https://doi.org/10.5194/egusphere-2025-1590
Preprint. Discussion started: 7 May 2025

Use of delayed ERA5-Land soil moisture products for improving landslide early warning

Nunziarita Palazzolo[1], Antonino Cancelliere[1], Robert D. Zofei[1], David J. Peres[1]

[1]Department of Civil Engineering and Architecture, University of Catania, Catania, 95125, Italy

Correspondence to: Nunziarita Palazzolo (nunziarita.palazzolo@unict.it)

Abstract. Previous studies have demonstrated that incorporating ECMWF ERA5-Land soil moisture products can improve the predictive performance of landslide-triggering thresholds. However, these data are released with a five-day latency, which limits their immediate operational use in Landslide Early Warning Systems (LEWSs). In this study, we investigate whether delayed soil moisture data—ranging from 0 to 15 days prior to rainfall events—can still effectively inform landslide-triggering conditions. Specifically, we develop artificial neural networks (ANNs) trained on various delay times and evaluate how detection performances vary with increasing lag. Focusing on Sicily, Italy, our results show that even delayed soil moisture data consistently outperform models based solely on rainfall (TSS = 0.68 vs. 0.59). Notably, TSS reduces only marginally, from 0.78 with no delay to 0.72 with five-day delay, and 0.67 with fifteen-day delay. This performance remains higher than that obtained using only soil moisture data (without precipitation and no delay, TSS = 0.53), as well as those achieved with a traditional power-law threshold based on rainfall intensity and duration (TSS = 0.50) and also through ANN model using rainfall intensity and duration (TSS = 0.59). These findings are, thus, promising for an operational use of ERA5-Land soil moisture products in LEWSs.

1 Introduction

Rainfall-induced landslides rank among the most devastating natural hazards, causing extensive loss of life and severe impacts on the environment, infrastructure, and the economy (Winter et al., 2016). Between 2004 and 2016, 4862 fatal landslides were recorded worldwide, leading to an estimated 55,997 fatalities (Froude and Petley, 2018). Territorial Landslide Early Warning Systems (LEWSs) aim to promptly alert at-risk populations, thereby significantly reducing casualties and injuries. Identifying the conditions that trigger landslides is crucial for developing effective LEWSs. Early efforts primarily focused on rainfall intensity–duration thresholds (Caine, 1980; Guzzetti et al., 2007), but more recent studies highlight the value of incorporating additional hydrological variables—particularly soil moisture—to better capture the mechanisms driving slope failure (Uwihirwe et al., 2022; Mirus et al., 2018a, b; Thomas et al., 2018; Segoni et al., 2018; Wicki et al., 2021; Bogaard and Greco, 2016, 2018; Reder and Rianna, 2021; Marino et al., 2020; Conrad et al., 2021; Palazzolo et al., 2023; Meneses et al., 2019).

Machine learning methods, particularly Artificial Neural Networks (ANNs), have proven effective in handling multiple predictors for landslide forecasting, often surpassing classical parametric threshold models (e.g., power-law relationships).

https://doi.org/10.5194/egusphere-2025-1590
Preprint. Discussion started: 7 May 2025

For instance, Distefano et al. (2022), in a study focused on Sicily (Italy), demonstrated that precipitation-based thresholds derived via ANNs outperform traditional power-law equations. Similar evidence has been reported for broader regions of Italy (Mondini et al., 2023) and China (Peng and Wu, 2024).

Building on these findings, Distefano et al. (2023) showed that incorporating ERA5-Land reanalysis soil moisture products (Muñoz-Sabater et al., 2021)—specifically the soil moisture at the onset of rainfall events—can significantly improve ANN-based landslide predictions. However, because ERA5-Land data (provided by the European Centre for Medium-Range Weather Forecasts, ECMWF) are released with a five-day delay, their direct use in operational real-time LEWSs is constrained. Against this backdrop, the present study investigates the feasibility of using delayed ERA5-Land soil moisture data for real-time landslide forecasting. In particular, we assess the extent to which different publication lags (up to 15 days) affect the performance of ANN-based landslide prediction. The models are evaluated using receiver operating characteristic (ROC) statistics, and the methodology is tested on the island of Sicily (Italy).

2 Material and methods

2.1 Dataset creation

The proposed methodology is briefly schematized in in Fig. 1.

