

Landslides (2021) 18:1191–1207
 DOI 10.1007/s10346-020-01599-w
 Received: 28 September 2019
 Accepted: 17 November 2020
 Published online: 5 January 2021
 © The Author(s) 2020

Gaetano Pecoraro · Michele Calvello

Integrating local pore water pressure monitoring in territorial early warning systems for weather-induced landslides

Abstract A methodology designed to integrate widespread meteorological monitoring and pore water pressure measurements is proposed. The procedure is tested in 30 hydrological basins highly susceptible to weather-induced landslides in Norway. The following data are used: a catalog of 125 weather-induced landslides in soils registered between January 2013 and June 2017, widespread meteorological monitoring data employed in a territorial warning model, and pore water pressure measurements retrieved from boreholes installed for a variety of geotechnical projects. The territorial warning model is initially applied to identify the warning events and the correspondent warning level in the test areas over the analysis period. Afterwards, a method for assessing the territorial warning events by analyzing the trends of the monitored pore water pressures is proposed. Finally, an augmented territorial warning model is calibrated and validated using statistical indicators widely adopted in literature. The analysis of the results reveals a satisfactory correspondence between days with landslides and the warning levels provided by the augmented territorial warning model. A final comparison between the results of the model calibration and the model validation highlighted the consistency of the model performance, once the three model parameters are adequately set.

Keywords Landslide risk · LEWS · Groundwater · Augmented warning model

Introduction

Weather-induced shallow landslides affecting superficial loose soil deposits with depth up to 2–3 m are a global natural phenomena occurring in many regions of Italy (Calvello and Pecoraro 2018), in Norway (Krøgli et al. 2018), in Switzerland (Springman et al. 2013; Fan et al. 2017), in Austria (Dietrich and Krautblatter 2017; Canli et al. 2018), in the USA (Baum and Godt 2010; Staley et al. 2013; Mirus et al. 2018), and in many parts of China (Lin and Wang, 2018; Zhang et al. 2018). The future frequency and the intensity of these weather-induced landslides are anticipated to increase due to climate change and the continuous urbanization of areas highly exposed to landslide risk (Barla and Antolini 2016). The main triggering mechanism for these landslides typically consists of rain or snow infiltration in shallow soil layers, resulting in an increase of the pore water pressure and a decrease of the soil shear strength (Caine 1980), as well as in the reduction of the matric suction in partially saturated soils (Sorbino and Nicotera 2013). This triggering process is characterized by a near total absence of warning signs. Although the mobilized volume is often limited in the triggering phase, in the propagation phase, the landslide mass may incorporate the soil material lying along the slope, significantly increasing its volume; and in steep channels, shallow slides may evolve into debris flows, destructive phenomena

characterized by extremely high velocities (Hungr et al. 2014). Shallow weather-induced landslides can occur frequently and often simultaneously over large areas; thus, they represent a widespread risk for local communities, structures, and infrastructures in many parts of the world (Petley 2012; Kirschbaum et al. 2015; Haque et al. 2016; Froude and Petley 2018).

Operational landslide early warning systems (LEWS) aim to reduce loss-of-life by informing exposed individuals, communities, and organizations, in order to enable them to act appropriately and in sufficient time (UNISDR 2006). LEWS have been recognized as important tools for landslide risk reduction and community resilience in many recent international initiatives (e.g., Sendai Framework for Disaster Risk Reduction 2015–2030, UN Agenda 2030 for sustainable development, European Climate Adaptation Platform). LEWS can be designed and managed at two different scales of operation: local systems (Lo-LEWS) address single landslides at slope scale (Pecoraro et al. 2019a); territorial systems (Te-LEWS) deal with the possible occurrence of multiple landslides over wide areas at regional scale (Piciullo et al. 2018). Te-LEWS are used to provide generalized warnings to authorities, civil protection personnel, and the population over appropriately defined homogeneous warning zones. Typically, these systems address weather-induced landslides through the monitoring and prediction of meteorological parameters (Segoni et al. 2018). It should be noted that the term weather is herein preferred to the term rainfall to refer to the broad spectrum of weather atmospheric time-dependent variables that may trigger landslides, such as intensity, duration, antecedent, and cumulated precipitation; snow cover and snow melting associated with temperature changes; and evapotranspiration and rainfall infiltration rates associated with atmospheric and soil conditions.

Meteorological monitoring alone does not consider critical soil properties controlling the initiation of the triggering process. Depending on these conditions, landslides may be triggered in response to a large variety of weather events. Therefore, although the integration of geotechnical monitoring (e.g., pore water pressure, soil water content, ground deformation) within warning models for weather-induced landslides may be very challenging, it can provide additional information to determine the likelihood of rainfall events actually triggering failure (Staley et al. 2013). On this issue, Stähli et al. (2015) proposed to combine local measurements of landslide precursors with model predictions at catchment scale. Recently, Calvello (2017) and Pecoraro et al. (2019b) suggested that local observations may provide fruitful indications on the local conditions within a wider warning zone. However, a standardized methodology for integrating geotechnical observations collected at local scale in a territorial warning model for weather-induced landslides still does not exist.

This study attempts at improving the prediction accuracy of a territorial warning model for weather-induced landslides using in-

situ pore water pressures data, and not only widespread hydrometeorological data. To this purpose, the proposed procedure is tested on the Norwegian national LEWS between January 2013 and June 2017, considering 30 hydrological basins as test areas.

Methodology

The conceptual framework

The procedure presented herein aims at integrating widespread meteorological monitoring data already adopted in a territorial warning model (TeWM) within an augmented territorial warning model (ATeWM) using local pore water pressure measurements. It builds upon the “strengths” of a territorial warning model already operative, which is based on widespread hydrometeorological data (i.e., relative water supply and relative soil saturation), and combines this with data from distributed positive pore water pressure measurements. To this aim, multiple in-situ measures must be available within any considered warning zone. Indeed, this ensemble of measures, processed and taken together as described in “Definition of the augmented territorial warning model” section, should be considered a proxy of significant weather-induced changes in the groundwater regime. The latter relates to possible shear strength changes along potential failure surfaces, and thus to possible variations in the factor of safety of potentially unstable soil volumes. The proposed procedure comprises three successive steps (Fig. 1): identification of the Representative Territorial Units, RTUs (phase I), application of the territorial warning model (phase II), and calibration and validation of the augmented territorial warning model (phase III). In phase I, the area of analysis is divided into territorial units of appropriate extent (i.e., the warning zones adopted in the model) considering the scale of analysis,

so that information from widespread monitoring and local observations can be combined. The territorial units are classified using two criteria: the occurrence of landslide events during the period of analysis and the availability of relevant information from monitoring instruments in the proximity of the landslide source areas. Following this classification, RTUs are identified. In phase II, the territorial warning model is applied in the selected areas. For each RTU, the warning events are identified, and the warning levels defined. In phase III, the calibration and the validation of the augmented territorial warning model are performed starting from the warning events issued by the territorial model. For calibration purposes, the pore water pressure observations are preliminarily analyzed to determine potential upward or downward trends. Then, the most appropriate parameters are identified, and the model is validated using statistical indicators derived from contingency tables.

Definition of the augmented territorial warning model

Pore water pressure measurements are analyzed to determine significant uptrends (and downtrends) that indicate local conditions which may lead (or not lead) to the triggering of a landslide within a territorial unit during a given weather event. Pore water pressure records are typically characterized by significant short-term variability and consequently, need to be statistically processed in order to smooth the short-term fluctuations and make potential trends identifiable. Moving averages are simple and common smoothing techniques widely used in time series analysis. They determine new time series, whose values are the average of a given number of observations in the original time series. Several types of moving averages—e.g., simple moving average (SMA), cumulative moving average (CMA), weighted moving

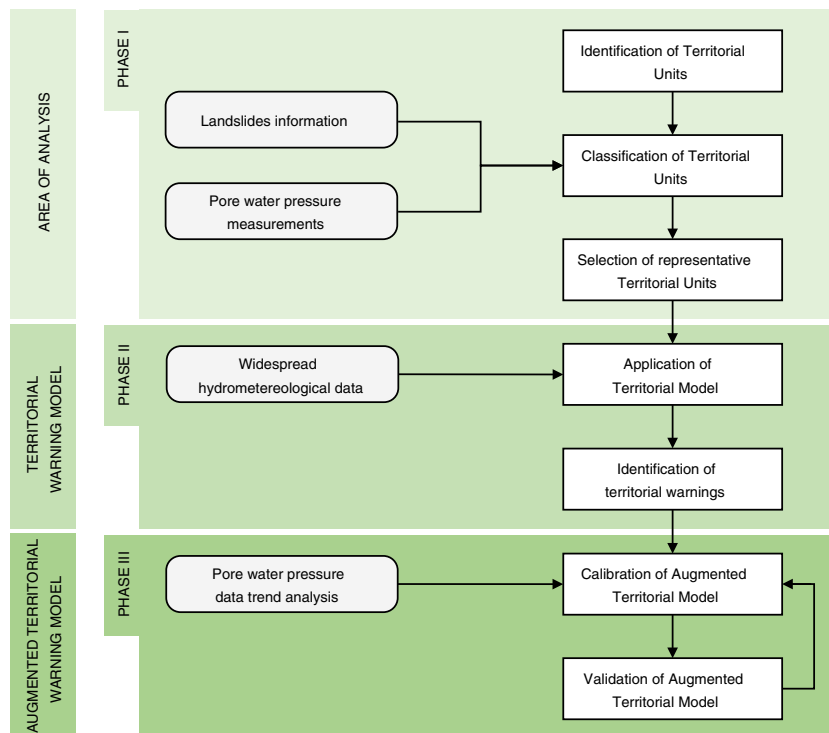


Fig. 1 Flowchart of the methodology proposed for the definition of an augmented territorial warning model for weather-induced landslides

average (WMA), and exponential moving average (EMA)—can be used, depending on the purpose of the analysis, the types of data, and the time periods.

In this study, the SMA of the recorded pore water pressures at a given day (u_i) is calculated for each piezometer, over the number of days of a specified time period (n), as follows:

$$u_i = \sum_{k=i-n+1}^i \frac{p_k}{n} \quad (1)$$

where p_k is the pore water pressure recorded at day k . Using these new time series, two indicators of pore water pressure variations are defined (Fig. 2), as follows:

$$\Delta u_i = u_i - u_{i-n} \quad (2)$$

$$\Delta u_i^* = \frac{u_i}{u_{i_{\max}}} \quad (3)$$

where Δu_i is the difference between the SMA calculated considering a length equal to n days and referring to days i and $i-n$; and Δu_i^* is the same difference normalized by $\Delta u_{i_{\max}}$, the maximum difference calculated over the time frame where pore water pressure records are available for each piezometer (L_1 , L_2 , and L_3 for the three piezometric datasets reported in Fig. 2).

