An approach for prospective forecasting of rock slope failure time

Leinauer, Johannes ; Weber, Samuel ; Cicoira, Alessandro ; Beutel, Jan ; Krautblatter, Michael

Abstract: Rock slope failures globally account for most single-event landslide disasters. Climatic changes in mountain areas boost failure activity and the demand for reliable failure time forecasts. State-of-the-art prediction models are often confused with high-frequency slope deformation data. Prospectively, they provide ambiguous forecasts as data filtering, starting point definition and forecast uncertainty remain arbitrary. Here, we develop a prospective failure time forecast model that applies multiple filtering and inverse velocity percentiles to minimize subjective decisions. We test the concept with 14 historic slope failures of 10^2-10^8 m^3 including 46 displacement datasets from different sensors. After automatic detection of the onset of acceleration, the failure time of all events is forecasted to within −1 ± 17 h for higher-frequency data and −1 ± 4 d for daily data with a final mean uncertainty of 1 ± 1 d and 7 ± 4 d that is estimated in real-time. This prospective approach overcomes previous long-standing problems by introducing a robust and uniform concept across various types of catastrophic slope failures and sensors.

ARTICLE

Check for updates

An approach for prospective forecasting of rock slope failure time

Johannes Leinauer ● [1✉], Samuel Weber ● [1,2,3], Alessandro Cicoira ● [4], Jan Beutel ● [5] & Michael Krautblatter ● [1]

Rock slope failures globally account for most single-event landslide disasters. Climatic changes in mountain areas boost failure activity and the demand for reliable failure time forecasts. State-of-the-art prediction models are often confused with high-frequency slope deformation data. Prospectively, they provide ambiguous forecasts as data filtering, starting point definition and forecast uncertainty remain arbitrary. Here, we develop a prospective failure time forecast model that applies multiple filtering and inverse velocity percentiles to minimize subjective decisions. We test the concept with 14 historic slope failures of 10^2-10^8 m^3 including 46 displacement datasets from different sensors. After automatic detection of the onset of acceleration, the failure time of all events is forecasted to within -1 ± 17 h for higher-frequency data and -1 ± 4 d for daily data with a final mean uncertainty of 1 ± 1 d and 7 ± 4 d that is estimated in real-time. This prospective approach overcomes previous long-standing problems by introducing a robust and uniform concept across various types of catastrophic slope failures and sensors.

[1] Landslide Research Group, TUM School of Engineering and Design, Technical University of Munich, Munich, Germany. [2] WSL Institute for Snow and Avalanche Research SLF, Davos Dorf, Switzerland. [3] Climate Change, Extremes and Natural Hazards in Alpine Regions Research Center, CERC, Davos Dorf, Switzerland. [4] Department of Geography, University of Zurich, Zurich, Switzerland. [5] Department of Computer Science, University of Innsbruck, Innsbruck, Austria. ✉email: johannes.leinauer@tum.de

Slope instabilities of varying magnitude frequently threaten the life, health or property of humans in all parts of the world[1–4]. Massive rock slope failures alone, including rock slides, rock avalanches and rock falls caused 58% of the global disasters from single landslide events (downslope mass movements in general) in the last millennium[5]. Under the current ongoing climatic changes, landslide risk is expected to increase critically[6] and therefore, cost-efficient mitigation measures with a wide applicability are needed. Such measures as alarm systems or evacuation of endangered zones require reliable forecasts of the imminent event[7,8]. Local monitoring systems provide important data that can be used to predict the time of slope failures[9]. However, to meet social, economic and scientific requirements, existing forecasting methods need amendments in the real-time processing of modern monitoring data and a conceptual basis that is valid for diverse failure processes.

All slope failures are preceded by precursory deformations of particular amplitude and duration[10]. The precursory signals, increasing deformation rates or seismic activity, originate within the unstable rock masses due to progressive crack propagation and non-elastic deformations prior to failure[11]. The measured prefailure surface displacement pattern can be described with a classical creep curve[9] (Fig. 1 top). In the secondary creep phase, under constant stress conditions, stabilizing and destabilizing forces balance each other and strain rates remain constant. Increasing stresses in the fatiguing material create a self-accelerating positive feedback setting of progressively failing retaining sectors[12–15]. As a result, in the final accelerating phase, the strain rate is in a power-law relation to time, fitting tertiary creep. Usually, the acceleration phase is shorter the smaller the volume and the more brittle the rock is[16], ranging from

seconds for small blocks up to days, weeks or months for bigger slope failures. In the accelerating phase, more or less accurate failure time forecasts are so far calculated with several methods[9,17–22]. The most common prediction method is the inverse velocity method after Fukuzono[19], possibly due to its relatively simple graphical solution[23–25]. Here, the inverse velocity of a failing mass is plotted against time during tertiary creep and the intersection of the extrapolated trend with the abscissa, i.e. at infinite velocity, indicates the predicted time of failure (Fig. 1 bottom).

While it can be easy, to find a linear inverse velocity fit that matches the actual time of a failure retrospectively post-failure[7–9,26–30], the challenge in real-time scenarios is to obtain reliable forecasts prospectively[31,32]. Modern monitoring systems are able to measure surface displacements in (near) real-time at high spatial and temporal resolution by in situ devices (e.g. crackmeter, tiltmeter, Global Navigation Satellite System), or by satellite or ground-based remote sensing techniques[29,33]. These data types create early warning opportunities, but present prediction models have not kept pace and are not ready to be used with high-frequency and therefore often noisy slope deformation data. Thus, they provide ambiguous forecasts.

Specifically, data filtering, starting point definition and forecast uncertainty remain unclear. Filtering the raw monitoring data can improve the forecasts, but the process is often arbitrary and involves subjective decisions[23,24,34,35]. Although the forecasted failure time is sensitive to the degree of data smoothing, most applications of the inverse velocity method in the literature include only one or two subjectively selected time windows. Furthermore, the outcome of the prediction model is highly sensitive to the definition of the transition point from secondary
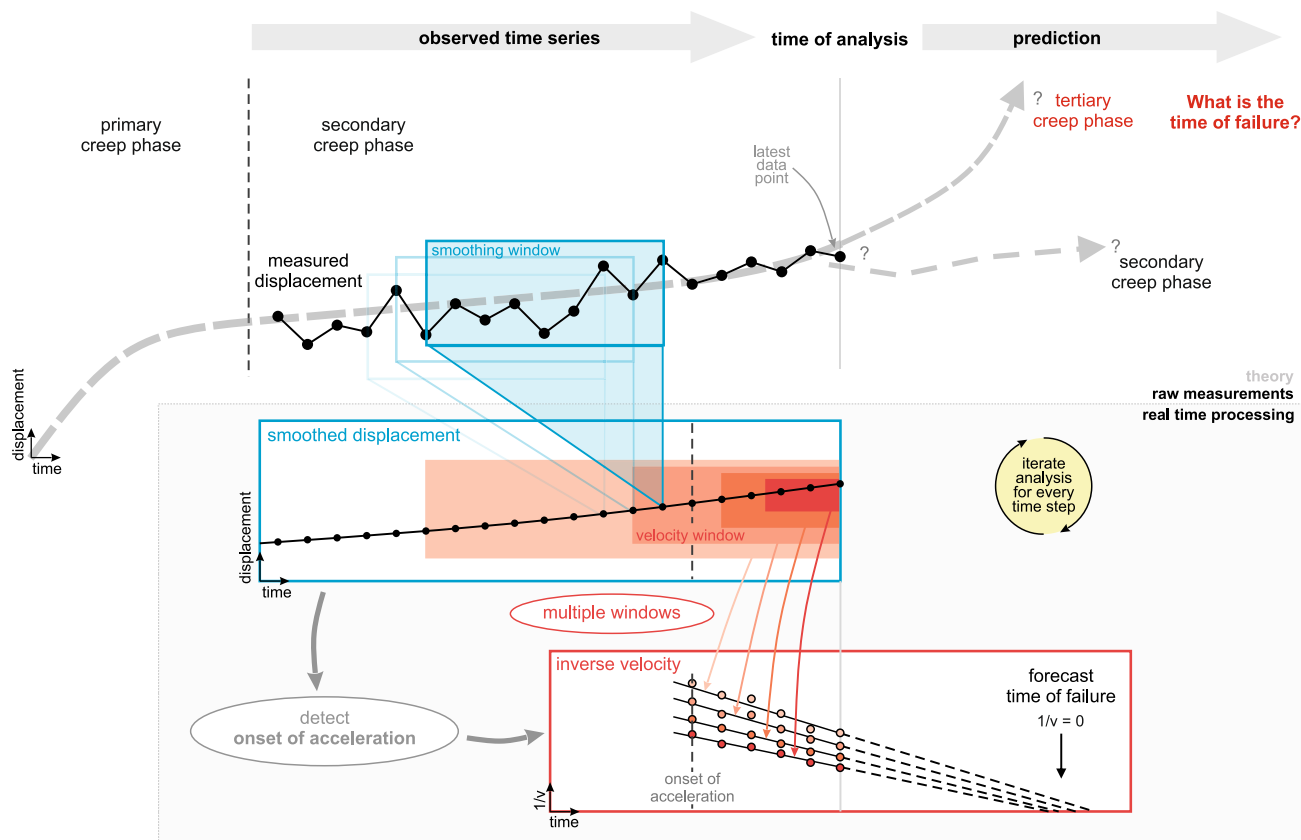
Fig. 1 Concept of the prospective failure time forecast model (PFTF). Displacements of unstable rock slopes follow a creep curve. Usually, monitoring starts in the secondary creep phase without knowing the transition point to tertiary creep (onset of acceleration, OOA). This point is used as starting point for failure time forecasting. The real time processing of PFTF uses multiple filtering windows to identify the OOA and to derive multiple inverse velocities that are used to forecast the time of failure and estimate the forecast uncertainty (for details see Methods).

