



Landslide displacement forecasting using deep learning and monitoring data across selected sites

Abstract Accurate early warning systems for landslides are a reliable risk-reduction strategy that may significantly reduce fatalities and economic losses. Several machine learning methods have been examined for this purpose, underlying deep learning (DL) models' remarkable prediction capabilities. The long short-term memory (LSTM) and gated recurrent unit (GRU) algorithms are the sole DL model studied in the extant comparisons. However, several other DL algorithms are suitable for time series forecasting tasks. In this paper, we assess, compare, and describe seven DL methods for forecasting future landslide displacement: multi-layer perception (MLP), LSTM, GRU, 1D convolutional neural network (1D CNN), 2xLSTM, bidirectional LSTM (bi-LSTM), and an architecture composed of 1D CNN and LSTM (Conv-LSTM). The investigation focuses on four landslides with different geographic locations, geological settings, time step dimensions, and measurement instruments. Two landslides are located in an artificial reservoir context, while the displacement of the other two is influenced just by rainfall. The results reveal that the MLP, GRU, and LSTM models can make reliable predictions in all four scenarios, while the Conv-LSTM model outperforms the others in the Baishuihe landslide, where the landslide is highly seasonal. No evident performance differences were found for landslides inside artificial reservoirs rather than outside. Furthermore, the research shows that MLP is better adapted to forecast the highest displacement peaks, while LSTM and GRU are better suited to model lower displacement peaks. We believe the findings of this research will serve as a precious aid when implementing a DL-based landslide early warning system (LEWS).

Keywords Landslide hazard · Remote sensing · Landslide early warning · Landslide forecasting · Artificial intelligence

Introduction

Every year, landslides, a major global geohazard, result in a tremendous number of fatalities and economic damage (Froude and Petley 2018). Predicting future landslide evolution is critical for risk assessment and the design of reliable early warning systems for landslides produced by external triggers such as rainfall, reservoir and groundwater fluctuations, earthquakes, and anthropogenic activities (Aleotti and Chowdhury 1999). It can be difficult to predict how landslides will behave in the future because their progression typically does not adhere to a linear pattern. Several external elements (above all, rainfall, and water level reservoir variations) coordinate and impact landslide future displacements with varying magnitudes and intensities. Physical-based and

data-driven techniques are the two main approaches for forecasting landslide displacement (Huang et al. 2017). The physical-based models mainly use creep theory, laboratory tests, and physical characteristics determined in situ to predict the landslide behavior. However, data-driven models are more widely used due to their straightforward methodology, precise predictions, lower costs, and scalability (Corominas et al. 2013). Recent breakthroughs in landslide displacement forecasting have demonstrated that artificial intelligence (AI) and, in particular, deep learning techniques are cutting edge. By accounting for contributing factors in addition to landslide displacement data, these approaches may accurately predict future velocity.

Landslides at the Three Gorges Reservoir have been investigated and utilized as benchmarks for the implementation, evaluation, and development of numerous AI-based techniques for landslide displacement forecasting (Yao et al. 2015). Several machine learning (ML) models have been used to this end such as artificial neural networks (ANN) (Du et al. 2013), support vector machines (SVM) (Zhu and Hu 2012; Zhou et al. 2016; Zhu et al. 2017; Wen et al. 2017; Miao et al. 2018; Ma et al. 2020; Wang et al. 2020, 2022; Han et al. 2021; Zhang et al. 2021), Gaussian process (Liu et al. 2014), and extreme learning machine (ELM) (Lian et al. 2015; Cao et al. 2016; Huang et al. 2017; Zhou et al. 2018a; Wang et al. 2019). The authors quite often used influencing factors in these situations which include previous landslide displacements, prior rainfall information, water level of reservoirs, and water level variations. In addition, Li et al. (2019) included a wavelet analysis-Volterra filter model (chaotic WA-Volterra model) based on chaos theory in SVM. The WA-Volterra model separates the high- and low-frequency components of cumulative displacement data. Chaos theory was used to recreate the spatial structure of each frequency, which was then used as input into the model. His paper investigates the possibility of chaotic characteristic identification of landslide displacements to be employed in machine learning.

Deep belief networks, LSTM-based architectures, gated recurrent unit neural networks (GRU), and other DL techniques have recently been employed in landslide displacement predictions (Yang et al. 2019; Xing et al. 2019; Li et al. 2020; Zhang et al. 2021; Guo et al. 2022). The relationship between landslide displacement and the primary triggering elements, for instance, was examined by Yang et al. (2019). In his paper, the periodical displacement was predicted using an LSTM model that considers the most crucial historical conditioning parameters. The LSTM model can gather insights from previous deformation time steps and create relationships between landslide conditions at different times. The results

demonstrated that the LSTM model performs better than the SVM techniques. Wang et al. (2022) tested various ML models, including particle swarm optimization (PSO)-extreme learning machine (PSO-ELM), PSO-kernel extreme learning machine (PSO-KELM), PSO-support vector machine (PSO-SVM), PSO-least-squares support vector machine (PSO-LSSVM), and LSTM. Their research found that LSTM and PSO-ELM performed superior single predictions, but PSO-KELM and PSO-LSSVM yielded higher mean accuracies across cases.

Although there is a large body of research that suggests methodologies for landslide forecasting in extremely unique study cases such as the Three Gorges Reservoir landslides, few studies have used ML to estimate landslide displacement in different geographical, geological, and hydrogeological contexts. Based on historical data from landslides, rainfall, and groundwater, a new Random Forest-based technique for forecasting landslide velocity was proposed by Krkač et al. (2017) in the largest landslide in the Republic of Croatia. The landslide velocity was modeled using the expected daily groundwater levels predicted by using the historical daily precipitation data. Lastly, Zhu et al. (2017) compared and evaluated the performances of two different configurations of LSSVMs in a rainfall-sensitive slow-moving landslide located in the Sichuan Province, China.

This literature review shows that despite a significant amount of literature addressing the prediction capabilities of various ML-based algorithms and approaches for landslides in the Three Gorges Reservoir, there are few studies in which ML is utilized to the same end for study cases located in diverse places. However, in the world, a huge number of landslides are located outside reservoir contexts. Several critical slow-moving landslides are not located in reservoirs, and their behavior is particularly sensitive to rainfall events. Lastly, so far, the authors often evaluate the performances of ML algorithms to the performance of a single DL model (usually

LSTM). However, several DL models are suited and employed successfully in a wide range of time series forecasting tasks.

In this paper, the abovementioned research gaps are addressed by evaluating, comparing, and discussing the predictive capabilities of seven state-of-art DL algorithms, in four study cases that differ in terms of geographical location, influencing factors, geological settings, time step dimensions, and measurement sensors.

Case studies and materials

Sant'Andrea and Lamosano landslides are both located in the Dolomites, an area of the Southern-Eastern Italian Alps, in the Province of Belluno (Veneto Region, NE Italy), in Perarolo di Cadore and Chies D'Alpago municipalities, respectively. The Baishuihe landslide is located in Three Gorges Reservoir in the Hubei province, Central China. The El Arrecife landslide is located within the western slope of the Rules Reservoir, in the Granada province (Southern Spain) (see Fig. 1).

Sant'Andrea landslide

The Sant'Andrea landslide affects the left-hand slope of the valley, overlooking the Boite river course, in the area just upstream of the Perarolo village (Fig. 2). The position of the landslide poses a significant risk to the inhabitants since a potential collapse of the unstable mass could result in temporary damming on the Boite river, which can cause flooding of the downstream area. The maximum temperature is 15 °C, while the minimum is −7 °C. Several surveys were conducted over several years to gather spatially distributed information on the geological characteristics of the landslide area. The distribution of lithological units, in particular, was performed by combining information gathered from on-site surveys as well as geological and geotechnical investigations (Brezzi et al. 2021). The

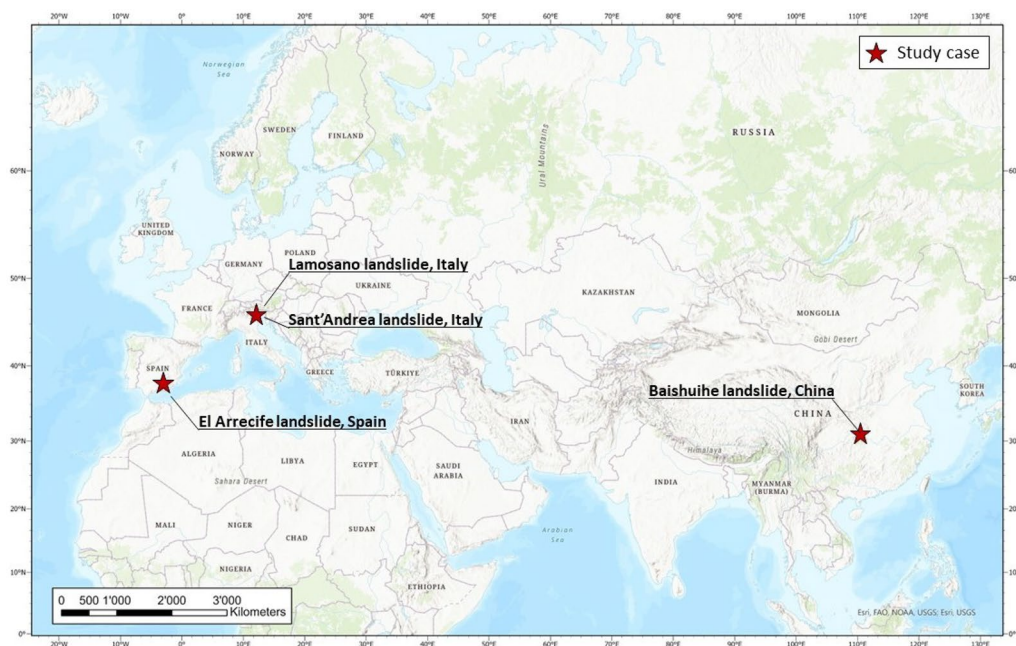


Fig. 1 The geographical location of the four different study cases investigated in this work

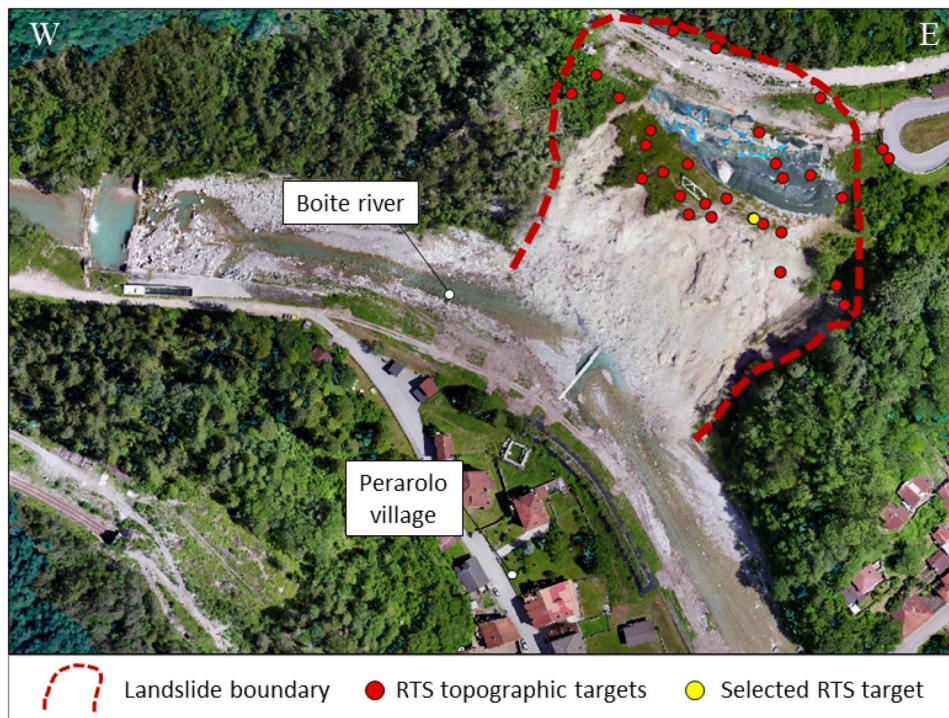


Fig. 2 Sant'Andrea landslide site. In the orthophoto, the targets of the topographic monitoring system are shown, as well as the boundary of the unstable area

Sant'Andrea landslide is a 30-m-thick deposit of clay-calcareous debris composed of heterogeneous materials with different grain sizes and geotechnical characteristics. The debris mass slides across the weathered part of the bedrock composed of a dolomitic lithology and folded layers rich in anhydrides and gypsum.

