

地震预测研究态势分析报告 *

——基于 SCI 论文数据 (2008—2018)

1. 数据来源及研究方法

● 分析数据集构建

本报告采用 SCI-E 和 CPCI-S 数据,通过主题检索来获取基础论文数据。经过 3 轮“数据检索→数据下载→关键词分析→检索结果浏览→检索式调整”的过程,辅以施引文献分析和补充,最终确定检索式如下:

TS=((earthquak*NEAR/3 precursor*) OR ("earth-quak*forecast*") OR ("earthquak*predict*") OR ("time depend* seism*hazard*") OR ("Time depend* probab* seism* hazard*") OR ((earthquak* NEAR/2 probab*) AND (probab* NEAR/2 (predict* or forecast*))))

检索数据库: SCI-EXPANDED, CPCI-S

检索时间: 2018 年 12 月 12 日 15:30

时间跨度: 2008—2018

特别说明:受数据库论文收录时滞的影响,在检索日期时,SCI-E 和 CPCI-S 数据库收录 2018 年论文并不完整。因此,日期跨度设置为 2008—2018 年,数据覆盖的时间范围为 10 年,但还要多一些。

以上述检索策略共获得论文记录 1720 条,通过文献

类型筛选,仅保留了“Article”、“Review”和“Proceedings Paper”,并且去掉“Retracted”论文,最终形成的分析数据集为 1674 条论文。进一步,从 WoS 平台将所有曾经引用过上述 1674 条论文的论文(施引文献)导出,共获得并下载施引文献 6066 条,形成施引文献数据集。

● 数据处理

数据集构建完成后,按如下顺序开展数据清洗和规范:

- 解析论文数据 1674 条。解析论文作者姓名及地址,共获得作者姓名数据 7352 条;解析论文的作者关键词字段,并对论文摘要进行切分、抽词,共获取关键词和主题词 2307 个;解析作者单位信息,共获取一级单位名称,以及二级、三级机构名称 879 个。
- 解析施引文献数据,共获得参考文献信息 313435 条。
- 关联论文与施引文献,共获得 11972 条关联记录。
- 规范作者机构名称 200 余个。基于机构的国别和城市信息,归并同一机构的多种不同写法和缩写,形成“机构名称规范表”。
- 规范学者姓名 730 个。基于作者机构,对作者

* 本文编写:中国科学技术信息研究所王立学、张卓然、王晓丽。

姓名的全拼和简写形式进行统一和补充,例如根据其他论文记录中标记的信息,将法国“INST PHYS GLOBE PARIS”的学者“Kossobokov, V (Kossobokov, Vladimir)”规范为“Kossobokov, VG (Kossobokov, Vladimir G.)”。

- 再次规范作者机构名称。以相同的作者全名信息为线索,检查同名作者的隶属单位信息,归并之前未识别的同一单位名称。
- 论文责任者认定。对于每一篇论文,仅选取一位作者作为责任者。选取的规则为:若有通讯作者,选取数据记录中的第一位通讯作者作为责任者,相应地,该通讯作者所标注的作者单位即为该论文的归属单位;若论文未标记通讯作者,则取第一作者作为责任者。值得说明的是,数据集中有两位学者是未挂靠任何工作单位的独立研究者,他们论文的归属单位为空。
- 规范作者关键词和从摘要中抽取出的主题词 2432 个,并将其合并成为地震预测领域术语表。

● 数据分析方法

本报告主要用到了以下三种文献计量方法:

- 合著分析。以论文中的作者署名信息为线索,两两抽取同一篇论文中的作者,可以获得学者之间合作发表论文的情况,进而在一定程度上反映学者之间的科研合作情况。通过关联作者以及他们的地址信息,可以在作者合著的基础上加以引申,分析揭示机构合作、国家合作等情况。合著关系图中,节点之间的连线粗细表示合作次数的相对多少,连线越粗表明合作次数越多。
- 共被引分析。以施引文献数据集中的参考文献数据为基础,两两抽取同一篇施引文献中的参考文献对,计为两篇论文的一次共被引。两篇论文共同被引用,可以认为是具有内在的主题关联。因此,论文共被引分析常常被用来识别和聚焦领域内基础性的核心论文(又称:知识基础)。将论文共被引关系引申至作者、机构、期刊等信息项上,即可获得更加丰富的共被引关系。共被引关系图中,节点之间的连线粗细表示共被引关系的相对多少,连线越粗表明研究主题相关程度越高。
- 关键词共现分析。多数论文中都会带有作者给出

的数个关键词,即作者关键词。由于作者关键词一般数量较少,所以在具体论文分析中,还经常从论文的题名和摘要中通过文本分析来切分、抽取关键词,即主题词。通常我们假设:作者关键词和主题词都可以反映论文的主要研究主题。因此,考察词语之间的两两共同出现的关系,可以在相当程度上揭示论文的研究主题。关键词共现关系图中,节点之间的连线粗细表示共现次数的相对多少,连线越粗表明关键词相关程度越高。

