September 2021

Third Year Activity Report (2020-2021)

Valerio De Rubeis

Introduction

This third year of activity of the Working Group concludes a data analysis and management process, which characterised the first two years, with the definition of a new process of homogenisation of macroseismic data, collected with different methods by the various institutions or agencies. The earthquake in Croatia, Mw = 6.3, on Dec 29, 2020, offered an excellent opportunity for this task. This seismic event involved several European and international groups (USGS and EMSC), producing different macroseismic datasets. This earthquake constituted an excellent example of the intensity data of a significant earthquake, collected with heterogeneous methods and analysis procedures. It was essential to carry out an evaluation and an intent to develop an integration method.

We presented this research at the recent 37th General Assembly of the European Seismological Commission.

With this third year of activity, the evaluation cycle on the problems relating to the homogenisation and integration of macroseismic data in Europe is closed. Indeed, the topic is by no means concluded nor exhausted; however, we believe that with this third year, the preliminary phase ends which, we hope, will open a new one based on all the experiences acquired.

Also, within this year, there have been difficulties mainly deriving from the COVID-19 pandemic: although the activity has continued, all interested parties will have to collect and make their own experiences deriving from this commitment.

We take this opportunity to thank all of those who have directly or indirectly participated in this project. Concerning this third and final year, we would particularly like to thank all those who have provided us with the intensity data relating to the Croatian earthquake: without their significant commitment to collecting and analysing, this analysis would not have had the slightest chance of being developed.

Merging different country-institution macroseismic data for the Croatian Mw=6.3, Dec 29, 2020 earthquake: a way to compare and attain mutuality between heterogeneous intensity datasets

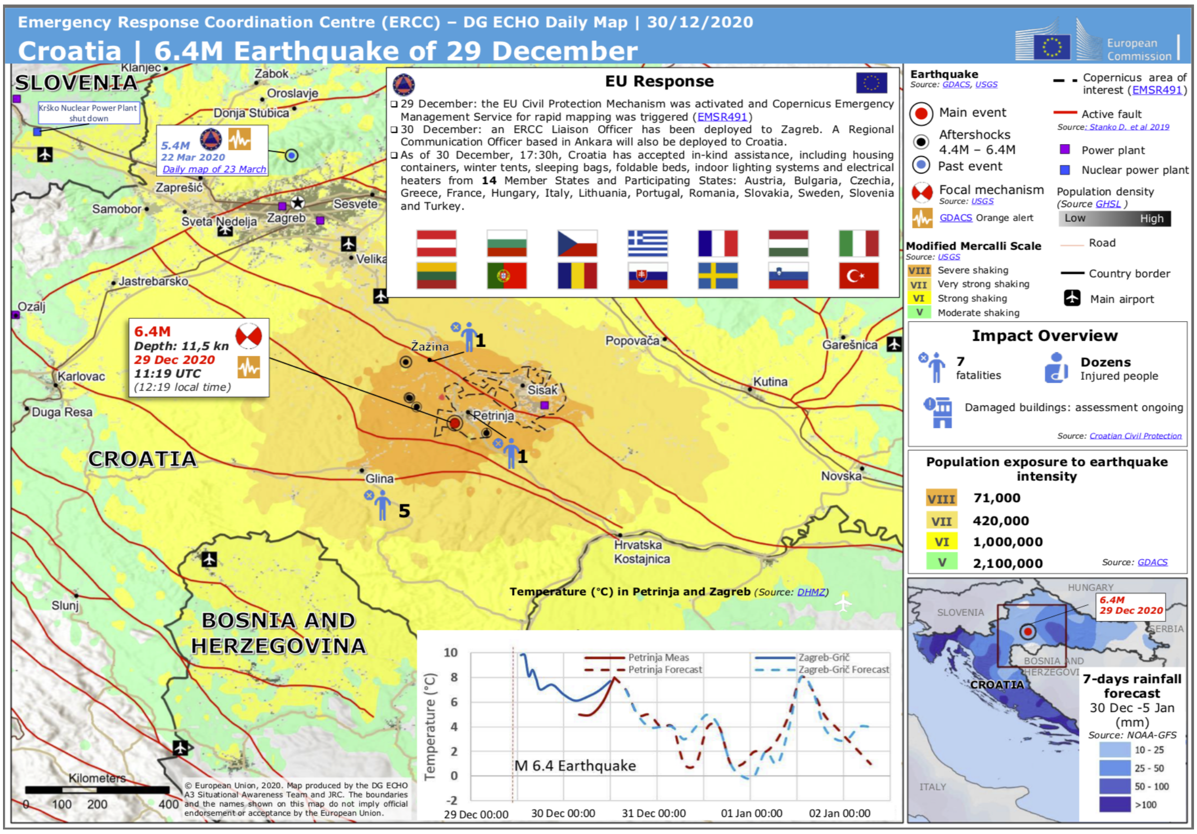


Fig.1 Emergency Response Coordination Centre (EU Civil Protection and Humanitarian Aid) map of the Croatia 6.4M earthquake of Dec 29 2020.

This section focuses on homogenising differing institution data for a specific Mw 6.3 earthquake in the Petrinja, Croatia region, on Dec 29, 2020. The reason to chose this particular earthquake was due to a magnitude great enough to involve several countries, thus the inclusion of numerous institutes-agencies, varying from strictly-country limited institutes such as KRSO (Kvesligethy Rad Seismological Observatory, Institute of Earth Physics and Space Science, Hungary), HSIT (Hai Sentito Il Terremoto - Did You Feel The Earthquake, Istituto Nazionale di Geofisica e Vulcanologia, Italy) and ARSO (Agencija Republike Slovenije Za Okolje, Slovenian Environmental Agency), as well as institutes acting internationally like USGS (United States Geological Service) for the whole world and EMSC

(European-Mediterranean Seismological Centre), mainly focused on Europe, and so for this earthquake, their data covers a broader area not limited to a specific country. On the other hand, local institutes contain data limited to their own respective country, in which they put all their experience in dealing with the public, as they are trusted. However, more importantly, they have direct experience in specific settings related to their countries, such as building structures and geological knowledge. However, their data is limited due to information only within their country, and in our case of a strong event, all limited country data maybe not homogeneous, and they may present problems when merged directly.

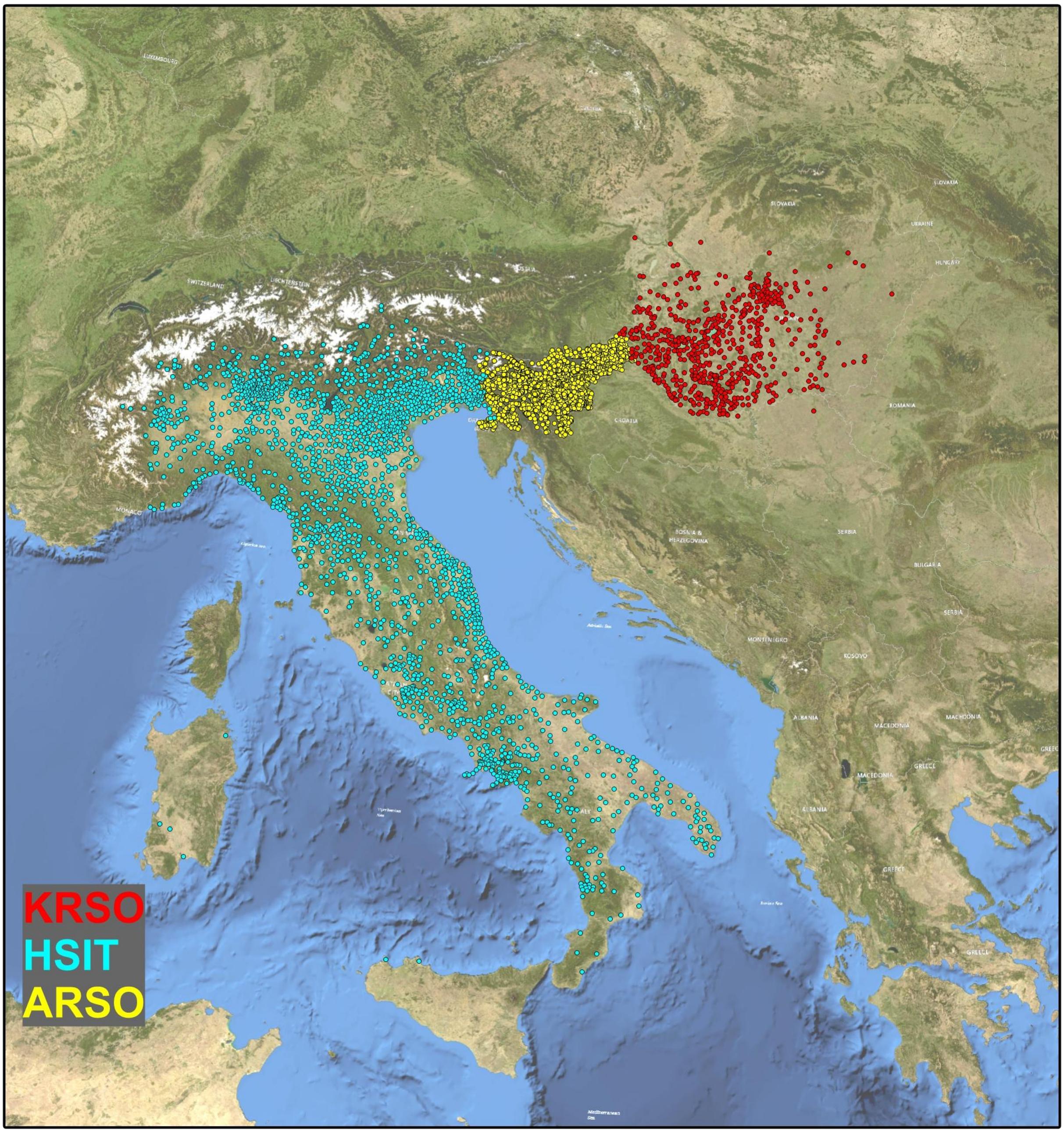


Fig. 2 Intensities distribution from country limited institutes such as KRSO (Kvesligethy Rad Seismological Observatory, Institute of Earth Physics and Space Science, Hungary), HSIT (Hai Sentito Il

Terremoto Did You Feel The Earthquake, Istituto Nazionale di Geofisica e Vulcanologia, Italy) and ARSO (Agencija Republike Slovenije Za Okolje, Slovenian Environmental Agency).

