CONSTITUTIVE PARAMETER CALIBRATION FOR STRUCTURAL STEEL: NON-UNIQUENESS AND LOSS OF ACCURACY

By R.J. Cooke^a and A.M. Kanvinde^b

^aForell Elsesser Engineers, San Francisco, CA

^bDepartment of Civil and Environmental Engineering, University of California, Davis, CA

Phone: +1 530 752 2605, Fax: +1 530 752 7872, Email: kanvinde@ucdavis.edu

ABSTRACT

Finite Element (FE) simulations used to characterize extreme limit states in steel structures require the calibration of numerous parameters. Calibration of these models for large strains (greater than 0.3 or so) cannot be performed using stress-strain curves on standard tests, since the stress state in these tests becomes non-homogenous due to necking or buckling which occur at lower strains. As a result, calibration is often performed by matching load-displacement curves of calibration specimens to those obtained through complementary FE simulations. In these situations, multiple parameter sets produce strain fields that match the measured load-displacement response, resulting in non-unique parameter fits. A series of 2400 FE simulations with 4 specimen geometries, 300 material parameters sets, and 2 loading histories indicates that multiple trial parameter sets produce excellent load-displacement match with the true material response, implying that the method is highly susceptible to non-unique fitting. All simulations use the Armstrong and Frederick constitutive model with a von Mises yield surface. The impact of non-unique fitting is assessed through FE simulations based on parameter sets that show excellent load-displacement match with calibration specimens. It is determined that the non-uniqueness does not significantly affect the prediction of peak force. However, it severely impacts the accuracy in prediction of internal plastic strains, with errors as large as 50% with respect to the true material. This has serious implications for FE simulation used to characterize extreme, strain-based limit states such as fracture. Strategies for mitigation of this inaccuracy are presented, along with limitations of the study.

KEYWORDS: Finite element simulation; Constitutive models; Calibration.

INTRODUCTION

Finite Element (FE) simulation is commonly used to characterize extreme limit states in structural steel components. Sophisticated FE simulations have recently been used to characterize the response of structural steel braces, connections as well as steel plate shear walls (Fell and Kanvinde, [1]; Myers et al. [2]; Purba and Bruneau, [3]). More specifically, the various objectives of these simulations include modeling the load-deformation response of structures and components (e.g., Berman et al. [4]), characterizing internal stress and strain distributions, and predicting failure modes such as fracture, based on these internal distributions (e.g., Chao et al. [5]). There is increasingly heavy reliance on FE models to quantitatively determine extreme limit states, within a FE-simulation framework that seeks to describe structural performance accurately.

Accurate representation of material constitutive response (i.e., the relationship between the true stress and strain) lies at the heart of these simulations. This is particularly important if the limit states of interest are associated with high levels of inelasticity. Structural FE models that simulate these extreme limit states rely on multi-axial cyclic plasticity models, which are able to simulate various aspects of material response including the Bauschinger effect, nonlinear isotropic and kinematic hardening, and ratcheting. For steel (and other polycrystalline metals), these constitutive models typically employ a von Mises yield surface, and the assumption of isochoric plastic flow (wherein volume is conserved during plastic deformation), Lubliner [6]. Some popular models used to simulate the inelastic response of steel include the bounding surface model (Dafalias and Popov, [7]), and the Armstrong–Frederick (AF) model (Armstrong

and Frederick, [8]). To replicate the various phenomena discussed above, these models require the calibration of numerous ($\approx 4-10$) parameters. The accurate calibration of these parameters is a challenging process for two reasons. First, standard uniaxial coupon specimens (ASTM E8, [9]) such as shown in Figure 1a (referred to henceforth as Cylindrical Tensile, or CT specimens) are ideal for calibrating “engineering” material properties such as yield stress to be used in design. However, these standard tests are able to produce a homogenous stress state in the specimen only prior to unstable necking, which initiates at strains on the order of 0.1-0.2 (i.e. 10%-20%) – e.g., see Ling [10]. After this, the localized (necked) specimen response is controlled by interactions between material hardening and the change of geometry. In this regime of loading, the strain and stress states are non-homogeneous, such that higher strains (as well as increased stress triaxiality or constraint) are encountered towards the center of the necked region. Bridgman [11] has examined this type of post-unstable-necking response extensively.