2.地震预测领域论文概况

在10年多的时间里,地震预测领域的论文及被引分布情况见表1和图1。可以看出,2011年发表的论文累积被引次数最高;2009年发表的论文篇均被引最高;2017年发表的论文至今已经被引用了368次,很接近2016年论文的394次,表明2017年发表的论文中有较受关注的研究成果。由于论文被引的累积效应,若仅仅考察论文被引次数时,通常都是发表年份较早的论文被引较高,难以反映领域内最新的发展动态。为此,本报告分析了近两年(2017—2018年)发表的论文对整个数据集中各个年份论文的引用情况,可以在一定程度上反映地震预测领域近期的研究和关注态势。

基于表1数据,若将数据归并为2008—2012年和2013—2017年两个时间段,结合论文数量、被引篇数和被引次数来计算篇均被引情况则可以获知:地震预测领域的研究者总体上略微偏好于引用5年内发表成果,表明领域内的研究内容具有与时俱进的特点;不过,发表时间更早的论文却获得了略大的引用强度,这意味着领域内已经积累下来了一批核心的经典论文。另外,2018年论文数量偏低,结合最近两年的发文数量及被引篇数来推测,SCI-E和CPCI-S两个数据库对领域内2018年论文的收录数量并未齐全。

3.国家/地区论文分布

3.1 国家/地区论文被引排名

地震预测领域的1674篇论文共由61个国家和地区发表,中国大陆、美国和意大利三个国家/地区发文数量较多;而按被引次数的排序则是美国、意大利和中国位列前三;瑞士发表的论文获得了最高的篇均被引;中国台湾地区的研究成果也较有显示度。论文累积被引次数超过100次的国家和地区的详情见表2。

表 1 地震预测领域论文发表及被引情况

年份	论文数量	累积被引	篇均被引	2017/2018 被引次数	2017/2018 篇均被引 a*	2017/2018 被引篇数	2017/2018 篇均被引 b*
2008	116	1447	12.5	247	2.1	58	4.3
2009	113	1528	13.5	287	2.5	59	4.9
2010	172	1914	11.1	410	2.4	108	3.8
2011	168	2036	12.1	544	3.2	108	5.0
2012	165	1430	8.7	464	2.8	106	4.4
2013	189	1520	8.0	633	3.3	128	4.9
2014	134	804	6.0	426	3.2	84	5.1
2015	161	681	4.2	469	2.9	123	3.8
2016	136	394	2.9	326	2.4	83	3.9
2017	182	368	2.0	364**	2.0	98	3.7
2018	138	77	0.6	—	—	—	—

注*：篇均被引a的数值计算为：2017/2018被引次数 ÷ 该年份论文数量；篇均被引b的数值计算为：2017/2018被引次数 ÷ 该年份在2017/2018年的被引篇数。

**：被引的本地统计值与WoS平台给出的被引值有出入，“364次”相对于第三列的“368次”的数值是因为无法从WoS平台浏览和下载对应的全部施引文献。

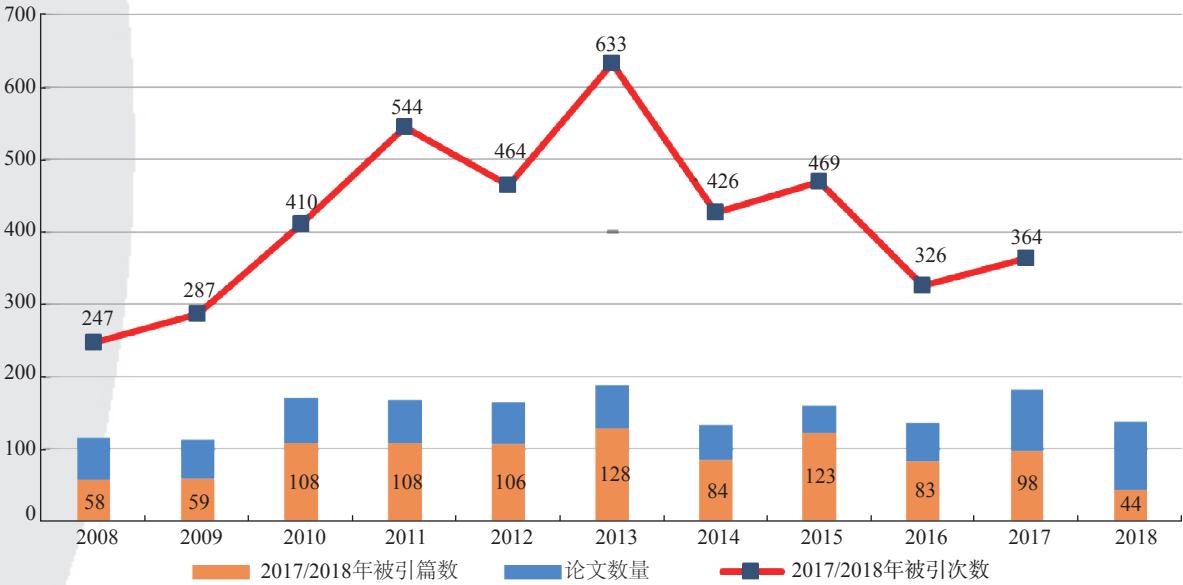


图1 地震预测领域论文发表及被引情况

3.2 国家/地区论文合作关系

图 2 和图 3 展现的是地震预测领域论文的国家 / 地区合作情况，两幅国家 / 地区合著关系图的节点位置相同，区别在于图 2 的节点大小表示国家 / 地区的通讯作者（未

