International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(3s) 2026 pp.482-492
https://doi.org/10.70917/ijcisim-2026-2345

Article

WAVELET–CNN–LSTM FRAMEWORK FOR ACCURATE EARTHQUAKE MAGNITUDE PREDICTION FROM SEISMIC WAVEFORMS

P. Devasudha[1], R. Ragupathy[2], M. Vadivukarassi[3]

[1]Department of Computer Science and Engineering, Annamalai University, Annamalainagar, Tamil Nadu-608002, India. Email: sudhajai2012@gmail.com
[2]Department of Computer Science and Engineering, Annamalai University, Annamalainagar, Tamil Nadu-608002, India. Email: cse_ragu@yahoo.com
[3]Department of Computer Science and Engineering, St. Martin's Engineering College, Secunderabad, Telangana-500100, India. Email: vadivume28@gmail.com

Abstract: Earthquake magnitude is a fundamental seismic parameter for assessing the severity of earthquake intensity and potential hazard. Early and precise magnitude estimation plays a vital role in facilitating rapid decision-making for earthquake disaster mitigation and emergency response. However, due to the spectral nature of seismic signals, it is difficult to achieve reliable prediction of earthquake. Traditional machine learning techniques, including random forest and gradient boosting, relied on manually engineered features and less temporal modelling, which often resulting in suboptimal performance as well as suffer from limitations such as high computational complexity, increased processing time and overfitting. To overcome these challenges this research proposed an innovative regression model which captures long term temporal dependencies and multi resolution frequency data for earthquake magnitude prediction by decomposing the seismic signals into one dimensional discrete wavelet transform to extract multi scale coefficients which is processed through convolution layers to identify patterns. Subsequently, Long Short-Term Memory layers are employed to model long range dependencies, while fully connected layers conduct nonlinear regression to predict the normalized magnitude. The model is trained using the Adam optimizer with early stopping and mean squared error as the loss function. Training and validation have done on a station-based subset of the Standard Earthquake Dataset (STEAD) comprising 1500 events with an 80/20 train–validation split. The proposed framework achieves RMSE of 0.328, an MAE of 0.253, and an R^2 of 0.567, surpassing baseline models based on Random Forest and Gradient Boosting. These findings imply that combining wavelet-based multi scale representation with CNN–LSTM temporal learning facilitates efficient, accurate and station-robust earthquake magnitude predictions, which is suitable for scalable real-time seismic monitoring systems.
Keywords: Earthquake magnitude prediction; Wavelet transform; CNN–LSTM; Seismic waveform analysis; STEAD dataset; Deep learning regression

1. INTRODUCTION

Earthquakes are sudden natural events that generate strong vibrations on the ground, causing structural damage, landslides and other hazards. They are unpredictable in nature, which frequently leads to a lot of financial losses and human damages which highlighting the critical need for rapid and accurate assessments [1]. Among seismic parameters, earthquake magnitude serves as a key indicator of event severity and is an effective solution for determining the critical functions in seismological monitoring and early warning systems [2]. Timely estimation of magnitude is essential for informed decision-making, allowing response to emergency response, evacuation, and disaster mitigation. In an Earthquake Early Warning system, the magnitude parameter must be calculated as soon as possible to allow warnings to be disseminated before the arrival of destructive earthquake waves [3]. In efforts to forecast earthquakes, such as their magnitude, location and occurance of time, numerous researchers have used physical techniques to explain and define them. Geological studies have also been conducted to identify earthquake predecessors and to better understand the underlying mechanisms of seismic activity [4-6]. Some of the indicators of earthquakes include those mentioned in previous studies, as intensity, magnitude and released seismic energy have been explored as explicit characteristics of seismic activity [7-8]. Although a variety of additional techniques have been proposed based on features and parameters influencing earthquakes, only a limited number of studies have achieved highly accurate predictions. This is limited due to earthquakes are inherently stochastic and complex phenomena, involving multiple interacting parameters which are hard to analyze comprehensively [9]. Therefore, a wide range of methodological approaches has been conducted in earthquake magnitude prediction studies, including mathematical modelling, precursor signal analysis, shallow machine learning algorithms, and deep learning techniques [10]. However, traditional methods for magnitude estimation, can be characterized by high costs of computation, time-consuming analyses, and overfitting. These limitations have inspired the use of deep learning approaches, which are capable of effectively learning and modeling complex spatiotemporal patterns inherent in seismic waveforms.

