Arab J Geosci (2015) 8:10893–10909
DOI 10.1007/s12517-015-1878-4

ORIGINAL PAPER

Worldwide trends in research on the San Andreas Fault System

Fabrizio Terenzio Gizzi[1]

Received: 22 September 2014 / Accepted: 12 March 2015 / Published online: 10 May 2015

Abstract This paper deals with the bibliometric study of the research concerning the San Andreas Fault System (SAFS), an extraordinary example of a complex boundary between the North of America and the Pacific plates that can be seen, followed and studied on land for hundreds of kilometers through California, in the USA. The bibliometric analysis discussed here considers the time span of 23 years, from 1991 to 2013. The bibliographic databases taken as a reference are the Science Citation Index Expanded (SCI-EXPANDED) and the Social Sciences Citation Index (SSCI) accessed via Web of Science Core Collection. After having selected the proper search terms to fix the query and extracting the useful information from the two Web of Science (WoS) databases, we performed the analysis of the outputs paying attention to the document typology and languages, journals, subject categories, authors, articles, countries, institutions, and keywords. The results of this study can (1) contribute to see how the research on SAFS has changed over the time and suggest clues about the future investigation trends, (2) help scientists or institutions to build a research network and find strategic partners, (3) suggest scientists and institutions the dissemination strategies, and (4) provide helpful information to researchers wishing to embark on work in this area.

Keywords San Andreas Fault System · California · Tectonic study · Earthquake · Bibliometry

✉ Fabrizio Terenzio Gizzi
f.gizzi@ibam.cnr.it

[1] Institute of Archaeological and Monumental Heritage (IBAM, National Research Council), 85050 Tito Scalo, PZ, Italy

Introduction

The San Andreas Fault System (SAFS) is a complex of faults having mainly right-lateral strike-slip movements that accommodate most of the relative N-S motion between the North America and the Pacific plates. It runs approximately for 1300 km, from the connection with the Mendocino triple junction fault zone on the north to the head of the Gulf of California to the southeast (Fig. 1) (Ellsworth 1990; Wallace 1990).

The dextral strike-slip San Andreas Fault zone is the main element of the San Andreas Fault System. The name San Andreas fault was adopted for the first time by Lawson (1895) even if the fault gained importance only after the 1906 great San Francisco earthquake (M=7.7) (Wallace 1990).

San Andreas Fault zone is divided into ten sections that are, from the north to the south, the Shelter Cove, North Coast, Peninsula, Santa Cruz Mountains, Creeping, Parkfield, Cholame-Carrizo, Mojave, San Bernardino Mountains, and Coachella sections (Bryant and Matthew 2002). Each of them consists of smaller elements such as branches or strand faults. For example, the San Bernardino Mountain section hosts Quaternary and Holocene faults such as Mission Creek, Mill Creek, and Banning faults; Mojave section includes Nadeau, Punchbowl, and Little Rock fault strands (Bryant and Matthew 2002). Other significant elements of the San Andreas Fault System are, from the north to the south, the Bartlett Springs, Maacama, Rodgers Creek, Green Valley, Calaveras, Hayward, San Gregorio, San Jacinto, Elsinore, and Imperial fault zones, in their turn including several tectonic lineaments (Bryant and Matthew 2002).

The motion of the North of America and the Pacific plates induces an elastic strain in SAFS that causes strong earthquakes once it is released. As a matter of fact, numerous seismic events happened, historically speaking, such as the Fort Tejon January 9, 1857 (M=7.8) and the San Francisco April 18, 1906 one (M=7.7), the two most significant earthquakes

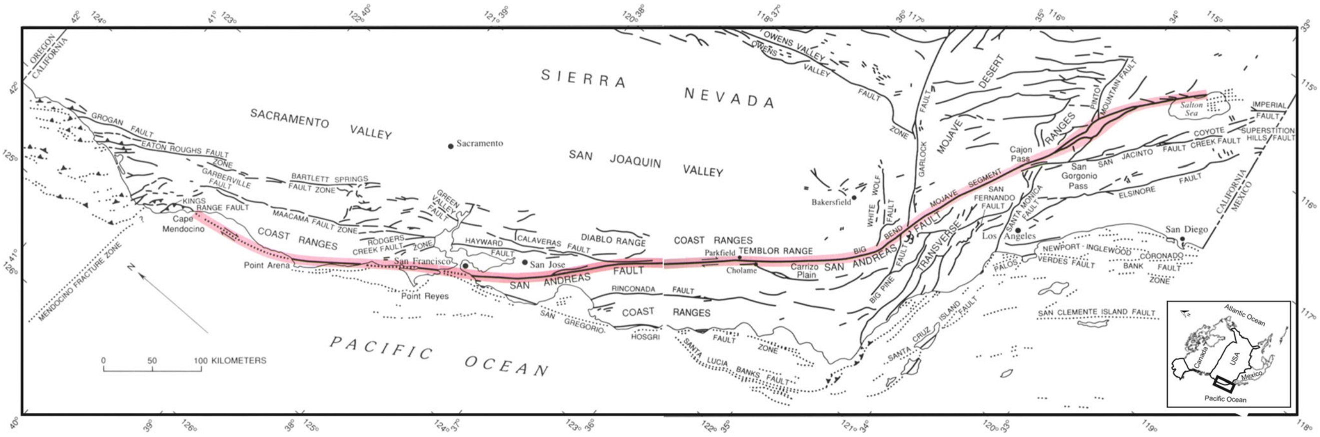


Fig. 1 Overview of the San Andreas Fault System in California. The red colour puts into evidence the San Andreas fault proper (after Wallace 1990, modified)

happened along the San Andreas fault proper, and other strong Hayward fault-generated earthquakes were those of June 10, 1836 (M=6.7) and October 21, 1868 (M=7.0) that both hit the city of San Francisco seriously (Ellsworth 1990).

Due to the unique features of the SAFS plate margin that can be studied and analyzed on land and the need to improve the knowledge of the California seismic hazard, SAFS has received the attention of numerous researchers and, as a result, has become the strike-slip fault system most in-depth studied all over the world. Despite the importance of such tectonic boundary, there are no previous studies aimed at analyzing systematically from the statistical view point, the worldwide scientific production concerning SAFS. Therefore, this work aims to fill this gap partially, dealing with the bibliometric analysis of the scientific production relating to the San Andreas Fault System considering the last 23 years (1991–2013).

The bibliometric studies of the literature based on the analysis of citations, countries, institutions, journals, keywords chosen by the authors or Keyword Plus, language of the papers is a method that is being considered more and more in many disciplines such as medicine, economy, chemistry, biology, engineering, geography, and earth sciences in order both to see how the research has changed over the time and get some clues about the future trends (e.g., Garcia-Ramon and Caballé 1998; Grant et al. 2000; Chiu and Ho 2007; Franceschini and Maisano 2010; Liu et al. 2012; Marx and Bornmann 2013; Niu et al. 2014).

With this in mind, this work analyzes the research on SAFS considering several aspects such as document typology and languages, publication trend and citations, subject categories and journals, productivity of the authors and their influence in the scientific community, the most cited articles, countries and institutions, author keywords and hot issues. The Thomson Reuters' Science Citation Index Expanded (SCI-EXPANDED) and the Social Sciences Citation Index (SSCI) accessed via the Web of Science Core Collection have been the bibliographic databases used to perform this study. It is necessary to emphasize that all the findings and discussions included below have to be strictly considered as the reflection of the content analysis of the two ISI databases taken as a reference here. For that reason, for example, relevant papers (with their citations) and leading topics may be omitted if they are not incorporated in the WoS indexes.

Data and methods

In order to perform the bibliometric study, we have used the SCI-EXPANDED and SSCI citation indexes. According to the Journal Citation Report (JCR), the SCI-EXPANDED database indexes 8478 journals, while SSCI indexes 3045 in 2013. These values were respectively 5550 and 1699 in 1999 with an associated mean annual increase of journals equal to 195.2 and 89.7 in the 1999–2013 period.

With the intention of pulling out the suitable records from the databases, we have considered some strings to be used to search for titles, abstracts, or keywords of the indexed papers.

To select the proper search keywords, we have taken advantage of the *Quaternary fault and fold database* of the USA, compiled by the US Geological Survey (http://earthquakes.usgs.gov/hazards/qfaults, accessed 24 April 2014). That database lists 11 major fault zones making part of SAFS (Bryant et al. 2002). For each of the fault zone, the "Complete report" supplied by the US Geological Survey database gives a literature overview including the names of the principal faults of the fault zone itself (Hart 1998; Treiman 1998; Treiman and Treiman 1998; Black and Hecker 1999; Treiman 1999; Treiman et al. 1999; Bryant and Cluett 1999a, b; Bryant 2000; Bryant and Cluett 2000; Hart and Bryant 2001; Bryant and Cluett 2002; Bryant and Matthew 2002). For example, the San Andreas Fault zone—San Bernardino section—includes, among the others, Mill Creek, Mission Creek, and the Banning faults; the Hayward fault zone—Northern Hayward section—incorporates Mission, Evergreen, Quimby, the Crosley, and Clayton faults. The reports also allowed to consider the different names given by scientists to a fault over the time, as for the Maacama fault-north section-designed as the Talmadge fault in the past (Hart et al. 2001).

The first search string used has been *Andreas fault**, followed by the names of the major faults as derived from the analysis of the US Geological Survey database (Table 1). On the whole, the search counts about one hundred terms. However, as much as≈86 % of the total documents or articles retrieved from the two databases came from the *mother search* that has made use only of the *Andreas fault** as string. Therefore, the other terms have contributed to the final results with about four results each, on average. In order to avoid possible fake results, the papers resulted in the search output have been analyzed manually each time that a new search string has been added to the previous.

Once the records have been selected from the two databases, they have been downloaded as plain text including the names of the authors, the address, the title, the publication year, the author keywords and keyword plus, the abstract, the journal name, the Web of Science categories of the paper, the citations, and the references.

