

**Structural Identification of a Complex Structure using both Conventional and
Multiple Model Approaches**

A Thesis

Submitted to the Faculty

of

Drexel University

by

Jeffrey Scott Weidner

in partial fulfillment of the

requirements for the degree

of

Doctor of Philosophy

June 2012

Acknowledgements

I would first like to extend my most sincere gratitude to my advisor, Dr. Frank Moon, who has been a constant source of support, guidance, and optimism from long before I began as a Ph.D. student. He constantly challenged me to improve and succeed, allowed me to stumble when the lesson was great and the consequences were minor, protected me when the opposite was true, and helped me foster a strong sense of confidence in myself. I consider these characteristics to be the paradigm of an academic and an advisor. I am fortunate to have had Frank as an advisor and am proud to count him as my friend.

I would also like to thank Dr. Aktan for dedicating his time and effort, and for putting his trust in me countless times. Throughout my research, he was an invaluable source of knowledge, intuition and experience and I have benefited greatly from his advice.

Similarly, my committee was comprised of a diverse group of experts from several disciplines, all of whom were sources of both support and challenge. I would like to thank Dr. Patrick Gurian, Dr. Alan Lau, Dr. Ivan Bartoli and Dr. Antonios Koutsos for dedicating their time by serving on my committee. I would also like to thank Dr. Anu Pradhan and Dr. Nick Dudley-Ward for providing guidance outside of my committee.

None of this would have been possible without my colleagues and friends, Dr. Nathan Dubbs, Dr. John Prader and Matt Yarnold. If I ever had a problem and no one else could help, I could always count on them. I would also like to thank the new class of students following us, John DeVitis, Dave Maseri, and Adrienne Deal, for providing a fresh perspective and keeping me on my toes.

My research and education were made possible through funding provided by the National Science Foundation (NSF) and the Federal Highway Administration (FHWA). The views presented in this work are those of the Author alone, and do not represent the views of NSF or FHWA.

My curiosity and my affinity for learning have been nurtured since I was a kid by my parents, Dr. Laura Weidner and Dan Weidner. Without their support and love, none of this would be possible.

Finally, though it is woefully inadequate, I would like to extend my deepest thanks and appreciation to my amazing wife, Dr. Jennifer Weidner, who made all of this possible through her unwavering support, timely motivation and endless patience. I am truly blessed to share my life with such a strong, intelligent and inspirational woman.

Table of Contents

Structural Identification of a Complex Structure using both Conventional and Multiple Model Approaches.....	i
Acknowledgements	ii
Table of Contents	iii
List of Figures	x
List of Tables.....	xviii
Abstract.....	xxi
Chapter 1: Introduction.....	1
1.1 Motivation and Context	1
1.2 Overview of Structural Identification.....	3
1.3 Thesis Objectives and Organization.....	5
1.4 Thesis Structure.....	7
Chapter 2: Background	11
2.1 Current Practice of Bridge Assessment	11
2.2 Heuristic and Analytical Load Ratings.....	13
2.3 Load Testing	14
2.4 Structural Identification	15
2.4.1 Step 1: Observation and Conceptualization.....	16
2.4.2 Step 2: A Priori Modeling.....	17
2.4.3 Step 3: Experimentation	19
2.4.4 Step 4: Data Reduction	21
2.4.5 Step 5: Model-Experiment Correlation	22
2.4.6 Step 6: Simulation for Bridge Management	23
2.5 Overview – Types of Model-Experiment Correlation.....	23
2.5.1 Direct Updating.....	23
2.5.2 Iterative Methods.....	24
2.5.3 Multiple Model Methods	26
2.6 Summary	32
Chapter 3: Observation and Conceptualization.....	34
3.1 Background of the International Bridge Study	34
3.2 Selection of the IBS Test Bridge	35
3.3 Overview of the Structure (Review of Design Drawings).....	37
3.4 Overview of Current Condition (Review of Inspection Reports)	41

3.5	Relation to Overall Bridge Population	42
3.6	On-site Observations (Initial Field Visit)	46
3.7	Preliminary Data Collection	49
3.8	Preliminary Load Rating.....	51
3.9	Qualitative Risk Assessment and Development of Critical Questions	53
3.10	Identification of the Relevance of St-Id and Potential Technologies.....	57
Chapter 4:	A Priori Modeling	59
4.1	Purpose of A Priori Modeling.....	59
4.2	Selection of Model Resolution.....	60
4.3	Development of Model Geometry.....	61
4.4	A Priori Model Construction.....	63
4.5	Model Error Screening	66
4.6	Benchmark Studies	67
4.6.1	Example Benchmark Study - Beam-Deck Compatibility.....	68
4.7	Utilization of the Completed Model for Sensitivity Studies.....	71
4.7.1	Identification of Model Building Blocks (Parameter Selection)	73
4.7.2	Sensitivity Study	74
4.7.3	MATLAB & Strand7 Interface	77
4.7.4	Parameter Sampling	78
4.7.5	Results	79
Chapter 5:	Live Load Test: Experimental Design and Results	83
5.1	Identification of Constraints	83
5.2	Design of Live Load Test Program	85
5.2.1	Load Levels.....	86
5.2.2	Load Configuration	88
5.2.3	Traffic Control Plan.....	89
5.3	Criteria for Instrumentation Plan.....	90
5.3.1	Requirements to Meet Test Objectives.....	91
5.3.2	Situational Awareness	93
5.4	Live Load Instrumentation Design.....	94
5.4.1	Instrumentation Plan.....	94
5.4.2	Displacement Measurements	94
5.4.3	Strain Measurements	99
5.4.4	Other Instrumentation.....	105

5.5	Centralized vs. Distributed Data Acquisition	108
5.6	IBS Case Study - Data Acquisition Hardware	110
5.6.1	Configuration for Outdoor Field Work	111
5.6.2	Software Design	114
5.7	Live Load Test: Execution, Processing and Results	120
5.7.1	General Guidance for Test Execution	120
5.7.2	IBS Test Execution.....	122
5.7.3	Data Reduction and Visualization.....	126
5.7.4	Framework for Data Interpretation	130
5.8	Peak Responses	133
5.9	Linearity Checks	134
5.9.1	General Linearity Checks	134
5.9.2	Observations from Linearity Checks - Softening of Girder #1.....	136
5.10	Global Behavior and Spatial Relationships of Data.....	140
5.10.1	Global Displacement Basin	140
5.10.2	Observations of Girder Line Profiles – Loss of Composite Action	145
5.11	Piercap Crack	150
5.12	Analysis of Lateral Load Transfer System	152
5.12.1	Diaphragm Performance.....	152
5.12.2	Wind Brace Performance.....	153
5.13	Test Data Reduction Not Completed.....	159
5.13.1	Crawl Speed Test Results	159
5.13.2	Ambient Traffic Data Reduction.....	159
5.14	Dynamic Test Results.....	161
Chapter 6:	Model-Experiment Correlation	164
6.1	Considerations for Model-Experiment Correlation Methods	164
6.1.1	Parameterization of a Model.....	164
6.1.2	Evaluation Method	165
6.1.3	Reconciliation of Information Formats	165
6.1.4	Importance (Weight)	166
6.1.5	Pitfalls	167
6.2	Pre-modification to A Priori Model	168
6.2.1	Problem Reduction	168
6.2.2	Discretization	169

6.2.3	Beam to Deck Connection	169
6.2.4	Remodeling of the Bearings to Represent Actual Kinematics.....	170
6.3	Manual Calibration.....	172
6.3.1	Parameterization of IBS Model for Manual Updating	172
6.3.2	Optimization Parameters	174
6.3.3	Calibration Results	175
6.3.4	Major Changes to the Model.....	177
6.3.5	Conclusions – Manual Calibration.....	183
6.4	Parameter Identification – Static Results.....	183
6.4.1	Parameterization of IBS for Automated Updating	184
6.4.2	Calibration Results - Displacements	186
6.4.3	Calibration Results – Additional Runs	188
6.5	Parameter Identification – Modal Properties	190
6.5.2	Calibration Results – Modal Properties	191
6.5.3	Comparison of Statically and Dynamically Calibrated Models	193
6.5.4	Load Ratings	195
6.5.5	Conclusions – Parameter Identification.....	197
Chapter 7:	Utilization of the Calibrated Model for Simulation	199
7.1	Potential Scenario Investigations for IBS	199
7.1.1	Mitigation of the Excessive Span Vibration	200
7.2	Sample Simulation – Removal of Wind Brace	202
7.2.1	Process of Wind Brace Removal.....	203
7.2.2	Wind Brace Removal Results	204
Chapter 8:	Multiple Model-Experiment Correlation.....	205
8.1	Summary of Previous Model-Experiment Correlation.....	205
8.2	Model-Experiment Correlation within Multiple Model St-Id	206
8.3	Markov Chain Monte Carlo	207
8.3.1	Selection of Proposal Functions.....	211
8.3.2	Identifying Burn-in.....	215
8.3.3	Assessing Convergence	216
8.4	Reversible Jump Markov Chain Monte Carlo	217
8.5	Multiple-chain Markov Chain Monte Carlo (MC3) Methods	220
8.5.1	Step 1: Development of a candidate population and parameter sets.....	220
8.5.2	Step 2: Determination of prior model weights.....	221

8.5.3	Step 3: Select a sampling method and an appropriate starting point	226
8.5.4	Step 4: Development of response predictions	228
8.6	Approach to Multiple Model Methodology Development	229
Chapter 9:	Numerical Study of Various Multiple-Model Approaches	231
9.1	Multiple Model Applications for a Numerical Example	231
9.2	Cantilever Beam System.....	232
9.3	Cantilever MCMC Analysis	236
9.4	Cantilever RJMCMC Analysis	243
9.4.1	Establishing the Problem	243
9.4.2	Scenario #1 – Actual model included in model population.....	245
9.4.3	Scenario #2 – Actual model not included in model population.....	248
9.4.4	Implementation of Alternative Proposal Functions	259
9.4.5	Discussion of RJMCMC for the Cantilever	261
9.5	Application of MC3 methods to the Cantilever Problem	262
9.5.1	Method #1 – MC3 Resampling of Individual Converged MCMC Chains	263
9.5.2	Method #2 – Real-time MC3	275
9.6	Assessment and Validation of Multi-Dimensional Bayesian Analyses	279
9.6.1	Verification of Negative Effects of Resetting Behavior in RJMCMC	279
9.6.2	Robustness of Predictions	283
9.6.3	Considerations for Application to Real, Constructed Systems	283
Chapter 10:	Multiple Model Structural Identification of IBS Bridge.....	285
10.1	Summary of IBS Bridge Model-Experiment Correlation.....	285
10.2	Problem Definition	286
10.2.1	Experimental Data.....	286
10.2.2	Finalization of Building Blocks and Spatial Variations	288
10.2.3	Determination of the Candidate Model Population.....	291
10.2.4	Modification of the Element Level Model	295
10.3	Development of Individual MCMC Chains.....	295
10.3.1	Optimization for Initial Likelihood and Starting Point Estimation.....	296
10.3.2	Burn-in the MCMC Analysis	298
10.3.3	Development of the Full MCMC Chains for the Selected Model Populations	308
10.3.4	Results from the Individual Chains	309
10.4	Mixing the Chains via MC3.....	321

Chapter 11: Results of MM St-Id of IBS Bridge.....	323
11.1 Uncertainty Regarding Predictions	323
11.2 Experimental Value vs. Mean Value vs. Maximum Likelihood.....	324
11.3 Translation from Measured Observations to Unmeasured Response Predictions – MCMC Chains	325
11.3.1 MCMC Chain Output – Displacements from Final Load Stage.....	326
11.3.2 MCMC Chain Output – Strains from the Final Load Stage	333
11.3.3 Comments on the MCMC Chain Response Predictions.....	342
11.4 Translation from Service to Proof Limit State –MCMC Chains	343
11.4.1 MCMC Chain Output – Displacement at Service Limit State.....	344
11.4.2 MCMC Chain Output – Strain at Service Limit State.....	349
11.4.3 MCMC Chain Output - Comparison of Service and Proof Limit State	354
11.5 Summary of Multiple Model Analysis for the IBS Bridge.....	357
Chapter 12: Comparison of Traditional and Multiple Model St-Id.....	359
12.1 Comparison of MCMC Response Predictions to Single Model Predictions	359
12.1.1 Frequency Response Comparisons.....	360
12.1.2 Displacement Prediction Comparisons.....	365
12.1.3 Strain Prediction Comparisons.....	370
12.2 Assessment of Parameter Interaction	374
12.2.1 Investigation of the Effects of Longitudinal Bearing Stiffness.....	378
12.3 Sample Load Rating and Factor of Safety	385
12.3.1 Sample Rating Truck Load and Position	386
12.3.2 Nominal Capacity and Dead Load Demand	387
12.3.3 RF for Midspan Stress – Girder #3	389
12.3.4 FS for Midspan Stress – Girder #3.....	390
12.3.5 Comparison of RF and FS to Single Model Prediction	391
12.4 Discussion St-Id Comparisons	393
Chapter 13: Conclusions, Recommendations and Future Work	396
13.1 Summary	396
13.2 Primary Conclusions	396
13.3 Conclusions Regarding the IBS Test Structure.....	397
13.4 Conclusions and Recommendations Regarding Short-term Testing (i.e., St-Id) of Bridges	398
13.5 Conclusions Regarding Single Model-Experiment Correlation	401

13.6 Conclusions Regarding the Cantilever Investigation into Multiple Model Methods	402
13.7 Conclusions Regarding the Development of MCMC Chains for the IBS Bridge	403
13.8 Conclusions Regarding Traditional and Multiple Model St-Id	406
13.9 Future Work	407
List of References	409
Vita	421

List of Figures

Figure 1-1 - Structural Identification Paradigm	4
Figure 1-2 - Thesis Organization.....	5
Figure 2-1 - Variation in Condition Rating (Phares, Washer et al. 2004)	12
Figure 3-1 - US202/NJ23 International Test Specimen in Wayne, NJ	36
Figure 3-2 - Plan View of Bridge	38
Figure 3-3 - Typical Wind Bracing for Fascia Girders.....	40
Figure 3-4 - Typical Diaphragms	40
Figure 3-5 - Wind Brace Connection	42
Figure 3-6 - Sub-population of the test bridges (All % of Total Population)	43
Figure 3-7 Distributions of the family of bridges based on (a) span length, (b) ADTT,...	45
Figure 3-8 - Fully Cracked Pintel.....	47
Figure 3-9 - Crack in SB Pier 1	48
Figure 3-10 - Acceleration Time History and Power Spectral Densities for Span 2 NB ..	50
Figure 3-11 - Acceleration Time History and Power Spectral Densities for Span 2 SB ...	50
Figure 4-1 - 3D Centerline Model vs. Actual Structure	62
Figure 4-2 - Strand7 Finite Element Model	64
Figure 4-3 - Modeling Assumptions for Connection to Piercap	66
Figure 4-4 - Modeling Techniques for Benchmark Study	69
Figure 4-5 - Benchmark Models (Extruded View) a. Shell model b. Frame element model	70
Figure 4-6 - Displacement Prediction Distributions.....	72
Figure 4-7 - Sensitivity of Girder #3 Displacement to % Composite Action	75

Figure 4-8 - Sensitivity of Girder #3 Displacement to Cross-brace Stiffness	75
Figure 4-9 - Sensitivity of Girder #3 Displacement to Rotational Bearing Stiffness.....	76
Figure 4-10 - Midspan Displacement Response Histogram	80
Figure 4-11 - Midspan Strain Response Histogram	80
Figure 5-1 - Actual Traffic Control Plan	89
Figure 5-2 - Instrumentation Design Flowchart.....	91
Figure 5-3 - Displacement Sensor, Mounting Plate and Configuration	95
Figure 5-4 - Creep Behavior of Fishing Line for Displacements.....	96
Figure 5-5 - Displacement Sensor Layout	98
Figure 5-6 - Installed Strain Gages.....	100
Figure 5-7 - Longitudinal Strain Gage Configuration	101
Figure 5-8 - Diaphragm Strain Gage Configuration (Elevation)	102
Figure 5-9 - Wind Brace Sensor Layout (Plan View)	103
Figure 5-10 - Strain Gage Locations - Span 2 SB	104
Figure 5-11 - Strain Gage Locations - Span 1 SB	105
Figure 5-12 - Piercap Crack	107
Figure 5-13- Installation of Clip Gage over Crack.....	108
Figure 5-14 - DAQ Box Layout	112
Figure 5-15 – Data Acquisition Network Diagram	113
Figure 5-16 - Software Architecture.....	115
Figure 5-17 - Real-time Visualization Setup	119
Figure 5-18 - Control Screen for Test Control.....	119
Figure 5-19 - Control Screen for Rapid Data Interpretation	120

Figure 5-20 - Span SB2 Fully Loaded with Six Trucks	123
Figure 5-21- Plot Matrix for Displacement	129
Figure 5-22 - Displacement Linearity Check.....	135
Figure 5-23 - Strain Linearity Check.....	136
Figure 5-24 – Disp. Linearity Check for Load at 1/4 Point	137
Figure 5-25 - Strain Linearity Check - Load at 1/4 Point.....	138
Figure 5-26 - Strain Linearity Check - Load at 3/4 Point.....	139
Figure 5-27 - Overall Displaced Shapes	140
Figure 5-28 - Girder Disp. Profiles.....	141
Figure 5-29 - Transverse Displacement Plot.....	143
Figure 5-30 - Strain Profile along Girders - Load at Midspan	144
Figure 5-31 - Strain Cross-section at Midspan	145
Figure 5-32 - Compatibility Plot - Girder #3 Midspan.....	146
Figure 5-33 - Extruded View of Benchmark Beam Model	147
Figure 5-34 - Axial/Moment Diagram for Benchmark - Full Transfer	148
Figure 5-35 - Axial/Moment Diagram for Benchmark – No Transfer b/w Girder and Deck	149
Figure 5-36- Piercap Rebar Layout	151
Figure 5-37 -Diaphragm Response under Ambient Traffic	153
Figure 5-38 - Divergence of Wind Brace Strains due to Temperature.....	155
Figure 5-39 - Three-truck Ambient Event	158
Figure 5-40 - Ambient Data Collection with Video	160
Figure 5-41 - Camera Setup for Ambient Traffic	161

Figure 5-42 - Experimental Mode Shapes	163
Figure 6-1 - Modeling the Bearings	171
Figure 6-2 - Composite Action Distribution	181
Figure 6-3 - Composite Action Sensitivity	182
Figure 7-1- Wind Bracing to be Removed.....	203
Figure 8-1 - General MCMC Procedure with Metropolis-Hastings Algorithm.....	209
Figure 8-2 - Proposal Distribution Schematic - Adapted from Hanson (2000)	211
Figure 8-3 - Proposal Distribution Study - Adapted from Dubbs (2012).....	212
Figure 8-4 - Case A with Adaptation to the Proposal Function	213
Figure 8-5 - Proposal Function Value with Adaptation	214
Figure 8-6 - Burn-in for Case B with Poor Starting Value	215
Figure 8-7 - RJMCMC Process with Metropolis-Hastings Algorithm	218
Figure 9-1 - Base Cantilever Model	232
Figure 9-2 - Generic Model Forms of Cantilever Beam.....	233
Figure 9-3 - Experimental Mode Shapes	235
Figure 9-4 - Location of Load and Unmeasurable Responses	236
Figure 9-5 - MCMC Output Chain - Model E - Parameters 1 through 4	238
Figure 9-6 - MCMC Output Chain - Model E - Parameters 5 through 8	239
Figure 9-7 - Gradual Drop in Likelihood - Model E	241
Figure 9-8 - CDF Plot for Tip Displacement - Model E.....	241
Figure 9-9 - Comparison of Model D and E - CDF of Tip Displacement.....	242
Figure 9-10 - Model D as the Base Case	246
Figure 9-11 - Scenario #1 - Chain of Accepted Models.....	247

Figure 9-12 - Accepted Model Chain - Scenario #1.....	249
Figure 9-13 - Partial Burn-in for Scenario #1	252
Figure 9-14 - CDF of Tip Displacement - Scenario #2	255
Figure 9-15 - Strain CDF Plots	256
Figure 9-16 - CDF Plots of Frequencies	257
Figure 9-17 - RJMCMC Comparison of Inter and Intra-model Moves	260
Figure 9-18 - Displacement CDF - MC3 with Uniform Weights - Actual Model Included	266
Figure 9-19 - Strain CDF - MC3 with Uniform Weights – Actual Model Included.....	267
Figure 9-20 - Displacement CDF - MC3 with BIC Weights	268
Figure 9-21 - Strain CDF - MC3 with BIC Weight	269
Figure 9-22 - Method #1 - Comparison of Different Weighing Schemes	271
Figure 9-23 - Displacement CDF - MC3 with Uniform Weight - Actual Model Excluded	272
Figure 9-24 - Strain CDF - MC3 with Uniform Weight - Actual Model Excluded.....	273
Figure 9-25 - Displacement CDF - MC3 with BIC - Actual Model Excluded	274
Figure 9-26 - Strain CDF - MC3 with BIC Weights - Actual Model Excluded	275
Figure 9-27 - Comparison of Real-time and Resampled MC3	278
Figure 9-28 - Comparison of Likelihoods for Model E – RJMCMC (first proposal functions) vs. MC3	280
Figure 9-29 - Likelihood Comparison - RJMCMC (updated proposal function) vs. MC3	280
Figure 9-30 - Comparison of Displacement CDF - Inclusion of Actual Model	281

Figure 10-1 - Composite Action Parameter Discretization.....	292
Figure 10-2 - Vertical Stiffness Parameters.....	293
Figure 10-3 - Accepted Parameter Chains - Model I	299
Figure 10-4 - Accepted Parameter Chains - Model I with Reduced Ranges	300
Figure 10-5 - Dummy Parameter Proposal Function	301
Figure 10-6 - Comparison of starting points for two MCMC runs of Model A	303
Figure 10-7 - Likelihood Comparison for Different Starting Points	304
Figure 10-8 - Parameter Means for all Models	310
Figure 10-9 - Parameter 5, 6, and 7 vs. Likelihood - Model C	312
Figure 10-10 - Plot Matrix of Parameters 5 and 6.....	313
Figure 10-11 - Parameter Chains for all Models (Red = Model C; Blue = Model H; Grey = Model I)	314
Figure 10-12 - CDF of 1st Frequency - All Models	315
Figure 10-13 - CDF of 2nd Frequency - All Models.....	316
Figure 10-14 – CDF of Third Frequency – All Models	317
Figure 11-1 - Sample Displacement Prediction Showing Mean, Experimental Value, ..	325
Figure 11-2 - Girder #1 Displacement Predictions – Six Full Trucks.....	327
Figure 11-3 - Girder #1 Vertical Bearing Stiffness	328
Figure 11-4 - Girder #3 Displacement Predictions – Six Full Trucks.....	329
Figure 11-5 - Girder #6 Displacement Predictions – Six Full Trucks.....	330
Figure 11-6 - Girder #8 Displacement Predictions – Six Full Trucks.....	331
Figure 11-7 - Girder #1 Strain Predictions – Six Full Trucks.....	335
Figure 11-8 - Girder #3 Strain Predictions – Six Full Trucks.....	336

Figure 11-9 - Girder #6 Strain Predictions - Six Full Trucks	337
Figure 11-10 – Girder #8 Strain Prediction – Six Full Trucks.....	338
Figure 11-11 - Truck Position for Three Empty Truck Case	344
Figure 11-12 – Girder #1 Displacement Predictions – Three Empty Trucks	345
Figure 11-13 - Girder #3 Displacement Predictions - Three Empty Trucks.....	345
Figure 11-14 - Girder #6 Displacement Predictions - Three Empty Trucks.....	346
Figure 11-15 - Girder #8 Displacement Predictions - Three Empty Trucks.....	346
Figure 11-16 - Girder #1 Strain Predictions - Three Empty Trucks	349
Figure 11-17 - Girder #3 Strain Predictions - Three Empty Trucks	350
Figure 11-18 - Girder #6 Strain Predictions - Three Empty Trucks	351
Figure 11-19 - Girder #8 Strain Predictions - Three Empty Trucks	352
Figure 12-1 - Comparison of MCMC and Single Model to Experimental Frequency #1	361
Figure 12-2 - Comparison of MCMC and Single Model to Experimental Frequency #2	361
Figure 12-3 - Comparison of MCMC and Single Model to Experimental Frequency #3	362
Figure 12-4 - Comparison of MCMC and Single Model to Experimental Frequency #4	362
Figure 12-5 - Comparison of MCMC and Single Model to Experimental Frequency #5	363
Figure 12-6 - Comparison of MCMC and Single Model to Experimental Frequency #6	363

Figure 12-7 - Comparison of MCMC and Single Model to Experimental Frequency #7	364
Figure 12-8 - Comparison of Girder #1 Displacements	366
Figure 12-9 - Comparison of Girder #3 Displacements	366
Figure 12-10 - Comparison of Girder #6 Displacements	367
Figure 12-11 - Comparison of Girder #8 Displacements	367
Figure 12-12 - Comparison of Girder #1 Strains	370
Figure 12-13 - Comparison of Girder #3 Strains	371
Figure 12-14 - Comparison of Girder #6 Strains	371
Figure 12-15 - Comparison of Girder #8 Strains	372
Figure 12-16 - Matrix Plot of Chain for Model H.....	375
Figure 12-17 - Parameter versus Displacement Matrix Plot.....	377
Figure 12-18 – Parameter versus Strain Matrix Plot	378
Figure 12-19 - Midspan Displacement of Girder #3	379
Figure 12-20 - Frequency versus Parameter 8 - Model H	380
Figure 12-21 - MAC Value versus Parameter 8 - Model H.....	381
Figure 12-22 - Histogram of Parameter 8 - Model H.....	382
Figure 12-23 - Parameter 8 versus Midspan Displacement - Girder #3	383
Figure 12-24 - Selected Rating Truck	386
Figure 12-25 - RF for Stress at Midspan of Girder #3	390
Figure 12-26 - Factor of Safety for Stress at Midspan of Girder #3	391
Figure 12-27 - Rating Factor Comparison	392
Figure 12-28 - Factor of Safety Comparison	393

List of Tables

Table 3-1 - General Bridge Information	38
Table 3-2 - Distribution of the sub-population based on age and maximum span length	44
Table 3-3 - Identified Frequencies - Preliminary Data.....	51
Table 3-4 - Load Rating Factors	52
Table 3-5 - Summary of Performance Limit States, Hazards, Vulnerabilities, and Exposures for Bridges.....	54
Table 4-1- Displacement Comparison for Benchmark Models	70
Table 4-2 - Parameter Bounds.....	76
Table 4-3 - A Priori Strain Results Summary	81
Table 4-4 - A Priori Displacement Results Summary	82
Table 5-1 - Centralized vs. Distributed Data Acquisition	109
Table 5-2 - cRIO Specifications.....	110
Table 5-3 - Test Timeline	123
Table 5-4 - Addressing Critical Questions through Data Visualization	132
Table 5-5 - Peak Responses	134
Table 5-6 - Modal Properties (September 2010).....	162
Table 6-1: Displacement Comparison	176
Table 6-2: Strain Comparisons.....	176
Table 6-3 - Displacement Comparison -Parameter ID	186
Table 6-4 - Strain Comparison - Parameter ID	187

Table 6-5 - Strain Predictions when Updating with Strains	188
Table 6-6 - Displacement Predictions when Updating with Strains	189
Table 6-7 - Modal Calibration Results	192
Table 6-8 - Final Parameter Values	193
Table 6-9 - Comparison of Static and Dynamic Model – Dynamic Response	194
Table 6-10 - Comparison of Static and Dynamic Model - Static Response	195
Table 6-11 - Rating Factor Comparison	196
Table 9-1 - "Experimental" Frequencies	234
Table 9-2 - Unmeasurable Responses	236
Table 9-3 - Scenario #1 - Modal Frequencies.....	246
Table 9-4 - Scenario #1 - Static Responses	246
Table 9-5 - Prior MCMC Model Requirements	265
Table 9-6 - Model Attempt and Acceptance Counts – Real-time MC3	277
Table 10-1 - Model Designation and Properties	291
Table 10-2 - Parameter Breakdown by Model.....	294
Table 10-3 - Likelihood and BIC Weight from Burn-in.....	307
Table 10-4 - MCMC Chain Results.....	309
Table 10-5 - Comparison of Frequency Prediction Mean to Experimental Values	318
Table 10-6 - Comparison of Percent Errors of Frequency Predictions	318
Table 11-1 - Displacement Predictions from Each Model Form – Six Full Trucks	332
Table 11-2 - Percent Errors for Mean Displacement Predictions – Six Full Trucks.....	332
Table 11-3 - Displacement Predictions - Coefficient of Variation	333
Table 11-4 - Strain Predictions in Each Model Form – Six Full Trucks	339

Table 11-5 - Percent Errors for Strain Predictions – Six Full Trucks	340
Table 11-6 - Strain Prediction Coefficient of Variation	341
Table 11-7 - Mean Displacement Predictions - Three Empty Trucks.....	348
Table 11-8 - Mean Displacement Errors - Three Empty Trucks.....	348
Table 11-9 - Mean Strain Predictions - Three Empty Trucks.....	353
Table 11-10 - Mean Strain Errors - Three Empty Trucks.....	353
Table 11-11 – Average Model Prediction Displacement Errors – Magnitude	355
Table 11-12 – Average Model Prediction Strain Error - Magnitude	356
Table 12-1 - Frequency Comparison - Single Model and MCMC Results	365
Table 12-2 - Mean MCMC and Single Model Displacement Predictions.....	368
Table 12-3 - Percent Error from Mean MCMC and Single Model Displacement Predictions.....	369
Table 12-4 - Magnitude of Displacement Errors	369
Table 12-5 - Mean MCMC and Single Model Strain Predictions.....	373
Table 12-6 - Percent Error from Mean MCMC and Single Model Strain Predictions....	373
Table 12-7 - Correlation Coefficients - Model H.....	376
Table 12-8 - Dead Load Demand (from nominal model)	388
Table 12-9 - Capacity for Ratings	388

Abstract

Structural Identification of a Complex Structure using both Conventional and Multiple Model Approaches

Jeffrey Scott Weidner

Franklin Moon, Ph.D.

The confluence of a debilitated, often obsolete bridge population nearing the end of its design life with a massive shortfall of funding for rehabilitation and replacement has led to an unsustainable balancing act for bridge owners. One tool for addressing this issue is Structural Identification (St-Id), a six-step process for effectively integrating subjective information from visual inspections, with analytical modeling and experimental investigations to provide a more robust and defendable foundation for decision-making regarding constructed systems.

In an effort to understand and document St-Id from a global perspective, the International Bridge Study (IBS) brought researchers from all over the world to demonstrate various techniques and technologies on a single test specimen in New Jersey. Through this application a comprehensive documentation of a best-practices application of St-Id was developed and numerous advances to each step of the process and their integration were achieved. These advances included formalization of the criteria for bridge selection as a candidate for St-Id, development of a framework for instrumentation design considering safety and test objectives, design and implementation of a distributed data acquisition system with a real-time visualization interface and spatially-correlated data interpretation.

As part of this application some shortcomings associated with conventional model-experiment correlation approaches were identified. The second phase of the research aimed to develop and validate rigorous methods to mitigate these shortcomings. This research examined multiple-model approaches such Markov Chain Monte Carlo (MCMC) and multi-dimensional MCMC methods through both a benchmark numerical problem and the IBS Bridge. The primary findings in this phase included the difficulty in framing the problem to fully map the model space using Reversible Jump MCMC, an alternative approach through a multi-dimensional MCMC method, the difficulty in applying MCMC to actual bridges due to the complexity of the likelihood space and that MCMC can implicitly account for interaction between parameters through covariance and correlation coefficients. In general, the response predictions from the MCMC chains were more robust and informative than the results from the single model correlation.

Chapter 1: Introduction

1.1 Motivation and Context

Since the collapse of the Minnesota I-35W Bridge in 2007, a great deal of media attention and government deliberation has been dedicated to the state of the nation's infrastructure system, particularly bridges. The confluence of a debilitated, often obsolete bridge population nearing the end of its design life with a massive shortfall of funding for rehabilitation and replacement has led to an unsustainable balancing act for bridge owners. Funding is allocated through a worst-first system where the most problematic structures get the most attention (AASHTO 2008). This approach to prioritization is informed predominately by visual inspection, and results in bridges that appear to be in poor shape receiving the most attention. Counter intuitively, these older structures may have safety margins in excess of new bridges due to the larger conservatism associated with past design procedures. Unfortunately, since such reserve capacity cannot be identified through qualitative and subjective assessment approaches, it is quite common for bridges to be posted for load levels lower than the legal limit based only on their physical appearance.

Conservative as this approach may be, it can come at an excessive cost. A bridge posting can throttle commerce, create traffic issues, limit the mobility of emergency vehicles, undermine the confidence of the traveling public, and generate weight enforcement problems for the bridge owner. Currently there are over 60,000 bridges that are, or should be, posted based on current methods and a total of over 150,000 bridges that are considered structurally deficient or functionally obsolete (AASHTO 2008). Conversely, there are typically less than 50 bridge failures, defined as a bridge being removed from

service due to safety concerns, in an average year (Wardhana and Hadipriono 2003).

Clearly there is a disconnect between the perception of bridge condition and the reality of bridge safety.

In order to stabilize this system, quantitative data is needed to inform decision-making and owners must begin to develop a mechanistic understanding of bridge performance (Aktan, Ellingwood et al. 2007). Such a transformation is not likely to happen quickly as the reliance on qualitative condition ratings is currently incompatible with more quantitative, objective performance metrics. As a result, it is important that any new information be (to the extent possible) placed into context with the current approaches and presented in a way to facilitate decision-making.

Structural Identification (St-Id) is perhaps the most promising paradigm in this regard.

Figure 1-1 illustrates the six primary steps of St-Id, which have been adopted by the ASCE's Structural Identification of Constructed Systems Committee (2012). In the most general terms, St-Id aims to infer unmeasurable but desired responses/attributes of a structure from correlating measured responses with a simulation model. Stated a slightly different way, St-Id, in one form or another, attempts to correlate one or more analytical models with measured data and then use the correlated models to both estimate current attributes and forecast future performance. St-Id is a valuable tool for aiding in decision-making because when done properly, it can estimate key metrics such as live load capacity, seismic vulnerability, etc. that cannot be directly measured or reliably estimated based on heuristics. While this approach is quite promising, there are many potential pitfalls that not only limit its effectiveness, but can also potentially result in highly

misleading and unreliable results. This is an unfortunate but undeniable part of the history of St-Id, and has fueled skepticism related to the value of this paradigm.

It is stressed that while the St-Id process provides general guidance for the various aspects of a proper technology application, it does not mitigate the need for sound heuristics, which is by far the most important ingredient for success.

1.2 Overview of Structural Identification

As shown in Figure 1-1, the first step of St-Id involves conceptualizing the structure and all concerns and issues that are driving the application. Although this step does not necessarily employ any advanced technologies, it has more influence over the ultimate success of the St-Id than any other as it develops the guiding questions/objectives. In the second step of the process the a priori model is developed and used to predict the bounds of structural responses to aid in the selection of appropriate experimental approaches, sensors, instrumentation plans, etc. This is followed by the actual execution of the experiment (Step 3) - the only objective, quantitative link to the constructed system of interest. While huge advances in sensing and information technologies have been made in the last decade, this step still involves making a series of trade-offs to ensure that the relevant data is acquired with sufficient reliability and without undue cost.

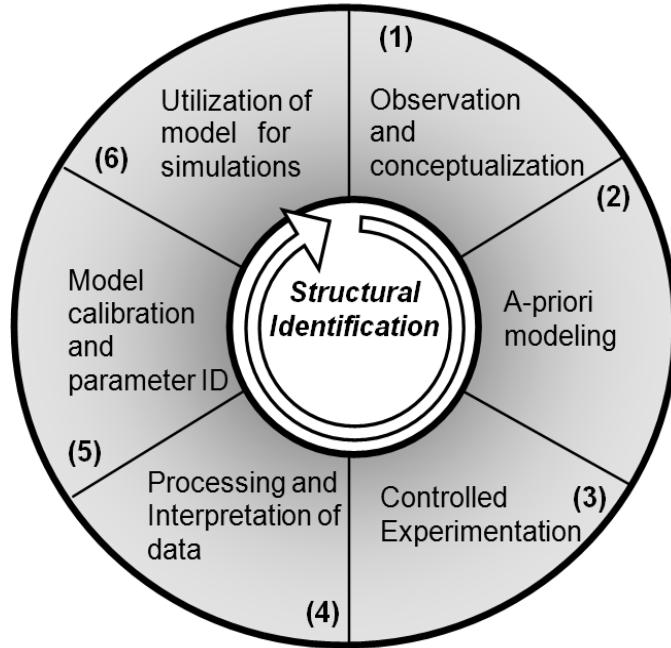


Figure 1-1 - Structural Identification Paradigm

The fourth step involves the processing and interpretation of the experimental data. This step invariably aims to remove data errors and noise (averaging, windowing, filtering), extract key response indices, and plot such indices temporally, spatially, and versus load position, load level, frequency, etc. to facilitate effective interpretation. Once the data has been effectively reduced and interpreted, a simulation model (often from Step 2) is calibrated or updated to minimize its discrepancies with the experiment. The general goal of this process is to reconcile the experiment and model to identify and explain the root causes of the observed data/responses.

The final step of the process involves using the calibrated model through scenario analysis, parametric studies, or what-if simulations, in order to inform decisions. Although it is tempting to think that St-Id directly produces decisions, the reality is that

decisions are necessarily made by balancing a diverse set of qualitative and quantitative information and, often times, multiple and competing objectives. While St-ID may provide necessary information to the decision-making process, it is not, by itself, sufficient. Further, it is important to recognize that St-ID remains an art and its value depends greatly on the talent, heuristics and insight of those that oversee and direct the application.

1.3 Thesis Objectives and Organization

Guided by the four primary scholarships of education, integration, application and discovery, this thesis will aim to advance the St-ID paradigm from two primary perspectives (as illustrated in Figure 1-2).

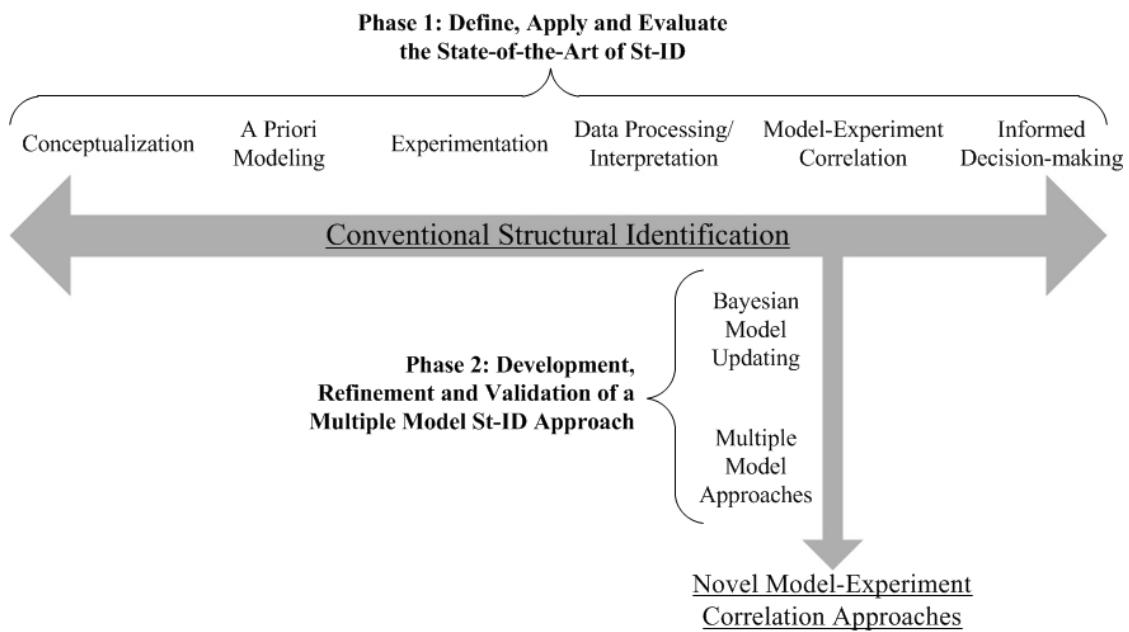


Figure 1-2 - Thesis Organization

The first phase of this research aimed to define, apply, and evaluate the current state-of-the-art of St-Id on an operating bridge. The primary significance of this phase was a series of advancements to the state-of-the-art that included procedural formalizations, improvements to each step of the St-Id process, and enhancements to their integration. This phase focused primarily on education, integration and application. The specific objectives of the first phase included:

- Document the process of conceptualization of a structure as a candidate for St-Id in terms of the context of the bridge in the population of structures, modeling considerations, site constraints, and test feasibility.
- Document and apply a generalized framework for a priori model construction and error screening in a St-Id application.
- Document a best practices experiment, instrumentation and data acquisition system design and implementation within the constraints of a real-life scenario.
- Develop a correlated finite element model using the state of the art in model-experiment correlation for use as a tool to inform decision-making.
- Provide recommendations to the bridge owner from both direct interpretation of the data and from the correlated finite element model.
- Critically assess the current best practices of St-Id and develop a series of recommendations for future developments.

The second phase of this research aimed to improve the robustness of the model-experiment correlation step within St-Id through the incorporation of multiple models to more rationally allow inferences to be made. This new approach was developed, demonstrated and compared to traditional approaches through both a benchmark

numerical study and a field application. This phase focused primarily on discovery. The specific objectives of this phase included:

- Develop a multiple model St-Id approach that allows for the incorporation of different model forms without inter-model parameter dependency so that the model space associated with each model form can be fully explored.
- Evaluate the merit of the developed method and contrast it to other competing approaches through a simple benchmark numerical study.
- Implement the developed multiple model approach on an operating bridge, assess its merit and feasibility, and contrast the outcomes to conventional single model approaches.
- Establish the correlation between various responses of an operating bridge and make recommendations related to how experiments should be structured to allow reliable inference of certain unmeasureable quantities to be made.

1.4 Thesis Structure

The thesis is structured as follows:

Chapter 1 highlights the disconnect between the perception of bridge condition and the reality of bridge safety and how this affects infrastructure decision-making. The paradigm of St-Id is introduced as a tool to help inform decision-making. The organization and objective of this thesis are presented in the context of two phases.

Chapter 2 provides a literature review of several key issues upon which this research is predicated. First the current practice of bridge condition assessment is discussed,

followed by an explanation of St-Id and how it fits into a condition assessment framework, as applied in Phase 1 of this research. Second, a discussion of various model-experiment correlation techniques and the role of this in St-Id is presented, with an emphasis on Multiple Model methods.

Chapter 3 introduces the first step of St-Id, where the structure is conceptualized in order to understand the problem(s) with the bridge, appropriately frame an experimental scenario, and develop reliable a priori models. This is demonstrated within the context of the IBS Bridge, from selection of the bridge, site visits, documentation review, preliminary data collection, and preliminary load rating. Based on this effort, critical questions regarding the bridge are developed which will guide the rest of the investigation.

Chapter 4 will present the process for development of an a priori model. This corresponds to step 2 of St-Id. The chapter includes discussion of high-level decisions like model form selection, resolution, and the extents (physical limits of the model), through specific decisions like element type and mesh density. Model error screening and benchmarking are discussed, leading to utilization of the model in preliminary sensitivity studies of the influence of uncertain parameters.

Chapter 5 presents the experimental program design, implementation and results. This corresponds to step 3 and 4 of St-Id. Experimental design includes framing the load test as a system of constraints which must be satisfied. Within these constraints, the loading characteristics, instrumentation layout, and data acquisition protocols are all designed and presented. A summary of the actual test of the IBS Bridge is presented, followed by a

discussion of the results, independent of any model correlation (i.e., directly interpretable results). Several key observations regarding the structure are presented.

Chapter 6 focuses on traditional model-experiment correlation techniques as a representation of the state of the art. Manual and automated model-correlation techniques are discussed and presented for the IBS Bridge. Finally, the correlated models are used as translators to predict responses that were measured but were not used to inform the correlation.

Chapter 7 completes the discussion of the traditional St-Id application for the IBS Bridge through utilization of the calibrated model from Chapter 6 to conduct simulations of scenarios of interest to the bridge owner.

Chapter 8 introduces multiple model methods for mode-experiment correlation. Discussions of common techniques, which are relatively new to structural engineering, are presented, including Markov Chain Monte Carlo (MCMC) for sampling parameters within a single model form, and Reversible Jump Markov Chain Monte Carlo (RJMCMC) for moving between model forms. A new multiple model method, termed Multiple Chain Markov Chain Monte Carlo, or MC3, is discussed.

Chapter 9 presents the multiple model methods discussed in Chapter 8 in the context of a simple cantilever beam. The methods are compared for robustness and reliability prior to application to a real structure.

Chapter 10 returns to the IBS Bridge case study. A population of candidate model forms is presented and then reduced considering their relative likelihood as a function of modal properties (i.e., frequencies and mode shapes). An MCMC analysis is run for each model

form in the final, reduced population. The results of these individual chains are presented. These analyses are then combined as a resampled MC3 analysis, and the subsequent results are presented.

Chapter 11 analyzes the results of both the individual MCMC analyses and the comprehensive MC3 analysis presented in Chapter 10. Comparisons of single model methods using deterministic and maximum likelihood estimates versus both MCMC and MC3 are provided. Similarly, investigations in the ability of these methods to translate between response modalities are presented. Finally, from the multiple model analyses, individual models will be extracted from the data sets to draw comparisons between models with similar likelihoods across model forms.

Chapter 12 closes the loop between traditional St-Id and multiple model St-Id via comparisons of response predictions from the IBS Bridge and predictions of unmeasurable quantities, specifically load ratings. The focus of this chapter is to highlight the potential impact of multiple model methods on decision-making regarding infrastructure.

Chapter 13 presents several key conclusions regarding both phases of this research. The major recommendations for implementation of St-Id are provided first, followed by conclusions stemming from the application of multiple model methods to the IBS Bridge. Conclusions are presented from comparisons of single and multiple model methods, different types of response modalities, and various model forms.

Chapter 2: Background

This chapter presents a review of the relevant literature in relation to bridge condition assessment, St-Id, and model-experiment correlation. Specifically, the background of bridge condition assessment as it is known today is presented, followed by a discussion of St-Id and its role in assessment. Following that, the topic of model-experiment correlation within the St-Id framework is discussed.

2.1 Current Practice of Bridge Assessment

As a result of the 1969 collapse of the Silver Bridge, in 1972 the U.S. implemented a National Bridge Inspection Standard (NBIS) which dictates that every structure must be inspected at least once every two years (NBIS 1996). Each component of the bridge (Deck, Substructure, and Superstructure) is rated on a 0-9 scale, and if any single element is rated a four or below then the entire bridge is deemed structurally deficient (the worst classification available). In addition to this classification, the condition rating is also used to compute a sufficiency rating (together with other metrics like detour length), which is used to determine whether a bridge is eligible to receive federal funding for either repair or replacement.

While implementation of the NBIS has certainly improved our knowledge of bridges and how they deteriorate, it remains an inherently subjective process. In one particular study this subjectivity was quantified, clearly highlighting the uncertainty with these methods (Phares, Washer et al. 2004). This study reported on an FHWA investigation (FHWA 2001) where 100 experienced inspectors were all asked to independently inspect the same

bridge. The results indicated that 95% of condition rating values from a visual inspection would vary within two points on a scale of 0-9, with a normal distribution (Figure 2-1). Considering the importance of this single metric in terms of prioritization and funding, even this variance is quite troubling.

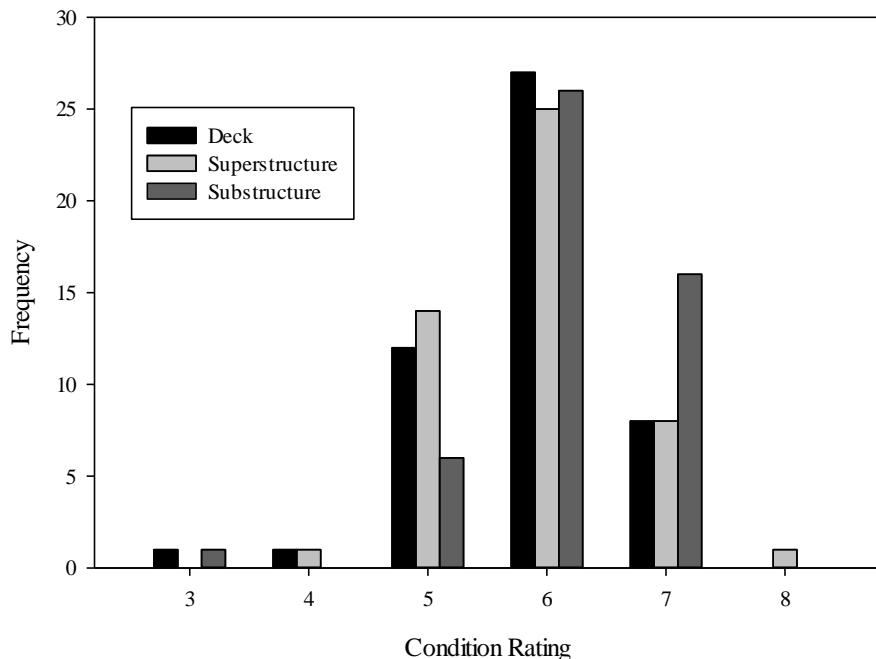


Figure 2-1 - Variation in Condition Rating (Phares, Washer et al. 2004)

Since its inception in the early 1970s, there have been some notable improvements to visual inspection that include fracture critical inspections (prompted by the Mianus River Bridge collapse), underwater inspections (prompted by the Schoharie Bridge collapse),

and element-level inspections (prompted by bridge management systems). In addition, there are several technologies such as ruggedized tablets that permit electronic data input and even direct sketching of deterioration on 3D CAD models. These improvements have made visual inspection more reliable and straightforward, but have not changed the basic framework for condition assessment. With all of these advances, visual assessment remains largely subjective and completely tied to visual appearance as opposed to a more mechanistic understanding of a bridge.

2.2 Heuristic and Analytical Load Ratings

Load rating is an assessment tool that compares capacity of the structure to the demand expected from live and dead load. These factors can be determined simply by engineering judgment, through analytical calculations using AASHTO guidelines, through load testing of the structure, or through utilization of an a priori or calibrated finite element model.

Typical load ratings are calculated using the methodologies and assumptions that were originally developed for new design. These assumptions include single-line girder analysis, where transverse effects are accounted for through distribution factors. Load ratings are often calculated for every member through the use of pre-packaged software products which are used in design as well. When a load rating is found to be inadequate, the options include posting the bridge, performing more refined analysis (e.g. FE model simulation), or performing a load test.

2.3 Load Testing

Load testing is essentially applying a substantial, known load to a structure and measuring responses. These responses can be used in load ratings directly as mentioned above, or for the calibration of a finite element model. Chajes and Shenton III (2005), Yost, Schulz et al. (2005) and Nielsen and McDowell (2006) are just three examples of the numerous applications found in the literature of where load testing was used to inform load ratings.

The AASHTO Manual for Bridge Evaluation (AASHTO 2008) provides an overview of all of the major aspects of conducting a load test. A broad range of test types, instrumentation types, and guidelines are discussed. Due to the nature of the document (meaning the need for widespread applicability) the discussion is non-specific, which allows for the guidelines provided to be applied to the majority of bridges. While this is a benefit in terms of applicability, the level of detail included does not provide adequate guidance to the practicing engineer interested in conducting a load test.

According to AASHTO guidelines, load testing can occur at two levels; diagnostic and proof levels. Diagnostic load level is a service load state, with a single loaded truck providing the static excitation. Proof level loading requires much higher load levels which often exceed the worst case static load level expected on the bridge. Naturally it can be inferred that diagnostic load testing requires extrapolation to higher load levels and proof-level load testing does not. Extrapolation is founded on the underlying assumption that the structure does not “change” as load increases (i.e., that the structure is linear). This is not the case for many structures.

In addition, the AASHTO Manual provides only minimal guidance on both the selection and use of an appropriate simulation models. For example, the entire approach is predicated on the identification of “critical members”. The reality however, is that given the significant uncertainties associated with constructed systems such as bridges, it is not possible to reliably identify the actual critical members in an a priori sense. It is likely that the actual critical members may not be identified at all through these procedures, and so comprehensive instrumentation and a rigorous approach to model-experimental correlation must be carried out if reliable results are to be expected. Unfortunately the AASHTO Manual completely ignores this reality and sets up a load rating framework that is likely to produce unreliable and potentially unconservative estimates of capacity. In addition, despite the numerical results, load rating procedures contain arbitrary condition factors based on visual appearance and engineering judgment, and therefore essentially remain subjective in nature even though many consider them to be objective.

2.4 Structural Identification

Structural Identification (St-Id), the state of the art in condition assessment, builds on the tools above but incorporates a rigorous correlation of the model and experiment. This process is a specific application of system identification which was first applied to civil structures by Liu and Yao (1978). Since then, there have been many applications and advancements (Aktan, Farhey et al. 1997; Aktan, Catbas et al. 1998; Aktan, Beck et al. 2000; Gentile 2006; Brownjohn, Fujino et al. 2008; Catbas, Moon et al. 2008; Morassi and Tonon 2008; He, Moaveni et al. 2009; Moon, Aktan et al. 2009), and St-Id remains a

very active area of research. The typical six step process, as described in Aktan, Farhey et al. (1997) was shown in Figure 1-1.

The process includes conceptualizing the structure, developing a priori models for response prediction, experimentation and data reduction, correlation and updating of the model using experimental data, and conducting simulations to predict unmeasurable quantities using the calibrated model.

2.4.1 Step 1: Observation and Conceptualization

The first step of St-Id, observation and conceptualization of the structure, is often given the least attention in a project and is potentially the most important. The overarching goal of this step is to maximize the potential for success by recognizing, estimating and minimizing the uncertainty surrounding the structure. Specifically this means assessing and estimating both the need and the limitations of St-Id for the given structure in addition to the potential benefits. To a large degree the success of this assessment is dependent on visual inspection, both by the regular engineer and the persons responsible for conducting the test. Conceptualization includes but is not limited to understanding loading, kinematics, load path, site constraints and the associated vulnerabilities of the structure. These elements are particularly critical for model development and can derail an entire St-Id effort if ignored.

2.4.2 Step 2: A Priori Modeling

The amount of information gained from an a priori model for a St-Id effort is directly tied to the type of model. Moon (2006) presents two classes of models for St-Id; physics-based and non-physics-based models. Physics-based models attempt to retain the structural engineering heuristics and intuition by explicitly modeling boundary and continuity conditions, equilibrium and kinematics while non-physics based models are concerned primarily with data and patterns within that data (Catbas, Aktan et al. 2009). This research focuses on physics-based models, and in particular, finite element (FE) models. The level of detail and depth of an FE model can vary greatly.

Phenomenological models are very basic representations of the structure meant to smear individual bridge element properties in order to obtain global characteristics like stiffness. This model form is not overly common in St-Id applications. Beck (2004) utilizes a phenomenological model of a ten-story shear building with localized springs and dampers. Similarly, Robert-Nicoud, Raphael et al. (2005) uses phenomenological models of a cantilever beam with various configurations of springs as boundary conditions to model bridge behavior. Catbas, Aktan et al. (2009) give an example of a phenomenological model of a suspension bridge which uses very simple elements.

On the other end of the spectrum are geometric replica models. These models are very in-depth, with little to no simplification of the geometry of the structure. Typically shell and solid elements are used, as these are capable of better representing the actual geometry of the structure. These models are not frequently used for St-Id because they are time consuming to build and error screen and are often very computationally intensive. An

example of a geometric replica model of a floor system for one panel of a steel truss bridge is presented in Catbas, Aktan et al. (2009).

Element-level models are the most common type of FE model used in St-Id. In an element-level model, linear elements are used to represent members like beams and braces, and 2D area elements are used to represent planes like bridge decks. The actual section properties are assigned to these elements and a global stiffness matrix can be calculated. Compatibility between elements is ensured mathematically in the model with various connections or constraints. Examples of element-level models are prevalent in the literature (Mottershead and Friswell 1993; Aktan, Farhey et al. 1997; Aktan, Catbas et al. 1998; Moller and Friberg 1998; Brownjohn and Xia 2000; Brownjohn, Xia et al. 2001; Friswell, Mottershead et al. 2001; Zhang, Chang et al. 2001; Teughels 2004; Jaishi and Ren 2005; Daniell and Macdonald 2007; Pan 2007; Zarate and Caicedo 2008; Catbas, Aktan et al. 2009). This research will utilize element-level models.

The main purpose of developing a mathematical or finite element model prior to experimentation is to provide response estimates. These predictions can directly be used to inform experiment and instrumentation design. For example, it is common for the a priori model from a load rating to be used to determine critical members which are the focus of local instrumentation for the purpose of directly updating the rating factor. Similarly, the required load level for a test may be estimated from the a priori model as well.

This process forces the engineer to develop an in-depth understanding of the structure, including the crucial building blocks for model construction. The amount of information

and insight that is gained during a priori model development is difficult to quantify but is certainly beneficial for the remainder of the St-Id process. On the other hand, the inherent danger to utilization of an a priori model is the fundamental assumption that the model is essentially correct and error free.

2.4.3 Step 3: Experimentation

Experimentation as part of St-Id can be subdivided into two parts; instrumentation/experiment design and experiment implementation, both of which are subject to the test constraints determined during Step 1.

The main goal of instrumentation and experiment design is to satisfy all of the test objectives while maintaining situational awareness and safety. This requires clear definition of the test objectives, safety concerns and anything else that needs to be accounted for directly through instrumentation. Typically an instrumentation plan is designed to either capture the global behavior or to understand the response of critical members. Design of the experiment includes load level, position and the number of cases considered as the key parameters around which an entire test protocol is built.

Implementation includes installation procedures, data acquisition protocols, data quality assurance, situational awareness and documentation of the experiment for future reference.

Any type of experiment could fit into the St-Id framework, but there are a few that are most prevalent. Static testing consists of applying a known load, generally with a truck,

to the bridge for an extended period of time (i.e., enough time for measurements to stabilize) and recording responses. Often more than one load level and position are used to check linearity of the structure. There are numerous examples of static testing available (Aktan, Farhey et al. 1997; Fang, Chen et al. 2004; Lucas, Virginia Transportation Research Council. et al. 2004; Krzmarzick and Hajjar 2006; Weidner, Prader et al. 2009).

Crawl testing uses the same load levels as static testing, but the trucks are continually moving at a slow speed to eliminate any dynamic effects and simply the estimation of truck position. These tests are very common for informing load ratings due to their rapid nature. Some examples of crawl tests include Chajes and Shenton III (2005), Shenton III and Chajes (2009), and more recently, Sanayei, Phelps et al. (2012).

Ambient vibration testing consists of monitoring vibration response (either acceleration or velocity) from traffic excitation, and extracting modal properties. Clearly in this scenario, the input is unknown. This makes processing of the data more difficult, but greatly simplifies the testing. Using ambient traffic as input minimizes the impact on traffic. For long span bridges, exciting the structure with live load is difficult in a static sense, meaning that ambient monitoring is often the only option. Ambient vibration testing is extremely common, with numerous examples in the literature (Brownjohn, Zasso et al. 1995; Doebling, Farrar et al. 1996; Brownjohn and Xia 2000; Catbas and Aktan 2000; Cunha, Caetano et al. 2001; Brownjohn, Moyo et al. 2003; Teughels 2004; Jaishi and Ren 2005; Grimmelmann 2006; Cigada, Caprioli et al. 2008; Schlune, Plos et al. 2009; Zhang, Prader et al. 2009; Bayraktar, Altunişik et al. 2010)

Forced vibration is the final common experimental method used for St-Id. In this case, the input is known. Often this requires more access and traffic control, but the results are correspondingly more reliable. There are some applications of forced vibration for highway bridges using shakers (Morassi and Tonon 2008) and using impact hammers and drop weights (Prader 2012).

2.4.4 Step 4: Data Reduction

The fourth step of St-Id is the reduction and interpretation of the experimental data. The data is error screened and improved wherever possible, and then compressed into a readily useable format. Much like instrumentation design, post-test data reduction is a very user-dependent process.

For a static load test this data compression includes reducing time histories into average values and plotting versus varying load levels and spatial distributions. Data is filtered to remove sensor noise prior to averaging. Offsets and linear trends, like drifting due to the increase in the temperature of a strain gage as current flows through the resistance element, is removed as well.

For dynamic data, the reduction is more complex. The data is first preprocessed which essentially accounts for the numerical issues that arise from conversion of an analog signal to a digital signal. Preprocessing includes averaging of signals, windowing, and filtering. The reduced data is used to develop Frequency Response Functions (FRFs) or something similar depending on the testing technique. From these FRFs, modal parameters (i.e., frequencies and mode shapes) are extracted using one of several

common model parameter estimation algorithms. A comprehensive discussion of preprocessing and modal estimation is provided in Prader (2012).

2.4.5 Step 5: Model-Experiment Correlation

Correlation of the model with the experimental results is the step that truly separates St-Id from typical state of the practice condition assessment. Typically in practice, the model is correlated to the experimental data by manually changing various parameters and comparing the results with the experimental data until a sufficient agreement is achieved (Brownjohn, Xia et al. 2001; Daniell and Macdonald 2007; Pan 2007; Schlune, Plos et al. 2009; Weidner, Prader et al. 2009). A more detailed discussion of manual updating is presented in Section 2.5.2.1).

There has been a great deal of attention paid to more automated approaches to parameter identification over the last 30 years (Moller and Friberg 1998; Xia and Brownjohn 2004; Jaishi and Ren 2005; Teughels and De Roeck 2005; Sanaye, Bell et al. 2006; Pan 2007; Bayraktar, Altunişik et al. 2009). A more comprehensive discussion of automated model updating is presented in Section 2.5.2.2.

Regardless of the approach selected, the inherent assumption employed is that all of the uncertainty associated with the model can be attributed to the parameters (or can be effectively compensated for by changing the parameter values). As will be discussed later, there is evidence that this assumption is unfounded and that a more rigorous approach to ensure the appropriateness of the a priori model is needed.

2.4.6 Step 6: Simulation for Bridge Management

This final step involves the use of the correlated model to better understand the structure and to inform decision making through scenario analysis. This step closes the loop between visual observation, quantitative response, and utilization for bridge management by allowing simulation of retrofits, permit loads, extreme events, and other cases of interest to the owner.

2.5 Overview – Types of Model-Experiment Correlation

Correlation of a finite element model with the experimental results is a crucial step in the structural identification process which enables the engineer to use the data from the experiment for more than simply presenting loads and responses. Using the experimental data to inform and validate a model removes some uncertainty regarding the model and the structure (Mottershead and Friswell 1993). There are numerous methods of model-experiment correlation, described in the following sections.

2.5.1 Direct Updating

Direct updating consists of manually manipulating the individual terms of the global stiffness matrix to force the model to represent the experimental data accurately (Friswell, Inman et al. 1998). This method originated before sophisticated finite element modeling software was available. For this reason, direct updating lends itself to less complex modeling. While this method will often lead to a model which accurately represents the measured behaviors, changing the stiffness matrix directly can lead to a

distortion of the physical reality of the structure of interest. Even if connectivity is maintained (through constraining the updating process) it is very different to identify the root causes and the plausibility of the optimized influence coefficients since these are composed of numerous parameters. Due to these shortcomings, the use of, and research into, direct updating methods has largely died out and it was not considered in this study.

2.5.2 Iterative Methods

While most model-experiment correlation is a repetitive process, two methods in particular are considered iterative in nature. These are manual calibration using heuristic knowledge, and parameter identification, which is an automated approach that uses various optimization algorithms to adjust parameters. While these two methods are described separately here, it is difficult in reality to divorce them. The metric for evaluating a manual calibration is often the same as parameter identification. Similarly, parameter ID is useless without sound heuristics as this is needed to select the appropriate parameters and their spatial distribution. The goal of iterative model-experiment correlation is to develop a single, calibrated model which represents the experimental data well and can be used for extrapolation.

2.5.2.1 Heuristic Calibration (Manual)

Heuristic calibration consists of selecting and updating parameters in a finite element model based on observations and knowledge of both the particular structure and bridges in general. Examples of manual model calibration are readily available, but are often

coupled with subsequent automated updating (Brownjohn, Xia et al. 2001; Daniell and Macdonald 2007; Pan 2007; Schlune, Plos et al. 2009; Weidner, Prader et al. 2009).

Heuristic calibration allows for simple implementation of changes in both magnitude and spatial variation of parameters. For example, an engineer may desire changing the modulus of the concrete deck in terms of the magnitude of the modulus and the location on the structure to represent localized cracking or delamination. The model is then rerun and the results are compared to the experimental observations again. While each iteration may be accompanied with the implicit or explicit evaluation of an objective function of some kind, the process is contingent on the input of the engineer (i.e., the engineer is directly in the loop). The parameters to be changed, how to change them both in magnitude and spatial variation, and the decision of when the model is considered calibrated all lie with the engineer.

2.5.2.2 Parameter Identification (Automated)

Parameter identification (ID) uses automated algorithms to select new parameter values. In this case, the user input again defines the parameters to be varied, as well as criteria for optimization like starting values, stopping points, step sizes and objective functions. Teughels (2003) provides a comprehensive discussion of parameter ID and model updating. Applications include Moller and Friberg (1998); Xia and Brownjohn (2004); Jaishi and Ren (2005); Teughels and De Roeck (2005); Sanayei, Bell et al. (2006); Bell, Sanayei et al. (2007); Bayraktar, Altunişik et al. (2009) and (Pan 2007).

The objective function describes the ability of the model to represent the actual system in quantitative terms. Typically the objective function will be a normalized error between the measured responses and the model responses, though other objective functions are available. The value of the objective function (or the change in value between steps) serves as the typical stopping criteria. Inherent in all automated parameter ID using gradient-based algorithms is the danger of missing the global minimum value of the objective function or optimizing to local minima. This can be addressed using optimization algorithms that are not gradient-based, though this is not directly implemented in this effort.

2.5.3 Multiple Model Methods

It is an established fact that model updating can lead to an unfavorable situation where more than one model can equally represent the experimental data. This problem of non-uniqueness of model updating is nearly unavoidable as the model is generally be far more detailed than the experimental information (Janter and Sas 1990).

Over the last 10 years, there has been some localized development of model-experiment correlation methodologies that explicitly recognize the uncertainty associated with a single model by incorporating, in one form or another, multiple models. The driving premise behind all multiple model methods is that the decision of the engineer to utilize either a particular model form or a particular set of parameters is somewhat arbitrary in an a priori sense, and that it therefore makes sense to attempt to evaluate these decisions prior to using the model to inform decision-making regarding the structure. While the related field of model selection is far more mature, it is distinct from multiple model

methods in that it still aims to identify a single model form as opposed to directly incorporating multiple model forms within the correlation process. There are several distinct approaches that utilize multiple models, described below.

2.5.3.1 Error Threshold

The Error Threshold approach has matured organically through the work of a research group led by Professor Ian Smith at the Swiss Federal Institute of Technology (EPFL). In this approach a candidate model population is developed based on certain criteria, particularly the ability of the model to predict measured responses. The candidate models vary typically in their parameter values, but can also vary due to their model form, element type used, simulation software, etc.

Once the population is developed, there are three potential outcomes; the model is within the prescribed error bounds and is kept, the model is removed from the population because it exceeds the bounds, or it is decided that more measurements must be recorded to distinguish between the models remaining in the population. Initially the population of models was developed using heuristic engineering knowledge about the structure (Robert-Nicoud, Raphael et al. 2005). Explicit accounting for all error in the modeling and experiment became the basis for the error thresholds which are used to distinguish between models (Robert-Nicoud, Raphael et al. 2005). They specifically quantify errors from three sources; 1) discrepancy between the model and the measured responses, 2) numerical error during computation, and 3) modeling errors including incorrect parameters and incorrect modeling assumptions (Raphael and Smith 2003). The model population was established through a non-gradient based algorithm (performance global

search Lausanne or PGSL), which is capable of performing without fixed constraints on the number of parameters (Smith and Saitta 2008). Manipulation of the model population was improved through data mining techniques like k-means and principal component analysis (Saitta 2008; Saitta, Kripakaran et al. 2008). Most recently, the error threshold method is being utilized to inform instrumentation design (Goulet, Kripakaran et al. 2009).

A principal criticism of this method is that the carefully crafted error bounds are both aggressive and somewhat arbitrary, and they often eliminate all models from the candidate population. Typically the solution is then to increase the bounds until there are models remaining in the population, which directly calls into question the ability to determine specific error bounds in the first place (Dubbs 2012).

2.5.3.2 Modeling to Generate Alternatives

Zarate and Caicedo (2008) implement a Modeling to Generate Alternatives (MGA) approach on a simple numerical problem, and on an existing cable-stayed bridge. The MGA approach, first applied in the field of land use planning (Brill, Chang et al. 1982), aims to develop a population of good solutions (local minima of an objective function space) from which the engineer can select one or more models based on heuristics and intuition.

To develop the candidate model population, first a global minimum is located through an objective function which compares experimental data with analytical model output. The optimization of this objective function is completed with a gradient-based algorithm

called sequential quadratic programming. From this point, a second objective function is optimized using this same algorithm. This objective function looks for secondary solutions which meet a constraint based on the magnitude of the total objective function of the global solution. The second optimization is started from several locations until no additional solutions are found.

The population is then reduced based on the experience and knowledge of the engineer. No rigorous criterion for model selection is provided. They explicitly state that depending on the application a single model or multiple models can be selected and that multiple models are particularly useful when the end result is a reliability analysis. They do not consider multiple model forms however, only multiple parameter sets within a model form.

2.5.3.3 Bayesian Model Updating

A Bayesian statistical framework was developed for civil structures in Beck and Katafygiotis (1998). This initial application was limited to within a given model class, but was explicitly capable of addressing the non-uniqueness problem associated with model updating and parameter identification. Several output-identical models are found, and their respective unobserved response predictions were compared and found to be different, necessitating multiple models. Note that in this early Bayesian framework, asymptotic approaches were used to estimate the posterior probability. Later these asymptotic approaches were replaced with a Markov Chain Monte Carlo (MCMC) approach for posterior estimation (Beck and Au 2002). The method was extended to serve as a model selection tool by accounting for variation of the degree of

parameterization between model forms (Beck 2004). Most recently, a hybrid MCMC approach is presented as a solution for high dimension problems where convergence is normally prohibitively slow (Cheung 2009).

Taking MCMC one step further, Reversible Jump MCMC or RJMCMC, accounts for uncertainty in both the parameters and the model form itself. The method was first presented by Green (1995) as a model selection tool. An MCMC chain is developed that can transition between model forms using jump equations which transform parameters between models. RJMCMC has seen extensive application in other statistics (Green 2002; Papathomas, Dellaportas et al. 2011; Scheel, Green et al. 2011), but has not until recently been applied to civil structures (Dubbs 2012).

The problem of model selection when the number of parameters is unknown, which is directly addressed using RJMCMC, has been addressed through several other techniques that are not as prevalent in the literature. One example is Metropolis-coupled MCMC or MCMCMC, which uses multiple simultaneous MCMC chains for the same model with different proposal functions (Geyer 1991). Moves are made between chains with the traditional MH acceptance criterion, which allows the overall chain to mix rapidly, at the potential expense of extra iterations given the need for several chains.

Carlin and Chib (1995) present a method using a joint conditional distribution across multiple model forms where the prior probabilities are actually considered pseudo-priors and are contingent upon a model indicator function. Developed around the same time as RJMCMC, these two methods were in direct competition, and most subsequent research was related to RJMCMC.

In a similar vein, Population MCMC (Fruhwirth-Schnatter 2001) is an extension of mixture models that moves between different model forms by using a composite posterior space, similar to Carlin and Chib (1995).

Sisson (2005) provides an overview of the development of trans-dimensional MCMC methods, including RJMCMC, MCMCMC and the pseudo-prior method previously discussed along with several other different approaches.

One constant throughout these Bayesian applications is that the applications are based on numerical examples, not constructed systems.

2.5.3.4 Multiple Model Structural Identification

Recently the concept of St-Id presented earlier in this chapter was extended to account for non-uniqueness through multiple models with a focus on response predictions (Dubbs 2012). This method attempts to explore and understand the model space in order to produce reliable, defined predictions of unmeasurable quantities. Specifically, investigations into sampling methodologies, weighting schemes, and model form selection were presented.

MCMC sampling was found to be the most efficient method for exploring the parameter space within a given model form. The method was applied to a long span steel truss bridge using ambient vibration data. The resulting model chain was first used to develop predictions of measured responses that were not included in the updating process. It was found that these measurements were contained in the prediction distributions. Next load

rating factors were calculated and finally several specific scenarios of interest (i.e. ship impact) were presented.

To investigate model selection, RJMCMC was investigated and applied to a laboratory grid model. The method was shown to produce response predictions based on multiple model forms which predicted measured responses with reasonable accuracy and allowed for comparison of the effectiveness of the model forms. RJMCMC was not applied to the steel truss bridge and has not been applied to any real structure to date.

A more detailed discussion of multiple model-experiment correlation is included in Chapter 8.

2.6 Summary

This chapter first discussed the state of the art in bridge condition assessment. This essentially amounts to visual inspection and load rating. Occasionally experiments are conducted in order to gain quantitative insight into the condition and performance of the structure. This fits naturally into a framework known as Structural Identification, or St-Id. St-Id is a well-known and accepted concept, but there is a lack of comprehensive, well-documented applications for real structures.

A crucial aspect of St-Id is the model-experiment correlation phase, in which the selected finite element model is updated using the experimental results to provide better correlation with the experiment. Some model-experiment correlation methods use a single model, while others use multiple models. To date, there has not been any

comparison of the benefits and drawbacks of these two types of applications, partially because there have not been many multiple model applications at all.

In the existing multiple model applications in the area of constructed systems, typically a single model form is chosen, and then the parameters within that form are varied to create multiple “models,” or multiple parameter sets. These methods have not been extended to account for and include multiple model forms, parameter sets, and even spatial distributions of parameters, particularly with a focus on predictions of unmeasureable quantities, which is the last step of a St-Id effort.

Chapter 3: Observation and Conceptualization

This chapter presents the initial steps in a St-Id effort. The candidate bridge is evaluated in terms of feasibility, relation to the overall bridge population and potential return on investment. The structure and site are analyzed considering access, testing and modeling and a set of critical questions regarding the structure are developed.

3.1 Background of the International Bridge Study

The Federal Highway Administration's (FHWA) Long-term Bridge Performance (LTBP) Program is a major strategic initiative, developed by the Office of Infrastructure Research and Development within FHWA, and designed to address the growing concern with aged and deteriorated infrastructure throughout the U.S. The project aims to develop a more accurate and timely picture of bridge health, to improve knowledge of bridge performance, and to promote the safety, longevity and reliability of the United States highway transportation system. As part of this flagship program, the FHWA launched an International Bridge Study (IBS) with the goal of establishing the worldwide "best practices" for the integration and application of technology to diagnose, perform prognosis, and design treatments to mitigate performance deficiencies for a given bridge. The general approach adopted for this study was to invite experts from around the world to demonstrate the 'best practices' associated with specific technologies on the same bridge. This type of round robin study not only allows for a complete documentation in the context of a real-world application, but it also permits the direct comparison between several different methods. The first round of applications focused primarily on short-term structural testing within the structural identification (St-Id) paradigm.

3.2 Selection of the IBS Test Bridge

In order to properly assess the value of various technologies to help inform decisions regarding the renewal, preservation or replacement of existing bridges, it was crucial to select a bridge that displayed common performance concerns whose root causes and effective mitigation approaches were unknown. It is important to recognize that the goal was not to select the poorest performing bridge (where replacement is probably the only feasible option), but rather to select a bridge with common, significant performance problems that are difficult to address. To accomplish this, the project team worked closely with the New Jersey Department of Transportation (NJDOT) staff to identify a series of candidate bridges.

In all a suite of 15 bridges was identified by the NJDOT as bridges for which there was no clear path forward (i.e. they met the primary goal above). The secondary criteria for the test bridge selection included (1) the significance and number of performance problems, (2) the commonality of the bridge type/form and of the performance problems, (3) the availability of documentation, (4) the significance of load rating and inspection challenges, and (5) the ease of access. To assess these issues for each of the candidate bridges the project team reviewed past inspection reports, examined their relation to the overall bridge population within NJ (through the LTBP Project Bridge Portal), and performed site visits.

Following this criteria sister multi-girder steel stringer bridges that carry US 202/NJ 23 through Wayne, NJ (approximately 30 minutes outside of New York City) were selected (Figure 3-1). These structures were constructed in 1983 and 1984 and currently display very common problems associated with approach settlement, bearing alignment/walking,

substantial vibrations and fatigue cracking. In addition, these structures have a variety of skew, partial skew and straight spans that traverse a park, which allows for unrestricted underside access. A more complete description of these structures is provided in Section 3.3.

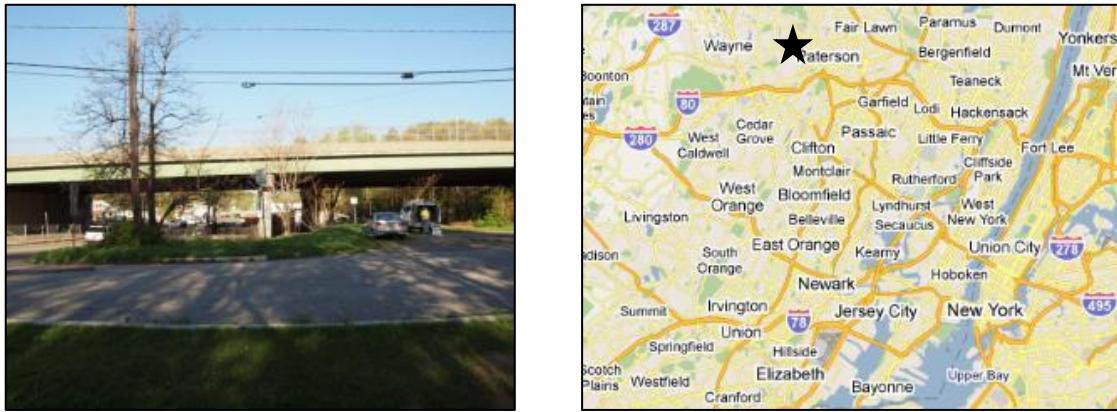


Figure 3-1 - US202/NJ23 International Test Specimen in Wayne, NJ

While in this case the test bridge was selected to best serve the goals of the IBS study, it does represent a very common approach to selecting structures for technology applications. That is, owners are not necessarily interested in testing their “worst” bridge, but rather a bridge that displays significant performance concerns for which the appropriate path forward is unclear. The other, and perhaps more important approach to selecting bridges for technology applications, occurs when a sizeable population of bridges have similar performance problems or there is significant uncertainty as to the best preservation/renewal practices for a population of bridges. In this case, a sample of bridges from the population in question may be selected with the goal of generating results that may be extrapolated and applied to the entire population. While such

applications have the potential to pay huge dividends, there are very few entities with the knowledge and ability to select appropriate sample populations (both size and distribution) and then extract generic information from this sample which may be applied to the overall population.

As discussed previously the observation and conceptualization step is of paramount importance as it drives all of the modeling and experimental activities throughout the St-Id. To perform this step for the US 202/NJ 23 Bridge (Figure 3-1) the project team obtained relevant documentation (including design drawings and the previous two inspection reports), mined NBI data to place the bridge in context, performed a preliminary ambient vibration test, and computed nominal rating factors. This all of led to a qualitative risk assessment and ultimately to the development of a series of critical questions to guide the entire St-Id. The details of these steps are provided in the following subsections.

3.3 Overview of the Structure (Review of Design Drawings)

Each of the test bridges (constructed in 1983 and 1984) are comprised of four simply supported spans using a standard steel stringer design of eight girders with variable section properties and geometries including straight, partially skewed and fully skewed spans. From south to north, the bridges span Mountainview Blvd, an open field, train tracks, and an exit ramp. Due to constraints imposed by the railroad, no underside access to Span 3 was provided to the project team at any time during this study. An overview of the general bridge statistics is provided in Table 3-1.

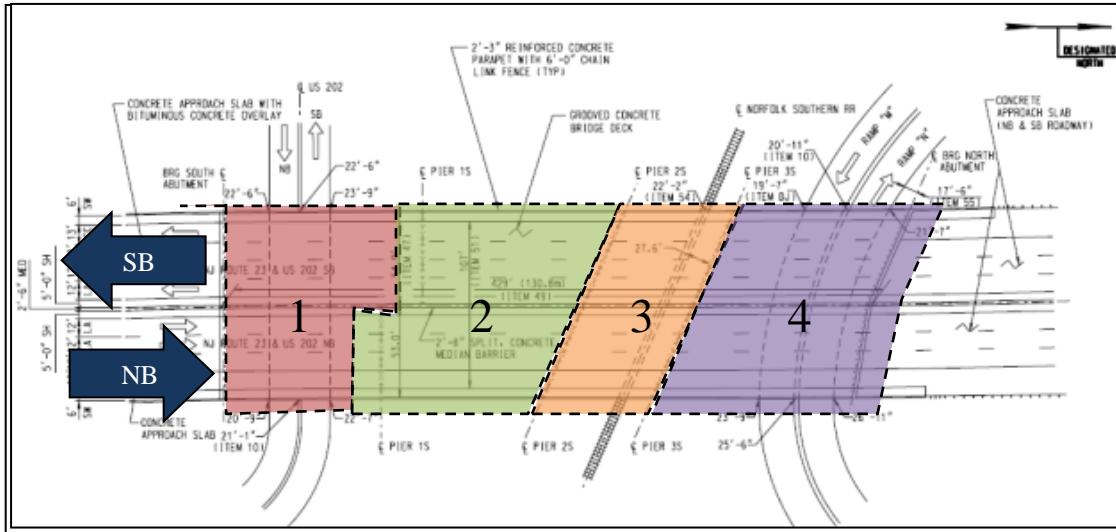


Figure 3-2 - Plan View of Bridge

Table 3-1 - General Bridge Information

The girders of both bridges are built up members with variable flange thicknesses. The change in flange thickness is a smooth, well-detailed transition that should not cause fatigue issues. The flange thickness varies from 1 in. to 2.5 in. depending on the girder length, with the top flange transitioning once and the bottom flange transitioning twice. This results in up to five different cross-sections on a given girder and adds to the overall complexity and irregularity caused by the varying skew conditions.

The decks of the two bridges were cast using stay-in-place forms which prevent any visual assessment of the condition of the concrete from the underside of the structure. Shear studs were provided in groups of two to three spaced at 15 in. to 21 in. along the girder, depending on the span and location. According to the design drawings, the studs are all 6 in. tall with a $\frac{3}{4}$ in. diameter.

The bridges contain diagonal wind braces (Figure 3-3) between the fascia and first interior girders on every span which are connected via a ‘Category E’ gusset-to-girder web detail, discussed in Section 3.4. The diaphragms are a standard truss-type composed of four single angles connected to the girders with bolted connections and gusset plates (Figure 3-4).



Figure 3-3 - Typical Wind Bracing for Fascia Girders

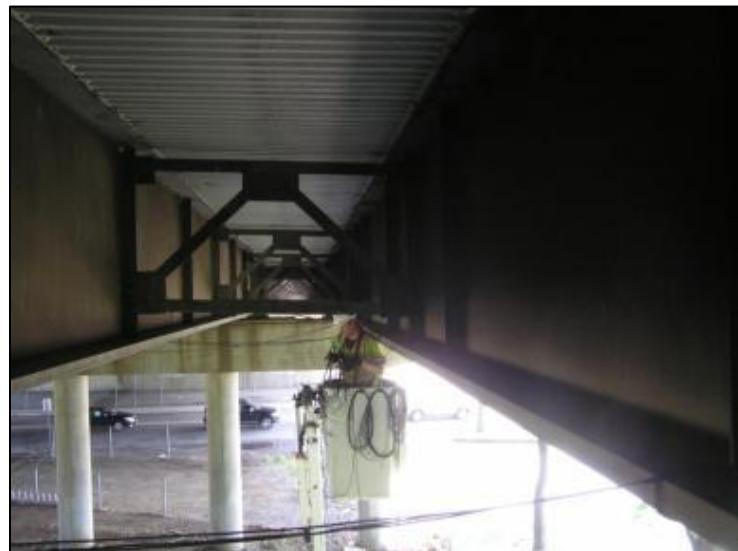


Figure 3-4 - Typical Diaphragms

Each span is simply supported, with pin and rocker bearings and preformed elastomeric compression seal joints.

3.4 Overview of Current Condition (Review of Inspection Reports)

According to the most recent inspection report available, the structures had an overall rating of 5 (Fair) due mainly to the condition of the superstructure. The inspection report pays particular attention to a series of fatigue cracks, which are present at the location of wind bracing-to-girder connections. These ‘Category E’ fatigue details, shown in Figure 3-5, have resulted in a total of 86 fatigue cracks, of which over 40 have propagated into the web. These 40 cracks were arrested by drilling a hole at the tip of the crack and inserting a bolt in the hole. Through the use of dye penetrant testing, it was determined that the arrested cracks have not progressed past where the hole was drilled.

In addition to the fatigue issues, the report notes bearing and joint deterioration and mentions a heavy vibration of the bridge under traffic loads. According to the inspection report, the bridge has an Average Daily Traffic (ADT) of 93,400 with 4% being truck traffic.

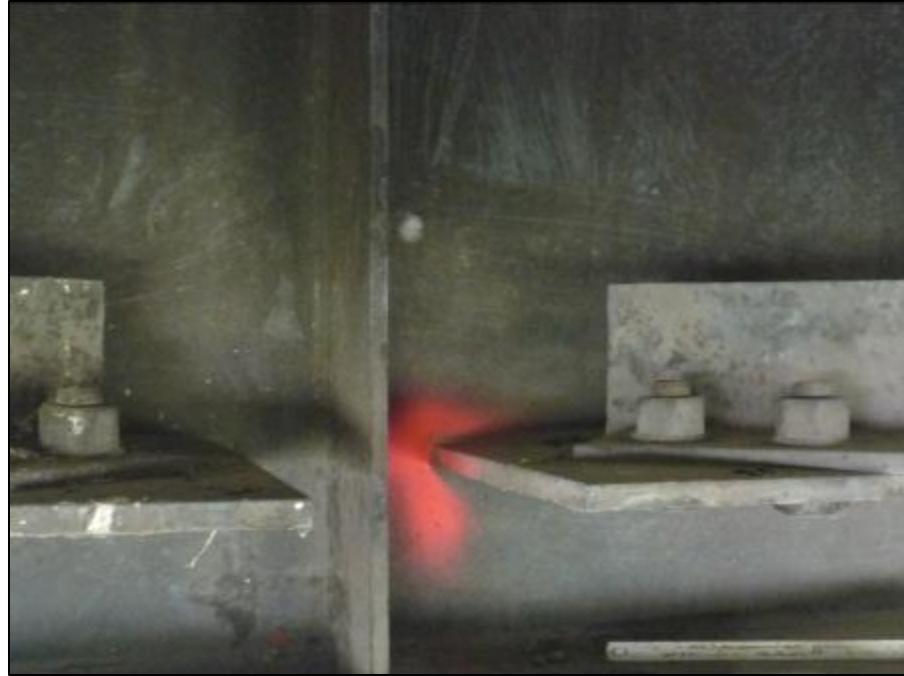


Figure 3-5 - Wind Brace Connection

3.5 Relation to Overall Bridge Population

To assess the relation between the test bridges and the overall bridge population with the state of NJ, the Nation Bridge Inventory (NBI) was mined using the LTBP Program's Bridge Portal. The Bridge Portal is a web-based data mining and visualization tool that allows easy access to the complete NBI (along with other data sources) and as such represents a powerful tool for assessing how a specific bridge fits within the overall population. To begin this process, the overall population is structured into a series of sub-populations based on parameters associated with structural form (referred to in NBI as "main design") and material (referred to in NBI as "main material"). It should be mentioned however that in the case of the material parameter the NBI includes

'continuous' as a possible modifier which has more to do with structural form than material. In any case, the sub-population that the test bridges fall within is simply-supported, steel stringer bridges (Figure 3-6), which represents a large portion (42%) of the total number of bridges within N.J.

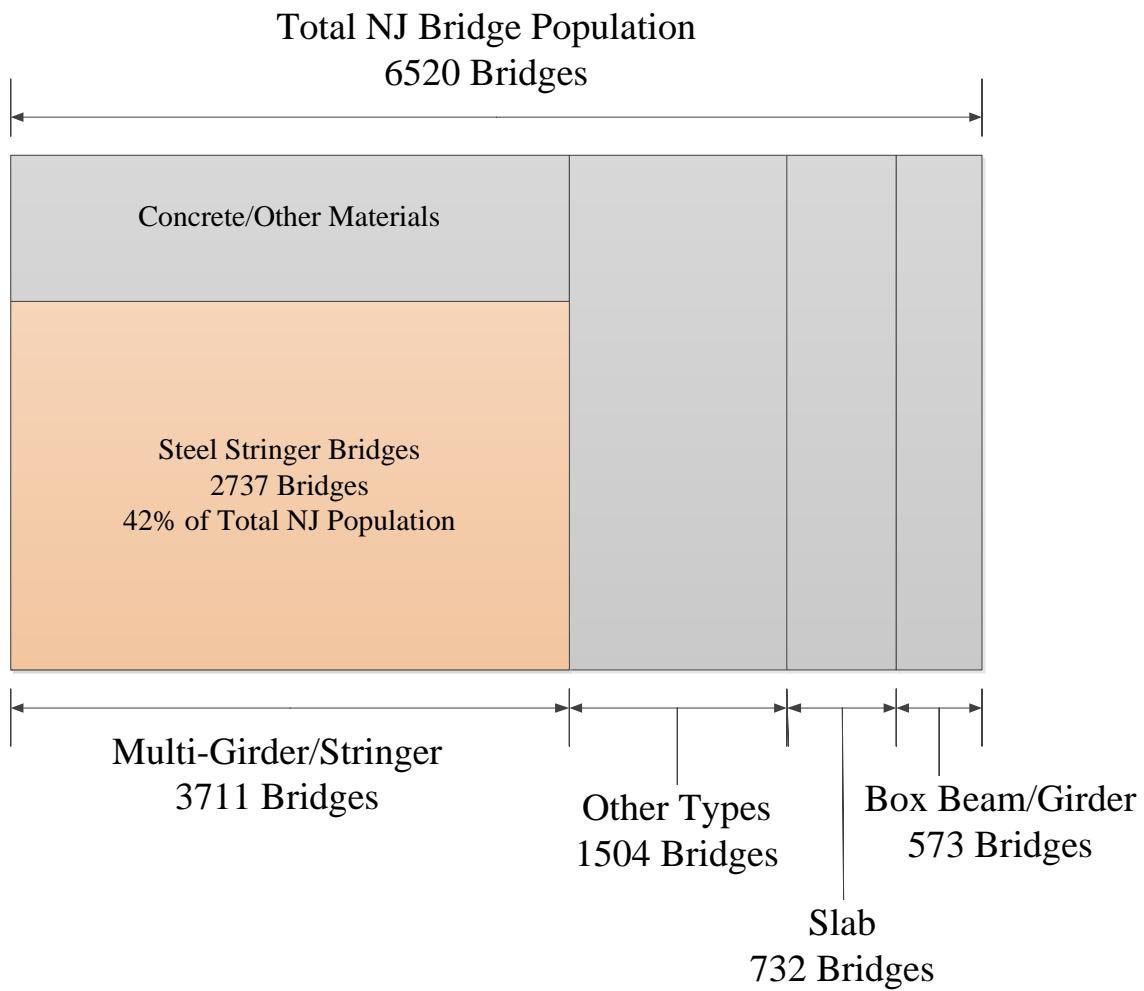


Figure 3-6 - Sub-population of the test bridges (All % of Total Population)

The second step of this process is to examine the distribution of key parameters within the relevant sub-population. This is accomplished by hypothesizing (based on heuristics) which parameters are expected to have the greatest influence on performance. One practical constraint on this step is that all influential characteristics must be represented in the NBI database (or another available database such as PONTIS). For the test bridges the parameters of age and maximum span length were selected, and the number of bridges in the various “bins” chosen for these parameters is shown in Table 3-2.

