Change-point analysis of geophysical time series: application to landslide displacement rate (Séchilienne rock avalanche, France)

D. Amorese$^{1,2}$, J.-R. Grasso$^3$, S. Garambois$^3$, M. Font$^4$

$^1$OVSG, IPGP, Sorbonne Paris Cité, UMR 7154 CNRS, France
$^2$Dept. Biology and Earth Sciences, Université de Caen-Normandie, Esplanade de la Paix, CS 14032 Caen Cedex 5, France
$^3$ISterre, Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, IFSTTAR, ISterre, 38000 Grenoble, France
$^4$Univ Caen, Caen, F-14000 France ; CNRS, UMR 6143 M2C, Caen, F-14000 France

Accepted 2018 February 13. Received 2018 January 12; in original form 2017 June 17

SUMMARY
The rank sum multiple change-point method is a robust statistical procedure designed to search for the optimal number and the location of change-points in an arbitrary continue or discrete sequence of values. As such, this procedure can be used to analyze time series data. Twelve years of robust data sets for the Séchilienne (French Alps) rockslide show a continuous increase in average displacement rate from 50 to 280 mm per month, in the 2004 - 2014 period, followed by a strong decrease back to 50 mm per month in the 2014 - 2015 period. When possible kinematic phases are tentatively suggested in previous studies, its solely rely on the basis of empirical threshold values. In this paper, we analyse how the use of a statistical algorithm for change point detection helps to better understand time phases in landslide kinematics. First we test the efficiency of the statistical algorithm on geophysical benchmark data, these data sets (stream flows, northern hemisphere temperatures) being already analysed by independent statistical tools. Second, we apply the method to 12-year daily time series of the Séchilienne landslide, for rainfall and displacement data, from December 2003 to December 2015, in order to quantitatively extract changes in landslide kinematics. We find two strong significant discontinuities in the weekly cumulated rainfall values: an average rainfall rate increase is resolved in April 2012 and a decrease in August 2014. Four robust changes are highlighted in the displacement time series (May 2008, November-December 2009-January 2010, September 2012 and March 2014), the 2010 one being preceded by a significant but weak rainfall rate increase (in November 2009). Accordingly, we are able to quantitatively define 5 kinematic stages for the Séchilienne rock avalanche during this period. The synchronisation between the rainfall and displacement rate, only resolved at the end of 2009 and beginning of 2010, corresponds to a remarkable change (fourfold increase in mean displacement rate) in the landslide kinematic. This suggests that an increase of the rainfall is able to drive an increase of the landslide displacement rate, but that most of the kinematics of the landslide is not directly attributable to rainfall amount. The detailed exploration of the characteristics of the 5 kinematic stages suggests that the weekly averaged displacement rates are more tied to the frequency or rainy days than to the rainfall rate values. These results suggest the pattern of Séchilienne rock avalanche is consistent with the previous findings that landslide kinematics is dependent upon not only rainfall but also soil moisture conditions (as known as being more strongly related to precipitation frequency than to precipitation amount). Finally our analysis of the displacement rate time series pinpoints a susceptibility change of slope response to rainfall, as being slower before the end of 2009 than after, respectively. The kinematic history as depicted by statistical tools opens new routes to understand the apparent complexity of Séchilienne landslide kinematic.

Key words: Creep and deformation – Europe – Geomechanics – Statistical methods – Time series analysis
1 INTRODUCTION

The mitigation of landslide hazard is receiving tremendous attention in many parts of the World. Two main approaches help to assess the risk: the static approach (prevention) consists in mapping of the landslide susceptibility, whereas the dynamic approach aims to predict the spatial-temporal progress of landslides through monitoring and mathematical/physical models. The role played by antecedent rainfall in the triggering of landslide is not a matter of debate (Van Asch et al., 1999; Godt et al., 2006): it prompts several authors (Ray & Jacobs, 2007; Ponziani et al., 2012; Brocca et al., 2012) to propose rainfall and soil moisture thresholds in the design of early warning system. Thus, most landslide warning systems use thresholds that are tied to recent precipitation amounts and durations. Baum & Godt (2010) in their compilation study for precipitation thresholds across the USA reported that, at every scale (including county and nation), there is no consensus on these values: variations of several orders of magnitude are frequently observed. Our study proposes a sligh different approach: we do not focus on rainfall events that resulted or did not result in landslide. Our attention is primarily focussed on surface landslide displacement values (and associated rainfall series). Indeed, it has been shown that a more accurate prediction of the occurrence of failure can be obtained by measuring the landslide displacement rates than by analysing rainfall intensity-duration diagrams (Federico et al., 2004). Actually, as for rainfall, there is no consensus on velocity threshold values for alerts: for instance, for creeping-type rockslides, the critical values may differ by an order of magnitude from one site to another, from tens to several hundreds mm/day (Crosta & Agliardi, 2003). We conduct an analysis that is related to this critical issue of the definition of thresholds, but assessing the reliability of the correlation between the probability of occurrence of a landslide and rainfall measurements is not our primary objective. When a landslide is monitored, there is interest in determining whether changes in displacement measurements are the results of random chaotic processes or indicate significant and lasting changes in the underlying physics (or in the way data are collected/processed). If the change is significant, then identifying when it occurs and quantifying its significance can be important breakthroughs toward the identification of the triggering processes and the onset of possible early warning strategies. Incidentally, analysing the dynamics of landslide is a prerequisite in the most sophisticated applications of the statistical impulse response model for the prediction of landslide velocities (Abellán et al., 2015; Bernardie et al., 2015). The analysis of the time evolu-