The articles from England, Northern Ireland, Scotland, and Wales have been reclassified as derived from the UK. The analysis of the collaboration patterns has been determined through the authors' address, and therefore, the phrase "single country article" has been assigned if the scientists' addresses were from the same country, while the term "internationally collaborated article" has been associated to the articles coauthored by researchers from multiple countries.

The traditional analysis of the data such as document typologies, languages, publication trends, subject categories, journals, authors, countries, institutions, and keywords has been made through the MS Excel software, the bibliographic coupling through *VOSviewer* software that builds distance-based maps (www.vosviewer.com; Waltman et al. 2010). Lastly, the network analysis of the international collaborations has been performed with the *IntColl.EXE* tool authored by Loet Leydesdorff (http://www.leydesdorff.net/software.htm, accessed 20 June 2014).

Results and discussion

Document typology and languages

The output of the search gives 3402 total publications during the time span of 23 years. Article, including articles published as proceeding papers or book chapters, was the most frequent document typology with 3126 records, amounting to about 92 % of the entire collection. Reviews reflected about 6 % (204) of the collected works. The other documents concerned Editorial materials (37; 1 %), news items (9; 0.3 %), corrections (7; 0.2 %), book reviews (5; 0.1 %), notes (5; 0.1 %), letters (4; 0.1 %), meeting abstracts (3; 0.1 %), and discussions (2; 0.1 %). Considering that the articles were predominant in the entire dataset, only this type of document has been used to perform the analysis discussed in the following sections. However, considering that most of the articles before 1991 do not include abstracts, we have restricted our analysis to the 1991–2013 period.

Table 1 Search terms used to extract the records from the SCI-EXPANDED and SSCI citation indexes

Search strings
Andreas fault*; San Andreas system; "Mill creek fault"; "Banning fault*"; "Punchbowl fault*"; "Nadeau fault*"; "Little Rock" fault*; "Mission Creek fault*"; "Garnet Hill" fault*; "Gandy Ranch" Fault*; "San Gorgonio Pass Fault*"; "San Gorgonio Fault*"; "Etsel Ridge fault*"; "Updegraff Ridge Fault*"; "Bartlett Springs thrust fault*"; "Bartlett Springs" fault*; " Hunting Creek-Berryessa fault*"; "Lake Mountain" fault*; "Talmadge fault*"; "Maacama fault*"; "Healdsburg fault*"; "Rodgers Creek fault*"; "Diablo Thrust fault*"; "Diablo fault*"; "Suisun fault*"; "Green Valley fault*"; "Paicines fault*"; "Calaveras fault*"; "San Benito" fault*; "Mission fault*" California; "Evergreen fault*"; "Quimby fault*"; "Crosley fault*"; "Clayton fault*" California Hayward fault*; "Coyote Creek fault*"; "Lytle Creek fault*"; "San Jacinto" fault*; "Glen Helen fault*"; "Rialto-Colton fault*"; "Claremont fault*"; "Casa Loma" fault*; "Clark fault*"; "Buck Ridge" fault*; "Coyote Mountain" fault*; "Superstition Hill" fault*; "Superstition Mountain" fault*; "Loma Linda" fault*; "Cucamonga fault*"; "Park Hill" fault*; "Anza Fault*"; "Horse Canyon" Fault*; "Santa Rosa Fault*"; "Arroyo Salada Fault*"; "San Gregorio Fault*"; "Seal Cove Fault*"; "Frijoles Fault*"; "Coastways Fault*"; "Greyhound Rock Fault*"; "Carmel Canyon Fault*"; "Denniston Creek Fault*"; "Año Nuevo Fault*"; "Sur Fault*"; "Palo Colorado Fault*"; "Palo Colorado-San Gregorio Fault*"; Imperial Fault*; "Elsinore Fault*"; "Main Street Fault*" California; "Chino Fault*" California; "Whittier Fault*"; "Fresno Fault*"; "Tin Mine" Fault*; "Main Street" Fault*; "Wildomar Fault*"; "Glen Ivy North Fault*"; "Glen Ivy South Fault*"; "Glen Ivy Fault*"; "Willard Fault*"; "Wolf Valley Fault*"; "Laguna Salada" Fault*; "Fresno-Eagle Fault*"; "Eagle Fault*"; "Earthquake Valley Fault*"; "San Felipe Fault*"; "San Felipe Valley Fault*"; Canyon Royo Fault*; "Murrieta Creek Fault*"; "Big Pine fault"; "Garlock fault"; "Pinto Mountain fault"; "Blue Cut fault"; "Fernando fault"; "Monica fault"; "Newport fault"; "Newport-Inglewood fault*"; "Inglewood fault"; "Cerro Prieto Fault*"; "Miguel fault*"

An overview of the publication language shows that English was widely predominant (3102 records, 99.2 %). This is consistent both with the fact that English is the predominant language in the scientific field and the SCI and SSCI index journals publish the full-text, or at least the abstract, in English (Gotze 1997). Moreover, the percentage of the articles published in English was particularly high because the San Andreas fault is located in the USA as well as the main institutions that investigated it, as we will see later in this paper.

Other four languages were represented in a very few documents, such as Chinese (15; 0.5 %), French (4; 0.1 %), Spanish (3; 0.1 %), and Russian (2; 0.1 %), which had a very low impact on the scientific community as emerges from the total number of citations of the above mentioned articles that, taken on the whole, did not reach the number of 50.

Publication trends and citations

In order to have a more detailed sight about the publications, we have analyzed the number of articles over the time. From Fig. 2 we can distinguish that the trend was characterized by two main periods. Indeed, after a linear growth (R^2=0.9993) in 13 years, from 1991 to 2003, where the total number of publications reached about 1200, the trend took an exponential form (R^2=0.9981) reaching the total number of outputs equal to 3126, that is 2.6 times of the total articles counted at the end of the linear time window. The exponential pattern roughly started when the number of scientists that contributed to the study of SAFS increased sharply (2002) (see Authors' productivity section and Fig. 6). This is probably due to the increase of the interdisciplinary studies on the subject, the significant development of new approaches and technologies to investigate SAFS, the increase of interests in some aspects of research, and the occurrence of earthquakes that stimulated new investigations, as we will see better in the next sections.

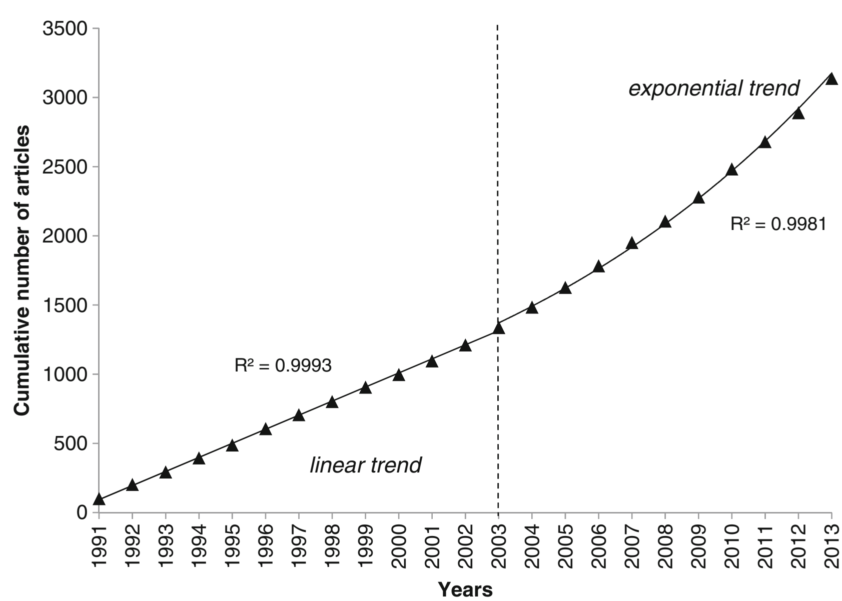
All the articles reached about 77,000 citations in Web of Science Core Collection, showing a trend that changed in quite a complex way over the time (Fig. 3). As a matter of fact, over the 23 years covered by our analysis, three different periods should be considered: 1991–2000, 2000–2005, and 2005–2013. For each of these, the number of citations fits a linear trend, but the rate of growth decreases as the publications become more recent. In particular, in the last part of the third period, especially 2011–2013, the number of citations decreased quickly because the time interval during which the most recent articles can be cited reduced progressively. This aspect can also be inferred from the number of the citations received from the articles published in 2012 and 2013 that, taken on the whole, did not reach 900 against the yearly average of about 3400 over the 23 years.

Subject categories and journals

The research on SAFS spanned over 59 ISI subject categories during the 1991–2013 period. *Geochemistry and Geophysics* was the most important category with 1971 items, covering 57 % of the entire categories. The classes that exceed 100 items were other three, *Multidisciplinary geosciences* (767; 22 %; second), *Geology* (201; 6 %; third), and *Multidisciplinary sciences* (114; 3 %; fourth).

These four categories covered about 88 % of the total from 1991 to 2013. As can be noticed from Fig. 4, the trend of the articles classified as belonging to *Multidisciplinary geosciences* and *Geology* categories was quite constant over time. Conversely, *Geochemistry and Geophysics* category had a sharp increase in the second part of the 23-year period under investigation. That increase seems acted as a guide of the

Fig. 2 Publication output over the 23-year time-window investigated in this paper



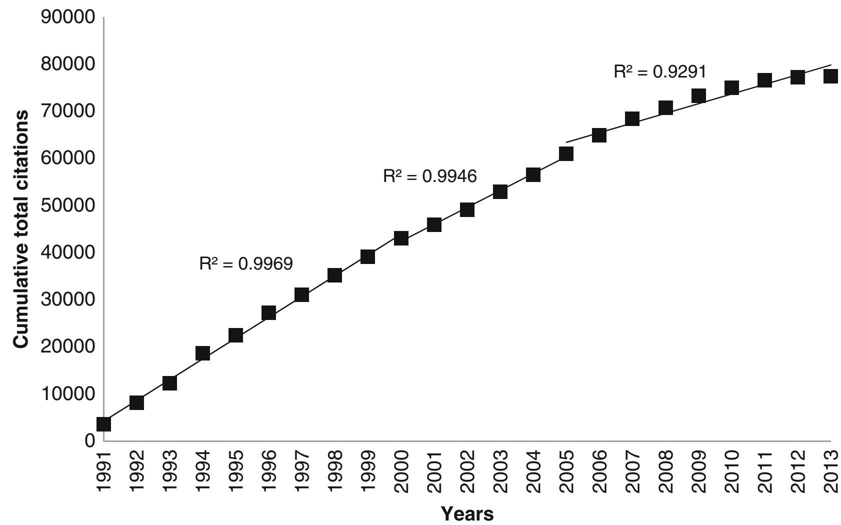


Fig. 3 Citation trends of the articles during 1991–2013

temporal trend observed in all the articles, as seen in the previous paragraph (see Fig. 2).

