Comparative Advances in AI-Driven Earthquake Intelligence: Machine Learning, Deep Learning, and Large Language Models for Prediction and Emergency Management

Mahmoud Shabrawy[1,*], **Nahla B. Abdel-Hamid**[1], **El-Sayed M. El-Kenawy**[2,3], **Mohamed M. Abdelsalam**[1]

[1]Computer Engineering and Control Systems Department, Faculty of Engineering Mansoura University, Mansoura, Egypt

[2]Department of Communications and Electronics, Delta Higher Institute of Engineering and Technology, Mansoura, 35111, Egypt

[3]Applied Science Research Center, Applied Science Private University, Amman, Jordan

Emails: mshabrawy@std.mans.edu.eg; nahla_bishri@mans.edu.eg; skenawy@ieee.org; mohmoawad@mans.edu.eg

Abstract

Prediction, hazard evaluation, and response to disasters remain severely problematic due to the nonlinear and multiscale nature of crustal behaviour on Earth and the relative sparsity, noise, and heterogeneity of observations. Even with significant improvements in seismology, conventional statistical and physical models still struggle to make short-term predictions, consistently identify precursors, and provide dynamic situational awareness of the state and post-seismic events. In turn, the rapid development of machine learning (ML), deep learning (DL), and large language models (LLMs) has created new opportunities to extract meaningful patterns from diverse datasets, integrate multimodal information, and enable real-time decision-making in earthquake-prone regions. The paper provides an overview of recent advances in AI-based earthquake studies, including environmental precursors, spatiotemporal seismic prediction, ground-motion prediction, multimodal structural damage, and LLM-based knowledge integration. We discuss developments in hydrochemical anomaly detection using ML models developed in the context of long-term hot spring monitoring and highlight improvements in anomaly detection, as well as the challenges posed by varying indicators and time-dependent instabilities. At the world scale, we consider deep architectures that use spherical convolutions and attention to model seismicity on the curved surface of the Earth, showing significant improvements in accuracy, recall, and long-term dependency modeling. Simultaneously, ensemble ML models for peak ground acceleration prediction and SARIMAX-based time-series models with exogenous variables demonstrate how data-driven models can supersede traditional attenuation relationships and capture some fundamental temporal behaviour of seismic processes. Beyond prediction, we consider the growing importance of LLMs as integrative reasoning systems that can combine heterogeneous streams of information, such

DOI: https://doi.org/10.54216/MOR.050101
Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

as textual reports, sensor logs, social media content, and visual signals. These paradigms support the new pipelines of building earthquake emergency knowledge graphs, performing retrieval-based logistics prediction, creating engineering-grade structural damage estimates, and providing real-time situational awareness based on citizen communication. Their increased utility, however, also creates new domain-grounding, bias, interpretability, and reliability issues in high-stakes settings. In these various uses, there are a few common barriers, such as limited model generalization to tectonic settings, insufficient high-magnitude events for training, physical constraints, and uncertainty quantification, all of which can be addressed. These results highlight that future systems are likely best built by blending physical knowledge with data-driven systems, using multimodal sources including seismic, environmental, satellite, geodetic, and social data, and using LLMs as embodiments of agents operating on transparent tools rather than opaque creators. At the end of the paper, the main directions for future research have been identified, including the need for standardized multimodal benchmarks, hybrid physics-ML designs, simulation-based training controls, robust uncertainty estimation methods, and governance systems that are transparent, fair, and reliable. These advances, combined, will no doubt lead to a new generation of AI-modified seismic forecasting and disaster-response structures that are scientifically defensible and operationally feasible, eventually making societies less susceptible to earthquake hazards.

Keywords: Earthquake Prediction; Machine Learning and Deep Learning; Large Language Models; Spatiotemporal Seismic Forecasting; Disaster Response and Decision Support

1 Introduction

Earthquakes are instant releases of the stored tectonic stress that are carried along the crust of the earth as seismic waves. These occurrences are very dangerous to human life, infrastructure, and socio-economic equilibrium especially in highly populated or geologically active areas [1]. Even though the underlying principles of earthquakes such as stress build up, strain release and fault breakage are conceptually well understood, their timing, magnitude and spatial distribution are historically hard to predict. This challenge is inherently present in the complexity and nonlinearity of geophysical processes: the distribution of stress across fault systems, the interactions among faults have cascading effects, and visible precursors are often not only ambiguous, subtle, and highly region-dependent. Consequently, prediction of earthquake is a unresolved scientific problem even after decades of seismology studies occurring in the field of seismology research [2].

Conventional seismology practices, such as statistical recurrence models, empirical relationships of attenuation and geophysical simulation methodologies have played a significant role in illustrating our knowledge on long-term seismic hazards. One example is the probabilistic seismic hazard analysis (PSHA) that offers powerful constructs to predict the probability of earthquakes in the multi-decadal period. Nevertheless, those approaches have difficulty in short-term or event-specific prediction since they presume statistical patterns incompleteness in the complex spatiotemporal behavior of seismicity. In addition, the predictive power of the conventional methods is further limited by observational constraints that include sparse sensor coverage, noisy observations, incomplete catalogues, as well as measurement uncertainty, etc. [3].

Machine learning (ML) and deep learning (DL) have been in the limelight in response to these challenges to seismic analysis [4]. ML provides versatile data-driven methods that have the capacity to discover

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

multifaceted correlations, nonlinear connections and latent structures in extensive and various datasets [5]. Scholars have used ML models to categorize seismic events, distinguish earthquakes and noise, identify spatiotemporal clusters, and estimate the ground-motion intensity. The field has been transformed particularly by deep learning, which allows them to automatically extract features directly out of raw waveforms and also to discover concealed temporal coupling of the seismic sequences in the sequences themselves [6]. Convolutional neural networks have found broad use in classification of waveforms and detecting events, whereas recurrent networks which include LSTMs and GRUs have shown good results in their ability to predict the time-varying characteristics of seismic wave patterns. More recently, architectures of transformers, which were initially developed to support natural language processing have demonstrated potential in long-range dependencies and multi-scale structure of seismic catalogs, allowing more fine grained descriptions of earthquake sequences to be made [7].

In addition to the analysis of the waveforms and catalog, the application of ML in the exploration of non-seismic precursors and the environmental indicators has also been used. A single research area studies predictive potential of fluid chemical anomalies in hydrothermal systems that are in active fault zones. As an illustration, a study that utilized hydrochemical monitoring in six hot springs along the southeastern coast of China in a period of two and a half years came up with a model of prediction that entailed the combination of six various ML algorithms used to predict earthquakes of magnitude $M \geq 5$ in Taiwan. This model has shown that the ML-based anomaly detection is superior to the older statistical methods on the hot-spring chemistry as well as shown that incorporating sampling time with them as an explicit feature can greatly improve their predictive capabilities. Simultaneously, the work raised some significant issues: the predictive capability among springs and types of indicators needs to be regularly fine-tuned, model parameters, including the rate of anomaly detection and the response-time threshold, need to be carefully configured, and the differentiation between pre- and post-seismic anomalies or the detection of the epicentral positions is still challenging to draw the line between these two concepts [8]. These results highlight the potential and the weaknesses of using the environmental precursors with the help of ML.

Parallel to such application-specific studies, comprehensive reviews have begun to map the broader landscape of ML usage in earthquake seismology. One such survey systematically examined advances across four main areas: catalog development, seismicity analysis, ground-motion prediction, and crustal deformation analysis [9]. In catalog development, ML has been deployed for event detection and classification, arrival-time picking, similar waveform searching, focal mechanism estimation, and even analysis of paleoseismic records [9]. In seismicity and risk analysis, models have been used to characterize spatiotemporal patterns, evaluate hazard, and support probabilistic forecasting. For ground-motion prediction, the review classified ML-based approaches according to whether they predict intensity measures or full time series, and whether they operate on hand-crafted features or directly on time-series input, noting that imbalanced training data remains a critical challenge. Finally, in crustal deformation studies, ML has been applied to geodetic data for clustering analysis and detection of signals associated with seismic and aseismic processes [9]. Collectively, this body of work illustrates that ML is now embedded across nearly the entire seismological workflow.

At the global scale, however, earthquake prediction faces additional challenges related to the representation of the Earth's curved geometry. Many DL models operate in Euclidean coordinates and thus suffer from spatial distortions when applied to global latitude–longitude grids. To address this, a recent study proposed a new architecture based on spherical convolutional LSTMs within a U-Net framework, explicitly designed for global-scale spatiotemporal earthquake forecasting [10]. In this model, spherical convolutional neural

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

networks are embedded into the LSTM structure, allowing the network to better respect spherical geometry and mitigate distortions that arise when standard convolutions are used on global maps. By generating earthquake distribution maps at high resolutions (e.g., 1920 and 3840 in map size) and formulating global prediction as a spatiotemporal series problem, the authors showed that their spherical ConvLSTM–U-Net model achieved improved precision and recall relative to previous methods, particularly for the lower map resolution where gains in precision were substantial [10]. These results highlight the importance of geometric fidelity in global DL models and demonstrate that architectural choices can meaningfully influence predictive skill.