A similar problem is encountered for cyclic loading (which may be used to calibrate parameters associated with cyclic hardening), such that compressive strains on the order of 5% (or even less) produce specimen buckling, thereby limiting the strain amplitudes that may be applied homogeneously to the specimens (Rodriguez et al. [12]). As a result, standard coupon tests are ill-suited for calibrating model parameters where the end application is characterizing material response over larger strain ranges (i.e., strain greater than ≈ 0.2). However, response of real structural components is often controlled by larger strain ranges over local regions (on the order of 0.8-1.0 prior to fracture; refer Rousselier, [13], and on the order of 0.3-0.4 prior to local buckling in tubular braces; refer Takeuchi and Matsui, [14]). The implication is that parameters

calibrated from standard coupon tests cannot be appropriately applied for the simulation of structural components.

To overcome these issues, a commonly used strategy (Brandonisio et al., [15]; Chi et al., [16], Kulkarni, et al., [17]) is to calibrate the material model parameters directly from the *load-displacement curve* of the specimen, rather than the measured stress-strain response. This involves the construction of finite element models complementary to (i.e., which replicate) the test specimen. The material parameters themselves are then calibrated indirectly through trial and error, by matching the load-displacement curve of the simulation to that obtained from the tests. Specimens used for this type of calibration may include CT specimens shown previously in Figure 1 (complemented by FE simulation of the post-unstable-necking response with non-homogenous strains) or circumferentially notched tension (CNT) specimens (Figure 1b) into which a sharp inhomogeneity can be introduced in a controlled manner, or other specimens, loaded in monotonic or cyclic loading.

While attractive at first glance, this approach introduces additional sources of inaccuracy and uncertainty. The main underlying problem is that the load displacement curves are the result of complex interactions between the material response (e.g. hardening) and geometric nonlinearity (such as localized necking or buckling). As a result, it may be argued that the inverse problem of calibrating material parameters from load displacement curves is ill-posed in general, since infinitely many deformation fields are admissible within the applied boundary conditions (i.e. the load and displacement response measured at the boundaries), and that these fields depend on the constitutive response. Other researchers (Yoshida et al., [18]; Yun and Shang, [19]; Gelin and

Ghouati, [20]) have identified the ill-posed nature of this problem, recognizing that it renders the calibration susceptible to non-unique parameter fits. This means that many sets of calibrated parameters may closely (or exactly) match the load-displacement data from the tests. The specific implications of this in the context of structural modeling are the following –

1. Since multiple constitutive parameter fits are theoretically possible for a set of load displacement calibration test data, these are liable to producing dissimilar load deformation response for *application* components that are dissimilar (in terms of stress distribution or loading history) as compared to the calibration data set. Moreover, it is impossible to determine *a priori* which of these parameter sets is the appropriate representation of constitutive response. This implies that the predictions of response based on the calibrated parameters are non-robust.
2. As discussed above, the non-unique fits are a result of multiple admissible strain fields within the applied boundary conditions. As a result, application simulations based on these fits may result in different strain fields (even if the load deformation response is not different). This has adverse implications when the stress/strain fields are used for prediction of fracture or other damage states utilizing criteria such as Stress Modified Critical Strain – SMCS (Hancock and Mackenzie, [21]) or the Void Growth Model (Rice and Tracey, [22]).

Motivated by these issues, the main objectives of this paper are –

1. To quantitatively examine the susceptibility of common calibration specimens (such as shown in Figures 1a and b) and procedures, to inaccuracies arising due to the above mentioned non-uniqueness of constitutive parameter calibration for structural steels.

