Final Project For ECE228 Track Number #1

Application of Long Short-Term Memory (LSTM) Networks-Based Surrogate Modeling for Nonlinear Structural Systems

Group Number #2: Abdoul Aziz Sandotin Coulibaly **Enrique Simbort Zeballos**

Yusuf Morsi **Ramin Sarange**

Abstract

Performance-based seismic design (PBSD) of structural systems relies on computationally expensive high-fidelity finite element (FE) models to predict how structures will respond to seismic excitation. For risk-based assessments, FE response simulations must be run thousands of times with different realizations of the sources of uncertainty. Consequently, data-driven machine learning (DDML) surrogate models have gained prominence as fast emulators for predicting seismic structural responses in probabilistic analyses. This paper leverages deep Long Short-Term Memory (LSTM) networks, known for their powerful and flexible framework for time series prediction tasks. The advantages of using LSTM networks include their ability to model continuous-time processes, adapt to varying temporal resolutions, maintain implicit memory of past information, model complex nonlinear dynamics, perform interpolation and extrapolation, handle noisy data robustly, and scale effectively to high-dimensional datasets. The effectiveness of the proposed method is validated through three proof-of-concept studies: one involving a linear elastic 2D 8-degree-of-freedom (DoF) shear building model, a nonlinear single degree of freedom system (NL-SDoF), and a 2D nonlinear 3DoF shear building model. The first and third models simulate a moment-resisting frame system. The findings indicate that the proposed LSTM network is a promising, dependable, and computationally efficient technique for predicting nonlinear structural responses. This paper offers a detailed and systematic introduction to the field of machine learning for structural engineers, aiming to demystify the often opaque processes and techniques involved.

The code implementation for this study can be accessed at this GitHub repository.

I certify that I have filled the evaluation.

1 Introduction

The earthquake events in Afghanistan, Morocco, and Turkey in 2023 led to 63,701 fatalities [Göçümlü, B. Ç., 2023]. These events underscore the necessity for a more refined approach to predicting the dynamic response behavior of structures subjected to large seismic events. This is particularly critical for complex civil infrastructure such as long-span bridges and concrete dams, where the consequences of failure are deemed unacceptable [Seed et al., 1973]. The 2021 report card released by the American Society of Civil Engineers (ASCE) assigned a C- grade to the U.S. infrastructure,

Preprint. Under review.

highlighting the need for structural assessment (SA) of critical civil infrastructure systems. However, evaluating the seismic response of long-span bridges or dams subjected to moderate to strong intensity ground motions poses a considerable challenge due to the lack of experimental data and the high computational cost associated with the validation process of state-of-the-art finite element (FE) models, which often rely on Bayesian inference.

2 Related work

Traditional methods for analyzing structural dynamic responses to natural hazards typically use physics-based models, such as the finite element method (FEM) (Di Paola and Navarra [2009]). FEM requires fine mesh discretization and small time steps for accuracy, making it computationally expensive, especially for nonlinear time history analysis and when numerous simulations are needed to account for stochastic uncertainties in external loads, such as Monte Carlo simulations (Zhang et al. [2013]) or incremental dynamic analysis (Vamvatsikos and Fragiadakis [2010]). Additionally, model assumptions and parameter uncertainties can lead to significant errors.

Alternative approaches, such as system identification (SI) methods, perform linear/nonlinear mapping of external excitations to structural responses using state-space or black-box models (Yousefianmoghadam et al. [2018], Brownjohn and Xia [2000]). However, these methods often struggle with nonlinear responses due to assumptions of stationary dynamics and linearity (Akaike [1969], Durbin [1959] , Said and Dickey [1984],). Artificial neural networks (ANNs), particularly multi-layer perceptrons (MLPs), have been used successfully for modeling structural responses under various loads but have limitations with complex time series data (Fishwick [1989], Ediger and Akar [2007], Moaveni et al. [2009]).

Recent advances in deep learning, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), offer better capabilities for modeling nonlinear structural responses (Sun et al. [2017], Wu and Jahanshahi [2019]). Despite CNNs' strength in data classification with grid-like topology, they struggle with large plastic deformations. RNNs are designed for sequential, time-varying data but face issues with long-range dependencies due to gradient vanishing and exploding (Jain et al. [2000], Mandic and Chambers [2001]).

This work reproduces an existing machine learning structural engineering application, addressing the limitations described above by developing a long short-term memory (LSTM) network for nonlinear structural response modeling extended to include strong motion datasets (Zhang et al. [2019]). Based on the studied paper, this work introduces one LSTM scheme for seismic response modeling. More specifically, this work methodically establishes the credibility of machine learning models, such as LSTM, as surrogate models for structural dynamics. By progressively increasing the level of nonlinearity in the structural dynamics models, we demonstrate the effectiveness and reliability of LSTM in accurately capturing complex behaviors. This step-by-step approach ensures a robust validation of machine learning techniques in handling intricate structural engineering problems. The effectiveness of the proposed method is validated by three proof-of-concept studies: 2D linear elastic 8 degree-of-freedom (DoF) shear building model, a nonlinear single degree of freedom system (NL-SDoF), and the third - a 2D nonlinear 3DoF shear building model. The second and the third models incorporate material nonlinearity only, and no effects of geometric nonlinearity are considered. These three models simulate real-world moment-resisting frame systems. The findings indicate that the proposed LSTM network is a promising, dependable, and computationally efficient technique for predicting nonlinear structural responses, with significant potential for assessing seismic risk and reliability analysis.

The paper is organized as follows: Section 1 introduces the motivation behind this work; Section 2 reviews previous related work and the evolution of methodologies prior to LSTM; Section 3 describes the methodology used in this study; Section 4 validates the methodology through proof-of-concept studies; and Section 5 discusses the limitations and presents the conclusions.