Machine learning has also been increasingly applied to ground-motion modeling, where one of the key tasks is predicting peak ground acceleration (PGA) at sites of interest. Traditional ground-motion prediction equations (GMPEs) are typically region-specific and based on parametric functional forms that may not capture complex dependencies across diverse tectonic regimes. To overcome these limitations, an ensemble model named SeisEML (Seismological Ensemble Machine Learning) has been proposed for cross-region PGA prediction [11]. SeisEML integrates hybridized models, kernel-based methods, tree-based regressors, and standard regression algorithms, and is trained on more than 20,000 accelerograms from the Japanese Kyoshin Network. When evaluated with MAE and RMSE metrics, SeisEML yields approximately half the error of conventional attenuation relations [11]. Its ability to reproduce iso-acceleration contour maps for several Japanese earthquakes and to generalize successfully to Iranian earthquake datasets demonstrates that ensemble ML models can serve as robust substitutes for GMPEs in both regional and cross-regional settings, provided that tectonic environments are sufficiently similar [11].

Alongside these ML and DL advances, classical time-series modeling continues to play a role in earthquake forecasting, particularly when exogenous geophysical variables are explicitly incorporated. An illustrative example is the application of Seasonal Autoregressive Integrated Moving Average with Exogenous Variables (SARIMAX) models to earthquake time-series forecasting [12]. In this framework, seismic activity is modeled as a temporal process influenced by additional covariates such as historical seismicity, geological characteristics, and geodetic measurements. By constructing and comparing multiple SARIMAX specifications and evaluating them using metrics like RMSE and MAE, the study demonstrated that such models can capture recurring temporal patterns and the influence of exogenous factors on seismic activity [12]. While SARIMAX remains more constrained than high-capacity DL models, its interpretability and explicit handling of external drivers make it a valuable baseline for understanding temporal dynamics and for informing decision-makers engaged in disaster preparedness and mitigation.

With these trends, the new frontier of utilizing large language models (LLMs) has been introduced that has significant consequences in earthquake science. LLMs are based on transformer architectures and trained on large-scale text corpora and have powerful reasoning, semantic interpretation, and knowledge integration abilities. The possible uses to which they can be put in seismology are many. LLMs may help to structure and interpret past earthquake alerts, process unstructured field survey or social media data, create domain-specific knowledge graphs, support emergency decision-making, facilitate multimodal input, and present natural-language explanations of model outputs. More significantly, LLMs are able to reason dynamically on inputs of heterogeneity, such as structured data, textual descriptions and multimodal signals, and are thus useful in both scientific and operational response. By integrating with the above-described [8, 9, 10, 11, 12] specialized ML and DL models, LLMs can become integrative engines in the sense that they can coordinate various predictive elements, put their results into context, and represent doubt in a way comprehensible to humans.

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

However, the process of using LLM as an instrument to predict and analyze earthquakes has its own challenges. Typical standard LLMs are not trained on specialized geophysical data, i.e. they might not have a physical law or foreground base, or empirical seismic patterns. In the absence of domain adaptation, they will be able to produce plausible-sounding but factually incorrect statements, and will fail to be consistent in their treatment of numerical data and time-series reasoning, and may need special fine-tuning, structured prompting, or integration with external computational units. More so, explainability and reliability are also important issues, particularly when the model outputs have a toxic effect on high-stakes choices involving public safety, hazard communication, or emergency resource distribution. The ethics of data governance, transparency and accountability arise in the case of the inclusion of the LLMs in the pipelines of the disaster management system, especially when the medium interacts between technical models and non-expert stakeholders.

The intersection of ML, DL, and LLMs thus offers an opportunity and a requirement: a necessity to come up with unified, data-driven, and physically aware models that exploit the advantages of contemporary AI and suppress the drawbacks of these models. The design of robust preprocessing and modeling pipelines is now becoming extremely important as the amount and variety of seismic data (waveforms, hydrochemical anomalies, accelerograms, geodetic measurements, climatic indicators, remote sensing measurements, and textual reports) keep on increasing, quickly and dramatically [8, 9, 10, 11, 12]. In this growing information environment, the problems of noise reduction, feature scaling, treatment of missing values, sensor time ranges, and hetereogeneous modalities fusion are not only technical specifics but also the factors predetermining the quality of downstream models and their stability. Badly curated or continuously preprocessed data can obscure faint precursory estimates, influence training model learning, and result in overconfident but unreliable forecasts, particularly in regimes with scanty training instances or with a high imbalance.

It is against this context that there is an increasing demand not only to have more powerful predictive architectures, but also to have intelligent frameworks that are sensitive enough to design, critique as well as to adapt preprocessing and modeling processes themselves. Their ability to encode methodological knowledge in seismology, statistics, machine learning, and reason over the semantic meaning of variables and metadata together with their capability to generate or refine analysis pipelines in response to changing data properties or research questions make LLLMs especially well-suited to perform such a role. This opens the opportunity of AI systems that are more than map input-output machine in the context of earthquake prediction and disaster management: they may choose suitable transformations, propose feature engineering plans, propose model families, and interpret their choice in a human comprehensible language. To bring this vision into reality, however, systematic analysis of the comparative capabilities of the various classes of models, such as classical ML, DL, and LLM-based reasoning systems, to leverage seismic and environmental information and act in a generalized fashion and survive in the real-world conditions is needed.

Within this expanded scientific and technological context, the current work is aimed at how the future of AI methods, and specifically, large language models can be used to improve the trustworthiness, interpretability, and feasibility of machine learning pipelines related to earthquakes, and in turn affect the future of earthquake prediction and disaster preparedness. In particular, we would like to place LLMs in a comparative context, in which their advantages and disadvantages are viewed in a complementary way with respect to traditional ML and DL models, particularly in the areas of complex data processing, multimodal learning, and explanation to experts. The synthesis of recent progress and a critical analysis of their assumptions, performance characteristics, and the challenges associated with their deployment, this paper will attempt to explain where the current AI-based solutions are, what are the limitations that need to be addressed, and

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

what research focus areas are most likely to yield successful and reliable, and useful AI-driven earthquake intelligence systems.

The rest of this paper is divided into three major sections, which expand upon the premises set in the introduction and are aimed at gradually reducing the focus of the issue to the narrow scope of theoretical progress and contemplation of the future. First, the section of the Literature Review presents a comparative and systematic review of the latest progress in machine learning, deep learning and large language model uses in earthquake tasks. This involves the research on seismic prediction, nowcasting, ground-motion observation, precursor and anomaly observation, social-media-enabled situational awareness, structural damage measurement, and intelligent decision support. Special emphasis is placed on the manner in which the various categories of models (classical ML, deep neural architecture, transformer based and LLM based) work with each different data modality (seismic catalogs, continuous waveforms, hydrochemical records, satellite observations, radon measurements, and textual or social media streams) and the performance metrics and evaluation protocols applied. This section reveals convergence patterns (e.g. the ubiquity of attention mechanism) and domain-specific changes (such as spherical convolutions or multimodal vision-language alignment) by categorizing the literature into thematic groups and prepositions a more organized comparison of capabilities and constraints across methods.

The second large part is the Discussion section that does not only follow the descriptive survey but a critical and integrative analysis of the reviewed studies. In this case we compare the strengths and weaknesses of the various AI paradigms against some cross-cutting dimensions: the ability of different paradigms to generalize across regions and tectonic environments, sensitivity to data quality and imbalance, robustness to noise and non-stationarity, the extent of physical interpretability, and the extent to which the various paradigms can be used in high-stakes operations. This part also directly previews the comparison nature of the work, comparing the performance of traditional ML models, deep neural networks, and the LLM-based systems on tackling similar classes of problems (e.g., magnitude prediction, nowcasting, precursor analysis or impact assessment), and the areas each set of methods performs best or worst. Moreover, we say something about the new, yet not well developed, purpose of LLMs to act as integrative reasoning systems that are capable of coordinating specialized predictive models, access and synthesize domain knowledge and even engage human specialists in natural language. The effects of these advances on the AI-driven seismology, real-time hazard surveillance, emergency logistics, and risk communication are discussed in depth, with a specific focus on the issues of trust, transparency, and regulation.

Lastly, the paper will conclude with the section of the literature review of the main findings derived by the paper and the discussion as a whole these findings into a logical collection of conclusions and research directions. This part also recaps the overall transformation of AI in terms of earthquake prediction and disaster management, as well as explains the limits of the existing functions and the potential dangers of overinterpretation. It then identifies specific directions of future research, such as the design of multimodal, physics informed architectures that can collectively utilise seismic, geodetic, environmental and socio-informational data; the creation of open, benchmark standards against which models and regions can fairly compare themselves; the addition of rigorous uncertainty quantification to AI based forecasts; and the development of notably less opaque LLMs as interpretable as well as domain-specific agents that engage with external tools and databases instead of being a black box. The section further highlights the importance of interdisciplinary cooperation and ethical standards to handle the fact that AI-driven seismic forecasting advances have to be transformed into reliable, fair, and workable systems capable of providing resiliency in societies against the threat of earthquakes.

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

2 Literature Review

Earthquakes are some of the most destructive natural hazards that have caused massive human losses, massive destruction of infrastructure and long-term socio-economic derailment. With the acceleration of urbanization and the increase in the interconnections of critical infrastructure networks, the need to have timely, accurate and actionable seismic intelligence has been exponentially expanded. Conventional methods of seismology, despite their robustness scientifically, have been in many cases unable to provide finely honed, operationally viable predictions or even to propagate fully the decision-making under uncertainty taken in real-time. Simultaneously, pattern recognition, knowledge extraction, and natural language understanding have been transformed in numerous fields, thanks to the development of artificial intelligence (AI) and deep learning, as well as large language models (LLMs). This has triggered a frenzied research in the field of intersecting earthquake science, disaster risk management, and AI.

