Strategies for Separation of Aleatory and Epistemic Uncertainties

Smitha D. Koduru
Research Engineer, C-FER Technologies, Edmonton, Canada

ABSTRACT: Practical applications for separation of aleatory and epistemic uncertainties are demonstrated with two examples applied to steel structures and pipelines. These examples show the need for separation of uncertainties, the application of existing strategies and their limitations. The first example pertains to the behaviour of a simple steel connection subjected to combined axial, shear and moment loading. Uncertainties in the material properties of the connection components are shown to have a significant influence on the model uncertainty corresponding to the connection moment capacity. The second example shows that the uncertainties in the geometric imperfections of a steel pipeline contribute to the total model uncertainty in the strain capacity of the pipe. More such case studies demonstrating the utility of separation of epistemic uncertainty are expected to lead to increased efficiency in future model development.

1. INTRODUCTION

Characterizing model uncertainty forms an important component of model-based (as opposed to historical data-based) risk analysis, as it has a significant influence on ranking the outcomes for risk management. When empirical or semi-empirical mathematical models are developed from experimental data, the model uncertainty is often assumed to be random and is represented by probabilistic distribution of the residuals of the model fitting. During validation, the discrepancy between the experimental and the model response is attributed entirely to the model error, and is considered as "total model uncertainty". However, experimental measurements themselves are subjected to variability due to inherent randomness in the material properties and geometric parameters of the experimental specimens, and measurement errors of the testing equipment.

The variation in the experimental results could be attributed to "aleatory uncertainty", i.e. uncertainty due to inherent randomness, as this uncertainty is irreducible despite repeated testing of nominally identical specimens. Therefore, a significant challenge in development of reliable mathematical models, and application of model-based risk analysis, is to identify and separate the aleatory uncertainty from the model uncertainty, also known as "epistemic uncertainty". As epistemic uncertainty is reducible with improvements to model form and parameters, it is desirable to identify the fraction of epistemic uncertainty in the total model uncertainty. Furthermore, when the uncertainty in the model material and geometric parameters is explicitly modeled in a risk analysis, employing the total model uncertainty would account for the uncertainties in model material and geometric parameters twice.

Assuming statistical independence and normal distribution for all the sources of uncertainty, the total variability in the response would be equal to sum of the variances from all sources of uncertainty in a linear model. Similarly, using the rule of total variance, the sum of the explained variance due to the parameter influence, and the unexplained variance – computed as the variance of the conditional expectation of the response – would be equal to the total variability in the response. Therefore, in theory, it is possible to isolate the epistemic model error once the variability in all other sources of uncertainty is known. However, these methods may not be practicable in reality due to the constraints on the type of data (e.g., direct measurements, population-wide statistical

measures, measurements with unquantified bias, and so on) available, and the sample sizes of available data.

In this paper, two case studies are presented where the estimation of epistemic model uncertainty is performed using the commonly available methods. The first case study pertains to the model for predicting the capacity of simple steel connections. The second case study pertains to the model for estimating compressive strain capacity of steel pipelines. For these two cases, the constraints on the available data and the difficulties in application of simple methods for separation of uncertainties are presented. Following this, strategies to overcome the limitations are also demonstrated in these applications.

2. CASE STUDY 1: CAPACITY OF SIMPLE STEEL CONNECTIONS

Shear-tab connections are one of the most commonly used beam to column connections to transfer shear load in steel frame buildings. These connections are designed assuming zero moment capacity and their behaviour under combined axial, shear, and moment loads is not well-known. Figure 1 shows the parts of a typical shear-tab connection. The steel plate, called "shear-tab", is welded to the column face and bolted to the beam web forming the load path for transfer of vertical loads from beam to column.

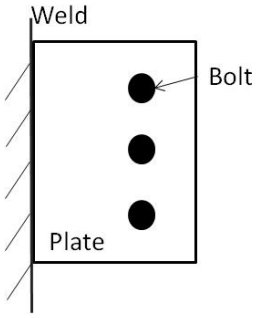

Figure 1: Parts of a typical shear-tab connection

A component-based model for predicting the capacity of single shear-tab connections subjected to a combined axial, shear, and bending loads has been recently proposed (Koduru & Driver 2014). The model is a combination of parallel and series system of several uni-axial springs representing weld fracture, bolt failure, plate yield, plate fracture and plate tear-out, where each spring model has the general form for maximum capacity as,

$$F_{spring} = C_i \cdot m_i \cdot g_i \qquad (1)$$

Where F_{spring} is the spring force, C_i represents empirical model coefficients, m_i represents material properties, and g_i represents geometric properties.

