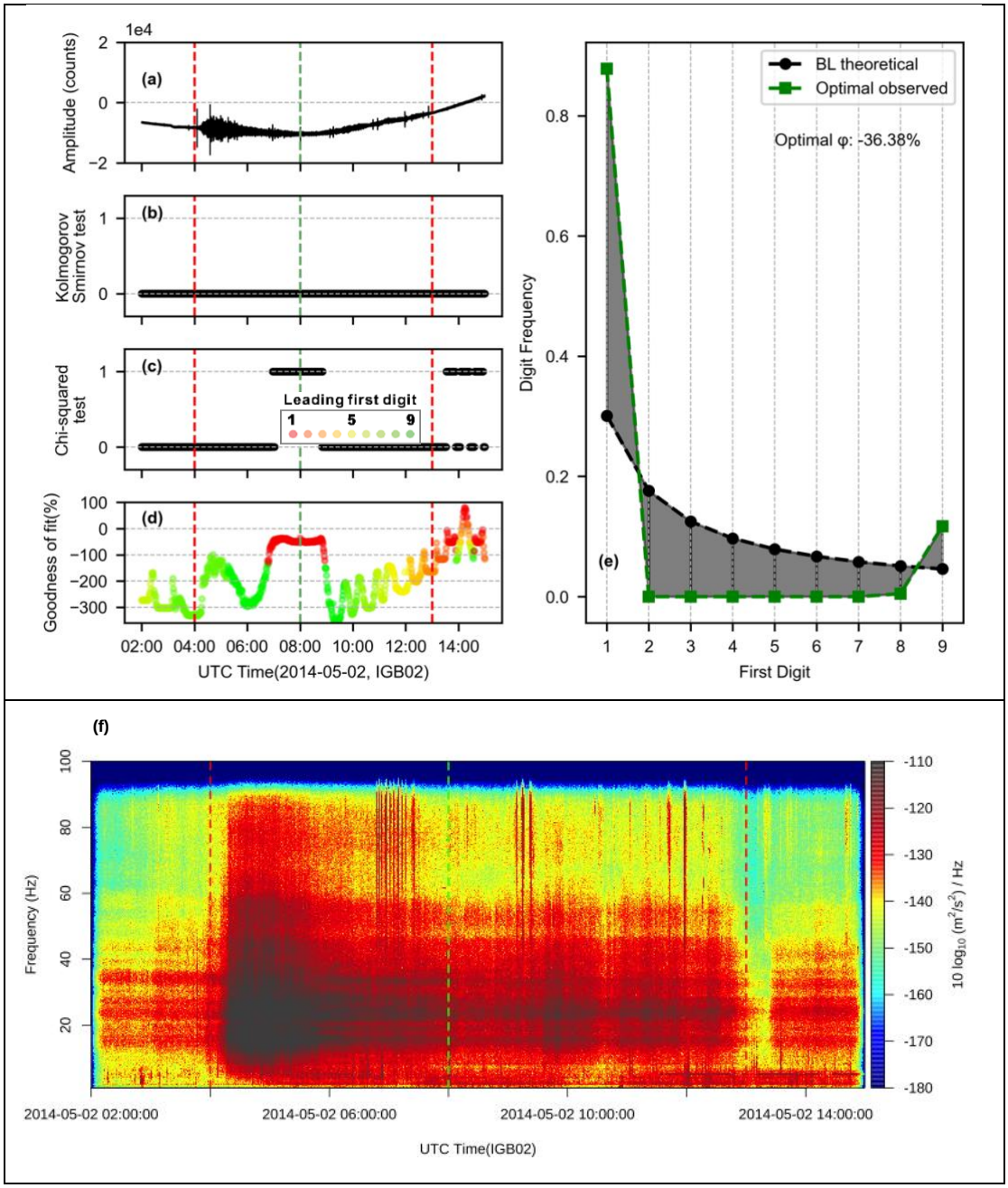


Figure S20. Raw waveform (a) and spectrogram (f) generated by a **debris flow** (**false-negative**, 2014-05-02, training events, GFZ label, IGB02 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

