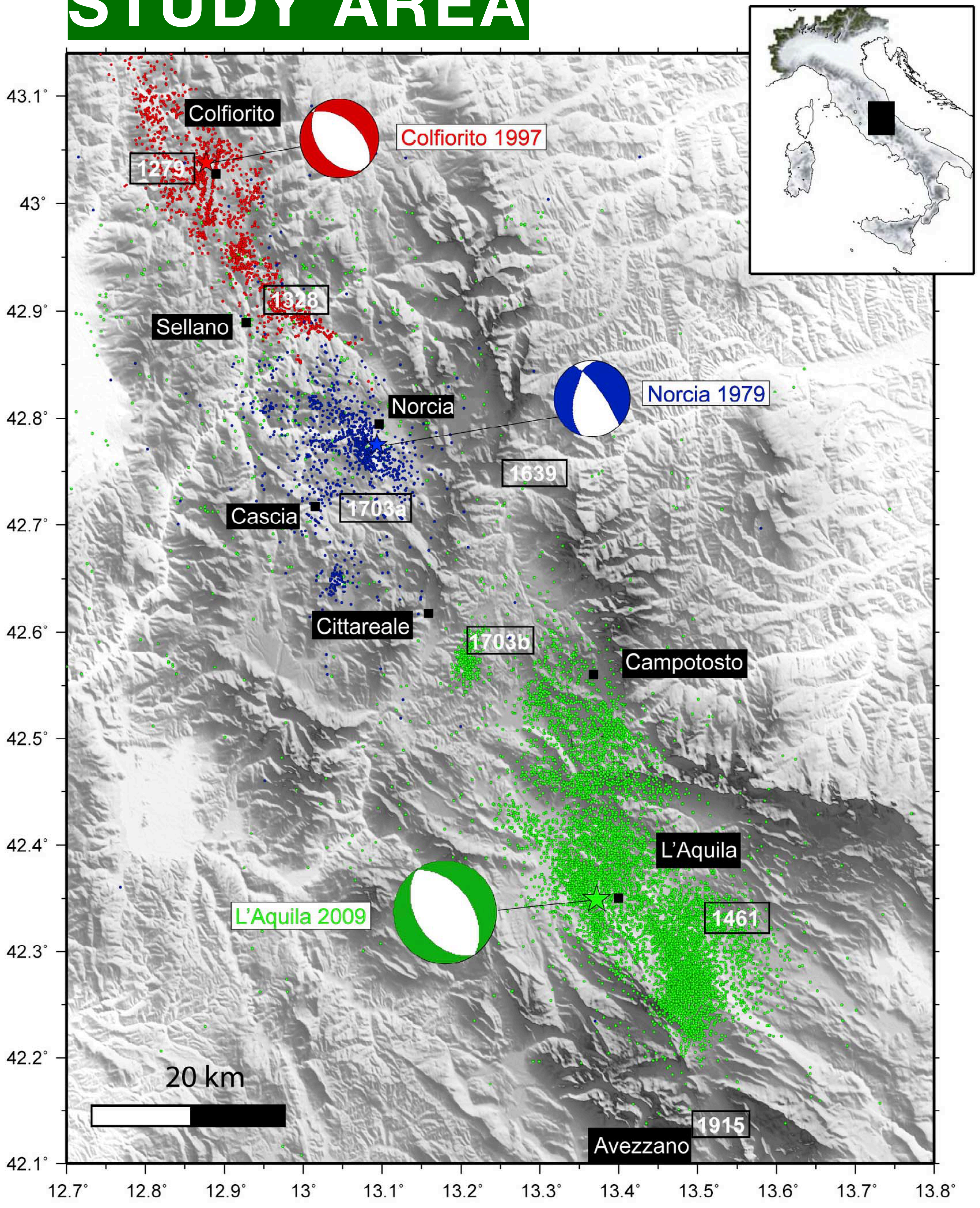
Extending near fault earthquakes catalogs using convolutional neural network and single-station waveforms

Josipa Majstorović, Sophie Giffard-Roisin and Piero Poli

Institut des Sciences de la Terre, Université Grenoble Alpes

EGU General Assembly 2020 online, May 6 2020

STUDY AREA



Our data consists of a three component seismograms from **1990** to **2019** recorded at **AQU** (42.354, 13.405) station near city of L'Aquila, central Italy.

Main study focus is **L'Aquila earthquake** (Mw 6.3) that occurred on **April 6th 2009**, 01:32 UTC right beneath the city of L'Aquila (Abruzzo region).

There have been several main shocks with an aftershock sequences.

The existence of continues records for an extended period of time and the seismically active study area are perfect for studying sesional and tidal effects on the local seismicity.

Chiaraluce et. al., 2011

INTRODUCTION

Far reaching goal

Starting point

To study the sesional effects on the local seismicity we need to **extend near fault earthquake (EQ) catalog**.

Method

This is achieved by using **a convolutional neural network (CNN)**.

Data

The CNN learns to recognise local uncatalogued events from **ALL catalog events**.

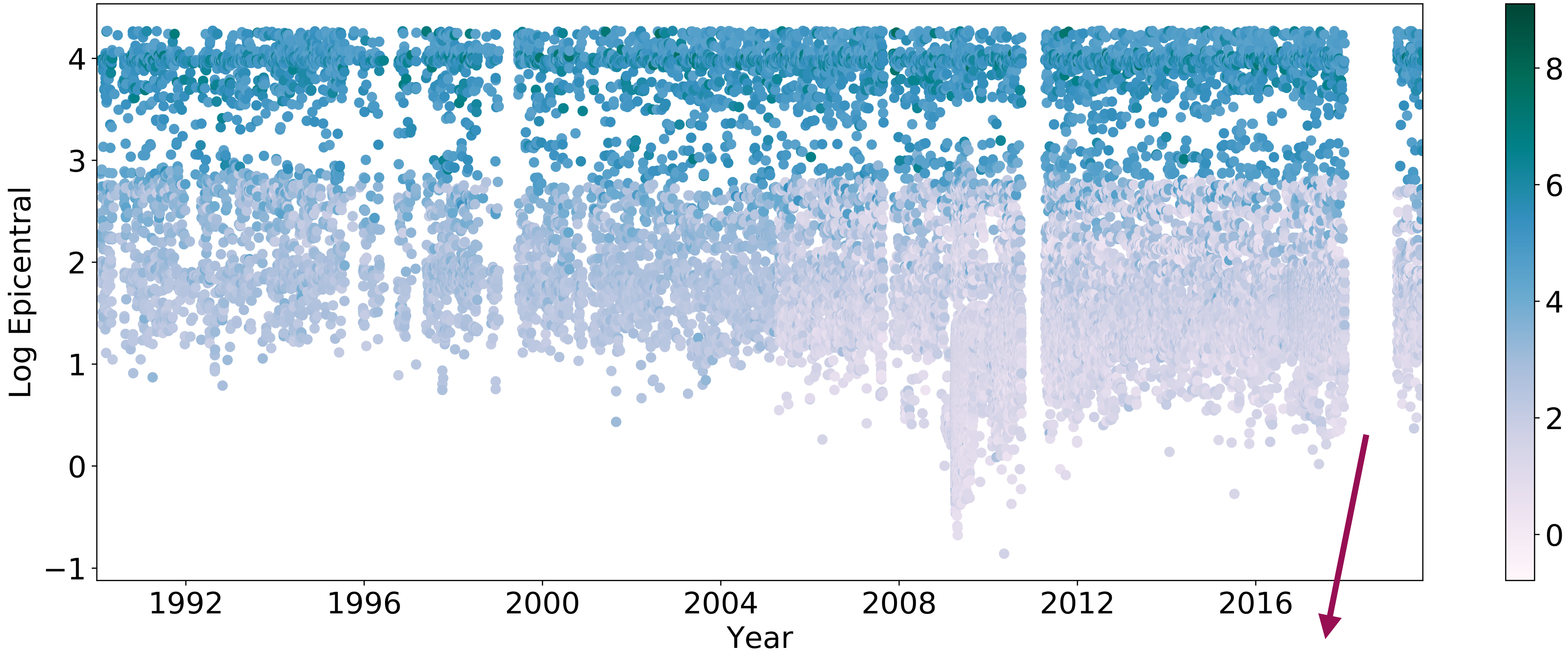
Goal

We want to **develop the CNN model** that recognise if something is an EQ and whether it is a local EQ.

DATA PREPARATION

The input catalogs are INGV, *Valoroso et. al.*, USGS (for M>4.5).

Only those events with SNR > 2 were accepted and this left us with **65 865** events.



Time evolution of the accepted events

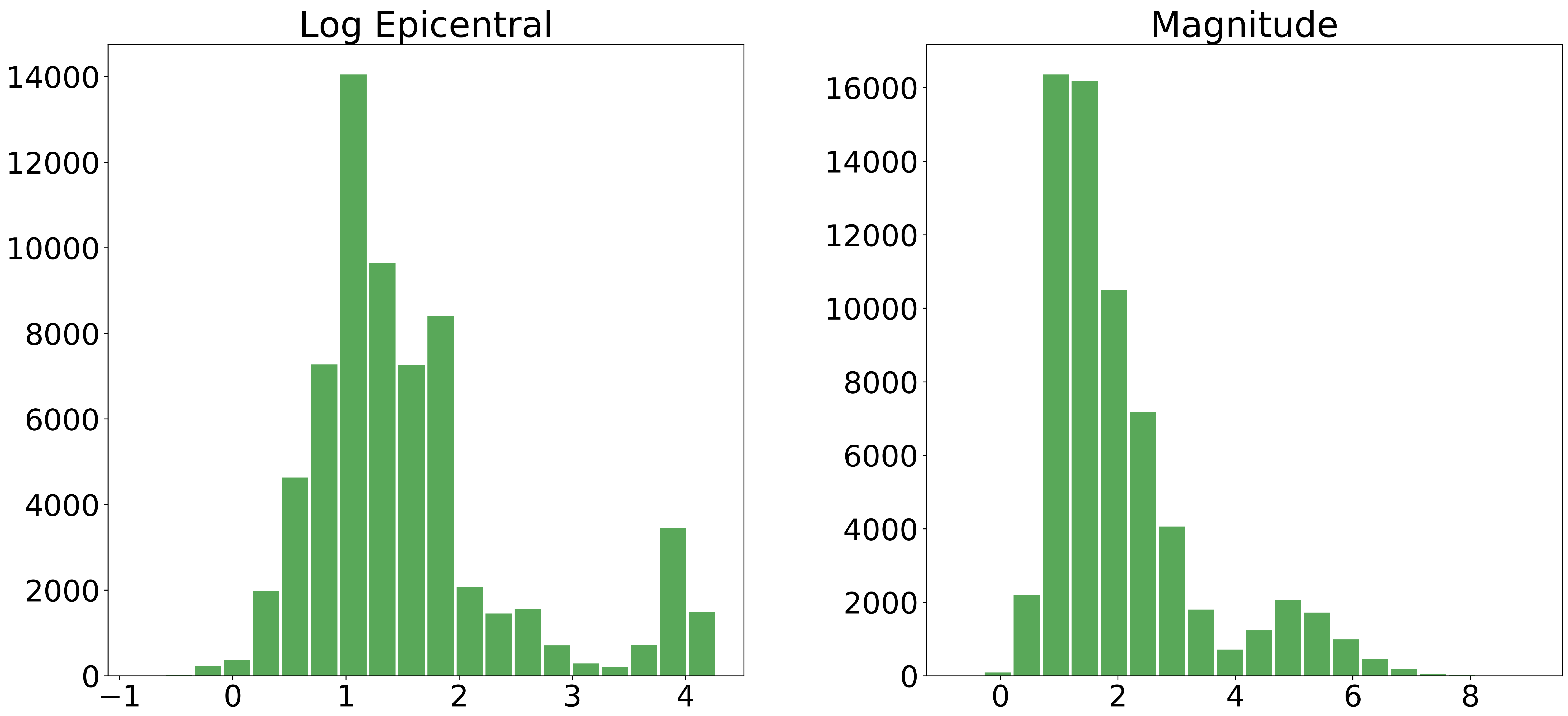
Station AQU was not recording.