标记通讯作者时取第一作者机构）发文数量；而图 3 的节点大小表示国家 / 地区的通讯作者（未标记通讯作者时取第一作者机构）发文的篇均被引。两幅图中，合作少于 5 次的关系未予显示。

表 2 地震预测领域论文的国家 / 地区分布 (被引 >100)

序号	国别 / 地区	论文数量	论文被引	篇均被引
1	USA	163	2392	14.7
2	Italy	153	1460	9.5
3	China	311	1108	3.6
4	Japan	107	1048	9.8
5	Russia	176	843	4.8
6	India	112	685	6.1
7	Taiwan,China	80	649	8.1
8	Switzerland	28	569	20.3
9	Greece	72	451	6.3
10	New Zealand	33	367	11.11
11	Germany	35	332	9.5
12	France	23	318	13.8
13	Iran	55	268	4.9
14	Spain	24	250	10.4
15	England	17	210	12.4
16	Turkey	45	208	4.6
17	Canada	19	152	8.0
18	Scotland	11	141	12.8

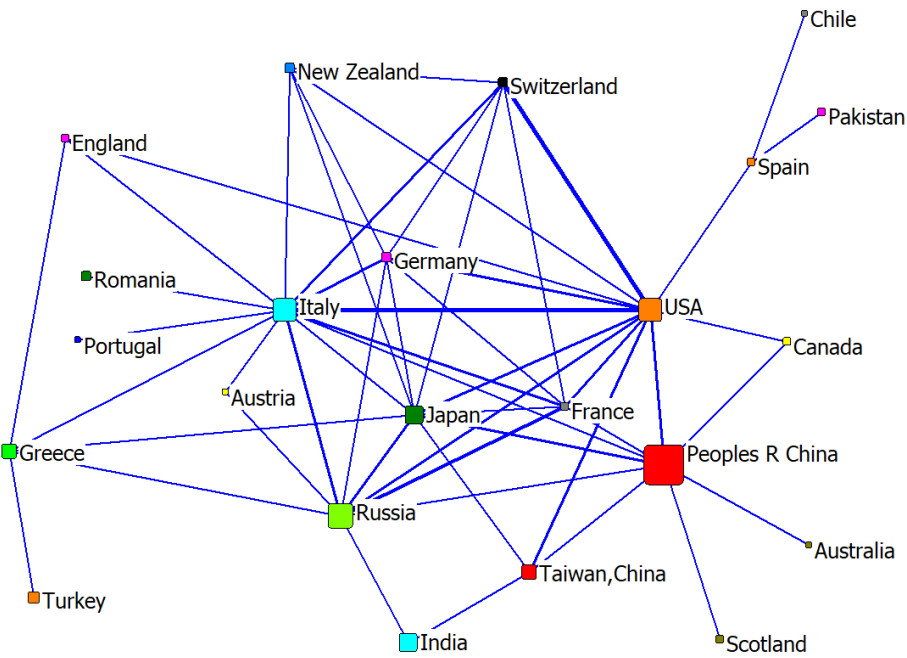


图2 国家/地区科研合作关联 (A)

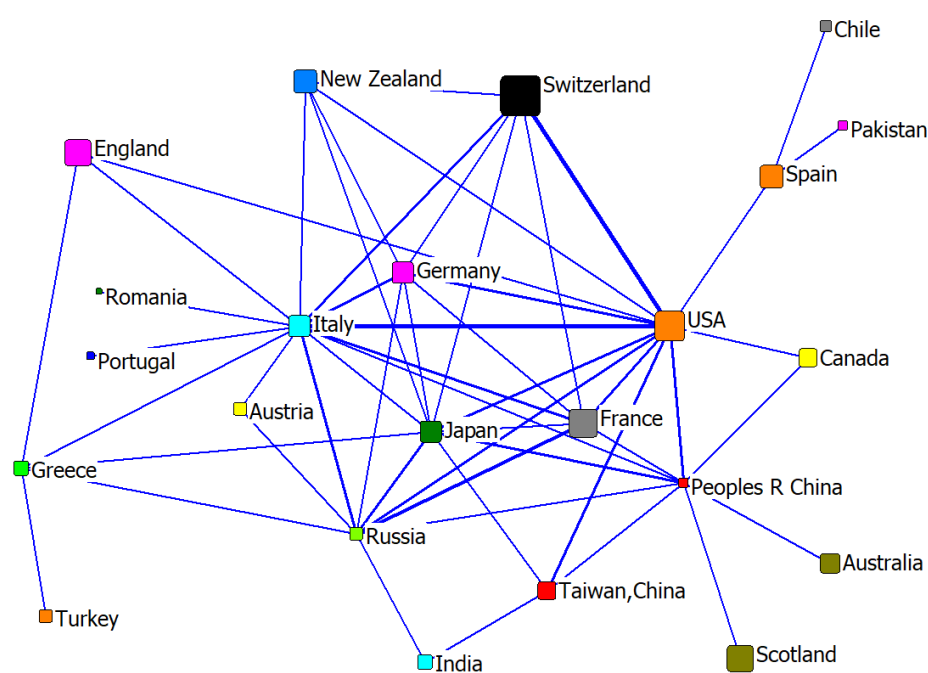


图3 国家/地区科研合作关联 (B)

