

Symbolic Neural Networks for Surrogate Modeling of Structures

Yiming Jia

PhD Candidate, Dept. of Civil and Env. Engineering, Northeastern University, Boston, United States

Mehrdad Sasaki

Professor, Dept. of Civil and Env. Engineering, Northeastern University, Boston, United States

Hao Sun

Associate Professor, Gaoling Sch. of Artificial Intelligence, Renmin University of China, Beijing, China

ABSTRACT: In order to evaluate urban earthquake resilience, reliable structural modeling is needed. However, detailed modeling of a large number of structures and carrying out time history analyses for sets of ground motions are not practical at an urban scale. Reduced-order surrogate models can expedite numerical simulations while maintaining necessary engineering accuracy. Neural networks have been shown to be a powerful tool for developing surrogate models, which often outperform classical surrogate models in terms of scalability of complex models. Training a reliable deep learning model, however, requires an immense amount of data that contain a rich input-output relationship, which typically cannot be satisfied in practical applications. In this paper, we propose model-informed symbolic neural networks (MiSNN) that can discover the underlying closed-form formulations (differential equations) for a reduced-order surrogate model. The MiSNN will be trained on datasets obtained from dynamic analyses of detailed reinforced concrete special moment frames designed for San Francisco, California, subject to a series of selected ground motions. Training the MiSNN is equivalent to finding the solution to a sparse optimization problem, which is solved by the Adam optimizer. The earthquake ground acceleration and story displacement, velocity, and acceleration time histories will be used to train 1) an integrated SNN, which takes displacement and velocity states and outputs the absolute acceleration response of the structure; and 2) a distributed SNN, which distills the underlying equation of motion for each story. The results show that the MiSNN can reduce computational cost while maintaining high prediction accuracy of building responses.

1. INTRODUCTION

To assess earthquake resilience for large-scale urban building clusters, there is a need for a detailed structural model of buildings to carry out reliable numerical simulations. Because buildings in an urban area experience different levels of ground motion (GM) severities (primarily due to fault rupture location and characteristics, soil condition, and the buildings' fundamental periods), the numerical simulations of a building need to be carried out under a series of selected GMs. More specifically, a considerable number of repeated numerical simulations are required to assess urban earthquake resilience, which is not practical. To address this issue, surrogate models can be used to expedite numerical simulations

while maintaining necessary engineering accuracy.

Surrogate models have drawn significant attention in civil, mechanical, and aerospace engineering, enabling computationally efficient analysis of complex structures. In performance-based structural engineering, surrogate models can promote efficient design, assessment, control, and optimization of engineering structures with reduced computational effort. Recent studies have shown that, owing to state-of-the-art advances in artificial intelligence, the use of deep learning (e.g., convolutional and recurrent neural networks) is a promising approach to establishing surrogate models for fast prediction of structural dynamic response (Wu and Jahanshahi, 2019; Oh

et al., 2020; Stoffel et al., 2020; Zhang et al., 2020a; Zhang et al., 2020b). Nonetheless, deep learning still has some limitations. Training a reliable deep learning model requires an immense amount of data that contains rich input-output relationships, which typically cannot be satisfied in most engineering problems. Commonly used nonlinear activation functions (e.g., sigmoid, hyperbolic tangent, and rectified linear unit) may drastically increase model complexity. Moreover, deep learning models are a “black box” and highly dependent on the representative quality of labeled data, leading to overfitting issues and limited extrapolation. Even with rich data, the resulting trained models are uninterpretable and may not make physical sense.

One approach to overcoming the limitations of deep learning is to develop a reduced-order surrogate model – model-informed symbolic neural network (MiSNN). MiSNN is essentially a symbolic neural network designed by leveraging domain-specific knowledge and fundamental principles of existing surrogate models (e.g., shear-beam for building seismic performance evaluation, Joyner and Sasani, 2020). MiSNN can discover the underlying closed-form formulations (differential equations) and accommodate use by engineers and practitioners who do not have knowledge of deep learning. Unlike deep learning models, symbolic neural networks use a combination of math operators as activation functions (e.g., absolute, sign, exponential, sinusoidal, cosine, square, cube, and multiplication). Recent studies show that symbolic neural networks are capable of finding parsimonious and interpretable mathematical expressions for generalized regression (Martius and Lampert, 2016; 2018; Kim et al., 2020). The domain-specific knowledge and fundamental principles of existing surrogate models are embedded in symbolic neural networks via variations of input data and selection of math operators. The embedded information can provide rigorous constraints to the parameters, alleviate overfitting issues, reduce the need for large training datasets, and thus, improve the robustness

of the trained MiSNN for more reliable prediction. Additionally, MiSNN can reduce the “black box” effect by making the model interpretable and providing physical meaning, thereby rendering it readily accessible for use by engineers and practitioners.