2. To quantify the effect of this non-uniqueness on performance prediction (i.e. simulation) of limit states (such as load deformation response, peak loads and critical strains) in structural steel components.
3. Based on (1) and (2) above, to characterize the inaccuracy in the simulation of structural steel components with the ultimate objective of providing support for methods and best practices to mitigate this inaccuracy.

It is emphasized here that the purpose of the paper is not to develop or refine a material constitutive model, or to assess the efficacy of an existing material model against experimental data. The purpose of the paper is *to assess the process* by which constitutive models are calibrated, and provide suggestions for enhancing these processes. The paper begins by summarizing the methodology of the research, which is based on a set of 2400 FE simulations. The simulations use a variant of the Armstrong and Frederick [8] constitutive model, and examine various strategies and tolerances for calibration, and the effect of these strategies on the accuracy of results. A critical analysis of the factors underlying the observed issues is then presented. The paper concludes by summarizing the results, and suggesting optimal strategies for robust calibration.

METHODOLOGY

Figure 2 schematically illustrates the methodology used to achieve the objectives outlined above. A brief overview of the methodology is first presented, followed by details of each component within it. The methodology relies on the notion of a “true” material, and the precept that the purpose of calibration is to determine material parameters by matching the load-displacement

curves of FE simulations of calibration specimens using trial parameters to the load-displacement curves of identical specimens which feature the true material. For the purposes of this study, a “true” material is generated synthetically, followed by the generation of 299 trial parameter sets. Finite Element simulations of various specimen types (discussed later) are conducted with the true material, as well as the trial parameter sets. In this way, trial parameter sets that produce an acceptable match with the true load-displacement curves (within specified tolerances) are identified. Analysis of these acceptable trial parameter sets is used to develop insights regarding the potential for non-unique fitting. Finally, the trial parameter sets are used to predict structural performance metrics (such as peak load and internal strains), to examine the effect of non-unique fitting on predictions of structural performance.

The methodology has four major components: (1) the constitutive model, (2) the parametric simulation process, (3) the specimen geometries and FE models, and (4) the development of error and fitting measures. Each of these is now described in detail.

Constitutive model

A variant of the material constitutive model proposed by Armstrong and Frederick [8] was used for the simulations in this study. The model (hereafter referred to as the AF model) was selected for the following reasons (1) the model has been used successfully by the authors and other researchers (Fell et al. [1]; Prinz and Nussbaumer, [23]) to simulate the response of structural steel components subjected to complex stress states and loading histories, (2) the model can simulate various important aspects of response, such as nonlinear isotropic and kinematic hardening with a relatively modest (i.e. 6) number of adjustable hardening parameters, and (3) it (or a variant) is implemented in FE platforms such ABAQUS [24] and ANSYS (2014)[25] such

that it is widely used by the steel research and practicing communities. The model itself is based on a von-Mises yield surface with associative plasticity (Lubliner, [6]). This implies that the plastic flow is isochoric, and insensitive to pressure (or mean-stress). The model is similar to bounding surface models (e.g. Dafalias and Popov [7]), but is more facile from the perspective of numerical implementation. The ABAQUS Theory Manual (ABAQUS, [24]) provides a detailed description of this model and implementation.

Figure 3 illustrates the reversed cyclic stress-strain response for a uniaxial stress condition to illustrate important aspects of response simulated by the model. As shown in the figure, the model is able to simulate nonlinear hardening, which may be decomposed into the isotropic part (expansion of the yield surface), and the kinematic part (translation of the yield surface, resulting in the Bauschinger effect). The figure also indicates associated translation of the yield surfaces in the π – plane (i.e., in the 2–dimensional projection of the principal stresses $\sigma_I, \sigma_{II}, \sigma_{III}$). Equations (1) – (4) below represent these types of hardening through evolution (i.e., expansion and translation) of the yield surface as a function of the equivalent plastic strain ε_p –