Table 3-2 - Distribution of the sub-population based on age and maximum span length

		Maximum Span Length (ft)			Total
		L < 60	61 < L < 150	151 < L	
Age (yr)	0 < Y < 20	31	97	14	142
	21 < Y < 40	60	680	49	789
	41 < Y < 60	309	775	19	1103
	61 < Y < 80	324	87	0	411
	81 < Y	269	23	0	292
	Total	993	1662	82	2737

As shown in bold in Table 3-2, the ‘family’ of the test bridges contains 680 bridges or just over 10% of the total bridge population within N.J. As a result, it appears that the test bridges belong to a fairly large family of bridges (as defined by form, material, age, and

span length). It is illustrative at this point to drill down and identify the ‘demographics’ of the bridge family to understand how the bridges of interest relate. It is useful to summarize this effort as a series of pie charts as shown in Figure 3-7 (with the test bridges shown as orange stars). Of course it is possible (and likely) that some parameters may have coupling effects, i.e., ADTT may become a more important parameter for older bridges. As a result, the “one-dimensional” representation shown may not suffice for all analyses.

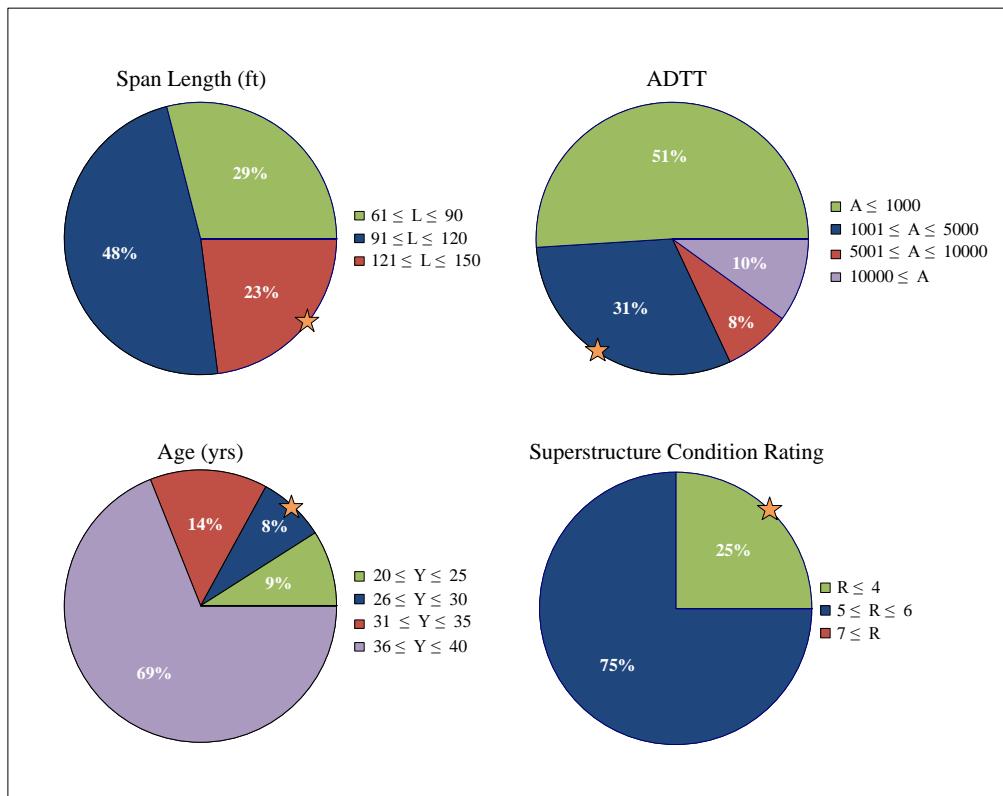


Figure 3-7 Distributions of the family of bridges based on (a) span length, (b) ADTT, (c) age, and (d) deck condition

Examining the pie charts shown in Figure 3-7, several observations about the test bridges can be made. First, while the bridges are quite young compared to the remainder of the family, they appear to be under-performing (based on superstructure rating). This is not surprising as these bridges were identified by NJDOT as outliers. In addition, it is observed that they have rather long span lengths and have a moderate to high ADTT compared to the rest of the family. While such information cannot be used to identify the root causes of the performance problems, it does place the bridges in context and may indicate how the results of the St-Id may be applied to other structures.

3.6 On-site Observations (Initial Field Visit)

During the first site visit, several important observations were made. While walking the span, heavy vibration of all spans was observed, though it was heaviest on Southbound Span 1 (SB1) and Southbound Span 2 (SB2). There was also transverse cracking of the top of the bridge deck at regular intervals visible on SB1. Otherwise the deck condition appeared satisfactory when viewed from above. However, without being able to see the deck from underneath, it is difficult to assess the quality of the deck and deterioration that may be limiting the ability of studs to enforce strain compatibility between the girders and deck. With the exception of the fatigue cracks previously discussed, no major performance problems were noted for the primary elements of the superstructure. This notwithstanding, there was one particular diaphragm connection that was missing two of the original three bolts and could be heard rattling under normal traffic loads.

Bearing condition varied significantly across the structure with some bearings exhibiting almost no corrosion or deterioration, while others appear so deteriorated that replacement

may be required. In particular, the fascia bearings showed extensive corrosion, and in one case a fully cracked pintel was observed (Figure 3-8) that allowed the bearing to rotate about its vertical axis.



Figure 3-8 - Fully Cracked Pintel

In addition, the majority of the joints were either filled with debris or bulging up into the roadway. Examination of the piercaps showed that most joints allowed water to drain directly through on top of the pier. In addition, the pier between SB1 and SB2 had a very large vertical crack between the bearings for Girder #1 and Girder #2 (Figure 3-9). There

was also some spalling and cracking visible on the piers, as well as extensive map cracking on the abutment walls.



Figure 3-9 - Crack in SB Pier 1

Also worth noting was the fact that noticeable ground vibrations were felt when standing under Spans 2 and 4. According to the design drawings the piers were founded on timber piles, but it was not possible to visually assess their condition. The approach slabs exhibited some settlement directly adjacent to the joint of the exterior spans of the bridge as well.

3.7 Preliminary Data Collection

Although not always required, it is very useful to perform a preliminary data collection effort during the site visit as this provides some quantitative information regarding bridge performance and can greatly help in developing the a priori model and designing the experimental study. Given the concerns surrounding the vibration of the spans, a preliminary ambient vibration monitoring of the bridge was conducted during one of the site visits. This study consisted of recording ten minutes of data from four accelerometers installed on the sidewalk of each span (so no traffic control was required). While a single setup is possible, this particular study was carried out by moving the four accelerometers from span to span and the total data collection effort took less than two hours.

Figure 3-10 and Figure 3-11 show typical acceleration time histories and power spectral densities for Span 2 northbound and south bound respectively. As apparent from these figures, the bridges experience significant acceleration under normal traffic loading (with 0.4 g the highest acceleration recorded). In addition, while both spans (which are nearly identical) have very well defined peaks (indicating low damping ratios), they display distinctly different natural frequencies (as shown in Table 3-3). In addition, the higher frequency modes of Span 2 northbound appear to be participating more than their counterparts for Span 2 southbound. While there are many potential reasons for this (such as the different locations of the sidewalks relative to the skew, or simply different traffic loading during the monitoring), at this point it is important to at least be aware that the vibration responses of the two bridges may have significant differences.

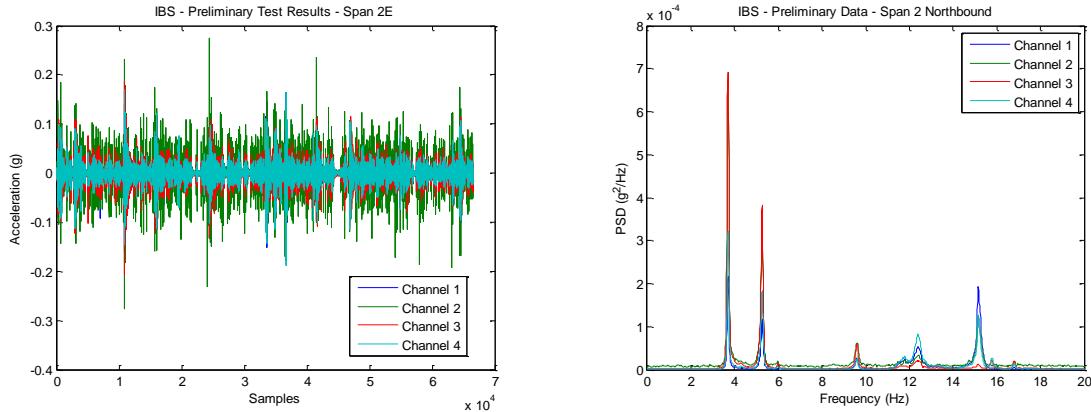


Figure 3-10 - Acceleration Time History and Power Spectral Densities for Span 2 NB

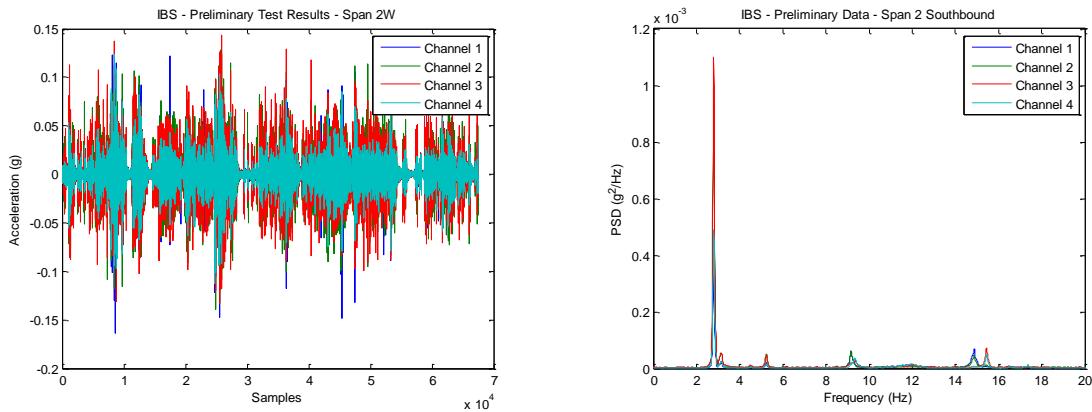


Figure 3-11 - Acceleration Time History and Power Spectral Densities for Span 2 SB

Table 3-3 - Identified Frequencies - Preliminary Data

Span 2 NB	Span 2 SB
3.699	2.808
5.261	3.174
5.981	5.249
9.595	9.167
12.380	14.880
15.150	15.440
20.860	

In addition to providing a quantitative basis for the conceptualization of the apparent vibration problem, this study also served to provide a sound basis for the error screening of the a priori model and the design of the vibration tests. Specifically, these results provide an indication of the bandwidth of interest and the level of ambient vibrations, which are important design parameters for the sensing and data acquisition protocols. In addition, given the large levels of vibration, it was noted that it will be difficult to perform forced vibration testing if the bridge is not completely closed as the excitation systems (impact or shaker) may not provide sufficient energy to overcome the ambient vibration level.

3.8 Preliminary Load Rating

The inspection reports provided by NJDOT include the most critical load rating values calculated using PennDOT's Bar7 Program (PennDOT 2008), which utilizes the Load Factor Rating (LFR) approach. The critical member listed in the inspection report is

Girder #5 on Span 1SB with ratings of 41 tons at inventory and 68 tons at operating for the HS-20 trucks. No access to the actual calculations or the software was provided.

As an independent check on the inspection report, preliminary load ratings were calculated using the AASHTO LRFD Bridge Design Specification (American Association of State Highway and Transportation Officials. 2007). The results of these ratings, in terms of rating factor, are presented in Table 3-4. Note that all ratings were governed by flexure.

Table 3-4 - Load Rating Factors

Span	HL-93		HS-20	
	Inventory	Operating	Inventory	Operating
SB1	1.50	1.95	2.21	2.86
SB2 (G8)	1.46	1.90	2.03	2.63
SB3	1.61	2.09	2.05	2.65
SB4	1.58	2.05	2.23	2.90
NB1	1.86	2.41	2.56	3.32
NB2 (G1)	1.29	1.68	1.90	2.46
NB3	1.61	2.09	2.05	2.65
NB4	1.50	1.95	2.13	2.76

As a direct comparison to the ratings provided in the inspection report, SB1 has ratings of 44.2 tons at inventory and 57.2 tons at operating. As these ratings were calculated using the LRFR approach, and not with Bar7 (LFR), it is likely that the discrepancy lies in this difference or in assumptions or simplifications made in either or both rating calculations. In fact, the LRFR calculations indicate that the most crucial member was Girder #1 of

Span 2NB. Incidentally this member has the same section properties as all of the girders on Span 1SB.

3.9 Qualitative Risk Assessment and Development of Critical Questions

Perhaps the most important portion of the first step of an assessment is the identification of potential limit states at which the structure may fail to perform in a satisfactory manner. Given that these limit states may include everything from functional issues through structural safety concerns, it is important that this be done in a risk-based manner. That is, the bridge in question (and the surrounding network, environment) should be examined to identify potential hazards (external inputs to the bridge), vulnerabilities (internal weaknesses), and the consequences associated with failures that may be mobilized by the identified hazards and vulnerabilities. Table 3-5 provides a partial listing of some common hazards, vulnerabilities, and exposures associated with various performance limit states.

Following this procedure a qualitative risk assessment was performed on the test bridges to identify potential risk factors. From a Safety: Geotechnical and Hydraulic view point it was determined that all hazards were quite low, and while the perceptible ground vibrations felt during the site visit indicated some loading of the foundation system, the potential for this to propagate into a safety issue was taken as negligible.

Table 3-5 - Summary of Performance Limit States, Hazards, Vulnerabilities, and Exposures for Bridges

Performance Limit States	Hazards	Vulnerabilities	Exposures
Safety: Geotechnical/ Hydraulic	<ul style="list-style-type: none"> • Flowing water • Debris and ice • Seismic • Vessel Collision • Flood 	<ul style="list-style-type: none"> • Scour/Undermining • Loss of support • Soil liquefaction • Unseating of superstructure • Settlement • Overtopping 	<ul style="list-style-type: none"> • Loss of human life • Replacement and repair costs • Impact of removal from service related to: <ul style="list-style-type: none"> • Safety – life line, • Economic • Social – mobility • Defense
Safety: Structural	<ul style="list-style-type: none"> • Seismic • Repeated loads • Trucks and overloads • Vehicle collision • Fire 	<ul style="list-style-type: none"> • Lack of ductility and redundancy • Fatigue and fracture • Overloads • Details and bearings 	
Serviceability, Durability and Maintenance	<ul style="list-style-type: none"> • Winter maintenance practices • Climate • Intrinsic Loads • Impact (Vertical) • Environment 	<ul style="list-style-type: none"> • Corrosion • Cracking/spalling • Excessive deflections/vibrations • Chemical attacks/reactions • Difficulty of maintenance 	<ul style="list-style-type: none"> • User costs • Maintenance costs <ul style="list-style-type: none"> • Direct • Indirect – delays, congestion, etc.
Functionality	<ul style="list-style-type: none"> • Traffic • Special traffic and freight demands 	<ul style="list-style-type: none"> • Network redundancy and adequacy • Geometry and roadway alignment 	<ul style="list-style-type: none"> • Loss of human life and property (accidents) • Economic and social impacts of congestion

In the case of Safety: Structural limit states a few risk factors were noted. The largest hazard noted was the very high traffic and truck traffic loading (nearly 100,000 and 4,000 ADT and ADTT, respectively). While there is no quantitative information on overloaded vehicles, during the site visit numerous heavily loaded trucks were observed crossing the bridge. In addition, three primary vulnerabilities were noted including (1) the observed fatigue cracking at the location of the poor connection details of the wind bracing members, (2) the type (rocker) and significant deterioration of the bearings, and (3) the

large single crack noticed in the pier cap, which likely indicates poor reinforcement details (due to the lack of distributed cracking). While the computed rating factors indicate that the bridge does not have a capacity issue, they do not provide any information about the potential for continued fatigue crack propagation, local bearing instabilities, or localized failures due to poor reinforcement details.

From a Serviceability, Durability and Maintenance standpoint, the primary hazards noted were associated with the climate (including freeze-thaw), the use of de-icing chemicals, and the repeated heavy truck loads the bridge experiences. On the vulnerability side, five primary issues were noted, including (1) the excessive vibrations under traffic loading (nearly 0.4g of vertical upwards acceleration), (2) the transverse cracking of the concrete deck, (3) the poorly performing bridge joints, (4) the corrosion of the girder ends and bearing assemblies, and (5) the map cracking within the abutments. While these risk factors are not likely to pose a safety issue, they do impose significant costs due to their negative impact of durability which shortens the service life of the bridge.

Finally, the functionality of the bridge was assessed and four risk factors were identified including (1) significant traffic levels (hazard), (2) approach slab settlement, which caused a significant bump at the beginning of the bridge (vulnerability), (3) relatively poor network redundancy (vulnerability), and (4) high economic importance as evidenced by the large traffic levels (exposure). While in many cases it is not possible to use sensor and simulation technologies to improve bridge functionality, it is important to place the safety and serviceability risks in context with functionality when making a decision related to preservation and renewal.

Given this qualitative risk assessment and the associated uncertainty, the following critical questions were identified:

1. What is the influence of the observed fatigue cracking on the load carrying mechanisms of the bridge? Can the fatigue sensitivity be mitigated by removing the wind bracing? Could this potentially cause additional problems?
2. Is the observed crack in the pier cap due to live load or the excessive vibration of the bridge? Is this crack active under service load? Does it indicate a safety issue?
3. Given the various skew arrangements of the bridge and the resulting bridge movements under temperature and live load (including potential uplift and rotation) are the current rocker bearings sufficiently reliable?
4. Is the excessive vibration a result of the initial design or is it being exacerbated by accumulated deterioration? What options are available to reduce the level of vibration?
5. What is the relationship between the observed deck cracking and the excessive vibrations of the bridge? Is it possible that the deck is being placed under net tension due to vibrations?

While there are clearly other questions related to (1) the performance and replacement of the joints, (2) the development of mitigation strategies for the observed approach settlements, and (3) the diagnosis of the map cracking of the abutments, the use of sensing and simulation technology will not provide meaningful input to these questions. It is noted that in the case of the map cracking there are various nondestructive evaluation techniques that may be appropriate to estimate the extent of the problem.

3.10 Identification of the Relevance of St-Id and Potential Technologies

There are many approaches to reliably answer the questions above, but they all involve the integrated use of experimental and simulation technologies together with sound bridge heuristics. While visual inspection may provide insight into condition or appearance, such information is not sufficient to reliably answer the critical questions posed. It is equally important to recognize that experimental or simulation technologies alone are insufficient. In the case of experiments, the questions posed cannot be directly answered through sensor measurements as the responses available provide only indirect information. That is, they may provide us with vibration responses (frequencies, mode shapes, etc.), but they do not explain what aspects of the bridge design (or deterioration) provide the most influence over those parameters. Simulation models also fall short, but in this case due to both the complexity of constructed systems and the many uncertain parameters (e.g. continuity and boundary conditions) that greatly influence model predictions and can only be identified through comparison with experimental data. As a result, *a priori* simulation models have routinely been shown to be off by significant margins and in many cases do not provide the reliability required to inform intervention decisions.

This situation is not unique to the US202/NJ23 Bridge and is, in fact, the primary motivation for the use of St-Id which integrates experimental and simulation technologies together with heuristics to compensate for the shortcomings of each individual approach.

Given the specific questions posed, it appears that in general terms, an understanding of (1) how the bridge carries live load (load paths, force-resisting mechanisms, continuity and boundary conditions, etc.) and (2) the bridge's vibration characteristics (frequencies,

mode shapes, damping, etc.) is required. To capture relevant information related to these responses it was decided to perform both a truck-load (crawl-speed and static) test and a series of dynamic tests. While theoretically any single approach to dynamic testing would be sufficient, given the round robin nature of this study two different approaches were employed; a multiple input-multiple output impact test and a series of ambient vibration tests.

Chapter 4: A Priori Modeling

This chapter presents the construction, development, error-screening and application of an a priori model of the IBS Bridge, including the selection of the model form, determination of model geometry, and benchmark feasibility investigations of several modeling techniques. Response predictions for the test load cases are developed.

4.1 Purpose of A Priori Modeling

The development of an a priori model within St-Id process serves to provide estimates of structural responses to (1) specialize the specific experimental approaches selected (sensor layout, load levels, etc.), and (2) ensure safety and provide a comparison point for real-time interpretation of data during the experiment. The actual modeling approach adopted is dependent on the objectives of the St-Id as well as the complexity of the structure being identified. In some cases simple phenomenological models are sufficient whereas in other cases structural models or high resolution geometric-replica finite element models may be justified. In all cases, the development of a priori models should follow established modeling practices. That is, the effect of modeling assumptions should be examined through comparison with benchmark problems and the sensitivity of the simulations to parameters with significant uncertainty should be established. In addition to error screening the model, these studies also serve to identify key structural responses and their bounds to ensure the experimental program is robust and reliable.

4.2 Selection of Model Resolution

Ideally the selection of model resolution should follow directly from the critical questions at the center of the effort, which allows the a priori model to be employed throughout the application. It is acknowledged however, that in many cases, time constraints only permit a very crude model at this stage that may provide only coarse response predictions to ensure safety during the test. In these cases a more refined model (commensurate with the critical questions) may have to be developed following the experimental program.

In the case of the US202/NJ23 Bridge, the focus on load carrying mechanisms and vibration characteristics indicated that a so-called structural model (or element-level model) would be appropriate as the primary a priori model. This type of model is perhaps the most common and employs both one-dimensional (plane or space frame elements) and two-dimensional elements (plate or shell elements). In an effort to remain consistent with the three dimensional geometry of the structure, various link elements, constraints, and rigid offsets are also included. The primary advantage of these models is their ability to simulate component-level responses, and allow the impact of various member-level continuity and boundary conditions on the overall response to be assessed.

This class of models is essentially a compromise between the high resolution geometric replica models and very low-resolution phenomenological models. Geometric replica models employ three-dimensional elements and minimize the need for geometric simplifications. Such models are quite appropriate in cases of very complex geometry or where the simulation of stress concentrations is of interest. For the US202/NJ23 Bridge a geometric replica model of the connection regions where fatigue cracks were identified may be useful during later stages of the process, but is not justified at this stage.

In contrast, phenomenological models typically only employ a few elements and significant geometric simplifications. As the name indicates these models are geared towards helping the analyst understand relevant phenomena instead of the accurate prediction of responses. While not always justified, in some cases the use of multiple a priori models in conjunction with one another can provide important insight and added reliability to the overall St-Id.

4.3 Development of Model Geometry

The first issue that must be addressed related to model geometry is determining the extent of the structure to be modeled. Obviously very refined simulation approaches would be desirable, but may be precluded for time and/or cost considerations. The experience of the analyst is very important in this regard. If one were to neglect or misrepresent important mechanisms, the model can become quite misleading, especially if it is distorted to compensate for such errors following the experiment. In the case of the US202/NJ23 Bridge, the initial a priori model explicitly included the substructures of the test span (Span SB2), but did not explicitly include the foundations of any other spans. While the initial model included the entire bridge, the eventual decision to model Span SB2 in greater detail than the rest of the bridge came about as a result of the constraints of the project, and in particular the experimental constraints which limited the amount of bridge access, time and money available. These constraints are discussed in-depth in Section 5.1. Since testing focused on Span SB2, the modeling focused there as well.

Typically, a geometric model consists of centerline representations of prismatic members like beams and columns, and 2D representations of planar components like the bridge

deck. Additionally, in some cases, connection members like link elements can be included in the geometric model for direct translation to the structural model in a FE software package. However, in many cases, connectivity must be added later as it is often represented through some form of master-slave relationship in the FE software as opposed to “physical” element. Note that special attention must be paid to the relative spatial relationships between bridge components. This often requires critical assessment of how real physical connections are made in a structure and how they can be simplified for the purposes of modeling. Figure 4-1 shows a photo of the bridge as compared with the geometric model of 1D elements.

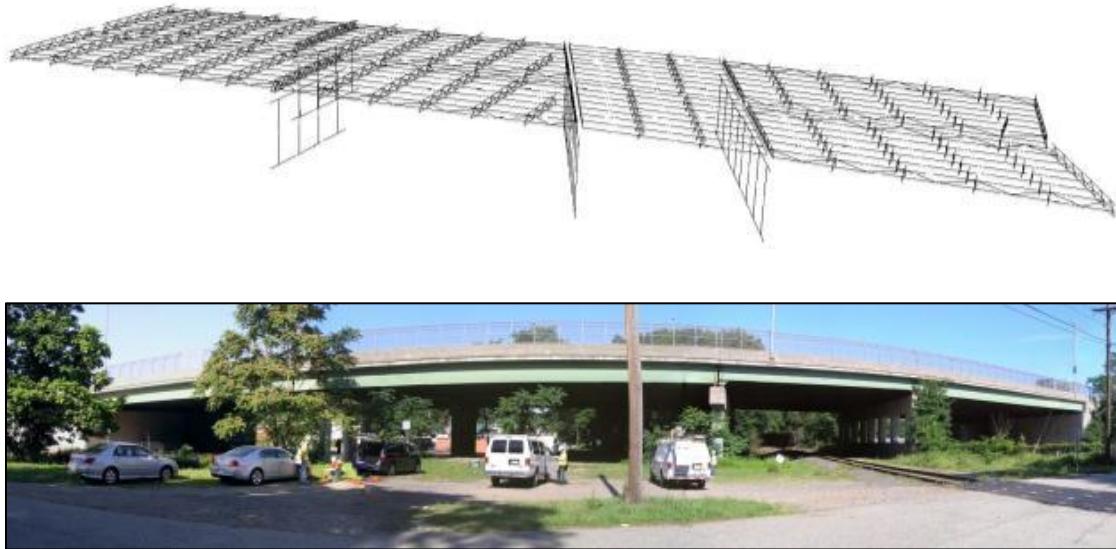


Figure 4-1 - 3D Centerline Model vs. Actual Structure

The primary purpose of this geometric model was as the basis for construction of the FE model, as described above. However, there were further intangible benefits to this type of model construction. In particular, the engineer forms a close understanding of the structure by investigating the plans to the level of detail required to develop such a model. The process forces one to answer, or at least recognize, questions that perhaps they never would have asked otherwise.

4.4 A Priori Model Construction

The selection of a simulation software package is an important step within the modeling process. While there are many reliable packages commercially available, there are differences between them and these differences should be understood and considered before settling on a single software package. In the case of the US202/NJ23 Bridge the Strand7 simulation software package was employed. While this software is not as well known in the United States as more traditional packages such as Staad or SAP2000, Strand7 has equivalent functionality (as determined through a rigorous comparison) and has the advantage that it can seamlessly interface with computational software packages such as MATLAB. As a result, MATLAB can be used to control Strand7, changing parameters, running analyses, and recording results in an automated fashion, which makes the preliminary sensitivity studies as well as post-test model calibration far more efficient. The model can be seen in Figure 4-2.

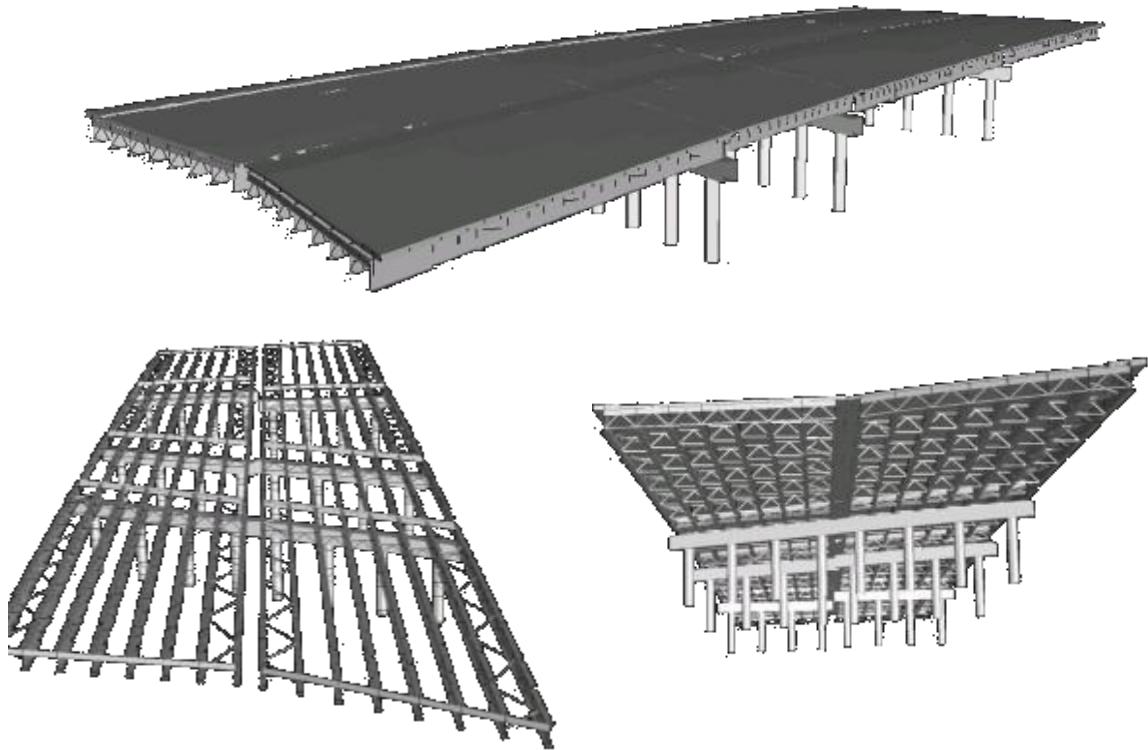


Figure 4-2 - Strand7 Finite Element Model

As previously discussed, the geometric model was imported into Strand7 as a centerline representation of the real prismatic and planar components of the structure with appropriate spatial relationships to one another. Within the FE model construction process, it is necessary to apply appropriate geometric properties to ensure consistency with the defined spatial relationships. For example, if the centroid of a girder is 40 in. from the midline of the deck, then that space must be accounted for by applying the actual section properties of the girder, deck, haunch etc. These section properties can be

calculated manually using mechanics and software like Microsoft Excel, using a third party software like Section Builder, or through internal utilities within the FE software package. In the case of US202/RT23, the section properties were generated directly within Strand7. Due to the varying flange thicknesses, vertical offsets had to be applied to the girder properties to ensure continuity of the web, as the centroid of the girder varies along the length (due to flange thickness changes).

The model includes a representation of the piers as well as the superstructure. Since, in reality, the girders are supported at their bottom flanges, rotational springs were included to account for the rotational restraint offered by this eccentricity. In general, connectivity of all components was achieved through link elements, which provide a physical connection and flexibility in terms of allowing the adjustment of stiffness. Links were used to connect the deck to the girders, bracing to the girders, and girders to the piers. The piers were supported by standard boundary conditions. Examples of these connections in the model can be seen in Figure 4-3.

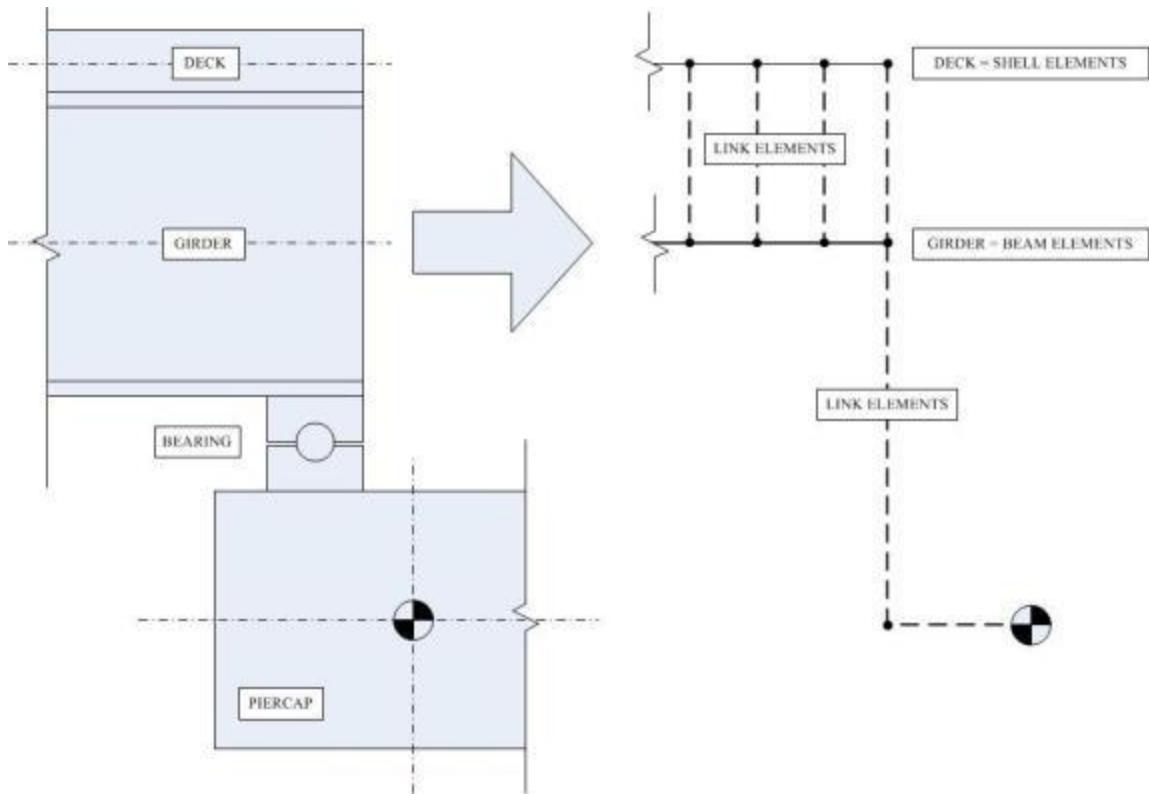


Figure 4-3 - Modeling Assumptions for Connection to Piercap

4.5 Model Error Screening

As with any attempt to model a real, complex constructed system, there is a sizeable potential for errors. Therefore it is necessary to constantly view any model with a critical eye. Through experience, a model error screening checklist for common issues has been developed to root out typical human errors associated with model construction. The general approach to model error screening is to check the model at various levels of completeness for (1) duplicate members, (2) plausible deflected shapes under dead load and simple live load patterns, (3) physically consistent mode shapes, (4) local

compatibility, and (5) local and global equilibrium. In the first step of the process the model was imported from AutoCAD and simple boundary conditions were applied. This stage occurred before section properties were assigned and details like cross-bracing and the deck were added to the model. Since each frame element had identical, default section properties, the responses (1) through (5) above were relatively straightforward to interpret and identify anomalies. This procedure was repeated after the secondary members were imported (e.g. link elements, wind bracing, diaphragms, deck) with all elements still employing default section properties. Finally the procedure was done a third time after all the appropriate section properties were assigned to the members.

4.6 Benchmark Studies

The IBS Bridge provides an interesting case from the modeling perspective. It was necessary to run several benchmark studies to understand the behaviors of the real structure and to identify how best to model the real physical components of the bridge. In general, there are three types of benchmark studies that are commonly employed to support FE modeling:

1. Comparison of the FE model to a closed-form, exact solution – This type of study is ideal for understanding the capabilities of certain elements and to support studies on relatively simple laboratory systems. Unfortunately, given the complexity of constructed systems this type of study is of only general use (as the actual geometry and loading tends to be far too complex to allow for a closed form solution).

2. Comparison of a geometrically simplified model to a more refined (geometric-replica or similar) model – This type of study aims to ensure that the geometric simplifications (1D, 2D elements and links) employed are justified and can properly simulate the actual geometry.
3. Comparison of an FE model (either structural or geometric-replica) with a very simple smeared model – This type of study aims to identify whether or not the refined model fits within reasonable bounds and has similar trends as a very simple model (which likely has fewer or no errors). This study acts as kind of a truth check and can be a good means of doing an overall error-screening.

Experienced analysts have no doubt performed numerous benchmark studies throughout their education and careers. As the library of an analysts benchmark studies increases it is very likely that new projects may require only reference to these previously completed studies. In addition to the past benchmark studies (associated with issues like mesh refinement, girder support modeling, etc.), in the case of the IBS study, an additional benchmark study was done to help the project engineers understand some of the idiosyncrasies of the structure.

4.6.1 Example Benchmark Study - Beam-Deck Compatibility

As discussed, the girders will be modeled with frame elements as opposed to explicitly modeling these members with shell elements. Before this assumption can be made comfortably, a benchmark study was conducted to verify its accuracy. The benchmark study included modeling a single girder and deck section out of shell elements, then

duplicating the exact geometry with beam and shell elements connected with rigid links.

This technique is much less computationally intensive, and is a known modeling

technique. The benchmark simply confirms that it is applicable to the IBS situation.

The model consisted of a 100ft beam with an element discretization of 1ft. The flanges were 1in thick and 12in wide, and the web had a thickness of 0.375in and a depth of 60in. The deck was constructed as a 4ft wide section of 6in thick concrete. Simple supports at the centerline of the beam were used as boundary conditions. Out of plane rotation or translation was restricted. The modeling details are shown schematically in Figure 4-4.

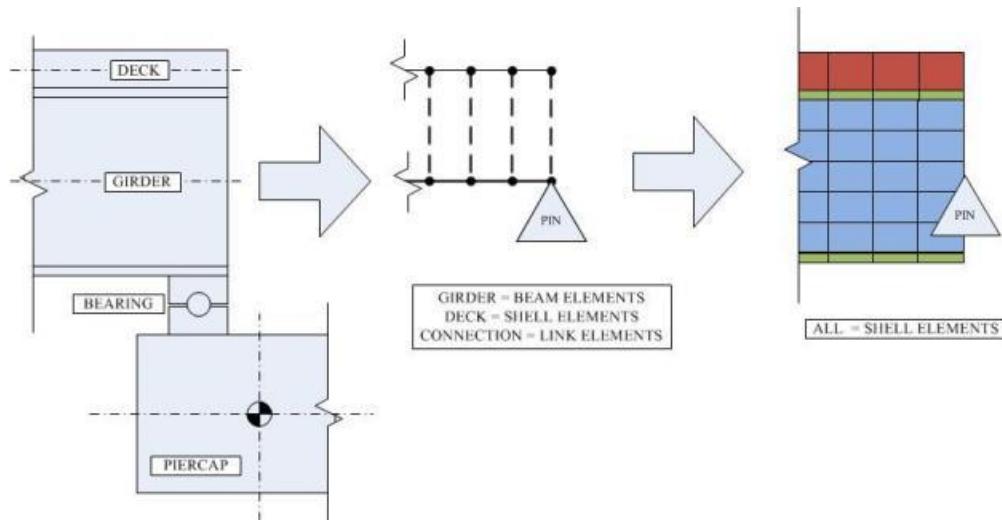


Figure 4-4 - Modeling Techniques for Benchmark Study

The benchmark models can be seen in Figure 4-5. In the extruded view it is difficult to see any differences in the models qualitatively.

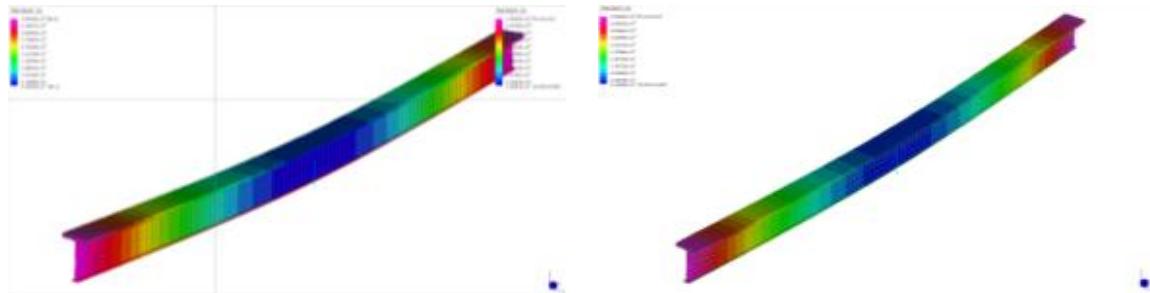


Figure 4-5 - Benchmark Models (Extruded View) a. Shell model b. Frame element model

The load consisted of a 10k point load applied at midspan, and the metric of comparison was midspan displacement. Table 4-1 shows the midspan displacement results for the two models and shows the percent difference of only 2.14%.

Table 4-1- Displacement Comparison for Benchmark Models

Shell Only	Beam-Shell	% Difference
0.240967	0.235814	2.14%

Considering these results it was deemed that modeling the IBS bridge with a combination of beam and shell elements was an acceptable modeling technique and therefore would be applied in the a priori model.

4.7 Utilization of the Completed Model for Sensitivity Studies

A major benefit of an a priori model is providing predictions for comparison to measured responses during the test. This can also be a detriment if inaccurate or unreliable model predictions are used to make decisions regarding the test. For this reason, it is recommended that uncertain parameters are varied to ensure that the response distributions for critical locations reflect the uncertainties present. Figure 4-6 is a schematic showing the possible displacement ranges for 12 different locations on the bridge as predicted by 1000 finite element models.

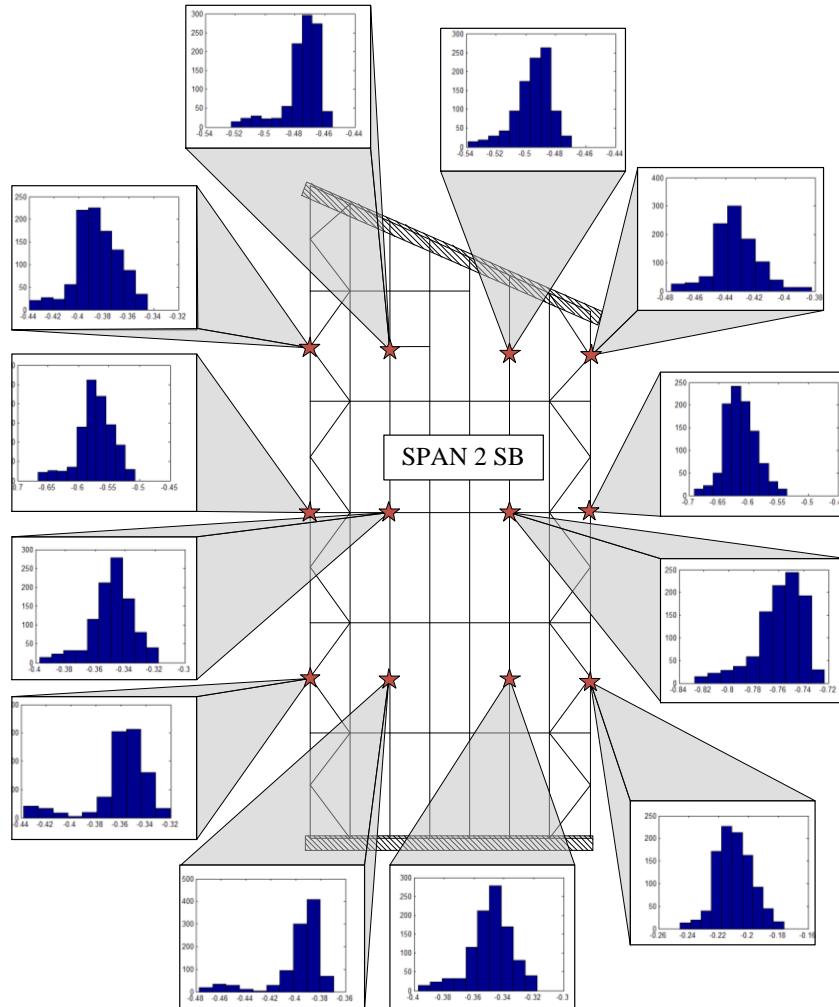


Figure 4-6 - Displacement Prediction Distributions

These response prediction histograms are the main result of the a priori model presented in this section. Interpreting these histograms can help inform the design of instrumentation and will serve as a real-time reference during the load test. Response magnitudes measured during the test can be immediately and easily compared to these distributions to ensure that the responses are within reasonable or expected bounds.

4.7.1 Identification of Model Building Blocks (Parameter Selection)

Before an a priori model can be used for response prediction for a controlled load test, an experienced user must identify what uncertain parameters exist within the model based on field visits to the actual structure and available inspection reports and prior test data (if applicable). Once a set of model building blocks have been identified (e.g. boundary conditions, continuity conditions and material properties), a user must establish whether these parameters are influential with respect to the desired predicted responses (displacements, strains and/or natural frequencies). This type of sensitivity study then provides estimated ranges of responses, based on the selected parameters.

If no experimental data exists for the structure being studied, it is highly recommended that a small subset of measurements be made so that the a priori model can be assessed for rough correlation. Preliminary data collection can assist in eliminating major model errors so that response predictions used during a controlled load test can be trusted with a higher degree of certainty. For the IBS project, preliminary vibration data was used in conjunction with multiple site visits and photo-documentation of key areas to construct a list of uncertain parameters. These parameters were grouped, due to the sparse nature of the preliminary test, into three main parameters for Span SB2:

- Stiffness of the bearings
- Stiffness of the transverse bracing
- Composite action of the stringers and deck

Before the controlled load test was carried out for the IBS project, a detailed sensitivity study was carried out for each identified parameter and a suite of models was generated

by sampling these parameters between established bounds to produce a histogram of response predictions. Since little information is available to help distinguish between parameter values of the a priori model, all parameter values and subsequent combinations within the model are considered equally likely. The following sections present the details of the sensitivity study.

4.7.2 Sensitivity Study

Sensitivity studies help to identify what the most influential parameters are with respect to the desired response. The sensitivity plots seen in Figure 4-7, Figure 4-8, and Figure 4-9 highlight how the change in each parameter independently influences the displacement at mid-span of Span SB2. It is seen in the plots that all parameters, especially rotational spring rigidity at the supports, are influential on mid-span displacement, and were subsequently used for more detailed analysis.

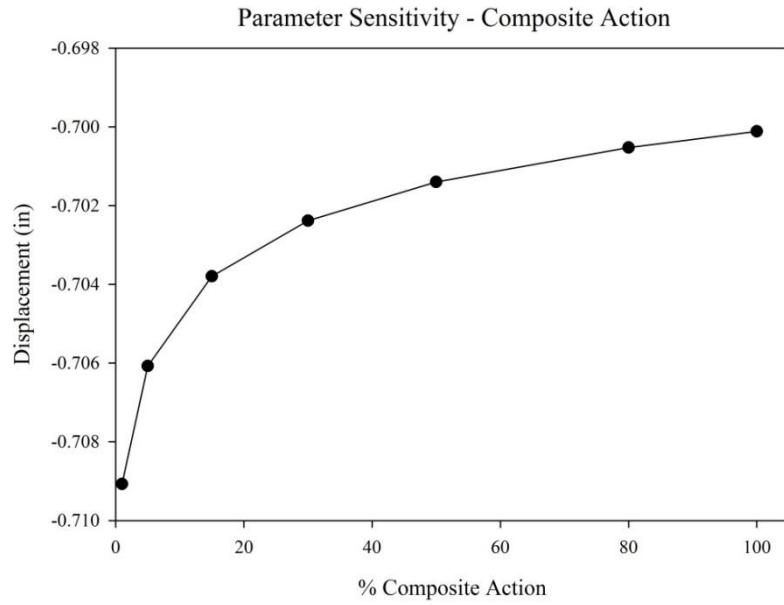


Figure 4-7 - Sensitivity of Girder #3 Displacement to % Composite Action

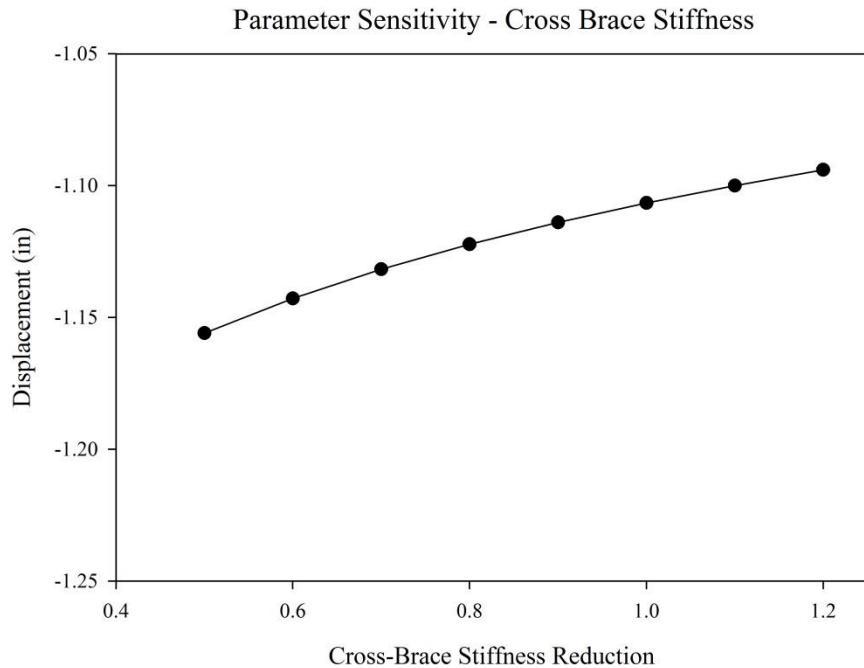


Figure 4-8 - Sensitivity of Girder #3 Displacement to Cross-brace Stiffness

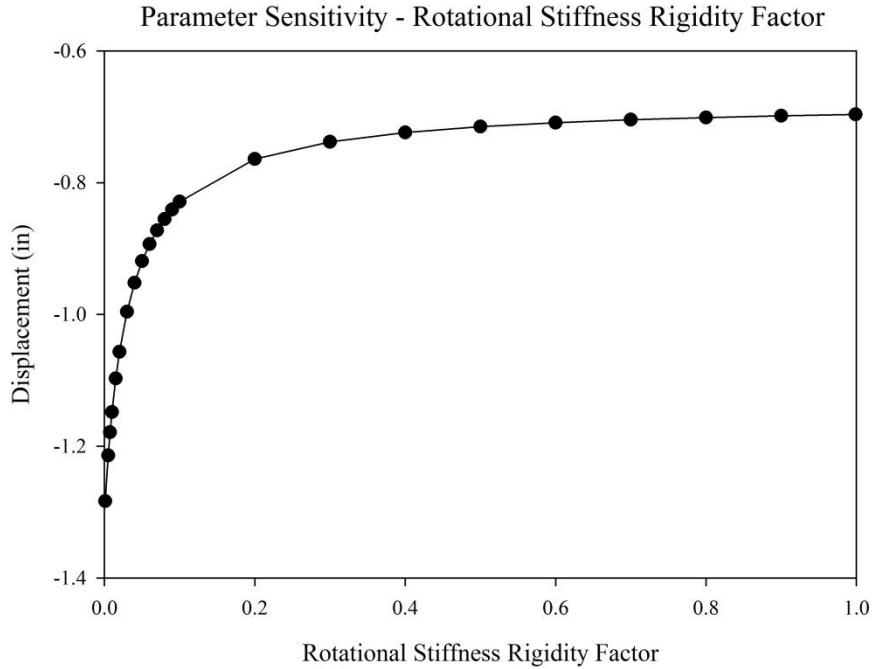


Figure 4-9 - Sensitivity of Girder #3 Displacement to Rotational Bearing Stiffness

The bounds used for the three parameters are listed in Table 4-2:

Table 4-2 - Parameter Bounds

	Lower Bound	Upper Bound
Stiffness Modification Factor of Cross Bracing	0.5	1.2
Composite Action Stiffness (kip/in)	1	100
Rotational Spring Stiffness Rigidity Factor	0	1

The bounds of the stiffness modification factor were established based on site visits and heuristics. There were no cross-bracing members that had significant section loss, and most of the uncertainty with the stiffness of these members existed in the connections. Bounds of 50% and 120% stiffness modification were therefore used to represent the cases where connections and members are not as rigid as anticipated and where the connections and members are even slightly more rigid than anticipated.

The bounds for composite action stiffness were determined initially using heuristics. The bounds were then refined using multiple sensitivity studies to determine where the responses converged to a single value after increasing the stiffness of the connection element between the stringer and the deck.

The bounds of the rotational spring rigidity factor were taken as zero (pin) and one (fixed). The rigidity factor is defined as:

$$f = \frac{1}{\left(1 + \frac{3EI}{KL}\right)}$$

where E is the elastic modulus, I is the moment of inertia and L is the length of the element whose node serves as the boundary.

4.7.3 MATLAB & Strand7 Interface

The ability to automatically modify parameter values and execute analyses is beneficial as relying on manual procedures would take significant time and cost. Therefore, a custom methodology was developed to utilize an API between MATLAB and Strand7 so that a large set of parameter values could be automatically generated, used in analyses,

and the results extracted and processed. The algorithm begins by generating samples of the uncertain parameters, as defined by the user. These parameters are then assigned to the model accordingly, and selected solvers are run. For the IBS project, the three parameters identified above were sampled (as discussed in the following section) and then assigned to the model. The model was then solved for both static deflections and strains under multiple load cases and for natural frequencies and mode shapes. The displacements and strains were extracted at locations where instrumentation was planned on being installed. The instrumentation design process was iterative based on the output from these studies, and is described in Section 5.4.

4.7.4 Parameter Sampling

Once the algorithm was developed to automatically generate and analyze many model iterations of Span SB2, it was necessary to decide on how to sample the parameters effectively. Initially, the parameters were sampled deterministically by choosing ten evenly spaced parameter values between the bounds identified above. This led to 10^3 or 1000 parameter combinations, using a full factorial approach, and subsequently 1000 models used for response prediction. The results of this initial study were poor, in that there appeared to be obvious gaps in predicted displacements indicating that the discretization of the parameters was not small enough to capture the sensitive range of one or more parameters. The deterministic sampling procedure employed was originally used for the parameter sensitivity studies and is included herein only to serve as a direct example of the potential inadequacy of this sampling procedure, however a comprehensive assessment of deterministic methods is available in Dubbs (2012).

It was found that the responses were very sensitive to the rotational stiffness factor between values of 0.0 and 0.1 and were not being accounted for in the deterministic sampling approach. This agrees with the sensitivity studies that were previously conducted. A more robust approach is to sample the parameters randomly each time a model was generated, using a Monte Carlo approach. To ensure that the most sensitive areas of each parameter were being sampled sufficiently, the rigidity factor was sampled from an exponential distribution as opposed to uniform, ensuring that the sampling was concentrated near the lower bound of the rigidity factor and not the higher bound. After revising the algorithm to incorporate the sampling modifications, the analysis was run again for 1000 parameter samples.

4.7.5 Results

The results of the revised sampling approach were much more comprehensive and provided a reasonable histogram of predicted results. Sample histograms can be seen in Figure 4-10 and Figure 4-11 which show the range of displacements and strains at midspan of Girder #3 for the different combinations of parameters within the sensitivity study. The breadth of the ranges indicates that the parameters selected were very influential, and that the uncertainty associated with these parameters is quite significant. . These types of histograms were developed for strain and displacement measurements at a regular grid of 12 locations, as described schematically in Figure 4-6.

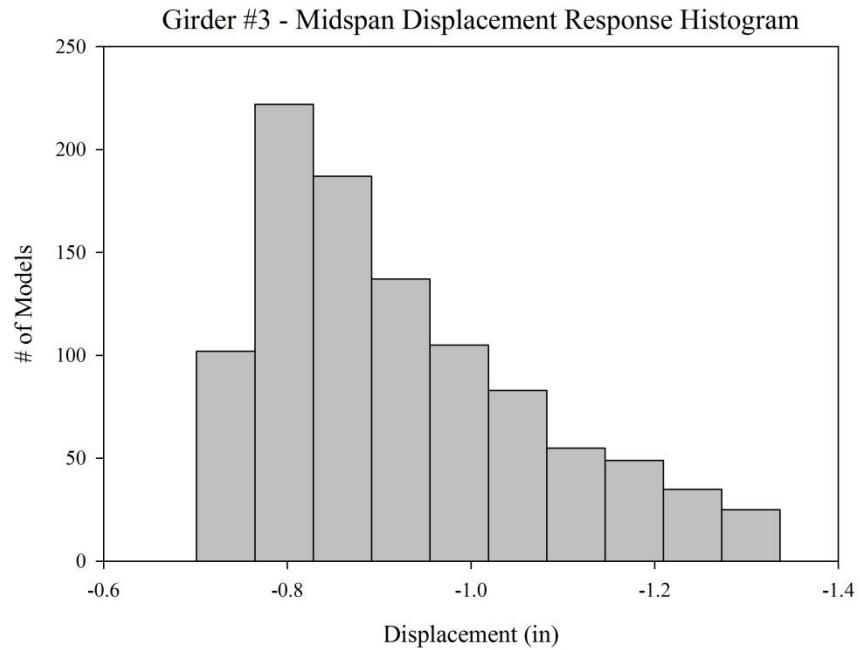


Figure 4-10 - Midspan Displacement Response Histogram

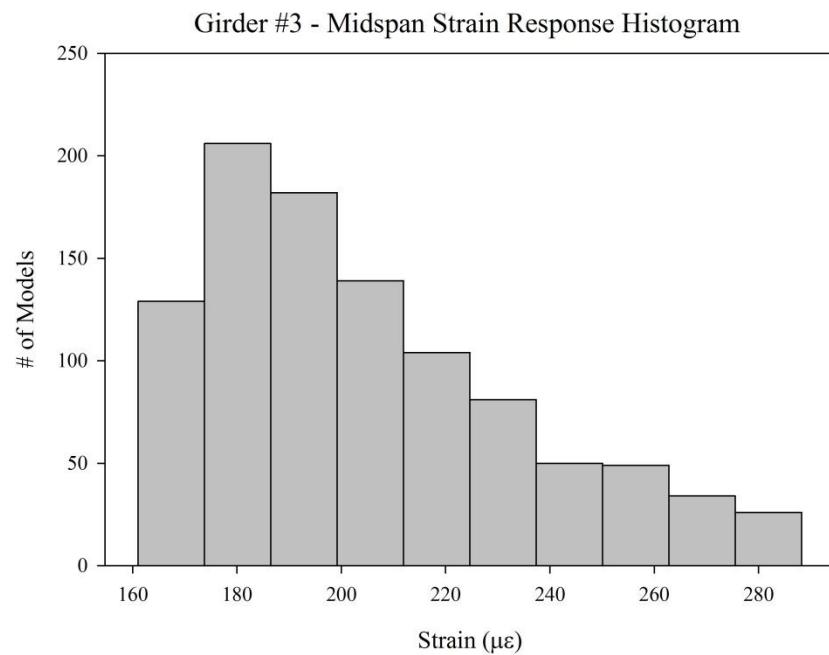


Figure 4-11 - Midspan Strain Response Histogram

Table 4-3 and Table 4-4 summarize the strain and displacement results, respectively, from the a priori model analysis. The static results (strain and displacement) correspond to Span SB2 under six legal truck loads, configured back to back and centered on the span.

Table 4-3 - A Priori Strain Results Summary

Beam	Location	Strain (uE) Min	Strain (uE) Max	Strain (uE) Range	Strain (uE) Mean
Beam 1	1/4 Span	15.30	103.66	88.37	45.48
Beam 3	1/4 Span	-5.42	86.71	92.13	26.47
Beam 6	1/4 Span	-0.94	101.32	102.26	33.82
Beam 8	1/4 Span	14.18	110.39	96.22	45.62
Beam 1	Mid Span	67.71	131.00	63.29	93.13
Beam 3	Mid Span	109.86	170.39	60.54	131.57
Beam 6	Mid Span	129.97	197.95	67.97	154.15
Beam 8	Mid Span	100.23	176.33	76.10	130.01
Beam 1	3/4 Span	20.94	72.83	51.89	38.30
Beam 3	3/4 Span	11.10	110.33	99.23	46.32
Beam 6	3/4 Span	29.58	107.27	77.69	59.70
Beam 8	3/4 Span	32.93	50.25	17.32	40.53

Table 4-4 - A Priori Displacement Results Summary

Beam	Location	Disp (in) Min	Disp (in) Max	Disp (in) Range	Disp (in) Mean
Beam 1	1/4 Span	-0.85	-0.35	0.50	-0.53
Beam 3	1/4 Span	-0.96	-0.46	0.50	-0.63
Beam 6	1/4 Span	-0.98	-0.48	0.50	-0.66
Beam 8	1/4 Span	-0.90	-0.41	0.49	-0.58
Beam 1	Mid Span	-1.15	-0.52	0.64	-0.75
Beam 3	Mid Span	-1.34	-0.70	0.64	-0.92
Beam 6	Mid Span	-1.29	-0.73	0.57	-0.93
Beam 8	Mid Span	-1.08	-0.56	0.52	-0.75
Beam 1	3/4 Span	-0.79	-0.32	0.47	-0.48
Beam 3	3/4 Span	-0.83	-0.36	0.46	-0.53
Beam 6	3/4 Span	-0.59	-0.32	0.28	-0.42
Beam 8	3/4 Span	-0.34	-0.18	0.15	-0.24

Chapter 5: Live Load Test: Experimental Design and Results

The ensuing sections describe a sound process to guide the design and implementation of a live load test, including discussions of constraints, load design, instrumentation design, and the design of the data acquisition protocols. While the process is presented in a linear manner, in reality it is quite iterative as each component has a large effect on the others.

5.1 Identification of Constraints

Once the need for a live load test has been established, the next step is to identify all relevant constraints that will influence the test design and implementation. To accomplish this, it is necessary to (1) visit the bridge and develop an understanding of the surrounding network and hard physical constraints, (2) work closely with the bridge owner to identify human side constraints (e.g. potential for lane closures, time-sensitive issue, etc.), and (3) examine the available budget and approximate the funds available for (a) personnel (preparation, on-site, data interpretation, reporting), (b) hardware purchases (sensors, mounting hardware, cabling, DAQ), and (c) traffic control and access equipment. In most cases there are many constraints that require difficult trade-offs to be made; however, if they prove severe enough to call the value of the entire effort into question, then it is necessary to either examine means of alleviating some constraints or abandon the entire effort. In the case of the US202/NJ23 Bridge, the relevant constraints identified are listed in the following:

Physical Constraints:

- There was no power available on the site. The bridge does have lights installed on the underside, but these have never been seen functioning, and it was not possible to tap into the supply for these lights. The closest business was unwilling to provide access to their power.
- The bridge spans two roads and a railroad. That leaves only one span with unobstructed underside access. Traffic control will be required to access the middle of the spans over the roadways, and proper permitting will be required to access the spans over the railroad.
- The bridge is 20+ feet off of the ground and will require ladders and bucket trucks to access. The terrain makes it possible to use bucket trucks without much difficulty. This height means that equipment installed on the underside of the bridge will be safe if left unattended overnight.
- Close proximity of a stone quarry which will allow the test trucks to be used in an empty configuration and then loaded with stone for the final load stages.

Human-Side Constraints:

- The road is an important artery and therefore requires that any single lane shutdowns must be limited to 9am to 3pm and 10pm to 4am on weekdays. The multi-lane shutdown for the actual test will be at night.
- The required project schedule did not accommodate the time needed to obtain the required permit to access the spans over the railroad, and thus no underside access to these spans will be provided.

- The bridge has a sidewalk, but it is not generally open to foot traffic, and any sidewalk work will require a shutdown of the adjacent lane/shoulder.

Budget Constraints:

- Support for three experienced personnel to be on-site for one week and for 150 person-hours of time for data processing, interpretation and reporting.
- Budget of approximately \$7500 for the purchase of permanent sensors, mounting hardware, cabling, etc.
- Sufficient budget to provide two bucket trucks for access during the one-week of site work, 12 hours of traffic control for Mountainview Blvd, and one 12 hour night closure of southbound US202/NJ23 for each round of testing.
- Sufficient budget to provide up to six tridem trucks for the night closure.

5.2 Design of Live Load Test Program

One of the first decisions that must be made, which essentially drives the remainder of the test, is the selection/design of the load levels and load configurations to be used.

Perhaps the most basic decision is whether or not the objectives require a controlled live load test or if an ambient (e.g. operating) live load test is sufficient. While the former is more expensive (due to the required lane closures, trucks, etc.) it has the distinct advantage of having a stationary, known input (truck weight and location) associated with all the measured responses, which is a significant benefit during the model-experiment correlation phase. In contrast, the ambient traffic test is less expensive while capturing the actual operating response of the bridge due to the traffic it normally carries

(typical truck weights, traffic speed, dynamic effects, etc.). In cases where data under ambient traffic is obtained, it is highly recommended that the topside of the bridge be videotaped to allow the rough correlation of truck load events to the largest responses measured.

For the US202/NJ23 Bridge, the primary objectives require that the force-resisting mechanisms and boundary/continuity conditions are established, which necessitate the full transfer function to be captured (e.g. input-output relationship). As a result, it was decided to primarily focus on a controlled live load test. However, given the presence of the instrumentation, a limited ambient traffic monitoring study prior to the controlled test was also conducted.

The following sections provide a discussion of the design of the controlled live load levels and configuration, within the constraints outlined in the previous section.

5.2.1 Load Levels

One of the most fundamental requirements of live load tests is that they provide several different load levels in order to assess the linearity of the response. While this requirement can be partially achieved through the use of multiple trucks at a single load level, this is not ideal as single and multiple truck configurations will result in different pattern loading that does not allow linearity to be directly assessed. It is recommended that the test begin with one or more empty trucks and proceed to fully loaded trucks as this provides a direct means to assess linearity. If the proximity of the bridge to stone or

salt deposits is large, then any single truck will likely not be able to be used in both empty and filled states and thus additional trucks will have to be employed.

While it varies from bridge to bridge, ensuring that the load levels induce appreciable responses is a primary concern for several reasons. First, it is important that the load level be capable of pushing the responses well above the noise floor and within the reliable range of the installed sensors. Second, the final load stage should be sufficient to overcome stick-slip mechanisms and fully seat the structure so that the apparent behavior is, in fact, reliable in the long-term. It is also recommended that bridge be “exercised” prior to load testing by repeatedly driving fully-loaded trucks across it. Third, the load level should well-exceed the legal limit to allow the identification of any low-level nonlinearities and force redistribution mechanisms, as well as to be sufficient to demonstrate clearly the actual capabilities of the structure. For example, while single-truck tests are somewhat common, this invariably requires the extrapolation of the results to higher load levels and thus drastically reduces the reliability and the value of the test.

By satisfying this requirement, the value of test is greatly enhanced, but the potential for safety concerns during the test is also raised. While cases where such tests have actually resulted in permanent damage are exceedingly rare, it is important that such concerns be taken seriously. In general, three approaches should be employed. First, the a priori modeling and sensitivity studies should be used to provide a lower bound of the likely capacity and upper bound estimates of the measured responses. Second, the data acquisition protocol should allow for real-time interpretation of the results and direct comparison with the a priori modeling results. Third, the use of numerous load levels allows the actually linearity of the bridge to be assessed and decisions made based on

how much softening is observed (note that all structures should be expected to display a degree of nonlinearity under proof-level loads). If any of these indicators show the potential for safety concerns (or even permanent damage to the bridge) the test should be concluded prematurely.

In the case of US202/NJ23 Bridge six tridem trucks were available. Based on the distance to the nearest quarry, it was decided that the test would start with three empty trucks and three fully loaded trucks, and then the empty trucks would travel to the quarry, be fully loaded and then return to the bridge for the final load stage of six fully load trucks. This allowed for testing with three empty trucks followed by three full trucks to directly examine linearity. In addition, this allowed for a final load stage consisting of six fully loaded trucks with total weight of approximately 460 kips.

5.2.2 Load Configuration

Load tests are often conducted using either crawling trucks, static trucks (in various configurations), or a combination of the two. Tests that employ crawling trucks have the advantages of being able to directly obtain influence lines and easily compute distribution factors. The primary shortcomings of this approach are that it is not possible to examine the responses at higher load levels and it is difficult to obtain exact position of the load. In contrast, static truck loads allow the precise positioning of several trucks (back-to-back, side-by-side) and facilitates data collection over long windows to allow averaging to remove noise (particularly important when some lanes remain open to traffic). The primary drawbacks of static tests include the time required to carry them out, and that they are not ideal for the computation of distribution factors or influence lines. In the case

of the US202/NJ23 Bridge, both crawl-speed and static tests were selected due to their complementary nature.

5.2.3 Traffic Control Plan

Traffic control was provided by outside contractors, organized by Parsons Brinckerhoff, who acted as on-site coordinators. The testing required a closure of three of the four lanes. Since three lanes were closed, the testing had to be completed at night. The closure included the exit lane (right lane) and the next two interior lanes, leaving only the first (left) lane open. The traffic control layout can be seen in Figure 5-1. There is an entrance ramp on the north side of the bridge and an exit ramp on the south side which limited the length of the traffic control closure. The red zone was shut down and available for testing, while the green area was open for traffic. The live load testing was focused on Span SB2.

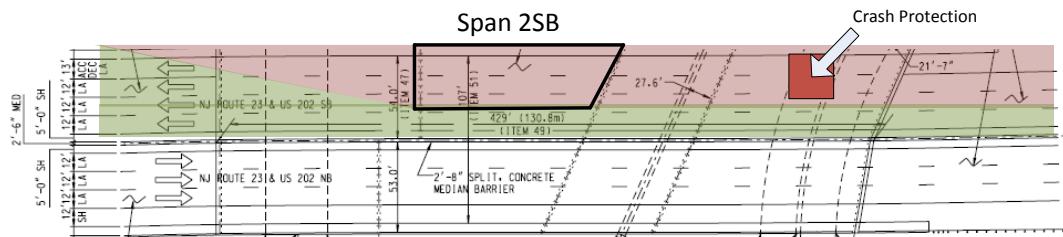


Figure 5-1 - Actual Traffic Control Plan

It is important to account for setup and breakdown time for traffic control. On average, for a heavily traveled road like US202/NJ23, it takes upwards of an hour to put in place all the required signage, cones and safety equipment. Considering that the first cone cannot be put down until the start of the closure window, this effectively shortens the useable closure window by at least two hours, barring any complications.

5.3 Criteria for Instrumentation Plan

Instrumentation utilized for static load testing must not violate the identified constraints and must satisfy two criteria: (1) satisfaction of the overall test objectives and (2) the ability to provide the user with situational awareness to maintain safety and assess/correct any data quality issues that may arise. Constraints, as discussed previously, arise from physical, human and budget realities. As in many cases these are non-negotiable, these are the first issues to be considered when designing an instrumentation plan. Next are the objectives of the test, which include the ability to understand the structure both in terms of model-based and direct interpretation of the data. Situational awareness is simply maintaining a properly operating sensing/data acquisition system and a safe testing environment. Like the overall test design, the process of instrumentation design is also iterative with the key issues to be considered shown in Figure 5-2.

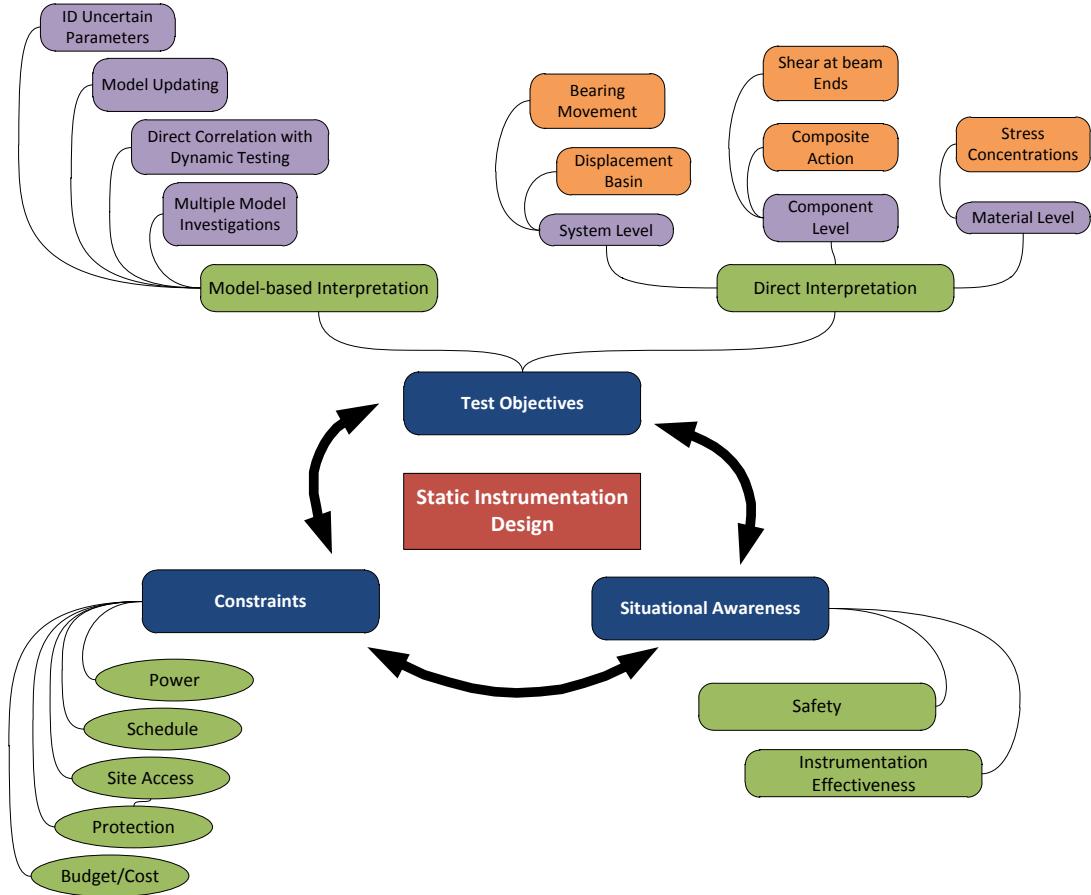


Figure 5-2 - Instrumentation Design Flowchart

5.3.1 Requirements to Meet Test Objectives

The objectives of the test are directly related to the critical questions previously listed. Excluding safety, the remaining objectives can be loosely grouped into two categories; attainable through model-based interpretation and attainable through direct or physical interpretation. Looking first at model-based interpretation, there are several procedures which can be completed post-test that require data of a certain structure. Model updating, Multiple Model Investigations and Direct Correlation with Dynamic Testing all require

regularly spaced instrumentation to capture the overall global behavior of the structure.

Working towards this, a core instrumentation plan is developed. The core instrumentation is essentially a dense grid of sensors; designed to be reduced if need be to satisfy other requirements. This instrumentation directly informs the St-Id process.

Direct interpretation of data often requires instruments that coincide with the core instrumentation. In particular, direct interpretation occurs at the system, component and material levels. Often the core instrumentation plan can inform system and component level requirements like displacement basins and composite action checks. However, some system and component level behaviors, like bearing movements, shear, and crack opening would not coincide with the core instrumentation. Assuming that the limit on the number of sensors in the core instrumentation was the capability of the data acquisition, then it will be necessary to sacrifice some of the core in order to accommodate sensors for direct interpretation. Often, depending on the situation, material level direct interpretation will be required. One example would be determining a stress concentration at a fatigue susceptible detail. Again, allowing for this measurement may require scaling back the core instrumentation. This revised instrumentation plan, based on the core plan but allowing for direct interpretation, is called the specific instrumentation plan.

It is worth noting that measurements for direct interpretation can still be utilized for model-based interpretation, whether it is as a part of the model updating, or as an independent check following model updating.

5.3.2 Situational Awareness

The last step in designing an instrumentation plan is ensuring situational awareness.

Essentially, this means safety first and foremost. The person responsible for the test has to make decisions about load placement, load level, traffic control and other logistics on the fly in a limited time window. By ensuring that the instrumentation directly informs these variables, it is easier to make confident, correct decisions. An example of instrumentation specifically used for situational awareness might be a clip gage (strain-based displacement transducer) which spans an active crack in a concrete pier. This gage will do little to inform direct or model based interpretation, but if the response gets too large at this location, the test may need to be stopped.

In most cases, the instrumentation for direct and model-based interpretation will be more than adequate for an overall safety assessment of the structure. Additionally, by observing the response under either ambient traffic or lower load levels, it is possible to develop an understanding of which gages tend to record the highest responses. While it is absolutely impossible to be sure that the gage is at the exact location of maximum response, it is possible to be confident based on relative relationships of responses that one gage is more critical than another. These critical gages can be used to ensure that the structure does not undergo any plastic deformation.

The balance between remaining in the linear range of the structure and achieving an appreciable level of load is a complex problem. Model predictions help in that they give the researcher a “feel” about how the bridge is responding. However, it is crucial that these predictions are easily accessible, interpretable, and directly comparable to measured

responses. This leads to the previously discussed sensitivity studies and response distributions.

5.4 Live Load Instrumentation Design

The procedure described above was used to design the instrumentation for the International Bridge static test. The final instrumentation design and how it was settled upon is presented in the following sections.

5.4.1 Instrumentation Plan

As previously mentioned, most test objectives are met through either model-based or direct interpretation. A core instrumentation plan was developed based on achieving model-based interpretation, and modified to ensure that direct interpretation is possible. Static instrumentation was limited to Spans SB2 and SB1 as a result of the constraints described previously.

5.4.2 Displacement Measurements

Displacements are a critical, globally referenced response. For this test, twenty displacements were measured in the vertical direction. This was limited by the number of displacement sensors available. The sensor used was the Celesco PT8510 string potentiometer. The sensor had a 5 in. total range over a 0 to 10V output range, accuracy of +/-0.28% full stroke, and an essentially infinite resolution. This is a particularly robust model which is resistant to environmental effects like moisture. The potentiometer was

mounted onto an aluminum plate with countersunk bolts, allowing it to rest flush against the underside of the girder flanges, and to be attached to the flange with c-clamps. The sensor, mounting plate and installation setup can be seen in Figure 5-3.



Figure 5-3 - Displacement Sensor, Mounting Plate and Configuration

The sensor was connected via a braided, high-strength fishing line to lead weights which rested on the ground. The fishing line was lab tested to determine the creep behavior. It was shown that by leaving the weight on the line for at least 10 hours before the test the majority of the creep would occur before the start of the test. After several hours, the rate of creep was such that it would not be noticeable over the time windows used during each stage of the test. This type of wire was preferred over piano wire, which does not creep,

but is notoriously difficult to work with in the field due to its propensity to tangle and kink. The creep behavior of the braided fishing line can be seen in Figure 5-4.

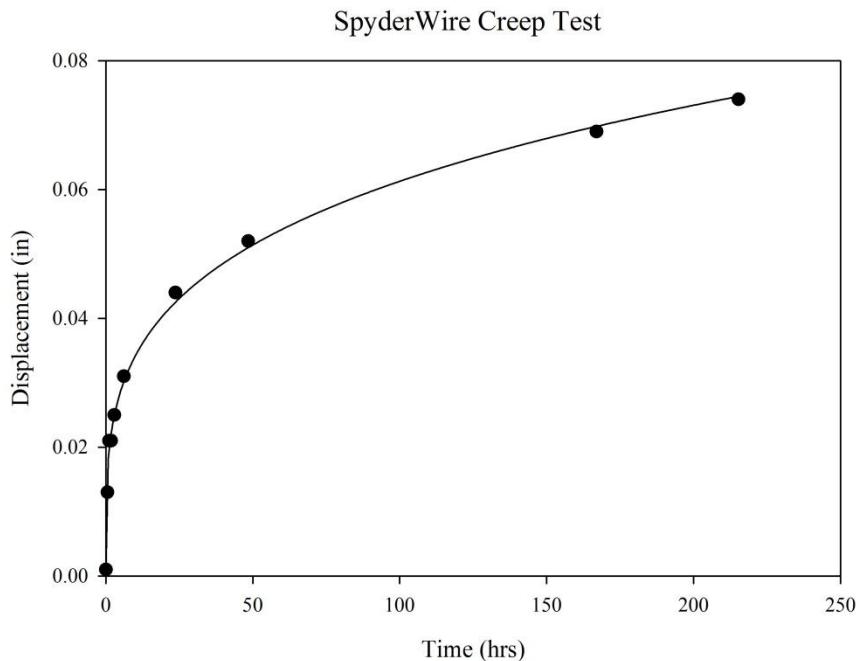


Figure 5-4 - Creep Behavior of Fishing Line for Displacements

Each sensor was mounted using the following procedure:

1. The sensor was attached to the plate on the ground, brought up to girder level, and connected to the DAQ via a pre-run cable with a connector installed.

2. The plate was then mounted to the flange at an exact location determined by using a laser tape referenced to the bearing on that girder.
3. The lead weight was tied to the fishing line using a no-slip knot. The fishing line was then run up to the sensor using a manlift.
4. The length of line required was determined by displacing the gage to 50% of its overall stroke.
5. The line was cut and attached to a hook at approximately the length required to achieve this 50% of range initial value.
6. The line was used to pick up the lead weight and use it as a plumb to ensure the wire was vertical.
7. The line was hooked to the gage, all while the response was being monitored.
8. The output of the gage was measured, and if adjustment of line length was required, then the lead weight location was raised up or lowered down until the response was near mid-stroke.

Prior to the test, each sensor was double-checked by displacing the wire downwards manually and verifying the response through the real-time visualization screens.

5.4.2.1 Displacement Sensor Locations

All sensors were installed on Girders 1, 3, 6, or 8. It was desirable to have a sensor on every girder at a given line, but because of the number of sensors available, this was not possible. The displacement sensor layout can be seen in Figure 5-5.

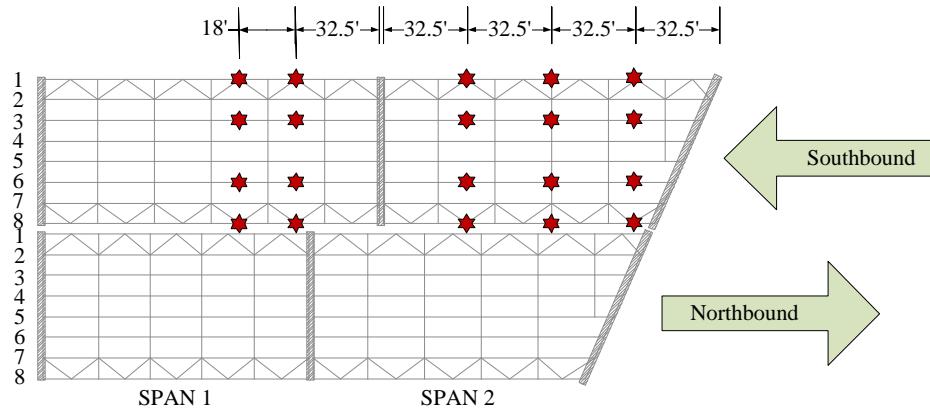


Figure 5-5 - Displacement Sensor Layout

Of the 20 displacement sensors used, 12 were located on Span 2 SB in a rectangular grid. This is in contrast to the partial skew of the span. A rectangular grid and a grid along the skew angle were both considered, but in the end it was decided that more information could be directly interpreted from a rectangular grid. In particular, comparisons between fascia girders on the long and short side of the skew would be easier to draw rapidly, giving some information on the skew. This would be directly comparable to a partial grid on SB1, which is a non-skewed span. The grid was limited to the region of SB1 that did not pass over the roadway, hence a partial grid. The second row of sensors was approximately 14' from the center of the span. Note that the displacement sensors were only placed in locations where no traffic control was required, as the wire and weight would impede traffic as long as it was in place.

These locations will coincide with other modalities of instrumentation including strains and accelerations.

5.4.3 Strain Measurements

Strain was measured in three different configurations on the bridge. The first configuration was longitudinally along the length of the girders. The next was a transverse measurement on the diagonal members of the diaphragms, and the last was on the wind bracing between the exterior girders.

The strain gages used throughout the test were the Hi-Tec weldable quarter-bridge strain gage, shown in Figure 5-6. The sensors had a 2 in. shim length with a 1 in. gage length.

They were attached to the structure via the following procedure:

- (1) The location of the gage was precisely determined using laser measurements referenced to the bearings.
- (2) The paint and any corrosion at the gage location were removed using a 4" angle grinder.
- (3) The gage was positioned and taped in place throughout installation.
- (4) The shim around the gage was attached to the raw steel via microdot welding.
- (5) Once securely welded, the cable was strain-relieved and run along the member to the designated data acquisition system.
- (6) The area of exposed steel was painted over to prevent corrosion

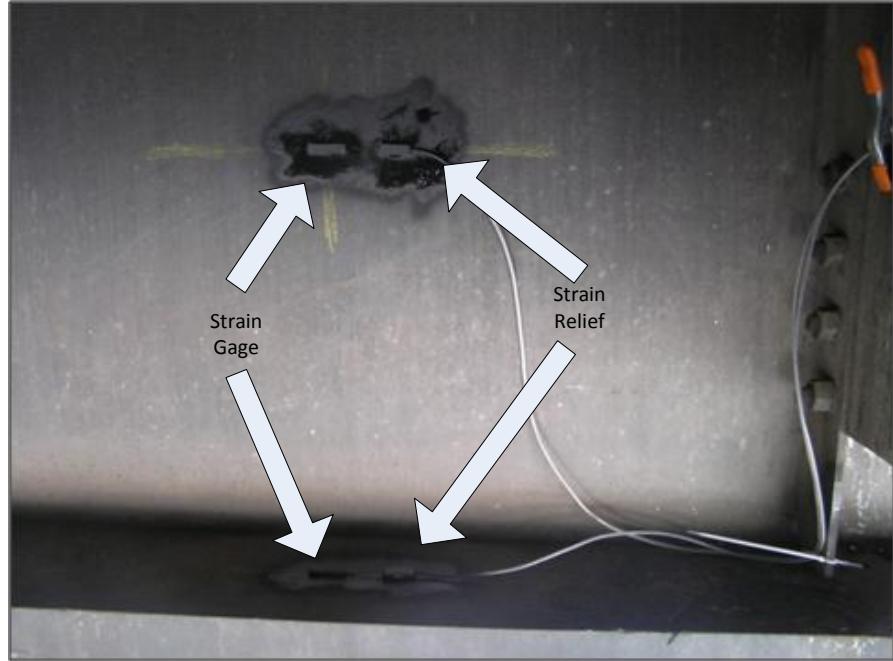


Figure 5-6 - Installed Strain Gages

These gages are effectively a permanent installation in that while they can be removed from the structure, they cannot be reused. Fortunately, the bond is typically very strong and the gage can be used confidently over a period of several years, particularly if it is located in a place that is not openly exposed to the elements. This facilitates future testing and ambient monitoring.

5.4.3.1 Strain Gage Locations

Figure 5-7 shows the typical longitudinal configuration of strain gages. This configuration consisted of two gages at the same longitudinal distance from the bearings. The first was mounted on the topside of the flange, while the second was mounted 20" up the web. The gages provide two data points in a strain profile of the cross section. In

select cases a third gage was installed at 30" from the flange to check linearity. The purpose of locating multiple sensors at a given cross-section was to use the strain profile to check compatibility between the girder and the deck. Obviously only two gages at a given cross-section can only predict a linear strain profile, so this assumption was used whenever only 2 sensors were used. Time and budget constraints prevented the use of more than two gages at every selected location, though this was desired.

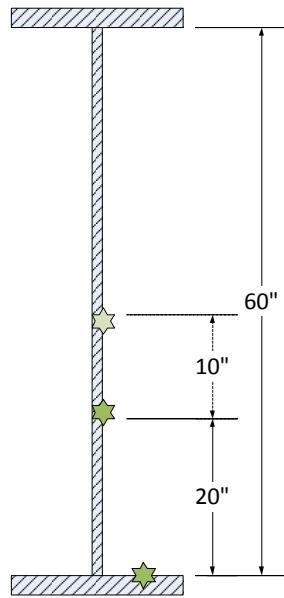


Figure 5-7 - Longitudinal Strain Gage Configuration

Note that no gages were installed on the upper half of the member, or on the concrete deck. This is because the neutral axis of bending should be located in that region, and the correspondingly small strain responses near the neutral axis would be unreliable at best.

The assumption of linearity within at least the steel girder will be justified by checking the locations where three sensors were installed. Ideally, in the cases where the third gage was used, it would have been installed away from the non-composite neutral axis, or approximately the center of the girder. However, due to the uncertainty in regards to the location of the neutral axis, as well as difficulty accessing the upper portion of the beam, 30 in. above the flange was used. If the structure is completely non-composite, meaning the neutral axis will be close to the third sensor, then the lower two gages should serve as an indicator.

The diaphragm configuration consists of two sensors, one installed on each of the diagonals of the diaphragm. The sensors were installed on the vertical leg of the angle member at $\frac{1}{2}$ of its total length, as shown in Figure 5-8. The sensors were installed using the same procedure as above.

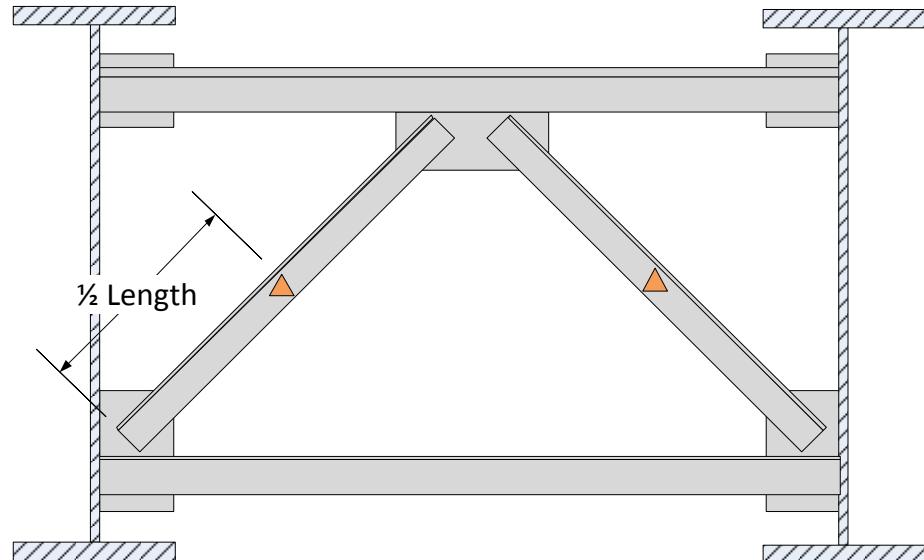


Figure 5-8 - Diaphragm Strain Gage Configuration (Elevation)

The wind brace configuration consisted of six sensors for a given pair of diagonal wind brace members. The diagonals are an inverted T-section, and one gage was placed at mid length of the member on each of the three legs of the T, for a total of six sensors. The configuration can be seen in Figure 5-9.

