

地震科技前沿快报

2023 年第 6 期（月刊总第 53 期）

中国地震局科学技术委员会
中国地震局科技与国际合作司

中国科学院兰州文献情报中心
中国科学技术信息研究所

本期概要

1、利用遥感技术进行滑坡检测、监测和预测综述及展望

检测、监测和预测是管理滑坡灾害风险的基础，遥感技术是这项研究中不可或缺的工具，通过空基平台和地基平台可提供系统、综合和经济有效的观测。文章综述了多光谱、干涉卫星、地面干涉测量、激光扫描和多普勒雷达等遥感技术，描述了这些遥感技术的功能和发展，并概述了它们在检测、监测和预测方面的应用、优势和局限性。

结果表明，每种遥感技术都有其优势和限制，取决于应用目的（检测、监测或预测）、关注区域的大小、滑坡类型、变形模式以及滑坡带来的风险；因此，需要根据实际情况，将各种技术进行整合使用。未来 5-10 年，不同技术的整合将成为研究重点；另外，在滑坡监测和管理中，仪器成本与实际使用间的匹配也至关重要。滑坡灾害管理需考虑技术和社会因素，通过良好的沟通策略和先进的监测和预警系统，将可降低滑坡灾害风险。滑坡灾害的遥感技术综述成果对地震灾害防控现有技术优缺点总结及展望具有参考意义。

2、通过基于物理信息的人工智能-深度学习方法模拟地壳形变

地壳和上地幔的运动和形变为地震过程的演变和未来的地震风险提供了重要信息。地壳形变可以用位错模型来模拟，该模型将地壳中的地震断层表示为连续介质中的缺陷。

在本研究中，研究人员提出了一种基于物理的深度学习新方法模拟地震引起的地壳形变。神经网络通过将控制方程和边界条件纳入损失函数，可以表示任意几何结构和岩石力学性质的连续位移场。通过引入极坐标系作为边界条件，精确地模拟了断层上的位移不连续。研究人员通过对走滑断层的实例分析，说明了这种方法的有效性和实用性。

与传统方法相比，这种方法具有潜在优势，因为它可以直接扩展到高维问题、非弹性问题、非线性问题和逆问题。

3、通过不完整极性数据的机器学习改善地震震源机制的确定

地震震源机制通常使用 P 波初动极性来确定，并且需要充足的地震记录支持。最新的一种方法较少依赖高信噪比记录，其利用类似波形的相对极性测量来确定地震对之间的关系。

利用这些相对极性测量，可以使用区域地震台网生成微震序列中聚簇的复合震源机制。美国地质调查局采用迭代随机森林进行极性估算，以代替不可靠的极性测量。使用插补后的极性数据，可将地震事件聚类为不同的家族。作为案例研究，研究人员将这种改进后的复合机制工作流应用于 2020 年发生在美国加州北部 **Maacama** 断层附近地震群的多台站模板匹配目录。结果发现，完善后的方法产生了更高质量的震源机制。

利用遥感技术进行滑坡检测、监测和预测

滑坡是一类广泛发生的事件，在靠近定居点和基础设施的地方发生时，有可能产生灾难性影响。检测、监测和预测是管理滑坡风险的基础，这依赖于遥感技术（RST），包括对地空间观测、激光扫描和地面干涉测量。在本技术评述中，研究人员描述了遥感技术在滑坡分析和管理中的应用。卫星遥感技术用于检测和测量滑坡位移，提供各种时空尺度上的综合观测。地面传感器（包括地基干涉雷达、多普勒雷达和激光雷达）用于监测较小的区域，但由于其准确性、高采集频率和配置灵活性，在滑坡的实时监测和预警中越来越常用。每种遥感技术都有其优势和限制，这取决于应用目的（检测、监测或预测）、关注区域的大小、滑坡类型、变形模式以及滑坡带来的风险，因此通常最好将各种技术进行整合使用。更有效的滑坡风险管理需要更好地利用大数据、更具战略性地使用监测资源，并与滑坡易发区居民进行更好的沟通。相关成果发表在 2023 年 1 月的 *Nature Reviews Earth & Environment*。

滑坡是由构造、气候和人类活动驱动的自然灾害，在斜坡环境中普遍存在。全球范围内发生的滑坡事件众多，尤其亚洲受影响最严重。全球变化可能加剧滑坡风险，如降雨频率和强度的增加、永久冻土退化等。因此，国际组织和机构提倡的滑坡风险减轻政策包括了解风险、增强灾害应对准备工作，以实现有效响应。为实现这些目标，准确检测和监测不稳定斜坡的存在、速度和稳定条件及其随时间的演变变得至关重要。遥感技术成为滑坡研究中不可或缺的工具，通过空基平台和地基平台可提供对地表的系统、综合和经济有效的观察。