Seven (12 %) other categories that exceed ten items were *Engineering geology*, *Engineering-civil*, *Geography-physical*, *Water resources*, *Oceanography*, *Physic-multidisciplinary*, and *Mineralogy*. Finally, as many as 18 (30 %) were the categories that counted only one item.

Articles were published in 231 journals, but those having a number of published works that exceeded 100, covering about 66 % of the whole data-set were only eight (3.5 %), thus showing that the hottest research studies are expected to be published in a few high-impact journals (Table 2).

The *Journal of Geophysical research—Solid earth*—published most of the articles (683; 21.8 %) followed by the *Bulletin of the Seismological Society of America* (443;14.2 %) and the *Geophysical Research Letters* with 226 items (7.2 %). Figure 5, which reports the annual amount of publications for each of these three journals, shows that JGR always published the highest number of articles, except for 1991, 2002, 2006, and 2008 when BSSA was the first journal.

GRL was the journal with a more constant publication trend until 2003, after which two significant peaks were recorded (2004 and 2007). From the comparison of the three publication patterns it can be observed that a maximum (relative or absolute) publication trend of one journal frequently occurred when a minimum publication pattern of the other/s journal/s was recorded. This can also be due to the publication of special issues, such as that published by BSSA in 2002 focused on the 16 October 1999 $M=7.1$ Hector Mine earthquake.

Regarding the citations, *Journal of Geophysical research—Solid earth*—held the highest number (21600), but *Nature* (84.32) and *Science* (77.36) had the highest average citations for article.

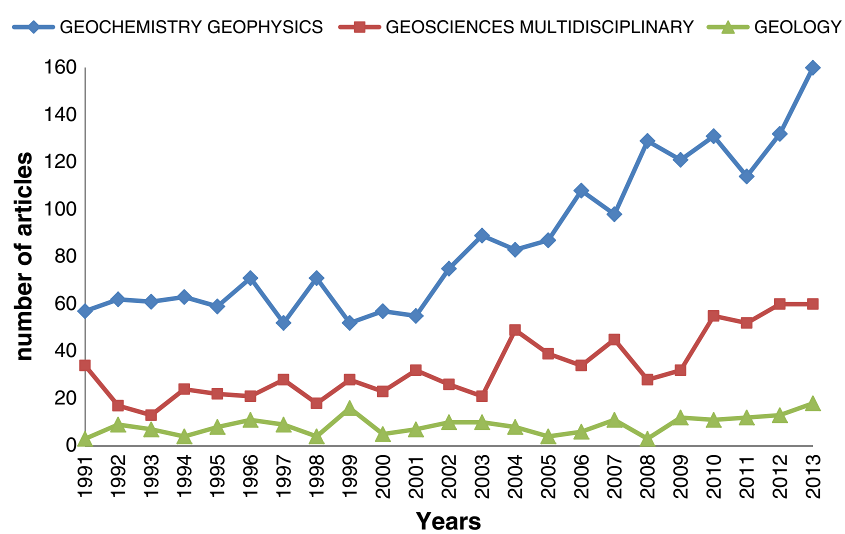


Fig. 4 Trend of publications of the three main WoS subject categories over the 1991–2013 time span

Table 2 Features of the top 30 most productive journals during 1991–2013

Journals	TA	TA(%)	TC	TC/TA	IF
Journal of Geophysical Research-Solid Earth	683	21.8	21600	31.63	3.174
Bulletin of the Seismological Society of America	443	14.2	11742	26.51	1.940
Geophysical Research Letters	226	7.2	4659	20.62	3.982
Geophysical Journal International	197	6.3	3482	17.68	2.820
Tectonophysics	138	4.4	2837	20.56	2.684
Geology	134	4.3	4942	36.88	4.087
Journal of Structural Geology	124	4.0	4167	33.60	2.285
Earth and Planetary Science Letters	113	3.6	1860	16.46	4.349
Geological Society of America Bulletin	80	2.6	1941	24.26	4.286
Tectonics	75	2.4	1867	24.89	3.487
Pure and Applied Geophysics	72	2.3	1377	19.13	1.617
Science	53	1.7	4100	77.36	31.027
Geochemistry, Geophysics, Geosystems	38	1.2	283	7.45	2.939
Nature	37	1.2	3120	84.32	38.597
Earth Planets and Space	30	1.0	229	7.63	2.921
International Geology Review	24	0.8	582	24.25	3.359
Physics of the Earth and Planetary Interiors	24	0.8	468	19.50	2.383
Geophysics	22	0.7	354	16.09	1.723
Geosphere	22	0.7	239	10.86	2.023
Earthquake Spectra	19	0.6	96	5.05	1.079
Geomorphology	19	0.6	183	9.63	2.552
Journal of Geodynamics	18	0.6	111	6.17	2.967
Journal of Geophys Research-Solid Earth and Planets	17	0.5	881	51.82	3.174
Seismological Research Letters	16	0.5	226	14.13	3.036
Chinese Journal of Geophysics-Chinese Edition	14	0.4	21	1.50	0.667
Lithosphere	14	0.4	32	2.29	2.169
Nature Geoscience	14	0.4	291	20.79	12.367
Annals of Geophysics	11	0.4	119	10.82	1.138
Terrestrial Atmospheric and Oceanic Sciences	11	0.4	74	6.73	0.705
Izvestiya, Physics of the Solid Earth	10	0.3	3	0.30	0.402

TA total articles, *TC* total citations, *IF* impact factor, 2012

Based on the analysis of the 2012 Journal Citation Report (JCR) Science edition, it emerges that among the top 30 journals, the most (28) had the impact factor (IF) equal or greater than 1; moreover, 44 % of these articles were

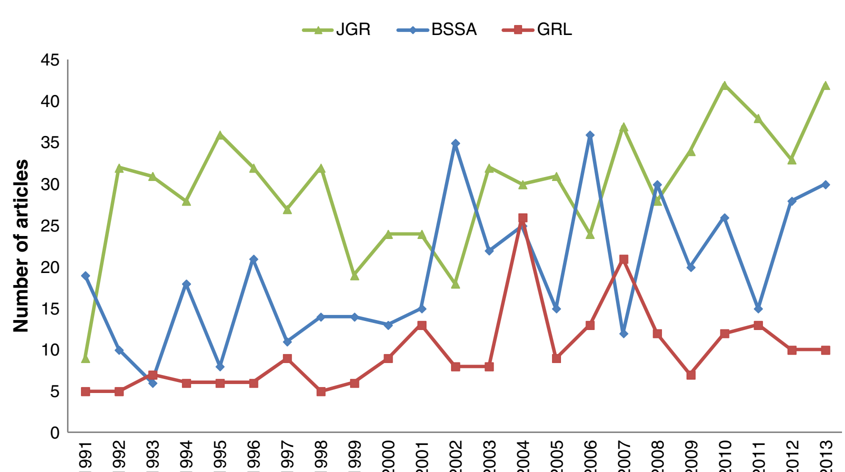


Fig. 5 Publications of the top three most productive journals during 1991–2013

published in journals having an IF greater than 3. The two journals with the highest IF were *Nature* (38.597) and *Science* (31.027), where 90 papers were published (2.9 %).

Among the other journals, *Geophysical Journal International* ranked fourth with 197 articles (6.3 %). Fifth to eighth positions were taken by *Tectonophysics* (138; 4.4 %), *Geology* (134; 4.3 %), *Journal of Structural Geology* (124; 4.0 %), and *Earth and Planetary Science Letters* (112; 3.6 %).

Productivity of the authors and their influence on the scientific community

The number of the single authors involved in the SAFS research was 5135, and the average of the collaboration index, that is the number of authors for article, was 3.3 in the overall time span considered.

The amount of the single-author number had an increase from 214, in 1991, up to 882, in 2013 (Fig. 6). In particular, the number of authors increased slightly until 2001, after which at least three sharp raises were recorded in 2002, 2010, and 2013, respectively.

The same analysis can be made for the total number of authors involved in the articles, whose trend was quite similar to that of the single authors, especially until 2001. The two patterns stayed comparable in the next years, but the increase in value difference from 2002 suggests more collaborations between the authors, as also indicated by the mean data of the collaboration index ranging from 2.31 in 1991 to 4.08 in 2013 (Fig. 7). However, this analysis can be influenced by problems affecting the authorship. As a matter of fact, two or more authors may have the same name or may use more than one name to sign their own articles (Tang and Walsh 2010).

