Hydrogeological threshold using effective rainfall and support vector machine (SVM) applied to a deep seated unstable slope (Séchilienne, French Alps)

Détermination d'un seuil hydrogéologique pour la déstabilisation profonde d'un versant (Séchilienne, Alpes) à partir des précipitations efficaces et des machines à vecteurs de support (SVM)

Journées 'Aléa Gravitaire' 2013

A. Vallet[1], C. Bertrand[1], D. Varron[2], J. Mudry[1]
[1]UMR6249 Chrono-Environnement, Besançon, France

[2]UMR6623 Laboratoire de Mathématiques, Besançon, France

ABSTRACT: Rainfall threshold is a widely used method for estimating minimum critical rainfall amount which can yield to slope failure. Literature reviews shows that most of the threshold studies are subjective and not optimal. In addition, for this study effective rainfall was considered for threshold definition. Support vector machine (SVM) and automatic events identification were used in order to establish an optimal and objective threshold for Séchilienne landslide. The method has been designed in order to be easily incorporated in a landslide warning system. Hydrogeological threshold results show similar performance and configuration for effective rainfall and raw rainfall. Accuracy of the two thresholds is comparable and the index combination antecedent/precedent is pretty much identical, 9D/4D and 10D/5D for effective and raw rainfall respectively. Accuracy of both thresholds defined for Séchilienne landslide (> 86%), makes it appropriate to be integrated in a landslide warning system.

1 INTRODUCTION

Rainfall threshold is a widely used method for estimating minimum critical rainfall amount which can yield to slope failure. Rainfall threshold can be defined either with empirical (statistical) or deterministic (physical based model) approach from local to worldwide scale (Terlien, 1998). To the best of our knowledge, no attempts of defining an empirical rainfall threshold to a deep seated unstable slope (>100m) have yet been successfully undertaken. Literature reviews shows that most of the threshold studies are subjective and not optimal, as thresholds are usually drawn visually or with poor mathematically/statistically bases (Guzzetti et al., 2007). Moreover, effective rainfall, which is the part of rainfall which recharges the aquifer, is relevant to consider instead of raw rainfall for deep seated landslides involving groundwater flow (Vallet et al., 2013). The aim of this study is to develop a new objective approach using effective rainfall to establish hydrogeological statistical threshold of a deep seated unstable slope. The method has been designed in order to be easily incorporated in a landslide warning system. Support vector machine (SVM) and automatic events identification were implemented for this study.

2 MONITORING NETWORK AND DATASET

The Séchilienne site is located on the external part of Belledonne crystalline range in the French Alps (mica-schist bedrock), south-east side of Grenoble city (France). The climate is mountainous with mean annual precipitations of 1200 mm. The Séchilienne site is a deep unstable slope on mica-shist bedrock. Precipitations are recorded at Mont Sec weather station (Météo France), including a rain gauge and a snow gauge. It allows estimating total precipitation (rainfall + snowfall), rainfall, snow cover, and snow melt in water equivalent. Effective rainfall was computed using a water balance method from precipitation, runoff and evapotranspiration with a soil available water storage of 35mm (Vallet et al., 2013). Displacement velocities of the Séchilienne unstable slope are monitored by several extensometers since 1990. The A13, A16 and C2 extensometers were selected as representative of the Séchilienne most active zone displacement. Daily precipitation, recharge, and displacement time series range from 1st January 1994 to 31 July 2012.

3 HYDROGEOLOGICAL THRESHOLD FOR DEEP SEATED LANDSLIDE

3.1 Threshold definition

Threshold matches with the separation boundary between rainfall conditions (rainfall index) which have and haven't caused slope destabilization. Threshold defines the minimum required amount of rainfall which yields to destabilization. Establishment of rainfall threshold is based on a two dimensional rainfall index. Literature presents various combination of index with the most common used are intensity-duration and daily-antecedent. Index selection depends mainly of study site settings (landslides type, climatic conditions, geomorphology, local or global threshold…).

Séchilienne unstable slope involves a large scale aquifer with water catchment by far larger than the unstable slope extension. Hydrosystem inertia and buffering properties are significant and smooth over short term events properties such intensity and duration. Unstable slope aquifer hydrodynamic response is then more influenced by antecedent rainfall which considers multiple rainfall events than a single event. Therefore for Séchilienne, threshold definition is based on antecedent and precedent rainfall index. Antecedent rainfall and precedent rainfall match with the accumulated amount of rainfall over a period of days before an event (unstable or not).