研究人员主要关注多光谱和干涉卫星、地面干涉测量、激光扫描和多普勒雷达等遥感技术，描述了这些遥感技术在滑坡科学中的功能和发展，

并概述了它们在检测、监测和预测方面的应用、优势和局限性。最后，提出了未来的研究和发展方向，以最大程度地改善它们在滑坡风险管理中的应用。

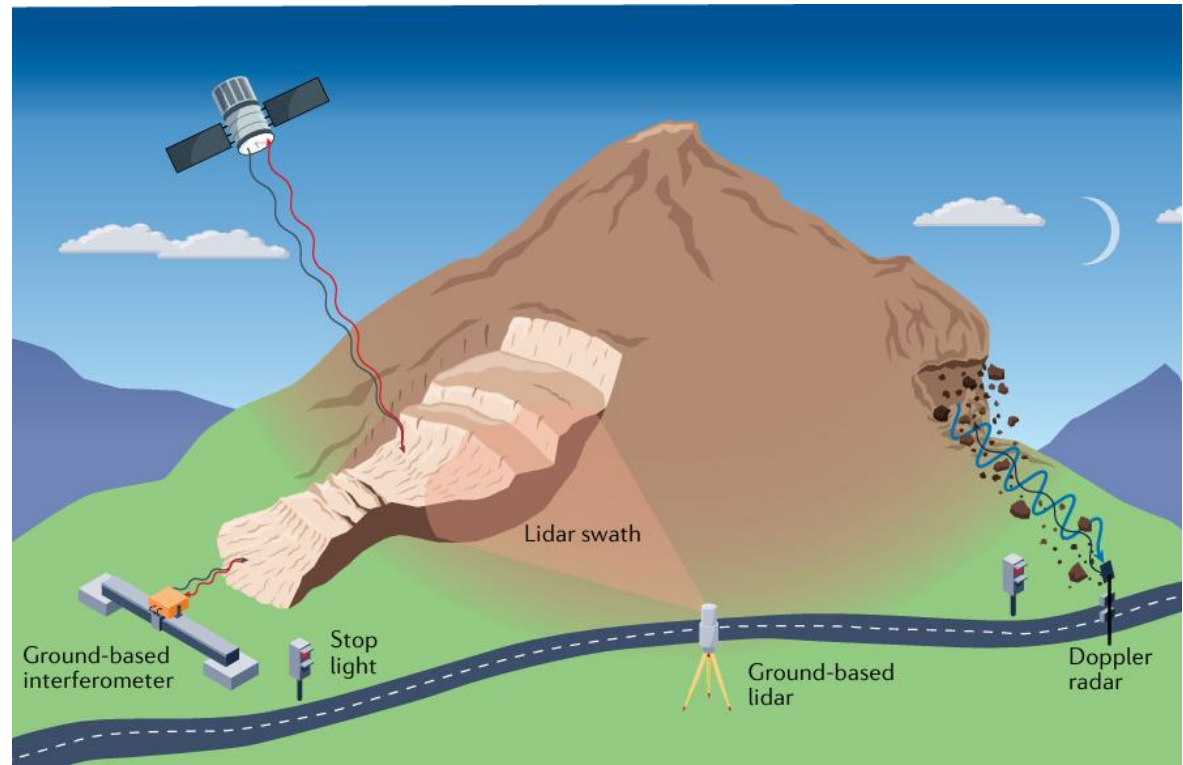


图 1 利用遥感技术进行现场监测

1. 空基技术

(1) 多光谱卫星传感器：多光谱卫星平台搭载被动式仪器，通过探测和测量地球表面的电磁辐射，提供了广泛的图像数据。这些卫星从太阳同步轨道垂直观测地球，确保阳光角度保持一致，使得可见光、近红外和短波红外光谱的图像获取成为可能。这些图像被广泛应用于制作滑坡清单地图和地表特征研究，而高分辨率卫星可以提供更详细的图像。在选择卫星数据时，需考虑重访时间（**revisiting time**）和幅宽（**swath width**）等因素，以满足滑坡研究的需求。当前可用的卫星平台增加了在滑坡事件或高

变形期间获取无云多光谱影像的可能性，但天气条件和日光仍然是主要限制因素。

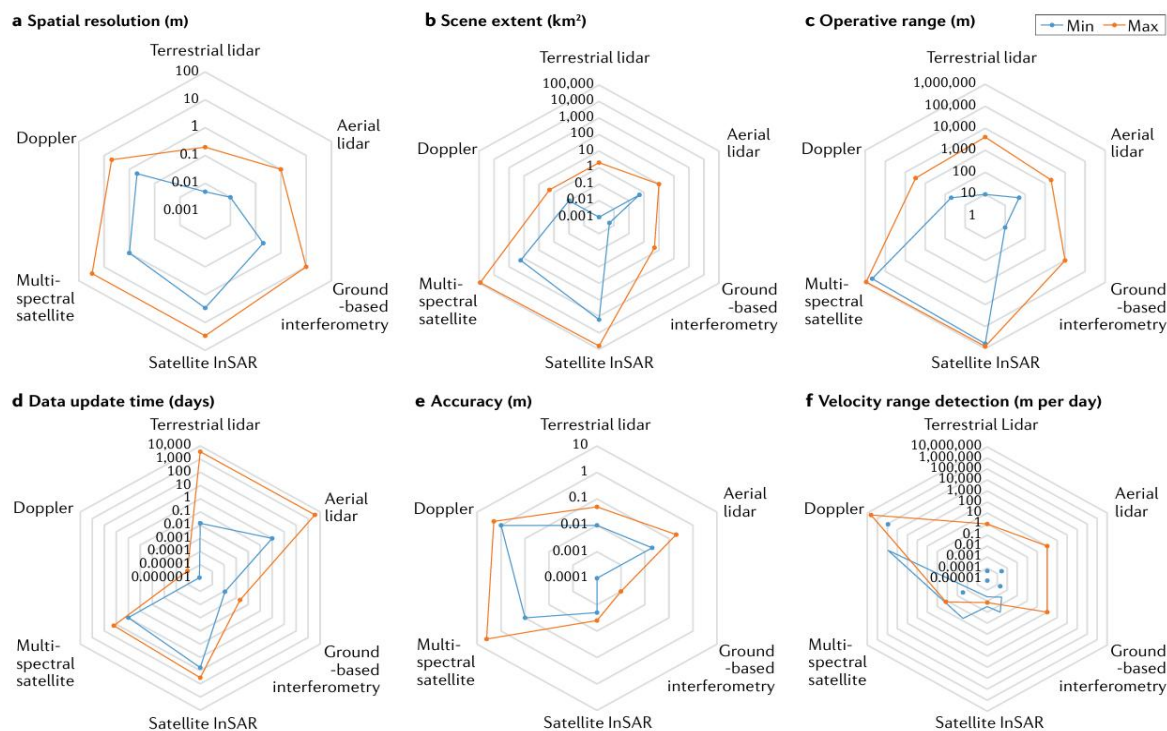


图2 不同遥感技术之间的比较

(2) 合成孔径雷达 (SAR) 传感器: SAR 技术利用主动传感器提供照明能量, 具有无视时间和季节限制的优势, 这对持续的滑坡监测和预测至关重要。SAR 图像以具有振幅和相位值的像素表示地面区域。差分干涉 SAR (DInSAR) 利用同一目标区采集的两幅连续雷达图像之间的相位差异测量地表运动, 可以测量 1 厘米到几十厘米或更小的地表运动 (取决于波长), 并在数十或数百平方千米的区域内具有适当的分辨率。多时相干涉 SAR (MTInSAR) 通过识别对应于单个像素或一组表现出稳定雷达特征的几个像素的测量点网格, 可增强干涉 SAR 的信息。在时间去相关和相干性丧失的情况下, 振幅分析可以替代干涉 SAR (InSAR) 进行表面变形成像, 提供更大范围的运动速率。

2. 地基技术

(1) 地面干涉测量：地面干涉测量技术是一种先进的监测方法，用于测量斜坡变形和人造结构物。通过合成孔径雷达和真实孔径雷达可实现实时位移数据和变形速率的获取，能够进行近乎实时的滑坡规模监测，并有可能进行预警。同时，这些数据还具有高精度（亚毫米）和高空间分辨率（1—10 m）的特征。该技术采集时间短，操作方便，并在几乎所有天气条件下可持续运作。目前，其已成功应用于滑坡监测、地质灾害预警、矿坑斜坡监测等领域。

(2) 多普勒雷达：多普勒雷达系统利用多普勒效应进行滑坡检测，如果滑坡体的尺寸和滑坡速度足够大，就会引起并检测到信号频率的偏差。多普勒雷达可提供快速滑坡事件（超过 1 m/s）的二维地图，包括滑坡速度、持续时间和轨迹等。同时，其还可发出警报，而警报可根据空间标准定制，并通过特定算法过滤误报。

(3) 激光雷达：激光雷达技术用于监测不稳定斜坡，主要因为其提供了所观测地表的详细数字表面模型。在该模型上，可投射其他空间信息，如干涉位移。如果比较在不同时间进行的多次扫描结果，则可以检测和监测位移场。全波形激光雷达系统能够数字化完整回波波形，提供高分辨率的多目标信息，因此可增加地面植被、地表形态等的细节信息。

3. 主要应用

(1) 检测：包括收集滑坡发生、分布和强度的数据与信息，以及识别受地表变形过程影响的元素。多光谱影像在已发生滑坡检测中被广泛使用，高分辨率（HR）和超高分辨率（VHR）传感器的增加可克服许多现有传感器的低空间分辨率限制。当检测不是针对缓慢移动的滑坡（尚未发生坍塌的区域）时，通常使用 InSAR 测量地面位移，其中 MTInSAR 在不同阶段的滑坡相关事件分析中非常有效。卫星应用（包括雷达和多光谱）

