

LA-UR-21-32326

Approved for public release; distribution is unlimited.

Title: Structural Health Monitoring: Using an Autoencoder to Identify Damage in a Bolted Joint

Author(s): Schelle, Caleb Van

Intended for: Capstone report for UCSD master's degree program

Issued: 2021-12-17



Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by Triad National Security, LLC for the National Nuclear Security Administration of U.S. Department of Energy under contract 89233218CNA000001. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

Structural Health Monitoring: Using an Autoencoder to Identify Damage in a Bolted Joint

Caleb Schelle

SE296: UCSD Capstone Project

December 9, 2021

Introduction

Los Alamos National Laboratory (LANL) is an important fixture in the United States Department of Energy's (DOE) National Nuclear Security Agency (NNSA) complex. LANL is one of the largest national laboratories in the country, and the laboratory's primary mission is to support the nation's nuclear stockpile. The lab functions as a design agencies for the NNSA and performs extensive testing on weapons as part of that mission. The shock and vibration test team at LANL utilizes electrodynamic shaker systems for important qualification testing in support of the laboratory's mission.

Modern engineering relies heavily on bolted joints to connect two objects. During these shaker tests, engineers depend on bolted joints to secure the test article to a fixture and the fixture to the table. The test article may be hazardous and contain high explosives which could create a safety issue if a bolted joint failed. A loss of preload in a bolt will affect the way energy is input to the system and may introduce nonlinearities as the joint opens and closes. The loss of preload can create challenges controlling the test and make acquired signals useless. Being able to monitor preload within bolted joints during testing can improve the quality of the data and keep workers safe.

Operational Evaluation

Bolts can lose preload in a variety of ways during dynamic environments, but sometimes, they never even achieve the designed preload. In the case of test engineering, an article will be subjected to numerous environments in three orthogonal directions requiring frequent reconfiguration of the article. This handling is done by engineers and technologists in the laboratory space, and forgetting to torque a bolt is a frequent concern. A typical fixture may have 20 bolts securing the fixture to the table and another dozen mounting the test article to the fixture. As two or even three people reconfigure the test article between axes, there is a chance that not all the bolts will be torqued and some may only be hand-tight.

Another common mechanism for missing preload comes from the torquing process. Measuring the preload in a bolt is not a simple task and requires advanced sensors. To circumvent instrumenting every bolt in a fixture, the shock and vibe team relies on applying a torque

specification correlated to the desired preload. There are many factors that may affect an applied torque load such as material, corrosion, or thread lubrication. The most common tool for torquing bolts is a torque wrench which has an associated error with it, usually 3-5%, and can result in a bolt with insufficient preload. An insufficiently torqued bolt may continue to lose preload under dynamic loading.

The experiments described in this report were not subject to common data acquisition constraints. The tests were conducted in an indoor, climate-controlled laboratory space. This laboratory space is not permitted to contain hazardous articles, so there were no safety concerns. The data acquisition unit used for testing was a Siemens device and had sufficient channels for the sensors used. The sensors were available from the shock and vibration team's inventory, but were not calibrated at the time of testing.

However, during an actual shaker test, accelerometers mounted near bolted joints on the test article and fixture will be subjected to operational environments more extreme than those completed during this benchtop testing. In these cases, the sensors will need to withstand the conditions while still taking measurements and transmitting data. The shakers used for qualifying complete systems are capable of high displacement, velocity, and acceleration outputs, which sensors will experience if located on the armature or any other portion of the moving mass. Some shock and vibration tests are performed while test articles are thermally conditioned. These combined thermal-mechanical tests may be run at temperatures between -54°C and 74°C. Sensors located on the moving mass of the shaker or the slip table would need to be able to operate under these temperature conditions.

Experimental Set-up

Testing was performed on a complete BARC (Box Assembly with Removable Component) structure manufactured at Los Alamos National Laboratory. The complete BARC structure is pictured in Figure 1. The removable component is bolted to two C channel components, each with a single #10-32 threaded fastener torqued to 5.7 Newton-meters (50 inch-pounds). Each C channel component is bolted to the box assembly with four #6-32 threaded fasteners torqued to a specified 2.3 Newton-meters (20 inch-pounds), seen in Figure 2. The bolted joint between the box assembly and the C channel is the joint being studied. The BARC was instrumented with two triaxial PCB 356A32 accelerometers, one on either side of one C channel/box assembly joint. These locations were chosen to monitor the transmissibility across the bolted joint and monitor for changes in the signal indicating a loose fastener.

