Program Sherlock: A Tool for Stochastic Finite Element Analysis and Field Assessment of Concrete Structures

by

Rui Lin Ma

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science

Department of Civil & Mineral Engineering
University of Toronto

Program Sherlock: A Tool for Stochastic Finite Element Analysis and Field Assessment of Concrete Structures

Rui Lin Ma

Master of Applied Science

Department of Civil & Mineral Engineering
University of Toronto

2018

Abstract

The use of stochastic analysis as a way to better simulate the uncertainty effects often encountered in structural evaluation was investigated. Unlike deterministic analysis where the sensitivity to material inputs cannot be captured, stochastic simulations with the addition of established material distributions and damage prediction models can provide a higher level of confidence in structural behaviour. Stochastic analysis capability was implemented into VecTor2, a nonlinear finite element analysis program for reinforced concrete structures, and simulations were performed on a series of reinforced concrete beams and shear walls. The simulation results were compared against the deterministic and experimental results and found to be an adequate substitute whilst providing a level of confidence consistent with each specimen.

In an attempt to narrow the confidence interval, the methodology of incorporating early stage field measurements with stochastic results was explored. Using early stage parameters such as deflections, crack widths, and reinforcement strains, the stochastic output was recalibrated to produce more accurate results. The implementation of such an analysis as a field assessment tool was examined.

Acknowledgements

I would like to thank my supervisor Professor Frank J. Vecchio for his support, guidance, and most importantly patience throughout not just the past two years during my M.A.Sc. project, but for the past seven years when he first gave me the opportunity to study under his tutelage. His wealth of knowledge with regard to reinforced concrete structures is unmatched as is his kindness and understanding.

I would also like to thank my supervisor Professor Fae Azhari for her guidance in areas of research that were unfamiliar to me. I want to thank her for welcoming me to her loving family and for showing me courage in the face of adversity.

Thank you to all my fellow colleagues for their friendship that supplemented my academic experience: Joaquin Acosta, Stamatina Chasiotti, Stanley Cheng, Frederic Fong, Siavash Habibi, Benard Isojeh, Scott Koshman, Allan Kuan, Simon Liu, Andac Lulec, Giorgio Proestos, Vahid Sadeghian, and Saif Shaban. I would like to especially thank Anca Jurcut for the countless number of questions she has answered, for the countless number of stories we have shared, and the countless number of hours she has lost because of me.

Last but not least, I want to thank my family. My parents, for their full support in whatever academic pursuits I embark on and for working harder than I've ever have to provide the opportunities that I have been blessed with. And Brittany Yap, my better half, for her love and for being by my side through everything. You hold me together when I crack, without you I fall apart, but together we can stand up to anything.

"If you know a thing only qualitatively, you know it no more than vaguely. If you know it quantitatively… you are beginning to know it deeply… Being afraid of quantification is tantamount to disenfranchising yourself, giving up on one of the most potent prospects for understanding and changing the world."

-Carl Sagan, *Billions & Billions (1997)*

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Background and Motivation

In the past 50 years, many cities around the world have experienced significant construction booms at one point or another. The combination of urbanisation, rising populations, and economic growth led to increases in residential and infrastructure related construction. In some cities, such as those in Asia, the boom has been more recent. In cities like Toronto, the rise has been prevalent across the past few decades as seen in Figure 1-1. Although many new construction projects are commissioned to meet continuing growth, structures built in the 1960s, following the post-World War II economic expansion, have begun to reach their 50-year service lives. As such, many structures will need to be structurally re-evaluated to determine whether they are adequate, require structural repair and rehabilitation, or damaged beyond repair.

Figure 1-1 (Top) Graphical growth in condo units in the 1970s, 1990s, and 2010s provided by the Globe and Mail. (Bottom) Visual growth in the Toronto skyline from the 1990's to 2010's provided by CBC

The challenge with the evaluation of existing structures arises from the additional uncertainty of the available information. According to the Profession Engineers of Ontario, in addition to the issue of building codes mainly focusing on new structures, "original design and construction documents of existing buildings are often not available" leading to "difficulties [in] estimating the reliability of existing buildings." (PEO, 2012) Usually, only the nominal material properties are available for analysis which can be quite inaccurate compared to the in-situ properties, especially

after long periods of time. Furthermore, the presence of structural damage and deterioration are bound to affect the integrity of the structure, introducing more uncertainty into the problem. Without accurate information and without an understanding of the reliability of the information, it is difficult to make a confident evaluation and structural assessment. Non-destructive testing could be utilized to assess material properties and damage, however in some cases testing options are not feasible and in some cases the testing procedures introduce large levels of uncertainty themselves.

When the difficulty of the analysis centers on the uncertainty of the inputs, stochastic analysis can be more advantageous than traditional deterministic analysis. Instead of using the best available estimates to compute a single definitive prediction, repeated simulation with probability distributions of material properties and structural damage effects can produce a statistical outcome. Despite not knowing the exact properties and extent of damage, the simulation of hundreds of trials will generate hundreds of results including, theoretically, the correct behaviour. Of course, the analyst will not know which trial is the most accurate, but the trials will form a distribution showing which results are most likely and the extent of possible failures. The output takes the form of a reliable range rather than a single prediction which better quantifies the risk involved in structural evaluation.

Stochastic analysis is most powerful for systems with large degrees of freedom by producing interactive effects unexpected in deterministic analysis. The use of Monte Carlo simulations has been heavily prevalent in many complex fields ranging from fluid dynamics to operations research to economics and finance, ever increasingly with the development in computing power. However, stochastic simulations of relatively simple structural analysis models will not provide the same level of benefit. Repeated substitution of variables into a closed form failure load equation will indeed produce a distribution of results. However, the complex mechanical interactions of the reinforced concrete material cannot be captured. Therefore, a non-linear finite element analysis program, VecTor2, was outfitted with stochastic analysis capabilities (Hunter, 2016). Finite element models apply fundamentally and empirically defined laws to hundreds or thousands of individual elements that interact with each other to produce the best possible results. Thus, these models are much more suited for stochastic simulations.

Due to large levels of uncertainty with the inputs and as well as with certain types of failure mechanisms, some analyses could produce excessively large ranges of results. For example, in 2015, the University of Toronto held a prediction competition of a 4000mm deep reinforced concrete slab strip. Out of the 66 entries from both universities and industry around the world, including code equations from various countries, the failure loads predicted ranged from 200kN to 3800kN at ultimate displacements of 10mm to 140mm as seen in Figure 1-2 (Quach, 2016). The experimental failure load was 685kN at 12mm. In this case, the material properties provided to all 66 analyses, or trials, were all the same. However, the modelling tools and assumptions used were different. Despite not capturing the effects of material input uncertainties, the uncertainties within the analytical models themselves were captured and seen as the large spread of results in Figure 1-2. However, what if the University of Toronto had provided the additional information that the reaction load at 5.35mm displacement was 377kN. Would there still have been predictions of 200kN or 3800kN? This was the idea behind the incorporation of field measurements to aid the prediction of structural failure. If the amount of conventional modelling inputs are inadequate, the inclusion of additional information such as current structural behaviour should help better predict ultimate behaviour.

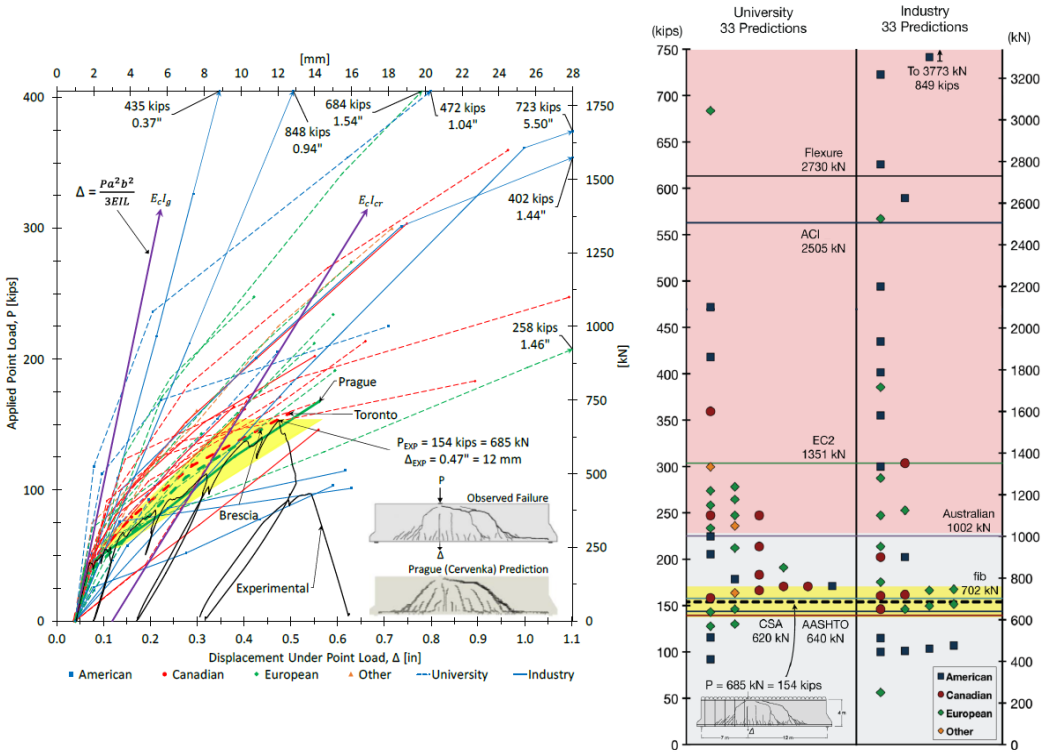


Figure 1-2 Predictions of load-deformation response and failure load for University of Toronto deep beam competition. (Taken from Quach, 2016)

1.2 Project Objectives and Research Significance

The focus of this thesis was divided between into two major objectives: to investigate the use of stochastic analysis on reinforced concrete structures, and to develop a methodology for incorporating early stage field measurements to help predict ultimate behaviour. The use of stochastic simulations influence the reliability analysis performed to determine safety factors in many building codes (Nowak and Szerszen, 2003; Razkozy and Nowak, 2014; Bartlett, 2007). However, to maintain simplicity, these safety factors are often used widespread regardless of the type of structure. In some cases these factors are over-conservative while in other cases they can be inadequate. The introduction of a tool that combines the well-established deterministic capabilities of a finite element program like VecTor2 with the capacity to perform stochastic simulations will hopefully result in more accurate reliability analysis of reinforced concrete structures. The work by Hunter (2016) introduced stochastic capabilities into VecTor2 and included stochastic analyses on a few specimens. This thesis examined 20 more specimens as well as developed a post-processing software named Sherlock.

Sherlock was also designed to facilitate the second object of this thesis: the incorporation of field measurements. As mentioned, structural re-evaluation encounters large amounts of uncertainty. Of course being able to capture the effects of uncertainty is crucial but being able to reduce the uncertainty is even more beneficial. With the belief that additional information from field measurements can aid in the prediction of ultimate behaviour, a methodology was developed and the results were compared against 23 experimental specimens. The effects of corrosion were not included in the scope of this thesis but left for future work. Nevertheless, if successful, such a tool will provide great utility in the field of structural rehabilitation.

1.3 Thesis Outline

This section summaries the following chapters of this thesis.

Chapter 2 reviews the work by Hunter (2016) in implementing stochastic capabilities into VecTor2. Basic regression analysis that was examined to develop the methodology for the incorporation of field measurements are also discussed.

Chapter 3 presents the methodology developed for the incorporation of field measurements as well as the post-processing capabilities of Sherlock.

Chapter 4 presents the results of the stochastic simulations for 23 specimens including the finite element models used, the stochastic simulation parameters, and an examination into each analysis. The overall distributions of material properties and failure loads were evaluated as well as the comparison between deterministic and stochastic results and experimental and stochastic results. Finally a statistical summary of the computational costs of stochastic analysis is provided.

Chapter 5 summarizes the results of the incorporation of field measurements. Trends of the measurement types, the measurement sensitivity, and the percentage of ultimate load at the point of measurements were analyzed. Following the analysis, the overall validity of this new methodology was examined.

Chapter 6 outlines the main conclusions as well as the major areas of improvement for future research to investigate.

Chapter 2 Literature Review

This chapter summarizes the literature required to conduct stochastic analysis using VecTor2 and to develop a method of incorporating field measurements to improve failure load predictions.

2.1 Stochastic Modelling Capabilities of VecTor2

The previous work by Hunter (2016) detailed the full implementation of stochastic analysis using VecTor2. In order to investigate the effectiveness of stochastic analysis, an understanding of the implementation was required. The overall idea was to generate sets of material modification factors from pre-defined or user-defined distributions and assign them to the materials in each trial. This section summarizes the material property distributions used, including important parameters such as mean, standard deviation, and type of statistical distribution.

2.1.1 Concrete Material Property Distributions

The three material properties examined were the compressive strength, the tensile strength, and the modulus of elasticity. Of the many literature sources reviewed by Hunter, the following were deemed the most reliable and practical.

Mirza et al. (1979)

The large variability in concrete compressive strengths can be attributed to the proportions of mix design, the mixing, placing and curing practices, the testing procedures of concrete cylinders such as the rate of loading, and the difference between test specimens and in-situ. From the data compiled by Mirza et al. (1979), the following equations summarize the statistical distribution of concrete compressive strengths.

$$f'_{c,is,R} = f'_{c,is}[0.89(1 + 0.08 \log R)] \quad (2\text{-}1)$$

$$f'_{c,is} = 0.675 f'_c + 7.584 \leq 1.15 f'_c \ (MPa) \quad (2\text{-}2)$$

$$V^2_{c,is,R} = V^2_{c,cyl} + 0.0084 \quad (2\text{-}3)$$

where $f'_{c,is,R}$ represents the mean compressive strength considering the loading rate and in-situ effects, $f'_{c,is}$ represents the mean compressive strength disregarding loading rate effects, R

represents the loading rate, f_c' represents the specified compressive strength, $V_{c,is,R}^2$ represents the squared coefficient of variation considering loading and in-situ effects, and $V_{c,cyl}^2$ represents the squared coefficient of variation for cylinder strengths. For cylinders cast with 4000psi (28MPa) or lower strength, the coefficient of variation can be taken as 0.10, 0.15, and 0.20 for excellent, average, and poor quality concrete respectively. For concrete above 4000psi, the standard deviation can be taken as 400psi (2.8MPa), 600psi (4.2MPa), and 800psi (5.6MPa) for excellent, average, and poor quality concrete. The distribution was determined to be normal.

In the VecTor2 formulation, the loading rate effects were neglected and $f_{c,is,R}'$ was taken as $f_{c,is}'$. For the calculation of the standard deviation for concrete strength using the Mirza et al. formulation, average quality concrete was assumed.

The variability in concrete tensile strengths can be largely attributed to the mix design in terms of the size and type of aggregate, air entrainment, curing conditions, water to cement ratio, cement content and the age of concrete at loading, as well as loading and in-situ effects. It is also commonly established that the tensile strength is related to the concrete compressive strength. Mirza et al. compiled data from split cylinder ($f_{sp,is,R}'$) and modulus of rupture ($f_{r,is,R}'$) tests and generated two different statistical relationships for the tensile strength of concrete as a function of the compressive strength using regressional analysis. The effect of in-situ against controlled specimens was not captured and thus the relationships represented are related to concrete strength considering in-situ effects ($f_{c,is}'$). The distribution was found to be normal.

$$f_{sp,is,R}' = \left(0.5314 {f_{c,is}'}^{0.5}\right)[0.96(1 + 0.11 \log R)]\ (MPa) \tag{2-4}$$

$$f_{r,is,R}' = \left(0.6892 {f_{c,is}'}^{0.5}\right)[0.96(1 + 0.11 \log R)]\ (MPa) \tag{2-5}$$

$$V_{sp,is,R}^2 = \frac{V_{c,cyl}^2}{4} + 0.0190 \geq V_{c,is,R}^2 \tag{2-6}$$

$$V_{r,is,R}^2 = \frac{V_{c,cyl}^2}{4} + 0.0421 \geq V_{c,is,R}^2 \tag{2-7}$$

However, in the VecTor2 formulation, the direct tensile strength ($f_{t,is}'$) is required, not the split cylinder nor the modulus of rupture. Therefore, Hunter developed a hybrid model between the

work by Mirza et al. and CSA A23.3. The coefficient of variation was assumed to be 0.127 and the distribution was normal. The mean value is calculated as below.

$$f'_{t,is} = 0.33\sqrt{f'_c}\ (MPa) \tag{2-8}$$

Finally, the modulus of elasticity was also examined by Mirza et al. (1979) similarly to the tensile strength. The distribution was compiled based on cylinder tests with strengths between 6.9MPa and 48.3MPa (1000psi and 7000psi). The relationship for the mean modulus value is presented below. The coefficient of variation of 0.08 was recommended by Mirza et al. The distribution was determined to be normal.

$$E_c = 5015.3\sqrt{f'_c}(MPa) \tag{2-9}$$

In the VecTor2 formulation, it was found that the formulation by Mirza et al. produced higher modulus of elasticity values than typically observed. Therefore, a second model was developed using the same recommended coefficient of variation but with the mean value calculated by CSA A23.3.

$$E_c = 3320\sqrt{f'_c} + 6900\ (MPa) \tag{2-10}$$

Bartlett and MacGregor (1996)

The relationship between the specified compressive strength of concrete and the actual in-situ strength of concrete was further broken down by Bartlett and MacGregor into two quantifiable factors: the variation between the specified and the concrete cylinder strength (F_1), and the variation between the cylinder and in-situ strength (F_2). F_1 was calculated as 1.25 for cast-in-place concrete with a COV of 0.104 and 1.19 for precast concrete with a COV of 0.05. F_2 was determined to be a function of the age of concrete in days (a), a Heaviside variable (Z_h) taken as 1 if the depth of the specimen exceeded 450mm and 0 otherwise, and the cement and fly ash content. The COV for F_2 was found consistently to be 0.14. A log normal distribution was assumed for both parameters. The final formula for the expected compressive strength of concrete considering both factors and relating the in-situ strength ($f'_{c,is}$) to specified strength (f'_c) is presented below. The coefficient of variation of 0.186 was recommended by Bartlett and Macgregor.

$$f'_{c,is} = \left[1.205 + 0.105Z_h + 0.125\ln\left(\frac{a}{28}\right)\right] f'_c\ (MPa) \tag{2-11}$$

If the concrete cylinder strengths were known instead of the specified strength, the factor F_1 can be disregarded in the derivation of Equation 2-11. Hunter developed the following equation for estimating the in-situ strength from the concrete cylinder strength ($f'_{c,cyl}$). Without F_1, the COV would then be 0.14.

$$f'_{c,is} = \left[0.9502 + 0.0863Z_h + 0.0985\ln\left(\frac{a}{28}\right)\right] f'_{c,cyl}\ (MPa) \qquad (2\text{-}12)$$

In the VecTor2 formulation, it was assumed that the specimen always exceeded 450mm and Z_h was always 1.

Nowak and Szerszen (2003)

The work by Nowak and Szerszen was based upon the calibration of the ACI-318 building code. The formulation developed between the specified concrete strength and the mean in-situ compressive strength is as given in Equation 2-13. The coefficient of variation was 0.101 and the distribution was determined to be normal. The VecTor2 formulation followed the above formula exactly.

$$f'_{c,is} = -0.0081(f'_c)^3 + 0.1509(f'_c)^2 - 0.9338(f'_c)^3 + 3.0649\ (ksi) \qquad (2\text{-}13)$$

Unanwa and Mahan (2014)

A large sample size of 3269 cylinders were collected by Unanwa and Mahan from the Caltrans Transportation Laboratory in California. The cylinders had specified compressive strengths of 25MPa, 28MPa, and 35MPa and were cast between 2007 and 2011. Through regressional analysis, relationships were generated to predict the cylinder strengths at a time t (days) for 25MPa, 28MPa, and 35MPa concrete as shown in Equations 2-14a to 2-14c respectively.

$$f^t_{c,cyl} = 8.763\ln(t) + 6.731\ (MPa) \qquad \text{For } f'_c = 25\ MPA \quad (2\text{-}14a)$$

$$f^t_{c,cyl} = 12.021\ln(t) - 0.309\ (MPa) \qquad \text{For } f'_c = 28\ MPA \quad (2\text{-}14b)$$

$$f^t_{c,cyl} = 13.033\ln(t) + 3.248\ (MPa) \qquad \text{For } f'_c = 35\ MPA \quad (2\text{-}14c)$$

The average factor between the specified and cylinder strength was 1.45 for 25MPa and 28MPa concrete and 1.33 for 35MPa concrete. The coefficients of variation were 0.19, 0.18, and 0.13, respectively.

To further develop the model to predict the in-situ compressive strength, Unanwa and Mahan proposed the formulation of three factors as seen in Equation 2-15: f_1, the relationship between the specified strength and the cylinder strength; f_2, the relationship between the cylinder strength and the in-situ strength at 28 days; f_3, the relationship between the 28-day in-situ strength to the strength at time t. f_1 can be taken from the findings presented above. f_2 was taken as 0.81 by Unanwa and Mahan. f_3 was developed as follows.

$$f'_{c,is} = f_1 f_2 f_3 f'_c \tag{2-15}$$

$$f_3 = e^{\left[0.3\left(1-\sqrt{28/t}\right)\right]} \tag{2-16}$$

The removal of f_1 produces the relationship between the 28-day cylinder strength and the in-situ strength.

In the VecTor2 formulation, it was assumed that f_1 always equated 1.45 with a COV of 0.18.

2.1.2 Steel Material Property Distributions

The three material properties examined were the yield strength, the ultimate strength, and the modulus of elasticity. Of the many literature sources reviewed by Hunter, the following were deemed the most reliable and practical.

Mirza and MacGregor (1979)

The work by Mirza and MacGregor examined the material properties of reinforcement bars used in reinforced concrete structures. In terms of the yield strength, five factors were determined to influence the variation. These factors were: the variation in steel material, the variation in rebar cross-sectional area, the loading rate, the bar size, and the strain at yield. It was found that the variability within one single bar and within the same batch of bars was small while the variability within bars from different batches was large (Mirza and MacGregor, 1979). The focus of the study was limited to Grade 40 (280MPa) and Grade 60 (420MPa) deformed bars. The distribution was determined to be best represented by beta distributions. The probability distribution functions (PDF) are presented below for Grade 40 and Grade 60 steel. The range for Grade 40 steel was found to be 227MPa to 428MPa (33ksi to 62ksi), the range for Grade 60 steel was found to be 372MPa to 704MPa (54ksi to 102ksi).

$$PDF(f_y) = 4.106\left(\frac{f_y-33}{29}\right)^{2.21}\left(\frac{62-f_y}{29}\right)^{3.82} \quad \text{For Grade 40} \quad (2\text{-}17a)$$

$$PDF(f_y) = 7.587\left(\frac{f_y-54}{48}\right)^{2.21}\left(\frac{102-f_y}{48}\right)^{6.95} \quad \text{For Grade 60} \quad (2\text{-}17b)$$

The study by Mirza and MacGregor also determined beta distributions for the ultimate strength of Grade 40 and Grade 60 steel rebars. The range for Grade 40 steel was between 51ksi to 96ksi (351MPa to 662MPa) and the range for Grade 60 steel was between 84ksi to 158ksi (579MPa to 1090MPa). The distributions for the yield strength and ultimate strength are shown in Figure 2-1.

$$PDF(f_u) = 2.646\left(\frac{f_u-51}{45}\right)^{2.21}\left(\frac{96-f_u}{45}\right)^{3.82} \quad \text{For Grade 40} \quad (2\text{-}18a)$$

$$PDF(f_u) = 4.922\left(\frac{f_u-84}{74}\right)^{2.21}\left(\frac{158-f_u}{74}\right)^{6.95} \quad \text{For Grade 60} \quad (2\text{-}18b)$$

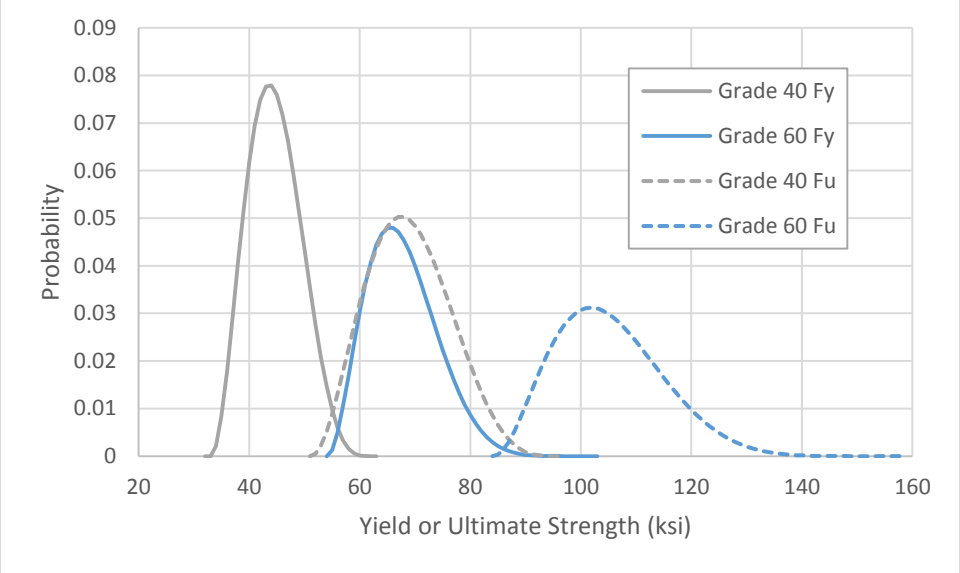


Figure 2-1 Probability distribution for reinforcement material properties by Mirza and MacGregor (1979)

In the VecTor2 formulation, the Grade 60 steel distributions for yield and ultimate strengths were used with a modification factor for other specified grades of steel. This method was cautioned by the Mirza and MacGregor as untested. According to Mirza and MacGregor, the distribution for the modulus of elasticity was normal with a mean value of 201,000MPa and a coefficient of variation of 0.033. The modulus of elasticity was modelled in VecTor2 the same as proposed.

Nowak and Szerszen (2003)

A compilation of test data was created by Nowak and Szerszen for Grade 60 (420MPa) bars ranging in size from No. 3 to No. 11 (nominal diameters of 9.53mm to 35.81mm). They found no noticeable trend between the bar size and the yield strength and the distribution can be modelled as normal with a mean bias factor of 1.145 and a coefficient of variation of 0.05.

In the VecTor2 formulation, this model was followed exactly with a modification factor used for other specified grades of steel.

Wisniewski et al. (2012)

A more recent study by Wisniewski et al. examined reinforcing steel from European manufacturers and the difference between more modern day variability in steel versus past manufacturers. The study concluded that the bias factor for the yield strength of old and new steel should be 1.20 and 1.15 respectively with coefficients of variations of 0.10 and 0.05. A distinct definition of old and new reinforcement was not provided in the study. The modulus of elasticity for steel was defined as normally distributed with a mean of 202,000MPa and a coefficient of variation of 0.04. In the VecTor2 formulation, this model remains a work in progress due to the lack of distinction between old and new reinforcement.

2.2 Past Stochastic Simulation of Reinforced Concrete Structures

This section summarizes stochastic simulations of reinforced concrete structures by previous researchers.

Ramsay et al. (1979)

Stochastic analysis focusing on short-term deflection of reinforced concrete beams was examined by Ramsay et al. (1979). The set of specimens included 10 T-beams and 10 one-way slabs. The analytical model used to predict deflections was based on elastic beam theory using an effective moment of inertia that is an approximation between the gross and the cracked moment of inertia. The mechanism of moment redistribution was captured through an iterative approach to converge on the effective moment of inertia. The variability introduced in the analysis was in the concrete material properties, steel material properties, and member dimensions. Monte Carlo simulations were conducted at four load stages with 500 trials each for all specimens. The results showed

increasing variability in deflection as the beam specimens approached cracking. The variation then decreased as the load increased further past the cracking load. It was concluded by Ramsay et al. that the largest form of uncertainty to the beam stiffness was due to concrete cracking and likely concrete material properties. For smaller beam specimens, the placement of bars also significantly affected the response.

Mirza and MacGregor (1982)

The stochastic study by Mirza and MacGregor (1982) was conducted to develop material resistance factors for the CSA A23.3 Code. The set of specimens included flexural members, combined flexural and axial members (both slender and squat), and prestressed flexural members. The analytical model used to predict ultimate load capacities was sectional analysis. For flexural only specimens, the moment curvature relationships were generated. For combined axial and flexural specimens, the axial load was increased incrementally to obtain interaction relationships. The concrete compressive post-peak behaviour used was Hognestad's parabola. The tensile behaviour was assumed linear elastic with brittle failure in tension. The reinforcement stress-strain response was assumed to be linear elastic perfectly plastic. Finally, shear resistance was calculated based on closed form code equations. The random inputs were the concrete and steel material properties based on the previous work by Mirza and MacGregor (1979). Variability was also considered for prestressing reinforcement and the specimen dimensions. The stochastic results were normalized against nominal resistances to create strength reduction factors used in the CSA A23.3 Code.

Mirza (1998)

Another set of stochastic analyses was conducted by Mirza (1998) to study the interaction behaviour of composite steel/concrete columns for CSA A23.3. The set of specimens included four short columns and six slender composite steel/concrete columns with varying concrete strengths, structural steel ratios, and slenderness ratios. The analytical model used was again section analysis but with consideration of second-order effects by using trial and error to predict a fourth-order deflected shape along the column and considering maximum eccentricities. The concrete compressive post-peak behaviour was modelled by the modified Park-Kent model (1982) considering confinement from stirrups. The tensile behaviour was linear elastic with brittle failure.

The reinforcement and structural steel were assumed to be linear elastic perfectly plastic. The material properties were the main sources of variability. 500 trial simulations were conducted for each specimen. The stochastic results were then compared to nominal resistances without material resistance factors. From the study, Mirza determined that the variability in structural resistance increases with more slender columns. Overall, a coefficient of variation of 0.14 to 0.15 was computed for compression governed failure including slender specimens, while a coefficient of variation of 0.06 to 0.08 was computed for flexural or tension governed failure.

Choi et al. (2004)

The work by Choi et al. (2004) was a continuation of the work by Ramsay et al. (1979) to examine beam deflection predictions. The sets of specimens were: six simply supported one-way slabs, three fixed end one-way slabs, and three two-span continuous T-beams. The analytical model was upgraded to a layered finite element model. The effects of tension stiffening were considered as were long-term effects such as creep and shrinkage. The analytical model was validated using other experimental data. The variability input included material properties, dimensional properties, and different loading histories. The different loading histories assumed different levels of significant construction load as instantaneous loading. The concrete compressive strengths were randomly generated while the modulus of rupture and modulus of elasticity were calculated from the concrete compressive strength for each trial. The stochastic results showed different variability across different specimens despite using the same variability inputs. For specimens that did not crack under the applied loading, the variability in behaviour was low, while the specimens subjected to cracking produced larger variability in the deflection values.

Vincent et al. (2011)

Vincent et al. (2011) also studied the deflection of two reinforced concrete beams and one prestressed slab. The analytical model was similar to that of Ramsay (1979) based on elastic beam theory and effective modulus of elasticity. The difference was that Vincent et al. computed confidence intervals for beam deflection calculations based on the stochastic results.

Ning and Li (2016)

An in-depth stochastic analysis was conducted on reinforced concrete squat shear walls by Ning and Li (2016). An extensive set of 182 squat shear walls was compiled from literature. The set

considered wide ranges of design and experimental properties including both cyclically and monotonically loaded specimens. The analytical model was based on a closed form formula that combined the mechanism of diagonal strut action, and horizontal and vertical mechanism from strut-and-tie analysis. The formula contains four unknown model parameters to capture stochastic behaviour in lieu of using material distributions from the literature. Using the Generalized Likelihood Uncertainty Estimation (GLUE) method developed by Beven and Andrew (1992), the distribution of the four unknown model parameters were determined. Essentially, these model parameters were assigned random values between 0.0 and 1.0 for more than 100,000 independent samples. Then, each sample is represented by a different closed-form formula in the calculation of shear wall resistance. This formula was then applied to each of the 182 specimens and the predicted resistance and the experimental resistance was compared through the Nash-Sutcliffe (NS) efficiency as a likelihood function (Nash and Sutcliffe, 1970) for the select set of unknown model parameters. Finally, using certain threshold values, the 100,000 samples were filtered and the distribution type and parameters of the four model parameters were determined. Once determined, the closed form formula was used to recalculate the shear resistance for each of the 182 specimens. The results are presented in Figure 2-2. According to Ning and Li (2016), the use of this model can produce appropriate confidence levels to guide the seismic design of squat shear in walls.

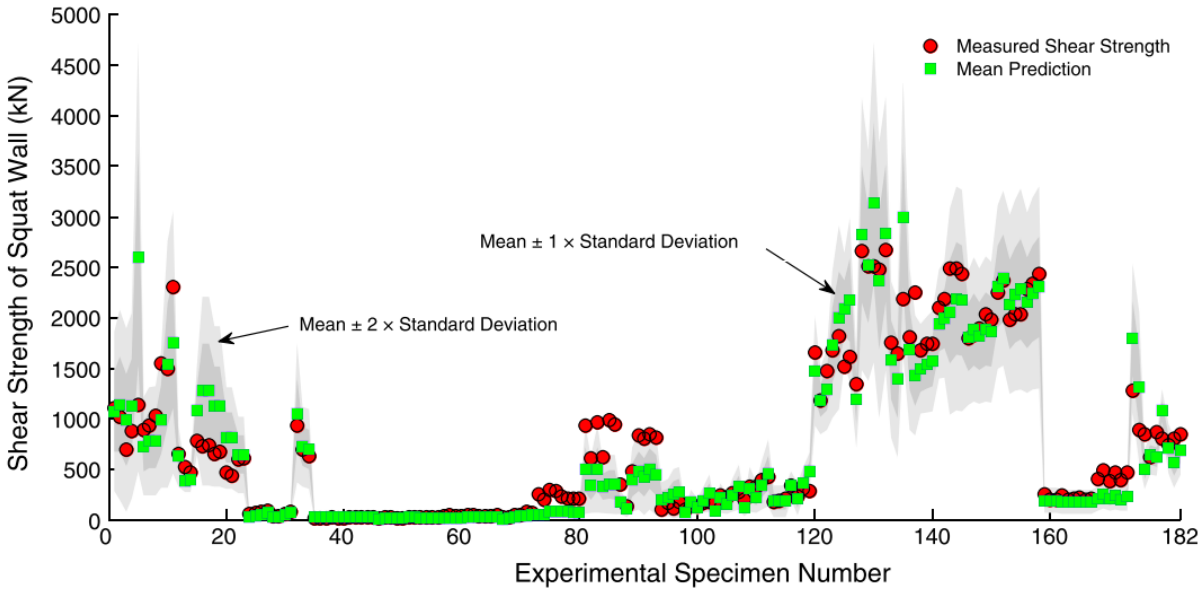


Figure 2-2 Stochastic predictions for shear strength of squat shear walls taken from Ning and Li (2016)

2.3 Total Least Squares Method

In order to match sets of field measurements against predicted trial relationships, an understanding of regression analysis methods was required. For an analysis considering two variables, or measurements, the goal is to essentially find the distance between a "point," the field measurements, and a field of "lines", the predicted responses. Initially, the Least Squares Method was considered where one variable would be considered the independent variable and the error would only be observed in the dependent variable. It was decided that the displacement values would be considered independent and the other measurements would be dependent. However, this was arbitrary and it was not justifiable to have no observed error in one type of measurement.

Therefore, the Total Least Squares Method was examined to consider errors in all variables. Figure 2-3 shows a simple comparison between the Least Squares Method and Total Least Squares Method. The difficulty with using the Total Least Squares Method arises when the variables are not of the same units. When calculating the shortest distance between a point and a line, the addition of values with different units could occur. Two approaches can be used to avoid such problems. The first method considers converting variables to dimensionless values (Markovsky, 2007) in a process called normalization or standardization. One common normalization method is to use measurement precisions. The other approach is to consider the residual as a product of the horizontal and vertical residual in a two variable case.

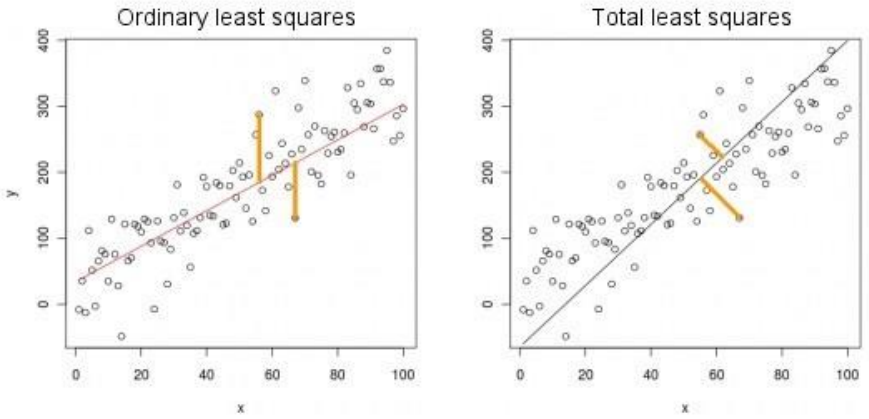


Figure 2-3 Comparison of Least Squares and Total Least Squares methods

Chapter 3 Sherlock Implementation

To facilitate analyses incorporating field measurements and post-processing of stochastic results, a software program called Sherlock was developed. Sherlock was designed to work with VecTor2, the program used to conduct stochastic analysis of reinforced concrete structures (Hunter, 2016).

3.1 Post-processing of Stochastic Analyses

Stochastic analyses can generate large volumes of data. Therefore, an efficient post-processor was required to easily examine the most important results. Sherlock was designed to extract nodal displacements, nodal reaction forces, elemental total strains, elemental reinforcement strains for both smeared and discrete reinforcement elements, and elemental crack information from any node or element of the finite element model, referenced by the node or element number as seen in Figure 3-1.

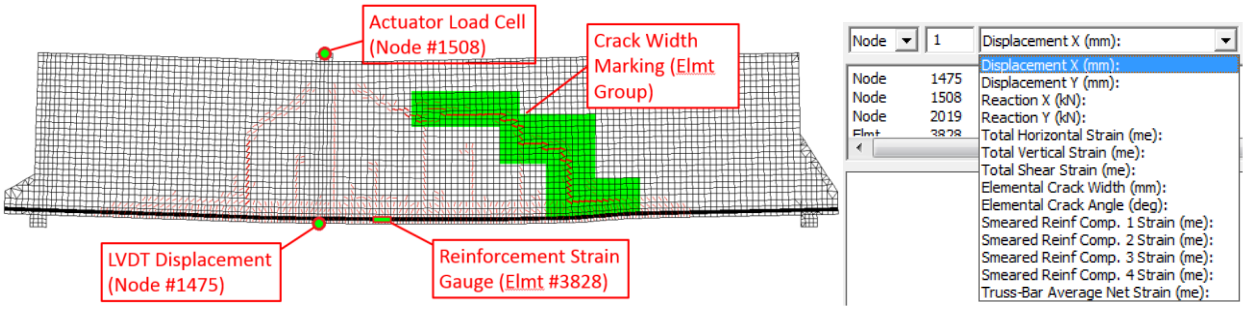


Figure 3-1 Sample outputs that can be processed by Sherlock

3.1.1 Crack Information Extraction

The extraction of crack information from analytical results was less trivial compared to other data. Displacements, loads, or strains measurements can all be extracted from a general location where a single node or element could be selected for extraction that would produce a value per load stage and per trial. However, cracks in experimental specimens propagate across distances represented by many elements in a finite element model. Furthermore, for analysis of shear-critical structures, the location where major concentrated cracks develop is random both physically and analytically. VecTor2 employs a smeared rotating crack model where each element can undergo cracking based on the stress state of the element. As elements without steel reinforcement crack, they lose elemental stiffness leading to a redistribution and concentration of cracks. Although the location

of these cracks will be near locations of high tensile stress, the exact path will be very different depending on the analysis, as seen in Figure 3-2. Therefore, to capture potential cracking areas, large regions or groups of elements would need to be selected for consideration as delineated in Figure 3-1. Multiple groups were allowed for tracking of multiple cracks.

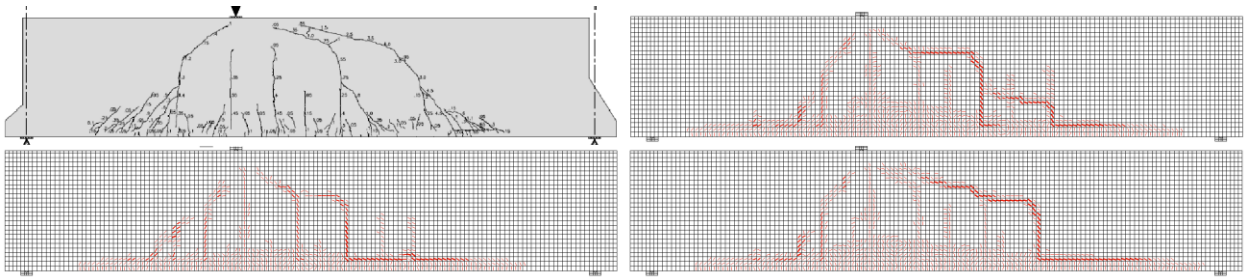


Figure 3-2 Comparison of crack patterns between experimental (top left) and finite element analyses

Finally, to ultimately match against field measured crack information, a single value must be extracted. The largest crack width found within the group of elements was extracted since experimental concrete cracks are always measured at their maximum width.

3.1.2 Ultimate Load Extraction

To determine the cracking and ultimate loads for each trial in the stochastic analysis, the specified support reactions and critical displacements, typically at the location of loading, were used. The algorithm for determining ultimate loads could typically be as simple as taking the maximum load value. However, in certain situations, an artificially generated second peak behaviour can be produced. In other cases, divergence at the end of analyses could generate unrealistically large values. Thus, the algorithm was refined to stop searching for ultimate load values after the load drops to 95% of the previously determined ultimate load but not before the peak crack load if specified. A similar approach of 90% was used by Hunter (2016), however, it was seen that the use of 95% produced more accurate extraction of failure loads.

Figure 3-3 shows the process for one trial as well as the final results for 500 trials. Sherlock allows the analyst to further refine the determination of ultimate loads upon visual inspection by specifying displacement ranges to search within. It was paramount that the correct ultimate loads were extracted since in stochastic analysis, where large amounts of data must be processed through automation, errors within the automation process could go undetected. Finally, the load values could then be statistically analysed to determine the distribution and important parameters such as the mean and standard deviations. Figure 3-4 summaries the basic methodology.

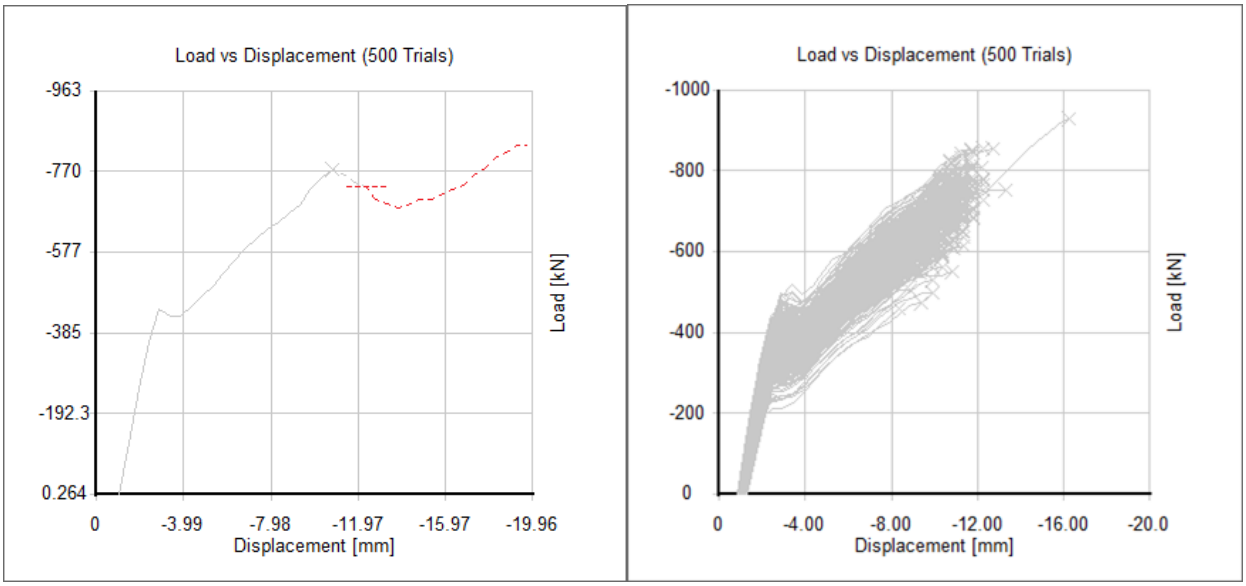


Figure 3-3 The determination of ultimate load for a) one trial and b) 500 trials.

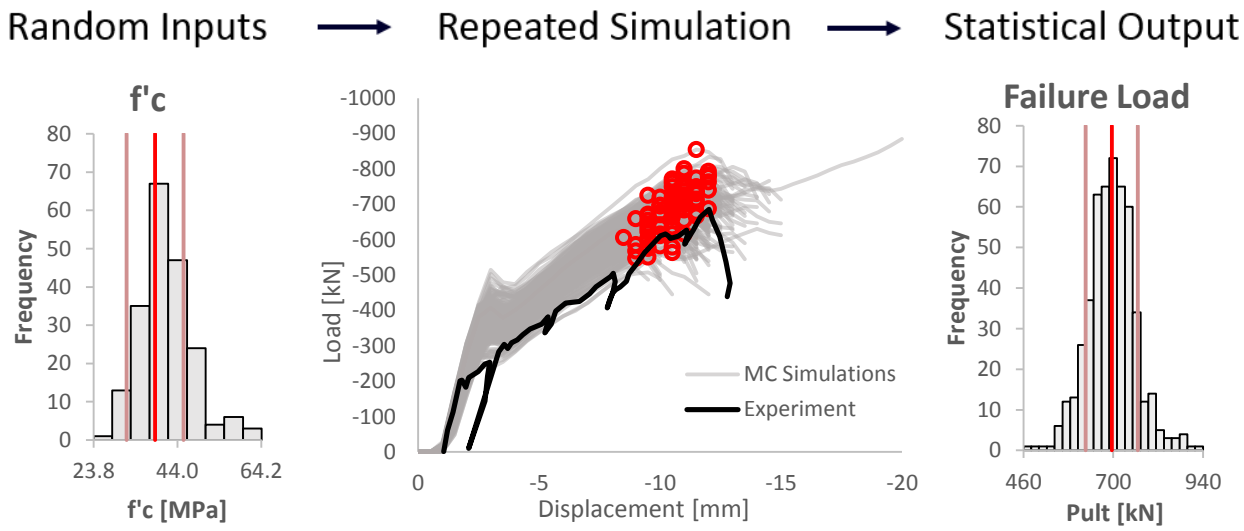


Figure 3-4 Basic workflow for stochastic analysis

3.2 Incorporation of Field Measurements

After the extraction and post-processing of all stochastic results, Sherlock was designed to match each trial against a set of field measurements. Possible field measurements included displacements from LVDTs, reactions from load cells, crack widths and angles from crack markers, strain values from reinforcement strain gages, and surface strains from LED targets or coupled LVDTs.

3.2.1 Field Measurement Sensitivity

The acquisition of field measurements typically produces a level of uncertainty that must be accounted for. These uncertainties can be due to inadequate calibration of measurement devices, human error in reading of analogue measurements, or imprecise location of measurement with regards to the structure. Field measurements themselves can be regarded as another form of modelling inputs; the same as material properties. Therefore, a deterministic acceptance of these parameters would contradict the stochastic philosophy and must be supplemented with more statistical information.

The simplest method would be to statistically sample the measurement and extract a mean and standard deviation value. Otherwise, the variance could be approximated by using values from the literature, or more effectively from expert opinion given the specific sensors or other measurement apparatus used.