The underlying closed-form formulations discovered by MiSNN can be solved using numerical integration methods, such as the fourth order Runge-Kutta method (RK4, Dormand and Prince, 1980). The Runge-Kutta method is an effective and widely used method for solving initial-value problems of differential equations (Zheng and Zhang, 2017). Compared with Newmark and Euler methods, which have the maximum orders of accuracy as second and first, respectively, the Runge-Kutta method is easy to implement and can achieve a higher order of accuracy.

In this paper, the MiSNNs embedded into RK4 are used to discover the equations of motion for a multi-degree-of-freedom (MDOF) system under seismic excitation.

2. MODEL-INFORMED SYMBOLIC NEURAL NETWORK FOR APPROXIMATING THE EQUATION OF MOTION

The equation of motion for an MDOF system under seismic excitation can be written as

$$\ddot{\mathbf{U}} + f(\mathbf{U}, \dot{\mathbf{U}}) = -\mathbf{\Gamma}a_g \quad (1)$$

where $\mathbf{U}$, $\dot{\mathbf{U}}$, and $\ddot{\mathbf{U}}$ = vectors that represent the displacements, velocities, and accelerations of n DOFs relative to ground; a_g = GM acceleration time series; $\mathbf{\Gamma}$ = influence vector, a $n \times 1$ vector with each element (γ); and $f(\mathbf{U}, \dot{\mathbf{U}})$ = mass-normalized internal restoring force vector learned by MiSNNs, which has the same dimension as $\mathbf{U}$, $\dot{\mathbf{U}}$, and $\ddot{\mathbf{U}}$.

The $\mathbf{U}$, $\dot{\mathbf{U}}$, and signs of interstory drift and velocity ($\Delta\mathbf{U}$ and $\Delta\dot{\mathbf{U}}$) are included as inputs to MiSNN. A $n \times 1$ vector containing sinusoidal functions of time with a phase shift ($\sin(\mathbf{T})$) learned from time via a fully-connected layer with weights and bias) is also included in MiSNN to

detect possible time-variant structural properties. The output of MiSNN is the absolute acceleration of a DOF, $-\ddot{u} - \gamma a_g$. For an MDOF system, instead of using one MiSNN with n outputs, it is better to use the same number of MiSNNs as the number of DOFs (Chen et al., 2022a). Each MiSNN has sufficient flexibility to learn the contributions from all DOFs' displacement and velocity states to the absolute acceleration of each DOF.

The proposed MiSNN includes two layers of multiplication to produce a polynomial function up to fourth order to learn the potential complex input-output relationship (see Figure 1). Analogous to Long et al. (2019), Chen et al. (2022a), and Chen et al. (2022b), instead of fully connected layers, MiSNN allows inputs (and outputs of hidden layers) to pass directly to the following layers. This can significantly reduce the number of parameters while maintaining high accuracy, as observed from parametric studies.

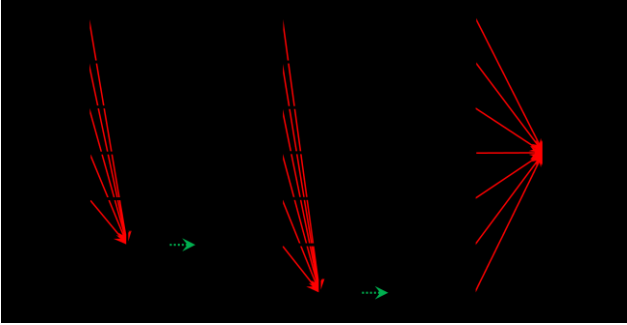


Figure 1: Standard MiSNN (solid thick black arrow = identity, solid thin red arrow = fully connected with weights and bias, dotted green arrow = multiplication).

