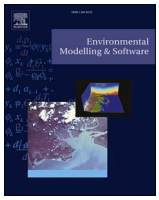
Environmental Modelling and Software 170 (2023) 105833

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: www.elsevier.com/locate/envsoft



Speech-recognition in landslide predictive modelling: A case for a next generation early warning system

Zhice Fang [a], Hakan Tanyas [b], Tolga Gorum [c], Ashok Dahal [b], Yi Wang [a], Luigi Lombardo [b,*]

[a] China University of Geosciences, Institute of Geophysics and Geomatics, Wuhan, China
[b] University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC), Enschede, Netherlands
[c] Eurasia Institute of Earth Sciences, Istanbul Technical University, Istanbul, Turkey

ARTICLE INFO

Keywords:
Landslide prediction
Precipitation
Speech recognition
Time series
Early warning system

ABSTRACT

Traditional landslide early warnings are based on the notion that intensity-duration relations can be approximated to single precipitation values cumulated over fixed time windows. Here, we take on a similar task being inspired by modeling architectures typical of speech-recognition tasks. We aim at classifying the Turkish landscape into 5 km grids assigned with dynamic landslide susceptibility estimates. We collected all available national information on precipitation-induced landslide occurrences. This information is passed to a Long Short-Term Memory equipped with the whole rainfall time series, obtained from daily CHIRPS data. We test this model: 1) by randomizing the presence/absence data to represent the slope instability over Turkey and over 13 years under consideration (2008–2020) and 2) by assessing the effect of different time windows used to pass the rainfall signal to the neural network. Results show that the inclusion of the full precipitation signal rather than its scalar approximation leads to a substantial increase in prediction power (approximately 20%). This may potentially pave the road for a new generation of speech-recognition-based landslide early warning systems.

1. Introduction

Landslide early warning systems (LEWSs, Guzzetti et al., 2020) focus on forecasting slope failures as a function of weather forecast data (Gariano and Guzzetti, 2016; Segoni et al., 2018). Regional and global scale LEWSs mainly follow a probabilistic framework (Pecoraro et al., 2019; Stanley et al., 2020) and numerically estimate occurrences of landslides by discriminating locations that previously experienced landsliding from those that did not fail (Bragagnolo et al., 2020; Guo et al., 2022). Overall, LEWSs issue landslide alerts solving a classification task, and specifically by estimating the dependence between landslide presence/absence data and precipitation proxies. This is different from traditional susceptibility studies, which are solely based on static terrain and geological characteristics (Reichenbach et al., 2018; Titti et al., 2022). The occurrence probability coming from LEWSs is therefore mosty interpreted as the temporally-dynamic element of the hazard definition (e.g., Ahmed et al., 2020; Steger et al., 2023).

At the core of any LEWSs is the notion of precipitation thresholds (Harilal et al., 2019; Rosi et al., 2021). These are essentially precipitation values above which an alert is issued to local communities to minimize the risk of being exposed to landsliding (Intrieri et al., 2012; Osanai et al., 2010). In the context of data-driven models, precipitation thresholds are mainly estimated by taking a variable selection approach, where a classification task is solved using the precipitation measured at the date of the landslide occurrence (Melillo et al., 2018). The classification performance is then stored, and the second step tests the same situation but takes the cumulative precipitation measured on the day of the landslide occurrence plus the day before it (Peruccacci et al., 2017). The performance is stored once again and the procedure continues backward, taking each time the cumulative precipitation value for a fixed window and checking whether its scalar use would lead to a performance increase (e.g., Afungang and Bateira, 2016; Wang et al., 2021). The routine usually stops when the maximum accuracy is reached. This allows one to intuitively visualize the intensity-duration relationship (Mathew et al., 2014), as a means to understand how much precipitation cumulated over time is most likely to trigger landslides for specific landscapes (Zhao et al., 2022).

Landslide types and weather characteristics can largely influence the length of the cumulative time window. For instance, LEWSs for shallow landslides usually define the precipitation threshold over a shorter amount of time (e.g., one week reported by Nikolopoulos et al., 2014) as compared to deep-seated failures (e.g., sixty days reported by Martelloni

* Corresponding author.
E-mail address: l.lombardo@utwente.nl (L. Lombardo).

https://doi.org/10.1016/j.envsoft.2023.105833
Received 11 August 2023; Received in revised form 18 September 2023; Accepted 20 September 2023
Available online 22 September 2023

et al., 2012). This is mainly because of the lag time between precipitation and failure. In other words, water infiltration could take longer for deep-seated landslides as compared to shallow ones and thus, requires a longer lag time to produce slope instabilities. However, this does not mean that the influence of long-term precipitation could be ignored while predicting shallow-seated landslides or vice versa for deep-seated ones. Both short- and long-term precipitation records are useful to better understand conditions leading to the genesis of landslides. For instance, prolonged dry periods followed by heavy rainfall could affect the shrinking-swelling response of clay-rich hillslopes and cause instabilities (Schulz et al., 2018; Tichavský et al., 2019). This implies that aggregating the precipitation discharged over time in LEWSs may not be the ideal way to take into account this complex dynamic process. And yet, the geoscientific community has mostly worked with single scalar rainfall values so far.

Using a single aggregated measure of precipitation over time neglects the potential information a data-driven model can gain from using the entire precipitation time series. This is actually the main motivation of this manuscript. We question whether we could pass the continuous precipitation signal carrying rich spatiotemporal information to a binary classifier to perform the very same task that traditional LEWSs solve.

This paper aims at proposing a novel approach for the prediction of precipitation-triggered landslides as a baseline that could lead to a new LEWS protocol. Our framework explore a third dimension, time and to do so, we exploit the entire precipitation time series as an explanatory variable of our predictive model.

We use speech recognition architectures based on a deep learning algorithm known as Long Short-Term Memory networks (LSTM, Zhou et al., 2016). These networks are capable of solving a classification task by using time series or functional data instead of a scalar information. Their invention has been originally proposed in the context of text reading (Cheng et al., 2016) or speech recognition (Sak et al., 2014). In the latter applications, they can discriminate two or multiple speakers as a function of the time series of vocal cord vibrations, recognizing unique voices, tones, and other characteristics. Similarly, we envision the use of speech recognition to discriminate landslide presences/absences in space and time, as a function of full precipitation time series. To test this assumption, we go through a large number of sources reporting landslides including technical reports, newspapers, and scientific articles, and gather a precipitation-trigger landslide catalog of Turkey covering the period between 2008 and 2020. We use this dataset to test our speech recognition idea in the context of landslide prediction.