Complex structural geology and induced groundwater hydrodynamic of deep seated landslide involves generally a complicated hydro-mechanical relationship with rainfall (Berti et al., 2012; Terlien, 1998). For deep seated landslides, threshold deterministic approach, based on hydro-mechanical model, are challenging to implement due to rarity of direct hydrodynamic parameters measured. Therefore, in case of large amount of available data, definition of a statistical local threshold (specific to a landslide) is preferred as it can implicitly take into account these relationships.

3.2 Hydrogeological index

Even though, Terlien (1998) has shown the importance of period extension selection in threshold definition, in literature, choice of period extension of antecedent rainfalls is mainly explored empirically, with no optimization to find the best number of days.

Hydrogeological threshold was estimated for all combinations of antecedent and precedent period varying from 1 to 60 days and from 1 to 10 days, respectively yielding to 545 combinations for each dataset (effective and raw rainfall). The two periods, antecedent and precedent, do not overlap and are always adjacent (Figure 1). SVM performance was used to assess the best combination of precedent/antecedent period extension which maximizes the discrimination between low and high instability events for both raw and effective rainfall.

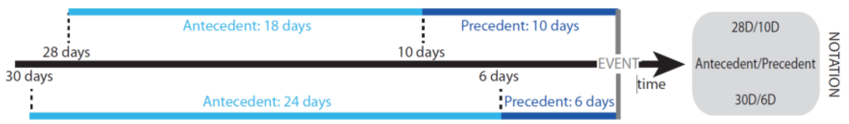

Figure1: Scheme of combinations of antecedent and precedent period for hydrogeological index and associated notation

4 EVENTS IDENTIFICATION

Although deep seated landslide destabilization is controlled mainly by rain triggered (short term component) ,time-dependent factors (long term component) such as rock weakening, slope groundwater permeability and connectivity modifications can be significantly influenced the destabilisation (Berti et al., 2012). Long term displacement monitoring, of the three extensometers, shows that displacement rate and amplitude have significantly increased with time and follows an exponential trend (Figure 2). Rainfall data series doesn't show any trend over the year, meaning that displacement trend is independent of the recharge amount. On Séchilienne, detrended displacement seasonal variations are clearly linked to the hydrocyle and to the recharge amount (Vallet et al., 2013). Observed trend is interpreted as the consequences of a progressive deterioration of rock mechanical properties (weakening) due to long term repetitive stress which has yielded to permanent deformation.

For this study, no stable or unstable events have been identified as Séchilienne unstable slope is constantly moving. Furthermore due to the long term increasing trend, it was not possible do define absolute landslide response from rainfall trigger. Therefore, choice has been made to define a critical statistical threshold which defines minimal amount of rainfall (raw or effective) which yield to a significant increase of slope destabilization (relative and not absolute variation) independently from the trend. Thus, displacement trend which is inherent to mechanic rock slope weakening was removed and not taken into account for events identification.

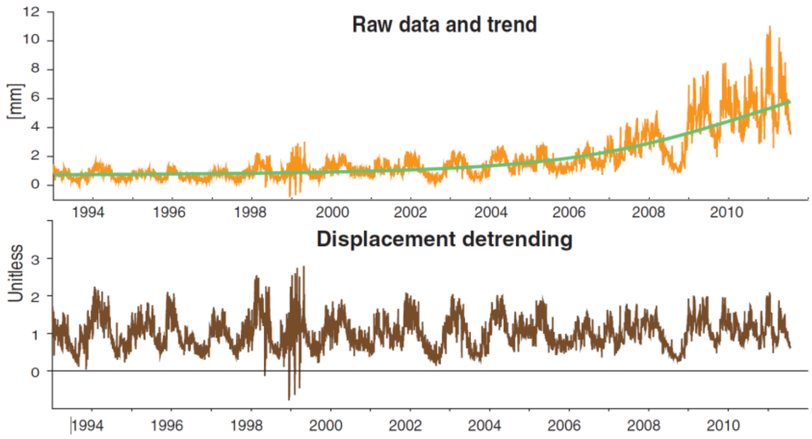

Figure2: Illustration of the displacement trend with extensometer A13 where raw displacement is in orange, exponential trend is in green and detrended displacement is in brown

4.1 Automatic events detection

Two types of events were defined, low (LDE) and high (HDE) destabilisation events and were assimilated respectively as local time series minima (Valley) and maxima (Peak). Most of peaks of displacement data series are asymmetrical with steep increasing side followed by a soft decreasing side (inverse for valleys), revealing system memory and reactivity. Neighbouring points of the peak/valley are then defined by a moving window cut into two parts back and front (named neighbour window). In addition, to take into account peak/valley width in the detection process, event location is also defined by a second moving window cut into two parts, included and centred on the first one, rather than only one point (named event window). The neighbour window does not overlap the event window and they have both the ability to be asymmetrical to fit to signal pattern. If event window average is superior to a general threshold and is superior to average of both neighbour window parts added with their own local threshold, then maximum value of event window is considered as a peak (figure2). For valley detection, inverse reasoning was adopted. General threshold is set as the data series third and second quartile, for peak and valley respectively. Because extensometers are a local measurement, an event is considered as a landslide representative event only if it is identified on the three extensometers selected with a shift less than 10 days (spatial and time coincidence). Event date is then defined as the barycentre of the three identified events.