In addition to domain-specific knowledge, the fundamental principle of shear-beam is also used to develop MiSNN. Shear-beam has been proven to be an effective approach to simplify building models for dynamic analysis while maintaining accuracy (Miranda and Taghavi, 2005; Khoshnoudian and Ehsan, 2013; Ganhavi et al., 2016; Escalona and Wong, 2018; Joyner and Sasani, 2020). This type of model idealizes each story's constitutive shear-drift relationship using

a single shear element, which can significantly reduce the computational cost (Joyner and Sasani, 2020). If a frame structure is considered as a shear-beam, the only inputs to MiSNN that are needed to estimate the absolute acceleration of a given DOF are those from that DOF and its adjacent DOFs. If the given DOF is either the first or last, the out-of-range states are assumed to be the same as the states of the given DOF.

3. FOURTH ORDER RUNGE-KUTTA INTEGRATION

The underlying closed-form formulations learned by MiSNNs can be treated as a system of second order differential equations, which can be solved using RK4. With displacement and velocity states at t_i ($\mathbf{U}_i$ and $\dot{\mathbf{U}}_i$), the states at t_{i+1} can be estimated as

$$\mathbf{U}_{i+1} = \mathbf{U}_i + \frac{1}{6}\Delta t(\mathbf{K}_1 + 2\mathbf{K}_2 + 2\mathbf{K}_3 + \mathbf{K}_4) \quad (2)$$

$$\dot{\mathbf{U}}_{i+1} = \dot{\mathbf{U}}_i + \frac{1}{6}\Delta t(\mathbf{L}_1 + 2\mathbf{L}_2 + 2\mathbf{L}_3 + \mathbf{L}_4) \quad (3)$$

where $\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_3$, and $\mathbf{K}_4$ = vectors consisting of slopes of displacement at each DOF (see Figure 2); $\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3$, and $\mathbf{L}_4$ = vectors consisting of slopes of velocity at each DOF (see Figure 2); and Δt = user-defined time interval.

4. MODEL TRAINING ALGORITHM

Since the underlying closed-form formulations discovered by MiSNNs are solved using RK4, the stability of RK4 must be considered. According to Hairer and Wanner (2010), the stability region of RK4 depends on Δt and the differential equations. More specifically, only a reasonably-selected Δt with well-trained MiSNNs can guarantee the stability of RK4. Training MiSNN independent to RK4 may trigger stability issues. To improve training efficiency, the MiSNNs are embedded into RK4 and trained using the Adam optimizer (Kingma and Ba, 2014) with the reducing learning rate on plateau strategy (i.e., reducing learning rate when the loss metric stops improving). The loss function is given by

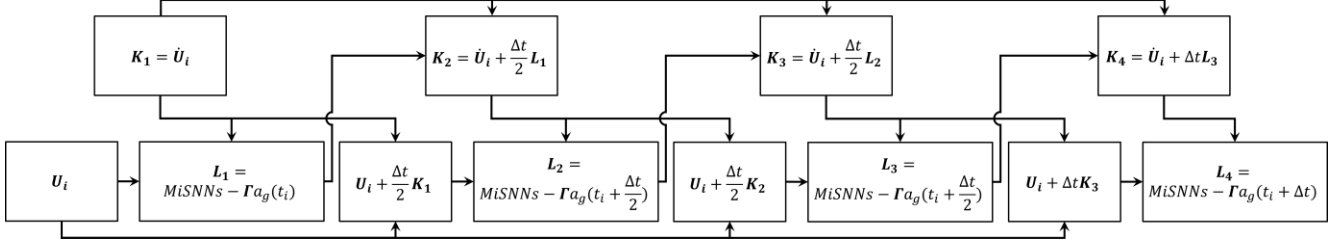


Figure 2: Slopes in RK4 ($\text{MiSNNs} = -f(\mathbf{U}, \dot{\mathbf{U}})$, $\Delta t = \text{time interval}$, $a_g(t_i + \frac{\Delta t}{2}) \approx \frac{a_g(t_i) + a_g(t_i + \Delta t)}{2}$).

