EGU General Assembly 2024

Session NH3.6
Space and time forecasting of landslides
EGU24-6201

Sharing is encouraged

DEEP LEARNING FORECAST OF RAINFALL-INDUCED SHALLOW LANDSLIDES IN ITALY

Fausto Guzzetti, Alessandro C. Mondini, Massimo Melillo
Consiglio Nazionale delle Ricerche

Consiglio Nazionale delle Ricerche CNR irpi imati

ASSUMPTIONS

- In a landscape, landslides occur (or do not occur) depending on rainfall and local terrain conditions.
- P(F | R, S)
  - L, probability of landslide occurrence; R, rainfall; S, susceptibility

Mondini et al 2023

ASSUMPTIONS

- In a landscape, landslides occur (or do not occur) depending on rainfall and local terrain conditions.
- $P(L \mid R, S) = P(L \mid R) \times P(L \mid S)$
  - L, probability of landslide occurrence; R, rainfall; S, susceptibility

Mondini et al 2023

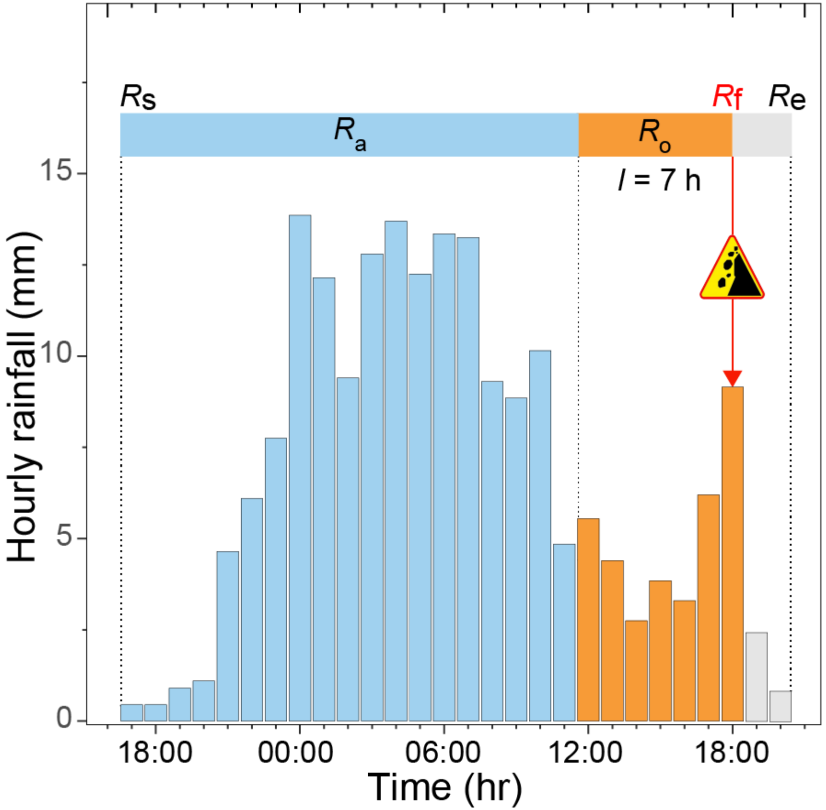
ASSUMPTIONS

- In a landscape, landslides occur (or do not occur) depending on rainfall and local terrain conditions.
- $P(L \mid R, S) = P(L \mid R) \times P(L \mid S) = P(L \mid R) \times c$
  - L, probability of landslide occurrence; R, rainfall; S, susceptibility
  - landslides can occur, $c = 1$; landslides cannot occur, $c = 0$

