The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific
and Cultural Organization

International Atomic
Energy Agency

World Year of PHYSICS 2005

H4.SMR/1645-9

"2nd Workshop on Earthquake Engineering for Nuclear Facilities: Uncertainties in Seismic Hazard"

14 - 25 February 2005

Analysis of Earthquake Catalogs

A. Peresan

Department of Earth Sciences
University of Trieste

2nd Workshop on Earthquake Engineering for Nuclear Facilities
Uncertainties in Seismic Hazard Assessment
Trieste - Italy, 14 - 25 February 2005

Seismological database:
Historical seismicity & earthquake catalogues

Analysis of Earthquake Catalogs

A. Peresan & G.F. Panza

DST
Department of Earth Sciences
University of Trieste

the abdus salam international centre for theoretical physics
united nations educational, scientific and cultural organization
international atomic energy agency

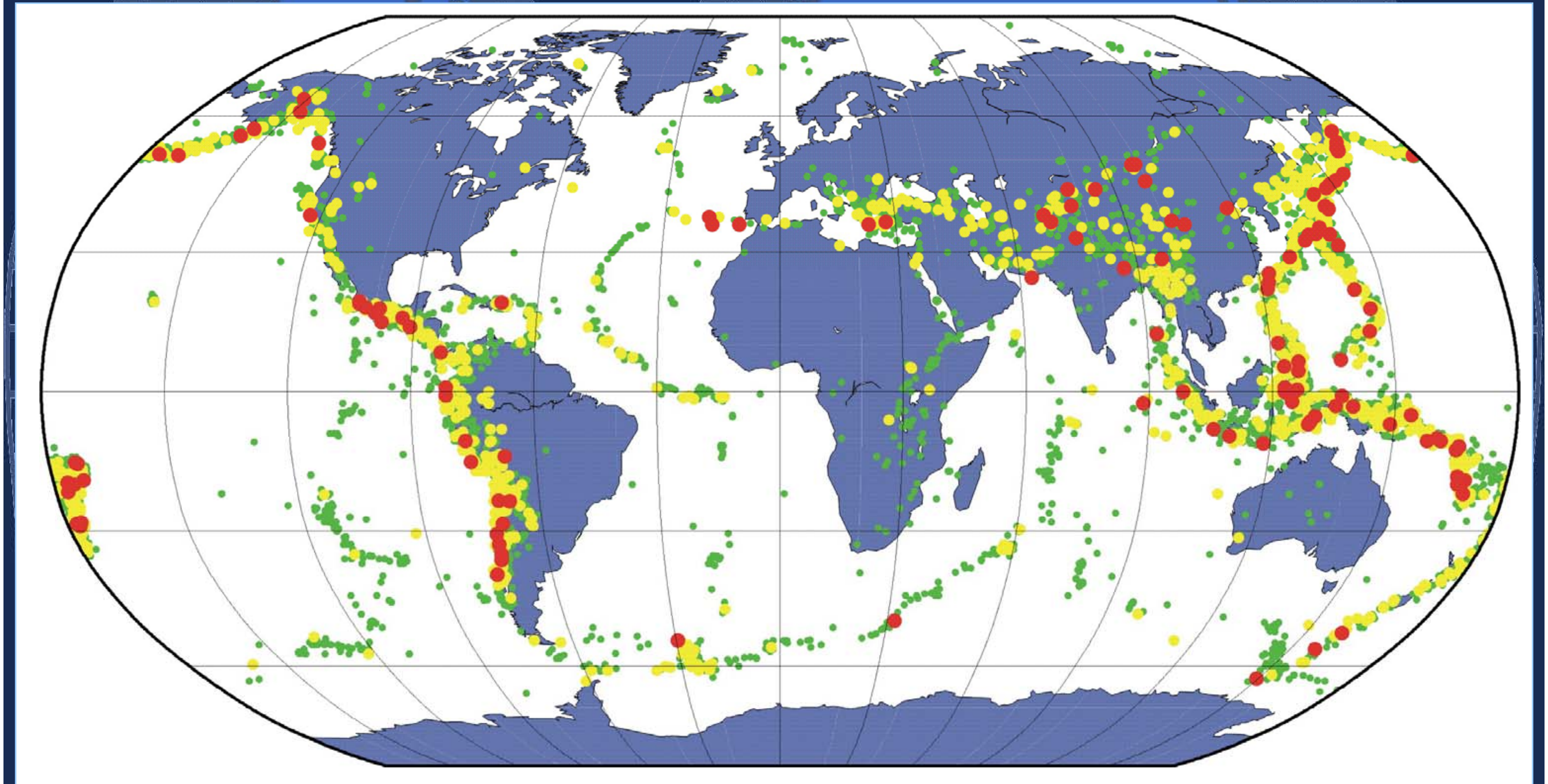
Outline

- Earthquakes and earthquakes catalogs
- Measuring the size of an earthquake
- Heterogeneity of earthquake catalogs
- Uncertainties and catalog errors
- Examples of catalogs analysis
- Detection of long lasting changes in reported magnitudes
- Effects of random errors in magnitude on intermediate-term middle-range earthquake predictions
- The Gutenberg-Richter law

What are earthquakes?

- Earthquakes are generally due to sudden fractures in the outermost fragile part of the Earth's lithosphere, the crust, that cause permanent deformations and radiate seismic waves.
- Seismic waves propagate from the source through the Earth in any direction, eventually causing detectable ground shaking at the surface.
- Although historical records on earthquakes are known from 2100 B.C., most of them before the middle of the 18th century are generally lacking description or are not reliable.

Most of the earthquakes happen under water which makes their detection and measurement even more difficult.
Some earthquakes produce tsunamis.



Earthquakes with magnitude M>6 (USGS-NEIC data, 1900-2004)

Measuring the size of an earthquake

- The information about the size of historical earthquakes is generally provided in terms of earthquake intensity, i.e. a quantitative estimation based on the observed damage.
- It was only in the 1930's that *Charles F. Richter*, a California seismologist, introduced the concept of earthquake magnitude.

Intensity scales

- The Mercalli scale was put forward by Mercalli in 1902. An elaboration of the Mercalli scale, was published by Sieberg in 1923.
- This form was in turn used as the basis for the Modified Mercalli (MM) Scale of 1931 by Wood and Neumann.
- Subsequently other intensity scales have been introduced by Mercalli, Cancani and Sieberg (MCS) and by Medvedev, Sponeuer and Karnik (MSK). More recently the EMS-1992 macroseismic scale has been proposed.

Intensity scales

- The existence of many different scales is a demonstration of the complexity of the problem of describing earthquake effects. The multiplicity of scales generates some problems in practical applications, that must therefore rely upon very conservative assumptions.

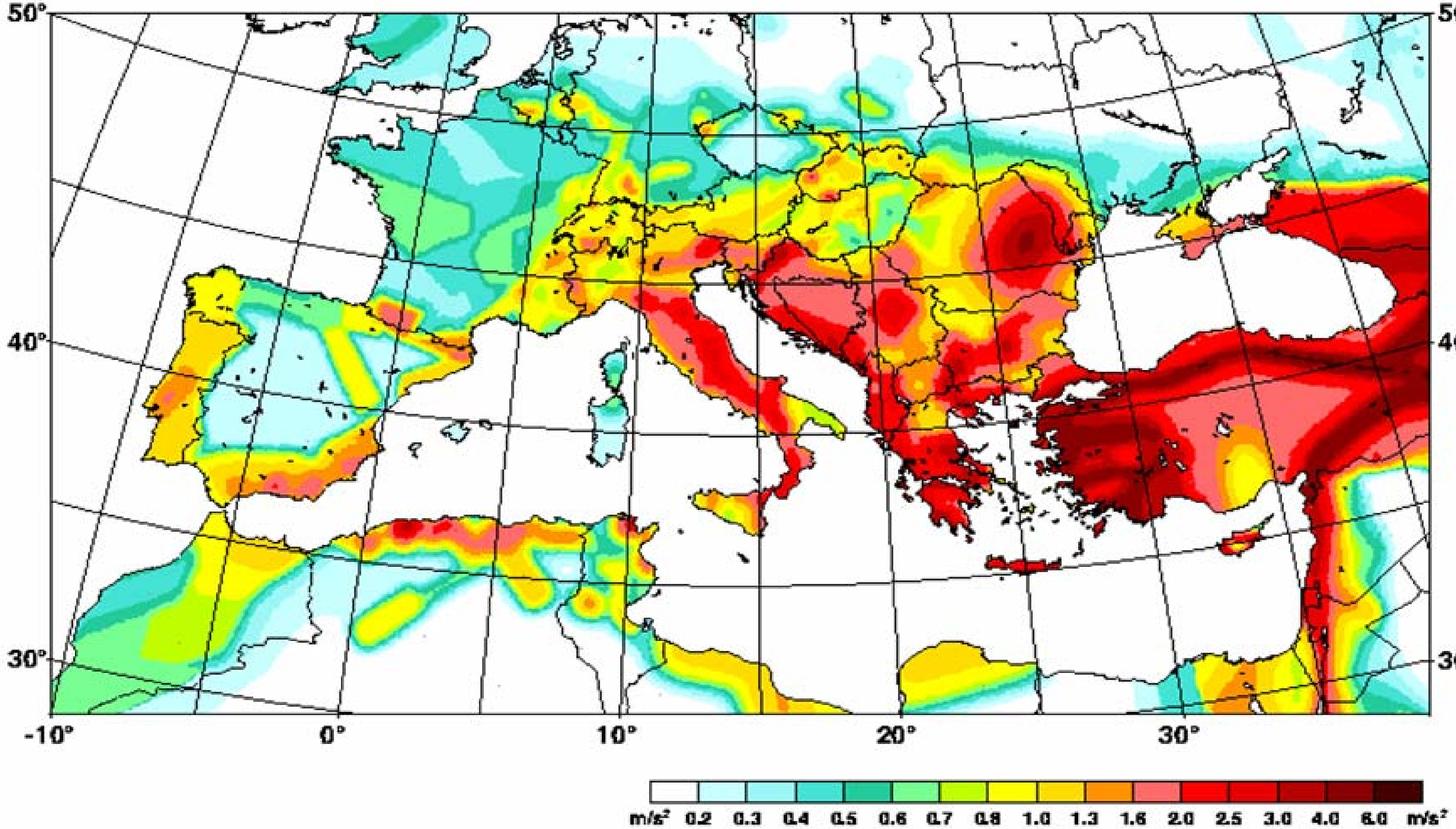
<table>
<tr><th>MM</th><th>RF</th><th>JMA</th><th>MCS</th><th>MSK</th></tr>
<tr><td rowspan="2">I</td><td rowspan="3">I</td><td></td><td rowspan="2">II</td><td rowspan="2">I</td></tr>
<tr><td rowspan="4">I</td></tr>
<tr><td rowspan="2">II</td><td rowspan="2">III</td><td rowspan="2">II</td></tr>
<tr><td>II</td></tr>
<tr><td>III</td><td>III</td><td>IV</td><td>III</td></tr>
<tr><td rowspan="2">IV</td><td>IV</td><td rowspan="2">II</td><td rowspan="2">V</td><td rowspan="2">IV</td></tr>
<tr><td rowspan="2">V</td></tr>
<tr><td rowspan="2">V</td><td rowspan="2">III</td><td rowspan="2">VI</td><td rowspan="2">V</td></tr>
<tr><td rowspan="2">VI</td></tr>
<tr><td rowspan="2">VI</td><td rowspan="2">IV</td><td rowspan="2">VII</td><td rowspan="2">VI</td></tr>
<tr><td>VII</td></tr>
<tr><td rowspan="2">VII</td><td rowspan="3">VIII</td><td rowspan="5">V</td><td>VIII</td><td rowspan="2">VII</td></tr>
<tr><td rowspan="3">IX</td></tr>
<tr><td rowspan="3">VIII</td><td rowspan="3">VIII</td></tr>
<tr><td rowspan="4">IX</td></tr>
<tr><td rowspan="2">X</td></tr>
<tr><td rowspan="2">IX</td><td rowspan="4">VI</td><td rowspan="2">IX</td></tr>
<tr><td rowspan="2">XI</td></tr>
<tr><td rowspan="2">X</td><td rowspan="4">X</td><td rowspan="2">X</td></tr>
<tr><td>XII</td></tr>
<tr><td>XI</td><td rowspan="2">VII</td><td rowspan="2"></td><td>XI</td></tr>
<tr><td>XII</td><td>XII</td></tr>
</table>

Comparison of seismic intensity scales:
MM – Modified Mercalli
RF – Rossi-Forel
JMA – Japanese Meteorological Agency
MCS – Mercalli-Cancani-Sieberg
MSK – Medvedev-Sponheuer-Karnik

Intensity scales

- Intensity provides a qualitative description of the earthquake size, based on the observation of the related damage. Hence, for a given earthquake, the intensity I can be different in different places.
- Intensity values are discrete; undue accuracy in related computations can be misleading.

The GSHAP probabilistic map at 475 years



Horizontal Peak Ground Acceleration seismic hazard map representing stiff site conditions for an exceedance or occurrence rate of 10% within 50 years for the Mediterranean region.

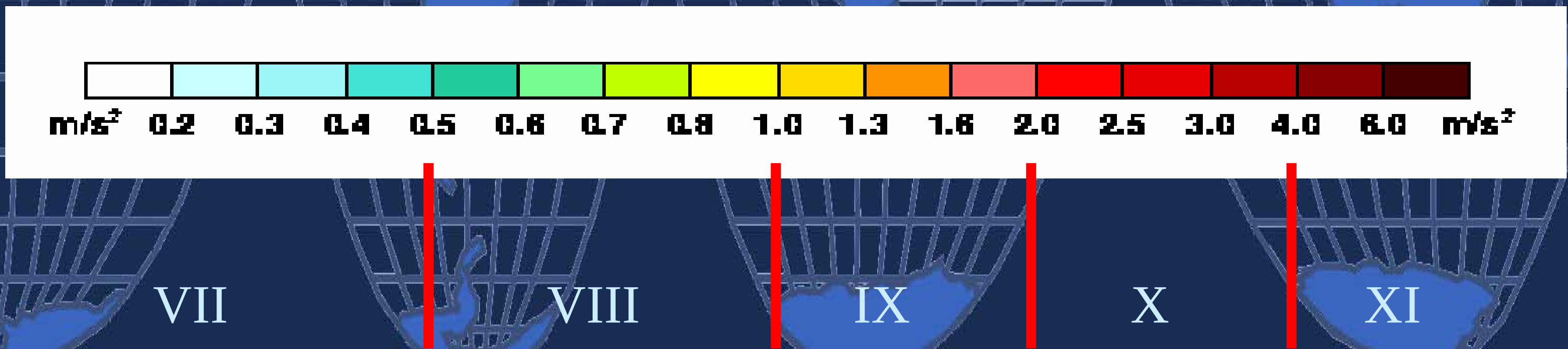
The log-linear regression between maximum observed macroseismic intensity, I (MCS), and peak values of ground motion has a slope close to 0.3 (see Panza et al., 1999; Shteinberg et al., 1993 and references therein). Hence one degree of intensity corresponds to a factor two in the values of ground motion:

DGA(I-1)/DGA(I)=2

PGV(I-1)/PGV(I)=2

PGD(I-1)/PGD(I)=2

Comparison between GSHAP scale used in the Mediterrnean, and MCS Intensity scale



Magnitude scales

- Richter's original magnitude scale (M_L) held only for California earthquakes occurring within 600 km of a particular type of seismograph (i.e., the *Woods-Anderson* torsion instrument). Then it was extended to observations of earthquakes of any distance and of focal depths ranging between 0 and 700 km.
- Because earthquakes excite both body waves, which travel into and through the Earth, and surface waves, which are constrained to follow the Earth's uppermost layers, two other magnitude scales evolved - the m_b and M_S

What is an earthquake catalog?

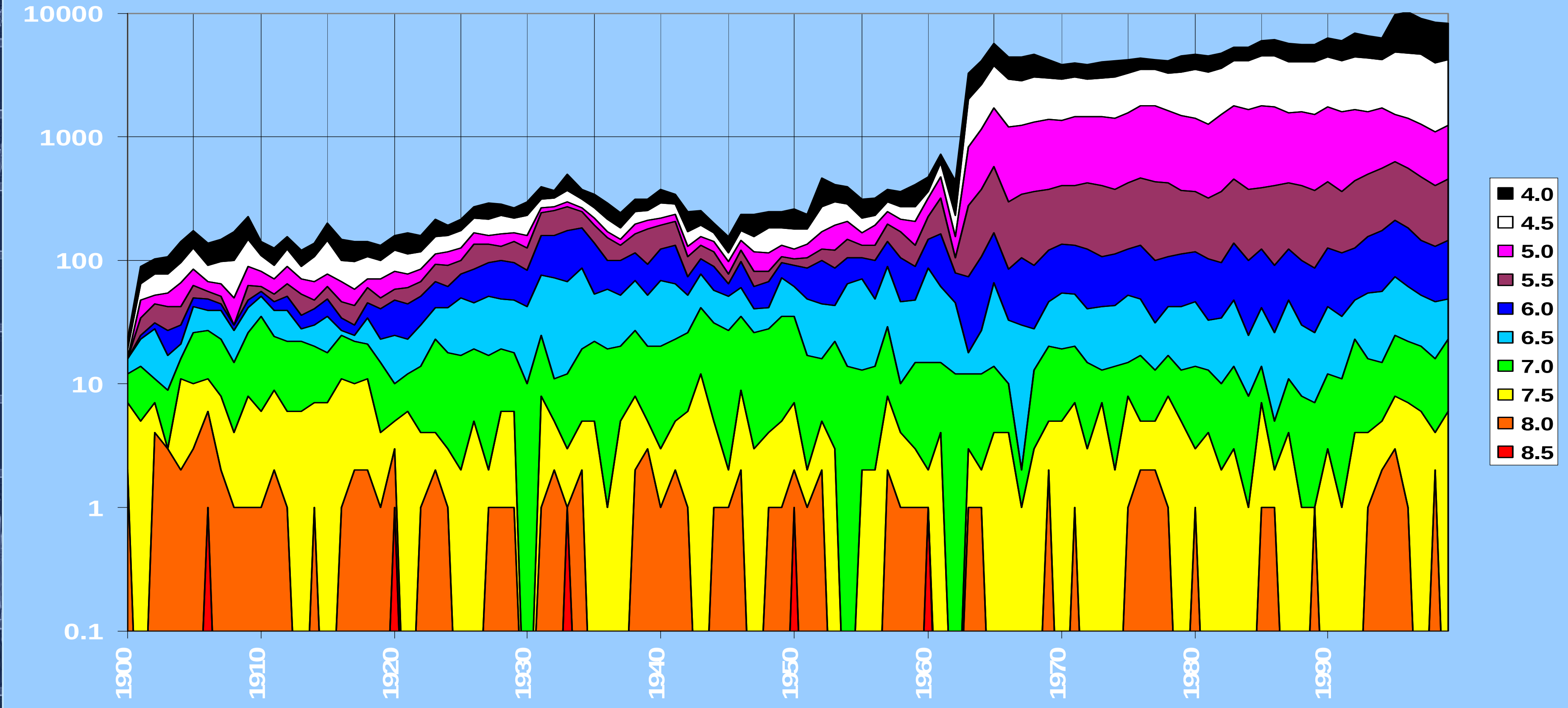
- An earthquake catalog is a collection of information about a set of seismic events, basically including:
  - Origin time
  - Location
  - Size of the earthquakes
- Additional information can be provided, ranging from related damage to seismic source parameters.
- A catalog may include several magnitude estimations, generally with a precision of one digit, even if values provided by different agencies may differ more than one unit.

What is an earthquake catalog?

- Catalogs are compiled for different purposes and by different agencies. Therefore they differ in:
  - geographical coverage
  - time span
  - level of detection
  - criteria of compilation
  - type and quality of earthquake data
- Consequence: no unique catalog for a given territory… but usually an heterogeneous set of catalogs (historical, instrumental, local, global, etc.), not always comparable, which may require different tools of analysis.
- A positive step forward: compilation of global catalogs (e.g. USGS-NEIC and ISC)

The USGS/NEIC Global Hypocenter Data Base

Global Number of Earthquakes vs. Time

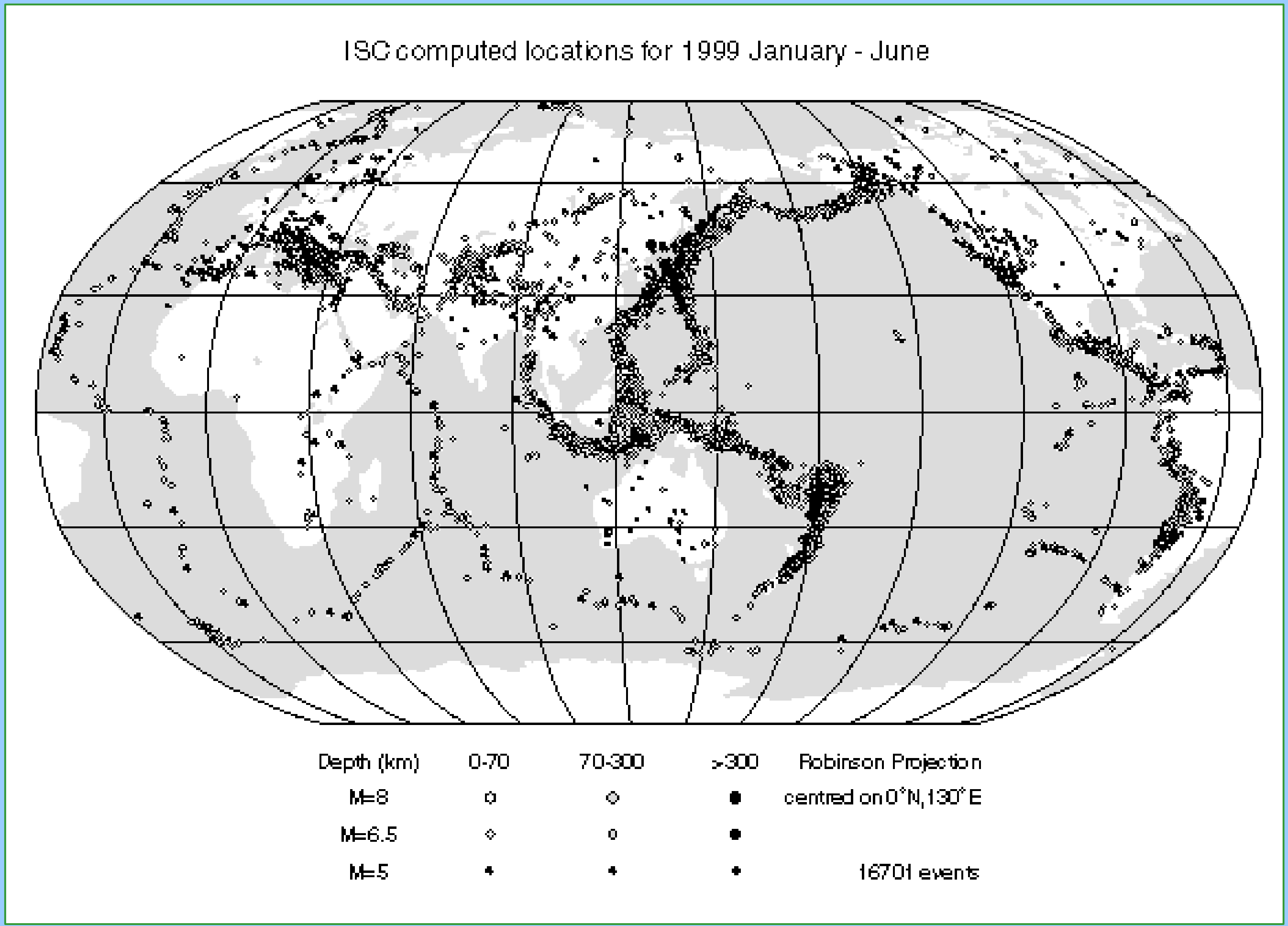


The ISC Global Data Base

Bulletin of the International Seismological Centre

Regional Catalogue of earthquakes

Felt and Damaging earthquakes



What can we learn from a catalog of earthquakes?

There are two extreme opinions on the subject:

- *Pessimistic*: "… in the case of seismic data, most of the observed variations are, in fact, related to changes in the system for detecting and reporting earthquakes and not to actual changes in the Earth."
- *Optimistic*: Among existing data seismic catalogs remain the most reliable record on distribution of earthquakes in space and time.

- All catalogs have errors, which may render invalid conclusions derived in a study based on a catalog of earthquakes.

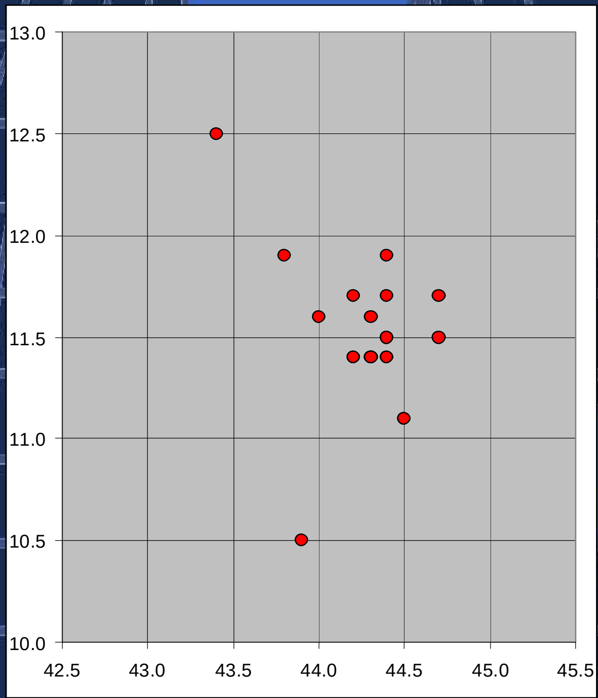
  Two ways to avoid the errors:

  - Postpone the analysis until the data are refined;

  - Use robust methods within the limits of their applicability.

  When different catalogs are available, a comparative analysis may allow to detect errors.

Uncertainties and errors

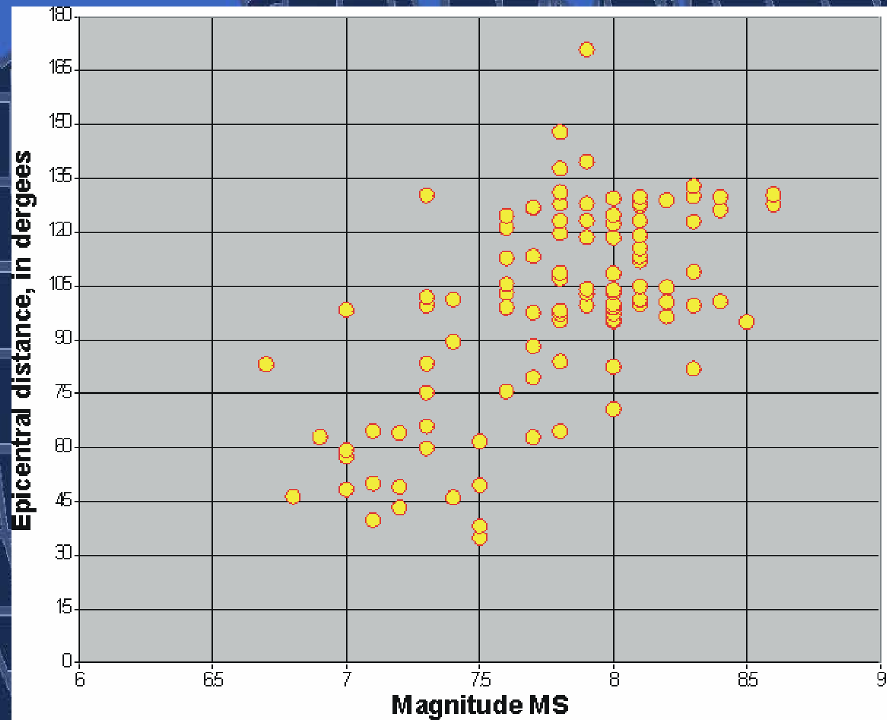


Uncertainty in preliminary epicentral determinations

Fast determinations of the epicenter for the 14 September 2003 earthquake in Northern Italy by different seismological agencies to European-Mediterranean Seismological Centre (EMSC)

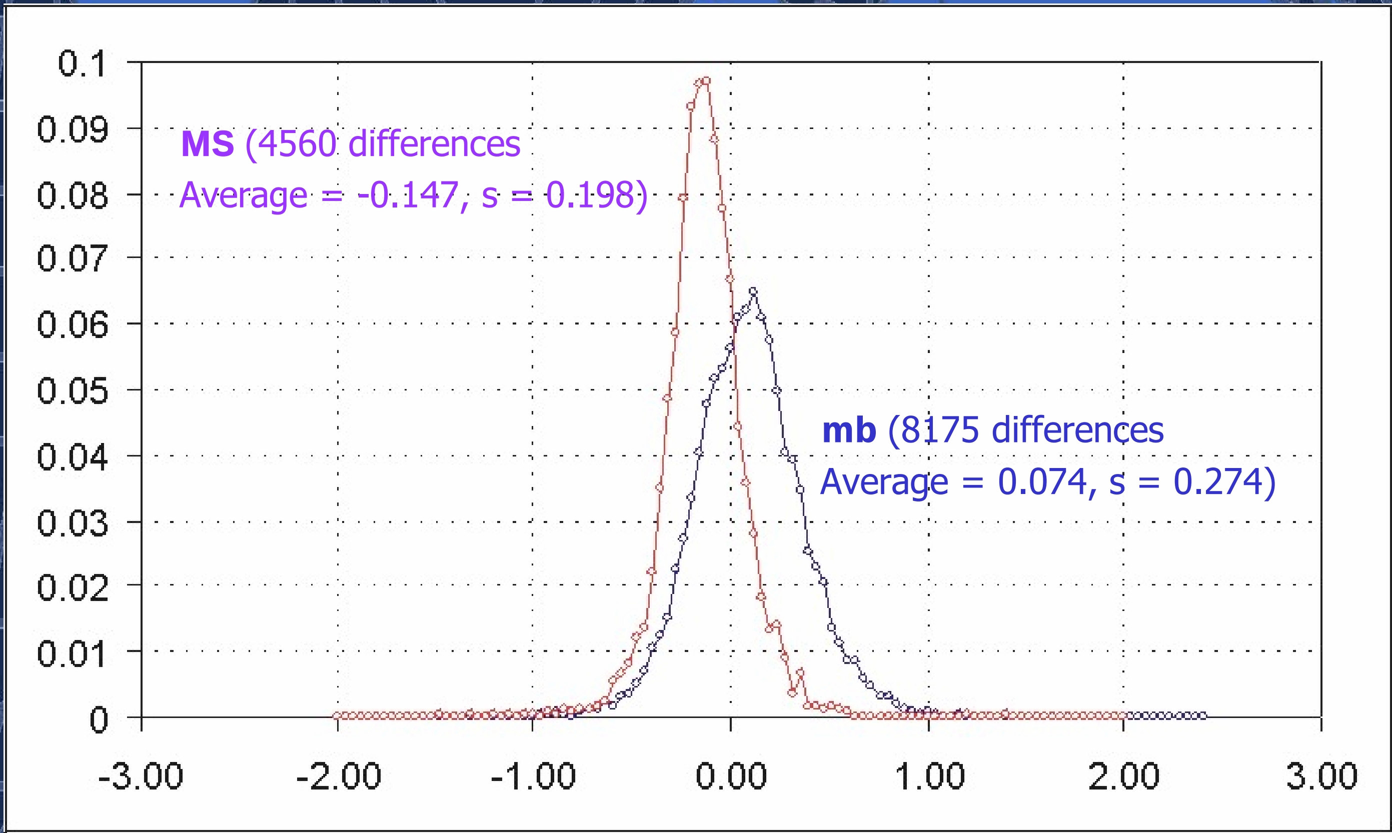
Uncertainty in magnitude determinations

Epicenter distance vs. Station magnitude for the 108 determinations for the 08 September 2002 earthquake NEAR NORTH COAST OF NEW GUINEA



The distribution of the difference between average magnitudes in epicenter and antipodal hemispheres

MCHEDR 1990-2000, all events that have three or more station magnitudes in each hemisphere



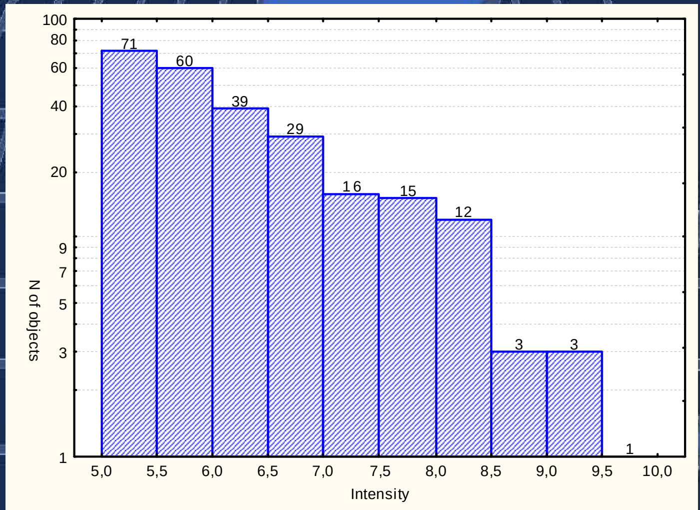
Herak, M. and Herak. D. (1993). Bull. Seism. Soc. Am., 83, 6, 1881-1892.

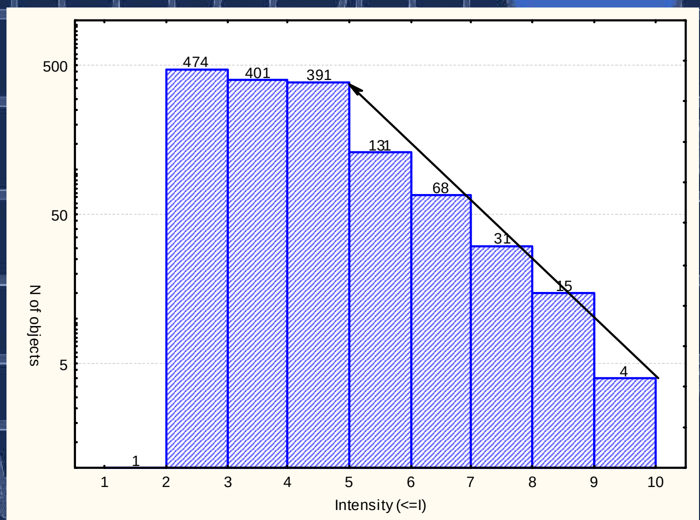
Historical and instrumental catalogs: the example of Spain

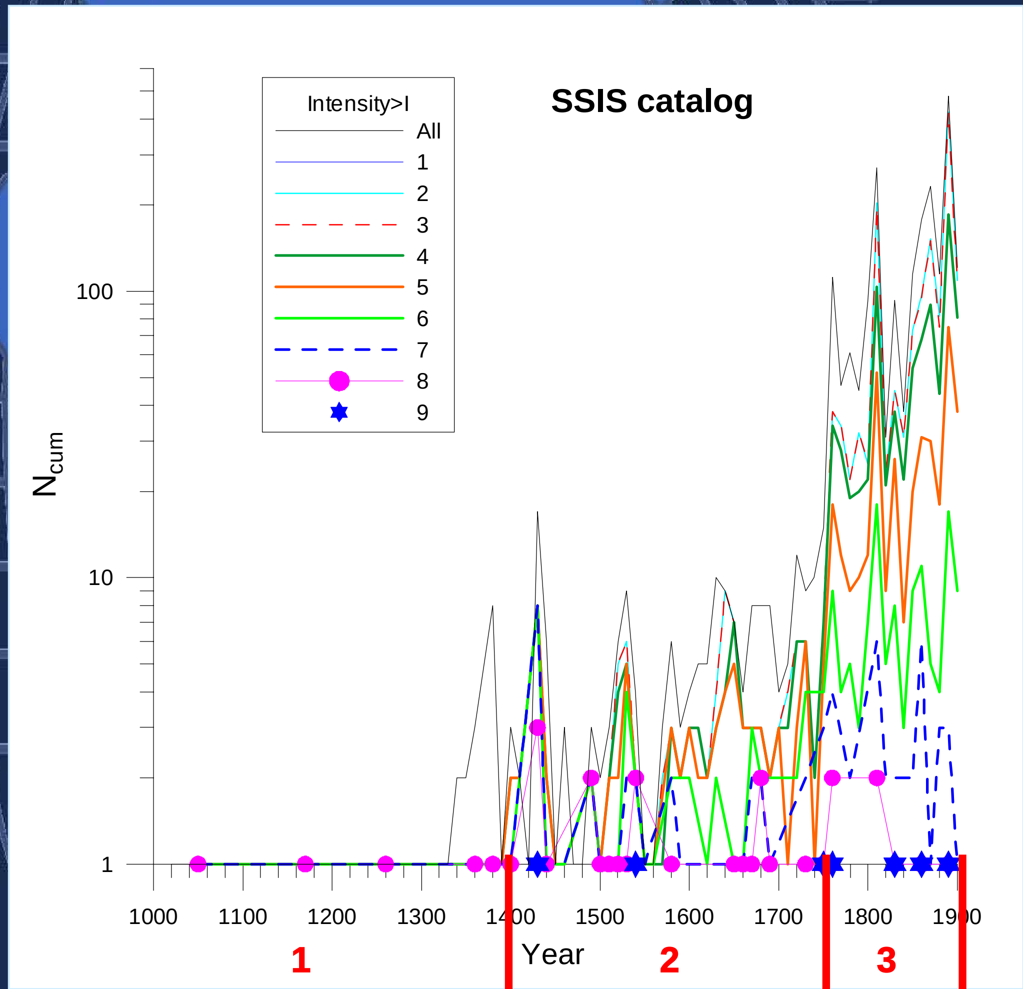
The catalog of seismicity for the Iberian peninsula (SSIS)

Catalogo sismico de la Peninsula Iberica (880 a.C.-1900)

Martinez Solares and Mezcua Rodriguez (2002)