Tension, shear, and moment capacity predictions from the model were validated against the experimental results from a variety of loading regimes. Table 1 shows the experimental data from full scale tests conducted by Thompson (2009) compared to the model predictions for tension and moment capacity of the shear tab connections.

Table 1: Comparison of experimental data with model predictions (Koduru & Driver 2014)

Moment(kN.m)		*Tension(kN)*	
Measured	*Predicted*	*Measured*	*Predicted*
17.78	*20.15*	*217.73*	*248.02*
16.35	*20.15*	*217.73*	*248.02*
19.10	*20.15*	*188.11*	*248.02*
20.64	*20.15*	*188.11*	*248.02*
36.72	*36.98*	*289.12*	*302.02*
38.41	*36.98*	*293.57*	*302.02*
36.72	*36.98*	*185.21*	*302.02*
42.37	*36.98*	*185.21*	*302.02*
31.07	*36.98*	*160.13*	*302.02*
33.89	*36.98*	*151.23*	*302.02*
64.20	*68.30*	*226.85*	*231.74*
70.61	*68.30*	*214.26*	*231.74*
61.17	*68.30*	*253.54*	*231.74*
70.00	*68.30*	*240.77*	*231.74*

The full scale tests include four nominally identical specimens each for a three-bolt and a five-bolt connection, and six nominally identical specimens of a four-bolt connection. The model predictions were computed from the nominal values of the specimen dimensions and strengths.

Therefore, these predictions do not represent the inherent randomness of the material and geometric parameters.

The total model uncertainty is computed as the ratio of measured to predicted values. The mean and coefficient of variation (CoV) of model uncertainty for tension are 0.82 and 0.24, while the mean and CoV for moment capacity are 0.96 and 0.093, respectively. As noted before, the model bias and variance for the predictions represented in Table 1 do not include the inherent randomness due to the input parameters. Therefore, in order to evaluate the epistemic uncertainty in the total model uncertainty, the variability due to input parameters must be separated.

Table 2 shows a partial list of the required input material and geometry parameters of the model and their variability expressed as CoV.

Table 2: Variance of input material and geometry parameters

Parameter	*CoV*
Weld tensile strength[a], X_u	*0.082*
Plate yield strength[b], f_{yp}	*0.054*
Plate tensile strength[b], f_{up}	*0.034*
Beam web yield strength[b], f_{yw}	*0.054*
Beam web tensile strength[b], f_{uw}	*0.046*
Bolt tensile strength[c], f_{ub}	*0.045*
Weld thickness[a], W_t	*0.154*
Plate thickness[b], t_p	*0.025*
Beam web thickness[b], t_w	*0.025*
Bolt cross-section area[d], A_b	*0.006*

[a]Kanvinde et al. (2008)
[b]Schmidt & Barlett (2002)
[c]Kulak (2005)
[d]Moore et al. (2008)

All the parameters in Table 2 have a lognormal distribution. As the coupon tests and other specimen measurements are not available for all the required parameters, the available CoVs – shown in Table 2 – are collected from the literature. It is noted here that it is often rare to find the details of all the specimen measurements for experimental results, and as such population-based variance measures available in the literature are often used in practical applications. Assuming that the input parameters are uncorrelated and the model uncertainty is multiplicative, the variance of total model uncertainty is computed as,

$$\sigma_t^2 = \sum_{i=1}^{n} \sigma_i^2 + \sigma_e^2 \qquad (2)$$

where σ_i^2 is variance of the i^{th} input parameter computed as $\sigma_i^2 = \log(1+\text{CoV}_i^2)$, n is the total number of input parameters influential on the model response, and σ_e^2 is variance of the epistemic model uncertainty.

