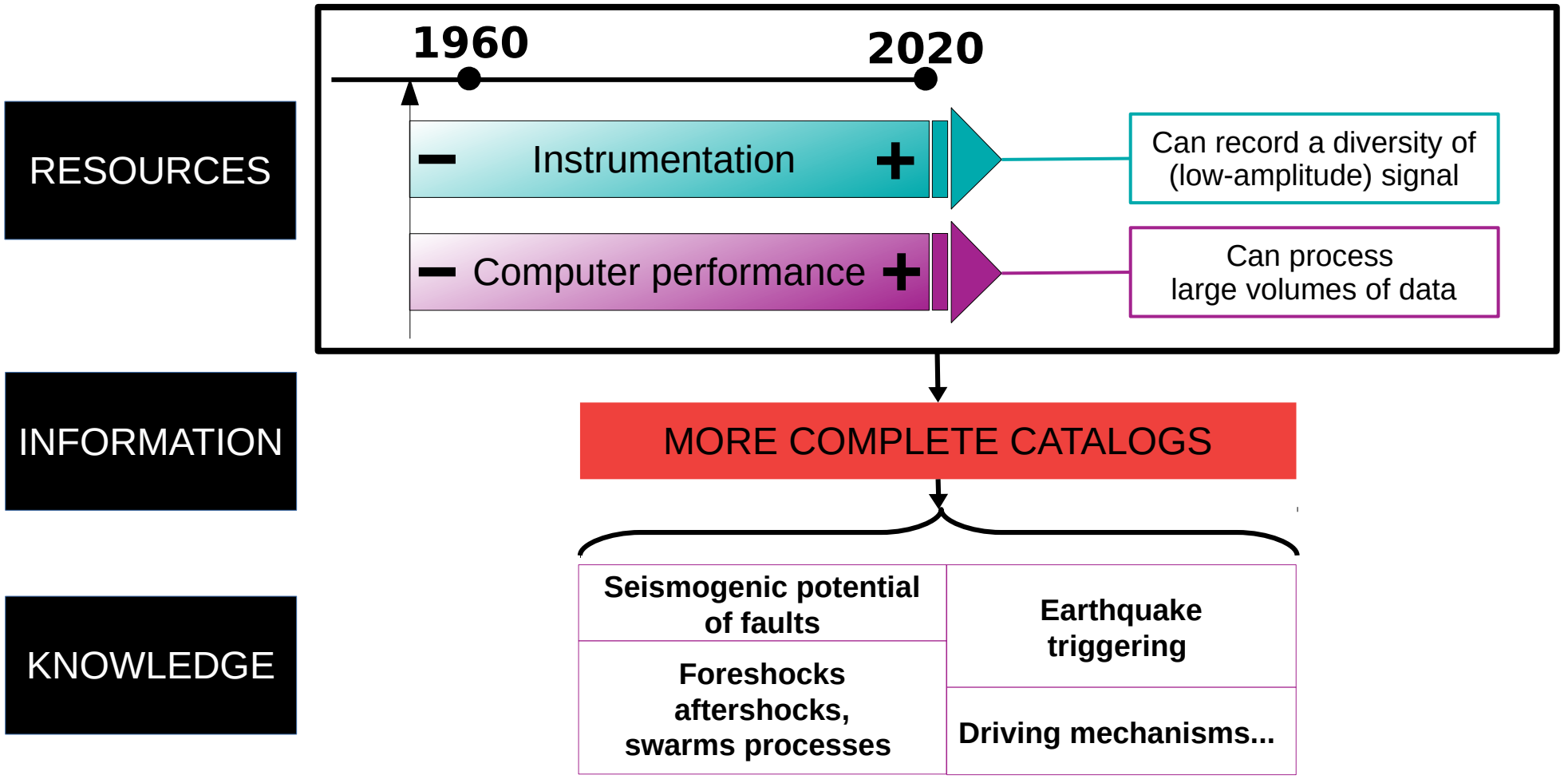
Apport de l'intelligence artificielle pour la détection automatique des séismes en contexte anthropogénique

A.Renouard, M. Grunberg, C. Doubre, A. Maggi, C. Hibert

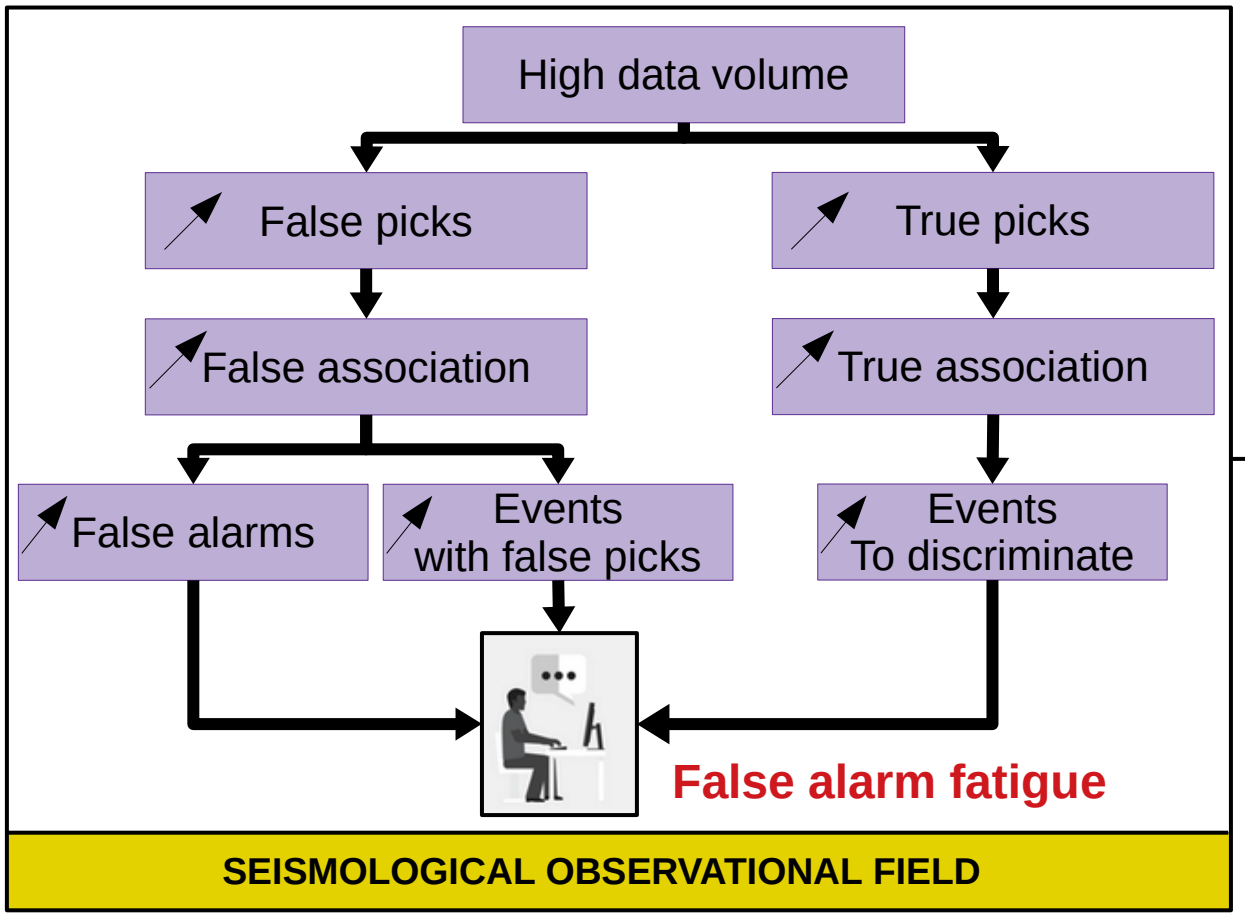
AlpArray-FR

Seismology is a data-driven science



But...