DATA PREPARATION

The input catalogs are INGV, *Valoroso et. al.*, USGS (for M>4.5).

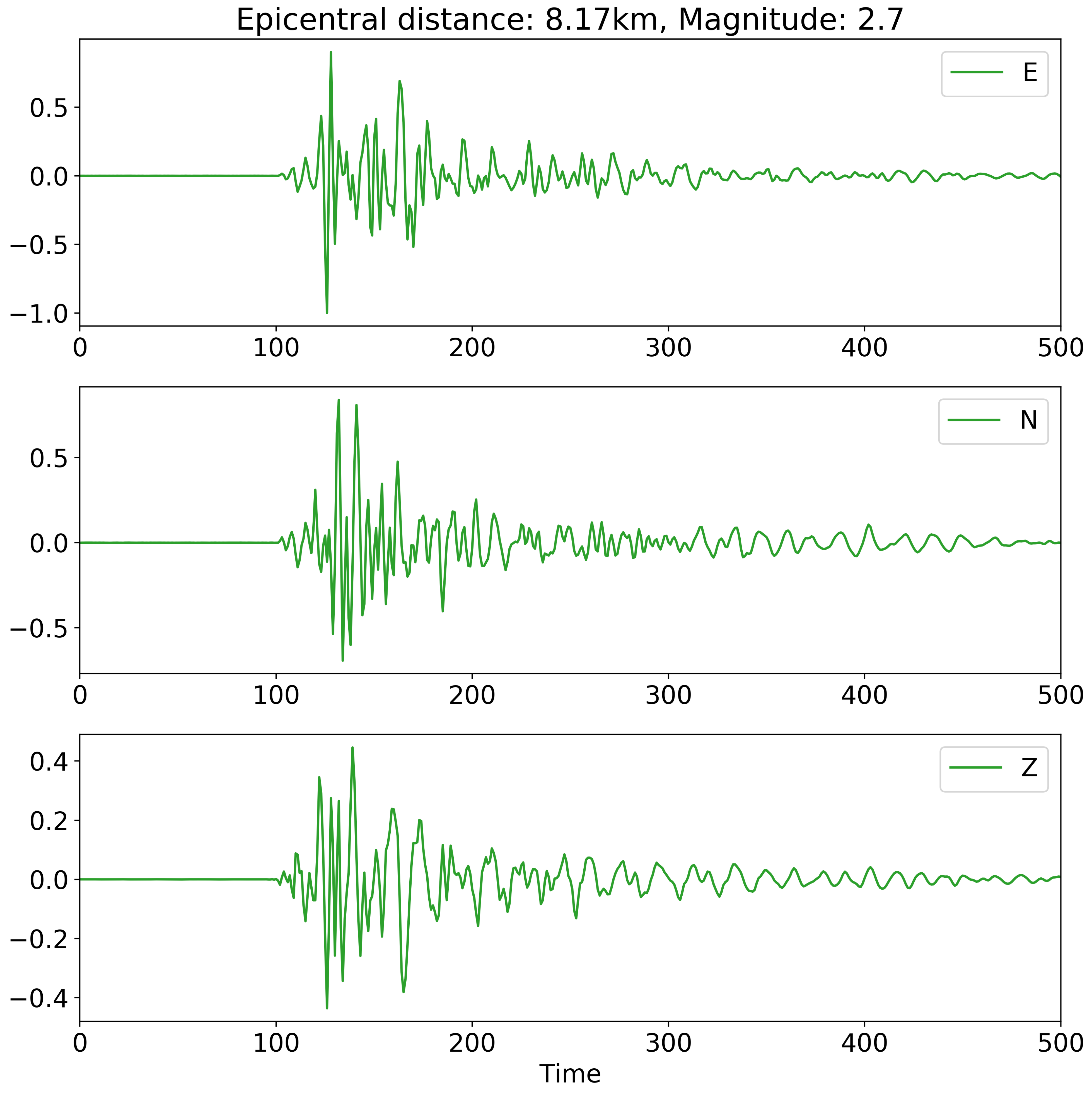
Only those events with SNR > 2 were accepted and this left us with **65 865** events.

Histograms of selected epicentral distance and magnitude



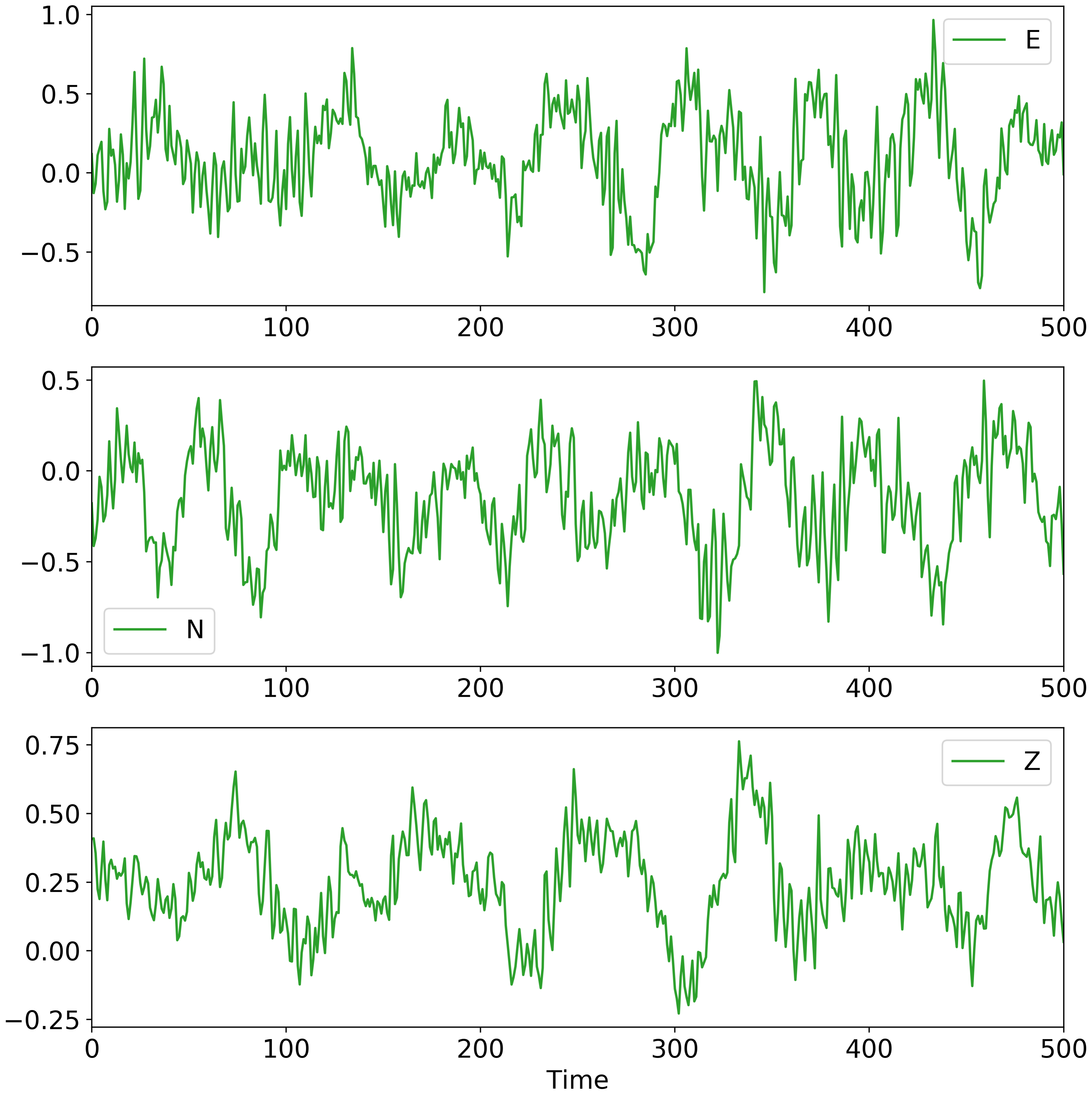
DATA PREPARATION

For all **65 865** events we extracted **25-s-long** EQ event windows.



positive sample

We also extracted **65 865** 25-s -long windows that do not contain EQ events - the noise events.



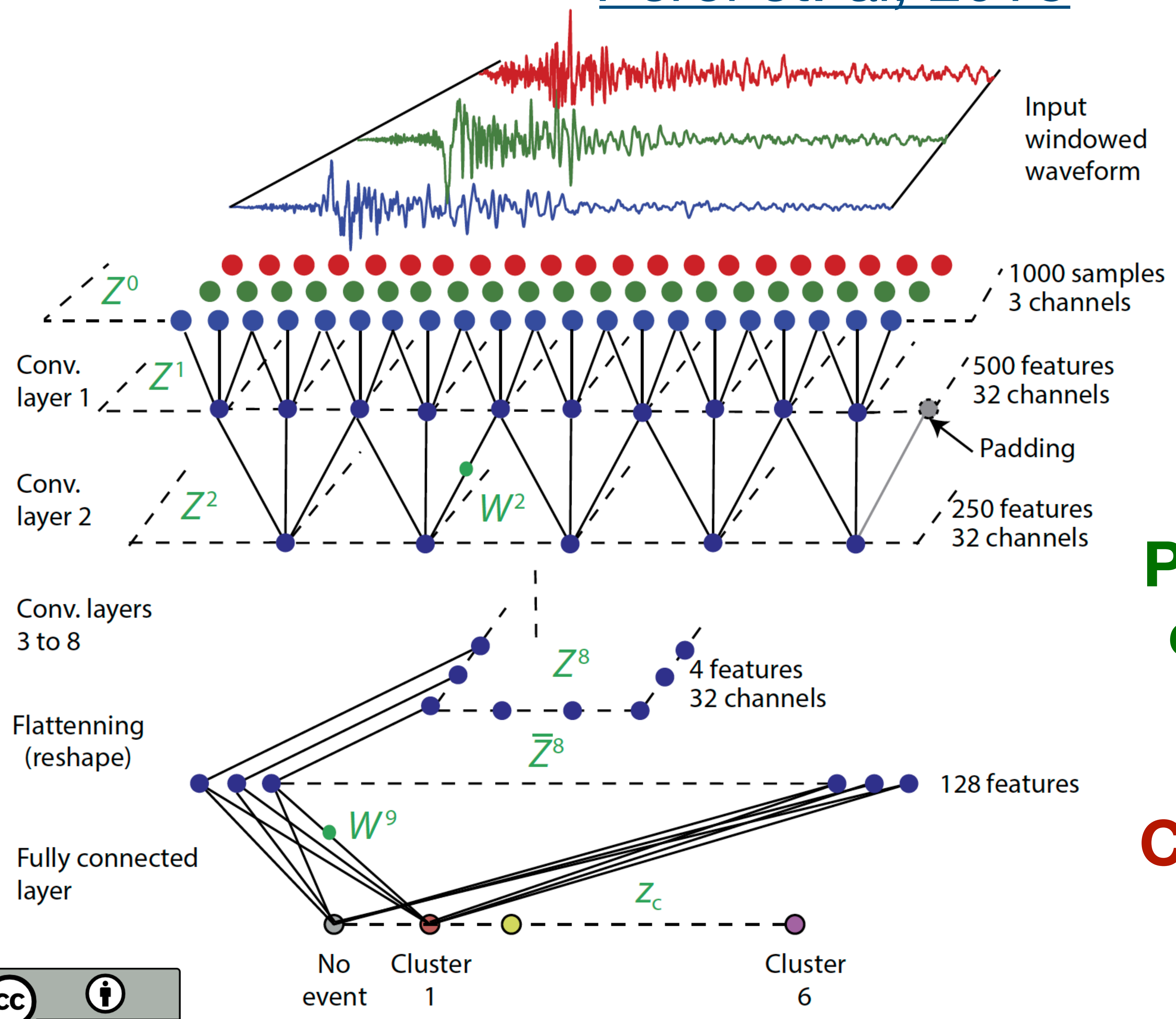
negative sample

MODEL DEFINITION

Following the statement: *We want to* ***develop the CNN model*** *that recognise if something is an EQ and whether it is a local EQ.*