tion and/or the cross-correlation of measurements collected on landslides is part of the ordinary course of research on solid flow hazards (Hungr et al., 2005; Hemsstetter & Garambois, 2010; Crosta et al., 2014; Rianna et al., 2014; Benoit et al., 2015; Schlögel et al., 2015; Confuorto et al., 2017). But so far, no reproducible method of analysis has been applied on these kind of data for the determination of successional stages. In this paper, a non-parametric technique for the detection of changes in times series, first introduced by Lanzante (1996), is applied to the time series of the Séchilienne landslide. The purpose of this study is to improve the description of the landslide kinematics (definition of motion acceleration or deceleration stages) and then to better assess which factors (average rainfall rate, rainfall frequency, intense rainfall event) influence the Séchilienne landslide evolution. In the first part of this work, the efficiency and outcomes of the methodology are tested on natural and synthetic data sets. Then, landslide displacement and rainfall series are analyzed. Eventually, this approach allows to sort out a list of features of the 5 kinematic stages of the Séchilienne rock avalanche in the 2003-2015 period.

2 STUDY SITE AND DATA

The Séchilienne landslide is located on the southern slope of the Mont-Sec Massif in the Dauphiné region (French Alps). The most unstable part of the landslide (estimated volume: 3 $10^6$ m$^3$) is threatening the main communication axis between major alpine cities, Grenoble and Briançon cities, and including key access to both regional ski resorts and northern Italy region. Formation of a landslide dam lake into the downslope Romanche river valley is another possible major issue of the earthfall. The survey of the site began in 1985, including dense displacement and weather measurements. The surface strain monitoring of the Séchilienne landslide in fully operated by the CEREMA (Centre d’Etudes et d’Expertise sur les Risques, l’Environnement, la Mobilité et l’Aménagement) of Lyon (Kasperski et al., 2010). In this study, we use displacement data from the most active zone of the landslide i. e., those acquired by extensometer #13 (Fig. 1). We investigate lengthening values that have been collected every day by the extensometer #13 from December 1, 2003 to December 31, 2015 for a total of 4,337 recordings. The Vizille weather station (latitude: 45.0802 N; longitude: 5.7694 E; elevation: 290 m) maintained by Sébastien Pierart (ROMMA association) is located about 3.5 km NW of the landslide and operated on the whole time period. It is also sampled at a daily rate. Several authors (Durville
Nevertheless, the WMW test is basically a test for the variability within each segment: the WMW test can be viewed as a median test. When the assumption of homogeneity of variances is not required in this application of the WMW statistics. The method can be called the rank-sum multiple change-point method (hereafter RSM-CPM).

At each point \( i \) in the series of \( n \) points, the sum of the ranks (\( SR_i \)) from the beginning of the series to that point is calculated. Since the sum of the ranks depends on the number of points, \( SR_i \) is adjusted. Then, an adjusted value for the sum of the ranks, \( SA_i \), is:

\[
SA_i = |2SR_i - i(n + 1)| \tag{1}
\]

In the right-hand side of equation (1), \( i(n + 1) \) is connected with

\[
E(W_i) = i(n + 1)/2 \tag{2}
\]

which is the expected value of the rank sum for the \( i \) first observed ranks out of a total of \( n \) points. The next step of the procedure is to find the maximum of \( SA_i \) to divide the series into two segments. The point \( n_1 \) being where the value of \( SA_i \) is maximum, the following variables are defined:

\[
W = SR_{n_1} \tag{3}
\]

and

\[
n_2 = n - n_1. \tag{4}
\]

After this, the WMW test is used to decide whether or not the null hypothesis (that there is no change in the sequence at \( n_1 \)) is rejected, in favor of the alternative hypothesis. In this study, the chosen level of significance is 5% as the generally accepted and expected alpha (type I error rate) value in most disciplines for statistical tests. The RSMCPM is applied to a given series as long as the statistical significance of each new change point is less than the specified significance level. For each iteration, a list of \( N \) change points is delivered that defined \( N + 1 \) segments. At each iteration, the series is adjusted by subtracting the median of its segment from each point.

Additionally to change point significance probabilities and following Lanzante (1996), a signal to noise ratio (SNR) which quantifies the magnitude of each discontinuity is computed. For a given change point, this ratio appraises the variability associated with the shift in level between the adjacent segments relative to the variability within each segment.
where $S_{CP}^2$ is the variance owing to the shift in level between the segments adjacent to the change point and $S_N^2$ is the noise variance. The value of $S_{CP}^2$ is (Lanzante, 1996):
\[
S_{CP}^2 = \frac{n_1(\bar{X}_1 - \bar{X})^2 - n_2(\bar{X}_2 - \bar{X})^2}{n - 1}
\]
where $n_1$ and $n_2$ are the number of values in the left and right segments (segments are right-closed and left-open intervals) respectively ($n_1 + n_2 = n$). $\bar{X}_1$ and $\bar{X}_2$ are the estimates of the mean values in these two segments. $\bar{X}$ is the overall mean:
\[
\bar{X} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n}.
\]

The noise variance, $S_N^2$, is the variance of the combination of the two segments after they were normalized by subtracting $\bar{X}_1$ or $\bar{X}_2$ from all of the values in the left or right segment. It should be noted that the length for determining the SNR can be customized: 50 points away from the change point is the value recommended by Lanzante (1996). Larger values are prone to weaken the force and usefulness of SNR values. The SNR can be used to eliminate change points which are too “weak”. Lanzante (1996) proposed a SNR value of at least 0.05 or 0.1 for “important” change points.