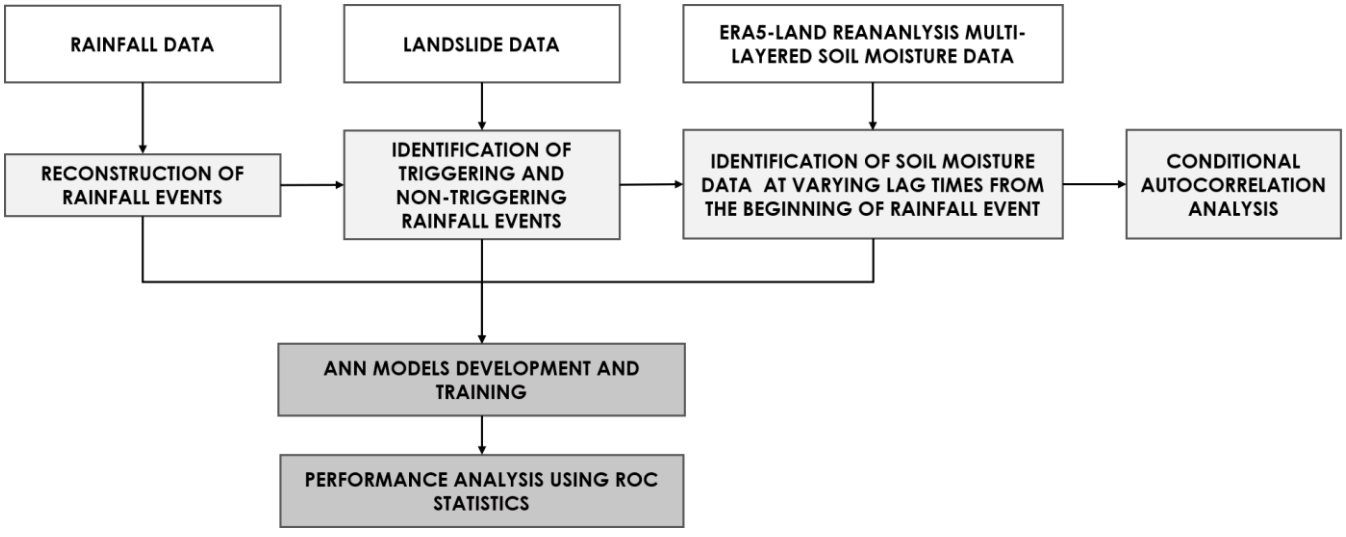

Figure 1: Scheme of the methodology applied to explore the potential use of delayed ERA5-Land soil moisture products for the identification of landslide triggering conditions

As illustrated, the first step of the methodology involves collecting rainfall and landslide data. Specifically, the FraneItalia inventory (Calvello and Pecoraro, 2018) is employed to gather information on observed landslides. This inventory documents historical landslides throughout Italy since 2010, encompassing both fatal and non-fatal events. This detailed dataset enables the identification and selection of landslide events whose main cause is intense and/or prolonged rainfall. The

https://doi.org/10.5194/egusphere-2025-1590
Preprint. Discussion started: 7 May 2025

CTRL-T (Calculation of Thresholds for Rainfall-Induced Landslides-Tool) code (Melillo et al., 2018) has then been applied to identify rainfall events most likely responsible for the observed slope failures. Briefly, the tool comprises different modules, each designed for a specific function. One of these modules reconstructs rainfall events in terms of duration (*D*, in hours) and cumulative event rainfall (*E*, in millimeters) by analyzing a continuous rainfall time series, along with various climate and spatial parameters. Another module, on the other hand, selects the most representative rain gauge for each landslide event. This tool allows, thus, the reconstruction of both triggering and non-triggering rainfall events. On the other side, ERA5-Land reanalysis dataset provides soil moisture values (ϑ [$m^3\ m^{-3}$]) at four distinct depths (i.e., 0–7 cm, 7–28 cm, 28–100 cm, and 100–289 cm), at the hourly scale and at a high spatial resolution of 0.1° x 0.1° corresponding to about 9 x 9 km grid at the latitudes of the case study area (see section Study area and data) (https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=download). Thus, multi-layered soil moisture data, from the ERA5-Land reanalysis project, at the onset of all rainfall events are then retrieved, together with the preceding values up to 15 days before.

2.2 Preliminary correlation analysis

An exploratory analysis is firstly conducted, in order to understand from a purely statistical standpoint, how delay affects the information content of soil moisture data from the ERA5-Land Reanalysis. To this aim, we computed a conditional autocorrelation, i.e. the autocorrelation between the soil moisture at the beginning of triggering and non-triggering rainfall events and its values in the preceding days, with time lags ranging from 0 to 15 days with respect to the rainfall event's timing. Specifically, for each cell corresponding to a rainfall station soil moisture at each depth and at onset of rainfall events is compared with values from preceding days. With reference to the generic cell corresponding to the rainfall station, let $S_{t(i)}$ with i=1, 2,…,N represent the soil moisture at the onset time $t(i)$ of the generic rainfall event i and k the lag time (in days). We have then computed the linear correlation coefficient $\rho(k)$ between soil moisture values at lag k and lag 0 as follows (Eq. 1):

$$\rho(k) = \frac{\sum_{i=1}^{N}(S_{t(i)}-\bar{S})(S_{t(i)-k}-\bar{S})}{\sqrt{\sum_{i=1}^{N}(S_{t(i)}-\bar{S})^2 \sum_{i=1}^{N}(S_{t(i)-k}-\bar{S})^2}} \quad (1)$$

where $S_{t(i)-k}$ denotes soil moisture lagged by k days relative to the onset time $t(i)$ and $\bar{S}$ is the mean of $S_{t(i)-k}$. Eq. (1) differs from the so-called autocorrelation for being conditioned to the beginning time of rainfall events. For this reason, we refer to $\rho(k)$ as a conditional autocorrelation. Eq. 1 has been applied with reference to all rainfall events, as well as to the triggering and non-triggering events.

2.3 Artificial Neural Network models (ANNs)

ANNs are a class of machine learning models designed to mimic the structure and functioning of neural systems, particularly how the human brain processes information (Haykin, 1994). They consist of numerous interconnected processing elements

https://doi.org/10.5194/egusphere-2025-1590
Preprint. Discussion started: 7 May 2025

(i.e., neurons) that work together to solve complex tasks. Similar to human learning, ANNs improve their performance through experience, adjusting the strength of connections between neurons (synapses) to improve their ability to recognize patterns and make predictions (Grosan and Abraham, 2011; Haykin, 1994). In this study, the following input variables were considered: (i) precipitation duration *D*, (ii) cumulative precipitation *E*, (iii) volumetric soil water content *S* at different time-lag relatively to the starting instant of rainfall events. Regarding the lagged soil water content, we have chosen to consider the data relative only to a given lag. In other words, the ANNs for a given lag of *k* days consider as input soil moisture data only for that lag, and not for larger lags. All input variables considered have been transformed to their natural logarithms – this has been empirically proven to improve the efficiency of ANN calibration (Distefano et al., 2022). Fig. 2 shows the architecture of the implemented ANN model and the input variables considered in this study.

