地震学报
ACTA SEISMOLOGICA SINICA
ISSN 0253-3782 CN 11-2021/P

基于四分量钻孔应变数据的神经网络地震活动性预测分析

于紫凝 李海峰 景锡龙 池成全 郑海永

Borehole strain data based seismicity prediction analysis using a neural network

Yu Zining, Li Haifeng, Jing Xilong, Chi Chengquan, Zheng Haiyong

引用本文:
于紫凝, 李海峰, 景锡龙, 池成全, 郑海永. 2024. 基于四分量钻孔应变数据的神经网络地震活动性预测分析. 地震学报[J], 46(2): 327–339. DOI: 10.11939/jass.20230122
Yu Z N, Li H F, Jing X L, Chi C Q, Zheng H Y. 2024. Borehole strain data based seismicity prediction analysis using a neural network. *Acta Seismologica Sinica*[J], 46(2): 327–339. DOI: 10.11939/jass.20230122

在线阅读 View online: https://doi.org/10.11939/jass.20230122

您可能感兴趣的其他文章

Articles you may be interested in

基于深度卷积神经网络的地震数据随机噪声压制
Random noise suppression of seismic data based on deep convolution neural network
地震学报. 2021, 43(4): 474–482 https://doi.org/10.11939/jass.20200135

利用卷积神经网络检测地震的方法与优化
Earthquake detection using convolutional neural network and its optimization
地震学报. 2020, 42(6): 669–683 https://doi.org/10.11939/jass.20200045

基于反向传播神经网络的闽台M_L震级偏差分析与修正
Analysis and revision of magnitude M_L deviation between Fujian and Taiwan based on BP neural network
地震学报. 2019, 41(6): 723–734 https://doi.org/10.11939/jass.20190038

基于样本增强的卷积神经网络震相拾取方法
Seismic phase identification using the convolutional neural networks based on sample enhancement
地震学报. 2020, 42(2): 163–176 https://doi.org/10.11939/jass.20190070

佘山台四分量钻孔应变仪记录的台风扰动特征分析
Signatures of typhoons on strain records of four-component borehole strain meter at Sheshan station
地震学报. 2017, 39(5): 725–737 https://doi.org/10.11939/jass.2017.05.008

基于短时傅里叶变换和卷积神经网络的地震事件分类
Classification of seismic events based on short-time Fourier transform and convolutional neural network
地震学报. 2021, 43(4): 463–473 https://doi.org/10.11939/jass.20200128

第 46 卷　第 2 期　　地　震　学　报　　Vol. 46, No. 2
2024 年 3 月　（327−339）　　ACTA SEISMOLOGICA SINICA　　Mar., 2024

于紫凝，李海峰，景锡龙，池成全，郑海永. 2024. 基于四分量钻孔应变数据的神经网络地震活动性预测分析. 地震学报，**46**(2)：327−339. doi：10.11939/jass.20230122.
Yu Z N, Li H F, Jing X L, Chi C Q, Zheng H Y. 2024. Borehole strain data based seismicity prediction analysis using a neural network. *Acta Seismologica Sinica*, **46**(2): 327−339. doi: 10.11939/jass.20230122.

基于四分量钻孔应变数据的神经网络地震活动性预测分析*

于紫凝[1)]　李海峰[1)]　景锡龙[1)]　池成全[2)]　郑海永[1)]

1）中国山东青岛 266404 中国海洋大学电子工程学院
2）中国海口 571158 海南师范大学信息科学技术学院

摘要　首先利用四分量钻孔应变数据独有的自洽特性，构建震前应变特征数据集；之后基于一维卷积神经网络框架，设计地震震级与方位的预测模型；然后通过混淆矩阵计算准确率、召回率以及 F_1 分数，对模型预测结果进行评价与修正；最后对我国西南地区的永胜、昭通、姑咱及腾冲四个台站的钻孔应变特征分别进行训练与验证，并讨论了不同特征窗长对预测效果的影响. 训练完成后的模型效果在测试集上均表现优异，四个台站对震级和方位预测的平均准确率分别可达 85% 和 80% 左右，说明四分量钻孔应变数据特征与地震的发生有着很强的相关性，通过卷积神经网络对地震前兆特征进行挖掘具有很大研究潜力，本文提出的预测策略也为未来短临地震的精确预测研究打下基础.

关键词　四分量钻孔应变　卷积神经网络　震级预测　方位预测　地震前兆
doi：10.11939/jass.20230122　　**中图分类号**：P315.72^{+}7　　**文献标识码**：A

Borehole strain data based seismicity prediction analysis using a neural network

Yu Zining[1)]　Li Haifeng[1)]　Jing Xilong[1)]　Chi Chengquan[2)]　Zheng Haiyong[1)]

1) *College of Electronic Engineering*, *Ocean University of China*, *Shandong Qingdao* 266404, *China*
2) *School of Information Science and Technology*, *Hainan Normal University*, *Haikou* 571158, *China*

Abstract: With the development of seismological observational techniques, a number of case studies indicate that the seismogenic process of major earthquakes is often accompanied by deformation anomalies. Strain data serve as an indicator of crustal deformation, which reflects changes in subsurface stress and holds significant importance for seismology research. However, the research on the extraction of pre-earthquake anomalies from borehole strain data is currently limited to the stages of case analysis and small-sample statistical analysis. Thus, it is very meaningful to use a new technique of data mining to analyze the association between

* **基金项目**　国家自然科学基金青年基金（42204005）、山东省自然科学基金青年基金（ZR2022QF130）、中央高校基本科研业务费专项（202213042）和海南省自然科学基金高层次人才项目（622RC669）共同资助.
收稿日期　2023-09-30 收到初稿，2024-03-14 决定采用修改稿.
作者简介　于紫凝，博士，讲师，主要从事地震前兆分析、震前异常提取与地震预测方面的研究，e-mail：yuzining@ouc.edu.cn

strain anomalies and earthquakes.

In order to dig out more strain information and correlation information between multiple strains, according to the scholars' many analyses of the correlation between areal strains, shear strains and the self-consistent coefficients of the four-component strains, it is found that the borehole strain may reflect the preparatory process of earthquake nucleation. Therefore, this study calculates the Pearson correlation coefficients and self-consistent coefficients between these strains to finally constitute a 24-dimensional feature dataset. Subsequently, we divide earthquake events based on magnitude into three classes: no earthquake, earthquakes with $3.0 \leqslant M_S < 5.0$, and earthquakes with $M_S \geqslant 5.0$. Simultaneously, these earthquakes are also classified into five groups based on their orientation relative to the borehole strainmeter, forming an earthquake orientation dataset. Thereby, the labels for earthquake samples are generated according to these two classification criteria.