2. Data and methods

2.1. Landslide database

We generate a landslide database covering the period between 2008 and 2020. To do so, we scan through numerous sources including national, local printed and digital media reports (e.g., https://egazete.cumhuriyet.com.tr/yayinlar) as well as research papers, incidents reports from government and emergency agencies (e.g., https://www.afad.gov.tr/afet-analiz and https://katalog.devletarsivleri.gov.tr). We search for several keywords in Turkish (e.g., landslide, slide, mass movement, slump) to identify landslide occurrences and their occurrence dates. For each reported landslide event, we also check the available optical images through Google Earth and/or PlanetScope (3–5 m), Rapid Eye (5 m) images acquired from Planet Labs (Planet Team, 2017). This is done to confirm the spatial and temporal accuracy of the information for each event before adding it to the landslide database we generate for this research. The final output of the procedure is a point-based landslide database, whose geolocation is associated with the day of occurrence.

For the specific task of this experiment, we couple landslide observations with Climate Hazards Group InfraRed Precipitation with Station Data (CHIRPS, spatial resolution is 0.05°; Funk et al., 2015). The spatial resolution of the CHIRPS dataset constitutes the main criterion we use to decide whether the spatial accuracy of a given reported landslide event is suitable for our research. This means that if a reported landslide event is identified with ~5 km accuracy, we include them in our database because it would still match the CHIRPS resolution. Any larger uncertainty in the landslide geolocation leads to a rejection instead. The same spatial structure, therefore, dictated the choice of our mapping unit. In other words, we partitioned the Turkish landscape in a regular lattice of approximately 5 km side.

2.2. Long Short-Term Memory networks

LSTM networks belong to a special class of recurrent neural networks that can address vanishing gradient problems (Hochreiter, 1998) and have been successfully employed in sequential data modeling (Lipton et al., 2016). Fig. 1 summarizes the structure of our LSTM model.

LSTM uses three key gate functions to control the information flow process (Hochreiter and Schmidhuber, 1997), including the input gate f_t, forget gate i_t, and output gate o_t.

Let $x = \{x_1, x_2, ..., x_N\}$ be a precipitation time series input. W_x, W_h, and b denote the weight of input, the weight of the hidden state, and the bias, respectively. The forget gate f_t determines whether the previous information is to be remembered or can be forgotten in the current time step t. It can be denoted as flows:

$$f_t = \sigma\left(W_{fx}x_t + W_{fh}h_{t-1} + b_f\right) \tag{1}$$

where σ is the sigmoid activation function, and h_{t-1} is the previous hidden memory state. The input gate i_t determines the input information updating and $\widetilde{c}_t$ memorizes the new information, which is denoted as follows:

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \tag{2}$$

$$\widetilde{c}_t = tan\ h(W_{cx}x_t + W_{ch}h_{t-1} + b_c) \tag{3}$$

where $tan\ h$ is the activation function. The new memory cell c_t the state is then updated as follows:

$$c_t = f_t \odot c_{t-1} + i_t \odot \widetilde{c}_t \tag{4}$$

where c_{t-1} is the previous memory cell state and $\odot$ is the Hadamard product. Finally, the output gate o_t controls the output activation, and the hidden memory state sent to the next time step is defined as follows:

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \tag{5}$$

$$h_t = o_t \odot tan\ h(c_t) \tag{6}$$

In the final time step, we apply a softmax activation function on the hidden state h_N, and regard the maximum score in the output as the final prediction.

2.3. Model implementation

We use the mapped landslide inventory to build the binary classification model. This model boils down to a space-time occurrence probability estimator whose association between landslide presence/absence data and precipitation makes use of the full precipitation time series. In other words, if traditional landslide early warnings are based on a scalar measure of precipitation (the sum over a fixed time window), our model does not summarize the time series in a single value but rather makes use of the raw and full temporal signal instead. The functional representation of the precipitation signal is extracted from CHIRPS data.

To do so, we first extract daily precipitation values from the day of landslide occurrence to the 60th day before the event. We then obtain precipitation time series (thus made of 60 sequential time-points) for all landslide locations in the inventory and replicate the same procedure for

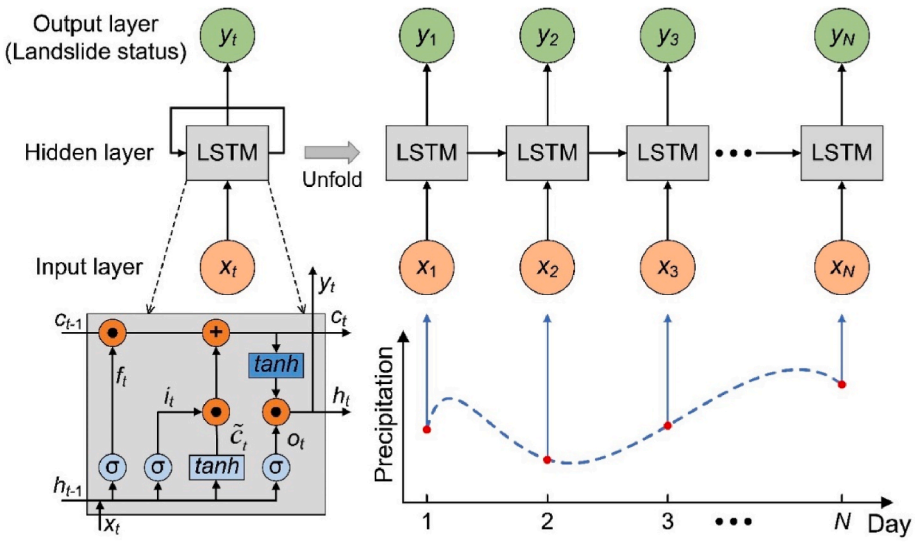


Fig. 1. Schematic diagram summarizing the structure of our LSTM model.

landslide-absence locations.

