Deep learning for identifying earthquake precursors: Applications and challenges in subsurface fluid signals

Bingfei Chu,[1,3,6] Yan Zhang,[1,3,6] Liyun Fu,[2] Shengwen Qi,[1,3,*] Gaoxiang Chen,[4] Zheming Shi,[5] Tianming Huang,[1,3,*] Zhonghe Pang,[1,3,*] and Huai Zhang[3,*]

[1]Key Laboratory of Shale Gas and Geoengineering, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China
[2]National Key Laboratory of Deep Oil and Gas, China University of Petroleum (East China), Qingdao 266580, China
[3]College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
[4]Zhejiang Lab, Hangzhou 311100, China
[5]MOE Key Laboratory of Groundwater Circulation and Environmental Evolution, China University of Geosciences, Beijing 100083, China
[6]These authors contributed equally
*Correspondence: qishengwen@mail.iggcas.ac.cn (S. Q.); tmhuang@mail.igcas.ac.cn (T. H.); z.pang@mail.iggcas.ac.cn (Z. P.); hzhang@ucas.ac.cn (H. Z.)

Received: May 18, 2024; Accepted: August 4, 2024; Published Online: August 14, 2024; https://doi.org/10.59717/j.xinn-geo.2024.100093

Citation: Chu B., Zhang Y., Fu L., et al., (2024). Deep learning for identifying earthquake precursors: Applications and challenges in subsurface fluid signals. The Innovation Geoscience **2(4)**: 100093.

Earthquakes, as one of the most destructive natural disasters, have caused immense damage and even irreversible losses to human society. Although technology advances, accurately predicting earthquakes still poses an unmet challenge in Earth sciences. Subsurface fluid monitoring plays a crucial role in earthquake prediction, with its abnormal changes serving as indicators of seismic precursors (e.g., the Haicheng earthquake in China).

Identifying reliable subsurface fluid precursor signals has always been the goal of hydro-seismologists. However, traditional monitoring methods struggle with detecting minor changes and dealing with noise.[1] Confronted with vast amounts of monitoring data and earthquake catalogs, traditional anomaly detection methods (such as correlation analysis, trend analysis, regression analysis, and multiscale wavelet decomposition) are significantly limited.

Development in deep learning technologies have brought new hope to this goal, based on the superior ability of the deep learning models in handling vast amounts of data and detecting subtle, complex, nonlinear signals that precede earthquakes. This article focuses on deep learning models and methods suitable for the identification of subsurface fluid precursor signals, and further discusses the future development of these models. This can enhance our understanding of subsurface fluid anomaly signals, providing new ideas and methods for subsurface fluid precursor monitoring and future earthquake prediction.

