Title: **Sliding mode observer with ensemble learning for state estimation of high-rate dynamic systems**

Authors : Matthew Nelson
Colin Wiersma
Vahid Barzegar
Simon Laflamme
Chao Hu
Austin Downey
Jason Bakos
Jacob Dodson

ABSTRACT

High-rate systems are defined as engineering systems that typically undergo accelerations of amplitudes greater than 100 g over a duration of less than 100 ms. Examples include adaptive airbag deployment systems, hypersonic vehicles, and active blast mitigation systems. The use of feedback mechanisms in these high-rate applications is often critical in ensuring their continuous operations and safety. Of interest to this paper are algorithms enabling high-rate structural health monitoring (HRSHM) to empower sub-millisecond decision systems. HRSHM is a complex task, because high-rate systems are uniquely characterized by 1) large uncertainties in the external loads; 2) high levels of nonstationarities and heavy disturbances; and 3) unmodeled dynamics generated from changes in system configurations that necessitate careful crafting of adaptive strategies. Here, we study the implementation of two ensemble predictive models for implementation of HRSHM, intending to define fundamental mechanisms required in employing neural networks when sub-millisecond performance is required. One is based on a long short-term memory architecture, and the other is based on a single-layer wavelet neural network architecture. Numerical simulations are conducted using experimental data generated by high-rate mechanisms. A comparison of performance shows that, while the ensemble of wavelet neural networks is capable of faster predictions, the ensemble of long short-term memory networks provides enhanced signal forecasting, highlighting the important trade-off between computation speed and accuracy for HRSHM applications. It is also shown that the use of neuro-predictions as inputs to the model reference adaptive system instead of pure measurements produces faster convergence to the state estimate, yet at the cost of significantly higher computation time.

Matthew Nelson, Department of Civil, Construction, and Environmental Engineering, Iowa State University, Ames, IA, USA . Email: mjn5@iastate.edu; Internet: ...
Colin Wiersma, Department of Civil, Construction, and Environmental Engineering, Iowa State University, Ames, IA,
Vahid Barzegar, Department of Civil, Construction, and Environmental Engineering, Iowa State University, Ames, IA, USA
Simon Laflamme, Department of Civil, Construction, and Environmental Engineering, Iowa State University, Ames, IA, USA
Chao Hu, Department of Mechanical Engineering, Iowa State University, Ames, IA, USA
Austin Downey, Department of Mechanical Engineering, University of South Carolina, Columbia, SC,

INTRODUCTION

High-rate dynamic systems are defined as engineering systems experiencing accelerations of high amplitudes, typically higher than 100 g_n, over durations often less than 100 ms [1]. Examples of such systems include blast mitigation mechanisms, advanced weaponry, and hypersonic vehicles. The field deployment and safe operation of high-rate systems require feedback capabilities in the sub-millisecond range and high-rate state estimation capabilities termed high-rate structural health monitoring (HRSHM) [2]. However, the development of HRSHM algorithms is a complex task given the unique characteristics of their dynamics that combine 1) large uncertainties in the external loads, 2) high levels of non-stationarity and heavy disturbances, and 3) unmolded dynamics generated from changes in the system configurations [3]. The complexity of the high-rate systems combined with their fast-changing working environments limits the application of accurate real-time physical modeling techniques [4]. There have been recent research efforts in developing data-based algorithms applicable to HRSHM. Some examples of data-based algorithms used for HRSHM include an ensemble of long short-term memory (LSTM) networks presented by Barzegar *et al.* [5] and wavelet neural network (WNN) with adaptive input capabilities presented by Hong *et al.* [6].

LSTM networks have recently gained popularity for the modeling of complex dynamics. These networks store and utilize previous data through various memory blocks and equations known as gates to help to preserve key patterns in the data and combat the vanishing gradient problem [7, 8]. Because of their processing power, LSTMs perform well on nonlinear systems and for real-time analysis. For example, Li *et al.* [9] studied real-time applications of LSTMs by predicting automotive crash risks in urban areas, and Zhang *et al.* [10] used LSTMs to track and predict nonlinear seismic events. However, computation requirements of LSTMs typically requires a large amount of memory, yielding slower computation time. Faster neural networks are those with fewer weights and biases and simpler architectures. Of interest to this paper are single layer WNNs, known for their universal approximation capabilities [11], yielding rich capabilities in learning non-stationary signals [12].