适用于扫描广泛区域并发现局部变形，特别是在野外活动具有挑战性或不可行的危险区域。在局部尺度上，地基技术具有高精度和灵活性，因此在滑坡检测中也越来越有帮助，而激光雷达则广泛用于滑坡表征和检测。

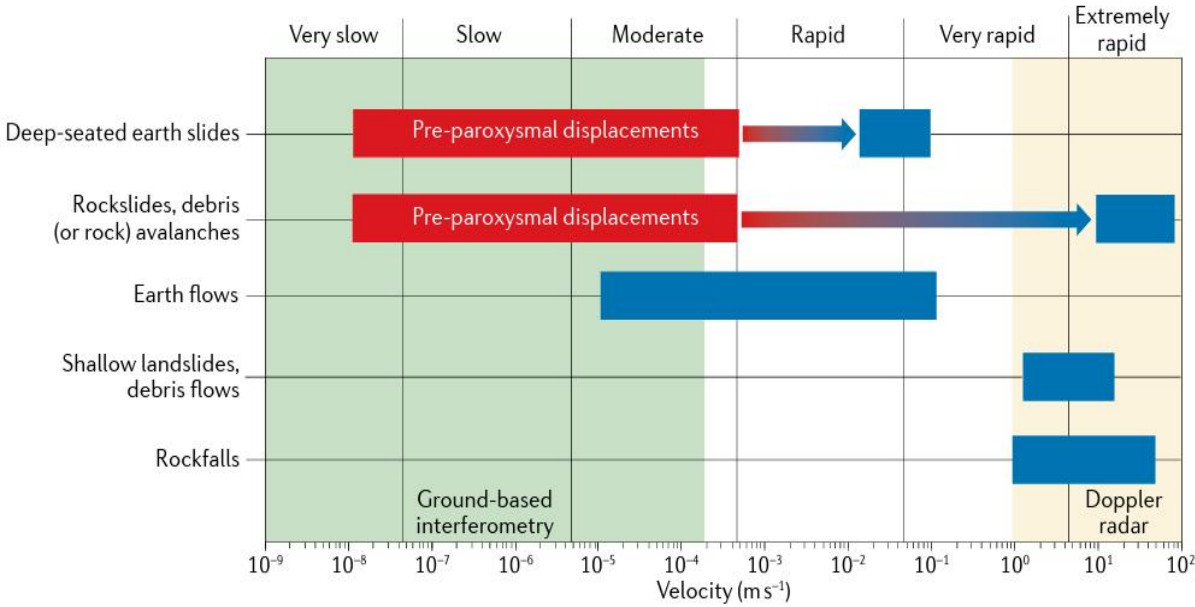


图3 地面干涉测量和多普勒雷达在常见滑坡中的作用范围

(2) 监测：滑坡发生后，测量关键参数（特别是随时间变化的表面位移）至关重要。多光谱成像适合滑坡运动的监测，超高分辨率卫星图像的出现和哨兵2号卫星的发射提高了多光谱成像监测的有效性。图像相关性分析可用于量化滑坡演化，推导亚像素位移。干涉测量适用于较慢速率的滑坡，多时相激光雷达可监测不同类型滑坡。卫星方法适合系统地追踪地面变形，而通过现场传感器可获取准确的高频数据，识别滑坡前兆迹象。

(3) 预测：滑坡监测和预测的目标是准确、及时地估计与滑坡发生相关的参数。滑坡的预测依赖于地面位移、降雨量和其他变量，位移数据在小规模斜坡上广泛使用，降雨在区域尺度上更适用。大规模滑坡通常在崩塌前显示出早期迹象，而坚硬岩石则倾向于突然坍塌。地基技术是预测滑坡时使用的主要方法，但哨兵1号卫星提升了 InSAR 在滑坡分析中的应

用，尤其是卫星多光谱数据的短周期重访时间等有助于大规模滑坡的预测和预警。

4. 总结和展望

遥感技术在滑坡的检测、监测和预测中起着关键作用。现在，新的和改进的技术能够收集多种类型边坡的不稳定性信息，以及多个尺度的信息。目前，滑坡监测和预测的瓶颈不仅在于收集滑坡数据，还在于如何充分地利用这些数据。科学界内更开放的共享政策以及更各界人士更密切的交流，可能是解决数据被低效利用的一种办法。总体而言，数据流通的越多，使用的可能性就越大。

除了更好地利用现有数据，滑坡监测的地理范围也必须扩大。在人口和/或滑坡易发性预计增加的发展中国家，以及难以采取减轻滑坡影响的预防政策的低收入国家，需要进行额外的研究和监测。

在未来的 5 到 10 年里，不同技术的整合应该成为滑坡科学的重点。另外，在滑坡监测和管理中，仪器成本与实际使用之间的匹配也至关重要。同时，建议在早期预警系统中整合基于三级蠕变的预测方法，以增加预警的准确性。滑坡管理需要考虑技术和社会因素，因此风险沟通和感知也至关重要。通过良好的沟通策略和先进的监测和预警系统，将可以提高风险意识，降低滑坡风险。

来源：Casagli N, Intrieri E, Tofani V, *et al.* Landslide detection, monitoring and prediction with remote-sensing techniques. *Nature Reviews Earth & Environment*, 2023, 4(1): 51-64.

通过基于物理信息的深度学习模拟地壳形变

地壳和上地幔的运动和形变为地震过程的演变和未来的地震风险提供了重要信息。地壳形变可以用位错模型来模拟，该模型将地壳中的地震断层表示为连续介质中的缺陷。在本研究中，研究人员提出了一种基于物理的深度学习方法来模拟地震引起的地壳形变。神经网络通过将控制方程和边界条件纳入损失函数，可以表示任意几何结构和岩石力学性质的连续位移场。通过引入极坐标系统作为边界条件，精确地模拟了断层上的位移不连续。研究人员通过对走滑断层的实例分析，说明了这种方法的有效性和实用性。与传统方法相比，这种方法具有潜在优势，因为它可以直接扩展到高维问题、非弹性问题、非线性问题和逆问题。相关研究成果发表在2022年11月的 *Nature Communications*。

使用不同的传感器和仪器（包括全球导航卫星系统）进行的大地测量产生了大量关于地壳形变的数据。对这种观测到的形变进行建模是理解地震过程中所蕴藏机制的基础。地震断层滑动的位错模型常用于同震、震后形变和地震周期的正、反演模拟。在过去的几十年里，线性流变学的解析和半解析方法得到了发展，并且，包括非线性力学特性在内的复杂结构的全数值方法已经建立起来。

由于大量的可用数据，机器学习，特别是深度学习技术不断取得进展。利用积累的地震记录，深度学习技术实现了在地球物理研究中的应用。同时，深度学习也促进了机器学习方法在物理系统中的应用，特别是在求解偏微分方程（PDE）方面。为优化神经网络（NN）而开发的自动微分法在导数的有效计算中发挥着核心作用。其中，物理信息神经网络（PINN）已被提出，用于统一求解偏微分方程的正、逆问题。PINN 表示没有离散化的连续解，并且可以通过将目标偏微分方程和边界/初始条件纳入损失函数

