1
Introduction to the *Database*

GIANLUCA VALENSISE and DANIELA PANTOSTI

Istituto Nazionale di Geofisica e Vulcanologia, Roma, Italy

1.1. Preface

This volume presents and describes Version 2.0 of the *"Database of Potential Sources for Earthquakes Larger than M 5.5 in Italy"* (also referred to as DISS, the acronym of the short form *Database of Italy's Seismogenic Sources*, or simply as *Database*) that was first conceived at the Istituto Nazionale di Geofisica e Vulcanologia of Rome (INGV; former Istituto Nazionale di Geofisica) in 1996. Aware of the potential of a such a database for seismic hazard assessment in any seismic-prone country, in 1997 we extended the Italian experience to the Southern European region by interacting with foreign scientists in the identification and detailed characterisation of faults of Greece, Spain and France. Along with several other European seismologists we launched a project denominated FAUST (FAUlts as a Seismologists' Tool), that was eventually funded by the European Union for three years (see http://faust.ingv.it/ for a summary of the project's main accomplishments and for browsing the European database). Since then the national and the European projects developed alongside and in July 2000 the structure and the data of Italian database reached a critical level that prompted its first official presentation to the Italian scientific community (20-21 July 2000, Version 1.0). A prototype of the entire *Database* and associated software was distributed on a CD-ROM within the INGV between July and December 2000. Finally, in February 2001 a new EU-funded project termed SAFE (Slow Active Faults in Europe; http://safeproject.free.fr/home.html) took on the challenge of extending some of the concepts developed for DISS to the slower but even more problematic active faults of the rest of Europe.

During over five years many workers participated in the construction of the *Database* with different roles and contributions. The first three years were largely devoted to developing the software with relatively few individuals involved, while most of the specific information was collected between 1998 and 2000 based on the work of over a dozen of *data-hunters*. The list includes geologists who contributed original data; geologists who compiled *Database* records based on existing literature; geologists and technicians that shaped up the structure of the *Database* and prepared the text and graphic material to be entered in it; and software engineers who interacted endlessly with the rest of the group to devise the fundamental structure of the *Database*, prepare and fine-tune the many tools and functions available to the users, and provide some of the non-geological data tables.

Appendix I lists an summarises all *seismogenic sources* that comprise the *Database* (see the following sections for a description of the different source types) and gives credit to all of the respective compilers. We would like to use this opportunity to thank all of them for a difficult, time-consuming, and not immediately rewarding effort. The project also benefited from computer resources and from funding from former ING and present INGV of Rome. For these opportunities we wish to thank Professor Enzo Boschi, who has encouraged and supported this project during its entire duration.

The following chapters describe the structure of the *Database* and of the specific software that governs it (Chapter 2), its main tools and commands along with practical examples (Chapter 3), and the procedures for extending and updating it (Chapter 4). In addition to Appendix I, that lists all sources contained in the present release, we supply also a list of the illustrations associated with each source (Appendix II), a list of fault maps compiled by previous workers (Appendix III) and a list of additional geological and geophysical datasets (Appendix IV). For any further inquiry the reader is encouraged to visit the *Database*'s web-page (http://www.ingv.it/ ~wwwpaleo/catalogosorgenti/) or to contact one of the authors. All throughout this presentation the reader may refer to the poster supplied with this volume. Clearly, the best way of exploring the *Database* is to browse it directly on a computer. This requires the *Database* CD-ROM and appropriate software to run it. The steps to be taken to receive one or more copies of the CD-ROM and to run the *Database* are summarised at the beginning of this volume.

1.2. State-of-the-art and rationale of the *Database*

The *Database* was designed as a permanent interface between the data providers, geologists and seismologists who identify and characterise seismogenic sources, and the final users, a vast category that includes other Earth scientists, earthquake engineers, planners and insurers. On the one hand, data providers will use the *Database* to put their observations and interpretations in the broad perspective offered by a huge amount of geological, historical and instrumental information, thus gaining confirmation of their hypotheses or suggestions on how to revise them. They will also see their results quickly incorporated and periodically released with proper referencing to potential users. On the other hand, different categories of end-users will benefit from homogeneously collected, frequently improved and periodically updated parametric and textual information on the country's main seismogenic sources. The incorporation of most current knowledge on the largest historical events in terms of earthquake sources and the simplicity of the scientific assumptions guarantee that the *Database* rests on solid foundations that are not likely to change significantly during the next few years. Nevertheless, although the *Database* comes after several years of analyses, tests and conceptual elaboration, we expect it to evolve continuously following the progress of understanding of the earthquake generation process, with special emphasis on rapid incorporation of any new insight on Italy from such diverse sources as geodetic networks, subsurface data, and geodynamic models.