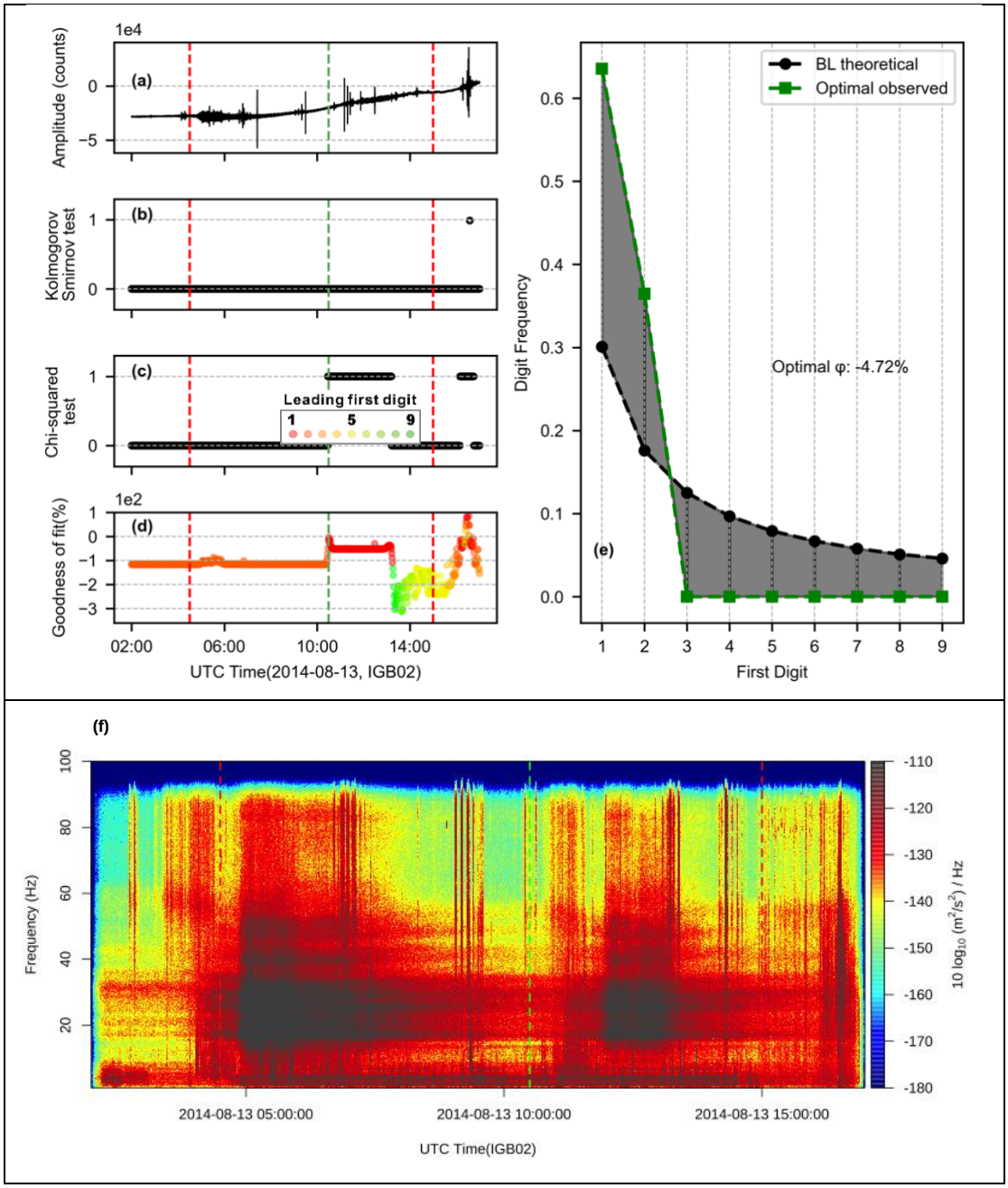


Figure S21. Raw waveform (a) and spectrogram (f) generated by a **debris flow** (**false-negative**, 2014-08-13, training events, GFZ label, IGB02 station). (b) and (c) results from the Kolmogorov-Smirnov test and the Chi-squared test. One means to accept the null hypothesis that observed first-digit distribution is similar to BL's theoretical value; otherwise, the value is zero. (d) the goodness of fit φ in different colors represents the leading first digit in each one-minute moving window. First-digit distribution (e) of BL theoretical and observed optimal periods. The red dashed vertical lines have manually marked the start and end times. The green dashed vertical line is the time for an optimal goodness of fit (e).

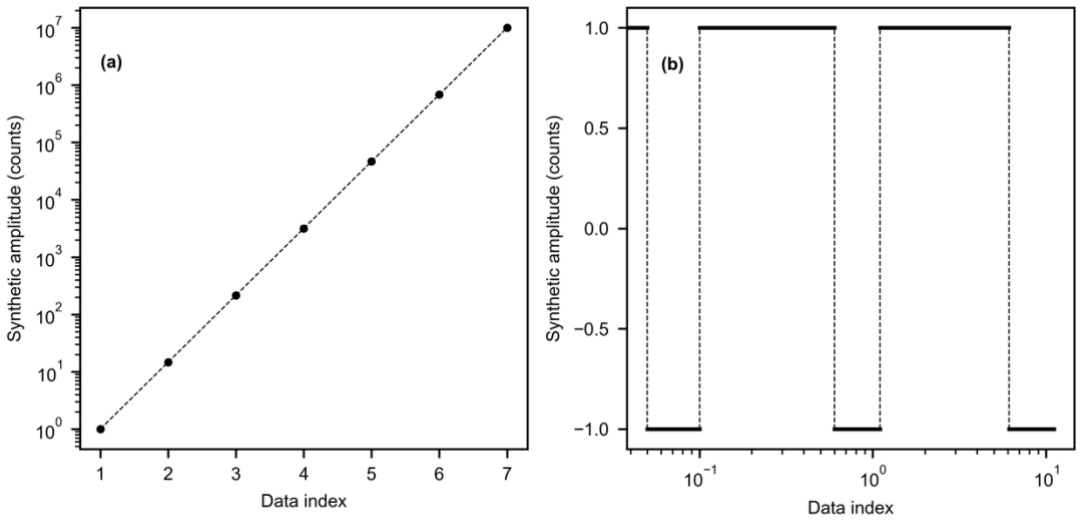


Figure S22. Synthetic amplitude dataset. (a) Datasets [1, 10, …, 10^7] with a first-digit of 1 that span 7 orders of magnitude but do not follow BL. (b) Datasets with a first digit of 1 that has a period of 0.1 Hz (data index from 0 to 0.1), 1 Hz (data index from 0.1 to 1.1), and 10 Hz (data index from 1.1 to 11.1) but do not follow BL.