to tertiary creep (from constant velocity to accelerating regime). This point is called onset of acceleration (OOA) and is used as starting point for forecasting (cp. Fig. 1)[31,34]. Monitoring usually starts in the secondary creep phase and in real-time it remains challenging to decide at what time the slope enters the tertiary creep phase (Fig. 1 top). An unfavourable OOA setting can result in misleading forecasts as the regression of inverse velocities may include points, which do not belong to the final failure phase in case of a too early OOA or may exclude important measurements in case of a too late OOA. Modern early warning systems deliver observations of multiple sensors at high temporal frequency. A prospective real-time application of forecasting methods therefore requires a predominantly automatic method for the starting point definition. However, the few recently published approaches for this[24,35,36] perform well in some cases but also detect non-optimal OOAs or false positives. Moreover, responsible decision-makers are dependent on a measure of reliability and uncertainty of the derived forecasts. In the past, the uncertainty has often not been evaluated e.g.[7,37]. More recently, uncertainty estimations via the distribution of inverse velocity regressions[38,39], via the calculation of a standard error[30] or via the calculation of a failure window during which a failure event is highly probable[24] have been put forward. However, there is no standard procedure for a statistical real-time uncertainty estimation.

Therefore, there is a lack of a forecasting concept that operates fully prospectively without post-tuning of the results, with a clear smoothing process, a non-arbitrary OOA detection and a systematic real-time uncertainty estimation. Here, we present a prospective failure time forecast model (PFTF) developed by analysing 46 data-sets from 14 slope failures with daily or higher observation frequency. We acquire a uniform robust concept across sensor and failure types overcoming previous major drawbacks. The onset of acceleration is detected sufficiently early and reliable forecasts, including uncertainty estimation can automatically be calculated in real-time. Overall, forecast accuracy is higher with higher-frequency data and closer towards the actual time of failure. This model can support decision-makers in various critical situations.

Conceptual framework

Our forecasting concept (Fig. 1) includes iterative real-time data processing, a two-step filtering process with multiple filtering window lengths (blue and red in Fig. 1), an automatic starting point definition and a statistical uncertainty estimation of every forecast. We implemented the concept into an operational prospective failure time forecast model (PFTF) complying with the requirements for real-time failure forecasting. PFTF iterates with every new available data point[24,25], including only past observations without knowing the time of failure. It exclusively uses preaccessible data in a forward model. Therefore, the open-source algorithm design (Fig. 2a) allows real-time prospective forecasting of imminent failures but also a prospective analysis of past events (simulated real-time). It is capable of processing multiple high-frequency observations (for details see Methods).

The first key element of PFTF is a two-step smoothing approach with multiple window lengths (see Prospective real-time forecasting with multiple data), as this enables the automatic detection of the onset of acceleration (OOA) (see Automatic onset of acceleration detection) and the statistical uncertainty estimation (see Converging forecasts). In the first filtering step, the displacement trend is clarified via a rolling mean of raw displacements over specific smoothing windows (w_smooth). In the second step, a variety of inverse velocities is calculated from each displacement data via linear regressions over multiple velocity windows (w_v, see Methods). The amount of data points used to derive the velocity depends on the length of the velocity windows (Fig. 1 red boxes and arrows).

Secondly, we implement an automatic OOA detection method that is based on four simple criteria and includes all data points within the latest smoothing window (w_smooth, Fig. 2b, for details see Methods). Criterion 1 requires an increasing displacement rate. Criterion 2 requires decreasing inverse velocities for all velocity windows (w_v), which is different from criterion 1 as the multiple inverse velocities can differ. Criterion 3 requires an overall accelerating regime, represented by a decreasing 50% inverse velocity quantile. Criterion 4 requires a short-term accelerating trend, detected by inverse velocities smaller than the 1% quantile, meaning that the latest velocity is greater than 99% of all velocities before. If all criteria are true over one complete smoothing window, the OOA is detected. This method is operating automatically within the algorithm.

Once the OOA has been detected, an expected time of failure is calculated with all inverse velocities after the OOA with the linear version ($\alpha = 2$) of Fukuzono's method[19]. We evaluate the forecast results based on consistently updated life expectancy plots where the predicted time until failure is plotted against the time when the forecast is made[24,34] with additional real-time statistical information (see Supplementary Fig. 1a and details in Methods. To give a real-time statistical uncertainty information for all diverse test sites, we enhanced the failure window approach[24] by using multiple smoothing windows simultaneously and adding the mean forecasted time to failure. The width of the failure window and the consistency of its limits are useful indicators for the reliability of the forecasts. Additionally, we show important updated statistical information in the boxplots (Supplementary Fig. 1b).

Results and discussion

Prospective real-time forecasting with multiple data filtering. The introduced concept allows the calculation of prospective failure time forecasts with divers rock slope monitoring data. With the presented multiple smoothing approach, we can diminish primary instrumental noise on the raw data and include all observed displacement data into the inverse velocity determination. The stability of the forecasts is improved without losing the sensitivity to changing trends[35] and the effects of various window lengths can directly be analyzed in the resulting plots in real-time (see Supplementary Fig. 1).

Typically, the window lengths vary between few hours and multiple days, but reasonable window lengths always depend on the following factors:

- signal-to-noise ratio/ quality of the data; controlled by

  - deformation rates (signal)
  - measurement technique (noise)
  - measurement frequency

- expected duration of precursory accelerations and the time needed for emergency actions (smoothing windows should be shorter than that)
- amplitude of other natural noise producing divergence of the measured data from a perfect linear inverse velocity trend[24].