The number of authors that contributed greatly to the SAFS research was a small percentage of the total researchers. The first 100 most prolific authors published about 54 % of the total items and the first 30 about 27 %, as can be seen from

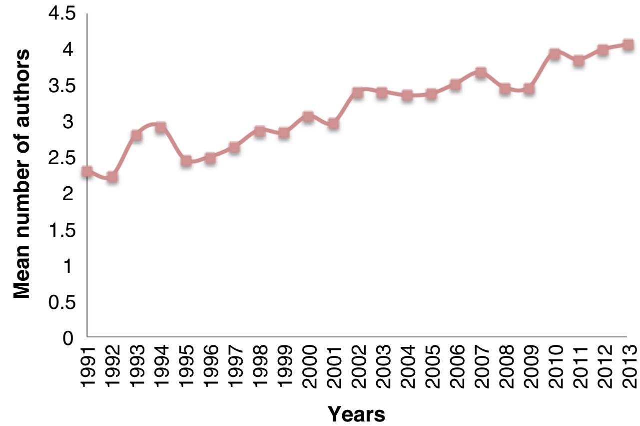


Fig. 7 Average of the number of authors for article involved in the SAFS research during the 1991–2013 period

Table 3. In order to face up the variation of personal spelling authorship problem, Table 3 (and Table 4) also shows the equivalence between different names used by the same author in his/her articles of the collection. To do this, we analyzed mainly the *Curriculum Vitae* (CV) of the authors under investigation available at the institutional web pages. For example, our collection of the articles on SAFS contains some papers published by Rockwell TK and others by Rockwell T. In order to make sure that the two different names conducted to the same author, we analyzed Rockwell TK's CV (available at http://www.earthconsultants.com/ resumes/Rockwell.pdf, accessed 11 August 2014) and we found that the articles authored by Rockwell TK and Rockwell T could be traced back to the same author. Therefore, the total number of the articles has been reported in Table 3 (and Table 4) under the double name attribution.

The most productive authors were Burgmann with 56 articles, closest followed by Ben-Zion with 52, Thurber with 43, Nadeau and Rockwell with 40 articles. The other authors followed with a number of articles that gradually tends to

Fig. 6 Number of authors involved in the SAFS research during the 1991–2013 period

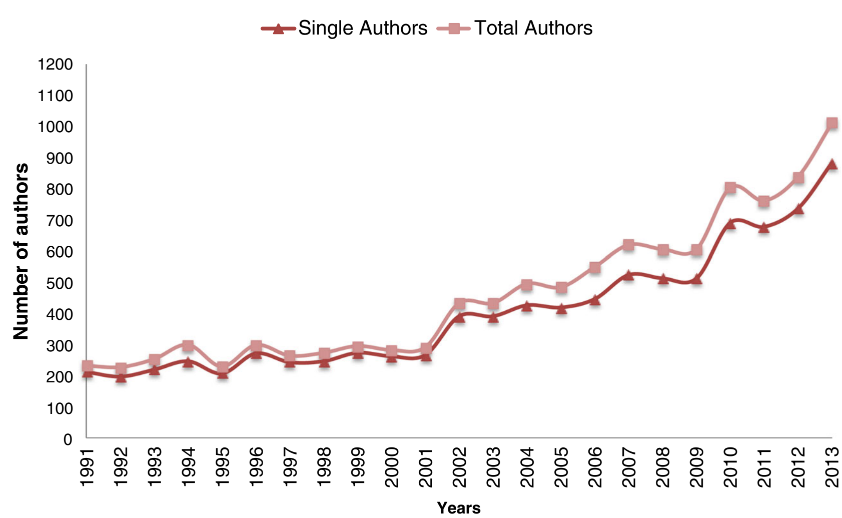


Table 3 Top 30 most productive authors involved in the research on SAFS during 1991–2013

Author	TA	SA	TC	TC/TA
Burgmann R	56	2	1816	32.4
Ben-Zion Y	52	3	1667	32.1
Thurber CH/C	43	2	1173	27.3
Nadeau RM/R	40	0	1660	41.5
Rockwell TK/T	40	0	668	16.7
Hauksson E	31	2	1443	46.5
Shearer PM/P	31	2	852	27.5
Arrowsmith JR/R	29	0	374	12.9
Parsons T	28	14	413	14.8
Zoback MD/M	28	1	1057	37.8
Savage JC	27	4	666	24.7
Segall P	26	1	1405	54.0
Simpson RW/R	26	0	1375	52.9
Ellsworth WL/W	25	0	2175	87.0
Peng ZG/Z	24	0	563	23.5
Zhang HJ/H	24	0	565	23.5
Furlong KP/K	23	1	469	20.4
Hardebeck JL/J	23	4	889	38.7
Fialko Y	22	4	644	29.3
Fuis GS/G	22	1	498	22.6
Li YG	22	1	669	30.4
Lockner DA/D	22	0	1210	55.0
Marone C	22	1	670	30.5
Pollitz FF/F	22	7	511	23.2
Rundle JB	22	1	357	16.2
Sandwell D/DT	22	0	402	18.3
K/KE	21	1	1000	47.6
Lienkaemper JJ	20	1	466	23.3
Sammis CG/C	20	0	1065	53.3
Wyss M	20	4	811	40.6

Authors in bold are also among top 30 most cited scientists (see Table 4)
TA total articles, *SA* single-author articles, *TC* total citations

Table 4 Top 30 most cited authors relating to the research on SAFS during 1991–2013

Author	TA	SA	TC	TC/TA
Stein RS/R	16	0	2345	146.6
Ellsworth WL/W	25	0	2175	87.0
Burgmann R	56	2	1816	32.4
Evans JP	18	0	1748	97.1
Lin J/JA	10	0	1725	172.5
Ben-Zion Y	52	3	1667	32.1
Nadeau RM/R	40	0	1660	41.5
King GCP/G	12	0	1616	134.7
Waldhauser F	14	0	1493	106.6
Hauksson E	31	2	1443	46.5
Segall P	26	1	1405	54.0
Eberhart-Phillips D	14	0	1383	98.8
Simpson RW/R	26	0	1375	52.9
Lockner DA/D	22	0	1210	55.0
Thurber CH/C	43	2	1173	27.3
Chester FM	10	1	1122	112.2
Sammis CG/C	20	0	1065	53.3
Zoback MD/M	28	1	1057	37.8
Reasenberg PA/P	7	0	1046	149.4
McEvilly TV	15	0	1014	67.6
Sieh K/KE	21	1	1000	47.6
Byerlee JD/J	13	2	957	73.6
Beroza GC	16	1	923	57.7
Rice JR	11	0	895	81.4
Hardebeck JL/J	23	4	889	38.7
Michael AJ/A	14	1	869	62.1
Jackson DD/D	15	0	867	57.8
Harris RA/R	10	1	857	85.7
Forster CB	4	0	856	214.0
Shearer PM/P	31	2	852	27.5

Authors in bold are also among top 30 most productive scientists (see Table 3)
TA total articles, *SA* single-author articles, *TC* total citations

20, the boundary of the 30th position. Most of these researchers collaborated with two or more authors. Very few were the articles with only one author (57, ~7 % of the list), and only one researcher stood out of the others for his high number of single-author papers (Parsons, 14).

About 47 % (14) of the top 30 most productive scientists were also among the most cited authors and as many as five occupied the first ten positions (Table 4). The number of the articles signed by only one scientist included in this list was still lower than those of the top 30 most productive author list (24, ~4 % against 57, ~7 %), confirming that collaborative papers are more likely to be cited.

Stein held the highest number of citations in the time span considered, while Forster held the highest average of citations of a paper of his (214) with a lower number of publications (4) on the list. Among the first three most cited authors, we can notice that the first was not included in the list of the most prolific authors, but his articles have a citation medium rate that fell among the first three of the most cited authors.

Analyzing the bibliographic coupling between the authors that signed or co-signed at least five articles, it can be noticed that the scientists can be grouped into 12 clusters, each of them with a different size ranging from 79 to 2 authors. Figure 8 shows this, also putting into evidence that the circle size is directly proportional to the number of the total articles of an author, the color represents the membership cluster, and the distance between each circle shows how strong the relationship with the other scientists is—the shorter is the distance, the

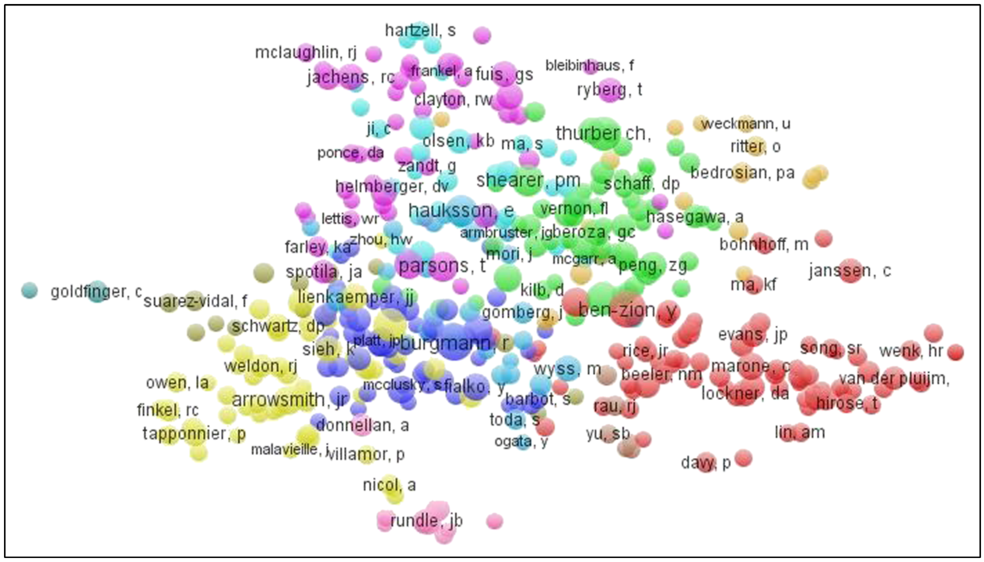


Fig. 8 Bibliographic coupling of the authors having a minimum number of articles equal to 5

higher is the connection. The first and the third wider clusters included the first two most productive authors—Burgmann (the third cluster with 59 authors) and Ben-Zion (the widest cluster, 79 authors). The map shows that the authors were grouped into clusters where they are close to each other, thus underlying that there is a good statistical probability that they treated a quite similar subject matter.

Most cited articles

The most cited article was *Static stress changes and the triggering of earthquake* published by King GCP et al., in 1994 in the Bulletin of the Seismological Society of America, and was cited 813 times (from Web of Science Core Collection) in the 1991–2013 time span. Observing the yearly-citations trend of this paper (Fig. 9), we notice that a sharp increase was recorded when the article had 2 years of life, and the number of citations increased constantly until the seventh year. The highest yearly value of citations (74) was reached 17 years after the publication.