The landslide activity shows an interplay of phases with slow displacements and accelerations that are primarily triggered by prolonged and intense rainfall. The characterization of the complex hydrogeological setting within the landslide was made possible by the information that was provided regarding the distribution of the geological units, which aids the interpretation of its behavior. Two circulation systems have been recognized within the unstable mass (Brezzi et al. 2021): a shallow groundwater flows in the upper layers of the debris deposits, whereas a deep one involves the upper part of the bedrock, mainly composed of altered and fractured gypsum. However, the role of water seems to be the main triggering factor of slope instability. On the one hand, the water circulations during the rainfall events cause the acceleration of the displacements. On the other hand, the active and deep circulation, coming from the upper part of the slope, induces slow displacements also in dry periods. This dynamic is related to the physical and chemical interaction between water and gypsum components of the upper part of the bedrock, as well as in the surficial debris layers, influencing the mechanical properties of rock mass (Brezzi et al. 2021). Hydration processes cause plastic rheology of the weak gypsum lithology, which drives the creep-like behavior of the unstable mass and the wide slope instability. Moreover, the dissolution induces an increase in the number of voids, resulting in the development of

karst cavities and millimeter- to centimeter-thick microcrack net both in the bedrock and in the altered gypsum of the fractured part. Consequently, the water circulation in the landslide area makes the mechanical behavior of the gypsum lithology of the bedrock quite unpredictable, leading to the hazard of a sudden collapse of the unstable mass.

The Sant'Andrea landslide has been monitored using a topographic system since the end of 2013. This system is composed of a robotic total station (RTS) and several reflective targets installed on the unstable slope (Fig. 2). The implementation of the RTS targets was prompted by the evolution of landslides as various regions demonstrated deterioration of the stability conditions over time. The P4 target was chosen for this study because it has a 4-year displacement time series and is situated in an area affected by significant displacements. Figure 3 illustrates the daily differential displacement that was recorded by the P4 target during the period 2014–2019. The increase in the displacement rate is strongly related to rainfall events and, as previously stated, the duration of the rainfall also affects the landslide activity. The original measured time step dimension is 1 h. However, we decide to use daily frequency since it offers benefits such as reduced noise, as daily data can smooth out short-term fluctuations, data aggregation, providing a higher-level overview of data patterns, and improved computational efficiency, as there are fewer data points to process. Lastly, being the forecasting range equal to one time step, daily forecasts are more relevant for decision-making and planning compared to hourly forecasts, which may be subject to short-term fluctuations and may not be as useful.

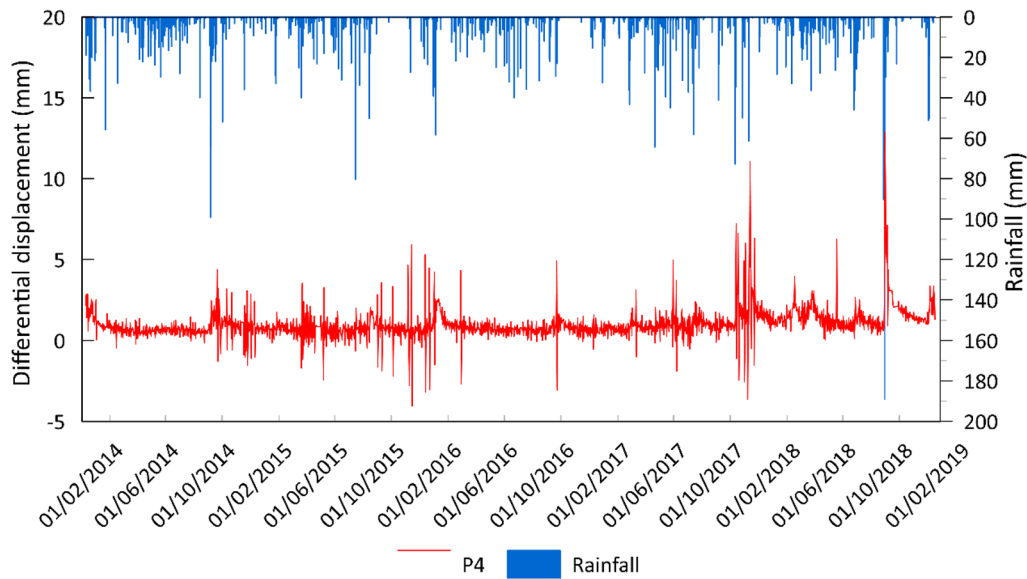


Fig. 3 Differential displacement of the selected RTS target for the monitoring of the Sant'Andrea landslide. In the same plot, the amount of precipitation is also reported. The time step dimension of the showed time series is 1 day

Lamosano landslide

The Lamosano landslide is located in the local village of Lamosano, in the Chies D'Alpago municipality (Fig. 4). This slope instability involves the area where several buildings have been built in the past, causing structural damages observed in recent years. The maximum temperature is 20 °C, while the minimum is −1 °C. The Lamosano landslide is classified as a rotational slow-moving landslide and its volume was estimated to be roughly $4.5 \times 10^6 \text{ m}^3$, currently

moving toward a WSW direction (Teza et al. 2008). Although a comprehensive characterization of the landslide is not available yet, the geological materials involved in the movement are mostly schistose clayey marls (upper layer of the bedrock), a grayish sandstone (lower layer of the bedrock), and a detrital cover of sandy gravel deposits (Pieraccini et al. 2006).

The landslide activity shows an interaction of phases with slow displacements and accelerations that, as previously described also for the Sant'Andrea case, are primarily triggered by prolonged and

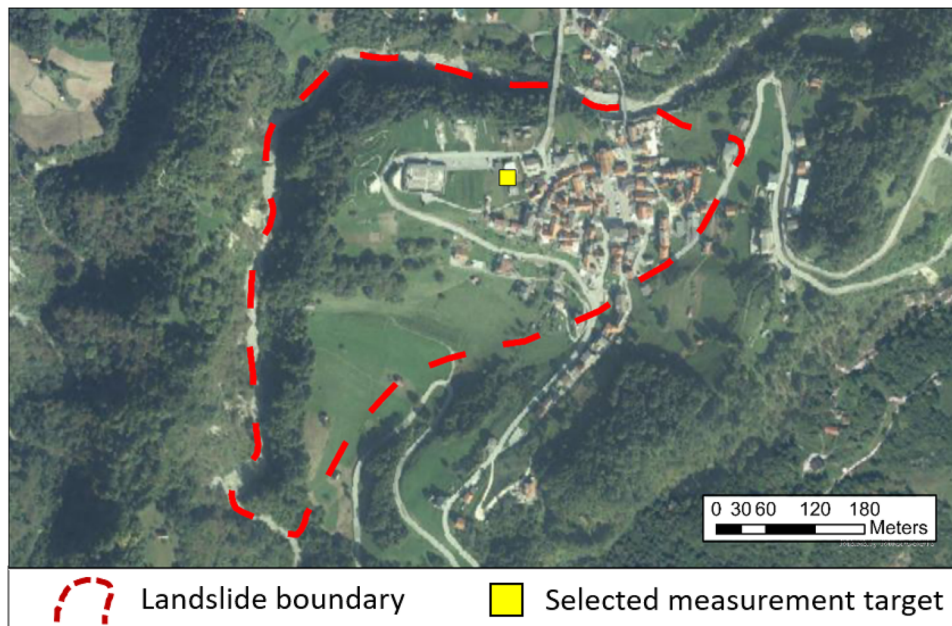


Fig. 4 Lamosano landslide site. The selected InSAR measurement point and the boundary of the unstable area are reported

intense rainfall. As reported in Fig. 5, these peaks of intense rainfall are related to changes in the horizontal displacement component detected by the InSAR remote sensing technique.

In this study, the landslide was monitored by using the C-band Sentinel-1 horizontal component from 30 March 2015 to 8 February 2020. Time series were extrapolated after processing 229 ascending and 249 descending images through the Small Baseline Subset (SBAS) algorithm (Berardino et al. 2002) in Sarscape software. We considered the horizontal component because it exhibits the most evident step-like movements related to rainfall events. Moreover, horizontal displacement is the moving component that has the most implications for the element at risk in the area. The time step dimension of the modeled time series is 11 days. This frequency is the smallest available for the selected landslide using the above-described InSAR processing method.

Baishuihe landslide

The Baishuihe landslide affects the right-hand slope of the Yangtze valley, in the Three Gorges Reservoir. Figure 6 reports the main topographical features of this slope instability. The highest elevation of the landslide boundary is 297 m a.s.l., with the crack acting as the rear barrier to the less unstable area (Song et al. 2020). The frontal edge is immersed in water, and its elevation is approximately 120 to 130 m a.s.l., which is always 145 m below the Yangtze river water level in the reservoir. The landslide measures 500 m in length from north to south and 430 m in width from east to west, encompassing an area of 215,000 m² with an average depth of around 30 m. The main sliding inclination is 20° and the volume has been estimated at 6.45 million m³ (Li et al. 2010). The landslide body is mainly composed of silty clay and gravelly soil, which appear alternately and its thickness ranges from 7.5 to 37.7 m. The slip surface is mainly composed of gravel-containing or breccia-containing silty clay; in addition, some parts are full of breccia and clay. The thickness of

the failure zone varies from 0.2 to 1.3 m with an average thickness of 0.7 m (Yang et al. 2019).

The landslide is located in the subtropical monsoon climate zone, which is characterized by substantial rainfall and distinct seasons. The maximum temperature is 42 °C instead the minimum is – 8.9 °C (Song et al. 2020), where the average temperature ranges between 17 and 19 °C. June through September is the flood season; instead, October to May is considered the non-flood season. When persistent and intense rainfall events occur during the flood season, the landslide deforms significantly. Consequently, precipitations can be considered one of the major factors triggering the deformation of the Baishuihe landslide. Moreover, the water level in the reservoir fluctuates periodically every year. Usually, the reservoir water level is maintained at about 175 m in November and December. From early January to May, the water level gradually drops from 175 to 145 m; during this period, the deformation of the Baishuihe increases (see Fig. 7). In June and July, the reservoir water level is at a low level which fluctuates slightly to about 145 m. Then, the reservoir water level begins to rise gradually at the end of July and early August. From October on, the water level rises again, and, at the same time, the velocity of the Baishuihe deformation slows down (Keqiang et al. 2008; Li et al. 2010; Miao et al. 2021). Likewise, the water level is an indispensable factor when studying the activity of the Baishuihe landslide.

Since June 2003, the Baishuihe landslide has been monitored by six manual GNSS monitoring stations, named XD-01, XD-02, XD-03, XD-04, ZG93, and ZG118 respectively (Li et al. 2008). The locations of the monitoring stations are reported in Fig. 6 and these allow a wide characterization of the landslide displacements. Middle and posterior areas of the landslide are the most unstable (Li et al. 2010). As shown respectively in Figs. 6 and 7, ZG118 has been utilized in this study for its comprehensive data. In addition, station ZG118 demonstrates a notable deformation with the longest time series and a considerably greater displacement. The time step dimension

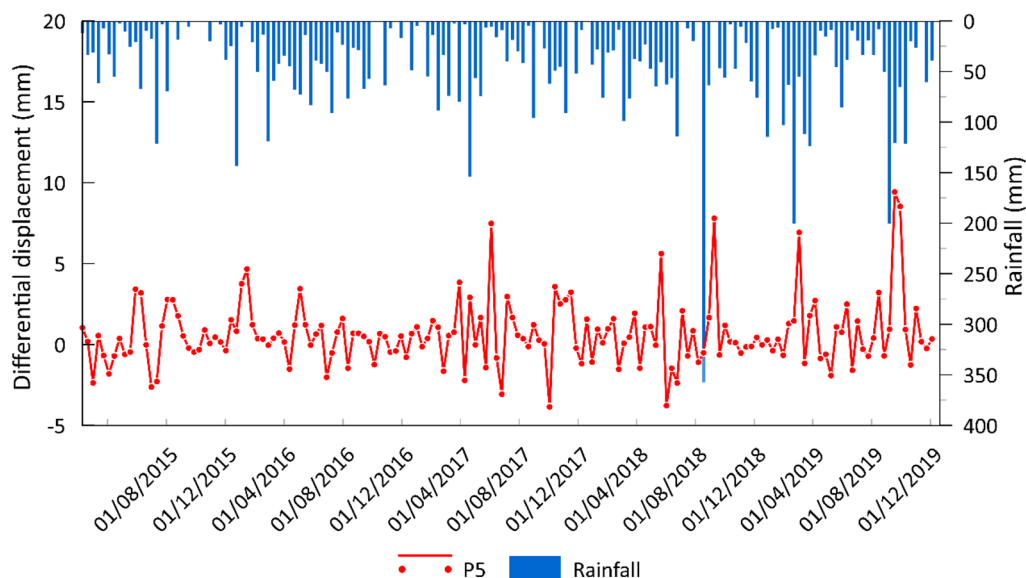


Fig. 5 Differential displacement trend of the selected InSAR monitoring point of the Lamosano landslide. In the same graph, the amount of rainfall is also reported. The time step dimension of the showed time series is 11 days