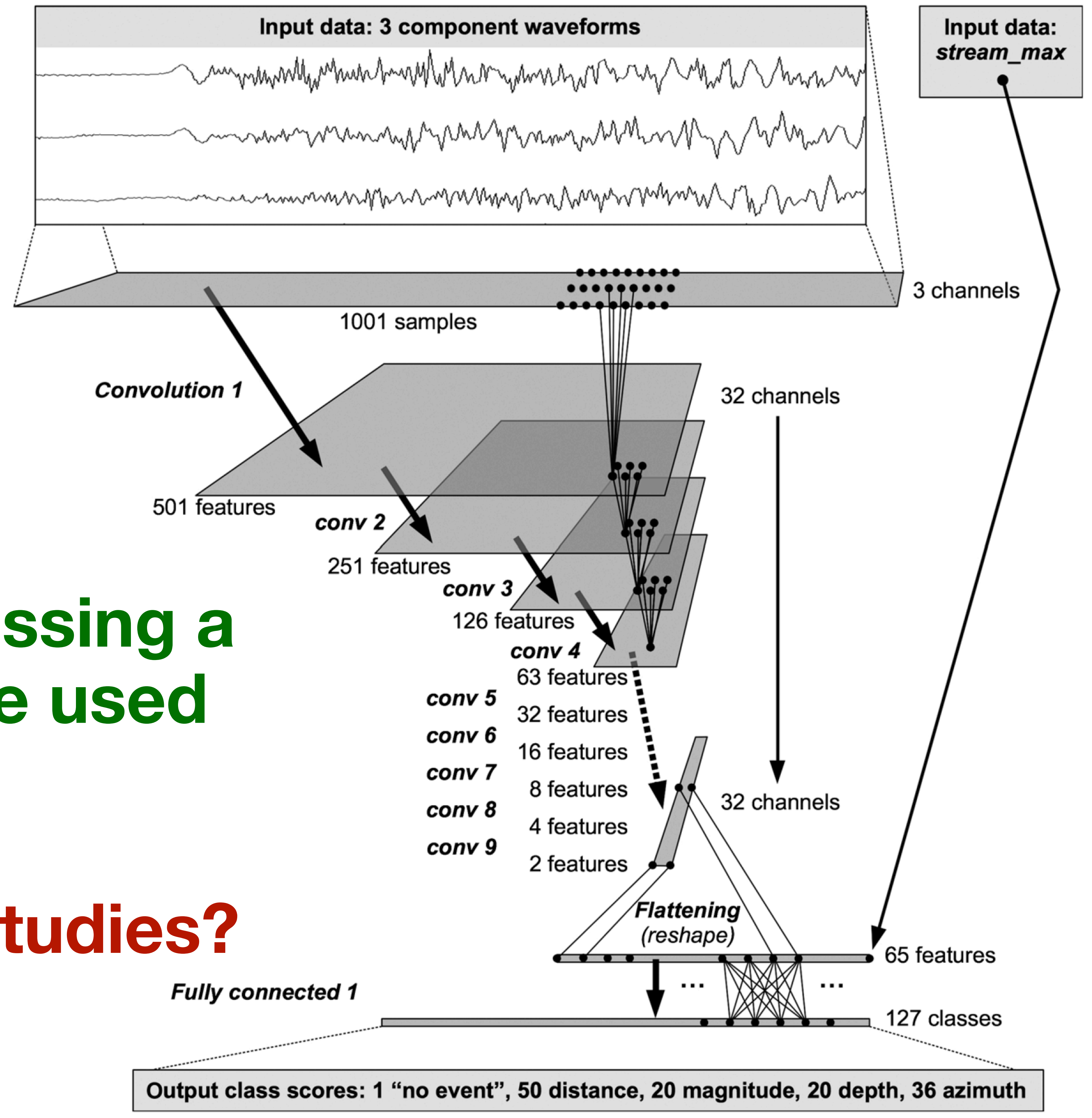
State of art

[Perol et. al, 2018](#)



[Lomax et. al., 2019](#)

CovNetQuake_INGV Architecture



Promising studies, but missing a careful explanation of the used model.

Can we elaborate those studies?

MODEL DEFINITION

Following the statement: *We want to* ***develop the CNN model*** *that recognise if something is an EQ and whether it is a local EQ.*

We worked on two models.

The CNN detector

The CNN model that is able to differentiate between the EQ event window and the noise window.

Two classes: earthquake and noise.

The CNN classifier

The CNN model that is able to classify the EQ event windows based on their epicentral distances and magnitudes.

Two classes: epicentral distance and magnitude.

FINDING THE BEST MODEL - TESTING

There were three sections of the performed tests:

1. Which is the best model?

How many subclasses should the CNN classifier have? Two classes for the epicentral distance and the magnitude or three classes? Three classes for the epicentral distance and only two for the magnitude.

. . .

2. The data normalisation

Which is the best normalisation for our dataset?

. . .

3. The hyperparameters of the the CNN model

We will focus further only on these tests…

The tests were performed by dividing the dataset in three parts:

Training 80%

Validation 10%

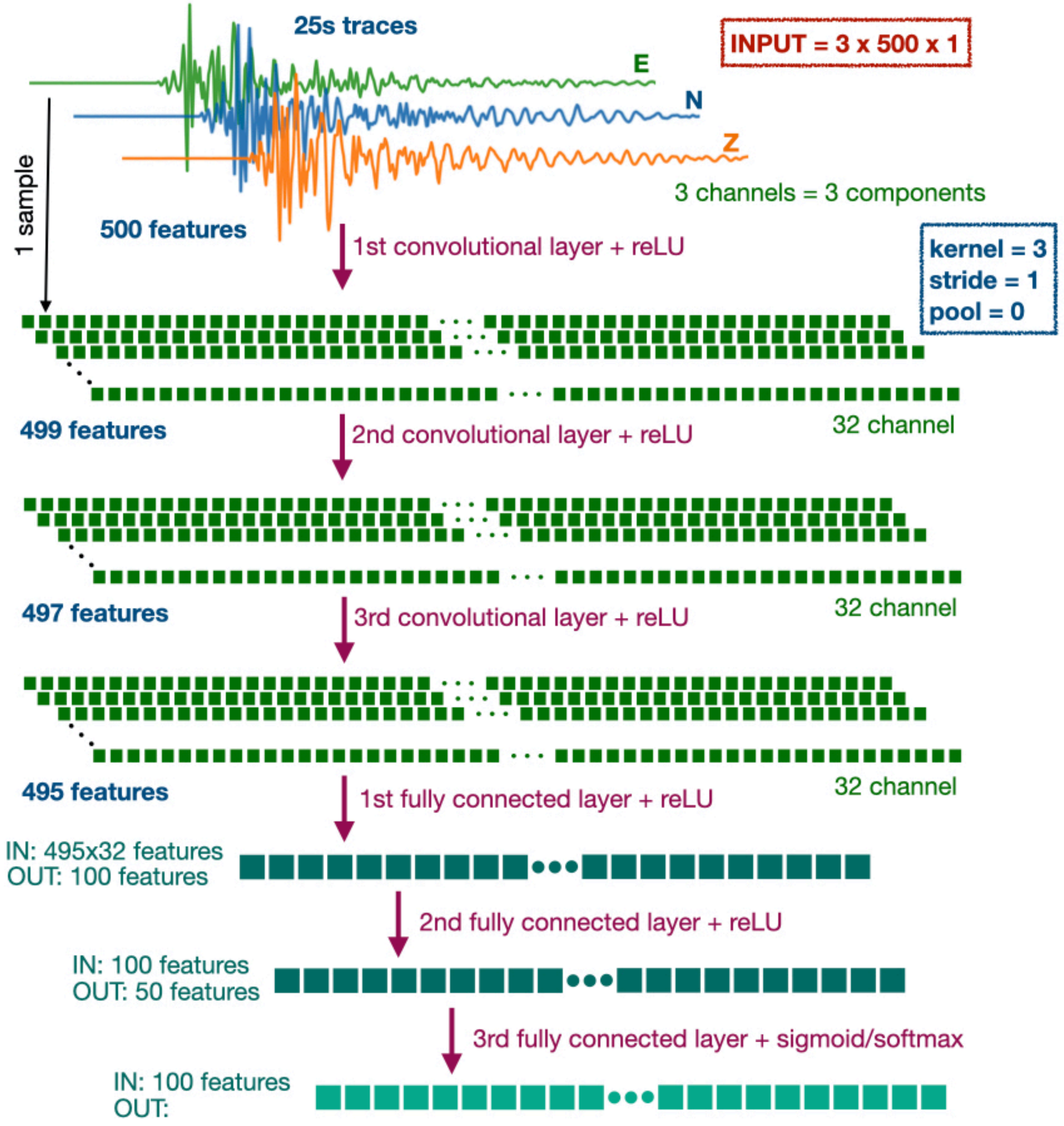
Evaluation 10%

TESTED HYPERPARAMETERS

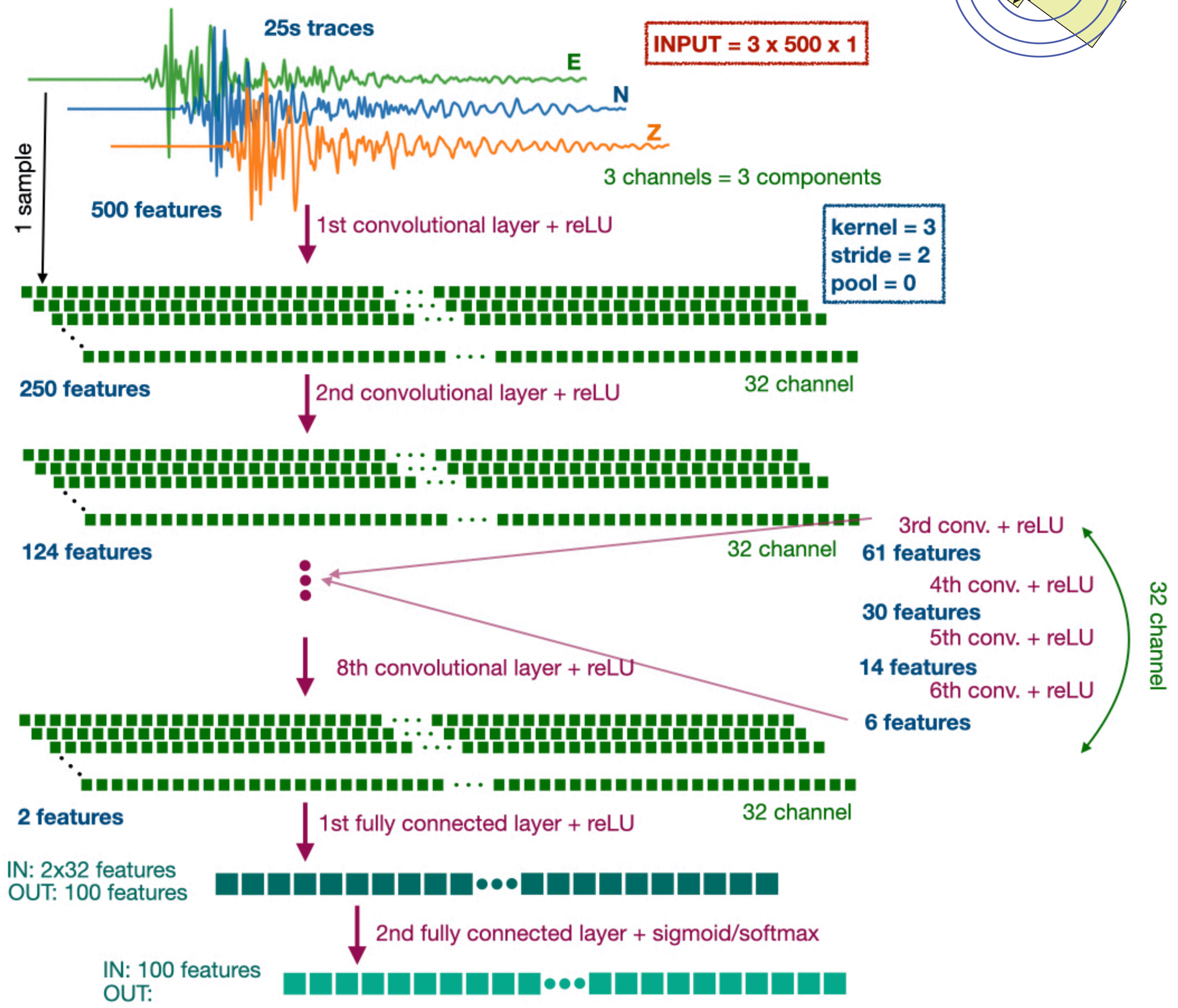
Number of the parameters to test	Description	Parameters
2	The number of neurons or the architecture of the model (kernel size, pooling, stride, number and type of layers)	2 models
Fixed	The activation functions	The sigmoid for the detector and the softmax for the classifier
2	Optimisation algorithms	SGD and ADAM
4	The learning rate and the momentum of the optimisation algorithms	Learning rate = [0.0001, 0.01] Momentum = [0.2, 0.9]
2	The mini batch size	N = 128, 512
Fixed	The number of epochs	Early stopping with patience of 50

There has been 2^5 = 32 preliminary tests based on the grid search for the two CNN models: the detector and the classifier.