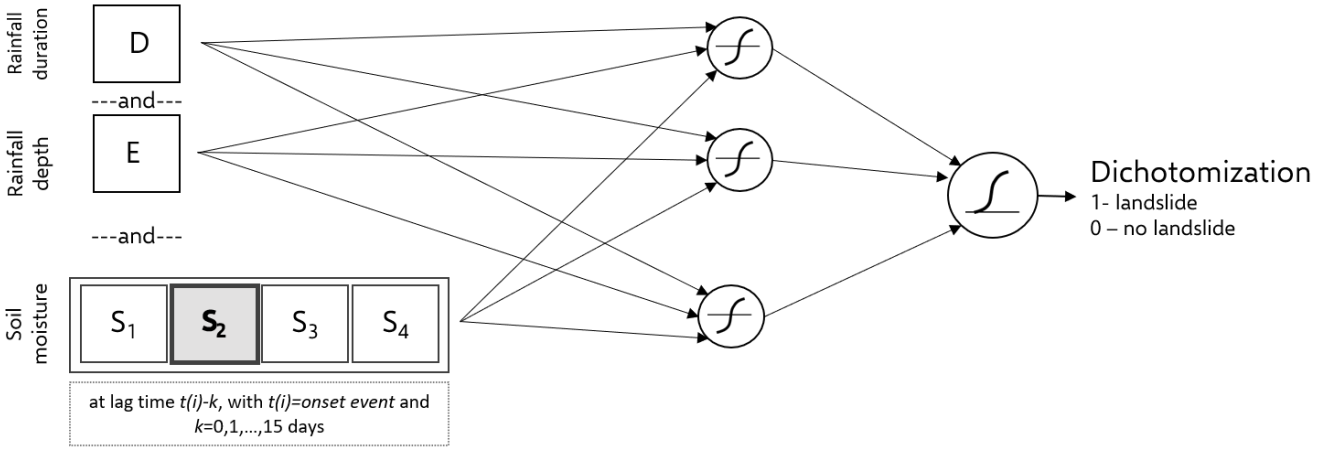

Figure 2: Sketch of the implemented ANN model. The subscripts of soil moisture indicate the soil layer (depth range) of the ERA5-Land reanalysis dataset. The second layer is highlighted with a different color, being the most influential one for the case study area as reported in Distefano et al. (2022).

The MATLAB® Deep Learning toolbox has been used to develop the ANN models. Specifically, the scaled conjugate gradient backpropagation algorithm (Møller, 1993) has been used to train the ANNs, while the "cross-entropy" function (Kline and Berardi, 2005) has been used as a measure of performance for training the networks. The entire dataset of both triggering and non-triggering rainfall events has been randomly split into three subsets: 70% for training, 15% for validation, and 15% for testing. In doing so, each network has been trained 30 times to assess the impact of random data subdivision on training, validation, and test samples. For each of these training sessions the number of neurons has been varied from 5 to 20, choosing at each session the number that corresponds to the highest performances; when presenting the results, the average, extended to the 30 training sessions of the best number of hidden neurons is shown.

The following simulations using different input datasets: (i) soil moisture at the second-layer depth only; (ii) rainfall duration (*D*, hours), rainfall depth (*E*, mm), and soil moisture at the second-layer depth; and (iii) rainfall duration (*D*), rainfall depth (*E*), and soil moisture at all depth layers. These input combinations were selected based on the findings of Distefano et al.

https://doi.org/10.5194/egusphere-2025-1590
Preprint. Discussion started: 7 May 2025

(2023), who reported that the highest performance was achieved when considering soil moisture from the second layer, with only slightly higher performances with all four layers.

To assess the accuracy of the ANNs' output, hence the performance of the model's prediction, ROCs (Receiver Operating Characteristics) analysis is carried out. Specifically, the True Skill Statistic (TSS) index is computed for each simulation, corresponding to every investigated combination of input variables. TSS is a measure of how much better a model performs compared to random guessing, making it a reliable measure of the model's ability, i.e., ANNs, to correctly classify both positive and negative cases, namely triggering and non-triggering rainfall events. Its computation is based on the confusion matrix which provides a detailed breakdown of the model's predictions compared to actual outcomes (Fawcett 2006). It is a square matrix where rows represent the actual classes and columns represent the predicted classes. Thus, for a binary classification problem, like the discrimination between triggering and non- triggering conditions, the matrix distinguishes: i) True Positives (TP) as triggering events correctly identified as triggering events; ii) True Negatives (TN) as non-triggering events correctly identified as non-triggering events; iii) False Positives (FP) as non-triggering events incorrectly classified as triggering events (false alarms); and iv) False Negatives (FN) as triggering events incorrectly classified as non-triggering events (missed alarms). Therefore, TSS (Eq. 4) index is defined as the difference between the True Positive Rate (TPR) (Eq. 2) and the False Positive Rate (FPR) (Eq. 3). The highest performances correspond to TSS=1 when the model correctly predicts for all the analysed data.

$$TPR = \frac{TP}{(TP+FN)} \quad (2)$$

$$FPR = \frac{FP}{(TN+FP)} \quad (3)$$

$$TSS = TPR - FPR \quad (4)$$

The TSS is computed after dichotomization of the original output of the ANN which is in the range [0 1], by maximizing the TSS.

3 Study area and data

Sicily (southern Italy) is selected as study area to perform our analyses (Fig. 3). Annual rainfall in the region ranges on average from 700 to 800 mm, with the majority occurring during the autumn and winter months. Red triangles, within the study area map, represent the 207 landslides retrieved by FraneItalia database from 2010 to 2018 along with their longitude–latitude coordinates (WGS84 datum) and the triggering time information, if the latter is available (otherwise, it is assumed at the end of the day). The map also shows the ERA5-Land grid at which extent soil moisture data are released. Blue circles, instead, represent the locations of the 306 gauges installed in Sicily and managed by the three main Sicilian gauging network, namely the Regional Water Observatory (Osservatorio delle Acque, OdA), the Sicilian Agrometeorological

https://doi.org/10.5194/egusphere-2025-1590
Preprint. Discussion started: 7 May 2025

Information Service (Servizio Informativo Agrometeorologico Siciliano, SIAS), and the Regional Department of Civil Protection (Dipartimento Regionale di Protezione Civile, DRPC). Rainfall data are, thus, provided at the hourly scale from 2009 up to 2018 (Distefano et al., 2022).

