UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia Galileo Galilei

Temporal and spatial analysis of earthquakes in Italy in the last century

Jamilov Javlon, Pirazzo Tommaso, Secco Benedetto

14 luglio 2025

What is this project about?

In this work we studied the temporal and spatial analysis of earthquakes in Italy in the period 1925-2025. To achieve the goal of giving a complete and exhaustive picture of the phenomenon, we followed these steps:

What is this project about?

In this work we studied the temporal and spatial analysis of earthquakes in Italy in the period 1925-2025. To achieve the goal of giving a complete and exhaustive picture of the phenomenon, we followed these steps:

1. Prepare and inspect the data, Gutenberg-Richter Law;

What is this project about?

In this work we studied the temporal and spatial analysis of earthquakes in Italy in the period 1925-2025. To achieve the goal of giving a complete and exhaustive picture of the phenomenon, we followed these steps:

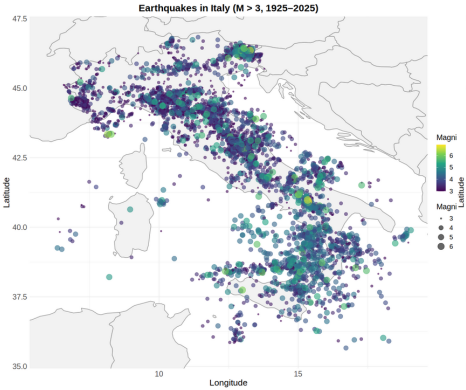
1. Prepare and inspect the data, Gutenberg-Richter Law;
2. Spatial analysis (Hazard map and hierarchical clustering);

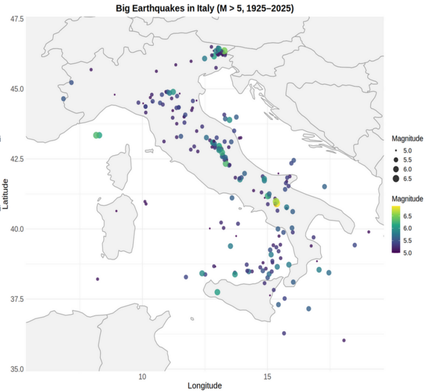
What is this project about?

In this work we studied the temporal and spatial analysis of earthquakes in Italy in the period 1925-2025. To achieve the goal of giving a complete and exhaustive picture of the phenomenon, we followed these steps:

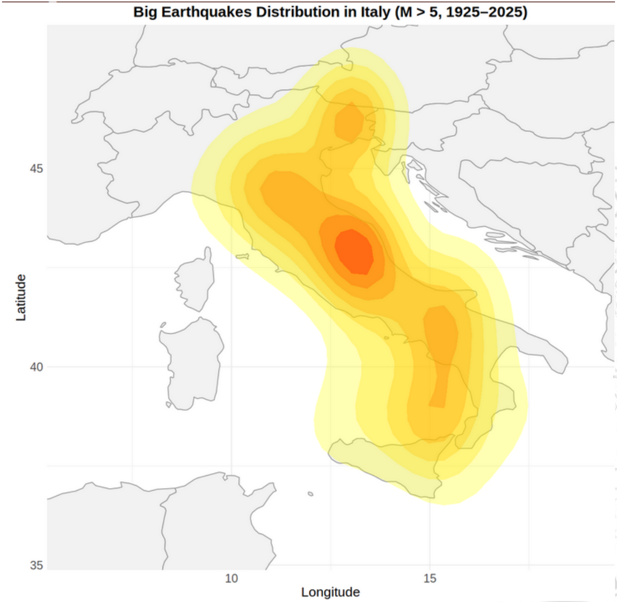
1. Prepare and inspect the data, Gutenberg-Richter Law;
2. Spatial analysis (Hazard map and hierarchical clustering);
3. Temporal analysis (Forecast with ARIMA and SSA model);