来进行训练，以符合物理定律。由于其简单的实现和对不同类型问题的适用性，PINN 在物理和工程上受到了极大关注。目前，PINN 已被应用于地震层析成像和全波形反演的综合模型。

本研究将 PINN 应用于地壳形变的位错模型。这些模型的一个基本特征是位移场在断层表面上是不连续的，不能直接用表示连续函数的神经网络来近似。为了解决这个困难，研究人员设置了一个合适的坐标系，以分离位移不连续两侧的值。这种方法可以精确地模拟地壳形变，包括近断层位置。研究表明，PINN 可以应用于复杂结构，并且很容易扩展到高维、非弹性和非线性问题，这是其相比于传统方法的一个潜在优势。

1. 物理信息神经网络的优势

具体而言，解析方法使用格林函数（GF）来研究任意滑动分布下的地壳形变，对于均匀半空间、层状半空间和层状球形地球等几种地壳结构，已经得到了弹性静态格林函数的精确表达式。然而，已知的解析解仅限于简单结构。半解析方法被用来模拟复合流变学和准静态地震周期，弯曲断层通常被表示为平面子断层的总和，一般地形通过全空间格林函数的复杂解析来处理。虽然在非线性准静态问题中，已建立了用格林函数表示滞弹性形变的一般表达式，但格林函数只能在齐次半空间中得到。

实际问题完全通过数值方法来解决，特别是有限元方法（FEM），其适用于模拟复杂的构造环境，如俯冲带、山区和非线性流变学。与有限元方法相比，PINN 直接对连续场进行建模，通过对任意数量的点进行训练，有可能达到较高的精度。神经网络可以在不存储所有训练点的情况下，进行小批量迭代训练，这在模拟真实的大规模问题时具有计算优势。PINN 的另一个优点是直接求解无限域问题，而不需要在模型边界上施加边界条

件。而有限元法只能解决有限域问题，并且通常需要建模区域远远大于目标区域，并仔细处理边界条件。

在神经网络在地壳形变正演模拟的开创性应用中，已有研究对神经网络进行了基于半解析方法模拟结果的训练，以对任意位置的解和模型参数（如震源深度和粘性）进行内插。这种方法大大加快了可应用半解析方法的简单问题的形变估计，而 PINN 可以在没有任何现有求解器的情况下处理复杂问题，尽管计算时间较长。

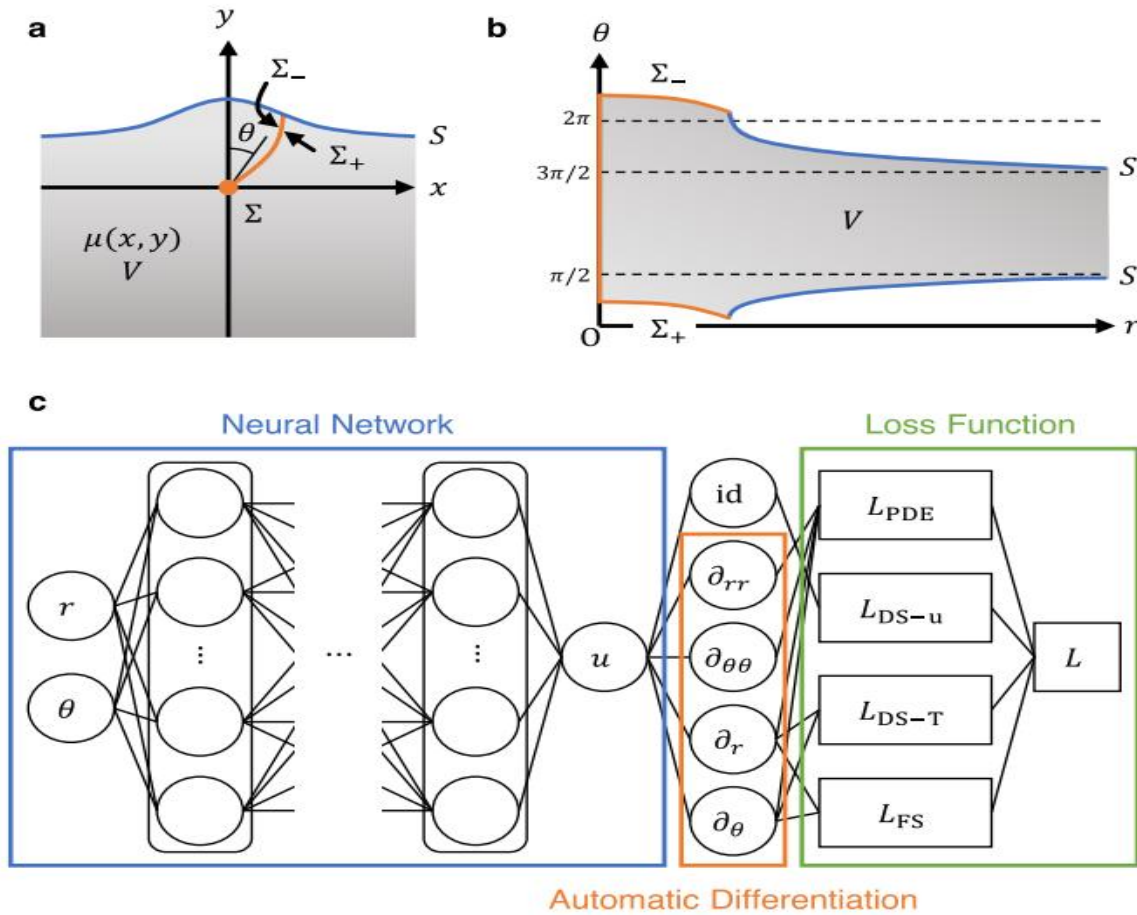


图 1 对反平面位错进行建模的物理信息神经网络

2. 物理信息神经网络建模

本研究以走滑断层为重点，为简单起见，研究人员假定在线弹性介质中有无限长的走滑断层。通过取水平方向的 x 轴和 z 轴，以及垂直方向的 y 轴，假设位移 $u(x, y)$ 平行于 z 方向且在 z 方向不变。研究人员将介质称为 V ，将断层（位错面，DS）称为 Σ ，将地球表面（自由面，FS）称为 S 。因为 $u(x, y)$ 在 Σ 上是不连续的，所以研究人员使用极坐标 (r, θ) 。如此， $u(r, \theta)$ 在整个区域内是连续的， Σ 上的滑移可以表示为 Σ_+ 和 Σ_- 之间的约束。

PINN 建模由三个构建块组成。首先，用神经网络代替连续位移场 $u(r, \theta)$ 。其次，使用自动微分法来评估输出 u 相对于输入变量 (r, θ) 的导数。最后，由控制方程和边界条件的残差平方和定义损失函数 L 。同时，利用随机梯度法对神经网络参数进行更新，以减小损失函数 L 。在本研究中，对于固定网格点，训练被迭代，直到 $L < 10^{-6}$ 。

3. 模拟分析

研究人员首先考虑均匀半空间中的垂直断层，其解析解是已知的。均方根误差（RMSE）分析显示，PINN 解通常是准确的，且 RMSE 比典型的 u 和 ϵ 值小得多。 u 的均方根误差大约是 ϵ 的五倍。这是因为主导的 PDE 约束的不是 u 本身，而是它的空间导数。与均匀滑动（1A 型）相比，分布滑动（1C 型）的形变是局部的，导致断层附近有较高的应变。