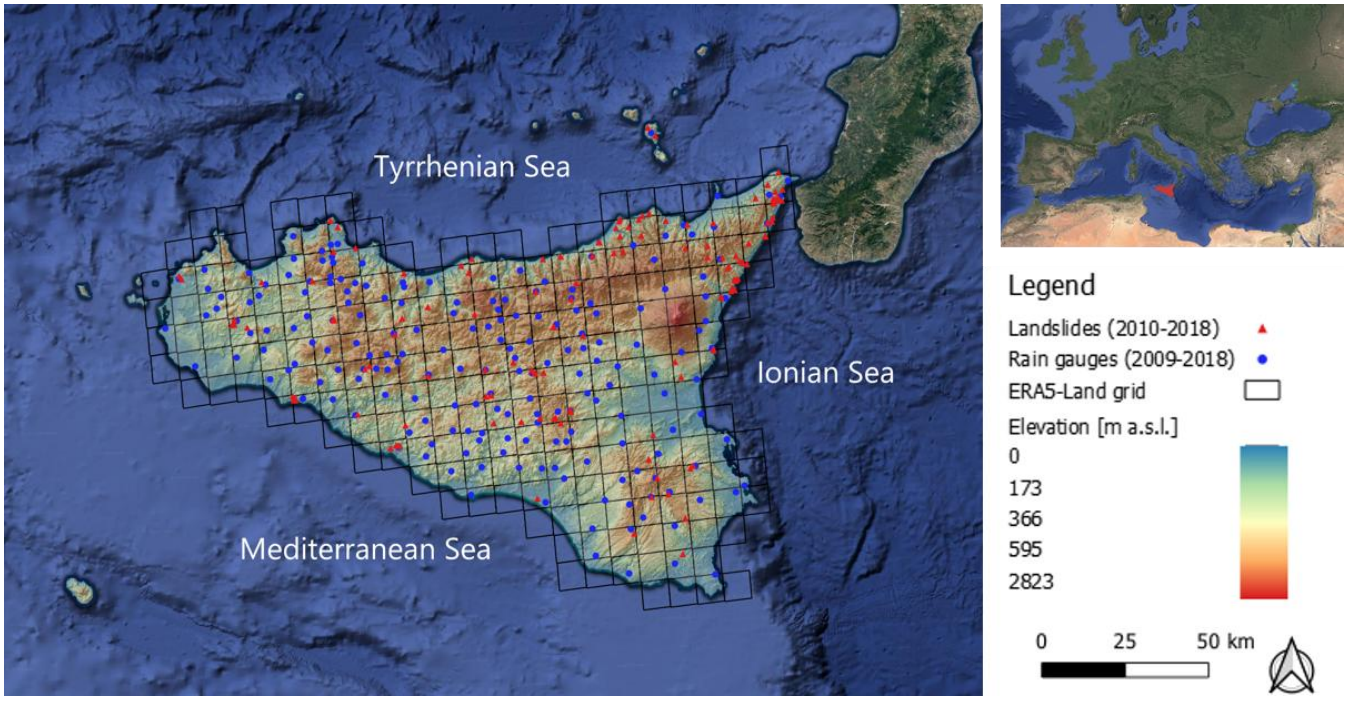

Figure 3: Study area map ((credit to http://www.sinanet.isprambiente.it/it/sia-ispra/download, last access: 01 April 2025, and ESRI, 2020).

As mentioned in the previous section, CTRL-T tool enables the identification of both triggering and non-triggering rainfall events for each rainfall station based the hourly rainfall, and by setting a combination of regional parameters. These latter are aimed at taking into account the seasonality of the climate, namely at defining the warm period C_W and the cold one C_C; the dry interval separating isolated rainfall measurements (P_1); the time periods used to remove irrelevant amounts of rainfall, (P_2), and (P_3); and the minimum dry period separating two rainfall events (P_4). Moreover, even the resolution of the rain gauge (G_S), as well the instrumental sensitivity of the rain gauge and the radius of the buffer to assign each landslide to the closest rain gauge (R_B) can be set. For further details on these parameters, readers are referred to Melillo et al. (2018). Table 1 summarize the parameters' values adopted for our analysis which are in line with Melillo et al. (2015). To ensure a consistent comparison with the findings of Distefano et al. (2023), this study adopts the same dataset, enriched with reanalysis multilayered soil moisture data at the cell closest to the precipitation station. More specifically, for each rainfall station, corresponding soil moisture data is associated to each rainfall event considering a delay ranging from 1 to 15 days.

https://doi.org/10.5194/egusphere-2025-1590
Preprint. Discussion started: 7 May 2025

Table 1: CTRL-T parameters for the reconstruction of the rainfall events used in the present study (after Distefano et al., 2022).

G_S	E_R	R_B	P_1 [h]		P_2 [h]		P_3 [h]		P_4 [h]	
[mm]	[mm]	[km]	C_w	C_c	C_w	C_c	C_w	C_c	C_w	C_c
0.2	0.2	16	3	6	6	12	1	1	48	96

4 Results and discussion

For all four soil layers, the conditional autocorrelation function has a decreasing trend, even though for the deepest layer the decay is minimal (Fig. 4).