The augmented territorial warning model is based on a 2-step procedure developed using the above-defined indicators (Fig. 3). In the first step, the differences between the simple moving averages referring to days i and $i-n$ are evaluated. In case they do not show a clear trend (i.e., the number of downtrends is equal to the number of uptrends in the datasets), the warning level issued by the territorial model (WL_{te}) remains unchanged. Otherwise, a second step is performed wherein the mean of the normalized simple moving average differences is calculated considering all the available piezometers, and this ensemble indicator is compared with predefined thresholds (i.e., a lower threshold, LT and an upper threshold, UT). Three final outcomes are possible: the warning level remains the same or is increased or decreased. No more than two warning level variations are allowed.

The calibration and the validation of the augmented territorial warning model are performed using statistical indicators and following procedures widely adopted in literature (e.g., Martelloni et al. 2012; Gariano et al. 2015; Calvello and Piciullo 2016). In a preliminary phase, the warnings issued and the landslides occurred are reported in a 2×4 correlation matrix, where the rows indicate the occurrence of the landslides (yes, no) and the columns, the corresponding occurrence of the warning levels (WL_1 , WL_2 , WL_3 , WL_4). It should be noted that the relative importance assigned to the different types of errors by the system managers is a key issue to consider to properly validate the model. Figure 4 reports a graphical representation of a comprehensive analysis of the correlation matrix based on a set of two performance criteria, both assigning a meaning to all the elements of the matrix.

The “alert classification” criterion (Fig. 4a) employs a classification scheme derived from a standard contingency table, and identifies correct alerts (CA), false alerts (FA), missed alerts (MA), and true negatives (TN). The presence of one of the three highest levels of warning (WL_2 , WL_3 , and WL_4) concurrently with the occurrence of at least one landslide is assumed as CA. FA and MA are incorrect predictions of the system: the first is related to the issuing of any of the three highest levels of warning (WL_2 , WL_3 , and WL_4) and the simultaneous absence of a landslide; the second refers to the occurrence of a landslide without any warning (WL_1). TN represents the absence of both landslides and warning occurrences. The “grade of accuracy” criterion (Fig. 4b) assigns a color code to the components of the correlation matrix in relation to the agreement between a given warning event and a given landslide event. For instance, if one of the two highest WLs is issued (i.e., WL_3 or WL_4) and no landslides occur, this could be considered a significant error of the warning model. Using this criterion, the elements are classified in three color-coded classes, as follows: green (Gre) for the elements which are assumed to be representative of the best model response, yellow (Yel) for elements representative of minor model errors, red (Red) for elements representative of significant model errors. Starting from the two performance criteria, several performance indicators can be derived (e.g., Piciullo et al. 2017a). Table 1 lists the indicators considered in this study.

Case study

Norwegian national LEWS

Norway is divided into 18 counties and 422 municipalities and covers an area of $\sim 385,000$ km² on the western and northern part of the Scandinavian Peninsula. Approximately 30% of the land area consists of mountains (with an average elevation of 460 m a.s.l.) and 6.7% of the country is covered by steep slopes. In geological terms, Norway is situated along the western margin of the Baltic shield covered by Caledonian nappes in the west. The bedrock of the Baltic shield is dominated by Precambrian basement rocks (e.g., granites, gneisses, amphibolites, and meta-sediments) in the southern and south-eastern part of the country (Fredin et al. 2013). Continuous till deposits cover large areas of the valley sides and floors, although fluvial and glaciofluvial deposits as well as marine clays are widespread.

Steep slopes, the presence of weak, loose sediments and favorable climatic conditions provide a basis for the triggering of a range of deep and shallow weather-induced landslides. These landslides are typically triggered by rainfall and snowmelt, or their combination, resulting in intense or long-duration water supply and high soil water content. Although the steep slopes covered by quaternary loose sediments are highly susceptible to landsliding, failures can also occur in gentle slopes covered by snow as well as in embankments along roads and railways. In addition, some failures are either triggered from or initiated as rockfalls or slush flows and may develop into debris flows as they propagate down-slope (Piciullo et al. 2017b). There are limited comprehensive estimates of human and economic losses related to shallow landslides in Norway (Krøgli et al. 2018). According to Furseth (2006), at least 230 fatalities can be associated to slope failures over the last 500 years. A recent report prepared by Haque et al. (2016)

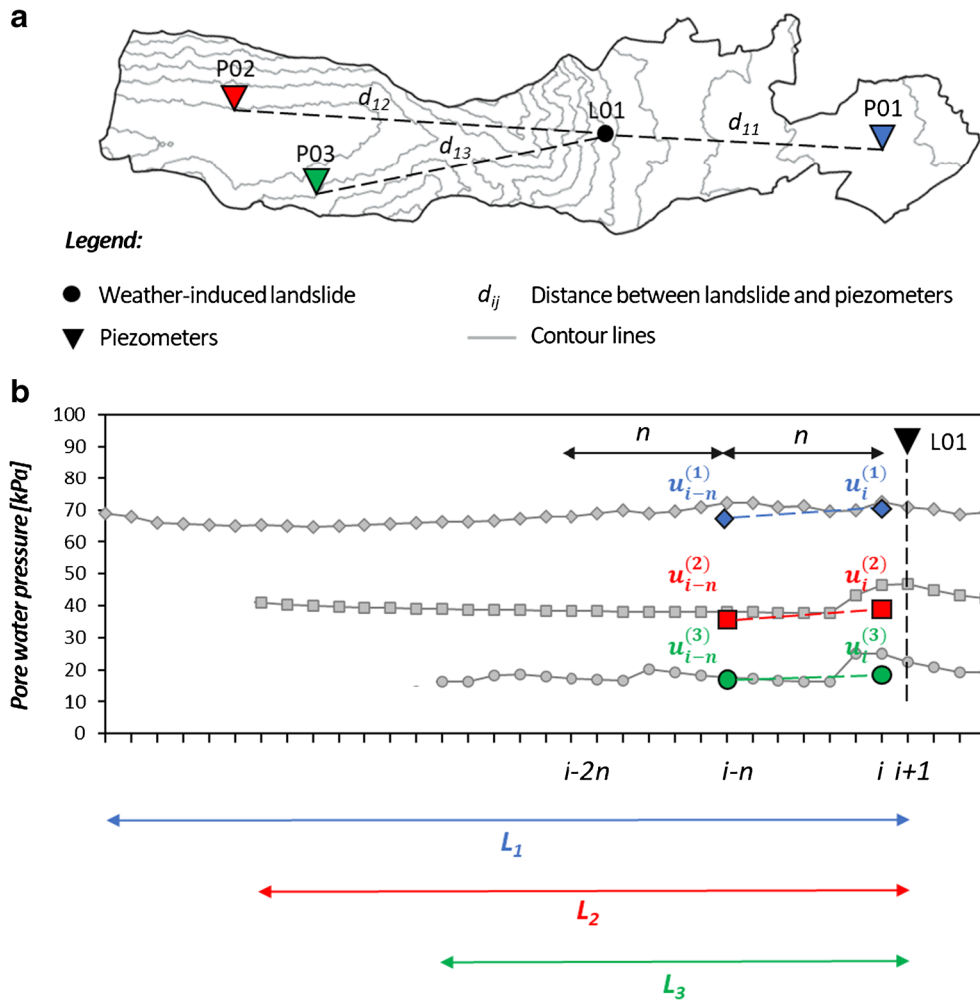


Fig. 2 Example of application of the proposed methodology to a test area: identification of weather-induced landslides and available piezometers (a) and pore water pressure trend analysis before a landslide event (b)

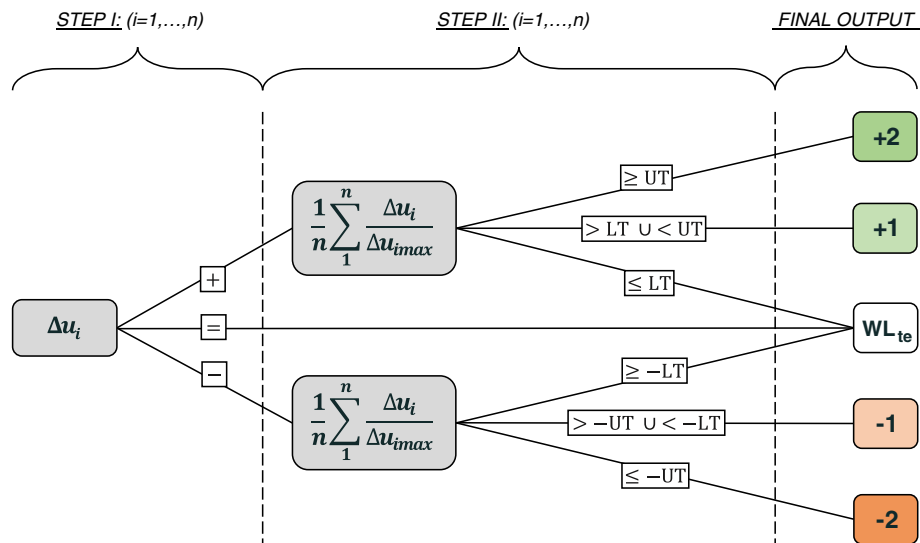


Fig. 3 Scheme of the procedure developed for analyzing pore water pressure observations. The numbers to the right indicate the change from the original warning level (WL_{te}) due to the augmented territorial warning model