Independent of the neural approach, a critical challenge with data-based techniques is that they do not provide any insights into the physical characteristics of the system. This impedes their applicabilities to HRSHM that requires on-time physical state-estimation to empower the decision system with actionable information. A solution is to integrate physical knowledge into the data-based approach, also known as physics-informed machine learning, or to simply use pure physical representation, yet typically at the cost of high computation time. The authors have previously proposed a Model Reference Adaptive System (MRAS)-based algorithm for HRSHM [13]. While demonstrated on a relatively simple dynamic system, the method was shown to be capable in the sub-millisecond range.

In this paper, we extend work on the MRAS-based algorithm in an attempt to further

USA

Jason Bakos, Department of Computer Science and Engineering, University of South Carolina, Columbia, SC, USA

Jacob Dodson, Air Force Research Laboratory, Munitions Directorate, Fuzes Branch, Eglin Air Force Base, FL, USA.

improve the convergence time of the state estimator. This is done by combining findings from previous work on an ensemble of LSTMs and WNNs to empower the MRAS algorithm with predicted time series data acting as the input. This architecture allows for predicted, instead of real-time, state-estimations resulting in quicker convergence times. Two different strategies are studied and compared. One uses an ensemble of LSTMs, the other uses an ensemble of WNNs. In prior work [14], the authors showed that using ensembles of short-sequence neural networks could dramatically accelerate computation time and enable multi-rate sampling to focus on unique features of the signal and thus predict nonlinearities more accurately. Ensemble learning has been extensively discussed in a general context by others [15].

BACKGROUND

This section provides an overview of the algorithmic architecture of the state estimator. First, the studied LSTMs and WNNs are used in an ensemble configuration, as illustrated in Figure 1. The ensemble consists of n neural cells, based on either a single layer LSTM cell [5], or a single layer WNN with Morlet activation function [6,16]. Each cell receives a unique input vector from the measured acceleration in the form of a delay vector.

$$\mathbf{x}_k^n = \begin{bmatrix} x_{k-d_i\tau_n} & x_{k-(d_n-1)\tau_n} & \cdots & x_{k-2\tau_n} & x_{k-\tau_n} \end{bmatrix} \tag{1}$$

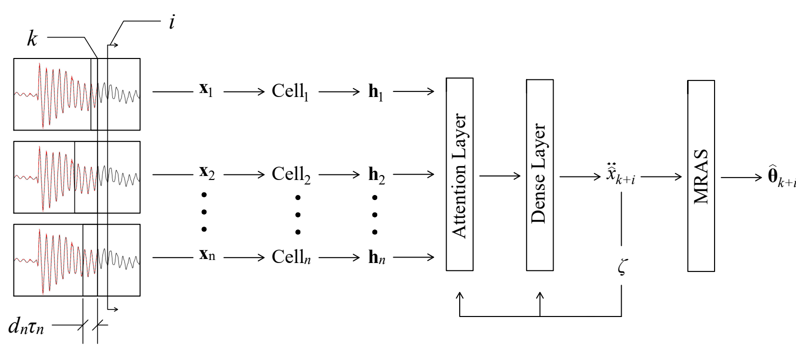


Figure 1. Ensemble architecture.

where τ and d are selected based on Takens' embedding theorem [17]. Each neural cell outputs a feature vector $\mathbf{h}$. After, each feature vector is assigned a weight using the attention layer. These weights are adapted in real-time to match the assembled features to those of the dataset [18]. Then, the dense layer produces the acceleration prediction $\ddot{x}_{k+i}$ where k is the current discrete time step, and i is the number of steps ahead. This ensemble architecture is presented in more detail in [5]. The neuro-prediction is back-propagated and weights adaptable based on a mean-squared error loss function and learning rate of ζ. Here, the prediction is used as an input to the MRAS algorithm.