Map of the data

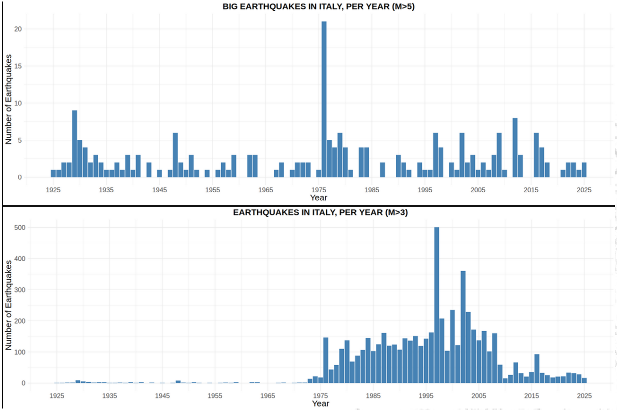




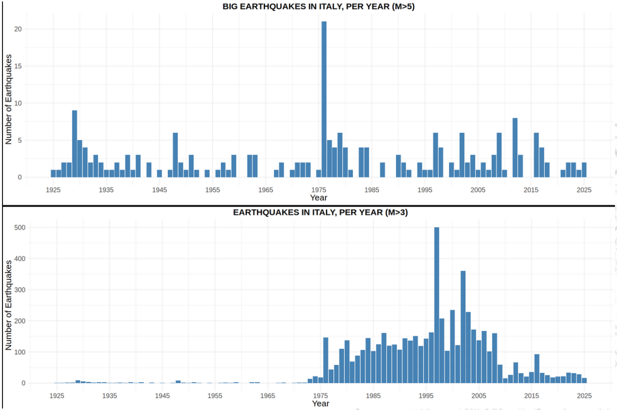
Focus on high magnitude events



Yearly distribution of events

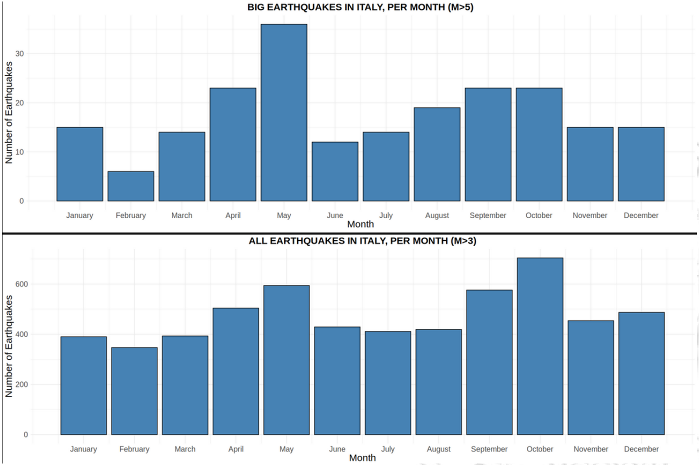


Yearly distribution of events

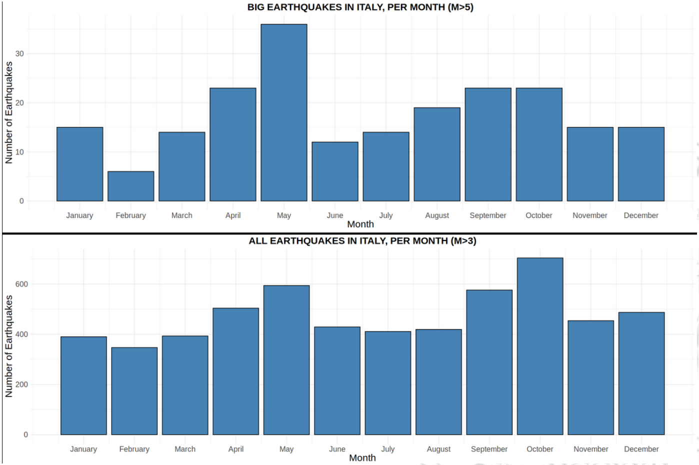


There are many more earthquake registered from 1973 onwards.
Our guess is that the reason is better technology.

Seasonality

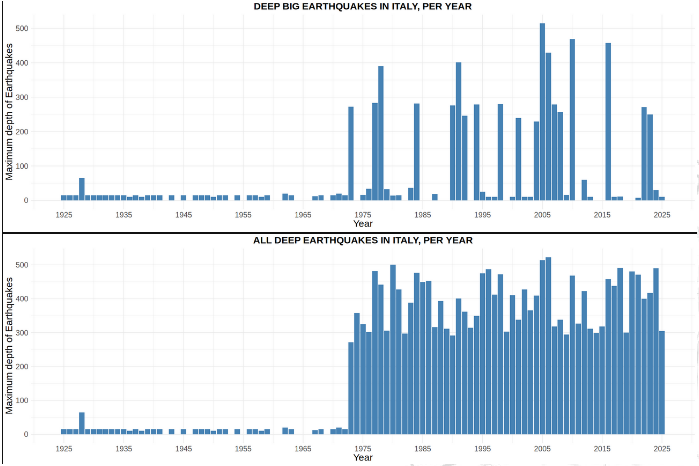


Seasonality

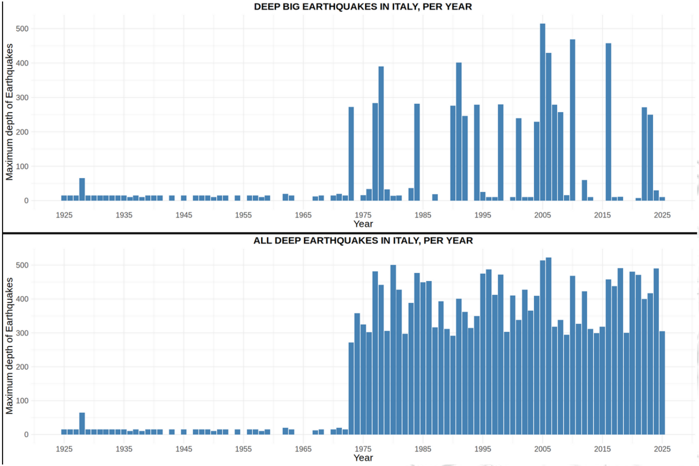


There is no evident seasonal ditribution.

Depth histograms

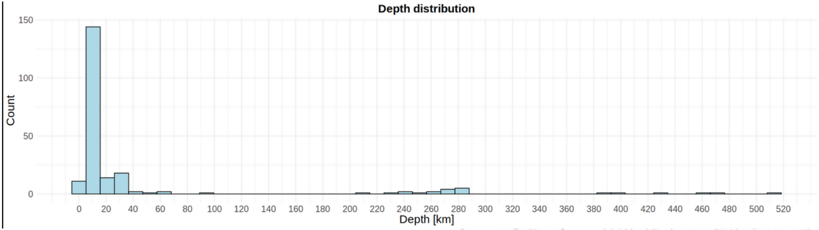


Depth histograms



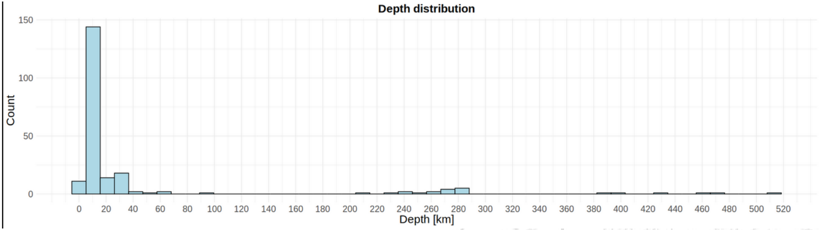
Deep earthquakes in Italy are only recorded from 1973 onward, likely reflecting improved seismic monitoring. Further analysis would be needed to determine whether this reflects a true onset of activity or increased detection capability.