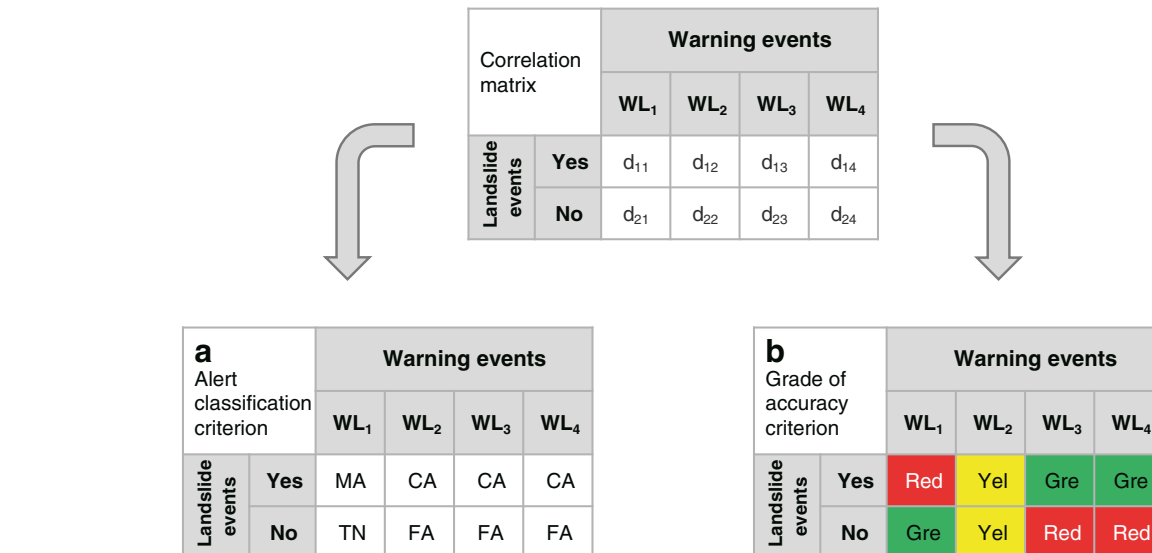


Fig. 4 Correlation matrix adopted to calibrate and validate the augmented territorial warning model: alert classification (a) and grade of correctness (b) performance criteria used for the analyses (modified from Calvello and Piccullo, 2016)

showed that 12 people died in Norway in the period 1995–2016 because of weather-induced landslides (slush flows in 7 cases and debris flows or debris avalanches in the remaining 5 cases). Economic consequences are mainly associated with disruption of road

Table 1 Performance indicators used for the analyses

Performance indicator	Symbol	Formula	Criterion	Range	Best value
Error rate (a)	ER _a	$\frac{MA+FA}{\sum_{ij} d_{ij}}$	a	[0, 100%]	0%
Error rate (b)	ER _b	$\frac{Yel+Red}{\sum_{ij} d_{ij}}$	b	[0, 100%]	0%
Missed alert rate	R _{MA}	$\frac{MA}{MA+CA}$	a	[0, 100%]	0%
False alert rate	R _{FA}	$\frac{FA}{FA+CA}$	a	[0, 100%]	0%
Index of severity of MA and FA	IS _{MA&FA}	$\frac{MA_{Red}+FA_{Red}}{MA+FA}$	a,b	[0, 100%]	0%
Hanssen and Kuipers (1965) skill score	HK	$\frac{CA}{CA+MA} - \frac{FA}{FA+TN}$	a	[− 100, 100%]	100%
Matthews correlation coefficient	MCC	$\frac{CA \times TN - FA \times MA}{\sqrt{(CA+FA)(CA+MA)(TN+FA)(TN+MA)}}$	a	[− 100, 100%]	100%
Euclidean distance (ROC classification)	δ _{ROC}	$\sqrt{\left(\frac{CA}{CA+MA} - 1\right)^2 + \left(\frac{FA}{FA+TN} - 0\right)^2}$	a	[0, 1]	0
Euclidean distance (precision-recall classification)	δ _{P-R}	$\sqrt{\left(\frac{CA}{CA+MA} - 1\right)^2 + \left(\frac{CA}{CA+FA} - 1\right)^2}$	a	[0, 1]	0

and railway networks. According to the Intergovernmental Panel on Climate Change (Hanssen-Bauer et al. 2017), the number of annual landslide events in Norway is expected to increase as Northern Europe will probably experience higher intensity and frequency of heavy precipitation in the future.

In 2009, the Norwegian Water Resources and Energy Directorate (NVE) started developing a national LEWS as part of a national program for landslide risk management; the system was officially launched in the autumn of 2013. The service is operative year-round performing daily regional scale landslide hazard assessments (i.e., for a county and/or group of municipalities). Predictions of the system are based on forecasting hydro-meteorological conditions leading to landslide initiation (Beldring et al. 2003). The model uses rainfall and temperature as input data and simulates many hydro-meteorological parameters, such as runoff, snowmelt, groundwater, soil saturation, and frost depth. In the development phase of the LEWS, hazard threshold levels were calibrated through a statistical analysis of historical landslides and different hydro-meteorological conditions. The thresholds currently adopted in the LEWS were proposed by Colletuille et al. (2010), by combining simulations of water supply (rainfall and snowmelt) and soil water content (both expressed as relative values normalized to annual averages and maxima over a 30-year reference period, respectively). Thresholds were statistically derived from empirical tree classification considering 206 landslides events from different parts of the country. Piciullo et al. (2017b) analyzed the warnings issued in Western Norway in the period 2013–2014, confirming an overall good performance of the adopted thresholds. Finally, it is worth mentioning that decision-making procedures are based upon hazard threshold levels that are supported by hydro-meteorological and real-time landslide observations as well as landslide inventories and susceptibility maps (Devoli et al. 2018).

Data

Data on landslide occurrences were retrieved from the national landslide database (www.skrednett.no), which contains more than 60,000 entries (represented by point locations) covering the whole country considering the last 500 years. A relevant number of registrations (16,346) are recorded in the period of analysis, i.e., from January 2013 to June 2017, among which 1481 can be considered weather-induced landslides in loose soils. Six hundred fifty-eight of these records (44.4%) are categorized as landslides in soil, not otherwise specified due to lack of further documentation; 654 (44.2%) are classified as debris flows, debris avalanches, or mudslides; 113 (7.6%) are reported as soil slides in artificial slopes (cuts and fillings along road and railway lines); 43 (2.9%) are slush flows, and the remaining 13 (0.9%) are clay slides. Registrations recorded by road and railway authorities are usually reported as points along transportation networks, thus often far away from the source area. In addition, further uncertainties may result from errors in the classification the type of landslide, lack of spatial and/or temporal information, and double registrations. As a consequence, each landslide event recorded in the national landslide database is affected by a certain degree of spatial uncertainty, varying from “exact” to “> 1000 m.” Uncertainty on landslide location always affects catalogs used for the calibration and validation of LEWS developed at regional scale (Segoni et al. 2018). Large uncertainties exist especially when assigning geographic

coordinates to a landslide event in case multiple landslides or affected areas were referenced in one report (Kirschbaum et al. 2010). However, the location of the large majority of the records used in this study is affected by a low spatial uncertainty (tens of meters); thus, the records can be considered adequate for analyses at catchment scale.

The pore water pressure measurements were collected at local scale-analyzing data from 41,706 boreholes installed by the Norwegian Geotechnical Institute (NGI) for a variety of geotechnical projects throughout Norway not specifically aimed at early warning purposes, such as geotechnical site characterization of soils; slope stability analysis; efficiency evaluation of surface drainage works; and monitoring of road and railway embankments. The piezometers considered representative of conditions which may lead to the triggering of landslides have been selected taking into account their spatial proximity to the landslides that occurred in the period of analysis, and the installation in the shallow soil layers or in areas characterized by the presence of loose sediments, according to a quaternary deposit map at 1:50,000 scale (www.ngu.no). It should be mentioned that only data from electric piezometers have been considered, as they provide longer and more reliable data series. Pore water pressure measurements used in this study were derived from a total of 240 boreholes installed between January 2013 and June 2017.

Test areas

The Norwegian hydrological basins have been considered as the most appropriate minimum territorial units for applying the augmented territorial warning model. Indeed, the catchment scale is an intermediate scale of analysis between small areas (e.g., single slopes) and very wide areas (e.g., a municipality, a region or a nation). Within these units, both the widespread meteorological monitoring data employed in the national LEWS and local pore water pressure observations may provide meaningful information for the definition of an efficient warning model. In addition, at catchment scale, the whole process area of the landslides is automatically considered. This allows reducing the uncertainty related to the location of the considered landslides. The test areas used in this study have been identified according to two main criteria: the occurrence of shallow landslides in loose soils and the availability of a relevant number of pore water pressure measurements.

According to the criteria of selection, 30 Norwegian hydrological basins have been identified as potentially useful for the analyses. Figure 5 shows that the majority of them (16 out of 30) are distributed along the western coast of Norway: three are situated in the southern part (SW in Fig. 5), twelve in the central part (CWa and CWb in Fig. 5), and one in the northern part (NW in Fig. 5). All of them are dominated by narrow fjords and steep mountain-sides and are characterized by the presence of shallow marine deposits covering weathered and altered bedrock. The remaining 14 basins are concentrated in the south-eastern part of Norway (SE in Fig. 5), an area highly prone to landslides due to long-term infiltration from large amounts of rain and/or snow in autumn and winter and presence of various shallow quaternary deposits (especially moraine materials). The main characteristics of the test areas are summarized in Table 2. The majority of the basins situated in the south-eastern part of Norway (10 out of 14) cover an area smaller than 50 km²; on the contrary, three out of the four basins with an area higher than 100 km² are distributed along the

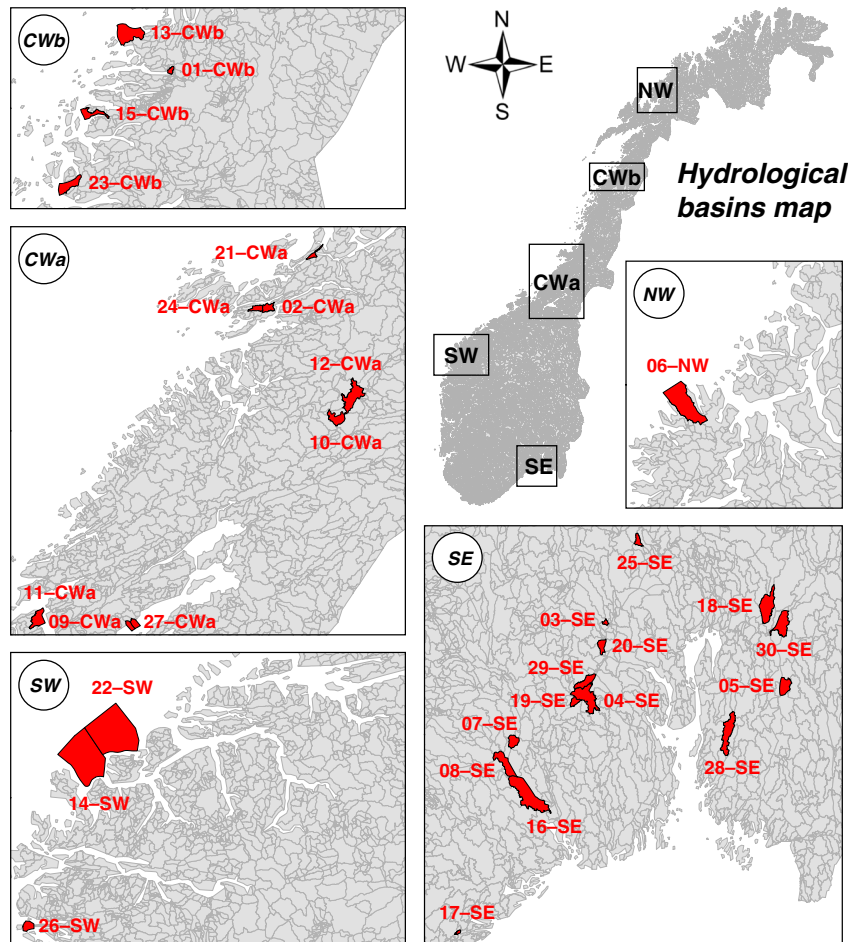


Fig. 5 Location of the 30 catchments used in this study as test areas

western coast. The shallow soil layers are significantly characterized by the presence of quaternary loose deposits highly prone to landslides, which cover more than half of the catchment surface in 11 cases out of 30. A total of 125 weather-induced landslides in soils occurred in the 30 test areas between January 2013 and June 2017: in the majority of them (25 out of 30), the number of landslides varies between 3 and 5, yet four basins were impacted by 6 landslides and the remaining one by 7 landslides. Pore water pressure measurements recorded in the period of analysis were derived from 240 boreholes, whose density within the test areas varies between 7 and 14.