从图2和图3来看,中国大陆和美国、意大利、俄罗斯、法国、日本、德国、瑞士等国家/地区共同组成了全球范围内地震预测领域的核心合作网络;中国大陆、美国 and 意大利三个国家/地区的发文数量大,体现得非常明显(图2);瑞士、美国、法国和英国等国家/地区论文的篇均被引较高(图3);而中国大陆地区的发文数量最多,但篇均被引并不突出;从合作网络中也可以看出,中国虽然是国际合作网络中的重要成员之一,但作为发文量最多的国

家,在合作网络中尚未成为中心节点。

3.3 中国大陆论文情况统计

与图1同等原理,中国大陆论文的发表和被引情况见图4。可能是受2008年汶川地震的冲击和影响,中国大陆在地震预测领域的研究论文数量在2010年达到峰值。值得注意的是,中国大陆地区2018年发表的论文已经获得了相对较高的引用。

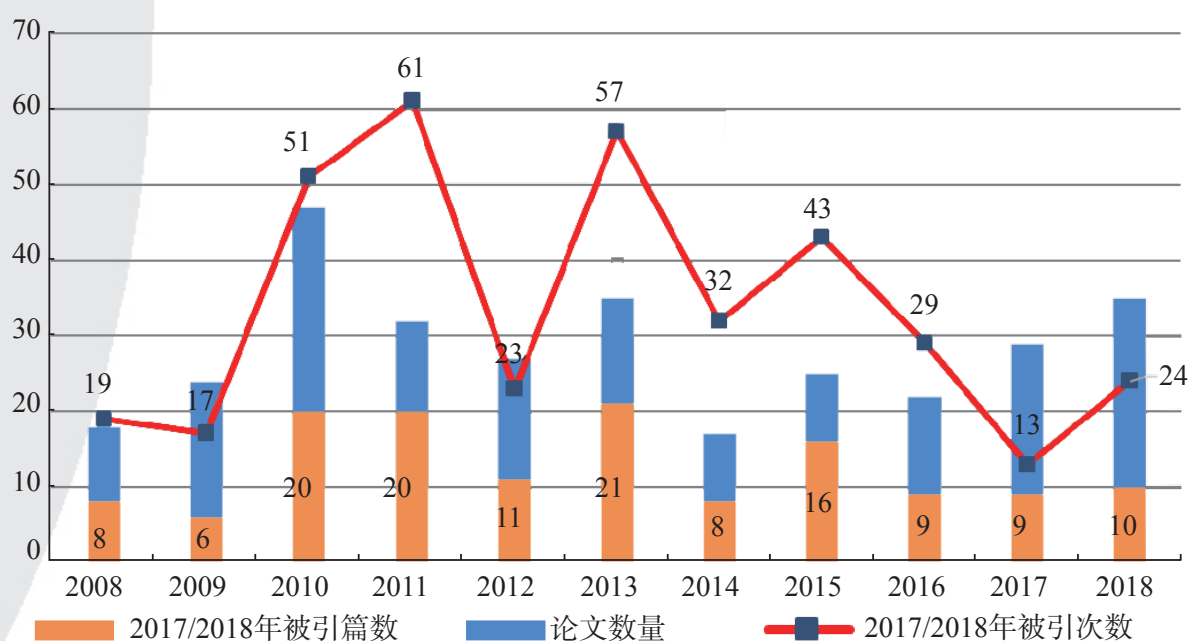


图4 地震预测领域中国论文的发表及被引情况

4.研究机构分布及关联

4.1 机构论文被引排名

地震预测领域的 1674 篇论文共由 61 个国家和地区的 572 个机构发表，被引次数超过 100 次的机构情况见表 3。可以看到，排名前三的机构分别是意大利的 INST NATL GEOPHYS VOLCANOL、俄罗斯的 RUSSIAN ACAD SCI 和瑞士的 ETH Zurich。中国大陆的机构 CEA CHINA（中国地震局）排名第五。