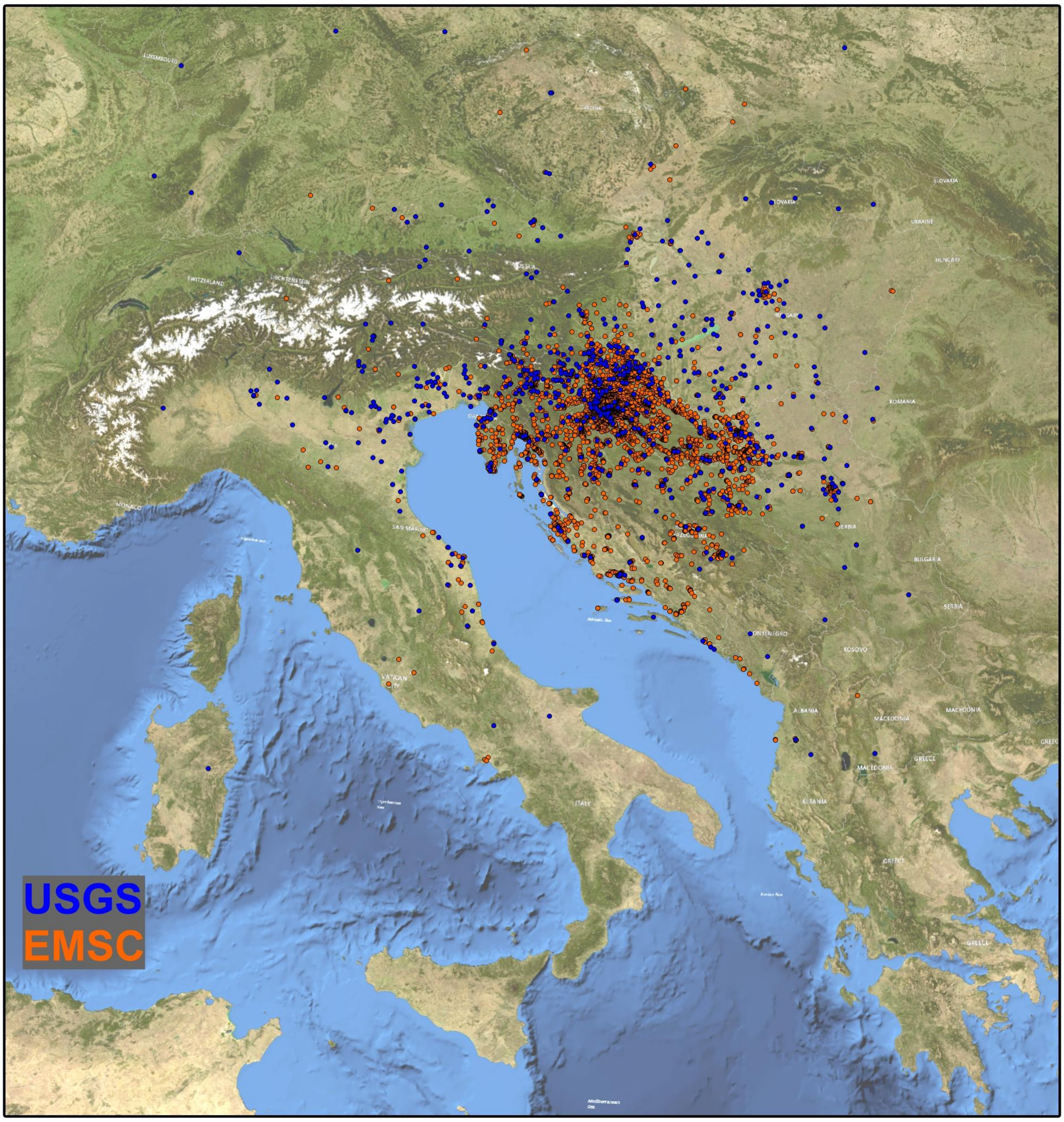


Fig. 3 Intensities distribution from institutes acting internationally like USGS (United States Geological Service) for the whole world and EMSC (European-Mediterranean Seismological Centre), mainly focused on Europe.

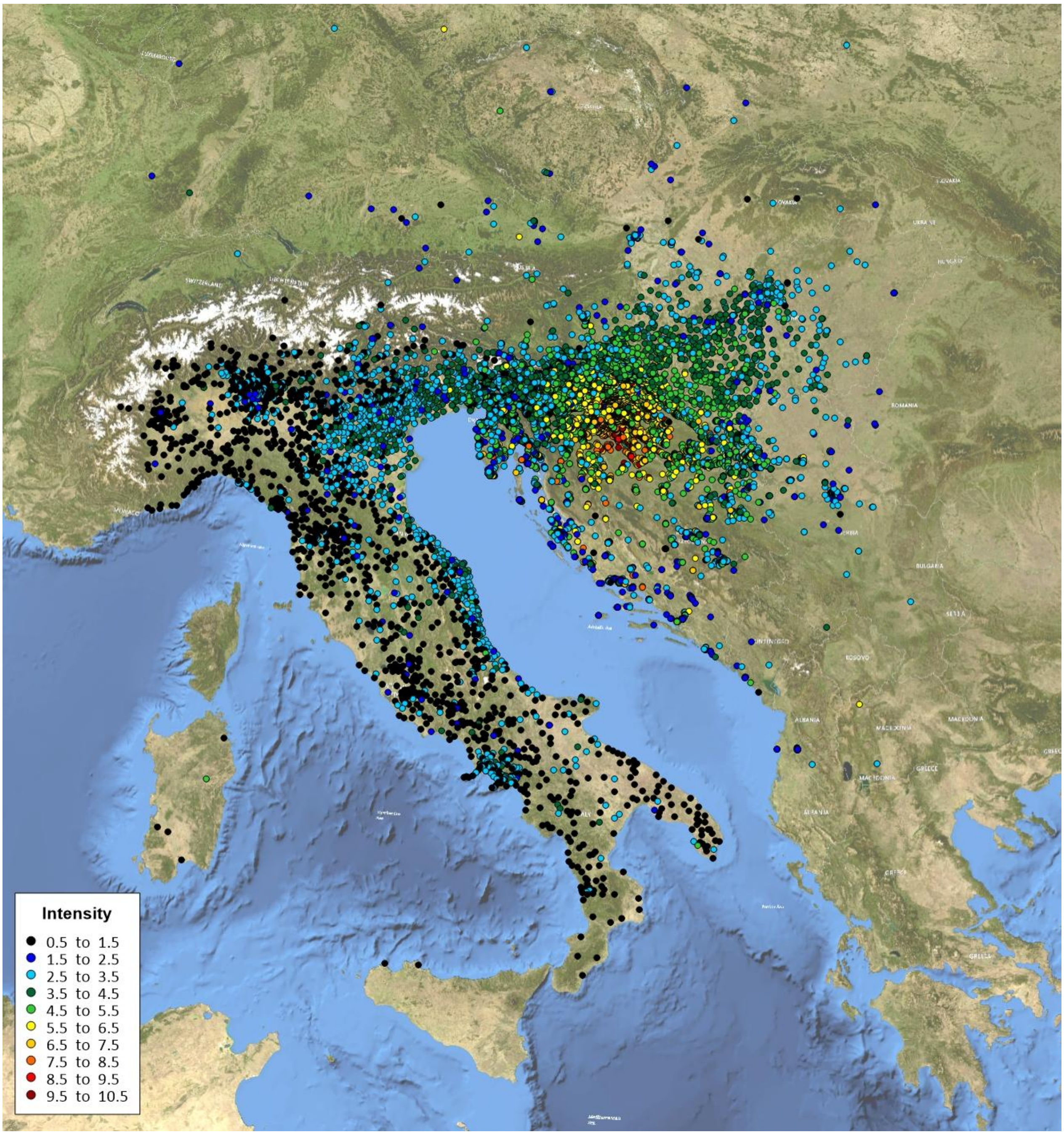


Fig. 4 All raw intensities as given by all considered Institutes. Colours indicate intensity degree.

This plot displays geographically distributed raw, unprocessed intensity data for all institutes. As shown, there is a clear pattern correctly exhibiting the epicentral area due to the higher intensities displayed. However, areas can be seen with higher intensities that do not conform to this general, first approximation rule, for example, the Padana Valley in Northern Italy and other limited areas.

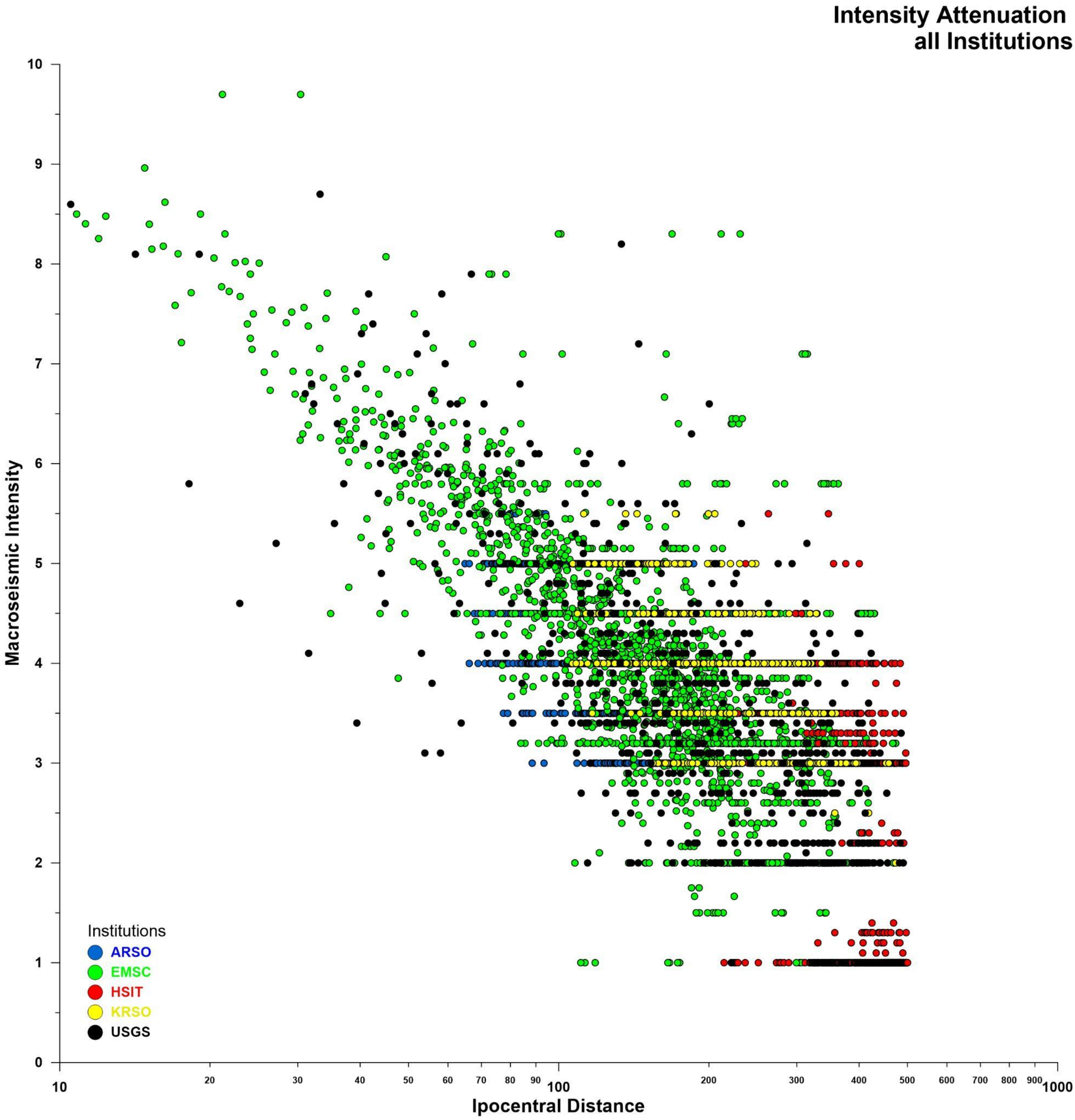


Fig. 5 Raw macroseismic intensities as function of log ipocentral distance (km) divided by Institutions.