The MRAS algorithm is used to convert acceleration predictions into physical state predictions, for example, predictions of stiffness and damping. Its architecture is illustrated in Figure 2 [13]. The adaptive system ($\dot{\hat{\mathbf{z}}}$) is an observer with the user-defined gain

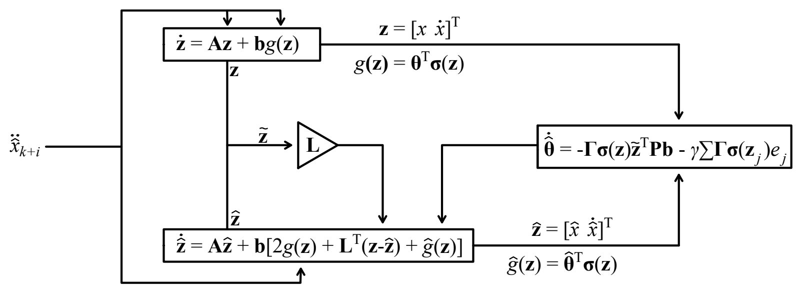


Figure 2. MRAS architecture.

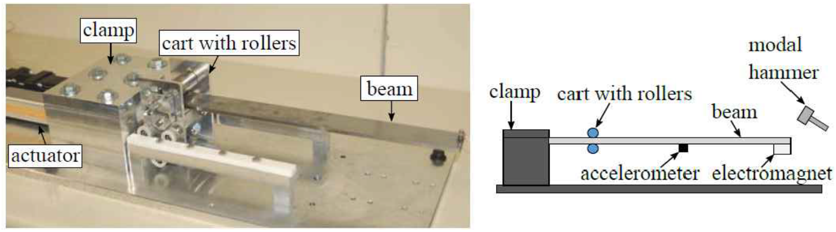
matrix $\mathbf{L}$ and is used to adapt the state estimation to obtain an estimated state vector $\hat{\mathbf{z}} = [\hat{x}\ \dot{\hat{x}}]^T$ that matches the reference model ($\dot{\mathbf{z}}$). The input is taken as unknown parameters $\boldsymbol{\theta} = [\theta_1\ \theta_2]$ and the basis function, $\boldsymbol{\sigma}(\mathbf{z})$, with $\hat{g}(\mathbf{z}) = \hat{\boldsymbol{\theta}}^T\boldsymbol{\sigma}(\mathbf{z})$ being the estimated input and $g(\mathbf{z}) = \boldsymbol{\theta}^T\boldsymbol{\sigma}(\mathbf{z})$ the true input. The reference model is a simplified physical representation of the system $\mathbf{Az}$, where $\mathbf{A}$ is the state space matrix, representing the known dynamics of the simplified reference model and $\mathbf{b}g(\mathbf{z})$, where $\mathbf{b}$ is the input matrix. The reference model state vector, $\mathbf{z}$ (i.e. true velocity and true displacement), was obtained through numerical integration of $\ddot{x}_{k+i}$ using the Newmark-Beta method with a recursive least squares (RLS) estimator to reduce drift in the displacement signal [19].

The simplified reference model assists the adaptive model in fine-tuning the unknown parameters, $\hat{\boldsymbol{\theta}} = [\hat{\theta}_1\ \hat{\theta}_2]$ to obtain the true values. The unknown parameters $\hat{\boldsymbol{\theta}}$ are tuned via sliding mode theory based on the state estimation error, $\tilde{\mathbf{z}} = \mathbf{z} - \hat{\mathbf{z}}$, and tracking error, $e_j = g(\mathbf{z}_j) - \hat{\boldsymbol{\theta}}^T\boldsymbol{\sigma}(\mathbf{z}_j)$. A particularity of the algorithm is that it uses concurrent learning to cope with the lack or persistence in the excitation, with J being the size of the history stack and j denotes the position within the history stack. Hyperparameters, $\boldsymbol{\Gamma}$, γ, and $\mathbf{P}$ are the learning rate matrix, learning rate, and user-defined positive-definite matrix, respectively. The algorithm is presented in more detail in [13].

METHODOLOGY

DROPBEAR Experimental Testbed

The vibration dataset obtained form the Dynamic Reproduction of Projectiles in Ballistic Environments for Advanced Research (DROPBEAR), shown in Figure 3, was used for validation [20]. DROPBEAR is a cantilever beam with repeatable and controllable changes in the dynamics mimicking sudden structural changes that occur in high-rate systems. This is done through a 1) sudden change of boundary condition produced using a moving cart, or 2) sudden change in mass produced using an electromagnet-activated mass drop, or 3) modal hammer impacts at the free end of the cantilever. Data analyzed in this paper focuses uniquely on changing boundary condition using static cart positions with the cart located 50, 100, 150, and 200 mm from the clamp. A modal hammer impact is used to excite the beam during the experiments. Measurements were filtered using a band pass filter of 5 Hz and 100 Hz cutoff frequencies.



