NAPLES | ITALY | 12–15 JUNE 2023

Università degli Studi della Campania *Luigi Vanvitelli*

# Landslide Early Warning System based on Machine learning and radar data

**AUTHORS**

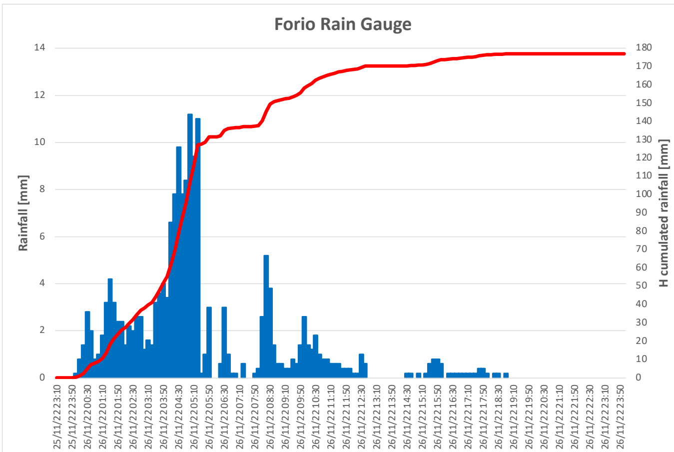
Giovanni Francesco Santonastaso, Fereshteh Taromideh, Pasquale Marino, Daniel Camilo Roman Quintero, and Roberto Greco

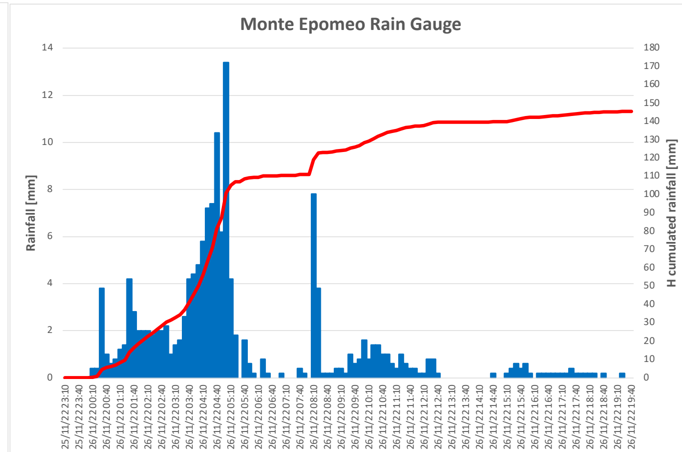
Università degli Studi della Campania "Luigi Vanvitelli", Department of Engineering, Aversa (CE) Italy (giovannifrancesco.santonastaso@unicampania.it)