An attenuation plot for original raw intensities indicates, on a log scale, roughly a first-degree fit. However, there is a large amount of spreading within data. We can think to reasonably model intensities constituted by different unrelated components:

1. a general log distance-attenuation first-degree decay,
2. local deviations in regions due to amplification and-or attenuation, like in the Padana valley (Italy),
3. the always present random noise component due, for example, to any error in data acquisition and processing, anthropogenic or very spatially limited, spatially independent geophysical factors.

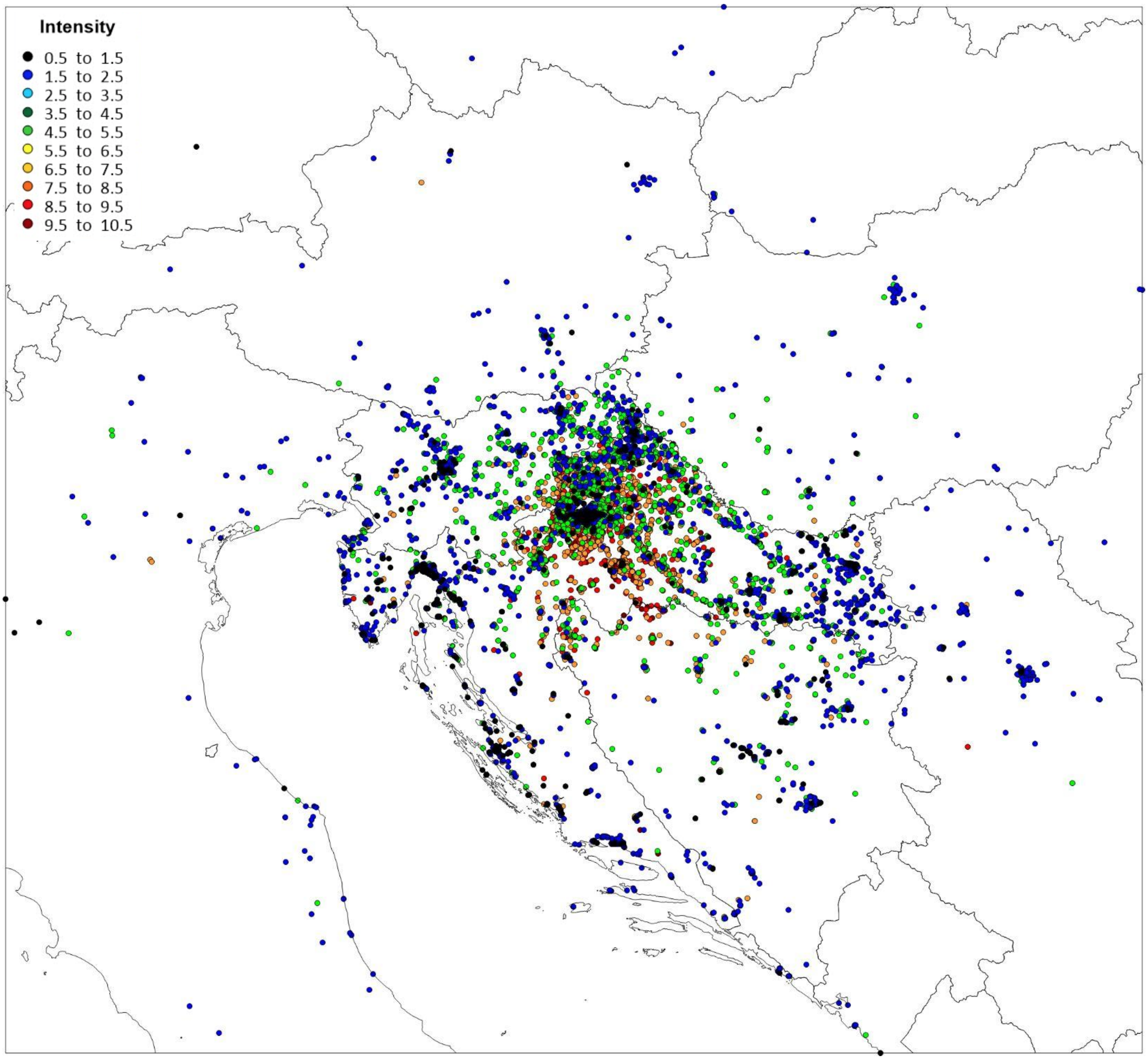


Fig. 6 EMSC raw intensities.

We analysed characteristics of specific data sets, whereby EMSC data includes many intensities characterised by high local variation and very different spatial density distribution. The main aim is to homogenise such data as noise reduction and more even geographical distribution.

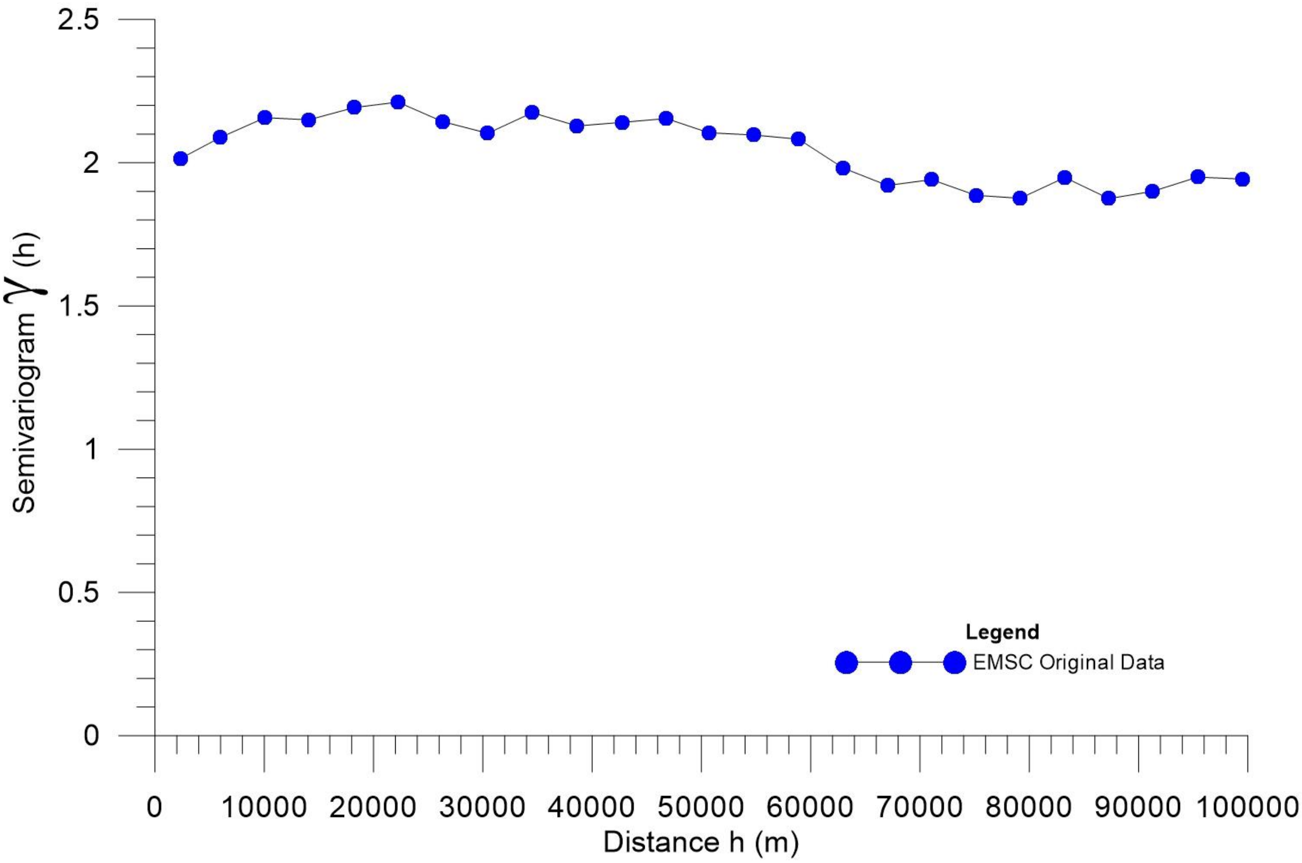


Fig. 7 Semivariogram as function of inter distances (m) among EMSC raw intensities.

Suppose we apply a semivariogram analysis to the macroseismic intensities as a function of their spatial inter distances. In that case, we see that the level of semivariance appears to be quite independent by the spatial closeness of intensities, indicating little or no spatial autocorrelation pattern. This result is typical of a nugget effect and, more generally, by the high content of the noise.

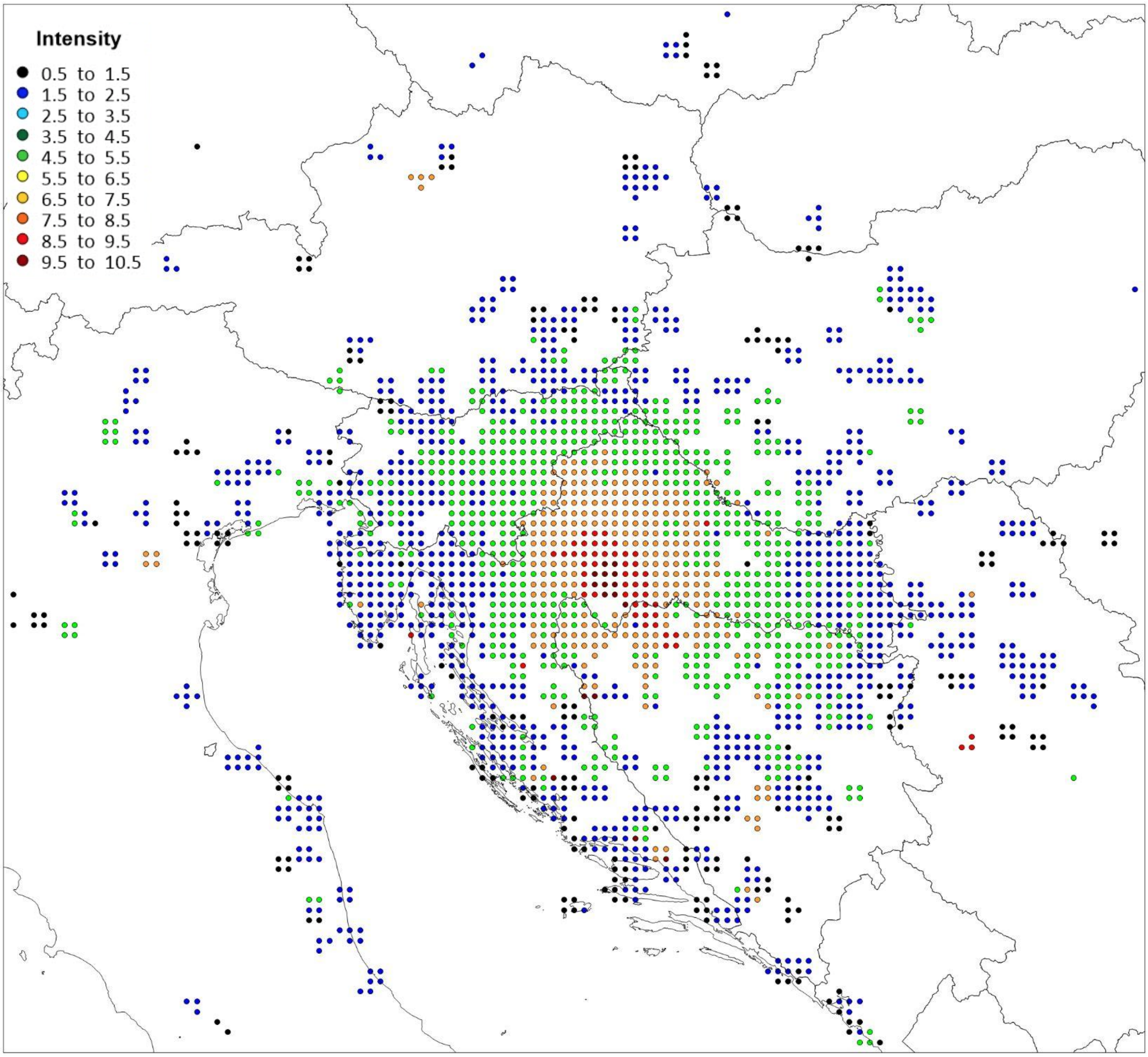


Fig. 8 EMSC filtered (windows averaged) intensities on a regular grid.