The essential guidelines of DISS stem from the awareness that the identification of Italian seismogenic faults has always been difficult and controversial even in high seismicity areas such as the Central and Southern Apennines. This condition and the exceptional availability of large amount of historical data conspired in keeping the Italian geological and seismological communities isolated from each other. As a consequence, the study of seismogenic sources in Italy has been traditionally based on the analysis of felt reports of historical earthquakes. This approach provides a satisfactory mapping of point-sources, which generally reveal the main active tectonic trends, and is good at constraining the size of the largest historical earthquakes. However, no information on the physical properties of the source of a specific event can be directly deduced from this approach (*e.g.*, length, width, dip, strike, etc.), little inferences can be made concerning the location and size of future large earthquakes, and nearly nothing can be done in the way of anticipating the characteristics of ground shaking deterministically.

From a geological-geomorphic standpoint, the traditionally available compilations that could be used to describe seismogenic sources are fault maps at different scales (many of which are included in the *Database* as background information). Because these maps were not always prepared in the framework of seismic hazard assessment programs, they tend to describe the entire suite of faults and major tectonic features of the region they are concerned with. In some cases, particularly with the most recent compilations, some faults are indeed described as active, but this is often only implicitly declared and hence left ambiguous. All the faults that are implicitly or explicitly tagged as "active" are generally considered as potential earthquake sources, although the equation "*active fault equal seismogenic source*" is clearly not straightforward and normally highly controversial. In general, no relation between faults and historical earthquakes is proposed with the exception of a few known recent earthquakes (*e.g.*, Bosi, 1975; Slejko *et al.*, 1987; Ortolani and Pagliuca, 1988; Castaldini and Panizza, 1991; all reported in the *Database*). In some other cases, potential sources could be indirectly deduced from neotectonic maps that contain tectonic elements as old as 5 My (Ciaranfi *et al.*, 1983; Ambrosetti *et al.*, 1987; Bigi *et al.*, 1989; all reported in the *Database*). In all of these cases, little attention is paid to the determination of the extent of faulting length-wise, and the potential end-user is supplied with the simple symbolic representation (a line on a map) of the fault that can be directly seen in the field, or of an inferred fault obtained by joining small bits of observed tectonic features. No indication is supplied as to the downward extent of faulting (depth-wise), such that very shallow faults and major regional faults are marked using the same symbols. Finally, some faults have unusually complex shape in map view (usually, convex towards the hanging wall for normal faults), and often even the sense of dip and the kinematics of the fault are not specified.

Problems of poor representation may to some extent be the result of pre-GIS, oversimplified mapping procedures. The fundamental problem of most existing compilations, however, is the extent of the match between active (or presumed active) faults and the actual location of seismogenic sources as revealed by historical seismicity. This circumstance and the intrinsic weakness of the older efforts is now becoming widely accepted; the faults shown in the most recent maps and inventories are largely unrelated with pre-existing compilations and interpretations and explicitly regarded as potential sources of historical and pre-historical earthquakes (*e.g.*, Galadini *et al.*, 2000 and Michetti *et al.*, 2000; both reported in the *Database*). But even in the most recent efforts the extent of the overlap between reported active faults and historical earthquakes that could be associated with them is minimal (less than 30% of faults correspond with historical earthquakes, at least tentatively), and the frequency of large earthquakes that would result from these fault patterns is largely incompatible with the observed seismicity. In addition, Valensise and Pantosti (2001) have shown that the concentration of active or "neotectonic" features inferred from current fault compilations is largest in areas that experienced very little seismicity (*e.g.*, along the Tyrrhenian coast of the peninsula), and *vice versa* the areas that experienced important seismicity show the least concentration of active or "neotectonic" features (*e.g.*, the Central-Southern Apennines). The inference is that "common sense" correlations between tectonic features and seismogenic sources are very ambiguous and may have highly adverse implications, such as anticipating the presence of major seismogenic faults in areas that are mildly active or not active at all, or "missing" major sources in areas characterised by rare large earthquakes or by dominant blind faulting. Considering that this type of faulting appears as a very common occurrence in Italy (out of 46 XX century earthquakes of magnitude ≥ 5.5 listed by Boschi *et al.* (2000), unequivocal surface faulting was reported for just three events, and only a few scattered testimonies are reported for older earthquakes by Valensise and Guidoboni (2000)), and that the repeat time of large earthquakes appears to be systematically much larger than the available historical record (2000-3000 years *versus* about seven centuries according to Valensise and Pantosti (2001)), the circumstances described above should certainly not be regarded as "extreme" or "unlikely".