1.1 problem statement and motivation

Earthquake magnitude prediction is a challenging task that aims at estimating the magnitude of seismic events based on the seismic data that is recorded in the past. Traditional approaches for magnitude estimation mainly based on statistical modelling, geological considerations, and historical seismic records to analyse patterns associated with seismic activity. These traditional approaches assume that variables have linear correlations, which do not necessarily represent the complex, nonlinear dynamics that influence earthquake magnitudes and occurrences. In addition, they frequently neglect to represent long-range correlations in time-series data because of their short-term dependency.

However, earthquake magnitude prediction can be improved using existing advancements in machine learning. In particular, random forests and gradient boosting have been explored for estimating seismic parameters. Although these techniques enhance empirical relationships, they often fail to capture hierarchical and temporal dependencies inherent in raw seismic signals. Traditional deep learning models are able to extract feature directly from waveform data, improving seismic signal analysis capabilities with extended ensemble framework for richer future extraction [11]. Although RNNs are inherently suitable to model sequential data, capturing long-term dependencies remains difficult because of the challenges in model training. However, this issue remains unsolved, even with the introduction of LSTMs and GRUs. Alternatively, this study presents a framework that combines wavelet transforms, convolutional Neural networks, and Long short term memory regression based model, which is tested on STEAD dataset to characterize seismic parameters.

The rest of this paper is structured as follows. Section 1 gives the introduction to earthquake prediction. Section 2 reviews related studies in the field. Section 3 oulines the proposed methodology for magnitude prediction. Section 4 describes the experimental setup used in the study. Section 5 is the results and discussion section, and the last section 6 is the conclusion of the paper and possible future research directions.

2. RELATED WORKS

2.1 machine learning approaches for estimating earthquake magnitude

Machine learning is a big deal for automated seismic signal analysis, especially in earthquake detection and source characterization. Traditional machine learning approaches used custom made features based on spectral and temporal characteristics of seismic waveforms. Apriani et al., [3] showed that machine learning can effectively estimate earthquake magnitude for limited seismic data, also suggested that hybrid approaches combining physics and regression based machine learning models can also improve the accuracy of earthquake magnitude prediction.

As an illustration, random forest and support vector machine models were used to distinguish between seismic and noise events to estimate source parameters with the help of engineered descriptors extracted from waveform windows [15, 16]. Although they perform well in limited environments, these approaches are feature based and are frequently unable to generalize across regions and noise conditions.

2.2 *deep learning seismic signal analysis*

Deep learning has greatly improved the analysis of seismic signal. CNN based models learn features directly from raw waveforms or spectrograms for tasks like detection, phase picking, and magnitude regression. Recurrent neural networks, particularly LSTM, capture temporal dependencies in seismic sequences. Time–frequency approaches using wavelets or spectrograms provide multi scale representations but do not offer sequential modeling. Hybrid CNN–LSTM architectures improve temporal consistency; however, explicit integration of wavelet multi resolution decomposition with deep sequential learning for magnitude estimation remains limited in literature. The proposed framework addresses this gap by unifying multi scale signal processing and temporal modeling in a single architecture.

Wu et al., [12] demonstrated that deep learning achieves high accuracy in phase association compared to other classical association algorithms, however the performance strongly influenced by the quality and density of seismic networks, to enhance both reliability and interpretability, the approach should be integrated with physics based constraints. Perol et al., [13] proposed ConvNetQuake, a convolutional neural network developed for detecting and locating earthquakes directly from raw seismic waveforms, and generalizes well to unseen waveforms with robustness against noise, however exploring hybrid architectures for temporal context may improve the methodology for unlabeled seismic data. Zhu et al., [14] demonstrated that PhaseNet, a deep neural network trained on a large amount of labelled seismograms, can reliably identify P and S wave arrivals. Nevertheless, interpretability studies of the learned features could provide insights into the physical basis of the networks decisions in seismology. These models exploit local temporal patterns in waveforms and achieve improved robustness compared to classical detectors. Transfer learning strategies have also been explored to adapt deep seismic models across geographic regions and instrumentation differences [15]. In addition to detection, deep neural networks have been used for the earthquake source property estimation, including magnitude and distance regression. Regression-oriented architectures using convolutional encoders have demonstrated that waveform amplitude evolution and frequency content can be mapped directly to magnitude estimates without explicit feature extraction. Such approaches enable quick magnitude estimation suitable for early warning applications. However, purely convolutional models may not sufficient to capture long-duration temporal dependencies present in seismic signals.