Next, the study employs a one-dimensional convolutional neural network (1D-CNN) framework to develop a short-term prediction model for the magnitude and location of earthquakes. The CNN can leverage the advantages of its convolutional layers' parameter-sharing mechanism to effectively capture local features in the data. This 1D-CNN model consists of three parts: the input layer, hidden layers, and output layer. Strain feature samples from the dataset are used as inputs to the model, which outputs predicted values for earthquake magnitude and location labels. The hidden layer of the model is divided into two parts: the convolutional region and the fully connected region. We apply the cross-entropy loss function and the Adam optimizer for model compilation, and a learning rate decay strategy is used to dynamically adjust the learning rate. Parameter optimization is conducted using a random search algorithm. To evaluate this prediction model, we use a confusion matrix to calculate accuracy, recall, and F_1 scores for examining the efficiencies for each class. Furthermore, the study adjusts the model's prediction window size for earthquakes above magnitude 5.0 to balance the distribution of samples across different magnitude classes.

Finally, the borehole strain data from stations Guzan, Yongsheng, Zhaotong, and Tengchong in southwest China are respectively utilized for training and testing in this study. The results indicate that the pre-earthquake strain features from stations Yongsheng, Guzan, Zhaotong, and Tengchong exhibited a high level of accuracy up to 80% in predicting both magnitude and direction. Furthermore, the predictive results are independently validated for five typical major earthquakes in China. The findings demonstrate that the predicted magnitude and direction labels for these five earthquakes corresponded to the actual events, suggesting that the model successfully captures valuable pre-earthquake strain information and possesses predictive capabilities. Finally, we analyze the impact of the time window and prediction window sizes of input data on the model's prediction accuracy across different magnitude classes. The results reveal that for all three magnitude classes, a longer time window leads to higher predictive accuracy of the model. Moreover, the results of major earthquakes are overall higher and more random than that of moderate earthquakes. It may reflect that the borehole strain has the short-term predictive capability for major earthquakes, and there are differences in pre-earthquake features among different major earthquakes.

The 1D-CNN models built in this study could effectively predict earthquake magnitude and approximate location using data from the respective stations. This demonstrates that the convolutional neural network architecture has the capability to extract meaningful pre-earthquake

anomaly features from borehole strain data. This provides new insights and methods for research in earthquake precursory observations, laying a foundation for accurate earthquake predictions in the future. However, considering practical applicability, there are limitations to the methodology of this study. Future research will devise more robust networks to address sample imbalances by combining labels and achieve simultaneous predictions for the three seismic elements. Alternatively, a combination of multiple strain stations could be utilized, incorporating additional earthquake location information (such as epicentral distance and angles) to enhance directional prediction accuracy through partitioning high-resolution spatial grids. Moreover, experimental validations will be conducted on strain observation stations in seismically weak areas or other multi-seismic regions to establish a robust short-term earthquake prediction model suitable for strong seismic events.

Key words: four-component borehole strain; convolutional neural network; magnitude predictions; orientation predictions; earthquake precursor

引言

我国属于地震频发国家，如果能够有效预测大地震的时间、位置和震级，便可及时采取应对措施，保护人民的生命财产安全，降低房屋、桥梁和道路等基础设施的破坏程度和经济损失，从而维护社会稳定和发展(陈运泰，2008)．然而，由于地震发生概率低，孕震机制复杂，地球深部观测技术受限，地震预测仍待长期持续的研究(陈运泰，2007)．

随着观测技术的不断进步，大量震例研究表明，大地震的孕育过程通常伴有地表异常，如 2008 年汶川 M_S8.0 地震(江在森等，2009；Zhu *et al*，2019)、2011 年日本东北 M_W9.0 地震(Hirose，2011)、2017 年九寨沟 M_S7.0 地震(高曙德，2020)等．GPS、应力和应变观测资料作为地壳形变指标，在多次大地震前显示出断层闭锁或变幅异常(张希等，2019；顾国华，王武星，2020；牛安福等，2021；Yu *et al*，2021a，b)．近年来，国内外学者研究发现，地震孕育过程中，岩石破裂还会产生压电效应、动电效应及压磁效应等一系列地震电磁效应(De Santis *et al*，2017；Parrot，2017；Pulinets *et al*，2018)．以芦山地震为例，姑咱台钻孔应变观测在主震前半年就记录到脉冲异常(池顺良等，2013)，除此之外，附近台站还观测到视电阻率、地电场异常(高曙德，2016)、地磁场异常(廖晓峰等，2018)等．由于电磁波的趋肤深度与震源深度相当，低频电磁辐射也被认为是非常有力的地震前兆(Ouyang *et al*，2018)．

应变数据可反映地下应力变化，对于地震研究具有重要意义．在青海玉树、日本东北和四川芦山三次大地震中，青海地区的 6 个钻孔应变台站均记录到了与震级和震源方位相符的应变阶跃(张敏等，2014)；针对汶川地震，邱泽华等(2009，2012)在剔除钻孔应变数据的周期信号后，采用超限率分析的方法提取了高频信号中的幅度异常，发现异常在震前开始增加、震后逐步减少，认为可能与汶川地震有关；对于芦山地震，邱泽华等(2015)通过分析姑咱台在地震数天前记录到的钻孔应变波形数据，排除了远震干扰和实地施工等人为干扰，认为剩余的幅度异常变化与芦山地震呈因果关系；池顺良等(2013)同样通过观察姑咱台的钻孔应变波形数据的幅度变化，以及计算观察到的应变异常方向与地震应力的主方向一致，证明了芦山地震前几天和震前 4 至 6 个月记录的异常现象可能反映了地震发震前地层的蠕滑；刘琦等(2014)利用 S 变换方法也对芦山地震前后的钻孔应变数据进行了时频分析，发现两簇高能量异常表现出与芦山地震较好的相关性；Chi 等(2019)采用变分模态分解方法对钻孔应变

观测数据进行分解，通过状态空间模型，采用频率对比方式去掉气压和钻孔水位的影响后，发现了芦山地震的震前异常；Zhu 等(2019)提出了钻孔应变数据的震前负熵异常提取方法，汶川和芦山地震的负熵结果分析显示，各台站负熵异常的累计值在震前均出现以幂律形式的加速增长；之后，Yu 等(2020)又采用多通道奇异谱分析联合分解了多台站钻孔应变数据，构造了多台站的空间拓扑网络，联合提取了川滇地区 13 次地震前的应变异常．这些研究都表明，大地震前钻孔应变数据能够反映地震孕育过程．

