Objective Performance Assessment Using Artificial Neural Networks

A thesis submitted by

Jordan C. Weinstein

in partial fulfillment of the requirements for the degree of

Master of Science

in

Civil and Environmental Engineering

Tufts University

February 2018

Co-Advisors: Masoud Sanayei, Ph.D. & Brian Brenner, P.E.

Abstract

Bridge behavior is used as an objective, data-driven indicator of the performance of bridges. A framework with which bridge behavior can be identified and learned is presented, and a method of long-term damage identification using the expected bridge behavior is introduced. At the Powder Mill Bridge (PMB) in Barre, Massachusetts, strains at each strain gage location are recorded during operational traffic events. Bridge behavior is defined as each sensor location's range of expected peak strain during a traffic event based on all other measured strains at the time at which it experiences its peak strain. Artificial neural networks (ANNs) are trained with operational data in a bootstrapping scheme to generate a probabilistic model of bridge behavior. When tested against new data, the ANN-learned model of bridge behavior is validated for a variety of traffic events with unknown loading conditions.

Structural damage is one way that bridge behavior, an indicator of performance, of a bridge can change. Damage scenarios are simulated in a finite element model (FEM) which is calibrated to PMB truck load test data. The effects of damage are extracted from FEM truck runs and applied to operational data to assess the capability of the proposed damage identification method through a series of trials. It is effective at detecting damage, with no Type I and no Type II errors when using a Wilcoxon rank-sum test of an appropriate significance level. Damage is effectively localized for two out of three damage scenarios.

Certificate of Fitness

Tufts
UNIVERSITY

ARTS, SCIENCES, AND ENGINEERING

CERTIFICATE OF FITNESS

This certifies that the undersigned, appointed to determine the fitness of

Jordan Weinstein

for the degree of Master of Science in Civil and Environmental Engineering

have examined the candidate's thesis/dissertation (or papers) on the subject:

Bridge Performance Assessment Using Artificial Neural Networks

and have found it satisfactory.

August 25, 2017

Date of Defense

This certifies further that the candidate this day has successfully passed the customary examination.
We recommend, therefore, that the degree be awarded under the usual conditions.

SIGNATURES	NAMES
Masoud Sanayei	Masoud Sanayei (Co-Advisor)
Chairperson, Examining Committee	*Chairperson, Examining Committee*
Brian Brenner	Brian Brenner (Co-Advisor)
Erin Bell	Erin Bell

Student Services
Dowling Hall
Medford, Massachusetts 02155
617.627.2000

Acknowledgements

First, I would like to thank my advisors, Professor Masoud Sanayei and Professor Brian Brenner. Their guidance and knowledge throughout this process has been invaluable, and I look forward to maintaining a close relationship with them in the future as I begin working. I would also like to thank one of my committee members, Professor Erin Bell. Her accommodation of my summer defense schedule, which can be difficult for her since she is not at Tufts, has made things much more manageable for me.

I would like to thank my officemates and friends, Connor Brown, Nate Davis, Sofia Puerto, and Wenjian Lin, for making all my long stretches in the office a more fun experience. I would also like to thank my brothers, Jonathan and David. They have been there for me my entire life and were a huge source of love and support throughout my two years in graduate school.

I most importantly want to thank my parents, Karen and Joe. I would not have completed this thesis without everything they have done for me. Beyond that I owe everything I have, and everything I am, to them. They have provided me with a better life than I could have ever asked for, and they raised me to be the self-motivated person I am today. They encouraged me to do whatever brings me happiness in life. Their unconditional love and support constantly fills me with a sense of comfort, happiness, and home.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Scope of Research

The Powder Mill Bridge (PMB) is used as a case study to assess a proposed method of structural health monitoring (SHM) using artificial neural networks (ANNs) as non-mechanistic modeling tools. The bridge is a three-span steel composite concrete stringer highway bridge crossing the Ware River in Barre, Massachusetts. The bridge has been fitted with an instrumentation system that responds to vehicular loading.

In this study ANNs provide a mathematical system model linking various components of the PMB structural response (strains in steel girders at several locations) to each other. A framework with which bridge behavior can be identified and learned is developed, and a method of long-term damage identification using the expected bridge behavior, which can be passively performed on a weekly basis, is introduced. ANNs are trained with operational data in a bootstrapping scheme to generate a probabilistic model of bridge behavior. When tested against new data, the ANN-learned model of bridge behavior is validated for a variety of traffic events with unknown loading conditions.

Damage scenarios are simulated in a finite element model (FEM) which is calibrated to the PMB. The effects of damage are extracted from FEM truck runs and applied to operational data. Several trials are generated to assess the damage detection and localization capabilities of the developed method. The approach is shown to be exceptionally effective at detecting damage, with no Type I errors (detecting damage when it is not present) and no Type II errors (failing to detect

damage when it is present) when tested with a Wilcoxon rank-sum test of a sufficient significance level. Damage is effectively localized for two out of three damage scenarios generated.

1.2 Summary of Chapters

This chapter introduces the thesis and its structure. Chapter 2 provides a broad summary of the overall research approach and results and is intended to be submitted as a stand-alone paper to the ASCE *Journal of Bridge Engineering*. Chapter 3 describes the removal of truck events not used in this research; Chapter 4 provides an overview and justification of the Artificial Neural Network classification used in the study; Chapter 5 demonstrates the research approach through a simplified example; Chapter 6 provides an analysis of the effects of temperature on the study; Chapter 7 gives additional details and justification of damage effects to supplement those in Chapter 2; Chapter 8 addresses the uncertainties of damage simulation with a sensitivity analysis; and Chapter 9 concludes the work.

Chapter 2. Bridge Damage Identification Using Artificial Neural Networks

Proposed Journal for Submittal: ASCE *Journal of Bridge Engineering*

2.1 Abstract

An objective, data-driven approach to evaluate the performance of a bridge for developing a structural health monitoring system is introduced as *bridge behavior*. A method of identifying structural damage through evaluation of response data from an instrumented bridge is proposed. Strains during operational traffic events at the Powder Mill Bridge (PMB) in Barre, Massachusetts are recorded at multiple locations along the bridge. *Bridge behavior* is defined as each and every sensor location's range of expected peak strain during a traffic event based on all other measured strains at the time at which it experiences its peak strain. Artificial neural networks (ANNs) are trained with operational bridge response data in a bootstrapping scheme to generate a probabilistic model of bridge behavior. When tested against new data, the ANN-learned model of bridge predicted behavior is proven effective and applicable to varying traffic events with unknown loading conditions. A method for long-term performance assessment using the expected bridge behavior is proposed. Structural damage is one way that bridge behavior, an indicator of performance, of a bridge can change. The simulated effects of structural damage are extracted from calibrated FEM truck runs and applied to operational strain data to assess the damage identification method. When assessed, the damage identification method is effective at detecting the presence of damage, with no Type I and no Type II errors when using a Wilcoxon rank-sum test of an appropriate

significance level. Damage is effectively localized for two out of three types of simulated damage.

Key Words: *Structural Health Monitoring, Artificial Neural Networks, Bridge Behavior, Damage Identification, Operational Strain Measurements, Response Only, Hypothesis Test*

Highlights:

- *A Framework is defined through which bridge health can be assessed.*
- *Artificial Neural Networks are trained using measured operational strains to represent bridge behavior.*
- *A trained model is used to imitate baseline bridge behavior for new truck loadings.*
- *Simulated damage is used to examine capability of trained model.*

2.2 Introduction

Bridge maintenance is supported by a program of inspection. In the United States, bridge inspections are required to be performed and documented at least once every two years. The current bridge inspection approach relies mostly on visual inspections with limited field measurement. This approach is labor intensive and somewhat subjective (Moore et al. 2001).

Developing technologies of instrumentation, data collection, remote sensing, and analysis provide the ability to evaluate a bridge's structural response in real time. These technologies can form the basis of a structural health monitoring system, which would provide a more objective evaluation of the overall performance and condition of a bridge. This would assist bridge owners in efficiently allocating finite resources for maintenance.

Since bridge structures are highly indeterminate, relative changes in local stiffnesses can alter the way load is carried. Shenton III and Hu (2006) used the distribution of dead load strains in an analytical structural model to develop a baseline, and then used the baseline to detect simulated damage. Chakraborty and DeWolf (2006) and Cardini and DeWolf (2009) used strain data to determine baseline neutral axis locations and live load distribution among bridge steel girders to be used for long-term monitoring. Reiff et al. (2016) used strain data from traffic events to statistically establish a bridge signature and detect potential damage using girder distribution factors (GDFs).

Artificial Neural Network (ANN) methods provide a non-mechanistic approach to evaluating baseline characteristics of bridges to be used for long-term monitoring. ANNs can be developed from gathered information, determine relationships from complex systems with noisy data, and generalize such relationships to be applicable to varying circumstances (Reed 2009). Lam et al. (2006) used an ANN to identify different types of damage patterns caused by known types of damage. Zhang et al. (2012) performed damage detection of deck delamination of bridges using ANNs as a key tool. Kromanis and Kripakaran (2014) used ANNs to learn the input/output relationship between induced temperature loads and strains of a laboratory truss and the input/output relationship between temperature readings and tilt measurements of a pedestrian bridge. Mehrjoo et al. (2008) used ANNs fed with the natural frequencies of bridge models and the corresponding mode shapes to detect damage at truss nodes.

The proposed approach in this study is to determine baseline performance characteristics of a structure by assembling response data from instrumentation. As it pertains to this study, the baseline performance characteristic of a structure is sensor measurements that capture the distribution of load being carried throughout the bridge. ANNs are used to determine the relationship between strains measured at several sensor locations on a bridge during traffic events. It is an output-only approach: it utilizes measured strains and does not require any information about the nature of the traffic. A baseline of performance is determined using the learned relationship between the many measured strains along the bridge during traffic events. The use of ANNs is effective at establishing said baseline.

Additionally, a separate approach of performance assessment is presented utilizing aspects of the established ANN models of the bridge behavior, using damage identification as one example of its application. This approach demonstrates how bridge performance would be evaluated. A calibrated FEM is used to simulate potential damage since the studied bridge is not damaged. The simulated effects of damage are extracted from the FEM and applied to measured operational data to assess the capability of the proposed damage identification method. In structural health monitoring, there are four commonly used aspects of damage identification: detection, localization, assessment, and consequence (Rytter 1993). The proposed method addresses the first two: it is successful at detecting damage, and damage is localized for most of the simulated damage scenarios. Suggestions are made for future work to improve upon these newly developed approaches.

2.3 Powder Mill Bridge, Instrumentation, and Measured Data

This study is performed with operational traffic data from the Powder Mill Bridge (PMB), in Barre, Massachusetts (bridge B-02-012). The PMB, which opened in 2009, is owned by the Town of Barre and carries traffic on Vernon Avenue over the Ware River in Barre, Massachusetts (Figure 1). It is a continuous, two-lane, three-span bridge with six steel girders that act compositely with a concrete deck. The northern span of the structure widens into an intersection adjacent to the bridge. The bridge is in a rural area, near a waste management station. This location is optimal for research with operational traffic data, because many of the live loads experienced by the bridge are isolated, heavy trucks. The crossing of heavy vehicles over the bridge are herein referred to as *truck events*.

Fig. 1. Powder Mill Bridge and DAQ System

The bridge has many permanent sensors and data acquisition (DAQ) boxes that have been continuously recording and storing data since 2009. Figure 2 shows an instrumentation plan adapted from Sanayei et al. (2012) which displays the sensor locations, girder numbering, and sensor station numbering. A total of 50 strain gages recording at 50 Hz are potentially available for this study: one on each side

of the top faces of the bottom flanges of all four interior girders, and one on the interior side of the top faces of the bottom flanges of the exterior girders – at stations 2, 4, 6, 8, and 10. Strains are averaged between adjacent gages for locations at which both gages function. If no gages function at a location, data from the location are not used. If only one of two gages functions at a location, then data at that location are used from only the working gage. Figure 2 designates the 27 usable sensor locations in this study.

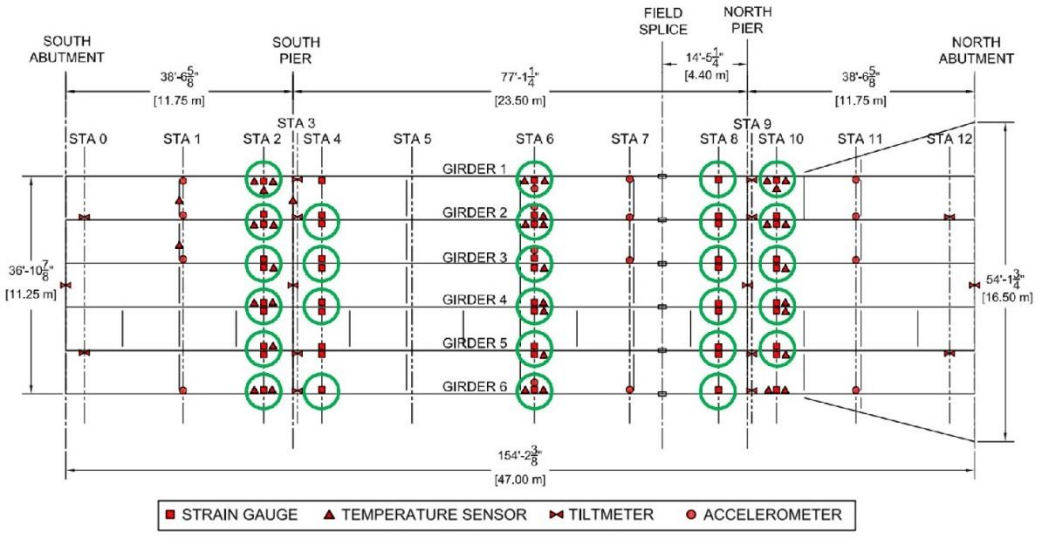

Fig. 2. Instrumentation Plan and Usable Sensor Locations (Sanayei et al. 2012)

A trigger program is written and implemented that automatically records traffic-induced strains when vehicles cross the bridge. The change in strain is used during a truck event, rather than the total strain. Since each truck event only lasts a few seconds and measurements are zeroed out for each truck event measurement, issues such as foil strain gage drift and temperature effects on gage wire resistance are negligible.

Truck events are used only when there is a single vehicle on the bridge at a time, because in this study the distribution of strains over the entire bridge caused by only a single vehicle load are assessed. The single vehicle must also be in either the northbound or southbound lane, as opposed to straddling the centerline of the road, because of way structural damage is simulated later in this work. Additionally, only truck events with heavy vehicles are used because the strain gage signal-to-noise ratios are higher. A moving average filter with a window of $7/50^{th}$ of a second, or 7 samples per second, is applied to strain signals. A program is written and implemented to automatically filter out truck events triggered by (a) a single light vehicle inducing maximum overall strain of less than 40 με, (b) two or more vehicles simultaneously on the bridge, or (c) a single vehicle straddling the centerline of the road.

2.4 Bridge Behavior Determination

2.4.1 Definition of Bridge Behavior

The framework through which the performance of the bridge is expressed in this study is called *bridge behavior*, because it considers the interrelationship between the structural response at different locations. Bridge behavior is defined as each and every sensor location's range of expected peak strain during a traffic event based on all other measured strains at the time at which it experiences its peak strain. For instance, the bridge behavior as observed at Girder 6 Station 2 and as observed at Girder 6 Station 4 are shown in Figures 3(a) and 3(b), respectively. The expected peak strain range ($\varepsilon_{\max}$) at Girder 6 Station 2 during a truck event can be

predicted based on the strains at all 26 other sensor locations ($\{\varepsilon\}$) at the instance in time at which Girder 6 Station 2 experiences ε_{max}. Likewise, the expected peak strain range at Girder 6 Station 4 during a truck event can be determined based on the strains at all 26 other sensor locations at the instance in time at which Girder 6 Station 4 experiences its own ε_{max}. This process as visualized in Figure 3 is applied to all 27 sensor locations on the PMB. If the peak strain values at a sensor location excessively deviate from their expected values, then it is postulated that the bridge behavior has changed. For simplicity, *bridge behavior* without reference to a specific sensor location herein refers to all 27 measured input/peak strain output ($\{\varepsilon\}/\varepsilon_{max}$) configurations. *Bridge behavior at Girder 6 Station 2*, for example, herein refers to only the measured input/peak strain output relationship in Figure 3(a).

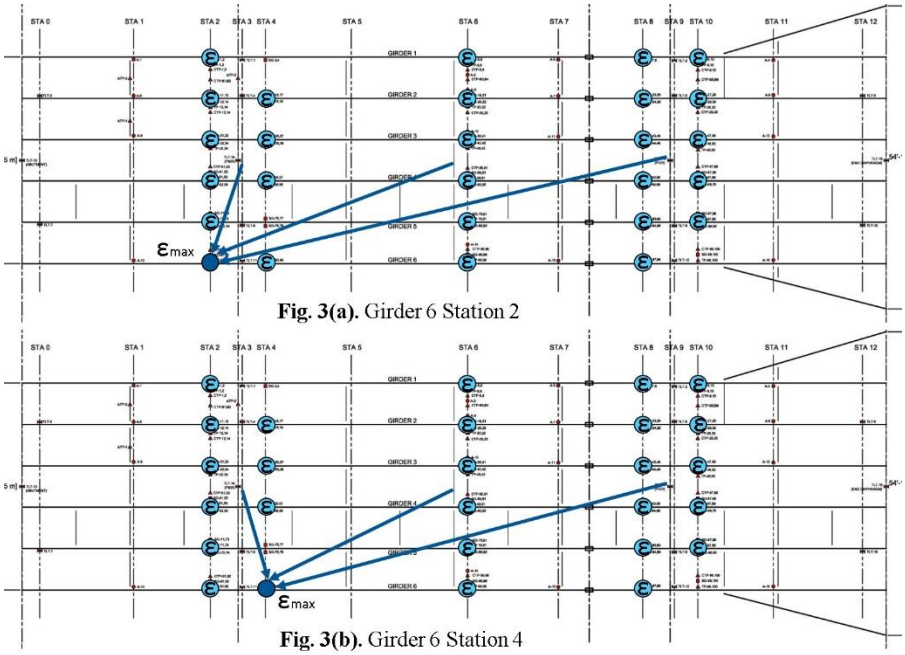

Fig. 3. Bridge Behavior as Observed at Girder 6 Stations 2 and 4

Moser and Moaveni (2011), Moaveni and Behmanesh (2012), and Alampalli (2000) determined that the modal properties of bridges vary with changing environmental operational conditions, such as temperature. Even though temperature can influence the physical bridge behavior of the PMB, thermistor data at the PMB was not used in this study because it did not have a significant effect on the developed bridge behavior model (as described later). When the model includes thermistor data as additional measured inputs, it performs similarly as it does when it omits thermistor data as additional inputs. Any actual temperature effects (or other environmental effects) that exist manifest themselves in this study by (a) creating a spread in data used to develop the baseline bridge behavior, (b) inducing more prediction error in the performance of the bridge behavior model, and (c) creating a spread in how accurately the damage effects simulated in the FEM (described later) represent the physical damage effects that are modeled.

2.4.2 ANN Use and Characteristics

The ANNs used in this study are feedforward perceptrons, which determine the optimal values of the parameters embedded in their framework to mimic the input/output relationship (Cao et al. 1998). Each data "point" used in the training of a perceptron can have multiple input terms and multiple output terms. In this study, the perceptrons train with data that has 26 measured input strains ($\{\varepsilon\}$) and 1 peak strain output ($\varepsilon_{\max}$). The perceptrons have two hidden layers: the first has a log-sigmoid function and the second has a linear function. The first hidden layer contains six neurons and the second hidden layer contains one, to match the number of output terms. Two hidden layers is common for modeling relationships that are

not overly complex, and the functions used for the neurons in each layer are also typically recommended because they can model almost any relationship when used in succession (Hagan et al. 1996).

The ANNs are trained with the Levenberg-Marquardt (LM) Algorithm with Bayesian Regularization (BR), as opposed to Steepest Descent, Newton's, or Conjugate Gradient Methods, for example. The combination of LM and BR is used to create the simplest possible model that captures the underlying trends of the data to prevent the model from overfitting the data (MathWorks 2017). The LM Algorithm is used because of its common use in the field and for its capability to train ANNs in a variety of circumstances (Kostić and Gül 2017, Hsieh and Mura 1995, and Hadi 2002). The objective function that the algorithm minimizes during its optimization is the sum of squared prediction error of all output terms. BR is applied to the LM algorithm to improve the ANNs' generalization capabilities and prevent overfitting by including an additional term summed in the objective function of the LM algorithm: the sum of squared weights of the ANN. Conceptually, weights represent how much an ANN values particular inputs. By penalizing an ANN for having more and larger weights (a complex model), BR incentivizes the ANN during training to become a simpler model which strikes a balance between fitting and generalization capabilities (Hagan et al. 1996). ANN training is conducted using the MATLAB Neural Network Toolbox.

2.4.3 ANN Model Training

The 27 measured input/peak strain output relationships that comprise the PMB bridge behavior are modeled independently from one another. The extraction of

data during a typical truck event is shown in Figure 4 for two of the configurations, along with the corresponding locations of the sensors designated as peak strain outputs: the light cyan points are the measured inputs and the dark blue points are the peak strain outputs. The data extraction shown in the figure is performed 27 times for each truck event, once for each sensor location as a single peak strain output with 26 measured strain inputs.

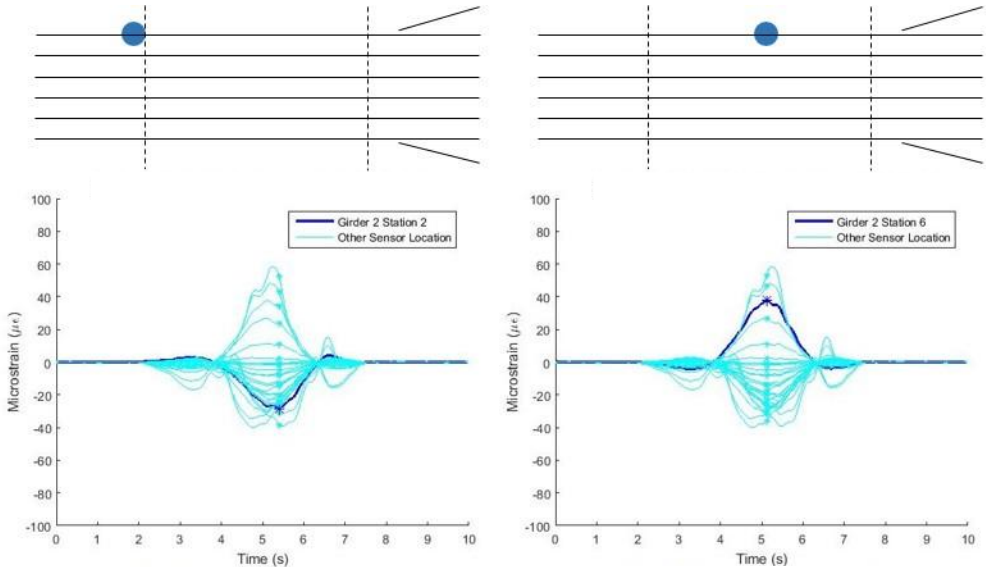

Fig. 4(a). Girder 2 Station 2 **Fig. 4(b).** Girder 2 Station 6

Fig. 4. Extraction of Measured Strain Inputs and Peak Strain Outputs

With the measured input and peak strain output extracted for all 27 configurations for all the recorded truck events, a bootstrapping scheme is implemented to determine the bridge behavior, using ANNs trained with the extracted inputs and outputs. Bootstrapping is a nonparametric numerical approach that can replace many analytical statistical methods (Efron and Tibshirani 1993). In structural health monitoring studies on the PMB, bootstrapping has been used as a resampling method that determines the confidence interval of a bridge parameter

value for the entire population of possible parameter values, given only a finite amount of sampled parameter values (Follen et al. 2014 and Reiff et al. 2016).

A benefit of bootstrapping is that it enables the evaluation of different subsets of data in an overall database. The approach can then better account for data error with adjustment for data outliers. The general process of bootstrapping is to (a) create multiple subsets of data from the total observed data set by random sampling with replacement, (b) calculate the value of a desired parameter for each subset of data, (c) create a distribution of the parameter value, with one instance in the distribution being determined from each subset of data, and (d) determine the confidence interval of the parameter value based on corresponding percentiles of the parameter value distribution. The bootstrapping scheme used in this study is described for determining only the expected bridge behavior at Girder 2 Station 6. However, the process is performed 27 times to determine the bridge behavior at all 27 sensor locations.

First, 420 of the 1,929 total recorded events are randomly set aside from model training as a control group to assess its performance after the bootstrapping scheme is completed. The same 420 events are set aside for the bridge behavior models at all sensor locations. Then, the remaining 1,509 events are randomly subdivided into two data sets 1,000 times. In each subdivision, 85% of the events are used to train an ANN, and 15% are set aside from training. Each trained ANN is a best fit of the measured data, based on the 85% of the data it was trained with. 15% is a typical portion of data to set aside from ANN training for model performance assessment (Hagan et al. 1996).