The second most cited article (728) was *A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California* by Waldhauser, F and Ellsworth, WL published in 2000 in the Bulletin of the Seismological Society of America, too. This article reached rapidly a high number of yearly citations and the highest values respectively 3 and 13 years after the publication.

The third (575 Citations) and the fourth (428 Citations) articles in the rank were *Fault zone architecture and permeability structure* by Caine et al., published in the Geology in 1996 and *Internal structure and weakening mechanism of the San Andreas Fault* by Chester et al., published in the Journal of Geophysical Research in 1993. All the four articles retained a high number of yearly citations testifying that the attention of the researchers to these issues was constant over the whole time span taken as a reference here.

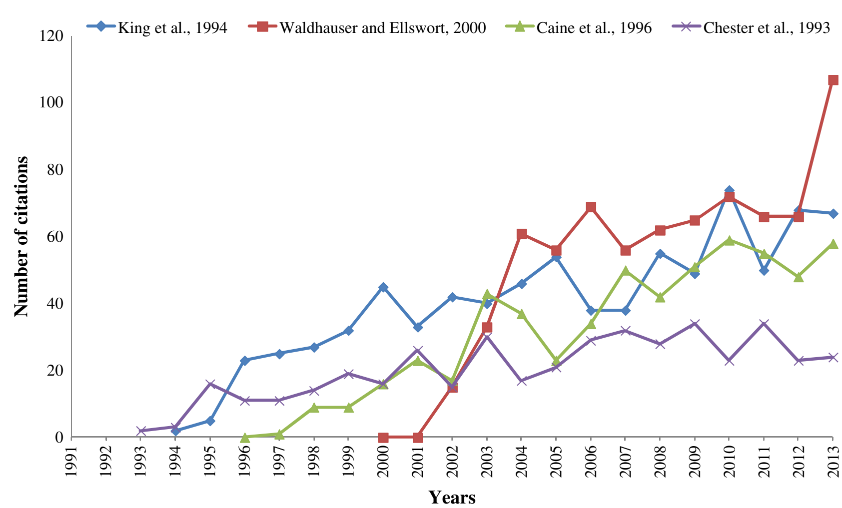


Fig. 9 The four most cited articles on the SAFS during 1991–2013. The data for each article starts from the year of publication

Where the publications were born: country and institution analysis

Among 3126 articles, 319 (10.2 %) have no author address information, and therefore, these data have been excluded both from the analysis of the publication country and the institutional distribution of research outputs.

Seventy-three countries/territories all around the world were involved in the SAFS research in the 1991–2013 period. Taking into consideration the most productive top 30 countries (Table 5), one can see that four continents were active in the SAFS research. Europe topped the rank with 13 countries, Asia counted 8 countries, America 5, and Oceania 2, while Africa was not ranked.

The country that held the leadership was the USA, with 1937 records (69.0 %). Among the European countries, France occupied the second place but significantly distant from the first, signing 234 papers (8.3 %). Considering the countries with a minimum number of 100 records, it emerges that France was followed by Japan with 209 articles (7.4 %), and Germany (147; 5.2 %), Italy (122; 4.3 %), China (120; 4.3 %), and the UK with 108 articles (3.8 %).

Table 5 Top 30 countries/territories involved in the research on SAFS during 1991–2013

Countries/territories	TA	TA(%)	IC	IC(%)	TC	TC/TA
USA	1937	69.0	464	24.0	52265	27.0
France	234	8.3	164	70.1	4231	18.1
Japan	209	7.4	92	44.0	3698	17.7
Germany	147	5.2	109	74.1	1659	11.3
Italy	122	4.3	70	57.4	2331	19.1
People's R China	120	4.3	78	65.0	1787	14.9
UK	108	3.8	65	60.2	2458	22.8
Canada	87	3.1	54	62.1	1354	15.6
Taiwan	74	2.6	52	70.3	827	11.2
New Zealand	60	2.1	56	93.3	1325	22.1
Switzerland	52	1.9	45	86.5	1789	34.4
Mexico	46	1.6	33	71.7	703	15.3
Australia	44	1.6	32	72.7	667	15.2
Turkey	44	1.6	36	81.8	1479	33.6
India	32	1.1	8	25.0	319	10.0
Norway	31	1.1	22	71.0	538	17.4
Israel	29	1.0	22	75.9	438	15.1
Spain	28	1.0	20	71.4	213	7.6
Greece	26	0.9	14	53.8	310	11.9
Netherlands	25	0.9	20	80.0	556	22.2
Russia	24	0.9	12	50.0	204	8.5
Iran	14	0.5	10	71.4	99	7.1
South Korea	12	0.4	7	58.3	243	20.3
Ireland	11	0.4	6	54.5	188	17.1
Chile	9	0.3	9	100.0	98	10.9
Brazil	8	0.3	3	37.5	111	13.9
Indonesia	7	0.2	7	100.0	143	20.4
Austria	6	0.2	3	50.0	94	15.7
Belgium	6	0.2	2	33.3	22	3.7
Denmark	5	0.2	0	0.0	81	16.2

TA total articles, *IC* internationally collaborated articles, *TC* total citations

The USA collaborated with the highest number of countries (64) even if the percentage of the international collaborated articles was the lowest (24.0 %). Other countries had an international collaborative percentage that ranged from 33 % of Belgium to 100 % of Chile and Indonesia, with a high mean of ~64 %. For example, France and Germany, which signed internationally collaborated articles with 42 and 36 countries respectively, had a percentage of multi-country items, ranging from 70 to 74 %. To have a clearer picture of these aspects, Fig. 10 shows the complex network of international collaborations between all 73 countries. All these data imply that the SAFS research calls for teamwork between countries, with a specific reference to the USA.

The analysis of the institution contribution has been performed considering the affiliation of at least one author of the articles. As said before, there were 319 articles without institution information. Therefore, the total number of articles considered for this analysis was 2807. Furthermore, the examination has left out of the investigation the subdivision of the institutions in branches.

The top 30 institutions active in the past 23 years are listed in Table 6 from which it emerges that the SAFS research studies were performed mainly by the institutions having the seat in the USA. As a matter of fact, among the first 30 most prolific worldwide institutions, as many as 22 (73.3 %) were from the USA. The most prolific institution was the *US Geological Survey (USGS)* with 495 articles (17.6 %) published in the 1991–2013 period. The *California Institute of Technology* (*Caltech)* followed with 163 (5.8 %) and five other universities of California. The first non-USA institution was the Italian *Istituto Nazionale di Geofisica e Vulcanologia (INGV)* that occupied the eighth position with 63 papers (2.2 %). Among the other EU-country institutions, we single out the *University of Grenoble 1* (France) ranked 16th with 47 articles (1.7 %) and the *GeoForschungsZentrum* (GFZ) of Potsdam (Germany) ranked 19th with 41 papers (1.5 %), while the first Asiatic institution was the *University of Tokyo* (Japan) that occupied the 10th position with 59 publications (2.1 %). The highest mean value of citations for article was held by *Harvard University* (48.1) followed by *Columbia University* (40.5), while the mean of the citations for article, considering all the 30 institutions, was about 25.

Author keywords

According to the previous experiences, the author keywords can help to get the trend regarding a specific research field (Chiu and Ho 2007; Wang et al. 2014). Therefore, we performed the

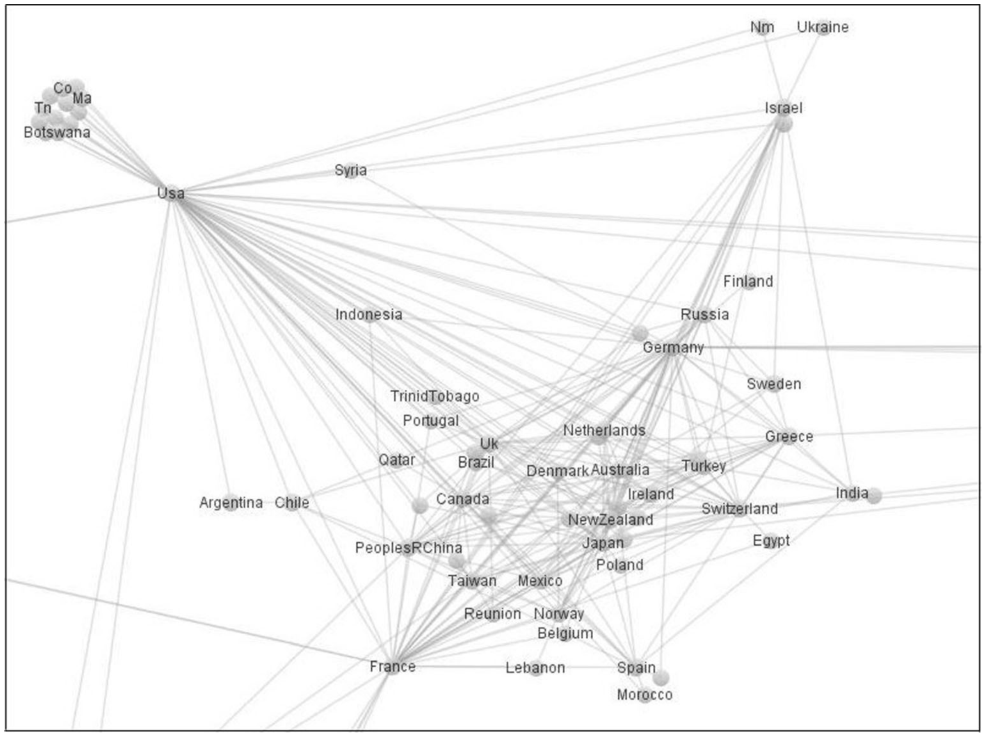


Fig. 10 Collaboration network between the countries during 1991–2013. The elaboration has been performed for all 73 countries using the *IntColl.exe* tool, and the visualization of the network has been obtained through the VOSviewer software (see text). The figure zooms on the core of the network

analysis of the author keywords to try to show up the most significant issues dealt with in the articles over the years. Therefore, we have considered the changes of the keywords in four 5-year periods, starting from 1994. On 2845 total articles, 1134 (39.9 %) included one or more keywords. In detail, 724 (63.8 %) of these had a number of keywords between five and twelve; the remainder (410) publications had a number varying from 1 and 4.