$$\sigma_{iso} = Q_{inf} \left(1 - \exp \left(-b * \varepsilon_p \right) \right) \quad (1)$$

$$\sigma_{kin,1} = \alpha_1 = C_1 * \varepsilon_p \quad (2)$$

$$\sigma_{kin,2} = \alpha_2 = \frac{C_2}{\gamma_2} \left(1 - \exp \left(-\gamma_2 * \varepsilon_p \right) \right) \quad (3)$$

$$\sigma_Y = \sigma|_0 + \sigma_{iso} + \sigma_{kin,1} + \sigma_{kin,2} \quad (4)$$

Equation (1) represents the isotropic component of the hardening (i.e. expansion of the yield surface). Two parameters, i.e. Q_{inf} (representing the maximum extent of yield surface expansion) and b (which controls the rate of this expansion) describe the evolution of the size of

the yield surface as a function of the equivalent plastic strain. The equivalent plastic strain itself is a scalar measure of accumulated strain, such that $\varepsilon_p = \int d\varepsilon_p = \int \sqrt{(2/3) \cdot d\varepsilon_{p,ij} \cdot d\varepsilon_{p,ij}}$. The equivalent plastic strain, defined in this way, is always positive and increases monotonically during loading. Equations (2) and (3) represent the kinematic translation of the yield surface, based on two backstress tensors (i.e., α_1 and α_2). Each backstress tensor represents a component of the center of the yield surface (and thus its translation). Referring to the equations above, the backstress evolves with respect to equivalent plastic strain according to the parameters C_1 , C_2 , and γ_2 . The instantaneous yield stress (σ_Y) is equal to the sum of the initial yield stress ($\sigma|_0$) and the hardening stresses due to isotropic and kinematic hardening. In general, additional kinematic hardening terms (i.e., backstress tensors) can be included to obtain improved agreement with experimental data. However, for most structural steels, using one or two backstresses is typically sufficient (e.g., Mahan et al. [26]). In fact, as discussed by Smith et al. [27], the use of additional backstresses results in more parameters that need calibration, exacerbating the non-uniqueness problem. In this paper, the two-backstress variant of the A-F model is used, such that a total of 6 parameters (i.e., Q_{inf} , b , C_1 , C_2 , γ_1 , and γ_2) must be calibrated to represent material constitutive response. This assumes that the elastic constants E (the Young's Modulus), ν (the Poisson's ratio), and $\sigma|_0$ (the initial yield stress) are assumed invariant, such that. $E = 29,000$ ksi, $\nu = 0.3$, and $\sigma|_0 = 50$ ksi.

Parametric simulation – based on true and trial parameter sets

The next step in the process is FE simulation of the four specimen types (discussed in the next section) for a range of material constitutive parameters. Using the AF constitutive model, a systematic process is adopted for parametric simulation. This involves the following steps -

1. A parameter set corresponding to a “true” material is established; this parameter set is now defined as the vector $\mathbf{X}_{true}$. Referring to prior discussion, the vector has 6 elements, which represent the parameters $Q_{inf} = 32$ ksi, $b = 9$, $C_1 = 15$ ksi, $C_2 = 525$ ksi, $\gamma_1 = 0$, and $\gamma_2 = 19$. Specifically, the values of the parameters are determined to provide a best fit with a large number of tests (on various specimen types) conducted by Smith et al. [27]. Note that the main purpose of developing this parameter set is to generate a synthetic benchmark material, whose response may be considered “true,” such that calibration errors may be assessed relative to it. To this end, the primary consideration is not replication of an exact material response, but rather a parameter set that represents a realistic material. The use of synthetic material data, instead of experimental material stress-strain response is advantageous for two reasons. First, because the “true” response can be replicated exactly by the constitutive model (since it is generated from it), the analysis of errors provides a focused assessment of the calibration approach, and the non-uniqueness problem is uncorrupted by the inadequacy of the material model to simulate experimental response. Similarly, the assessment is not affected by experimental noise, which is always present in measured data. Second, and perhaps more importantly, the synthetic material (when applied to FE simulations of common specimen geometries – see next subsection) results in simulated “test” data, which provides the opportunity to perform the error analysis using internal stress and strain fields (which are available for the FE simulations), rather than limited measurements (such as boundary loads and displacements) that are typically the only data streams commonly available for physical experiments. This is important because in addition to the load-deformation response, FE simulations are often used to quantify local stress or strain fields,

