The Active Faults of Eurasia Database (AFEAD): the ontology and design behind the continental-scale dataset

Egor Zelenin[1], **Dmitry Bachmanov**[1], **Sofya Garipova**[1], **Vladimir Trifonov**[1], **and Andrey Kozhurin**[2,1]

[1]Laboratory of Neotectonics and Recent Geodynamics, Geological Institute, Russian Academy of Sciences, Moscow, 119017, Russia
[2]Laboratory of Active Tectonics and Paleoseismology, Institute of Volcanology and Seismology, Petropavlovsk-Kamchatsky, 683006, Russia

Correspondence: Egor Zelenin (egor.zelenin@ginras.ru)

Abstract. Active faults are those faults on which movement is possible in the future. This draws particular attention to active faults in geodynamic studies and seismic hazard assessment. Here, we present a high-detail continental-scale geodatabase: The Active Faults of Eurasia Database (AFEAD). It comprises 48 205 objects stored in shapefile format with spatial detail sufficient for a 1 : 1 000 000 map scale. The fault sense, a rank of confidence in activity, a rank of slip rate, and a reference to source publications are provided for each database entry. Where possible, this information is supplemented by a fault name, fault zone name, abbreviated fault parameters (e.g., slip rate, age of the last motion, and total offset), and text information from the sources. The database was collected from 612 published sources, including regional maps, databases, and research papers.

AFEAD facilitates a spatial search for local studies. It provides sufficient detail for planning a study of a particular fault system and guides deeper bibliographical investigations. This scenario is particularly significant for vast central and northern Asian areas, where most studies are available only in Russian and hard copy. Moreover, the database model provides the basis for regional- and continental-scale integrative studies based on geographic information systems (GISs).

The database is available at https://doi.org/10.13140/RG.2.2.25509.58084 (Bachmanov et al., 2022) and via web map at http://neotec.ginras.ru/index/mapbox/database_map.html (last access: 5 May 2022). Database representations and supplementary data are hosted at http://neotec.ginras.ru/index/english/database_eng.html (last access: 5 May 2022).

1 Introduction

The concept of an "active fault" emerged in order to distinguish a specific group of faults with present tectonic movement and, hence, anticipated activity in the nearest future. The term active fault and its synonym "living fault" were introduced in the late 1940s to 1950s by both American and European authors (e.g., Wallace, 1949, and the topic issue of *Geologische Rundschau*, 1955). This group of faults has a particular significance in two aspects of geological studies. First, the slip at the fault produces an earthquake; therefore, active faults are a crucial component of seismic hazard assessment (e.g., Ulomov et al., 1993; Basili et al., 2013; Walker et al., 2021). Second, active faulting occurs simultaneously across the Earth and, thus, provides a basis for studies of recent geodynamics (e.g., Rukieh et al., 2005; Schellart and Lister, 2005; Kozhurin and Zelenin, 2017).

The first global-scale inventory of active faults was the Project II-2 World Map of Major Active Faults of the International Lithosphere Program (ILP), which was initiated in 1989 and included in the ILP Global Seismic Hazard Assessment Program in 1993. Project II-2 joined more than 70 sci-

entists from 50 countries lead by two representatives from the Eastern (Vladimir G. Trifonov) and Western (Michael S. Machette) hemispheres. The Geological Institute of the Russian Academy of Sciences, Moscow, hosted the data on active faults of Eurasia provided by project members. These source data were obtained at different scales and in various formats (maps, tables, descriptions, and papers); thus, the first database (DB96) of active faults of Eurasia (Ioffe et al., 1993; Ioffe and Kozhurin, 1996; Trifonov, 1997) was intended to store digitized fault geometry at a uniform 1 : 5 000 000 scale and with a unified set of attributes. Recent advances in tectonics and information technology have highlighted the limitations of DB96: the outdated database scheme became incompatible with modern geographic information system (GIS) architectures, the fault locations lacked accuracy, and many recently studied faults were to be incorporated.

All of these issues required the creation of a conceptually new database, and this work was initiated based on DB96 some 15 years ago (Bachmanov et al., 2017). The authors have designed a new database and GIS for data processing that inherited all of the strengths of DB96 but provided far more opportunities for parameterization and analysis. The result of this work is the Active Faults of Eurasia Database (AFEAD), which is presented in this paper (Fig. 1) and distributed as a shapefile (Bachmanov et al., 2022). During the last decade, active-fault databases have been published for some countries (e.g., Atanackov et al., 2021; Ganas, 2021; Jomard et al., 2017) and subcontinental regions (e.g., Danciu et al., 2018, for the Middle East; Mohadjer et al., 2016, for Central Asia). A global collection of active faults was also compiled within the Global Earthquake Model project (Christophersen et al., 2015; Styron and Pagani, 2020). However, none of the published datasets cover the entire Eurasian area with uniform detail, and a great lack of data exists for areas of the former Union of Soviet Socialist Republics (USSR). AFEAD addresses this lack, as it includes a number of sources published in Russian or hard copy only (or both) – ca. 350 publications in total.

AFEAD provides an active-fault pattern for the continent of Eurasia and its adjacent waterbodies north of 20° S and between 30° W and 180° E. The spatial detail of AFEAD is uniform and equal to a 1 : 1 000 000 hard copy map across the entire dataset in order to maintain a balance between a large amount of data for well-studied regions (such as the Mediterranean) and scarce data for less-studied areas (such as northeastern Siberia). There may be a significant time lag between the release of recent publications and their inclusion into AFEAD, and the reader is advised to support AFEAD data with a query to academic databases.

2   The concept of an active fault

Each object of the database represents an active fault. The meaning of this term varies significantly among studies;

therefore, we consider it crucial to discuss what kind of data comprise the AFEAD.

From the most general approach, active faults are those faults on which movement is possible in the future (see the discussion in Slemmons and dePolo, 1986). Movements at the fault are typically intermittent, with strong earthquakes and long quiescence between them. The repose period is generally much longer than a human's life; hence, sole present-day observations cannot resolve uncertainty in fault activity, making it necessary to study the geological history of the fault.

The expectancy of future movement is what makes a fault active. All other fault parameters (such as kinematics, an average rate of movement, or any other) are intrinsic to faults in general, regardless of their age, and cannot affect the problem of activity. The crucial question is "What evidence could provide the basis for expecting future movements?". This question has traditionally been solved via the determination of a "critical" time span back from the present during which time at least one fault movement could have occurred. It is assumed that the fault should be considered active if such a movement has occurred. Estimations of the critical period were summarized by Galadini et al. (2012), and all of the cited intervals fall within the range of 10 kyr to 1 Myr. However, even longer intervals have been discussed (e.g., 2.6 Ma by Atanackov et al., 2021).

The concept of a critical interval implies that a slip may occur at an active fault after long quiescence. Paleoseismological studies, however, provide recurrence periods of hundreds to thousands of years, with maximum values of the order of tens of thousands of years. For example, Umehara Fault in Japan has a recurrence interval of 14–15 kyr (Kumamoto, 1998). Indeed, further studies could reveal longer intervals, but even doubling or tripling this value limits the quiescence period within the Holocene and the Late Pleistocene. Even in slowly deforming continental regions, substantiated recurrence intervals fall within 20–30 kyr (Bollinger et al., 2021). This estimate is close to that of Trifonov and Machette (1993) for the World Map of Major Active Faults. Late Quaternary deformations remain on the Earth's surface and could be unambiguously distinguished when found at late Quaternary landforms. Therefore, remote sensing interpretation of recent landforms is sufficient for regional-scale mapping or paleoseismological fieldwork planning.

Identifying and mapping active faults must precede any detailed research, and many studies report results of sole remote sensing interpretation. We account for this kind of information after the verification following the guidelines discussed by Trifonov and Kozhurin (2010): a studied object is unlikely to be created by sole non-tectonic surface processes; it offsets landforms in the same direction, which is consistent with the fault pattern of the region, and the offset is larger on older landforms.