To address temporal dynamics, recurrent neural networks and sequence models have been used for seismic waveform modeling. Gated recurrent architectures have demonstrated improved performance in capturing evolving waveform envelopes and source characteristics over long time windows. Hybrid convolutional–recurrent models have additional advantages for sequential seismic tasks including event characterization and ground-motion prediction, a combination of local feature extraction and temporal context integration.

2.3 Wavelet based multiscale seismic representation

In parallel, multiscale signal representations are significant in seismology due to the broadband and nonstationary character of earthquake waveforms. Wavelet transforms provide localized time–frequency analysis and have been widely used for seismic denoising, arrival detection, and feature extraction. Multiresolution wavelet features have been shown to enhance classification and characterization performance under low signal-to-noise conditions. Recent research indicates that incorporating multiscale representations can be used to improve deep learning performance by explicitly preserving scale-dependent information in seismic signals [16-18]. Although these advances have been made, limited work has been done to investigate the integration of explicit wavelet-based multiscale representations using hybrid convolutional–recurrent networks for earthquake magnitude regression. The majority of deep seismic regression models are based on raw waveform inputs or spectrogram representations, which may not fully exploit structured multiscale temporal–frequency information. This discontinuity inspires the proposed wavelet-integrated CNN–LSTM framework for improved seismic magnitude estimation.

Even though prior studies have shown the effectiveness of CNNs in deriving spatial features and LSTMs in learning temporal dependencies as well as hybrid architectures for magnitude estimation [19-22], a unified framework that explicitly incorporates wavelet-based multiresolution decomposition within a deep convolutional–recurrent architecture is still underexplored. This gap can be filled to potentially enhance robustness, interpretability, and predictive accuracy in earthquake magnitude estimation.

3. METHODOLOGY

3.1 STEAD Description

This study utilizes a station-based subset of the STEAD dataset [23], a large scale labeled seismic waveform repository designed for machine learning applications in seismology. The subset contains 1500 seismic events recorded across multiple stations with earthquake magnitudes greater than or equal to 1.0.

Magnitudes are normalized as:

$$\mathbf{Mag_{norm}} = \frac{\textbf{source magnitude}}{\mathbf{10}} \tag{1}$$

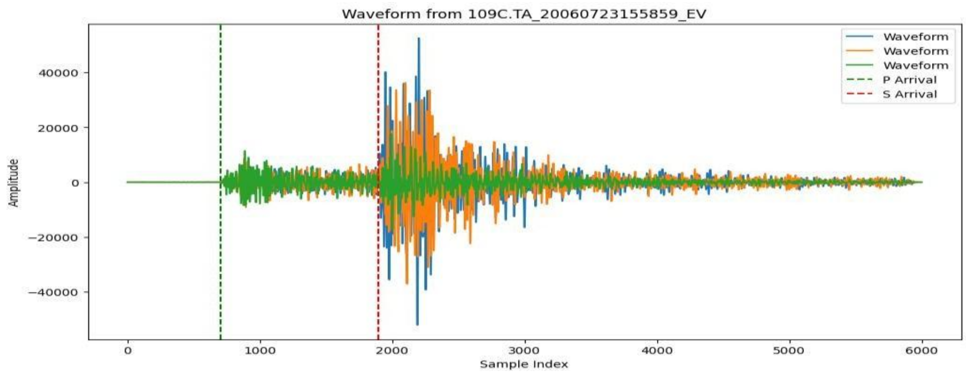

Figure 1 Sample waveform data in STEAD

3.2 Wavelet Feature extraction

Seismic waveforms are inherently nonstationary and contain energy distributed across multiplefrequency bands. To capture both temporal and spectral characteristics, a one-dimensional Discrete Wavelet Transform (DWT) is applied to each waveform channel using Daubechies-4 (db4) wavelets, which are widely adopted for seismic signal analysis due to their compactsupport and similarity to transient seismic pulses.

$$[C_A, C_D] = DWT(x(t)) \tag{2}$$

The coefficients and are concatenated to form input features for the CNN layers. This multiscale representation is used as input to the deep learning architecture to enhance feature discriminability for magnitude estimation.