which are used in conjunction with fracture (or damage) criteria to determine ultimate limit states. Consequently, assessment of errors with respect to these fields is critical.

2. Once the true parameter set is established, 299 additional trial parameter sets (where each parameter vector may be denoted as $\mathbf{X}_i$; $i = 1:300$) are generated in the space of the 6 parameters. These are generated to span the 6-space (described by the 6 parameters) centered on the true parameter set, i.e. $\mathbf{X}_{true}$. These parameters are each varied within a $\pm 10\%$ range of their values in $\mathbf{X}_{true}$. Within these ranges, a process of random sampling is used for the generation of these trial parameters, wherein all the parameters are independently sampled from uniform distributions within the specific range for that parameter. In contrast to more traditional ways of sampling high-order parameter spaces (such as a full factorial parameter set generation, which would generate n^6 parameter sets if n realizations were used for each of the 6 parameters), the random sampling method is more efficient, and yet spans the parameter space.
3. Using each trial parameter set (including $\mathbf{X}_{true}$), FE simulations are conducted for the specimen geometries and loading histories described in subsequent sections.

Specimen Geometries and Finite Element Simulations

Four specimen geometries were selected for the FE simulation. These are illustrated schematically in Figures 4a-c.

1. Cylindrical Tension (CT) specimen (Figure 4a): This specimen is commonly used for characterizing stress-strain response of steel and other materials, and features in standard test protocols such as ASTM E8 (ASTM, [9]). However, this specimen is susceptible to tension necking and the resulting non-homogeneity in the stress and strain field. As per standard

practice, these specimens may be used to calibrate material stress-strain response in the pre-necking phase based on the test data itself (i.e., without complementary simulation). Alternatively, if FE simulations are conducted, then the post-unstable-necking phase of these experiments may be utilized along with the complementary FE simulations to characterize the constitutive response over an extended strain range. For this study, a standard CT specimen with diameter $d_{CT} = 0.5$ inches, and a length of 3 inches was used for the simulations.

2. Cylindrically Notched Tension (CNT) specimens (Figure 4b): These specimens are commonly used for calibrating local fracture toughness properties of metals because they provide a triaxial stress state in the notched region, promoting fracture (Kanvinde and Deierlein, [28]; Panontin and Sheppard, [29]; Prinz and Nussbaumer, [23]). However, they are also convenient for calibrating material constitutive response (when paired with complementary FE simulations), because unlike the CT specimens (which undergo severe necking deformations and therefore large changes in stress state over the duration of loading), the CNT specimens maintain a fairly constant stress field over the entire loading process. In this study, CNT specimens with two notch diameters and $d_{notch} = 0.125$ (denoted CNT1) $d_{notch} = 0.0625$ (denoted CNT2) were used.
3. Grooved Plate (GP) specimen (Figure 4c): This specimen was included as representative example of a structural application; the stress state in this type of specimen approaches plane strain with a relatively high stress triaxiality (i.e., $\sigma_m/\sigma_e = 1.3$), which is similar to that encountered in structural components of interest, such as beam flanges in the regions of connections or plates with notches or holes in them. Finite element models were built for each of these specimens.