In the database, authors make efforts to keep a balance between a unified representation of data and the intention
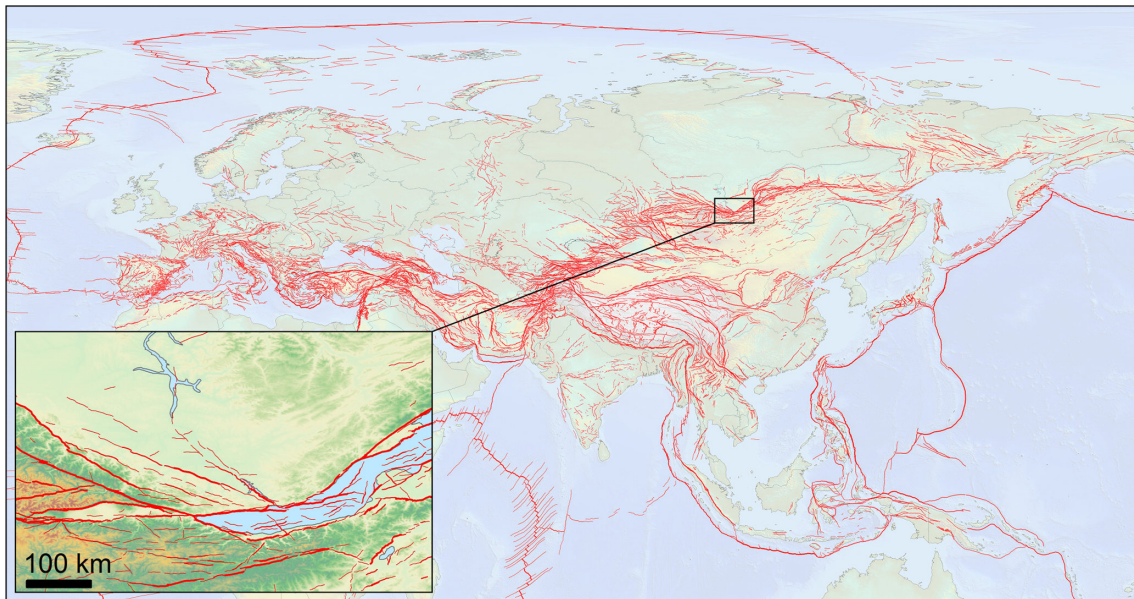
Figure 1. Overview and detail of the Active Faults of Eurasia Database (AFEAD).

to store all available data. Therefore, we propose a database model capable of indicating cases with significantly different study approaches.

3 Database model

The Active Faults of Eurasia Database is a geodatabase in shapefile format, which is an open standard for geospatial vector data. It stores the spatial location of fault lines on the Earth's surface and their attributes (Fig. 2). Within the attribute table, the field order streamlines the workflow of the database population; thus, the fields organically form two groups. The first group contains fault parameters transferred "as is" from sources and is accompanied by a complete bibliographic reference. Unification was applied to these data to ensure the uniform glossary, spelling of fault names, and spatial detail. The attributes of this group have a broad domain of values that may hamper queries but allow for textual analysis. The second group of attributes, generated by the database owners, provide consistent nomenclature to simplify GIS processing and querying.

3.1 Geometry

Every AFEAD record has a two-dimensional linear shape stored as a polyline. In most cases, it represents a fault line crossing the Earth's surface and traced by scarp or a linear deformation of landforms, whereas some objects represent the presumed intersection of a fault plane with the Earth's surface when an active fault was revealed by geophysical or seismological studies below a fold at the surface. The spatial data populated from sources are processed to comply with the topography, the database model, and the target map with a 1 : 1 000 000 scale. Usually, an active fault deforms the Earth's surface and may be traced on satellite imagery and digital elevation models. We utilized the Shuttle Radar Topography Mission (SRTM) global digital elevation model and Landsat 7 Enhanced Thematic Mapper Plus (ETM+) imagery for georeferencing and spatial adjustment of published maps. In Soviet publications, the geographic location of studied objects was classified; therefore, the processing of such data required deeper investigations and sometimes even the reinterpretation of remote sensing data.

To provide an accurate presentation at the target scale of 1 : 1 000 000, all of the objects in the database were redrawn at a scale of 1 : 500 000 with uniform spacing between vertices whenever possible, disregarding geometry accuracy in source data. Most of the fault lines bear attributes that vary along them. In such cases, fault lines were split at the point where any attribute changes its value.

3.2 Attributes transferred from source data

Collected data are stored in five fields: AUTH, FAULT_NAME, ZONE_NAME, PARM, and TEXT. Their order provides a workflow for data population, starting from identifying references and ending with auxiliary text data. This group of fields is supposed to store as much relevant data as possible and retain authors' interpretation. Still, some unification is applied to ensure the consistency of definitions and naming. If studies provide contradictory data, all of the provided parameters are recorded with a relevant citation. The low formalized string fields of this group are designed to store different estimates of the same parameter
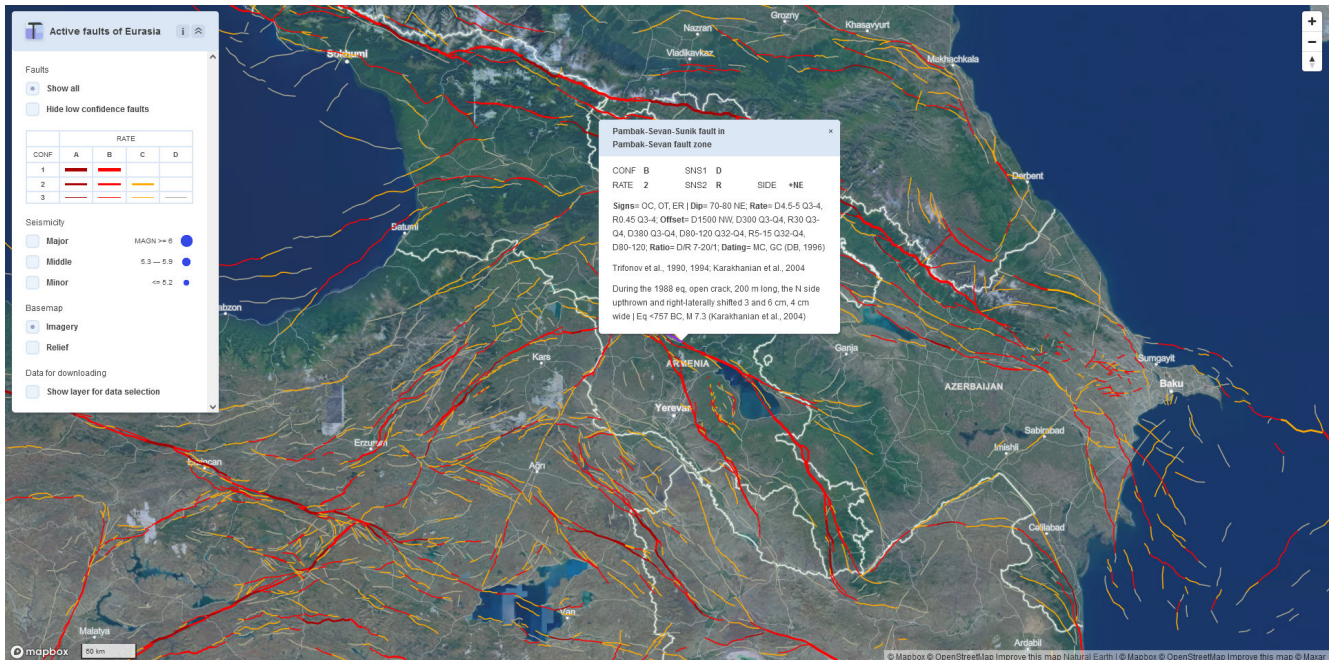
Figure 2. Spatial pattern of AFEAD faults in Transcaucasia with an attribute table for an object within the Pambak–Sevan fault. Snapshot of the AFEAD web interface at http://neotec.ginras.ru/index/mapbox/database_map.html (last access: 5 May 2022).

and relevant references; any normalized representation of these parameters would either eliminate the context or increase the complexity of the database dramatically. As a shapefile may store up to 254 symbols in a field, exceeding the field length is resolved by omitting the least relevant data. As the present-day active-fault studies provide rare cases of abundant yet different estimates of fault parameters, the length limit does not hamper the use of the shapefile standard. The abovementioned five fields are briefly outlined in the following:

– *AUTH* is a non-nullable field that stores brief English references to the studies that consider a specific object. The field format complies with the reference list (see Sect. 3.4) for further bibliographic retrieval.

– *FAULT_NAME* and *ZONE_NAME* store proper names of a fault and a fault zone, respectively, if available. In most cases, a fault zone includes a group of faults, either named or nameless. The name of a zone may originate from the name of its main fault, and an object with equal values of FAULT_NAME and ZONE_NAME should be interpreted as a main fault within a zone. If a fault has an ambiguous attribution to several fault zones, all of them are listed in ZONE_NAME (separated using a comma). Uniform name spelling and designation of objects to a zone or a single fault are maintained in the database, even if they vary among studies.

– *PARM* is a formal description of fault parameters. It has a dictionary-like structure (Fig. 2; Appendix A)

with mnemonic keys ending with "=" and a comma-separated list of values after it. A sequence of keys from a single source is separated by semicolon, whereas different authorship is indicated with a vertical bar "|" and provided with the reference. For brevity, units of measurement are omitted, with length reported in meters (m), depth reported in kilometers (km), rates reported in millimeters per year (mm yr^{-1}), age reported in years BP or units of the geological timescale (e.g., "N1" for the Miocene or "Q4" for the Holocene), and location reported in decimal degrees (°), unless otherwise explicitly stated. Acronyms are applied for directions and sense of slip (Table 1; Appendix A). "Signs=", and "Dating=" have coded values, with the complete form provided in Appendix A. These keys originate from DB96 attributes (Trifonov and Machette, 1993), although they are concatenated into a single synthetic field. DB96 had an excessive number of null values and could not handle different estimates of the same parameter. Therefore, a single denormalized field is more effective than the prior structure.

– *TEXT* contains free-form comments on the fault characteristics and supplements other fields of this group.

3.3 Derivative attributes

The derivative attributes are those produced by the database owners based on collected data. The data domain is defined

Table 1. Acronyms for fault sense used in the database.