(a) (b)

Figure 3. (a) DROPBEAR experimental testbed; and (b) schematic the testbed with key components [13].

Neural Architectures

Using the known physics of the system, four cells ($n = 4$) were selected to construct the ensemble, with each cell representing a different static cart position. Each WNN and LSTM cell, both with a single hidden layer, were trained offline to model the dynamics of a single degree-of-freedom system with an oscillating frequency equal to that of DROPBEAR under each considered static cart position, i.e. 17.8, 21, 25, and 31 Hz. The WNN had adaptive nodal weights, centers, and bandwidths, with the learning rates for the nodal centers and bandwidths kept constant at 1e-4 and 1e-1. The rest of the offline hyperparameters are listed in Table I

TABLE I. OFFLINE HYPERPARAMETERS.

Cell	Cart Position (mm)	Frequency (Hz)	τ (steps)	d	LSTM Learning rate	WNN Learning rate (weights)
1	50	17.8	332	2	0.015	0.9
2	100	21	275	3	0.015	0.2
3	150	25	220	4	0.015	0.9
4	200	31	183	2	0.02	0.5

RESULTS

For this preliminary investigation, damping (θ_2) is assumed available. Figures 4(a) and (b) plot the typical acceleration signals of DROPBEAR, taken with the cart positioned at 150 mm away from the clamp and under free vibration after being struck by the modal hammer. These plots compare the response estimated by the LSTMs (Figure 4(a)) and WNNs (Figure 4(b)) for a 300 step-ahead (0.12 ms) prediction. The DROPBEAR datasets are taken at a sampling frequency of 25,600 samples per second. A moving average filter was applied with a size of 0.1% of the sampling frequency to reduce the prediction noise. The MRAS inputs are these filtered prediction, and is constructed with the following hyperparameters: $\mathbf{\Gamma} = [.003\ 0]$, $\gamma = 1$, $\mathbf{L} = [1,000\ 50]$, and $J = 20$. Results show that the ensembles quickly catch up to the true values as the initial acceleration damps. For both the LSTMs and WNNs, the neural networks converge to the true values after damping the initial conditions. Figures 4(c) and (d) plot the stiffness estimated by the MRAS algorithm using predictions from LSTM and WNN, respectively. Both neural networks yield fast convergence, where convergence is taken when the oscillations remain within 10% of the average value.

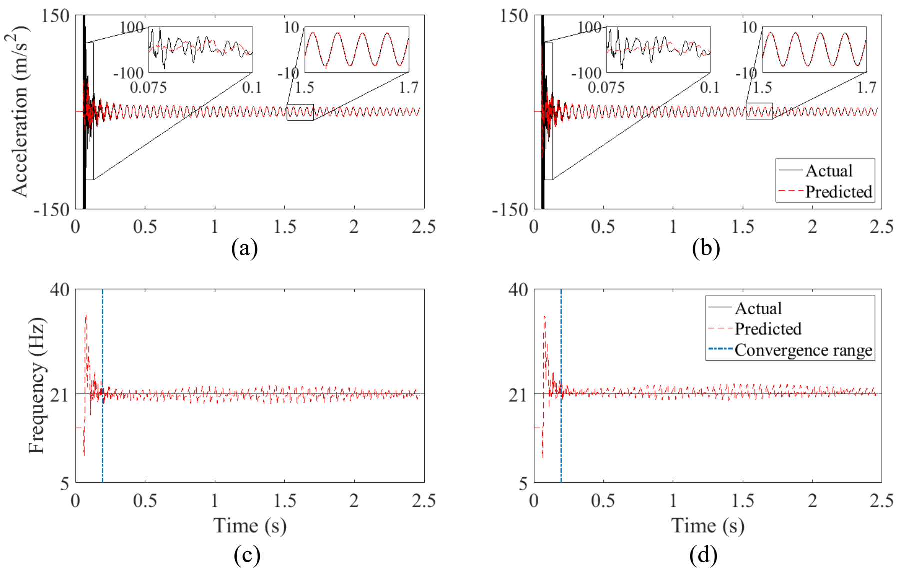


Figure 4. Time series responses of DROPBEAR comparing acceleration state estimation for the LSTMs (a) and WNNs (b), along with the corresponding estimated stiffnesses by MRAS using the predicted output from the LSTMs (c) and WNNs (d).