In the comparison process of a part of the times series with what has occurred before (or what will occur after), it is easy to understand that the center of a segment which has a trend can be falsely identified as a change point, since the trend alters mean values. To overcome this drawback, Lanzante (1996) proposed a straightforward and effective solution: after each new change point is identified, both the $S_N^2$ (as described above) and a so-called “trend noise variance” are computed and compared. If the trend slope is b, the trend noise variance, $S_T^2$, is the variance of the combination of the two segments after they were normalized by subtracting b($x_i - x_{CP}$) from all of the values $y_i$ in the left and the right segments ($x_i$ and $x_{CP}$ are the x-values of $y_i$ and of the suspected change point, respectively). If the $S_T^2$ value is larger than the $S_N^2$ value, then the change point is validated since the variability remains apparent in the trend noise variance (which is meant to be corrected from the trend bias). Otherwise ($S_T^2 < S_N^2$), a trend reduction is applied (by subtracting b($x_i - x_{CP}$) from $y_i$) before the suspected change point is tested again.

Several parameters can be customized in the RSMCPM. The results of this study were obtained using values that we have defined as “default values” for the algorithm:

- The default value for the significance level is 0.05 (Fisher, 1925).
- The default maximum number of points for the calculation of the SNR value is 50 points on both sides of the change point, which means a maximum number of 100 points for the calculation.
- The default maximum number of iterations is 5. The maximum number of iterations is unrelated to the possible number of change points: it concerns the maximum number of adjustments that are allowed until a new change point is detected.
- The default minimum number of points at each end of the series is 2.
- The default number of points (inter-change point gap) between two consecutive change points is 2.
- The default threshold difference between $S_{CP}^2$ and $S_T^2$ is 0. This means that the trend reduction is applied whenever the discontinuity noise variance is larger than the trend noise variance. The larger the value of this threshold, the more restricted is the application of the trend reduction.

Since a decade, many approaches and software have addressed the determination of change points in univariate and/or multivariate time series. Some Bayesian methods require independent Gaussian observations (Barry & Hartigan, 1993; Erdman & Emerson, 2007). As we will show in the next section, the condition of normality is not matched for our extensometer displacement data series. This limitation precludes the sustainable usage of the Erdman et al.’s method (2007) for our data set, whereas our technique works properly. Both type and shape of the probability distributions are not problematic for the RSMCPM. Nevertheless, this technique cannot be applied to continuous cumulative data, because, in common with many other methods of statistical inference, the WMW test also requires independence within groups (Hollander & Wolfe, 1973). As a final technical remark, one must note that the RSMCPM does not perform better than others technique when trying to detect change points close to the ends of the series. Actually, if the sum of the size of the two samples under comparison is smaller than 10, no matter how much the two groups differ, the WMW test loses power (ability to reject a false null hypothesis): in this case, it fails to reach statistical significance at the level $\alpha = 0.01$ for the two-sided problem. Thus, even a very strong change point may not be detected if it occurs in the ends of the series under study. Nevertheless, we guess the RSMCPM does not worse than other statistical methods when dealing with series ends.
3.2 Method validation on geophysical time series

Before processing our landslide data, we validate the use of our rank-sum change point detection algorithm on benchmark data: (i) the synthetic combination of normal distributions, investigated by James & Matteson (2014), (ii) the annual January to June streamflow amounts for the Romaine river in Quebec, that has been explored by means of a Bayesian change point approach in Perreault et al. (2000), and (iii) the Northern Hemisphere temperature from proxy data, A.D. 200-1995 (Jones & Mann, 2004; Matyasovszky, 2011).

Figure 2 shows the results of the RSMCPM, as applied to a synthetic combination of normal distributions. This is a modified version of the data example from James & Matteson (2014)’s study: a sequence of 100 independent samples from normal distributions \( N(0, 1), N(0, 3), N(2, 1), \) and \( N(2, 4) \). The notation \( N(\mu, \sigma) \) means normally distributed with mean \( \mu \) and standard deviation \( \sigma \). We slightly upgrade this synthetic test by adding an extra \( N(0, 3) \) very short (10 samples) segment at the end of the initial sequence. This extra tip is added (Fig. 2a) in order to assess the detection capability for a breakpoint close to series’ end, where an edge effect may be significant. In Figure 2b, a 5% slope is added to the synthetic series of Figure 2a to simulate a series with upward trend. As can be seen (Fig. 2a), the method has difficulty identifying changes in standard deviation (the first breakpoint is missed), but it is perfectly successful for the detection of the change in mean at index 200. Even close to the series’ end (Fig. 2a), the final breakpoint (index 400) is detected. Nevertheless, the results for this point is not accurate (index 389). When the series shows a trend (Fig. 2b), the first breakpoint is mislocated and the final breakpoint is missed, but there is no detection of extra change points. As comparison across approaches, a divisive hierarchical estimation algorithm for multiple change point analysis (James & Matteson, 2014) is applied and gives the results displayed in figures 2c and 2d. By doing so, we mimic one of the examples proposed by James & Matteson (2014) to illustrate the use of their divisive algorithm, except that these authors did not examine the case in which the data show a trend. We show that the James & Matteson (2014)’s algorithm misses the second and the last breakpoint in Figure 2c and it detects many fictitious breakpoints in the series that shows a trend (Figure 2d).