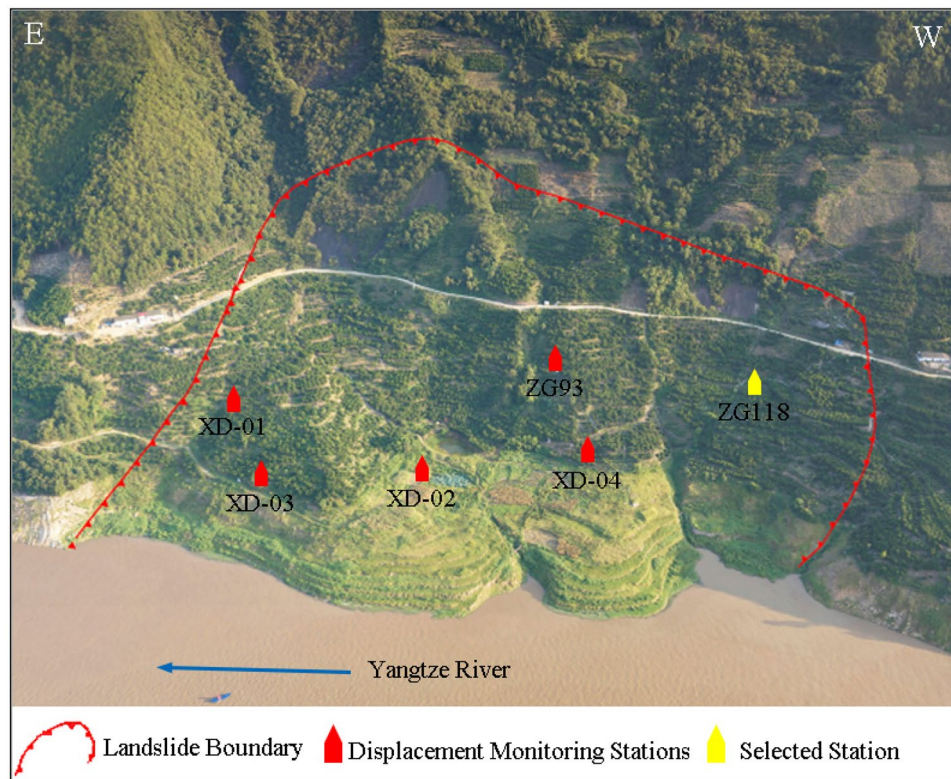


Fig. 6 Baishuihe landslide site. The GNSS monitoring stations and the boundary of the unstable area are reported (Guo et al. 2022)

of the modeled time series is 1 month. This frequency is the highest available for the selected landslide.

El Arrecife landslide

The El Arrecife landslide affects the slope directly overlooking the Rules Reservoir. This landslide has been recently identified and characterized by Reyes-Carmona et al. (2021, 2020). It is classified as a translational landslide with a sliding surface dipping 21° toward $N120^\circ E$, parallel to the mean orientation of the slope, while the landslide's foot is affected by several smaller-scale rotational slides. The landslide involves an area of 473.107 m^2 , with a mean thickness of 31.1 m , and a volume of 14.7 million m^3 , which makes the El Arrecife to be considered an extremely large landslide (see Fig. 8). Its most notable feature is that it lacks well-defined landslide morphologies (e.g., large head scarps or lateral scarps) and is difficult to identify in the landscape. Furthermore, geological variables such as lithology and geological structure contribute to the formation of the El Arrecife landslide. The slope is composed of phyllites, which are coarse-grained metamorphic rocks with a planar fabric and a low friction angle. Because of these characteristics, the slope has a high potential for instability and landsliding. As the landslide is situated near a reservoir, it poses a threat not only to the reservoir itself but also to other facilities (such as highways, viaducts, and powerlines) and the local inhabitants. Due to its translational nature, it is impossible to rule out the possibility of a critical acceleration of the entire landslide mass and a catastrophic failure of the slope. If an impulse wave is generated and the reservoir's dam

is breached, resulting in a downstream flash flood, this scenario would have catastrophic implications. Despite this, it is more likely that the smaller-scale rotating slides at the base of the landslide may cause greater damage to other infrastructures, such as the N-323 National Road, which crosses the El Arrecife landslide.

The landslide activity exhibits a linear movement pattern with small accelerations in the lower sector of the unstable area which are related to a decrease in the water level of the Rules Reservoir. Two slight accelerations were observed during two periods of water level decline from 2017 to 2019 (Reyes-Carmona et al.), whereas times of no change or rise in the water level resulted in the movement stabilizing or not accelerating. In contrast, rainfall seems unrelated to landslide accelerations since it leads to reservoir filling and, thus, relative slope stabilization. Therefore, the reservoir's water level variation is the only factor that triggers the movement of the El Arrecife landslide, as it has a greater effect on the movement than rainfall.

The landslide was first monitored through the InSAR technique from March 2015 to September 2018 (Reyes-Carmona et al. 2020). A mean surface ground displacement of 25 mm/year and a maximum of 55 mm/year were observed, as well as a cumulative displacement of 10 cm during that period (approx. 3.5 years). Later, the InSAR monitoring has been also extended to 2014, providing the landslide activity from December 2014 to March 2020 and resulting in a mean displacement of up to 20 mm/year (Reyes-Carmona et al. 2021). Figure 9 reports a frame of the InSAR monitoring, from October 2016 to March 2020, in which the relationship between the landslide displacement and the reservoir water table is recognized.

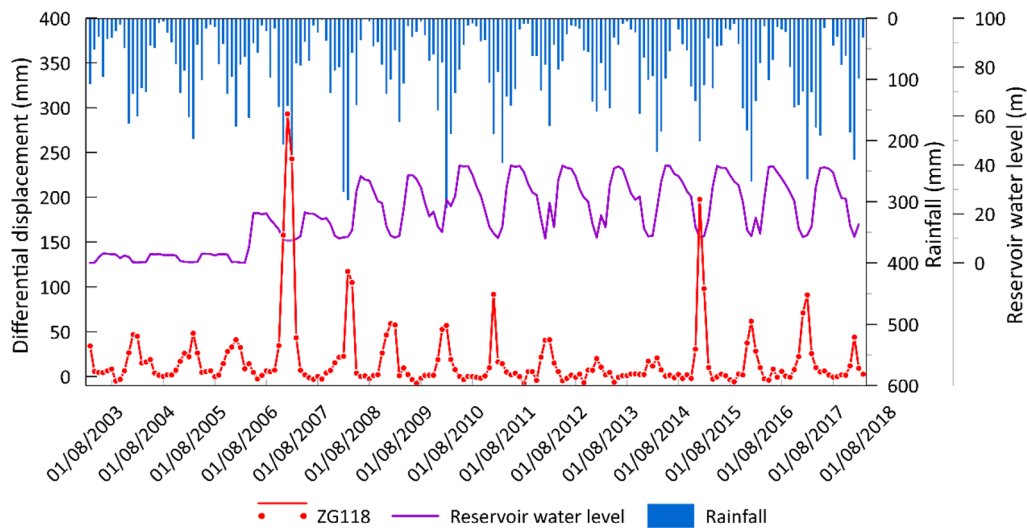


Fig. 7 Differential displacement of the selected GNSS monitoring station in the Baishuihe landslide. In the same graph, the amount of rainfall is also reported, as well as the water level of the reservoir (0 corresponds to the first value of the level of the reservoir). The time step dimension of the showed time series is 1 month

Additionally, a ground penetrating radar (GPR) study has been performed from 1997 to 2020, evidencing that the landslide has been active for the last 22 years: as a result, a vertical ground subsidence of up to 23 mm/year was estimated.

In this study, the landslide was monitored through the C-band Sentinel-1B satellite from 30 September 2015 to 13 March 2020. Time series with a temporal resolution of 12 days were extrapolated after processing 101 ascending orbit images by using the Parallel Small Baseline Subset (P-SBAS) algorithm (Casu et al. 2014) implemented in the European Space Agency’s (ESA) geohazards exploitation platform (GEP) (De Luca et al. 2015). The time step dimension of the modeled time series is 12 days. This frequency is the highest

available for the selected landslide using the above-described InSAR processing.

Methodology

Data pre-processing

In reservoir landslide displacement prediction studies using machine learning, cumulative landslide displacement is frequently divided into a trend term and a periodic component (Du et al. 2013; Zhou et al. 2018b; Yang et al. 2019). In these cases, the evolutionary tendency of the landslide is represented by a so-called trend-term displacement, whereas a periodic-term displacement refers to

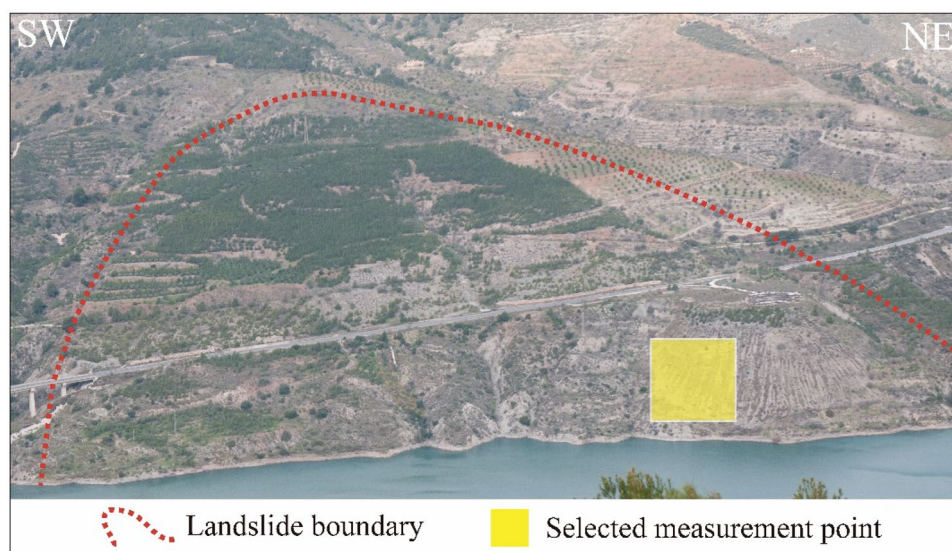


Fig. 8 El Arrecife landslide site. The selected InSAR measurement point is reported, as well as the boundary of the unstable area

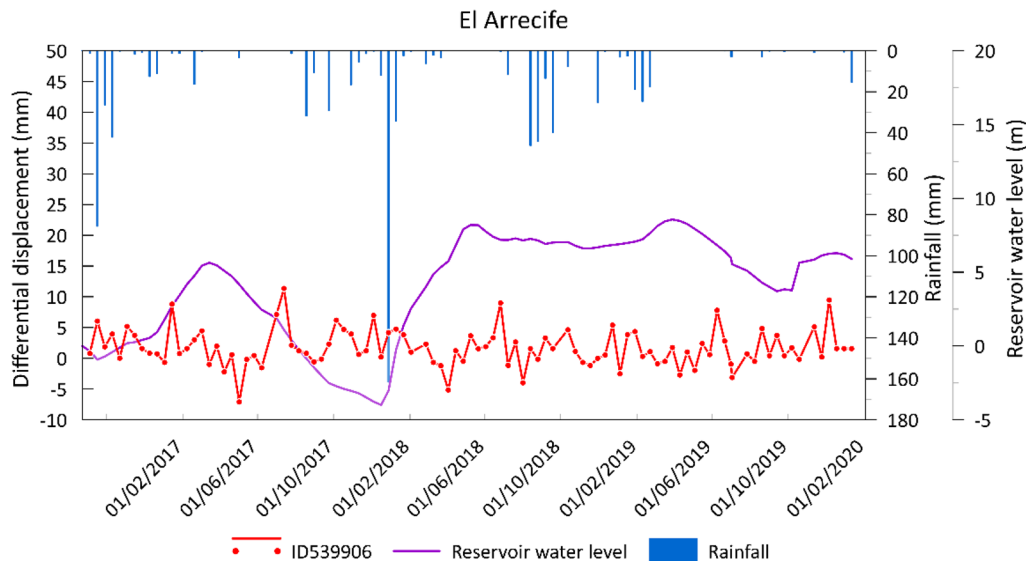


Fig. 9 Differential displacement trend of the selected InSAR measurement point in the El Arrecife landslide. In the same graph, the amount of rainfall is also reported, as well as the water level variation in the reservoir. The time step dimension of the showed time series is 2 weeks

variations in displacement due to periodic triggers. The two terms (trend and periodic displacements) are then forecasted independently, and the final expected cumulative displacement is produced by combining the two predictions. However, this method works on the assumption that the trend term is dependent just upon itself, and no external trigger influences it. However, depending on the decomposition strategy, this assumption might not always be true. Moreover, by removing level fluctuations in a time series and so decreasing trend and seasonality, differencing can help to stabilize the mean of a data series, and therefore, it helps to improve the prediction capabilities of data-driven models (Montesino Pouzols and Lendasse 2010). For this reason, we detrend both the displacement and reservoir water level time series by differencing. In detail, the value at the current time step is determined as the difference between the original observation and the observation at the previous time step. This is valid for all the timesteps, except the first of the series. By avoiding decomposition and opting for differencing as a detrending approach, the research aims to minimize potential uncertainties and complexities associated with decomposition strategies, while maintaining simplicity and effectiveness in trend removal from the landslide displacement time series data.