We have identified 3037 unique author keywords, with a total of 5632 occurrences. Few (68, 2.2 %) were the words that appeared more than or equal to ten times. Two thousand and three hundred forty (77.0 %) were the words that occurred only once showing a possible episodic nature of some research activities or the lacking of standardization of the terms chosen by the authors (Chuang et al. 2007).

In order to get a closer view, we have selected the 40 (1.3 %) most frequent keywords that, taken on the whole, occurred 932 times, equal to 16.5 % of the total occurrences.

The 40 most frequent keywords can be divided into two main groups: those reflecting the *tectonic* studies and those that can be referred to as *seismological* works (Table 7). The former can be identified, for example, by the words *fault(s)*, *faulting*, *dynamic and mechanic of faulting*, *active fault*, *fractures and faults*, *fault gouge*, the latter by *earthquake dynamics*, *wave propagation*, *body waves*, *paleoseismology*, *seismic anisotropy*, and *seismic hazard*. A third category regards the remote-sensing tools such as *InSAR* or *GPS*.

However, the keywords had a different importance over the years. As a matter of fact, excluding *San Andreas fault*, *earthquake/s*, *seismicity*, and *friction*, that retained high positions in the rank all over the four temporal windows considered, most of the top 40 keywords were not classified nor had little relevance in the 1994–1998 period. This can be due, for example, to the fact that some studies had pioneering feature at that time. For example, the *InSAR* technique that was not ranked in the 1994–1998 period was used for the first time in the 1960s to observe the planets of the solar system (Rogers and Ingalls 1969). The 1980s saw the applications of InSAR to get DEMs of the Earth's surface, but the technique was improved only in the 1990s. Just in these years, there was a strong rise of the InSAR applications, among which those in the tectonic field that were inaugurated with the analysis of the co-seismic displacement brought about by the 1992 Landers earthquake (Massonnet et al. 1993; Goldstein et al. 1998). A few years after that earthquake and from the end of the 1990s, InSAR was considered more extensively for the study of SAFS and was also used to monitor creep and strain accumulation (e.g., Lyons and Sandwell 2003).

Other words included in the list were *paleoseismology*, *GPS*, *tomography*, *magnetotellurics*, *seismic anisotropy*, *wave propagation*, and *SAFOD*. The GPS, jointly with the InSAR technique or seismicity-derived stress orientations were used, for example, to get data relating to the vertical motions in the San Francisco Bay Area or the fault slip rates of the tectonic lineaments in the southern California plate-boundary region (Becker et al. 2005; Burgmann et al. 2006). The second half of the last decade also saw the start of the publication of the studies related to the activity of the *San Andreas Fault*

Table 6 Top 30 institutions involved in the research on SAFS during 1991–2013

Institutions	TA	TA(%)	TC	TC/TA
USGS, USA	495	17.6	14630	29.6
Caltech, USA	163	5.8	5062	31.1
University of Southern California, USA	146	5.2	3907	26.8
University of California, Berkeley, USA	137	4.9	3804	27.8
Stanford University, USA	136	4.8	3725	27.4
University of California, San Diego, USA	105	3.7	3145	30.0
University of California, Los Angeles, USA	79	2.8	2230	28.2
INGV, Italy	63	2.2	867	13.8
State University of San Diego, USA	63	2.2	1221	19.4
University of Wisconsin, USA	61	2.2	1400	23.0
University of Tokyo, Japan	59	2.1	909	15.4
University of California, Santa Barbara, USA	58	2.1	2141	36.9
University of Nevada, USA	58	2.1	1504	25.9
Columbia University, USA	53	1.9	2147	40.5
China Earthquake Administration, China	51	1.8	545	10.7
MIT, USA	50	1.8	1750	35.0
University of California, Davis, USA	48	1.7	905	18.9
University of Grenoble 1, France	47	1.7	687	14.6
Pennsylvania State University, USA	45	1.6	975	21.7
Lawrence Livermore National Laboratory, USA	44	1.6	1442	32.8
University of Colorado, USA	44	1.6	1158	26.3
GFZ, Potsdam, Germany	41	1.5	838	20.4
Kyoto University, Japan	37	1.3	723	19.5
Oregon State University, USA	36	1.3	691	19.2
Arizona State University, USA	35	1.2	511	14.6
Harvard University, USA	34	1.2	1637	48.1
Insti Phys Globe, France	33	1.2	896	27.2
University of Oregon, USA	33	1.2	788	23.9
University Montpellier 2, France	32	1.1	454	14.2
University of California Santa Cruz, USA	30	1.1	938	31.3

TA total articles, *TC* total citations

Observatory at Depth (SAFOD) whose drilling activities began in 2004 and ended in 2007 (http://www-icdp.icdp-online.org/front_content.php?idcat=712, accessed on 14 July 2014). The investigation results regarded the rock or fluid sample age or physical, geochemical, mineralogical features, or the velocity model identification around the SAFOD site placed close to the town of Parkfield (e.g., Bradbury et al. 2007; Wiersberg and Erzinger 2007; Zhang et al. 2009; Janssen et al. 2011; Spencer et al. 2012).

Hot issues

As said before, only 39.9 % of the articles had the author keywords, and still a smaller percentage of the articles had a number of keywords sufficient to represent their content. This can lead to terms indexing problems (Garfield 1990). Therefore, in order to overcome these limitations and capture the hot issues and main research trends, we have performed a joint analysis of the author keywords, keywords plus, and title words thus supplying the top 40 most frequent term list (Table 8). As for the author keywords, we have considered the changes of the terms in four 5-year periods, starting from 1994 so as to be able to perform a comparison between the two groups of terms.

The terms *Slip (5th)*, *Parkfield* (12th), *Basin* (24th), *Hayward fault* (30th), *heat-flow* (39th), *Punchbowl fault* (36th) were some of the words that got in the list, while *paleoseismology*, *InSAR*, *magnetotelluric*, and *SAFOD* were some of the author keywords that lost significance and, consequently, went out of the list. *Slip* retained a high importance in all the four 5-year time windows analyzed as well as *Parkfield* that reached the highest position (11th) in the 2004–2008 period. These terms also coincided with the hot issues dealt with by the three most cited authors. As a matter

Table 7 Top 40 keywords used by authors in their own articles during 1994–2013 and four 5-year periods

	Total		1994–1998		1999–2003		2004–2008		2009–2013	
Word	Cnts	Rank	Cnts	Rank	Cnts	Rank	Cnts	Rank	Cnts	Rank
San Andreas fault	92	1	8	1	20	1	34	1	30	1
Earthquakes	59	2	5	2	18	2	19	2	17	4
Seismicity	45	3	5	2	11	3	11	4	18	3
Earthquake	29	4	3	4	6	7	9	6	11	9
Faults	28	5	4	3	5	8	12	3	7	13
Friction	28	5	2	5	6	7	10	5	10	10
Crustal structure	27	6	2	5	3	10	7	8	15	6
Fault	26	7	–	–	8	6	7	8	11	9
GPS	26	7	–	–	4	9	12	3	10	10
Paleoseismology	26	7	–	–	10	4	11	4	5	15
Tectonics	26	7	–	–	9	5	9	6	8	12
Faulting	25	8	2	5	8	6	7	8	8	12
Seismic tomography	24	9	–	–	4	9	4	11	16	5
Tomography	24	9	2	5	3	10	9	6	10	10
Fault zone	22	10	1	6	6	7	9	6	6	14
Seismicity and tectonics	22	10	–	–	–	–	–	–	22	2
California	21	11	1	6	11	3	4	11	5	15
Crustal deformation	21	11	–	–	5	8	12	3	4	16
Dynamics and mechanics of faulting	20	12	–	–	–	–	4	11	16	5
Fractures and faults	19	13	–	–	–	–	1	14	18	3
INSAR	19	13	–	–	4	9	7	8	8	12
Magnetotellurics	19	13	–	–	1	12	10	5	8	12
Earthquake dynamics	18	14	–	–	–	–	3	12	15	6
Seismic anisotropy	18	14	1	6	1	12	6	9	10	10
Wave propagation	18	14	–	–	2	11	2	13	14	7
Body waves	17	15	–	–	2	11	1	14	14	7
Seismic hazard	17	15	1	6	3	10	9	6	4	16
Strike-slip fault	17	15	–	–	2	11	10	5	5	15
Active fault	16	16	1	6	2	11	7	8	6	14
Creep	16	16	2	5	5	8	4	11	5	15
Permeability	16	16	–	–	4	9	4	11	8	12
Earthquake interaction	15	17	–	–	2	11	2	13	11	9
Geodesy	15	17	–	–	–	–	8	7	7	13
Slip rate	15	17	–	–	2	11	9	6	4	16
Stress	15	17	–	–	6	7	6	9	3	17
Transform faults	15	17	2	5	–	–	4	11	9	11
Fault gouge	14	18	2	5	2	11	3	12	7	13
Fluids	14	18	–	–	2	11	7	8	5	15
Mechanics	14	18	–	–	2	11	2	13	10	10
SAFOD	14	18	–	–	–	–	2	13	12	8

Cnts, Counts

of fact, Stein RS cosigned a paper investigating the *slip* caused by the 1989 Loma-Prieta earthquake comparing the leveling surveys performed before and after the quake itself (Marshall et al. 1991). The same author analyzing the 1983 M=6.5 Coalinga and M=6.0 Nuñez earthquakes inferred about the shear and Coulomb stress change induced on the *Parkfield* segment so as to calculate the probability of an event with a magnitude similar to six Parkfield earthquakes during the 2001–2011 period (Toda and Stein 2002). Moreover Ellsworth WL, taking