Mondini et al 2023

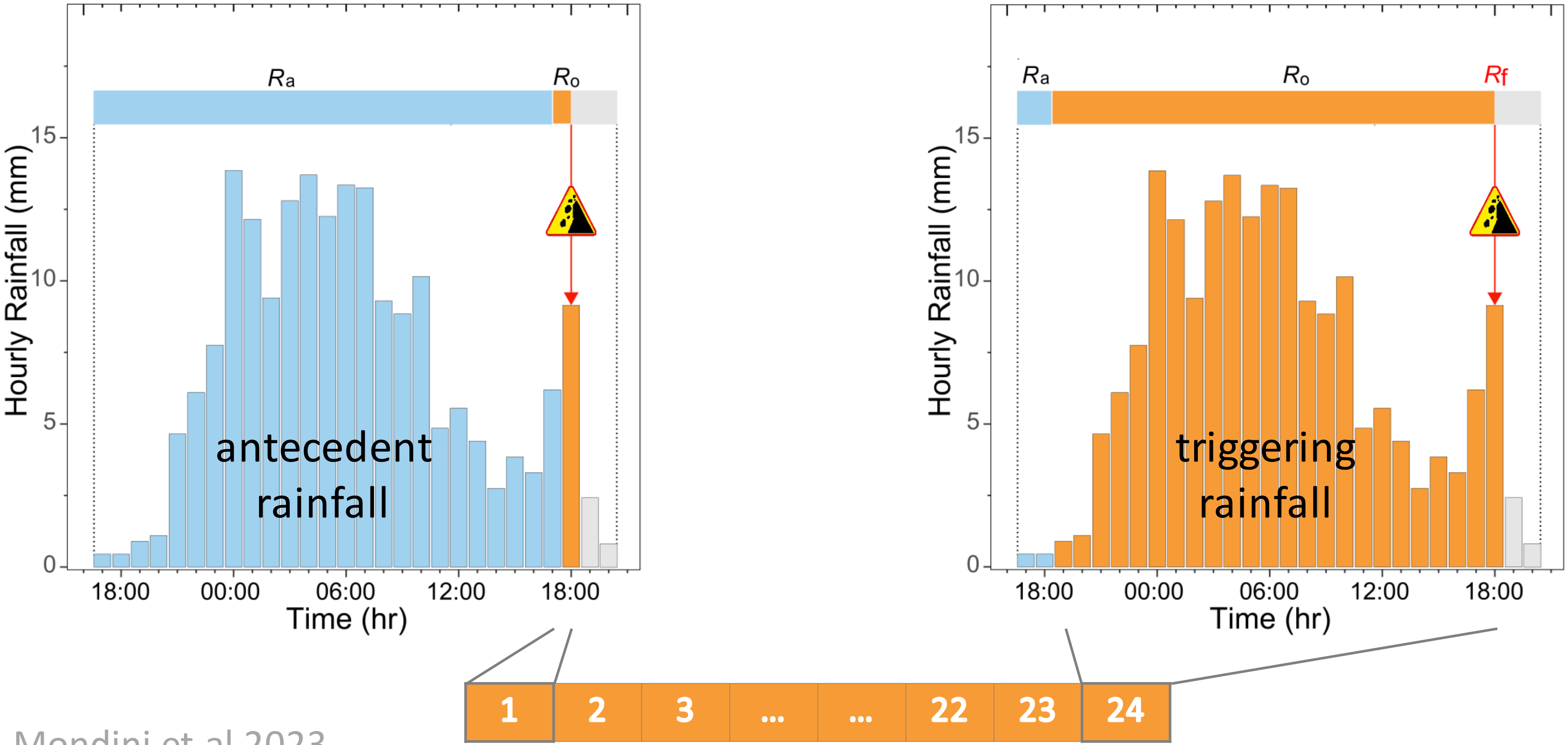
ASSUMPTIONS

- A rainfall event that can generate landslides has an antecedent part and a triggering part