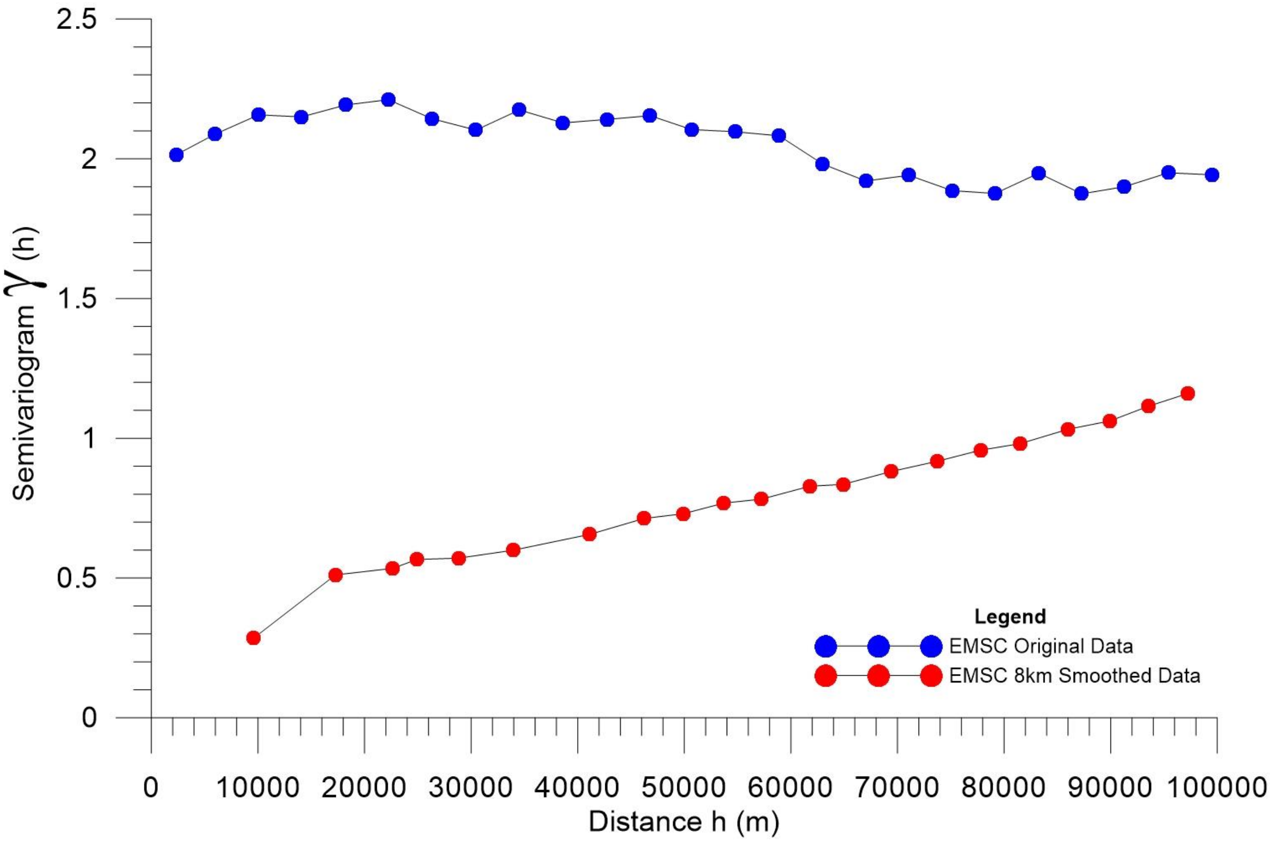


Fig. 9 Semivariograms as a function of inter distances (m) among EMSC raw intensities (blue dots) and filtered intensities (red dots).

We dealt with such EMSC data problems to resolve uneven spatial distribution and high levels of variance issues. A chosen resolution was an 8 km window width average to reduce the number of data and limit the variance and correct the uneven spatial distribution. The newly transformed data on the semivariogram expresses the disappearance of the nugget effect (Fig. 9), indicating the intensity spatial autocorrelation which we could expect in such data.

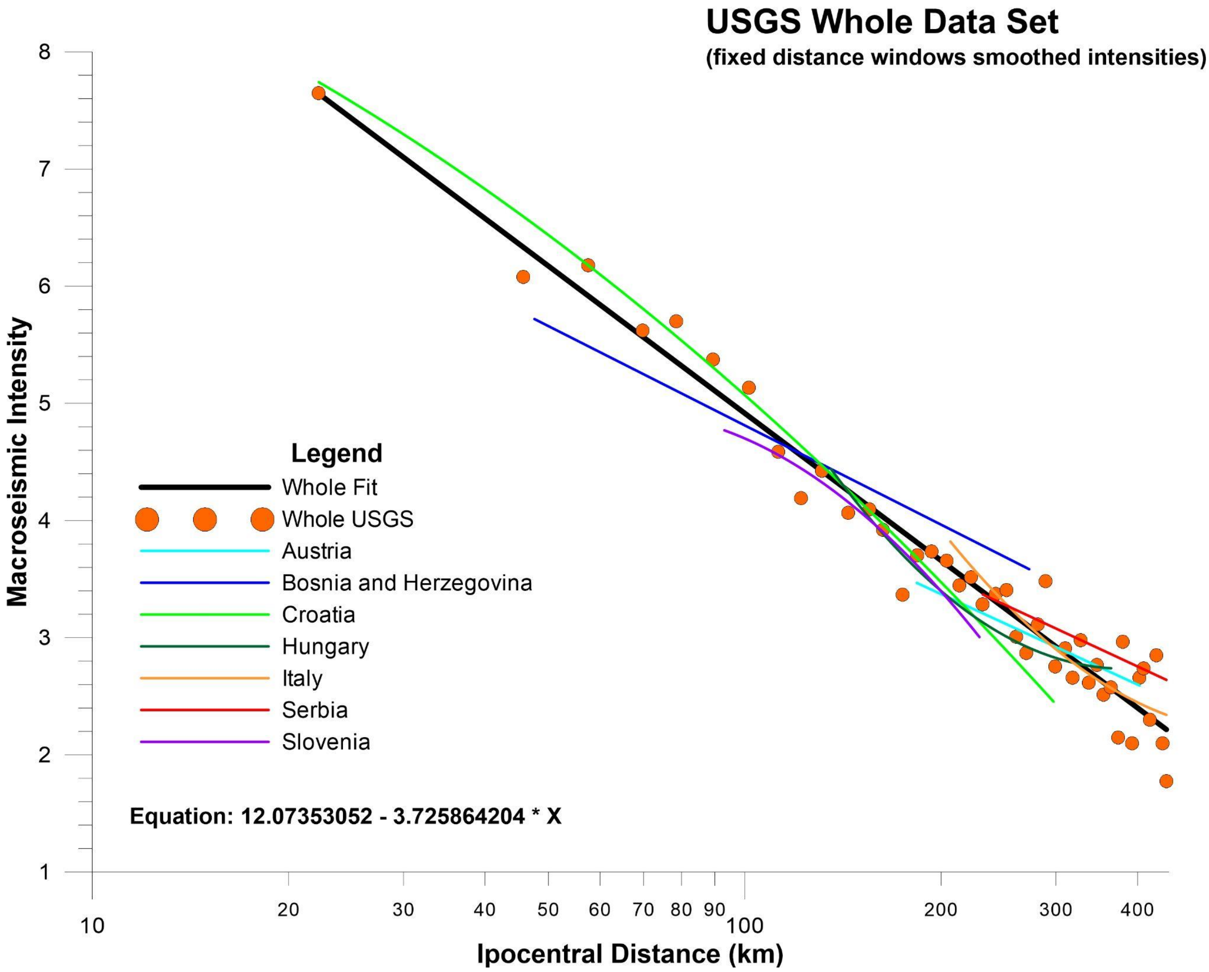


Fig. 10 USGS intensities data represented as a whole (orange dots and black first-degree polynomial fit) compared with the country separated independent fits.

We needed to work out if, dealing with intensity attenuation by source distance patterns, the presence of same distance range differences in intensity are due to specific areas or by methods of analysing within institutes. In order to limit the random noise component, we decided to average all USGS intensities inside hypocentral distance windows. It resulted in an evident general log first-degree fit.

Using the same USGS data as before but splitting them according to their belonging country, we see that differences are not so high, suggesting the country of origin may be less effective than differences introduced by different institutions methods (Fig. 10).