As the total capacity of the connection is governed by the weakest component of the connection, the variance in the capacities are driven by the variance in the parameters associated with the dominant mode of failure under a given loading regime. The sensitivity analyses in Koduru & Driver (2014) have also shown that the capacity of shear-tab connections is significantly dependent on the mode of failure. For the experimental results considered in Table 1, the modes of failure observed are, bolt shear failures, edge tear-out of the plate, fracture of the plate, and a combination of bolt failure with plate failure modes. Therefore, Eq. (2) is applied to compute epistemic model uncertainty separately for each failure mode and a combination of bolt and plate failure modes. In the component-based model, the spring capacity to resist bolt shear failure is modelled as,

$$F_{bolt} = C_b \cdot f_{ub} \cdot A_b \qquad (3)$$

where C_b, is the empirical model coefficient to convert bolt tensile strength to shear strength. The spring capacity to resist plate tear-out failure is modelled as,

$$F_{tear-out} = C_e \cdot f_{up} \cdot t_p \cdot L_e \qquad (4)$$

where C_e, is the empirical model coefficient and L_e is the distance from center of the bolt to the free edge of the plate. Finally, the spring capacity to resist plate fracture is modelled as,

$$F_{fracture} = f_{up} \cdot t_p \cdot L_{nt} \qquad (5)$$

where L_{nt}, is the net section length in tension. For combined failure modes, the total capacity will be a summation of capacities for each failure mode. The parameters influencing each failure mode and the computed epistemic model uncertainty for moment and tension capacities are listed in Table 3.

Table 3: Epistemic model uncertainty

Failure mode	*Parameters*	*Epistemic CoV*	
		Moment	*Tension*
Bolt failure	f_{ub}, A_b	*0.082*	*0.231*
Plate failure	f_{up}, t_p	*0.083*	*0.232*
Bolt +Plate failure	f_{ub}, A_b, f_{up}, t_p	*0.078*	*0.230*

Table 3 shows that the epistemic model uncertainty for moment capacity is dependent on the observed failure modes. For single failure modes, the reduction in total model uncertainty (expressed as CoV) due to the separation of aleatory uncertainty is between 11-13%. In contrast, for the combined failure mode, aleatory uncertainty due to material and geometric parameter variability accounts for 17% of the total model uncertainty. For the tension capacity, the material and geometric parameter variance explains only 2-3% of the total model uncertainty. Therefore, the uncertainty in the capacity model coefficients for bolt shear, and plate tear-out failure modes are included to compute the remaining epistemic model uncertainty.

The CoV of model coefficient for bolt capacity, C_b, is taken as 0.048 (Kulak 2005), and the CoV of model coefficient for plate tear-out, C_e is taken as 0.083 (Koduru & Driver 2012). Table 4 shows the portion of epistemic model uncertainty for moment and tension capacity with the inclusion of uncertainties in model coefficients of dominant failure modes. For moment capacity, 90% of the total model uncertainty is accounted by the known uncertainties in the material, geometry, and model coefficients for plate failure mode. Under the combined failure mode, 94% of the uncertainty is explained by the known uncertainties. Therefore, selection of variance to represent model uncertainty for moment capacity must be connection-specific and dependent on the expected failure mode. For tension capacity, although the epistemic uncertainty reduced by 4-10% with the inclusion of uncertainties in model coefficients, there is potential for further reduction. In future studies, consideration of uncertainties in connection-specific geometric parameters (e.g., spacing between the bolts, the distance between bolt and the free edge of the plate, and bolt-hole diameter) have the potential to reduce total model uncertainty.

Table 4: Epistemic model uncertainty for tension considering variance of model coefficients

Failure mode	*Parameters*	*Epistemic CoV*	
		Moment	*Tension*
Bolt failure	f_{ub}, A_b, C_b	*0.066*	*0.226*
Plate failure	f_{up}, t_p, C_e	*0.010*	*0.216*
Bolt +Plate failure	$f_{ub}, A_b, f_{up}, t_p, C_b, C_e$	*0.005*	*0.213*

3. CASE STUDY 2: COMPRESSIVE STRAIN CAPACITY OF STEEL PIPELINES

Buried onshore pipelines are often subjected to compression due to ground movements caused by slope movements, frost heave, thaw settlement and seismic loads. Excessive compression on the pipeline either due to bending or due to axial loads causes local buckling, which is a threat to pipeline integrity. Therefore, accurate assessment of compressive strain capacity of buried pipelines forms a significant part of strain-based design of buried pipelines.