One ANN is trained for each of the 1,000 data subsets, the inputs ($\{\varepsilon\}$) and outputs ($\varepsilon_{\max}$) of which are as visualized in Figure 4(b) for Girder 2 Station 6. These 1,000 ANNs are herein referred to as *individual ANN models* for simplicity. All individual ANN models are slightly different from one another because each one was trained with different data. The performances of all individual ANN models are also slightly different from one another because (a) the models themselves are slightly different from one another, and (b) the performances the models are assessed with a different 15% of truck events set aside from training. An average of all 1,000 individual ANN models is taken to be the *final ANN model*. It is calculated as:

$$\overline{ANN}\left(\{\varepsilon\}\right) = \frac{1}{1000}\sum_{i=1}^{1000} ANN_i\left(\{\varepsilon\}\right) \tag{1}$$

where $\overline{ANN}\left(\{\varepsilon\}\right)$ is the final ANN model, and $ANN_1\left(\{\varepsilon\}\right)$ through $ANN_{1000}\left(\{\varepsilon\}\right)$ are the individual ANN models.

The final ANN model gives the mean expected peak strain during a truck event at Girder 2 Station 6 given the strains at all other sensor locations at the time at which Girder 2 Station 6 experiences its peak strain. To determine the *range* of expected peak strains, a prediction error analysis is performed on the 1,000 individual ANN models. Prediction errors for each model, for the 15% of data points set aside from training, are calculated as $E = \varepsilon_{\max} - ANN_i\left(\{\varepsilon\}\right)$, where E is the error, $\varepsilon_{\max}$ is the measured peak strain value, and $ANN_i\left(\{\varepsilon\}\right)$ is the predicted

value from the i^{th} individual ANN model. A distribution is then created that pools the prediction errors of all the individual ANN models of which the final ANN model is composed (Figure 5). This effectively represents the distribution of how much the measured data – which has not been used to train an individual ANN model – deviates from said trained individual ANN model.

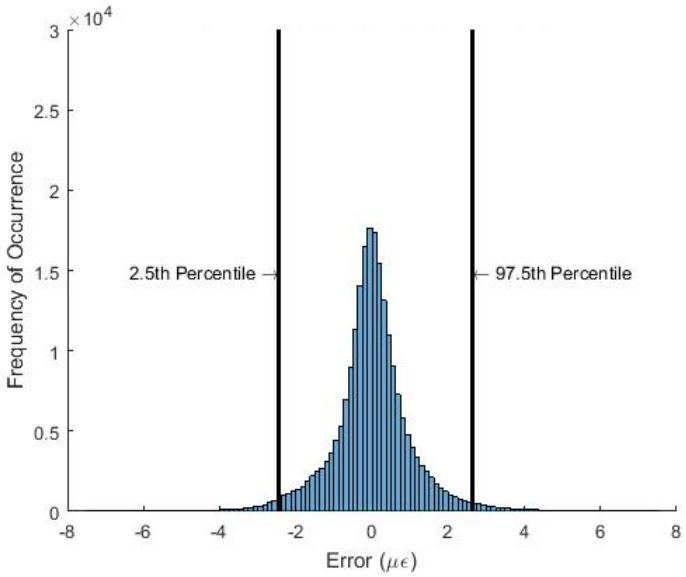

Fig. 5. Pooled Prediction Error Distribution at Girder 2 Station 6

The pooled error distribution is normal. Therefore, 95% of the prediction errors lie between the 2.5^{th} and 97.5^{th} percentiles, which are shown as vertical lines. The measured truck event data at this point is now bootstrapped to be representative of the entire population of all possible truck event data. Therefore, the 2.5^{th} and 97.5^{th} percentiles of the pooled prediction errors represent the range within which there is a 95% confidence that any newly measured peak strains will deviate from the final ANN model. Based on this assertion, the 95% confidence interval for the peak

strain at Girder 2 Station 6 is calculated for any newly recorded truck event as follows:

$$\begin{aligned} \overline{ANN}\left(\{\varepsilon\}\right)_{LB} &= \overline{ANN}\left(\{\varepsilon\}\right) + \varepsilon_{2.5} \\ \overline{ANN}\left(\{\varepsilon\}\right)_{UB} &= \overline{ANN}\left(\{\varepsilon\}\right) + \varepsilon_{97.5} \end{aligned} \tag{2}$$

where $\overline{ANN}\left(\{\varepsilon\}\right)_{LB}$ and $\overline{ANN}\left(\{\varepsilon\}\right)_{UB}$ are the lower and upper bounds of the 95% confidence interval of the final ANN model's peak strain prediction, respectively, and $\varepsilon_{2.5}$ and $\varepsilon_{97.5}$ are the 2.5^{th} and 97.5^{th} percentiles of the pooled prediction errors, respectively. This 95% confidence interval is the expected bridge behavior of the PMB at Girder 2 Station 6. Strain readings during many truck events outside this confidence interval would suggest that structural damage has occurred.

This process is valid under the assumption that the final ANN model is similar to the individual ANN models of which it is comprised, because the pooled error distribution that determines the confidence interval is derived from the performance of the individual ANN models, not the final ANN model. Each individual ANN model is trained with a random 85% of the available data, so there is not much potential for differences between the individual ANN models. This process is also valid under the assumption that the amount that peak strains deviate from the final ANN model is independent of the measured input values.

2.4.4 Expected Bridge Behavior

All confidence interval ranges are summarized in Table 1. Most sensor locations have small confidence interval ranges for their peak strain values – within ± 2 microstrains – which is only about twice the filtered ambient noise level. G6-

S8 is the largest outlier, potentially since that location is under the sidewalk and does not respond much to the roadway loadings. Other locations with large confidence intervals, such as G1-S6 and G3-S2, most have final ANN models that do not capture the structural behavior precisely. Overall, the confidence intervals imply that the peak strain values along the bridge are usually predicted reliably at most sensor locations.

Table 1. Peak Strain Confidence Interval Ranges at All Sensor Locations

Sensor Location	Confidence Interval Lower Bound (με)	Confidence Interval Upper Bound (με)
G1-S2	-0.864	0.794
G1-S6	-3.995	4.160
G1-S8	-0.836	0.722
G1-S10	-0.946	0.906
G2-S2	-1.363	1.253
G2-S4	-1.749	1.807
G2-S6	-2.439	2.620
G2-S8	-1.154	1.043
G2-S10	-1.078	1.038
G3-S2	-2.790	2.807
G3-S4	-0.981	0.947
G3-S6	-1.942	1.913
G3-S8	-1.056	0.979
G3-S10	-1.166	1.062
G4-S2	-1.282	1.154
G4-S4	-0.868	0.731
G4-S6	-1.853	1.998
G4-S8	-0.995	0.958
G4-S10	-1.348	1.231
G5-S2	-0.530	0.520
G5-S6	-1.464	1.552
G5-S8	-0.925	0.900
G5-S10	-3.039	2.865
G6-S2	-0.573	0.569
G6-S4	-0.542	0.511
G6-S6	-1.417	1.288
G6-S8	-8.024	8.024

The peak strain values and the corresponding 95% confidence intervals at all sensor locations for the 420 PMB truck events set aside from model training are predicted. The bridge behavior model performs well. A typical performance of the bridge behavior model is demonstrated with the successful peak strain predictions for all sensor locations, shown in Figure 6 for one of the 420 set-aside events. The predictions (blue), which are shown in the figure at each sensor station separately, match the measured values (green) well. Additionally, all measured values are within their confidence intervals.

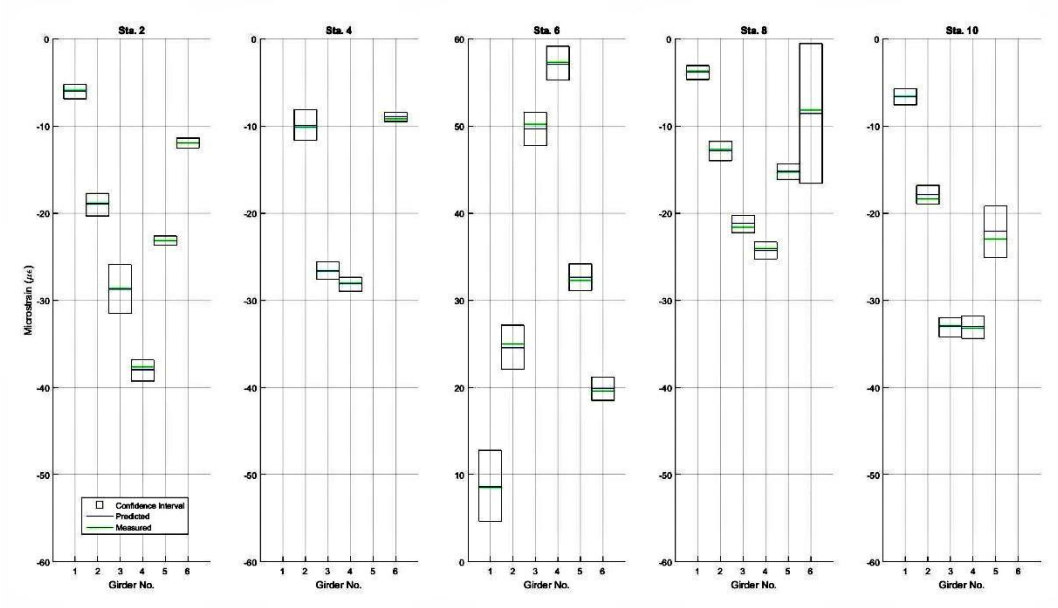

Fig. 6. Bridge Behavior Model Peak Strain Predictions for Typical Truck Event

The performance of the bridge behavior model is assessed with all 420 set-aside truck events, and the model is validated with the results (Table 2). The middle and right columns display the mean error magnitude and the error standard deviation, respectively, over all the 420 set-aside events at each location. The prediction errors of the peak strains at most locations are remarkably small, indicating that peak

strains at most of the used sensor locations on the bridge can be sufficiently predicted without any knowledge of the traffic that induced them.

Table 2. Bridge Behavior Model Performance

Sensor Location	Mean Error Mag. (με)	Error Std. Dev. (με)
G1-S2	0.319	0.594
G1-S6	1.201	4.308
G1-S8	0.301	0.603
G1-S10	0.349	0.713
G2-S2	0.453	0.816
G2-S4	0.526	0.895
G2-S6	0.771	1.524
G2-S8	0.384	0.655
G2-S10	0.413	1.393
G3-S2	0.957	1.422
G3-S4	0.329	0.536
G3-S6	0.694	1.453
G3-S8	0.339	0.743
G3-S10	0.421	0.647
G4-S2	0.362	0.560
G4-S4	0.247	0.414
G4-S6	0.653	1.489
G4-S8	0.356	0.950
G4-S10	0.422	0.839
G5-S2	0.193	0.494
G5-S6	0.397	0.858
G5-S8	0.255	0.405
G5-S10	0.805	1.927
G6-S2	0.169	0.243
G6-S4	0.154	0.206
G6-S6	0.384	0.725
G6-S8	1.683	3.152

2.4.5 Sufficiency of Number of Truck Events to Use Bootstrapping

In the bootstrapping scheme, it is desired for each of the 1,000 randomly resampled data subsets of truck events to sufficiently represent the variety and distribution of the entire population of all possible truck events. A data subset is considered stable and representative of the population when its distribution does

not change when extra data from the same population are added to it. The GDF of each girder as measured at Station 6 is the parameter used in this study to express the variety of truck events. Ghosn et al. (1986) defined the GDF of a girder as the strain experienced by the girder divided by the sum of strains experienced by all girders at the same transverse location:

$$GDF_i = \frac{\varepsilon_i}{\sum_{j=1}^{N} \varepsilon_j} \tag{3}$$

In this study, the strains used in the GDF_i calculation are sampled at the instant when Girder i experiences its peak strain.

One by one, the Girder 3 GDF (as an example) is calculated from a truck event randomly sampled from all 1,929 truck events without replacement, and is added to a pool of calculated Girder 3 GDFs. As each randomly sampled GDF gets added to the pool, the 0.2, 0.4, 0.6, and 0.8 quantiles of the GDF pool are calculated to assess the variety of truck events sampled up to that point. When the quantiles no longer change with the addition of extra truck events, then the amount of truck events is considered stable. A typical quantile plot is shown in Figure 7; the quantiles are sufficiently stable after 300 truck events. In this study, 420 events are set aside from training, and each individual ANN model used in the bootstrapping process trains with 85% of the 1,509 remaining truck events, which is 1,283 events. 1,283 truck events is greater than 300; therefore it is sufficient. The same conclusion is reached when this assessment is performed for each of the six girder GDFs.

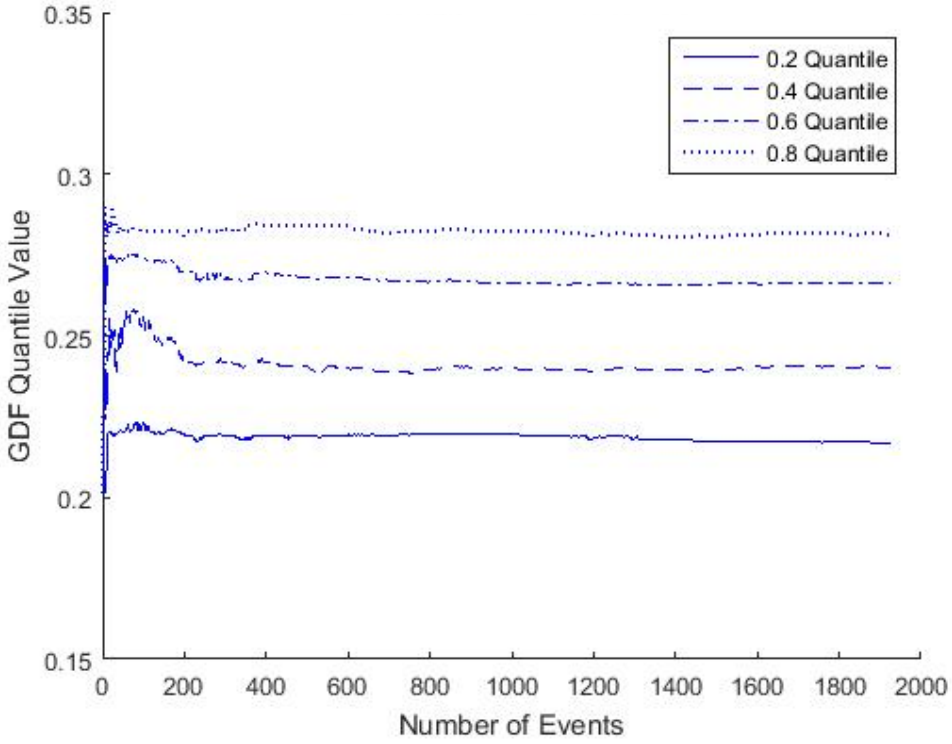

Fig. 7. Girder 3 GDF Quantile Stability Plot

2.5 Damage Effects on Operational Data

There are several ways a bridge's performance can change, such as a change in relative stiffnesses of bridge components caused by the removal of a utility pipe. However, in this study, structural damage is used as the example of altered performance. To assess the effectiveness of the method of damage identification proposed later, several damage scenarios are generated (*damage identification* herein refers to the detection and localization of damage). Since the PMB has no known damage, the effects of damage are simulated in a CSiBridge FEM which was created and calibrated by Sanayei et al. (2012). In this research, hypothetical damage scenarios are developed via the FEM data to illustrate how the proposed method can identify structural damage. The types of simulated damage scenarios

and naming scheme of said scenarios are based on those used in previous research on the PMB.

The first damage scenario is Case U, which contains no imposed damage (baseline condition). The second damage scenario, Case A, is interior girder fracture. It is modeled as a 2.5 mm-wide (longitudinally) section of Girder 2 from the midheight of the girder through the bottom flange that is altered to have an elastic modulus close to zero. The modeled girder crack was located at a diaphragm connection, close to the midspan of the girder. The third damage scenario, Case B, is fascia girder corrosion. It is modeled as a 30% reduction of the elastic modulus of the web and bottom flange of Girder 1 along its entire length. The last damage scenario, Case D, is deck delamination, which is modeled as a 35% reduction in the elastic modulus of deck concrete. The delaminated deck area in the FEM is 16 m in the longitudinal direction and transversely spans the entire southbound portion of the bridge over which trucks can drive. The delaminated area is centered longitudinally on the middle span of the model. Reiff et al. (2016) provided justification of the types, severities, and modeling of damage scenarios almost identical to those used in this research. The justification of their Case A can be traced to WisDot and FHWA (2001), Chajes et al. (2005), Kaufmann et al. (2004), and Farrar et al. (1994); the justification of their Case B can be traced to Enright and Frangopol (2000) and Miller et al. (2001); and the justification of their case D can be traced to Warhus et al. (1995).

For each damage case, two simulated HS20 trucks are run separately over the bridge FEM: one northbound truck centered in the northbound lane and one

southbound truck centered in the southbound lane. The effects of damage are quantified in a way that is applicable to the framework of bridge behavior defined in this study.

Damage effects in this study are defined as the percent change in measured inputs and peak strain outputs of all the bridge behavior models (bridge behavior at all sensor locations) between a bridge with damage and a bridge without damage. Several FEM truck runs are simulated to determine these damage effects. The measured input/peak strain output sampling process shown in Figure 4 is performed in the FEM for the bridge behavior model at all sensor locations, for all damage cases, for a northbound truck and then a southbound truck. For simplicity, *damage effects* herein refers to the damage effects for all 27 measured input/peak strain output configurations. Similarly, the *damage effects at Girder 2 Station 6*, for example, herein refers to the damage effects for the configuration with Girder 2 Station 6 as the designated peak strain output.

Since the calculated damage effects extracted from the FEM depend on the lane path of the truck, the damage effects applied to each operational PMB truck event depend on whether the truck at the PMB was southbound or northbound. A truck event is considered northbound when the peak strain at Girder 5 Station 6 is greater than that of Girder 2 Station 6.

Since the operational PMB truck events on which damage effects are imposed are grouped simply as being northbound events or southbound events, it is assumed that damage effects as observed during a truck event with a truck centered in a lane is sufficiently representative of the damage effects as observed during a truck event

with a truck transversely located anywhere in that lane. To test this assumption, an FEM sensitivity analysis is performed in which damage effects as observed during truck events with a variety of transverse truck locations are compared to each other. The results of the sensitivity analysis validate the assumption.

2.6 Damage Identification

2.6.1 Damage Identification Trials

The proposed use of damage detection in this study is largely for long-term monitoring of a bridge but can also be used for the rapid detection of sudden damage. Most structural deterioration is gradual. Ideally, a functioning bridge SHM system will provide long-term objective feedback which a bridge engineer can use to evaluate overall structural health and trends.

Fifty trials of a week's worth of simulated data are generated for each of the damage cases to assess how effectively damage is identified. The data of each trial is generated by randomly sampling, without replacement, 20 of the 420 truck events set aside from model training and applying damage effects to the 20 events. 20 events are sampled because approximately that many truck events (which can be utilized) are recorded at the PMB each week. The remaining 400 set-aside truck events are left as they are initially recorded, with no known damage. The 20 events of each trial represent approximately a week's worth of new passively collected data, and the remaining 400 events represent data that have been previously recorded when there was no known damage on the bridge, that was set aside from the model training.

First, the significance of a week's worth of events with respect to damage detection is assessed. Similar to the process shown in Figure 7, a quantile stability plot is produced for the peak strain prediction error magnitudes, averaged across all 27 sensor locations, for the 420 set-aside truck events (Figure 8). The number of events at which the quantiles stabilize (i.e., an almost constant quantile value with only small variations from additionally added data) represents the number of events required to fully convey how much the true PMB behavior deviates from the model – at least in its healthy state. Thus, it also represents the number of events required to decidedly determine whether damage is present. At the weekly average of twenty events, the quantiles are almost stable, but not completely so, since the quantile values change slightly when additional prediction errors are added to the pool of the first 20 prediction errors. This implies that the number of events in a typical week is not enough to *decidedly* detect damage rapidly in the case of sudden damage. However, it is enough to give an engineer a good impression of the behavior of the bridge, which is true because the damage detection method (assessed below) is extremely sensitive to structural changes. Additionally, if an engineer is observing slow deterioration over several years, as opposed to rapidly detecting sudden damage, then the number of events per week is sufficient because enough data is presented to the engineer over time.

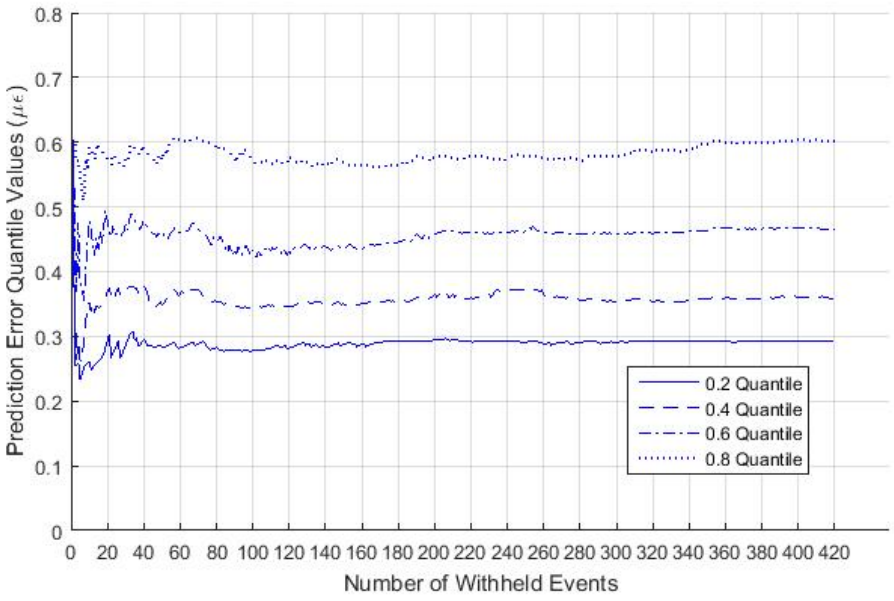

Fig. 8. Prediction Error Quantile Stability Plot

2.6.2 Damage Detection

The first component of damage identification is detection (Rytter 1993). Since there is a bridge behavior model at each of the 27 sensor locations, damage detection is performed individually for each model. The method of damage detection used in this study is herein described for only the bridge behavior at Girder 2 Station 6; however, the process described below is performed 27 times (once for the bridge behavior at each sensor location).

If damage exists in a generated trial of 20 "newly measured" truck events, then peak strains at Girder 2 Station 6 for those 20 events should deviate further from the final ANN model predictions than do sampled peak strains of events from its corresponding control group of 400 events. This comparison of deviation is done with the Wilcoxon rank-sum test. It is a nonparametric approach to determine the difference in medians between two sets of data (Wilcoxon, 1945). A one-sided test

is performed when there is an expectation of which data set is potentially larger; and a two-sided test is performed when it is unknown which data set is potentially larger (Helsel and Hirsch 2002). In this study, a two-sided test is used because the effects of damage on the measured data is unknown *a priori*. Additionally, when the smaller of the two data sets being compared is less than 10 samples, the exact method must be used (Helsel and Hirsch 2002); however, since the data sets being compared in this example both have more than 10 samples, the approximate method is used.

In a Wilcoxon rank-sum test, the null hypothesis, H_0, is that the medians of both sets of data are the same. The parameter p is the probability, when the null hypothesis is true, that the two data sets are at least as different from one another as they exist when compared in the rank-sum test. The value of p ranges from 0 to 1. The parameter α is the significance level. Any calculated p value below α demonstrates a rejection of H_0, and any calculated p value above α demonstrates a failure to reject H_0. The parameter h is calculated to be either 0 or 1. In this example, h would be 1 if H_0 is rejected and would be 0 if there is a failure to reject H_0. The compared parameter of the two sets of data in each of the 50 trials is the deviation of the peak strains at Girder 2 Station 6 from the final ANN model predictions. A Wilcoxon rank-sum test is performed on the two pools of data with a significance level of 0.1%, or 0.001. For each of the 50 trials, p and h values are calculated to determine whether damage exists according to the bridge behavior model at Girder 2 Station 6.

A summary of these values for all trials is provided in Table 3 for each damage case for Girder 2 Station 6, as well as all the other sensor locations. Values of p are averaged across all trials; in the table, these mean values demonstrate that damage is detected ($h = 1$) at that location on average if they are less than 0.001. The lower the mean p value, the more significant damage is on average according to the bridge behavior model at that location. For each scenario with induced damage, bridge behavior models at many locations are severely affected, as shown by the low mean p values. The table also displays the percentage of the 50 trials for each damage scenario in which damage is detected at each location. Damage is reliably detected at many locations when there is damage, and damage is not detected at all when there is no damage. Damage is considered detected in general for a week's worth of events if it is detected according to the bridge behavior model for at least one sensor location independently, because the bridge behavior model at each location is one way in which bridge behavior in general is defined and observed. Based on this assertion, damage is detected in 100% of the trials for cases in which there is damage, and damage is detected in 0% of the trials for Case U (no damage). This means there are no Type I or Type II errors in any of the trials with a significance level of 0.1%.