$$\text{Loss} = \frac{1}{M} \sum_{j=1}^M \frac{1}{W} \sum_{i=1}^W (\mathbf{Y}_{ij} - \hat{\mathbf{Y}}_{ij})^2 \quad (4)$$

where $\mathbf{Y}_{ij}$ = matrix including all recorded $\mathbf{U}_i$ and $\dot{\mathbf{U}}_i$; $\hat{\mathbf{Y}}_{ij}$ = matrix including all $\hat{\mathbf{U}}_i$ and $\hat{\dot{\mathbf{U}}}_i$ estimated by MiSNNs embedded into RK4; W = window, which is the number of time steps used in RK4 calculation; and M = number of windows. In light of the different magnitudes of displacement and velocity, in Eq. (4), $\mathbf{U}_i$ and $\dot{\mathbf{U}}_i$ are assigned more weight by multiplying them by a factor of 10 to match the magnitudes of $\hat{\mathbf{U}}_i$ and $\hat{\dot{\mathbf{U}}}_i$. In addition to embedding MiSNN into RK4, during the training, Δt can be multiplied by factors of 0.5, 0.25, 0.125, and 0.065 to avoid a potential stability issue. The corresponding GM acceleration is linearly interpolated.

5. APPLICATIONS

The linear and nonlinear dynamic analysis results of 3-story detailed reinforced concrete (RC) buildings under a series of selected GMs are used to evaluate the proposed MiSNNs embedded into RK4.

5.1. Model and Data

5.1.1. 3D detailed reinforced concrete building model

A 3-story representative RC building located in Financial District, San Francisco, California, is designed according to ASCE 7 (ASCE, 2022) Design Level as a Risk Category II building. The typical floor plan of the designed building is shown in Figure 3. The story heights are 4.27 m and 3.66 m for the first story and all other stories, respectively. 3D detailed linear and nonlinear models of the designed building are developed

using OpenSees (McKenna et al., 2010). The linear model is developed using elastic beam-column elements. For the nonlinear model, distributed plasticity is accounted for using nonlinear beam-column elements with sections discretized into concrete core, concrete cover, and steel fibers. The buckling and bar-slip effects of reinforcing bars are also included. More details about building design and the corresponding 3D detailed model development can be found in Chan Esquivel et al. (2023).

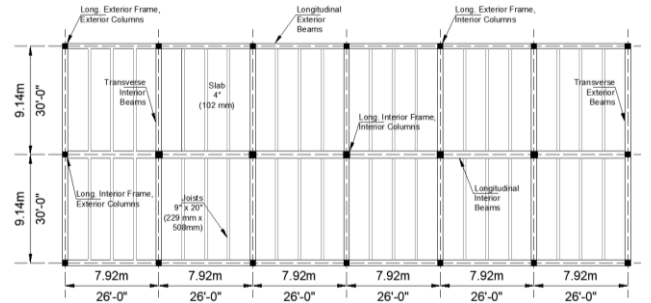


Figure 3: Plan view of the designed RC buildings

5.1.2. Ground motion selection and numerical simulation

Using spectral acceleration as the GM intensity measure, the GMs are selected using conditional spectra (Baker, 2011; Lin et al., 2013; Baker and Lee, 2018) at the building's fundamental period. 20 GMs are selected to represent the median and logarithmic standard deviation of the conditional spectrum (Chan Esquivel et al., 2023). Because pulse-type GM has a larger damage potential than ordinary GM, the selection procedure accommodates a portion of the selected GM records being pulse-type (NIST, 2011). The spectral acceleration of the 20 selected GMs and 4 (out of the 20) GM acceleration time series are

shown in Figures 4 and 5, respectively. The numerical simulations of 3D detailed linear and nonlinear models of the 3-story RC building under the selected GMs are carried out using OpenSees (McKenna et al., 2010). Seismic excitation is applied to the designed building in a direction parallel to the longitudinal frame.

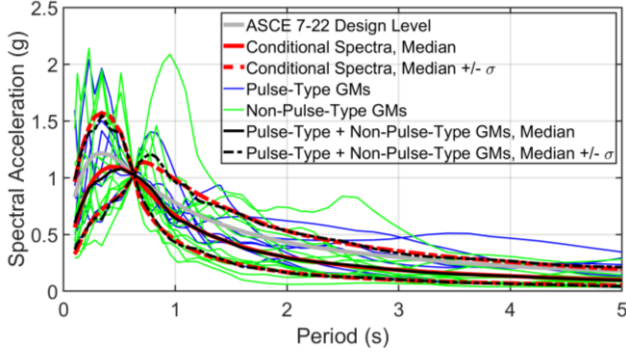


Figure 4: Spectral accelerations of the selected 20 GMs.