Table II summarizes results from all cart positions, listing three performance metrics: the average computation time per time step, the MRAS convergence time for the stiffness, and the root mean squared error (RMSE) on the acceleration prediction. The computation time is measured as the amount of time in milliseconds to perform signal prediction and state estimation of a single input vector. Performance is assessed for both the LSTMs and WNNs, and also for the MRAS using pure measurements instead of neuro-predictions (labeled "MRAS"). Results show the LSTMs are more accurate (RMSE) than the WNNs at forecasting the dataset, but require computation time approximately 50% longer. Following that, the LSTMs yield faster convergence speed than the WNN under every case except for the 100 mm cart position, likely attributed to increased accuracy of the ensemble of LSTM's predictions. Also, a cross-comparison between the neuro-predictors and the pure MRAS reveals that the use of a neuro-prediction as an input to the MRAS ("LSTM" or "WNN") produces faster convergence to the state estimate under all cart positions, yet yielding significantly higher computation time.

TABLE II. PERFORMANCE RESULTS

	50 mm			100 mm			150 mm			200 mm		
	LSTMs	WNNs	MRAS	LSTMs	WNNs	MRAS	LSTMs	WNNs	MRAS	LSTMs	WNNs	MRAS
Runtime/step (ms)	6.40	3.67	0.92	5.94	3.60	0.92	6.35	3.60	0.92	6.53	3.62	0.93
k convergence (ms)	169	193	207	200	198	211	99	108	117	112	127	130
RMSE (m/s^2)	7.87	7.89	N/A	9.15	14.52	N/A	7.67	16.9	N/A	7.26	10.7	N/A

CONCLUDING REMARKS

This paper evaluated the use of neuro-predictors to accelerate the convergence of an adaptive physical model known as the Model Reference Adaptive System (MRAS) to estimate the stiffness of a dynamic system, with applications to HRSHM. The study compared the performance of ensembles of long short term memory (LSTMs) neural networks and wavelet neural networks (WNNs) in predicting an acceleration signal. Those predictions were used as step-ahead predictions in the MRAS model. Validation was conducted on experimental data collected from the Dynamic Reproduction of Projectiles in Ballistic Environments for Advanced Research (DROPBEAR) testbed. The ensemble was constructed using four cells, each associated with a different DROPBEAR dynamics corresponding to the four cart positions under study. Offline training was conducted using simplified single-degree-of-freedom systems of fundamental frequency consistent with that of DROPBEAR under each cart position. Results show that adding predicted values from the ensemble networks to the MRAS produced quicker convergence to the estimated state under all cart positions, but adding substantial computation time. Comparing the neuro-predictors, the LSTMs yielded faster convergence and prediction accuracy compared to the WNNs, yet the WNNs yielded faster computation speed. This result demonstrates a key trade-off in high-rate structural health monitoring (HRSHM) between computation time and accuracy.

Future work is to include improved algorithmic capabilities for pure acceleration feedback and dynamic changes in states. The physics-informed machine learning algorithm's capabilities will be challenged to augment offline learning techniques, and online learning capabilities will be improved to perform multi-step ahead state estimation of more complex systems.

ACKNOWLEDGMENTS

The work presented in this paper is funded by the National Science Foundation under award number CISE-1937460. Their support is gratefully acknowledged. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsor. The authors also acknowledge Dr. Janet Wolfson and Dr. Jonathan Hong for providing the experimental data.

REFERENCES

1. Dodson, J., B. Joyce, J. Hong, S. Laflamme, and J. Wolfson. 2017. "Microsecond State Monitoring of Nonlinear Time-Varying Dynamic Systems," in *Volume 2: Modeling, Simulation and Control of Adaptive Systems; Integrated System Design and Implementation; Structural Health Monitoring*, American Society of Mechanical Engineers, doi:10.1115/smasis2017-3999.
2. Dodson, J., A. Downey, S. Laflamme, M. Todd, A. G. Moura, W. Yang, Z. Mao, P. Avitabile, and E. Blasch. 2021. "High-Rate Structural Health Monitoring and Prognostics: An Overview," *Civil, Construction and Environmental Engineering Conference Presentations andProceedings*.

