EGU General Assembly 2021

OGS National Institute of Oceanography and Applied Geophysics

INGV Istituto Nazionale Di Geofisica e Vulcanologia

An optimized online version of NESTORE software package for the forecasting of strong aftershocks: an application to Italian clusters

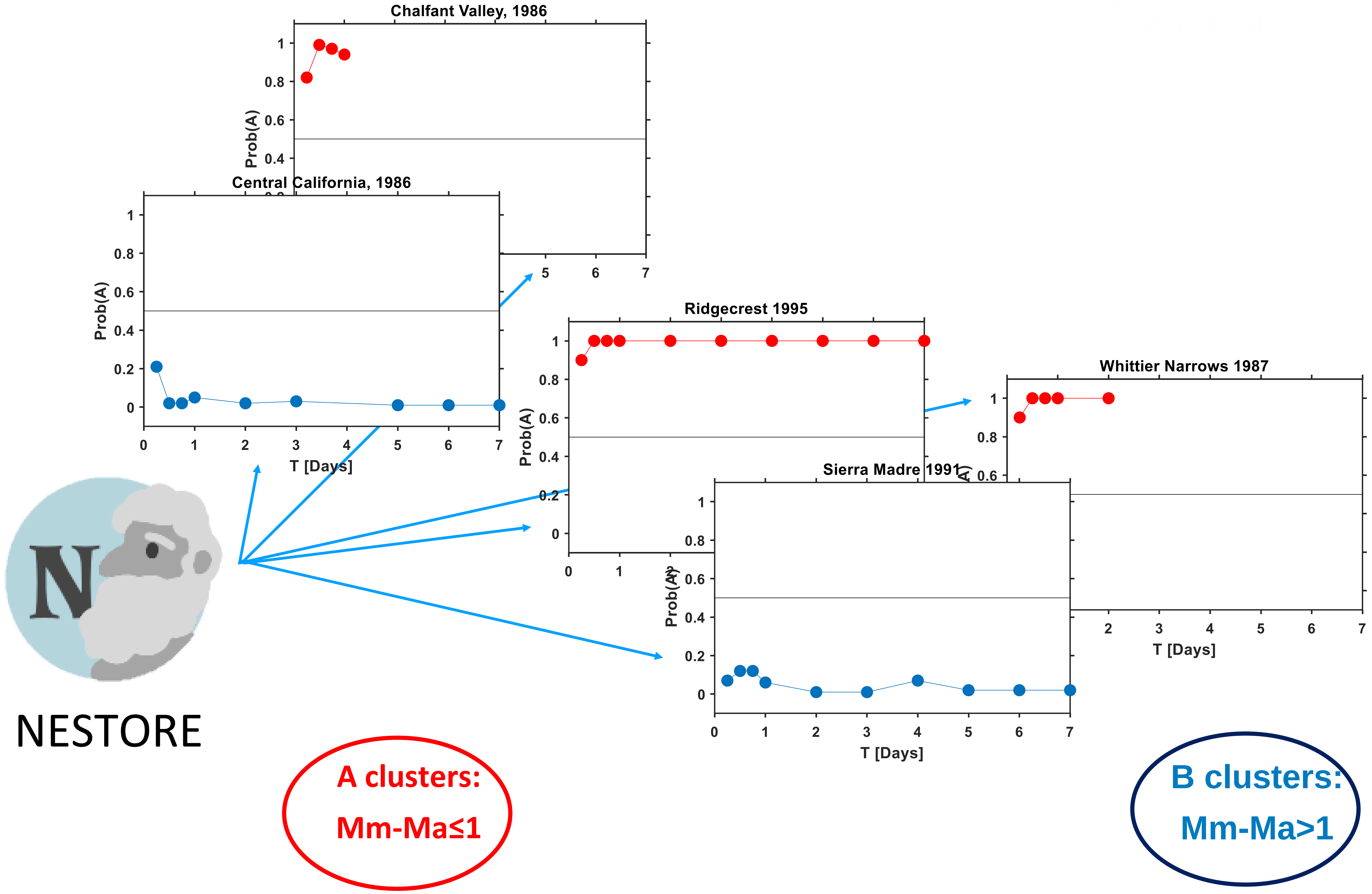
Stefania Gentili, Rita di Giovambattista, Piero Brondi

NESTORE

- **NESTORE** – (*NExt STrOng Related Earthquake*) is an algorithm based on machine learning approach

- It divides the clusters into two classes based on the difference in magnitude Dm between the o-mainshock (the first large shock e.g. M>4) and the strongest following earthquake
(Type A: Dm≤1 Type B: Dm>1)

- NESTORE goal is type A clusters probabilistic forecasting based on features extracted from seismic catalogues in the first hours/days after the o-mainshock



OGS INGV

Ministry of Foreign Affairs and International Cooperation

Funded by a grant from the Italian Ministry of Foreign Affairs and International Cooperation

NESTOREv1.0 Features

NUMBER AND SPATIAL DISTRIBUTION OF EVENTS

$$N_2(i) = \sum_i n_i \quad \text{if } m_i > M_m - 2$$

$$Z(i) = \frac{mean(10^{0.69m_i - 3.22})}{mean(r_{ij})}$$

SOURCE AREA TREND

$$S(i) = \sum_i 10^{(m_i - M_m)}$$

$$SLCum(i) = \sum_i abs[S(t_i) - S(t_{i-1})\frac{i \cdot dt}{(i-1) \cdot dt}$$

$$SLCum2(i) = \sum_i abs[S([s_1 + (i-1) \cdot dt, s_1 + i \cdot dt]) - S([s_1 + (i-1) \cdot dt, s_1 + (i-1) \cdot dt + d\tau])\frac{dt}{d\tau}]$$

ENERGY AND MAGNITUDE TREND

$$Q(i) = \frac{\sum_i E_i}{E_m}$$

$$QLCum(i) = \sum_i abs[Q(t_i) - Q(t_{i-1})\frac{i \cdot dt}{(i-1) \cdot dt}$$

$$V_m(i) = \sum_i |m_i - m_{i-1}|$$

$$QLCum2(i) = \sum_i abs[Q([s_1 + (i-1) \cdot dt, s_1 + i \cdot dt]) - Q([s_1 + (i-1) \cdot dt, s_1 + (i-1) \cdot dt + d\tau])\frac{dt}{d\tau}]$$