TESTED HYPERPARAMETERS - 2 MODELS

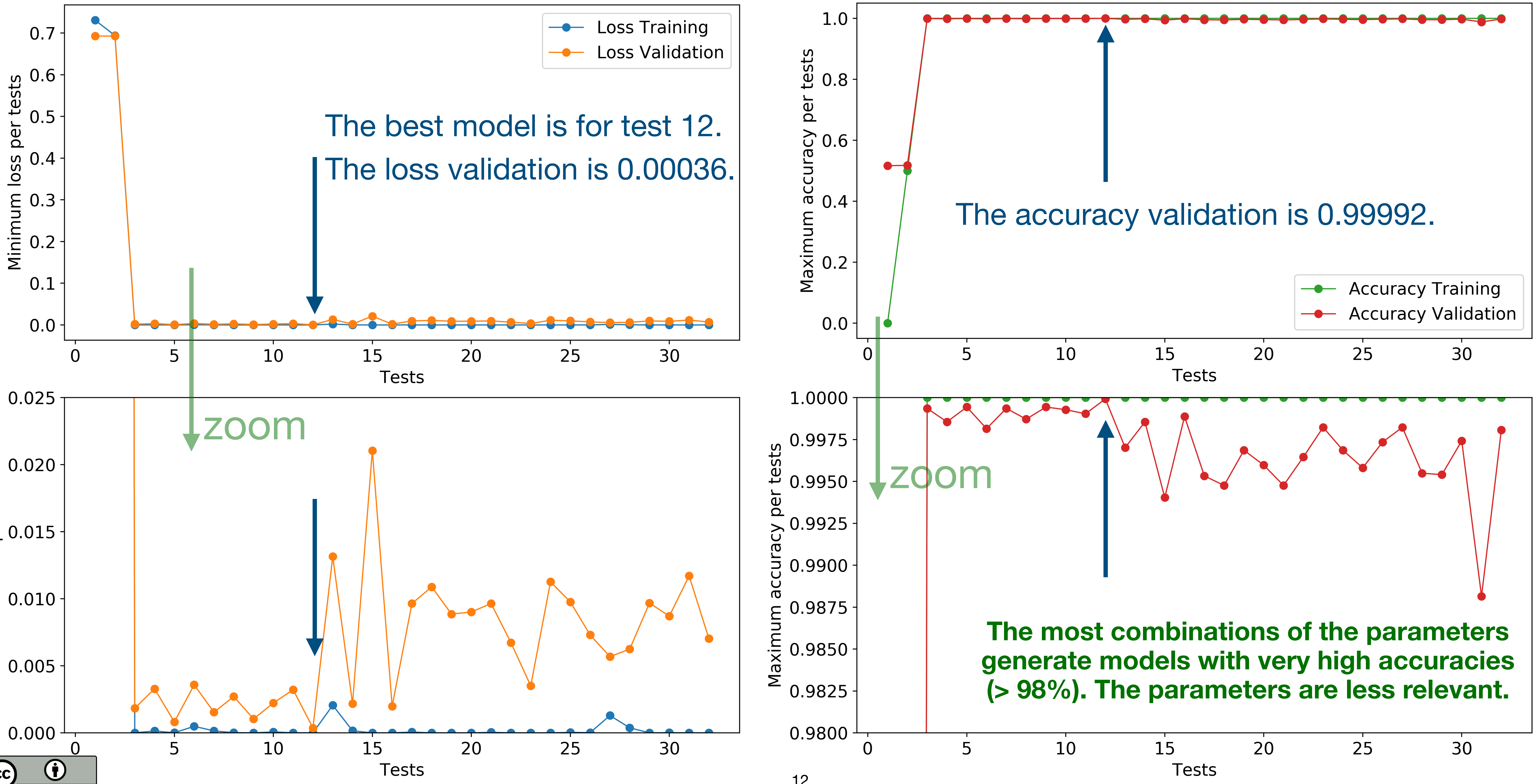


1 595 729 trainable weights **The first model**

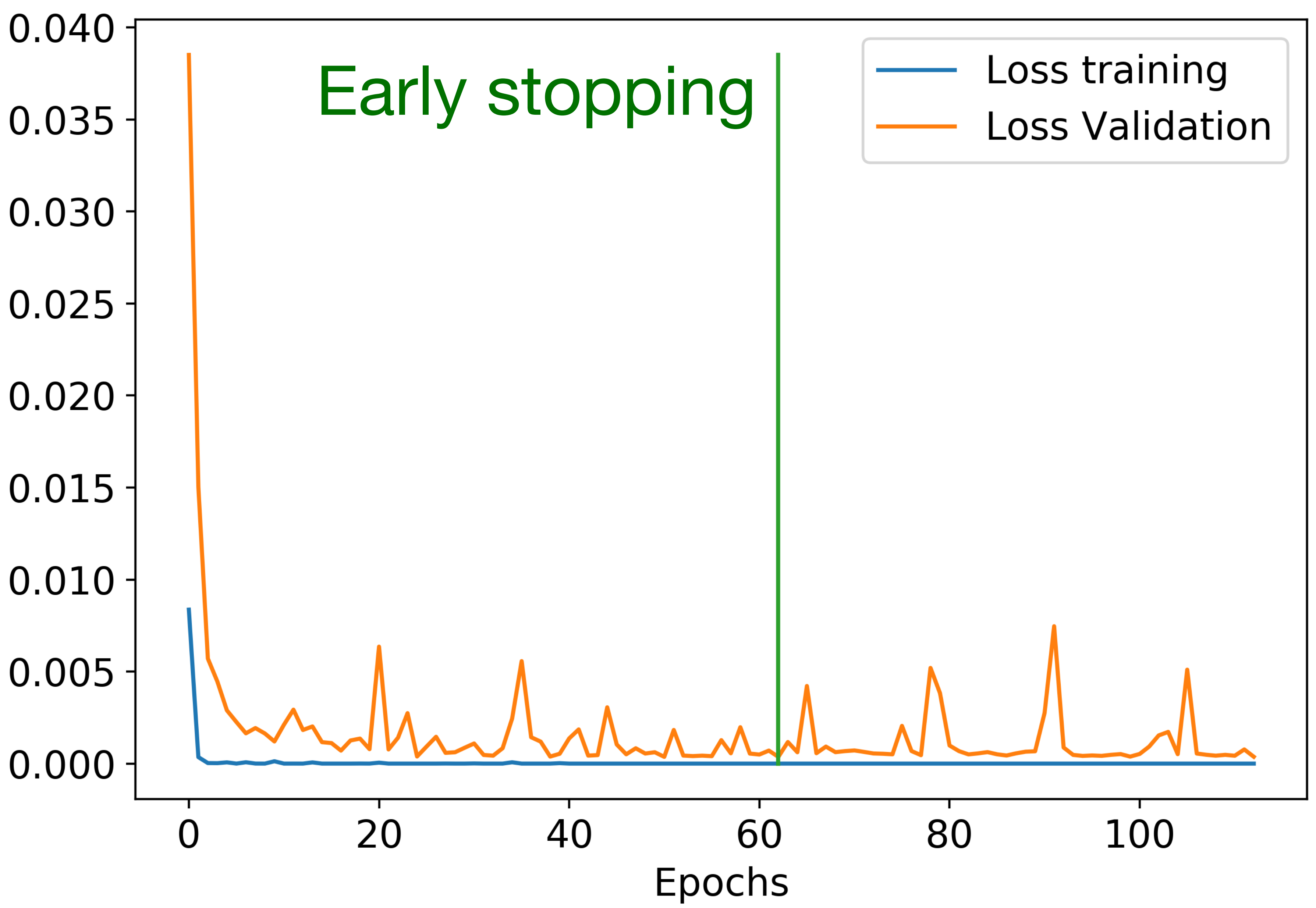


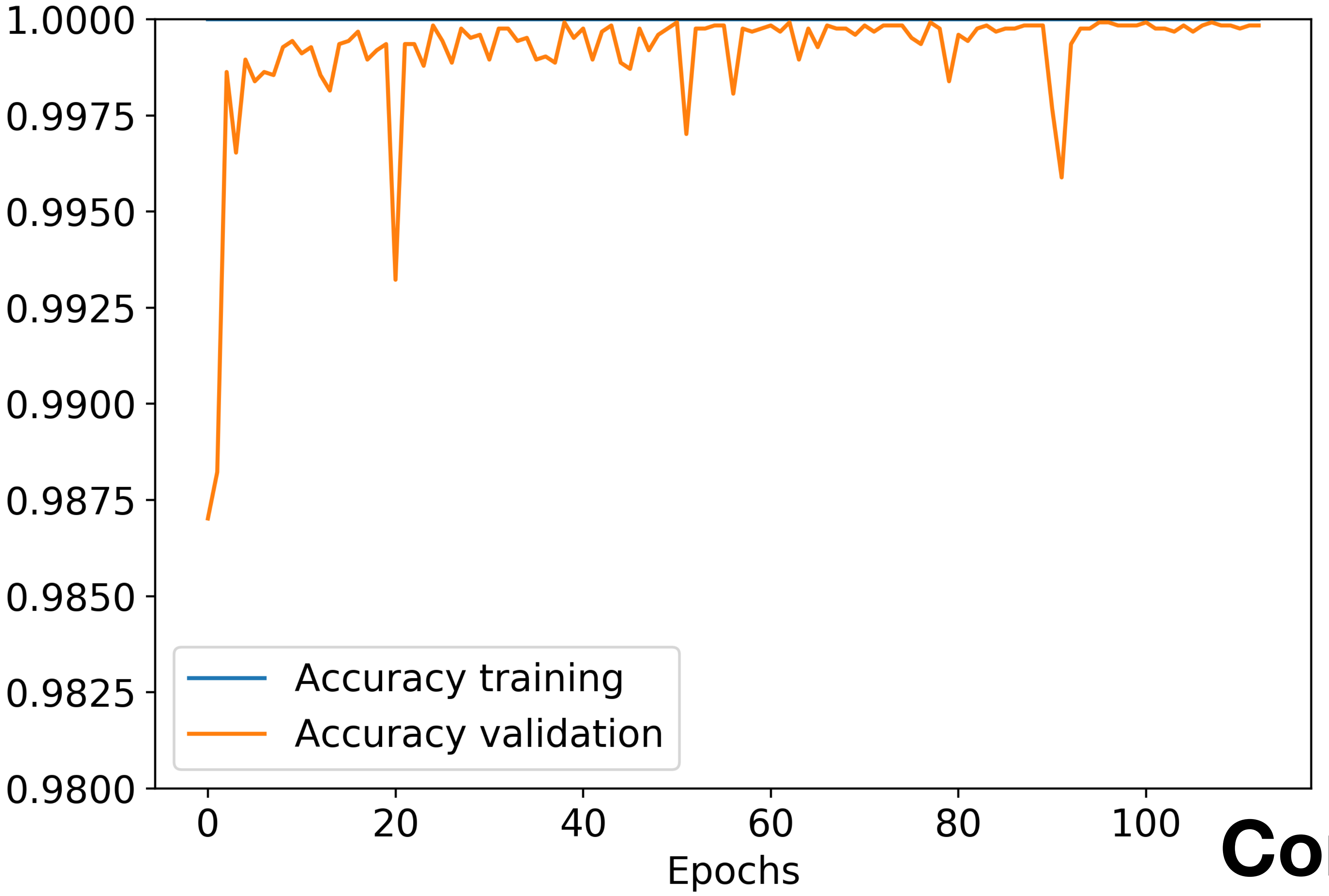
30 545 trainable weights **The second model**

THE DETECTOR RESULTS ALL TESTS



THE DETECTOR RESULTS TEST 12

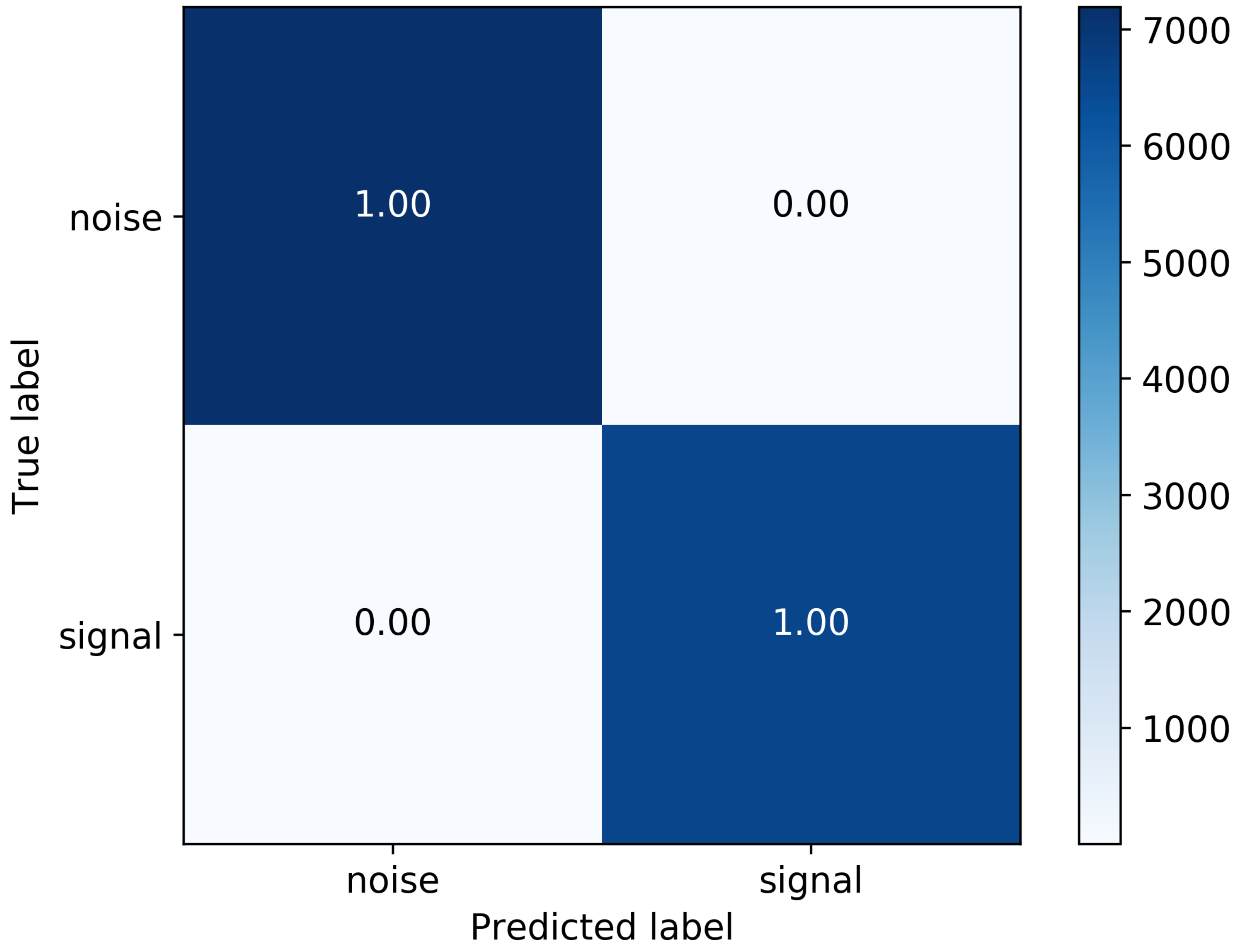




Hyperparameters

Model	First one
Opt. algorithm	Adam
Learning rate	0.01
Momentum	0.9
Batch size	512