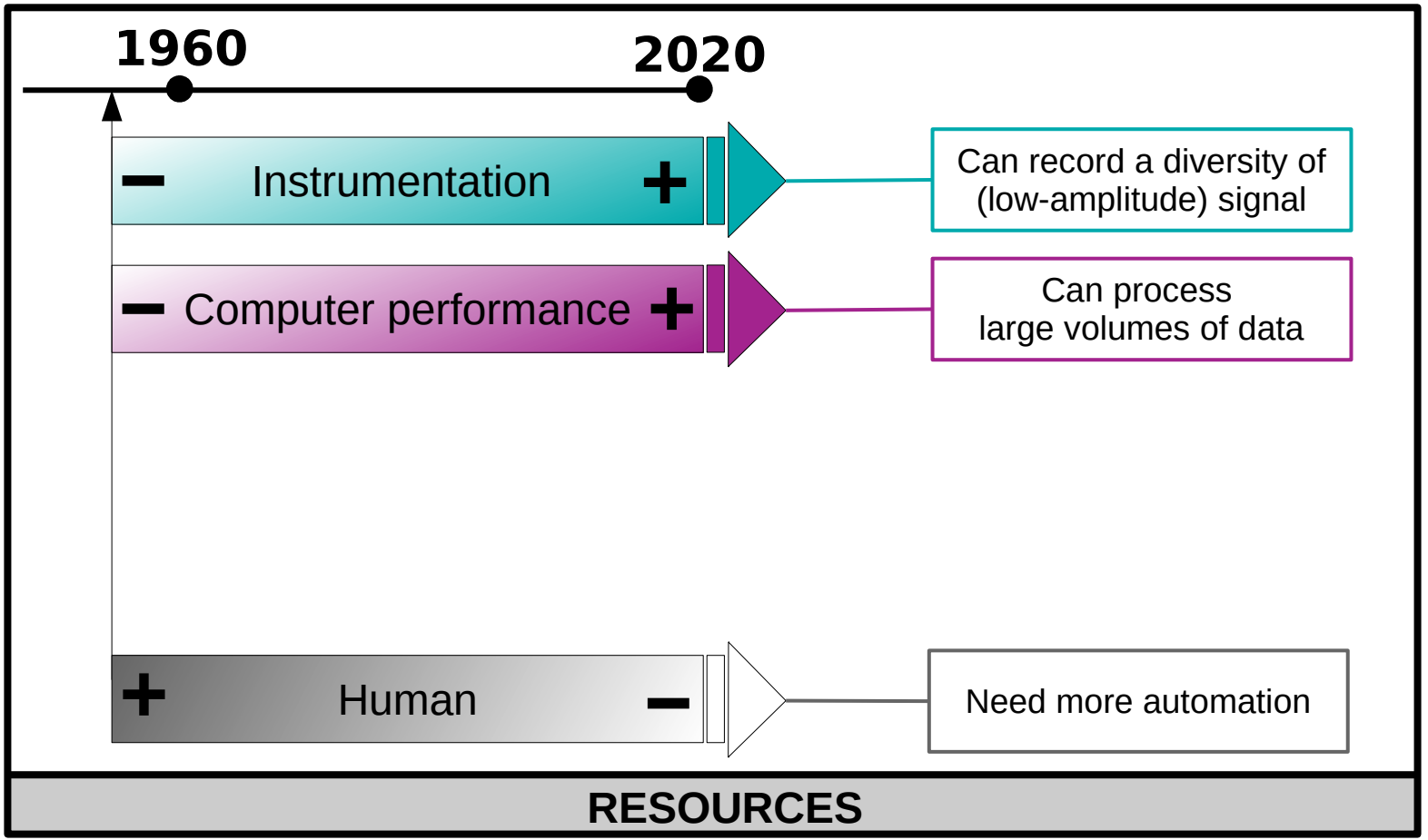
Expected improvements quite limited in seismological observatories



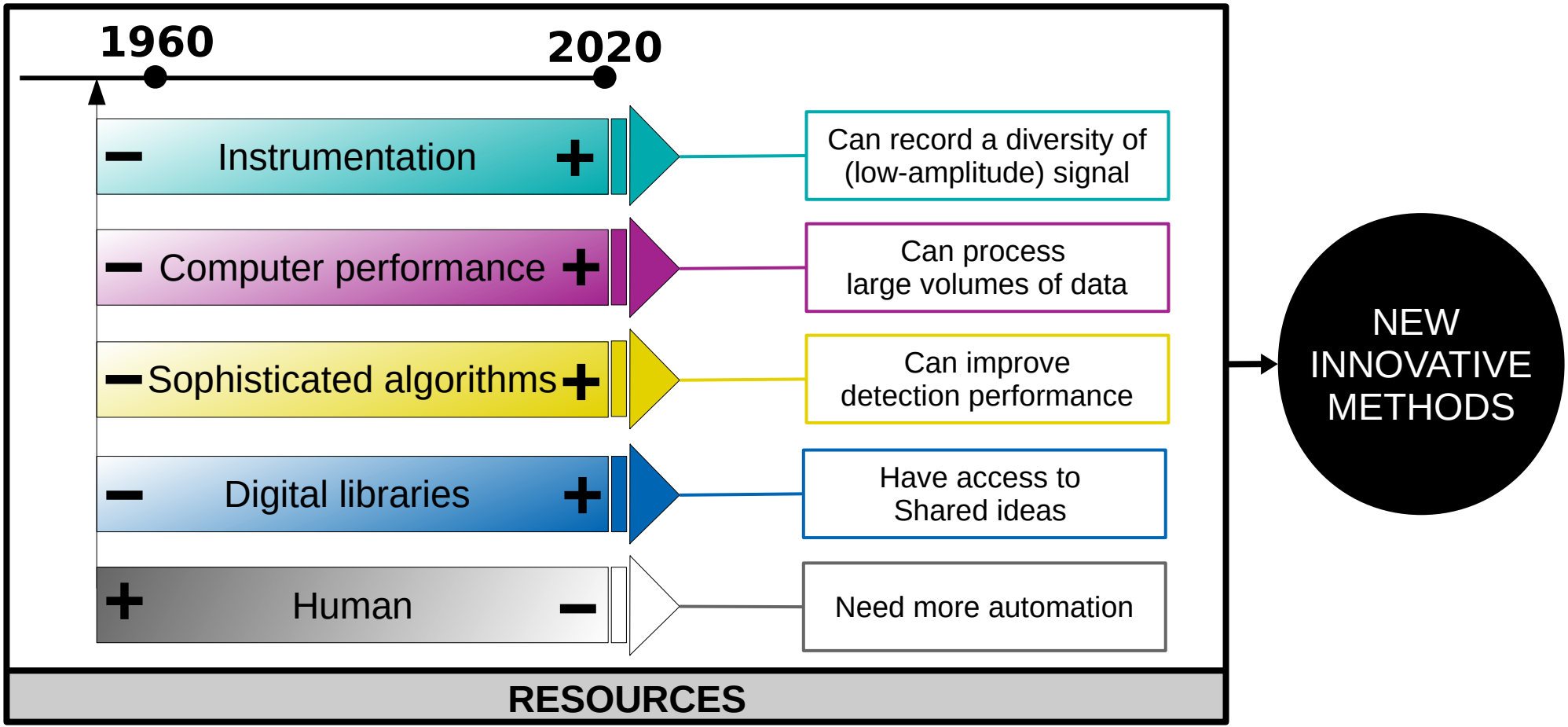
Referent catalog with uncovered events

How to improve the routine Detection procedure?