N	Normal fault
R	Reverse fault
T	Thrust fault (reverse fault dipping $<45°$)
D	Dextral (right lateral) strike-slip fault
S	Sinistral (left lateral) strike-slip fault
E	Extensional fracture
V	Sense unknown, vertical offset
U	Slip unknown, no vertical offset

for the fields of this group (Table 2), thereby providing a basis for classification and spatial analysis. All of them, except for SENS2 and SIDE, are non-nullable. The respective fields are briefly outlined in the following:

- *CONF* indicates a level of confidence that a particular object once identified as an active fault meets the conventional definition of an active fault. Its variations represent the fact that the definition of an active fault varies between studies depending on their objectives and the reasoning behind active-fault mapping. In AFEAD, "confidence" is a qualitative measure of the expectance that an independent researcher would support the hypothesis of fault activity considering evidence from sources provided in the AUTH field. We are aware that it is a highly ambiguous criterion, although still being a crucial attribute of an active fault, as this concept varies significantly in meaning among studies. Confidence in activity is unrelated to age or level of source publication and cannot be technically derived from it. A brief outline of the guidelines proposed for CONF value assignment is now given. "A" represents that a fault has been proven active by either a series of published historical and instrumental evidence or paleoseismological studies and that multiple crustal earthquakes occurred at the fault line consistent with the fault sense. "B" denotes unambiguous surface deformations, although no recent slips have been described yet, and the attribution of earthquakes is questionable. "C" represents the fact that few pieces of evidence of activity are known as well as the lack of seismicity, the lack of surface deformations, or both. "D" represents a fault that was once declared active, but the evidence provided for such opinion is insufficient or even absent. CONF is reconsidered if new data are obtained for the studied object as well as if new data for neighboring faults support or doubt the activity of the studied object.

- *RATE* is a rank of deformation rate with estimated boundary values of 1 and 5 mm yr^{-1}, which may not be considered as actual measured slip rates. The slip rate is unlikely to be justified well at low-confidence faults so that RATE is meaningless for such faults. Therefore, boundary values of RATE have been chosen to primar-

ily affect high-confidence faults with minimal rate estimation error. However, measurement techniques and their accuracy may vary greatly, even in well-studied regions, thereby making RATE a qualitative indicator that is acceptable for regional-scale visualization and spatial analysis. RATE source data are point measurements stored in the PARM field of the relevant fault segment; thus, RATE is propagated along the fault line considering topography. Even though this approach implies ambiguous rank assignment, it is the most effective way of representing rather scarce data on deformation rates.

- *SENS1* and *SENS2* represent sense slip motion by its respective primary and secondary components (Table 1); SENS2 has values only for faults with identified oblique slip. Objects with poorly constrained slip are identified as SENS1 = "U" (for "unknown") or SENS1 = "V" (for "vertical") where vertical offset is evident along the scarp.

- *UPSIDE* is a direction of an upthrown side of the fault and is recorded as a cardinal direction.

3.4 Reference list

Most of the entries in the database have been collected from published studies, and they comprise 621 references and 63 unpublished pieces of data. The reference list represents the database sources as a tab-separated text file with citations equal to the domain of the "AUTH" field and a complete bibliographic reference to the source publication in English as well as in the language of publication. The reference list is supplemented with a many-to-many relational table that connects FAULT_ID with primary keys of the bibliographic reference table, represented in a "citations" field. AFEAD references the initial studies, even though they may be considered outdated, and the studies that provided significant contributions to each object. For data compiled for the World Map of Major Active Faults (Trifonov and Machette, 1993; Trifonov, 2004) and not previously published, the reference is supplemented by the name of a researcher responsible for a region containing the fault (e.g., Kozhurin, data 1996). Recent unpublished contributions have a reference to the responsible researcher (or "working material" for contributions of the database owners) and the year of update.

4 Source data

The actual structure and contents of AFEAD is defined by the style of data presentation used in recent active-fault studies. AFEAD generally inherits the approach of the World Map of Major Active Faults (scale 1 : 5 000 000; Trifonov and Machette, 1993), although it is radically improved with respect to spatial detail and fault parameterization. These improvements required the compilation of heterogeneous sources,

Table 2. Values of derivative attributes.

Attribute	Domain	Comments
FAULT_ID	Integer	Unique ID number of the object
RATE	{1, 2, 3}	1 – high, estimated rate exceeds 5 mm yr^{-1}; 2 – medium, estimated rate from 1 to 5 mm yr^{-1}; 3 – low, no estimates or estimated rate below 1 mm yr^{-1}
SENS1	{N, R, T, D, S, E, V, U}	Described in Table 1
SENS2	{N, R, T, D, S, E, V, null}	Described in Table 1
UPSIDE	{N, NE, E, SE, S, SW, W, NW, null}	Direction of the uplifted side
CONF	{A, B, C, D}	A – *proven active* by either a series of published historical and instrumental evidence or paleoseismological studies; B – unambiguous *surface deformations, but no recent slips* have been described yet; C – few pieces of evidence of activity are known; D – fault once declared active, but the *evidence is insufficient* or even absent.

mostly regional maps of active faults and research papers. Due to the amount of processed data, we stopped updating a fault zone when we reached sufficient detail, and we resumed the processing in the rare case of a contradiction between the database and newly published data. We exploited several sources of fault characteristics with varying data processing methods, and we present them following the workflow of data population in AFEAD.

The first attempt to collect a global-scale inventory of active faults was the World Map of Major Active Faults, which was initiated in 1989 and approved by the International Lithosphere Commission in 1990 as Project II-2 of the International Lithosphere Program (ILP). The project was supported by the United Nations Educational, Scientific and Cultural Organization (UNESCO) as a contribution of the ILP to the International Decade of Natural Disaster Reduction, launched by the United Nations (UN), and was included in the ILP Global Seismic Hazard Assessment Program in 1993. During that project, the AFEAD authors proposed data representation methods (Trifonov and Machette, 1993) and the database DB96 model for the parametrization of active faults (Ioffe and Kozhurin, 1996) based on the software design of Ioffe et al. (1993). This first database of active faults of Eurasia contained about 10 000 objects. These data were published as the active-fault maps of Eurasia and northern and eastern Africa (1 : 10 000 000, Fig. 3; Trifonov, 2004), Eurasia (1 : 5 000 000; Trifonov, 1997), and in table format (Trifonov et al., 2002).

In AFEAD, we use the spatial database of the World Map of Major Active Faults as a base layer, referred to as DB96. Its entries were redrawn to comply with the target scale, the attributes were updated according to the actual database scheme and the references were revised, thereby providing a framework for further data collection. During AFEAD development, much of the DB96 data have been replaced by more relevant information. However, there are some regions that cannot be updated due to the absence of studies after the publishing of DB96.

Among the rest of the sources, a large number of data were collected from region-scale maps and databases of active faults (e.g., Hessami et al., 2003; Basili et al., 2013), both digital and published in hard copy. Such maps usually bear sufficient information on slip direction and rate as well as generalization level, but they lack reasoning. Therefore, their processing includes georeferencing, verification of their spatial location against topography, and population of attributes (Fig. 4). CONF is to be set uniform and low, and its elevation requires additional studies at individual faults.

Research papers are the most comprehensive sources considering a particular fault or a fault zone. A standard structure of a research paper provides both parameters of active faults and rationale for it, thereby allowing us to assess the reasoning and choose the most relevant and accurate data among that published. However, the methods and definitions significantly vary by source. Therefore, they require a thorough analysis and normalization before data population (Fig. 5).

Spatial adjustment and verification of populated fault parameters require complementary sources. The global digital elevation model SRTM V3 and Landsat 7 ETM+ imagery serve as base maps for any spatial processing. Earthquakes are strong evidence of fault activity, so we utilized catalogs of the National Earthquake Information Center (NEIC) of the United States Geological Survey (https://earthquake.usgs.gov/, last access: 5 May 2022) and the International Seismological Centre (ISC, http://www.isc.ac.uk/, last access: 5 May 2022) to refine the confidence in fault activity (a value of "CONF" field) and, where needed, to specify fault plane geometry and slip sense.