Depth histograms



From the histogram we can identify three main depth zones:

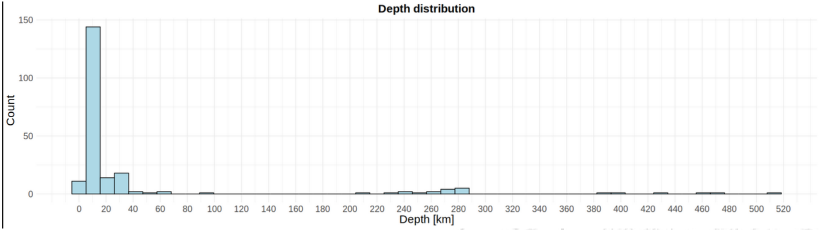
Depth histograms



From the histogram we can identify three main depth zones:

- $0 - 100$ km, with a peak at low depth

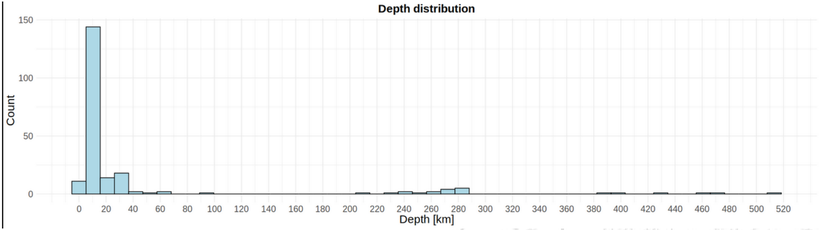
Depth histograms



From the histogram we can identify three main depth zones:

- $0 - 100$ km, with a peak at low depth
- $200 - 300$ km

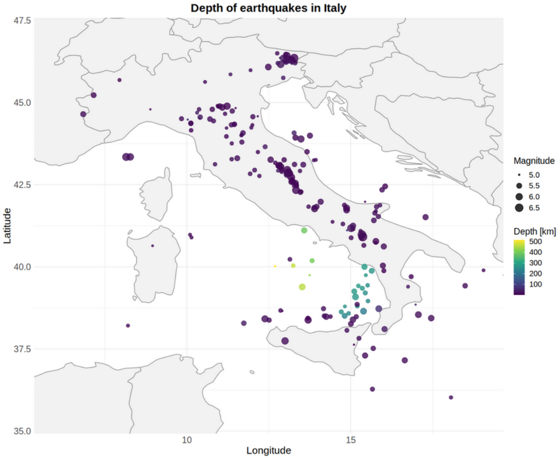
Depth histograms



From the histogram we can identify three main depth zones:

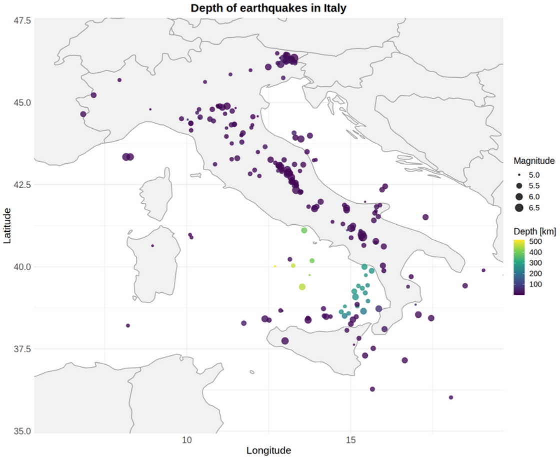
- $0 - 100$ km, with a peak at low depth
- $200 - 300$ km
- $380 - 520$ km

Depth map



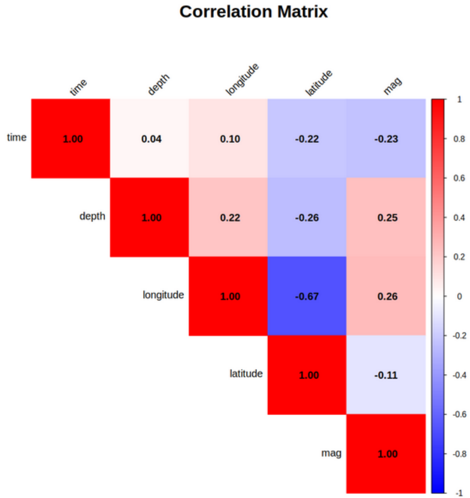
Depth of earthquakes in Italy
47.5
45.0
42.5
40.0
37.5
35.0
Latitude
10
15
Longitude
Magnitude
5.0
5.5
6.0
6.5
Depth [km]
500
400
300
200
100

Depth map

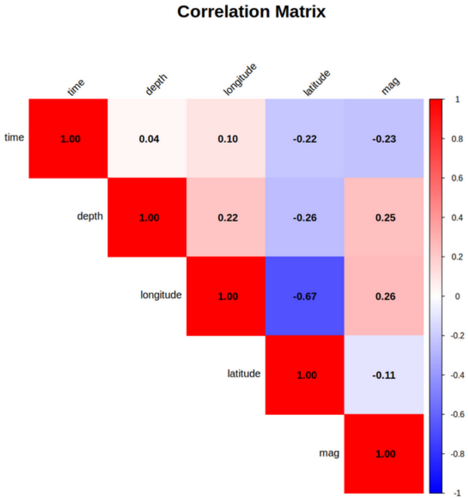


The last two zones coincide exactly with the high depth zone seen in the map. The deep earthquakes we see in the sea northwest of Sicily and west of Calabria are almost certainly due to the subduction of the Ionian plate beneath the Calabrian arc.