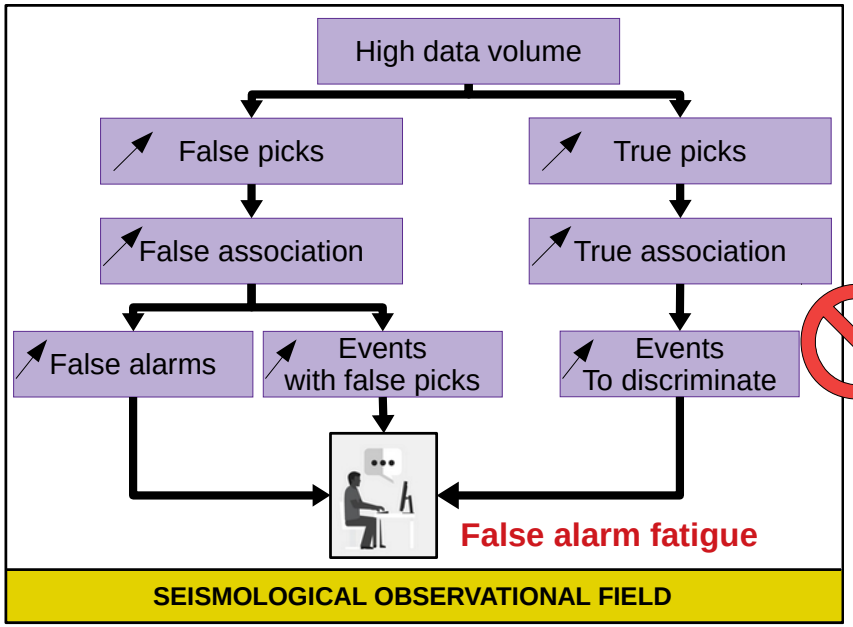
Increase the number of analysts ?



Leverage new innovative methods ?



Leverage new innovative methods ?



Algorithm incompatibility
High computational cost
Do not use STA/LTA

ConvNetQuake_INGV (Lomax et al., 2019)

CPIC (Zhu et al., 2019)

DeepPick (Dickley et al., 2019)

PhaseLink (Ross et al., 2019)

ConvNetQuake (Pérol et al., 2018)

DeepDetect (Wu et al., 2018)

CRED (Mousavi et al., 2018)

PhaseNet (Zhu et al., 2018)

FAST (Yoon et al., 2015)

CONEDEP (Zhou et al., 2017)

...