然而，对于钻孔应变数据的震前异常提取研究，目前还停留在震例分析与小样本统计分析阶段．Yu 等(2020)分析了 2010—2017 年多台站网络度的异常，也只是研究了 29 次能量较大的地震，对于 M3.0—5.0 中等地震的震前异常目前尚未涉及．另外，大多数研究只关注了应变数据异常与地震时间上的关联，对异常的空间分布以及震源方位的分析研究甚少．

本文拟依据大量钻孔应变观测数据，发挥数据驱动方法的优势，采用卷积神经网络框架，设计地震震级与方位的短临预测模型；根据四分量钻孔应变仪原理，计算分量应变、面应变、剪应变间的相关系数和自洽系数，构造应变特征数据集以作为模型的输入；模型输出分别为震级以及方位，按震级大小将震级划分为三类，根据钻孔应变仪方位角将方位划分为五类；通过混淆矩阵计算准确率、召回率以及 F_1 分数评价与修正模型．最后，通过对我国西南地区的姑咱、永胜、昭通和腾冲台的钻孔应变数据进行训练与测试，探索应变数据与地震发生大小、方位的内在联系．

1 数据简介

1.1 四分量钻孔应变观测原理

四分量钻孔应变仪一般放置于地下数十米断裂带周围的基岩中，通过内部传感器的双环模型测量孔径变化量．孔径变化量与原孔径的比值为 S_θ，表示沿孔径变化方向 θ 即传感器方向直接观测到的孔径变化．根据该模型，在相对远处有均匀水平应变 ε_1 和 ε_2 时，有(李进武，邱泽华，2014)

$$S_\theta=A(\varepsilon_1-\varepsilon_2)+B(\varepsilon_1+\varepsilon_2)\cos[2(\theta-\phi)], \tag{1}$$

式中：ϕ 为 1 分量的方位角；A，B 为耦合系数，其大小受周围水泥和套筒的尺寸及材质的共同影响．

四分量钻孔应变仪的四个传感器按相邻夹角 45°布置，根据式(1)，有四分量观测值：

$$\begin{cases} S_1=S_\theta=A(\varepsilon_1-\varepsilon_2)+B(\varepsilon_1+\varepsilon_2)\cos[2(\theta-\phi)], \\ S_2=S_{\theta+\pi/4}=A(\varepsilon_1-\varepsilon_2)-B(\varepsilon_1+\varepsilon_2)\sin[2(\theta-\phi)], \\ S_3=S_{\theta+\pi/2}=A(\varepsilon_1-\varepsilon_2)-B(\varepsilon_1+\varepsilon_2)\cos[2(\theta-\phi)], \\ S_4=S_{\theta+3\pi/4}=A(\varepsilon_1-\varepsilon_2)+B(\varepsilon_1+\varepsilon_2)\sin[2(\theta-\phi)], \end{cases} \tag{2}$$

在式(2)的基础上可作变换，令 $S_{13}=S_1-S_3$ 和 $S_{24}=S_2-S_4$，S_{13} 和 S_{24} 称为剪应变．

四分量钻孔应变观测还有一个重要的特性就是观测自洽，它可以判断元件的性能，也可评价观测数据的准确性(于紫凝，2022)．自洽方程为

$$S_1+S_3=k(S_2+S_4), \tag{3}$$

式中，k 为自洽系数，若 $k\geqslant 0.95$，即认为数据满足自洽．实践证明，若应变仪探头布设于完整岩石中时，其观测曲线应大体上符合观测自洽．自洽系数 k 也常作为震前数据异常的重要

依据，在汶川、鲁甸和康定地震前，姑咱台和昭通台均观测到了失恰现象（池顺良等，2014）.

1.2　钻孔应变特征数据集设计

四分量钻孔应变仪每隔 45° 布置一路应变传感器，四路应变分别为 S_1，S_2，S_3 和 S_4，采样间隔为 1 min. 为了挖掘到更多的应变信息以及多路应变之间的关联信息，本文又构造剪应变 S_1-S_3，S_2-S_4 及自洽方程两端面应变 S_1+S_3 和 S_2+S_4，共得到 8 维应变数据. 池顺良等（2014）多次对面应变、剪应变之间的相关性以及四路应变的自洽系数进行分析，发现汶川、鲁甸、康定地震前，钻孔应变可能反映了地震成核的过程. 所以，本文将这些应变间的皮尔逊相关系数与自洽系数作为特征数据集，共 24 维，其物理含义如表 1 所示.

表 1　24 维应变数据特征

Table 1　Features of 24-dimensional strain data

序号	特征名称	物理含义
1	corr01_02	S_1与S_2间的皮尔逊相关系数
2	corr01_03	S_1与S_3间的皮尔逊相关系数
3	corr01_04	S_1与S_4间的皮尔逊相关系数
4	corr02_03	S_2与S_3间的皮尔逊相关系数
5	corr02_04	S_2与S_4间的皮尔逊相关系数
6	corr03_04	S_3与S_4间的皮尔逊相关系数
7—10	corr01_13—corr04_13	S_1-S_4与剪应变S_1-S_3间的皮尔逊相关系数
11—15	corr01_24—corr04_24	S_1-S_4与剪应变S_2-S_4间的皮尔逊相关系数
16—19	corr01+13—corr04+13	S_1-S_4与面应变S_1+S_3间的皮尔逊相关系数
20—23	corr01+24—corr04+24	S_1-S_4与面应变S_2+S_4间的皮尔逊相关系数
24	corr13+24	自洽系数

注：S_1，S_2，S_3和S_4为钻孔应变四分量观测值.

2 基于钻孔应变的一维卷积神经网络地震活动性预测模型

2.1　一维卷积神经网络模型原理

为防止模型参数过多导致训练过拟合，本文采用一维卷积神经网络（convolutional neural network，缩写为 CNN），提取震前应变数据特征训练模型，实现地震震级及方位的预测. 卷积神经网络利用其卷积层参数共享机制的优势，可以有效捕捉到数据中的局部特征（卢宏涛，张秦川，2016）. 如图 1 所示，模型由输入层、隐藏层、输出层三部分组成. 模型输入为单输入，输入层 24 个神经元节点 $x_1, x_2, \cdots, x_{24}$ 分别对应 24 种特征. 模型输出为多输出多分类：多输出是指模型输出具有多个预测值 $f(x_1, x_2, \cdots, x_{24})=(y_1, y_2)$，输出层 1 和输出层 2 分别对应震级标签 y_1 和方位标签 y_2 两个输出预测值；多分类是指每个预测值可以取多个类别，$y_1 \in \{y_{11}, y_{12}, y_{13}\}$ 分别对应无震、$3.0 \leqslant M_S < 5.0$ 及 $M_S \geqslant 5.0$ 三个类别，$y_2 \in \{y_{21}, y_{22}, y_{23}, y_{24}, y_{25}\}$ 分别对应无地震及划分地震方位的四个方向，共五