3 Methodology

3.1 Structural Analysis and Computational Mechanics

Consider the following system of second-order differential equations of motion governing the response of a multi-degree of freedom (MDF) system to earthquake-induced ground motion as follows:

$$\mathbf{M}\ddot{\mathbf{U}}(t) + \mathbf{C}\dot{\mathbf{U}}(t) + \mathbf{R}(\mathbf{U}, \dot{\mathbf{U}}, t) = -\mathbf{M}\mathbf{L}\ddot{\mathbf{u}}_g(t), \tag{1}$$

with the initial conditions

$$\mathbf{U} = \mathbf{U}(0) \quad \text{and} \quad \dot{\mathbf{U}} = \dot{\mathbf{U}}(0), \tag{2}$$

where $\mathbf{M} \in \mathbb{R}^{n\times n}$ is the symmetric positive-definite mass matrix of the structure, i.e. $\mathbf{M} = \mathbf{M} > 0$, $\mathbf{C} \in \mathbb{R}^{n\times n}$ is the structure damping matrix, based on Rayleigh damping model (3), $\mathbf{R}(\mathbf{U}, \dot{\mathbf{U}}, t) \in \mathbb{R}^n$ is nonlinear path-dependent restitutive force vector, $\mathbf{L} \in \mathbb{R}^n$ is the influence vector (Chopra [2012]), $\mathbf{U} \in \mathbb{R}^n$ is the relative displacement vector with respect to the base of the structure, $\dot{\mathbf{U}}(t)$ and $\ddot{\mathbf{U}}(t)$ are the first and second-time derivatives of $\mathbf{U}(t)$, respectively, and $\ddot{\mathbf{u}}_g(t) \in \mathbb{R}^K$ is the input ground acceleration.

$$\mathbf{C} = \alpha\mathbf{M} + \beta\mathbf{K} \in \mathbb{R}^{n\times n} \tag{3}$$

where α and β can be determined from specified damping ratios ξ_i and ξ_j for the ith and jth modes, respectively. The system in (1) is regarded as the full-order model (FOM). As was stated in Section 2 the formulation of the nonlinear differential equations, with path-dependent nonlinear terms, is computationally demanding. The nonlinear path-dependent restitutive force vector must be reformulated at each time instant from the element force vectors corresponding to the present deformation and its path dependence and for a large structure this process must be repeated for thousands of structural elements.

The objective is given the FOM described by (1) to find a reduced-order model based on the deep LSTM networks approach (Zhang et al. [2019]). With the aim of evaluating, and extending the methodology proposed in (Zhang et al. [2019]), we present three study examples in Section 4 (see Figure 1).

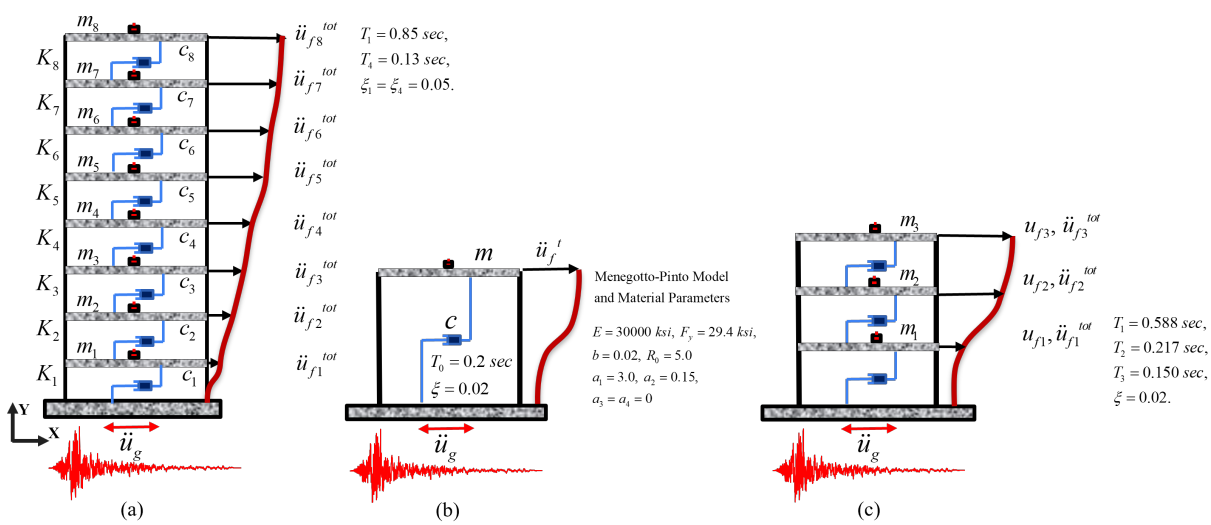


Figure 1: Test-bed structures for surrogate modeling study

3.2 Long Short-Term Memory (LSTM) Network

3.2.1 *The FOM implementation*

In this section we consider a FOM with n = 8 DOFs to investigate the numerical performance of the LSTM Network. The system under study is governed by Equation (1) with initial conditions given by