Previous studies by Yoosef-Ghodsi et al. (2014) and Zhang et al. (2014) have shown Dorey model (Dorey et al. 2001) as one of the recommended models for predicting compressive strain capacity of pipelines. Dorey model has four equations for predicting the strain capacity

of pipe body and girth-weld with or without Luder's plateau in the stress strain relationship of the pipe steel. The equations for compressive strain capacity with Luder's plateau are,

$$\varepsilon_p = f_p \cdot (1 - a \cdot \frac{\sigma_h}{f_y})^{-1} \cdot \left(\frac{E}{f_y}\right)^{0.8} \cdot \left[b - \frac{h_g}{t}^{0.15}\right] \quad (6)$$

$$\varepsilon_w = f_w \cdot (1 - c \cdot \frac{\sigma_h}{f_y})^{-1} \cdot \left(\frac{E}{f_y}\right)^{0.7} \cdot \left[d - \frac{h_m}{t}^{0.09}\right] \quad (7)$$

where ε_p and ε_w are the strain capacities for plain pipe and girth weld, respectively, a, b, c and d are model constants, f_p and f_w are constants for a given pipe wall thickness and diameter, σ_h is the hoop stress due to internal pressure, f_y is the hoop yield strength, E is the elastic modulus, t is the pipe wall thickness, h_g is pipe-body geometric imperfection, and h_m is girth-weld high-low misalignment.

This model has the least number of required input parameters and has the lowest standard deviation for the total model uncertainty compared to the available models for compressive strain capacity (Zhang et al. 2014). From the experimental data for 61 full scale tests summarized in Zhang et al. (2014), the mean and CoV for total model uncertainty are 1.01 and 0.27, respectively. However, these values are computed assuming median values for two critical input parameters of the model, namely, pipe-body geometric imperfections, h_g, and girth-weld high-low misalignment, h_m.

Geometric imperfection in the pipe body is caused by surface undulations on the outer surface of the pipe and is expressed as the height between peak and valley of the undulation, measured perpendicular to pipe longitudinal axis. Figure 2 shows the quantification of h_g.

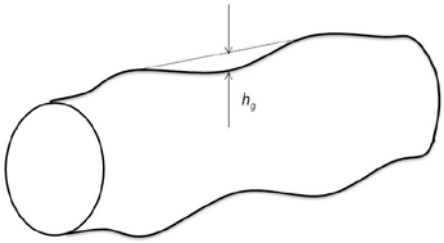

Figure 2: Geometric imperfections of pipe body (after Zhang et al. 2014)

Girth-weld high-low misalignment is caused due to misalignment in pipe sections at the cross-section weld joints. This misalignment reduces the compressive strain capacity near the girth weld of the pipe. Table 5 shows the CoVs for selected material and geometry parameters of Dorey model. The remaining input parameters, which are, pipe wall thickness, pipe diameter, internal pressure, and Young's modulus are not considered due to the lack of uncertainty characterizations.

Table 5: Variance of input material and geometry parameters for Dorey model

Parameter	*CoV*
Hoop yield strength[a], f_y	*0.035*
Pipe body imperfections[b], h_g	*0.548*
Weld misalignment[b], h_m	*0.721*

[a]CSA Z662, Annex O. (CSA 2011)
[b]Derived from data in Zhang et al. (2014)

Assuming lognormal distributions for the parameters in Table 5, CoV for epistemic model uncertainty is estimated separately for pipe body and for girth-weld compressive strain capacities. Based on the model form of Dorey equations, total model uncertainty for pipe body capacity is expressed as,

$$\sigma_t^2 = (0.8^2 + 2)\sigma_{f_y}^2 + 0.15^2\sigma_{h_g}^2 + \sigma_e^2 \quad (8)$$

and the total model uncertainty for girth-weld capacity is expressed as,

$$\sigma_t^2 = (0.7^2 + 2)\sigma_{f_y}^2 + 0.09^2\sigma_{h_m}^2 + \sigma_e^2 \quad (9)$$

Based on Eq.s (8) and (9), the CoV for epistemic model uncertainties for pipe body and girth-weld capacities are assessed as 0.251 and 0.257 respectively. This indicates that epistemic model uncertainty forms a significant part of the total model uncertainty and it is possible to further reduce this uncertainty in future studies with the availability of detailed measurements of input model parameters.