3.3 Wavelet-CNN-LSTM model

The proposed model integrates convolutional and recurrent neural networks to jointly learn spatial features and temporal relationships from wavelet-domain seismic signals.

1. Convolutional Layers: Local temporal patterns in the wavelet representation are extracted using one-dimensional convolutional layers. Local temporal patterns extracted by

1D convolutions:

$$X_{conv} = ReLU(Conv1D(X_{wavelets})) \tag{3}$$

The convolutional filters learn localized waveform motifs such as amplitude transitions and frequency bursts associated with seismic energy release.

2. Maxpooling

Downsampling to reduce dimensionality:

To reduce feature dimensionality and improve robustness to small temporal shifts, max-pooling is applied:

$$\boldsymbol{X_{pool}} = \boldsymbol{MaxPool1D}(\boldsymbol{X_{conv}}) \tag{4}$$

This operation retains dominant activations while compressing the temporal resolution of feature maps.

3. LSTM Layers

Earthquake magnitude is strongly related to the temporal evolution of seismic energy. Long Short-Term Memory (LSTM) layers are employed to analyze long-range dependencies in the pooled feature sequence.

Captures long-term temporal dependencies:

$$h_t, C_t = LSTM(X_{pool}) \tag{5}$$

The LSTM gating mechanism preserves relevant historical information and captures wave-form duration and envelope growth patterns indicative of earthquake size.

4. Fully connected layers

The temporal representation is mapped to a continuous magnitude estimate through fully connected layers:

Produces regression output:

$$\hat{y} = W_2(\, ReLU(W_1 h_r + b_1)) + \, b_2 \tag{6}$$

A linear output layer is used to produce normalized magnitude predictions.

3.4 Proposed Model

The suggested framework introduces a multi scale-spatiotemporal learning strategy to estimate the magnitude of earthquakes. In contrast to conventional seismic deep learning approaches that only accept raw waveform convolution or spectrogram-based inputs, the proposed model combines wavelet-based multi resolution decomposition with sequential deep learning. This hierarchical integration of multi scale signal processing and time modeling improves magnitude regression accuracy and stability, particularly under limited training data conditions. Moreover, the architecture is computationally efficient because of dimensionality reduction introduced by wavelet decomposition, making it suitable for scalable seismic monitoring applications. The proposed Wavelet-CNN-LSTM model therefore offers a seismically interpretable and resource-efficient alternative to transformer-based architectures for predicting magnitude.

Figure 2 emphasizes the Wavelet-CNN-LSTM framework, highlighting the transformation of raw seismic waveforms into multi scale features, spatial temporal representations, and final magnitude estimates.

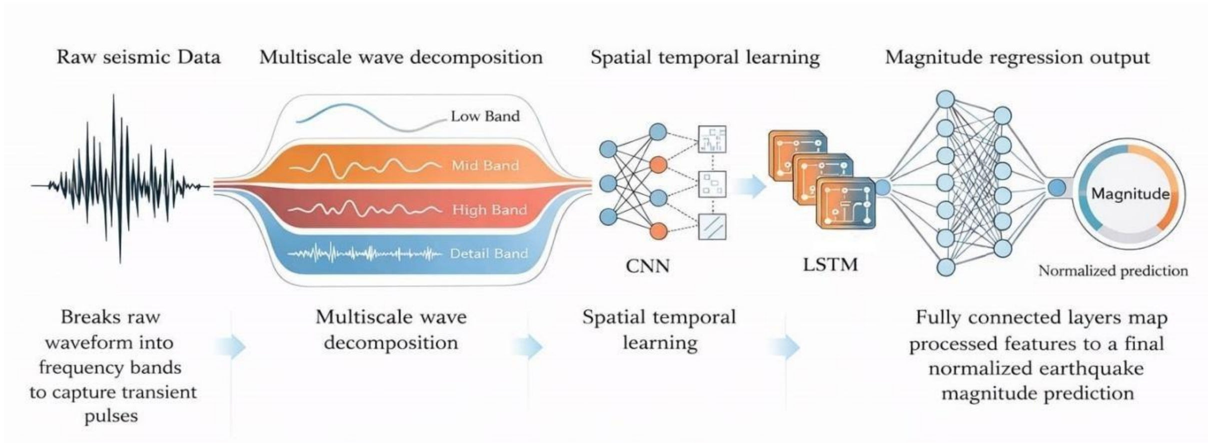

Figure 2 Work flow of Wavelet-CNN-LSTM architecture

The architecture incorporates multiresolution signal decomposition with spatial-temporal feature learning to predicate earthquake magnitude from raw seismic waveforms. The general framework consists of four sequential stages: wavelet decomposition, convolutional feature extraction, temporal sequence modeling, and regression mapping.