Equation (2). In this first step, we analyze the linear version of Equation (1) which can be written as follows:

$$\mathbf{M}\ddot{\mathbf{U}}(t) + \mathbf{C}\dot{\mathbf{U}}(t) + \mathbf{K}\mathbf{U}(t) = -\mathbf{M}\mathbf{L}\ddot{\mathbf{u}}_g(t), \tag{4}$$

with the initial conditions

$$\mathbf{U} = \mathbf{U}(0) \quad \text{and} \quad \dot{\mathbf{U}} = \dot{\mathbf{U}}(0), \tag{5}$$

where $\mathbf{M}, \mathbf{K} \in \mathbb{R}^{n\times n}$ are the symmetric positive-definite mass and stiffness matrices of the structure, i.e. $\mathbf{M} = \mathbf{M} > 0$, $\mathbf{K} = \mathbf{K} > 0$, $\mathbf{C} \in \mathbb{R}^{n\times n}$ is the structure damping matrix, based on Rayleigh damping model (3), $\mathbf{L} \in \mathbb{R}^n$ is the influence vector (Chopra [2012]), $\mathbf{U} \in \mathbb{R}^n$ is the relative displacement vector with respect to the base of the structure, $\dot{\mathbf{U}}(t)$ and $\ddot{\mathbf{U}}(t)$ are the first and second time derivatives of $\mathbf{U}(t)$, respectively, and $\ddot{\mathbf{u}}_g(t) \in \mathbb{R}^K$ is the input ground acceleration. The Rayleigh damping parameters, α and β, are computed based on damping ratios $\xi_1 = 0.05\%$ and $\xi_4 = 0.05\%$ for the 1st and 4th modes of vibration, respectively. The with first and the fourth natural periods of vibration are $T_1 = 0.85$ s. and $T_4 = 0.13$ s. According to this $\alpha = 0.64$ and $\beta = 0.0018$. The FOM was developed in OpenSees (McKenna et al. [2010]) as 2-D linear elastic model of this shear building model system, using fiber-section force-based Bernoulli-Euler beam-column elements (SPACONE et al. [1996]). A step-by-step earthquake response analysis for the first 20 seconds of each selected ground motion was performed. The FOM equations of motion are numerically integrated using the Newmark-beta method, namely the Constant Average Acceleration time stepping scheme, with $(\gamma = 1/2, \beta = 1/4)$. As integration time step, $\Delta t = 0.01$ s. was used, which is the same as the sampling time of the earthquake record. The modal and material properties of models two and three are presented in Figure 1(a) and (b), respectively.

Models Two and Three are governed by the same equations of motion described above. Figure (1(b) and (c)) illustrates all material properties and parameters defining the nonlinear behavior of the uniaxial material model employed to simulate the structure's nonlinear response. As it can be observed from Figure 1(b), the second test-bed structure has natural period of vibration $T_1 = 0.2$ s. and it was modeled considering a damping ratio of $\xi_1 = 0.02\%$. The third test-bed structure (see Figure 1(c)) considered an equal value of damping ratio, $\xi_1 = 0.02\%$ for the 1st and 2th modes of vibration. The with first and the second natural periods of vibration are $T_1 = 0.588$ s. and $T_2 = 0.217$ s.

3.2.2 *The LSTM cell*

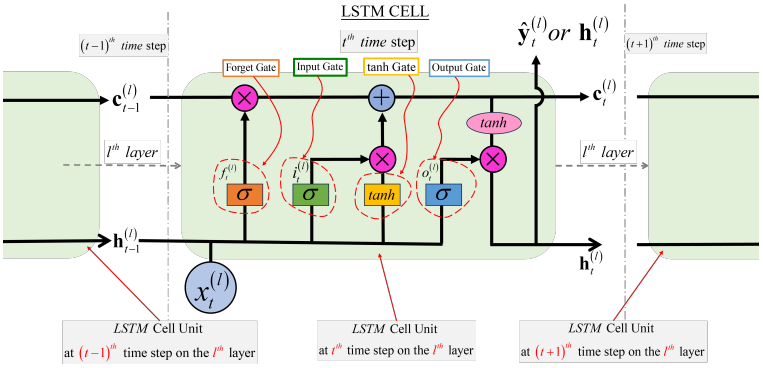
The LSTM cell is the fundamental building block of the LSTM network (See Fig). It consists of several gates that control the flow of information, making it capable of learning long-term dependencies. An LSTM Cell at a specific time step τ receive simultaneously an input vector $\mathbf{x}_\tau$, the hidden state output of the previous step, $\mathbf{h}_{\tau-1}$ and the cell memory state of the previous time step $\mathbf{c}_{\tau-1}$. At τ , the input gate, the forget gate, output gate, and cell state are denoted respectively as $\mathbf{f}_\tau$,$\mathbf{o}_\tau$, $\mathbf{c}_\tau$. We also coin the word candidate state, which we denote as $\tilde{\mathbf{c}}_{\tau-1}$. Mathematical relationships between the gates and states are presents as follow:

$$\begin{aligned}
f_\tau^l &= \sigma\left(\mathbf{W}_{hf}^l \mathbf{h}_{\tau-1}^l + \mathbf{W}_{xf}^l \mathbf{x}_\tau^l + \mathbf{b}_f^l\right) \\
i_\tau^l &= \sigma\left(\mathbf{W}_{hi}^l \mathbf{h}_{\tau-1}^l + \mathbf{W}_{xi}^l \mathbf{x}_\tau^l + \mathbf{b}_i^l\right) \\
\tilde{c}_\tau^l &= \tanh\left(\mathbf{W}_{h\tilde{c}}^l \mathbf{h}_{\tau-1}^l + \mathbf{W}_{x\tilde{c}}^l \mathbf{x}_\tau^l + \mathbf{b}_{\tilde{c}}^l\right) \\
c_\tau^l &= f_\tau^l \odot c_{\tau-1}^l + i_\tau^l \odot \tilde{c}_\tau^l \\
o_\tau^l &= \sigma\left(\mathbf{W}_{ho}^l \mathbf{h}_{\tau-1}^l + \mathbf{W}_{xo}^l \mathbf{x}_\tau^l + \mathbf{b}_o^l\right) \\
h_\tau^l &= o_\tau^l \odot \tanh\left(c_\tau^l\right)
\end{aligned}$$