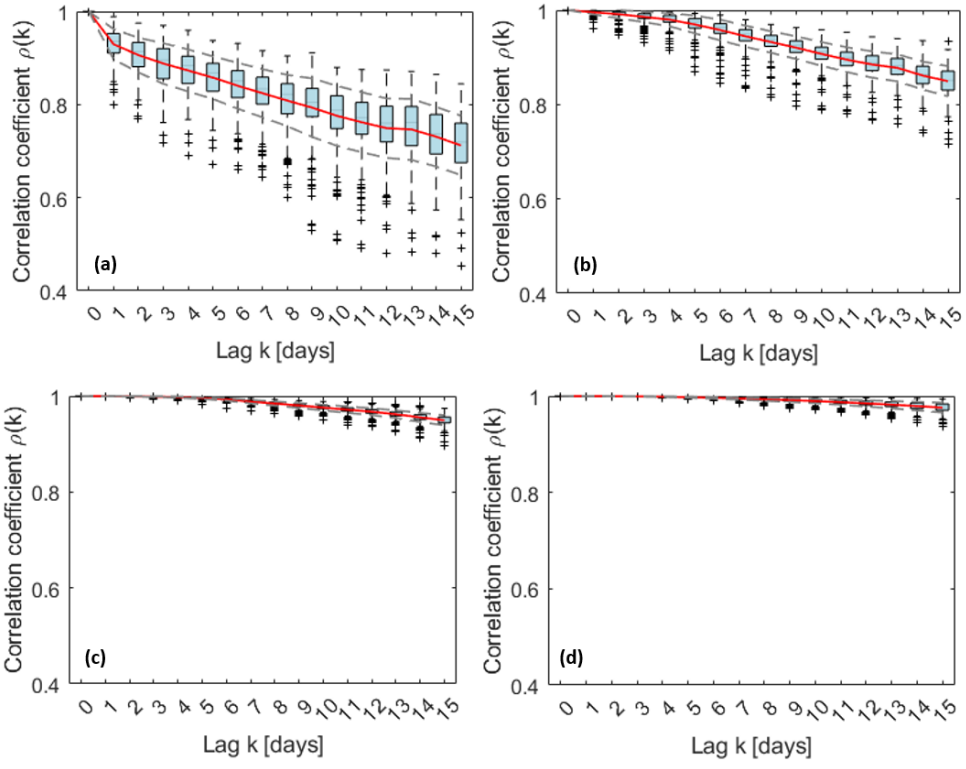

Figure 4: Conditional autocorrelation function for soil moisture at the four ERA5-Land depths (time lags from 1 to 15 days) (a) 0-7 cm; (b) 7-28 cm; (c) 28-100 cm; (d) 100-239 cm.

The shallowest layers, particularly the first one, appear to be the most responsive to the effect of lag, showing a notable drop in conditional autocorrelation as early as lag = 1. This is consistent with the fact that shallow layers are mostly sensitive to the forcing of even small rainfall events that induce high frequency fluctuations in soil moisture, with a consequent reduction of short short-term memory in the process. As deeper layers are considered, high frequency forcings tend to be dampened by the presence of the above layers resulting in a more stable pattern in time and stronger autocorrelation, at least for the investigated lags.

https://doi.org/10.5194/egusphere-2025-1590
Preprint. Discussion started: 7 May 2025

Moving to the results of ANNs for the prediction of landslide triggering conditions, starting with those using only soil moisture data without precipitation (Table 2), a TSS of 0.53 is achieved when no lag is considered. This performance is slightly higher than that obtained with a traditional approach (i.e., the power-law threshold based on rainfall intensity and duration, which has a TSS of 0.50; Palazzolo et al., 2023) but slightly lower than that achieved by Distefano et al. (2023), where an ANN model using rainfall intensity and duration lead to significantly higher performances (TSS = 0.59). When a 5-day lag is considered, TSS drops further to 0.43.

Table 2: Performances of ANNs using exclusively soil moisture at the second layer (7-28 cm).

Lag [days]	**Hidden neurons** [mean]	**TSS_{all}** [mean]	**TSS_{train}** [mean]	**TSS_{val}** [mean]	**TSS_{test}** [mean]	**TPR_{all}** [mean]	**FPR_{all}** [mean]
0	9.17	**0.53**	0.57	0.50	0.48	0.74	0.21
1	5.50	0.48	0.52	0.45	0.45	0.72	0.24
2	12.40	0.44	0.48	0.42	0.42	0.73	0.29
3	15.40	0.44	0.47	0.42	0.42	0.78	0.33
4	15.60	0.44	0.47	0.41	0.41	0.78	0.35
5	6.20	**0.43**	0.46	0.39	0.39	0.71	0.28
6	8.93	0.42	0.46	0.38	0.38	0.71	0.29
7	11.30	0.40	0.44	0.36	0.38	0.68	0.28
8	13.53	0.39	0.44	0.35	0.34	0.64	0.25
9	17.70	0.41	0.44	0.36	0.38	0.66	0.26
10	13.87	0.38	0.42	0.34	0.35	0.64	0.27
11	9.27	0.36	0.40	0.33	0.32	0.61	0.25
12	10.13	0.35	0.38	0.33	0.32	0.73	0.38
13	12.00	0.34	0.39	0.34	0.31	0.70	0.36
14	11.57	0.35	0.39	0.32	0.33	0.73	0.38
15	11.73	0.36	0.40	0.32	0.34	0.72	0.36

As expected, including rainfall duration (*D*, [hours]) and rainfall depth (*E*, [mm]) besides soil moisture leads to a significant improvement in performance (Fig. 5). In more detail, when *D*-*E*-S_2 dataset is considered as input to the model, a TSS of 0.76 is achieved, while incorporating the entire *D*-*E*-S_{all} dataset increases the TSS to 0.78. Nevertheless, in both cases, performance reduces as the time-lag increases. Focusing on a 5-day time delay a slight drop in performance is observed, with TSS values of 0.72 and 0.70, respectively. Although these values are lower, they are still above the ANN model based solely on rainfall data. As the time delay increases, performance continues to marginally decrease, reaching a TSS of around 0.68 at the 15-day time lag in both analyzed scenarios, namely a decrease of just 0.1 points.

https://doi.org/10.5194/egusphere-2025-1590
Preprint. Discussion started: 7 May 2025