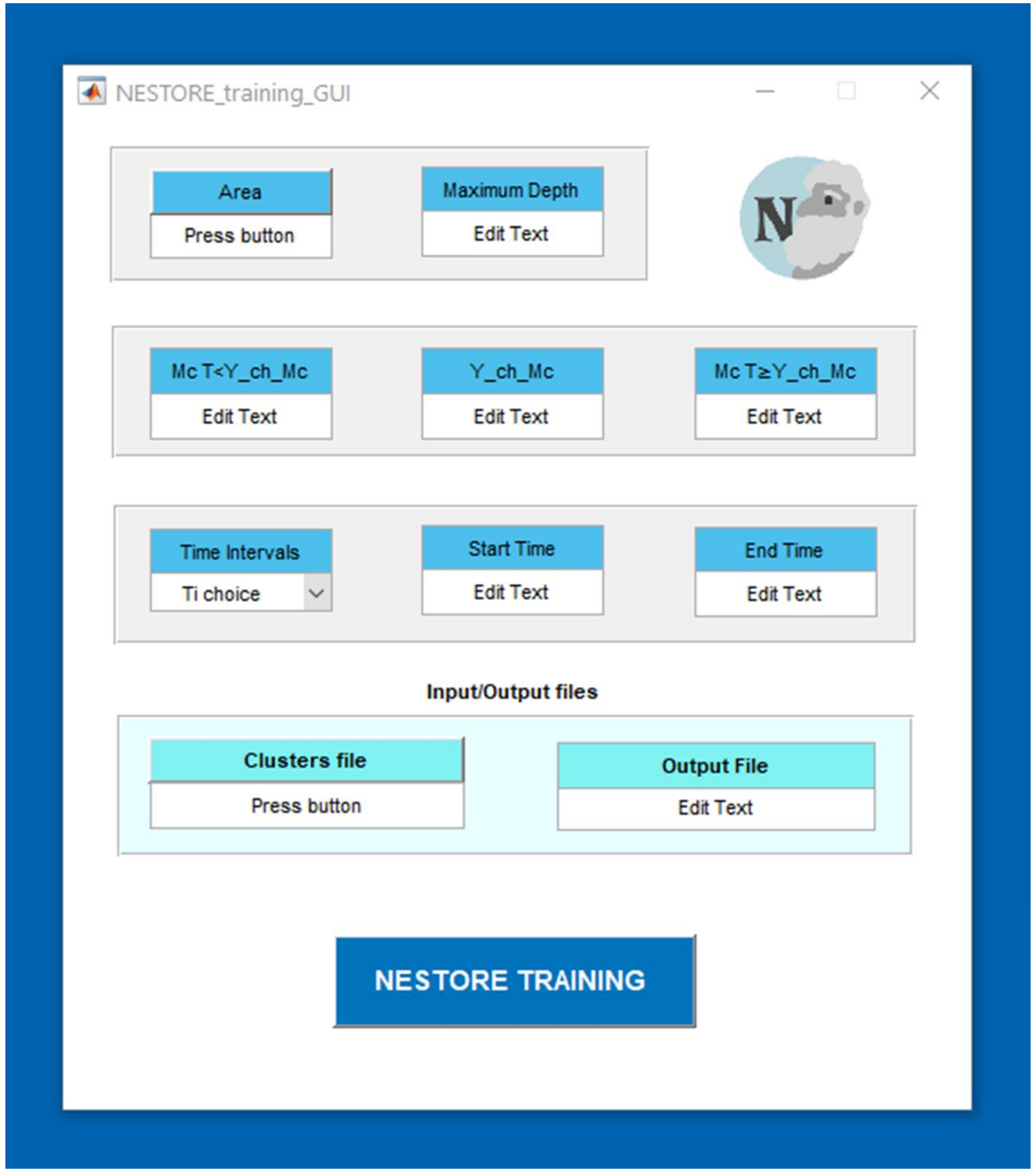
Ministry of Foreign Affairs and International Cooperation

Funded by a grant from the Italian Ministry of Foreign Affairs and International Cooperation

NESTOREv1.0 modules

- **NESTOREv1.0** software package consists of four main parts:
  - **cluster identification module**
  - **training module**
  - **testing module**
  - **NRT classification (1 cluster)**

- All modules can be launched separately by command line, a dedicated user interface (GUI) is being developed too

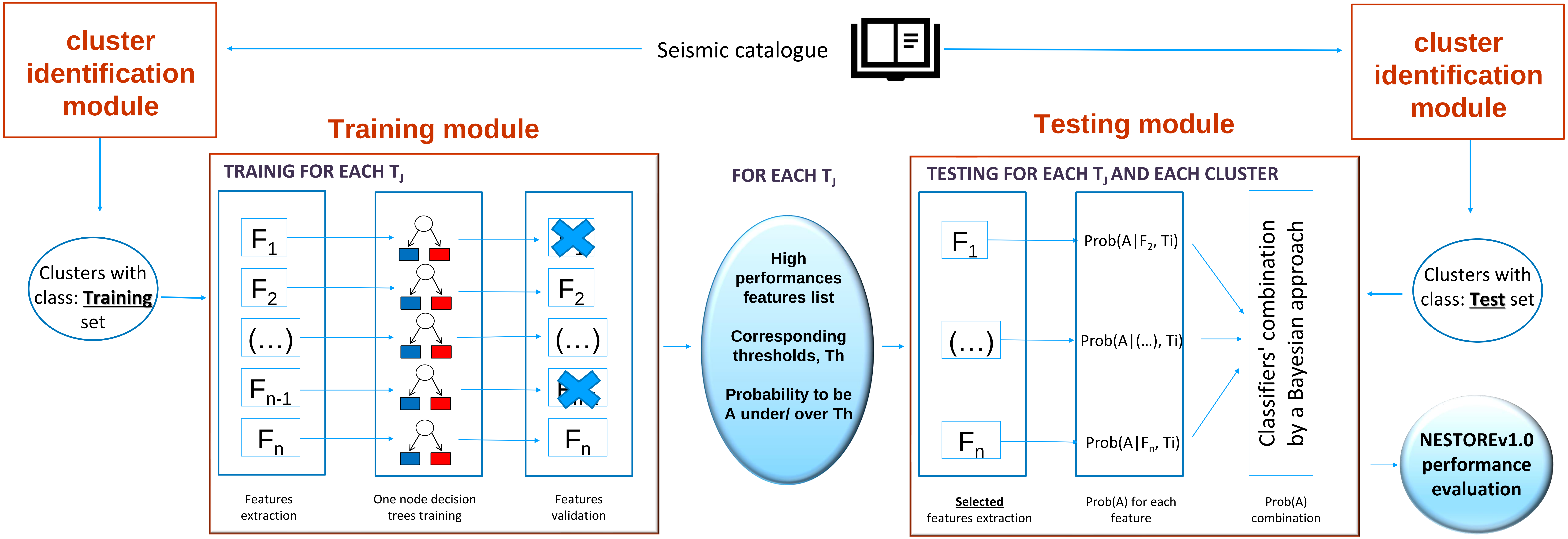


AVAILABLE ONLINE IN THE NEXT FEW MONTHS

Ministry of Foreign Affairs and International Cooperation
Funded by a grant from the Italian Ministry of Foreign Affairs and International Cooperation