Mondini et al 2023

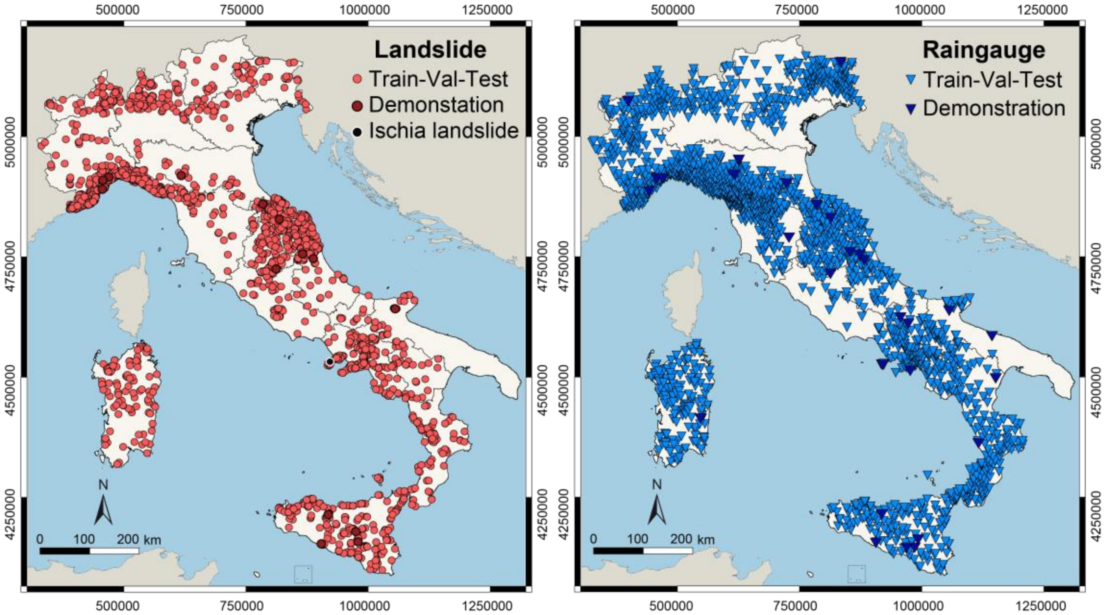
ASSUMPTIONS



Mondini et al 2023

DATA

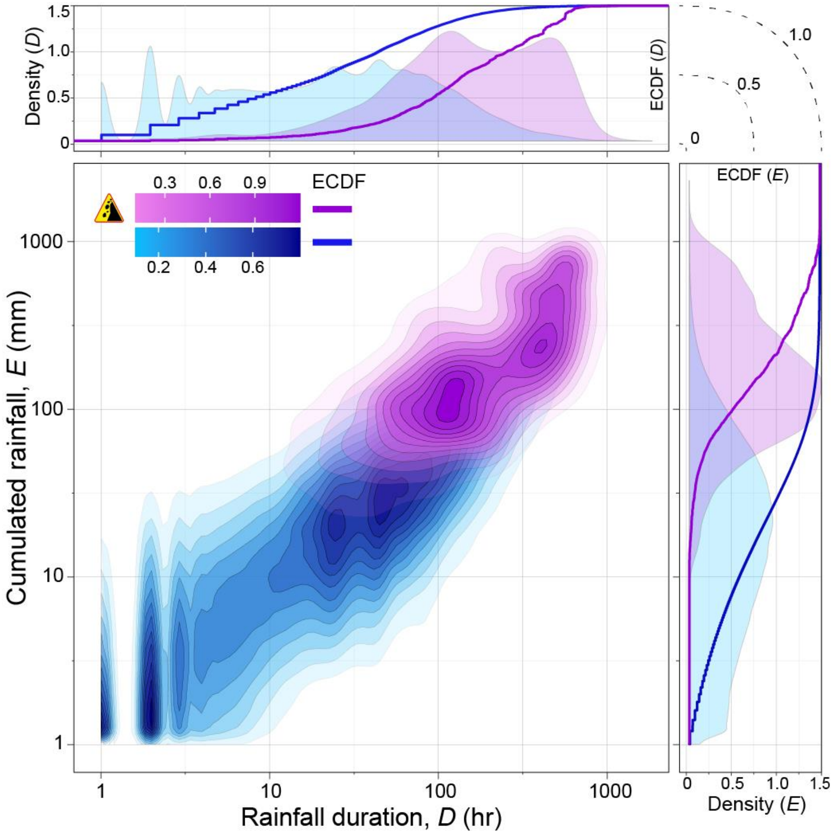
2486 landslides & hourly rainfall from 2096 rain gauges from February 2002 to December 2020



Mondini et al 2023

RAINFALL EVENTS

2472 with landslides &
778,294 without landslides
from February 2002
to December 2020