Variables correlation

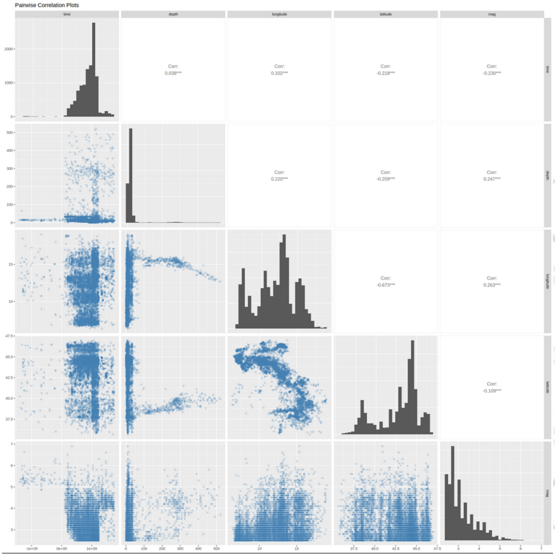


Variables correlation

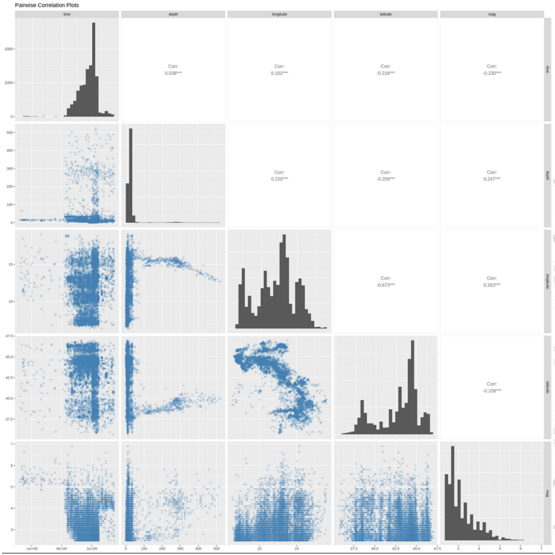


The strongest (anti)correlation is observed between latitude and longitude, likely reflecting the geographical alignment of the Apennine mountain range, which follows a tectonic plate boundary.

Variables correlation plots



Variables correlation plots



- Most of the data were recorded after 1973.
- Most earthquakes occur near the surface.
- Longitude is correlated with latitude, as mentioned earlier.
- Most earthquakes have low magnitudes.

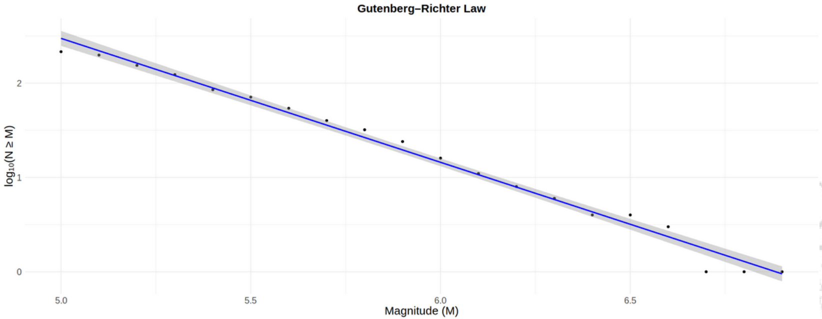
The Gutenberg-Richter Law

In seismology, the Gutenberg–Richter Law (GR law) expresses the relationship between the magnitude and the total number of earthquakes in any given region and time period of at least that magnitude, and is given by:

$$log_{10} N = a - b \cdot M$$

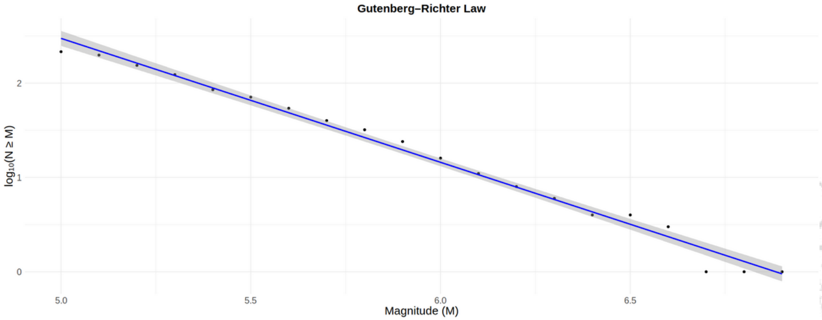
where N is the number of events having a magnitude M ,a and b are constants. The values of a and b may vary significantly from region to region or over time.

The Gutenberg-Richter Law



	Value	σ
a	9.0	0.2
b	1.31	0.03

The Gutenberg-Richter Law



	Value	σ
a	9.0	0.2
b	1.31	0.03

A b-value around 1 is typical globally; this slightly higher value suggests a higher proportion of small earthquakes compared to large ones.

Spatial analysis

Spatial analysis

In order to obtain a sismic map based on the Gutenberg-Richter law We followed these steps region by region:

Spatial analysis

In order to obtain a sismic map based on the Gutenberg-Richter law We followed these steps region by region:

1. Divide the region into grid cells;

Spatial analysis

In order to obtain a sismic map based on the Gutenberg-Richter law We followed these steps region by region:

1. Divide the region into grid cells;
2. Count the number of events above magnitude 4 (Trade-off) for each cell;

Spatial analysis

In order to obtain a sismic map based on the Gutenberg-Richter law We followed these steps region by region:

1. Divide the region into grid cells;
2. Count the number of events above magnitude 4 (Trade-off) for each cell;
3. Fit the G-R law using linear regression on $log_{10}(N)$ vs. M with the local data;

Spatial analysis

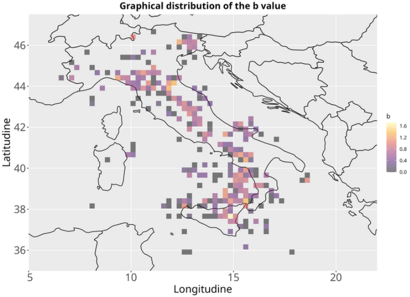
In order to obtain a sismic map based on the Gutenberg-Richter law We followed these steps region by region:

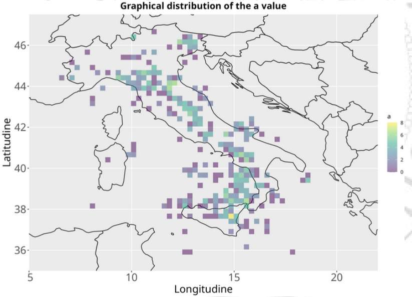
1. Divide the region into grid cells;
2. Count the number of events above magnitude 4 (Trade-off) for each cell;
3. Fit the G-R law using linear regression on $log_{10}(N)$ vs. M with the local data;
4. Derive local a and b values;

Spatial analysis

In order to obtain a sismic map based on the Gutenberg-Richter law We followed these steps region by region:

1. Divide the region into grid cells;
2. Count the number of events above magnitude 4 (Trade-off) for each cell;
3. Fit the G-R law using linear regression on $log_{10}(N)$ vs. M with the local data;
4. Derive local a and b values;

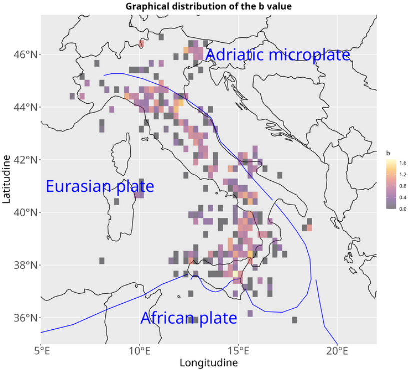




Adding tectonic plates boundaries

```
# https://www.usgs.gov/programs/earthquake-hazards/google-earthtmkml-files -> page where you can find the file for the edge of the plates

plates <- st_read("doc.kml", layer = "Plate Interface")
                ...

# add tectonic plates boundaries
a_plot_plates <- ggplot() +

        # borders
        borders("world",
        regions = c(reg),
        fill = "gray80", colour = "gray10", alpha = 0) +

        # tectonic plates boundaries layer
        geom_sf(data = plates, color = "blue", size = 2) +
                ...
```

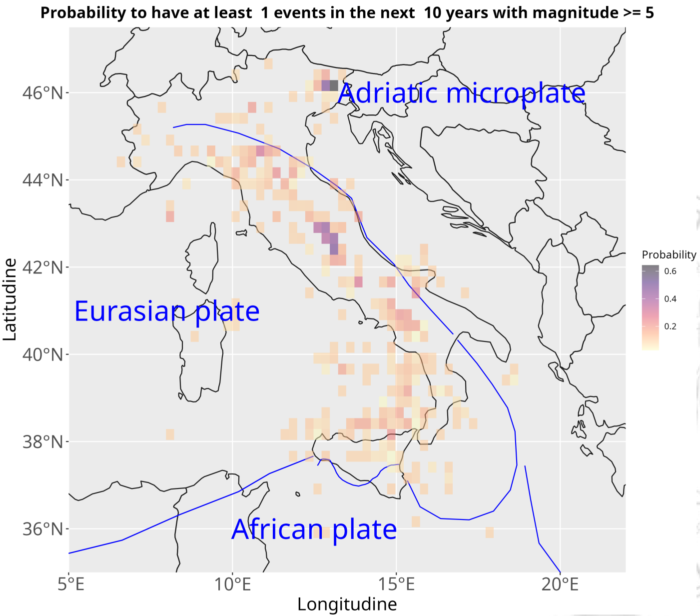
Adding tectonic plates boundaries



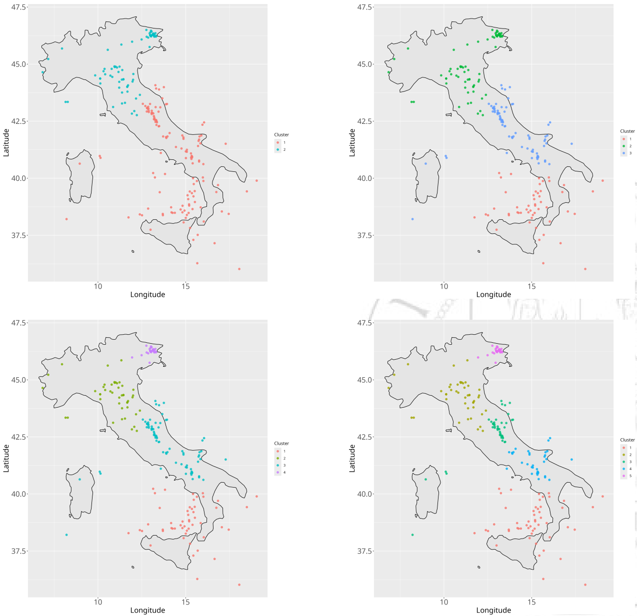
A simple model for sismic hazard

In order to provide an exstimation of the sismic hazard in Italy, We calculated the probability to have at least 1 high magnitude event in the next 10 years assuming that the intense events follw a Poisson distribution.

Naive model for sismic hazard



Hierarchical clustering



Descriptive Statistics of AEM and MEM

Statistic	AEM	MEM
Min	2.952	5.100
Max	5.935	7.000
Mean	4.530	5.854
Median	4.375	5.770
Standard Deviation	0.993	0.443

Table: Descriptive statistics for Average (AEM) and Maximum (MEM) Earthquake Magnitudes.

AEM and MEM Trend Analysis

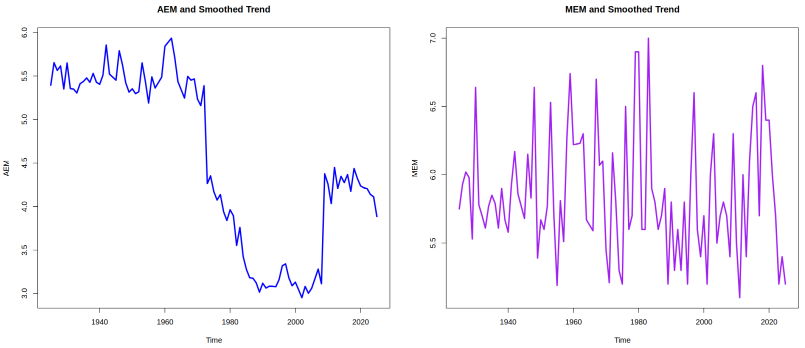


Figure: The smoothed trends of AEM (left) and MEM (right) magnitudes across years. AEM shows a significant long-term decreasing trend, especially after the 1970s. MEM remains more volatile, with frequent high-magnitude spikes throughout the century.

