# Rainfall nowcasting with machine-learning for landslide early warning

**Fereshteh Taromideh, Giovanni Francesco Santonastaso, and Roberto Greco**

**Dipartimento di Ingegneria, Università degli Studi della Campania "Luigi Vanvitelli", Italy**
**fereshteh.taromideh@unicampania.it**
**https://doi.org/10.5194/egusphere-egu24-11137**

Sharing is encouraged

EGU General Assembly 2024

Università degli Studi della Campania *Luigi Vanvitelli* — *Dipartimento di Ingegneria*

## Methodology

### 1 Problem statement

- Machine learning model for rainfall nowcasting

**Rainfall nowcasting:**
- The accurate prediction of rainfall on short-term lead time.
- Lead time: (10, 30, 60, 120, 180 minutes)
- Lag time: a period before the present time

### 2 Case study

- Island of Ischia in the Gulf of Naples, Italy

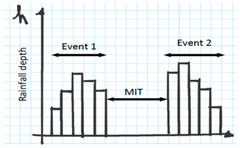
**Data collection:**

**Table 1.** Data sources and their characteristics.

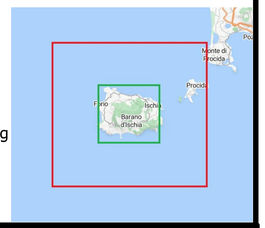
| Factor | Type | Data range | Sample rate |
|---|---|---|---|
| Rainfall radar data | Grid size | November 2020 to December 2023 | 5 minutes (resampled to 10 minutes) |
| Rain gauge information | Dataframe | | 10 minutes |

### 3 Preprocessing

**Rain gauge information**
- Rain gauge stations :Forio
- Calculating cumulative rainfall
- Rain_event, The Minimum Inter-arrival Time (MIT)





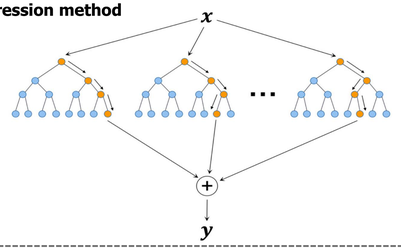



**Rainfall radar data**
- Radar data in the two-monitoring area with different grid sizes:
- 10x10 $km^2$ (green square area),
- 30x30 $km^2$ (Red square area ),

## Machine learning technique

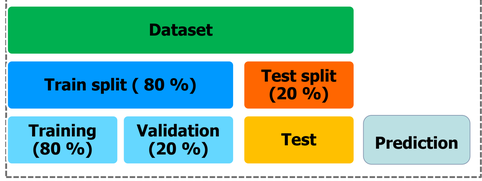
### 4 Random Forest

- **Regression method**




- **Evaluation metric**
- Root-Mean-Square Error (RMSE)
- Cross-Validation: a 5-fold cross-validation is set

**Train-Test Split**




- **Feature Importance**
- Feature importance represents the relative importance of each input variable in predicting the target variable.
- In each fold of the cross-validation process, a column of importance values is created for the features in the dataset.

## Complexity analysis

- **To avoid overfitting**
- Conducting experiments with various combinations of:

**Number of trees (n_estimators )** | **Maximum depth of the trees (max_depth)**

**Table 2.** Effect of RF complexity on RMSE.

▪ **Rainfall radar data 10 x 10 Km²**
▪ **Lag time for the rain gauge information= 0 , -10**
▪ **Lag time for the rainfall radar data= 0, -30**

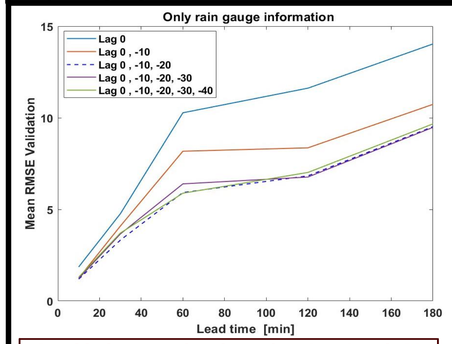
| Number of trees | Max_depth | Lead time (min) | Mean RMSE Validation |
|---|---|---|---|
| 10 | None (no limit) | 120 | 3.893 |
| 20 | | 120 | 3.586 |
| 30 | | 120 | 3.566 |
| 40 | | 120 | 3.479 |
| 50 | | 120 | 3.451 |

▪ **Rainfall radar data 10 x 10 Km²**
▪ **Lag time for the rain gauge information= 0 , -10**
▪ **Lag time for the rainfall radar data= 0, -30**

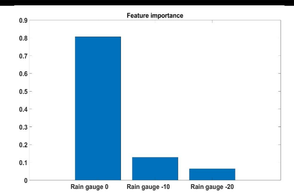
| Number of trees | Max_depth | Lead time (min) | Mean RMSE Validation |
|---|---|---|---|
| 100 | 6 | 120 | 5.191 |
| | 8 | 120 | 4.077 |
| | 10 | 120 | 3.610 |
| | 12 | 120 | 3.466 |
| | 14 | 120 | 3.422 |

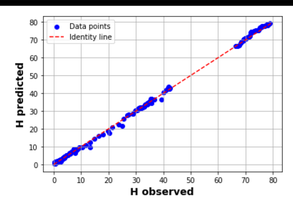
**Suitable values:**
**Number of trees: 30 Max_depth: 10**

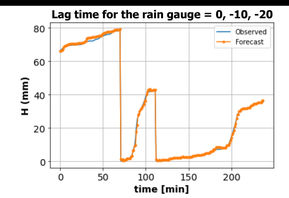
## Only rain gauge input data



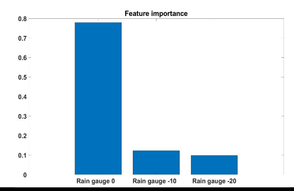

**Considering more lag times than 0, -10 and -20 minutes does not improve significantly RMSE.**

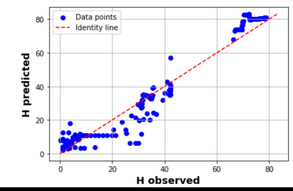


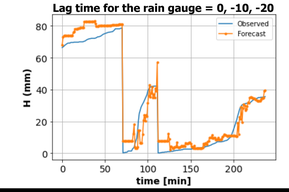




**Lead time: 10 min**




**Lead time: 120 min**

## Rain gauge and rainfall radar input data 10 x 10 Km²




**10 x 10 Km² rainfall radar data input adds valuable prediction skill.**










**Lead time: 10 min**










**Lead time: 120 min**

## Rain gauge and rainfall radar input data 30 x 30 Km²




**30 x 30 Km² rainfall radar data input does not introduce valuable information.**










**Lead time: 10 min**










**Lead time: 120 min**