To create a reliable presence dataset, we remove any landslide point for which the rainfall was zero. Taking away presence data with no precipitation on the day the slope failure ensures removing any potential noise due to either inaccurate landslide reports or precipitation data. As for the absences, the situation requires a number of additional steps. For instance, the number of potential absences to sample from the whole Turkish territory and repeated on a daily basis for 13 years, is extremely large. We thus start calculating the mean and standard deviation of the landslide-presence precipitation time series and set the mean plus 2 standard deviations as the minimum threshold to filter in landslide-absence samples. We only retain landslide-absence samples as those that have at least one extreme precipitation day (i.e., at least one day above the $\mu+2\sigma$ obtained from the presences) across the 60-day time series. This operation is included to remove any trivial information. In other words, if we would not include this criterion, the most common situation in an arid territory such as the Turkish one, would potentially extract time series where no rainfall occurred for large time windows. Such information is not what we seek to model, as it should be obvious that no precipitation-induced landslide can take place without precipitation in the first place. To avoid simply learning the difference between dry and wet conditions, extracting rainfall time series with at least one extreme would challenge the model to discriminate the influence of different precipitation regimes.

The network structure and hyperparameters are important for deep learning-based modeling. In this study, the model architecture has a LSTM hidden layer, further containing 16 hidden units. The model follows an output layer with two neurons, each one equipped with a softmax function. During the optimization, we use the categorical cross-entropy as the loss function and ADAM is the optimizer of our choice (e.g., Kingma, 2014; Li et al., 2020). We set the batch size and epoch as 64 and 30, respectively (Brownlee, 2018). We implemented the LSTM model using Python language under the Keras framework (Gulli and Pal, 2017).

To assess the model prediction performance, we use three different performance estimators: (1) the receiver operating characteristic curve (AUC, Fawcett, 2006), (2) F1-score (Singhal, 2001), and (3) Cohen's Kappa coefficient (Kraemer, 2014). Ultimately, we apply a 10-fold cross-validation procedure to validate our model based on the balanced dataset.

2.4. Model benchmark

We also compare our approach against an alternative structure typical of traditional Early Warning Systems (Segoni et al., 2018). These are universally based on an intensity-duration relationship, reflected in the model by cumulating the rainfall amount over a given time window and thus obtaining a single scalar value. In doing so, we allow our version of LEWS to be validated against a consolidated benchmark, comparing the effect of using the whole time series and a single representative value of the same. The performance metrics (AUC, F1-Score, and Cohen's Kappa coefficient) mentioned above are also used in this case.

3. Results

The landslide inventory we put together featured 2380 individual precipitation-triggered landslides that occurred between 2008 and 2020 in Turkey (Fig. 2). Looking at their co-location within the 5 km lattice defined by the CHIRPS data, and after filtering out dry occurrence dates, the resulting presence database was reduced to 680 landslide-presence grids. The absences were then generated by selecting a balanced and random sample from the remaining 5 km grids that fulfilled the criterion of hosting at least an extremely rainy day. We would like to stress here that such data construction was also iteratively repeated to obtain a representative sample of the suitable absence instances across the whole space-time domain under consideration.

We then merged the presences with the random subsets of absence instances to build the LSTM networks. Fig. 3 shows a graphical example of the precipitation time series associated with randomly selected landslide and no-landslide locations in 2020. This is something we show to highlight how complex the differentiation between the two classes is if done through visual comparison. In fact, no evident difference emerges between the two classes. It is for this reason that a neural network architecture is particularly appealing as it is capable of differentiating between the two sets of presence/absence time series (Karim et al., 2018).

We tested the sensitivity of our model using AUC, F1-score, and Cohen's Kappa coefficient, these being computed from two separate experiments. The first one revolves around iteratively combining 10 random batches of 680 absences from a large database, together with the 680 presences. Each one of these ten newly formed datasets will then undergo the same LSTM procedure, whose performance is assessedvia a random10-fold cross-validation. This operation ensures testing for the

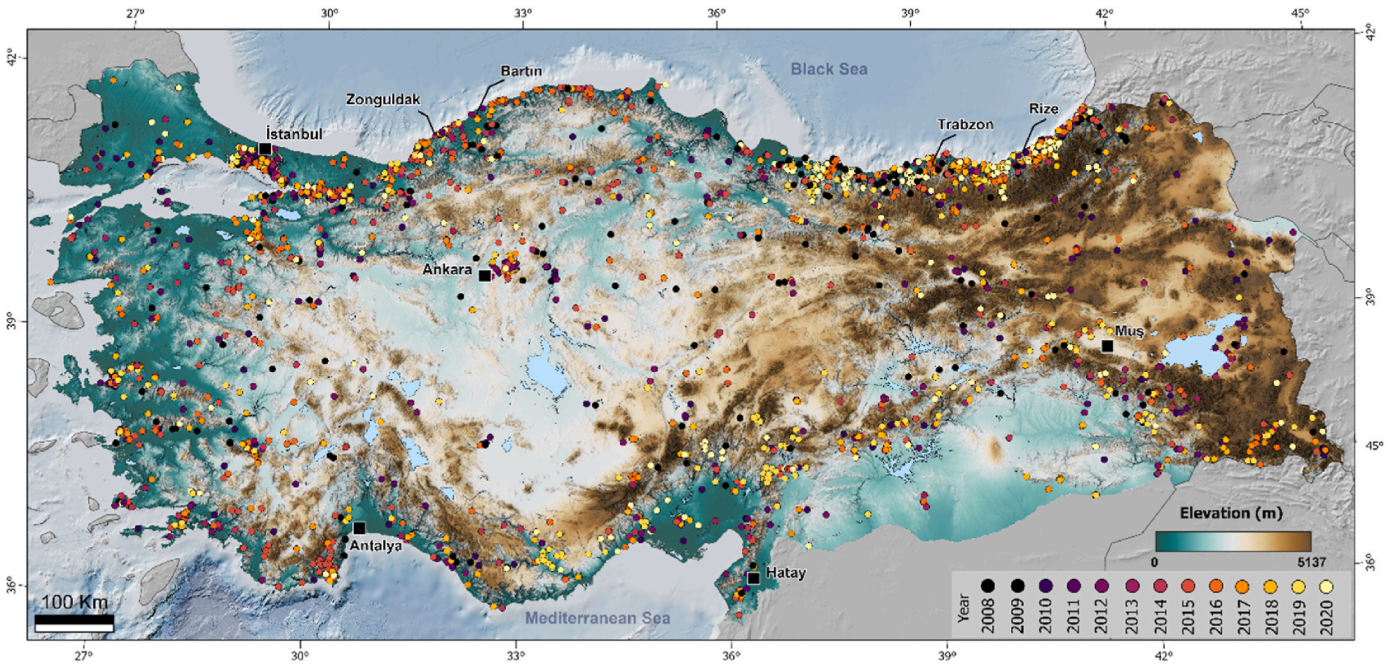


Fig. 2. Spatial distribution of landslides occurred in Turkey between 2008 and 2020.

