

张盛峰,张永仙,李莎. 2025. CSEP 2.0 阶段工作研究进展及中国测试区设计与模型试验[J]. 地震研究, 48(2):188–198, doi:10.20015/j.cnki.ISSN1000–0666.2025.0020.

Zhang S F, Zhang Y X, Li S. 2025. Phase 2.0 of CSEP work: Progress, design and model forecasting experiment in the China test region [J]. *Journal of Seismological Research*, 48(2):188–198, doi:10.20015/j.cnki.ISSN1000–0666.2025.0020.

CSEP 2.0 阶段工作研究进展及中国测试区 设计与模型试验*

张盛峰^{1,2}, 张永仙¹, 李莎³

(1. 中国地震局地震预测研究所, 北京 100036; 2. 新西兰地质与核科学研究所, 新西兰 下哈特 5040;
3. 广西壮族自治区地震局, 广西 南宁 530022)

摘要: 由美国南加州地震中心和美国地质调查局共同支持的 RELM 工作组以及由此进一步发展而来的“地震可预测性国际合作研究”(CSEP)计划开始之后, 研究人员围绕地震预测模型和统计检验方法内容的不同测试中心对地震可预测性问题开展了广泛研究。在前期针对 CSEP 1.0 阶段介绍的基础上, 介绍 CSEP 2.0 阶段的相关内容, 包含该阶段发起的背景、发展的总体设计、主要进展以及不同测试区的重点关注内容。针对我国当前的参与情况, 初步设计了参与该阶段工作的中国 CSEP 2.0 测试区范围以及使用图像信息学(PI)算法开展的预测试验。

关键词: CSEP 2.0; 中国测试区; PI 算法; 效能评估

中图分类号: P315.72 **文献标识码:** A **文章编号:** 1000–0666(2025)02–0188–11

doi:10.20015/j.cnki.ISSN1000–0666.2025.0020

0 引言

20 世纪 80 年代末以来, 基于统计物理学的理论和方法与地震学问题相结合, 成为了地震学和物理学之间的重要交叉领域(吴忠良, 陈运泰, 2002), 尤其是在针对地震可预测性问题进行的国际讨论的驱动下, 近年来结合统计学和地震学问题的统计地震学手段愈加受到重视。在美国南加州地震中心(South California Earthquake Center, SCEC)的主导下, 区域地震概率模型(Regional Earthquake Likelihood Models, RELM)工作组鼓励发展多种模型, 并采用类似竞技的方式吸纳不同类型的地震预测模型开展预测试验。2009 年 8 月 1 日正式开始的地震可预测性国际合作研究计划(Collaboratory for the Study of Earth-

quake Predictability, CSEP)倡导在全球不同区域以 CSEP 测试中心的方式开展地震可预测性研究, 采取统一的数据来源和预报规则, 并注重对预测结果的统计检验(Field, 2007)。目前发展较为成熟的机构包含美国南加州地震中心(SCEC)、瑞士苏黎世高等工业学校(ETH Zurich)、新西兰地质与核科学研究所(Institute of Geological and Nuclear Science, GNS, New Zealand)、日本东京大学地震研究所(Earthquake Research Institute, University of Tokyo)等, 主要对测试区域包含美国加州地区、西北太平洋地区、日本地区、意大利地区、新西兰地区和全球尺度范围的预测模型开展严格评估(Schorlemmer *et al*, 2007)。从这些发展中也可以看出, 针对地震的可预测性问题研究, 随着基于不同时间尺度、不同数据类型和研究区域的预测模型不断出现和完善, 相应的统

* 收稿日期: 2024–06–30.

基金项目: 国家自然科学基金(42004038); 中国地震局地震预测研究所基本科研业务专项(CEAIEF2022030206); 地震预测开放基金(XH24005D); 中国地震局和留学基金委英才计划 CSC 项目(202204190019).

第一作者简介: 张盛峰(1988–), 副研究员, 博士, 主要从事地震预测及统计地震学研究.

E-mail: 085012104@163.com.

计检验方法为随机模型或者标准参考模型的效能评估工作提供了良好的手段 (Schorlemmer *et al*, 2007)。

针对预测模型的发展以及不断提出的新的科学问题, 目前 CSEP 计划已由 1.0 阶段发展至 2.0 阶段, 模型研发专家对算法改进、亟需进一步研究的问题和方法进行了广泛的讨论, 且不同测试区也基本确定了新的研究方向 (Michael, Werner, 2018; Schorlemmer *et al*, 2018)。2018 年 5 月召开的汶川地震十周年国际研讨会暨第四届大陆地震国际研讨会正式宣布了中国地震科学实验场的成立, 为我国的地震预测研究提供了一个新的国际舞台 (吴忠良等, 2021)。同时, 在我国的 CSEP 2.0 工作即将开展阶段, 中国地震科学实验场也被选定为这一阶段的测试区, 在前期 CSEP 1.0 工作的基础上继续进行不同模型的预测试验和平台搭建 (Zhang *et al*, 2019; 张盛峰, 2019)。由于本项目在全球不同测试中心基本上均采用“边探索边研究”的模式, 虽然大致研究方向基本相同, 但研究的具体内容会由于不同测试区的实际要求和研究进程而存在差异, 及时追踪与这一阶段内容相关的研究趋势对于我国下一步开展好相关工作显得尤其重要。相关研究人员已针对 CSEP 工作的总体路线和 1.0 阶段产出成果进行了介绍 (张

盛峰, 张永仙, 2021), 因此, 本文针对 CSEP 2.0 阶段的相关内容进行了梳理, 通过借鉴其它测试区或模型研发人员的先进经验, 把握好国际研究趋势, 提高中国在这一国际合作项目中的参与程度; 同时为了保证与其它测试中心研究标准和策略一致, 本文介绍了 CSEP 2.0 中国测试中心的标准化设置以及在此基础上初步开展的预测试验, 期望为后续其它模型参与该项国际合作提供有效的参考。

1 CSEP 不同阶段总体设计

CSEP 计划发展至目前, 其规划设计大致被划分成了 3 个阶段, 一是单服务器阶段 (2007—2017 年), 即形成了具有统一计算流程和代码库的单服务器 CSEP 检验系统; 二是多服务器阶段 (2017—2018 年), 即保证在不同的服务器上可以运行相同的处理软件, 并产出类似的结果, 在多服务器上可以产出独立预测结果的 CSEP 新系统; 三是发展具有成熟多服务器工作流程的 CSEP 系统 (2018 年至今), 这一系统将提高自动化的容错能力, 可便捷地获取和处理输入数据, 并能同时包含回溯性预测、向前预测及效能评估一体化的工作流程 (图 1)。