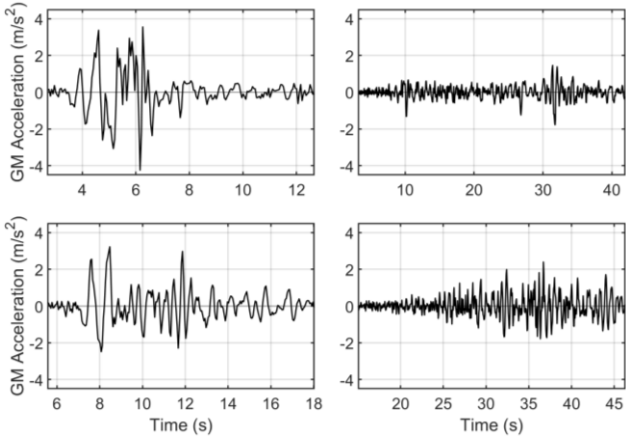


Figure 5: 4 GM acceleration time series.

5.2. Linear response estimation

The linear dynamic analysis results of the 3D detailed model of the 3-story RC building are used to train the MiSNNs embedded into RK4. The building responses and the corresponding GM accelerations are interpolated to a time interval of 0.05 s using a piecewise cubic hermite interpolating polynomial (Kreyszig, 2020). The training setups are listed in Table 1.

Table 1. Training setups

Training dataset	18 GMs ^a
Test dataset	2 GMs
Window size	25
Time in each window	0:0.05:1.20s ^b
Epochs	500
Initial learning rate	0.001
Initial learning rate decay	25% ^c

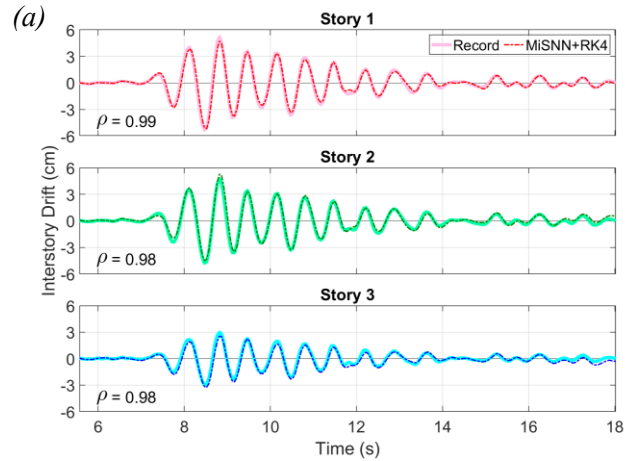
^a 18 GMs are randomly selected.

^b Using relative time in each window can result in higher model accuracy than using recorded time (as observed by parametric study).

^c The initial learning rate decays by 25% whenever the training loss stops decreasing for 4 epochs.

5.2.1. Standard MiSNN

The standard MiSNNs embedded into RK4 are trained using the setups specified in Table 1. The comparisons between the building responses estimated by these MiSNNs and those recorded by numerical simulation for one test GM are shown in Figure 6. The correlation coefficient (ρ) between these two building response estimates is used as a numerical goodness-of-fit measurement (see Figure 6). Both the graphical and numerical results indicate that the standard MiSNNs embedded into RK4 can capture the dynamic behaviors of 3-story RC buildings very well.