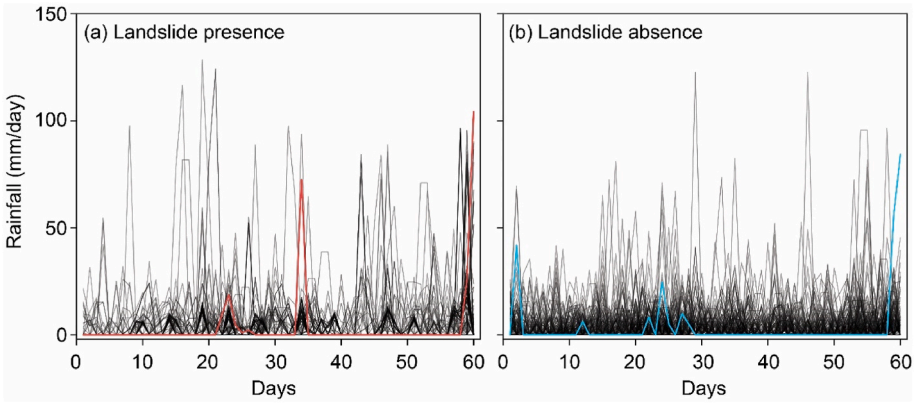


Fig. 3. Randomly selected precipitation time series for landslide-presence and landslide-absence conditions in 2020. The red and blue lines are just two examples we picked to highlight the similarity between the two datasets. Even the absence time series host extreme rainfall discharges comparable to those shown in the presence counterpart.

model robustness varying the representative precipitation regimes over the Turkish landscape and also over time. The second experiment repeats the same operation but shrinks the time window that defined the rainfall time series, from an initial length of 60 days–10 days for a total of 6 nested tests (60, 50, 40, 30, 20, 10 days).

Results in Fig. 4 show that the AUC, F1-score, and Cohen's s coefficient respectively range between 0.86 and 0.96, 0.83–0.92, and 0.63–0.80, considering all experiments described above. The AUC is a performance diagnostic commonly employed for classification purposes. For this reason, its values have been classified to summarize the results into acceptable ($0.7 < AUC < 0.8$), excellent ($0.8 < AUC < 0.9$), and outstanding ($0.9 < AUC < 1$), with 0.5 and 1.0 respectively defining the purely random and perfect classification cases (see, Hosmer and Lemeshow, 2000). For this reason, an overall mean AUC of 0.920 across all tests indicates an outstanding performance of our proposed approach that relied on the full precipitation time series. As for the best and worst models (see Fig. 5a), what stands out is that using the 60-day precipitation time series produces an AUC of 0.934, whereas achieving the lowest AUC of 0.906 is associated with the 20-day precipitation time series. The same panel shows all three performance metrics together, where an incremental performance is achieved essentially across all of them at increasing lengths of the considered time windows.

We consider this a reasonable result because some additional information may be contained over larger periods. But, it should also be stressed that the power of using the whole time series is already visible simply within 10 days, where the performances are particularly high nonetheless.

For any new experimental design and associated numerical solution to be valid, one has to benchmark the results against a scientific standard. In our case, this was done by keeping the LSTM architecture but exploring the variation in performance obtained by using cumulated precipitation, as per Guzzetti et al. (2020), instead of the whole time series. The results are shown in Fig. 5b, c, and 5d where a drastic drop stands out across the whole spectrum of performance diagnostics. Looking at the AUC values estimated for different time windows, the results obtained with a scalar representation of the precipitation signal appear very far from those estimated with our full signal approach. In fact, the maximum AUC (approx. 0.75) in the scalar case is reached at a