3.4.1 Wavelet Decomposition

Assuming an input three-component seismic waveform

$$X \in R^{6000*3} \tag{7}$$

Each channel is independently subjected to a 4-level discrete wavelet transform (DWT) is applied independently to each channel to get multiscale representations. The decomposition divides low- and high-frequency components, allowing the model to capture both long-period and transient seismic characteristics. The resulting coefficients are concatenated by channels and scales to create

$$X_w \in R^{750*12} \tag{8}$$

This stage decreases the temporal resolution but maintains salient frequency information relevant to earthquake magnitude.

3.4.2 Convolutional Feature Extraction

The multiscale coefficients are processed by a 1-dimensional convolutional neural network to learn local waveform patterns:

$$F_{CNN} = CNN\ (X_w) \tag{9}$$

The CNN contains two convolutional layers with 32 and 64 filters, separated by temporal pooling. This stage records amplitude envelopes, waveform shapes, and localized seismic signatures. the feature tensor after convolution and pooling is,

$$F_{CNN} = R^{94*64} \tag{10}$$

3.4.3 Temporal Modeling Using LSTM

The temporal dependence of earthquake signals are multiscaled, the CNN features are inputted into an LSTM network to model these sequential correlations:

$$H = LSTM\ (F_{CNN}) \tag{11}$$

The LSTM has 64 hidden units and takes the sequence of length 94. The last hidden state

$$h_r \in R^{64} \tag{12}$$

encodes the global temporal dynamics of the seismic event.

3.4.4 Magnitude Regression

The trained temporal representation is trnsformed to earthquake magnitude through fully connected layers:

$$\hat{y} = W_2(\sigma\ (W_1 h_r + b_1)) + b_2 \tag{13}$$

where $\sigma(\cdot)$ denotes the ReLU activation function. The network provides a continuous magnitude estimate

$$\hat{y} \in R \tag{14}$$

The average squared error (MSE) loss:

$$L_{MSE} = \frac{1}{N}\sum_{i=1}^{N} (\hat{y}_i - y_i)^2 \tag{15}$$

4. EXPERIMENTAL SETUP

4.1 Training strategy

The Adam optimizer is used to train the network with a learning rate set at 0.001 and a batch size of 16. The dataset is divided into 80% for training and 20% for validation. To avoid overfitting, early stopping is applied with a patience of 5 epochs,by halting training when there is no improvement in validation loss. The model parameters are randomly set and updated through backpropagation until convergence.

4.2 Loss function

Mean Squared Error (MSE):

The parameters of the model are optimized to minimize Mean Squared Error (MSE) between predicted and true normalized magnitudes:

$$MSE = \frac{1}{N}\sum_{i=1}^{N} (\hat{y}_i - y_i)^2$$

MSE is commonly used in seismic magnitude regression tasks due to its sensitivity to large

prediction errors.

5. RESULTS AND DISCUSSION

5.1 Regression performance

The performance of the proposed approach is validated using some of the regression metrics like Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R-squared. The performance values are illustrated in Table 1. Further training and testing losses and error also validated.

Table 1: Regression values on test set.

Metric	**Value**
RMSE	0.328
MAE	0.253
R2	0.567

The performances of suggested method is compared with some of the existing magnitude prediction frameworks including Gradient boosting and random forest to prove the efficacy of the proposed approach. Furthermore, the training accuracy and loss of presented approach is illustrated in Figure 3. These curves demonstrated that a consistent decline across 21 epochs, it conforms that the model learns successfully the data distribution. The relative gap between validation and training loss further suggests that overfitting is minimal.

The scatter plot comparing predicted values against actual targets provides evidence of strong predictive accuracy as shown in Figure 4. The greater part of data points are closely aligned with the reference diagonal, representing the ideal case of perfect prediction. This clustering implies that the model captures the relationship between input features and target outputs effectively. Deviations from the diagonal are relatively small and evenly distributed, indicating the absence of systematic bias in predictions.