In the formulation above, the symbol $\odot$ in our notation represents the Hadamard product, also known as the element-wise product. $\sigma(\cdot)$ and $\tanh(\cdot)$ are respectively sigmoid and hyperbolic tangent activation functions, with equations as follows:

$$\sigma(y) = \frac{1}{1+e^{-y}} \quad \text{and} \quad \tanh(y) = \frac{e^y - e^{-y}}{e^y + e^{-y}}.$$

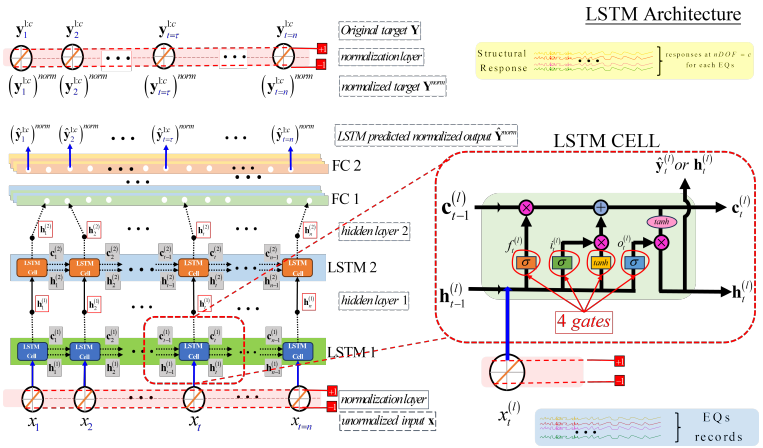
Here, we intend not to provide too much details of the goal of each gate as it will lengthen our final report. We direct interested readers to Olah [2015]



3.2.3 *The LSTM network architecture: Full sequence to sequence LSTM network (LSTM-f)*

The LSTM network architecture is designed to handle sequential data by maintaining a memory of previous inputs. This architecture includes multiple LSTM cells arranged in layers to capture complex temporal patterns.

The full sequence-to-sequence LSTM network (LSTM-f) is an advanced configuration that processes entire sequences of data from input to output. It is particularly effective for tasks requiring accurate prediction of time-series data. See the figure below that shows the LSTM architecture configuration.



4 Experiments

4.0.1 *Experimental Setup*

4.0.2 *Model Training*

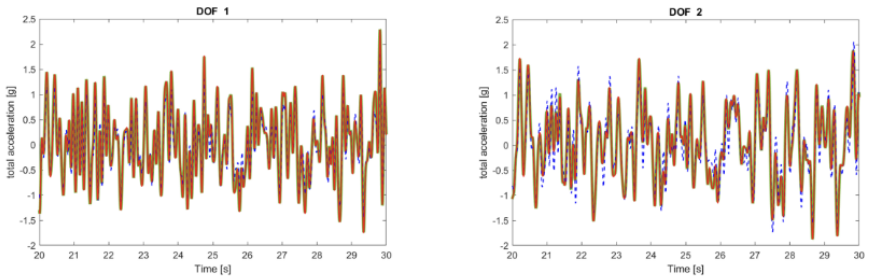
For preparing our data, we had to scale both the input and output data using the MinMaxScaler to normalize the values between -1 and 1 (this was done to prevent bias from data magnitude). Subsequently, our input and output data were reshaped and scaled to fit the Long Short-Term Memory (LSTM) network requirements shape requirements from Tensorflow Abadi et al. [2015]. The model architecture was essentially two LSTM layers (each with 100 units), in addition to two dense layers to map the LSTM outputs to the target space. We used the common Adam optimizer and configured our model with a learning rate of 0.001. We also decided to implement the mean squared error (MSE) loss function. Using a batch of 5 and running a variety of epochs numbers, we monitored the training and validation losses. We also shuffled the dataset before every epoch and saved the best-performing model based on validation loss in order to increase the robustness of the model and prevent overfitting. Overall, this process exemplified the LSTM model's ability to find temporal dependencies in seismic response data.

We also went through the process of leveraging a high-performance computing environment for machine learning tasks. The process begins with requesting 1 node and 4 NVIDIA Quadro RTX

5000 GPUs for computational tasks at the Texas Advanced Computing Center (TACC), highlighting the robust computational resources utilized in this workflow. The model is then executed on the Frontera supercomputer, which is noted as the fastest academic supercomputer in the world. This step is critical for handling complex computations and large datasets. The Machine Learning Frameworks used include Python and popular machine learning libraries such as Keras and TensorFlow. In terms of post-processing, after execution, the results are post-processed using MATLAB for further analysis and visualization.

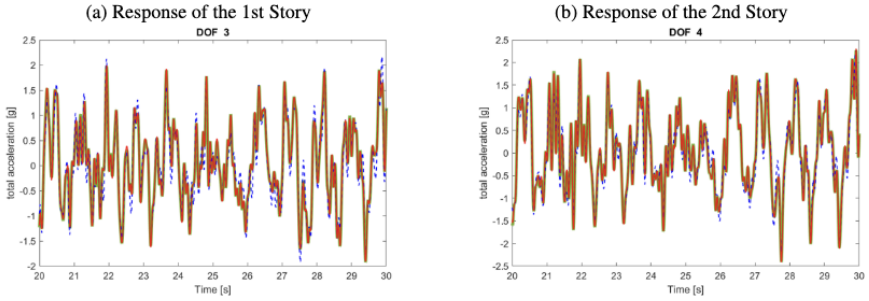
4.0.3 *Results for the Linear Elastic 8-Story Shear Building*