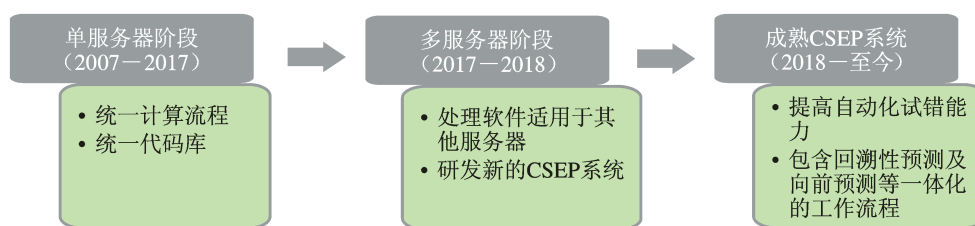


图 1 CSEP 计划发展规划的大致发展阶段

Fig. 1 The general development stages of CSEP project

CSEP 2.0 计划将数据的前期处理分析、预测模型预测试验及评估、完善的自动化处理能力等不同方面建设为现代化的工作流程, 因此经过不同测试区研发专家的广泛讨论, 对当前阶段提出了具体的工作内容, 主要有: ① CSEP 2.0 阶段应该研发产出可对这一阶段所包含预测模型效能进行评估的软件系统和硬件平台; ② 研发符合 CSEP 2.0 阶段工作方向和指导原则的新技术; ③ 将以往 CSEP 处理过程

的受控环境设计为容器操作; ④ 向公众展示成果的途径应设计为在 Web 网页上可查看结果; ⑤ 相关程序应该开源并可评估; ⑥ 针对处理流程的可再现性特征, 应该设计为容器内的工作流程; ⑦ 进一步讨论预测和效能评价容易脱钩的问题; ⑧ 建议向社区开放源代码试验数据; ⑨ 应该重视地震目录数据的版本问题; ⑩ 引入容量较小且可行性较高的产品在容器中运行 Open - SHA 算法。

2 CSEP 2.0 阶段研究进展

CSEP 1.0 阶段在不同测试区预测试验、检验评估等方面开展了很多工作 (张盛峰, 张永仙, 2021), 但其工作流程已无法满足于当前时期地震学家们针对最新科学方法和模型的期望水平, 主要表现为: ①预测规范过于严格; ②工作流程灵活度不高; ③新的且计算成本高的模型已经可用, 但尚未应用到 CSEP 工作中, 如统一的加利福尼亚州地震破裂预测第 3 版 (UCERF3)、三维 + 有限破裂尺度传染型余震序列 (3D + Finite ETAS) 模型。因此, 由美国、中国、日本、欧洲等不同测试中心共同参与的 CSEP 2.0 计划试图发展新的地震预测模型, 包含美国地质调查局倡导的可操作的余震预测模型和 California 倡导的统一的加利福尼亚州地震破裂预测第 3 版 - 传染型余震序列 (UCERF3 - ETAS) 模型 (Schorlemmer *et al.*, 2018)。

2018 年召开的美国南加州地震中心 (SCEC) 学术年会包含与 CSEP 工作相关的专题研讨会^①, 与

会代表来自 SCEC 测试区、全球其它 CSEP 测试区、美国地质调查局 (USGS) 和其他负责向公众通报地震危害的国际政府机构。会议初步决定通过整合现有的 CSEP 检验中心, 于 2018 年 9 月启动 CSEP 2.0 计划, 以进一步推进地震预测国际合作, 扩展当前的模型和数据空间, 并对 CSEP 平台现有的评估软件提出了新的要求。该研讨会主要围绕不同模型开展实验, 以便对现有的模型进行深入的评估, 讨论主题主要包括: 新的回溯性和前瞻性试验蓝图的构建、对竞技模型进行评估和比较所需的技术方法、与 CSEP 相关的软件设计、国际合作和软件开发、CSEP 1.0 数据和产权问题。2018 年在《地震研究快报》(Seismological Research Letters) 出版了与 CSEP 工作相关的特辑, 包含了对 CSEP 1.0 相关工作进行梳理的 9 篇文章, 并对 CSEP 2.0 的工作进行了展望 (Christophersen *et al.*, 2018; Michael, Werner, 2018; Schorlemmer *et al.*, 2018; Strader *et al.*, 2018; Taroni *et al.*, 2018)。通过梳理以上相关成果, 本文总结了 CSEP 2.0 阶段备受关注的的工作内容及研究方向, 见表 1。

表 1 CSEP 2.0 阶段的重要工作内容和方向
Tab. 1 Key directions and contents of CSEP 2.0 stage

工作方向	重要内容
数据空间的扩展	<ol style="list-style-type: none"> 1. 数据空间的扩展: 其他具有良好地震目录的地区 + 余震序列 + 古地震记录 2. 数据质量的提高: 发震时刻、震中位置误差、震级的不确定性 3. 解决实验最短持续时间问题: 基于各分量的检验/模型重构
模型类型的扩展	<ol style="list-style-type: none"> 1. 3D 模型: 日本关东地区 3D 实验场 2. 混合/杂交模型: 模型组合/增加新变量 3. 基于断层的模型: 对 UCERF3 进行检验 4. 基于事件的模型: 每次事件后进行预测更新, 而非按照常规时间间隔更新 5. 基于物理的模型: 类似 RSQsim 模拟器, 加入断裂物理学概念 6. 完整的概率模型: 利用模型对所有不确定性进行完整描述 7. 地震动和危险性模型: 利用权威来源的破裂参数, 首先针对日本国家地震危险性模型 8. 前兆模型: 评估类似大地测量、电磁异常等可能的前兆模型 9. 外部预测: 如政府机构发布的可操作预测或无法安装在内部的前兆模型预测
关键假设和问题的检验	<ol style="list-style-type: none"> 1. 大地震与小地震在丛集性、尺度或长期行为上的本质区别 2. 单一断层上的地震震级分布特征 3. 强震发生对附近其他主震发生概率的影响 4. 中等强度地震对大地震的触发概率 5. b 值的预测能力 6. 小地震空间位置对未来大震位置的指示作用 7. 前震识别

① <https://www.scec.org/workshops/2018/csep>.