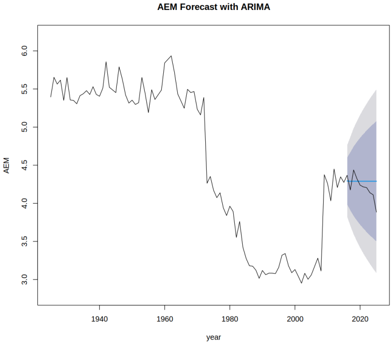
What is ARIMA?

- **ARIMA** stands for **AutoRegressive Integrated Moving Average**.
- It is one of the most widely used models in time series forecasting.
- ARIMA combines three components:
  - **AR (AutoRegressive)**: uses the dependency between current and past values.
  - **I (Integrated)**: applies differencing to make the time series stationary.
  - **MA (Moving Average)**: uses the relationship between an observation and residual errors from previous steps.
- In this project, ARIMA is used to forecast future earthquake magnitudes (AEM and MEM).

Forecasting Strategy with ARIMA

- To test ARIMA model performance, we trained it using two different data ranges:
  1. **Model 1:** Trained on data from 1925 to 2015, then forecasted 2016–2025.
  2. **Model 2:** Trained on full data from 1925 to 2025, then forecasted 2026–2035.
- This setup allows us to:
  - Evaluate the model's accuracy by comparing the first forecast (2016–2025) with real observed values.
  - Produce long-term projections from the full historical data.
- Forecasts are shown for both AEM (average magnitude) and MEM (maximum magnitude).

ARIMA Forecast: Until 2015



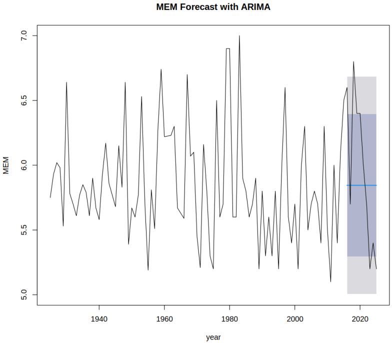
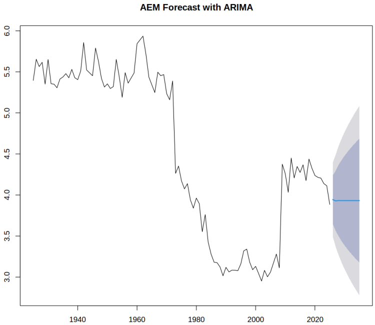


Figure: ARIMA forecast for AEM and MEM using data from 1925 to 2015. The model predicts the next 10 years (2016–2025), which can be compared with actual observed data in that period.

ARIMA forecast: until 2025



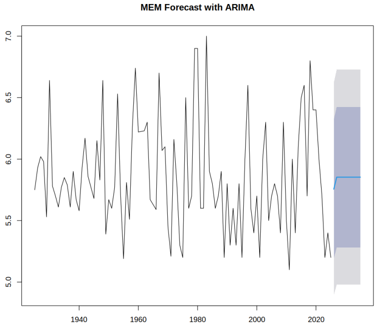


Figure: ARIMA forecast for AEM and MEM using full historical data up to 2025. This forecast projects earthquake magnitudes for 2026–2035.

Residual Diagnostics for ARIMA(1,1,0) — AEM

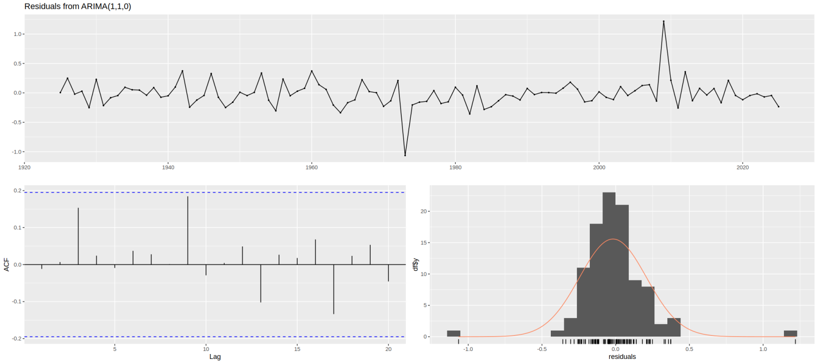


Figure: Residual analysis of ARIMA(1,1,0) fitted to AEM. **Top:** Residuals over time, showing mostly random variation around zero. **Bottom Left:** Autocorrelation (ACF) of residuals remains within confidence bounds, suggesting no significant autocorrelation. **Bottom Right:** Histogram of residuals approximates a normal distribution. This indicates the model assumptions are reasonably satisfied.

Residual Diagnostics for ARIMA(0,0,1) - MEM

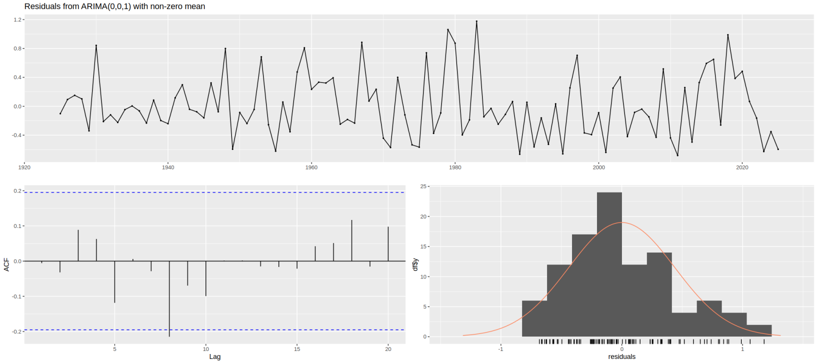


Figure: Residual analysis of ARIMA(0,0,1) fitted to MEM. **Top:** Residuals show some variation but remain centered around zero. **Bottom Left:** ACF indicates slightly more structure than desired, suggesting mild autocorrelation may remain. **Bottom Right:** Histogram shows approximate normality but with heavier tails. Model may require further refinement.