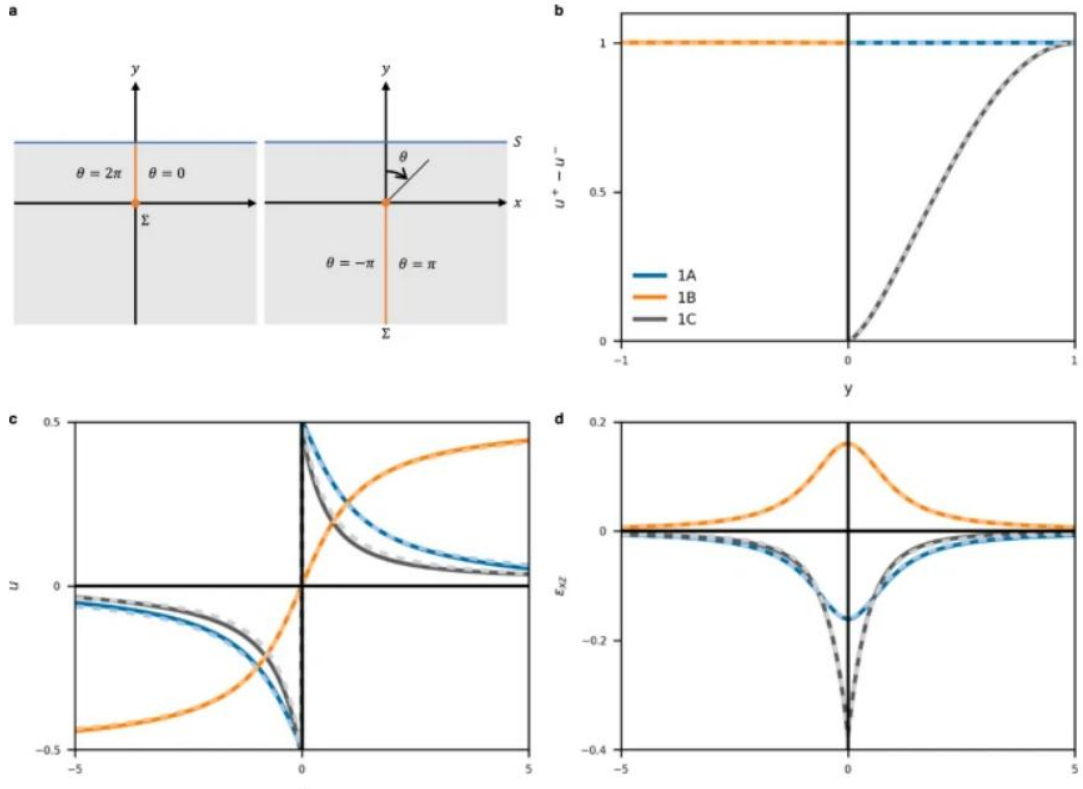


图 2 齐次半空间中的模型和估计结果

对位错模型的垂直截面分析发现，PINN 在模拟复杂地壳结构方面具有优势。具体而言，连续几何和力学性质的变化可以在没有任何近似的情况下表示，例如级数展开和离散化。

地壳是由具有不同力学性质的各种岩石组成的，应变在容易发生地震的物质边界处积累。有些地震断层位于相当深的破裂带内，因此，不连续介质的建模具有重要的实际意义。在本研究中，在单个材料区域（V1 和 V2）中使用两个神经网络来模拟位移场。

研究人员假设埋藏的垂直断层两侧的力学性质变化是不连续的，分析发现，PINN 解在具有断层不连续性的模型中是准确的，但在具有柔性断层的模型中，在距离断层较远的地方显示出系统性的高估。这将是材料边界 B 将 V2 与施加了位移不连续（即狄利克雷边界条件）的断层 Σ 隔

离，这导致了长距离的误差累积。这表明，多个材料的不连续性会使 PINN 的收敛性质复杂化。

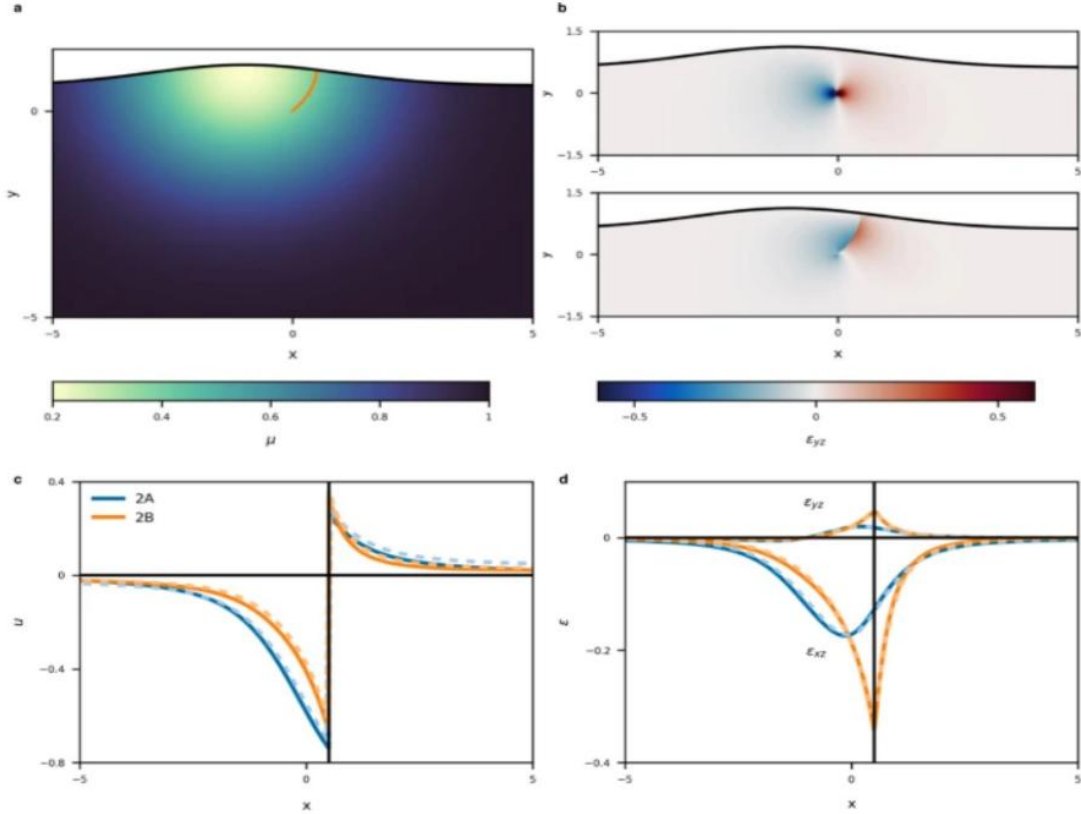


图 3 弯曲、非均质结构的模型和估计结果

对计算成本的分析发现，达到期望精度所需的迭代次数在很大程度上取决于模型结构。几何和力学性质的影响较小，而分布式断层滑动导致多次迭代。材料不连续性不会显著增加迭代次数，但会使每次迭代的计算时间增加约 1.6 倍。