The histogram of prediction errors reveals a distribution centered around zero, with a symmetric spread as shown in Figure 5. This pattern confirms that the model does not consistently overestimate or underestimate outcomes. The concentration of errors near zero highlights that most predictions are close to the true values, while the balanced tails of the distribution reflect robustness across varying input conditions.

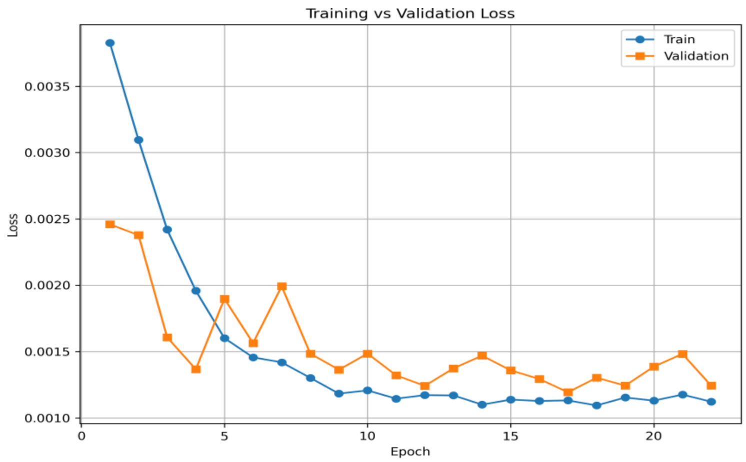

Figure 3 Training and validation curves

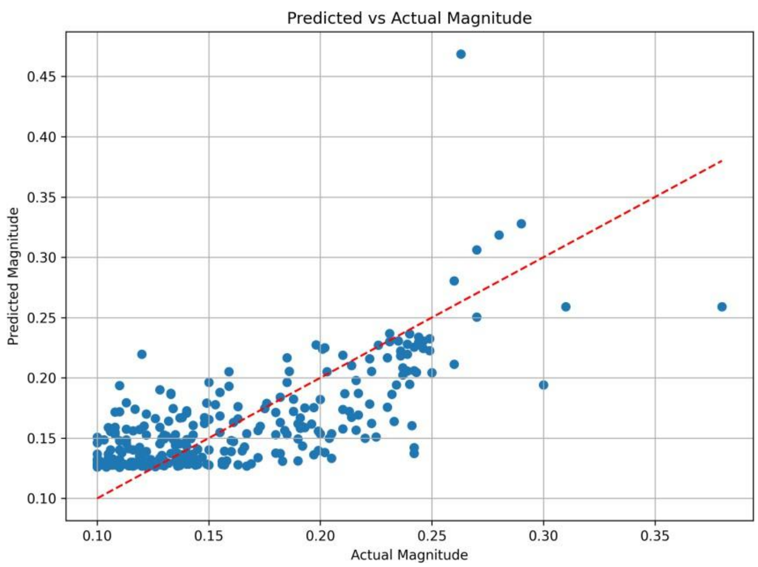

Figure 4 Predicted Vs. actual earthquake magnitude

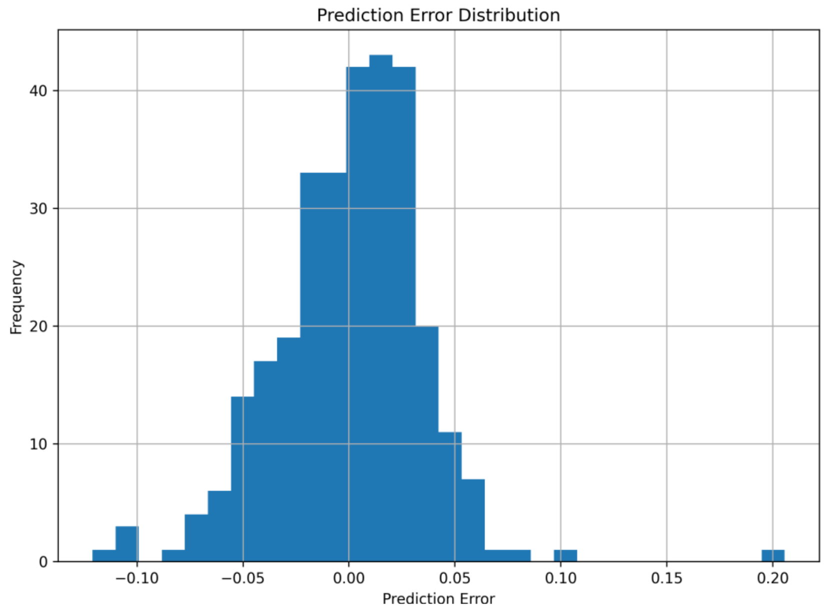

Figure 5 Prediction error distribution

5.2 Performance evaluation

5.2.1 Model Performance Comparison

The performance of proposed methodology is compared with RMSE, MAE and R2 values of Random forest and Gradient boosting are depicted in table 2.