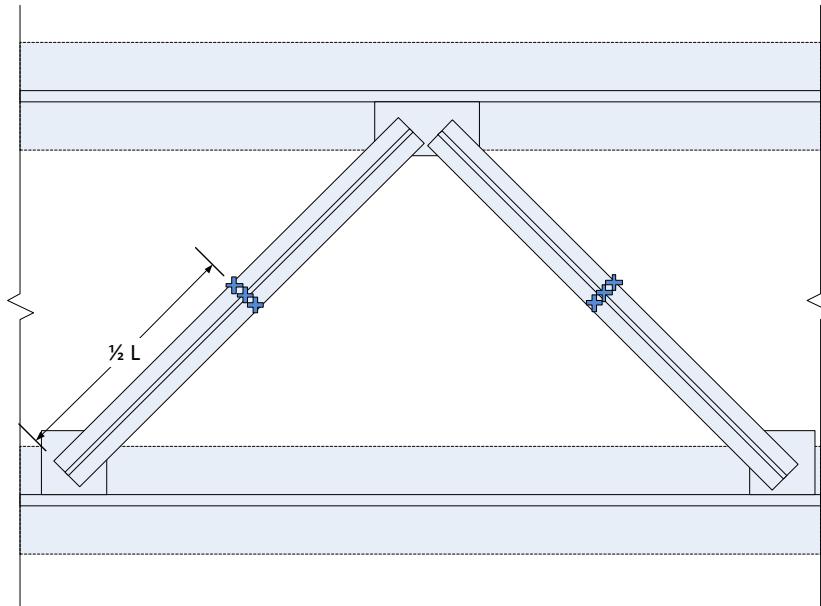


Figure 5-9 - Wind Brace Sensor Layout (Plan View)

The strain gage layout for Span SB2 is shown in Figure 5-10. The longitudinal configuration coincides with the locations where displacement was measured and is represented by green stars. The diaphragm and wind brace configurations are concentrated on a particular transverse line that runs between the $\frac{1}{4}$ and $\frac{1}{2}$ point and are represented by orange triangles and blue crosses, respectively.

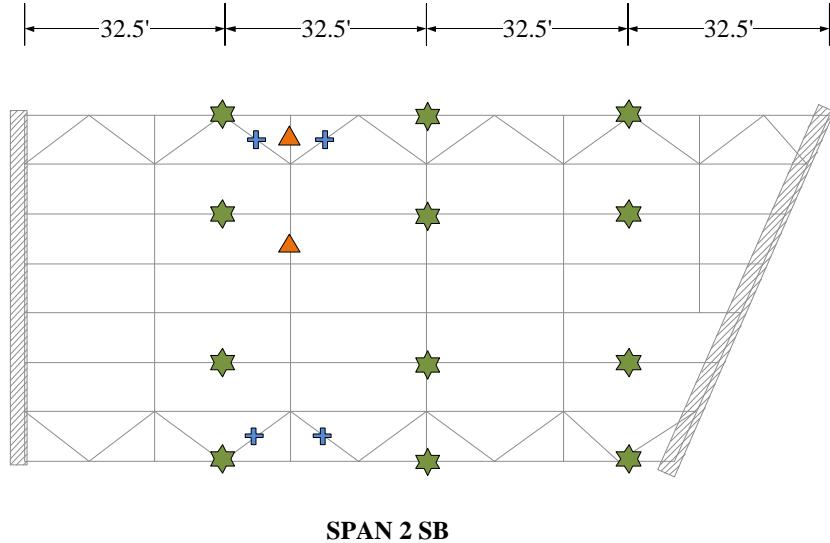


Figure 5-10 - Strain Gage Locations - Span 2 SB

The strain gage layout for Span SB1 is shown in Figure 5-11. The longitudinal configuration with two sensors coincides with the locations where displacement was measured and is represented by green stars. The longitudinal configuration with three sensors is represented by the teal stars.

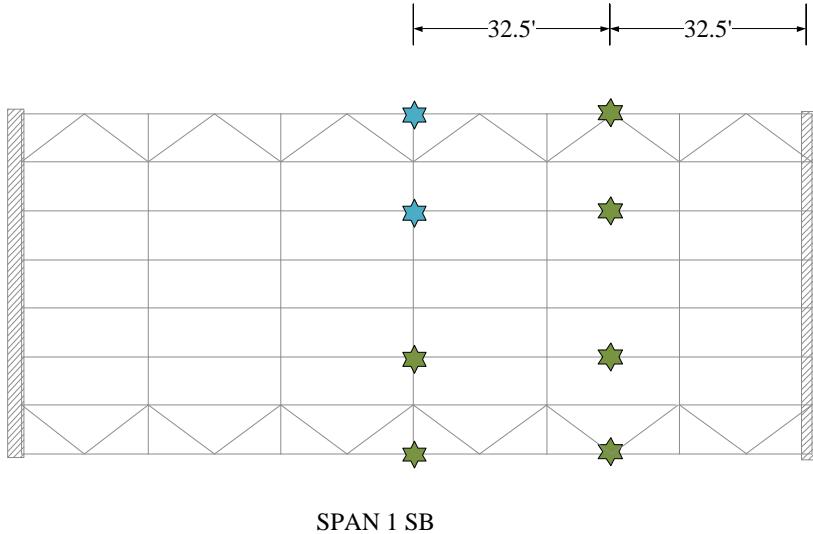


Figure 5-11 - Strain Gage Locations - Span 1 SB

Minimal instrumentation in the form of strain gages at midspan of Girders 1, 3, 6 and 8 was installed on both Spans NB1 and NB2, though no static test was conducted on the northbound side. Ambient traffic data was collected for both directions however.

5.4.4 Other Instrumentation

In addition to the core instrumentation described in the previous sections, there were several other responses that were deemed crucial for the static load testing effort.

5.4.4.1 Longitudinal Bearing Movement

The rocker bearings on the north end of Span 2 SB were instrumented in the longitudinal direction using TML strain-based displacement transducers. These sensors have a full

range of 25mm, and were positioned in the middle of that range to record the displacement of the bearings in the lateral direction. The sensors were placed there primarily as a safety check and as a directly-interpretable response relating to the global behavior of the entire span. Due to the complex nature of the movement of the bearings, it will be difficult to use this single response to compare to model output for the behavior of the bearing. However, the magnitude and distribution of the response provides immediate feedback about how the bridge is behaving.

5.4.4.2 Piercap Crack Opening

The piercap on SB2 had a particularly large crack which was showing evidence of rebar corrosion. The crack, highlighted in red, is shown in Figure 5-12. It is worth noting is the large crack on the underside of the piercap, highlighted in blue. This second crack extends under the cap along the connection to the column.



Figure 5-12 - Piercap Crack

The vertical crack was instrumented with a clip gage to see if it widened under traffic loading, or if it was simply due to weather related issues at this point. Figure 5-13 shows the attempted installation of the clip gages on the upper half of the piercap. Ideally, multiple gages would have been used, but much of the concrete on the face of the piercap was not sound enough to mount the gages. In the end, a single gage was mounted on the lower portion of the piercap where the concrete was sounder.



Figure 5-13- Installation of Clip Gage over Crack

5.5 Centralized vs. Distributed Data Acquisition

In the past, the standard for load testing was a centralized data acquisition system for field tests. Most companies which build data acquisition systems provide robust centralized systems in terms of acquisition capability. However these systems tend to be designed for laboratory use and are sensitive to outdoor conditions like dust and precipitation. Centralized DAQ may simplify power requirements, but it also complicates sensor cabling. Running all sensor cables to a single location not only requires a substantial amount of extra cable, but also tends to limit the location of the

centralized DAQ. Long cable runs can cause degradation of the signal as well. Table 5-1 shows the pros and cons of both centralized and distributed data acquisition

Considering these conclusions, drawn from numerous prior field test experiences, it was decided to use a distributed approach to data acquisition. The system consists of several small DAQs mounted on the structure, as opposed to a single DAQ on the ground. This type of system is employed by other entities who do load testing, though these systems are typically limited in channel count. By digitizing the data remotely on the bridge and bringing that digital signal back to a central location through a single cable, a substantial amount of signal cable is conserved.

Table 5-1 - Centralized vs. Distributed Data Acquisition

	Centralized Data Acquisition	Distributed Data Acquisition
Pros	<ul style="list-style-type: none"> • Simple, highly accurate synchronization • Easy to power • Countless ready-made systems available • Off-the-shelf software is available • Configurable inputs 	<ul style="list-style-type: none"> • Small form factor • Rugged, robust system • Hardwired into junction box • Military connectors • Ground locally to the bridge • Only ethernet cable run off the bridge to control laptop • Custom software is very flexible and powerful • Configurable inputs • Expandable via ethernet network
Cons	<ul style="list-style-type: none"> • Often designed for laboratory usage • Cumbersome size and setup • Very long sensor cable runs to DAQ location • Messy, unsightly cable routing • Grounding difficulties • Screw terminals are difficult in the field • Often a limit on channels 	<ul style="list-style-type: none"> • Separate systems require synchronization • Each DAQ box needs power • Off-the-shelf solutions are not affordable • Custom box construction (internally hardwired with screw terminals) • Custom software is time-consuming

5.6 IBS Case Study - Data Acquisition Hardware

The following list of criteria was used in selecting the type of data acquisition hardware to use:

- 1) Capable of sampling at high rates (kHz if needed)
- 2) Environmental robustness
- 3) Compatibility with at least voltage, resistance and current based gages
- 4) At least 16 bit A/D; Preferably 24 bit
- 5) Input flexibility (ability to customize cards/modules)
- 6) Compact form factor
- 7) Reasonable power requirements (capable of running on battery if needed)
- 8) Software flexibility

Considering these requirements, the National Instruments CompactRIO (cRIO) model line was selected for the basis of the distributed data acquisition system employed for the IBS experiment. The specifications for the cRIO system can be seen in Table 5-2.

Table 5-2 - cRIO Specifications

Processor	533 MHz RT Processor
Storage	2 GB
Memory	256 MB DRAM
Inputs	Ethernet/USB/Serial
Input Voltage Range	9V to 35V
Temperature Rating	-40° to 70°
Shock Rating	50 g
Power Consumption	35 W

5.6.1 Configuration for Outdoor Field Work

One benefit of the cRIO platform is the substantial increase in robustness and durability.

However, the system still requires a bare wire connection, as opposed to a connector.

Additionally, the cRIO needs an external power supply to regulate the voltage coming into the controller. Using bare wire connections on the bridge would be cumbersome, as would mounting both a DAQ and power supply. Also, while the system is much more robust than typical centralized systems, it is not impervious to weather.

In order to remedy these issues, it was decided that the cRIO and power supply would be mounted into an enclosure. Connectors for data cables, power and ethernet would be mounted on the outside of the enclosure and hardwired to the system on the interior. Essentially this provides an easily transportable, mountable data acquisition system which, once configured, need not be opened and accessed at all in the field. Connections are made through the exterior of the box. The final DAQ box layout can be seen in Figure 5-14.



Figure 5-14 - DAQ Box Layout

Note the addition of grounding terminals and a ground lug on the exterior of the box as well. The Pelican™ carrying case provided a weatherproof, resilient enclosure system which was easily modified to house the DAQ. Six-pin military connectors were used on the outside for sensor connections because of the versatility of having six pins as well as the environmental robustness of the military grade connections. The total distributed data acquisition system used for this test consisted of five of these boxes, an Ethernet switch and a single laptop. A schematic of the network setup for the system is shown in Figure 5-15.

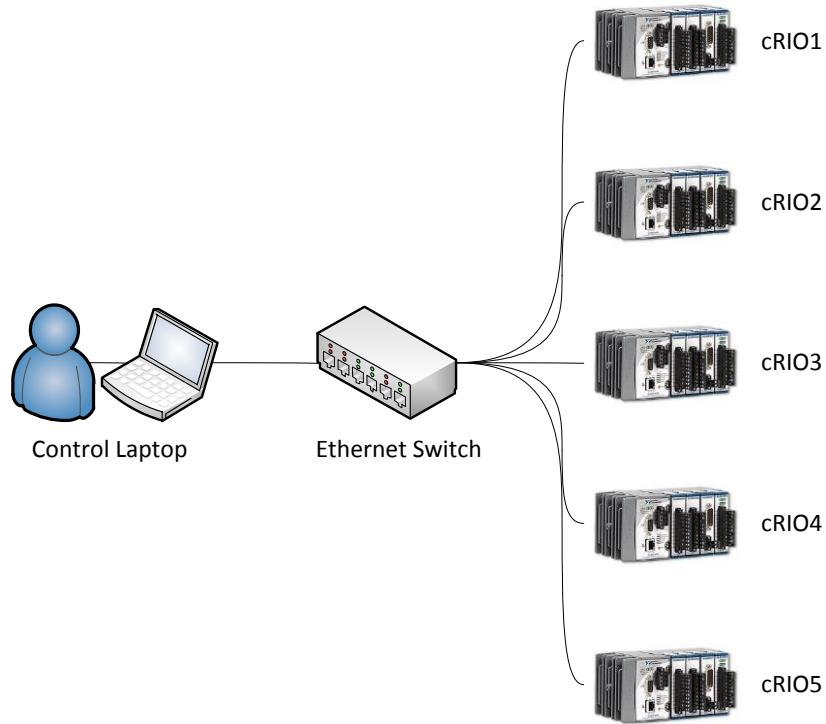


Figure 5-15 – Data Acquisition Network Diagram

The cRIO platform does not require a substantial amount of power. It was found that the entire system could be powered off of one generator, centralizing the point at which power cables would come down from the span to the ground. Note that it was necessary to maintain the gas level in the generator as any loss in power would result in a loss of data. The system used did not have any battery backup.

Since it was desired that the laptop used for visualization be located on the topside of the bridge during the test, only a single Ethernet cable and a single power cable were run up from the underside of the bridge. If a centralized system were used, every single data

cable would have been run up. Otherwise, the control laptop would have to be on the ground making test coordination more difficult.

5.6.2 Software Design

A major benefit of the National Instruments hardware is the flexibility of the Labview software which is used to control the hardware. While Labview has a substantial learning curve, the software offers unparalleled flexibility and power in terms of data acquisition, processing and visualization.

Labview is a programming language like C+ or Basic, except that it is visual as opposed to text-based. A program written in Labview can be deployed to run on any NI hardware as well as on a PC. There are commercial products on the market which are programs written in Labview that are designed to control NI hardware. Essentially, a refined user-interface analogous to the type of software that comes with most other data acquisition systems, these products are generally geared towards centralized data acquisition, and thus were not applicable to this situation.

There are also companies which work directly with NI and their customers to provide software solutions. At the present it was decided that this type of investment was not required. Instead the data acquisition software was developed by the Drexel researchers.

5.6.2.1 Software Architecture

While it was theoretically possible to use a single program located on a laptop to run all five data acquisition systems, this was found to be highly inefficient and in the end,

infeasible due to data transfer rates. Instead a hierarchy of programs was utilized. The software architecture mirrors the physical distributed data acquisition idea, and can be seen in Figure 5-16.

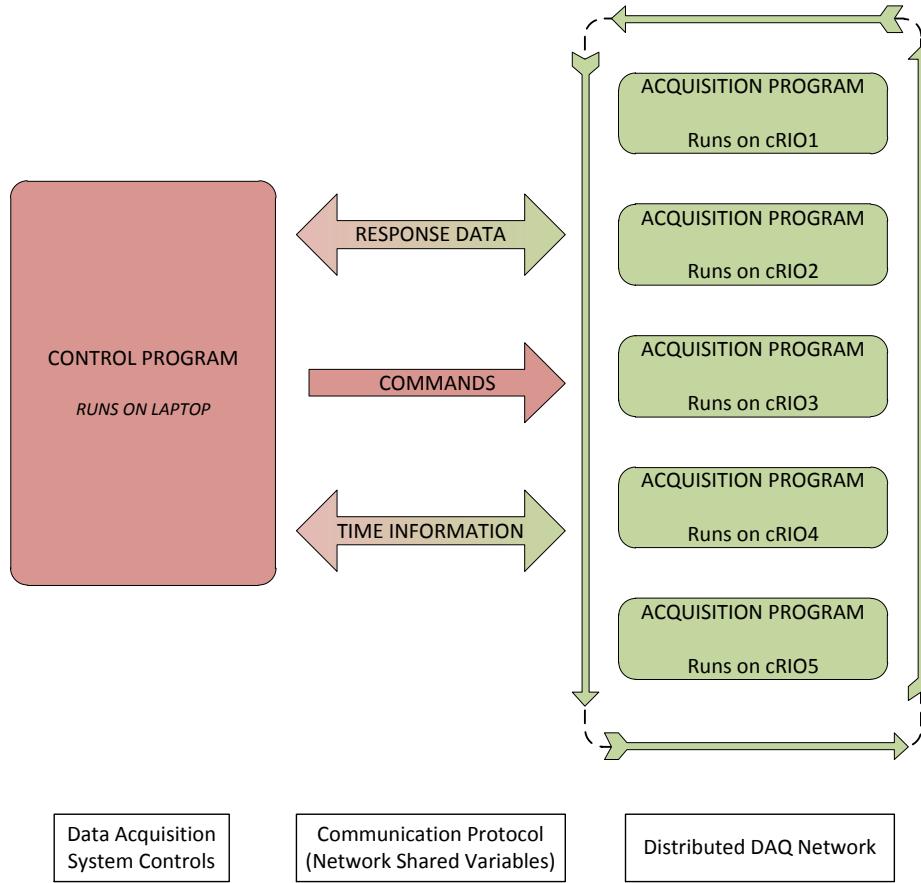


Figure 5-16 - Software Architecture

The Control program running on the laptop supplies commands to the Acquisition program running on each cRIO. Control polls the Acquisition program for response data in order to provide real-time feedback on the bridge behavior. Time information is also exchanged in order to synchronize the software.

5.6.2.2 Software Synchronization

The use of five data acquisition systems provided an interesting and previously unexpected challenge of synchronization. Five systems physically spread apart from each other and not physically connected directly needed to be synchronized to a level deemed acceptable for static and crawl testing. This threshold level of synchronization was decided to be 0.05 seconds based on previous experience and the static/quasi-static nature of the test.

As described above, the cRIOS were networked via ethernet cable and a gigabit switch. Some research on network timing indicated that there are two general methods for synchronizing over a network. One method is Standard Network Timing Protocol (SNTP) and the other is 1588 Network Timing Protocol (1588). 1588 provides nanosecond accuracy, but is very expensive in terms of the required hardware. The system is based off of GPS timestamps, and requires a 1588-ready ethernet switch.

SNTP uses a machine on the network as the time source and syncs all other machines to that source. In this case, the laptop was used as the timing source, and the five cRIO systems synced to the laptop. This provided accuracy on the order of 0.01 seconds which surpassed the threshold level. This timing protocol was embedded into the boot-up

routine of the cRIO systems, and a real-time check on timing accuracy was incorporated into the Acquisition and Control programs.

5.6.2.3 Data Storage

Since each cRIO will be recording data independently, a given test run will have five test files. These files are recorded locally on the cRIO. Transmitting the data, combining into a single data file, and recording it locally on the laptop were investigated, but it was found that this severely limited the sampling rate of the data acquisition. Each cRIO has 2GB of storage space, so for a short term application, storing data locally on the cRIO was an acceptable option. Therefore the Acquisition program locally stores data on the cRIO and sends selected data to the Control program for visualization. The stored data could be accessed at any time from a PC on the local area network via FTP protocol. After each round of testing, the data was downloaded from each cRIO to the laptop and backed up.

5.6.2.4 Real-time Data Visualization

The main reason that software development was done in-house was to incorporate a much more robust system for real-time data visualization. On past tests, it was found that quickly assessing the “quality” of the data as well as the overall performance of the bridge was difficult at best, and often not possible. At minimum, this quick quality assessment requires plotting time histories of crucial gages. However, a review of every single time history is time consuming if done individually, and unreliable if done en masse. Therefore, additional parameters can be used to verify the quality of data. One

parameter is the spatial relationship of the sensors to one another, the loading, and the bridge. An experienced engineer has a heuristic sense of the expected behavior of the structure for a given load condition. By looking at groups of gages with a simple spatial relationship (i.e. girder displacement profile), the engineer can confidently determine if the data appears to be reasonable for several gages at a time, as opposed to a single gage. This also instills confidence in moving forward with higher load levels, or potentially indicates that the test should be temporarily, or permanently, halted.

Substantial efforts were made during software development to alleviate this problem by providing intuitive, real-time visualization of the data during the test, including spatial variation. As such, the user interface was designed utilizing two screens; a laptop screen and an extra monitor. The laptop screen was purely used for controlling inputs, verifying DAQ connectivity, synchronization accuracy and gage functionality. The secondary monitor housed real-time plots of response of related gages, including the spatial component which enabled the use of the heuristics of an experienced bridge engineer to verify the bridge is behaving “as expected.” The data acquisition setup, including both screens can be seen in Figure 5-17. Details of the screens can be seen in Figure 5-18 and Figure 5-19.



Figure 5-17 - Real-time Visualization Setup

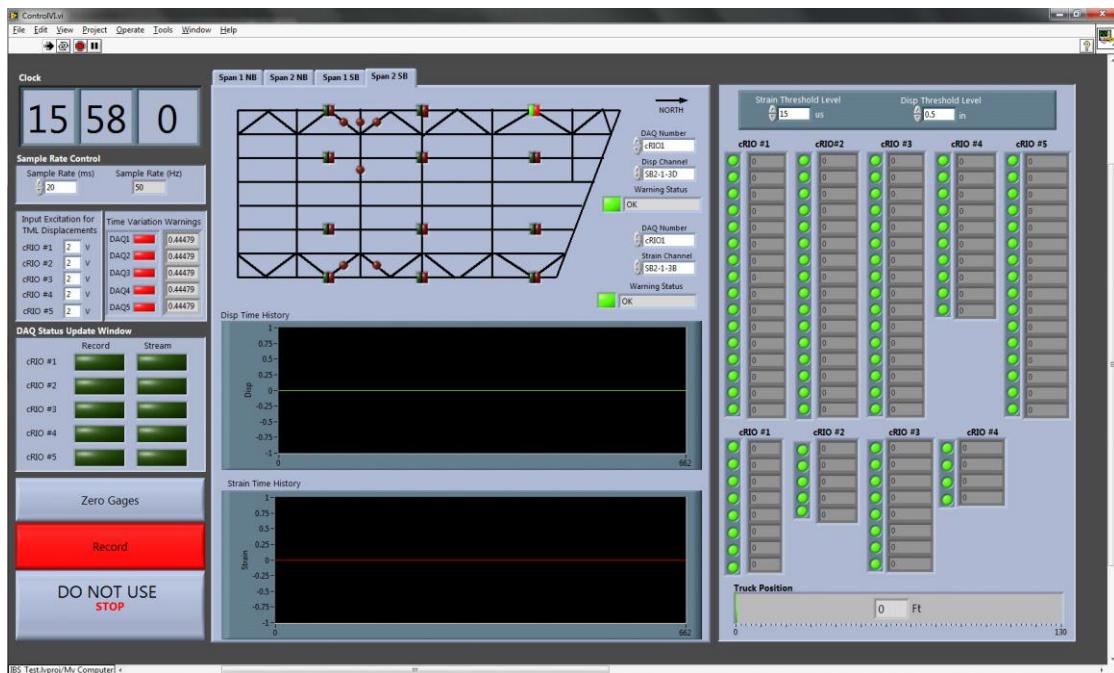


Figure 5-18 - Control Screen for Test Control

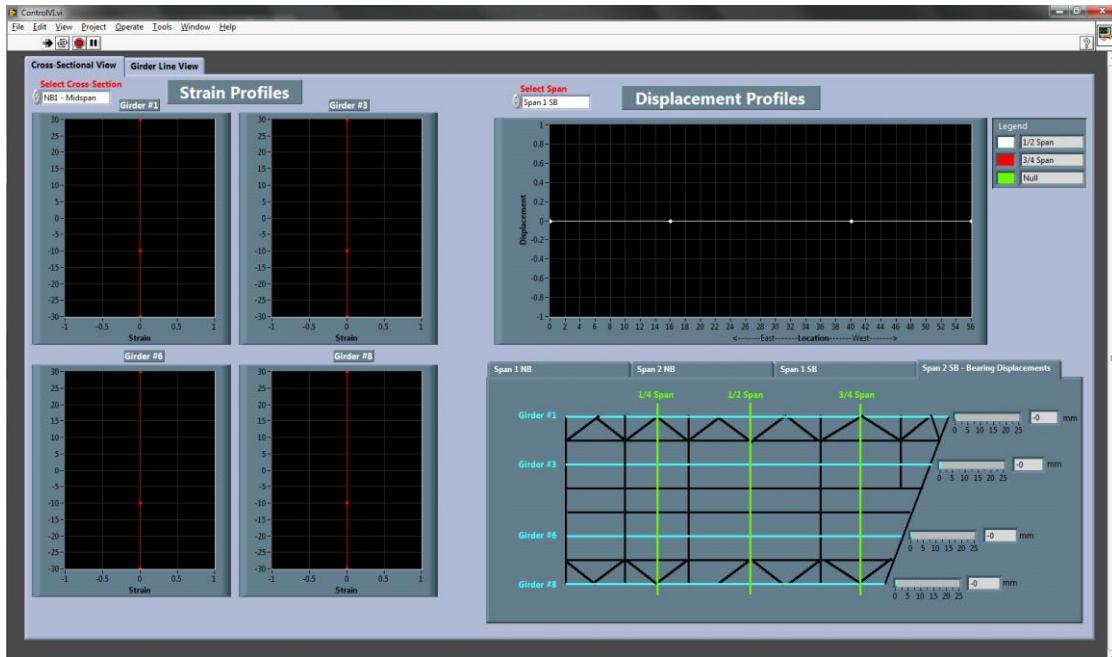


Figure 5-19 - Control Screen for Rapid Data Interpretation

5.7 Live Load Test: Execution, Processing and Results

The following sections discuss the execution of the live load test, the processing, reduction and visualization of the data, and the results as they relate to the critical questions identified in Section 3.9.

5.7.1 General Guidance for Test Execution

While a sound testing plan is imperative, it is also important to have both contingency plans and access to pertinent information to be able to make informed decisions regarding when to implement a contingency plan. The most obvious decision during a live load test is whether or not to proceed to each ensuing load level (which requires real-time data visualization). However, there are many other issues that often arise related to schedule

and data acquisition/sensors. From a schedule standpoint, it is not uncommon for traffic control or trucks to be delayed and thus a summary test plan should be developed beforehand in the case that the full testing program is not feasible due to delays.

From a hardware standpoint, it is imperative that the test engineers have a continuous understanding of the proper operation of all sensors and data acquisition units as well as an understanding of the “expected” range of responses. Whenever possible, data collection during ambient traffic prior to the test allows for troubleshooting of the system. If certain sensors are providing suspect data they may be supplemented or double checked prior to the final load stages. Similar to the scaled-back load testing program, a scaled-back instrumentation plan (or at least a hierarchy of sensor importance) should also be developed in case there are a significant number of malfunctioning sensors. It is good practice to always keep 10% additional sensors on hand to either add or swap out as needed. If all sensors are required, then the hierarchy of importance for the sensors may be required to rearrange sensors in between test stages to ensure that the most important responses be captured.

To ensure data quality, it is important to develop a series of data templates to facilitate the rapid visualization of data once it is downloaded following the test. While the real-time data visualization is quite important it is not sufficient to ensure that the data recorded (as opposed to scanned) is of sufficient quality. In addition, it is difficult to zoom or compare data sets from different load stages in real-time. Once the test stage is concluded the data should be immediately backed up and transferred to a separate computer for error screening. Prior to the test execution, a labeling scheme for all data files should be developed and all data should be stamped with date and time.

It is important to keep a test log during the test to record any anomalies identified or issues that arise so that they can be examined later. Towards that end it is useful to have personnel both top and bottom side of the bridge in communication through cell phones or radios to allow quick error screening of sensors and ensure the identification of any anomalies. In addition, it is good practice to always log the time at which each test stage started and ended (as well as general notes/observations) in case there is confusion with the file labeling scheme (as all data should be time-stamped).

5.7.2 IBS Test Execution

The static testing of US202/NJ23 was originally scheduled for the night of Thursday, September 30th but was delayed due to heavy rains. Instead the testing occurred on Friday, October 1st and lasted throughout the night and into Saturday Morning, October 2nd. There were complications which delayed the start of the testing on the southbound side of the bridge from 10pm until approximately 4am. A detailed time line of the entire night, developed based on the test log, is included in Table 5-3. Once initiated, the testing went smoothly, reaching the final load stage around 8am. The bridge, fully loaded with six trucks, can be seen in Figure 5-20.

Table 5-3 - Test Timeline

Time	Event
9:00 PM	Trucks and PB Representatives Arrive
10:00 PM	Trucks Measured; NB closure begins
11:00 PM	SB closure team arrives, but cannot set up
11:30 PM	Dynamic testing begins on NB
1:00 AM	SB team sent home; NB closure team moves to SB side
4:00 AM	SB closure is set up
4:15 AM	Three empty truck load stages
5:00 AM	Three full truck load stages
6:00 AM	Dynamic testing
7:50 AM	Six full truck load stage
9:00 AM	Complete

**Figure 5-20 - Span SB2 Fully Loaded with Six Trucks**

5.7.2.1 Discussion of Test Execution and On-site Adjustments

There are several logistical procedures and some unexpected variations from the designed test plan that are worth discussion.

On the evening of the test, well before traffic control was in place, the project team met to discuss everyone's roles throughout the night. By assigning responsibilities prior to the test, the engineer responsible for running the test was free to focus on preparing the data acquisition system, confident that everything else that needs to happen is being handled.

Some major personnel responsibilities include:

1. Truck Documentation
2. Test Log and Photography
3. Truck Positioning
4. Underside Sensor/Cable Troubleshooting
5. Power Maintenance (Gas for Generators)
6. Food/Coffee for Entire Team

The importance of some of these tasks may seem minimal, but during an overnight test, it is easy to forget to eat, or to check the gas level in a generator. These types of oversights can be very detrimental to the test and the on-site engineers.

In order to expedite the useable time in the traffic control windows, it was requested that the trucks arrive to the site adjacent to the structure thirty minutes early. This allowed the test engineers to measure and weigh the trucks in detail without impeding traffic. These measurements include all wheel spacing and individual tire weights. Additionally, the trucks are all clearly numbered so that each truck's position during the load test can be readily determined and recorded.

Traffic control installation began at 10PM for the northbound direction, but southbound started later (11PM) due to the late arrival of the crew. The southbound crew was plagued

with equipment issues which prevented a safe closure of the desired lanes, and in the end, were not able provide the closure at all. This was officially decided at 1AM after exhausting all alternative options, which set back the testing efforts a great deal.

Considering the focus of the effort was on the static testing of the southbound side, it was decided that dynamic testing on the northbound side would be stopped, and the crew from that closure would move to the southbound immediately. This transfer was complete at 4AM, meaning that access to the span for the static testing began a full 6 hours later than expected. Additional complications arose from the fact that the only remaining crew did not have the exact equipment required for the southbound shutdown as they were hired to provide the closure on the northbound span. Therefore, the closure plan varied from the original design. In particular, it was not possible for them to provide space off the spans of the bridge for maneuvering and aligning trucks prior to loading. As a result, the zeroing of the sensors took place with the trucks on Span 1 and Span 3 as opposed to being entirely off of the bridge, which would have been preferable.

The first several load stages progressed without event, and the empty trucks were sent to be loaded. In this time window, dynamic testing of the southbound side was conducted.

See Prader (2012) for an in-depth discussion on the dynamic testing of the IBS Birdge.

Due to the shift in schedule, there was substantially more traffic in the single open lane, which has an adverse effect on the quality of dynamic data.

When the three loaded trucks returned, dynamic testing was completed and the final load stage was conducted. Throughout the static loading process, some truck positions were held for extended periods of time to allow for the visiting researchers to make non-contact displacement measurements. Static testing was concluded at 9AM which was the

latest the test could be run, considering traffic control had to be removed by 10AM at the latest.

5.7.3 Data Reduction and Visualization

Perhaps one of the most under-appreciated aspects of experimental studies is the processing, data reduction and data visualization that typically takes place “behind the scenes”. The goal of these aspects is to transform and organize the data obtained during the test (which is typically in extremely large arrays) to better facilitate interpretation. To put into perspective the amount of data collected for a static test, the data for final load stage of the IBS test consisted of five separate data files of over 140,000 data points each. The general approach is to reduce or compress the huge amount of data obtained to the most critical response indices, which may be a specific strain or displacement under a specific load, an influence line or coefficient, or modal parameters in the case of dynamic testing. Once these critical responses have been extracted from the data they may be plotted/visualized in many ways to allow trends and relationships to be identified spatially, temporally, or across different load levels/configurations. In addition, these response indices are ideally suited for model-experiment correlation. Accomplishing this typically involves (1) data processing (including zeroing, averaging, filtering, etc.) and feature extraction, and (2) plotting.

5.7.3.1 Data Processing and Feature Extraction

The following is a general description of the process used to reduce the large data files from truck testing into a manageable format. These steps generally occur in MATLAB because of the size of the files and amount of data being manipulated.

1. Visual assessment of time histories – The time histories for related sensors are reviewed on a single plot for any glaring inconsistencies or errors. Examples of inconsistencies include large spikes in the data, pre- and post-loading offsets, and time lags between sensors.
2. Determination of relevant data windows – During the test, it is possible to record data windows during each load stage in a preliminary fashion. This can save time after the test, as these windows only need to be verified.
3. Manual zeroing of data – If there is any offset in the data, this is manually removed by subtracting a constant from the entire array of data, or depending on the size of files, the relevant data windows.
4. Averaging of relevant time windows – Each data window corresponding to a particular load stage is averaged to reduce the time history data to a single data point.
5. Tabulation of averaged static data – All of the reduced data is compiled into a tabular format which can be directly exported in Excel if desired.
6. Filtering crawl speed test data – The time history for crawl speed tests are kept intact, but are cleaned up through filtering or a moving average to make the data more readily interpretable.

7. Maximum and minimums for crawl tests – Time history data can be automatically reduced to maximum and minimum values for a given run. This serves as an error screening on the quality of the data, as well as a very quick classification of the responses.
8. Event analysis of ambient data – Ambient windows can be automatically processed to tabulate the number of events within a given response range, giving information about service loads and the cross-section of typical traffic.

5.7.3.2 Plotting and Organization

As discussed in relation to real-time visualization of data, the spatial relationships of data provide insight not only during the test, but also through careful examination post-test. The same families of plots are used during data reduction and analysis as are used during the test, with the exception that more in-depth and useful comparisons can be drawn. These relationships are given an added temporal dimension after the test as well. The most important time-sensitive variable is the loading position and magnitude, but there are other time-varying considerations as well. These other time varying conditions are especially crucial in dynamic testing and will be discussed in the ensuing section.

Considering the number and type of sensors, truck loads levels, and truck positions, there is a huge number of combinations of parameters which could be selected to describe the bridge response through several different relationships. Considering this, an organized and accessible approach to developing the plots of these relationships is required. In order to facilitate drawing conclusions from groups of related plots, a matrix plotting style was utilized. Essentially, families of plots were grouped together into a matrix, $[M_{ij}]$ with

term M_{ij} corresponding to different combinations of parameters. For example, in the case of the family of plots showing point displacements across load levels, M_{ij} might equate to load at position i with response at position j . This is shown pictorially in Figure 5-21. The plots on the diagonal are the cases where the load and sensors were at the same position, which are the most readily interpretable.

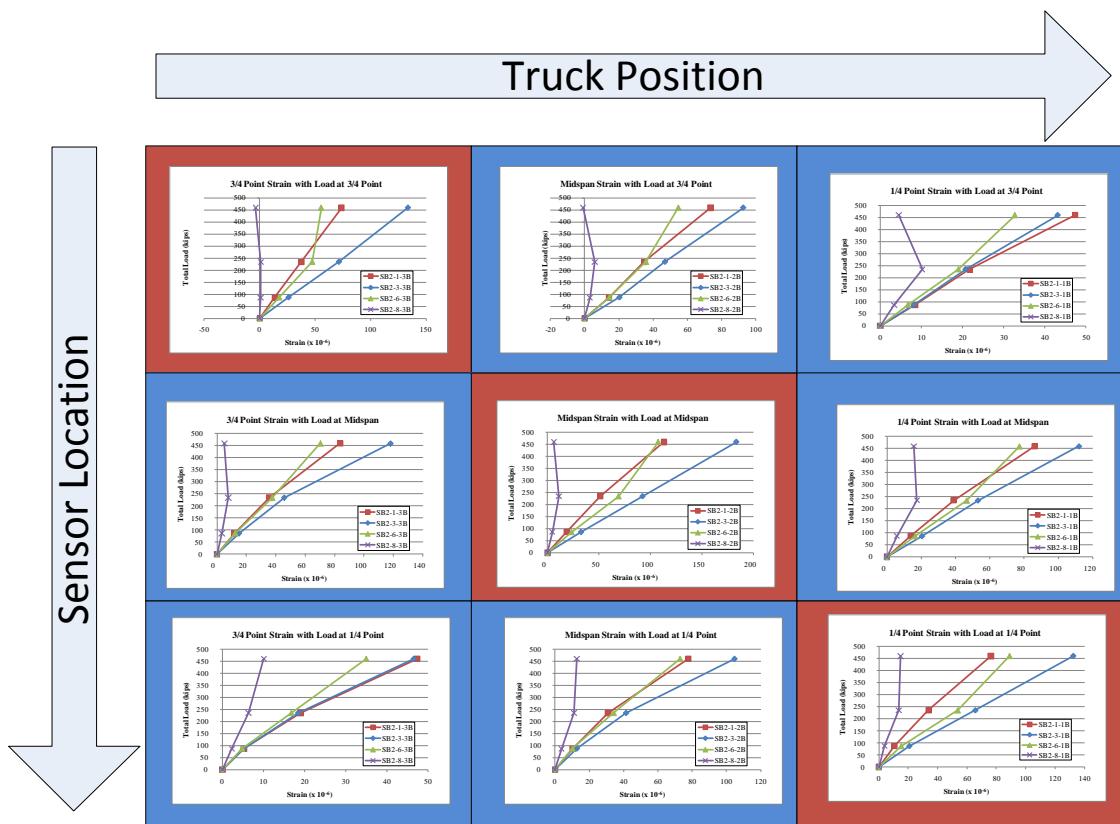


Figure 5-21- Plot Matrix for Displacement

Arranging similar plots in this way allows for rapid, visual interpretation of relationships between several parameter combinations at the same time. The following relationships were explored after the test to examine the quality of the data as well as to interpret and draw conclusions in regards to the performance of the structure:

- Linearity of strains through various load stages
- Linearity of displacements through various load stages
- Girder-line strain profiles
- Girder-line displacement profiles
- Transverse strain profiles
- Transverse displacement profiles
- Cross-sectional strain profiles
- Strain to displacement ratios through various load stages
- Displacement basin (3D plot of displacement)
- Strain basin
- Time histories

5.7.4 Framework for Data Interpretation

Translating the reduced data into valid and defendable conclusions regarding the performance of the structure requires that the critical questions developed previously be addressed by any and all relevant data sets. For the IBS investigation, the following critical questions were developed in Section 3.9 based on documentation reviews, on-site investigations and a qualitative risk assessment of the structure.

1. What is the influence of the observed fatigue cracking on the load carrying mechanisms of the bridge? Can the fatigue sensitivity be mitigated by removing the wind bracing? Could this potentially cause additional problems?
2. Is the observed crack in the pier cap due to live load or the excessive vibration of the bridge? Is this crack active under service load? Does it indicate a safety issue?
3. Given the various skew arrangements of the bridge and the resulting bridge movements under temperature and live load (including potential uplift and rotation) are the current rocker bearings sufficiently reliable?
4. Is the excessive vibration a result of the initial design or is it being exacerbated by accumulated deterioration? What options are available to reduce the level of vibration?
5. What is the relationship between the observed deck cracking and the excessive vibrations of the bridge? Is it possible that the deck is being placed under net tension due to vibrations?

These critical questions inform the data interpretation process by providing a framework through which the data can be utilized to directly address the critical issues surrounding the bridge. This allows for the integration of the different types of tests and data (i.e. static, crawl speed, and ambient) to draw unified conclusions. The relationships mentioned in Section 5.7.3.2 are organized based on which critical questions they directly inform, resulting in a robust data interpretation framework, outlined below. Each question is listed, followed by the relevant data relationships in Table 5-4.

Table 5-4 - Addressing Critical Questions through Data Visualization

Critical Questions	Test Type	Plot Utilized	Plot Usage
Fatigue cracking and wind bracing details	Static	Wind Bracing Response vs. Truck Position	Member force under static load
	Ambient	Wind Bracing Response - Ambient Traffic	Activation of wind bracing under non-lateral service loads
	Crawl	Diaphragm Response vs. Truck Position	Transverse load distribution
	Ambient	Diaphragm Response - Ambient Traffic	Transverse load distribution - service loads
Pier cap cracking	Ambient	Crack Opening - Service Response	Crack activation under service loads
	Static	Crack Opening vs. Load Level	Linearity of crack opening response
	Static	Girder Displacement vs. Load Level	Linearity of displacement response
	Static	Girder Strain vs. Load Level	Linearity of strain response
Bearing reliability	Static	Bearing Displacement vs. Load Level	Linearity of bearing movement
	Static	Girder Displacement vs. Load Level	Linearity of displacement response
	Static	Strain Response vs. Load Level	Linearity of strain response
	Static	Displacement Basin	Global system response
	Static	Strain Basin	Global system response
Excessive vibration	Static	Composite Action	Neutral axis location
	Ambient	Girder Response - Ambient Traffic	Response under service loads
	Static	Displacement Basin	Global system response
	Static	Girder Strain vs. Load Level	Linearity of strain response
Deck cracking	Static	Composite Action	Neutral axis location

5.8 Peak Responses

Under the final load case of 6 full trucks, the measured peak values of strain and displacement were $183\mu\varepsilon$ and -0.845 inches respectively. Both of these measurements were located at the midspan location of Girder #3, as was the load. There is no guarantee that these responses were the peak responses for the structure. In fact, because of the skew, it is likely that these were not the actual peak responses of the structure. This is an important distinction to make, as it can be dangerous to assume that one measured the actual peak response simply because heuristics and theory indicate the response would be maximum at the selected location. The recommended deflection criterion of L/800 corresponds to a displacement of 1.83 inches which is much greater than the actual measured response of 0.845 inches. This displacement over the length of Girder #3 (122ft) corresponds to a deflection ratio of L/1732, which exceeds even the more stringent L/1200 criteria to which the structure was designed. Table 5-5 shows the peak responses of all four instrumented girders.

Table 5-5 - Peak Responses

	Strain (x10⁻⁶)	Disp (in)
SB2-1-3	83.9	-0.467
SB2-1-2	113.1	-0.639
SB2-1-1	86.2	-0.625
SB2-3-3	133.7	-0.530
SB2-3-2	183.6	-0.846
SB2-3-1	132.3	-0.616
SB2-6-3	70.4	-0.249
SB2-6-2	107.7	-0.447
SB2-6-1	89.0	-0.358
SB2-8-3	10.1	-0.024
SB2-8-2	12.6	-0.085
SB2-8-1	17.4	-0.077

5.9 Linearity Checks

It is important to determine whether the bridge remained in the linear range of load-response behavior throughout the test. This is checked by comparing a single response at several varying load levels. These checks should be carefully considered because of the variation of the number and exact location of the trucks as well as changing ambient conditions.

5.9.1 General Linearity Checks

Figure 5-22 shows the linearity check of load vs. displacement for the midspan location of all four instrumented girders under the load at midspan. Girders #1 and #3 show some slight softening. Girders #6 and #8 exhibit apparent stiffening. As load increases, nonlinear behavior due to any type of plastic deformation would lead to a flattening of

these plots as seen with Girders #1 and #3. However, #6 and #8 show the opposite. This indicates that the nonlinear behavior was likely due to a change in how the load was distributed at the higher levels, as opposed to plastic deformation.

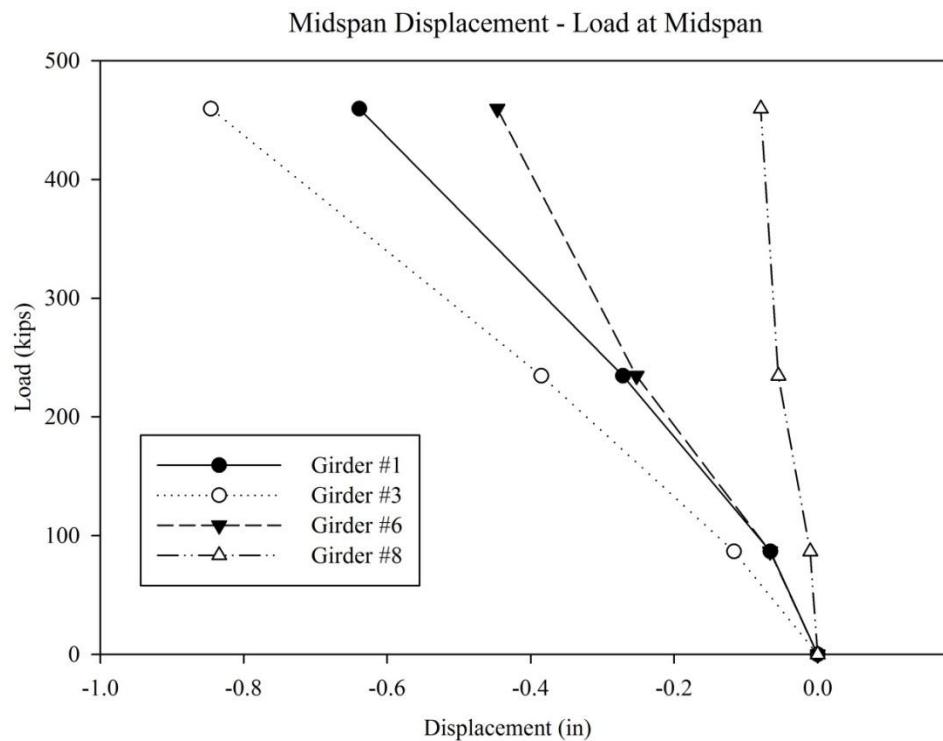


Figure 5-22 - Displacement Linearity Check

Figure 5-23 shows the same relationships are present when comparing bottom flange strains with load. Girders #1 and #3 present a small amount of softening, while #6 and #8 show a stiffening behavior which is caused by the phenomenon described previously.

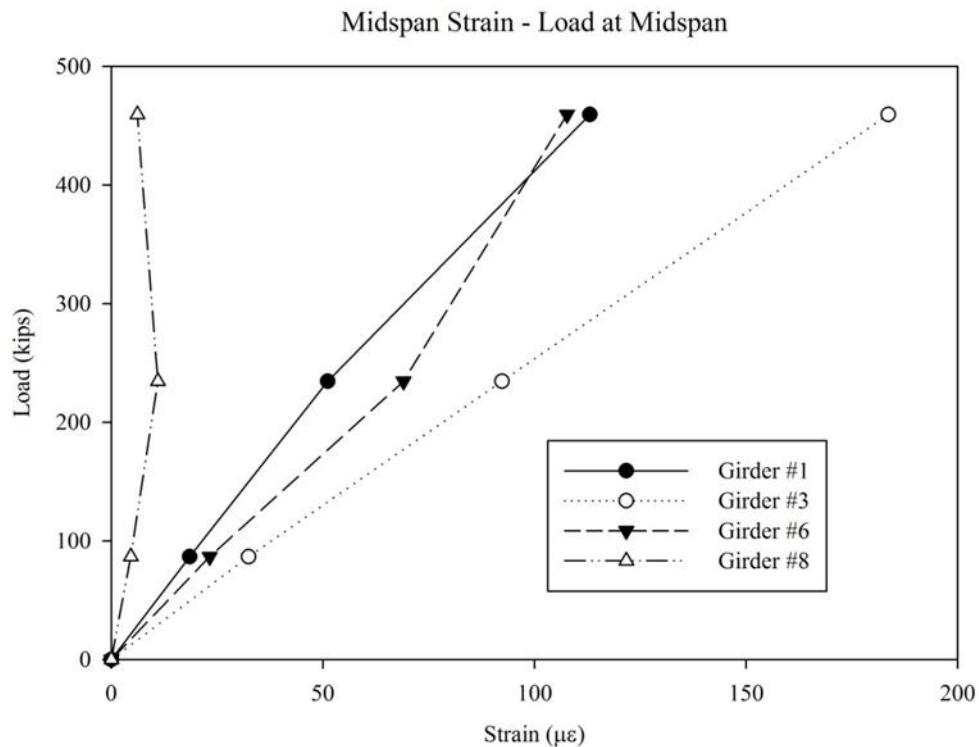


Figure 5-23 - Strain Linearity Check

5.9.2 Observations from Linearity Checks - Softening of Girder #1

Other load configurations and positions showed more apparent softening. In particular, when the six trucks were positioned at the one-quarter line, the displacement on Girder

#1 more softening than can reasonably be attributed to load redistribution. This can be seen in Figure 5-24. The other three girders behaved similarly to the previous case, showing the same stiffening/softening behavior.

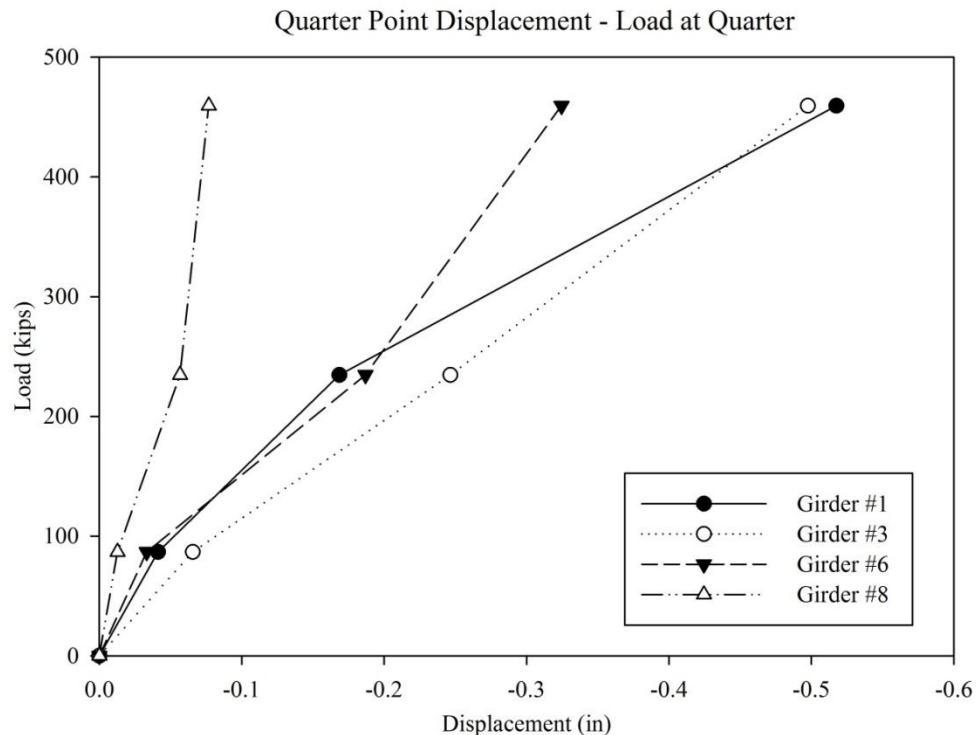


Figure 5-24 – Disp. Linearity Check for Load at 1/4 Point

In this case the strain behavior, shown in Figure 5-25, does not directly mirror the displacement behavior. Girder #1 shows only a slight amount of softening. Considering these two plots together, it is safe to assume that the major softening seen in the

displacement data was likely due to a support settlement of some type, as opposed to a softening of the structure. A support settlement would manifest particularly in displacement response since it is a rigid body motion, not an elastic deformation. It is hypothesized that this displacement response was a result of the crack in the piercap near Girder#1.

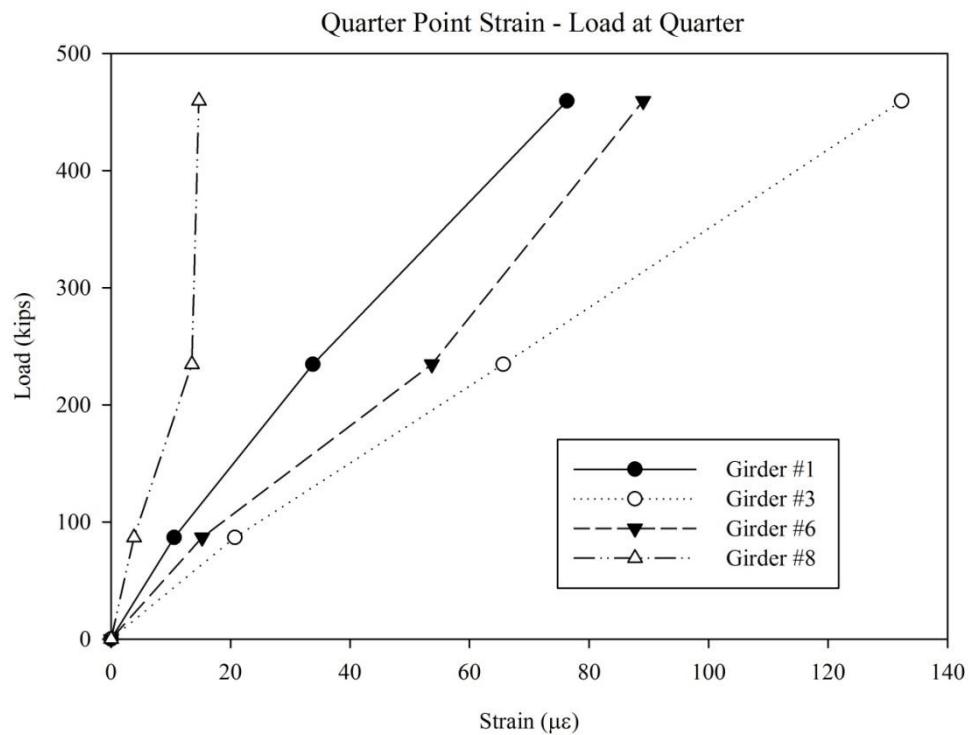


Figure 5-25 - Strain Linearity Check - Load at 1/4 Point

5.9.2.1 Observations of Linearity Checks - Uplift of Girder #8

Another phenomenon which was observed in the linearity data is a small amount of sign reversal, which could correspond to uplift, particularly in relation to the skewed end of the bridge. While it is unlikely that the span would overcome the dead load reaction at the bearing, it certainly could be unweighted. Figure 5-26 shows the strain response of the gages at the $\frac{3}{4}$ line with the load at the same location. Girders #1 and #3 are nearly linear, while Girder #6 shows the apparent stiffening seen previously. Girder #8 begins to show compression response on the bottom flange in the highest load stage, indicating negative bending. This is predominately an effect of the skew, and will be explored thoroughly in the model correlation process.

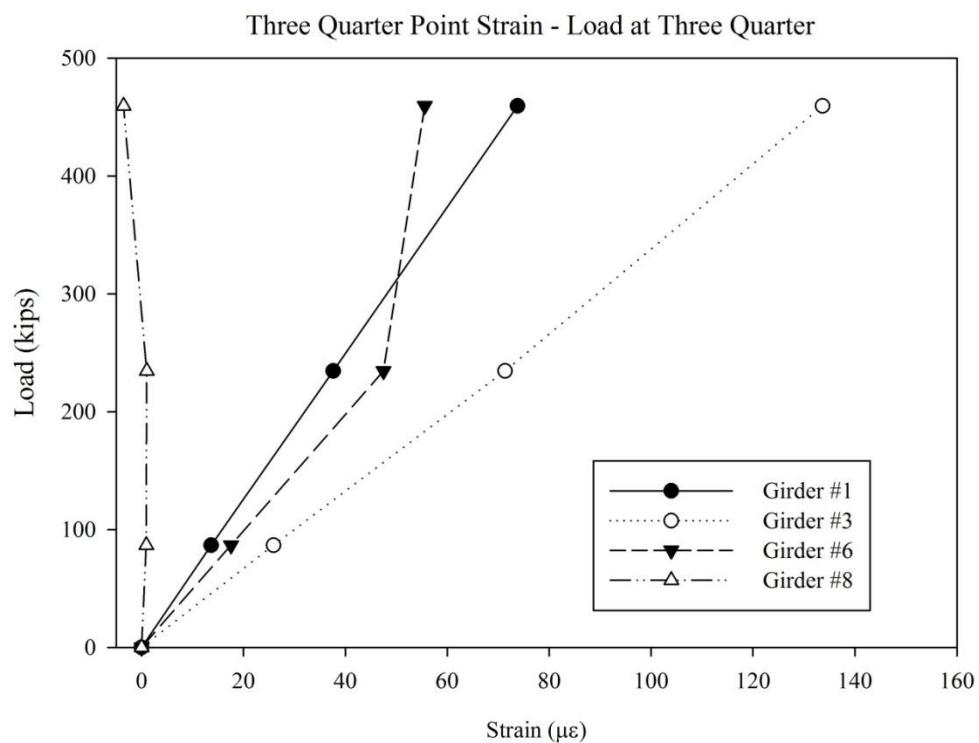


Figure 5-26 - Strain Linearity Check - Load at 3/4 Point

5.10 Global Behavior and Spatial Relationships of Data

5.10.1 Global Displacement Basin

The overall displacement basin for the entire structure with the load at midspan is shown in Figure 5-27. The points representing the locations of the bearings were not directly measured and are assumed for the purpose of plotting the entire displaced shape. The displacement basin shows behavior as one would expect for a typical simply supported structure. The lack of appreciable response at Girder #8 can be attributed to several factors. First the load was applied mostly over Girder #2 through #6. Also, this Girder #8 is directly beneath a concrete barrier on the roadway, which adds stiffness. Lastly, Girder #8 is 25' shorter than Girder #1 because of the skew, meaning that the displacement would be lower on this girder regardless, and that the load is located towards one end as opposed to the middle.

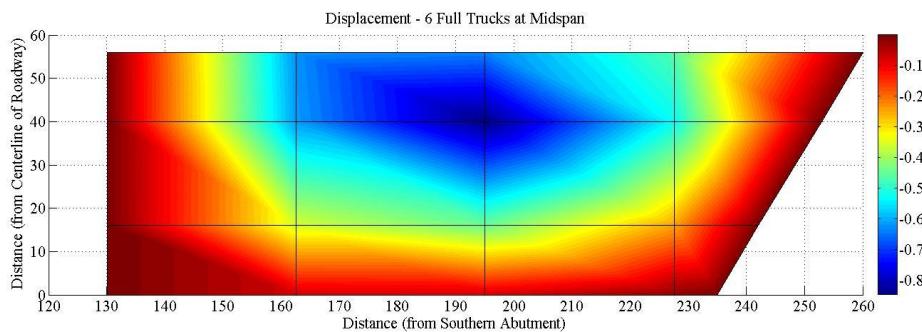


Figure 5-27 - Overall Displaced Shapes

5.10.1.1 Girder Line Displacement Profiles

The displacement profiles for each girder under all load positions were plotted and reviewed. Figure 5-28 shows the profile lines of all 4 instrumented girders with the load at midspan, the same case as the displacement basin previously shown. The distinction between the length of the girders and the maximum responses can be seen clearly in this Figure 5-28. Girder #3 has the largest displacement response, despite the fact that Girder #1 is longer, as a result of transverse load position. Also, note that Girder #1 has a somewhat asymmetric response. This could be a result of the apparent softening of Girder #1 as discussed previously.

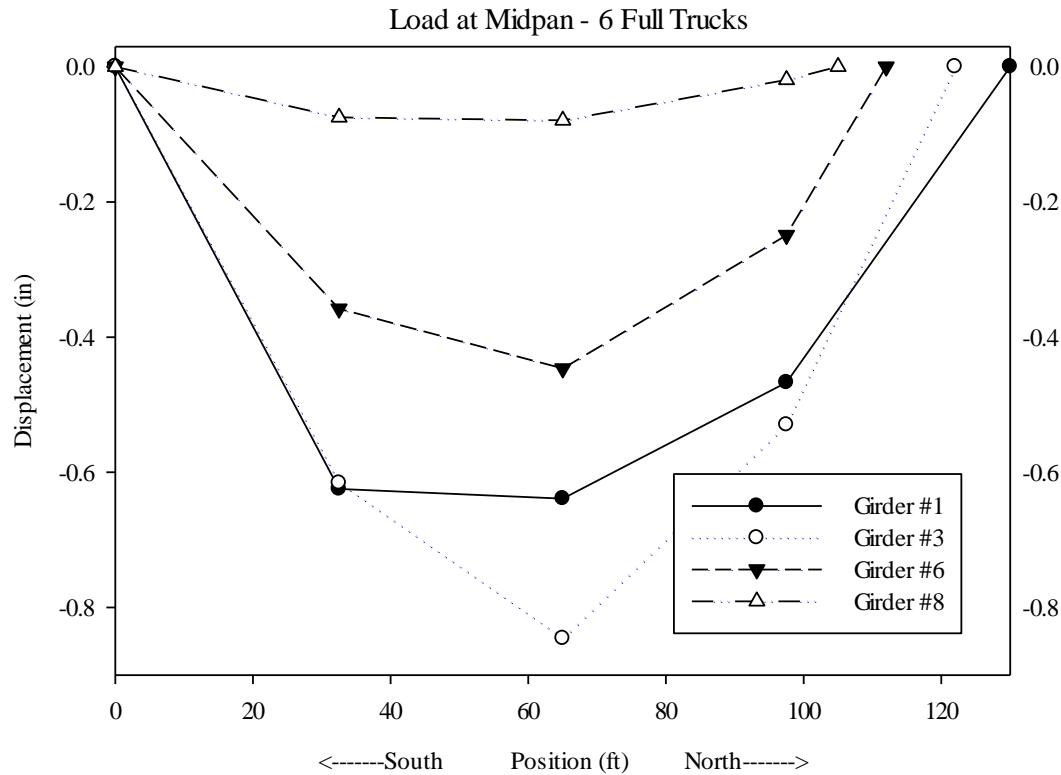


Figure 5-28 - Girder Disp. Profiles

5.10.1.2 Transverse Displacement Cross-sections

The conclusion that the load was redistributed, creating an apparent stiffening of two instrumented girders and softening of the other two, is also supported by the plots of transverse displacement. Figure 5-29 shows the transverse section at the $\frac{1}{4}$ line when the loads were at the same position. Qualitatively it is immediately clear by the change in the shape of the displacement profile between the load stages that the load has been redistributed. Quantitatively, the percent increase in each response does not remain constant across the four different locations measured between load stages. In other words, the response of Girder #1 increases by 0.16 inches from 3 empty trucks to 3 full trucks, and 0.41 inches from 3 full trucks to 6 full trucks. This represents a 156% increase. Conversely, Girder #6 increased by 0.15 inches and 0.17 inches respectively, representing a 13% change. Additionally, Girder #8 actually shows a decrease in the amount of additional displacement, decreasing from 0.04 inches to 0.01 inches. This trend indicates that Girder #6 (and Girder #8) is taking a smaller percentage of the total load in the final load stage.

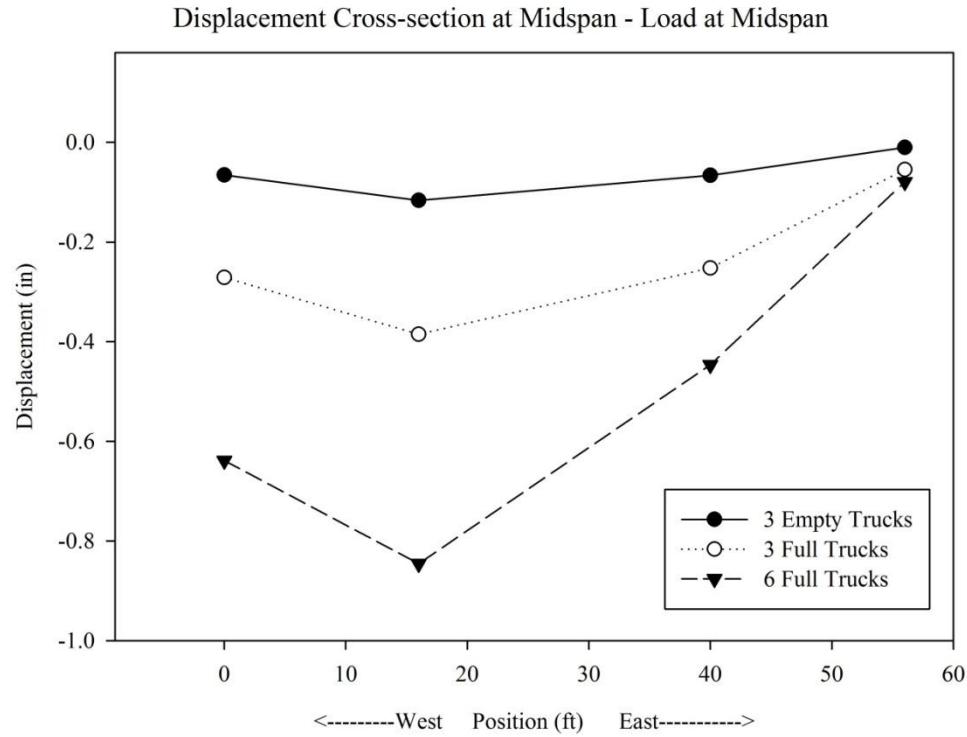


Figure 5-29 - Transverse Displacement Plot

5.10.1.3 Girder Line Strain Profiles

Figure 5-30 shows the strain profiles along the length of the four instrumented girders on Span 2 SB. These can be compared qualitatively to the shape of the analogous plots of displacement. In this case, it is clear the strain profiles appear to be nearly symmetric, while the displacements were asymmetric. This reinforces the conclusion that the softening of Girder #1 is due to some rigid body movement at the support, as the strain seems unaffected.

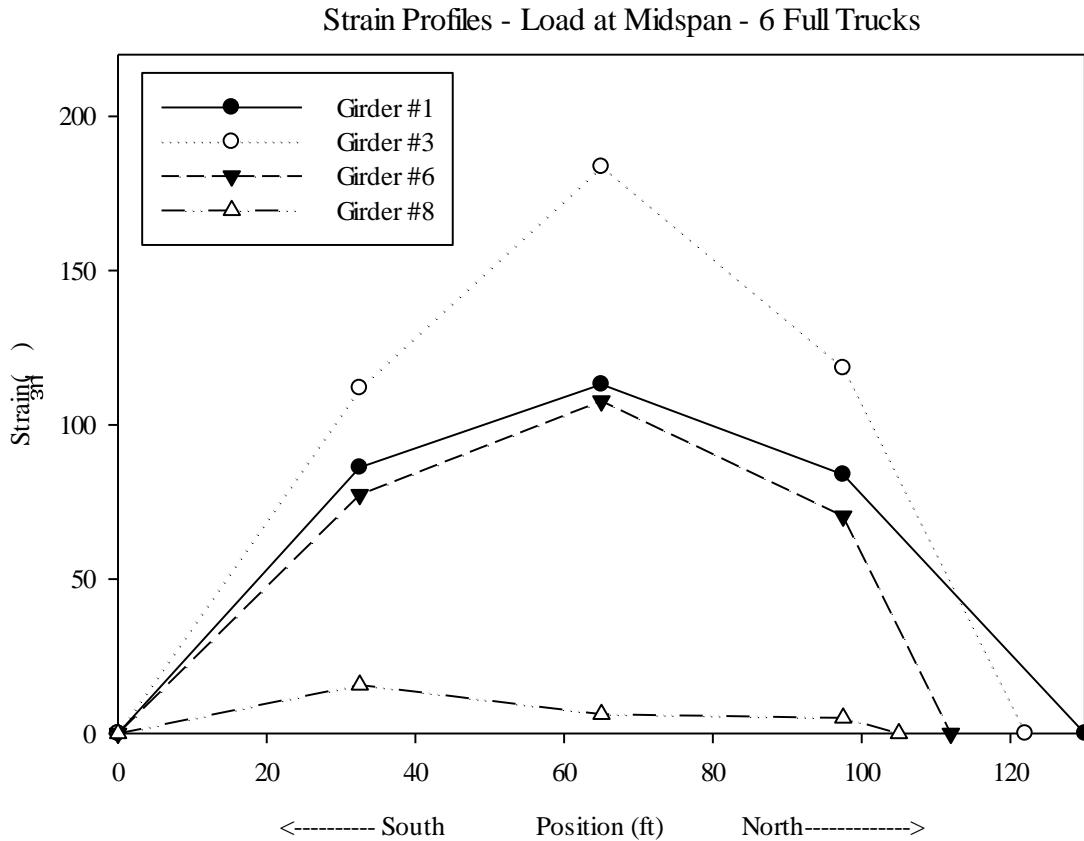


Figure 5-30 - Strain Profile along Girders - Load at Midspan

5.10.1.4 Transverse Strain Cross-sections

The strain cross-section, seen in Figure 5-31, is similar in qualitative shape to the displacement plots, except for the sign reversal. Again, Girder #3 shows the largest response, and the relative difference between load cases indicates a redistribution of load.

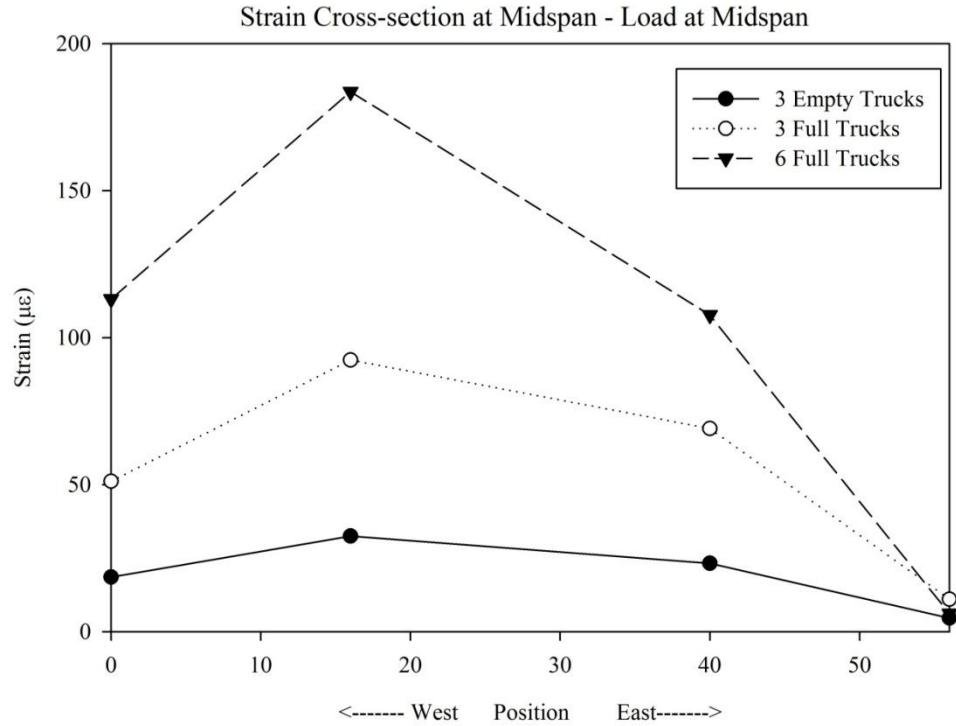


Figure 5-31 - Strain Cross-section at Midspan

5.10.2 Observations of Girder Line Profiles – Loss of Composite Action

Each instrumented location had two longitudinal strain gages at different heights along the cross-section of the girder, allowing for linear interpolation and a prediction of the neutral axis of the composite section. This value was then compared to the theoretical non-composite and fully composite neutral axes. In general, the neutral axis was located between the expected bounds, however some interesting behaviors were observed. The strain profile for every instrumented location was plotted and compared to these theoretical bounds, as in Figure 5-32.

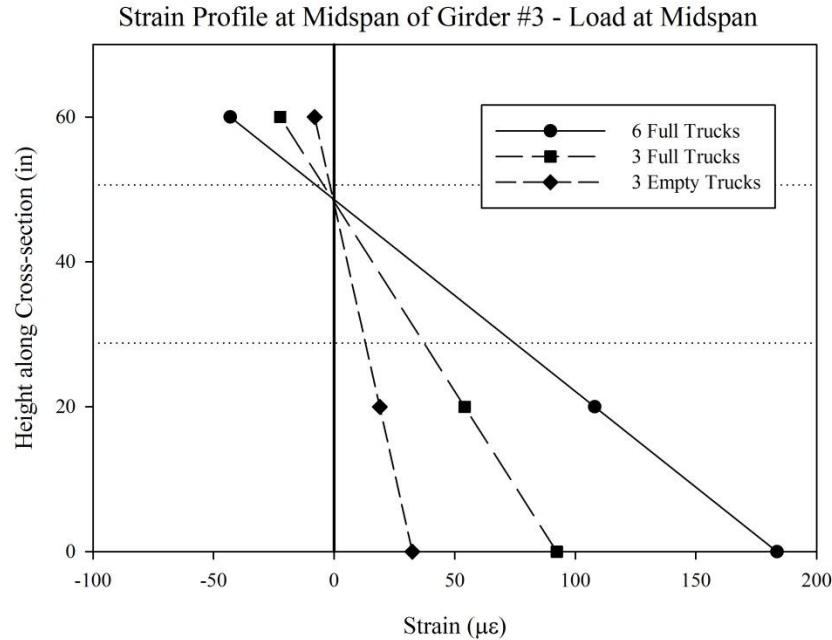


Figure 5-32 - Compatibility Plot - Girder #3 Midspan

5.10.2.1 Movement of the NA

Several instrumented locations showed movement of the neutral axis depending on the position of the load. Typically the neutral axis was located closer to the fully composite location when the load was placed directly over the sensor location. As the load was moved farther from that location, the axis shifted down towards the non-composite axis location. This indicates that there can be substantial load transfer between deck and girders due to friction when there is a large normal force present. In the cases where the section was fully composite, there was no movement of the axis location as the load moved at all.

5.10.2.2 Neutral Axis Located Above the Fully Composite Axis

Several locations showed a neutral axis location above the theoretical fully composite axis location. Some variability is expected due to the empirical assumptions required (Modulus of elasticity of concrete, effective width) and unknown parameters (haunch height) which add uncertainty to the calculated values. However, one particular case showed a very large net tensile force in the girder which was too great to explain with assumptions.

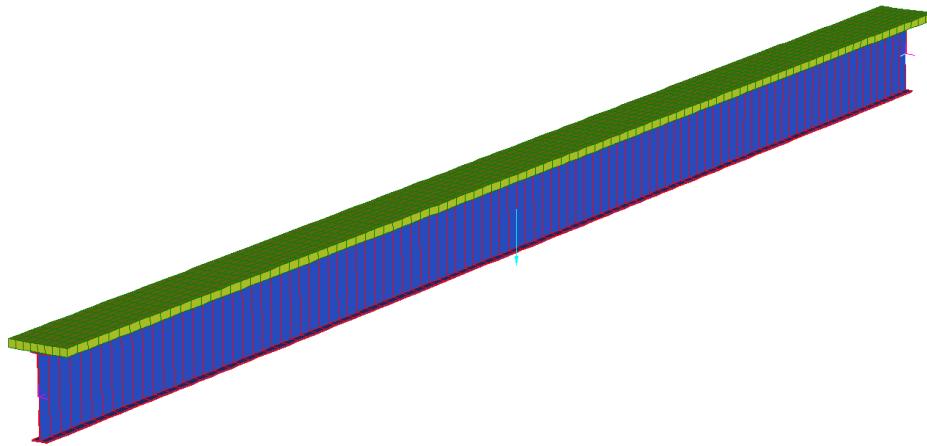


Figure 5-33 - Extruded View of Benchmark Beam Model

To investigate this behavior, a simple finite element model of a single girder and deck with a typical effective width was developed. The model, shown in Figure 5-33, was the

same simple beam-shell model connected with links that was used in prior benchmark studies. As a check on the validity of the model, a tributary portion of the total truck load applied to the model for comparison. The model response for displacement is reasonable when compared with the load test results. Figure 5-34 shows the axial and moment diagrams for full transfer between the deck and girders. Note that axial is on the top and moment on the bottom. The two mirror each other because the only loading present is bending, and the section is full composite.

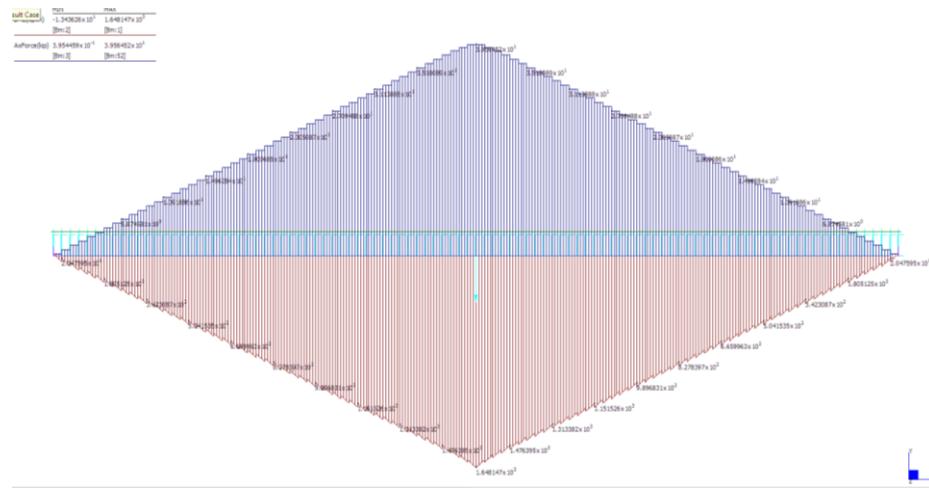


Figure 5-34 - Axial/Moment Diagram for Benchmark - Full Transfer

In order to try and replicate the desired behavior, the out-of-plane stiffness of the vertical links was removed for varying lengths of the girder. It was found that approximately 1/3 of the length of the girder had no composite action towards one end of the beam,

excluding the diaphragm connection at the end of the beam; the behavior seen during the test can be replicated.

The area of reduced connectivity is highlighted in Figure 5-35. In this case there is a region of constant axial force and linearly varying moment. Depending on the load level and position, it would be possible for the axial strain to greatly exceed the bending strain, resulting in the behavior seen on IBS.

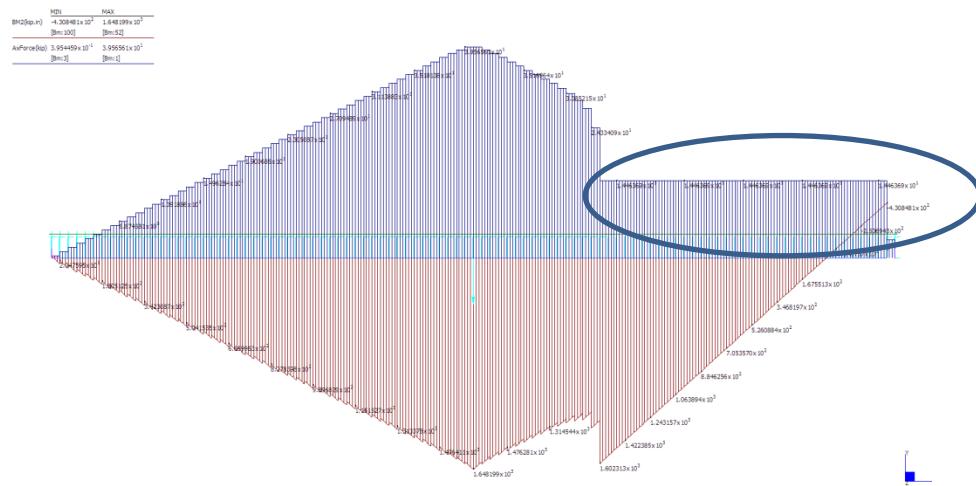


Figure 5-35 - Axial/Moment Diagram for Benchmark – No Transfer b/w Girder and Deck

Considering this, a possible explanation for the large net tensile force in Girder #3 at the $\frac{1}{4}$ line is a lack of composite action in the surrounding region, excluding the diaphragm at the end of the beam near the joint, around which the deck was cast.

5.11 Piercap Crack

Girder #1 is located directly adjacent to the large crack in the piercap shown previously in Figure 3-9 and bears directly on the cantilevered portion of the piercap just outside of the crack. This crack was monitored using a clip gage to track any response. The data on this gage was highly affected by the interaction with the adjacent span as well as the input from traffic in the open lane, but it does show activity during static loading and crawl tests as well as ambient traffic. The gage was installed on the lower portion of the pier cap, where the opening might not be as large as upper potion, considering the negative bending that could be expected. While directly tying the magnitude of the response of this clip gage to the softening of Girder #1 is not possible, the accumulation of plastic deformation of the crack is a reasonable explanation for the nonlinear behavior.

In order to explain the presence of this crack, the detailing in the original drawings was investigated. The rebar layout of the piercap, according to the drawings, is shown in Figure 5-36.

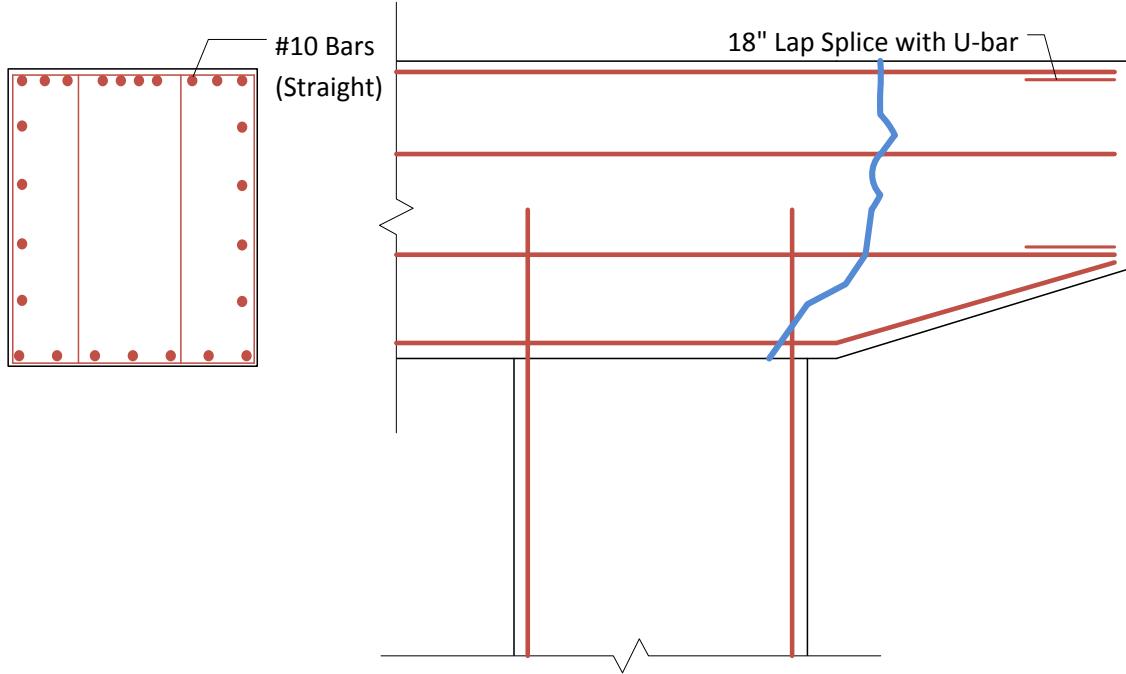


Figure 5-36- Piercap Rebar Layout

The top of the cap is reinforced with straight #10 bars, which do not terminate in a bend. There is a U-shaped bar with lap splices at the ends of the cap, but no mention of any mechanical connection between these bars. Therefore, due to the size of the bar and lack of any bends, it is clear that the bar was not fully developed. In other words, the concrete is unable to transfer load to the reinforcing bars. The calculated development length, using the ACI code, is 90.5in while the distance from the end of the piercap overhang to the center of the column is only 72in.

Considering that Girder #1 is not directly loaded by traffic, the root cause of the crack in terms of load is not known. It is possible that simply the dead load of the girder, deck,

sidewalk and barrier was enough to cause the crack to form. Likewise, a single, excessive load during construction or operation could be the reason the crack form. However, the heavy vibration of the span is the likely cause for the crack. Ambient data indicates that even when a truck passes in the interior lanes, Girder #1 tends to vibrate most heavily after the truck is off the span. This conclusion is supported by the first mode shape of vibration, presented in Section 5.10.1. Regardless of the root cause of the crack, it is clear that some type of retrofit should be put in place to mitigate the issue.

5.12 Analysis of Lateral Load Transfer System

The lateral system of this bridge is comprised of diaphragms that are directly perpendicular to the girders, and wind bracing which laces the exterior girder with the first interior girder on both sides of the bridge.

5.12.1 Diaphragm Performance

Two diaphragm connections were instrumented as described in Section 5.4.3.1. The instrument showed that the diaphragm was carrying substantial load in the diagonal member, and therefore functioning as designed. The two diaphragm responses under ambient traffic can be seen in Figure 5-37. The response magnitudes are nearly equal and opposite, which is expected. However, the redistribution of the load at the highest static load levels indicates that the diaphragms are a redundant system which allows for another load transfer mechanism (most likely the deck).

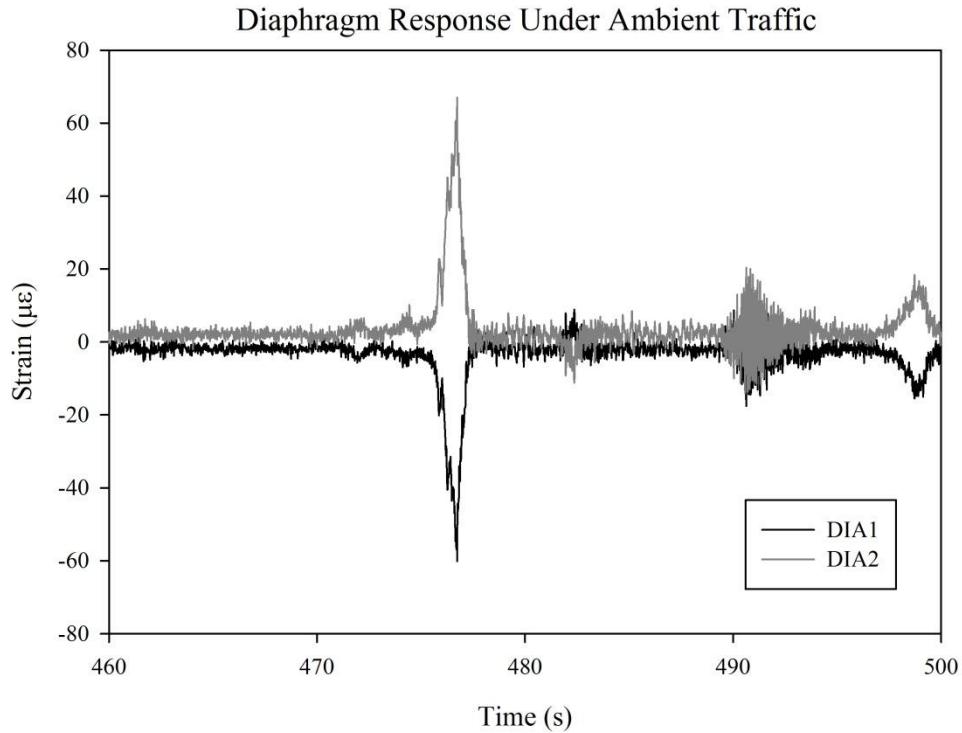


Figure 5-37 -Diaphragm Response under Ambient Traffic

5.12.2 Wind Brace Performance

The wind bracing contains the detail which is susceptible to fatigue cracking so it was of interest to understand how these members interact with the rest of the structure. As described previously, gages were installed on both the web and flange of the inverted T-section at the middle of the member. To develop an understanding for the general level of stress and force in these members, the peak strain value was used as a worst case scenario for the loading of this member. Under the final load case, the member itself, not the connection to the girders, experienced a peak strain value of around $50\mu\epsilon$ which corresponds to 1.46 ksi. It is difficult to determine from this stress level alone how much

the wind bracing is participating as a part of the load path of the bridge and if these members are really necessary at this point, or if the deck serves as enough lateral stiffness. Further analysis requires using the finite element model for extrapolation and feasibility studies for retrofits.

It is worth noting that the peak responses above were experienced under static loading in the final load stage of six full trucks. Interestingly, similar and occasionally larger peak responses were observed under traffic loading. Often these members would exhibit near resonant vibration under normal traffic loading, which could exacerbate the fatigue sensitive detail coming into the girder webs.

Additionally, if the entire time window for the final load stage is observed, the wind bracing between Girder #1 and #2 shows substantial response due to temperature change. This can be seen by the divergence of the response of the two strain gages mounted on a single wind-brace member in Figure 5-38. Note that the additional spikes in the data correspond to ambient traffic on the bridge during the static test.

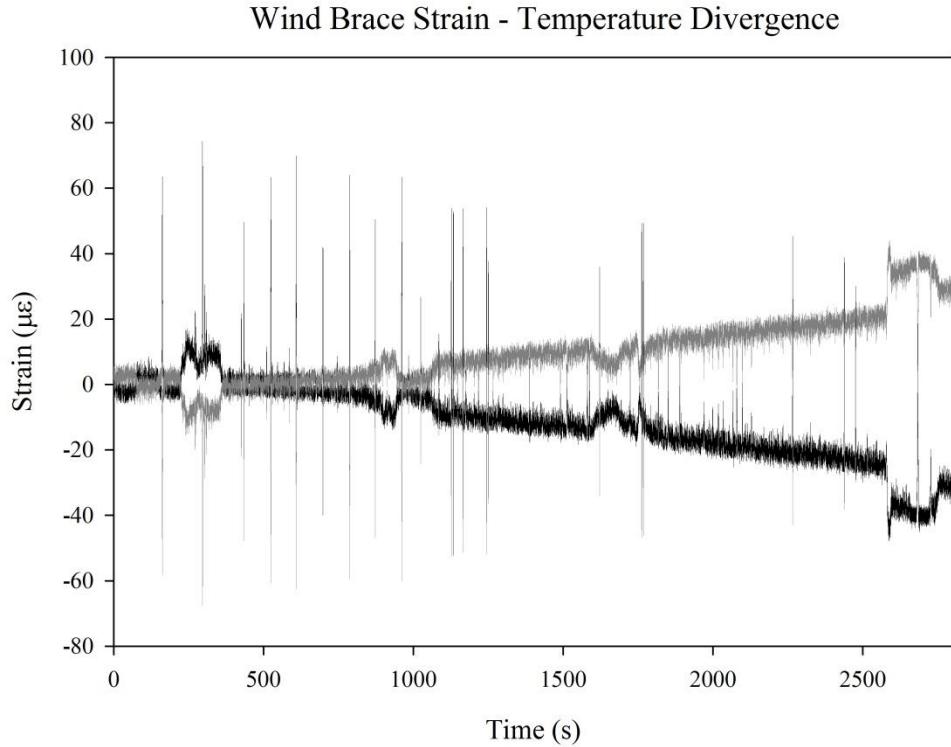


Figure 5-38 - Divergence of Wind Brace Strains due to Temperature

5.12.2.1 Fatigue Detail – Wind Brace Connection

In Section 3.4, the presence of fatigue cracks on the exterior and first interior girders of all four spans in both directions was mentioned as the controlling factor for the superstructure rating of the bridge. The bridge experiences an ADTT_{SL} of 3040 which is less than the theoretical infinite life of 3545 cycles for this detail (AASHTO Table C6.6.1.2.5-1). Considering the ADTT does not exceed the theoretical infinite life of the structure, the presence of numerous fatigue cracks throughout the structure indicate the presence of one or more other mechanisms which decrease the fatigue life by increasing either the number of cycles, the nominal stress level, or stresses induced by distortion.