The collection of works reviewed here demonstrates how AI and LLM-based methods can support the entire earthquake risk chain: from knowledge representation and emergency decision support, through social-media-based situational awareness and multimodal structural damage assessment, to seismic forecasting/nowcasting, precursor analysis, and resource planning. For clarity, we label the provided studies as LR01–LR18 and use these labels as citation keys. We organize the literature into thematic clusters: (i) LLM-based knowledge graphs and decision support for emergency management [13, 14, 15], (ii) social media analytics for rapid impact assessment [16, 17], (iii) multimodal LLMs for structural damage assessment [18], (iv) AI for emergency resource planning and logistics [15], (v) deep learning and transformer-based models for prediction and nowcasting [19, 20, 21, 22, 23, 24, 25, 26, 27], (vi) non-seismic precursors and remote sensing [28, 29], and (vii) explainability and bias in LLM applications to earthquake-related texts [30]. We conclude with a synthesis of research gaps and future directions.

A core difficulty in earthquake emergency management is the integration and semantic interpretation of heterogeneous information originating from official reports, sensor networks, historical archives, expert guidelines, and ad hoc case descriptions. The study in [13] directly addresses this challenge by proposing a domain knowledge extraction method based on a large language model augmented with a three-level prompt engineering system (TPES-LLM): instruction fine-tuning, demand awareness, and case matching. Instead of treating the LLM as a generic text generator, the authors deploy a local QWEN2.5-7B model using LangChain and adapt it to the earthquake domain via LoRA-based fine-tuning with expert classifications and relevant industry standards. This explicitly injects domain knowledge into the model, constraining its generative behavior and improving its ability to recognize earthquake-related entities, relations, and event structures.

One of the contributions that can be identified in the major study, [13], is that the multi-head attention is optimized based on the co-occurrence statistics of historical earthquake entities. The model is biased towards realistic patterns of relationships found in historical disasters by weighting the attention based on the empirically determined co-occurrence frequency. The demand-awareness step further enhances the extraction process to find the significant textual features and segments that have highest impact on the downstream knowledge extraction, which in effect directs the attention of the model to the information of high value. The system can also be trained on 36 known earthquake disaster events and learn latent patterns of association among entities (e.g., locations, magnitudes, casualties), relationships (e.g., cause-effect chains) and high-level events (e.g., cascading failures, secondary hazards).

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

To turn extracted knowledge into an operational decision-support tool, [13] introduces a bidirectional graph attention network (Bi-GAT) that enables information to propagate both ways across the graph and dynamically aggregates node features. This is complemented by a path confidence constraint algorithm (PCCA), designed to enforce plausible multi-step semantic connections between nodes, thereby capturing deep structural associations within the earthquake knowledge graph. Implemented using Neo4j, the resulting emergency knowledge graph is evaluated on several major real events, including the 2022 Luding M6.8, 2024 Jishishan M6.2, and 2025 Dingri M6.8 earthquakes. The reported intelligent question–answering accuracy (ranging from roughly 87% to over 90%) suggests that the system can reliably support query-driven emergency decision-making [13]. Importantly, this study illustrates an end-to-end pipeline: domain adaptation of LLMs, structure induction via graph learning, and practical application to real-world emergencies.

Although the context of [13] is narrowed down to earthquakes and the creation of emergency knowledge graphs, in the article by [14] the authors expand the range of thought by suggesting an AI-enhanced framework of combined natural disaster prevention and response, which is based on the DeepSeek LLM. The conceptualized three principal technical directions of LLMs application to geohazard scenarios by the authors include: (1) knowledge-graph-based dynamic risk modelling, (2) reinforcement-learning-based emergency decision system optimization, and (3) secure local deployment architectures. Unlike the entirely text-driven systems, the DeepSeek model is characterized as the one that uses a hybrid reasoning process that integrates both semantic-based textual input comprehension and geospatial patterns recognition, thus, allowing more comprehensive data assortments to be integrated, including historical disaster data, real-time IoT sensor measurements, and socio-environmental signals.

The framework in [14] is modular, where (a) is automated construction of domain specific knowledge graphs through unsupervised learning of physical relationships, (b) is scenario adaptive resource allocation by risk analysis through simulation, and (c) is the federated learning of decentralized but coordinated model updating in multi stakeholder settings. One of them is the focus on the data governance: the local deployment paradigm is approached to honoring privacy and security limitations, particularly in a cross-border or multi-jurisdictional environment, as well as in accordance with the FAIR (Findable, Accessible, Interoperable, Reusable) principles. In comparison to [13], which is more concerned with immediate emergency decision support provided by question -answering on a knowledge graph, at a larger scale, the perspective of LLM in [14] is that of a more broad, long-term geoscience-AI convergence, both in terms of prevention and response.

The use of social media (or microblogging services, in particular) provides the richest, real-time information sources at the time of earthquakes. Nevertheless, this data is infamously loud, it includes rumors, irrelevant content, and repeated posts. The contribution in the article of [16] is the first domain-specific LLM, QuakeBERT, which is directly adapted to the classification and filtering of earthquake-related microblogs to quickly evaluate their impact. The authors begin with a brief, yet useful taxonomy of categories of the physical impacts (e.g., reports of damage, casualties, infrastructure disruption) and social impacts (e.g., fear, panic, support requests). They subsequently create a dataset of 7282 microblogs of twenty different earthquakes in various places in order to have variety as far as the use of language, culture and events are concerned.

QuakeBERT is obtained by fine-tuning a transformer model on this curated dataset, and the study systematically analyzes how data diversity and data volume influence performance. The results show that both factors are critical: expanding the diversity of events and the volume of labeled data yields up to a 27% improvement in macro F1 scores [16]. When compared against traditional CNN- and RNN-based baselines,

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

QuakeBERT significantly outperforms them, raising the macro F1 from approximately 60.87% to 84.33%. These gains are especially important because classifying microblogs accurately into physically and socially relevant categories is a prerequisite for reliable downstream analyses, such as estimating damage severity, mapping affected areas, and measuring public sentiment.

On top of classification, [16] is an integrated approach (i) based on the analysis of the trend of public opinion to monitor the changes in the topic over time (ii) sentiment analysis to understand how the masses feel and (iii) quantification of physical impacts associated with a keyword to estimate how serious the physical effects are. The use on two similar earthquakes with similar magnitudes and focal depths shows that the suggested method can point at the variation in perceived impact and community response and note that magnitude is not a sufficient quality to describe the disaster magnitude. QuakeBERT allows more precise and on-time post-disaster estimations and policy-making through filtering noisy microblogs and concentrating on high-informing material to build resilient cities.

Beyond this LLC center collection, [17] discusses a framework of ML that utilizes Twitter data to respond to earthquakes, in which the main focus will be to derive geospatial information. The authors do not use a domain-specific transformer and rather use the traditional NLP techniques and machine learning classifiers to filter relevant tweets and get location references. The severity maps are then created by the extraction of the locations and using them to derive affected areas. Combining these maps into a WebGIS infrastructure, a visual interface is availed by [17] that can be utilized by emergency operators, government and non-governmental organizations and be useful to locate areas of foci and arrange resource allocation.

Even though the models behind the results in literature in [17] are simpler than the LLM-based QuakeBERT, the article demonstrates that even more traditional pipelines can be valuable in the operational setting, particularly when computation power is constrained or where no pre-trained domain-specific LLM can be found. Historically, the article being referenced as [17] is the predecessor of social-media-based situational awareness tools, and the article by [16] demonstrates the superior performance and the improved analytic capabilities that can be expected with the help of special LLMs. Combined with other works, these papers follow a path of moving away of rule-based or shallow-based or shallow learners to more semantically aware and disaster-related social media analysis models.

Quick and efficient evaluation of structural damage is crucial to emergency rescue missions as well as longer-term reconstruction. Conventional in situ damage detection depends on the visual inspection of teams of engineering experts, or on single-modal deep learning classifications of damage based on imagery only. Both methods are pragmatically problematic: the state inspection by experts is resource-consuming and time-intensive, whereas textualizing and context-driven engineering ideas might not be represented in purely visual models.

The SDA-Chat model of the article in [18] solves those issues, embracing the multimodal vision language paradigm premised on the visual question answering (VQA). Instead of merely classifying images by categories of damage, SDA-Chat participates in multi-round interactions of VQA, producing professional textual descriptions of structural damage properties that are present in the images. This design is capable of not just reporting that damage has happened, but also represents elaborated evaluation (e.g. describing cracking patterns, spalling, or mechanisms of collapse) that are more akin to engineering practice.

Methods Methodologically, to train on structured and expert-like answers, it is suggested in cite LR03 that the training is conducted in three stages, with the third stage being an instruction fine-tuning step, where

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

the model is trained to respond to domain-specific cues. One technical advancement is the cross-modality projector which employs dimension reshaping and parallel network structure to match visual characteristics (damage images) and linguistic characteristics (textual descriptions) in an efficient and accurate manner. The goal of this architecture is to overcome the famous bottleneck of multimodal alignment, in which inconsiderations between image and text representations can worsen VQA performance.

The data set created by the authors consists of 8195 pairs of image-text samples of structural damage, more than 8,000 of which have been curated to represent a broad range of damage types and severities. The evaluations of several of the advanced LLMs show that SDA-Chat is able to perform seven different tasks, such as classification, description, reasoning about the evolution of the damage, and possibly, suggesting what to do now. The most successful set up has a question answering accuracy of 83.04 %and a generation rate of 435.31 tokens/s, which means that the system is not only accurate but efficient enough to be used in field or near-real-time scenario [18]. Also variants of the models with a high degree of precision and the lightweight feature are developed, which allows implementation on other hardware platforms. The presented work demonstrates an example of how multimodal LLM can be used to convert the raw visual observation into the structured, decipherable, and high-professional textual evaluation.