NESTOREv1.0 – TRAINING and TESTING

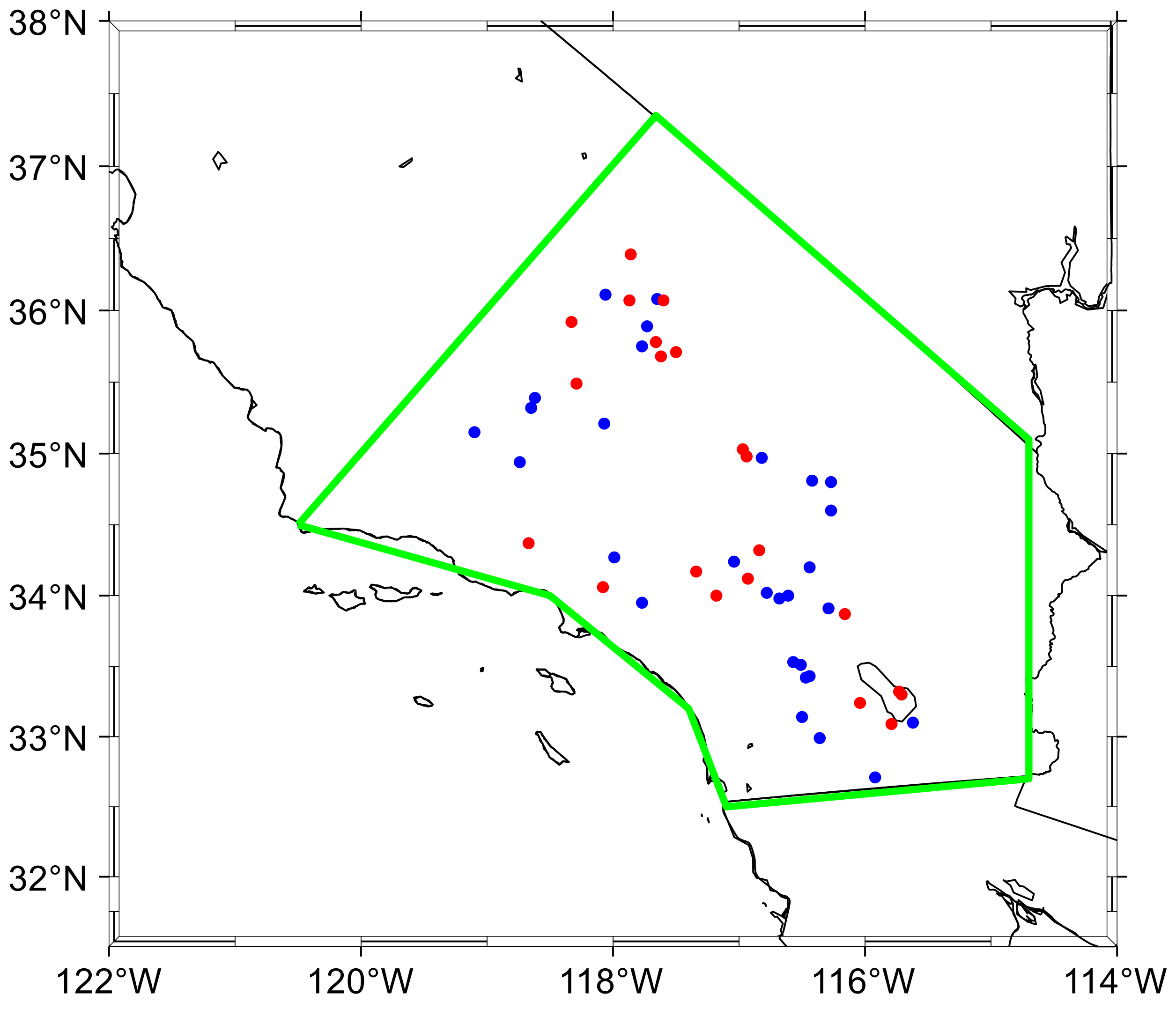
- NESTOREv1.0 analyses the seismic data at increasing time intervals T=[0.25,0.5,0.75,1,2,…,7] days (T_i) after the mainshock.
- After the training NESTORE can supply the probability to have an A cluster at time intervals T_i for new clusters



Funded by a grant from the Italian Ministry
of Foreign Affairs and International Cooperation

Ministry of Foreign Affairs and International Cooperation

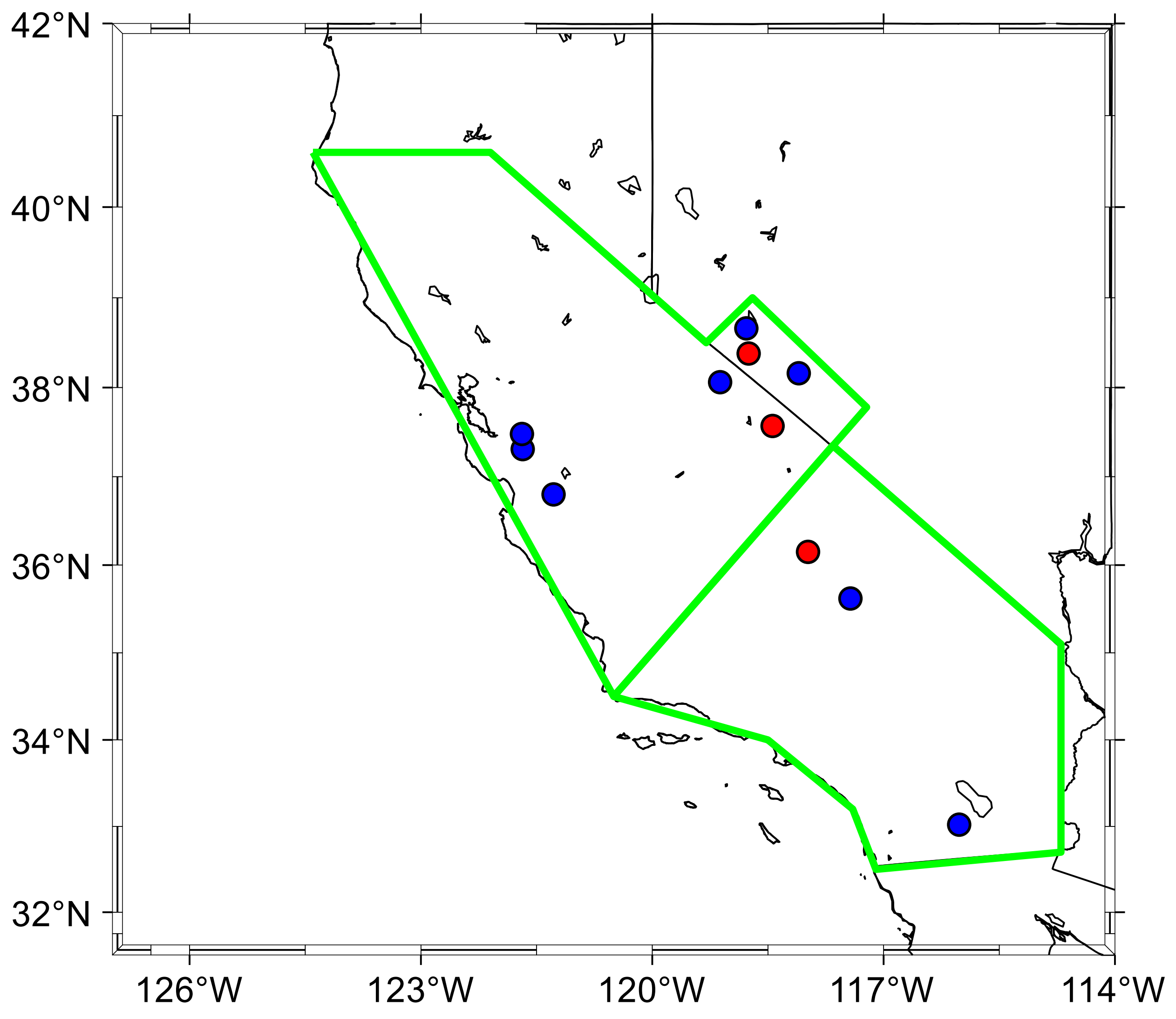
Training and testing NESTORE v1.0 in California