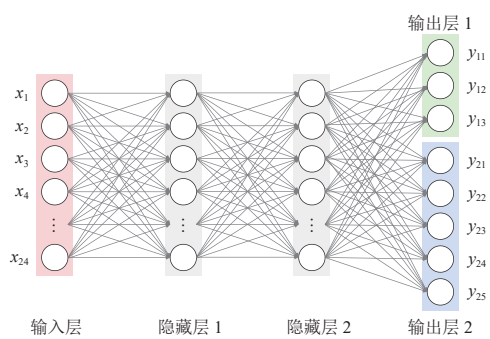

图 1　地震预测模型层级示意图

Fig. 1　Schematic diagram of structure earthquake prediction model

个类别．隐藏层 1 和隐藏层 2 分别为卷积区和全连接区．

卷积区由多个卷积层、批量标准化层、最大池化层组成，其结构如图 2 所示．输入为 335×24 的应变特征矩阵，第一个卷积层使用 64 个卷积核提取出 329×64 的特征图．特征图在批量正则化层进行归一化处理，并输入最大池化层进行降采样，得到输出为 164×64 的特征图，并将其作为第二个卷积层输入，进行第二轮卷积、归一化、池化操作．两个卷积层采用不同大小的卷积核，具有参数共享和局部连接的特性，可以减少模型的参数量，增强模型的表征能力（周飞燕等，2017）．在两轮卷积、归一化、池化操作后得到 51×64 的特征图，并将其输入全连接区．

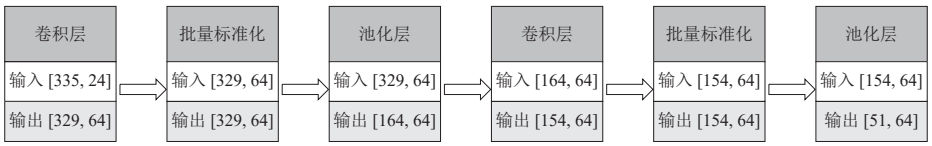

图 2 卷积层架构图

Fig. 2 The architecture diagram of convolutional layer

全连接区由全局平均池化层、全连接层和丢弃层组成（图 3）．卷积区输入 51×64 的特征图，通过全局平均池化层对每个通道特征图进行平均运算，得到该通道特征图的平均强度，64 个通道特征图取得了长度为 64 的特征向量并输入全连接层．两个全连接层分别采用 32 个和 16 个神经元，用于提取特征并进行分类．每个全连接层后添加丢弃层，防止模型过拟合．

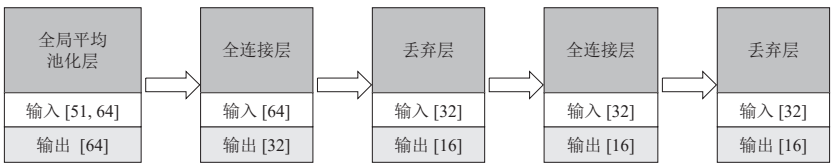

图 3 全连接层架构图

Fig. 3 The architecture diagram of fully connection layer

输出层使用 Softmax 激活函数，可以将输出层 1 和输出层 2 的特征向量映射为(0, 1)区间内的预测值，表示震级和方位的类别概率分布，并将概率最大类别作为震级和方位类别的预测结果．

整体网络设置中，ReLU 激活函数应用于每个卷积层和全连接层，并在每个卷积层和全连接层采用 L2 正则化方法．L2 正则化通过在损失函数中添加与参数平方和成比例的惩罚项来约束模型复杂度，在卷积层采用$\lambda=0.01$的权重，在全连接层采用$\lambda=0.001$的权重．最小化损失函数和 L2 正则化项的加权和可以降低模型的复杂性，以避免过拟合．计算方法为

$$J(w)=\min_{w}\left[L(w)+\lambda\|w\|_2^2\right], \tag{4}$$

式中，J 为加入正则化项的网络代价函数，w 为神经网络权重参数，L 为网络的损失函数．

2.2 模型编译与参数优化

模型编译使用多元交叉熵损失函数和 Adam 优化器，并采用学习率衰减策略动态调整学习率．学习率衰减策略根据验证集损失值判断模型训练状态，当验证集损失值在一定轮次

(10 轮次)内未下降时，学习率则乘以衰减因子(0.1)以使学习率缩小为原来的十分之一，直至学习率达到最小值(1×10^{-8})，从而避免由学习率过小或过大导致的收敛缓慢、梯度爆炸等问题，并可提高模型的收敛速度和准确性.

为了优化模型超参数并提高模型性能，本文使用随机搜索算法寻找最优超参数，构造包含卷积核大小及数量、池化窗口、全连接层神经元个数、L2 正则化系数、丢弃层丢弃率、样本批次等参数的参数空间. 在参数空间随机选择一组超参数组合，计算此时模型的验证集准确率，并向着模型最优的方向不断动态调整超参数. 依次持续迭代，直到模型的性能达到稳定且最优水平为止. 选择该组超参数作为模型超参数.

$$x^{*}=\underset{x\in X}{\operatorname{argmax}} f(x)\,, \tag{5}$$

式中，x 为超参数集合，X 为超参数空间，$f(x)$ 为模型验证集的性能指标函数，x^{*} 为最优的超参数向量. 本文构建的 1D_CNN 模型共两层，超参数数量适中，随机搜索过程基本可以覆盖全部参数空间.

2.3　模型预测结果评价指标

混淆矩阵是用于评估分类模型性能的一种矩阵(表 2). 根据模型的预测结果与真实标签的一致性进行统计得到真正例 TP、假正例 FP、真反例 TN 和假反例 FN 四个指标. 在混淆矩阵的基础上可以得到准确率 P_{acc}、精确率 P_{pcs}、召回率 P_{rec} 等常用的模型评估指标. 各指标的意义及计算方法为：准确率，预测正确样本数占总样本数的比例，

表 2　混淆矩阵

Table 2　Confusion matrix

	实际正例	实际负例
预测正例	真正例(TP)	假正例(FP)
预测负例	假反例(FN)	真反例(TN)

$$P_{\mathrm{acc}}=\frac{TP+TN}{TP+TN+FP+FN}; \tag{6}$$

精确率，预测为正例中正确预测的样本比例，

$$P_{\mathrm{pcs}}=\frac{TP}{TP+FP}; \tag{7}$$

召回率，正例中被正确预测的样本比例，

$$P_{\mathrm{rec}}=\frac{TP}{TP+FN}; \tag{8}$$

F_1 分数(F_1-score)，精确率和召回率的调和平均数，综合评价分类器的性能，

$$F_1=\frac{2P_{\mathrm{pcs}}P_{\mathrm{rec}}}{P_{\mathrm{pcs}}+P_{\mathrm{rec}}}. \tag{9}$$