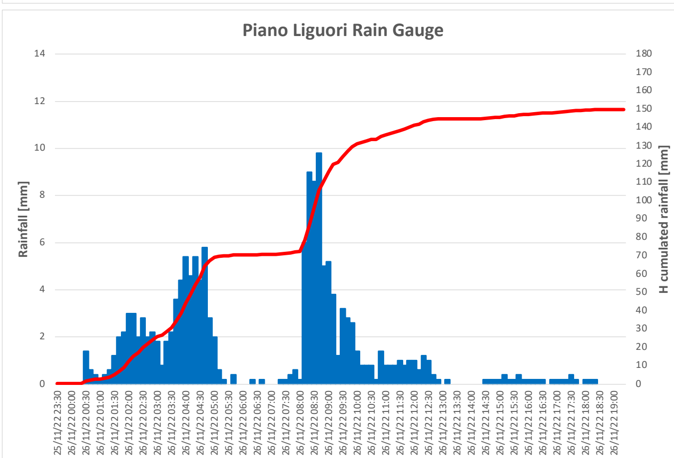
European Geosciences Union

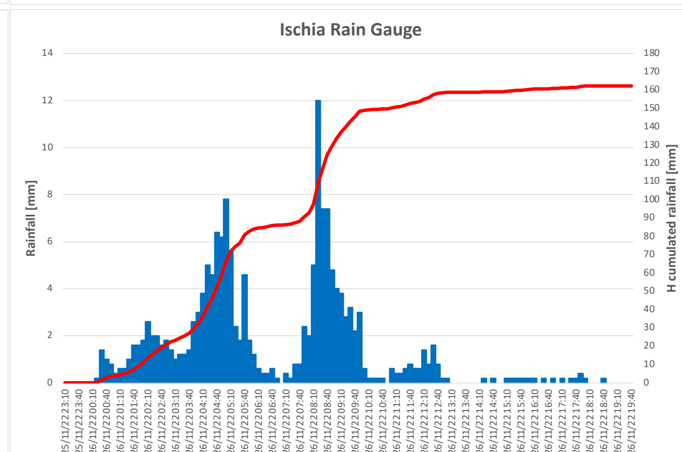
## *Introduction*

In the island of Ischia, on November 26, 2022, heavy rain triggered landslides that killed people and caused great damage to buildings and roads. The 6-hour cumulative rainfall (between 00:00 on 25/11 and 06:00 on 26/11) resulted 126 mm. The storm occurred can be considered a convective event.