As a second example we partly revisited the 1970-2000 streamflow time series for Northern Quebec Labrador region (Perreault et al., 2000). These authors used a Bayesian change point analysis to suggest a change in the average streamflow of rivers that occurred in 1984. Figure 3a displays the marginal posterior probability density function of the change point for the Romaine river, as obtained from the Bayesian change point analysis (Perreault et al., 2000). The RSMCPM (Fig. 3b) achieves similar results than Perreault et al. (2000)’s results but we resolve an additional breakpoint in 1976 with a significant signal-to-noise ratio (0.32). One must note that the 1984 change point is quantified by a strong SNR value (2.35).

As a third test, we use the temperature time series (A. D. 200-1995) for the Northern hemisphere (Jones & Mann, 2004; Matyasovszky, 2011). The list of abrupt changes detected in the Northern Hemisphere temperature values for the period 200-1995 includes the years 825, 1296, 1387, 1656, 1749 and 1883 (Fig. 4a). These values are obtained from the kink point analysis conducted by Matyasovszky (2011). The coldest and the warmest intervals in the past centuries are conventionally labeled the Little Ice Age (LIA) and the Medieval Warm Period (MWP). There is no broadly accepted definition of these climate epochs (Bradley et al., 2003). Matyasovszky (2011) pinpoints two phases in the LIA: a stronger and longer period (1387-1656) and a weaker and shorter time interval (1749-1883). For Matyasovszky (2011), the MWP lies between years 795 to 1120. Our method detects 6 breakpoints (Fig. 4b). They show rather compelling (larger than 0.05) SNR values (they are ranging from 0.06 to 0.97). Nevertheless, a warmer temperature segment is found from 780 to 1119 and the possibility of three distinct phases in the LIA (1252-1444, 1444-1718 and 1718-1845) is not ruled out. The results from the RSMCPM agree well with the definition of the LIA proposed by Matyasovszky (2011). Moreover, our results are highly consistent with the definitions suggested by Yan et al. (2015). For these authors the beginning of the MWP is A.D. 800 and they suggest that the transition between MWP and LIA should be linked to the timing of 1246 A.D. being the last time in which the perihelion of the Sun-earth orbit coincided with Northern Hemisphere winter solstice.

The 3 comparisons show that the RSMCPM algorithm is able to yield useful and convincing results. The method provides 3 main advantages for landslide displacement rate analysis:

- It holds few assumptions, especially the non-Gaussian character of data does not affect results.
- The existence of a trend in the serial data is not a critical issue.
- SNR values are helpful in inferring hierarchies in change points.

The Séchilienne rock avalanche time series
4 RESULTS

The cumulative displacement curve from extensometer #13 at Sèchilienne (Fig. 5) shows very gradual daily variations: the average displacement rate is about 3.5 ± 2.9 (SD) mm/day. The median is 2.4 mm/day. The average rate is rather small compared to its variability. This high variability does not result from extreme values: the interquartile range, which is a resistant estimate of the dispersion, is still 3.2 mm/day for the Sèchilienne data, as to compare with the 3.5 and 2.4 mm/day average and median values. Extensometer measurements may be daily disturbed by the weather (wind, frost) and it is hardly surprising that these data show relatively large dispersion values. Daily variations of displacements are overall very gradual but the series show local abrupt changes (Fig. 5). In the daily displacement curve (Fig. 6a), 4 change points are identified, respectively located at 2008.400 (2008-05-26), 2010.812 (2010-10-24), 2012.728 (2012-09-23) and 2014.193 (2014-03-21). The SNR values for these change points are 1.01, 0.40, 0.49 and 3.63 respectively (Table 1). These results suggest that all the change-points are important (SNR values larger than 0.05). In order to smooth down the daily fluctuations we apply the RSMCPM to weekly (7 days) displacement data (Fig. 6a). In the weekly displacement curve (Fig. 6b and Table 2), 4 change points are identified, respectively located at 2008.400 (2008-05-26), 2009.914 (2009-11-30), 2009.971 (2009-12-21) and 2012.712 (2012-09-17). One of these change points is minor, as it shows an SNR value of 4.1 × 10−3 (2009.914) (Table 2). One must note that the change points from daily windows, with the 2 highest SNR values, do fit the one of the weekly window (May 2008, September 2012). As an aside, for daily window data, Shapiro-Wilk normality tests (Shapiro & Wilk, 1965) reveal that the data of each segment are unlikely to be Gaussian (the p-value for the tests are all smaller than 2 × 10−10). This finding provides sound justification for using the RSMCPM rather than a change point detection method that requires data normality.