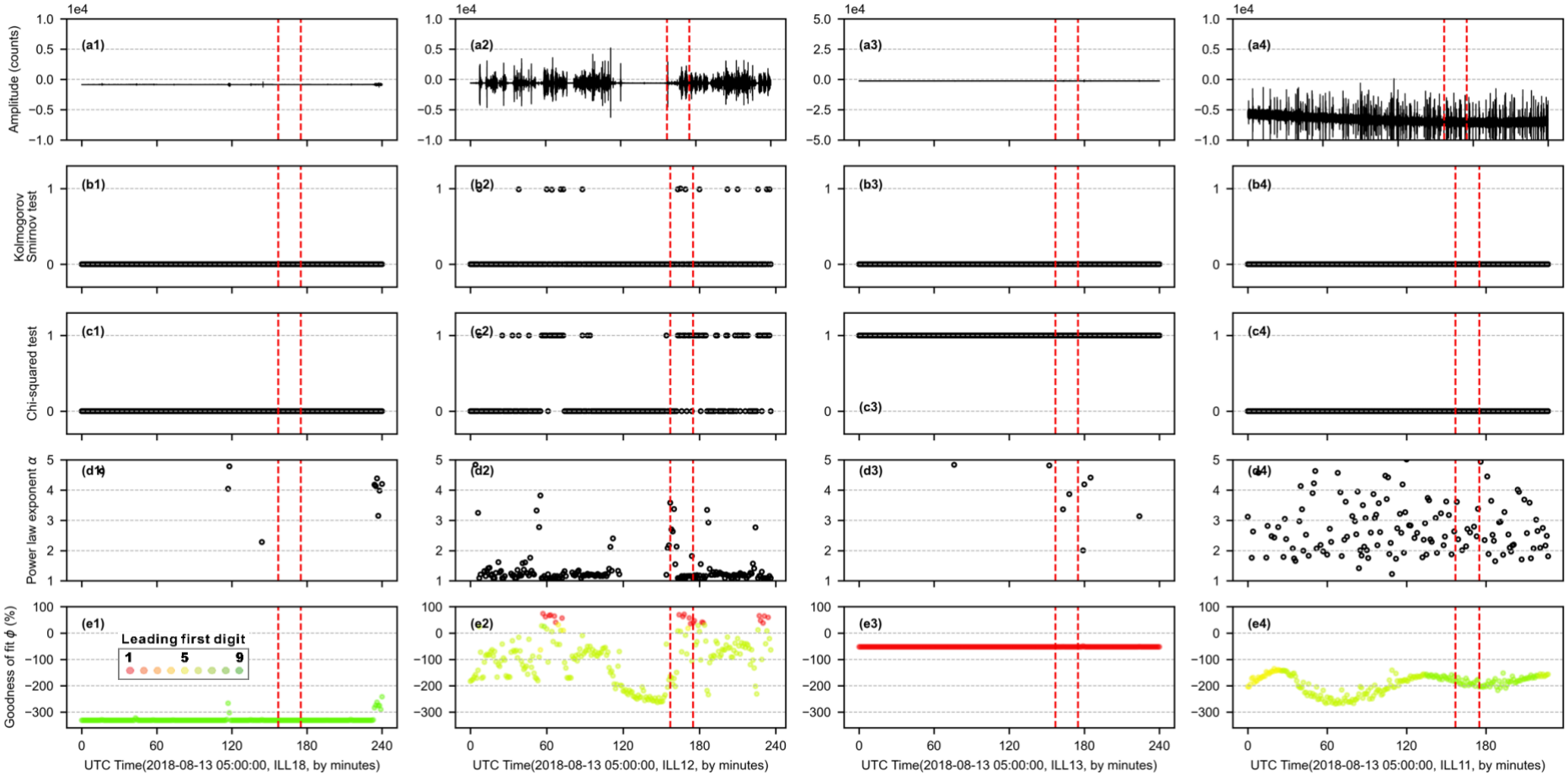


Figure S23. BL features of seismic signal generated by **Actual negative** class (2018-08-13, validation case). It was labeled as **Predicted positive** by our detector **(False positive)**. Raw waveform (row a), Chi-squared test (row b) and Kolmogorov-Smirnov test (row c) results of first-digit distribution (one means to follow BL), power-law exponent (row d), and goodness of fit φ (row e) of different seismic stations (column 1-4). The red dashed vertical lines are event start and time (2018-08-13 07:47:00, 08:06:00) of our detector marker based on ILL12. The upstream ILL18 and downstream ILL11 do not capture the same trend, this process could be caused by instruments noise or a local event.