After the initial population of the database, new data may contradict AFEAD. As there is no direct relation between the
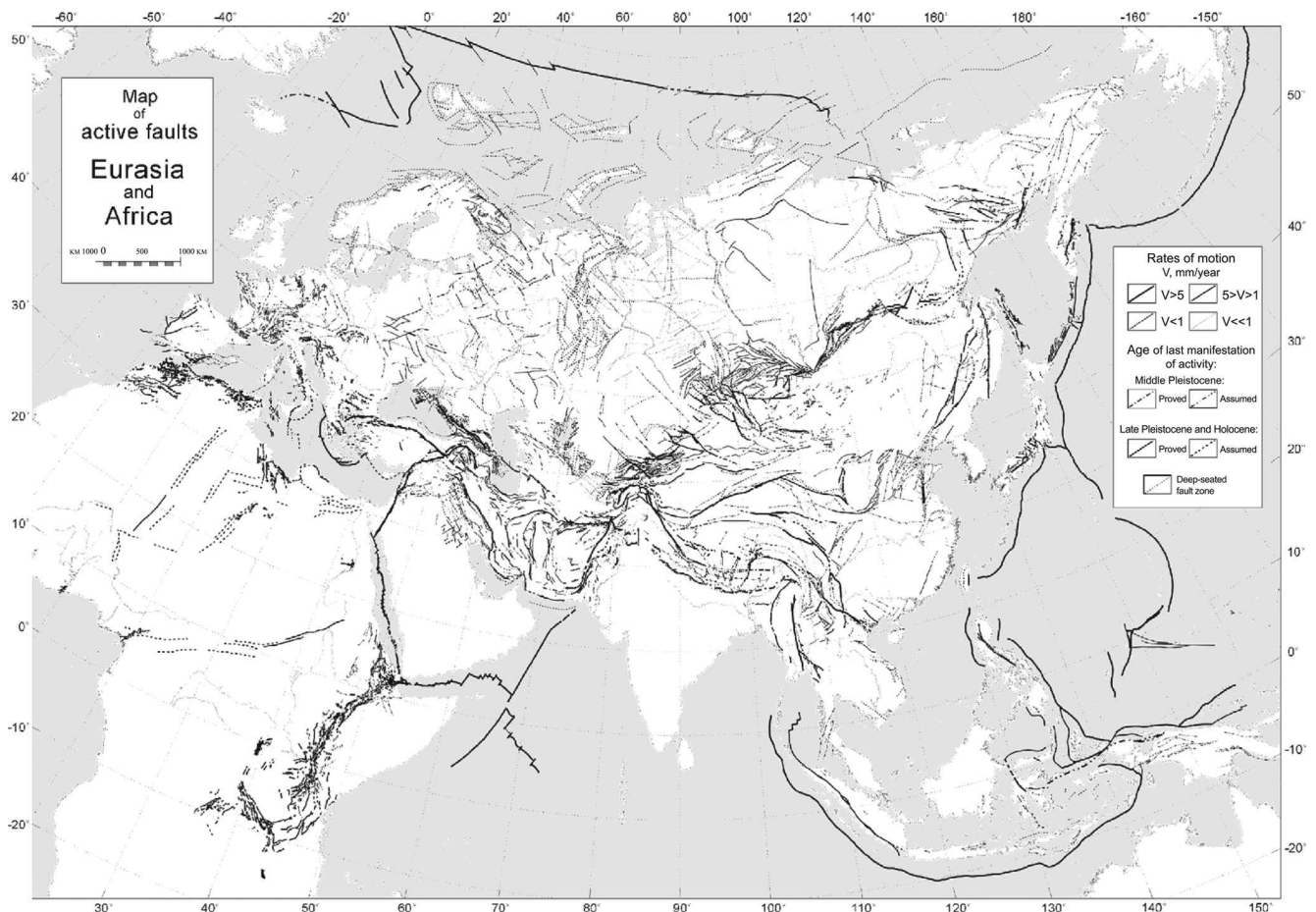
Figure 3. A selection from the World Map of Major Active Faults for Eurasia and Africa (Trifonov, 2004).

recency of the information and its accuracy, the assimilation of recent data requires the comparison of older and recent studies. The result of the comparison affects CONF, resulting in either its elevation or decrease up to deletion from the database.

Published data have to be supplemented by the additional interpretation of remote sensing data to facilitate the uniform detail of the database by adding unpublished objects. In most cases, additional research was required at the spatial boundaries of cited studies (e.g., national borders or the limits of a tectonic structure). Unfortunately, a significant number of fault zones still lack published data. The absence of relevant published information could be confused with the absence of active faults; therefore, the AFEAD authors collect unpublished data, after thorough consideration and the assignment of a level of confidence. This contribution significantly improves the spatial pattern of active faults in remote areas, mainly in North Asia. Such unpublished entries will be updated as soon as new studies are published, although this is unlikely to occur in the near future for all of the entries.

5 Overview of the dataset

5.1 General characteristics

The database comprises 48 205 objects – active faults and their segments. Most of them (ca. 44 000) belong to mainland Eurasia, whereas the rest are located on islands or underwater. Active faults tend to group in broad belts (Fig. 1) incorporating minor plates and crustal blocks at margins of tectonic plates. However, active faults were also mapped elsewhere across tectonic plates. Individual faults are indistinguishable on the continent-scale map, but a fault density map (Fig. 6) shows an even greater contrast between active belts and cratons.

The mean length of database objects is 22 km, and 90 % of them belong to the range from 5 to 60 km (Fig. 7). Most of the faults longer than 100 km are underwater, where no detailed data are available.

Within the dataset, 81 % of objects have been previously published, including 36 % with two and more sources. The PARM field represents the amount of data yielded by studies. Aside from unpublished entries, 31 % of entries have an empty PARM field, with most of them originating from maps
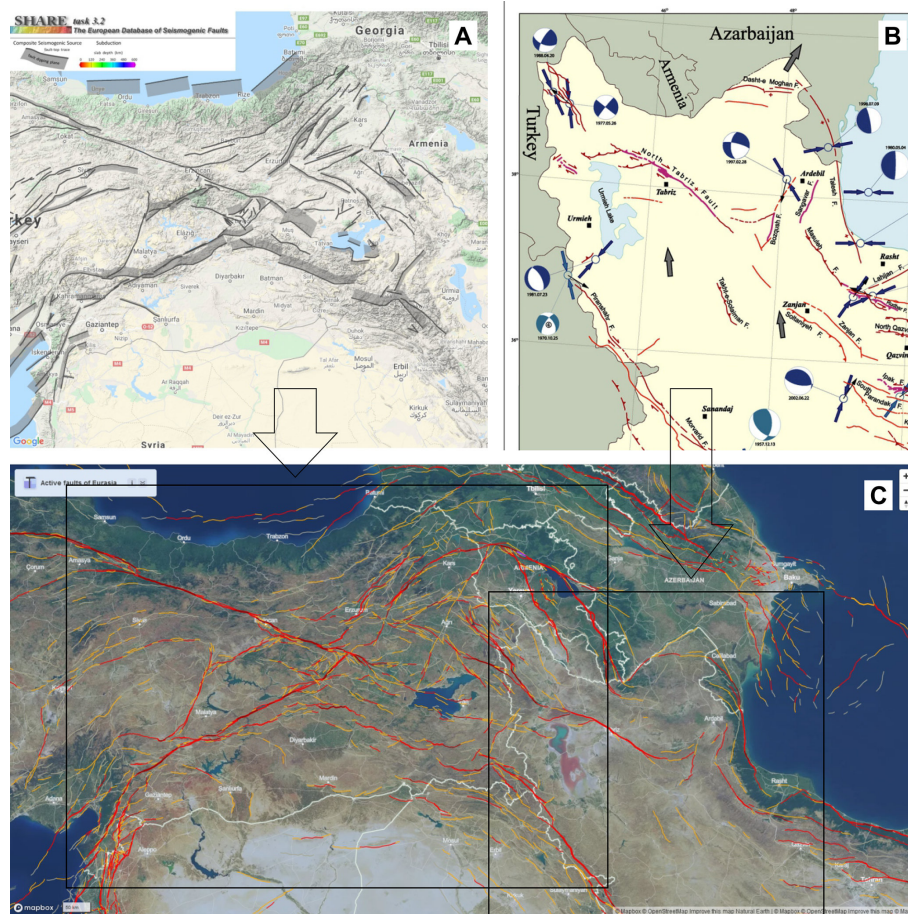
Figure 4. Active faults in Transcaucasia, showing **(a)** the web interface of the European database of Seismogenic Faults (Basili et al., 2013), **(b)** a fragment of the Major Active Faults of Iran Map (Hessami et al., 2003), and **(c)** the AFEAD web interface (http://neotec.ginras.ru/ index/mapbox/database_map.html, last access: 5 May 2022) for this area.

with no parametrization. Complete descriptions are common in the western Pacific, Baikal Rift, and Alpine–Himalayan Belt. In the latter, prominent clusters are located in the eastern Mediterranean and Anatolia, Transcaucasia, the Iranian mountains, and the Tibetan region. On the contrary, active faults within cratons are typically poorly studied and low confidence, representing the different approaches of active-fault studies.

The intensity of motions at a fault affects two parameters: RATE and CONF (Fig. 8). The former is a natural characteristic, whereas the latter represents the quality and quantity of studies that considered the fault (Table 2). Therefore, subsequent research efforts may elevate CONF, even at slow-moving faults. Most AFEAD entries (94 %) are slow-moving faults (RATE = 3); all low-confidence (CONF = "D") and most CONF = "C" faults fall within this class. Deformation rates exceeding $1\,\mathrm{mm\,yr^{-1}}$ are identified for 6.2 % of entries; faults exceeding $5\,\mathrm{mm\,yr^{-1}}$ comprise 1 % of entries. The amount of both CONF = "A" and CONF = "B" faults in

RATE classes gradually decreases from slow-moving to fast-moving faults.