2019 年召开的 SCEC 年会^①主要讨论了 CSEP 2.0 工作的进展,尤其是围绕 UCERF3 – ETAS 模型和其它可操作的地震预测模型对 Ridgecrest 地震序列展开分析,并讨论了对 CSEP 软件在全球范围内进行研发的管理工作和未来发展规划,会议期间,意大利、新西兰、日本和中国测试区的 CSEP 专家进行了工作介绍。在 2020 年 9 月召开的 SCEC 年会^②上,与 CSEP 2.0 阶段相关的地震可预测性问题讨论、软件研发及不同预测模型评估等学术报告吸引了众多关注,如 Werner (2020) 根据 CSEP 计划中采用的严格统计检验方法,针对地震前观测或计算得到的若干前兆指标的评估进行了讨论,其中提到了这项工作前期需具备的若干工作基础。Savran 等 (2020) 针对采用 UCERF3 – ETAS 模型对 2019 年 Ridgecrest 地震序列的分析进行了评估。

在模型发展研究方面,Shcherbakov (2020) 利用 Bayesian 方法对与余震预测与检验的工作进行了讨论,认为基于 ETAS 模型的贝叶斯分析预测方法比基于 Omori 公式极值分布的传统预测方法效果更好;Ogata 和 Omi (2020) 使用统计地震学方法对地震发生早期进行实时监视以及早期预测,利用时空传染型余震序列模型 (ETAS) 给出震前在空间上呈显著平静、活跃或者发生地震活动迁移的区域,并建议将类似的异常现象加入到对未来发生大地震的统计检验中。Zhuang (2020) 介绍了最新研发的加入震源机制解成分的 ETAS 模型版本,以期更好地符合实际地震的空间展布情况,通过分析实际数据,发现可以有效改进模型拟合水平。可以看出,无论在针对地震可预测性研究的科学问题和具体应用方面,还是在针对某种模型加以改进和扩展方面,与 CSEP 2.0 相关的工作内容在 1.0 阶段的基础上均有了新的进展。

3 全球不同测试区 CSEP 2.0 阶段主要工作内容

CSEP 工作是当前 SCEC 日常研究工作的重要

组成部分,在近年来的若干次 SCEC 年会中,与 CSEP 内容相关的研讨会为 CSEP 研发专家提供了讨论和交流的平台。

表 2 总结归纳了意大利、日本、新西兰、中国等 4 个 CSEP 测试区在 CSEP 2.0 阶段提出的相关工作内容。其中,意大利 CSEP 测试中心工作主要针对加入多种模型的复合模型进行研究,形成可操作的地震预报的工作流程,并将 CSEP 1.0 阶段的系统代码移植到新的开发平台 (Pace, Peruzza, 2009; Falcone *et al*, 2017)。日本 CSEP 中心则试图将研究对象进一步扩大,即在原有基础上对地震的时间、空间和震级等不同要素发展相应的预测评估方法,并对早期地震活动行为和地震丛集进行自动检测 (Ogata, Omi, 2020)。值得一提的是,日本测试区于 2019 年 8 月在日本箱根顺利召开了第 11 届国际统计地震学研讨会,该研讨会主要针对以下几个问题进行了讨论:①地震活动性分析统计模型和方法的发展;②地震物理学;③地震预报、检验和减灾;④将统计地震学扩展到地震活动之外的研究;⑤利用贝叶斯方法分析地震大数据的研究进展。新西兰 CSEP 测试中心则在原有工作基础上深入研究已有模型包含参数的物理意义,进一步发展基于震前地震活动的预测模型,在 N – test、L – test 等统计检验中去除 Poisson 似然模型,采用类似 K – S 检验的方法对实际情况与模型给出的目标地震发生率情况进行对比 (Christophersen *et al*, 2018)。新西兰测试区未来将在数据质量控制、经费筹集和与其他测试区的国际合作方面采取新的措施,相比可以提供严格检验模型效能的平台,该测试区将更倾向于帮助模型研发人员构建性能更好的预报模型。中国测试区 CSEP 工作团队将在南北地震带地区 CSEP 1.0 工作的基础上,整合研发团队建设新的 CSEP 2.0 检验平台,并采用更多的预报模型和统计检验方法对中国地震科学实验场开展试验 (Zhang *et al*, 2019)。可见,不同 CSEP 测试中心在 2.0 阶段中提出了明确的研究目标,并与 CSEP 1.0 阶段做到了良好的工作衔接。

① <https://www.scec.org/meetings/2019/am>.

② <https://www.scec.org/meetings/2020/am>.

表 2 不同 CSEP 2.0 测试中心主要工作方向
Tab. 2 Main work directions of CSEP 2.0 testing centers in different countries

国家	CSEP 2.0 工作内容
意大利	<div><div>1. 针对 2014—2019 年进行的复合模型预测试验开展测试评估</div><div>2. 基于 RISE (Real-time earthquake risk reduction for a resilient Europe) 项目发展新的测试工具</div><div>3. 基于 RISE 项目发展新的地震预测模型</div><div>4. 形成具有可操作地震预报的工作流程</div><div>5. 移植所有的系统代码</div><div>6. 移除地震活动时间和空间上呈现 Poisson 分布的假设</div><div>7. 所研发模型需要包含由于科研人员认知水平的差异而导致的不确定性</div><div>8. 需要更多模型参与研究空间网格的离散化问题</div><div>9. CSEP 2.0 阶段应该在 CSEP 1.0 基础上做出最小的更改</div><div>10. 应对合成地震目录更加关注</div></div>
日本	<div><div>1. 针对时间、空间、震级等不同地震要素发展不同的评估方法以测试相应预测模型的效能</div><div>2. 对日本内陆地区开展 4~5 级重复地震的预报研究</div><div>3. 加强不同单位合作, 根据自动化的目录处理进行实时预测</div><div>4. 对早期地震活动行为的自动估计</div><div>5. 利用稳态 ETAS 模型进行地震丛集的自动检测</div><div>6. 深入在 Kanto 地区的 3D 模型研究, 包含 3D 球形模型、3D 时空 ETAS 模型和 3D HIST-ETAS 模型</div><div>7. 完善对加入震源机制解的 ETAS 版本的研究</div></div>
新西兰	<div><div>1. 深入理解已有模型包含参数的意义</div><div>2. 进一步发展 EEPAS (Every Earthquake a Precursor According to Scale) 模型</div><div>3. 数据质量存在问题, 如 GeoNet 数据</div><div>4. 为服务一定时长的研究人员提供经费存在困难</div><div>5. 探讨如何与其它测试中心进行更加广泛的合作</div><div>6. 从自动运行的 CSEP 系统中分离出检验模块</div><div>7. 发展基于已有模型、警报函数和地球物理数据等不同方面的综合模块</div><div>8. 提高科学界对 CSEP 活动的参与程度</div></div>
中国	<div><div>1. 建设新的 CSEP 2.0 测试平台, 倡导提交不同类型的模型进行严格评估</div><div>2. 应用更多的预报模型, 如将 ETAS 模型、加卸载相应比 (LURR) 算法等模型应用于中国地震科学试验区域</div><div>3. 引入和扩展其他统计检验方法在 CN-CSEP 测试区的应用</div></div>