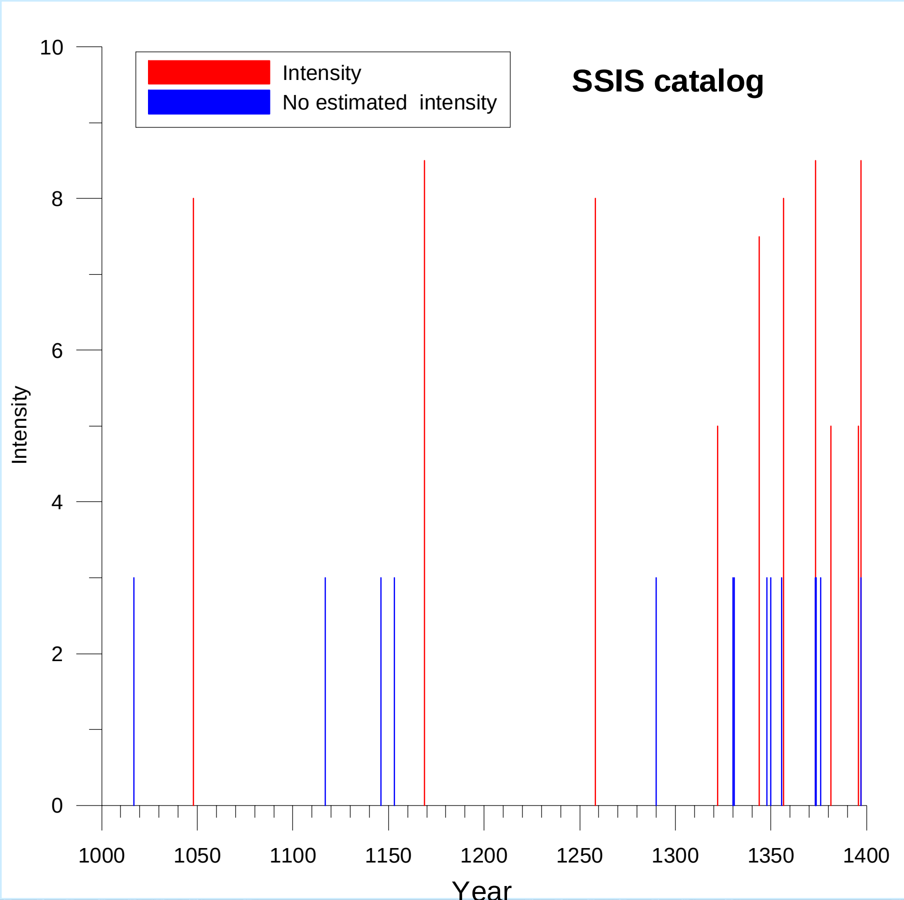


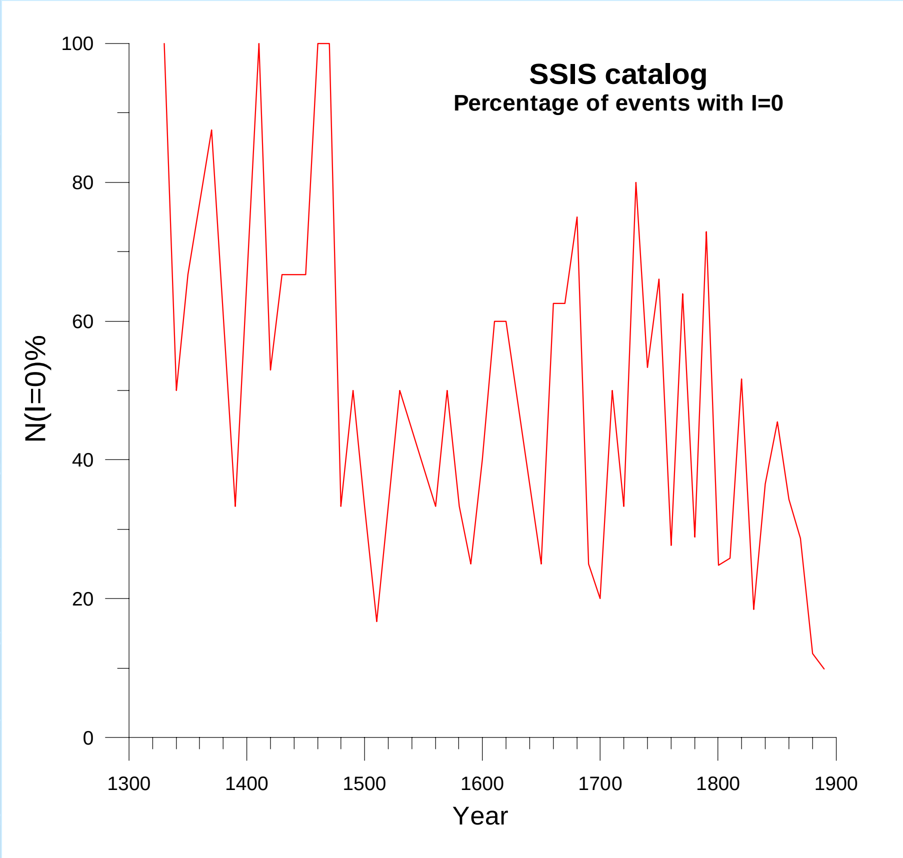


Cumulative number of events vs. time and intensity

The catalog of seismicity for the Iberian peninsula (SSIS)

Time interval: 1000-1400



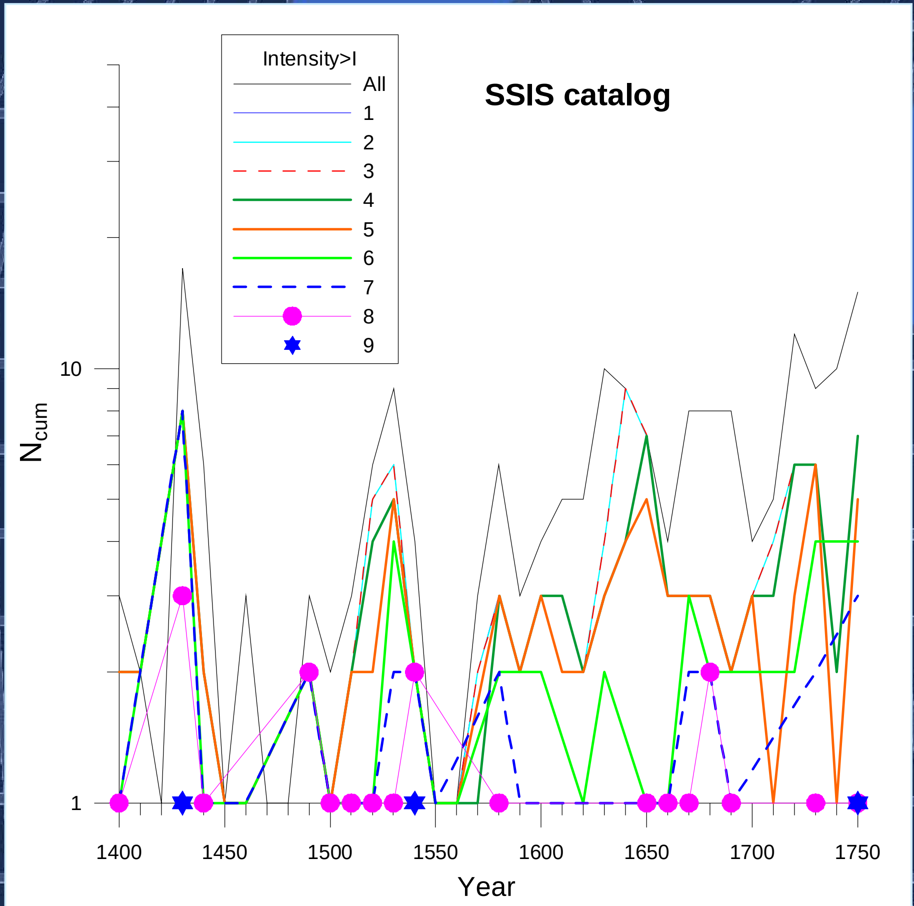


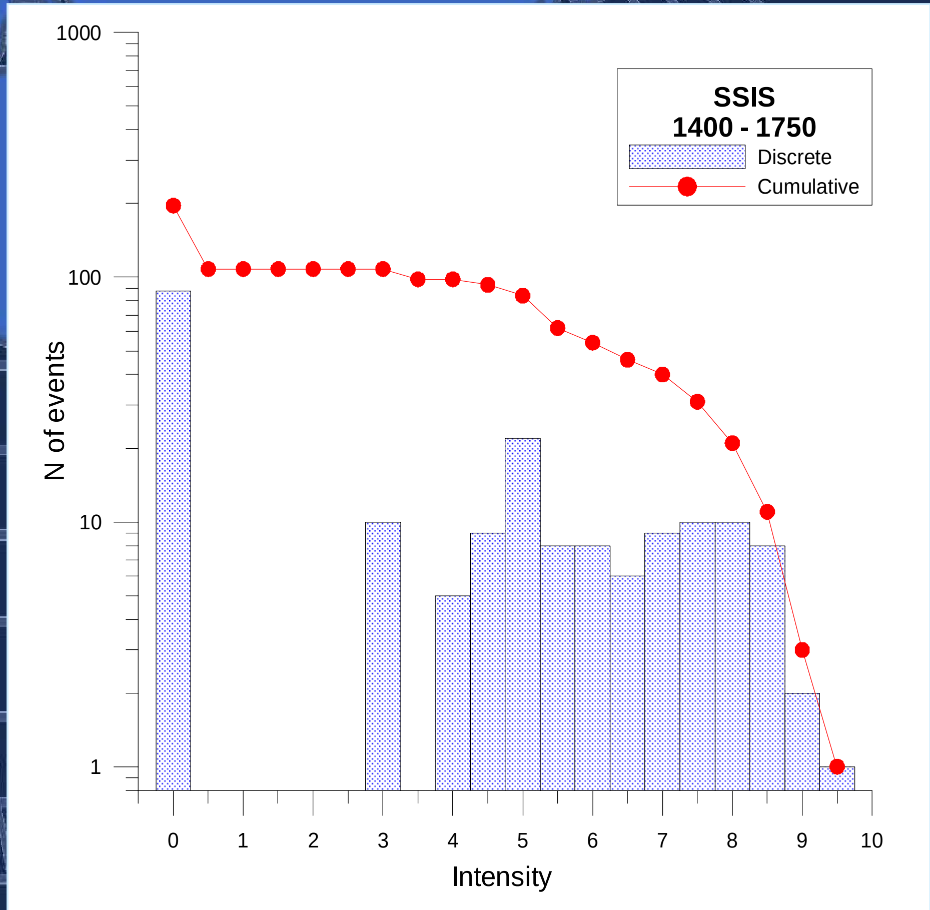
High percentage of events without any intensity estimation in the SSIS catalogue, with some sensible reduction only since 1500

Before 1400 the completeness threshold cannot be clearly identified. The catalogue is not complete for I>8 up to 1360

The catalog of seismicity for the Iberian peninsula (SSIS)

Time interval: 1400-1750

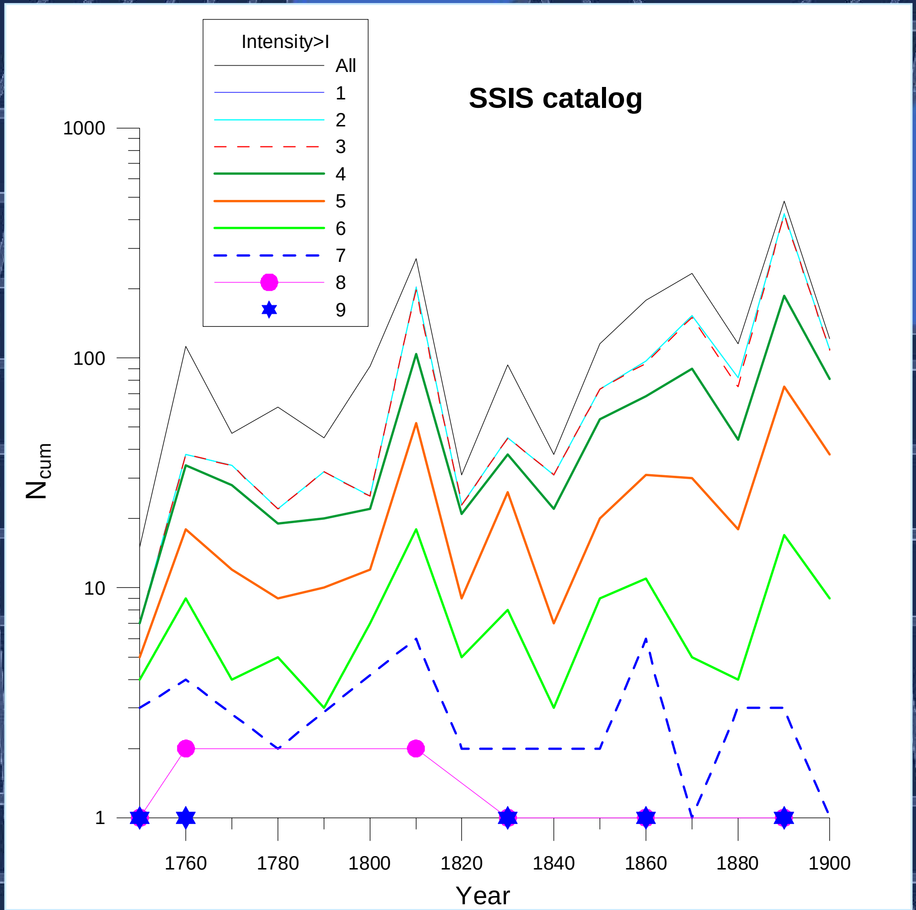


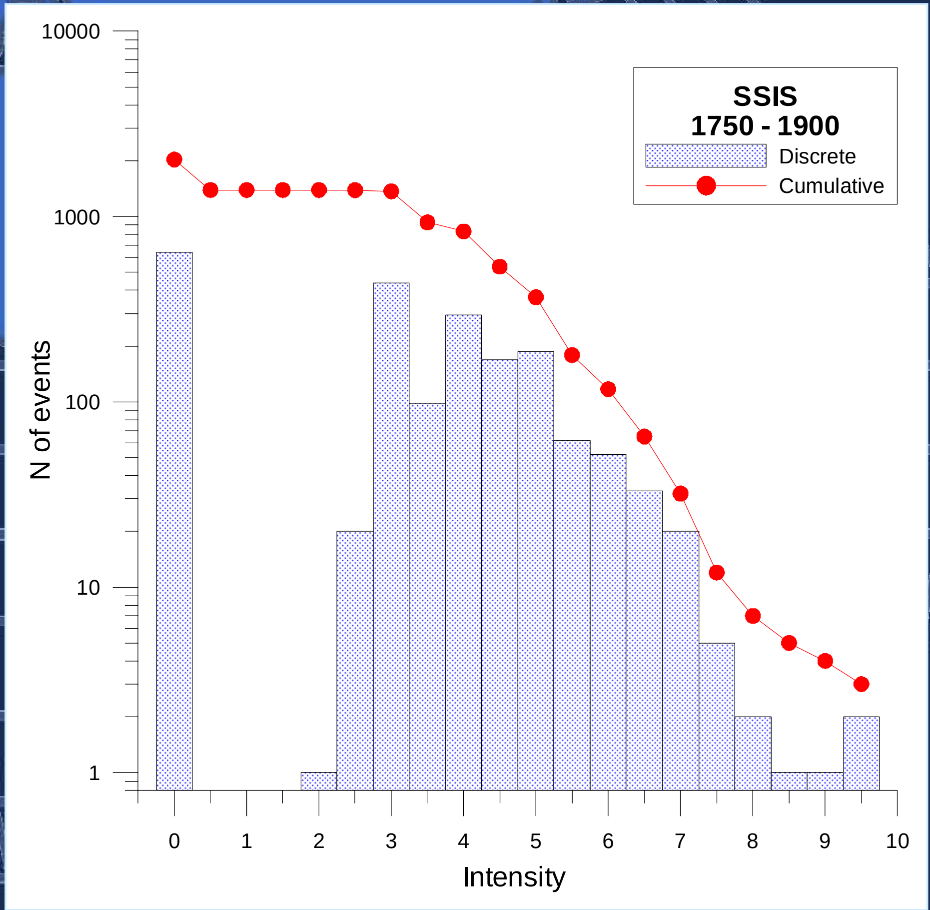


The catalogue appears complete for I>8 during the period 1400-1750

The catalog of seismicity for the Iberian peninsula (SSIS)

Time interval: 1750-1900

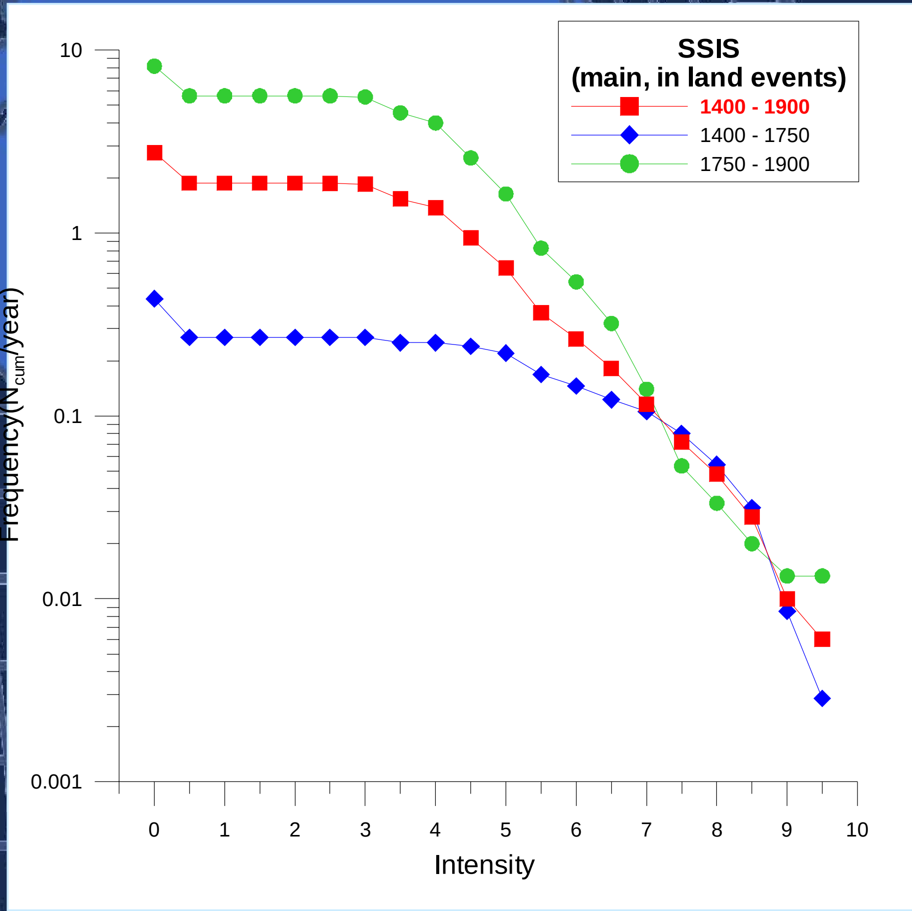




The catalogue appears complete for I>6 during the whole period 1750-1900 and for I>5 (and eventually for I>4) since 1986.

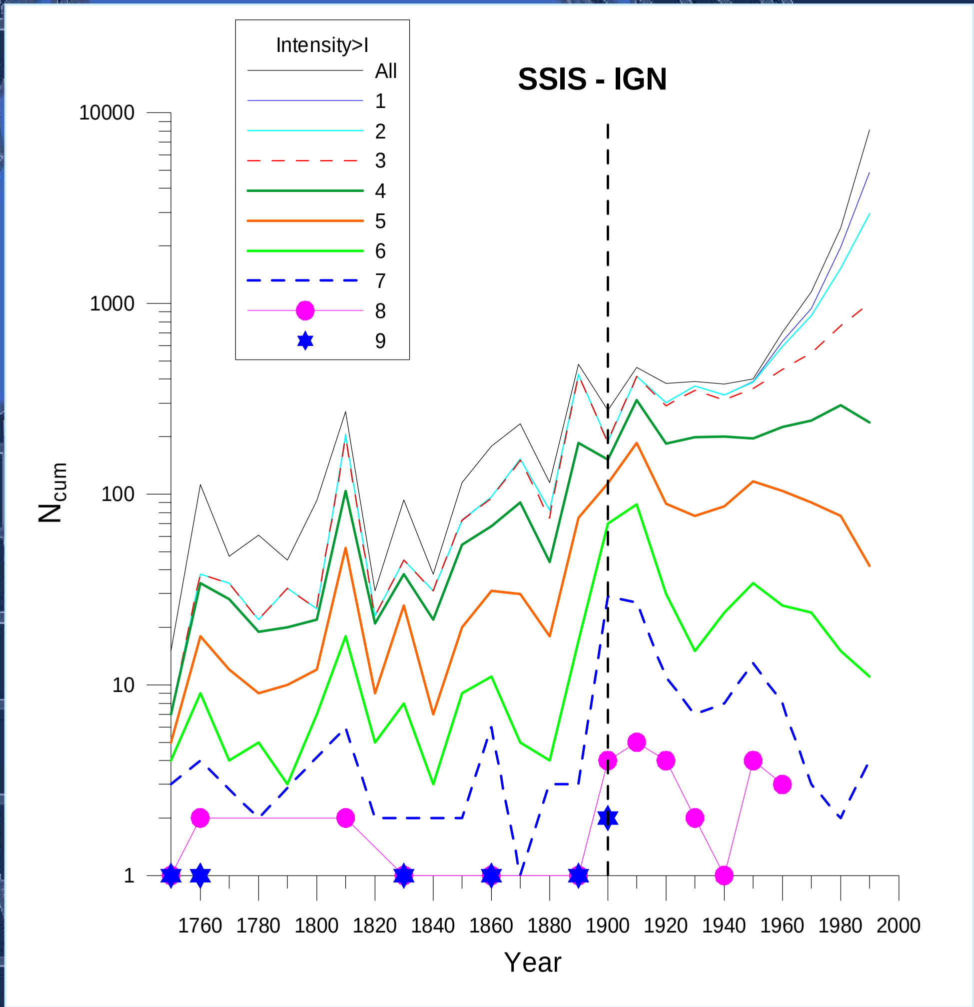
The catalog of seismicity for the Iberian peninsula (SSIS)

- The frequency-intensity distributions, obtained for the SSIS catalog over three different time intervals, are comparable only in the intensity range 7.0-8.5.
- The cumulative distribution for the whole time interval 1400 – 1900 appears preferable for I>7.
- The frequency of the events with I>9 appears much larger during the period 1750-1900 than for 1400-1750.
- The catalog seems to be quite complete and homogeneous for intensities I>6 after 1750; nevertheless the time interval 1750-1900 alone is not enough to characterise the frequency of the large events.



The catalog of seismicity for the Iberian peninsula (SSIS)

The overall increase in the number of earthquakes seems not due to an increased "detection" level, that should not affect the largest intensities, but would rather imply a progressive increase of the number of small events.
We argue that the intensities reported in the two catalogs SSIS and IGN are not homogeneous and, if used all together, do not provide a consistent picture of seismicity.



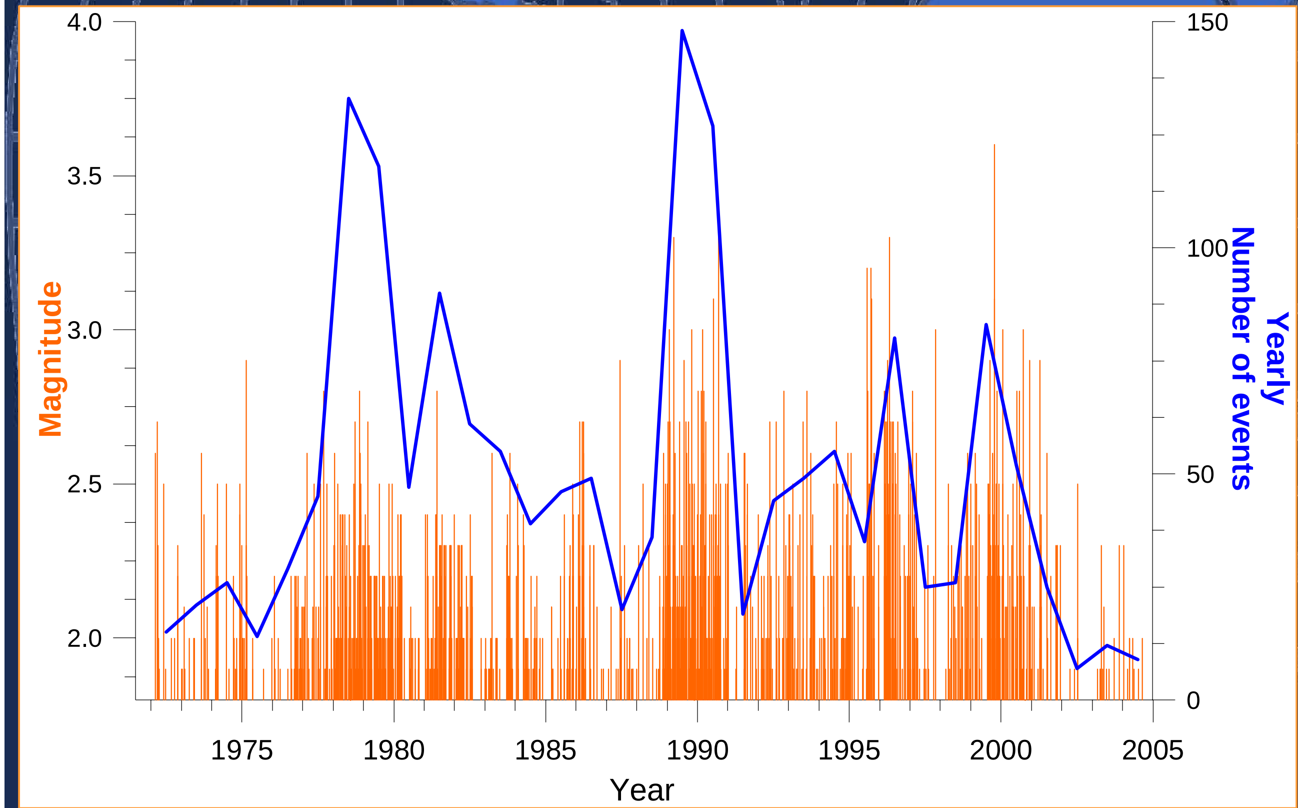
Cumulative number of events vs. time for the SSIS catalog (1750 –1900) and for the IGN catalog (1900 – 2000). For the events reported without any intensity in the IGN catalogue, the intensity recalculated from IGN magnitude mb (e.g. Lopez-Casado et al., 2000**) is considered.**

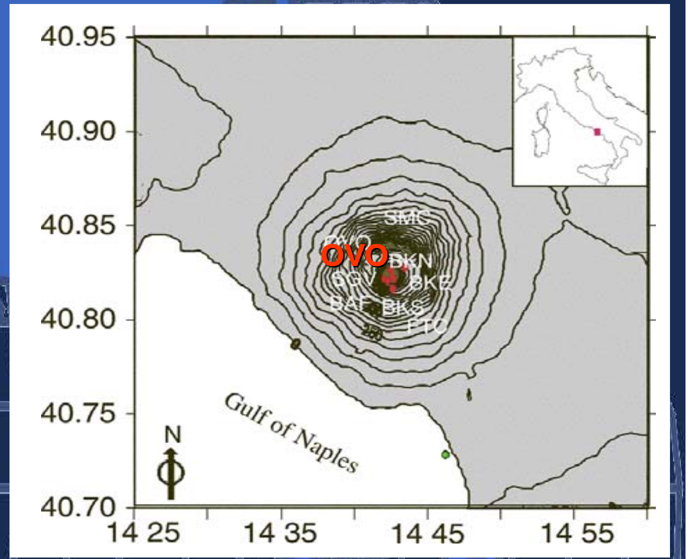
Comparing low magnitude local catalogs: seismicity at Mt. Vesuvius

The OVO earthquake catalog

Data used: catalog of volcanic earthquakes recorded at the station OVO (Osservatorio Vesuviano – INGV, Naples)

Time period: 1972-2004

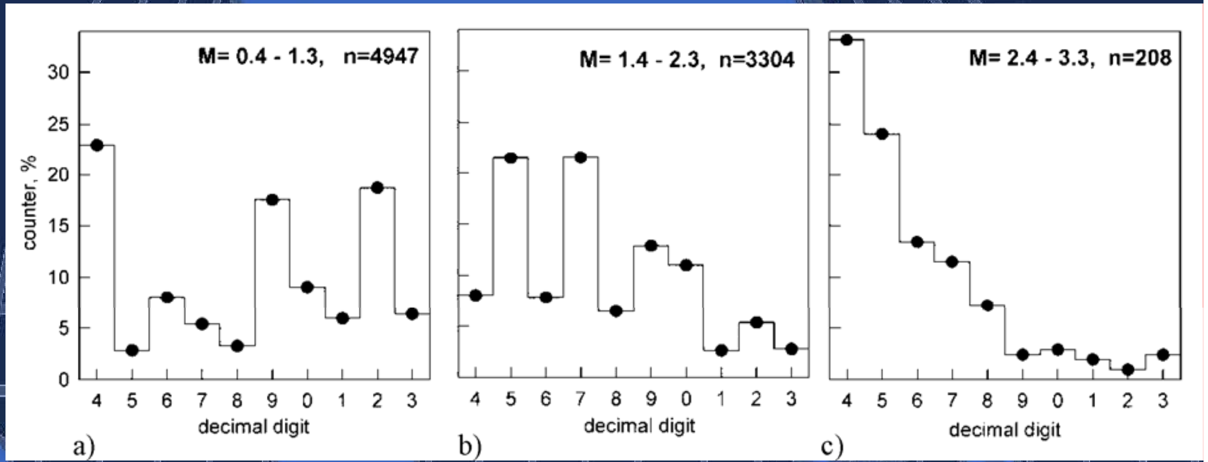




Time sequence of the events M(t) and yearly number of earthquakes with $M \geq M_C = 1.8$ reported in the OVO catalog

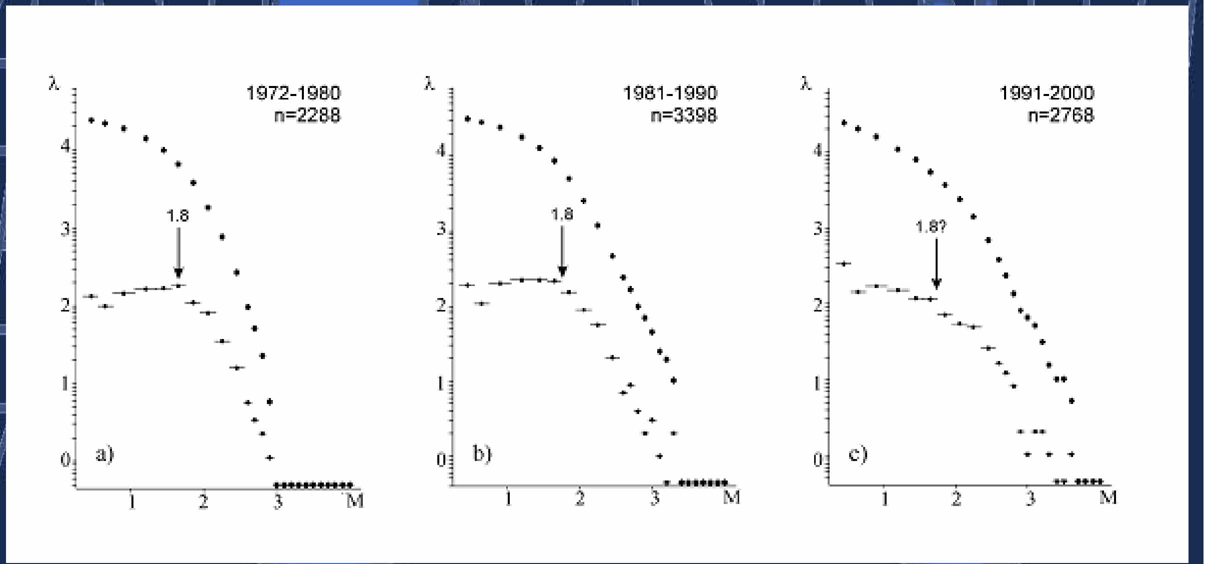
The OVO earthquake catalog

Magnitude grouping: The analysis evidences whether there are dominating values of magnitudes. It permits to choose the appropriate intervals of magnitude grouping ΔM to be considered for the frequency-magnitude distribution.



Catalog completeness: The completeness of the catalog is determined from the frequency-magnitude distribution λ(M), where λ is the number of earthquakes within each magnitude grouping interval ΔM, normalized to the space-time-magnitude volume unit V=[1000 km2 x 1 year x 1 M].

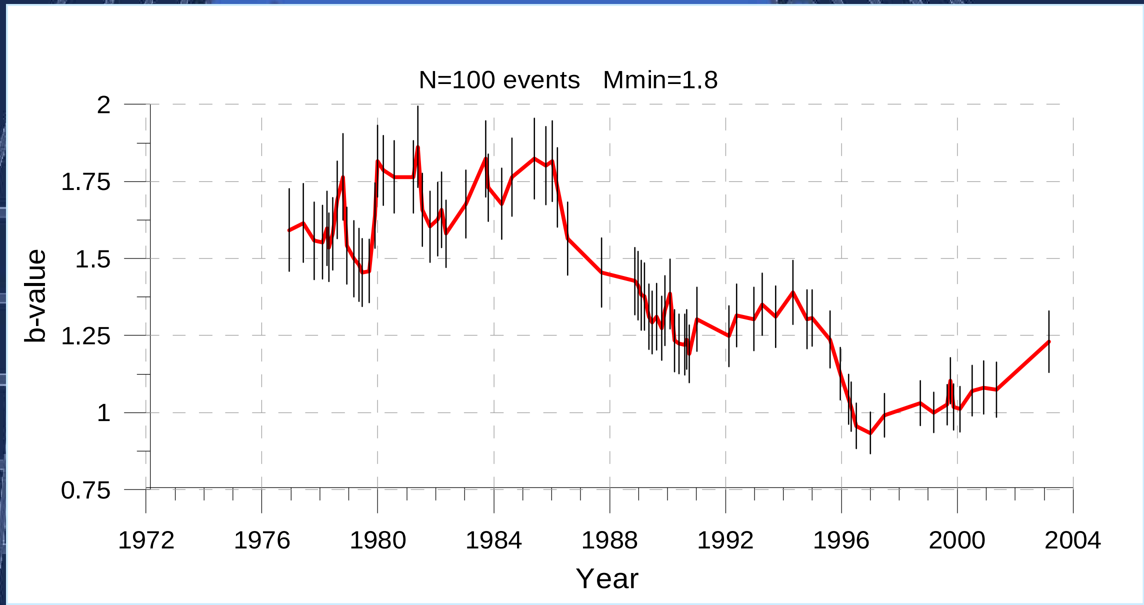
$M_{completeness} = M_C = 1.8$



Time changes in seismic activity: analysis of the OVO catalog

- **The time variations of the b-value in the Gutenberg-Richter (GR) law, are analysed and show that it decreases progressively from 1.8, before 1986, to about 1.0 in 1996.**

(Maximum likelihood estimation by Wiemer & Zuniga, ZMAP software)



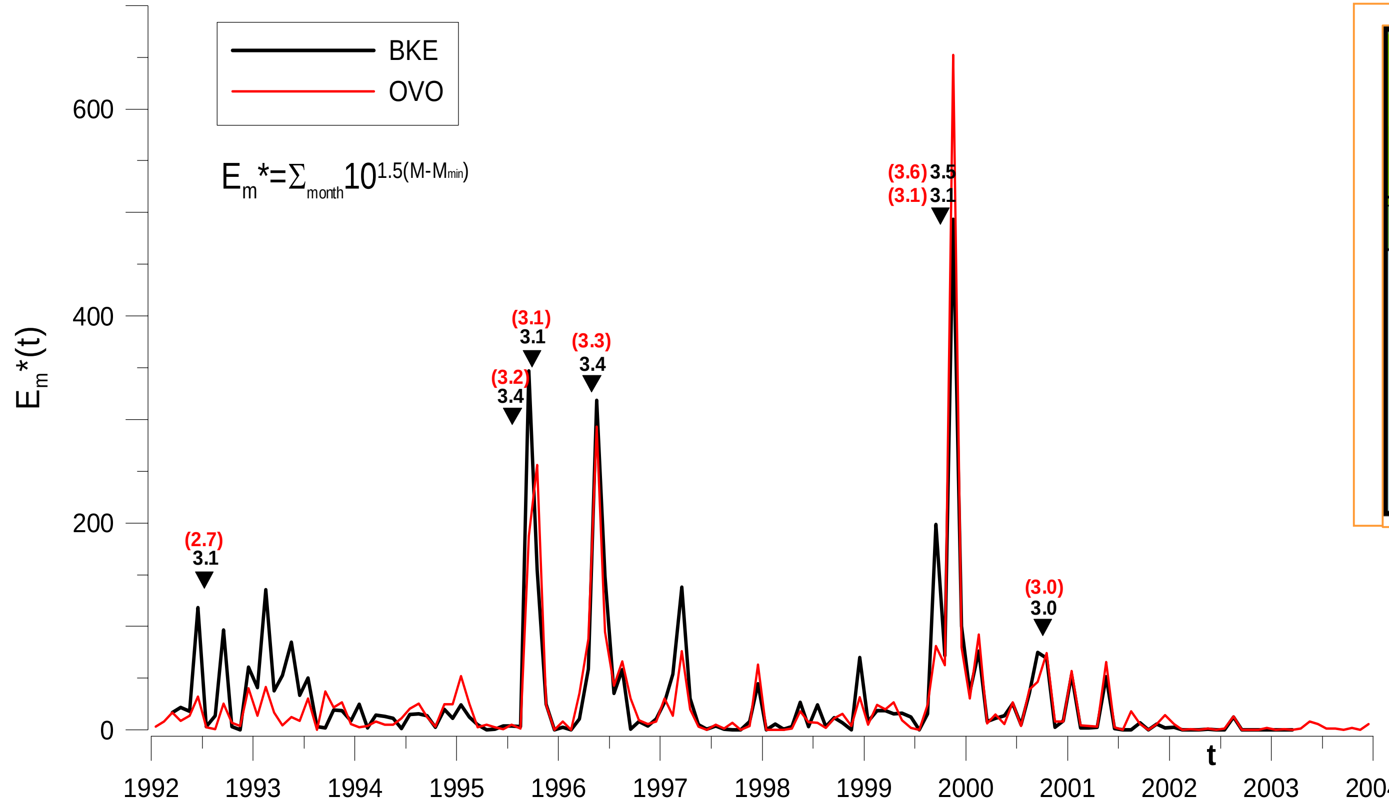
- **The seismic energy release is studied considering the quantity *E**, energy normalised to the minimum magnitude event, computed from magnitude according to the formula:**

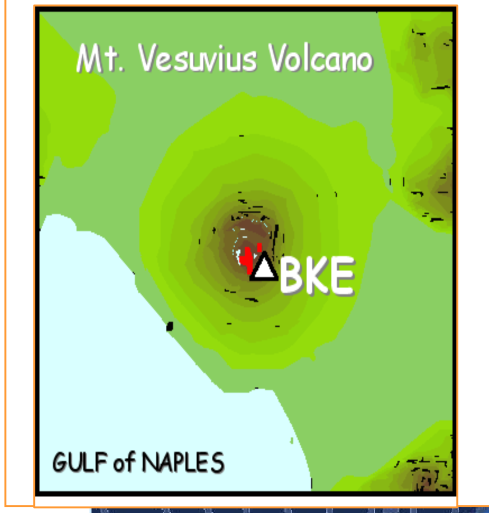
$$E^* = 10^{d(M - M_{\min})}$$

d = const

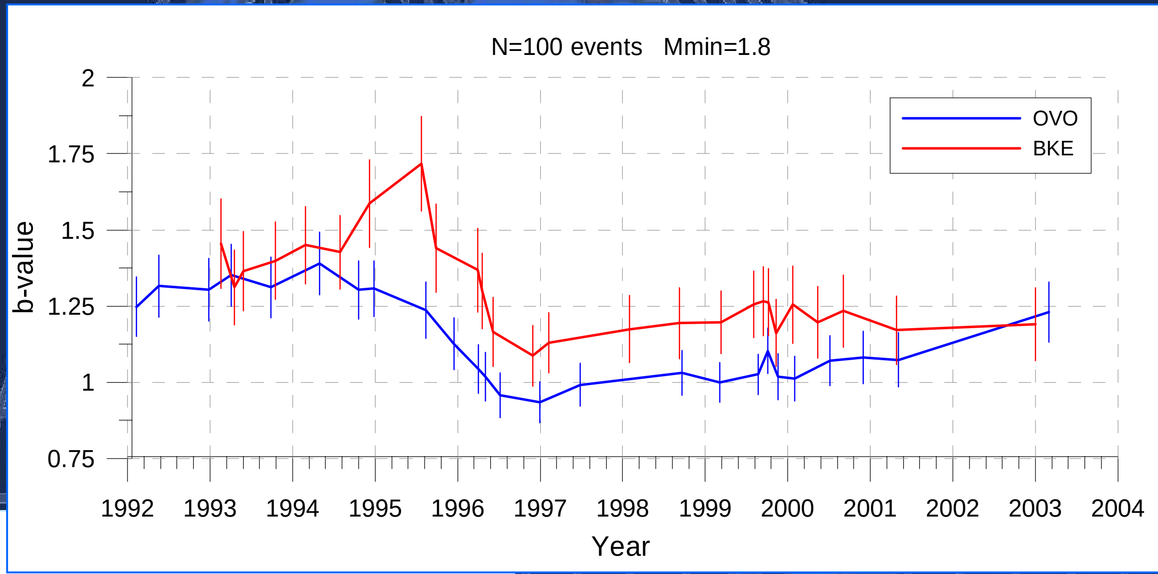
Time changes in seismic activity: comparison with the BKE catalog

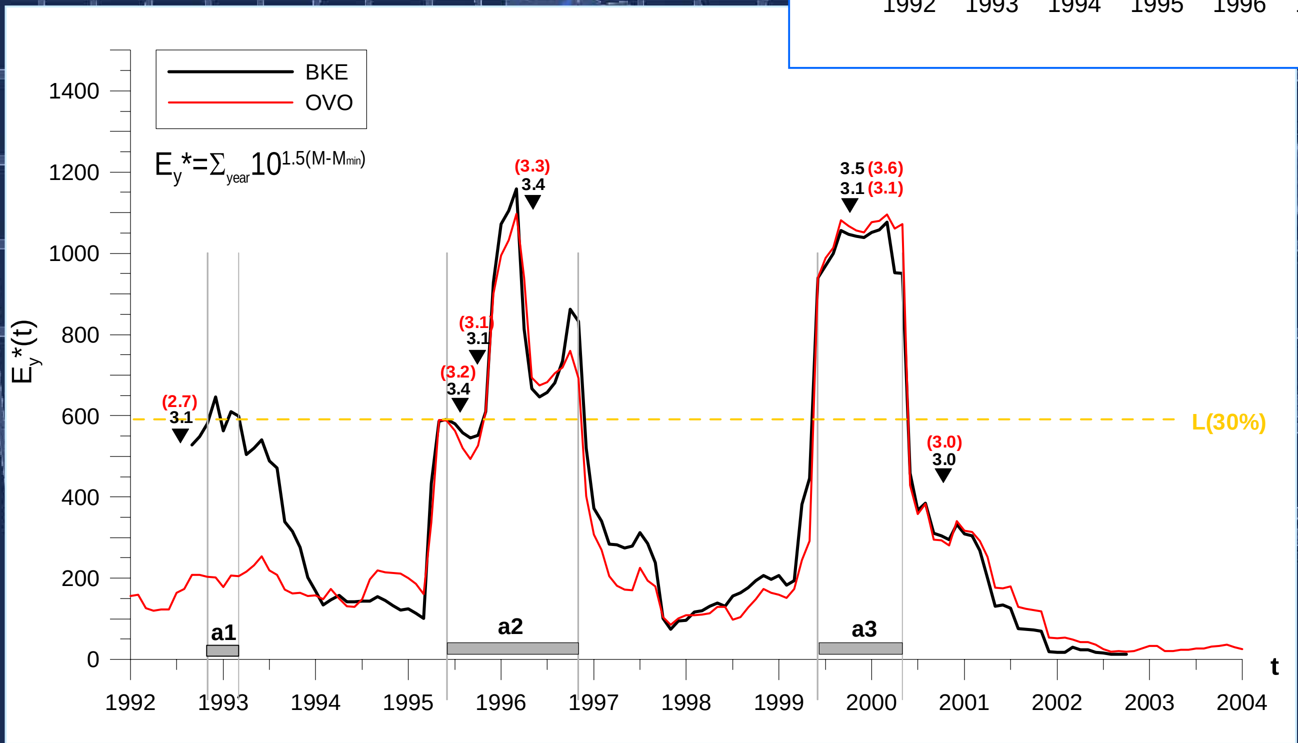
The time variations of the b-value and seismic energy release, observed for the OVO catalog, are checked performing a similar analysis with a different catalog of Vesuvian earthquakes, as compiled from the records at the BKE station during the period 1992-2003.