Figure 5a shows an illustration of a deformed FE mesh for a CNT1 specimen. Simulations for the CT and CNT specimens were axisymmetric, and utilized two-dimensional CAX8R elements (i.e. 8 node elements with reduced integration). Approximately 1000 elements were used for the CT specimen, whereas 1500 elements were used for the CNT specimen owing to mesh refinement in the notched region. On the other hand, the simulations for the GP specimens (see Figure 5b) were 3-dimensional models, featuring 13500 20-node hexahedral elements with reduced integration. The element types were selected to mitigate spurious modes such as element locking and hourglassing while preserving fidelity of the simulation. The 3-d model for the GP specimens was a one-eighth model utilizing the 3 planes of symmetry. Figure 5b shows a vertical section of the model through its center. The visible portion model itself is visually replicated by mirroring the one-eighth model about two other axes of symmetry. The mesh within each model was refined to represent the stress and strain gradients with accuracy. For this purpose, a standard mesh sensitivity analysis with a 1% error threshold (i.e., refinement of meshes did not change stress or strain by more than 1%) was conducted. These meshes have been demonstrated to replicate various aspects of experimental response by Smith et al. [27] with reasonable fidelity, once the parameters of the constitutive model have been tuned.

The FE models generated in this way were executed using each of the 300 constitutive model parameter sets $\mathbf{X}_i$; $i = 1:300$. In addition, two loading histories were applied in the form of displacement controlled boundary conditions at the ends of the specimen. The important features of these loading histories are –

1. The monotonic loading history comprised a single tensile displacement Δ_{max} , which is different for each specimen, such that the Δ_{max} values for CT, CNT1, CNT2, and GP specimens are 0.2", 0.02", 0.015", and 0.04" respectively. The magnitude of this displacement reflects typical displacements corresponding to fracture (based on experimental data by Smith et al. [27], Myers et al. [2]) for specimens that are geometrically identical to the ones simulated. Since the simulated specimen never fractures, the intent is to preserve realism in the magnitude of strains and deformations observed in the various specimens.
2. The cyclic loading history consisted of two reversed cycles followed by a pull to fracture. The final pull to fracture for each specimen was equal to the Δ_{max} for the respective specimens. The displacement limits at which reversals were applied correspond to the displacement associated with the peak load, i.e., Δ_{F-max} observed in the specimen of interest when simulated with the parameter set $\mathbf{X}_{true}$. This serves as a convenient reference displacement for each specimen type, and facilitates the comparison of trends across specimens (discussed later), because unstable necking marks the onset of localized deformation, and an associated rapid increase geometric nonlinearity and associated strains. As a result, evaluating the efficacy of the calibration approaches is especially informative in the context of the necking deformation.

To illustrate the response of the FE models, Figure 6 shows the load displacement curves obtained from the CT specimens using the $\mathbf{X}_{true}$ parameter set; the figure includes both monotonic and cyclic curves. Referring to the figure, the response consists of an initial elastic region followed by a phase of strain hardening. This is followed by the onset of necking instability, where the load begins to drop after attaining a maximum value of F_{max} . This regime

continues until the simulation is terminated. The cyclic load displacement curve follows the “backbone” of the monotonic curve. Qualitatively, the curves for the other specimens (i.e. CNT and GP) are similar. It is interesting to note that although the FE simulation for the CT specimen does not include a notch imperfection, it still exhibits a necking instability. While somewhat counterintuitive (because the stress state in the specimen is perfectly homogenous), this may be attributed to the boundary conditions at the ends which prevent lateral (i.e. radial) contraction as the specimen elongates. This results in lower stress constraint (and thereby higher proclivity towards yielding) at the center of the specimen, where the instability occurs.