The most frequent fault sense in AFEAD is reverse (21 %), even considered separately from thrusts (7 %); the normal sense is identified in 17 % of objects; right-lateral and left-lateral faults are equally common (13 % and 12 %, respectively); and dip-slip is identified in 18 % of entries. The sense of slip remains unknown for the rest (11 %) of the entries. A secondary component is provided for 22 % of entries (Fig. 9).

5.2 Regional examples

To provide a deeper insight into the AFEAD structure and usability, we present two contrasting examples from the compressional setting of the Caucasus and the transtensional setting of the Baikal region, corresponding to K-38 Tbilisi (http://neotec.ginras.ru/index/datamap/AFEAD_ K38_Map.html, last access: 5 May 2022) and N-49 Chita (http://neotec.ginras.ru/index/datamap/AFEAD_N49_Map, last access: 5 May 2022) $4° \times 6°$ map tiles, respectively.
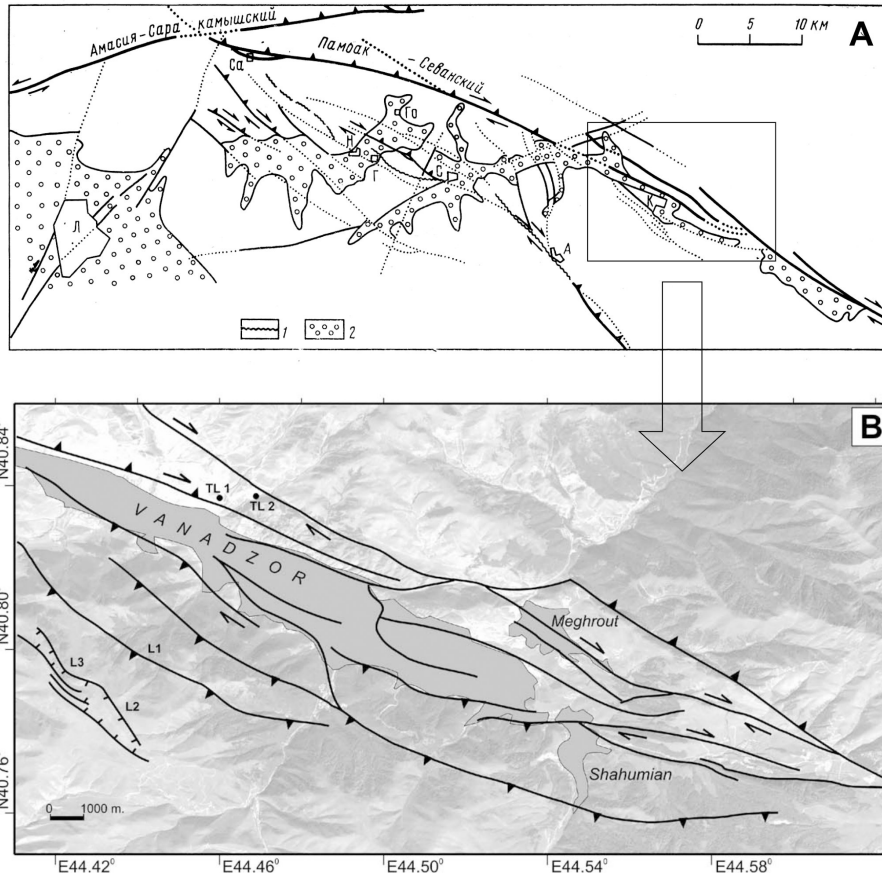
Figure 5. Sources for mapping the Pambak–Sevan fault near Vanadzor city, Armenia. Panel **(a)** displays an initial study by Trifonov et al. (1990), where geographical objects and coordinates had been omitted due to Soviet legal requirements; the location of Vanadzor is labeled with "K" (representing its former name "Kirovakan") within the box. Panel **(b)** shows a more recent georeferenced map (Karakhanian et al., 2004). Both papers affected the location and attributes of the Pambak–Sevan fault in AFEAD, including the object highlighted in Fig. 2.
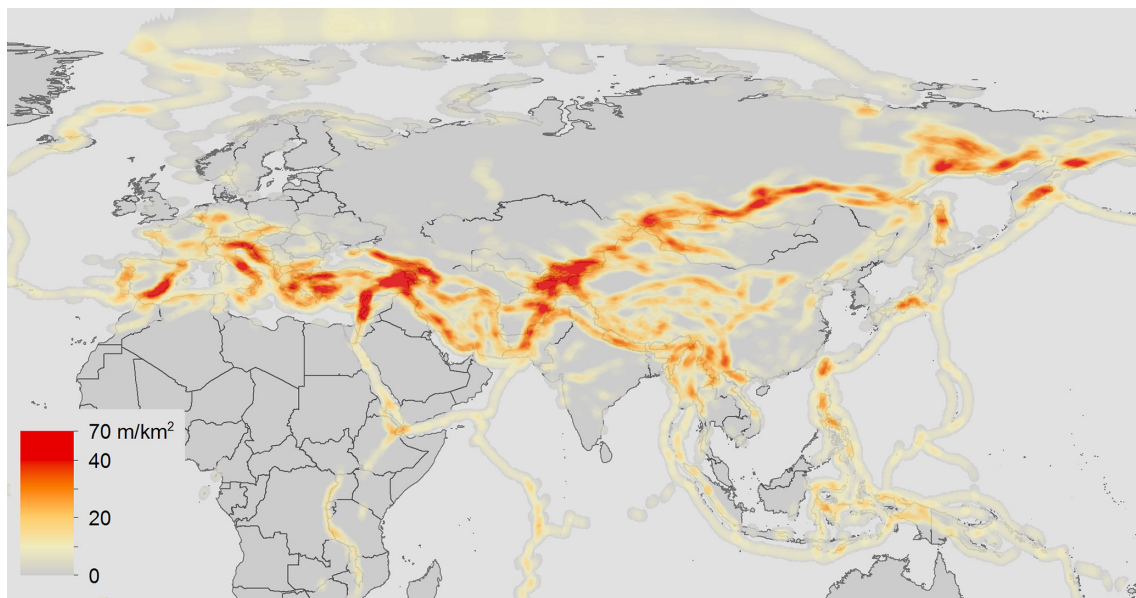
Figure 6. The density of active faults in AFEAD (in meters per square kilometer). Low-confidence faults (CONF of "D") are omitted.
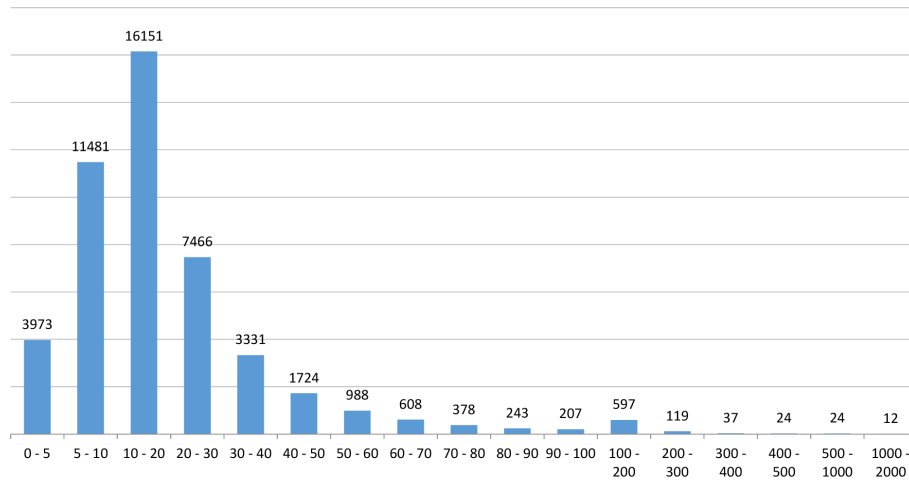
Figure 7. Distribution of AFEAD objects by length.
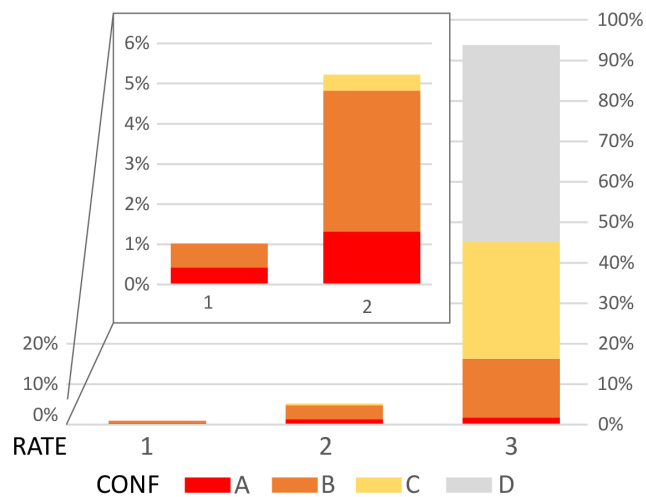
Figure 8. Distribution of AFEAD objects by RATE (horizontal axis) and CONF (color) classes. The inset shows a close-up of RATE classes 1 and 2.