All in all, the investigation of seismogenic sources in Italy is a difficult task that requires a multidisciplinary approach, a homogeneous strategy of data collection and interpretation, the development of innovative ways of exploiting geological data, the enforcement of firm geometric, kinematic and geodynamic constraints, and the broadest possible appreciation for what has been done by previous workers (the *References* section of the *Database* lists over 1250 papers dealing with the recent tectonics and seismicity of Italy, from which nearly 400 images and over 40 fault maps were taken). Identifying individual earthquake sources is universally recognised as the fundamental contribution for more accurate assessment of regional seismic hazard and of the hazard associated with critical facilities, for effective urban planning and for determining realistic priorities in the enforcement of a risk mitigation plan. Our *Database* intends to create such foundations for Italy.

1.3. Beyond basic fault mapping: new strategies for old data

The *Database* fully acknowledges the difficulties and intrinsic ambiguities described in the previous section, and challenges them with a twofold strategy:

1) each of the seismogenic sources contained in the *Database* is clearly and unequivocally described by one or more compilers following a pre-defined scheme. Each source record is hence based on the personal judgement, and compiled under the responsibility, of the scientist(s) that sign the source record. In contrast, previous compilations often listed alternative (and anonymous) options for the same seismogenic source, in some instances even if geometrically or kinematically incompatible with each other, thus transferring onto the end-user all the open issues;
2) each of the seismogenic sources is backed by the raw material that was used by the compiler(s) to derive its main parameters and achieve the given conclusions, including published numerical estimates and images. The most critical parameters are keyed to the relevant paper by the logics of the *Database*, or explicitly declared as originally derived by the compiler(s). Any user may therefore trace back the compiler's reasoning and the data upon which it was based, form his/her own personal judgement and keep or modify the source parameters accordingly.

This strategy is extended to all sources, including those that were derived from earthquake intensity data exclusively (see Section 1.5.). For these sources the *Database* supplies both the raw data from the selected historical catalogue and the software used to derive their essential parameters (*Boxer* code, included in the CD-ROM under the heading *Related Software*). The approach is similar to that taken by the *Catalogue of Strong Italian Earthquakes* (Boschi *et al.*, 1995, 1997 and 2000), that is commonly referred to as a "new generation earthquake catalogue" in that it supplies the original texts and data along with synthetic comments and analytical parameters. The ultimate goal of both compilations, the *Catalogue of Strong Italian Earthquakes* and our *Database*, is to become a stable (*i.e.* printed) reference for the state of knowledge at the time of their publication and a valuable basis for further work.

One of the main characteristics of the new *Database* is the joint use of all geological and historical evidence for seismogenic faulting, with constraints obtained from instrumental seismicity and large-scale geodynamic and geomorphic features. We feel that, compared with the strategies used for previous efforts and compilations, such multi-disciplinary approach is extremely robust and truly mandatory in countries where the geologic evidence alone is ambiguous if not misleading, such as Italy. So far there have been only sporadic attempts to integrate the historical data with geological and geomorphic investigations of the surface expression of a seismogenic structure. Some of these attempts have resulted in very detailed descriptions of the earthquake source even for earthquakes of the early-instrumental era (*e.g.*, Valensise and Pantosti, 1992; Pantosti *et al.*, 1993, 1996; Michetti *et al.*, 1996; Collier *et al.*, 1998; Galadini and Galli, 1999; Pantosti *et al.*, 2001), but no integration on a regional scale is available to date except for the notable case represented by Ambraseys and Jackson (1990). Although the past 20 years have witnessed major advancements in the investigation of potential earthquake sources using geological methods, for the purpose of seismic hazard assessment most European countries continued to treat seismogenic sources as points on a map derived from historical seismicity, with very little or no input from field data.