表 3 地震预测领域论文的机构分布（被引 >100 次）

序号	国别 / 地区	机构名	论文数量	论文被引	篇均被引
1	Italy	INST NATL GEOPHYS VOLCANOL	63	696	11.0
2	Russia	RUSSIAN ACAD SCI	135	529	3.9
3	Switzerland	ETH Zurich	27	521	19.3
4	USA	UNIV SO CALIF	10	495	49.5
5	China	CEA CHINA	122	464	3.8
6	USA	US GEOL SURVEY	22	354	16.1
7	USA	UNIV CALIF LOS ANGELES	23	273	11.9
8	New Zealand	GNS SCI	22	258	11.7
9	Taiwan,China	NATL CENT UNIV	22	255	11.6
10	Greece	UNIV ATHENS	26	238	9.2
11	Germany	GFZ GERMANY	19	177	9.3
12	Japan	UNIV TOKYO	20	170	8.5
14	China	CHINESE ACAD SCI	39	165	4.2
13	Italy	CNR ITALY	17	165	9.7
15	Russia	INST APPL GEOPHYS	1	144	144.0
16	USA	BRIDGEWATER STATE COLL	1	139	139.0
17	Japan	JAPAN ACAD	4	131	32.8
18	Scotland	UNIV EDINBURGH	8	129	16.1
19	Iran	UNIV TEHRAN	20	118	5.9
20	Japan	UNIV ELECTROCOMMUN(UEC)	12	111	9.3
21	Italy	UNIV TRIESTE	8	107	13.4
22	Japan	TOKAI UNIV	2	102	51.0

4.2 机构发文主题关联

图 5 展现的地震预测领域论文的机构研究主题关联，由论文共被引数据分析获得。图中，节点大小为该机构作为通讯作者（未标记通讯作者时取第一作者机构）的发文数量，共被引次数少于 38 次的关系未予显示。

从图 5 中可以看出，意大利 INST NATL GEOPHYS VOLCANOL、瑞士 ETH Zurich、美国 UNIV SO CALIF 和

UNIV CALIF LOS ANGELES、新西兰 GNS SCI 等是机构研究主题关联网络图中的核心节点，意味着上述机构的论文研究主题相关度较强；中国地震局（CEA CHINA）和俄罗斯科学院（RUSSIAN ACAD SCI）发文量很大但并未成为关联网络中的核心节点，意味着两所机构的研究主题可能相对地另有侧重，也可能意味着语种的差异仍然是阻碍地震预测领域科研成果交流和传播的一大障碍。