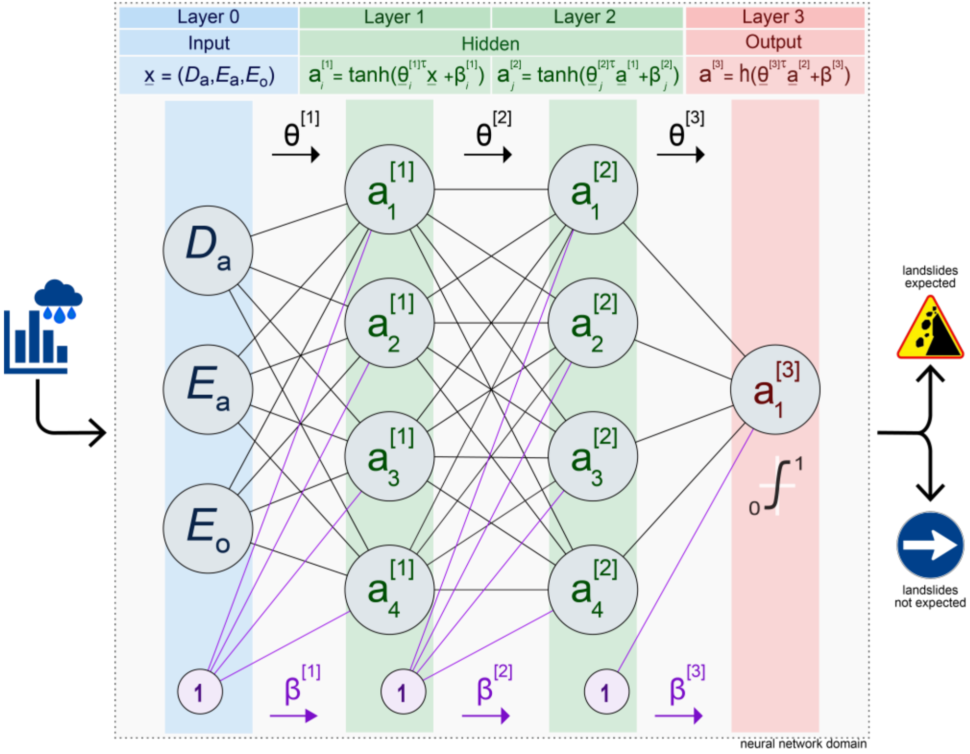


Mondini et al 2023

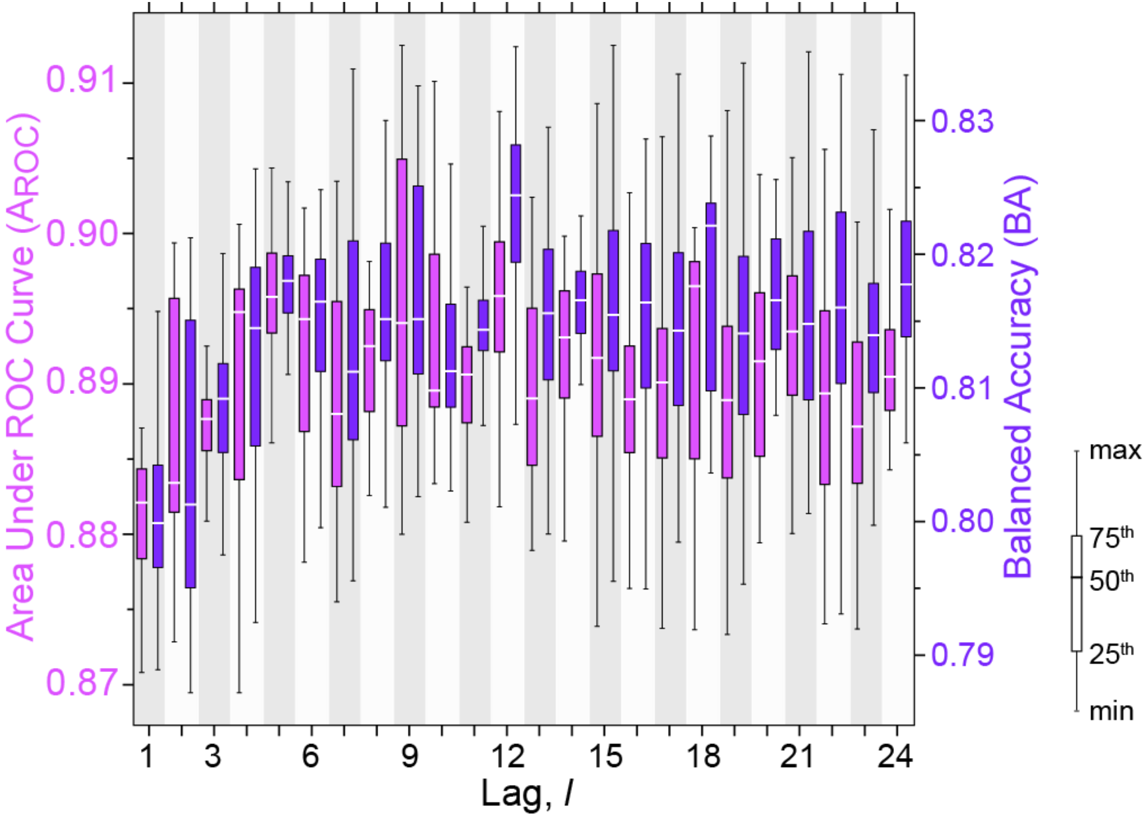
DEEP LEARNING

- 24 sets of fully connected, 4-layer neural networks
- trained with 3 rainfall variables, $\{D_a, E_a, E_o\}$