5.12.2.2 Nominal Stress Level

The gusset plate for the wind bracing connects to the web 5" above the top of the bottom flange where the longitudinal strain gages were located. Another strain gage at each location was typically at 20" above the top of the bottom flange. Therefore, it is possible to linearly interpolate between these two sensors and calculate the nominal strain under the static loading as well as ambient traffic loading at the location of the fatigue sensitive category E detail. For the purposes of a fatigue investigation, the nominal stress under ambient traffic is most useful. The assumption that the maximum and minimum live load stress measured during the recorded ambient traffic is the typical stress cycle for this structure will be used to estimate the fatigue life of the category E detail. This approximate stress range was calculated as 2.87 ksi (-0.82 to +2.05ksi). Using equation 7.2.5.1-1 from The Manual for Bridge Evaluation, shown below, the finite evaluation fatigue life would be 46.7 years. This is shorter than the design life of the structure, but cracks have begun to appear well before this age was reached. The most conservative value possible, the minimum life, is 35.9 years, which is still substantially longer than the lifespan at which the cracks began.

5.12.2.3 Contribution of Excess Vibration

Examination of the ambient traffic data recorded prior to the static load test indicates that often a single truck causes more than one stress cycle because of the vibration of the span. AASHTO uses a multiplier for cycles for short spans where one set of wheels is off of the span before the next one is on, causing two cycles. This is not a factor for the IBS

study since the span is 105' at its shortest. AASHTO also lists a multiplier of 5x for cantilever girder systems because as the truck leaves, that type of structure tends to vibrate. Although nothing is explicitly mentioned about typical structures which are not cantilever girder, it seems in the case of IBS that the span vibrates so heavily under truck traffic that a multiplier may be required.

A typical twenty second window of strain response under ambient loading is shown in Figure 5-39. Notice that there are several cycles of vibration after the first truck has passed, with the magnitude of response on Girder #1 (blue line) becoming larger than Girder #3 (red line) despite the fact that Girder #3 was directly loaded. This is due to the vibration of the span, which can be noticeably felt when standing on the structure. The second truck, in lane #3, does not excite the first mode of vibration as well as truck #1, and the third truck has little to no effect whatsoever. This indicates that the type of truck, speed of truck, and position on the bridge can all have an effect on the excess vibration experienced by the bridge.

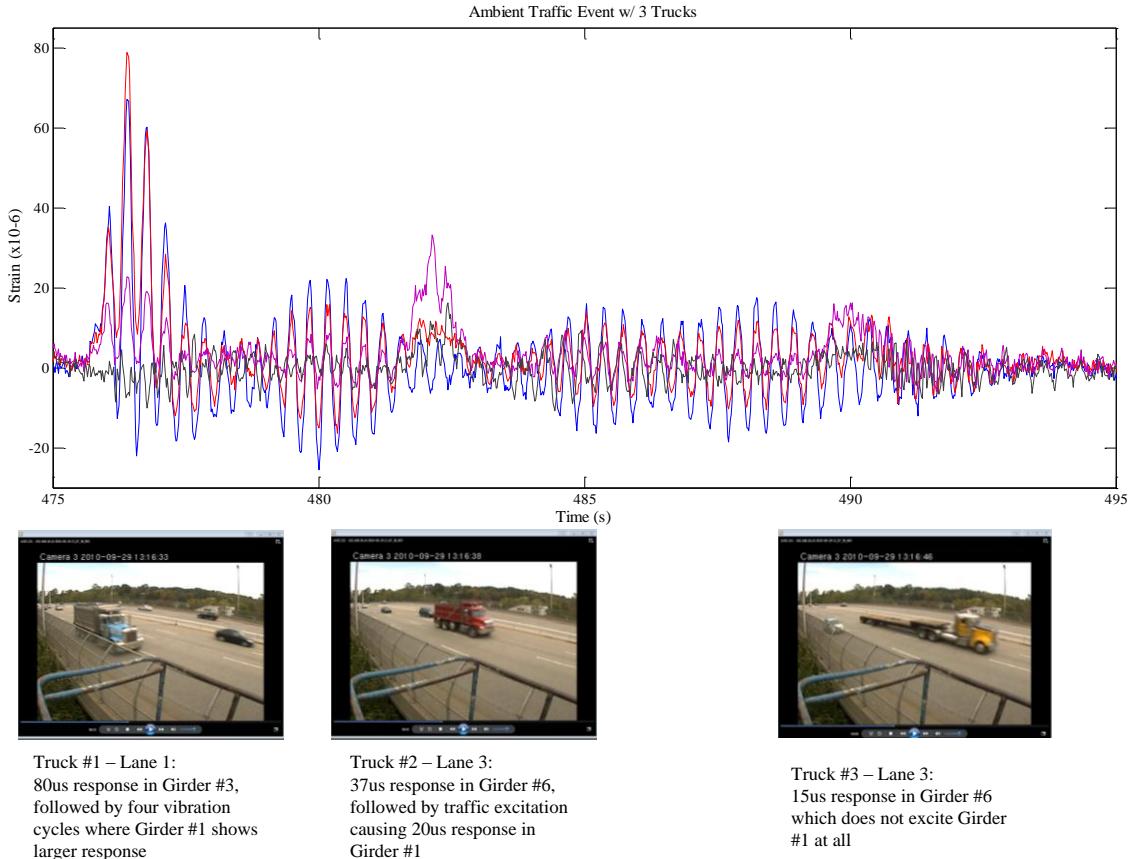


Figure 5-39 - Three-truck Ambient Event

5.12.2.4 Distortion-Induced Fatigue

Both the static testing and ambient monitoring showed considerable differential displacement between the instrumented girders. However, Girder #2 was not instrumented, as the exact differential displacement was not anticipated as a requirement in the design of the instrumentation plan. It is safe to assume that Girder #1 and Girder #2 displace relatively different. The wind bracing which ties these two girders together induces out of plane distortion into the web of the girder through the gusset plate

attachment as a result of this differential displacement. Note that this plate is located 5" above the bottom flange and within 2" of a vertical stiffener. The added local stiffness from the flange and stiffener will serve to increase the local distortion stress.

5.13 Test Data Reduction Not Completed

5.13.1 Crawl Speed Test Results

The crawl speed tests using an empty and a full truck in each lane have not yet been processed due to complications tracking the position of the truck. The laser system functioned but required adjustment as the truck moved due to the curvature of the roadway. During the testing, there were not enough personnel to dedicate someone to adjusting the position of the laser sensor. Also, there were safety concerns with the laser interfering with the traffic in the open lane. While the system shows potential for real-time, continuous, immediate feedback on truck position, it requires more refinement before application. There was a backup system for truck position in the form of video. However, this will require manual correlation to track truck position, which has not been completed yet.

5.13.2 Ambient Traffic Data Reduction

Over the week prior to the test, during final sensor installation, ambient traffic data was recorded. The day before and the day of the test in particular were heavily focused on collecting ambient data because all of the sensors were installed. In total, nearly 60

datasets of approximately 20 minutes in length were collected. Many of these sets can be directly linked with video of traffic, as seen in Figure 5-40.



Figure 5-40 - Ambient Data Collection with Video

After all sensors were installed, a camera was mounted to the manlift and the bucket was raised above the roadway level to record traffic. The camera was manually synced with the data via timestamps. The camera setup can be seen in Figure 5-41. The camera was not installed on the roadway because of the weather conditions and lack of traffic control at the time.



Figure 5-41 - Camera Setup for Ambient Traffic

5.14 Dynamic Test Results

The dynamic testing consisted of ambient vibration and forced excitation. A comprehensive discussion of all dynamic testing (including subsequent tests that occurred in the summer of 2011) on the IBS Bridge is presented in Prader (2012). For the purposes of the investigations in this thesis, the ambient results from the September 2010 testing window are utilized. These results are presented in Table 5-6 and Figure 5-42.

Table 5-6 - Modal Properties (September 2010)

Mode Number	Frequency (Hz)	Description
1	2.89	1 st Bending
2	3.79	1 st Torsion
3	5.23	1 st Butterfly
4	9.47	2 nd Torsion
5	11.62	2 nd Bending
6	12.25	2 nd Butterfly
7	15.11	3 rd Torsion

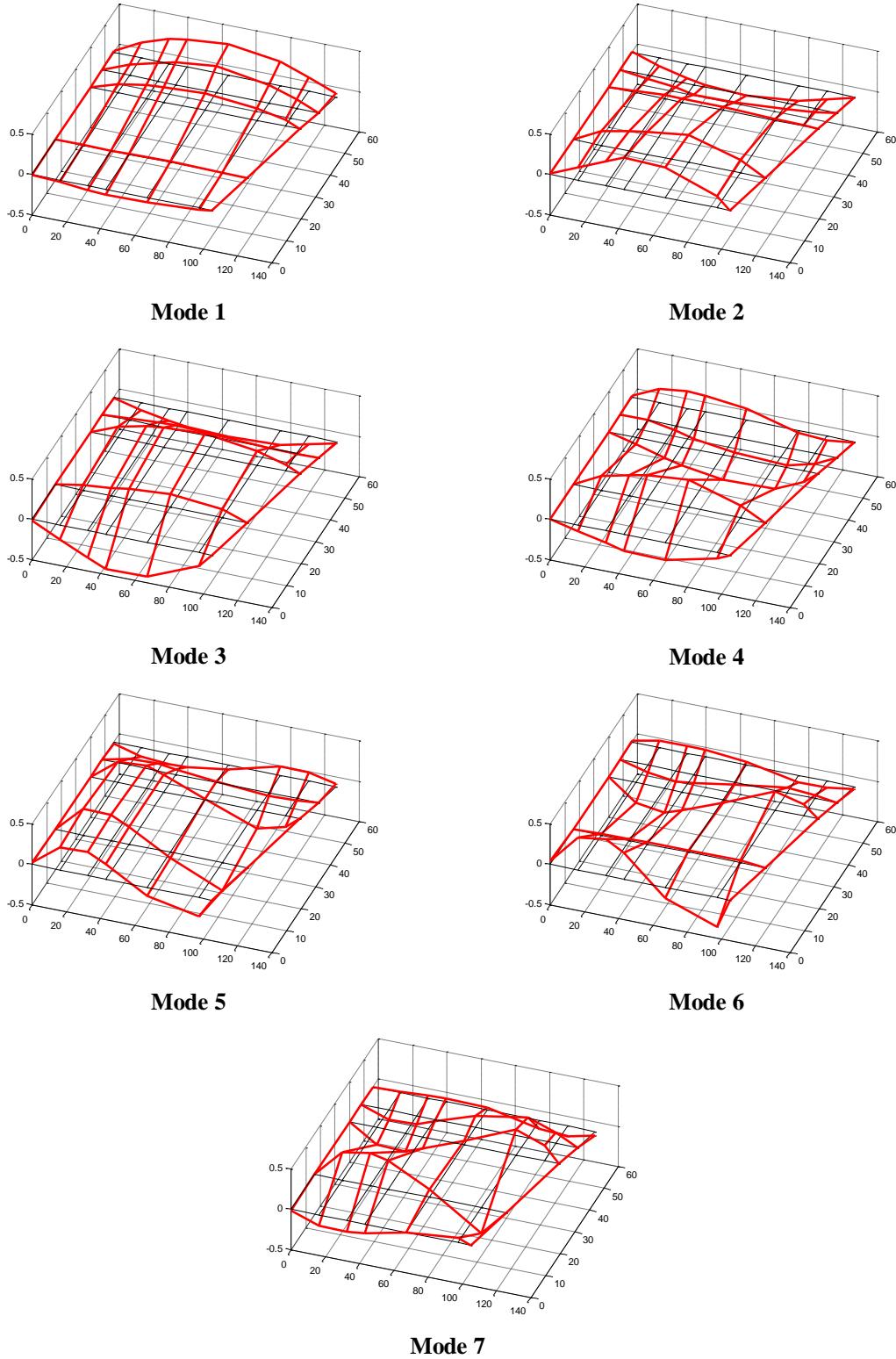


Figure 5-42 - Experimental Mode Shapes

Chapter 6: Model-Experiment Correlation

This chapter presents the post-test correlation between the experiment described in Chapter 5 and a finite element model. The original a priori model is modified to account for subsequent observations made during the extensive time spent on-site. The model was first calibrated manually, and then in an automated fashion using multiple modalities of data.

6.1 Considerations for Model-Experiment Correlation Methods

6.1.1 Parameterization of a Model

Constructed systems are continuous while models are discretized. This distinction forces the engineer to make assumptions and concessions when developing a model to represent a real constructed system. Parameterization is often a totally implicit part of the process of model construction. However, it only becomes explicit when the engineer begins to select and modify parameters based on visual or experimental observations.

An easy way to conceptualize parameterization versus basic modeling construction is to think of parameters as knobs with infinite resolution between bounds, where typical model construction is more like switches. For example, a support for a beam in a basic model may be fixed, pinned or a roller. The reality of boundary conditions is far from the typical idealized representations. When the model is parameterized, the same support would be represented with a pin and a rotational spring, the stiffness of which can be adjusted. This stiffness can be tuned to better represent experimental results.

During any model-experiment correlation effort, observations made during field work often serve as the basis for defining these parameters. For this reason, careful observation and documentation during the field testing should be conducted. It is important to understand the physical mechanisms that can cause the numerical difference of a given parameter from the ideal or assumed value.

6.1.2 Evaluation Method

The evaluation of how accurately a model can represent experimental data is usually done by way of an objective function. An objective function compares all of the experimental data to the model output. The function is generally a summation, the value of which is minimized (either manually or automatically) to find the “best” model. Though there are numerous types of objective functions, this effort will focus on a summation of percent difference, or error, of each measurement. The value of the objective function is often referred to as a residual.

6.1.3 Reconciliation of Information Formats

Any measured bridge response with a numeric value can be associated to a model response, though some may be much simpler to directly compare. For example a strain gage located on a steel beam provides a strain reading during the test. However, a traditional beam-shell model does not output strain but rather member actions. These include bending, shear and axial forces. Using mechanics, strain can be calculated from these forces.

In addition to transforming model outputs to comparable responses, objective functions typically provide superior performance when all of the response inputs are on the same order of magnitude. In other words, if an objective function includes strain and displacement and the displacements are being measured in fractions of an inch, then the strains should be converted from units of strain to units of microstrain to bring the order of magnitude of the responses closer.

Similarly, the engineer should consider the magnitude of the response when looking at percent error. A very small measurement might have 100% error, but the discrepancy is within noise range of the sensor. In this case, there may be justification for excluding the response from the objective function or assigning a small weight to that response.

6.1.4 Importance (Weight)

Weighing the terms of the objective function by some rational, defendable means can help guide the model-experiment correlation in a reasonable, predictable direction. In a parameter ID, the objective function will try to minimize the global error, or residual, without any consideration as to how the error values for each iteration were obtained.

An example of a weighing methodology which would be employed for dynamic test results is factoring the Modal Assurance Criterion or MAC value by the discrepancy in measured and calculated frequency. In other words, the MAC value is reduced by the percent error in the corresponding frequency. This would force the objective function to account directly for the correlation between frequency and mode shape. There are

numerous other ways to assign importance to information; even as basic as heuristically assigning importance, though this may not be as defendable.

6.1.5 Pitfalls

A common pitfall associated with model-experiment correlation is improper parameterization of the model. If an important parameter is missing, the selected parameters used will often assume unreasonable values to compensate. Many problems associated with model-experiment correlation will stem from this mistake.

An exception would be problems with local minima. This is particularly associated with automated parameter ID. Gradient based methods can often find a solution which cannot be improved upon with small steps in any parameters. This solution may however be a local minimum, meaning that the global minimum, or correct solution, is located elsewhere in the model space. This can be mitigated through using numerous starting points.

Both of these pitfalls are manifestations of the most significant problem related to model-experiment correlation, known as non-uniqueness. Non-uniqueness essentially means that, for a given model form, any number of parameter sets and values can represent the experimental data equally well, making it difficult to draw conclusions based off of a single, correlated model.

6.2 Pre-modification to A Priori Model

In Section 4.4 the construction and manipulation of an a priori model was presented. At the time, the focus and direction of the project was not completely settled upon. The original model was based on a geometric replica model of the entire bridge used for planning and visualization purposes. This model was available and it appeared to serve as a good starting point for finite element model construction. As the test progressed, the focus of the study became Span SB2. Post-test model-experiment correlation requires numerous iterations of running the model solver, comparing results, and changing parameters. The a priori model was large and cumbersome to work with, requiring a substantial run time per iteration. For this reason, it was decided that a smaller, simplified model would be created to drastically reduce the computational expense. The specific reasons and modifications are discussed in the ensuing sections.

6.2.1 Problem Reduction

The model including all eight spans was very computationally expensive, particularly regarding the dynamic or modal solver. The static test and most dynamic testing were focused on Span SB2 so the reduced model would also focus here. At first, this span was simply removed from the full model. However, further investigations indicated that the manner in which the entire model was oriented (true north versus Cartesian coordinates) would cause issues with parameterization of boundary conditions and composite action. A new, simplified model of Span SB2 was constructed. This model did not include the piers, the crown of the roadway, the vertical curvature of the span and the effects of the adjacent spans. It was decided that these simplifications were acceptable for this structure

based on observation of bridge condition and behavior during the test. In the future, these additions may be included in the model.

6.2.2 Discretization

Another required change was the regularization of the discretization of the model. Recall that the girders change section properties along their lengths. These changes, along with the skew, prevent a simple division of the elements. The first model was discretized manually after importing into Strand7 and was non-uniform and inconsistent. The simplified model was discretized during geometric construction and is uniform.

6.2.3 Beam to Deck Connection

The a priori model was built in an effort to represent partial composite action. This was achieved by using what is called a connection element in Strand7. Ideally the connection element would act as a link element with variable stiffness in each of the six degrees of freedom at each node. However, the connection element spread out the flexibility of the element across the entire element length, meaning that a long element (as was originally used) would not behave in the desired manner; mainly a loss of stiffness at the interface between the deck and the beam. The new model used a rigid link to cover 99% of the distance between the line representing the beam and the area representing the deck in the model. The last 1% was a connection element. This forced the software to lump the stiffness reduction at the correct location. Note that the deck was added to the model and offset upwards, meaning that the link element would actually connect to the deck at the

bottom face of the deck, and not its centerline. This is a better representation of the real structure.

6.2.4 Remodeling of the Bearings to Represent Actual Kinematics

Another major modeling change was to adjust the model method of the bearings.

Originally the bearings were also represented with a connection element. The justification was that any unintended bearing stiffness due to deterioration or any other cause could be lumped into the stiffness terms of the connection element. However, based on the observations leading to the requirement to localize the connection element for the deck connection, it was decided that the bearing should be represented in a different way.

The kinematics of the actual rocker bearing is a complex problem. The pin serves as the instantaneous center of rotation for the bearing, and the rocker contact point moves back and forth as the bearing moves. Therefore the movement is a combination of rotation and translation. Additionally, the drawings indicate that all bearings except what those marked as alignment bearings have up to 1/8th inch lateral play. This means that the span could theoretically shift laterally under vertical loading. The alignment bearings have much less lateral play. Also, none of the bearings have the ability to resist uplift.

In the end the rocker bearings were modeled with a combination of rigid and pin link elements which closely, though not perfectly, represent the kinematics of the bearing. The non-expansion bearings are modeled as a single rigid link with a pin support. Both types are shown in Figure 6-1.

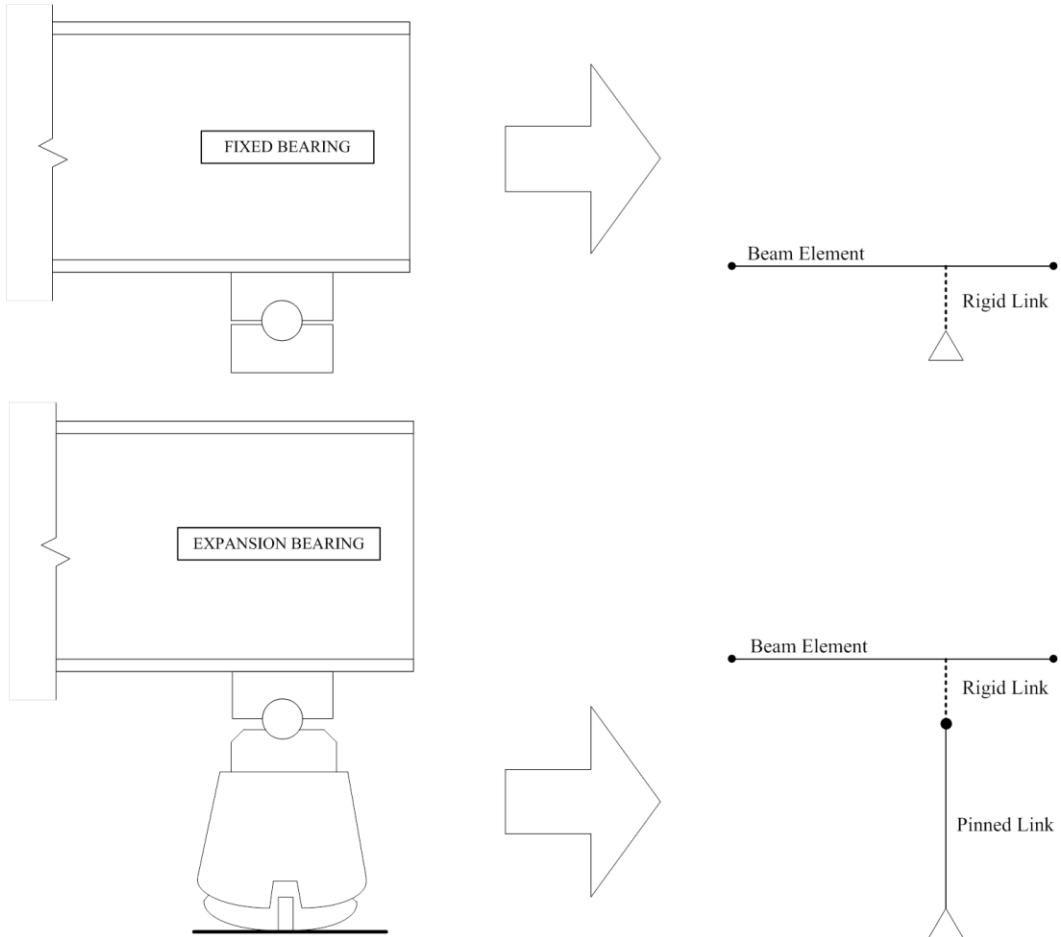


Figure 6-1 - Modeling the Bearings

Note that the beams were modeled referencing the middle of the bottom flange, so that geometrically, the bearings connected to the beam at the bottom flange, not the centerline. This allowed the connection to form a couple and develop axial force which would be impossible if the beam was supported at its centerline.

6.3 Manual Calibration

Numerous modalities of response were collected during the load testing of the IBS structure. For the purposes of manual calibration, the twelve displacement locations will generally be used as the metric for the performance of the model. The strains at the same twelve locations will be compared as well, but in a secondary sense. To save time for the manual calibration, the interface between Matlab and Strand7 was once again utilized. In this case, changes were made manually to the model in Strand7. The model was then run via Matlab, which also extracts the relevant responses from the results file and compiles them into arrays. These values were then put into a spreadsheet which automatically calculated percent error. The percent error results for each sensor were automatically color coded based on the absolute value of the error. It was possible to qualitatively see how the model improves every iteration with the manual changes to parameters. The entire manual calibration process occurred four times for the IBS bridge, the first three ending with a discovery about the model that required adjustment and a restart of the calibration.

6.3.1 Parameterization of IBS Model for Manual Updating

During the construction of the model, care was taken to make the model easily parameterized. In other words, the construction of the model typically included springs, grouping for spatial variations of properties and other simplifications.

The simplified model started out with substantial error as compared to the experimental displacements. While this is expected, it is an interesting exercise to consider what should have happened had it accurately predicted the responses without any change. Direct

observations of the test data indicate several issues which should be reflected in the model (i.e. piercap cracking) that are not explicitly accounted for in the model. Had the model been initially accurate, the engineer may have simply proceeded on to using the model for simulation. Alternatively, he may have assumed that because an observation like a piercap crack needs to be included in the model, and the effect of that model must be compensated for by another error in the model.

6.3.1.1 Selected Parameters

The parameters selected for calibration of the model manually included the following values:

1. Boundary Stiffness in all degrees of freedom
2. Compatibility between the deck and the girders
3. Diaphragm stiffness
4. Wind brace stiffness
5. Parapet involvement

These parameters tend to have a substantial effect on the global response of the structure. It is interesting and perhaps very important to note that throughout the process, there were numerous occasions which the model could very accurately predict the responses but the model was considered “unrealistic” and could not be used. An example would be decreasing the stiffness of the wind bracing to almost negligible levels. Making this change to the model at a certain stage in the calibration process reduced the percent error on almost every sensor to reasonable levels. However, what this means in a physical

sense is that the wind brace cannot transfer any load. This was investigated in the real data and it was seen that the braces carry load even under only temperature loading. Therefore the idea of artificially decreasing the stiffness of the braces was invalid despite the accuracy of this particular model.

Each parameter was bounded using sensitivity studies similar to those conducted before the load test. Typically the calibration stayed within these bounds. The sensitivity study helped to limit the amount of options for changing the model for a given iteration. If a major change to the model occurred, the sensitivity studies were quickly redone to verify that efficacy of the parameter in the new configuration.

6.3.2 Optimization Parameters

The Strand7-Matlab API interface was used to conduct the optimizations relative to numerous modalities of response measurements. The actual optimization used the LSQNONLIN function in Matlab. This function finds the minimum of the sum of squares of error terms set forth in an objective function. The algorithm used is the trust-region-reflective algorithm which can utilize bounds, but cannot be underdetermined. The simple objective function used for static results was:

$$f(x) = \sum_{i=1}^n \left(\frac{\text{Analytical}_i - \text{Measured}_i}{\text{Measured}_i} \right)^2 \quad \text{Eqn. 6-1}$$

where n is the number of measured responses. A similar function was used for dynamic properties, except that for the MAC value, the difference between the MAC and one was minimized.

6.3.3 Calibration Results

The final model was calibrated to within 0.09 inches at any given displacement response. For some of the smaller response locations, this accounted for up to 33% error in response measurement. The fact that the response was only 0.07 inch means that this 33% error corresponded to a mere .025 inches, not enough to invalidate the model. Table 6-1 shows a comparison of the measured and model displacement responses corresponding to the twelve point grid used throughout the test. Also included is the actual difference and percent difference between the responses. Considering these low error levels in conjunction with the complexity of the model, at this point, the model was considered manually calibrated.

Table 6-1: Displacement Comparison

Displacement Comparisons				
Location	Measured	Model	Difference	% Difference
G1-1/4	-0.625	-0.598	0.027	-4.3%
G3-1/4	-0.616	-0.651	-0.035	5.7%
G6-1/4	-0.358	-0.387	-0.029	8.0%
G8-1/4	-0.075	-0.104	-0.029	39.2%
G1-2/4	-0.639	-0.700	-0.061	9.6%
G3-2/4	-0.846	-0.836	0.010	-1.1%
G6-2/4	-0.447	-0.457	-0.010	2.3%
G8-2/4	-0.079	-0.104	-0.024	30.7%
G1-3/4	-0.467	-0.513	-0.046	9.8%
G3-3/4	-0.530	-0.567	-0.037	7.0%
G6-3/4	-0.249	-0.247	0.003	-1.0%
G8-3/4	-0.019	-0.024	-0.005	26.3%

Table 6-2: Strain Comparisons

Strain Comparisons				
Location	Measured	Model	Difference	% Difference
G1-1/4	86.21	88.53	2.3	2.7%
G3-1/4	111.92	102.87	-9.0	-8.1%
G6-1/4	77.36	78.28	0.9	1.2%
G8-1/4	15.67	43.93	28.3	180.3%
G1-2/4	113.11	122.97	9.9	8.7%
G3-2/4	183.63	193.11	9.5	5.2%
G6-2/4	107.71	112.53	4.8	4.5%
G8-2/4	6.16	35.17	29.0	470.9%
G1-3/4	83.87	79.62	-4.2	-5.1%
G3-3/4	118.36	107.91	-10.5	-8.8%
G6-3/4	70.35	58.81	-11.5	-16.4%
G8-3/4	4.98	19.01	14.0	281.7%

Table 6-2 shows the comparison of strains to measured responses. It is clear that the strain responses do not compare as well as the displacements. In particular, Girder #8 has very large percent errors. This will be addressed later. The strains are calculated using mechanics, accounting for both bending and axial force. The gages were installed on the topside of the flange, offset from the web. This meant that the weak axis bending was very influential. In the case of Girder #8, the weak axis bending tended to have as large or a larger magnitude component of the perceived longitudinal strain as did the strong axis bending. In the end, it was decided that a beam-shell model form was not able to account for the complex interaction of forces seen by Girder #8, particularly coupled with the fact that the uncertainty associated with exact placement of the gage on the flange has an enormous effect on the response.

6.3.4 Major Changes to the Model

During the calibration process, parameters are perturbed and the effect on the comparison to the results is quantified. This section presents the resulting major changes to the model in terms of boundary conditions, material properties, composite action and wind brace participation.

6.3.4.1 Boundary Conditions

As noted in the experimental results, there was a softness of Girder #1 on the south side. It was assumed that this was a result of the piercap crack, a conclusion supported by the activity of the piercap crack during the testing and under ambient traffic. The model

results also support this conclusion. In order to achieve roughly the same magnitude of response on Girder #1 at the quarter and midpoint, the support stiffness at the south end in the vertical direction was softened to an almost negligible value of 50k/in. In effect this forced the model to distribute the load elsewhere.

Initial investigations of the structure, and in particular, the bearings, left the engineers with the belief that the bearings were stiff laterally. However, further review of the plans indicated that the system was designed with a single “alignment” bearing with very tight tolerances. The other bearings had upwards of 0.25” of lateral play. This became crucial during the initial model runs, when each of the bearings was developing abnormally and unacceptably large lateral force, on the order of 70 kips. The bearings were then allowed to move laterally at the location of the pin in order to release these forces. After each run, the lateral displacements were checked to ensure that none of the movements were outside of the range of the individual bearings. The alignment bearings were supported with a rigid restraint.

The manner in which the bearings were modeled, described previously, was chosen so that rotational stiffness could be added to bearing easily. However, the functionality of the pin element chosen to represent the bottom portion of the bearing is not compatible with rotational springs. Therefore, rotational restraint was represented with a longitudinal spring at the location of the pin in the bearing. The bearings on the expansion side of the bridge each had a longitudinal spring included. Girders #1 and #8 had a stiffness value of 350k/in while all of the interior girders were 100k/in. The exact reason for the increased stiffness at the exterior girders cannot be determined with certainty. However, there are numerous plausible explanations. The exterior girder bearings are the most exposed, and

show a correspondingly large amount of corrosion and deterioration. Also, investigation of all the bearings on the span showed cracking of the pintels on many of the exterior bearings. This could be indicative of the bearing attempting to rotate. Direct contact with the pintels could provide restraint as well. Lastly, the interface of the spans and more likely parapet and barrier may not be completely independent from span to span. Additional restraint due to this connection would certainly manifest in the bearings on the exterior girders.

Lastly, Girder #8 tended to develop a substantial uplift reaction. The design of the bearing relies on downward compression force, meaning that vertical restraint upwards is not possible. Therefore, the restraint at this location, for this load condition only, was removed. The caveat that this load condition only causes this effect is important, because under ambient traffic, or temperature loading, this will likely not be the case, particularly considering the contribution of dead load. Therefore, any extrapolation using this model should reinstate this boundary condition, and should be checked for negative reactions.

6.3.4.2 Material Properties

Only minor adjustments to material properties were made in the manual calibration process. The deck modulus was left at the nominal value (3122ksi) as indicated in the plans. This parameter is variable and could be better represented only after material testing. In terms of the model, increasing the concrete modulus tended to have a similar effect to increasing the lateral boundary stiffness. Since these parameters seem to compete, it is entirely likely that the current values are both incorrect and are

compensating for one another. This is an accepted reality with model calibration that can only ever be mitigated, not eliminated.

The modulus of the parapets and barriers were increased to 4000 ksi. No specific concrete information regarding these components was included in the plans, but it can be assumed that a similar mix to what was used in the deck. Therefore the additional stiffness is attributed not directly to a difference in the material property, but likely the simplifications made during modeling regarding the section properties.

6.3.4.3 Composite Action

As noted in the experimental results, there was evidence of some loss of composite action. This was incorporated into the model calibration through three levels of stiffness. The stiffness in four of the six degrees of freedom at a given connection was varied. The only values to remain fixed were the axial stiffness of the connection and the rotational stiffness about the axis transverse to the span, the main bending axis. This ensured that the deck translated and rotated together with the beams, as would be the case as long as the deck and beams were in contact. The other degrees of freedom were set at either 10,000 k/in, 1000 k/in or 100 k/in. The distribution of these three values was based on the experimental results, where it was believed that composite action may have been lost. The approximate distribution can be seen in Figure 6-2.

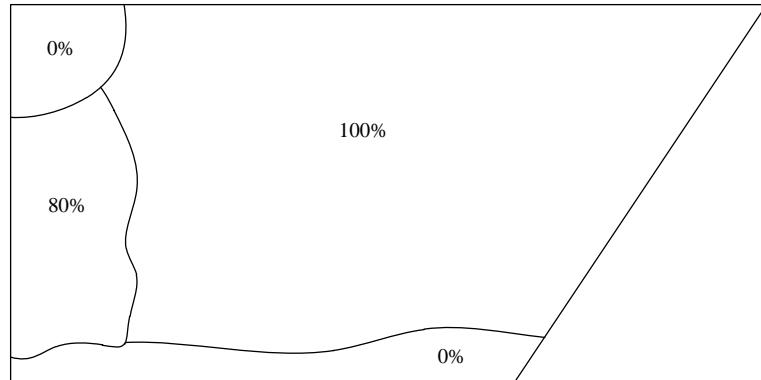


Figure 6-2 - Composite Action Distribution

In order to understand what exactly these values mean in a physical sense, another benchmark study was conducted on a single beam with the exact same setup as the model. Convergence of the largest displacement value of the girder was achieved at a stiffness of 10,000 k/in which was considered to be 100% composite action. The intermediate value was 1000 k/in which corresponds to 83% composite action. Finally 100 k/in corresponds to a near total loss of composite action. The sensitivity plot can be seen in Figure 6-3. Considering the substantial uncertainty regarding both the distribution and the amount of composite action that was lost, only the three values specified above were used.

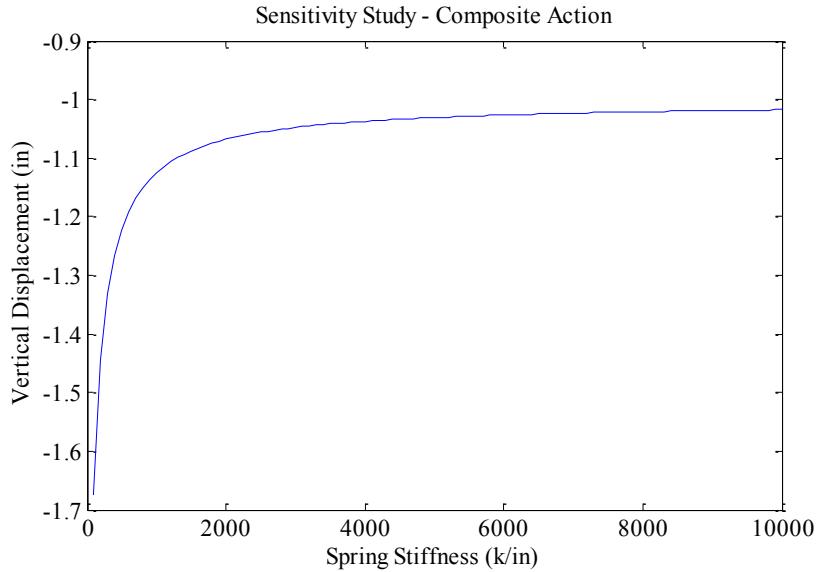


Figure 6-3 - Composite Action Sensitivity

6.3.4.4 Wind Brace Axial Stiffness

It was observed in the model behavior that the wind braces tended to add a substantial amount of axial force to the beam. The wind brace connection is modeled as a rigid connection when in reality there is some flexibility. Considering also the presence of fatigue cracks at the connection of most wind brace elements, the axial stiffness of the wind bracing was reduced slightly to 95% of its full value. This particularly had a large effect on the expansion side of Girder #8 where the strain values from the model were much larger than the experimental values.

6.3.5 Conclusions – Manual Calibration

The manually calibrated model represents a potential scenario of real-life conditions which are represented in the model as to satisfy the constraints imposed by the reality of the structure, observations made while on site, and the measured responses. Adapting to these constraints is where the heuristics of manual calibration are required. The model can be considered a representation of the real structure and used for future scenario analysis with the caveat that there is uncertainty associated with the model and any potential repair or retrofit scenario should be carefully vetted through secondary means.

6.4 Parameter Identification – Static Results

Parameter identification is often used as the only method of model-experiment correlation after a load test. In this case a series of parameter identifications were utilized in conjunction with the manual calibration. The manually calibrated model serves as the starting point for parameter ID. This is a more defendable approach. A simple metaphor would be to consider parameter identification as a metal detector. Without the guidance of a map (manual calibration), the metal detector is likely not useful for finding buried treasure.

Again the global response of displacement is used at the main input value for the objective function, though strains alone and a combination of responses were also explored. The process used and final results are presented in the following sections.

6.4.1 Parameterization of IBS for Automated Updating

6.4.1.1 Grouping of Parameters

In the manual calibration, the engineer serves as the algorithm, deciding which parameters to change and how to change it. In parameter ID, the algorithm is implicit in the processing software, like MATLAB. For this reason, it becomes necessary to group parameters in order to save computation time. For example, each spring on the sixteen boundaries was updated individually in the manual calibration. This is not advisable in the automated case for several reasons. First, as was implied above, there is a strong correlation between number of parameters and computation time. Secondly, the distribution of sensors dictates the distribution of parameters in an automated updating. In other words, a parameter directly related to a girder which was not instrumented, like a boundary spring on a girder without any sensors, is not directly related to any responses. If included, these parameters will tend to hit the bound values as there is no information available to help find a better solution. To avoid these issues, parameters are often grouped or smeared spatially.

6.4.1.2 Selection of Parameters

The selection of parameters was generally similar to that of the manual calibration, except for the addition of grouping. The parameters for parameter ID included:

1. Boundary Stiffness in all degrees of freedom
2. Compatibility between the deck and the girders
3. Parapet involvement

Each direction of boundary stiffness was lumped together as a single parameter for all eight bearing locations on each side of the span. Specifically, the degrees of freedom used were the vertical stiffness on the fixed end, and the translational stiffness on the expansion end. The deck to girder compatibility could not be distributed spatially in an automated manner, so a single uniform configuration was used which smeared the effect of any loss of compatibility. Parapet involvement is essentially the material properties of the parapet.

After the first several iterations of the parameter ID, it became clear that the displacement response was very close to the experiment with the exception of the responses on the corners of G1 and G8. During the manual calibration, it was seen that a softening of the support at G1 tended to bring both responses closer to the measured values. Therefore, the vertical reaction at G1 was made its own parameter and the analyses were rerun with this parameter set.

6.4.1.3 Starting points

Every iteration of the parameter ID process used several different starting points. This was done in an effort to ensure that the final result was not a local minimum as opposed to the global minimum of the entire modeling space. This is not a perfect solution for this issue, but it does provide some confidence that the updated parameters are a true minimum.

6.4.2 Calibration Results - Displacements

Table 6-3 shows the comparison of experimental and model displacements after parameter identification of the six parameters described previously, using an objective function based on displacement. Qualitatively, the color scheme indicates a substantial improvement over the manually calibrated model. In fact, the largest percent difference is 12%. The largest magnitude difference is 0.041 inches. Clearly the optimization is able to locate a better overall result than the manual calibration.

In the same manner, a qualitative assessment of the strains would at first glance indicate that the strains are not improved at all. In order to verify this, the value of the objective function using strains for both cases was calculated, and the parameter ID case was overall more accurate. The manual calibration tended to lump the errors at Girder #8 while the parameter ID spread the error out over the entire span.

Table 6-3 - Displacement Comparison -Parameter ID

Displacement Comparisons				
Location	Measured	Model	Difference	% Difference
G1-1/4	-0.625	-0.583	0.041	-6.6%
G3-1/4	-0.616	-0.648	-0.032	5.3%
G6-1/4	-0.358	-0.364	-0.006	1.6%
G8-1/4	-0.075	-0.079	-0.004	5.4%
G1-2/4	-0.639	-0.673	-0.034	5.4%
G3-2/4	-0.846	-0.849	-0.004	0.4%
G6-2/4	-0.447	-0.436	0.010	-2.3%
G8-2/4	-0.079	-0.070	0.010	-12.0%
G1-3/4	-0.467	-0.486	-0.019	4.0%
G3-3/4	-0.530	-0.556	-0.026	5.0%
G6-3/4	-0.249	-0.228	0.021	-8.4%
G8-3/4	-0.019	-0.020	-0.001	4.9%

Table 6-4 - Strain Comparison - Parameter ID

Strain Comparisons				
Location	Measured	Model	Difference	% Difference
G1-1/4	86.21	88.53	-20.1	-23.3%
G3-1/4	111.92	102.87	-28.0	-25.0%
G6-1/4	77.36	78.28	-20.2	-26.1%
G8-1/4	15.67	43.93	9.0	57.4%
G1-2/4	113.11	122.97	-11.3	-10.0%
G3-2/4	183.63	193.11	-9.7	-5.3%
G6-2/4	107.71	112.53	-15.6	-14.5%
G8-2/4	6.16	35.17	11.4	185.4%
G1-3/4	83.87	79.62	-24.0	-28.6%
G3-3/4	118.36	107.91	-37.6	-31.8%
G6-3/4	70.35	58.81	-42.6	-60.5%
G8-3/4	4.98	19.01	0.0	0.1%

6.4.2.1 Distributed versus Localized Errors

The discrepancies between Girder #8 in the manually calibrated model are likely a result of the complex interactions of the structure in that vicinity compounded by the generally lower level of response, and lack of direct loading. In other words, the model and the test are not really capable of fully informing a calibration this portion of the span. This is not immediately clear in the parameter ID because the algorithm accounts for these large errors by slightly increasing other errors in order to lower others. The smearing of the error provides excellent numerical results, but depending on the application it may not be the best result in terms of informing the project engineers.

6.4.3 Calibration Results – Additional Runs

The model was updated also using strains, though it should be noted that no sensitivity studies of strain response versus the selected parameters were run. The results appear to be indicative of this, as neither the strains nor the displacements predict the experimental results noticeably better. Intuitively, one would expect the strains to be more accurate.

The comparison tables can be seen below.

Table 6-5 - Strain Predictions when Updating with Strains

Strain Comparisons				
Location	Measured	Model	Difference	% Difference
G1-1/4	86.21	62.19	-24.0	-27.9%
G3-1/4	111.92	78.42	-33.5	-29.9%
G6-1/4	77.36	55.50	-21.9	-28.3%
G8-1/4	15.67	22.93	7.3	46.3%
G1-2/4	113.11	98.94	-14.2	-12.5%
G3-2/4	183.63	167.88	-15.7	-8.6%
G6-2/4	107.71	89.43	-18.3	-17.0%
G8-2/4	6.16	16.87	10.7	173.9%
G1-3/4	83.87	57.33	-26.5	-31.7%
G3-3/4	118.36	76.10	-42.3	-35.7%
G6-3/4	70.35	23.53	-46.8	-66.6%
G8-3/4	4.98	4.35	-0.6	-12.7%

Table 6-6 - Displacement Predictions when Updating with Strains

Displacement Comparisons				
Location	Measured	Model	Difference	% Difference
G1-1/4	-0.625	-0.499	0.126	-20.1%
G3-1/4	-0.616	-0.618	-0.002	0.4%
G6-1/4	-0.358	-0.355	0.003	-0.8%
G8-1/4	-0.075	-0.076	-0.001	1.7%
G1-2/4	-0.639	-0.604	0.035	-5.5%
G3-2/4	-0.846	-0.798	0.048	-5.7%
G6-2/4	-0.447	-0.417	0.030	-6.6%
G8-2/4	-0.079	-0.069	0.010	-12.6%
G1-3/4	-0.467	-0.440	0.027	-5.8%
G3-3/4	-0.530	-0.519	0.011	-2.1%
G6-3/4	-0.249	-0.215	0.034	-13.7%
G8-3/4	-0.019	-0.020	-0.001	5.7%

The results in terms of displacements are quite similar to the previous iteration where displacement was in the objective function, with the exception of Girder #1. The strain measurement on G1 did not exhibit any effects from the piercap crack, while the displacement did. For this reason, this iteration did not adjust the spring on G1 in the same way the previous iteration did.

What is actually happening is that the strain responses are not very sensitive to the selected parameters over the specified ranges. This is further supported by an iteration of the parameter identification in which both strains and displacements were included in the objective function. The end results were essentially the same as the displacement only case, meaning adjusting the model according to displacement was the only way for the algorithm to reduce the modeling error.

6.4.3.1 Verification using a Globally Known Variable

In order to verify the efficacy of the model calibrated based on displacement values, a single parameter was allowed to optimize unconstrained using the base model as a starting point. The single parameter chosen was the modulus of steel of the beams. The purpose of this is to test whether the modulus, a fairly deterministic parameter, will remain constant, or vary greatly.

Additionally, a second similar test was run where the concrete deck was divided longitudinally into three different sections. The moduli of these three sections were also left free to optimize.

The results of the first test show that no matter the starting point for the optimization, the end result was a value of modulus for steel of 29,200 ksi, which is within 1% of the deterministic value typically accepted. This indicates that there is not a major error in the model for which this parameter could be made to compensate. The second test had similar results, with the modulus of the concrete remaining within 10% of the specified value. There is no guarantee that this value is correct, but it is at least stable. As mentioned before, a discrepancy in this value could be compensated for through other parameters.

6.5 Parameter Identification – Modal Properties

Parameterization of the model for updating using modal properties was similar to the static case, except that the parameter set needed to be able to account for mass as well as stiffness. For the IBS case study, the mass-related parameter used was the density of the

concrete deck. This could account for a variation in material properties or a discrepancy in the thickness of the deck compared to the drawings. The other parameters included were the modulus of the barrier and parapet concretes, and the translational spring on the expansion side of the structure.

6.5.1.1 Model Changes to Accommodate Dynamic Model Calibration

The static calibration of the model resulted in some boundary conditions which were not applicable in the dynamic sense. In particular, under the relatively low levels of dynamic load applied to the bridge during impact testing, the lateral movement of the bearings and the vertical displacement of the fixed bearings would likely not be activated. Therefore, for the dynamic calibration, the boundary conditions were returned to the initial pin-roller state, with the added translational spring on the expansion side. All other parameters from the static calibration were left constant.

6.5.2 Calibration Results – Modal Properties

The experimental frequencies are included in all subsequent tables for comparison purposes. The final optimized modal property set is presented below in Table 6-7.

Table 6-7 - Modal Calibration Results

Mode #	Exp Freq	Model Freq	% Diff	MAC
1	2.89	2.85	-1.3%	0.924
2	3.79	3.71	-2.2%	0.938
3	5.23	5.54	6.0%	0.950
4	9.47	8.99	-5.1%	0.880
5	11.61	10.61	-8.6%	0.822
6	12.25	13.13	7.2%	0.742
7	15.12	14.75	-2.5%	0.530
8	20.64	18.13	-12.1%	0.649

Modes 1-3 show excellent correlation with the experimental modes and percent errors with measured frequencies that are 6% or less. Modes 4-6 show some correlation with the experimental modes, though not as much as the first three modes. Modes 7-8 are not very well correlated.

Part of the reason for the discrepancy in the higher modes is the presence of local modes in the model, occurring along the unsupported cantilever sections outside of the exterior girders. Also, the higher modes are affected more by stiffness than mass, while lower modes are the opposite. This indicates that the calibration has accurately predicted the mass, but may still have some discrepancy in stiffness, which could be due to any of several nonlinear behaviors of the structure, as discussed in the static calibration section.

The final parameter set is presented in Table 6-8.

Table 6-8 - Final Parameter Values

Parameter	Initial	Final	
Density of Concrete	150	142.2	pcf
Modulus of Barrier	3500	3031	ksi
Modulus of Parapet	3500	4394	ksi
Longitudinal Spring	50	93.2	k/in

Note that none of the values hit a bound, meaning that the solution is stable. This is supported also by the fact that this solution was reached from several starting points.

6.5.2.1 Verification using a Globally Known Variable

In the case of the dynamic calibration, the verification of the model solution through the stability of a globally known parameter occurred during the normal updating. The density of concrete varied only slightly meaning that it was in fact a stable parameter. Similar to the static case, this indicates that the model is not compensating for errors via alteration of what should be a stable parameter.

6.5.3 Comparison of Statically and Dynamically Calibrated Models

It was intimated in the discussion of the calibration to dynamic data that in actuality, two separate models were required to represent both static and dynamic data. This is because the small load levels present during dynamic loading do not activate all the same mechanisms of the bridge that the heavy static load does. Table 6-9 shows the discrepancy in modal properties between the experiment, the dynamically calibrated

model and the static calibrated model. In some cases, the static model may have a slightly improved MAC value or a somewhat more accurate frequency value, but overall, the static model cannot adequately represent the dynamic properties. Similarly, Table 6-10 shows the opposite case, where the dynamic model cannot represent the displacement data well. Based on these results, it can be concluded that two models were indeed required. This highlights the importance of achieving substantial load levels during static testing in order to activate all of these mechanisms.

Table 6-9 - Comparison of Static and Dynamic Model – Dynamic Response

	Frequencies			Mac Values	
	EXP.	DYN.	STATIC	DYN.	STATIC
Mode 1	2.89	2.85	2.60	0.924	0.955
Mode 2	3.79	3.71	3.80	0.938	0.734
Mode 3	5.23	5.54	5.21	0.950	0.860
Mode 4	9.47	8.99	8.69	0.880	0.780
Mode 5	11.61	10.61	9.78	0.822	0.623
Mode 6	12.25	13.13	11.58	0.742	0.437
Mode 7	15.12	14.75	13.06	0.530	0.809
Mode 8	20.64	18.13	16.98	0.649	0.657

Table 6-10 - Comparison of Static and Dynamic Model - Static Response

	Displacements		
	EXP.	DYN.	STATIC
G1-1/4	-0.62	-0.53	-0.56
G3-1/4	-0.62	-0.69	-0.65
G6-1/4	-0.36	-0.46	-0.37
G8-1/4	-0.07	-0.19	-0.08
G1-2/4	-0.64	-0.69	-0.66
G3-2/4	-0.85	-0.93	-0.85
G6-2/4	-0.45	-0.55	-0.44
G8-2/4	-0.08	-0.17	-0.07
G1-3/4	-0.47	-0.53	-0.48
G3-3/4	-0.53	-0.63	-0.56
G6-3/4	-0.25	-0.30	-0.23
G8-3/4	-0.02	-0.06	-0.02

6.5.4 Load Ratings

In an effort to provide quantitative feedback in a usable form that is familiar to owners and engineers, load ratings for several models throughout the model-experiment correlation process were calculated. The load ratings were calculated based on stress in the girders due to the truck loading, the worst case of which was typically Girder #4 near midspan. Note that this is not necessarily the rating factor for the entire bridge, as it may not be the most critical position. The midspan of the girder is often assumed to be the critical position for flexural load ratings however, and for that reason is used in this exercise. Four models chosen throughout the process were used to calculate load ratings. The equation used was AASHTO 6A.4.2.1-1, shown qualitatively below:

$$RF = \frac{Capacity - Dead Load Demand}{Live Load Demand} \quad \text{Eqn. 6-1}$$

The actual process presented in AASHTO 6A.4.2 was used including all relevant factors and values. The load rating results for each model are shown in Table 6-11. One important assumption to note was that the stress in the model due to live load tended to be a combination of axial and flexural stress. For the purposes of this simple exercise, the two were both assumed to be flexural.

Table 6-11 - Rating Factor Comparison

Model	RF
Original	2.86
Manual	3.54
Static	3.58
Dynamic	2.99

The expected pattern of a slightly improved load rating with model calibration is seen. Often this is not the case, and the load rating will decrease. The dynamic case shows a slightly lower rating factor than the other two calibrated models, likely because it employed simple boundary conditions as a result of the load applied load levels used to excite the structure for that case.

In reality, the complex combinations of forces that occur near the supports, particularly in the skewed corners, would likely provide lower load ratings. The diaphragms on the skewed end were very sensitive to the boundary condition changes, and would provide a much higher load rating, although not one necessarily directly correlated with safety of the bridge. Relaxation of the boundary conditions under the extreme load case tended to relieve these stresses, indicating that there were some mechanisms of energy dissipation in these bearings which were activated under the largest load case.

6.5.5 Conclusions – Parameter Identification

Three calibration scenarios were conducted including a manual calibration, and parameter identification using both static and dynamic data. The manual calibration utilized heuristics and field observation to develop a fairly reliable model, including what the engineer feels are the crucial aspects of the structure. Particular attention was paid to the modeling of the bearings and the deck-to-beam connections.

The manual calibration was used as the starting point for updating based on displacements, producing a stable, reliable representation of the Span 2 of the IBS structure. This parameter identification was conducted using both strains and displacements, but in the end displacements were used because of the more reliable, global nature of these measurements.

Finally, this model served as the basis for a calibration to dynamic data collected under ambient and forced impact loading. Several changes were made to the boundary conditions in order to force this model to behave as the actual structure would under

small, transient loads akin to those used during the dynamic testing. The final dynamically calibrated model, while different from the static model, was able to accurately represent the modal properties of the structure observed during testing.

The models have been through a rigorous error screening and troubleshooting process, guided by heuristics and real-life observations. From this point, these models can serve as the basis for numerous hypothetical scenarios to aid the owner in management of the structure.

Chapter 7: Utilization of the Calibrated Model for Simulation

This chapter introduces several scenarios that may be of interest to the bridge owner, one of which is investigated using the calibrated finite element model. The true benefit of an effort like the IBS study is not simply to calibrate the model to predict responses that were measured in the field. The calibration of the model provides the engineer with confidence regarding the reliability and applicability of the model. The model can then be used to extrapolate for many various, hypothetical scenarios that may be of interest to the bridge owner.

7.1 Potential Scenario Investigations for IBS

In the case of IBS there were numerous proposed retrofits and modifications to the bridge that the owner may want to consider in an effort to help mitigate some of the issues that have arisen regarding the bridge. For this thesis, only one scenario will be conducted in a cursory sense. It should be stressed that while the model is more reliable than an uncorrelated model and can be used for these simulations, the engineer must not disregard engineering heuristics, design principles, and code guidelines. The calibrated model is simply another tool available to the engineer.

Through scenario analysis the calibrated model can serve as an excellent proof-of-concept regarding potential retrofits, permit loads, fatigue life, and many other situations that would require extrapolation. Some of these scenarios are described qualitatively here including what steps would be taken to conduct the analysis, how the model would be modified, and how the results could inform the bridge owner. One example scenario, removal of the wind bracing, is completed.

7.1.1 Mitigation of the Excessive Span Vibration

It has already been stated that the bridge, under the static load case, did not exceed the deflection criteria of L/1200 specified in the drawings. However, it does have a serviceability issue in terms of vibration. There are numerous ways to address this issue, some of which are highlighted below.

7.1.1.1 Rendering the Span Continuous

One potential solution to address the heavy span vibration observed in the field would be to render the span continuous. This would potentially have several other benefits for the bridge as well. In particular, the removal of the joints would help extend the life of the bearings, girders, and piercaps by preventing water intrusion to the superstructure. The rocker bearings would likely be replaced as well to allow for the different types and ranges of movement that would result from a continuous span. The added stiffness would ideally push the frequency range of the structure away from that of the traffic, mitigating the vibration problems as well. All of these changes would most likely extend the life of the bridge greatly.

In the case of this particular scenario, the model would require some very specific changes. Currently the model includes only span SB2. To render it continuous would require modeling of the rest of the spans, or at least representing them numerically in the model via springs. Additionally the deck would need at least spring connections as well as the girders. The load case used for the static test would be informative, but an ambient load input case would also be required. Detailed analysis of the connections used to render the spans continuous may need to be analyzed at the micro level as well, using

actions from the macro model as inputs. Pre and post-retrofit load ratings can be calculated to numerically represent the change in behavior.

Assuming that the modeling indicates that rendering the span continuous will help in terms of the vibration issues, the owner can comfortably move forward with the design of a retrofit by a professional engineer.

7.1.1.2 Adding Mass to the Structure via Asphalt Overlay

Another way to move the primary frequencies away from that of traffic would be to add mass to the structure. The best way to uniformly do this would be via an asphalt overlay on the deck. This would not only add mass, but also some stiffness, protection for the concrete from salt penetration, and a generally smoother roadway travel surface.

In all scenarios, an important distinction to consider is not only how well a retrofit may work in adjusting or mitigating the undesirable behavior of the bridge, but also predicting any unexpected effects of the change. In the case of an asphalt overlay, one such potential effect would be the change in material properties with temperature. Essentially, the added mass may be offset by added stiffness when the temperature is low.

For this model, the engineer should consider preprocessing the model with a temperature change, or alternatively, running the model at various temperature levels to study the effects on the asphalt. Clearly this information would help the owner make future maintenance decisions.

7.1.1.3 Stiffening the Girders Longitudinally

A final method of adjusting the dynamic properties of the structure would be to directly stiffen the longitudinal girders. This would be accomplished through the addition of cover plates on the bottom flanges. This is a commonly applied retrofit using in bridge engineering and could be considered for this situation.

Of particular importance would be dealing with the changes in the bottom flange thickness. These changes are carefully transitioned to prevent stress concentrations, and the welded cover plate attachment would need to consider this geometry. The macro model could be used to determine the change in the nominal stress due to the retrofit, but an additional micro level model would be required to determine the approximate effects of the transition areas with the cover plates.

7.2 Sample Simulation – Removal of Wind Brace

Throughout the observation and testing of the IBS structure, it was often hypothesized that the wind bracing (Figure 7-1) was really serving no purpose and was in fact detrimental. Physically, this is obviously supported by the location of the fatigue cracks at the connection of the wind brace to the beam webs. The large differential displacement between adjacent girders compounded with the relatively short life of the bridge thus far indicates that this cracking may be distortion induced. Since the bridge has a substantial deck and diaphragms providing lateral stiffness, a potential retrofit to mitigate the fatigue cracking problem was to totally remove the wind bracing.

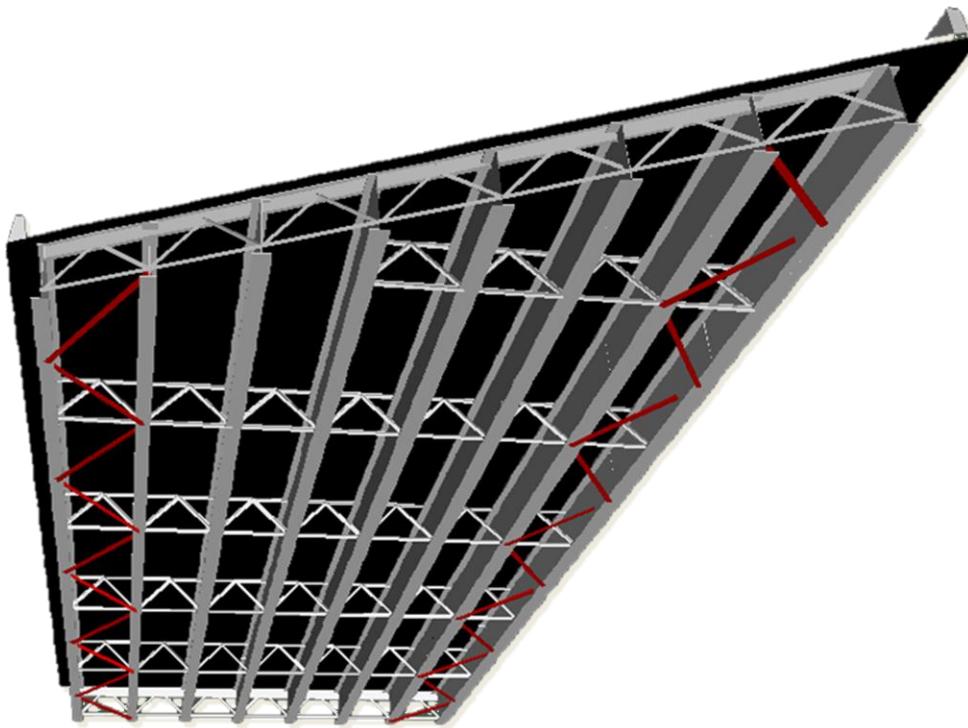


Figure 7-1- Wind Bracing to be Removed

7.2.1 Process of Wind Brace Removal

The analysis will be conducted on the model updated both through manual calibration then through parameter identification. In order to maintain the node and element numbering convention used in the prior analyses, the wind brace will not be physically deleted from the model. Rather the section properties will be reduced to such a level that they will not support any load. This concession will allow the engineer to use the MATLAB/Strand7 interface to conduct the analysis and extract results.

7.2.2 Wind Brace Removal Results

Before discussing the results of the wind brace removal model simulation, it is important to note the preliminary nature of these results. The loading and ambient conditions were not varied as of yet. Also, the response comparisons for the six loaded trucks are used a comparison when in reality it is unlikely that scenario will reoccur.

In addition to the comparison of the measured responses, the model was scrutinized for major changes in behavior, including reaction changes and support movements. The evaluation of the wind brace removal will be primarily based on these metrics.

7.2.2.1 Evaluation of Wind Brace Removal Scenario

Removal of the wind brace did not cause substantial changes in any of the measurements except those located on Girder #8. In both the strain and displacement case there was a substantial drop in response. This indicates that the response seen in Girder #8 in the model was largely due to the wind bracing, and likely its interaction with the skewed diaphragm. The reactions and nodal displacements at the boundaries did not exhibit any substantial change as well. It can be concluded that in order to mitigate the fatigue cracking issue moving forward, the owner should strongly consider removal of the wind bracing.

Chapter 8: Multiple Model-Experiment Correlation

This chapter transitions between the St-Id application of the IBS Bridge and the in-depth investigation into multiple model St-Id methods. The justification of the need for multiple models is presented first followed by a summary of current multiple model methods. Finally, a new multiple model method is outlined.

8.1 Summary of Previous Model-Experiment Correlation

The model-experiment correlation conducted as a part of the application of best practices St-Id for the IBS Bridge was exhaustive and comprehensive from a single-model standpoint. This process included:

- Robust error screening
- Sensitivity studies for all considered parameters
- Manual calibration to multiple modalities of data (displacement, strain etc.)
- Parameter identification using gradient-based optimizations

While the results presented in the prior chapters represent the best that can be achieved through traditional St-Id, they are still subject to the same non-uniqueness issue that plagues all modeling efforts for constructed systems. To reiterate, it is entirely likely that there are numerous other combinations of parameters and modeling techniques that are acceptable and capable of predicting the experimental results equally well. This, in and of itself is not necessarily a problem, but for the fact that these different combinations can produce drastically different predictions does pose a significant problem. This is not a reflection on the modeler, the software or the bridge, though all play a part. A finite

element model, or any type of model, is predicated on assumptions and is always an incomplete and limited representation of a real structure. Moon (2006) provides a comprehensive discussion on the characteristics of constructed systems that lead to uncertainties associated with modeling and model-experiment correlation. For example, constructed systems are continuous, heterogeneous, redundant systems with non-idealized boundary conditions. All of these characteristics are principle sources of uncertainty regarding the ability of the model to represent the real structure, and they require the adoption of underlying assumptions that are difficult (if not impossible) to fully validate. Given this, it is beneficial to focus not on the identification of the “best” parameters (as these are really only the best parameters given all other modeling assumptions) but rather on the effects that changes in these parameters have on the prediction of immeasurable quantities. This leads directly into the idea that information from a single calibrated model gains much more relevancy when it is placed in the context of similar information from numerous different models, as this would show the effects of parameter and modeling assumption changes.

8.2 Model-Experiment Correlation within Multiple Model St-Id

The general approach to using multiple models during St-Id starts with the development of a candidate population of models. For the purposes of this research, the term model is used interchangeably to describe a set of parameter values within a given model form (e.g., a range of boundary rotational stiffnesses), different sets of parameters and their subsequent values within a model form (e.g., pinned vs. fixed vs. rotational spring boundaries), different spatial variations of parameters within a model form (e.g., a single

value for deck modulus vs. a different value at each girder), or entirely different model forms (e.g., element-level vs. shell models).

Given a candidate model population, the models are sampled using one of many approaches. This research will focus on a few particular methods, described below, but a high-level discussion of sampling approaches is presented in Chapter 2.

The sampled models are weighed based on their ability to predict a set of observed responses. Again there are many approaches to weighing, though this research focuses on using the likelihood of the model given the observations. Other weighting methods are discussed in Chapter 2.

Finally the weighted models are used to predict unobserved responses. These responses are typically of interest to the bridge owner as they inform decision-making regarding the structure in a much more direct and understandable manner.

This section will present more background on Markov Chain Monte Carlo, Reversible Jump Markov Chain Monte Carlo, and a third approach using Multiple MCMC chains.

8.3 Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) is a well-known and documented method for numerical integration, similar to Monte Carlo integration, but through the use of Markov Chains. The method was first proposed by Metropolis and Ulam (1949) and was later generalized by (Hastings 1970). In this context, MCMC is used to approximate the posterior probability of the model space. Bayes Theorem is defined in Eqn. 8-1 as:

$$P(\theta|D) = \frac{P(\theta)f(D|\theta)}{\int P(\theta)f(D|\theta)d\theta} \quad \text{Eqn. 8-1}$$

where $P(\theta)$ is the prior distribution, and $f(D|\theta)$ is the likelihood of the observation, D, given the parameter, θ . The likelihood calculation assumes a normal probability distribution with a mean equivalent to the model prediction, and an assumed standard deviation of 4%. The 4% assumption is based on prior experience, but remains an arbitrary choice. The likelihood equation is shown in

$$f(D|\theta) = \prod_{j=1}^n \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{(x_j - u_{ij})^2}{2\sigma_j^2}} \quad \text{Eqn. 8-2}$$

In this context, these equations apply to parameters within a model, but can be modified to apply to multiple model forms. The most difficult part of Bayesian inference is estimating the denominator of this equation as the joint distribution is not generally known in advance, and therefore cannot be readily integrated. MCMC consists of generating samples from a proposal function which, after a large enough number of samples are drawn, approximates the posterior distribution, $P(\theta|D)$. This process is shown schematically in Figure 8-1.

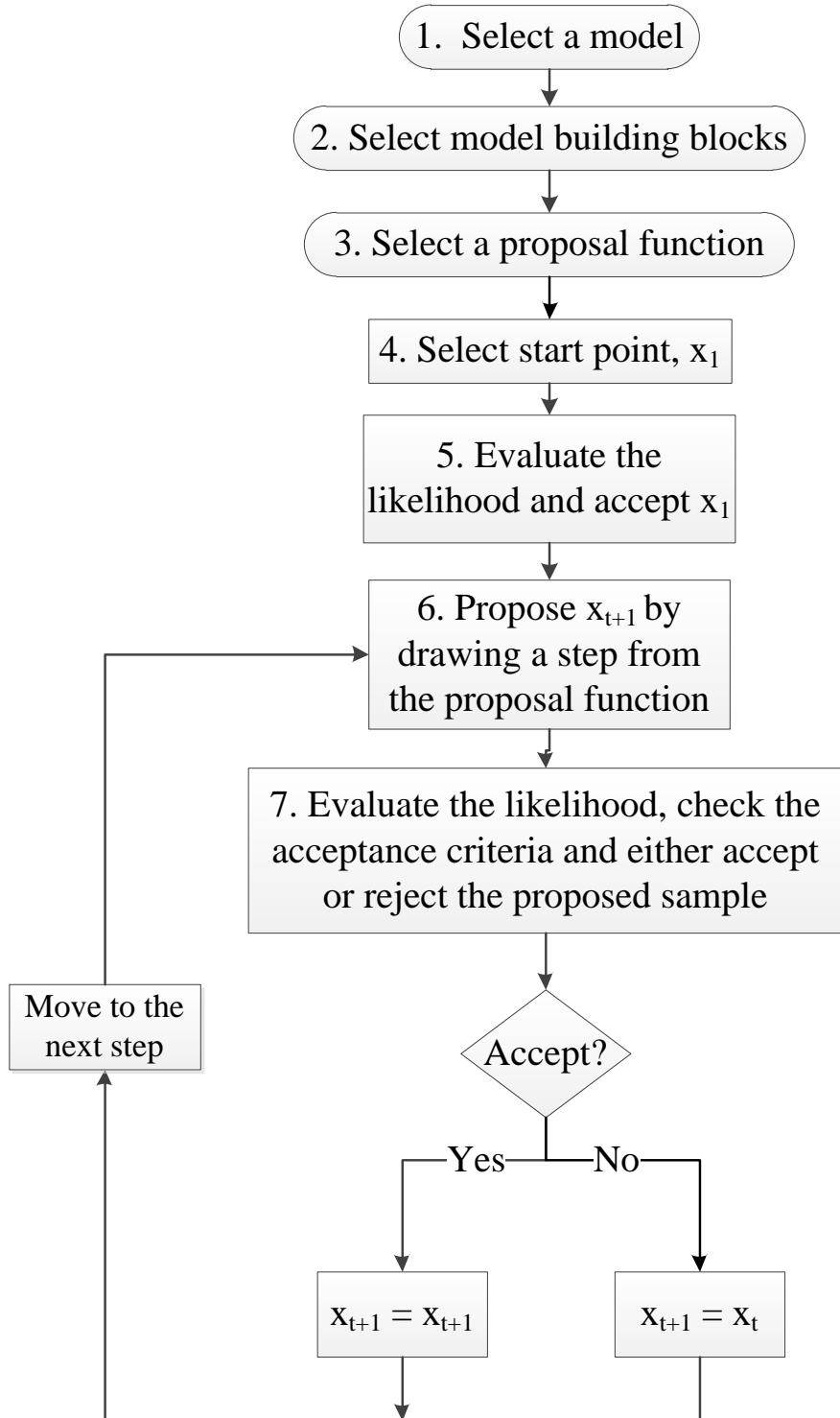


Figure 8-1 - General MCMC Procedure with Metropolis-Hastings Algorithm

These particular applications of MCMC will utilize the Metropolis-Hastings algorithm (Hastings 1970). In this algorithm, acceptance or rejection of a given sample is determined by an acceptance ratio (Eqn. 8-3)

$$\alpha = \frac{P(\theta_{i+1}|D) \cdot q(\theta_i|\theta_{i+1})}{P(\theta_i|D) \cdot q(\theta_{i+1}|\theta_i)} \quad \text{Eqn. 8-3}$$

where q is the assumed proposal function. In the case where the proposal function is symmetric (as applied in this research), these terms cancel out.

The general procedure of the algorithm, as shown above, is as follows:

1. Randomly select parameter values, θ , run and accept the parameter set
2. Generate a sample from the proposal function, q , and step each parameter
3. Evaluate the given model form with the selected parameter set
4. Determine the acceptance ratio, α , using Eqn. 8-3
 - a. If $\alpha > 1$, then accept the model
 - b. If $\alpha < 1$, accept with a probability of α , and reject otherwise
 - c. If rejected, return to the previous state
5. Repeat until converged

8.3.1 Selection of Proposal Functions

Selection of the proposal function is a difficult task which has several documented issues (Haario, Saksman et al. 1999). Figure 8-2 shows a conceptual schematic of proposal distributions relative to a target distribution (Hanson 2000). The proposal function is being used to map out the entire target distribution so naturally it is beneficial for the proposal function to be approximately the same as the target distribution (Case B). If a proposal function is too small (Case A), convergence will be slow as the chain will mix very slowly due to very high acceptance. Similarly, if the proposal function is too large, the chain will also mix slowly because of a large number of rejected samples (Case C).

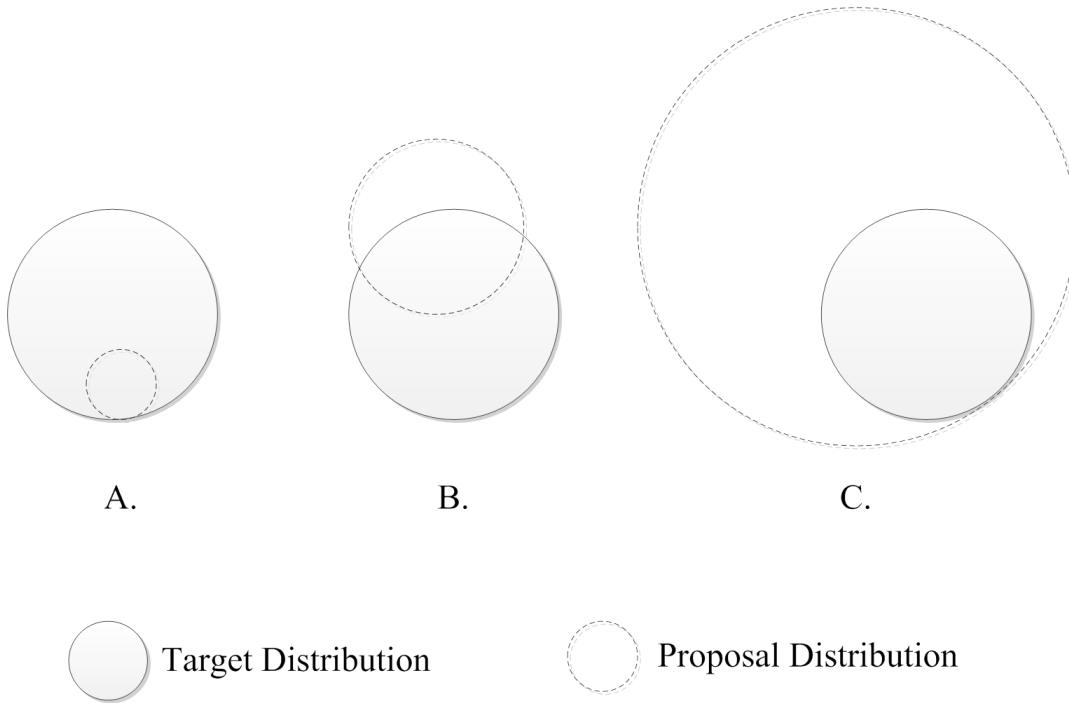


Figure 8-2 - Proposal Distribution Schematic - Adapted from Hanson (2000)

This phenomenon is shown qualitatively through a simple example presented in Dubbs (2012) and included here. The target distribution is taken as the standard normal probability density function with a mean of zero and a standard deviation of one. The proposal standard deviations of Case A, B, and C will be 0.01, 1 and 100, respectively. The results of each case are shown below in Figure 8-3.

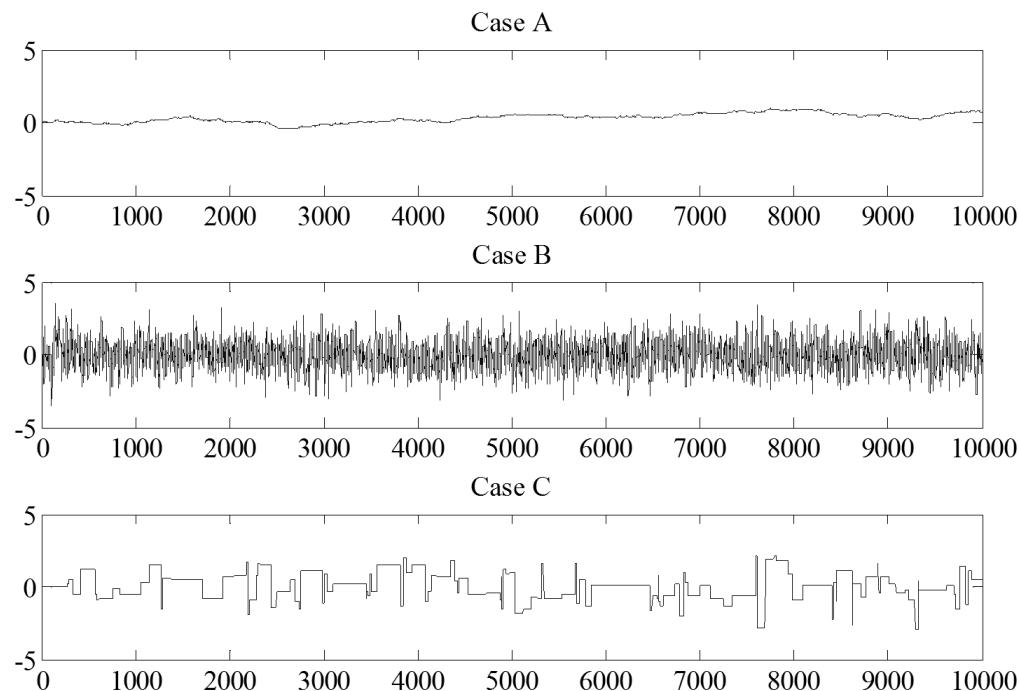


Figure 8-3 - Proposal Distribution Study - Adapted from Dubbs (2012)

Since typically the target distribution is not known in advance, this uncertainty can be addressed by using an adaptive proposal function which periodically evaluates the acceptance ratio of the chain, and accordingly adjusts the proposal to achieve a better acceptance (Haario, Saksman et al. 1999; Haario, Saksman et al. 2001). Figure 8-4 shows Case A from above, with adaptation to the proposal distribution every 1000 iterations. Qualitatively, it is observed that the chain begins to resemble Case B. Figure 8-5 shows the actual proposal function value, which approaches the ideal value of 1.

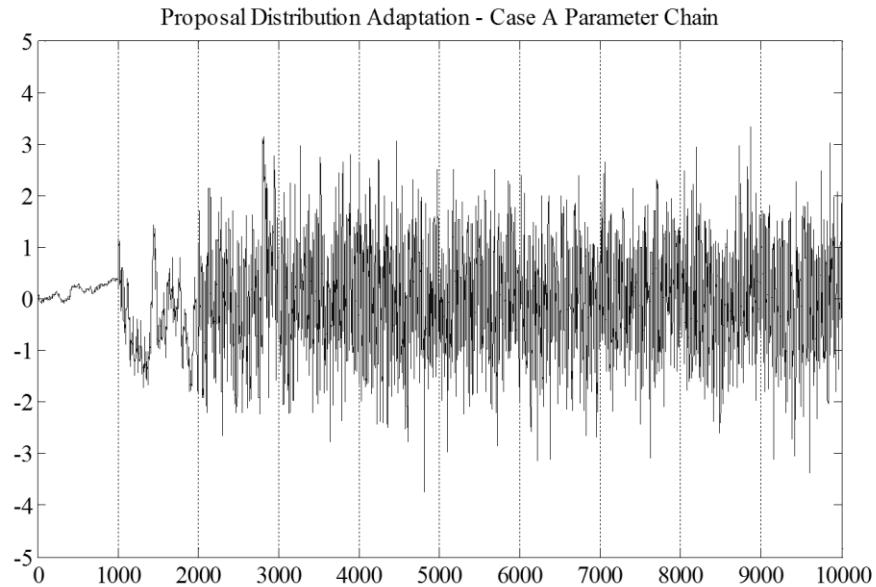


Figure 8-4 - Case A with Adaptation to the Proposal Function

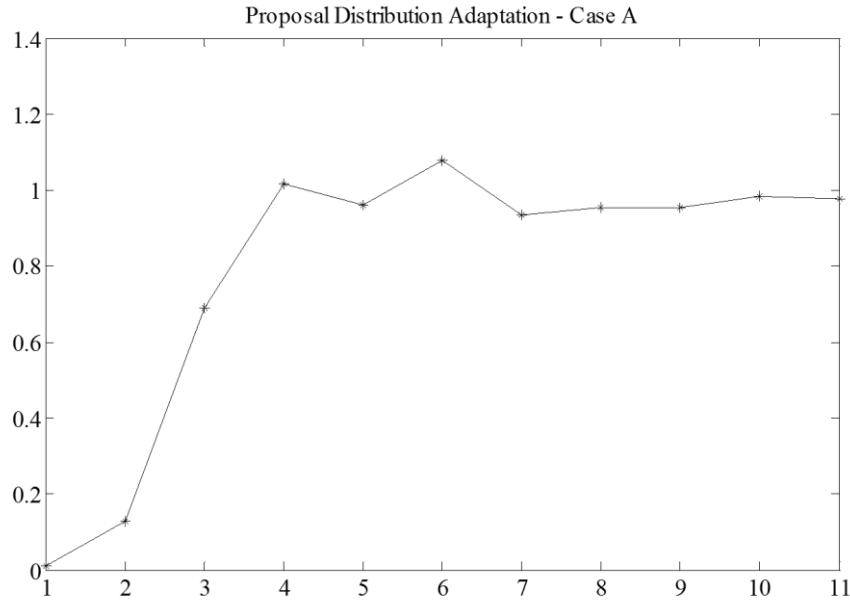


Figure 8-5 - Proposal Function Value with Adaptation

The major downside of adaptive proposal functions is that any time there is a change in the proposal function the chain no longer remains a true MCMC chain, and must effectively restart. Also adaptive methods are sensitive and can easily become unstable.

For this research, two methods of adaptation were explored. The first was a greedy start method, meant to rapidly locate areas of high likelihood on the posterior. This method adapts the proposal function based only using the unique accepted models of the chain, over a certain update frequency. The second method, termed adaptive proposal by Haario, Saksman et al. (1999), updates over a set frequency, but includes the entire chain, not just the unique instances of accepted models. The adaptive proposal method was eventually implemented fully because it was more stable than greedy starts.

8.3.2 Identifying Burn-in

MCMC methods require a certain number of samples to “burn in.” Burn-in is achieved when the parameter samples are no longer dependent on the starting point of the chain (Geyer 1991). Determining the required length of the burn-in period is not a trivial task. The chain output can be interpreted directly, or through a function of the output. Direct interpretation is essentially visually assessing the parameter chains for an apparent stabilization. To illustrate this, the first 1000 iterations of Case B are shown in Figure 8-6 with a poor starting value. Qualitatively, the burn-in in this case would be approximately the first 200 iterations, though a longer burn-in is conservative.

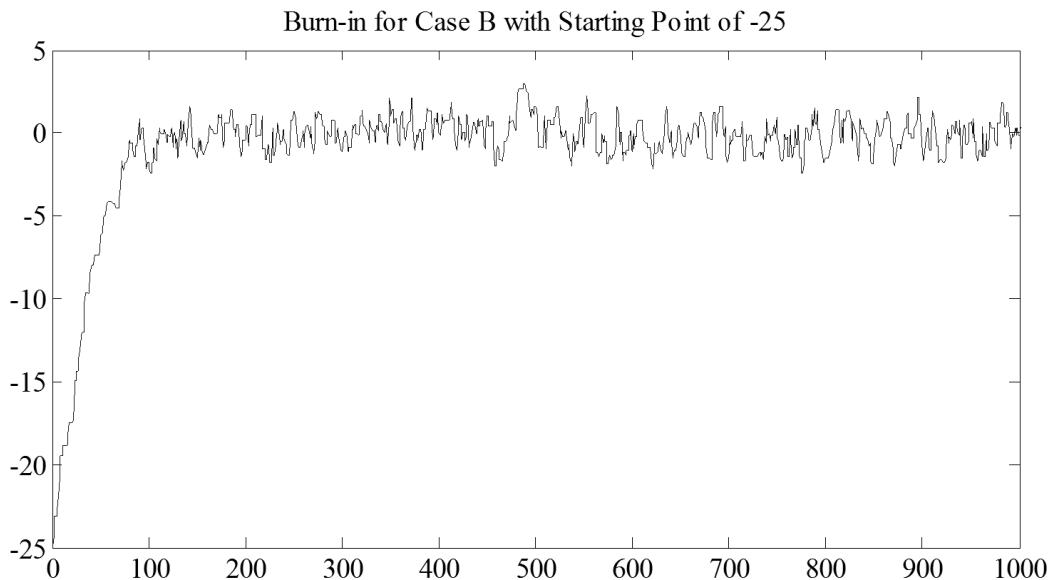


Figure 8-6 - Burn-in for Case B with Poor Starting Value

For a more quantitative approach, Geyer (1991) suggests that burn-in has occurred when the autocovariance of the chain has reached negligible levels, which he states is typically 1-2% of the entire run length.

8.3.3 Assessing Convergence

It is similarly difficult to determine when the convergence of the proposal distribution to the posterior distribution has occurred. Traditionally, for single model MCMC, convergence is based on stabilization of the parameters, but it is often recommended that several chains be run in parallel with different starting points (Gilks, Richardson et al. 1996). An overview of convergence diagnostics was provided in Cowles and Carlin (1996) as well. Many of these diagnostics can be directly applied to burn-in as well as the overall convergence. Other newer convergence techniques have been proposed, including a Hybrid Markov Chain Monte Carlo procedure developed by Cheung (2009) with particular burn-in convergence criteria associated with it.

For this effort, the convergence checks implemented in Dubbs (2012) are used for single MCMC runs. These include checking the convergence of means and standard deviations of the parameter chains.

These three aspects of MCMC methods, (identifying burn-in, assessing convergence, and selecting proposal functions) are absolutely fundamental to the process and are still very difficult to address despite substantial research. Further discussion of these aspects is provided in the context of the investigations in later chapters.

8.4 Reversible Jump Markov Chain Monte Carlo

RJMCMC methods allow for the development of a true Markov Chain across model forms of different dimensions by linking the parameters using jump equations. The method was proposed by Green (1995). In many cases, the parameters are nested, and the jump equations are simple arithmetic. In other cases, the transition between models can be much more complex. The general outline of the RJMCMC sampling strategy is shown in Figure 8-7.

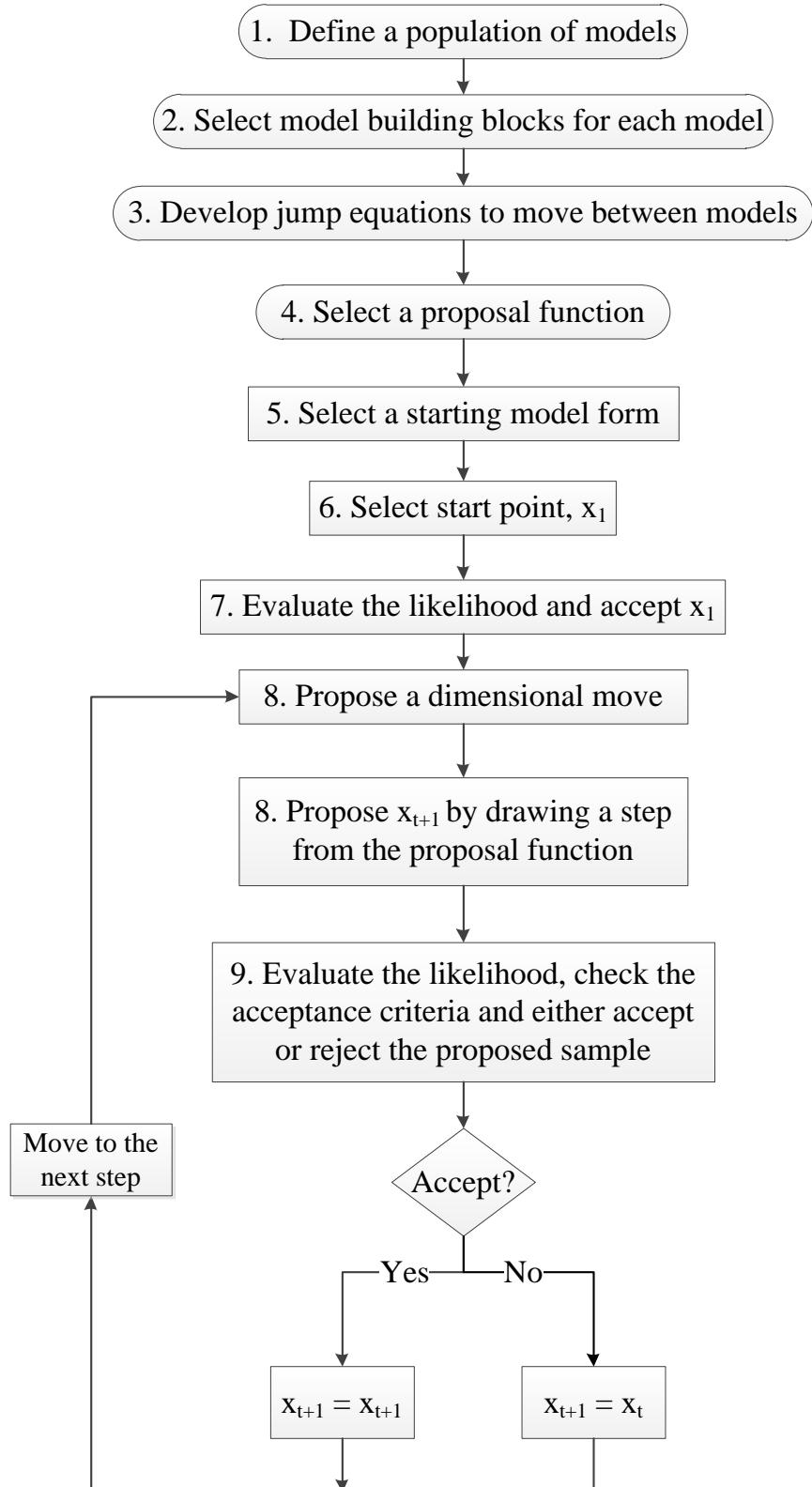


Figure 8-7 - RJMCMC Process with Metropolis-Hastings Algorithm

The process is described below:

1. Select the applicable model forms for the analysis
2. Assign a prior probability to each model form
3. Run the RJMCMC analysis using the MH Algorithm
 - a. Randomly select, run and accept a model
 - b. Propose a dimensional move and generate a parameter sample using jump equations
 - c. Evaluate the given model form with the selected parameter set
 - d. Determine the acceptance ratio, α , using Eqn. 8-3
 - i. If $\alpha > 1$, then accept the model
 - ii. If $\alpha < 1$, accept with a probability of α , and reject otherwise
 - iii. If rejected, return to the previous state
4. Repeat these steps until the analysis burns in after which convergence of the response predictions has occurred

For the RJMCMC analyses, convergence will be based on the stabilization within reasonable limits of the mean and standard deviation of the response predictions. The reason for using response predictions as opposed to parameters is two-fold. First, the response predictions are the unknown quantity of interest and they stay consistent across model forms where parameters may change. Second, the parameters in a model that is not sampled frequently (low prior probability) or not accepted frequently (relatively low likelihood) will not stabilize for an excessively large number of total iterations.