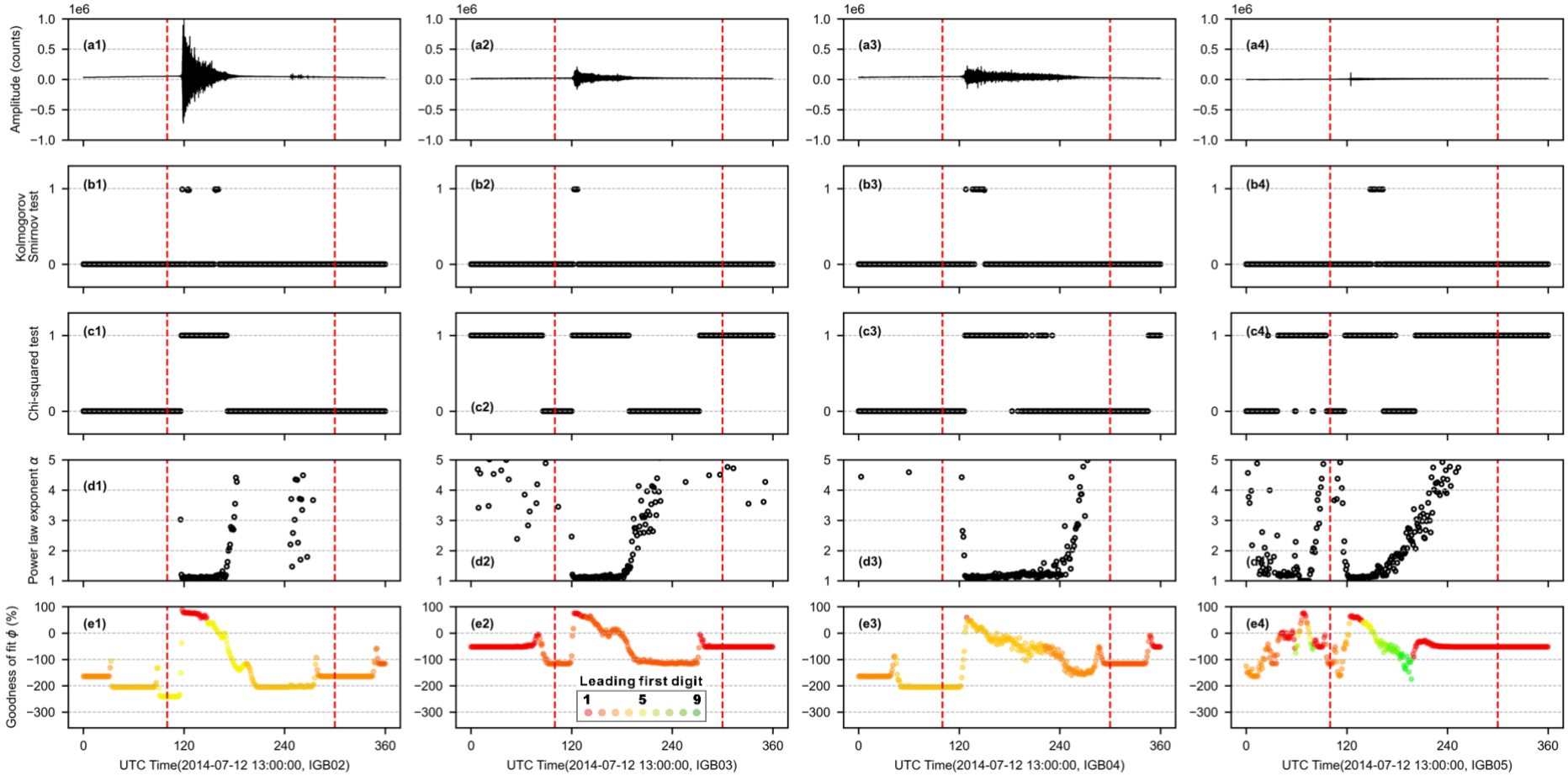


Figure S24. BL features of seismic signal generated by **Actual positive** (2014-07-12, training case). It was labeled as **Predicted positive** by our detector **(True positive)**. Raw waveform (row a), Chi-squared test (row b) and Kolmogorov-Smirnov test (row c) results of first-digit distribution (one means to follow BL), power-law exponent (row d), and goodness of fit φ (row e) of different seismic stations (column 1-4). The red dashed vertical lines are manually labeled start and end time (2014-07-12 14:40 and 18:00). IGB02 are closer to channel, and more than one station can detect this event through BL.