- 100 models for each hourly period

Mondini et al 2023



PERFORMANCE



Mondini et al 2023

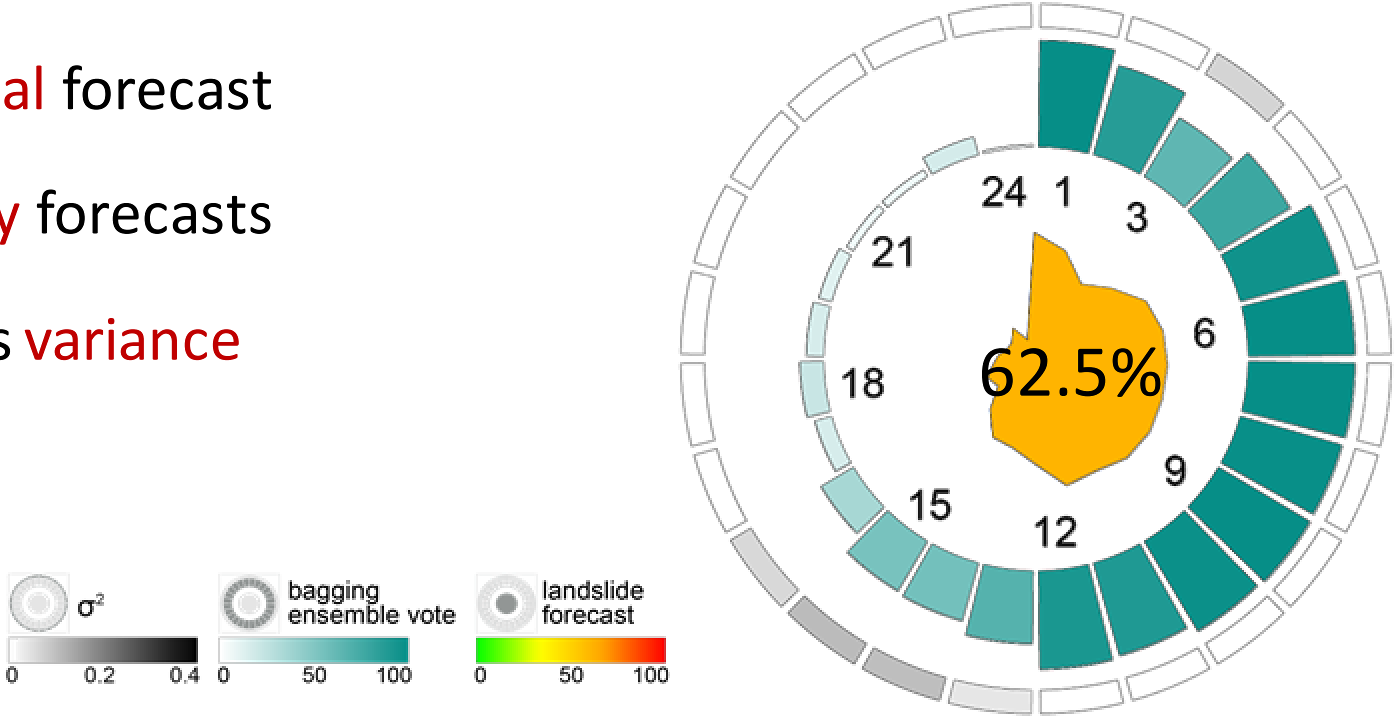
WHICH FORECAST?

- The system produced 24 forecast sets with the associated uncertainties.