Time changes in seismic activity: comparison with the BKE catalog

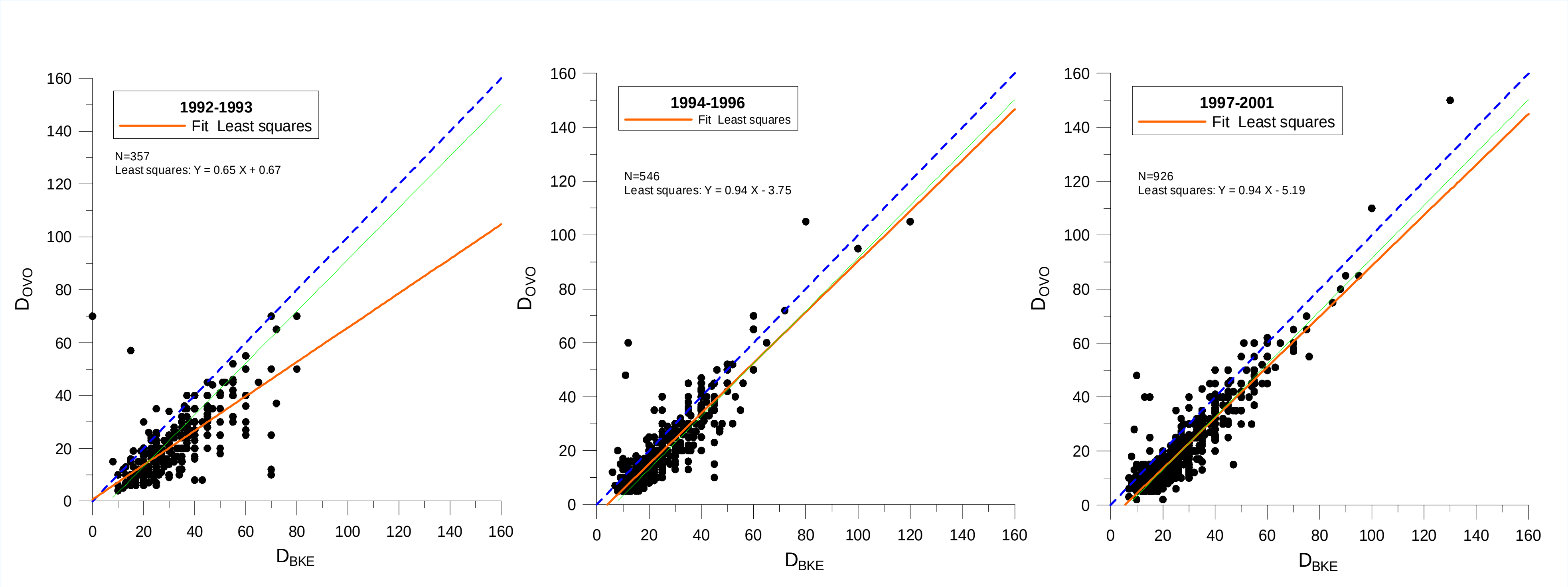




The analysis performed using the catalog compiled for the BKE stations confirms the b-value decrement and the identification of the periods of quiescence and activity, since 1994.

Time changes in seismic activity: comparison with the BKE catalog

The differences observed during the period 1992-1994 are well explained by a certain overestimation of BKE durations during such period of time, as shown by the comparison of OVO and BKE durations for the common events



Problem: identification of the common events in catalogs characterised by low magnitude and highly clustered events

Analysis of Earthquake Catalogs for Earthquake Prediction purposes: the Case of Italy

Algorithms for middle-range intermediate-term prediction

Algorithms fully formalized and globally tested for prediction are:

- **CN algorithm** (*Gabrielov et al., 1986; Rotwain and Novikova, 1999*)
- **M8 algorithm** (*Keilis-Borok and Kossobokov, 1987; Kossobokov et al., 1999*)

They allow to identify the TIPs (Times of Increased Probability) for the occurrence of a strong earthquake within a delimited region

Algorithms for middle-range intermediate-term prediction

The algorithms are based on a set of empirical functions to allow for a quantitative analysis of the premonitory patterns which can be detected in the seismic flow:

- Variations in the seismic activity
- Seismic quiescence
- Space-time clustering of events

These methods make use of detectable inverse cascade of seismic process, at different space and time ranges, to reduce consecutively space and time limits where a disastrous earthquake has to be expected.

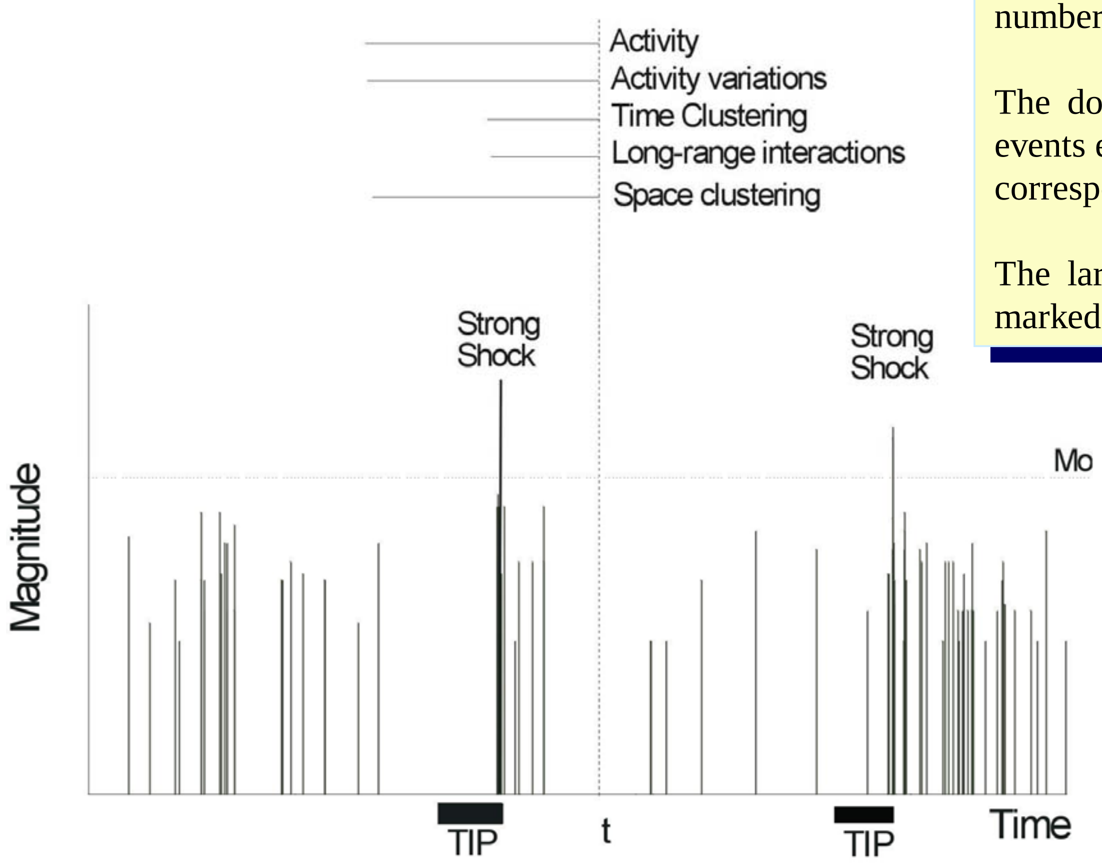
•**DST:**

2) Quiescence:

The sign + indicates that the sum includes only the positive terms; therefore only the time intervals (t-s,t) where the number of earthquakes is less than the average are considered.

The dotted horizontal line indicates the average number of events expected in the time interval of length s. The grey areas correspond to the periods of quiescence.

The larger is the value assumed by the function, the more marked and prolonged is the quiescence.



Seismic flow

Time sequence of the earthquakes occurred within a delimited region

TIP

Time of Increased Probability

Interval of time when the probability for the occurrence of a strong earthquake, within a delimited region, increases with respect to the normal conditions

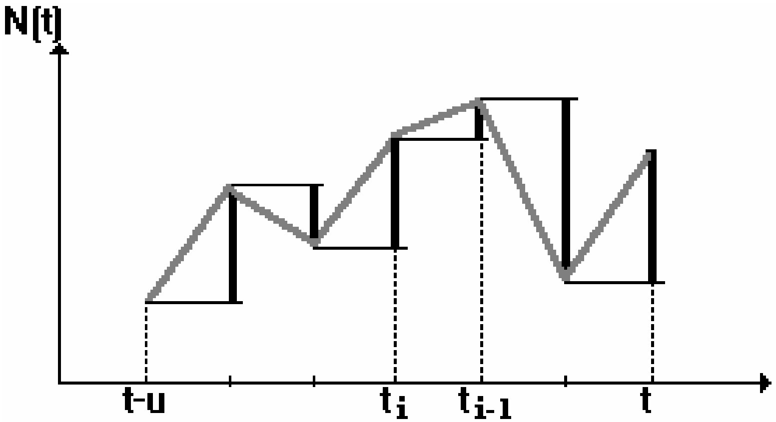
Functions of the seismic flow

DST:

Variation of seismic activity

Is the sum of the differences between the numbers of earthquakes in two consecutive time intervals.

It corresponds to the sum of the differences between the numbers of earthquakes in two consecutive time intervals, with fixed length s.
The moments belong to the time interval (t–u,t), where u is multiple of s.
Usually s is equal to one year.



$$V(t|\underline{M},s,u) = \sum_i \left|N(t_i|\underline{M},s) - N(t_{i-1}|\underline{M},s)\right|$$

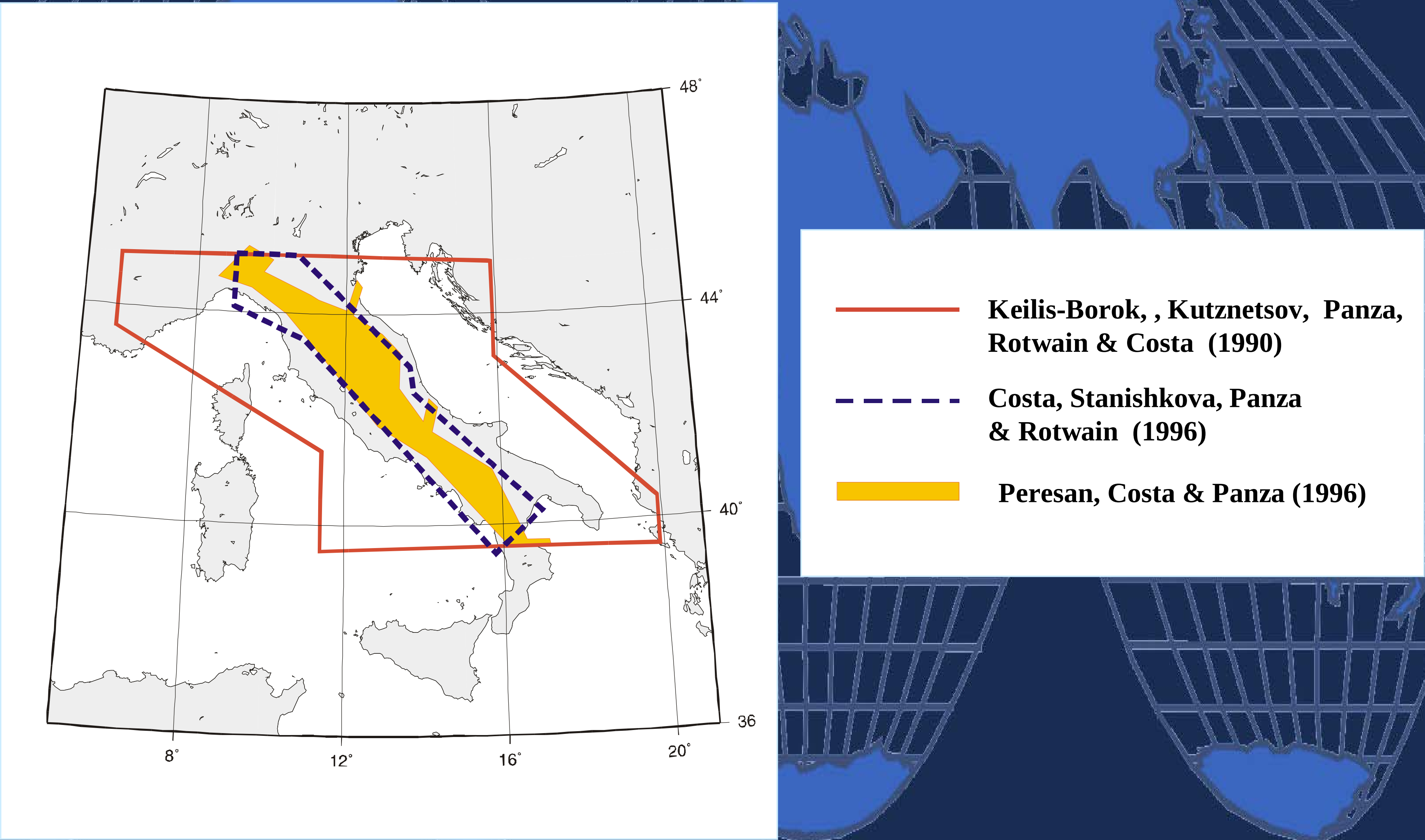
$i=1,\dots,n \qquad u=ns$

Variation of seismic activity

$$V(t|\underline{M},s,u)$$

It is a measure of the cumulative changes in time of the earthquake number

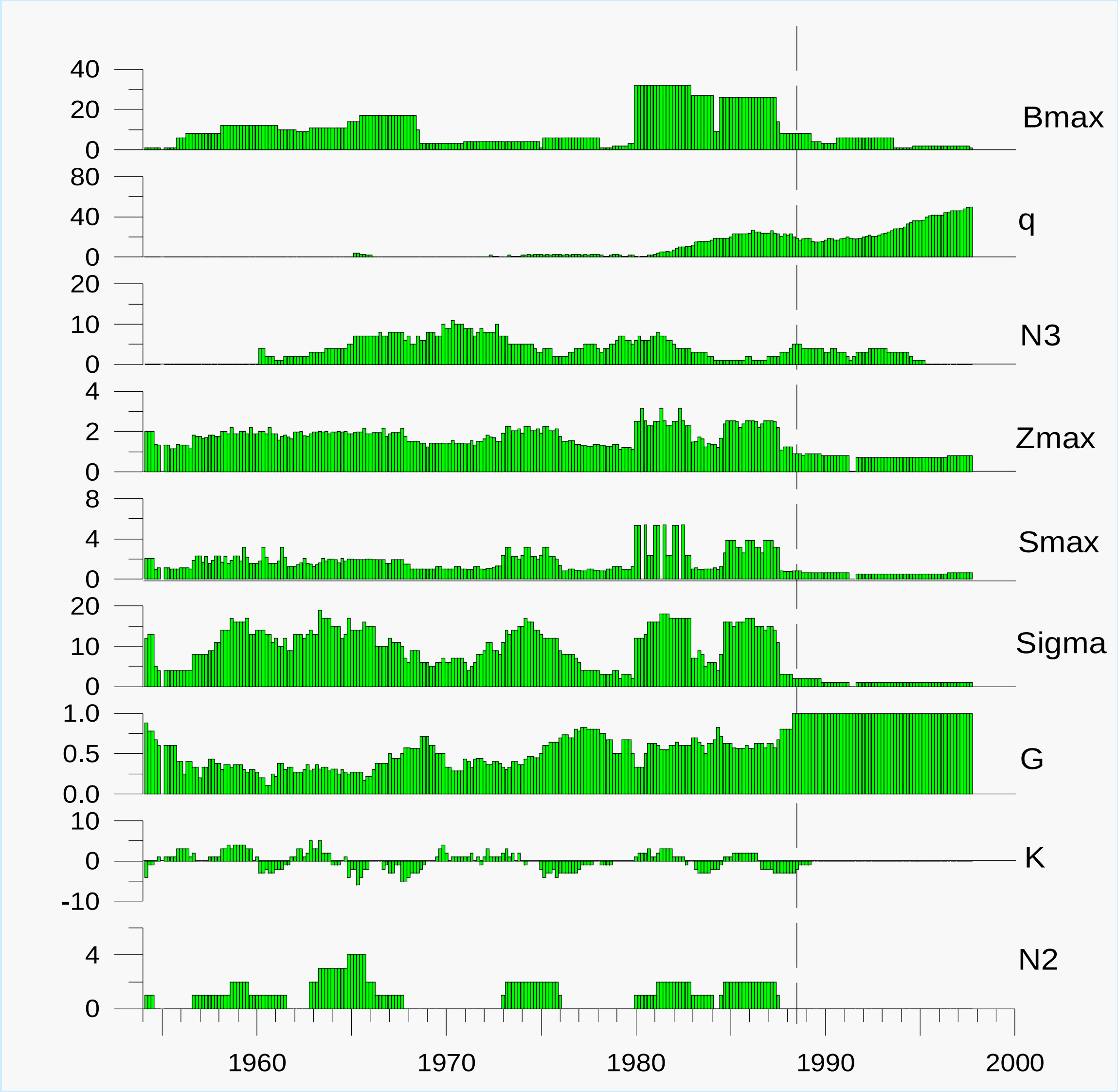
CN algorithm and long lasting changes in reported magnitudes



CN algorithm and long lasting changes in reported magnitudes

Time diagrams of the standard CN functions obtained for the Central region

(Peresan et al., 1999)



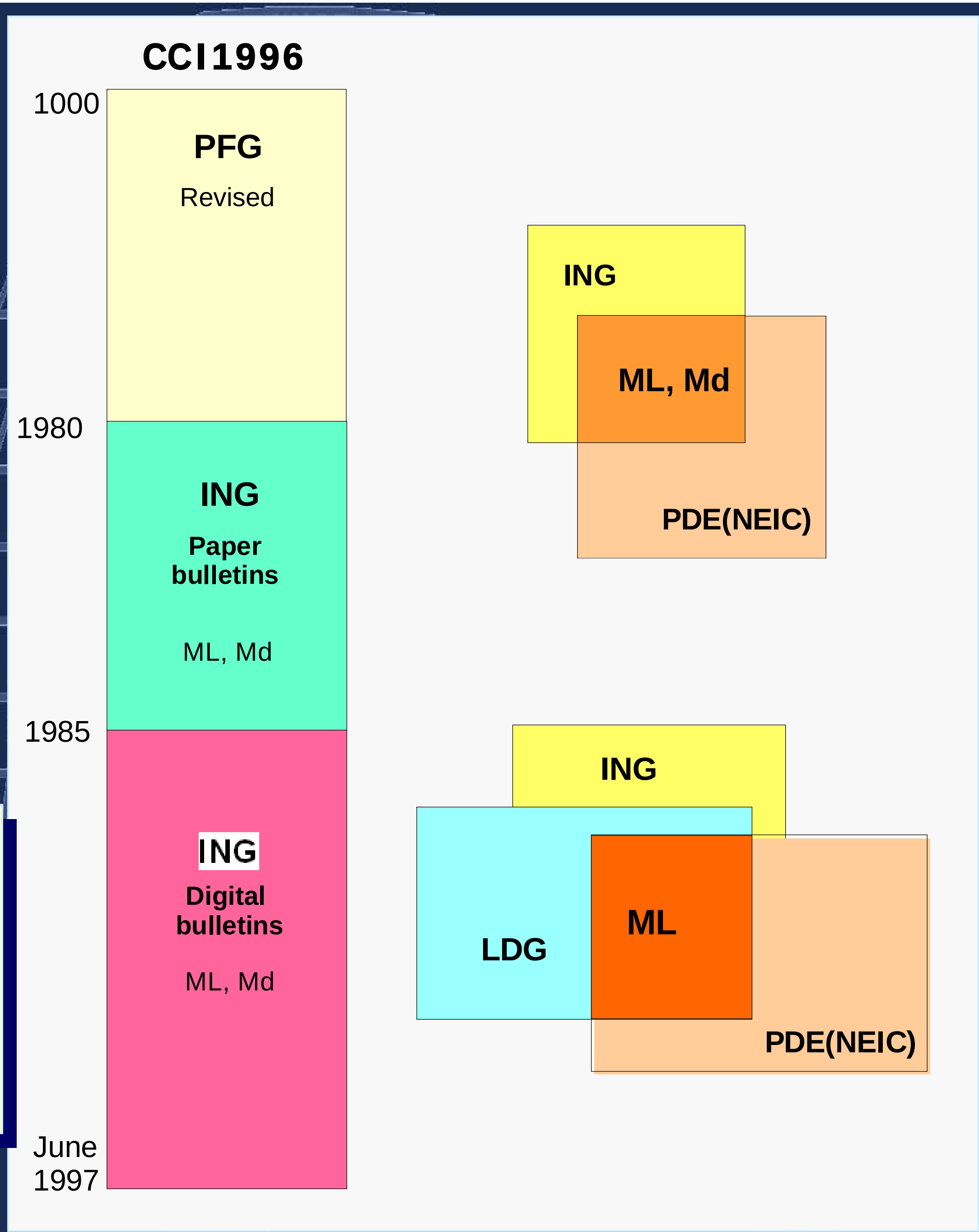
Functions *Sigma*, *Smax* and *Zmax* and are evaluated for 4.2≤M≤4.6, functions *K*, *G*, *N3*, *q* for M≥4.5 and functions N2 for M ≥5.0

Catalogues comparison

Common events:
ΔT≤1min

ΔLat, ΔLon≤1°
(Storchak, Bird and Adams, 1998)

•DST:
In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.



Magnitude comparison: Central Region

•DST:
In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.

Duration Magnitude

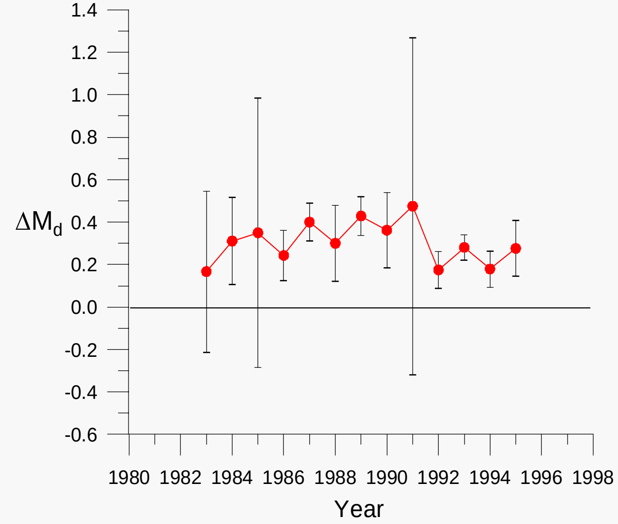
M_d(NEIC)-M_d(ING)

$M_d \geq 3.0$

(1983-1997)

ΔM_d=0.30±0.04

Yearly Average



Yearly Average differences NEIC-ING

Local Magnitude

M_L(NEIC)-M_L(ING)

M_L

$M_L \geq 3.0$

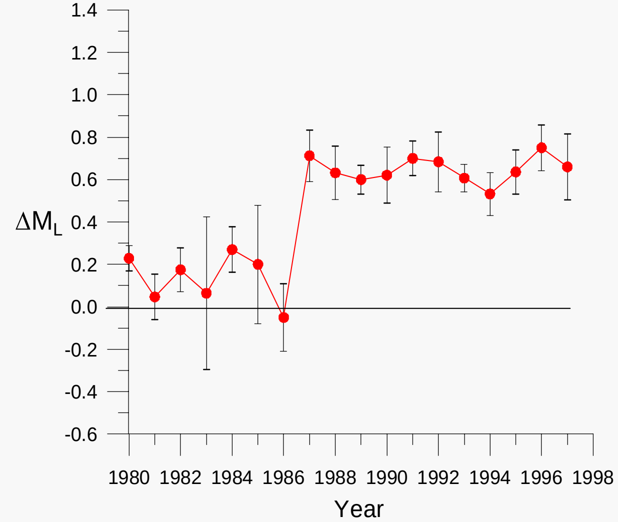
(1980-1986)

ΔM_L=0.13±0.05

(1988-1997)

ΔM_L=0.64±0.04

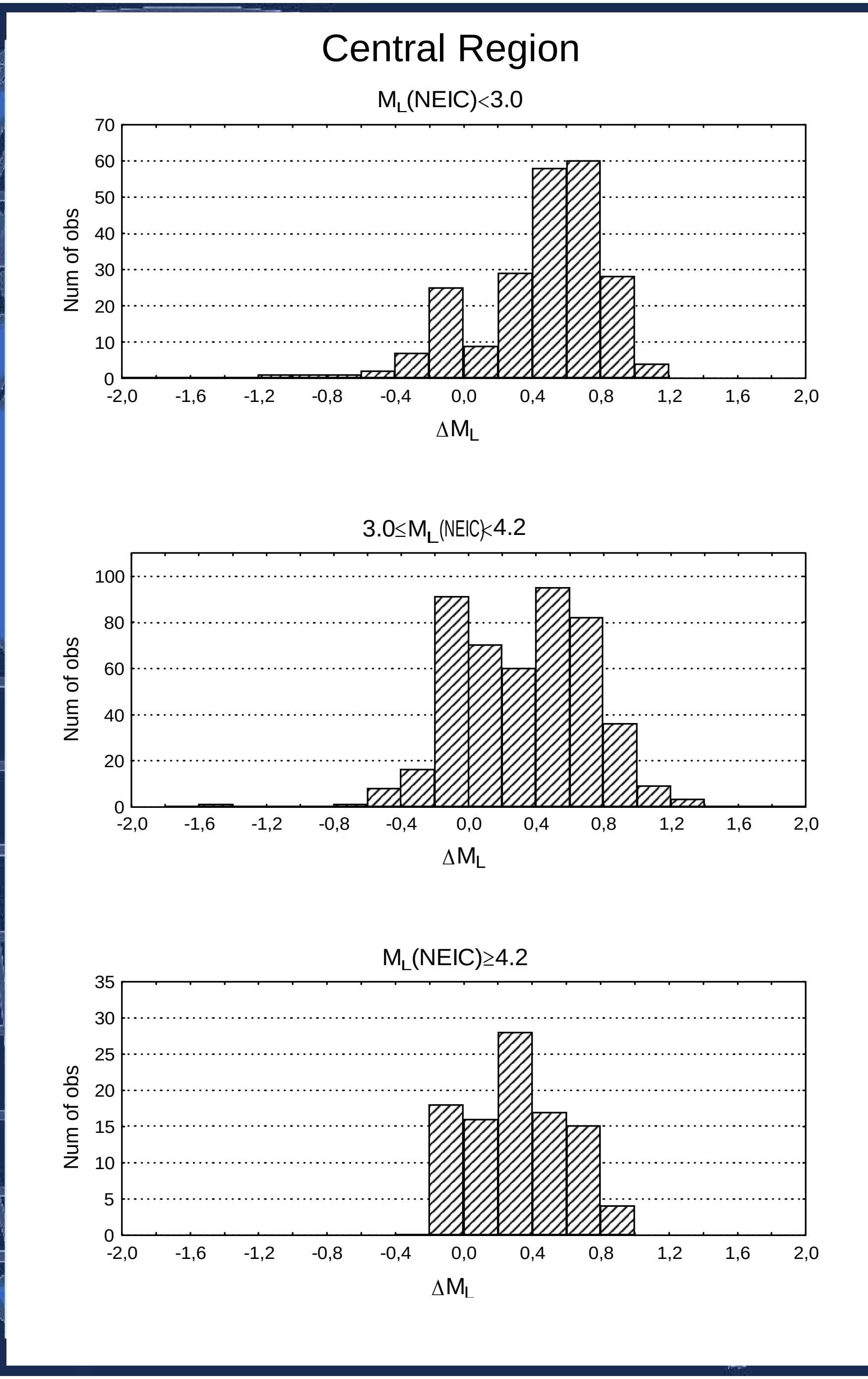
Yearly Average



Magnitude comparison: Central Region

Local Magnitude Differences

$$\Delta M_L = M_L(\text{NEIC}) - M_L(\text{ING})$$



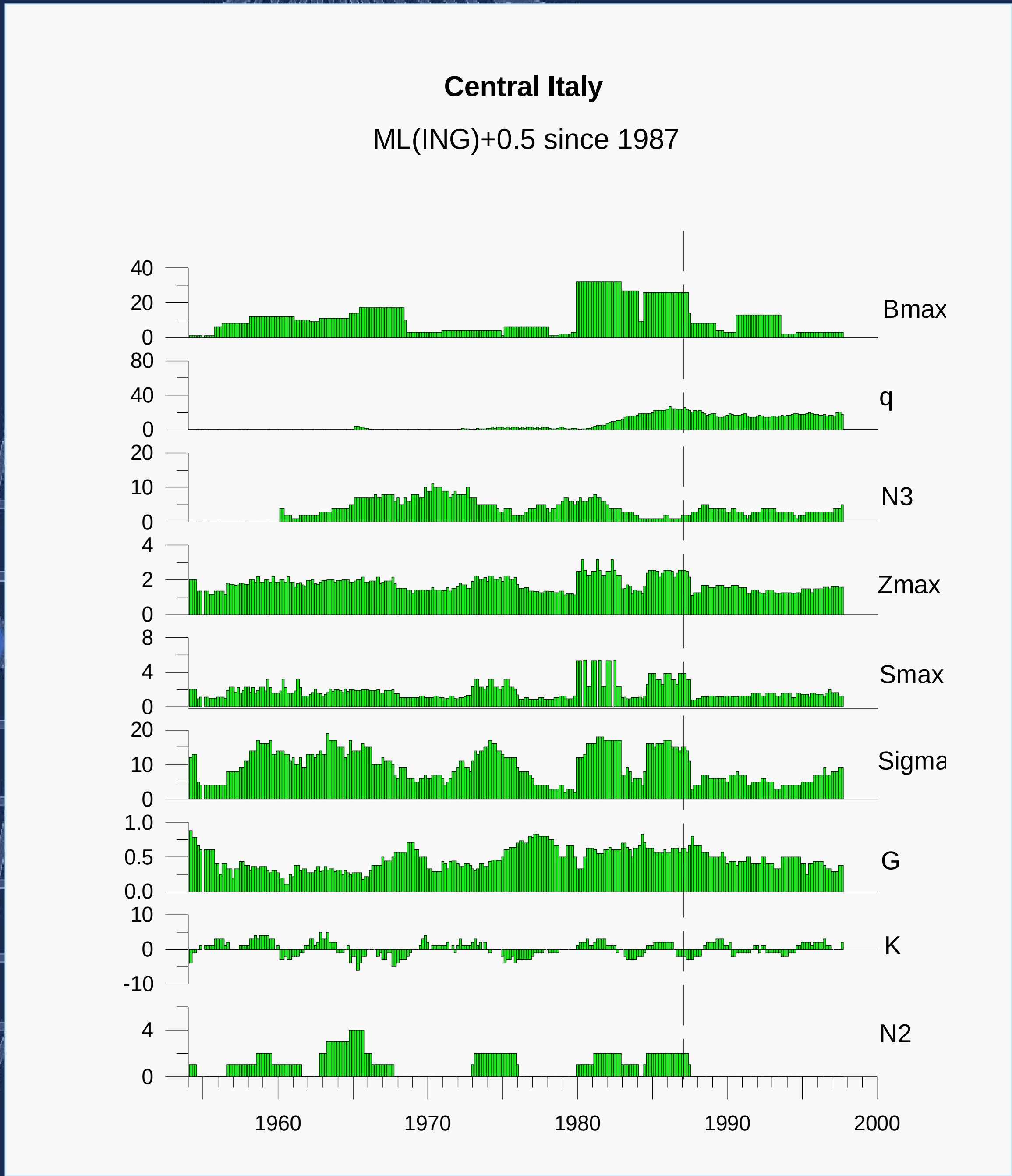
•DST:

In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.

CN algorithm and long lasting changes in reported magnitudes

Time diagrams of the standard CN functions obtained for the Central region *(Peresan et al., 1999)* :

M=ML(ING)+0.5



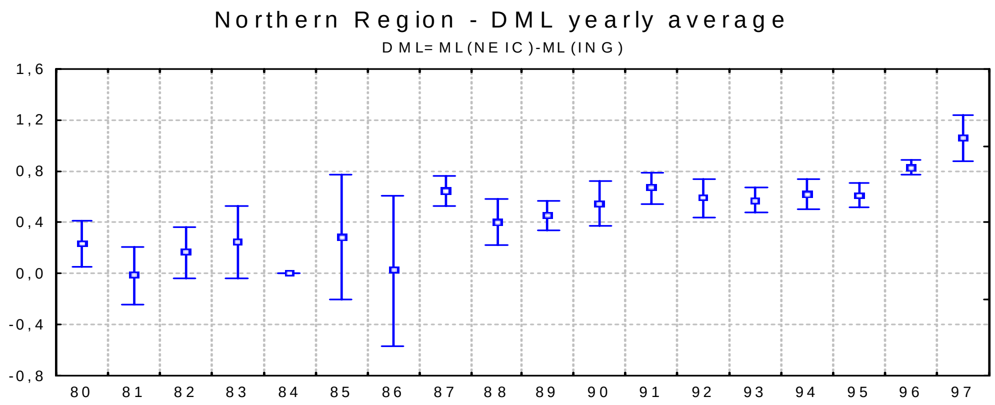
Magnitude comparison: Northern Region

Yearly Average differences NEIC-ING

Local Magnitude

ΔML=ML(NEIC)- ML(ING)

M_L (NEIC) ≥ 3.0



Average difference:

(1980-1986) $\Delta M_L = 0.16 \pm 0.10$

(1988-1997) $\Delta M_L = 0.64 \pm 0.04$

•DST:

In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.

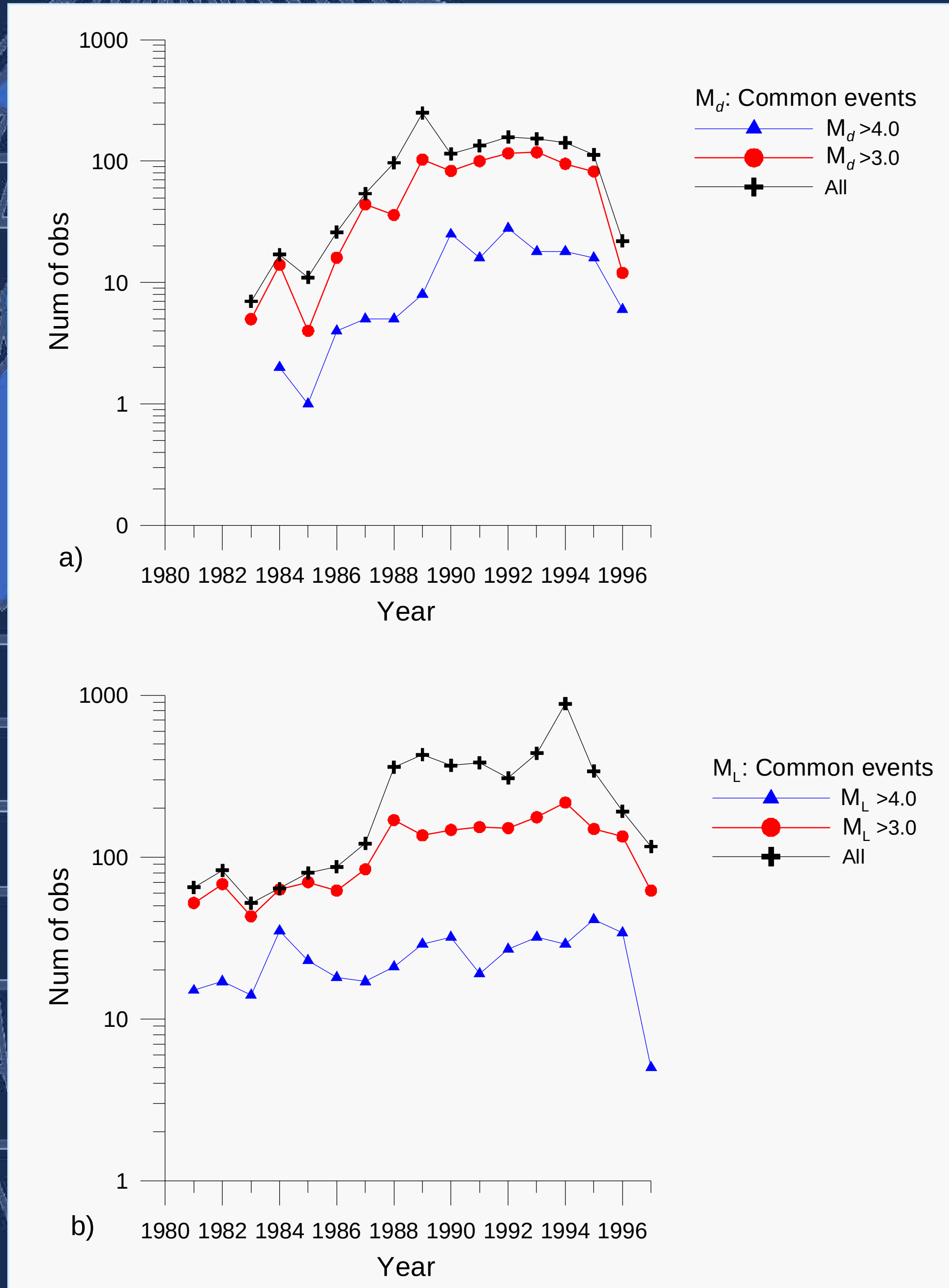
Magnitude comparison: Italian territory

•DST:
In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.