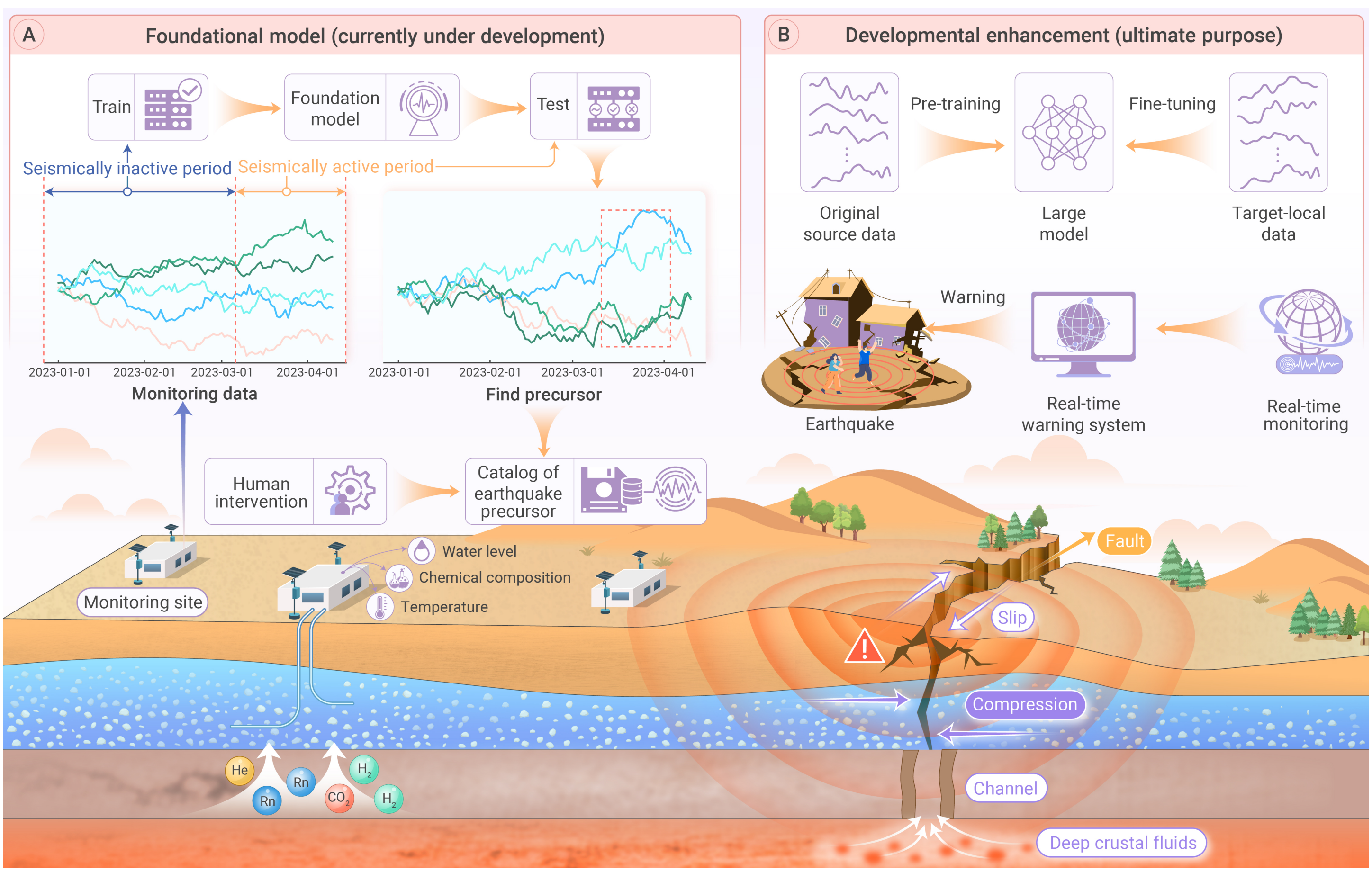

Figure 1. The schematic diagram showing the earthquake precursor prediction and the machine learning model.

WHY CAN SEISMIC PRECURSOR SIGNALS BE DETECTED IN SUBSURFACE FLUIDS?

Groundwater systems can greatly amplify minor tectonic deformations and seismic stress field changes, which are reflected in changes in well pore pressures, water levels (heads), and the chemical composition of water and gases.[1] By observing changes in water levels, water temperature, chemical composition of water, and soil gases, information about the internal dynamic processes of the Earth's crust can be gathered. Undoubtedly, changes in subsurface fluids are closely associated with regional stress-strain responses and fault activity.

Hypotheses that might cause deviated mechanisms during the earthquake generation process include the development of fracture volumes near the epicenter, the propagation of deformation wavefronts, and seismic slip along parts of fault planes. For example: Bletery et al. (2023), through the analysis of high-resolution data from the Global Positioning System (GPS), observed a precursory slip phase at the onset of major earthquakes.[2] These hypotheses and observed phenomena provide reliable evidence for changes in the stress-strain field before major earthquakes. Thus, anomalous changes in subsurface fluids can act as effective indicators of short-term earthquake precursors.

In China, subsurface fluid monitoring is widely used in earthquake precursor studies; by analyzing fluid anomalies, it provides valuable information for identifying earthquake hazard zones and tracking imminent precursor anomalies. Therefore, subsurface fluid monitoring is an indispensable tool in the field of earthquake prediction, significantly enhancing the accuracy and scientific basis of earthquake forecasts.

CHALLENGES AND OPPORTUNITIES

Anomaly detection in time series typically involves two methods: data-based reconstruction and prediction. An anomaly is identified when the model's predicted or reconstructed values significantly deviate from observed data beyond a set threshold. It's crucial to note that not all anomalies indicate earthquake precursors, which requires professional verification and exclusion, but undoubtedly, the deep learning method can quickly locate and identify possible potential precursor information. Those relies on the model's ability to effectively reconstruct or predict normal data (Figure 1A). Hence, it is necessary to ensure that the model's training set contains as few anomalies as possible.

To ensure the model is well-trained, it is crucial to effectively segment the monitoring data, such as into periods of seismically active (SA) and seismically inactive (SI).[3] Historical data from periods of seismically inactive are utilized for model training, and data from periods of seismically active are used for model testing (Figure 1A). This approach hinges on the hypothesis that there is a significant shift in stress-strain exclusively prior to a major earthquake, as outlined in the previous section. More and more research, including studies by Bletery et al. (2023), validates the possibility of this phenomenon.[2] This provides hope and theoretical support for utilizing deep learning to identify precursors in subsurface fluids.