Design of a New Scalable Detection System

Minimum computational cost

Real-time processing

Large datasets

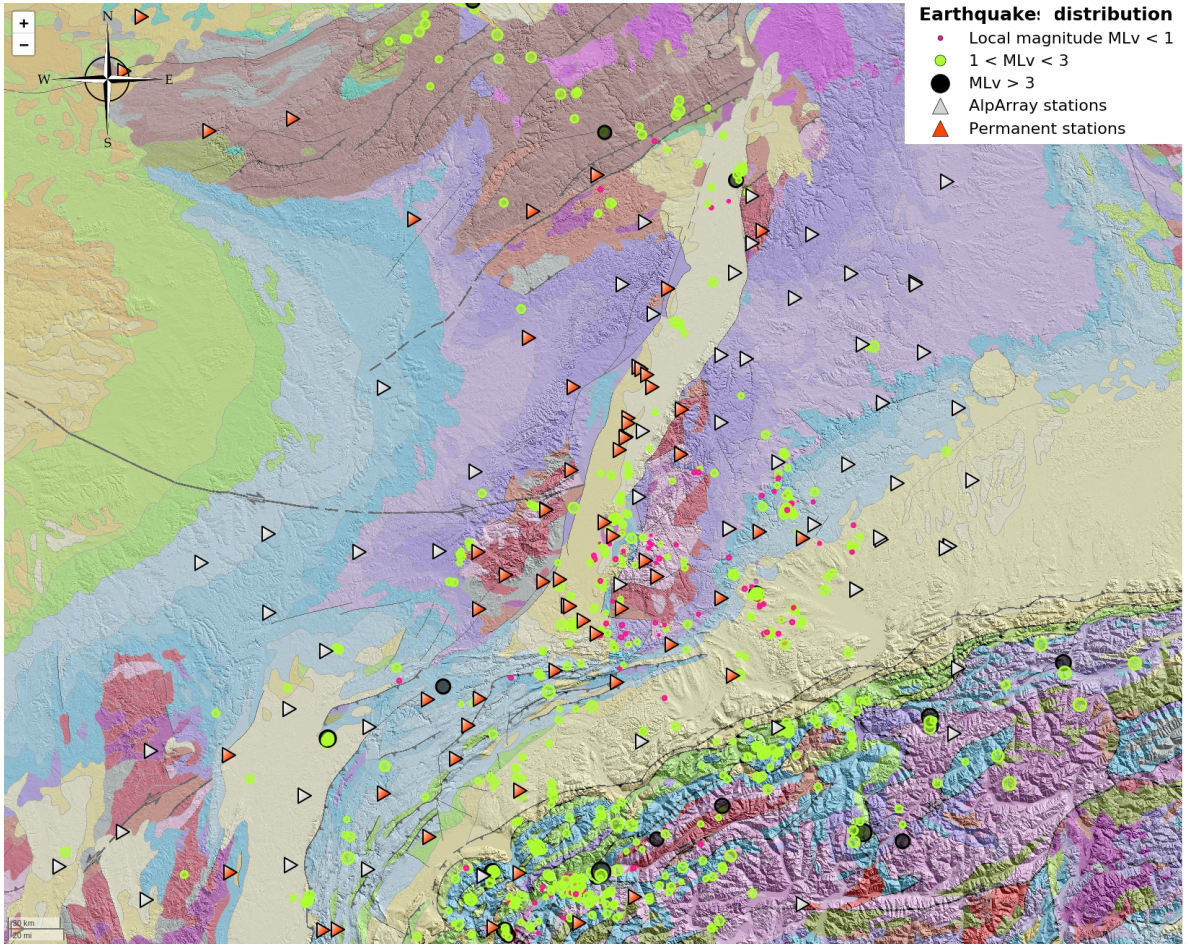
+

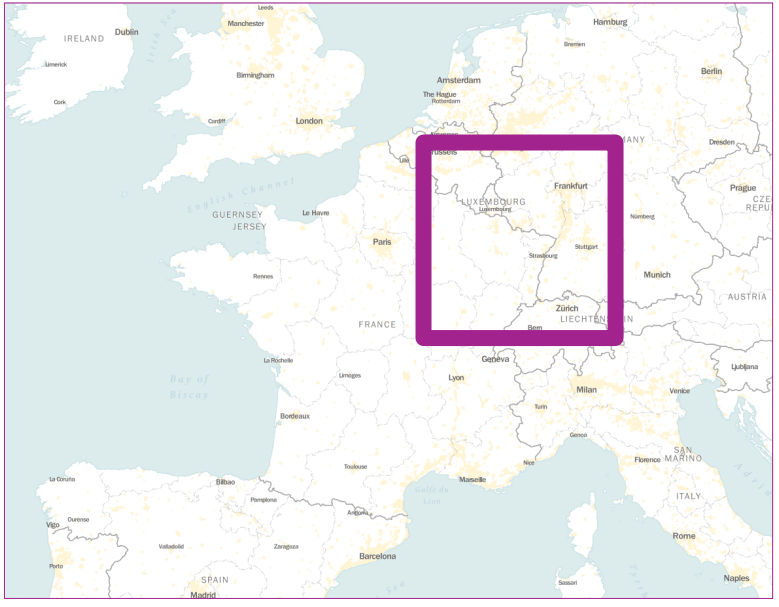
Low detection threshold

High event discriminability

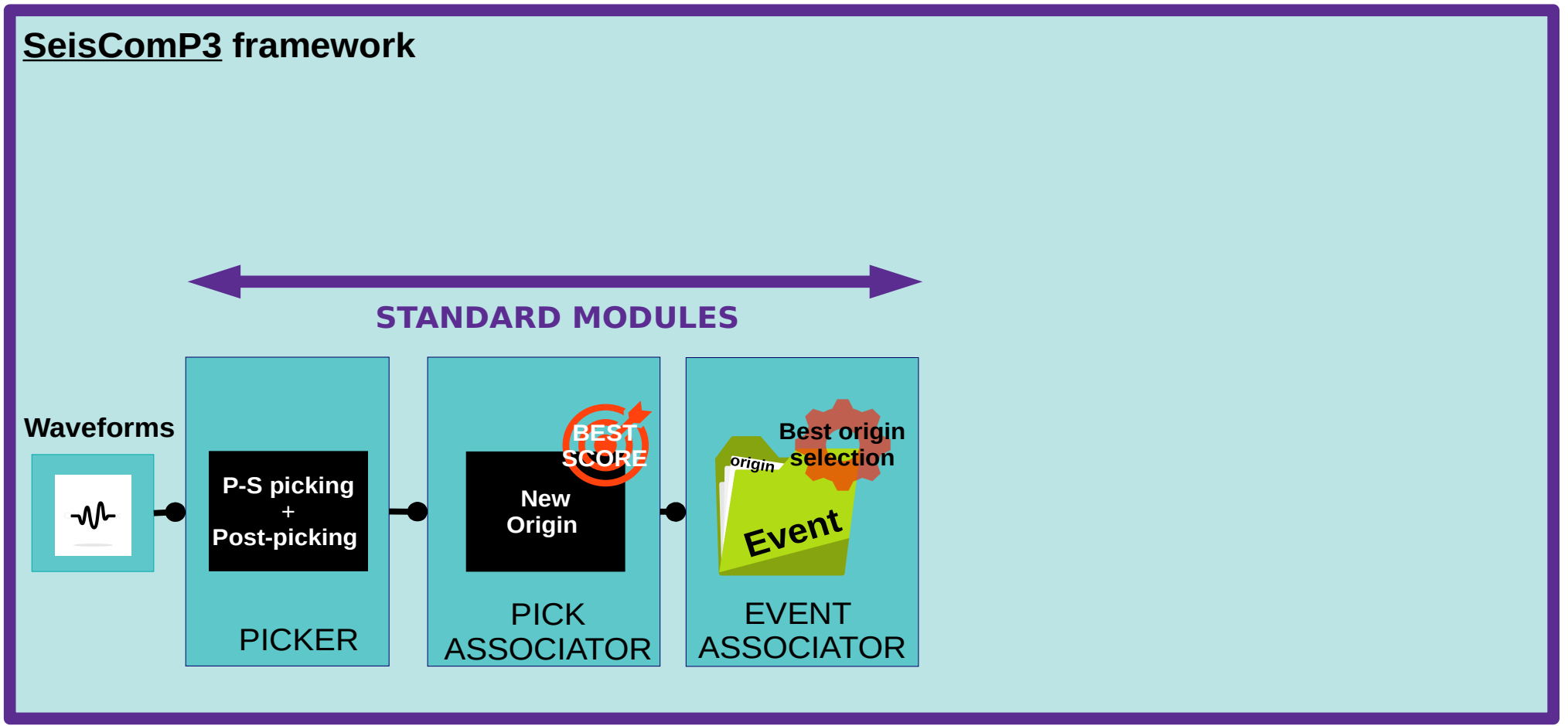
Low rates of false alarms

Testing Region and Time Period : the Rhine Graben Area between 2016-2019

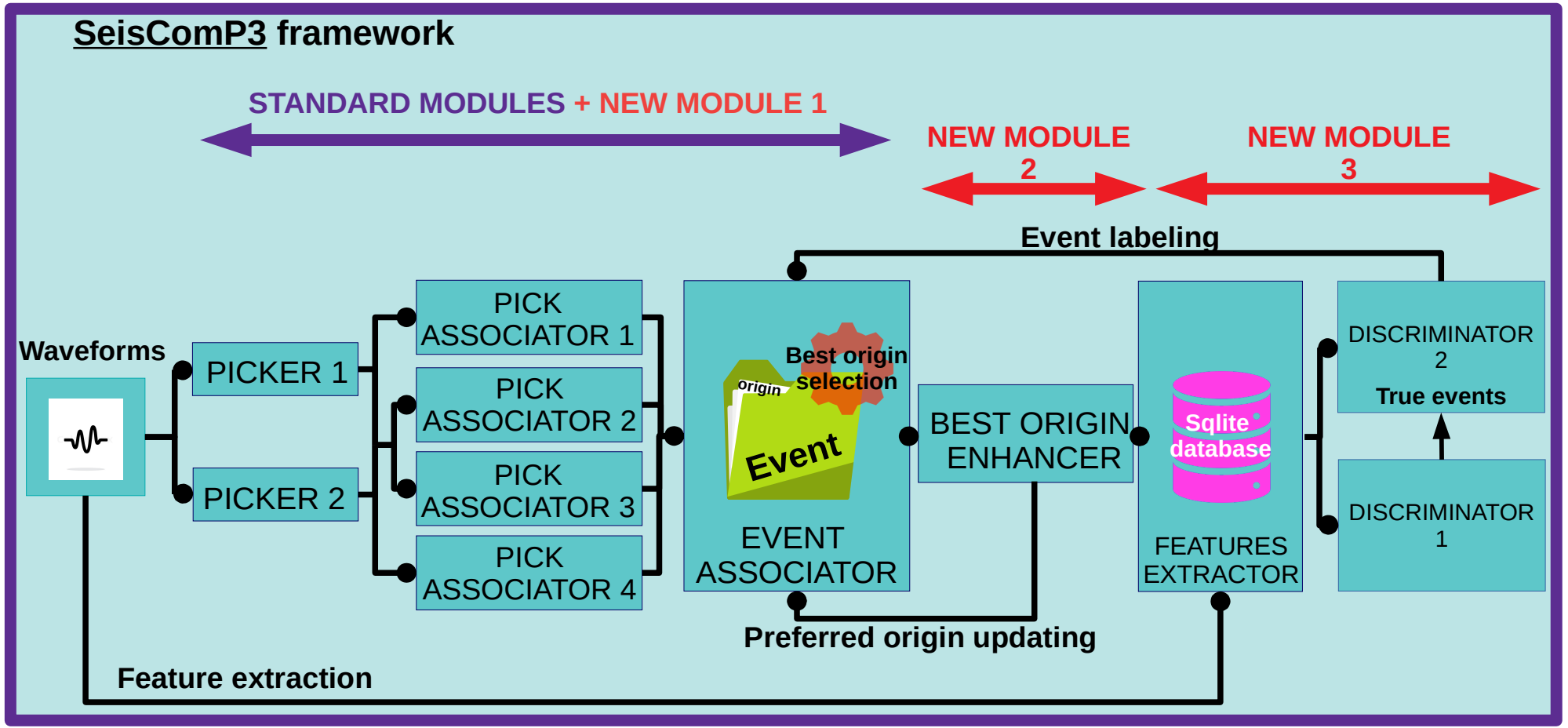




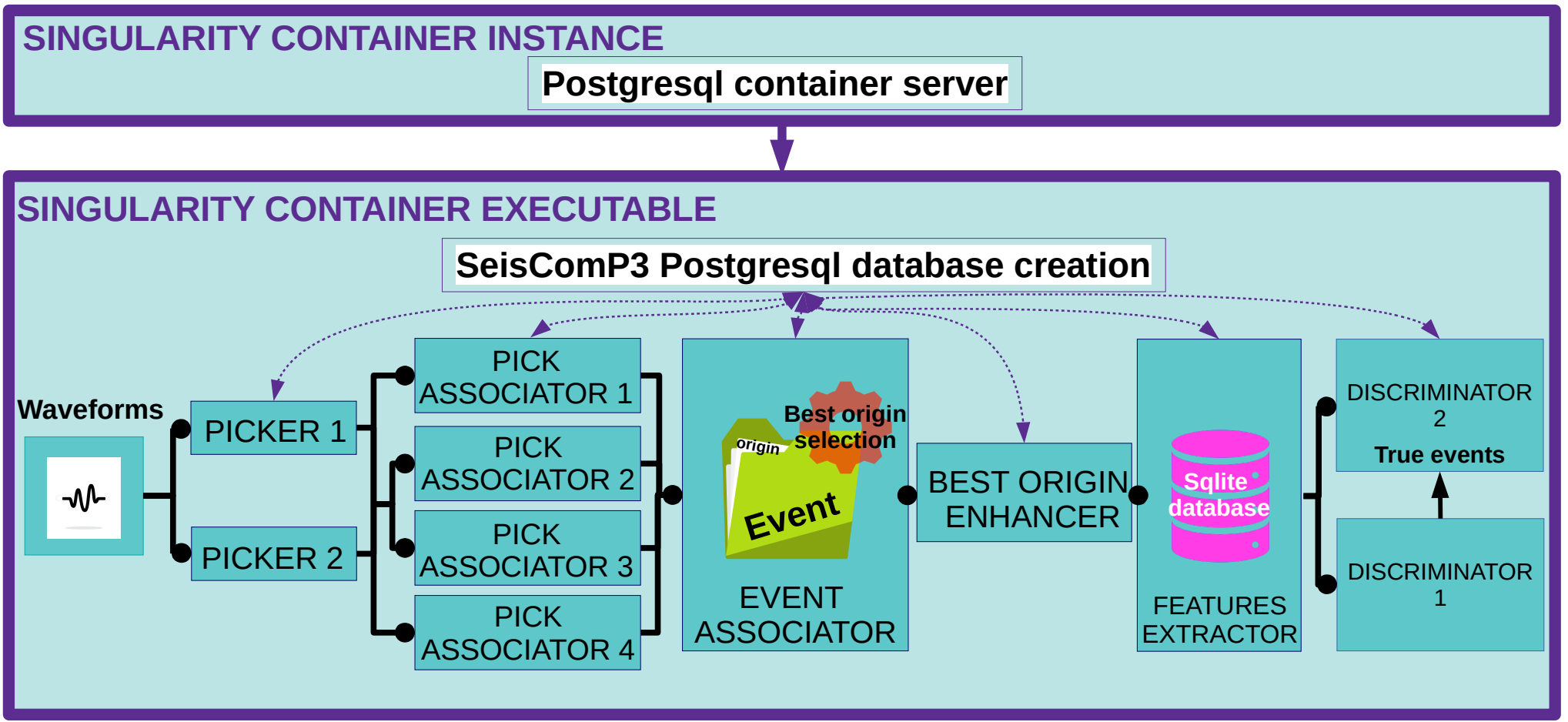
SeisComP3 Classical Detection system



SeisComP3 New Scalable Detection System



SeisComP3 New Scalable Detection System