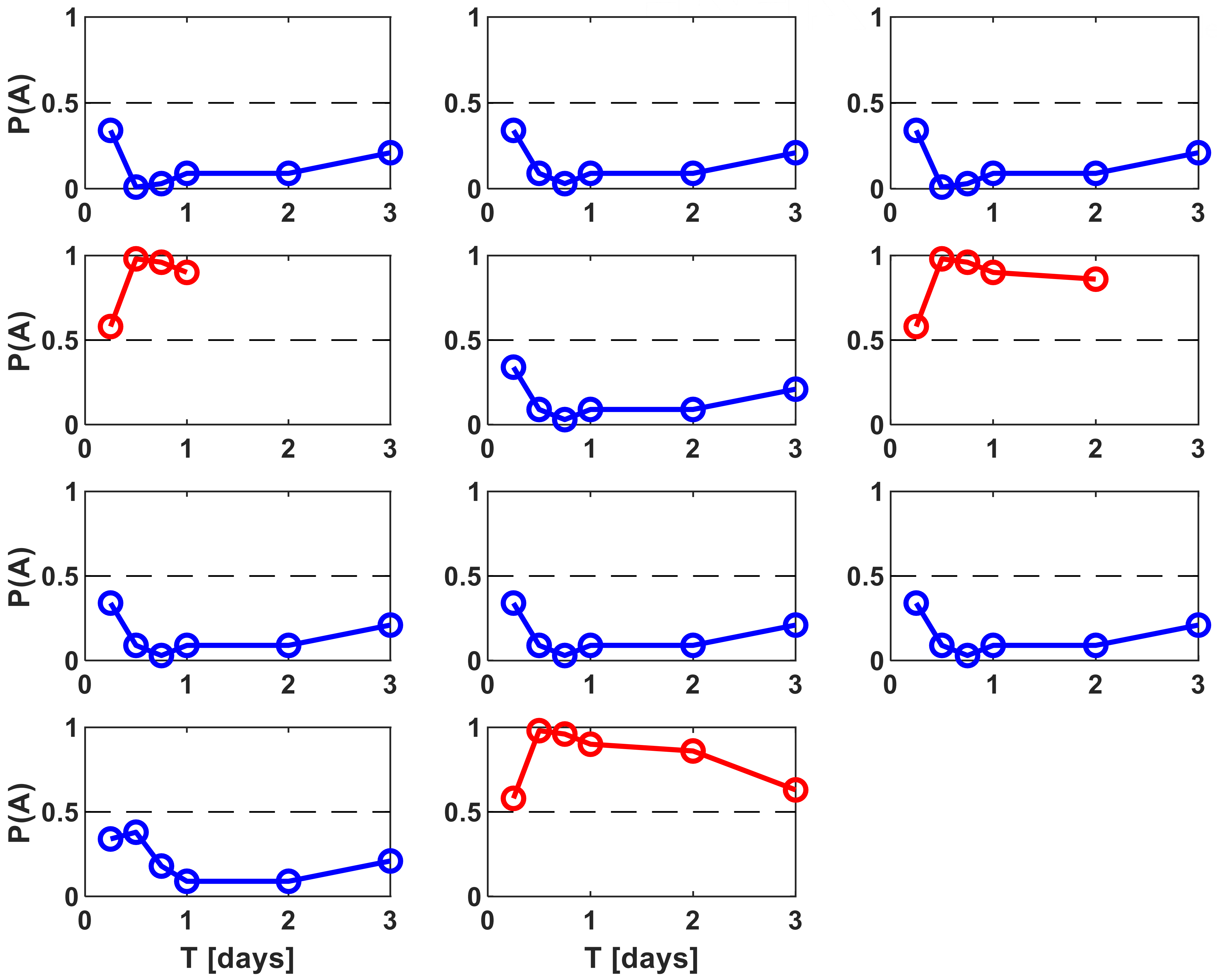


Dots: clusters mainshock location (50 clusters); red: type A clusters; blue: type B clusters.