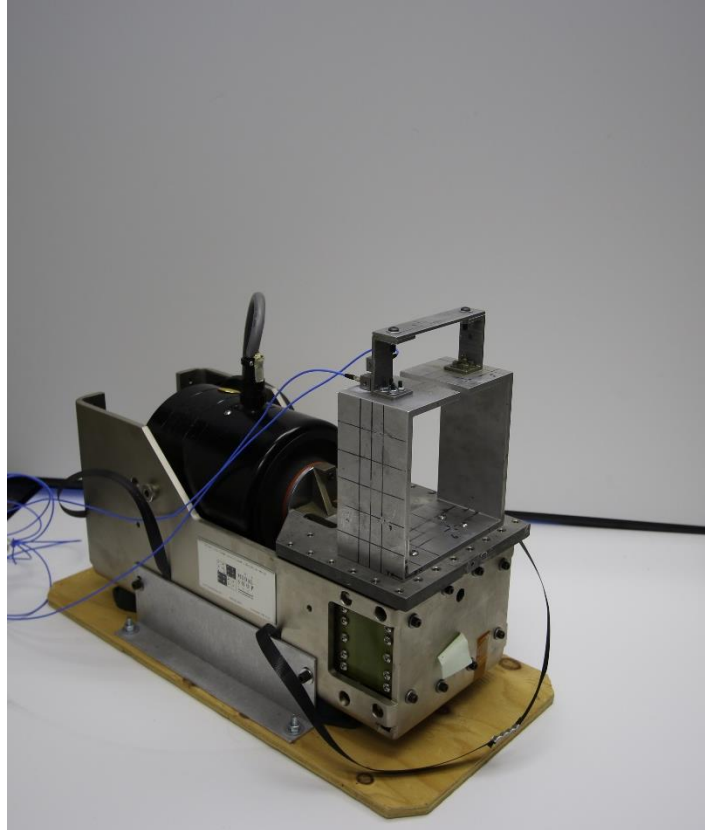


Figure 1 – BARC assembly mounted to slip table

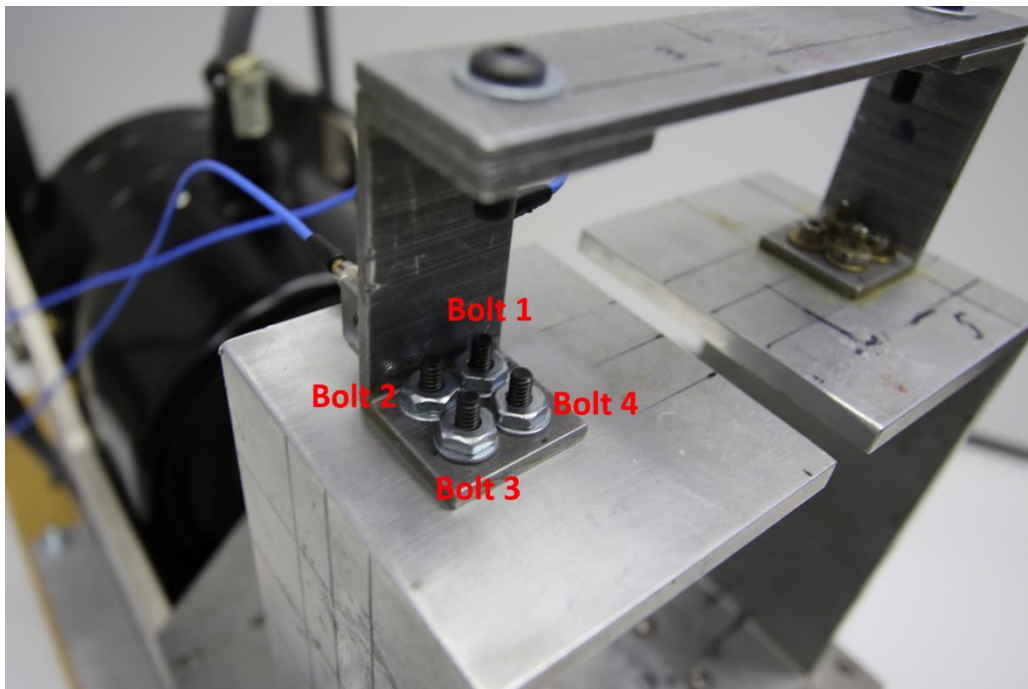


Figure 2 – Bolted joint being studied and bolt numbers for identification

The input to the shaker was a flat, broadband random excitation with 1g RMS. The input frequency range was defined over 10-3000 Hz with a constant amplitude of 0.004 g²/Hz. This broadband profile was chosen to excite non linearities within the bolted joint and is pictured below in Figure 3.

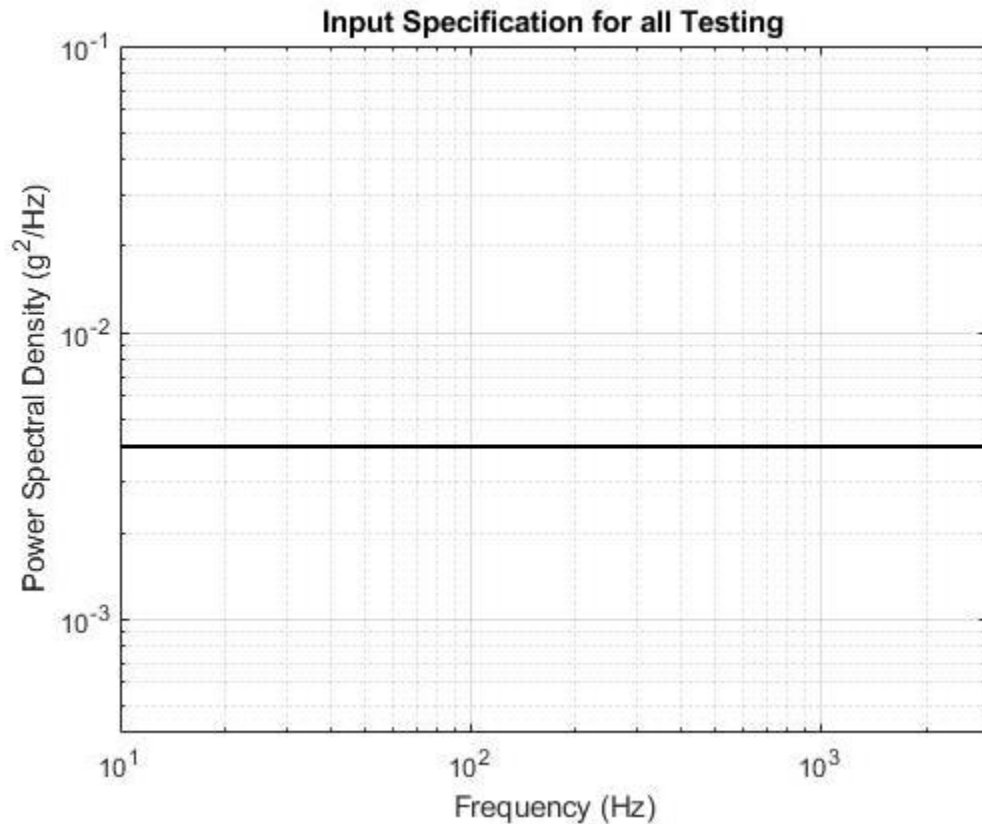


Figure 3 - Shaker input to excite BARC structure.

The damage cases are listed below in Table 1. The types of damage being studied in this experiment are loose and missing bolts. The table displays the percentage of the specification each bolt was torqued to for each individual case. The torque specification for all four bolts is the same, 2.3 Newton-meters (20 in-lbs) and the 50% torque value is 1.15 Newton-meters (10 in-lbs). The reduced torque was applied by loosening the bolt completely, then re-tightening to the damaged specification.

Table 1. Definition of Damage Cases.

Case ID	Bolt 1 Torque	Bolt 2 Torque	Bolt 3 Torque	Bolt 4 Torque
1 (Healthy)	100%	100%	100%	100%
2	100%	100%	100%	50%
3	100%	100%	100%	0% (removed)
4	100%	100%	50%	50%

5	100%	50%	50%	50%
6	50%	50%	50%	50%
7	50%	50%	100%	100%
8	50%	50%	50%	100%

Data Acquisition

Shaker

The shaker used to excite the BARC structure was the Modal Shop model 2075E, a 334 Newton (75 pound force) electrodynamic shaker. It is rated from 8-5000Hz and has a peak force output of 334 Newtons (75 pounds) [1]. The shaker was fixed to the K2075E-HT, a horizontal slip table 152.4 millimeters (6 inches) long and 190.5 millimeters (7.5 inches) wide. The slip table has a 25.4 millimeter (1 inch) grid hole pattern for attaching fixtures or test articles. The maximum acceleration of the bare slip table is 196 m/s^2 (20 g) and is reduced when a test article is mounted to the table [2]. The shaker and slip table are pictured in Figure 1.

Sensors

The experimental setup consisted of three accelerometers. One uniaxial accelerometer, an Endevco 2250AM1-10, was located at the front edge of the slip table farthest from the shaker connection. The uniaxial accelerometer was only used to control the experiment and no measurements were used in the model for determining the structure's health. The uniaxial accelerometer had a nominal sensitivity of 1.02 mV/m/s^2 (10 mV/g) and was secured to the slip table with superglue and allowed 24 hours to cure.

Two triaxial accelerometers were located on either side of the bolted joint being studied as seen in Figure 4. Both triaxial accelerometers were PCB model 356A32 with nominal sensitivities of 10.2 mV/m/s^2 (100mV/g) in each axis. Both triaxial accelerometers were mounted to the BARC structure after the surface was abraded with 220 grit sandpaper and cleaned with acetone. The adhesive used was superglue which was allowed to cure for 24 hours before testing began. Superglue is a common adhesive used to mount sensors in ambient environments.