The following are the results for a linear elastic 8-story shear building:



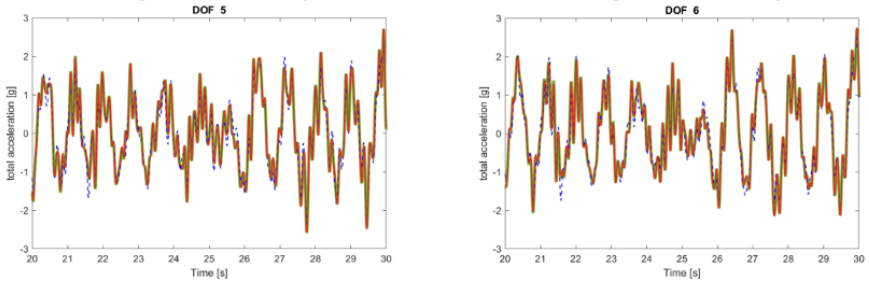
(a) Response of the 1st Story

(b) Response of the 2nd Story



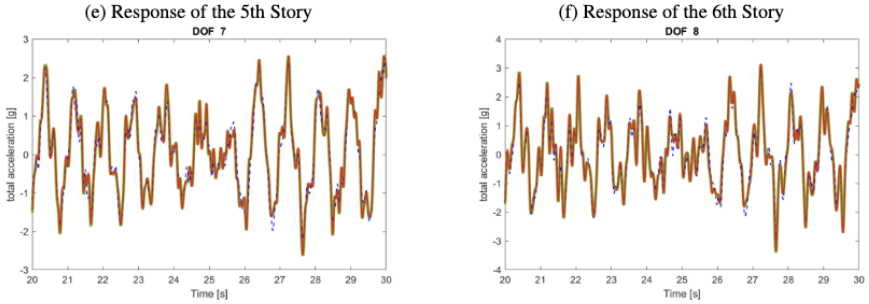
(c) Response of the 3rd Story

(d) Response of the 4th Story



(e) Response of the 5th Story

(f) Response of the 6th Story



(g) Response of the 7th Story

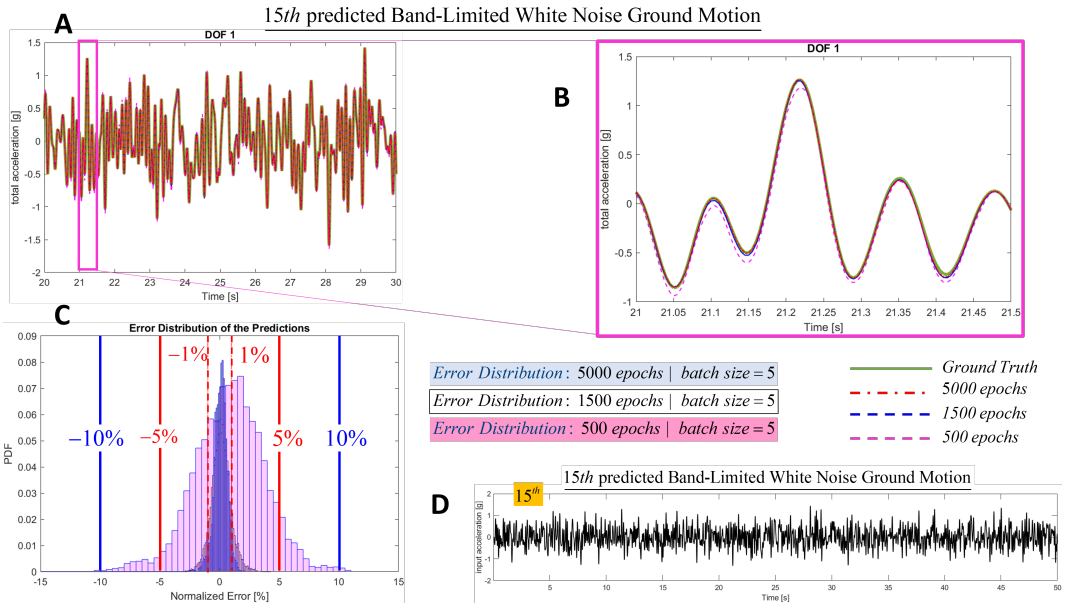
(h) Response of the 8th Story

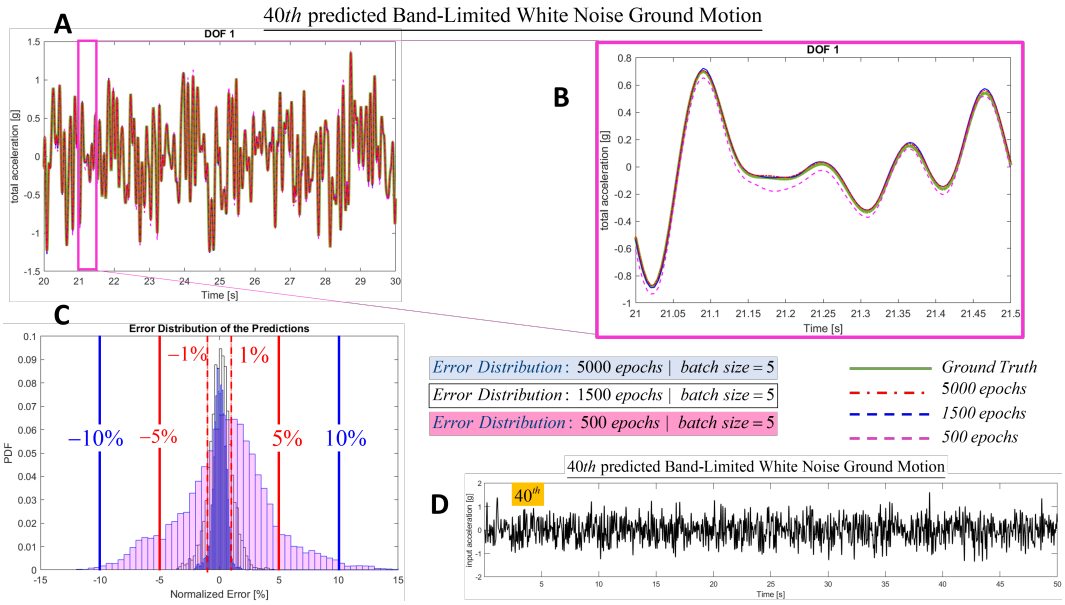
Total Acceleration [g] vs. Time [sec]: Ground Truth in GREEN (data generated using a numerical solver) | LSTM-Seq-2-Seq with 3000 epoch in RED | LSTM-Seq-2-Seq with 1500 epoch in dotted BLUE