Table 8 Top 40 words as derived by the elaboration of author keywords, keywords plus, and title words during 1994–2013 and four 5-year periods

	Total		1994–1998		1999–2003		2004–2008		2009–2013	
Word	Cnts	Rank	Cnts	Rank	Cnts	Rank	Cnts	Rank	Cnts	Rank
California	1457	1	297	1	315	1	400	1	445	2
San-Andreas fault	1401	2	235	2	260	2	349	2	557	1
Fault	1183	3	194	3	224	3	329	3	436	3
Earthquake	890	4	143	4	160	4	270	4	317	4
Slip	603	5	100	8	112	6	163	5	228	5
Zone	563	6	84	10	110	8	158	6	211	6
Deformation	549	7	107	6	111	7	147	8	184	7
Southern(−)California	519	8	104	7	127	5	135	9	153	10
Earthquakes	456	9	94	9	87	10	121	10	154	9
Stress	431	10	117	5	87	10	102	13	125	12
Andreas	347	11	46	18	62	12	110	12	129	11
Parkfield	334	12	48	17	50	17	116	11	120	13
System	304	13	42	21	67	11	92	14	103	15
Seismic	293	14	49	16	51	16	85	15	108	14
Seismicity	272	15	58	13	48	18	79	16	87	19
Model	260	16	51	15	54	14	74	18	81	20
Evolution	258	17	52	14	56	13	62	22	88	18
Rupture	256	18	44	19	47	19	68	19	97	16
Faults	236	19	43	20	39	21	57	23	97	16
Tectonics	201	20	51	15	52	15	54	24	44	34
Structure	185	21	25	32	39	21	65	20	56	25
Implications	174	22	38	24	32	23	34	35	70	21
Zones	167	23	40	22	32	23	43	29	52	28
Basin	163	24	36	25	41	20	41	30	45	33
Region	161	25	27	31	29	26	51	26	54	26
Inversion	155	26	39	23	30	25	43	29	43	35
Strain	146	27	44	19	27	28	36	33	39	38
Recurrence	144	28	24	33	26	29	53	25	41	36
Models	143	29	22	35	30	25	34	35	57	24
Hayward fault	141	30	–	–	15	39	63	21	63	22
Strike-slip	139	31	17	39	27	28	39	32	56	25
Friction	138	32	20	36	28	27	40	31	50	30
Crustal	135	33	29	29	27	28	32	37	47	32
Velocity	130	34	13	43	22	32	43	29	52	28
Japan	126	35	15	41	16	38	45	28	50	30
Shear	126	35	23	34	34	22	30	39	39	38
Punchbowl fault	124	36	3	53	28	27	49	27	44	34
Valley	118	37	32	26	28	27	33	36	25	51
Motion	117	38	20	36	32	23	39	32	26	50
Heat-flow	116	39	28	30	14	39	34	35	40	37

Cnts, Counts

advantage of the high-quality measurements obtained after the 28 September 2004 *Parkfield* earthquake, dealt with new earthquake models in order to better estimate the size and the location of the future ground shaking (Bakun et al. 2005). Finally, Burgman R put into evidence, for example, how powerfully the fault interaction can influence the *slip* gradients (Burgmann et al. 1994).

As concerns the *Hayward fault*, one can show that it gained importance during the 1999–2003 period, when it ranked 39th and in the subsequent two periods when it was classified at least 22nd. The same analysis can be made for the *Punchbowl fault*, against the trend that *inversion* had, which diminished in importance over the time.

A closer analysis of the temporal changes of some of the new terms introduced by the joint analysis can be gained from Fig. 11, from which one can see that the trend is quite constant for all the words, except for *Hayward fault* whose related papers gained weight from 2002, and *slip* and *Parkfield* which had an increase after 2004 and 2003, respectively.

Regarding the terms excluded from the list, *paleoseismology* ranked 100th over the entire period. Its rank was 48th in the 1994–1998 period, during which the term was not indexed in the author keyword top ranked list, increasing slightly in significance in the 1999–2003 period and gaining the 43rd position. In these years, the paleo-seismological studies were performed, for example, on the San Andreas fault or its branches in order to determine the earthquake average recurrence interval (e.g., Dolan et al. 2000; Dawson et al. 2003). The last two time-windows saw a progressive light decrease of this research field down to the 55th position of the 2009–2013 period, confirming partially the general trend that appeared from the only analysis of the author keywords. In its turn, *InSAR* ranked 98th in the multi-term analysis list against 13th as derived from the author keyword analysis. The term gained importance in the last 15 years when it passed progressively from the 50th position of the 1999–2003 period to the highest position (27th) recorded in the 2009–2013 period.

Analogous considerations can be made for *SAFOD* that ranked 80th on the joint-term analysis list. The word had an increase up to the 41th position in the 2004–2008 period with an increase in the ranking (29th) in the 2009–2013 period, similarly to the outputs of the analysis restricted only to the author keywords.

Conclusions

The paper has presented an overview of the global research on the San Andreas Fault System during the 1991–2013 period based on the Thomson Reuters' Science Citation Index Expanded (SCI-EXPANDED) and the Social Sciences Citation Index (SSCI) accessed via the Web of Science Core Collection. The aspects analyzed have been the typology and languages, the journals, the subject categories, the authors, the articles, the countries, the institutions, and the hot issues.

The study has highlighted that 3126 articles were published. The publication trend over time indicates that the increase in publication number occurred since 2003, when an exponential tendency replaced the linear growth. The exponential pattern, which roughly started when the number of researchers that contributed to the studies on SAFS increased sharply, if used as future forecast may suggest that about 2000 articles will be published until 2020.

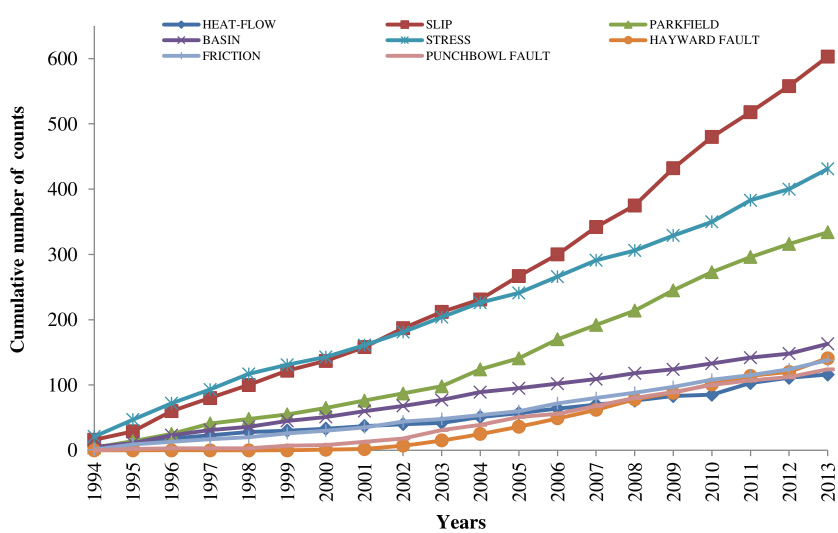
The *Geochemistry and Geophysics* was the most important subject category covering 57 % of the entire categories. With regard to the journals, the *Journal of Geophysical research—Solid earth*—published most of the articles followed by the *Bulletin of the Seismological Society of America* and the *Geophysical Research Letters*.

The most cited article was *Static stress changes and the triggering of earthquake* published by King GCP et al., in 1994 in the Bulletin of the Seismological Society of America.

Seventy-three countries/territories all around the world contributed to the SAFS research in the 1991–2013 period. The country that held the leadership was the USA that also collaborated with the highest number of countries.

Regarding the institutions, there was a clear predominance of those having the seat in the USA. The most prolific institution was the *US Geological Survey (USGS)* followed by the *California Institute of Technology* (*Caltech)*.

Fig. 11 Visualization of some of the most significant terms ranked in the top 40 list obtained through the joint analysis of the author keywords, keywords plus, and title words (total data as in Table 8)



The hottest and most significant items, as derived from the joint analysis of author keywords, keywords plus, and title words, were *earthquake/s*, *fault*, *slip*, *deformation*, *stress*, *Parkfield*, *model*, *evolution*, and *rupture*, which maintained high significance all over the time span investigated. Furthermore, a change in research perspective seems to be revealed by the increase of interest for the seismic hazard studies over the last 10 years especially for the Parkfield area hit by the significant 28 September 2004 earthquake. The last decade also saw major attention toward the study both of the Punchbowl fault, an inactive strand of the San Andreas, and the Hayward fault, a significant element of SAFS. Similarly, the increase of interest in the same time span was also dedicated by the scientists to the analysis of the velocity structure of the fault zones and crust/mantle in California, without neglecting the investigation of the changes in seismic velocity caused by earthquakes, such as the variations brought about by the 2004 Parkfield seismic event on the San Andreas Fault. All that may suggest that the topics above mentioned will be among the mainstreams of the future research studies, too.

Summarizing, the outcome of this study can contribute both to get an insight into the changes made in the SAFS study over the time and supply clues about the research trends so as to better address the future efforts improving the knowledge of an area, such as California, characterized by a high seismic hazard. In addition, scientists can find this study outputs useful to build a network of partners that complement each of their own strength to participate in research projects, choosing colleagues and institutions/countries that are active in performing research on SAFS. Besides, the paper results can influence the publication strategies, supplying a backed to evaluate where to publish in order to obtain high visibility. Finally, considering all the aspects dealt with, the paper can supply a support to scientists that wish to go aboard on work in this research area or perform a review about the subject.