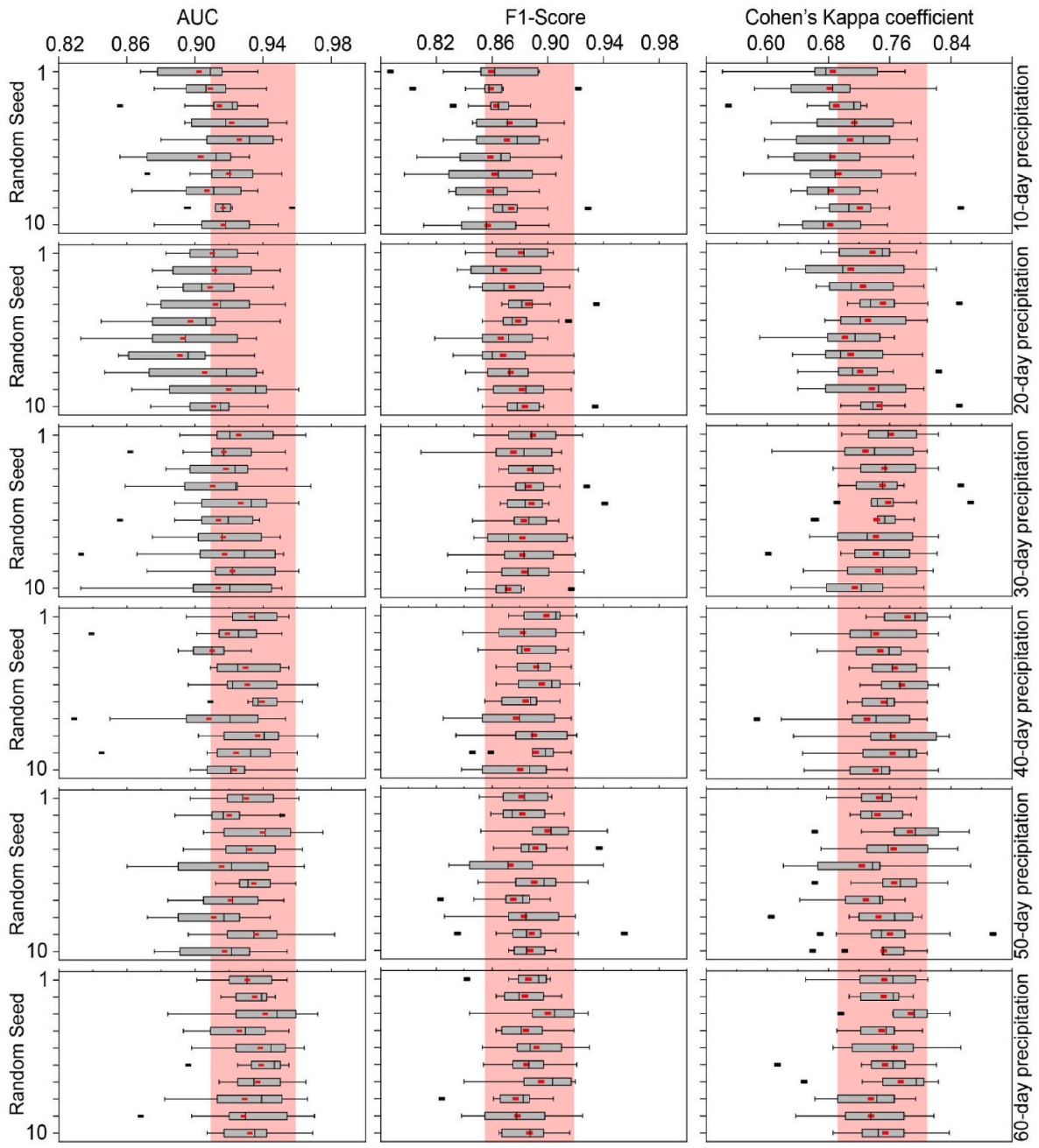


Fig. 4. Model prediction performance for different sampling schemes. Each boxplot shows the range of AUC based on 10-fold cross-validation in a given length of precipitation time series and a given random seed of absence samples selection. The red points in the boxplot refer to the mean value, whereas the red box indicates the limits of lower and upper quartiles calculated using 60-day precipitation.

cumulated rainfall of 50 days. This almost corresponds to a 20% drop in performance with respect to the maximum AUC reached for a 60-day time window, when using the entire precipitation series. The very same pattern is retrieved for the remaining two performance indicators.

Having demonstrated the gain in predictive power, we opted to include a hindcast example, converting the results to map form. To do so, we focused on the North-Eastern sector of Turkey, where rainfall is frequently responsible for landslide occurrences, even fatal ones (Görüm and Fidan, 2021). There, we simulated the daily dynamism in the susceptibility patterns for June 2020, following our approach and specifically for the model defined over the 60-day time series (because it proved to be numerically the best). The occurrence of a reference landslide in the area is marked on the 13th of June in our inventory. This is also the day in our simulation when the maximum occurrence probability is reached for the single 5 km × 5 km grid where the failure took place. Another interesting element to be examined in Fig. 6 corresponds to the right panel at the bottom. There, for the unstable grid mentioned above, we plot the sequence of daily susceptibility values, together with the rainfall time series. Despite the model takes into account 60 days of precipitation, it is still sensitive to impulsive rainfall discharges, with trends in the probability that follow potential spikes in precipitation discharges. This is quite typical of the LSTM architecture as it progressively "forgets" signals from the past and assigns larger weights as close as possible to the date of interest. This characteristic is quite suitable in the landslide context because as a community, we usually discriminate between preparatory and triggering rainfall (Mondini et al., 2023; Steger et al., 2023).

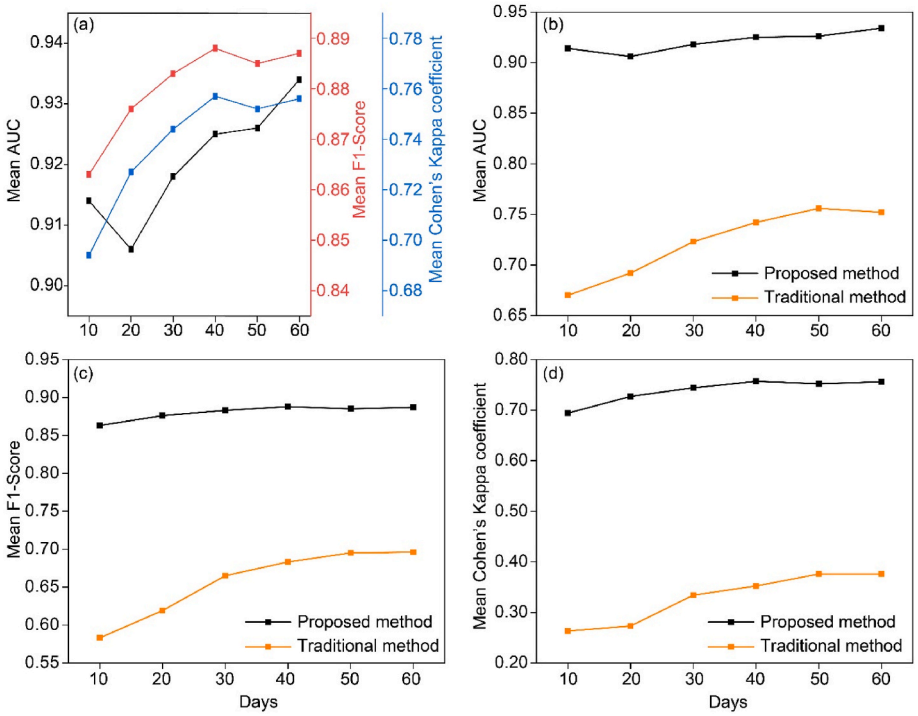


Fig. 5. Panel (a) reports the three performance metrics computed for the LSTM architecture fed with the continuous precipitation signal. Panels (b), (c), and (d) show the same information (black lines) but overlaid to a 2-D space where the same metrics are obtained for an LSTM fed with a scalar representation of the precipitation (orange lines).