4.0.4 *Results for the Nonlinear Inelastic Single Degree Of Freedom (SDOF) Structure*

We increase the level of nonlinearity with this SDOF system which in our case is nonlinear inelastic. we present two predictions with bandlimited white noise as an input excitation, which our model has never seen before (as it was the case for the previous estimations), and total acceleration as the output response. We have 50 of those predictions, which the 15th and the 40th are presented below. Image A show the predictions for 3 epochs categories (5000 epochs in dotted-red, 1500 epochs for dotted-blue, and 500 epoch in dotted-pink) with the green line being the ground truth value. Image B is a zoomed-in version of image A between 21 seconds and 21.5 seconds, and Image C shows the probability distribution function corresponding to the 3 epochs categories.

Image D shows the time series of the corresponding Band-limited White noise ground motion.

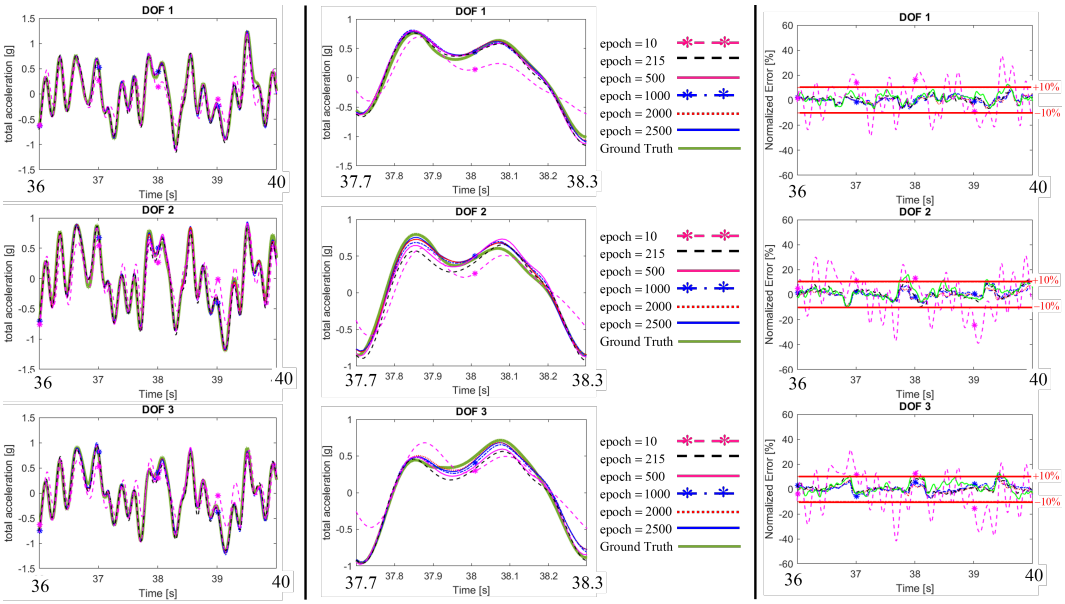




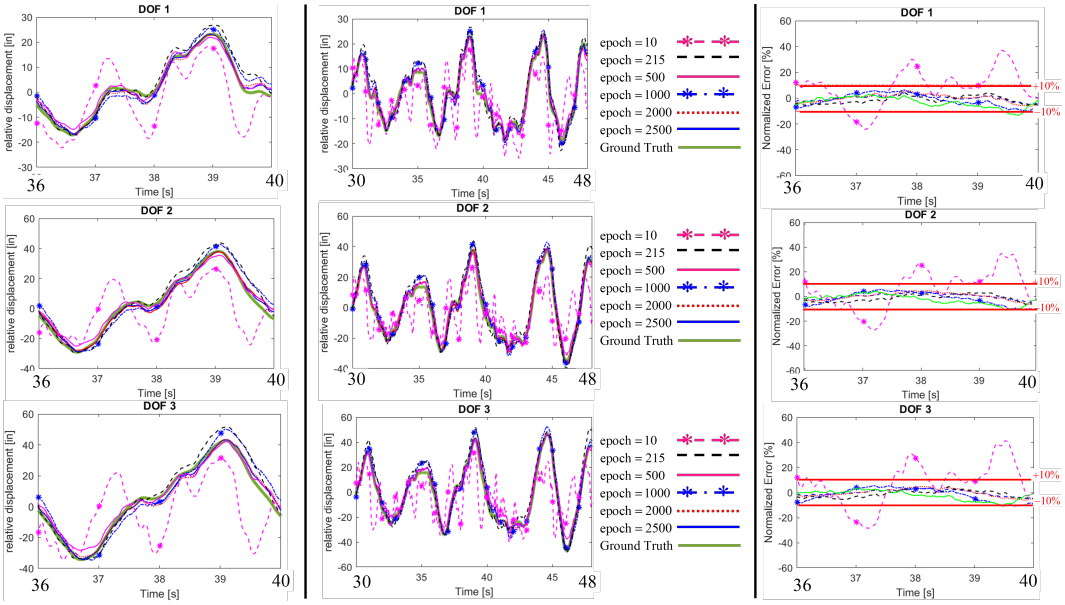
4.0.5 *Results for the Nonlinear Inelastic Multi Degree Of Freedom (MDOF) Structure*