From a methodological standpoint, the novelty represented by the *Database* can be summarised by two main points:

1) the *Database* deals explicitly with the structural complexity of the Italian peninsula, with the widespread phenomenon of blind faulting and with the ensuing difficulties in the identification of the source of large earthquakes. This problem is dealt with using two different strategies: a) through innovative schemes of analysis using non conventional evidence for the presence of active tectonic structures, such as the drainage pattern, the recent evolution of the landscape, the distribution of recent deposits; b) through a new scheme for parameterising large historical earthquakes, that uses the detailed intensity patterns offered by new generation earthquake catalogues to derive the essential spatial and geometric properties of the respective seismogenic sources;
2) the *Database* takes a clear stand in the difficult problem of relating surface tectonic structures, which are more easily accessible but often smaller and discontinuous, with the deeper (*i.e.* not directly accessible) through-going structures that are responsible for the bulk of seismic release. To deal with this potential ambiguity the *Database* proposes to use the *seismogenic source* as a fundamental physical entity and archiving unit. The seismogenic source is intended as the three-dimensional, simplified and georeferenced representation of a fault plane that is assumed to be capable of primary slip during a large earthquake. All active tectonic features lying in the vicinity of the source are described, georeferenced and linked to the relevant deep source by the logics of the *Database*, but are kept separated from it in further elaborations. The *Database* relies on the expert judgement of the compiler(s) to reduce the complexity and indeterminacy in the description of each source and to decide which surface tectonic features are or may be linked to what source.

The first of these two operating strategies allowed poorly known/understood sources to be included into a homogeneous and internally coherent tectonic scheme. Sources that were bound to remain spatially and geometrically unresolved have been assigned a tentative description, which allowed them to be immediately checked against neighbouring (possibly better identified) sources for consistency of their location, orientation and size. This is the case with the sources of various important yet poorly understood earthquakes of the past century such as the Messina Straits (1908, M_w 7.0), Garfagnana (1920, M_e 6.4), Senigallia (1930, M_e 5.9) and Irpinia (1930 M_a 6.7).

In its turn, the second strategy basically involved a reduction of the complexity of the real geological world, up to a level where all sources could be described and parameterised according to the same standard scheme. This step also contributed to reduce drastically the existing gap between the community of geologists, the geological approach and even the geological jargon on the one hand, and the world of seismology on the other hand. Seismologists are used to deal with simplified but rigorous models of the earthquake generation process, and it was indeed one of the goals of the geologists that created the *Database* to try to adjust their view of the world to that of the potential end-users. Although we must certainly wait a bit longer to see if this goal was reached, it is worth mentioning three typical seismological applications that already benefited from the existence of early versions of the *Database*:

a) studies of stress transfer and fault interaction applied to a suite of seismogenic sources of the Southern Apennines (Nostro *et al.*, 1997);
b) innovative, and hopefully more realistic, estimates of seismic hazard based on the so-called "Hybrid methods" (Peruzza *et al.*, 1997);
c) comparison between peak ground acceleration estimates obtained with conventional methods and those obtained with extended seismogenic sources such as those supplied by the *Database*, performed within the EU-funded project FAUST (Mucciarelli, 2000).

The availability of a homogeneous database of the country's seismic sources has also attracted interest from ongoing projects for improved seismic hazard assessment both at national (project

"Terremoti probabili in Italia tra l'anno 2000 e il 2030: elementi per la definizione di priorità degli interventi di riduzione del rischio sismico", coordinated by A. Amato and funded by the Italian Gruppo Nazionale per la Difesa dai Terremoti) and European level (EU-funded project SAFE, Slow Active Faults in Europe, coordinated by M. Sebrier, and the recently submitted project DASHER, Data and Analyses on Seismic Hazard in the Euro-Mediterranean Region, coordinated by M. Stucchi).

1.4. THE *SEISMOGENIC SOURCE*

As anticipated in the previous section, the fundamental entity of the *Database* is the *seismogenic source*. This choice was not meant to increase the terminology chaos that already exists in this field at the border between seismology and geology. On the contrary, ours is an attempt to bring concepts and needs together, simplify the essential definitions and create a clear and unequivocal standard for future efforts. In general, the adjectives *active*, *seismogenic*, *capable* are placed next to the term *fault* or *structure* depending on a series of reasons and applications that remains generally ambiguous to the user; in some extreme cases the same term is used to describe widely different objects (*e.g.*, Erice consensus statement, Boschi *et al.*, 1996; Machette, 2000). In addition, in most cases these terms are representative and descriptive of geologic elements that provide no objective background to describe the true physical object that is responsible for coseismic slip at depth. Faults that are commonly considered *active*, *seismogenic* or *capable* may indeed coincide with the surface expression of the seismogenic source or part of it, may be mechanically associated with it (*i.e.* a surface fault may slip sympathetically even though it does not lie on the projection of the source to the surface), or may simply be the evidence of gravitational processes induced by the coseismic strain or by the ground shaking. Finally, a specific and common problem is represented by blind sources, particularly those that have already produced a large historical earthquake and yet can not be represented through traditional geological schemes and categories.