Events used for M_d analysis

Yearly number of common events used for the comparison between ING and NEIC catalogues

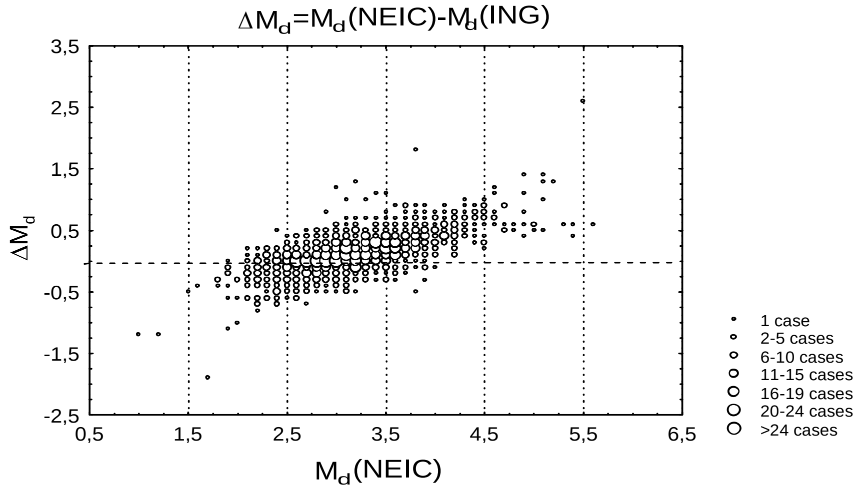
Events used for M_L analysis



Magnitude comparison: Italian territory

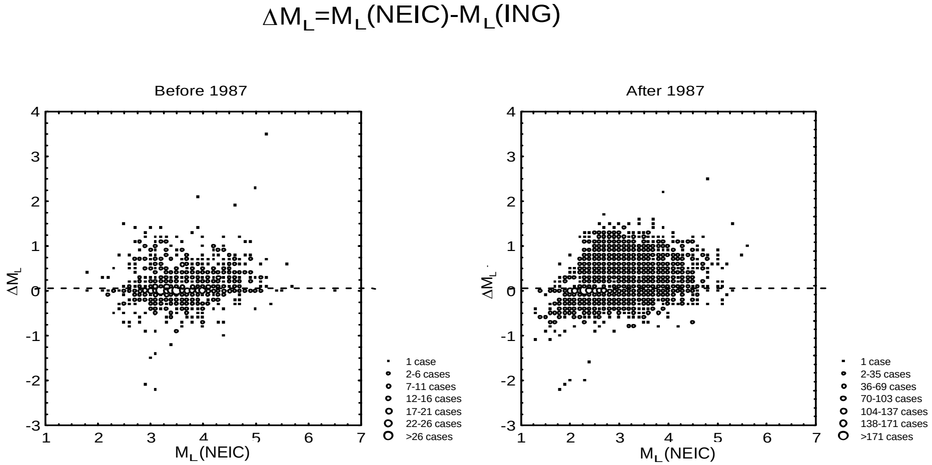
•DST:

In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.



ΔM_d

Frequency scatter-plots of ΔM_d and ΔM_L versus the corresponding NEIC magnitude

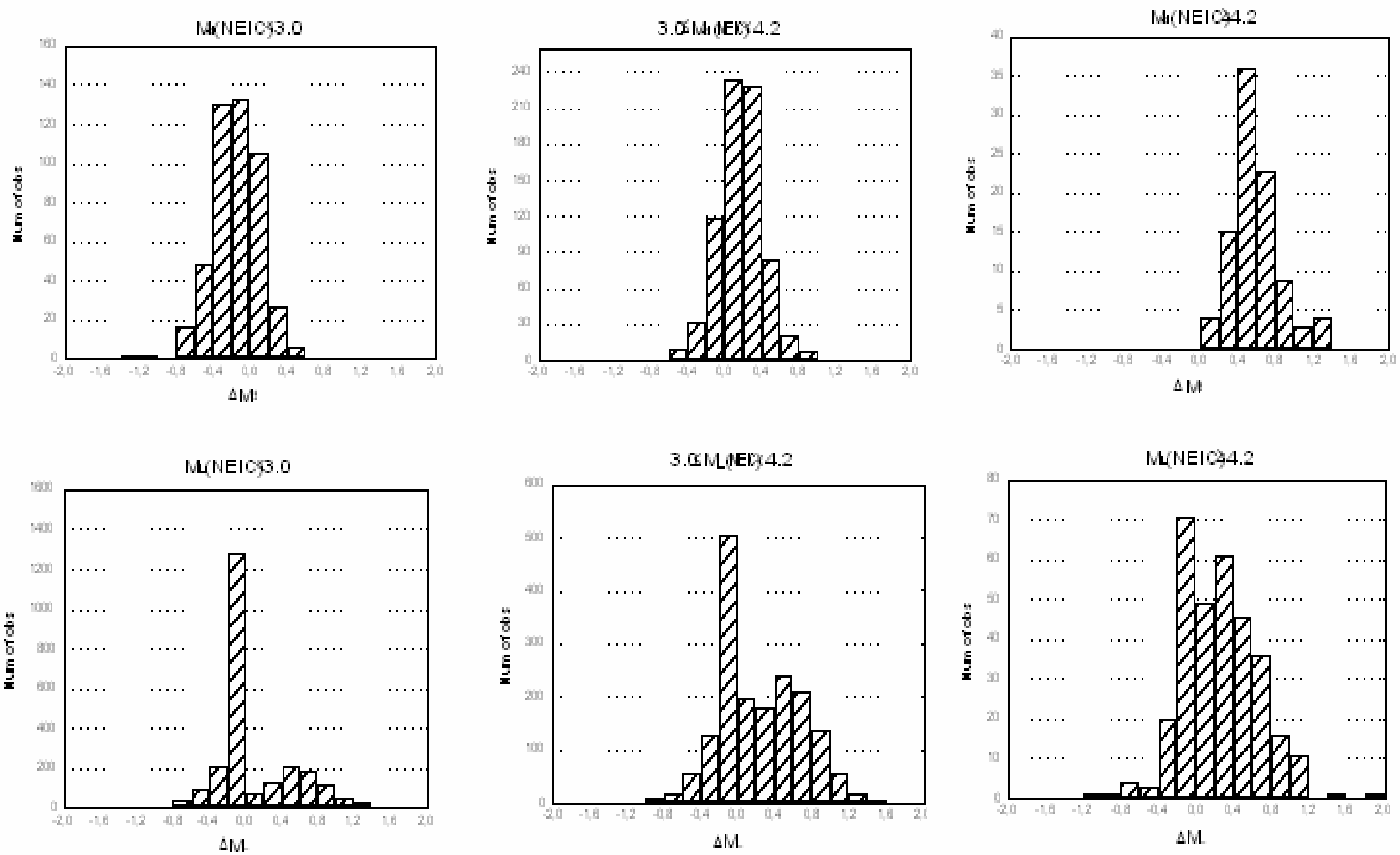


ΔM_L

Magnitude comparison: Italian territory

•DST:

In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.

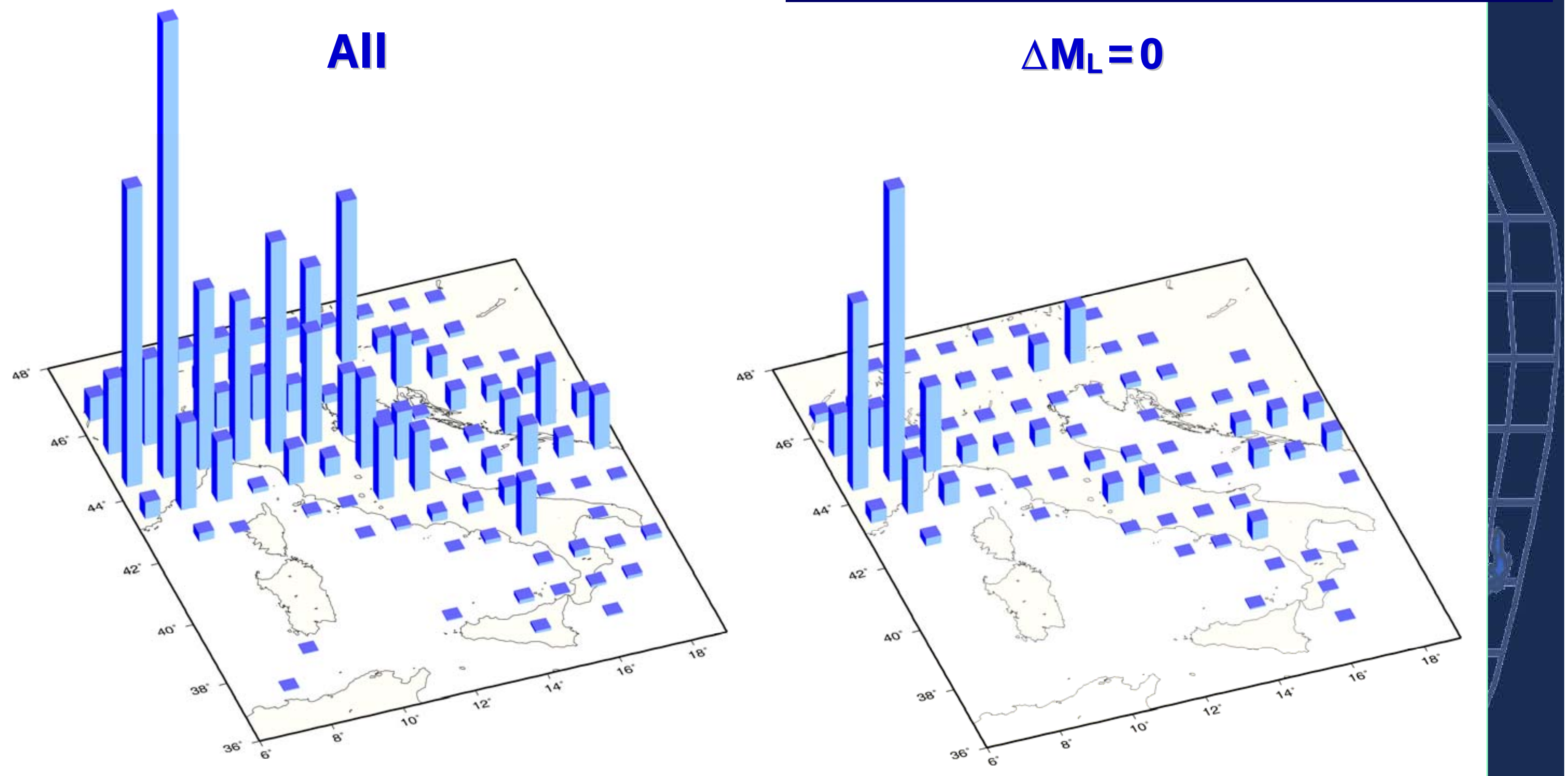


ΔM_d

ΔM_L

Magnitude comparison: Italian territory

•DST:
In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.



Magnitude comparison: Italian territory

•DST:
In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.

Duration Magnitude NEIC - ING

M_D(NEIC)-M_D(ING)
$M_D \geq 3.0$

(1983-1986)
$\Delta M_D = 0.30 \pm 0.02$

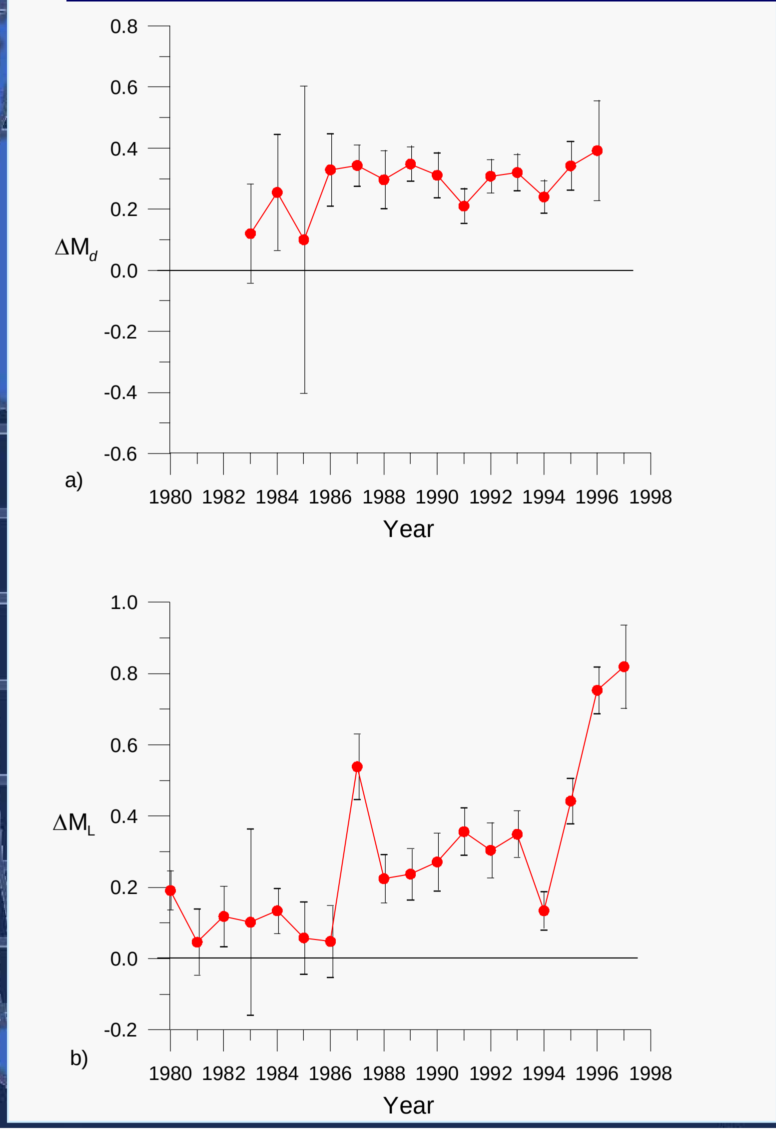
Yearly Average

Local Magnitude NEIC - ING

M_L(NEIC)-M_L(ING)
M_L (NEIC)≥ 3.0

(1980-1986)
$\Delta M_L = 0.08 \pm 0.05$
(1988-1993)
$\Delta M_L = 0.30 \pm 0.04$
(1995-1997)
$\Delta M_L = 0.77 \pm 0.06$

Yearly Average



Magnitude comparison: Italian territory

•DST:
In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.

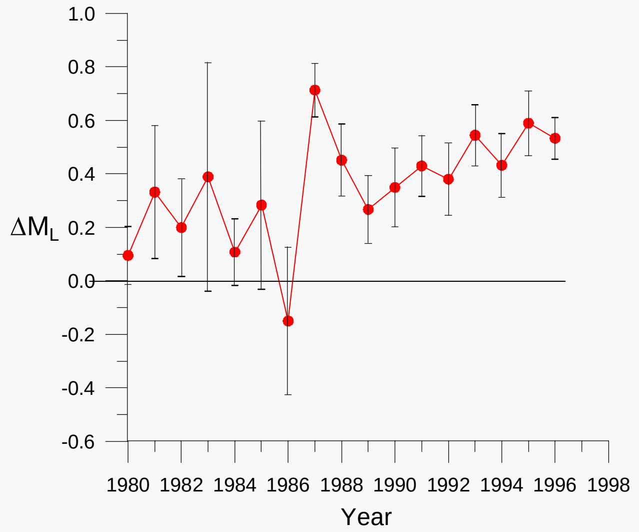
LDG - ING

M_L(LDG)-M_L(ING)

$M_L \geq 3.0$

(1980-1986)
$\Delta M_L = 0.18 \pm 0.08$
(1988-1996)
$\Delta M_L = 0.44 \pm 0.04$

Yearly Average



Local Magnitude: Yearly Average differences

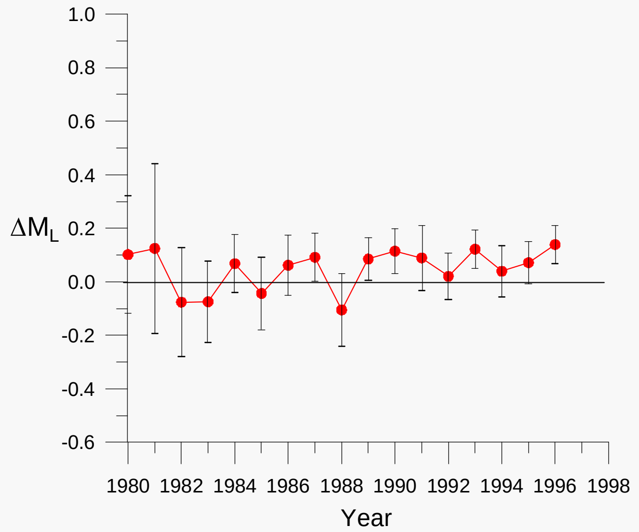
NEIC - LDG

M_L(NEIC)-M_L(LDG)

$M_L \geq 3.0$

(1980-1986)
$\Delta M_L = 0.03 \pm 0.06$
(1988-1996)
$\Delta M_L = 0.08 \pm 0.03$

Yearly Average



CN algorithm and long lasting changes in reported magnitudes

- **The analysis of CN functions in Central Italy allowed us to detect a relevant long lasting change in the reported magnitudes.**
- **The comparison of individual magnitudes, reported by ING and NEIC, indicates, since 1987, an average underestimation of about 0.5 in the Local Magnitude provided by ING.**

•DST:

The analysis of CN functions in Central Italy allowed us to detect a relevant long lasting change in the reported magnitudes.

The comparison of individual magnitudes, and , reported by ING and NEIC catalogues indicates, since 1987, an average underestimation of about 0.5 in the provided by ING.

The hypothesis of general local magnitude underestimation in the Italian ING bulletins is substantiated by the cross-comparison performed between ING, LDG and NEIC catalogues.

The presence of the evidenced magnitude change prevents the use of ING bulletins for CN algorithm application and makes necessary to use global data like NEIC.

CN algorithm and long lasting changes in reported magnitudes

•**DST:**

The analysis of CN functions in Central Italy allowed us to detect a relevant long lasting change in the reported magnitudes.

The comparison of individual magnitudes, and , reported by ING and NEIC catalogues indicates, since 1987, an average underestimation of about 0.5 in the provided by ING.

The hypothesis of general local magnitude underestimation in the Italian ING bulletins is substantiated by the cross-comparison performed between ING, LDG and NEIC catalogues.

The presence of the evidenced magnitude change prevents the use of ING bulletins for CN algorithm application and makes necessary to use global data like NEIC.

- **The presence of a general local magnitude underestimation in the Italian ING bulletins is substantiated by the cross-comparison performed between ING, LDG and NEIC catalogues.**

(Peresan, Panza & Costa, GJI 2000)

(Gasperini, Vannucci & Orlanducci: "Rivalutazione della magnitudo per i terremoti italiani nel periodo post 1980". In: "Catalogo strumentale dei terremoti Italiani dal 1981 al 1996" – 2001)

Compilation of a homogeneous updated catalogue for CN monitoring in Italy

•**DST:**

The analysis of CN functions in Central Italy allowed us to detect a relevant long lasting change in the reported magnitudes.

The comparison of individual magnitudes, and , reported by ING and NEIC catalogues indicates, since 1987, an average underestimation of about 0.5 in the provided by ING.

The hypothesis of general local magnitude underestimation in the Italian ING bulletins is substantiated by the cross-comparison performed between ING, LDG and NEIC catalogues.

The presence of the evidenced magnitude change prevents the use of ING bulletins for CN algorithm application and makes necessary to use global data like NEIC.

Databases available to us:

CCI1996: PFG revised+ING bulletins (Italian catalogue, available up to July 1997)

Priority: M_L, M_d, M_I

NEIC: PDE Preliminary Determinations of Epicenters from NEIC (global catalogue).

Priority: to be defined (available M: m_b, M_S, M_1, M_2)

ALPOR: Catalogo delle Alpi Orientali (local catalogue for eastern Alps)

Priority: M_L, M_I

Compilation of a homogeneous updated catalogue for CN monitoring in Italy

•**DST:**

The analysis of CN functions in Central Italy allowed us to detect a relevant long lasting change in the reported magnitudes.

The comparison of individual magnitudes, and , reported by ING and NEIC catalogues indicates, since 1987, an average underestimation of about 0.5 in the provided by ING.

The hypothesis of general local magnitude underestimation in the Italian ING bulletins is substantiated by the cross-comparison performed between ING, LDG and NEIC catalogues.

The presence of the evidenced magnitude change prevents the use of ING bulletins for CN algorithm application and makes necessary to use global data like NEIC.

Procedure:

•Study of the completeness of PDE catalogue;

•Study of the relations between different kind of magnitudes reported in the CCI1996 and PDE catalogues;

•Formulation of a rule for the choice of magnitude priority in PDE, similar to the priority used for CCI1996;

•Construction of the Updated catalogue, integrating CCI1996, ALPOR and NEIC data (compatibly with the completeness of NEIC).

Operating magnitude selection

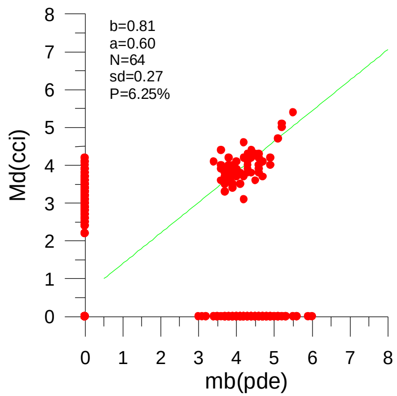
•DST:

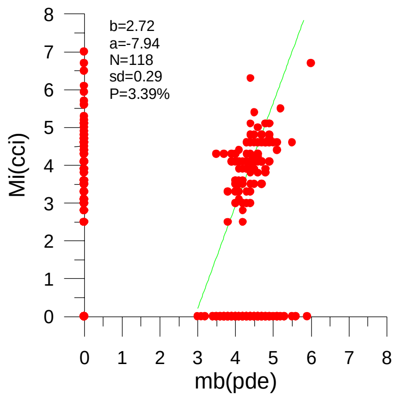
The analysis of CN functions in Central Italy allowed us to detect a relevant long lasting change in the reported magnitudes.

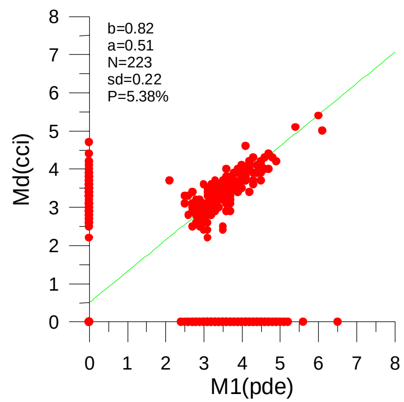
The comparison of individual magnitudes, and , reported by ING and NEIC catalogues indicates, since 1987, an average underestimation of about 0.5 in the provided by ING.

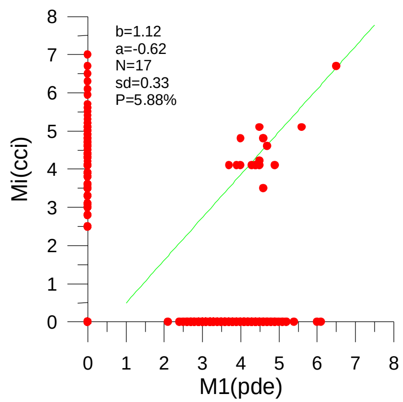
The hypothesis of general local magnitude underestimation in the Italian ING bulletins is substantiated by the cross-comparison performed between ING, LDG and NEIC catalogues.

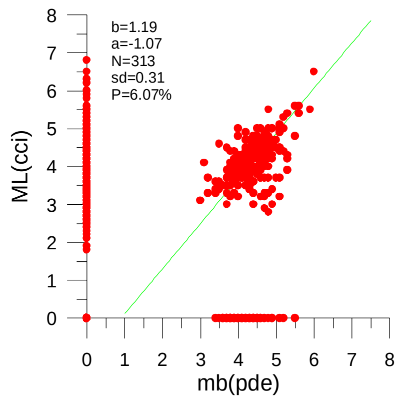
The presence of the evidenced magnitude change prevents the use of ING bulletins for CN algorithm application and makes necessary to use global data like NEIC.

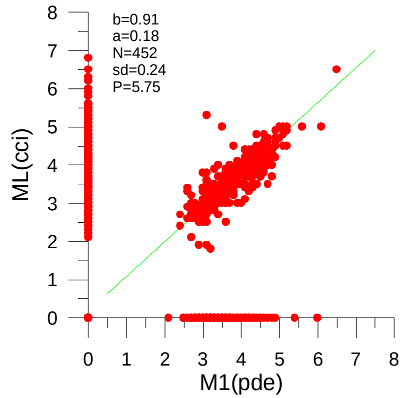












M(CCI1996) – mb(PDE)

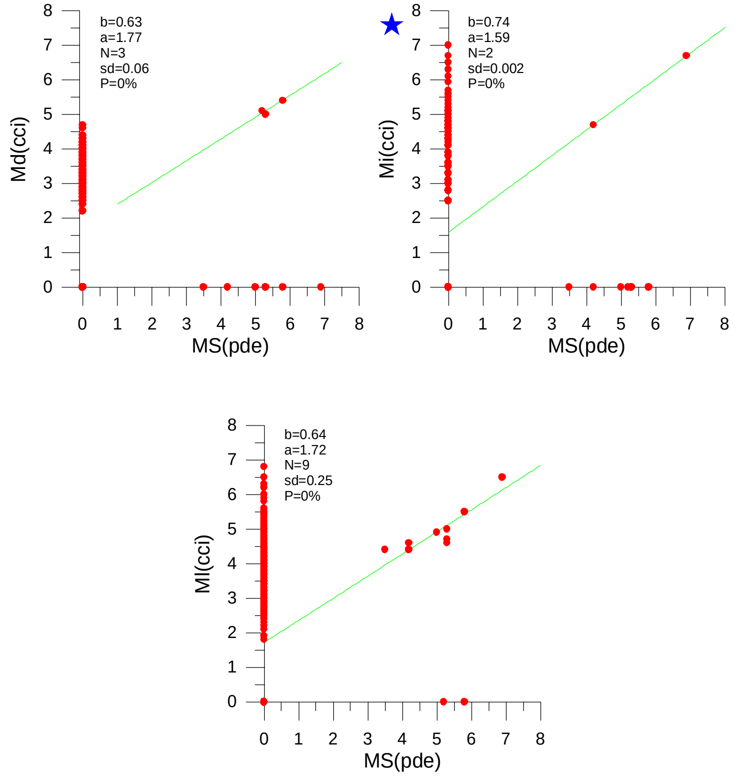
M(CCI1996) – M1(PDE)

Operating magnitude selection

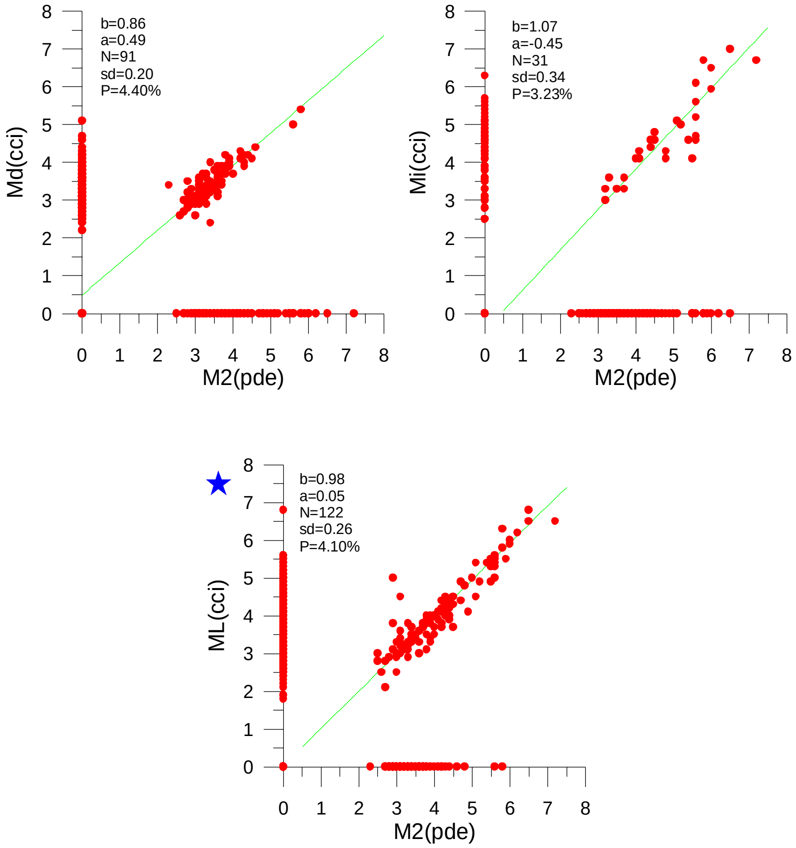
•DST:

The analysis of CN functions in Central Italy allowed us to detect a relevant long lasting change in the reported magnitudes.

The comparison of individual magnitudes, and , reported by ING and NEIC catalogues indicates, since 1987, an average underestimation of about 0.5 in the provided by ING.
The hypothesis of general local magnitude underestimation in the Italian ING bulletins is substantiated by the cross-comparison performed between ING, LDG and NEIC catalogues.
The presence of the evidenced magnitude change prevents the use of ING bulletins for CN algorithm application and makes necessary to use global data like NEIC.



M(CCI1996) – MS(PDE)



M(CCI1996) – M2(PDE)

Operating magnitude selection

•DST:

The analysis of CN functions in Central Italy allowed us to detect a relevant long lasting change in the reported magnitudes.

The comparison of individual magnitudes, and , reported by ING and NEIC catalogues indicates, since 1987, an average underestimation of about 0.5 in the provided by ING.

The hypothesis of general local magnitude underestimation in the Italian ING bulletins is substantiated by the cross-comparison performed between ING, LDG and NEIC catalogues.

The presence of the evidenced magnitude change prevents the use of ING bulletins for CN algorithm application and makes necessary to use global data like NEIC.

$M1(PDE) \longrightarrow M_d(CCI)$

$M2(PDE) \longrightarrow M_L(CCI)$

$M_S(PDE) \longrightarrow$ Poor statistic

$m_b(PDE) \longrightarrow$ Not representative of any of CCI magnitudes

$$M_{CCI}(M_L, M_d, M_I) \Longrightarrow M_{PDE}(M2, M1, M_S)$$

Operating magnitude selection

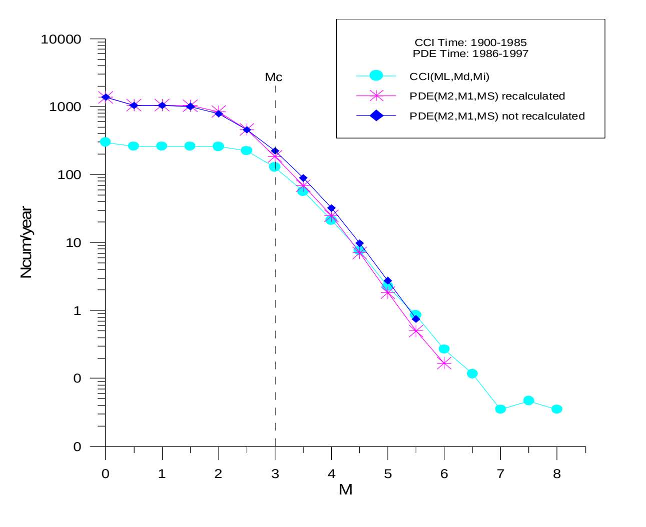
•DST:

The analysis of CN functions in Central Italy allowed us to detect a relevant long lasting change in the reported magnitudes.

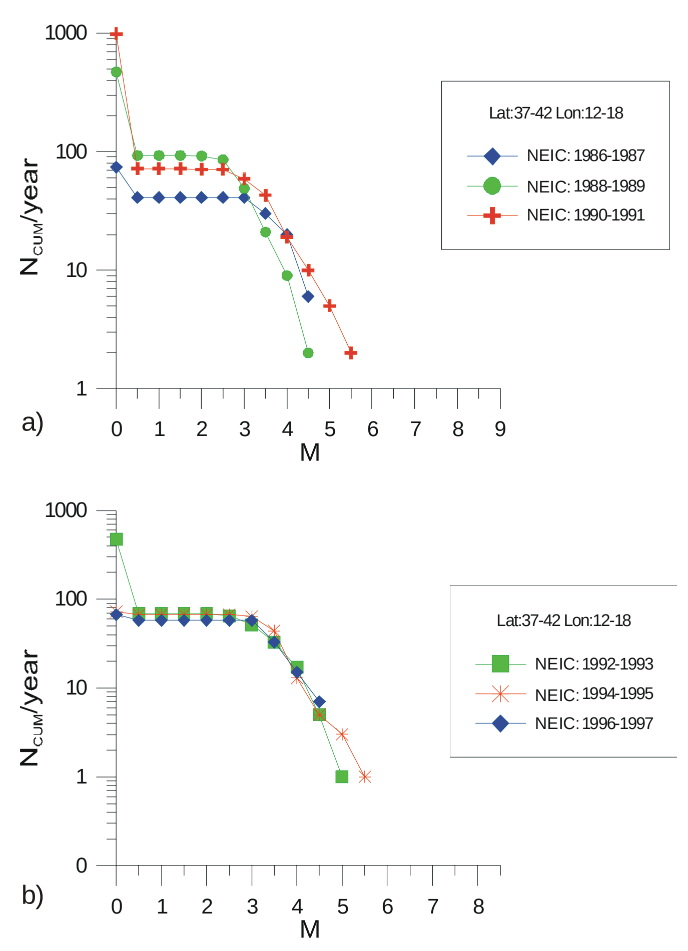
The comparison of individual magnitudes, and , reported by ING and NEIC catalogues indicates, since 1987, an average underestimation of about 0.5 in the provided by ING.

The hypothesis of general local magnitude underestimation in the Italian ING bulletins is substantiated by the cross-comparison performed between ING, LDG and NEIC catalogues.

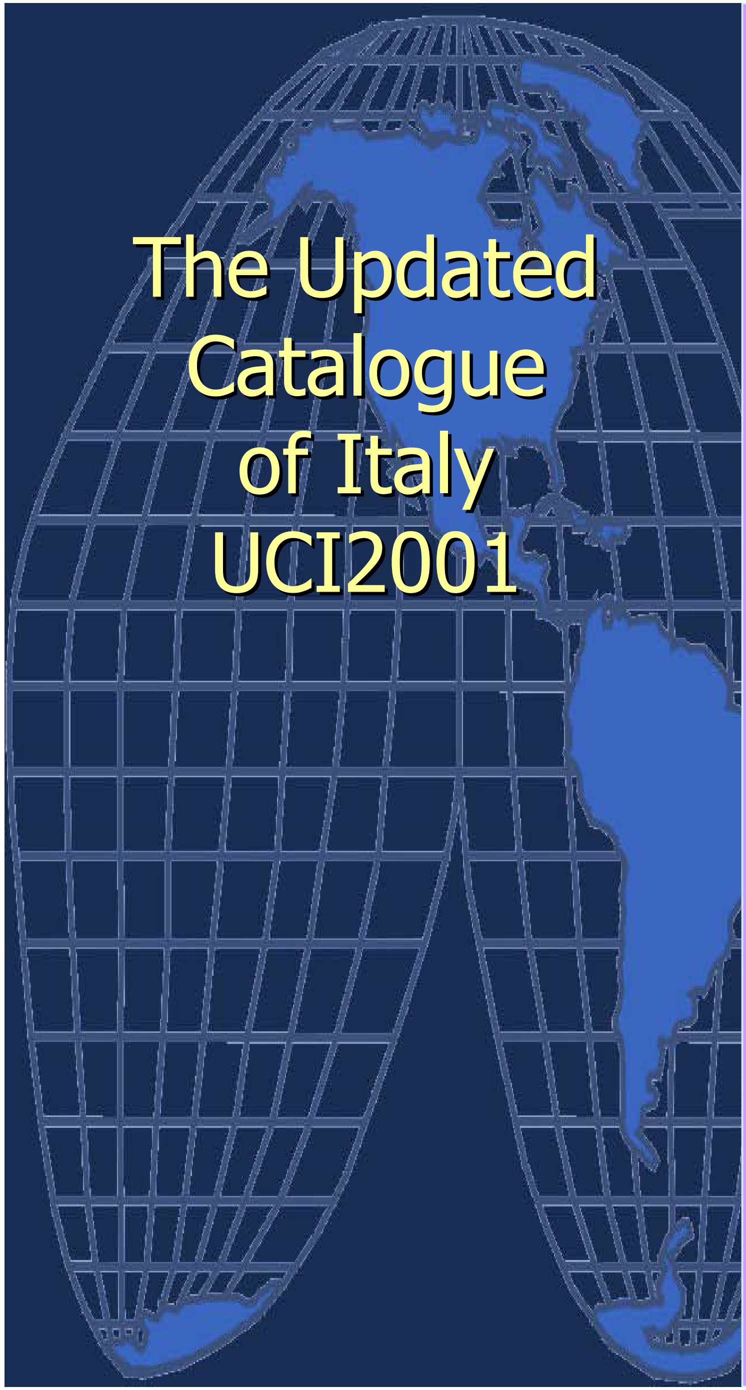
The presence of the evidenced magnitude change prevents the use of ING bulletins for CN algorithm application and makes necessary to use global data like NEIC.