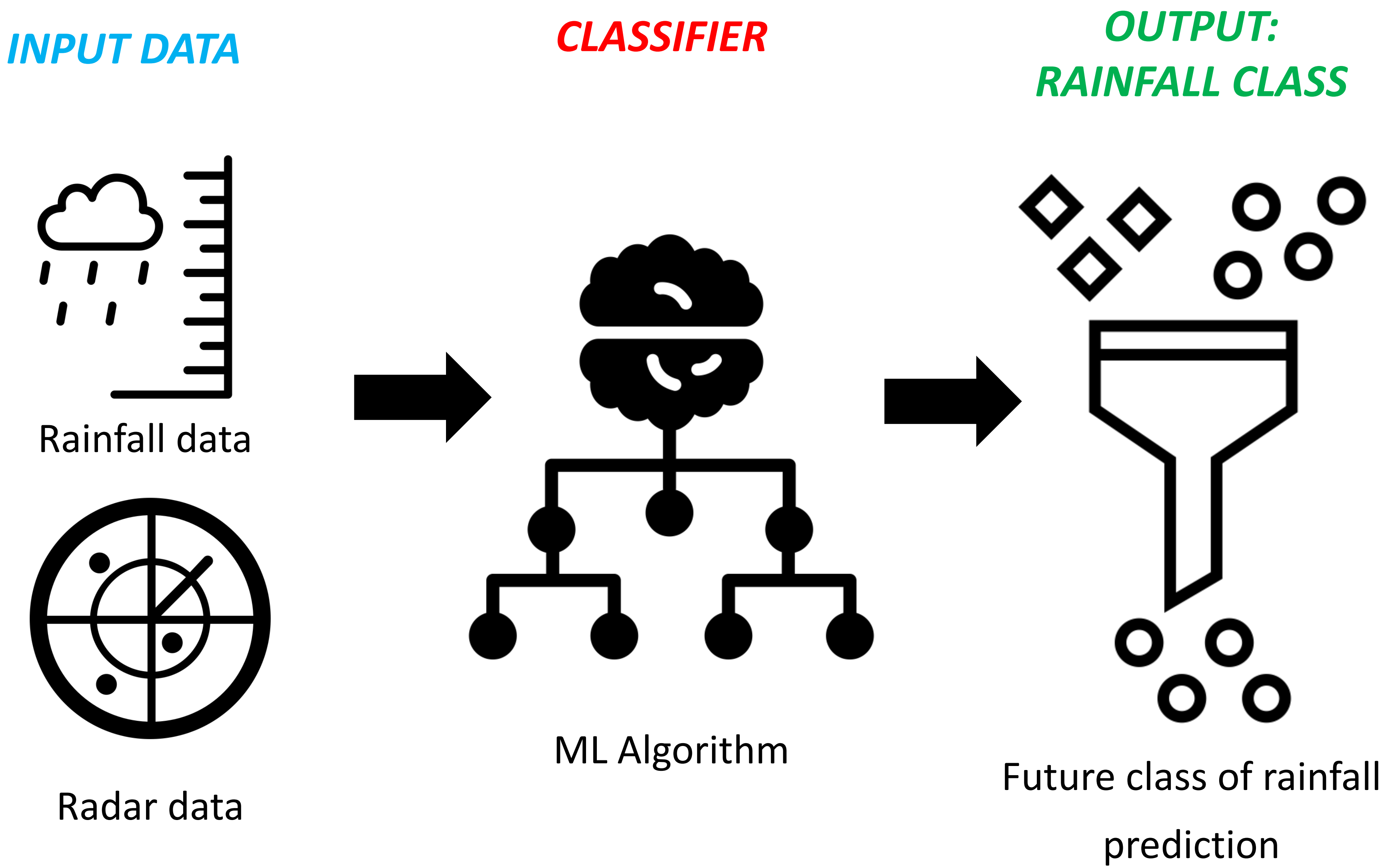


# Landslide Early Warning System (LEWS)



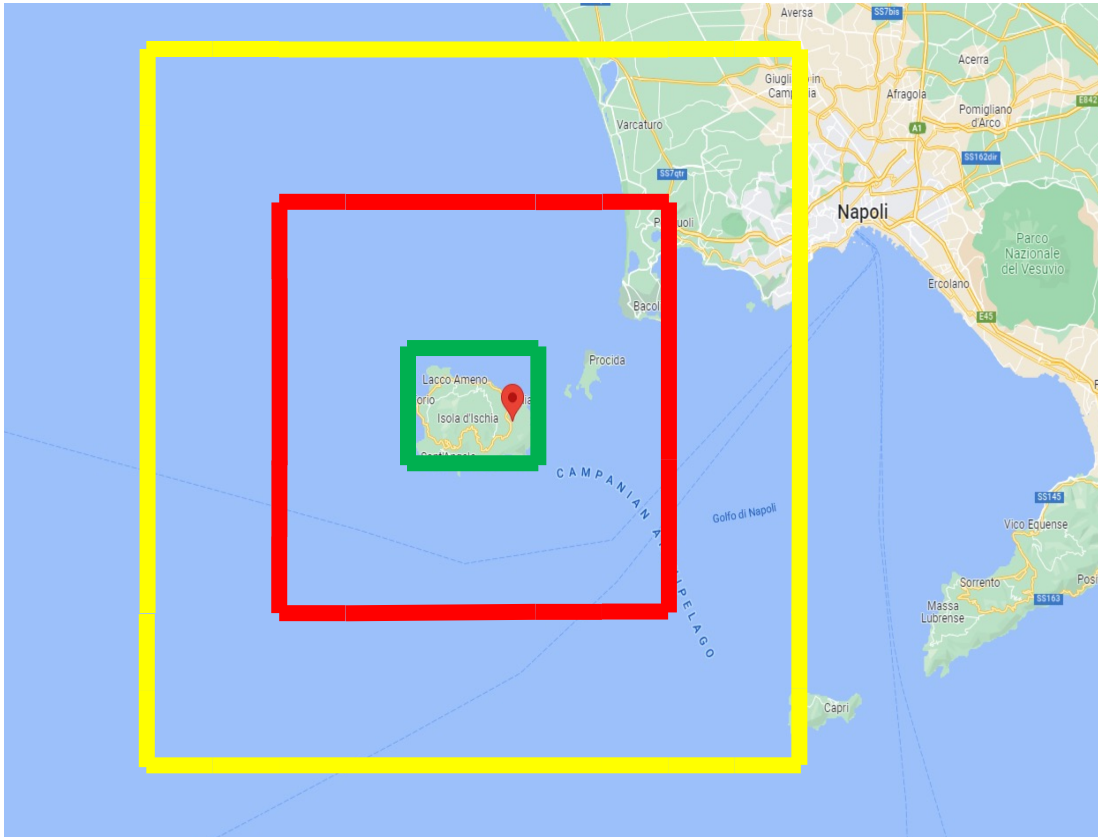
## *Input Data: Radar Data*

- ✓ Radar Data cover the entire national territory. Spatial resolution resolution: **1kmx1km;** Temporal resolution: **5 minutes**