In summary, the *seismogenic source* is a simplified, conventional three-dimensional representation of the real world

- *that is seen in action by instrumental seismologists and geodesists*;
- *whose dynamic and physical effects are later observed by historical seismologists*;
- *whose physical effects are seen/predicted by earthquake geologists*;
- *whose recurrence properties can be investigated by paleoseismologists*,

but that is also

- *easily perceived by most end-users, and even by the media and the population*;
- *easily used as input for deterministic calculations of expected ground shaking*;
- *easily incorporated into the new-generation seismic hazard schemes*;
- *suitable for assessing residual uncertainties and determining future research goals.*

The entire data bank rests on the basic assumption that each seismogenic source tends to generate repeatedly and exclusively its largest allowed earthquake, that is, the assumption of "characteristic" behaviour (in the sense of Schwartz and Coppersmith, 1984) for what concerns fault location, geometry and size. Three important physical implications of this assumption are that:

a) all sources can only repeat the same-size event, although they may trigger nearby sources dynamically (as shown by the sequence of the 1783 Calabrian earthquakes, each of which is assigned to an individual, physically separated and non-overlapping source);
b) the uncertainty in the association of a historical earthquake to a fault is entirely due to poor understanding of the sources themselves, that is to say, some sources are simply unknown. The overlap between adjacent sources, that is to say, the possibility of an earthquake rupturing a portion of a larger source, is not permitted;

c) the entire suite of seismogenic sources is controlled or "modulated" by pre-existing tectonic structures, referred to as "tectonic lineaments", which act as segment boundaries. The violation of such boundaries is seldom observed with historical earthquakes and not permitted for quiescent sources.

Although seemingly crude, these assumptions are receiving support in tectonic provinces dominated by dip-slip faulting away from the main active plate boundaries, such as Europe (*e.g.*, Roberts and Jackson, 1991, for Greece; Valensise and Pantosti, 2001, for Italy). We realise that departures from this expected behaviour are entirely possible, particularly as the earthquake magnitude decreases, but we contend that for all practical purposes our assumptions are sufficiently realistic. A typical case is represented by seismic hazard calculations based on *renewal*; these will always focus on regions that have not released a large earthquake in the past millennium before attempting to explore if there is any potential left within, or on the tails of, a historical rupture.

At any rate, to reduce the adverse impact of these assumptions we established a lower threshold of *M* 5.5 for the definition of an individual source. This threshold:

a) roughly coincides with a rupture length of about 5 km, which is close to the limit of resolution of any method of geologic investigation of earthquake faulting;
b) corresponds to the minimum magnitude for which the technique used for parameterising historical sources (*i.e.* based only on intensity data) returns meaningful results;
c) is similar to the lower magnitude limit of earthquakes contained in the *Catalogue of Strong Italian Earthquakes*;
d) is commonly considered as the minimum earthquake magnitude that may produce structural damage in standard tectonic settings.

The present version of the *Database* also allows the user to plot, view and analyse in the framework of the identified sources many smaller historical and instrumental earthquakes. For the time being these are not included in any particular category, for instance in homogeneous seismogenic zones, but dealing with background seismicity is indeed one of the goals of future efforts and of future releases of the *Database*.

Finally, each source is parameterised based on geologic observations, instrumental data and historical data (processed using the analytical approach described below), or more frequently, of all or a combination of the above. Our ultimate goal was to describe internally consistent sources that would benefit from all existing constraints both at local (terminations of adjacent sources/location of important segment boundaries) and regional scale (continuity of regional faulting trends, uplift patterns, etc.). Unknown parameters (*e.g.*, the width of the fault when the identification of the source is based only on direct surface faulting evidence) are obtained by scaling with the other parameters using appropriate empirical laws (*e.g.*, Wells and Coppersmith, 1994), unless unquestionable experimental observations are available. The reader is encouraged to refer to workbook *FaultMapper* in the *Related Software* section of the *Database* for details on how these decisions and assumptions were practically implemented.

1.5. Types of seismogenic sources

The approach to the identification and systematisation of Italian seismogenic sources is based on a fundamental subdivision between sources for which there exists substantial geological and geophysical evidence (*Geologic/Geophysical*) and sources obtained exclusively from intensity data (belonging to one of the three categories *Well Constrained, Poorly Constrained, Deep*). A full description of the structure of these sources, both from the point of view of their software structure and of their information content, is given in § 2.2.3 and in § 3.2.5. and 3.3.2., respectively. A separate section discusses the problem of source reliability and rating. Finally, a list of all sources is presented in Appendix I.