Results

Application of the territorial warning model

The TeWM employed within the Norwegian LEWS has been applied to the 30 test areas identified in “Data” section. The model has been partially modified to consider the dimension of the test areas, the definition of the territorial warning events, and the meaning of the warning levels. Indeed, the Norwegian LEWS employs variable minimum territorial units, varying from a small group of municipalities (hundreds or thousands of km²) to several administrative regions (tens of thousands of km²). Thus, extent and position of warning zones are dynamic and may change day

by day, depending on hydrometeorological conditions. Conversely, the minimum territorial units adopted herein are the Norwegian hydrological basins, which are predefined static warning areas ranging from few km² to more than 100 km². In the Norwegian LEWS, relative water supply (rain and snowmelt) and relative soil saturation/groundwater conditions are combined for the definition of a hydro-meteorological index, which is compared with statistically defined thresholds. Then, a daily qualitative assessment of landslide warning levels is performed by an expert on duty supported by susceptibility maps and real-time observations. On the contrary, in this study, the hydrometeorological index is directly compared with the warning thresholds without any further qualitative assessment. In both cases, four warning levels are adopted: green (WL₁), yellow (WL₂), orange (WL₃), and red (WL₄), although the classification criteria are rather different. In the Norwegian LEWS, the principle behind the warning criteria is that rare hydrometeorological conditions are expected to cause more landslides and possibly higher damages. In addition, the criteria contain information on the expected number of landslides per area, as well as hazard signs indicating landslide activity. On the other hand, the classification criteria adopted in this study are not expected to be correlated to the number of expected landslides and the extension of the hazardous area in each territorial unit, but to provide an indication on the probability of landslide

Table 2 Name, area, presence of loose sediments, landslides, and piezometers available within the 30 test areas

ID	Name	Area (km ²)	Loose sediments (%)	Weather-induced landslides	Piezometers
01–CWb	Fykanåga/Glomfjorden, Gåsværfjorden og Sørfjorden	5.58	23	3	7
02–CWa	Horvereidelva	14.04	55	4	6
03–SE	Simoa/Drammensvassdraget	3.69	65	4	8
04–SE	Numedalslågen	64.00	38	3	10
05–SE	Glommavassdraget	27.13	59	5	8
06–NW	Senja Vest	136.66	17	6	8
07–SE	Skienassdraget	18.21	52	4	9
08–SE	Skienassdraget	38.88	53	5	8
09–CWa	Trondheimsfjorden Vest	8.08	63	3	6
10–CWa	Namsen	26.48	31	4	7
11–CWa	Ytre Trondheimsfjorden: Agdenes Fyr-geitaneset	30.02	22	4	7
12–CWa	Namsen	56.51	47	3	7
13–CWb	Fykanåga/Glomfjorden, Gåsværfjorden og Sørfjorden	65.47	29	3	6
14–SW	Hareidlandet og Gurskøya	269.15	7	3	8
15–CWb	Vestre Svartisen og Rødøy kommune	23.08	19	4	8
16–SE	Skienassdraget	104.55	51	6	12
17–SE	Vegårsvassdraget og Gjerstadvassdraget/kyst Kragerø-Tromøya	2.45	52	4	7
18–SE	Leira/Nitelva/Glommavassdraget	57.76	62	7	13
19–SE	Numedalslågen	10.09	39	3	8
20–SE	Drammensvassdraget	14.38	66	4	8
21–CWa	Vikna og Kyst Foldfjorden-Bindalsfjorden	7.64	72	3	7
22–SW	Hareidlandet og Gurskøya	307.18	10	4	6
23–CWb	kyst Utskarpen-nesna-tonnes	33.45	17	3	8
24–CWa	Horvereidelva	10.58	41	5	7
25–SE	Simoa/Drammensvassdraget	9.43	69	5	6
26–SW	Oselvassdraget/Flora kommune	15.25	27	3	6
27–CWa	Trondheimsfjorden Vest	9.83	78	3	6
28–SE	Mosseassdraget	51.39	36	6	10
29–SE	Vestfosselva/Drammensvassdraget	25.08	30	5	9
30–SE	Glommavassdraget	43.46	46	6	14

occurrence, i.e., null (WL₁), moderate (WL₂), high (WL₃), and very high (WL₄).

The assessment of the warning events resulting from the application of the territorial warning model within a territorial unit can be schematized into two phases. Firstly, the daily forecasts of water supply and relative soil water content are retrieved from the open-access web portal www.xgeo.no, where they are displayed as raster data at 1-km² resolution. Successively, hydro-meteorological indexes are evaluated calculating the average values of all the grid cells covering each territorial unit. Finally, the daily average values

are compared with the warning levels employed by the territorial model, to identify the days with warnings as well as to define, in these cases, the warning level. The results are reported in a 2 × 4 correlation matrix (in Table 3 displayed as a simple 1 × 8 row), which lists the landslide events (i.e., the presence or absence of a landslide) in relation to the warning events (WL₁, WL₂, WL₃, and WL₄). Twenty-six of the 125 landslide events in the test areas between January 2013 and June 2017 occurred when the warning model was in level 1, i.e., when no warnings were issued. For the other 99 occurrences, the warning model was in level 2 in 59 cases,

Table 3 Results obtained applying the TeWM in the 30 test areas from January 2013 to June 2017

Label	Landslides WL ₁	WL ₂	WL ₃	WL ₄	No landslides WL ₁	WL ₂	WL ₃	WL ₄
Total	26	59	35	5	41,622	492	104	0
Calibration dataset	13	37	30	2	27,191	327	61	0
Validation dataset	13	22	5	3	14,431	165	43	0

in level 3 in 35 cases and in level 4 in the remaining 5 cases. In the period of analysis, 695 daily warnings have been issued: the majority being level 2 “moderate warnings” (551 events), 139 events being level 3 “severe warnings”, with the rest of them being level 4 “very severe warnings” (5 events).

The 695 warnings issued by applying the territorial model were divided into two subsets: (i) a calibration set used to define the augmented territorial model, listing 457 warnings issued within 20 territorial units, and (ii) a validation set used to validate the model, listing 238 warnings issued within the remaining 10 territorial units. The two subsets have been defined with similar characteristics in terms of area of the territorial units, number of landslides, and availability of piezometers.

Augmented territorial warning model calibration and performance evaluation

The ATeWM presented in Fig. 3 has been calibrated using the 457 warnings issued by the territorial warning model. To this aim, a parametric analysis has been carried out for identifying the most appropriate time period (n) for calculating the simple moving average indicators. Then, a performance evaluation through statistical indicators has been developed for setting the values of the two thresholds (LT and UT) to be adopted in the second step of the procedure.

The possible trends of Δu_i have been evaluated over time periods (n) of 1, 2, 3, 4, 5, 6, 7, and 14 days. Only relatively short time periods have been considered in the analyses because pore water pressure variations leading to possible landslide initiation in shallow loose sediments are typically recorded few days before a landslide event. Besides, a short-term moving average allows identifying possible significant trends without any significant temporal lag between the original and the average data series caused by longer window lengths. Table 4 shows the comparison between the moving average trends considering the 69 warning events during which landslides occurred and the 388 warning events that are not associated with known landslides. It is worth highlighting that in the first case, an “uptrend” can be considered as a correct indicator (i.e., a trend suggesting that an increase of the warning level may be appropriate), whereas in the second case, a “downtrend” is a sign of coherence with the data (i.e., a trend suggesting that a decrease of the warning level or a withdraw of the warning may be appropriate). In both cases, “no trend” indicates that the moving average differences do not show a clear trend. Regarding “severe” (WL₃) and “moderate” (WL₂) warning events associated to landslides, a number of moving average differences (i.e., Δu_4 , Δu_5 , Δu_6 , Δu_7 , and Δu_{14} for WL₃; Δu_3 , Δu_4 , Δu_5 , Δu_6 , and Δu_{14} for WL₂) provide similar results, with a percent difference between correct and incorrect indicators greater than 50%. Conversely, considering both “severe” (WL₃) and “moderate” (WL₂) warning events not associated to landslides, the best performance is clearly

obtained using Δu_{14} , as this time period produces the maximum number of downtrends (41 for WL₃ and 206 for WL₂) as well as the minimum number of uptrends (17 for WL₃ and 80 for WL₂). “Very severe” warning events (WL₄) cannot be evaluated in this case because only two events were issued in the considered time frame, both of them associated to the occurrence of a landslide.

The trends of each moving average difference have been grouped according to the correctness of the indication provided, regardless of the warning level issued by the territorial model. On this issue, correct instances are represented by uptrends associated to warning events issued when landslides occurred, and downtrends associated with warning events issued when landslides did not occur. Conversely, incorrect instances are represented by uptrends associated with warning events issued when landslides did not occur and downtrends associated with warning events issued when landslides occurred. “No trend” instances refer, once again, to moving average differences that do not show a clear trend. Figure 6 highlights that the percent difference between correct and incorrect instances increases with the moving average time period. Indeed, Δu_{14} shows the highest value of this difference (42%), as 301 instances out of 457 are correct (66%), 110 are incorrect (24%), and the remaining 46 do not show any trend (10%). It should be pointed out that the performed pore water pressure trend analysis is conducted in addition to the evaluation of the territorial warning model, i.e., the Norwegian national LEWS, already defined combining hydrometeorological indexes likely to trigger landslides rather than using, as most of the territorial LEWS, only rainfall thresholds (Boje et al. 2014; Krogl et al. 2018). This suggests that the trigger of shallow landslides in Norway cannot be adequately predicted by short-term rainfall combinations alone. In this context, the information provided by the ATeWM (i.e., simple moving averages of recorded pore water pressures with a time period of 14 days) is precious and can be considered as representative of important hydrological preconditions increasing (or decreasing), at any given day, the average susceptibility to landslides within the warning zone.