- ✓ Radar data were resampled with a time interval of 10 minutes in order to fit the same time interval of rain gauge

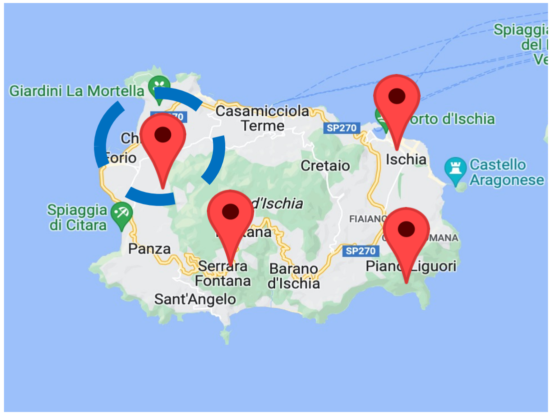
- ✓ The average value of Surface Rainfall intensity were calculated in the three monitoring areas:

1. **Yellow square area** (located 20km from Ischia): **ASR20km**;
2. **Red square area** (located 10 km from Ischia): **ASR10km**;
3. **Green square area** (located at border of Island): **ASRIschia**.

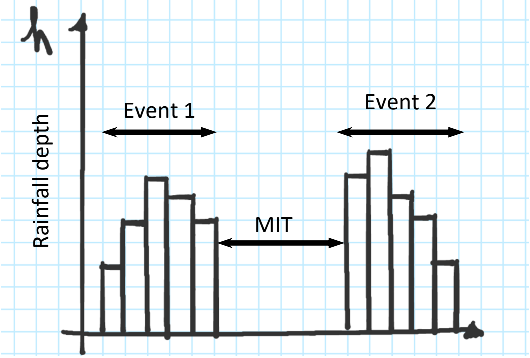


## *Input Data: Rainfall data*

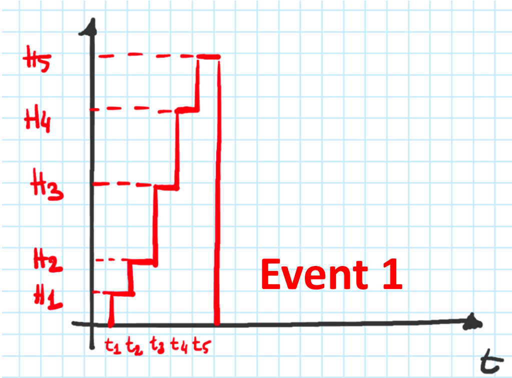
✓ 4 Rain Gauges installed in Ischia with a temporal resolution of **10 minutes;**

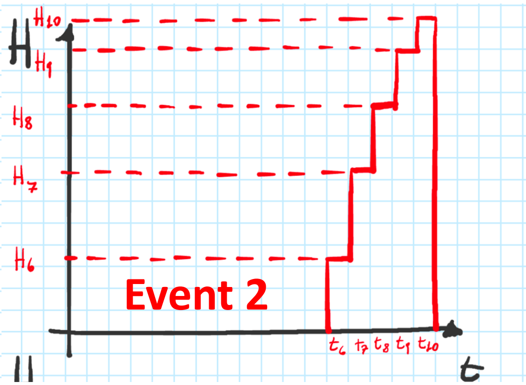



✓ MIT criteria was adapted for defining rainfall events: a rainfall event is defined as a period of rainfall with preceding and succeeding dry periods more than MIT (≧ 2 hours);




✓ For each event the cumulated rainfall, **$H_i$**, was computed;