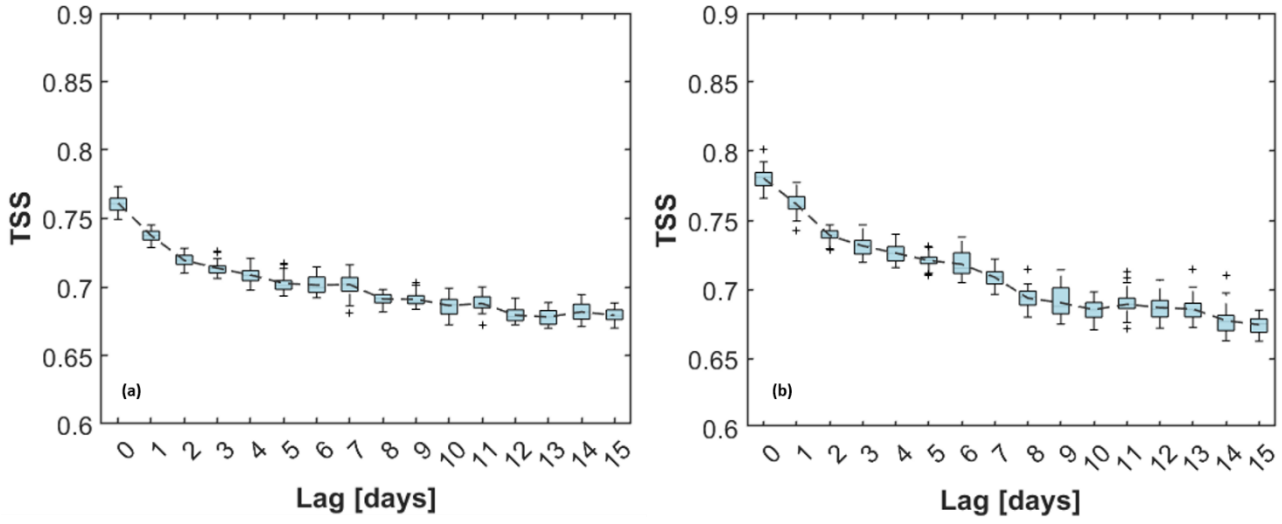

Figure 5: TSS values from 30 ANN training runs, with the dashed grey line representing the mean TSS across lag time. (a) simulation set with rainfall duration (D, [hours]), rainfall depth (E, [mm]) and soil moisture at the second-layer depth as input dataset; (b) simulation set with rainfall duration (D, [hours]), rainfall depth (E, [mm]) and soil moisture at all -layer depth as input dataset.

Overall, box-plots present a narrow interquartile range, supporting the adequacy of the input dataset and confirming robustness of the results. More details about simulations carried out in correspondence of the investigating scenarios are summarized within Table 3 and 4.

Table 3: Performances of ANNs using rainfall duration, depth and soil moisture at the second layer (7-28 cm).

Lag [days]	Hidden neurons [mean]	TSS_{all} [mean]	TSS_{train} [mean]	TSS_{val} [mean]	TSS_{test} [mean]	TPR_{all} [mean]	FPR_{all} [mean]
0	13.23	**0.76**	0.78	0.73	0.73	0.89	0.12
1	13.33	0.74	0.76	0.71	0.71	0.88	0.15
2	14.83	0.72	0.74	0.69	0.70	0.88	0.16
3	13.33	0.71	0.74	0.69	0.68	0.88	0.17
4	14.37	0.71	0.74	0.69	0.68	0.88	0.17
5	14.30	**0.70**	0.73	0.68	0.67	0.87	0.16
6	14.60	0.70	0.73	0.69	0.67	0.88	0.18
7	15.03	0.70	0.73	0.67	0.68	0.88	0.18
8	15.20	0.69	0.72	0.67	0.66	0.87	0.18
9	13.63	0.69	0.72	0.67	0.65	0.85	0.16
10	15.10	0.69	0.71	0.67	0.66	0.85	0.16
11	15.63	0.69	0.72	0.67	0.65	0.86	0.17

https://doi.org/10.5194/egusphere-2025-1590
Preprint. Discussion started: 7 May 2025

12	14.53	0.68	0.71	0.65	0.64	0.85	0.17
13	15.07	0.68	0.71	0.64	0.67	0.87	0.19
14	16.60	0.68	0.71	0.66	0.66	0.88	0.20
15	15.10	0.68	0.71	0.66	0.65	0.86	0.18

Table 4: Performances of ANNs using rainfall duration, depth and soil moisture at all four layers.

Lag [days]	**Hidden neurons** [mean]	**TSS_{all}** [mean]	**TSS_{train}** [mean]	**TSS_{val}** [mean]	**TSS_{test}** [mean]	**TPR_{all}** [mean]	**FPR_{all}** [mean]
0	15.20	**0.78**	0.81	0.76	0.75	0.89	0.11
1	14.07	0.76	0.79	0.74	0.74	0.89	0.13
2	15.53	0.74	0.77	0.72	0.72	0.90	0.16
3	16.80	0.73	0.76	0.71	0.70	0.89	0.15
4	15.47	0.73	0.75	0.71	0.70	0.89	0.17
5	13.93	**0.72**	0.75	0.71	0.70	0.90	0.18
6	14.60	0.72	0.74	0.70	0.69	0.89	0.17
7	16.30	0.71	0.74	0.69	0.67	0.88	0.17
8	15.10	0.69	0.72	0.67	0.67	0.86	0.16
9	15.60	0.69	0.72	0.67	0.67	0.85	0.16
10	14.37	0.69	0.71	0.68	0.67	0.85	0.16
11	15.33	0.69	0.72	0.68	0.66	0.85	0.16
12	14.30	0.69	0.72	0.66	0.66	0.85	0.17
13	14.60	0.69	0.71	0.67	0.66	0.88	0.20
14	16.30	0.68	0.71	0.66	0.65	0.89	0.21
15	15.23	0.67	0.70	0.65	0.65	0.88	0.20

In these two simulation sets, the results indicate that, on average, using 15 neurons provides the best performance, providing a suitable balance between model complexity and generalization capabilities.

Notably, a lower number of neurons may lead to underfitting, limiting the network's capacity to capture the patterns in the data. Conversely, increasing the neuron beyond 15 does not yield significant improvements and may introduce unnecessary complexity, potentially leading to overfitting. In this regard, looking at the simulation set exclusively using S_2 as input, the number of hidden neurons decreases at 9, on average, reasonably due to the lower complexity of the input dataset.