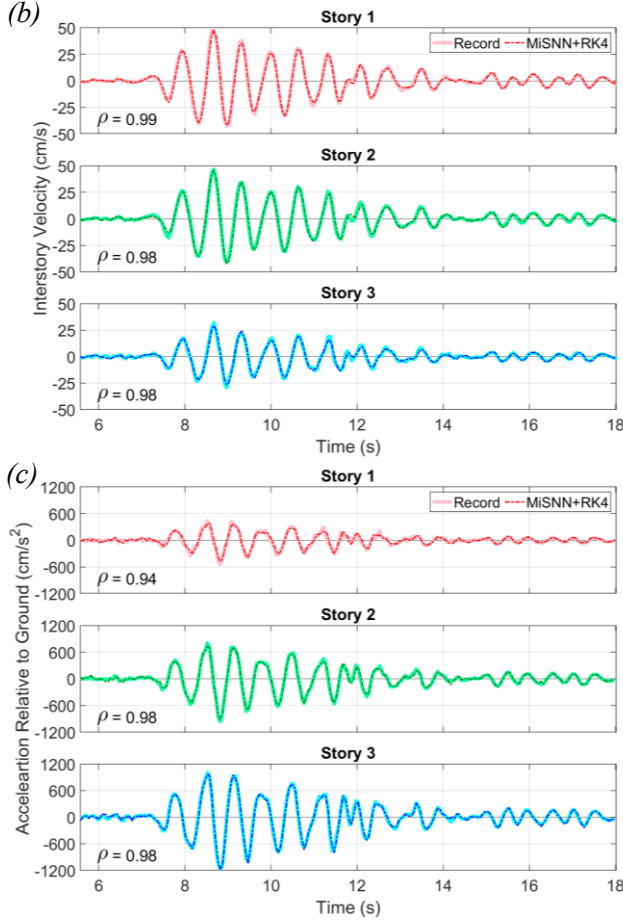


Figure 6: Comparison between the linear responses of the 3-story RC building under one test GM estimated by the standard MiSNNs embedded into RK4 and recorded by numerical simulation: (a) interstory drift, (b) interstory velocity, and (c) acceleration relative to ground.

5.2.2. Shear-beam-based MiSNN

Shear-beam-based MiSNNs embedded into RK4 are also trained following Table 1. The comparisons for one test GM show that these MiSNNs also can predict the linear dynamic responses of 3-story RC buildings very well (see Figure 7). The correlation coefficients obtained for shear-beam-based MiSNNs are slightly lower than those obtained using the standard MiSNNs. This difference is observed because the data used for shear-beam-based MiSNNs training are not as extensive as those used for standard MiSNNs training due to the out-of-range states.

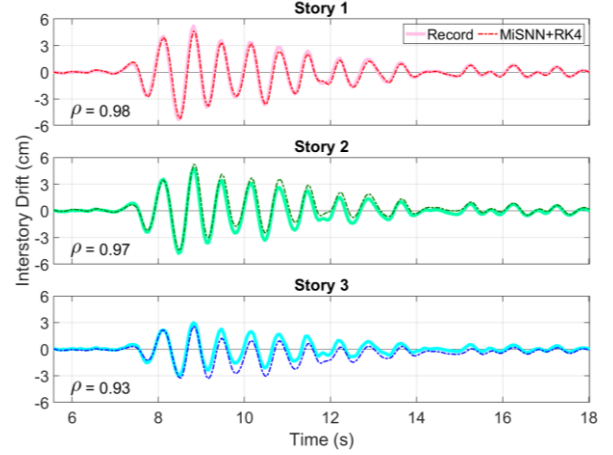


Figure 7: Comparison between the linear responses (interstory drifts) of the 3-story RC building under one test GM estimated by the shear-beam-based MiSNNs embedded into RK4 and recorded by numerical simulation.

5.3. Nonlinear response estimation

Considering the complexity of nonlinear behavior, two more multiplications and the sinusoidal functions of $\mathbf{W}\mathbf{U}_i + \mathbf{B}$ and $\mathbf{W}\dot{\mathbf{U}}_i + \mathbf{B}$ (where $\mathbf{W}$ and $\mathbf{B}$ are an $n \times n$ weight matrix and an $n \times 1$ bias vector, respectively, learned during MiSNN training) are included in the MiSNN architecture (see Figure 8). Both the standard and shear-beam-based MiSNNs are embedded into RK4 and trained using the 3-story RC building nonlinear dynamic responses obtained for 12 non-pulse-type GMs. This procedure was implemented because the pulse-type and non-pulse-type GMs have different damage potentials for RC buildings (NIST, 2011), and 13 sets of building responses under non-pulse-type GM (12 for training and 1 for test) can provide sufficient data for MiSNNs to learn the equation of motion. The training setups are the same as those shown in Table 1 but with one randomly selected GM for test purposes. Table 2 lists the improvements in model accuracy by comparing the nonlinear responses estimated using the MiSNNs designed for linear and nonlinear cases (see Figures 1 and 8). Although the model accuracy is improved for MiSNNs designed for nonlinear cases, the MiSNN shown in Figure 8 does not reach a high prediction accuracy (say $\rho > 0.95$), which

indicates that further studies on MiSNN architecture design are needed.