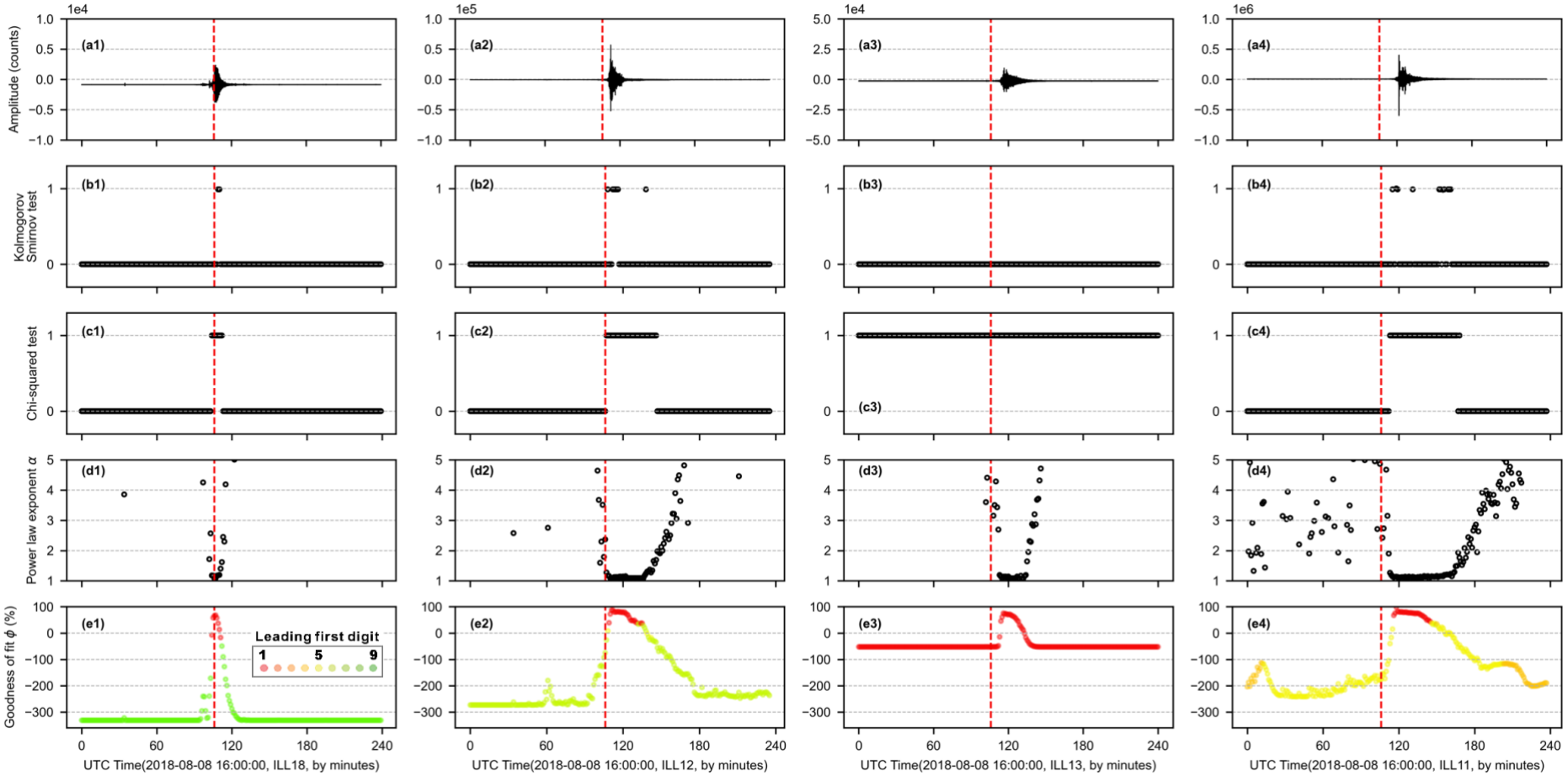


Figure S25. BL features of seismic signal generated by **Actual positive** (2018-08-08, validation case). It was labeled as **Predicted positive** by our detector **(True positive)**. Raw waveform (row a), Chi-squared test (row b) and Kolmogorov-Smirnov test (row c) results of first-digit distribution (one means to follow BL), power law exponent (row d), and goodness of fit φ (row e) of different seismic stations (column 1-4). The red dashed vertical line is debris flow front arrive at CD1 (2018-08-08 17:49:25). From upstream ILL18 to downstream ILL11, it is clearly seen that BL is captured sequentially in the time series, which indicates an event of movement along the trench.

Table S1. Seismometer Parameters

Year	Sensor	Data Logger	Signal Bandwidth	Sampling Frequency	Stations	Operated	Format
2013-2014	Trillium Compact TC120s	Omnirecs Cube 3ext with breakout box	1-100 Hz	200 Hz	IGB02	GFZ[1]	SAC
2017	Lennartz LE-3D/5S	Nanometrics Centaur	1-50 Hz	100 Hz	ILL02	WSL[2]	MiniSEED
2018-2019	Lennartz LE-3D/5S	Nanometrics Centaur	1-50 Hz	100 Hz	ILL12	WSL[2]	MiniSEED

[1]GFZ, German Research Centre for Geosciences;

[2]WSL, Swiss Federal Institute for Forest, Snow and Landscape Research.

Table S2. Training Events from 2013 to 2014

N	Source	Type	Start time (UTC)	End time (UTC)	SNR	Available Station	BL
N1	WSL	F	2013-05-03 07:45	2013-05-03 08:45	1.3	01, 06, 07, 10	
N2	WSL	DF	2013-08-08 10:30	2013-08-08 12:30	1.7	01, 04, 07, 08, 10	
N3	WSL	DF	2013-08-24 13:30	2013-08-24 15:00	1.4	01, 04, 07, 08, 10	
N4	WSL	DF	2013-09-08 19:00	2013-09-08 21:00	1.7	01, 04, 07, 08, 10	
1	WSL	DF	2013-07-22 16:30	2013-07-22 18:30	4.9	01, **02**, 10	Y
2	WSL	F	2013-07-29 07:30	2013-07-29 15:00	4.2	01, **02**, 03, 07, 08, 10	Y
3	GFZ	DF	2014-04-08 08:00	2014-04-08 18:00	42.4	**02**, 03, 05, 06, 07	Y
4	GFZ	DF	2014-04-26 22:30	2014-04-27 05:00	12.9	**02**, 03, 05, 06	Y
5	GFZ	DF	2014-04-29 09:00	2014-04-29 15:00	5.9	**02**, 05, 06, 07	Y
6	GFZ	DF	2014-04-30 10:00	2014-04-30 15:00	27.0	**02**, 05, 06, 07	Y
7	GFZ	DF	2014-05-02 03:00	2014-05-02 13:00	0.1	**02**, 03, 05, 06, 07	N
8	WSL	DF	2014-05-07 15:30	2014-05-07 18:00	15.2	**02**, 03, 05, 06, 07	Y
9	GFZ	DF	2014-05-23 03:00	2014-05-23 15:00	57.0	01, **02**, 04	Y
10	GFZ	DF	2014-05-24 09:00	2014-05-24 15:00	16.1	01, **02**, 04, 05	Y
11	GFZ	DF	2014-05-27 05:30	2014-05-27 16:00	22.2	01, **02**, 03, 04, 05	Y
12	GFZ	DF	2014-06-19 14:00	2014-06-19 16:00	6.2	01, **02**, 04, 05	Y
13	GFZ	DF	2014-06-23 18:00	2014-06-23 20:00	10.1	01, **02**, 04, 05	Y
14	WSL	DF	2014-07-08 02:00	2014-07-08 14:00	10.1	01, **02**, 03, 04, 05	Y
15	WSL	DF	2014-07-12 14:40	2014-07-12 18:00	20.6	01, **02**, 03, 04, 05	Y
16	WSL	DF	2014-07-20 20:00	2014-07-21 03:00	12.7	01, **02**, 03, 04, 05	Y
17	WSL	DF	2014-07-23 23:00	2014-07-24 05:00	65.6	01, **02**, 04	Y
18	WSL	DF	2014-07-28 16:00	2014-07-28 20:50	54.3	01, **02**, 03, 04, 05	Y
19	WSL	DF	2014-07-28 20:55	2014-07-29 06:00	19.5	01, **02**, 03, 04, 05	Y