Instabilities with high movement rates usually require shorter smoothing windows, while slow-moving instabilities and noisy data require stronger smoothing[24]. Also, higher measurement frequencies require filtering over a greater number of observations as more background noise is recorded. In contrast, noise has less influence when a sparse monitoring frequency (e.g. 1 d) is applied[24]. This might also be the reason for the increasing prediction accuracy with greater measurement intervals reported by Iwata & Sasahara (2021)[40]. However, the analysis of the sites Galterengraben, La Saxe, Veslemannen and Weissmiess from this
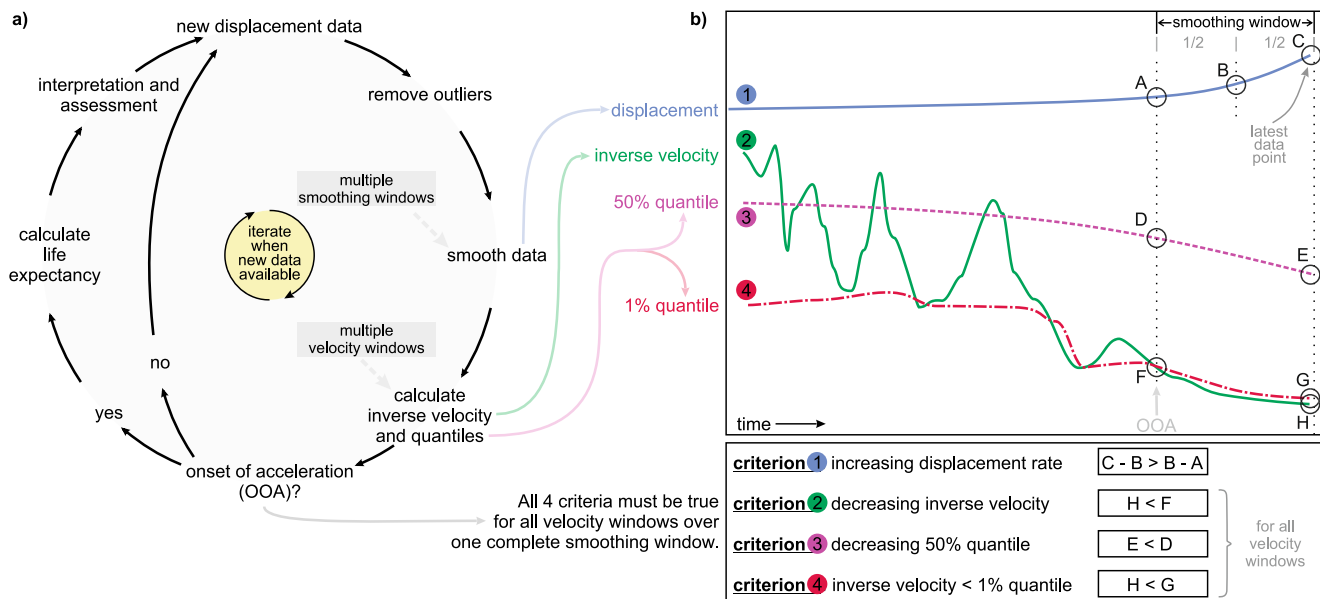
Fig. 2 Flowchart of the PFTF model and OOA detection criteria. a Flowchart of the PFTF model steps (for explanation see Methods). **b** Criteria for the automatic detection of the onset of acceleration (OOA) that is used as starting point for forecasting (for details see Methods).

study show that if the monitoring frequency is too low, precursory accelerations can be over-seen or too few data points exist to perform reliable forecasting. This is most relevant in geomechanically strong rock slopes consisting of brittle hard rock where the acceleration phase can be quite short[24]. Therefore, we identified a high monitoring frequency as a key factor, especially in situations with short acceleration phases[35]. Hence, we suggest using a sufficiently high sampling frequency, e.g. daily or faster for volumes $> 10^6 \, m^3$ and hourly or faster for smaller volumes, in combination with adequate data filtering, depending on frequency and signal-to-noise ratio (cp. Supplementary Fig. 2). Many monitoring methods allow higher sampling frequencies without additional costs or effort. Although longer smoothing windows can diminish the uncertainty of the forecasts, excessive smoothing using too long smoothing windows distorts the calculations and adds a lag to the time series (compare Converging forecasts). As a result, forecasts are shifted towards unsafe predictions (forecast after failure) which should be avoided. Therefore, the smoothing windows should be as short as possible but as long as necessary.

Due to the specific complexity of each individual slope instability, it is impossible to define a universal rule for optimal window lengths at this stage. However, for the sake of comparability between the sites in this study, we use three sets of fixed window lengths depending on the data resolution and the expected acceleration phase duration (see Supplementary Fig. 2). In real-time cases, a detailed knowledge of the unstable slope, its geological process and the site-specific data quality is needed. The operating specialist has then to decide on reasonable window lengths and might even calculate with multiple windows simultaneously, although the interpretation of the results can become more complex. Also, the smoothing process is not necessarily fixed over all time. In cases where conditions change, e.g. when decelerations appear after the OOA but failure is assumed imminent, the smoothing windows might need adaption (see Supplementary Fig. 2 and the sites La Saxe and Veslemannen).

Automatic onset of acceleration detection. The automatic determination of the OOA minimizes the subjectivity of critical real-time decisions under the pressure of imminent failure and could also be combined with previously published methods.

Our approach is robust for all 14 tested sites in this study (see Supplementary table and details in Converging forecasts). The OOA is prospectively detected sufficiently early (0.1-200 d prior the actual time of failure) while the timing of the OOA is dependent on failure volume and process (see Converging forecasts). Analogue to the smoothing windows, a restart of the OOA detection analysis is necessary in case of trend updates.

False alarms should always be kept to a minimum, but can never be ruled out due to periodic accelerations and natural variations. From the 14 tested sites, the PFTF model detects false OOAs in two cases (La Saxe and Veslemannen). A typical life expectancy plot of such a setting is shown in Supplementary Fig. 3. A false detection of the OOA can be identified as such as soon as a consistent deceleration with increasing life expectancies appears. Then, a restart of the analysis only with data after the deceleration phase gives the best results. It is worth noting that a false OOA detection is not necessarily leading to a false alarm immediately. After evaluation of the forecast results by the early warning system operators, either a manual intervention in case of a false OOA or the triggering of an alarm in case of a threshold exceedance are possible. On the other hand, if no clear OOA point can be detected sufficiently early before the actual failure, the model fails leading to a missed alarm. This is the case in 4% of the analyses. All difficulties leading to a bad or non-representative signal-to-noise ratio (noise fluctuations in the order of the acceleration signal hiding or substantially distorting it) can contribute to such a case. Then, an improvement of the monitoring quality and smoothing window lengths, the use of different methods, and a better understanding of the instability's process dynamics can assist.

Converging forecasts. Our real-time uncertainty estimation, enabled by the usage of multiple windows, is expressed by the plotting of failure window and mean in the life expectancy plot. A converging failure window corresponds to an increasing confidence of the forecasts. In operating systems, predefined alarm limits can be graphically drawn as horizontal lines (e.g., at a life expectancy of 24 h). Different warning levels or required mitigation measures could be triggered when the failure window or the mean reaches specific alarm thresholds. The site-individual

alarm time before failure is dependent on the required time for mitigation measures like evacuation and the desired level of safety.

Additionally, with the updated boxplot, the reader can easily evaluate the range of all past forecasts and the influence of the applied smoothing. With ongoing updated iterations, anomalous forecasts, mainly appearing before the forecasts converge, turn into outliers. The over-plotting of the latest forecasted time of failure per velocity window (w_v, red diamonds) is helpful when evaluating the latest calculations and trends. Unstable and inconsistent forecasts can be detected. A stable position of the marker above the box represents a trend to a later expected time of failure and vice versa. Although the boxplots must be assessed with care, they present more valuable information than just the life expectancy plots usually presented by most authors, as the most recent forecast is not necessarily the most accurate one[8].