The application of deep learning in earthquake prediction is rapidly evolving, yet it remains in its initial stages; the limitation is due to the existing datasets needing further exploration.[4] Similarly, most of the known subsurface fluid monitoring data currently lacks explicit labeling information. Applying unsupervised learning might be a more effective strategy, employing algorithms to identify patterns in data without prior labels, a practice widely used (e.g., in precursor signals for major landslides and tsunamis).[4]

In recent years, many deep learning algorithms have demonstrated broad application potential in anomaly detection tasks across various fields. For instance, prediction-based methods such as LSTM, Informer, and Transformer, etc. Similarly, data reconstruction methods like LSTM-AE, TSMAE, ANOMALYBERT, etc. These methods have been successfully applied to anomaly detection in time-series data. When handling complex and high-dimensional subsurface fluid data, developing hybrid models that combine traditional signal analysis methods with deep learning techniques may be more effective. For instance, using PCA for data dimension reduction, and employing wavelet transform and EEMD to decompose time series data are promising directions worth exploring.

Given that current research is still in the exploratory stage, facing challenges such as noise in the data and the lack of clear labels, it is particularly important to develop unsupervised deep learning models that are highly robust and have strong generalization abilities (robustness refers to the model's ability to handle noise and disturbances; generalization refers to the model's ability to perform well on new data).

OUTLOOK

In the future, deep learning algorithms will find increased applications in subsurface fluid monitoring. Through human-computer interaction, we can further develop and refine catalogs of subsurface fluid seismic precursor information, including: labels for seismically active and inactive periods, precursor anomaly labels, corresponding earthquake magnitudes, epicentral distances, and times from precursor occurrence to earthquake onset.

Once the subsurface fluid precursor monitoring model has matured and a comprehensive catalog of seismic precursor information is established, the base model can be expanded to a larger model to enhance its accuracy and versatility (Figure 1B). Drawing on the underlying logic similar to that of ChatGPT, the model uses large-scale subsurface fluid data for self-supervised learning during the training phase to identify precursor anomalies and predict earthquakes. Based on this, by fine-tuning with labeled data for specific regions, the model is further optimized through supervised learning to ensure it precisely adapts to specific subsurface fluid monitoring tasks.

The ultimate goal is to establish an automated, real-time subsurface fluid precursor monitoring and early warning system. To achieve this goal, it is necessary to further optimize the model architecture, including improving data storage mechanisms and processing speed.[5] In contrast to conventional earthquake forecasting systems, by real-time monitoring and analyzing subsurface fluid anomalous signals, it can provide longer lead times for earthquake warnings, thus reducing potential earthquake damages (Figure 1B). The system's design will focus on increasing the accuracy of warnings and reducing false alarm rates, ensuring the practicality and effectiveness of the alert system. Implementing this system will also require in-depth research and development across multiple technical fields.

CONCLUSION

Monitoring and identifying anomalies in subsurface fluids are crucial for earthquake precursor research. As data volume grows rapidly, traditional analysis and monitoring methods struggle with information barriers and dataset integration. Advances in artificial intelligence, particularly deep learning, offer new solutions. Deep neural networks enhance anomaly detection, and the development of large models is set to overcome traditional limitations, enabling true data integration and revealing more earthquake precursors. By creating a comprehensive catalog of subsurface fluid data, developing foundational models (especially unsupervised learning), and advancing to large models, we can achieve systematic real-time monitoring and early warning.

REFERENCES

1. Wang, C.-Y. and Manga, M. (2021). Hydrologic precursors, in: Water and earthquakes, Lecture notes in earth system sciences. Springer, Cham, pp: 343–368. DOI:10.1007/978-3-030-64308-9_13.
2. Bletery, Q. and Nocquet, J.-M. (2023). The precursory phase of large earthquakes. Science, 381: 297–301. DOI: 10.1126/science.adg2565.
3. Yan, X., Shi, Z., Wang, G., et al. (2021). Detection of possible hydrological precursor anomalies using long short-term memory: A case study of the 1996 Lijiang earthquake. J Hydrol. 599: 126369. DOI: 10.1016/j.jhydrol.2021.126369.
4. Bergen, K.J., Johnson, P.A., De Hoop, M.V., et al. (2019). Machine learning for data-driven discovery in solid Earth geoscience. Science 363: eaau0323. DOI: 10.1126/science.aau0323.
5. Kuang, W., Yuan, C., and Zhang, J. (2021). Real-time determination of earthquake focal mechanism via deep learning. Nat. Commun. 12: 1432. DOI: 10.1038/s41467-021-21670-x.

FUNDING AND ACKNOWLEDGMENTS

This research was supported by the National Natural Science Foundation of China (Grants 42141009） and the Joint Funds of the National Natural Science Foundation of China (No. U2239205).

AUTHOR CONTRIBUTIONS

Conceptualization: B.C., Y.Z.. Investigation: B.C., Y.Z.. Writing – Initial: B.C., Y.Z.. Writing - Review, Editing: B.C., Y.Z., L.F., H.Z., G.C., Z.S., S.Q., T.H., Z.P.. Funding acquisition: S.Q., H.Z..

DECLARATION OF INTERESTS

The authors declare no competing interests.