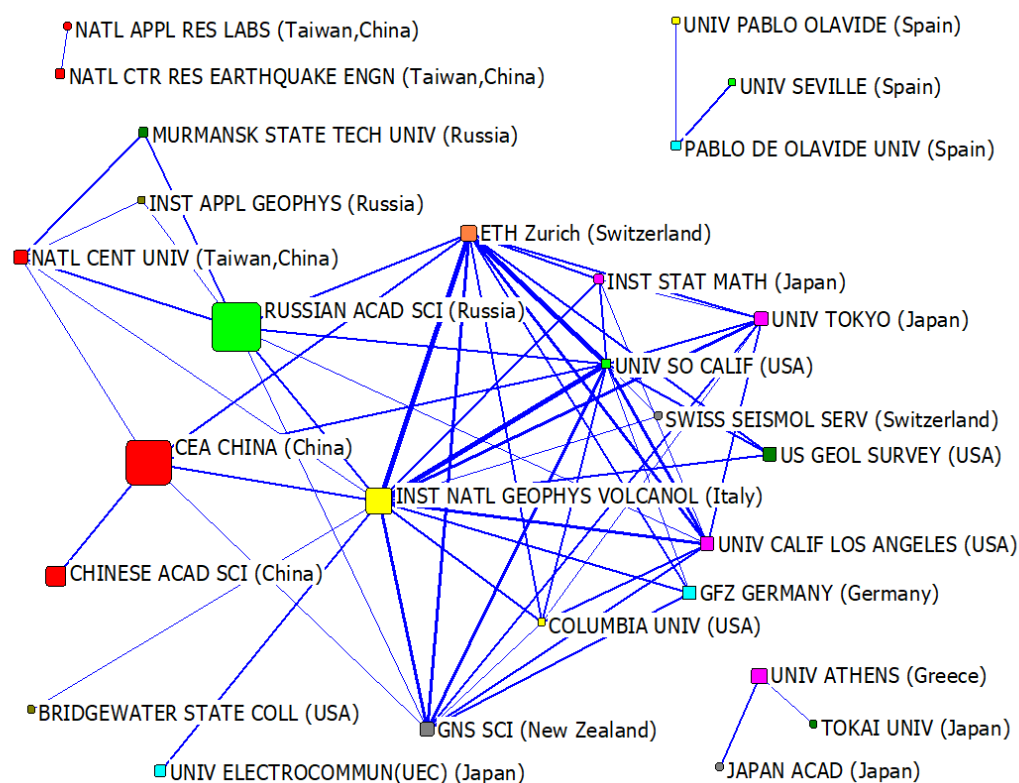


图5 机构研究主题关联

4.3 机构科研合作关系

图6和图7展现的是地震预测领域论文的机构合作情况，两幅机构合作关系图的节点位置相同，区别在于图6的节点大小表示机构的通讯作者（未标记通讯作者时取第

一作者机构）发文数量；而图7的节点大小表示机构的通讯作者（未标记通讯作者时取第一作者机构）发文的篇均被引。两幅图中，合作少于7次的关系未予显示。

从图6和图7中可以看出，地震预测领域的机构合作

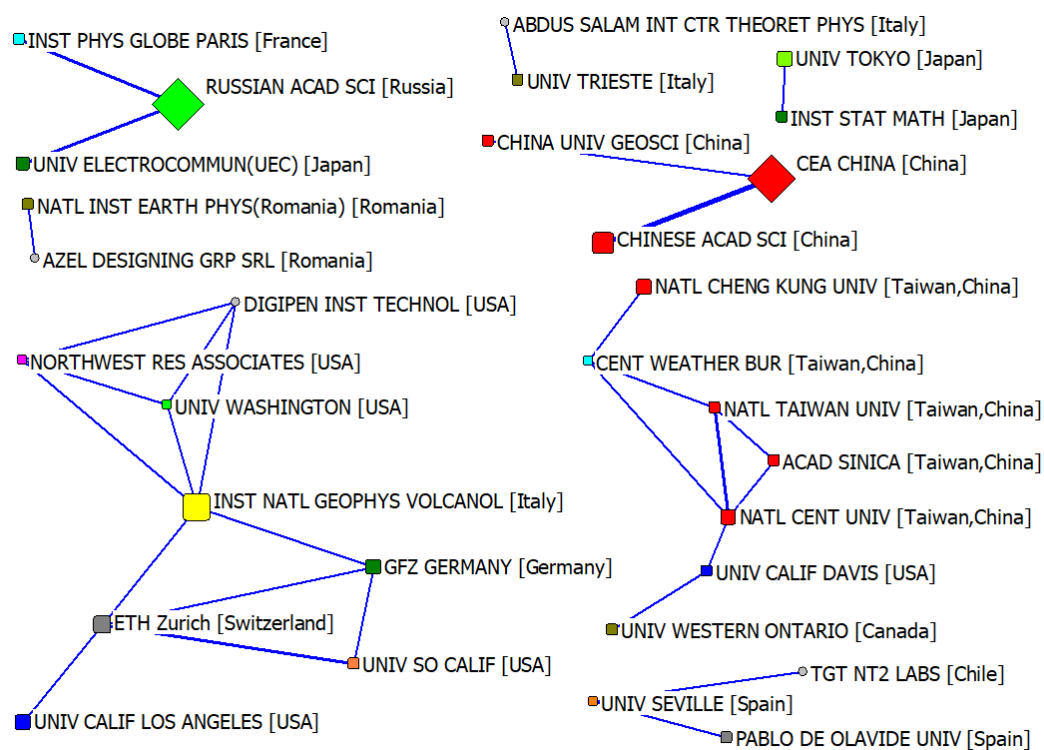


图6 机构科研合作关系 (A)

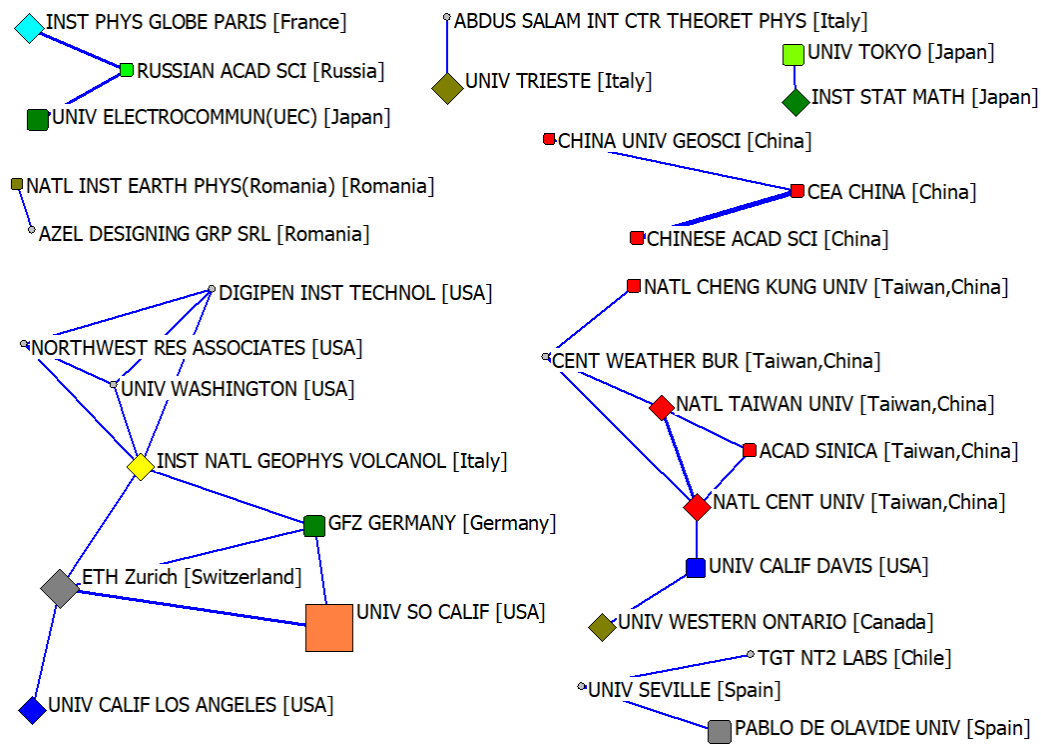


图7 机构科研合作关联（B）

关系，呈现出同一国家内部合作与跨国合作并存的局面：如中国地震局（CEA CHINA）和中科院是典型的国内机构合作，可能意味着中国大陆的研究机构之间有着较为相近且固定的研究方向，能够形成稳定且大量的合作；而意大利 INST NATL GEOPHYS VOLCANOL、瑞士 ETH Zurich、美国 UNIV SO CALIF 和德国 GFZ GERMANY 则形成了较为突出的跨国合作，可能意味着上述机构共同开展了较多的全球化研究。