The average rainfall rate is about 2.5 ± 4.8 (SD) mm/day. For this data set too, the variability is high compared with the mean value: the coefficient of variation (cv, ratio of the standard deviation to the mean) of the series is about 1.9. Values in the daily rainfall data series are varying quickly (Fig. 7a). A Wald-Wolfowitz test (Wald & Wolfowitz, 1943), designed to check independence and stationarity, confirms that the daily series is not stationary. Change points (or trends) are signs of nonstationarity. Nevertheless, local homogeneity in the time series, as evidenced by the Wald-Wolfowitz test, is required to allow their detection. This may explain the RSMCPM fails in detecting any change point in the daily rainfall series (the figure is not shown). In the weekly rainfall curve (Fig. 7b), the RSMCPM is successful and detects 5 change points located at 2007.289 (2007-04-16), 2007.864 (2007-11-12), 2009.875 (2009-11-16), 2012.253 (2012-04-02) and 2014.590 (2014-08-04) respectively (Table 2). 3 points are increases and 2 are decreases. Only one increase (2009.875), the smallest one, correlates in time with a displacement rate change. It suggests other control parameters for the 2008 and 2012 change points we resolved in weekly displacement rates.

Data from another extensometer (extensometer #16; see Fig. 1 for location) help to improve our understanding of the results: our procedure highlights similar change-points in the two daily and weekly displacement series for both extensometers (Fig. 6, Fig. 8 and Tables 1-2). This outcome provides further evidence that results of our procedure are robust. Nevertheless, the series for extensometer #16 show fewer change-points than the series for extensometer #13: for the daily series, only two minor change-points are detected, located in March 2014 (2014-03-08 and 2014-03-11). For the weekly series, the extensometer #16 show paramount change-points: they both show SNR values larger than 1 (1.17 and 1.59 respectively). The first change-point is perfectly coincident with one change-point of the weekly series for extensometer #13 (2009-11-30). The second change-point (2014-03-10) is concomitant with the change-points of March 2014 in the daily series. To summarize, 4 pivotal times are highlighted in the Sèchilienne displacement time series: May 2008, November-December 2009, September 2012 and March 2014. The strongest change points (SNR values greater than 1) are always common to several series (extensometers #13 and #16 or daily and weekly series): 2008.400, 2009.914, 2014.188 and 2014.193. The largest SNR values for these points are 1.01, 1.17, 1.59 and 3.63 respectively. Several change points in the daily and weekly series for the extensometer #13 (Table 1-2) may be considered as local features: 2010.812 (2010-10-24), 2012.712 (2012-09-17 in the weekly series) and 2012.728 (2012-09-23 in the daily series).

5 DISCUSSION AND CONCLUSION

For the three calibration datasets we perform, the RSMCPM is undoubtedly efficient in addressing properly the double challenge of the multiple change point problem: finding the optimal number and finding the location of change points. The procedure is effective even under adverse conditions: the first (synthetic) example shows that the procedure is successful in detecting
change points close to a series’ end and when there is a trend in the series. The second example illustrates how the technique is conducive in detecting change points in sparse series. The last example reveals that the method may be helpful when change points in a series are subject to various interpretations, as is the case for historical climate data. The main objective of the study is to apply the change point method to a landslide displacement and rainfall time series and to compare the results with previous analysis. For the time range spanning from 2003 to 2013, Chanut et al. (2013) labelled 5 different phases in the landslide displacement rate: 3 slow ones (called R2, R3 and R5, showing displacement rates about or less than 50 mm/month) and 2 fast ones (R4 and R6 showing displacement rates equal or larger than 100 mm/month). Chanut et al. (2013) set the start times of R3, R4, R5 and R6 on January 2005, 2008, 2009 and 2010, respectively (Figure 4 in Chanut et al. (2013)). No indication is given by Chanut et al. (2013) on the way these phases are identified. The RSMCPM partly mirrors the segmentation of the series proposed by Chanut et al. (2013) (Fig. 9, tables 2-3), as the two study periods do not overlap exactly (2003-2013 and 2004-2015). Our S1 slow stage covers the timespan of the Chanut et al. (2013)’s R2 and R3 slow phases. Our moderately fast S2 stage can be related to the R4 and R5 phases (however, any slow phase such as R5 is not highlighted by the RSMCPM during the S2 stage). The fast S3 stage begins almost at the same time as the R6 phase. Since the study series is longer than Chanut et al. (2013)’s and continues beyond 2013, we detect two additional stages: the very fast stage S4, beginning at the end of 2012 and the slow phase S5, beginning at the start of 2014. Figure 9 summarizes this stage analysis.

The RSMCPM does not have, as any statistical tools, the capacity to explain the origin of a shift in the study time series. Nevertheless, it provides information on key stages for temporal change in the landslide kinematics. Except for the 2009-11-16 weak breakpoint in the rainfall series, there is no relationship between the rainfall and the displacement breakpoints: the 2012-04-02 change point in the rainfall data highlights an increase of 39% for the biweight mean (Tukey, 1960) of rainfall but there is no equivalent change in displacement rate for the weekly series comparable to that in rainfall. The 2014-08-04 change-point in the rainfall data is located long after the 2014-03-10/2014-03-12 decrease in displacement rate (from 190-270 mm/month to 40-50 mm/month). Minor change points in the rainfall data like the 2007-04-16 and the 2007-11-12 discontinuities are not associated with any change point in the displacement series (Fig. 9, table 2-3). In the same way, the 2008-05-26 increase in displacement rate is observed with no significant changes in rainfall to be related to (Fig. 5-6). All this suggests that an increase of the rainfall is able to induce an increase of the landslide displacement rate (that is the case for the November-December 2009 breakpoint), but that most of the dynamics of the landslide is not directly attributable to rainfall increase/decrease.