Deep learning algorithms

This paper evaluates seven cutting-edge deep learning architectures for time series forecasting and compares them. Among them, MLP, LSTM, GRU, and 1D CNNs are commonly used in several forecasting tasks. However, to the author's knowledge, although 1D CNNs are quite common for several time series forecasting tasks (Kiranyaz et al. 2021), (Kiranyaz et al. 2021), such a model has never been used for landslide displacement forecasting. We decided to evaluate also the bidirectional LSTM since it has already been used for landslide displacement forecasting (Lin et al. 2022). Moreover, some configurations based

on the combination of the abovementioned models are proposed. For instance, the 2xLSTM model is nothing but the combination of two LSTM layers. This configuration is expected to perform better than a single LSTM layer when dealing with complex problems. Another proposed model is the Conv-LSTM, which performs well in several time series forecasting tasks. Further details of the used models can be found in the Supplementary materials.

Multi-layer perceptron (MLP) (see Fig. S1) The ANN model is one of the most widely used models that has been effectively utilized in various applications, including time series modeling and forecasting. The appeal of ANN models stems from extrapolating the underlying data flow without constraints on the model structure. Another notable aspect of ANNs is that they are universal approximators capable of accurately approximating a wide range of functions. In the literature, there are indeed a variety of ANN model architectures. Whereas ANNs have a similar structure, distinctions among different algorithms have been made based on how they are designed. The so-called multi-layer perceptron (MLP), a three-layer feed-forward network (input, hidden, and output), is the most widely used ANN architecture for time series forecasting (Khashei and Hajirahimi 2019).

Recurrent neural networks (RNNs) Three types of RNNs are used in this research, namely long short-term memory (LSTM), gated recurrent unit (GRU), and bidirectional LSTM. Traditional neural networks are fully linked to the input-implicit-output layer. There is no connection between the positions in the sequences. As a result, time series prediction cannot be performed using a standard neural network (Chen and Chou 2012). The recurrent neural network (RNN), which enables feedback in the networks (Sak et al. 2014), retains the prior knowledge and applies it to the current output computation.

LSTM (see Fig. S2): A unique variety of RNNs called LSTM neural networks may successfully address the issue of “gradient explosion” (Hochreiter and Schmidhuber 1997). The forgetting gate, input gate, and output gate of the LSTM assure data discovery and long-term memory. The three gate functions offer a reliable nonlinear control mechanism for the input and deletion of control information. Following forward propagation, the backpropagation through time (BPTT) method transfers the accumulated error back in time and determines the gradient of error-associated parameters. Finally, the stochastic gradient descent method updates the weights and thresholds. Two architectures, the first with one LSTM layer and the second with two LSTM layers (2xLSTM), are evaluated.

Bidirectional LSTM (see Fig. S4): Even though the LSTM model fixes the gradient vanishing issue with RNNs, it can only learn knowledge from the past and cannot consider future data. Concerning the training phase, landslide motion is connected to both past and future displacement information, which may be used as a supplement while training the model since the causes impacting it is repeated in time. Architecture-wise, Bi-LSTM has forward and backward LSTM layers. The backward LSTM extracts forthcoming data regarding the current time step while in the opposite direction from the forward LSTM model, which leverages the current time step to retrieve past information. The Bi-LSTM architecture has been utilized effectively in several forecasting applications, such as solar radiation (Peng et al. 2021), well log (Shan et al. 2021), and tourism demand (Kulshrestha et al. 2020).

GRU (see Fig. S3): Recurrent unit is an RNN variation that performs better while being more streamlined (Cho et al. 2014; Zhao et al. 2018). The GRU is similar to a long short-term memory (LSTM) with a forget gate, but it has fewer parameters since it eliminates an output gate. GRU outperformed LSTM on specific tasks such as audio modeling, voice signal modeling, and natural language analysis (Ravanelli et al. 2018).

1D convolutional neural network (Conv) (see Fig. S5) Convolutional neural networks (CNNs), a subset of ANNs, are widely used in a wide range of applications, such as image and video recognition,

image classification, and natural language processing. CNN utilizes filters of comparable form to the data processed and dependent on the shape of the data. Because time series (Gamboa 2017) are one-dimensional data, 1-dimensional CNNs (Amarasinghe et al. 2017; Kiranyaz et al. 2021) are used for the landslide displacement forecasting task. A conv1D layer’s output sequence is shorter than its input sequence. This is corrected by padding the input sequence with zeros at both ends, also known as the same padding. A filter is moved over the input sequence using a Conv1D layer. One stride means that the filter is moved across the input sequence one step at a time. This is the first research in which a 1D convolutional-based architecture is used to forecast landslide displacement.

Conv-LSTM model (see Fig. S6) The Conv-LSTM model used in this research is composed of one 1D convolutional layer, two LSTM layers, and one dense hidden layer. In this model, the feature extraction of the input data is performed by a CNN layer; then, the output values are fed to the LSTM and Dense layers to predict the landslide displacement with the extracted feature data. The combination between CNN and LSTM achieved the highest prediction accuracies in several forecasting tasks (Xue et al. 2019; Lu et al. 2020; Livieris et al. 2020). However, this is the first research in which a Conv-LSTM architecture is used to forecast landslide displacement.

Training strategy and optimization

Present and past landslide displacement information and its triggers might influence future displacements through nonlinear relationships. In this research, we use seven DL models to forecast one single displacement time step in the future by showing the model n past time steps (look back) of both displacement and trigger variables. Several look-back windows (3, 5, 7, 9, 12) are evaluated for all the abovementioned models across all four case studies, to define the best one for each study case. A visual scheme of the approach is shown in Fig. 10.

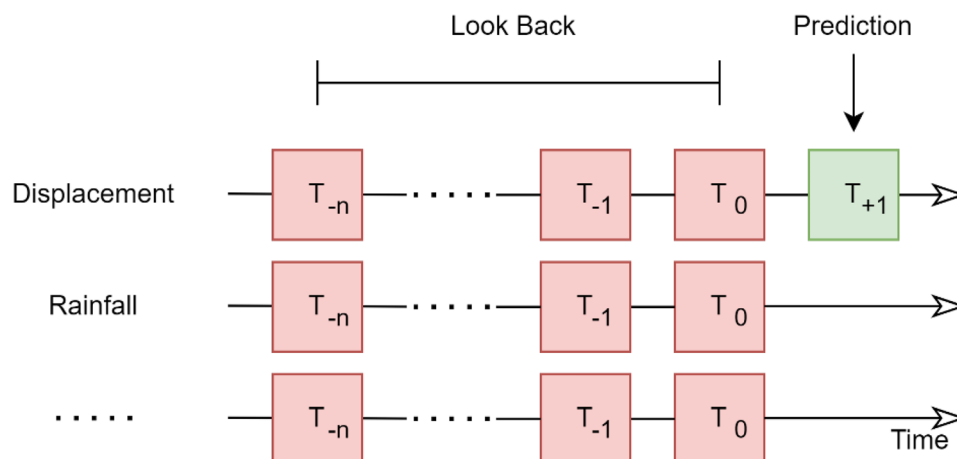


Fig. 10 The forecasting approach used. The models are trained to predict one single displacement time step in the future by looking at all the variables (differential displacement and triggers) considered for n time steps. Several look-back windows are evaluated

The displacement time series are divided into 80% and 20% for training and testing, respectively. Twenty percent of the training set is used to monitor the model performances during training, as well as to create checkpoints (save model weights) corresponding to the lower validation loss. The division is performed temporally, meaning that the training set is the oldest, followed by the validation set, and finally the test set. All sets are further divided into small chunks of time series, having the length of the look-back value plus the ground truth of the predicted step. Meaning that if the look back is set equal to 3, the chunks' length is four time steps. Huber loss was used as the loss function (Holland and Welsch 1977; Huang and Wu 2021). The Huber loss combines and optimizes both the mean square error (MSE) and mean average error (MAE), and it is defined as follows:

$$L_{\delta}(y, f(x)) = \begin{cases} \frac{1}{2}(y - f(x))^2 & \text{for } |y - f(x)| \leq \delta, \\ \delta|y - f(x)| - \frac{1}{2}\delta^2 & \text{otherwise.} \end{cases} \quad (1)$$

where $\delta = 1$.

Therefore, for loss values greater than δ , MAE is used, while for lower loss values MSE is used. We reduce the importance of outliers by using the MAE for bigger loss values, which nevertheless results in a well-rounded model. At the same time, we preserve a quadratic function close to the center by using the MSE for the smaller loss values. The number of training epochs is set to 1000, while, to reduce tuning time, an early stopping strategy is used to stop the training when the validation loss does not decrease after 20 subsequent epochs. For each architecture, weights are automatically saved in the epoch corresponding to the lowest validation loss value.

Furthermore, the appropriate combinations of hyperparameters must be used while training such deep learning models to optimize the model and deliver the best results. As a result, we iteratively train the model using a variety of combinations of batch sizes (9, 18, 36, 74, 144), learning rates (10e-3, 5e-3, 10e-4, 5e-4, 10e-5, 5e-5) and the number of layer nodes (8, 16, 32, 64, 128, 256). Therefore, for each architecture, we train 180 different hyperparameter combinations per each look back, for a total of 900 combinations per architecture.

The process is built in Python programming language. All experiments were executed on a Windows operating system computer with a 3rd Gen Ryzen Threadripper 3990X CPU and NVIDIA RTX 3090 GPU with 10,496 CUDA cores.

Model evaluation

After model training, the prediction models must be evaluated on an unseen test dataset. The prediction performance is estimated using root mean squared error (RMSE) (2), normalized RMSE (NRMSE) (3), and coefficient of determination (R^2) (17), which are defined as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (P_i - M_i)^2}{N}} \quad (2)$$

$$R^2 = \left[\frac{\sum_{i=1}^N (M_i - \bar{M}) \cdot (P_i - \bar{P})}{\sqrt{\sum_{i=1}^N (M_i - \bar{M})^2 \cdot \sum_{i=1}^N (P_i - \bar{P})^2}} \right]^2 \quad (3)$$

where P and M stand for the predicted and measured displacements, respectively. \bar{M} is the mean of observed values and \bar{P} refers to the average of the predicted displacements. The number of samples is represented by N .

The R^2 statistic evaluates the connection between measured and predicted values. The RMSE displays the difference between actual and predicted values. An effective model has a low value of RMSE and a high value of R^2 .

Landslide displacement forecasting

Landslides outside artificial reservoir contexts

As stated above, few studies in which ML approaches are applied for landslide displacement forecasting outside reservoir contexts exist. In this section, we show the forecasting results of two rainfall-triggered slow-moving landslides. In those cases, the main factor that influences the landslide displacement is the groundwater level, that, in turn, is further conditioned by rainfall. Since for both cases groundwater level time series are not available, just rainfall is used as a triggering factor.

Sant'Andrea landslide

In the Sant'Andrea landslide, the above-explained complex hydrogeological settings make it difficult to directly associate the daily rainfall with daily displacement. The landslide body behavior is heavily controlled both by the shallow and karst-deep groundwater circulation systems. Therefore, the same rainfall event might prolong landslide accelerations by evolving into two slightly temporally shifted groundwater flows. Since groundwater level time series data are not available, in this case, it is not possible to depict the exact time shift between rainfall and groundwater level raise. Therefore, we decide to feed to the model several cumulated rainfalls, from daily cumulate to 7 days cumulate. In this case, gray relational analysis (Kuo et al. 2008; Kayacan et al. 2010) is used to study the correlation between the derived cumulates and the landslide differential displacement. Here, gray relational coefficient is applied to investigate the correlation of the cumulates. Since the coefficient does not show consistent correlation changes through the seven cumulates, we decide to feed all the derived rainfalls as model inputs, along with landslide differential displacement. Moreover, an iterative evaluation of the rainfall derivatives shows that not increasing the number of cumulates, nor decreasing it was improving the performance of the models.