Sherlock was set up to allow the analyst to input field measurements as a mean value plus/minus one standard deviation. This would give the analyst freedom to provide the most accurate measurement sensitivity. A full in-depth study could have been conducted on sensors and other measurement devices to provide Sherlock with a built-in database to suggest sensitivity values. However, this was beyond the scope of this thesis. Nevertheless, based on preliminary research, a suggestion of 15% C.O.V. was made for displacement values based on Mishra (2005), 5% for load values based on Blakeborough (2001), and 10% for strain values. For crack marking information, a constant sensitivity of 0.15mm was used since the lowest denomination of typical crack marking apparatus is 0.05mm and extra width was considered for human error.

3.2.2 Trial Matching

With the inclusion of field measurements and analytical results from stochastic analysis, Sherlock was designed to assess each trial and compute a probability that the select trial was a match for the field measurements. The computed probability values would then allow Sherlock to rank and determine which trials would be deemed high or low confidence trials. The probability was computed based on how far the measured values deviated from the trial values. In more direct terms, the shortest distance between the field measurement ordinate to the trial results in n dimensions was calculated, where n equals the number of field measurements considered. Figure

3-5 (a) shows the analysis results of two select trials with two (n=2) field measurements; mid-span deflection and reaction load. The field measurement ordinate (5.35mm, 377kN) is shown as the blue 'x' with the measurement sensitivity (±0.5mm, ±20kN) represented by the blue ellipse around the 'x.' Figure 3-5 shows the trial in grey having a larger minimum distance to the field measurements than the trial in green and thus the second trial is a better, more confident match to the field measurements.

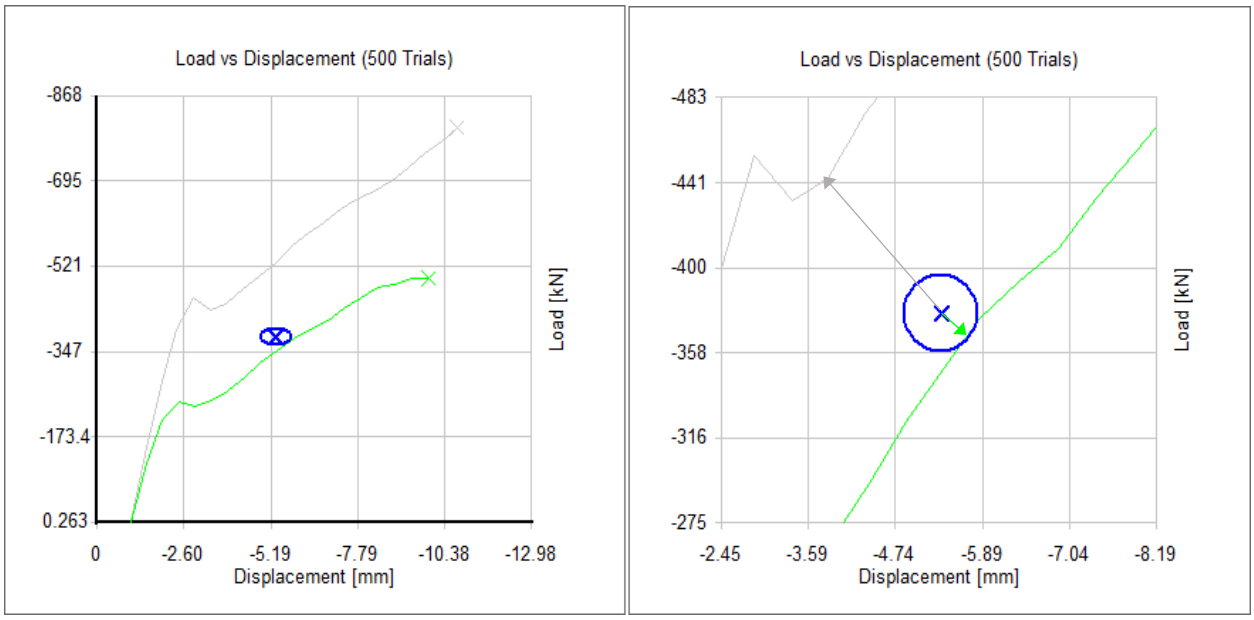


Figure 3-5 Determination of trial confidence based on the deviation between the field measurements (x marker) and analytical responses and the trial sensitivity (blue circle) for two sample trials

However, in carrying out the analysis for measurements with different units (mm, kN), the distance calculation becomes meaningless due to the addition of values with different units. Furthermore, scaling up one measurement (from mm to m) would change the distance calculations entirely. Thus, Sherlock was programmed to convert and normalize the measurements by the sensitivity values to become unitless. This would map the displacement versus load coordinate space to a displacement deviation versus load deviation coordinate space where both values are unitless and the distance calculation is valid. The root sum of squares of each field measurement deviation effectively captures the influence of each measurement. Normalizing against measurement sensitivity values also provides other benefits. A 5kN offset and a 5mm offset can be ambiguous in terms of which measurement was more erroneous. However, after being normalized against sensitivity values of ±0.5mm and ±20kN, 10 standard deviations in displacement is much greater

than 0.25 standard deviations in load. This also effectively allows the analyst to assign weighting to different field measurements. This method is commonly seen in total least squares regression methods which is similar to least squares regression except errors are accounted for in both dependent and independent variables (Markovsky, 2007). After determining the minimum deviation for each trial, a probability can be calculated based on a normal distribution. If a deviation of 1.0 is found, there is a 32.7% probability that the trial is still a good estimate assuming a normal distribution shown in Equation 3.1. While a deviation of 1.96 would correlate to a probability of 5%. All variables are treated as independent measurements. A bivariate normal distribution centered at the measurement ordinate with the provided standard deviations and zero covariance value is shown in Figure 3-6.

$$P(X \leq \mu - 1\sigma) + P(X \geq \mu + 1\sigma) = 0.327 \tag{3-1}$$

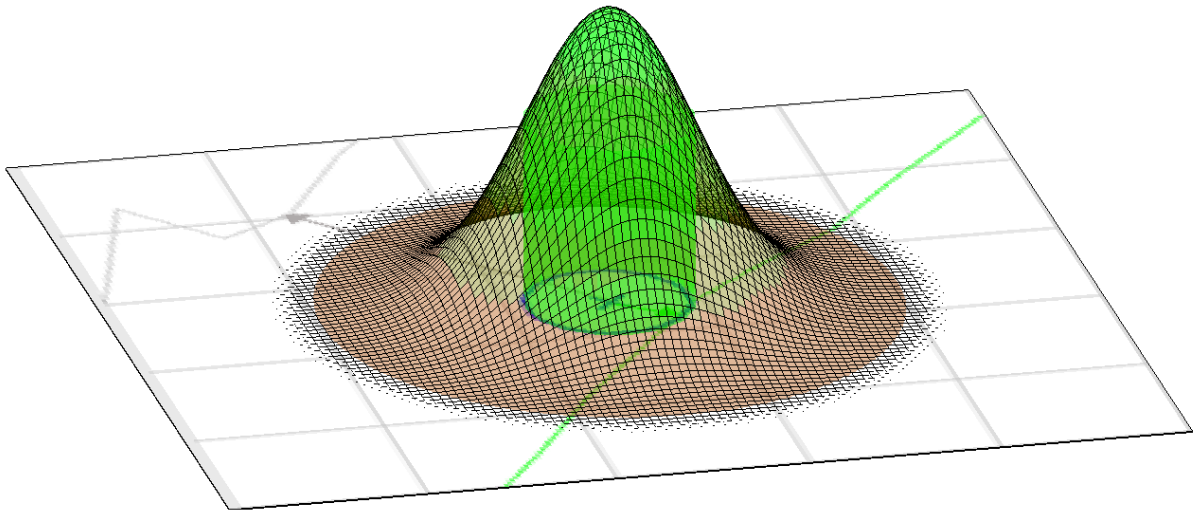


Figure 3-6 A bivariate normal distribution centered at the measurement ordinate showing the relationship between the deviation to select trial and the probability the select trial is a good estimate.

Figure 3-7 shows a full analysis of 300 trials with three field measurements: displacement, load, and crack widths. The green trials represent trials more likely to capture the correct behaviour while the red trials represent less likely trials. A three-dimensional visualization of select trials can be seen in Figure 3-8 with the 2-D projections in Figure 3-7 also illustrated.

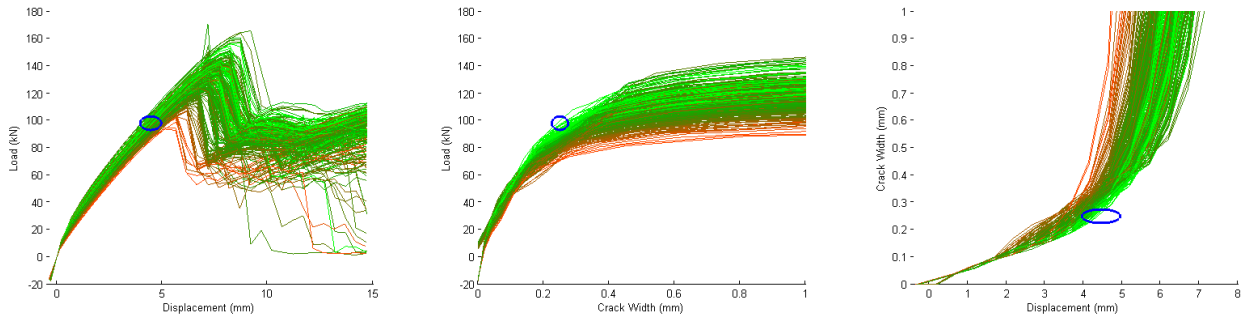


Figure 3-7 Results of trial match analysis with three types of field measurements (load, displacement, crack width) shown as two measurement projections

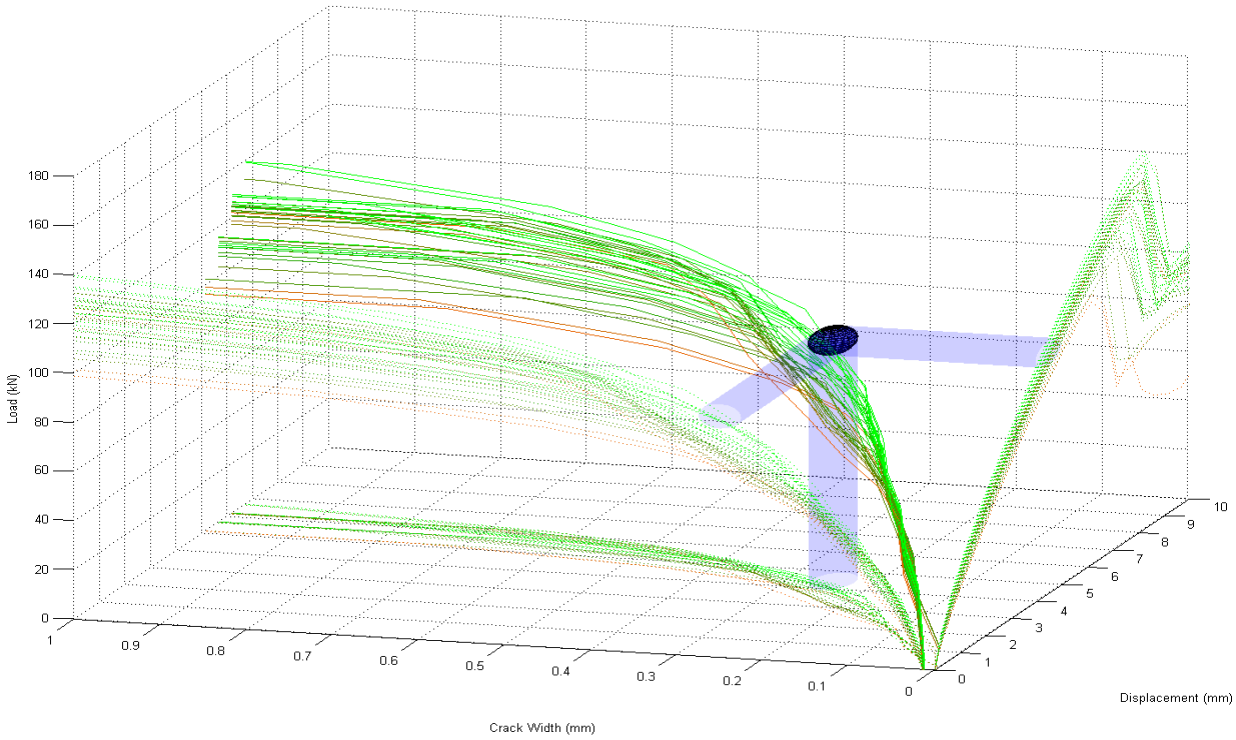


Figure 3-8 Results of trial match analysis with three types of field measurements (load, displacement, crack width)

3.2.3 Modified Stochastic Output

The final step in incorporating field measurements with stochastic simulations was to alter the stochastic output. The previously generated distribution was modified by eliminating trials that were deemed too unlikely to be good estimates. This would concentrate the distribution towards the trials that were exhibiting structural behaviours that matched the field measurements. The new subset of trials would provide a new distribution of results to examine. The cut-off for eliminating trials was set at 32.7% probability or within 1.0 deviation from the field measurements.

Graphically, this would mean that all accepted trials must pass through the ellipse (or ellipsoid in 3D, etc.) created by the field measurements and the measurement sensitivities. This further reinforced the consideration of all field measurements. A certain trial could have a very close match in terms of the load and displacement measurements, but if the crack width was matched poorly, that trial will not be considered. The updated methodology can be seen in Figure 3-9.

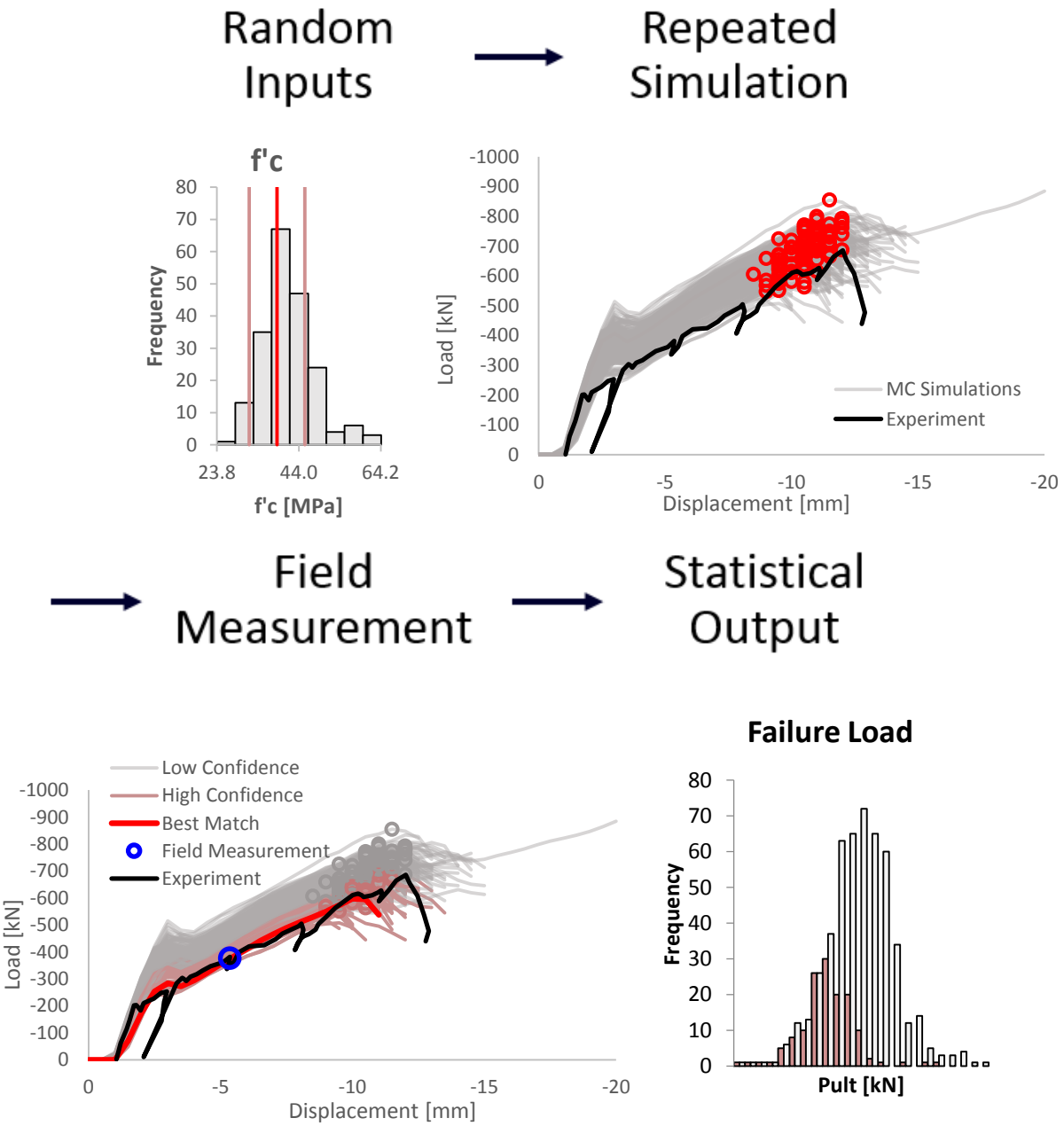


Figure 3-9 Updated methodology for the incorporation of field measurements

Chapter 4 Stochastic Simulation Results

This chapter summarizes the specimens that were analyzed with stochastic simulations and present their results are shown. The ideal experiments to examine for this thesis would have been aged structural components subjected to possible corrosion and other damage effects as well as detailed records of experimental data. In lieu of such experiments which were not readily available, conventional reinforced concrete experiments were used with a focus on different structural elements and detailed data records.

4.1 Specimens Overview

The work by Hunter (2016) examined four beams from the Toronto size effect series (Stanik et al., 1998; Yoshia et al., 2000; Cao et al., 2001; Sherwood et al., 2008; Quach et al., 2016). This thesis examines eight beams from a series of experiments (Shim et al., 2004) modelled after the classic series of reinforced concrete beams by Bresler and Scordelis (1963), eight shear walls by Lefas et al. (1990) and the Portland Cement Association (Oesterle et al., 1976), in addition to seven beams from the Toronto size effect series by Stanik et al. (1998) and Quach et al. (2016).

The series of beams by Bresler and Scordelis (1963) has long been used as a benchmark for data calibration and verification of finite element models. The series exhibited a wide range of failure modes from using various spans and reinforcement ratios. The recreation by Shim et al. (2004) at the University of Toronto was selected because it successfully replicated the original experiments, was more available in terms of detailed data, and used more modern mix designs. To obtain how stochastic analyses behaved with different types of structures, shear wall experiments also were examined. However, the scope was limited to monotonically loaded shear walls only. The proposed methodology of incorporating field measurements by mapping against stochastic results would have been excessively chaotic if applied to cyclically loaded experiments. Furthermore, cyclic analyses require far greater numbers of load stages, into the hundreds. As Section 4.2 and Section 4.6 will show, the large computational time and large data storage required can become intractable. One of the few modern shear wall experiment series tested monotonically was that of Lefas et al. (1990) and thus it was selected. In addition, out of the series of shear walls tested by the Portland Cement Association (Oesterle et al., 1976), only one wall, Specimen B4, was tested

monotonically. Finally, more beams from the Toronto Size Effect series were examined to compare with the analyses by Hunter but also to extend the sample size of specimens. The experiments by Stanik (1998) and Quach (2016) investigated the effects of varying depth on beams and slabs without shear reinforcement. Analysis of shear-critical beams containing no shear reinforcement is complex and often requires established finite element programs like VecTor2. Well-kept data was also retrievable for both sets of experiments.

4.2 Trial Number Sensitivity Analysis

Before stochastic analysis was conducted on all specimens, a sensitivity analysis was undertaken on the number of trials used. In any Monte Carlo simulation, large amounts of trials are required to generate as many different outcomes as possible. However, more trials lead directly to more computational time and data storage. Therefore, an optimization is ideal. In the previous work by Hunter, the four stochastic simulations conducted used 200, 300, 400, and 175 trials. The numbers were chosen based on model simplicity, variation in failure mode, and computation time as well as the importance of fringe results (Hunter, 2016). A final recommended default was set at 200 trials minimum.

The sensitivity analysis was conducted on Specimen PLS4000, tested by Quach, for 100, 200, 350, and 500 trials. The finite element mesh was created using FormWorks (Wong, 2013), the pre-processor to VecTor2, and is shown in Figure 4-1 with the material and model properties summarized in Section 4.3. The stochastic analysis parameters are also discussed in Section 4.4.

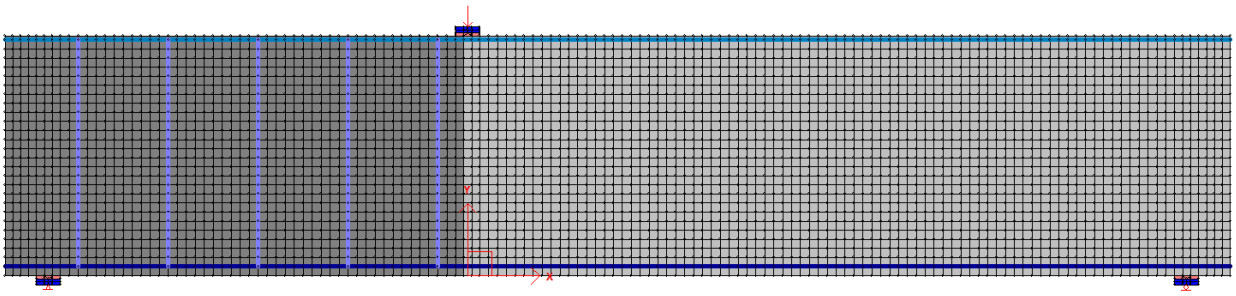


Figure 4-1 Finite element mesh for PLS4000

The stochastic results are shown in Figure 4-2 and summarized in Table 4-1. The load-displacement plots are shown in Figure 4-3. Based on the results, the recommendation set by Hunter of 200 trials seemed to have merit. From examining the ultimate loads, the results appeared to converge with more trials in terms of the average and the shape of the distributions. In contrast,

the 100 trial analysis deviated from the convergence and produced a more scattered distribution. The probable reason that the 100 trial analysis produced relatively different failure loads was due to the different distribution of concrete tensile strengths that was produced. PLS4000 was a shear-critical deep beam and thus the tensile strength of concrete was an important property. Lastly, the 100 trial analysis produced a higher coefficient of variation which converged as more trials were used.

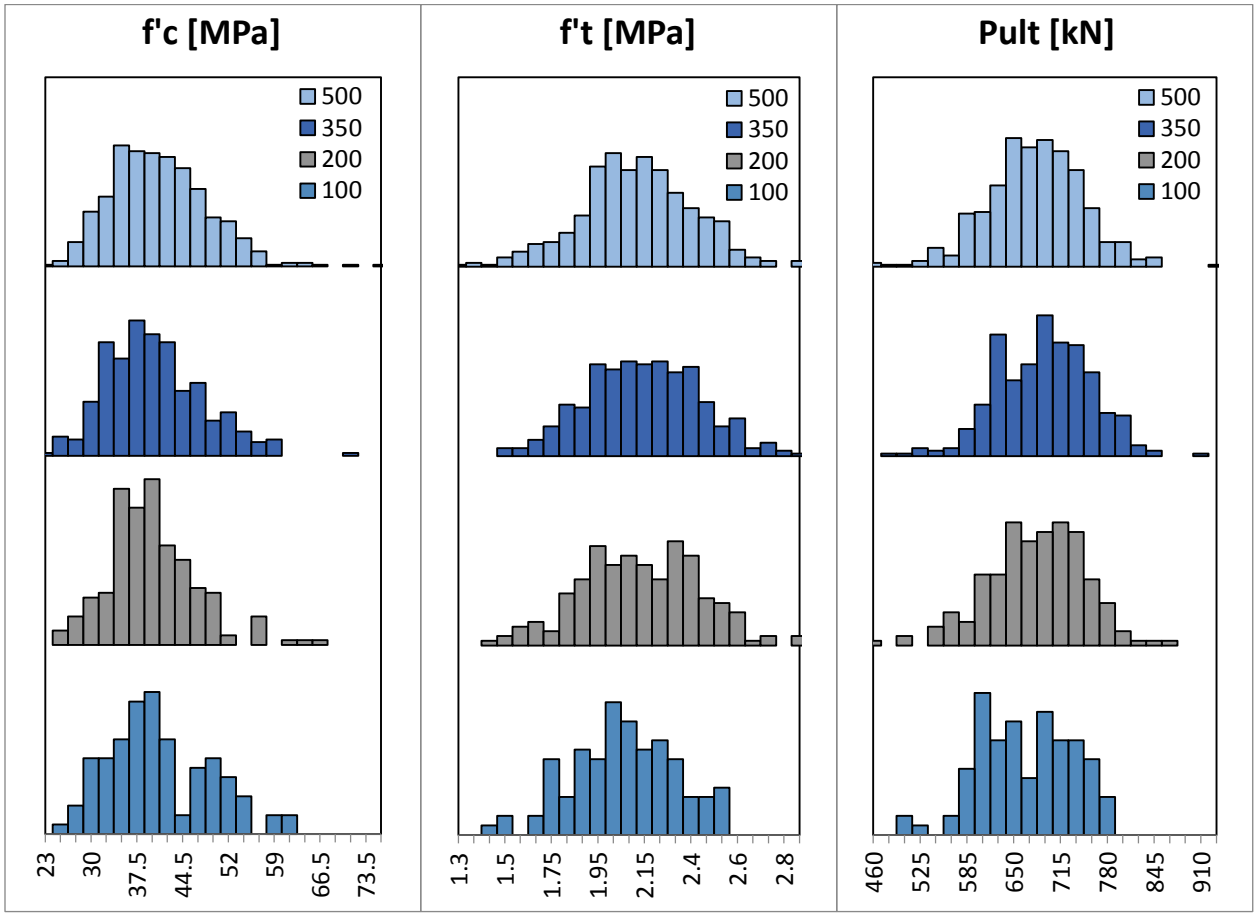


Figure 4-2 Stochastic results in terms of concrete properties and ultimate distributions for analysis of PLS4000 using 100, 200, 350, and 500 trials

Table 4-1 Stochastic averages and coefficient of variation for the trial sensitivity analysis of PLS4000

Number of Trials	f'_c [MPa]		f'_t [MPa]		P_{ult} [kN]		Time	Data
	Avg.	COV	Avg.	COV	Avg.	COV	[hr]	[GB]
100	39.7	20%	2.04	12%	678.2	9%	16	8.8
200	38.6	18%	2.11	13%	692.0	10%	30	17.6
350	38.9	19%	2.13	12%	700.2	9%	55	30.8
500	39.7	19%	2.09	13%	691.5	9%	81	44.1

As the results converged and produced marginal differences overall, the computational time and data storage required both increased linearly. Using an Intel® Core™ i7-3770 CPU @ 3.40 GHz processor with 8.00 GB of RAM, the analyses consistently averaged 16 hours per 100 trials. 8.8 GB of storage was also required per 100 trials. Therefore, 200 trials appeared to be an optimal number of trials since the addition of more trials did not provide meaningingly different results.

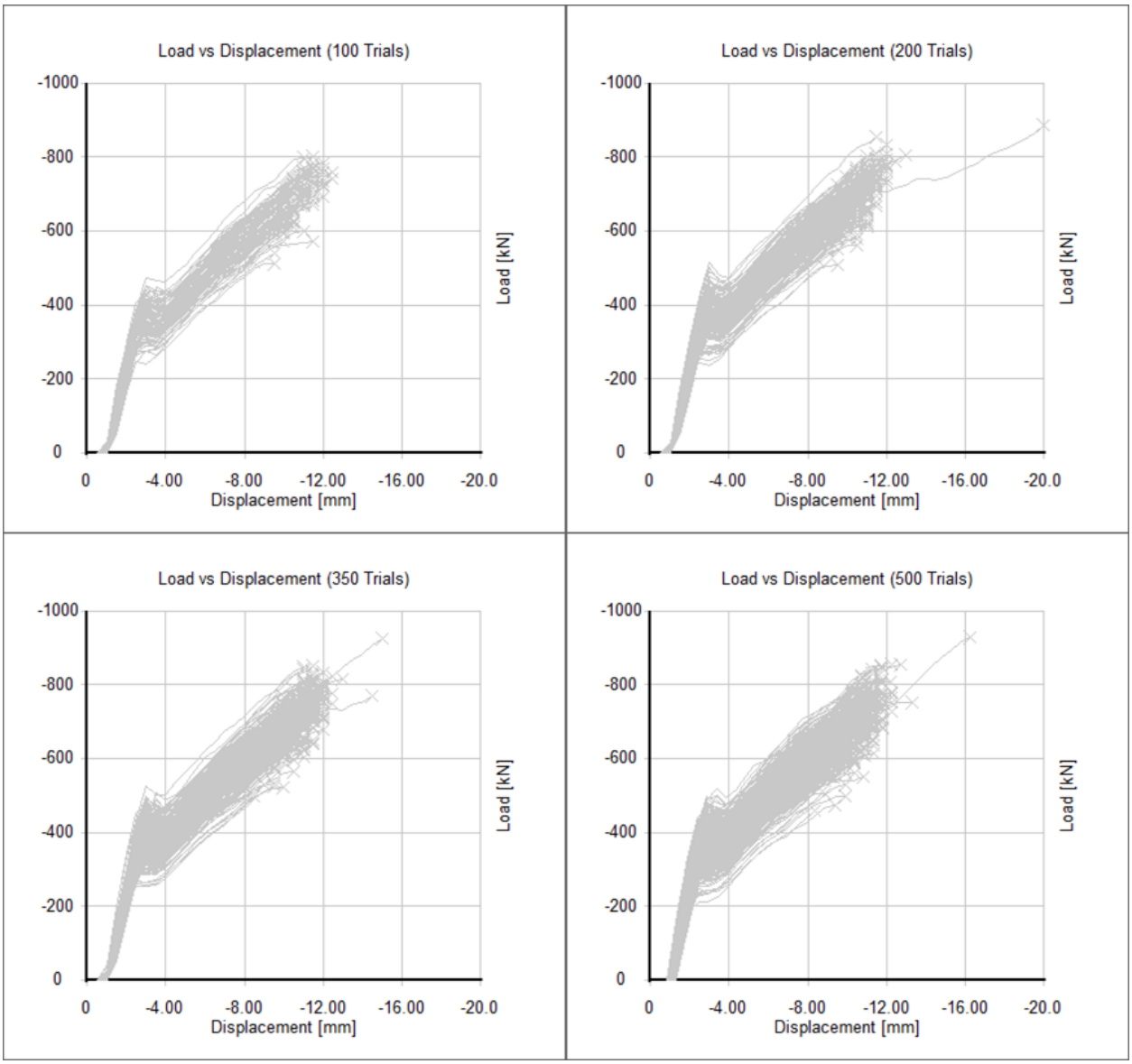


Figure 4-3 Load-displacement plots for the trial sensitivity analysis of PLS4000

It should be noted that other factors such as the number of elements in the model and the number of load stages used for each trial contributed to the large amounts of time and data required. PLS4000 was a relatively complicated model with approximately 4000 elements, two to four times

more than typical VecTor2 analyses. A summary of computational time and data storage required for all analyses is provided in a Section 4.6.

Ultimately, however, it was decided to use 250-300 trials for subsequent analysis. Although 200 trials appeared to provide adequate results, early testing of the field measurement incorporation methodology showed that more trials were required. As discussed with more evidence in Chapter 5, the methodology involves taking subsets of all trials. In general, the larger the total set, the larger the desired subset. As some analysis with 200 trials resulted in extremely small subsets, it was found that more trials were needed.

4.3 Finite Element Models

For each specimen, a finite element model was created using FormWorks, and both a deterministic analysis using test day material properties and a stochastic analysis using specified or 28-day cylinder strengths as reference strengths was performed. The selection of reference strength was determined by data availability. The default constitutive models in VecTor2 were used for all analyses as shown in Table 4-2.

Table 4-2 Default VecTor2 Material Models

Concrete Constitutive Models			
Compression Pre-Peak	Hognestad(Parabola)	**Dilation**	Variable-Isotropic
Compression Post-Peak	Modified Park-Kent	**Cracking Criterion**	Mohr-Coulomb (Stress)
Compression Softening	Vecchio 1992-A	**Crack Stress Calculation**	Basic (DSFM/MCFT)
Tension Stiffening	Modified Bentz 2003	**Crack Width Check**	Agg/2.5 Max Crack Width
Tension Softening	Bilinear	**Crack Slip Calculation**	Walraven
FRC Tension	SDEM-Monotonic	**Creep and Relaxation**	Not Considered
Confined Strength	Kupfer / Richart	**Hysteretic Response**	Nonlinear w/ Plastic offsets
Reinforcement Constitutive Models			
Hysteretic Response	Bauschinger Effect (Seckin)	**Buckling**	Akkaya 2012
Dowel Effect	Tassios (Crack Slip)	**Concrete Bond**	Eligehausen

Material properties other than the concrete compressive strengths and the reinforcement yield and ultimate strengths were input as the VecTor2 default values shown in Table 4-3 unless specified otherwise for select specimens. The philosophy was to conduct analyses with nominal properties despite more in-depths data from the experiments being available. The major material properties

are summarized later in Section 4.4. More details of each specimens can be found in their respective references.

Table 4-3 Default VecTor2 material property values

Concrete Models				
Thickness (mm)	Required Input	**C_c (/°C)**		10×10^{-6}
f'_c (MPa)	Required Input	**Max. agg. size (mm)**		20
f'_t (MPa)	$0.33\sqrt{f'_c}$	**Density (kg/m3)**		2400
E_c (MPa)	$3320\sqrt{f'_c} + 6900$	**K_c (mm2/s)**		1.2
ε_o (mm/m)	$1.8 + 0.0075 \cdot f'_c$	**Max Crack Spacing**	**S_x (mm)**	1000
ν_o	0.15		**S_y (mm)**	1000
Reinforcement Constitutive Models				
A_s (mm^2)	Required Input	**E_s (MPa)**		200000
D_b (mm)	Required Input	**ε_{sh} (mm/m)**		10
F_y (MPa)	Required Input	**ε_u (mm/m)**		180
F_u (MPa)	Required Input	**C_s (/°C)**		10×10^{-6}

4.3.1 Shim Beams

For the eight beams tested by Shim, each finite element model was generated in the following manner. First, half-span models were created to conserve computational time. This was justifiable due to the simple symmetrical construction and loading, and it was also the method elected by Shim (2004). The longitudinal reinforcement was modelled as discrete truss elements while any shear reinforcement was smeared in both the y- and z-directions. The loading plates were modelled as a layer of steel material with a layer of bearing material between the plate and the concrete to avoid artificial confinement to surrounding concrete regions. Anchorage plates at the ends of bottom longitudinal reinforcement were modelled using structural steel elements to match the experiments (Shim, 2004).

The depths of all beams were consistently around 560mm while the spans varied from 4100mm (series OA, and A) to 5010mm (series B) to 6840mm (series C) as seen in Figure 4-4. 16 elements were used through the depths of each beam with approximate element sizes of 40x40mm. This resulted in models with approximately 700, 800, and 1100 elements depending on the specimen span.

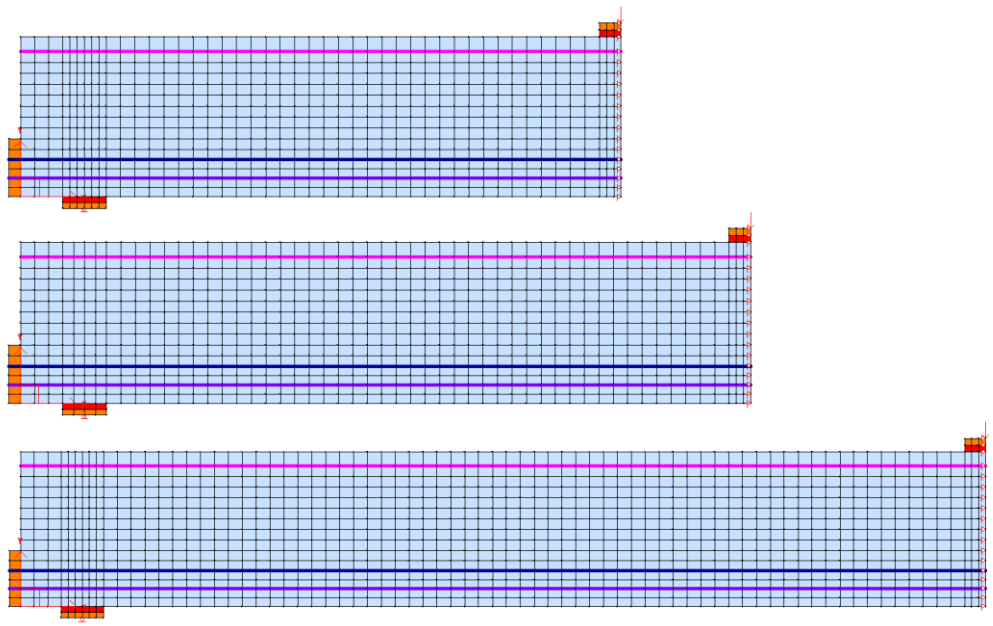


Figure 4-4 Typical finite element meshes for Shim beams: series A and series OA (top), series B (middle), series C (bottom)

The deterministic results of the eight beams by Shim can be seen in Figure 4-5. In most cases the match was very good especially Specimens OA3, A3, B2, B3, and C1. Specimen OA1 contained no shear reinforcement and failed in diagonal tension. Hence, the failure was heavily reliant on the tensile strength of concrete. Given that the formulation used for estimating the tensile strength of concrete tends to be conservative according to CSA A23.3, the premature failure can be explained. Figure 4-6 (a) shows the diagonal tension failure. The analysis for Specimen A1 can be somewhat misleading. Two peaks appear in the deterministic analysis with the second peak matching well with the experiment. However, after the first peak, significant diagonal cracking had occurred which did not lead to numerical failure but would have led to physical failure. Instead, the model developed a double curvature behaviour with vertical cracking near the top of the beam above the support seen in Figure 4-6 (b). This was due to the separation of the top and the bottom of the beam; this resulted from a lack of shear transfer where significant shear cracks developed but not in the uncracked regions near mid-span and supports. Nevertheless, the experiment never exhibited the artificial behaviour. The discrepancy between the first peak and the experiment, especially the post-crack stiffness, could be attributed to material properties. Specimen C2 failed due to extensive damage in the compression zone. Although the analytical model also showed extensive horizontal cracking in the compression zone, the failure load was much higher than observed.

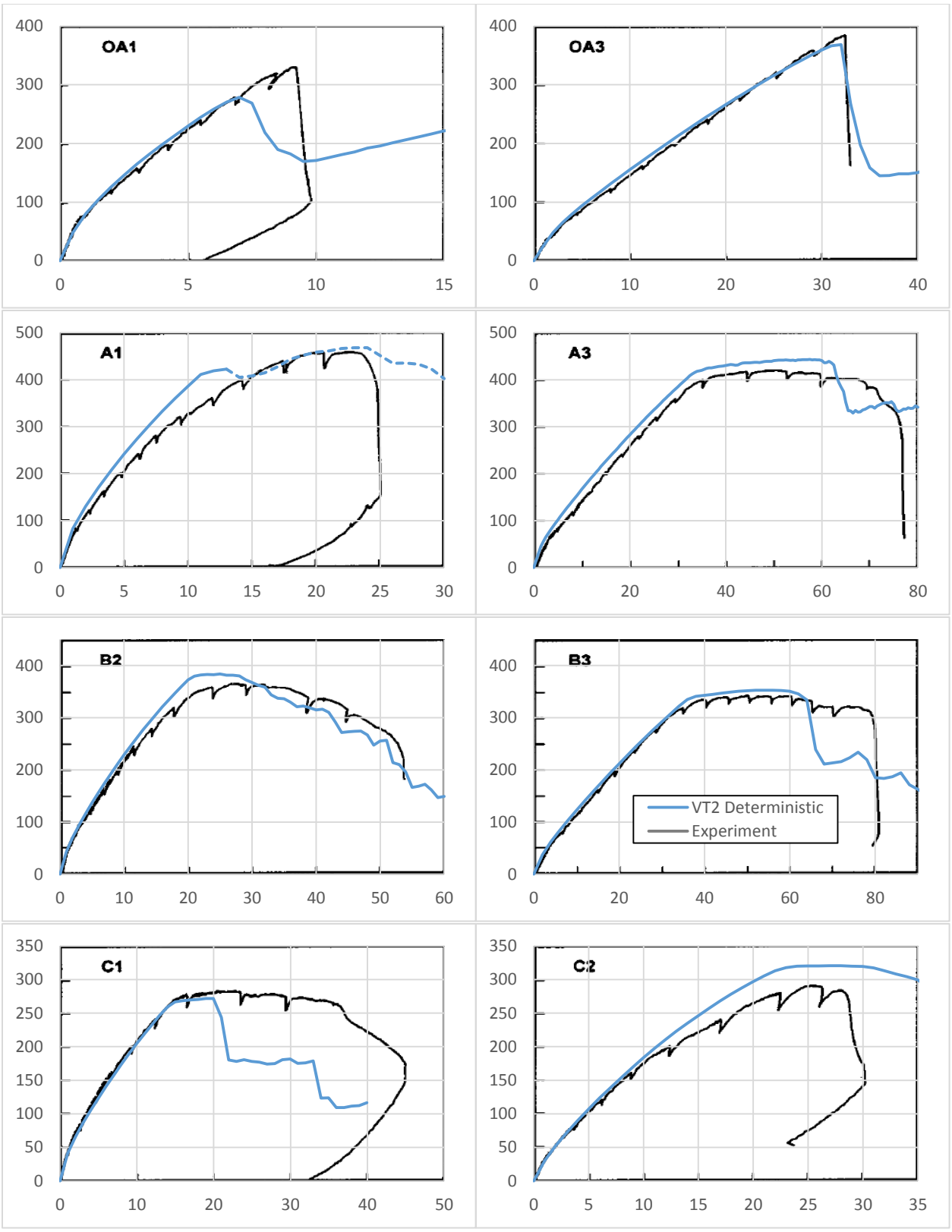


Figure 4-5 Deterministic load displacement results of all beams by Shim et al. (vertical axes: Load [kN]; horizontal axes: Displacements [mm]).

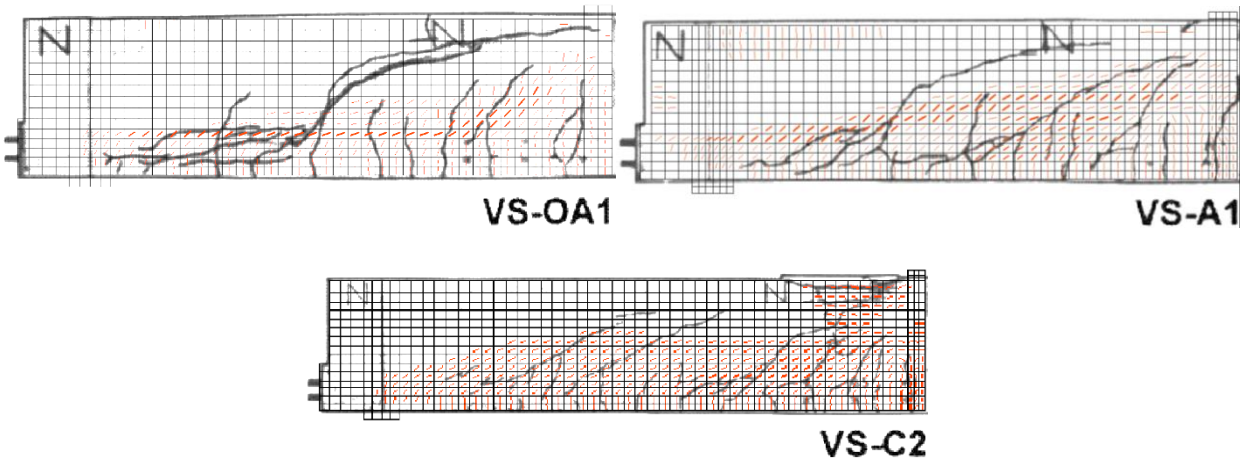


Figure 4-6 Comparison of experimental and analytical crack patterns at failure for Specimens VSOA1, VSA1, and VSC2

4.3.2 Lefas et al. and Oesterle et al. Shear Walls

Out of the eight shear walls analyzed, seven were by Lefas et al. (1990) and one by Oesterle et al. (1976). The seven shear walls by Lefas et al. were relatively small scale models with dimensions of 750 x 750 x 70mm (SW11, SW12, SW15, SW16) and 650 x 1300 x 65mm (SW21, SW22, SW23). The walls were rectangular in section with confined end zones. The specimens also contained both an upper and lower beam which were modelled as well. In the experimental setup, the bottom beam was anchored by two cross beams, on the sides of the shear wall, which were then bolted to the ground, as seen in Figure 4-7. To match the boundary conditions, pin supports were placed in groups on the sides of the shear wall at the top surface of the bottom beam to represent the cross bars (200mm or 8in. wide), and pin supports were placed along the bottom surface to represent the floor seen in Figure 4-8. A layer of bearing was added between the bottom surface and the pin supports to allow for uplift of the bottom beam from the floor, because the bottom beam was only 200mm thick and not extremely stiff in comparison to the shear wall. All reinforcement, vertical, horizontal, and hooped, was modelled as smeared reinforcement in the y-, x-, and z- directions. All models used approximately 30x30mm elements for the shear wall portion and 30x50mm elements for the upper and bottom beams. For the SW1 series of walls, 26 elements were used across the width and 1130 elements were used in total. For the SW2 series of walls, 23 elements were used across the width and 1480 in total. The vertical reinforcement provided was high-tensile bar that did not exhibit a distinct yield plateau. Thus, the steel material properties were altered from the default to match the provided stress-strain profiles. Finally, Specimens SW12, SW15, SW16, SW22, and SW23 were also loaded axially in addition to the

horizontal load. These loads were applied as nodal loads along the top beam in the portion above the shear wall.

Specimen B4 tested by Oesterle et al. (1976) is also shown in Figure 4-8 but at a much different scale. The wall dimensions were 1910 x 4570 x 102mm (75 x 15ft. x 4in.); approximately 3 to 5 times larger than the Lefas shear walls. The specimen had a barbell cross section with 305mm (12in.) square end zones. The bottom beam was anchored to the floor at four locations along the length of the bottom beam both in front of the shear wall as well as behind. The bottom beam was also sufficiently stiff (1220mm or 4 ft. thick) compared to the shear wall. Thus the boundary condition was modelled by pin supports along the entire bottom surface only. Approximately 100x100mm elements were used in the mesh with 22 elements across the width of the shear wall totalling 1250 elements.