Mondini et al 2023

VOTING SCHEME

- one global forecast
- 24 hourly forecasts
- forecasts variance

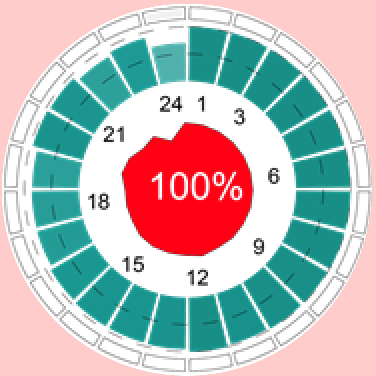


Mondini et al 2023

DEMONSTRATION

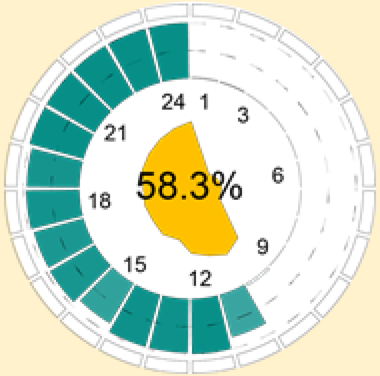
11

landslide forecasted
landslide occurred



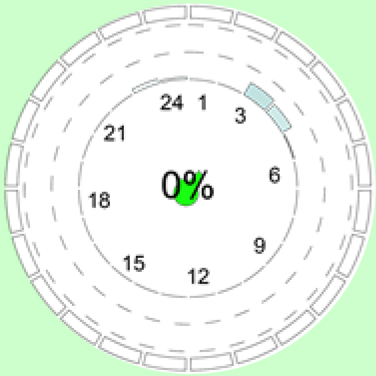
16

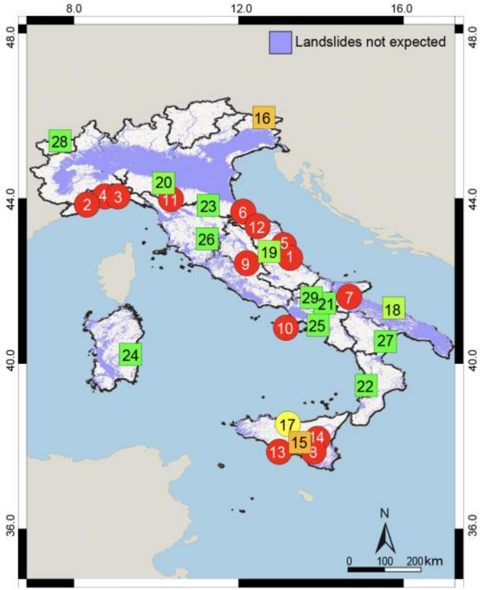
landslide forecasted
no landslide occurred



21

no landslide forecasted
no landslide occurred





~90% correct and ~10% incorrect forecasts

Mondini et al 2023

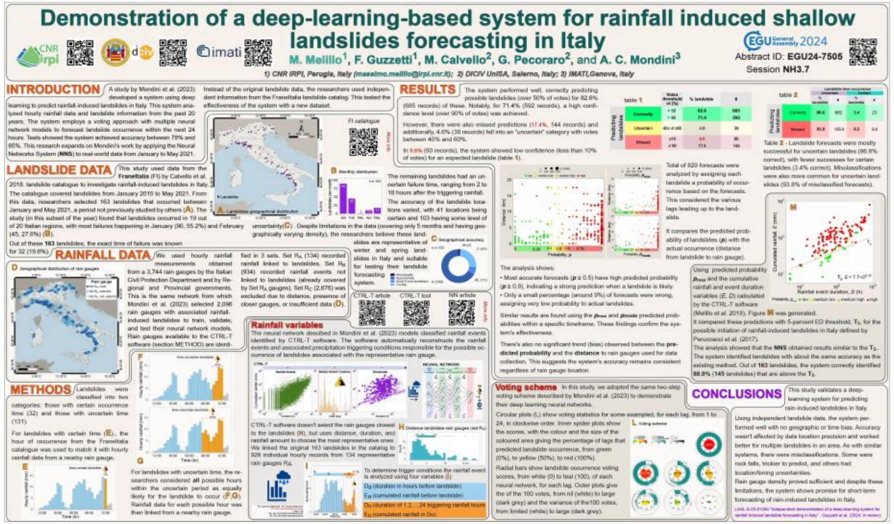
FURTHER DEMONSTRATION

X4.140 | EGU24-7505 | NH3.7

Demonstration of a deep-learning-based system for rainfall induced shallow landslides forecasting in Italy