Table 3. Damage Detection Trial Results

Sensor Location	Mean p Value				% of Trials w/ Detected Dmg.			
	Case U	Case A	Case B	Case D	Case U	Case A	Case B	Case D
G1-S2	5.3E-01	3.0E-01	9.9E-05	2.4E-01	0	4	98	14
G1-S6	5.4E-01	5.2E-12	1.2E-07	6.0E-06	0	100	100	100
G1-S8	5.3E-01	5.5E-10	2.6E-01	1.5E-01	0	100	12	30
G1-S10	5.3E-01	5.8E-12	1.6E-12	3.7E-03	0	100	100	70
G2-S2	5.3E-01	2.8E-01	2.6E-01	5.0E-04	0	16	16	92

Sensor Location	Mean p Value				% of Trials w/ Detected Dmg.			
	Case U	Case A	Case B	Case D	Case U	Case A	Case B	Case D
G2-S4	4.5E-01	1.1E-04	1.1E-12	6.9E-13	0	96	100	100
G2-S6	4.3E-01	4.7E-14	1.9E-09	9.8E-02	0	100	100	36
G2-S8	4.9E-01	5.7E-04	1.7E-01	2.7E-08	0	92	20	100
G2-S10	4.8E-01	3.1E-01	2.6E-01	1.1E-01	0	4	16	22
G3-S2	5.1E-01	1.4E-01	1.3E-01	4.3E-01	0	24	26	2
G3-S4	5.4E-01	1.8E-01	5.5E-03	3.0E-01	0	26	96	4
G3-S6	5.0E-01	4.8E-13	5.0E-03	1.3E-02	0	100	78	68
G3-S8	4.9E-01	5.0E-09	4.6E-12	3.5E-01	0	100	100	6
G3-S10	5.0E-01	1.2E-02	4.9E-13	2.9E-01	0	70	100	8
G4-S2	5.3E-01	1.3E-02	6.9E-14	8.7E-03	0	60	100	58
G4-S4	5.1E-01	1.0E-04	3.2E-01	7.4E-11	0	98	6	100
G4-S6	4.1E-01	4.8E-13	2.4E-02	3.7E-01	0	100	78	4
G4-S8	5.0E-01	1.4E-12	5.9E-02	1.2E-06	0	100	68	100
G4-S10	5.0E-01	8.7E-02	3.7E-01	2.3E-03	0	30	4	86
G5-S2	5.0E-01	2.7E-01	2.7E-12	3.1E-02	0	12	100	42
G5-S6	4.9E-01	3.1E-01	1.1E-10	1.4E-01	0	10	100	18
G5-S8	5.0E-01	5.1E-01	3.7E-01	1.6E-08	0	0	4	100
G5-S10	4.5E-01	2.9E-08	1.4E-07	3.7E-01	0	100	100	4
G6-S2	5.3E-01	2.1E-03	2.9E-05	4.2E-02	0	80	98	28
G6-S4	5.4E-01	2.7E-01	7.0E-03	1.6E-02	0	6	74	52
G6-S6	4.7E-01	8.6E-02	2.4E-01	5.7E-02	0	24	10	36
G6-S8	5.0E-01	2.5E-01	1.9E-02	3.2E-01	0	6	50	0

2.6.3 Damage Localization

The second component of damage identification is localization (Rytter 1993). Figure 9 displays average results for the way damage is attempted to be visually localized for each damage case for one trial out of the fifty that are performed. The black dots are the sensor locations, and the vertical axes are the peak strain values of each of the 20 events in the trial, zeroed with respect to the bridge behavior model predictions. The predicted peak strain values are zero on the vertical axes, and the vertical black bars are the confidence intervals of the predictions (see Table 1). Each circle stemming out from zero on the vertical axis is the median deviation of a location's measured peak strains from their predicted values over the 20 events in the trial. A circle is green if its value is within the location's normalized

confidence interval, and a circle is red if its value is outside of it. They imply whether or not, respectively, the bridge behavior at each location is as expected. The authors hypothesize *a priori* that damage is localized based on how much each location's measured peak strains deviate from the predicted values, because the structural effects of damage are most severe close to the damage.

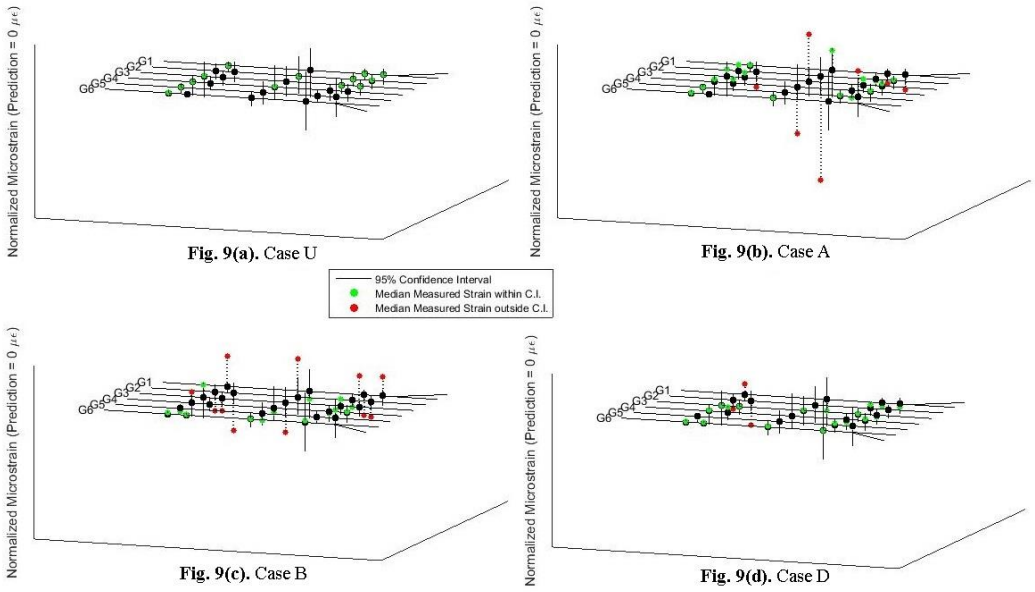

Fig. 9. Typical Week's Damage Localization for (a) No Damage, (b) Fracture, (c) Corrosion, and (d) Delamination

The proposed visual method successfully localizes the damage of Case A (Girder 2 fracture) both longitudinally and transversely, as shown in the typical trial. The measured peak strains deviate from the predicted values most severely at Girder 2 Station 6, which is adjacent to the damage, and then next most severely near that sensor location. The typical trial shown also demonstrates that the damage of Case B (Girder 1 corrosion) is successfully localized. Most of the median measured values that significantly deviate from their predicted values are along

Girder 1, which is corroded, and other locations of significant deviations are on the southbound half of the bridge.

The simulated damage is not effectively localized for Case D (deck delamination). Even though the sensor locations whose median peak strains fall outside of their normalized confidence intervals are on the delaminated side of the bridge, they are concentrated longitudinally at the south pier of the bridge, as opposed to longitudinally distributed over the delaminated portion. This may be because the damage was not significant enough to force enough locations' peak strain measurements outside of their confidence intervals, even though it was significant enough for several locations to detect damage. Lastly, as expected, the measured values for Case U (no damage) do not significantly deviate from the predictions.

2.6.4 Further Discussion of Results

The 0.1% significance level used in the Wilcoxon rank-sum tests is arbitrary. The value used in this study is one that effectively emphasizes the contrast in bridge behaviors between an undamaged and damaged structure. This issue is further addressed later.

Even though the effects of non-catastrophic damage on the PMB would not be physically "experienced" very much at sensors far from the actual damage, Table 3 demonstrates that damage is still typically detected in this study at most of the sensor locations. It is hypothesized that this occurs because of the way in which bridge behavior is defined: strains at sensor locations are not observed in isolation; they are observed in relation to one another. There are two ways in which damage

can be detected in this framework. (1) All the measured inputs that a final ANN model deems relevant are far from the damage, therefore their values are “normal.” The predicted peak strain output value is thus “normal” as well. The output sensor location is close to the damage, so the measured value is “not normal” and thus deviates from the model prediction. (2) Some of the measured inputs that the trained ANN model deems relevant are near the damage, therefore their values are “not normal.” The predicted peak strain output value is thus “not normal” as well. The output sensor location is far from the damage, therefore the measured value is “normal” and thus deviates from the model prediction.

A phenomenon similar to that may also explain why measured peak strains during damage scenarios significantly deviate from their predicted values when they are close to damage, but do not necessarily deviate in the way that is expected. The typical damage localization trial for Case B (corrosion) exemplifies this. It is expected that all the median measured peak strains along Girder 1 would be lower than their predictions, since Girder 1 carries less load due to a loss of relative stiffness when it is corroded. However, in the damage localization plot in Figure 9(c), all the Girder 1 median measured peak strains are higher than their predictions.

2.7 Conclusion

When tested against new data, the ANN-learned model of bridge behavior is shown to be effective and applicable to a variety of traffic events with unknown operational loading conditions. The proposed method of damage identification is remarkably successful at detecting damage, and damage is effectively localized for two out of three damage scenarios.

There is still much work to be done to enhance this new approach to quantifying and assessing bridge behavior. However, the preliminary results of this study demonstrate that it has the capability to be effective in the future for engineers to easily determine the baseline performance of bridges and receive automated weekly updates on bridges' statuses. This would lead to more informed decision-making, which means that time and funding would be allocated more efficiently to maintain and improve US infrastructure.

2.8 Future Work

One issue that can be addressed in future work is the *a priori* determination of an appropriate significance level for damage detection for any bridge like the PMB, potentially based on the characteristics of the bridge and its instrumentation. Future work can also assess other input/output relationships, such as temperature inputs and strain outputs at all sensor locations when no trucks are on the bridge. Only one final ANN model would be needed to define the bridge behavior, rather than 27, because all the temperatures across the bridge can be used as direct predictors of all the strains across the bridge.

2.9 References

Alampalli, S. (2000). "Effects of Testing, Analysis, Damage, and Environment on Modal Parameters." *Mechanical Systems and Signal Processing*, 14(1), 63-74.

American Society of Civil Engineers (2017). "2017 Infrastructure Report Card." <https://www.infrastructurereportcard.org/wp-content/uploads/2017/01/Bridges-Final.pdf> (Jun. 28, 2017).

Cao, X. et al. (1998). "Application of artificial neural networks to load identification." *Computers & Structures*, 69(1), 63-78.

Cardini, A.J., and DeWolf, J.T. (2009). "Long-term Structural Health Monitoring of a Multi-girder Steel Composite Bridge Using Strain Data." *Structural Health Monitoring*, 8(1), 47-58.

Chajes, M. et al. (2005). "Steel Girder Fracture on Delaware's I-95 Bridge over the Brandywine River." *Proc., ASCE Structures Congress,* ASCE, New York.

Chakraborty, S., and DeWolf, J.T. (2006). "Development and Implementation of a Continuous Strain Monitoring System on a Multi-Girder Composite Steel Bridge." *J. Bridge Eng.*, 10.1061/(ASCE)1084-0702(2006)11:6(753).

Efron, B., and Tibshirani, R. (1993). *An Introduction to the Bootstrap*, Chapman & Hall, Inc.

Enright, M.P., and Frangopol, D.M. (2000). "Survey and Evaluation of Damaged Concrete Bridges." *J. Bridge Eng.*, 10.1061/(ASCE)1084-0702(2000) 5:1(31).

Farrar, C.R. et al. (1994). *Dynamic Characterization and Damage Detection in the I-40 Bridge Over the Rio Grande*, Los Alamos National Laboratory, Los Alamos.

Follen, C.W. et al. (2012). "Statistical Bridge Signatures." *J. Bridge Eng.*, 19(7), 10.1061/(ASCE)BE.1943-5592.0000596.

Ghosn, M. et al. (1986). "Evaluation of Steel Bridges Using In-Service Testing." *Transportation Research Record*, 71-78.

Hadi, M. (2003). "Neural networks applications in concrete structures." *Computers & Structures*, 81(6), 373-381.

Hagan, M. et al. (1996). *Neural Network Design*, 2nd Ed., PWS Publishing Co., Boston.

Helsel, D.R., and Hirsch, R.M. (2002). *Statistical Methods in Water Resources Techniques of Water Resources Investigations*, Book 4, Ch. A3, U.S. Geological Survey.

Hsieh, S.C., and Mura, T. (1993). "Design of a Statically Indeterminate Truss with Specified Creep Deflections." *Computers & Structures*, 54(5), 921-924.

Kaufmann, E. et al. (2004). *Failure Analysis of the US 422 Girder Fracture*, ATLSS Reports, Bethlehem.

Kostić, B., and Gül, M. (2017). "Vibration-Based Damage Detection of Bridges under Varying Temperature Effects Using Time-Series Analysis and

Artificial Neural Networks." *J. Bridge Eng.*, 10.1061/(ASCE)BE.1943-5592.0001085.

Kromanis, R., and Kripakaran, P. (2012). "Predicting thermal response of bridges using regression models derived from measurement histories." *Computers & Structures*, 136, 64-77.

Lam, H. et al. (2006). "Structural Health Monitoring via Measured Ritz Vectors Utilizing Articial Neural Networks." *Computer-Aided Civil and Infrastructure Engineering*, 21(4), 232-241.

MarketsAndMarkets (2017). "Structural Health Monitoring Market by Solutions (Hardware: Sensors, Data Acquisition System; Software & Services), Technology (Wired and Wireless), End Users and Geography – Global Forecast to 2022." <http://www.marketsandmarkets.com/Market-Reports/structural-health-monitoring-market-101431220.html?gclid=CjwKEAjw1a3KBRCY9cfsmdmWgQ0SJAATUZ8bp9ax5YeUoLWhO7AaDtPqae4HgiV5-zugdEWLlKzIhBoC6KTw_wcB> (Jul. 15, 2017).

MathWorks (2017). "Improve Neural Network Generalization and Avoid Overfitting." <https://www.mathworks.com/help/nnet/ug/improve-neural-network-generalization-and-avoid-overfitting.html> (Jun. 7, 2017).

Miller, T. et al. (2001). "Strengthening of a Steel Bridge Girder Using CFRP Plates." *J. Bridge Eng.*, 10.1061/(ASCE)1084-0702(2001) 6:6(514).

Mehrjoo, M. et al. (2008). "Damage detection of truss bridge joints using Artificial Neural Networks." *Expert Systems with Applications*, 35(3), 1122-1131.

Moaveni, B., and Behmanesh, I. (2012). "Effects of changing ambient temperature on finite element model updating of the Dowling Hall Footbridge." *Engineering Structures*, 43, 58-68.

Moore, M. et al. (2001). *Reliability of Visual Inspection for Highway Bridges, Volume I: Final Report*. Federal Highway Administration.

Moser, P., and Moaveni, B. (2011). "Environmental effects on the identified natural frequencies of the Dowling Hall Footbridge." *Mechanical Systems and Signal Processing*, 25(7), 2336-2357.

Reed, S. (2009). *Encyclopedia of Structural Health Monitoring*, John Wiley & Sons, Ltd., Chichester.

Reiff, A. et al. (2016). "Statistical bridge damage detection using girder distribution factors." *Engineering Structures*, 109, 139-151.

Rytter, A. (1993). *Vibrational Based Inspection of Civil Engineering Structures*, Department of Building Technology and Structural Engineering, Aalborg University, Aalborg.

Sanayei, M. et al. (2012). "Instrumentation, Nondestructive Testing, and Finite-Element Model Updating for Bridge Evaluation Using Strain Measurements." *J. Bridge Eng.*, 10.1061/(ASCE)BE.1943- 5592.0000228.

Shenton III, H.W., and Hu, X. (2006). "Damage Identification Based on Dead Load Redistribution: Methodology." *J. Struct. Eng.*, 10.1061/(ASCE)0733-9445(2006)132:8(1254).

Warhus, J.P. et al. (1995) "Imaging radar for bridge deck inspection." *SPIE*, 2456, 296-305.

Wilcoxon, F. (1945). "Individual Comparisons by Ranking Methods." *Biometrics Bulletin*, 1(6), 80-83.

Wisconsin Department of Transportation and Federal Highway Administration (2001). *Hoan Bridge Forensic Investigation Failure Analysis Final Report*.

Zhang, G. et al. (2012). "Automatic Delamination Detection of Concrete Bridge Decks Using Impact Signals." *J. Bridge Eng.*, 10.1061/(ASCE)BE. 1943-5592.0000326.

Chapter 3. Truck Event Selection Process

3.1 Overview of Truck Events Selection

The two main objectives of this study are to (a) train ANNs to mimic the way a bridge creates a distribution of traffic-induced strains from a single vehicle and to (b) assess the damage identification capability of an approach that utilizes the trained ANNs. These objectives require the removal of many types of events from the pool of operational data with which the ANNs can potentially be trained. Undesirable traffic events are identified herein by assessing the distribution and timing of the strains at the bottom flanges of all six girders at Station 6, which is close to the midspan of the middle span of the bridge. SG-6, SG-22, SG-42, SG-62, SG-82, and SG-96 are used for Girders 1 through 6, respectively. Sanayei et al. (2012) provided a detailed summary of the instrumentation of the bridge. The strains at these sensor locations are assessed in all the following truck event selection rules, unless otherwise noted, as they are used in this study: with a moving average filter with a window of $7/50^{th}$ of a second applied to the raw strain data. The order in which undesirable truck events are identified and removed are the same as it is presented herein. 33,985 total events are removed, the majority of which according to the requirements described in Section 3.2.

3.2 Removal of Light Vehicle Events

The types of vehicles that cross the PMB are divided into two categories: light and heavy (Follen at al. 2014). Only data from heavy vehicles are used, because the signal-to-noise ratio is higher. Based on their study, events are deemed as heavy truck events when the maximum overall strain experienced by any girder at Station

6 is greater than 40 microstrains. Figure 10 shows the strains of all the girders at Station 6 for a light vehicle event.

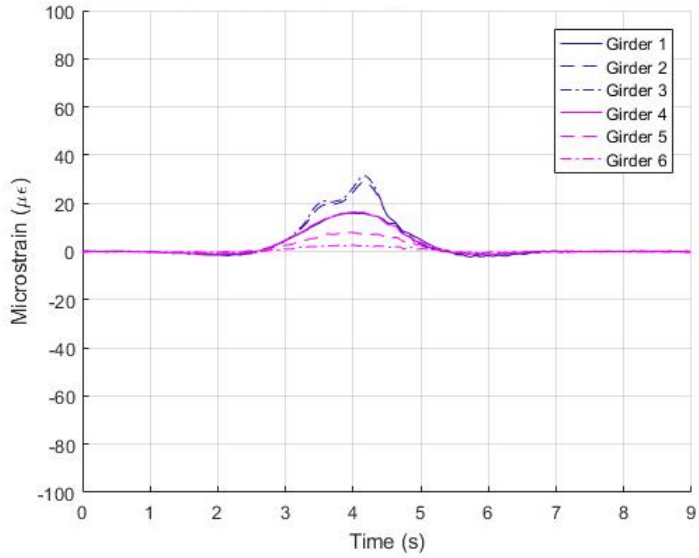

Fig. 10. Typical Event with Light Vehicle

3.3 Removal of Events with Excessive Head or Tail Energy

Truck events with excessive strain energy anywhere on the bridge (but observed at Station 6) at the beginning or end of the event caused by a second vehicle are removed, because the objective of the ANN training is to simulate how the bridge distributes a single vehicle load – not several vehicle loads – to many sensor locations. Events are flagged as such when any of the girders have a variance of strain values in the first or last three seconds of the recorded event that is greater than 8 microstrains. A typical event flagged for this reason is plotted in Figure 11 for Station 6. In the figure below, the strain values dipping below zero after the main truck event on the center span has ended signify that a new truck has crossed

onto the bridge – causing negative bending at Station 6 – before the main truck completely exits the bridge.

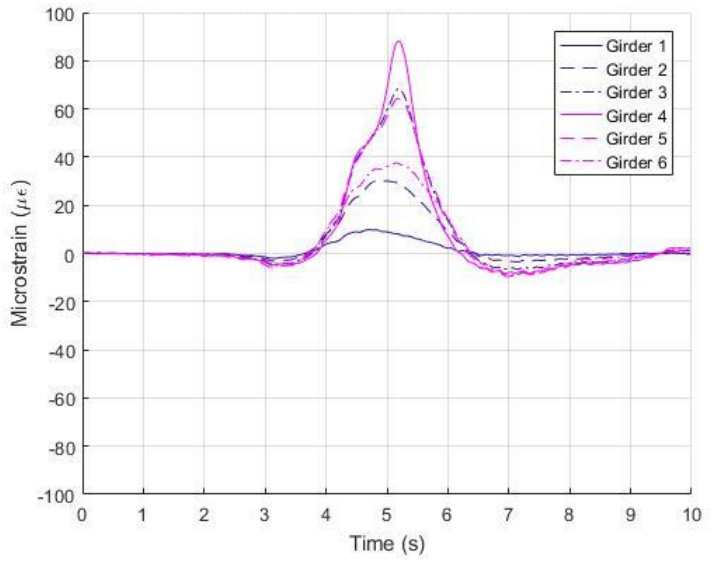

Fig. 11. Typical Event with Excessive Head or Tail Energy

3.4 Removal of Events with Centered Truck or Two Opposing Trucks at Midspan

Events are removed if they are triggered by either (a) two trucks travelling in opposite directions that pass each other close to midspan of the middle span or (b) a single truck that straddles both lanes. Reason (a) is necessary because of the same reason described in Section 3.3. Reason (b) is necessary because damage is simulated in this study only for trucks in the northbound lane and in the southbound lane, not for trucks straddling the centerline. An event is flagged under this rule if it meets all the following criteria: (a) the peak strain of Girder 3 is the maximum overall strain of all the girders; (b) the smaller of the peak strains of Girders 2 and 4 is greater than the larger of the peak strains of Girders 1 and 5; and (c) the larger

of the peak strains of Girders 2 and 4 is greater than 2/3 of the peak strain of Girder 3. A typical event removed for this reason is shown in Figure 12 for all the girders at Station 6. Since Girder 3 clearly experiences the maximum strain, the truck event in the figure must be triggered by either a truck straddling Girder 3 or two trucks that meet close to midspan of the middle span.

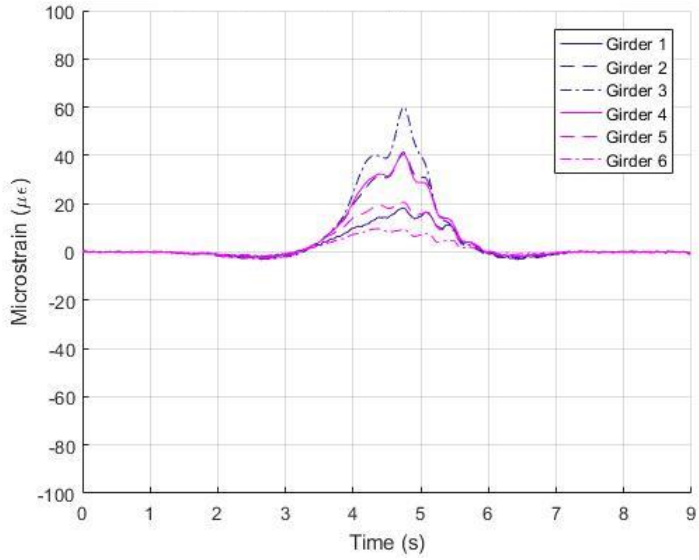

Fig. 12. Typical Event with Centered Truck or Two Opposing Trucks at Midspan

3.5 Removal of Events with Two Aligned Trucks in Succession

For the same reason described in Section 3.3, events are removed when they are triggered by two successive trucks in the same lane. To identify these events, definitions of key strain values – which pertain to this particular rule – must first be laid out. A *peak* is any strain data point whose two adjacent points are of a lower value than it is. A *main peak* is the peak of highest value of a single girder during an event. A *secondary peak* is any peak that is not a main peak. A *valley* is any strain data point whose two adjacent points are of a higher value than it is. A peak's

prominence is the absolute difference between the peak's value and the highest-value valley adjacent to it. A valley's *prominence* is the absolute difference between the valley's value and the lowest-valued peak adjacent to it.

For this particular truck event removal rule, a moving average filter with a window of $17/50^{th}$ of a second is applied to the raw strain data. This window ensures that any strain peaks and valleys caused by noise in the data are removed, but ensures that the relationships between all the important peaks and valleys are preserved. Events are flagged for removal if the girder with the maximum overall strain of all the girders has a secondary peak that meets all the following criteria: (a) it is at least 40% of the magnitude of the girder's main peak; (b) it occurs between 0.75 seconds and 8 seconds before or after the occurrence of the girder's main peak; and (c) it has any valley between itself and the girder's main peak that has a prominence of at least 45% of the prominence of the girder's main peak. An event removed for this reason is shown in the Figure 13, with the default moving average window of $7/50^{th}$ of a second applied for viewing consistency. In the figure, the two "bumps" signify two trucks crossing the bridge. The figure also confirms that both trucks are travelling in the same direction because percentage of load distributed to each girder is approximately the same for both bumps.