All tested sites show a similar pattern of converging forecasts towards failure. As an example of the achieved results, we present a time lapse video animation of the Preonzo site (Reflector 9) showing how the PFTF result appears to a responsible operator or decision-maker in real-time in the Supplementary Movie. Figure 3 displays the final three weeks before failure of the same data and boxplots at two paradigmatic time steps close to the OOA and close to failure. The resulting PFTF outcome can be divided into three phases. In Phase 1 shortly after the OOA, the calculated life expectancies (Fig. 3d) are scattered. The mean is unstable with no clear trend 45° downward and the failure window is wide. Subfigure 3e shows a corresponding boxplot from April 29. The boxes and the failure window cover a wide time range and the latest forecasts (red diamonds) are spread over 10 days. In this phase, the uncertainty is rated high. In the second phase (point A in Fig. 3d), the forecasts converge and the failure window becomes narrower. The exact position of this point is dependent on the degree of smoothing and the position of the OOA. In the last phase of 6-10 days prior to failure, the life expectancies of all w_v lie within a short time span, the mean plots stable 45° downward, and the failure window becomes small (ca. 2 d). Subfigure 3f shows a corresponding boxplot from May 12. The boxes and the failure window cover a narrow time range and the latest forecasts are consistent. In this phase the uncertainty is rated low. In fact, the mean forecasted time of failure is 1 day too early, while the failure on May 15 lies within the failure window.

The application of the PFTF model to all 46 data-sets proves it to be a powerful and robust concept for prospective forecasting of slope failures. It is essential to asses the forecast results relative to the lead time which is the time between warning and the impact of an event[41]. The application as an early warning tool premises that the forecasts must be early enough (life expectancy < lead time) and accurate enough (uncertainty < lead time). Our single PFTF concept is successful with a variety of monitoring methods, different rock failure processes in different geological settings and over seven failure magnitudes. The mean life expectancy at the actual time of failure of all sites is -1 ± 17 h for sub-daily and -1 ± 4 d for daily data sets respectively, representing a good fit of the forecasts to reality. The mean failure window width is 24 ± 23 h and 7 ± 4 d, which is in the range of an earlier estimated general forecast uncertainty in 74 pit wall failures of up to 21 h[23]. In general, our results suggest a better performance of the forecasting method with higher frequency data. Yet, forecast uncertainties of less than a day, or a few days respectively, are usually sufficient compared to the respective lead time (Fig. 4g;[41]).

All forecast results are summarized in Fig. 4a–f, separated into two groups of daily (d-f) and higher (a-c) data frequencies. The most probable failure time is plotted as black diamond and the coloured bars represent the width of each failure window during

which a failure is likely. Differences between the sites or data-sets can be connected to inhomogeneous deformations of the unstable body and the positioning of the sensors. Prospectively, these variations cannot be ruled out. The operating experts must evaluate them statistically, taking into account which trends are reflected by which sensors combined with their knowledge of the process dynamics and how the site is monitored. This can be challenging (cp. see Predictability across sensor and failure types). However, the results are robust considering the variability of input data. Note that stronger smoothing, i.e. the lowest group in each subplot, is usually shortening the failure window but also shifting the forecasts to later expected times of failure, possibly towards unsafe predictions (see Subfig. 4ab+de). Generally, the mean forecast errors are decreasing towards failure. Due to the lower data resolution and longer acceleration phase in the daily data-sets, this effect is less distinct here when comparing the last five days before failure (Fig. 4c+f). Overall, failure windows and uncertainties are bigger in the daily data-sets, and, therefore higher sampling frequencies are preferable (compare Fig. 4a–c with d–f).

The comparison of the acceleration time between automatically detected OOA and failure with the detached volumes is generally in line with the assumption[16] that bigger volumes tend to have longer acceleration phases, although a significant correlation cannot be proven with the available data. We calculated a log-log fit of $OOA = 0.76 volume^{0.25}$ based on the 14 analysed sites ($R^2 = 0.30$, p = 0.07). Due to the individuality of each geological setting, strong deviations exist (Fig. 4g). At the biggest analysed failure of Vaiont, the OOA has been detected ca. 50 d before failure after the latest deceleration in 1963 (trend update after February 1963, see Supplementary Discussion 1). The shortest acceleration phase of 3-10 h has been detected at Galterengraben although it is not the smallest volume. Here, the relatively small volume and the brittleness of the hanging sandstone block seem to contribute to a rapid final acceleration. At the smallest analysed event at Grabengufer, the acceleration phase of 5 d is relatively long. In fact, this was a failure of a small block that drifted on a sackung (deep-seated gravitational deformation) of ca. 400,000 m³. The complex superposition of different failure processes and volumes might contribute to the duration of final acceleration. The extraordinary long acceleration phase of up to 200 d at Nevis Bluff is possibly connected to a very consistent increase of acceleration without or with only constant external drivers.