3. Hong, J., S. Laflamme, J. Dodson, and B. Joyce. 2018. "Introduction to State Estimation of High-Rate System Dynamics," *Sensors*, 18(2):217, doi:10.3390/s18010217.
4. Downey, A., J. Hong, J. Dodson, M. Carroll, and J. Scheppegrell. 2020. "Millisecond model updating for structures experiencing unmodeled high-rate dynamic events," *Mechanical Systems and Signal Processing*, 138:106551, doi:10.1016/j.ymssp.2019.106551.
5. Barzegar, V., S. Laflamme, C. Hu, and J. Dodson. 2021. "Multi-time Resolution Ensemble LSTMs for Enhanced Feature Extraction in High-Rate Time Series," *Sensors*.
6. Hong, J., S. Laflamme, L. Cao, J. Dodson, and B. Joyce. 2018. "Variable input observer for nonstationary high-rate dynamic systems," *Neural Computing and Applications*, 32(9):5015–5026, doi:10.1007/s00521-018-3927-x.
7. DiPietro, R. and G. D. Hager. 2020. "Deep learning: RNNs and LSTM," in *Handbook of Medical Image Computing and Computer Assisted Intervention*, Elsevier, pp. 503–519, doi: 10.1016/b978-0-12-816176-0.00026-0.
8. Hassaballah, M. 2020. *Deep learning in computer vision : principles and applications*, CRC Press, Taylor and Francis, Boca Raton, ISBN 9781351003827.
9. Li, P., M. Abdel-Aty, and J. Yuan. 2020. "Real-time crash risk prediction on arterials based on LSTM-CNN," *Accident Analysis & Prevention*, 135:105371, doi: 10.1016/j.aap.2019.105371.
10. Zhang, R., Z. Chen, S. Chen, J. Zheng, O. Büyüköztürk, and H. Sun. 2019. "Deep long short-term memory networks for nonlinear structural seismic response prediction," *Computers & Structures*, 220:55–68, doi:10.1016/j.compstruc.2019.05.006.
11. Zhang, Q. and A. Benveniste. 1992. "Wavelet networks," *IEEE Transactions on Neural Networks*, 3(6):889–898, doi:10.1109/72.165591.
12. Zhang, D. 2019. "Wavelet Transform," in *Texts in Computer Science*, Springer International Publishing, pp. 35–44.
13. Yan, J., S. Laflamme, J. Hong, and J. Dodson. 2021. "Online parameter estimation under non-persistent excitations for high-rate dynamic systems," *Mechanical Systems and Signal Processing*, 161:107960, doi:10.1016/j.ymssp.2021.107960.
14. Barzegar, V., S. Laflamme, C. Hu, and J. Dodson. 2022. "Ensemble of recurrent neural networks with long short-term memory cells for high-rate structural health monitoring," *Mechanical Systems and Signal Processing*, 164:108201, doi:10.1016/j.ymssp.2021.108201.
15. Zhou, Z.-H. 2021. "Ensemble Learning," in *Machine Learning*, Springer Singapore, pp. 181–210.
16. Shao, H., M. Xia, J. Wan, and C. D. Silva. 2021. "Modified Stacked Auto-encoder Using Adaptive Morlet Wavelet for Intelligent Fault Diagnosis of Rotating Machinery," *IEEE/ASME Transactions on Mechatronics*:1–1, doi:10.1109/tmech.2021.3058061.
17. Takens, F. 1981. "Detecting strange attractors in turbulence," in *Dynamical systems and turbulence, Warwick 1980*, Springer, pp. 366–381.
18. Chowdhury, F. A. R. R., Q. Wang, I. L. Moreno, and L. Wan. 2018. "Attention-Based Models for Text-Dependent Speaker Verification," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, doi:10.1109/icassp.2018.8461587.
19. Zheng, W., D. Dan, W. Cheng, and Y. Xia. 2019. "Real-time dynamic displacement monitoring with double integration of acceleration based on recursive least squares method," *Measurement*, 141:460–471, doi:10.1016/j.measurement.2019.04.053.
20. Joyce, B., J. Dodson, S. Laflamme, and J. Hong. 2018. "An Experimental Test Bed for Developing High-Rate Structural Health Monitoring Methods," *Shock and Vibration*, 2018:1–10, doi:10.1155/2018/3827463.