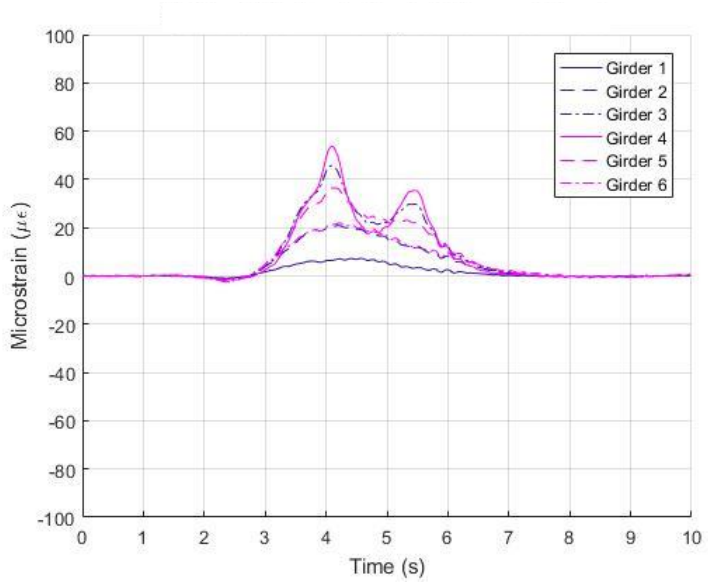

Fig. 13. Typical Event with Two Successive Aligned Trucks

3.6 Removal of Events with Two Opposing Trucks not at Midspan

For the same reasons described in Section 3.3, events are removed when they are triggered by two trucks travelling in opposite directions that pass each other somewhere other than near the midspan of the middle span of the bridge. To identify these events, the same definitions of key strain values used in Section 3.5 are used here. Additionally, the same moving average window used in that section is used here, with the same reasoning as well. A truck event is removed if any girder whose main peak ranks 3rd, 4th, or 5th in value among all the girders' main peaks has a main peak that meets both of the following criteria: (a) it occurs between 1 second and 8 seconds before or after the occurrence of the overall main peak of all the girders; and (b) it has a magnitude of at least 15 microstrains when its raw strain data is filtered with a moving average window of 7/50th of a second. Such an event is shown in Figure 14 with a moving average window of 7/50th of a second for

viewing consistency. Similar to those in Figure 13, the two tallest "peaks" from Girder 2 signify two trucks on the bridge, with at least one being in the southbound lane. However, the peak strain of Girder 5 is offset from the other peaks by about 1.5 seconds, and Girder 3 has the overall maximum strains of all the girders. This implies that one of the two trucks is traveling northbound.

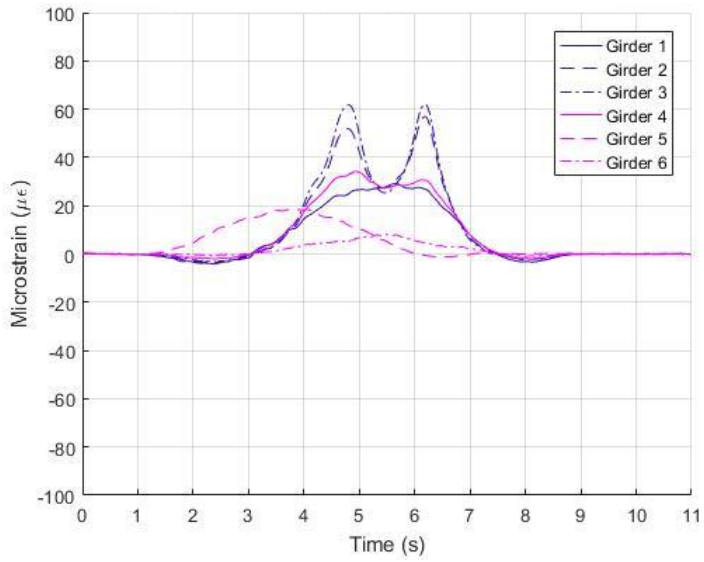

Fig. 14. Typical Event with Two Opposing Trucks not at Midspan

3.7 References

Follen, C.W. et al. (2012). "Statistical Bridge Signatures." *J. Bridge Eng.*, 10.1061/(ASCE)BE.1943-5592.0000596.

Sanayei, M. et al. (2012). "Instrumentation, Nondestructive Testing, and Finite-Element Model Updating for Bridge Evaluation Using Strain Measurements." *J. Bridge Eng.*, 10.1061/(ASCE)BE.1943- 5592.0000228.

Chapter 4. Artificial Neural Network Use

4.1 Classification and Architecture

The type of ANN used in this study is the feedforward perceptron, which learns the relationship between a set of inputs and outputs by determining the optimal values of the parameters embedded in its framework that form the input-output relationship (Cao et al. 1998). An idealized perceptron provided by Hagan et al. (1996) is shown in Figure 15. Each data "point" used in the training of a perceptron can have multiple input terms and multiple output terms.

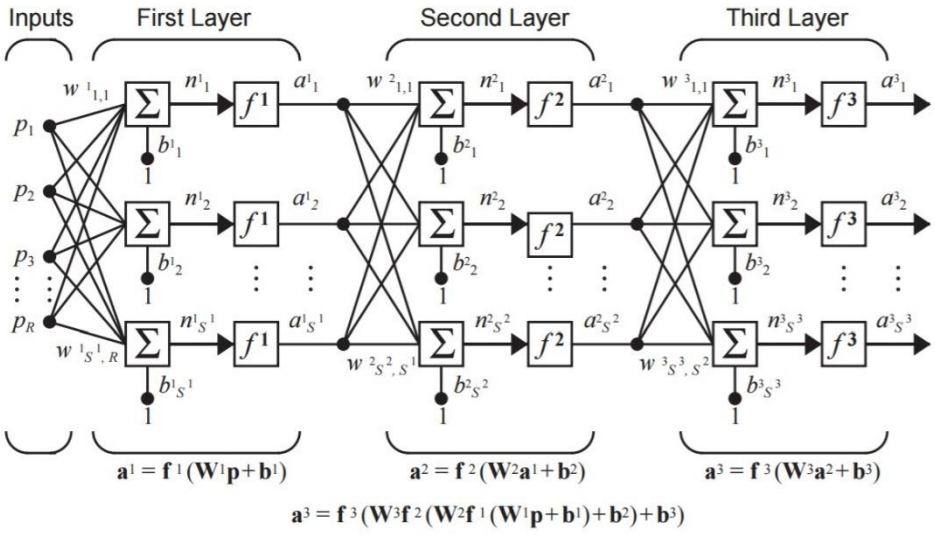

Fig. 15. Idealized Perceptron (Hagan et al. 1996)

Output terms are predicted based on input terms with a complex net of functions that flow from layer to layer. The input of a single data "point" according to the idealized model is an R x 1 vector, with R input terms. The first neural layer has s^1 neurons. Each neuron in the first layer (a) multiplies each input by a unique weight, (b) sums the products of the inputs and their corresponding weights, (c)

adds a bias to the sum, and (d) passes the final sum through a function. At the end of the first layer, there are s^1 intermediate terms, which are designated as a^1_1 through $a^1_{s^1}$. All the intermediate terms can be represented as the s^1 x 1 vector $\{a^1\}$. This process is repeated for each neural layer, where the input of each layer is the intermediate output of the previous layer. Each layer can have a unique number of neurons with their own unique weights and biases. The transfer function can vary between layers too. This process is repeated until the final layer, the output, is calculated. The output of a single data "point" according to the idealized model is an s^3 x 1 vector, with s^3 output terms. The number of neural layers is also customizable.

The perceptrons used in this study have 26 input terms and 1 output term. In general, an ANN will have better generalization capabilities with less neural layers and less neurons in each layer. The objective of the neural network architect is to create the simplest possible model that captures the underlying trends of the data (MathWorks 2017). Each ANN in this study has two hidden layers, which are neural layers in between the input layer and the output layer. The first hidden layer has 6 neurons, so it receives 26 inputs from the input layer and has 6 intermediate outputs. The second hidden layer has 1 neuron, so it receives 6 intermediate inputs from the first hidden layer and has 1 output to the output layer. The first-layer neurons use a log-sigmoid function, and the second-layer neuron uses a linear function. Two hidden layers is common for modeling relationships that are not overly complex, and the functions used for the neurons in each layer are also typically recommended (Hagan et al. 1996).

4.2 Optimization and Regularization

ANN training is an iterative process by which the ANN model improves to fit the training data as effectively as possible. The iterative process is a loop of assessing how well the ANN fits the data with its current set of weights and biases and adjusting the weights and biases for the next iteration based on its performance in the current iteration. ANNs can update their parameters with many different optimization algorithms. The one used in this study is the Levenberg-Marquardt Algorithm. The LM Algorithm is used because of its common use in the field and for its capability to train ANNs in a variety of circumstances (Kostić and Gül 2017, Hsieh and Mura 1995, and Hadi 2002). Hagan et al. (1996) provided a detailed explanation of it.

Bayesian Regularization is implemented to affect the way in which the Levenberg-Marquardt Algorithm is used: it ensures a balance between data-fitting and generalization capabilities. Without Bayesian Regularization, the objective function used with the Levenberg-Marquardt Algorithm is the sum of squared errors of all output terms. Bayesian Regularization introduces an additional objective function term that is added to the sum of the squared errors: the sum of the squared weights. By penalizing the network for having more weights, the ANN is incentivized in its training to make more of its weights closer to zero. Weights with values closer to zero effectively eliminates parameters from its model, which simplifies the ANN and improves its generalization capability. This helps prevent the ANN from overfitting the data it trains with. Hagan et al. (1996) provided an in-depth statistical justification of the use of this regularization.

4.3 References

Cao, X. et al. (1998). "Application of artificial neural networks to load identification." *Computers & Structures*, 69(1), 63-78.

Hadi, M. (2003). "Neural networks applications in concrete structures." *Computers & Structures*, 81(6), 373-381.

Hagan, M. et al. (1996). *Neural Network Design*, 2nd Ed., PWS Publishing Co., Boston.

Hsieh, S.C., and Mura, T. (1993). "Design of a Statically Indeterminate Truss with Specified Creep Deflections." *Computers & Structures*, 54(5), 921-924.

Kostić, B., and Gül, M. (2017). "Vibration-Based Damage Detection of Bridges under Varying Temperature Effects Using Time-Series Analysis and Artificial Neural Networks." *J. Bridge Eng.*, 10.1061/(ASCE)BE.1943-5592.0001085.

MathWorks (2017). "Improve Neural Network Generalization and Avoid Overfitting." <https://www.mathworks.com/help/nnet/ug/improve-neural-network-generalization-and-avoid-overfitting.html> (Jun. 7, 2017).

Chapter 5. Simplified Analogy of Bridge Behavior and Damage Detection

5.1 ANN Model Training

ANN models in Chapter 2 are developed to mimic the structural behavior of the PMB. However, the nature and development of the ANN models are complex. Therefore the training, assessment, and use of the models is performed herein with a simplified analogy of the PMB operational data to provide a conceptual understanding.

A hypothetical structure is assumed to have an input-output relationship of $y = 2\sqrt{x} + z$, as measured by an observer, where $y = 2\sqrt{x}$ is the analytical relationship, and z is a randomized, normally distributed noise term with a mean of 10 and a standard deviation of 7. The input-output relationship represents the relationship between two measured variables, such as the response at two different locations as the structure experiences loading. The expected output value based on an input value is the definition of the structure's behavior. The "measured" data are the individual data points in Figure 16.

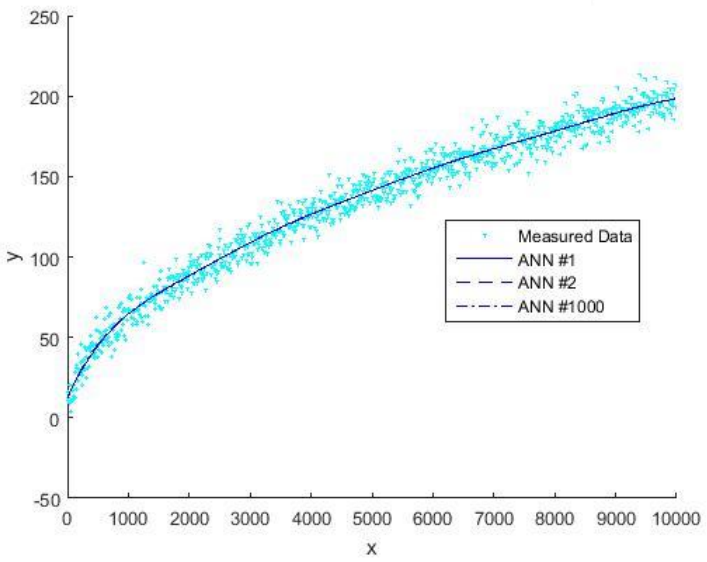

Fig. 16. Measured Data and Individual ANN Models

There are 1,000 measured data points. The input variable x spans from 0 to 10,000 in uniform increments of 10. It is also shown that the noise term z introduces uncertainty of the input-output relationship of the structure. It is desired to determine the range of expected behavior of the structure, so a bootstrapping scheme is used.

In this analogy, it is not desired to determine the range of expected values of y regardless of x; rather, it is desired to determine the range of expected values of any single y given a single value of x upon which the value of y depends. First, the existing set of 1,000 measured data points is randomly subdivided into two data sets 1,000 times. In each subdivision, 85% of the measured data points are used to train an ANN, and 15% are set aside from training to assess the performance of the trained ANN. Each trained ANN is a best fit of the measured data, based on the 85% of the data it was trained with. Each data "point" used in the training of a

perceptron can have multiple input terms and multiple output terms; however, this simplified analogy only relates one input term and one output term: x and y.

One ANN is trained for each of the 1,000 data subsets. These ANNs are herein referred to as *individual ANN models* for simplicity. The curves of a few of the individual ANN models are shown in Figure 16. All the models are slightly different from one another because each individual model was trained with different data. The performances of all individual ANN models are also different from one another because (a) the models themselves are different from one another, and (b) the performances the models are assessed with a different 15% of the data points set aside from training.

As shown in Figure 16, the individual ANN models do not differ much from one another. An average of all the individual ANN models is taken to be the *final ANN model* of the input-output relationship of the structure, which is ultimately very similar to the individual ANN models from which it is derived. It is calculated as:

$$\overline{ANN}(x) = \frac{1}{1000}\sum_{i=1}^{1000} ANN_i(x) \tag{4}$$

where $\overline{ANN}(x)$ is the final ANN model, and $ANN_1(x)$ through $ANN_{1000}(x)$ are the individual ANN models.

5.2 Expected Behavior

The final ANN model gives the expectation of y given any x value. To determine the range of expected y values given any x value, a residuals analysis

is performed on the 1,000 individual ANN models. Each model has prediction errors for the 15% of data points set aside from training, which are calculated as $\varepsilon = y - ANN_i(x)$, where ε is error, y is the measured value, and $ANN_i(x)$ is the predicted value from the i^{th} individual ANN model.

A distribution is then created that pools the prediction errors of the 15% of set-aside data points of all the individual ANN models of which the final ANN model is composed (Figure 17). This effectively represents the distribution of how much the measured data – which has not been used to train an individual ANN model – deviates from said trained individual ANN models.

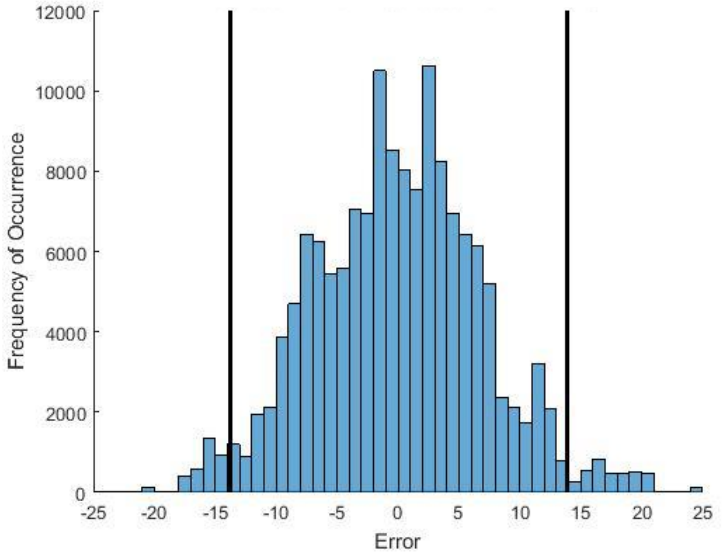

Fig. 17. Pooled Prediction Error Distribution and 95% Confidence Interval

The pooled error distribution is approximately normal. Therefore, 95% of the prediction errors lie between the 2.5$^{\text{th}}$ and 97.5$^{\text{th}}$ percentiles, which are shown as vertical lines. The measured data at this point has now been bootstrapped to be representative of the entire population of all possible data points. Therefore, the

2.5th and 97.5th percentiles of the pooled prediction errors representing the 95% confidence range that any newly measured data point may deviate from the final ANN model. Based on this assertion, the 95% confidence interval for the value of y given any newly measured value of x is calculated as follows:

$$\begin{aligned} \overline{ANN}(x)_{LB} &= \overline{ANN}(x) + \varepsilon_{2.5} \\ \overline{ANN}(x)_{UB} &= \overline{ANN}(x) + \varepsilon_{97.5} \end{aligned} \tag{5}$$

where $\overline{ANN}(x)_{LB}$ and $\overline{ANN}(x)_{UB}$ are the lower and upper bounds of the 95% confidence interval of the final ANN model's prediction, respectively, and $\varepsilon_{2.5}$ and $\varepsilon_{97.5}$ are the 2.5th and 97.5th percentiles of the pooled prediction errors, respectively (shown in Figure 18). This is effectively the 95% confidence interval of the expected behavior of the structure.

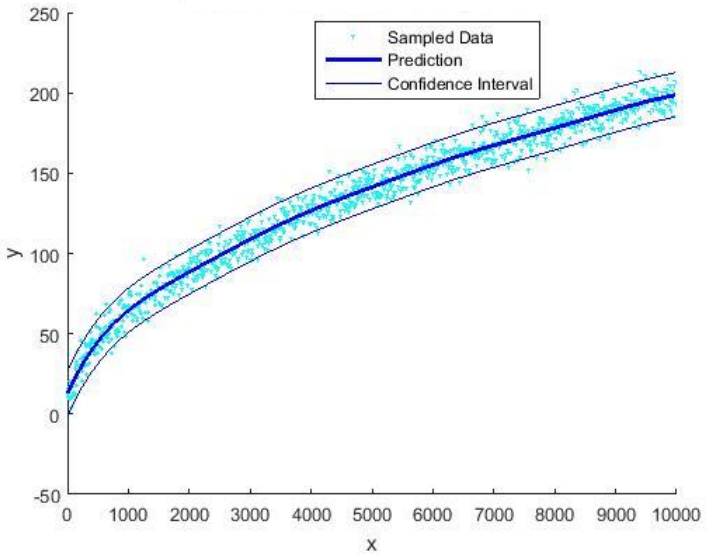

Fig. 18. 95% Confidence Interval of y Prediction Given x

This calculation is valid under the assumption that the amount that newly measured data points deviate from the final ANN model is independent of x. It is also valid under the assumption that the final ANN model is similar to the individual ANN models of which it is comprised, because the pooled error distribution that determines the confidence interval is derived from the performance of the individual ANN models, not the final ANN model.

5.3 Damage Detection

Hypothetical damage is imposed on the structure after the data points have been sampled from the undamaged structure. The input-output relationship with damage is represented mathematically as $y = 2\sqrt{x} + 10 + z$, where 10 is an offset caused by damage. Without the prior knowledge that damage exists, new samples must be measured to determine whether damage is present. A total of 50 new samples are taken after damage is imposed (Figure 19(a)). This sampling is represented by the newly introduced input-output equation, where x is uniformly spaced between 0 and 10,000.

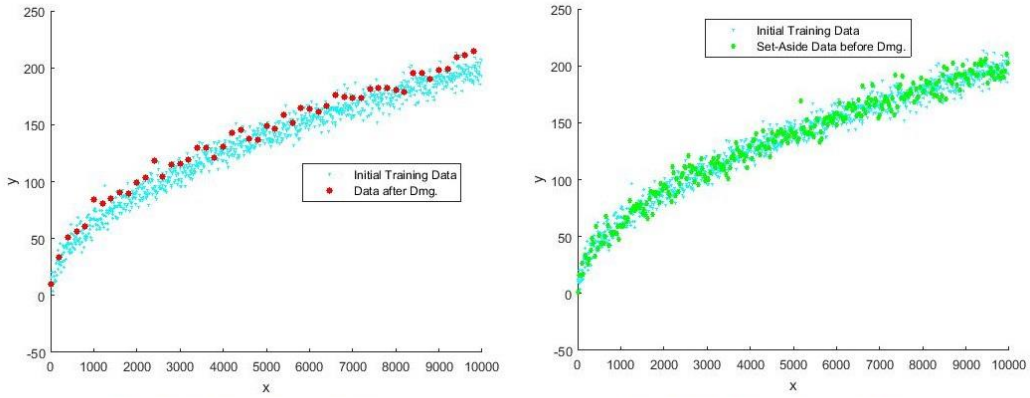

Fig. 19(a). After Imposed Damage **Fig. 19(b).** Before Imposed Damage

Fig. 19. Additional Samples (a) after and (b) before Imposed Damage

If damage exists, then the newly sampled data should deviate further from the final ANN model than does sampled data taken prior to the imposition of damage. The newly sampled data, however, cannot be compared to the data shown in Figure 16, which has been used to develop the final ANN model. Data used to develop the final ANN model will inherently deviate less from the model than will any newly sampled data, regardless of the presence of damage. Therefore, 300 additional samples are taken before the imposition of damage, shown in Figure 19(b). These 300 additional samples have not been used to develop the final ANN model. This sampling is represented by the input-output equation without the damage term of 10, where x is uniformly spaced between 0 and 10,000.

The 50 data points sampled after the damage imposition are compared to the 300 data points sampled additionally when it is known that the structure is undamaged. This comparison is done with the two-sided approximate Wilcoxon rank-sum test with a significance level (α) of 0.1%, or 0.001. A calculated p value below 0.001 demonstrates a rejection of the null hypothesis H_0 with a 0.1% significance level, and a calculated p value above 0.001 demonstrates a failure to reject H_0 with a 0.1% significance level. The parameter h is calculated to be either 0 or 1. In this example, h would be 1 if H_0 is rejected and would be 0 if there is a failure to reject H_0.

The compared parameter of the two sets of data is the deviation of the data points from the final ANN model. The resultant value of p is 4.3829×10^{-12}, which is much less than the α value of 0.001. This yields an h value of 1, which implies

a rejection of the null hypothesis H_0 with a significance level of 0.1%. Therefore, it is concluded with 99.9% certainty that the structure is damaged during the sampling of the 50 new data points.

Chapter 6. Temperature Sensitivity Analysis

6.1 Effects of Temperature

Moser and Moaveni (2011) and Moaveni and Behmanesh (2012) determined that the natural frequencies of a footbridge depend on the temperature. Clinton et al. (2006) observed the effects of weather – such as wind, rain, and temperature – on the fundamental frequency of a building. Alampalli (2000) showed that modal properties of a highway bridge are affected by operational conditions. The bridge behavior framework defined in this study assesses the relationship between measured strains at different locations. It does not consider environmental effects on the characteristics of the structure and how such potential changes in characteristics could affect the relationship between measured strains.

The consideration of environmental effects would manifest itself in this study in two ways. Firstly, to predict the bridge behavior of the PMB more precisely, ANN models trained to mimic the bridge behavior of the PMB would need to include environmental effects as inputs in the measured input-peak strain output relationships defined earlier in this work. Secondly, to simulate the effects of damage more accurately, the damage effects which are simulated in the FEM would need to be determined for many different environmental conditions and then applied to operational truck events based on the corresponding operational conditions during the events. However, that would require accurate prior knowledge of environmental effects on the structure to properly simulate the environmental effects on the FEM structure in the first place. Therefore, only the first consideration is assessed in the rest of this chapter.

6.2 Sensitivity Analysis

It is assumed that the two factors that affect the bridge behavior of the PMB most are (a) temperature and (b) whether the ground supporting the piers and abutments are frozen, which is indirectly a function of temperature. To assess the effect of environmental conditions on this study, a comparison is made between the performances of two developed bridge behavior models. One model is trained with the same measured input/peak strain output relationships as shown in Figure 3 earlier in this work. The other model includes the temperatures of Girders 1 through 6 and the ambient temperature as measured inputs to supplement the inputs shown in Figure 3. Both models are structured and developed the same way as the utilized bridge behavior model is in this study, except that the model that includes temperature has seven additional input terms.

Of the 1,929 accepted truck events used in this study, only 502 have available temperature data and thus are the only events used in this sensitivity analysis. 50 of the 502 events are randomly set aside from model training to assess the performances of the two bridge behavior models. The same randomly selected events are set aside for both models. Each individual ANN in the bootstrapping process trains with a random 85% of the 452 events remaining for model training. A histogram of the ambient temperature during all the events used in this sensitivity analysis, shown below, demonstrates that there is enough variety in temperature for potential environmental effects on model performance to be observed.