Successively, a performance evaluation has been conducted using Δu_{14} to identify the two thresholds to be employed in the second step of the procedure. In particular, six combinations have been compared (Table 5), considering three values of the lower threshold, LT (i.e., 5%, 10%, and 15%) and two values of the upper threshold, UT (i.e., 25% and 30%). Table 6 shows the results obtained for the six combinations and the territorial warning model considering the elements of each correlation matrix in terms of alert classification and grade of accuracy. Firstly, the analysis reveals that the model is extremely sensitive to the variations of LT, resulting in large differences among the three pairs of combinations $\Delta u_{14,5,25}^* - \Delta u_{14,5,30}^*$, $\Delta u_{14,10,25}^* - \Delta u_{14,10,30}^*$, and $\Delta u_{14,15,25}^* - \Delta u_{14,15,30}^*$. This can be explained considering that the large majority of the level transitions in the period of analysis are

Table 4 Number of uptrends (up), downtrends (down), and no trends (nt) for SMA differences of various time periods considering warning events that resulted and did not result in landslides. The number of warning events issued for each warning level is reported in round brackets

Indicator	Landslides WL ₂ (37) up	down	nt	WL ₃ (30) up	down	nt	WL ₄ (2) up	down	nt
Δp_1	25	7	5	17	8	5	0	1	1
Δu_2	26	8	3	19	7	4	0	2	0
Δu_3	25	8	4	22	5	3	0	1	1
Δu_4	27	6	4	23	4	3	0	2	0
Δu_5	29	5	3	21	5	4	0	2	0
Δu_6	29	4	4	21	4	5	0	2	0
Δu_7	28	4	5	20	6	4	0	2	0
Δu_{14}	30	6	1	24	5	1	0	2	0
Indicator	No landslides WL ₂ (327) up	down	nt	WL ₃ (61) up	down	nt	WL ₄ (0) up	down	nt
Δp_1	138	120	69	25	25	11	0	0	0
Δu_2	133	126	68	22	29	10	0	0	0
Δu_3	127	130	70	27	28	6	0	0	0
Δu_4	130	130	67	27	28	6	0	0	0
Δu_5	118	141	68	26	30	5	0	0	0
Δu_6	117	147	63	24	31	6	0	0	0
Δu_7	100	171	56	22	29	10	0	0	0
Δu_{14}	80	206	41	17	41	3	0	0	0

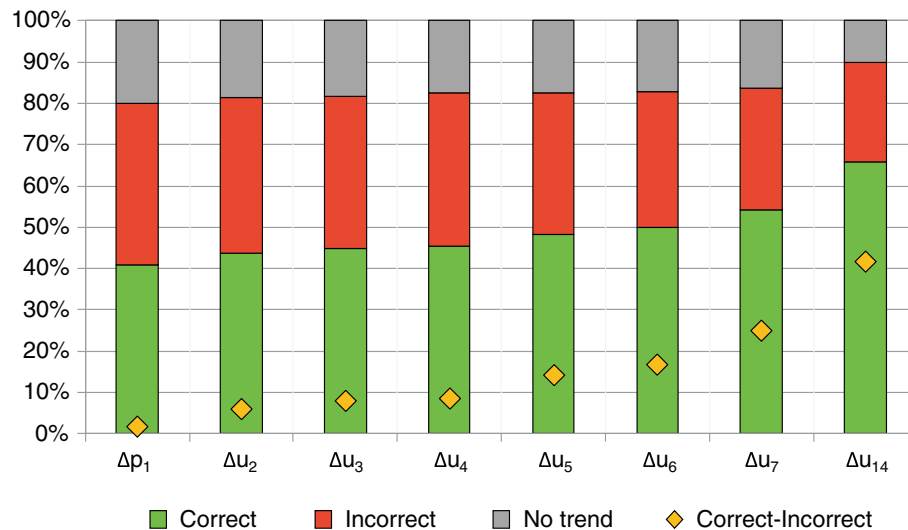


Fig. 6 Correctness of the indication provided by the SMA differences calculated using time periods (n) of 1, 2, 3, 4, 5, 6, 7, and 14 days

1-level transitions caused by the exceedance of LT. Higher values of TN and Gre and lower values of FA and Yel were obtained considering the pair of combinations $\Delta u_{14,5,25}^* - \Delta u_{14,5,30}^*$. This behavior is due to the high number of relocations from WL_2 to WL_1 in the second row of the matrix. However, a low value of the LT resulted in a high number of MA (19 for $\Delta u_{14,5,25}^*$ and 18 for $\Delta u_{14,5,30}^*$) and Red errors (85 and 84 respectively). On the other hand, a high value of the LT (15%) implied a lower number of positive level transitions from the territorial model when compared to the other two thresholds combinations (especially in terms of TN, FA, and Yel). For these reasons, the best-performing pair is $\Delta u_{14,10,25}^* - \Delta u_{14,10,30}^*$, which allows to significantly increase the number of TN and Gre, both minimizing the number of MA (15 for $\Delta u_{14,10,25}^*$ and 14 for $\Delta u_{14,10,30}^*$ against 13 of the territorial model) and significantly reducing the number of FA, Yel, and Red.

To confirm this finding, Table 7 and Fig. 7a, b show the results in terms of performance indicators for the six different threshold combinations and for the territorial warning model. Table 7 also reports the relative variations of the performance indicators from the territorial model for each threshold combination (\hat{x}), calculated as follows:

$$\hat{x} = \frac{x_i - x_0}{x_0} \quad (4)$$

where x_i is the performance indicator associated to a threshold combination, and x_0 is the same performance indicator associated to the territorial model.

Among the error indicators, the error rates (ER_a and ER_b) are very low (about 1% in both cases) for all the thresholds combinations as well as for the territorial warning model, due to the high numbers of TN and Gre. The overall smallest values are associated to the augmented territorial warning model with LT of 5% and 10% (more than 30% less than TeWM in both cases). The missed alert rate (R_{MA}) and the false alert rate (R_{FA}) are function of, respectively, days of MA and FA. Regarding the R_{MA} , it is worth mentioning that the augmented territorial warning model cannot decrease, for how it is defined, the number of MA recorded in the territorial model. The lowest increase of R_{MA} occurs for the threshold combinations employing LT of 10% and 15%, with values similar to that of the territorial model. The R_{FA} in the augmented territorial warning model is generally lower than in the territorial model and the lowest values are recorded employing LT of 5% and 10%. The lower threshold is affected by a non-negligible probability of having severe model errors, ($IS_{MA\&FA}$ is more than 30% in both the cases), whereas adopting 10% as LT of 10%, the values are quite similar to that of the territorial model (17.63% and 17.14%, respectively).

Regarding the success indicators, the Hanssen and Kuipers (1965) skill score (HK) shows a moderate but general decrease

Table 5 Combinations considered to identify the optimal values of LT and UT

Combination ID	Lower threshold (LT)	Upper threshold (UT)
$\Delta u_{14,5,25}^*$	5%	25%
$\Delta u_{14,5,30}^*$	5%	30%
$\Delta u_{14,10,25}^*$	10%	25%
$\Delta u_{14,10,30}^*$	10%	30%
$\Delta u_{14,15,25}^*$	15%	25%
$\Delta u_{14,15,30}^*$	15%	30%

Table 6 Number of correlation matrix elements considering the “alert classification” (a) and “grade of accuracy” (b) criteria

Element	Criterion	TeWM	$\Delta u_{14,5,25}^*$	$\Delta u_{14,5,30}^*$	$\Delta u_{14,10,25}^*$	$\Delta u_{14,10,30}^*$	$\Delta u_{14,15,25}^*$	$\Delta u_{14,15,30}^*$
CA	a	69	63	64	67	68	67	68
TN	a	27,191	27,349	27,347	27,326	27,324	27,279	27,277
MA	a	13	19	18	15	14	15	14
FA	a	388	230	232	253	255	300	302
Gre	b	27,223	27,405	27,403	27,383	27,381	27,327	27,325
Yel	b	364	171	174	229	232	269	272
Red	b	74	85	84	49	48	65	64

when the augmented territorial warning model is applied. This can be explained considering that this indicator is strongly influenced by the number of CA. However, the Matthews correlation coefficient (MCC) highlights an overall good performance of the augmented territorial warning model, especially for the pairs of combinations $\Delta u_{14,10,25}^* - \Delta u_{14,10,30}^*$ (positive relative variations higher than 16% respect to the territorial model). MCC can be considered a more balanced indicator for our analyses, as it is not influenced by the different orders of magnitude existing between the number of TN and the other elements of the correlation matrix.

Similarly, the Euclidean distance from the perfect ROC classification in the ROC analysis (δ_{ROC}) is lower in the territorial model (once again, as a consequence of the number of CA), whereas the Euclidean distance from the perfect classification in the Precision-Recall analysis (δ_{P-R}) is lower for the augmented territorial warning model regardless of the adopted thresholds. Based on this, $\Delta u_{14,10,25}^* - \Delta u_{14,10,30}^*$ are herein considered the best performing pair of combinations.

Augmented territorial warning model validation

The results of ATeWM have been tested with a validation dataset (i.e., 238 territorial warning events) in order to adopt the optimal combination of thresholds. Table 8 shows the results obtained for the six combinations and the TeWM considering the elements of each correlation matrix in terms of alert classification and grade of accuracy. The combination $\Delta u_{14,5,25}^* - \Delta u_{14,5,30}^*$ exhibits the best performance for 4 indicators out of 7 (i.e., TN, FA, Gre, and Yel). However, the remaining three indicators (i.e., CA, MA, and Red) produce values far lower than those of the territorial warning model. An increase of the LT to 15% resulted, in terms of days, in an increase of FA, Yel, and Red with respect to the two other pairs of threshold combinations. Therefore, once again, the best-performing pair of threshold combinations is $\Delta u_{14,10,25}^* - \Delta u_{14,10,30}^*$. Both combinations indeed produce a significant reduction of severe model errors (Red), and they do not increase the number of MA of the territorial model.