Massimo Melillo, Fausto Guzzetti, Michele Calvello, Gaetano Pecoraro, and Alessandro C. Mondini

Abstract ID: EGU24-7505 | Posters on site: Thu, 18 Apr, 10:45-12:30 | Hall X4



Demonstration of a deep-learning-based system for rainfall induced shallow landslides forecasting in Italy

M. Melillo[1], F. Guzzetti[1], M. Calvello[2], G. Pecoraro[2], and A. C. Mondini[3]

1) CNR IRPI, Perugia, Italy (massimo.melillo@irpi.cnr.it); 2) DICIV UniSA, Salerno, Italy; 3) IMATI,Genova, Italy

EGU General Assembly 2024
Abstract ID: **EGU24-7505**
Session **NH3.7**

INTRODUCTION A study by Mondini et al. (2023) developed a system using deep learning to predict rainfall-induced landslides in Italy. This system analyzed hourly rainfall data and landslide information from the past 20 years. The system employs a voting approach with multiple neural network models to forecast landslide occurrence within the next 24 hours. Tests showed the system achieved accuracy between 78% and 85%. This research expands on Mondini's work by applying the Neural Networks System (**NNS**) to real-world data from January to May 2021.

Instead of the original landslide data, the researchers used independent information from the FraneItalia landslide catalog. This tested the effectiveness of the system with a new dataset.

LANDSLIDE DATA This study used data from the **FraneItalia** (FI) by Calvello et al. 2018. landslide catalogue to investigate rainfall-induced landslides in Italy. The catalogue covered landslides from January 2010 to May 2021. From this data, researchers selected 163 landslides that occurred between January and May 2021, a period not previously studied by others (A). The study (in this subset of the year) found that landslides occurred in 19 out of 20 Italian regions, with most failures happening in January (90, 55.2%) and February (45, 27.6%) (B).

Out of these **163** landslides, the exact time of failure was known for 32 (19.6%).

The remaining landslides had an uncertain failure time, ranging from 2 to 18 hours after the triggering rainfall.

The accuracy of the landslide locations varied, with 41 locations being certain and 103 having some level of uncertainty (C). Despite limitations in the data (covering only 5 months and having geographically varying density), the researchers believe these landslides are representative of winter and spring landslides in Italy and suitable for testing their landslide forecasting system.

RAINFALL DATA We used hourly rainfall measurements obtained from a 3,744 rain gauges by the Italian Civil Protection Department and by Regional and Provincial governments. This is the same network from which Mondini et al. (2023) selected 2,096 rain gauges with associated rainfall-induced landslides to train, validate, and test their neural network models. Rain gauges available to the CTRL-T software (section METHOD) are identified in 3 sets. Set R_A (134) recorded rainfall linked to landslides. Set R_B (934) recorded rainfall events not linked to landslides (already covered by Set R_A gauges). Set R_C (2,676) was excluded due to distance, presence of closer gauges, or insufficient data (D).

METHODS Landslides were classified into two categories: those with certain occurrence time (32) and those with uncertain time (131).

For landslides with certain time (E), the hour of occurrence from the FraneItalia catalogue was used to match it with hourly rainfall data from a nearby rain gauge.

For landslides with uncertain time, the researchers considered **all** possible hours within the uncertain period as equally likely for the landslide to occur (F,G). Rainfall data for each possible hour was then linked from a nearby rain gauge.

Rainfall variables
The neural network described in Mondini et al. (2023) models classified rainfall events identified by CTRL-T software. The software automatically reconstructs the rainfall events and associated precipitation triggering conditions responsible for the possible occurrence of landslides associated with the representative rain gauge.

CTRL-T software doesn't select the rain gauges closest to the landslides (H), but uses distance, duration, and rainfall amount to choose the most representative ones. We linked the original 163 landslides in the catalog to 928 individual hourly records from 134 representative rain gauges R_A.

To determine trigger conditions the rainfall event is analyzed using four variables (I):

- D_B (duration in hours before landslide)
- E_B (cumulated rainfall before landslide)
- D_D (duration of 1,2,...,24 triggering rainfall hours)
- E_D (cumulated rainfall in D_D)

RESULTS The system performed well, correctly predicting possible landslides (over 50% of votes) for 82.6% (685 records) of these. Notably, for 71.4% (592 records), a high confidence level (over 90% of votes) was achieved.