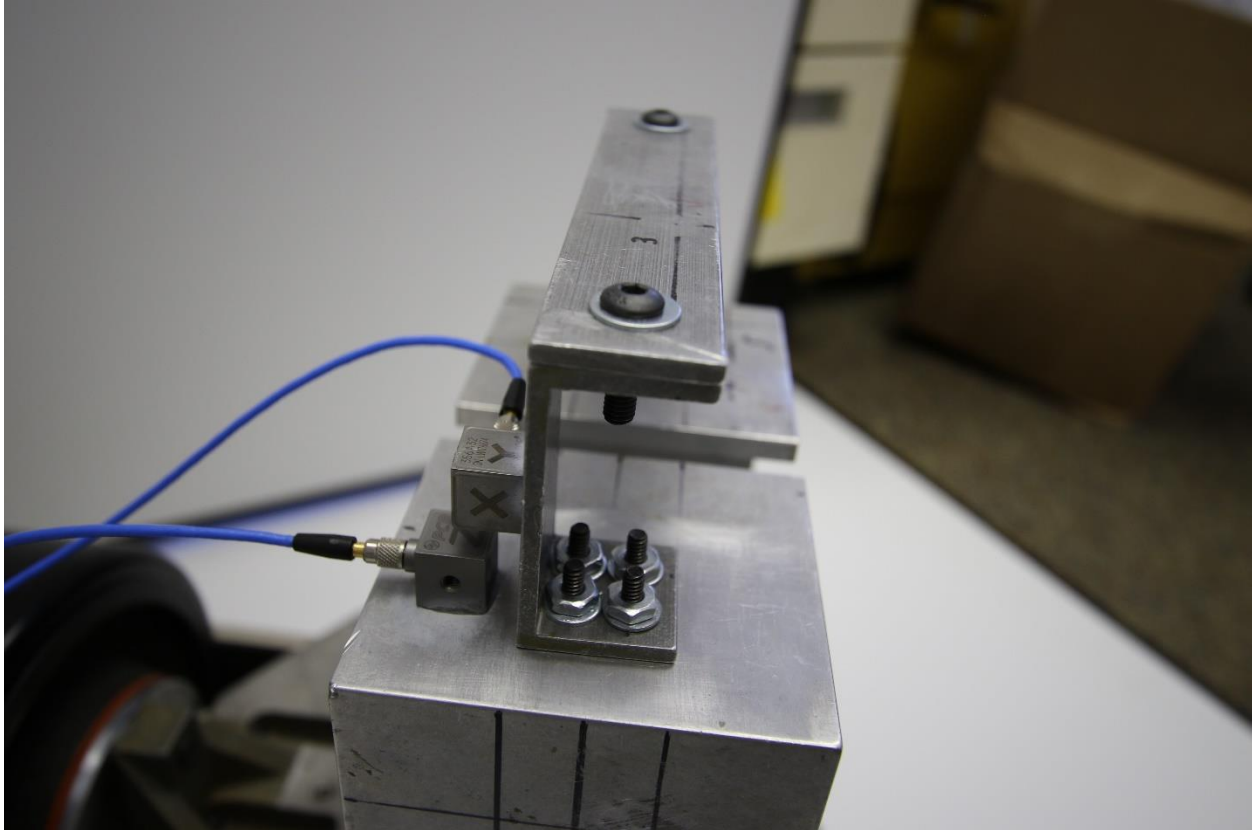


Figure 4 – Sensor locations and orientations

Data Acquisition Unit

The control and data acquisition software is Siemens Simcenter Testlab, and the hardware used was the Simcenter SCADAS Mobile. The data acquisition unit can accommodate up to 40 channels, but the experiment only used seven channels. This software and hardware is standard in LANL's shock and vibration test facilities.

Control Computer

A standard laptop PC running Windows operating software was set up next to the shaker to control the test and allow the test operator to monitor output data. No special computing capabilities were required to control the test and collect data and most current engineering PCs have more than enough storage and computing power to analyze the collected data.

Feature Extraction & Normalization

Measurements and Signal Conditioning

The sensors mounted to the BARC structure were triaxial piezoelectric accelerometers which record voltage and convert it to acceleration. In this study, only the data from channels in

the same axis as the excitation input are used for analysis. This criteria means that the -Z axis data from the top accelerometer and the +X data from the bottom accelerometer were used for analysis. The recorded time histories are transformed into the frequency domain, and then a transmissibility is calculated between the in-axis channels of the two accelerometers. This transmissibility data is referred to as a frequency response function (FRF). The accelerometer above the joint was treated as the response output and the accelerometer on the box assembly was treated as the input.

The data collected during testing were sampled at 25,600 Hz to avoid aliasing. Transmissibility was calculated averaging 15 one-second windows with zero overlap. Each window has a Hanning function applied to prevent leakage. This operation resulted in 240 transmissibility functions for the healthy system and 20 transmissibility functions for each of seven damage cases. The measured transmissibility data had a frequency range between 10 and 3000 Hz to capture higher frequency modes that are more sensitive to non linearities being introduced into a system, such as a joint opening and closing.

Data Normalization

This experiment was conducted in a controlled environment where minimal data normalization was required for the model to separate healthy from damaged data. After the shaker and control software completes its self-check, the controller steps up in levels from -9dB to -6dB and -3dB before finally inputting the nominal 0dB environmental specification. These build-up levels allow the software to develop and adjust the control strategy, but are not quality data that should be considered when calculating the transmissibility functions for training the model. The initial 20 seconds of each time history was trimmed to remove these lower level inputs and normalize the data.

Statistical Model Development

Model Selection

The statistical model chosen to analyze the FRF for indications of decreased preload in a bolted joint is a special type of neural network called an autoencoder. Autoencoders are designed to match the output to the input via unsupervised learning. The model takes frequency response functions as input data, reduces the input data dimension to just a few values capturing the most important features of the FRF, and then attempts to reconstruct the original FRF from the compressed latent space.

Autoencoders are used to capture salient information during training and are an excellent choice when noise is present in the measurements. Another benefit of autoencoders is the ability to learn non linear transformations using multiple layers and non linear activation functions. The ability to learn non linear transformations made the autoencoder more desirable than principal component analysis, which is constrained to learning linear transformations.

Autoencoders are only useful for evaluating data similar to what the model was trained with. In the case of this research, the model was trained on healthy data measured from the undamaged BARC system. Training the model only on healthy data meant the model would be

sensitive to any changes to the input data, such as a measurement taken from a damaged BARC system. The final iteration of the model trained over 1000 epochs and the individual neuron weights in the autoencoder were optimized to reconstruct healthy FRFs. Once the input data were modified as a result of damage in the system, the weights are no longer optimal and our model produces an output FRF with significant error compared to the input.

This experiment was performed in a laboratory space where there were minimal changes to the operational conditions of the BARC structure. For a given test article undergoing shock and vibration environmental testing combined with thermal conditioning, measurements would need to be taken of the healthy structure across a spectrum of these conditions. The autoencoder is well suited to this scenario where there are more than one operating conditions because of the model's ability to learn non linear transformations. For example, varying the temperature may not result in a linear increase of error between the original and reconstructed FRFs. Additionally, the interactions between changing temperature and different damage cases are unlikely to be linear, but the autoencoder can account for the interaction.

Model Anatomy

An autoencoder has two parts: an encoder and a decoder. The encoder takes the input data and reduces the dimension down to what is called the latent space. The latent space is a low level representation of the input data and may be only a few neurons wide. The decoder uses the latent space as the input and expands the data in an attempt to replicate the original input to the encoder. The decoder output is compared to the input and the loss is calculated. The loss is backpropogated through the model and the weights are updated to perform better on the next input data set. This process is repeated as defined by the user until the training time is excessive or a suitable level of error is achieved.

The latent space can appear abstract and may not contain any physical meaning but represents the important features of the input data with only a few neurons. The decoder is a mirror of the encoder such that the entire network looks like an hourglass on its side as seen in Figure 5.