4. 小结

本研究的重点是走滑断层的位错模型，然而 PINN 也可以对倾滑断层和准静态过程进行建模，只是神经网络的输入和输出变量略有增加。与传统的线性求解相比，PINN 的一个主要优点是实现了简单的反演分析。事实上，PINN 目前在逆向建模方面比在正演建模更成功。由于地球物理数

据通常是有噪声的，因此可能需要使用贝叶斯方法来进行稳定的反演。通过改变基于控制方程的损失项，PINN 也可以应用于一般的流变学，如粘弹性、孔隙弹性。区域流变性的差异可以通过定义每个分区域的损失函数来处理，这类似于不连续介质。

该研究以简单的神经网络结构和优化方法，成功地将 PINN 用于反平面问题中的地壳形变建模。然而，现实问题需要更大、更高维的模型空间，各种材料特性和流变性需要越来越多的、与之相对应的神经网络。人们已经认识到，在依赖时间的建模中，PINN 有时会收敛到错误的解。为了优化损失函数以达到足够的精度，这些因素将导致更多的计算成本。稳定、快速的优化方法是将 PINN 应用于大规模地球物理问题的关键。

来源：Okazaki T, Ito T, Hirahara K, *et al.* Physics-informed deep learning approach for modeling crustal deformation. *Nature Communications*, 2022, 13, 7092.

通过不完整极性数据的机器学习改善地震震源机制的确定

地震震源机制通常使用 **P** 波初动极性来确定，并且需要充足的地震记录支持。最新的一种方法较少依赖高信噪比记录，其利用类似波形的相对极性测量来确定地震对 (**earthquake pair**) 之间的关系。利用这些相对极性测量，可以使用区域地震台网生成微震序列中聚簇的复合震源机制。然而，缺失或置信度低的极性测量限制了人们计算高质量复合震源机制的能力。美国地质调查局的研究人员采用迭代随机森林（一种无监督的机器学习方法）进行极性估算，以代替不可靠的极性测量。使用插补后的极性数据，可将地震事件聚类为不同的家族 (**family**)。作为案例研究，研究人员将这种改进后的复合机制工作流程应用于 **2020** 年发生在美国加州北部 **Maacama** 断层附近地震群的多台站模板匹配目录。结果发现，完善后的方法产生了更高质量的地震家族 (**earthquake family**) 和改进的复合震源机制，**94%** 的家族的断层面不确定度小于 **35°**，而使用先前的方法，只有 **34%** 的家族的断层面不确定度小于 **35°**。相关研究成果于 **2023** 年 **1** 月发表在 *Seismological Research Letters*。

聚类是一种无监督机器学习方法，其目标是将数据分为包含相似特征的群组。在地球科学中，具有缺失值的数据集经常出现。在进行聚类之前，必须以某种方式解决这些缺失数据。本研究的目标之一是使用无监督机器学习将缺失数据（“插补”）替换为可靠的估值，以便进行更有意义的聚类分析。

地震震源机制对断层的几何形态和运动学提供了重要限制。利用区域地震台网，通过格林函数与观测数据的比较，可以将矩张量反演用于更大的地震 ($M > 3.5$ — 4.0)。使用本地地震台网，类似方法也可以用于约束较小震级地震的节点平面。对于小震级地震不发生在本地地震台网附近的情

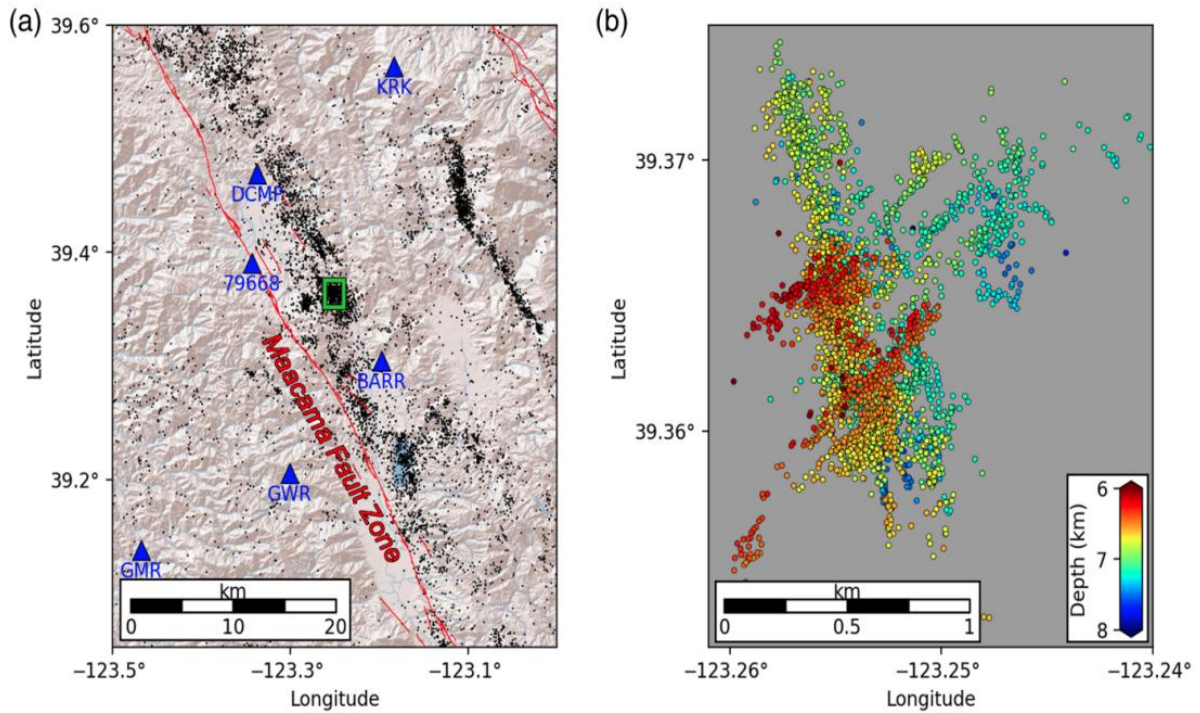


图 1 Maacama 断裂带的震中位置

况，常使用 P 波初动极性获得震源机制。虽然 P 波极性通常是手动识别的，但最近开始使用深度学习方法在大型地震目录中识别初动方向。使用 P 波初动产生受限的断层面解仍需要地震台网（具有足够的方位角覆盖）对地震活动进行很好的记录。因此，利用这些方法确定微地震的可靠震源机制通常需要密集的本地地震台网。

波形交叉相关分析是从相似信号和反向相似信号中提取信息的有力工具，它可以帮助人们对微震序列进行表征。先前研究还表明，利用交叉相关可以产生可靠的微震震源机制。生成震源机制的方法一般包括三个步骤：①协调地震-台站对的 P 波和 S 波的相对极性；②用协调后的相对极性权重将地震分组成家族；③使用 HASH (Hardebeck, 2002) 确定复合机制解。

在本研究中，研究人员重新评估了利用极性测量来产生具有类似震源机制的家族的第二步。具体过程包括使用机器学习方法替换缺失和低置信度的极性测量值，并分类聚簇的相对极性。这些缺失数据可以被插补，因

为初动极性是地震的描绘性特征。如果其中一次地震在某个站点上缺少极性测量，可以使用其它地震的极性来推断缺失的极性。通过事件之间的许多（反）相似性测量，可将此方法扩展到由各种机制及其相应极性组成的目录。