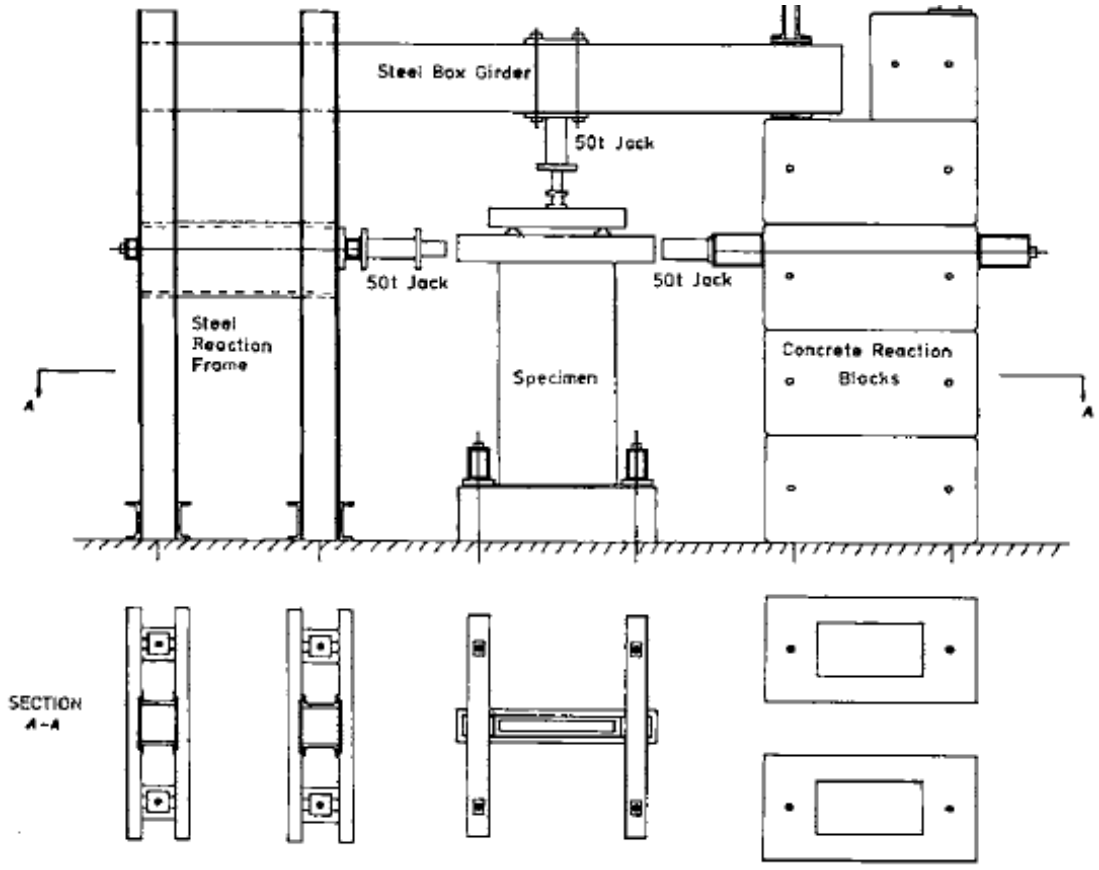


Figure 4-7 Experimental setup of Lefas et al. (1990) shear walls taken from Lefas et al. (1990)

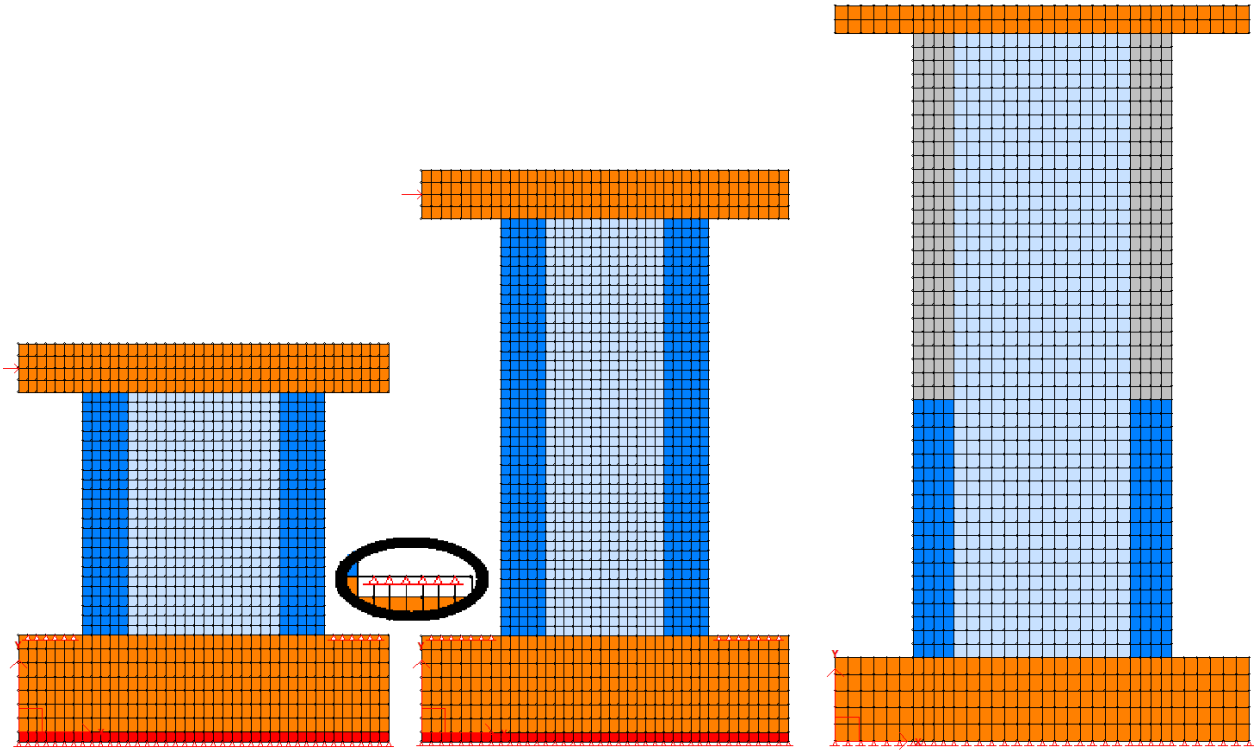


Figure 4-8 Typical finite element meshes for shear wall specimen series SW1 (Lefas et al., 1990), SW2 (Lefas et al., 1990), and Specimen B4 (Oesterle et al. 1976)

The results of all shear walls can be seen in Figure 4-9. According to Lefas et al., all walls failed due to failure of the compression zones. The four squat wall specimens (SW1 series) failed in diagonal compression while the more slender walls (SW2 series) failed more in flexural compression as seen in the distribution of concrete principal compressive stresses in Figure 4-10. Thus, the results of each analysis would depend on the concrete compressive strength. The test-day compressive strength was given in both cube strengths and cylinder strengths as seen in Table 4-4. The cylinder strengths were used for these analyses because it was the preferred option for VecTor2. According to the British Standard BS 1881-120 (1983), the typical cube to cylinder strength ratio is 1.2. In actuality, Specimens SW12, SW15, and SW22 had much higher ratios and produced some of the worst matches while Specimens SW11 and SW16 had the best matches. It could be possible that the cylinder tests conducted for the former were less accurate.

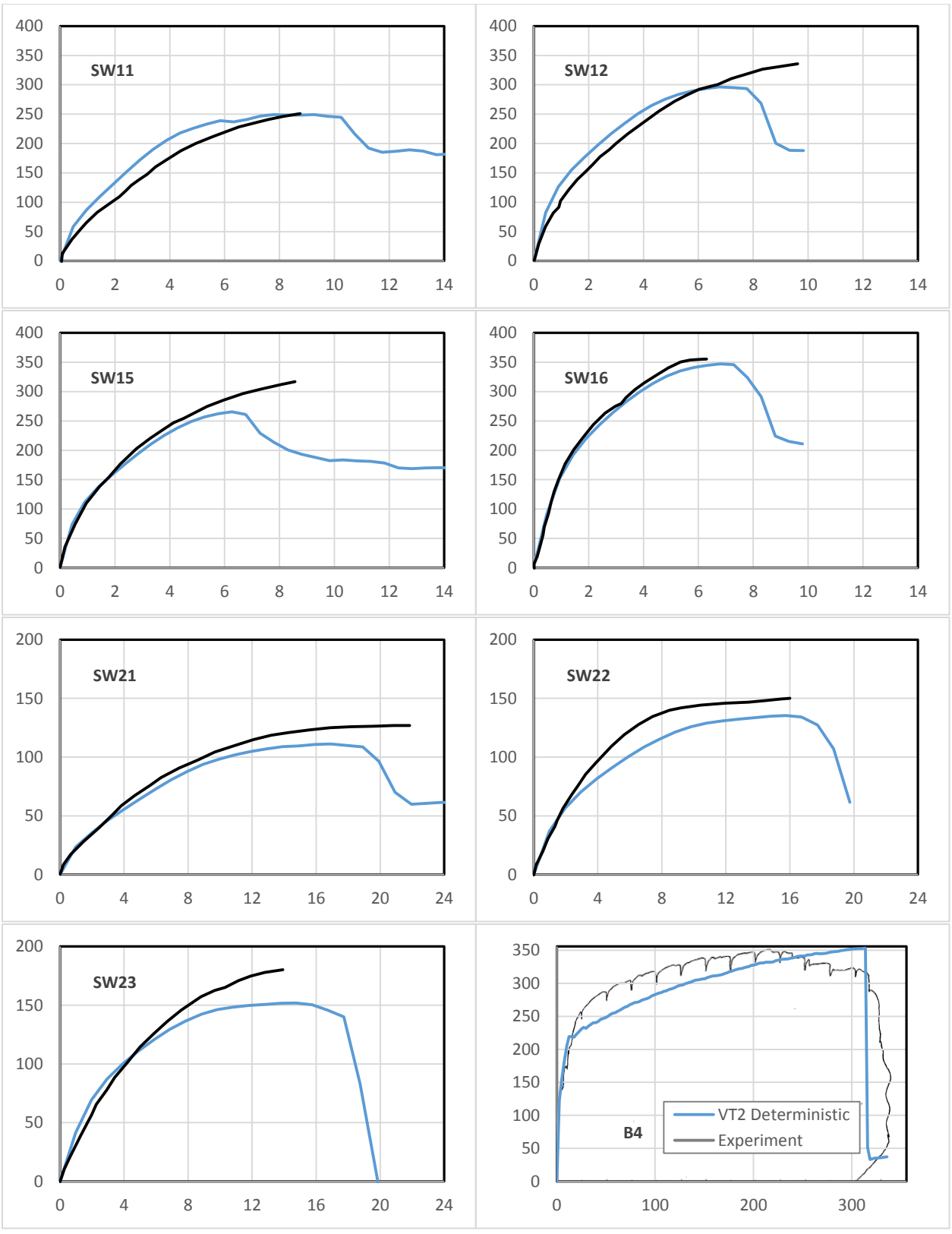


Figure 4-9 Deterministic load-displacement results for shear wall specimens by Lefas et al. and Oesterle et al. (vertical axes: Load [kN]; horizontal axes: Displacements [mm]).

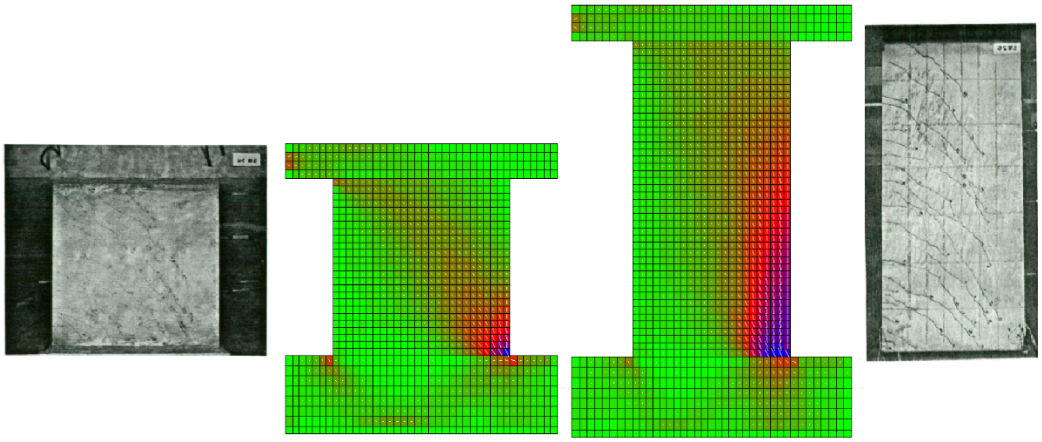


Figure 4-10 Analytical compressive stress patterns at failure for Lefas et al. shear walls compared to experimental crack patterns taken from Lefas et al. (1990)

Table 4-4 Concrete compressive strength of Lefas et al. (1990) shear walls

#	Specimen	Cube Strength [MPa]	Cylinder Strength [MPa]	Cube / Cylinder
9	**SW11**	52.3	44.2	1.18
10	**SW12**	53.6	39.4	1.36
11	**SW15**	43.3	30.5	1.42
12	**SW16**	51.7	43.6	1.19
13	**SW21**	42.8	34.3	1.25
14	**SW22**	50.6	35.3	1.43
15	**SW23**	47.8	38.5	1.24
			Average:	1.29

Specimen B4 failed due to rupture of the vertical reinforcement in the web region and end zones. It was reported that at the maximum horizontal load, the first reinforcement in the web fractured. As the displacement increased, more bars fractured corresponding to decreases in load with each fracture. Ultimately, the main vertical reinforcement in the tension zone ruptured resulting in a complete loss of load caring capacity. In the analysis, a flexural failure was also seen. However instead of a progressive failure, the smeared vertical reinforcement in the tension zone as well as a third of the smeared reinforcement in the web region all ruptured at the same time. This resulted in the instant loss of maximum load capacity rather a step wise behaviour. This was possibly due to the modelling of smeared reinforcement with uniform properties versus the experiment that had discrete rebars with varying ultimate strengths.

The overall behaviour was a better match than appeared. The behaviour through first cracking and first yield matched well. The point of full yield, however, symbolized by the dramatic decrease in

stiffness, occurred earlier in the analysis and led to dramatically different load-displacement plots. Nevertheless, the post-yield stiffness and the ultimate capacity were both well matched. If the yield strength was actually stronger than reported, then an analysis with a higher value could have produced a later yield point while maintaining a similar post-yield stiffness and matching the experiment more accurately.

4.3.3 Quach and Stanik

Six specimens tested by Stanik et al. (1998) and one by Quach et al. (2016) were examined. In all cases, except Specimen BM100 where minimal stirrups were provided, no shear reinforcement was provided. PLS4000 did have shear reinforcement in the west span; however, the failure was first reached in the east span. The finite element models utilized were similar to those for the Shim specimens. Except for PLS4000 which was asymmetrical, all beams were modelled as half-spans with discrete truss elements for longitudinal reinforcement. Shear reinforcement in PLS4000 was modelled as discrete truss elements while the shear reinforcement in BM100 was smeared. Specimen BN100D and BH100D also contained skin reinforcement which was smeared as well. The meshes used for Specimen BN100, BN100D, BH100D, and BM100 were identical and only differed in material specification and smeared reinforcement. Approximately 50x50mm elements were used with 21 elements through the depths of 1000mm, totaling 1314 elements. The same can be said for Specimens BN50 and BH50. Approximately 30x30mm elements were used with 17 elements through the depth of 500mm, totalling 902 elements. Specimen PLS4000 used approximately 50x50mm elements with 27 elements through the depth of 4000mm, totalling 4186 elements. It should be noted that the maximum crack spacing was changed from the default 1000mm to 4035mm. The crack spacing parameter influences crack width calculations which influence shear capacity across cracks. For concrete beams without shear reinforcement, the crack spacing value becomes similar to the shear depth. The default for the max spacing of 1000mm was not adequate for such a large specimen. The value of 4035mm was calculated based on the CIB-FIB code and the recommendation from Collins and Mitchell (1997) regarding unreinforced concrete beams. Figure 4-11 show the meshes used for typical specimens.

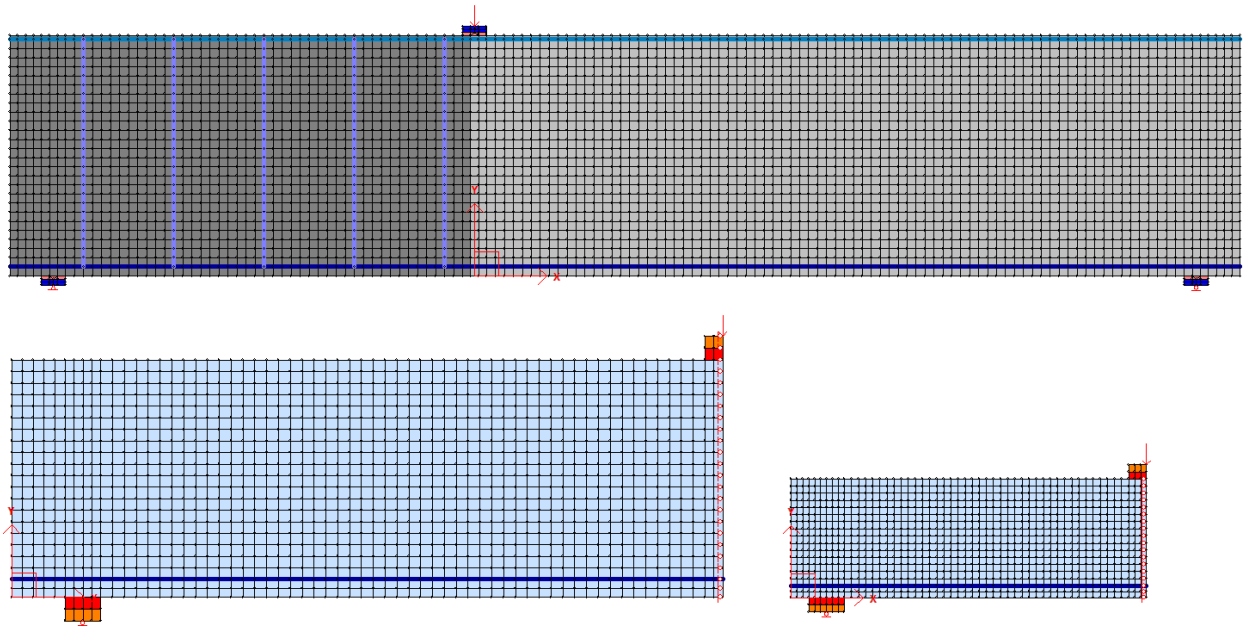


Figure 4-11 Typical finite element meshes used for Quach et al. (2016) beams (top) and Stanik et al. (1998) (bottom)

The results for the specimens are shown in Figure 4-12. As expected, all beams failed by diagonal tension failure which was also observed in the experiment. For many specimens, Stanik et al. reported that the dominant diagonal crack eventually propagated to the point of load application in one direction and along a plane parallel to the flexural reinforcement in the opposite direction. These results were reflected in the analyses as seen in Figure 4-13. The analyses of Specimens BH50 and BH100D produced much higher failures. Due to the nature of the failure, the tensile strength of concrete was highly influential. No direct tension tests were performed for the experiment but modulus of rupture (f_{cr}) tests were. For the normal strength concrete (BN and BM series), the reported f_{cr} was 3.2MPa. For the high strength concrete (BH series), the reported f_{cr} was 4.0MPa. The formula for predicting f_{cr} according to CSA A23.3 is $0.6\sqrt{f_c'}$, which would give conservative values of 3.65MPa and 5.97MPa respectively given the test-day concrete cylinder strengths. Although not a direct reflection on the direct tensile strength, it shows that the experimental tensile properties were much lower than the analytical values, especially for the BH series.

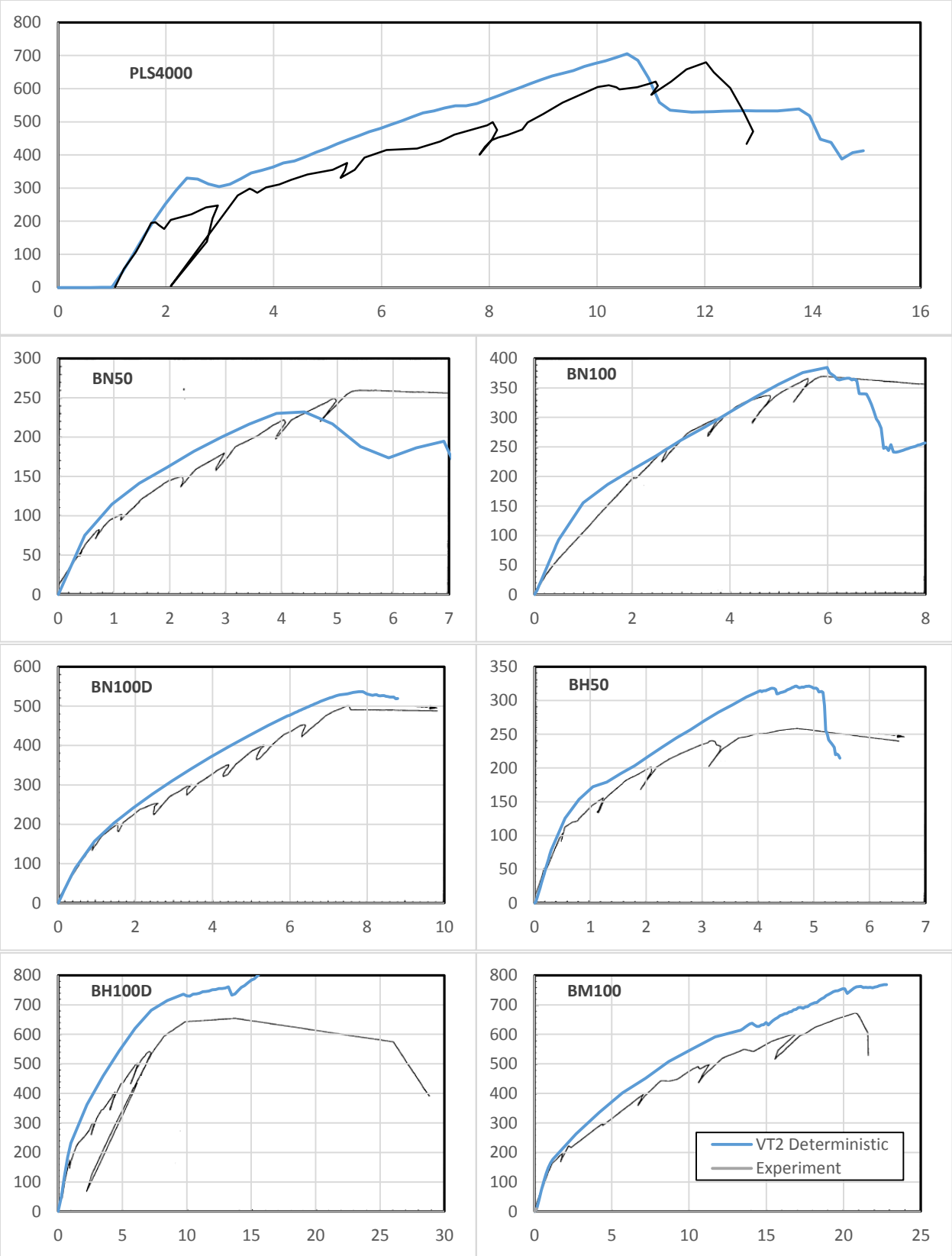


Figure 4-12 Deterministic results for beams by Stanik et al. (1998) and Quach et al. (2016) (vertical axes: Load [kN]; horizontal axes: Displacements [mm]).

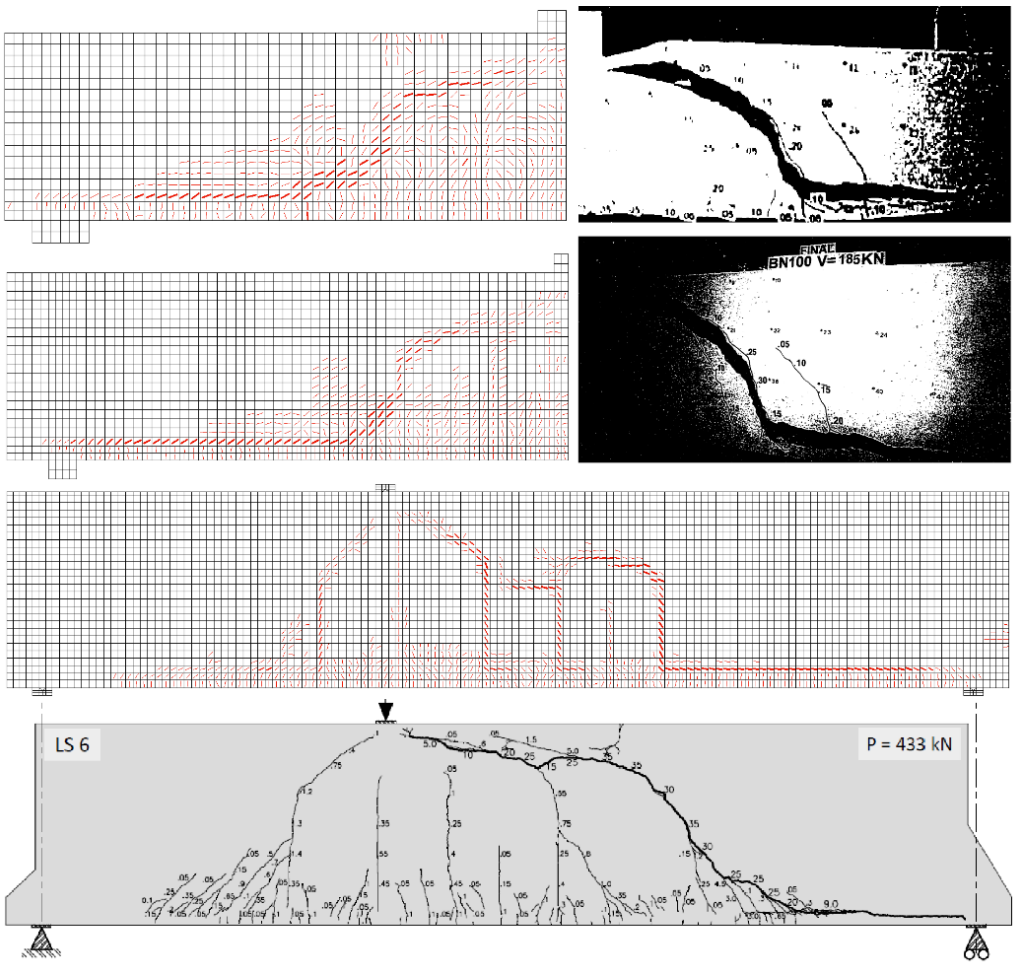


Figure 4-13 Crack pattern for Specimen BN50, BN100 taken from Stanik et al. (1998) and PLS4000 taken from Quach et al. (2015) at failure

Varying degrees of match were seen in the deterministic analyses. In most cases, the overall behaviour was well captured in terms of failure mechanisms and progression of nonlinear behaviour. In terms of the ultimate load comparisons, the maximum absolute error produced was 24% while the average of all 23 specimens was 9.5%. The average ratio between analyses to experiments was 0.993 as some specimens were overestimated while some were underestimated. However, despite these results, improvements could be made in many cases perhaps with different material property values. Conversely, different values could very well produce the opposite effect. In either case, the deterministic analyses did not illustrate how reliable and how sensitive each analysis was.

4.4 Stochastic Analysis

Essentially the same finite element models used for the deterministic analyses were used for the stochastic analyses. The new inputs required for the stochastic analyses are shown in Table 4-5.

Table 4-5 Default stochastic analysis parameters

Stochastic Analysis Parameters		
Stochastic Analysis Type		Monte Carlo Simulation (MCS)
Number of Simulations (Trials)		300 / 250 / 500
Concrete Material Property Distribution Models	**Compressive Strength**	Bartlett and MacGregor (1996)
	Tensile Strength	Mirza et al. (1979)
	Elastic Modulus	Hybrid Mirza (1979) + CSA
Steel Material Property Distribution Models	**Yield Strength**	Nowak and Szerzen (2003)
	Ultimate Strength	Mirza et al. (1979)
	Elastic Modulus	Mirza et al. (1979)
Reference Strength		Specified / 28-Day Cylinder*
Age (Days)		Required Input

**For specimens tested by Shim*

For this study, the material distribution models used were those set as default by Hunter (2016). The scope was limited to Monte Carlo simulations in lieu of Latin Hypercube and Correlated sampling. Spatial variation was also not considered. The focus was to examine the default type of stochastic analysis first.

The only other difference between the finite element models was the material properties. Table 4-6 shows the major material properties of all 23 specimens. For each property, the test day values were used for deterministic analysis. For concrete tensile strengths, the deterministic value was calculated using the equation $f_t' = 0.33\sqrt{f_c'}$. For each stochastic analysis, the material property distributions were generated by using the specified/28-day cylinder values as reference strengths, and the age of the specimen. The mean and coefficient of variation are shown for the generated distributions.

In the specimens tested by Shim et al. (2004), the specified concrete strengths were not provided as the mixes were designed to match the original experiments exactly. The 28-day cylinder strengths were provided and used instead. In all other cases, the specified strength was used as the preferred concrete reference strength. Also, the specimens by Shim et al. and Stanik et al. did not provide specified steel strengths and reported rebar coupon results only. For those specimens, the

steel material properties were kept deterministic. In hindsight, distributions with the mean set to the provided steel strengths could have been used.

Table 4-6 Material properties of all specimens

#	Specimen Name	f'_c [MPa]						f'_t [MPa]			f_y [MPa]			
		Spec	28 Day	Test Day	Age [days]	Stochastic Avg.	Stochastic COV	0.33√f'c	Stochastic Avg.	Stochastic COV	Spec	Test Day	Stochastic Avg.	Stochastic COV
1	**OA1^{Sh}**	-	22.5	22.6	38	24.2	13%	1.57	1.63	12%	-	440	440	-
2	**OA3^{Sh}**	-	39.0	43.5	127	45.8	14%	2.18	2.25	12%	-	440	440	-
3	**A1^{Sh}**	-	22.5	22.6	38	24.0	15%	1.57	1.64	12%	-	440	440	-
4	**A3^{Sh}**	-	39.0	43.5	127	46.1	14%	2.18	2.25	12%	-	440	440	-
5	**B2^{Sh}**	-	25.9	25.9	51	28.1	13%	1.68	1.77	14%	-	440	440	-
6	**B3^{Sh}**	-	39.0	43.5	127	46.7	14%	2.18	2.26	13%	-	440	440	-
7	**C1^{Sh}**	-	22.5	22.6	38	23.9	14%	1.57	1.63	12%	-	440	440	-
8	**C2^{Sh}**	-	25.9	25.9	51	28.4	14%	1.68	1.76	13%	-	440	440	-
9	**SW11^{L}**	37.5	-	44.2	87	54.6	19%	2.19	2.44	13%	460	470	525	5%
10	**SW12^{L}**	37.5	-	39.4	77	54.2	18%	2.07	2.43	13%	460	470	527	4%
11	**SW15^{L}**	37.5	-	30.5	48	50.9	19%	1.82	2.37	13%	460	470	530	5%
12	**SW16^{L}**	37.5	-	43.6	46	52.4	18%	2.18	2.35	13%	460	470	527	5%
13	**SW21^{L}**	37.5	-	34.3	55	49.4	18%	1.93	2.33	12%	460	470	525	5%
14	**SW22^{L}**	37.5	-	35.3	49	51.2	18%	1.96	2.40	13%	460	470	524	6%
15	**SW23^{L}**	37.5	-	38.5	46	52.4	19%	2.05	2.39	13%	460	470	524	5%
16	**B4^{Oe}**	41.4	-	45	68	59.0	18%	2.21	2.55	13%	414	450	475	5%
17	**PLS4000^{Q}**	30	-	40	49	39.7	19%	2.09	2.09	13%	500	573	570	5%
18	**BN50St**	25	-	37	35	33.8	18%	2.01	1.93	12%	-	550	550	-
19	**BN100St**	25	-	37	42	34.2	19%	2.01	1.95	13%	-	550	550	-
20	**BN100DSt**	25	-	37	40	33.6	20%	2.01	1.92	13%	-	550	550	-
21	**BH50St**	60	-	99	33	79.7	20%	3.28	3.01	12%	-	550	550	-
22	**BH100DSt**	60	-	99	35	79.7	19%	3.28	2.99	14%	-	550	550	-
23	**BM100St**	25	-	46	119	37.6	17%	2.24	2.05	13%	-	550	550	-

4.4.1 Material Property Distribution Predictions

Examining the initial distributions of the material properties, some specimens exhibited better matches than others. The ratios of the average stochastic concrete compressive strengths to the test day concrete strengths are shown in Figure 4-14. For the first eight specimens, tested by Shim et al., the average ratio was 1.07. For the shear walls tested by Lefas et al. (1990) and Oesterle et al. (1976), the average ratio was considerably higher at 1.38, showing a large discrepancy between

the predicted strengths and the actual strengths. Finally, the beams tested by Stanik et al. (1998) and Quach et al. (2016) had an average ratio between the average stochastic concrete compressive strengths to the test-day concrete strengths of 0.82. Among the seven specimens, the normal strength concrete (below 40MPa) had an average ratio of 0.91 while the high strength concrete had an average ratio of 0.81.

The immediate explanation could be drawn from the experiments themselves. The tests by Shim et al., Stanik et al., and Quach et al. were all conducted in the 2000s and were all with Canadian concrete. Meanwhile the tests by Lefas et al. were conducted in 1990 at Imperial College in the United Kingdom and the tests by Oesterle et al. were conducted much earlier in 1976 in the United States. The distribution model used for concrete compressive strengths was by Bartlett and MacGregor from 1996 and deduced from experiments on Canadian concrete. The experiment involved 3756 cylinder tests from 108 concrete mixes produced in Alberta, Canada between 1988 and 1993 (Bartlett and MacGregor, 1996). Also, the study was conducted for concrete with strengths less than 55MPa which could explain the inconsistency with the high strength concrete used by Stanik et al.

Furthermore, the large ratios in the specimens by Lefas et al. could be due to the conversion between cylinder and cube strengths. The specified strength provided was a cube strength of 45MPa. As mentioned, the test-day strengths however, were provided for both cube and cylinder strengths in Table 4-4. In the deterministic analyses, the cylinder strengths were used as the test-day strengths. For stochastic analysis, the only usable reference strength was the cube strength of 45MPa. Therefore, a conversion from dividing by a factor of 1.2 (BS 1881-120, 1983) was required. However, this factor may not always be accurate. The average ratio of the seven specimens was actually 1.29 as seen in Table 4-4. However, the 1.2 ratio was maintained to represent a purely nominal point of view.

4.4.2 Stochastic Analysis Ultimate Load Distributions

The summary of the ultimate load distributions for all 23 specimens can be seen in Figure 4-14. The three dashes indicate the average, and plus/minus one standard deviation, ultimate loads divided by the experimental ultimate load. Also presented are the ratio of the deterministic analysis results over those from the experiment. And finally, the ratios of the average stochastic prediction of concrete strength versus the test day strength are given to show the general idea of how accurate

the material predictions were for each analysis. The results are also shown in Table 4-7 with the general failure mode of each specimen. Finally, the load-displacement plots for each specimen are shown in Figure 4-15 and Figure 4-16. Each plot shows all trials generated stochastically with the failure points determined, the experimental result, and the deterministic result.

For specimens that had good material matches with the test-day values, such as the eight tests by Shim et al., the difference between the deterministic analyses and the average of the stochastic analyses were understandably minimal. For the seven specimens by Lefas et al., where the predicted material distributions were much higher, a similar trend of higher ultimate loads was seen. In the case of Specimens SW12, SW15, SW21, SW22, and SW23, this produced more accurate stochastic predictions than the deterministic ones. Although this would suggest the possibility of errors in the provided concrete cylinder strengths, there was not enough evidence to support such a claim since the stochastic analyses for Specimens SW11 and SW16 were unconservative. A similar but opposite effect was seen in the high strength concrete tests by Stanik et al. The material properties were under-predicted which led to stochastic analyses that produced weaker, and perhaps by happenstance, more accurate responses.

In general, the deviation between the deterministic analysis and the average stochastic analysis followed the deviation between the test-day material properties and the stochastically predicted distributions of the material properties. This was an obvious conclusion and not one of much importance. The average failure load was not the most important output of the stochastic analysis. Instead it was the overall distribution and the combination of material properties producing different responses. However, it was promising that the mean error in the average stochastic predictions was 7.1% versus the 9.5% for the deterministic analyses. The mean experimental ratio was 1.037. Furthermore, the gap was much larger for deterministic analysis with specified properties, which produced a mean error of 16.6% and mean experimental ratio of 0.857.

The results showed that the material distribution models using nominal properties as references were adequate and represented a major improvement to deterministic analysis using only nominal material properties. Of course, in the structural evaluation of reinforced concrete structures, strengths change is only one aspect of uncertainty with others being deterioration effects such as corrosion and concrete expansion. However, the results showed a promising validation before the combination of other factors.

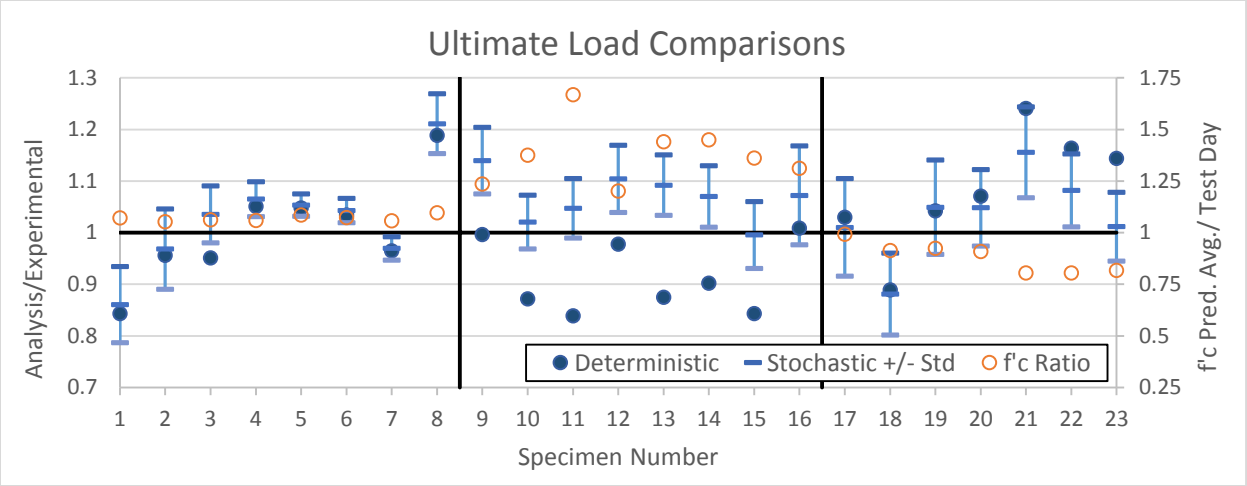


Figure 4-14 Comparison of deterministic and stochastic ultimate loads against experiment. Also shown are the ratios between the stochastically generated concrete strength over the test day strengths

Table 4-7 Summary of experimental (Exp.), deterministic (Det.) failure loads, and stochastic failure load distributions (Mean. and COV) along with failure modes

#	Specimen Name	P_{ult} [kN]				Failure Mode
		Exp.	Det.	Mean	COV	
1	**OA1^{Sh}**	165	139	142	8.6%	Diagonal Tension, Splitting along Tension Reinforcement
2	**OA3^{Sh}**	193	184	186	8.0%	
3	**A1^{Sh}**	229	217	237	5.3%	Concrete Crushing due to Diagonal Shear Compression
4	**A3^{Sh}**	211	222	225	3.2%	Concrete Crushing due to Flexural Compression
5	**B2^{Sh}**	183	192	193	2.1%	Concrete Crushing due to Flexural and Shear Compression
6	**B3^{Sh}**	171	176	178	2.3%	Concrete Crushing due to Flexural Compression
7	**C1^{Sh}**	141	136	137	2.4%	Concrete Crushing due to Flexural and Shear Compression
8	**C2^{Sh}**	145	172	175	4.8%	
9	**SW11^{L}**	250	249	285	5.7%	Concrete Crushing due to Diagonal Shear Compression, Cover Spalling
10	**SW12^{L}**	340	296	347	5.1%	
11	**SW15^{L}**	317	266	332	5.5%	
12	**SW16^{L}**	355	347	392	5.9%	
13	**SW21^{L}**	127	111	139	5.4%	Concrete Crushing due to Flexural Compression, Cover Spalling
14	**SW22^{L}**	150	135	161	5.6%	
15	**SW23^{L}**	180	152	179	6.5%	
16	**B4^{Oe}**	347	350	372	8.9%	Flexural Failure, Rupture of Tension Reinforcement
17	**PLS4000^{Q}**	685	705	692	9.3%	Diagonal Tension, Splitting along Tension Reinforcement
18	**BN50St**	130	116	115	9.0%	
19	**BN100St**	184	192	193	8.7%	
20	**BN100DSt**	251	268	263	7.1%	
21	**BH50St**	130	161	150	7.6%	
22	**BH100DSt**	327	380	354	6.5%	
23	**BM100St**	336	384	340	6.6%	

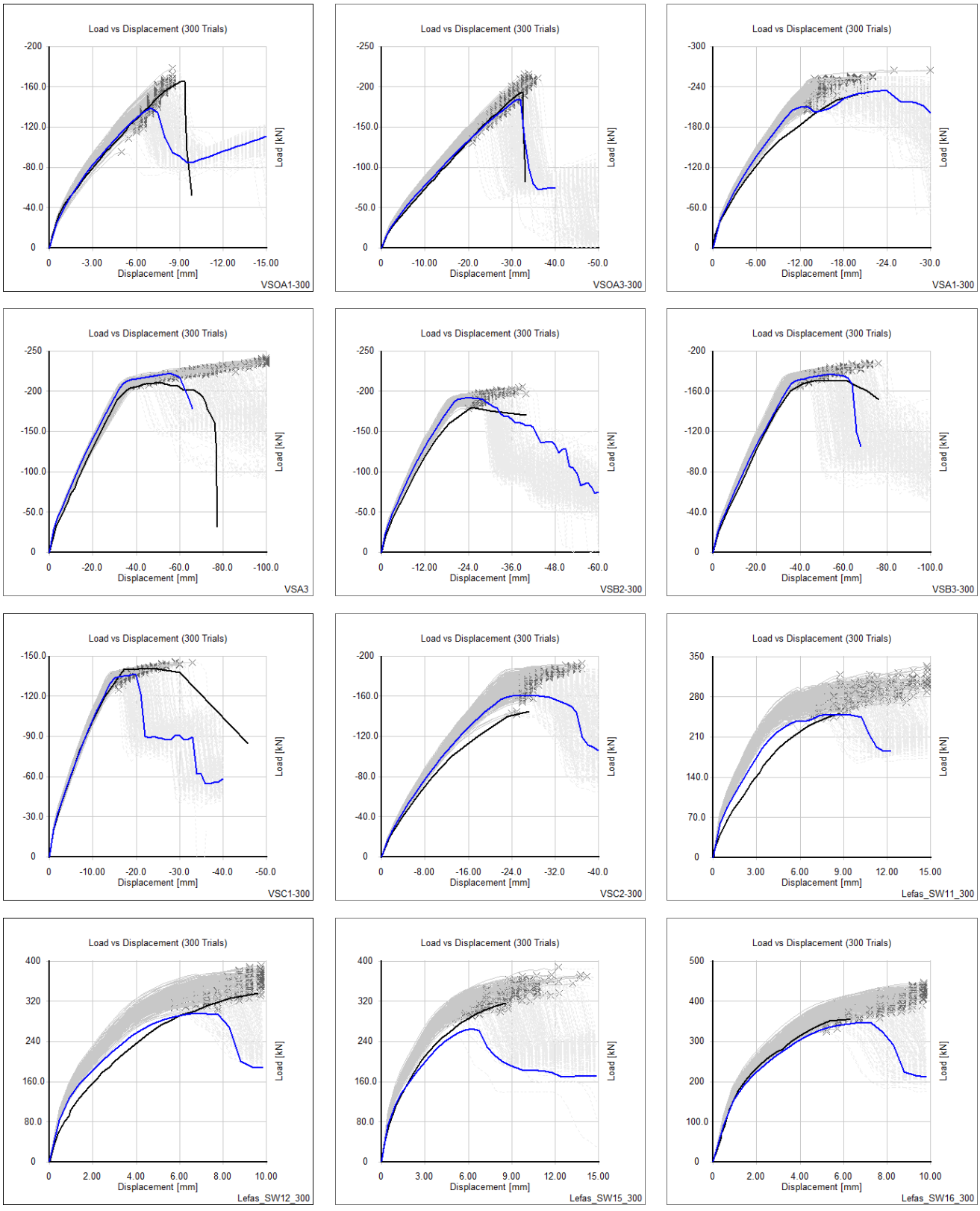


Figure 4-15 Load-displacement results of all specimens including experimental (black) and deterministic (blue) and stochastic (grey) results

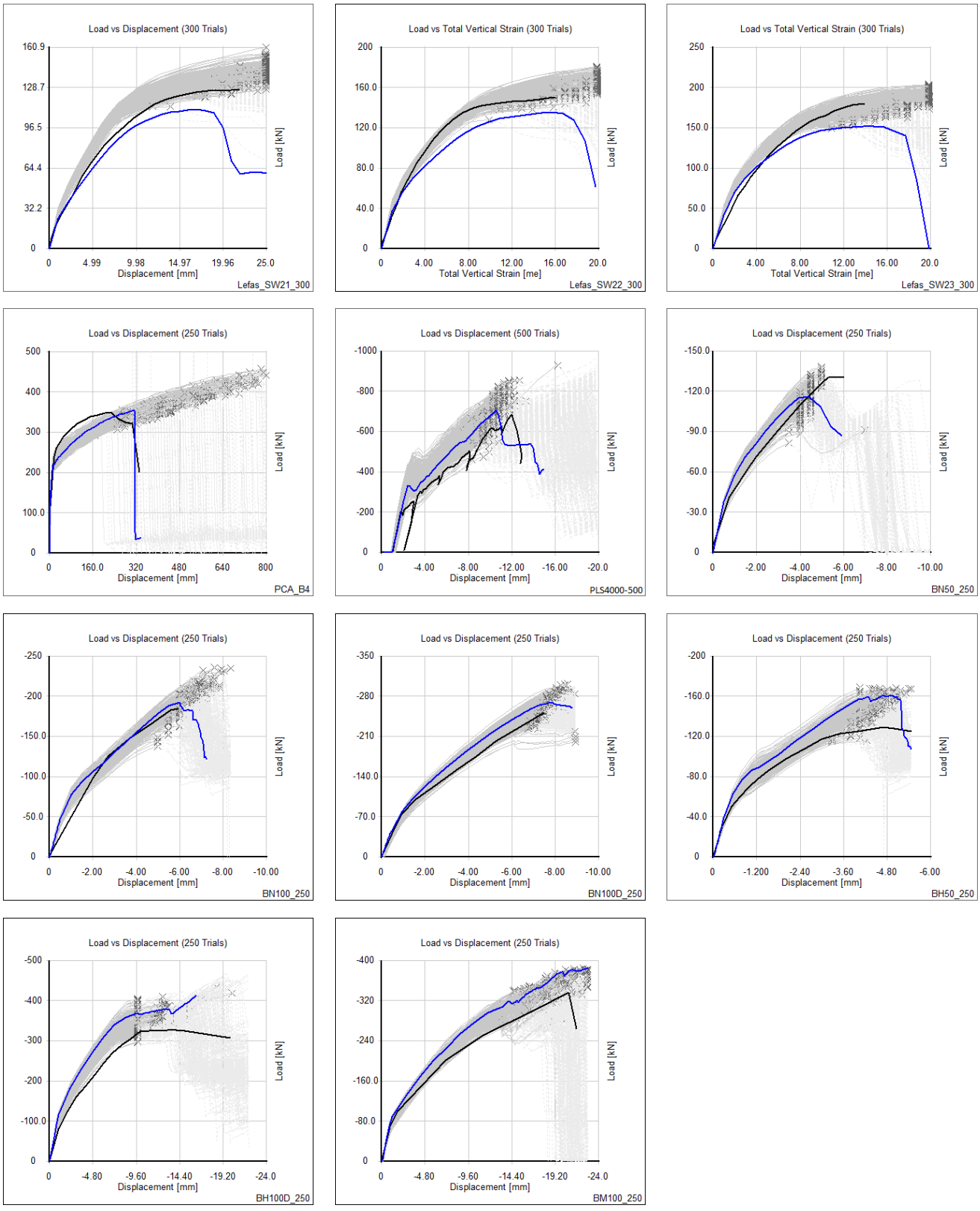


Figure 4-16 Load-displacement results of all specimens including experimental (black) and deterministic (blue) and stochastic (grey) results (cont'd)

4.4.3 Examination of Failure Modes

The results showed that different specimens produced different variances in their failure loads. The only variance in the model inputs were the material properties and, across different specimens, the variation was largely the same as seen in Table 4-8. Therefore, the type of specimen, or more importantly the type of failure, must be dictating the variance, or reliability, of the predictions. Table 4-7 showed the dominant failure modes associated with each experiment. Four major modes were identified as: diagonal tension failure, compression failure due to shear, compression failure due to flexure, and flexural failure due to reinforcement rupture. Table 4-8 summarizes the average coefficient of variance of each type of failure.

Table 4-8 Aggregation of specimens by failure modes

Failure Mode	Mean COV	Count
Diagonal Tension	7.9%	9
Compression due to Shear	4.6%	8
Compression due to Flexure	4.6%	5
Flexural Failure due to Rupture	8.9%	1

Diagonal tension failures are common in shear-critical specimens without shear reinforcement. As a reinforced concrete specimen is loaded in shear, tensile stresses develop perpendicular to the compression stress fields resisting the loading. As the tensile stresses exceed the tensile strength of concrete, major diagonal cracks begin to form. Without shear reinforcement providing tensile support across the cracks, the cracks widen and propagate. Although shear forces can be carried across concrete cracks, the capacity is severely reduced as the cracks develop. The failure ultimately results from the propagation of the diagonal cracks and loss of load carrying capacity of the specimen. The analysis of such failures can be quite challenging due to the many mechanisms involved. The development of diagonal shear cracks, including the path and crack angles, requires robust finite element analysis tools. Simple closed-form solutions are often inadequate. The shear stress carrying capability across concrete cracks has produced disagreements amongst researchers in determining the best models. Therefore, it was expected that this type of failure would produce relatively larger variances.