4.2 Method calibration

This method has 6 parameters to estimate (three for each peak/valley side: local threshold (t_1, t_2), neighbour window half-width (w_1, w_2), event window half-width (ω_1, ω_1) which can yield to complex manual calibration (Figure 2). To overstep this constrain, a supervised learning method was implemented to calibrate the input parameters. Year 2001 to 2004 of the studied interval was chosen to train the algorithm. For these four years, events were identified manually based on the three extensometers measurements (Figure 2). Events detection method input parameters were estimated thanks to an algorithm of optimization by maximizing, (i) the proportion of automatic events matching with manual events and minimizing (ii) the date average difference between manual and automatic matching events and minimizing (iii) the proportion of automatic events which do not match with manual events. Simulated annealing, which is a global optimization method, was used for minimization process. Method was calibrated separately for peak and for valley detection.

5 OBJECTIVE AND OPTIMAL THRESHOLD DEFINTION: SUPPORTS VECTOR MACHINE

Establishment of a rainfall threshold requires defining a linear frontier between two categories of two dimensions points, labelled as stable or unstable, where misclassification is authorized for few points of the dataset. There are infinite possibilities of linear lines which can separate two classes of points. Support vector machines (SVM) are a widely used two-class linear classifier, belonging to supervised learning models and kernel methods (i.e. dot products) (Hastie, 2009).

5.1 SVM

Soft margin SVM formulation, by allowing some points to be in the margin or even to be misclassified in order to achieve a greater margin, is adapted to threshold context. A soft margin cost parameter (C) need to be specified by the user. Small values of C will produce a classifier with larger margin but will allow a higher proportion of misclassified sample data. Optimal hyperplane is obtained by solving the optimization problem of equation 4. The effectiveness of SVM classification depends on calibration of the soft margin parameter. SVM analysis was performed with the Matlab® package LIB-SVM (Chang and Lin, 2011). Data scaling (or normalization) is strongly recommended to enhance SVM performance and is performed for this study.

5.2 Calibration and performance

Cross-validation is a statistical method which assesses model performance, specially adapted to learning predictive model such as SVM (Hsu et al., 2003). Cross-validation assesses the accuracy of the model to predict new data by estimating the average classification rate in percent, i.e. for instance if 55 points are well classified on a total of 60, then accuracy is equal to 91.5%. Cross-validation is also recommended for SVM calibration and can avoid model overfitting (Hsu et al., 2003). For our purpose, leave-one-out cross-validation was chosen.

6 RESULTS AND DISCUSSIONS

6.1 Automatic events identification: calibration

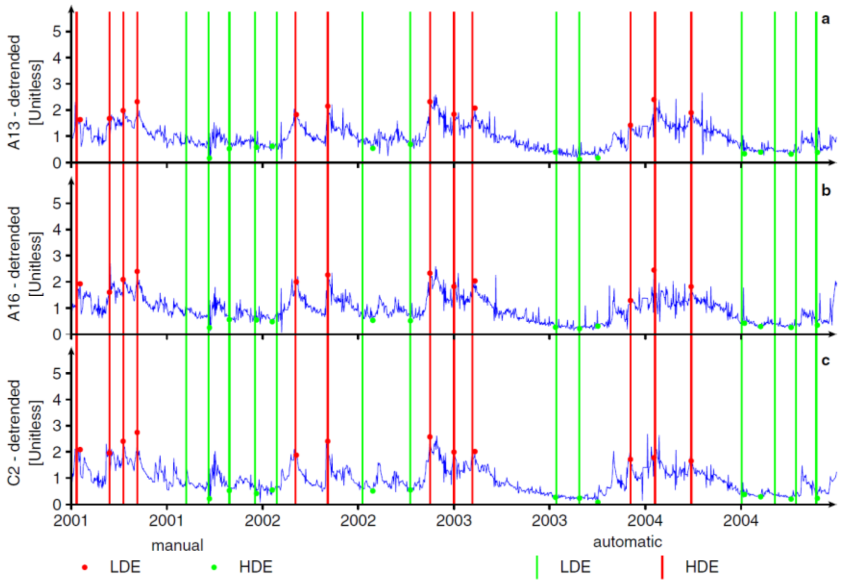

Figure 3: Calibration results of automatic of automatic events detection with (a) extensometer A13, (b) extensometer A16 and (c) extensometer C2. LDE (Low destabilization events) and HDE (High destabilization events)