Training set:
SCSN [1981, 2020)
50 clusters

Test set:
(ComCat) [1981, 2021] +
SCSN [2020, 2021]
11 clusters





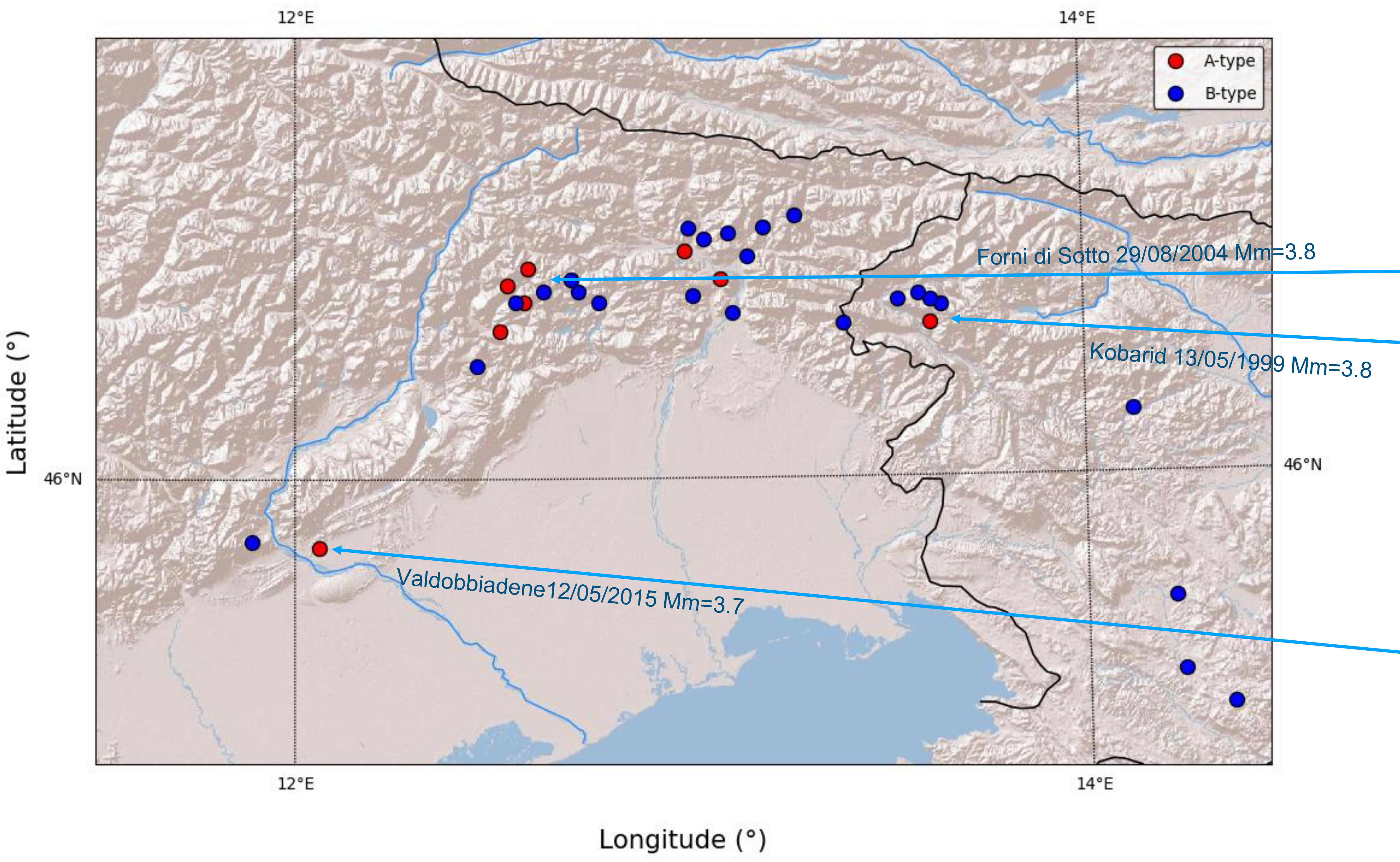
Performances during time

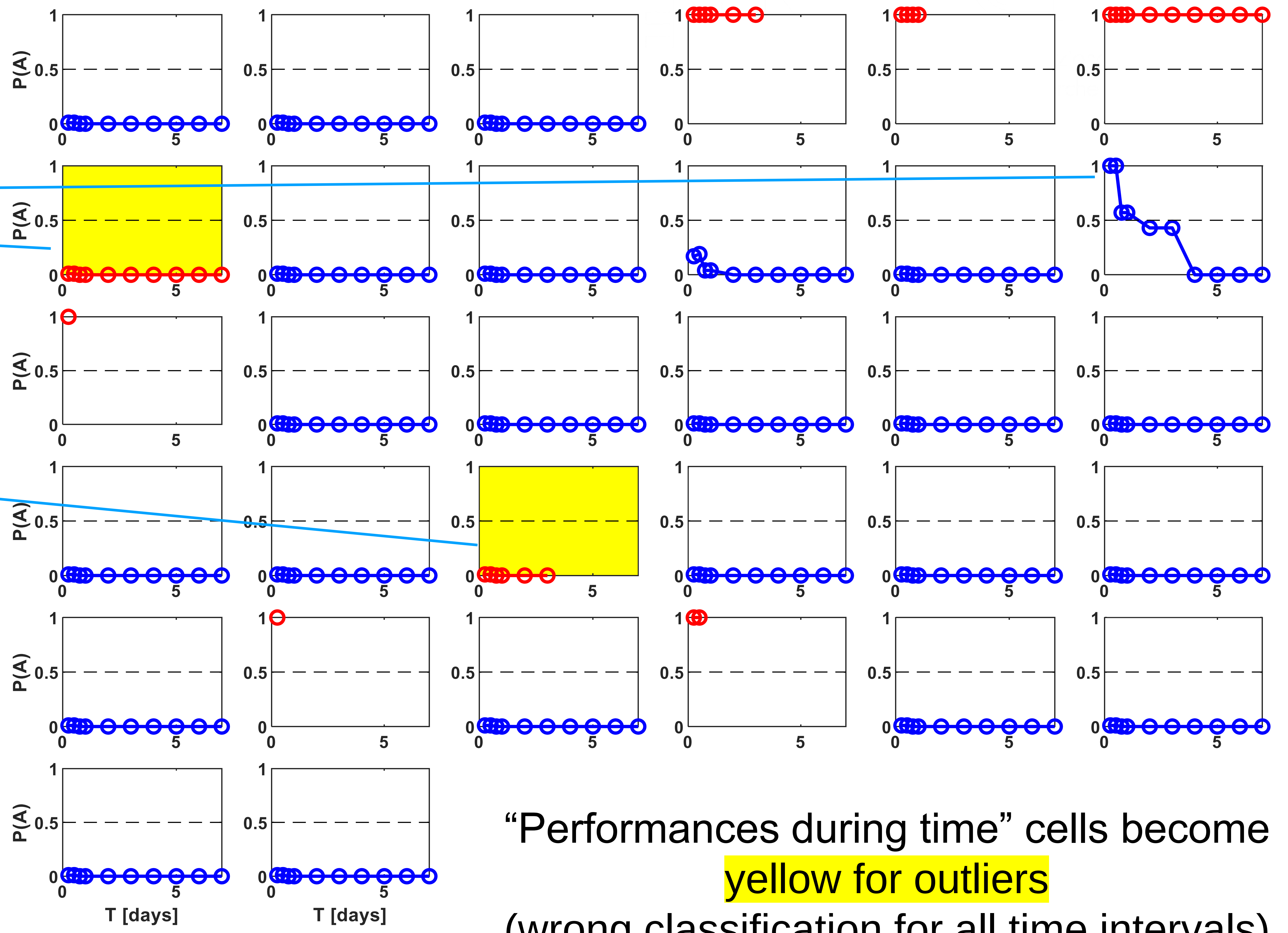
OGS INGV

Ministry of Foreign Affairs and International Cooperation

Funded by a grant from the Italian Ministry of Foreign Affairs and International Cooperation

Self-test on NE Italy - Western Slovenia





Training AND testing: OGS catalogue [1977, 2021]
32 clusters

"Performances during time" cells become yellow for outliers
(wrong classification for all time intervals)

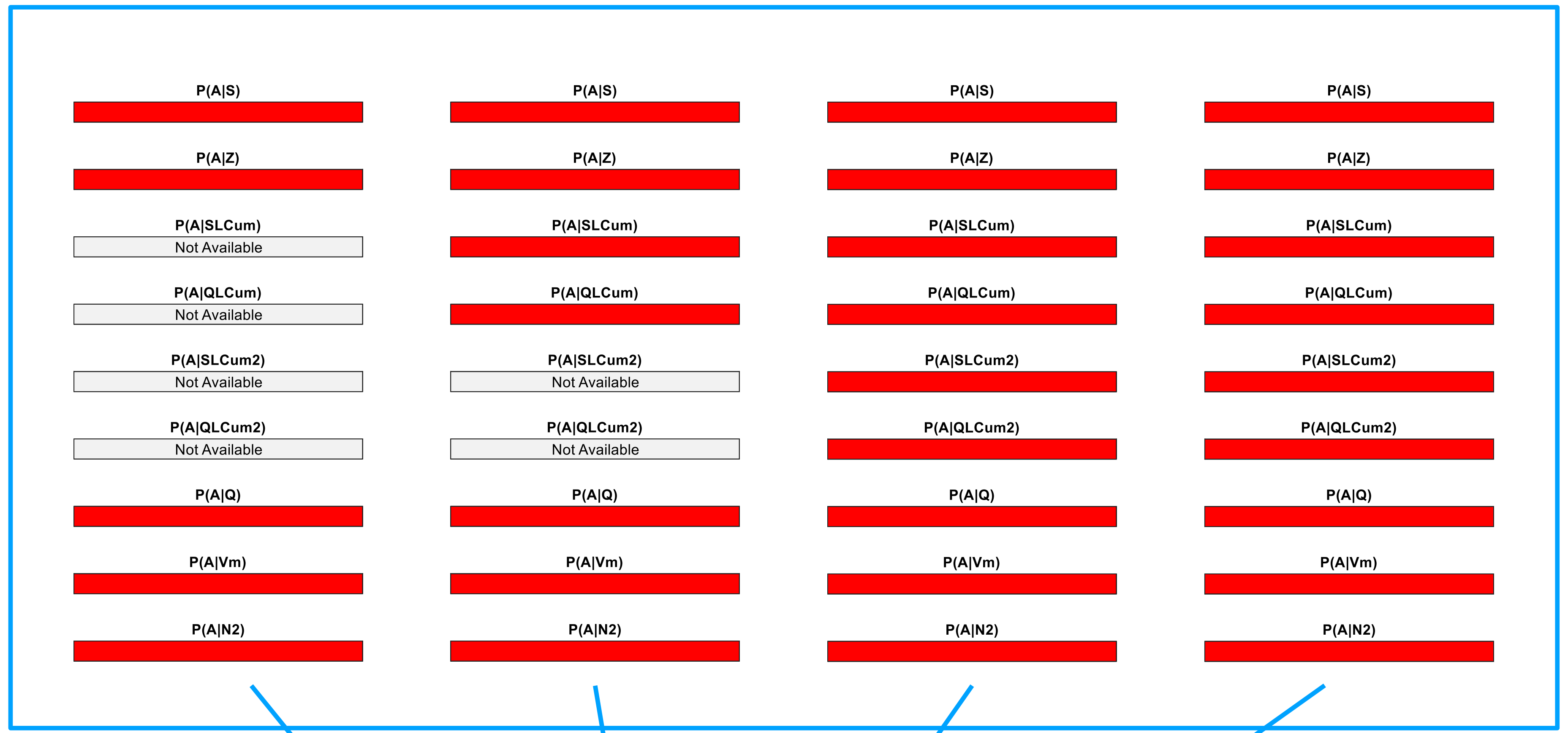
Dots: clusters mainshock location (50 clusters); red: type A clusters; blue: type B clusters.