The following images show the total acceleration response of the MDOF system at each floor (3 floors in our case). The first column shows the response of the structure when using a numerical solver

(Ground truth in green) and prediction when varying the number of epochs (dotted-pink-with-stars for 10 epochs, dotted-black-for 215 epochs, pink line for 500 epochs, dotted-blue-with-stars for 1000 epochs, dotted-blue for 2500 epochs). The second column shows a zoomed-in version (between 37.7 seconds to 38.3 seconds) of the first column. The third column is an error quantification of our response quantities.



We then present the prediction of another response quantity, which is the displacement responses at each DOF.

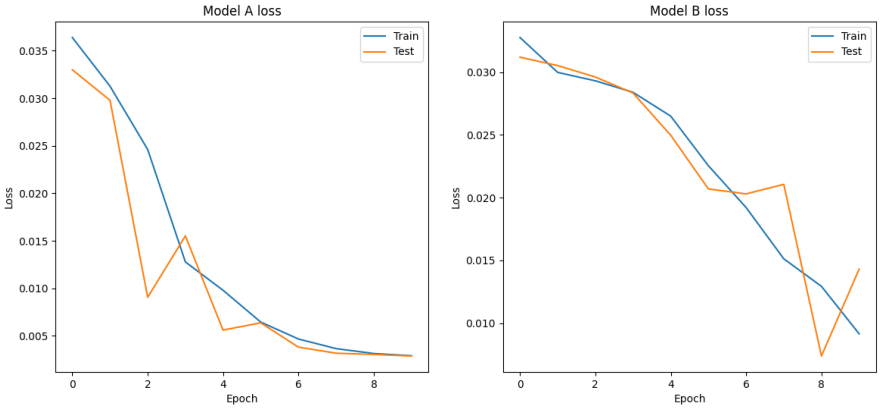


4.0.6 *A/B Testing*

To evaluate our Long Short-Term Memory (LSTM) network, we decided to conduct A/B testing where we compared two distinct LSTM models, namely 'A' and 'B.' The former was a more straightforward architecture with a single LSTM layer and ReLU activation functions; meanwhile, Model B was a bit more complex with two LSTM layers and Tanh activation functions. Model A had an LSTM layer with 100 units with a dense layer with 100 neurons, which would give us an output layer with the proper dimensions. The reason we designed Model A this way was to create a baseline

for performance and computational efficiency. When creating Model B, however, we had the two LSTM layers have 50 units each, combined with Tanh activation functions, and a dense layer with 50 neurons prior the the output layer. This model was made to use deeper sequential learning and different activation functions to potentially improve our prediction accuracy and the robustness of our model.

We decided to train and validate both models on the same dataset, where we had the train/test ratio of 70/30. For performance metrics, we used the Mean Squared Error (MSE) on the test data, training loss, and validation loss. We ran 10 epochs (as seen in the graph below) and observed that Model A shows a steady decrease in both training and validation losses, showing a final test loss of 0.00287 MSE. We saw very few fluctuations in the validation loss curve, which showed us that we had a stable learning process and effective generalization on unseen data. Model B also showed a consistent decrease in training and validation losses, but with slightly more fluctuation in the validation loss curve, which suggests that there may be some overfitting or even some sensitivity to our input data; our final test loss for Model B was 0.0143 MSE.



The analysis conducted with A/B testing shows that while both models do well in learning the nonlinear structural response patterns, Model A was overall more stable and had a lower validation loss, which tells us that it is a better choice for predicting seismic structural responses. This experiment teaches the importance of model architecture and activation function selection for use-case-specific LSTM network optimization.

5 Conclusion

The proposed LSTM architecture consists of several LSTM layers and fully connected layers to model the time dependency and causality of the input and output sequences. One scheme of LSTM was implemented, namely, the LSTM-f with different input/output formats. The performance of this scheme is illustrated via two studies: one involving a linear elastic 2D shear building model and the other a 2D shear building model with material nonlinearity. Both models simulate a moment-resisting frame system.

The technical challenges associated with implementing Long Short-Term Memory (LSTM) network-based surrogate modeling for nonlinear structural systems are multifaceted. Ensuring that the LSTM-based models accurately capture the behavior of the FE models across a range of conditions, including varying levels of ground motion intensities and selected testbed structures, poses a significant challenge. Optimizing these models to achieve substantial computational speedup while maintaining acceptable levels of accuracy requires meticulous design of the network architecture and training process. Additionally, ensuring that the LSTM-based models generalize well to unseen data and can effectively emulate the behavior of FE models for the selected testbed structures is crucial.

Data availability is another significant hurdle, as acquiring sufficient and relevant data for training and validating the models, especially given the scarcity of experimental data for seismic responses, is essential. Furthermore, ground motion characteristics exhibit regional variations due to diverse geological and tectonic conditions, which include differences in duration, intensity, frequency

content, and amplitude of seismic waves. Addressing this variability necessitates clustering ground motions based on their geographic regions to ensure the training data comprehensively captures these characteristics, leading to more accurate prediction outcomes. Lastly, the inherent damping ratio, a crucial parameter in structural dynamics, varies across different structures and significantly influences their dynamic response to seismic events. Effectively incorporating this variability into the training process is essential for accurately predicting the structural response to seismic loads.