The Caucasus is located at the northern flank of the Alpine–Himalayan collision belt. It has been experiencing compression since the Oligocene (e.g., Nikishin et al., 1998) that, along with mantle geodynamics, built up the high mountain range of the Greater Caucasus and the highlands of Transcaucasia south of it. Recent deformation in the area is concentrated at the main thrust of the Greater Caucasus and an arcuate system of Zheltorechensky–Sarykamysh, Pambak–Sevan, Garni, and Khanarassar strike-slip fault zones in Transcaucasia (Fig. 10). Minor faults generally follow their pattern, although they are scattered across large areas up to the northern foothills of the Caucasus. In the Greater Caucasus, most of the faults had been identified by the 1980s and published in monographs on broad geological topics (e.g., Milanovsky, 1968; Kogoshvili, 1970) incorporated into

DB96. Few works were carried out after the compilation of DB96; thus, the Greater Caucasus area appears to lack state-of-art paleoseismological studies. Transcaucasia had been much less studied until the infamous 1988 Spitak earthquake ($M_s = 6.7$; Bommer and Ambraseys, 1989). The subsequent extensive studies (e.g., Trifonov et al., 1990; Karakhanian et al., 2004) revealed spatial patterns and geodynamic settings of active faulting. Hence, AFEAD entries in Transcaucasia bear many more attributes than those at the Greater Caucasus.

Another case of different tectonic settings and research history is the Baikal Rift zone (Fig. 11). It is a linear system of grabens bounded by normal or transtensional faults that have been developing since the Oligocene (e.g., Logatchev and Zorin, 1992). General features of this zone had been identified by the 1980s (Sherman and Levi, 1978; Solonenko, 1977; Logachev, 1984), but intermittent seismicity keeps drawing constant attention to the active faults of the region. In addition to research papers, recent studies have been published as regional data collections: an inventory of paleoseismic sites by Smekalin et al. (2010), the database of southern East Siberian Pliocene–Quaternary faults (Lunina et al., 2014), and a seismotectonic map of eastern Siberia (Imaeva et al., 2015). A compilation of these sources provides uniform input for the AFEAD.

6 Update strategy

Since the initial population of the database, the database owners have been monitoring scholarly literature for relevant data. After the acquisition of new data, the data reference is appended to the reference list, and a graphical representation of faults is georeferenced and traced (if needed). Their location is checked against AFEAD fault patterns, so that the location of the new fault line may either differ from AFEAD objects or match some of them. In the first case, the fault
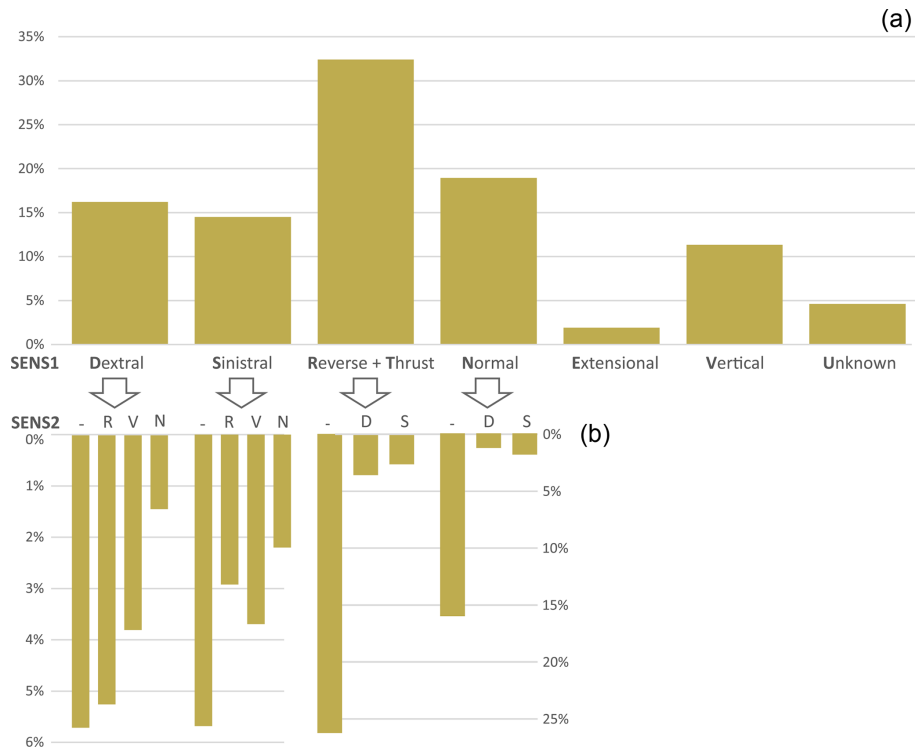
Figure 9. Distribution of AFEAD objects by fault sense, showing **(a)** SENS1 and **(b)** SENS2 for each SENS1 value. For the sense indices, see Table 1.
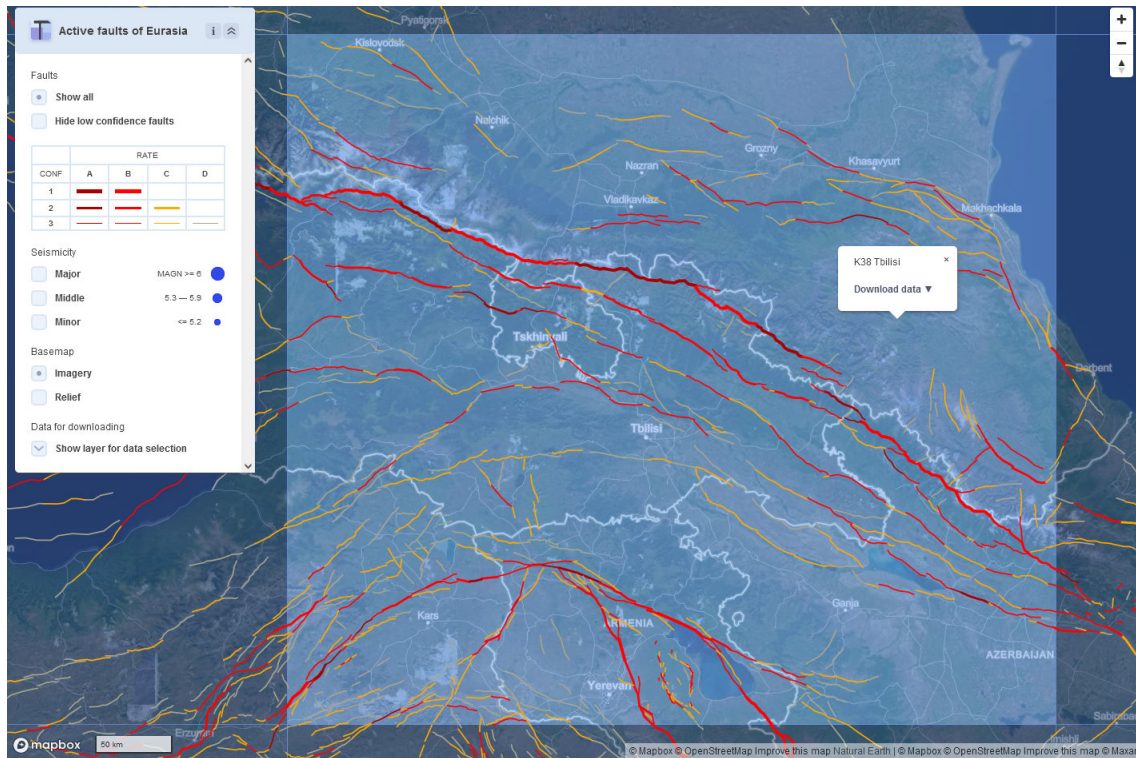