Ministry of Foreign Affairs and International Cooperation

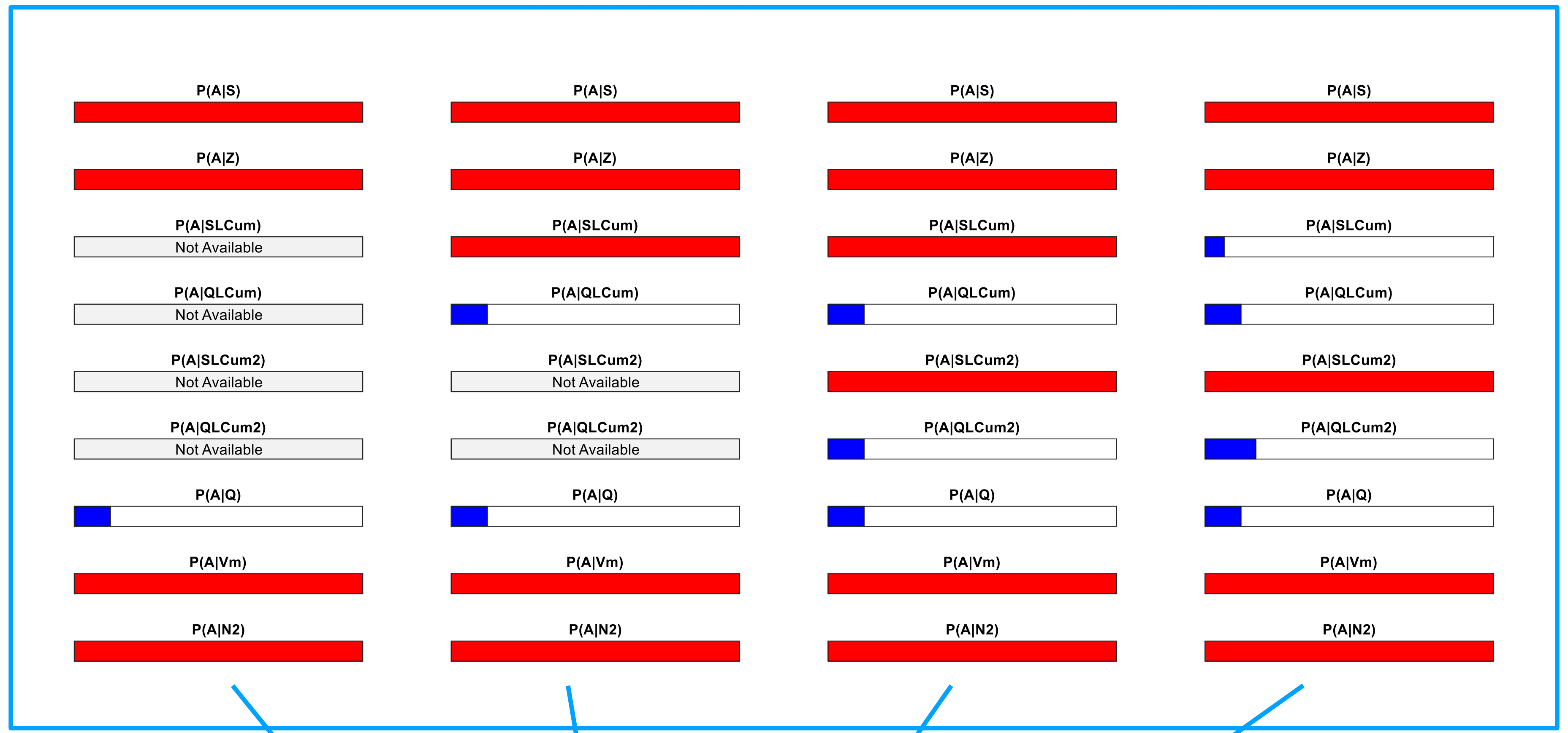
Funded by a grant from the Italian Ministry of Foreign Affairs and International Cooperation