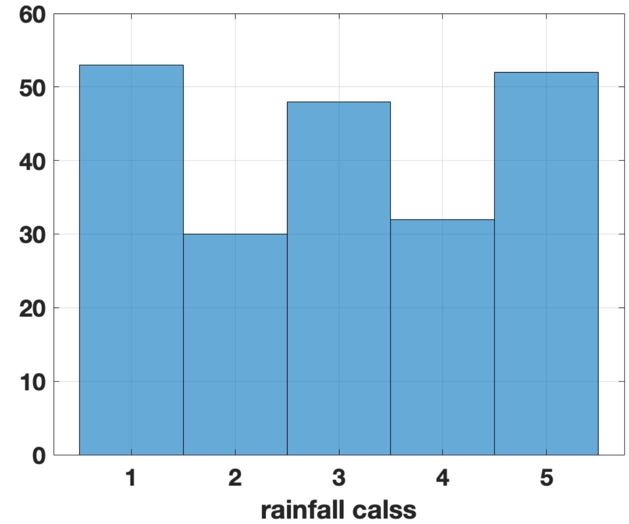




✓ Based on 5 thresholds values the flowing rainfall classes are defined and :

35 mm ≦ C1 < 50 mm
50 mm ≦ C2 < 70 mm
70 mm ≦ C3 < 100 mm
100 mm ≦ C4 < 160 mm
C5 ≧ 160 mm

## *Input Data: input Time Serie for classifier model*

- ✓ The data are collected from ***19/10/2020*** to ***31/12/2022***. 365 rainfall events are identified;
- ✓ The input array used to predict the class of rainfall at time ***t + Δt*** , where ***Δt*** interval of prediction; is reported in the following table:



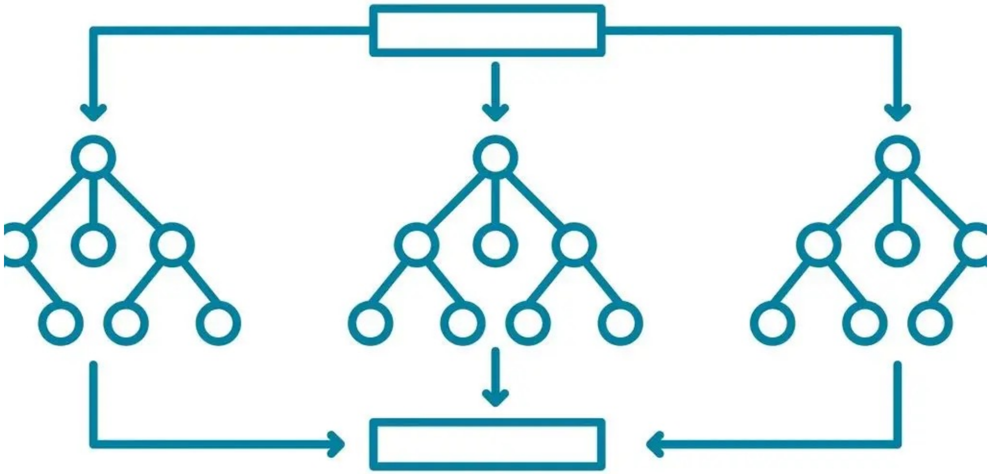

| $H_i$ | ASR20km | ASR10km | ASRIschia |
|---|---|---|---|
| $H_1$ | $ASR20km_1$ | $ASR10km_1$ | $ASRIschia_1$ |
| $H_2$ | $ASR20km_2$ | $ASR10km_2$ | $ASRIschia_2$ |
| $H_3$ | $ASR20km_3$ | $ASR10km_3$ | $ASRIschia_3$ |
| ⋮ | ⋮ | ⋮ | ⋮ |
| ⋮ | ⋮ | ⋮ | ⋮ |
| ⋮ | ⋮ | ⋮ | ⋮ |
| $H_{N-\Delta t}$ | $ASR20km_{N-\Delta t}$ | $ASR10km_{N-\Delta t}$ | $ASRIschia_{N-\Delta t}$ |

## *Classifier*

***Random forest (RF)*** is a machine learning algorithm that uses an ensemble of decision trees to classify data. Each decision tree is trained on a subset of the data and a random set of features, and the final prediction is made by combining the results of all trees.