Figure 10. Representation of active faults in the Caucasus by the AFEAD web interface (http://neotec.ginras.ru/index/mapbox/database_map.html, last access: 5 May 2022).
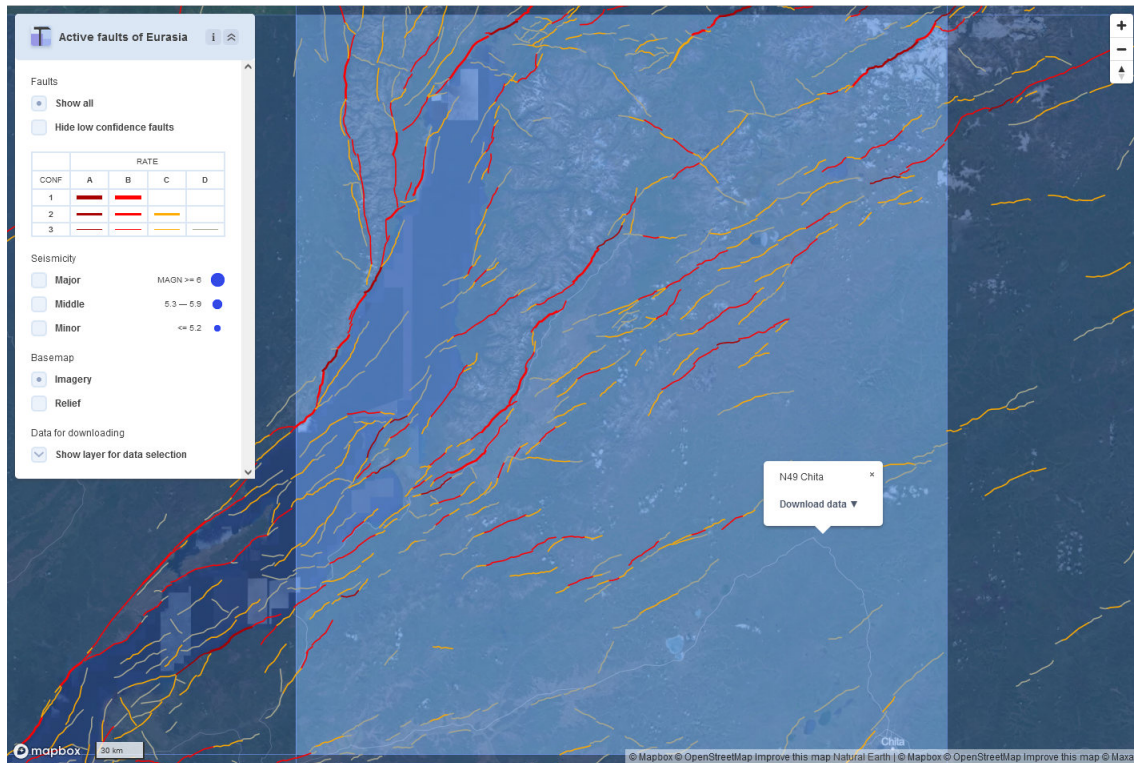
Figure 11. Representation of active faults at the Baikal rift zone by the AFEAD web interface (http://neotec.ginras.ru/index/mapbox/database_map.html, last access: 5 May 2022).

line is redrawn to comply with the topography and target map scale of 1 : 1 000 000, and primary attributes are populated: AUTH, FAULT_NAME, ZONE_NAME, PARM, and TEXT. Derivative attributes of kinematics (SENS1, SENS1, and UPSIDE) are set according to them and topography. The confidence of activity relies both on reported evidence of activity (Table 2) and on coherence of fault parameters with parameters of adjacent AFEAD objects. If data for neighboring faults support the activity of the studied object, it may elevate CONF for the entire zone; otherwise, if faults contradict each other, CONF is reduced for the new fault or pre-existing objects, or for both, depending on the reasoning behind each object. Finally, RATE is set according to the actual slip measurement at the fault and its confidence, if possible, or propagated from other objects in the fault zone. The second case of matching locations causes an update to the existing AFEAD object. Its location is then adjusted according to the reasoning in the new source and AFEAD. The new citation is then appended to AUTH, and newly acquired parameters are appended to PARM along with comments to TEXT. If these affect attributes of kinematics (SENS1, SENS1, and UPSIDE), their values are reconsidered. New reported evidence of activity may also elevate the CONF value.

The presented database AFEAD v.2022 has reached target detail, and no major revisions in the database model are planned after the completion of this study. However, new ver-

sions will be released after the acquisition of recently published data. To ensure data consistency, no direct external contribution to the database is possible. The authors encourage researchers to inform us about missing or recently obtained data via the email (to the corresponding author).

7 Data access

The main access point to the most recent version of AFEAD is a web map available at http://neotec.ginras.ru/index/mapbox/database_map.html (last access: 5 May 2022). The current dataset (v.2022) is available from https://doi.org/10.13140/RG.2.2.25509.58084 (Bachmanov et al., 2022). A variety of up-to-date database representations, supplemented by a reference list and explanatory notes, are hosted at http://neotec.ginras.ru/index/english/database_eng.html (last access: 5 May 2022); this includes a raster overview map, raster map tiles designed for print, and .kmz and .shp vector tiles.

Studies considering the AFEAD scheme or its reasoning may refer to this study, the DOI of the database, or the initial publication on this topic from Bachmanov et al. (2017). None of them would be an acceptable reference for studies considering fault locations or their parameters; instead, the

researcher is advised to cite source studies provided in the AUTH field.

8 Conclusion

AFEAD is the largest and the most comprehensive collection of active faults, comprising ∼ 48 000 entries spanning the entirety of Eurasia and the adjacent seas. For each entry in the database, its spatial location and characteristics of motions are provided. All spatial data have uniform detail equal to hard copy maps with a 1 : 1 000 000 scale. The attributes of faults store relevant information transferred from sources and derivative parameters generated by the database owners.

The database makes a spatial search possible for local studies. It provides sufficient detail for planning a study of a particular fault system and guides deeper bibliographical investigations. This scenario is particularly significant for vast central and northern Asian areas, where most studies are available only in Russian and hard copy. Moreover, the database model provides a foundation for GIS-based regional- and continental-scale integrative studies. The authors suggest that the use of this database will support geodynamic and paleoseismological studies in Eurasia.

Appendix A: Data domain of the PARM field

Age – age of the latest dated slip in years BP or units of the geological timescale (e.g., "N1" for the Miocene or "Q4" for the Holocene).
Signs – signs of recent fault motion:
DF,   drape fold;
DT,   sharp change in recent deposits' thickness;
EC,   en echelon array of compressional structures;
EQ,   earthquake hypocenter;
ER,   surface seismic ruptures;
ET,   en echelon array of extensional structures;
FD,   surface folding;
FM,   earthquake focal mechanism;
FR,   linear group of fractures;
GA,   gas and hydrochemical anomalies;
GD,   geodetic surveys;
GP,   geophysical data;
HS,   historical or archeological data;
HT,   hydrothermal springs;
LS,   linear group of landslides or rockfalls;

MV,   mud volcanism;
OC,   offset river channels;
OD,   offset recent deposits;
OT,   offset river terraces or alluvial fans;
PS,   paleoseismic sites;
TS,   sharp change in tectonic structure;
VC,   volcanic chain.
Dating – dating techniques:
AR,   archeological;
CR,   radiocarbon;
GC,   geological correlation;
HI,   historical;
IN,   instrumental;
LH,   lichenometry;
MC,   geomorphological correlation.
Layers – faulted layers of the lithosphere:
S,   sedimentary cover;
UC,   upper crust;
LC,   lower crust;
M,   upper mantle.
Dip – dip angle (in degrees) and dip direction (cardinal directions), occasionally supplemented with site coordinates or fault part (cardinal directions, "C" for central) for which data are relevant.
Depth – fault depth in kilometers.
Offset – measured offset at the surface, occasionally supplemented with site coordinates or fault part (cardinal directions, "C" for central) for which data are relevant.
Rake – the angle between the slip direction and the strike line.
Rate – average slip rate ($mm\,yr^{-1}$) supplemented with time span and, occasionally, site coordinates or fault part (cardinal directions, "C" for central) for which data are relevant. An asterisk (*) indicates geodetic measurements, and a double asterisk (**) indicates seismological measurements.
Ratio – ratio of strike slip to dip slip.
Seism – parameters of an earthquake that occurred at the fault: magnitude, name, date, depth.
SeismDepth – depth range of earthquakes at the fault.
SeismRecur – mean recurrence interval of earthquakes at the fault.
Sense – fault slip sense, abbreviated (as in Table 1).
Side – a direction of an upthrown side of the fault.

References

Atanackov, J., Jamšek Rupnik, P., Jež, J., Celarc, B., Novak, M., Milanič, B., Markelj, A., Bavec, M., and Kastelic, V.: Database of active faults in Slovenia: Compiling a New Active Fault Database at the Junction Between the Alps, the Dinarides and the Pannonian Basin Tectonic Domains, Front. Earth Sci., 9, 604388, https://doi.org/10.3389/feart.2021.604388, 2021.

Bachmanov, D. M., Kozhurin, A. I., and Trifonov, V. G.: The Active Faults in Eurasia Database, Geodynamics and Tectonophysics, 8, 711–736, https://doi.org/10.5800/GT-2017-8-4-0314, 2017.

Bachmanov, D. M., Trifonov, V. G., Kozhurin, A. I., and Zelenin, E. A.: AFEAD v.2022, ResearchGate [data set], https://doi.org/10.13140/RG.2.2.25509.58084, 2022.

Basili, R., Kastelic, V., Demircioglu, M. B., Garcia Moreno, D., Nemser, E. S., Petricca, P., Sboras, S. P., Besana-Ostman, G. M., Cabral, J., Camelbeeck, T., Caputo, R., Danciu, L., Domac, H., Fonseca, J., García-Mayordomo, J., Giardini, D., Glavatovic, B., Gulen, L., Ince, Y., Pavlides, S., Sesetyan, K., Tarabusi, G., Tiberti, M. M., Utkucu, M., Valensise, G., Vanneste, K., Vilanova, S., and Wössner, J.: The European Database of Seismogenic Faults (EDSF) compiled in the framework of the Project SHARE, https://doi.org/10.6092/INGV.IT-SHARE-EDSF, 2013.

Bollinger, L., Klinger, Y., Forman, S., Chimed, O., Bayasgalan, A., Munkhuu, U., Davaasuren, G., Dolgorsuren, T., Enkhee, B. and Sodnomsambuu, D.: 25,000 Years Long Seismic Cycle in a Slow Deforming Continental Region of Mongolia, Scientific Reports, 11, 17855, https://doi.org/10.1038/s41598-021-97167-w, 2021.