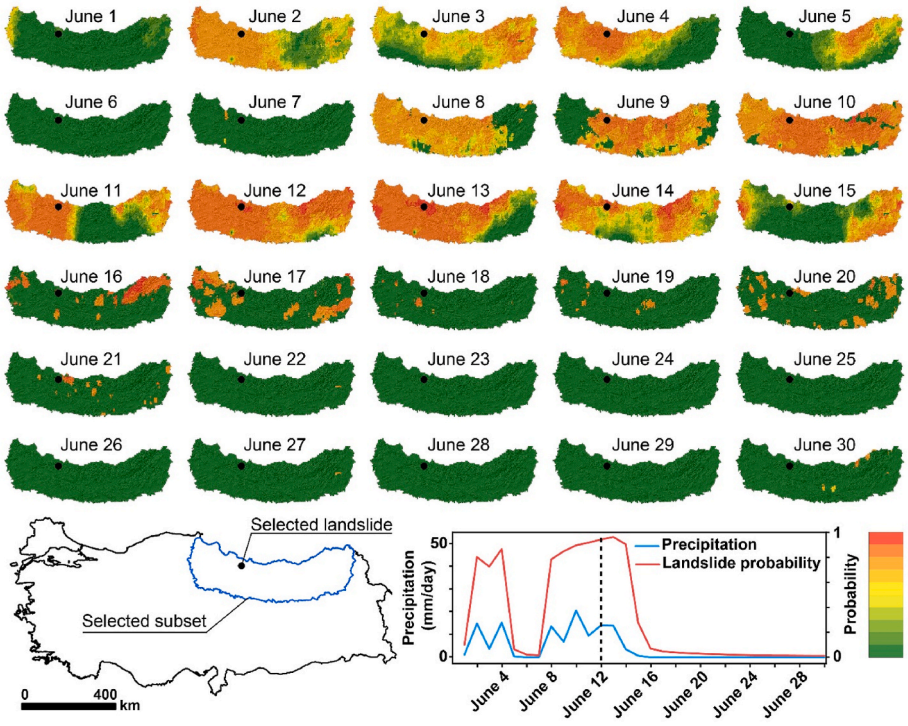


Fig. 6. A hindcast example demonstrating predicted landslide probably varying over time with respect to precipitation time series.

4. Discussion

The proposed approach returned outstanding performances across the whole spectrum of tests and model diagnostics, whether we included random tests performed by changing the absence instances, or for all tests run at different time windows. Overall, this is a positive indication that the idea of using speech recognition in landslide predictive studies may constitute the foundation for next-generation early warning systems in the future. However, this method still requires additional tests and below we will share our vision of what still needs to be done before it may become an operational tool.

The most intuitive element to be added corresponds to the use of

terrain characteristics in the model. The current model we propose stems from the traditional setting where rainfall intensity and duration are responsible for the landslide occurrence probability, without featuring landscape characteristics typical of landslide susceptibility studies. However, time, technological advances and data availability have reached a level of maturity where space-time models could finally feature all this information at once. An example already exists where terrain and rainfall characteristics (albeit in a scalar form) are simultaneously regressed against dynamic landslide occurrence data (Stanley et al., 2021; Steger et al., 2023).

Another important aspect to be improved could involve moving away from a coarse mapping unit. The 5 km mapping unit we chose could be further downscaled to provide valuable information for regional or even catchment-scale applications. However, to do so, satellite-based precipitation may not be sufficient as they are mostly coarse and affected by a number of potential biases and uncertainties. One possible solution may involve the use of weather stations equipped with radar sensors. This would certainly ensure much finer spatial and temporal resolutions. Notably, radar data constitutes the backbone of many state-of-the-art weather forecast systems. In its current state, our speech-recognition approach is exclusively relying on CHIRPS and specifically on past precipitation estimates. Therefore, it can only be used for hindcasting purposes. This is something that still needs to be validated and compared against other analogous products. A suitable candidate in this sense may be the MSWEP V2 (Beck et al., 2019) as it both brings high spatial and temporal details, with a level of accuracy proven to be better than most satellite-based alternative products. However, aside from the hindcasting aspects, we envision future efforts to expand beyond the limitations of looking backward in time and rather test it for nowcast and forecast applications. This would require re-training the LSTM architecture on equally nowcasted and forecasted precipitation measures. Ground-based radar could certainly offer a potential solution, but this is also valid for other space-born satellite missions. For instance, the model-based Goddard Earth Observing System-Forward Processing (GEOS-FP) precipitation (Molod et al., 2012; Suarez et al., 2008) offers a forecast which has already been exploited as a tool to develop an early warning system for precipitation-induced landslides (Khan et al., 2022). The use of such precipitation forecast instead of CHIRPS may open up a LEWS tool. This of course brings us back to the spatial resolution issue, albeit this would come with an increase in the temporal details. The topic of temporal resolution is something particularly interesting, especially because of the speech-recognition architecture we propose. In fact, the current model scans through the precipitation signal expressed daily. Therefore, it could be scientifically interesting to see whether adding sub-daily information (3-hourly, hourly, or even sub-hourly) would lead to different results and explore why. Going back to the topic related to the uncertainty in the precipitation products, better accuracy could also lead to variation in the model output. For instance, in the current settings, we had to remove part of the landslide presence information because some instances reported zero or negligible precipitation on the day of the landslide occurrence. Relying on better precipitation products, even from interpolated discharges measured over a rain-gauge network could potentially retain useful slope instability information. Ultimately, another potential improvement we envision relies on the use of a classified landslide inventory. Currently, we used a national-level landslide information and as such, it did not report which landslide type each data point belonged to. However, at least theoretically, one should expect that different landslide types should behave differently form a pure hydrological perspective. Therefore, by testing the current modeling architecture on specific landslide types should lead to even better results as compared to those obtained in this work.

5. Conclusions

This research offers a different take on the typical modeling structure for predicting precipitation-induced landslides. Specifically, as an alternative method to the traditional accumulated precipitation proxies, here we explore the use of a Neural Network architecture typical of speech recognition. This offers the ability to pass to the model a functional representation of the precipitation signal rather than a scalar approximation of it. To support this complex experimental design, we developed a completely new landslide inventory exclusively reporting precipitation-induced landslides, that occurred between 2008 and 2022 across Turkey.