20	GFZ	DF	2014-08-02 17:30	2014-08-02 21:30	45.7	01, **02**, 03, 04, 05	Y
21	GFZ	DF	2014-08-08 18:30	2014-08-08 21:00	16.8	01, **02**, 04, 05	Y
22	GFZ	DF	2014-08-11 02:00	2014-08-11 18:00	7.7	01, **02**, 03, 04, 05	Y
23	GFZ	DF	2014-08-13 04:30	2014-08-13 15:00	0.4	01, **02**, 04, 05	N
24	WSL	DF	2014-09-08 19:30	2014-09-08 23:00	175.6	01, **02**, 04, 05	Y

N: Event number (N1-N4 were non-used events because data from IGB02 is not available). **Type**: DF is debris flow, F is flood. **Start time and End time** are labeled manually based on data from the IGB02 station. The format is *yyyy-mm-dd hh:mm* format. **SNR**, signal-to-noise ratio (N1-N4 none-used event by IGB10, 1-24 used event by IGB02). **Available station**, IGB01 abbreviation is 01. The total duration of the 24 used events is approximately 166 hours. The calculating window size is 60 seconds with 12000 data points. **BL**, from the **Start time** to the **End time** of the event, if the output of any time window from Chi-squared and Kolmogorov-Smirnov tests accepts the null hypothesis, that event will be marked as Y (Yes, follow BL); otherwise, N (No, does not follow BL).

Table S3. Validation Events from 2017 to 2019

N	Type	CD1 Time (UTC)	Start Time (UTC)	End Time (UTC)	Volume (m^3)	Data Source	BL
25	DF	2017-05-19 11:41:00	2017-05-19 10:00	2017-05-19 13:00	No Data	No Data	No data
26	DF	2017-05-29 16:58:31	2017-05-29 15:00	2017-05-29 19:00	100000	ILL02	Y
27	DF	2017-06-03 20:23:07	2017-06-03 18:00	2017-06-03 22:30	No Data	ILL02	Y
28	DF	2017-06-03 23:27:38	2017-06-03 22:40	2017-06-04 02:00	25000	ILL02	Y
29	DF	2017-06-14 19:30:48	2017-06-14 18:00	2017-06-14 23:30	No Data	ILL02	Y
30	DF	2018-06-11 10:46:39	2018-06-11 09:00	2018-06-11 14:00	35000	ILL12	Y
31	DF	2018-06-12 18:29:16	2018-06-12 17:00	2018-06-12 22:00	No Data	ILL12	Y
32	DF	2018-07-25 16:56:40	2018-07-25 15:00	2018-07-25 20:00	<50000	ILL12	Y
33	DF	2018-08-08 17:49:25	2018-08-08 16:00	2018-08-08 20:00	<100000	ILL12	Y
34	DF	2019-06-10 17:02:51	2019-06-10 16:00	2019-06-10 20:00	3300	ILL12	N
35	DF	2019-06-10 22:01:17	2019-06-10 21:00	2019-06-11 03:00	6600	ILL12	Y
36	DF	2019-06-20 09:12:17	2019-06-20 07:00	2019-06-20 11:00	No Data	ILL12	N
37	DF	2019-06-21 19:34:42	2019-06-21 19:00	2019-06-21 22:00	83000	ILL12	Y
38	DF	2019-07-01 23:00:29	2019-07-01 22:00	2019-07-02 04:00	78000	ILL12	Y
39	DF	2019-07-02 22:09:28	2019-07-02 21:00	2019-07-02 23:30	39000	ILL12	Y
40	DF	2019-07-03 16:43:15	2019-07-03 15:00	2019-07-03 20:00	No Data	ILL12	N
41	DF	2019-07-15 03:40:21	2019-07-15 02:00	2019-07-15 06:00	16000	ILL12	N
42	DF	2019-07-26 17:33:12	2019-07-26 16:30	2019-07-26 19:30	64000	ILL12	Y
43	DF	2019-08-11 17:02:34	2019-08-11 16:00	2019-08-11 19:00	53000	ILL12	Y
44	DF	2019-08-20 16:40:59	2019-08-20 15:00	2019-08-20 18:00	13000	ILL12	Y
45	DF	2019-10-09 11:45:28	2019-10-09 10:30	2019-10-09 13:30	No Data	ILL12	Y

46	DF	2019-10-15 16:10:50	2019-10-15 15:00	2019-10-15 23:00	No Data	ILL12	N

N: Event number. **CD1** time is the arrival time at CD1 and **Volume** is the integrated sum of discharge over the entire debris-flow wave (Chmiel et al., 2021). **The start time and End time** are labeled manually from ILL02 or ILL12 station. The format is *yyyy-mm-dd hh:mm* format. The calculating window size is 60 seconds with 6000 data points. **BL**, same as training events, Y (Yes, follow BL); N (No, does not follow BL).