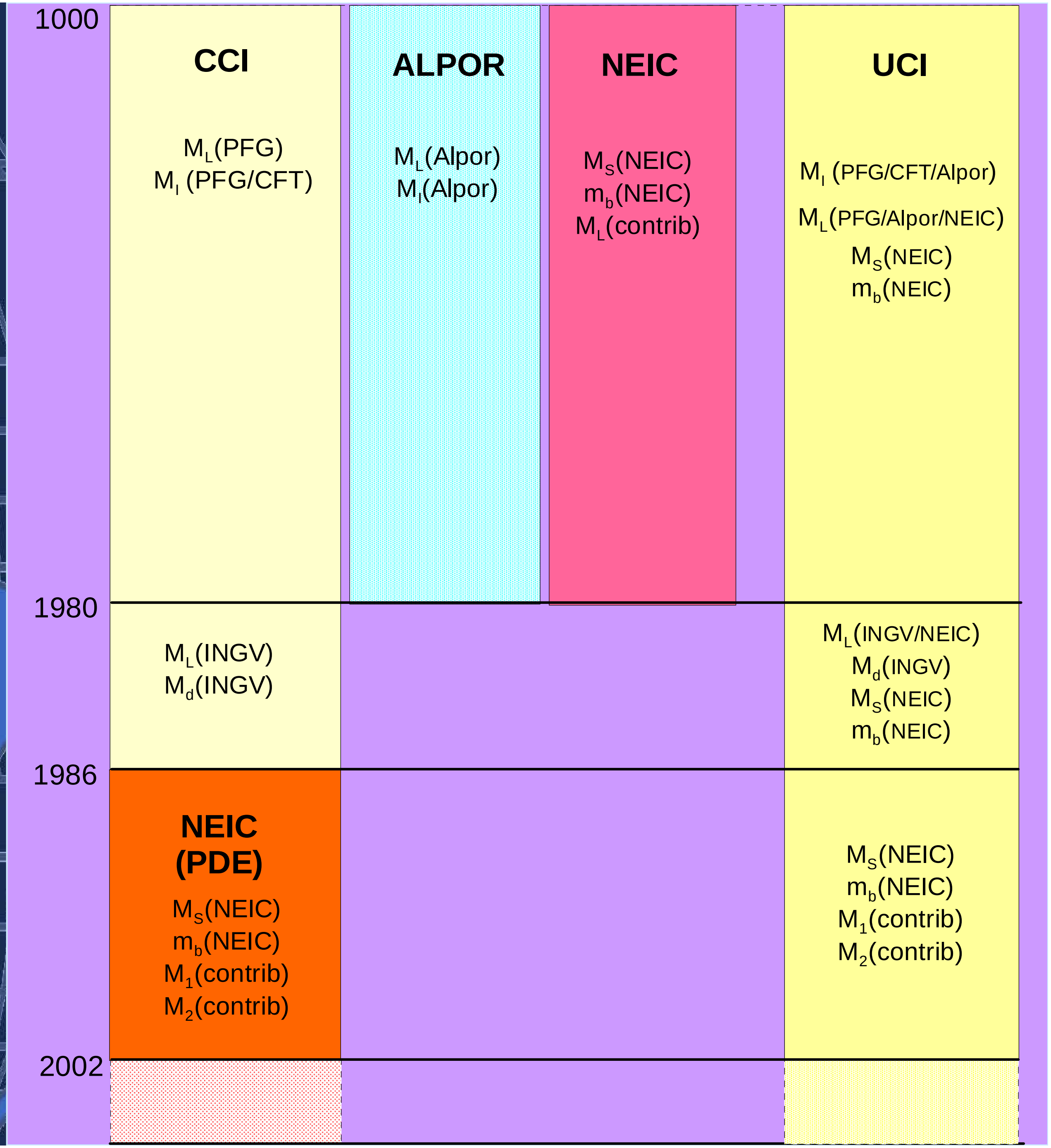


Central Italy

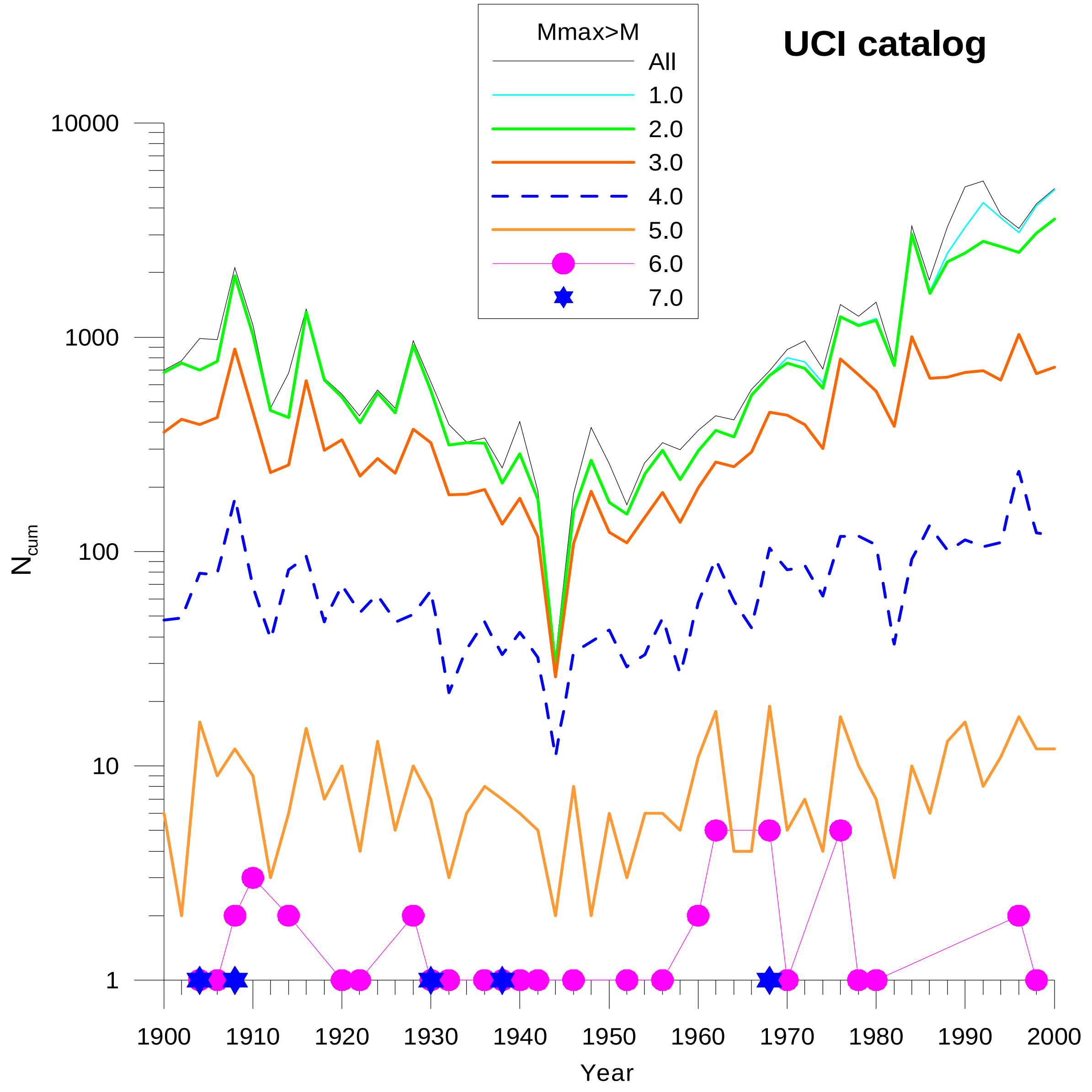


Southern Italy


The Updated
Catalogue
of Italy
UCI2001


1000
1980
1986
2002
CCI
M_L(PFG)
M_I (PFG/CFT)
ALPOR
M_L(Alpor)
M_I(Alpor)
NEIC
M_S(NEIC)
m_b(NEIC)
M_L(contrib)
UCI
M_I (PFG/CFT/Alpor)
M_L(PFG/Alpor/NEIC)
M_S(NEIC)
m_b(NEIC)
M_L(INGV)
M_d(INGV)
M_L(INGV/NEIC)
M_d(INGV)
M_S(NEIC)
m_b(NEIC)
NEIC
(PDE)
M_S(NEIC)
m_b(NEIC)
M_1(contrib)
M_2(contrib)
M_S(NEIC)
m_b(NEIC)
M_1(contrib)
M_2(contrib)

The Updated Catalogue of Italy



Stability of intermediate-term earthquake predictions with respect to random errors in magnitude: the case of Central Italy

CN is essentially based on the information contained in the **earthquake catalogues**.

It analyses the seismic flow using: origin time, epicentral coordinates and **magnitudes** of earthquakes.

All catalogues are inevitably affected by errors.

Magnitudes are characterised by the most significant errors, conditioning both aftershocks removal and seismic flow.

Systematic Errors | **Random Errors**

How random errors on reported magnitudes affect CN prediction results?

Randomised Magnitude

$$M_R = M_C + \Delta M_I + \Delta M_D$$

M_c : operating magnitude

ΔM_I : measurement error ΔM_D : error of discretisation

Measurement random errors

$P(\Delta M_I) = F^{TR}(\Delta M_I)$

Truncated normal probability distribution with: $\sigma = M_{\max}/3$

$F^{TR}(\Delta M_I) = F(\Delta M_I) / [2F(\Delta M_{max})-1]$

M_{max} = maximum assumed error on magnitudes

$$|\Delta M_I| \leq \Delta M_{max}$$

Errors of magnitude discretisation

$P(\Delta M_D)$ = Uniform probability distribution

Interval: $[-d/2; d/2]$

d=discretization step=0.1

•**DST:**

The analysis of CN functions in Central Italy allowed us to detect a relevant long lasting change in the reported magnitudes.

The comparison of individual magnitudes, and , reported by ING and NEIC catalogues indicates, since 1987, an average underestimation of about 0.5 in the provided by ING.

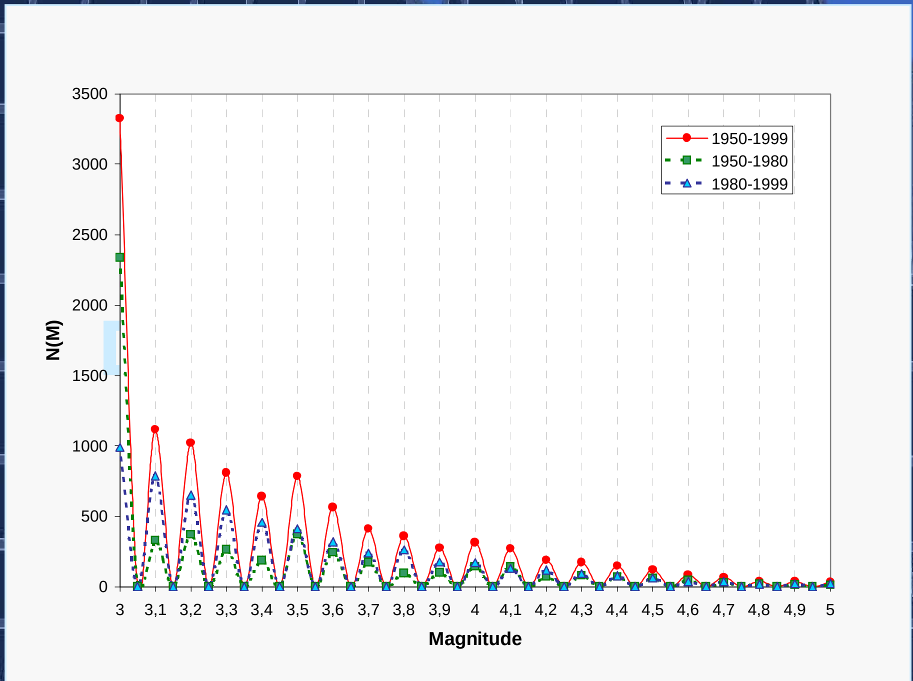
The hypothesis of general local magnitude underestimation in the Italian ING bulletins is substantiated by the cross-comparison performed between ING, LDG and NEIC catalogues.

The presence of the evidenced magnitude change prevents the use of ING bulletins for CN algorithm application and makes necessary to use global data like NEIC.

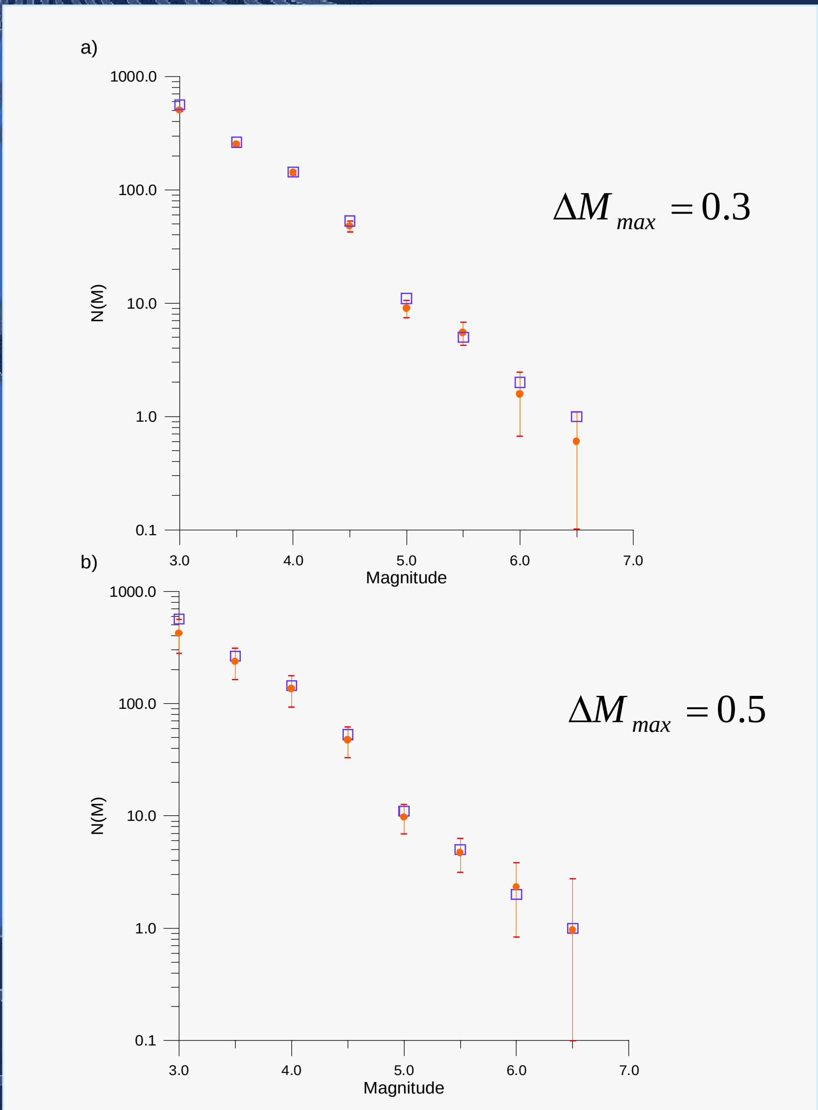
Magnitude randomisation

•DST:

In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.



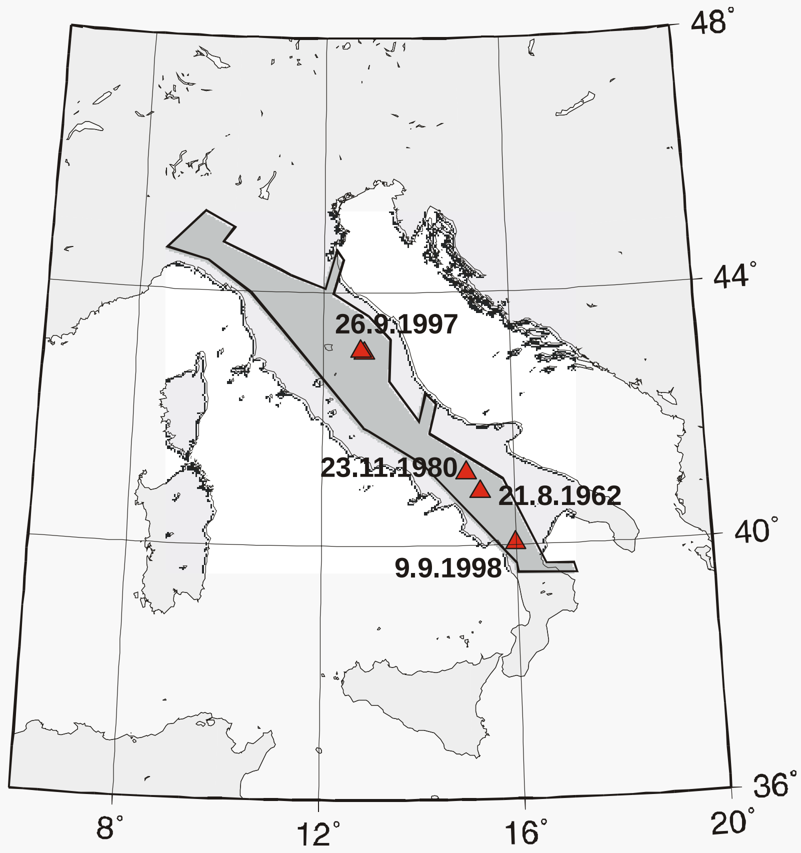
Discretisation step: *d*=0.1



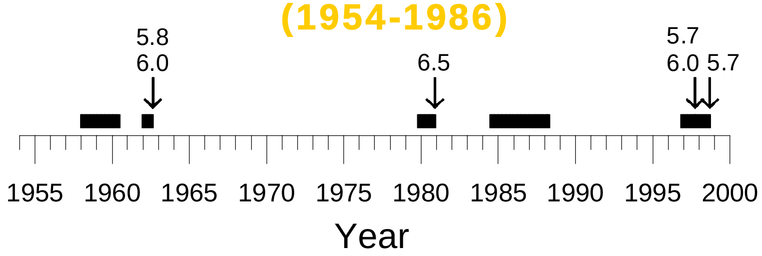
Distribution of the number of events versus magnitude for the original and the randomised catalogues

Results obtained with the original catalogue

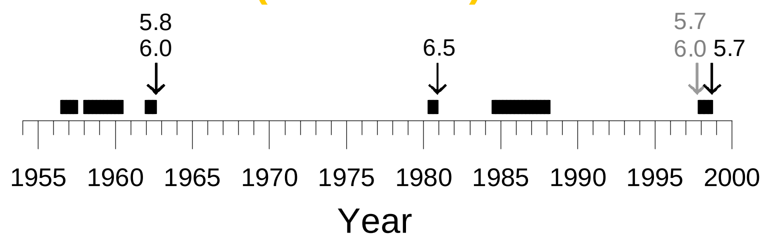
•DST:
In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.



Short Learning
(1954-1986)



Long Learning
(1954-1998)



TIPs obtained with the original catalogue with a different length of the "thresholds setting period" (Learning)

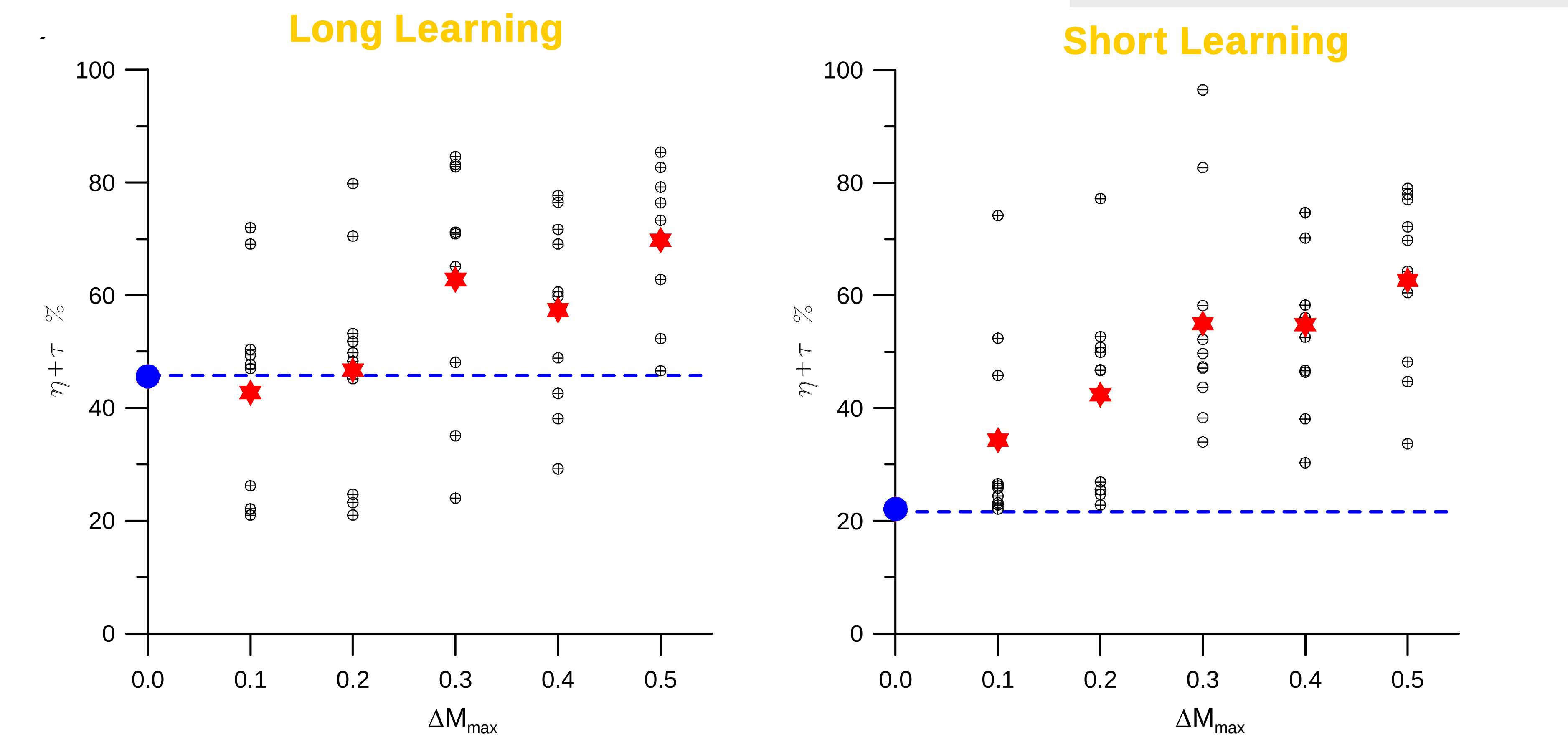
Prediction of the events with M≥5.6

Results obtained with the randomised catalogue

•DST:

In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.

- Original catalogue
- All randomised catalogues
- Average of randomised

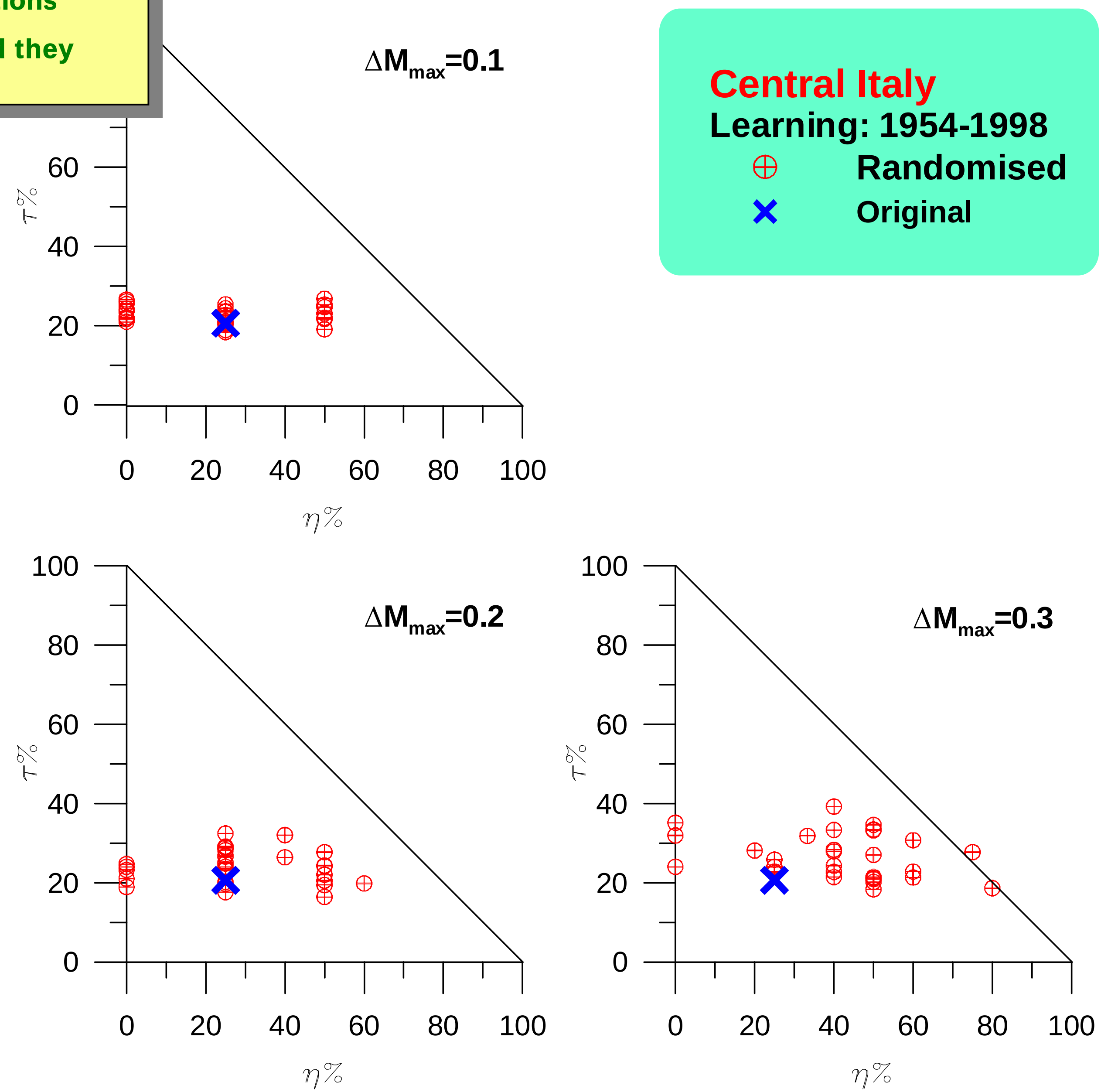


- ***N*** **is the number of strong earthquakes occurred during the time period *T* covered by predictions**
- **The alarms cover altogether the time *t* and they have missed *n* strong events**

Results obtained with the randomised catalogue

$\eta = n/N$: the rate of failures-to-predict

$\tau = t/T$: the rate of time of alarms



Results obtained with the randomised catalogue:
variation of target events

•**DST:**

In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.

Event date	*M*	N_S, %	N_P, %
23.11.1980	6.5	100.0	66.7
21.8.1962	6.0	100.0	93.3
26.9.1997	6.0	100.0	40.0
9.9.1998	5.7	83.3	56.0
19.9.1979	5.5	23.3	0.0
5.5.1990	5.5	12.0	0.0
7.5.1984	5.4	0.0	0.0

N_S : percentage of times each event is a strong event ($M \geq M_0 = 5.6$)

N_P : percentage of times each event is predicted (catalogues randomised with $\Delta M_{max} = 0.3$)

List of earthquakes with
$M \geq 5.3$
Central Italy: 1950-1999

Results obtained with the randomised catalogue

•DST:
In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.

Average differences in prediction errors

ΔM_{max}	$<\Delta\tau>$, %	$<\Delta\eta>$, %	$<\Delta(\eta+\tau)>$, %	$<\Delta N_{SE}>$, %
0.1	2.2	-2.5	-0.3	0.0
0.2	3.5	3.8	7.3	2.5
0.3	5.9	15.6	21.5	13.4

OC
Original Catalogue

RCs
Randomised Catalogues

$<\Delta\tau>$, $<\Delta\eta>$
average variation of prediction errors for the RCs

$$<\Delta N_{SE}> = < |N_R - N_0| / N_0 >$$
average variation of the number of strong events

N_0 - number of events to be predicted in the OC
N_R - number of events to be predicted in the RC

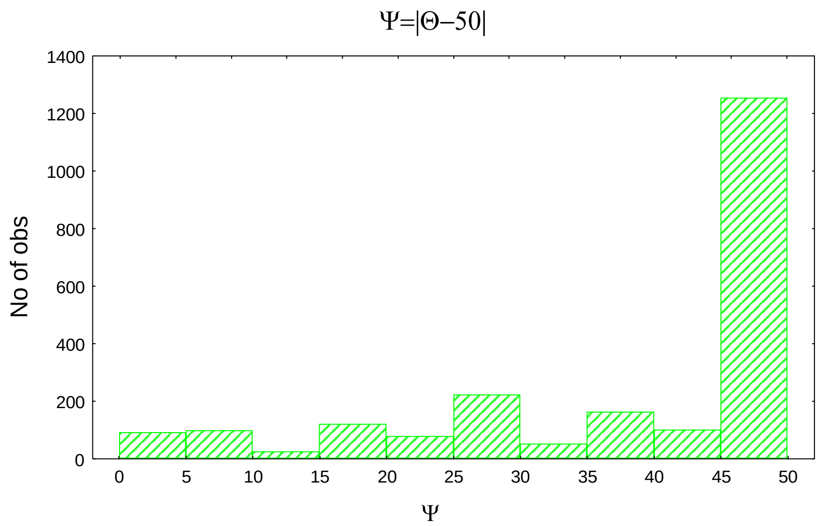
Stability of TIPs diagnosis

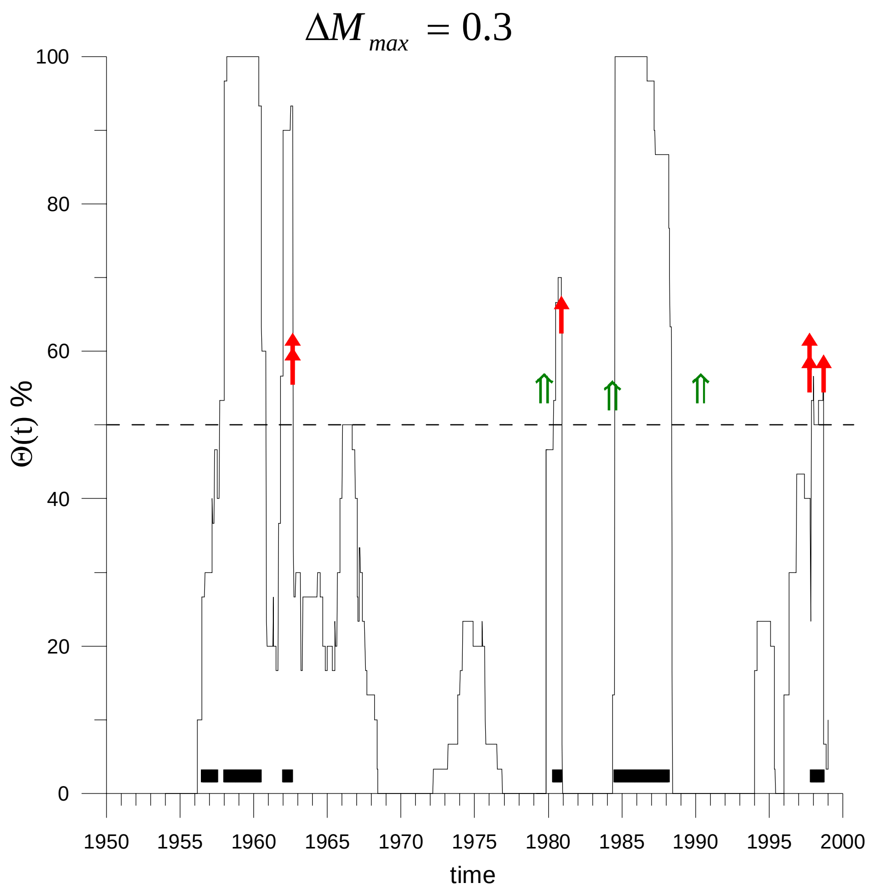
Central Italy
(Long threshold setting period)

•DST:
In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: a) time difference 1 minute; b) epicentral distance: 1° for the comparison with the global catalogue (Storchak, Bird and Adams, 1998). No limitation is imposed to magnitude or depth differences.

Ψ: **percentage of tests for which the recognition of the time *t* does not change with respect to its average value**

0≤Ψ≤50





Θ(t)**: percentage of tests for which the time *t* belongs to a TIP**

Stability of CN predictions with respect to random errors in magnitude

•DST:

The analysis of CN functions in Central Italy allowed us to detect a relevant long lasting change in the reported magnitudes.

The comparison of individual magnitudes, and , reported by ING and NEIC catalogues indicates, since 1987, an average underestimation of about 0.5 in the provided by ING.

The hypothesis of general local magnitude underestimation in the Italian ING bulletins is substantiated by the cross-comparison performed between ING, LDG and NEIC catalogues.

The presence of the evidenced magnitude change prevents the use of ING bulletins for CN algorithm application and makes necessary to use global data like NEIC.

- **The results of prediction remain stable for $\Delta M_{max} < 0.3$.**
- **To guarantee the stability of the results, the thresholds setting period must be long enough to include a significant sample of dangerous and non dangerous intervals of time.**
- **The quality of predictions is mainly controlled by the percentage of failures to predict, which depends on the changes in the number of strong earthquakes.**
- **The identification of TIPs is very stable during most of the time and the randomisation does not introduce spurious alarming patterns associated with the occasionally strong events.**

Intermediate-term middle-range earthquake prediction experiment in Italy

CN algorithm (*Keilis-Borok et al., 1990; Peresan et al., 2004*)
M8S algorithm (*Kossobokov et al, 2002)*

Main features:

- Fully formalized algorithms and computer codes available for independent testing;
- Use of published & routine catalogues of earthquakes;
- Worldwide tests ongoing for more than 10 years permitted to assess the significance of the issued predictions.

Italy:

- Stability tests with respect to several free parameters of the algorithms *(e.g. Costa et al., 1995; Peresan et al., GJI, 2000; Peresan et al., PEPI, 130, 2002);*
- CN predictions are regularly updated every two months since January 1998.
- M8s predictions are regularly updated every six months since January 2002.

→ **Real time prediction experiment started in July 2003**

Intermediate-term middle-range earthquake prediction experiment in Italy

The experiment, launched starting on July 2003, is aimed at a *real-time test* of M8S and CN predictions in Italy.

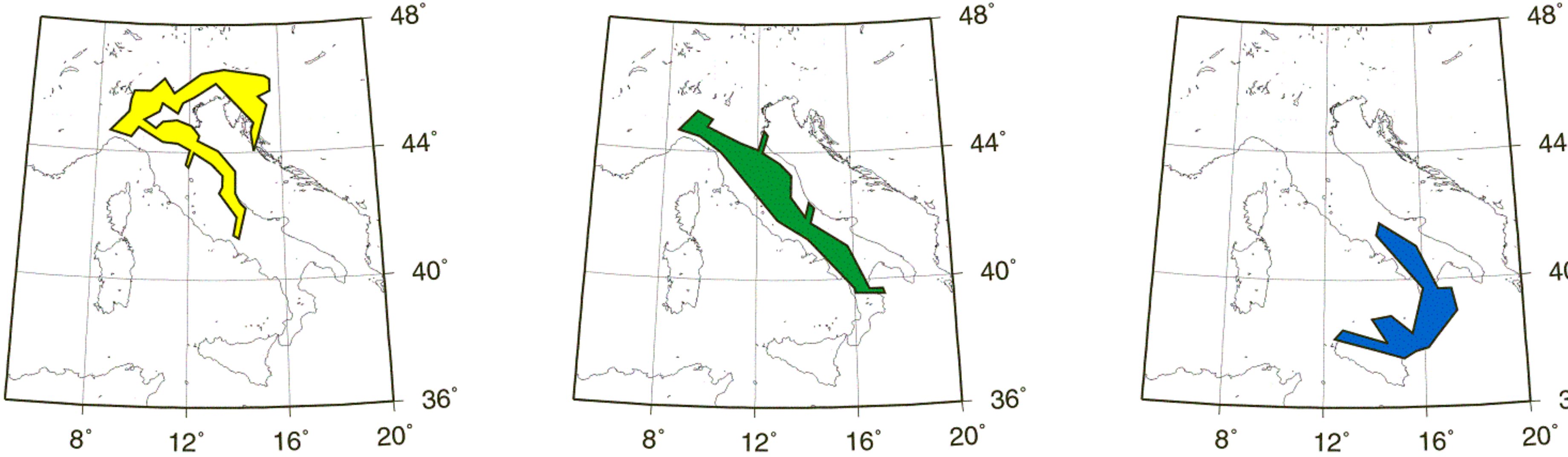
Updated predictions are regularly posted at:

"http://www.ictp.trieste.it/www_users/sand/prediction/prediction.htm"

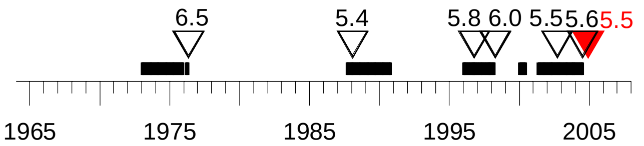
A complete archive of predictions is made accessible to a number of scientists, with the goal to accumulate a collection of correct and wrong predictions, that will permit to validate the considered methodology.

Current predictions are protected by password. Although these predictions are intermediate-term and by no means imply a "red alert", there is a legitimate concern about maintaining necessary confidentiality.

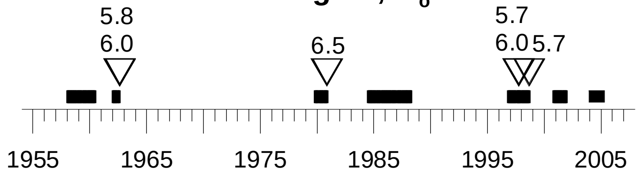
CN algorithm in Italy



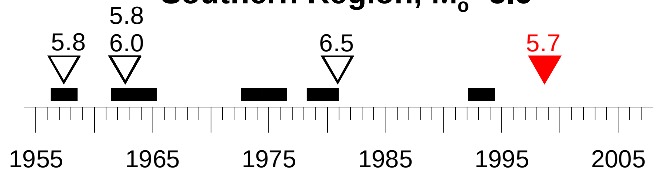
Northern Region, M_o=5.4



Central Region, M_o=5.6



Southern Region, M_o=5.6



Predictions as on: 1-1-2005 (subject to update on: 1-3-2005)

Intermediate-term middle-range earthquake prediction CN

Space-time volume of alarm in CN application in Italy

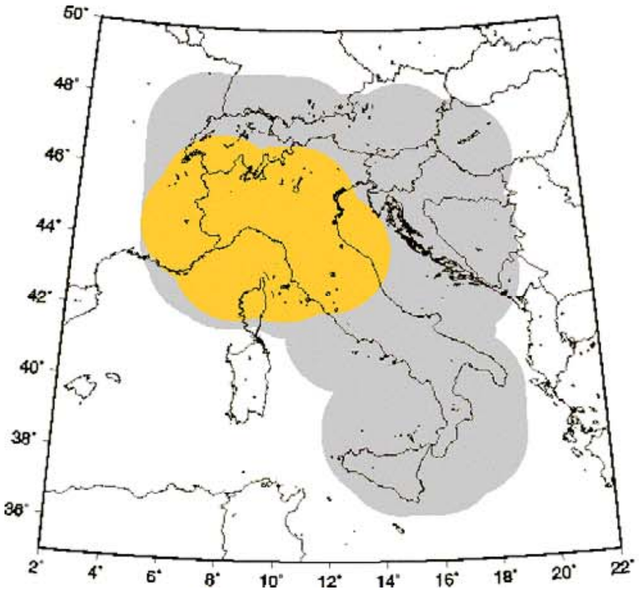
Experiment	Space-time volume of alarm (%)	n/N	Confidence level (%)
Retrospective* (1954 – 1963)	41	3/3	93
Retrospective (1964 – 1997)	27	5/5	>99
Forward (1998 – 2004)	41	4/5	90
All together (1954 – 2004)	32	12/13	>99

* Central and Southern regions only

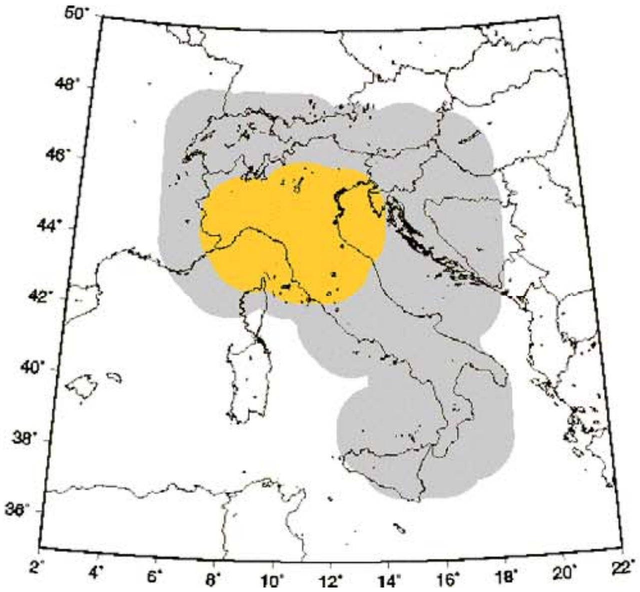
Algorithm CN predicted 12 out of the 13 strong earthquakes occurred in the monitored zones of Italy, with 32% of the considered space-time volume occupied by alarms.