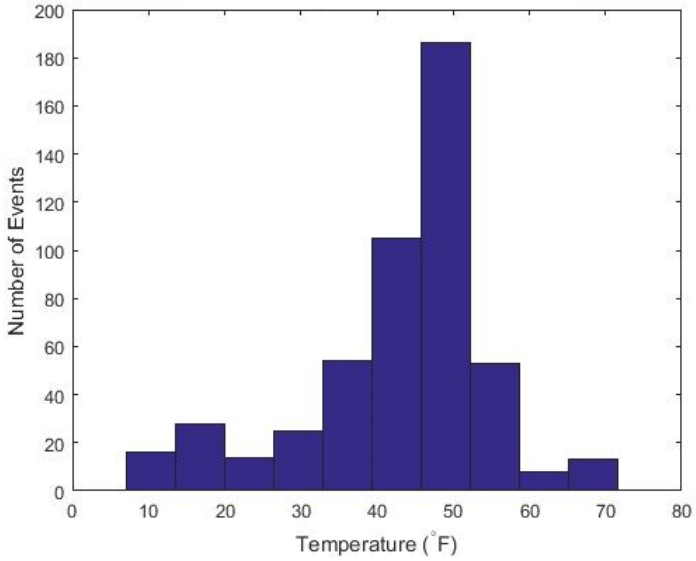

Fig. 20. Ambient Temperatures of Truck Events in Sensitivity Analysis

The performances of the two bridge behavior models are assessed with a residuals analysis. The model predictions of the peak strains are compared to the actual peak strains at all sensor locations for the 50 truck events set aside from model training.

The residuals analysis for the 50 set-aside truck events is summarized in Table 4. The three columns under *Mean Error Mag. (με)* display the mean of the prediction error magnitudes of the bridge behavior models not accounting for temperature and accounting for temperature, as well as the differences between the two. The three columns under *Error Std. Dev. (με)* display the error standard deviations of the bridge behavior models not accounting for temperature and accounting for temperature, as well as the differences between the two. When assessed at all sensor locations, the difference in performance between the two bridge behavior models is negligible. Therefore, environmental effects are not

accounted for in this study. Any actual environmental effects that exist manifest themselves in this study by (a) creating a spread in data used to train the final ANN models of bridge behavior, (b) inducing more prediction error of the final ANN models when they are assessed with set-aside truck events, and (c) creating a spread in how accurately the damage effects modeled in the FEM represent the actual damage effects that should be applied to each individual truck event.

Table 4. Performances of Bridge Behavior Models with and without Temperature

Sensor Location	Mean Error Mag. (με)			Error Std. Dev. (με)		
	W/O Temp.	W/ Temp.	(W/)-(W/O)	W/O Temp.	W/ Temp.	(W/)-(W/O)
G1-S2	0.226	0.210	-0.016	0.258	0.251	-0.006
G1-S6	0.801	0.676	-0.125	1.070	0.924	-0.145
G1-S8	0.265	0.228	-0.037	0.249	0.267	0.019
G1-S10	0.375	0.313	-0.062	0.454	0.397	-0.057
G2-S2	0.424	0.353	-0.071	0.481	0.413	-0.068
G2-S4	0.601	0.452	-0.149	0.592	0.405	-0.187
G2-S6	0.721	0.727	0.006	1.011	0.989	-0.022
G2-S8	0.504	0.425	-0.078	0.436	0.419	-0.017
G2-S10	0.363	0.351	-0.012	0.440	0.452	0.012
G3-S2	0.913	0.680	-0.233	1.129	0.838	-0.291
G3-S4	0.358	0.273	-0.085	0.420	0.347	-0.073
G3-S6	0.475	0.505	0.031	0.620	0.682	0.062
G3-S8	0.325	0.323	-0.002	0.446	0.430	-0.016
G3-S10	0.339	0.329	-0.010	0.387	0.380	-0.007
G4-S2	0.188	0.207	0.019	0.287	0.272	-0.015
G4-S4	0.192	0.165	-0.028	0.225	0.206	-0.019
G4-S6	0.259	0.281	0.022	0.357	0.396	0.039
G4-S8	0.169	0.173	0.004	0.226	0.231	0.005
G4-S10	0.223	0.268	0.045	0.268	0.338	0.071
G5-S2	0.074	0.082	0.008	0.092	0.107	0.015
G5-S6	0.118	0.102	-0.016	0.130	0.132	0.002
G5-S8	0.155	0.159	0.004	0.197	0.205	0.008
G5-S10	0.141	0.148	0.006	0.174	0.175	0.001
G6-S2	0.116	0.100	-0.016	0.148	0.113	-0.035
G6-S4	0.151	0.139	-0.012	0.164	0.161	-0.003
G6-S6	0.229	0.186	-0.043	0.532	0.273	-0.259
G6-S8	7.588	7.593	0.006	7.779	7.787	0.007

6.3 References

Alampalli, S. (2000). "Effects of Testing, Analysis, Damage, and Environment on Modal Parameters." *Mechanical Systems and Signal Processing*, 14(1), 63-74.

Clinton, J.F. et al. (2006). "The Observed Wander of the Natural Frequencies in a Structure." *Bulletin of the Seismological Society of America*, 96(1), 237-257.

Moaveni, B., and Behmanesh, I. (2012). "Effects of changing ambient temperature on finite element model updating of the Dowling Hall Footbridge." *Engineering Structures*, 43, 58-68.

Moser, P., and Moaveni, B. (2011). "Environmental effects on the identified natural frequencies of the Dowling Hall Footbridge." *Mechanical Systems and Signal Processing*, 25(7), 2336-2357.

Chapter 7. Justification and Modeling of Damage Cases

7.1 Case A: Interior Girder Fracture

Damage Case A is an interior girder fracture. It is the first of three damage scenarios that contain simulated damage. The type, severity, and modeling of the damage for this scenario is justified herein (and is also done so for Cases B and D in subsequent sections).

There are documented instances of similar bridge girder fractures, particularly near the midspan of bridges close to steel details. Several cracks were detected on the Hoan Bridge in 2000 on a southern approach span, which were believed to be initiated at constrained weld details on the webs (WisDOT and FHWA 2001). The Delaware I-95 Bridge over the Brandywine River had a brittle crack occur at a stiffener-web connection (Chajes et al. 2005). A crack was found at a lateral gusset plate connection of a girder in the Pennsylvania US-422 Bridge over the Schuykill River when retrofits were being performed (Kaufmann et al. 2004). For two of the three bridges above, girder cracks were not detected until routine inspection or maintenance. Due to the potentially catastrophic nature of girder cracks, it is beneficial to be able to detect them quickly after they develop.

Farrar et al. (1994) analyzed the effects of a girder crack by inducing damage on the I-40 bridge over the Rio Grande River in New Mexico before it was scheduled to be demolished. To induce damage in the bridge, torch cuts of approximately 3/8” wide were incrementally applied starting at midheight of the web, progressing downward through the bottom flange.

In this research, to simulate a realistic girder fracture at the PMB, a 2.5 mm-wide (longitudinally) section of Girder 2 in the FEM was altered to have an elastic modulus close to zero. The modeled girder crack was located at a diaphragm connection, close to the midspan of the girder. The crack is modeled from the midheight of the web through the bottom of the bottom flange. Reiff et al. (2016) implemented a reduction in the elastic modulus to simulate damage – for Case A, as well as Cases B and D (described below). Because Girder 2 has a decrease in stiffness from the crack, it carries less of the load than it does when there is no damage, and other girders carry a larger portion of the load than they do when there is no damage.

7.2 Case B: Fascia Girder Corrosion

Damage Case B is corrosion in the fascia girder. Fascia girder corrosion may be caused when de-icing salts mix with runoff water, which then runs off the side of the bridge onto exterior girders (Enright and Frangopol 2000). The resultant long-term corrosion causes section loss in the affected girder. Miller et al. (2001) tested corroded fascia girders that were removed from a bridge, and determined that their global loss in stiffness ranged from 13% to 32%. They also noted that most of the section loss occurred in the webs and bottom flanges.

Based on their work, fascia girder corrosion is modeled in the calibrated FEM as a 30% reduction of the elastic modulus of the web and bottom flange of Girder 1 along its entire length. Because the flexural stiffness of Girder 1 is decreased, Girder 1 takes less load from a truck when corroded than it normally would when

undamaged, and the rest of the girders carry additional load when Girder 1 is corroded.

7.3 Case D: Southbound Deck Delamination

Damage Case D is southbound deck delamination, in which bond between concrete and reinforcement is lost. Deck delamination is often caused by deck reinforcement corrosion, usually in the top reinforcement layer, which causes the concrete to decouple from the reinforcement bars. This reduces the strength of the structure (Warhus et al. 1995). Deck damage repairs are particularly costly (FHWA 2002), so it would be beneficial to be able to more quantitatively identify deck damage on bridges in order for decision makers to make well-informed decisions regarding asset management.

There are many factors that impact the severity of the effects of delamination on a structure, such as reinforcement ratio since it pertains to lost contributions of debonded bars, for instance (Jnaid and Aboutaha 2014). In general, a deck's contribution to flexural stiffness when undamaged and the severity of its delamination ultimately will affect how much deck delamination reduces the flexural stiffness of part of a bridge. To most effectively work with the configuration and properties of the deck solid elements in the FEM, Reiff et al. (2016) modeled deck delamination as a 35% reduction in the elastic modulus of deck concrete at the PMB, which can be considered an advanced case of delamination. The delaminated deck area in the FEM is 16 m in the longitudinal direction and transversely spans the entire southbound portion of the bridge over which trucks can drive. The delaminated area is centered longitudinally on the

middle span of the model. Because the delaminated portion of the deck has less flexural stiffness, the bridge superstructure is less able to transversely distribute truck loads between the girders over which the delamination is imposed. Additionally, the southbound half of the bridge does not have as much load distributed to it as the northbound half does because it is relatively less stiff.

7.4 References

Chajes, M. et al. (2005). "Steel Girder Fracture on Delaware's I-95 Bridge over the Brandywine River." *Proc., ASCE Structures Congress*, ASCE, New York.

Enright, M.P., and Frangopol, D.M. (2000). "Survey and Evaluation of Damaged Concrete Bridges." *J. Bridge Eng.*, 10.1061/(ASCE)1084-0702(2000)5.

Farrar, C.R. et al. (1994). *Dynamic Characterization and Damage Detection in the I-40 Bridge Over the Rio Grande*, Los Alamos National Laboratory, Los Alamos.

Federal Highway Administration (2002). *Corrosion Costs and Preventative Strategies in the United States*, Federal Highway Administration.

Jnaid, F., and Aboutaha, R.S. (2014). "Residual Flexural Strength of RC Beams with Unbonded Reinforcement." *Structural Journal*, 111(6), 1419-1430.

Kaufmann, E. et al. (2004). *Failure Analysis of the US 422 Girder Fracture*, ATLSS Reports, Bethlehem.

Miller, T. et al. (2001). "Strengthening of a Steel Bridge Girder Using CFRP Plates." *J. Bridge Eng.*, 10.1061/(ASCE)1084-0702(2001)6:6(514).

Reiff, A. et al. (2016). "Statistical bridge damage detection using girder distribution factors." *Engineering Structures*, 109, 139-151.

Warhus, J.P. et al. (1995) "Imaging radar for bridge deck inspection." *SPIE*, 2456, 296-305.

Wisconsin Department of Transportation and Federal Highway Administration (2001). *Hoan Bridge Forensic Investigation Failure Analysis Final Report*.

Chapter 8. Uncertainty Introduced by FEM Damage Simulation

8.1 Assumptions and Uncertainty in Damage Simulation

In this research, assumptions are made when damage is simulated, when damage effects are extracted and processed from FEMs, and when the processed damage effects are applied to operational data from the PMB. Firstly, it is assumed that the structural behavior of the calibrated FEM is sufficiently reflective of the behavior of the real bridge. The results of the work of Sanayei et al. (2012) validate this assumption.

Secondly, it is assumed that the damage effects observed in the calibrated FEM with an HS20 truck are sufficiently representative of the damage effects that would be observed during a truck event that meets the requirements laid out in Chapter 3. Note that the HS20 is an idealized design truck with an axle spacing, axle width, and an axle weight distribution that will not exactly reflect actual vehicles crossing the PMB. Because the trained bridge behavior model in this study accurately and effectively predicts the PMB structural behavior during operational truck events that the model has not yet seen (see Section 2.4.4), the variations of the axle spacings, axle widths, and axle weight distributions of the trucks at the PMB did not significantly hinder the performance of the model's prediction capability. Therefore, even though the accepted PMB truck events are not those of an HS20 truck, an HS20 truck is an acceptable representative truck to determine the effects of damage in the calibrated FEM.

In addition, it is assumed that trucks at the PMB stay in the same lane as they cross the bridge. This assumption is made because (a) damage effects are

determined in the calibrated FEM for northbound events and southbound events, in which the truck is centered in the appropriate lane and maintains a linear path parallel to the longitudinal axis of the bridge, and (b) truck events at the PMB are determined to be northbound or southbound – and have damage appropriately applied based on it – by assessing the distribution of strains at sensor locations at only one station along the length of the bridge. This assumption is valid due to the short length of the bridge, the fact that only one lane exists in each direction, and the assumed reasonable driving habits of truck drivers.

However, it is acceptable in this study to use PMB strain data from truck events in which truck transverse positions within the lanes themselves vary, both between different truck events and over the course of each individual event. For instance, it is acceptable for one northbound truck to remain flush to the road centerline for its entirety on the bridge and for another northbound truck to enter onto the bridge flush to the road centerline and exit the bridge centered in the lane. Damage effects as observed during a northbound truck centered in the northbound lane are applied to all northbound PMB truck events, regardless of where the trucks actually are in the northbound lane. The same approach is used for southbound PMB truck events. Even though the damage effects applied to truck events do not perfectly represent what the actual damage effects would be for each individual truck event, the discrepancy is proven to be acceptable with a sensitivity analysis. The sensitivity analysis is composed of multiple tests and is provided below to determine how much the transverse position of trucks within a lane affects the accuracy of the damage effects applied to truck events.

8.2 Lane Variation Sensitivity Analysis

Damage effects in this study are calculated as the percent changes in measured strain input and peak strain output between no-damage and induced-damage FEM truck runs. For northbound events at the PMB, damage effects from the FEM simulations with the northbound HS20 centered in the northbound lane are applied. Likewise, for southbound events at the PMB, damage effects from the southbound FEM simulations are applied. A sensitivity analysis is performed to determine how much the transverse position of trucks within each lane affects the accuracy of the damage effects applied to truck events at the PMB.

One damage scenario was used in the sensitivity analysis: Case A, interior girder fracture. For the sensitivity analysis only, the effects of damage are qualitatively defined as the percent difference in the strain response of the undamaged bridge and damaged bridge during the same exact truck event, over the course of the whole event, at a location. This percent difference in response between a no-damage case and an induced-damage case is the parameter used to assess the similarity of damage effects observed during truck events with trucks in different lane paths. 8 tests are performed in this assessment, using 12 different simulated truck events in the calibrated FEM and two different sensor locations at which the strain responses are observed. The truck paths used are shown in Figure 21, and the FEM truck events run are shown in Table 5.

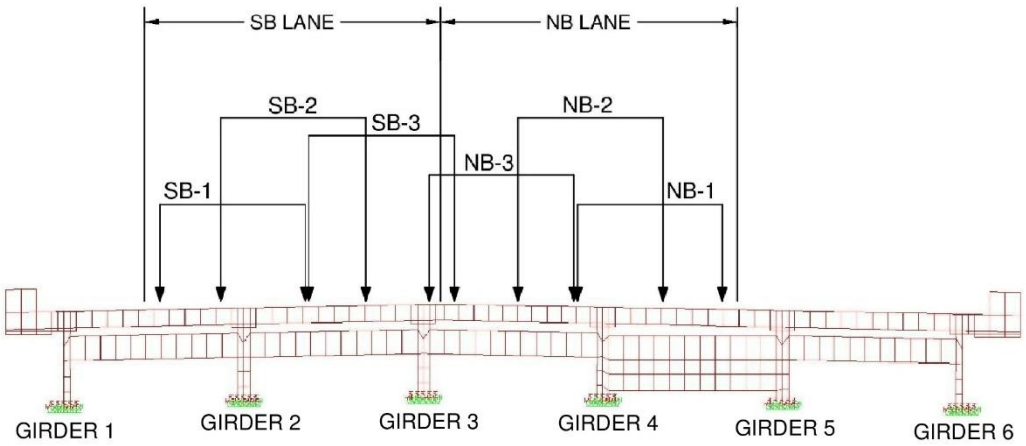

Fig. 21. Lane Paths for Sensitivity Analysis

Table 5. FEM Truck Runs in Sensitivity Analysis

Run No.	Lane Path	Damage
1	NB-1	N
2	NB-1	Y
3	NB-2	N
4	NB-2	Y
5	NB-3	N
6	NB-3	Y
7	SB-1	N
8	SB-1	Y
9	SB-2	N
10	SB-2	Y
11	SB-3	N
12	SB-3	Y

Lane paths NB-2 and SB-2 are centered in the northbound and southbound lanes, respectively. NB-1 and SB-1 are located such that the outer wheels of the HS20 truck are 0.21 m from the outer edge of their respective lanes; and NB-3 and SB-3 are located such that the inner wheels of the truck cross over the road

centerline by 0.16 m. NB-1, SB-1, NB-3, and SB-3 are assumed to be reasonable extents of typical truck events.

The damage effects observed in the sensitivity analysis are determined to be the percent difference in strain response between Run 1 and Run 2, the percent difference in strain response between Run 3 and Run 4, the percent difference in strain response between Run 5 and Run 6, and so on for all runs. With the damage effects defined for each truck lane path, tests are performed to determine how well the damage effects as observed during a truck event with a truck centered in a lane (NB-2 and SB-2) represents the damage effects as observed during a truck event with a truck anywhere in that lane (NB-1, SB-1, NB-3, and SB-3). The tests performed are summarized in Table 6. In the table, Girder 2 Station 6 is the designated sensor location near the damage, and Girder 6 Station 2 is the designated sensor location far from the damage.

Table 6. Sensitivity Analysis Tests

Test No.	Lane Paths of Compared Damage Effects		Sensor Location at which Comparison is Made
1	NB-2	NB-1	Near Dmg.
2	NB-2	NB-3	Near Dmg.
3	NB-2	NB-1	Far from Dmg.
4	NB-2	NB-3	Far from Dmg.
5	SB-2	SB-1	Near Dmg.
6	SB-2	SB-3	Near Dmg.
7	SB-2	SB-1	Far from Dmg.
8	SB-2	SB-3	Far from Dmg.

The results of the sensitivity analysis are explained in detail below for only Test #1, and the rest of the test results are summarized. The FEM simulated truck event

strains at Girder 2 Station 6 are shown in Figure 22 for a truck moving along NB-2 (centered in the northbound lane) and for a truck moving along NB-1 (outer portion of the northbound lane), both with and without the imposed girder fracture. As shown in the figure, the strains experienced at the sensor location decrease when there is damage. This is because the sensor location is on the fractured girder, near the fracture. Because Girder 2 has a loss in flexural stiffness when it becomes damaged, it takes less load under such conditions.

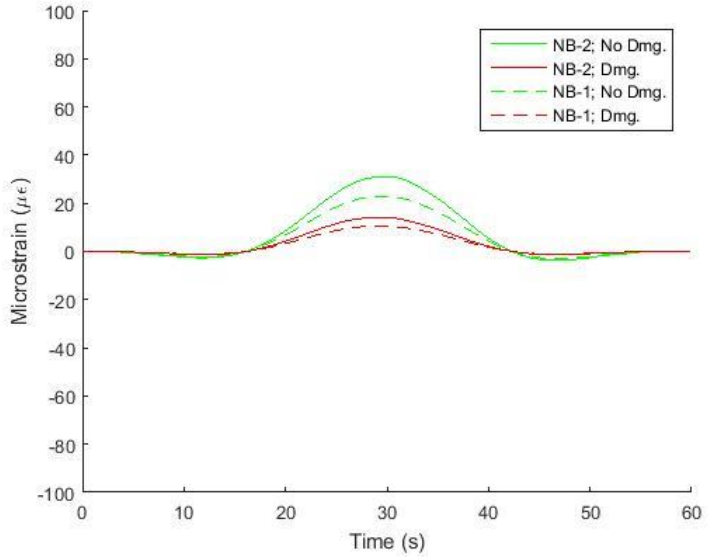

Fig. 22. Strains near Damage for NB-2 and NB-1 with and without Damage

The percent change in response used to represent damage in this sensitivity analysis, shown as Δ/ε in Figure 23, is calculated for the strain response over entire truck events as:

$$\Delta/\varepsilon = \frac{\varepsilon_D - \varepsilon_U}{\varepsilon_U} \tag{6}$$

where ε_D and ε_U are the strain responses with and without damage, respectively. The discrepancy between $[\Delta/\varepsilon]_{NB-2}$ and $[\Delta/\varepsilon]_{NB-1}$, the damage effects for lane paths NB-2 and NB-1, respectively, is shown in Figure 23 as $[\Delta/\varepsilon]_{NB-1} - [\Delta/\varepsilon]_{NB-2}$ and is also calculated as such.

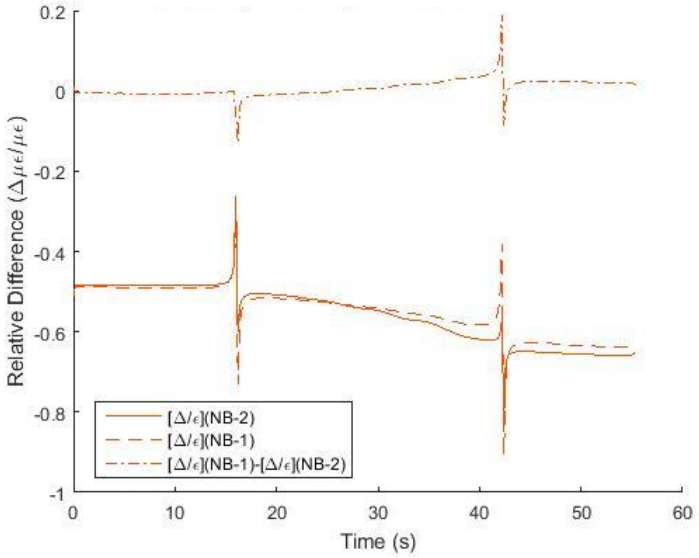

Fig. 23. Damage Effects near Damage for NB-2 and NB-1

It is desired to determine not simply the difference between damage effects for NB-1 and NB-2, but rather the magnitude of that difference with respect to the magnitudes of the damage effects themselves. The effects of damage for NB-1 and NB-2 are similar to one another for most of the duration of truck events. The only points at which the magnitude of the discrepancy of the damage effects is large compared to the magnitude of the damage effects themselves are at approximately 16 seconds and 43 seconds into the truck events. However, during those two times, the magnitude of the strain readings are almost zero (Figure 22). Therefore, if damage effects are ever applied at these times, then the applied effects would still

be negligible for the sensor location in this sensitivity analysis because the strain value to which they are applied is almost zero.

The results for all the other tests summarized in Table 6 are consistent with the results of Test #1. It is therefore concluded that lane path location within lanes does not have a significant impact on damage effects regarding how they are applied in this study. The assumptions stated earlier in this chapter are valid.

8.3 References

Sanayei, M. et al. (2012). "Instrumentation, Nondestructive Testing, and Finite-Element Model Updating for Bridge Evaluation Using Strain Measurements." *J. Bridge Eng.*, 10.1061/(ASCE)BE.1943- 5592.0000228.

Chapter 9. Conclusion

9.1 Summary of Work

The live load responses of the PMB are used as a case study to assess a proposed statistical modeling method using ANNs for the development of structural health monitoring. Bridge behavior is a framework developed in this study with which bridge performance can be identified and learned. A method of long-term performance assessment using the expected bridge behavior, which can be passively evaluated over time, is introduced. Hypothetical damage is used as one way that bridge performance can change, and damage identification is conducted to assess performance and determine whether or not it changed. The bridge behavior model is assessed with truck events set aside from its development. The effects of damage are extracted from FEM truck runs and applied to operational data to assess the damage identification capabilities of the proposed method.

When tested against new data, the ANN-learned model of bridge behavior is proved to be effective and applicable to a variety of traffic events with unknown loading conditions. Damage is also detected efficiently when an appropriate significance level is used for Wilcoxon rank-sum testing. Damage is localized for two out of the three generated scenarios in which damage is imposed on the bridge.

Several sensitivity analyses and literature surveys are performed to justify components of the work performed in this thesis. A new method of categorizing truck events and filtering out undesirable truck events is presented. A brief overview of ANNs as they pertain to this study is provided.

9.2 Concluding Remarks

This non-mechanistic framework of bridge behavior is powerful because it assesses performance with no knowledge of the traffic that induces the strains used in this study; it relates strains measured over the whole bridge during a variety of truck events of unknown nature. In addition, it is sensitive to the effects of damage because how interconnected the measured strains are in the ANN models trained to relate them to one another. Additional work is still required to make the methods in this study more effective at defining bridge behavior and identifying damage. For instance, new temperature input/strain output configurations can be assessed, and a method to determine *a priori* an appropriate Wilcoxon rank-sum test significance level can be developed.