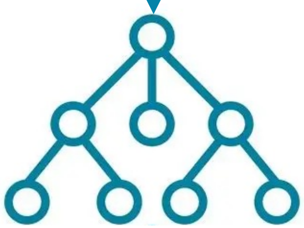
***Advantages***:

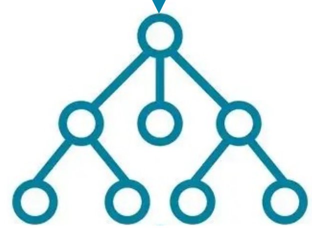
1. RF can handle large datasets without overfitting;
2. RF can capture complex relationships between variables;
3. Finally, it can provide insights into feature importance, showing which variables are most relevant for classification.

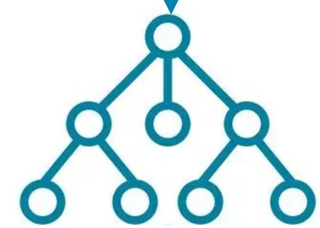


## *Classifier*

| $H_i$ | ASR20km | ASR10km | ASRIschia |
|---|---|---|---|
| $H_1$ | $ASR20km_1$ | $ASR10km_1$ | $ASRIschia_1$ |
| $H_2$ | $ASR20km_2$ | $ASR10km_2$ | $ASRIschia_2$ |
| ⋮ | ⋮ | ⋮ | ⋮ |
| $H_{N\text{-}\Delta t}$ | $ASR20km_{N\text{-}\Delta t}$ | $ASR10km_{N\text{-}\Delta t}$ | $ASRIschia_{N\text{-}\Delta t}$ |







| Predicted class |
|---|
| $C1_2$ |
| $C1_3$ |
| ⋮ |
| $C3_N$ |

*Size of Dataset, N=225*

***70% of data are used as training dataset***

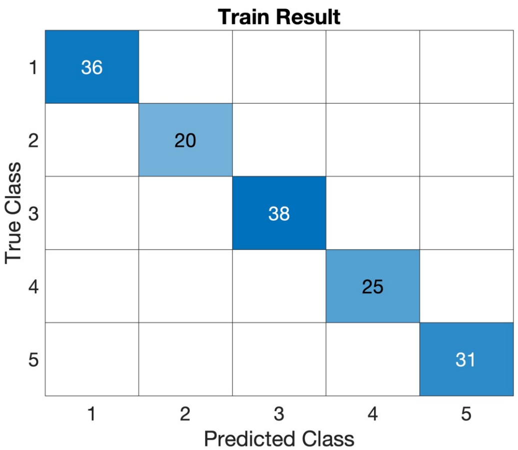
***30% of data are used as validation dataset***

# Results

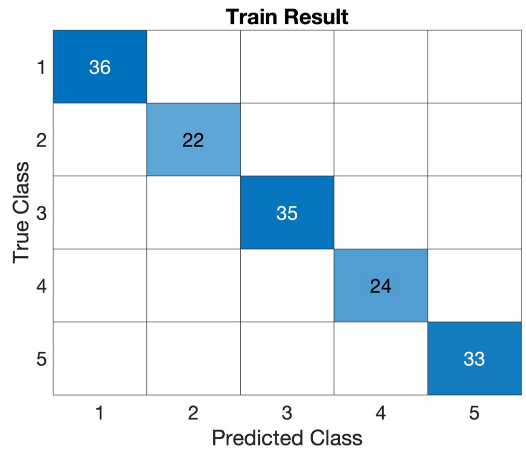
3 models of Random forest are trained using different ***Δt*** interval of prediction:

M1 with *Δt=0.5h*; M2 with **Δ**t=1h; M3 with **Δ**t=3h.