8.5 Multiple-chain Markov Chain Monte Carlo (MC3) Methods

The MC3 method will be very similar to RJMCMC, except that the parameters are not transformed when a dimensional move to a different model occurs. This distinction specifically addresses a shortcoming to RJMCMC that is discussed in Section 9.4.5. The general procedure for any multi-dimensional Bayesian analysis is outlined below:

1. Development of a candidate population and parameter sets
2. Determination of prior model weights
3. Selection of a sampling methodology and if required, an appropriate starting point
4. Calculation of response predictions

This framework can be used with many combinations of sampling and weighing methods. As a frame of reference, Monte Carlo sampling will be investigated, but generally, MCMC sampling will be implemented because of the documented increase in efficiency.

8.5.1 Step 1: Development of a candidate population and parameter sets

Step 1 consists of selecting applicable model forms and parameters. This step involves potentially the most heuristic structural engineering knowledge of the entire process.

While any Bayesian process implicitly penalizes selection of inappropriate models, there is no recourse for accidental exclusion of appropriate models. Therefore a sound basis in structural engineering and modeling is required in order to develop a reliable candidate model population. Developing an adequate candidate model population includes both

selection and spatial distribution of parameters in the model as well as modeling techniques.

These decisions should be addressed at least in a preliminary sense through the manual calibration of a single model. It is important to consider that the goal is not to select only the parameters that will allow the model to accurately represent the experimental data, but also to include parameters that the modeler believes are crucial to understanding the response predictions of interest. Specifically, this means that a parameter may not be at all informed by the experimental observations, but if it is important to the response prediction, it must be included. An example would be density of concrete as a parameter with static displacements as experimental observations and dynamic properties as the desired response predictions.

It is recognized that there are a theoretically infinite combination of parameters and spatial distributions which is why the heuristic knowledge of the engineer is so important. Regardless, one criticism of these methods is that there is no way to be 100% sure that there is not a better model form/parameter set that is not included in the candidate model population. While this is true, recognizing this shortcoming and at least attempting to address it by using multiple models is a more honest and defendable approach.

8.5.2 Step 2: Determination of prior model weights

Recall Eqn. 8-1 for Bayesian Inference, which shows that posterior probability of a model given some experimental data is equal to the ratio of the prior probability of the model multiplied by the likelihood of the data given the model over the summation of

these quantities for the entire modeling space. Bayesian methods are predicated on the idea that existing knowledge is incorporated into the estimation of the probability of an event in part via prior probabilities. In traditional single-model methods, the model prior probability is simply the product of the parameter prior probabilities within that model. When multiple model forms with different numbers of parameters are included, assigning a prior is difficult. One simple option would be to equally weight the models. The weight of each model would be:

$$P(M_i) = \frac{1}{N} \quad \text{Eqn. 8-4}$$

where N is the number of models used in the analysis.

Another approach, with the potential to be much more efficient, is to assign the model weight through the calculation of a Bayesian Information Criteria (BIC) factor for each model form. BIC attempts to approximate the denominator of Eqn. 8-1. The typical use of BIC is for comparison and selection of model forms prior to an analysis.

Theoretically assigning a weight based on BIC is defendable because it addresses several key issues which arise when trying to assign weights to models with different dimensions. There are three key perspectives regarding model weight to be addressed:

1. Goodness of fit compared to the observations,
2. The number of observations as compared with the number of free parameters

3. The influence of the included parameters in terms of both observations and the desired response predictions.

The BIC calculation explicitly attempts to account for the first two. The formula for BIC is shown in Eqn. 8-5.

$$BIC = -2 * \ln(MLE) + k * \ln(n) \quad \text{Eqn. 8-5}$$

where MLE is the maximum likelihood estimate from an optimization technique, n is the number of free parameters in the model form, and k is the number of known responses or pieces of data. There are several other criteria, including the Akaike Information Criterion (Akaike 1974) which does not explicitly include the number of observations, and a method developed by Beck which incorporates a factor to explicitly account for parameterization (Beck 2004). None of these criteria reflect the need for the included parameters to influence the response predictions.

8.5.2.1 Determination of MLE

The MLE estimate for each model form is determined through a nonlinear least squares optimization (LSQNONLIN) using the MATLAB optimization toolbox. This function requires an objective function, parameter bounds, and a starting point as inputs. For model updating and parameter identification, objective functions to be minimized are associated with residuals of percent errors between measured and analytical responses. In

this case, because the maximum likelihood was a desired output from the process, the following objective function was employed:

$$obj = -\ln(f(D|M)) \quad \text{Eqn. 8-6}$$

where $f(D|M)$ is the likelihood of the experimental data given the model form and parameter set. By using this objective function, as opposed to percent errors of responses, the MLE value required for BIC calculation was readily available.

The LSQNONLIN algorithm requires parameter bounds as an input. If the selected bounds are not adequate, the optimization will pin a value at the bound and optimize the rest of the parameters accordingly, which is not a true optimized solution or global minima. Often bounds are determined by the physical realities of the world. In other cases, as is typical with springs representing boundary restraints in a finite element model, the parameter is representing a quantity that is somewhere between zero and infinity. These types of parameters are typically asymptotic to a certain value, and the upper bound should be defined based on when that asymptote is “close enough” to its limit. To determine this point, sensitivity studies are conducted. These studies examine how sensitive the quantity of interest is to changes in a parameter, with all other parameters held constant at nominal values.

Deterministic sampling is the easiest and most transparent way to conduct these studies, though it is not without flaws. A range for a parameter and a number of desired samples is defined, and the range is divided up evenly. The model is rerun at each parameter

value, with all other parameters remaining constant, and the objective function of interest is evaluated. Traditionally, the objective function would utilize the observations which will be used in the optimization. However, in the case where the desired result is not the parameter values, but the response predictions, it is important to investigate the sensitivity of these predictions over the specified ranges. Based on analysis of the results of these studies, the range may be adjusted and the analysis rerun or it may be deemed an acceptable range.

The notion that the parameter ranges should inform both the chosen observations and the desired response predictions leads naturally to an experimental design approach. The model serves as a translation between what can be measured and what is desired and can therefore be used to help select the most informative response measurements relative to the selected unmeasurable predictions for an experiment.

In cases where multiple parameters cause similar changes in the objective function, several sets of nominal values for the stationary parameters may be required. Ideally a full factorial sensitivity study could be utilized, though this is not practical or efficient. For the cantilever beam, the parameters all had the same bounds, which were defined based on physical requirements.

The optimization process is very sensitive to starting point, as it is quite possible, if not likely, to locate a local minimum on modeling space. Therefore, to provide the most reliable estimation of maximum likelihood, each optimization was run ten times, with randomly generated starting points. Using these results, the MLE value was taken as the most likely of the ten iterations.

8.5.2.2 Determination of weights using BIC values

The absolute magnitude of the BIC value is not necessarily of that much value. The relative relationship between two BIC values for two different models indicates which model is more suitable to represent the experimental data. Therefore, BIC is appropriate to use as a starting point to derive the prior probability for each model form being examined.

MLE ranges from 0 to 1 and the first term of Eqn. 8-5 approaches zero as the MLE goes to 1. This indicates that a lower BIC value is tied with a higher likelihood model. Similarly, the second term decreases as the ratio of n/k increases. This term penalizes the BIC for models with a large number of parameters, relative to the number of observations.

The prior model weight using BIC is derived as follows:

$$P(M_i) = \frac{\frac{1}{BIC_i}}{\sum_i \frac{1}{BIC_i}} \quad \text{Eqn. 8-7}$$

8.5.3 Step 3: Select a sampling method and an appropriate starting point

All multi-dimensional Bayesian methods assume it is logical to compare different model forms and parameter sets. For RJMCMC, different model forms are linked through both the acceptance criteria and the parameter values, which are translated between model

forms using jump equations. The MC3 approach will only link model forms using the acceptance criteria, and thus each model will have its own parameter chains. At first glance this is a violation of the requirements of a Markov chain, in that the subsequent sample is not specifically dependent on the prior sample (i.e., when a step is taken to a different model form, the next sample generated is based on the parameter values in the chain corresponding to the current model, not the prior model). However, since the goal is not to determine parameter values, but rather response predictions, the use of acceptance criteria to select models could be viewed as a method of weighting, meaning that strictly adhering to the requirements of Markov chains for any resampling may not be required.

For traditional Monte Carlo sampling, the starting point is not a concern. The starting point is only of importance when MCMC sampling is used because this approach aims to draw samples directly from the posterior distribution, which is not consistent with the use of an arbitrary starting point. Rather, a burn-in period is required in which the MCMC sampling essentially “forgets” or is no longer conditioned by the starting point. When MCMC sampling is used for multiple models, each model form must be burned in individually. The individual model burn-in can be adaptive; meaning that the proposal function is periodically updated based on the parameter chains. This helps to accelerate the process of burn-in. The most efficient MCMC analysis uses the actual posterior probability distribution as the proposal function, so from an efficiency standpoint, it makes sense to update the proposal as more information is gained regarding the posterior distribution.

Determination of convergence of the burn-in period is achieved through a combination of several factors. The mean and standard deviation of the parameters of the accepted models should be stabilized. Similarly, the mean of the likelihood should stabilize. There is an intuitive, heuristic aspect to assessing burn-in as well. Qualitatively, the parameter chains can be assessed visually to determine if the analysis is meandering, indicating a small proposal function, or if it is flat-lining, indicating too large of a proposal function. If desired, convergence for a single chain could also be the point at which the response prediction mean and standard deviation have converged. The reason that this may be expensive is that in the case of finite element models, typically a different solver is used for different types of results. If modal properties are used in the updating and strains are used as a response prediction, this would require twice as many model iterations, which is potentially not feasible simply for checking convergence.

8.5.4 Step 4: Development of response predictions

The final step of any multi-dimensional Bayesian analysis is to develop robust response predictions which implicitly account for uncertainty in the model parameters and the model forms. This is achieved by sampling the model form (using prior weights) and then the parameters within that model (using the prescribed sampling methodology). Convergence of the multi-dimensional Bayesian analysis cannot be checked via stabilization of parameters, as that would have to be done on a model by model basis and certain models will be sampled very sparsely. Instead convergence will be dependent on response predictions, as these are the quantities of interest.

8.6 Approach to Multiple Model Methodology Development

Three different perspectives will be used to approach multiple model methods. The context for multiple models could be as different values of parameters in a fixed model form, as different, but dependent parameter sets across model forms or as independent parameters sets and model forms.

The approach for the remainder of this research is to investigate these perspectives as follows:

1. Exploratory investigations using a numerical study of a transparent, intuitive cantilever beam – The aforementioned complications that arise with constructed systems even in single model-experiment correlation prevent exploratory investigations using an actual structure, like the IBS Bridge. Therefore, a simple cantilever beam will serve as a test bed for several different multiple model methods. The best approach for application to the IBS Bridge will be selected based on the cantilever investigations. (Chapter 9)
2. Application to the IBS Bridge – The selected method(s) will be applied to the IBS Bridge problem considering all three perspectives described previously. The single-model St-Id and the subsequent heuristic knowledge base developed during that process will serve as a starting point for this analysis, meaning that any information or intuition developed since the conclusion of the traditional St-Id will be included here. (Chapter 10 and Chapter 11)
3. Comparison to traditional, single-model St-Id – In order to quantify and understand the differences between multiple model methods and traditional St-Id, a series of comparisons will be developed. These comparisons will focus on

response predictions including a portion of the experimental results “held back” for verification purposes and other unmeasurable quantities of interest, like load ratings. The end goal of these methods is aid in owner decision-making, and metrics like load rating are more important and intuitive to a bridge owner than peak response values or mode shapes. (Chapter 12)

Chapter 9: Numerical Study of Various Multiple-Model Approaches

This chapter will examine, evaluate and compare, in quantitative terms, the most promising MM St-Id approaches using a numerical study. The numerical study allowed for consideration of multiple approaches because the model analysis time is substantially faster than a full finite element model execution using an independent, third party software package. Again, this study was approached through three perspectives.

9.1 Multiple Model Applications for a Numerical Example

The first perspective consisted of traditional Markov Chain Monte Carlo (MCMC) sampling of a fixed parameter set and model form. The second perspective was Reversible Jump Markov Chain Monte Carlo (RJMCMC), where the Markov Chain moves through different model forms via jump equations. The third perspective was a Multiple-chain Markov Chain Monte Carlo (MC3) method. This method is similar to the MCMCMC method described by Geyer (1991) except that the chains are different model forms. The application could be considered a type of population MCMC as well, with a two-part sampler (random to select the chain and Metropolis-Hastings to select the parameters). This approach differs from RJMCMC because while the acceptance of a given model is still dependent upon comparison of likelihoods through an acceptance criterion, in this case the parameter chains remain independent of one another across model forms. The latter two can be broadly categorized as trans-dimensional MCMC approaches. All three perspectives will be presented in the context of an analytical cantilever beam.

9.2 Cantilever Beam System

A cantilever beam was selected for the preliminary investigations into multiple model methods because it is simple and readily interpretable. It is comprised of a variable height section with a uniform width of 1 inch and a length of 120 inches. Both the number of individual sections and the heights of these sections are considered parameters in this analysis. Since this is an entirely numerical study, the “experimental” results were arbitrarily selected as the output from the model shown in Figure 9-1.

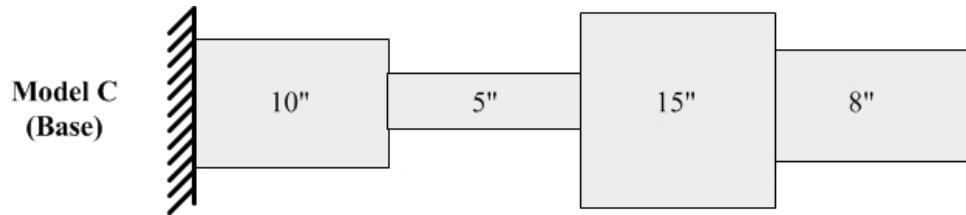


Figure 9-1 - Base Cantilever Model

The FE model used for the study is comprised of a total of 24 elements over which the height can be varied. The different model forms to be investigated were defined by varying the spatial distribution of the section height parameters along the length of the beam. The model forms investigated are shown in Figure 9-2. From this point, each model will be referred to as a designation from Model A through Model F, where Model C is the base case described above.

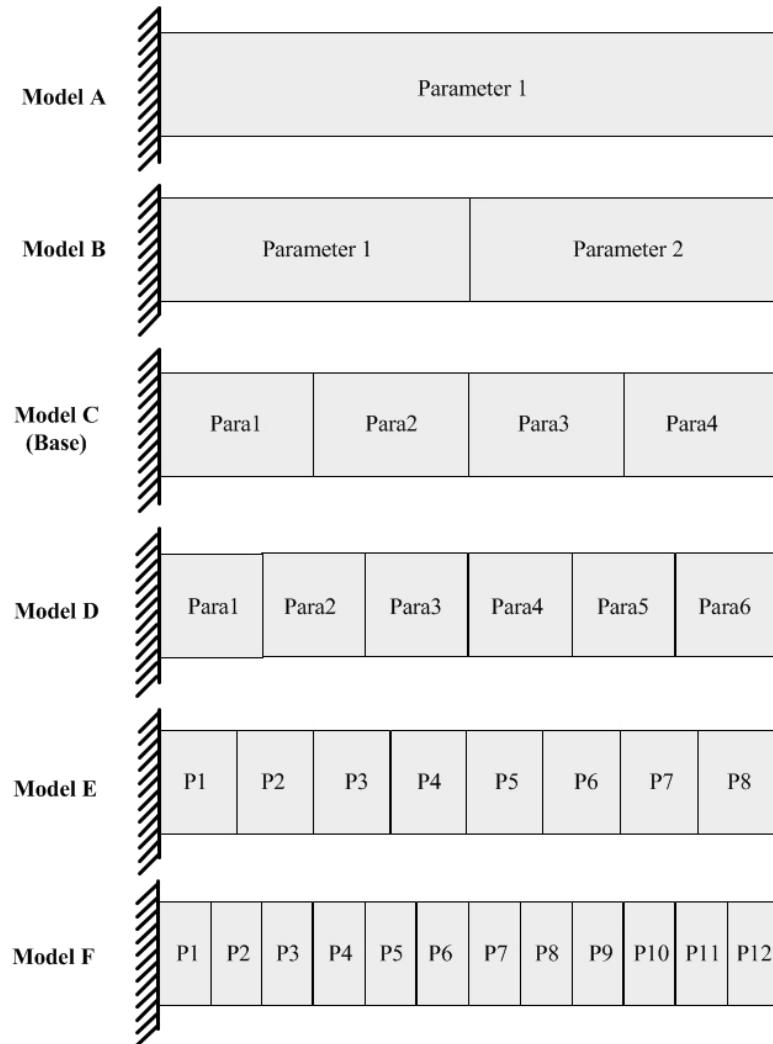


Figure 9-2 - Generic Model Forms of Cantilever Beam

The simplicity of the structure allows for the use of a MATLAB-based finite element code as opposed to a traditional finite element software program. This software greatly reduced computational time and allowed for a full investigation and validation of several different potential multiple model solutions, through a large number of model executions.

In order to assess the applied approaches, a framework based on informing the analysis with one set of data and using the results to predict another set of data is implemented. This translates well to the real-world scenario typical to a bridge where the response of interest (capacity/load rating) is unmeasureable, and the responses that are measureable (strains, displacements, accelerations) may or may directly inform the unmeasureable quantity.

Modal data was selected as the observed data because it is typically easiest to collect, and it is influenced by both mass and stiffness. Similarly, the selected parameter, h , has an effect on both mass and stiffness. Note that the structure was considered planar and axially rigid. The experimental results (from Model C) are shown below in Table 9-1.

Figure 9-3 shows the experimental mode shapes.

Table 9-1 - "Experimental" Frequencies

Mode #	Frequency (hz)
1	13.714
2	113.00
3	363.00
4	553.92
5	1142.5

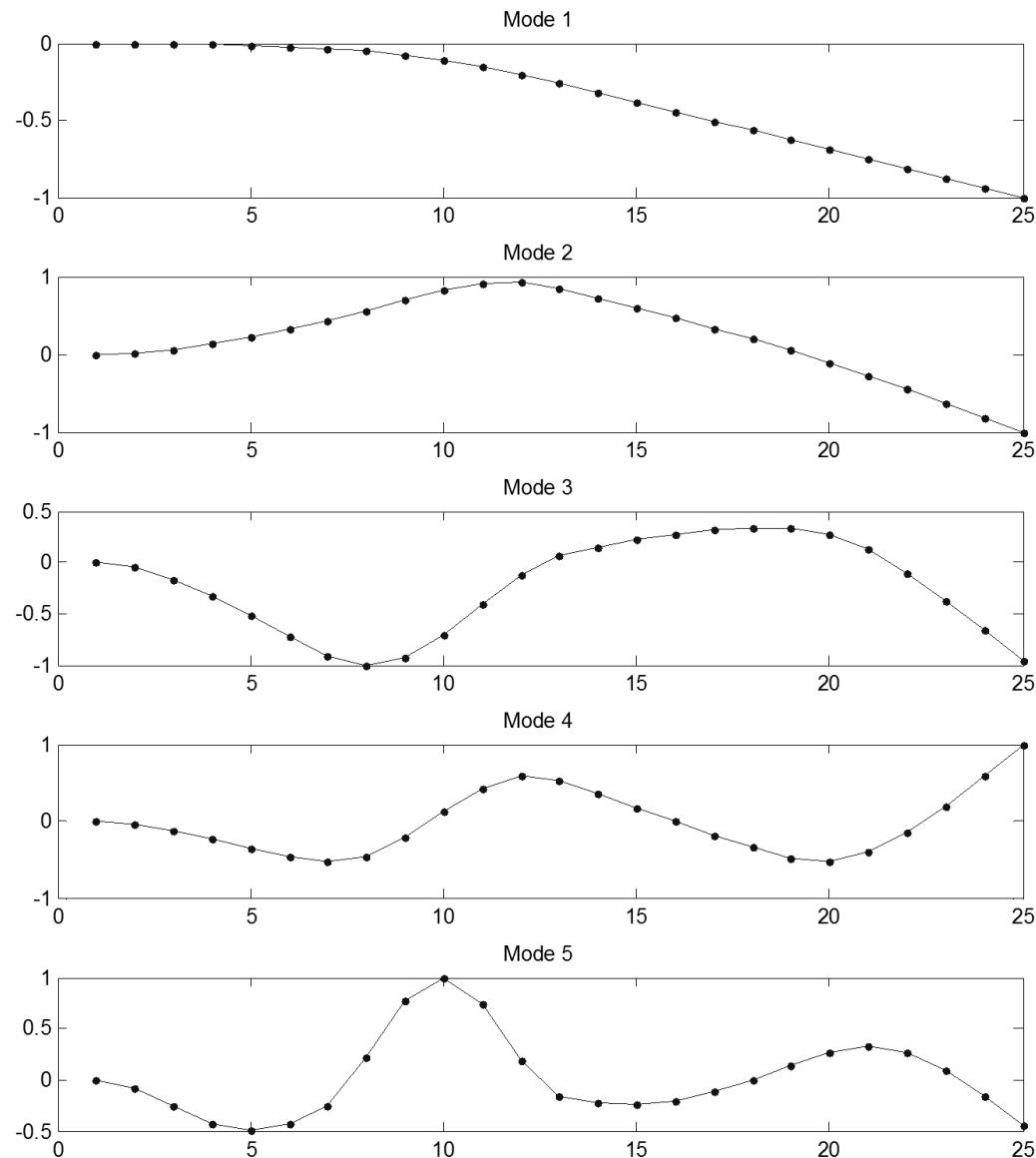


Figure 9-3 - Experimental Mode Shapes

Tip displacement and strain at two particular locations (Figure 9-4) due to a point load at the tip were chosen as the unmeasureable data. These unmeasureable responses were at

least partially tied with the experimental data, but the extent of how well they can be informed was not known. The unmeasureable data is presented below in Figure 9-4.

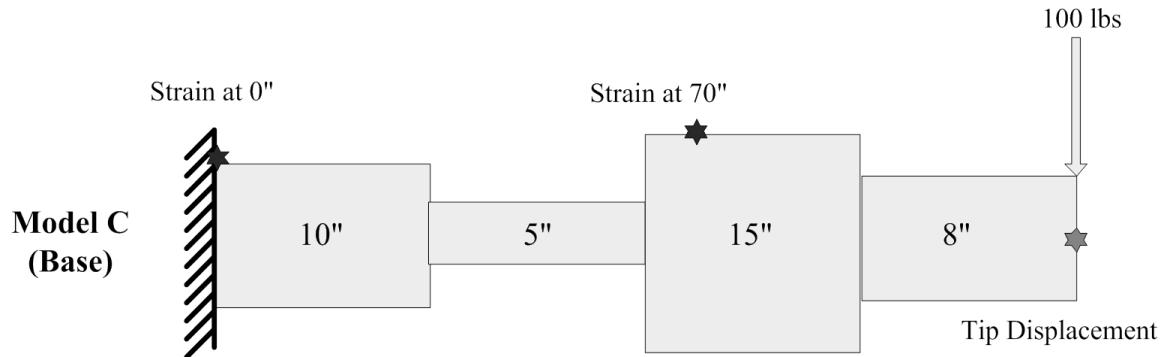


Figure 9-4 - Location of Load and Unmeasurable Responses

Table 9-2 - Unmeasurable Responses

Response	Location	Magnitude
Displacement	Tip	-.072 in
Strain1	0"	24.83 $\mu\epsilon$
Strain2	70"	4.60 $\mu\epsilon$

9.3 Cantilever MCMC Analysis

Each cantilever beam setup shown in Figure 9-2 served as a base model form for a single model MCMC analysis. For discussion, the results from Model E (8 parameter case) are included here though every model was analyzed.

Figure 9-5 and Figure 9-6 show the MCMC output from each parameter of Model E (chain) over a total of 200,000 simulations. The transparent nature of the numerical model led to very short burn-in periods of less than 500 model iterations. The analysis resulted in 22,292 accepted models, meaning the acceptance rate for this model form was approximately 11%.

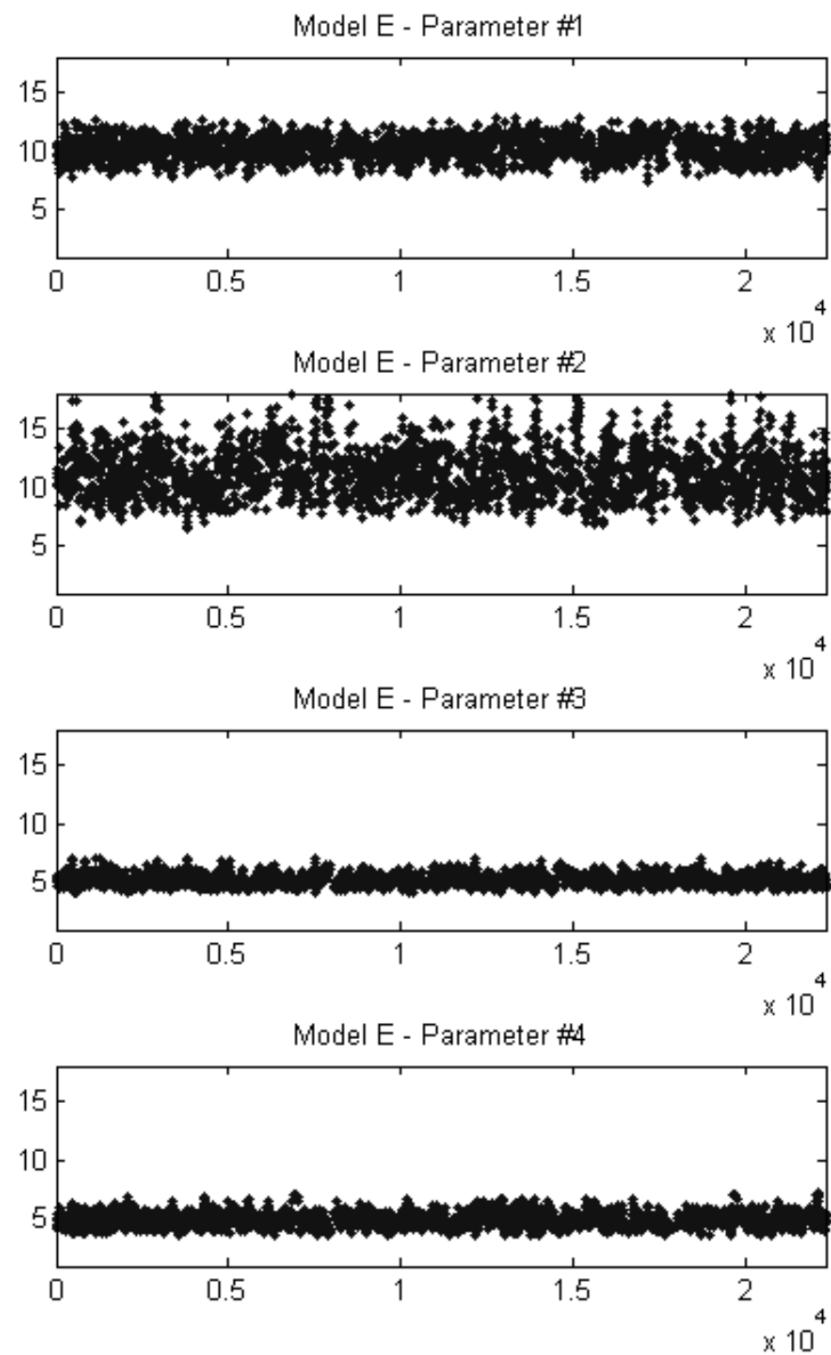


Figure 9-5 - MCMC Output Chain - Model E - Parameters 1 through 4

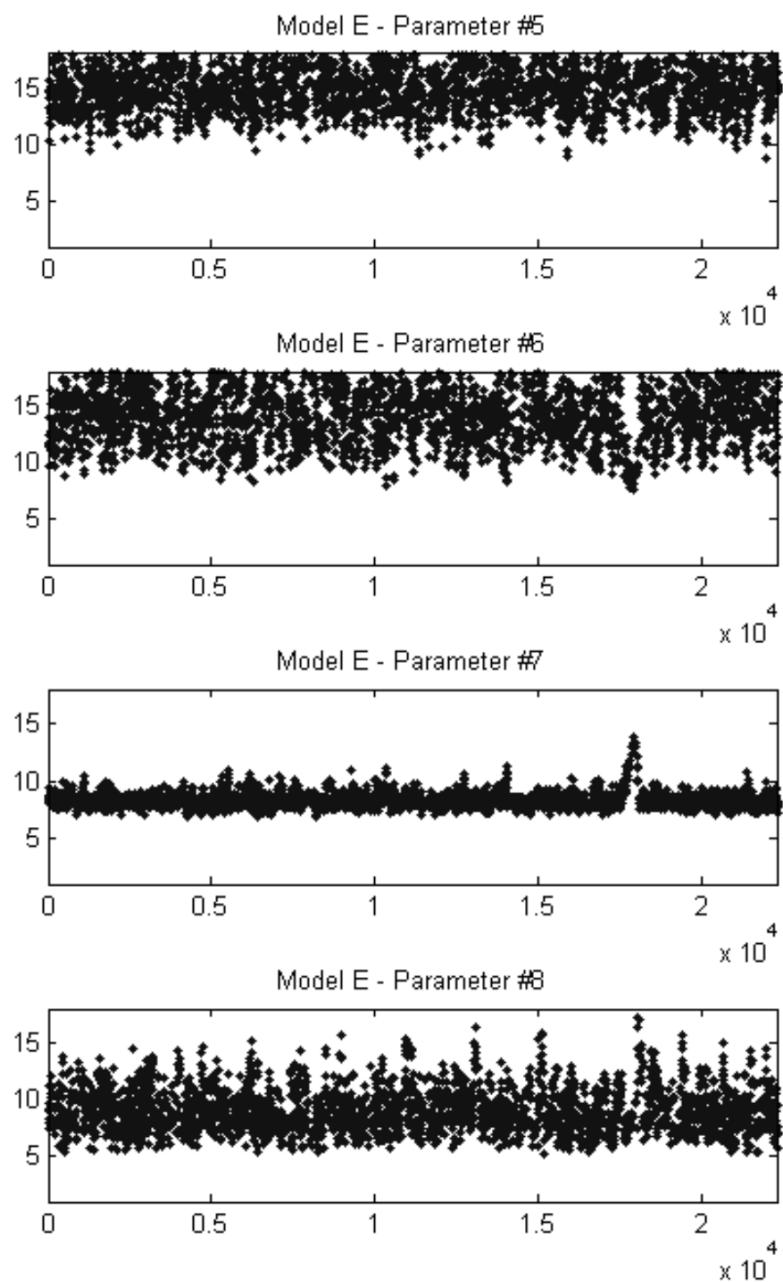


Figure 9-6 - MCMC Output Chain - Model E - Parameters 5 through 8

Note that this eight parameter case is capable of exactly replicating the base case model of four parameters. In order to do that, the pair of parameter 1 and parameter 2 would be equivalent, as would 3 and 4, and so on. In the case of parameters 3 and 4, as well as 5 and 6, the chains look nearly identical. In the other two pairs, substantial differences are observed. This is a direct reflection on the ability to identify the height parameter using only global dynamic properties. At certain locations along the length of the cantilever beam, either the stiffness or the mass had a particularly large effect on the modal properties. In these regions (for example, parameter 3 and 4) the chain is very tight. In areas of the beam where both mass and stiffness affect the modal properties, the chain is much wider, proportional to the uncertainty.

Model E was one of the more complex models investigated in the cantilever study, and the stability of the MCMC algorithm reflects this complexity. Figure 9-7 shows a section of the likelihood of all models from the MCMC sampling of Model E. There are clear instances of the sampling moving gradually, over several samples, from areas of higher likelihood to areas of lower likelihood. This does not occur with the more simple models, which have a clearly defined region of maximum likelihood that MCMC can fully explore and represent without difficulty. Even for the cantilever beam, the complex likelihood space makes fully representing the posterior space difficult.

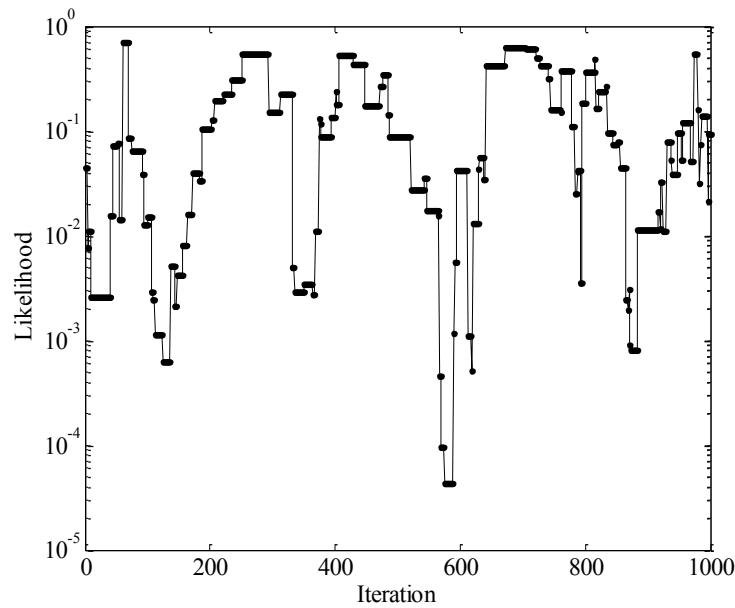


Figure 9-7 - Gradual Drop in Likelihood - Model E

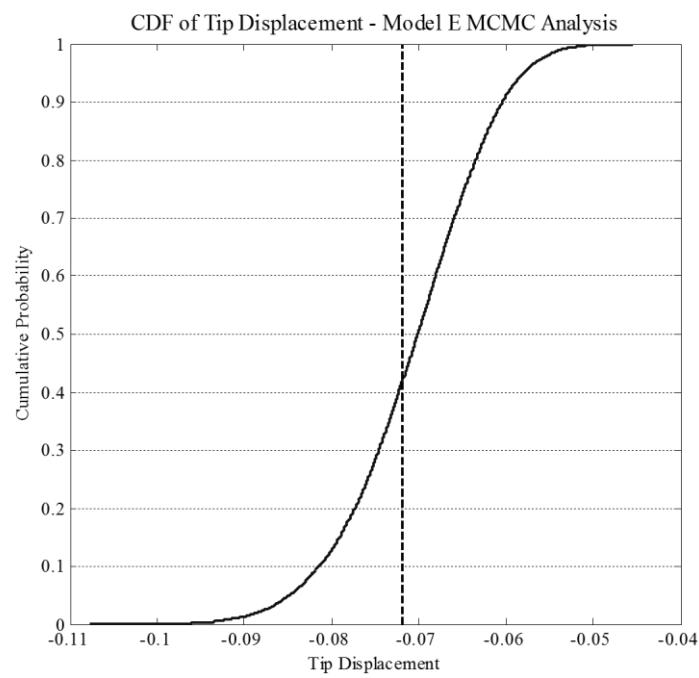


Figure 9-8 - CDF Plot for Tip Displacement - Model E

Figure 9-8 shows a cumulative distribution function of the displacement predictions based on the chain of accepted models. Included on the plot is the experimental value as well. The analysis produces a reasonable prediction displacement as compared with the experimental results.

For the sake of comparison, the CDF for tip displacement for Model D, the 6 parameter case, is shown in Figure 9-9. Model D is not directly able to replicate the base case, and is less accurate in response prediction. Additionally, there is more variability in the response predictions as well.

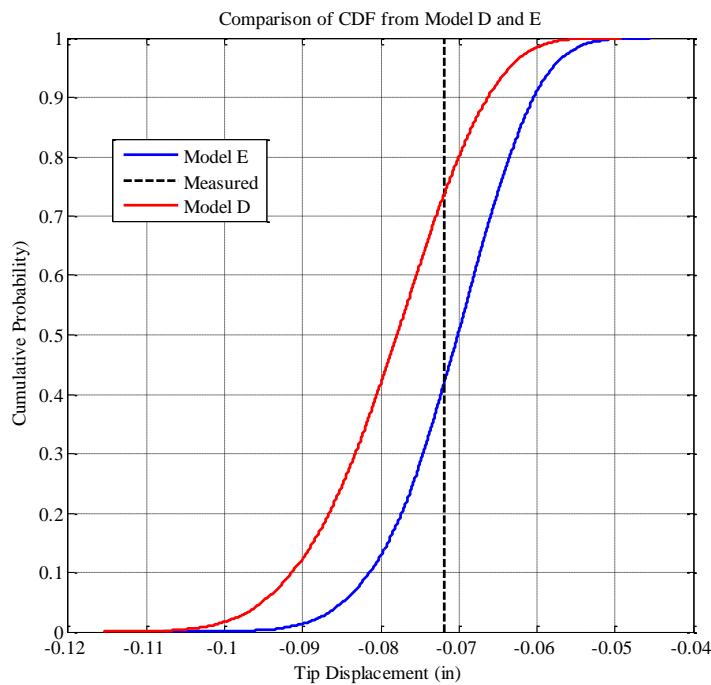


Figure 9-9 - Comparison of Model D and E - CDF of Tip Displacement

Considering the changes that occur in the response predictions based on minor re-parameterization of the simple cantilever model, it is clear that, just like traditional single model-experiment correlation methods, a single model MCMC chain is very dependent on the selected model form and building blocks. Where MCMC differs from single model approaches is that it explicitly represents and accounts for the variation in the parameters and how that affects the response predictions. The process could be improved by extending it to explicitly consider the effects of model form as well, which is the motivation for the RJMCMC and MC3 methods.

9.4 Cantilever RJMCMC Analysis

9.4.1 Establishing the Problem

The RJMCMC analysis of the cantilever benchmark problem consisted of two scenarios. This study used Models A, B, C, D and E. Recall that the form of Model C can only be physically replicated by Model E, though Model E has twice as many parameters. The two scenarios examined the effects of including the actual model in the population of candidates versus not including it.

In the first scenario the experimental data was the output from a single parameter set in Model D. This scenario was included to see the effects of including a model which was very capable of representing the experimental data in the analysis.

In the second scenario, the experimental data was from Model C, but the candidate population excluded this model. This scenario explicitly represented the case where the

actual model was *not* in the model population which is the situation for real structures where models can only approximate the real system.

Since the various models of the cantilever system were nested (i.e. the only variation was associated with the spatial distribution of parameters) the jump equations were fairly straightforward to develop and implement. All of the parameters have the same units, the same bounds, and were on the same order of magnitude. Therefore the jump equations were simple mathematical combinations of the parameters. For example, when moving from Model A to Model B, the single parameter in A would just be applied to both parameters in Model B. Conversely, a move from Model B to Model A would simply be an average of the two parameters in Model B. A simplistic version of jump equation to move from Model A and Model B is shown in Eqn. 9-1. A similar equation to move back to Model A from Model B is shown in Eqn. 9-2.

$$h_{1|Model\ A} = h_{1|Model\ B} = h_{2|Model\ B} \quad \text{Eqn. 9-1}$$

$$h_{1|Model\ A} = \frac{h_{1|Model\ B} + h_{2|Model\ B}}{2} \quad \text{Eqn. 9-2}$$

The jump equations were formulated into matrices and had to satisfy the requirements of reversibility. In other words, it must be possible for the sampler to return to the exact same state it just came from. Also, typically a random step is taken after the move to a new dimension in order to save computational time, which was also included in these

equations. A more rigorous discussion of this formulation is available in Dubbs (2012).

The movement between models, averaging when moving to lower parameterized models and splitting when moving to higher parameterized models, is partially a function of this particular structure, but is also based on common sense. Construction of this type of move in other scenarios which are not as trivial as this problem may be much more difficult.

For the sake of simplicity, it was assumed that there was no prior knowledge regarding the relative validity of these four models, which means the prior probability of the models was all equal. The prior probabilities of the parameters were selected as uniform distributions. The bounds of these distributions represented the amount of prior knowledge about the parameters that was available. If a sample exceeded a bound, then it was assigned a probability of zero automatically.

The proposal function was kept constant for all parameters, and was not changed or adapted during the procedure. This is a preliminary application of RJMCMC and as such, it was decided that the more complex adaptive methods would be excluded initially. A small proposal function was chosen so that the acceptance rate of the analysis would be reasonable. It was known that this would not result in rapid mixing of the chain, but considering that the model is computationally inexpensive, this was deemed acceptable.

9.4.2 Scenario #1 – Actual model included in model population

Scenario #1 included the actual model in the population of candidate models. In order to limit the number of times that jump equations would be developed, a different set of

experimental data was used for the case with the model in the population. In this case, the experimental data was based on Model D with parameters as shown in Figure 9-10 and experimental properties listed in Table 9-3 and Table 9-4.

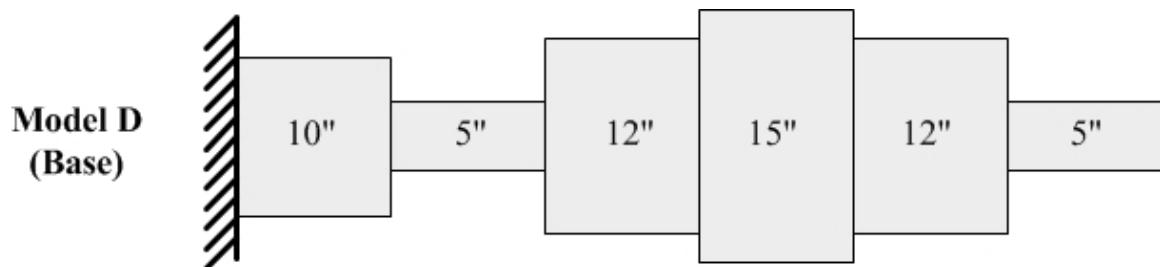


Figure 9-10 - Model D as the Base Case

Table 9-3 - Scenario #1 - Modal Frequencies

Mode #	Frequency (hz)
1	13.862
2	155.72
3	385.07
4	689.39
5	1023.9

Table 9-4 - Scenario #1 - Static Responses

Response	Location	Magnitude
Displacement	Tip	-.0682 in
Strain1	0"	24.83 $\mu\epsilon$
Strain2	70"	5.058 $\mu\epsilon$

The same proposal function was used for all parameters and all models. The process was not adaptive. The plot of the chain of accepted models in each model form is shown in Figure 9-11. The long flatline behavior indicates regions where no models were accepted in that particular form

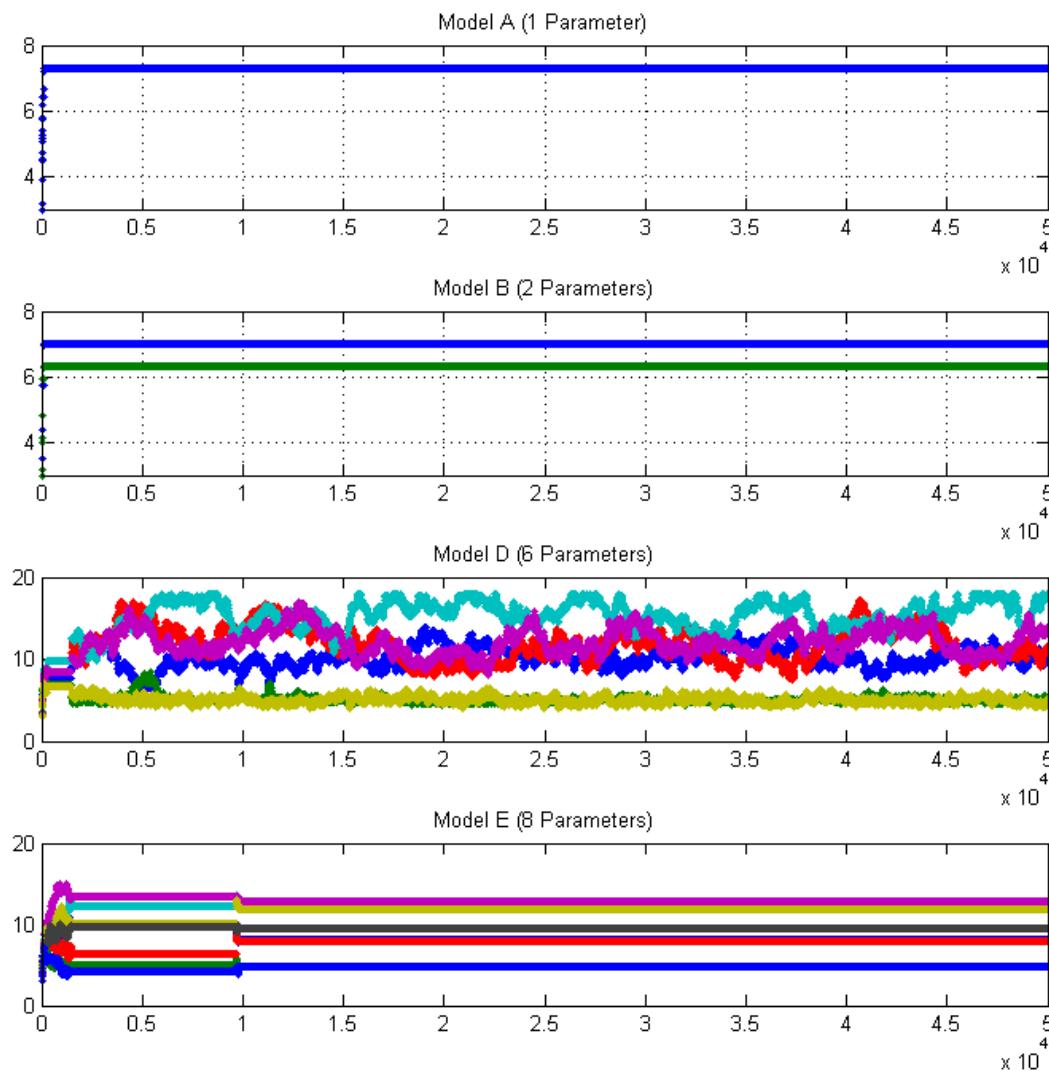


Figure 9-11 - Scenario #1 - Chain of Accepted Models

9.4.2.1 Model Form Selection

In the case where the exact model was included in the candidate model population, the RJMCMC algorithm selected samples almost exclusively from that model form. The discrepancy in the likelihood of the actual model predictions versus the other models is great enough that really only one model is accepted. This will not be the case when a real structure is being modeled, as it is impossible to have the actual model in the candidate model population.

9.4.3 Scenario #2 – Actual model not included in model population

The first scenario utilized Models A, B, D, and E with the experimental data based on Model D. This second scenario uses the same models, but the experimental data is based on Model C, which is not included in the candidate population. With the simple cantilever model, it was easier to run a substantial number of models and manually check convergence as opposed to conducting more frequent checks on smaller runs. The first scenario was run for a total of 50,000 iterations when convergence was checked. It was found that the mean and standard deviation of the response predictions had stabilized to within 5% of the final value.

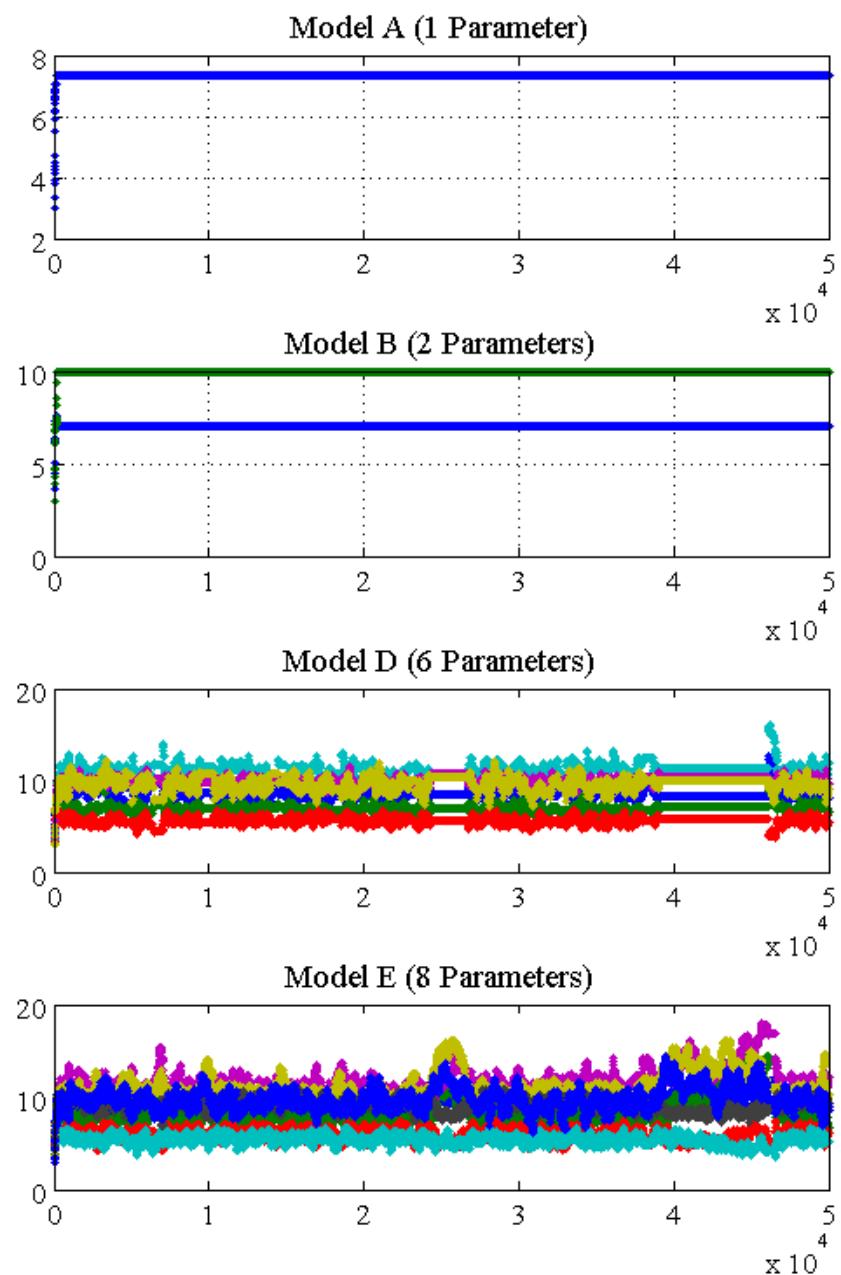


Figure 9-12 - Accepted Model Chain - Scenario #1

Figure 9-12 shows the chains parameter values (accepted and rejected) in each model form versus the global iteration number. There are several observations which can be made directly related to the chains as discussed in the following sections.

9.4.3.1 Resetting behavior

Model E has two regions where the parameters begin to move independently, at approximately 25,000 and 40,000 samples. The parameters reset whenever a sample from Model D is accepted. Based on this observation, it is clear that the simplest reasonable model conditions the parameters of the more complex models, in this case by way of the jump equations. The jump equations for moving from a more complex to a simpler model involve averaging the parameters over a larger spatial range to satisfy the smaller number of parameters in simpler model. Therefore whenever a model is accepted in a simple model, like Model D as compared to Model E, it will serve to “reset” the higher parameter models. This resetting phenomenon can be seen clearly from samples 40,000 to 46,000 in the full chain plot in Figure 9-11. This resetting occurs throughout the RJMCMC process.

The resetting appears to be related to the proposal function. For this analysis, a small proposal function was used and was not adapted. When a step was taken to a simpler model, the jump equations would average the parameters from the more complex model and then take a step in the new parameter layout. Similarly, when a step to more complex model was taken, the previous parameters would be split up and then step was taken in the new parameters based on the proposal function. Since the proposal function was small, these steps after a transition would not change the model substantially, allowing

the RJMCMC algorithm to step back and forth between model forms without rapid mixing and exploration of the entire model space.

9.4.3.2 Proclivity towards Model D and E

It is readily apparent that the majority of analysis was focused on Models D and E. Considering that all of the models have the same prior probability, the preference of the algorithm to Models D and E can be directly attributed to the acceptance probability. In other words whenever a step in Model A or B was attempted after the initial burn-in, it was rejected because it was so much less likely than the previous step. In this way, RJMCMC has inherent model selection attributes. Model selection is a well-documented use of RJMCMC in statistics and other fields, but is a new concept for civil structures. (Gilks, Richardson et al. 1996)

9.4.3.3 Rapid Burn-in Period

Through observation of the first 250 samples, it is clear how the analysis burned in, and how RJMCMC methods can potentially accelerate this process. A portion of the burn-in period is shown in Figure 9-13. All four models were started with all parameters at a single height of 3 in., and therefore were essentially the same model. As the burn-in began, models were most frequently accepted in Model A and B, and when models were accepted in D or E, the parameters were closely clustered, meaning they are essentially identical to the simpler models.

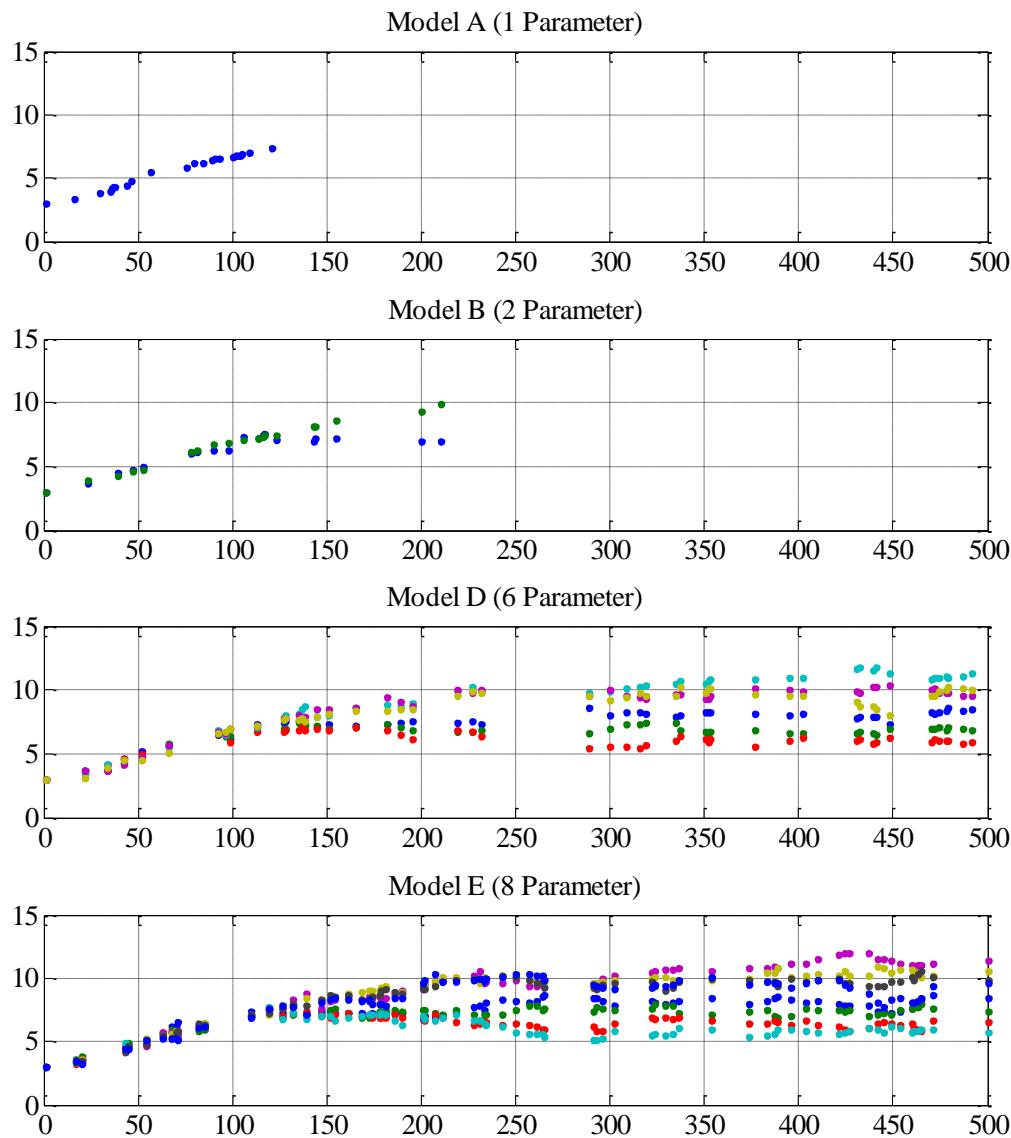


Figure 9-13 - Partial Burn-in for Scenario #1

After around 150 iterations, Model D and E became more prevalent and the parameters began to move apart. This was because Model A and B were no longer able to represent

the actual structure even remotely as well as Model D or E. However, given the random nature of the process, it is intuitively logical that these models with lower parameters burn-in individually faster. With fewer parameters, simpler models move toward a better solution faster and therefore they drive the burn-in of the RJMCMC process until they are no longer adequate (in a relative sense) to represent the experimental data. The observation can be particularly useful in a scenario involving nested parameters, like the cantilever, where before any individual parameters can be fine-tuned the modeler must be certain that the family of parameters is in the correct “ballpark.” This is analogous to the comparison of manual calibration to optimization and parameter identification presented in Chapter 6.4.

9.4.3.4 Justification for Multiple Model Methods

Contrary to what one might expect intuitively, the RJMCMC process for Scenario #2 did not heavily favor Model E, despite the fact that this model is capable of replicating the exact experimental setup. The main reason for this was the nature of the experimental data. Mode shapes and frequencies are global, system level behaviors, and while Model E is capable of representing the experiment better than Model D, it is also true that Model D can represent the experiment adequately and in many cases, just as well as a given parameter set in Model E. This type of behavior is precisely the justification for multiple model methods. Without any further information, it is very difficult to distinguish between Model D and Model E as a predictor for tip displacement.

The argument for multiple models as opposed to single models has already been made, and is only presented here as example. Given the population of accepted models, now the predictions of responses are investigated.

9.4.3.5 Response Predictions

Figure 9-14 shows cumulative distribution of displacement predictions for Scenario #2 and the actual “experimental” value. The response was contained within the distribution, but the prediction is not located exactly at the mean of the distribution. The discrepancy here can be attributed to the nature of the experimental responses, which are affected both by stiffness and mass changes, as compared to the unmeasureable response, displacement, which is only affected by changes in stiffness.

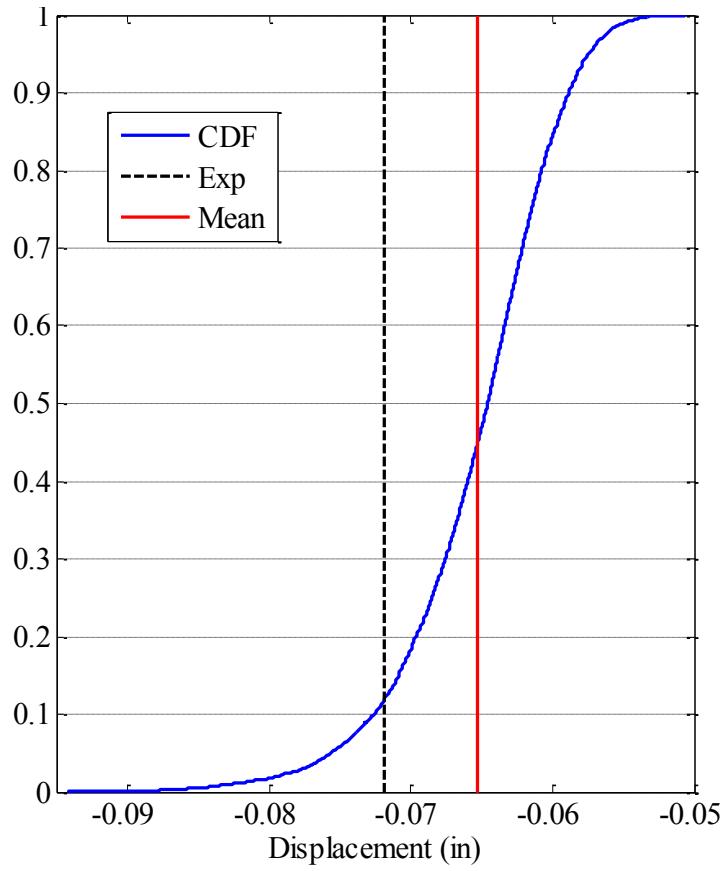


Figure 9-14 - CDF of Tip Displacement - Scenario #2

Figure 9-15 shows the CDF of the strains at the two locations specified in Figure 9-4.

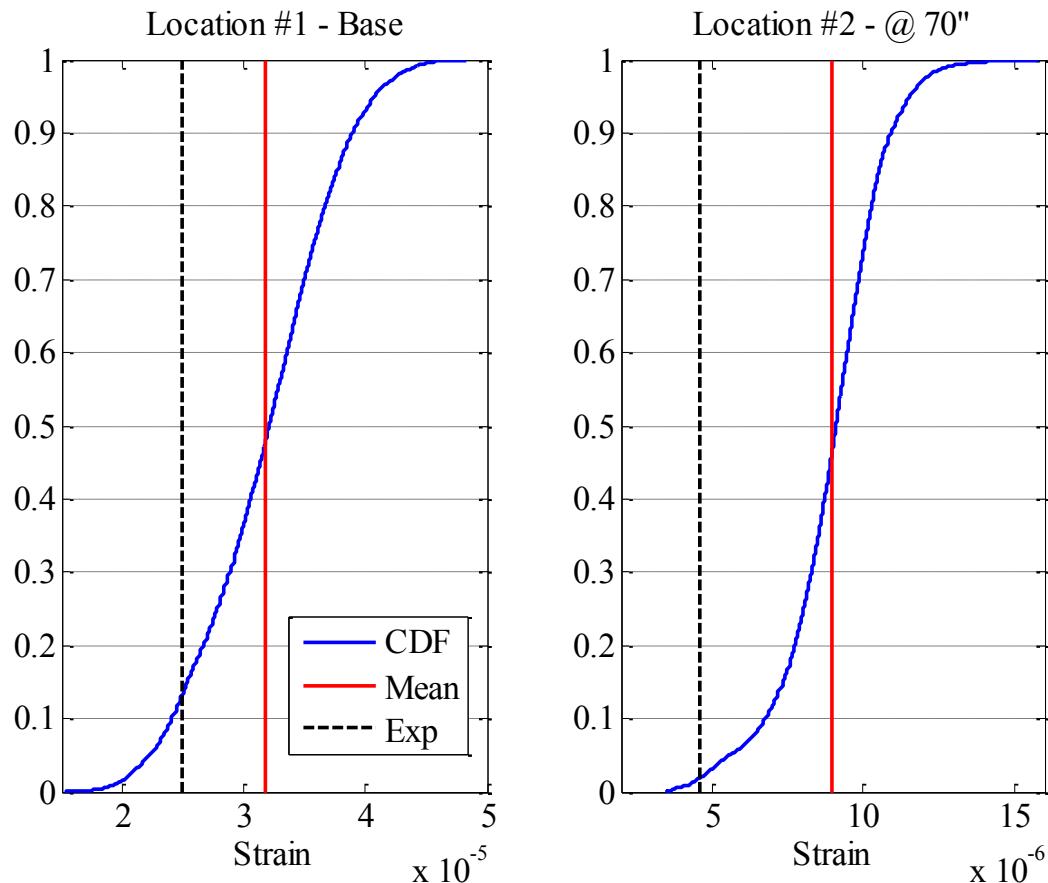


Figure 9-15 - Strain CDF Plots

The strain responses are directly sensitive to the height of the section where the response is extracted, unlike modal properties or displacements which are more global responses. This sensitivity is evident in the strain predictions. The actual responses are included in

the distribution, but the analysis over-predicts the response by as much as 51% in the case of Location #2.

Though it is somewhat trivial to investigate, the cumulative distributions of the frequencies are shown in Figure 9-16. The first frequency is estimated fairly well, but at the higher frequencies the discrepancies become larger.

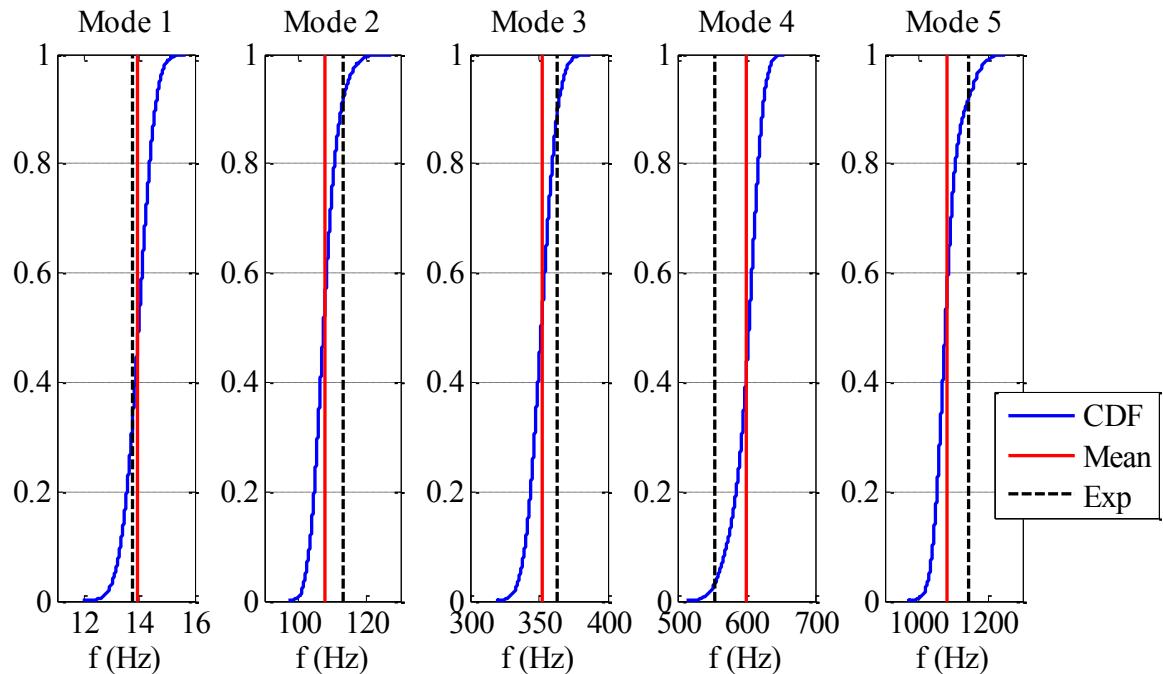


Figure 9-16 - CDF Plots of Frequencies

Since all of these pieces of information as well as the mode shapes are included in the calculation of the likelihood of a given model, they can balance one another some extent. Considering that the frequencies are all equally weighted in the likelihood calculation, it is not surprising that it appears as though the 4th and 5th frequencies are effectively

competing. The 4th frequency is on the lower end of its distribution and the 5th frequency is on the higher end of its distribution.

Clearly the frequencies were not identified perfectly, even though the sampling process was directly informed by these values. This is due to the inherent standard deviation applied to the experimental data that is incorporated into the likelihood calculation. A value of 4% was assumed and for this scenario, it was apparently too large. However, as this value is decreased, the RJMCMC process would eventually select only models with the exact correct parameter set. As this case was designed to show the scenario where there is not an exact representation in the candidate model population, it does an excellent job of highlighting the importance of the inherent standard deviation.

In general, Scenario #2 provided reasonable predictions of unmeasureable responses and matched the experimental data within reason. The mean of the displacement prediction was within 9% of the experimental value. The mean strains were within 28% at Location #1 and 51% at Location #2. The larger errors in the strains are directly a function of the local nature of strain responses mentioned previously. This reinforces the conclusion from Section 6.5.5 that using global responses, like mode shapes, to inform a model updating process where local response predictions are the goal is not necessarily a reliable course of action.

The algorithm predominately sampled two of the four model forms, which had similar likelihood values, or in other words, were equally acceptable in terms of predicting the experimental results.

9.4.4 Implementation of Alternative Proposal Functions

The resetting phenomenon was exacerbated by the selection of a small proposal function. This highlights the importance of the selection of the proposal function, which is not a trivial task. The best way to select an appropriate proposal function for each parameter in a model is through exploratory MCMC runs, which are time consuming.

To investigate the resetting behavior, better proposal functions for each parameter and model form were extracted from the individual MCMC analyses and applied to the RJMCMC process. The resetting behavior appears to be mitigated, meaning that after a transition, the algorithm can take a larger step and better explore the model space. The problem was not entirely eliminated, however. The resetting does not manifest as obviously as with the small proposal function, but it can be observed through the comparison of intra-model parameter values versus inter-model. Figure 9-17 shows the parameter chains from Model E which was comprised of eight parameters. The chains are split into inter-model (red points) moves and intra-model (black points) moves. Specifically in the case of parameters 2, 4 and 6, the inter-model moves tend to sample a smaller range than the intra-model moves. The effect of this limitation is very small, particularly compared with the smaller proposal function previously presented, but it was present nonetheless. After a move between models, there would be a short period, similar to burn-in, that must occur within the model form to begin to sample the full range again. It should be noted that for most applications of RJMCMC in the literature (particularly outside of structural engineering), this is not really considered a drawback of the procedure. Additionally, the effects would be mitigated with more strict prior probabilities. It is only for the application to civil structures, where the parameters can

affect both observations and unmeasureable response predictions, that the ability of the method to sample the full range of the parameter is important.

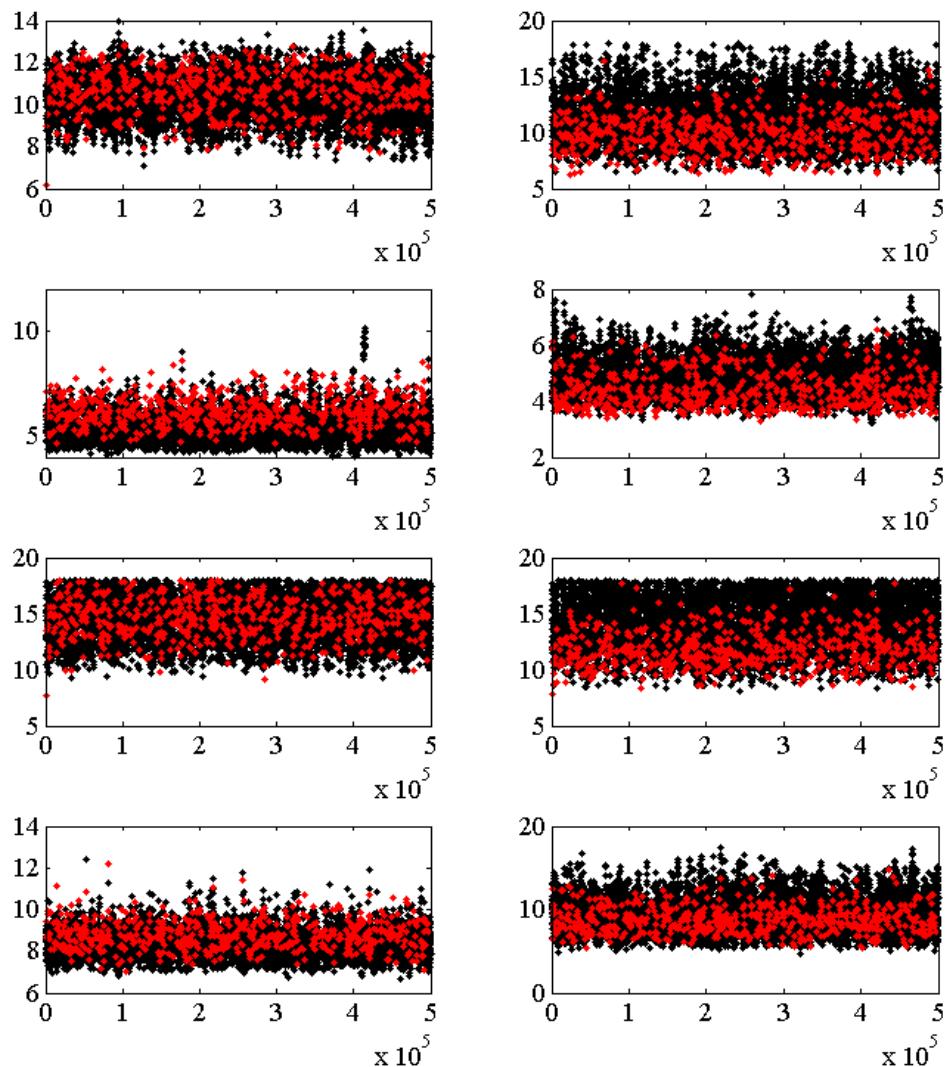


Figure 9-17 - RJMCMC Comparison of Inter and Intra-model Moves

Irrespective of the resetting issue, the acceptance rate fell to around 2% to 5% depending on the proposal functions selected. For a trivial model like the cantilever beam, this type of acceptance rate can be managed through an extremely large number of simulations. For a more complex bridge model with a substantial execution time, an acceptance rate this low would make RJMCMC time prohibitive, particularly considering that the included proposal functions were extracted from full MCMC analyses which would also be time consuming.

9.4.5 Discussion of RJMCMC for the Cantilever

The RJMCMC study conducted on the simple cantilever was a preliminary investigation designed to assess this process as a tool for application to large, complex constructed systems. The jump equations serve to maintain a true, single Markov Chain throughout the entire process, and therefore can directly inform model selection and response prediction which are both valuable to structural engineering. The development of the jump equations for the cantilever was fairly simple, but would be exponentially more difficult for a real structure where nested parameters are not the standard, and the difference between model forms is much more drastic than this simple case.

Considering the initial selected proposal function, the conditioning of parameters by the least-complex model provides some information about the ability of the model forms to represent the data. One could draw the conclusion that the acceptance of Model D and the subsequent conditioning of Model E indicate that Model D is the best solution given the data. However the higher parameter models never really burn-in and explore the

parameter space, so to conclude that it simply cannot represent the data is not necessarily true and potentially detrimental to the reliability of the response predictions.

When more appropriate proposal functions were used, the acceptance rate was very low, indicating that an application with a more complex model and a substantially longer execution time would be time prohibitive.

9.5 Application of MC3 methods to the Cantilever Problem

In order to mitigate the difficulties observed for the cantilever beam study using RJMCMC, an alternative method of sampling parameters across model forms was desired. The general methodology, termed MC3, was outlined in Section 8.5.

An MCMC sampling method and both uniform and BIC prior model weights were implemented within a two-step framework where each model form has a fully converged MCMC chain that is resampled to combine model forms. Monte Carlo sampling was investigated as a base line despite the fact that MCMC sampling is considered much more efficient.

The ability to implement MC3 in real-time was also investigated. Real-time refers to whether or not the individual chains are developed as the procedure moves between models or if the full MCMC chain for each model is developed in advance and resampled. A description of each method and the application of the method to the cantilever beam problem are provided in the following section. Evaluation was based on computational efficiency and viability of response predictions.

Similar to the RJMCMC study presented in Section 9.4, a scenario where the actual model is included in the population, and a scenario without the actual model in the population were both analyzed.

9.5.1 Method #1 – MC3 Resampling of Individual Converged MCMC Chains

This method required the individual generation of MCMC chains for each model form. The individual chain was run until converged, as indicated by the stability of the mean and standard deviation of the parameters. The entire chain (accepted and rejected samples) excluding the burn-in period was kept and then resampled using the Metropolis-Hastings sampler with either uniform or BIC prior model weights. Recall that the acceptance criteria includes both the prior probability associated with the model, and the likelihood of that model and selected parameter set considering the experimental observations. The parameters all had uniform prior probabilities and cancel out.

9.5.1.1 Uniform Prior Model Weights

For uniform weighing, the length of the secondary MCMC resampling analysis was dictated by the length of the a priori chain directly because the accepted and rejected models are included. This means the secondary analysis will be limited by the length of the shortest prior chain. The number of models required was approximately equal to the total number of secondary iterations divided by the number of model forms.

9.5.1.2 BIC Prior Model Weights

For BIC weighing, the length of the individual chains was informed by the model weights. Again the total secondary analysis length was used, but now the requirement of the prior MCMC chain length is based on the BIC weight. There is no way to know in advance how many models of the secondary MCMC resampling will be accepted, meaning that the initial MCMC runs may have to be appended if convergence has not occurred.

9.5.1.3 Application of Method #1 to the Cantilever – Actual Model Included

The first application of this method to the cantilever was for the scenario where the actual model was included in the candidate model population. Again this is not a realistic scenario for constructed systems, but it is informative during the development of this process. Both weighing schemes were applied to MCMC resampling of a set of prior MCMC chains developed for each model form. Table 9-5 lists the weight, preliminary acceptance ratio, and total preliminary MCMC iterations required for both weight schemes considering a secondary MCMC analysis of 200,000 iterations. Note that the actual prior MCMC chains were much longer than required.

Table 9-5 - Prior MCMC Model Requirements

Form	Prior MCMC Acceptance	Uniform Weights	BIC Weights	Required # of Models (Uniform)	Required # of Models (BIC)
A	41.2%	.1667	.030	33,333	6,000
B	54.9%	.1667	.068	33,333	13,600
C	34.9%	.1667	.398	33,333	79,700
D	16.6%	.1667	.171	33,333	34,300
E	11.1%	.1667	.199	33,333	39,800
F	1.0%	.1667	.133	33,333	26,600
Total	-	1	1	200,000	200,000

Method #1 was applied first with uniform model weights. The analysis converged at approximately 5000 accepted models, which corresponded to a total number of iterations of 67,616. Of the 5000 accepted models, 3978 were in Model C, the actual model. Again, since both the acceptance rate of the secondary MCMC analysis and the number of models required for convergence are unknown, it was not possible to predict how many models were required in advance. Note that after the initial MCMC chains are developed, the secondary assessment is trivial in terms of computational time since the finite element model does not need to be re-executed. Figure 9-18 shows the CDF of tip displacement as well the experimental value, considered unknown in this case. The mean of the chain was almost exactly the same as the prediction.

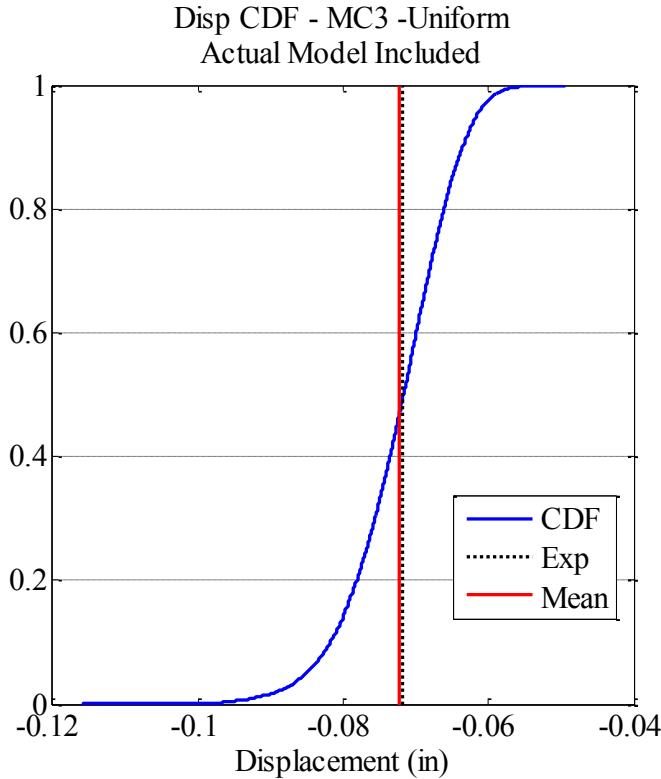


Figure 9-18 - Displacement CDF - MC3 with Uniform Weights - Actual Model Included

Similarly, the mean values of the strains were also very close to the unmeasured experimental values, as seen in Figure 9-19. The response at the base was predicted as accurately as the displacement. A slight discrepancy is apparent in the strain at 70", but the magnitude of this response was much smaller as well. The distribution for strain at 70" was skewed because the sampling was occurring near a parameter bound which imposes a bound on the strain response.

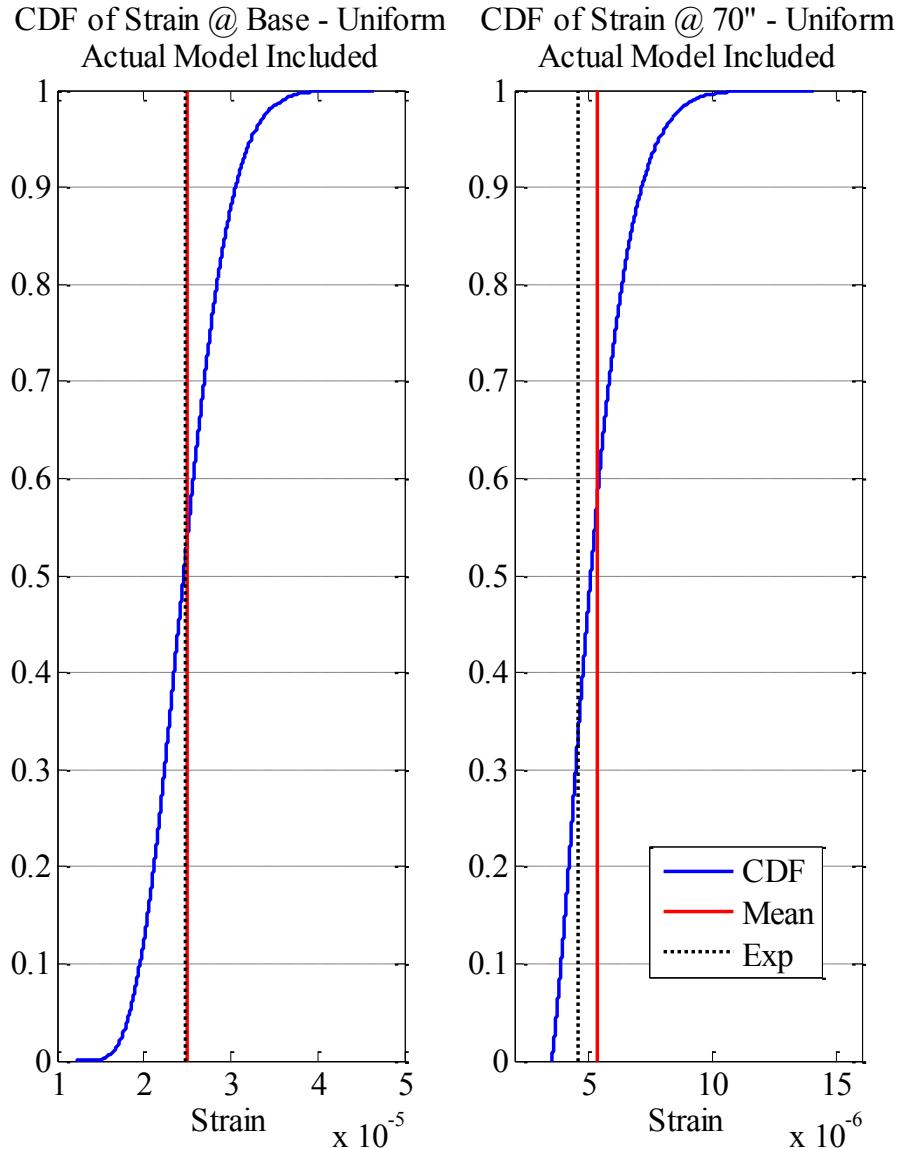


Figure 9-19 - Strain CDF - MC3 with Uniform Weights – Actual Model Included

Though it is difficult to quantify, there is some information regarding the ability of the MC3 procedure to inform the responses at the particular locations relative to one another. The section height parameter at the base was very influential on the modal properties, particularly on the stiffness of the structure. The parameter which directly informs strain

at 70" from the base was not as influential in terms of stiffness, but was likely more influential in terms of mass. However, there was less certainty here, and subsequently, the response prediction was not as accurate.

Method #1 was also applied using BIC weights. Again the resampling analysis was 200,000 total iterations. Convergence occurred at approximately 5000 accepted models which corresponded to 26,566 total iterations. In this case, 4558 of the 5000 accepted were in Model C. Figure 9-20 shows the CDF of tip displacement. The mean response and the unmeasured observation are exactly the same and lie in the middle of the distribution.

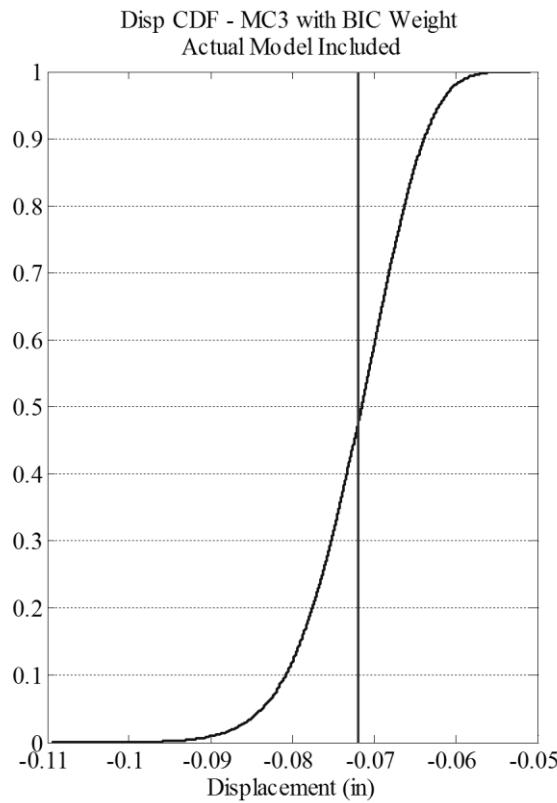


Figure 9-20 - Displacement CDF - MC3 with BIC Weights

Figure 9-21 shows the CDF plots of the strain at the base and at 70" from the base.

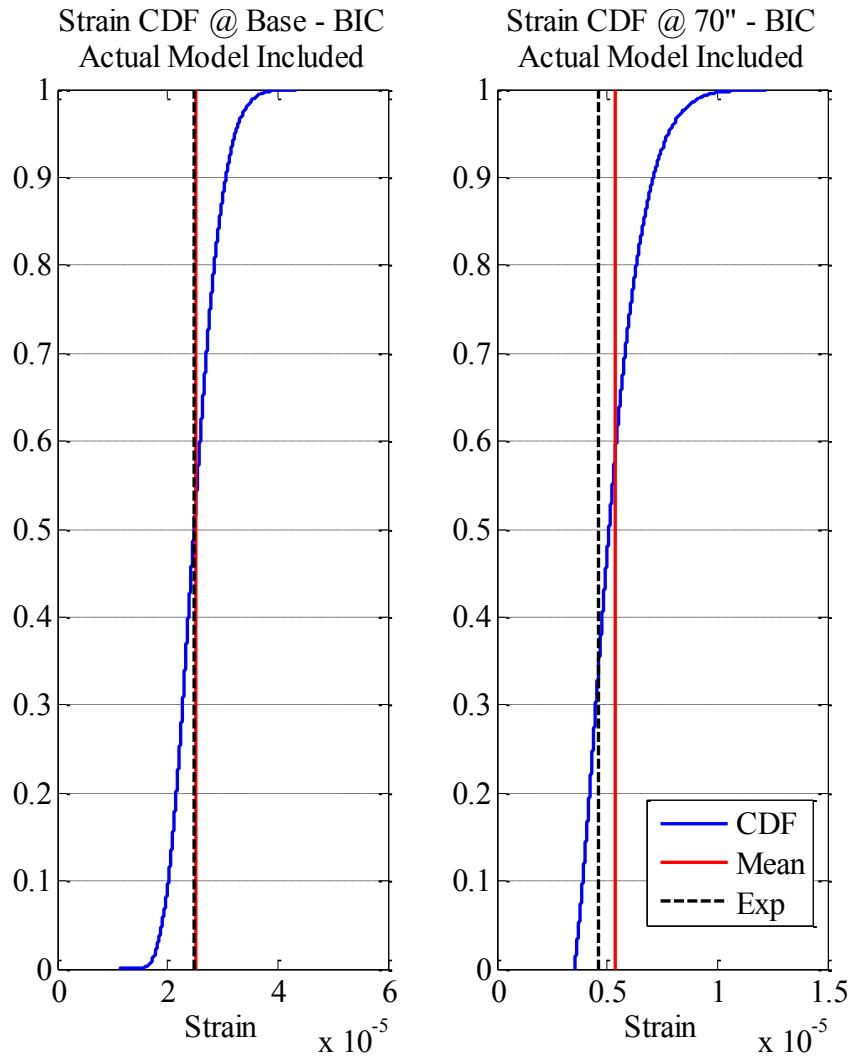


Figure 9-21 - Strain CDF - MC3 with BIC Weight

Since the experimental scenario was explicitly included in the analysis, it is intuitively logical that the both analyses would converge when the “actual” model (Model C) reached a certain number of accepted instances. This phenomenon will not continue when real constructed systems and experimental observations are being modeled. However, it does serve to indicate that when a model which is substantially better at representing the experimental observations is included in the population, uniform weighting is inefficient.

Figure 9-22 shows the CDF plot for both weight schemes. The responses are nearly identical. Method #1 is clearly able to accurately predict the tip displacement response independent of how the model forms are weighed. However, BIC weights appear to be much more efficient based on the number of total iterations until convergence, with slightly less variability.

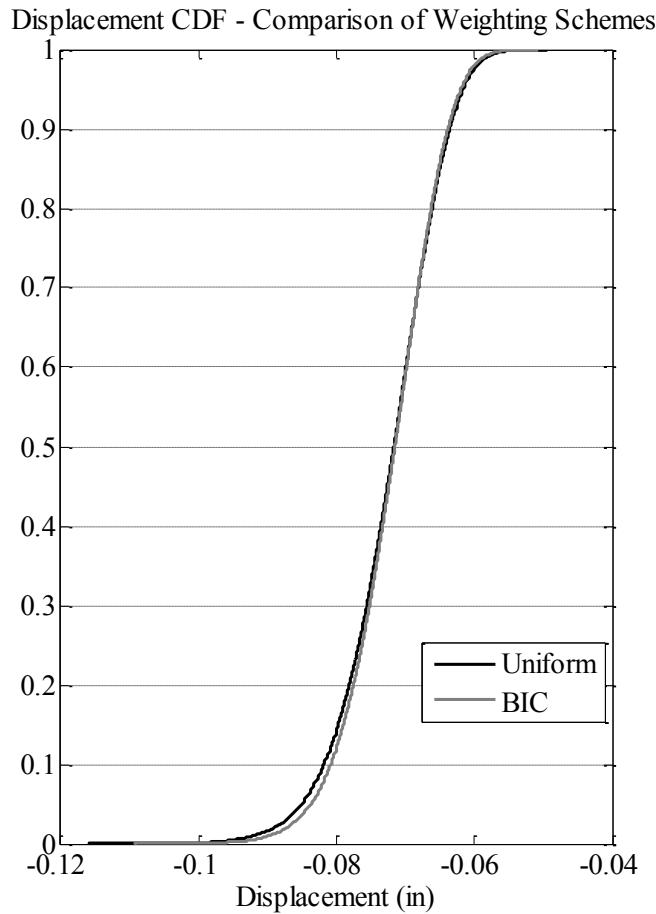


Figure 9-22 - Method #1 - Comparison of Different Weighing Schemes

9.5.1.4 Application of Method #1 to the Cantilever – Actual Model Not Included

The more realistic scenario of the candidate population which does not contain the actual model is presented here. Again a total of 200,000 iterations between Models A, B, D, E and F were run with both uniform and BIC weight. In general, much lower acceptance rates were observed, which was expected. Comparisons highlight the effects of including versus excluding the actual model are presented in Section 9.6.1.1.

9.5.1.5 Uniform Weights

Figure 9-24 shows the CDF of tip displacement for Method #1 with uniform weights when the actual model is excluded from the population.