(updated to December 31 2004)

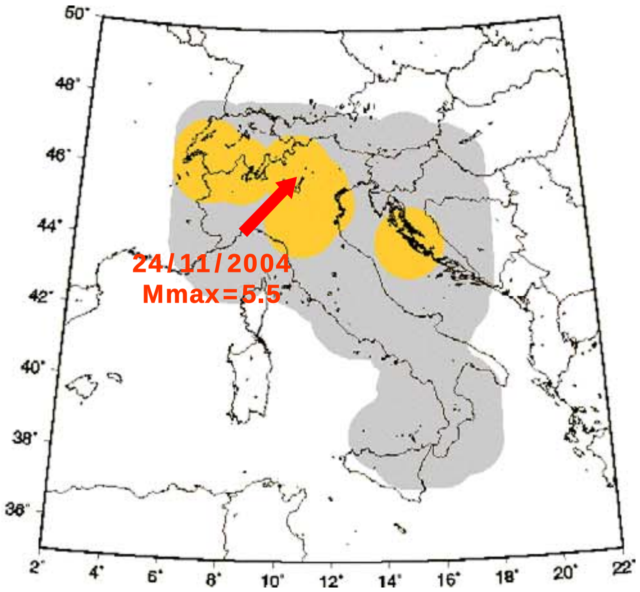
M8S algorithm in Italy



M≥6.5



M≥6.0



M≥5.5

Monitored region

Alerted region

Predictions as on: 1-7-2004 (subject to update on: 1-1-2005)

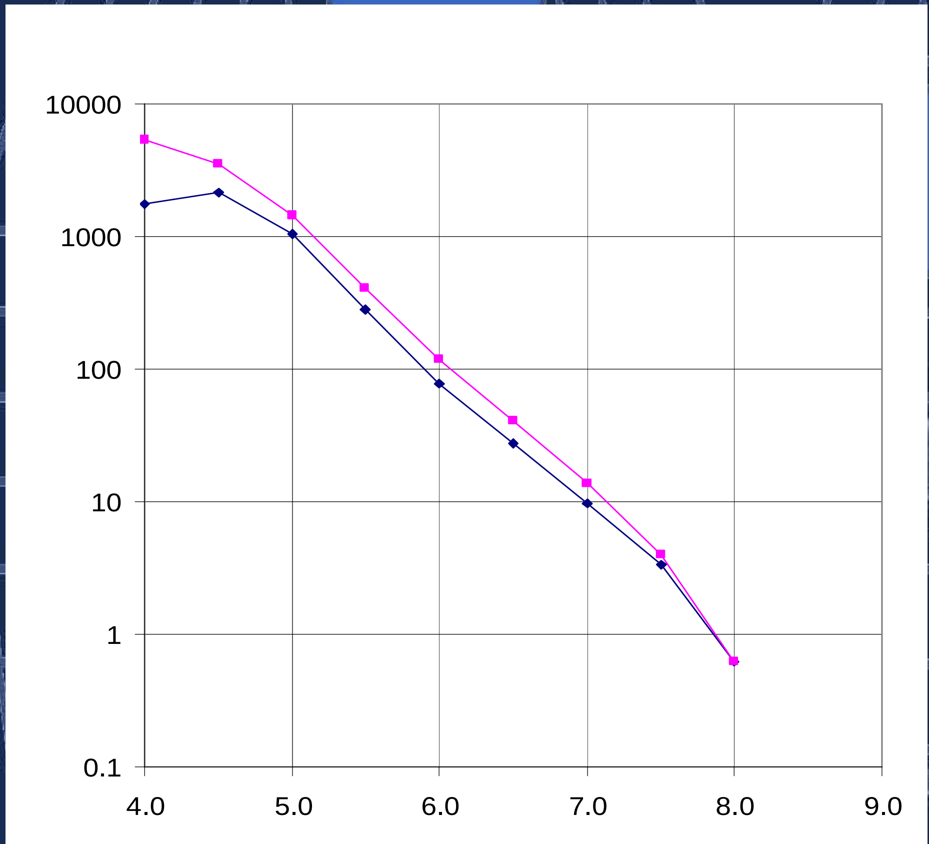
Space-time volume of alarm in M8s application in Italy

Experiment	M6.5+		M6.0+		M5.5+	
	Space-time volume, %	n/N	Space-time volume, %	n/N	Space-time volume, %	n/N
Retrospective (1972-2001)	36	2/2	40	1/2	39	9/14
Forward (2002-2004)	49	0/0	43	0/0	25	4/8
All together (1972-2004)	37	2/2	40	1/2	38	13/22

Algorithm M8S predicted 62% of the events occurred in the monitored zones in Italy, i.e. 16 out of 26 events occurred within the area alerted for the corresponding magnitude range (updated to December 31 2004).

The confidence level of M5.5+ predictions since 1972 has been estimated to be about 97%; no estimation is still possible for M6.0+ and M6.5+.

The Gutenberg-Richter law



- Averaged over a large territory and time the number of earthquakes equal or above certain magnitude, N(M) scales as

$$\log_{10} N(M) = a - bM$$

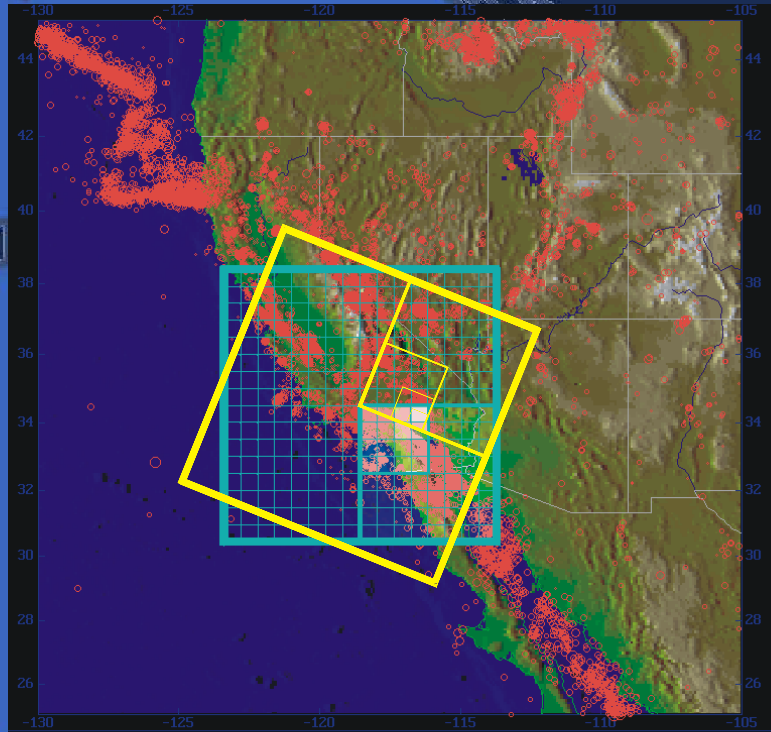
This general law of similarity establishes the scaling of earthquake sizes in a given space time volume but gives no explanation to the question how the number, N, changes when you zoom the analysis to a smaller size part of this volume.

Which are the limits of validity of the power-law scaling expressed by the Gutenberg-Richter law?

- The analysis of global seismicty shows that a single Gutenberg-Richter (GR) law is not universally valid and that a multiscale seismicity model (*Molchan, Kronrod & Panza, BSSA, 1997*) can reconcile two apparently conflicting concepts the Characteristic Earthquake (CE) the Self-Organized Criticality (SOC).
- The multiscale seismicity model, implies that only the set of earthquakes with dimensions that are small with respect to the dimensions of the analysed region can be described adequately by the Gutenberg-Richter law.
- This conditon, fully satisfied in the study of global seismicity made by Gutenberg and Richter, has been violated in many subsequent investigations.

The scheme for box-counting

The counts in sets of cascading squares, “telescopes”, estimate the natural scaling of the spatial distribution of earthquake epicenters and provide evidence for rewriting the Gutenberg-Richter recurrence law the generalised form:



$$\log_{10}N = A + B\cdot(5 - M) + C\cdot\log_{10}L$$

where N = N(M, L) is the expected annual number of earthquakes with magnitude M in an area of linear dimension L.

The first results

The method was tested successfully on artificial catalogs and applied in a dozen of selected seismic regions from the hemispheres of the Earth to a certain intersection of faults.

12 KOSOBOKOV AND MAZHKENOV: SIMILARITY IN SPATIAL DISTRIBUTION OF SEISMICITY

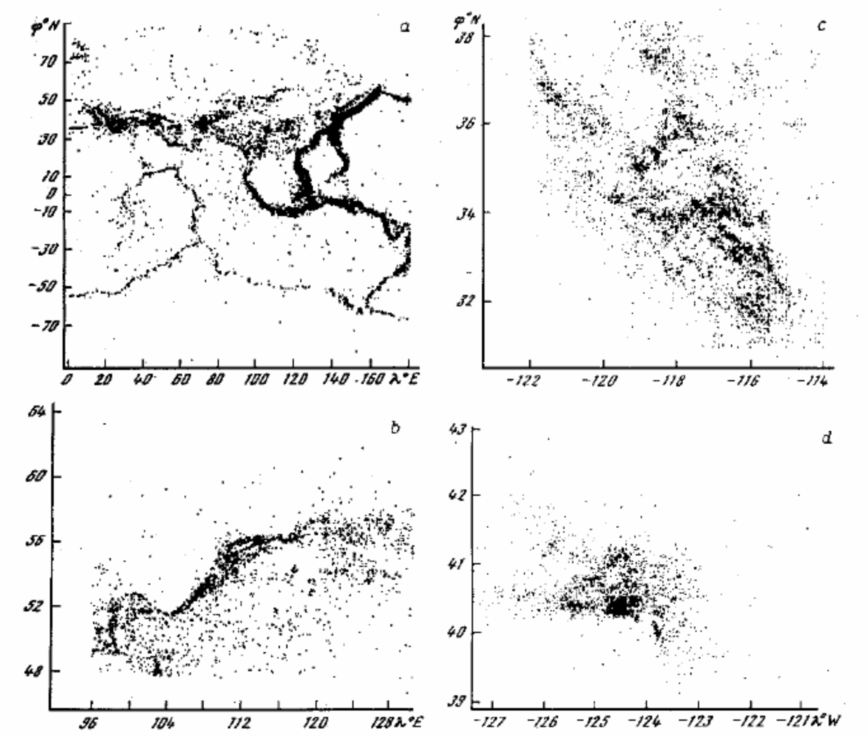


Fig. 2. Examples of spatial distribution of epicenters from catalogs of mainshocks. (a) Eastern Hemisphere. (b) Lake Baikal area. (c) Southern California. (d) The Cape Mendocino vicinity.

14 KOSOBOKOV AND MAZHKENOV: SIMILARITY IN SPATIAL DISTRIBUTION OF SEISMICITY

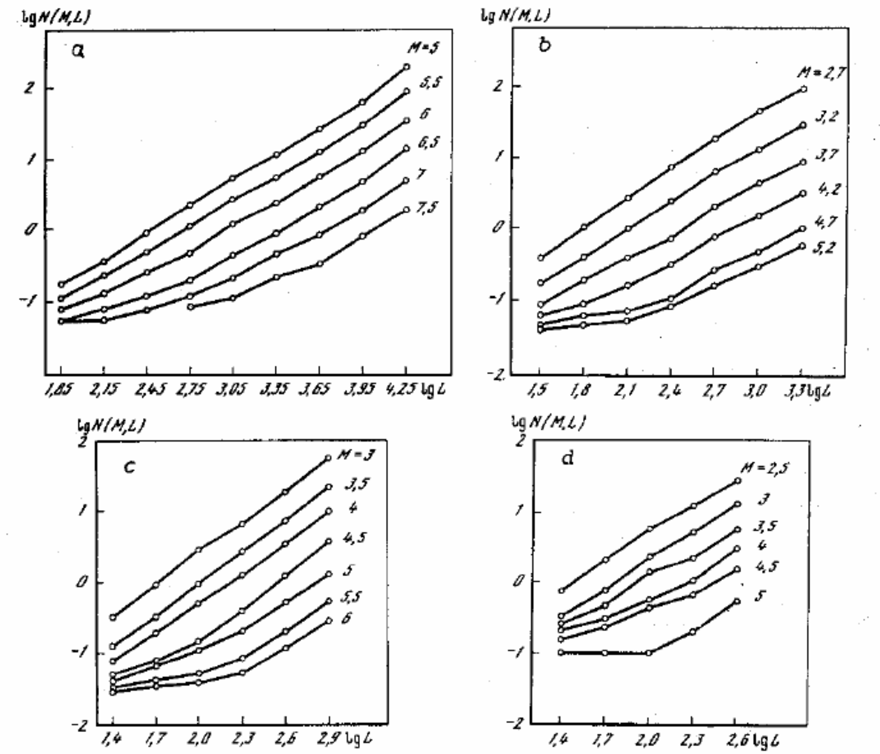
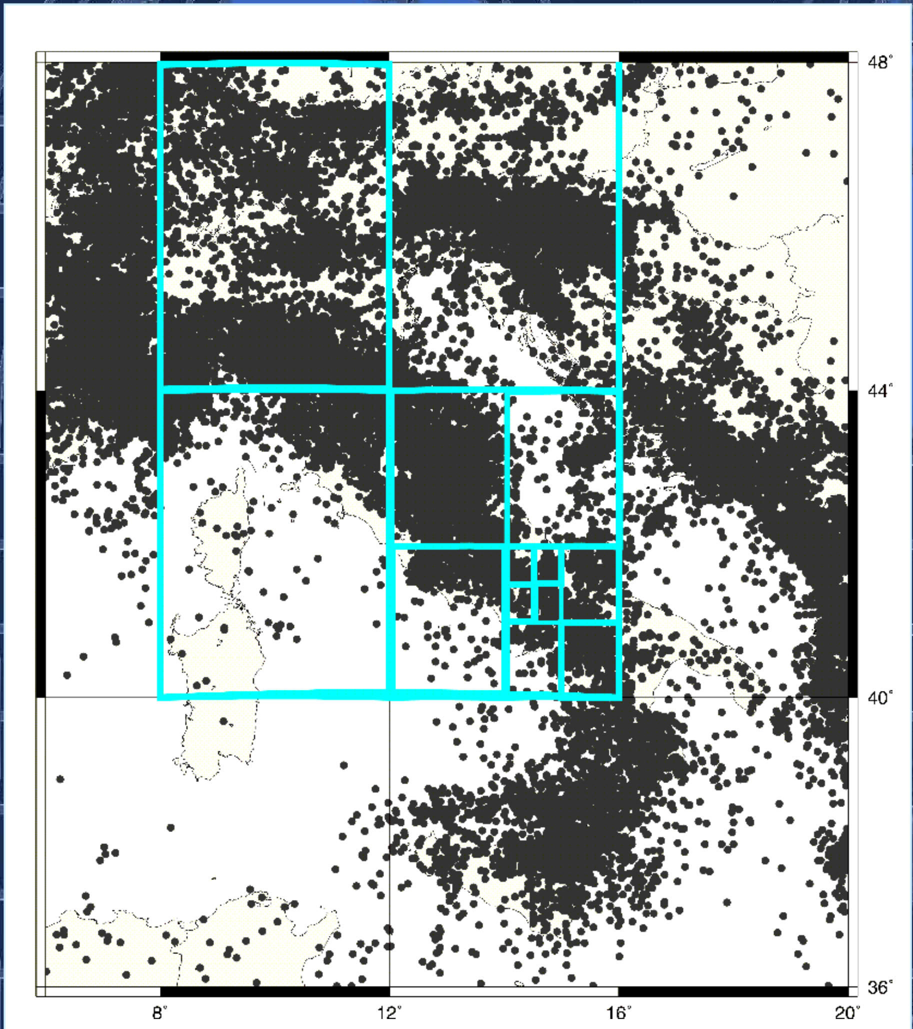


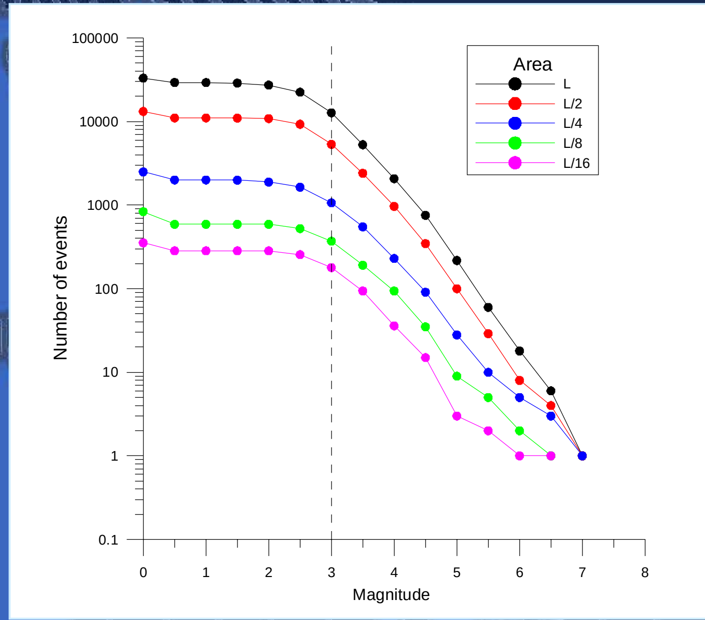
Fig. 3. Examples of log $N(M, L)$ graphs. (a) Eastern Hemisphere. (b) Lake Baikal area. (c) Southern California. (d) The Cape Mendocino vicinity.

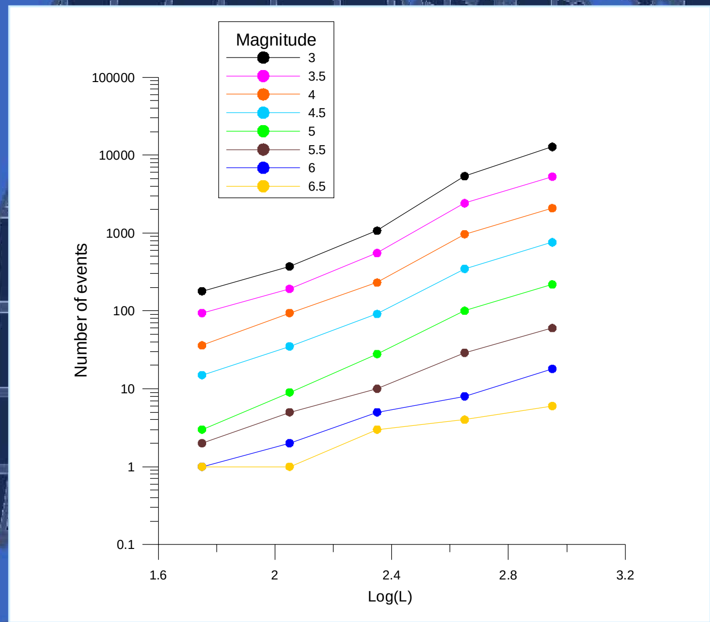
(*Kossobokov and Mazhkenov, 1988*)

Multiscale seismicity model



UCI2001 catalogue
1900-2002





ELSEVIER

Physics of the Earth and Planetary Interiors 130 (2002) 117–127

PHYSICS OF THE EARTH AND PLANETARY INTERIORS

www.elsevier.com/locate/pepi

Stability of intermediate-term earthquake predictions with respect to random errors in magnitude: the case of central Italy

A. Peresan[a,*], I. Rotwain[b], I. Zaliapin[b], G.F. Panza[a,c]

[a] Department of Earth Sciences, University of Trieste, via E. Weiss 1, 34127 Trieste, Italy
[b] International Institute of Earthquake Prediction Theory and Mathematical Geophysics, Russian Academy of Sciences, Warshavskoe sh.79, kor.2, Moscow 113556, Russia
[c] The Abdus Salam International Centre for Theoretical Physics, SAND Group ICTP, 34100 Trieste, Miramare, Italy

Received 22 December 2000; received in revised form 26 November 2001; accepted 3 December 2001

Abstract

The influence of random magnitude errors on the results of intermediate-term earthquake predictions is analyzed in this study. The particular case of predictions performed using the algorithm CN in central Italy is considered. The magnitudes of all events reported in the original catalog (OC) are randomly perturbed within the range of the expected errors, thus generating a set of randomized catalogs. The results of predictions for the original and the randomized catalogs, performed following the standard CN rules, are then compared. The average prediction quality of the algorithm CN appear stable with respect to magnitude errors up to ±0.3 units. Such a stable prediction is assured if the threshold setting period corresponds to a time interval sufficiently long and representative of the seismic activity within the region, while if the threshold setting period is too short, the average quality of CN decreases linearly for increasing maximum error in magnitude.

Keywords: Earthquake prediction; Algorithm CN; Magnitude error; Randomization; Italy

1. Introduction

A generation of intermediate-term earthquake prediction algorithms was developed and exhaustively tested during the past two decades (Gabrielov et al., 1986; Keilis-Borok, 1990; Minster and Williams, 1992; Keilis-Borok and Shebalin, 1999; Kossobokov et al., 1999; Rotwain and Novikova, 1999; Vorobieva, 1999). The empirical nature of these algorithms, however, makes it difficult to evaluate their efficiency and strength in a formal way, due to the long time required for the tests in real predictions and to the possible over-fitting in retrospective studies. This stimulated the development of different specific methods for the evaluation of algorithms quality (Habermann and Creamer, 1994; Minster and Williams, 1992; Keilis-Borok and Shebalin, 1999). One relevant question that still needs to be answered is: to what extent are predictions influenced by the unavoidable errors affecting the input data?

Earthquake catalogs represent the most widely available geophysical data, containing systematically collected information about seismicity. This is why most of the studies concerning precursory phenomena, and therefore earthquake predictions, are based on the analysis of earthquake catalogs. The catalogs contain errors which can be distinguished into systematic and random ones (this is comprehensively

* Corresponding author. Tel.: +39-40-6762-129; fax: +39-40-6762-111.
E-mail address: anto@dst.univ.trieste.it (A. Peresan).

PII: S0031-9201(01)00311-9

discussed by Kossobokov, 1995). Systematic errors can be associated to changes in the data acquisition system or in the methods for the determination of earthquake parameters. Random errors correspond generally to the uncertainty of determination and to possible mistakes made during the data input process. The presence of a systematic error in magnitude may hamper prediction results and is generally quite difficult to detect (Habermann, 1991; Habermann and Creamer, 1994; Kossobokov and Shebalin, 1995); some aspects of this problem for the Italian catalog have been considered by Peresan et al. (2000).

The goal of this work is to study the influence on the results of intermediate-term earthquake prediction of random errors in magnitude determination. We analyze here the particular case of earthquake predictions performed by CN algorithm (Keilis-Borok and Rotwain, 1990). The algorithm uses origin time, hypocentral coordinates and magnitude of earthquakes. Among these parameters, magnitude is the most significant source of errors in the results of predictions, because it enters in the determination of the values of the functions describing the seismic sequence as well as in the definition of the strong earthquakes (Rotwain and Novikova, 1999).

Our analysis is based on the predictions performed for central Italy (Peresan et al., 1999b). To establish the dependence of the prediction results on possible random errors in magnitude, the algorithm CN is applied to several "randomized" catalogs (RC). These RC are obtained by random modification, within the range of the assumed errors, of the magnitudes of all events. The results of predictions obtained for the original catalog (OC) and the RCs are compared, providing useful information about the stability and the expected performances of the algorithm.

2. General scheme of prediction with CN algorithm

The algorithm CN has been designed for the prediction of *strong earthquakes*, which are the events with magnitude greater or equal to a fixed threshold M_0. The algorithm is based on the analysis of a set of empirical functions describing the earthquake flow. These functions are normalized by thresholds in magnitude, which are selected on the basis of the average return period of events observed during the *thresholds setting period*. The functions are discretized into small, medium and large values, accordingly to the level of seismic activity in the considered region, and the thresholds for discretization are selected by the retrospective analysis of seismicity within the thresholds setting period. The discretization of functions causes some loss of information, but makes the algorithm more robust with respect to fluctuations in the data. The thresholds setting period must correspond to an interval of time long enough to provide a representative sample of the seismic activity within the considered region, including periods of quiescence as well as periods of high activity (Keilis-Borok and Rotwain, 1990).

The algorithm CN identifies the times of increased probability (TIPs) for the occurrence of strong earthquakes. When a strong event occurs during a TIP, then it is indicated as a *successful prediction*, otherwise it is referred as *failure to predict*. If no strong earthquake occurs during a declared TIP, then the TIP is called a *false alarm*.

According to Molchan (1990), the results of a prediction can be characterized by two types of errors. The first one is the percentage η of failures to predict: $\eta = F/N$, where F is the number of failures to predict and N the number of events to be predicted. The second one is the percentage τ of the total duration of alarms: $\tau = A/T$, where A is the total duration of alarms and T the length of the whole time interval considered. The strength of a prediction is estimated by the analysis of the *error diagram*, collecting information on both types of errors. According to Molchan (1990, 1996), in order to characterize the quality of predictions in terms of the errors η and τ, it is possible to consider any convex function $\Omega = f(\eta, \tau)$. Among the several possible functions, the sum of errors appears to be the most straightforward and suitable for the evaluation of the outcomes, as recommended by Molchan (1996). Hence, in the present analysis the quality of predictions will be quantified by the sum of errors: $\Omega = \eta + \tau$. Since the random prediction gives $\Omega = 1$ (Molchan, 1990), one can roughly estimate the quality of prediction by the deviation of Ω from unity (or from the corresponding percentage $\Omega = 100\%$).

3. Magnitude randomization procedure

The procedure of magnitude randomization simulates possible random magnitude errors in the analyzed catalog. In the present study, we concentrate on two types of random errors: the error of *measurement* and the errors of magnitude *discretization*.

The value of magnitude reported in the catalog is usually the result of estimations made from several stations recording the occurrence of an earthquake. The error of measurement represents the several factors (radiation pattern, local effects, etc.) that may influence the records of an event at different stations and consequently affect the magnitude estimation. We use here a rough assumption that the final measurement error ΔM_M is normally distributed. Since the observed magnitude is finite, we use for ΔM_M the truncated normal distribution $F^{TR}(x)$, that is defined on the interval $x \in [-\Delta M_{max}, \Delta M_{max}]$ as

$$F^{TR}(x) = \frac{F(x) - F(\Delta M_{max})}{F(\Delta M_{max}) - F(-\Delta M_{max})} = \frac{F(x) - F(\Delta M_{max})}{2F(\Delta M_{max}) - 1} \quad (1)$$

where $F(x)$ denotes the cumulative normal distribution with 0 mean and S.D. $= \Delta M_{max}/3$. The measurement error ΔM_M thus becomes normally distributed on the interval $[-\Delta M_{max}, \Delta M_{max}]$ while the values outside this interval are disallowed.

The values of magnitudes are calculated as real numbers with several digits, but their precision hardly exceeds the first decimal digit (e.g. the magnitudes calculated from intensity, which is a discrete scale, exhibit a discrete distribution); for this reason the measured values are usually rounded, i.e. they are discretized to fit some predefined lattice. Let M_C be the operating magnitude selected from the catalog and k the step of the discretization lattice, when discretizing (i.e. rounding) the magnitude, an error as large as $\pm k/2$ can be introduced. Since the measured magnitude may correspond to any of the values in the magnitude interval $[M_C - (k/2), M_C + (k/2))$, then we assume that the error of magnitude discretization ΔM_K has a uniform distribution within the interval $[-k/2, k/2]$.

Considering the quantities M_C, ΔM_M, and ΔM_K defined earlier, we introduce the randomized magnitude M_R

$$M_R = M_C + \Delta M_M + \Delta M_K \quad (2)$$

where the measurement error ΔM_M and the error of discretization ΔM_K are independent.

4. Data analysis

4.1. Numerical parameters for randomization

The randomization procedure, introduced in Section 3, depends on two parameters: k and ΔM_{max}. The parameter k is uniquely determined by the used catalog, which is fully described by Peresan et al. (1999a).

The catalog is composed by the CCI1996 (Peresan et al., 1997) for the period 1900–1985, and is updated using the NEIC preliminary determinations of epicentres (PDE) since 1986. The operating magnitude in the catalog CCI1996 is selected according to the following priority order: M_L, M_d, M_I (Molchan et al., 1997), where M_L is the local magnitude, M_d the duration magnitude and M_I the magnitude from intensities. A corresponding priority choice has been defined for the magnitudes in the PDE catalogue as follows: $M2$, $M1$, M_s. The magnitude from the surface waves estimated by NEIC is given by M_s, while $M1$ and $M2$ correspond to magnitudes of different kind, supplied by different agencies, mainly corresponding to local and duration magnitudes (Peresan et al., 1999a and references therein).

The number of earthquakes in different magnitude intervals and for three time periods (1950–1980, 1980–1999 and for the entire interval 1950–1999) is given in Fig. 1. It is possible to observe that since 1950 the magnitudes are determined to the first decimal digit, hence the discretization step is $k = 0.1$.

The second parameter of randomization, ΔM_{max}, is a variable one. According to Båth (1973), from empirical observations we can expect errors as large as ± 0.3 units in reported magnitudes; such a value is confirmed by theoretical arguments (Panza and Calcagnile, 1974; Herak et al., 2001). Hence, in the present study we will assume $\Delta M_{max} = 0.3$ as an upper bound for realistic errors in magnitude. Larger values for ΔM_{max} are considered in order to evaluate the dependence on ΔM_{max} of the quality of the results. The randomization procedure is applied to the OC, varying the parameter

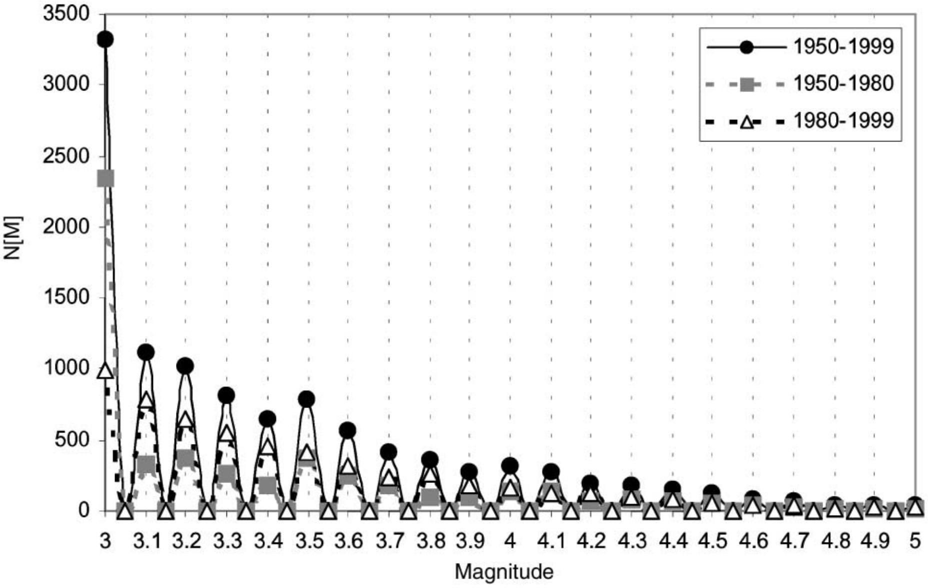


Fig. 1. Number of earthquakes with $3 \leq M \leq 5$ reported in the OC (Peresan et al., 1999a) for three different periods of time: 1950–1980, 1980–1999 and 1950–1999.

$\Delta M_{\max}$, and a set of 110 RCs is generated: 30 catalogs for $\Delta M_{\max} = 0.1$, 0.2 and 0.3, plus 10 catalogs for $\Delta M_{\max} = 0.4$ and 0.5.

It is necessary to check that the randomization procedure does not affect the basic features of the earthquake sequence, the magnitude–frequency relation being the most important. The magnitude–frequency relations for the OC and for the RCs with $\Delta M_{\max} = 0.3$ and $\Delta M_{\max} = 0.5$ are shown in Fig. 2. The distributions are quite similar and the differences with respect to the OC become relevant only at the largest magnitudes, due to the small number of events. The linearity of the frequency–magnitude relation is preserved even for $\Delta M_{\max}$ as large as 0.5 and the *b*-value does not change significantly, since it is mainly controlled by the small and intermediate magnitude events.

4.2. Predictions based on the original catalog

The application of CN algorithm for the intermediate-term earthquake prediction in central Italy (Fig. 3) is described in detail by Peresan et al. (1999b). Six strong events with $M \geq M_0 = 5.6$ occurred during the considered time interval. In two cases (21 August 1962 and 26 September 1997) two strong earthquakes occurred in the same day and very close in space. Such "coupled" events cannot be distinguished at the time scale characteristic of the algorithm, since CN predictions are performed with a time step of 2 months. These earthquakes will be associated to the same TIP, hence they must be counted just as a single event (i.e. instead of six strong earthquakes, there are just four events to be predicted).

The results of predictions obtained for two different thresholds setting periods are shown in Fig. 4. In the first case the thresholds setting period, referred as *long period*, lasts from 1 January 1954 to 31 December 1998. In the second case, it lasts from 1 January 1954 to 31 December 1985, and it is referred as *short period*. The results of predictions are summarized in the following paragraphs.

For the long thresholds setting period (Fig. 4a), three out of the four strong events are predicted, with the percentage of the total alarm duration $\tau \approx 21\%$. The coupled event ($M = 5.7$ and 6.0), occurred on 26 September 1997, is a failure to predict; hence $\eta = 25\%$ and $\Omega \approx 46\%$. For the short thresholds setting period (Fig. 4b), TIPs occupy about 22% of the total time and precede all the four strong events, hence $\eta = 0$, $\tau \approx$

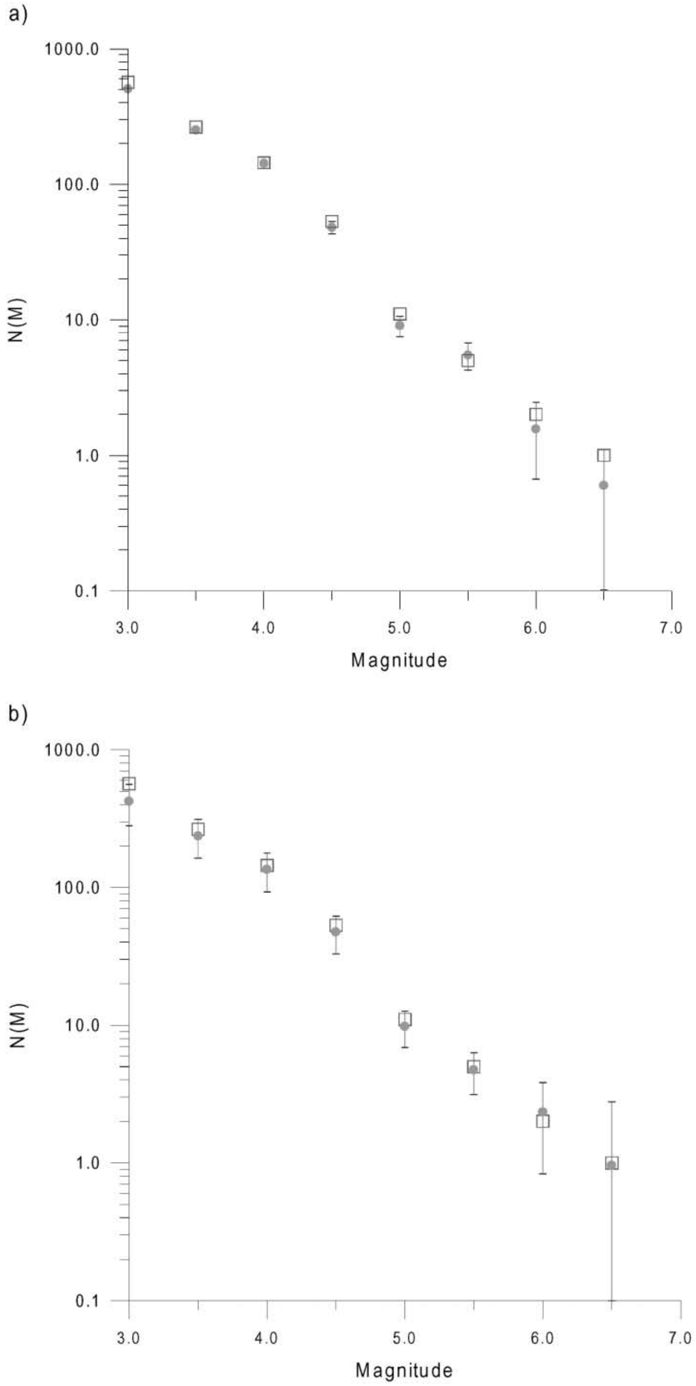


Fig. 2. Distribution of the number of events versus magnitude for the RCs with (a) $\Delta M_{max} = 0.3$ and (b) $\Delta M_{max} = 0.5$. For each magnitude interval the average number of events (filled dots) and its S.D. (vertical bars) are given for the RCs, together with the number of events in the OC (open squares).

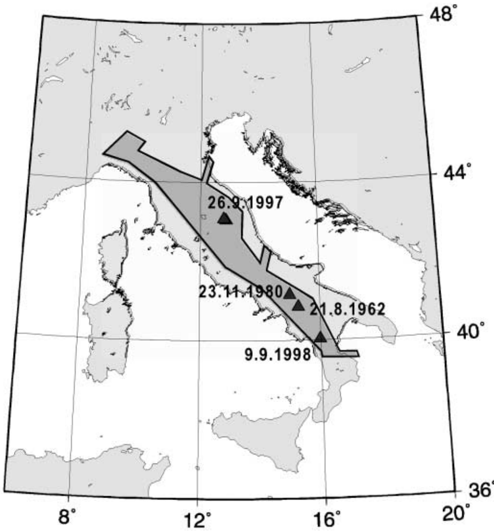


Fig. 3. Region used for the routine application of the algorithm CN in central Italy (Peresan et al., 1999b). The epicentres of the strong events are given together with their occurrence time. Note that the events occurred on 21 August 1962 and 26 September 1997, are "coupled" events (i.e. two strong earthquakes occurred in the same day and close in space).

22% and $\Omega \approx 22\%$. Therefore, predictions associated with the short thresholds setting period seem to provide a better score, despite of the reduced information considered. However, the analysis of the stability of these two results, described in Section 4.3, will show that the long thresholds setting period supplies more stable results.

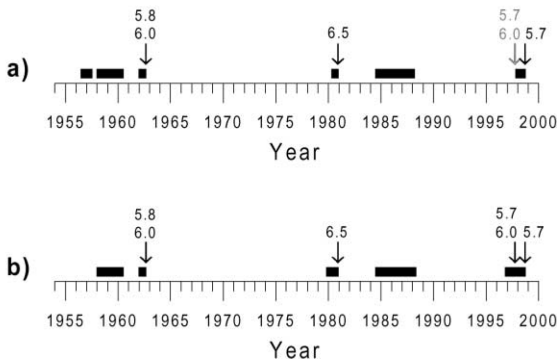


Fig. 4. TIPs obtained with the OC for (a) the long thresholds setting period and (b) the short thresholds setting period. Black boxes indicate the TIPs (periods of alarm). Arrows with a number above indicate the time of occurrence of strong earthquakes and their magnitude; failures to predict are given in gray.

4.3. Predictions based on the randomized catalogs

The algorithm CN is applied to each of the RCs independently, following the standard rules. The threshold M_0 for the selection of the events to be predicted is kept equal to 5.6, the same as for the OC, in order to check the stability of the set of events to be predicted with respect to this threshold.

The results of predictions obtained for 50 RCs—10 for each $\Delta M_{\max}$ from 0.1 to 0.5—are synthetically described by means of the sum of errors $\Omega = \eta + \tau$. The values of Ω versus $\Delta M_{\max}$ are shown in Fig. 5.

With the long thresholds setting period (Fig. 5a), individual results can be comparable or even better than the original one for $\Delta M_{\max}$ up to 0.4; the values of $\langle \Omega \rangle$, averaged for each $\Delta M_{\max}$, are approximately equal to the original value ($\Omega = 46\%$) for $\Delta M_{\max} \leq 0.2$. The exact values of the average errors of prediction, $\langle \tau \rangle$, $\langle \eta \rangle$ and $\langle \eta + \tau \rangle$ (together with its S.D. σ) for the different $\Delta M_{\max}$, are given in Table 1. The original result obtained with the long thresholds setting period appears quite stable with respect to magnitude errors up to ± 0.2.

For the short thresholds setting period (Fig. 5b), one can observe an increasing trend of $\langle \Omega \rangle$, showing that the quality of predictions decreases almost linearly with $\Delta M_{\max}$. In particular, the average values $\langle \Omega \rangle$ are significantly larger than the $\Omega = 22.1\%$ corresponding to the OC and all the results obtained using the RCs appear worse than the original one. Hence, this result appears very sensitive to possible magnitude errors. Its good quality, obtained at the price of stability, seems to indicate some over-fitting to the original data, and casts some doubts about the real predictive capability of CN, when using the short thresholds setting period.

The analysis described earlier allows us to single out the instability of the prediction results related to the insufficient length of the thresholds setting period. In fact, due to the small number of strong events occurred during the short thresholds setting period (only two events), the information provided to the algorithm about the seismic activity preceding the strong events is limited, and hence it strongly depends on the few given cases.