Bommer, J. J. and Ambraseys, N. N.: The Spitak (Armenia, USSR) earthquake of 7 December 1988: a summary engineering seismology report, Earthq. Eng. Struct. D., 18, 921–925, 1989.

Christophersen, A., Litchfield, N., Berryman, K., Thomas, R., Pagani, M., Henshaw, P., Wyss, B., Wallace, L., Ries, W., Hayes, G., Haller, K., Toshikazu, Y., Koehler, R. D., Clark, D., Wolfson-Schwehr, M., Boettcher, M., Villamor, P., Horspool, N., Ornthammarath, T., Zuniga, R., Langridge, R., Stirlig, M., Goded, T., Basili, R., Stein, R., Costa, C., and Yeats, R.: Development of the Global Earthquake Model's neotectonic fault database, Nat. Hazards, 79, 111–135, https://doi.org/10.1007/s11069-015-1831-6, 2015.

Danciu, L., Sesetyan, K., Demircioglu, M., Gulen, L., Zare, M., Basili, R., Elias, A., Adamia, S., Tsereteli, N., Yalçin, H., Utkucu, M., Khan, A., Sayab, M., Hessami, K., Rovida, A., Stucchi, M., Burg, J.-P., Karakhanian, A., Babayan, H., and Giardini, D.: The 2014 Earthquake Model of the Middle East: seismogenic sources, Bull. Earthq. Eng., 16, 3465–3496, 2018.

Galadini, F., Falcucci, E., Galli, P., Giaccio, B., Gori, S., Messina, P., Moro, M., Saroli, M., Scardia, G., and Sposato, A.: Time intervals to assess active and capable faults for engineering practices in Italy, Eng. Geol., 139, 50–65, 2012.

Ganas, A., Oikonomou, I. A., and Tsimi, C.: NOAfaults: a digital database for active faults in Greece, Bull. Geol. Society. Greece, 47, 518–530, 2021.

Geologische Rundschau: 43, 1, 1955.

Hessami, K., Jamali, F., and Tabassi, H.: Major active faults of Iran, IIEES, Tehran, 2003.

Imaeva L. P., Imaev, V. S., Smekalin, O. P., Grib, N. N., Koz'min, B. M., and Chilizubov, A. V.: Seismotectonics map of Eastern Siberia, edited by: Gusev, G. S., and Kolodeznikov, I. I., ESDB repository, GCRAS, Moscow, https://doi.org/10.2205/ESDB-VONZ-125-map, 2015.

Ioffe, A. I. and Kozhurin, A. I.: Database of active faults of Eurasia, Earthq. Pred. Res., 5, 431–435, 1996.

Ioffe, A., Govorova, N., Volchkova, G., and Irifonov, R.: Data base of active faults for the USSR area, Geoinformatics, 4, 289–290, 1993.

Jomard, H., Cushing, E. M., Palumbo, L., Baize, S., David, C., and Chartier, T.: Transposing an active fault database into a seismic hazard fault model for nuclear facilities – Part 1: Building a database of potentially active faults (BDFA) for metropolitan France, Nat. Hazards Earth Syst. Sci., 17, 1573–1584, https://doi.org/10.5194/nhess-17-1573-2017, 2017.

Karakhanian, A. S., Trifonov, V. G., Philip, H., Avagyan, A., Hessami, K., Jamali, F., Bayraktutan, M. S., Bagdassarian, H., Arakelian, S., Davtian, V., and Adilkhanyan, A.: Active faulting and natural hazards in Armenia, eastern Turkey and northwestern Iran, Tectonophys., 380, 189–219, 2004.

Kogoshvili, L. V.: Active tectonics of Georgia and its influence to relief, Metsienerba publishing house, Tbilisi, 220 pp., 1970 (in Russian).

Kozhurin, A. and Zelenin, E.: An extending island arc: The case of Kamchatka, Tectonophys., 706, 91–102, 2017.

Kumamoto T.: Long term conditional seismic hazard of Quaternary active faults in Japan, J. Seismol. Soc. Japan, 50, 53–71, 1998 (in Japanese).

Logachev, N. I. (Ed.): Geology and seismicity of the Baikal-Amur railway zone. Neotectonics, Nauka, Novosibirsk, 1984, 208 pp., 1984 (in Russian).

Logatchev, N. A. and Zorin, Y. A.: Baikal rift zone: structure and geodynamics, Tectonophys., 208, 273–286, 1992.

Lunina, O. V., Caputo, R., and Gladkov, A. S.: Southern East Siberia Pliocene-Quaternary faults: database, analysis and inference, Geosci. Front., 5, 605–619, 2014.

Milanovsky, E. E.: Neotectonics of the Caucasus, Nedra, Moscow, 484 pp., 1968 (in Russian).

Mohadjer, S., Ehlers, T. A., Bendick, R., Stübner, K., and Strube, T.: A Quaternary fault database for central Asia, Nat. Hazards Earth Syst. Sci., 16, 529–542, https://doi.org/10.5194/nhess-16-529-2016, 2016.

Nikishin, A. M., Cloetingh, S., Brunet, M. F., Stephenson, R. A., Bolotov, S. N., and Ershov, A. V.: Scythian Platform, Caucasus and Black Sea region: Mesozoic-Cenozoic tectonic history and dymanics, in: Peri-Tethys Memoir 3: stratigraphy and evolution of Peri-Tethyan platforms, edited by: Crasquin-Soleau, S. and Barrier, E., Mem. Mus. Natn. Hist. Nat., 177, Paris, 163–176, 1998.

Rukieh, M., Trifonov, V. G., Dodonov, A. E., Minini, H., Ammar, O., Ivanova, T. P., Zaza, T., Yusef, A., Al-Shara, M., and Jobaili, Y.: Neotectonic map of Syria and some aspects of Late Cenozoic evolution of the northwestern boundary zone of the Arabian plate, J. Geod., 40, 235–256, 2005.

Schellart, W. P. and Lister, G. S.: The role of the East Asian active margin in widespread extensional and strike-slip deformation in East Asia, J. Geol. Soc., 162, 959–972, 2005.

Sherman, S. I. and Levi, K. G.: Transform faults in the Baikal rift zone and seismicity on its flanks, in: Tectonics and seismicity of the continental rift zones, Nauka, Moscow, 7–185, 1978 (in Russian).

Slemmons, D. B., and dePolo, C. M.: Evaluation of active faulting and associated hazards, in: Active tectonics, edited by: Wallace, R. E., National Academy Press, Washington, D.C., 45–62, ISBN 978-0-309-07395-0, 1968.

Smekalin, O. P., Chipizubov, A. V., and Imayev, V. S.: Paleoearthquakes in the Pribaikalie: methods and results of dating, Geod. Tectonophys., 1, 55–74, 2010 (in Russian).

Solonenko, V. P. (Ed.): Seismic zonation of the Eastern Siberia and its geologic-geophysical basis, Nauka, Novosibirsk, 304 pp., 1977 (in Russian).

Styron, R. H. and Pagani, M.: The GEM Global Active Faults Database, Earthq. Spectra, 36, 160–180, https://doi.org/10.1177/8755293020944182, 2020.

Trifonov, V. G.: World map of active faults, their seismic and environmental effects, in: Historical and Prehistorical Earthquakes in the Caucasus, 1st edn., edited by: Giardini, D. and Balassanian, S., Kluwer Acad. Publ., Dordrecht, 169–180, ISBN 978-0-7923-4614-2, 1997.

Trifonov, V. G.: Active faults in Eurasia: general remarks, Tectonophys., 380, 123–130, 2004.

Trifonov, V. G. and Kozhurin, A. I.: Study of active faults: Theoretical and applied implications, Geotectonics, 44, 510–528, 2010.

Trifonov, V. G. and Machette, M. N.: The world map of major active faults project, Ann. Geophys., 36, 225–236, 1993.

Trifonov, V. G., Karakhanian, A. S., and Kozhurin, A. I.: The Spitak earthquake as a manifestation of active tectonics, Geotectonika, 6, 46–60, 1990 (in Russian).

Trifonov, V. G., Soboleva, O. V., Trifonov, R. V., and Vostrikov, G. A.: Recent Geodynamics of the Alpine-Himalayan Collision Belt, GEOS, Moscow, 225 pp., ISBN 5-89118-279-3, 2002 (in Russian).

Ulomov, V., Strakhov, V., and Giardini, D.: Seismic hazard assessment in Northern Eurasia, Ann. Geophys., 36, 83–92, 1993.

Walker, J. F., Boncio, P., Pace, B., Roberts, G., Benedetti, L., Scotti, O., Visini, F., and Peruzza, L.: Fault2SHA Central Apennines database and structuring active fault data for seismic hazard assessment, Scientific data, 8, 87, https://doi.org/10.1038/s41597-021-00868-0, 2021.

Wallace, R. E.: Structure of a portion of the San Andreas rift in southern California, Geol. Soc. Am. Bull., 60, 781–806, 1949.