For consistency with the calibration analysis, the results are also reported in terms of performance indicators for the six

Table 7 Performance indicators computed for the TeWM and for six thresholds combinations of the ATeWM. Relative variations from the TeWM target value are reported in round brackets

Indicator	TeWM	$\Delta u_{14,5,25}^*$	$\Delta u_{14,5,30}^*$	$\Delta u_{14,10,25}^*$	$\Delta u_{14,10,30}^*$	$\Delta u_{14,15,25}^*$	$\Delta u_{14,15,30}^*$
ER _a	1.45%	0.90% (− 37.91%)	0.90% (− 37.66%)	0.97% (− 33.17%)	0.97% (− 32.92%)	1.14% (− 21.45%)	1.14% (− 21.20%)
ER _b	1.58%	0.93% (− 41.55%)	0.93% (− 41.10%)	1.01% (− 36.53%)	1.01% (− 36.07%)	1.21% (− 23.74%)	1.21% (− 23.29%)
R _{MA}	15.85%	23.17% (46.15%)	21.95% (38.46%)	18.29% (15.38%)	17.07% (7.69%)	18.29% (15.38%)	17.07% (7.69%)
R _{FA}	84.90%	78.50% (− 7.54%)	78.38% (− 7.68%)	79.06% (− 6.88%)	78.95% (− 7.01%)	81.74% (− 3.72%)	81.62% (− 3.86%)
IS _{MA&FA}	16.89%	33.20% (96.53%)	32.56% (92.71%)	17.63% (4.33%)	17.14% (1.47%)	19.46% (15.19%)	19.05% (12.74%)
HK	82.74%	76.00% (− 8.15%)	77.21% (− 6.69%)	80.79% (− 2.36%)	82.00% (− 0.89%)	80.62% (− 2.56%)	81.83% (− 1.10%)
MCC	35.29%	40.36% (14.36%)	40.80% (15.61%)	41.07% (16.40%)	41.50% (17.60%)	38.31% (8.55%)	38.73% (9.74%)
δ_{ROC}	0.1592	0.2319 (45.68%)	0.2197 (38.02%)	0.1832 (15.08%)	0.1710 (7.43%)	0.1832 (15.14%)	0.1711 (7.49%)
δ_{P-R}	0.8637	0.8185 (− 5.24%)	0.8139 (− 5.76%)	0.8115 (− 6.04%)	0.8077 (− 6.48%)	0.8377 (− 3.01%)	0.8339 (− 3.45%)

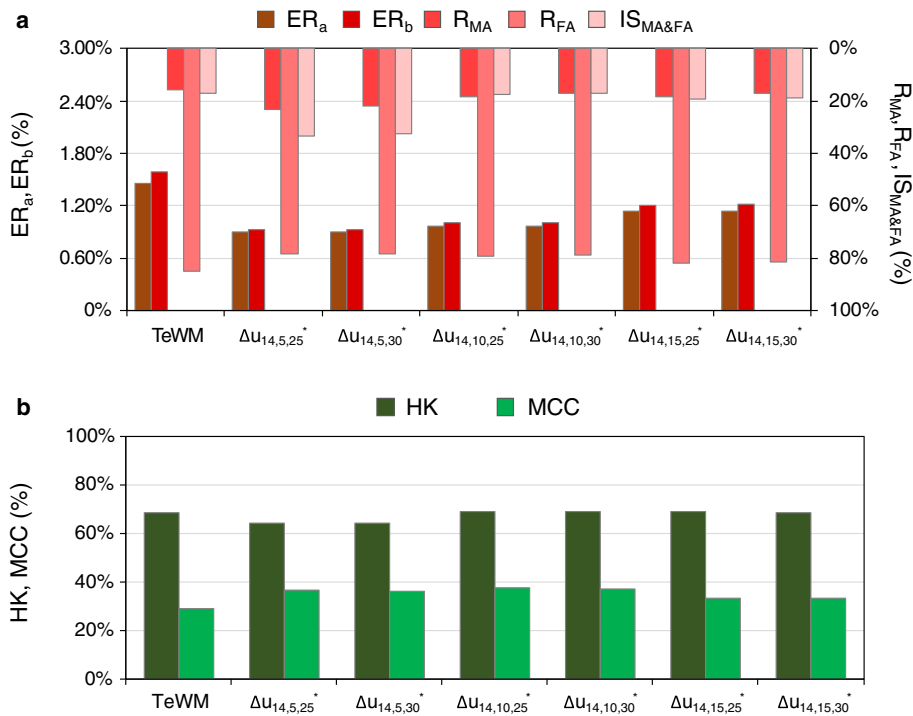


Fig. 7 Bar charts showing the values of error (a) and success (b) indicators for the TeWM and for each combination of thresholds adopted in the ATeWM

different threshold combinations and the territorial warning model (Table 9). Looking at the error rates (ER_a and ER_b), the application of the augmented territorial warning model results in a general decrease of the errors, regardless of the thresholds adopted (ER_a and ER_b become lower than 1% adopting a LT of 5% or 10%). Regarding the R_{MA}, it could be considered satisfactory that two pairs of combinations (i.e., $\Delta u_{14,10,25}^* - \Delta u_{14,10,30}^*$, and $\Delta u_{14,15,25}^* - \Delta u_{14,15,30}^*$) show the same value in the model. The R_{FA} and the IS_{MA&FA} follow a general trend similar to the calibration phase: the first indicator is lower than the territorial model for all six combinations, and the best values are recorded employing a LT of 5%, even though the differences are not substantial when LT is set to 10%. Moreover, $\Delta u_{14,10,25}^* - \Delta u_{14,10,30}^*$ allow minimizing the number of severe model errors, as the IS_{MA&FA} do not show relevant relative variations in respect to the territorial model (4.33% and 1.47%, respectively). Concerning the success indicators, a positive effect can be observed on the HK by adopting a LT of 10% and 15%, as the MA do not show any increase respect to the

territorial model and the FA decrease. The MCC follows a trend similar to the calibration phase, confirming that the pair $\Delta u_{14,10,25}^* - \Delta u_{14,10,30}^*$ represent a good compromise for reducing the FA without increasing the number of MA.

Figure 8 a reveals that δ_{ROC} exhibits almost the same values, apart from the pair $\Delta u_{14,5,25}^* - \Delta u_{14,5,30}^*$ because of the increase of the MA. On the other hand, the application of the augmented territorial warning model results in a reduction of the δ_{P-R} regardless of the adopted thresholds (Fig. 8b). Final comparisons are also shown for the type of errors within the FA (distinguishing between Red and Yel in Fig. 8c) and on the grade of correctness of the CA (distinguishing between Yel and Gre in Fig. 8d). Although a LT set to 5% allows minimizing of the number of FA, the severe errors remain almost the same, reducing marginally from 43 to 40. On the contrary, a significant reduction can be observed raising the LT to 10%, as the Red errors are reduced to 23. Moreover, looking at the CA, a LT of 10% leads to a good number of transitions from WL₂ to WL₃ and WL₄ (14), without missing any landslides. From

Table 8 Number of correlation matrix elements in terms of “alert classification” (a) and “grade of accuracy” (b) criteria

Element	Criterion	TeWM	$\Delta u_{14,5,25}^*$	$\Delta u_{14,5,30}^*$	$\Delta u_{14,10,25}^*$	$\Delta u_{14,10,30}^*$	$\Delta u_{14,15,25}^*$	$\Delta u_{14,15,30}^*$
CA	a	30	28	28	30	30	30	30
TN	a	14,431	14,532	14,530	14,522	14,520	14,485	14,483
MA	a	13	15	15	13	13	13	13
FA	a	208	107	109	117	119	154	156
Gre	b	14,439	14,557	14,555	14,544	14,542	14,496	14,494
Yel	b	187	70	72	102	104	137	139
Red	b	56	55	55	36	36	49	49

Table 9 Performance indicators computed for the TeWM and for six thresholds combinations of the ATeWM. Relative variations from the TeWM target value are reported in round brackets

Indicator	TeWM	$\Delta u_{14,5,25}^*$	$\Delta u_{14,5,30}^*$	$\Delta u_{14,10,25}^*$	$\Delta u_{14,10,30}^*$	$\Delta u_{14,15,25}^*$	$\Delta u_{14,15,30}^*$
ER _a	1.51%	0.83% (− 44.80%)	0.84% (− 43.89%)	0.89% (− 41.18%)	0.90% (− 40.27%)	1.14% (− 24.43%)	1.15% (− 23.53%)
ER _b	1.66%	0.85% (− 48.56%)	0.87% (− 47.74%)	0.94% (− 43.21%)	0.95% (− 42.39%)	1.27% (− 23.46%)	1.28% (− 22.63%)
R _{MA}	30.23%	34.88% (15.38%)	34.88% (15.38%)	30.23% (0.00%)	30.23% (0.00%)	30.23% (0.00%)	30.23% (0.00%)
R _{FA}	87.39%	79.26% (− 9.31%)	79.56% (− 8.96%)	79.59% (− 8.93%)	79.87% (− 8.62%)	83.70% (− 4.23%)	83.87% (− 4.03%)
IS _{MA&FA}	23.05%	44.00% (96.53%)	43.31% (92.71%)	26.09% (4.33%)	25.71% (1.47%)	26.34% (15.19%)	26.06% (12.74%)
HK	68.35%	64.39% (− 5.80%)	64.37% (− 5.82%)	68.97% (0.91%)	68.95% (0.89%)	68.72% (0.54%)	68.70% (0.52%)
MCC	29.25%	36.45% (24.64%)	36.18% (23.71%)	37.43% (28.00%)	37.18% (27.12%)	33.38% (14.13%)	33.20% (13.50%)
$\bar{\sigma}_{ROC}$	0.3027	0.3489 (15.28%)	0.3489 (15.28%)	0.3024 (− 0.08%)	0.3024 (− 0.07%)	0.3025 (− 0.05%)	0.3025 (− 0.05%)
$\bar{\sigma}_{P-R}$	0.9248	0.8660 (− 6.36%)	0.8687 (− 6.06%)	0.8514 (− 7.93%)	0.8540 (− 7.66%)	0.8899 (− 3.77%)	0.8915 (− 3.59%)