Compression or concrete crushing failure was the most dominant form seen in the specimens examined. This form of failure occurs when the compressive stress in critical zones exceed the concrete compressive strength and crushing of the concrete occurs. There are other mechanisms

also involved. Compression softening is the reduction in compressive strength and stiffness due to coexisting transverse cracking and tensile straining (Vecchio and Collins, 1986). Confinement of the compression zones can be thought of as the opposite effect where confinement stresses, due to reinforcement or adjacent boundary elements, reduce transverse tensile effects and enhance the strength and ductility of concrete. In the specimens examined, two types of compression failures were seen: compression due to diagonal strut action to resist shear forces, and compression due to bending moments. In both cases the failure mechanism is the same. The difference is the effective area in which the compressive forces are carried. In the flexural case, the formulation for the compression zone depth has been well developed in the application of sectional analysis. For shear governing behaviour, the use of strut and tie models have been developed. However, the formulation for strut width and strut angle, such as the one in CSA A23.3, is more arbitrary, capable of producing a wide range of results. This is because struts, uniaxially stressed truss members, are just simplified representations of complex stress flows in disturbed regions (Collins and Mitchell, 1997). Thus, it was expected that different geometries and different material properties would provide a wider range of shear compression failures compared to flexural compression failures. Specimens A3 and B3 were the most flexural-dominant cases with their large span to depth ratios and did produce two of the smallest COVs. However, the largest COVs of the specimens that failed in compression were produced by the slender shear walls SW21, SW22, and SW23 which also produced flexural-dominant behaviour. It should also be noted that the low COV values of Shim et al. specimens were definitely affected by the deterministic modelling of the longitudinal reinforcement. In many cases, however, the type of compression failure was not as distinct. Shim et al. noted that for specimens such as Specimen C2, the behaviour was shear-flexural in nature (Shim, 2004). Figure 4-17 shows varying combinations of flexural and shear behaviour.

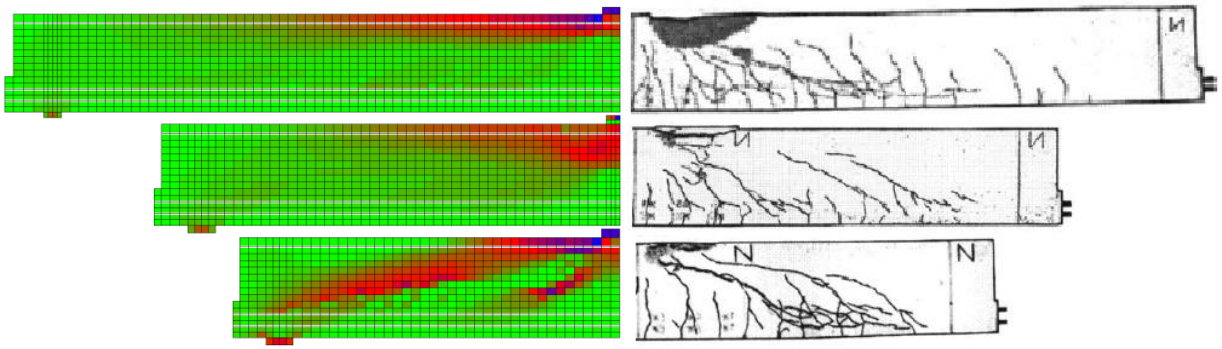


Figure 4-17 Variations of flexural compression and shear compression failures taken from Shim et al. (2004)

The last failure mode examined was the flexural failure of Specimen B4. Despite flexural response being well defined, the variance for Specimen B4 was one of the largest. This was because although the failure mechanism was well defined, the strain hardening and rupturing behaviour of reinforcement bars can vary greatly. For main vertical reinforcement in the tension zone, the distribution model by Mirza and MacGregor (1979) produced an average ultimate strength of 726MPa with a COV of 10.6% using a reference strength of 600MPa. The reported ultimate strength was 707MPa.

These results showed that with the same material property prediction models, different failure load distributions can be produced due to the types of failures. This is an important observation to the value of stochastic analysis. With the deterministic analyses, there was no information regarding how reliable the results were. With the stochastic results, it can be seen that a diagonal tension failure is harder to predict than a flexural compression failure. In certain cases, the conclusions may appear obvious. But it must be noted that these laboratory experiments were designed to produce certain types of failures with idealized loading and boundary conditions. For more complex analyses, the resulting variance will not be so self-evident. Instead it should be seen as the validation of the first step in steadily introducing more levels of uncertainty. Furthermore, it allows the analyst to quantify the reliability of each specimen individually rather than relying on code factors generated to encapsulate all cases. However, much more evidence will need to be provided.

4.4.4 Combination of Failure Modes

It was hoped that through stochastic analysis, different failure modes could be produced for the same specimen. There lies the power of stochastic analysis in generating possible but unexpected

outcomes. Although distributions of different failure loads were seen, in all cases, the failure mechanisms were the same and the shapes of the load-displacement response were the same with slight variations in stiffness values and critical points.

The first reason for the lack of different failure modes was the level of variability or degrees of freedom input into the analysis. Although material properties influence failure, there is still a reasonable range of variation from the reference strength. Other effects such as corrosion and shrinkage could produce more variations in response. Boundary conditions are another source of major uncertainty that could produce different failure modes. The second reason was that the specimens examined were all laboratory experiments designed for specific failure modes. Undesired failure modes would have been overdesigned for, whereas in design, more optimization would be expected. For example, Specimen VSOA1 had a flexural capacity of 404.1kNm and would have failed at 441.6kN compared to the observed shear failure at 330kN. Therefore, it would have been unlikely to produce a trial with properties that delayed the shear failure while maintaining the same flexural failure. Hence, very little combination of failure modes was seen.

4.4.5 Best Fit Trials

Despite the lack of failure mode combinations, the random combination of material properties in the stochastic analyses did generate trials that better fit the experiments than the deterministic analyses. For Specimens VSOA3, VSA3, VSB2, VSB3, VSC1, SW16, BN100, and BN100D, the majority of trials fit the experiment well despite producing different ultimate loads. From the examination of the load-displacement responses, the characteristic behaviours and shapes were very similar to the experiments but perhaps scaled up or down, as seen in Figure 4-15 and Figure 4-16. However, for Specimens VSOA1, VSA1, VSC2, SW11, SW12, SW15, SW21, SW22, SW23, B4, PLS4000, BN50, BH50, BH100D, and BM100, the general behaviour of most trials differed from the experiments. However, through random combinations of the material properties, trials with better fits were produced.

For Specimens VSA1, SW11, SW12, and SW21, the stochastic analyses did not produce any trials that modelled the experiment significantly better upon visual inspection. For Specimen B4, the governing properties were the yield and ultimate strength of the vertical reinforcement in the tension zone. As mentioned, the yield strength determined the yield point in the response and the ultimate strength determined the failure. In the deterministic analysis, the yield point was premature, but the failure load was well predicted. In the stochastic analysis as seen in Figure 4-18, many trials over-predicted the ultimate load while under-predicting the yield point. However, four trials matched very well due to the combinations of yield and ultimate strength as seen in Figure 4-18. The trials produced yield strengths between 533MPa and 508MPa from a distribution that averaged 475 ± 23MPa, and ultimate strengths between 670MPa and 713MPa from a distribution that averaged 726 ± 64MPa. The unexpected combinations of material properties resulted in the best predictions and would have been unlikely considered using deterministic analysis.

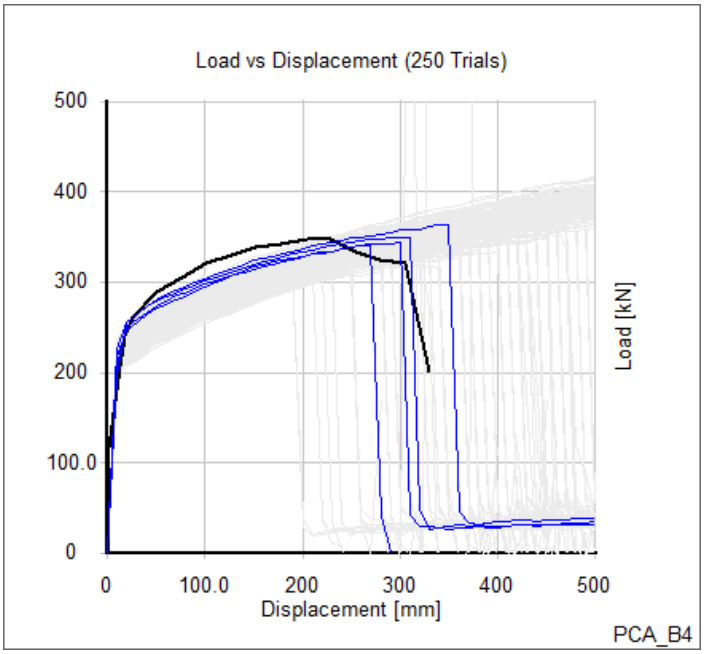


Figure 4-18 Best matched trials for Specimen B4

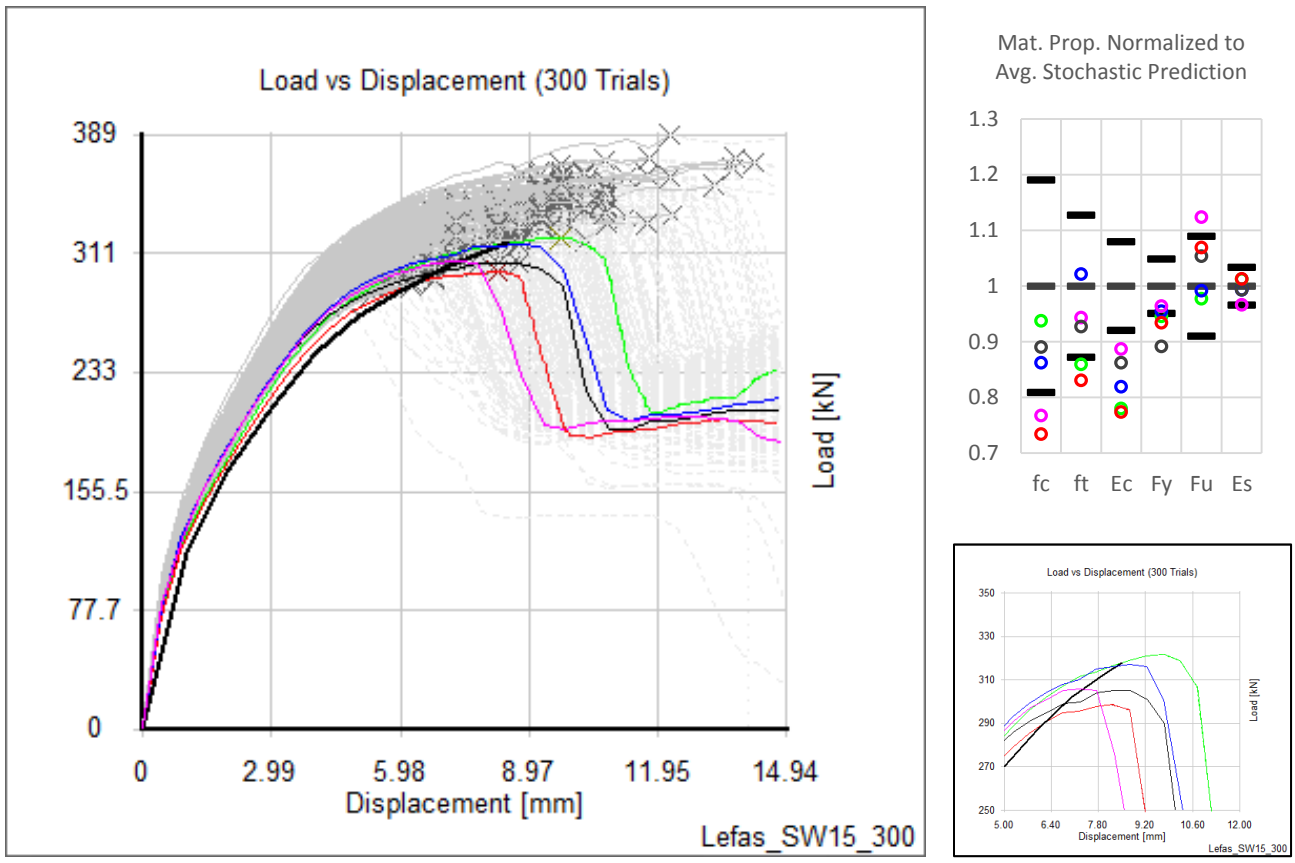


Figure 4-19 (a) Load-displacement (b) material properties of best matched trials for Specimen SW15 (c) Closer examination at failure

Many of the trials in the analysis of Specimen SW15 and SW23 produced responses with stiffer initial stiffness and softer behaviour towards failure compared to the experiment. Trials with more accurate initial stiffness tended to fail at lesser loads while trials with more accurate failure loads had poorly predicted stiffnesses. Five trials produced the best overall fit for SW15, highlighted in Figure 4-19. Figure 4-19 (b) shows the material properties of the five trials plotted against the stochastic average and standard deviation generated for each property. The five best trials had low values of elastic modulus of concrete replicating the softer stiffness of the experiment. Three of the trials placed in the top five trials with the lowest modulus out of 300 trials, and all five placed in the top 21. The best trials were the trials shown in green and blue which had low moduli values but relatively higher compressive strengths. A summary is provided in Table 4-9.

Table 4-9 Summary of best matched trials for Specimens SW15 and SW22

SW15	Test Day	Reference	Stochastic	Trial 41 (Green)	Trial 177 (Blue)
f'_c [MPa]	30.5	37.5	50.9 ± 9.7	47.7	43.9
E_c [MPa]	25 235	27 230	30 797 ± 2 451	24 017	25 215
SW23	**Test Day**	**Reference**	**Stochastic**	**Trial 106 (Magenta)**	**Trial 78 (Green)**
f'_c [MPa]	38.5	37.5	52.4 ± 9.8	47.7	44.4
E_c [MPa]	27 500	27 230	30 607 ± 2 451	27 203	29 054

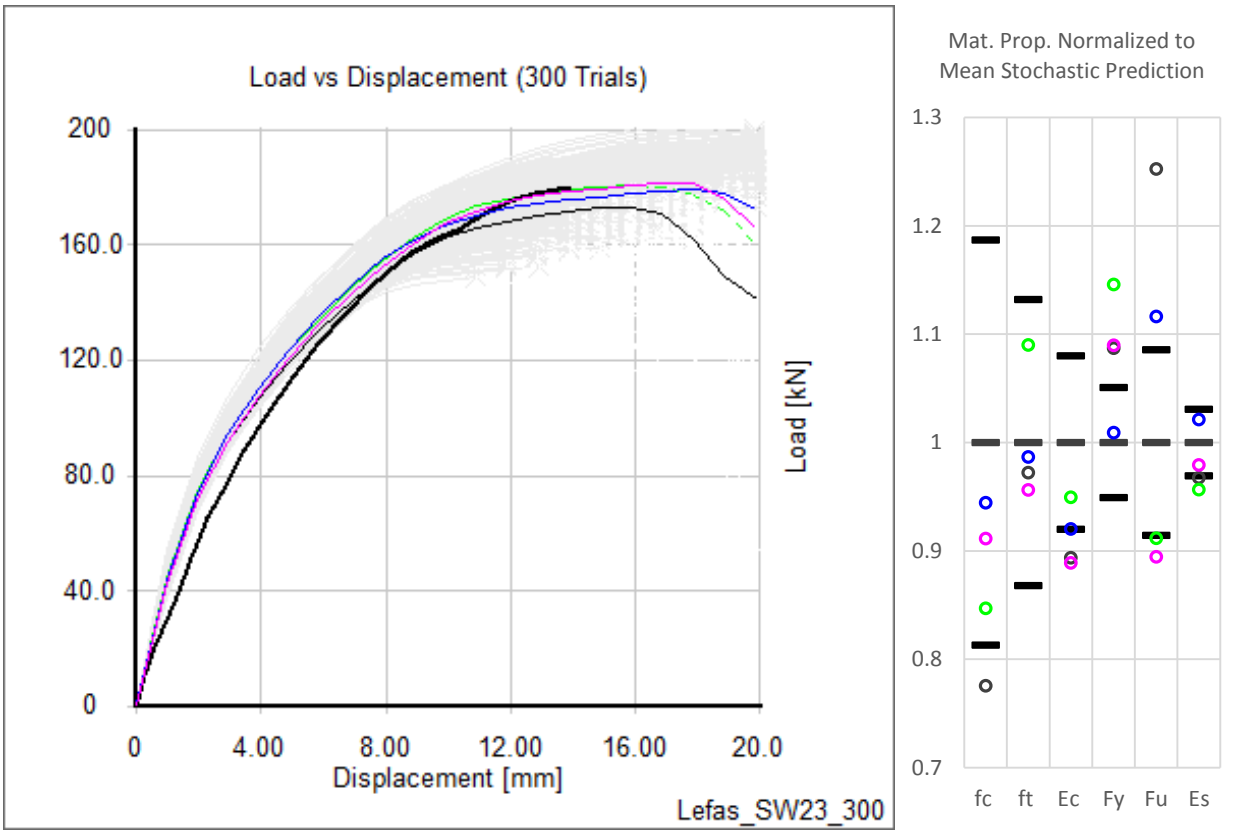


Figure 4-20 (a) Load-displacement (b) material properties of best matched trials for Specimen SW23

A similar effect was seen to an even larger extent in Specimen SW23 as seen in Figure 4-20. Four trials were identified as producing the best fit. Once again, these trials had low elastic modulus values while maintaining a certain level of compressive strength. The magenta trial provided the best fit with the 25th lowest modulus. Trials with lower modulus provided better estimates of the initial stiffness but not better estimates of failure load. Finally, out of the four trials highlighted, the green trial had the third largest compressive strength by a non-trivial difference and yet produced similar ultimate loads. This could be due to the sufficiently high tensile strength of the vertical reinforcement providing added flexural resistance and delaying compression failure.

The difficulties in matching the stiffness and failure of SW15 and SW23 as well as the other specimens by Lefas et al. (SW11, SW12, and SW21) could be due to a variety of reasons. The boundary conditions may have been improperly modelled. The concrete constitutive model used perhaps was not the most appropriate. Or perhaps there were anomalies and technical difficulties encountered in the laboratory tests. This is also true for Specimen B4. However, without the ability to easily but rigorously re-examine the experiment, the prediction cannot be properly refined. This scenario is even more applicable for structural evaluation projects where even less information would be available. However, with the stochastic analyses performed, more accurate predictions were produced. Even though the likelihood of concrete with 47.7MPa concrete strength and 24000MPa elastic modulus was low, the correct behaviour was replicated where a deterministic analysis could not do so.

Finally, for Specimens VSOA1, SW22, PLS4000, BN50, BH50, BH100D, and BM100, large numbers of trials produced adequate fits such that examination of individual trials was tedious. Instead, an algorithm was developed to automate the isolation of "well-matched" trials against experimental data. The algorithm would calculate a root mean square error term to judge the fit up to failure. Then, the percentage errors of the ultimate load and displacement predictions for each trial were applied to the overall error such that a bad estimate of ultimate load or displacement would increase the total error calculated while a good estimation would not. Thus, both the fit before the failure was accounted for as well as the failure prediction. This final value was only meant to rank the trials to isolate the better matched trials. The values themselves were not interpreted to have significant meaning.

Figure 4-21 shows the results for SW22. The trials highlighted in green were the trials judged to have produced the best fits. Only trials below a certain error threshold were highlighted. This threshold was determined by visual inspection. The histograms show the material property distribution of the overall analysis against the distribution of the highlighted trials in green. Similar to the other shear wall specimens, a certain combination of elastic modulus and compressive strengths were required to match the experiment. Except, in this case, the experiment was stiffer in response but failed relatively earlier. Thus, the best trials had relatively high elastic moduli within the predicted distribution and a certain range of compressive strengths below the average

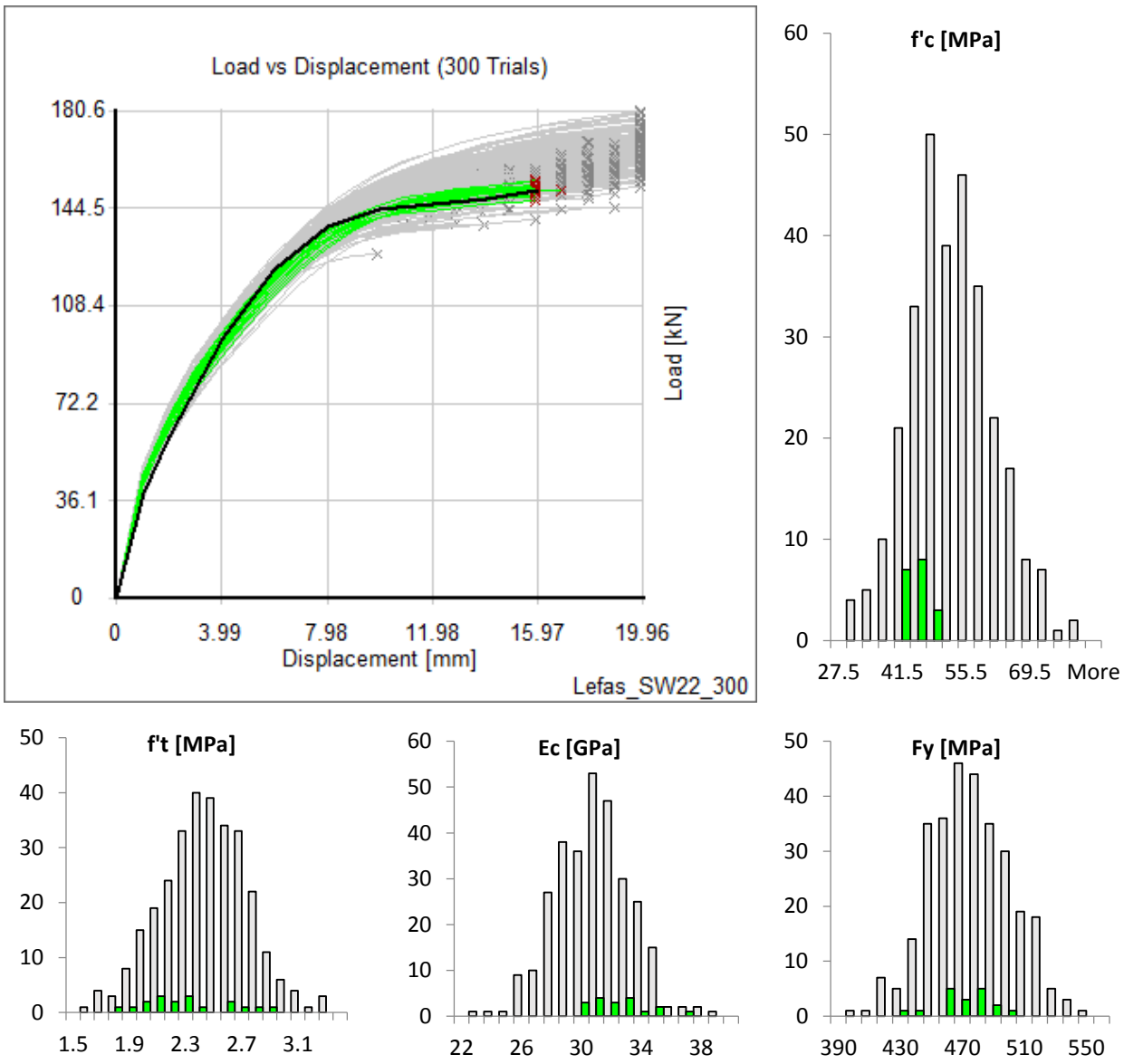


Figure 4-21 (a) Load-displacement (b)-(e) material properties of best matched trials for Specimen SW22

compressive strength. The concrete tensile strength and yield strength of the vertical reinforcement did not appear to have significant effects.

For Specimen VSOA1, no trial was able to capture both the ultimate load and the fit prior to failure. This was because both were dictated by the concrete tensile strength rather than a combination of parameters. Trials with higher, more accurate failure loads also had higher cracking loads. The delay of cracking resulted in parallel offsets between the experimental and analytical post-crack behaviour which was processed as poor fits. If the cracking load was well-predicted, the post-crack behaviour would fit well with the experiment, but the ultimate load would be under-predicted.

Since both phenomena were controlled by the tensile strength, no trial produced the optimal combination. The highlighted trials appeared to favour trials with lower elastic moduli. However, this was because softer stiffnesses only provided an apparent better fit when the cracking point was delayed.

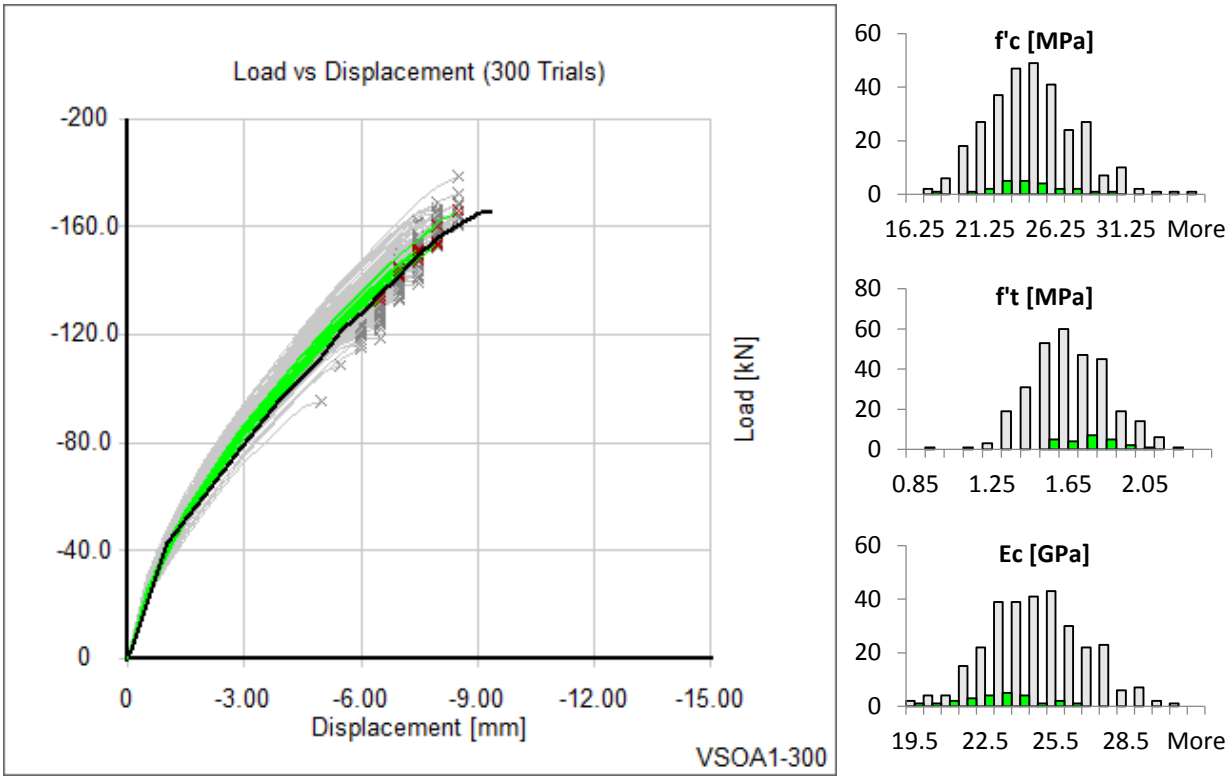


Figure 4-22 (a) Load-displacement (b)-(c) material properties of best matched trials for Specimen VSOA1

This effect was seen clearer in PLS4000. Because VSOA1 had a high flexural reinforcement ratio of 1.72%, the difference between the elastic and post-crack stiffness was smaller compared to that of PLS4000 which had a reinforcement ratio of 0.66%. As seen in Figure 4-22, there was a much clearer distinction between the elastic and post-crack behaviour. This led to more accurate calculations of fit as trials with larger concrete tensile strengths, and failure loads, would produce larger offsets to the experimental response during post-crack behaviour.

Nevertheless, once again no trials produced both the correct fit and the correct failure load because both relied on the tensile strength and the generation of stochastic simulations could not produce an optimal result. Here lies the weakness of stochastic simulation when the results are dependent

on one parameter and the combination of other degrees of freedom, such as other material properties, have little effect. The same band of results in Figure 4-23 (a) can be produced with 20MPa or 50MPa concrete as seen in Figure 4-23 (b). Thus, more uncertainty effects should be incorporated as inputs to the stochastic analysis.

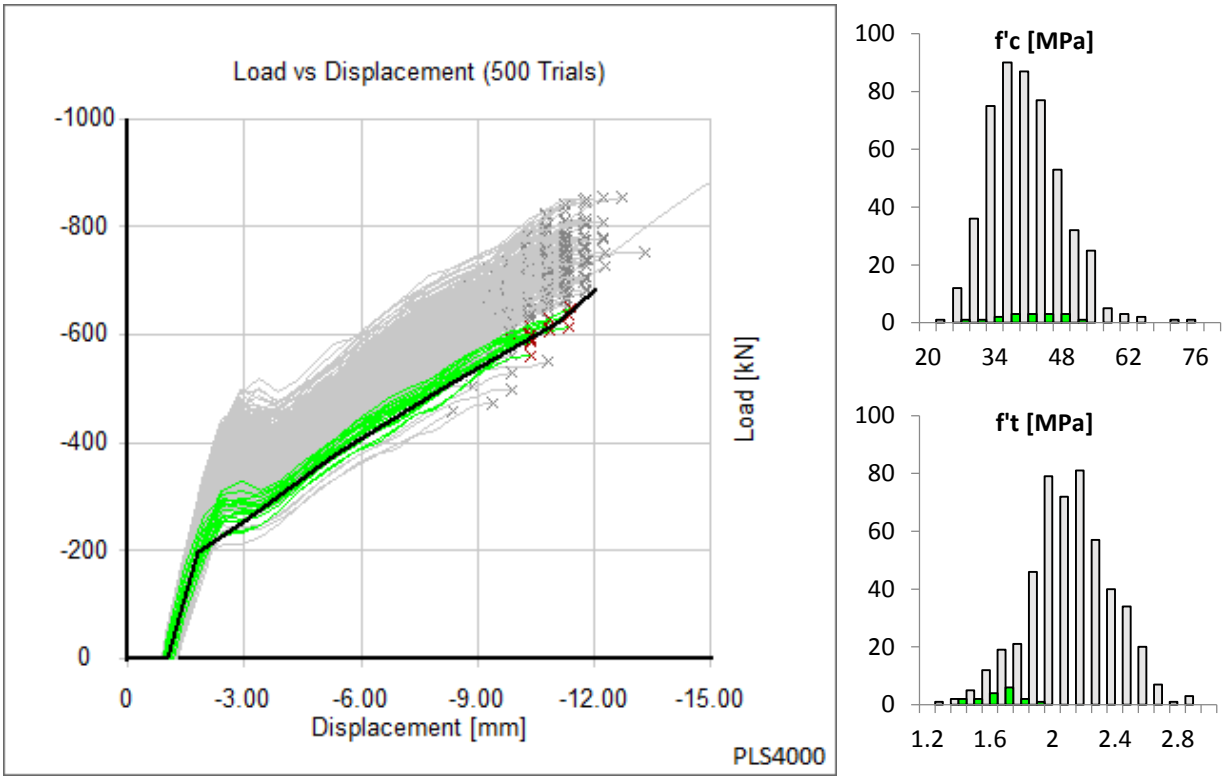


Figure 4-23 (a) Load-displacement (b)-(d) material properties of best matched trials for Specimen PLS4000

Finally, the stochastic results for Specimen BH50 and BH100D, as seen in Figure 4-24 and Figure 4-25, also showed the same behaviour. The failure loads appear to be better predicted for BH50, however, it may be due to the unexpected experimental response. For BH50, the post-crack behaviour was more nonlinear with decreasing stiffness compared to usual linear post-crack response of beams without shear reinforcement. This behaviour was not seen in the trials generated. For BH100D, the inclusion of skin reinforcement, which controls concrete cracks including diagonal shear cracks, perhaps made the prediction of failure more accurate. However, out of the best fit trials, the same trend of well-predicted failure load but poor cracking load, and vice versa, was seen.

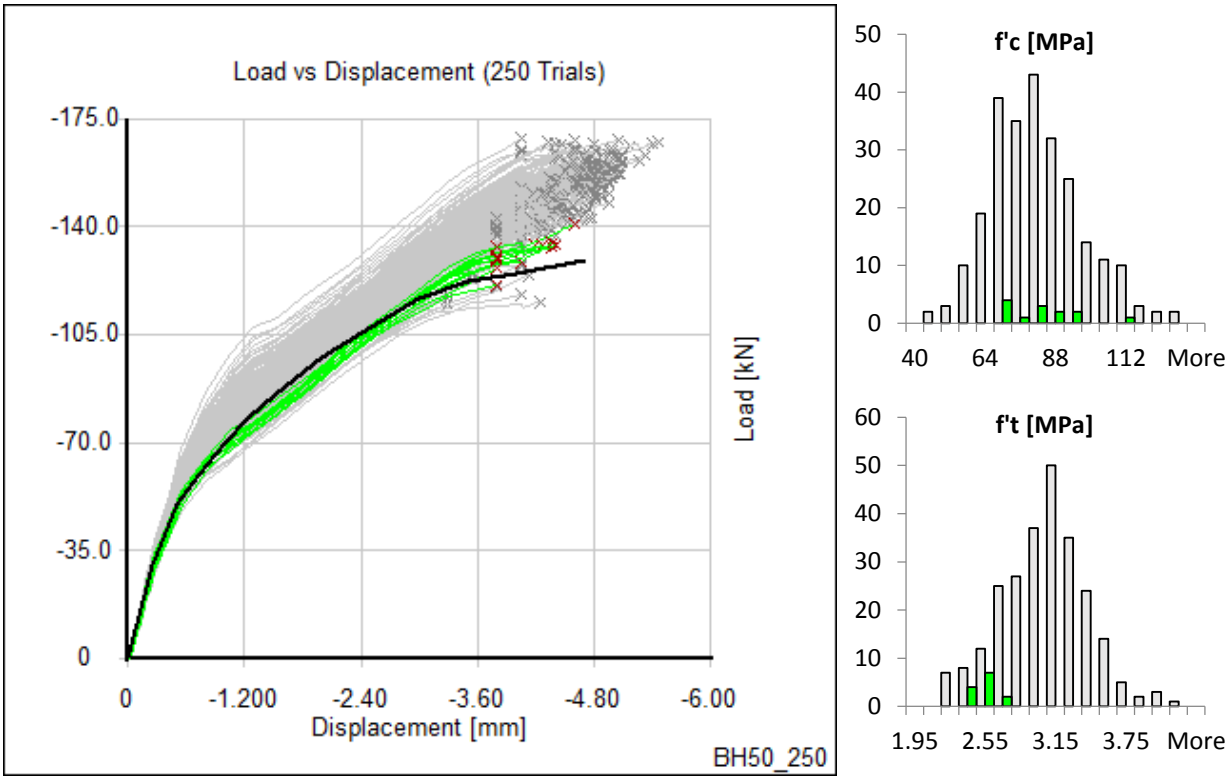


Figure 4-24 (a) Load-displacement (b)-(c) material properties of best matched trials for Specimen BH50

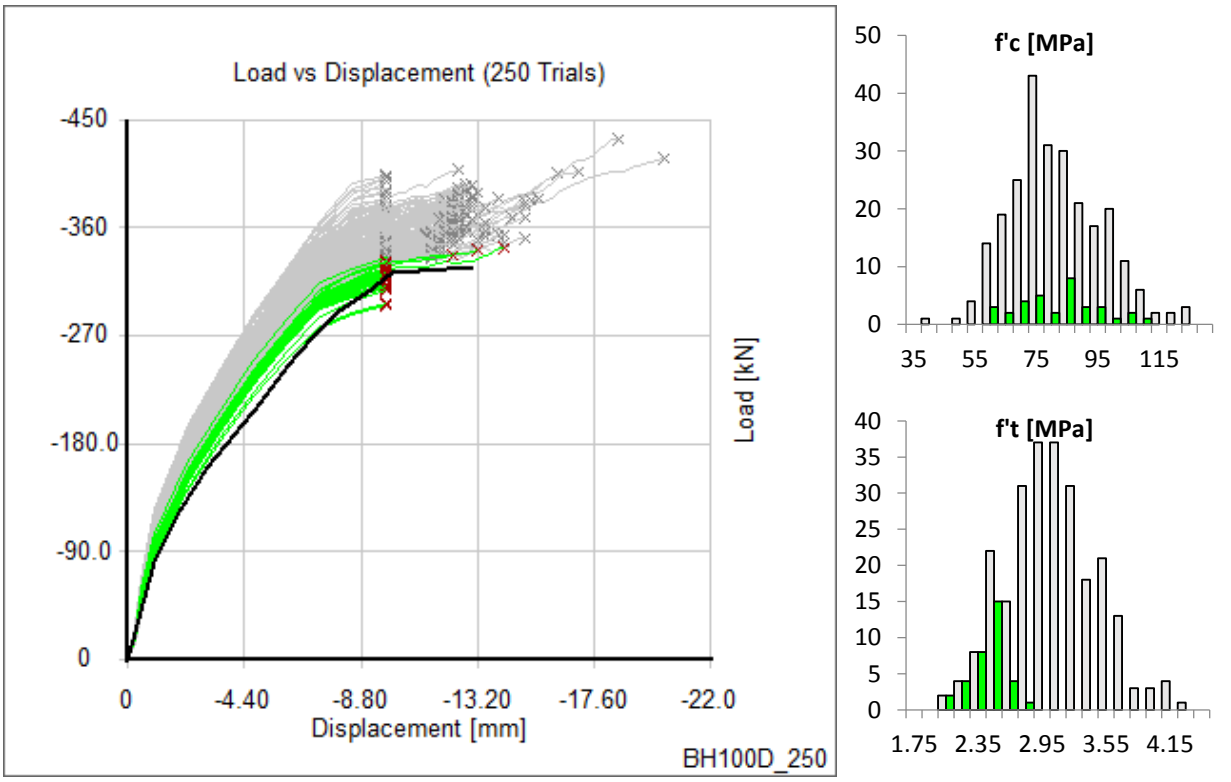


Figure 4-25 (a) Load-displacement (b)-(c) material properties of best matched trials for Specimen BH100D

4.5 Computational Statistics

Table 4-10 shows the computational cost in terms of time and data storage space for each stochastic analysis. The major parameters affecting computational cost were the number of trials, the number of elements, and the number of load stages.

Table 4-10 Summary of computational statistics for all specimens

Specimen	#Trials	#Elements	#Load Stages	Time [hr]	Data [GB]	Processor	Hrs per 1000E 100LS 200T	GBs per 1000E 100LS 200T	Hr/GB
OA1	300	749	31	5.0	4.0	A	14.36	11.48	1.25
OA3	300	1225	51	13.5	10.2	A	14.41	10.88	1.32
A1	300	790	31	5.5	5.0	A	14.97	13.61	1.10
A3	300	1292	51	13.0	11.2	A	13.15	11.33	1.16
B2	300	972	61	14.5	11.2	A	16.30	12.59	1.29
B3	300	1378	51	15.0	13.6	A	14.23	12.90	1.10
C1	300	874	41	8.0	7.1	A	14.88	13.21	1.13
C2	300	972	41	9.0	8.0	A	15.06	13.38	1.13
SW11	300	1130	31	6.2	9.0	B	11.80	17.13	0.69
SW12	300	1130	21	4.1	6.5	B	11.52	18.26	0.63
SW15	300	1130	31	6.2	9.0	B	11.80	17.13	0.69
SW16	300	1130	21	4.1	6.4	B	11.52	17.98	0.64
SW21	300	1480	26	7.3	10.2	B	12.65	17.67	0.72
SW22	300	1480	21	5.9	8.5	B	12.66	18.23	0.69
SW23	300	1480	21	5.8	8.4	B	12.44	18.02	0.69
B4	250	1250	101	11.7	18.1	B	7.41	11.47	0.65
PLS4000	500	4186	51	81.0	44.1	A	15.18	8.26	1.84
BN50	250	954	21	7.4	3.4	C	29.55	13.58	2.18
BN100	250	1376	71	54.5	13.6	C	44.63	11.14	4.01
BN100D	250	1376	81	17.8	16.5	B	12.78	11.84	1.08
BH50	250	954	85	33.5	11.1	C	33.05	10.95	3.02
BH100D	250	1376	65	29.2	12.9	C	26.12	11.54	2.26
BM100	250	1376	81	65.3	20.1	C	46.87	14.43	3.25

A = Intel® Core™ i7-3770 CPU @ 3.40 GHz processor with 8.00 GB of RAM
B = Intel® Core™ i7-6700 CPU @ 3.40 GHz processor with 16.00 GB of RAM
C = Intel® Core™ i7-3770 CPU @ 3.40 GHz processor with 8.00 GB of RAM

A total of three computers were used labeled as processor A, B, and C. Processor B contained the best specifications and produced the fastest results. The specification for processor A and C were the same, yet processor C produced drastically slower and unreliable results. This can be seen in

the normalized computational time and space costs. For each analysis, the time and data space were normalized to an analysis with 1000 elements, 100 load stages, and 200 trials to represent what an average stochastic analysis would require. The average storage space required for a typical analysis was estimated to be 13.8GB. It was not expected that the processor would affect this value. In terms of analysis time, processor A averaged 14.7hrs, processor B averaged 11.6hrs, and processor C averaged 36hrs despite the average data generated on the analyses for processor A, B, and C being 12GB, 16.4GB, and 12.3GB respectively.

Chapter 5 Results of Field Measurements Incorporation

This section presents the results from the analyses with the incorporation of field measurements. The major parameters considered in each analysis included the sensitivity of each field measurement, the number of field measurements used, the types of field measurements used, and the percentage of failure load the measurements were taken at. The result of each analysis was a different distribution of failure loads. The performance of this method was judged upon the ability to improve the predicted distributions. The four field measurements used were displacements, reaction loads, reinforcement strains, and crack widths.

5.1 Default Sensitivities

In the method implemented to incorporate field measurements, the measurement sensitivities used for each measurement can be quite important. The sensitivity value represents the perceived reliability of the measurement by the analyst and is treated as one sample standard deviation. As the results of the analyses will show, the sensitivity values used can heavily affect the final results. Initially, a calibration exercise was attempted to compute sensitivity values for each measurement type that resulted in the most improved predictions. However, given what the values represented, it was deemed inappropriate to derive default values from the performance results of the proposed method. Instead, they should be derived from direct testing of the sensors used in the collection of field measurements. Unfortunately, this was out of the scope of this thesis. Instead, the values were set based on preliminary research.

For laboratory experiments, measures of displacements are most commonly obtained by Linear Variable Differential Transformers (LVDT). However, despite the name, many sensors display nonlinear characteristics and conventional design of LVDTs high linearity can be difficult to achieve due to many unpredictable and uncertain factors including environmental and aging effects (Mishra, 2005). According to Mishra, who developed a system to compensate for LVDT error, conventional LVDT response could produce errors up to approximately 15% as seen in Figure 5-1, due to the assumed linear relationship between the output voltage and the actual displacement. Therefore, a default sensitivity of 15% was used for displacement measurements. A percentage value was used instead of a fixed displacement value because different experiments operate on

different scales. For different measurements within the same experiment, a percentage value was still justified because as Figure 5-1 suggests, the error increases as the LVDT measurement increases.

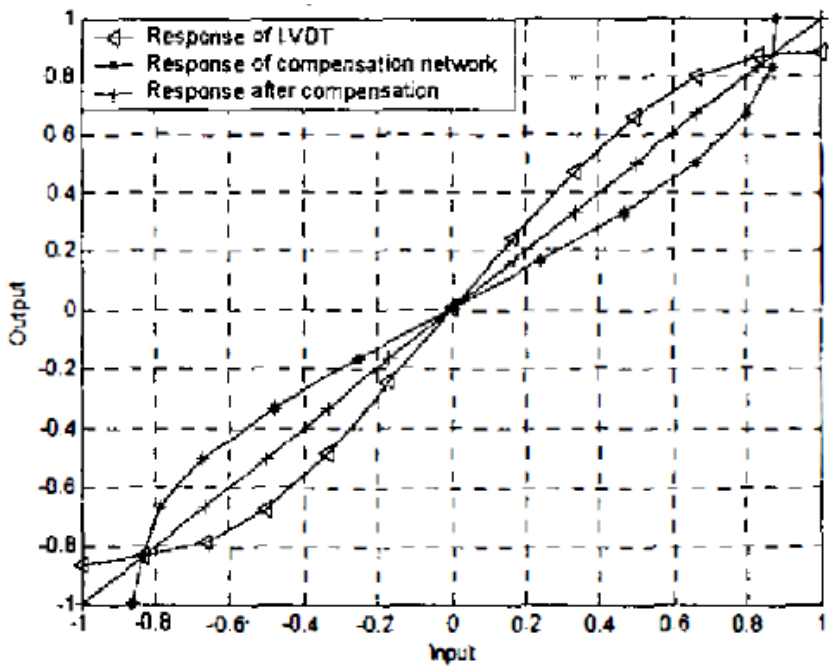


Figure 5-1 Nonlinear behaviour of typical LVDT provided by Mishra (2005)

The default sensitivity for strain readings was set to 10%. According to Micro-Measurements, a manufacturer of precision strain and stress measurement products, the error of moderate precision strain gages is between 2-5% (Micro-Measurements). Added uncertainty was included due to the placement of strain gages, and damage and dislocation during concrete casting.

Load cells typically obtain readings by measuring deformations of a load cell component due to the applied force using high precision strain gauges. Many of the same factors also affect load cell accuracy such as non-linearity, hysteresis, zeroing, and temperature effects. A value of 5% was assigned as the default load cell sensitivity.

Typical concrete crack markers provide denominations of 0.05mm. It is then up to the measurer to determine the correct crack width. Therefore, the initial value of crack width sensitivity was set to 0.15mm to allow for human error. Since most crack widths values were within a reasonable range of up to 0.1mm, and the values were more discrete rather than continuous, a fixed value was used instead of percentage like the previous parameters.

5.2 Summary of Results

5.2.1 Specimen VSOA1

Three sets of measurements were used, each including a displacement, reaction load, and crack width measurement. The three sets were taken at loads corresponding to 36%, 48%, and 61% of the failure load. Figure 5-2 shows each set of analysis in terms of the load-displacement response. The trials highlighted in green are the accepted trials with higher confidence, the grey trials have lower confidence, and the experiment is in black. A histogram of the failure loads is also shown.