Therefore, historical differential displacement and eight different cumulates of rainfall (1 to 7 days) are used to predict the landslide displacement (mm) 24 h in the future. We train the models on 4 years of data from 09 February 2014 to 15 February 2018, for a total of 1468 daily time steps. One entire year is chosen as the test set, for a total of 365 time steps, from 16 February 2018 to 15 February 2019. The best metrics are yielded by the MLP, LSTM, and GRU models, while the worse are yielded by the 1D CNN model (see Table 1).

Figure 11 illustrates the predicted values for the top three models at the test set's highest peak. The peak and the start of increasing displacement can be accurately predicted by MLP, LSTM, and GRU. For the remaining models, this is not accurate. Although LSTM and GRU significantly underestimate the peak, MLP achieves the closest predictions to the peak. Figures 12 and 13 depict alternative scenarios where the accelerations are milder than those in the precedent

Table 1 Comparison of performance of the seven models in Sant'Andrea landslide. Values in bold emphases are the scores of the best three models

Model	Batch size	Learning rate	Nodes	lags	RMSE (mm)	R ²
MLP	36	0.005	16	3	0.706	0.5928
LSTM	144	0.005	64	12	0.704	0.5947
GRU	144	0.001	128	7	0.700	0.6000
2xLSTM	72	0.0005	32	3	0.830	0.4365
Bi-LSTM	144	0.0005	64	5	0.805	0.4700
Conv	72	0.01	32	3	0.953	0.2580
Conv-LSTM	144	0.0005	256	3	0.827	0.4415

The three best results are in bold

cases, both in terms of the acceleration gradient and magnitude. It is clear from these examples how both LSTM and GRU can predict all of the accelerations and peaks with accuracy. Instead, MLP is unable to predict the peak in this instance.

Lamosano landslide

Historical horizontal differential displacement and rainfall are used to predict the landslide displacement (mm) 11 days in the future. All the variables' time series used have 11 days as the dimension of the time steps. We train the models for almost 5 years from 10 April 2015 to 03 February 2020, for a total of 161 time steps. Almost 2 years are chosen as the test set, for a total of 61 time steps, from 14 April 2018 to 03 February 2020. Generally, in this study case, all models except Conv and Conv-LSTM

yield competitive results. The best metrics are yielded by the MLP, GRU, and BI-LSTM models (with 3, 7, and 9 as look-back dimensions, respectively), while, ones more, the worse is yielded by the Conv-LSTM and 1D CNN models (see Table 2). However, evaluation scores of LSTM and 2xLSTM are close to the ones of the best three models.

However, we can notice substantially different predictive behaviors across the best three models. For instance, in the first displacement peak (Fig. 14), the MLP model can perfectly predict both the displacement peak and onset of acceleration. On the other hand, GRU and BI-LSTM predict correctly just the onset of acceleration, while they anticipate/underestimate the displacement peak. However, in the second and third displacement peaks (Fig. 15), all three models heavily underestimate the displacement, while, once again, they correctly predict the onset of acceleration.

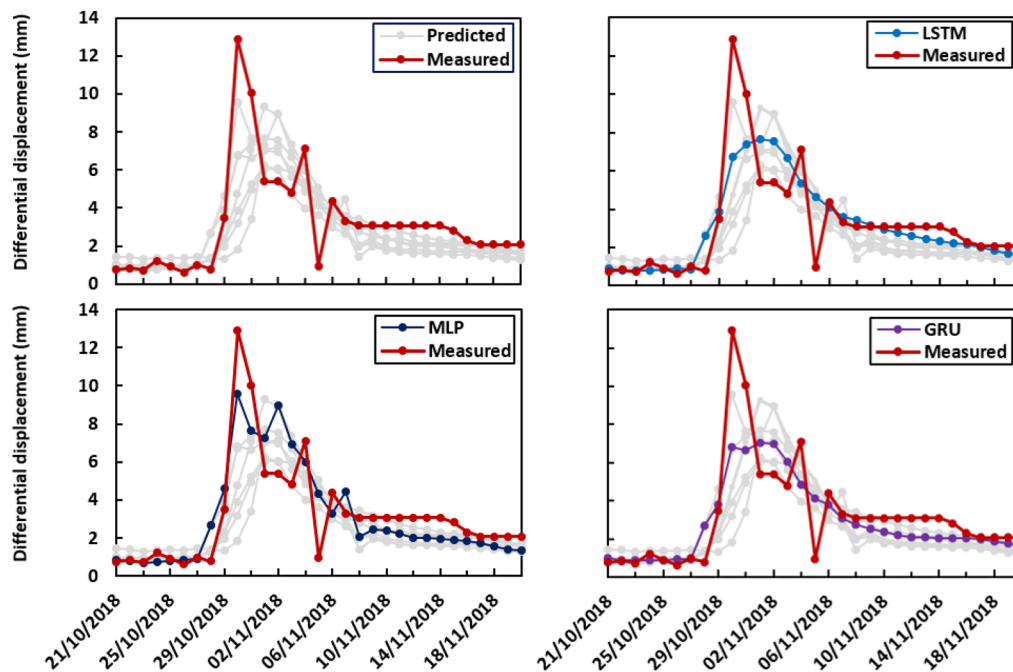


Fig. 11 Forecasting results of the highest displacement peak by the best three models (MLP, LSTM, GRU)

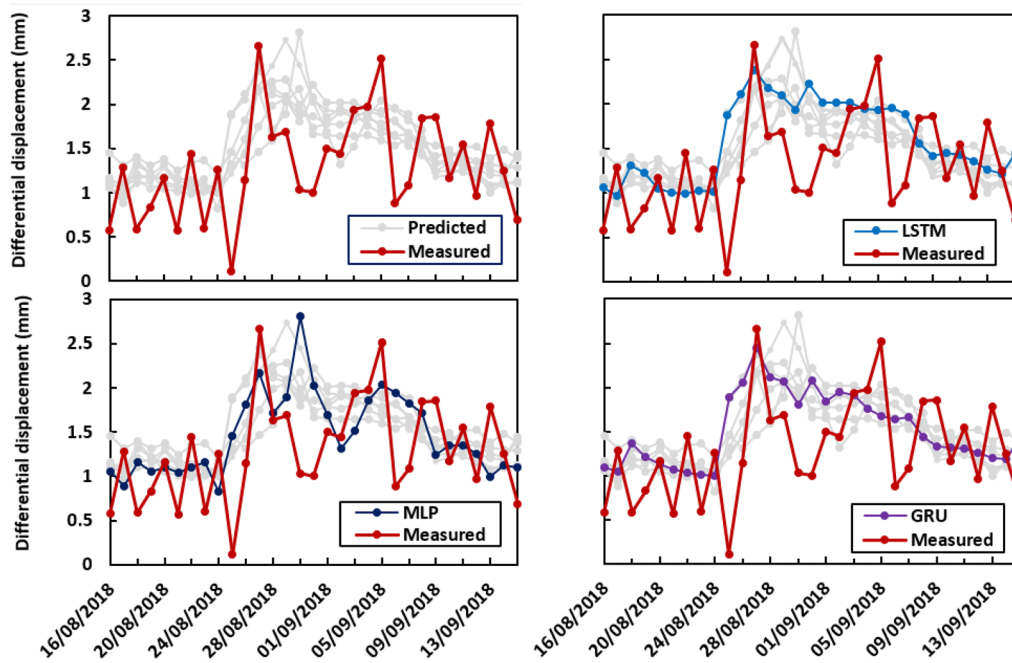


Fig. 12 Forecasting results of the first small displacement peak by the best three models (MLP, LSTM, GRU)

Landslides in artificial reservoirs

Numerous pieces of research show that reservoir landslide's displacement is mostly influenced by reservoir water level changes, rainfall, and its preceding displacement (Huang et al. 2017; Zhou et al. 2018b; Wang et al. 2019, 2020; Reyes-Carmona

et al. 2021). Therefore, rainfall, antecedent displacement, and differential reservoir water level are included in the multivariate modeling in both the reservoir landslides investigated in this research, namely Baishuihe (China) and El Arrecife (Spain). According to Wang et al. (2022), several variables derived from rainfall and reservoir water level have been used

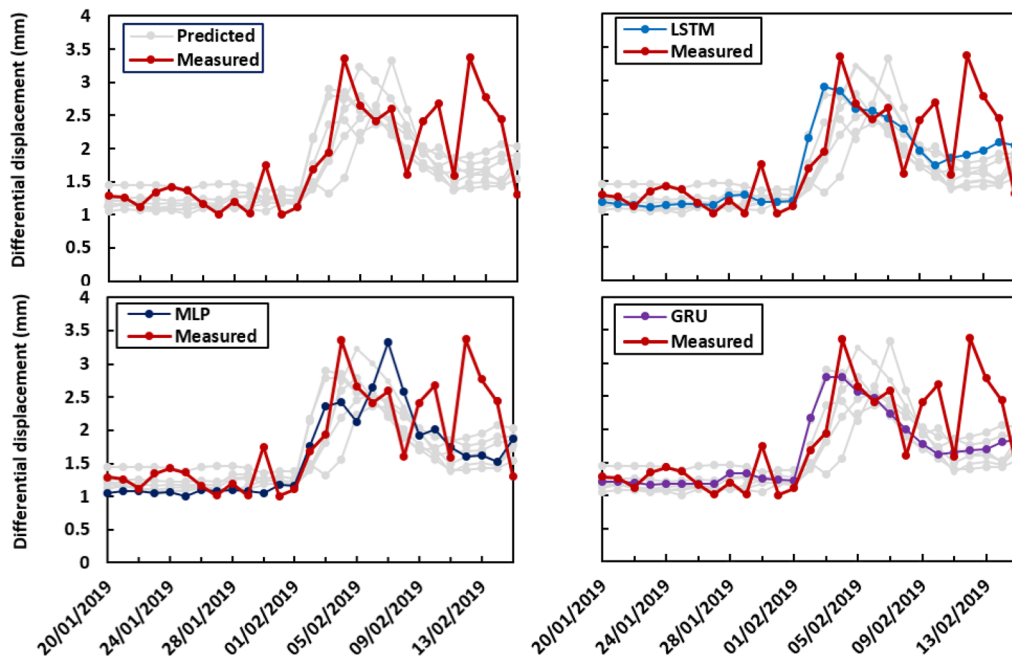


Fig. 13 Forecasting results of the second small displacement peak by the best three models (MLP, LSTM, GRU)

Table 2 Comparison of performance of the seven models in Lamosano landslide. Values in bold emphases are the scores of the best three models

Model	Batch size	Learning rate	Nodes	Lags	RMSE (mm)	R ²
MLP	144	0.01	32	3	2.080	0.2782
LSTM	9	0.0005	256	3	2.176	0.2097
GRU	72	0.005	16	7	2.142	0.2347
2xLSTM	144	0.005	256	12	2.152	0.2272
Bi-LSTM	144	0.005	256	9	2.070	0.2850
Conv	72	0.01	128	7	2.424	0.0198
Conv-LSTM	144	0.0001	64	3	2.257	0.1503

The three best results are in bold

by researchers as model inputs. However, most of them differ from the previous time steps that are taken into consideration while modeling. Moreover, the strategies adopted by the authors to select the candidate variables change from research to research, and therefore, different variables may be selected for the same case studies. We decide, therefore, to avoid a statistical-based selection criterion, but to select displacement, rainfall, and changes in reservoir water level as monthly (Baishuihe) or 2 weeks (El Arrecife) information, according to the highest displacement time step resolution available. The influence of previous months' triggers is then evaluated through the look-back windows. No moving cumulates are considered in this case, based on the assumption that the look-back window of up to 12 months can get all the information necessary for reliable forecasting.

Baishuihe landslide

Historical monthly differential displacement, monthly rainfall, and monthly changes in reservoir water level are used to predict the landslide displacement (mm) 1 month in the future. We train the models for 13 years from 31 August 2003 to 31 August 2016, for a total of 157 monthly time steps. In this case, two entire years are chosen as a test set, for a total of 24 time steps, from 30 September 2016 to 31 August 2018.