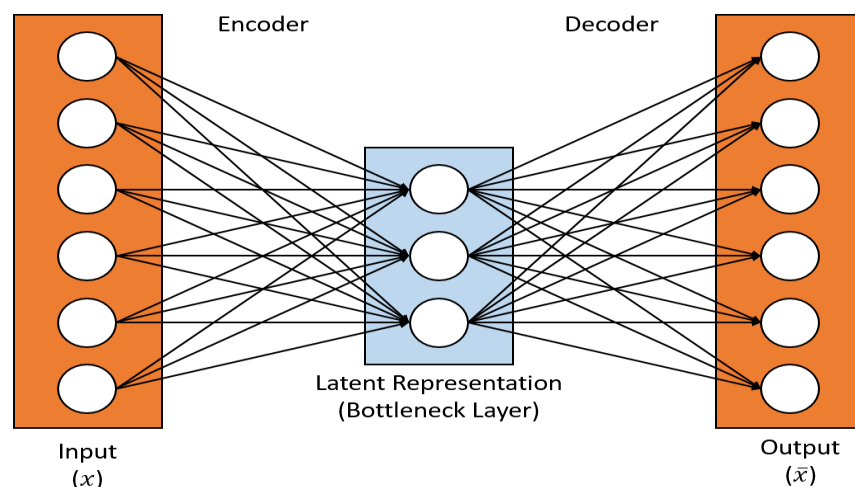


Figure 5 - Graphical representation of simplified autoencoder

The healthy data was split 50/50 into two sets, one for training and one for testing. Each data set contained 120 healthy system FRFs. The model was trained on the first data set and then tested on the second data set as well as all seven damage cases. Figure 6 below shows a diagram capturing the flow of data through the model. In the encoder, the input FRFs are multiplied by a weight layer and a bias term is added to the product before the logsig transfer function is applied. The data output from the logsig function is the latent space representation of the data, only 10 neurons. These 10 neurons contain the salient information learned during training that result in minimized reconstruction error as it is passed through the decoder. The decoder performs a similar process to the encoder, taking the latent space weights as input data, multiplying by another layer of weights, adding a bias term, and applying the decoding transfer function, also logsig.

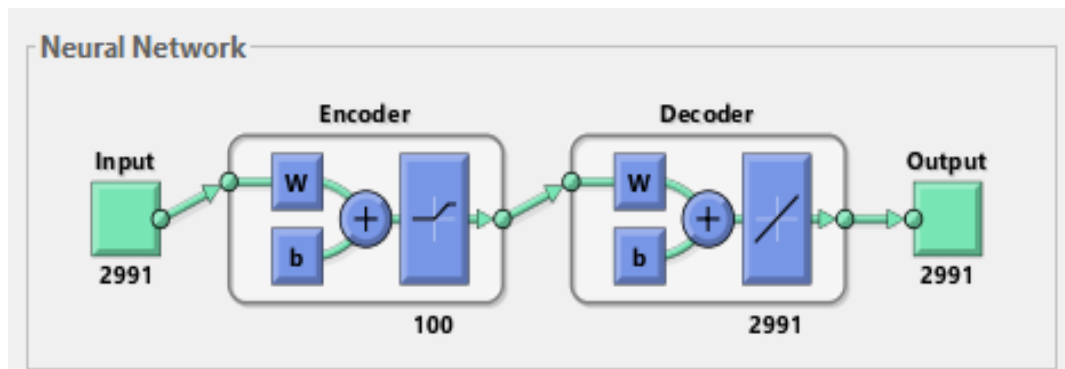


Figure 6 - Example of data operations during model training.

Selected Model Hyperparameters

All data processing was completed in Matlab 2021b. The only required toolbox was the Deep Learning toolbox which simplified building the autoencoder model and allowed for hyperparameter tuning. Table 2 details the hyperparameters used in the final model. The first hyperparameter, MaxEpochs, refers to the maximum number of times the model will train on the entire data set before completion. The transfer function used in both the encoder and decoder was a logistic sigmoid function, or logsig. A plot of the logsig function is displayed in Figure 7.

Table 2. Hyperparameter Settings for Autoencoder.

Hyperparameter	Value
MaxEpochs	1000
L2WeightRegularization	0.005
SparsityRegularization	4

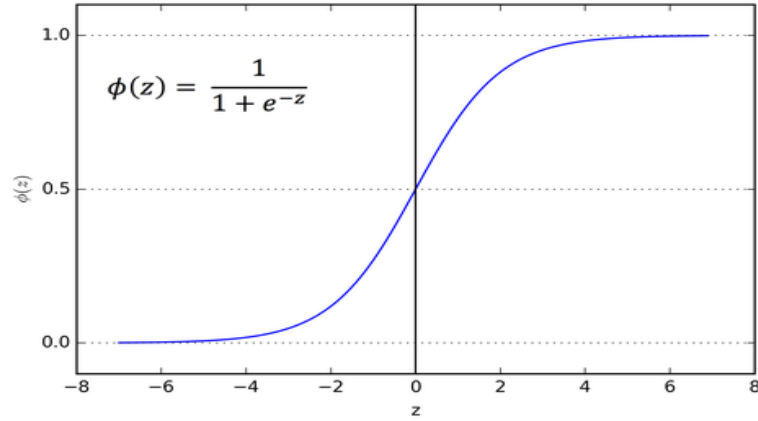


Figure 7 - Logistic sigmoid activation function used in encoder and decoder [3].

The L2WeightRegularization was increased slightly from the default value to help avoid overfitting the model to the training data. If a model begins to overfit the data, it will achieve 100% accuracy in training, but will adapt poorly to new data during testing. While overfitting may sound like a beneficial feature for the application of detecting loose bolts as any damaged data would have sizeable reconstruction error, it would likely also mislabel new healthy data as damaged. To avoid overfitting the model to the training data, the L2 weight regularization term is increased.

The SparsityRegularization hyperparameter was also increased from the default value. The sparsity regularization term refers to a constraint placed on the sparsity of the output from the hidden layer. Increasing the value encourages sparsity in the hidden layer output and produces better results classifying damaged data [4]. The model's cost function is a modified mean squared error function that includes the L2 regularization and sparsity regularization.

Results

Measured and Model Reconstructed Transmissibility

Figure 8 compares the measured FRF to the FRF reconstructed by the model for the healthy system. The blue data is the FRF calculated from the measured time data. The orange data is the FRF reconstructed by the model. Several of the peaks in the plot show that the reconstructed FRFs had a lower amplitude than the measured FRFs. This result is caused by the autoencoder's structure when the data is compressed down to the latent space, and is then reconstructed to the full length FRF from 10 to 3000 Hz. Some smaller details in the data are sacrificed to preserve the most important features in the training data, resulting in the slight differences between the measured and reconstructed FRFs of the healthy system.

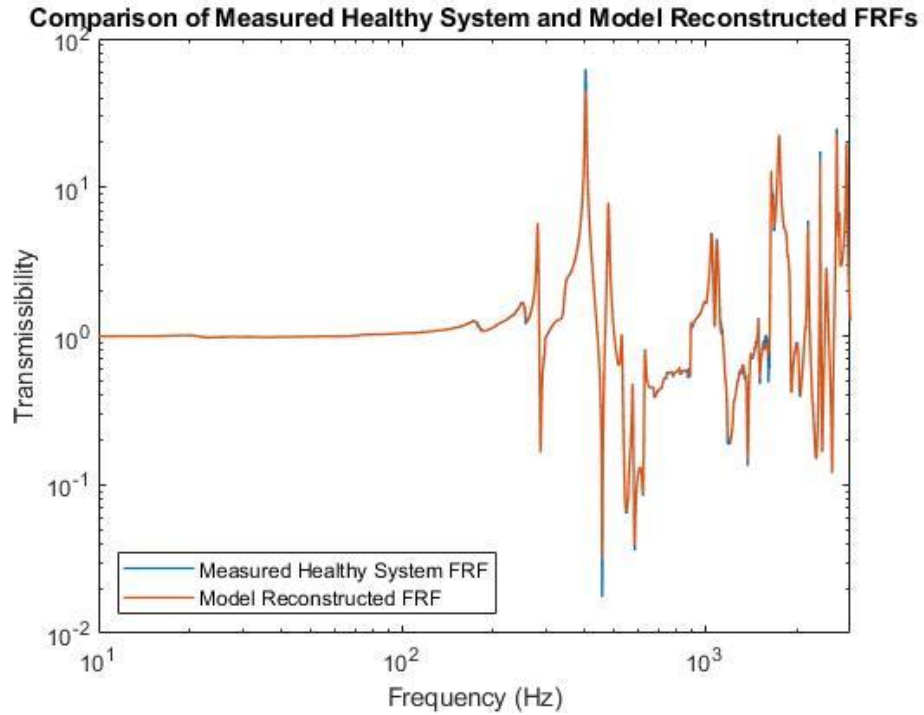


Figure 8 - Comparing measured and reconstructed FRFs from healthy system.