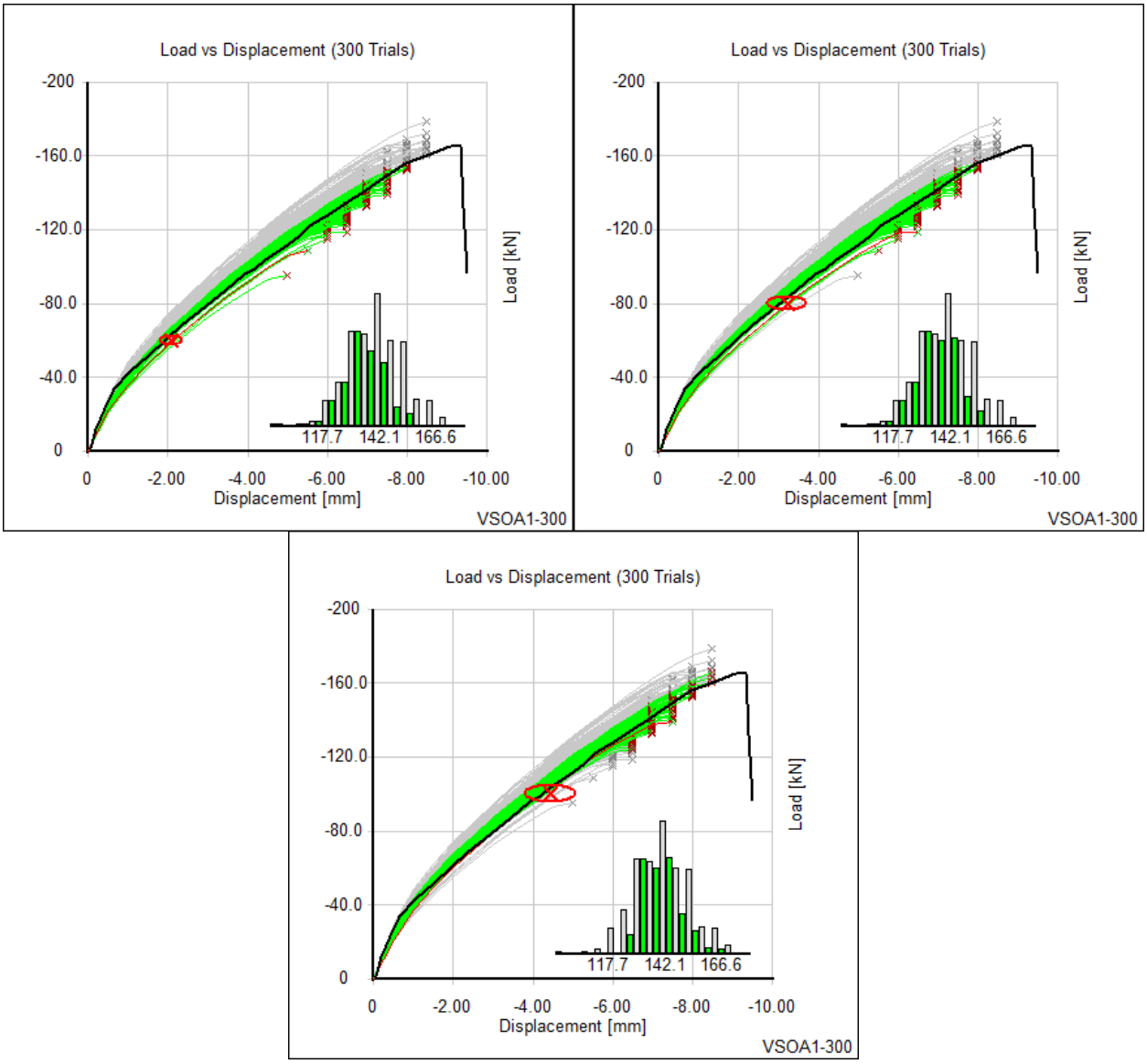


Figure 5-2 Load-displacement response of Specimen VSOA1 considering reaction loads, mid-span displacements, and crack widths at (a) 36% (b) 48% (c) 61% ultimate load

The average failure loads, from the three sets of analyses were 134.4kN, 135.8kN, 139.7kN respectively while the average failure load of all trials without considering field measurements was 142.1kN; the experimental failure load was 165.1kN. It would appear that the use of field measurements produced more incorrect results. However, this was again attributed to the modelling capabilities of VecTor2 of shear-critical beams without shear reinforcement. With the information Sherlock was provided, it was able to highlight the trials that matched the experiment the best. However, in terms of comparing failure loads, the highlighted trials failed earlier. The use of measurements closer to failure appeared to help the prediction as expected.

Figure 5-3 (a) shows the relationship between maximum crack widths and specimen displacement in terms of the experiment and all trials. In general, the analytical crack width values can skyrocket in comparison, especially towards failure. VecTor2 predicts crack widths by multiplying the principal strain by an estimated crack spacing parameter. However, in a finite element implementation, certain elements can produce high levels of strain and unrealistic values of crack widths. Nevertheless, the early crack widths values matched the experiment well. Unfortunately, the variation in crack widths was very low within the uncertainty involved in crack width measurement (0.15mm). Thus, the crack width measurement had no effect in the first two sets of measurements because all trials were deemed likely in terms of crack width. For the third set, the crack width measurement was turned off to investigate the effects of crack width. The analysis considering only mid-span displacement and reaction load is shown in Figure 5-3 (b). In this analysis, the average failure was 137.5kN. Although a minor difference, the positive effect of considering crack width was seen. The experimental load-displacement relationship was towards the bottom of the band of trial results despite ultimately failing at a relatively higher load. As a result, Sherlock would consider weaker trials to be more likely as seen in Figure 5-3 (b). However, the weaker trials would crack earlier and produce larger crack widths while the experiment exhibited relatively narrow cracks. Thus, the crack width measurement eliminated the weakest trials and pushed the average behaviour closer to the experiment. Figure 5-3 (c) shows the analysis with a tighter crack width sensitivity of 0.05mm. This eliminated more trials and raised the average failure load to 142.6kN. This also produced a narrower band of results around the experiment.

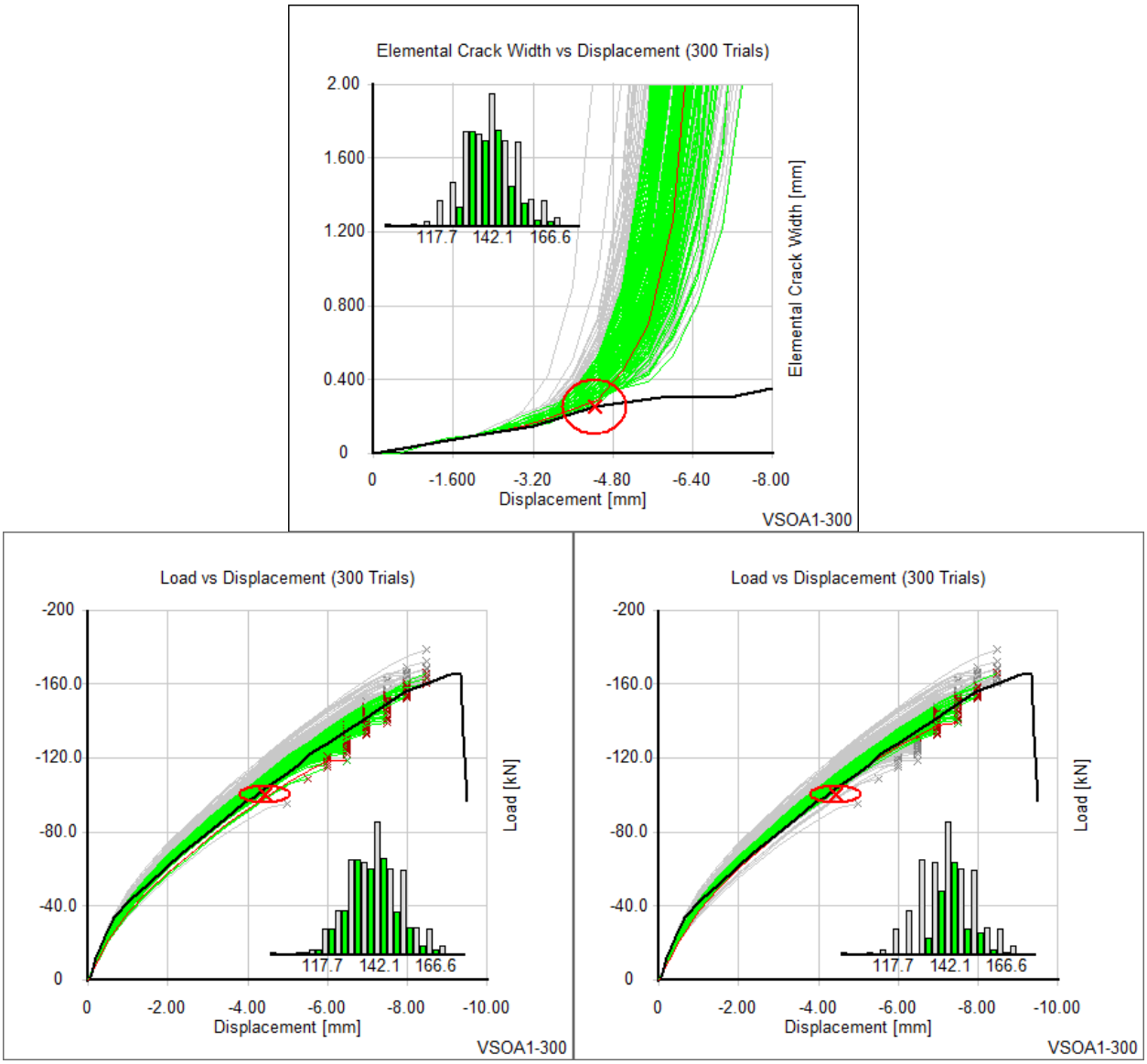


Figure 5-3 (a) Crack width displacement relationship for analysis considering all three measurements (b) Analysis considering only load and displacement (c) Analysis considering crack width with a tighter sensitivity on crack width

Finally, the reaction load measurement was excluded to investigate analyses without load measurements which can be hard to obtain in structural assessment. In the early stages, the crack widths measurements considered all trials as high confidence and no real results were obtained. For the last set of measurements of only displacement and crack width, the average failure load was 145.3kN with the load-displacement plots seen in Figure 5-4 (a). This essentially removed a few of the weaker trials only. The strongest trials were no longer eliminated by the relatively low load measurement of the experiment. However, in this case, only a few trials were eliminated, and the analysis matched the stochastic analysis without considering field measurements. If the

sensitivities were reduced to 0.2mm for the displacement and 0.05mm for the crack width, the average failure load would be 153kN with the load, displacement, and crack width relationships shown in Figure 5-4 (b) and (c). Although the failure load was more accurate due to the delayed experimental failure, the fit of the response to the experiment was less accurate.

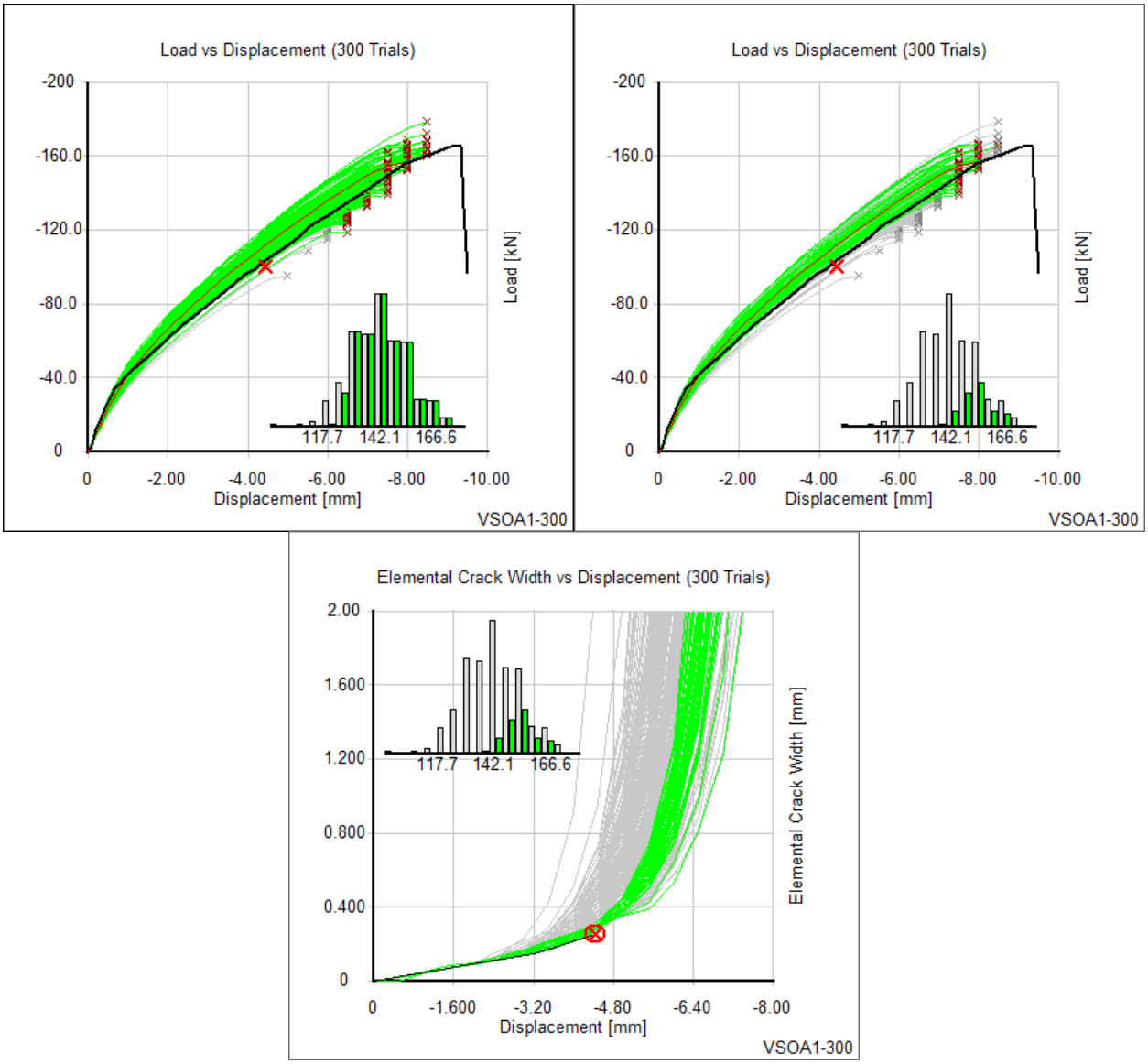


Figure 5-4 Analysis considering displacement and crack width with (a) default sensitivities (b), (c) tighter sensitivities

5.2.2 Specimen VSOA3

Three sets of measurements including displacements, loads, and crack widths were used at three different load stages. However, for Specimen VSOA3, the predicted crack widths values were quite over-estimated compared to the experimental values; so much so that the last two load stages

examined produced no confident trials. The analysis at the second load stage examined used a larger crack width sensitivity of 0.25mm instead of 0.15mm to capture more trials. Figure 5-5 (a) and (b) show the load-displacement plots for the first two load stages at 41% and 52% of the ultimate load respectively producing average failure loads 137.2kN and 158.0kN against the average failure of all trials of 186.3kN, and the experimental failure of 192.5kN.

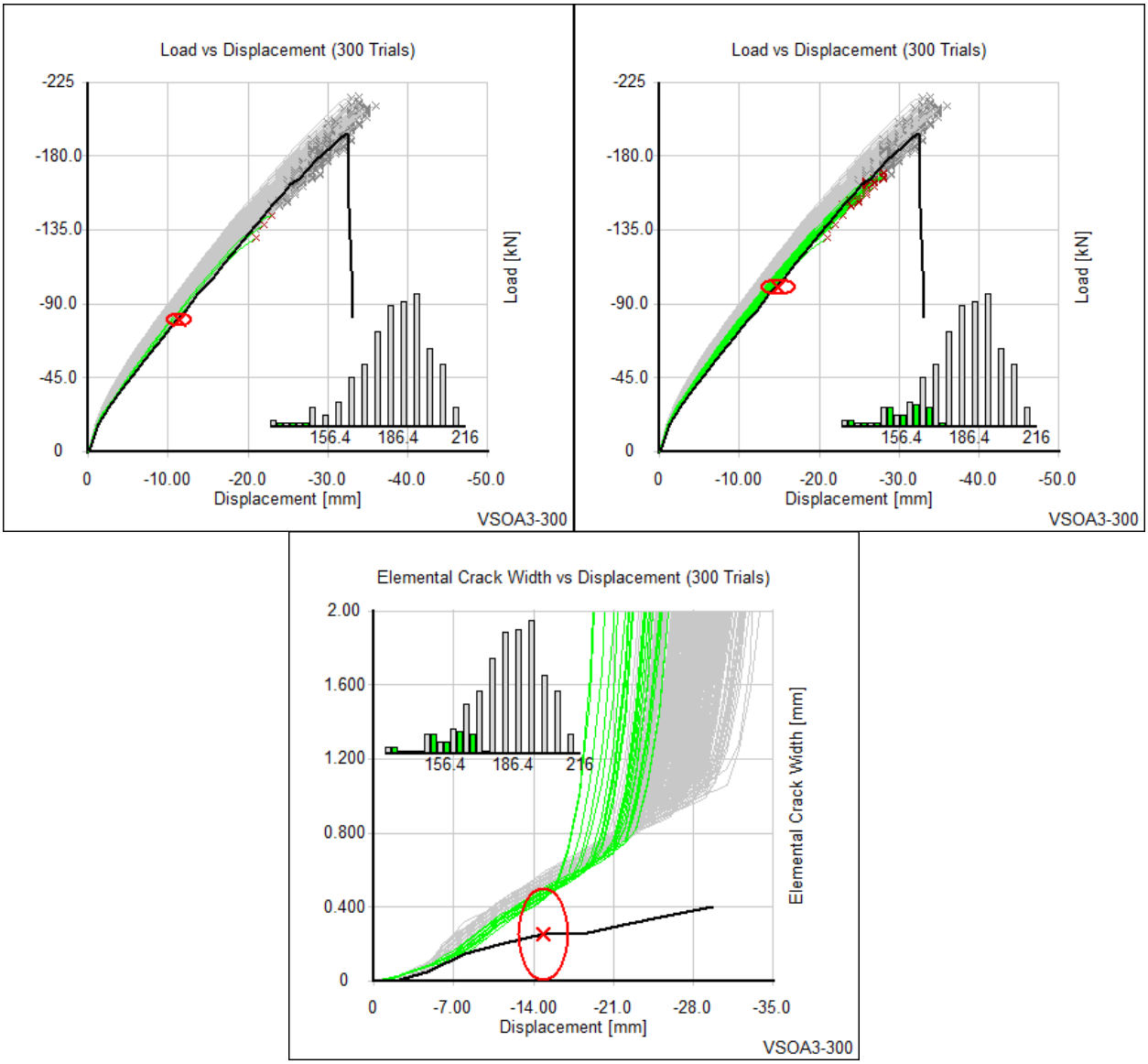


Figure 5-5 Analysis considering load, displacement, and crack width measurements at (a) 41% (b), (c) 52% of ultimate load

Figure 5-5 (c) shows the displacement crack width relationship for the second load stage and how, in subsequent load stages, the experimental crack widths and the predictions become more than twice the experimental. Figure 5-6 shows the analyses without any crack width producing average

failure loads of 179.0kN, 182.7kN, and 185.2kN. The last load stage examined was at 61% ultimate load.

Similarly to Specimen VSOA1, the early stage measurements suggested a lower failure load. However, the VecTor2 predicted trials were better compared to Specimen VSOA1 and it was expected that the inclusion of crack widths would once again eliminate some of the weaker trials that matched with the softer experimental load-displacement response. Instead, the opposite was seen. In comparing Figure 5-5 (b) and Figure 5-6 (b), the inclusion of crack widths unexpectedly removed many of the stronger trials.

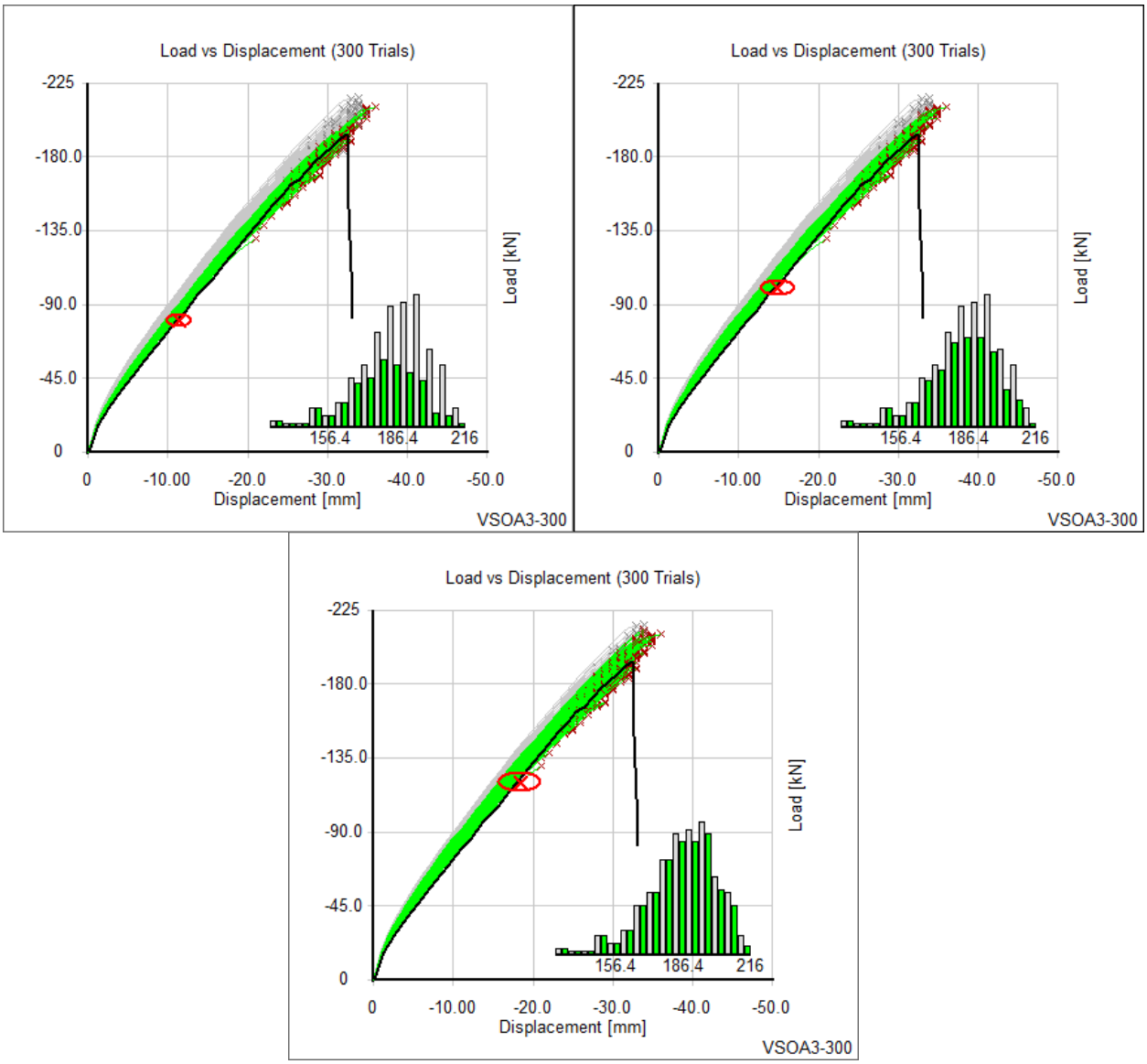


Figure 5-6 Analysis without considering crack width measurements at (a) 41% (b) 52% (c) 61% of ultimate load

Figure 5-5 (c) shows that trials with narrower cracks early on became the first ones to develop major cracking unlike Specimen VSOA1. The explanation for this behaviour was that Specimen VSOA3 was a more slender beam that developed more flexural cracks before developing major diagonal shear cracks. Weaker trials with lower tensile strengths, such as Trial 47, had deeper cracks within the cross section compared to stronger trials with higher tensile strengths, such as Trial 162. The crack width distribution for both trials at the same displacement of 12mm can be seen in Figure 5-7. First, it can be seen that the largest cracks are flexural. Second, the cracks in the cross section near mid-span are more evenly distributed for Trial 47 than Trial 162. Because the crack widths and therefore strains are more concentrated in Trial 162, the extreme element produced the largest crack width across both trials. Figure 5-7 also shows the region defined to capture cracking behaviour which did include the extreme element of Trial 147. As the load progresses, shear cracks dominate and develop earlier in weaker trials. In fact, in Figure 5-7, shear cracks have already begun to develop in Trial 46 but not as much in Trial 147. This explains the unexpected behaviour and a possible change required for how Sherlock extracts crack widths.

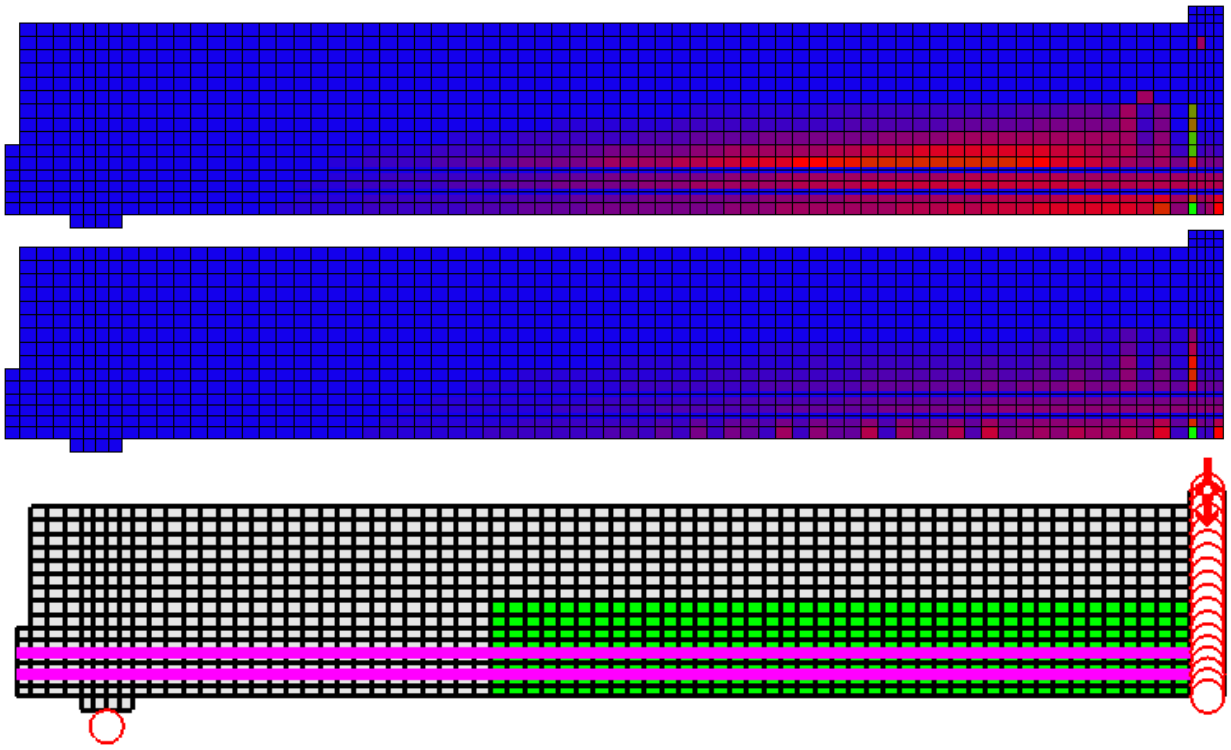


Figure 5-7 Crack width distribution of Trial 47 (top) Trial 162 (middle) and the elements selected to extract crack widths for analysis (bottom)

An analysis was rerun without considering the cracks in the bottom cover of the analytical model for the load stage at 61% ultimate load. If the crack width sensitivity was set to 0.05mm, the average failure was 195kN with the results seen in Figure 5-8.

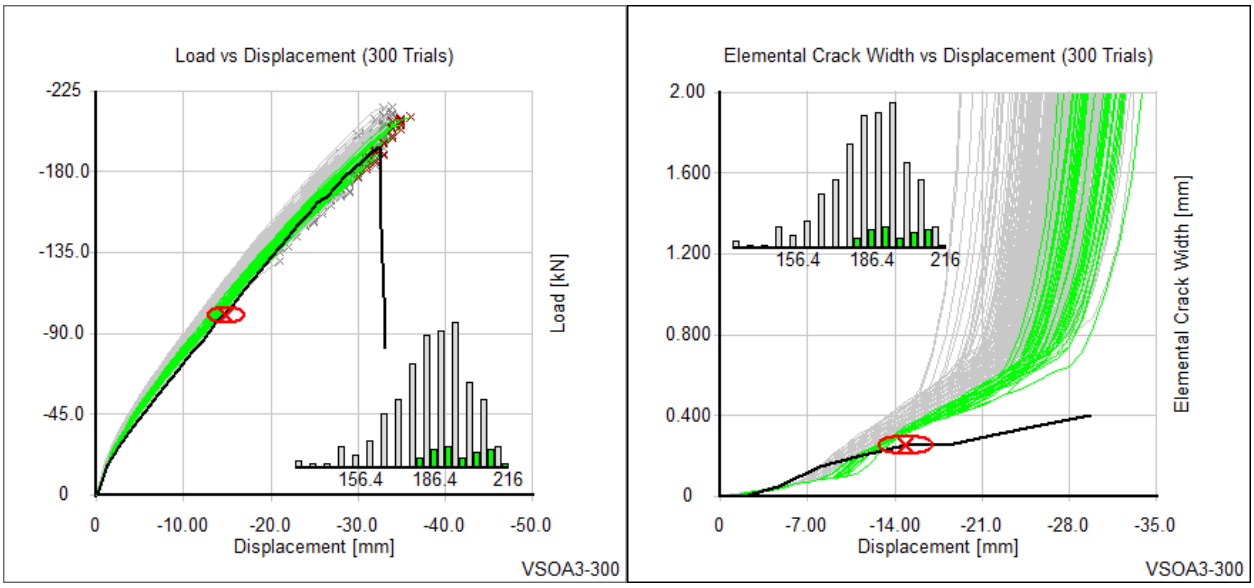


Figure 5-8 Analysis with refined crack width sensitivities

Overall with respect to the original analysis, the band of load-displacement response was relatively narrow which resulted in the distribution of matched trials being the same as the distribution of all trials. This is particularly true in the use of only two types of measurements in which the pinpointing effect is lower, and particularly in the latter load stages as the measurement sensitivities increase proportionally to the measurements while the variation in the predicted trials became constant.

5.2.3 Specimen VSA1

The crack width measurements used in the analysis for Specimen VSA1 did not produce significant effects. Figure 5-9 shows the load-displacement plots for three load stages at 44%, 53%, and 61% ultimate load, with only displacement and load measurements. The average failures produced were 225.6kN, 223.7kN, and 222.6kN respectively with the average failure for all trials being 236.6kN; 228.5kN was measured in the experiment. Figure 5-10 shows the displacement versus crack width relationships of analyses considering crack widths measurements.

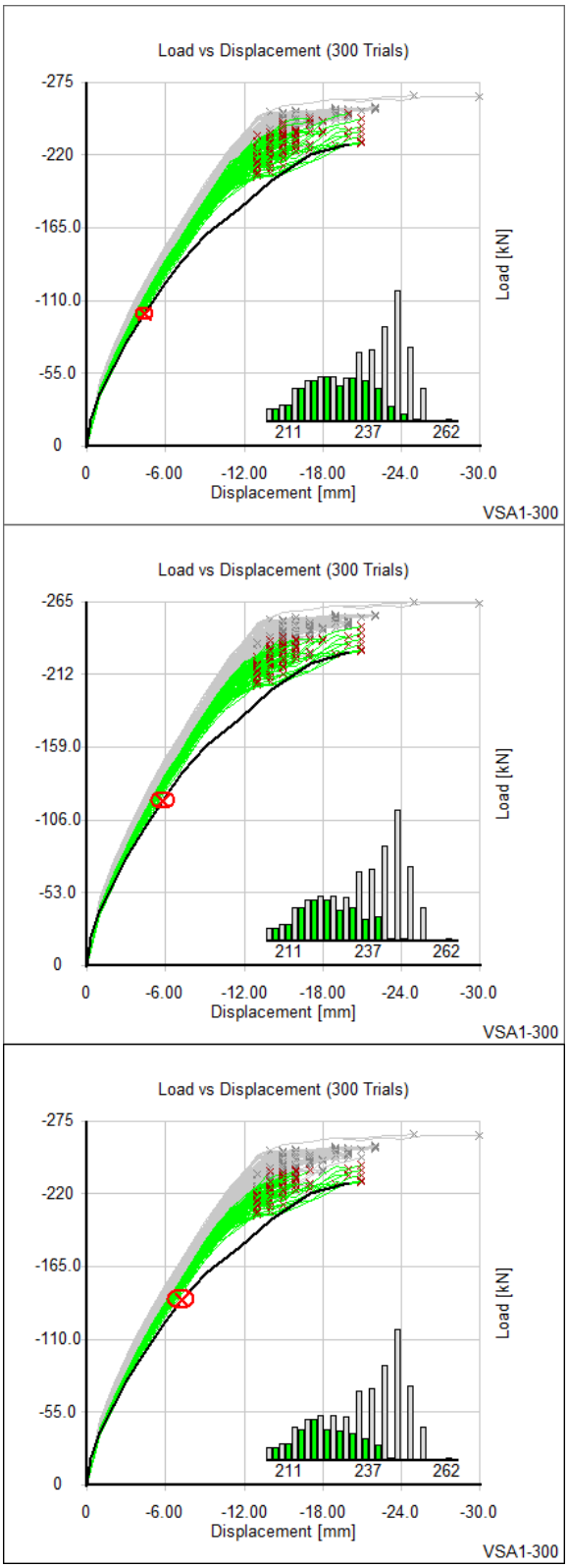


Figure 5-9 Analysis with load and displacement measurements at (a) 44% (b) 53% (c) 61% ultimate load

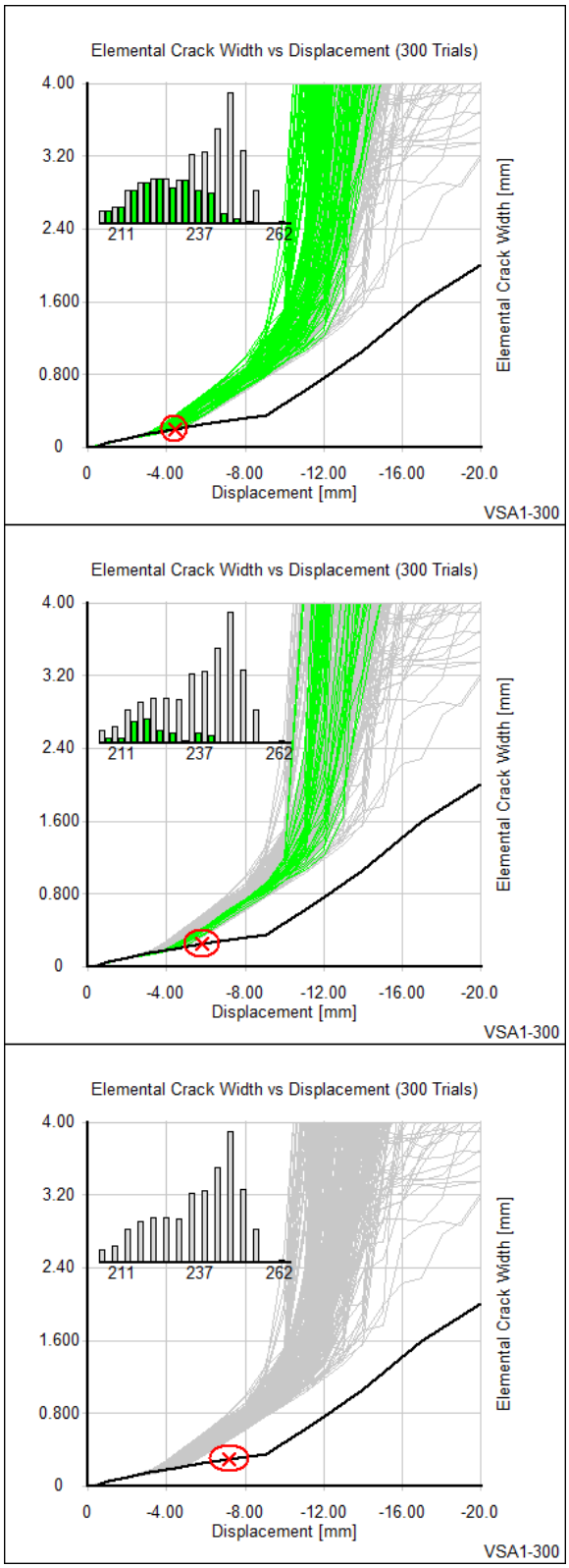


Figure 5-10 Analysis with load, displacement, and crack width measurements at (a) 44% (b) 53% (c) 61% ultimate

Figure 5-10 (a) shows that for the first load stage examined, there was very little variation within the predicted crack width. Figure 5-10 (c) shows that for the last load stage examined, the predicted crack width values were much higher than the experimental. Figure 5-10 (b) shows that for the second load stage examined, the inclusion of crack width eliminated trials uniformly and did not affect the overall failure load average or distribution. The reason for the difference compared to Specimens VSOA1 and VSOA3 is that Specimen VSA1 contained shear reinforcement and did not fail due to diagonal tension. Cracking did not contribute as much to the ultimate failure.

The benefit of field measurements was well demonstrated in this analysis. Sherlock was able to extract the trials with the best overall fit and accurate failure loads resulting in tighter and more accurate failure load predictions. The success of this analysis was largely due to the more accurate modelling of VecTor2 in terms of the overall response where similar early stage measurements, between the experiment and the predictions, resulted in similar failure loads.

5.2.4 Specimens VSA3, VSB2, VSB3, VSC1

The four specimens all exhibited a similar trend; the generated predictions produced very little variability and the overall experimental behaviour was matched very well. The combination of the two resulted in minimal to no improvement of the failure load distribution. In terms of the crack width measurement, Specimen VSA3 captured the experimental well within the examined load stages as seen in Figure 5-12 (a). However, the variability was less than the sensitivity of the crack measurement and the inclusion of crack width did not have a profound effect. For the Specimens VSB2, VSB3, and VSC1, the crack measurements were overestimated and also did not produce any effects.

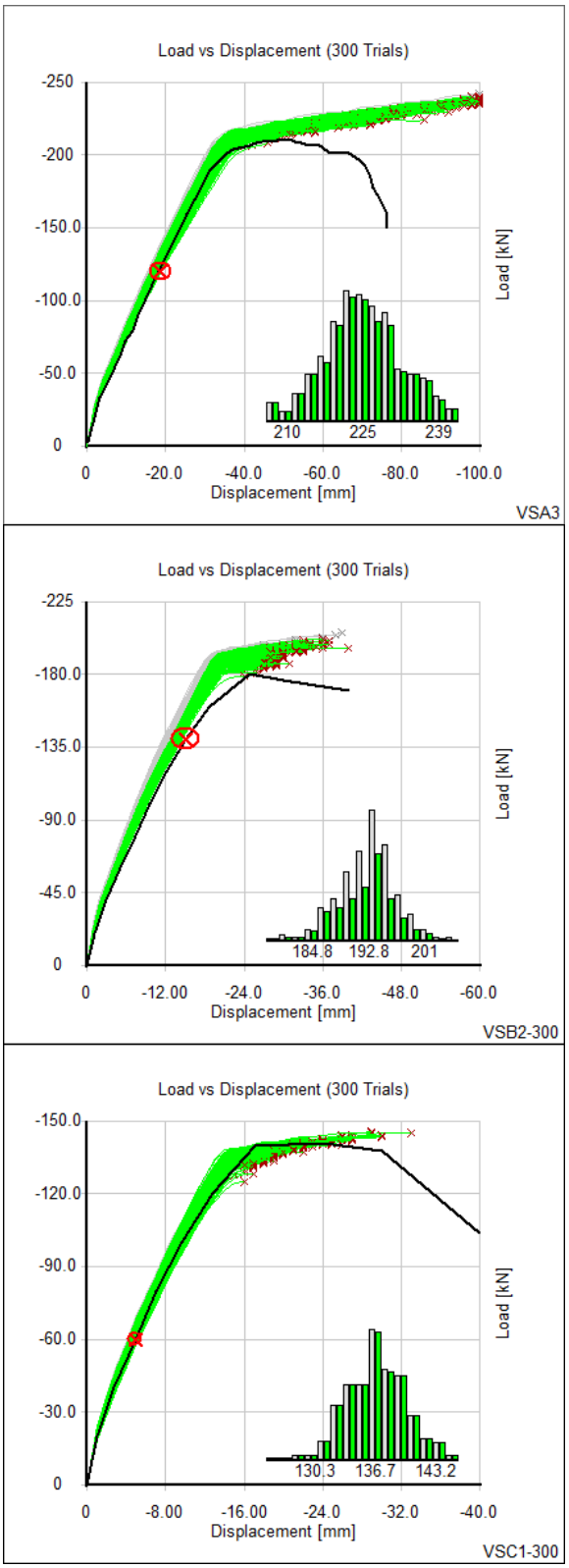


Figure 5-11 Analysis of Specimens (a) VSA3 at 57% (b) VSB2 at 78% (c) VSC1 at 43% ultimate load

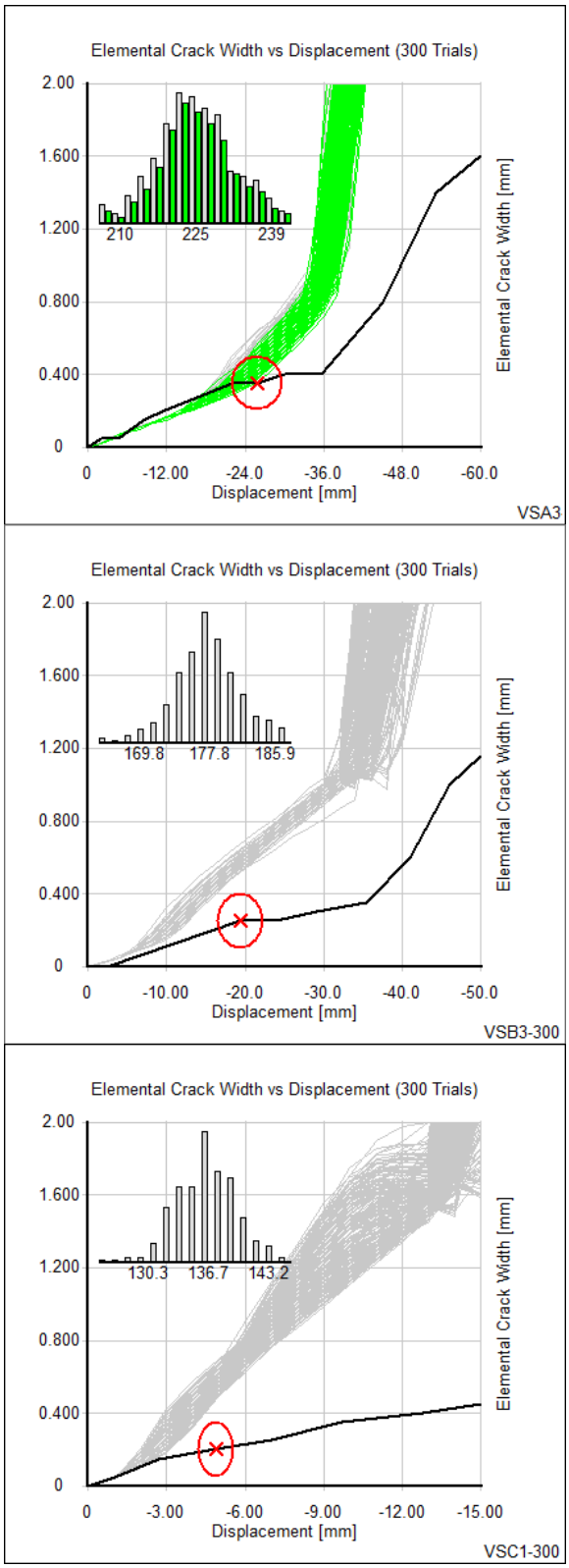


Figure 5-12 Crack width displacement relationship of Specimens (a) VSA3 (b) VSB3 (c) VSC1

5.2.5 Specimen VSC2

Once again, the crack measurements were overestimated for Specimen VSC2 and analysis with only mid-span displacement and reaction load measurements were examined. Figures 5-13 (a) – (d) show the load displacement results for four load stages at 43%, 57%, 71%, and 85% ultimate load which produced average failure loads of 173.5kN, 173.1kN, 168.2kN, and 153.3kN respectively with the average failure of all trials being 175kN; the experimental failure was 141.0kN. The crack width measurement at the first load stage examined (0.15mm) was still considered close to the predicted behaviours. However, comparison of failure load histograms in Figures 5-13 (a) and (e) shows that the inclusion did not affect the overall distribution, quite possibly because the failure mode was not diagonal tension and cracking was not overly influential. Figure 5-13 (f) shows that for subsequent load stages examined, the crack widths were overestimated quite possibly due to the same reasons as Specimen VSOA3.

It appears that the use of latter load stages improves the prediction. In the first two load stages, the trials isolated by Sherlock eventually "fanned" out at failure producing little change to the overall failure distribution. This is a result of the earlier stage behaviour being more dictated by the initial elastic modulus of the concrete while the failure in this specimen was dictated by the compressive strength. These results show that a relatively softer initial response does not guarantee a weaker failure. However, as the load progressed towards failure, only trials with a certain range of concrete compressive strength were captured and were thus more reflective of the failure.

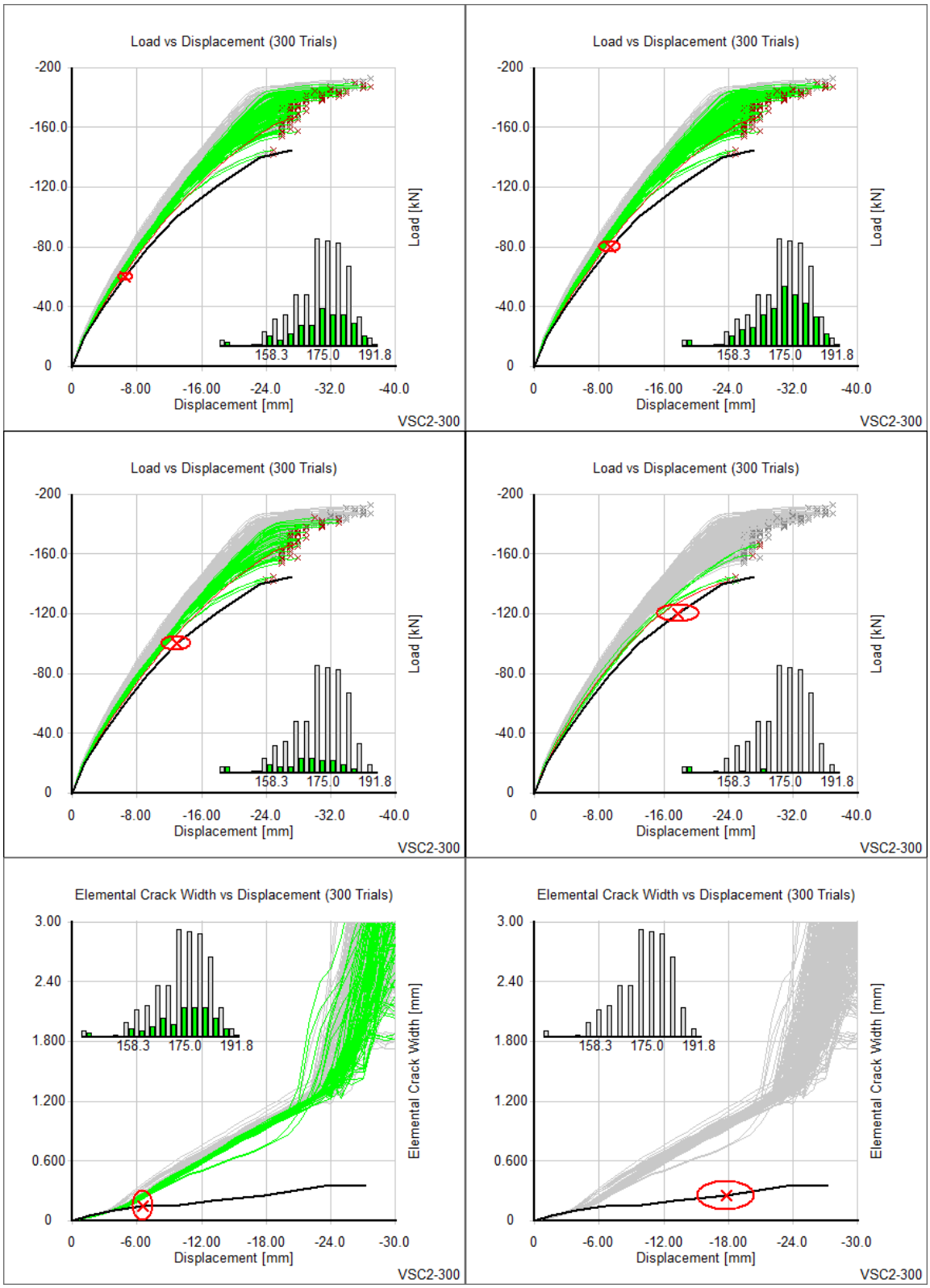


Figure 5-13 Load-displacement results of the analysis of Specimen VSC2 considering only load displacement measurements at (a) 43% (b)57% (c) 71% (d) 85% ultimate load. Crack displacement results of the analysis considering crack width measurement at (e) 43% (f) 85% ultimate load

5.2.6 Specimens SW11 and SW12

For these two specimens, the experimental behaviours and the predicted behaviours deviated largely. As such, the use of early stage field measurements did not identify any trials as likely. In these cases where all predicted trials poorly match the experiment, whether due to modelling deficiencies or experimental error, Sherlock cannot produce helpful conclusions. However, it could be concluded that a significant error had occurred in the modelling of the analysis or a major parameter was overlooked. Figure 5-14 shows sample results from Specimens SW11 and SW12.