4 中国 CSEP 2.0 阶段测试区的初步设计与模型试验

4.1 中国 CSEP 2.0 测试区初步设计

由于 CSEP 测试中心的主要功能为针对不同尺度的预测模型结果开展效能检验, 分析预测结果与实际观测的一致性以及模型间的信息增益 (Jordan, 2006; Schorlemmer *et al*, 2006), 因此在正式向全球模型研发专家征集不同种类的预测模型之前, 需要针对测试区的空间范围、基础数据格式、预测结果等情况进行统一设置, 以减少未来

由于不同团队提供的数据规范差别而造成的混乱 (Bayona *et al*, 2022)。CSEP 1.0 阶段以中国南北地震带地区作为测试区, 相关学者针对预测模型所需的空间网格、震级下限、模型自身的参数设置等进行了诸多研究 (韩立波等, 2012)。自 CSEP 2.0 阶段将中国地震科学实验场作为新的测试区以来, 通过“中国地震科学实验场的地震可预测性国际合作研究”国际合作项目的实施, 我国已初步建设了 CSEP 地震可预测性检验的中国分中心, 实现了与该项目全球网络基础设施的对接, 目前包含加卸载响应比、态矢量、地震综合概率、图像信息学算法、传染型余震序列模型、

b 值等多种预测模块以及 R 值、ROC 等预测检验模块。张盛峰 (2019) 前期已对与该区域对应的地震目录数据、测震台站分布、历史地震目录等情况进行了概括性的总结。为了与不同测试中心的具体工作进一步接轨, 本工作按照国际 CSEP 测试中心的设置经验对参加 CSEP 2.0 阶段的中国测试区进行了网格化定义, 并简要分析该区域地震监测能力的时空分布情况 (Gerstenberger, Schorlemmer, 2007)。图 2 给出了中国 CSEP 2.0 测试区、数据收集区的空间位置, 具体为: 中国地震科学实验场区域边界 (红框) 外围 50 km 范围定义为测试区范围 (黑框), 表示为预测模型给出预测结果需要覆盖的空间范围; 测试区外围 50 km 范围表示数据收集区 (灰色线框), 主要作为预测模型计算所需数据的空间收集范围, 目的是为了完全覆盖测试区域并保证数据资料在空间分布方面的完整性情况。空间网格设置为 $0.1^\circ \times 0.1^\circ$, 且使用网格中心点的位置定义每一网格。

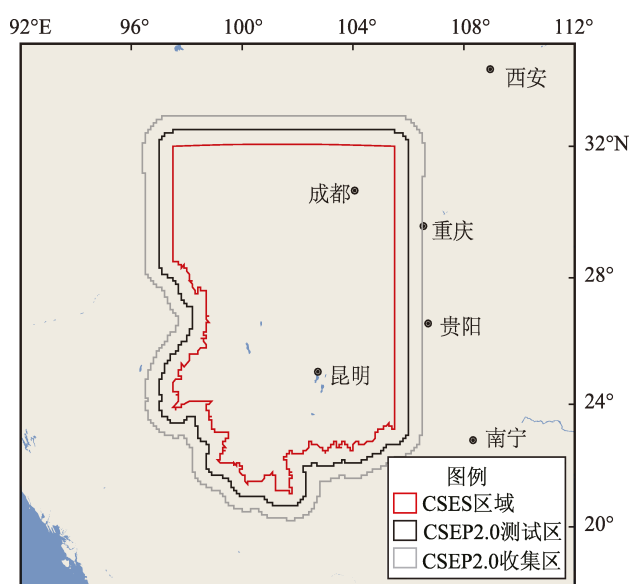


图 2 CSEP 2.0 阶段中国测试区的空间区域定义

Fig. 2 Spatial definition of the China CSEP 2.0 testing region

地震目录的完整性水平是预测模型中选取震级下限参数的重要参考指标。为了在模型预测试验中更准确的把握 CSEP 2.0 阶段中国测试区内地震目录的完整性情况, 本文给出了这一区域 1970 年以来地震目录完整性水平的演化 (图 3) 和空间分布情况 (图 4)。图 3 中深绿色曲线为使用

Entire Magnitude Range (EMR) 技术 (Woessner, Wiemer, 2005) 对测试区地震目录进行扫描得到的完整性震级时间变化, 红色横线和相应文字表示大致划定的不同时段可选取的完整性水平。从图 3a 中可以看出, 主要由于该区域地震监测台站在不同时期的建设情况, 地震事件记录的水平在不同时段存在较大的差异。在需要设置震级下限以保证数据完整性的模型中, 按照以上参考则可基本保证数据在相应时段内的完整性要求 (韩立波等, 2012)。从图 3b 可看出, 由于一些大震的发生, 通常导致震后短时间内发生余震事件的漏记情况, 因此在进行余震预测或者短期预测试验时, 震级下限的设置需要更加保守的选取 (Zhang *et al.*, 2015, 2023)。从图 4 所示的由 BMC (Bayesian Magnitude of Completeness) 方法 (Li *et al.*, 2023) 计算得到的 M_c 空间分布来看, 测试区西部的青藏高原区域由于监测水平较低, M_c 为 3.0~4.0; 高于 4.0 的区域主要为境外区域; 对于测试区内部区域, 大部分区域 M_c 为 1.0~2.0, 仅西部少数区域为 2.0~2.5。

4.2 图像信息学 (PI) 算法预测及效能检验

传统的图像信息学 (PI) 算法主要是一种根据背景参考时间窗和异常学习时间窗内包含的地震活动, 给出预测时间窗内目标地震较高发生趋势的中长期预测算法。其预测窗口一般为 5 年, 预测结果常以显示的“热点”型预测结果表示 (Holliday *et al.*, 2007; Zhang *et al.*, 2015; 张盛峰等, 2017)。CSEP 1.0 阶段, 在美国南加州、日本、中国南北地震带等地区进行了很多试验, 结果表明这一算法对于分析地震活动的涨落特征以及在中长期地震趋势预测方面具有一定的优势 (蒋长胜, 吴忠良, 2008; Jiang, Wu, 2011; 张小涛等, 2014)。本文为了与国际上其它 CSEP 测试中心工作框架下的地震预测模型预测结果形式一致, 且便于后期使用多种针对这一预测结果的检验方法进行效能评估, 尝试使用基于地震发生率预测结果的 PI 算法对上述定义的 CSEP 2.0 区域开展预测试验, 并检验模型效果 (Holliday *et al.*, 2007)。本文选取震级下限 3.0 级, PI 算法中的背景参考时间窗为 1999 年 1 月 1 日至 2008 年 12 月 31 日, 异常学习时间窗为 2009 年 1 月 1 日至 2013 年 12 月 31 日。图 5 为使用 PI 算法计算得到的 CSEP 2.0