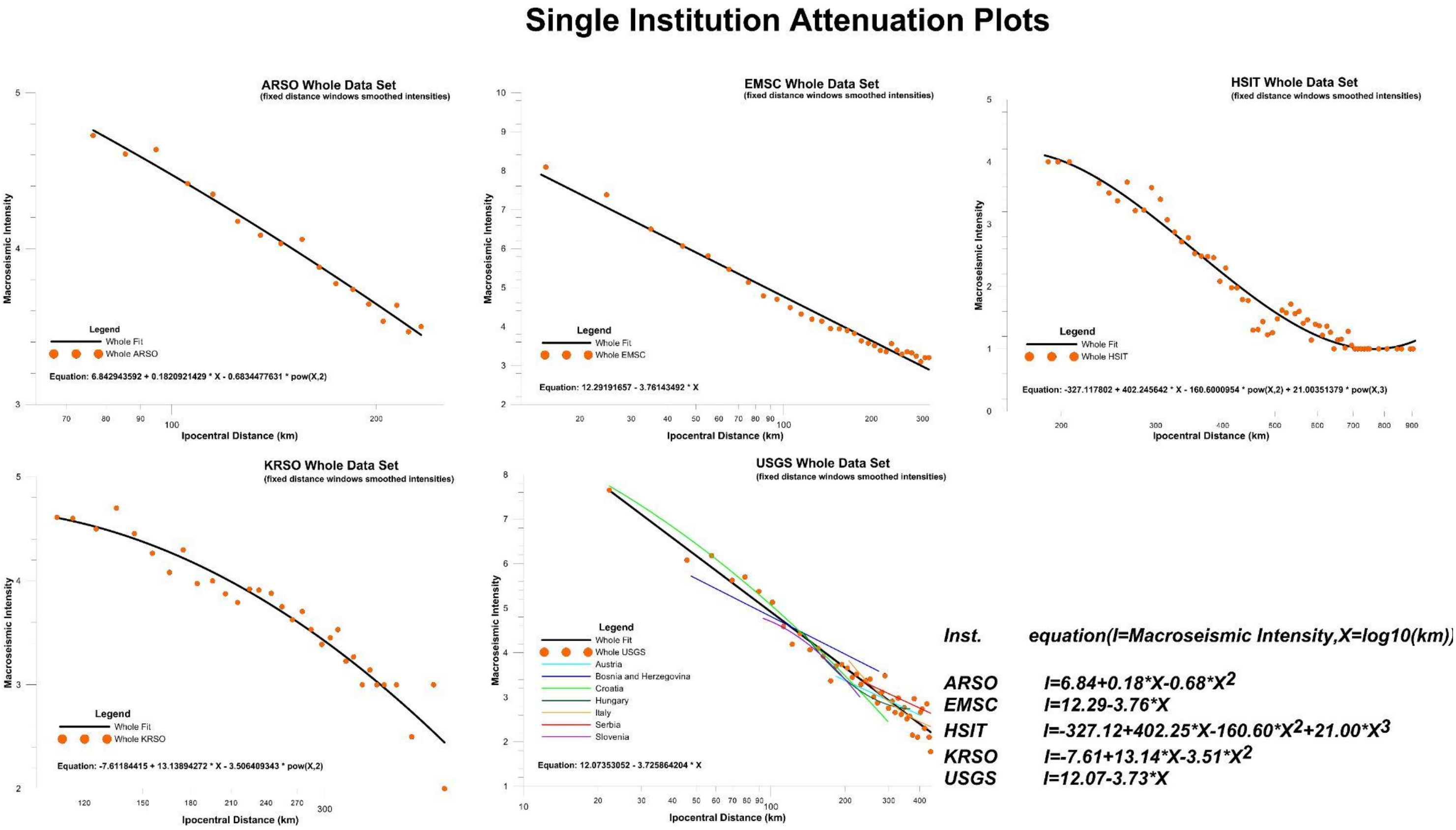


Fig. 11 All Institution's specific intensities versus hypocentral distance fits and relative fits polynomial equations.

We introduced attenuation analysis applying window smoothing for each Institution separately to remove high variation within data and, most importantly, model institution-specific attenuation law.

The attenuation fit equations are different, suggesting that a method to homogenise all of them could use this information.

Using our primary method to homogenise data, we selected USGS as the target institution because it was present in every country for the earthquake chosen, and it has a well-established method to assess intensity. It is possible to translate other, different local-country or widespread institution intensities using the equations previously stated. As an example, we aimed to transform and homogenise Italian HSIT data to USGS. We defined the attenuation fits of each Institution and calculated the differences between the two equations. We were then able to determine HSIT transformed data by adding the previously defined difference to original HSIT data.

Method to Homogenise Data

- Choose a Target Institution: USGS, and use its attenuation equation as a reference.
- For the other Institution, define D: Difference as D = Target Institution Att. Equation -

Source Institution Att.Equation

- Transformed Source Institution = Source Institution Data + D

Example: HSIT data transformed to USGS target

USGS Att. Eq.: ***Int.= 12.07-3.73*x***
x=Log₁₀ ipocentral distance
HSIT Att. Eq.: ***Int.= -327.12+402.25*x-160.60*x²+21.00*x³***

$$D=12.07-3.73*x - (-327.12+402.25*x-160.60*x^2+21.00*x^3)$$

HSIT transformed Intensity data $Tr.Int.(x) = Int.(x) + 12.07-3.73*x - (-327.12+402.25*x-160.60*x^2+21.00*x^3)$

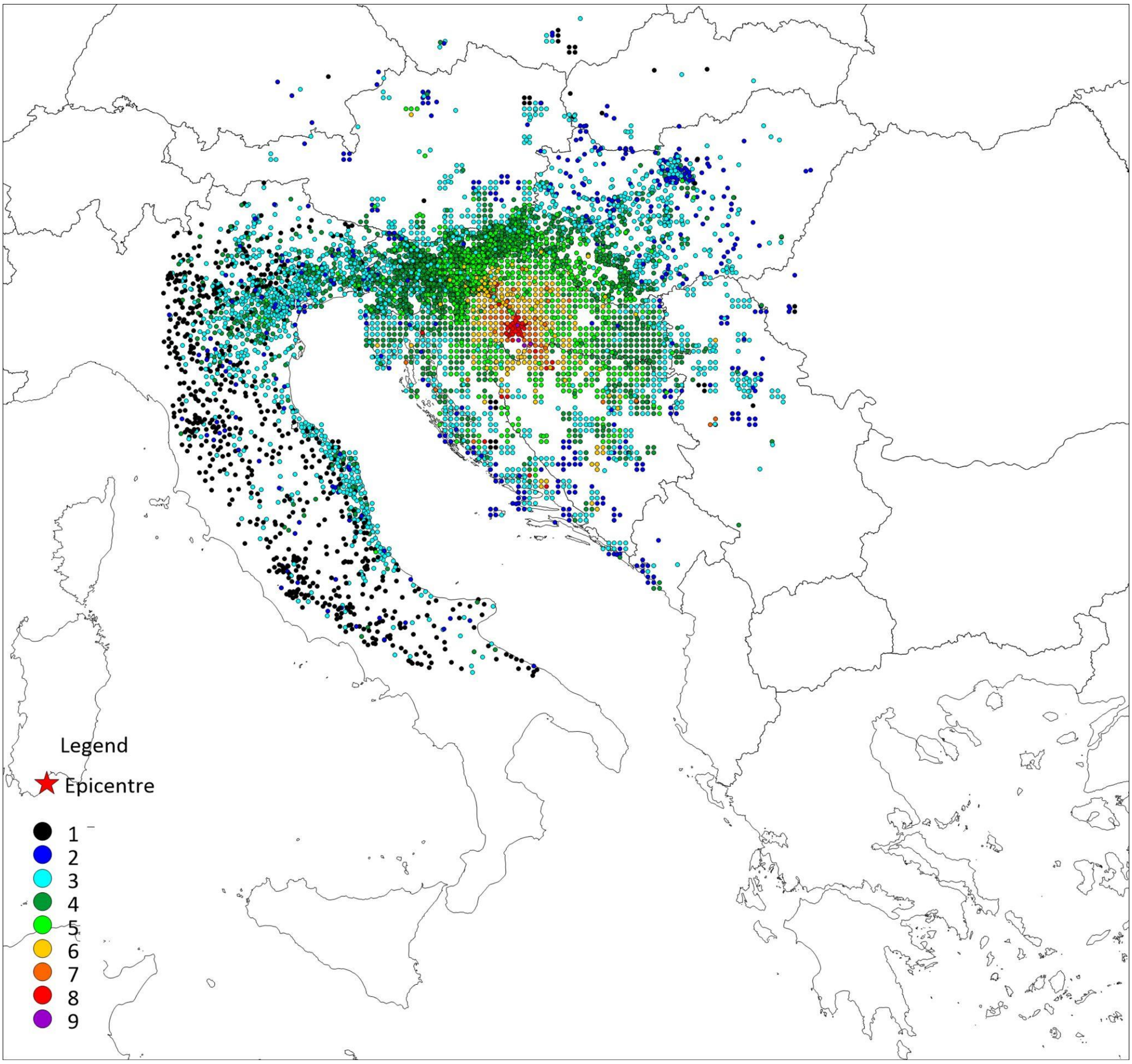


Fig. 12 USGS transformed intensities from all data collected.

This plot shows USGS transformed intensities from all data we collected (Fig. 12). We aimed to interpolate the intensities recognising and separating source, regional and local components and filtering out random components.

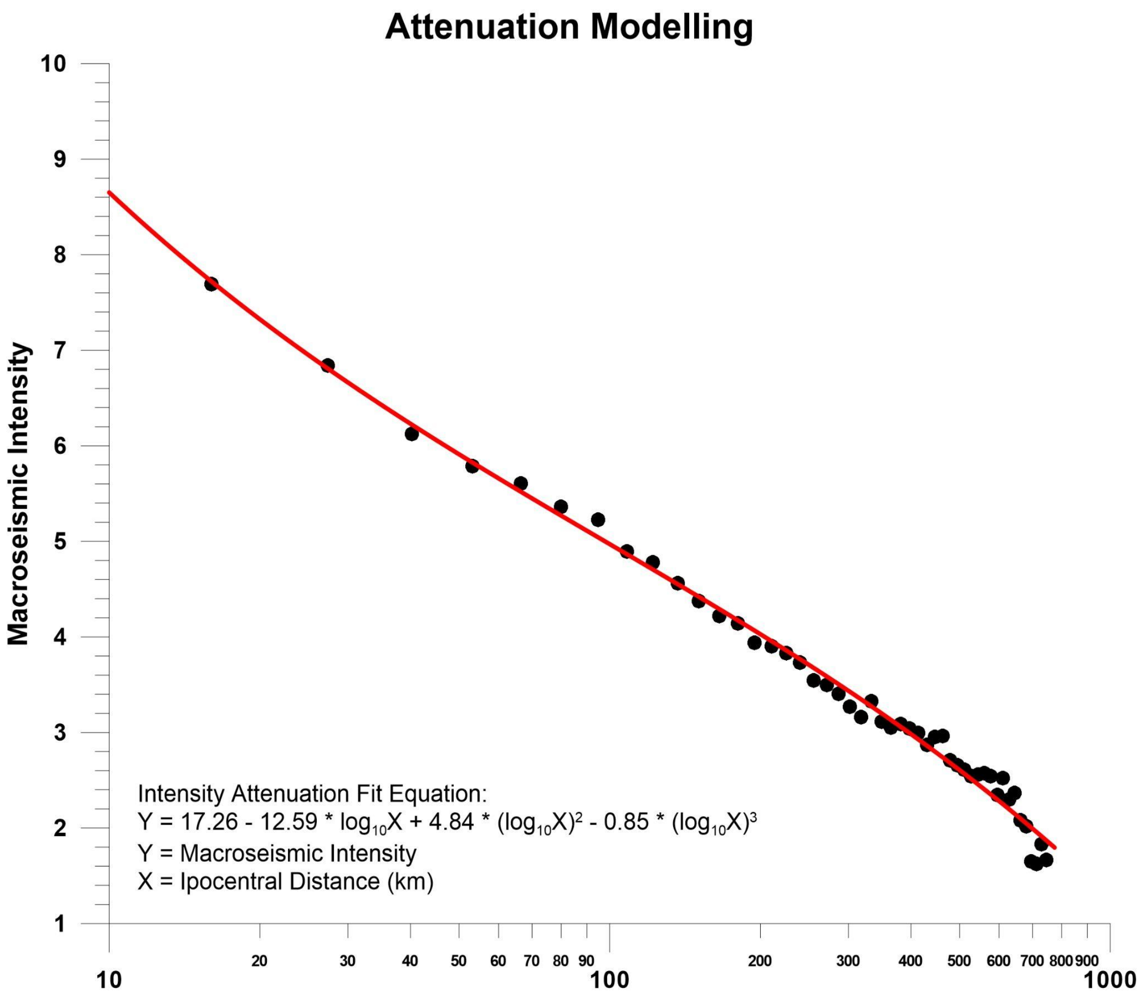


Fig. 13 Moving 10 km window smoothed intensity attenuation as a function of hypocentral distance (black dots) and polynomial attenuation fit (third-degree equation) plotted as the red line.

One main component is attenuation modelling, where we retrieved the equation from the polynomial fit, remembering that we provided a window filtering made-average of 10Km window distance width. This fit is a third-degree polynomial, and it is very close to point values (Fig. 13).