Random Forest, a supervised learning algorithm

Input data : labeled raw dataset

LABELS

Y1 : [1]
Y2 : [0]
Y3 : [2]
Yn : [0]

0 = false alarm
1 = quarry blast
2 = earthquake

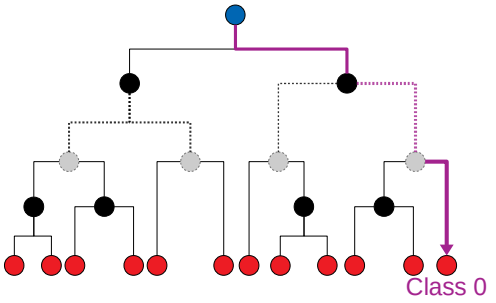
DATA

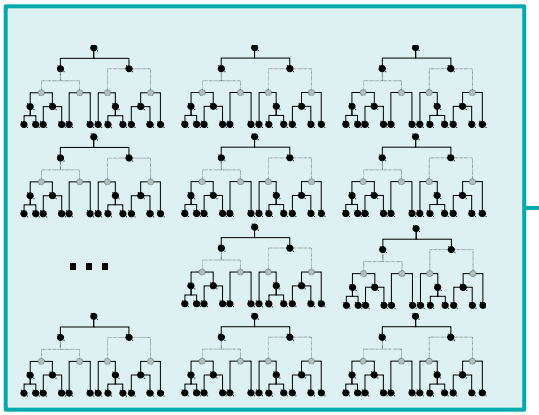
X1 : [[7],[0.09],[1.57],[0.85]]
X2 : [[6],[0.22],[2.82],[3.99]]
X3 : [[16],[0.19], [1.42], [0.37]]
Xn : [[5], [0.18], [3.04],[7.62]]

Feature 1 **Feature 2** **Feature 3** **Feature n**

Statistical features : waveforms, spectrum, spectrogram, IMFs, network geometry, magnitudes, event parameters, polarities...