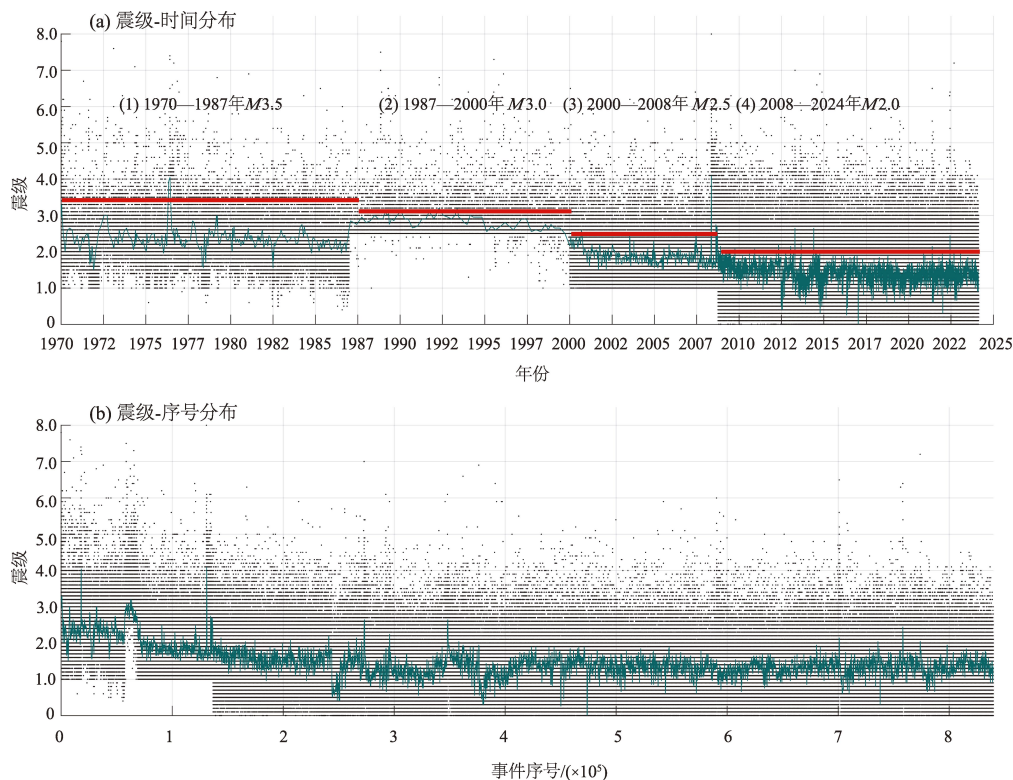


图3 CSEP 2.0 中国测试区地震目录时序分析

Fig. 3 Temporal analysis of the earthquake catalog in the China CSEP 2.0 testing region

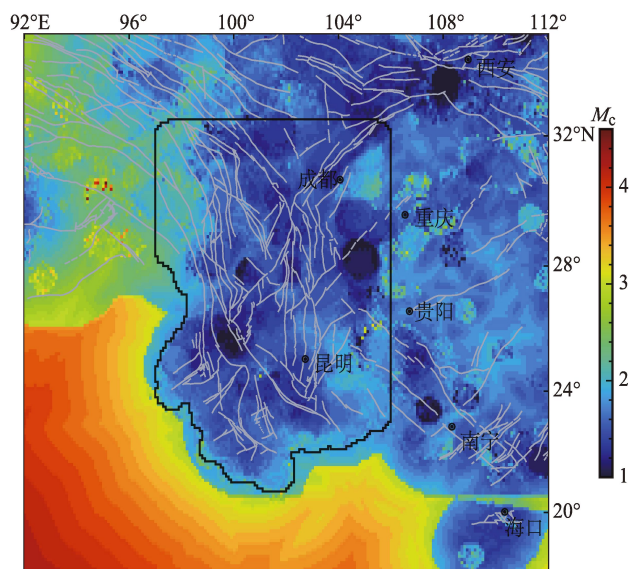


图4 CSEP 2.0 中国测试区地震监测能力空间分布

Fig. 4 Spatial distribution of the earthquake monitoring capability in the China CSEP 2.0 testing region

中国测试区 2014 年 1 月 1 日至 2018 年 12 月 31 日 $M \geq 6.0$ 地震的发生率预测结果, 图中黄色圆点为实际发生的 $M \geq 6$ 地震。相比基于“热点”显示的警报型预测结果, 基于地震发生率的预测结果在

每个空间网格内均有显示, 而非仅显示一定水平以上的热点网格, 因此也更加便于使用不同的统计检验技术对所有网格进行标准化效能检验。

本文选取常用的针对空间发生率预测结果的 ROC (Receiver Operating Characteristic) 检验、Molchan 图表法、空间 (Spatial) S 检验和地震数 (Number) N 检验方法对以上预测结果进行评估。ROC 检验按照地震发生率递减顺序排列单元格后的归一化累计发生率和归一化的累计面积分别计算击中率和虚报率, 可以评估在不同阈值水平下预报结果的特异性 (或平稳性) 的高低。Molchan 图表法则与 ROC 检验类似, 通过计算漏报率以及异常时空占有率来评价预测与实际观测的一致性。N 检验对比实际观测地震数目与模型预测结果, 通过设置一定的置信水平 (如 0.05), 计算发生至少和不多于 N_{obs} 个地震的两个参数值 (δ_1 和 δ_2) 来给出检验结果 (Schorlemmer *et al*, 2007; Zechar *et al*, 2010a)。S 检验则将每个空间网格的期望发生率与震级分段相加, 以分离出预测的空间部分, 并将得到的空间发生率归一化为实际观测地震的总数。通过评估每个单元的 Poisson 似然函数,

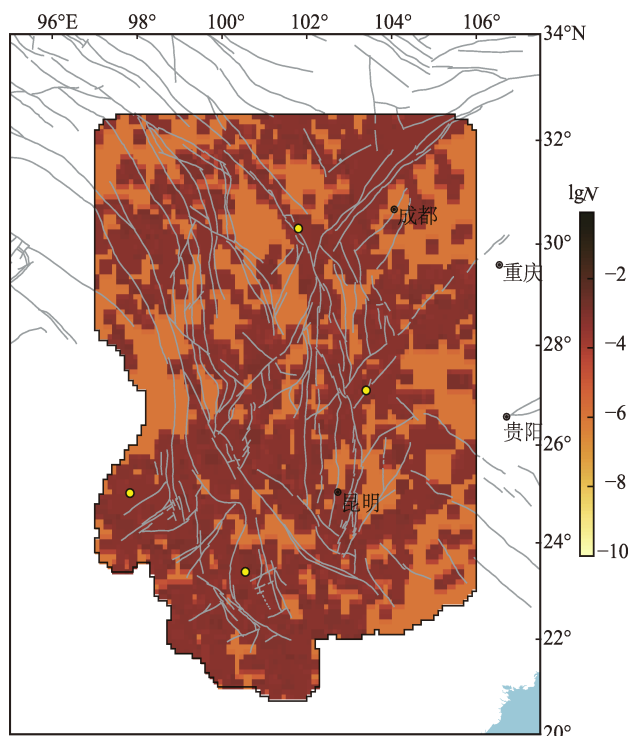


图5 基于PI算法得到的CSEP 2.0中国测试区5年尺度(2014年1月1日至2018年12月31日) $M \geq 6.0$ 地震发生率预测结果

Fig. 5 Forecasting result of earthquake occurrence rate for the China CSEP 2.0 testing region over a 5-year scale based on the PI algorithm from January 1, 2014 to December 31, 2018