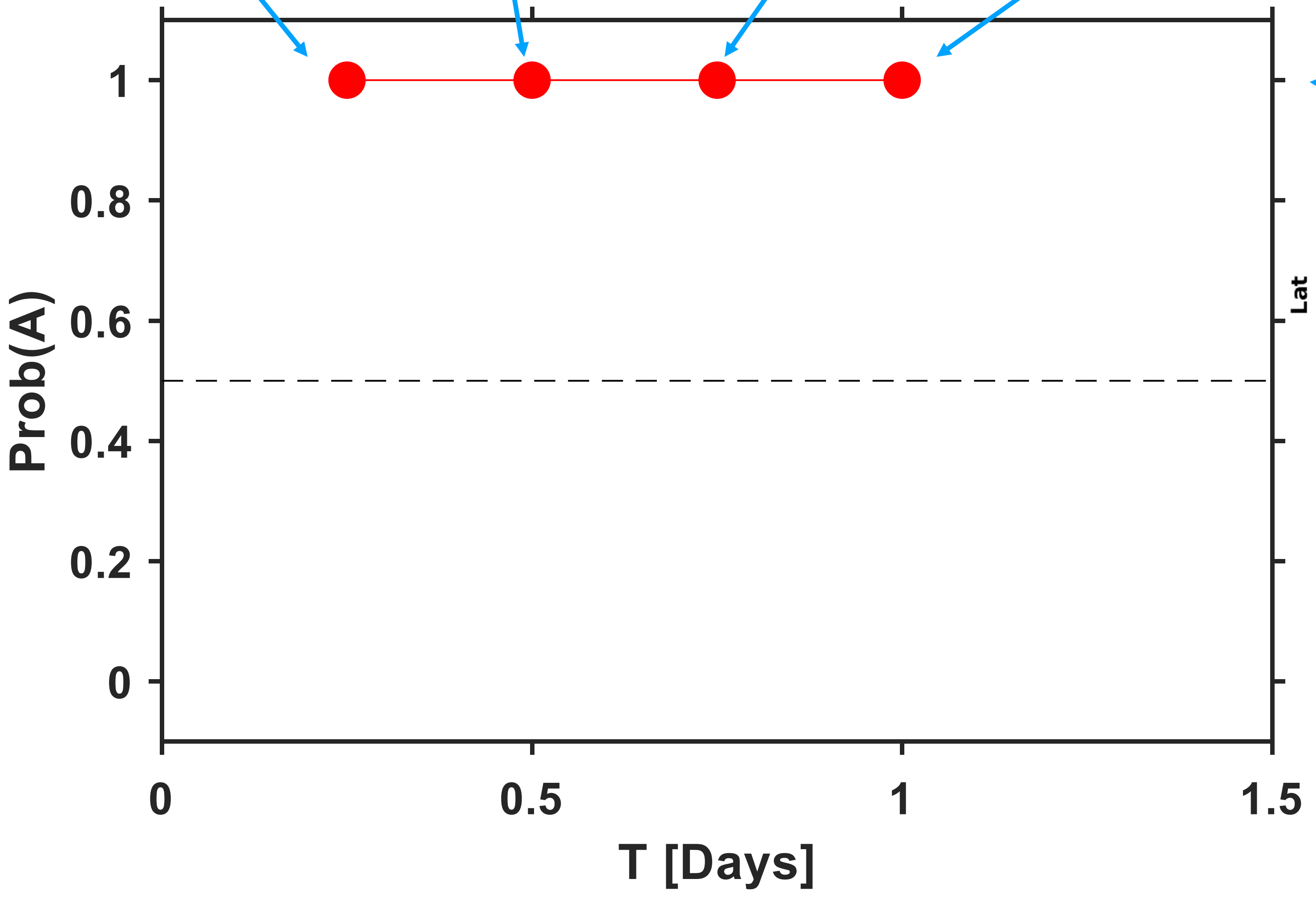
NRT classification Italy

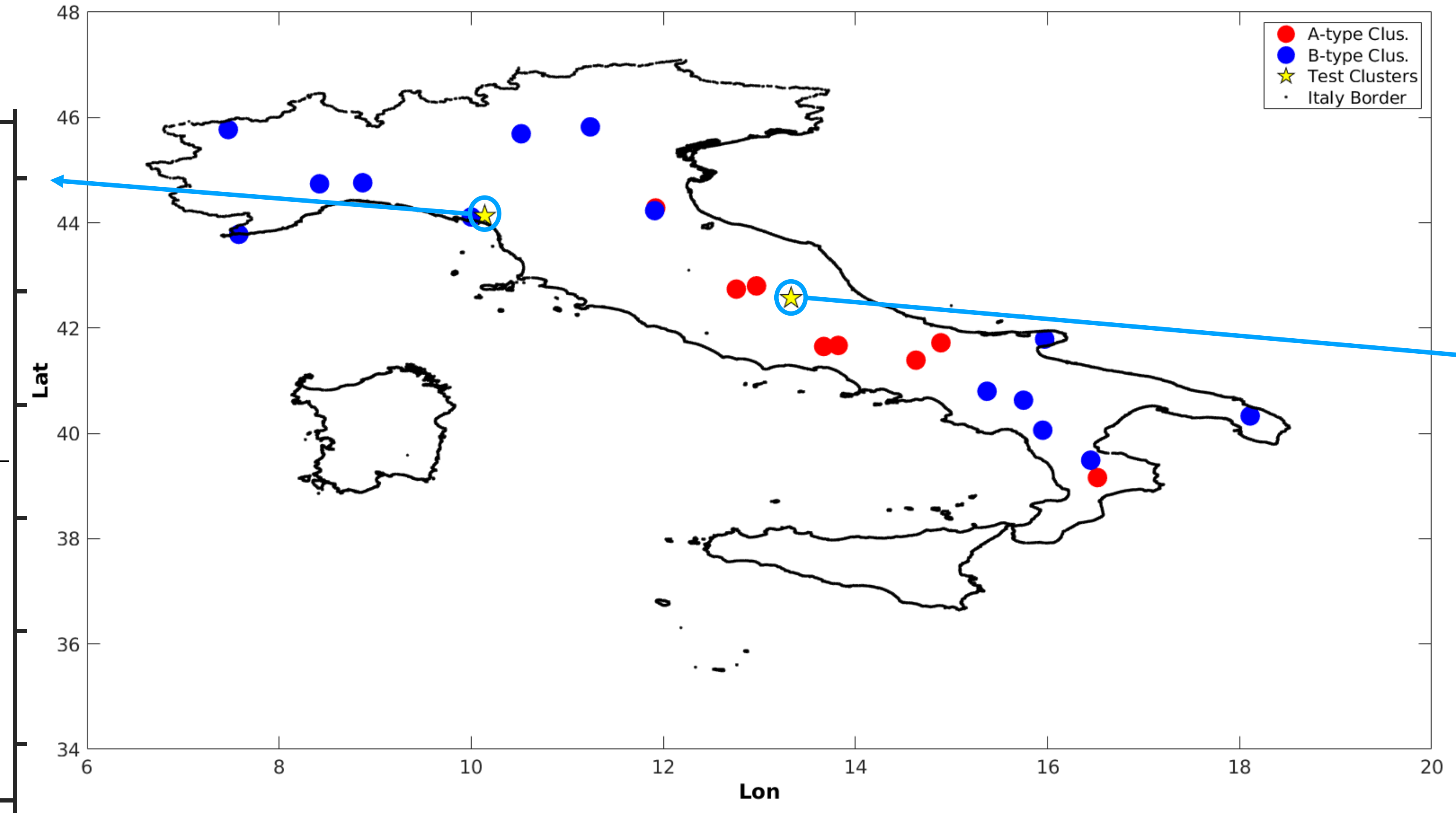
June 21, 2013

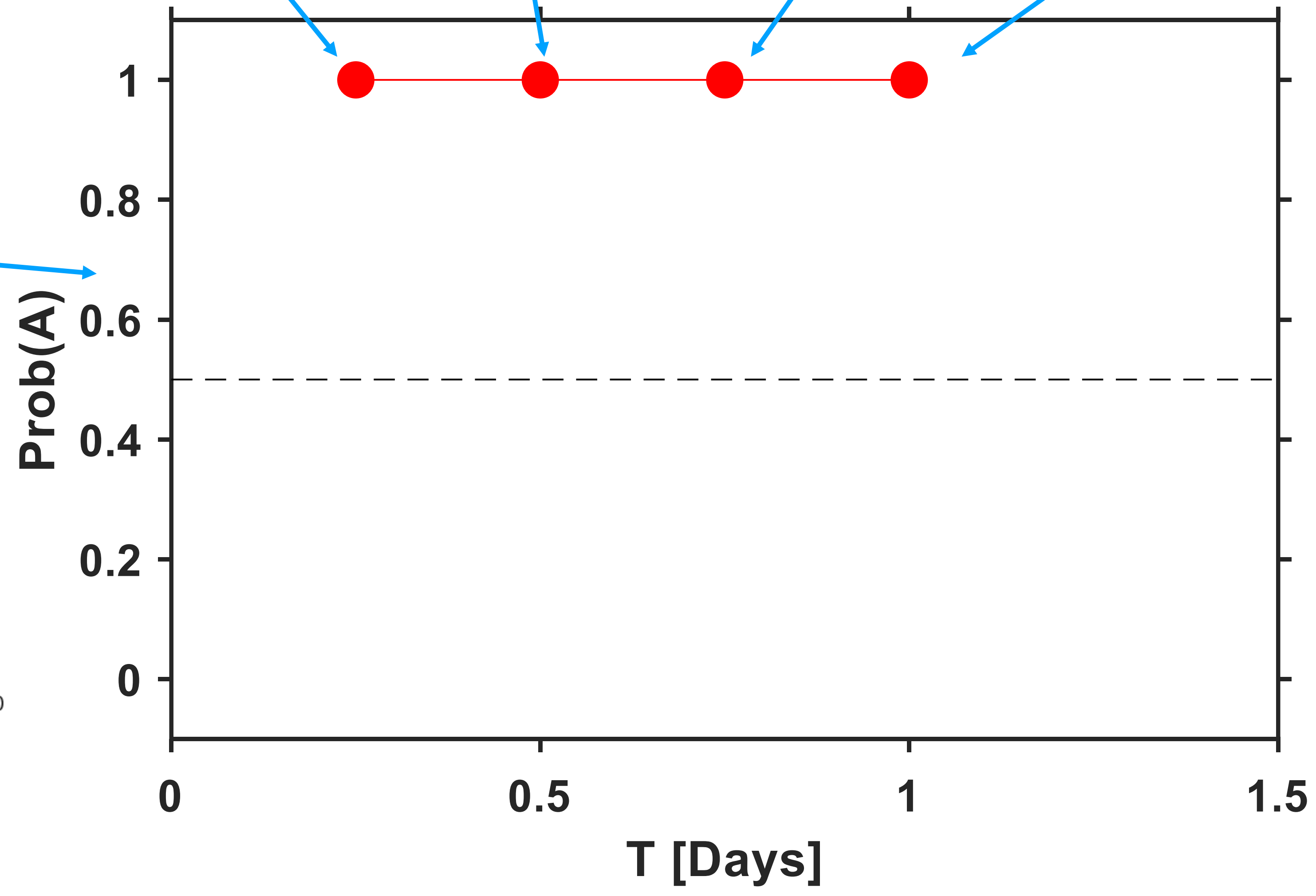


July 22, 2017









Training set:
Lolli & Gasperini catalogue [1980-2004] + ISIDe [2005-2009]