However, the preliminary results of this study demonstrate that future adaptations of this work could be promising. This global approach to bridge behavior can potentially be applied any type of bridge, with adjustments made to the types of data used as inputs and outputs of the ANN models. This framework for bridge behavior can be useful to the long-term objective of improving infrastructure management.

9.3 Future Work

One issue that can be addressed in future work is the *a priori* determination of an appropriate significance level for damage detection for any bridge like the PMB. This can be potentially based on many factors, such as the characteristics of the

bridge structure, its instrumentation, and the nature of the traffic that the bridge carries.

Additionally, future work can determine how well different input/output relationships of the bridge behavior model are able to capture and convey changes in bridge performance. For example, the strains of multiple sensor locations can be clustered together to create input/output relationships that have strains in one region of the bridge act as predictors of strains in another region of the bridge. The way sensors can be clustered may end up depending on the bridge performance requirements established by bridge owners and carried out in design by engineers. Similarly, there is potential to assess other input/output relationships, such as temperature inputs and strain outputs at all sensor locations when no trucks are on the bridge. For said example, only one final ANN model would be needed to define the bridge behavior (rather than 27) because all the temperatures across the bridge can be used as direct predictors of all the strains across the bridge. In addition to temperature data, accelerometer and tiltmeter data at the PMB can potentially be incorporated into the input/output relationships.

Also regarding ANNs, this research is intended as a proof of concept using an ANN design that is considered to most likely be effective according to literature. A sensitivity analysis can be performed to determine optimal ANN architectures and optimization algorithms, depending on what data the researcher uses to define bridge behavior.

Further into the future, the methods in this study can be used to characterize damage on bridges similar to the PMB. The PMB is a small stringer bridge, which

is extremely common, so there is a lot of potential data to utilize and many bridges to which the findings can be applied. If bridges comparable to the PMB all have similar sensor instrumentations, the ways that damages manifest themselves in the different bridges' behavior models can be studied on a broad scale. If conclusions could be drawn about how certain types of damage become apparent in bridge behavior models, then engineers responsible for bridges similar to the PMB can know what types of changes in bridge performance to look for when they are monitoring conditions in the long-term.

Although the methods used in this research only apply to bridges that experience enough isolated live loads caused by a single heavy vehicle, the concepts introduced can be applied to bridges with more continuous traffic loads. For example, instead of defining bridge behavior as the interrelationships between instantaneous strain measurements, bridge behavior for large bridges with high-volume traffic can be defined as the interrelationships between sensor locations' statistical distributions of measurements recorded throughout a period of time.

Appendix

A.1 Malfunctioning Strain Gage Detection and Truck Event Removal

Some of the 50 bottom flange strain gages are observed as malfunctioning during recorded truck events. To address this issue, a two-step process is implemented. First, gages that malfunction for more than 30 accepted truck events (as described in Section 2.3) are removed from the study. The remaining gages are known as *usable gages*. Then, any accepted truck events during which a usable gage malfunctions is removed from this study. This two-step process ensures that enough sensors are used to define a sufficient bridge behavior model, and that enough accepted truck events are available to adequately train and assess the ANN models.

First, strain gages are flagged as malfunctioning during an event if they have flat readings. A gage is flagged for this reason if the absolute difference between its maximum and minimum readings over the entire truck event is less than 0.05 microstrains. The strains of various gages at Station 6 are plotted in Figure 24 to visualize what the readings of strain gages with this malfunction are relative to what those of functioning strain gages are. Then, strain gages are flagged as malfunctioning during an event if they have sudden jumps in readings. A gage is flagged for this reason if it has a change in reading of at least 15 microstrains within $1/25^{th}$ of a second throughout the entire event. There are three typical ways this occurs. The strains of various gages at Station 6 are plotted in Figures 25 through 27 to visualize what the readings of strain gages with this malfunction are relative to what those of functioning strain gages are.

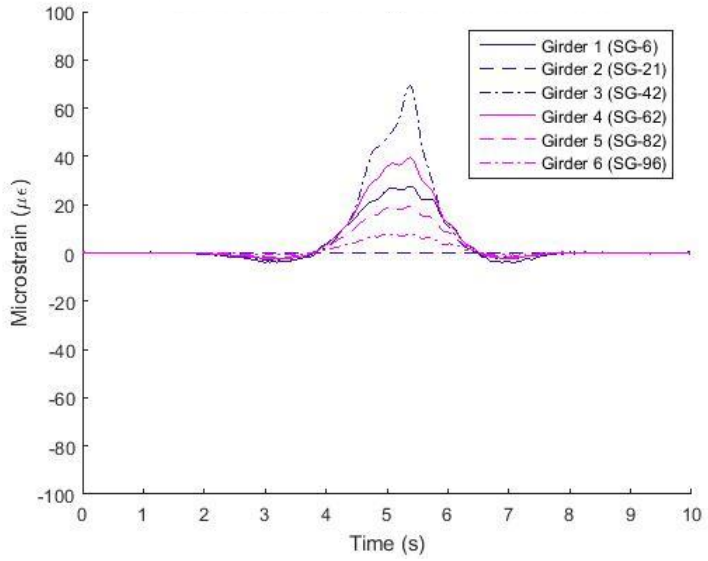

Fig. 24. Station 6 Strains for Typical Event with Flat Reading

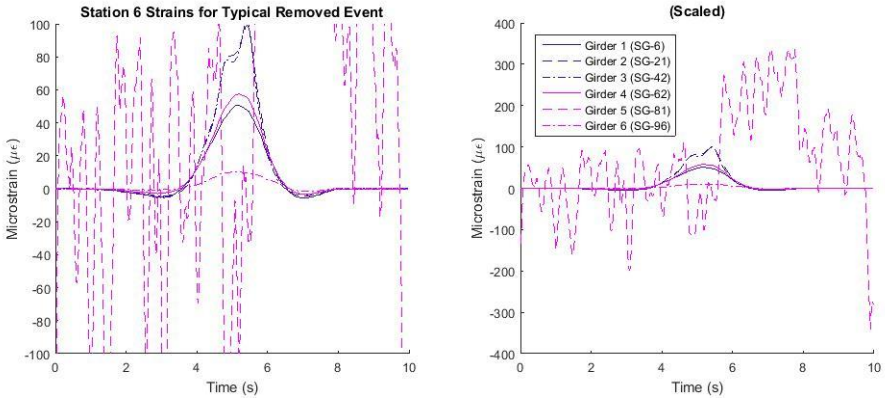

Fig. 25. Station 6 Strains for Typical Event with Jump in Data (1)

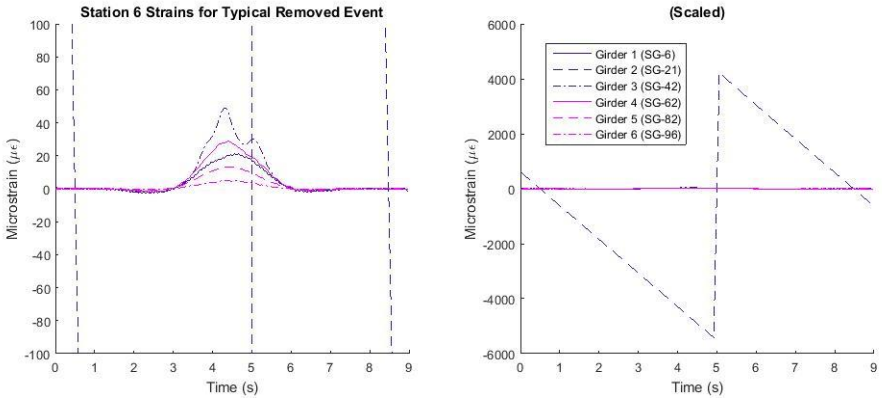

Fig. 26. Station 6 Strains for Typical Event with Jump in Data (2)

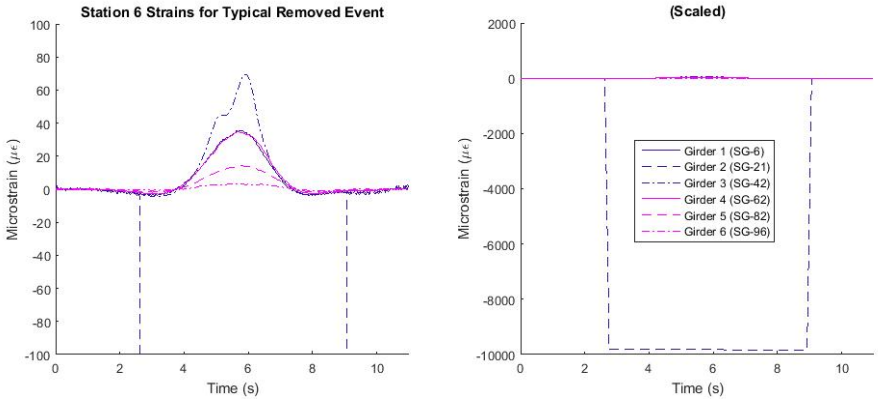

Fig. 27. Station 6 Strains for Typical Event with Jump in Data (3)

The strain gages entirely removed from the study are SG-4 (Girder 1 Station 4), SG-30 (Girder 2 Station 10), SG-33 (Girder 3 Station 2), SG-50 (Girder 3 Station 10), SG-77, (Girder 5 Station 4), SG-78 (Girder 5 Station 4), SG-81 (Girder 5 Station 6), and SG-100 (Girder 6 Station 10).

A.2 All Girder GDF Quantile Stability Plots

Figure 9 earlier in this work is used as an example to convey that there is a sufficient number of truck events to properly use the bootstrapping scheme to develop a probabilistic model of bridge behavior. The Girder 3 GDF quantile stability plot in the figure must also be supplemented with the quantile stability plots of all GDFs to completely convey that enough truck events are used.

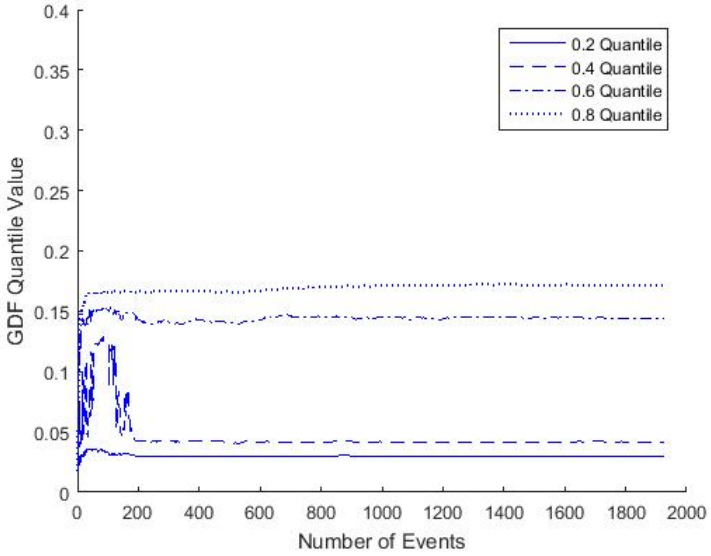

Fig. 28. Girder 1 GDF Quantile Stability Plot

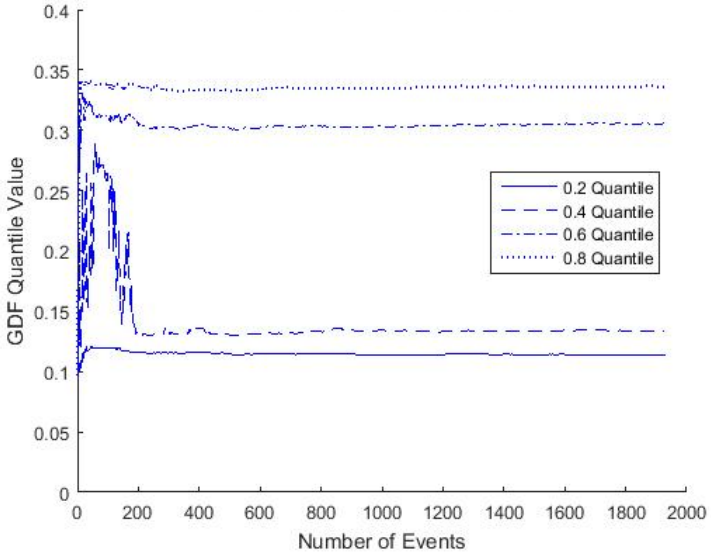

Fig. 29. Girder 2 GDF Quantile Stability Plot

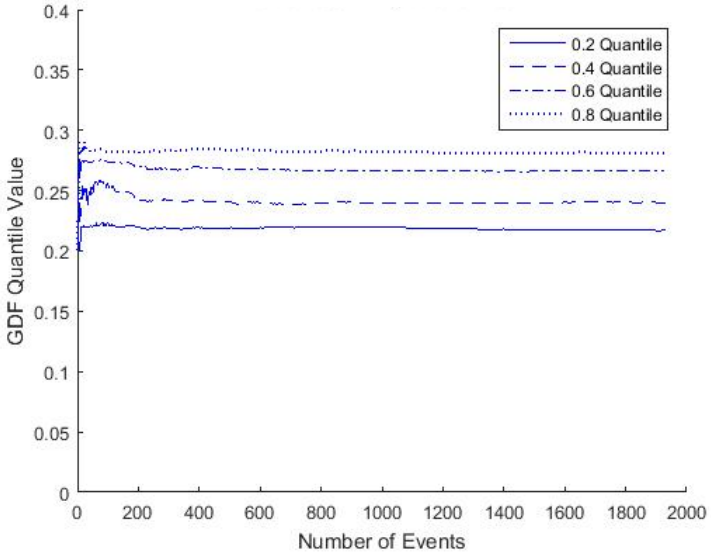

Fig. 30. Girder 3 GDF Quantile Stability Plot

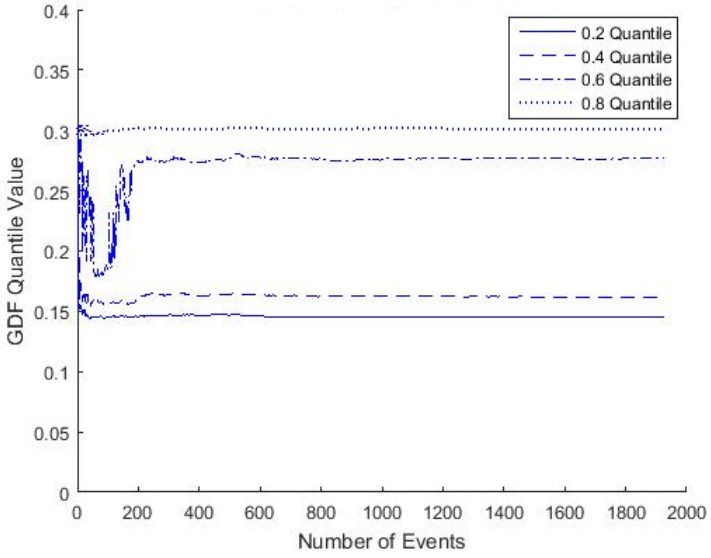

Fig. 31. Girder 2 GDF Quantile Stability Plot

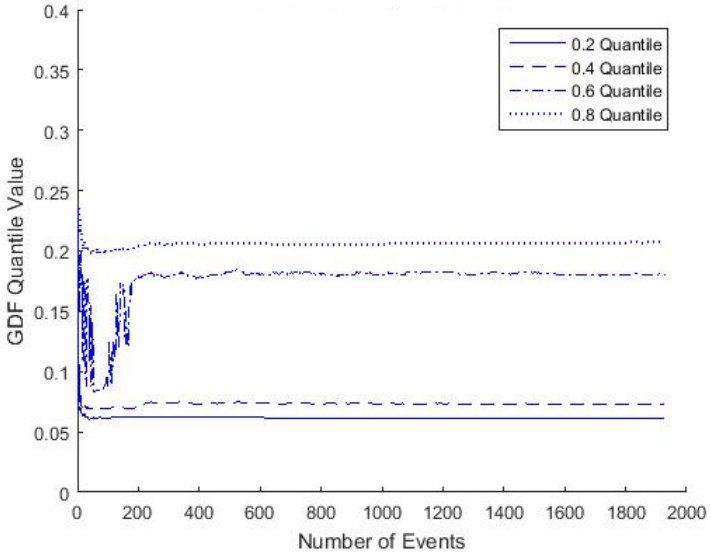

Fig. 32. Girder 5 GDF Quantile Stability Plot

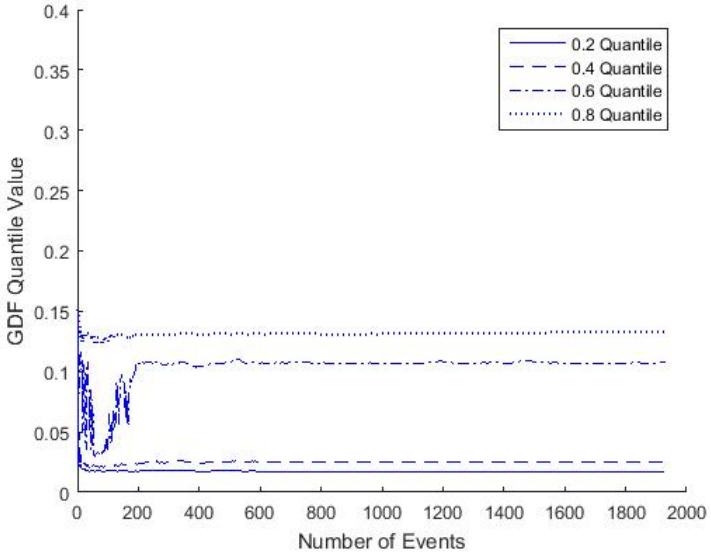

Fig. 33. Girder 6 GDF Quantile Stability Plot

A.3 Confidence Interval Formulation for All Output Sensor Locations

Figure 10 earlier in this work demonstrates how a confidence interval of the peak strain prediction of one sensor location is determined. Below are the training error distributions, like that of Figure 10, for the 1,000 individual ANN models of

all 27 ANN bridge behavior models. The 2.5th and 97.5th percentiles of the following distributions, shown by the black bars in the histograms, determine the confidence interval range of any peak strain prediction at the corresponding sensor location.