Other than the extraction process and damage analysis, efficient disaster management must have powerful logistics and resource planning. Practically, the material requirements after the earthquake (e.g., medical supplies, shelters, food, water, rescue equipment) are diverse as they are affected by season, geography, population density, infrastructure and emergency protocols specifics. The current literature usually revolves around a few basic materials and heavily depends on the experience of experts and fixed assumptions, resulting in the discrepancy between the forecasted and actual requirements.

The article in [15] addresses this issue, presenting the reasoning-enhanced framework of the LLM used in conjunction with the Retrieval-Augmented Generation (RAG) to predict the demand of the material in case of an earthquake after the event. The authors develop an emergency knowledge base, which combines standardized response plans, past data of earthquake cases, and scenario features identified on digital sources. Using RAG, relevant documents and passages are retrieved in this knowledge base and then the reasoning of the LLM is based on domain-specific content that is proven. This prevents the propensity of generic LLMs to hallucinate, as well as makes predictions grounded in actual guidelines and empirical evidence.

In order to facilitate a more open and professional line of reasoning, Chain-of-Thought prompting presented in [15] is used, which persuades the model to explain the reasoning processes in between when generating demand forecasts. The framework generates differentiated material demand schemes which detail the type of the required supplies as well as the amount of demand measures per-capita considering the contextual conditions including season, local infrastructure, and population. One such feature is the dynamic updating mechanism of the system: with constant consumption of post-disaster data (online sources, e.g. updated number of victims, changing weather conditions, infrastructure damage reports, etc.) the model can adjust demand estimates on the fly.

Two methods of evaluation are expert evaluation based on the 2013 Ya'an Lushan earthquake and further simulated cases and actual implementation in the 7 January 2025 Dingri M6.8 earthquake. The findings have shown that the predictions made by the system are similar to the expert judgment and can be applied practically to help in the material allocation allocation of materials that are available to them at hand in practice [15]. This LLM+RAG framework is more flexible and explainable than the traditional optimization-based logistics models, further providing the simulation of decision making patterns of human experts.

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

In their predicting model of earthquakes presented in [19] the authors of the EPT model show an example of data-driven based model of earthquake prediction utilizing global seismic catalog and excluding the use of local historical data only. The authors claim that most prior schemes disregard the basic crustal and plate motions trends as they focus on localized data. EPT aims at identifying patterns of the world scale and transferring this information to regional mainshock prediction. The model to this effect employs gated feature extraction blocks (GFEB) whose role is to extract latent representations of crustal motion and plate dynamics based on historical catalogs throughout the world. These are in turn exploited to forecast mainshock in the individual provincial regions.

Another key component of [19] is the use of multi-headed self-attention to capture long-term dependencies in regional time series. Traditional LSTM networks struggle when tasked with learning over very long sequences due to vanishing gradients and limited memory capacity. By contrast, self-attention mechanisms can directly model interactions across distant time points, enabling the model to consider long-term precursory patterns. To tackle the common issue of imbalanced data—particularly the scarcity of large-magnitude events—[19] employs a gradient harmonization mechanism classification (GHMC) loss function, which reweights gradients to pay more attention to hard or underrepresented samples. Experiments on five provincial datasets in mainland China show that EPT can achieve accuracies exceeding 90% and improve prediction performance by up to 50% compared to baseline approaches [19]. These results underscore the value of combining global pattern mining, attention mechanisms, and loss-function engineering in seismic prediction tasks.

Earthquake nowcasting—estimating the changing probability of large events in a region over relatively short time horizons—has gained attention as an intermediate goal between long-term hazard assessment and elusive deterministic prediction. The QuakeGPT model in [20] adapts an attention-based science transformer to this task. Building on earlier work that used Receiver Operating Characteristic (ROC) methods and Shannon information analysis to quantify the information content of earthquake catalogs, [20] recognizes that the relatively short duration of reliable observational catalogs (spanning only a few decades) is insufficient for training large transformers. To overcome this, they rely on the ERAS (Earthquake Rescaled Aftershock Seismicity) simulation model, a simplified variant of ETAS models with only two adjustable parameters, to generate long synthetic catalogs for training.

QuakeGPT is trained on ERAS-generated catalogs and subsequently evaluated on an ERAS validation set. The model is assessed both for its ability to reproduce known patterns in the simulated data and for its potential applicability to real observed catalogs [20]. The authors report strong performance in capturing seismic patterns and promising, though not yet robust, results in near-future prediction experiments. This work illustrates both the opportunities and limitations of simulation-driven training: synthetic data can overcome observational scarcity, but domain shift between simulated and real seismicity remains a key obstacle.

In a related but more empirical study, [21] systematically evaluates a spectrum of deep learning architectures and pre-trained foundation models for earthquake nowcasting in Southern California. Here, the task is formulated as predicting the seismic energy release (in logarithmic scale) for the next 14 days in 0.1-degree spatial bins, using data from 1986 to 2024. The authors introduce two models: Multi Foundation Quake and GNNCoder. Multi Foundation Quake integrates outputs from pre-trained foundation models as auxiliary streams into a bespoke architecture, while GNNCoder emphasizes graph-based encoding of spatial relationships. Their experiments reveal that the performance of foundation models varies substantially depending on the characteristics of their pre-training datasets, emphasizing the importance of dataset selection and domain

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

adaptation. The Multi Foundation Quake approach, which fuses custom-designed patterns with foundation model predictions, achieves the best overall performance [21]. Together, [20] and [21] highlight the growing role of transformers and foundation models in seismic forecasting, while candidly acknowledging the challenges of generalization and interpretability.

Several works focus specifically on predicting earthquake magnitudes using deep neural networks and, in some cases, climate variables. The study in [22] posits that global warming and associated climatic variations (e.g., temperature, precipitation, snow cover) may influence seismicity through changes in stress and pore-fluid pressure. Using global temperature as a key climatic variable, along with eight mathematically derived seismic parameters, the authors train LSTM, Bi-LSTM, and transformer models to predict the magnitude of the next probable earthquake in three seismic regions: Japan, Indonesia, and the Hindu-Kush Karakoram Himalayan (HKKH) region. Their evaluation, based on MAE, MSE, log-cosh loss, and MSLE metrics, indicates that all models converge to low error values, suggesting that the combined climate–seismic feature space contains nontrivial predictive information [22]. This work is particularly notable for bringing climate data into the predictive framework, raising interesting questions about climate–seismic coupling mechanisms.

In [25], a similar triad of models (LSTM, Bi-LSTM, transformer) is applied to eight seismic indicators derived from earthquake catalogs for the same three regions, but without explicitly including climate variables. Using held-out test datasets, the authors show that model performance is region-dependent: LSTM achieves the best metrics for Japan (e.g., MAE = 0.060, MSE = 0.006), Bi-LSTM performs best for Indonesia, and the transformer excels for the HKKH region (e.g., MAE = 0.062, MSE = 0.006) [25]. These results imply that both the temporal structure of seismicity and the complexity of tectonic regimes may influence which architecture is most suitable.

Extending this line of research to the Horn of Africa, [26] formulates the prediction task as multivariate time series regression, forecasting magnitudes ($M \geq 3$) for the next three months. LSTM, BiLSTM, BiLSTM with attention (BiLSTM-AT), and transformers are compared, with transformers again obtaining the best results across MAE, MSE, RMSE, and MAPE metrics (e.g., MAE = 0.276, MSE = 0.147) [26]. This suggests that transformers are particularly effective in regions with complex or sparse seismicity patterns, where long-range dependencies and irregular event occurrences are prevalent.

From a different regional perspective, [24] focuses on the North Anatolian Fault Zone (NAFZ) in Turkey between 2007 and 2010 and compares multi-layer perceptrons (MLPs), standard LSTMs, and attention encoder–decoder LSTM models. The attention-based model shows markedly superior performance, reflecting the value of selectively focusing on the most informative segments of the input sequence when modeling complex seismic patterns [24]. In summary, [22, 24, 25, 26] collectively demonstrate that deep neural networks—especially those incorporating attention—are capable of achieving low prediction errors for regional magnitude prediction, at least for moderate magnitudes, while leaving open the question of reliability for rare, high-impact events.

Most of the predictive models discussed thus far operate on catalog-based features (e.g., magnitudes, inter-event times, spatial bins). In contrast, [23] directly analyzes continuous seismic waveforms using the Wav2Vec 2.0 self-supervised framework originally developed for speech recognition. The authors consider the 2018 caldera collapse at Kīlauea volcano in Hawai‘i and treat seismic waveforms as an analog to audio

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

signals. By pre-training Wav2Vec 2.0 on caldera seismic waveforms and augmenting the architecture to predict contemporaneous surface displacement (a proxy for fault slip), they show that the model can accurately infer displacement solely from seismic recordings [23].

The results firmly support the notion that earthquake faults emit seismic signatures that encode information about their evolving state, in a manner comparable to how speech encodes linguistic information. When the model is adapted to near-future prediction tasks, it demonstrates hints of predictive ability, though not yet at a level suitable for operational deployment. Nonetheless, [23] highlights the promise of transferring advances in self-supervised audio representation learning to earthquake science and fault monitoring, opening the door to new techniques for probing fault behavior in near real time.