What is SSA?

- **Singular Spectrum Analysis (SSA)** is a powerful, non-parametric method for time series decomposition and forecasting.
- SSA breaks a time series into interpretable components:
  - **Trend**
  - **Seasonality / Periodic Patterns**
  - **Noise (Residuals)**
- It uses linear algebra techniques such as:
  - Embedding the time series into a matrix (trajectory matrix)
  - Performing **Singular Value Decomposition (SVD)**
  - Reconstructing components based on grouped eigenvectors
- SSA is especially useful for irregular or noisy time series — like earthquake magnitudes.

Why SSA for Earthquake Magnitudes?

- Earthquake time series are often irregular and non-stationary.
- Traditional models (like ARIMA) assume linear structure and stationarity.
- SSA:
  - Does not require a statistical distribution
  - Captures hidden patterns such as slow-moving trends or sudden shocks
  - Can provide more flexible and accurate long-term forecasts
- SSA is well-suited for detecting subtle trends in average (AEM) and extreme (MEM) magnitudes over time.

W-Correlation Matrix (AEM)

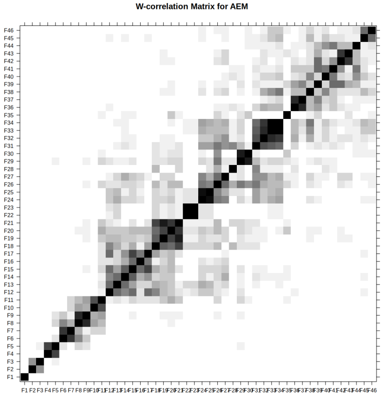
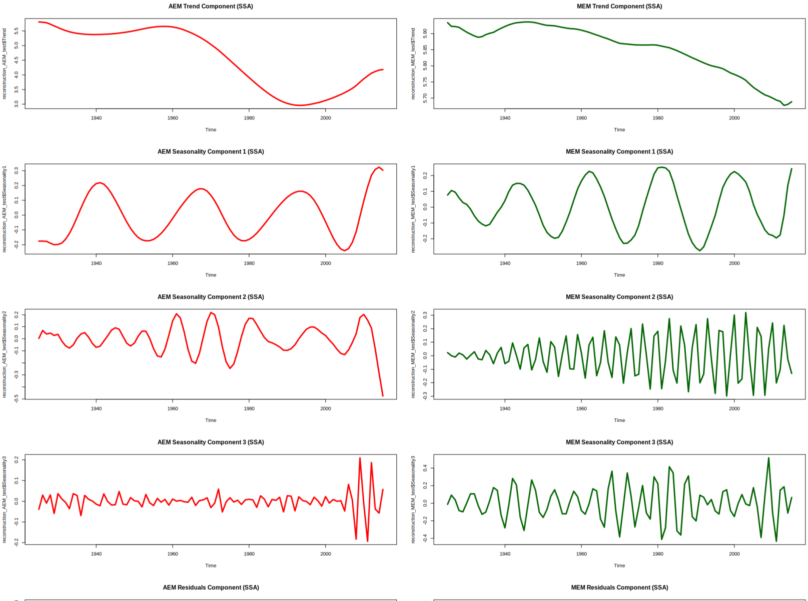


Figure: W-correlation matrix for AEM components (F1–F46) using data until 2015. Darker blocks along the diagonal indicate strong correlation between consecutive components. This helps identify groups for reconstructing trend and seasonality, noise. Component pairs (e.g., F1–F2, F3–F4) were grouped accordingly.

SSA Component Decomposition



SSA Forecast: 2016–2025

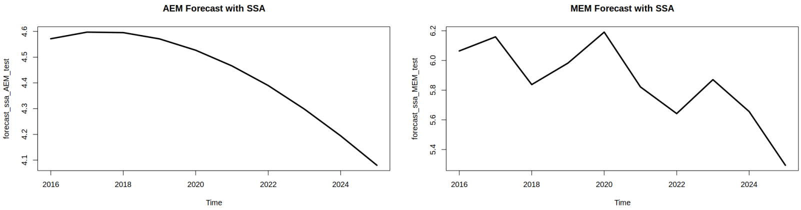
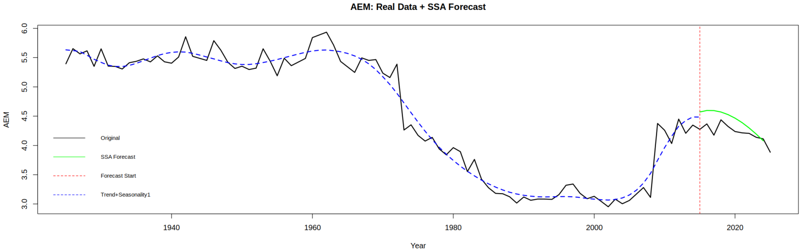
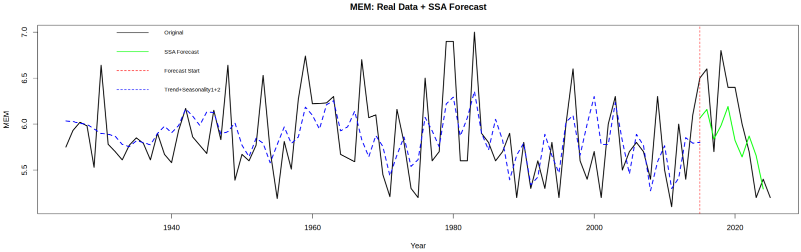