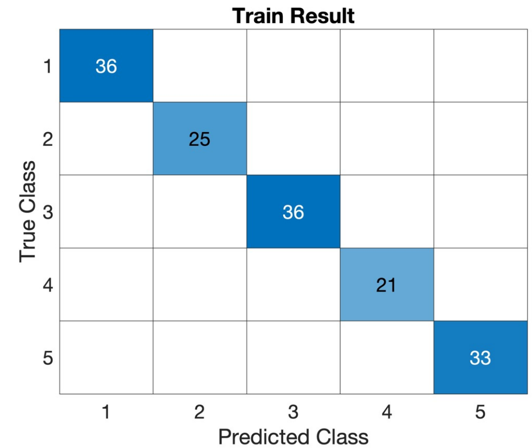
**Accuracy train =100%**



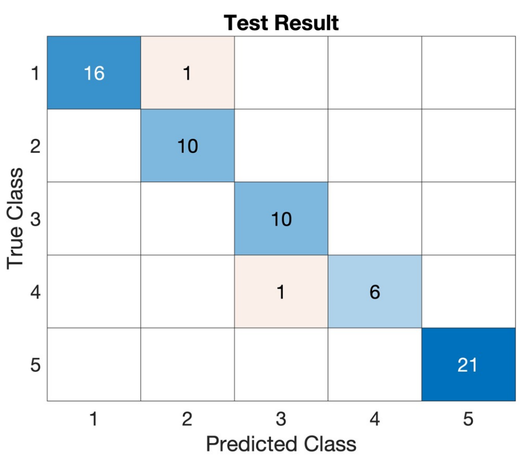

**Accuracy train =100%**




**Accuracy train =100%**

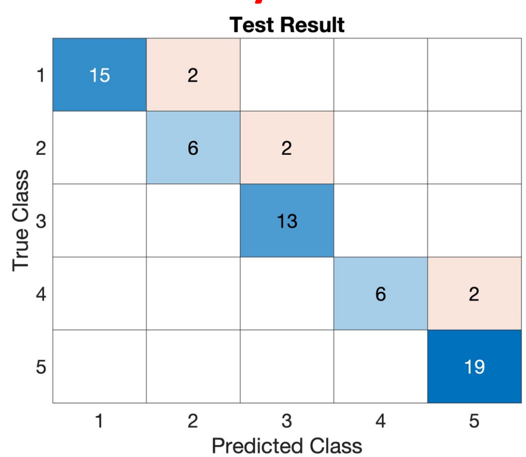



**Accuracy test=97%**



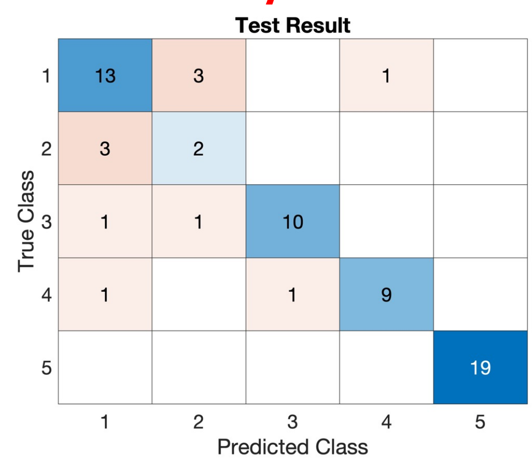

M1 with *Δt=0.5h*

**Accuracy test=91%**




M2 with *Δt=1h*

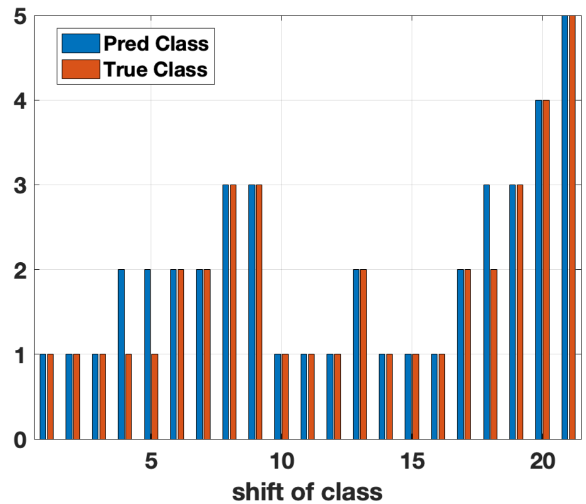
**Accuracy test=82%**




M3 with *Δt=3h*

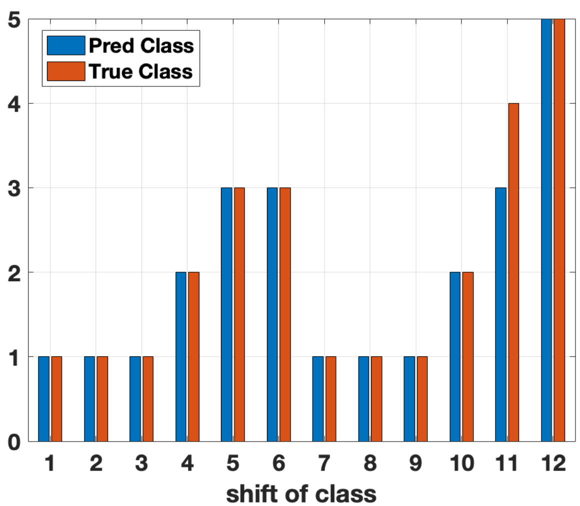
## *Results*

What happens if the model has to predict the precipitation class in case of a class shift, i.e. if the class of the input cumulative precipitation is different from the one to be predicted?

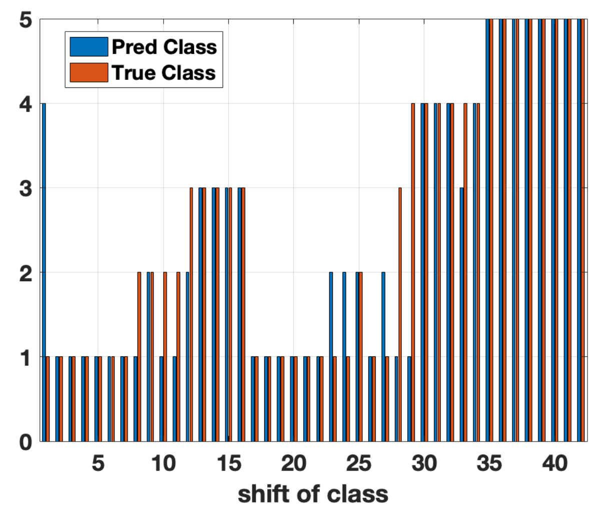



**M1 with *Δt=0.5h***
***Number of shifts=12***
***Error prediction = 1***
***Lower class pred= 0***
***Greater class pred=1***

**M3 with *Δt=3h***
***Number of shifts=42***
***Error prediction = 10***
***Lower class pred= 6***
***Greater class pred=4***




**M2 with *Δt=1h***
***Number of shifts=21***
***Error prediction = 3***
***Lower class pred= 0***
***Greater class pred=3***



## *Conclusions*

- ✓ A novel approach to LEWS, based on machine learning, radar and rainfall data is proposed;
- ✓ 3 Random forest models with different predictive power are trained (M1 with *Δt=0.5h*; M2 with **Δ**t=1h; M3 with **Δ**t=3h);
- ✓ The results show good performance for both 3 models with the lowest accuracy value of 82%; all 3 models also perform well with a class change;
- ✓ All tree models perform better at predicting higher classes (more important in the case of LEWS) than lower classes.

## *Future works*

- ✓ ***Improve the model by adding more information as input data to include not only the distance from the island, but also the direction of the storm;***
- ✓ **Use the RF model as an interpolation tool to combine radar and precipitation data as hydrologic input.**

***Thank you for***

***your attention!***