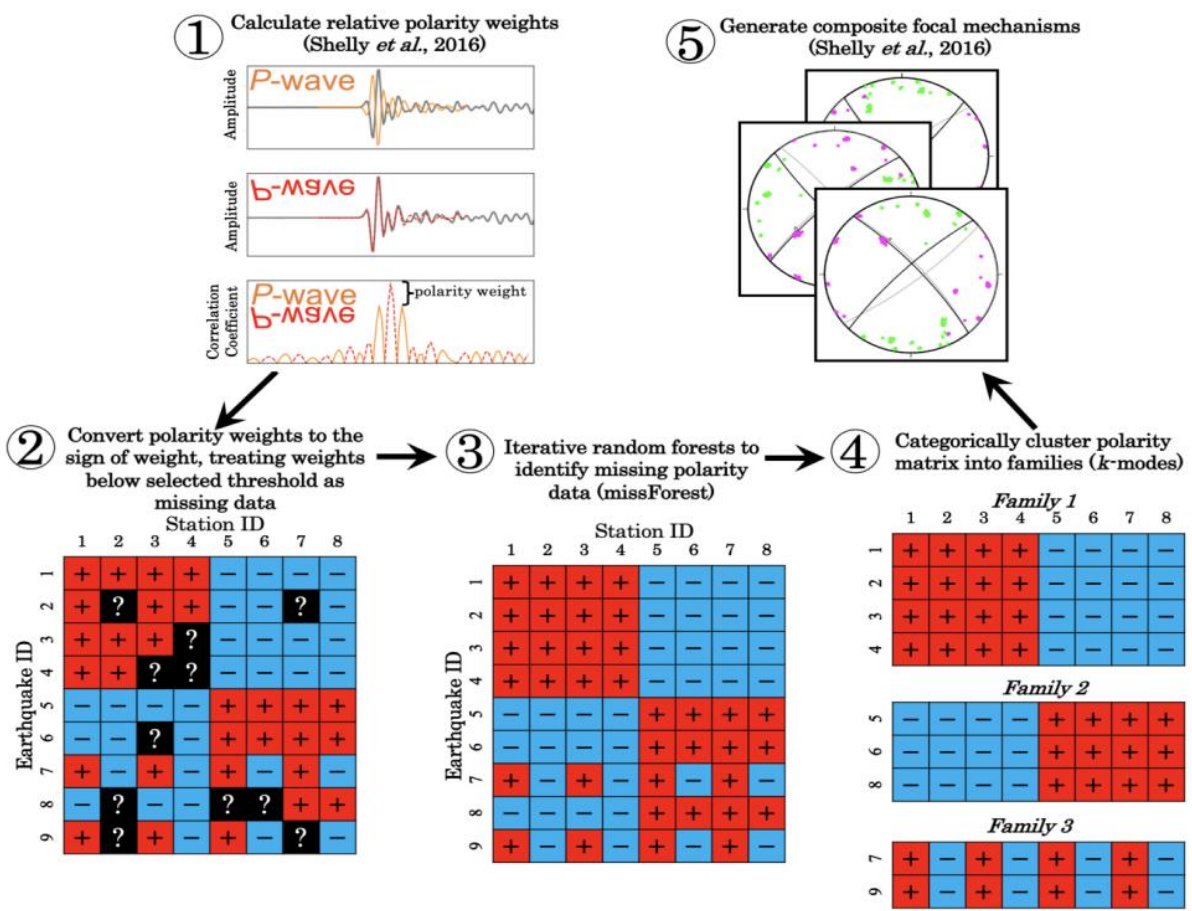


图2 研究所采取工作流的主要步骤

基于 2020 年 Maacama 地震群 3520 个已重新定位地震的相位到达，研究人员能够进行 152245 次协调极性测量。这些测量的绝对极性权重约在 10^{-22} 到 0.5 之间。为了评估数据集中极性权重的质量，并测试可靠地确定缺失极性的能力，研究人员使用 0.00~0.01 之间的各种极性权重阈值，重复了迭代随机森林插补。低于给定阈值的相对极性测量被视为“缺失”数

据，而超过阈值的测量被视为优质数据。研究人员只考虑了通过交叉相关确定的、大于质量协调后的极性权重（quality reconciled polarity weight）一定次数的地震，使用这个地震子目录中的极性，通过省略额外 10%~50% 的极性，来评估每次测试的成功情况。对于给定的极性权重阈值，每省略 10% 的额外极性测量，重复极性测量的准确性就会下降约 0.2%。随着增加相对极性阈值，可靠地重复极性测量的能力得到了改善。极性权重阈值的选择是在可靠地估计缺失极性测量的能力和尽可能考虑更多地震之间的平衡。从这些测试中，研究人员选择了 0.001 的极性权重阈值进行分析，这对应于约 95% 的替换准确性。这样将能够分析 1629 个地震，大约是重新定位目录中事件数量的一半。