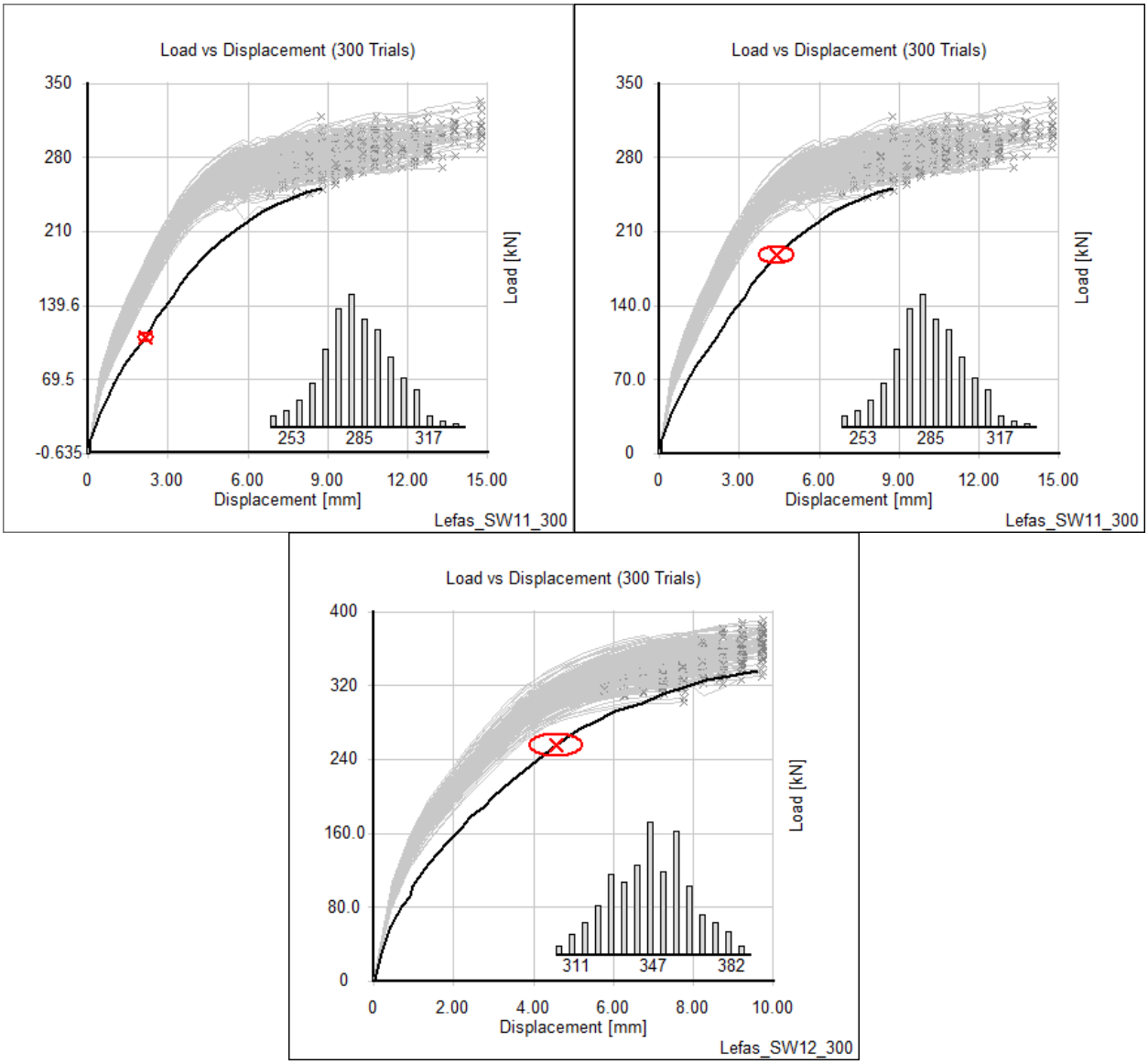


Figure 5-14 Sample results of Specimens SW11 and SW12

5.2.7 Specimen SW15

There was once again a deviation between the experimental and predicted behaviours for Specimen SW15. However, the difference was much smaller which allowed Sherlock to identify a few high confidence trials. Figure 5-15 shows the results at load stages corresponding to 43%, 63%, and 79% of the ultimate load using only reaction load and horizontal roof displacement measurements. The analysis produced averaged failure loads of 327.7kN, 320.7kN, and 298.9kN respectively while the average of all trials was 332.1kN; the experimental failure load was 320.0kN.

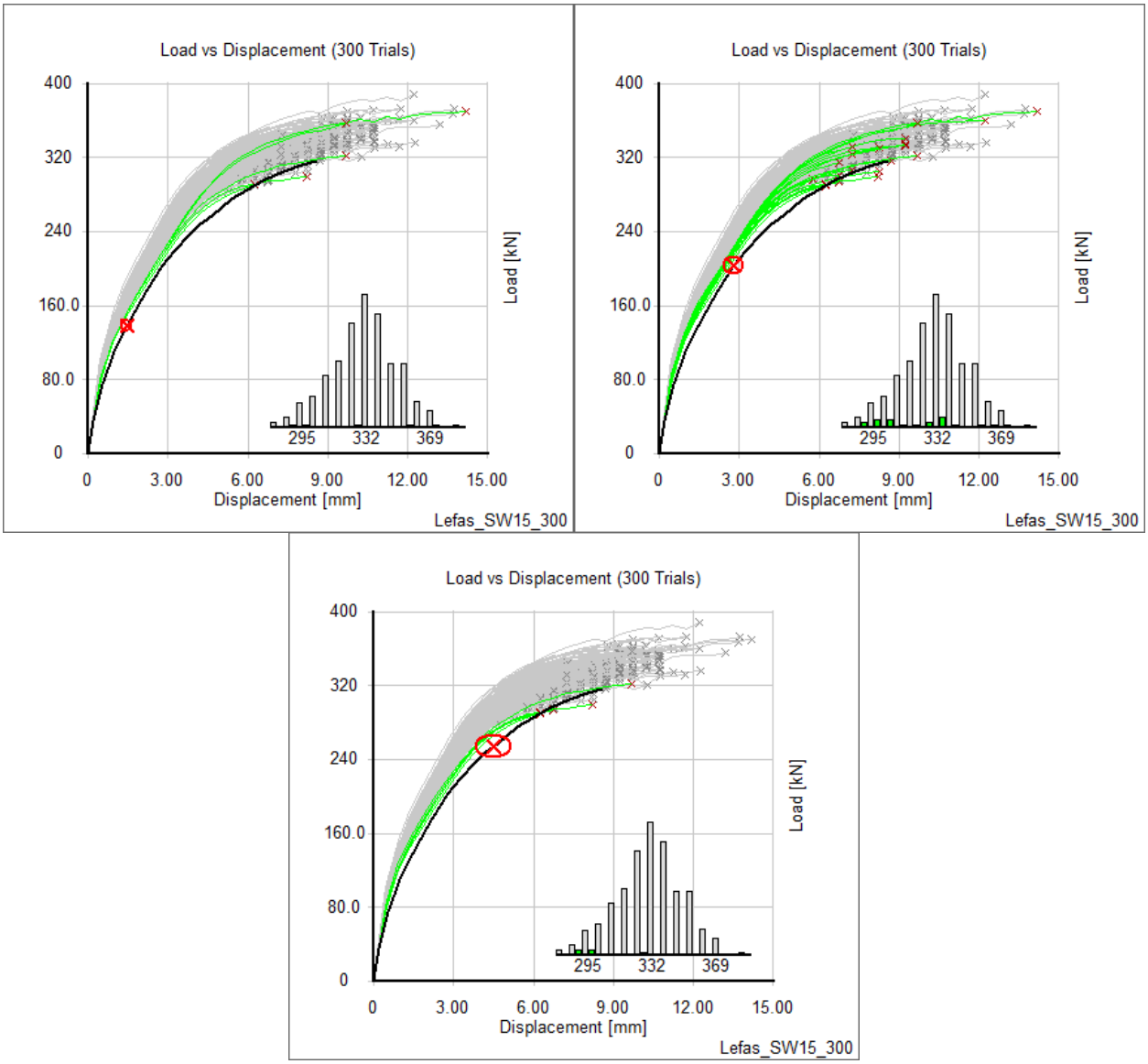


Figure 5-15 Analysis with load and displacement measurements at (a) 43% (b) 63% (c) 79% ultimate load

The number of trials considered high confidence increased from the first load stage examined versus the second because the measurement sensitivities increased faster than the variation of the load displacement responses. At the third load stage examined, there was a larger deviation between the experiment and the predicted trial behaviours which led to less accepted trials. As the load stage progressed, the band of load-displacement behaviour of the accepted trials became narrower, as seen in Figure 5-15 shown in green. However, because of the different shape of the experimental curve, the more focused band in the last load stage produced a lower average failure. In general, when the overall experiment differs largely from the predicted behaviour, Sherlock can be misled by experimental measurements. This is especially true for analyses with only reaction load and displacement measurements. With the inclusion of more measurements, Sherlock can detect failure mechanisms that occurred to better identify failure, such as the use of crack width measurements to pinpoint the progression of diagonal tension failure in VSOA1 and VSOA3.

A second displacement at two-thirds height of the shear wall was also used. Figures 5-16 (a) – (c) show the analyses rerun with consideration of the second displacement. Figure 5-16 (d) shows the locations of the primary displacement measurement at the top of the wall and the secondary displacement. Figure 5-16 (e) shows the load versus the secondary displacement relationship at the second load stage examined and Figure 5-16 (f) shows the primary displacement (horizontal axis) versus the secondary displacement (vertical axis) relationship. All trial results are shown in grey with the high confidence trials shown in green and the experimental results shown in black.

Figure 5-16 (f) shows that the ratio between the two displacements was well captured, as was the deflected shape of the shear wall. The new average failure loads were 298.7kN, 323.5kN, and 294.3kN for the first, second, and last load stages examined respectively. Examination of Figure 5-15 (b) and Figure 5-16 (b) did not show a systematic elimination of trials but rather a uniform elimination within the band of highlighted trials. This further suggests that in this case, the second displacement measurement did not effectively change the analysis. Although the average failure did change, this was probably due to the low number of high confidence trials, 20 trials without considering the second displacement measurement and only 8 trials with. With such few samples, a change of a few trials can lead to relatively large changes in the average failure load. Finally Figure 5-16 (f) shows that analysis at any load stage before failure without the consideration of the reaction load measurement would identify all trials as high confidence and thus inconclusive.

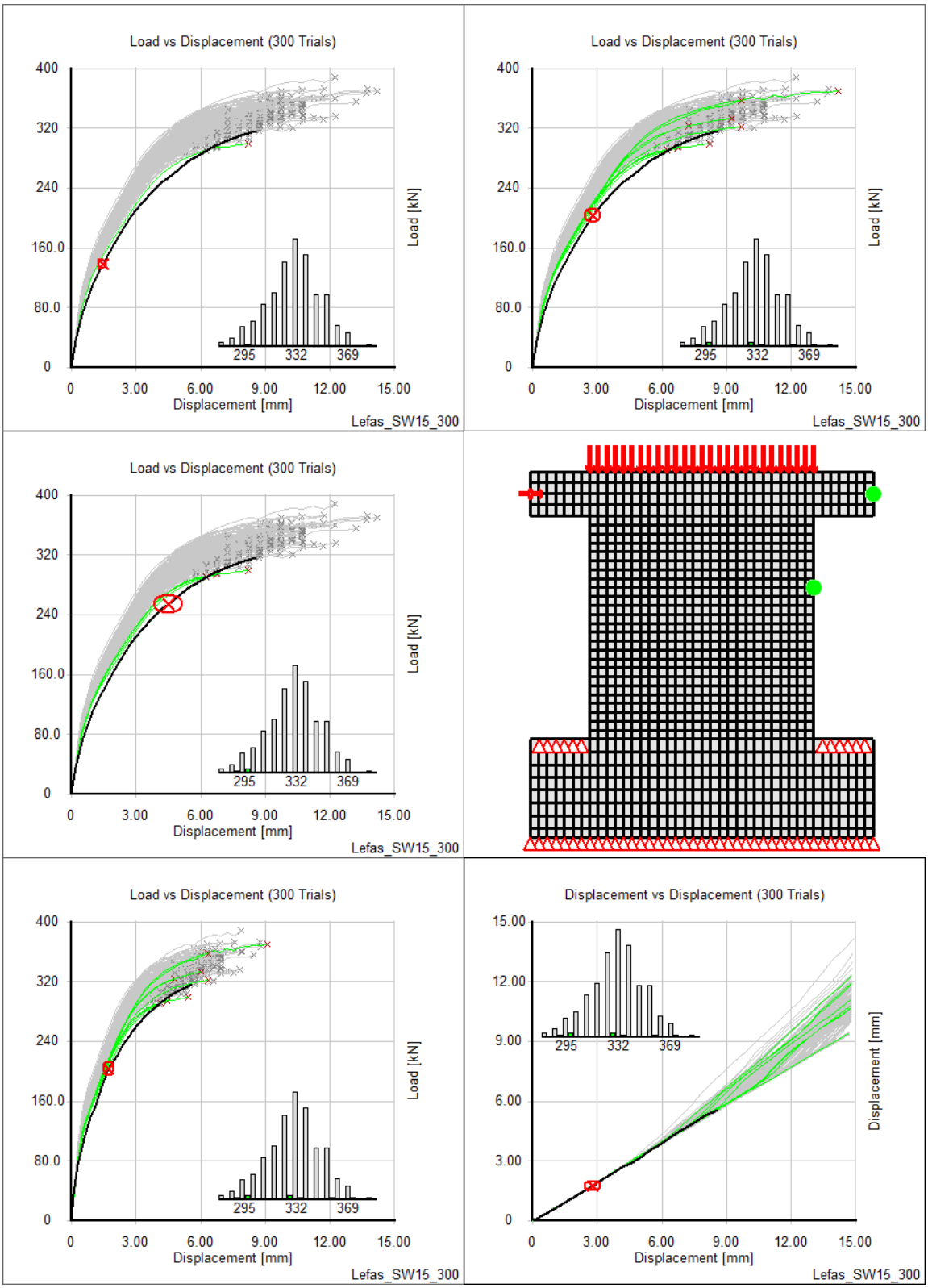


Figure 5-16 (a) - (c) Analysis considering a second displacement measurement (d) the location of the two displacement measurements (e) reaction load versus second displacement relationship (f) top displacement versus second displacement

5.2.8 Specimen SW16

For this specimen, the secondary displacement at the same height as Specimen SW15 was also included along with the top displacement and reaction load measurements. Figures 5-17 (a) – (c) show the analyses at 37%, 74%, and 85% of the ultimate load which produced average failure loads of 384.9kN, 385.8kN, and 380.7kN respectively against the average of all trials being 391.9kN; the experimental failure load was 355.1kN. To investigate the effect of the secondary displacement, Figures 5-17 (d) – (f) show the analyses without the secondary displacement which produced average failure loads of 387.0kN, 388.3kN, and 381.4kN respectively.

Analysis with consideration of field measurements produced more accurate average failure loads compared to the average for all trials. Analysis considering the second displacement measurement produced more accurate average failure loads than analysis without. And analysis with later load stages produced more accurate average failure loads than early load stages as expected. However in all cases, the improvement was minimal. The best prediction was 380.7kN, with the use of all field measurements at 85% of the failure load, and the worst prediction was 391.9kN, without consideration of field measurements, while the experiment failed at 355kN. Furthermore, Figure 5-17 shows that there were likely several trials that matched the experiment well and yet the analysis could not pinpoint the failure more accurately.

The poor performance of the analysis of Specimen SW16 is similar to the analyses of Specimens VSA3, VSB2, VSB3, and VSC1. The sensitivities of the measurements used were too large compared to the variation in the predicted trials such that the majority of trials were considered likely. This led to average failure loads that were similar to the average of all trials. Furthermore, the early stage behaviour was more dictated by the concrete elastic moduli while the failure was more dictated by the concrete compressive strength as Specimen SW16 also failed due to compression. Therefore, early measurements did not correlate to failures. Measurements at later load stages correlated better, however the sensitivities were much too large to produce more accurate results. If the sensitivities for all three measurements (reaction load, primary displacement, and secondary displacement) were reduced by half, the average failure load would have been 372kN, as seen in Figure 5-18.

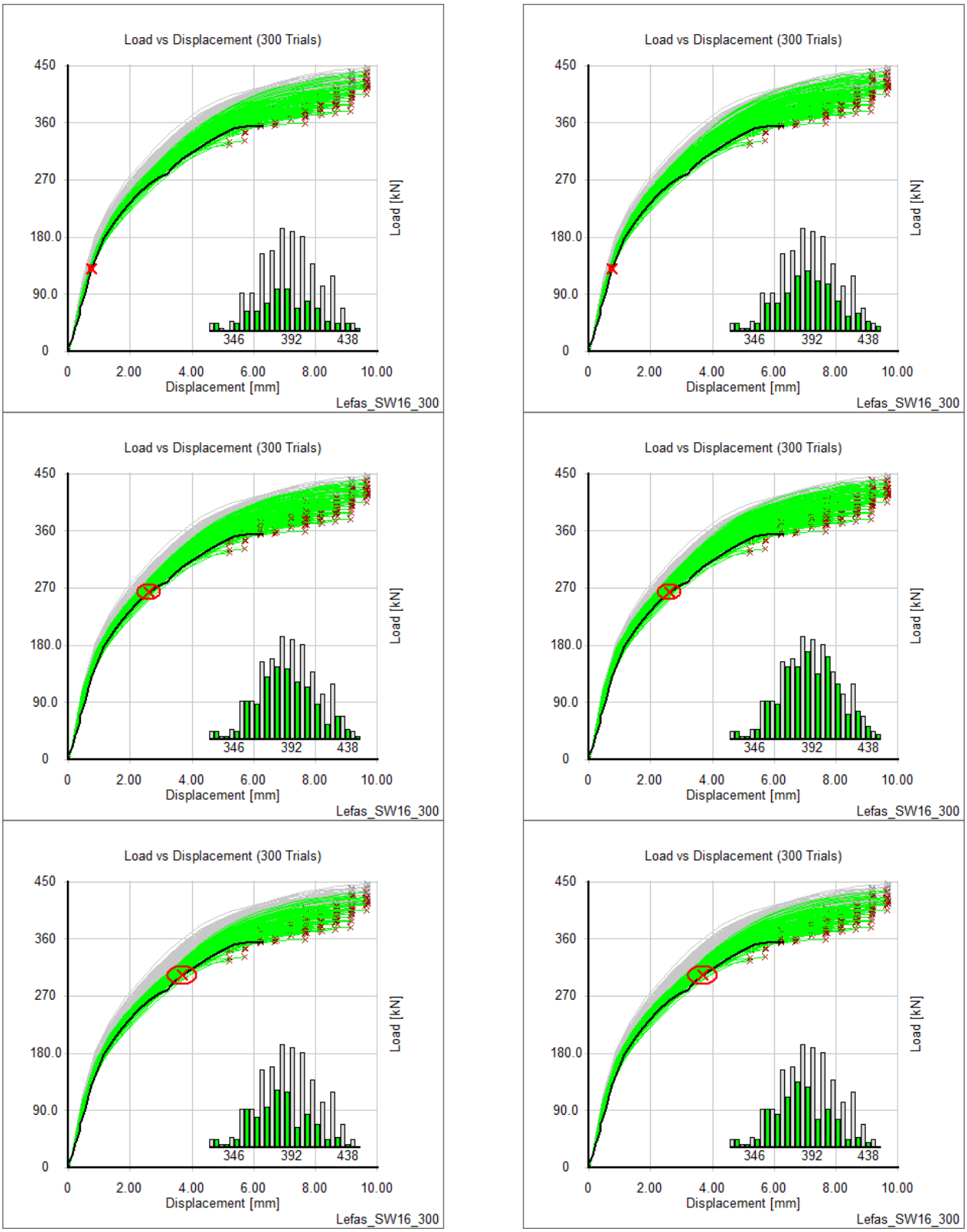


Figure 5-17 Analysis of Specimen SW16 considering the secondary displacement at (a) 37% (b) 74% (c) 85% ultimate load. (d) – (f) Analysis not considering the secondary displacement

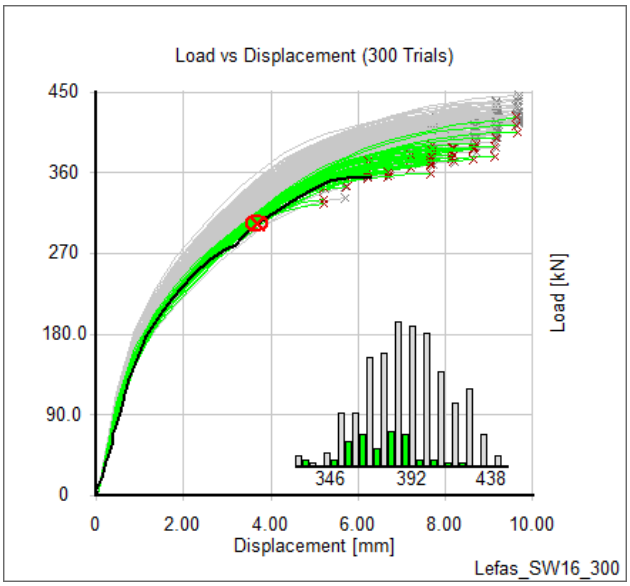


Figure 5-18 Analysis with reduced sensitivities on all three measurements

However, this should not necessarily be considered a negative result, similar to the results of Specimens VSA3, VSB2, VSB3, and VSC1. If the sensitivities were well chosen based on the sensors used and similar results were obtained, it simply means that the experimental behaviour was well captured by VecTor2 and confirmed by early stage field measurements.

5.2.9 Specimen SW21

Figures 5-19 (a) – (c) show the results of analyses at 53%, 65%, and 82% of the ultimate load using three measurements: reaction load and two displacements, one at the top and one at mid-height of the shear wall. The three analyses produced average failure loads of 139.1kN, 134.6kN, and 125.2kN respectively, while the average failure load for all trials was 138.6kN and the experimental failure was 126.9kN. Figures 5-19 (d) – (f) show the results of analyses without the second displacement measurement, yielding average failure loads of 139.0kN, 138.5kN, and 130.8kN.

Once again, the first load stage examined produced very little change in the overall distribution of failure loads. The later load stages, especially the last load stage examined, produced a highly pinpointed band of results and a close average failure to the experiment. Comparison of Figures 5-19 (a) – (c) against Figures 5-19 (d) – (f) shows that the inclusion of the secondary displacement eliminated many trials and resulted in more accurate results. Figure 5-20 offers a possible explanation.

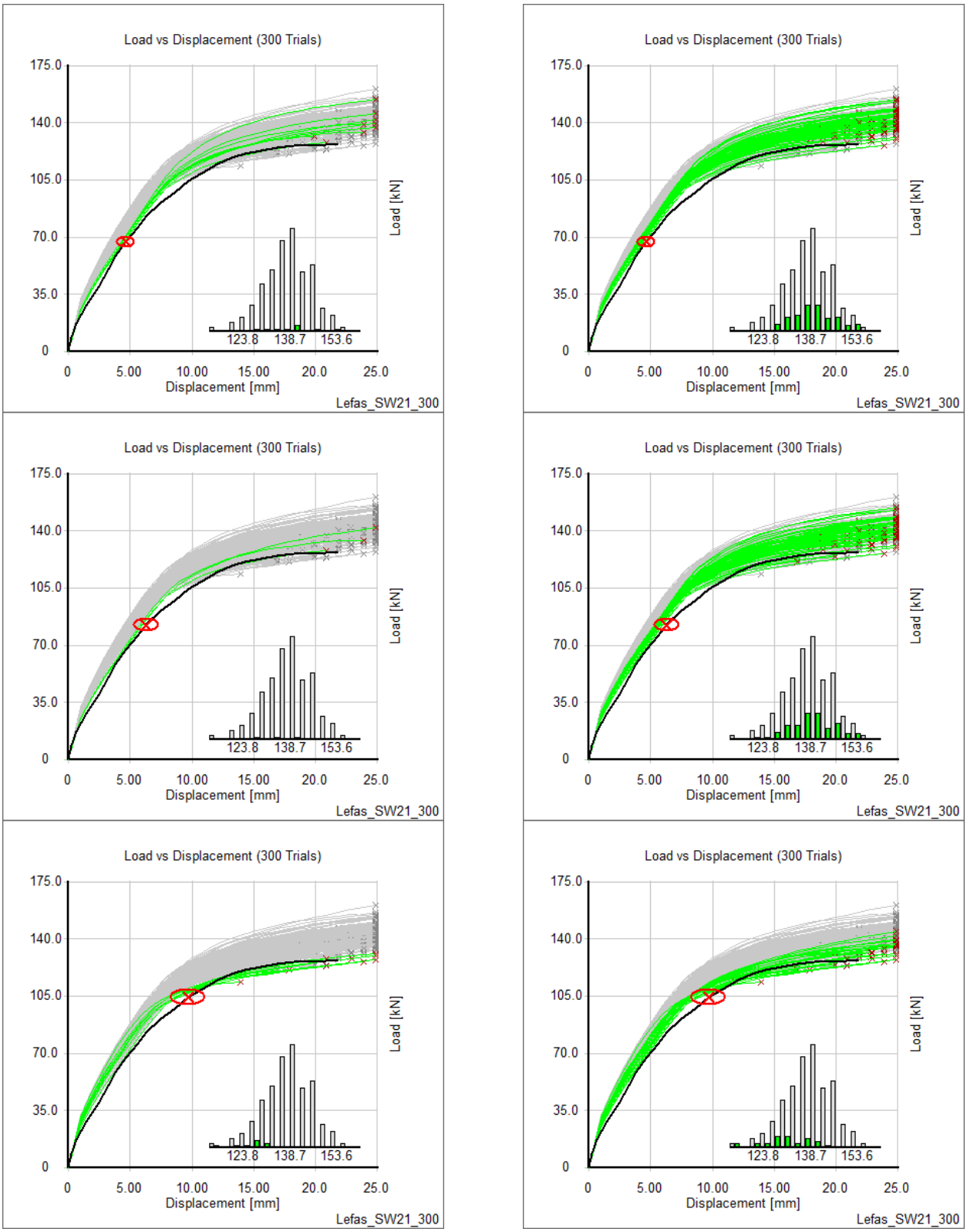


Figure 5-19 Analysis of Specimen SW21 considering the secondary displacement at (a) 37% (b) 74% (c) 85% ultimate load. (d) – (f) Analysis not considering the secondary displacement

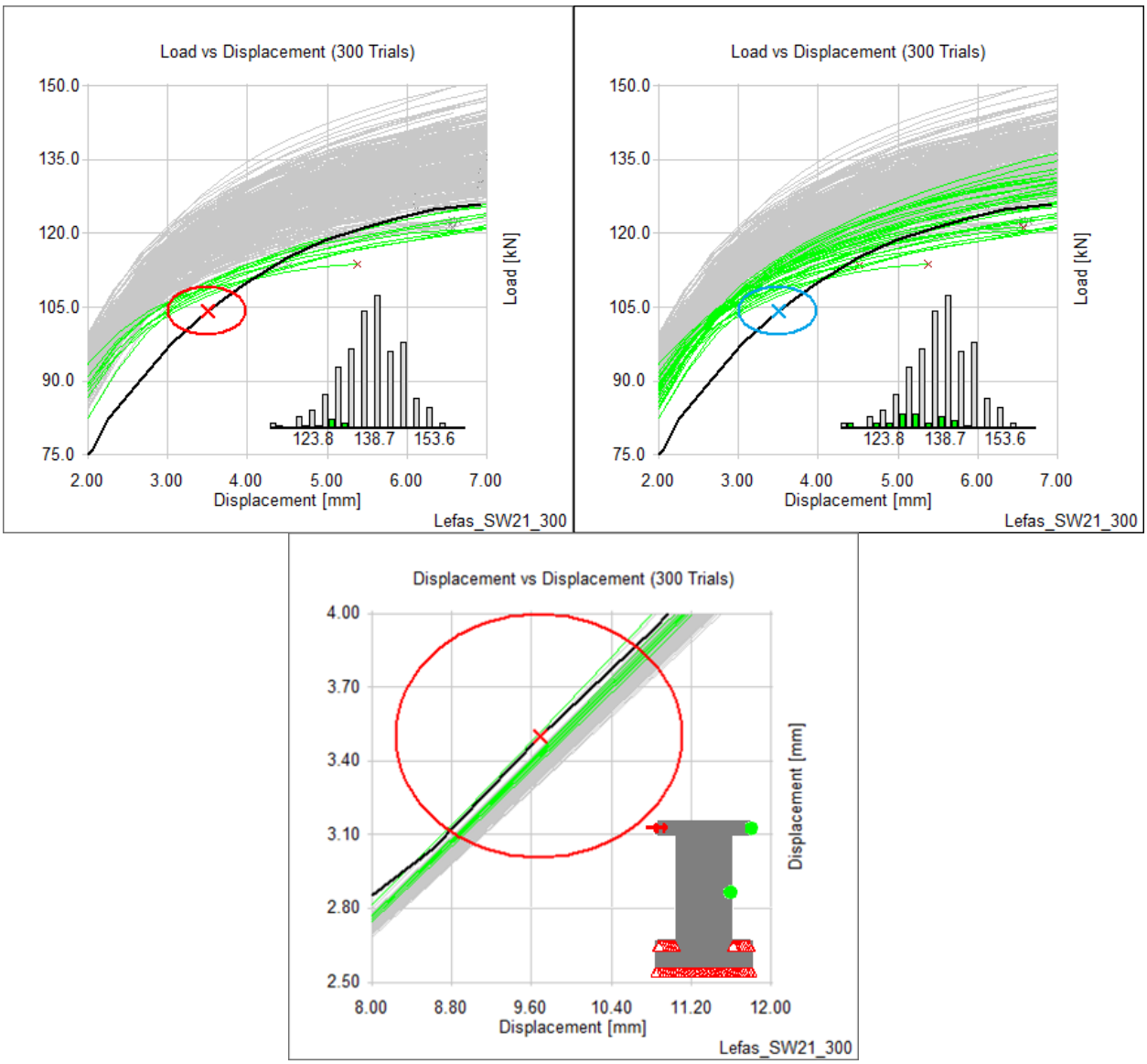


Figure 5-20 Analysis at 85% ultimate (a) considering secondary displacement (b) without secondary displacement measurement but shown for comparison. (c) Relationship of top vs secondary displacement

Figure 5-20 (a) shows the reaction load versus the secondary displacement relationship of the analysis at 82% ultimate load with all three measurements considered. Figure 5-20 (b) shows the same analysis without considering the mid-height displacement but with the measurement sensitivity shown for comparison. Figure 5-20 (c) shows the relationship between the top displacement (horizontal axis) and the mid-height displacement (vertical axis). This shows that for the same top displacement, the experimental mid-height displacement was larger than predicted. This suggests that the base of the wall in the experiment exhibited less fixity than modelled thus producing a softer response than predicted. The inclusion of the larger mid-height displacement

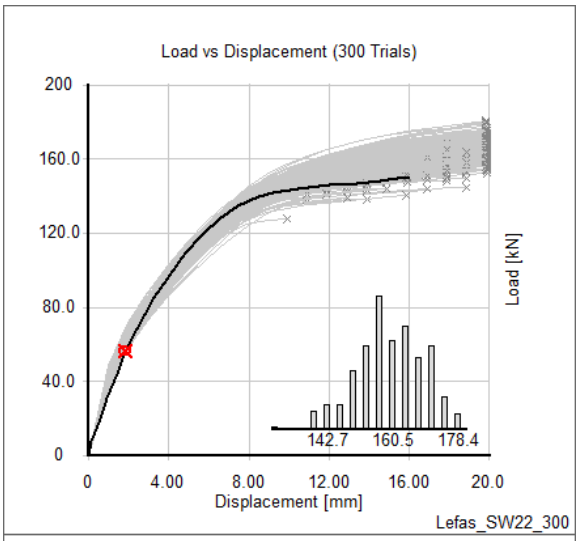
captures, as best as it could, this weaker base behaviour and identifies trials more similar to the experimental boundary conditions.

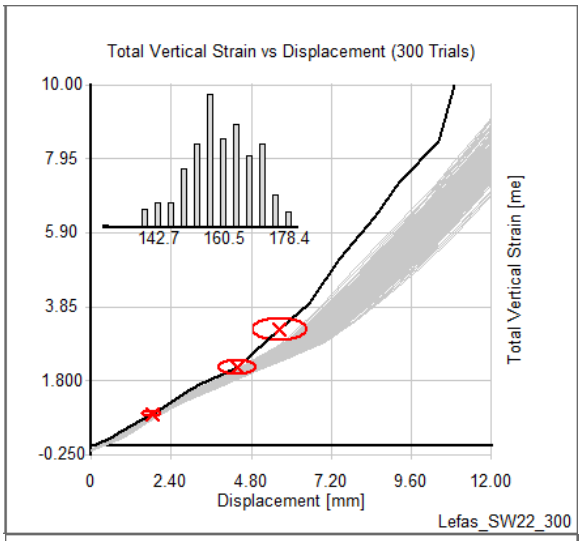
5.2.10 Specimen SW22

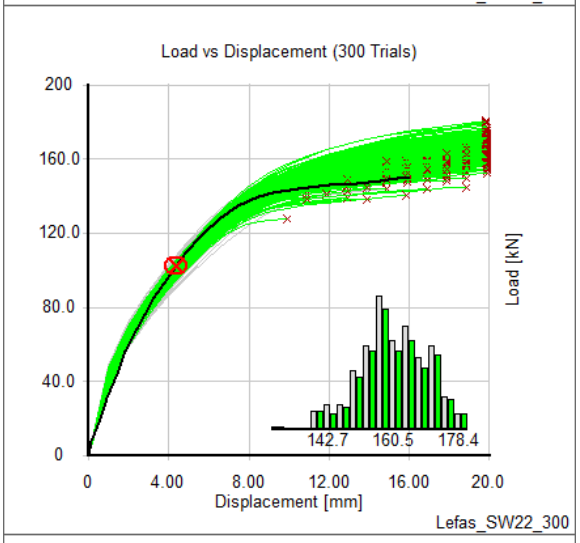
Four measurements were examined for Specimen SW22: reaction load, top displacement, mid-height displacement, and strain in the vertical reinforcement in the tension zone of the shear wall. Since the reinforcement was smeared in this analysis, the elemental vertical strain was taken as the steel strain. The location of all four measurements are shown in Figure 5-21 (f). With the use of four measurements, the ability to match a significant amount of trials was reduced.

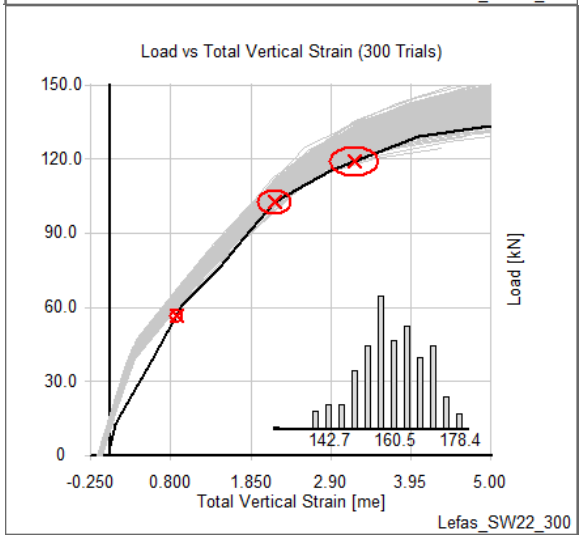
Figures 5-21 (a) – (c) show the analyses at 38%, 68%, and 79% of the ultimate load. Analysis of the first load stage examined produced no trials with high confidence. The second load stage examined considered most trials as high confidence resulting in an average failure load of 160.5kN while the last load stage examined identified a few trials as high confidence resulting in an average failure of 150.6kN. The average failure load of all trials was 160.5kN while the experiment failed at 150.0kN.

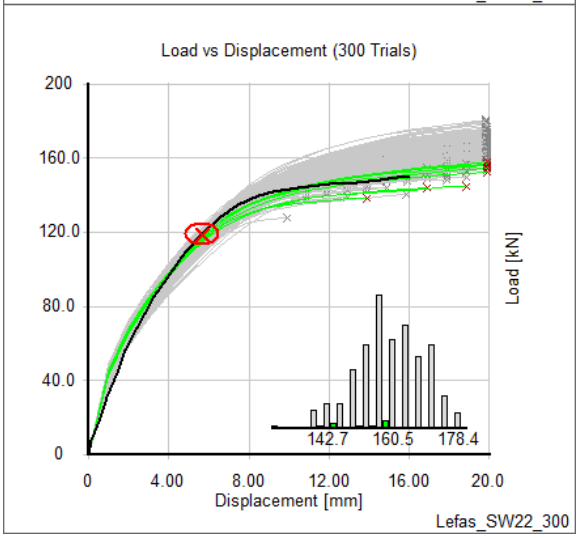
The sporadic behaviour between the three load stages examined can be attributed to the reinforcement strain measurement. At the first load stage, the experimental load versus reinforcement strain relationship in Figure 5-21 (d) sharply deviated from predicted. It was possible that the strain gages were zeroed after the application of the axial load on the shear wall since no initial compressive strain was reported. At the second load stage examined, the sensitivity of the measurements increased faster than the variation in the predicted responses and resulted in most trials being considered high confidence. Finally, at the last load stage examined, the experimental reinforcement strain versus top displacement relationship, seen in Figure 5-21 (e), deviated from the predicted responses and resulted in the elimination of many trials.











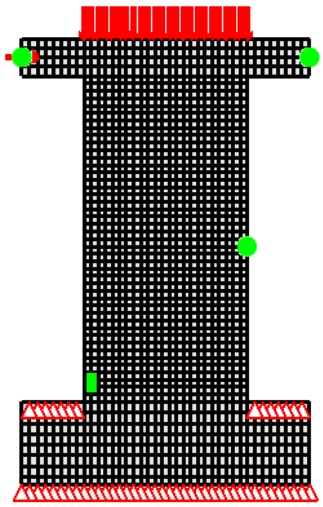


Figure 5-21 Analysis of Specimen SW22 considering load, top displacement, secondary displacement, and vertical reinforcement strain at (a) 38% (b) 68% (c) 79% ultimate load. (d) Load vs reinforcement strain relationship e) Top displacement vs reinforcement strain relationship (f) Locations of measurements taken

To investigate the effects of each type of measurement, three more analyses were conducted at the last load stage examined with different combinations of measurements used. Table 5-1 shows the three combinations and resulting average failure loads.

Table 5-1 Combinations of measurements used and failure load results

Measurements Used in Analysis						
Top Displacement	Reaction Load	Mid Height Displacement	Reinforcement Strain	Trials Identified	Failure Load [kN] Avg.	COV
✔	✔	✔	-	281	161.0	5.3%
✔	✔	-	✔	10	149.2	4.9%
✔	-	✔	✔	19	152.4	4.3%

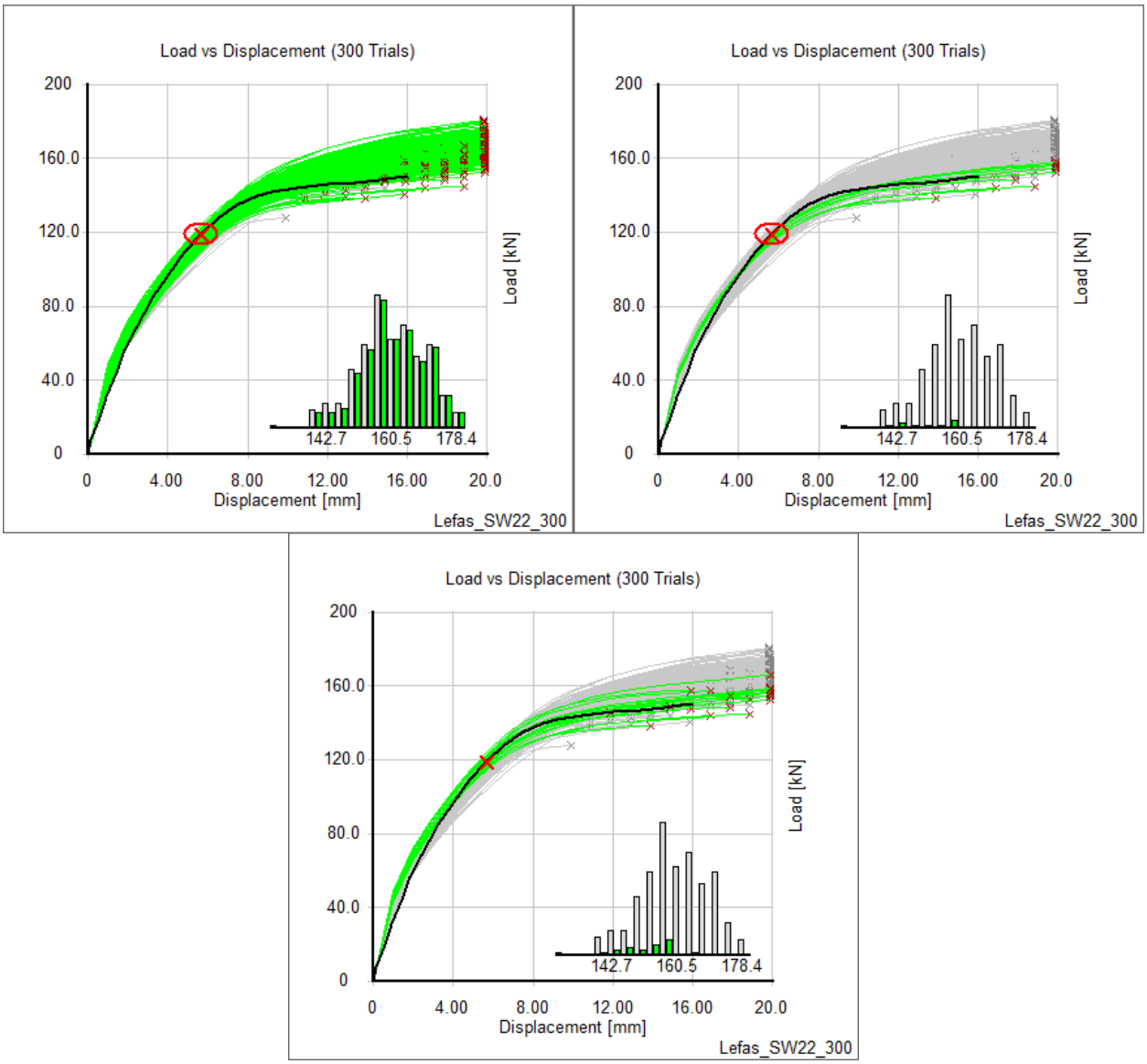


Figure 5-22 Analysis of Specimen SW22 without considering (a) reinf. strains (b) mid height displacements (c) reaction loads

Reinforcement strains appeared to be the most influential parameter. For the analyses that included reinforcement strains, the top matched trials had the lowest values of reinforcement yield strengths. These trials exhibited yielding relatively earlier while having relatively stiffer initial responses. Both characteristics matched the experiment and thus the trials identified with the inclusion of reinforcement strains produced the most accurate average failure loads. Finally, the analyses showed that the reaction load was not required to produce accurate results in this case. Similar investigations were conducted at the first two load stages examined. However, in most cases the early stage behaviour did not significantly alter the failure load distribution.

5.2.11 Specimen SW23

Three types of measurements, reaction load, top displacement, and vertical reinforcement strain in the tension zone, were used for Specimen SW23. Unfortunately, the reinforcement strains obtained differed largely against the predicted strains as seen in Figures 5-23 (a) and (b). Once again, the strains appeared to have been zeroed after the application of axial load resulting in offsets between the experimental and predicted results. In this case, the effect was larger due to the larger axial load compared to SW22 (Lefas et al., 1990). Therefore, the use of early stage reinforcement strain measurements did not change the results.

Furthermore, the use of reaction load and top displacement measurements also did not provide any conclusive results as the experimental load displacement response was significantly different from the predicted responses as seen in Figure 5-23 (c). The first two load stages examined at 31% and 49% of the ultimate load resulted in little to no trials matched while the last load stage examined at 76% matched too many trials without significantly changing the failure load distribution.

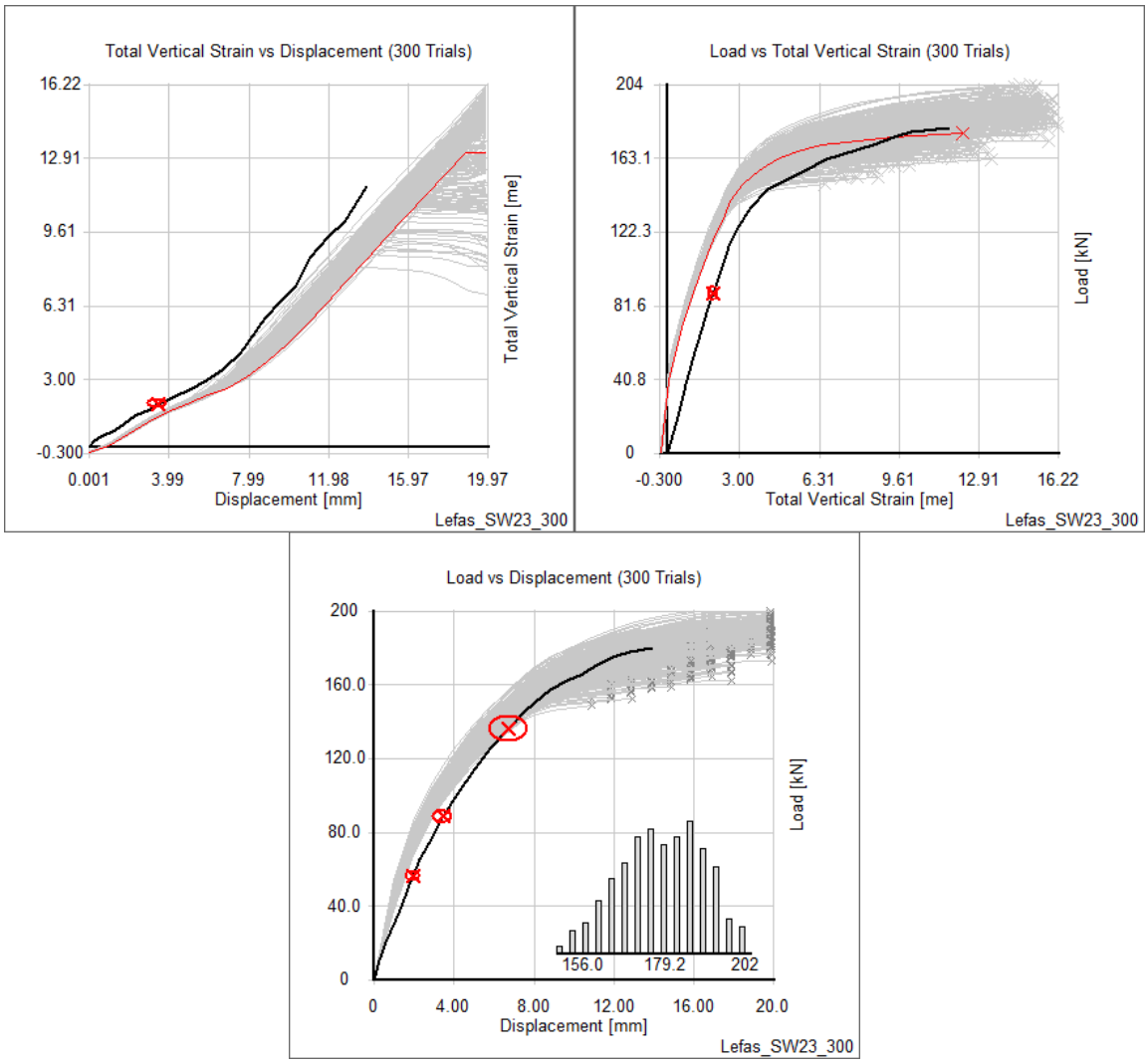


Figure 5-23 (a) Reinforcement strain vs top displacement relationship (b) Reaction vs top displacement (c) Load displacement relationship for Specimen SW23

5.2.12 Specimen B4

The stochastic results produced for Specimen B4 were inadequate for analysis considering early stage field measurements. The experimental behaviour and especially many of the predicted behaviours were extremely ductile since the failure was due to eventual rupture of tensile reinforcement. The structural analysis using VecTor2 was displacement controlled. Thus, to completely fail certain trials, final displacements of up to 1000mmm had to be reached. Since analysis parameters such as the number of load stages and the load step had to be set for all trials, relatively large load steps (10mm) were chosen to save on computational costs. However, due to the ductile nature of the specimen, the experiment reached 75% of the ultimate load within 20mm. Therefore, within the range of loading where most field measurements would be taken, the

predicted responses were very coarse, composed of only two data points. Without sufficient detail in the region that requires it the most, the analysis would not be conclusive.