Confusion matrix

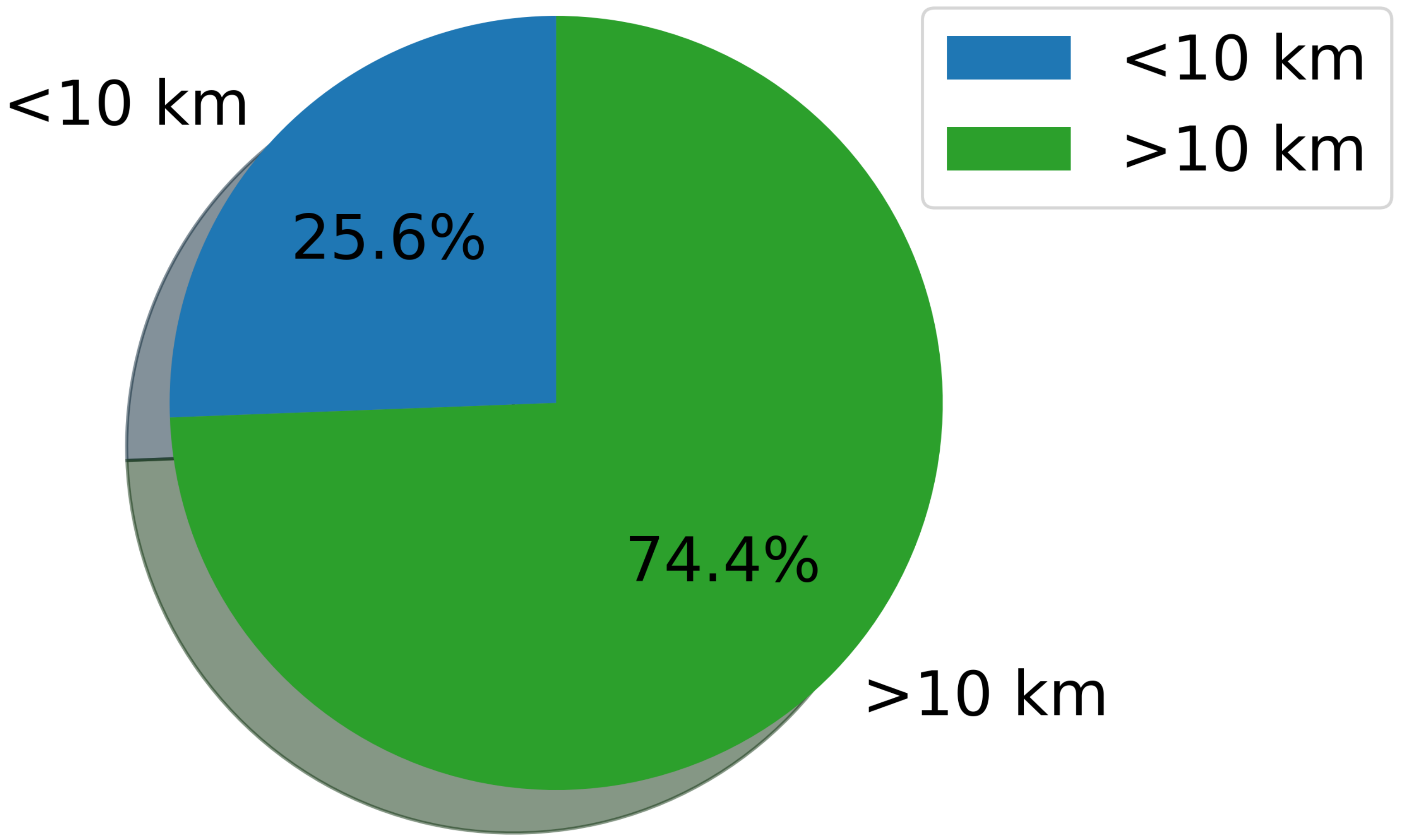


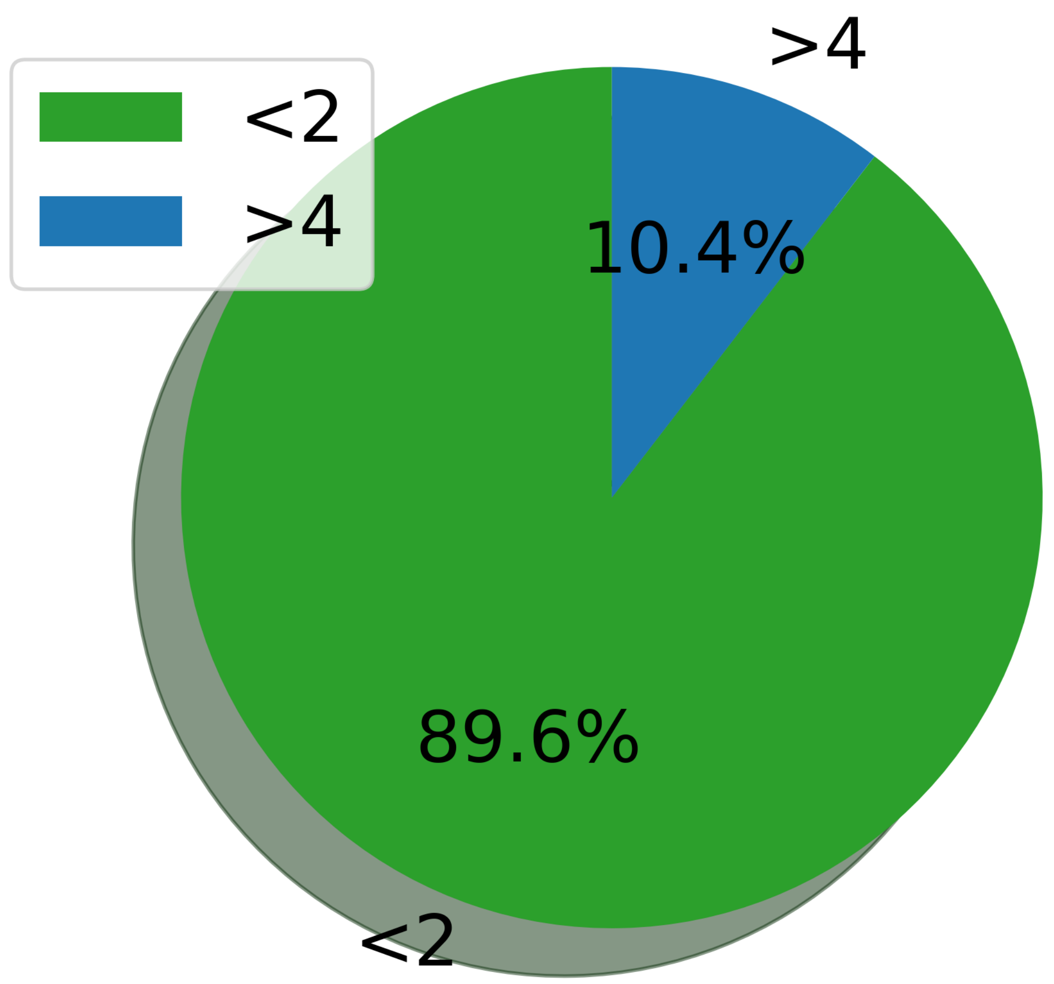
Testing the selected model 12 on the evaluation data (data that model has never seen before) and accuracy is still 100%.

The model is capable of differentiating between the earthquake and the noise event windows.

THE CLASSIFIER RESULTS

First, we choose to work with the CNN model that classify the EQ events windows into **two subclasses** for epicentral distance and **two subclasses** for magnitude.

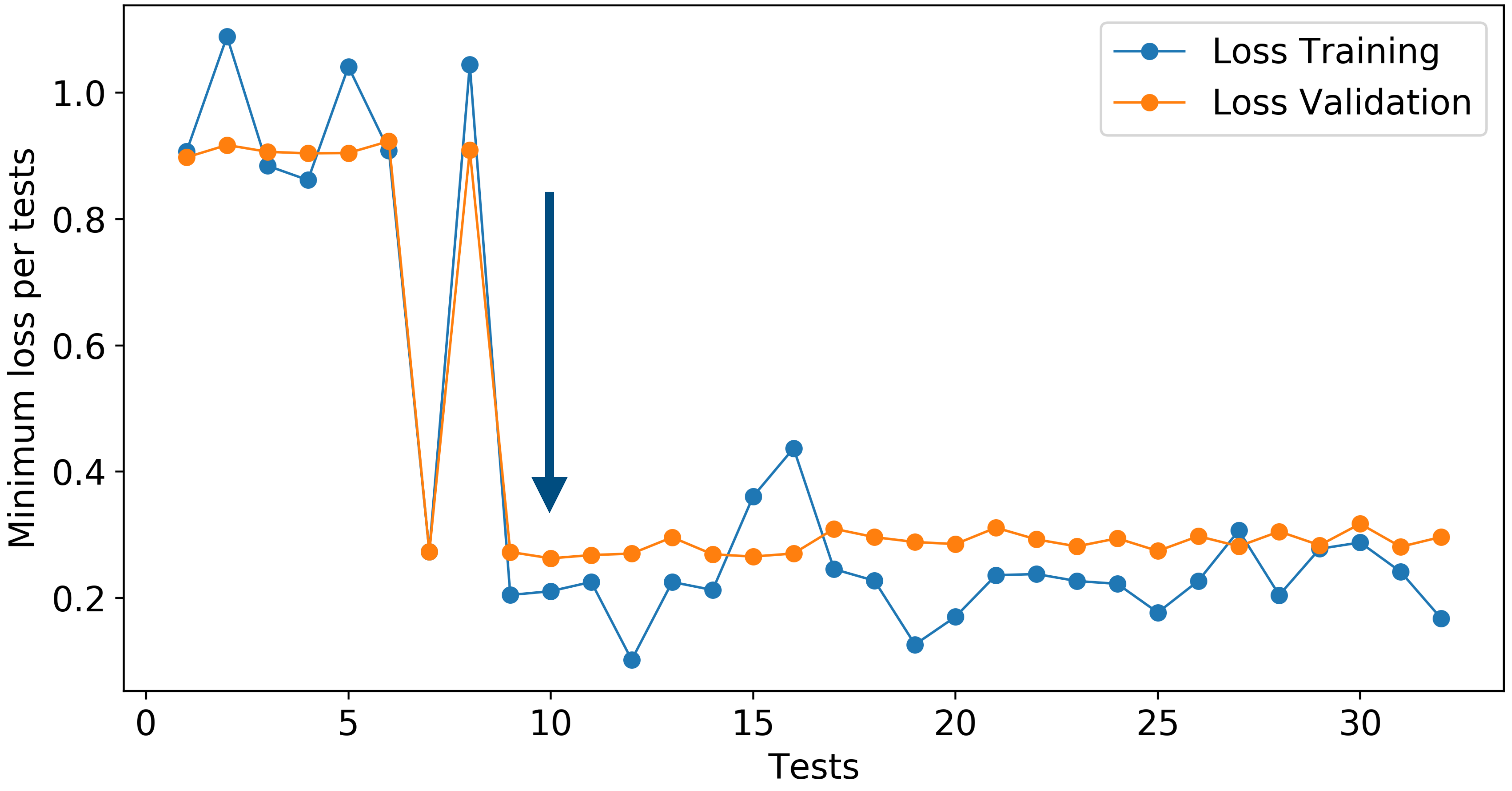




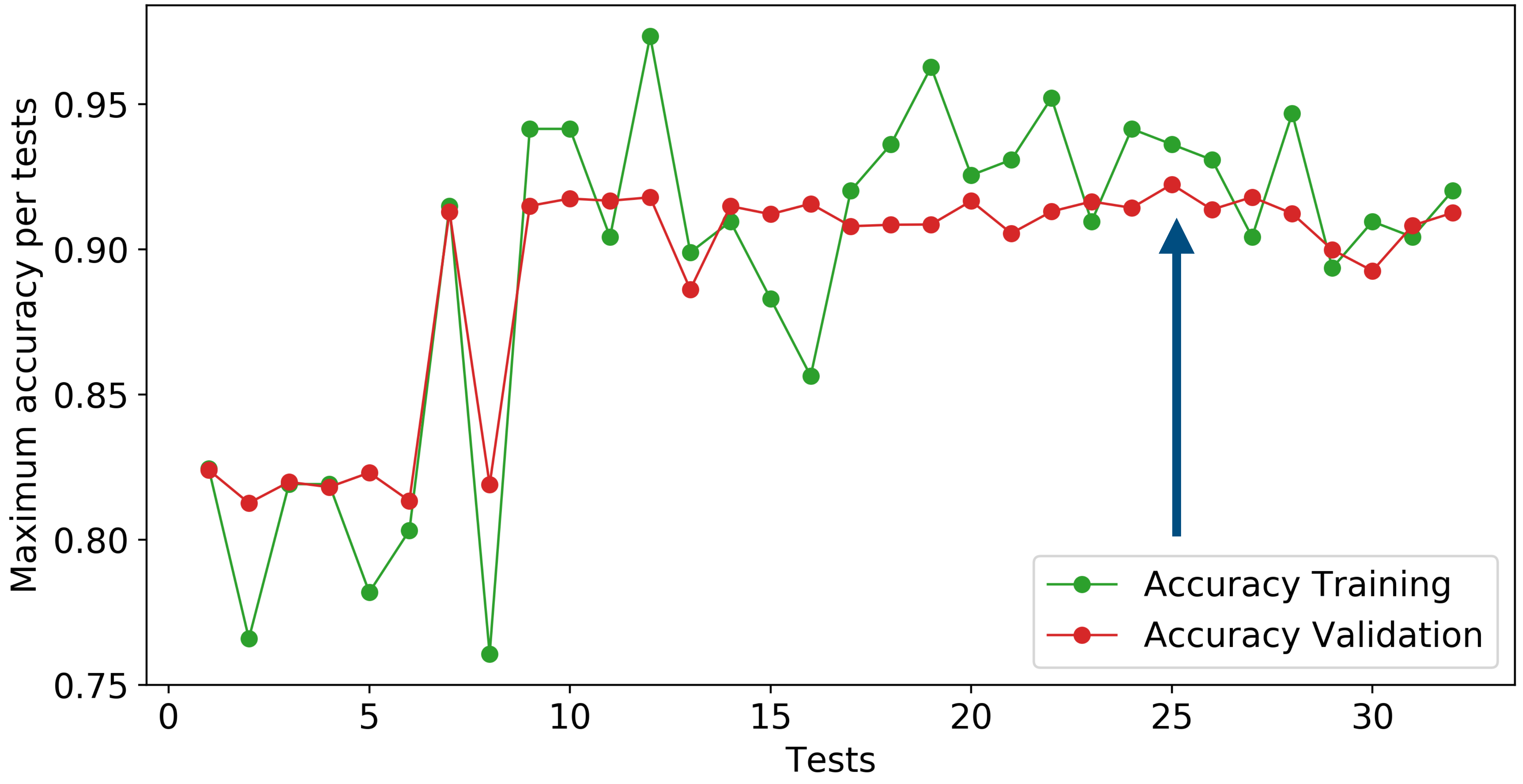
	Epicentral distance	Magnitude
1 subclass	< 10 km	< 4
2 subclass	> 10 km	> 4

THE CLASSIFIER RESULTS

The minimum loss value is achieved for test 10.