Finally, [27] examines earthquakes in Los Angeles, California, using a diverse set of machine learning and neural network models to predict the maximum potential earthquake category within the next 30 days. The authors construct a detailed feature matrix that aggregates multiple seismic indicators and other predictive features derived from the literature, then evaluate sixteen algorithms. Among these, the Random Forest model provides the best balance of accuracy and robustness, achieving high performance in predicting the maximum earthquake category in short-term windows [27]. This work underscores that, in some scenarios, well-tuned ensemble methods with carefully engineered features can rival or complement deep learning architectures, particularly when data volumes are moderate and interpretability is a concern.

In addition to seismic and social media data, remote sensing is a great source of possible earthquake precursors. The survey of satellite-based earthquake prediction is presented in a general sense in the review of [28]. It observes that initial efforts to measure precursors with in situ ground measurements (e.g., of temperature, of gas emissions, of ionospheric parameters) had been impeded by low spatial resolution, low time resolution, and high prices with limited advancement and inconclusive results. As remote sensing satellites have been invented and launched in large numbers, though, the number and types of possible precursory signals have increased exponentially.

As noted in [28] numerous statistical studies have reported such anomalous response in physical and chemical parameters as thermal anomalies, ionospheric disturbances and change of atmospheric composition in the time window of about 1-30 days prior to the occurrence of strong earthquakes. The authors state that the future of satellite based earthquake prediction is in multi-precursor analysis, whereby multiple independent signals are analyzed together with classic and intelligent anomaly detection algorithms. They stress the value of multi-method analysis (e.g. a combination of statistical methods, machine learning and physics-based methods) and fusion systems, which can coherently combine varying precursor data into coherent alerts or warning levels.

Current technological advances like cloud-based geospatial data services and smart interfaces are identified as some of the enablers that can democratize the access to complex remote sensing data and analysis services. In general, the article by [28] is a rather conservative viewpoint on the topic: the reliable low-uncertainty earthquake prediction is still out of reach, but the conjunction of multi-precursor satellite data, enhanced anomaly detection, fusion frameworks, and enhanced computational infrastructure can lead to the real warning systems.

Another class of possible precursors involves the radioactive gas radon (^{222}Rn), whose concentration in soil, water, and air can be affected by crustal deformation and fluid flow. The study in [29] revisits radon

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

anomalies as earthquake precursors, focusing on atmospheric radon measurements from Kobe Pharmaceutical University (KPU) prior to the 1995 Kobe earthquake and ionization currents at Fukushima Medical University (FMU) before the 2011 Tohoku-oki earthquake. Rather than relying on subjective parameter choices for anomaly detection, [29] introduces an objective, data-driven methodology using Random Forest (RF) regression.

The approach consists of modeling the typical annual pattern of atmospheric radon concentration using RF, which learns from historical data to produce daily or seasonal baseline predictions. Anomalies are then defined as instances where the difference between observed and predicted values exceeds three times the standard deviation, a standard statistical threshold for outliers. Applying this method to the KPU and FMU datasets, [29] finds that significant anomalies occurred prior to both major earthquakes, with deviations surpassing the three-sigma threshold. These findings suggest that RF-based modeling of radon time series can offer a more objective and statistically grounded framework for precursor analysis than traditional ad hoc methods. While not sufficient for stand-alone prediction, radon-based anomaly detection could form part of a multi-precursor warning system in combination with other indicators.

While the majority of the reviewed studies focus on improving predictive performance or operational utility, [30] takes a different but equally important perspective by examining bias and explainability in LLM-based processing of historical earthquake-related documents. Using multilingual newspaper coverage of the 1908 Messina earthquake (German, English, and French sources) as a case study, the authors develop an explainability-driven framework to evaluate model design bias in multilingual news extraction.

By systematically evaluating six state-of-the-art models, [30] identifies three recurrent bias patterns. First, contextual integration bias refers to the way models integrate local information into broader event narratives, potentially overemphasizing certain contexts and underrepresenting others. Second, overconfidence bias arises when models assign high confidence to extracted information even when the underlying data is ambiguous or conflicting. Third, preference bias reflects systematic tendencies to favor specific linguistic forms, stylistic expressions, or source types. Notably, the study concludes that these biases are more strongly linked to alignment and fine-tuning procedures (e.g., reinforcement learning from human feedback, prompt design) than to limitations in the underlying training corpus. This insight has broad implications: it suggests that the deployment and adaptation phase, rather than pre-training alone, is a critical locus for introducing or amplifying bias.

The methodological framework in [30] thus provides valuable guidance for the responsible deployment of LLMs in digital humanities and, by extension, in disaster risk communication and historical analyses of past events. In earthquake-related contexts, biased or overconfident model outputs could distort historical narratives, misrepresent affected populations, or amplify particular viewpoints, with ethical and societal consequences. Consequently, models used in operational settings should be subjected to rigorous bias and reliability assessments in addition to traditional performance evaluations.

Taken together, the works [13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30] illustrate a broad and rapidly evolving landscape of AI applications in earthquake science and disaster management. At the information and decision-support layer, LLM-driven knowledge extraction and graph construction [13], integrated disaster-prevention architectures [14], and RAG-based reasoning for logistics planning [15] demonstrate that LLMs can serve as powerful engines for organizing domain knowledge, answering complex queries, and simulating expert decision-making. Social media analysis has advanced from traditional

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

NLP-based location extraction and mapping [17] to sophisticated domain-specific LLMs that jointly capture physical impacts and public sentiment [16]. Multimodal models such as SDA-Chat [18] show how vision–language systems can transform raw images into rich engineering narratives, supporting both rapid triage and detailed structural assessment.

On the physical and predictive side, deep learning models span a continuum from globally informed, attention-based catalog models such as EPT [19] to transformer-based nowcasting systems trained on simulations [20] and multi-stream foundation model architectures [21]. Regional magnitude prediction has been explored using LSTMs, Bi-LSTMs, transformers, and attention encoder–decoder architectures across diverse tectonic environments [22, 24, 25, 26, 27], with consistently low error metrics for moderate magnitude ranges. Waveform-based approaches leveraging self-supervised audio models like Wav2Vec 2.0 [23] highlight a promising avenue for continuously monitoring fault states and potentially detecting subtle changes before major events. Complementary non-seismic precursors—satellite-observed anomalies [28] and radon concentration fluctuations [29]—expand the data landscape further, suggesting that truly robust prediction or early warning systems will likely require the integration of multiple, heterogeneous data streams.

Despite these substantial advances, several fundamental challenges and open research questions remain:

- **Generalization and Transferability.** Many models are trained and evaluated on specific regions or datasets, and their performance under cross-regional transfer or in previously unseen tectonic settings is not fully characterized. The extent to which models such as EPT [19], QuakeGPT [20], or Multi Foundation Quake [21] generalize beyond their training regimes remains an important topic for future work.
- **Data Scarcity for Rare, Extreme Events.** The predictive performance reported in [22, 24, 25, 26, 27] is generally strongest for moderate magnitudes (e.g., $3.5 \leq M \leq 6.0$). However, the events of greatest societal concern—large, rare earthquakes—are intrinsically underrepresented in historical catalogs. Simulation-based approaches [20] can partially mitigate this, but robust modeling of tail behavior and extreme events remains challenging.
- **Physical Interpretability and Constraints.** Most deep learning and LLM-based models operate as black boxes, raising concerns about physical interpretability and trustworthiness. Incorporating physical constraints, domain knowledge, or hybrid physics–ML architectures could improve reliability and make model outputs more interpretable to seismologists and engineers. The physically motivated feature engineering in [19, 22, 24, 25, 26, 27] and the emphasis on realistic simulation in [20] are important steps in this direction, but more systematic frameworks are needed.
- **Multi-source Data Fusion.** Several works hint at the need to combine diverse data sources—seismic catalogs, continuous waveforms, climate records, remote sensing, social media, radon measurements—yet fully integrated, end-to-end systems are still rare. Future research could explore multimodal architectures that simultaneously ingest and fuse these data types, extending the multimodal paradigm of SDA-Chat [18] to broader earthquake prediction, monitoring, and response tasks.
- **Uncertainty Quantification and Risk Communication.** For operational use, especially in early warning or nowcasting, it is not enough for models to output point predictions or class labels; they must also provide calibrated uncertainty estimates and support risk-based decision-making. ROC-based evaluations and information-theoretic analyses [20] are valuable, but more work is needed on

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

probabilistic forecasting, ensemble methods, and robust uncertainty quantification for AI-based earthquake models.

- **Ethics, Bias, and Governance.** The bias analysis in [30] underscores that LLMs can exhibit nontrivial design biases even when trained on high-quality data. Similar issues may arise in other contexts, such as social media analysis [16, 17] or historical document processing. Ensuring that AI systems used in disaster management do not exacerbate inequalities, misrepresent marginalized groups, or propagate misinformation is a central ethical challenge. Governance frameworks that address data sovereignty, privacy, and accountability are equally important, especially for systems like those proposed in [14, 15] that operate in cross-border or multi-institutional environments.

Finally, based on the analyzed literature, it is possible to note that AI and methods based on LLM are starting to infiltrate almost all areas of the earthquake risk management cycle: the interpretation and forecasting of seismic events, the assessment of the consequences, the management of resources, and the study of history. The further development will probably depend on the close cooperation between AI researchers, seismologists, structural engineers, social scientists, and policy-makers; on the responsible combination of data-driven and physics-driven models; and on the long-term focus on such aspects as transparency, reliability, and implications to society. Provided the challenges can be resolved, AI can also become one of the cornerstones of next-generation earthquake resilience and disaster risk reduction systems.