3 基于钻孔应变数据的地震活动性预测结果及分析

3.1　地震震级与方位数据集的构造

本文选择我国西部永胜、昭通、姑咱和腾冲四个台站 2010—2017 年的应变数据构造震前应变特征数据集. 根据震级大小将地震分为无震、$3.0\leqslant M_S<5.0$ 及 $M_S\geqslant5.0$ 三类，再根据震源相对于钻孔应变仪的方向，将方位分为五类，制作样本标签. 以永胜台为例，台站及地震筛选分布以及震源预测的方位示意图如图 4所示. 震源方位是以钻孔应变仪 1 分量为基准，顺

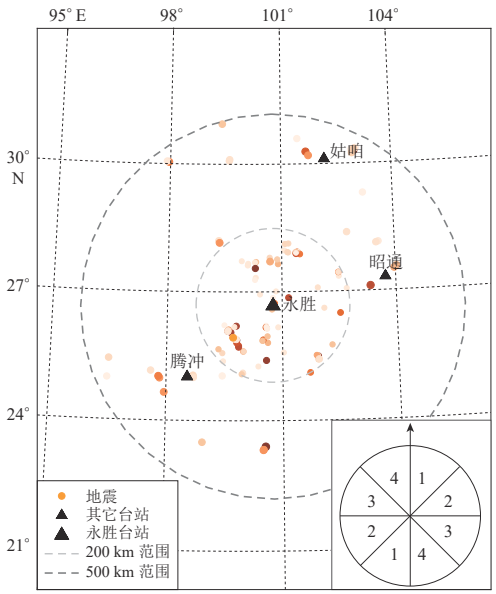

图 4 钻孔应变台站位置和地震分布

右下角为震源方位划分示意图

Fig. 4 Location of borehole strain stations and earthquakes distribution map

The inset at the lower right is the schematic diagram of epicenter orientation delineation

时针每隔 45°划分为一类，与无震时的样本一起将震源方位划分为五类.

以四个台站为中心，筛选地震样本基本情况(表 3)，包括样本数量、最大震级、最小距离、最大距离等. 可以看到 $M5.0$ 以上地震较少，大约是 $M3.0$—5.0 地震的 1/7，而大地震的影响范围大，孕震时间又长. 为了增加大地震的样本数量，对于 $M5.0$ 以上地震的选择，本文首先将空间半径扩大到 500 km. $M3.0$—5.0 地震的选择范围则是距离中心台站 200 km 以内. 其次，对于 $M5.0$ 以上地震，本文使用窗长 7 天、步长 1 天的滑动时间窗口来预测之后 7 天的地震大小和方位. 也就是说，大地震作为预测窗口的第一天和第七天，可构造出 7 个样本. 这样既增加了样本的数量，同时也可保证样本的真实性和有效性. 另外，由于无震时的样本数目众多，本文从中均随机抽取 300 个作为该类别的应变特征数据集.

3.2 模型训练与预测结果分析

本文以震前一周的钻孔应变特征预测之后一周内的地震震级和方位. 以永胜台为例，图 5 给出了震级预测和方位预测的模型损失率曲线. 模型训练集和验证集的损失率曲线均在 150 轮训练后平稳收敛. 昭通、腾冲和姑咱三个台站的模型训练结果与永胜台均一致，四个台站验证集损失率则稳定在 0.1—0.3 之间.

表 3 永胜、昭通、姑咱和腾冲四个台站筛选的地震基本情况

Table 3 Basic information about the earthquake samples from the stations Yongsheng, Zhaotong, Guzan and Tengchong

台站	M_S3.0—5.0地震数量	M_S≥5.0地震数量	最大震级M_S	最小距离/km	最大距离/km
永胜	99	47	7.0	5.96	491.86
昭通	138	31	7.0	31.09	497.47
腾冲	172	52	7.6	6.44	498.75
姑咱	287	36	7.0	13.89	431.65

模型训练完成后，本文在测试数据集上分析模型的预测效果. 永胜、昭通、姑咱和腾冲四个台站的数据集测试结果列于表 4 和表 5. 总的来说，在震级和方位的预测上，模型均表现出了较高的准确率. 尤其对于 0 类标签预测，即非地震标签，有着非常高的准确率和召回率，F_1 值均接近或达到 0.85，这说明正常情况下钻孔应变数据特征是比较容易捕捉的. 实际上，钻孔应变仪在没有外界干扰时，其波形数据中记录着的光滑且有周期性的固体潮是数据中最显著的特征(邱泽华等，2015).

对永胜台来说，在震级预测上，模型对 1 类样本($3.0 \leqslant M_S < 5.0$)的识别有较高的准确率、

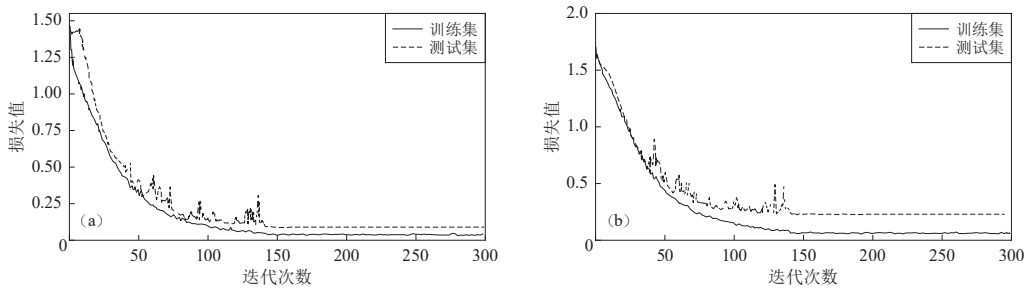

图 5　永胜台预测模型的损失率曲线

(a) 震级预测损失曲线；(b) 方位预测损失曲线

Fig. 5　Loss curves of the prediction model at Yongsheng station

(a) Loss curves for the magnitude prediction；(b) Loss curves for the orientation prediction

表 4　永胜、昭通、姑咱和腾冲台震级预测结果表

Table 4　Magnitude prediction results for the stations Yongsheng, Zhaotong, Guzan and Tengchong

	标签	精确率	召回率	准确率	F_1值	样本数		标签	精确率	召回率	准确率	F_1值	样本数
永胜台	0	0.903	0.823		0.861	147	昭通台	0	0.813	0.875		0.843	144
	1	0.829	0.907		0.866	107		1	0.836	0.825		0.830	154
	2	0.720	0.766		0.742	47		2	0.810	0.567		0.677	30
				0.844		301(总)					0.823		328(总)
腾冲台	0	0.837	0.860		0.848	143	姑咱台	0	0.788	0.748		0.768	139
	1	0.888	0.864		0.876	184		1	0.868	0.909		0.888	318
	2	0.880	0.772		0.822	57		2	0.840	0.636		0.724	33
				0.852		384(总)					0.845		490(总)