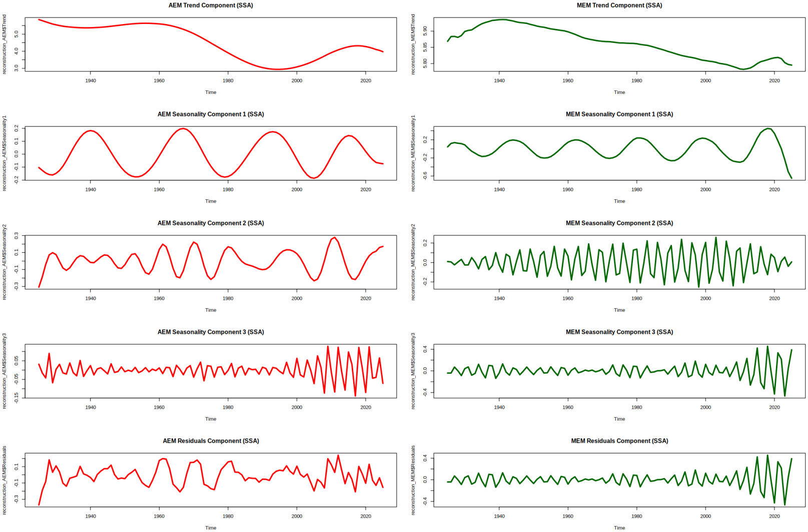
Figure: Forecasted AEM and MEM values for 2016–2025 using SSA. **Left:** AEM shows a steady decrease, indicating a potential reduction in average earthquake magnitude. **Right:** MEM forecast remains variable but generally declining, with signs of future high-magnitude events reducing in strength.

SSA Forecast vs Real Data





SSA Components 2025



SSA Forecast: 2026–2035

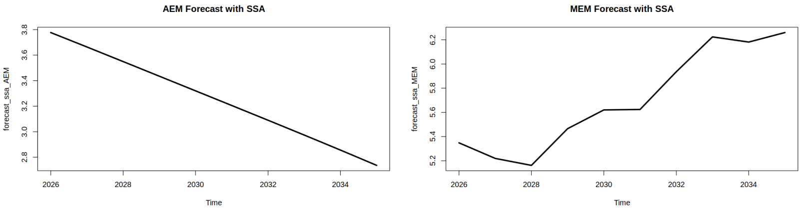
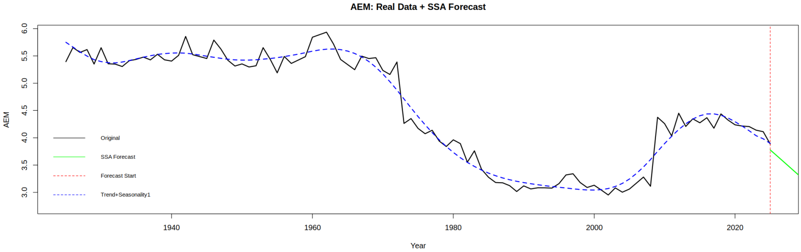
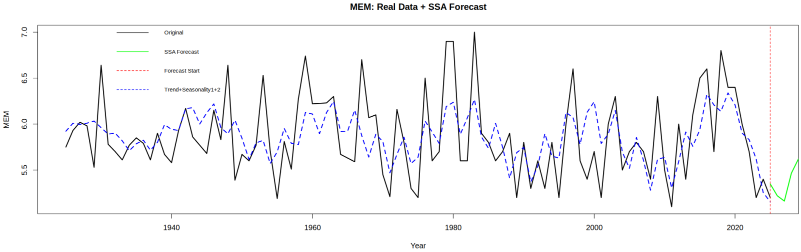


Figure: Forecasted AEM and MEM magnitudes for 2026–2035 using SSA trained on data up to 2025. AEM continues its decreasing trend, falling below 3.0. MEM shows an upward rebound after 2028, possibly indicating increasing seismic energy release in coming years.

SSA Forecast with Historical Data





Comparison of the Two Methods Using RMSE

RMSE (Root Mean Square Error) is used to evaluate the accuracy of forecasting models.

For n data points:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

- y_i: actual observed value
- $\hat{y}_i$: predicted value

Interpretation:

- Lower RMSE indicates a better model fit.
- Assumes residuals behave like white noise (uncorrelated and centered around 0).

RMSE Comparison: ARIMA vs SSA

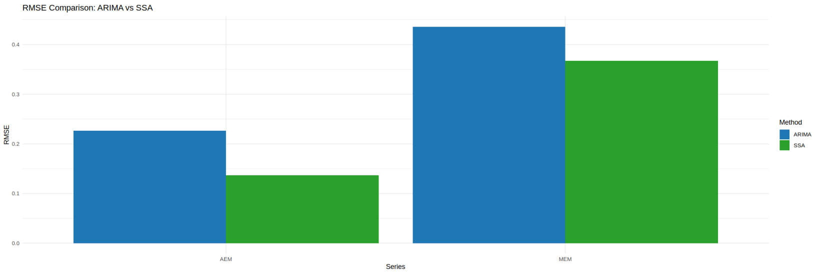


Figure: Root Mean Square Error (RMSE) comparison between ARIMA and SSA models. SSA consistently outperformed ARIMA with lower RMSE values for both AEM and MEM series. This demonstrates SSA's higher predictive power and robustness in modeling earthquake magnitudes.

Conclusion

- We analyzed the temporal characteristics of earthquake magnitudes in Italy from 1925 to 2025.
- Two forecasting models were applied:
  - **ARIMA**: classic, linear, good short-term performance.
  - **SSA**: non-parametric, capable of capturing complex structures.
- **Key Findings:**
  - SSA achieved lower RMSE than ARIMA for both AEM and MEM.
  - AEM is projected to gradually decrease further, while MEM remains volatile but shows possible decline.
  - SSA offers more reliable long-term forecasts in seismic time series with irregular patterns.

Thanks for your attention!