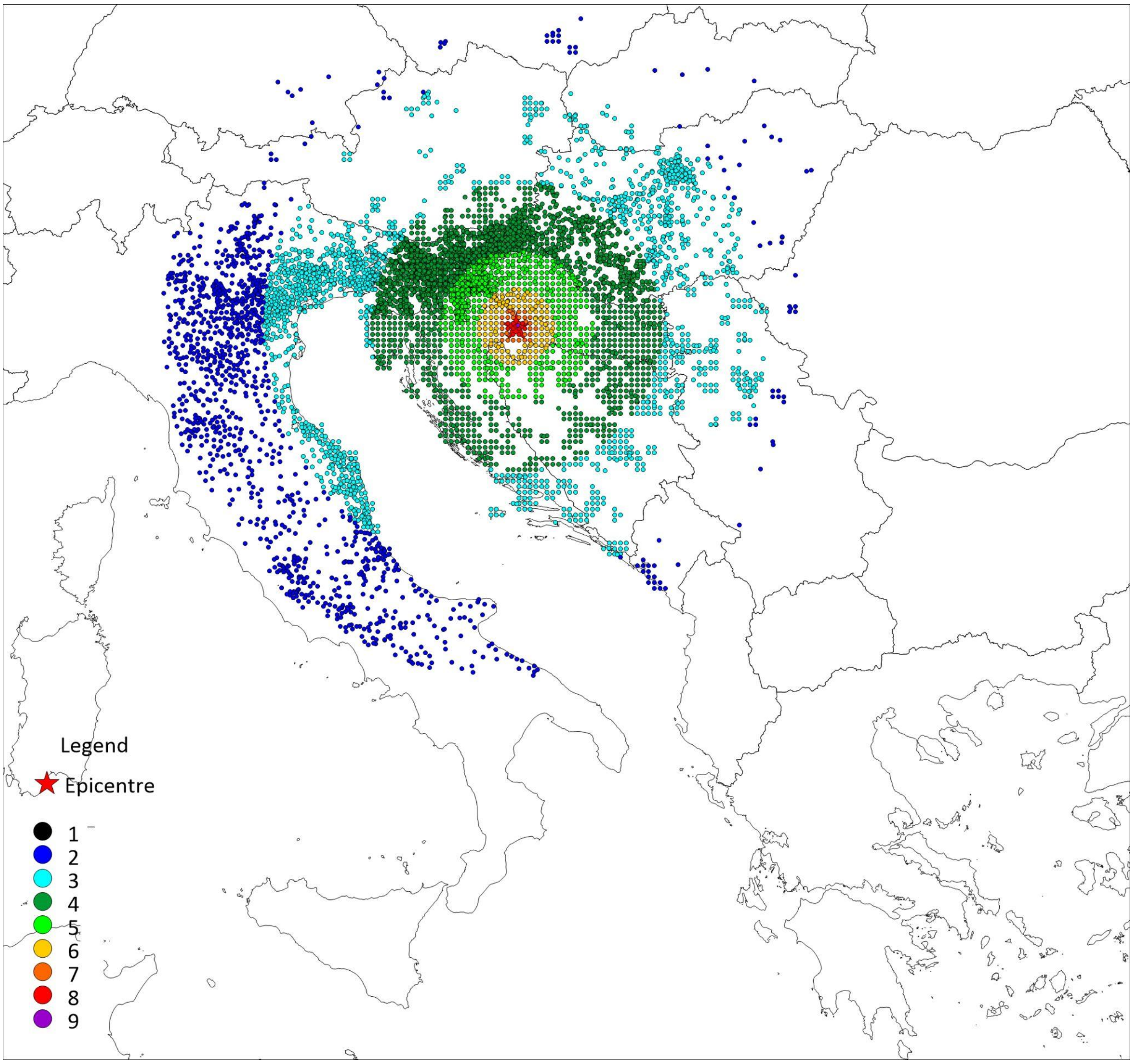


Fig. 14 Intensity spatial distribution from third-degree polynomial fitted equation.

Using the third-degree polynomial fit function previously calculated, we could then translate this data component into space, where we have the figure above represented (Fig. 14).

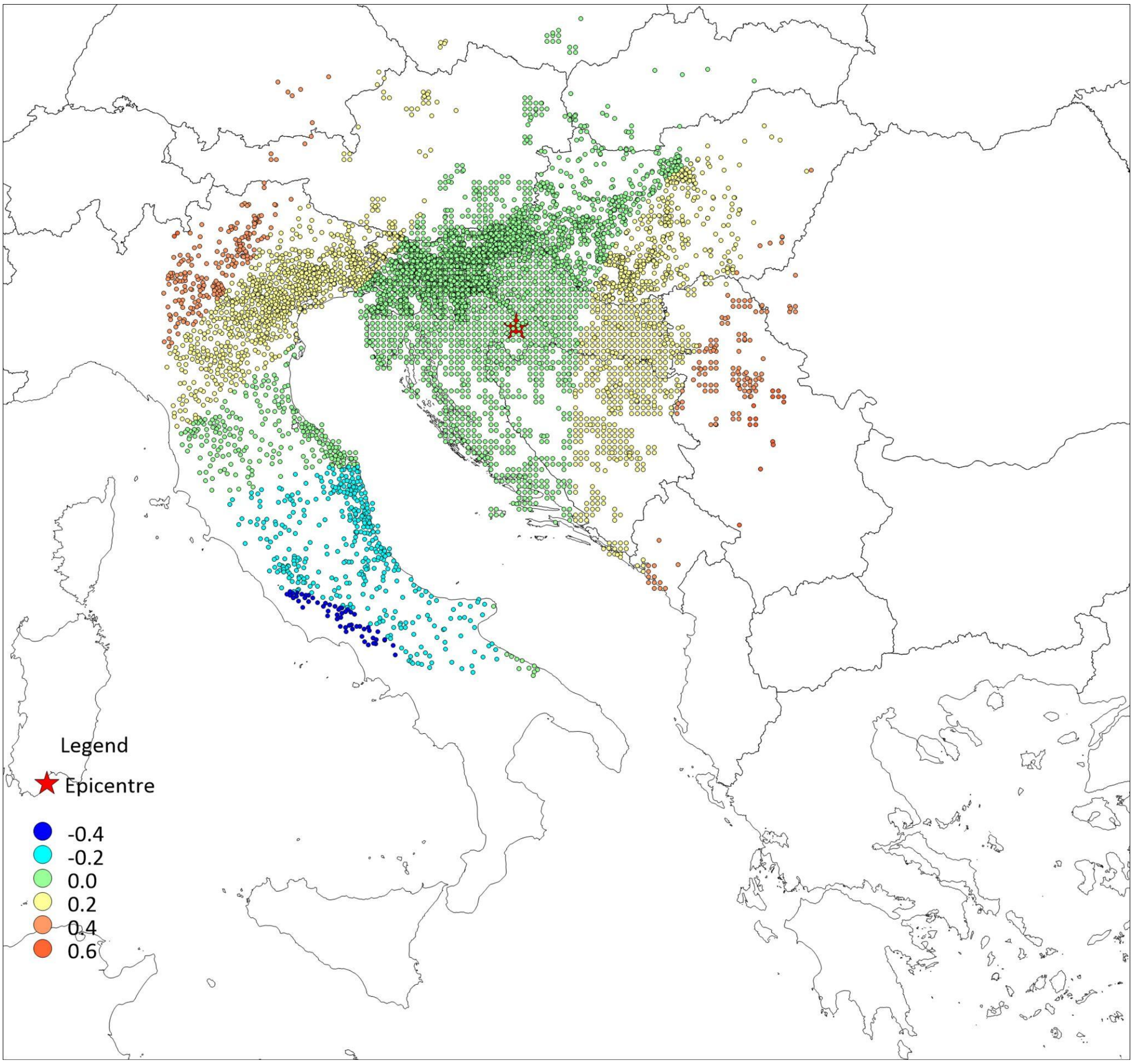


Fig. 15 Intensity spatial representation of a polynomial function (2nd order polynomial surface), representing a spatial trend not evidenced by the main attenuation fit.

A check on residuals calculated from the removal of USGS transformed data, with an attenuation third polynomial fit previously obtained, showed that data were affected by a regional trend constituted by a 2nd order polynomial function of latitude and longitude, transformed into y and x UTM coordinates (Fig. 15). Therefore, we removed this second trend. The need to remove this regional trend is due to a correct application of kriging and related to semivariogram analysis, involving proper interpolation and noise removal.

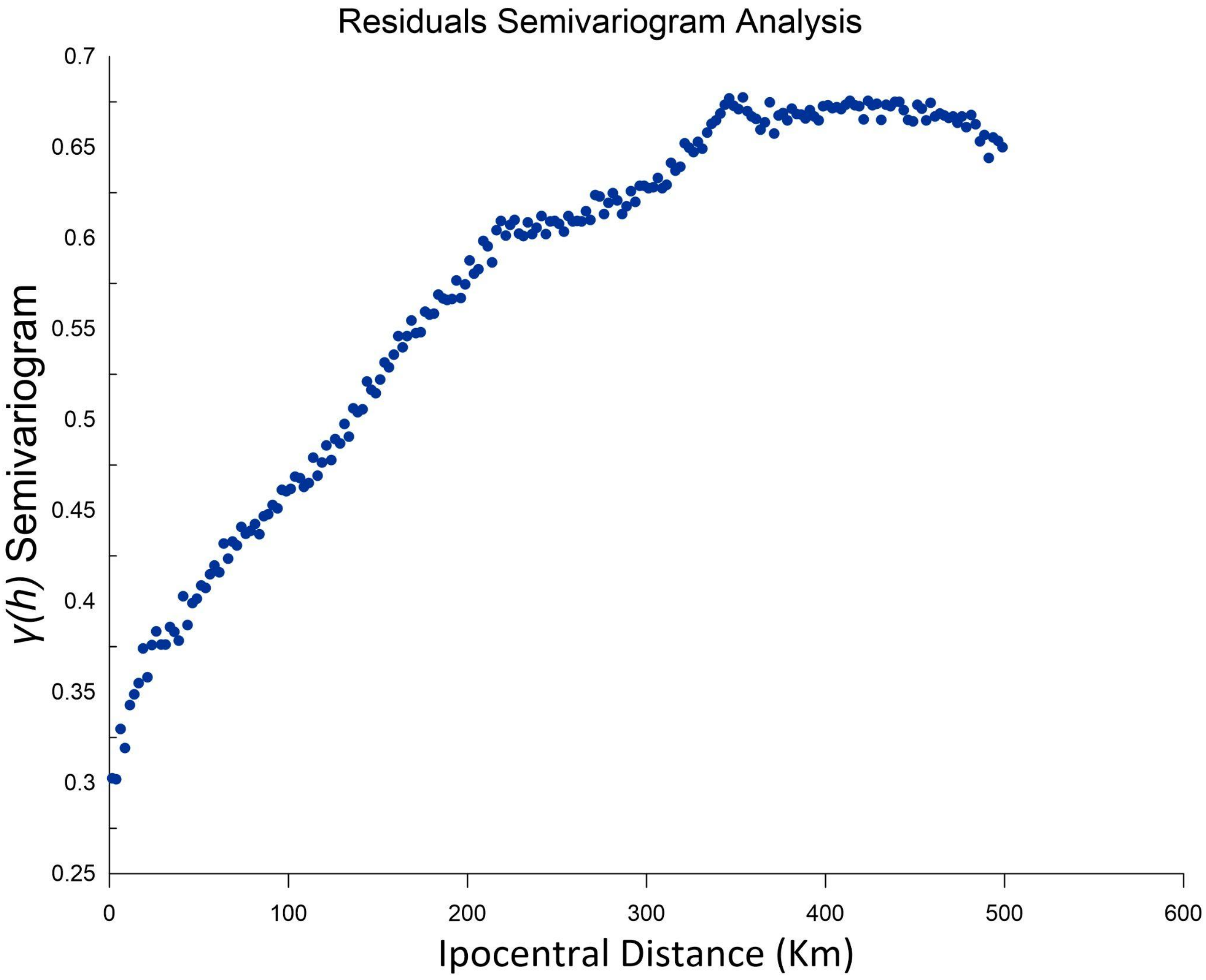


Fig. 16 Residuals semivariogram as a function of inter distances (km) among EMSC intensities after two steps detrending.