Overall, the findings indicate that the proposed LSTM network is a promising, dependable, and computationally efficient technique for predicting nonlinear structural responses, with significant potential for assessing seismic risk and reliability analysis.

6 Code Availability

The implementation code for the LSTM-based surrogate modeling discussed in this paper is available on GitHub. This repository includes all scripts, data, and instructions to reproduce the experiments and results presented in this paper. You can access the code at the following link: https://github.com/ymorsi7/LSTMsForNonlinearStructuralSystems.

References

M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL `https://www.tensorflow.org/`. Software available from tensorflow.org.

H. Akaike. Fitting autoregreesive models for prediction. In *Selected Papers of Hirotugu Akaike*, pages 131–135. Springer, 1969.

J. M. W. Brownjohn and P.-Q. Xia. Dynamic assessment of curved cable-stayed bridge by model updating. *Journal of Structural Engineering*, 126(2):252–260, 2000. doi: 10.1061/(ASCE)0733-9445(2000)126:2(252). URL `https://ascelibrary.org/doi/abs/10.1061/%28ASCE%290733-9445%282000%29126%3A2%28252%29`.

A. Chopra. *Dynamics of Structures*. Pearson Education, 2012. ISBN 9780133072693. URL `https://books.google.com/books?id=eRcvAAAAQBAJ`.

M. Di Paola and G. Navarra. Stochastic seismic analysis of mdof structures with nonlinear viscous dampers. *Structural Control and Health Monitoring*, 16(3):303–318, 2009. doi: https://doi.org/10.1002/stc.254. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/stc.254`.

J. Durbin. Efficient estimation of parameters in moving-average models. *Biometrika*, 46(3/4): 306–316, 1959.

V. Ş. Ediger and S. Akar. Arima forecasting of primary energy demand by fuel in turkey. *Energy policy*, 35(3):1701–1708, 2007.

P. A. Fishwick. Neural network models in simulation: A comparison with traditional modeling approaches. In *Proceedings of the 21st conference on Winter simulation*, pages 702–709, 1989.

L. C. Jain et al. Recurrent neural networks: design and applications. Technical report, Crc Press, 2000.

D. P. Mandic and J. Chambers. *Recurrent neural networks for prediction: learning algorithms, architectures and stability*. John Wiley & Sons, Inc., 2001.

F. McKenna, M. H. Scott, and G. L. Fenves. Nonlinear finite-element analysis software architecture using object composition. *Journal of Computing in Civil Engineering*, 24(1):95–107, 2010. URL `https://doi.org/10.1061/(ASCE)CP.1943-5487.0000002`.

B. Moaveni, J. P. Conte, and F. M. Hemez. Uncertainty and sensitivity analysis of damage identification results obtained using finite element model updating. *Computer-Aided Civil and Infrastructure Engineering*, 24(5):320–334, 2009. doi: https://doi.org/10.1111/j.1467-8667.2008.00589.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8667.2008.00589.x.

C. Olah. Understanding lstm networks, 2015. URL http://colah.github.io/posts/2015-08-Understanding-LSTMs/.

S. E. Said and D. A. Dickey. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599–607, 1984.

E. SPACONE, F. C. FILIPPOU, and F. F. TAUCER. Fibre beam–column model for non-linear analysis of r/c frames: Part i. formulation. *Earthquake Engineering & Structural Dynamics*, 25(7): 711–725, 1996. doi: https://doi.org/10.1002/(SICI)1096-9845(199607)25:7<711::AID-EQE576>3.0.CO;2-9. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291096-9845%28199607%2925%3A7%3C711%3A%3AAID-EQE576%3E3.0.CO%3B2-9.

S.-B. Sun, Y.-Y. He, S.-D. Zhou, and Z.-J. Yue. A data-driven response virtual sensor technique with partial vibration measurements using convolutional neural network. *Sensors*, 17(12):2888, 2017.

D. Vamvatsikos and M. Fragiadakis. Incremental dynamic analysis for estimating seismic performance sensitivity and uncertainty. *Earthquake Engineering & Structural Dynamics*, 39(2):141–163, 2010. doi: https://doi.org/10.1002/eqe.935. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/eqe.935.

R.-T. Wu and M. R. Jahanshahi. Deep convolutional neural network for structural dynamic response estimation and system identification. *Journal of Engineering Mechanics*, 145(1):04018125, 2019.

S. Yousefianmoghadam, I. Behmanesh, A. Stavridis, B. Moaveni, A. Nozari, and A. Sacco. System identification and modeling of a dynamically tested and gradually damaged 10-story reinforced concrete building. *Earthquake Engineering & Structural Dynamics*, 47(1):25–47, 2018. doi: https://doi.org/10.1002/eqe.2935. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/eqe.2935.

J. Zhang, C. Wan, and T. Sato. Advanced markov chain monte carlo approach for finite element calibration under uncertainty. *Computer-Aided Civil and Infrastructure Engineering*, 28(7):522–530, 2013. doi: https://doi.org/10.1111/j.1467-8667.2012.00802.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8667.2012.00802.x.

R. Zhang, Z. Chen, S. Chen, J. Zheng, O. Büyüköztürk, and H. Sun. Deep long short-term memory networks for nonlinear structural seismic response prediction. *Computers Structures*, 220:55–68, 2019. ISSN 0045-7949. doi: https://doi.org/10.1016/j.compstruc.2019.05.006. URL https://www.sciencedirect.com/science/article/pii/S0045794919302263.