For HDE, 13 events were automatically detected on the 13 manually identified (100%) and 0 events were detected whereas no manual events were identified (Figure 3). For LDE, 10 events were automatically detected on the 13 manually identified (77%) and 3 events were detected whereas no manual events were identified (Figure 3). Performances of calibration is better for HDE as peak time series patterns are better constrained (sharp width and amplitude: short wave) having one HDE per peak whereas multiple LDE can be identified for one valley pattern (long wave).

6.2 Automatic events identification: results

57 high destabilisation events (HDE) and 55 low destabilisation events (LDE) were identified for the whole studied interval on the extensometers detrended displacement and are plotted on the raw displacement time series (Figure 4). Events detections are relatively well spread on the time series extension, although for some years (1997, 1999, 2003 and 2006) only few events are identified. Number of identified events for each class is balanced and the total events number is sufficient to perform SVM analysis.

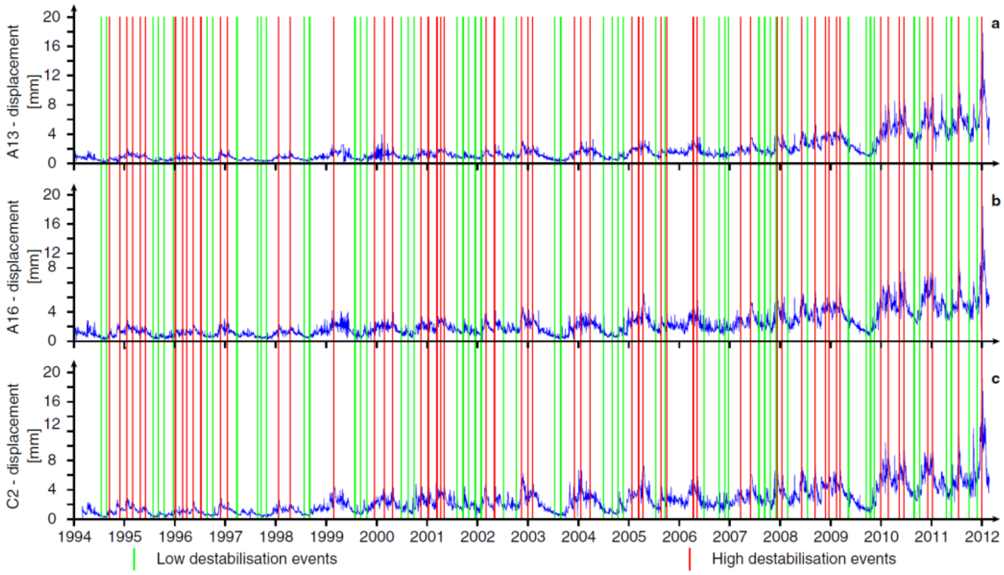

Figure 4: Results of automatic events detection for extensometers (a) A13, (b) A16 and (c) C2.

6.3 Hydrogeological threshold

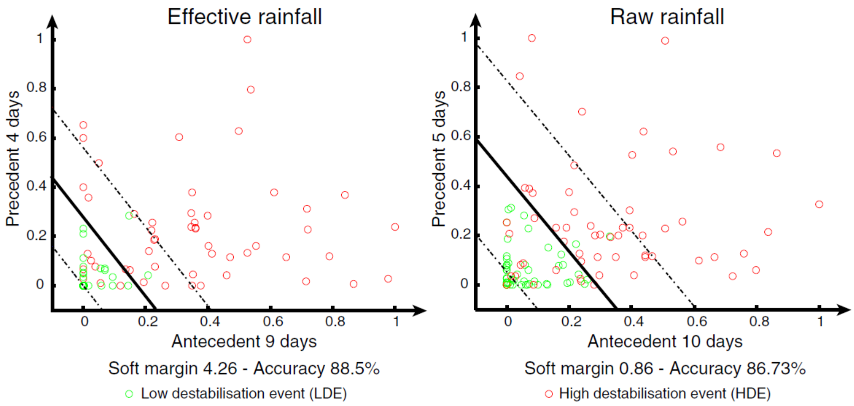

Figure 5: Results of best SVM performance for effective rainfall and raw rainfall. Both dataset are scaled from 0 to 1.