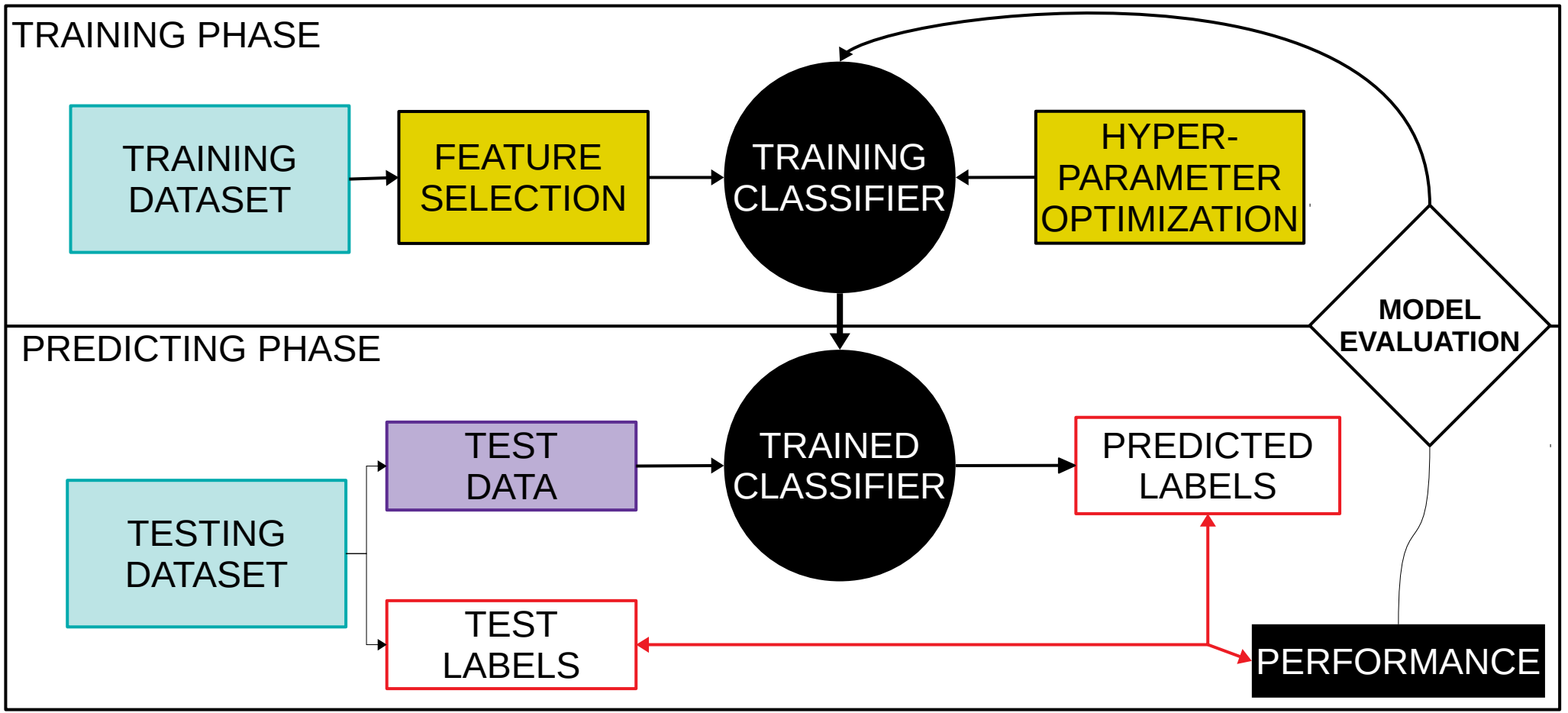
Random Forest algorithm : ensemble of decision trees



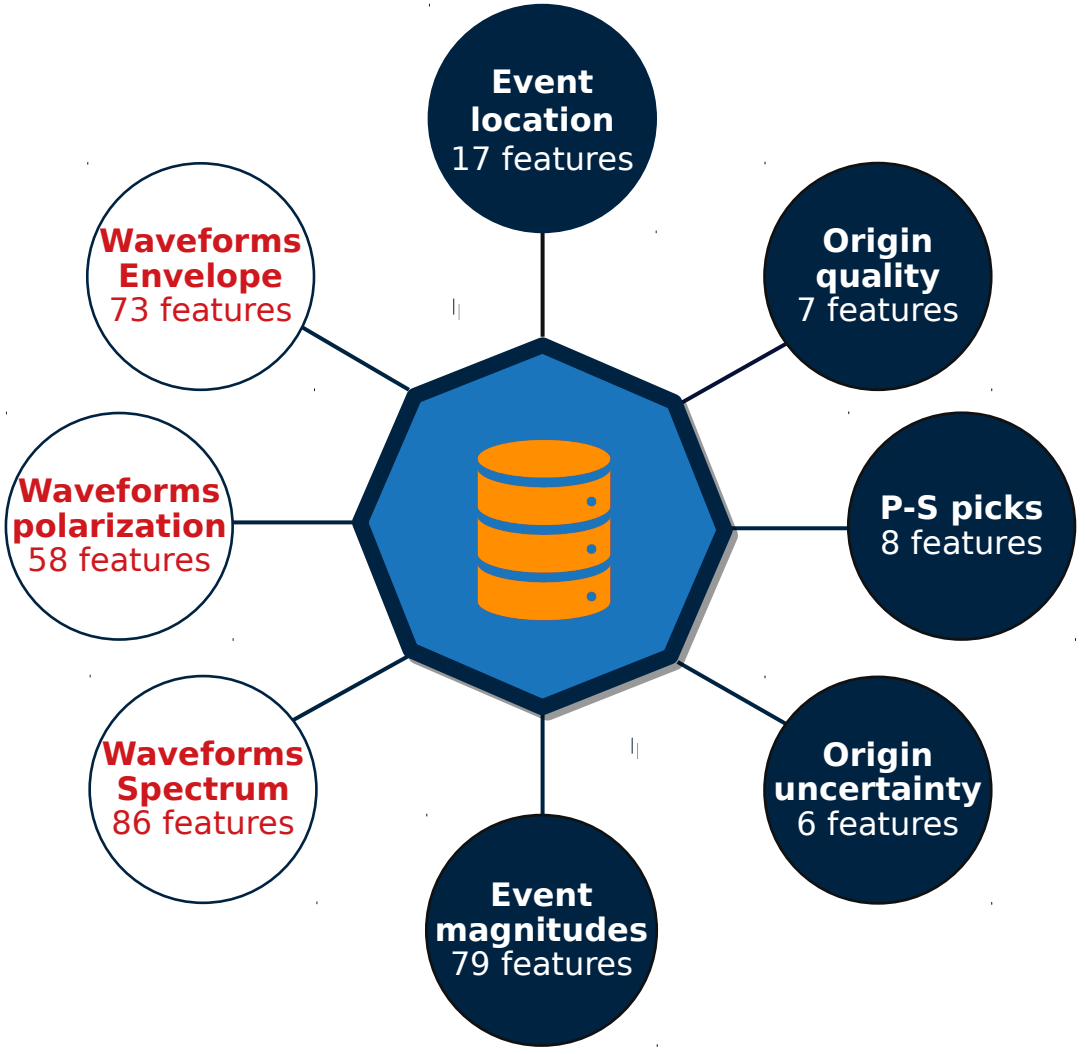


Majority voting

Random Forest, a supervised learning algorithm



Type of features computed in the Sqlite database



Results for Fake vs True Event Discrimination

Training model performance
(4 years 2015-19)
Results from Random Forest Algorithm