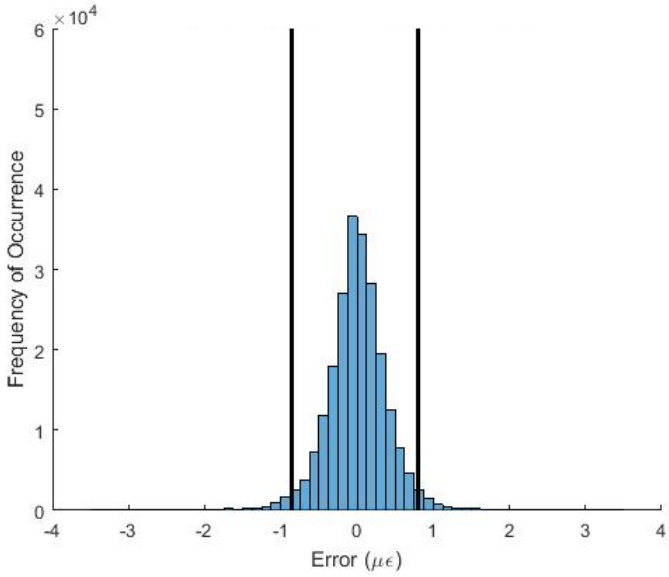

Fig. 34. Pooled Prediction Error Distribution at Girder 1 Station 1

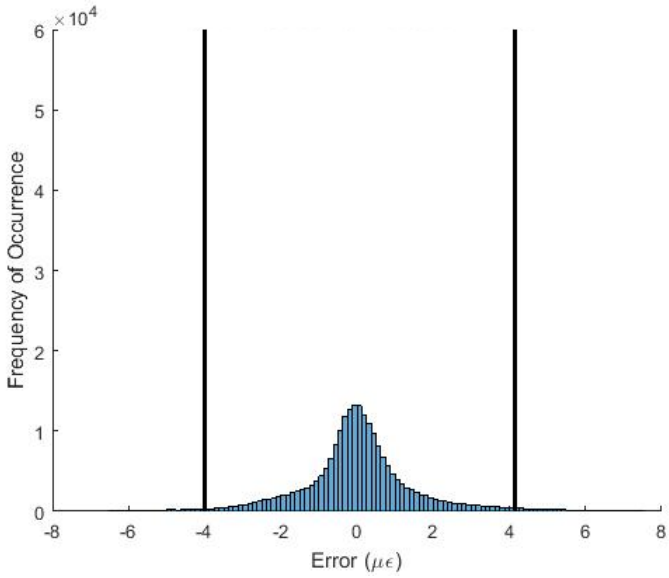

Fig. 35. Pooled Prediction Error Distribution at Girder 1 Station 6

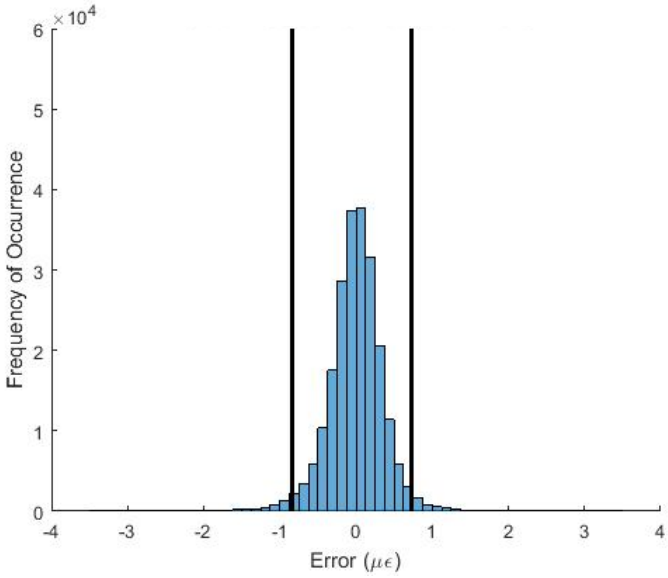

Fig. 36. Pooled Prediction Error Distribution at Girder 1 Station 8

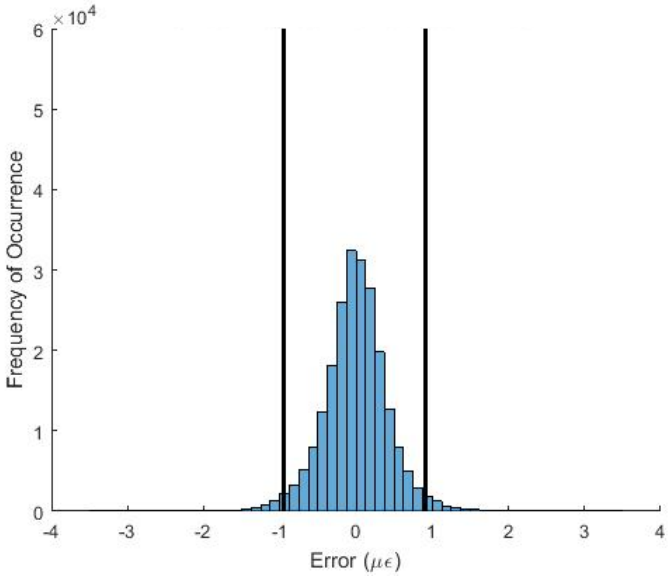

Fig. 37. Pooled Prediction Error Distribution at Girder 1 Station 10

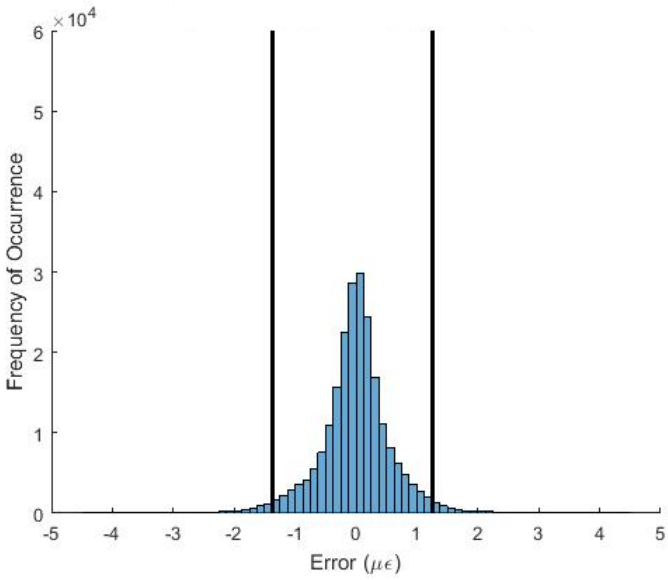

Fig. 38. Pooled Prediction Error Distribution at Girder 2 Station 2

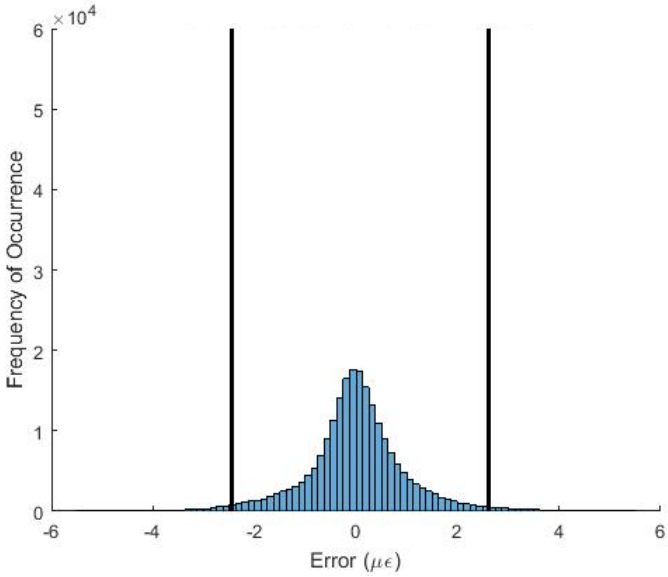

Fig. 39. Pooled Prediction Error Distribution at Girder 2 Station 6

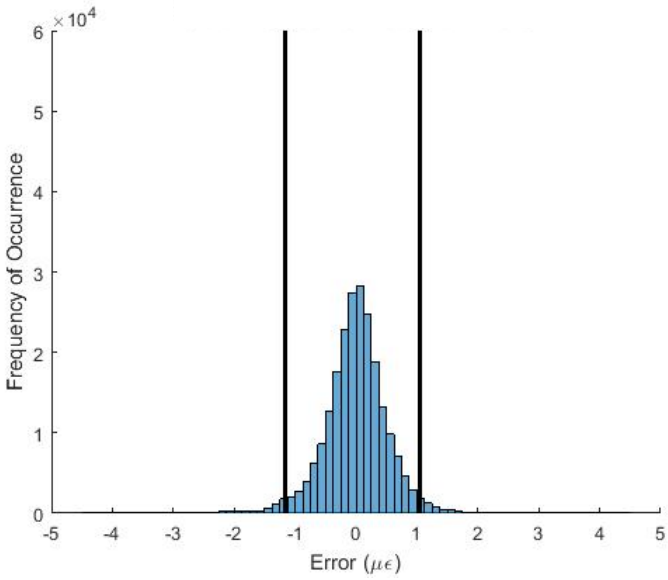

Fig. 40. Pooled Prediction Error Distribution at Girder 2 Station 8

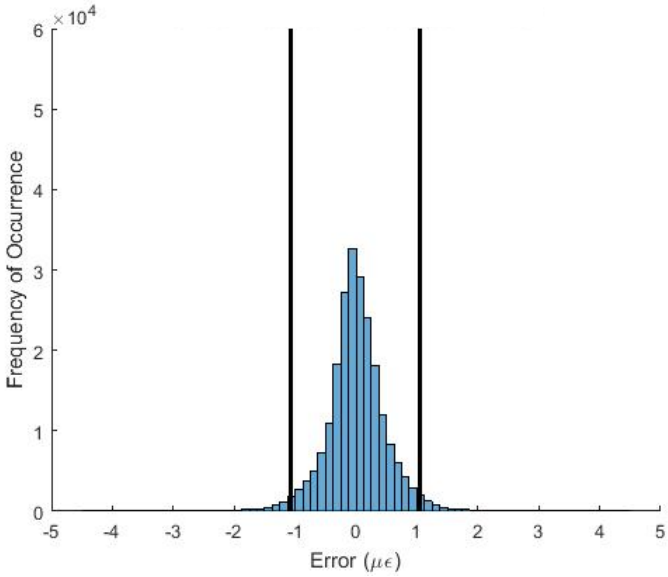

Fig. 41. Pooled Prediction Error Distribution at Girder 2 Station 10

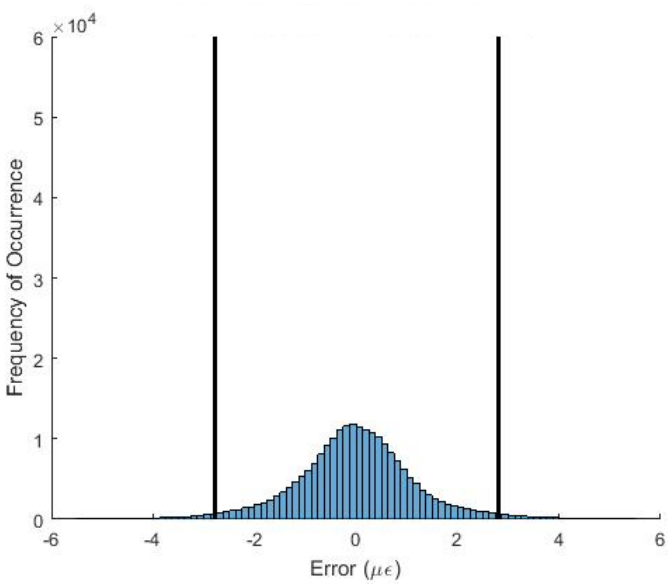

Fig. 42. Pooled Prediction Error Distribution at Girder 3 Station 2

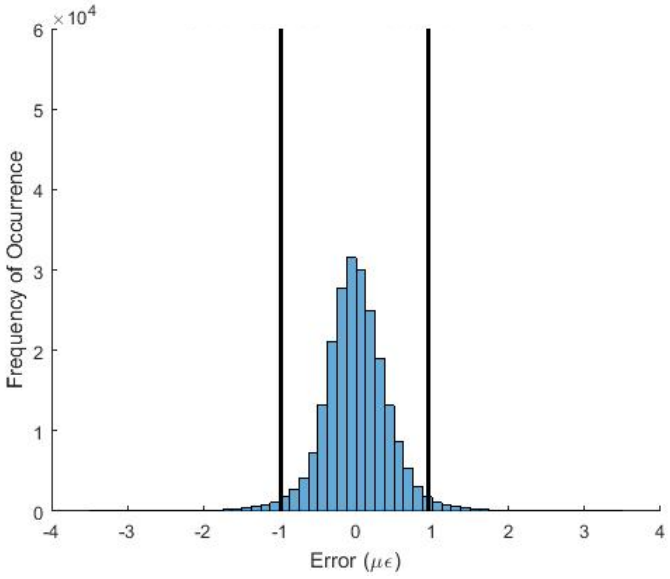

Fig. 43. Pooled Prediction Error Distribution at Girder 3 Station 4

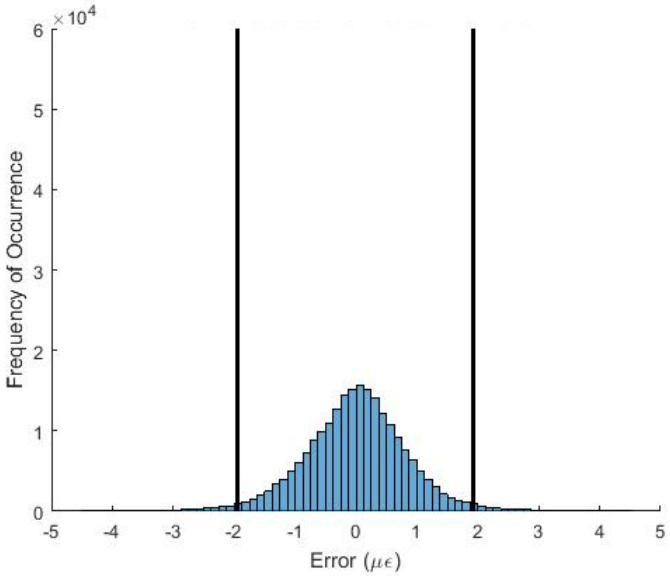

Fig. 44. Pooled Prediction Error Distribution at Girder 3 Station

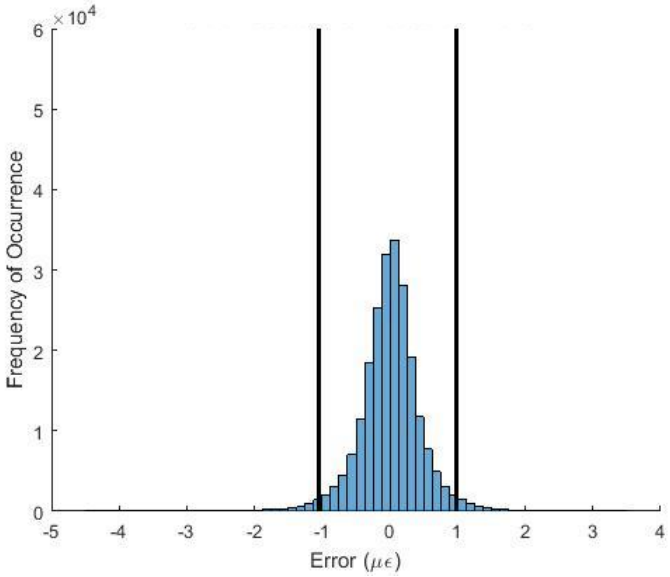

Fig. 45. Pooled Prediction Error Distribution at Girder 3 Station 8

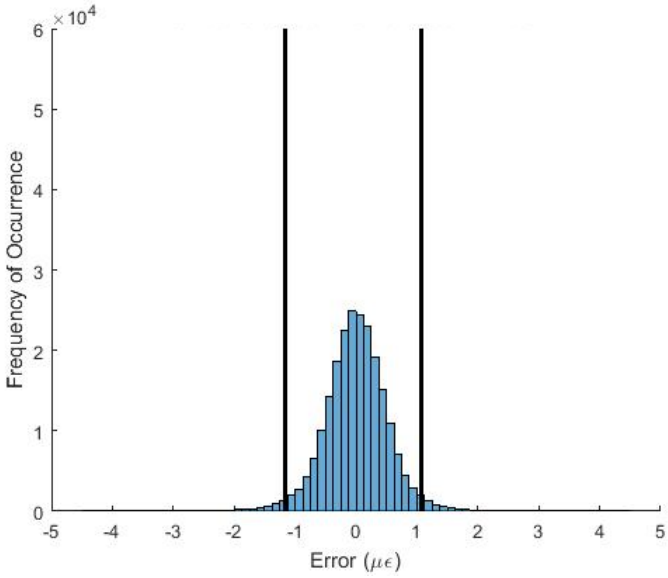

Fig. 46. Pooled Prediction Error Distribution at Girder 3 Station 10

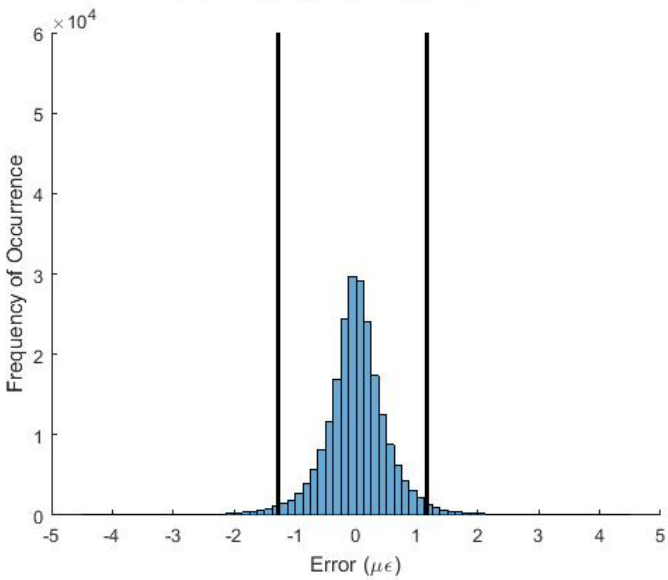

Fig. 47. Pooled Prediction Error Distribution at Girder 4 Station 2

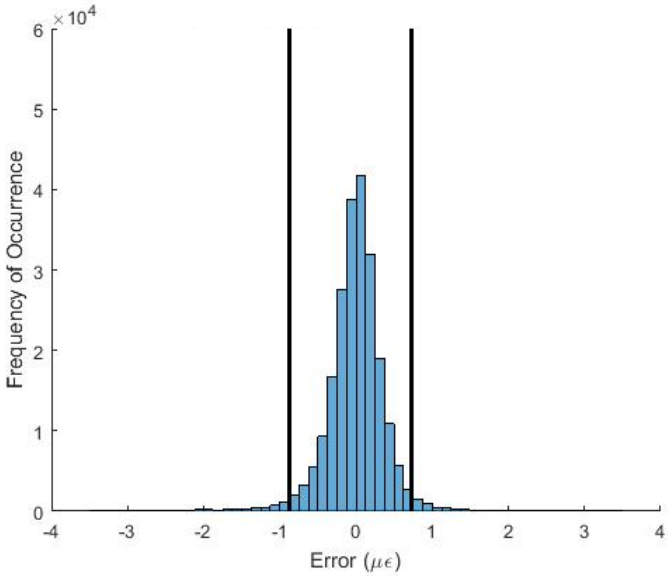

Fig. 48. Pooled Prediction Error Distribution at Girder 4 Station 4

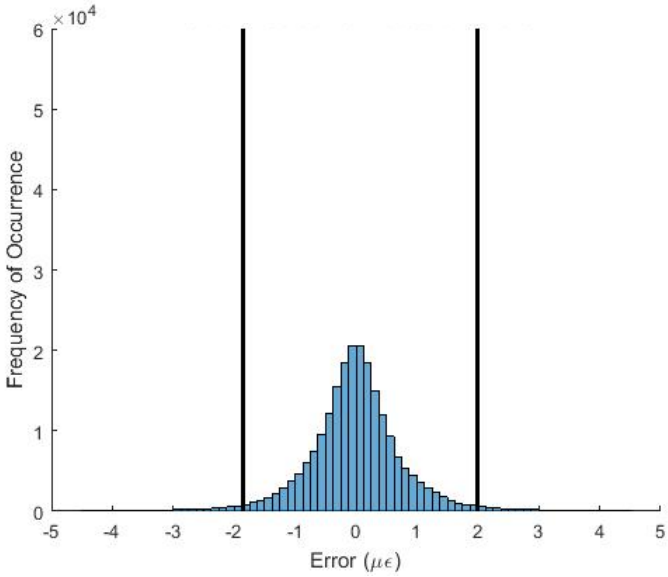

Fig. 49. Pooled Prediction Error Distribution at Girder 4 Station 6

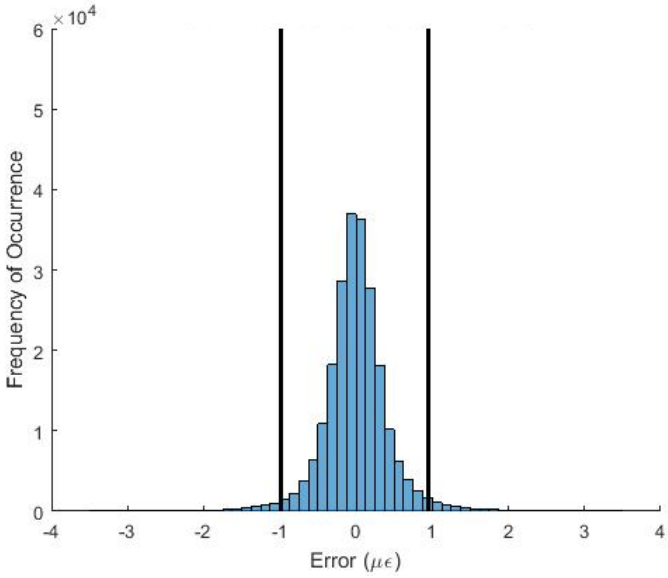

Fig. 50. Pooled Prediction Error Distribution at Girder 4 Station 8

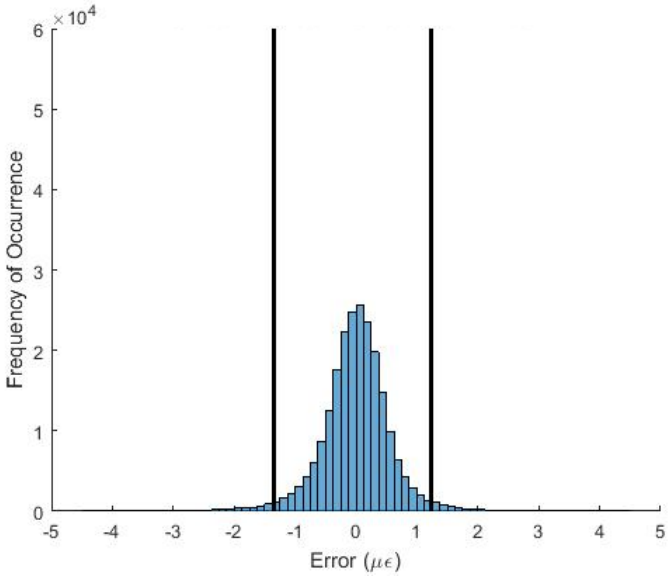

Fig. 51. Pooled Prediction Error Distribution at Girder 4 Station 10

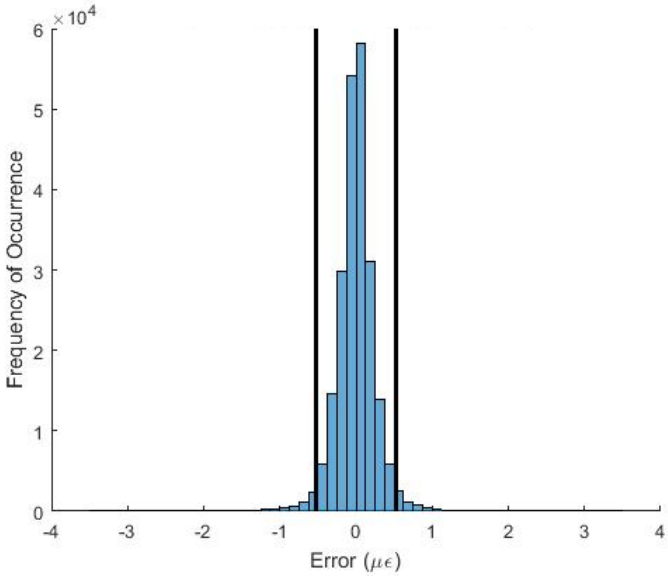

Fig. 52. Pooled Prediction Error Distribution at Girder 5 Station 2

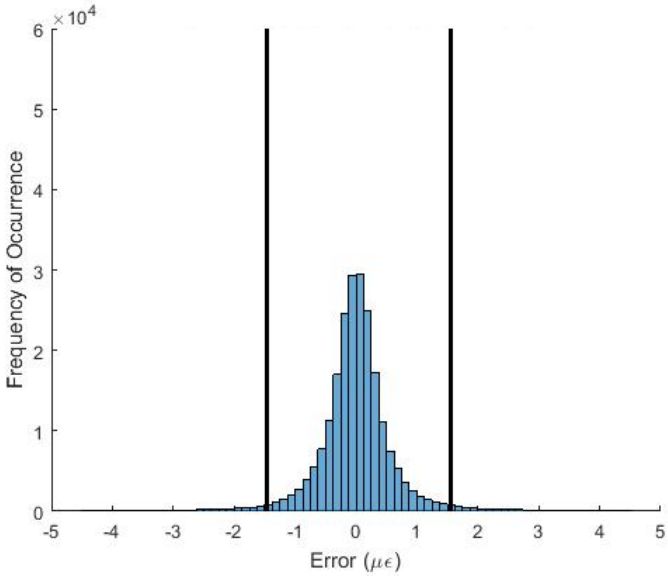

Fig. 53. Pooled Prediction Error Distribution at Girder 5 Station 6

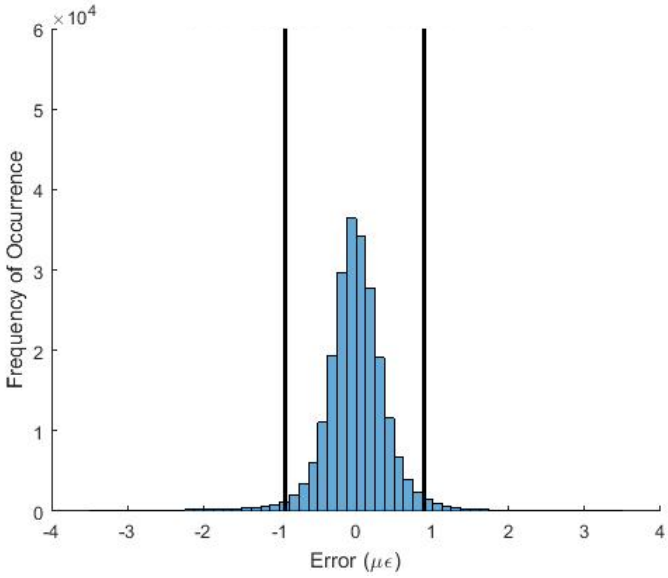

Fig. 54. Pooled Prediction Error Distribution at Girder 5 Station 8

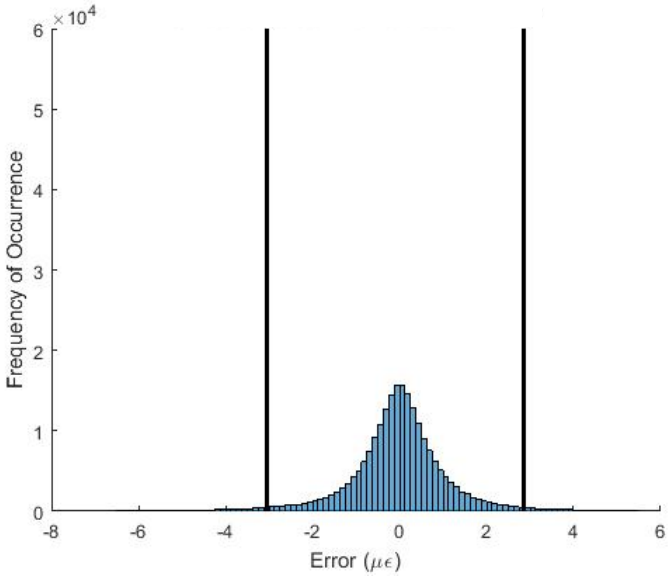

Fig. 55. Pooled Prediction Error Distribution at Girder 5 Station 10

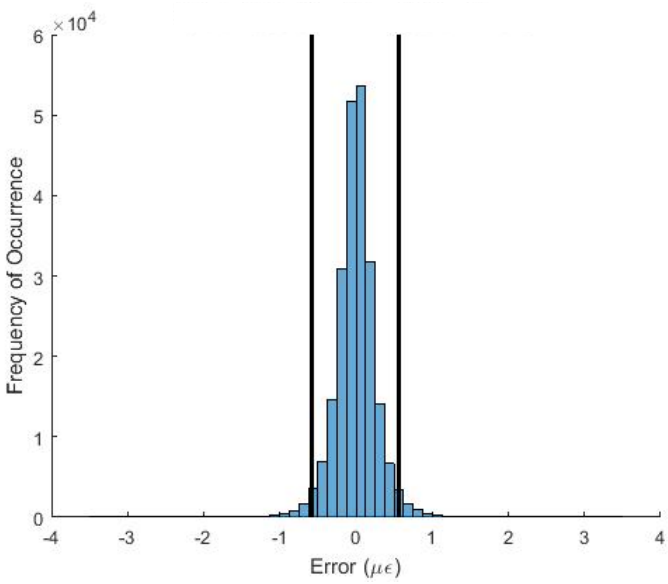

Fig. 56. Pooled Prediction Error Distribution at Girder 6 Station 2

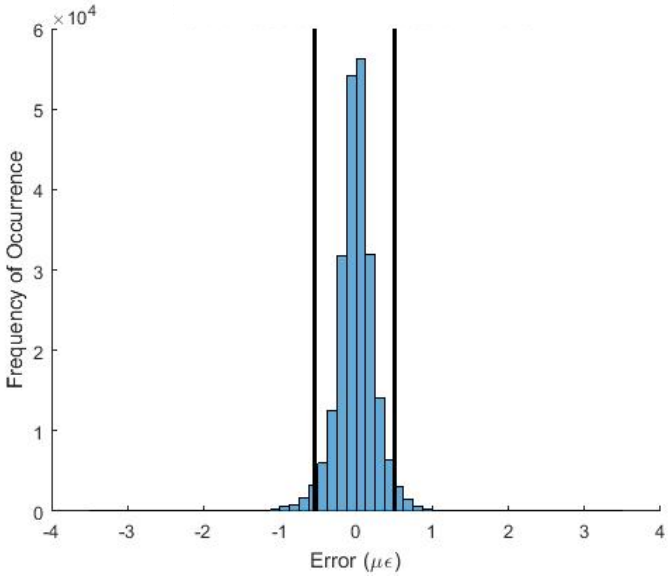

Fig. 57. Pooled Prediction Error Distribution at Girder 6 Station 4

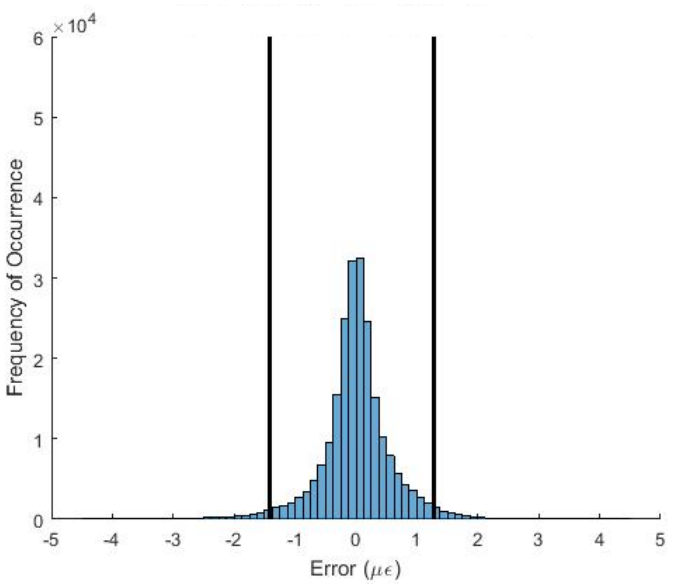

Fig. 58. Pooled Prediction Error Distribution at Girder 6 Station 6

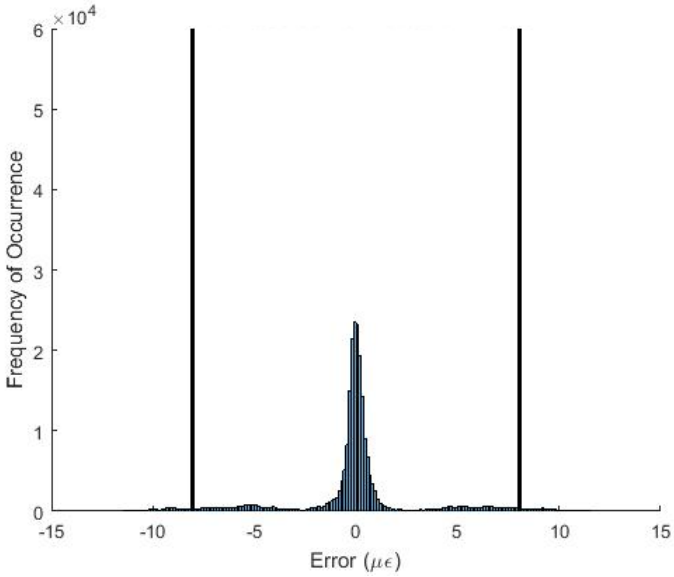

Fig. 59. Pooled Prediction Error Distribution at Girder 6 Station 8

A.4 Strain Extraction from Finite Element Models

The calibrated FEM in CSiBridge contains plane stress elements for the steel girders of the bridge. Stresses are extracted at the element nodes closest to each strain gage location. Stresses are extracted at such nodes at the top of the elements,

because the strain gages were installed on the tops of the bottom flanges of the girders. The nodal stresses of all four elements that share such nodes are averaged. Longitudinal strain, which is also the strain measured by the gages, is calculated as follows:

$$\varepsilon = \frac{\sigma_x}{E} + \upsilon \frac{\sigma_y}{E} \tag{7}$$

where σ_x and σ_y are the longitudinal and transverse stresses, respectively. The elements from which the nodal stresses are extracted are plane stress, so there are no stresses in the z direction that contribute to the longitudinal strain.