表 5　永胜、昭通、姑咱和腾冲台震源方位预测结果表

Table 5　Orientation prediction results for the stations Yongsheng, Zhaotong, Guzan and Tengchong

	标签	精确率	召回率	准确率	F_1值	样本数		标签	精确率	召回率	准确率	F_1值	样本数
永胜台	0	0.879	0.837		0.857	147	昭通台	0	0.801	0.868		0.833	144
	1	0.857	0.686		0.762	35		1	0.879	0.886		0.883	140
	2	0.754	0.897		0.819	58		2	1.000	0.467		0.636	15
	3	0.781	0.862		0.820	58		3	0.882	0.682		0.769	22
	4	0.000	0.000		0.000	3		4	0.571	0.571		0.571	7
				0.827		301(总)					0.838		328(总)
腾冲台	0	0.810	0.867		0.838	143	姑咱台	0	0.728	0.770		0.748	139
	1	0.750	0.632		0.686	19		1	0.840	0.750		0.792	28
	2	0.743	0.728		0.735	103		2	0.593	0.574		0.583	61
	3	0.824	0.792		0.808	106		3	0.775	0.811		0.793	106
	4	0.833	0.769		0.800	13		4	0.818	0.776		0.796	156
				0.794		384(总)					0.755		490(总)

召回率和 F_1 值，均可达到 0.82 以上．2 类样本($M_S \geqslant 5.0$)的 F_1 值达到 0.74 以上，结果也处于相对高的水平，但较低于 1 类样本．可能原因有二：一是根据古登堡—里克特定律，随着震级的增大，地震数量呈指数减少，2 类样本也就是 $M_S \geqslant 5.0$ 地震的样本数量较少，模型没有足够的样本特征去学习，从而导致了一些误判；二是大地震的发生本身就具有一定的复杂

性，地震类型、传播途径等因素都会导致震前特征复杂多变，这也增加了模型的预测难度.

对于震源方位预测，与震级预测类似，0 类预测的评价指标都很高，可达到 0.87 以上. 除此之外，其它类别预测结果的准确率、召回率和 F_1 指标都与样本数量有直接关系，如永胜台 4 类结果，该类样本数很少，因此模型未得到充分训练，导致模型评价指标相对较低.

昭通、姑咱和腾冲三个台站的预测结果也与永胜台类似，其震级预测的准确率分别可达到 0.823，0.845 和 0.852，震源方位预测的准确率分别可达 0.838，0.755 和 0.794. 这说明，三个台站钻孔应变数据的震前特征也都能够由本文提出的 1D-CNN 模型挖掘得到，模型在地震发生前一周的应变数据中学习到了十分有效的震前特征.

针对我国五次典型的大地震，本文进一步单独给出了实际预测结果(表 6). 从表 6 可以看出，5 次地震震级和方位的预测标签均与实际相符，说明本文模型捕捉到了有效的震前应变信息，有一定的预测能力.

此外，为了观察不同震级类别所对应的最优参数范围，本文还分析了不同震级类别下，不同时间窗口及预测窗口对模型预测效率的影响. 由于四个台站的结果相差不大，本文依旧以永胜台站为例，结果如图 6 所示.

表 6 五次典型强震的实际预测效果

Table 6 Prediction results of five typical strong earthquakes

序号	发震日期	发震地点	M_S	预测震级标签	实际震级标签	预测方位标签	实际方位标签
1	2010-04-14	青海玉树	7.3	2	2	1	1
2	2013-07-22	甘肃岷县	6.7	2	2	2	2
3	2014-08-03	云南鲁甸	6.6	2	2	3	3
4	2016-10-17	青海杂多	6.3	2	2	1	1
5	2017-08-08	四川九寨沟	7.0	2	2	2	2

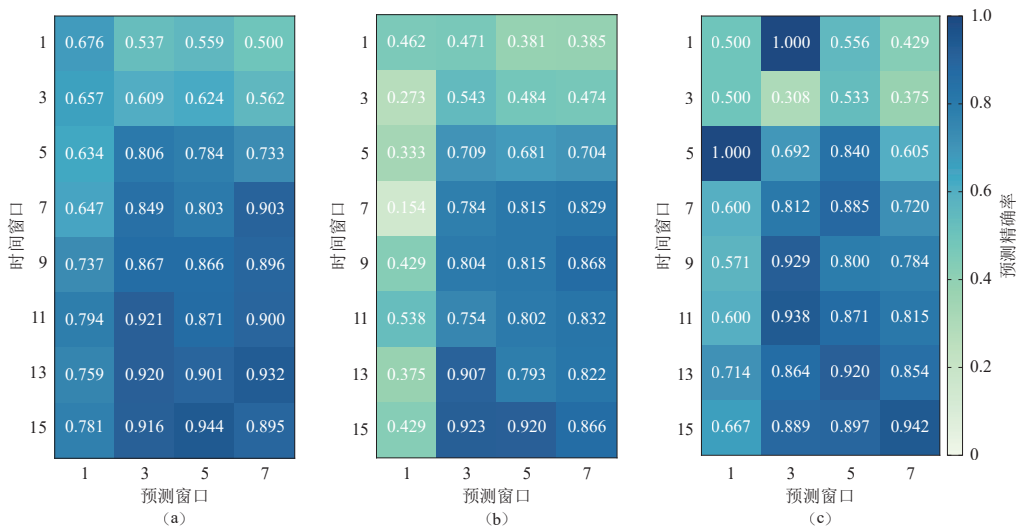

图 6 各震级标签下不同时间窗口及预测窗口下的预测精确率

(a) 0 类标签；(b) 1 类标签；(c) 2 类标签

Fig. 6 Prediction accuracies under different time windows and prediction windows for each class

(a) Class 0；(b) Class 1；(c) Class 2

整体上，三个震级的预测精确率都随着时间窗口的增大而上升．这是由于时间窗口越长，含有的应变特征信息就越多，准确率自然就会上升．首先，0 类预测基本上在时间窗口为 5、预测窗口为 3 后，精确率就基本上稳定在 0.80 以上．其次，1 类和 2 类的准确率规律性略差，这可能是因为震前特征丰富，增加了随机性．其中，1 类是在时间窗口为 7、预测窗口为 3 后，精确率可达到 0.80 以上．其余方案准确率都很低．而 2 类整体上比 1 类高一些，随机性也更大．这可能反映了钻孔应变数据对大地震具有一定的短临预测能力，同时也说明了不同大地震间的震前特征存在差异，可能有些特征显著、有些则难以提取，有些异常特征持续长、有些则短．因此，如何设计网络提取样本量小、差异化又大的大地震震前异常特征，也需要试验和讨论．最后，由于本文使用了滑动窗口来构造样本，因此在两个连续样本中可能会存在部分重复特征的情况，而时间窗口越长，重复就越多，因此图 6 右下角的精确率可能会比实际情况偏高一些．