In order to provide an overview of the broad range of literature reviewed in the literature review section, Tables 1 and 2 give a tabular summary of the key studies in machine learning, deep learning, and the application of large language models to earthquake prediction, monitoring and disaster management. Considering the scope and variety of what is currently available, such as knowledge graph construction and social-media-based situational awareness, deep seismic forecasting models, multimodal structural assessment systems, and multi-precursor anomaly detection, the given tabulated summaries allow bringing much more straightforward comparison between methodological, research goal and main findings. The tables contribute to the discovery of the thematic patterns, accentuating the innovativeness of methods, as well as displaying the changing role of artificial intelligence in seismological studies, by grouping the studies in the similarity of their evaluative dimensions.

To create better readability and clarity, the summarized works are spread over the two tables. Table 2 lists the works of researchers including the following: [13] through [21] most of which concentrate on the knowledge extraction based on LLM, emergency decision-support architecture, social media analytics, multimodal structural damage detection, and modern deep learning-based models of regional and global earthquakes. Table 2 includes studies by [30] to[29] on bias and explainability in LLM applications, climate-seismic coupling, regional magnitude prediction by advanced neural architectures, waveform-based displacement inference, ensemble learning to short-term hazards estimation, and non-seismic precursor detection like satellite anomalies and atmospheric radon. Collectively, these tables provide a coherent and comparative understanding that supplements the narrative review of the chance that sheds light on the merits as well as the shortcomings of the current AI-based methodologies of seismic hazard and disaster preparedness.

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

Table 1: Summary of AI-Driven Earthquake Studies (Part I)

Ref.	Objective	Methodology	Key Findings
[13]	Construct an earthquake emergency knowledge graph.	Domain-adapted QWEN2.5-7B, three-level TPES-LLM prompting, LoRA fine-tuning, Bi-GAT reasoning, Neo4j graph storage.	Produces high-accuracy emergency Q&A and reveals strong entity–relation extraction for operational decision support.
[14]	Develop an AI-assisted disaster-prevention and emergency-response framework.	DeepSeek LLM, dynamic risk modeling, RL-based decision optimization, federated learning, FAIR-aligned local deployment.	Shows LLMs can guide long-term geohazard risk management, resource allocation, and data-secure multi-agency coordination.
[15]	Predict post-earthquake emergency material demand.	RAG-enhanced LLM, emergency knowledge base, Chain-of-Thought prompting, dynamic updates from real-time disaster info.	Generates accurate, expert-like material-demand schemes validated in real earthquake deployments.
[16]	Classify and assess earthquake-related microblogs.	QuakeBERT, transformer fine-tuning on 7,282 posts across 20 events; sentiment and trend analysis pipeline.	Achieves large gains in F1 score, filters noise effectively, and improves physical and social impact assessment.
[17]	Provide geospatial situational awareness from Twitter.	Classical NLP + ML classifiers for tweet filtering and geolocation; severity mapping in WebGIS.	Delivers useful real-time damage localization where LLMs are not available.
[18]	Automate structural damage assessment.	SDA-Chat multimodal VQA framework, instruction tuning, cross-modality projector, 8,195 image–text pairs.	Generates engineering-level descriptions with high QA accuracy and fast inference.
[19]	Improve regional mainshock prediction using global seismic patterns.	EPT model with gated feature extraction, multi-head attention, GHMC loss; trained on global catalogs.	Achieves ¿90% accuracy and outperforms baselines by up to 50%.
[20]	Develop a transformer-based earthquake nowcasting model.	QuakeGPT trained on ERAS synthetic catalogs; evaluated on simulation and real data.	Captures key seismic patterns; demonstrates early predictive potential despite simulation–reality domain shift.
[21]	Benchmark deep architectures and foundation models for regional nowcasting.	Multi Foundation Quake (fused foundation models) and GNNCoder; prediction of 14-day seismic energy in S. California.	Fused models outperform single architectures; performance strongly depends on pre-training data.

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

Table 2: Summary of AI-Driven Earthquake Studies (Part II)

Ref.	Objective	Methodology	Key Findings
[30]	Assess bias and explainability in LLM extraction of historical earthquake news.	Evaluation of six multilingual models; analysis of contextual, overconfidence, and preference bias.	Finds that alignment and fine-tuning steps introduce major biases; highlights ethical risks in disaster communication.
[22]	Examine climate–seismic coupling for magnitude prediction.	LSTM, Bi-LSTM, transformer trained on temperature + eight seismic indicators across three regions.	Models converge to low errors, suggesting predictive value in climate–seismic feature interactions.
[25]	Compare DL models for regional magnitude prediction.	LSTM, Bi-LSTM, transformer applied to eight catalog-derived seismic indicators.	Optimal architecture varies by tectonic region; transformer excels in complex settings.
[26]	Forecast magnitude in the Horn of Africa using multivariate time series.	LSTM, BiLSTM, attention BiLSTM-AT, and transformer.	Transformer outperforms all baselines across MAE, MSE, RMSE, and MAPE.
[24]	Evaluate AI models for the North Anatolian Fault Zone.	Comparison of MLP, LSTM, and attention encoder–decoder LSTM.	Attention-based model provides superior performance for complex temporal patterns.
[23]	Infer displacement and near-future behavior from waveform data.	Adapted Wav2Vec 2.0 self-supervised audio model; displacement prediction from waveforms.	Accurately estimates displacement and shows potential for early fault-state monitoring.
[27]	Predict maximum earthquake category in Los Angeles.	Evaluation of 16 ML/NN algorithms on engineered features; emphasis on Random Forest.	Random Forest gives best accuracy and robustness for 30-day category prediction.
[28]	Review satellite-based multi-precursor earthquake prediction.	Survey of thermal, atmospheric, and ionospheric anomalies; fusion and cloud-based tools.	Highlights importance of multi-precursor systems and modern geospatial infrastructures.
[29]	Analyze atmospheric radon anomalies before major earthquakes.	Random Forest regression to model annual radon cycles; anomaly detection via 3-sigma threshold.	Identifies significant pre-event anomalies, supporting radon as a contributing precursor.

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

3 Discussion

The body of work reviewed in this paper collectively demonstrates that machine learning, deep learning, and, increasingly, large language models have moved from peripheral tools to central components of modern earthquake science. Across applications ranging from hydrochemical anomaly interpretation [8] and catalog enrichment [16, 17] to global-scale forecasting [10, 19, 20, 21] and ground-motion prediction [11], data-driven models consistently outperform traditional statistical baselines or physics-free empirical relations. At the same time, these models are not replacements for classical seismology but rather extensions of it: they leverage decades of observational data and physical insight, yet they operate in regimes—high-dimensional feature spaces, multimodal fusion, and complex temporal dependencies—where conventional techniques are limited. The integration of ML and DL into the seismological workflow has thus shifted the research emphasis from purely parametric hazard characterization toward rich, data-centric representations of seismic processes and impacts [9, 19, 22, 25]. This shift has opened new avenues for prediction and nowcasting, but it has also revealed structural weaknesses related to data quality, generalization, interpretability, and operational readiness.

A central theme emerging from the reviewed studies is the tension between impressive local or regional performance and uncertain generalization across space, time, and magnitude ranges. Many models report strong accuracy or low error metrics on specific datasets—such as EPT's high accuracy on five Chinese provincial catalogs [19], the transformer-based models' performance in Southern California nowcasting [21], or the low magnitude-prediction errors achieved in Japan, Indonesia, and the HKKH region [22, 25, 26]. Similarly, SeisEML achieves roughly half the MAE and RMSE of conventional attenuation relations on Japanese and Iranian PGA datasets [11], and SARIMAX models show that interpretable time-series structures can be captured when exogenous variables are carefully chosen [12]. Yet, the very conditions under which these models excel—well-characterized regions, specific magnitude ranges (often $3.5 \leq M \leq 6.0$), and curated datasets—limit their proven applicability to rare, extreme events and to tectonic settings beyond their training domain [19, 20, 21, 27]. Simulation-driven training, as in QuakeGPT with ERAS catalogs [20], and cross-region experiments such as those in SeisEML [11], represent important steps toward broader generalization, but they also highlight the persistent problem of domain shift between training and deployment environments.

Data quality and heterogeneity are presented as equally serious performance and reliability limitations of the model. Seismic catalogs are never complete or noisy; waveform data is of poor quality, site conditions and environmental noise can be detected and unclear precursor signals include hydrochemical anomalies, radon levels, thermal or ionospheric anomalies detected by satellites and so on are sparse, noisy, and ambiguous [8, 28, 29]. Indicatively, the fluid-chemistry-based prediction framework in [8] demonstrates that predictive abilities of the ML models are very sensitive to the springs and indicator type selected, and that the model parameters of anomaly detection thresholds and response-time windows need close tuning. The anomaly detection experiment with radon in [29] shows that despite the high level of sophistication in defining baselines, such as the use of Random Forests, an anomaly needs to be analyzed carefully and preferably alongside other signals. Surveys of catalog development and ground-motion modeling point out that data imbalance (especially, the dearth of large earthquakes and strong-motion recordings) can bias models and requires special loss functions or reweighting measures [9, 19, 21]. These problems support the importance of standardized and high-quality datasets and benchmarks, and generating systematic investigations into the influence of data curation, preprocessing, and labeling choices on downstream predictions.