Henceforth, we consider only the long thresholds setting period. For $\Delta M_{\max} = 0.4$ and 0.5, 10 tests permit already to evidence the instability of results; therefore it seems not necessary to investigate further

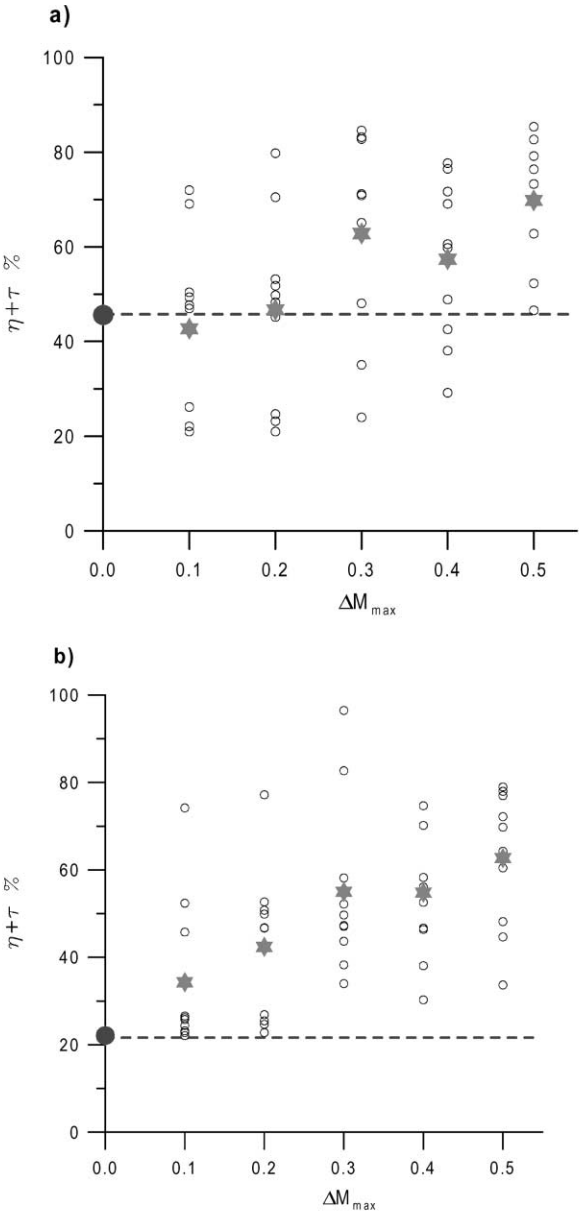


Fig. 5. Diagram showing the dependence of prediction errors $\Omega = \eta + \tau$ on ΔM_{max} for (a) the long thresholds setting period and (b) the short thresholds setting period. The circles indicate the $\eta + \tau$ values for the 10 RCs used for each ΔM_{max}; stars show their average values. The horizontal dashed line and the full dot indicate the $\eta + \tau$ value for the OC. The quantities η and τ are defined in the text.

Table 1
Average errors of prediction for different ΔM_{max}

ΔM_{max}	$\langle\tau\rangle$ (%)	$\langle\eta\rangle$ (%)	$\langle\eta + \tau\rangle \pm \sigma$ (%)
0.1	22.8	22.5	42.7 ± 19.1
0.2	24.1	28.8	46.7 ± 19.6
0.3	26.5	40.6	62.8 ± 20.9
0.4	27.1	30.3	57.4 ± 17.0
0.5	28.6	41.2	69.8 ± 12.7

the effects of such large errors. For smaller errors, instead, additional tests are needed, in order to properly evaluate the stability of the results, as well as to improve the significance of the analysis. A comprehensive analysis is then performed for 90 RCs, 30 each for $\Delta M_{max} = 0.1$, 0.2 and 0.3, respectively. In fact, $\Delta M_{max} = 0.3$ is a realistic estimate of the upper bound for the error which can affect magnitudes (Båth, 1973; Panza and Calcagnile, 1974; Herak et al., 2001) and it corresponds to the value of ΔM_{max} for which predictions begin to be unstable (Fig. 5a).

The stability of the results is evaluated by means of the error diagrams (Molchan et al., 1990), compiled for the results obtained with the OC and the RCs (Fig. 6). Each point in the (η, τ) plane corresponds to the prediction obtained for a given catalog; the diagonal line defined by the equation $\Omega = \eta + \tau = 100\%$ corresponds to the results of a random guess. The clustering of points indicates the stability of predictions, while the increasing distance from the diagonal line indicates the increasing quality of the predictions. Fig. 6 shows that, for $\Delta M_{max} \leq 0.2$, the results are quite well clustered around the one obtained with the OC, while for $\Delta M_{max} \leq 0.3$ they appear much more scattered. The error τ, indicating the percentage of time occupied by alarms, does not vary significantly with respect to the original predictions. Therefore, the changes in the quality of the results are mainly due to a larger percentage of failures to predict η, which is strictly related to the changes in the number of events to be predicted. This aspect emerges quite clearly from Table 2, showing the average difference between the CN results obtained using the OC and those obtained using the RCs: $\langle\Delta\tau\rangle = \langle\tau\rangle - \tau_O$ and $\langle\Delta\eta\rangle = \langle\eta\rangle - \eta_O$, where τ_O and η_O are the prediction errors for the OC (Fig. 6). The average variation of the number of strong events N_{SE} has been evaluated by means of the relation: $\langle\Delta N_{SE}\rangle = \langle|N_R - N_O|/N_O\rangle$, where N_O and

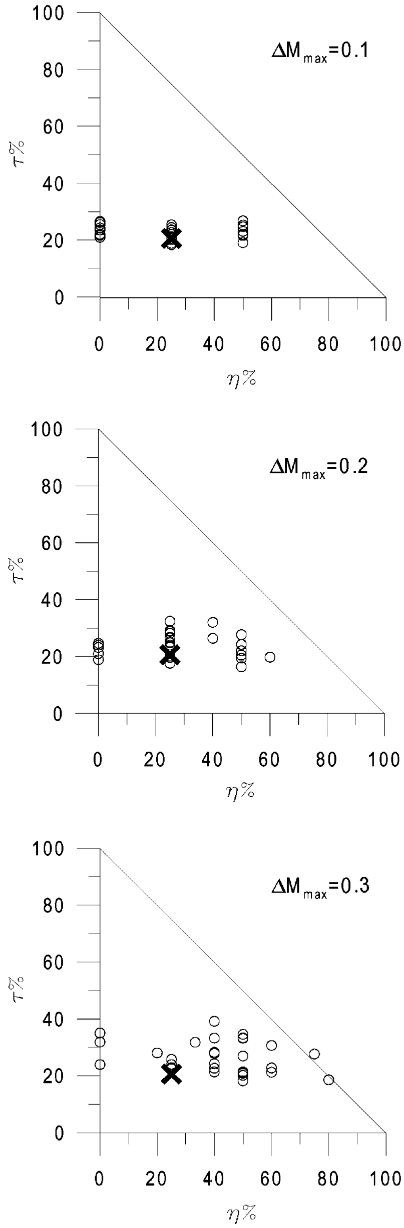


Fig. 6. Error diagrams for the results of TIPs diagnosis using the OC (bold cross) and the RCs (circles) with long thresholds setting period. The diagonal line corresponds to the results of a random guess. Thirty tests have been performed for each ΔM_{max}.

N_R are the numbers of events to be predicted (with $M \geq 5.6$) in the OC and in the RCs, respectively; for $\Delta M_{max} = 0.3$ both $\langle \Delta N_{SE} \rangle$ and $\langle \Delta\eta \rangle$ increase significantly.

Table 2
Average difference between the results of CN obtained using the OC and those obtained using the RCs

ΔM_{max}	$\langle \Delta\tau \rangle$ (%)	$\langle \Delta\eta \rangle$ (%)	$\langle \Delta(\eta + \tau) \rangle$ (%)	$\langle \Delta N_{SE} \rangle$ (%)
0.1	2.2	−2.5	−0.3	0.0
0.2	3.5	3.8	7.3	2.5
0.3	5.9	15.6	21.5	13.4

$\langle \Delta\tau \rangle$, $\langle \Delta\eta \rangle$: average variation of prediction errors for the RCs; $\langle \Delta N_{SE} \rangle = \langle |N_R - N_O|/N_O \rangle$: average variation of the number of strong events; N_O: number of events to be predicted in the OC; N_R: number of events to be predicted in the RC.

A list of the earthquakes, ordered by decreasing magnitude, with $M \geq 5.3$ in the OC, occurred in central Italy (Fig. 3) during the period 1950–1999, is provided in Table 3. This table contains information about all the events whose randomized magnitude may exceed the threshold $M_0 = 5.6$ for $\Delta M_{max} = 0.3$. The percentage of times each earthquake turns out to be a strong event, N_{SE}, and the percentage of times it is predicted (N_P) are provided as well, considering 30 RCs. The table shows that the earthquakes with original magnitude less than 5.6, sporadically becoming strong events, are never predicted. This explains the drastic increase of the error η for $\Delta M_{max} = 0.3$, and at the same time indicates that the randomization does not introduce spurious alarming patterns, as it clearly appears from the analysis of the stability of TIPs identification.

To evaluate the stability of predictions we have analyzed also the variation of TIPs. Considering a set of predictions made for the same time interval, the quantity $\Theta(t)$ has been defined as the percentage of cases where the instant t is identified as a TIP.

Table 3
List of earthquakes with $M \geq 5.3$, central Italy 1950–1999

Event date	M	N_{SE} (%)	N_P (%)
23 November 1980	6.5	100.0	66.7
21 August 1962[a]	5.8/6.0	100.0	93.3
26 September.1997[a]	5.7/6.0	100.0	40.0
9 September 1998	5.7	83.3	56.0
19 September 1979	5.5	23.3	0.0
5 May 1990	5.5	12.0	0.0
7 May 1984	5.4	0.0	0.0

[a] “Coupled” events, occurred in the same day and very close in space. The magnitudes of both the strong earthquakes are provided.

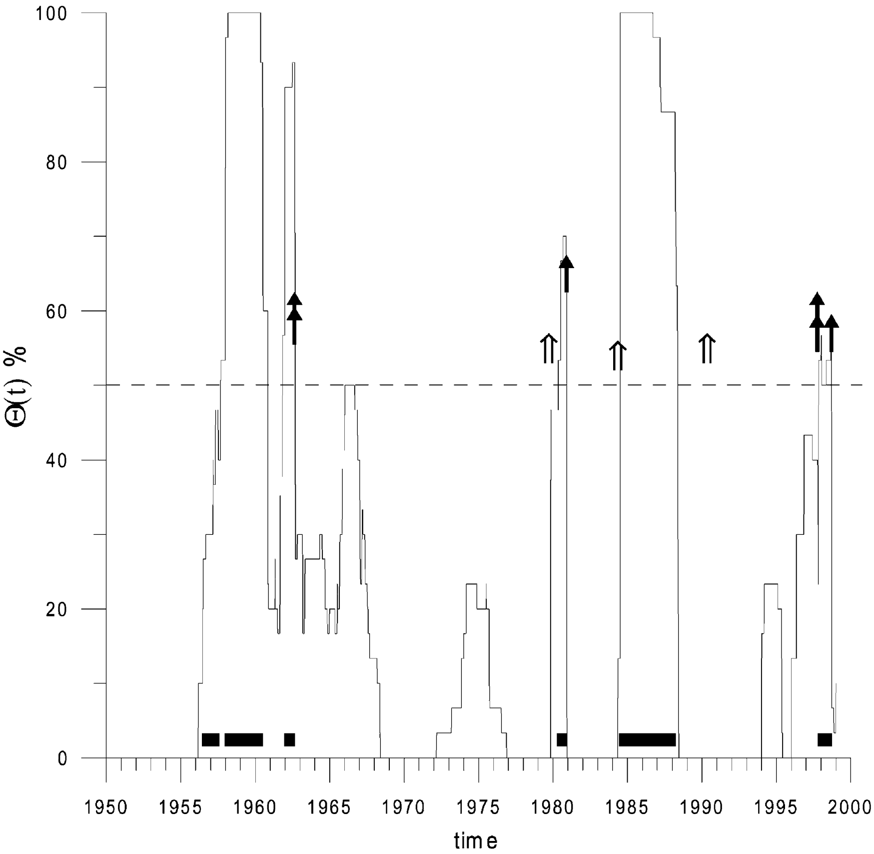


Fig. 7. The percentage $\Theta(t)$ is the number of tests for which the time (t) belongs to a TIP. Black boxes represent the TIPs declared using the OC. Black arrows indicate the time of occurrence of strong earthquakes ($M \geq M_0 = 5.6$) in the OC. Grey open arrows indicate additional strong earthquakes sporadically appearing in the RCs. The distance of the arrows from the dashed line is proportional to the magnitude reported in the OC. Thirty tests have been performed with $\Delta M_{\max} = 0.3$.

Fig. 7 shows the function $\Theta(t)$ based on the results of predictions for the 30 RCs with $\Delta M_{\max} = 0.3$.

All the original TIPs appear to be very stable; moreover, TIP is never declared during considerable part of the time. This confirms the conclusion about the high stability of the results.

5. Discussion and conclusions

The stability of the intermediate-term predictions with respect to random errors in magnitudes has been evaluated considering as an example the CN prediction for central Italy. The algorithm has been applied, following the standard rules, to a set of catalogs with magnitudes randomly modified (RCs) within the range of the assumed errors, and the outcomes of these predictions have been compared with the results obtained with the OC. Our analysis shows that the results of prediction remain stable for $\Delta M_{\max} < 0.3$. The quality of predictions seems to be mainly controlled by the percentage of failures to predict, which depends on the changes in the number of strong earthquakes ($M \geq M_0$), almost negligible for $\Delta M_{\max} \leq 0.2$. The strong events generated by the randomization are never predicted, thus increasing the percentage of failures to predict.

The procedure used here for magnitude randomization has been defined on the basis of quite rough and sketchy arguments. This study, however, is not aimed to provide neither the optimal randomization procedure nor the best definition of random magnitude errors and certainly, the appropriate construction of the optimal model for such errors represents an open problem.

Prediction algorithms have generally a set of fitting parameters that, in the case of CN algorithm, correspond to the thresholds for normalization and discretization of functions. These parameters are fixed by the retrospective analysis of seismicity during the thresholds setting period; once they are assigned from the analysis of past seismicity the algorithm is applied for prediction, using such fixed parameters. The randomization procedure introduced in this paper can help both to evaluate the stability of the algorithm and to prevent data over-fitting. The tests performed with the RCs show that, with a short thresholds setting period, the average quality of results decreases linearly for increasing magnitude error, while with a longer thresholds setting period it appears to be constant, at least for $\Delta M_{\max} < 0.3$. This indicates that, in order to guarantee a certain stability of the results, the thresholds setting period must be long enough to include a significant sample of dangerous and non-dangerous intervals of time, otherwise the assigned parameters of the algorithm will strongly depend on the few given cases. This would lead to an excessive sensibility to possible data errors and hence to worst prediction results.

Therefore, from now onward, the application of the algorithm CN in Italy will be performed using the parameters fixed with the analysis of the OC with the long thresholds setting period. In this case, the results should be robust at least for magnitude errors lower than ± 0.3, with an expected score around $\Omega = 55\%$. This value of Ω is in good agreement with the average performance of CN in 22 regions of the world (Rotwain and Novikova, 1999). The results of routine predictions for the central Italy region (Fig. 3), updated to 1 September 2001, indicate a TIP, lasting from November 2000 up to 1 September 2002, both considering the short and the long thresholds setting period.

Acknowledgements

We are grateful to M.G. Shnirman and I. Kuznetsov for the precious suggestions and for the useful discussions. During this study AP was supported by MURST (Cofinanziamento and 60% funds) and by CNR (contract 98.03238.PF54), IR and IZ were supported by a sucontacted with Cornell University, Geological Sciences, under EAR-9804859 from the US National Science and administered by the US Civilian Research & Development Foundation for the Independent States of the former Soviet Union (CRDF) and by the Abdus Salam International Centre for Theoretical Physics, Trieste, Italy.

References

Båth, M., 1973. Introduction to seismology. Birkhauser, Basel, 365 pp.

Gabrielov, A.M., Dmitrieva, O.E., Keilis-Borok, V.I., Kossobokov, V.G., Kutznetsov, I.V., Levishina, T.A., Mirzoev, K.M., Molchan, G.M., Negmatullaev, S.Kh., Pisarenko, V.F., Prozorov, A.G., Rinheart, W., Rotwain, I.M., Shelbalin, P.N., Shnirman, M.G., Schreider, S.Yu., 1986. Algorithms of long term earthquakes' prediction, International School for Research Oriented to earthquake Prediction Algorithms. Software and Data Handling. Lima, Peru.

Habermann, R.E., 1991. Seismicity rate variations and systematic changes in magnitudes in teleseismic catalogs. Tectonophysics 193, 277–289.

Habermann, R.E., Creamer, F., 1994. Catalog errors and the M8 earthquake prediction algorithm. Bull. Seismol. Soc. Am. 84, 1551–1559.

Herak, M., Panza, G.F., Costa, G., 2001. Theoretical and observed depth corrections for M_s. Pure and Appl. Geophys. 158, 1517–1530.

Keilis-Borok, V.I. (Ed.), 1990. Intermediate-term Earthquake Prediction: Models, Algorithms, Worldwide Tests (special issue). Phys. Earth Planet. Int. 61.

Keilis-Borok, V.I., Rotwain, I.M., 1990. Diagnosis of time of increased probability of strong earthquakes in different regions of the world: algorithm CN. Phys. Earth Planet. Int. 61, 57–72.

Keilis-Borok, V.I., Shebalin, P.N. (Eds.), 1999. Dynamics of Lithosphere and Earthquake Prediction (special issue). Phys. Earth Planet. Int. 111 179–330.

Kossobokov, V.G., 1995. Analysis of earthquake catalogs. In: Proceedings of the Third Workshop on Non-Linear Dynamics and Earthquake Prediction. International Centre for Theoretical Physics, Trieste, Italy.

Kossobokov, V.G., Shebalin, P.N., 1995. Comment on Catalog Errors and the M8 Earthquake Prediction Algorithm. In: Habermann, R.E., Creamer, F (Eds.), Proceedings of the Third Workshop on Non-Linear Dynamics and Earthquake Prediction. International Centre for Theoretical Physics, Trieste, Italy.

Kossobokov, V.G., Romashkova, L.L., Keilis-Borok, V.I., Healy, J.H., 1999. Testing earthquake prediction algorithms: statistically significant real-time prediction of the largest earthquakes in the Circum-Pacific, 1992–1997. Phys. Earth Planet. Int. 111 (3/4), 187–196.

Minster, J.B., Williams, N.P., 1992. The M8 intermediate term earthquake prediction algorithm: an independent assessment. EOS (Transactions, AGU) 73, S22B-06.

Molchan, G.M., 1990. Strategies in strong earthquake prediction. Phys. Earth Planet. Int. 61, 84–98.

Molchan, G.M., 1996. Earthquake prediction as a decision-making problem. PAGEOPH 147 (1), 1–15.

Molchan, G.M., Dmitrieva, O.E., Rotwain, I.M., Dewey, J., 1990. Statistical analysis of the results of earthquake prediction, based on bursts of aftershocks. Phys. Earth Planet. Int. 61, 128-139.

Molchan, G.M., Kronrod, T.L., Panza, G.F., 1997. Multiscale seismicity model for seismic risk. Bull. Seismol. Soc. Am. 87 (5), 1220–1229.

Panza, G.F., Calcagnile, G., 1974. Some remarks about the focal effect on the magnitude determination from Rayleigh waves. In: Proceedings of the XIIIth General Assembly of the European Seismological Commission, Tech. Econ. Stud. 10, 327–334

Peresan, A., Costa, G., Vaccari, F., 1997. CCI1996: the current catalog of Italy, Internal report IC/IR/97/9. International Centre for Theoretical Physics, Trieste, Italy.

Peresan, A., Costa, G., Panza, G.F., 1999a. A proposal of regionalization for the application of the CN earthquake prediction algorithm to the Italian territory. Ann. Geofisica. 42, 281–306.

Peresan, A., Costa, G., Panza, G.F., 1999b. Seismotectonic model and CN earthquake prediction in Italy. PAGEOPH 154, 281–306.

Peresan, A., Panza, G.F., Costa, G., 2000. CN algorithm and long lasting changes in reported magnitudes: the case of Italy. Geophys. J. Int. 141, 425–437.

Rotwain, I.M., Novikova, O., 1999. Performance of the earthquake prediction algorithm CN in 22 regions of the world. Phys. Earth Planet. Int. 111, 207–213.

Vorobieva, I.A., 1999. Prediction of a subsequent large earthquake. Phys. Earth Planet. Int. 111 (3/4), 197–206.

Geophys. J. Int. (2000) **141**, 425–437

CN algorithm and long-lasting changes in reported magnitudes: the case of Italy

A. Peresan,[1] G. F. Panza[1,2] and G. Costa[1]

[1] *Department of Earth Sciences, University of Trieste, via E. Weiss 1, 34127 Trieste, Italy. E-mail: anto@geosun0.univ.trieste.it*

[2] *International Centre for Theoretical Physics, SAND Group, ICTP, 34100 Trieste, Miramare, Italy*

Accepted 1999 December 1. Received 1999 November 9; in original form 1999 July 14

SUMMARY

Prediction methods based on seismic precursors, and hence assuming that catalogues contain the necessary information to predict earthquakes, are sometimes criticised for their sensitivity to the unavoidable catalogue errors and possible undeclared variations in the evaluation of reported magnitudes. We consider a real example and we discuss the effect, on CN predictions, of a long-lasting underestimation of the reported magnitudes.

Starting approximately in 1988, the CN functions in Central Italy evidence an anomalous behaviour, not associated with TIPs, that indicates an unusual absence of moderate events. To investigate this phenomenon, the magnitudes given in the catalogue used, which since 1980 is defined by the ING bulletins, are compared to the magnitudes reported by the global catalogue NEIC (National Earthquake Information Centre, USGS, USA) and by the regional LDG bulletins issued at the Laboratoire de Detection et de Geophysique, Bruyeres-le-Chatel, France.

The comparison is performed between the ING bulletins and the NEIC catalogue, considering the local, M_L, and duration, M_d, magnitudes, first within the Central region, and then extended to the whole Italian territory. To check the consistency of the conclusions drawn from ING and NEIC data, the comparison of local magnitudes is extended to a third data set, the LDG bulletins.

The differences between duration magnitudes M_d that are reported by ING and NEIC since 1983 appear quite constant with time. Starting in 1987, an average underestimation of about 0.5 can be attributed to M_L reported by ING for the Central region; this difference decreases to about 0.2 when the whole Italian territory is considered. The anomalous behaviour of the CN functions disappears if a magnitude correction of +0.5 is applied to M_L reported in the ING bulletins. However, such a simple magnitude shift cannot restore the real features of the seismic flow, and ING bulletins are not suitable for CN algorithm application.

Key words: earthquake catalogues, earthquake prediction, Italy, regionalization.

INTRODUCTION

CN is an intermediate-term earthquake prediction algorithm based on the quantitative analysis of premonitory phenomena, which can be detected in the seismic flow preceding the occurrence of strong earthquakes (Gabrielov *et al.* 1986; Keilis-Borok & Rotwain 1990). The quantification of the properties of the seismic flow is performed by means of a set of functions of time (Table 1), which evaluate variations in the seismic activity, seismic quiescence and space–time clustering of events. The normalization of the functions allows us to apply CN to regions with different seismic activity (Keilis-Borok 1996; Rotwain & Novikova 1999).

The CN algorithm has been applied to the monitoring of seismicity in Central Italy since 1990 (Keilis-Borok *et al.* 1990; Costa *et al.* 1996; Peresan *et al.* 1998a). The analysis of the time behaviour of CN functions for the different regionalizations defined for Central Italy (Fig. 1) allowed us to observe the common anomalous flat values of some functions (see Z_{max}, S_{max}, Sigma, K and G in Fig. 2), starting approximately in 1988. The flat trend of the functions, never observed before, indicates the absence of moderate events and hence evidences an unusual decrease in the seismicity rate, suggesting the need to check for possible changes in the magnitudes reported by the catalogue used.

Until July 1997 the catalogue used for CN monitoring in Italy was the CCI1996 (Peresan *et al.* 1997). This catalogue is composed of the revised PFG catalogue (Postpischl 1985) for the period 1000–1979, and since 1980 we have updated it with the bulletins distributed by the Istituto Nazionale di

Table 1. Definition of the time functions used in the CN algorithm for the quantification of the properties of the seismic flow (from Keilis-Borok *et al.* 1990). The magnitude thresholds m_1, m_2, m_3 that allow the normalization of the functions are fixed according to the average yearly frequency of the main shocks that occurred within the region during the learning period (1954–1986). For the Central region (in dark grey in Fig. 1) $m_1 = 4.2$, $m_2 = 4.5$, $m_3 = 5.0$, corresponding to the standard yearly average frequencies $n_1 = 3.0$, $n_2 = 1.4$, $n_3 = 0.4$.

$N_2(t)$	Number of main shocks with $M \geq m_3$ that occurred in the time interval $(t-3$ yr, $t)$.
$K(t)$	$K(t) = K_1 - K_2$, where K_i is the number of main shocks with $M_i \geq m_2$ and origin time $(t - 2j\ \mathrm{yr}) \leq t_i \leq [t - 2(j-1)\ \mathrm{yr}]$.
$G(t)$	$G(t) = 1 - P$, where P is the ratio between the number of the main shocks with $M_j \geq m_2 (m_2 > m_1)$ and the number of the main shocks with $M_j \geq m_1$. Only main shocks with origin time t_j in the interval $(t - 1\ \mathrm{yr}) \leq t_j \leq t$ are considered.
Sigma(t)	Sigma$(t) = \Sigma\, 10^{\beta(Mi-\alpha)}$; the main shocks with $m_1 \leq M_i \leq M_0 - 0.1$ and origin time $(t - 3$ years$) \leq t_i \leq t$ are included in the summation; $\alpha = 4.5$, $\beta = 1.00$.
$S_{max}(t)$	$S_{max}(t) = \max\{S_1/N_1, S_2/N_2, S_3/N_3\}$, where S_j is calculated as Sigma(t) for the events with origin time $(t - j\ \mathrm{yr}) \leq t_i \leq [t - (j-1)$ years$]$, and N_j is the number of earthquakes in the sum.
$Z_{max}(t)$	$Z_{max}(t) = \max\{Z_1/N_1^{2/3}, Z_2/N_2^{2/3}, Z_3/N_3^{2/3}\}$, where Z_j is calculated as S_j, but with $\beta = 0.5$ and N_j is the number of earthquakes in the sum.
$N_3(t)$	Number of main shocks with $M \geq m_2$, which occurred in the time interval (t-10 years, t-7 years)
$q(t)$	$q(t) = \Sigma_{j=1}^{6} \max\{0{,}6a_2 - n_j\}$, where a_2 is the average annual number of main shocks with $M_j \geq m_2$, n_j is the number of main shocks with $M_j \geq m_2$ and origin time $[t - (8+j)\ \mathrm{yr}] \leq t_i \leq [t - (2+j)\ \mathrm{yr}]$.
$B_{max}(t)$	Maximum number of aftershocks for each main shock counted within a radius of 50 km for the first 2 days after the main shock.

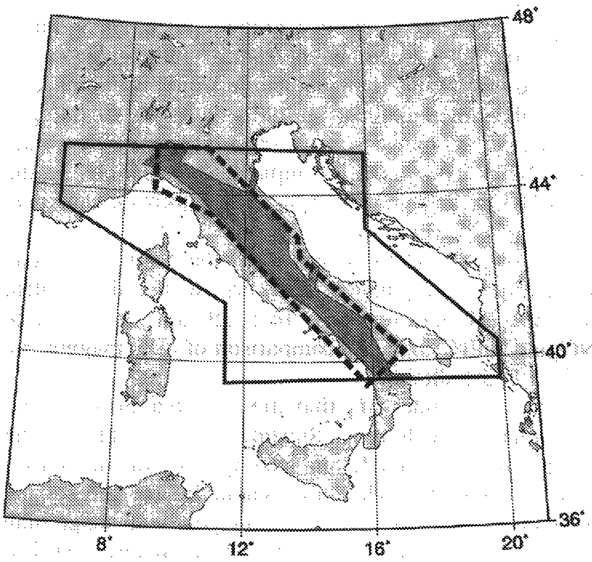


Figure 1. Different regionalizations defined for CN application to Central Italy. The continuous line delimits the region defined by Keilis-Borok *et al.* (1990), while the dotted line shows the region proposed by Costa *et al.* (1995). The region currently used for CN monitoring, defined strictly following the seismotectonic model (Peresan *et al.* 1998a), corresponds to the dark grey area.

Geofisica (ING). For the years 1980–1985 we use the ING paper bulletins, while from 1986 the upgrading is performed with the digital ING bulletins made available via ftp until July 1997. In order to check a possible change in reported magnitudes, the ING data are compared with the following catalogues (Table 2):

the Preliminary Determinations of Epicentres (PDE) distributed by NEIC, USGS, for the time period 1980–1997;

the Bulletins compiled at the Laboratoire de Detection et de Geophysique (CEA, Bruyeres-le-Chatel, France), referred to as LDG in the following, from January 1980 to December 1996.

We do not use the ISC catalogue since it does not provide revised M_L and M_d.

Table 2. Data set used for the catalogue comparison. For each agency the following are indicated: the period of time, the kind of catalogue and how the data are made available.

ING: Istituto Nazionale di Geofisica		
1980–1984	Revised ING bulletins	printed
1985–1986	Digital ING bulletins	floppy disk
1987–1997	Digital ING bulletins	ftp
LDG: Laboratoire de Detection et de Geophysique		
1980–1996	LDG Bulletins	Auto DRM
NEIC: National Earthquake Information Centre, USGS		
1980–1989	Global Hypocentres Data Base	cd-rom
1990–1997	Earthquake Hypocentres Data Files	ftp

The ING bulletins contain two estimations of magnitude: the local magnitude M_L and, since 1983, the duration magnitude M_d. The NEIC global catalogue reports the magnitudes m_b and M_S, both computed by NEIC, plus two values, $M1$ and $M2$, that correspond to magnitudes of a different kind contributed by different agencies. From a previous analysis of the NEIC catalogue (Peresan & Rotwain 1998) we observed that, for the Italian area, both $M1$ and $M2$ are mainly M_d and M_L, and that M_L is 10 times more frequent than M_d. Furthermore, ING is among the contributors to the PDE, and it supplied information for more than 600 events, from 1987 to 1997, as can be observed by listing the events with network code ROM reported in the PDE catalogue. Most of these events have magnitudes below 4.0, especially when M_d is considered, while about 100 of them have $M_L > 4.0$. The bulletins distributed by LDG contain two magnitude values, mainly corresponding to M_L and M_d.

In order to perform the magnitude comparison, the events common to the different catalogues are identified according to the following rules: (a) time difference $\Delta t \leq 1$ min; (b) epicentral distance $\Delta\mathrm{Lat} = \Delta\mathrm{Lon} \leq 1°$ for the comparison with the global catalogue (Storchak *et al.* 1998). No limitation is imposed on magnitude or depth differences.

The analysis is performed by evaluating, for a fixed type of magnitude, the quantities

$$\Delta M = M(\mathrm{C1}) - M(\mathrm{C2}), \qquad (1)$$

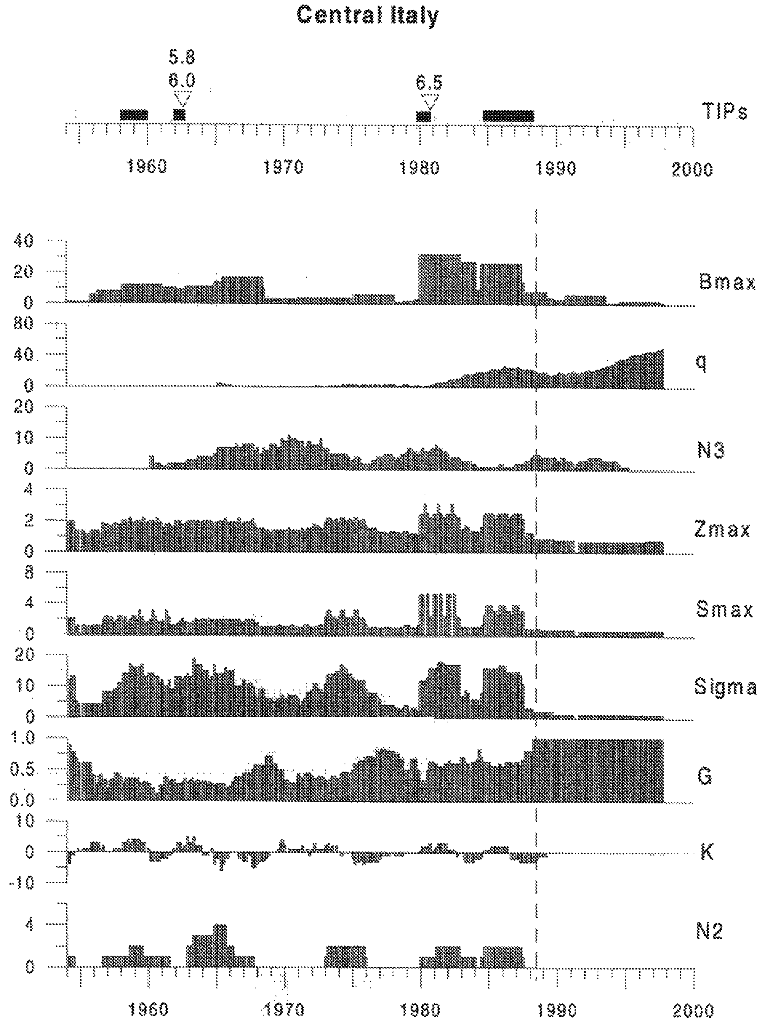


Figure 2. Time diagrams of the standard CN functions obtained for the Central region shown in Fig. 1. Functions Sigma, S_{max} and Z_{max} are evaluated for $4.2 \leq M \leq 4.6$, functions K, G, N_3, q for $M \geq 4.5$ and function N_2 for $M \geq 5.0$; magnitude thresholds have been selected according to the general rules for normalization of functions (Keilis-Borok & Rotwain 1990). The corresponding diagram of TIPs (times of increased probabilities) obtained using the CCI1996 catalogue is given at the top of the figure (triangles indicate the occurrence of strong events). The dotted line indicates the beginning of the anomalous behaviour of functions.

which are the differences between magnitudes of the same type reported in the catalogues C1 and C2 for each of the common earthquakes.

The comparison between ING and NEIC estimations is performed considering M_L and M_d separately among the events for which M_L and M_d are reported in both the catalogues. The events contributed to NEIC by ING, which represent a relatively small fraction of the set of common events (less than 10 per cent), are obviously excluded from the analysis. Initially, the comparison is focused on the Central region (Fig. 1) and the yearly average values ΔM_L and ΔM_d are evaluated from the common events contained in the area monitored using the CN algorithm. Subsequently, the comparison between the ING and NEIC catalogues is enlarged to the whole Italian territory and its surroundings, as shown in Fig. 9.

To check the consistency of the conclusions drawn from ING and NEIC data, the comparison of M_L is extended to a third catalogue, and the ING and NEIC M_L are compared directly with the M_L reported by the LDG bulletins. Since the LDG is among the NEIC contributors for the area analysed, the NEIC events with magnitude code LDG are obviously excluded when performing the comparison between LDG and NEIC data.

CHANGES IN REPORTED MAGNITUDES FOR CENTRAL ITALY

The analysis of the behaviour of CN functions in Central Italy allows us to identify the anomalous flat trend of some of the functions (Fig. 2), starting approximately in 1988. Such a flat trend indicates an unusual absence of moderate events.

To look for an explanation for this anomaly we focus our attention on the magnitude variations within the Central region currently used for the monitoring of seismicity (in dark grey in Fig. 1). The subcatalogue of earthquakes common to ING and NEIC contains about 800 events. The operating magnitude for CN monitoring is chosen from the Italian catalogue CCI1996, and hence from ING bulletins, according to the priority order M_L, M_d (Costa *et al.* 1996; Peresan *et al.* 1998a); therefore, local magnitudes play a relevant role in the CN analysis of seismicity. Hence, as a first stage, we study the discrepancies among the M_L values reported in the two catalogues, i.e. the quantity

$$\Delta M_L = M_L(\text{NEIC}) - M_L(\text{ING}). \qquad (2)$$

The histograms of ΔM_L are plotted for three contiguous ranges of magnitude (Fig. 3), chosen to correspond to the CN magnitude thresholds for Central Italy. The events with $M_L < 3$ are not used by CN, the events with $3.0 \leq M_L < 4.2$ are included only in the counting of aftershocks, and those with $M_L \geq 4.2$ can enter into the calculation of functions. For most of the events, $\Delta M_L > 0$, while a secondary peak around $\Delta M_L = 0$ can be seen in Fig. 3 for the smaller events.

In order to detect a possible undeclared long-lasting change in the estimation of the reported M_L, the time behaviour of the yearly average of ΔM_L is analysed considering only earthquakes with $M_L(\text{NEIC}) \geq 3.0$. The yearly number of such events is around 20–25, with two exceptions: there were 83 earthquakes in 1980 (mainly associated with the Irpinia event of 1980 November 23) and only four events in 1987.

The time distribution of ΔM_L yearly averages, shown in Fig. 4(a), indicates the presence of a major discontinuity in 1987. The average ΔM_L, estimated using eq. (2) for two

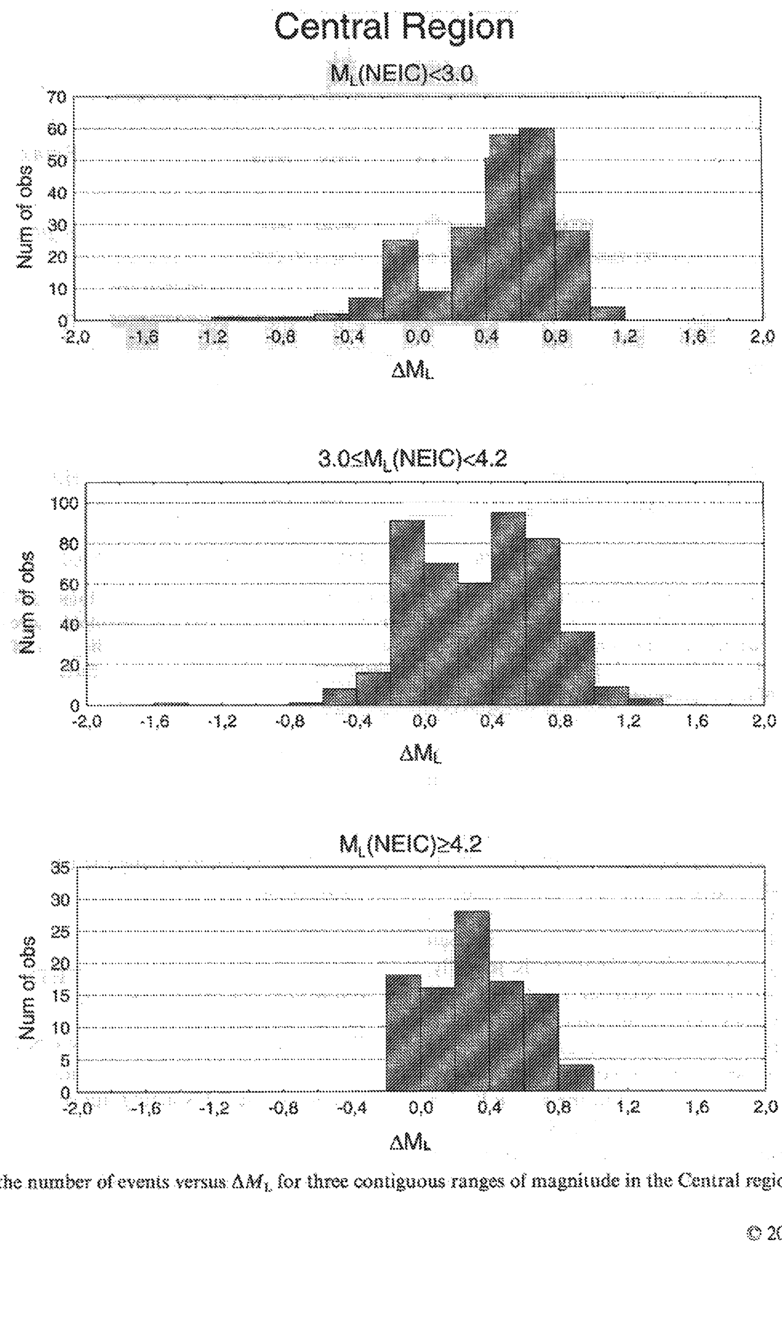


Figure 3. Histograms of the number of events versus ΔM_L for three contiguous ranges of magnitude in the Central region (dark grey area in Fig. 1).