4 讨论与结论

面对地震短临预测这一科学难题，本文突破传统方法的局限，提出了采用深度学习手段进行地震震级和方位预测的策略．通过钻孔应变仪四分量数据构建震前应变特征数据集，采用 1D-CNN 搭建了地震震级和方位的预测模型，其中，地震震级划分为无震、$3.0 \leqslant M_S < 5.0$ 以及 $M_S \geqslant 5.0$ 三类，震源方位标签根据四分量钻孔应变仪 1 分量方位角方向划分为四类．结果显示，震前应变特征在震级和方位预测上的预测准确性可高达 80% 以上．这表明，对于不同方向放置的钻孔应变仪，本文搭建的 1D-CNN 模型都能利用该台站数据预测地震震级与大致方位，也说明了卷积神经网络结构有能力从钻孔应变数据中挖掘到有效的震前异常特征．这为地震前兆观测研究提供了新的思路和方法．

然而，本文将复杂的地震预测问题抽象成深度学习中的分类任务，这与实际地震预测业务还有一定差距．实际预测中，从业人员要根据全新的观测数据给出地震预测意见，而三要素的预测报准率、虚报率、漏报率都会在决策中起到重要影响．因此，地震预测仍是一个科学难题．若要使预测任务更加贴近实际应用，本文仍有一些待改进之处．例如：由于震级及空间位置的双方面限制会导致大地震的样本非常少，本文将震级与方位视为两个独立的标签，想要实现震级与方位的同时预测还存在一定挑战；本文网络模型在四川云南地区的地震预测中表现良好，然而在其它地区是否具有普适性还有待验证．考虑到实际应用价值，未来将设计更完善的网络，将两种标签组合起来，解决样本不平衡问题，实现地震三要素预测；或采用多个应变台站联合，加入震中距、角度等更多的地震位置信息作为预测标签，通过划分分辨率较高的空间网格来提高方位预测的精度；对弱震区或其它多震区的应变观测台站进行试验验证，构建稳健、适用强的短临地震预测模型．

参　考　文　献

陈运泰．2007．地震预测：进展、困难与前景 [J]．地震地磁观测与研究，**28**(2)：1–24．

Chen Y T．2007．Earthquake prediction：Progress，difficulties and prospects[J]．*Seismological and Geomagnetic Observation and Research*，**28**(2)：1–24 (in Chinese)．

陈运泰．2008．地震预测要知难而进 [J]．求是，**15**：58–60．

Chen Y T．2008．Predicting earthquakes requires perseverance and determination[J]．*Qiushi*，**15**：58–60 (in Chinese)．

池顺良，刘琦，池毅，邓涛，廖成旺，阳光，张贵萍，陈洁．2013．2013 年芦山 M_S7.0 地震的震前及临震应变异常 [J]．地震

学报，**35**(3)：296–303.

Chi S L，Liu Q，Chi Y，Deng T，Liao C W，Yang G，Zhang G P，Chen J. 2013. Borehole strain anomalies before the 20 April 2013 Lushan M_S7.0 earthquake[J]. *Acta Seismologica Sinica*，**35**(3)：296–303 (in Chinese).

池顺良，张晶，池毅. 2014. 汶川、鲁甸、康定地震前应变数据由自洽到失洽的转变与地震成核[J]. 国际地震动态，(12)：3–13.

Chi S L，Zhang J，Chi Y. 2014. Failure of self-consistent strain data before Wenchuan，Ludian and Kangding earthquakes and its relation with earthquake nucleation[J]. *Recent Development in World Seismology*，(12)：3–13 (in Chinese).

高曙德. 2016. 深井地电观测技术在地震监测中的应用探讨[J]. 地球物理学进展，**31**(5)：2078–2088.

Gao S D. 2016. Discussion on the deep well geoelectric observation technique applied in earthquake monitoring[J]. *Progress in Geophysics*，**31**(5)：2078–2088 (in Chinese).

高曙德. 2020. 四川九寨沟7.0级地震前震情跟踪概述及震后总结[J]. 地球物理学进展，**35**(4)：1250–1260.

Gao S D. 2020. Overview of tracking of Sichuan Jiuzhaigou M_S7.0 earthquake in 2017 and its post earthquake precursor anomaly summary[J]. *Progress in Geophysics*，**35**(4)：1250–1260 (in Chinese).

顾国华，王武星. 2020. 2016年日本本州东岸近海M_S7.2地震前后的地壳运动[J]. 地震学报，**42**(2)：196–204.

Gu G H，Wang W X. 2020. Crustal movements of the eastern Honshu offshore M_S7.2 earthquake in Japan in 2016[J]. *Acta Seismologica Sinica*，**42**(2)：196–204 (in Chinese).

江在森，方颖，武艳强，王敏，杜方，平建军. 2009. 汶川8.0级地震前区域地壳运动与变形动态过程[J]. 地球物理学报，**52**(2)：505–518.

Jiang Z S，Fang Y，Wu Y Q，Wang M，Du F，Ping J J. 2009. The dynamic process of regional crustal movement and deformation before Wenchuan M_S8.0 earthquake[J]. *Chinese Journal of Geophysics*，**52**(2)：505–518 (in Chinese).

李进武，邱泽华. 2014. 钻孔应变仪观测的面应变潮汐因子初步分析[J]. 地球物理学进展，**29**(5)：2013–2018.

Li J W，Qiu Z H. 2014. Analysis on strain tidal factor observed borehole strainmeters[J]. *Progress in Geophysics*，**29**(5)：2013–2018 (in Chinese).

廖晓峰，何康，张明东，何畅，魏强. 2018. 基于平滑伪魏格纳-维勒时频分析的地磁数据研究[J]. 地震，**38**(1)：107–116.

Liao X F，He K，Zhang M D，He C，Wei Q. 2018. Time-frequency analysis of geomagnetic data based on smooth pseudo-Wigner-Ville distribution[J]. *Earthquake*，**38**(1)：107–116 (in Chinese).

刘琦，张晶，池顺良，闫伟. 2014. 2013年芦山M_S7.0地震前后姑咱台四分量钻孔应变时频特征分析[J]. 地震学报，**36**(5)：770–779.