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

The role of architectural decisions in the effectiveness with which models utilize the available information is decisive as well as the consideration that the underlying physics and geometry of the problem. The geometric distortion of using Euclidean convolutions on global latitude-longitude grids directly addressed by the spherical ConvLSTM-U-Net architecture in [10] shows concrete improvements in precision and recall of global-scale prediction. On the same note, other models such as EPT [19] use gated feature extraction and multi-headed self-attention to learn long-term, global-scale dependencies that would be hard to learn by standard LSTMs. The architectures Multimodal Multimedia Multimodal architectures like SDA-Chat that cross-modality projectors connect visual and textual representations demonstrate that the combination of various types of data can significantly benefit structural damage assessment [18]. Self-supervised models that are derived using speech processing, e.g. Wav2Vec 2.0 tested on continuous seismic waveforms [23] show that models which are trained to learn other tasks can also be trained to learn rich latent representations of fault state under suitable adaptation. All these studies point to the fact that selecting architecture which is geometry-aware, modality-aware and architecture which is able to model multi-scale dependencies is not simply an engineering feature but a scientific decision which can radically change the patterns which can be learnt on the basis of seismic data.

Massive language models take a unique role in this landscape in that they are more than predictive engines, but also reasoning, integration and communication layers that overlay or co-exist with specialized ML and DL models. In other areas like emergency knowledge graph construction [13], RAG-based logistics planning [15], and integrated geohazard management frameworks [14] the LLMs are orchestrators; ingesting heterogeneous inputs, recalling pertinent domain documents, and producing structured outputs; graph updates to material demand plans, and approximate expert reasoning. Examples of domain-specific models such as QuakeBERT indicate that with a LLM that has been fine-tuned on non-noise microblog collections, it is possible to filter noise, classify physical and social effects, and feed rich situational awareness pipelines at the heart of these tools [16]. Simultaneously, the bias analysis in [30] keeps in mind that LLMs can suffer the non-trivial contextual integration, overconfidence, and preference bias which are especially common when using multilingual historical sources or when affected by their alignment procedures with distorting effects on the behavior. These results indicate that although LLMs have great potential in semantic integration and explanation throughout the earthquake risk chain, they still need to be adapted, based on domain data, and thoroughly test under bias and reliability standards before they can be trusted in operational application.

For the methodology, the studied literature indicates the need to transition to integrated, multimodal and physics-informed models, rather than isolated and single-purpose models. There is a large body of literature that implicitly recommends a combination of various data types of hydrochemical indicators and accelerograms and catalogs and waveforms and geodetic measurements and satellite observations and radon time series and social media streams but complete systems that consume these modalities together are infrequent [8, 9, 11, 12, 18, 23, 28, 29]. The effectiveness of SDA-Chat in multimodal structural damage assessment [18], the proven usefulness of exogenous variables in SARIMAX models, and the usefulness of incorporating synthetic and real data in QuakeGPT all are indicative of the advantages of integrating complementary channels of information. These architectures (by embedding physical constraints either with physics-inspired loss functions or hybrid mechanistic/ML models, or physically modeling stress transfer and fault geometry) may be also made more interpretable and robust. Concurrently, uncertainty quantification and risk communication should be made components of model design especially where nowcasting and early-warning models need to be made where decisions depend on probabilistic predictions [20, 21, 27]. Even technically correct models can not help to make effective decisions without clear communication plans and explicit uncertainty estimates.

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

Finally, the translation of these AI advances into practice raises cross-cutting issues of governance, ethics, and interdisciplinary collaboration. Models deployed in real-time hazard monitoring, emergency logistics, or public communication must operate within institutional, legal, and societal constraints that are only partially captured in technical metrics. Frameworks like the federated, privacy-preserving architectures discussed in [14] illustrate how data governance principles such as FAIR can be woven into the design of AI systems, but many practical questions remain: Who owns and maintains the models? How are responsibilities allocated when predictions are wrong? How are biases monitored and mitigated over time, especially as data distributions shift? Addressing these questions will require closer collaboration between AI researchers, seismologists, structural engineers, social scientists, emergency managers, and policy-makers. The literature reviewed here shows that AI and LLM-based methods can significantly enhance our ability to understand and manage earthquake risk; the challenge now is to embed these methods in robust, transparent, and socially responsible frameworks that can support real-world decision-making before, during, and after seismic crises.

4 Conclusion and Future Work

This paper discusses how machine learning, deep learning, and large language models have become increasingly central to earthquake science, focusing on their use for prediction, ground-motion estimation, precursor analysis, and decision support. Across a broad range of literature, the data-driven approach has demonstrated high performance when conventional statistical models and empirical relations have failed. They have enhanced the identification of faint patterns in seismic catalogs, the better exploitation of non-seismic precursors (e.g., hydrochemical anomalies and radon concentrations), and peak ground acceleration prediction, and have made time-series forecasting with exogenous variables more expressive. Meanwhile, such developments have also demonstrated the underlying issues with the idea of generalization, data quality, interpretability, and operational deployment to the extent that AI-based models are not neutralized systems but rather components of a larger scientific and decision-making ecosystem. Among the fundamental findings, the most significant performance improvement is observed in well-characterized, domain-specific environments. Models for regional prediction have been trained on longer, higher-quality catalogs, achieving astounding accuracy in predicting magnitudes within a specific range. Deep models on a global scale, explicitly learning the Earth's geometry, are more accurate and improve spatiotemporal forecast recall. Ensemble ground-motion prediction methods have been shown to outperform traditional attenuation relations across various datasets. In contrast, time-series models that include exogenous variables have revealed recurrent timing patterns and the effects of the geophysical driver. These achievements indicate that when models are well-adapted to the data regime and physical setting, machine learning and deep learning can enhance the performance of classical seismological techniques. Nonetheless, they also show that these models have constraints, as performance can deteriorate on untrained distributions of tectonic setting, on undersampled magnitude ranges, or under changing observational states. Information quality and inconsistency appear throughout as underlying determinants of model behavior. Seismic catalogs have varying completeness, completeness magnitude, and noise levels; waveform data are prone to instrumentation variations and environmental interference; precursor measurements, whether chemical, geodetic, or atmospheric, are often sparse and noisy; and social media streams contain a lot of irrelevant or misleading information. Models trained on such data inherit these flaws. Predictive performance based on hydrochemical data, for example, is sensitive to the springs being monitored, the indicators used, the definition of anomalies, and the time window over which responses are checked. In the same manner, radon precursor analyses should be able

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

to differentiate between honest crustal reactions and seasonal or meteorological effects. Given class imbalance in ground-motion and catalog-based predictions, which require significant events, it is necessary to use specialized loss functions, resampling techniques, or ensemble designs. These facts underscore the need to improve algorithm design with a strict focus on data curation, standardization, and documentation. Another determinant of the success of AI-based earthquake models is their design. Geometrical convolutional networks that leverage recurrent networks have enabled geometry-aware architectures, which are required to model global seismicity without introducing distortions caused by naive projections. Attention-based models have enhanced the ability to learn long-range temporal dependencies in seismic sequences, overcoming the expected limitations of recurrent networks. Using vision and language, we have encoded raw structural damage images into rich, professional textual judgments, demonstrating the usefulness of a heterogeneous combination of data types. Self-trained encoders trained in other applications (including speech processing) have shown that continuous seismic waveforms contain much more information about fault state and deformation than is typically found in catalog summaries. As these examples demonstrate, the selection of neural architecture is not a purely computational issue but a scientific one, determining which properties of the underlying physics can be represented and learned effectively. In this landscape, large language models hold a special place and play an ever-growing role. Instead of being mainly numerical predictors, LLMs are integrators, organizers and interpreters of knowledge. They can read and organize unstructured reports, build and query knowledge graphs, make sense of social media content to build situational awareness, drive resource-allocation decisions using retrieval-augmented reasoning, and provide natural-language explanations of detailed model results. Their heterogeneous-thinking capacity, which extends to include technical documents and sensor descriptions, as well as human narratives and policy prescriptions, places them in a high-level orchestration role that can sit above technical seismic and geophysical modelling. Nonetheless, the new risks are also presented by the LLMs. As they are usually trained on general-purpose text corpora, they lack innate support for geophysical laws. When tasked with reasoning about seismic processes, they can produce plausible but erroneous statements. They may introduce biases from alignment or fine-tuning, inconsistently treat numerical information, and, in some cases, exaggerate their confidence. These problems indicate that, in high-stakes settings like earthquake early warning, hazard communication, or strategic planning, domain-adapted LLMs need to be evaluated and implemented rigorously within well-thought-out guardrails. In future research, there are numerous directions to take, and prospects are encouraging. To start with, a more urgent demand is for standard, open, and multimodal benchmark datasets that cover different regions, tectonic environments, magnitude scales, and observation types. These benchmarks are expected to cover not only seismic catalogs and waveforms but also hydrochemical parameter readings, radon time series, satellite-based parameters, geodetic time series, and socio-informational measurements such as social media reports and textual reports. This would provide a systematic, reproducible comparison of the architectures of models, training procedures, and evaluation metrics, and would assist in understanding the marginal benefit of including each data modality. Second, explicitly incorporating physical knowledge into AI models is a significant opportunity. Physics-informed neural networks, which combine mechanistic simulators with learned components, hybrid models based on fault geometry and stress transfer relations, and models based on fault geometry and stress transfer relations, may also help constrain the models, enhance interpretability, and guard against unphysical predictions. The other necessary frontier is uncertainty quantification. In most of the tasks under discussion, such as short-term prediction, nowcasting, early warning, and resource planning, the usefulness of a model, as well as its predictive performance, is also determined by how well it describes uncertainty in an objectively, decision-relevant manner. Further studies are advised to incorporate further probabilistic forecasting models, ensemble models, Bayesian deep learning models, and the information-theoretic metrics into models and model evaluation. These software tools help convey risk in messages that are practical for emergency managers and policymakers, rather than making predictions sound

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

true. Simultaneously, the topic of large language models as tool-using agents warrants further discussion. The LLMs might be tasked with invoking specific seismic or geophysical models, accessing the appropriate scientific literature, comparing their results against physical constraints, and generating explanations for a variety of audiences, including technical specialists and laypeople. Lastly, these technical advances require strong governance, interdisciplinary collaboration, and sustained stakeholder engagement to translate them into benefits for society. Earthquake models based on AI do not exist in a vacuum, but they live within institutional cultures, regulatory frameworks, and systems of communication within the community. Questions of responsibility, transparency, and accountability arise in predictions that drive high-consequence choices. Further effort in this area should include not only algorithmic innovation but also co-design with seismologists, engineers, planners, and emergency managers, as well as with at-risk communities. This involves creating interfaces that ensure model behavior and constraints are transparent, protocols for validating and revising models in the future, and ethical guidelines for data use and model implementation. When taken together, these endeavors can help ensure that machine learning, deep learning, and large language models not only advance academic research but also build more resilient, better-equipped societies in the face of earthquake hazards.