As a whole, results show that delayed ERA5-Land soil moisture information can be considered a useful predictor for landslide prediction, although the performance of the prediction model slightly decreases as the time-lag increases. Indeed, the improved performance, even with lagged data, can be explained by the slow variation of soil moisture during dry periods

https://doi.org/10.5194/egusphere-2025-1590
Preprint. Discussion started: 7 May 2025

preceding rainfall events. Consequently, soil moisture at the onset of a rainfall event does not differ significantly from that recorded five or more days earlier. Thus, while soil moisture at the beginning of rainfall events provides the highest performance, incorporating delayed soil moisture information still results in an improved performance and, most importantly, can be readily incorporated for operational landslide prediction.

5 Conclusions

This study proposed an ANN-based framework to evaluate whether ERA5-Land reanalysis soil moisture data can effectively serve as real-time data for landslide prediction. Building on the work of Distefano et al. (2023), our investigation focused on the classification capabilities of ANNs to identify triggering conditions using rainfall duration, rainfall depth, and delayed multi-layered soil moisture information. Specifically, ERA5-Land data are released with an approximately 5-day delay, limiting their immediate availability for real-time landslide prediction. The analysis was conducted with a focus on Sicily, Italy. Three simulation sets were examined, progressively increasing in complexity based on the number of variables considered, with performance evaluated in terms of TSS. The highest performance was achieved when rainfall duration, rainfall depth, and all available soil moisture information were used as input into ANNs. In this case, results demonstrated that soil moisture lagged data did not significantly undermine landslide prediction performance, as the TSS decreased from 0.78 with lag=0 to 0.72 with a 5-day lag. Even with a lag of 15 days, the prediction performance of ANNs is still significantly higher than those models using only rainfall event characteristics, corresponding to a decrease of just 0.06 points. These findings underscore the potential of integrating ERA5-Land multi-layered soil moisture data into LEWSs, despite their delayed release. Indeed, their regular availability in both time and space enhances their suitability as a reliable tool for these systems. Further developments of this study will consider the use of rainfall forecasts at various temporal horizons in order to assess effective the operational predictive capability of the investigated tools.

Author contributions

Conceptualization was done by D.J.P., A.C. and N.P.; formal analysis by D.J.P., A.C. and N.P.; investigation by D.J.P. and N.P.; methodology by D.J.P. and N.P.; coding by D.J.P., N.P. and R.D.Z.; writing the original draft by D.J.P., A.C., N.P. and R.D.Z. All authors have read and agreed to the published version of the paper.

Data availability

The FraneItalia landslides catalog is available at https://doi.org/10.17632/zygb8jygrw.2 (Calvello and Pecoraro, 2018). Rainfall measurements are available at the website of the Servizio Informativo Agrometeorologico Siciliano (SIAS) (http://www.sias.regione.sicilia.it/, SIAS, 2025) and at the Osservatorio delle Acque

https://doi.org/10.5194/egusphere-2025-1590
Preprint. Discussion started: 7 May 2025

(http://www.bio.isprambiente.it/annalipdf/, ISPRA, 2025). Reanalysis soil moisture data are available from https://doi.org/10.24381/cds.e2161bac (Muñoz-Sabater et al., 2021)

Acknowledgements

The authors would like to thank Valerio Santoro for his valuable input in the early stages of this work.

References

Bogaard, T. and Greco, R.: Invited perspectives: Hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: Proposing hydro-meteorological thresholds, Nat. Hazards Earth Syst. Sci., 18, 31–39, https://doi.org/10.5194/NHESS-18-31-2018, 2018.

Bogaard, T. A. and Greco, R.: Landslide hydrology: from hydrology to pore pressure, Wiley Interdiscip. Rev. Water, 3, 439–459, https://doi.org/10.1002/WAT2.1126, 2016.

Caine, N.: The Rainfall Intensity: Duration Control of Shallow Landslides and Debris Flows, Geogr. Ann. Ser. A, Phys. Geogr., 62, 23, https://doi.org/10.2307/520449, 1980.

Calvello, M. and Pecoraro, G.: FraneItalia: a catalog of recent Italian landslides, 5, 1–16, https://doi.org/10.1186/S40677-018-0105-5/FIGURES/11, 2018.

Conrad, J. L., Morphew, M. D., Baum, R. L., and Mirus, B. B.: HydroMet: A New Code for Automated Objective Optimization of Hydrometeorological Thresholds for Landslide Initiation, Water 2021, Vol. 13, Page 1752, 13, 1752, https://doi.org/10.3390/W13131752, 2021.

Distefano, P., Peres, D. J., Scandura, P., and Cancelliere, A.: Brief communication: Introducing rainfall thresholds for landslide triggering based on artificial neural networks, Nat. Hazards Earth Syst. Sci., 22, 1151–1157, https://doi.org/10.5194/NHESS-22-1151-2022, 2022.

Distefano, P., Peres, D. J., Piciullo, L., Palazzolo, N., Scandura, P., and Cancelliere, A.: Hydro-meteorological landslide triggering thresholds based on artificial neural networks using observed precipitation and ERA5-Land soil moisture, 20, 2725–2739, https://doi.org/10.1007/S10346-023-02132-5/TABLES/6, 2023.

https://doi.org/10.5194/egusphere-2025-1590
Preprint. Discussion started: 7 May 2025

ESRI: "Ocean" [basemap], Scale Not Given, "World Ocean Base map", https://services.arcgisonline.com/arcgis/rest/services/ Ocean/World_Ocean_Reference/MapServer (last access: 20 March 2025) 2020.

Froude, M. J. and Petley, D. N.: Global fatal landslide occurrence 2004 to 2016, https://doi.org/10.5194/NHESS-2018-49, 2018.