We wish to stress that the discussion presented here was deliberately reduced to the essential background information. The best way to grasp the real differences between the different categories of sources, to view them in their geologic and geodynamic context and to evaluate the material behind them is to refer to Chapter 3 and to the CD-ROM.

Sources from Geologic/Geophysical Data

These are the best quality sources with a potential for an earthquake of M 5.5 or larger. In several cases they benefit from instrumental and/or historical data. In either case, there has been geological work aimed at identifying field evidence for the given source, or at least at providing geological, tectonic or geodynamic constraints to identify and locate it with the smallest possible uncertainty. Full details on sources of this category are given in § 2.2.3.1. The full source list is given in Appendix Ia.

These sources are described tri-dimensionally by the rectangular rupture plane that best approximates the expected, observed or inferred earthquake rupture. In the cartographic interface they are represented by a rectangle, that corresponds to the surface projection of the rupture plane, and by a line drawn next to the rectangle, that corresponds to the geometric intersection of the rupture plane with the local topography. For surface-breaking faults this line will appear very close or coincident with one of the edges of the rectangle, while for blind faults there will be a gap between the line and the rectangle; the size of this gap is a function of dip and depth of the fault plane. All rectangle and line edges are georeferenced using their latitude and longitude. The inferred length, width, depth, strike, dip and kinematics (rake) of the source conclude its description.

The reliability of each source is expressed by a single *Quality* parameter that ranges from A (most reliable) to D (least reliable). The detailed description of the rating classes for this type of sources is given in § 2.2.3.1.

The recurrence properties of the source are optionally described by its slip-rate, average recurrence interval, timing of the latest and penultimate earthquakes, and time elapsed since the latest event. Based on the consistency among the observed slip rates, and on the basis of geodynamic considerations, all unknown slip rates are assumed to fall in the range 0.1 to 1.0 mm/yr. Similarly, unknown recurrence intervals are set on the basis of the slip rate and average displacement and of historical constraints. A minimum elapsed time of 700 years (presumed length of the historical record for large earthquakes) is imposed for all sources not associated with a historical earthquake.

The description of sources of this category is optionally complemented by:

a) a list of *References*, that includes all papers that mention the source specifically plus a number of papers that deal with the geology, tectonics and historical seismicity of the region surrounding it;
b) a text named *Summaries of the Main Studies*, that summarises all the relevant work carried out on the given source by previous workers and the level of knowledge achieved before the implementation of the *Database*;
c) a text named *Comments and Open Questions*, that presents the basis of some of the decisions made in the definition of the source parameters and summarises aspects of the source that still need to be investigated or await confirmation.

Sources from Historical Data, Well Constrained

These sources were obtained exclusively from the intensity data of an individual large earthquake ($M \geq 5.5$) through the *Boxer* code (Gasperini *et al.*, 1999). This code is available for download in the *Related Software* section of the *Database*, along with a manual and sample data that illustrate how it works. For some sources of this category the compilers of the *Database* brought together support information (essentially references and pictures) and prepared *Summaries of the Main Studies* and *Comments and Open Questions* files, similarly to what is done for sources based on geologic and geophysical information. The additional information, however, was not enough to allow the compiler to turn these sources into *Geologic/Geophysical* sources, and for this reason they are referred to as *Historical - Well Constrained with Geological Background*. Full details on sources of this group are given in § 2.2.3.2. and 2.2.3.3. The full source list is given in Appendix Ib.

The geometric representation of these sources is pseudo-tridimensional. The *Boxer* code determines the size and orientation of the source, but does not provide any hint on its dip direction, dip angle, depth and kinematics. Hence in the cartographic interface each of these sources is still represented with a rectangle, that is intended to mark the surface projection of the rupture plane, but no indication is given concerning the intersection of this plane with the surface. For a more realistic representation, the width of the rectangle is shown as if the rupture plane had a dip of 45°. All rectangle edges are georeferenced using their latitude and longitude. The inferred length and width of the source conclude its description.

The reliability of each source is expressed by two *Quality* parameters that range from A (most reliable) to E (least reliable). The detailed description of the rating classes for this type of sources is given in § 2.2.3.2.

The recurrence properties of a source derived from intensity data can obviously be represented only by the time elapsed since the latest event.