5.2.13 Specimen PLS4000

Five measurements were used for the analysis of PLS4000: the displacement under the point load, the reaction load, the crack width, the tensile reinforcement strain, and the shear reinforcement strain. Unfortunately, the use of five measurements resulted in no trials being matched. The major cause appeared to be the tensile reinforcement strain as seen in Figures 5-24 (b) and (c). Furthermore, the strain gages also appear to have been zeroed after self-weight unlike the predictions. However, this would have further increased the deviation between the experimental and predicted strains. A possible explanation could be the horizontal and vertical location of the strain gauges versus the model, or the strain gages were reading concentrated levels of strain due to concrete bonding effects versus the smooth distribution of strains in the analytical model.

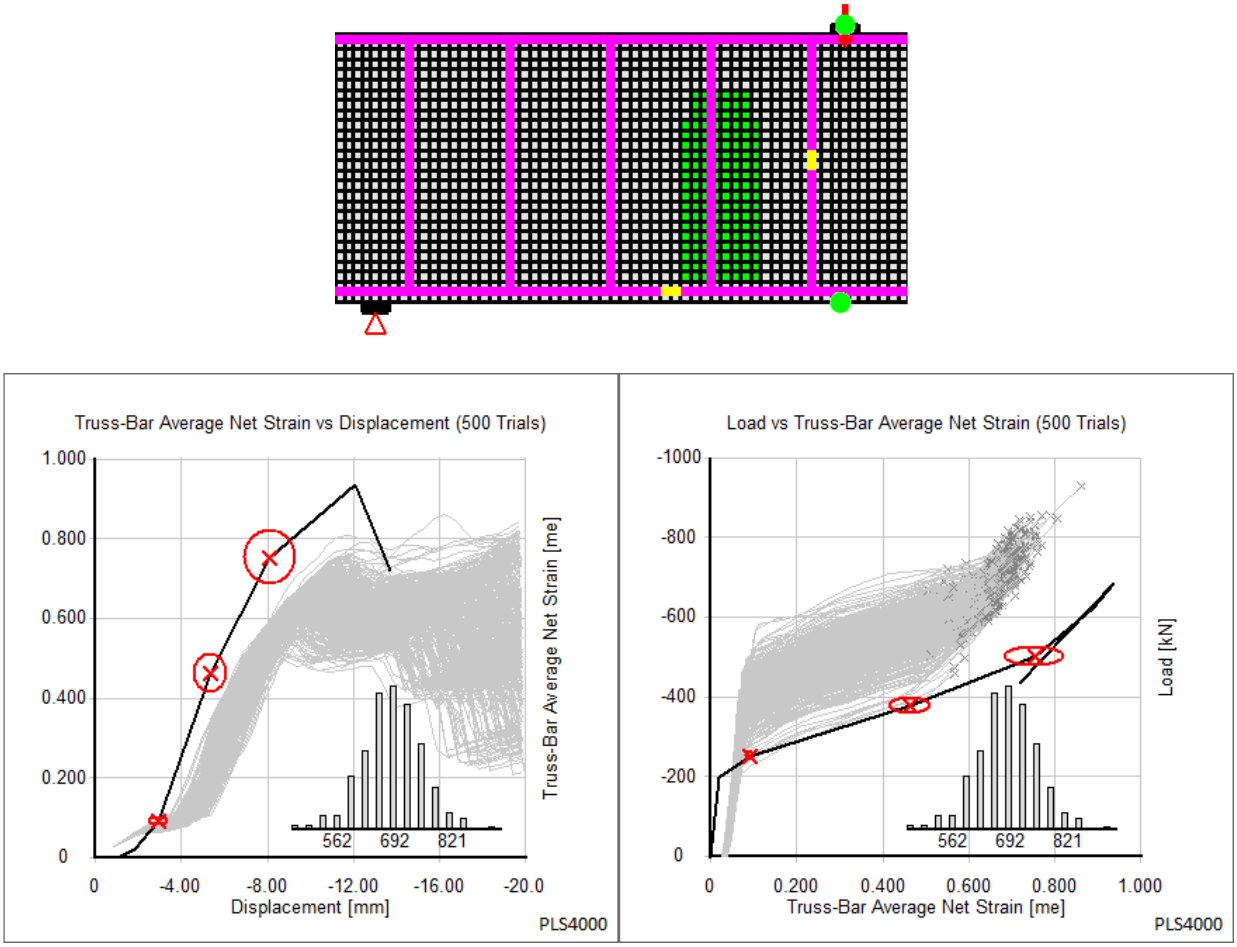


Figure 5-24 (a) Location of measurements for Specimen PLS4000 (b), (c) Tensile reinforcement strain relationships

The analyses were rerun without the tensile reinforcement strain, using four measurements. However, these only produced at most three matched trials out of 500 total trials. Instead, the number of measurements used was further reduced to combinations of three or two measurements. Analyses were conducted at three load stages corresponding to 37%, 55%, and 73% of the ultimate load. A summary of the combinations and results are shown in Table 5-2. The average failure load of all trials was 691.5kN; the experimental failure load was 684.0kN.

Table 5-2 Combination of measurements used and failure load results for Specimen PLS4000

Measurements Used				37% Ultimate Load			55% Ultimate Load			73% Ultimate Load		
Disp	Load	Crack Width	Reinf. Strain	Trials ID'd	Failure [kN] Avg.	COV	Trials ID'd	Failure [kN] Avg.	COV	Trials ID'd	Failure [kN] Avg.	COV
✔	✔	-	-	6	555.0	12.6%	82	610.6	7.3%	78	608.1	7.0%
✔	✔	✔	-	0	-	-	33	620.0	6.5%	2	604.5	8.7%
✔	✔	-	✔	3	532.2	13.8%	30	586.2	5.8%	61	603.5	7.2%
✔	-	✔	✔	280	716.4	10.0%	80	693.8	9.2%	5	676.4	11.1%

As expected, the comparison between the experimental and the predicted response for Specimen PLS4000 was similar to other shear-critical beams without shear reinforcement (Specimens VSOA1 and VSOA3). Specimen PLS4000 contained shear reinforcement in the west span from which measurements were taken but not in the east span where the failure was designed to occur. The use of reaction load and displacement measurements identified trials that fit the experiment well until failure. However, all the identified trials failed prematurely, as seen in Figures 5-25 (a) – (c).

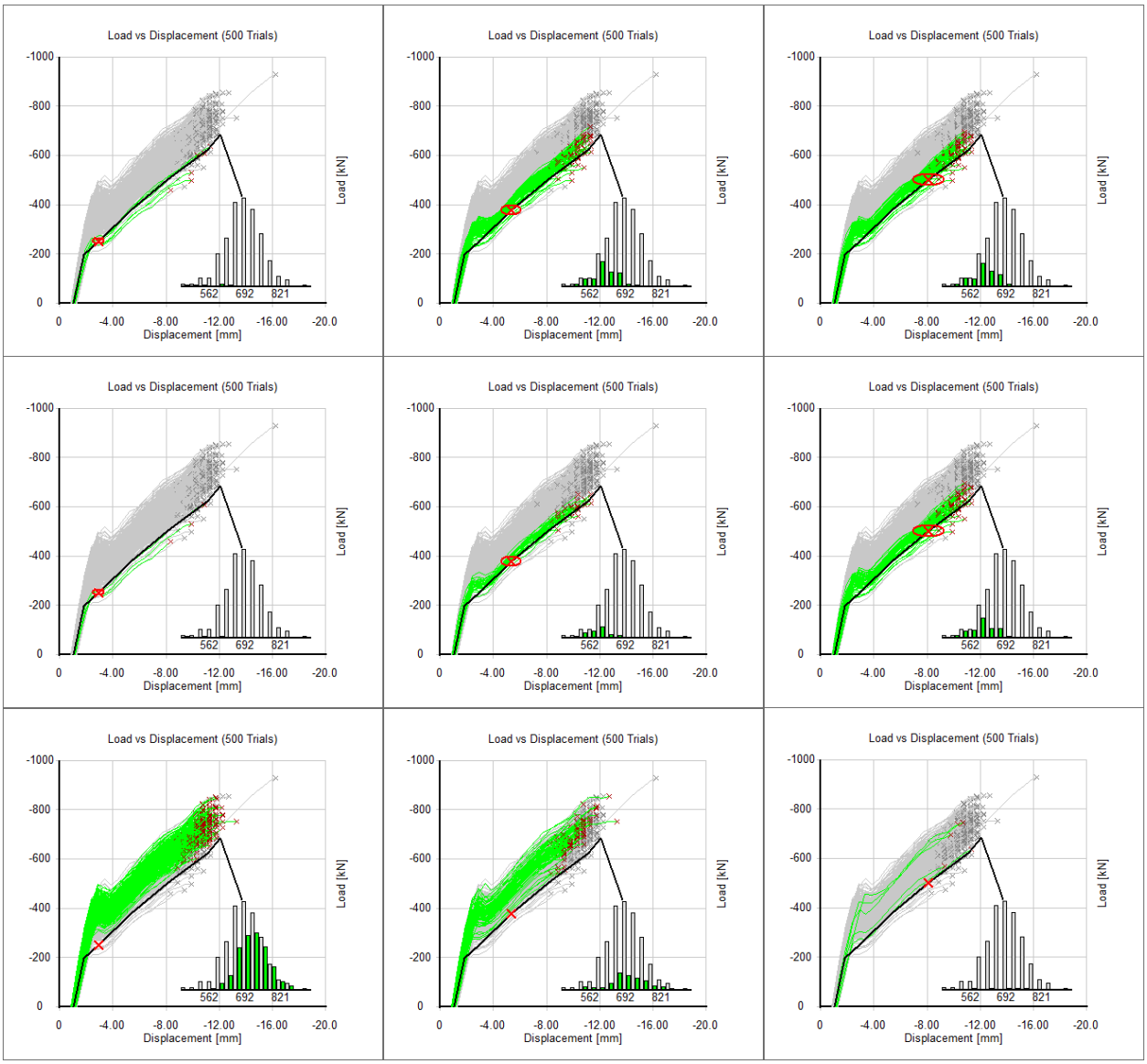


Figure 5-25 Analysis considering load displacement measurements at (a) 37% (b) 55% (c) 73% ultimate load. (d) - (f) Also considering crack widths. (g) - (i) Considering displacements, crack widths, and reinforcement strains

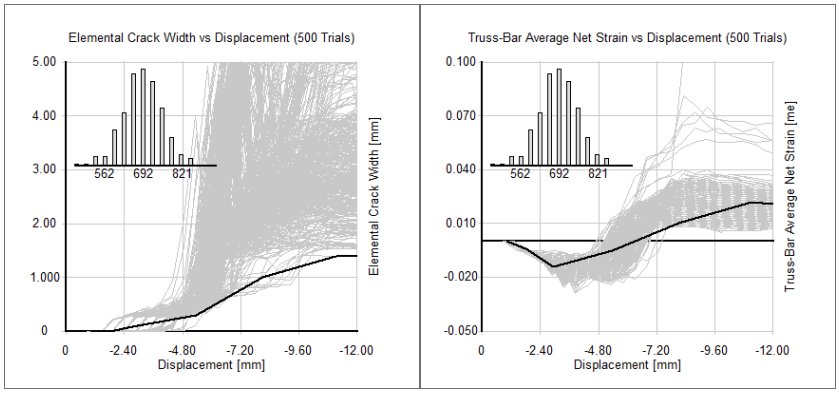


Figure 5-26 Crack width relationships for Specimen PLS4000

The use of crack width measurements, and in this case shear reinforcement strains, eliminated weaker trials since both types of measurements indicate the extent of diagonal crack progression, as seen in Figures 5-25 (d) – (f). In the trials that failed early, larger cracks and strains were seen compared to the experiment. Therefore, the consideration of these measurements eliminated such trials. For Specimen PLS4000, the shear reinforcement measurement was a better indicator of diagonal tension failure compared to crack widths due to more stability in the predicted outputs. The crack width versus displacement and the shear reinforcement strain versus displacement relationships are seen in Figure 5-26. Unfortunately, trials that failed at more accurate, higher failure loads cracked later as well, resulting in mismatches compared to the experiment, as seen in Figures 5-25 (g) – (i) where the load measurements were not used. Such a pattern is not unusual from a VecTor2 analysis of deep shear-critical beams without shear reinforcement. None of the analyses by Sherlock was able to produce accurate average failure loads simply because no trials existed that matched the early stage behaviour as well as the failure load.

5.2.14 Specimen BN50

Four measurements were used for the analysis of Specimen BN50: reaction load, mid-span deflection, crack width, and tensile reinforcement strain. However, the crack width measurements did prove unhelpful for the analyses at the load stages examined corresponding to 39%, 58%, and 70% ultimate load. At the first load stage examined, there was little deviation in the trial responses while in the latter two load stages, the crack widths were over-estimated. It is likely that the method in obtaining predicted crack width values captured concentrated flexural cracks similar to Specimen VSOA3. Similarly, the reinforcement strain measurements also contributed little improvement. At the first load stage examined, there was low variation in the predicted response, and at the latter load stages the experimental reinforcement strains diverted from predicted behaviour. Finally, with only the load-displacement measurements to use, the analysis produced the same trend as previous shear-critical beams without shear reinforcement. Overall, no improvements to the failure load distribution could be obtained with Specimen BN50.

5.2.15 Specimen BN100

The use of reaction load, mid-span deflection, crack width, and tensile reinforcement strain measurements was attempted for the analysis of Specimen BN100 at the load stages corresponding to 54%, 68%, and 80% of the ultimate load. At the load stages examined, the crack widths were

over-predicted similar to Specimens BN50 and VSOA3. Fortunately, the reinforcement strain measurement matched better. Figure 5-27 shows the results of the analyses without crack width measurements. The average failure loads were 182.0kN, 197.1kN, and 195.6kN respectively, with the average of all trials being 193.5kN; the experimental failure load was 184.4kN.

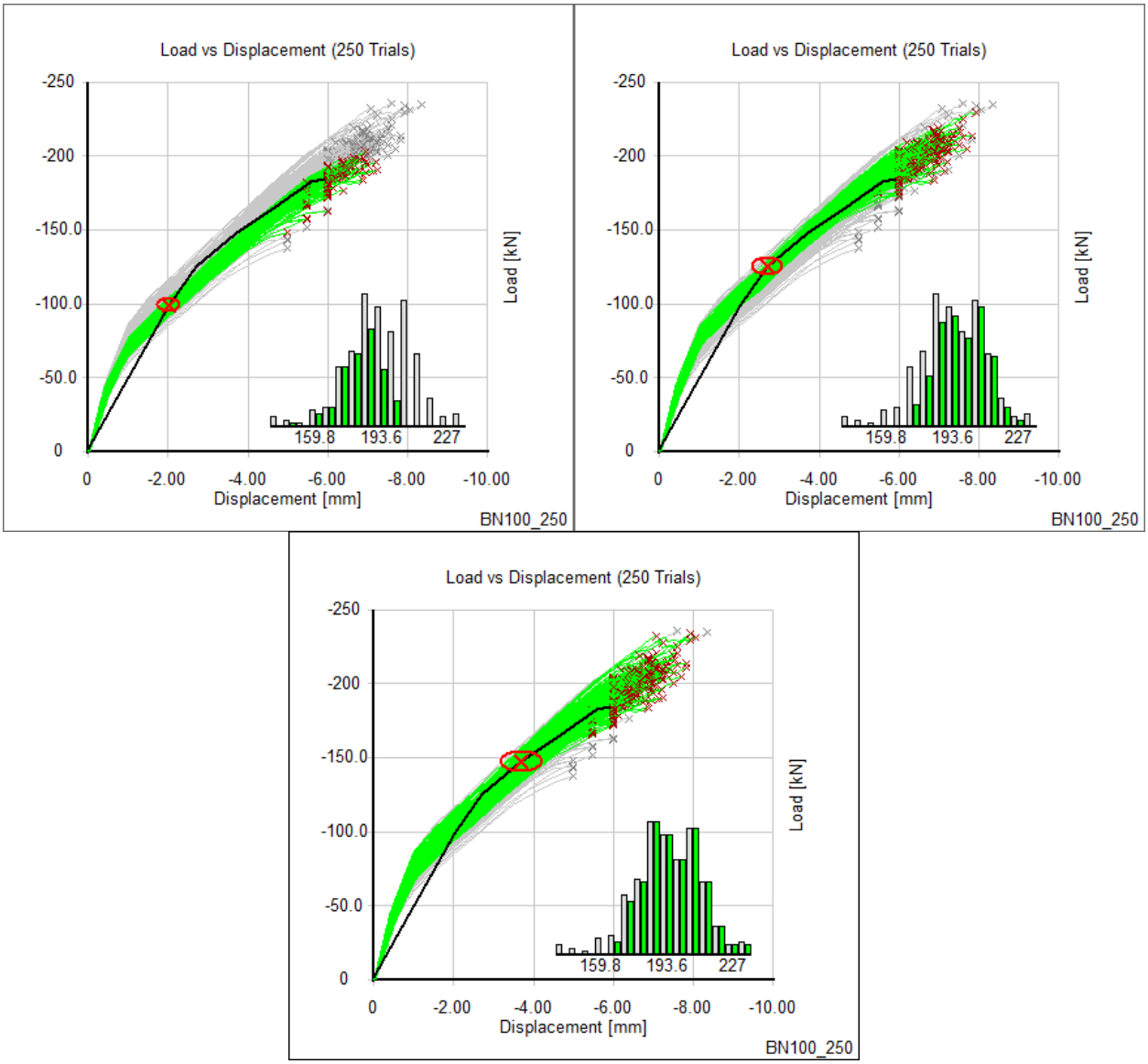


Figure 5-27 Analysis of Specimen BN100 considering load, displacement, and reinforcement strain measurements at (a) 54% (b) 68% (c) 80% ultimate load

Although the average failure load at the first load stage examined was the most accurate, the initial experimental stiffness and load-displacement response did not appear natural. The behaviour certainly did not match the predicted behaviour and the measurement was simply at an ideal location. At the latter two load stages, the experimental behaviour adjusted back to a more natural

progression and was a better match to the predicted trials. However, similar to Specimens VSA3, VSB2, VSB3, and VSC1, the experimental response was located in the middle of the band of trial results. In addition, the measurement sensitivities swelled to cover most of the generated variation resulting in very little change to the overall failure mode distribution as seen in Figures 5-27 (b) – (c). The reinforcement strains were also well predicted as seen in Figure 5-28.

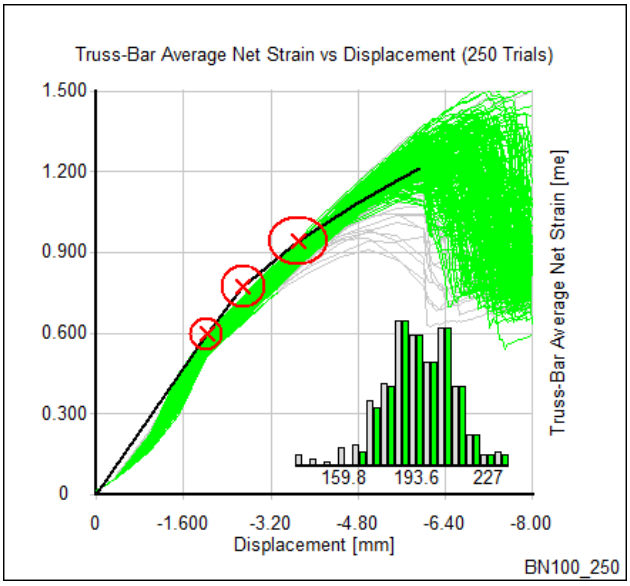


Figure 5-28 Displacement vs reinforcement strain relationship for Specimen BN100

5.2.16 Specimen BN100D

Reaction load, mid-span displacement, crack width, and tensile reinforcement strain measurements were used for the analysis of Specimen BN100D. Normally, the reinforcement strain at mid-span was used. For Specimen BN100D, the strain gage data appeared sporadic as seen in Figure 5-29 (a) and thus the next available strain gage was used as seen in Figure 5-29 (c).

At load stages corresponding to 40%, 60%, and 70% of the ultimate load, seven combinations of the four measurements were investigated, summarized in Table 5-3. The average failure load of all trials was 262.7kN while the experimental failure load was 250.6kN. Figure 5-30 shows the analysis with all four measurements at 70% ultimate load.

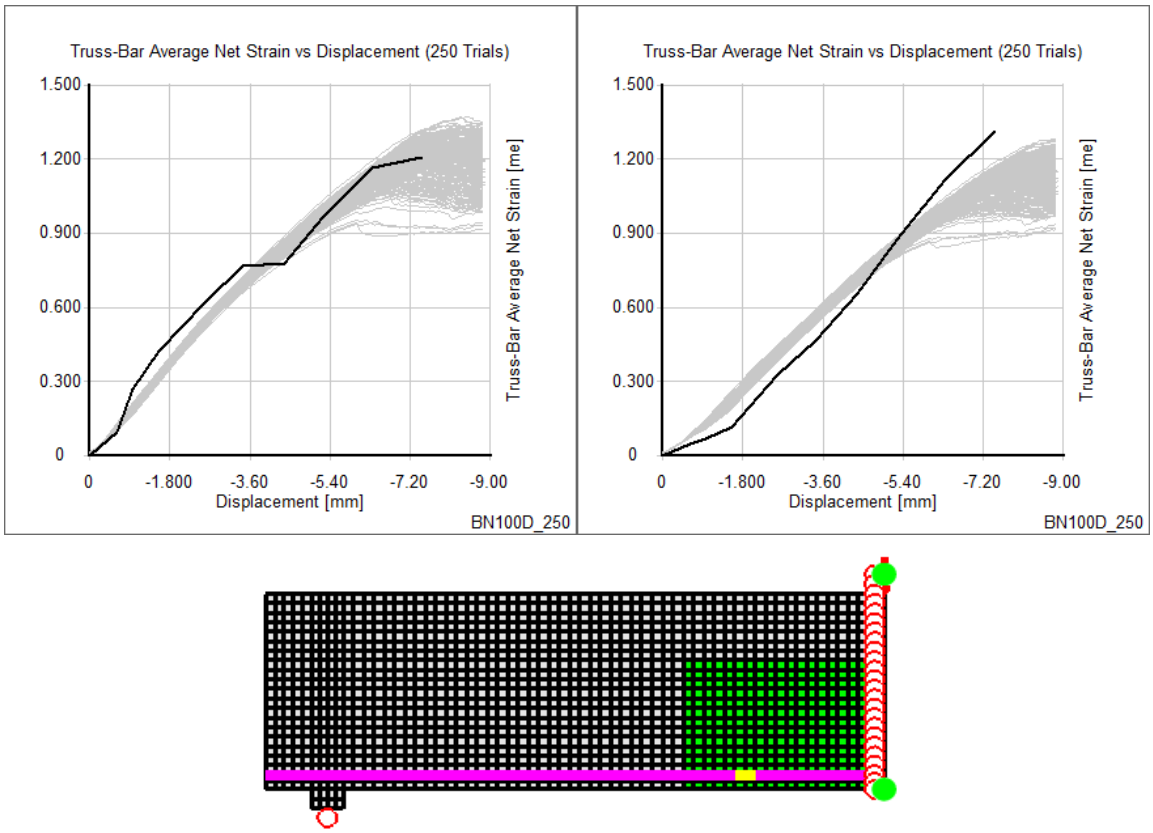


Figure 5-29 (a) Mid-span reinforcement strain (b) At next strain gage (c) Location of all measurements taken for Specimen BN100D

Table 5-3 Combinations of measurements used and failure load results

Measurements Used				40% Ultimate Load			60% Ultimate Load			70% Ultimate Load		
Disp	Load	Crack Width	Reinf. Strain	Trials ID'd	Failure [kN] Avg.	COV	Trials ID'd	Failure [kN] Avg.	COV	Trials ID'd	Failure [kN] Avg.	COV
✔	✔	✔	✔	-	-	-	18	269.8	2.4%	57	257.1	3.2%
✔	✔	✔	-	215	260.3	6.2%	129	251.3	5.5%	76	260.8	3.8%
✔	✔	-	✔	-	-	-	22	269.4	2.7%	133	250.5	5.3%
✔	-	✔	✔	-	-	-	53	273.2	5.3%	188	270.0	4.8%
✔	✔	-	-	225	261.4	6.3%	137	250.9	6.2%	152	252.5	6.3%
✔	-	✔	-	250	262.7	7.1%	250	262.7	7.1%	201	269.2	4.9%
✔	-	-	✔	-	-	-	72	273.9	5.1%	250	262.7	7.1%

At the first load stage examined, consideration of the reinforcement strain rules out all trials as high confidence as seen in Figures 5-30 (c) and (d), while consideration of only other measurements consider most trials as high confidence and not affecting the overall failure load distribution. From the second load stage examined, consideration of only displacement and load measurements produced the best result. The consideration of crack measurements favoured trials with less cracking and delayed failure due to significant diagonal cracks, as seen in Figure 5-30

(b), and raised the average failure load away from the experimental failure load. Unlike previous shear-critical specimens, Specimen BN100D contained distributed reinforcement that controlled concrete cracking. It is possible that the uniformly modelled distributed steel delayed significant cracking more than the experiment. The initial crack measurements identified trials with stronger concrete tensile strength. However at failure, which considers the combined resistance of the concrete and the distributed reinforcement, the trials with weaker concrete, and numerically stronger crack control effects, actually matched the experiment better. This resulted in more accurate results not using crack width measurements.

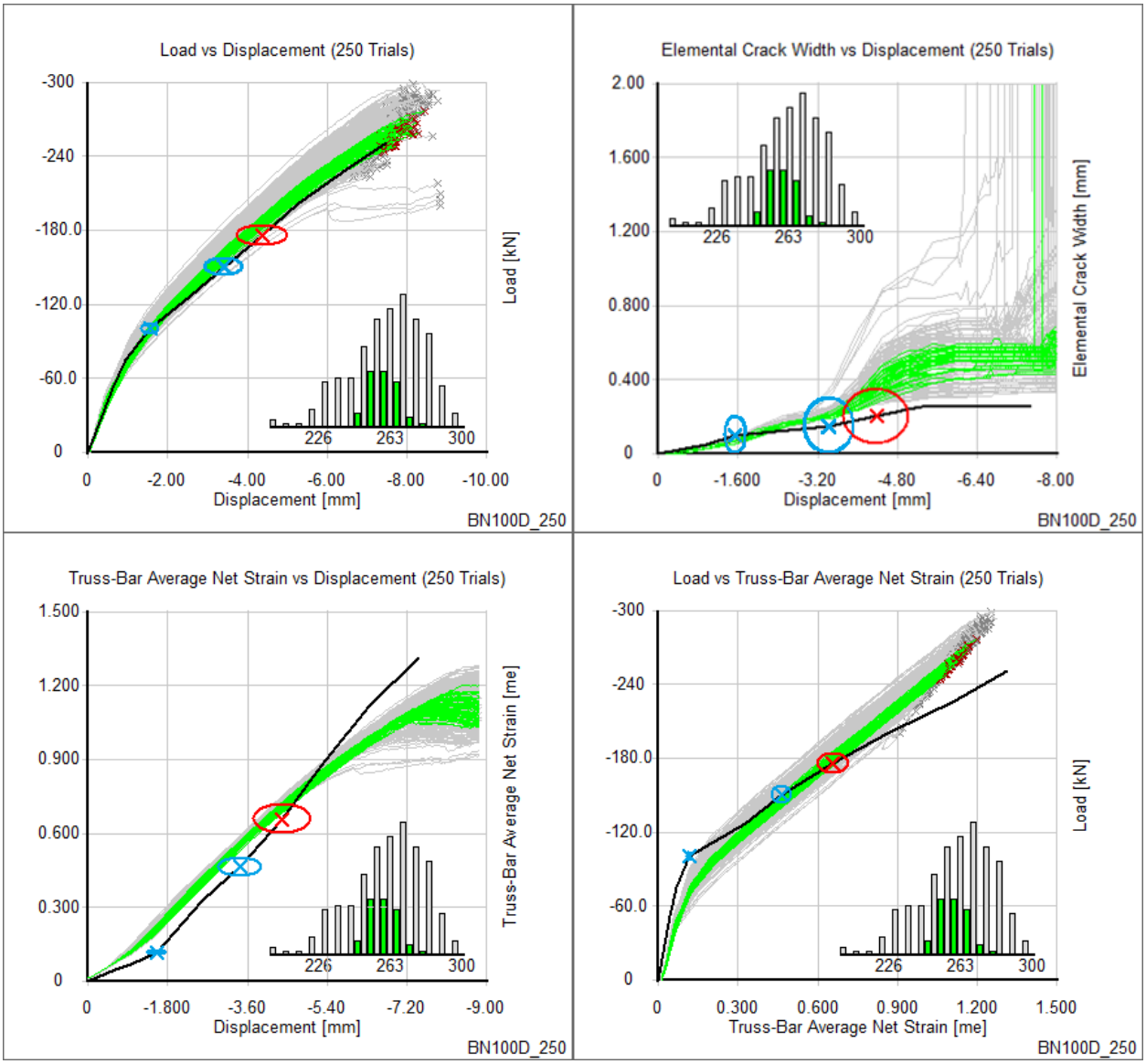


Figure 5-30 Analysis of Specimen BN100D considering all measurements at 70% ultimate load. Measurements from the first two load stages are also shown

The reinforcement strain measurement favoured trails with less initial elongation. However, as seen in Figure 5-30 (d), the experimental reinforcement exhibited much more yielding at higher loads compared to the predictions. The effects of reinforcement strains appear to be more detrimental compared to the crack widths. Figure 5-30 (c) shows that the second load stage marker completely eliminated trials outside its sensitivity boundary. While Figure 5-30 (b) shows that the crack width marker did not definitively eliminate any trials, however, trials with larger deviations were more likely to be eliminated based on combined effects from other measurements. At the last load stage examined, similar effects were seen as in the second load stage. However, the magnitudes of the effects from crack width and reinforcement strain measurements switched based on location of the measurements against the trial distributions.

5.2.17 Specimen BH50

Specimen BH50 was another specimen that produced unexpected experimental behaviour. The reaction load, mid-span deflection, crack width, and tensile reinforcement strain measurements were used to improve the prediction. Three load stages were examined at 40%, 60%, and 78% of the ultimate load. The reinforcement strains, however, were underestimated and resulted in no matched trials. Using the remaining three measurements, the average failure loads were 148.9kN and 144.6kN for the first and second load stage examined. The last load stage examined did not produce any matches due to overestimated crack width values. The average failure load of all trials was 149.8kN and the experimental failure load was 130.1kN. Figure 5-31 shows the results.

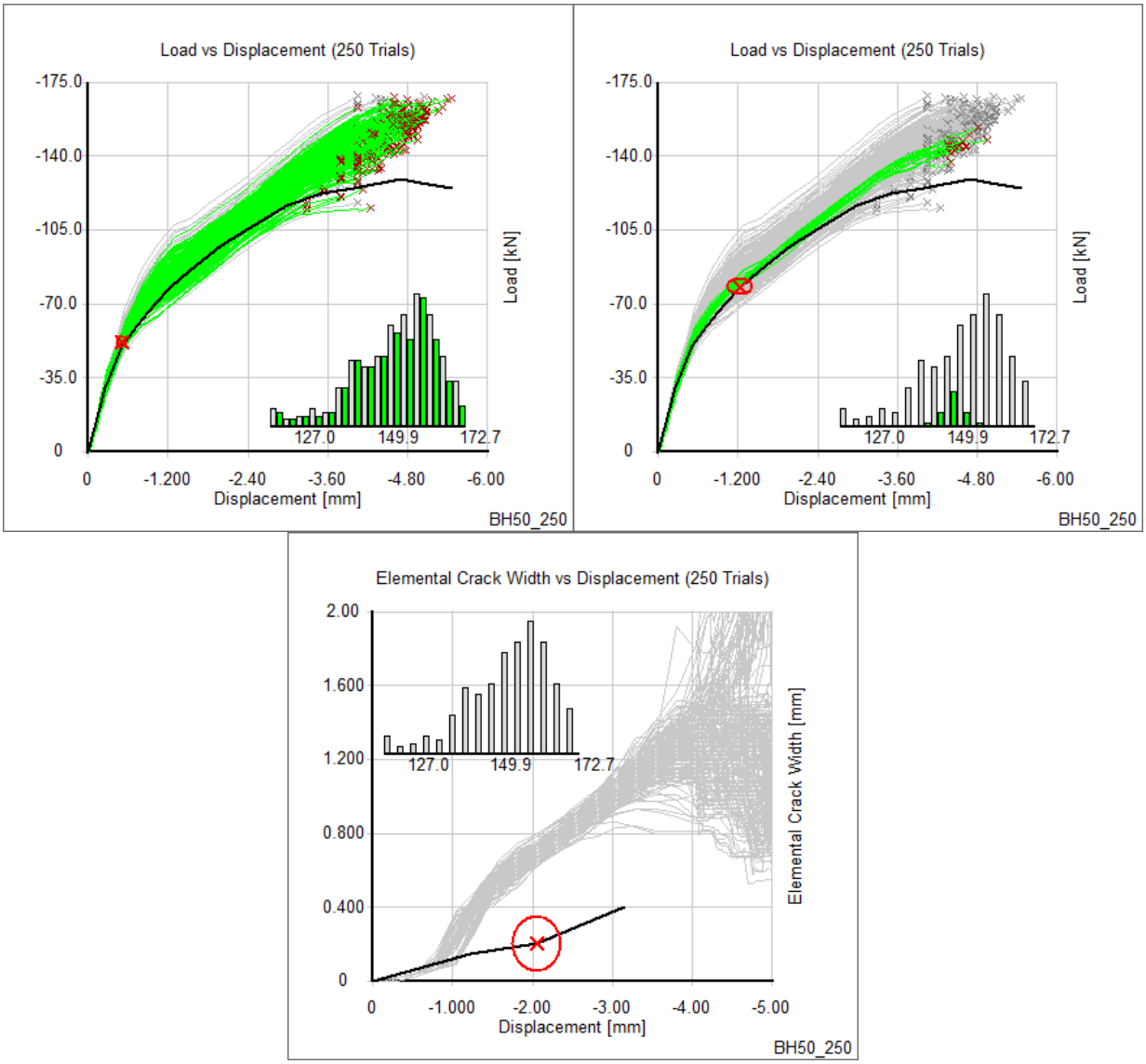


Figure 5-31 Analysis of Specimen BH50 considering load, disp., and crack width at (a) 40% (b) 60% (c) 78% ultimate load

At the first load stage examined, there was very little variation overall and most trials were deemed likely. At the second load stage, a very tight band was identified that matched the experimental behaviour very well until failure. This was achieved due to the inclusion of the crack width measurement eliminating weaker trials. Eventually however, the experiment diverged unexpectedly at failure and the average failure load was far from the experimental failure load.

5.2.18 Specimen BH100D

Four measurements were used for the analysis of Specimen BH100D. Like Specimen BN100D, the tensile reinforcement strain measurements were taken one-eighth of the span away from mid-

span due to inconsistent strain gauge data at mid-span. Analyses with all measurements produced no matched trials at load stages corresponding to 46%, 62%, and 77% of the ultimate load. The mismatch was due to the combination of reaction load and reinforcement strain seen in Figure 5-32 (a). The crack width measurements also did not affect the analysis greatly. As seen in Figure 5-32 (b), the experimental results fit in the middle of the band of trials with most of the variation captured within the measurement sensitivity and thus all trial would be considered likely.

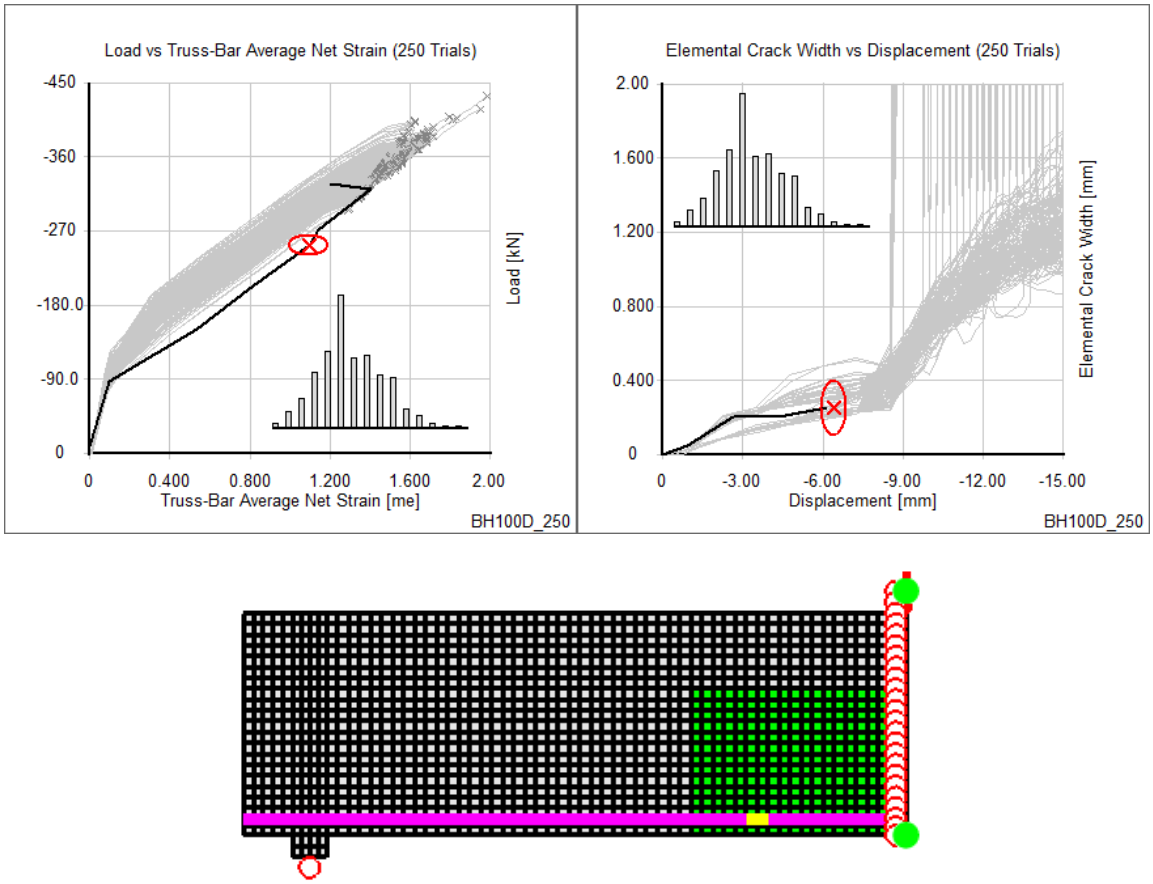


Figure 5-32 (a) Reinforcement strain (b) crack width relationship for Specimen BH100D. (c) Location of all measurements

Using the load-displacement measurements, the average failure loads were 315.5kN, 317.0kN, and 311.5kN with the average of all trials being 353.6kN; the experimental failure load was 326.8kN. Figure 5-33 shows the results from all three load stages.

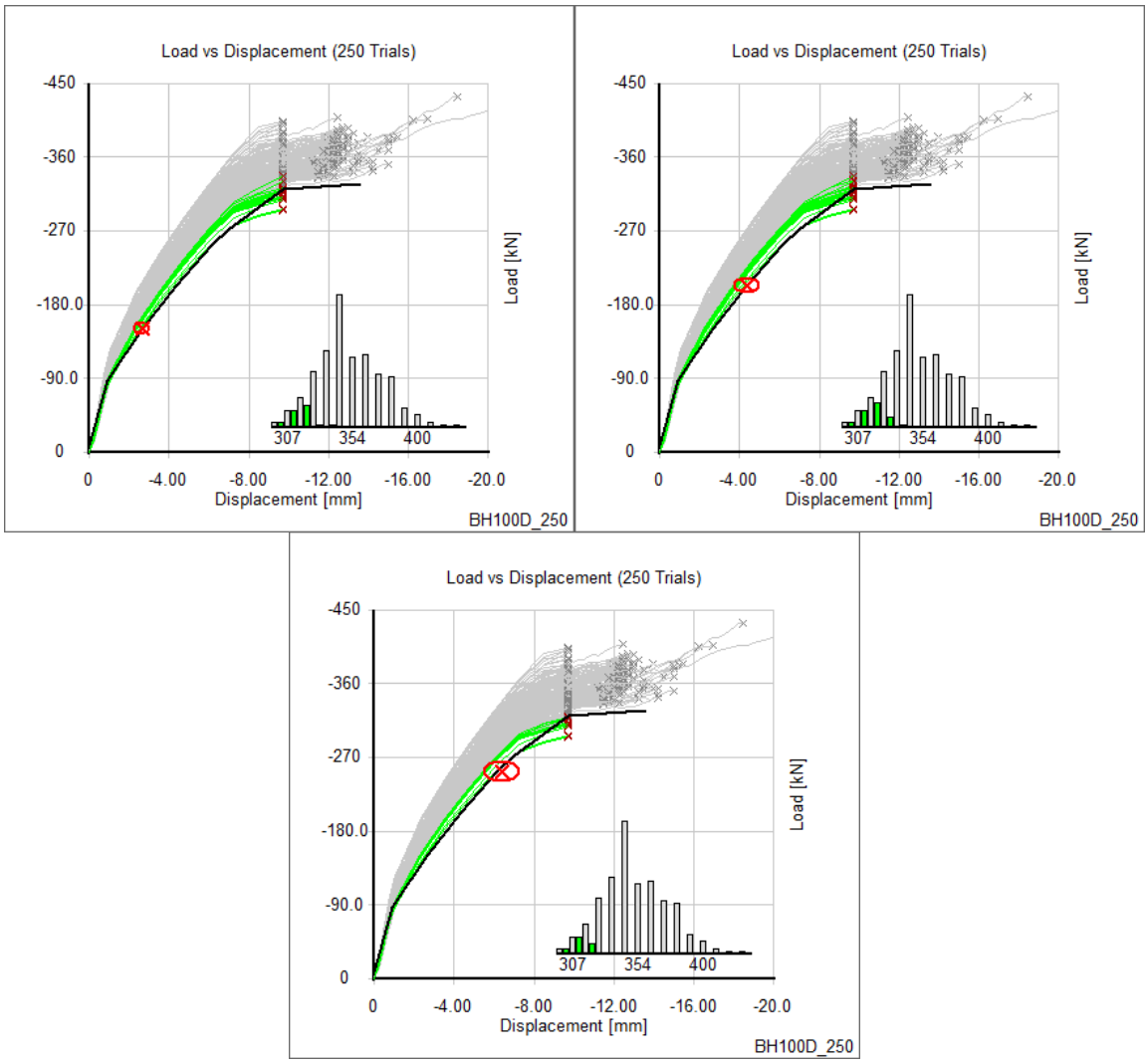


Figure 5-33 Analysis without considering crack widths at (a) 46% (b) 62% and (c)77% of ultimate load

To investigate the effectiveness of analysis without reaction load measurements, a set of analyses was conducted with the other three measurements producing average failure loads of 329.0kN, 354.0kN, and 358.1kN. The results for the first load stage are shown in Figure 5-34. The first load stage examined produced a very accurate estimate of failure load. However, Figure 5-34 (b) shows that trials with larger initial cracking resulted in smaller cracks towards failure and Figure 5-34 (c) shows the same trend with initial and final strains. Both behaviours were unexpected and yet led to accurate predictions. At the latter two load stages, the sensitivities became too large and most trials were considered likely, resulting in little change to the overall distribution.

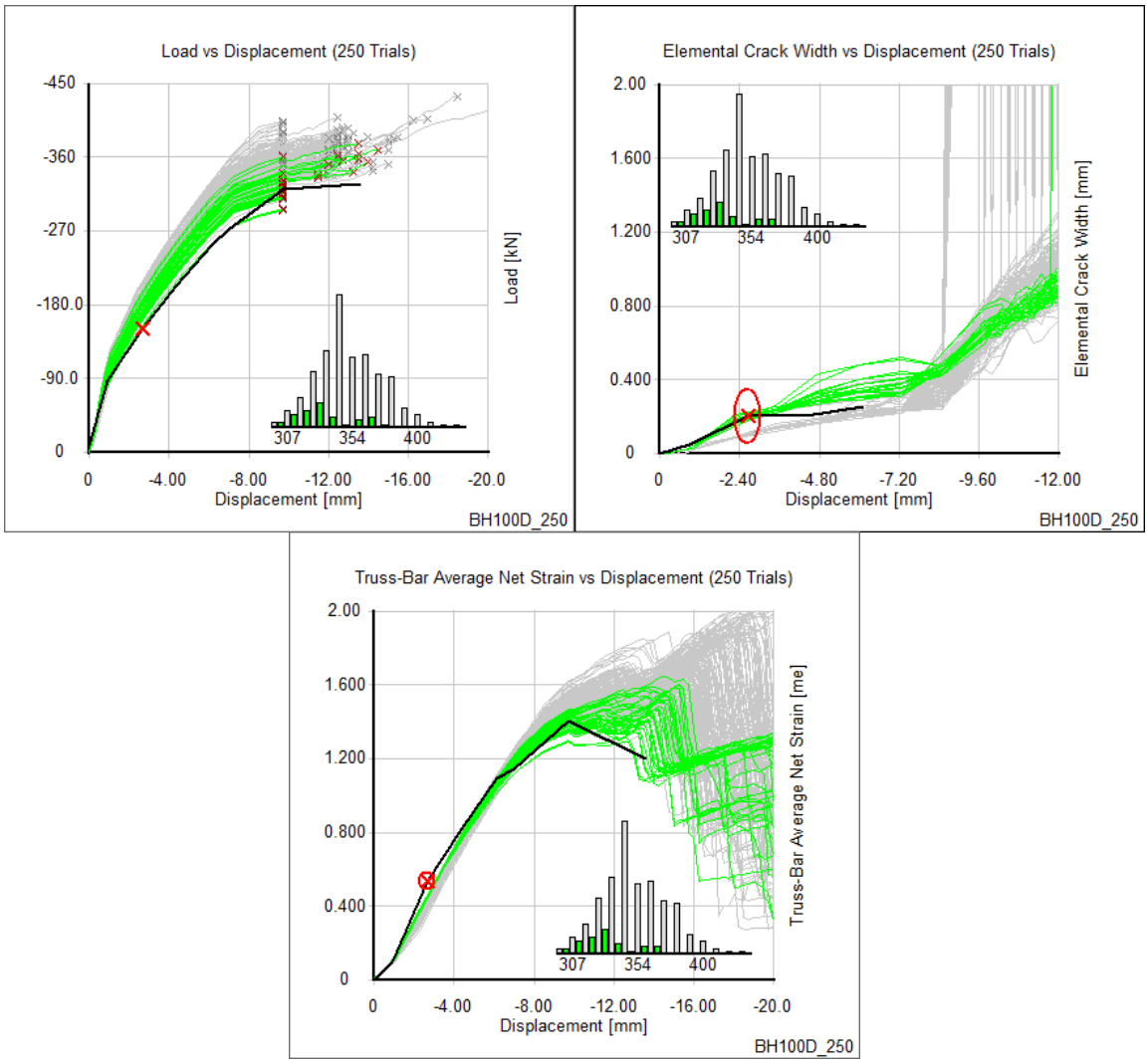


Figure 5-34 Analysis without considering reaction loads at 46% ultimate load

5.2.19 Specimen BM100

For Specimen BM100, the crack width measurements were largely overestimated. Using reaction load, mid-span displacement and tensile reinforcement strain measurements, analyses were conducted at load stages corresponding to 30%, 60%, and 75% of the ultimate load. Figure 5-35 shows the load displacement results of each analysis while Table 5-4 summarizes the average failure loads. The average failure load was 339.7kN and the experimental failure load was 335.6kN.