We then studied the final residuals through the semivariogram analysis. This plot shows the semivariance characterised by a clear spatial autocorrelation relationship, with a very reduced, if not absent, nugget effect (Fig. 16).

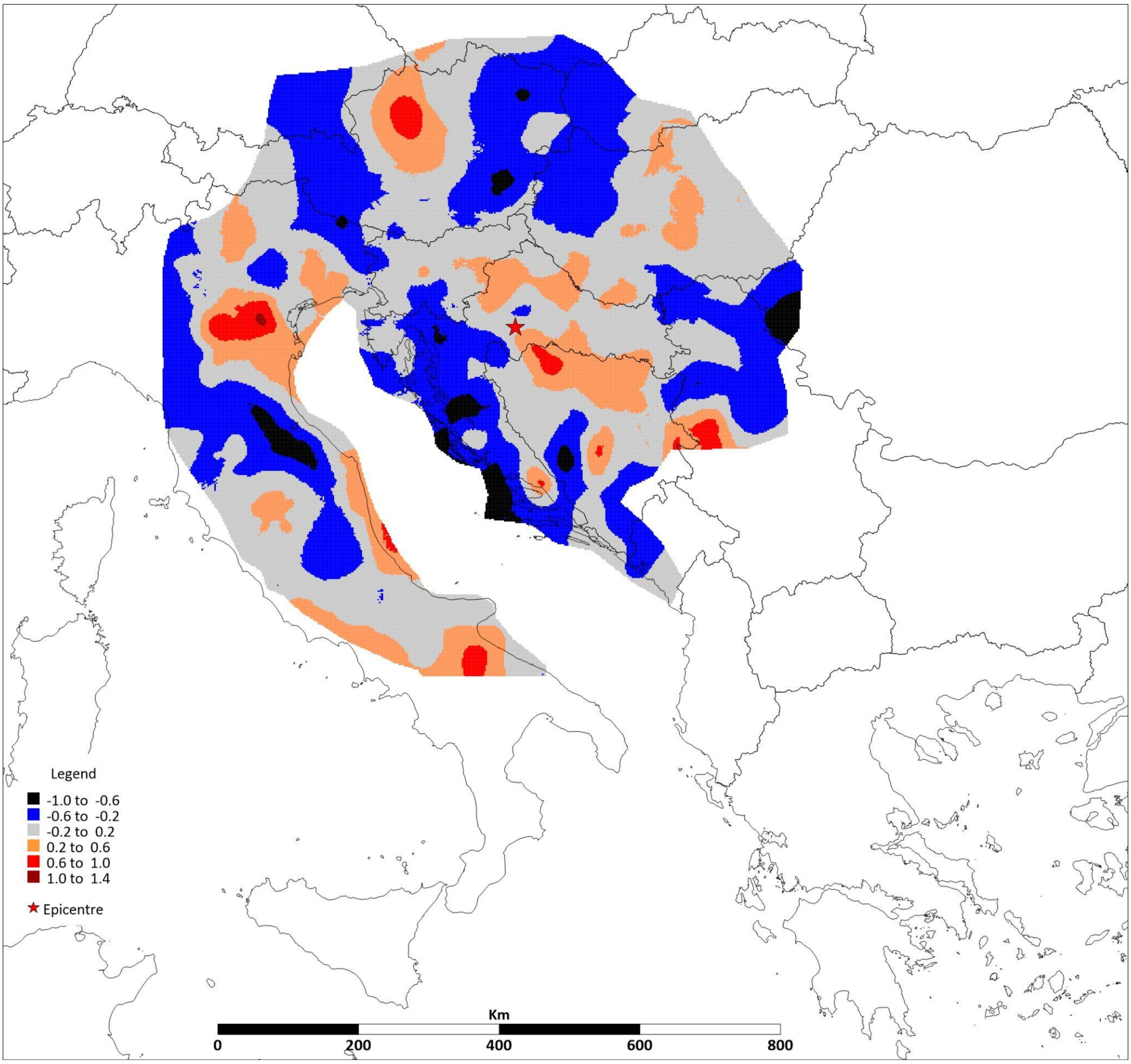
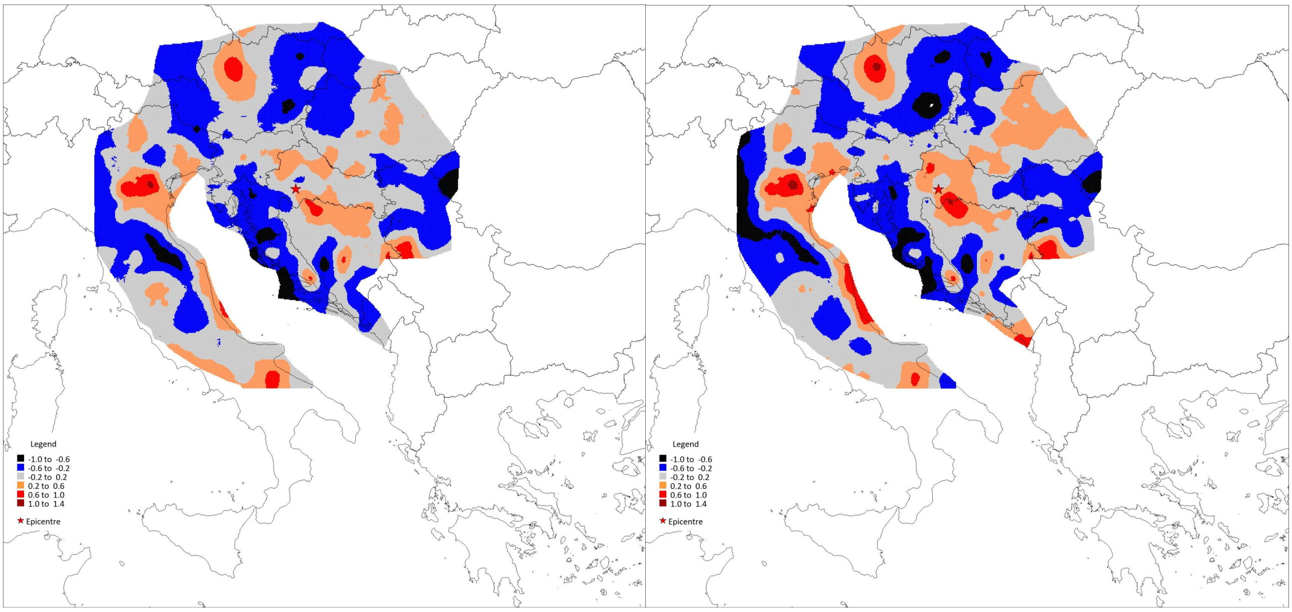


Fig. 17 Kriging interpolation of USGS transformed intensity residuals: orange-red areas indicate amplification, blue-black areas represent attenuation regarding trend (attenuations) intensity behaviour.

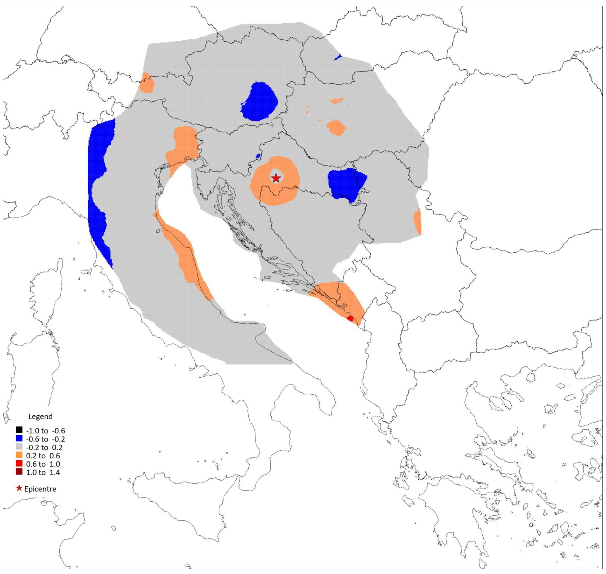
After the semivariogram analysis, we applied kriging interpolation to obtain residual intensities, separated by random noise. Areas with a grey colour indicate no anomalies; therefore, the field has the same composition as the trend components presented previously. A blue/black colour individuates attenuation areas, whereas orange/red areas individuate amplification patterns (Fig. 17).



USGS Transformed data   Original data

Fig. 18 Comparison between USGS transformed intensity data (residuals) and raw intensities (residuals).

The figure on the left denotes the USGS transformed data. However, the figure on the right includes the same analysis and processes calculated before, but with original data rather than USGS transformed. As we can see, the original data plot includes certain areas that are more enunciated (Fig. 18).



The difference between the two previous plots relates to the figure on the left. Specific regions such as the epicentre and the East coast of Italy are highlighted and can be located easily. These areas put in evidence the risks in using unhomogenised intensities, with almost one intensity degree difference (Fig 19).

Fig. 19 Residuals kriging interpolation difference between raw and USGS transformed data (Res=Raw-USGS trans.).