Figures 9a and 9b compare the measured FRF to the FRF reconstructed by the model for the damaged system. In both figures, the blue data is the FRF calculated from the measured time data while the system was damaged, in this example, Case 5. The orange data is the FRF reconstructed after the data has been compressed and reconstructed by the autoencoder. Figure 9b is a zoomed view of the same plot in Figure 9a, just over the frequency range 250-3000 Hz which makes the differences more visible. In this plot, error can be seen in peak shifts, such as at 300 Hz. This error is a result from trying to replicate a damaged system FRF with a model that was optimized for healthy system FRFs. The difference is then used to calculate MSE and compared against the damage threshold to assess the state of the system's health.

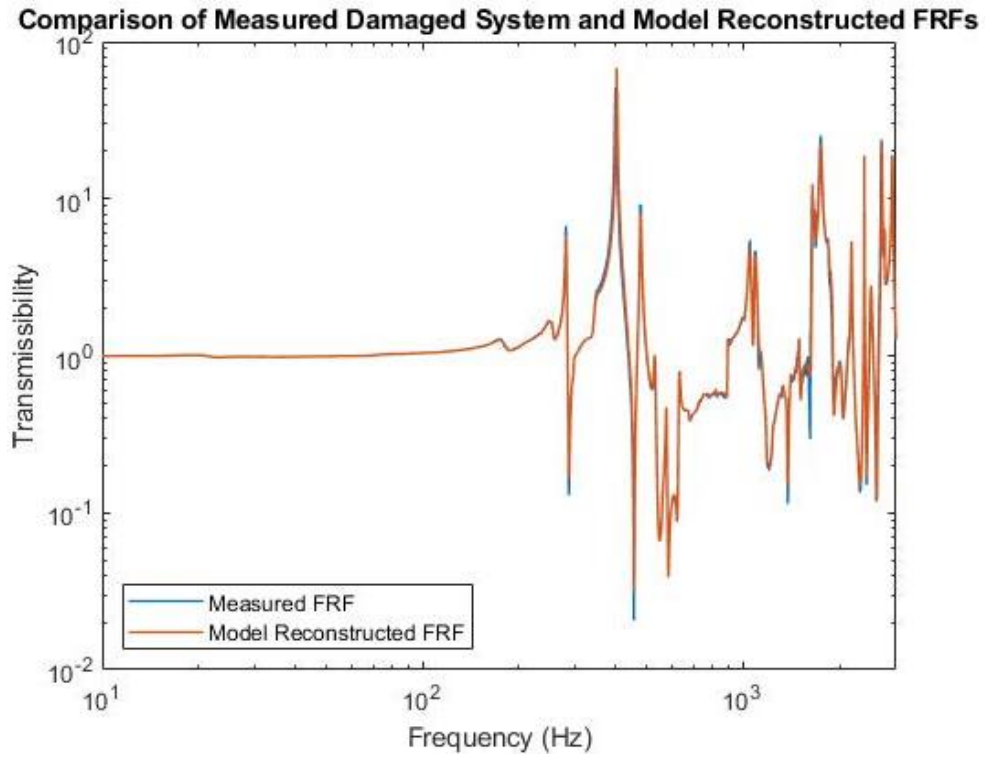


Figure 9a - Comparing Measured and Reconstructed FRFs from Damaged System.

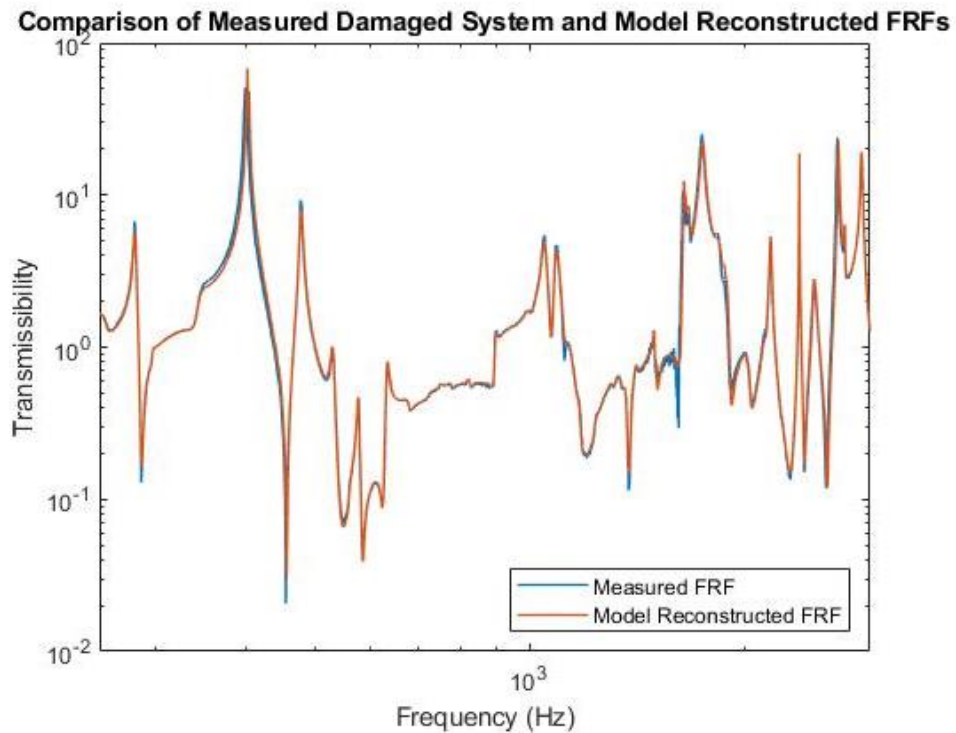


Figure 9b - Comparing Measured and Reconstructed FRFs from Damaged System (zoomed in from 250-3000Hz).

Damage Detection and Classification

Figure 10 plots a histogram of four unique damage cases compared to healthy system data. The horizontal axis is MSE, mean squared error, which is calculated between the input FRFs and the reconstructed output FRFs from the testing data set. Each bin width is 0.5. The vertical axis is occurrence frequency for each MSE bin. The vertical axis has been normalized such that the sum of probabilities for each individual test case is equal to one. The blue data represents the healthy system data which is located farthest left on the plot, indicating low MSE. Low MSE is expected as the model was trained to minimize the error between the measured input FRFs and reconstructed FRFs from healthy system data.