EXPECTED \ PREDICTED	False alarms	True detections
False alarms	915	13
True detections	12	2818

Accuracy : 99.33 +/- 0.01 %

Predicted results on new data
(4 months of automatic catalog
July-October 2016)

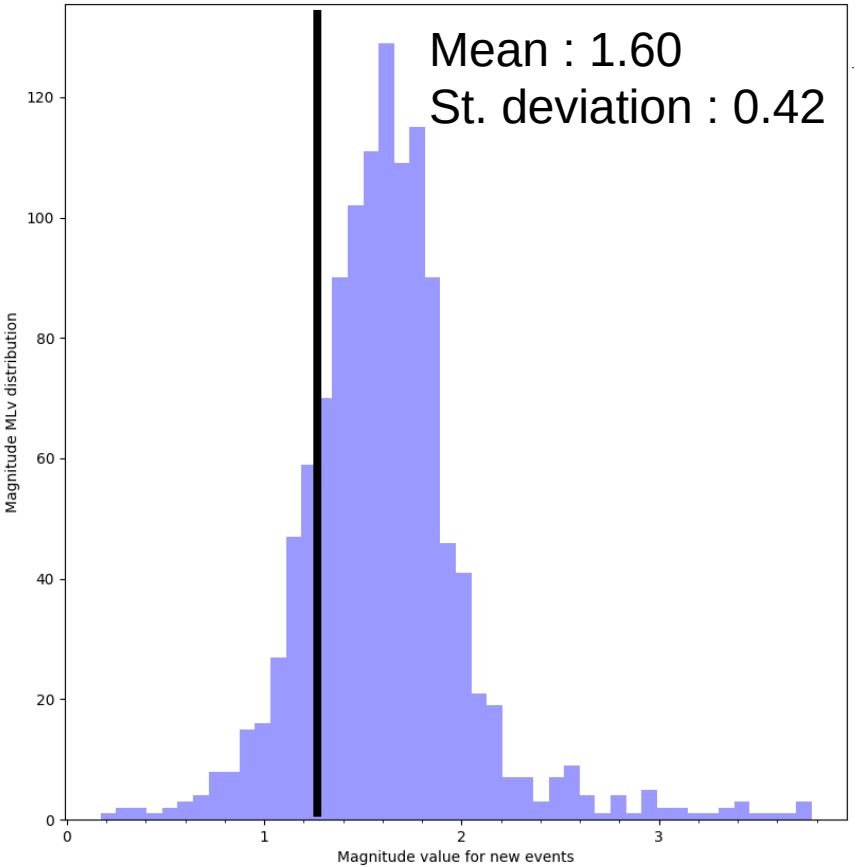
EXPECTED \ PREDICTED	False alarms	True detections
False alarms	49617	418
True detections	106	2170

Accuracy : 97.58% (Manual verification)

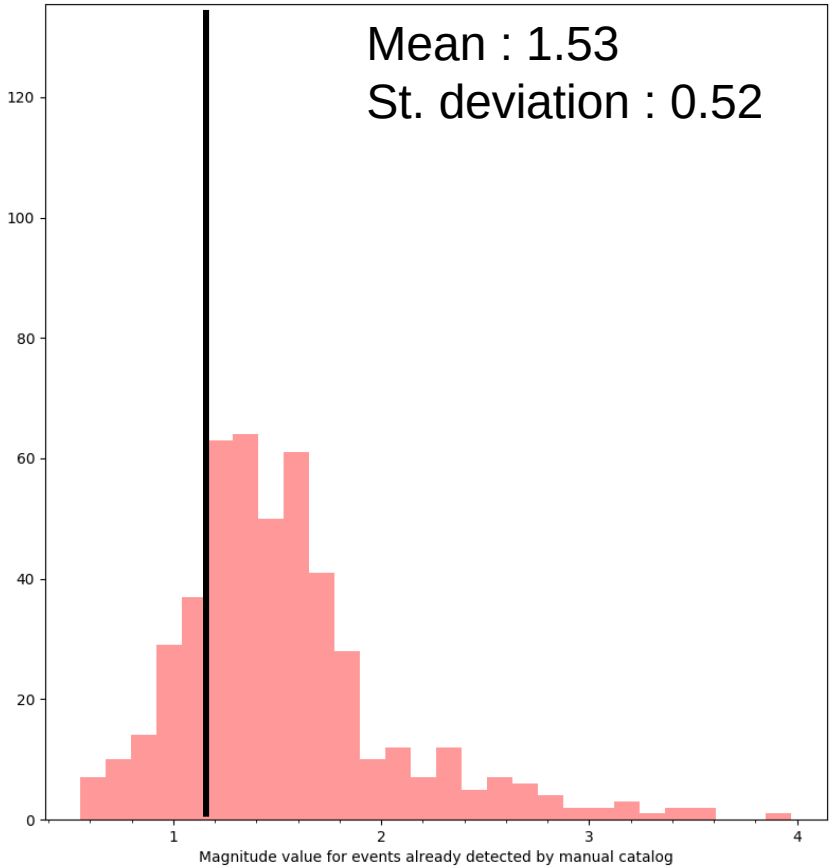
Magnitude distribution

Magnitude Mlv distribution

New events : 1203



Events already detected by BCSF-RéNaSSS : 480



Results for Natural vs Anthropogenic Event Discrimination

Training model performance
(4 years 2015-2019)

EXPECTED \ PREDICTED	Earthquakes	Quarry blasts
Earthquakes	2521	64
Quarry blasts	72	1740