Table S4. Signals Generated by Different Mass Movements and Fluvial Processes

N	Type	Location	Date (UTC)	Station	Sensor	Data Logger	Sampling Frequency	Window Size	Ref	BL
47	Rockfall 1	Illgraben, Switzerland	2013-07-22 16:24	IGB01	Trillium Compact TC120s	Omnirecs Cube 3ext with breakout box	200 Hz	10 s	Burtin et al., 2016	N
48	Rockfall 2 Event 30 in ref	Lauterbrunnen, Switzerland	2015-04-06 13:23	Funny Rain-TC120s-Cube3extBOB	Trillium Compact TC120s	Omnirecs Cube 3ext with breakout box	200 Hz	10 s	Dietze, Mohadjer, et al., 2017; Dietze, Turowski, et al., 2017	N
49	Landslide	Askja, Iceland	2014-07-21 23:24	MOFO	Güralp CMG-6TD	Nanometrics Taurus	100 Hz	60 s	Schöpa et al., 2018	Y
50	Hurricane-induced Lahar	Volcán de Colima, Meixco	2015-10-24 01:00	SHK2	Sercel SG-10	No info.	250 Hz	60 s	Capra et al., 2018	Y
51	Glacial-lake-outburst flood	Bhotekoshi, Nepal	2016-07-05 13:30	Hindi (NEP08)	Trillium Compact TC120s	Omnirecs Cube 3ext with breakout box	200 Hz	60 s	Cook et al., 2018	Y
52	Bedload transport	Liwu, Hualien	2021-10-12 01:05	TA64	Sensor Nederland PE-6/B	Omnirecs Cube 3ext	100 Hz	60 s		N

BL, from the **Start time** to the **End time** of the event, if the output of any time window from Chi-squared and Kolmogorov-Smirnov tests accepts the null hypothesis, that event will be marked as Y (Yes, follow BL); otherwise, N (No, does not follow BL).

Table S5. Coefficients of the Exponential Fitting

N	$\varphi_{optimal}$	Time (UTC)	Data Length	R^2	*a*	*b*	*c*	*α*	BL
1	36.62	2013-07-22 16:52:00	2	0.767	8.776E+02	2.885E-02	2.602E+04	1.10	Y
2	20.60	2013-07-29 07:35:00	5	0.899	2.600E+01	2.209E-02	1.061E+03	1.14	Y
3	77.97	2014-04-08 13:34:00	5	0.246	-2.902E+05	-1.319E-05	2.905E+05	1.00	Y
4	81.90	2014-04-26 23:58:00	1	0.771	4.071E+00	6.973E-02	7.531E+03	1.11	Y
5	81.76	2014-04-29 11:00:00	4	0.808	8.628E+01	1.649E-02	1.091E+03	1.00	Y
6	79.82	2014-04-30 12:44:00	5	0.052	7.613E-02	7.834E-02	4.425E+02	1.00	Y
7	-36.38	2014-05-02 08:00:00	4	0.771	2.639E-08	8.310E-02	1.980E+02	5.22	N
8	80.78	2014-05-07 15:53:00	5	0.940	1.690E+02	1.626E-02	-3.612E+02	1.12	Y
9	84.18	2014-05-23 04:42:00	5	0.920	7.927E+01	2.083E-02	1.217E+03	1.11	Y
10	84.98	2014-05-24 09:36:00	3	0.171	3.535E-54	5.886E-01	1.561E+04	1.11	Y
11	78.36	2014-05-27 13:01:00	2	0.747	9.090E-21	1.262E+01	4.595E+02	1.00	Y
12	70.16	2014-06-19 14:25:00	1	0.692	3.480E+02	5.768E-02	2.879E+04	1.00	Y
13	74.50	2014-06-23 18:12:00	3	0.736	7.486E+02	1.854E-02	1.370E+04	1.12	Y
14	80.25	2014-07-08 10:13:00	2	0.618	6.075E+07	3.890E-06	-6.072E+07	1.10	Y
15	87.97	2014-07-12 14:58:00	5	0.944	2.274E+00	3.252E-02	1.494E+03	1.10	Y
16	79.11	2014-07-20 20:45:00	5	0.905	3.962E+03	5.455E-03	2.568E+03	1.12	Y
17	93.10	2014-07-23 23:03:00	5	0.868	1.573E-01	3.858E-02	1.902E+03	1.10	Y
18	87.46	2014-07-28 17:09:00	1	0.838	1.775E+00	9.960E-02	6.508E+04	1.09	Y
19	80.08	2014-07-29 01:00:00	2	0.904	3.577E+02	3.034E-02	2.800E+04	1.10	Y
20	82.17	2014-08-02 19:17:00	5	0.882	4.056E+02	2.646E-02	6.460E+03	1.10	Y
21	79.50	2014-08-08 19:18:00	1	0.832	3.044E+00	8.648E-02	2.808E+04	1.10	Y
22	82.95	2014-08-11 17:02:00	2	0.930	3.404E-13	9.804E-01	7.724E+02	1.15	Y
23	-4.72	2014-08-13 10:30:00	3	0.199	4.305E-14	4.446E-01	2.564E+02	32.77	N
24	89.14	2014-09-08 19:58:00	5	0.842	9.514E-01	3.641E-02	3.316E+03	1.11	Y
25	No data	2017-05-19							
26	92.75	2017-05-29 17:00:00	5	0.894	1.125E-02	3.508E-02	1.322E+02	1.00	Y
27	80.60	2017-06-03 20:30:00	5	0.794	6.382E+01	5.378E-03	2.322E+01	1.00	Y
28	86.79	2017-06-03 23:30:00	5	0.814	3.505E-01	2.558E-02	1.594E+02	1.14	Y