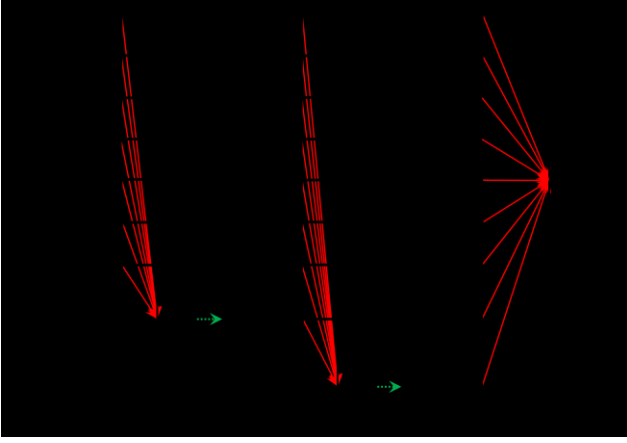


Figure 8: Standard MiSNN designed for nonlinear building response estimation (solid thick black arrow = identity, solid thin red arrow = fully connected with weights and bias, dotted green arrow = multiplication).

Table 2. Model accuracy improvement for the test GM

Response	Story	$\frac{(\rho_{NL} - \rho_L)^a}{\rho_L}$ (%)
Interstory drift	1	6
	2	11
	3	81
Interstory velocity	1	6
	2	9
	3	46

^a ρ_{NL} and ρ_L are the correlation coefficients between the nonlinear responses of the 3-story RC building under the test GM estimated by the standard MiSNNs designed for nonlinear (NL, see Figure 8) and linear (L, see Figure 1) cases embedded into RK4 and recorded by numerical simulation.

6. CONCLUSIONS

The MiSNN is designed by leveraging domain-specific knowledge and fundamental principles of existing surrogate models. Embedding MiSNN into RK4 can successfully establish a reduced-order surrogate model to estimate the response of MDOF systems under seismic excitation with low computational cost while maintaining high accuracy. Compared with deep learning models,

the architecture of MiSNN can significantly reduce model complexity while increasing model interpretability. MiSNNs embedded into RK4 are a user-friendly approach and can be implemented easily by engineers and practitioners.

The following conclusions are drawn based on the applications:

- Including the signs of inter-story drift and velocity as inputs and sinusoidal function of time with a phase shift in MiSNN architecture can improve the model accuracy.
- The shear-beam-based MiSNN can predict RC buildings' responses with high accuracy.
- Training MiSNN embedded into RK4 can effectively avoid a potential stability issue and speed the training process.
- Further studies on MiSNN architecture design are needed to improve the accuracy of nonlinear response estimation.

The MiSNN presented in this paper can also be solved using numerical integration methods other than RK4 and has the potential to be applied more widely to other engineering problems.

7. ACKNOWLEDGEMENTS

This paper is based upon research supported by the National Science Foundation under Grant No. CMMI-2053741.

8. REFERENCES

- ASCE. (2022). *Minimum design loads for buildings and other structures*, ASCE/SEI 7-22, Reston, Virginia.
- Baker, J. W. (2011). "Conditional mean spectrum: Tool for ground motion selection." *Journal of Structural Engineering*, ASCE, 137(3), 322–331.
- Baker, J. W., and Lee, C. (2018). "An improved algorithm for selecting ground motions to match a conditional spectrum." *Journal of Earthquake Engineering*, 2(4), 708–723.
- Chan Esquivel, S., Jia, Y., and Sasani, M. (2023). "Earthquake resilience of spatially distributed building clusters: Methodology and application." *Engineering Structures*, under review.