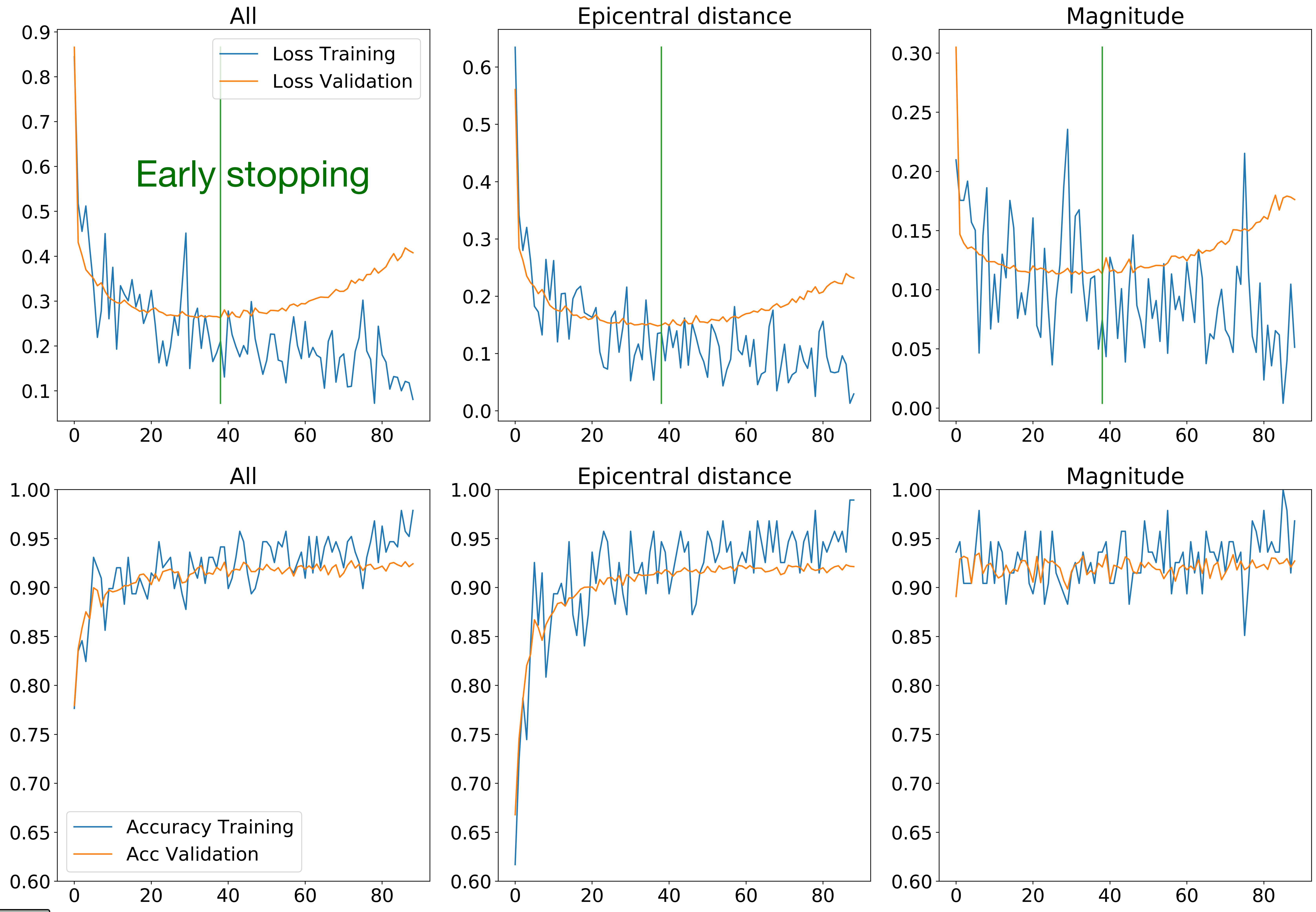
The maximum accuracy values is achieved for test 25.



	Loss Validation	Accuracy validation
Test 10	0.26297	0.91750
Test 25	0.27458	0.26297

The parameters are more relevant when it comes to the models' accuracies, due to the fact that the classification problem is more complex than the detection problem.

THE CLASSIFIER RESULTS TEST 10

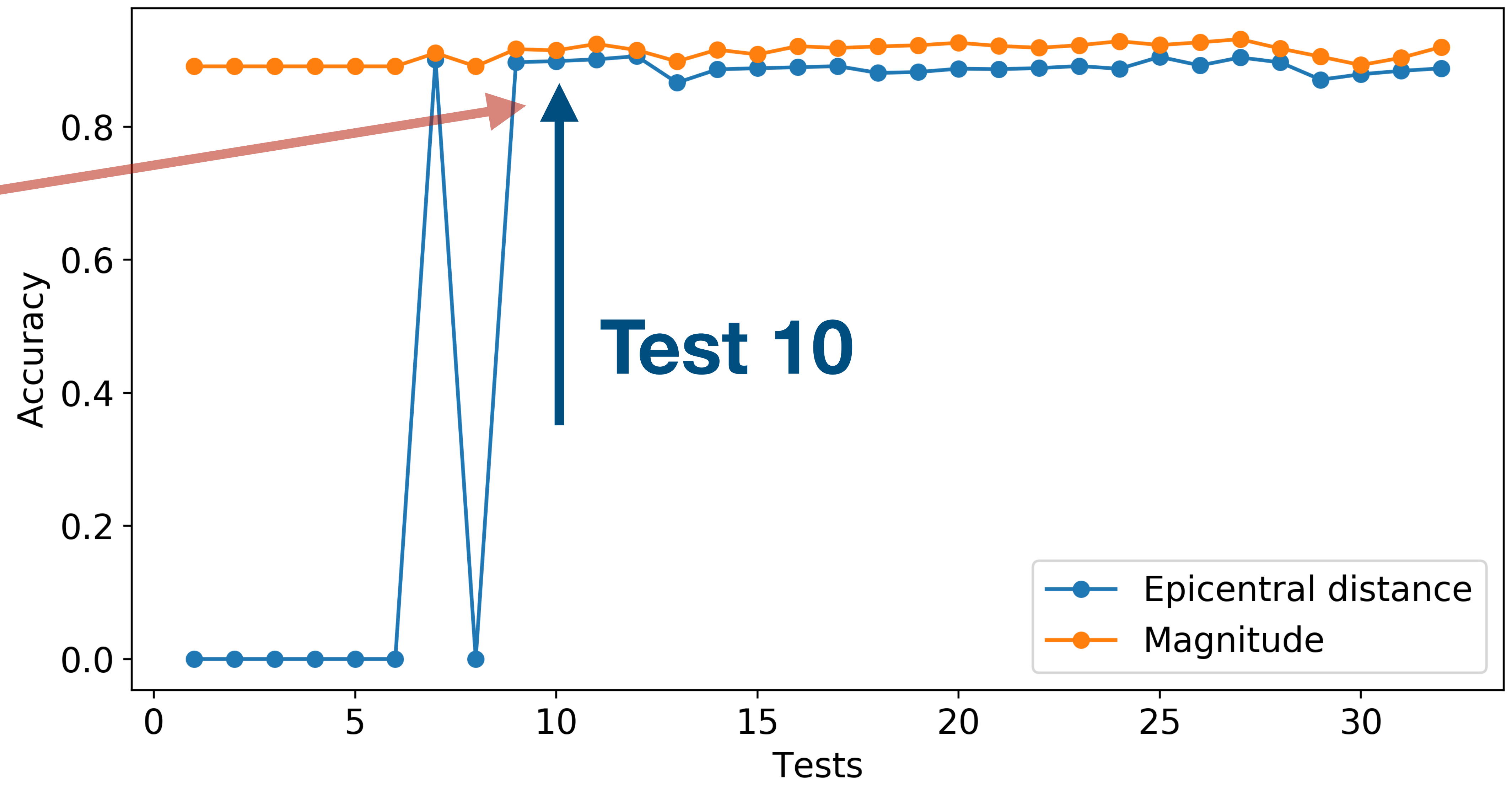


Hyperparameters

Model	First one
Opt. algorithm	Adam
Learning rate	0.0001
Momentum	0.2
Batch size	512

THE CLASSIFIER RESULTS EVALUATION

For the test 10 the change in the optimisation algorithm is introduced.

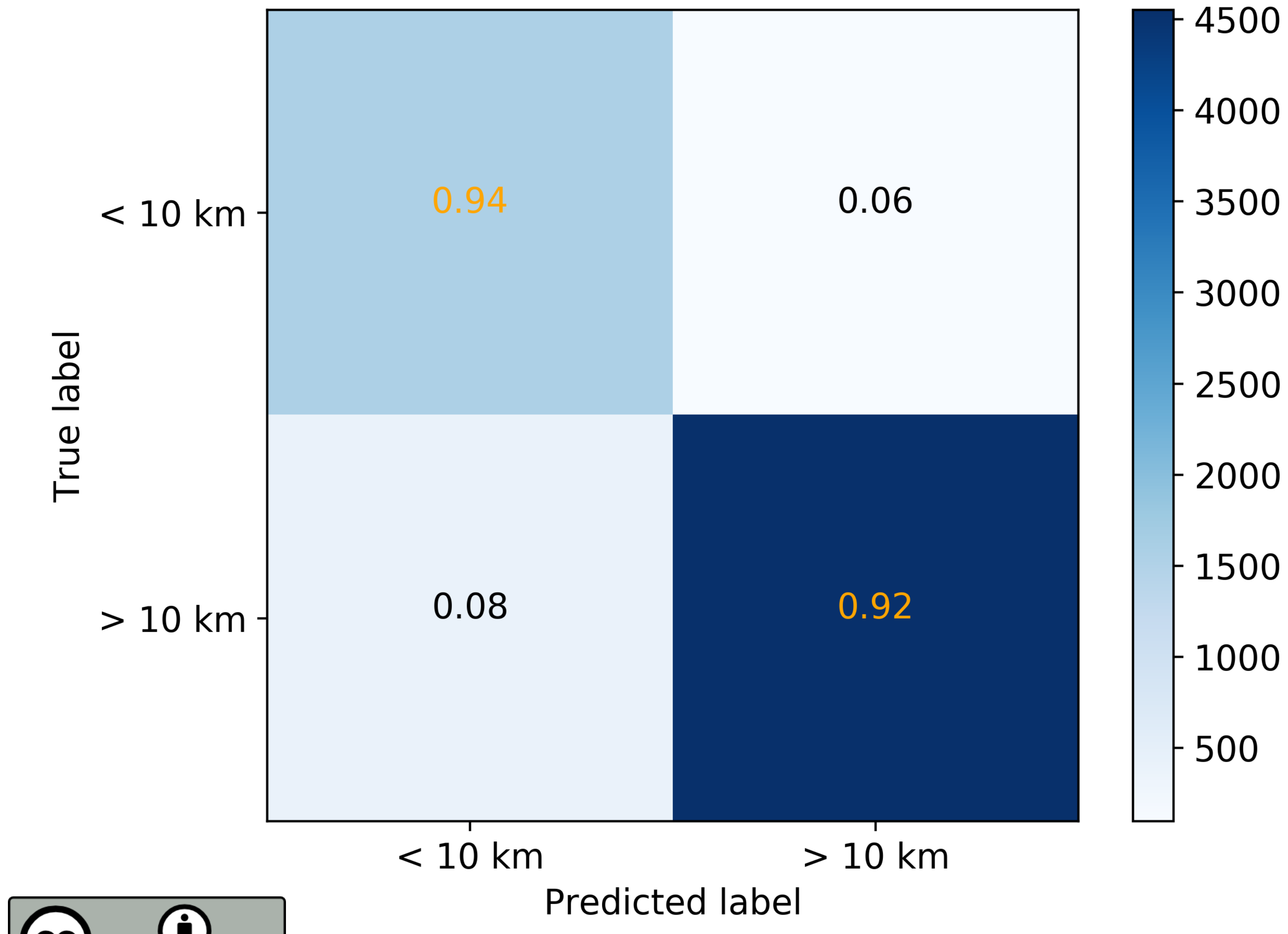


The epicentral distance accuracy for first 9 tests is unstable and after it stabilises around 90%.