References

Bakun WH, Aagaard B, Dost B, Ellsworth WL, Hardebeck JL, Harris RA, Ji C, Johnston MJS, Langbein J, Lienkaemper JJ, Michael AJ, Murray JR, Nadeau RM, Reichle MS, Reasenberg PA, Roeloffs EA, Shakal A, Simpson RW, Waldhauser F (2005) The 2004 Parkfield (CA) earthquake: implications for prediction and hazard assessment. Nature 437:969–974

Becker TW, Hardebeck JL, Anderson G (2005) Constraints on fault slip rates of the southern California plate boundary from GPS velocity and stress inversions. Geophys J Int 160:634–650. doi:10.1111/j.1365246X.2004.02528.x

Black BD, Hecker S, compilers (1999) Fault number 2470, Elsinore fault (fold), in Quaternary fault and fold database of the United States: U.S. Geological Survey website, http://earthquakes.usgs.gov/hazards/qfaults. Accessed 23 Jul 2014

Bradbury KK, Barton DC, Solum JG, Draper SD, Evans JP (2007) Mineralogic and textural analyses of drill cuttings from the San Andreas Fault Observatory at Depth (SAFOD) boreholes: initial interpretations of fault zone composition and constraints on geologic models. Geosphere 3(5):299–318. doi:10.1130/GES00076.1

Bryant WA, compiler (2000) Fault number 29a/c, Bartlett Springs fault system, Bartlett Springs section, in Quaternary fault and fold database of the United States: U.S. Geological Survey website, http://earthquakes.usgs.gov/hazards/qfaults. Accessed 23 Jul 2014

Bryant WA, Matthew LM, compilers (2002) Fault number 1a-j, San Andreas fault zone, all sections, in Quaternary fault and fold database of the United States: U.S. Geological Survey website, http://earthquakes.usgs.gov/hazards/qfaults. Accessed 23 Jul 2014

Bryant WA, Cluett SE, compilers (1999a) Fault number 54a/b/c/d, Calaveras fault zone, Central Calaveras fault section, in Quaternary fault and fold database of the United States: U.S. Geological Survey website, http://earthquakes.usgs.gov/hazards/qfaults. Accessed 23 Jul 2014

Bryant WA, Cluett SE, compilers (1999b) Fault number 60a/b, San Gregorio fault zone, Sur Region section, in Quaternary fault and fold database of the United States: U.S. Geological Survey website, http://earthquakes.usgs.gov/hazards/qfaults. Accessed 23 Jul 2014

Bryant WA, Cluett SE, compilers (2000) Fault number 55a/b/c, Hayward fault zone, Southeast Extension section, in Quaternary fault and fold database of the United States: U.S. Geological Survey website, http://earthquakes.usgs.gov/hazards/qfaults. Accessed 23 Jul 2014

Bryant WA, Cluett SE, compilers (2002) Fault number 37, Green Valley fault, in Quaternary fault and fold database of the United States: U.S. Geological Survey website, http://earthquakes.usgs.gov/hazards/qfaults. Accessed 23 Jul 2014

Burgmann R, Pollard DD, Martel SJ (1994) Slip distributions on faults: effects of stress gradients, inelastic deformation, heterogeneous hostrock stiffness, and fault interaction. J Struct Geol 16:1675–1690

Burgmann R, Hilley G, Ferretti A, Novali F (2006) Resolving vertical tectonics in the San Francisco Bay area from permanent scattered InSAR and GPS analysis. Geology 34:221–224

Chiu WT, Ho YS (2007) Bibliometric analysis of tsunami research. Scientometrics 73:3–17

Chuang KY, Huang YL, Ho YS (2007) A bibliometric and citation analysis of stroke-related research in Taiwan. Scientometrics 72(2):201–212

Dawson TE, McGill SF, Rockwell TK (2003) Irregular recurrence of paleoearthquakes along the central Garlock Fault near El Paso Peaks, California. J Geophys Res 108(B7):2356. doi:10.1029/2001JB001744

Dolan JF, Sieh K, Rockwell TK (2000) Late Quaternary activity and seismic potential of the Santa Monica fault system, Los Angeles, California. Geol Soc Am Bull 112:1559–1581

Ellsworth WL (1990) Earthquake history, 1769-1989. In Wallace RE (Ed.) The San Andreas fault system, California. (pp. 153–190). U.S. Geological survey professional paper 1515. Washington

Franceschini F, Maisano D (2010) A survey of quality engineering-management journals by bibliometric indicators. Qual Reliab Eng Int 26(6):593–604. doi:10.1002/qre.1083

Garcia-Ramon MD, Caballé A (1998) Situating gender geographies: a bibliometric analysis. Tijdschr Econ Soc Geogr 89(2):210–216

Garfield E (1990) KeyWords plus: ISI's breakthrough retrieval method, part 1, expanding your searching power on current contents on diskette. Curr Contents 32:5–9

Goldstein R, Zebker H, Werner C (1998) Satellite radar interferometry: two-dimensional phase unwrapping. Radio Sci 23:713–720

Gotze H (1997) The English language in scientific publishing. Publ Res Q 13(1):52–72

Grant J, Cottrell R, Cluzeau F, Fawcett G (2000) Evaluating 'payback' on biomedical research from papers cited in clinical guidelines: applied bibliometric study. Br Med J 320. doi:10.1136/bmj.320.7242.1107

Hart EW, compiler (1998) Fault number 32, Rodgers Creek fault, in Quaternary fault and fold database of the United States: U.S. Geological Survey website, http://earthquakes.usgs.gov/hazards/qfaults. Accessed 23 Jul 2014

Hart EW, Bryant WA, compilers (2001) Fault number 30a/b, Maacama fault zone, northern section, in Quaternary fault and fold database of the United States: U.S. Geological Survey website, http://earthquakes.usgs.gov/hazards/qfaults. Accessed 23 Jul 2014

Janssen C, Wirth R, Reinicke A, Rybacki E, Naumann R, Wenk HR, Dresen G (2011) Nanoscale porosity in SAFOD core samples (San Andreas Fault). Earth Planet Sci Lett 301:179–189

Lawson AC (1895) Sketch of the geology of the San Francisco peninsula: U.S. Geol Surv Ann Rep 15:439–473

Liu X, Zhan FB, Hong S, Niu B, Liu Y (2012) A bibliometric study of earthquake research: 1900-2010. Scientometrics 92(3):747–765. doi:10.1007/s11192-011-0599-z

Lyons S, Sandwell D (2003) Fault creep along the southern San Andreas from interferometric synthetic aperture radar, permanent scatterers, and stacking. J Geophys Res 108. doi:10.1029/2002JB001831

Marshall GA, Stein RS, Thatcher W (1991) Faulting geometry and slip from coseismic elevation changes: The 18 October 1989, Loma Prieta, California, earthquake. Bull Seismol Soc Am 81:1660–1693

Marx W, Bornmann L (2013) The emergence of plate tectonics and the Kuhnian model of a paradigm shift: a bibliometric case study based on the Anna Karenina principle. Scientometrics 94(2):595–614. doi:10.1007/s11192-012-0741-6

Massonnet D, Rossi M, Carmona C, Adragna F, Peltzer G, Feigl K, Rabaute T (1993) The displacement field of the Landers earthquake mapped by radar interferometry. Nature 364(6433):138–142

Niu B, Hong S, Yuan J, Peng S, Wang Z, Zhang X (2014) Global trends in sediment-related research in earth science during 1992–2011: a bibliometric analysis. Scientometrics 98(1):511–529

Rogers AE, Ingalls RP (1969) Venus: mapping the surface reflectivity by radar interferometry. Science 165(3895):797–799

Spencer JQG, Hadizadeh J, Gratier JP, Doan ML (2012) Dating deep? Luminescence studies of fault gouge from the San Andreas Fault Zone 2.6 km beneath Earth's surface. Quat Geochronol 10. doi:10.1016/j.quageo.2012.1004.1023

Tang L, Walsh J (2010) Bibliometric fingerprints: name disambiguation based on approximate structure equivalence of cognitive maps. Scientometrics 84:763–784

Toda S, Stein RS (2002) Response of the San Andreas fault to the 1983 Coalinga-Nuñez earthquakes: an application of interaction-based probabilities for Parkfield. J Geophys Res 107(B4):11029–11044

Treiman JJ, compiler (1999) Fault number 132, Imperial fault, in Quaternary fault and fold database of the United States: U.S. Geological Survey website, http://earthquakes.usgs.gov/hazards/qfaults. Accessed 23 Jul 2014

Treiman JJ, compilers (1998) Fault number 126a/c/d/e/g, Elsinore fault zone, Glen Ivy section, in Quaternary fault and fold database of the United States: U.S. Geological Survey website, http://earthquakes.usgs.gov/hazards/qfaults. Accessed 23 Jul 2014

Treiman JA, Treiman JJ, compilers (1998) Fault number 126b/f, Elsinore fault zone, Chino section, in Quaternary fault and fold database of the United States: U.S. Geological Survey website, http://earthquakes.usgs.gov/hazards/qfaults. Accessed 23 Jul 2014

Treiman JA, Lundberg MM, compilers (1999) Fault number 125a/b/c/d/e/f/g, San Jacinto fault, Borrego Mountain section, in Quaternary fault and fold database of the United States: U.S. Geological Survey website, http://earthquakes.usgs.gov/hazards/qfaults. Accessed 23 Jul 2014

Wallace RE (1990) General features. In Wallace RE (Ed.) The San Andreas Fault System, California (pp. 3–12). U.S. Geological survey professional, paper 1515. Washington

Waltman L, Van Eck NJ, Noyons E (2010) A unified approach to mapping and clustering of bibliometric networks. J Inf 4(4):629–635

Wang Q, Yang Z, Yang Y, Long C, Li H (2014) A bibliometric analysis of research on the risk of engineering nanomaterials during 1999-2012. Sci Total Environ 473–474:483–489. doi:10.1016/j.scitotenv.2013.12.066

Wiersberg T, Erzinger JA (2007) Helium isotope cross-section study through the San Andreas Fault at seismogenic depths. Geochem Geophys Geosyst 88(1):Q01002. doi:10.1029/2006GC001388

Zhang H, Thurber C, Bedrosian B (2009) Joint inversion of Vp, Vs, and Vp/Vs at SAFOD, Parkfield, California. Geochem Geophys Geosyst 10, Q11002