Predictability across sensor and failure types. Our forecast model gives useful results with a variety of failure processes including rock slides, rock falls, toppling, ice fall, complex as well as different monitoring methods including Global Navigation Satellite System, total stations, displacement meters, inclinometers, InSAR and volumes from 10^2 to 10^8 m³). A quantitative comparison of the results based on failure type and monitoring method is inhibited by the diversity of the 14 test sites. Differences in the accuracy of the PFTF model might also result from other factors like the properties of the unstable mass, external drivers like precipitation, snow or pore water pressure, and the location of the sensors. The reason why the PFTF model can deal with a great variety of slope failures nonetheless, is that it uses the phenomenological effect of slope displacement which is the result of all influencing processes. Compared to previously published studies from the same sites[7,28,39,42,43], the PFTF results are comparable or better especially in the sense of prospective real-time application, uncertainty estimation and uniform robustness (see also black points of other studies in Fig. 4c+f). Our uncertainty estimation gives additional statistical information, that helps the responsible decision-maker to judge the informative
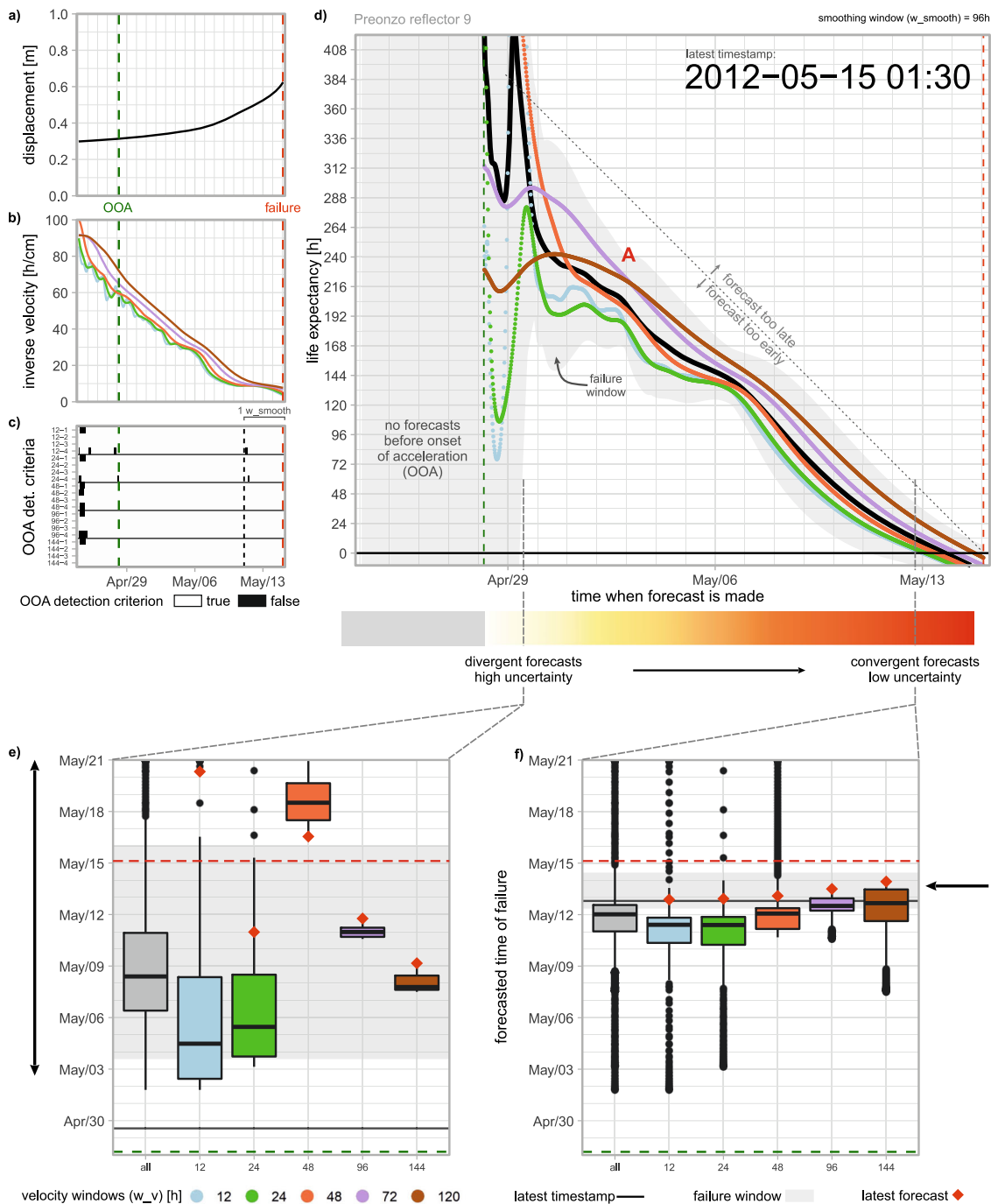
Fig. 3 Final outcome of PFTF at Preonzo (Reflector 9) for a 96 h smoothing window. See the Supplementary Movie for a time-lapse real-time animation of the prospective forecast and Supplementary Fig. 1 for details on the visualization. Colors represent velocity windows (see legend). The dashed dark green line marks the onset of acceleration (OOA). The dashed red line marks the actual time of failure. **a** Displacement. **b** Decreasing inverse velocities. **c** Criteria for OOA detection. The OOA is detected 17 d before the failure. Labels on the vertical axis are "velocity window - criterion number". The black dashed line marks one smoothing window (w_smooth, for details see Methods). **d** Life expectancy plot showing divergent and highly uncertain forecasts during the first days after the OOA. 12 d before failure (point A), the forecasts converge and the failure window (grey area) becomes narrower indicating a low uncertainty. The black line shows the mean of all forecasts at each time step and the grey area represents the failure window during which a failure is likely. **e** Boxplot from April 29 with distributed latest forecasts (red diamonds) and a wide failure window (grey area). The horizontal black line marks the latest timestamp. **f** Boxplot from May 12 with convergent forecasts. Boxplots show the median as line, the interquartile range as box, 1.5 times the interquartile range or the minimum or maximum data point as whiskers and outliers as circles.

value of the forecasts itself. Yet, it must be emphasised that the PFTF must be complemented with geological knowledge and that profound understanding of the unstable site is always needed. Two aspects affecting the predictability can be identified. First, some simplifications are already introduced by the linear inverse

velocity method itself and secondly there are external factors that influence the outcome of the model.

Our PFTF model uses a common[25] linear simplification of Fukuzono's method. Although this can theoretically lead to mistakes, the linear application is mathematically least complex
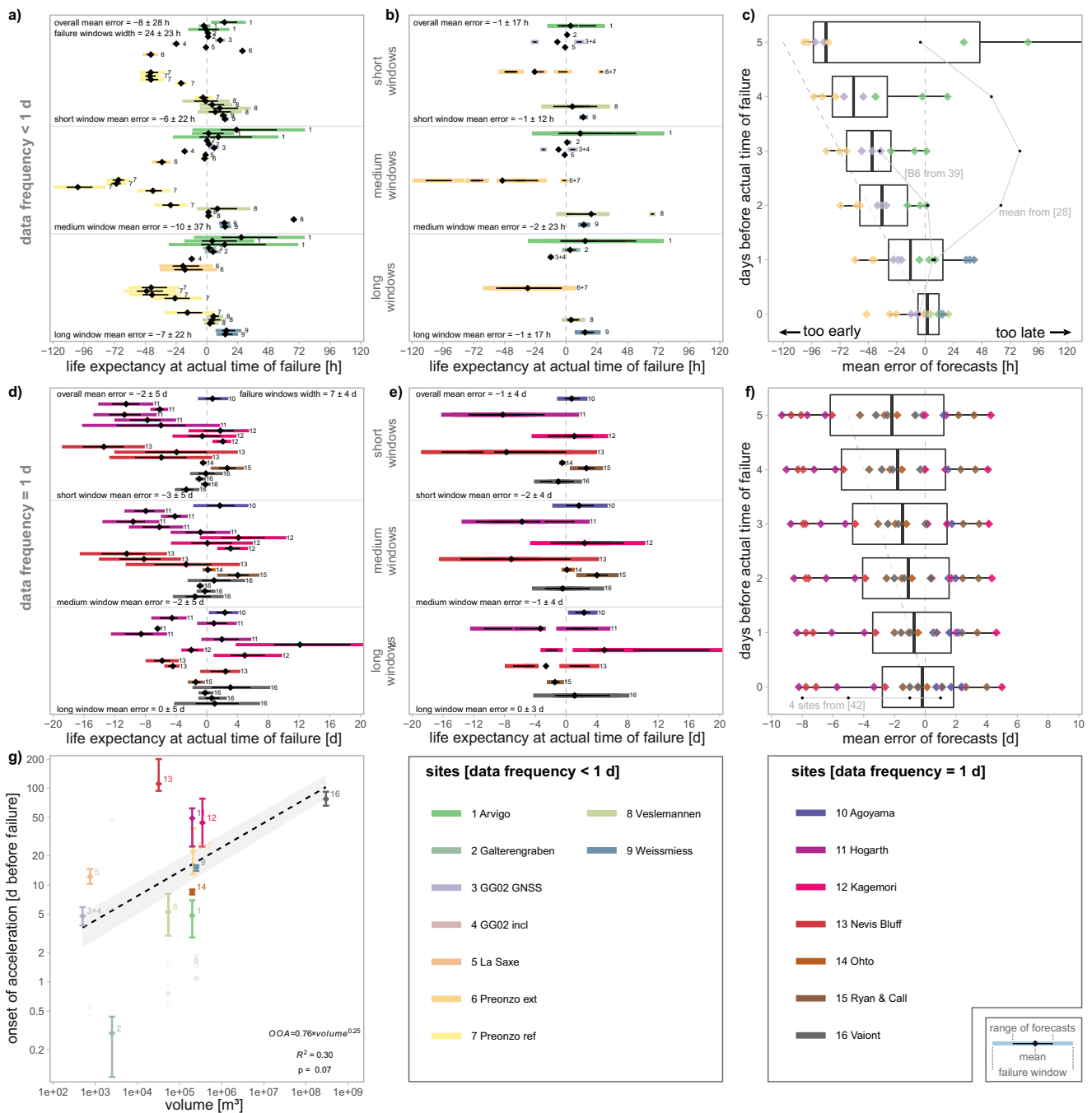
Fig. 4 Forecast results of all tested data. Subplots a-c show high-frequency data-sets; subplots d-f show daily data-sets. Note the different axis scales between the two groups. Boxplots show the median as line, the interquartile range as box, 1.5 times the interquartile range or the minimum or maximum data point as whiskers and outliers as circles. **a** Life expectancy at actual time of failure for short, medium and long window sets (see Methods). A value of 0 means a perfect fit of the forecast to reality. Negative values indicate too early forecasts, positive values too late forecasts. For each sensor, the range of the forecasts (black lines), their mean (black diamonds) and the failure window (in color corresponding to the different sites, see legend) are shown. The mean failure window length is 24 ± 23 h. **b** Forecasts merged per site. The mean merged life expectancy is -1 ± 17 h. **c** Decreasing mean forecast errors in the last 5 d before actual time of failure. **d** Life expectancy at actual time of failure for daily data-sets. The mean failure window length is 7 ± 4 d. **e** Forecasts merged per site for daily data-sets. The mean merged life expectancy is -1 ± 4 d. **f** Decreasing mean forecast errors in the last 5 d before actual time of failure for daily data-sets. **g** OOA detection time against failed volume with black dashed regression line. Larger instabilities can usually be detected earlier due to a longer accelerating phase. The grey area is delimited by the linear regressions of earliest and latest OOA per site, represented by the error bars. Open circles mark secondary OOAs (see Methods).