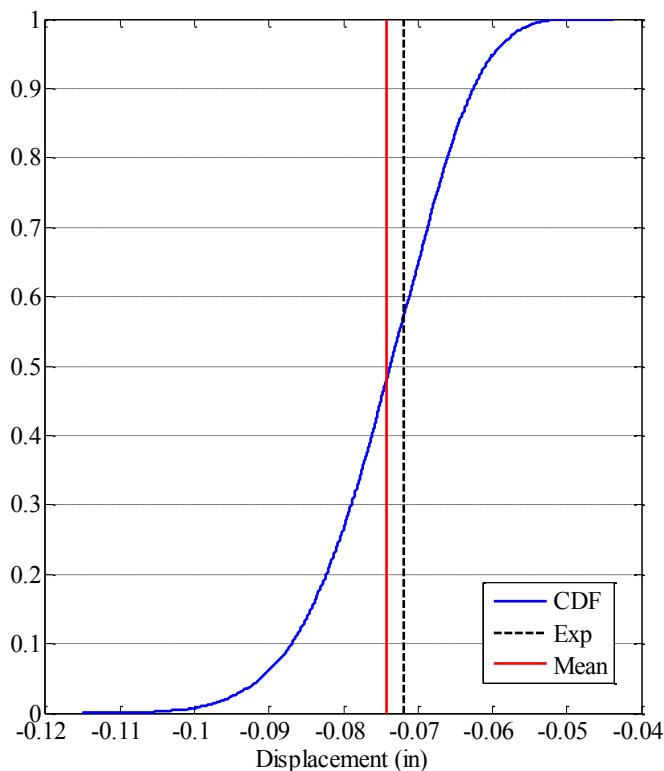


Figure 9-23 - Displacement CDF - MC3 with Uniform Weight - Actual Model Excluded

This scenario still gave a reliable prediction, with the experimental value falling in the middle of the distribution. However, the mean of the predictions was no longer exactly the same as the actual unmeasured value. Similarly, in Figure 9-24 the strain predictions showed slightly more variability.

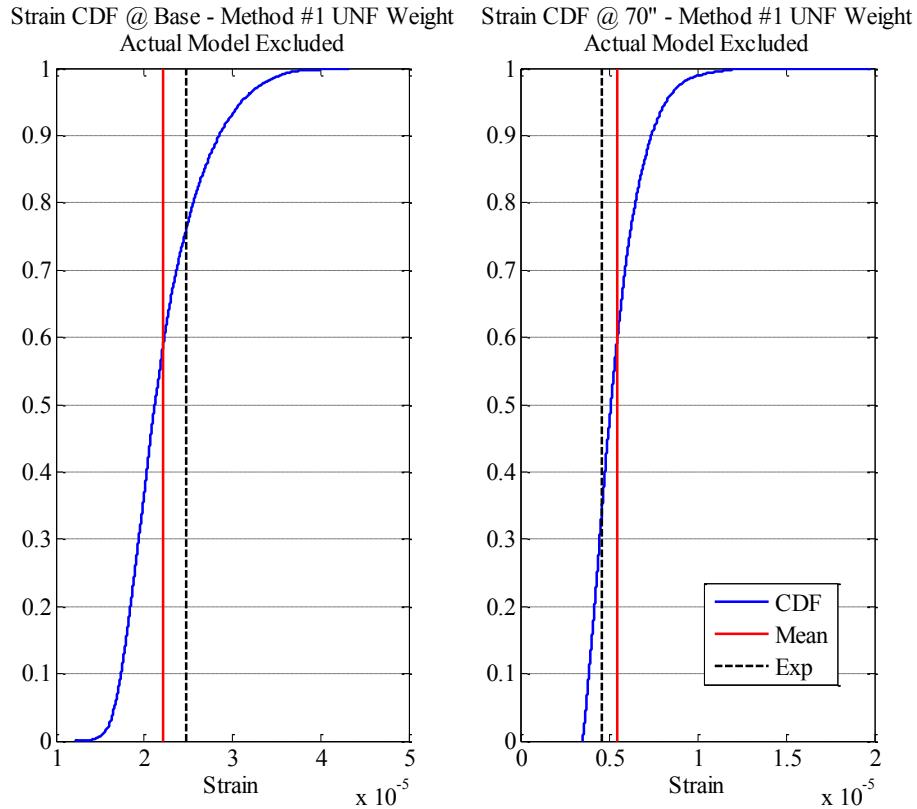


Figure 9-24 - Strain CDF - MC3 with Uniform Weight - Actual Model Excluded

9.5.1.6 BIC Weights

Figure 9-25 shows the CDF of tip displacement for the MC3 case with BIC weighing where the actual model is excluded from the model population. Similar to what was observed with the uniform weights, the mean prediction was not exactly the same as the unmeasured experimental value.

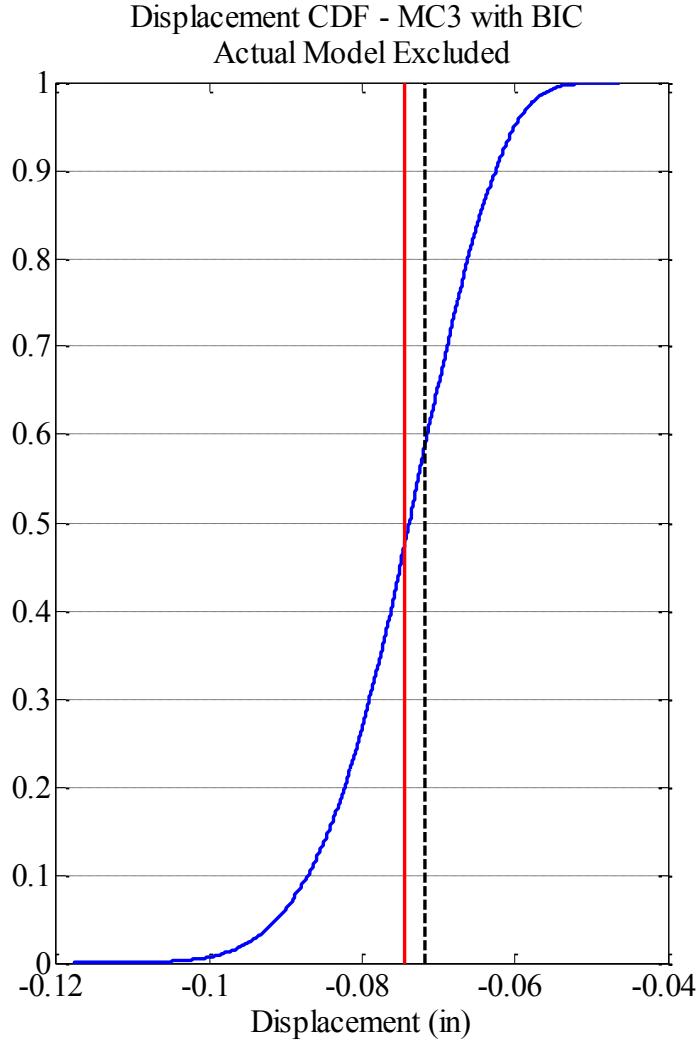


Figure 9-25 - Displacement CDF - MC3 with BIC - Actual Model Excluded

Similar results are observed in the strain CDF plots shown in Figure 9-26. The strain distributions were both heavily skewed. In the case of the strain at the base, the experimental value falls towards the tail of the distribution, but was still well within the distribution.

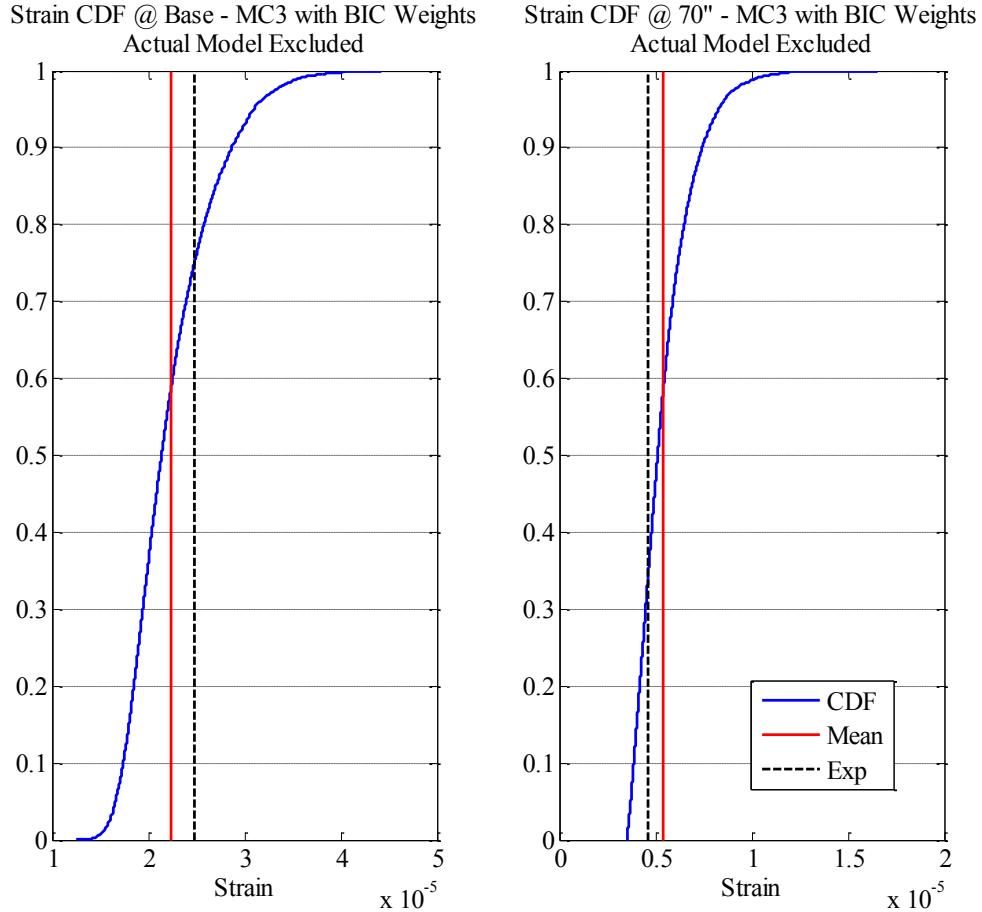


Figure 9-26 - Strain CDF - MC3 with BIC Weights - Actual Model Excluded

9.5.2 Method #2 – Real-time MC3

MC3 analysis is essentially identical to Method #1 except that the individual chains are maintained in real-time, not run in advance. This provides a massive improvement in computational efficiency, as no models outside of the burn-in period that are not either explicitly accepted or rejected in the analysis need to be run. This negates the need for any estimation of model iteration requirements. However, in order for any inter-model MCMC chain comparisons to be valid, the individual MCMC chains must be properly

burned in. This is analogous to removing the burn-in from the individual chains in Method #1. Again convergence of the burn-in is based on stabilization of the mean and standard deviation of the parameters, and stabilization of the proposal function if an adaptive method is used.

Using BIC weighting can be aggressive in the acceptance or rejection of model forms and requires a reliable estimate of MLE to inform the BIC calculation. This means that at a minimum, the optimization procedure used to estimate BIC should be very robust with numerous starting points and iterations. However, it is entirely likely that during the burn-in period, a sample will be drawn that exceeds the MLE from the optimization. If this occurs, then this new sample should be used in the calculation of BIC. This is not possible when real-time sampling is used, as the prior weight is set beforehand.

Conversely, the dependence on reliable a priori MLE values can be mitigated by using uniform weights for each model. This comes at the cost of computational efficiency, as models with relatively low likelihoods and thus little chance of acceptance will be sampled more frequently.

9.5.2.1 Application of Method #2 to the Cantilever – Actual Model Included

The example presented here is a proof-of-concept for the application of MC3 in real-time. It was previously noted that the Method #1 and #2 were essentially identical after convergence, and that BIC weighing tends to be more efficient than uniform. Table 9-6 shows the total number of iterations run to burn in, the number of attempts, and the

number of accepted models in each model form for a set analysis length of 200,000 models.

Table 9-6 - Model Attempt and Acceptance Counts – Real-time MC3

Model Form	Burn-in Required		Attempted Models		Accepted Models	
	Uniform	BIC	Uniform	BIC	Uniform	BIC
A	200	200	33,400	5,911	0	0
B	300	300	33,273	13,257	0	0
C*	700	700	33,186	79,951	11,523	27,630
D	1,100	1,100	33,485	34,348	1,109	1,218
E	1,200	1,200	33,295	39,828	2,126	2,419
F	1,200	1,200	33,367	26,441	153	148
Total	4,700	4,700	200,000	200,000	14,911	31415

* Base Case model takes this form

For this preliminary study, specific convergence criteria were not included. Rather the analysis was run for a sufficiently long amount of time to ensure that the response predictions had stabilized. The results were then analyzed to determine when convergence would have occurred.

Convergence of the response predictions, in this case tip displacement, served as the stopping criterion. The models were started in a high likelihood region and were burned in individually; meaning the mean and standard deviation of the response predictions tended to converge quickly. For example, a bound of $\pm 2.5\%$ of the current value was arbitrarily used a check point for validation of convergence of the standard deviation.

Using this bound, both analyses converged in approximately 2000 accepted models.

These correspond to 27,109 and 12,400 total models, respectively. Note that of these total required models for convergence in each weighing method, Model C accounted for 4587 and 5037 models, respectively. Figure 9-27 shows the comparison of the displacement CDF for resampled MC3 versus real-time MC3. The results are nearly identical, highlighting the potential benefits of eventually implementing this process in a real-time manner. However, in this development phase, the resampling method will be used as this allows for interpretation and error-screening at many levels.

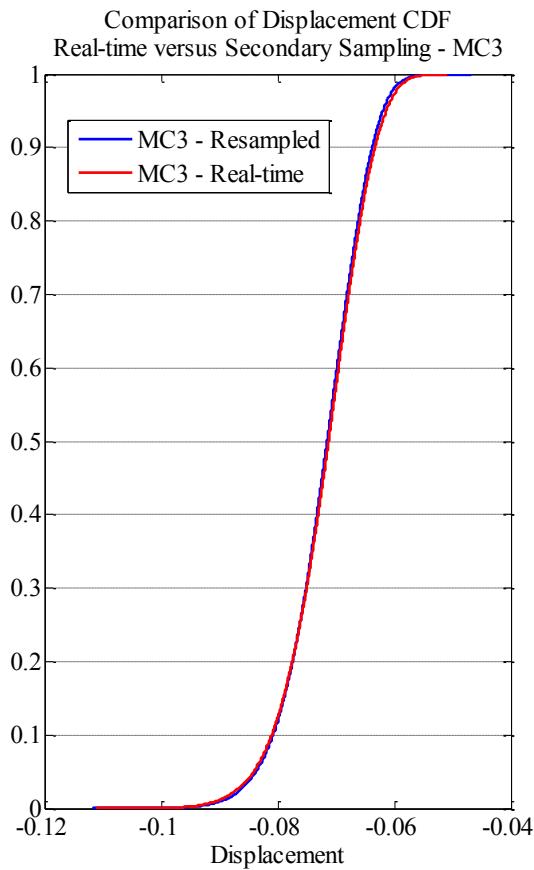


Figure 9-27 - Comparison of Real-time and Resampled MC3

9.6 Assessment and Validation of Multi-Dimensional Bayesian Analyses

Each of the methods described in the previous section were applied to the cantilever problem presented in Section 9.2. The methods were evaluated considering computational expense and robustness of predictions. The response predictions of interest were tip deflection and strain response at two locations under a 100 lb. point load.

9.6.1 Verification of Negative Effects of Resetting Behavior in RJMCMC

A comparison of the accepted instances of Model E in both the MC3 and RJMCMC was conducted to verify the importance of a suitable proposal function to negate the resetting behavior. In particular, the likelihoods of the accepted samples from Model E in the MC3 analysis and the RJMCMC analysis with both proposal functions are plotted on the same axes in Figure 9-28. This figure clearly highlights how Model E in RJMCMC with the small proposal function never consistently explored the area of high likelihood located in the MC3 example. The RJMCMC analysis with the larger proposal function does explore the full area, but at the cost of an enormous number of iterations. To achieve 2000 accepted models in that particular model form, over 136,000 iterations were required, and Model E was the most frequently sampled form.

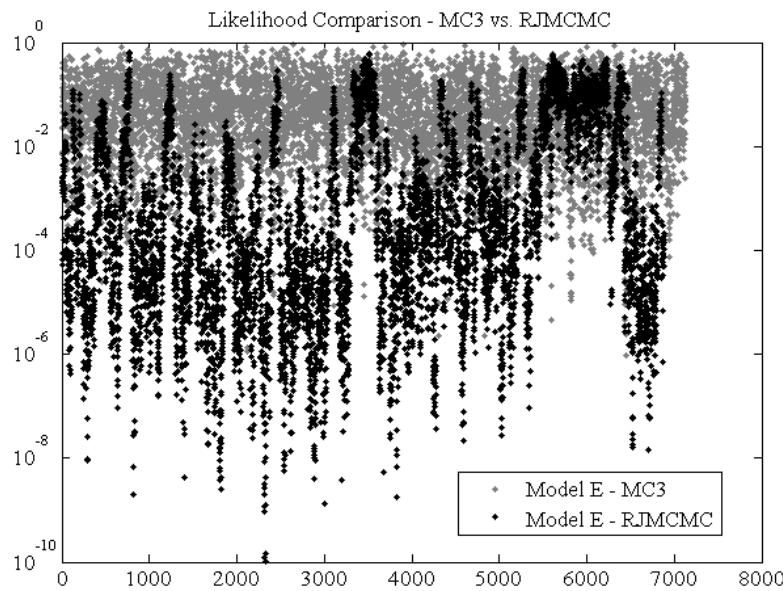


Figure 9-28 - Comparison of Likelihoods for Model E – RJMCMC (first proposal functions) vs. MC3

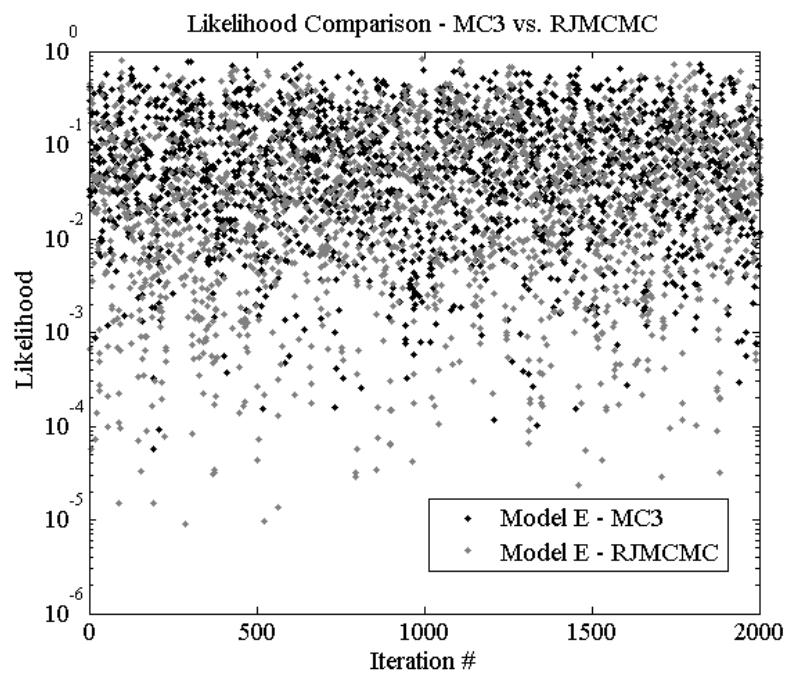


Figure 9-29 - Likelihood Comparison - RJMCMC (updated proposal function) vs. MC3

9.6.1.1 Inclusion of the Actual Model

Figure 9-30 shows the displacement CDF for the MC3 method with the actual model included, and with it excluded. This figure highlights the increase in variability that results from removing the actual model from the candidate population. In this case, the change in the mean values was minor, but distribution is wider. For constructed systems, this increased variability will always be included in the response predictions and there will not be any frame of reference.

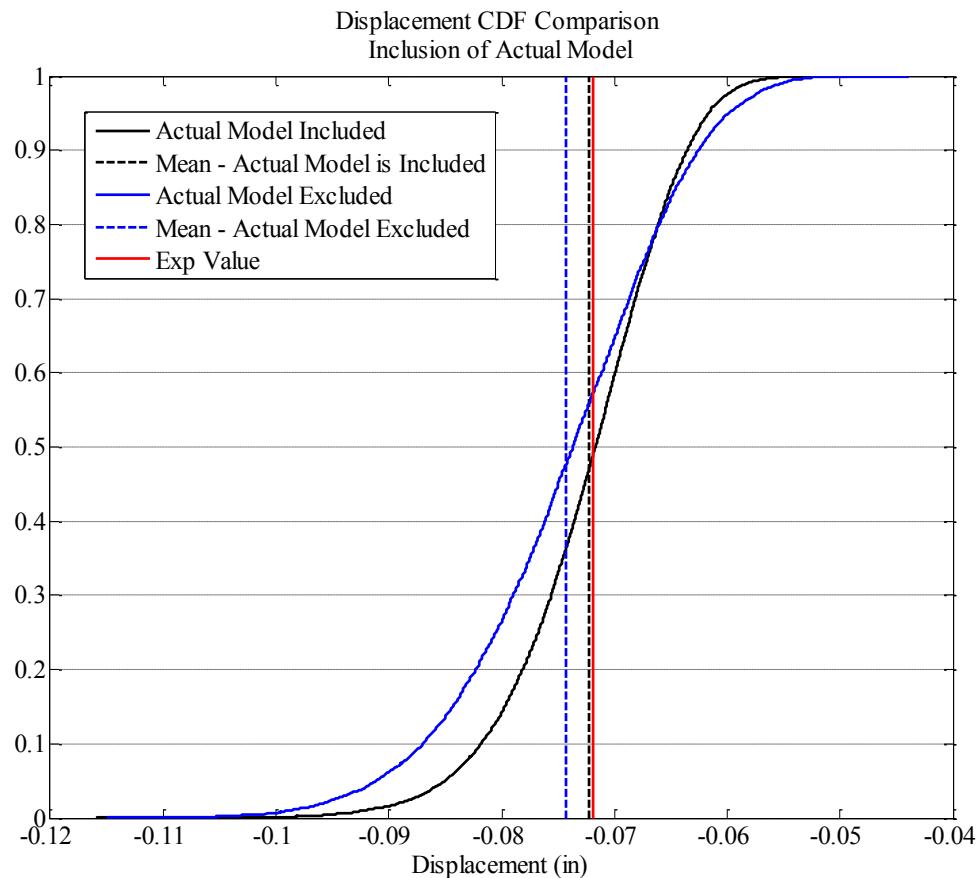


Figure 9-30 - Comparison of Displacement CDF - Inclusion of Actual Model

9.6.1.2 Real-time versus Resampling

It has been shown that resampling and real-time MC3 provide essentially the same final results. There can be a substantial difference in the number of total model executions required to achieve this result, but this is predicated on the assumption that the engineer has an accurate method of stopping the real-time analysis. Convergence of response predictions as a stopping criterion for real-time MC3 analyses was shown to be somewhat subjective and arbitrary. This uncertainty lessens the value of the reduction in computational expense provided by real-time MC3.

From an analysis standpoint, resampling has the added benefit of leaving the engineer the option of directly interpreting the individual chains. This not only provides better MLE values for calculation of BIC-based weights, but it also can provide insight into the structure via the value and the variability of the parameters.

Considering this it makes sense that as this method is being developed and initially applied to constructed systems, the full chains for each model should be run and then resampled.

9.6.1.3 BIC versus Uniform Weighing

As expected, BIC weighing was more efficient than uniform weighing for this situation. This is a trivial conclusion and is only included herein because improvement in efficiency will not be standard across all problems. It is by definition a function of the model population, meaning that any missing model building blocks or grossly incorrect MLE estimates would skew the results. Therefore, when prior knowledge is limited, uniform weighing provides a safe, but less efficient option. Conversely, when the modeler is

confident that all model building blocks are explicitly accounted for, then BIC weighing can save substantial computational time.

9.6.2 Robustness of Predictions

The MC3 approach appears to provide reasonable response predictions for the cantilever study. For this example, the selection of the model population was fairly intuitive and straightforward. It can be confidently stated the model population was reasonable. Model A and Model B were missing critical parameter, which was apparent considering how infrequently, if at all, they were sampled and accepted. Models C, D, E and F were all capable of representing the observed data with reasonable likelihood values on approximately the same order of magnitude. For this reason, they were all sampled and the predictions then reflect the uncertainty of the distribution of the parameter as well as the uncertainty associated with predicting a response that was not explicitly measured.

9.6.3 Considerations for Application to Real, Constructed Systems

The analytical cantilever example provides a clear, approachable scenario through which it was possible to explore and understand multiple dimension Bayesian analysis methods. RJMCMC is promising as a model selection technique but the necessity of the jump equations, the difficulty in proposal function selection and the potentially long analysis times limit its versatility. MC3 methods, which sample from multiple independent Markov Chains but do not tie the parameters between the chains, provide insight into the effects of the model form, selected parameters, and the bounds. MC3 methods are

particularly applicable to constructed systems when there are so many modeling assumptions required, and so much uncertainty regarding parameters, mechanisms, and material properties.

Specifically, when applying any of these methods to a constructed system, one should consider the following:

- Prior Model Weight – BIC weights can be more efficient, but they also can mask substantial differences in magnitude of likelihood. As a tool for developing prior model weights, BIC should be used cautiously and alternatives should be considered.
- Real-time versus Resampling – There is an obvious benefit to real-time sampling across chains, as it greatly reduces the amount of model iterations required, but this requires sound heuristics and a priori knowledge. While this method is in development, resampling of parameters should be used.
- Parameter versus Model Form – This investigation only considered changes in parameter spatial distribution. However, it is entirely possible to approximate the uncertainty associated with a given model form, as compared to another. For example, the cantilever could have been modeled with beam elements, shell elements and solid elements and compared as well. This would theoretically work with both RJMCMC and all proposed MC3 methods, but the construction of the jump equations may be very difficult to develop if RJMCMC were implemented.

Chapter 10: Multiple Model Structural Identification of IBS Bridge

This chapter presents the application of Bayesian model updating to the IBS Bridge. The focus is on examining multiple models that have different spatial variations of parameters. To begin, nine model forms were considered and burned in. Based on their relative likelihoods three models were selected from the original nine and for these models full MCMC chains were developed. This chapter presents comparisons to the modal properties, which were selected as the observations used to inform the model updating.

10.1 Summary of IBS Bridge Model-Experiment Correlation

In Chapter 6 a robust model-experiment correlation for the IBS Bridge was conducted. This included manual calibration and parameter identification using multiple modalities of measured data, and resulted in several generations of the FE model form. For the purposes of this research, this effort is considered the state of the art as would be conducted by a typical practicing engineer or academic researcher. As discussed, there is still the problem of non-uniqueness of solutions. To address this issue, several multiple model approaches were applied to an analytical cantilever beam in Chapter 9. Though a simple structure, this study indicated the potential for multiple model methods to at least recognize and quantify the non-uniqueness problem. This chapter discusses how these methods were employed on the IBS Bridge to consider spatial variations of parameters for the element-level FE model.

Note that the goal of this effort was not to find an automated replacement for the structural engineer/modeler but rather to gain insight about the model forms, model

classes, parameters, and the actual structure. Throughout this chapter, numerous conclusions are presented that have perhaps been established in other fields of science or engineering, but are new to structural engineering. Other conclusions are simply made to attempt to quantify the extremely uncertain nature of model-experiment correlation.

10.2 Problem Definition

The investigation into spatial variation of several parameters using an element-level FE model employed several key components. These included the experimental data to which the models were compared, the unobservable responses, the model form (in this case defined as the model building blocks of interest and their spatial resolutions and distributions).

10.2.1 Experimental Data

Recall that the IBS experimental setup predominately included strains, displacements, and accelerations. This research attempted to quantify the effect of model form and response modality in the prediction of unmeasureable quantities of interest. Specifically, the scenario considered was observation of modal parameters in order to predict static displacements and strains. In Dubbs (2012), the inaccuracy of using one modality of response to inform a model and then predict another modality of response was highlighted. Again, in Section 6.4 of this document a similar conclusion was drawn using traditional, single model approaches. This is a fundamentally different point of view than traditional model updating which uses a single model and updates with the same type of

response than the model subsequently predicts. The reason this scenario is employed is that in reality, often it is not possible to measure the desired unknown on a structure. In this case, it was possible to close the loop by measuring strain and displacement, and then hold these responses in reserve as a validation data set.

The modal properties (frequencies and mode shapes) calculated from the ambient vibration testing conducted at the time of the original static test were used as the experimental observations. Using dynamic data was a reasonable assumption to make as it is often the simplest type of data to collect. There is no need for an external reference to the ground, as the sensor is referenced to gravity. The sensor is installed with magnets which are attached much faster than displacements and strain gages. Using dynamic properties for model calibration and estimation of flexibility are both very common in the research community and have the potential to be used routinely in practice in the future.

A total of seven frequencies and seven corresponding mode shapes were included in the analysis. The frequencies are shown in Table 5-6 and the mode shapes can be seen in Figure 5-42.

The unobservable responses were the displacements and strains, an identical scenario to what was used for the cantilever numerical study. The strains are calculated based on the member actions in the frame elements of the element-level FE model. This allows for the inclusion of the effects of any out-of-plane bending, which is important to consider because the strain gages installed on the top-side of the bottom flange were eccentric relative to the weak-axis centroid.

In Chapter 12, flexural load ratings are presented in the context of developing owner deliverables as a comparison to traditional, single-model St-Id.

10.2.2 Finalization of Building Blocks and Spatial Variations

Throughout the entire single model-experiment correlation process, several model building blocks or parameters were varied both manually and through parameter identification. A large subset of this entire group was included in these analyses, and was varied spatially. Recall that this set of parameters had to be sensitive to either the observations used to calculate the likelihood of the model, or to the prediction of the unmeasurable responses. These parameters did not necessarily have to inform both, however. The parameters included in this study were:

- Composite Action – For the manual calibration this parameter was limited to one of three values to attempt to reduce the size of the problem. In this case, a full range of values was used, but three specific areas were identified and varied either individually or together, depending on the model. These areas were identified during the manual calibration and represent one of many possible distributions of composite action properties.
- Axial Diaphragm Stiffness – The modulus of steel of the diaphragms was previously used as a model building block to consider the uncertainty associated with the transverse stiffness of the bridge. This parameter was still used, but in its inverted form (i.e., flexibility) as discussed in Section 10.2.4.

- Vertical Bearing Stiffness – In Chapter 6, the vertical bearing stiffness was included for the static calibration efforts and excluded for the calibration to dynamic data. This was because at the dynamic load levels the mechanisms which cause softness at the bearings would likely not be active. For multiple model investigations however, whatever parameters are believed to affect the desired response predictions should be included in the analysis. Instead of the parameters being excluded because they are not sensitive to the data with which the model is being compared, they were included because they were sensitive to the desired outputs. There are three regions for this parameter; the south end of girder #1, where the crack in the pier cap was located; the rest of the southern bearings; and all eight northern bearings.
- Longitudinal Bearing Stiffness – As discussed in Chapter 6, this parameter smeared together longitudinal restraint at the expansion end as well as rotational restraint, because the spring was offset from the neutral axes of the composite section. This was located only on the northern (expansion) side of the bridge.
- Transverse Bearing Stiffness – This parameter, like the vertical stiffness, was excluded from the manual calibration to dynamic data. However it was shown to have a significant effect on the predictions of static response and so it was included for the multiple model investigations. This parameter was only on the northern (expansion) side of the bridge.
- Modulus of the Parapet and Barrier – This was a smeared parameter meant to represent not only the uncertainty regarding the material properties of the barrier and the parapet, but also their connectivity to the deck.

There are some notable parameters that were not included in this study. The first and most obvious is the modulus of the deck. The reason for excluding this parameter is that composite action and diaphragm stiffness affect the model response (for the observations and predictions) in a very similar manner as the deck modulus would, which could have led to a situation where the sampling methodology would first pin one parameter to a bound before finding a region of high likelihood to sample for the other parameter. As a result, the composite action parameter can be considered to include the smeared effects of the deck modulus.

Modulus of the steel girders and the density of the concrete deck are both excluded because they are essentially deterministic. The deck density could have been included as an influence on the mass, but was not because it would tend to be a catch-all for other errors in the model, and in reality, is likely very close to the prescribed 150 pcf. It is recognized that some have intentionally included deterministic parameters in order to evaluate to what degree such parameters are distorted in the updating process. While this can be a powerful approach in practice, for this study the scenario allows a direct comparison with the actual measured results and so the additional dimensions to the updating problems were not needed. In addition, the issue of what prior probability distribution to use for such parameters is not an easy one. If a tight distribution is selected, then the sampling of the parameter is essentially constrained to the known area and it is of little value in diagnosis of the results. In contrast, if a broad prior probability is selected, then one could argue that ignoring known information can bias the results since an unrealistic covariance with other, truly uncertain parameters may appear.

10.2.3 Determination of the Candidate Model Population

Table 10-1 lists the final models used, including both spatial variation of parameter and model form. These designations are used throughout this chapter and the next.

Table 10-1 - Model Designation and Properties

Model Designation	Model Type	Number of Parameters
A	Element	6
B	Element	7
C	Element	8
D	Element	7
E	Element	8
F	Element	8
G	Element	9
H	Element	9
I	Element	10

As indicated above, the options ranged from six parameters in Model A to ten parameters in Model I. All cases with more than six parameters were simply different spatial variations of the base parameters in Model A. Figure 10-1 shows the three locations where the composite action parameters were delineated as discussed.

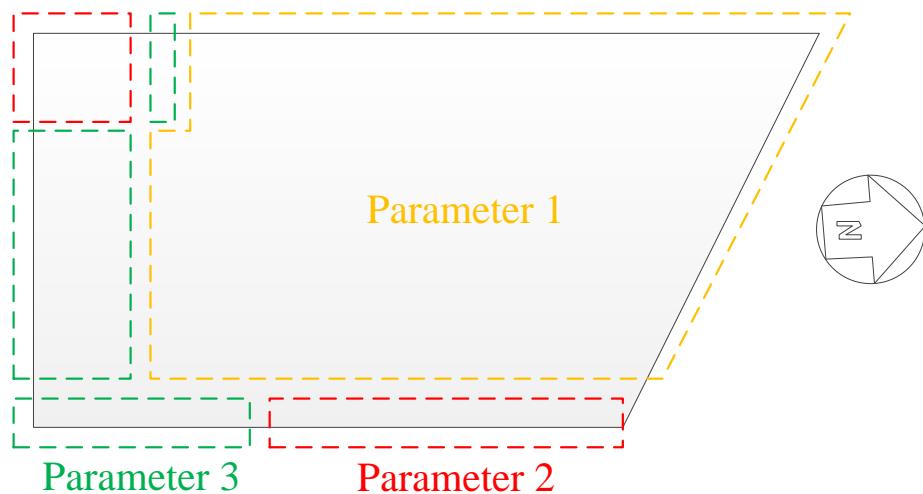


Figure 10-1 - Composite Action Parameter Discretization

This discretization was based on the results of the traditional, single model updating. It serves as a feasible representation of how a continuous parameter with substantial uncertainty in the real structure may be approximated for a model based on observation and knowledge gained during instrumentation of the bridge. This is not the “correct” distribution as this is not known, nor is it the only feasible distribution of this parameter.

Similarly, Figure 10-2 shows the location of the vertical stiffness parameters. Again the distribution of this parameter is an informed choice based on heuristics and observation.

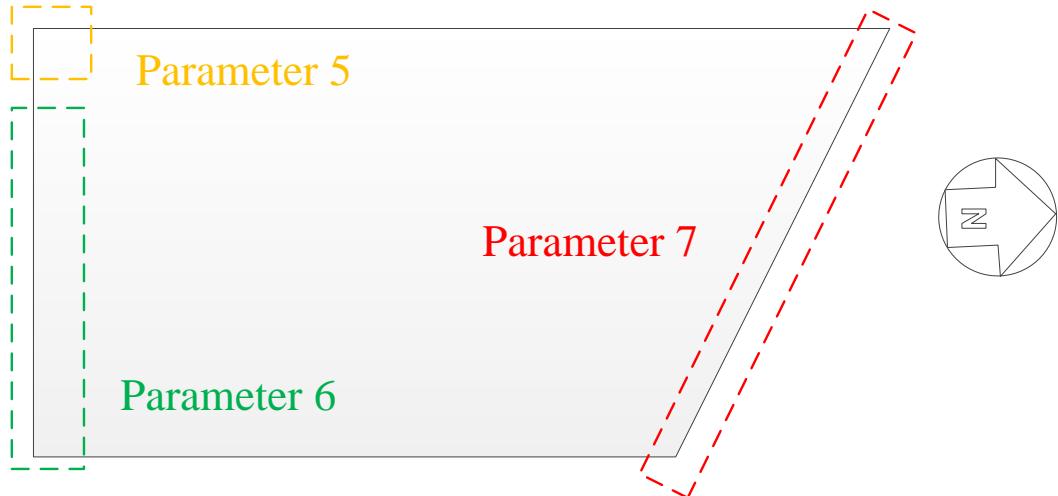


Figure 10-2 - Vertical Stiffness Parameters

Table 10-2 shows how the parameters are incorporated in each model form. The designation for each parameter in the table refers to that specific model form. Globally the parameters are referred to by the column headers 1-10.

Table 10-2 - Parameter Breakdown by Model

	Comp. Act #1	Comp. Act #2	Comp. Act #3	Axial Dia. Stiff	Vert. Stiff. G1-South	Vert. Stiff. Rest of South	Vert. Stiff. North	Long. Stiff. North	Lateral Stiff. North	Barrier and Parapet Modulus
	1	2	3	4	5	6	7	8	9	10
A	Para1			Para2	Para3			Para4	Para5	Para6
B	Para1	Para2	Para1	Para3	Para4			Para5	Para6	Para7
C	Para1	Para2	Para3	Para4	Para5			Para6	Para7	Para8
D	Para1			Para2	Para3		Para 4	Para5	Para6	Para7
E	Para1			Para2	Para3	Para4	Para5	Para6	Para7	Para8
F	Para1	Para2	Para1	Para3	Para4		Para5	Para6	Para7	Para8
G	Para1	Para2	Para1	Para3	Para4	Para5	Para6	Para7	Para8	Para9
H	Para1	Para2	Para3	Para4	Para5	Para5	Para6	Para7	Para8	Para9
I	Para1	Para2	Para3	Para4	Para5	Para6	Para7	Para8	Para9	Para10

Clearly the selection of these parameters and spatial distributions was heavily influenced by the information learned during the manual calibration and parameter identification presented in Section 6.3. Considering these multiple model methods are in the preliminary application phases for constructed systems, it was decided to use prior information and inferences from the single model calibration. There would have been substantial uncertainty regarding the model-experiment correlation which would have been indistinguishable from the uncertainty resulting from the method itself. In the future, when multiple model methods have been optimized and vetted, then evaluation from an uninformed viewpoint may be more appropriate.

10.2.4 Modification of the Element Level Model

The end rotations of the frame elements comprising the diaphragms were released in the element-level FE model. These elements were designed primarily as axial members, and the rotation release forces them to act axially. Considering this change, the axial stiffness became a smeared parameter representing partial rotational restraint at the gussets for the diaphragms.

The only other changes to the element-level FE model were minor improvements involving the connectivity between shells of the deck. Essentially this means slight adjustments of the mesh to “clean” up the model.

10.3 Development of Individual MCMC Chains

After selecting the model forms and parameters, the general procedure for development of a single model MCMC analysis was as follows:

1. Locate a reasonable starting point through optimization
2. Burn-in the model using an adaptive proposal function
3. Fix the proposal function and start the development of the Markov Chain
4. Stop the analysis considering the convergence of the mean and standard deviation of the parameters and the response predictions
5. Considering the likelihoods calculated in the chains, eliminate inadequate models from further consideration

10.3.1 Optimization for Initial Likelihood and Starting Point Estimation

Observations from the individual optimizations and MCMC burn-in of all the models are included in order to provide insight and guidance into the process of establishing a sound foundation upon which to apply such methods to constructed systems. Only a portion of the models listed in Table 10-1 are analyzed using full, individual MCMC runs, but every model was optimized and burned in before either being selected or excluded.

10.3.1.1 Selection of Parameter Bounds

For the numerical study presented in Chapter 9, the cantilever beam model had very clear physical bounds which made determining the initial starting point and proposal function very straightforward. The analysis was started using the most likely values from a number of optimizations all started from different points. Typically these optimizations all found approximately the same optimal solutions. This indicates that the objective function, in this case the likelihood of the model given the experimental data, was not a complex, peaked shape which would lead to frequent identification of local maxima.

For the multiple model investigation of the IBS element-level FE model, parameter bounds were defined based on the preliminary sensitivity studies similar to those conducted in Chapter 4. A simple objective function using the likelihood (Eqn. 8-6) calculated with modal properties was used, and the bounds over which the parameters appeared to influence this objective function were selected.

Note that simply running sensitivity studies of the effects of parameter changes on the selected observations is inadequate. The sensitivity study, and the subsequent selected

bounds, must also include the range of the parameter in which the desired unmeasureable responses are also sensitive. This implicitly includes any uncertainty with translation between the observations and the predictions.

Sensitivity studies represent the effect of changing a parameter considering all other parameters at fixed, arbitrary values. This creates uncertainty regarding the relevance of parameter bounds at a different set of arbitrary values. As such, these bounds were subsequently increased considering that the sampling algorithm would find the sensitive areas.

10.3.1.2 Interpretation of the Optimization Results to Establish Starting Points

In contrast to the simple numerical example, the optimization results for the IBS models were much less consistent. This is indicative of a much more complex likelihood space with numerous peaks. This type of behavior makes gradient-based optimization algorithms difficult to implement. To mitigate this to a degree, each model was optimized ten times with different starting points that were randomly drawn within the selected bounds of each parameter.

The optimization results with the largest likelihood were selected as the initial starting point for the burn-in of each model form. The proposal function was arbitrarily started at the same position for each parameter and each model form. This value was fairly small as compared with the parameter ranges and starting values in general. The logic for this was that since the burn-in period uses an adaptive proposal function, that proposal would adapt at least somewhat monotonically to a better value.

10.3.2 Burn-in the MCMC Analysis

Using the starting points and proposal functions determined via optimization, the burn-in for each model was run. An adaptation interval of 500 iterations was utilized. In this case an iteration considers acceptance or rejection of a model via the acceptance criteria (Eqn. 8-3), or a parameter bound exceedance.

Each parameter is assigned a uniform prior probability to reflect the lack of prior knowledge regarding the parameters.

The burn-in was typically run in a piecewise fashion. Convergence was determined based on a qualitative assessment of the resulting parameter chains, the mean and standard deviations of these parameters, and the likelihoods of the accepted models. Note that in all cases, parameter sets with likelihoods that were higher than the optimized values were sampled.

10.3.2.1 Effects of Parameter Bounds on Burn-in Analysis

It was observed that as the adaptation of the proposal function continued, the mean and standard deviation of certain parameters did not seem to converge after the rest of the parameters had converged. The mean and standard deviation of these parameters tended to increase monotonically, while the proposal function oscillated around a typically very large value as it updated. Initially it was believed that the proposal function selected was inadequate and that the analysis had not truly burned in. However, after allowing the analysis to run for a sufficiently long time, the mean and standard deviation of every parameter stabilized. It was observed that the mean of the parameters with the unstable

proposal function stabilized at a value equivalent to half of the parameter range, and the sampling occurred over the entire range.

Figure 10-3 shows the chains of the accepted parameter values for Model I during burn-in. Note that parameter 1, 5, and 6 seem to continue to rise and expand until essentially the entire range is being sampled.

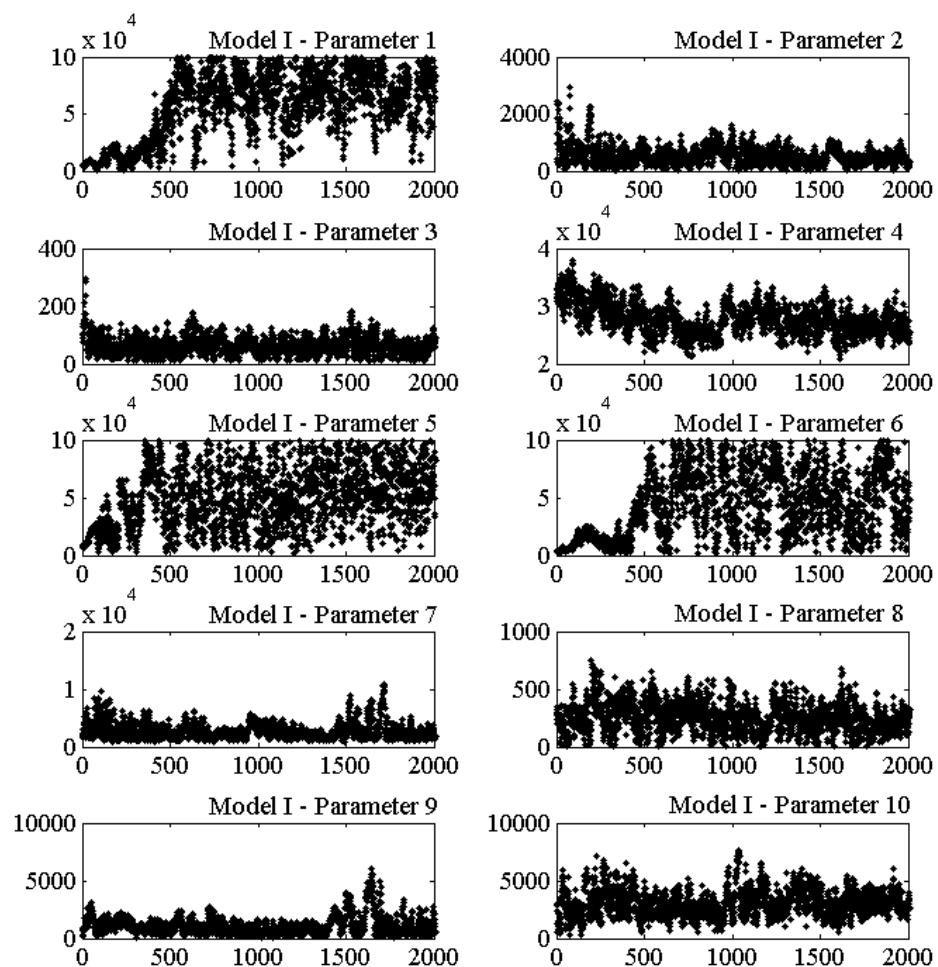


Figure 10-3 - Accepted Parameter Chains - Model I

To investigate this phenomenon, a second burn-in was run with a reduced range. The behavior was observed except over the reduced range, as seen in Figure 10-4.

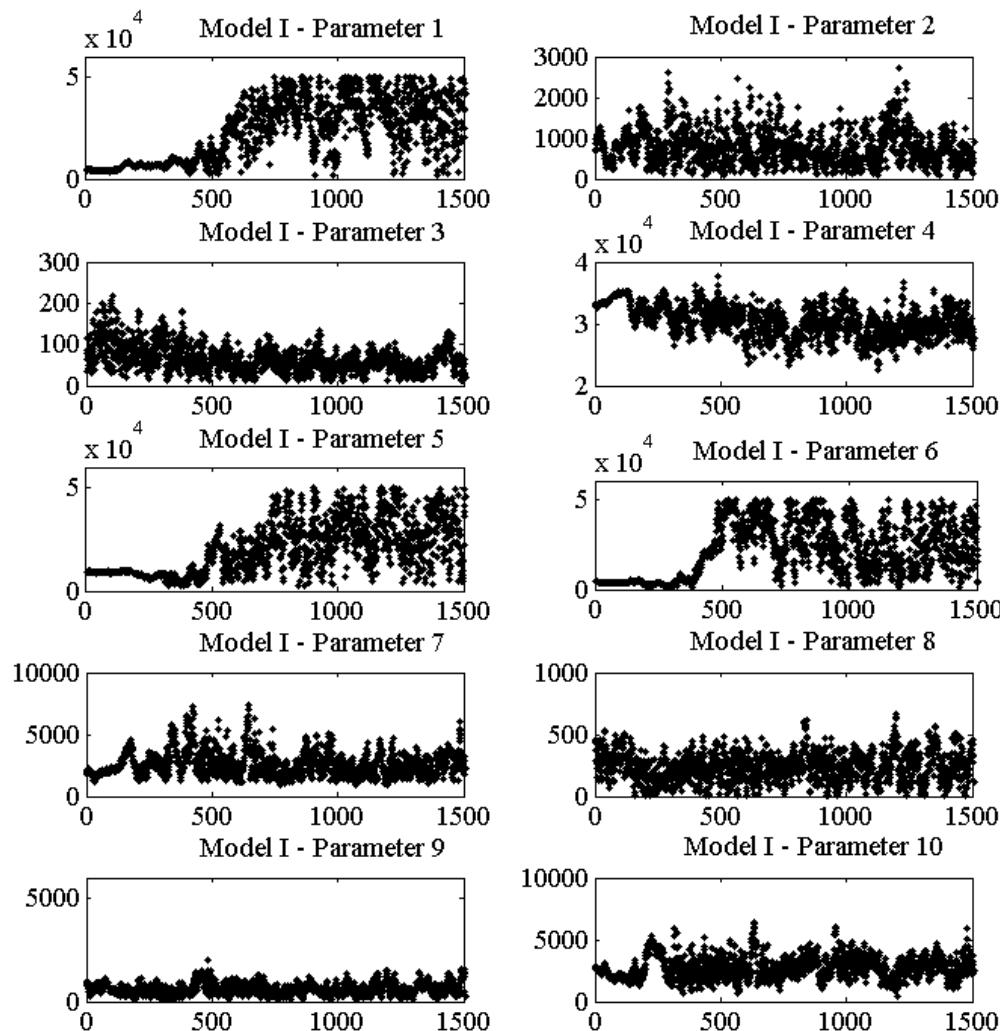


Figure 10-4 - Accepted Parameter Chains - Model I with Reduced Ranges

An investigation using a “dummy” parameter was also run. This proposal function was updated using the results from the burn-in of the same model, but the dummy parameter was completely independent of the observations (i.e., it was not included in the model or calculation of the likelihood). The observed behavior was similar to what is observed with the parameters previously mentioned. Figure 10-5 shows the proposal function for the dummy parameter, and how it expands with time. This expansion can only be contained by bounds.

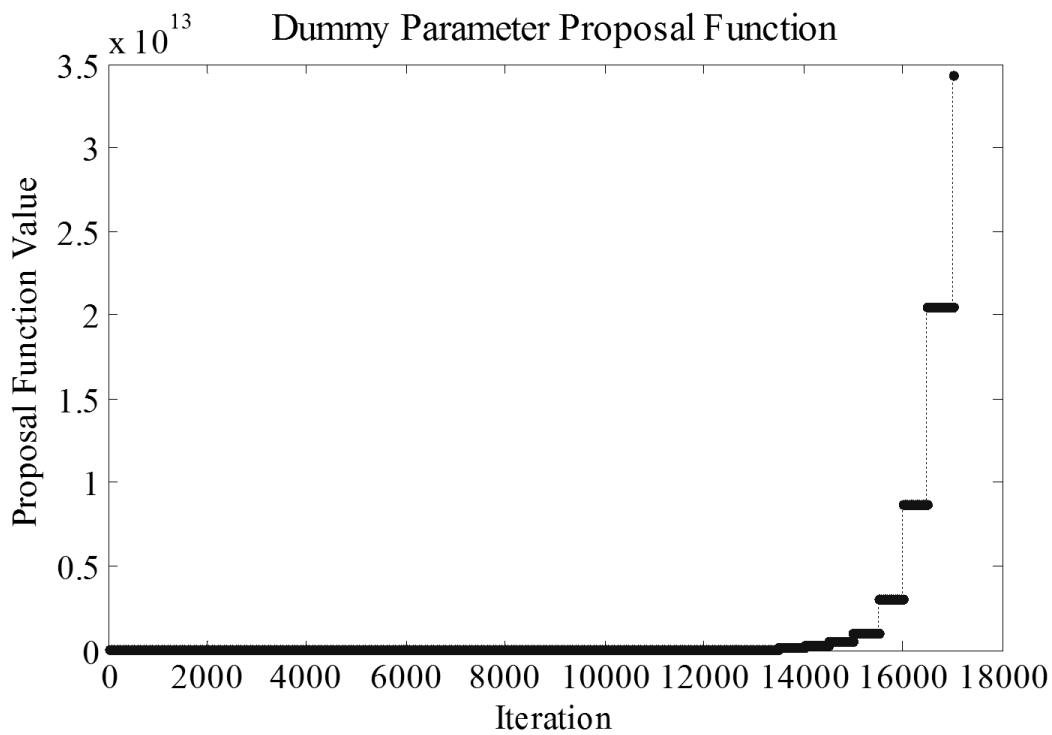


Figure 10-5 - Dummy Parameter Proposal Function

It was concluded that these parameters, similar to the dummy parameter, were not able to directly inform the observations used (i.e. modal properties) except over the very small range located during the sensitivity studies. However there exists a direct tie between these parameters (composite action and vertical bearing stiffness) and the desired unobservable response predictions which did not factor into the initial sensitivity studies. It was decided that even though these parameters are not able to inform modal properties very well, they were important for response predictions and therefore were included in the analysis. Any subsequent increased variability of the response predictions was then a reflection on the viability of the experiment as a tool to inform these particular response predictions.

10.3.2.2 Effect of Parameter Starting Point on Burn-in Analysis

Recall that the starting point for each model form was the best result from a suite of ten optimization analyses. It is very likely that the analysis started at a place where there is a local maximum in the likelihood space.

To investigate this, a second starting point was implemented for every model. This coincided with the well-accepted notion that the most effective method of establishing burn-in is by running simultaneous chains and verifying that they converge at the same position (Gilks, Richardson et al. 1996). This would have been implemented from the start if not for the high computational cost, and should be standard in any future MCMC analyses, despite this cost.

The second starting point often showed different behavior than the first. Figure 10-6 shows the chains for all parameters of Model A for the original starting point and the new starting point. The original chain is shown in grey, and the second starting point is black.

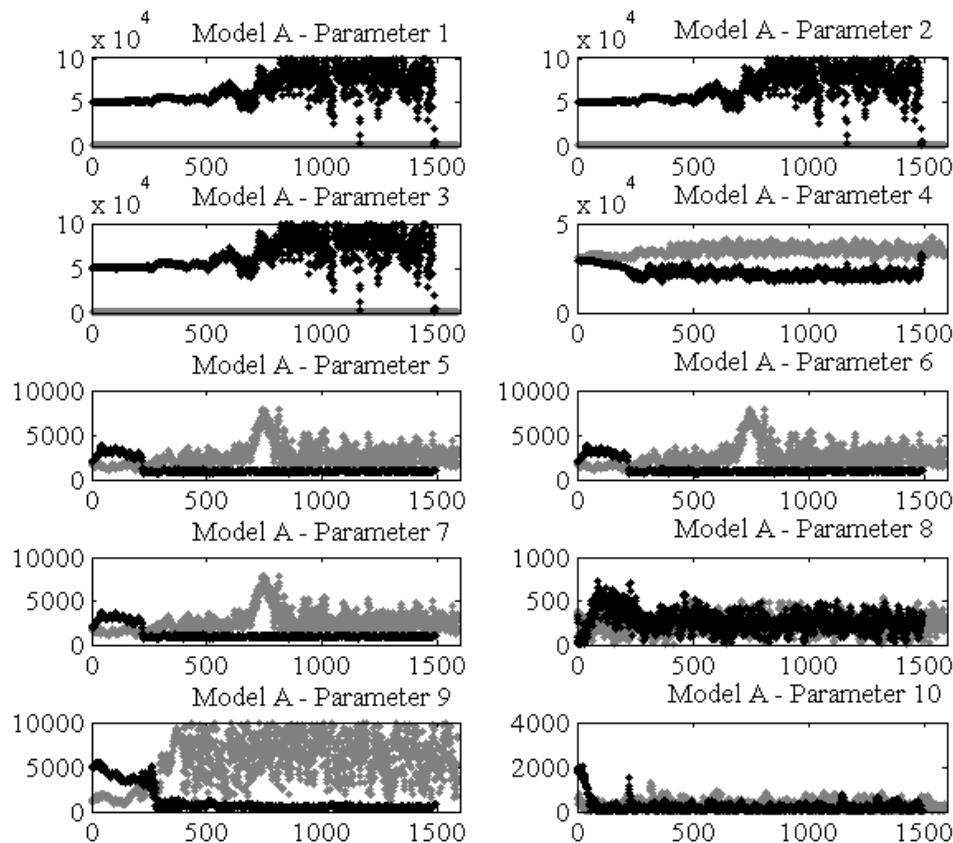


Figure 10-6 - Comparison of starting points for two MCMC runs of Model A

In the latter case, the Parameter #1 behaves as expected based on the rest of the model population, sampling in the higher portion of the parameter range. Some other

parameters, particularly 5, 6 and 7, move into unexpected areas, indicating that the interaction between the value and the spatial variation of these parameters is critical.

Figure 10-7 shows the comparison of the likelihoods of the accepted models for both starting points.

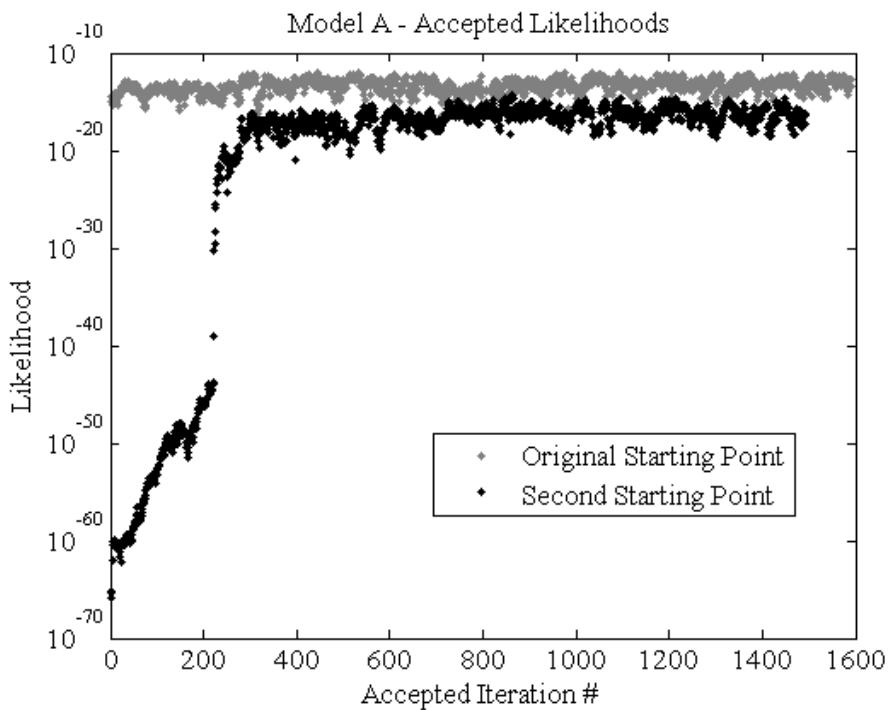


Figure 10-7 - Likelihood Comparison for Different Starting Points

The second starting point is clearly much poorer than the first, and while the analysis moved towards a better value, it never reached the same level as the first starting point. This indicates that despite the discrepancies in specific parameters, the area of posterior distribution that was being sampled from the first starting point was a better area, for that

model form, than the area that is more similar to what was observed in other model forms, as located during the second run.

It is acknowledged that these two chains did not burn in to the same point. In order to be considered converged, these two chains would have to have been run until they represent the same posterior distribution, which is clearly not the case here. This is because the adaptation of the proposal function can be aggressive, leading to small values which in some cases prevent the MCMC algorithm from rapidly finding the appropriate location on the model space. This is what occurred with the second starting point of Model A. If one assumes that these two chains were in different locations on the model space, then if the analysis of Model A from the second starting point were run infinitely, eventually a model would have been sampled and accepted in the same region of the model space as the first starting point, and future samples would converge to the initial starting point samples. The ability of the MCMC algorithm to find the correct posterior distribution with enough iterations is the most substantial benefit of the algorithm. It is simultaneously a weakness because the required run-time is always uncertain and there is no guaranteed method for completely ensuring that convergence has occurred or that it has occurred in the highest likelihood region.

10.3.2.3 Discussion of Optimization and Burn-in

The process of developing a candidate model population, identifying reasonable starting points and burning the models in to establish a fair framework for comparison was difficult and should be considered the most critical step of any multiple model method. There are numerous ways to inadvertently pollute the results. Also there is no way to be

completely sure that burn-in has actually occurred. The decision to consider the model burned in and to move on to a traditional MCMC analysis with a fixed proposal function is based heavily on the user and is dependent on structural engineering knowledge as well as mathematics.

Based on what was observed for the IBS Bridge, gradient-based optimizations to locate a better starting point were helpful, but were not adequate to eliminate the need for multiple burn-in analyses. Global optimization techniques could be used to determine potential starting points for MCMC analyses. A similar concept is Population MCMC, as discussed in Nichols, Moore et al. (2011).

Perhaps the most troubling aspect is that despite all of these second checks and anecdotal confirmation that the burn-in had reached the area of the best likelihoods, there was simply no way to be completely sure that there is not another area not sampled yet with better results. On the other hand, this notion puts into sharp relief yet again the inadequacy of single model approaches.

10.3.2.4 Eliminate Inappropriate Models

The multiple burn-in analyses with varying starting points were analyzed, and based on the relative magnitudes of the maximum likelihood values it was evident that some models could likely be eliminated prior to running full analyses. For multi-dimensional analyses, it may be important to assign a weight to each model form, as discussed in previous chapters. To investigate the BIC weighting method discussed in previous

chapters on a real application, BIC weights for each model were calculated. The results are shown in Table 10-3.

Table 10-3 - Likelihood and BIC Weight from Burn-in

Model	Max Likelihood	BIC	BIC Weight
A	1.0300e-12	80.29	.1058
B	6.4607e-12	78.78	.1078
C	5.4993e-9	65.86	.1289
D	1.2296e-12	82.09	.1035
E	1.8640e-13	87.73	.0968
F	1.3570e-12	83.76	.1014
G	1.3018e-12	85.50	.0993
H	4.8577e-8	69.32	.1225
I	4.7683e-8	36.37	.1340

Clearly the BIC was not sensitive enough to changes in likelihood on this scale, because the resulting weights were essentially uniform. Therefore, a different tactic was needed if the clear discrepancies in likelihood for each model were to be properly reflected in the weights.

For the current study a manual weighing approach was employed. If the chain for a given model form did not have any parameter sets with likelihoods within two orders of magnitude of the model form that had the highest likelihood values, then it was assigned a weight of zero. This was primarily done because such a model would not have had a meaningful acceptance rate for any of the multi-dimensional MCMC approaches considered. It would have been accepted with a probability equivalent to that ratio of

those likelihoods, roughly 1:1000 or more. In addition, by assigning a weight of zero, the full MCMC chain did not have to be run at all, resulting in a large computational savings.

Of the nine element-level models selected for the IBS Bridge, only three were included in the full MCMC analysis. These were Model C, H and I. The common characteristic of these three models in particular was that the composite action parameters are all independent with the single spatial variation discussed previously. This does not mean that distribution of composite action was exactly this on the real bridge. It simply means that given the chosen parameters and the selected spatial distributions of composite action, this distribution was better able to represent the experimental data.

10.3.3 Development of the Full MCMC Chains for the Selected Model Populations

For the three selected model forms, a fixed proposal function was extracted from the adaptive burn-in results. Typically this was an average of the steady-state adaptation of the proposal function after the model was burned in. Alternatively, the last value of the proposal function could have been used as well. The entire burn-in period was adaptive, meaning the chain was Markovian only over the range where the proposal function was constant (between adaptations). This means that irrespective of how the fixed proposal function was determined, the full MCMC analysis has another burn-in period.

The goal for each MCMC analysis was to achieve 3000 accepted models after burn-in. This allowed, assuming equal weights, for an MC3 analysis of approximately 12,000 models. Table 10-4 shows, for each model form, the total number of MCMC iterations, the number of bound exceedances, and the number of accepted models.

Table 10-4 - MCMC Chain Results

Model	Total Iterations	Number Accepted	Bound Exceedance	Acceptance
C	23,000	3,351	10,671	27.2%
H	34,000	3,093	20,481	22.8%
I	25,000	3,225	14,736	31.4%

The large numbers of bound exceedances seems inefficient but considering the prior probabilities of the parameters are uniform over a set of bounds; this was expected particularly when the mean of a parameter ended up being near a bound. The model was not actually executed when a bound was exceeded, so there was very little computational expense associated with this event.

The acceptance was calculated using the number of accepted models versus the number of model executions, or the total iterations less the number of bound exceedances.

10.3.4 Results from the Individual Chains

To ensure the chains were acceptable, the converged parameters, predictions of observations and predictions of unmeasureable responses from the individual chains were investigated. Comparisons between model forms as well as with the experimental data are presented in this section.

10.3.4.1 Convergence of Parameter Values

Since the MCMC chains for each model form were run individually, it is possible to estimate convergence to the posterior distribution based on the stability of the mean and standard deviation of the parameter values.

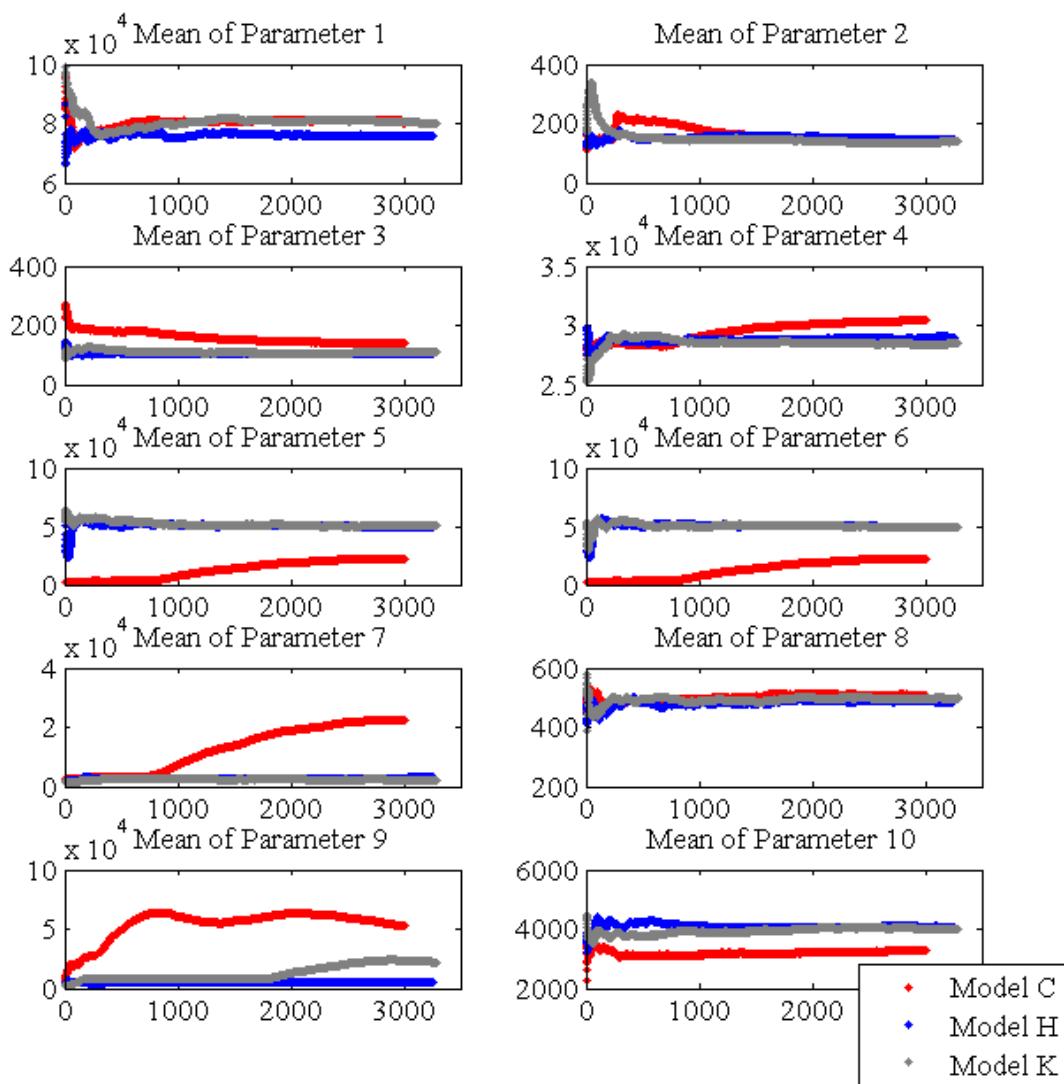


Figure 10-8 - Parameter Means for all Models

Figure 10-8 shows the mean of the parameters of the accepted models as the analysis was run. When this mean flatlines for every parameter in a given model, the model can be considered converged. With a few exceptions, the parameters are clearly converged.

For parameters 5, 6, and 7, Model C appears to grow. In this case, the proposal function was small, and the parameter tended to walk once it moved away from the low range of the assigned bound. Note that the range in which the parameter (vertical boundary stiffness) was walking was high, which essentially means that the boundary was fixed in the vertical direction, and therefore had no effect. The parameter behavior and the likelihood do not appear to be correlated (Figure 10-9), but there are two distinct regions which were sampled. The higher likelihood region was in the lower range of the parameter, and the lower likelihood region was in the higher range of the parameter. The shift to a lower likelihood area indicates the instability of this particular analysis. It was concluded that this spatial resolution of the vertical stiffness parameter is inadequate to represent the real structure.

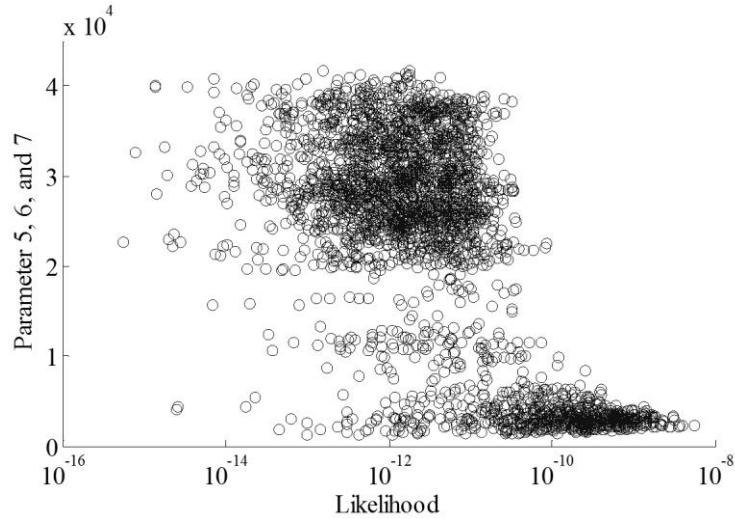


Figure 10-9 - Parameter 5, 6, and 7 vs. Likelihood - Model C

Regarding Parameters 5 and 6, the vertical bearing stiffness on the south side of the span, Figure 10-10 shows the plot matrix of these two parameters for Model I. The diagonal terms are histograms and the off-diagonal terms are scatter plots comparing the two chains. There is no indication of correlation between these two parameters. Both appear to be very similar, and to sample over the entire range. In Model H, these parameters are the same value, but still sample the entire range, effectively making Model H and Model I the same model.

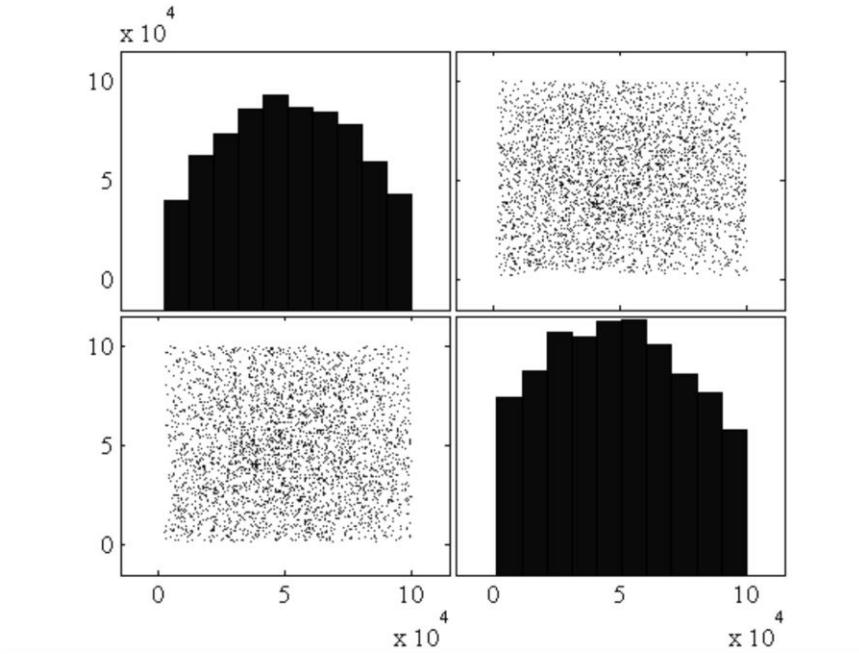


Figure 10-10 - Plot Matrix of Parameters 5 and 6

Figure 10-11 shows the accepted chains for all three models in the candidate population plotted together. The behavior of Parameters 5, 6, and 7 was already discussed. Note that for Parameter 9 in Model C and Model I, the chain appears to walk. Typically this indicates that the proposal function was too small. In this case the parameter is not particularly sensitive to the observations and therefore the proposal function was defined by the parameter bounds, as indicated previously with the dummy parameter study. The reason that the proposal function was too small was a simple error in the size of the bounds for the full MCMC analysis. The bounds were intended to be 0-1000 kip-in, but were included as 0-10,000 kip-in. The small proposal function relative to these bounds allowed the analysis to walk around the entire range, but there was no appreciable

correlation with the likelihood of the models. Since there was no effect, the MCMC analyses were not rerun.

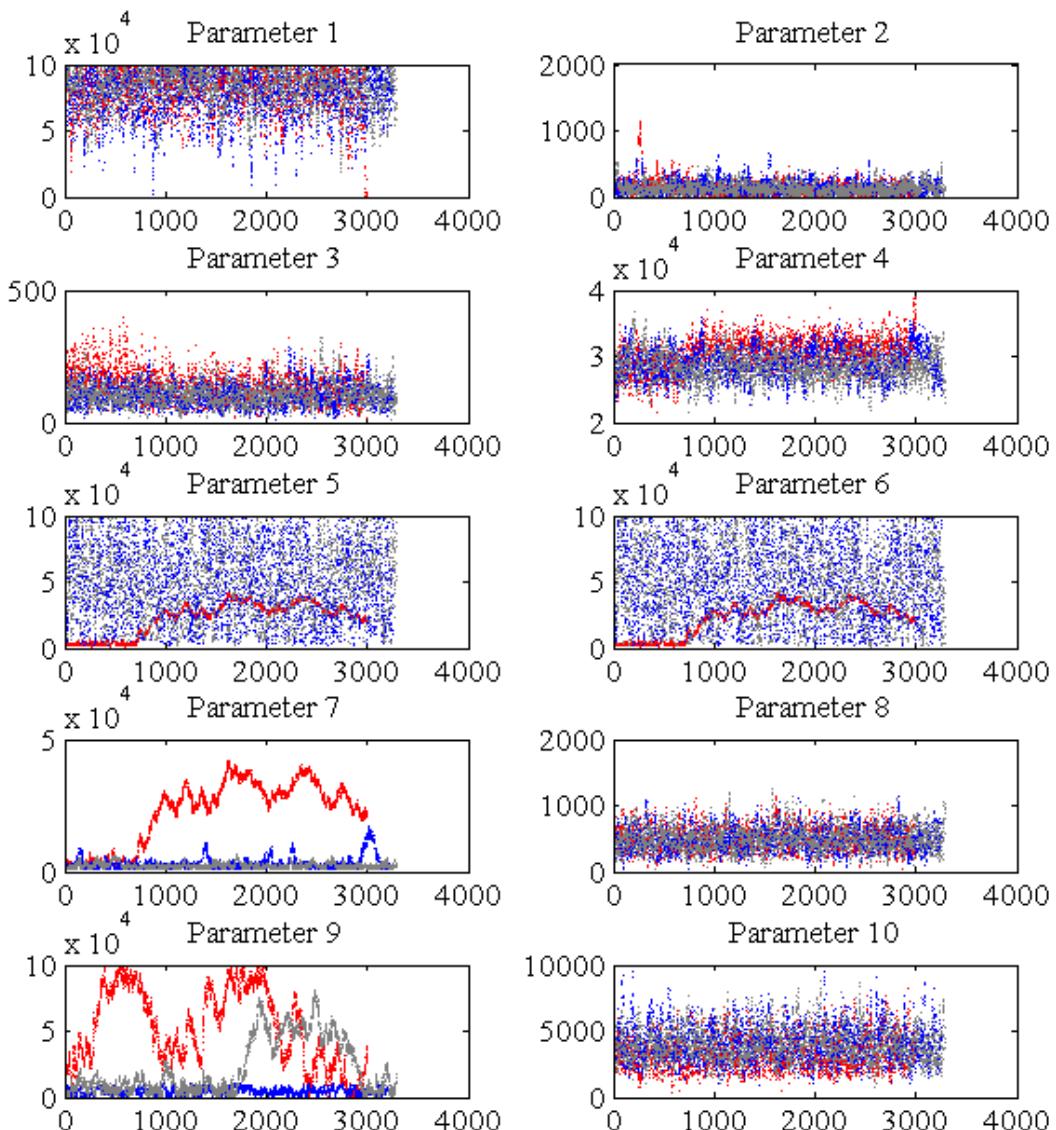


Figure 10-11 - Parameter Chains for all Models (Red = Model C; Blue = Model H; Grey = Model I)

10.3.4.2 Individual MCMC Chain Output – Frequencies

Frequencies and MAC values were used to inform the generation of the MCMC chain.

Figure 10-12 shows the CDF plot of the first frequency from each model form, and the experimental frequency value. The experimental value is shown as a vertical line at the measured value for clarity, but could be represented as a normal distribution with the mean at the measured value and an assumed standard deviation of 4% since that assumption has been made. When comparing between model forms, it is immediately clear that in terms of this frequency, Model H and Model I produce essentially identical distributions. Model C is slightly different. The first mode is very global and is likely not as sensitive to the smaller changes in parameters which supports this conclusion.

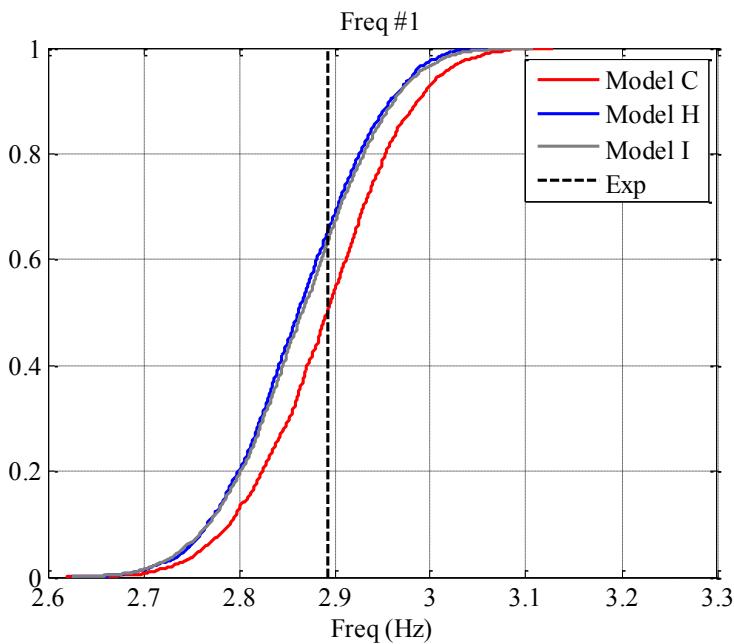


Figure 10-12 - CDF of 1st Frequency - All Models

When comparing how the models predict the experimental response, it is clear that there is some overlap in the distributions, and the mean of the experimental response is well within the model response distributions. The mean analytical frequency was approximately 2.88 Hz while the experimental response was 2.89 Hz.

Figure 10-13 shows the same plot for the second frequency. In this case, there are more discrepancies between the model forms, but the experimental prediction is worse. The measured value was 3.79 Hz while the model prediction mean was approximately 3.98 Hz. This corresponds to an error of 5% which is generally deemed acceptable for a model calibration effort.

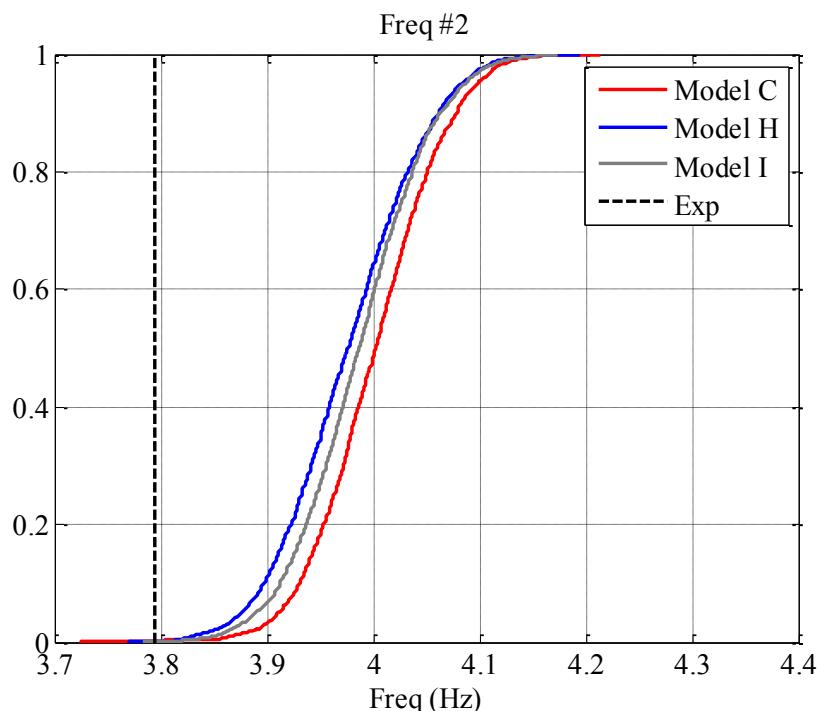


Figure 10-13 - CDF of 2nd Frequency - All Models

With the exception of the third frequency (Figure 10-14), all of the percent errors were within 5%. Mode 3 had percent errors as large as 9% which was the largest percent error observed in the comparison of analytical and experimental frequencies. The distinction between Model C and the other models is much clearer for this frequency.

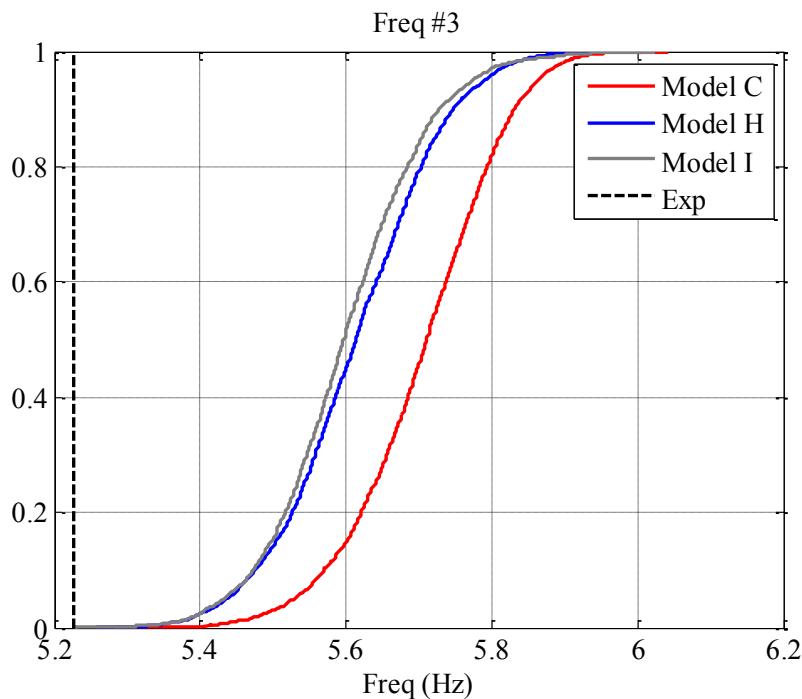


Figure 10-14 – CDF of Third Frequency – All Models

Table 10-5 shows the comparison of all frequencies in each model form to the measured experimental value. Table 10-6 shows the percent errors of the mean predictions with the

measured values. With the exception of the second mode, the average percent error is less than 6%.

Table 10-5 - Comparison of Frequency Prediction Mean to Experimental Values

Mode	Exp (Hz)	Mod C (Hz)	Mod H (Hz)	Mod I (Hz)
1	2.89	2.89	2.86	2.87
2	3.79	4.00	3.98	3.99
3	5.23	5.71	5.61	5.60
4	9.47	9.90	9.73	9.71
5	11.62	11.18	11.04	11.03
6	12.25	12.61	12.38	12.38
7	15.12	14.98	14.65	14.62

Table 10-6 - Comparison of Percent Errors of Frequency Predictions

Mode	Mod C (% Err)	Mod H (% Err)	Mod I (% Err)	Avg % Err
1	0.09%	1.04%	0.93%	0.69%
2	-5.48%	-4.82%	-5.07%	-5.12%
3	-9.19%	-7.43%	-7.17%	-7.93%
4	-4.55%	-2.70%	-2.48%	-3.24%
5	3.75%	4.93%	5.04%	4.57%
6	-2.92%	-1.10%	-1.06%	-1.69%
7	0.88%	3.09%	3.28%	2.42%

When compared to the experimental values, the frequency predictions ranged from negligible levels of percent error, up to approximately 9% error in the worst case. The

errors in the frequency are not all in one direction (i.e. all larger or all smaller) indicating that the discrepancies do not result from a bias error. The constraints within the process (e.g., that only certain parameters were updated in certain patterns) effectively resulted in a situation where the model was unable to fully replicate the observations. The most-sampled models were the ones that split the difference between all observations.

Predictions of unmeasureable responses are presented and discussed in Chapter 11.

10.3.4.3 Discussion of Individual MCMC Chains

The following conclusions were drawn by comparing the individual chains to one another and to the experimental observations.

1. Every model form showed larger likelihood values from the MCMC chain than were located during the optimization. This indicates that the complex likelihood space that results from a model with many, often correlated, parameters is not conducive to traditional, gradient-based optimization techniques. Rather, a global optimization technique, like simulated annealing or genetic algorithms, to locate regions of high likelihood may provide a better starting point for MCMC chains.
2. The three selected models all have the same parameters except for the vertical bearing stiffness which is different between these model forms. The frequency distributions from the models indicate that Model H and Model I have very similar results, while Model C is slightly different. This is supported by the unstable behavior of Model C discussed earlier. Model C has a single vertical

bearing stiffness value for the entire structure. This spatial distribution of this parameter is inadequate and invalidates Model C as a potential candidate.

3. Model H and I are very similar. The only difference between the two models is the vertical bearing stiffness of Girder #1 (Parameter 5) on the southern side of the span. In Model H, this bearing is lumped in with the rest of the southern bearings, and in Model I it is assigned a separate stiffness, to account for the piercap crack. In Figure 10-11, the chains for each parameter were shown. Parameter 5 for both Model H and Model I was sampled over the entire range assigned, and therefore the models (given the set of experimental observations) are essentially identical. This is further supported by investigating Parameters 5 and 6 within Model I. Figure 10-10 shows a plot matrix of these two parameter chains. There is no correlation between them, and they sample the full range just as the single parameter samples the full range in Model H. The end result is essentially one model form.
4. In terms of computational effort, developing full MCMC chains even for a reduced candidate model population was very time consuming and difficult. It is anticipated that some of the difficulty resulted from the learning curve that comes with a new process like MCMC using FE models for constructed systems. Also, these efforts did not make use of any parallel computing or processing power available through super computers. The ease of the process could be greatly improved by migrating the computing to more powerful computers. Regardless, in its current form, this process is not a candidate for application to every structure.

There is an immense amount of knowledge to be gained from these analyses, but there is a cost of time and effort that are not justified for every structure.

10.4 Mixing the Chains via MC3

In the previous section, a model population of ten models was reduced to three candidates and individual MCMC chains were run on each model. Of these three models, one (Model C) was determined to be inadequate and is only presented for comparison. The two remaining models were shown to be essentially identical. This means that mixing these particular models with MC3 is not relevant, and therefore MC3 was not be applied to the IBS Bridge.

This is not a reflection on the MC3 method, which was shown to be effective for the cantilever beam in Chapter 9. Rather, this indicates that the selected model population was stratified to enough of a degree that only one model of the candidate population was viable. This does not mean that there are no other viable model forms which would result in similar responses to Model H and Model I, but simply that they were not include in the original model candidate population.

The procedure applied to determine the models to be included in the MC3 analysis instead functioned as a robust method of model selection. Through a direct numerical comparison of likelihoods, which is a reflection of the model's ability to predict observations, some candidates were eliminated. Of the remaining population, heuristics regarding the behavior of certain parameters, given a set distribution, was used to further eliminate candidates. Finally, the remaining two models were determined to be

essentially identical, leaving a single model form from the original set of candidate models.

The ability of this model form to predict unmeasured responses will be assessed in Chapter 11, and a comparison of this model form to traditional model updating will be presented in Chapter 12.

Chapter 11: Results of MM St-Id of IBS Bridge

This chapter is a discussion of the results for the individual MCMC chains introduced in Chapter 10. The MCMC chains are compared against the experimental responses for displacements and strains as well as each other. The focus of the assessment is the ability of the model to translate one type of observation into a prediction of a different response quantity. Similarly, the veracity of predictions at an extreme (i.e., nonlinear) load level versus within the elastic range of the structure are compared. In the previous Chapter, Model C was shown to be inadequate to represent the structure, but the response predictions of this model are included here for the sake of comparison.

11.1 Uncertainty Regarding Predictions

Recall that an inherent standard deviation of 4% was assumed to encompass the limitations of our ability to predict measurements using FE models. This standard deviation factored in to the calculation of the likelihood of the observations given the model, for Bayesian inference. Here the standard deviation was again factored into the response predictions considering the same logic. For each resultant prediction value from the chain, a set of 100 additional samples was randomly generated. To generate these samples, the resultant prediction value was assumed to be the mean of a normal distribution with a standard deviation of 4% of the mean. The added samples were included in the construction of the distribution plots to explicitly represent the additional assumed 4% inherent uncertainty.

11.2 Experimental Value vs. Mean Value vs. Maximum Likelihood

The ensuing sections investigate the response predictions of displacements and strains from the individual MCMC chains for the candidate model population (i.e., the accepted models for each model form). The response predictions are presented through cumulative distribution functions, or CDFs. Typically, the response predictions are plotted considering the inherent 4% standard deviation, and are compared with the measured experimental value. Recall that these experimental results were not used in any way during the sampling and weighing of the models.

There are two other quantities of interest (mean prediction and prediction from the most likely model) that are not included in the plots, simply because the plots became too complicated and confusing. A sample displacement prediction including these values is shown in Figure 11-1. The mean prediction value is not shown on the plots typically, but is tabulated at the end of each section. For a normal distribution, the mean would cross the CDF plot at a cumulative probability of 0.5. Most of the response predictions are not perfectly normal, but the mean tends to be towards the center of the distribution.