The detailed exploration of the characteristics of the 5 kinematic stages (Fig. 10) suggests that the displacement rate is tied in with the rainfall intensity but clearly there is no one-to-one correspondence between the displacement rate and a given rainfall amount (as an example, points S2 and S3 show almost the same rainfall rates for very different displacement velocities in Fig. 10a). This is consistent with displacements weakly correlated with rainfall on Séchilienne site (Chanut et al., 2013; Klein et al., 2013).

The relation of the mean displacement rate with the rainfall frequency is more pronounced than with the rainfall intensity (Fig. 10a and b). In Figure 10a, point S3 (stage S3) shows a larger displacement rate for a lower rainfall intensity than point S2 (stage S2). The relation becomes monotonic for the frequency of rainy days (Fig. 10b). Except for the last stage (S5), it appears that the mean displacement rate is directly connected to the frequency of rainy days (Fig. 10b). This observation appears consistent with the previous finding that landslide kinematics is dependent upon not only rainfall but also soil moisture conditions (Ponziani et al., 2012; Brocca et al., 2016; Greco & Bogaard, 2016; Bogaard & Greco, 2017). The correlation between the frequency of rainfall events and soil moisture levels is an unremarkable fact and several environmental studies confirmed that soil moisture is more strongly related to precipitation frequency than precipitation amount (e.g. Piao et al., 2009; Wu et al., 2012). The last stage (stage S5) seems to be located in an unexpected place of our bivariate plot: it shows almost the same mean displacement rate than stage S1 for a quite larger frequency of rainfall events. We can guess that the fracture opening within the rockslide and the resulting fall in pore pressure hinders the displacement during this last stage.

A remarkable shift in the displacement rate from S2 to S3 is evident in Figure 10: The Séchilienne rock avalanche reached a new level of displacement rates in 2010. It is important to recognize that in September 2009, the French government declared a state of drought emergency for the administrative region in which the Séchilienne rock avalanche is situated. In this way, stage 2 is associated with the period billed as the driest period for the study time range (2004-2015). This
is consistent with our rain information during S2 showing both small rainfall and small rainy day frequency (Fig. 10). There can be questions about a possible non-linear response of the landslide to precipitation trigger after the drought. The change point that starts S3 stage is strong. The occurrence of a stage (S3) with very high displacement rates specifically just after the period of severe drought may be not without significance. Whereas Cappa et al. (2014) underlined a background seasonality of dry and wet seasons on displacement rates, our change-point analysis emphasizes the impact of an exceptional dry year on the landslide kinematics. As such the second part of 2009 is a period of weak brittle deformation (few quakes and rockfalls) as suggested by Figure 2 (page #28) in the SLAMS project’s final report (Garambois, 2014). Figure 11 clearly shows that the response of the landslide to precipitation trigger is stronger after the 2009 drought.

Rainfall is a well-identified cause of landslide triggering (Hufschmidt & Crozier, 2008; Tatard et al., 2010; Greco & Bogaard, 2016). For the Séchilienne landslide, we already know that rockfalls and micro-seismicity are correlated with a vanishing value, if any, for time lag with rainfall (Helmstetter & Garambois, 2010; Klein et al., 2013). Our comparison of rainfall and displacement rate breakpoints is deficient (it is likely that the RSMCPM may not pick up all the possible breakpoints in the rainfall series, as the Wilcoxon–Mann–Whitney test is implemented to reduce the risk of false positive), nevertheless it is noticeable that very intense rainfall events broke out:

- just before the 2008-05-26 increase in displacement, from May 24th to May 26th (64 mm collected over 3 successive days);
- on November and December 2009, in the time period of the 2009-11-30 and the 2009-12-21 increases in displacement rates. From November 28th to November 29th and from December 21st to December 24th, about 90 mm were collected (46 mm and 43 mm during each rainfall event respectively);
- on September 22nd, before the 2012-09-23 displacement rate breakpoint (about 100 mm rainfall were collected over 5 successive days, from September 22nd to September 26th, 2012).

This is confirmed in an Event Coincidence Analysis (ECA) performed using the CoinCalc R package (Siegmund et al., 2017). We do not fail to reject at the 0.05 level (the p-value for the precursor/trigger coincidence significance test, from randomly shuffled series, is 0.001) the null hypothesis that the observed number of coincidences between displacement breakpoints and heavy rainfall can be explained by two independent series of randomly distributed events (Fig. 12). The coincidence of the starts of displacement phases S2, S3 and S4 with heavy rainfall events is unquestionable (Fig. 12).

At any rate, our result are conclusive, as being able to quantitatively define stages in an apparent continuous divergence of the cumulative displacement. The ability to define stages in the 10 year duration of the acceleration of Séchilienne rock avalanche displacement (from 2004 to 2014) demonstrates the process is still susceptible to external forcing (e.g. rainy day frequency). The results obtained from the retrospective use of our change-point analysis show that averaging on different time windows helps in understanding the kinematics of the Séchilienne rock avalanche. The odds are pretty good this kind of innovative approach will be also profitable for studies of the kinematics of volcanic or seismic areas. When applied to the Séchilienne rock avalanche, the RSMCPM identifies and quantifies 5 kinematic stages in the 2003-2015 period with the following features:

- during these stages the displacement rate values correlate both with rainfall rate value and the rainy day frequency (Fig. 10). The observed relationships allows us to suggest that the frequency of rainy days is a key parameter that controls the displacement rates, more clearly than the rainfall amount does;
- the durations of stages are longer than 1.5 year, in accordance with the strong inertia of the rockslide displacement;
- the onsets of stages (accelerating step) always coincide with heavy rainfall episodes;
- the interrelationship between displacement and rainfall rates evolves over stages and time (Fig. 11). Accordingly, we suggest that the beginning of 2010 marks a crucial moment in the acceleration of the kinematics of the Schilienne rock avalanche;
- in March 2014 (onset of the last stage), the possible fracture opening within the rockslide and the resulting fall in pore pressure slows the rockslide back to displacement rates of 2003.