and therefore more likely used in practise[16]. Furthermore, most cases do not deviate much from the linear case (empirical value $\alpha = 2$)[9,43] validating this simplification. Deviations from the linear behaviour resulting in a slightly convex or concave setting can appear (e.g. Grabengufer site, Supplementary Fig. 4 or

Weissmiess site) due to different internal and external reasons. While in concave cases linear forecasts appear to be on the conservative and safer side (too early), convex cases can be more problematic. To handle this prospectively, the resulting plots must be interpreted carefully. For example, a slightly flatter than

45° slope of the mean life expectancy line and a position of the latest forecasts above the boxes can indicate a shift to a later expected time of failure (concave case; compare Supplementary Fig. 4b). However, in doubtful cases with real danger, it is always on the safer side to expect the failure too early rather than too late, although higher interruption costs might result.

Among external parameters that influence the predictability of a slope failure are the measured or derived displacement used as input, the monitoring technique and frequency, and the failure type and volume. The displacement used as input for the PFTF is not always equal to the true movement of the unstable mass when the sampling direction (e.g. the line of sight of a radar) is not in the direction of the movement or when the derived relative displacement neglects curvy three-dimensional movements. However, we could achieve good results with all tested sensor types irrespective of their sampling direction. The PFTF also works with angle differences, which has not often been used[11], but is cheap and effective to measure[44]. In fact, the forecast results from the Global Navigation Satellite System position data and the inclinometer angle data from the Grabengufer block differ by only about 1 d (see Supplementary Fig. 4), which is mainly due to the application of the same fixed window lengths to the two kinds of data. Another challenge in real-time and at complex instabilities is to judge, if the used sensor data such as point measurements are representative for the later failure. If not, the forecasts can become unreliable[38]. E.g. only a part of an instability could fail as in the La Saxe site[39], or the location of the measurement could be inappropriate. However, the PFTF model can only calculate reasonable forecasts, if the input data are representative.

The selection of optimal input data as well as the interpretation of the data-driven forecast results still depend on experienced users[39] and further research is needed here. One possibility to constrain better, which sensors are most representative of the controlling part of an instability is to use areal measurements like radar, laser scanning or photogrammetry to find the hotspots. Nevertheless, complex cases like the La Saxe site remain challenging, especially if decision-makers have to decide conservatively on the safe side (see the Supplementary Discussion 2).

In any case, higher sampling frequencies potentially enable better results, irrespective of the monitoring technique. Those higher frequencies often necessitate smoothing over a greater number of observations, but from our analysis we infer that higher observation frequencies provide more accurate forecasts. Then, the use of one short-term and one long-term window already brings several advantages[24,34], but applying multiple windows with differing lengths minimizes the subjectivity and enables a direct analysis of the effects of various window lengths in the resulting plots (see Supplementary Fig. 1).

Although the data possibly indicate a relation between the failed volume and the acceleration time between OOA and failure, it is not the goal to achieve this relation in prospective forecasting. The goal is to detect an optimal OOA as early as possible to enable timely forecasting. Concurrently, the expected duration of the acceleration phase is relevant for the determination of the smoothing window lengths. E.g. if the acceleration phase is only several hours, smoothing windows of multiple days are useless. Two sites from this study are paradigmatic for a challenging kinematic behaviour with false OOA detection that complicates interpretations[8,27,39]. The La Saxe instability shows a diverse velocity regime and a strong deceleration at April 17 shortly before the partial failure. At Veslemannen, a dominant environmental forcing controls periodic accelerations. In such cases, experience, detailed knowledge of the site specific processes and the determination of secondary OOAs (circles in Fig. 4g) support successful forecasting. By detecting trend update points[24,34] and adjusting the OOA accordingly, reliable prospective forecasts can still be calculated. The only site in this study where all failure windows lie on the unsafe side (forecast after failure) is the Weissmiess ice fall. Three factors could be a reason for this: (i) the radar line of sight is not in the direction of the movement, (ii) the radar observations started after the OOA and thus no optimal OOA can be found, and (iii) the ice fall behaves different than rock slope failures. A slightly convex behaviour could result from the high strain rates increasing from 1 to over 3.5 md^{-1} during the last two days before failure. However, this site supports that the concept of increasing accelerations is not limited to fracturing[9] nor to rock and earth material.

Our results show, that the prospective failure time forecast concept (PFTF) is able to overcome the major drawbacks of current retrospective failure time forecasting methods achieving reliable results with a variety of slope failure processes, volumes, materials and sensor types. It can detect the onset of acceleration automatically and provides uniform forecast information including an uncertainty estimation. In combination with the ongoing progress in wireless (near) real-time monitoring and internal fracturing observations[45], the PFTF can become a key element for future reliable and quantitative real-time natural hazard management. Its application might not be limited to rock slopes, but the underlying physical principle might also be valid at earth slopes, man-made slopes, artificial structures and glaciers, thus supporting decision-makers in a multitude of critical situations.

Methods

Iterative prospective forecasting. The algorithm of the PFTF model calculates prospective real-time (or simulated real-time) forecasts iteratively with every new available data point. The concept of the PFTF model is shown in Fig. 1. Fig. 2a gives an overview of the algorithm steps which include:

- reading new displacement data with every iteration
- removing outliers/ erroneous measurements
- smooth raw displacement data (see Data smoothing)
- calculate velocities and inverse velocities with multiple velocity windows (see Inverse velocity calculation)
- calculate 1% and 50% quantiles of inverse velocities (see Starting point definition)
- automatically detect the onset of acceleration (OOA) which is used as starting point to calculate forecasts (see Starting point definition)
- calculate and plot the time to failure (life expectancy) and time of failure (see Forecast and uncertainty presentation)

For the technical implementation of this workflow in an open-source R-Code, (RStudio Version 2021.09.0 and R Version 4.1.3[46] see the code availability statement[47].

Data preparation. The PFTF model reads displacement data as input. Thus, position data have been converted to relative displacements and inclination data to relative angle displacement respectively. All data have been formatted as time series with equal time steps and no gaps, so that every time step contains one displacement data point or missing values (NA) in time steps without measurements. For details on the monitoring method see the Supplementary Table.

Data smoothing. The PFTF model uses a two-stage smoothing approach with multiple windows to enable OOA detection and statistical uncertainty estimation. First, the raw displacement is smoothed by a moving mean over the last n values while n is specified by the length of the given smoothing windows (w_smooth). To minimize the effect of subjective decisions, we propose the application of multiple smoothing windows. Displacements (Fig. 3a) are plotted after smoothing. A second step of data smoothing is performed during the calculation of the inverse velocity via linear regression over multiple velocity windows (w_v) (see Inverse velocity calculation).