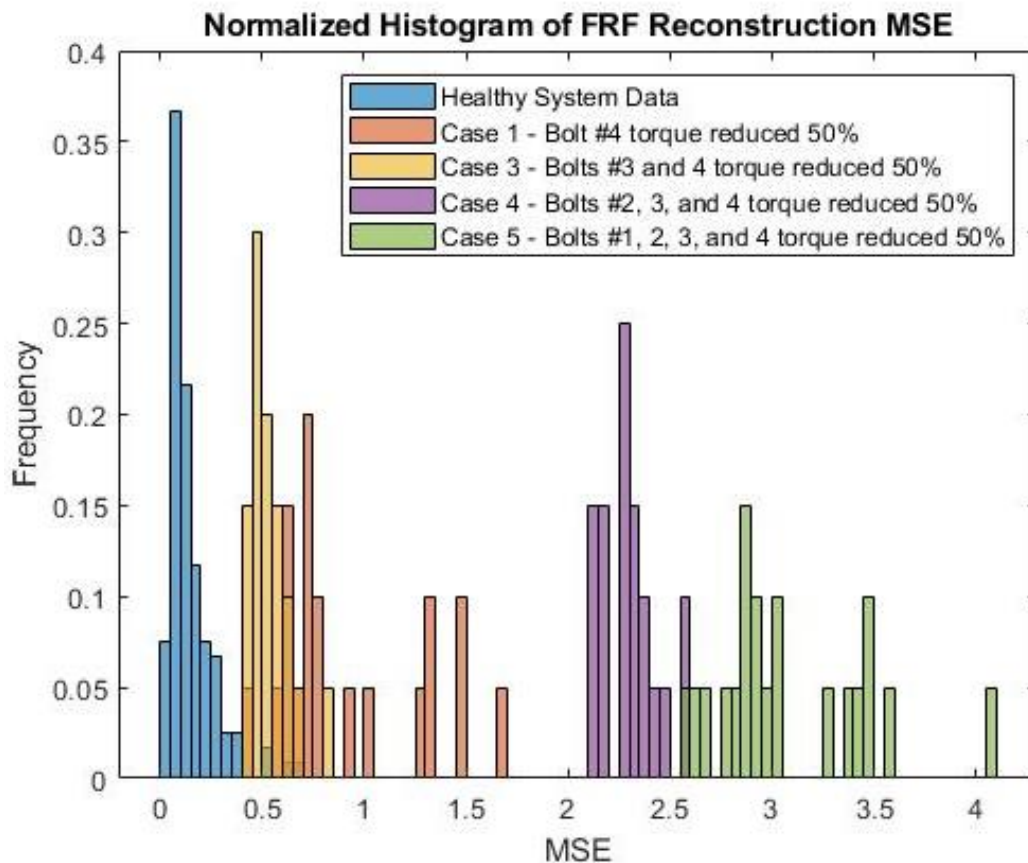


Figure 10 – Histogram comparing MSE from healthy and damage cases as more bolts lose preload.

Cases 1, 3, 4, and 5 show incrementally more damage present in the bolted joint. Each case has an additional bolt torqued to 50% of the specified value. Cases 4 and 5 have high MSE, on the order of four to seven times the MSE present in the healthy system data. The two cases are also distinguishable from one another, Case 5 exhibiting higher MSE than Case 4. Case 5's higher MSE is to be expected as there is more damage present in the system with an additional loose bolt.

However, Cases 1 and 3 display opposite behavior where the less damaged joint, Case 1, has higher MSE than the more damaged joint, Case 3. This phenomenon is counter intuitive and may be an artifact of the autoencoder's latent space reconstruction. The model's weights

were optimized to encode and decode the healthy system FRFs, but it's possible that these weights also reconstructed the Case 3 FRFs with less error than Case 1, despite the greater damage in the bolted joint.

Figure 11 compares the results of a loose bolt versus the same bolt completely removed. The other three bolts were torqued to the specification in both Cases 1 and 2. The trend is obvious in the plot as the case with a missing bolt exhibited a much higher MSE compared to the case with the same bolt torqued to 50% of the specification. The clear separation between these two classes are promising as the system can be identified for damage and the type of damage can be classified. In test engineering, a loose bolt may not warrant pausing a test, but identifying a missing bolt would certainly stop work.

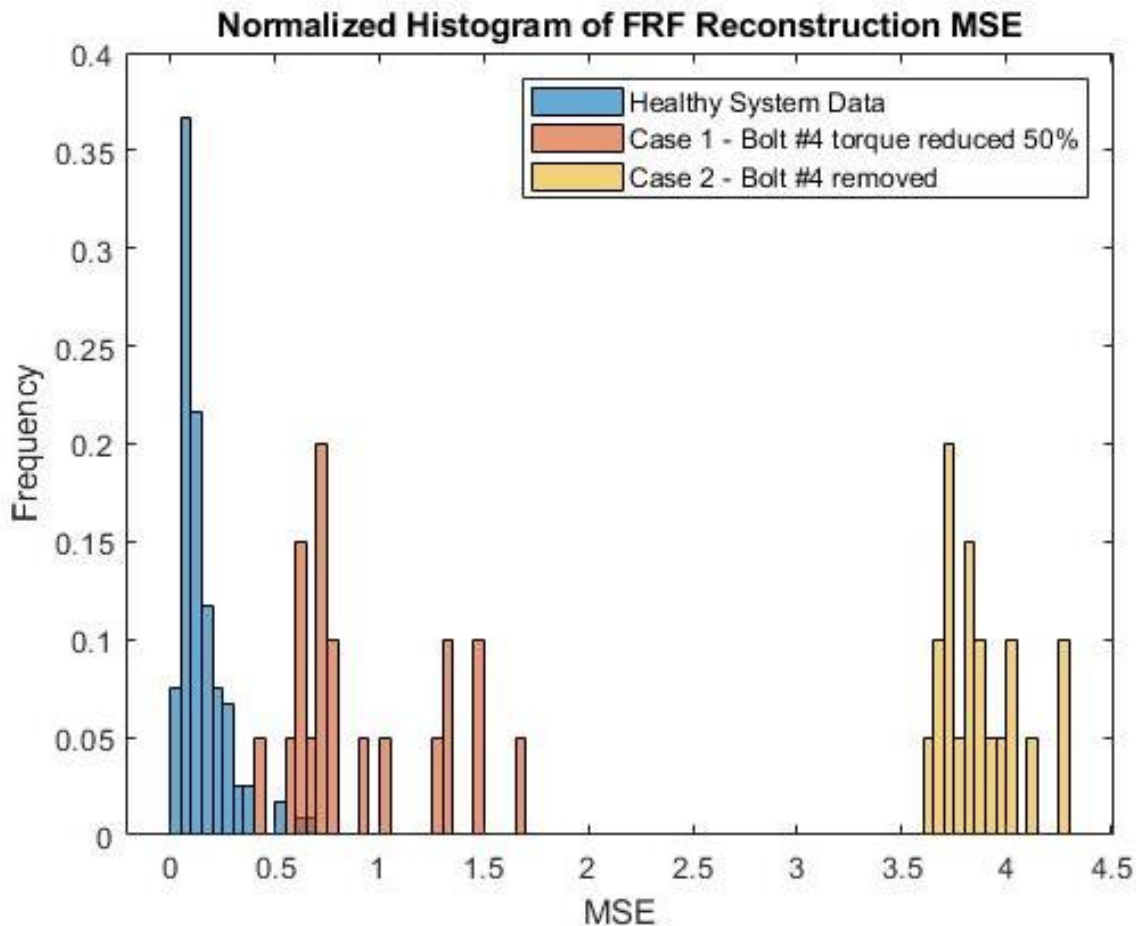


Figure 11 – Histogram comparing MSE from damage cases with a loose bolt versus a missing bolt.

Figure 12 explores reciprocity in the experiment. In Case 3, bolts 1 and 2 were torqued to specification while bolts 3 and 4 were torqued to 50% of the specification. In Case 4, the scenario was flipped such that bolts 3 and 4 were torqued to specification while bolts 1 and 2 were torqued to 50% of the specification. Case 3 exhibits less error than Case 6.