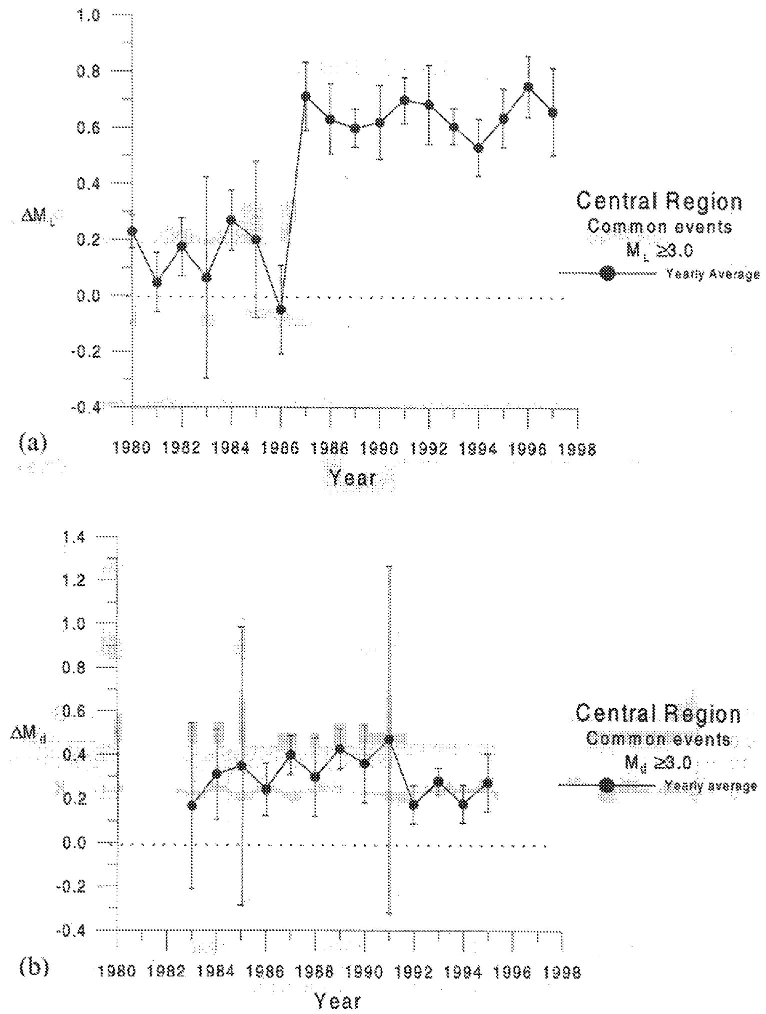


Figure 4. Yearly average of (a) ΔM_L and (b) ΔM_d obtained for the NEIC and ING catalogues, considering the common events that occurred within the Central region (Fig. 1). Error bars correspond to the 95 per cent confidence interval of the mean.

subsequent periods of time, excluding the year of transition, 1987, are as follows (the error corresponds to the 95 per cent confidence interval of the mean):

$$(1980\text{–}1986) \quad \Delta M_L = 0.13 \pm 0.05,$$
$$(1988\text{–}1997) \quad \Delta M_L = 0.64 \pm 0.04.$$

According to these average results, assuming M_L(NEIC) as a uniform reference value, an underestimation of about 0.5 can be assigned to the M_L values reported by ING since 1987.

A similar analysis, performed by replacing M_L with M_d in eq. (2), does not evidence a significant change for M_d(ING). The relevant uncertainty associated with the value of ΔM_d (Fig. 4b) for the years 1985 and 1991 is mainly due to the reduced sample size (only two events in 1985 and four in 1991). The average magnitude difference for the whole period 1983–1995 for which the sample is available is estimated to be $\Delta M_d = 0.30 \pm 0.04$.

CN: A DETECTOR OF ANOMALOUS VARIATIONS IN REPORTED MAGNITUDES

In order to understand whether the variations found in reported magnitudes can account for the anomalous behaviour of the CN functions observed in the Central region, the quantity $D = 0.5$ is added to the M_L reported by the ING bulletins, beginning in 1987. M_d values do not need to be modified because no significant time variation has been detected. CN is then applied to the Central region using the 'corrected' catalogue and following the standard procedure of forward monitoring of seismicity: learning is not repeated and the parameters are kept unchanged. The time diagram obtained is shown in Fig. 5 and clearly indicates that the anomalous behaviour of some CN functions, shown in Fig. 2, is no longer present.

Obviously, this magnitude transformation cannot be used to correct the catalogue and the magnitude revision must be

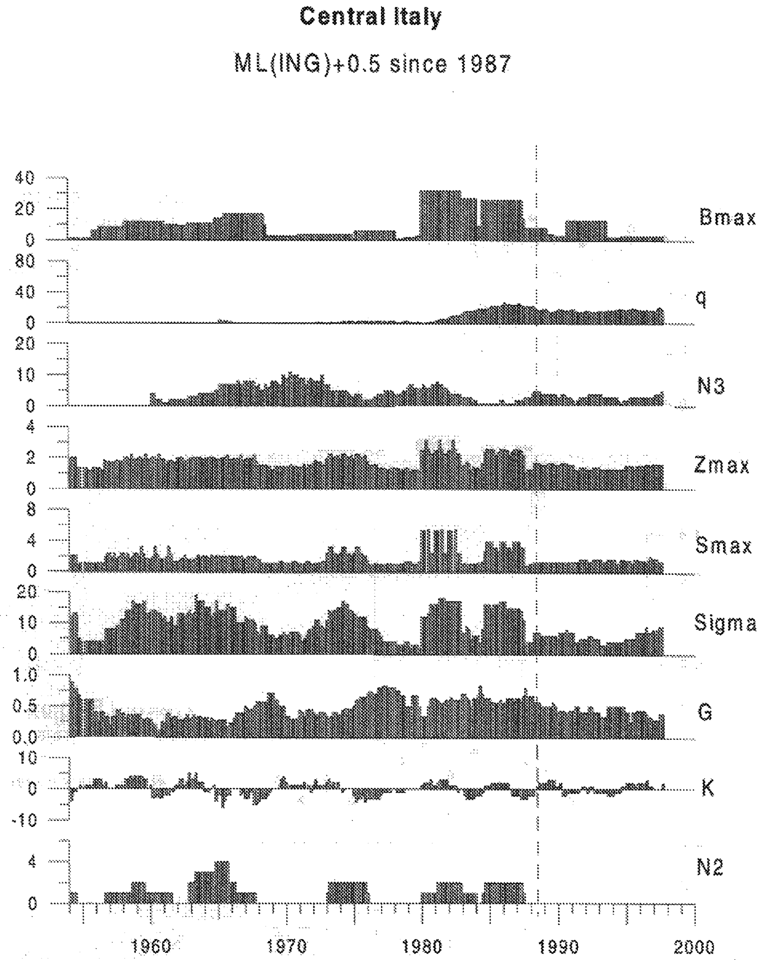


Figure 5. Time diagrams of the CN functions obtained for the Central region using the 'corrected' catalogue, in which the quantity $D=0.5$ is added to M_L(ING) beginning in 1987.

performed using all the available information (especially concerning variations in the acquisition system), not only that provided by the catalogue itself. Furthermore, a simple magnitude shift, estimated from a limited sample, cannot restore all the properties of the real seismic sequence.

Several tests performed by systematically increasing or decreasing the operating magnitude in the catalogue used for CN monitoring (Peresan & Rotwain 1998) show that the functions G, Sigma, Z_{max} and S_{max} (Table 1) are sensitive to long-lasting major magnitude underestimations of about half a magnitude unit: they became abnormally constant for relatively long periods of time, while the function q keeps very high values, but do not cause any TIP activation. On the other end, magnitude overestimations lead to unusually high values, especially for the functions N_2 and N_3, that can be used to identify and therefore discard possible TIPs declared by CN.

EXTENSION OF THE ANALYSIS TO THE WHOLE ITALIAN REGION

The magnitude differences have also been analysed within the Northern and Southern regions defined for the application of CN to the Italian territory (Peresan *et al.* 1998a). In the Northern region, the results are in very good agreement with those obtained for the Central region and, on average, an increase of +0.5 is observed for ΔM_L in 1987. The variation in reported M_L does not affect the CN functions in the Northern region as clearly as in the Central region because the Italian catalogue (Postpischl 1985) covers an area that, towards the north, follows the Italian border and consequently is incomplete for CN application. This incompleteness has been filled in by Costa *et al.* (1996) and Peresan *et al.* (1998a) with data provided by two other catalogues: ALPOR (Catalogo delle Alpi Orientali) (1987) and NEIC, thus reducing the influence of M_L(ING) in the computation of CN functions in the Northern region. The small number of common events, and hence the insufficient sample size, does not allow any conclusive analysis in the Southern region.

The analysis of the NEIC catalogue performed by Peresan & Rotwain (1998) for the Italian area showed that for the magnitudes M_d and M_L contributed to NEIC by other agencies, M_L is 10 times more frequent than M_d. From Fig. 6 it is seen that the total yearly number of common events varies quite significantly with time. The number of common events

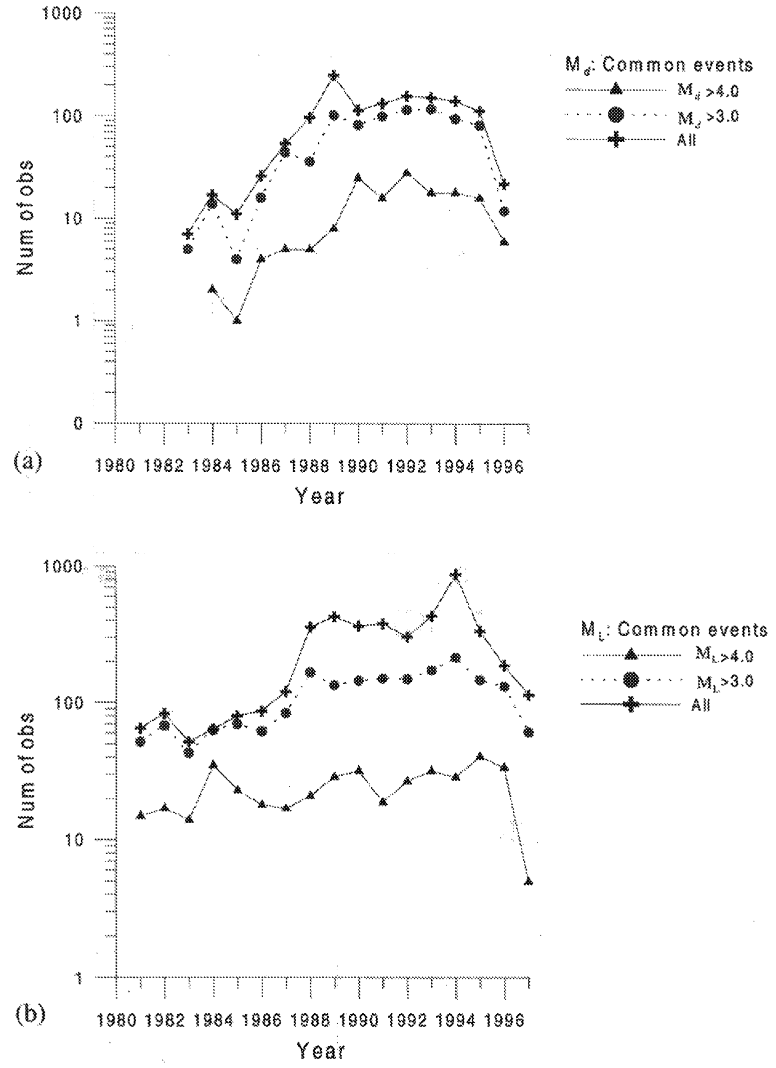


Figure 6. Yearly number of common events used for the comparison between the ING and NEIC catalogues. (a) Events used for M_d analysis; (b) events used for M_L analysis.

considerably increases after 1988, for both M_L and M_d, especially when the smaller earthquakes are considered.

The frequency distributions of ΔM_L and ΔM_d versus NEIC magnitude are analysed to evaluate their possible correlation with the earthquakes size (Fig. 7). The linear correlation between ΔM_L and M_L(NEIC) appears quite weak, while the correlation is significant for ΔM_d versus M_d(NEIC), the correlation coefficient being about 0.7 (significant at $P < 0.05$). The distributions of ΔM_L and ΔM_d are rather different, as can easily be seen from their histograms constructed for three contiguous intervals of magnitude (Fig. 8). The values of ΔM_d appear normally distributed around mean values increasing with M_d. However, the histograms of ΔM_L are centred around $\Delta M_L = 0$, with a tail towards positive values. It seems that the set of common events can be divided into two subsets: (a) events with ΔM_L distributed around zero; and (b) events with ΔM_L distributed around 0.5.

A detailed analysis, suggested by the bimodal distribution of ΔM_L, shows that the events giving $\Delta M_L \equiv 0$ are fairly localized in space (Fig. 9). The peak in the ΔM_L histograms is due to the coincidence of M_L(ING) with the M_L contributed to NEIC by some local networks, mainly from GEN (IGG network, Dipartimento Scienze della Terra, Università di Genova, Italy), LDG (Laboratoire de Detection et de Geophysique, Bruyeres-le-Chatel, France), TTG (Seismological Institute of Montenegro, Podgorica, Yugoslavia) and TRI (OGS, Osservatorio Geofisico Sperimentale, Trieste, Italy), following the standard station codes used by NEIC. Indeed, the data reported by some local networks are used by ING to integrate the information collected by the Italian network (Fig. 8).

Fig. 6 indicates that the size of the sample becomes relatively stable for magnitudes larger than 3.0, although the yearly number of common events generally increases in 1988. Hence, in this step of the analysis also, the time behaviour of the

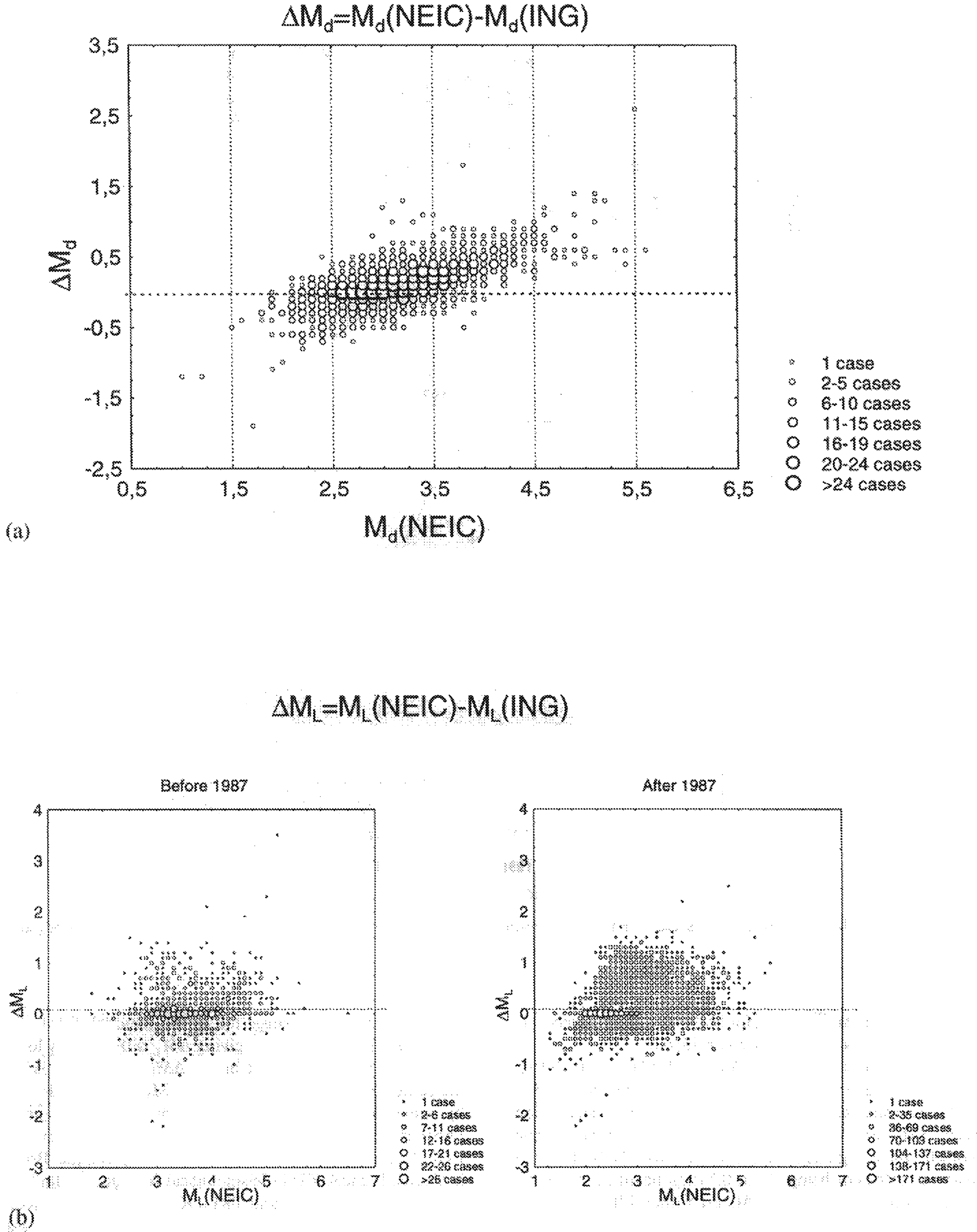


Figure 7. Frequency scatter plots of (a) ΔM_d and (b) ΔM_L versus the corresponding NEIC magnitude.

yearly average of ΔM_L and ΔM_d is evaluated using only earthquakes with NEIC magnitude larger than 3.0.

The yearly average values of ΔM_L and ΔM_d are shown in Fig. 10. The remarkable uncertainties on the average value of ΔM_L during the year 1983 and, similarly, of ΔM_d in 1985 are due to the large dispersion of the reported values rather than to the sample size. For the whole period 1983–1997, the yearly average of ΔM_d appears almost constant around a mean value

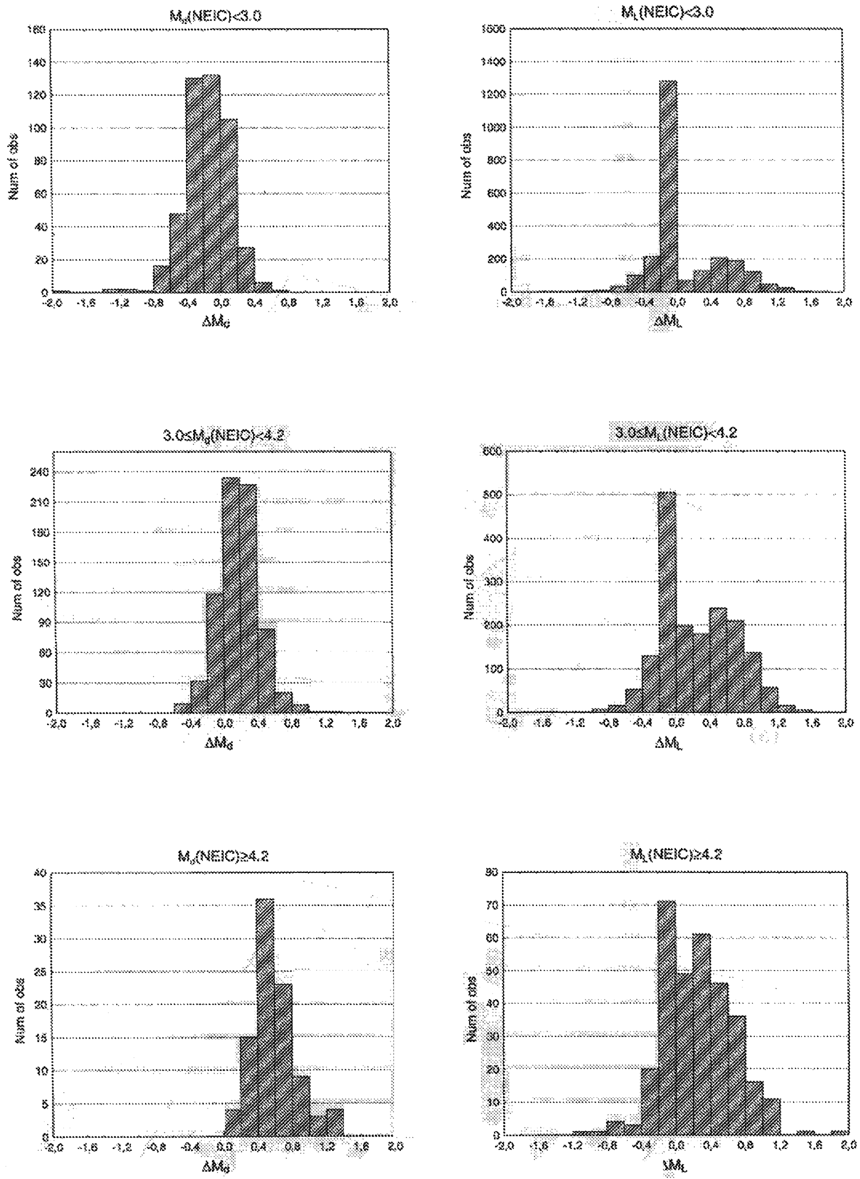


Figure 8. Histograms of the number of events versus ΔM for three contiguous ranges of magnitude for (a) ΔM_d and (b) ΔM_L. Events with ΔM lower than or equal to the upper boundary are counted in each interval.

of 0.30 ± 0.02 (Fig. 10a), in very good agreement with the results obtained for the Central region. Therefore, this analysis seems to confirm that since 1983, when they started to be reported, there have been no changes in the M_d values provided by ING. A linear relation between the M_d reported by the two agencies can be estimated by orthogonal regression of M_d(ING) versus M_d(NEIC) using the set of common events, as follows:

$$M_d(\mathrm{ING}) = 0.7M_d(\mathrm{NEIC}) + 0.8\,. \qquad (3)$$

According to this relation, the events with $M_d(\mathrm{ING}) \geq 3.0$ are on average underestimated with respect to M_d(NEIC), while smaller events are overestimated.

The diagram of the yearly average ΔM_L (Fig. 10b), however, seems to indicate the presence of two main discontinuities: the first in 1987 and the second in 1994. The average ΔM_L, estimated for the three contiguous periods of time, are as follows (the error corresponds to the 95 per cent confidence interval of the mean):

$$(1980\text{–}1986) \quad \Delta M_L = 0.08 \pm 0.05,$$
$$(1988\text{–}1993) \quad \Delta M_L = 0.30 \pm 0.04,$$
$$(1995\text{–}1997) \quad \Delta M_L = 0.77 \pm 0.06.$$

The ΔM_L increase observed during 1987 appears less relevant within the whole Italian area than for the Central region

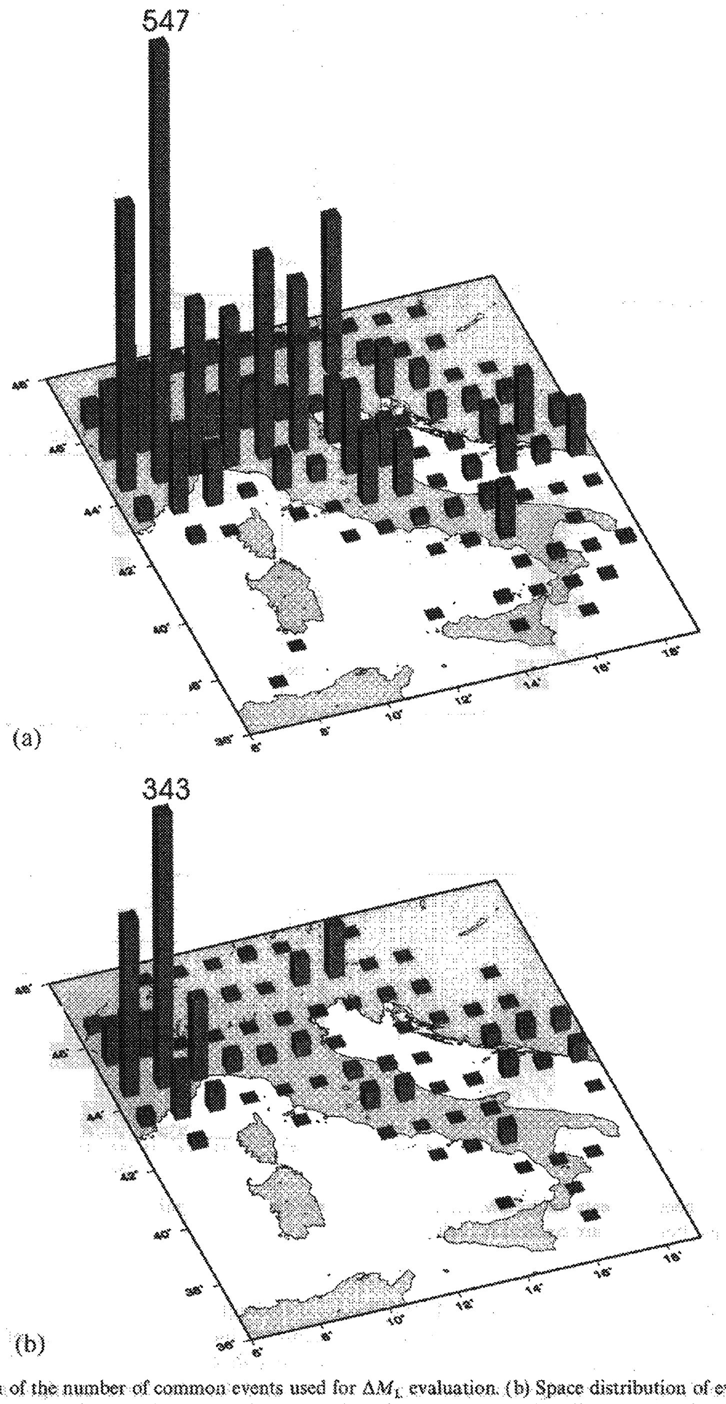
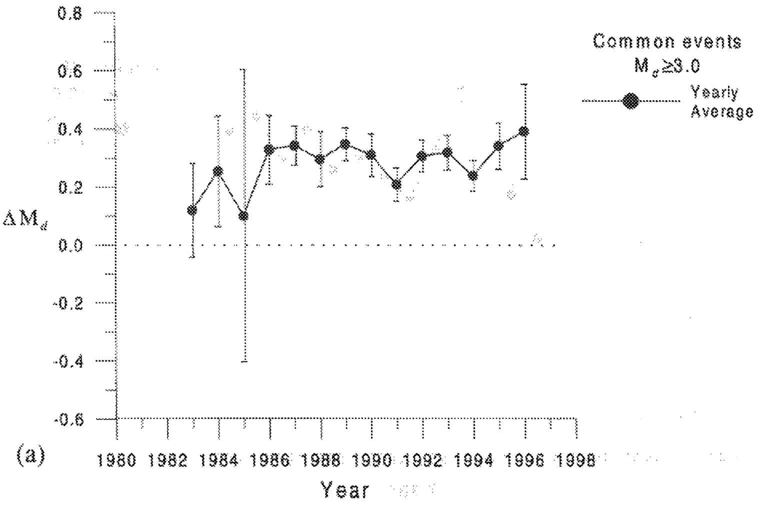


Figure 9. (a) Space histogram of the number of common events used for ΔM_L evaluation. (b) Space distribution of events with $\Delta M_L = 0$. The two histograms are plotted using the same linear scale. The maximum number of common events is indicated as a reference.

(Figs 10b and 4b). This reduction of ΔM_L can be explained by the inclusion of the M_L values contributed to both NEIC and ING by some of the neighbouring local networks, located near to the French and Slovenian borders and along the Croatian coast.

COMPARISON WITH MAGNITUDES FROM LDG BULLETINS

The use of eq. (2) for M_L reported by the catalogues ING and NEIC gives positive values for ΔM_L. To check the conclusions



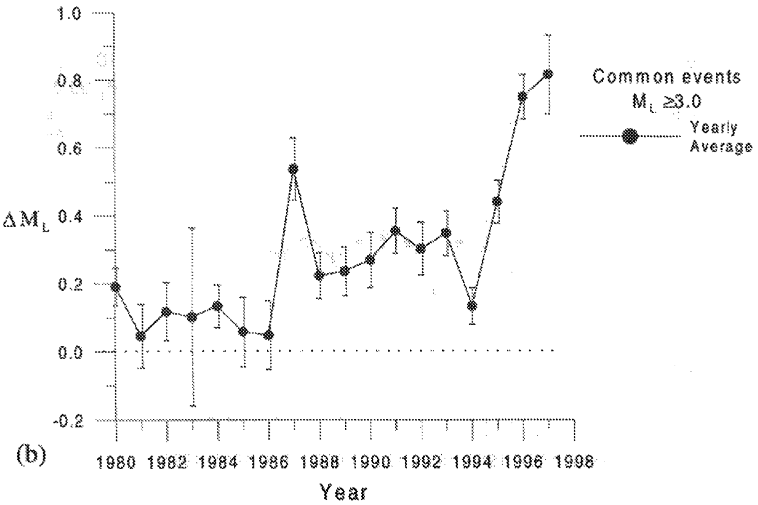


Figure 10. Yearly average of (a) ΔM_d and (b) ΔM_L for the NEIC and ING catalogues. Only events with magnitude greater than 3.0 have been considered. Error bars correspond to a 95 per cent confidence range on the calculated average. The ΔM_L minimum in 1994 is explained by the very large number of events with magnitudes coinciding with those provided by the local networks, mainly the IGG network.

drawn from the analysis of ING and NEIC data, the comparison of M_L is extended to the LDG bulletins.

The comparison between ING local magnitudes and those reported by LDG bulletins is performed within the time interval 1980–1996. About 1000 common events are selected from these regional catalogues according to the following rules: (a) time difference $\Delta t \leq 1$ min; (b) epicentral distance $\Delta\text{Lat} = \Delta\text{Lon} \leq 0.1$.

The bimodal distribution of ΔM_L observed in the comparison with the NEIC catalogue (Fig. 8) becomes even more marked when the ING and LDG magnitudes are considered. Nevertheless, most of the events with $\Delta M_L = 0$ have M_L(LDG) lower than 3.0. Hence, considering only events with magnitude larger than 3.0 allows us to exclude a large part of such events, whose magnitudes have very probably been provided by the same agency, while permitting us to keep events for which magnitude determinations can be considered quite reliable in regional catalogues.

The yearly average values of ΔM_L for the pairs of catalogues LDG–ING and NEIC–LDG have been estimated and are plotted in Fig. 11. The number of common events used for such estimations increases in time from about 10–15 events per year up to 30–40 events per year, and this is also apparent from the corresponding reduction of uncertainties. The average values obtained from eq. (2) for the pair of catalogues LDG–ING is always significantly greater than zero, even

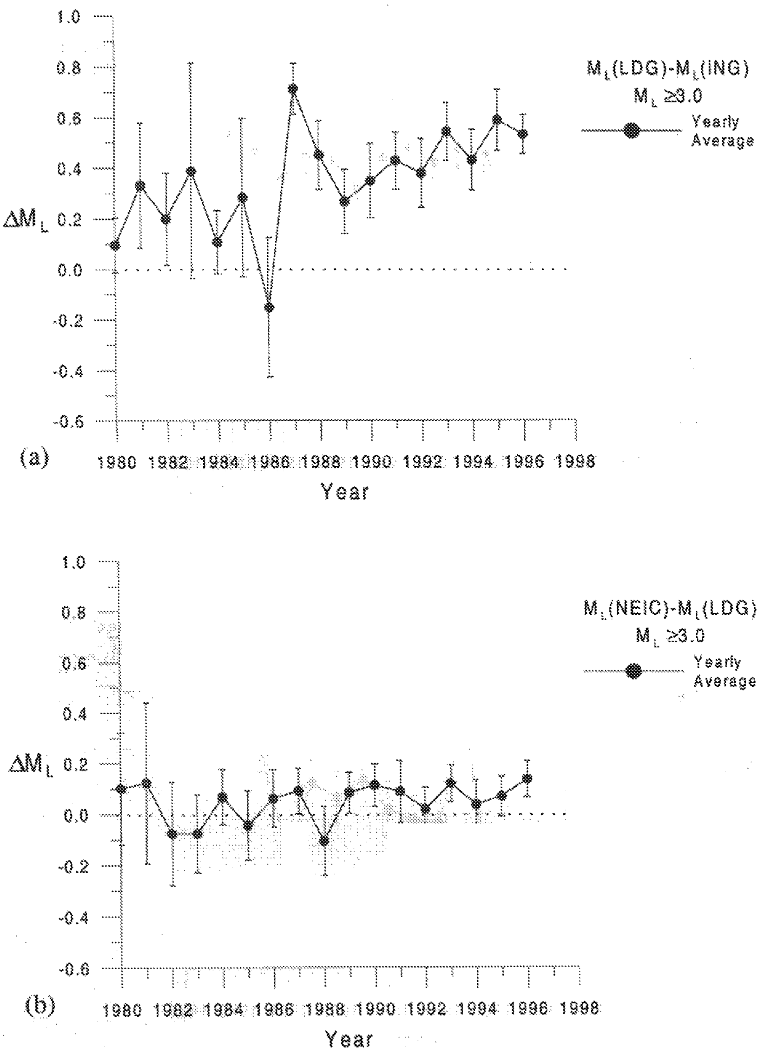


Figure 11. Yearly average of ΔM_L for (a) LDG and ING bulletins and (b) for the NEIC catalogue and LDG bulletins. Error bars indicate the 95 per cent confidence interval of the average.

with fluctuations in time (Fig. 11a). The differences ΔM_L estimated for the pair of catalogues LDG–ING and for the two intervals of time indicated in brackets give the following average values:

(1980–1986) $\Delta M_L = 0.18 \pm 0.08$,
(1988–1996) $\Delta M_L = 0.44 \pm 0.04$.

These values are in good agreement with those computed, for the whole Italian territory, comparing M_L from the NEIC and ING catalogues.

The average values ΔM_L calculated for the global catalogue NEIC and the regional bulletins LDG (about 1200 common events) are always close to zero (Fig. 11b) and, on average, are

(1980–1986) $\Delta M_L = 0.03 \pm 0.06$,
(1988–1996) $\Delta M_L = 0.08 \pm 0.03$.

This comparison seems to confirm the relative uniformity of the reference catalogues NEIC and LDG, despite the heterogeneous origin of M_L(NEIC).

A series of magnitude comparisons focused on the Central region, excluding from NEIC the events contributed by LDG or comparing directly ING and LDG, essentially confirms observations made comparing the ING and NEIC catalogues.

According to Bath (1973), we have to expect errors as large as ± 0.3 units in a calculated magnitude; nevertheless, the differences ΔM_L between the ING and the two catalogues considered have been, even after averaging, equal to or larger than $+0.3$ since 1987. Giardini *et al.* (1997) stated that local magnitudes are generally of poor quality with respect to the seismic moment, and this study indicates that they can even be inhomogeneous within the same bulletins. Unfortunately, M_L is the basic instrumental magnitude in the Italian catalogue, while M_d has only been reported since 1983.

CONCLUSIONS

Prediction methods based on seismic precursors are sometimes criticised for their sensitivity to the unavoidable catalogue errors and undeclared changes in the evaluation of the reported magnitudes (Habermann 1991; Habermann & Creamer 1994; Peresan *et al.* 1998b). This study provides a real example, showing the effect of a long-lasting systematic magnitude underestimation on CN predictions.

The absence of moderate events detected by CN functions and consequently the unusual decrease of the seismicity rate within the Central region used for the CN monitoring in Italy lead us to check for possible systematic errors in the reported magnitudes.

A detailed comparative analysis, focused on M_L and M_d, has been performed between ING and NEIC catalogues, within the area corresponding to the Central region. The magnitude differences ΔM_d appear quite stable in time and small, while a variation of about 0.5 has been found in ΔM_L, starting in 1987. This difference decreases to about 0.2 when the analysis is extended to a wider area including the whole Italian territory, but there is always an underestimation of the M_L values given by ING with respect to NEIC. The comparison extended to a third catalogue, the LDG bulletins, confirms such underestimation.

The robustness of the CN algorithm has been successfully tested with respect to the partial replacements in the catalogue, provided the homogeneity of data is preserved (Peresan & Rotwain 1998), and with respect to the short-term inadvertent increase in reported magnitude indicated by Zuniga & Wyss (1995) for the Italian catalogue, which does not seem to affect the results of predictions (Peresan *et al.* 1998a).

Therefore, our study indicates that a careful analysis of CN functions allows us to find major long-lasting undeclared changes in the reported magnitudes and may permit us to separate such effects from the anomalies in the seismic flow that define the times of increased probability (TIPs) for the occurrence of a strong event. The results of our analysis cannot be used for catalogue correction; therefore, the ING catalogue cannot be used for CN monitoring and one has to make use of a different data set such as the NEIC catalogue.

ACKNOWLEDGMENTS

We are grateful to Profs I. M. Rotwain and V. I. Kossobokov for their relevant contribution in the catalogue analyses. We wish to thank Profs G. Molchan and T. Kronrod for the useful discussions and their invaluable suggestions. This research has been supported by MURST (40 per cent and 60 per cent funds), by CNR (contracts numbers 96.02968.PF54 and 97.00507.PF54) and by INTAS (94–0232).

REFERENCES

ALPOR, 1987. *Catalogue of the Eastern Alps,* Osservatorio Geofisico Sperimentale, Trieste, Italy (computer file).

Bath, M., 1973. *Introduction to Seismology,* Birkhauser, Basel.

Costa, G., Panza, G.F. & Rotwain, I.M., 1995. Stability of premonitory seismicity pattern and intermediate-term earthquake prediction in central Italy, *Pure appl. Geophys.,* **145,** 259–275.

Costa, G., Stanishkova, I.O., Panza, G.F. & Rotwain, I.M., 1996. Seismotectonic models and CN algorithm: the case of Italy, *Pure appl. Geophys.,* **147,** 1–12.

Gabrielov, A.M. *et al.,* 1986. *Algorithms of Long-Term Earthquakes' Prediction,* International School for Research Oriented to Earthquake Prediction-algorithms, Software and Data Handling, Lima, Perù.

Giardini, D., Di Donato, M. & Boschi, E., 1997. Calibration of magnitude scales for earthquake of the Mediterranean, *J. Seism.,* **1,** 161–180.

Habermann, R.E., 1991. Seismicity rate variations and systematic changes in magnitudes in teleseismic catalogs, *Tectonophysics,* **193,** 277–289.

Habermann, R.E. & Creamer, F., 1994. Catalog errors and the M8 earthquake prediction algorithm, *Bull. seism. Soc. Am.,* **84,** 1551–1559.

Keilis-Borok, V.I., 1996. Intermediate term earthquake prediction, *Proc. Nat. Acad. Sci. USA,* **93,** 3748–3755.

Keilis-Borok, V.I. & Rotwain, I.M., 1990. Diagnosis of time of increased probability of strong earthquakes in different regions of the world: algorithm CN, *Phys. Earth planet. Inter.,* **61,** 57–72.

Keilis-Borok, V.I., Kutznetsov, I.V., Panza, G.F., Rotwain, I.M. & Costa, G., 1990. On intermediate-term earthquake prediction in Central Italy, *Pure appl. Geophys.,* **134,** 79–92.

Peresan, A., Costa, G. & Vaccari, F., 1997. CCI1996: the current catalogue of Italy, *Internal Rept IC/IR/97/9,* International Centre for Theoretical Physics, Trieste, Italy.

Peresan, A., Costa, G. & Panza, G.F., 1998a. Seismotectonic model and CN earthquake prediction in Italy, *Pure appl. Geophys.,* **154,** 281–306.

Peresan, A. & Rotwain, I.M. & Panza, G.F., 1998b. Evaluation of the stability of algorithm CN with respect to random errors in magnitude: Central Italy, *Ann. Geophys., Abstract Suppl.,* **16,** 1091.

Peresan, A. & Rotwain, I.M., 1998. Analysis and definition of magnitude selection criteria for NEIC (PDE) data, oriented to the compilation of a homogeneous updated catalogue for CN monitoring in Italy, *Internal Rept,* International Centre for Theoretical Physics, Trieste, Italy.

Postpischl, D., 1985. *Catalogo dei Terremoti Italiani dall'anno 1000 al 1980,* CNR–Progetto Finalizzato Geodinamica, Bologna, Italy.

Rotwain, I. & Novikova, O., 1999. Performance of the earthquake prediction algorithm CN in 22 regions of the world, *Phys. Earth planet. Inter.,* **111,** 207–213.

Storchak, D.A. & Bird, A.L. & Adams, R.D., 1998. *Location Discrepancies,* XXVI General Assembly of ESC, Tel Aviv, Israel.

Zuniga, F.R. & Wyss, M., 1995. Inadvertent changes in magnitude reported in earthquake catalogs: their evaluation through the b-value estimates, *Bull. seism. Soc. Am.,* **5,** 1858–1866.