- Chen, Z., Liu, Y., and Sun, H. (2022a). "Symbolic deep learning for structural system identification." *Journal of Structural Engineering, ASCE*, 148 (9), 04022116: 1–14.
- Chen, Z., Liu, Y., and Sun, H. (2022b). "Forecasting of nonlinear dynamic based on symbolic invariance." *Computer Physics Communications*, 277, 108382: 1–16.
- Dormand, J. R., and Prince, P. J. (1980). "A family of embedded Runge-Kutta formulae." *Journal of Computational and Applied Mathematics*, 6(1), 19–26.
- Escalona, O., and Wong, J. (2018). "Rigid body response & performance based design of seismically isolated structures." *11th U.S. National Conference on Earthquake Engineering*, Los Angeles, California.
- Ganjavi, B., Hajirasouliha, I., and Bolourchi, A. (2016). "Optimum lateral load distribution for seismic design of nonlinear shear-buildings considering soil-structure interaction." *Soil Dynamics and Earthquake Engineering*, 88, 356–368.
- Hairer, E., and Wanner, G. (2010). *Solving ordinary differential equations II: Stiff and differential-algebraic problems*, 2nd Ed., Springer, Berlin.
- Joyner, M. D., and Sasaki, M. (2020). "Building performance for earthquake resilience." *Engineering Structures*, 210, 110371: 1–14.
- Khoshtoulian, F., and Ahmadi, E. (2013). "Effects of pulse period of near-field ground motions on the seismic demands of soil-MDOF structure systems using mathematical pulse model." *Earthquake Engineering Structural Dynamics*, 42, 1565–1582.
- Kim, S., Lu, P., Mukherjee, S., Gilbert, M., Jing, L., Ceperic, V., and Soljagic, M. (2021). "Integration of neural network-based symbolic regression in deep learning for scientific discovery." *IEEE Transaction on Neural Networks and Learning Systems*, 32 (9), 4166–4177.
- Kingma, D. R., and Ba, J. (2014). "Adam: A method for stochastic optimization." arXiv:1412.6980.
- Kreyszig, E. (2020). *Advanced engineering mathematics*, 10th Ed., Wiley, New York.
- Lin, T., Harmsen, S. C., Baker, J. W., and Luco, N. (2013). "Conditional spectrum computation incorporating multiple causal earthquakes and ground motion prediction models." *Bulletin of the Seismological Society of America*, 103(2A), 1103–1116.
- Long, Z., Lu, Y., and Dong, B. (2019). "PDE-Net 2.0: Learning PDEs from data with a numeric-symbolic hybrid deep network." *Journal of Computational Physics*, 399, 108925: 1–17.
- Martius, G., Lampert, C. H. (2016). "Extrapolation and learning equations." arXiv:1610.02995.
- McKenna, F., Scott, M. H., and Fenves, G. L. (2010). "Nonlinear finite-element analysis software architecture using object composition." *Journal of Computing in Civil Engineering, ASCE*, 24(1), 95–107.
- Miranda, E., and Taghavi S. (2005). "Approximate floor acceleration demands in multistory buildings I: formulation." *Journal of Structural Engineering, ASCE*, 131(2): 203–211.
- NIST. (2011). *Selecting and scaling earthquake ground motions for performing response-history analyses*, NIST, Gaithersburg, Maryland.
- Oh, B. K., Park, Y., and Park, H. S. (2020). "Seismic response prediction method for building structures using convolutional neural networks." *Structural Control and Health Monitoring*, 27(5), e2519: 1–17.
- Stoffel, M., Bamer, F., and Markert, B. (2020). "Deep convolutional neural networks in structural dynamics under consideration of viscoplastic material behaviour." *Mechanics Research Communications*, 108, 103565: 1–6.
- Wu, R.-T., and Jahanshahi, M. R. (2019). "Deep convolutional neural network for structural dynamic response estimation and system identification." *Journal of Engineering Mechanics, ASCE*, 145(1), 04018125: 1–25.
- Zhang, R., Liu, Y., and Sun, H. (2020a). "Physical-guided convolutional neural network (PhyCNN) for data-driven seismic response modeling." *Engineering Structures*, 215, 110704: 1–13.
- Zhang, R., Liu, Y., and Sun, H. (2020b). "Physics-informed multi-LSTM networks for metamodeling of nonlinear structures." *Computer Methods in Applied Mechanics and Engineering*, 369, 113226: 1–16.
- Zheng, L., and Zhang, X. (2017). *Modeling and analysis of modern fluid problems*, Academic Press, New York.