A.5 Lane Variation Sensitivity Analysis Results for All Comparisons

The observations made from Figures 22 and 23 partly validate the assumption that the transverse location of a truck within a lane does not significantly impact the accuracy of the damage effects applied to the data during that truck event. The observations made in the figures are for only one of the tests performed in the sensitivity analysis. The same types of figures are shown below for all the eight sensitivity analysis tests summarized in Table 6. The observations made from Figures 22 and 23 are also made for the following figures, which wholly validates the assumption that damage effects are not sensitive to lane path variation within a lane.

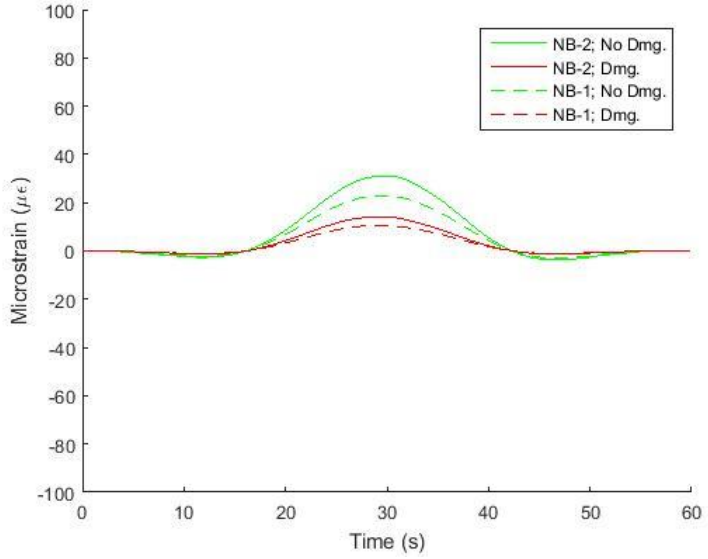

Fig. 60. Strains near Damage for NB-2 and NB-1 with and without Damage

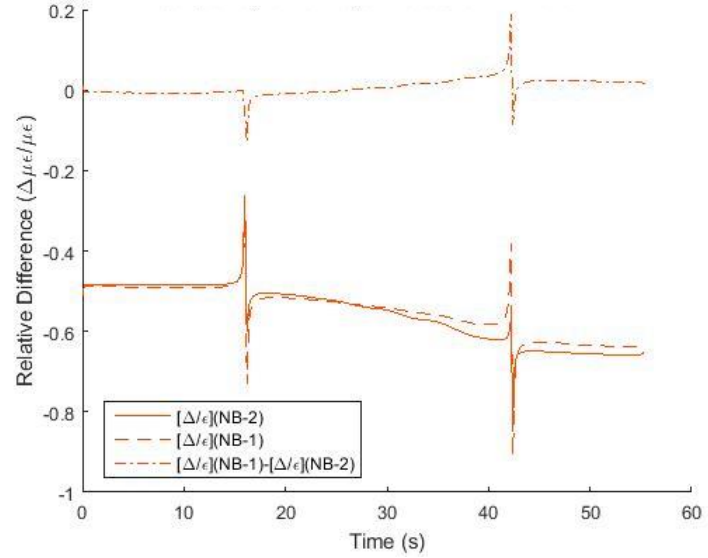

Fig. 61. Damage Effects near Damage for NB-2 and NB-1

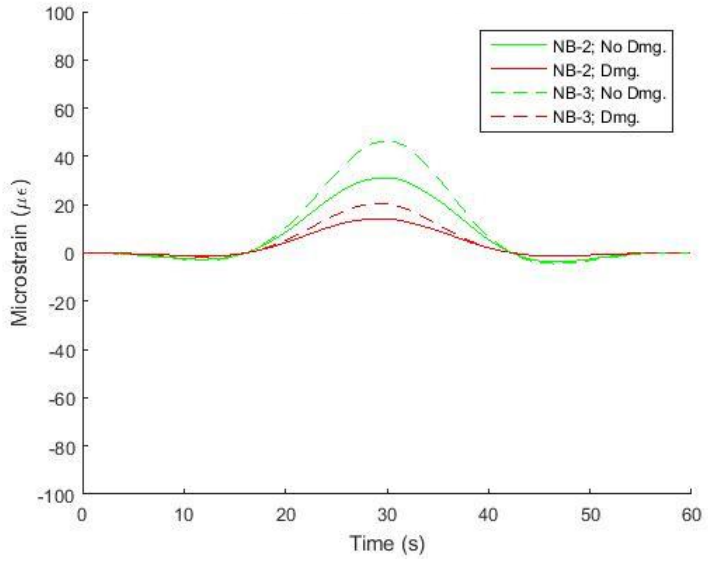

Fig. 62. Strains near Damage for NB-2 and NB-3 with and without Damage

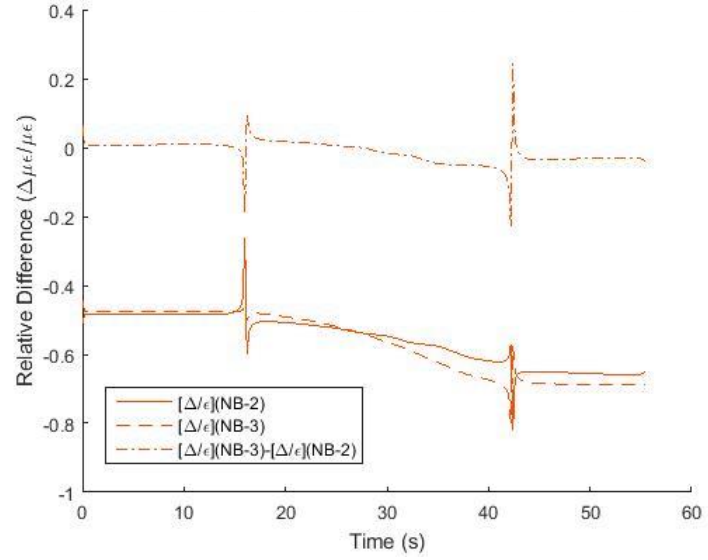

Fig. 63. Damage Effects near Damage for NB-2 and NB-3

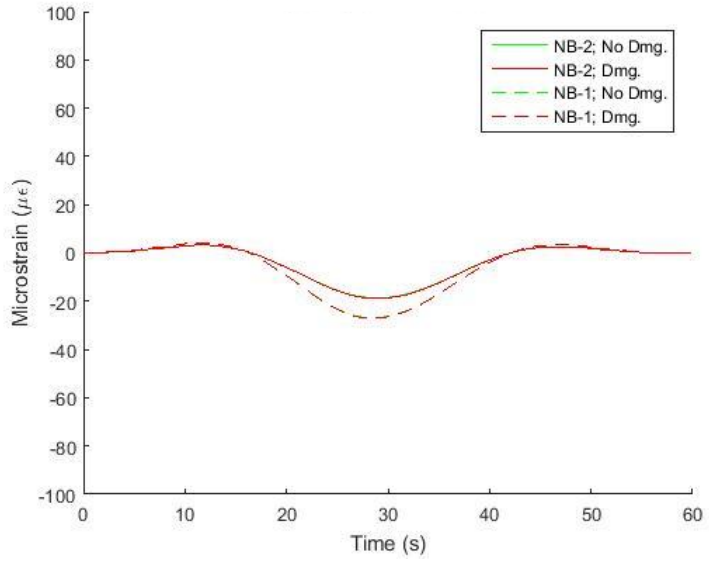

Fig. 64. Strains far from Damage for NB-2 and NB-1 with and without Damage

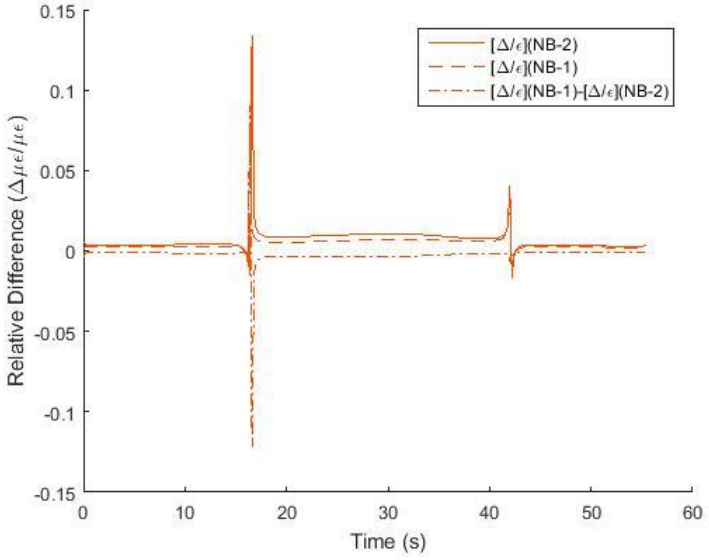

Fig. 65. Damage Effects far from Damage for NB-2 and NB-1

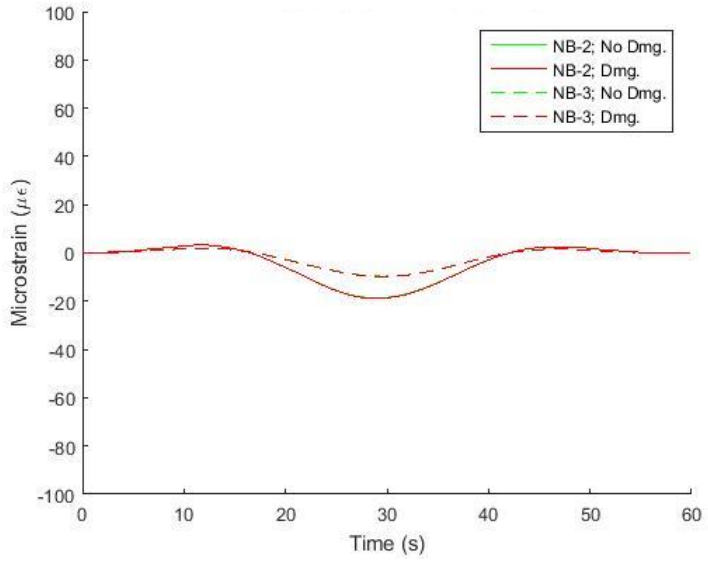

Fig. 66. Strains far from Damage for NB-2 and NB-3 with and without Damage

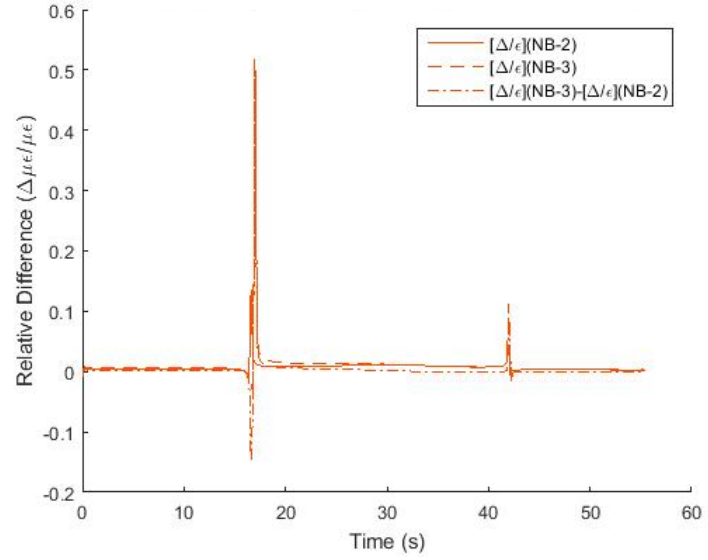

Fig. 67. Damage Effects far from Damage for NB-2 and NB-3

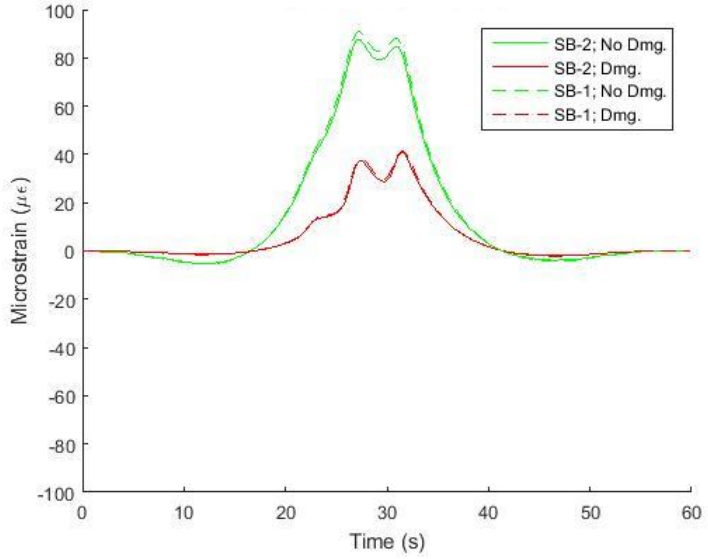

Fig. 68. Strains near Damage for SB-2 and SB-1 with and without Damage

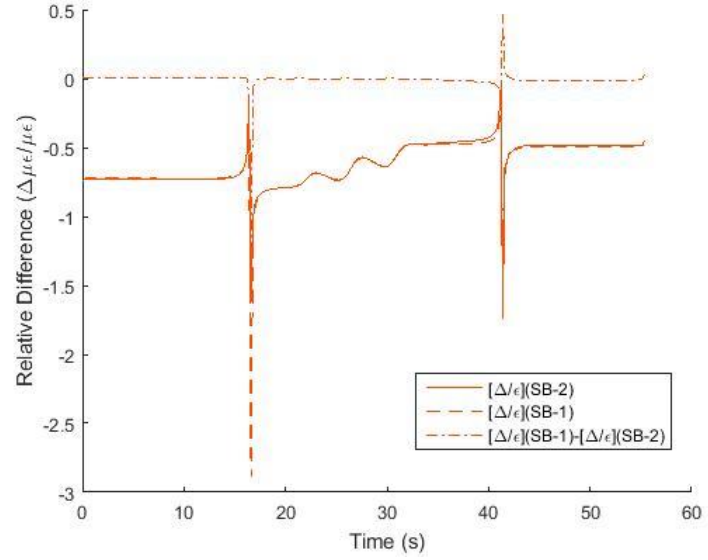

Fig. 69. Damage Effects near Damage for SB-2 and SB-1

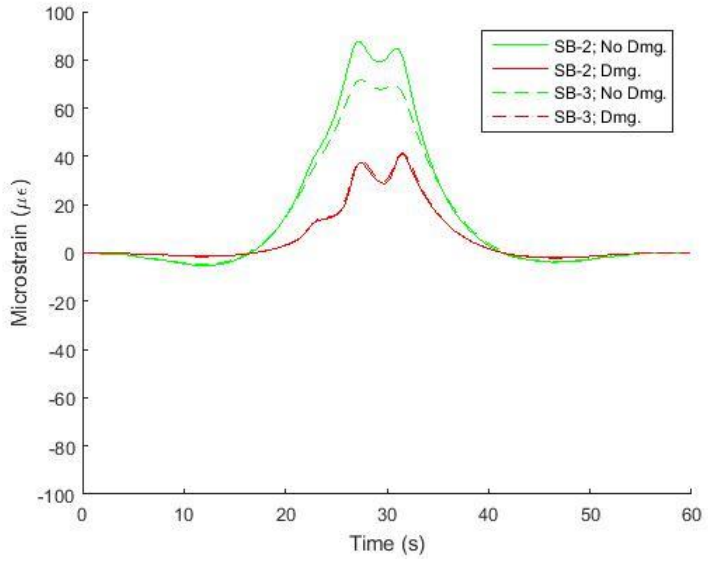

Fig. 70. Strains near Damage for SB-2 and SB-3 with and without Damage

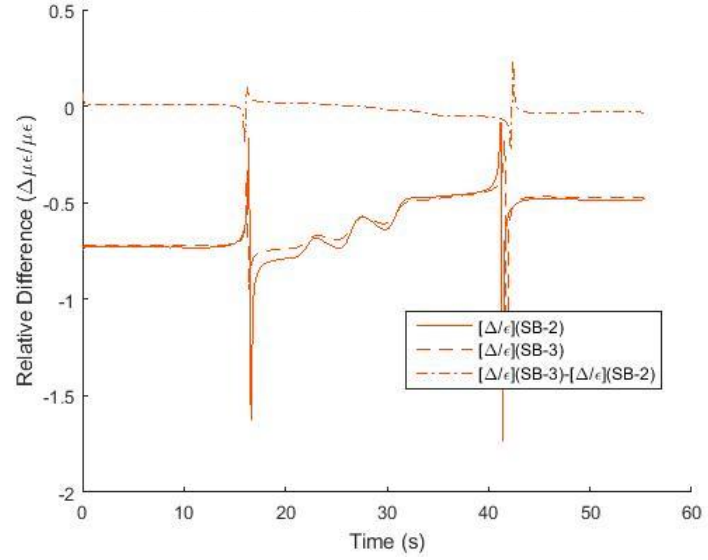

Fig. 71. Damage Effects near Damage for SB-2 and SB-3

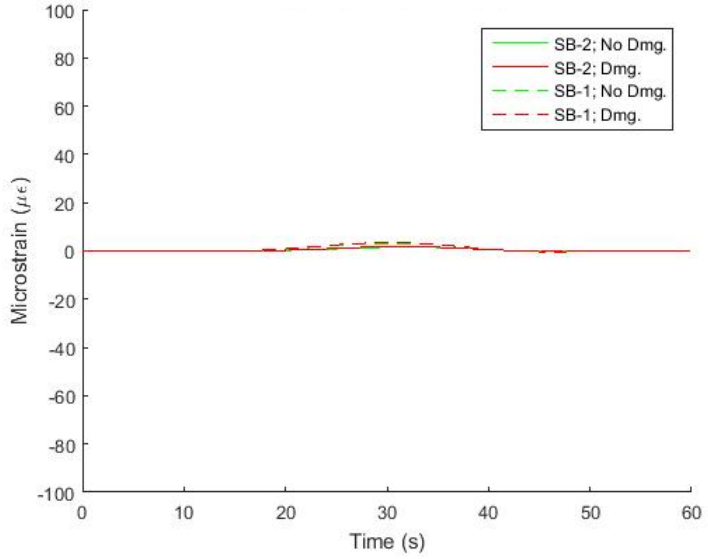

Fig. 72. Strains far from Damage for SB-2 and SB-1 with and without Damage

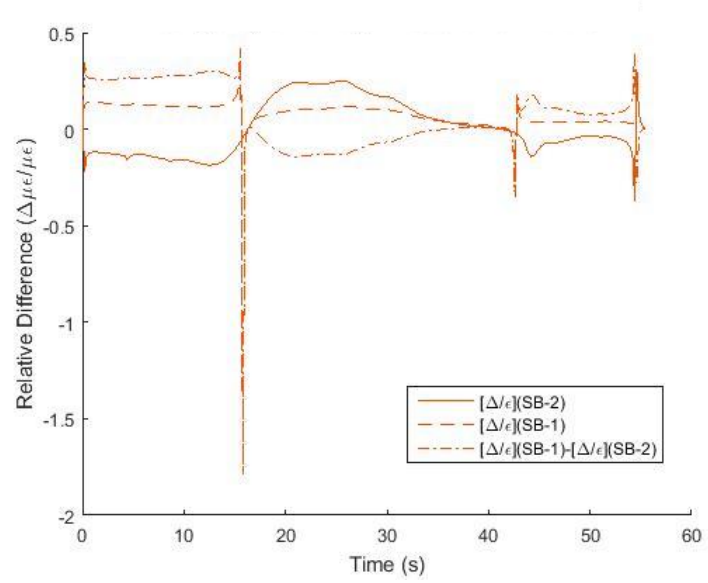

Fig. 73. Damage Effects far from Damage for SB-2 and SB-1

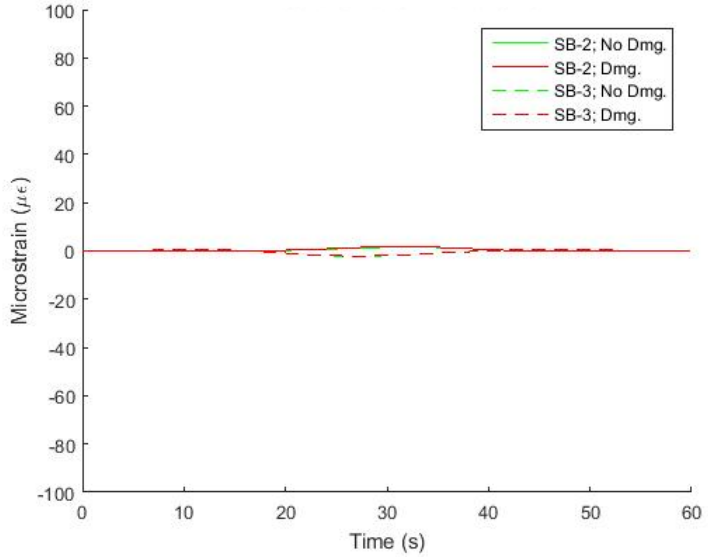

Fig. 74. Strains far from Damage for SB-2 and SB-3 with and without Damage

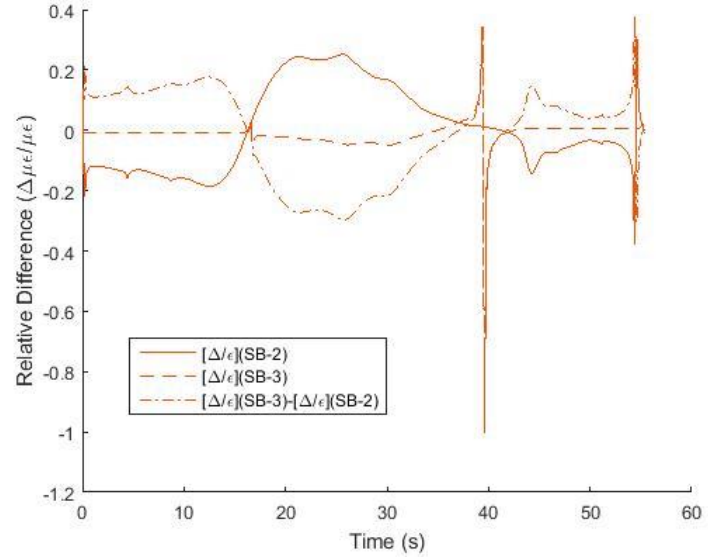

Fig. 75. Damage Effects far from Damage for SB-2 and SB-3