To create comparability between all sites in this study, we defined three sets of standard windows that we applied to the analysed data based on a simple decision schema (Supplementary Fig. 2). Note that other data might require different smoothing window lengths.

Inverse velocity calculation. Velocities are calculated from the smoothed displacement using a linear regression model over the last n values while n is specified by the length of the given velocity windows (w_v). This method includes all displacement measurements within w_v and gives the best fitting velocity over the last w_v. For the standard window sets (Supplementary Fig. 2), we used five velocity

windows in each case, ranging from $\frac{1}{8}w\_smooth$ to $5w\_smooth$. Inverse velocities are calculated as $velocity^{-1}$ and plotted in Fig. 3b.

Starting point definition. All inverse velocity data after the defined starting point (onset of acceleration, OOA) are included in the extrapolated linear regression to forecast the time of failure. The real-time detection of the OOA is based on four criteria (Fig. 2b). If all criteria are true over one complete smoothing window (w_smooth), the OOA is detected at $1w\_smooth$ before current time.

- **criterion 1**: increasing displacement rate. The displacement difference over the second half of the last w_smooth must be greater than over the first half of the last w_smooth.
- **criterion 2**: decreasing inverse velocity. The inverse velocity must decrease over the last w_smooth for all velocity windows w_v. This is different from criterion 1 as the different w_v can have differing trends.
- **criterion 3**: decreasing 50% quantile. The 50% quantile must decrease over the last w_smooth for all velocity windows w_v. This represents a general long-term accelerating trend.
- **criterion 4**: inverse velocity smaller than 1% quantile. The latest inverse velocity must be smaller than the 1% quantile of all inverse velocities. Respectively, the current velocity must be greater than 99% of all velocities before. This represents a short-term accelerating trend.

If the displacement data show a trend update due to a deceleration phase or a changing process regime, the analysis is restarted only with data after the trend update (compare Supplementary Fig. 2) and secondary OOAs are determined.

Forecast and uncertainty presentation. Forecasts are calculated with the linear version ($\alpha = 2$) of Fukuzono's method[19]. The intersection of an extrapolated linear trend line of all inverse velocities after the OOA with the horizontal time axis corresponds to the forecasted time of failure. The results are plotted in a life expectancy plot[24,34], where the predicted time until failure (life expectancy) is plotted against the time when the forecast is made on the abscissa. This allows simple updating with every new data point and a good visualization of all historic forecasts. The axes are scaled 1:1 for easy identification of changing forecast trends. Completely consistent new forecasts appear 45° downward while a shift to later expected failure times appears flatter and vice versa. The PFTF creates one life expectancy plot per w_smooth and one coloured point for every w_v used (see Supplementary Fig. 1).

To evaluate the statistical uncertainty of the forecasts, we calculate a failure window during which failure is highly probable after[24]. It contains the time between the earliest and the latest forecast of each time step and half of that time span added on each side. In the life expectancy plot, we show the mean life expectancy of all w_v per time step as black line and the failure window as grey area.

Additionally, updated standard boxplots of all forecasts since the OOA are produced. The plot includes one box per w_v (colour) and one for all w_v (grey). The latest forecasted times of failure per w_v are plotted as red diamonds. A horizontal grey line indicates the latest timestamp. The failure window is marked as grey area (see Supplementary Fig. 1).

Statistical evaluation of results. For the statistical analysis of the results of all sites in Fig. 4, the life expectancies, their mean, and the failure window were extracted from the forecasts five to zero days before actual time of failure. The mean error is calculated as average of all forecast means. Uncertainties are expressed as standard deviation. For the Subfigures 4b+e, all forecasts per site are merged into one line. The merged mean is the average of all means per site. The time of OOA detection in Subfigure 4g is plotted as mean and range of all values per site. The linear regression is calculated over the $\log_{10}$ values of mean OOA detection time and volume. The uncertainty area is delimited by the regression lines of minimum and maximum OOA detection times.

Site selection. The tested historical slope failure events have been selected based on availability and documentation. All sites are rock slope failures, except Weissmiess which is an ice fall. All sites have a daily or higher sampling frequency, and include measurements starting before the onset of acceleration and until failure. The 14 sites represent a variety of monitoring methods and volumes. Details on the data sets, sources and where data have been extracted from plots are listed in Supplementary Table.

Data availability
The source of all used monitoring data is indicated in the Supplementary Table. Data of ten sites are taken from published studies. Original data of four sites have been provided by third party persons or institutions (Arvigo, Galterengraben, Preonzo, and Weissmiess). A collection of all data is available under https://doi.org/10.14459/2023mp1688868[48].

Code availability
The R-Code of the PFTF model including instructions is completely open-source and available under https://doi.org/10.5281/zenodo.8010361[47].

References
1. Lacasse, S. & Nadim, F. Landslide risk assessment and mitigation strategy. In Sassa, K. & Canuti, P. (eds.) *Landslides – Disaster Risk Reduction*, 31–61 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2009).
2. Petley, D. Global patterns of loss of life from landslides. *Geology* **40**, 927–930 (2012).
3. Guzzetti, F. Landslide fatalities and the evaluation of landslide risk in Italy. *Eng. Geol.* **58**, 89–107 (2000).
4. Guzzetti, F. On the prediction of landslides and their consequences. In Sassa, K. et al. (eds.) *Understanding and Reducing Landslide Disaster Risk*, ICL Contribution to Landslide Disaster Risk Reduction, 3–32 (Springer International Publishing, Cham, 2021).
5. Evans, S. G. Single-event landslides resulting from massive rock slope failure: characterising their frequency and impact on society. In Evans, S. G., Mugnozza, G. S., Strom, A. & Hermanns, R. L. (eds.) *Landslides from Massive Rock Slope Failure*, vol. 49 of *NATO Science Series*, 53–73 (Springer Netherlands, Dordrecht, 2006).
6. Picarelli, L., Lacasse, S. & Ho, K. K. S. The impact of climate change on landslide hazard and risk. In Sassa, K. et al. (eds.) *Understanding and Reducing Landslide Disaster Risk*, ICL Contribution to Landslide Disaster Risk Reduction, 131–141 (Springer International Publishing, Cham, 2021).
7. Kilburn, C. R. & Petley, D. N. Forecasting giant, catastrophic slope collapse: lessons from Vajont, Northern Italy. *Geomorphology* **54**, 21–32 (2003).
8. Intrieri, E. & Gigli, G. Landslide forecasting and factors influencing predictability. *Nat. Hazards Earth Syst. Sci.* **16**, 2501–2510 (2016).
9. Intrieri, E., Carlà, T. & Gigli, G. Forecasting the time of failure of landslides at slope-scale: A literature review. *Earth-Sci. Rev.* **193**, 333–349 (2019).
10. Federico, A., Popescu, M. & Murianni, A. Temporal prediction of landslide occurence: a possibility or a challenge? *IJEGE* **15**, 41–60 (2015).
11. Petley, D. N. The evolution of slope failures: mechanisms of rupture propagation. *Nat. Hazards. Earth Syst. Sci.* **4**, 147–152 (2004).
12. Amitrano, D. & Helmstetter, A. Brittle creep, damage, and time to failure in rocks. *J. Geophys. Res.: Solid Earth* **111**, n/a–n/a (2006).
13. Main, I. G. A damage mechanics model for power-law creep and earthquake aftershock and foreshock sequences. *Geophys. J. Int.* **142**, 151–161 (2000).
14. Corcoran, J. & Davies, C. M. Monitoring power-law creep using the failure forecast method. *Int. J. Mech. Sci.* **140**, 179–188 (2018).
15. Kemeny, J. The time-dependent reduction of sliding cohesion due to rock bridges along discontinuities: A fracture mechanics approach. *Rock Mech. Rock Eng.* **36**, 27–38 (2003).
16. Rose, N. D. & Hungr, O. Forecasting potential rock slope failure in open pit mines using the inverse-velocity method. *Int. J. Rock Mech. Min. Sci.* **44**, 308–320 (2007).
17. Saito, M. & Uezawa, H. Failure of soil due to creep. *Proc. 5th Int. Conf. on Soil Mech. Found. Engineering, Paris* **1**, 315–318 (1961).
18. Saito, M. Forecasting time of slope failure by tertiary creep. *Proc. 7th Int. Conf. Soil Mech. Found. Eng. Mexico* **2**, 677–683 (1969).
19. Fukuzono, T. A method to predict the time of slope failure caused by rainfall using the inverse number of velocity of surface displacement. *J. Japan Landslide Soc.* **22**, 8–14 (1985).
20. Voight, B. A relation to describe rate-dependent material failure. *Science (New York, N.Y.)* **243**, 200–203 (1989).
21. Voight, B. Materials science law applies to time forecasts of slope failure. *Landslide News* **3**, 8–11 (1989).
22. Mufundirwa, A., Fujii, Y. & Kodama, J. A new practical method for prediction of geomechanical failure-time. *Int. J. Rock Mech. Min.* **47**, 1079–1090 (2010).
23. Cabrejo-Liévano, A. G. Analysis of failures in open pit mines and consideration of the uncertainty when predicting collapses. In Dight, P. M. (ed.) *Slope stability 2013*, 483–498 (ACG, Nedlands, 2013).
24. Carlà, T. et al. Guidelines on the use of inverse velocity method as a tool for setting alarm thresholds and forecasting landslides and structure collapses. *Landslides* **14**, 517–534 (2017).
25. Casagli, N. et al. Monitoring and early warning systems: Applications and perspectives. In Casagli, N., Tofani, V., Sassa, K., Bobrowsky, P. T. & Takara, K. (eds.) *Understanding and Reducing Landslide Disaster Risk*, ICL Contribution to Landslide Disaster Risk Reduction, 1–21 (Springer International Publishing, Cham, 2021).
26. Petley, D. N. & Petley, D. J. On the initiation of large rockslides: Perspectives from a new analysis of the Vaiont movement record. In Evans, S. G., Mugnozza, G. S., Strom, A. & Hermanns, R. L. (eds.) *Landslides from Massive Rock Slope Failure*, vol. 49 of *NATO Science Series*, 77–84 (Springer Netherlands, Dordrecht, 2006).