However, there were also missed predictions (17.4%, 144 records) and additionally, 4.6% (38 records) fell into an "uncertain" category with votes between 40% and 60%.

In 9.6% (80 records), the system showed low confidence (less than 10% of votes) for an expected landslide (table 1).

table 1

Predicting landslides	Votes threshold vt (%)	% landslide	#
Correctly	> 50 > 90	82.6 71.4	685 592
Uncertain	40≤ vt ≤60	4.6	38
Missed	≤10 < 50	9.6 17.4	80 144

table 2

Predicting landslides	Landslide time occurrence: Uncertain % landslide	#	Certain % landslide	#
Correctly	96.6	652	3.4	23
Missed	93.8	135.0	6.2	9.0

Table 2 - Landslide forecasts were mostly successful for uncertain landslides (96.6% correct), with fewer successes for certain landslides (3.4% correct). Misclassifications were also more common for uncertain landslides (93.8% of misclassified forecasts).

Total of 829 forecasts were analyzed by assigning each landslide a probability of occurrence based on the forecasts. This considered the various lags leading up to the landslide.

It compares the predicted probability of landslides (*p*) with the actual occurence (distance from landslide to rain gauge).

The analysis shows:

- Most accurate forecasts (*p* ≥ 0.5) have high predicted probability (*p* ≥ 0.9), indicating a strong prediction when a landslide is likely.
- Only a small percentage (around 9%) of forecasts were wrong, assigning very low probability to actual landslides.

Similar results are found using the p_{max} and *pmode* predicted probabilities within a specific timeframe. These findings confirm the system's effectiveness.

There's also no significant trend (bias) observed between the **predicted probability** and the **distance** to rain gauges used for data collection. This suggests the system's accuracy remains consistent regardless of rain gauge location.

Using predicted probability p_{max} and the cumulative rainfall and event duration variables (*E*, *D*) calculated by the CTRL-T software (Melillo et al. 2018), Figure M was generated.
It compared these predictions with 5-percent ED threshold, T_5, for the possible initiation of rainfall-induced landslides in Italy defined by Peruccacci et al. (2017).
The analysis showed that the **NNS** obtained results similar to the T_5.
The system identified landslides with about the same accuracy as the existing method. Out of **163** landslides, the system correctly identified **88.9%** (**145** landslides) that are above the T_5.

Voting scheme In this study, we adopted the same two-step voting scheme described by Mondini et al. (2023) to demonstrate their deep learning neural networks.

Circular plots (L) show voting statistics for some exampled, for each lag, from 1 to 24, in clockwise order. Inner spider plots show the scores, with the colour and the size of the coloured area giving the percentage of lags that predicted landslide occurence, from green (0%), to yellow (50%), to red (100%).

Radial bars show landslide occurrence voting scores, from white (0) to teal (100), of each neural network, for each lag. Outer plots give the of the 100 votes, from nil (white) to large (dark grey) and the variance of the100 votes, from limited (white) to large (dark grey).

CONCLUSIONS This study validates a deep-learning system for predicting rain-induced landslides in Italy.

Using independent landslide data, the system performed well with no geographic or time bias. Accuracy wasn't affected by data location precision and worked better for multiple landslides in an area. As with similar systems, there were misclassifications. Some were rock falls, trickier to predict, and others had location/timing uncertainties.
Rain gauge density proved sufficient and despite these limitations, the system shows promise for short-term forecasting of rain-induced landslides in Italy.

"Independent demonstration of a deep-learning system for rainfall induced landslide forecasting in Italy", Guzzetti et al. (2024, in review)

Guzzetti et al (in review)

LESSONS LEARNT

- Rainfall-induced shallow landslides occur where and when it rains.
- Rainfall-induced shallow landslides can be forecasted using rainfall measurements.
- Are we prepared to use a system that we do not why it works?

Mondini et al 2023

Session NH3.6
Space and time forecasting of landslides
EGU24-6201

Sharing is encouraged

THANK YOU!

Fausto Guzzetti, Alessandro C. Mondini, Massimo Melillo
Consiglio Nazionale delle Ricerche