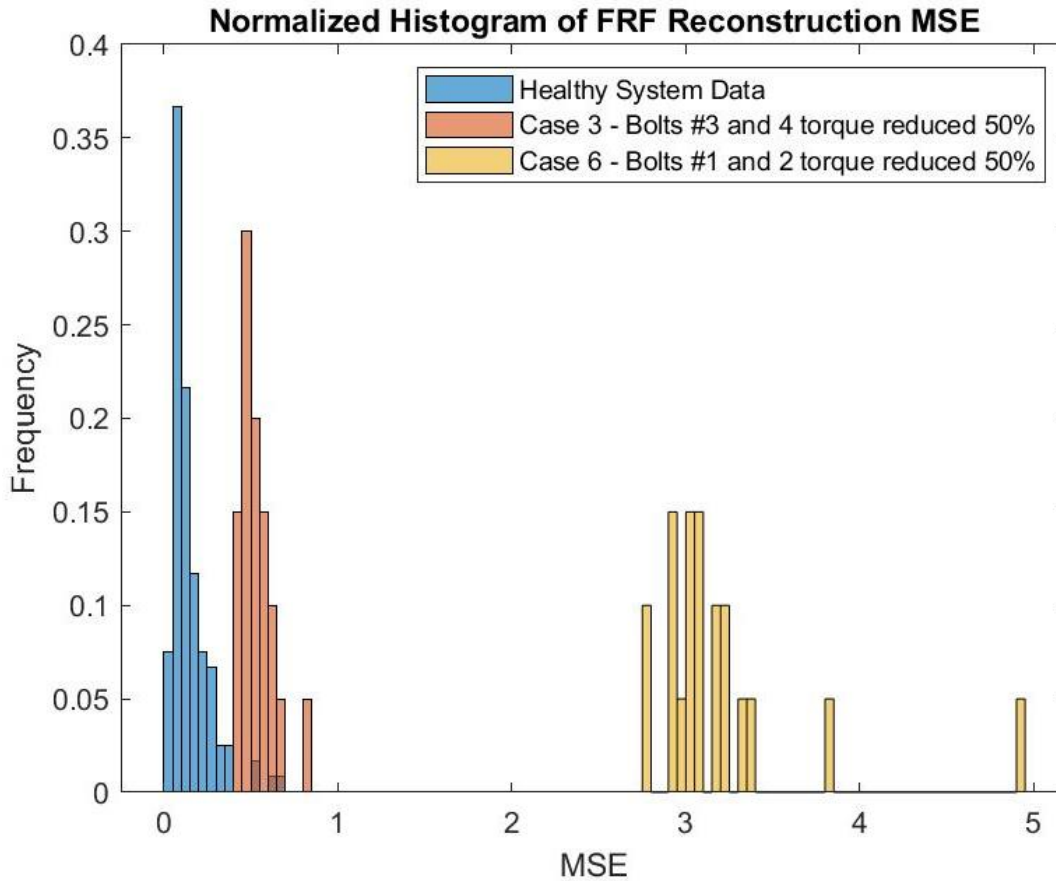


Figure 12 – Histogram comparing MSE from damage cases with mirrored bolts missing preload.

Previously, it was hypothesized that Case 3 had low error because the weights of the model may have also represented this damage case well. However, Case 6 explored the same scenario with two different bolts loosened and exhibited higher error. As seen in Figure 4, the two sensors collecting measurements were spatially located closer to bolts 1 and 2 which could be a factor contributing to the increased error. However, the relationship is unclear and more data would need to be collected and analyzed before a conclusion could be made.

The below figure, Figure 13, is a histogram plot with mean squared error on the horizontal axis and normalized frequency on the vertical axis. The bin width is 0.05. The black dashed line represents the average MSE of the healthy system's reconstructed FRFs and the red dashed line is three standard deviations from the mean MSE. The majority of the data falls within three standard deviations, but several reconstructed FRFs resulted in error higher than the three standard deviations. The damage threshold was defined to be three standard deviations under the assumption that the distribution of the healthy data histogram could be approximated at Gaussian. Three standard deviations would capture 99.7% of data in a Gaussian distribution. The threshold could be set higher, such as five standard deviations, but would increase the number of false negatives the model detects. In Figure 10, there is visible overlap between healthy and damaged cases, so while a higher damage threshold would minimize false positives, it would also increase false negatives.

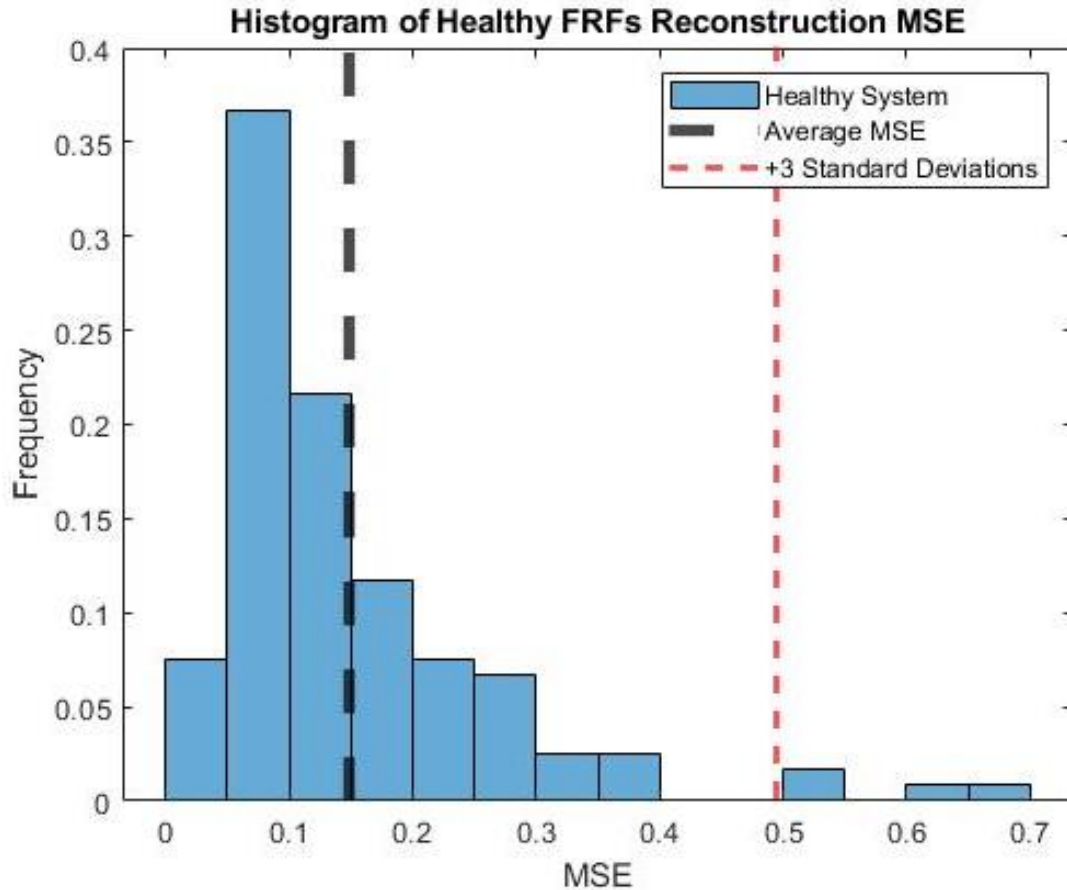


Figure 13 – Histogram of healthy data with established damage threshold.

As seen in previous plots, several damage cases have instances where the MSE overlaps the healthy system outliers and fall within the three standard deviation mark. This overlap creates a challenge defining a distinct threshold for defining damage in the bolted joint system. The overlap can be seen in Figure 14. The healthy system data color was changed to green for easy identification among the seven damage cases and the damage threshold line is plotted. As in Figure 13, the false positives can be identified where the green bars exceed the damage threshold. Similarly, false negatives are visible for three damage cases, Cases 1, 3, and 7, where blue, yellow, and red bars fall underneath the damage threshold. The other damage cases can be easily classified as damage because all of the tested instances exceed the damage threshold. Table 3 prints the resulting confusion matrix with the damage threshold set to three standard deviations from the healthy system's average MSE.