27. Mazzanti, P., Bozzano, F., Cipriani, I. & Prestininzi, A. New insights into the temporal prediction of landslides by a terrestrial SAR interferometry monitoring case study. *Landslides* **12**, 55–68 (2015).

28. Sättele, M., Krautblatter, M., Bründl, M. & Straub, D. Forecasting rock slope failure: how reliable and effective are warning systems? *Landslides* **13**, 737–750 (2016).

29. Carlà, T. et al. Perspectives on the prediction of catastrophic slope failures from satellite InSAR. *Sci. Rep.* **9**, 14137 (2019).

30. Grebby, S. et al. Advanced analysis of satellite data reveals ground deformation precursors to the Brumadinho Tailings Dam collapse. *Communications Earth & Environment* **2** (2021).

31. Intrieri, E. et al. Early warning systems in Italy: State-of-the-art and future trends. In Sassa, K. et al. (eds.) *Understanding and Reducing Landslide Disaster Risk*, ICL Contribution to Landslide Disaster Risk Reduction, 537–543 (Springer International Publishing, Cham, 2021).

32. Kristensen, L. et al. Movements, failure and climatic control of the Veslemannen rockslide, Western Norway. *Landslides* (2021).

33. Chae, B.-G., Park, H.-J., Catani, F., Simoni, A. & Berti, M. Landslide prediction, monitoring and early warning: a concise review of state-of-the-art. *Geosci. J.* **21**, 1033–1070 (2017).

34. Dick, G. J., Eberhardt, E., Cabrejo-Liévano, A. G., Stead, D. & Rose, N. D. Development of an early-warning time-of-failure analysis methodology for open-pit mine slopes utilizing ground-based slope stability radar monitoring data. *Can. Geotech. J.* **52**, 515–529 (2015).

35. Bozzano, F., Mazzanti, P. & Moretto, S. Discussion to: 'Guidelines on the use of inverse velocity method as a tool for setting alarm thresholds and forecasting landslides and structure collapses' by T. Carlà, E. Intrieri, F. Di Traglia, T. Nolesini, G. Gigli and N. Casagli. *Landslides* **15**, 1437–1441 (2018).

36. Valletta, A., Carri, A. & Segalini, A. Definition and application of a multi-criteria algorithm to identify landslide acceleration phases. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* 1–15 (2021).

37. Krähenbühl, R. Der Felssturz, der sich auf die Stunde genau ankündigte. *Bull. angew. Geol.* **11**, 49–63 (2006).

38. Manconi, A. & Giordan, D. Landslide early warning based on failure forecast models: the example of the Mt. de La Saxe rockslide, northern Italy. *Nat. Hazards Earth Syst. Sci.* **15**, 1639–1644 (2015).

39. Manconi, A. & Giordan, D. Landslide failure forecast in near-real-time. *Geomatics Nat. Hazards Risk* **7**, 639–648 (2016).

40. Iwata, N. & Sasahara, K. Influence of intervals measuring surface displacement on time prediction of slope failure using Fukuzono method. In Casagli, N., Tofani, V., Sassa, K., Bobrowsky, P. T. & Takara, K. (eds.) *Understanding and Reducing Landslide Disaster Risk*, ICL Contribution to Landslide Disaster Risk Reduction, 307–313 (Springer International Publishing, Cham, 2021).

41. Hermle, D., Keuschnig, M., Hartmeyer, I., Delleske, R. & Krautblatter, M. Timely prediction potential of landslide early warning systems with multispectral remote sensing: a conceptual approach tested in the Sattelkar, Austria. *Nat. Hazards Earth Syst. Sci.* **21**, 2753–2772 (2021).

42. Loew, S., Gschwind, S., Gischig, V., Keller-Signer, A. & Valenti, G. Monitoring and early warning of the 2012 Preonzo catastrophic rockslope failure. *Landslides* **14**, 141–154 (2017).

43. Segalini, A., Valletta, A. & Carri, A. Landslide time-of-failure forecast and alert threshold assessment: A generalized criterion. *Eng. Geol.* **245**, 72–80 (2018).

44. Cicoira, A. et al. In situ observations of the swiss periglacial environment using gnss instruments. *Earth Syst. Sci. Data* **14**, 5061–5091 (2022).

45. Weber, S. et al. Quantifying irreversible movement in steep, fractured bedrock permafrost on Matterhorn (CH). *Cryosphere* **11**, 567–583 (2017).

46. R Core Team. R: A language and environment for statistical computing https://www.R-project.org/ (2022).

47. Leinauer, J., Weber, S. & Cicoira, A. R-tool for prospective failure time forecasting of slope failures using inverse velocity https://doi.org/10.5281/zenodo.8010361 (2023).

48. Leinauer, J., Weber, S., Cicoira, A., Beutel, J. & Krautblatter, M. Collection of slope deformation data until failure with a timely resolution of daily or better https://doi.org/10.14459/2023mp1688868 (2023).

Author contributions

JL wrote the manuscript and the R-code, designed the figures, collected monitoring data and performed the data analysis. SW had the initial idea for the concept. SW, AC and JB provided monitoring data and improved the concept substantially. SW, AC and MK contributed to the interpretation of the results and improved figures and manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-023-00909-z.

Correspondence and requests for materials should be addressed to Johannes Leinauer.

Peer review information *Communications Earth & Environment* thanks David Toll and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Joe Aslin. A peer review file is available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.