The performance we retrieve appears to be significantly better than the traditional counterpart (i.e., 20% higher prediction power). Therefore, our results indicate that the full-time series may carry more useful information when classifying locations likely to undergo landsliding or not. Despite the achieved results, many other tests are required before a speech-recognition could actually make its way into landslide early warning systems. Among these tests, other study areas are certainly required, but the most important element to be addressed lies in the use of forecasted rather than hindcasted precipitation. If such data would also confirm its success over the traditional scalar version of the same, a new generation of alert system may be developed, based on this innovative speech-recognition idea.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Luigi Lombardo reports financial support was provided by National Natural Science Foundation of China.

Data availability

https://doi.org/10.5281/zenodo.8103276

Acknowledgement

This work was supported by the Joint Funds of the National Natural Science Foundation of China (U21A2013) and the Fundamental Research Funds for National Universities, China University of Geosciences (Wuhan). Initial idea proposed by Ashok Dahal. Experimental design developed by Zhice Fang, Luigi Lombardo and Hakan Tanyas. The same authors have prepared scientific illustrations and written the manuscript. Zhice Fang has also worked on the implementation. Tolga Gorum provided the data. Yi Wang secured the above mentioned funds.

References

Afungang, R.N., Bateira, C.V., 2016. Temporal probability analysis of landslides triggered by intense rainfall in the Bamenda Mountain Region, Cameroon. Environ. Earth Sci. 75, 1032. https://doi.org/10.1007/s12665-016-5835-7.

Ahmed, B., Rahman, M.S., Sammonds, P., Islam, R., Uddin, K., 2020. Application of geospatial technologies in developing a dynamic landslide early warning system in a humanitarian context: the Rohingya refugee crisis in Cox's Bazar, Bangladesh. Geomatics, Nat. Hazards Risk 11, 446–468. https://doi.org/10.1080/19475705.2020.1730988.

Beck, H.E., Wood, E.F., Pan, M., Fisher, C.K., Miralles, D.G., Van Dijk, A.I., McVicar, T.R., Adler, R.F., 2019. MSWEP V2 global 3-hourly 0.1 precipitation: methodology and quantitative assessment. Bull. Am. Meteorol. Soc. 100 (3), 473–500.

Bragagnolo, L., da Silva, R.V., Grzybowski, J.M.V., 2020. Landslide Susceptibility Mapping with R. Landslide: A Free Open-Source GIS-Integrated Tool Based on Artificial Neural Networks, vol. 123. Environmental Modelling & Software, 104565.

Brownlee, J., 2018. What is the difference between a batch and an epoch in a neural network. Mach. Learn. Mastery 20.

Cheng, J., Dong, L., Lapata, M., 2016. Long Short-Term Memory-Networks for Machine Reading, 06733 arXiv Prepr. arXiv1601.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recogn. Lett. 27, 861–874. https://doi.org/10.1016/j.patrec.2005.10.010.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci. Data 2, 150066. https://doi.org/10.1038/sdata.2015.66.

Gariano, S.L., Guzzetti, F., 2016. Landslides in a changing climate. Earth Sci. Rev. 162, 227–252. https://doi.org/10.1016/j.earscirev.2016.08.011.