另外值得注意的是，对比图 6 和图 7 可以看出，中国地震局（CEA CHINA）和俄罗斯科学院（RUSSIAN

ACAD SCI）论文的篇均被引较低，可能意味着两所研究机构仍需耐心磨砺，努力引发研究成果在国际同行间的共鸣。

4.4 中国大陆科研机构情况统计

进一步，中国大陆科研机构按照发文数量、被引次数和篇均被引情况排序的 Top10 分别见表 4、表 5 和表 6。无论是从发文数量上还是被引次数上来看，中国地震局都是绝对的“领头羊”，中科院紧随其后，排名第二；篇均被引方面，有多所机构因为发文数量少而篇均被引高。

表 4 地震预测领域中国大陆机构：发文量 Top10

序号	机构名	论文数量	论文被引	篇均被引
1	CEA CHINA	122	464	3.8
2	CHINESE ACAD SCI	39	165	4.2
3	CHINA UNIV GEOSCI	10	24	2.4
4	CHINA RES INST RADIOWAVE PROPAGAT	8	47	5.9
5	NANJING UNIV INFORMAT SCI TECHNOL	8	32	4.0
6	CHINA UNIV MINING TECHNOL	8	23	2.9
7	PEKING UNIV	7	67	9.6
8	HUAZHONG UNIV SCI TECHNOL	5	24	4.8
9	SHANDONG UNIV SCI TECHNOL	5	20	4.0
10	WUHAN UNIV	5	5	1.0

表 5 地震预测领域中国大陆机构：论文被引 Top10

序号	机构名	论文数量	论文被引	篇均被引
1	CEA CHINA	122	464	3.8
2	CHINESE ACAD SCI	39	165	4.2
3	PEKING UNIV	7	67	9.6
4	CHINA RES INST RADIOWAVE PROPAGAT	8	47	5.9
5	HONG KONG UNIV SCI TECHNOL	2	39	19.5
6	CHINA METEOROL ADM	1	39	39.0
7	MOE BEIJING NORMAL UNIV	1	33	33.0
8	NANJING UNIV INFORMAT SCI TECHNOL	8	32	4.0
9	CHINA UNIV GEOSCI	10	24	2.4
10	HUAZHONG UNIV SCI TECHNOL	5	24	4.8

表 6 地震预测领域中国大陆机构：篇均被引 Top10

序号	机构名	论文数量	论文被引	篇均被引
1	CHINA METEOROL ADM	1	39	39.0
2	MOE BEIJING NORMAL UNIV	1	33	33.0
3	HONG KONG UNIV SCI TECHNOL	2	39	19.5
4	NANJING FORESTRY UNIV	1	11	11.0
5	PEKING UNIV	7	67	9.6
6	CHENGDU UNIV TECHNOL	1	9	9.0
7	CHINA RES INST RADIOWAVE PROPAGAT	8	47	5.9
8	HUAZHONG UNIV SCI TECHNOL	5	24	4.8
9	NANJING UNIV AERONAUT ASTRONAUT	4	17	4.3
10	CHINESE ACAD SCI	39	165	4.2

5.研究主题分布态势

5.1 词频分布态势

基于对地震预测领域论文的解析和统计，分别获得了作者关键词的词频列表及年度分布，见表 7 和表 8。在表 7 的词频列表中，对于本报告主题来说相对通用的“earthquake”、“earthquake prediction”、“earthquake precursor”或“earthquake forecast”等词语的词频最高，相对专指一些的“radon”、“ionosphere”和“probability model”等词语紧随其后。在表 8 的词频年度分布中，隐去了对于本专题来说意义较为通用的词语，仅展现了相对具体的词语的年度分布。可以看出，汶川地震的相关研

究成果，在 2010 — 2012 年集中发表见刊。

地震预测领域论文的主题词词频列表及年度分布见表 9 和表 10。在表 9 中，出现次数较多的仍然是意义不太专指的“giant earthquake”、“earthquake prediction”、“earthquake precursor”或“earthquake activity”等词语。经过筛选，在表 10 中列出了意义较为专指的词语的年度分布，如“radon concentration”、“time series”和“tec anomaly”等。

表7 地震预测领域论文作者关键词：词频Top30

序号	关键词	文档词频	序号	关键词	文档词频
1	earthquake	212	16	b-value	25
2	earthquake prediction	197	17	groundwater	24
3	earthquake precursor	193	18	seismicity and tectonics	22
4	earthquake forecast	71	19	gps	21
5	statistical seismology	62	20	neural network	21
6	radon	57	21	remote sensing	17
7	ionosphere	55	22	tec data	17
8	probability model	39	23	time series	16
9	forecast	40	24	acoustic emission	15
10	earthquake hazard	37	25	tec	15
11	precursor	36	26	time-series analysis	15
12	seismicity	31	27	radon anomaly	14
13	2008 wenchuan earthquake	30	28	earthquake interaction, forecasting, and prediction	14
14	earthquake excitation	28	29	earthquake activity	13
15	prediction	28	30	tir anomaly	13