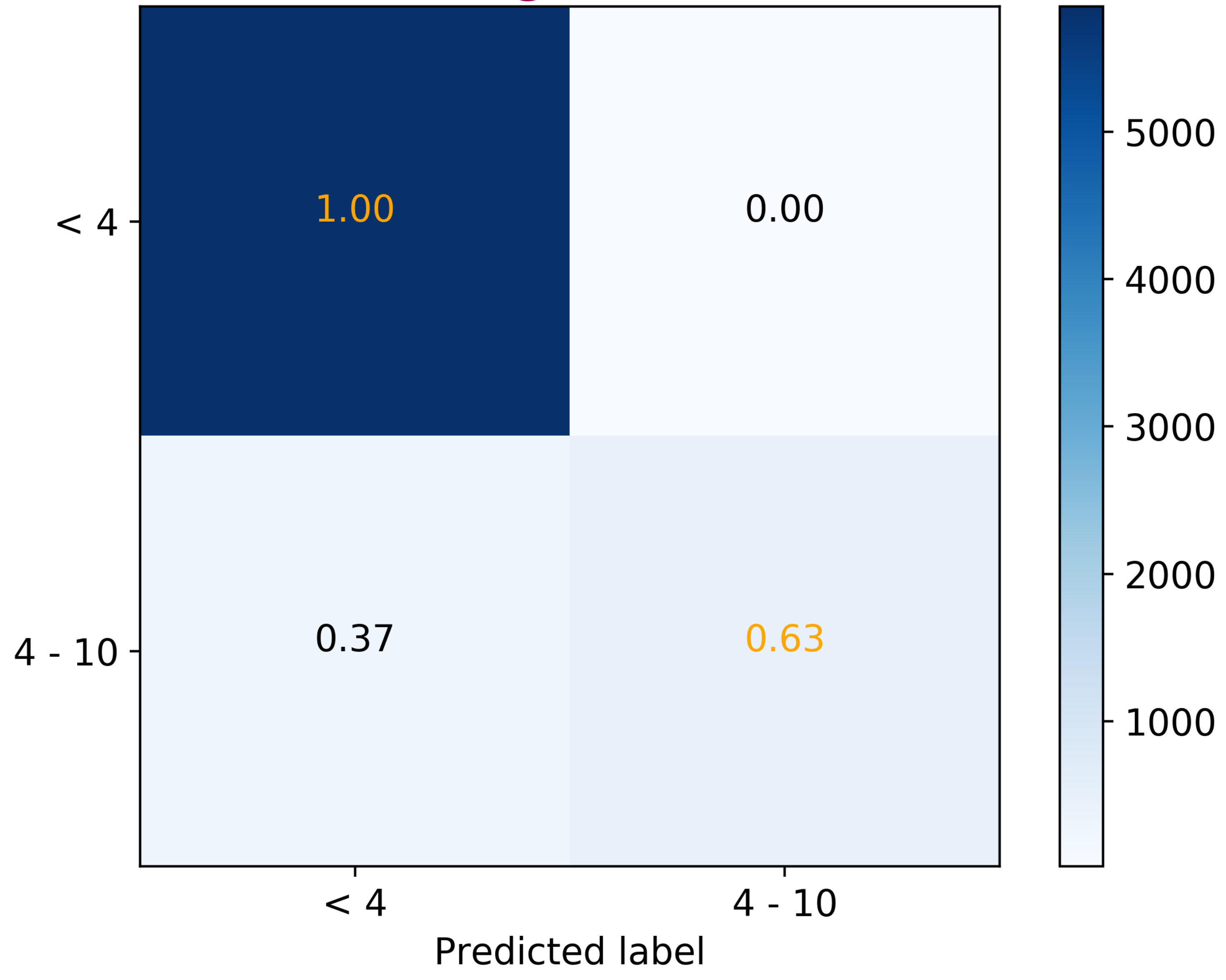
The magnitude accuracy for all tests is around 90%.

Test 10

Epicentral distance



Magnitude



For individual test 10 the model is able to separate between the two epicentral classes successfully.

For the magnitude classes model from test 10 is able to differentiate all events with magnitude < 4. For events with magnitude >4 the accuracy is 63%.

CONCLUSIONS - SUMMARY OF ALL TESTS

T = training
V = validation

		The detector				The classifier							
Tests	Models	T Loss	V Loss	T Acc	V Acc	T Loss	V Loss	T Acc	V Acc	Opt.	Learning rate	Momentum	Batch size
1	#1	7.31e-01	6.93e-01	0.00e+00	5.16e-01	0.91	0.90	0.82	0.82	SGD	1E-04	0.2	128
2	#1	6.94e-01	6.93e-01	5.00e-01	5.18e-01	1.09	0.92	0.77	0.81	SGD	1E-04	0.2	512
3	#1	3.52e-06	1.83e-03	1.00e+00	9.99e-01	0.88	0.91	0.82	0.82	SGD	1E-04	0.9	128
4	#1	1.40e-04	3.28e-03	1.00e+00	9.99e-01	0.86	0.90	0.82	0.82	SGD	1E-04	0.9	512
5	#1	8.23e-06	8.12e-04	1.00e+00	9.99e-01	1.04	0.90	0.78	0.82	SGD	1E-02	0.2	128
6	#1	4.86e-04	3.58e-03	1.00e+00	9.98e-01	0.91	0.92	0.80	0.81	SGD	1E-02	0.2	512
7	#1	1.47e-04	1.54e-03	1.00e+00	9.99e-01	0.27	0.27	0.91	0.91	SGD	1E-02	0.9	128
8	#1	1.48e-05	2.71e-03	1.00e+00	9.99e-01	1.04	0.91	0.76	0.82	SGD	1E-02	0.9	512
9	#1	0.00e+00	1.04e-03	1.00e+00	9.99e-01	0.20	0.27	0.94	0.91	ADAM	1E-04	0.2	128
10	#1	7.14e-05	2.22e-03	1.00e+00	9.99e-01	0.21	0.26	0.94	0.92	ADAM	1E-04	0.2	512
11	#1	5.42e-06	3.21e-03	1.00e+00	9.99e-01	0.23	0.27	0.90	0.92	ADAM	1E-04	0.9	128
12	#1	0.00e+00	3.64e-04	1.00e+00	1.00e+00	0.10	0.27	0.97	0.92	ADAM	1E-04	0.9	512
13	#1	2.06e-03	1.31e-02	1.00e+00	9.97e-01	0.23	0.30	0.90	0.89	ADAM	1E-02	0.2	128
14	#1	1.57e-04	2.17e-03	1.00e+00	9.99e-01	0.21	0.27	0.91	0.91	ADAM	1E-02	0.2	512
15	#1	0.00e+00	2.10e-02	1.00e+00	9.94e-01	0.36	0.27	0.88	0.91	ADAM	1E-02	0.9	128
16	#1	0.00e+00	1.98e-03	1.00e+00	9.99e-01	0.44	0.27	0.86	0.92	ADAM	1E-02	0.9	512
17	#2	6.00e-05	9.64e-03	1.00e+00	9.95e-01	0.25	0.31	0.92	0.91	SGD	1E-04	0.2	128
18	#2	0.00e+00	1.09e-02	1.00e+00	9.95e-01	0.23	0.30	0.94	0.91	SGD	1E-04	0.2	512
19	#2	8.94e-07	8.85e-03	1.00e+00	9.97e-01	0.13	0.29	0.96	0.91	SGD	1E-04	0.9	128
20	#2	0.00e+00	9.01e-03	1.00e+00	9.96e-01	0.17	0.29	0.93	0.92	SGD	1E-04	0.9	512
21	#2	4.17e-05	9.64e-03	1.00e+00	9.95e-01	0.24	0.31	0.93	0.91	SGD	1E-02	0.2	128
22	#2	0.00e+00	6.73e-03	1.00e+00	9.96e-01	0.24	0.29	0.95	0.91	SGD	1E-02	0.2	512
23	#2	0.00e+00	3.50e-03	1.00e+00	9.98e-01	0.23	0.28	0.91	0.92	SGD	1E-02	0.9	128
24	#2	0.00e+00	1.13e-02	1.00e+00	9.97e-01	0.22	0.29	0.94	0.91	SGD	1E-02	0.9	512
25	#2	3.49e-05	9.76e-03	1.00e+00	9.96e-01	0.18	0.27	0.94	0.92	ADAM	1E-04	0.2	128
26	#2	1.42e-05	7.30e-03	1.00e+00	9.97e-01	0.23	0.30	0.93	0.91	ADAM	1E-04	0.2	512
27	#2	1.29e-03	5.68e-03	1.00e+00	9.98e-01	0.31	0.28	0.90	0.92	ADAM	1E-04	0.9	128
28	#2	3.78e-04	6.25e-03	1.00e+00	9.95e-01	0.20	0.31	0.95	0.91	ADAM	1E-04	0.9	512
29	#2	0.00e+00	9.68e-03	1.00e+00	9.95e-01	0.28	0.28	0.89	0.90	ADAM	1E-02	0.2	128
30	#2	1.24e-05	8.70e-03	1.00e+00	9.97e-01	0.29	0.32	0.91	0.89	ADAM	1E-02	0.2	512
31	#2	0.00e+00	1.17e-02	1.00e+00	9.88e-01	0.24	0.28	0.90	0.91	ADAM	1E-02	0.9	128
32	#2	0.00e+00	7.02e-03	1.00e+00	9.98e-01	0.17	0.30	0.92	0.91	ADAM	1E-02	0.9	512

CONCLUSIONS

We performed an extensive and rigorous study on the parameters' tuning for the two different CNN models, which gives us a starting point to better understand the advantages and disadvantages of these models used in the seismological community.

We succeeded in training the CNN model to successfully differentiate between the earthquake and the noise event windows.

We also succeeded in training the CNN model to classify earthquakes based on two labels: the epicentral distance and the magnitude.

Using our models we should be able to detect new events in the continuous data and successfully label them as local earthquakes.