4. DISCUSSION

In this paper, two case studies are presented to demonstrate the strategies for separation of epistemic uncertainty in the total model error. The limitations for separation of epistemic uncertainty are primarily due to:

- Lack of relevant specimen measurements corresponding to experimental results.
- Limited number of experimental samples which prevent accurate characterization of material and geometry parameter uncertainties.
- Difficulty in obtaining precise physical measurements for certain model input parameters, e.g., surface undulations, weld throat thickness, and weld misalignment and so on.
- Discrepancy in definitions of measured model response, e.g., gauge length for averaging compressive strains close to local buckling, tension capacity of steel capacity measured as horizontal force or along the axis of the connecting beam, which prevent the meta-analysis of all available experimental data.

Despite these limitations, it is possible to separate the epistemic model uncertainty by the following strategies:

- Use population-based variance measures for material and geometry parameters.
- For applications where population-based variance is disproportionately large, it is advantageous to use a subset of the data that corresponds to the specimen characteristics. For example, finding the data set with similar steel grade, similar thickness, and so on.
- When model parameters are expected to be correlated and the correlation structure is unknown, assume a perfect correlation and represent the variance of only one of the correlated variables. Otherwise, ignoring the correlation leads to under-estimation of epistemic model uncertainty.
- A sensitivity analysis of the model response to the input parameters provides a guideline for selecting the most important parameters that are likely to have contributed to the total model uncertainty.

5. CONCLUSION

Employing mathematical models for representation of structural system behaviour is common in engineering practice. Often the response from these models is validated against the experimental results to characterize the uncertainties and approximations associated with the mathematical models. However, ignoring the uncertainties due to material and geometry parameters would lead to an overestimation of model uncertainty. In the long run, this leads to inefficient effort to improve the model predictions as the proportion of aleatory uncertainty in the total model uncertainty increases.

The present study shows the strategies to separate aleatory and epistemic uncertainties for two practical applications. Despite the highlighted limitations due to the low sample sizes of the measurement data, and uncertainty characterization of the input parameters, employing the proposed strategies has shown which models would benefit from further improvements and which models should focus on proper characterization of aleatory uncertainty. In future, more such case studies demonstrating the utility of separation of epistemic uncertainty are expected to lead to increased efficiency in engineering model developments.

6. REFERENCES

CSA (2011), *Pipeline Systems Operations, Quality and Integrity Management: Annex O. CSA Z662*, Canadian Standards Association, Missisauga, ON, Canada

Dorey, A.B., Cheng, J.J.R. and Murray, D.W. (2001). *Critical Buckling Strains for Energy Pipelines*. Structural Engineering Report No. 237, Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Albera.

Kanvinde, A., Grondin, G., Gomez, I. and Kwan, Y. K. (2008). *Strength and Ductility of Welded Joints Subjected to Out-of-Plane Bending*, AISC Report No. 3237, American Institute of Steel Construction, Chicago, IL.

Koduru, S.D., and Driver, R.G. (2012). "Uncertainty modelling of shear tab connections," In *Proceedings of 3rd International Structural Speciality Conference*, Canadian Society for Civil Engineering Annual Conference, June 4-6, Edmonton, AB.

Koduru, S.D., and Driver, R.G., (2014), "Generalized component-based model for shear tab connections," ASCE *Journal of Structural Engineering,*140 (2), 04013041

Kulak, G. L. (2005). *High Strength Bolting for Canadian Engineers*, Canadian Institute of Steel Construction, Markham, ON.

Moore, A. M., Rassati, G. A. and Swanson, J. A. (2008). *Evaluation of the Current Resistance Factors for High Strength Bolts*, Research Council on Structural Connections, Chicago, IL.

Schmidt, B. J. and Bartlett, F. M. (2002). "Review of resistance factor for steel: data collection." *Canadian Journal of Civil Engineering*, Canadian Science Publishing, 29: 98-108.

Thompson, S. L. (2009). *Axial, Shear and Moment Interaction of Single Plate "Shear Tab" Connections*, M.S. Thesis, Milwaukee School of Engineering, Milwaukee, WI

Yoosef-Ghodsi, N., Ozkan, I. and Chen, Q. (2014), "Comparison of compressive strain limit equations". In *Proceedings of the 2014 10th International Pipeline Conference, IPC 2014*, September 29 – October 3, 2014, Calgary, Alberta, Canada

Zhang, Z., Yu, Z., Liu, M., Kotian, K. and Zhang, F. (2014), "Application of compressive strain capacity models to multiple grades of pipelines". In *Proceedings of the 2014 10th International Pipeline Conference, IPC 2014*, September 29 – October 3, 2014, Calgary, Alberta, Canada