Table 5-4 Combinations of measurements used and failure load results

Measurements Used			30% Ultimate Load			60% Ultimate Load			75% Ultimate Load		
Disp	Load	Reinf. Strain	Trials ID'd	Failure [kN] Avg.	COV	Trials ID'd	Failure [kN] Avg.	COV	Trials ID'd	Failure [kN] Avg.	COV
✔	✔	✔	46	350.4	5.3%	12	322	5.4%	27	323.5	6.7%
✔	✔	-	226	341.1	6.5%	184	335.2	6.7%	95	327.4	6.8%
✔	-	✔	177	341.4	6.9%	56	322.7	5.3%	31	322.3	7.1%

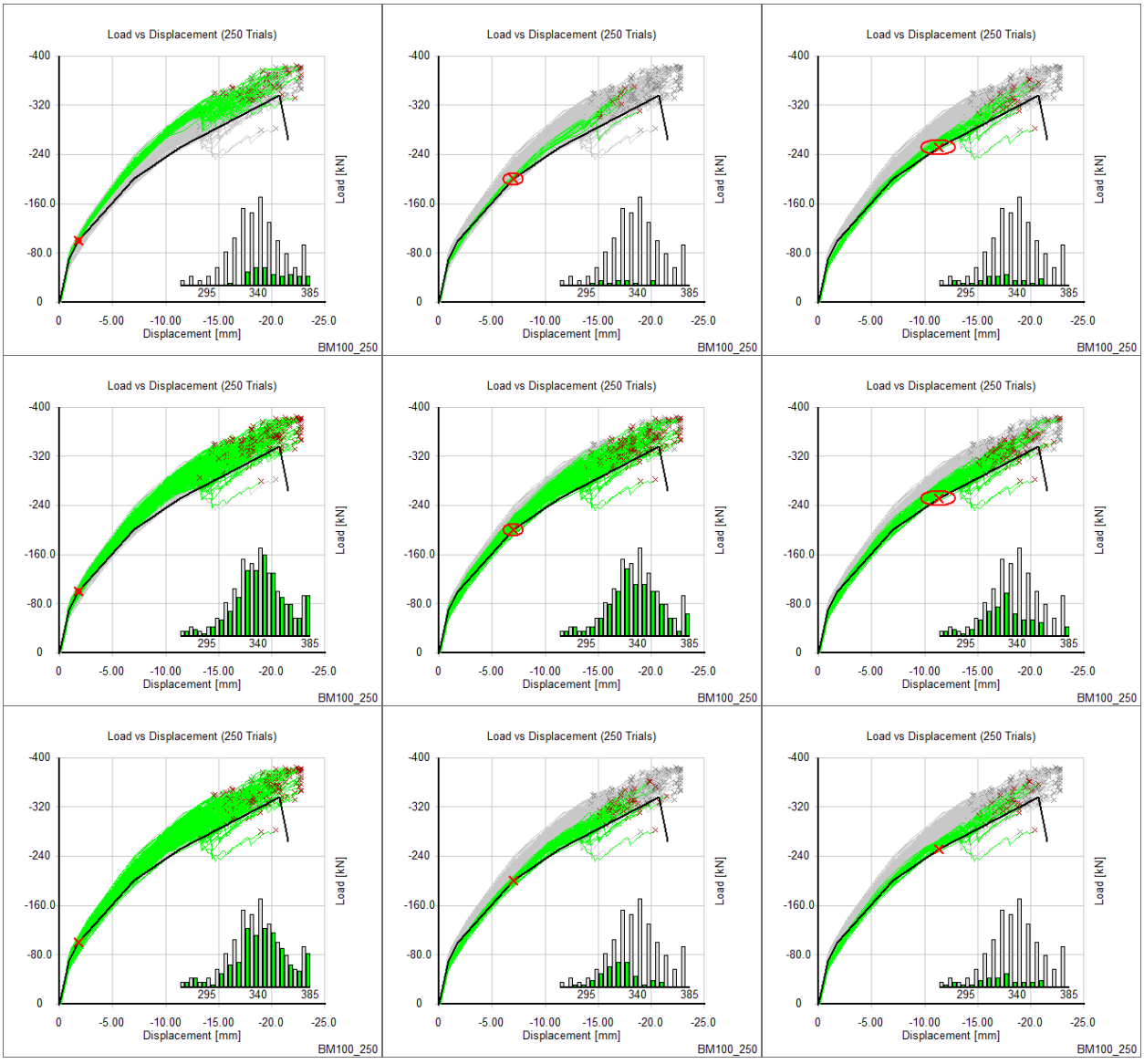


Figure 5-35 Analysis of Specimen BM100 considering load, displacement, and reinforcement strain measurements at (a) 30% (b) 60% (c) 75% ultimate load. (d) - (f) without considering reinf. strain (g) - (i) without considering reaction load

Despite containing shear reinforcements, Specimen BM100D was still a shear-critical beam with minimal shear reinforcement and the failure was due to diagonal shear after the rupturing of the

stirrups. Therefore, similar to previous specimens, trials that better matched the load-displacement response failed relatively earlier compared to the experiment as seen in Figures 5-35 (b) and (c). The consideration of longitudinal reinforcement strains appeared to eliminate stronger trials and better match the experiment. Even in the analyses without considering reaction load measurements as seen in Figures 5-35 (h) and (i), the identified trials matched the experiment particularly well. The reinforcement strain results are also shown in Figure 5-36 for the analysis considering all measurements at 75% ultimate load.

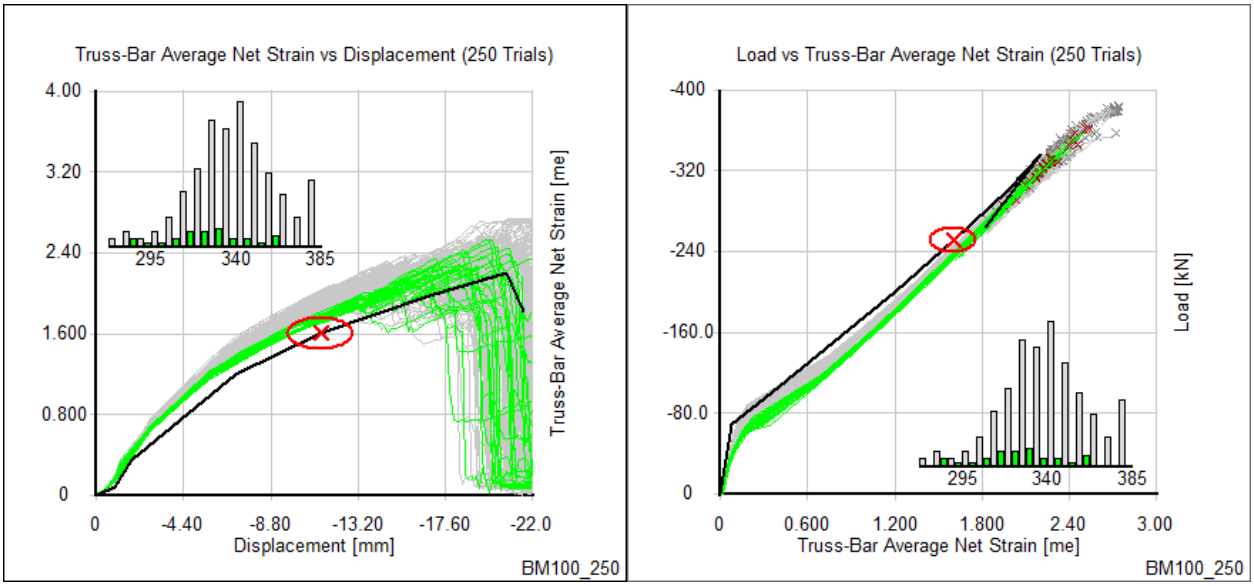


Figure 5-36 Reinforcement strain relationships for Specimen BM100

In general, the fit to the experiment improved with later load stages. However, once again, this was not correlated with improvements to the failure load prediction.

5.3 Summary of Observed Trends

In total, 176 sets of analyses were undertaken considering various combinations of field measurements and taken at load stages corresponding to minimum 30% to maximum 85% of the ultimate load. Out of the 176 analyses, 35 analyses resulted in zero matched trials. For the remaining 141 analyses, each with a predicted average failure load, the absolute error to the experimental failure loads were calculated. The average error was 7.6% which was comparable to the 7.1% average error of the stochastic analysis without field measurements but unfortunately larger. Furthermore, 35 analyses resulted in no failure load estimates which were not included in the average error calculation. In addition, the difference between the absolute errors of the

stochastic analyses without considering field measurements and the errors considering field measurements for each analysis were computed to represent the improvement with the use of field measurements. The average improvement was actually -0.99% which means that on average, the consideration of field measurements resulted in slightly worse failure load predictions.

There were a few major reasons for the lack of improvement overall. Table 5-5 shows the results of all analyses considering field measurements (FM) broken down by specimen. The experimental failure load, deterministic failure load, the average failure load from the stochastic analyses without consideration of field measurements are also shown for reference. It should be noted that the errors presented for analyses considering field measurements are not calculated as the percentage difference between the experimental failure load and the average failure load of all analyses considering field measurements for a certain specimen. Because a number of analyses were conducted with different combinations of measurements at different load stages for each specimen, the average of the absolute percentage differences for each analysis was taken as the errors presented in Table 5-5. For specimens where all analyses produced strictly over- or under-estimated failures, the resulting error term would be the same. However, in other cases, the presented error would be larger, and more appropriate, than the percentage difference between the experimental and average failure load of all analyses.

The first reason was the types of specimens used for this study. From Table 5-5, Specimens VSOA3, PLS4000, and BN50 produced the worst improvement scores. All three of these specimens failed due to diagonal tension, a failure mechanism that is expected to be under-estimated by VecTor2. Therefore, trials generated by VecTor2 that match well with early stage measurements will also produce premature failure as discussed. In fact, most other specimens that failed due to diagonal tension such as VSOA3, BN100, BN100D, and BM100 resulted in zero or negative improvement. Specimens BH50 and BH100D, however, performed well. This was possibly because the experimental tensile behaviour was weaker than modelled which cancelled out the conservative nature of VecTor2 on shear-critical beams without shear reinforcement.

Another reason for the lack of overall improvement was the group of specimens that did not generate sufficient amounts of variability. For specimens VSA3, VSB2, VSB3, VSC1, and to some extent B4, any consideration of field measurements would identify all predicted trials as likely because all trials exhibited extremely similar behavior. The result would then be equivalent to the

stochastic analysis without considering field measurements and produce zero improvement. Further break down of the results are presented in the following sections.

Table 5-5 Break-down of Sherlock performance by specimen. All failure loads expressed in kN

Specimen	Experiment	Deterministic		Stochastic		Stochastic + FM		Improvement using FM
		Failure	Error	Failure	Error	Failure	Error*	
VSOA1	165	139	15.7%	142	13.9%	140	16.0%	-2.1%
VSOA3	193	184	4.4%	186	3.2%	168	12.5%	-9.3%
VSA1	229	234	2.5%	237	3.5%	226	2.7%	+0.9%
VSA3	211	222	5.1%	225	6.5%	225	6.5%	0.0%
VSB2	183	192	4.8%	193	5.4%	192	5.1%	+0.3%
VSB3	171	176	3.4%	178	4.3%	177	4.0%	+0.3%
VSC1	141	136	3.5%	137	3.0%	137	3.0%	0.0%
VSC2	145	172	18.9%	175	21.1%	167	15.2%	+5.5%
SW11	250	249	0.4%	285	14.0%	-	-	-
SW12	340	308	9.3%	347	2.1%	-	-	-
SW15	317	280	11.7%	332	4.7%	313	3.6%	+1.2%
SW16	355	347	2.2%	392	10.4%	385	8.4%	+2.0%
SW21	127	111	12.5%	139	9.2%	135	6.5%	+2.8%
SW22	150	135	9.8%	161	7.0%	157	4.7%	+2.3%
SW23	180	152	15.7%	179	0.4%	176	2.1%	-1.6%
B4	347	317	8.8%	372	7.2%	373	7.5%	-0.3%
PLS4000	685	623	8.9%	697	1.8%	620	10.9%	-9.8%
BN50	130	116	11.1%	115	11.9%	106	18.8%	-7.2%
BN100	184	192	4.2%	194	4.9%	194	5.2%	-0.3%
BN100D	251	268	7.1%	263	4.8%	262	4.7%	+0.2%
BH50	130	161	24.1%	150	15.6%	148	13.5%	+1.7%
BH100D	327	380	16.4%	354	8.2%	331	5.0%	+3.2%
BM100	336	384	14.4%	340	1.2%	332	2.9%	-1.6%

Table 5-6 Average ratios of experiment to predicted failure loads

	Deterministic	Stochastic	Stochastic + FM
Pcalc / Pexp	1.007	0.964	1.006
COV	11.3%	8.1%	12.9%

Finally, for some specimens, the early stage behaviour did not dictate failure behaviour. Most notably for the shear wall specimens, the early stage behaviour was more related to stiffness properties while the failure was more related to strength properties. Therefore, many cases showed

distinctive bands of highlighted results at early stages that fanned out in terms of the failure stages and covered the entire failure load distribution. This then led to minimal improvement against stochastic analysis without consideration of field measurements.

Table 5-6 shows the average ratios and coefficients of variation between the experimental and predicted failure loads for all specimens. Although the closer ratios for the sets of deterministic and stochastic analyses considering field measurements suggest better performance, it is believed this was observed because of the set of specimens chosen.

5.3.1 Load Stage Progression

As expected, the effectiveness of considering field measurements changed depending on what percentage of ultimate load the measurements were taken at.

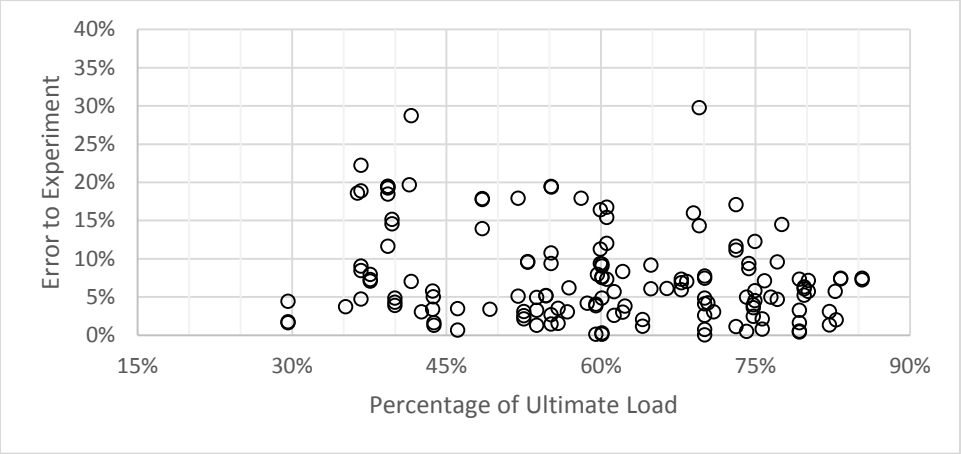


Figure 5-37 Experimental error vs percentage of ultimate load measurements taken for 141 analyses

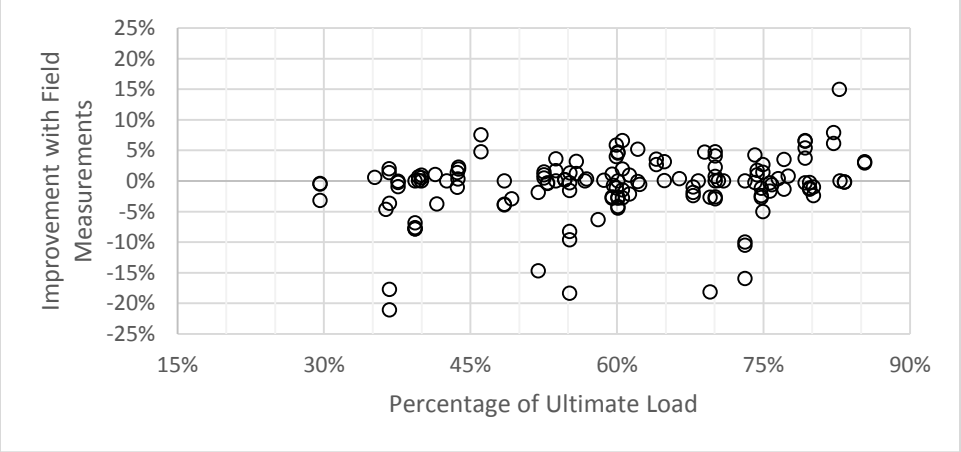


Figure 5-38 Percentage improvement vs percentage of ultimate measurements were taken for 141 analyses

Figure 5-37 shows a gradual decrease in error to the experimental failure load as measurements are taken closer to failure. Figure 5-38 shows a gradual improvement in the comparison between consideration of field measurements and without consideration. In analyses at lesser ultimate loads, the variation in predicted response tends to be small and changes, either positively or negatively, tend to be small as well. However, it would still appear that many analyses produced minimal improvements likely because the field measurements did not change the failure load distribution. Table 5-7 shows the aggregate averages by percentage of ultimate load. It also shows the break-down of percentage of matched trials which suggest that mismatches are actually more likely to occur with earlier stage data. This can be attributed again to low early stage variation and failure to match against field measurements.

Table 5-7 Error and improvement broken down by the percentage of ultimate load measurements at which field measurements (FM) were taken

%Ultimate Load FM Taken	Number of All Analyses	Number of Matched Analyses	Percentage Matched	Error to Experiment	Improvement with FM
30% - 45%	44	31	70.5%	10.3%	-3.3%
45% - 60%	43	35	81.4%	7.8%	-1.2%
60% - 75%	58	50	86.2%	7.0%	-0.9%
75% - 90%	31	25	80.6%	5.2%	+2.0%
	176	141	80.1%	7.6%	-1.0%

Figures 5-39 (a) and (b) represent the same information as Figure 5-37 and Figure 5-38 except analyses with the exact same combination of measurements are connected together to more accurately show the trend of results against later load stages. Also, the sets of analyses that produced worse results at latter load stages are highlighted as outliers. Finally, in Figure 5-39 (a), the signed errors are shown just to provide a less crowded presentation.

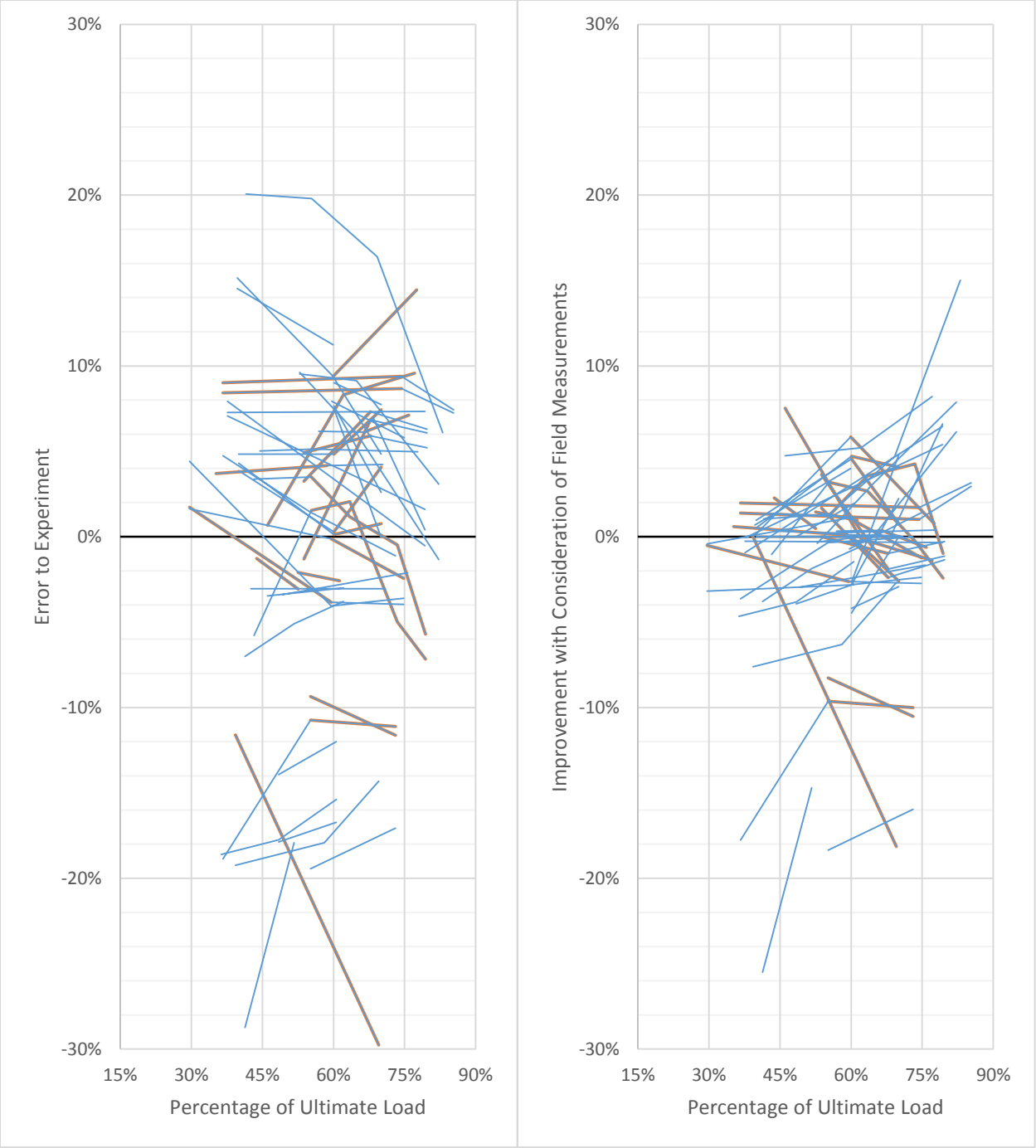


Figure 5-39 (a) Experimental error (b) Percentage improvement with consideration of field measurements vs percentage of ultimate load measurements taken. Sets of analyses with worse results at later load stages highlighted

5.3.2 Number of Measurements Used

The results were also aggregated based on the number of measurements used. Firstly, Table 5-8 shows that as the number of measurements increased, trial matching became harder to achieve. Logically, this was expected, however not to the degree where no matches were obtained when

using five measurements. On the other hand, Figure 5-40 show the results in histograms. For analyses with only two measurements, a large percentage of results produced very little positive or negative improvement. This suggests that using two measurements made it more difficult to alter the overall failure load distribution due to over-matching. Although both extremes were expected to some degree, the methodology produced results in the form of unaltered failure load distributions or no distributions, neither of which were very helpful. In terms of accuracy to the experiment, no significant change was seen contrary to the expectation that more measurements would led to more accurate predictions. However, it could have been due to the small sample size of four and five measurement analyses; especially since those analyses had more difficulty matching.

Table 5-8 Break-down of Sherlock performance by number of measurements used

Number of Measurements	Number of All Analyses	Number of Matched Analyses	Percentage Matched	Error to Experiment	Improvement with FM
5	6	0	-	-	-
4	18	8	44.4%	9.8%	-4.2%
3	69	58	84.1%	7.6%	-1.1%
2	83	75	94.1%	7.4%	-0.5%
	176	141	80.1%	7.6%	-1.0%

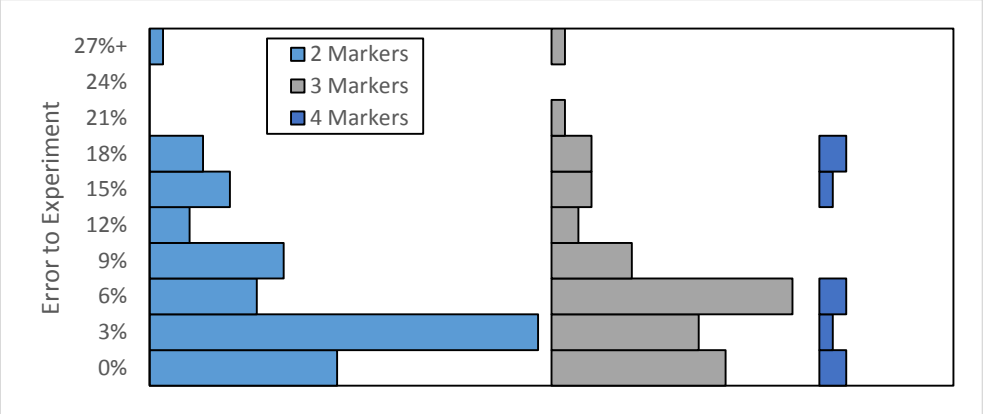


Figure 5-40 Histograms of absolute error for anallyses with different number of measurements

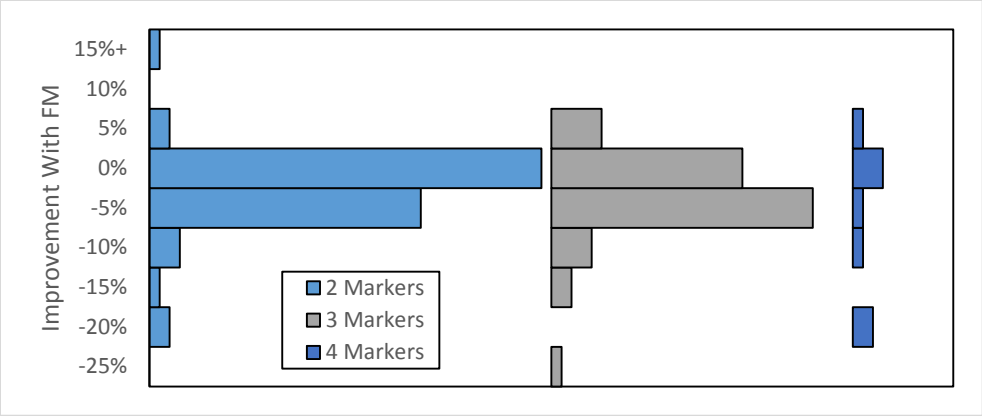


Figure 5-41 Histogram of improvement considering field measurements for analyses with different number of measurements

5.3.3 Types of Measurements

The break-down of performance based on the use of types of measurements was examined. In this study, the displacement measurement was used in all analyses as it was considered in some ways the dependent variable. It should also be pointed out that the structural analyses conducted by VecTor2 were displacement controlled. However, this does not mean displacement measurements are always needed. Table 5-9 shows the breakdown of performance in terms of the other measurements used: reaction loads, crack widths, and reinforcement strains. The number of analyses with crack widths and without crack widths in Table 5-9 do not sum up to 176 analyses because crack widths measurements were not available for all specimens. Those analyses were not counted as no crack widths consideration to maintain the same set of specimens for comparison. The same applies to reinforcement strains. Overall, the aggregate average performances do not reveal conclusive trends and some results can be misleading. A more in-depths analysis of each measurement type is provided in the following section.

Table 5-9 Break-down of Sherlock performance by types of measurement used

Number of Measurements	Number of All Analyses	Number of Matched Analyses	Percentage Matched	Error to Experiment	Improvement with FM
With Load	139	108	77.7%	7.8%	-1.0%
No Load	37	33	89.2%	6.8%	-1.1%
With Crack	59	41	69.5%	9.2%	-2.6%
No Crack	63	59	93.7%	8.0%	-1.5%
With Reinf. Strain	69	44	63.8%	7.2%	-2.1%
No Reinf. Strain	46	41	89.1%	7.9%	-1.3%
	176	141	80.1%	7.6%	-1.0%

Reaction Loads

Since often the primary goal is to better predict the ultimate failure load, the reaction load measurement was expected to be the most effective measurement to consider. However, in cases where the experimental and the predicted behaviour differed substantially, the load measurement was actually a negative influence. This was particularly true in shear-critical beams without shear reinforcement. Table 5-9 suggests that analyses without load measurements not only produced slightly more accurate results, but also had a higher match percentage. This is an encouraging result because reaction load measurements would be the most difficult to obtain outside of laboratory experiments. However, analyses including reaction loads were conducted for all specimens at all load stages while analyses not considering reaction loads were conducted for most specimens but not all and usually at latter load stages. Therefore, it's possible the set of analyses without load consideration were better predicted overall. This could explain how not including reaction loads produced less error, but slightly worse improvement compared to stochastic analysis. Nonetheless, the performance of analyses without considering reaction loads in predicting the failure load was still promising.

Crack Widths

There was a large drop-off in match percentage when crack widths were considered. In most cases, the crack widths were over-estimated. It is hypothesized that the reason is the same as that discussed for Specimen VSOA3. Certain elements experience concentrated strain development and in turn unrealistically large crack predictions. To combat this, a new methodology should be considered in the extraction of crack widths instead of taking the maximum encountered.

The relatively poorer performance compared to analyses without crack widths consideration is misleading. A detailed summary further broken down by specimen is provided in Table 5-10. None of the shear wall specimens provided detailed crack width data and were not included in the comparison. Table 5-10 shows that for several specimens, no match could be obtained considering crack widths. For these specimens, analyses without crack widths obtained relatively lower errors as well as little improvement. This suggests that had the analyses considering crack widths resulted in matches, the errors would have been small as well. But since the analyses were not able to obtain matches, their relatively lower errors were not included in the average, thus raising the average

error for analyses considering crack widths. If only specimens that produced matches with and without crack width measurements were considered, the average errors and improvements in Table 5-10 show that the consideration of crack widths is beneficial overall. This was more seen in shear-critical specimens where the crack width progression was a good predictor of diagonal tension failure.

Table 5-10 Breakdown of results by specimen and by the usage of crack width measurements

Specimen	With Crack Widths			Without Crack Widths		
	Matched Analyses	Error	Improvement	Matched Analyses	Error	Improvement
VSOA1	5	15.5%	-1.6%	2	17.3%	-3.4%
VSOA3	2	23.3%	-20.1%	3	5.3%	-2.1%
VSA1	3	3.3%	0.2%	2	2.3%	+1.2%
VSA3	3	6.5%	0.0%	3	6.5%	+0.0%
VSB2	0	-	-	3	5.0%	+0.3%
VSB3	0	-	-	3	4.0%	+0.3%
VSC1	0	-	-	3	3.0%	+0.0%
VSC2	0	-	-	4	15.6%	+5.5%
PLS4000	8	8.4%	-7.3%	4	15.7%	-14.6%
BN50	2	19.0%	-7.3%	6	18.7%	-7.1%
BN100	0	-	-	9	5.2%	-0.3%
BN100D	10	5.2%	-0.4%	7	3.8%	+1.0%
BH50	3	14.1%	1.2%	3	13.0%	+2.2%
BH100D	5	5.0%	3.2%	1	4.7%	+3.5%
BM100	0	-	-	9	2.9%	-1.6%
Adjusted Avg.	41	9.2%	-2.6%	31	10.3%	-3.0%

Reinforcement Strains

The rebar strain data also had trouble obtaining reliable matches according to Table 5-9. The nature of strain gauge readings can be unstable due damage during the casting or loading processes. At other times, slip between the rebar and the concrete could result in less strain than the conditions modelled. Finally, strain gauges are typically zeroed after the application of self-weight. For the shear walls tested by Lefas et al. (1990), it appeared the shear walls were zeroed after the application of axial load as well. This sometimes led to offsets in strain measurements since the strains are not zeroed in the analytical models. This was seen in large specimens such as Specimen PLS4000 and in shear walls with large axial loads such as Specimen SW23.

In terms of the performance, a detailed breakdown by specimen is presented in Table 5-11. The average had to be adjusted once again since Specimen BH50 did not obtain any matches with reinforcement strains considered. Overall, there does not appear to be conclusive evidence to suggest whether consideration of reinforcement strains was beneficial or not.

Table 5-11 Breakdown of results by specimen and by the usage of rebar strains

Specimen	With Rebar Strains			Without Rebar Strains		
	Matched Analyses	Error	Improvement	Matched Analyses	Error	Improvement
SW22	7	4.0%	3.0%	2	7.3%	-0.3%
SW23	1	0.8%	-0.3%	2	2.8%	-2.3%
B4	3	7.0%	0.2%	3	8.1%	-0.9%
PLS4000	6	11.0%	-9.9%	6	10.7%	-9.6%
BN50	4	20.1%	-8.5%	4	17.5%	-5.9%
BN100	6	5.0%	-0.1%	3	5.6%	-0.7%
BN100D	8	6.1%	-1.3%	9	3.4%	1.4%
BH50	0	-	-	6	13.5%	1.7%
BH100D	3	6.2%	2.0%	3	3.7%	4.5%
BM100	6	3.6%	-2.4%	3	1.4%	-0.2%
Adjusted Avg.	44	7.2%	-2.1%	35	6.9%	-1.9%

5.3.4 Measurement Sensitivity

As expected, the measurement sensitivities had a large impact on the analysis results. In some cases, the sensitivities were too large and, in some cases, too small. Two main trends were observed. First, analyses with more measurements required larger sensitivities to obtain matches and analyses with less measurements would have benefitted more from smaller sensitivities. Second, taking sensitivities as percentages of the measurement values resulted in a linear increase in sensitivity values. However, although the variation in predicted responses increased at progressive load stages, the rate of change did not increase linearly. Therefore, in many cases, the measurement sensitivities at later stages became too large in comparison.

However, regardless of the results, it should be reminded that the sensitivities represent how reliable the field measurements obtained are. Therefore, results that obtain no matches because the sensitivities are too small should not be regarded as null results. Instead, they should be interpreted as possible errors in the analytical model of the experiment. On the other hand, results that considered all trials as likely are not null results either. Instead they should be seen as complete

agreement between the analytical model and early stage field measurements. In this study, the most mismatched specimens were shear wall specimens. The specimens that had the most analyses where most trials were considered likely, were the non-shear-critical beams by Shim et al. On one hand, the shear walls were difficult to model because of the uncertain base conditions used in the experiment, on the other hand, the near flexural dominant beams were very simple to model. Nevertheless, those conclusions could be difficult to infer when the sensitivities used were only assumptions. Detailed investigations into more specific sensors used would provide more conclusive answers.

Finally, perhaps the phenomenon of overmatching can be avoided altogether. The current methodology considers a discrete cut-off between trials with high confidence, matched trials, and trials with low confidence. This gives the same weight, in terms of calculating average failure loads, to trials that matched the field measurements perfectly and trials that were barely matched. For overmatched analyses, this results in very little change to the predicted average failure load. Instead, a weighted average could be used based on the actual match of each trial. Figure 5-42 show examples where weighted averages could have provided different results rather than the same failure load distribution as the stochastic analyses without considering field measurements.

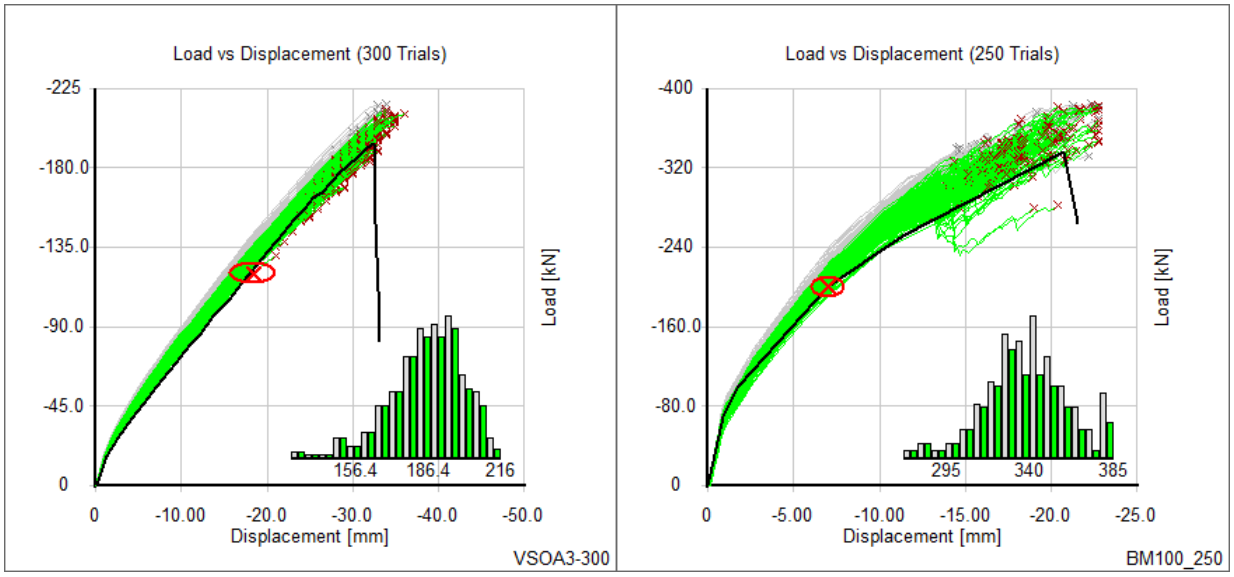


Figure 5-42 Candidate analyses for the consideration of weighted average failure loads

Chapter 6 Conclusions and Future Work

This chapter summarizes the findings of this study and the opportunities for future research. The objectives of this study were to investigate the effectiveness of stochastic analysis for the performance assessment of reinforced concrete structures, and to develop and test a methodology for incorporating field measurements to better predict ultimate behaviour.

6.1 Verification of Stochastic Analysis

6.1.1 Summary and Conclusions

1. For the specimens examined in this study, the average failure loads calculated through stochastic analysis with specified material properties were more accurate (7.1% average absolute error) than deterministic analysis with test day material properties (9.5% average absolute error) and especially more accurate than deterministic analysis with specified material properties (16.6% average absolute error). This showed that the use of stochastic analysis is a viable alternative to more commonly used deterministic methods, with potential improvements in accuracy where errors of up to 10% are typically considered adequate in analysis of reinforced concrete structures.
2. Good accuracy was obtained because of the accurate material property models. However, better results were obtained with Canadian mixed concrete, more contemporary mixed designs, and normal strength concrete.
3. Specimens with different failure modes produced different levels of variability in the ultimate load distributions. Diagonal tension failures produced the largest coefficient of variation at 7.9%, while compression failure due to shear and flexural-dominant responses varied at 4.6%. The one specimen that failed due to rebar rupture also had a high variability of 8.9%.
4. The use of 250 to 300 trials per stochastic analysis was found to be adequate. However, it is estimated that a typical analysis would require 15hrs for a typical processor (Intel® Core™ i7-3770 CPU @ 3.40 GHz with 8.00 GB of RAM).
5. Stochastic analyses did not typically return much variation in the predicted failure mode, but were able to capture unique responses due to combinations of material properties. The

introduction of more variation into the analysis, for example with the added consideration of corrosion, should result in different failure modes.

6. Stochastic results can be used to conduct sensitivity analysis on the material properties and failure loads.
7. The developed post-processor developed, Sherlock, was able to facilitate the extraction of stochastic results.

6.1.2 Future Work

1. More variation should be implemented into stochastic analyses from environmental factors such as corrosion. The work by Habibi et al. (2017) implemented stochastic analysis considering corrosion; although not considered in this study, it should introduce much more variability.
2. Shrinkage can affect the overall response of certain structures and is often not considered due to its uncertain behaviour. In VecTor2, shrinkage strains can be applied as elemental loads and could be varied from trial to trial.
3. Support and boundary conditions can also largely influence analytical results and cannot always be modelled exactly. This would be especially true for structural components in the field. Perhaps variable boundary elements, to represent joints or slabs, can be modelled in VecTor2 with variable values of stiffness for each trial.
4. In this study, only the default Monte Carlo simulation was examined. According to Hunter (2016), the use of Latin Hypercube Sampling can reduce the number of trials required in comparison. This could save on computational time, especially if more variations are employed.
5. More material property models should be compiled and catalogued in terms of the date and location of concrete mixes.

6.2 Incorporation of Field Measurements

6.2.1 Summary and Conclusions

1. Stochastic analyses incorporating field measurements gave overall results that showed little improvement (-0.99%) in prediction capabilities compared to stochastic analysis without

field measurements. This could be attributed to the selection of specimens examined, with the majority being shear-critical structures.

2. More accurate failure load predictions were obtained with measurements taken closer to failure.
3. Analyses with more field measurements (four or more) typically encountered difficulty in identifying likely trials. Analyses with few field measurements (three or less) resulted in the identification of most trials as likely resulting in small changes to the failure load distribution obtained from stochastic analysis without considering field measurements.
4. No distinct conclusions could be made on the effectiveness of each type of measurement examined (reaction loads, crack widths, reinforcement strains).
5. Analyses without considering reaction load measurements still produced good estimates of failure. This was promising as reaction loads can be difficult to obtain outside of laboratory experiments.
6. Analyses considering crack width measurements had more difficulty identifying likely trials. Crack widths were often over-estimated possibly due to elements in the analytical model that obtained artificially high crack width predictions.
7. Analyses considering reinforcement strain measurements also had difficulty identifying likely trials possibly due to different assumptions on when strain gauges are zeroed and damage and slip effects during loading. Also, large variations between average and local strains in reinforcement, as influenced by proximity to a crack, contributed to the difficulty.
8. The default measurement sensitivities used were often too large for later load stages. More representative/accurate sensor reliabilities should be used.

6.2.2 Future Work

1. A more in-depth and larger sample of specimens should be examined featuring more equal representation of different failure modes.
2. A study into different sensors should be undertaken to obtain more accurate field measurement sensitivities.
3. A weighted average should be used to calculate the modified failure load distribution with respect to how close each trial is to the field measurements.

4. The extraction of crack widths from the analytical results should be modified to avoid concentrated crack widths that over-estimate experimental values.
5. The post-processor, Sherlock, should be better automated to consider different combinations of measurements at different load stages and produce resulting distributions of failure loads for all combinations. This will save time for the analyst to consider each specimen and increase the set of specimens examined for a future study.

References

Bartlett, F. M. (2007). Canadian Standards Association standard A23.3-04 resistance factor for concrete in compression. Canadian Journal of Civil Engineering Can. J. Civ. Eng., 34(9), 1029-1037.

Bartlett, F. M., & MacGregor, J. G. (1996). Statistical analysis of the compressive strength of concrete in Structures. ACI Materials Journal MJ, 93(2).

Beven, K., & Andrew, B., (1992). The future of distributed models: model calibration and uncertainty prediction. Hydrological Processes, 6(3), 279-298.

Blakeborough, A., Clement, D., Williams, M. S., Woodward, N. (2001). Novel load cell for measuring axial force, shear force, and ending movement in large-scale structural experiments. Experimental Mechanics, 42(1), 155-122.

Bresler, B., and Scordelis, A. C. (1963). Shear strength of reinforced concrete beams. Journal of American Concrete Institute, 60(1), 51–72.

British Standards Institution (1983), BS 1881-120: Testing concrete. Method for determination of the compressive strength of concrete cores.

Cao, S. (2001). Size effect and the influence of longitudinal reinforcement on the shear response of large reinforced concrete members. (Master's Thesis). University of Toronto.

Choi, B., Scanlon, A., & Johnson, P. A. (2004). Monte Carlo simulation of immediate and time dependent deflections of reinforced concrete beams and slabs. SJ ACI Structural Journal, 101(5), 633-641.

Collins, M. P., & Mitchell, D. (1997). Prestressed concrete structures. Toronto: Response Publications.

Collins, M. P., Bentz, E. C., Quach, P. T., & Proestos, G. T. (2015). The challenge of predicting the shear strength of very thick slabs. Concrete International, 37(11), 29-37.

Habibi, S. (2017). Finite element modelling of corrosion damaged reinforced concrete structures. (Master's Thesis). University of Toronto.

Hunter, M. D. (2016). Towards stochastic finite element analysis of reinforced concrete structures. (Master's Thesis). University of Toronto.

Lefas, I. D., (1998) Behaviour of reinforced concrete walls and its implication for ultimate limit state design. (PhD Thesis). Imperial College, University of London.

Lefas, I.D., Kotsovos, M. D., Ambraseys, N. N. (1990) Behaviour of reinforced concrete structural walls: strength, deformation characteristics, and failure mechanism. ACI Structural Journal, 87(1), 23-31.

Markovsky, I., & Van Huffel, S. (2007). Overview of total least squares methods, Signal Processing, 87(2007), 2283-2302.

Mirza, S. A., & MacGregor, J. G. (1979a). Variability of mechanical properties of reinforcing bars. Journal of the Structural Division, 105, 921-937.

Mirza, S. A., & MacGregor, J. G. (1979b). Variations in dimensions of reinforced concrete members. Journal of the Structural Division, 105, 751-766.

Mirza, S. A., & MacGregor, J. G. (1982). Probabilistic study of strength of reinforced concrete members. Canadian Journal of Civil Engineering Can. J. Civ. Eng., 9(3), 431-448.

Mirza, S. A. (1998). Monte Carlo simulation of dispersions in composite steel-concrete column strength interaction. Engineering Structures, 20(1-2), 97-104.

Mishra, S. K., Panda, G., Das, D. P., Pattanaik, S. K., Meher, M. R. (2005). A novel method of designing LVDT using artificial neural network. Proceedings of Intelligent Sensing and Information Processing 2005, 223-227.

Nash, J. E., & Sutcliffe, J.V. (1970). River flow forecasting through conceptual models part I – a discussion of principles. Journal of Hydrology, 10(3), 282-290.

Ning, C., & Li, B. (2016). Probabilistic development of shear strength model for reinforced concrete squat walls. The Journal of the International Association for Earthquake Engineering, 46, 877-897.

Nowak, A. S., & Szerszen, M. M. (2003). Calibration of design code for buildings (ACI 318): Part 1—Statistical models for resistance. SJ ACI Structural Journal, 100(3), 378-382.

Oesterle, R. G. Fiorato, A. E., Johal, L. S., Carpenter, J. E., Russell, H. G., Corley, W. G. (1976). Earthquake Resistant Structural Walls-Tests of Isolated Walls. Report to National Science Foundation, Construction Technology Laboratories, Portland Cement Association.

Professional Engineers of Ontario, (2012). Professional practice bulletin regulatory No. 2, Association of Professional Engineers of Ontario, Department of Policy and Professional Affairs.

Quach, P. (2016). Understanding and safely predicting the shear response of large-scale reinforced concrete structures (Master's thesis). University of Toronto.

Ramsay, R. J., Mirza, S. A., & MacGregor, J. G. (1979). Monte Carlo study of short time deflections of reinforced concrete beams. ACI Journal Proceedings JP, 76(8), 897-918.

Rakoczy, A., & Nowak, A. (2013). Resistance model of lightweight concrete members. ACI Materials Journal, 110(1), 99-109.

Sherwood, E. G., Collins, M. P., & Bentz, E. C., (2008). Where is shear reinforcement required? Review of research results and design procedures. SJ ACI Structural Journal, 105(5), 590-600.

Shim, W., (2002). Analysis of beams with low shear reinforcement. (Master's thesis). University of Toronto

Shim, W., & Vecchio, F. J., (2004). Experimental and analytical reexamination of classic concrete beam tests. ASCE Journal of Structural Engineering, 130(3), 460-469.

Stanik, B. A., (1998). The influence of concrete strength, distribution of longitudinal reinforcement, amount of transverse reinforcement and member size on shear strength of reinforced concrete members. (Master's thesis). University of Toronto.

Unanwa, C., & Mahan, M. (2014). Statistical analysis of concrete compressive strengths for california highway bridges. J. Perform. Constr. Facil. Journal of Performance of Constructed Facilities, 28(1), 157-167.

Vecchio, F. J. and Collins, M. P., (1986). The modified compression field theory for reinforced concrete elements subject to shear. ACI Journal, 83(2), 219-231.

Vecchio, F. J. (2000). Distributed Stress Field Model for Reinforced Concrete Formulation. ASCE Journal of Structural Engineering, 126(9), 1071-1077.

Vincent, T., Ozbakkaloglu, T., Seracino, R., Kaggwa, W. (2011). Influence of variations in concrete material properties on the serviceability of reinforced and prestressed concrete flexural members. Engineering Structures, 33(1), 99-106.

Wiśniewski, D. F., Cruz, P. J., Henriques, A. A., & Simões, R. A. (2012). Probabilistic models for mechanical properties of concrete, reinforcing steel and pre-stressing steel. Structure and Infrastructure Engineering, 8(2), 111-123.

Wong, P., Trommels, H., & Vecchio, F. J. (2013). VecTor2 and FormWorks user's manual

Yoshida, Y. (2000). Shear reinforcement for large lightly reinforced concrete members. (Master's thesis). University of Toronto.