计算出每个网格的空间联合对数似然并将所有空间网格对数似然相加。为了评估观测到的对数似然比得分是否由预测产生,需要通过模拟得到与预测一致的空间对数似然比得分分布。为了评估地震的实际发生位置与空间预测之间的一致性,通常设置置信水平 α 为0.05来检查观测值在模拟值分布范围内的位置(Zechar *et al.*, 2010b)

图6为使用以上几种评估方法对PI算法预测的评估结果,ROC检验和Molchan检验均显示PI算法优于随机预测,这一结果基本与针对以往得到的“热点”型预测结果的检验情况一致。这是因为两种检验方法均是通过将预测结果排序后划定为最终不同阈值下的警报型结果,进而计算相应的评估指标。S检验和N检验可以给出其它角度的评估结果,如S检验显示PI算法预测结果在空间上与观测结果具有较高的一致性,而地震数N检验显示预测与实际观测情况一致性较差,预测

的地震数目相比实际观测偏少,在95%置信水平下没有通过检验。可以看出,对于某一预测结果,不仅需要使用传统固有的检验技术检验其结果,同时需要从其它角度对这一结果进行多次评估,进而获得更多检验信息增益或者截然相反的认识。

5 讨论

无论是已经结束的CSEP 1.0阶段,还是2018年开始的CSEP 2.0阶段,该计划本意是对于不同种类的地震预测模型或算法开展“竞赛式”的效能评估,而这种效能评估所采取的标准则是基于严格统计检验技术的理论方法,虽然采用统一的预测规则和数据来源是该计划的基础要求,但由于不同模型给出的预测结果差异,如可以给出“警报型”或者“概率型”的目标地震发生趋势,因此难以使用固定的某一种评估方法进行评估(Schorlemmer *et al.*, 2007)。对于某种具体的预测算法,其研发过程通常会受到研发人或者研究人员主观因素的影响,如在算法应用前期开展的回溯性预测试验中,常常要设置研究区范围、区域网格化设置、模型初始条件和参数、地震活动拟合次数以及最终预报结果的呈现方式等,而模型的最终效果往往对以上因素具有较强的依赖性。因此,研发过程中如何有效地减弱研究人员主观因素的影响,或者如何在模型影响因素和效能评估结果之间达到良好的平衡,可能是研究人员需要考虑的问题。但目前看来,通过尽量减弱这些“非随机因素”的比重,按照“一步一个脚印”的工作思路对模型原理和运行环境进行改进,才能与CSEP研究计划的工作理念和初衷相一致(Jordan, 2006)。作为CSEP计划的积极参与者,近年来我国在针对南北地震带地区若干预测模型和检验方法的研究试验方面取得了很多认识,同时在不同预测模型和评估方法应用的丰富度或者针对某一模型的研究深度方面仍有发展的空间。当前我国建立的CSEP 2.0测试中心平台包含了传染型余震序列(ETAS)模型、图像信息学(PI)算法、相对强度(RI)算法、加卸载响应比(LURR)等国内外常用的地震预测方法以及相对应的ROC检验、Molchan图表法等检验方法,与国际上其它测试中心包含的模型相比,在操作机制、结果产出

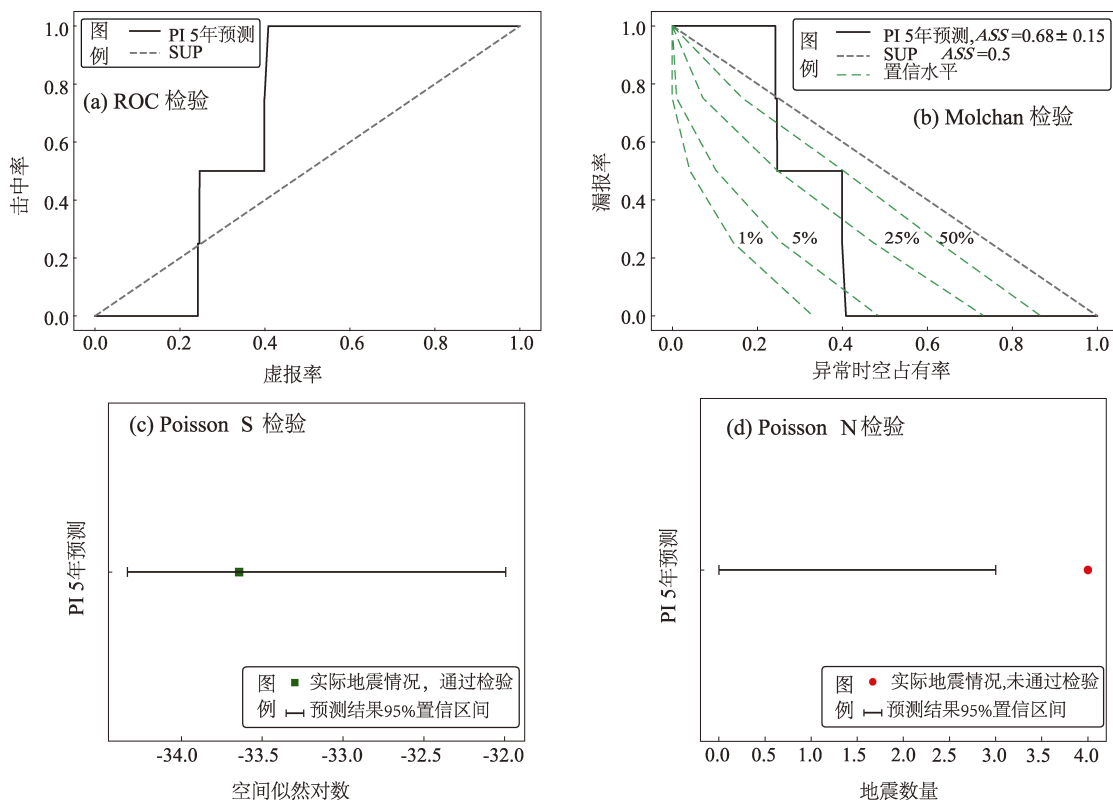


图6 对PI地震发生率预测结果进行的多种检验方法的效能评估结果

Fig. 6 Performance evaluation of multiple testing methods for the forecasting result of the PI earthquake occurrence rate