Among the antecedent/precedent combinations tested, hydrogeological threshold results, for effective and raw rainfall, show similar performance and configuration. Accuracy of the two thresholds is comparable, 88.5% and 86.73% and the index combination antecedent/precedent is pretty much identical, 9D/4D and 10D/5D, for effective rainfall and raw rainfall respectively (Figure 5). Accuracy of both thresholds is sufficiently high to be then integrated as a tool in an early warning system. Both thresholds only require a maximum of 10 days (on the 60 days tested) to discriminate low from high destabilisation stages.

Although, Vallet et al. (2013) have shown that seasonal variation of Séchilienne displacement signal is rather correlated to effective rainfall than raw rainfall, effective rainfall does not improve significantly the threshold performance. Indeed, with a soil available water storage of 35 mm and a evapotranspiration potential daily average of 2 mm, sum on a low extension period (here ≤ 5 days for each index component) yield to low difference between effective and raw rainfall. In addition, establishment of a hydrogeological threshold does not take into account amplitude of destabilisation as it only classifies binary data labelled. Results show that the precipitations, until 10 days have a great influence on the occurrence or not of a destabilisation stages. This short term component shows the high reactivity of Séchilienne slope to water input with a fast transit time. Nevertheless, effective rainfall seems to be the best water input variable to predict displacement amplitude fluctuations (Chanut et al., 2013) which are more dependent on the long term component of hydrosystem inertia (saturation state). Destabilisation increasing could be the consequences of a high satured state of hydrosystem simultaneous with high rainfall events. This assumption, deduced from these study preliminary results, will have to be confirmed by further analyses.

7 CONCLUSION

SVM and automatic events detection have produced good results for the establishment of a hydrogeological threshold for Séchilienne landslide. Accuracy of threshold defined for Séchilienne landslide, makes it appropriate to be integrated in a landslide warning system. Effective rainfall does not improve significantly threshold performance and raw rainfall seems to be the best variable to use as no computation is required. Further analyses are required in order to confirm these study preliminary results.

8 ACKNOWLEDGMENTS

This research was fund by the program SLAMS (Séchilienne Land movement: Multidisciplinary Studies) of the Agence Nationale de la Recherche. This work was performed in partnership with OMIV and OSU THETA. The meteorological data used were provided by MétéoFrance, EdF and CETE Lyon. Displacement data were supplied by CETE Lyon. The authors gratefully acknowledge the support of Marie-Aurélie Chanut and Jean-Pierre Duranthon from the CETE Lyon.

9 REFERENCES

Berti, M., Martina, M. L. V., Franceschini, S., Pignone, S., Simoni, A. and Pizziolo, M.: Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach, J. Geophys. Res. Earth Surf., 117(F4), n/a–n/a, doi:10.1029/2012JF002367, 2012.

Chang, C.-C. and Lin, C.-J.: LIBSVM: A library for support vector machines, ACM Trans. Intell. Syst. Technol., 2(3), 1–27, doi:10.1145/1961189.1961199, 2011.

Chanut M-A, Dubois L, Duranthon J-P, Durville J-L, (2013), Mouvement de versant de Séchilienne : relations entre précipitations et déplacements, Actes de First International Conference on Landslides Risk, Tunisie, 14-16 mars 2013

Guzzetti, F., Peruccacci, S., Rossi, M. and Stark, C. P.: Rainfall thresholds for the initiation of landslides in central and southern Europe, Meteorol. Atmospheric Phys., 98(3-4), 239–267, doi:10.1007/s00703-007-0262-7, 2007.

Hastie, T.: The elements of statistical learning: data mining, inference, and prediction, 2nd ed., Springer, New York, NY., 2009.

Hsu, C.-W., Chang, C.-C. and Lin, C.-J.: A Practical Guide to Support Vector Classification, Department of Computer Science, National Taiwan University. [online] Available from: http://www.csie.ntu.edu.tw/cjlin/papers.html, 2003.

Terlien, M. T. J.: The determination of statistical and deterministic hydrological landslide-triggering thresholds, Environ. Geol., 35(2-3), 124–130, doi:10.1007/s002540050299, 1998.

Vallet, A., Bertrand, C. and Mudry, J.: Effective rainfall: a significant parameter to improve understanding of deep-seated rainfall triggering landslide - a simple computation temperature based method applied to Séchilienne unstable slope (French Alps), Hydrol. Earth Syst. Sci. Discuss., 10(7), 8945–8991, doi:10.5194/hessd-10-8945-2013, 2013.