The description of sources of this category is optionally complemented by:

a) a list of *References,* that includes the intensity databases from which the intensity data were obtained and any additional papers dealing with the geology, tectonics and historical seismicity of region surrounding the source;
b) texts named *Summaries of the Main Studies* and *Comments and Open Questions*, that describe the source of the data used in the elaboration, decisions made in the selection of the input data and other relevant pieces of information. These texts are substantially more extended for sources *with Geological Background.*

Sources from Historical Data, Poorly Constrained

These sources were also obtained exclusively from the intensity data of an individual large earthquake ($M \geq 5.5$) through the *Boxer* code (Gasperini *et al.*, 1999). Similarly to the previous case, this category includes a limited number of sources that are referred to as *Historical - Poorly Constrained with Geological Background*. Full details on sources of this group are given in § 2.2.3.4. and 2.2.3.5. The full source list is given in Appendix Ic.

The geometric representation of these sources is necessarily pseudo-bidimensional. Due to the lower quality of the available input data, for a source of this category the *Boxer* code determines only the size. The cartographic interface represents it as a circle having the diameter equal to the source length determined for the two previous source types. The circle is intended to represent the envelope of an oriented source for which any given orientation is equally probable *(e.g.*, the envelope formed by a box rotating about a vertical axis). The location of the source is expressed by the latitude and longitude of its centre (coincident with the earthquake epicentre).

The reliability is expressed by two *Quality* parameters. The first is always equal to E, while the second ranges from A (most reliable) to E (least reliable). The detailed description of the rating classes for this type of sources is given in § 2.2.3.2.

The description of the recurrence properties and the additional information supplied for sources of this category are similar to those seen for sources of the previous category.

Sources from Historical Data, Deep

These sources are entirely similar to those of the previous category, but for them the compiler(s) hypothesised a depth larger than ordinary (usually below 10 km). Full details on sources of this category are given in § 2.2.3.6., while their full list is given in Appendix Id.

The geometric representation of these sources is necessarily pseudo-bidimensional. Due to the opinion that the source is deeper than ordinary and to the known difficulty in determining the correct orientation of a deeper source from intensity data, for sources of this category the *Boxer* code determines only the size. The cartographic interface represents them with hexagons having the diameter equal to the source length determined for oriented sources. The location of the source is expressed by the latitude and longitude of its centre (coincident with the earthquake epicentre). Notice that sources are assigned to this category regardless of whether the formal uncertainties in the solution obtained would have allowed to derive a rectangular source or simply a circular source.

Integrated Source Dataset

The *Integrated Source Dataset* represents the fundamental output of the *Database* and was expressly designed as the most practical interface with the end-users (see enclosed poster). This dataset contains the "preferred" parameters for those sources for which more than one solution are available (*e.g.*, the source of a large recent earthquake for which both geologic/geophysical and historical information is available). A flag ("T" or "F") that appears at the far right of Appendix Ib and Ic indicates whether the given intensity-based source is or is not included in this dataset. In contrast, all *Geologic/Geophysical* and *Deep* sources are automatically included in the dataset.

Source rating criteria

The rating of the reliability of the inferred seismogenic sources is an important step of the entire procedure of construction of the *Database*. This parameter is fundamental for driving decisions when more than one source exist in a given area of interest, or when there exist competing options for the source associated with a large historical earthquake.

The rating of *Geologic/Geophysical* sources is based on expert judgement by the compiler of each individual source. The process of rating is made very complex by the large number and the diverse nature of uncertainty types involved in the identification and characterisation of any individual source. For example, some of the sources carry an *epistemic uncertainty* associated with their mere existence, or with their attitude towards generating their maximum-size earthquake *versus* releasing multiple smaller events. These types of non-conventional uncertainties are difficult to be incorporated into a single rating parameter and for the time being this task has been entirely left to the expert judgement of the individual compilers. A more accurate incorporation of all types of uncertainties into the source rating is indeed one of the goals of future releases of the *Database*. Nevertheless, additional information on the reliability of the proposed source parameters is also generally contained in the *Comments and Open Questions* text associated with these sources.

We wish to remark that the greater reliability of sources rated A and B demands that the remaining two classes be given special attention; on the one hand by the compilers of future releases of the *Database*, who should do their best to improve the present level of knowledge; on the other hand by the end-users, who must take special care in assessing the impact of the existing uncertainties on their estimates.

The rating of all intensity-based sources is expressed by two parameters based exclusively on the quality of the solution supplied by *Boxer*. Although the two parameters are assigned automatically, important intensity-based sources are looked at in detail, are often turned into *Geologic/Geophysical* sources and subsequently rated by expert judgement. The end-user should also consider that by definition all intensity-based sources were active historically and therefore are not regarded as immediate candidates for large future earthquakes. The importance of reliable historical sources in conjunction with well constrained *Geologic/Geophysical* sources rests essentially in their capacity to delineate homogeneous tectonic trends and highlight areas of historical quiescence.