29	88.78	2017-06-14 19:32:00	5	0.947	1.507E-01	2.829E-02	1.659E+02	1.14	Y
30	77.82	2018-06-11 11:00:00	5	0.924	1.257E+02	6.730E-03	-1.350E+01	1.00	Y
31	73.20	2018-06-12 18:46:00	5	0.439	4.804E+05	1.561E-06	-4.800E+05	1.00	Y
32	83.37	2018-07-25 16:59:00	5	0.930	6.260E+00	1.854E-02	8.504E+01	1.00	Y
33	87.17	2018-08-08 17:52:00	5	0.910	7.524E+00	2.069E-02	2.320E+02	1.00	Y
34	68.99	2019-06-10 17:30:00	3	0.290	2.464E+01	1.999E-02	1.725E+03	1.13	N
35	77.29	2019-06-10 22:54:00	5	0.918	6.911E+01	1.106E-02	6.430E+02	1.14	Y
36	72.41	2019-06-20 09:23:00	5	0.840	2.176E+03	3.061E-03	-2.148E+03	1.15	N
37	80.18	2019-06-21 19:39:00	5	0.878	6.205E+03	2.586E-03	-6.752E+03	1.13	Y
38	80.54	2019-07-01 23:35:00	4	0.527	9.053E+02	6.602E-03	2.048E+03	1.13	Y
39	75.99	2019-07-02 22:16:00	5	0.847	6.091E+06	1.954E-06	-6.091E+06	1.13	Y
40	48.59	2019-07-03 17:08:00	3	0.511	6.445E+00	2.704E-02	9.103E+02	1.00	N
41	67.80	2019-07-15 03:49:00	5	0.854	9.492E+05	9.374E-06	-9.491E+05	1.13	N
42	81.66	2019-07-26 17:38:00	5	0.905	1.367E+02	1.071E-02	-2.983E+01	1.13	Y
43	87.39	2019-08-11 17:04:00	5	0.899	8.506E+00	2.044E-02	1.775E+02	1.12	Y
44	79.34	2019-08-20 16:49:00	5	0.894	3.303E+02	9.328E-03	2.485E+02	1.18	Y
45	74.21	2019-10-09 11:57:00	5	0.915	1.345E+02	1.019E-02	5.004E+02	1.00	Y
46	-24.42	2019-10-15 16:19:00	5	0.888	1.275E+02	6.050E-03	1.586E+01	1.22	N

N: Event number. $\boldsymbol{\varphi_{optimal}}$ is the optimal goodness of fit between manually labeled start and end time (unit by %). **Time** corresponds to the moment when $\boldsymbol{\varphi}$ is at its optimal. **Data length** is the length of the data used to fit the exponential function (before φ_{optima}, in minutes). $\mathbf{R^2}$ is the coefficient of determination of exponential fitting. $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are the coefficients of the exponential fitting in $S(t) = a \times e^{bt} + c$ (Text S3). $\boldsymbol{\alpha}$ is the power law exponent at optimal goodness of fit time.

Code S1.

All codes are available in
https://github.com/Nedasd/Benfords-law-in-environmental-seismology.git

Reference

Belli, G., Walter, F., McArdell, B., Gheri, D., & Marchetti, E. (2022). Infrasonic and Seismic Analysis of Debris-Flow Events at Illgraben (Switzerland): Relating Signal Features to Flow Parameters and to the Seismo-Acoustic Source Mechanism. *Journal of Geophysical Research: Earth Surface*, *127*(6), 1–20. https://doi.org/10.1029/2021jf006576

Burtin, A., Hovius, N., McArdell, B. W., Turowski, J. M., & Vergne, J. (2014). Seismic constraints on dynamic links between geomorphic processes and routing of sediment in a steep mountain catchment. *Earth Surface Dynamics*, *2*(1), 21–33. https://doi.org/10.5194/esurf-2-21-2014

Burtin, A., Hovius, N., & Turowski, J. M. (2016). Seismic monitoring of torrential and fluvial processes. *Earth Surface Dynamics*, *4*(2), 285–307. https://doi.org/10.5194/esurf-4-285-2016

Capra, L., Coviello, V., Borselli, L., Márquez-Ramírez, V. H., & Arámbula-Mendoza, R. (2018). Hydrological control of large hurricane-induced lahars: Evidence from rainfall-runoff modeling, seismic and video monitoring. *Natural Hazards and Earth System Sciences*, *18*(3), 781–794. https://doi.org/10.5194/nhess-18-781-2018

Chmiel, M., Walter, F., Wenner, M., Zhang, Z., McArdell, B. W., & Hibert, C. (2021). Machine Learning Improves Debris Flow Warning. *Geophysical Research Letters*, *48*(3), 1–11. https://doi.org/10.1029/2020GL090874

Cook, K. L., Andermann, C., Gimbert, F., Adhikari, B. R., & Hovius, N. (2018). Glacial lake outburst floods as drivers of fluvial erosion in the Himalaya. *Science*, *362*(6410), 53–57. https://doi.org/10.1126/science.aat4981

Dietze, M. (2018). The R package "eseis"-a software toolbox for environmental seismology. *Earth Surface Dynamics*, *6*(3), 669–686. https://doi.org/10.5194/esurf-6-669-2018

Dietze, M., Mohadjer, S., Turowski, J. M., Ehlers, T. A., & Hovius, N. (2017). Seismic monitoring of small alpine rockfalls-validity, precision and limitations. *Earth Surface Dynamics*, *5*(4), 653–668. https://doi.org/10.5194/esurf-5-653-2017

Dietze, M., Turowski, J. M., Cook, K. L., & Hovius, N. (2017). Spatiotemporal patterns, triggers and anatomies of seismically detected rockfalls. *Earth Surface Dynamics*, *5*(4), 757–779. https://doi.org/10.5194/esurf-5-757-2017

Schöpa, A., Chao, W. A., Lipovsky, B. P., Hovius, N., White, R. S., Green, R. G., & Turowski, J. M. (2018). Dynamics of the Askja caldera July 2014 landslide, Iceland, from seismic signal analysis: Precursor, motion and aftermath. *Earth Surface Dynamics*, *6*(2), 467–485. https://doi.org/10.5194/esurf-6-467-2018