The Conv-LSTM model achieved the best forecasting performances with 12 as the look-back window (1 year), followed by MLP and LSTM with 9 and 12 as the look-back window sizes, respectively. In this case, it is evident how Conv-LSTM outperforms all the other models by far, yielding 8.6 mm of RMSE and 0.85 of R², while the second-best model, the MLP, yielded an RMSE of 13.55 and an R² of 0.65. In this case, the best look-back

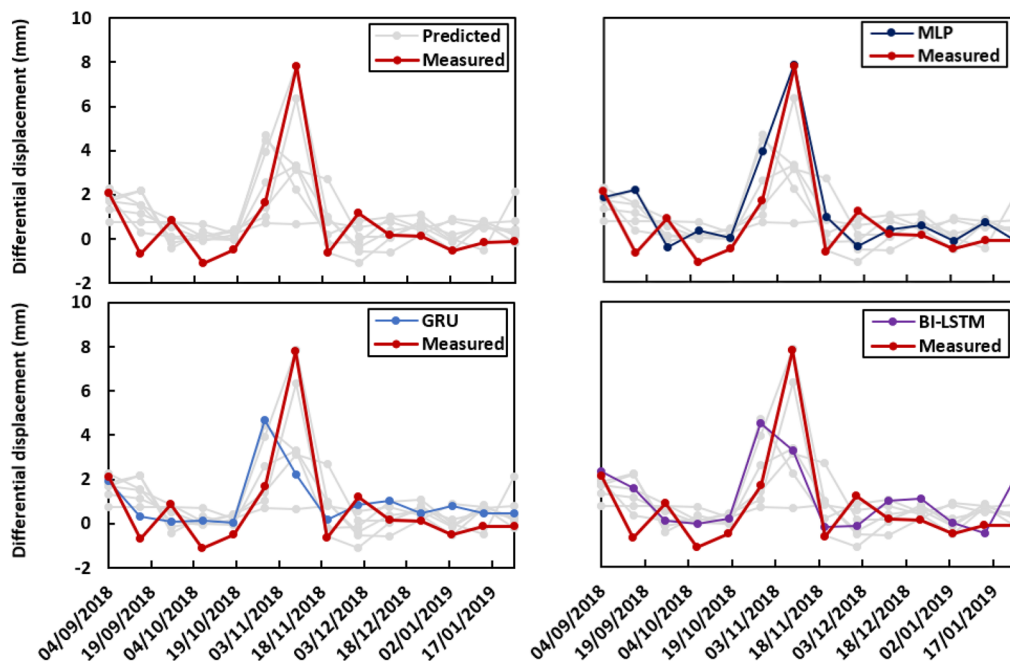


Fig. 14 Forecasting results of the first displacement peak by the best three models (MLP, GRU, BI-LSTM)

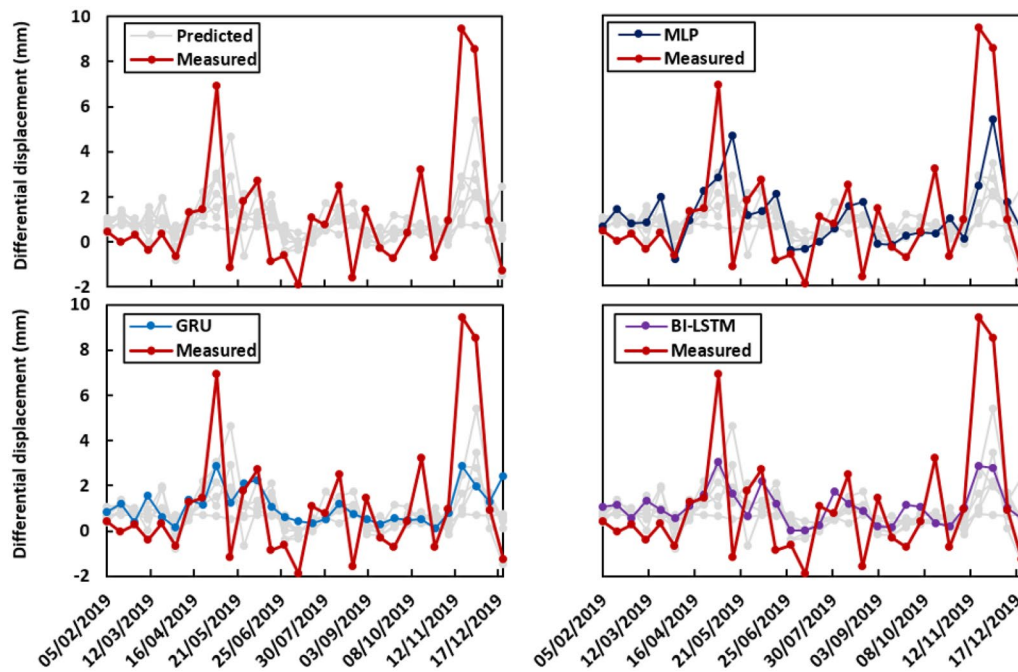


Fig. 15 Forecasting results of the second and third displacement peaks by the best three models (MLP, GRU, BI-LSTM)

window remains quite stable across all the architectures, ranging from 9 to 12 time steps (Table 3).

Figure 16 shows the predictions of the best three models for the unseen 2 years of the test set. This set shows two peaks, at a distance of 11 months. The first peak in July 2017 shows 91 mm of monthly differential displacement, while the second one, in June 2018, shows a lower displacement, with a 30-mm peak. By looking at the model predictions, it is evident how all the models can precisely predict the onset of acceleration in time, except for the Conv model. However, GRU and bi-LSTM anticipate and “flatten” the onset of acceleration of both peaks. The general tendency of all models is to underestimate the first peak while overestimating the second one. However, Conv-LSTM shows in this case outstanding predictive capabilities, underestimating the first by 19 mm while overestimating the second just by 3.5 mm, while the second best (MLP) instead underestimates the first by 29 mm and overestimates the second by 5 mm.

El Arrecife landslide

Historical differential displacement, rainfall, and changes in reservoir water level with 12-day time step dimension are used to predict the landslide displacement (mm) 12 days in the future. We train the models for 3 years from 05 November 2016 to 18 May 2019, for a total of 72 time steps. In this case, around 1 year is chosen as the test set, for a total of 24 time steps, from 30 May 2019 to 13 March 2020.

In this case, the landslide behavior is quite different from the previous three cases shown in the paper. For instance, the accelerations (and decelerations) are gentler, and the overall velocity of the landslide is lower. Moreover, the influence of the rainfall is unclear and not simple to quantify since in correspondence with the most intense rainfalls the reservoir water level rises, with a consequent deceleration of the landslide movement. In this study case, all models achieve acceptable RMSE values, with GRU, MLP, and LSTM yielding the best scores (Table 4).

Table 3 Comparison of performance of the seven models in Baishuihe landslide

Model	Batch size	Learning rate	Nodes	Lags	RMSE (mm)	R^2
MLP	36	0.005	128	9	13.555	0.6562
LSTM	18	0.005	96	12	14.478	0.6079
GRU	36	5.00E-05	16	9	17.909	0.4000
2xLSTM	72	0.001	96	9	14.879	0.5858
Bi-LSTM	9	0.005	128	9	15.806	0.5326
Conv	18	5.00E-05	16	9	19.811	0.2658
Conv-LSTM	36	0.005	96	12	8.672	0.8593

The three best results are in bold

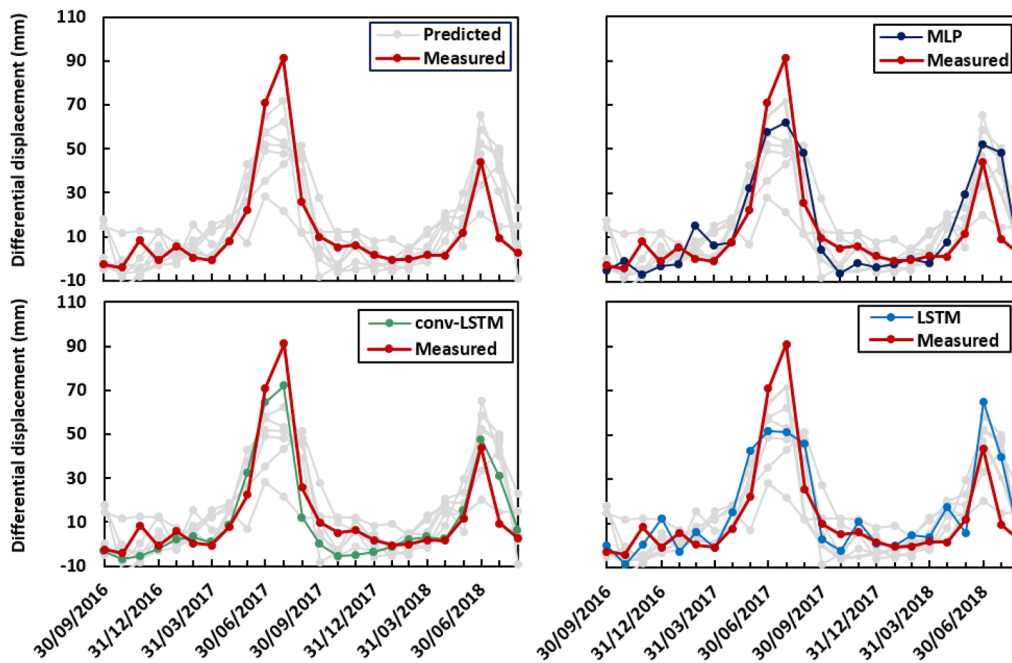


Fig. 16 Forecasting results by the best three DL models (MLP, LSTM, Conv-LSTM)

Figure 17 shows the prediction of all models for the unseen test set. It is visible how all predictions are very similar, and they all fit the measured displacement. A gentle acceleration is predicted around 31 October 2019, when the water level is close to the local minimum. All the rest of the test series presents a linear behavior, which makes the forecasting less challenging than in the other cases.

Discussion

The results show that no model yields the best results across all four study cases. However, generally, MLP, LSTM, and GRU achieve competitive results in all four landslides, proving to be able to perform well in different scenarios, yielding the most consistent cross-site predictions. Moreover, the models do not show any behavior dependent on the type of landslide or triggering factor of the same. No evident performance differences were found when modeling landslides inside artificial reservoirs rather than outside. The Conv

model yielded in all cases the worst predictions. Therefore, the use of just convolutional layers for landslide displacement forecasting is found ineffective in our investigation. The BI-LSTM model yielded competitive predictions just in the Lamosano landslide, while the 2xLSTM model in these four study cases did not yield competitive results. The Conv-LSTM model achieves the best performance just in Baishuihe, while MLP, LSTM, and GRU yield the best results in the other three cases. Although the Conv-LSTM model can outperform the other models in Baishuihe, its predictions in the other cases are quite imprecise. The explanation for this behavior could be found in the characteristics of the kinematic of the Baishuihe landslide. In fact, inside the investigation time lag, the landslide displacement time series shows 15 major accelerations which occur periodically with a yearly frequency (either in June or July). This is due to both the strong seasonality of the rainfall cycles as well as to the reservoir water levels. In such cases, we expect that the same,

Table 4 Comparison of performance of the seven models in El Arrecife landslide

Model	Batch size	Learning rate	Nodes	Lags	RMSE (mm)	R^2
MLP	144	0.005	16	12	2.541	0.2418
LSTM	36	0.005	256	9	2.606	0.2023
GRU	18	0.01	128	9	2.518	0.2556
2xLSTM	18	0.01	256	5	2.733	0.1230
Bi-LSTM	9	0.01	32	7	2.658	0.1702
Conv	72	0.0005	16	3	2.775	0.0651
Conv-LSTM	18	0.001	128	9	2.617	0.1956

The three best results are in bold

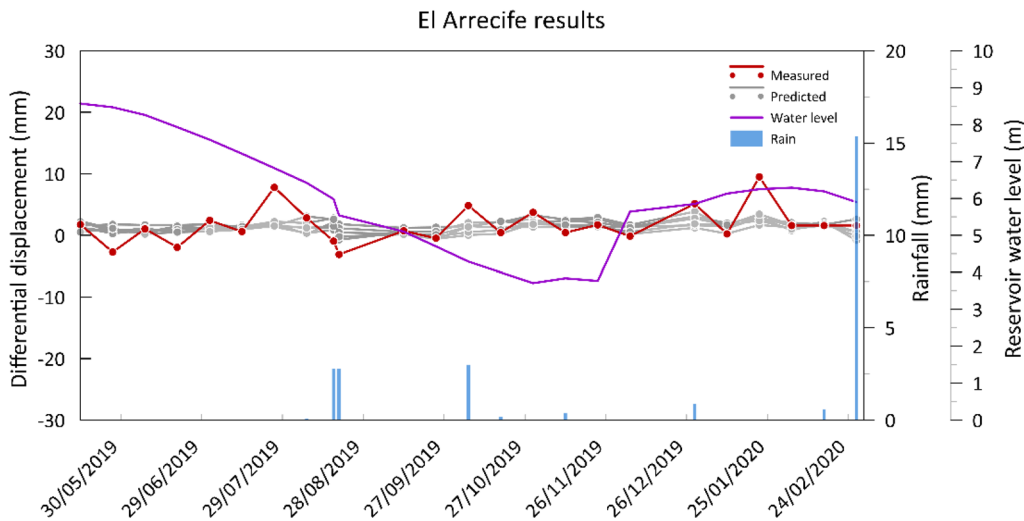


Fig. 17 Forecasting results by all models and measured differential displacement. Reservoir water level and rainfall are shown along with the predictions

or a slightly modified version of the proposed Conv-LSTM model, with a look-back and a kernel size of 1 year can reliably capture and model the periodicity of the displacement. Therefore, the combination of CNN and LSTM layers represents a newly introduced suitable approach for displacement forecasting in landslides that present a strong and constant seasonality.