Table 2: Comparison of Wavelet-CNN-LSTM with traditional ML models

Model Names	RMSE	MAE	R2
Random Forest	0.591	0.453	0.515
Gradient Boosting	0.528	0.412	0.542
Wavelet-CNN-LSTM	**0.328**	**0.253**	**0.567**

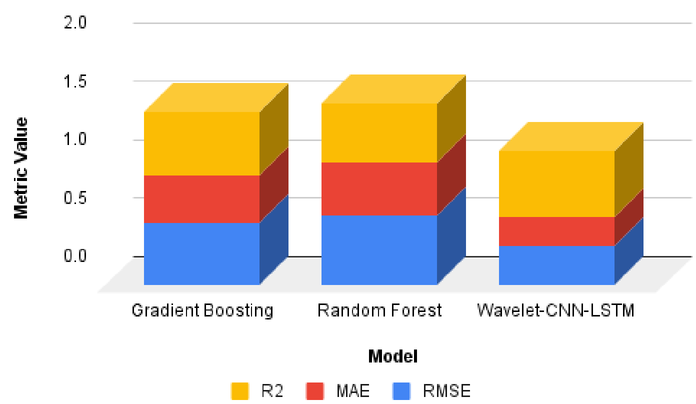

Figure 6 Performance evaluation of proposed approach with existing models

The performance of the suggested approach have found to be effective. The recommended model attained R-squared value of 0.567 while existing attained 0.515 and 0.542 for random forest and gradient boosting model. Further the mean absolute error and Root mean squared error of presented model is comparatively low as attained 0.253 and

0.328 , which indicates the proposed model predicted the magnitude in better manner as shown in Figure 6. In general lesser error rate significantly increases the performances.

5.3 Interpretation of Results

Overall, these findings indicate that the proposed model demonstrates robust performance in regression tasks. The loss curve validates effective training dynamics, the predicted vs. actual plot demonstrates high predictive fidelity, and the error distribution confirms unbiased error behavior. Collectively, these findings indicate methodological robustness and support the suitability of the model for earthquake-related prediction tasks.

6. CONCLUSION

This study presented a deep learning framework for earthquake magnitude estimation using seismic waveform data from the STEAD dataset. The proposed model integrates wavelet-based time–frequency representation with convolutional neural networks (CNN) and long short-term memory (LSTM) layers to simultaneously capture spectral characteristics and temporal dependencies of seismic signals. A station-based subset selection strategy was adopted to ensure computational feasibility while preserving waveform diversity. The results show that the proposed Wavelet–CNN–LSTM regression model achieves reliable magnitude estimation performance with an RMSE of 0.33 and coefficient of determination (R2) of 0.57 on the selected STEAD subset. The training and validation loss curves indicate stable convergence with effective generalization enabled by early stopping. The predicted versus actual magnitude distribution further confirms the capability of the model to capture magnitude trends across the considered seismic range. The error distribution analysis shows that most prediction errors are concentrated near zero, indicating unbiased estimation behavior.

Compared with conventional machine learning approaches that rely on handcrafted seismic features, the proposed framework automatically learns hierarchical representations directly from raw waveforms. The integration of wavelet decomposition enhances multi-resolution feature extraction, while CNN layers capture local waveform patterns and LSTM units model temporal evolution. This hybrid architecture provides an effective end-to-end solution for seismic magnitude regression under limited computational resources. Overall, the findings confirm that combining time–frequency analysis with deep temporal modeling significantly improves earthquake magnitude prediction from seismic waveforms. The proposed approach offers a scalable and computationally efficient methodology suitable for real-time seismic monitoring systems and data-driven seismic hazard assessment. In future, the proposed model's generalizability is explored using the Indian and its surrounding region dataset.