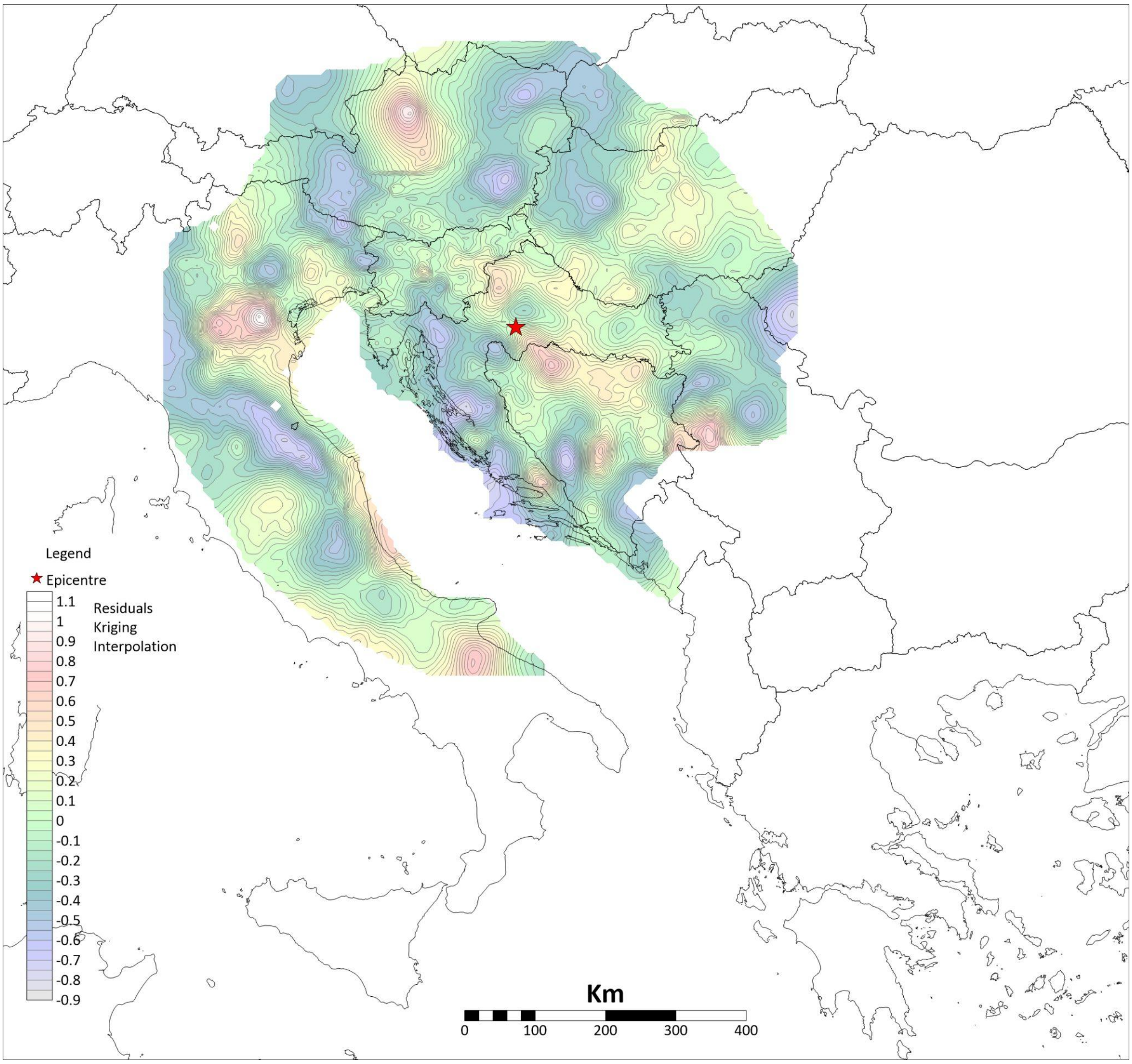


Fig. 20 Kriging interpolation of USGS transformed intensity residuals: orange-red areas indicate amplification, blue-black areas attenuation regarding trend (attenuations) intensity behaviour. As Fig 17 but with more detailed contouring.

This figure above equals the USGS homogenised plot shown in figure 17, containing the same residuals. However, we want to highlight residuals in a more detailed format, so we have adjusted the contouring accordingly.

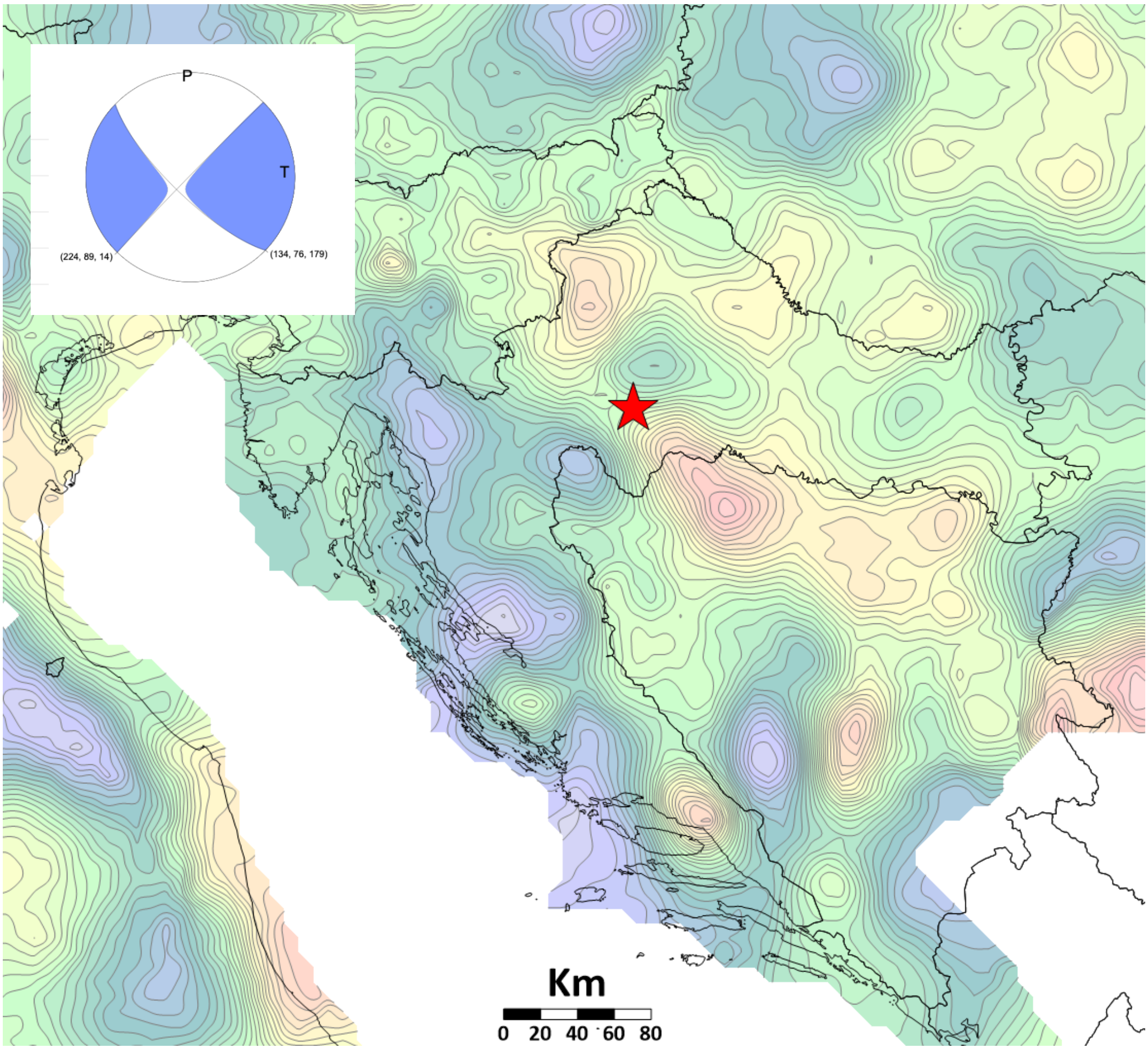


Fig. 21 As figure 20 but a more detailed scale highlights the epicentral area and the focal mechanism.

This figure (Fig. 21) indicates the same plot but displays an enlargement centred on the radiation source pattern epicentral area. As we can see, in a North-East to South-West divide, there is attenuation in these areas compared to amplification, which we can see in a North-West to South-East divide. A focal mechanism can be displayed too.

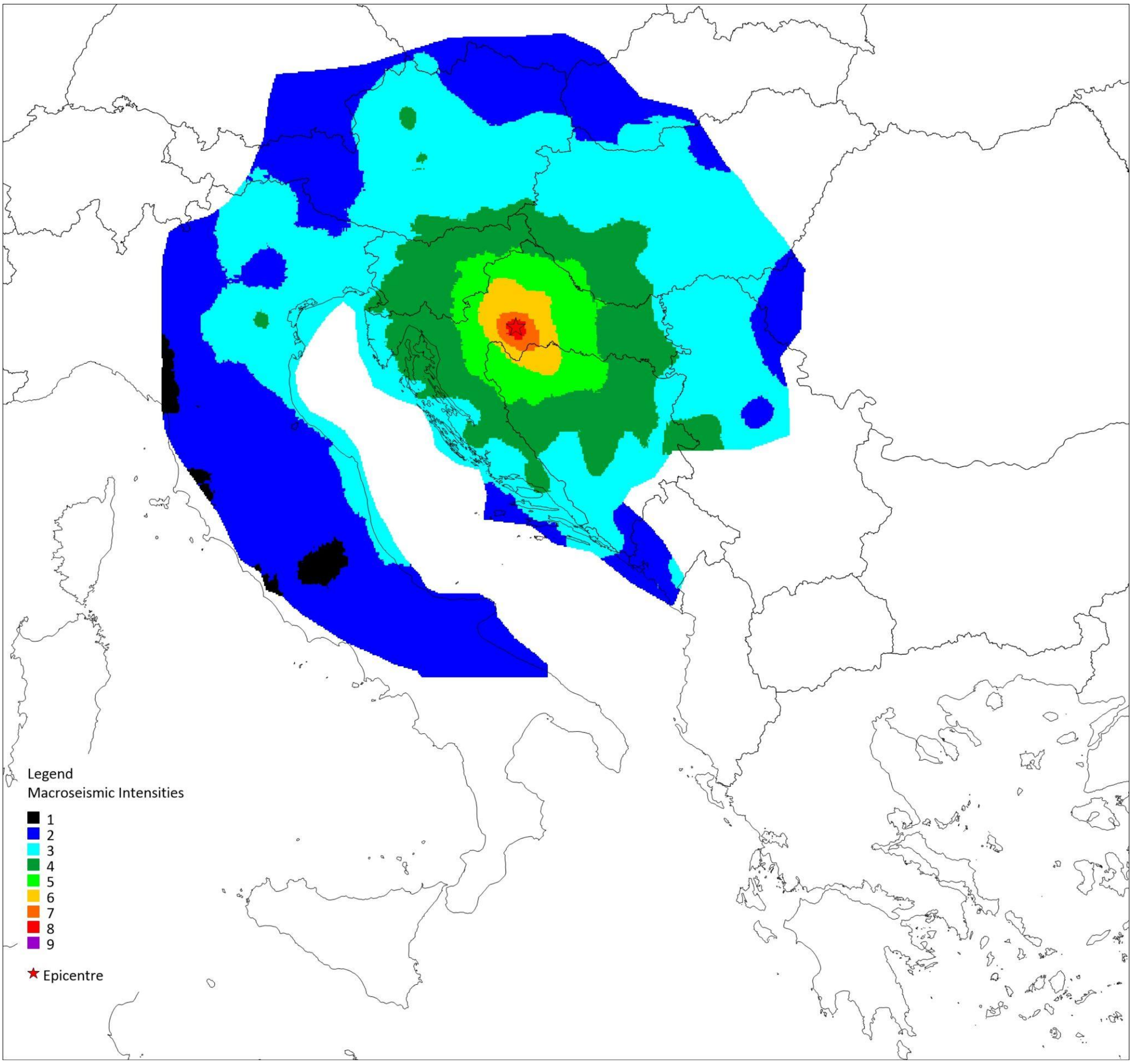


Fig. 22 Complete kriging interpolated macroseismic intensity field of USGS transformed data.

A final, complete USGS transformed data, kriging interpolated, macroseismic field plot is displayed. It is composed by adding all two trend components to kriging residuals, discarding the sole random components (Fig 22).

This analysis shows how, in the presence of a source of intensity data covering the entire area affected by the earthquake, it is possible to make homogeneous data relating to specific areas (generally limited to the country). This process makes it possible to have a data set made up of all the available data. We achieved an improvement in the statistical robustness of the data. We based this homogenisation criterion on the analysis of the isotropic attenuations of the individual data sources. In this way, it is possible to identify regional trends, anomalies (local attenuation or amplification) and the separation of the random component of the data. Last but not least, using kriging, we also have the confidence limits evaluation for each interpolated intensity in the territory. These limits define the estimate's uncertainty, but it depends on the quantity and spatial variability of the original intensity data.