Considering Figure 11-1, it is apparent that the mean prediction and the prediction from the model of maximum likelihood are not the same. In most cases for this investigation, this held true. The magnitude of the difference varied, but the concept that the most likely model based on the observations would be associated with the mean of the predictions is not true in general.

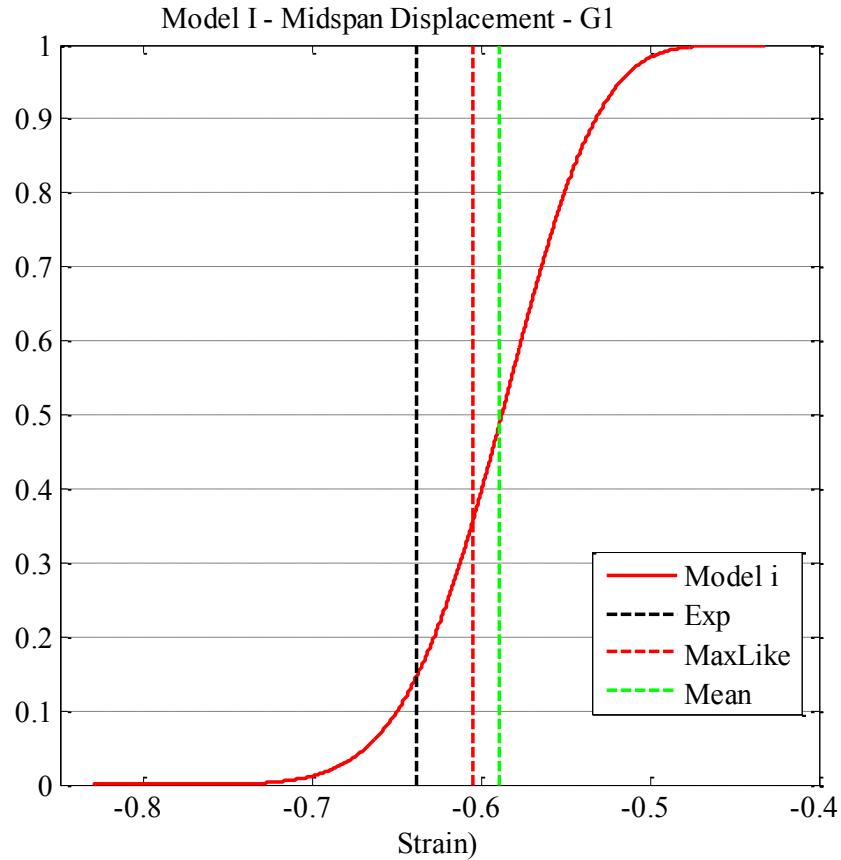


Figure 11-1 - Sample Displacement Prediction Showing Mean, Experimental Value, and Maximum Likelihood

11.3 Translation from Measured Observations to Unmeasured Response Predictions – MCMC Chains

In Chapter 10, the frequency outputs from the models which comprised the individual chains were shown. Recall that frequencies and mode shapes were the observations used to inform the calculation of likelihood for MCMC sampling. The ability of MCMC sampling within a model form to translate between observation of modal properties and prediction of other responses is assessed in the following sections.

11.3.1 MCMC Chain Output – Displacements from Final Load Stage

The displacement measured for several load cases serves as a verification data set to assess the ability of the MCMC sampling using modal observations to develop a series of models that accurately represent the bridge of interest. The displacement sensor layout is shown in Figure 5-5. A grid of twelve responses, based on the longest girders dimensions, was recorded throughout the truck testing, and the responses from the final load stage (i.e., six full trucks) are used here.

Figure 11-2 shows the CDF of the three displacements along Girder #1 for each model, and the experimental value. Similar to what was observed with the frequencies; the displacement predictions indicated that despite the individual bearing stiffness under the south bearing of Girder #1, Model H and Model I produced similar response predictions. Note that Model C also produced similar result predictions, despite the fact that the spatial variation of bearing stiffness of this model is questionable.

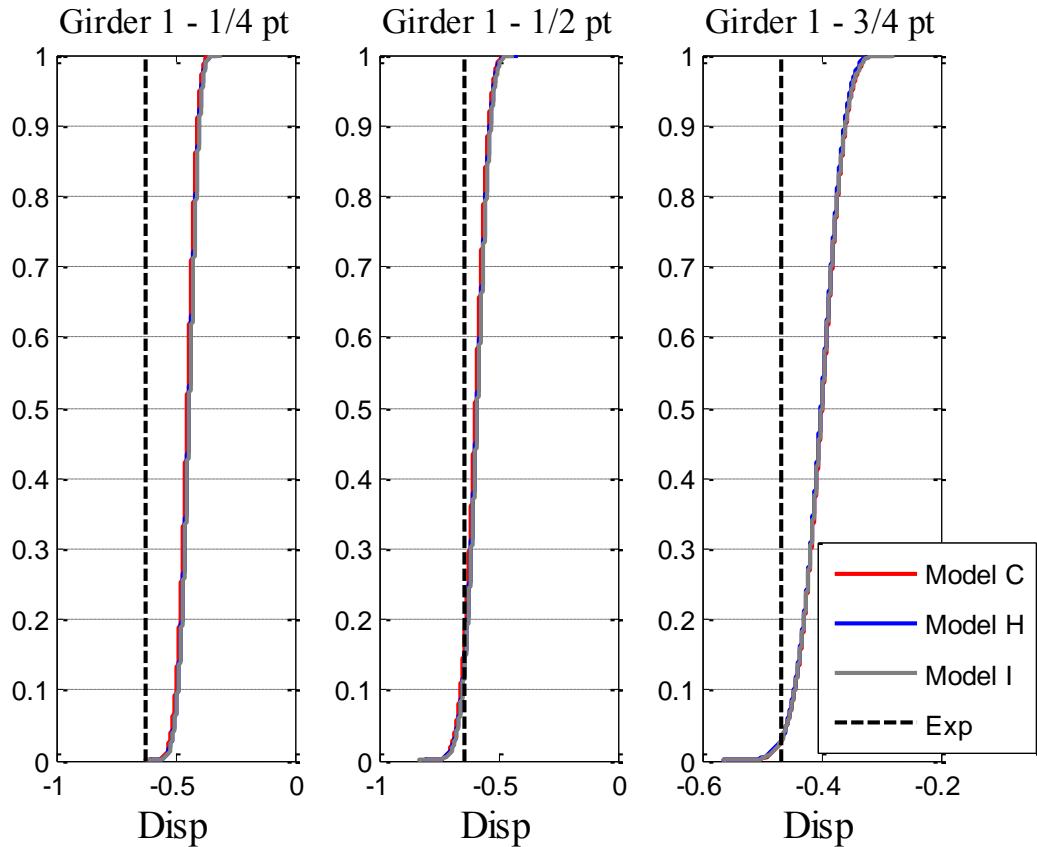


Figure 11-2 - Girder #1 Displacement Predictions – Six Full Trucks

The midspan and three-quarter span locations have reasonable predictions of displacement, but the mean prediction at the one-quarter span location has approximately 29% error. This location corresponds to the softness that was attributed to the softness of the piercap discussed in Section 5.11. Recall that in some models, Girder #1 was given a specific vertical bearing stiffness as a parameter to allow for the explicit inclusion of the piercap crack. The inability to predict that response indicates that the observed modal parameters did not reflect the softness of the pier cap or that this information was drowned out during the MCMC sampling. This is verified by examining the chain for that

parameter in Model I, where the G1 vertical stiffness parameter was free to move about.

The chain shown in Figure 11-3 indicates that values were sampled throughout the available range for the parameter, and that it had little effect on informing the observations.

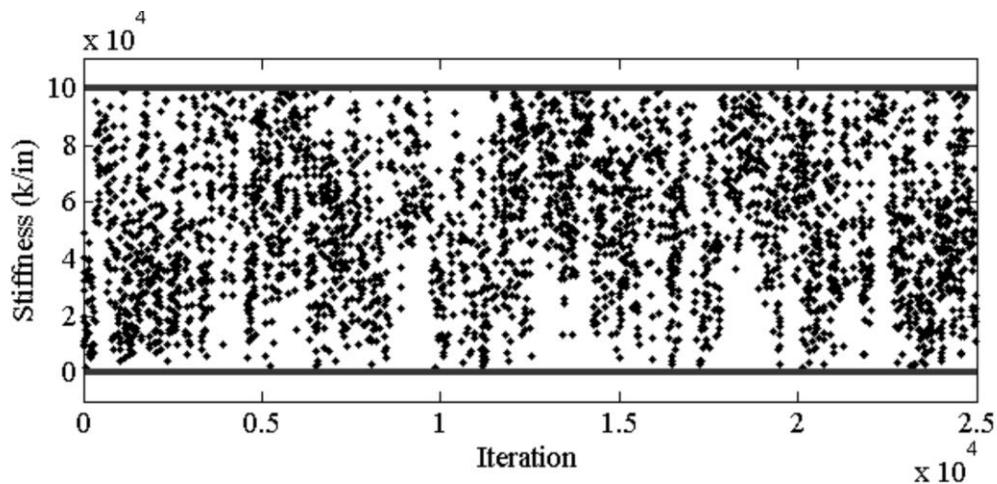


Figure 11-3 - Girder #1 Vertical Bearing Stiffness

Figure 11-4 shows the CDF of the three displacements along Girder #3 for each model, and the experimental value. In this case, the model forms all provide very similar predictions, which slightly under-predict the experimental displacements. Similarly, Girder #6 showed consistent predictions, as seen in Figure 11-5. In this case, the mean predictions at the one-quarter and midspan points had errors around 12% but the three-quarter point had an error of 17%.

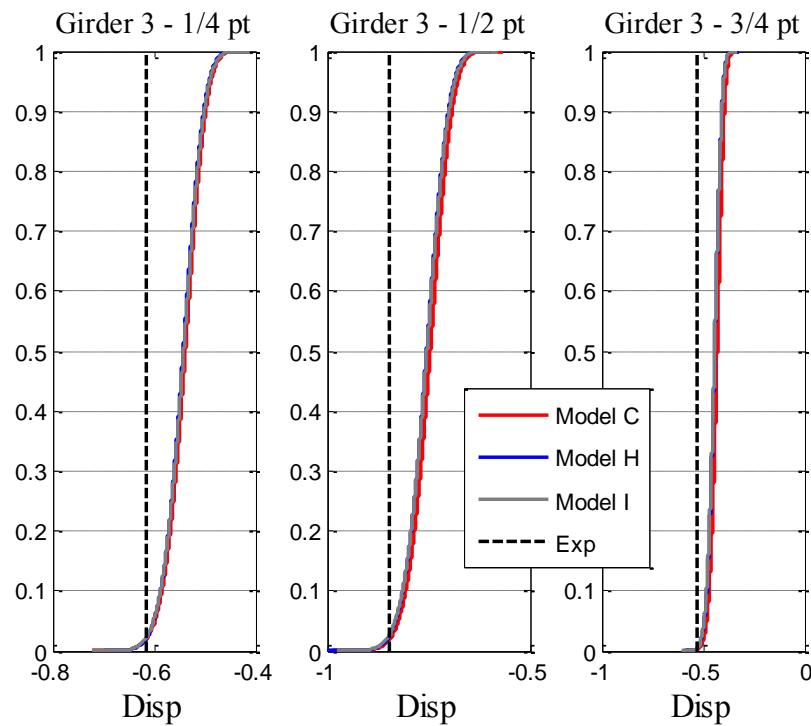


Figure 11-4 - Girder #3 Displacement Predictions – Six Full Trucks

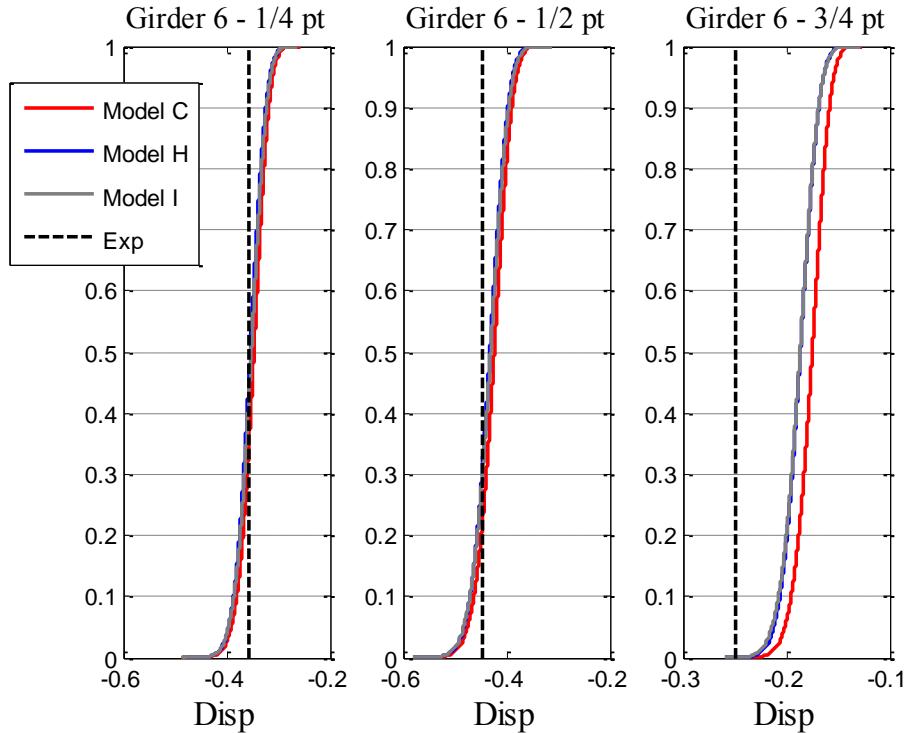


Figure 11-5 - Girder #6 Displacement Predictions – Six Full Trucks

The predictions for Girder #8 are shown in Figure 11-6. This location was not directly loaded during the static load test, and the resulting response magnitudes were very small. These measured values are not very reliable in this region simply because of the small magnitude of response, particularly relative to the sensor resolution. Also, recall that the lane over these sensors was open to traffic during the static test, which introduces noise and uncertainty. The experimental responses were contained within the distributions, but the mean prediction errors ranged from 15% to 32% because of the small magnitude of the response.

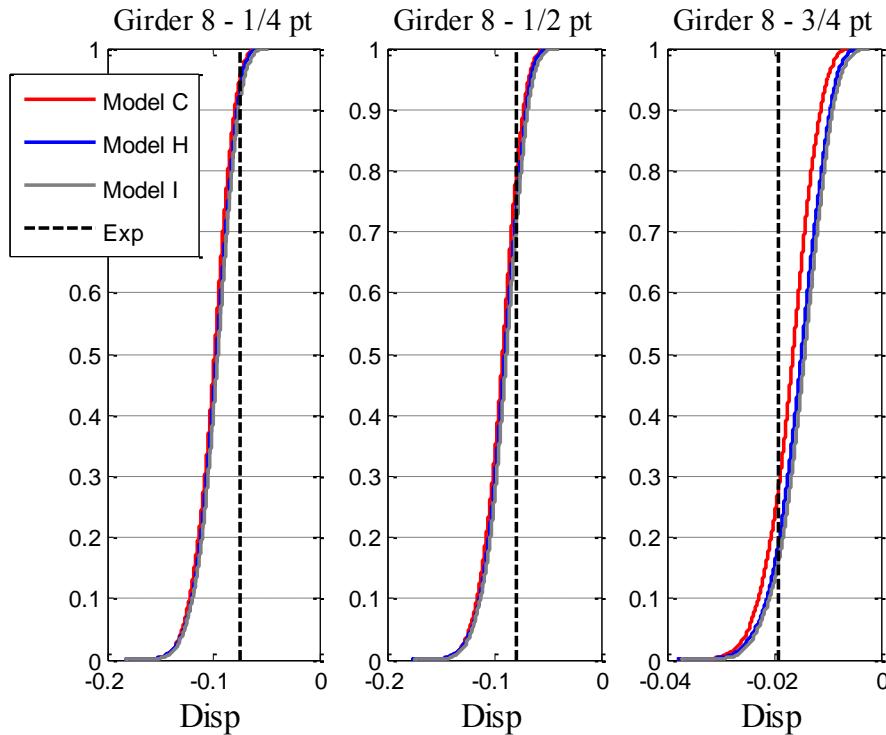


Figure 11-6 - Girder #8 Displacement Predictions – Six Full Trucks

Table 11-1 shows the mean response predictions color-coded with contours corresponding to the largest displacements as green and the smallest as red. Table 11-2 shows the absolute values of percent errors for all 12 response locations with the minimum error as green and the maximum as red. Note that while the largest percent errors can be seen on Girder #8 (Location 10), in fact the largest magnitude error is Girder #1 (Location 1). The magnitude of the error at Location #1 is at minimum 0.173 in.

Table 11-1 - Displacement Predictions from Each Model Form – Six Full Trucks

Loc	Girder	Exp (in)	Mod C (in)	Mod H (in)	Mod I (in)
1	1 - 1/4 pt	-0.625	-0.452	-0.444	-0.443
2	1 - 1/2 pt	-0.639	-0.597	-0.591	-0.590
3	1 - 3/4 pt	-0.467	-0.401	-0.402	-0.401
4	3 - 1/4 pt	-0.616	-0.540	-0.544	-0.543
5	3 - 1/2 pt	-0.846	-0.748	-0.756	-0.756
6	3 - 3/4 pt	-0.530	-0.432	-0.444	-0.445
7	6 - 1/4 pt	-0.358	-0.348	-0.355	-0.353
8	6 - 1/2 pt	-0.447	-0.425	-0.433	-0.432
9	6 - 3/4 pt	-0.249	-0.176	-0.187	-0.187
10	8 - 1/4 pt	-0.075	-0.100	-0.098	-0.097
11	8 - 1/2 pt	-0.079	-0.093	-0.091	-0.089
12	8 - 3/4 pt	-0.019	-0.017	-0.015	-0.015

Table 11-2 - Percent Errors for Mean Displacement Predictions – Six Full Trucks

Loc	Girder	Mod C (% Err)	Mod H (% Err)	Mod I (% Err)	Avg % Err
1	1 - 1/4 pt	28%	29%	29%	29%
2	1 - 1/2 pt	6%	7%	8%	7%
3	1 - 3/4 pt	14%	14%	14%	14%
4	3 - 1/4 pt	12%	12%	12%	12%
5	3 - 1/2 pt	12%	11%	11%	11%
6	3 - 3/4 pt	18%	16%	16%	17%
7	6 - 1/4 pt	3%	1%	1%	2%
8	6 - 1/2 pt	5%	3%	3%	4%
9	6 - 3/4 pt	29%	25%	25%	26%
10	8 - 1/4 pt	34%	31%	30%	32%
11	8 - 1/2 pt	17%	15%	13%	15%
12	8 - 3/4 pt	12%	21%	24%	19%

Table 11-3 shows the coefficient of variation (i.e., the standard deviation normalized by the mean) of the displacement predictions. The table is color-coded such that large coefficients of variation are red and small values are green. Comparing individual girders, clearly all model forms have consistent predictions for Girder #3 and much more variable predictions for Girder #8. Again this is also a factor of the magnitude of the response at Girder #8.

Table 11-3 - Displacement Predictions - Coefficient of Variation

Loc	Girder	Model C	Model H	Model I
1	1 - 1/4 pt	0.067	0.068	0.069
2	1 - 1/2 pt	0.065	0.066	0.066
3	1 - 3/4 pt	0.072	0.071	0.072
4	3 - 1/4 pt	0.048	0.046	0.047
5	3 - 1/2 pt	0.044	0.042	0.043
6	3 - 3/4 pt	0.055	0.052	0.054
7	6 - 1/4 pt	0.057	0.056	0.057
8	6 - 1/2 pt	0.055	0.053	0.055
9	6 - 3/4 pt	0.072	0.067	0.070
10	8 - 1/4 pt	0.160	0.166	0.171
11	8 - 1/2 pt	0.176	0.184	0.189
12	8 - 3/4 pt	0.264	0.304	0.316

11.3.2 MCMC Chain Output – Strains from the Final Load Stage

Recall that longitudinal strains were measured at the same locations as the displacements, as originally shown in Figure 5-10. Figure 11-7 shows the CDF of the strain predictions

on Girder #1 under the final, six-truck load case as compared with the measured response. As previously discussed during the manual calibration, strains are a predominately local phenomenon and therefore are typically not as well correlated with modal properties as are displacements (Dubbs 2012). The models produced similar predictions to one another along Girder #1, and the experimental values were contained in the prediction distributions. The error between mean prediction and measured experimental strain at midspan was only 6%. The models predicted the midspan strain very well, which indicates that the strain at this location is predominately influenced by the global behavior of the span, and that there are no substantial local influences. However, the error was as large as 29% at the quarter spans. This trend continued on Girder #3 and #6 as well.

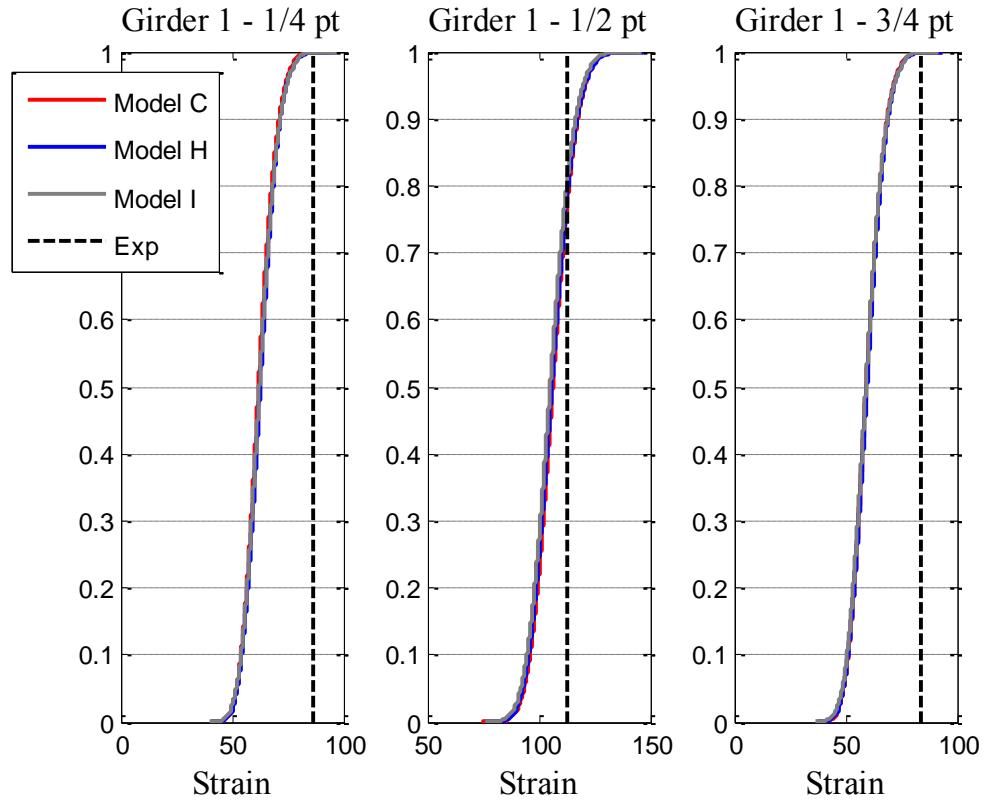


Figure 11-7 - Girder #1 Strain Predictions – Six Full Trucks

Figure 11-8 shows the responses for Girder #3, where maximum strain was measured. In this case the model predictions were very similar to one another but miss the three-quarter point experimental observation by an average of nearly $56 \mu\epsilon$.

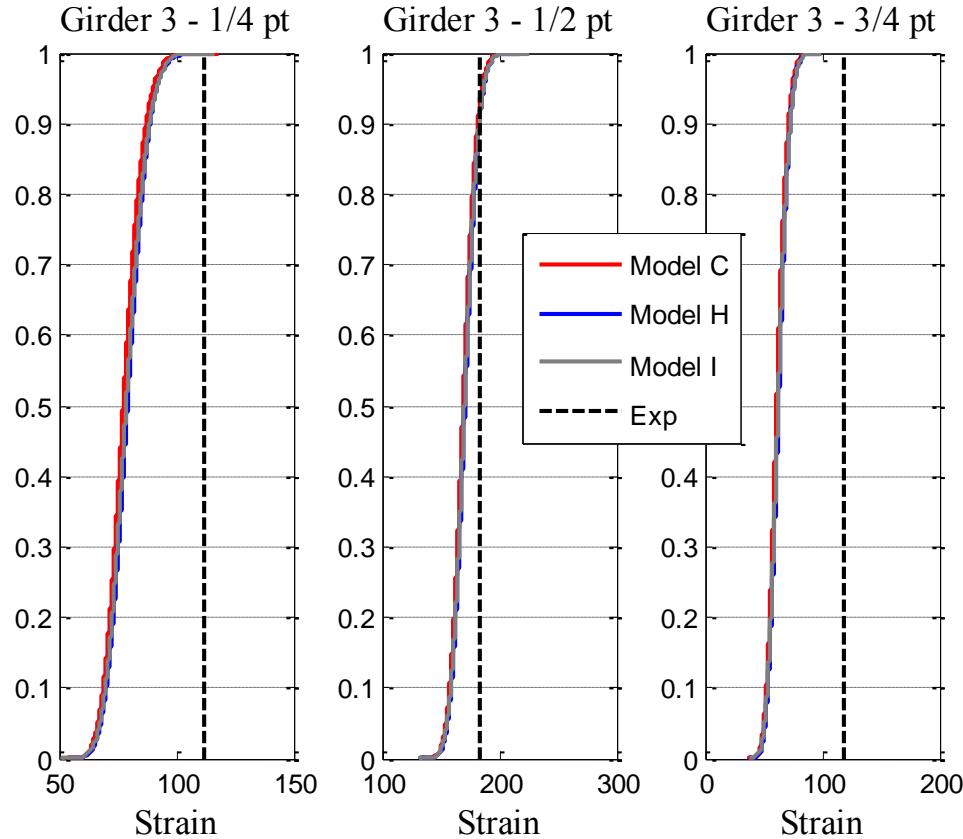


Figure 11-8 - Girder #3 Strain Predictions – Six Full Trucks

Figure 11-9 shows the strain response CDFs for Girder #6. Again the models tended to produce similar predictions. The midspan strain prediction was within 2% on all models. Note that at the quarter points, the predictions were all substantially smaller than the experimental value again, similar to Girders #1 and #3. This could be due to many reasons including that the strains in these areas are influenced by local mechanisms that do not influence the global modal parameters. The composite action parameter is much more influential on strain values away from midspan, because the shear force carried by the beam gets larger.

Also, as evidenced by the crack in the pier cap, there is certainly non-linearity present in the structure, particularly at larger static load levels. The extent to which any nonlinearity was activated under the ambient loading is not known. It is likely that there was some nonlinearity under ambient traffic, but not as much as under the static loading.

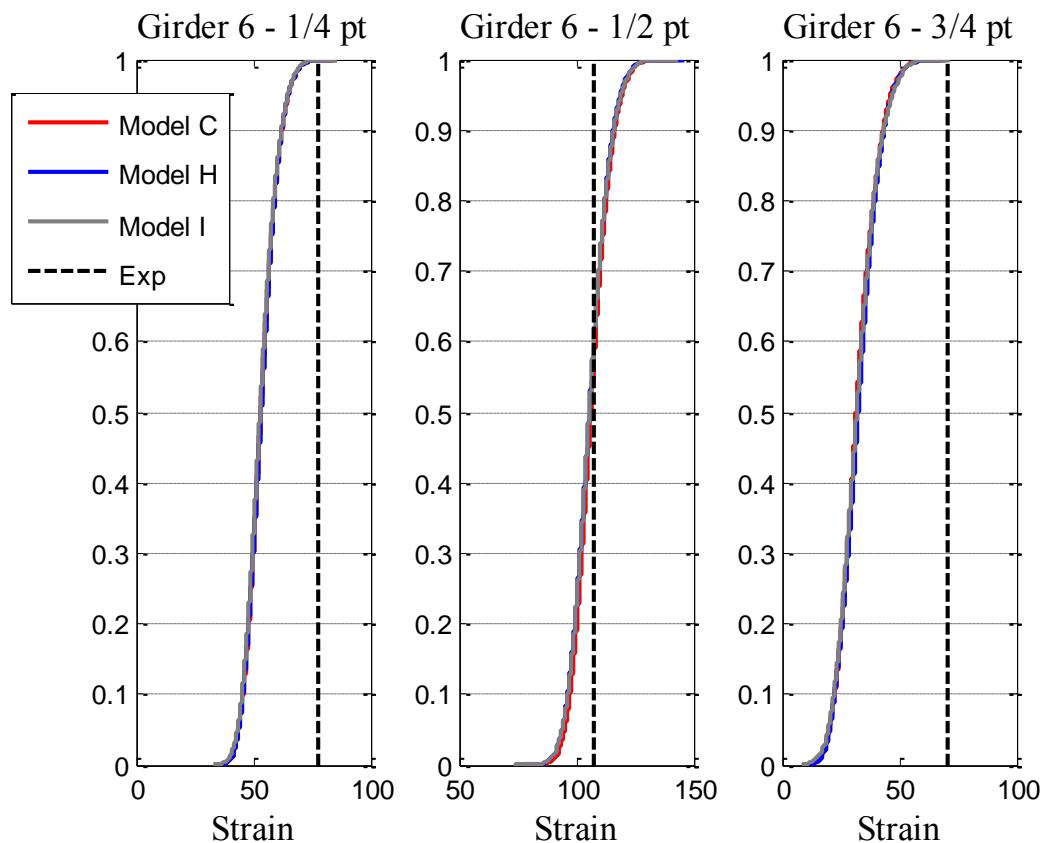


Figure 11-9 - Girder #6 Strain Predictions - Six Full Trucks

The strain predictions for Girder #8 under six full trucks are shown in Figure 11-10. For this girder, the one-quarter and midspan locations over-predicted the response substantially, while the three-quarter location was under-predicted and actually indicated compression in the bottom flange. The average strong-axis moment demand at the three-quarter point on Girder #8 was -220 kip-in., which corroborates the negative values.

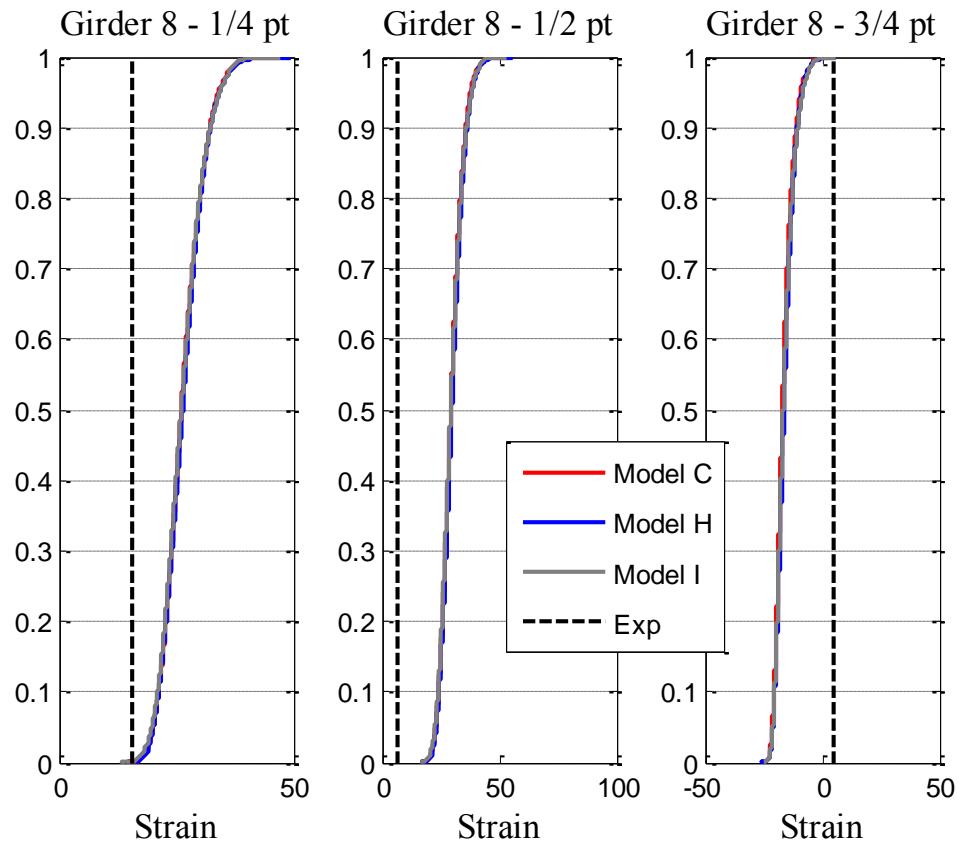


Figure 11-10 – Girder #8 Strain Prediction – Six Full Trucks

Table 11-4 shows the response predictions with the maximum responses color-coded green and the minimum as red. Table 11-5 shows the subsequent absolute value of percent errors for all locations in each model form with the maximum errors as red, and the minimum as green.

Table 11-4 - Strain Predictions in Each Model Form – Six Full Trucks

Loc	Girder	Exp ($\mu\epsilon$)	Mod C ($\mu\epsilon$)	Mod H ($\mu\epsilon$)	Mod I ($\mu\epsilon$)
1	1 - 1/4 pt	86.2	62.3	62.9	62.6
2	1 - 1/2 pt	113.1	106.6	106.2	105.4
3	1 - 3/4 pt	83.9	59.6	59.9	59.4
4	3 - 1/4 pt	111.9	77.7	79.3	78.9
5	3 - 1/2 pt	183.6	169.5	170.9	170.6
6	3 - 3/4 pt	118.4	61.6	62.9	62.6
7	6 - 1/4 pt	77.4	53.7	53.7	53.3
8	6 - 1/2 pt	107.7	106.6	105.9	105.9
9	6 - 3/4 pt	70.4	32.3	32.9	32.4
10	8 - 1/4 pt	15.7	26.5	26.6	26.4
11	8 - 1/2 pt	6.2	29.7	30.1	29.7
12	8 - 3/4 pt	5.0	-16.3	-15.8	-16.0

Table 11-5 - Percent Errors for Strain Predictions – Six Full Trucks

Loc	Girder	Mod C (% Err)	Mod H (% Err)	Mod I (% Err)	Avg % Err
1	1 - 1/4 pt	28%	27%	27%	27%
2	1 - 1/2 pt	6%	6%	7%	6%
3	1 - 3/4 pt	29%	29%	29%	29%
4	3 - 1/4 pt	31%	29%	29%	30%
5	3 - 1/2 pt	8%	7%	7%	7%
6	3 - 3/4 pt	48%	47%	47%	47%
7	6 - 1/4 pt	31%	31%	31%	31%
8	6 - 1/2 pt	1%	2%	2%	1%
9	6 - 3/4 pt	54%	53%	54%	54%
10	8 - 1/4 pt	69%	70%	68%	69%
11	8 - 1/2 pt	382%	388%	382%	384%
12	8 - 3/4 pt	427%	417%	420%	421%

The most inaccurate locations tended to be at the three-quarter points of each girder.

Multiple model methods provide a direct measure of this variability within the model population. Table 11-6 shows the coefficient of variation of strain predictions at each location in each model form, without the inherent 4% standard deviation factored in the predictions. Similar to the displacements, the table is color-coded to highlight the magnitude of the coefficient of variation. Looking at the individual girders, it is clear that the midspan response for each girder has the smallest variability, and the three-quarter point has the largest variability. The one-quarter point also had substantial variability when compared to the midspan.

Table 11-6 - Strain Prediction Coefficient of Variation

Loc	Girder	Model C	Model H	Model I
1	1 - 1/4 pt	0.082	0.085	0.087
2	1 - 1/2 pt	0.058	0.062	0.062
3	1 - 3/4 pt	0.086	0.092	0.094
4	3 - 1/4 pt	0.066	0.070	0.071
5	3 - 1/2 pt	0.035	0.035	0.036
6	3 - 3/4 pt	0.089	0.092	0.095
7	6 - 1/4 pt	0.092	0.096	0.098
8	6 - 1/2 pt	0.053	0.056	0.056
9	6 - 3/4 pt	0.164	0.177	0.183
10	8 - 1/4 pt	0.110	0.120	0.122
11	8 - 1/2 pt	0.108	0.117	0.119
12	8 - 3/4 pt	0.452	0.400	0.392

Recall that the three-quarter point was the side of the bridge with the skewed edge. The manner in which the bridge distributes load transversely in this region of the bridge is complex. Two parameters, composite action and diaphragm stiffness both directly influence transverse distribution as well. Clearly predictions from updating the model using modal parameters as observations are unable to represent the true behavior of this region as there is likely significant local turbulence. This could also be a function of the different in the magnitude of loading and any subsequent nonlinearities that occur at the high static load level that were simply not present at lower load levels, like under ambient traffic.

11.3.3 Comments on the MCMC Chain Response Predictions

There are several conclusions regarding the response predictions of the MCMC chains for the individual model forms.

1. The displacement predictions were generally smaller than the experimental predictions, but within reason. The notable exceptions were the displacement at the one-quarter point of Girder #1, and all of Girder #8 (discussed below). At the one-quarter point of Girder #1, all models under-predicted the displacement by nearly 30%. In the case of Model I, this occurred despite the inclusion of a vertical stiffness parameter for Girder #1 which was directly intended to inform this response. More discussion of this is provided in Chapter 12 in the comparison to traditional St-Id.
2. The strain predictions were highly variable, consistent with what was seen with single model structural identification efforts where the model was updating with dynamic properties and subsequently used to predict strains. The most variable region appears to be the northern portion of the span (i.e., three-quarter span), which has substantially wider distributions of response predictions. At the midspan of all girders except Girder #8, the mean strain prediction was as accurate as the mean displacement prediction at that location. Under the centered loading, there is little shear force at midspan so the composite action parameter is not as directly influential in this region.
3. The strain response on Girder #8 had very high percent errors. The displacement percent errors were also large, but the magnitude of these errors (about 0.025 in.) was reasonable. These types of errors were observed during manual calibration of

the single model as well. The magnitudes of responses were very low at this location because no load was directly applied there. More discussion of this is provided in Chapter 12 with a comparison to traditional St-Id.

11.4 Translation from Service to Proof Limit State –MCMC Chains

In the previous section, the response predictions for the final, six truck load stage were presented and comparisons regarding the ability of one type of observation to predict a different type of response were drawn. The observations were modal parameters, which were recorded at much lower dynamic load levels than the final static load stage (although it should be mentioned that they were recorded under operating traffic that contained numerous trucks. This section will investigate the response predictions at a lower static load level, the three empty truck load condition. The total load at this stage was 86.7 kips.

In Chapter 6, it was concluded that there was some nonlinearity at the crack in the piercap located adjacent to the southern bearing of Girder #1. This was evident in the strain and displacement data from that girder, particularly at the final load stage. The effect of this nonlinearity is smeared into the models used in this analysis through linear parameters. The displacement predictions for Girder #1 at the final load stage did not capture the effects of this nonlinearity well. In addition, there was some load redistribution between the girders observed at the higher load level.

This section focuses on the individual MCMC chains and the subsequent response predictions at the three empty truck load stage (Figure 11-11). Both displacement and strain predictions are presented.

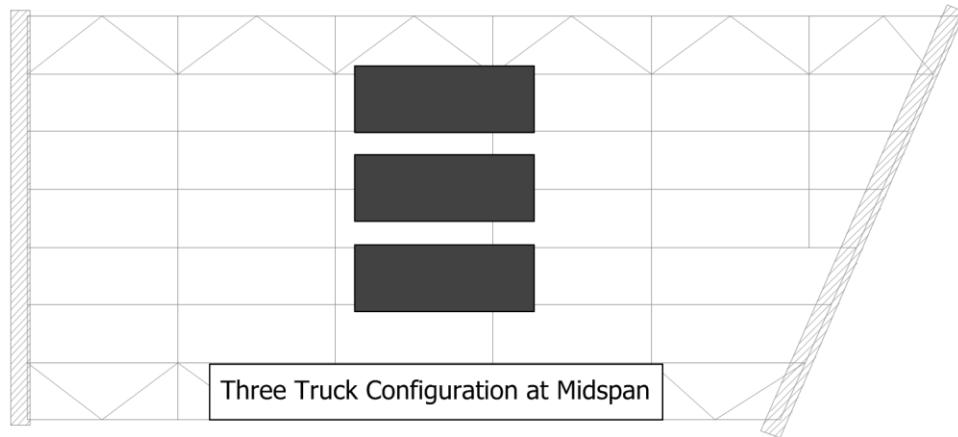


Figure 11-11 - Truck Position for Three Empty Truck Case

11.4.1 MCMC Chain Output – Displacement at Service Limit State

Figure 11-12 shows the displacement predictions at the load stage using three empty trucks. All three displacements along Girder #1 were over-predicted by at least 80%. In terms of magnitude this corresponds to only about a 0.055 in. difference. The trend of over prediction continued for all girders and locations, as seen in Figure 11-13, Figure 11-14, and Figure 11-15.

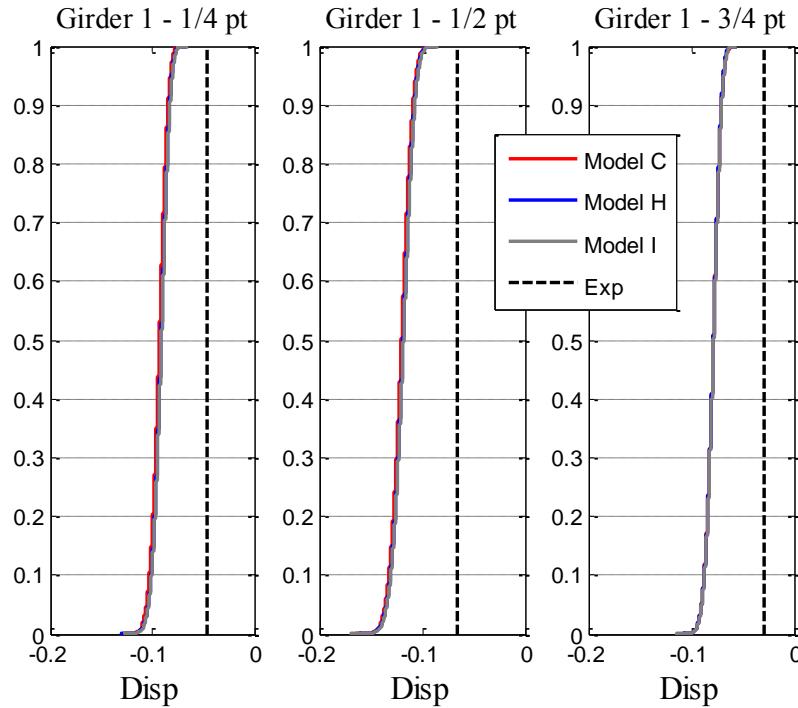


Figure 11-12 – Girder #1 Displacement Predictions – Three Empty Trucks

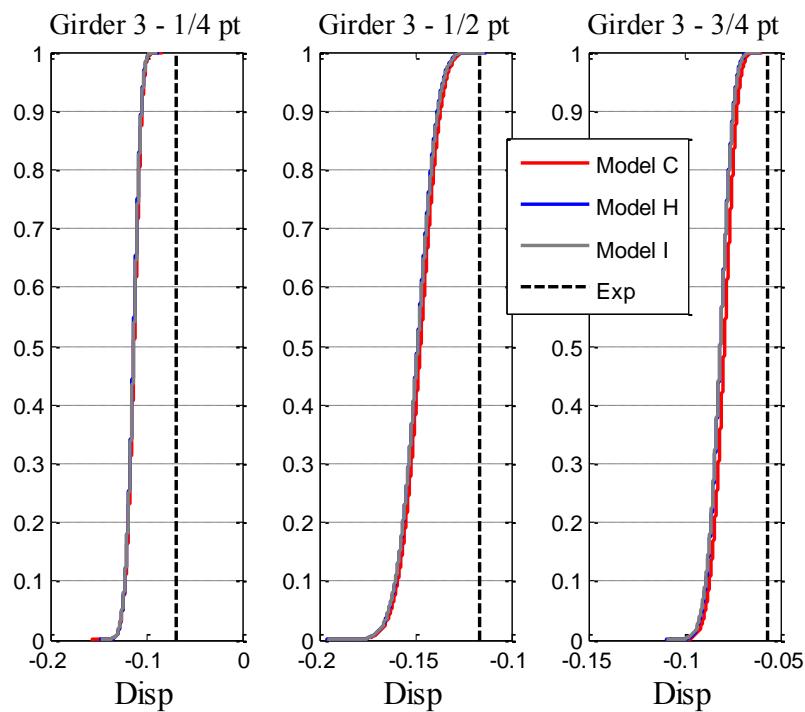


Figure 11-13 - Girder #3 Displacement Predictions - Three Empty Trucks

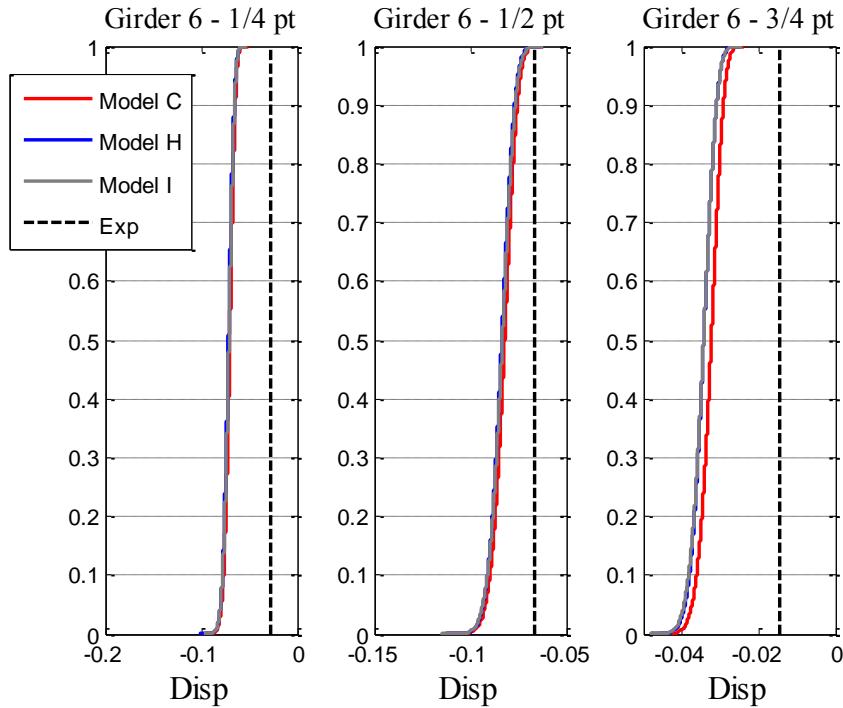


Figure 11-14 - Girder #6 Displacement Predictions - Three Empty Trucks

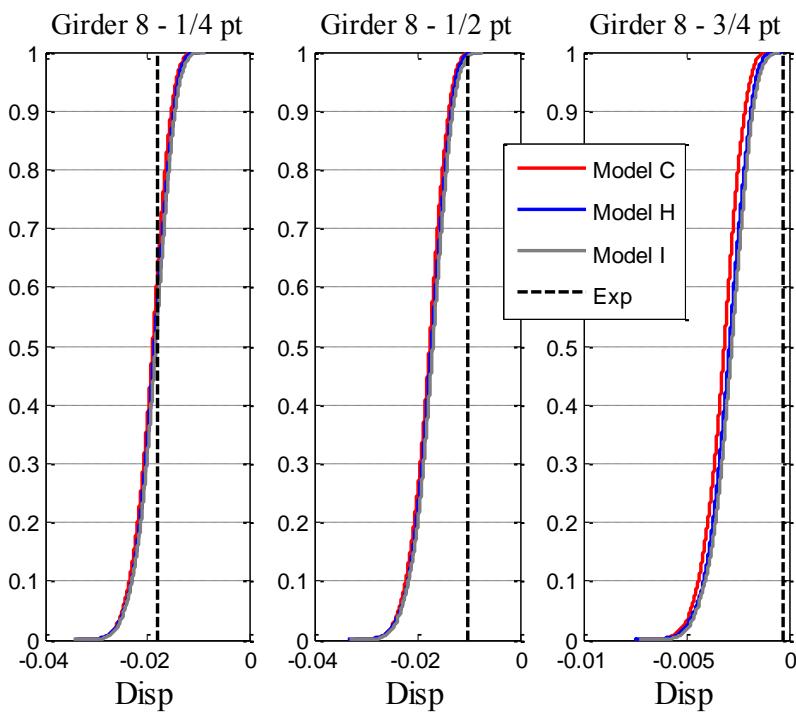


Figure 11-15 - Girder #8 Displacement Predictions - Three Empty Trucks

Table 11-7 shows the mean displacement predictions for the candidate model population under the three empty truck load case. Table 11-8 shows the absolute value of the corresponding percent errors. Clearly the percent errors are much larger for this case, ranging between 30% and 200%, but this is exacerbated by the small magnitudes of response. The actual magnitudes of the discrepancies between measured and analytical predictions of displacement are approximately the same order of magnitude at this load level as they are for the six full trucks load level.

One constant in all of these predictions is the assumed inherent standard deviation of 4%. In terms of displacements, 4% may be conservative for lower load levels when the recorded responses are not much larger than the noise threshold of the sensor. At these types of load levels, strain gauges may be more reliable to use as a prediction quantity, as the relative magnitude of the response is much larger compared to the noise threshold as it is for displacement measurements.

Table 11-7 - Mean Displacement Predictions - Three Empty Trucks

Loc	Girder	Exp (in)	Mod C (in)	Mod H (in)	Mod I (in)
1	1 - 1/4 pt	-0.046	-0.093	-0.092	-0.091
2	1 - 1/2 pt	-0.066	-0.120	-0.119	-0.119
3	1 - 3/4 pt	-0.030	-0.079	-0.079	-0.079
4	3 - 1/4 pt	-0.068	-0.113	-0.114	-0.113
5	3 - 1/2 pt	-0.117	-0.148	-0.149	-0.149
6	3 - 3/4 pt	-0.057	-0.080	-0.081	-0.082
7	6 - 1/4 pt	-0.028	-0.070	-0.072	-0.071
8	6 - 1/2 pt	-0.066	-0.082	-0.084	-0.084
9	6 - 3/4 pt	-0.015	-0.033	-0.034	-0.034
10	8 - 1/4 pt	-0.018	-0.019	-0.019	-0.019
11	8 - 1/2 pt	-0.010	-0.018	-0.017	-0.017
12	8 - 3/4 pt	0.000	-0.003	-0.003	-0.003

Table 11-8 - Mean Displacement Errors - Three Empty Trucks

Loc	Girder	Mod C (% Err)	Mod H (% Err)	Mod I (% Err)	Avg % Err
1	1 - 1/4 pt	101%	97%	97%	98%
2	1 - 1/2 pt	83%	81%	80%	81%
3	1 - 3/4 pt	164%	164%	164%	164%
4	3 - 1/4 pt	66%	67%	67%	67%
5	3 - 1/2 pt	27%	28%	28%	28%
6	3 - 3/4 pt	39%	43%	43%	42%
7	6 - 1/4 pt	153%	157%	156%	155%
8	6 - 1/2 pt	24%	26%	26%	26%
9	6 - 3/4 pt	123%	136%	136%	131%
10	8 - 1/4 pt	6%	4%	3%	5%
11	8 - 1/2 pt	74%	71%	68%	71%
12	8 - 3/4 pt	1004%	916%	876%	932%

11.4.2 MCMC Chain Output – Strain at Service Limit State

Figure 11-16 shows the strain predictions for the bottom flange of Girder #1 under the three empty truck load case. The mean midspan prediction has an error of 14%, but the quarter and three-quarter span locations are within 10%.

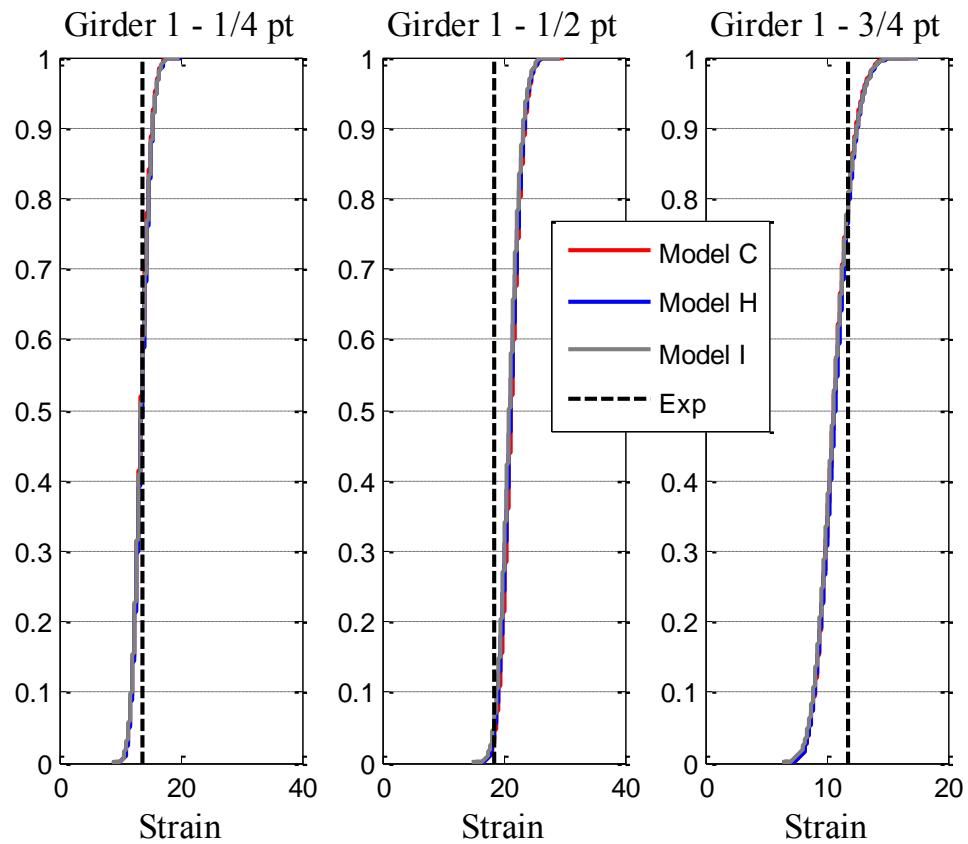


Figure 11-16 - Girder #1 Strain Predictions - Three Empty Trucks

Figure 11-17 shows the strain predictions for Girder #3. In this case, the models slightly over-predicted at midspan, but under-predicted at the quarter and three-quarter points. For Girder #6 (Figure 11-18), the measured value for the three-quarter point was completely independent of the response prediction distribution, and all three locations are under-predicted. The percent errors ranged from 14% to 63%.

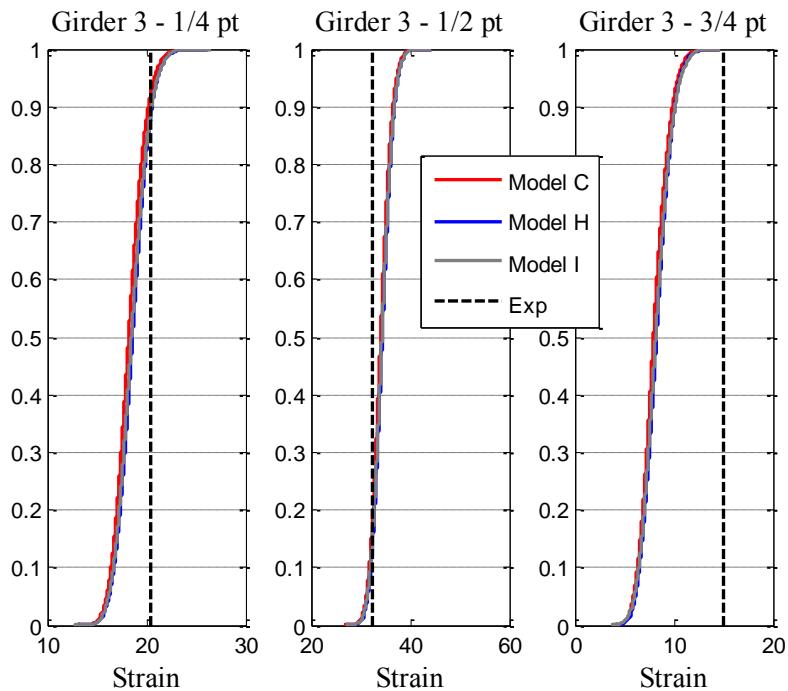


Figure 11-17 - Girder #3 Strain Predictions - Three Empty Trucks

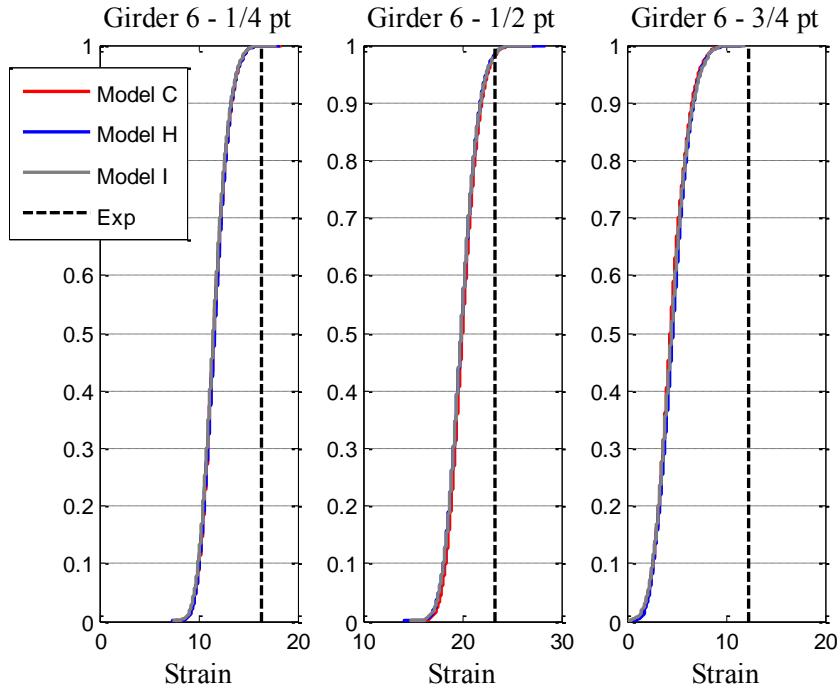


Figure 11-18 - Girder #6 Strain Predictions - Three Empty Trucks

Figure 11-19 shows the predictions for Girder #8. In this case, the majority of the prediction distribution at the three-quarter point showed negative or compressive strains on the bottom flange. This indicates that there was negative bending in that region of the bridge, or large out-of-plane bending that dominated the strain response. A check of the mean moment actions at that location indicates that the mean primary bending moment was -42 k-ft which would put the bottom flange in compression.

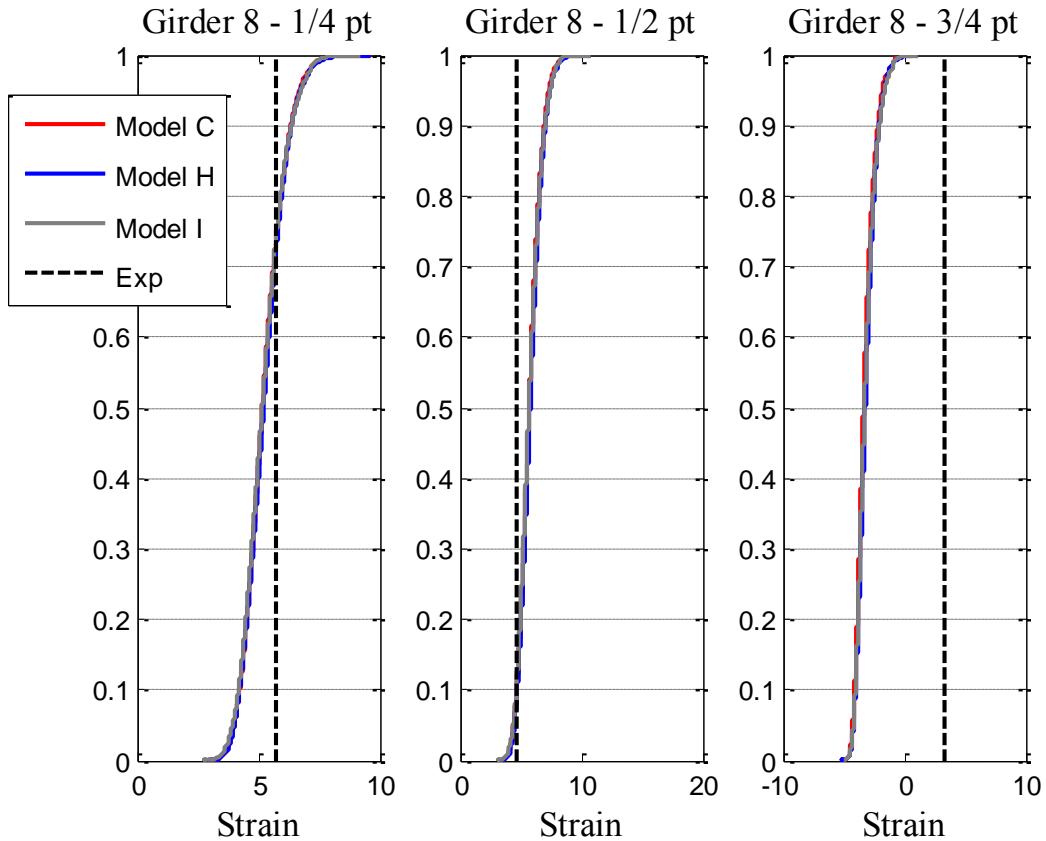


Figure 11-19 - Girder #8 Strain Predictions - Three Empty Trucks

Table 11-9 shows the mean strain predictions from the candidate model population under the three empty truck load case. Table 11-10 shows the corresponding percent errors.

Table 11-9 - Mean Strain Predictions - Three Empty Trucks

Loc	Girder	Exp ($\mu\epsilon$)	Mod C ($\mu\epsilon$)	Mod H ($\mu\epsilon$)	Mod I ($\mu\epsilon$)
1	1 - 1/4 pt	13.8	13.6	13.7	13.6
2	1 - 1/2 pt	18.6	21.3	21.2	21.1
3	1 - 3/4 pt	11.8	10.7	10.7	10.7
4	3 - 1/4 pt	20.4	18.3	18.6	18.5
5	3 - 1/2 pt	32.5	34.1	34.4	34.3
6	3 - 3/4 pt	15.0	8.0	8.3	8.2
7	6 - 1/4 pt	16.4	11.7	11.7	11.6
8	6 - 1/2 pt	23.3	20.1	19.9	20.0
9	6 - 3/4 pt	12.3	4.5	4.7	4.6
10	8 - 1/4 pt	5.8	5.2	5.3	5.2
11	8 - 1/2 pt	4.6	5.8	5.8	5.8
12	8 - 3/4 pt	3.3	-3.2	-3.1	-3.2

Table 11-10 - Mean Strain Errors - Three Empty Trucks

Loc	Girder	Mod C (% Err)	Mod H (% Err)	Mod I (% Err)	Avg % Err
1	1 - 1/4 pt	1%	1%	1%	1%
2	1 - 1/2 pt	15%	14%	14%	14%
3	1 - 3/4 pt	10%	9%	10%	10%
4	3 - 1/4 pt	10%	9%	9%	9%
5	3 - 1/2 pt	5%	6%	6%	6%
6	3 - 3/4 pt	46%	45%	45%	45%
7	6 - 1/4 pt	29%	29%	29%	29%
8	6 - 1/2 pt	14%	14%	14%	14%
9	6 - 3/4 pt	63%	62%	63%	63%
10	8 - 1/4 pt	9%	8%	9%	9%
11	8 - 1/2 pt	25%	26%	25%	25%
12	8 - 3/4 pt	198%	194%	195%	196%

11.4.3 MCMC Chain Output - Comparison of Service and Proof Limit State

The suite of candidate models was used to develop response predictions for twelve displacements and twelve strains under two different load cases. The first load case was six full trucks, corresponding to a total of 460 kips. The second load case was three empty trucks, corresponding to a total of 87 kips.

Table 11-11 shows the average absolute value of magnitude of error across all model forms. If both the structure and the model are completely linear, the magnitude of error would increase correspondingly with the increase in load. The general trend is observed with a few notable exceptions. At Location #1, the one-quarter point of Girder #1, the average magnitude of error between the two load stages increases by nearly four times. At the three-quarter points of Girders #3 and #6, the increase in the magnitude of error was 3.75 and 3.42 times, respectively. Most other locations showed increases of around 1.5. The quarter-span location of Girder #6 had much larger average errors for the Three Truck load case.

The variation in the changes of errors indicates that the discrepancy between model and experiment is not a bias error in the model which is consistent across both load stages. Rather, this indicates that the structure is exhibiting nonlinearity in one or more ways that cannot be replicated by the linear model. The piercap crack has already been identified as a non-linearity, as well as the load redistribution discussed in Section 5.9.1, both of which could influence the magnitudes of errors between load stages.

Table 11-11 – Average Model Prediction Displacement Errors – Magnitude

Loc	Girder	Exp (in) Six Trucks	Exp (in) Three Trucks	Difference (in.) Six Trucks	Difference (in.) Three Trucks
1	1 - 1/4 pt	-0.6246	-0.046	0.178	0.046
2	1 - 1/2 pt	-0.6389	-0.066	0.046	0.054
3	1 - 3/4 pt	-0.467	-0.030	0.066	0.049
4	3 - 1/4 pt	-0.6157	-0.068	0.073	0.045
5	3 - 1/2 pt	-0.8457	-0.117	0.092	0.032
6	3 - 3/4 pt	-0.5298	-0.057	0.089	0.024
7	6 - 1/4 pt	-0.358	-0.028	0.006	0.043
8	6 - 1/2 pt	-0.4466	-0.066	0.017	0.017
9	6 - 3/4 pt	-0.2492	-0.015	0.066	0.019
10	8 - 1/4 pt	-0.0747	-0.018	0.024	0.001
11	8 - 1/2 pt	-0.0793	-0.010	0.012	0.007
12	8 - 3/4 pt	-0.0192	0.000	0.004	0.003

Table 11-12 shows the absolute value of the average prediction error magnitudes for the two load stages. In this case, the expected trend of larger magnitude errors with larger magnitude responses was generally observed. Table 11-12 also clearly indicates the large errors present at the three-quarter point, particularly in Girder #3 and Girder #6 for the six truck case.

Table 11-12 – Average Model Prediction Strain Error - Magnitude

Loc	Girder	Exp ($\mu\epsilon$) Six Trucks	Exp ($\mu\epsilon$) Three Trucks	Difference ($\mu\epsilon$) Six Trucks	Difference ($\mu\epsilon$) Three Trucks
1	1 - 1/4 pt	86.2	13.788	11.420	0.177
2	1 - 1/2 pt	113.1	18.556	8.854	2.660
3	1 - 3/4 pt	83.9	11.835	11.465	1.134
4	3 - 1/4 pt	111.9	20.379	18.364	1.894
5	3 - 1/2 pt	183.6	32.463	1.852	1.829
6	3 - 3/4 pt	118.4	14.973	44.173	6.791
7	6 - 1/4 pt	77.4	16.351	12.111	4.713
8	6 - 1/2 pt	107.7	23.251	5.497	3.259
9	6 - 3/4 pt	70.4	12.286	25.234	7.683
10	8 - 1/4 pt	15.7	5.756	18.875	0.513
11	8 - 1/2 pt	6.2	4.618	31.377	1.163
12	8 - 3/4 pt	5.0	3.308	5.098	0.142

In general, it is difficult to draw conclusions about different load stages when the response predictions are also a variable. Since the chosen observation was modal parameters, and the response predictions were displacements and strains, it was impossible to divorce the uncertainty associated with that translation between response predictions from the translation between load stages. Looking only at percent errors, it is clear that the magnitude of the ambient traffic load observed when determining modal properties is somewhere in between the magnitude of the two static cases. The displacement predictions are generally better at the higher load level which indicates that the ambient load level must have activated a global nonlinearity that was present in the six truck load stage, but not the three truck load stage. The strain predictions are generally better at the lower load level. Recall that at the six truck load level, large errors

in mean strain prediction were observed at the one-quarter and three-quarter points, while the midspan location had relatively small errors. This is likely due to some nonlinearity between the deck and the beam that was not active at the lower load level.

In the future, the MCMC analysis should be informed by observations of strains or displacements, and the predictions should be over the same modality, but at different load levels.

11.5 Summary of Multiple Model Analysis for the IBS Bridge

The following specific conclusions were made in this chapter and are reiterated here.

1. The mean prediction from a model form and the prediction from the model with the maximum likelihood are not the same.
2. Despite the inclusion of the vertical stiffness parameter at Girder #1, the model forms were not able to capture that behavior as it did not manifest itself within the observed experimental responses (i.e., the modal parameters).
3. At the locations investigated (i.e., the measurement locations) the predictions responses across all model forms were similar, further supporting the conclusion from Chapter 10 that the models are essentially the same.
4. The static load cases used for the comparison did not create enough response at Girder #8 to draw meaningful conclusions about the parameters and their association with the behavior of that girder.
5. It is difficult to draw conclusions about the response predictions at different static load stages from the models developed using modal observations. The modal

properties observed came from ambient vibration results, meaning that the applied load level was on average somewhere in between a light static load level and the large load level seen during the six truck case, which fully activated the non-linearity of the piercap crack. In addition, during the ambient vibration monitoring, the load was transient in nature, whereas during the truck test the loading was largely static (except for the lane which was open to traffic).

6. Overall, using modal properties to predict other responses like strains and displacements is a difficult prospect considering the difference in load magnitude between dynamic and static loading, and the subsequent activation of non-nonlinear mechanisms at higher load levels.

Chapter 12: Comparison of Traditional and Multiple Model St-Id

This chapter closes the loop between the two phases of this research. Phase 1 focused on the documentation of a best practices application of St-Id, including a comprehensive model-experiment correlation. Phase 2 focused on investigating and improving multiple model techniques for model-experiment correlation as part of the St-Id framework considering the recommendations regarding Multiple Model St-Id presented in Dubbs (2012). In this chapter, comparisons between response predictions from the traditional model-experiment correlation presented in Chapter 6 and the multiple model analysis in Chapter 11 are presented. A discussion of the influence of parameter interaction and interdependency is provided. Finally a true unmeasureable quantity, rating factor, is compared between traditional and multiple model approaches.

12.1 Comparison of MCMC Response Predictions to Single Model Predictions

This section presents a comparison of traditional single model predictions to the multiple model prediction distributions presented in Chapter 11, as well as the experimental predictions. The calibrated model that was updated with modal properties during the original manual calibration effort was compared. The comparison included the modal properties of the model and the response predictions for strains and displacements. This model represents the best effort for single model calibration using heuristics and parameter identification. At the time of construction of the this model, it was decided that modal properties would not be informed by the composite action parameters that were believed to be a result of the large static loading, so a single, smeared composite action

parameter was used. This corresponds directly to Models A, D, and E in the MCMC analysis, all of which had a single composite action parameter and were subsequently eliminated from the full MCMC chain development stage.

The other parameters include barrier and parapet stiffness, vertical bearing stiffness on the southern edge (split into Girder #1 and the remaining Girders), and longitudinal stiffness on the northern edge of the span. Otherwise, the nominal properties for everything else were exactly the same as the models used through Chapter 10 and Chapter 11.

12.1.1 Frequency Response Comparisons

Figure 12-1 through Figure 12-7 shows the distributions of each frequency for the MCMC chains, the experimental value, and the single model output from the calibration using modal properties presented in Chapter 6. Note that the experimental frequency is shown as a vertical line representing the mean and a CDF generated using the assumed 4% standard deviation. Typically, the mean of the distribution from the MCMC chain is a better representation of the observed data than the single model output. In the case of Frequency #6, the values are nearly the same, and in the case of Frequency #3, the single model is slightly better. In no case are any of the predictions outside of the distribution of the experimental data assuming the 4% standard deviation.

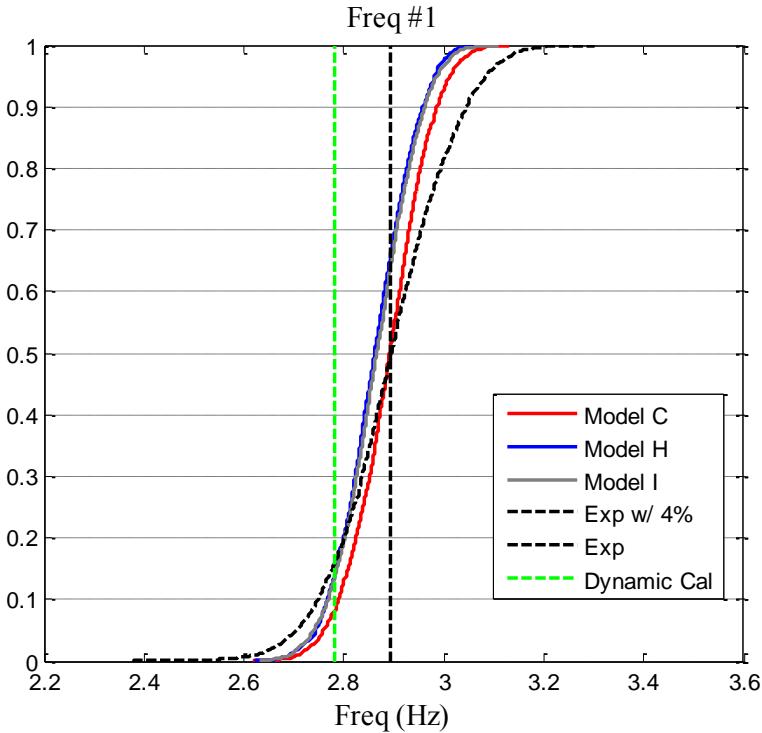


Figure 12-1 - Comparison of MCMC and Single Model to Experimental Frequency #1

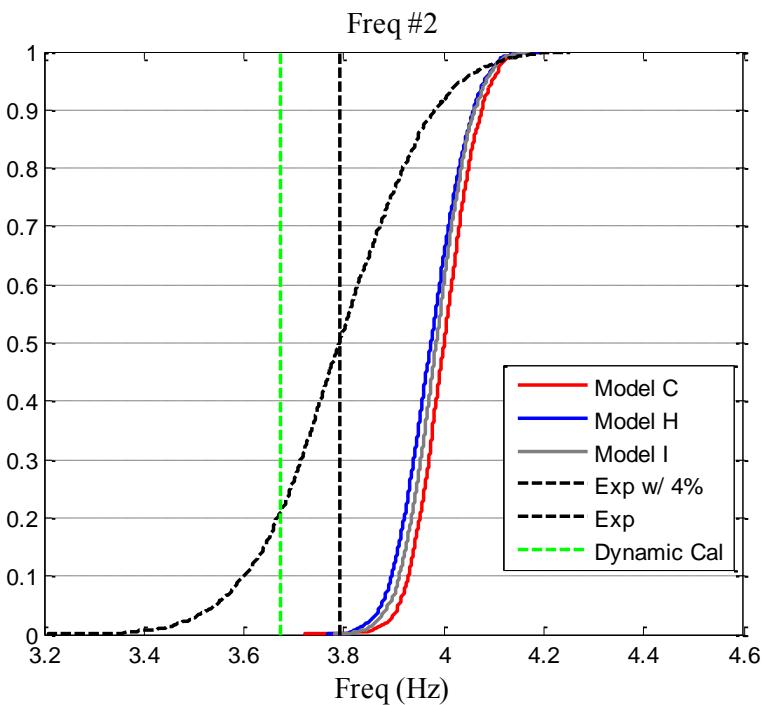


Figure 12-2 - Comparison of MCMC and Single Model to Experimental Frequency #2

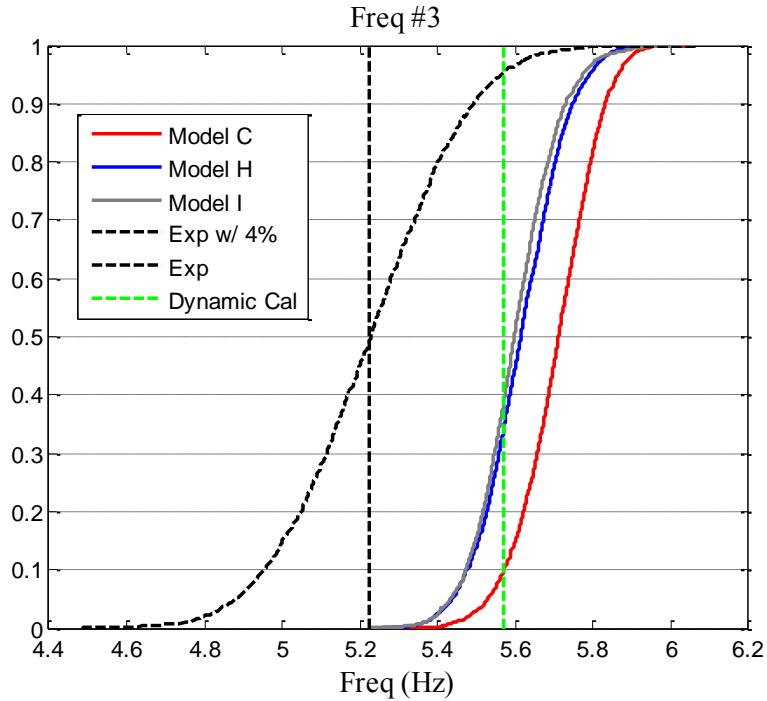


Figure 12-3 - Comparison of MCMC and Single Model to Experimental Frequency #3

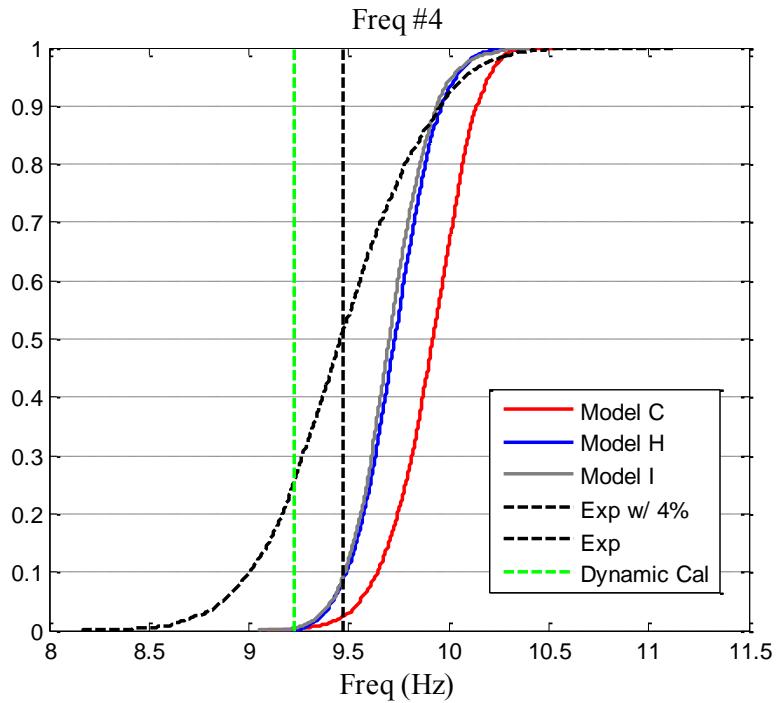


Figure 12-4 - Comparison of MCMC and Single Model to Experimental Frequency #4

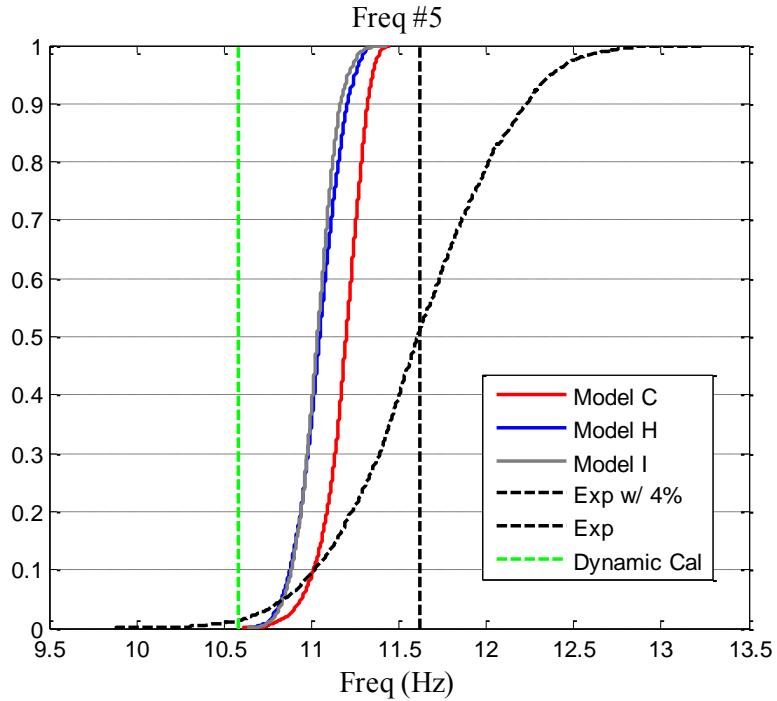


Figure 12-5 - Comparison of MCMC and Single Model to Experimental Frequency #5

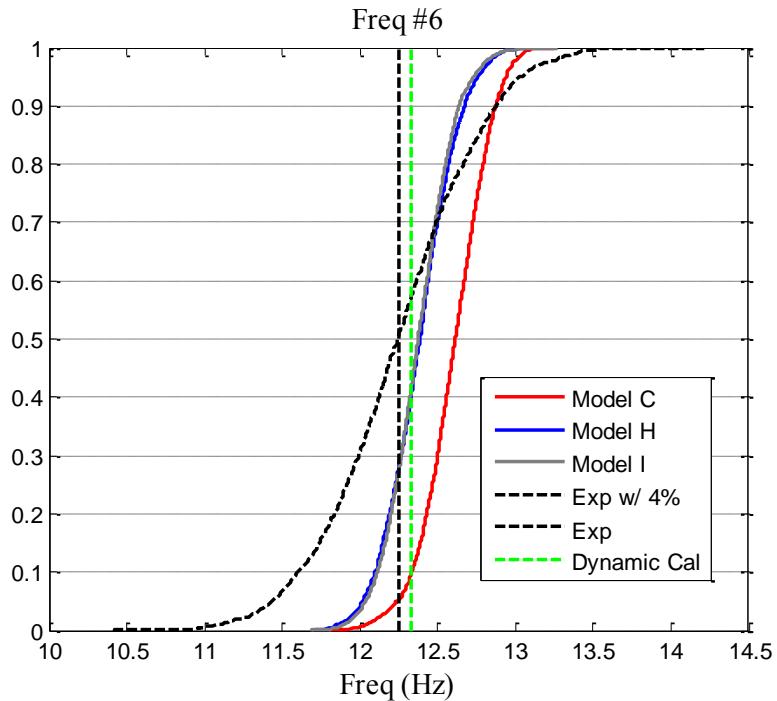


Figure 12-6 - Comparison of MCMC and Single Model to Experimental Frequency #6

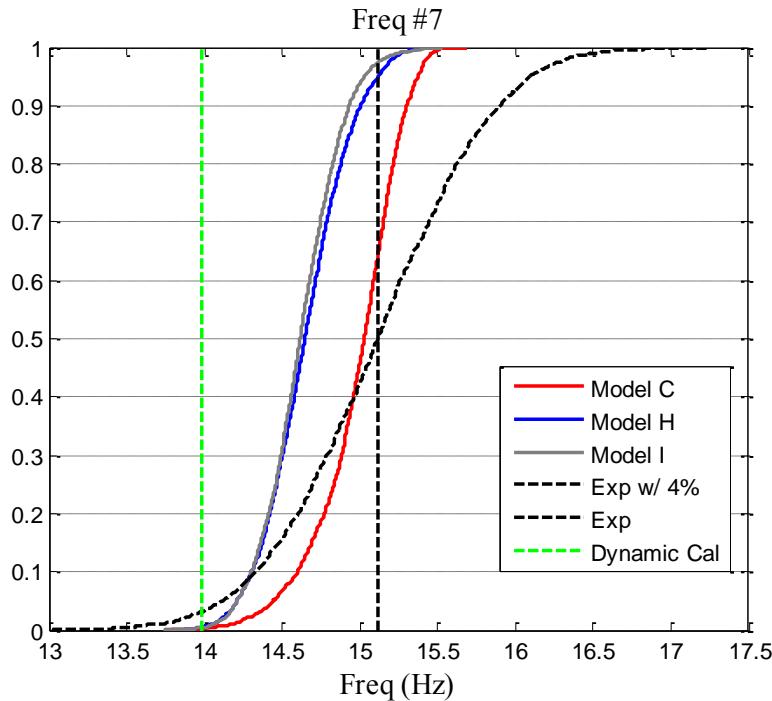


Figure 12-7 - Comparison of MCMC and Single Model to Experimental Frequency #7

Table 12-1 shows the frequency output from the single model updated with modal properties, the mean of the MCMC frequency distributions, the experimental frequencies, and the respective absolute value of percent errors.

Table 12-1 - Frequency Comparison - Single Model and MCMC Results

Mode	Exp (Hz)	Mod C (Hz)	Mod C (% Err)	Mod H (Hz)	Mod H (% Err)	Mod I (Hz)	Mod I (% Err)	Single Model	Single Model (% Err)
1	2.9	2.73	5.61%	2.74	5.27%	2.74	5.37%	2.78	3.78%
2	3.8	3.75	1.13%	3.81	0.40%	3.79	0.19%	3.67	3.19%
3	5.2	5.55	6.21%	5.51	5.37%	5.49	5.09%	5.57	6.54%
4	9.5	9.63	1.65%	9.57	1.00%	9.54	0.74%	9.23	2.59%
5	11.6	11.13	4.19%	11.03	5.08%	11.02	5.15%	10.58	8.92%
6	12.2	12.38	1.08%	12.25	0.05%	12.23	0.13%	12.33	0.63%
7	15.1	14.55	3.75%	14.43	4.56%	14.38	4.88%	13.98	7.51%

As expected, the percent errors for frequency for both methods are similar. Both the single model optimization and the MCMC analysis used model properties in the updating process, so they should be able to reasonable predict these responses.

12.1.2 Displacement Prediction Comparisons

Figure 12-8 through Figure 12-11 show the distributions of displacement predictions for each girder for the MCMC chains, the experimental values, and the single model predictions from the calibration using modal properties, presented in Chapter 6. The general trend is that the displacement predictions from the single model are typically larger than the measured experimental value and the MCMC prediction distributions.

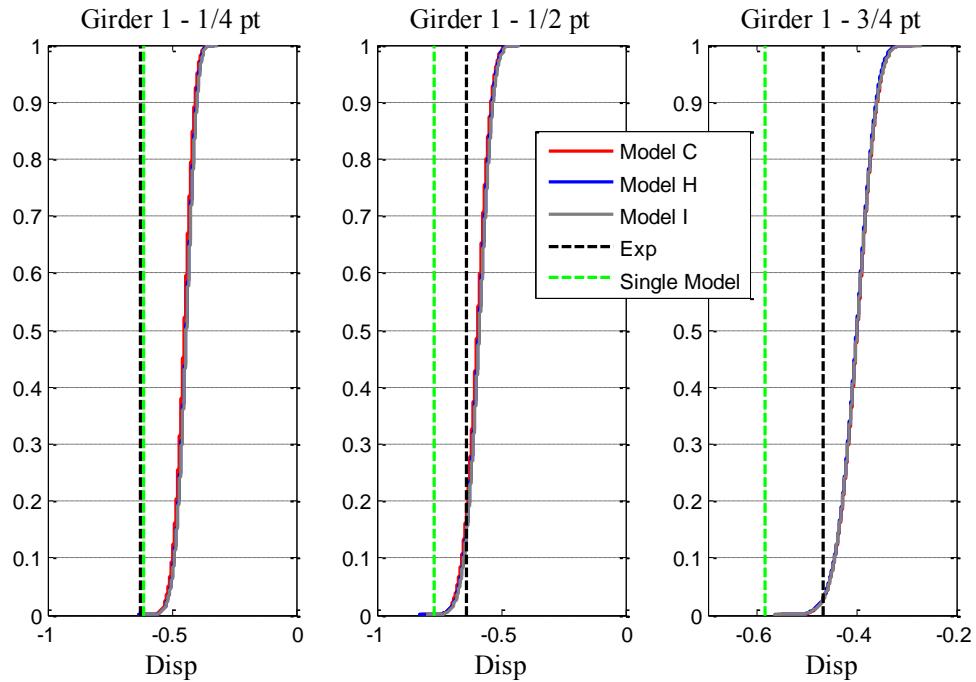


Figure 12-8 - Comparison of Girder #1 Displacements

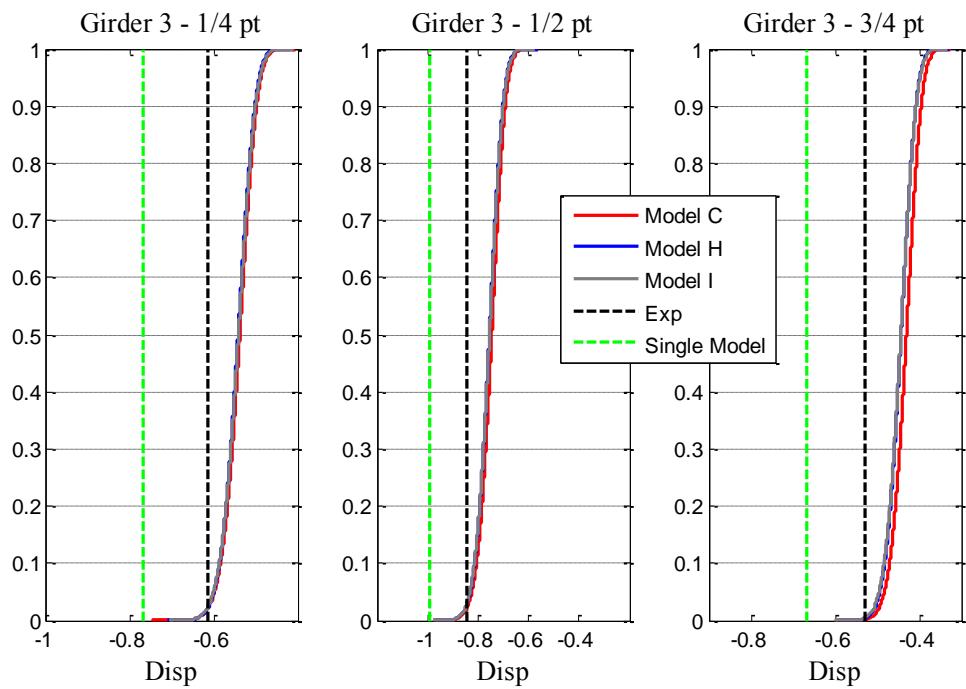


Figure 12-9 - Comparison of Girder #3 Displacements

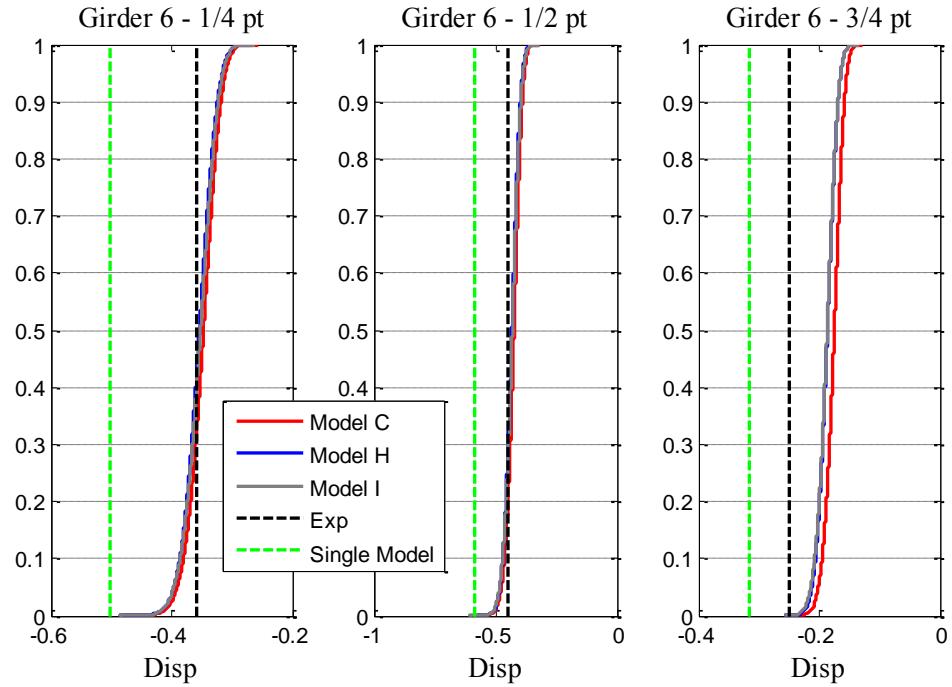


Figure 12-10 - Comparison of Girder #6 Displacements

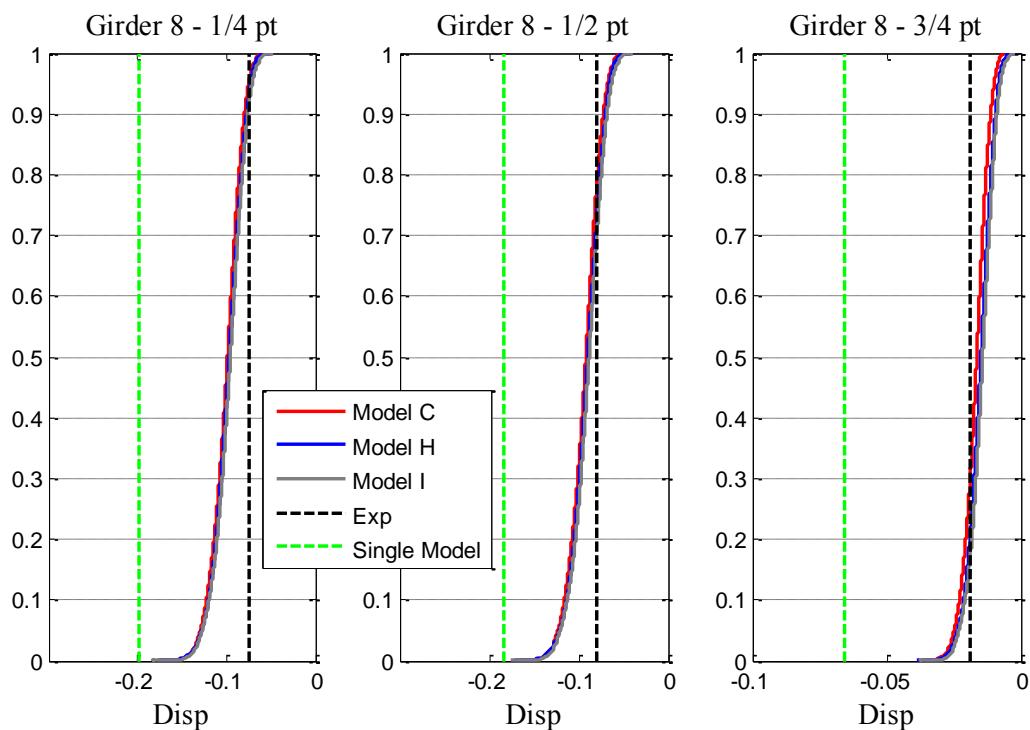


Figure 12-11 - Comparison of Girder #8 Displacements

Table 12-2 shows the mean displacement predictions from each of the MCMC chains, and the single model displacement predictions. Table 12-3 shows the percent errors from each of these models. The single model predictions tend to have a more uniform distribution of error, with the exception of Girder #8, where the magnitude of the responses magnifies the percent errors. Table 12-4 shows the actual magnitudes of errors in displacement.