Liu Q，Zhang J，Chi S L，Yan W. 2014. Time-frequency characteristics of four-component borehole strain at Guzan station before and after 2013 Lushan M_S7.0 earthquake[J]. *Acta Seismologica Sinica*，**36**(5)：770–779 (in Chinese).

卢宏涛，张秦川. 2016. 深度卷积神经网络在计算机视觉中的应用研究综述[J]. 数据采集与处理，**31**(1)：1–17.

Lu H T，Zhang Q C. 2016. Applications of deep convolutional neural network in computer vision[J]. *Journal of Data Acquisition and Processing*，**31**(1)：1–17 (in Chinese).

牛安福，张凌空，章静，闫伟，赵静，岳冲，苑争一. 2021. 汶川地震前远场形变异常及其与地震相关性研究[J]. 中国地震，**37**(1)：15–21.

Niu A F，Zhang L K，Zhang J，Yan W，Zhao J，Yue C，Yuan Z Y. 2021. Far field deformation anomaly and its correlation with earthquake before Wenchuan earthquake[J]. *Earthquake Research in China*，**37**(1)：15–21 (in Chinese).

邱泽华，唐磊，周龙寿，阚宝祥. 2009. 四分量钻孔应变台网汶川地震前的观测应变变化[J]. 大地测量与地球动力学，**29**(1)：1–5.

Qiu Z H，Tang L，Zhou L S，Kan B X. 2009. Observed strain changes from 4-component borehole strainmeter network before 2008 Wenchuan earthquake[J]. *Journal of Geodesy and Geodynamics*，**29**(1)：1–5 (in Chinese).

邱泽华，唐磊，张宝红，宋茉. 2012. 用小波-超限率分析提取宁陕台汶川地震体应变异常[J]. 地球物理学报，**55**(2)：538–546.

Qiu Z H，Tang L，Zhang B H，Song M. 2012. Extracting anomaly of the Wenchuan earthquake from the dilatometer recording at NSH by means of wavelet-overrun rate analysis[J]. *Chinese Journal of Geophysics*，**55**(2)：538–546 (in Chinese).

邱泽华，杨光，唐磊，郭燕萍，张宝红．2015．芦山 *M*7.0 地震前姑咱台钻孔应变观测异常 [J]．大地测量与地球动力学，**35**(1)：158–161．

Qiu Z H，Yang G，Tang L，Guo Y P，Zhang B H．2015．Abnormal strain changes prior to the *M*7.0 Lushan earthquake observed by a borehole strainmeter at Guzan[J]．*Journal of Geodesy and Geodynamics*，**35**(1)：158–161 (in Chinese)．

于紫凝．2022．钻孔应变观测数据的震前异常提取与评价方法研究 [D]．长春：吉林大学：16–17．

Yu Z N．2022．*Pre-Earthquake Anomaly Extraction and Evaluation of Borehole Strain Observations*[D]．Changchun：Jilin University：16–17 (in Chinese)．

张敏，赵燕杰，文勇，孟鑫．2014．青海地区钻孔应变同震响应特征分析 [J]．高原地震，**26**(3)：52–56．

Zhang M，Zhao Y J，Wen Y，Meng X．2014．Characteristics analysis on coseismic response of borehole strain in Qinghai area[J]．*Plateau Earthquake Research*，**26**(3)：52–56 (in Chinese)．

周飞燕，金林鹏，董军．2017．卷积神经网络研究综述 [J]．计算机学报，**40**(6)：1229–1251．

Zhou F Y，Jin L P，Dong J．2017．Review of convolutional neural network[J]．*Chinese Journal of Computers*，**40**(6)：1229–1251 (in Chinese)．

张希，秦姗兰，贾鹏，李瑞莎．2019．2016 年门源 M_S6.4 地震孕育-发生的地形变异常特征 [J]．地震，**39**(4)：27–38．

Zhang X，Qin S L，Jia P，Li R S．2019．Anomalies on characteristics of crustal deformation during the pregnant process of the Menyuan M_S6.4 earthquake[J]．*Earthquake*，**39**(4)：27–38 (in Chinese)．

Chi C Q，Zhu K G，Yu Z N，Fan M X，Li K T，Sun H H．2019．Detecting earthquake-related borehole strain data anomalies with variational mode decomposition and principal component analysis：A case study of the Wenchuan earthquake[J]．*IEEE Access*，**7**：157997–158006．

De Santis A，Balasis G，Pavón-Carrasco F．J，Cianchini G，Mandea M．2017．Potential earthquake precursory pattern from space：The 2015 Nepal event as seen by magnetic Swarm satellites[J]．*Earth Planet Sci Lett*，**461**：119–126．

Hirose H．2011．Tilt records prior to the 2011 off the Pacific coast of Tohoku earthquake[J]．*Earth Planet Space*，**63**(7)：655–658．

Ouyang X Y，Zong Q G，Bortnik J，Wang Y F，Chi P J，Zhou X Z，Yue C，Hao Y Q．2018．Nightside ULF waves observed in the topside ionosphere by the DEMETER satellite[J]．*J Geophys Res：Space Phys*，**123**(9)：7726–7739．

Parrot M．2017．Events linked to the lithosphere-atmosphere-ionosphere coupling observed by DEMETER[J]．*USRI Radio Sci Bull*，**2017**(360)：75–79．

Pulinets S，Ouzounov D，Karelin A，Davidenko D．2018．Lithosphere-atmosphere-ionosphere-magnetosphere coupling：A concept for pre-earthquake signals generation[C]//*Pre-Earthquake Processes：A Multi-Disciplinary Approach to Earthquake Prediction Studies*．Washington D.C.：AGU and John Wiley & Sons Inc：79–98．

Yu Z N，Hattori K，Zhu K G，Chi C Q，Fan M X，He X D．2020．Detecting earthquake-related anomalies of a borehole strain network based on multi-channel singular spectrum analysis[J]．*Entropy*，**22**(10)：1086．

Yu Z N，Zhu K G，Hattori K，Chi C Q，Fan M X，He X D．2021a．Borehole strain observations based on a state-space model and ApNe analysis associated with the 2013 Lushan earthquake[J]．*IEEE Access*，**9**：12167–12179．

Yu Z N，Hattori K，Zhu K G，Fan M X，Marchetti D，He X D，Chi C Q．2021b．Evaluation of pre-earthquake anomalies of borehole strain network by using receiver operating characteristic curve[J]．*Remote Sens*，**13**(3)：515．

Zhu K G，Yu Z N，Chi C Q，Fan M X，Li K Y．2019．Negentropy anomaly analysis of the borehole strain associated with the M_S8.0 Wenchuan earthquake[J]．*Nonlinear Process Geophys*，**26**(4)：371–380．