Grosan, C. and Abraham, A.: Artificial Neural Networks, Intell. Syst. Ref. Libr., 17, 281–323, https://doi.org/10.1007/978-3-642-21004-4_12, 2011.

Guzzetti, F., Peruccacci, S., Rossi, M., and Stark, C. P.: Rainfall thresholds for the initiation of landslides in central and southern Europe, Meteorol. Atmos. Phys., 98, 239–267, https://doi.org/10.1007/S00703-007-0262-7, 2007.

Haykin, S.: Neural networks: a comprehensive foundation, Prentice hall PTR, 1994.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Q. J. Roy. Meteor. Soc., 146, 1999– 2049, https://doi.org/10.1002/qj.3803, 2020.

Kline, D. M. and Berardi, V. L.: Revisiting squared-error and cross-entropy functions for training neural network classifiers, Neural Comput. Appl., 14, 310–318, https://doi.org/10.1007/S00521-005-0467-Y/TABLES/2, 2005.

Marino, P., Peres, D. J., Cancelliere, A., Greco, R., and Bogaard, T. A.: Soil moisture information can improve shallow landslide forecasting using the hydrometeorological threshold approach, 17, 2041–2054, https://doi.org/10.1007/S10346-020-01420-8/FIGURES/3, 2020.

Melillo, M., Brunetti, M. T., Peruccacci, S., Gariano, S. L., and Guzzetti, F.: An algorithm for the objective reconstruction of rainfall events responsible for landslides, 12, 311–320, https://doi.org/10.1007/S10346-014-0471-3/FIGURES/7, 2015.

Melillo, M., Brunetti, M. T., Peruccacci, S., Gariano, S. L., Roccati, A., and Guzzetti, F.: A tool for the automatic calculation of rainfall thresholds for landslide occurrence, Environ. Model. Softw., 105, 230–243, https://doi.org/10.1016/J.ENVSOFT.2018.03.024, 2018.

https://doi.org/10.5194/egusphere-2025-1590
Preprint. Discussion started: 7 May 2025

Meneses, B. M., Pereira, S., and Reis, E.: Effects of different land use and land cover data on the landslide susceptibility zonation of road networks, Nat. Hazards Earth Syst. Sci., 19, 471–487, https://doi.org/10.5194/NHESS-19-471-2019, 2019.

Mirus, B. B., Morphew, M. D., and Smith, J. B.: Developing Hydro-Meteorological Thresholds for Shallow Landslide Initiation and Early Warning, Water 2018, Vol. 10, Page 1274, 10, 1274, https://doi.org/10.3390/W10091274, 2018a.

Mirus, B. B., Becker, R. E., Baum, R. L., and Smith, J. B.: Integrating real-time subsurface hydrologic monitoring with empirical rainfall thresholds to improve landslide early warning, 15, 1909–1919, https://doi.org/10.1007/S10346-018-0995-Z/FIGURES/6, 2018b.

Møller, M. F.: A scaled conjugate gradient algorithm for fast supervised learning, 6, 525–533, https://doi.org/10.1016/S0893-6080(05)80056-5, 1993.

Mondini, A. C., Guzzetti, F., and Melillo, M.: Deep learning forecast of rainfall-induced shallow landslides, Nat. Commun. 2023 141, 14, 1–10, https://doi.org/10.1038/s41467-023-38135-y, 2023.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J. N.: ERA5-Land: A state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349–4383, https://doi.org/10.5194/ESSD-13-4349-2021, 2021.

Palazzolo, N., Peres, D. J., Creaco, E., and Cancelliere, A.: Using principal component analysis to incorporate multi-layer soil moisture information in hydrometeorological thresholds for landslide prediction: an investigation based on ERA5-Land reanalysis data, Nat. Hazards Earth Syst. Sci., 23, 279–291, https://doi.org/10.5194/NHESS-23-279-2023, 2023.

Peng, B. and Wu, X.: Optimizing rainfall-triggered landslide thresholds for daily landslide hazard warning in the Three Gorges Reservoir area, Nat. Hazards Earth Syst. Sci., 24, 3991–4013, https://doi.org/10.5194/NHESS-24-3991-2024, 2024.

Reder, A. and Rianna, G.: Exploring ERA5 reanalysis potentialities for supporting landslide investigations: a test case from Campania Region (Southern Italy), 18, 1909–1924, https://doi.org/10.1007/S10346-020-01610-4/FIGURES/11, 2021.

Segoni, S., Tofani, V., Rosi, A., Catani, F., and Casagli, N.: Combination of rainfall thresholds and susceptibility maps for dynamic landslide hazard assessment at regional scale, Front. Earth Sci., 6, 375062, https://doi.org/10.3389/FEART.2018.00085/BIBTEX, 2018.

Thomas, M. A., Collins, B. D., and Mirus, B. B.: Assessing the Feasibility of Satellite-Based Thresholds for Hydrologically Driven Landsliding, Water Resour. Res., 55, 9006–9023, https://doi.org/10.1029/2019WR025577, 2018.

https://doi.org/10.5194/egusphere-2025-1590
Preprint. Discussion started: 7 May 2025

Uwihirwe, J., Hrachowitz, M., and Bogaard, T.: Integration of observed and model-derived groundwater levels in landslide threshold models in Rwanda, Nat. Hazards Earth Syst. Sci., 22, 1723–1742, https://doi.org/10.5194/NHESS-22-1723-2022, 2022.

Wicki, A., Jansson, P. E., Lehmann, P., Hauck, C., and Stähli, M.: Simulated or measured soil moisture: Which one is adding more value to regional landslide early warning?, Hydrol. Earth Syst. Sci., 25, 4585–4610, https://doi.org/10.5194/HESS-25-4585-2021, 2021.

Winter, M. G., Shearer, B., Palmer, D., Peeling, D., Harmer, C., and Sharpe, J.: The Economic Impact of Landslides and Floods on the Road Network, Procedia Eng., 143, 1425–1434, https://doi.org/10.1016/J.PROENG.2016.06.168, 2016.