ACKNOWLEDGMENTS

We acknowledge OMIV, the CEREMA (Centre d’Etudes et d’Expertise sur les Risques, l’Environnement, la Mobilité et l’Aménagement) of Lyon and Sébastien Pierart (ROMMA association) for providing displacement and meteorological data (thanks to CEREMA to make the data available through the SNO-INSU OMIV portal: https://omiv.osug.fr/data.html). Graphics were performed using R (3.3.2) version statistical programming.
References


Siegmund, J. F., Siegmund, N., & Donner, R. V., 2017. Coincalca new r package for quantifying simultane-
ities of event series, Computers & Geosciences, 98, 64–72.
Yan, H., Soon, W., & Wang, Y., 2015. A composite sea surface temperature record of the northern South China Sea for the past 2500 years: A unique look into seasonality and seasonal climate changes during warm and cold periods, Earth-Science Reviews, 141, 122–135.
Figure 1. The Séchilienne rock avalanche location map. A- Location of the studied area in SW of France. B- Location of Séchilienne landslide on IGN RGEAlti 5m dem. C- SPOT 2014 image of the Séchilienne landslide. The extensometers #13 and #16 are displayed in white circles. The unstable area is in yellow line. The projection used is Lambert Conformal Conic 93 and the Geographic Coordinate System is the GCS RGF 1993.
Figure 2. Simulated independent Gaussian observations with changes in mean or variance a) Change point locations estimated by the rank sum method. Solid vertical lines indicate the true change point locations. Estimated change point locations are marked by numbers (1, 2 and 3). b) Change point locations estimated by the rank sum method when a 5% trend is added to the series in a). c) Change point locations estimated by the E-Divisive method. In this part of the figure, dashed vertical lines indicate the estimated change point locations. For details, please refer to Figure 1 in James & Matteson (2014). d) Change point locations estimated by the E-Divisive method when a 5% trend is added to the series in a).
Figure 3. Annual January to June streamflow amounts for the Romaine river in Quebec. a) Change point locations estimated by the Bayesian change point analysis. The “Posterior Means” displays the data along with the posterior mean of each position. The peak value of the posterior probability marks the change point location. For details, please refer to Erdman & Emerson (2007). b) Change point locations estimated by the rank sum method. Estimated change point locations are marked by numbers (1 and 2).
Figure 4. Northern Hemisphere temperature from proxy data, A.D. 200-1995. a) Change point locations estimated by the kink point analysis. A cubic smoothing spline fits the data (solid line) and reveals the kink points. Dashed vertical lines indicate the estimated change point locations. b) Change point locations estimated by the rank sum method. The gray shaded zones represent the Medieval Warm Period (MWP) and the Little Ice Age (LIA) as defined as the 800-1246 and 1246-1850 time intervals respectively.
Figure 5. Cumulative displacement plot at Séchilienne extensometer #13.
Figure 8. Displacement plots for the Séchilienne extensometer #16. a) Change point detected by the rank sum method in the daily displacement time series. The change points are respectively located at 2014.182 (2014-03-08) and 2014.190 (2014-03-11). b) Change point detected by the rank sum method in the weekly displacement time series. The change points are respectively located at 2009.914 (2009-11-30) and 2014.188 (2014-03-10).
Figure 9. The stage timeline. Filled triangles just above the x axis show the start times of the kinematic stages suggested by Chanut et al. (2013). The dashed line indicates the breakpoint sequence for the daily series of extensometer #13.
Figure 10. Bivariate plots of the relations between displacement rates and rain information for each displacement stage (S1-S5). Arrows follow the time sequence. a) Mean displacement rate-Mean rainfall. b) Mean displacement rate-Rainy day frequency. We define as a rainy day a day on which the rainfall is larger than zero.
Figure 11. Lagged scatterplot of displacement rates and rainfall (weekly series) for each displacement stage (S1-S5). The lag is $h = -1$ (7 days). For S3-S5 stages the fitting line (in red) is steeper than for S1-S2 stages (in blue). The black line is the regression line for all the points (S1-S5 stages).
Figure 12. Graphical comparison of the starts of the displacement phases (top) and the heavy rainfall events (bottom). Above the time arrow, the starts of displacement phases S2, S3 and S4 are marked (dates and inverted triangles). Below the arrow, each vertical bar marks a day of heavy rainfall: we have chosen to define a day of heavy rainfall as a day when rainfall daily amounts are more than 17.6 mm. This value is the 0.98 quantile of the distribution of the study rainfall data set. The coincidence of the starts of displacement phases with heavy rainfall events is unquestionable.
Table 1. Statistics of the change points for each series (daily bins). $\beta$ is for the "biweight mean".