等方面仍存在诸多不同。本文试图以 CSEP 2.0 中国测试区的确定以及 PI 算法预测试验为开端, 按照类似的思路为其它更多的预测模型步入国际竞技轨道提供参考。此外, 在当前不同测试中心 CSEP 2.0 工作步入正轨之际, 针对中国地震科学实验场区域, 除了努力发展与我国地震活动特点相适应的自有预测模型和检验方法以外, 与国外同行开展广泛的科技合作, 学习借鉴优秀的模型提出、发展和应用的思路经验, 对于我国 CSEP 2.0 阶段工作的下一步开展将具有重要意义。

6 结论

本文总结了 CSEP 2.0 阶段工作的背景、不同阶段设计、主要进展及不同测试区工作内容, 与 CSEP 1.0 阶段工作相比, CSEP 2.0 研究思路总体框架并未发生很大变化, 但对不同测试区所包含模型的研究深度、CSEP 分析系统计算能力及优化方面均试图进行升级, 在当前计算机技术和模型算法水平不断得到提高的背景下, 这一国际合作项目同样也表现

出了对于技术和工作流程进行更新的需求。

本工作针对 CSEP 2.0 中国测试区进行了规范化的定义, 并采用 PI 算法得到了 5 年尺度地震发生率的预测结果, 使用 ROC 检验、Molchan 检验、S 检验和 N 检验进行了效能评估, 结果显示相比随机预测, PI 算法预测效果较好, 除预测地震数较实际情况偏少以外, 其它检验方法均显示预测结果相比实际观测可通过严格的一致性检验。而针对某一种预测模型的预测结果, 可以看出, 类似本工作中的采用多种检验技术进行效能评估, 而非采用固定的仅代表一种检验角度的方法, 将是 CSEP 2.0 阶段工作的重要特点。

中国地震台网中心提供了分析所需的地震目录数据; 南方科技大学风险分析预测与管控研究院李佳威博士提供了基于 BMC 方法计算的 M_c 结果; 英国 Bristol 大学 Maximilian J. Werner 教授、Toño Bayona 博士、Francesco Serafini 博士, 新西兰地质与核科学研究所 (GNS) David Rhoades 教授、Kenny Graham 博士, 波茨坦德国地球科学研究中

心 (GFZ) Pablo Iturrieta Rebolledo 博士, 中国地震局地震预测研究所崔子建副研究员等与作者针对 CSEP 2.0 工作进行了相关讨论; CSEP 1.0 阶段相关专家为作者参与 CSEP 研究工作和程序计算提供了指导, 在此一并表示感谢。

参考文献:

- 韩立波, 蒋长胜, 李艳娥, 等. 2012. 用于地震可预测性 CSEP 计划的南北地震带地区地震最小完整性震级 M_c 研究[J]. 地震, 32(1): 17–27.
- Han L B, Jiang C S, Li Y E, *et al.* 2012. Minimum magnitude of completeness in the North–South seismic belt for collaborative study of earthquake predictability[J]. Earthquake, 32(1): 17–27. (in Chinese)
- 蒋长胜, 吴忠良. 2008. 对地震预测的一个统计物理算法在川滇地区的回溯性预测检验[J]. 中国科学: 地球科学, 38(7): 852–861.
- Jiang C S, Wu Z L. 2008. A statistical physics algorithm for earthquake prediction in Sichuan-Yunnan region[J]. Scientia Sinica Terrae, 38(7): 852–861. (in Chinese)
- 吴忠良, 陈运泰. 2002. 地震预测与统计物理[J]. 物理, 31(6): 365–371.
- Wu Z L, Chen Y T. 2002. Earthquake prediction and statistical physics[J]. Physics, 31(6): 365–371. (in Chinese)
- 吴忠良, 王龙, 车时, 等. 2021. 中国地震科学实验场: 认识与实践[J]. 地球与行星物理评论, 52(3): 348–352.
- Wu Z L, Wang L, Che S, *et al.* 2021. China seismic experimental site (CSES): planning and test[J]. Reviews of Geophysics and Planetary Physics, 52(3): 348–352. (in Chinese)
- 张盛峰, 郑建常, 蒋长胜, 等. 2017. 图像信息学 (PI) 算法计算参数优化分析——以山东及相邻地区为例[J]. 地球物理学报, 60(12): 4633–4643.
- Zhang S F, Zheng J C, Jiang C S, *et al.* 2017. Optimal analysis of parameter settings in pattern informatics (PI) algorithm: An example of Shandong and adjacent areas[J]. Chinese Journal of Geophysics, 60(12): 4633–4643. (in Chinese)
- 张盛峰, 张永仙. 2021. “地震可预测性国际合作研究”——1.0 阶段工作理念及成果[J]. 地震, 41(4): 203–217.
- Zhang S F, Zhang Y X. 2021. Collaboratory for the study of earthquake predictability——working philosophy and achievements of CSEP 1.0[J]. Earthquake, 41(4): 203–217. (in Chinese)
- 张盛峰. 2019. 中国地震科学实验场暨 CSEP–CN 计划的若干统计地震学问题[D]. 北京: 中国地震局地球物理研究所.
- Zhang S F. 2019. Statistical seismological problems associated with CSEP–CN testing region in the context of CSES[D]. Beijing: Institute of Geophysics, China Earthquake Administration. (in Chinese)
- 张小涛, 张永仙, 夏彩韵, 等. 2014. 利用图像信息方法研究芦山 $M_s 7.0$ 地震前川滇及附近地区的图像异常[J]. 地震学报, 36(5): 780–789.
- Zhang X T, Zhang Y X, Xia C Y, *et al.* 2014. Anomalous seismic activities in the Sichuan-Yunnan region and its adjacent areas before the Lushan $M_s 7.0$ earthquake by the pattern informatics method[J]. Acta Seismologica Sinica, 36(5): 780–789. (in Chinese)
- Bayona J A, Savran W H, Rhoades D A, *et al.* 2022. Prospective evaluation of multiplicative hybrid earthquake forecasting models in California[J]. Geophysical Journal International, 229(3): 1736–1753.
- Christophersen A, Rhoades D A, Gerstenberger M C, *et al.* 2018. Highlights from the first ten years of the New Zealand earthquake forecast testing center[J]. Seismological Research Letters, 89(4): 1229–1237.
- Falcone G, Marzocchi W, Murru M, *et al.* 2017. Earthquake forecasting system in Italy[R/OL]. (2017–12–01) [2024–06–30]. <https://ui.adsabs.harvard.edu/abs/2017AGUFM.S53B0696F>.
- Field E H. 2007. Overview of the working group for the development of regional earthquake likelihood models (RELM)[J]. Seismological Research Letters, 78(1): 7–16.
- Gerstenberger M C, Schorlemmer D. 2007. RELM testing center[J]. Seismological Research Letters, 78(1): 30–36.
- Holliday J R, Chen C C, Tiampo K F, *et al.* 2007. A RELM earthquake forecast based on pattern informatics[J]. Seismological Research Letters, 78(1): 87–93.
- Jiang C S, Wu Z L. 2011. PI forecast with or without de-clustering: An experiment for the Sichuan-Yunnan region[J]. Natural Hazards and Earth System Science, 11(3): 697–706.
- Jordan T H. 2006. Earthquake predictability, brick by brick[J]. Seismological Research Letters, 77(1): 3–6.
- Li J W, Mignan A, Sornette D, *et al.* 2023. Predicting the future performance of the planned seismic network in Chinese mainland[J]. Seismological Research Letters, 94(6): 2698–2711.
- Michael A J, Werner M J. 2018. Preface to the focus section on the collaborative for the study of earthquake predictability (CSEP): New results and future directions[J]. Seismological Research Letters, 89(4): 1226–1228.
- Ogata Y, Omi T. 2020. Statistical monitoring and early forecasting of earthquake sequence: Case studies after the 2019 $M_6.4$ Searles Valley Earthquake, California[R/OL]. (2020–08–01) [2024–06–30]. <https://central.scec.org/meetings/2020/am/poster/098>.
- Pace B, Peruzza L. 2009. An earthquake rupture forecast model for central Italy submitted to CSEP project[R/OL]. (2009–04–19) [2024–06–30]. <http://meetings.copernicus.org/egu2009>.
- Savran W H, Werner M J, Marzocchi W, *et al.* 2020. Pseudoprospective evaluation of UCERF3–ETAS forecasts during the 2019 ridgecrest sequence[J]. Bulletin of the Seismological Society of America, 110(4): 1799–1817.
- Schorlemmer D, Gerstenberger M, Wiemer S, *et al.* 2007. Earthquake likelihood model testing[J]. Seismological Research Letters, 78(1): 17–29.
- Schorlemmer D, Jordan T H, Zecher J D, *et al.* 2006. Collaboratory for the study of earthquake predictability[R/OL]. (2006–12–01) [2024–06–30]. <https://ui.adsabs.harvard.edu/abs/2006AGUFM>.