OGS INGV

Ministry of Foreign Affairs and International Cooperation

Funded by a grant from the Italian Ministry of Foreign Affairs and International Cooperation

Funded by a grant from the Italian Ministry
of Foreign Affairs and International Cooperation

Ministry of Foreign Affairs
and International Cooperation

For further details

- Gentili S. and Di Giovambattista R. (2017). Pattern recognition approach to the subsequent event of damaging earthquakes in Italy. Physics of the Earth and Planetary Interiors, 266, 1-17. https://www.researchgate.net/publication/314126252_Pattern_recognition_approach_to_the_subsequent_event_of_damaging_earthquakes_in_Italy

- Gentili S. and Di Giovambattista R. (2020). Forecasting strong aftershocks in earthquake clusters from northeastern Italy and western Slovenia. Physics of the Earth and Planetary Interiors, 303, 106483. https://arxiv.org/ftp/arxiv/papers/2005/2005.02779.pdf

- Gentili S. and Di Giovambattista R. (2022). Forecasting strong subsequent earthquakes in California clusters by machine learning. Physics of the Earth and Planetary Interiors, 327, 106879. https://www.sciencedirect.com/science/article/pii/S0031920122000401

Ministry of Foreign Affairs and International Cooperation

Funded by a grant from the Italian Ministry
of Foreign Affairs and International Cooperation

Confusion matrix

True class

		p	n
Hp class	Y	TP	FP
	N	FN	TN
totals		P	N

Confusion matrix

$$Recall = TPR = \frac{TP}{P}$$

$$Precision = \frac{TP}{Y}$$

$$Accuracy = \frac{TP + TN}{P + N}$$

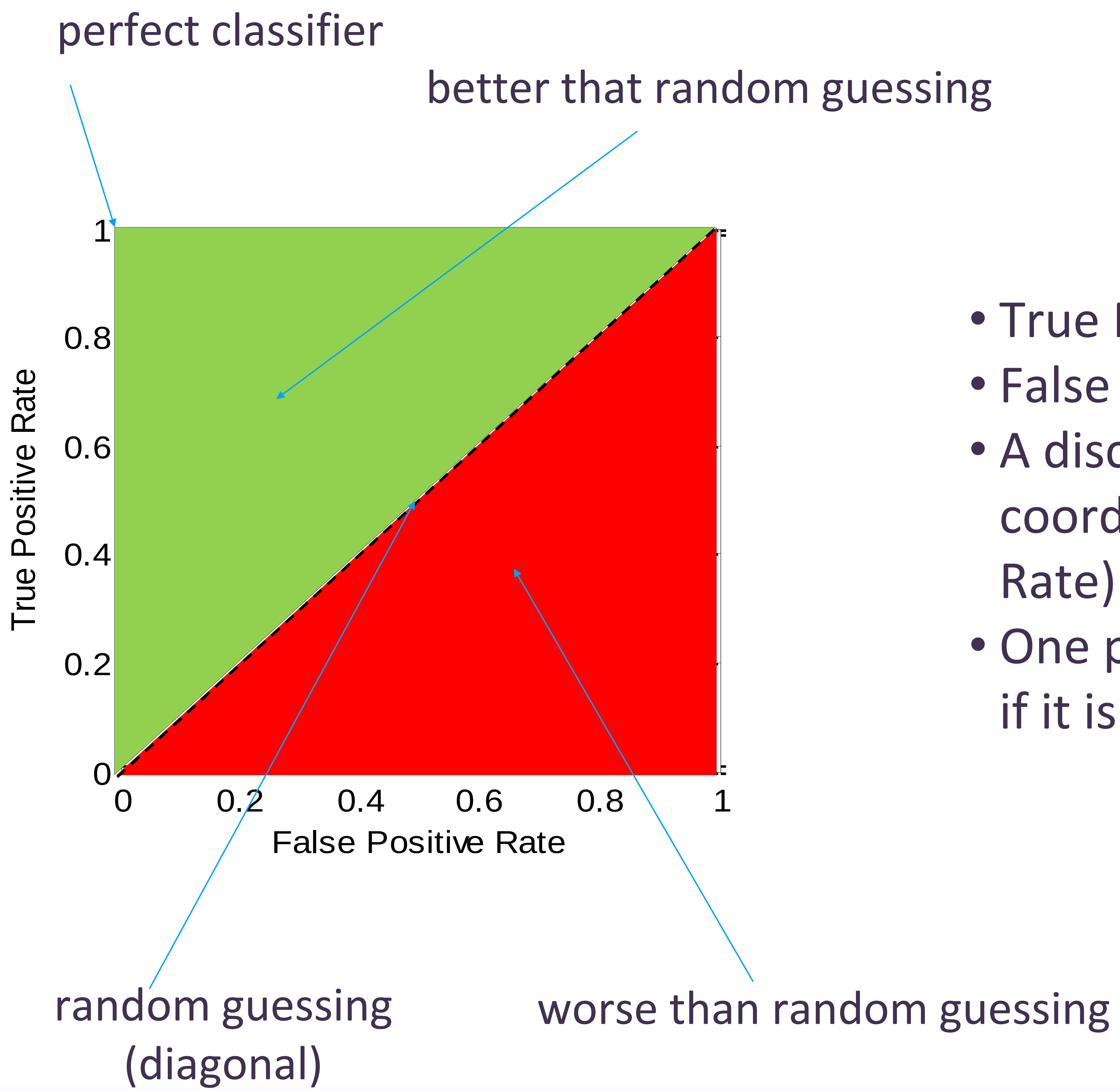
$$FPR = \frac{FP}{N}$$

$$Informedness = TPR - FPR$$

Ministry of Foreign Affairs and International Cooperation

Funded by a grant from the Italian Ministry of Foreign Affairs and International Cooperation

ROC



- True Positive Rate= True(A) / All(A)
- False Positive Rate= False(B)/ All(B)
- A discrete classifier produces a point of coordinates (False Positive Rate, True Positive Rate) in the ROC graph.
- One point in ROC space is better than another if it is closer to the point (0,1)

OGS INGV

Ministry of Foreign Affairs and International Cooperation

Funded by a grant from the Italian Ministry of Foreign Affairs and International Cooperation

Classifiers combination

- For each T_j, only the feature selected during the training set are extracted
- The thresholds are compared with the value of the feature to estimate the probability pi that the cluster is A for each feature
- The values of p_i are combined by using a Bayesian approach

$$P(A|D_1 \dots D_N) = \frac{[N(B)]^{N-1} \prod_{n=1}^{N} p_n}{[N(B)]^{N-1} \prod_{n=1}^{N} p_n + [N(A)]^{N-1} \prod_{n=1}^{N} (1 - p_n)}$$

OGS INGV

Ministry of Foreign Affairs and International Cooperation

Funded by a grant from the Italian Ministry of Foreign Affairs and International Cooperation