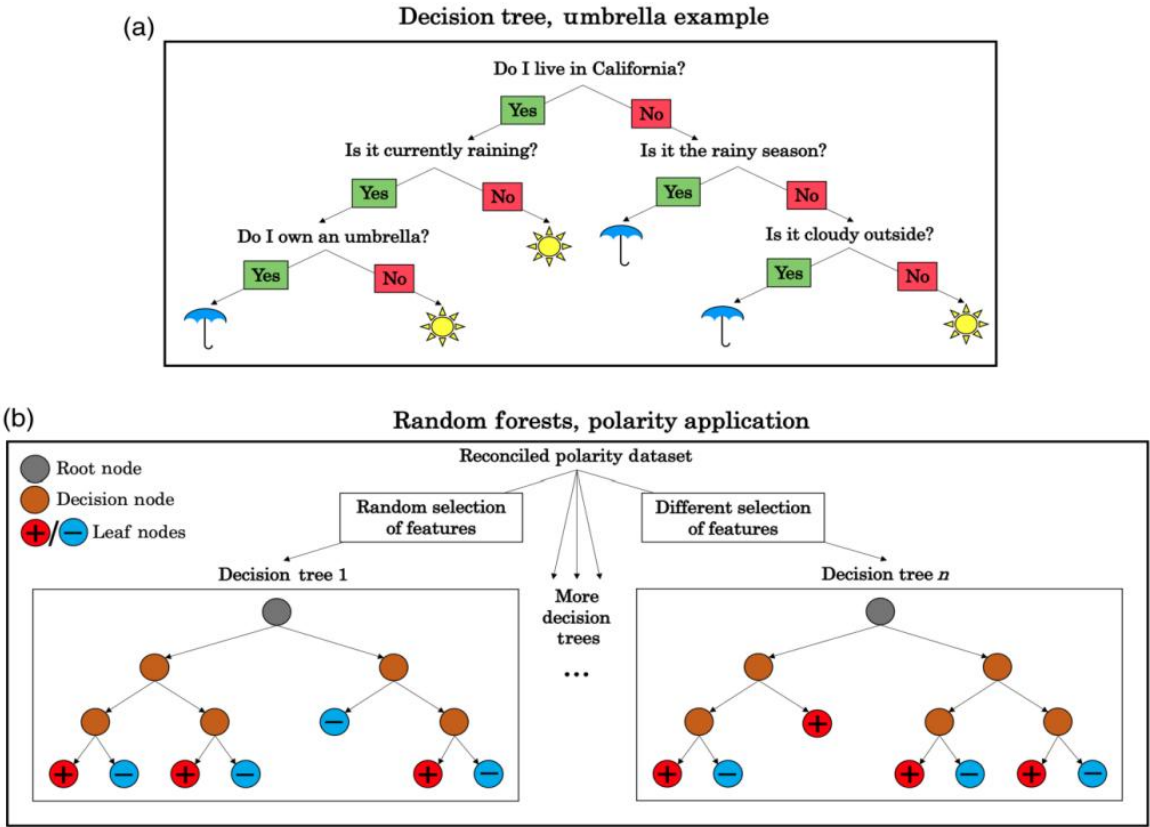


图 3 决策树和随机森林被应用于填补缺失和低置信度极性测量值的过程

在这 1629 次地震中，每次地震都有 ≥ 25 个极性测量值超过了 0.001 的阈值，而研究人员总共分析了 89685 个极性测量值。通过考虑这个地震子目录，并使用随机森林插补，研究人员填充了 135117 个缺失或低置信度的极性测量值，从而生成了一个“完整”的二进制分类数据集，其中包含地震-台站对（earthquake-station pair）的正负值。

为了评估该方法的相对质量，研究人员使用 Shelly 等（2016）的方法重新进行了分析，并将结果分成 50 个家族。然后，计算了目录中地震的 S 波和 P 波的振幅比，即 S/P 比。根据家族-台站对内 S/P 比的一致性，可以估计该研究中极性家族的质量。此外，研究人员还使用 HASH 计算来比较结果。

研究发现，在家族-台站对内，使用新方法计算的 $\log(S/P)$ 比的标准偏差比 Shelly 等（2016）的方法小，复合震源机制的不确定度也更小。使用新方法，94%（47 个家族）的复合震源机制具有小于 35° 的断层面不确定度，而 Shelly 等人（2016）的方法只有 34%（17 个家族）。可以得出结论，修改后的方法产生了更高质量的地震家族，并更好地约束了复合震源机制。

虽然插补可以约束复合震源机制，但这种方法也存在局限性。插补的极性可能不完美，如果数据集中存在不可靠的协调极性测量，插补错误数据的概率会增加。丢弃极性测量并评估被丢弃极性的准确性是一个很好的方法。为了准确地插补缺失的极性数据，目录中应包括每种预期机制类型的多个事件。理想情况下，会有多个地震产生类似的震源机制，且震中位置相近，因为交叉相关极性也取决于记录的波形是否（反）相似。在该方法中，任何未达到每日交叉相关系数的 7 倍 MAD（中位数绝对偏差）阈值的波形对都会被丢弃。如果目录中存在机制与其他事件不同的地震，则

数据插补和聚类可能不会正确进行。因此，这种方法最适用于反映研究区域不同机制的大型地震目录。

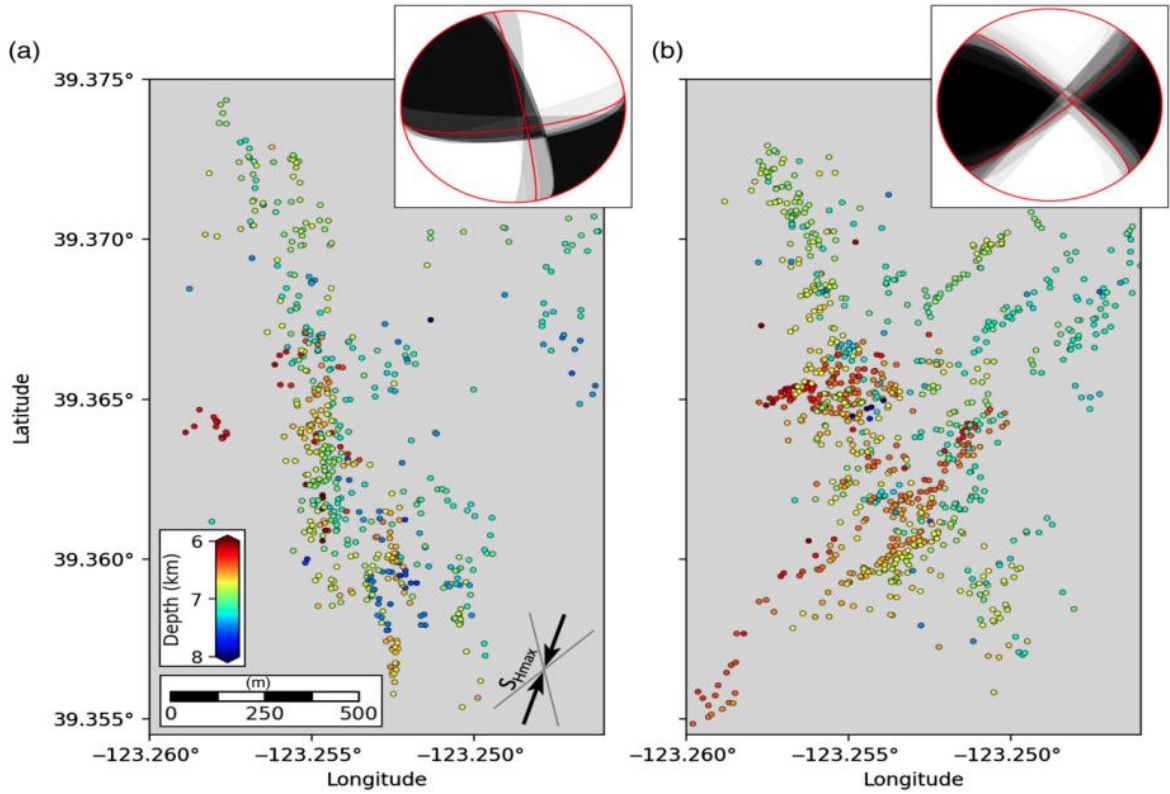


图4 具有复合震源机制的地震家族的地震分布：(a)走向 N11°W/N81°E；(b)走向 N44°W/N49°E

使用基于交叉相关的极性特别有利于信噪比低的事件，因为其提供了一些台站（如不使用深度学习相位拾取器的台站）确定初动极性的机会。输入缺失的极性测量值和使用聚类，可更好地约束震源机制，但更广泛的震源机制反演技术的局限性仍然存在。

通过丢弃在目录中仅具有少量优质极性测量的地震，研究人员将插补和聚类分析限制在最初所考虑地震数量的大约一半。尽管这些被丢弃的事件没有用于生成家族或震源机制，但如果需要，这些事件仍然可以与生成的复合震源机制相关联。之前被丢弃的地震的协调极性可以被评估，以查

看它们是否与一个或多个家族的极性一致。如果之前被丢弃的地震和家族之间的极性一致，那么可以暂时将该地震与该家族相关联。这可能是未来工作的主题，但需要注意的是，成功的关联需要这些事件的机制由确定的复合震源机制中的一种来表示。

来源：Skoumal R J, Shelly D R, Hardebeck J L. Using Machine Learning Techniques with Incomplete Polarity Datasets to Improve Earthquake Focal Mechanism Determination. Seismological Society of America, 2023, 94(1): 294-304.