Accuracy : 97.00 +/- 0.01 %

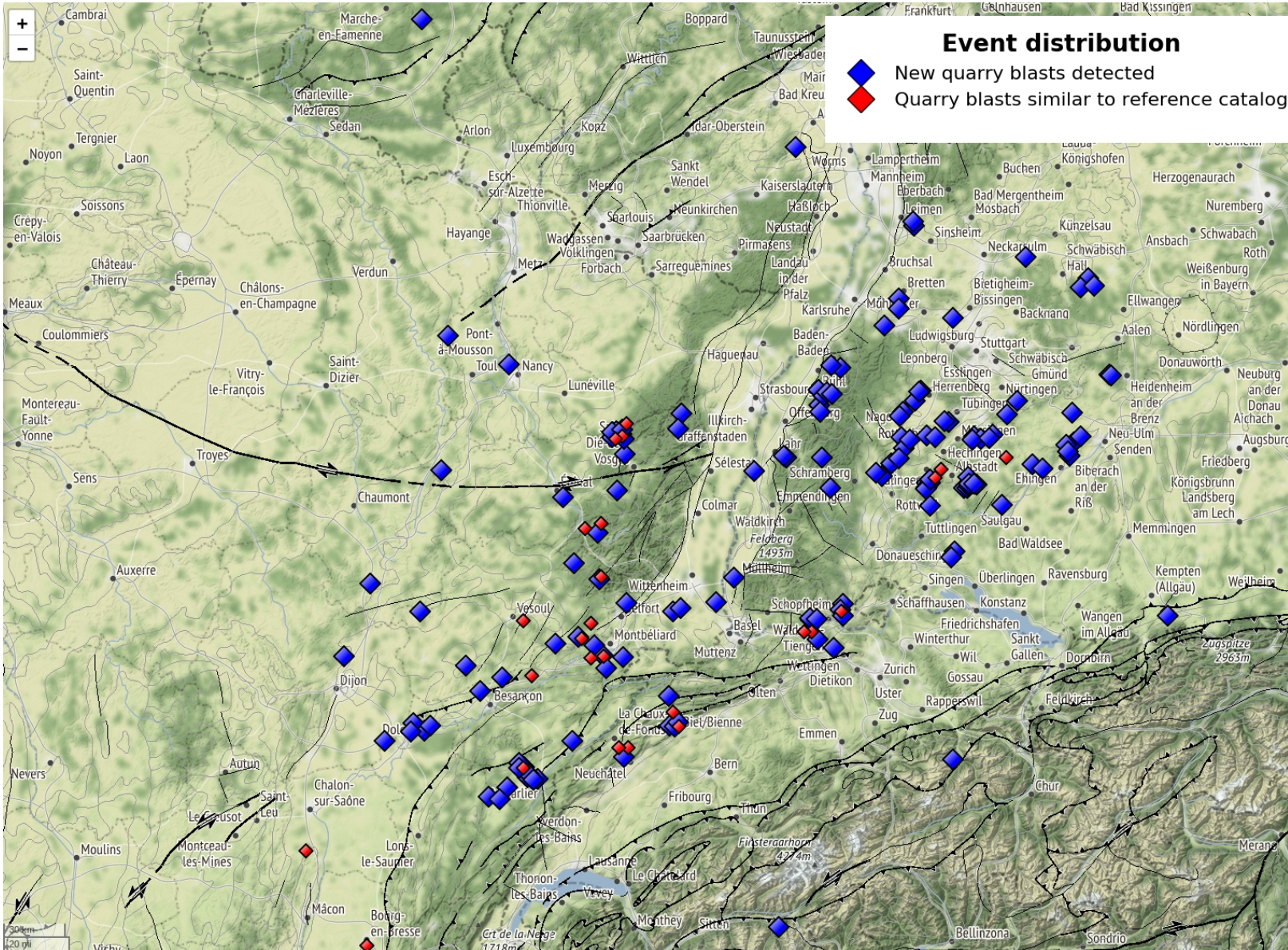
Predicted results on new data
(1 month of automatic catalog July 2016)

EXPECTED \ PREDICTED	Earthquakes	Quarry blasts
Earthquakes	122	16
Quarry blasts	13	161

Accuracy : 90.70 % (manual verification)

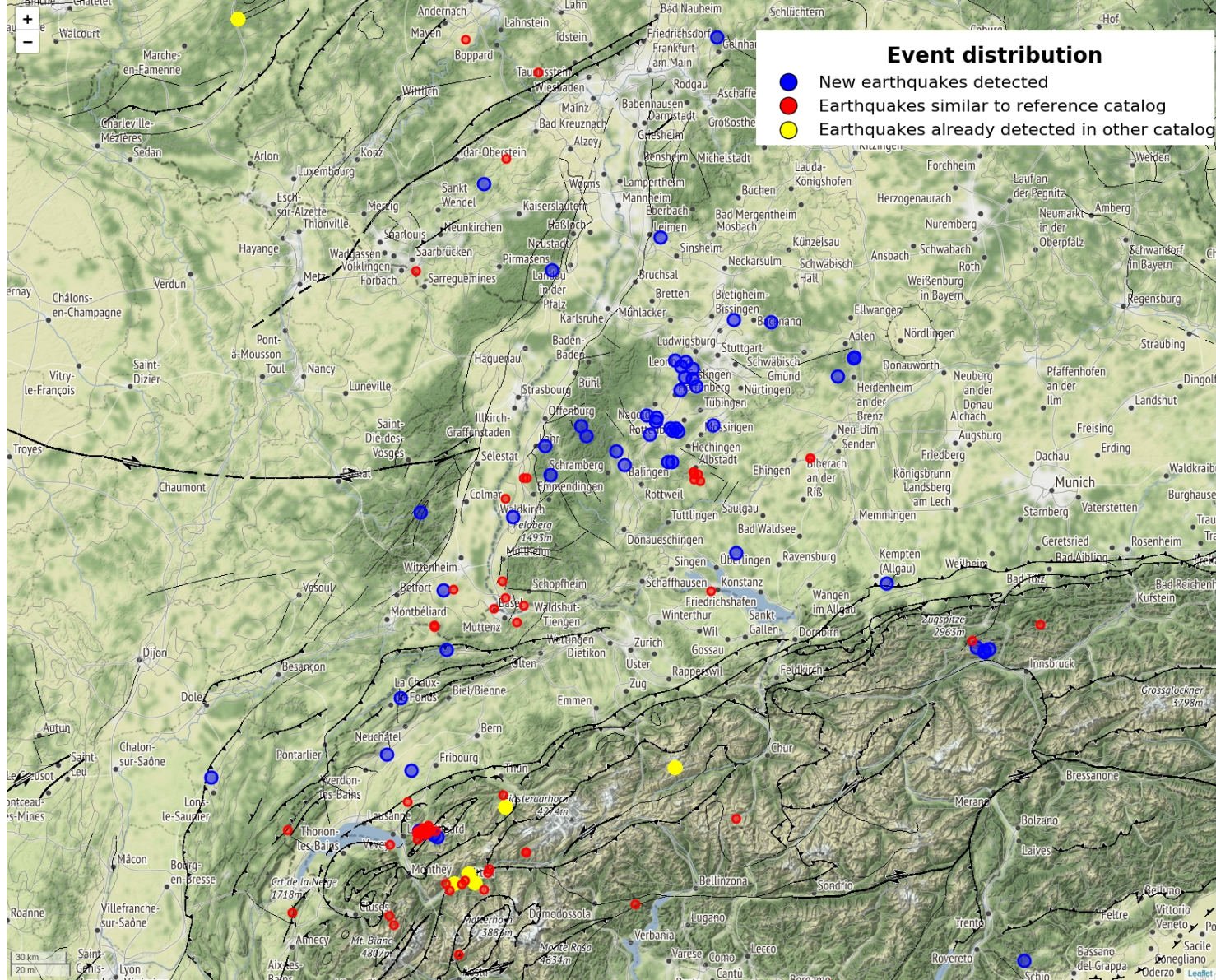
Quarry Blast Distribution (July 2016)

Reference catalog :
BCSF-RéNaSS



(Natural) Earthquake Distribution (July 2016)

Reference catalog :
BCSF-RéNaSS



Conclusions

Reduce of completeness magnitude

Large decrease of initial false alarms

Quite robust small event discrimination in a high anthropogenic context

Methodology designed to work with large dataset (waveforms, events, features) : High Performance Computing

New integrate modules to the SeisComP3 real-time routine detection