- S12A.01S.
- Shcherbakov R. 2020. Bayesian framework for aftershock forecasting and testing[R/OL]. (2020-08-01)[2024-06-30]. <https://central.scec.org/meetings/2020/am/poster/097>.
- Strader A, Werner M, Bayona J, *et al.* 2018. Prospective evaluation of global earthquake forecast models; 2 yrs of observations provide preliminary support for merging smoothed seismicity with geodetic strain rates[J]. *Seismological Research Letters*, 89(4):1262-1271.
- Taroni M, Marzocchi W, Schorlemmer D, *et al.* 2018. Prospective CSEP evaluation of 1-day, 3-month, and 5-yr earthquake forecasts for Italy[J]. *Seismological Research Letters*, 89(4):1251-1261.
- Werner M J. 2020. Experimental design for testing hypotheses of earthquake precursors[R/OL]. (2020-08-03)[2024-06-30]. <https://southern.scec.org/publication/10245>.
- Woessner J, Wiemer S. 2005. Assessing the quality of earthquake catalogues; estimating the magnitude of completeness and its uncertainty[J]. *Bulletin of the Seismological Society of America*, 95(2):684-698.
- Zechar J D, Gerstenberger M C, Rhoades D A. 2010a. Likelihood-based tests for evaluating space-rate-magnitude earthquake forecasts[J]. *Bulletin of the Seismological Society of America*, 100(3):1184-1195.
- Zechar J D, Schorlemmer D, Liukis M, *et al.* 2010b. The collaboratory for the study of earthquake predictability perspective on computational earthquake science[J]. *Concurrency and Computation: Practice and Experience*, 22(2):1836-1847.
- Zhang S F, Wu Z L, Jiang C S. 2015. The central China North-South seismic belt; seismicity, ergodicity, and five-year PI forecast in testing[J]. *Pageoph*, 173(1):245-254.
- Zhang S F, Wu Z L, Zhang Y X. 2023. Is the September 5, 2022, Luding M_s 6.8 earthquake an 'unexpected' event? [J]. *Earthquake Science*, 36(1):76-80.
- Zhang Y X, Wu Z L, Zhang X D, *et al.* 2019. CSEP 2 from China Perspective[R/OL]. (2019-08-15)[2024-06-30]. <https://southern.scec.org/publication/9711>.
- Zhuang J C. 2020. An ETAS model incorporated with focal mechanisms[R/OL]. (2020-08-01)[2024-06-30]. <https://central.scec.org/meetings/2020/am/poster/101>.

Phase 2.0 of CSEP Work: Progress, Design and Model Forecasting Experiment in the China Test Region

ZHANG Shengfeng^{1,2}, ZHANG Yongxian¹, LI Sha³

(1. Institute of Earthquake Forecasting, China Earthquake Administration, Beijing 100036, China)

(2. GNS Science, Avalon 5011, Lower Hutt 5040, New Zealand)

(3. Earthquake Agency of Guangxi Zhuang Autonomous Region, Nanning 530022, Guangxi, China)

Abstract

Since the establishment of the Regional Earthquake Likelihood Models (RELM) working group, jointly initiated by the Southern California Earthquake Center (SCEC) and the U. S. Geological Survey (USGS) in 2000, and the subsequent development of the Collaboratory for the Study of Earthquake Predictability (CSEP) project, various earthquake forecast models and statistical evaluation methods for assessing their performance have been widely applied. Based on the introduction of Phase 1.0 of CSEP work, this paper primarily introduces the aspects of CSEP 2.0, including its background, overall development design, several key conference discussions, and work plans for different test centers. Additionally, regarding the current participation of China, this paper preliminarily designs the scope of the China CSEP 2.0 test region and conducts a pilot experiment using the Pattern Informatics (PI) algorithm. It is hoped that this work can provide a reference for other models to participate in China CSEP 2.0 work and for the next step of our international co-operation research to be carried out in an orderly manner.

Keywords: CSEP 2.0 testing centers; China testing area; PI algorithm; CSEP performance testing techniques