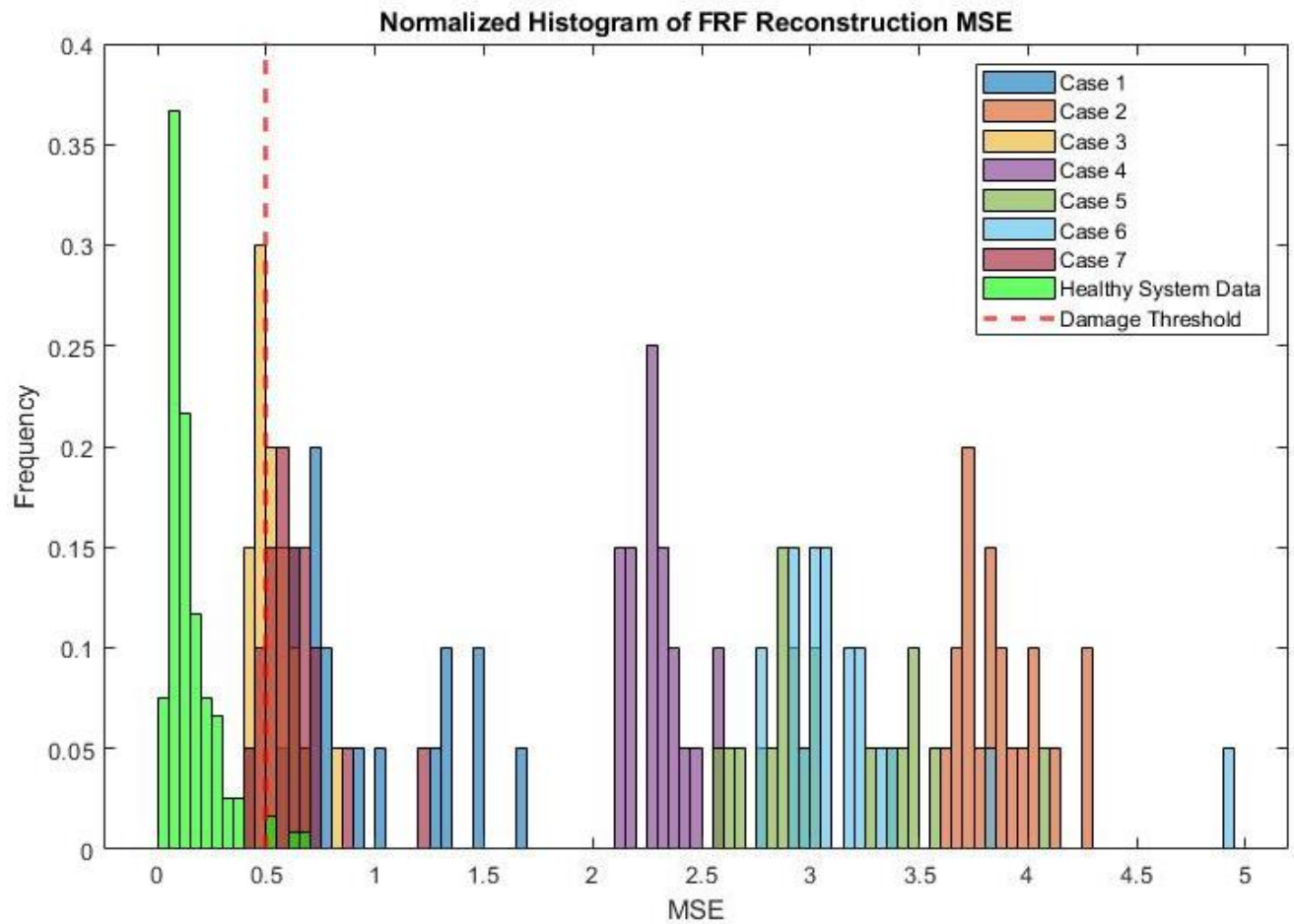


Figure 14 – Histogram of reconstruction error for all test cases with damage threshold visible.

Table 3. Confusion matrix for classification results.

n = 260	Predicted: NO	Predicted: YES	
Actual: NO	TN = 116	FP = 4	120
Actual: YES	FN = 13	TP = 127	140
	129	131	

The resulting accuracy of this damage threshold is 93.5% with a precision of 96.9%. The miss rate is 6.5%. This damage threshold is a basic approach and could be improved. With more data to work with, outliers might be identified and can be eliminated during the data normalization process, reducing the false positives. A more sophisticated statistical approach would be required to separate the healthy system from the damage cases that have overlapping instances, like Cases 1, 3, and 7. An example may be closer analysis of how the distributions are shaped and classifying instances based on other statistical features.

Implementation Issues and Recommended Improvements

If a test engineer were to implement this system on a real environmental test at a LANL facility, they would face several issues. As mentioned above, real environmental tests experience high acceleration, velocity, displacement, and possibly thermal conditioning. These environments create an expansive space of conditions that will cause a healthy system to exhibit changing FRF signals. However, the changes in these healthy system signals do not indicate damage, and the model must learn to fit the entire operating condition space. Additionally, there is the possibility that re-torquing bolts in between configurations could lead to the model falsely identifying damage.

Further data normalization would become crucial for the implementation of this sensor system and statistical model into a live environmental test. As the system is excited by different inputs and subjected to different temperatures, healthy signals may look different when compared to one another but do not indicate damage. The initial scope seems intimidating, but because these environmental tests are conducted in controlled settings and the fact that nearly-identical tests are often conducted more than once, it would not be impossible to normalize data for the entire operational space of the test.

This experiment had rudimentary cross validation and would benefit from an in-depth analysis. The hyperparameters were individually varied, usually increased or decreased by an order of magnitude, to determine if the change improved the model's performance. A more sophisticated method may randomly sample different combinations of all the hyperparameters between predetermined bounds and use that to achieve the most accurate model.

Traditional structural health monitoring systems require procurement of ample equipment including sensors, data acquisition hardware, and a computer. However, by leveraging LANL's testing capabilities, all of the equipment is already available. There are various sensors, spare channels on data acquisition units, and capable computers at the facility that could be used to implement a SHM system on various test articles with bolted joints. Even if this experiment was scaled up to a full system, the number of channels would not likely exceed the existing capabilities within testing facilities in place.

Conclusion

This experiment showed that autoencoders are a good statistical model for monitoring the health of a bolted joint. The model was able to detect decreased preload in several different combinations of bolts by comparing the model's reconstructed FRF to the measured FRF. When the model is trained on healthy signals, and a healthy signal is tested, the reconstructed signal has very little error. When a damaged signal is introduced, the reconstructed signal has significant error and can be classified as damage.

The threshold for damage was defined to be three positive standard deviations from the mean MSE of the healthy system data. This simple threshold was able to detect damage with an accuracy of 93.5% and a miss rate of 6.5%. There are several ways that this threshold could

be improved to achieve better accuracy and miss rate, but would rely on more sophisticated statistical methods.

Additional data and a wider variety of damage cases would help train the model to be more robust and allow it to be deployed in more scenarios. The results from this experiment indicate that studying a single joint with accelerometers located in close proximity can be successful. However, studying joints with more bolts with sensors located sparsely throughout the system could create a more challenging detection problem. The autoencoder model appears to be well suited for these problems, especially if trained on large datasets that encompass various healthy states of the system. The hyperparameters can be tuned to be increasingly sensitive to changes in the input FRF data and would make the damage easier to identify in a sparsely instrumented system.

References

[1] Model 2075E 75 lbf Dual Purpose Shaker. The Modal Shop, 2021.
<https://www.modalshop.com/excitation.asp?ID=251>

[2] Model K2075E-HT and K2110E-HT Vibration Shaker Horizontal Table. The Modal Shop, 2021. <https://www.modalshop.com/excitation/Horizontal-Table?ID=580>

[3] Sharma, Sagar. "Activation Functions in Neural Networks". (2017)
<https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6>

[4] Mathworks. (2021). Deep Learning Toolbox: User's Guide (r2021b).
https://www.mathworks.com/help/deeplearning/index.html?s_tid=CRUX_lftnav