Reliability of DL-based landslide displacement forecasting systems

The validation and deployment of forecasting models for landslide displacement forecasting tasks have been approached so far as a common time series forecasting problem. However, to successfully calibrate and validate, a reliable landslide displacement forecasting model is not sufficient to check the quality of the predictions through the usual common metrics. When forecasting the landslide displacement to develop a reliable EWS or to support existing ones, it is fundamental that a model is capable of predicting both the onset of acceleration and maximum displacement peaks reliably, while maintaining accurate predictions in all the “stable” displacement states. As seen in the results (especially in Sant’Andrea), similar values of RMSE and R^2 do not always describe the same predictive behavior. R^2 is used in this research just to compare models in the same study case (on the same set of data) rather than compare predictions across cases since R^2 is strictly dependent upon the set of data used. Moreover, in the existing literature, most of the test sets set by scientists are usually composed of a single acceleration. However, as proved by results in the Sant’Andrea and Lamosano study cases, this validation strategy might have serious shortcomings since the same model might be suitable for stronger accelerations rather than minor ones, and vice versa. Although the metrics of MLP, LSTM, and GRU are similar, the predictive behavior of the networks is quite different. For instance, by taking into consideration the predictions related to the highest displacement peak in Sant’Andrea (Fig. 11), although all models generally tend to underestimate the displacement, the closest prediction to the highest peak (12.88 mm) is yielded by

the MLP, with a difference of 3.31 mm, while GRU underestimates it by 6.08 mm and LSTM by 6.16 mm. By considering instead the minor accelerations in Figs. 12 and 13, LSTM and GRU yielded the closest peaks prediction, while in both cases, MLP strongly underestimates the first peak, while it predicts a second one where no real peaks are measured. For the Lamosano landslide, we chose three displacement peaks as the test set. In this case, several models yield reliable predictions of the first one, while they severely underpredict the second and third ones.

The standard deviation (STD) of the time series is calculated on stable measurement points/stations available for each study case to obtain the error margin of the measurements (see Table S1). If we compare the same with the RMSE yielded by our models, we can notice that in Sant’Andrea and El Arrecife landslides the prediction error is lower than the measurement uncertainty. This means that our models, in those cases, are quite reliable. Nevertheless, in the Lamosano study case, the error yielded by the model results higher than the error margin, confirming the poor predictions of the models.

Another interesting aspect is the relation between the time step dimension of the displacement time series and the modeled displacement. To develop a reliable forecasting model, we believe that it is necessary to have in-depth physical knowledge of the landslide we wish to model. The Baishuihe landslide, for instance, is modeled by using monthly time series. In this case, the assumption is that rainfall, water level, and displacement measured in the previous months contribute and are the only cause of the displacement in the following month. However, this assumption might not be true for all landslides. For instance, in Sant’Andrea few previous look-back days are sufficient to have a well-calibrated model when forecasting 1 day in the future, while in Baishuihe the best result is obtained with a look back of 1 year.

Contribution of DL-based landslide displacement forecasting models to EWS

For landslide early warning, the model’s predictive ability is extremely important (Sassa 2009). However, several of the existing

landslide EWS are based on the historical interaction between a certain landslide and its influencing factors. The most popular early warning method for slow-moving landslides is based on critical rainfall thresholds (Crosta et al. 2017; Xu et al. 2020). This technique is utilized extensively for sending emergency warnings because of its simplicity in application and interpretation as well as its dependability (Du et al. 2013; Segoni et al. 2014). However, although this method can deliver reliable early warnings, the magnitude of the expected landslide displacement and the timing of the maximum displacement are often missing or poorly estimated. The methodology investigated in this paper instead has the potential to deliver in advance both the onset of acceleration timing and the precise timing of the displacement peak. However, while the rainfall thresholds method considers just forecasted rainfall as the only triggering variable, the investigated approach does not consider future covariates, and it works on the assumption that just present and past triggers influence future landslide behavior. Therefore, to get the most from such methodologies, a combination of both might be suggested, where the rainfall threshold approach gives alerts based on future rainfall forecasting data, and the DL-based forecasting model predicts the magnitude of the acceleration based on present and past measurements.

Conclusion

This study tested the efficacy of seven deep learning techniques for predicting landslide displacement using four different landslides that varied in terms of geographic location, influencing factors, geological settings, time step dimensions, and measurement sensors and provides insights on their performance.

The study reveals that the study case had a small impact on how well the seven techniques performed. In fact, the models do not show any behavior dependent on the type of landslide or triggering factor of the same. The Conv model yielded instead in all cases the worst predictions. Three models, MLP, LSTM, and GRU, demonstrated the ability to produce reliable predictions in each of the four scenarios. Moreover, in the Baishuihe study scenario, where the landslide had a high seasonality, the suggested Conv-LSTM model outperformed the other models. In contrast, the GRU model performs well in Sant'Andrea and El Arrecife but is unable to accurately predict the displacement of the Baishuihe landslide. Instead, the dimensions of the look-back window are closely tied to the particular modeled event. MLP, GRU, and LSTM are thus advised while tackling a landslide displacement forecasting task. When the displacement exhibits a high seasonality, the combination of 1D CNN and LSTM layers (Conv-LSTM) must be considered. Several look-back windows must always be considered in all scenarios. Finally, we advise against using 1D CNNs and bidirectional LSTMs. Even if the results of this work showed the reliability of using DL algorithms for landslide displacement forecasting, some improvements are still needed before their use in operational EWS.

For instance, the support of future influencing factors (weather forecasting) might improve the prediction accuracy of the models, especially in cases in which the resolution of the time series is low (monthly). Nevertheless, using weather forecasting predictions as a covariate might introduce further uncertainties to the model since forecasted rainfall has per se some degree of uncertainty, that must

be properly evaluated. Other parameters as the seasonal variations of deformations in response to the triggering factors, a combination of different meteorological data (e.g., rainfall and temperature), or the lithological pattern of the landmass could be evaluated. Moreover, ensemble modeling could combine the strengths of different prediction algorithms, allowing the models to output more accurate predictions. In the Sant'Andrea study case, it would be interesting to ensemble MLP (suitable for high displacement peaks) and LSTM (suitable for small displacement peaks) models, to build a model able to predict both stronger and more gentle accelerations. In summary, this study has shown that deep learning (DL) can be successfully applied to landslide early warning systems (EWS). However, it has also emphasized that there is no one-size-fits-all model or configuration that is universally optimal. Instead, a site-specific calibration approach should be employed for the best results.

Funding

Open access funding provided by Università degli Studi di Padova within the CRUI-CARE Agreement. The authors wish to thank the Veneto Region and Veneto Strade for providing monitoring data and partially funding the research with the Grant "SUPPORTO SCIENTIFICO PER L'OTTIMIZZAZIONE, IMPLEMENTAZIONE E GESTIONE DEL SISTEMA DI MONITORAGGIO CON AGGIORNAMENTO DELLE SOGLIE DI ALLERTAMENTO DEL FENOMENO FRANOSO DI SANT'ANDREA – PERAROLO DI CADORE (BL)" and the Spanish Grant "SARAI, PID2020-116540RB-C21," funded by "MCIN/AEI/10.13039/501100011033" and "RISKCOAST" for having provided the InSAR displacement data of the El Arrecife landslide. Access to the Geohazard Exploitation Platform (GEP) of the European Space Agency (ESA) was provided by the NoR Projects Sponsorship (Project ID: 63737).

Data availability

Associated code required for fine-tuning the seven deep learning models used in our research. The link to the GitHub repository: <https://github.com/lorenzonavaga96/Landslide-Displacement-Forecasting-using-seven-Deep-Learning-architectures-and-monitoring-data>.

Declarations

Competing interests The authors declare no competing interests.