Table 12-2 - Mean MCMC and Single Model Displacement Predictions

Loc	Girder	Exp (in)	Mod C (in)	Mod H (in)	Mod I (in)	Single Model (in)
1	1 - 1/4 pt	-0.625	-0.452	-0.444	-0.443	-0.616
2	1 - 1/2 pt	-0.639	-0.597	-0.591	-0.590	-0.772
3	1 - 3/4 pt	-0.467	-0.401	-0.402	-0.401	-0.586
4	3 - 1/4 pt	-0.616	-0.540	-0.544	-0.543	-0.768
5	3 - 1/2 pt	-0.846	-0.748	-0.756	-0.756	-0.993
6	3 - 3/4 pt	-0.530	-0.432	-0.444	-0.445	-0.668
7	6 - 1/4 pt	-0.358	-0.348	-0.355	-0.353	-0.503
8	6 - 1/2 pt	-0.447	-0.425	-0.433	-0.432	-0.585
9	6 - 3/4 pt	-0.249	-0.176	-0.187	-0.187	-0.317
10	8 - 1/4 pt	-0.075	-0.100	-0.098	-0.097	-0.198
11	8 - 1/2 pt	-0.079	-0.093	-0.091	-0.089	-0.183
12	8 - 3/4 pt	-0.019	-0.017	-0.015	-0.015	-0.066

Table 12-3 - Percent Error from Mean MCMC and Single Model Displacement Predictions

Loc	Girder	Mod C (% Err)	Mod H (% Err)	Mod I (% Err)	Single Model (% Err)
1	1 - 1/4 pt	28%	29%	29%	1%
2	1 - 1/2 pt	6%	7%	8%	21%
3	1 - 3/4 pt	14%	14%	14%	26%
4	3 - 1/4 pt	12%	12%	12%	25%
5	3 - 1/2 pt	12%	11%	11%	17%
6	3 - 3/4 pt	18%	16%	16%	26%
7	6 - 1/4 pt	3%	1%	1%	41%
8	6 - 1/2 pt	5%	3%	3%	31%
9	6 - 3/4 pt	29%	25%	25%	27%
10	8 - 1/4 pt	34%	31%	30%	165%
11	8 - 1/2 pt	17%	15%	13%	131%
12	8 - 3/4 pt	12%	21%	24%	243%

Table 12-4 - Magnitude of Displacement Errors

Loc	Exp (in)	Mod C (in)	Mod H (in)	Mod I (in)	Single Model (in)
1	-0.62462	0.173	0.180	0.181	0.008
2	-0.63891	0.041	0.048	0.049	-0.133
3	-0.46695	0.066	0.065	0.066	-0.119
4	-0.61568	0.076	0.072	0.073	-0.152
5	-0.84571	0.098	0.090	0.090	-0.147
6	-0.52977	0.098	0.086	0.085	-0.139
7	-0.35797	0.010	0.003	0.005	-0.145
8	-0.44657	0.021	0.014	0.014	-0.139
9	-0.24921	0.073	0.062	0.062	-0.067
10	-0.07469	-0.025	-0.023	-0.022	-0.123
11	-0.07932	-0.014	-0.012	-0.010	-0.104
12	-0.0192	0.002	0.004	0.005	-0.047

12.1.3 Strain Prediction Comparisons

Figure 12-12 through Figure 12-15 show the distributions of strain predictions for each Girder for the MCMC chains, the experimental values, and the single model predictions from the calibration using modal properties, presented in Chapter 6. Again the errors in the single model prediction are all from the result of predictions that are larger than the experimental values. It is interesting to note that at least for Girder #1 and #3, the errors are larger at midspan than at the one-quarter and three-quarter points.

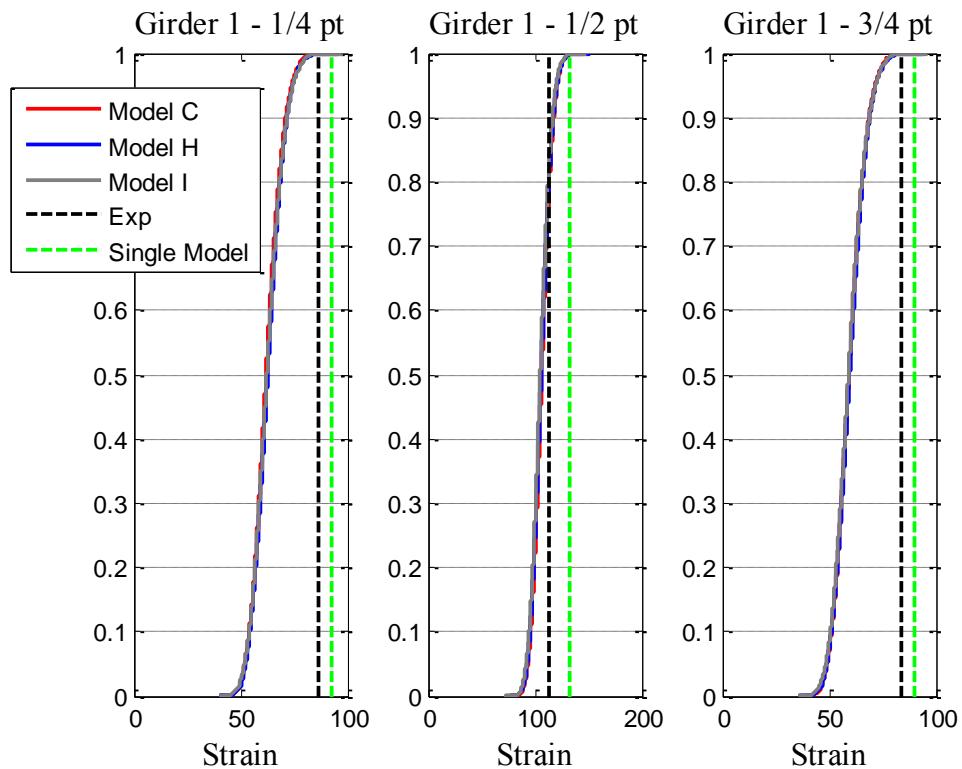


Figure 12-12 - Comparison of Girder #1 Strains

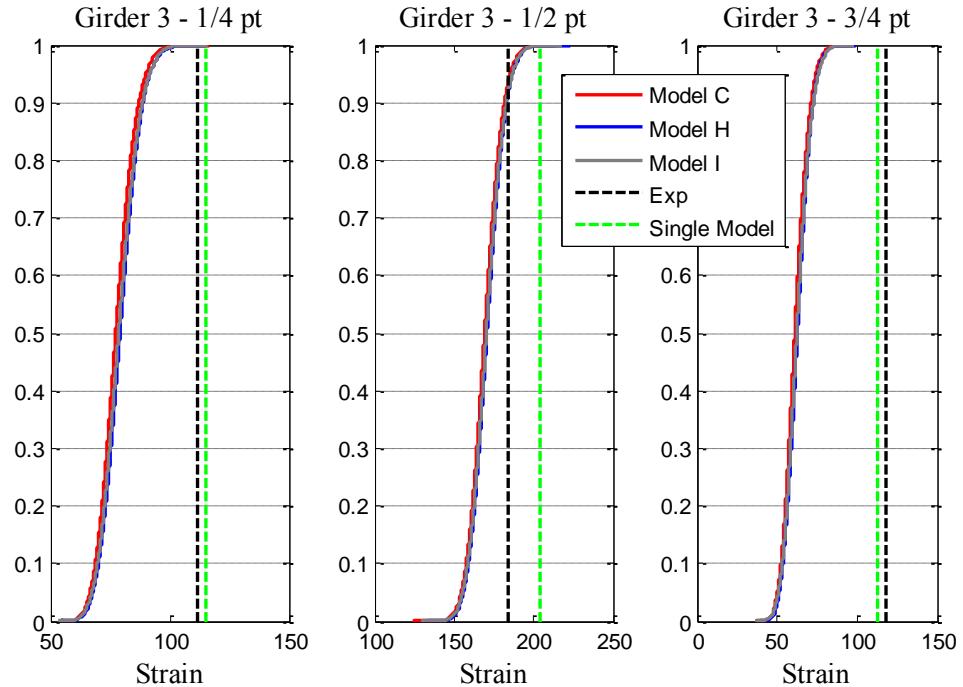


Figure 12-13 - Comparison of Girder #3 Strains

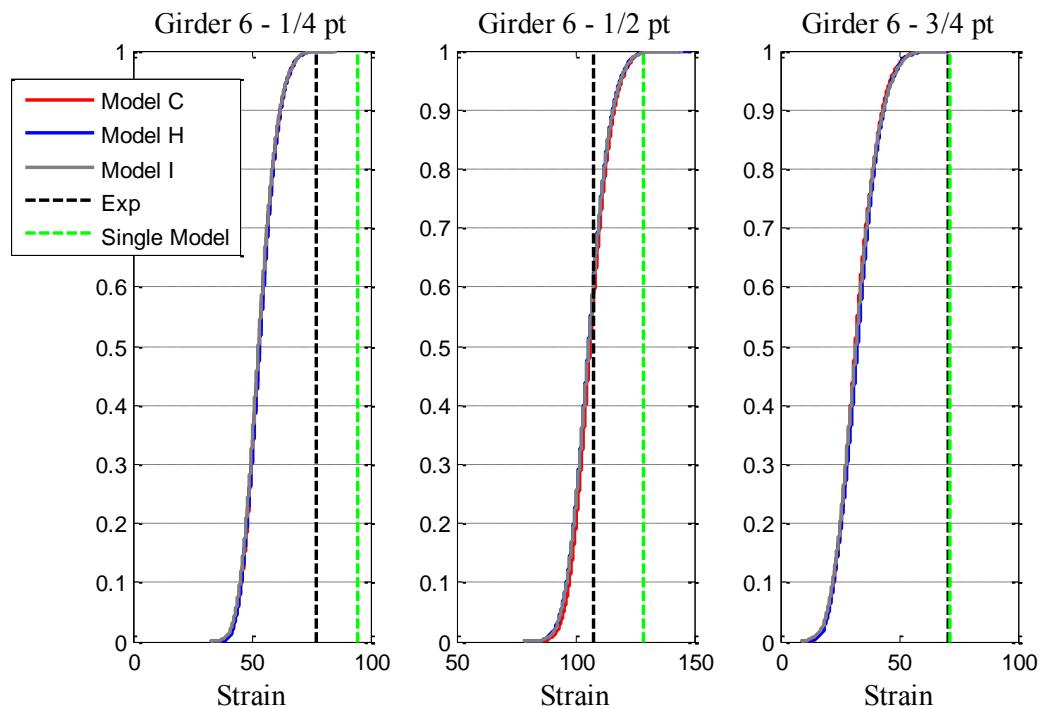


Figure 12-14 - Comparison of Girder #6 Strains

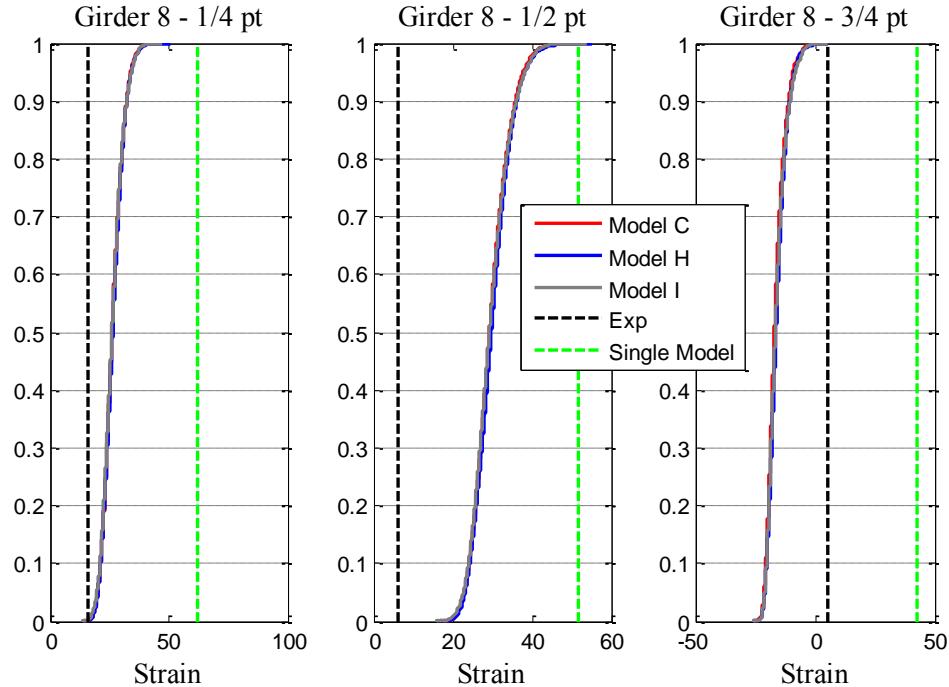


Figure 12-15 - Comparison of Girder #8 Strains

Table 12-5 shows the mean predictions of strain from the MCMC chains and the predictions from the single model dynamic calibration. Table 12-6 shows the percent errors for the strain predictions. In general, the strains are all too large, but with the exception of midspan locations, are generally more accurate than the MCMC mean predictions. Girder #8, per usual, has very inaccurate predictions.

Table 12-5 - Mean MCMC and Single Model Strain Predictions

Loc	Exp ($\mu\epsilon$)	Mod C ($\mu\epsilon$)	Mod H ($\mu\epsilon$)	Mod I ($\mu\epsilon$)	Single Model ($\mu\epsilon$)
1	86.2	62.3	62.9	62.6	92.7
2	113.1	106.6	106.2	105.4	131.7
3	83.9	59.6	59.9	59.4	90.1
4	111.9	77.7	79.3	78.9	115.2
5	183.6	169.5	170.9	170.6	204.3
6	118.4	61.6	62.9	62.6	112.8
7	77.4	53.7	53.7	53.3	94.3
8	107.7	106.6	105.9	105.9	128.6
9	70.4	32.3	32.9	32.4	71.1
10	15.7	26.5	26.6	26.4	62.2
11	6.2	29.7	30.1	29.7	51.9
12	5.0	-16.3	-15.8	-16.0	42.5

Table 12-6 - Percent Error from Mean MCMC and Single Model Strain Predictions

Loc	Mod C (% Err)	Mod H (% Err)	Mod I (% Err)	Single Model (% Err)
1	28%	27%	27%	7%
2	6%	6%	7%	16%
3	29%	29%	29%	7%
4	31%	29%	29%	3%
5	8%	7%	7%	11%
6	48%	47%	47%	5%
7	31%	31%	31%	22%
8	1%	2%	2%	19%
9	54%	53%	54%	1%
10	69%	70%	68%	297%
11	382%	388%	382%	742%
12	427%	417%	420%	753%

12.2 Assessment of Parameter Interaction

One of the main benefits of multiple model methods is the ability to assess the influence of parameters both on the structure and the corresponding responses, but also on one another. Parameter interaction is very difficult to comprehensively understand in traditional single model methods, and is often completely ignored. Figure 12-16 shows a matrix plot of the parameter chain from Model H, plotted against itself. The diagonal terms of this visual matrix show histograms of each parameter, while the off-diagonal terms are scatter plots of parameter plotted against one another. This format makes drawing conclusions about inter-parameter relationships straightforward.

Note that in the case of Model H, Parameter 5 and 6 were identical, hence the strong linear correlation. The other pairs where correlation can be readily seen are shown in red. These correlations prevent sampling across the entire model space. The plots highlighted in black are affiliated with Parameter 8, which will be discussed in more detail.

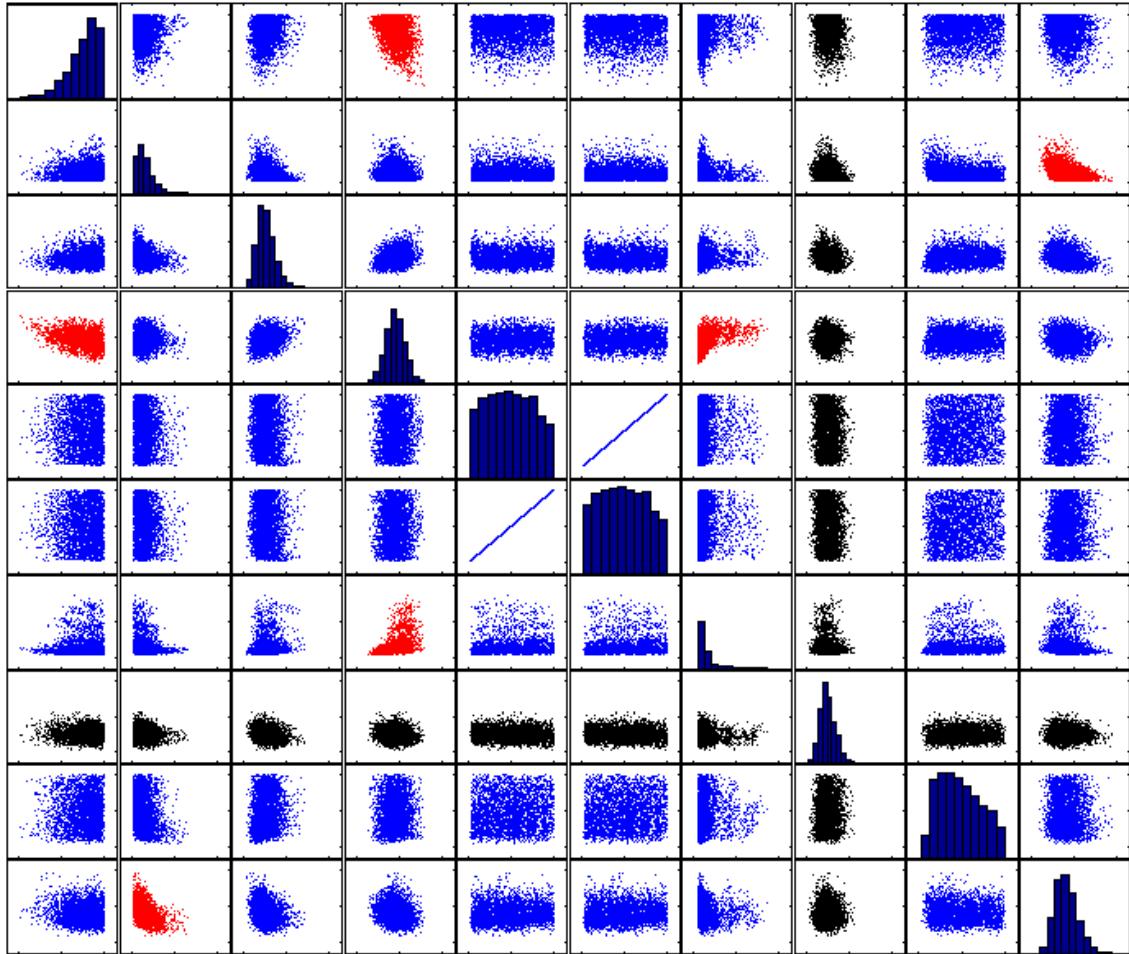


Figure 12-16 - Matrix Plot of Chain for Model H

Another way to investigate this is through correlation coefficients, which is a numerical representation of what is shown visually above. Correlation coefficients are scaled from -1 to 1, with -1 indicating a negative correlation and 1 indicating a positive correlation. Table 12-7 shows the correlation coefficients for the same Model H chain. The largest positive correlation is 0.44, and it is between the modulus of the diaphragms, and the vertical bearing stiffness of the northern bearings. The strongest negative correlation,

between parapet stiffness (Parameter 10) and composite action (Parameter 2) is roughly the same magnitude.

Table 12-7 - Correlation Coefficients - Model H

Parameter	1	2	3	4	5	6	7	8	9	10
1	0.00	0.04	0.08	-0.36	0.02	0.02	0.06	-0.02	-0.03	-0.12
2	0.04	0.00	-0.24	-0.02	-0.08	-0.08	-0.13	-0.08	-0.26	-0.41
3	0.08	-0.24	0.00	0.29	-0.04	-0.04	0.10	-0.23	0.07	-0.26
4	-0.36	-0.02	0.29	0.00	0.06	0.06	0.44	-0.13	-0.02	-0.13
5	0.02	-0.08	-0.04	0.06	0.00	0.00	0.06	-0.06	0.01	0.03
6	0.02	-0.08	-0.04	0.06	0.00	0.00	0.06	-0.06	0.01	0.03
7	0.06	-0.13	0.10	0.44	0.06	0.06	0.00	-0.09	-0.04	-0.07
8	-0.02	-0.08	-0.23	-0.13	-0.06	-0.06	-0.09	0.00	0.12	-0.03
9	-0.03	-0.26	0.07	-0.02	0.01	0.01	-0.04	0.12	0.00	-0.06
10	-0.12	-0.41	-0.26	-0.13	0.03	0.03	-0.07	-0.03	-0.06	0.00

Figure 12-17 shows a matrix plot of the chain for Model H versus the displacement predictions.

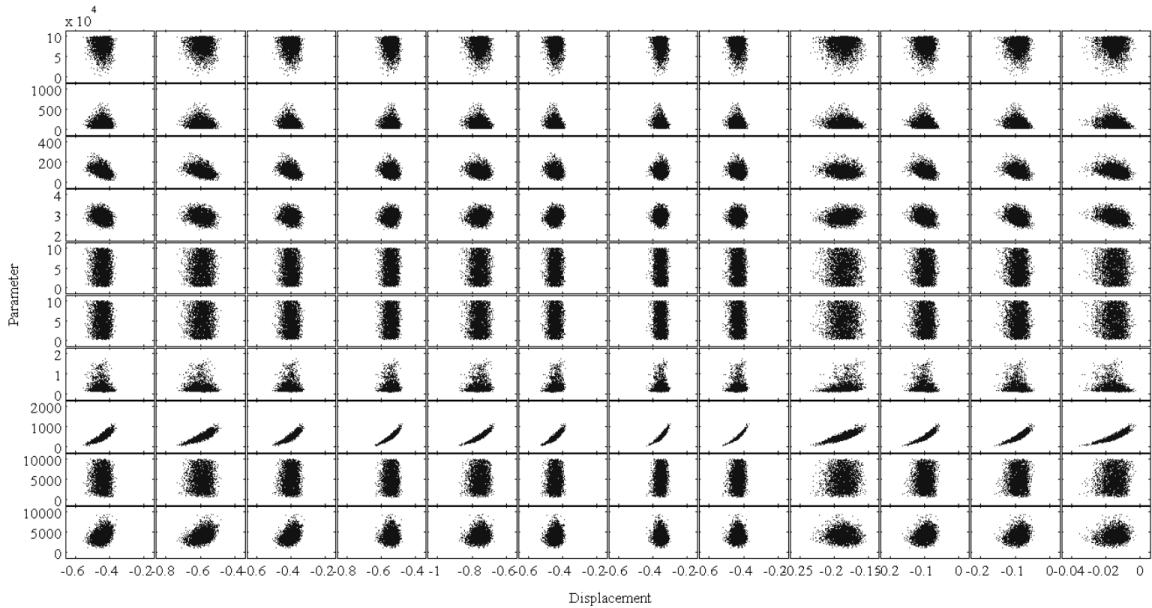


Figure 12-17 - Parameter versus Displacement Matrix Plot

It is immediately clear that Parameter 8 strongly influences all of the displacement predictions. Parameter 8 is the longitudinal stiffness of the expansion bearings on the northern side which would be expected to have a large effect on the vertical displacement. One would also expect that the vertical bearing stiffness would have a large effect on vertical displacement, but since the sampling occurred in the higher portion of the range, the portion of the total bridge displacement that was a direct result of the bearing springs was relatively small.

Figure 12-18 shows matrix plot of parameters versus strains. Again Parameter 8 seems to have the strongest correlation with the response predictions.

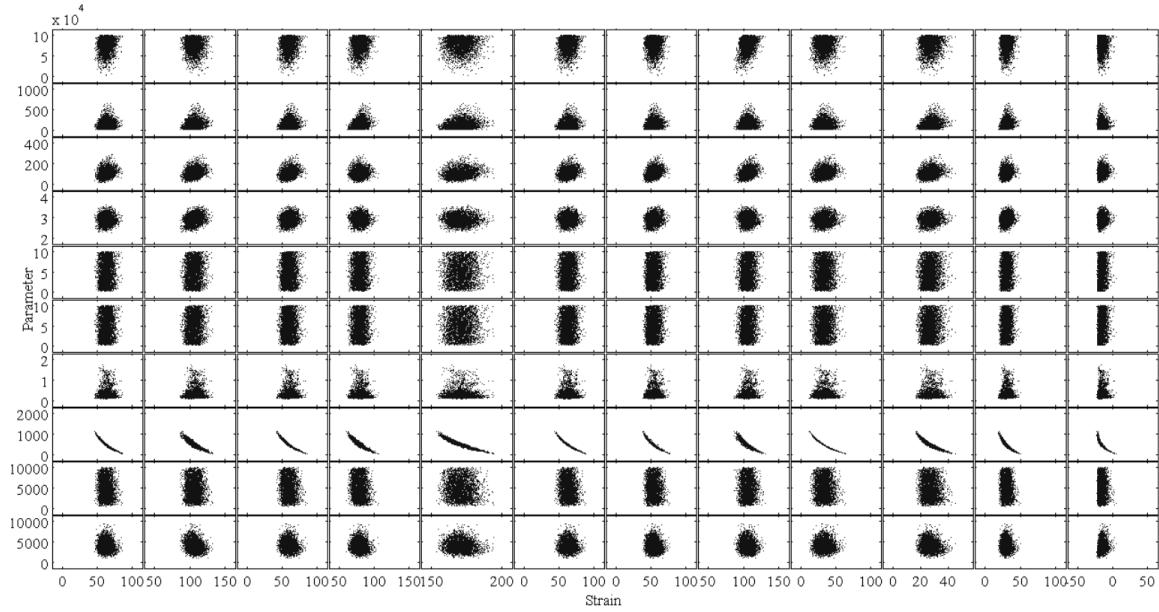


Figure 12-18 – Parameter versus Strain Matrix Plot

12.2.1 Investigation of the Effects of Longitudinal Bearing Stiffness

For the ensuing investigation, only the displacement at midspan of Girder #3 will be considered. Figure 12-19 shows the experimental value, single model prediction, and MCMC distribution of deflection from Model H at midspan of Girder #3. Model H predicted a fairly wide range of displacements, from -0.55 to -0.95 in., reflecting the uncertainty in the prediction of this response using modal observations. The single model prediction was -0.99 in. and the actual experimental value was -0.83 in.

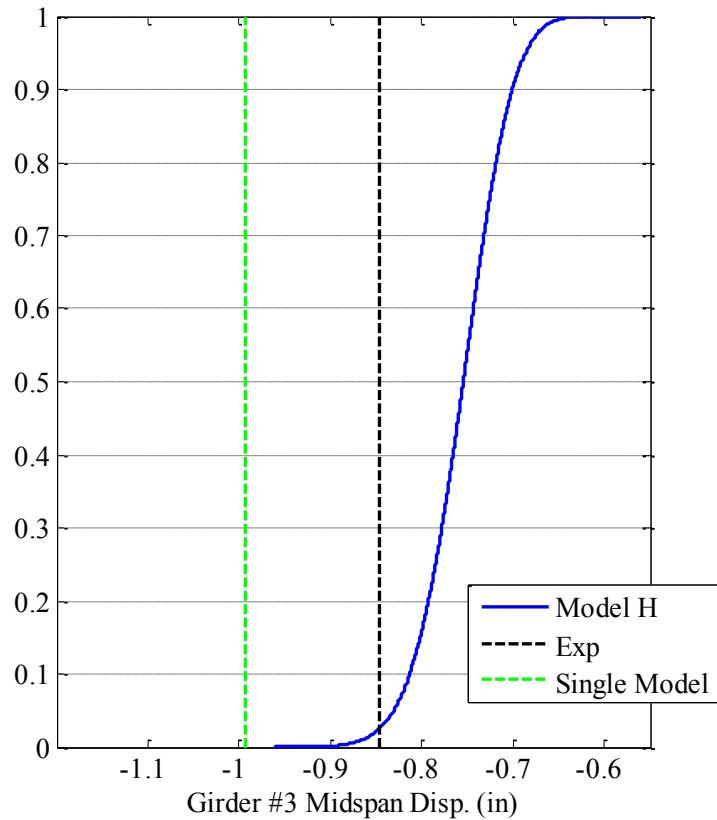


Figure 12-19 - Midspan Displacement of Girder #3

Figure 12-17 and Figure 12-18 show that Parameter 8 seems to have a large, similar influence on the all of the response predictions. Figure 12-20 shows the correlation of Parameter 8 with the frequencies. Only the first frequency is strongly correlated, meaning that while this parameter does influence the model observations to a degree, much of this influence would be drowned out in the calculation of the likelihood, which is a product several numbers ranging from 0 to 1. Similarly, for the single model, the objective function minimizes the total residual error, which can bastardize a single response in order to achieve a slightly lower total residual error.

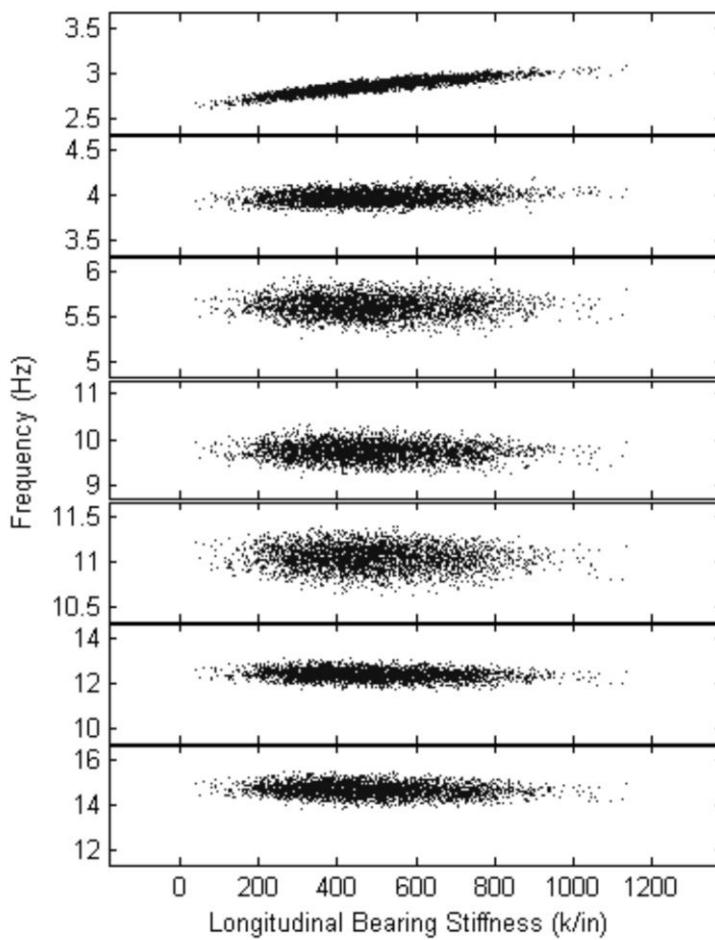


Figure 12-20 - Frequency versus Parameter 8 - Model H

Figure 12-21 shows a similar plot, but with the MAC values. In this case, Parameter 8 affects the second mode the most, and is correlated with the third mode as well. Interestingly, these two modes are correlated in opposite directions (i.e., when a change in the parameter increases the MAC in mode 2, it decreases the MAC in mode 3 and vice versa). So in summary, of the fourteen terms in the calculation of the likelihood, Parameter 8 has a large effect on three of them, and two appear to cancel one another out.

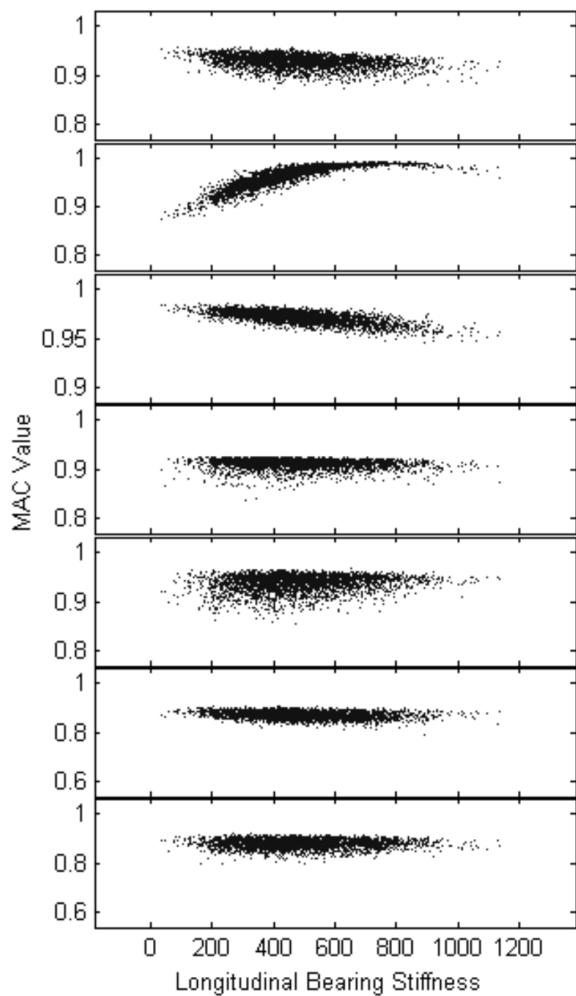


Figure 12-21 - MAC Value versus Parameter 8 - Model H

Figure 12-22 shows the histogram of Parameter 8 for the accepted models from the MCMC chain of Model H. Using the modal observations, with an assigned prior parameter distribution that was uniform over a range of 0 to 10000 kip/in, the MCMC sampling resulting in a posterior parameter distribution that is essentially normal over a range of 0 to 1200 kip/in. This is a marked improvement over the full uniform

distribution, but considering the sensitivity of the response predictions to this parameter, the range still results in the highly variable predictions shown in Figure 12-19.

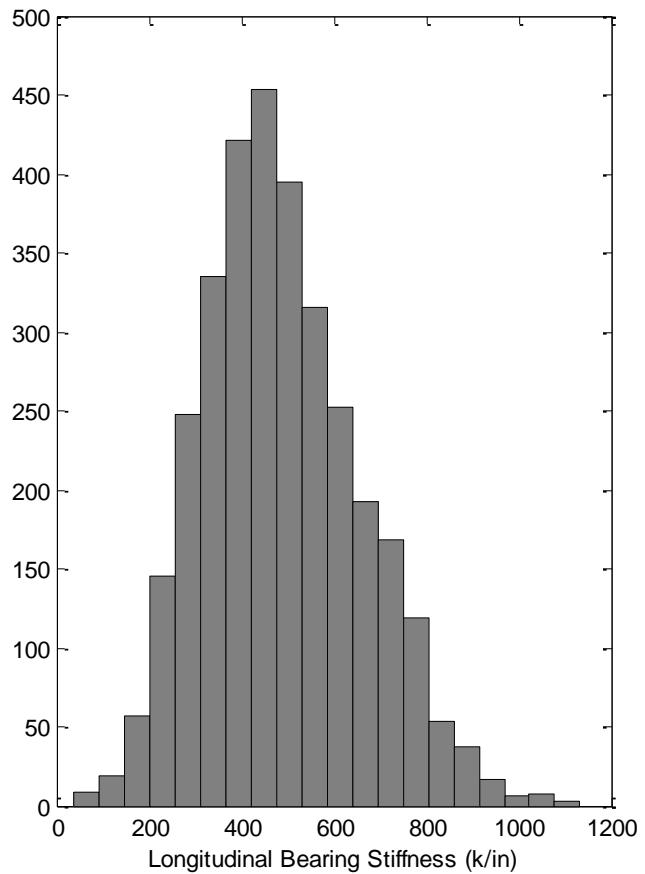


Figure 12-22 - Histogram of Parameter 8 - Model H

Figure 12-23 is the scatter plot of midspan displacement of Girder #3 versus Parameter 8, and the single model value. Considering the parameter value, the MCMC sampling did sample values approximately the same as what was used in the single model. However, the corresponding displacements from the same parameter values in the Model H were 30% smaller than the single model prediction.

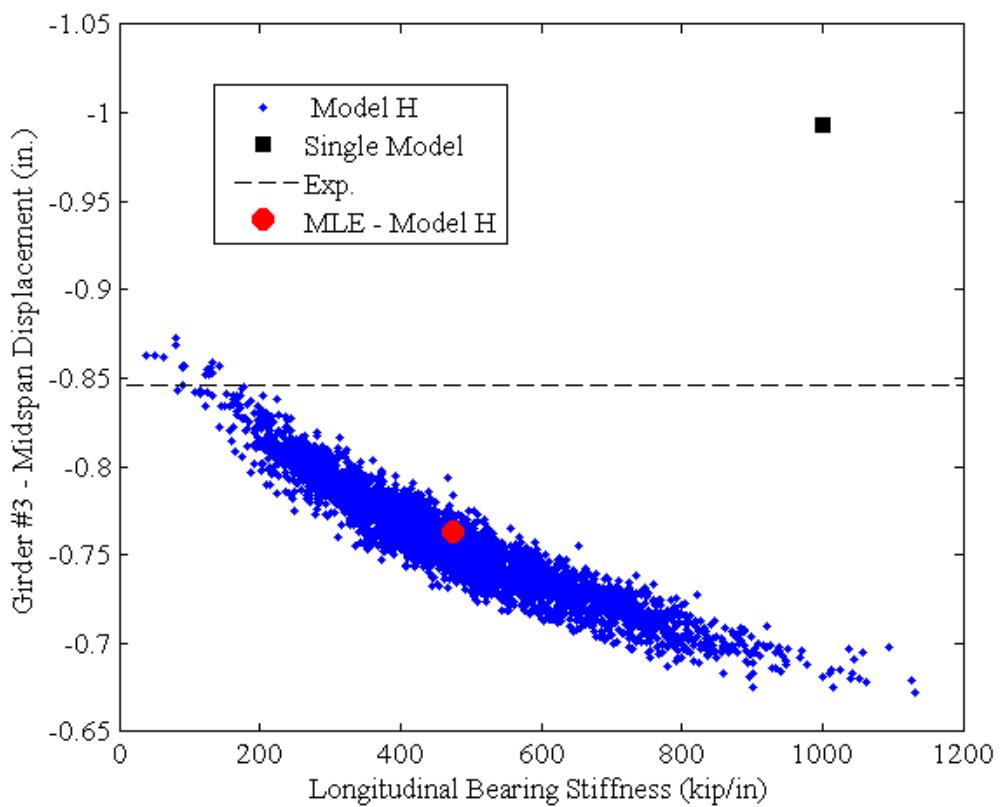


Figure 12-23 - Parameter 8 versus Midspan Displacement - Girder #3

In the single model optimization, the bounds for this parameter were set from 0 to 1000 kip/in. and the final value was pinned to the bound. The bounds were determined through sensitivity studies holding the other parameters constant, considering the range where the modal properties (i.e., the selected observations) were sensitive. After this parameter was pinned to the bounds, the other parameters were optimized and a solution was found.

This comparison is perhaps somewhat unfair, as it the single model optimization was unstable using the same parameter distributions as Model H. However, given the nature of the MCMC chain it was possible to extract a single model prediction as represented by the maximum likelihood from the chain. This model prediction is shown as the single dot in the heart of the scatter plot from Model H. While the MLE is close to the mean of the distribution, the bounds on both the parameter value and the response prediction are substantial, and the discrepancy between the MLE and the experiment value is nearly 10%. This clearly highlights the inherent danger in single model approaches which may provide seemingly good estimates that are not necessarily reliable. The multiple model approach accounts for the uncertainty in parameters and the ability of the observations to inform predictions inherently providing more robust predictions.

There is currently no equivalent for the single model approach, but one could consider rerunning sensitivity studies on both the observations and the response predictions after the optimization of the single model to the chosen observation. This could potentially indicate that the optimized result was in an area where the observations were not very sensitive and the responses predictions were, like the situation above. This would provide insight to the engineer regarding the optimized single model results that would be otherwise unavailable.

12.3 Sample Load Rating and Factor of Safety

One of the most important and readily understandable quantities for bridge condition evaluation is the rating factors. Rating factors (RF) were introduced in Chapter 2. In Chapter 3, preliminary load ratings for the IBS Bridge were calculated according to AASHTO (2008). In Chapter 6, RFs for the six truck static load case were presented for the a priori, the manually updated, and automatically optimized models. In this chapter, a basic load rating that uses a generalized version of the rating equation from AASHTO, but not a code-based rating, is presented. The purpose of this rating is to serve as a frame of reference for how multiple model methods could be incorporated into rating procedures. The equation used is shown below:

$$RF = \frac{Capacity - Dead\ Load\ Demand}{Live\ Load\ Demand} \quad \text{Eqn. 12-1}$$

In addition to RF, Factors of Safety (FS) will be calculated as well. FS is also not a code-based quantity, but it is perhaps more intuitive and readily understandable than a rating factor. The equation for FS is:

$$FS = \frac{Capacity}{Dead\ Load\ Demand + Live\ Load\ Demand} \quad \text{Eqn. 12-2}$$

12.3.1 Sample Rating Truck Load and Position

This scenario used for the calculation of RF and FS will be a single 72 kip truck with wheel loads and spacing shown in Figure 12-24. The spacing and loading is based on the trucks used during the load test, not on a code-specified truck. This truck was used because the concentrated rear axle loads amplify responses at their location.

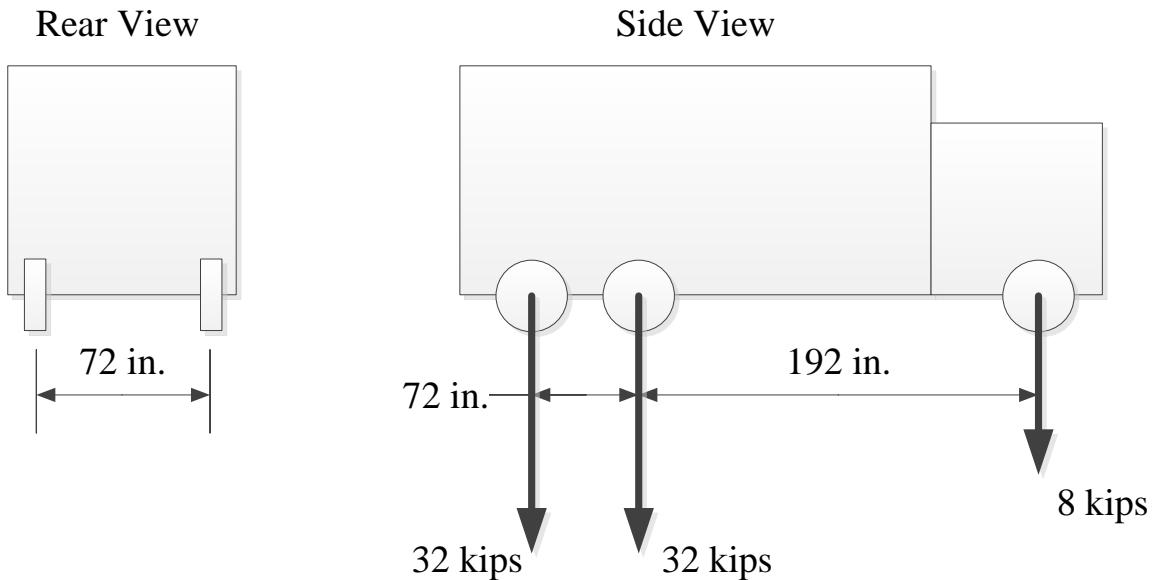


Figure 12-24 - Selected Rating Truck

The truck was placed in the center of the second lane from the west edge of the bridge (i.e., from the long girder) with the centroid of the two rear axles located 65 ft. from the south end of the span. The truck will be facing south, in the direction of traffic. This

position is chosen to center the rear axles over the location where maximum strain and displacement were recorded during the test on Girder #3. Note that this was not the rating case used in Chapter 6, and therefore the RF and FS for those manually updated models were developed with the new load case here.

12.3.2 Nominal Capacity and Dead Load Demand

The capacity and dead load demands were taken from a nominal model. Specifically, this means that any changes in material properties and boundary conditions did not factor into the dead load demand. The nominal model for dead load demand had pin-roller boundary conditions and non-composite properties. The dead load demand was not factored and is applicable for rating factor calculations for all model forms, as it was determined from a nominal model. Table 12-8 lists the dead load demand for each location. The highlighted row refers to the location presented here.

Table 12-8 - Dead Load Demand (from nominal model)

Location	Girder	Dead Load Demand (ksi)
1	1 - 1/4 pt	8.7
2	1 - 1/2 pt	11.8
3	1 - 3/4 pt	9.0
4	3 - 1/4 pt	8.2
5	3 - 1/2 pt	10.7
6	3 - 3/4 pt	7.0
7	6 - 1/4 pt	7.9
8	6 - 1/2 pt	9.4
9	6 - 3/4 pt	5.8
10	8 - 1/4 pt	7.2
11	8 - 1/2 pt	8.1
12	8 - 3/4 pt	5.0

The example rating factor is based on stress, so capacity is assumed to be the yield stress of 36 ksi. A factor of 0.55 was applied to the capacity. Table 12-9 lists the capacity used for the rating calculations.

Table 12-9 - Capacity for Ratings

	Capacity (ksi)
Unfactored	36
Factored	19.8
Determined From:	Drawings

The model or models were used to estimate live load demand. Member actions were extracted from the model and stress at the bottom flange was calculated based on these actions. The RF and FS presented are for stress in Girder #3 at 65 ft. from the south pier.

It should be noted that since the variability associated with dead load demand and capacity is ignored (i.e., both are assumed to be deterministic) the variability of both the RF and FS presented in the following sections should be considered a lower bound.

12.3.3 RF for Midspan Stress – Girder #3

This section presents the RFs for the midspan location of Girder #3 of the IBS Bridge. The RFs are presented for the individual MCMC chains for the three candidate models presented in Chapter 11. Figure 12-25 shows the CDF of RF for these models. The range of possible rating factors spans from 6.8 to 10.5 with a mean of approximately 8.3. This rating factor obviously indicates that there is not a capacity issue with this structure, which had been stated previously in this thesis. However, the range of possible rating factors, if scaled to a lower mean, is indicative of the benefits of multiple models for rating factors.

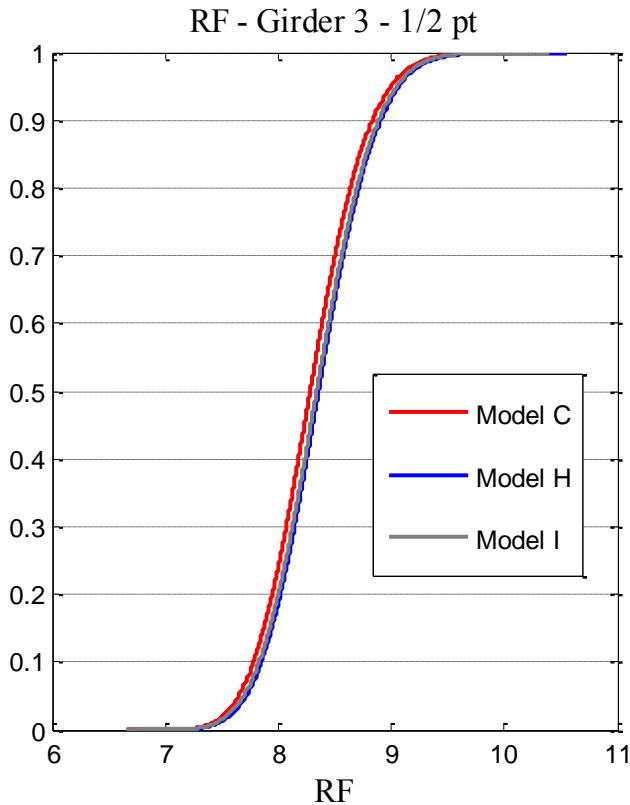


Figure 12-25 - RF for Stress at Midspan of Girder #3

12.3.4 FS for Midspan Stress – Girder #3

The FS for stress at the midspan of Girder #3 is shown in Figure 12-26. In contrast to the RF, the bounds for the FS are much tighter. The FS ranges from 1.86 to 1.98, indicating that the total load including live and dead load is using just over 50% of the available capacity. This criteria tends to be more stable than RF, as it is not so heavily dependent on the magnitude of the live load demand. In addition, it is not as sensitive to the variability of live load demand since the live load and dead load are summed together when computing this metric.

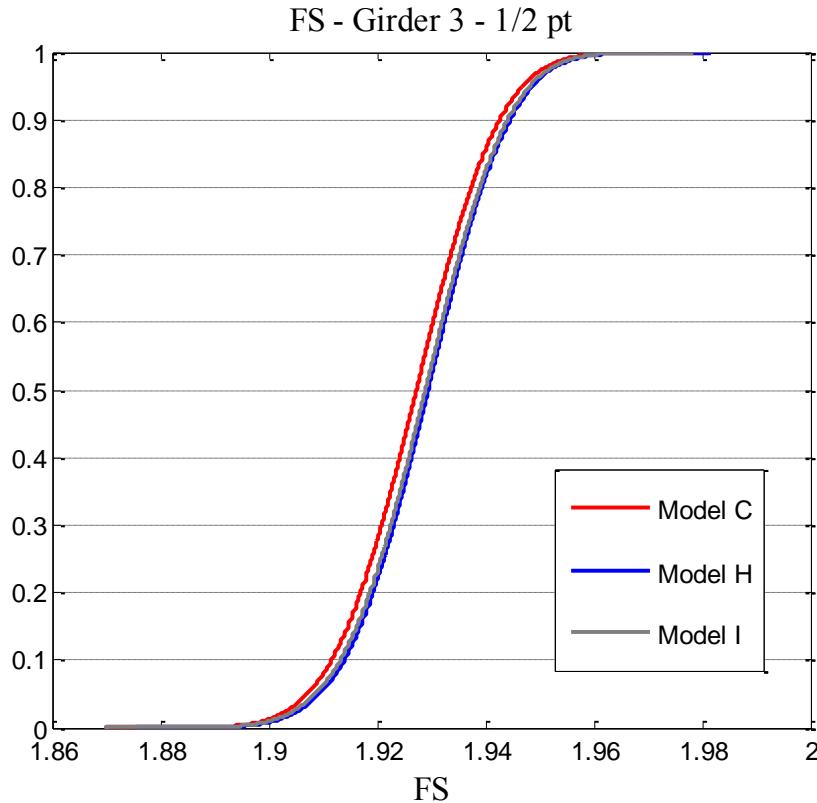


Figure 12-26 - Factor of Safety for Stress at Midspan of Girder #3

12.3.5 Comparison of RF and FS to Single Model Prediction

Recall that in Chapter 6, a rigorous model calibration was conducted using manual calibration and parameter identification. The resulting model from these analyses was reanalyzed using the rating case. Figure 12-27 shows the RF distributions from the three candidate model forms, and the RF from the original dynamic parameter identification. The RF from the original dynamic calibration is 7.38, and the mean of the distributions is 8.37.

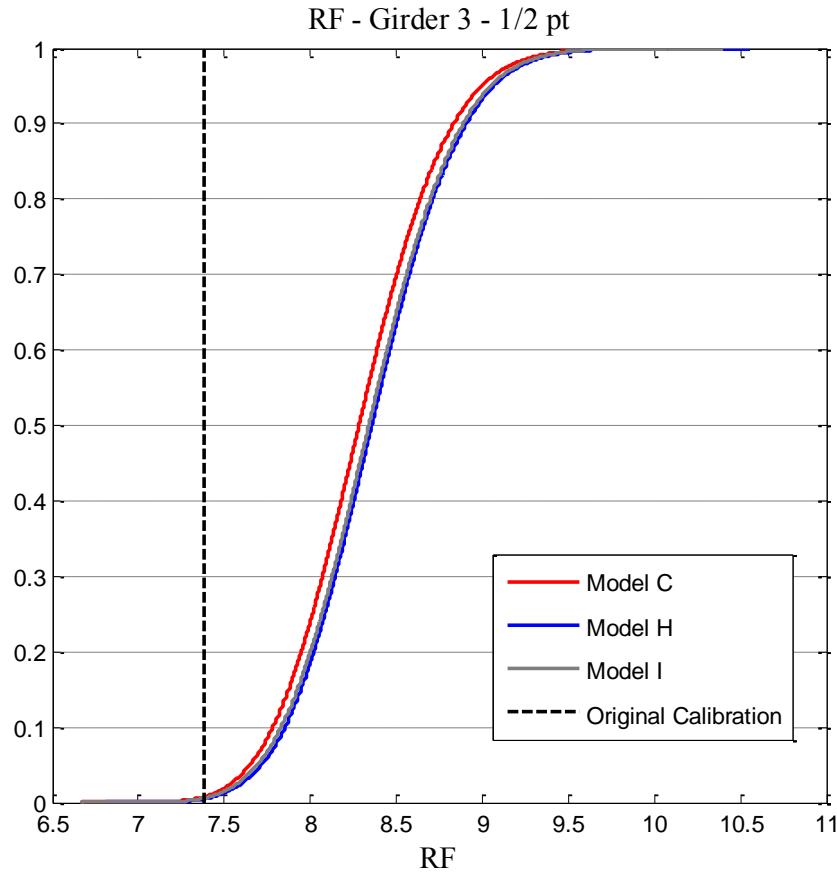


Figure 12-27 - Rating Factor Comparison

From the perspective of a bridge owner, a 7.38 provides less confidence than a distribution with a mean of 8.37 over a range of 6.75 to 9.5. Again the magnitudes of these particular RFs mitigate this effect, but the same scenario could be applied to RFs closer to 1, where an understanding of the variability associated with the predicted RF is far more important.

Similarly, the FS from the MCMC chains, and the original dynamic calibration, is shown in Figure 12-28.

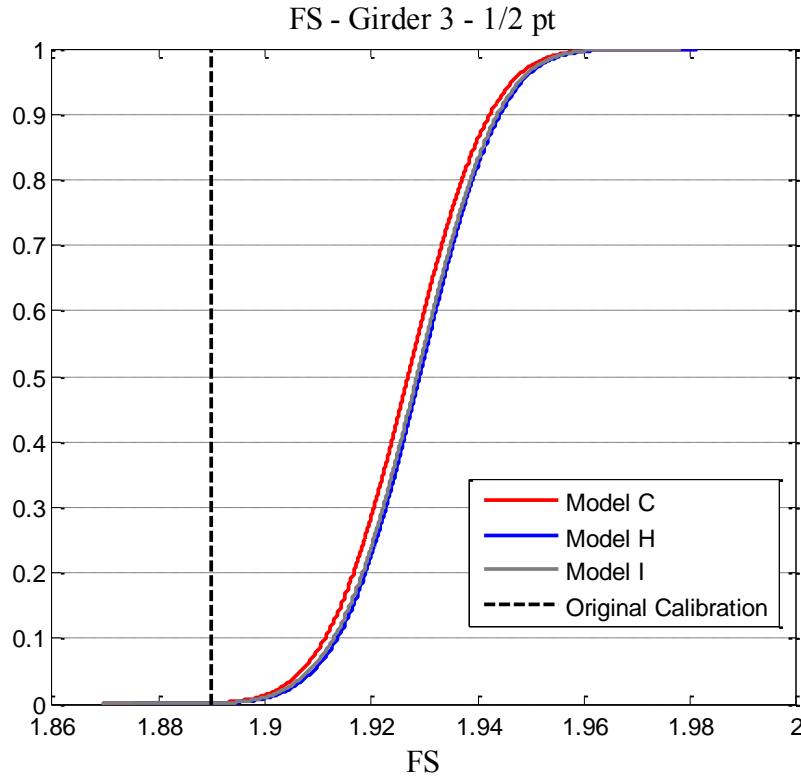


Figure 12-28 - Factor of Safety Comparison

12.4 Discussion St-Id Comparisons

There are several important conclusions that can be drawn considering the response predictions, parameter studies and rating factors presented in the previous sections.

1. Both the single model and the MCMC chains had similar frequency predictions, which were expected since they were both updated using modal properties.
2. The bias of the displacement predictions occurred because the optimization attempted to match the observed modal properties as closely as possible. It resulted in parameter values which adversely affect the response predictions.

When the optimization for the single model was initialized, the parameter bounds were based on sensitivity studies relative to the modal properties only, while the MCMC sensitivity studies used larger bounds to account for sensitivity to both the response predictions and the modal properties. While it is theoretically possible to do open the bounds up for gradient-based optimization, the fixed step size and the influence of starting point make this difficult. Global optimization may provide a better solution.

3. Considering modal properties as the observation, then the single model method and the mean of the MCMC analyses provide essentially equally likely models, which provide distinctly different response predictions. This is the main reason for multiple model investigations.
4. When the single model was calibrated using the dynamic properties, a single composite action was a valid and defendable assumption for the model. Through the implementation of the multiple model investigation it was found that the forms with this distribution were not nearly as reliable as the other, more discretized forms, and for this reason it was eliminated. Without multiple models, there is no metric, relative or absolute, by which model forms can be compared for model selection.
5. The investigation into the longitudinal bearing stiffness clearly highlighted the benefit of multiple models over traditional single model methods. A potential solution if single model approaches must be used is to rerun sensitivity studies using the optimized parameters relative to both the observations and the response predictions in order to understand if the optimized results are in an

area of high sensitivity to either the observations, the response predictions or both. This would provide the engineer with some of the insight that would come from a multiple model analysis.

Chapter 13: Conclusions, Recommendations and Future Work

13.1 Summary

This chapter presents the conclusions from both phases of this research including a best practice application of St-Id, the assessment of the test structure, the dangers of relying solely on a single model, the efficacy of different multiple model methods, and finally the comparison between traditional and multiple model St-Id. Recommendations for application of these methods and for future work are included.

13.2 Primary Conclusions

Throughout this thesis, there are numerous conclusions and recommendations, the majority of which are reiterated here. The primary, over-arching conclusions are as follows:

1. There were several key discoveries regarding the IBS Bridge stemming from both phases of this research. From the first phase, these include the nonlinearity at the pier cap, the distortion-induced fatigue at the wind-brace connections, the unsuitable nature of the bearings and the load redistribution at higher load levels. From the second phase, the importance of the longitudinal stiffness on the responses and the influence of the load magnitude even in different modalities of responses were quantified. These findings highlight the tangible benefits of St-Id for informing decision-making.
2. Even in the context of a best practice application of St-Id, single model-experiment correlation is limited and problematic. Multi-dimensional MCMC

methods were explored as an alternative and show promise as a solution to avoid the pitfall of limiting the analysis to a single model form.

3. The required time and computational power prevent multi-dimensional MCMC methods from being widely applicable currently, but when the scenario dictates extensive investigation, an immense amount of valuable information can be extracted from the multi-dimensional results, providing a much more comprehensive and robust analysis than traditional, single mode approaches.

13.3 Conclusions Regarding the IBS Test Structure

A comprehensive static and dynamic test of one span of the IBS Bridge was conducted, as well as several other smaller investigations on other spans. This test was part of an international collaborative effort as part of the Long Term Bridge Performance Program. The results of the static portion of this test conducted by Drexel University were presented herein. Critical observations and conclusions include:

1. Span 2 on the southbound side of the bridge was loaded with six legal trucks over 3 lanes, with a deflection which did not exceed the design deflection limit of L/1200 for a single truck
2. As the load level increased, the forces within the structure redistributed. In general, there was less lateral transfer under higher loads, meaning that the directly loaded girders appeared to soften while the adjacent girders appeared to stiffen
3. Girder #1 showed substantially more softening than the other girders when considering displacements, but not strains. This indicates a rigid body movement

of some type and was consistent with the observed crack in the pier cap support Girder #1.

4. The fatigue cracks that are prevalent throughout the structure are likely distortion induced as opposed to load induced cracks. The differential displacement of the girders causes the wind bracing to locally distort the web of the girders. This area is additionally stiffened by the flange and a vertical stiffener, exacerbating the stress condition.
5. The compatibility between the deck and girders appears to vary based on location on the bridge, location of the load and load magnitude during testing.

13.4 Conclusions and Recommendations Regarding Short-term Testing (i.e., St-Id) of Bridges

The thesis describes in detail each of the six accepted steps of St-Id. Numerous recommendations and guidelines were put forth, the most important of which are summarized here:

1. Step 1: Observation and Conceptualization
 - a. The candidate structure must have a valid concern that can be addressed through St-Id. Given the cost of a St-Id application, it can only be justified if conventional engineering practice fails to provide a compelling path forward, and the issue at hand has significant financial, safety or functionality implications. (Section 3.2)
 - b. The location of the structure within the surrounding infrastructure network must be understood and, if possible, utilized. Any information about the

population in the area can potentially help guide a St-Id effort, and vice versa. (Section 3.5)

- c. The feasibility of conducting the test must be considered from the outset as this can impact the cost-benefit balance of the application, and may render it impractical. This includes more than installation of sensors. Traffic control, availability of load, site accessibility and numerous other issues can hamper a St-Id effort from the start. (Section 3.3; Section 5.1)

2. Step 2: A Priori Modeling

- a. Choose a model resolution which provides maximum benefit while remaining useable, modifiable and defendable. Often this is an iterative process with increasingly complex models as the investigation continues. (Section 4.2)
- b. Use an iterative error screening procedure utilizing benchmark models when required. The old adage of “Garbage in – Garbage out” is most applicable in this stage of St-Id. (Section 4.5)
- c. Sensitivity studies for the model parameters should be conducted to investigate the influence of the parameters on not only the measured responses but also the response predictions.
- d. The a priori model should be used to provide prediction ranges for comparison during the test. Model output is not reliable without an experiment and vice versa. (Section 4.7)

3. Step 3: Experimentation

- a. The entire test program should be designed within the identified project and site constraints. (Section 5.1)
- b. Instrumentation design should stem from adherence to three requirements: constraints, situational awareness and test objectives. (Section 5.3)
- c. Consider distributed data acquisition to simplify instrumentation cabling and logistics. (Section 5.5)

4. Step 4: Data Reduction and Analysis

- a. Careful feature extraction from large datasets is required for static testing. (Section 5.7.3.1)
- b. Use observations and initial questions about the structure to guide data reduction. The framework for what needs to come from the experimental data is already in place prior to the test in St-Id. (Section 5.7.4)
- c. Use visual organization of plots and figures to relate information spatially. The ability to draw conclusions from visual assessment of data is crucial, and spatial relationships are an intuitive way to relate the data. (Section 5.7.3.2)

5. Step 5: Model-Experiment Correlation

- a. No single model-experiment correlation approach is fool-proof. A combination of heuristic and automated model correlation is required. (Section 2.5.2)
- b. Parameterization of the model is required in order to expedite any manual or automated model calibration. (Section 6.2; Section 6.3.1)

- c. Multiple modalities of response and large variations in load level may necessitate distinct models. (Section 6.5.3)
- 6. Step 6: Utilization of the Model for Simulation
 - a. The true value of the entire process is utilizing the model for simulations. The prediction of unmeasureable quantities with some confidence is the most important deliverable to a bridge owner. (Section 7.1)
 - b. These simulations serve as another tool for the engineer and owner to consider in the management of the bridge (Section 7.1.1)

13.5 Conclusions Regarding Single Model-Experiment Correlation

The non-uniqueness problem related to model updating and St-Id is well documented. It was most recently shown again in Dubbs (2012) that non-uniqueness in St-Id can call into question the validity of the results from a given model. In this research, further examples of this were shown (most notably in Chapter 12). To summarize, there is clearly information to be gained by investigating any model of a structure, even if that information is simply that the model is inappropriate. However, selecting a single model without investigating other model forms and spatial distributions of parameters, and using that model to predict other modalities of responses and other load levels is potentially disastrous. Simply put, if St-Id and model updating require sound heuristics when using a single model, then using the same heuristics to investigate more than one model form or parameter set can only be beneficial. It is also important to note that including multiple models can in no way replace or function without the engineering knowledge and heuristics required for single model approaches.

13.6 Conclusions Regarding the Cantilever Investigation into Multiple Model Methods

In Chapter 9, MCMC, RJMCMC and MC3 sampling methods were applied to a simple cantilever beam. The beam had varying section height and discretization of this section height along the length. For the methods which move between model forms (MC3 and RJMCMC) two scenarios were considered. The first included the exact model in the population and the second did not. The following conclusions were extracted from this study:

1. In most MCMC applications in other fields of science, the models used predict the same quantity that is being used to inform the sampling. It may be a data set from a different time, or under conditions, but is generally the same type of response. By using finite element models, more options open up to use the model as a translator for predictions of quantities that are of more interest.
2. RJMCMC moves between model forms in a single chain, making for an efficient method of sampling multiple model forms, particularly when compared to developing single MCMC chains for each form. However, conditioning an RJMCMC problem is difficult considering the need to develop jump equations. In certain cases (small proposal functions), the simpler models can condition the more complex models which bias the results toward these simple models.
3. MC3 provides a multi-dimensional solution which compares between model forms without transforming parameters between when a transition to a new model form is made. The method was developed by resampling fully developed and burned-in MCMC chains, which was shown to be nearly identical to moving between the chains in real-time, for the cantilever beam.

4. When the actual model is included in the population, it dominates the sampling.

This only occurs in numerical settings, as one can never produce the actual model for a real constructed system.

5. For the cantilever beam, weighting each model form using BIC was an appropriate method, particularly because the range of the MLE values for the model forms was large. If the range of the MLE values was not as broad, then BIC would not be as effective.

13.7 Conclusions Regarding the Development of MCMC Chains for the IBS Bridge

The MCMC and MC3 methods investigated in Chapter 9 were applied to the IBS Bridge in Chapter 10 and 11. Modal properties were used to inform the model updating process, and strains, displacements and rating factors were considered unmeasureable responses. A population of candidate models was developed then reduced for the full analysis. The reduced population was very similar, negating any benefit of running the MC3 analysis using this population.

1. MC3 showed potential in Chapter 9 with the cantilever study to be a robust method of combining model forms to better understand structural responses. In Chapter 10, preparation for the MC3 method ended up being a structured and defendable approach to model selection, essentially eliminating irrelevant candidate models. The end goal would be to have more than one model form in the final population, but this was not the case for the selected initial model forms. For this reason, MC3 was not actually applied to the IBS Bridge.

2. The candidate model population was reduced by comparing the maximum likelihoods of the model forms and eliminating models which would not be sampled with any regularity in a multiple model method where the acceptance of a given model is based on the likelihood of that model.
3. The burn-in process for each model returned higher likelihood values than any gradient based optimization that was done to locate starting points for the MCMC chains.
4. While an MCMC analysis will always converge to actual posterior distribution after enough samples, using an adaptive proposal function for an initial period of time can greatly improve the suitability of the proposal function and starting point for the full analysis.
5. The scenario of interest is to predict unmeasured responses by updating the model with measured observations. The parameters must reflect this translation. In other words, if a parameter is influential for the response predictions, but not for the observations, it still must be included, even if the only effect of this parameter is to widen the distributions of response predictions. This is crucial finding regarding the adequacy of the experiment to inform the prediction of the desired responses.
6. Parameters that do not affect the selected observations used for updating the model tended to “blow up” and sample the entire assigned range. This was validated through the inclusion of a dummy parameter which had no effect on the model whatsoever, which exhibited the same behavior.

7. A specific bearing stiffness was included to represent the pier cap cracking softness discovered through the analysis of the static test data. However, this was not mobilized in the modal observations, and therefore did not manifest in the model predictions at all.
8. The investigation into response predictions showed that the most likely model according to the selected observations was never the same model as the mean of the response prediction distributions. This is a common misconception.
9. The modal properties were recorded under ambient conditions for which the level of loading was unknown. The accuracy of the predictions at the two static load stages was variable, indicating that ambient loading was somewhere in between the static magnitudes.
10. The bridge exhibited local nonlinearity at the pier cap crack, as previously discussed. There was also likely some global nonlinearity which was at least partially active under the ambient loading, but not under the low level static loading. This is evident through comparison of the accuracy of the displacement predictions at the two static load stages. At the high load level, displacements were much more accurate than at the lower load level.
11. MCMC chains allow for direct interpretation of the interdependency of parameters. This can be achieved through covariance, correlation coefficients, or through visual assessment of scatter plots. This is a far superior alternative to what is available in single model approaches; mainly setting one of the dependent parameters at a value and letting the other vary, or worse letting both vary.

12. The development of the individual MCMC chains was time consuming and difficult, which would prevent these methods from being a regular solution for everyday structures. The return on the time and effort must be worth the investment, and until running and error screening MCMC chains becomes much faster, then application should be limited.

13.8 Conclusions Regarding Traditional and Multiple Model St-Id

In order to close the loop with the first phase of this research, the MCMC chain predictions were compared with the original calibrated FE model from the traditional St-Id effort on the IBS Bridge. The following conclusions were drawn based on these results:

1. Though the IBS Bridge does not have a capacity issue, the comparison of the rating factor distribution versus the single rating factor highlights the potential benefits of using multiple models to develop rating factors. In the situation where the distribution is located near the point at which the bridge owner would consider posting the bridge, the statistical properties of the distribution would be much more informative than a single rating factor.
2. Factor of Safety is a similar quantity to rating factor, but is not as heavily influenced by the magnitude of the live load demand. The resulting distribution is much tighter than a distribution of rating factors. This quantity should be considered in conjunction with rating factors when doing multiple model analyses.

3. Both the single model, and the mean prediction from the MCMC chains had similar levels of error for modal properties.
4. The displacement predictions from the single model calibrated with modal properties tended to have a large bias towards the flexible side. This is likely due to the stringent ranges applied to parameters based on their influence on the modal properties. MCMC methods use large ranges, and allow movement throughout the full range during the initial period particularly, when the proposal function is being adapted.
5. The single model calibrated using modal properties had a single composite action parameter, while the MCMC chains had three distinct regions of composite action. There were candidate models in the original population with a single composite action parameter, but they were eliminated prior to running the full MCMC chains. Without even this relative comparison between model forms, there was no way to know that the single composite action parameter was not adequate for the initial single model calibration effort.

13.9 Future Work

The following areas of future work should be considered:

1. Population MCMC and other methods which use a combined posterior space should be investigated as a more robust and mathematically defendable approach to sampling multiple model forms, as MC3 does not maintain a true Markov Chain.

2. Other model forms, including different types of model construction (i.e., shell or solid models) should be investigated for MCMC, RJMCMC and other multi-dimensional methods.
3. Different parameters and spatial distributions of these parameters should be considered, particularly for the IBS Bridge, where the end result thus far was a single model form.
4. Further study should be invested into the covariance of parameters within a model and the most reliable ways of presenting and extracting information from these values.
5. Other load levels and limit states, including non-linear models and collapse behavior should be considered as there is clearly a disconnect between real bridge behavior under substantial static and dynamic loading and finite element models which assume elastic behavior.

List of References

AASHTO (2008). Bridging the Gap: Restoring and Rebuilding the Nation's Bridges.

AASHTO (2008). The manual for bridge evaluation. Washington, DC, American Association of State Highway and Transportation Officials.

Akaike, H. (1974). "A new look at the statistical identification model." IEEE Transactions on Automatic Control **19**: 716-723.

Aktan, A. E., J. Beck, et al. (2000). The State of the Art in Structural Identification of Constructed Facilities. S. Doebling and C. Farrar, ASCE: 223.

Aktan, A. E., F. N. Catbas, et al. (1998). "Structural Identification: Analytical Aspects." Journal of Structural Engineering **124**(7): 817-829.

Aktan, A. E., B. Ellingwood, et al. (2007). "Performance-Based Engineering of Constructed Systems." Journal of Structural Engineering.

Aktan, A. E., D. Farhey, et al. (1997). "Structural Identification for Condition Assessment: Experimental Arts." Journal of Structural Engineering **123**(12): 1674-1684.

American Association of State Highway and Transportation Officials. (2007). AASHTO LRFD bridge design specifications. Washington, D.C., American Association of State Highway and Transportation Officials.

Bayraktar, A., A. C. Altunişik, et al. (2010). "Finite-Element Analysis and Vibration Testing of a Two-Span Masonry Arch Bridge." Journal of Performance of Constructed Facilities **24**(1): 46.

Bayraktar, A., A. C. Altunişik, et al. (2009). "Modal Testing, Finite-Element Model Updating, and Dynamic Analysis of an Arch Type Steel Footbridge." Journal of Performance of Constructed Facilities **23**(2): 81-89.

Beck, J. (2004). "Model Selection Using Response Measurements: Bayesian Probabilistic Approach." Journal of Engineering Mechanics **130**(2).

Beck, J. and S. K. Au (2002). "Bayesian Updating of Structural Models and Reliability using Markov Chain Monte Carlo Simulation." Journal of Engineering Mechanics **128**(4): 380-391.

Beck, J. and L. S. Katafygiotis (1998). "Updating Models and Their Uncertainties: Bayesian Statistical Framework." Journal of Engineering Mechanics **124**(4): 455-461.

Bell, E. S., M. Sanaye, et al. (2007). "Multiresponse Parameter Estimation for Finite-Element Model Updating Using Nondestructive Test Data." Journal of Structural Engineering **133**(8): 1067-1079.

Brill, E. D., S.-Y. Chang, et al. (1982). "Modeling to Generate Alternatives: The HSJ Approach and an Illustration Using a Problem in Land Use Planning." Management Science **28**(3): 221-235.

Brownjohn, J., Y. Fujino, et al. (2008). Structural Identification of Constructed Systems: Experimental Considerations. Structures Congress 2008: Crossing Borders.

Brownjohn, J., P. Moyo, et al. (2003). "Assessment of Highway Bridge Upgrading by Dynamic

Testing and Finite-Element Model Updating." Journal of Bridge Engineering **8**(3): 162-172.

Brownjohn, J. and P.-Q. Xia (2000). "DYNAMIC ASSESSMENT OF CURVED CABLE-STAYED BRIDGE BY MODEL UPDATING." Journal of Structural Engineering **126**(2): 252-260.

Brownjohn, J., P.-Q. Xia, et al. (2001). "Civil structure condition assessment by FE model updating: methodology and case studies." Finite Elements in Analysis and Design **37**: 761-775.

Brownjohn, J., A. Zasso, et al. (1995). "Analysis of experimental data from wind-induced response of a long span bridge." Journal of Wind Engineering and Industrial Aerodynamics **54/55**: 13-24.

Carlin, B. P. and S. Chib (1995). "Bayesian Model Choice via Markov Chain Monte Carlo Methods." Journal of the Royal Statistical Society. Series B (Methodological) **57**(3): 473-484.

Catbas, F. N. and A. E. Aktan (2000). Modal Analysis as a Bridge Health Monitoring Tool. ASCE Structures Congress 2000.

Catbas, F. N., A. E. Aktan, et al. (2009). Structural Identification of Constructed Facilities. ASCE, ASCE: 257.

Catbas, F. N., F. L. Moon, et al. (2008). Structural Identification of Constructed Systems: An Integrated Approach by the ASCE Committee. Structures Congress 2008: Crossing Borders.

Chajes, M. and H. W. Shenton III (2005). Using Diagnostic Load Tests for Accurate Load Rating of Typical Bridges. Structures Congress 2005. New York, NY, ASCE. **171**.

Cheung, S. H. (2009). "Bayesian Model Updating Using Hybrid Monte Carlo Simulation with Application to Structural Dynamic Models with Many Uncertain Parameters." Journal of Engineering Mechanics **135**(4).

Cigada, A., A. Caprioli, et al. (2008). "Vibration Testing at Meazza Stadium: Reliability of Operational Modal Analysis to Health Monitoring Purposes." Journal of Performance of Constructed Facilities **22**(4): 228-237.

Cowles, M. K. and B. P. Carlin (1996). "Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review." Journal of the American Statistical Association **91**(434): 883-904.

Cunha, A., E. Caetano, et al. (2001). "Dynamic Tests on large Cable-Stayed Bridge." Journal of Bridge Engineering **6**(1): 54-62.

Daniell, W. and J. Macdonald (2007). "Improved finite element modelling of a cable-stayed bridge through systematic manual tuning." Engineering Structures **29**(3): 358-371.

Doebling, S., C. Farrar, et al. (1996). Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review. A. Mauzy, Los Alamos National Laboratory: 136.

Dubbs, N. (2012). Development, Validation and Assessment of a Multiple Model Structural Identification Method. Civil, Architectural and Environmental Engineering. Philadelphia, PA, Drexel University. **Doctor of Philosophy:** 862

Fang, I. K., C.-R. Chen, et al. (2004). "Field Static Load Test on Kao-Ping-Hsi Cable-Stayed Bridge." Journal of Bridge Engineering **9**(6): 531-540.

FHWA (2001). Reliability of Visual Inspection for Highway Bridges, FHWA.

Friswell, M. I., D. J. Inman, et al. (1998). "The Direct Updating of Damping and Stiffness Matrices." American Institute of Aeronautics and Astronautics Journal **36**(3): 491-493.

Friswell, M. I., J. E. Mottershead, et al. (2001). "Finite-element model updating using experimental test data: parametrization and regularization." Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **359**(1778): 169-186.

Fruhwirth-Schnatter, S. (2001). "Markov Chain Monte Carlo Estimation of Classical and Dynamic Switching and Mixture Models." Journal of the American Statistical Association **96**(453): 194-209.

Gentile, C. (2006). "Modal and structural identification of a R.C. arch bridge." Structural Engineering and Mechanics **22**(1): 53-70.

Geyer, C. J. (1991). Markov chain Monte Carlo maximum likelihood. Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface, Fairfax, VA, Interface Foundation.

Gilks, W. R., S. Richardson, et al. (1996). Markov chain Monte Carlo in practice. London, Chapman & Hall.

Goulet, J., P. Kripakaran, et al. (2009). Considering sensor characteristics during measurement-

system design for structural system identification. Computing in Civil Engineering.

Green, P. J. (1995). "Reversible jump Markov Chain Monte Carlo Computation and Bayesian model determination." Biometrika **82**(4).

Green, P. J. (2002). Trans-dimensional Markov chain Monte Carlo. SAMSI Stochastic Computation.

Grimmelsman, K. (2006). Experimental Characterization of Towers in Cable-Supported Bridges by Ambient Vibration Testing. Department of Civil, Architectural and Environmental Engineering, Drexel University. **Doctor of Philosophy**: 465.

Haario, H., E. Saksman, et al. (1999). "Adaptive proposal distribution for random walk Metropolis algorithm." Computational Statistics **14**(3): 375-395.

Haario, H., E. Saksman, et al. (2001). "An adaptive Metropolis algorithm." Bernoulli **7**(2): 223-242.

Hanson, K. (2000). Tutorial on Markov Chain Monte Carlo. Gif-sur-Yvette, France, Bayesian and MaxEnt Workshop.

Hastings, W. K. (1970). "Monte Carlo Sampling Methods Using Markov Chains and Their Applications." Biometrika **57**(1): 97-109.

He, X., B. Moaveni, et al. (2009). "System Identification of Alfred Zampa Memorial Bridge Using Dynamic Field Test Data." Journal of Structural Engineering **135**(1): 54-66.

Jaishi, B. and W.-X. Ren (2005). "Structural Finite Element Model Updating Using Ambient Vibration Test Results." Journal of Structural Engineering **131**(4): 617-628.

Janter, T. and P. Sas (1990). "Uniqueness Aspects of Model Updating Procedures." American Institute of Aeronautics and Astronautics Journal **28**(3): 538-543.

Krzmarzick, D. P. and J. F. Hajjar (2006). Load Rating of Curved Composite Steel I-Girder Bridges through Load Testing with Heavy Trucks. Structures Congress 2006, St. Louis, Missouri, USA, ASCE.

Liu, S. and J. Yao (1978). "Structural Identification Concept." ASCE Journal of the Structural Division, **104**(ST12): 1845-1858.

Lucas, J. L., Virginia Transportation Research Council., et al. (2004). Final report : structural load testing and flexure analysis of the Route 701 Bridge in Louisa County, Virginia. Charlottesville, Va., Virginia Transportation Research Council.

Metropolis, N. N. and S. S. Ulam (1949). "The Monte Carlo method." Journal of the American Statistical Association **44**(247): 335-341.

Moller, P. and O. Friberg (1998). "Updating Large Finite Element Models in Structural Dynamics." American Institute of Aeronautics and Astronautics Journal **36**(10).

Moon, F. L. (2006). "Impacts of Epistemic (Bias) Uncertainty on Structural Identification of Constructed (Civil) Systems." The Shock and Vibration Digest **38**(5): 399-420.

Moon, F. L., A. E. Aktan, et al. (2009). Structural Identification of Various Constructed Systems to Inform Decisions. Don't Mess with Structural Engineers: Expanding Our Role, ASCE.

Morassi, A. and S. Tonon (2008). "Dynamic Testing for Structural Identification of a Bridge." Journal of Bridge Engineering **13**(6): 573-585.

Mottershead, J. E. and M. I. Friswell (1993). "Model Updating in Structural Dynamics- A Survey." Journal of Sound and Vibration **167**(2): 347-375.

NBIS (1996). National Bridge Inspection Standards: Code of Federal Regulations. Washington, D.C. , U.S. Government Printing Office.

Nichols, J. M., E. Z. Moore, et al. (2011). "Bayesian identification of a cracked plate using a population-based Markov Chain Monte Carlo method." Computers & Structures **89**(13–14): 1323-1332.

Nielsen, R. J. and N. J. McDowell (2006). Developing Bridge Rating Factors for the I. B. Perrine Bridge, St. Louis, Missouri, USA, ASCE.

Pan, Q. (2007). System Identification of Constructed Civil Engineering Structures and Uncertainty. Civil, Architectural and Environmental Engineering, Drexel University. **Doctor of Philosophy:** 243.

Papathomas, M., P. Dellaportas, et al. (2011). "A novel reversible jump algorithm for generalized linear models." Biometrika **98**(1): 231-236.

PennDOT (2008). "Software Downloads." from <http://penndot.enrprograms.com/home/>.

Phares, B., G. Washer, et al. (2004). "Routine Highway Bridge Inspection Condition Documentation Accuracy and Reliability." Journal of Bridge Engineering **9**(4): 403-413.

Prader, J. (2012). Rapid Impact Modal Testing for Bridge Flexibility: Towards Objective Condition Evaluation of Infrastructures Civil, Architectural and Environmental Engineering. Philadelphia, PA, Drexel University. **Ph.D.** .

Raphael, B. and I. Smith (2003). Fundamentals of Computer-Aided Engineering. New York, Wiley.

Robert-Nicoud, Y., B. Raphael, et al. (2005). "Model Identification of Bridges Using Measurement Data." Computer-Aided Civil and Infrastructure Engineering **20**: 118-131.

Robert-Nicoud, Y., B. Raphael, et al. (2005). "System Identification through Model Composition and Stochastic Search." Journal of Computing in Civil Engineering **19**(3): 239-247.

Saitta, S. (2008). Data Mining Methodologies for Supporting Engineers During System Identification. EPFL - Switzerland, EPFL - Switzerland. **Doctor of Philosophy:** 187.

Saitta, S., P. Kripakaran, et al. (2008). "Improving System Identification Using Clustering." Journal of Computing in Civil Engineering **22**(5): 292-302.

Sanayeи, M., E. S. Bell, et al. (2006). "Damage Localization and Finite-Element Model Updating Using Multiresponse NDT Data." Journal of Bridge Engineering **11**(6): 688-698.

Sanayeи, M., J. E. Phelps, et al. (2012). "Instrumentation, Nondestructive Testing, and Finite-Element Model Updating for Bridge Evaluation Using Strain Measurements." Journal of Bridge Engineering **17**(1): 130-138.

Scheel, I. D. A., P. J. Green, et al. (2011). "A Graphical Diagnostic for Identifying Influential Model Choices in Bayesian Hierarchical Models." Scandinavian Journal of Statistics **38**(3): 529-550.

Schlune, H., M. Plos, et al. (2009). "Improved bridge evaluation through finite element model updating using static and dynamic measurements." Engineering Structures **31**(7): 1477-1485.

Shenton III, H. W. and M. J. Chajes (2009). Experiences in Testing and Modeling for Bridge Maintenance and Rehabilitation, Austin, Texas, ASCE.

Sisson, S. A. (2005). "Transdimensional Markov Chains: A Decade of Progress and Future Perspectives." Journal of the American Statistical Association **100**(471): 1077-1089.

Smith, I. and S. Saitta (2008). "Improving Knowledge of Structural System Behavior through Multiple Models." Journal of Structural Engineering **134**(4): 553-561.

Teughels, A. (2003). Inverse Modelling of Civil Engineering Structures Based on Operational Model Data, KATHOLIEKE UNIVERSITEIT te LEUVEN. **Doctor of Philosophy:** 298.

Teughels, A. (2004). "Structural damage identification of the highway bridge Z24 by FE model updating." Journal of Sound and Vibration **278**(3): 589-610.

Teughels, A. and G. De Roeck (2005). "Damage Detection and Parameter Identification by Finite Element Model Updating." Archives of Computational Methods in Engineering **12**(2): 123-164.

Wardhana, K. and F. C. Hadipriono (2003). "Analysis of Recent Bridge Failures in the United States." Journal of Performance of Constructed Facilities **17**(3): 144-150.

Weidner, J., J. Prader, et al. (2009). Structural Identification of Bridges to Assess Safety and Performance. Structures 2009: Don't Mess with Structural Engineers, ASCE.

Xia, P.-Q. and J. Brownjohn (2004). "Bridge Structural Condition Assessment Using Systematically Validated Finite-Element Model." Journal of Bridge Engineering **9**(5): 418-423.

Yost, J. R., J. L. Schulz, et al. (2005). Using NDT Data for Finite Element Model Calibration and Load Rating of Bridges, New York, New York, ASCE.

Zarate, B. and J. Caicedo (2008). "Finite element model updating: Multiple alternatives." Engineering Structures **30**(12): 3724-3730.

Zhang, J., J. Prader, et al. (2009). "Challenges in Experimental Vibration Analysis for Structural Identification and Corresponding Engineering Strategies." UNKNOWN.

Zhang, Q. W., T. Y. P. Chang, et al. (2001). "FINITE-ELEMENT MODEL UPDATING FOR THE KAP SHUI MUN CABLE-STAYED BRIDGE." Journal of Bridge Engineering **6**(4): 285-293.

Vita

Jeffrey Weidner was born in Annapolis, MD on May 14th, 1982. He graduated with honors from Severna Park High School in Severna Park, MD in 2000. In the fall he matriculated at Drexel University in Philadelphia, PA. In 2005, he graduated from Drexel with a Bachelor of Science Degree in Architectural Engineering. He then attended Lehigh University in Bethlehem, PA and studied Structural Engineering, funded as a teaching assistant. He completed his Master's research, focused on earthquake-resistant building design, in the fall of 2007, and returned to Drexel University to begin a doctorate in Structural Engineering, focused on bridge assessment. During his time at Drexel, Jeffrey was lead field coordinator for a multi-disciplinary, international bridge research effort funded through the Long-Term Bridge Performance program, the flagship research effort of the Federal Highway Administration. He defended his research in the spring of 2012 and after graduation he plans to continue working in bridge condition assessment, eventually moving to academia.