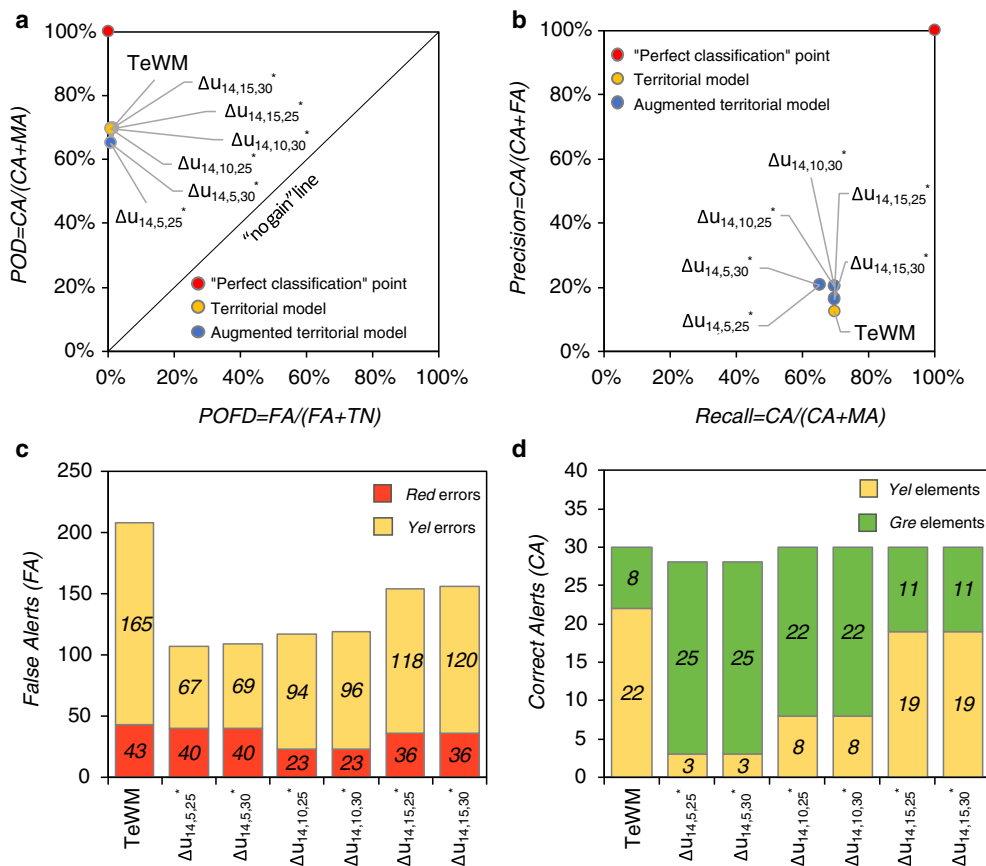


Fig. 8 Performance analysis of the TeWM and the six thresholds combinations of the ATeWM considering: (a) ROC classification, (b) Precision-Recall classification, (c) severity of errors of FA, and (d) grade of correctness of CA

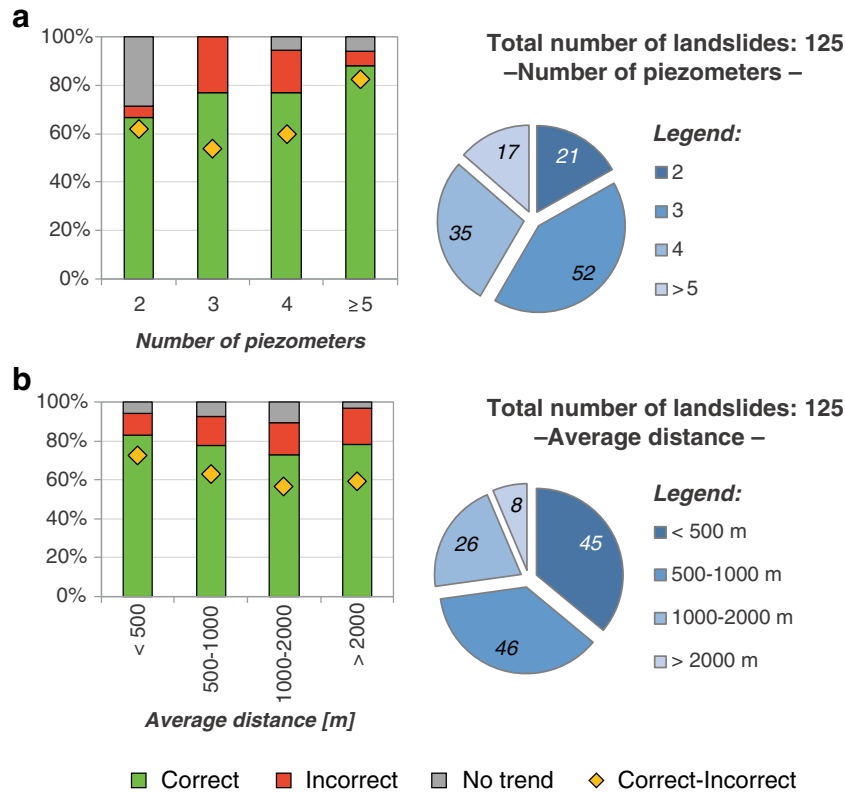


Fig. 9 Influence on the correctness of the results of the number of piezometers (a) and of the average distance between landslides and networks of piezometers (b). The total numbers of landslides belonging to each numerosity and average distance class are also reported

this, it is possible to conclude that the performed validation analyses confirm that the pair $\Delta u_{14,10,25}^* - \Delta u_{14,10,30}^*$ are the best performing threshold combinations.

Influence of numerosity and location of piezometric data

The influence of the piezometers network configurations (i.e., numerosity of the piezometers and average distance from the landslides) on the reliability of the predictions has been evaluated in order to assess their correlation with the accuracy of the results produced by the ATeWM.

A first parametric analysis has been conducted to assess the influence of the numerosity of the piezometric locations on the accuracy of the results. The number of piezometers available for each landslide that occurred in the period of analysis varies between 2 and 7. They have been grouped in four numerosity classes (2, 3, 4, and 5 or more piezometers) so that a reasonable number of data is available in each class for supporting the analyses. The moving average trends have been evaluated in order to identify the uptrends (correct indicators) associated with the landslides for each class, following the same procedure adopted for the parametric analysis carried out in “Application of the territorial warning model” section aimed at identifying the most appropriate time period (n) for calculating the simple moving average indicators. Figure 9a highlights that most of the landslides are related to uptrends, regardless to the numerosity of piezometric data. Yet, the correctness of the predictions increases passing from 2 (66%) to 5 or more piezometers (88%). Moreover, the upper class is characterized by the highest percent difference between correct

and incorrect indicators (82%), as only in 2 cases out of 17 there is not coherence between landslides and piezometric data. This seems to suggest a positive correlation between the reliability of the predictions and the numerosity of piezometric data.

Another parametric analysis has been carried out to assess whether and how the reliability of the results depends on the average distance between the networks of piezometers and the landslides. Indeed, the significance of piezometric data collected at a certain distance from the landslide is a crucial issue to investigate before using them for early warning purposes. To this end, the networks of piezometers have been grouped in four rather homogenous classes depending on their average distance from the landslides. It should be pointed out that “< 500 m” has been considered as the lower class in order to minimize the effects of the spatial uncertainties affecting the records contained in the landslide database (see “Data” section for further details on this issue). Following the same criterion already adopted for the two other parametric analyses, the moving average trends have been evaluated in order to assess the number of correct indicators (i.e., uptrends) for each class (Fig. 9b). As expected, the networks of piezometers installed in the proximity of the landslides (i.e., at a distance shorter than 500 m) provide the best results (83%), and only in 3 cases the occurrence of a landslide is not associated with an uptrend. However, the percent differences between correct and incorrect indicators observed for the four defined classes do not show significant variations, especially for average distances longer than 500 m. Indeed, the performed analyses revealed that the remaining piezometer classes provide correct indications in the

majority of the cases: 78% (“500–1000 m” class), 73% (“1000–2000 m” class), and 78% (“> 2000 m” class).

These results suggest that the piezometric data available in each territorial unit can be usefully adopted in the ATeWM, regardless of their distance from the landslides. These aspects are crucial for evaluating the applicability of this method to different territorial units.

Conclusions

Landslide early warning systems are used to assess the probability of occurrence of weather-induced landslides over large areas through the prediction and monitoring of meteorological variables. Weather measurements alone, however, are often insufficient to produce reliable predictions of landslide occurrences in complex geomorphological environments. In such circumstances, locally collected geotechnical data can provide additional information to allow warning levels to be dynamically adjusted before they are issued by systems managers. This study presented an innovative augmented territorial warning model aimed at integrating pore water pressure measurements in a territorial warning model for weather-induced landslides in Norway.

In the case study, the ATeWM has been tested and applied considering, as territorial units, the hydrological basins for which an ensemble of shallow local pore water pressure measures may be considered a proxy of significant weather-induced changes in the groundwater regime over significant parts of the basins. The model has been applied considering several types of landslides that are already under surveillance in the Norwegian national LEWS and for which pore water pressure measurements could provide indications. These landslides are generally triggered by a rise of pore water pressure due to rain infiltration or snow melting and usually occur in quaternary shallow soil layers overlying bedrock. The proposed methodology was successfully applied to 30 Norwegian hydrological basins in the period between January 2013 and June 2017, exploiting a landslide catalog containing 125 weather-induced landslides. Moreover, two parametric analyses carried out in order to evaluate the influence of the different network configurations of the piezometers (i.e., numerosity and average distance from the landslides) on the correctness of the predictions, clearly demonstrated the reliability of the pore water pressure trend analysis at catchment scale.

In the model, local pore water pressures have been processed and combined in an ensemble of measures within a procedure aimed at improving the performance of a territorial warning model that is already operational at national scale. Modeling the soil-atmosphere interaction for soil stability purposes at slope scale was outside the scope of this study. Therefore, these measures should be considered proxies of significant weather-induced changes in the groundwater regime at the scale of the territorial units used as warning zones. The specific results of this study are only applicable to the territorial warning model used for Norway, as different (possibly shorter) time periods may be adopted to optimize the performance of a model in other locations. Despite this, the proposed methodology could be applied to other case studies, upon considering adequate contextual differences and limitations. In general, the field of application of an augmented territorial warning model is

limited by the field of application of the original territorial warning model. Of course, the predictions of the ATeWM are also related to landslide phenomena for which the geotechnical measures collected at local scale, and considered in the model, can provide indications. In addition, the size of the territorial units used as warning zones needs to be appropriately defined. Indeed, a number of local measurements need to be available in each adopted territorial unit, as they are potentially useful for analyses at territorial scale only if the monitoring data are processed and taken together as an ensemble of measures.

In summary, the major advantages of the proposed methodology are the following:

- The simplicity and speed of the forecasting allow an easy implementation into operational early warning systems.
- The number of extra input data needed is limited, and consists of available local geotechnical monitoring data.
- The calibration and validation procedures, based on standard statistical indicators, allow the periodic performance assessment of the ATeWM.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10346-020-01599-w>.

Acknowledgments

The authors are grateful to the Norwegian Geotechnical Institute for providing pore water pressure data during a research period of Gaetano Pecoraro in WP3.4-Early warning systems in the context of the project Klima 2050-Risk reduction through climate adaptation of buildings and infrastructure.