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

- Aleotti P, Chowdhury R (1999) Landslide hazard assessment: summary review and new perspectives. *Bull Eng Geol Env* 58:21–44. <https://doi.org/10.1007/s100640050066>
- Amarasinghe K, Marino DL, Manic M (2017) Deep neural networks for energy load forecasting. In: 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE). IEEE pp 1483–1488
- Berardino P, Fornaro G, Lanari R, Sansosti E (2002) A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. *IEEE Trans Geosci Remote Sens* 40:2375–2383. <https://doi.org/10.1109/TGRS.2002.803792>
- Brezzi L, Carraro E, Pasa D et al (2021) Post-collapse evolution of a rapid landslide from sequential analysis with FE and SPH-based models. *Geosciences (basel)* 11:364. <https://doi.org/10.3390/geosciences11090364>
- Cao Y, Yin K, Alexander DE, Zhou C (2016) Using an extreme learning machine to predict the displacement of step-like landslides in relation to controlling factors. *Landslides* 13:725–736. <https://doi.org/10.1007/s10346-015-0596-z>
- Casu F, Elefante S, Imperatore P et al (2014) SBAS-DInSAR parallel processing for deformation time-series computation. *IEEE J Sel Top Appl Earth Obs Remote Sens* 7:3285–3296. <https://doi.org/10.1109/JSTARS.2014.2322671>
- Chen S-Y, Chou W-Y (2012) Short-term traffic flow prediction using EMD-based recurrent Hermite neural network approach. In: 2012 15th International IEEE Conference on Intelligent Transportation Systems. IEEE, pp 1821–1826
- Cho K, van Merriënboer B, Gulcehre C et al (2014) Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation
- Corominas J, van Westen C, Frattini P et al (2013) Recommendations for the quantitative analysis of landslide risk. *Bull Eng Geol Env*. <https://doi.org/10.1007/s10064-013-0538-8>
- Crosta GB, Agliardi F, Rivolta C et al (2017) Long-term evolution and early warning strategies for complex rockslides by real-time monitoring. *Landslides* 14:1615–1632. <https://doi.org/10.1007/s10346-017-0817-8>
- De Luca C, Cuccu R, Elefante S et al (2015) An on-demand web tool for the unsupervised retrieval of earth's surface deformation from SAR data: the P-SBAS service within the ESA G-POD environment. *Remote Sens (basel)* 7:15630–15650. <https://doi.org/10.3390/rs71115630>
- Du J, Yin K, Lacasse S (2013) Displacement prediction in colluvial landslides, Three Gorges Reservoir, China. *Landslides* 10:203–218. <https://doi.org/10.1007/s10346-012-0326-8>
- Froude MJ, Petley DN (2018) Global fatal landslide occurrence from 2004 to 2016. *Nat Hazard* 18:2161–2181. <https://doi.org/10.5194/nhess-18-2161-2018>
- Gamboa JCB (2017) Deep learning for time-series analysis
- Guo L, He K, Zhou Y et al (2022) Physical prediction model and application of compound hydrodynamic displacement response ratio with dynamics increment in reservoir landslide. *J Eng Geol* 30(5):30:1561–1572
- Han H, Shi B, Zhang L (2021) Prediction of landslide sharp increase displacement by SVM with considering hysteresis of groundwater change. *Eng Geol* 280:105876. <https://doi.org/10.1016/j.enggeo.2020.105876>
- Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9:1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Holland PW, Welsch RE (1977) Robust regression using iteratively reweighted least-squares. *Commun Stat Theory Methods* 6:813–827. <https://doi.org/10.1080/03610927708827533>
- Huang F, Huang J, Jiang S, Zhou C (2017) Landslide displacement prediction based on multivariate chaotic model and extreme learning machine. *Eng Geol* 218:173–186. <https://doi.org/10.1016/j.enggeo.2017.01.016>
- Huang S, Wu Q (2021) Robust pairwise learning with Huber loss. *J Complex* 66:101570. <https://doi.org/10.1016/j.jco.2021.101570>
- Kayacan E, Ulutas B, Kaynak O (2010) Grey system theory-based models in time series prediction. *Expert Syst Appl* 37:1784–1789. <https://doi.org/10.1016/j.eswa.2009.07.064>
- Keqiang H, Xiangran L, Xueqing Y, Dong G (2008) The landslides in the Three Gorges Reservoir Region, China and the effects of water storage and rain on their stability. *Environ Geol* 55:55–63. <https://doi.org/10.1007/s00254-007-0964-7>
- Khashei M, Hajirahimi Z (2019) A comparative study of series arima/mlp hybrid models for stock price forecasting. *Commun Stat Simul Comput* 48:2625–2640. <https://doi.org/10.1080/03610918.2018.1458138>
- Kiranyaz S, Avci O, Abdeljaber O et al (2021) 1D convolutional neural networks and applications: a survey. *Mech Syst Signal Process* 151:107398. <https://doi.org/10.1016/j.ymssp.2020.107398>
- Krkač M, Špoljarić D, Bernat S, Arbanas SM (2017) Method for prediction of landslide movements based on random forests. *Landslides* 14:947–960. <https://doi.org/10.1007/s10346-016-0761-z>
- Kulshrestha A, Krishnaswamy V, Sharma M (2020) Bayesian BiLSTM approach for tourism demand forecasting. *Ann Tour Res* 83:102925. <https://doi.org/10.1016/j.annals.2020.102925>
- Kuo Y, Yang T, Huang G-W (2008) The use of grey relational analysis in solving multiple attribute decision-making problems. *Comput Ind Eng* 55:80–93. <https://doi.org/10.1016/j.cie.2007.12.002>
- Lian C, Zeng Z, Yao W, Tang H (2015) Multiple neural networks switched prediction for landslide displacement. *Eng Geol* 186:91–99. <https://doi.org/10.1016/j.enggeo.2014.11.014>
- Li D, Yin K, Chen L, Chen W (2008) Application of GPS monitoring data in landslide prediction. In: Li D, Gong J, Wu H (eds). p 728545
- Li D, Yin K, Leo C (2010) Analysis of Baishuihe landslide influenced by the effects of reservoir water and rainfall. *Environ Earth Sci* 60:677–687. <https://doi.org/10.1007/s12665-009-0206-2>
- Li H, Xu Q, He Y et al (2020) Modeling and predicting reservoir landslide displacement with deep belief network and EWMA control charts: a case study in Three Gorges Reservoir. *Landslides* 17:693–707. <https://doi.org/10.1007/s10346-019-01312-6>
- Lin Z, Sun X, Ji Y (2022) Landslide displacement prediction based on time series analysis and double-BiLSTM model. *Int J Environ Res Public Health* 19:2077. <https://doi.org/10.3390/ijerph19042077>
- Liu Z, Shao J, Xu W et al (2014) Comparison on landslide nonlinear displacement analysis and prediction with computational intelligence approaches. *Landslides* 11:889–896. <https://doi.org/10.1007/s10346-013-0443-z>
- Livieris IE, Pintelas E, Pintelas P (2020) A CNN-LSTM model for gold price time-series forecasting. *Neural Comput Appl* 32:17351–17360. <https://doi.org/10.1007/s00521-020-04867-x>
- Li Y, Sun R, Yin K et al (2019) Forecasting of landslide displacements using a chaos theory based wavelet analysis-Volterra filter model. *Sci Rep* 9:19853. <https://doi.org/10.1038/s41598-019-56405-y>
- Lu W, Li J, Li Y et al (2020) A CNN-LSTM-based model to forecast stock prices. *Complexity* 2020:1–10. <https://doi.org/10.1155/2020/6622927>
- Ma J, Niu X, Tang H et al (2020) Displacement prediction of a complex landslide in the Three Gorges Reservoir Area (China) using a hybrid computational intelligence approach. *Complexity* 2020:1–15. <https://doi.org/10.1155/2020/2624547>
- Miao F, Wu Y, Li L et al (2021) Triggering factors and threshold analysis of Baishuihe landslide based on the data mining methods. *Nat Hazards* 105:2677–2696. <https://doi.org/10.1007/s11069-020-04419-5>
- Miao F, Wu Y, Xie Y et al (2018) Prediction of landslide displacement with step-like behavior based on multialgorithm optimization and a support vector regression model. *Landslides* 15:475–488. <https://doi.org/10.1007/s10346-017-0883-y>
- Montesino Pouzols F, Lendasse A (2010) Effect of different detrending approaches on computational intelligence models of time series. In: The 2010 International Joint Conference on Neural Networks (IJCNN). IEEE, pp 1–8
- Peng T, Zhang C, Zhou J, Nazir MS (2021) An integrated framework of bi-directional long-short term memory (BiLSTM) based on sine cosine algorithm for hourly solar radiation forecasting. *Energy* 221:119887. <https://doi.org/10.1016/j.energy.2021.119887>
- Pieraccini M, Noferini L, Mecatti D et al (2006) Integration of radar interferometry and laser scanning for remote monitoring of an urban site built on a sliding slope. *IEEE Trans Geosci Remote Sens* 44:2335–2342. <https://doi.org/10.1109/TGRS.2006.873574>
- Ravanelli M, Brakel P, Omologo M, Bengio Y (2018) Light gated recurrent units for speech recognition. *IEEE Trans Emerg Top Comput Intell* 2:92–102. <https://doi.org/10.1109/TETCI.2017.2762739>

- Reyes-Carmona C, Barra A, Galve J et al (2020) Sentinel-1 DInSAR for monitoring active landslides in critical infrastructures: the case of the rules reservoir (Southern Spain). *Remote Sens (basel)* 12:809. <https://doi.org/10.3390/rs12050809>
- Reyes-Carmona C, Galve JP, Moreno-Sánchez M et al (2021) Rapid characterisation of the extremely large landslide threatening the Rules Reservoir (Southern Spain). *Landslides* 18:3781–3798. <https://doi.org/10.1007/s10346-021-01728-z>
- Sak H, Senior A, Beaufays F (2014) Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition
- Sassa K (2009) Report of the 2008 First World Landslide Forum on 18–21 November 2008 at UNU, Tokyo. *Landslides* 6:167–179. <https://doi.org/10.1007/s10346-009-0161-8>
- Segoni S, Rossi G, Rosi A, Catani F (2014) Landslides triggered by rainfall: a semi-automated procedure to define consistent intensity–duration thresholds. *Comput Geosci* 63:123–131. <https://doi.org/10.1016/j.cageo.2013.10.009>
- Shan L, Liu Y, Tang M et al (2021) CNN-BiLSTM hybrid neural networks with attention mechanism for well log prediction. *J Pet Sci Eng* 205:108838. <https://doi.org/10.1016/j.petrol.2021.108838>
- Song J, Ma L, Her Y, Li Y (2020) Immediate influences of a large dam construction on local storm event patterns and weather variables: a case study of the Three Gorges Project. *Weather* 75:99–103. <https://doi.org/10.1002/wea.3410>
- Teza G, Atzeni C, Balzani M et al (2008) Ground-based monitoring of high-risk landslides through joint use of laser scanner and interferometric radar. *Int J Remote Sens* 29:4735–4756. <https://doi.org/10.1080/01431160801942227>
- Wang Y, Tang H, Huang J et al (2022) A comparative study of different machine learning methods for reservoir landslide displacement prediction. *Eng Geol* 298:106544. <https://doi.org/10.1016/j.enggeo.2022.106544>
- Wang Y, Tang H, Wen T, Ma J (2020) Direct interval prediction of landslide displacements using least squares support vector machines. *Complexity* 2020:1–15. <https://doi.org/10.1155/2020/7082594>
- Wang Y, Tang H, Wen T, Ma J (2019) A hybrid intelligent approach for constructing landslide displacement prediction intervals. *Appl Soft Comput* 81:105506. <https://doi.org/10.1016/j.asoc.2019.105506>
- Wen T, Tang H, Wang Y et al (2017) Landslide displacement prediction using the GA-LSSVM model and time series analysis: a case study of Three Gorges Reservoir, China. *Nat Hazard* 17:2181–2198. <https://doi.org/10.5194/nhess-17-2181-2017>
- Xing Y, Yue J, Chen C (2020) Interval estimation of landslide displacement prediction based on time series decomposition and long short-term memory network. *IEEE Access* 8:3187–3196. <https://doi.org/10.1109/ACCESS.2019.2961295>
- Xue N, Triguero I, Figueredo GP, Landa-Silva D (2019) Evolving deep CNN-LSTMs for inventory time series prediction. In: 2019 IEEE Congress on Evolutionary Computation (CEC). IEEE pp 1517–1524
- Xu Y, Lu Z, Schulz WH, Kim J (2020) Twelve-year dynamics and rainfall thresholds for alternating creep and rapid movement of the Hooskanaden landslide from integrating InSAR, pixel offset tracking, and borehole and hydrological measurements. *J Geophys Res Earth Surf* 125. <https://doi.org/10.1029/2020JF005640>
- Yang B, Yin K, Lacasse S, Liu Z (2019) Time series analysis and long short-term memory neural network to predict landslide displacement. *Landslides* 16:677–694. <https://doi.org/10.1007/s10346-018-01127-x>
- Yao W, Zeng Z, Lian C, Tang H (2015) Training enhanced reservoir computing predictor for landslide displacement. *Eng Geol* 188:101–109. <https://doi.org/10.1016/j.enggeo.2014.11.008>
- Zhang L, Shi B, Zhu H et al (2021) PSO-SVM-based deep displacement prediction of Majiagou landslide considering the deformation hysteresis effect. *Landslides* 18:179–193. <https://doi.org/10.1007/s10346-020-01426-2>
- Zhao R, Wang D, Yan R et al (2018) Machine health monitoring using local feature-based gated recurrent unit networks. *IEEE Trans Industr Electron* 65:1539–1548. <https://doi.org/10.1109/TIE.2017.2733438>
- Zhou C, Yin K, Cao Y et al (2018a) A novel method for landslide displacement prediction by integrating advanced computational intelligence algorithms. *Sci Rep* 8:7287. <https://doi.org/10.1038/s41598-018-25567-6>
- Zhou C, Yin K, Cao Y et al (2018b) Displacement prediction of step-like landslide by applying a novel kernel extreme learning machine method. *Landslides* 15:2211–2225. <https://doi.org/10.1007/s10346-018-1022-0>
- Zhou C, Yin K, Cao Y, Ahmed B (2016) Application of time series analysis and PSO-SVM model in predicting the Bazimen landslide in the Three Gorges Reservoir, China. *Eng Geol* 204:108–120. <https://doi.org/10.1016/j.enggeo.2016.02.009>
- Zhu CH, Hu GD (2012) Time series prediction of landslide displacement using SVM model: application to Baishuihe landslide in Three Gorges Reservoir Area, China. *Appl Mech Mater* 239–240:1413–1420. <https://doi.org/10.4028/www.scientific.net/AMM.239-240.1413>
- Zhu X, Xu Q, Tang M et al (2017) Comparison of two optimized machine learning models for predicting displacement of rainfall-induced landslide: a case study in Sichuan Province, China. *Eng Geol* 218:213–222. <https://doi.org/10.1016/j.enggeo.2017.01.022>

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10346-023-02104-9>.

Lorenzo Nava (✉) · **Silvia Puliero** · **Kushanav Bhuyan** · **Ascanio Rosi** · **Mario Floris** · **Sansar Raj Meena** · **Filippo Catani**
Machine Intelligence and Slope Stability Laboratory, Department of Geosciences, University of Padova, Padua, Italy

Lorenzo Nava
Email: lorenzo.nava@phd.unipd.it

Silvia Puliero
Email: silvia.puliero@phd.unipd.it

Kushanav Bhuyan
Email: kushanav.bhuyan@phd.unipd.it

Ascanio Rosi
Email: ascanio.rosi@unipd.it

Mario Floris
Email: mario.floris@unipd.it

Sansar Raj Meena
Email: sansarraj.meena@unipd.it

Filippo Catani
Email: filippo.catani@unipd.it

Edoardo Carraro
Department of Geography and Regional Research, Geomorphological Systems and Risk Research, University of Vienna, Universitätsstraße 7, 1010 Vienna, Austria
Email: edoardo.carraro@univie.ac.at

Cristina Reyes-Carmona · **Jorge Pedro Galve**
Departamento de Geodinámica, Universidad de Granada, Avda. del Hospicio, s/n, 18010 Granada, Spain
Email: cristinarc@ugr.es

Jorge Pedro Galve
Email: jpgalve@ugr.es

Oriol Monserrat
Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Geomatics Research Unit, Barcelona, Spain
Email: omonserrat@cttc.cat