表8 地震预测领域论文作者关键词：年度分布

关键词	词频	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
statistical seismology	62	2	2	8	7	9	5	7	3	6	9	4
radon	57	3	4	6	5	1	8	9	7	3	6	5
ionosphere	55	4	5	2	5	1	10	2	11	2	10	3
probability model	39	2	1	3	2	4	5	5	2	6	7	2
earthquake hazard	37	2	2	2	3	6	4	4	3	5	2	4
2008 wenchuan earthquake	30	0	2	6	10	5	2	0	1	1	1	2
earthquake excitation	28	0	1	1	2	7	6	3	1	1	6	0
b-value	25	1	0	1	2	5	4	1	4	3	3	1
groundwater	24	2	2	4	2	0	3	0	1	2	4	4
seismicity and tectonics	22	1	0	2	3	5	2	2	0	3	4	0
gps	21	1	2	3	3	1	5	0	4	1	1	0
neural network	21	0	1	1	3	3	2	1	1	3	4	2
remote sensing	17	0	1	0	0	5	1	4	1	1	3	1
tec data	17	1	2	1	0	2	3	0	4	0	3	1
time series	16	0	3	1	1	1	3	1	3	0	3	0
acoustic emission	15	0	2	0	2	1	2	0	2	1	2	3
tec	15	0	2	0	2	0	4	1	3	2	0	1
time-series analysis	15	0	0	0	4	2	3	2	0	1	2	1
radon anomaly	14	0	1	4	1	1	0	1	2	0	2	2
tir anomaly	13	1	1	0	1	0	0	3	3	1	1	2

表 9 地震预测领域论文主题词：词频 Top30

序号	主题词	文档词频	序号	主题词	文档词频
1	giant earthquake	103	16	earthquake forecast	14
2	earthquake prediction	70	17	radon activity	12
3	earthquake precursor	46	18	fault zone	12
4	earthquake activity	44	19	tir anomaly	10
5	radon concentration	37	20	radon anomaly	10
6	main shock	28	21	etas model	10
7	earthquake event	27	22	earthquake catalogue	10
8	main earthquake	23	23	seismicity rate	10
9	time series	23	24	soil gas	9
10	tec anomaly	22	25	earthquake cycle	8
11	earthquake hazard	21	26	tec data	8
12	electric field	20	27	tectonic force	7
13	neural network	20	28	gps	7
14	2008 wenchuan earthquake	15	29	earthquake parameter	7
15	study region	14	30	ground motion	7

表 10 地震预测领域论文主题词：年度分布

主题词	合计	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
radon concentration	37	1	2	2	2	3	3	5	5	0	8	6
time series	23	1	0	0	3	2	5	1	5	1	2	3
tec anomaly	22	0	1	1	1	0	8	1	3	1	4	2
electric field	20	1	2	3	3	1	3	2	1	0	0	4
neural network	20	1	2	1	2	3	4	1	0	2	3	1
2008 wenchuan earthquake	15	1	0	6	5	0	1	0	0	1	0	1
radon activity	12	0	2	0	2	0	0	1	2	1	3	1
fault zone	12	0	0	2	1	2	2	1	0	0	2	2
tir anomaly	10	0	1	0	0	1	2	1	2	2	0	1
radon anomaly	10	0	1	1	0	3	1	1	1	0	0	2
etas model	10	0	0	1	0	0	1	3	1	1	2	1
earthquake catalogue	10	0	1	3	0	1	1	1	0	0	1	2
seismicity rate	10	0	0	2	0	3	0	1	1	1	1	1
soil gas	9	1	0	2	0	1	0	0	1	0	3	1
earthquake cycle	8	2	0	0	1	2	0	0	1	1	1	0
tec data	8	0	1	0	0	0	1	0	3	0	1	2
tectonic force	7	0	0	2	1	0	0	1	1	0	0	2
gps	7	0	0	0	2	1	2	1	1	0	0	0
earthquake parameter	7	0	0	0	1	2	1	0	0	2	1	0
ground motion	7	0	2	1	1	1	1	0	0	0	1	0

5.2 共现关联态势

地震预测领域作者关键词共现关联图见图 8。图中，共现次数低于 5 次的共现关系未予显示。由图中可以看出，“earthquake prediction”、“earthquake precursor”和“statistical seismology”是三个核心概念，它们各自形成了相对成型的小主题网络，并且相互关联；“earthquake

析能够比作者关键词分析获得更多的领域术语(关键词)。从图 9 中的共现关系可以看出,“radon concentration”是与地震活动、地震前兆密切相关的主题词,这一概念已经形成了较为成型的主题概念网络。共现关联图的更多内涵留待地震预测领域的专业研究人员解读。

涵留待地震预测领域的专业研究人员解读。