<table>
<thead>
<tr>
<th>Data set</th>
<th>Date (decimal/ymd)</th>
<th>SNR value</th>
<th>Left $\beta$</th>
<th>Right $\beta$</th>
<th>Left SNR</th>
<th>Right SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ext 13</td>
<td>2008.400/2008-05-26</td>
<td>1.01</td>
<td>1.6</td>
<td>48.2</td>
<td>1.6</td>
<td>100.0</td>
</tr>
<tr>
<td>Ext 13</td>
<td>2010.812/2010-10-24</td>
<td>0.40</td>
<td>3.4</td>
<td>100.0</td>
<td>0.3</td>
<td>60.0</td>
</tr>
<tr>
<td>Ext 13</td>
<td>2012.728/2012-09-23</td>
<td>0.49</td>
<td>5.1</td>
<td>100.0</td>
<td>0.2</td>
<td>266.2</td>
</tr>
<tr>
<td>Ext 15</td>
<td>2014.182/2014-03-08</td>
<td>&lt;3.1 $10^{-5}$</td>
<td>4.0</td>
<td>122.6</td>
<td>3.7</td>
<td>111.5</td>
</tr>
<tr>
<td>Ext 16</td>
<td>2014.190/2014-03-11</td>
<td>0.63</td>
<td>5.7</td>
<td>113.5</td>
<td>1.4</td>
<td>41.9</td>
</tr>
<tr>
<td>Ext 15</td>
<td>2014.193/2014-03-12</td>
<td>3.63</td>
<td>8.7</td>
<td>206.2</td>
<td>1.6</td>
<td>48.2</td>
</tr>
</tbody>
</table>
Table 2. Statistics of the change points for each series (weekly bins). \( \beta \) is for the “biweight mean”.

<table>
<thead>
<tr>
<th>Date</th>
<th>Date</th>
<th>SNR ratio</th>
<th>Left ( \beta )</th>
<th>Right ( \beta )</th>
<th>Left ( \beta )</th>
<th>Right ( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rainfall</td>
<td>2007-05-01</td>
<td>0.00</td>
<td>12.5</td>
<td>12.0</td>
<td>59.7</td>
<td>21.6</td>
</tr>
<tr>
<td>Rainfall</td>
<td>2007-06-11</td>
<td>0.01</td>
<td>18.0</td>
<td>19.7</td>
<td>64.4</td>
<td>148.5</td>
</tr>
<tr>
<td>Eut 13</td>
<td>2008-04-08</td>
<td>0.47</td>
<td>11.2</td>
<td>19.1</td>
<td>95.1</td>
<td>31.4</td>
</tr>
<tr>
<td>Rainfall</td>
<td>2008-07-28</td>
<td>0.0007</td>
<td>15.7</td>
<td>16.4</td>
<td>71.4</td>
<td>148.5</td>
</tr>
<tr>
<td>Eut 13</td>
<td>2009-01-26</td>
<td>0.009</td>
<td>19.1</td>
<td>31.5</td>
<td>65.9</td>
<td>109.3</td>
</tr>
<tr>
<td>Eut 16</td>
<td>2009-01-30</td>
<td>0.13</td>
<td>19.5</td>
<td>23.5</td>
<td>63.9</td>
<td>109.3</td>
</tr>
<tr>
<td>Eut 13</td>
<td>2009-07-08</td>
<td>0.05</td>
<td>21.5</td>
<td>21.5</td>
<td>81.9</td>
<td>137.0</td>
</tr>
<tr>
<td>Eut 13</td>
<td>2012-01-10</td>
<td>0.05</td>
<td>15.8</td>
<td>22.6</td>
<td>93.9</td>
<td>137.0</td>
</tr>
<tr>
<td>Eut 16</td>
<td>2012-03-17</td>
<td>0.09</td>
<td>39.0</td>
<td>25.1</td>
<td>108.3</td>
<td>108.3</td>
</tr>
<tr>
<td>Eut 16</td>
<td>2014-03-10</td>
<td>0.79</td>
<td>106.3</td>
<td>9.9</td>
<td>108.3</td>
<td>108.3</td>
</tr>
<tr>
<td>Rainfall</td>
<td>2014-04-04</td>
<td>0.05</td>
<td>22.8</td>
<td>19.2</td>
<td>79.3</td>
<td>79.3</td>
</tr>
<tr>
<td>Stage</td>
<td>Starting time</td>
<td>Mean rate (mm/month)</td>
<td>Rate std (mm/month)</td>
<td>Rainy day Frequency</td>
<td>Duration (years)</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>---------------</td>
<td>----------------------</td>
<td>--------------------</td>
<td>---------------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td>49</td>
<td>17</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>May 2008</td>
<td>65</td>
<td>40</td>
<td>0.54</td>
<td>1.51</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>November-December 2009</td>
<td>104</td>
<td>45</td>
<td>0.57</td>
<td>2.80</td>
<td></td>
</tr>
<tr>
<td>S4</td>
<td>September 2012</td>
<td>276</td>
<td>182</td>
<td>0.66</td>
<td>4.00</td>
<td></td>
</tr>
<tr>
<td>S5</td>
<td>March 2014</td>
<td>52</td>
<td>33</td>
<td>0.55</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Kinematics stages in the last decade (displacement values are from the extensometer #13).