1.6. Tectonic and geodynamic insight gained from the *Database*

By providing a homogeneous portrait of the distribution of seismogenic sources of Italy, the *Database* may also form the basis for a new and updated understanding of Italy's ongoing tectonic processes. The new insight has clear implications for seismic hazard assessment, but may also contribute to a better understanding of the geodynamics of the entire peninsula and of the Central Mediterranean. Here we list very briefly a few major findings that we feel deserve special attention and that could become the object of specific papers in the near future. For a better understanding of what is being described the reader should refer to the poster or directly to the CD-ROM.

1) The first insight that can be derived by viewing through the GIS cartographic interface the *Integrated Source Dataset* (see poster), or simply the sources derived from geological and geophysical data, is the regularity in the patterns of present tectonic/coseismic deformation. The seismic moment is mainly released along three principal segmented seismogenic zones. Each of these zones is organised as a series of aligned individual sources. The main of these zones straddles the crest of the Apennines range, more or less following the regional drainage divide between Lunigiana and the Messina Straits (see inset in the lower middle of the poster), and is characterised by prevalent normal faulting and large seismic moment release, especially in its central and southern part. The other two zones are located along the Northern Apennines margin between Ancona and Piacenza, and along the eastern part of the Alpine piedmont, and both accommodate range-perpendicular compression. The three trends are somehow delineated also by the distribution of the largest historical earthquakes, with the notable exceptions of the foreland areas of the Gargano Promontory and the Hyblean Plateau and of the areas of recent volcanism (see inset in the lower left corner of the poster).
2) In addition to the kinematics and the general structural setting, sources belonging to the same zone share many essential characteristics: the source size, the maximum depth of faulting, the slip rate and the average recurrence time, all appear to be consistent along each of the three zones described above.
3) Preexisting tectonic lineaments play a definite role in controlling the segmentation of the seismogenic zones (see inset in the lower middle of the poster). Transverse lineaments seem to delineate a fundamental structural fabric that prevents very long earthquake ruptures: source length never exceeds 40 km and averages 20 km all along the Apennines, *versus* 10-15 km along the other zones. This observation alone calls for maximum expected magnitudes not larger than 7.0 in the Central and Southern Apennines, 6.5 along the Alpine piedmont and 6.0 along the Northern Apennines margin.
4) Although the Apennines seismogenic zone releases most of the seismic moment, observed slip rates are generally below 1 mm/yr. The Fucino Basin and the Messina Straits sources slightly exceed this value. Because of the consistency among the observed slip rates, and on the basis of geodynamic considerations, we assumed that all unknown slip rates lie in the range 0.1 to 1.0 mm/yr.

5) Directly estimated recurrence times are systematically longer than a millennium. This finding stresses the limitations of a description of the country's seismicity based on the instrumental/historical record only. The *Database* brings together different types of information and thus the recurrence interval, when the age of paleoearthquakes is unknown, is set on the basis of other lines of evidence and integrated with historical constraints.
6) Finally, a major insight that may have a critical impact on seismic hazard assessment and planning of future research and civil defence efforts is the identification of open gaps along all three seismogenic zones. Some of the gaps may be completely unexplored or unaffected by historical seismicity, while others may be the locus of clearly established sources that have not ruptured historically but for which the minimum elapsed time may be approaching the estimated recurrence interval. The *Database* assumes a quasi-periodic recurrence model along with a characteristic earthquake behaviour (*sensu* Schwartz and Coppersmith, 1984), which together predict that: a) a source normally ruptures for its entire extent preserving its boundaries and rupture parameters for several seismic cycles; b) the occurrence of the earthquakes it generates is quasi-periodic through time. Based on these assumptions, we consider these gaps as having the highest probability of experiencing a large earthquake in the future. On the one hand, some sources have been clearly recognised as *potential seismic gaps* because of their seismic history and behaviour (they are shown in the inset in the upper right corner of the poster). On the other hand, we realise that the *Database* may have missed some significant sources because they were silent during historical times and at the same time they could not be identified with geological tools due to their limited geomorphic evidence. These unknown sources would belong too to the *potential seismic gap* category, but in this case we can only locate suspect areas without being able to estimate the parameters of the expected earthquakes. One of these areas coincides with the Tuscan Apennines, that appears as a major gap in the Apennines seismogenic belt (see poster).