Funding

Open access funding provided by Università degli Studi di Salerno within the CRUI-CARE Agreement. This work was financially supported by the Project of Relevant National Interest (PRIN) 2015 on “Protecting the Cultural Heritage From Water Soil Interaction” (code: 300393PRN2015CASI) and by the Inter University Consortium of Risk and Hazard management (CUGRI).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Barla M, Antolini F (2016) An integrated methodology for landslides' early warning systems. *Landslides* 13:215–228. <https://doi.org/10.1007/s10346-015-0563-8>
- Baum RL, Godt JW (2010) Early warning of rainfall-induced shallow landslides and debris flows in the USA. *Landslides* 7(3):259–272. <https://doi.org/10.1007/s10346-009-0177-0>
- Beldring S, Engeland K, Roald LA, Sælthun NR, Voksø A (2003) Estimation of parameters in a distributed precipitation runoff model for Norway. *Hydrol Earth Syst Sci* 7:304–316. <https://doi.org/10.5194/hess-7-304-2003>
- Boje S, Colletuille H, Cepeda J, Devoli G (2014) Landslide thresholds at regional scale for the early warning system in Norway. *Proceeding 3rd World Landslide Forum*, 2–6 June 2014, Beijing, 2014
- Caine N (1980) The rainfall intensity-duration control of shallow landslides and debris flows. *Geografiska Annal* 62A:23–27
- Calvello M (2017) Early warning strategies to cope with landslide risk. *Riv It Geotecnica* 2:63–91. <https://doi.org/10.19199/2017.2.0557-1405.063>
- Calvello M, Pecoraro G (2018) Franelitalia: a catalog of recent Italian landslides. *Geoenviron Disasters* 5(13). <https://doi.org/10.1186/s40677-018-0105-5>
- Calvello M, Piciullo L (2016) Assessing the performance of regional landslide early warning models: the EDuMaP method. *Nat Hazards Earth Syst Sci* 16:103–122. <https://doi.org/10.5194/nhess-16-103-2016>
- Canli E, Mergili M, Thiebes B, Glade T (2018) Probabilistic landslide ensemble prediction systems: lessons to be learned from hydrology. *Nat Hazards Earth Syst Sci* 18:2183–2202. <https://doi.org/10.5194/nhess-18-2183-2018>
- Colletuille H, Haugen LE, Beldring S (2010) A forecast analysis tool for extreme hydrological conditions in Norway. Poster presented in: *Proc of 6th world FRIEND 2010*, Flow regime and international experiment and network data, Fez, Morocco
- Devoli G, Tiranti D, Cremonini R, Sund M, Boje S (2018) Comparison of landslide forecasting services in Piemonte (Italy) and in Norway, illustrated by events in late spring 2013. *Nat Hazards Earth Syst Sci* 18:1351–1372. <https://doi.org/10.5194/nhess18-1351-2018>
- Dietrich A, Krautblatter M (2017) Evidence for enhanced debris-flow activity in the Northern Calcareous Alps since the 1980s (Plansee, Austria). *Geomorphology* 287:144–158. <https://doi.org/10.1016/j.geomorph.2016.01.013>
- Fan L, Lehmann P, McArdell B, Or D (2017) Linking rainfall-induced landslides with debris flows runoff patterns towards catchment scale hazard assessment. *Geomorphology* 280:1–15. <https://doi.org/10.1016/j.geomorph.2016.10.007>
- Fredin O, Bergström B, Eilertsen R, Hansen L, Longva O, Nesje A, Sveian H (2013) Glacial landforms and Quaternary landscape development in Norway. *Quaternary Geology of Norway* In: Olsen L, Fredin O, Olesen O (eds.), Geological Survey of Norway, Special Publication, Trondheim, pp 5–25
- Froude MJ, Petley DN (2018) Global fatal landslide occurrence from 2004 to 2016. *Nat Hazards Earth Syst Sci* 18:2161–2181. <https://doi.org/10.5194/nhess-18-2161-2018>
- Furseth A (2006) Slide accidents in Norway. *Tun Forlag*, Oslo, pp. 207 (In Norwegian)
- Gariano SL, Brunetti MT, Iovine G, Melillo M, Peruccacci S, Terranova O, Vennari C, Guzzetti F (2015) Calibration and validation of rainfall thresholds for shallow landslide forecasting in Sicily, southern Italy. *Geomorphology* 228:653–665. <https://doi.org/10.1016/j.geomorph.2014.10.019>
- Hanssen AW, Kuipers WJA (1965) On the relationship between the frequency of rain and various meteorological parameters. *Koninklijk Nederlands Meteorologisch Instituut Meded Verhand* 81:2–15
- Hanssen-Bauer I, Førland EJ, Haddeland I, Hisdal H, Mayer S, Nesje A, Nilsen JEØ, Sandven S, Sandø AB, Sorteberg A, Ådlandsvik B (2017) Climate in Norway 2100, a knowledge base for climate adaptation. *NCCS report 1* <http://www.miljodirektoratet.no/Documents/publikasjoner/M741/M741.pdf>
- Haque U, Blum P, da Silva PF, Andersen P, Pilz J, Chalov SR, Malet JP, Auflič MJ, Andres N, Poyiadji E, Lamas PC, Zhang W, Peshevski I, Pétursson HG, Kurt T, Dobrev N, García-Davalillo JC, Halkia M, Ferri S, Gaprindashvili G, Engström J, Keellings D (2016) Fatal landslides in Europe. *Landslides* 13(6):1545–1554. <https://doi.org/10.1007/s10346-016-0689-3>
- Hungr O, Leroueil S, Picarelli L (2014) The Varnes classification of landslide types, an update. *Landslides* 11:167–194. <https://doi.org/10.1007/s10346-013-0436-y>
- Kirschbaum DB, Adler R, Hong Y, Hill S, Lerner-Lam AL (2010) A global landslide catalog for hazard applications: method, results and limitations. *Nat Hazards* 52:561–575. <https://doi.org/10.1007/s11069-009-9401-4>
- Kirschbaum DB, Stanley T, Zhou Y (2015) Spatial and temporal analysis of a global landslide catalog. *Geomorphology* 249:4–15. <https://doi.org/10.1016/j.geomorph.2015.03.016>
- Krogli IK, Devoli G, Colletuille H, Boje S, Sund M, Engen IK (2018) The Norwegian forecasting and warning service for rainfall- and snowmelt-induced landslides. *Nat Hazards Earth Syst Sci* 18:1427–1450. <https://doi.org/10.5194/nhess-18-1427-2018>
- Lin Q, Wang Y (2018) Spatial and temporal analysis of a fatal landslide inventory in China from 1950 to 2016. *Landslides* 15:2357–2372. <https://doi.org/10.1007/s10346-018-1037-6>
- Martelloni G, Segoni S, Fanti R, Catani F (2012) Rainfall thresholds for the forecasting of landslide occurrence at regional scale. *Landslides* 9:485–495. <https://doi.org/10.1007/s10346-011-0308-2>
- Mirus BB, Morphet MD, Smith JB (2018) Developing hydro-meteorological thresholds for shallow landslide initiation and early warning. *Water* 10(9):1274. <https://doi.org/10.3390/w10091274>
- Pecoraro G, Calvello M, Piciullo L (2019a) Monitoring strategies for local landslide early warning systems. *Landslides* 16:213–231. <https://doi.org/10.1007/s10346-018-1068-z>
- Pecoraro G, Calvello M, Cepeda JM (2019b) Using local monitoring data for regional forecasting of weather-induced landslides in Norway. *Proceedings of the XVII ECSMG-2019*, Reykjavik, Iceland, 1–6 September 2019. <https://doi.org/10.32075/17ECSMG-2019-0141>
- Petley D (2012) Global patterns of loss of life from landslides. *Geology* 40(10):927–930. <https://doi.org/10.1130/G33217.1>
- Piciullo L, Calvello M, Cepeda JM (2018) Territorial early warning systems for rainfall-induced landslides. *Earth Sci Rev* 179:228–247. <https://doi.org/10.1016/j.earscirev.2018.02.013>
- Piciullo L, Dahl M-P, Devoli G, Colletuille H, Calvello M (2017b) Adapting the EDuMaP method to test the performance of the Norwegian early warning system for weather-induced landslides. *Nat Hazards Earth Syst Sci* 17:817–831. <https://doi.org/10.5194/nhess-17-817-2017>
- Piciullo L, Gariano SL, Melillo M, Brunetti MT, Peruccacci S, Guzzetti F, Calvello M (2017a) Definition and performance of a threshold-based regional early warning model for rainfall-induced landslides. *Landslides* 14:995–1008. <https://doi.org/10.1007/s10346-016-0750-2>
- Segoni S, Piciullo L, Gariano SL (2018) A review of the recent literature on rainfall thresholds for landslide occurrence. *Landslides* 15:1483–1501. <https://doi.org/10.1007/s10346-018-0966-4>
- Sorbino G, Nicotera MV (2013) Unsaturated soil mechanics in rainfall-induced flow landslides. *Eng Geol* 165:105–132. <https://doi.org/10.1016/j.enggeo.2012.10.008>
- Springman SM, Thielem A, Kienzler P, Friedel S (2013) A long-term field study for the investigation of rainfall-induced landslides. *Geotechnique* 63(14):1177–1193. <https://doi.org/10.1680/geot.11.P.142>
- Stähli M, Sättele M, Huggel C, McArdell BW, Lehmann P, Van Herwijnen A, Berne A, Schleiss M, Ferrari A, Kos A, Or D, Springman SM (2015) Monitoring and prediction in early warning systems for rapid mass movements. *Nat Hazards Earth Syst Sci* 15:905–917. <https://doi.org/10.5194/nhess-15-905-2015>
- Staley DM, Kean JW, Cannon SH, Schmidt KM, Laber JL (2013) Objective definition of rainfall intensity-duration thresholds for the initiation of post-fire debris flows in southern California. *Landslides* 10:547–562. <https://doi.org/10.1007/s10346-012-0341-9>
- UNISDR (2006) Available at: <http://www.unisdr.org/2006/ppew/info-resources/ewc3/Global-Survey-of-Early-Warning-Systems.pdf>
- Zhang S, Zhao L, Delgado-Tellez R, Bao H (2018) A physics-based probabilistic forecasting model for rainfall-induced shallow landslides at regional scale. *Nat Hazards Earth Syst Sci* 18:969–982. <https://doi.org/10.5194/nhess-18-969-2018>

G. Pecoraro · M. Calvello 

Department of Civil Engineering,
University of Salerno,
84084, Fisciano, Italy
Email: mcalvello@unisa.it

G. Pecoraro
e-mail: gpecoraro@unisa.it