Görüm, T., Fidan, S., 2021. Spatiotemporal variations of fatal landslides in Turkey. Landslides. https://doi.org/10.1007/s10346-020-01580-7.
Gulli, A., Pal, S., 2017. Deep Learning with Keras. Packt Publishing Ltd.
Guo, Z., Torra, O., Hürlimann, M., Abancó, C., Medina, V., 2022. FSLAM: a QGIS plugin for fast regional susceptibility assessment of rainfall-induced landslides. Environ. Model. Software 150, 105354.
Guzzetti, F., Gariano, S.L., Peruccacci, S., Brunetti, M.T., Marchesini, I., Rossi, M., Melillo, M., 2020. Geographical landslide early warning systems. Earth Sci. Rev. 200 https://doi.org/10.1016/j.earscirev.2019.102973.
Harilal, G.T., Madhu, D., Ramesh, M.V., Pullarkatt, D., 2019. Towards establishing rainfall thresholds for a real-time landslide early warning system in Sikkim, India. Landslides 16, 2395–2408. https://doi.org/10.1007/s10346-019-01244-1.
Hochreiter, S., 1998. The vanishing gradient problem during learning recurrent neural nets and problem solutions. Int. J. Uncertain. Fuzziness Knowledge-Based Syst. 107–116. https://doi.org/10.1142/S0218488598000094, 06.
Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9, 1735–1780.
Hosmer, D.W., Lemeshow, S., 2000. Applied Logistic Regression. John Wiley & Sons, New York.
Intrieri, E., Gigli, G., Mugnai, F., Fanti, R., Casagli, N., 2012. Design and implementation of a landslide early warning system. Eng. Geol. 147–148, 124–136. https://doi.org/10.1016/j.enggeo.2012.07.017.
Karim, F., Majumdar, S., Darabi, H., Chen, S., 2018. LSTM fully convolutional networks for time series classification. IEEE Access 6, 1662–1669. https://doi.org/10.1109/ACCESS.2017.2779939.
Khan, S., Kirschbaum, D.B., Stanley, T.A., Amatya, P.M., Emberson, R.A., 2022. Global landslide forecasting system for hazard assessment and situational awareness. Front. Earth Sci.
Kingma, D.P., 2014. A Method for Stochastic Optimization. ArXiv Prepr.
Kraemer, H.C., 2014. Kappa Coefficient. Wiley StatsRef Stat. Ref online 1–4.
Li, S., Chen, H., Wang, M., Heidari, A.A., Mirjalili, S., 2020. Slime mould algorithm: a new method for stochastic optimization. Future Generat. Comput. Syst. 111, 300–323. https://doi.org/10.1016/j.future.2020.03.055.
Lipton, Z.C., Kale, D.C., Elkan, C., Wetzel, R.C., 2016. Learning to Diagnose with LSTM Recurrent Neural Networks. BT - 4th International Conference on Learning Representations. Conference Track Proceedings. ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016.
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.
Mathew, J., Babu, D.G., Kundu, S., Kumar, K.V., Pant, C.C., 2014. Integrating intensity–duration-based rainfall threshold and antecedent rainfall-based probability estimate towards generating early warning for rainfall-induced landslides in parts of the Garhwal Himalaya, India. Landslides 11, 575–588. https://doi.org/10.1007/s10346-013-0408-2.
Melillo, M., Brunetti, M.T., Peruccacci, S., Gariano, S.L., Roccati, A., Guzzetti, F., 2018. A Tool for the Automatic Calculation of Rainfall Thresholds for Landslide Occurrence, vol. 105. Environmental Modelling & Software, pp. 230–243.
Molod, A., Takacs, L., Suarez, M., Bacmeister, J., Song, I.-S., Eichmann, A., 2012. The GEOS-5 Atmospheric General Circulation Model: Mean Climate and Development from MERRA to Fortuna.
Mondini, A.C., Guzzetti, F., Melillo, M., 2023. Deep learning forecast of rainfall-induced shallow landslides. Nat. Commun. 14, 2466. https://doi.org/10.1038/s41467-023-38135-y.
Nikolopoulos, E.I., Crema, S., Marchi, L., Marra, F., Guzzetti, F., Borga, M., 2014. Impact of uncertainty in rainfall estimation on the identification of rainfall thresholds for debris flow occurrence. Geomorphology 221, 286–297. https://doi.org/10.1016/j.geomorph.2014.06.015.
Osanai, N., Shimizu, T., Kuramoto, K., Kojima, S., Noro, T., 2010. Japanese early-warning for debris flows and slope failures using rainfall indices with Radial Basis Function Network. Landslides 7, 325–338. https://doi.org/10.1007/s10346-010-0229-5.
Pecoraro, G., Calvello, M., Piciullo, L., 2019. Monitoring strategies for local landslide early warning systems. Landslides 16, 213–231. https://doi.org/10.1007/s10346-018-1068-z.
Peruccacci, S., Brunetti, M.T., Gariano, S.L., Melillo, M., Rossi, M., Guzzetti, F., 2017. Rainfall thresholds for possible landslide occurrence in Italy. Geomorphology 290, 39–57. https://doi.org/10.1016/j.geomorph.2017.03.031.
Planet Team, 2017. Planet application program interface. San Francisco, CA. In: Space for Life on Earth. https://api.planet.com [WWW Document].
Reichenbach, P., Rossi, M., Malamud, B.D., Mihir, M., Guzzetti, F., 2018. A review of statistically-based landslide susceptibility models. Earth Sci. Rev. 180, 60–91. https://doi.org/10.1016/j.earscirev.2018.03.001.
Rosi, A., Segoni, S., Canavesi, V., Monni, A., Gallucci, A., Casagli, N., 2021. Definition of 3D rainfall thresholds to increase operative landslide early warning system performances. Landslides 18, 1045–1057. https://doi.org/10.1007/s10346-020-01523-2.
Sak, H., Senior, A., Beaufays, F., 2014. Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition arXiv Prepr. arXiv1402.1128.
Schulz, W.H., Smith, J.B., Wang, G., Jiang, Y., Roering, J.J., 2018. Clayey landslide initiation and acceleration strongly modulated by soil swelling. Geophys. Res. Lett. 45, 1888–1896. https://doi.org/10.1002/2017GL076807.
Segoni, S., Piciullo, L., Gariano, S.L., 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.
Singhal, A., 2001. Modern information retrieval: a brief overview. IEEE Data Eng. Bull. 24, 35–43.
Stanley, T.A., Kirschbaum, D.B., Sobieszczyk, S., Jasinski, M.F., Borak, J.S., Slaughter, S. L., 2020. Building a Landslide Hazard Indicator with Machine Learning and Land Surface Models, vol. 129. Environmental Modelling & Software, 104692.
Stanley, T.A., Kirschbaum, D.B., Benz, G., Emberson, R.A., Amatya, P.M., Medwedeff, W., Clark, M.K., 2021. Data-driven landslide nowcasting at the global scale. Front. Earth Sci.
Steger, S., Moreno, M., Crespi, A., Zellner, P.J., Gariano, S.L., Brunetti, M.T., Melillo, M., Peruccacci, S., Marra, F., Kohrs, R., Goetz, J., Mair, V., Pittore, M., 2023. Deciphering seasonal effects of triggering and preparatory precipitation for improved shallow landslide prediction using generalized additive mixed models. Nat. Hazards Earth Syst. Sci. 23, 1483–1506. https://doi.org/10.5194/nhess-23-1483-2023.
Suarez, M.J., Rienecker, M.M., Todling, R., Bacmeister, J., Takacs, L., Liu, H.C., Gu, W., Sienkiewicz, M., Koster, R.D., Gelaro, R., 2008. The GEOS-5 Data Assimilation System-Documentation of Versions 5, 0. 1, 5.1. 0, and 5.2. 0.
Tichavský, R., Ballesteros-Cánovas, J.A., Šilhán, K., Tolasz, R., Stoffel, M., 2019. Dry spells and extreme precipitation are the main trigger of landslides in central Europe. Sci. Rep. 9, 14560 https://doi.org/10.1038/s41598-019-51148-2.
Titti, G., Napoli, G.N., Conoscenti, C., Lombardo, L., 2022. Cloud-based interactive susceptibility modeling of gully erosion in Google Earth Engine. Int. J. Appl. Earth Obs. Geoinf. 115, 103089.
Wang, N., Lombardo, L., Gariano, S.L., Cheng, W., Liu, C., Xiong, J., Wang, R., 2021. Using satellite rainfall products to assess the triggering conditions for hydro-morphological processes in different geomorphological settings in China. Int. J. Appl. Earth Obs. Geoinf. 102, 102350 https://doi.org/10.1016/j.jag.2021.102350.
Zhao, L., Liu, M., Song, Z., Wang, S., Zhao, Z., Zuo, S., 2022. Regional-scale modeling of rainfall-induced landslides under random rainfall patterns. Environ. Model. Software 155, 105454.
Zhou, P., Shi, W., Tian, J., Qi, Z., Li, B., Hao, H., Xu, B., 2016. Attention-based bidirectional long short-term memory networks for relation classification. In: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, 2, pp. 207–212. Short Papers).