References

[1] Guofu Luo, Yingcai Xu, Hengzhi Luo, Wenjun Li, and Bingzheng Hou. Spatiotemporal characteristics of the energy field and disaster cause of the 2023 gansu jishishan m 6.2 earthquake. *Geomatics, Natural Hazards and Risk*, 16(1):2569822, 2025.

[2] Attila Gergely, Tamás Sándor Biró, Ferenc Járai-Szabó, and Zoltán Néda. Statistics of earthquakes based on the extended lggr model. *Physica A: Statistical Mechanics and its Applications*, 650:129983, 2024.

[3] Xiangli He, Zhaoning Chen, Qing Yang, and Chong Xu. Advances in earthquake and cascading disasters. *Natural Hazards Research*, 5(2):421–431, 2025.

[4] Wangxin Zhang, Jianian Wen, Huihui Dong, Qiang Han, and Xiuli Du. Post-earthquake functionality and resilience prediction of bridge networks based on data-driven machine learning method. *Soil Dynamics and Earthquake Engineering*, 190:109127, 2025.

[5] Khairul Adib Yusof, Syamsiah Mashohor, Mardina Abdullah, Mohd Amiruddin Abd Rahman, Nurul Shazana Abdul Hamid, Kasyful Qaedi, Khamirul Amin Matori, and Masashi Hayakawa. Earthquake prediction model based on geomagnetic field data using automated machine learning. *IEEE Geoscience and Remote Sensing Letters*, 21:1–5, 2024.

[6] Ying Zhang, Chengxiang Zhan, Qinghua Huang, and Didier Sornette. Seismically informed reference models enhance ai-based earthquake prediction systems. *Journal of Geophysical Research: Solid Earth*, 129(3):e2023JB028037, 2024.

[7] Qiyue Wang, Yekun Zhang, Jiaqi Zhang, Zekang Zhao, and Xijun He. On the use of vmd-lstm neural network for approximate earthquake prediction. *Natural Hazards*, 120(14):13351–13367, 2024.

[8] Ruijie Zhu, Fengtian Yang, Xiaocheng Zhou, Jiao Tian, Yongxian Zhang, Miao He, Jingchao Li, Jinyuan Dong, and Ying Li. Anomaly detection using machine learning in hydrochemical data from hot

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

springs: Implications for earthquake prediction. *Water Resources Research*, 60(6):e2023WR034748, 2024.

[9] Hisahiko Kubo, Makoto Naoi, and Masayuki Kano. Recent advances in earthquake seismology using machine learning. *Earth, Planets and Space*, 76(1):36, 2024.

[10] Zhongchang Zhang and Yubing Wang. A global earthquake prediction model based on spherical convolutional lstm. *IEEE Transactions on Geoscience and Remote Sensing*, 62:1–10, 2024.

[11] Anushka Joshi, Balasubramanian Raman, C Krishna Mohan, and Linga Reddy Cenkeramaddi. Application of a new machine learning model to improve earthquake ground motion predictions. *Natural Hazards*, 120(1):729–753, 2024.

[12] Marat Nurtas, Zhumabek Zhantaev, and Aizhan Altaibek. Earthquake time-series forecast in kazakhstan territory: Forecasting accuracy with sarimax. *Procedia Computer Science*, 231:353–358, 2024. 14th International Conference on Emerging Ubiquitous Systems and Pervasive Networks / 13th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (EUSPN/ICTH 2023).

[13] Wentao Zhou, Meng Huang, Shuai Liu, Qiao You, and Fanxin Meng. Research on the construction and application of earthquake emergency information knowledge graph based on large language models. *IEEE Access*, 13:127742–127757, 2025.

[14] Chenchen Xie, Huiran Gao, Yuandong Huang, Zhiwen Xue, Chong Xu, and Kebin Dai. Leveraging the deepseek large model: A framework for ai-assisted disaster prevention, mitigation, and emergency response systems. *Earthquake Research Advances*, 5(4):100378, 2025.

[15] Song Zhang, Meng Huang, Shuai Liu, Fanxin Meng, Yingyao Xie, Xirui Ren, Yuanwang Zhang, and Wenbo Shao. Ai-driven post-earthquake emergency material demand prediction: Integrating rag with reasoning large language model. *IEEE Access*, 13:100630–100646, 2025.

[16] Jin Han, Zhe Zheng, Xin-Zheng Lu, Ke-Yin Chen, and Jia-Rui Lin. Enhanced earthquake impact analysis based on social media texts via large language model. *International Journal of Disaster Risk Reduction*, 109:104574, 2024.

[17] Deep Patel, Panthadeep Bhattacharjee, Amit Reza, and Priodyuti Pradhan. Earthquake response analysis with ai. In Ngoc Thanh Nguyen, Tokuro Matsuo, Ford Lumban Gaol, Yannis Manolopoulos, Hamido Fujita, Tzung-Pei Hong, and Krystian Wojtkiewicz, editors, *Recent Challenges in Intelligent Information and Database Systems*, pages 18–30, Singapore, 2025. Springer Nature Singapore.

[18] Yongqing Jiang, Jianze Wang, Xinyi Shen, and Kaoshan Dai. Large language model for post-earthquake structural damage assessment of buildings. *Computer-Aided Civil and Infrastructure Engineering*, 2025.

[19] Bo Zhang, Ziang Hu, Pin Wu, Haiwang Huang, and Jiansheng Xiang. Ept: A data-driven transformer model for earthquake prediction. *Engineering Applications of Artificial Intelligence*, 123:106176, 2023.

[20] John B. Rundle, Geoffrey C. Fox, Andrea Donnellan, and Lisa Grant Ludwig. *Nowcasting Earthquakes with QuakeGPT: Methods and First Results*, pages 113–138. Springer Nature Singapore, Singapore, 2025.

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025

[21] Alireza Jafari, Geoffrey Fox, John B. Rundle, Andrea Donnellan, and Lisa Grant Ludwig. Time series foundation models and deep learning architectures for earthquake temporal and spatial nowcasting. *GeoHazards*, 5(4):1247–1274, 2024.

[22] Bikash Sadhukhan, Shayak Chakraborty, Somenath Mukherjee, and Raj Kumar Samanta. Climatic and seismic data-driven deep learning model for earthquake magnitude prediction. *Frontiers in Earth Science*, 11:1082832, 2023.

[23] Christopher W Johnson, Kun Wang, and Paul A Johnson. Automatic speech recognition predicts contemporaneous earthquake fault displacement. *Nature Communications*, 16(1):1069, 2025.

[24] Sevim Bilici, Fatih Külahcı, and Ahmet Bilici. Predicting the unpredictable: advancements in earthquake forecasting using artificial intelligence and lstm networks. *Geomagnetism and Aeronomy*, 64(5):760–771, 2024.

[25] Bikash Sadhukhan, Shayak Chakraborty, and Somenath Mukherjee. Predicting the magnitude of an impending earthquake using deep learning techniques. *Earth Science Informatics*, 16(1):803–823, 2023.

[26] Ewnetu Abebe, Hailemichael Kebede, Mickus Kevin, and Zelalem Demissie. Earthquakes magnitude prediction using deep learning for the horn of africa. *Soil Dynamics and Earthquake Engineering*, 170:107913, 2023.

[27] Cemil Emre Yavas, Lei Chen, Christopher Kadlec, and Yiming Ji. Improving earthquake prediction accuracy in los angeles with machine learning. *Scientific Reports*, 14(1):24440, 2024.

[28] Mehdi Akhoondzadeh. Earthquake prediction using satellite data: Advances and ahead challenges. *Advances in Space Research*, 74(8):3539–3555, 2024.

[29] Mayu Tsuchiya, Hiroyuki Nagahama, Jun Muto, Mitsuhiro Hirano, and Yumi Yasuoka. Detection of atmospheric radon concentration anomalies and their potential for earthquake prediction using random forest analysis. *Scientific Reports*, 14(1):11626, 2024.

[30] Sarah Oberbichler, Johanna Mauermann, The Trung Tran, and Carlos-Emiliano González-Gallardo. Studying model design biases in llms for multilingual historical newspaper extraction; the messina earthquake case study. In Wolf-Tilo Balke, Koraljka Golub, Yannis Manolopoulos, Kostas Stefanidis, and Zheying Zhang, editors, *Linking Theory and Practice of Digital Libraries*, pages 263–286, Cham, 2025. Springer Nature Switzerland.

DOI: https://doi.org/10.54216/MOR.050101

Received: June 05, 2025 Revised: August 07, 2025 Accepted: November 20, 2025