References:

1. Yousefzadeh, Mohsen, Seyyed Ahmad Hosseini, and Mahdi Farnaghi. "Spatiotemporally explicit earthquake prediction using deep neural network." *Soil Dynamics and Earthquake Engineering* 144 (2021): 106663.
2. X. Zhang, Y. Wu, and Z. Wang, "Machine learning approaches for earthquake magnitudeestimation," Geophysical Journal International, vol. 218, pp. 1234–1248, 2019.
3. Apriani, M., S. K. Wijaya, and Daryono. "Earthquake magnitude estimation based on machine learning: Application to earthquake early warning system." *Journal of Physics: Conference Series*. Vol. 1951. No. 1. IOP Publishing, 2021.
4. Jena, Ratiranjan, Biswajeet Pradhan, Sambit Prasanajit Naik, and Abdullah M. Alamri. "Earthquake risk assessment in NE India using deep learning and geospatial analysis." *Geoscience Frontiers* 12, no. 3 (2021): 101110.
5. Renouard, Alexandra, Alessia Maggi, Marc Grunberg, Cécile Doubre, and Clément Hibert. "Toward false event detection and quarry blast versus earthquake discrimination in an operational setting using semiautomated machine learning." *Seismological Society of America* 92, no. 6 (2021): 3725-3742.
6. Lv, Hao, Xiangfang Zeng, Feng Bao, Jun Xie, Rongbing Lin, Zhenghong Song, and Gongbo Zhang. "ADE-Net: A deep neural network for DAS earthquake detection trained with a limited number of positive samples." *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022): 1-11.
7. Bao, Zhenyu, Jingyu Zhao, Pu Huang, Shanshan Yong, and Xin'an Wang. "A deep learning-based electromagnetic signal for earthquake magnitude prediction." *Sensors* 21, no. 13 (2021): 4434.
8. Joshi, Anushka, 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, no. 1 (2024): 729-753.
9. Kavianpour, Parisa, Mohammadreza Kavianpour, Ehsan Jahani, and Amin Ramezani. "A CNN-BiLSTM model with attention mechanism for earthquake prediction." *The Journal of Supercomputing* 79, no. 17 (2023): 19194-19226.
10. Mir, Adil Aslam, Fatih Vehbi Çelebi, Hadeel Alsolai, Shahzad Ahmad Qureshi, Muhammad Rafique, Jaber S. Alzahrani, Hany Mahgoub, and Manar Ahmed Hamza. "Anomalies prediction in radon time series for earthquake likelihood using machine learning-based ensemble model." *IEEE Access* 10 (2022): 37984-37999.

11. Joshi, Anushka, Chalavadi Vishnu, and C. Krishna Mohan. "Early detection of earthquake magnitude based on stacked ensemble model." *Journal of Asian Earth Sciences: X* 8 (2022): 100122.
12. Y. Wu and L. Zhao, "Earthquake phase association using deep learning," Geophysical Journal International, 2019.
13. T. Perol, M. Gharbi, and M. Denolle, "Convolutional neural network for earthquake detection and location," *Science Advances*, vol. 4, no. 2, 2018.
14. W. Zhu and G. C. Beroza,"PhaseNet: A deep-neural-network-based seismic arrival-time picking method," *Geophysical Journal International*, vol. 216, no. 1, pp. 261–273, 2019.
15. H. Sun, X. Chen, and J. Wu, "Deep learning for ground motion prediction using CNN–RNN models," *Soil Dynamics and Earthquake Engineering*, vol. 132, 2020.
16. I. Daubechies, *Ten Lectures on Wavelets*, SIAM, 1992.
17. S. Mallat, *A Wavelet Tour of Signal Processing*, 3rd ed., Academic Press (Elsevier), 2008.
18. P. S. Addison, *The Illustrated Wavelet Transform Handbook*, 2nd ed., CRC Press, 2017.
19. S. Kiranyaz, T. Ince, and M. Gabbouj, "Wavelets in deep learning: A review," *IEEE Signal Processing Magazine*, vol. 38, no. 6, pp. 59–76, 2021.
20. S. Gaci, "Wavelet-based seismic signal classification in noisy environments," *Digital Signal Processing*, vol. 48, pp. 46–56, 2016.
21. S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
22. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.
23. M. Schimmel, L. Yu, A. Patoine, and Z. E. Ross, "STEAD: A global seismic waveform dataset for deep learning," *Scientific Data*, vol. 7, 2020.\