Based on the above discussion, a total of 2400 FE simulations ($= 4$ specimen geometries $\times 2$ loading histories $\times 300$ parameter sets) was conducted. Three data streams were recovered from each of these specimen runs. Specifically, these are (1) the load, (2) the displacement recorded at the specimen boundaries (representative of load displacement data that would be measured in an experiment), and (3) the equivalent plastic strain ϵ_p at one critical location inside the specimen. For each of the specimens, ϵ_p is recorded at the center of the specimen, as shown in Figure 5. Numerous experimental and computational studies (e.g., Hancock and Mackenzie, [21]; Rousselier, [13]) have shown that fracture initiates at this location in each of the specimen types, such that the strain at this location is used in predictive models for fracture. As a result, this quantity is of interest within the context of this study, to examine the effect of non-uniqueness on its estimation.

Error and fitting measures

Once the three data streams discussed above are recovered for each of the simulations, they may be utilized to compute meaningful error estimates for each of the simulations (i.e., for each

combination of parameter set, loading history and specimen type). The error estimates are calculated as the difference between the response for a given parameter set $\mathbf{X}_i$, and response for the true parameter set $\mathbf{X}_{true}$. By definition, all errors are zero for all simulations corresponding to $\mathbf{X}_{true}$. From the three data streams in the previous subsection, three error measures are generated. Two of these (which correspond to estimates of structural response) reflect the error in the maximum computed load, and the equivalent plastic strain at the center of the specimens. The third is a calibration error which characterizes the match between the load-displacement curve corresponding to any trial parameter set $\mathbf{X}_i$ and the curve corresponding to $\mathbf{X}_{true}$. For the plastic strain ε_p and the peak load F_{max} , the errors may be calculated as follows –

$$\epsilon_{\varepsilon_p}^i = |\varepsilon_p^{\mathbf{X}_i} - \varepsilon_p^{\mathbf{X}_{true}}| / \varepsilon_p^{\mathbf{X}_{true}} \quad (5)$$

$$\epsilon_{F_{max}}^i = |F_{max}^{\mathbf{X}_i} - F_{max}^{\mathbf{X}_{true}}| / F_{max}^{\mathbf{X}_{true}} \quad (6)$$

In the above equations, the terms on the right hand side refer to the response metric measured in the simulation for true parameter set, e.g. $\varepsilon_p^{\mathbf{X}_{true}}$ and a corresponding value measured for the parameter set $\mathbf{X}_i$, i.e. $\varepsilon_p^{\mathbf{X}_i}$. For the equivalent plastic strain, the resulting quantity $\epsilon_{\varepsilon_p}^i$ is a data stream over the entire duration of the loading history. On the other hand, $\epsilon_{F_{max}}^i$ is a discrete value since it is measured only at one instant during the loading history, when the load is at its maximum value. The error measures defined in Equations (5) and (6) above are generic descriptors, and may be specified for specimen type or loading history. For example, $\epsilon_{F_{max}}^{i,CNT-1,M}$ denotes the error in the maximum load for the simulation parameter set $\mathbf{X}_i$ used in conjunction with the CNT-1 model subjected to monotonic loading. Similarly, $\epsilon_{\varepsilon_p}^{i,CT,C}$ denotes the error in equivalent plastic strain at the center of the CT specimen (at any instant in the loading history), for the parameter set $\mathbf{X}_i$, where the specimen is subjected to cyclic loading.

In addition to the error measures for the maximum load and equivalent plastic strain, an additional error measure is established to characterize the match between the load-deformation curve generated by a trial parameter set $\mathbf{X}_i$, and the load displacement curve generated by $\mathbf{X}_{true}$. This error measure is particularly important since it enables an objective characterization of fitting error, in contrast to visual matching (or “eyeball fitting”) which is commonly used in the trial and error process for material parameters. The error measure is defined as follows –

$$\epsilon_{F\Delta}^i = \frac{\int |F_{\mathbf{X}_i} - F_{\mathbf{X}_{true}}| |d\Delta_p|}{\int |F_{\mathbf{X}_{true}}| d\Delta_p} \quad (7)$$

The above error measure is a discrete value, calculated for each simulation run. The terms within the integral, the terms $F_{\mathbf{X}_i}$ and $F_{\mathbf{X}_{true}}$ refer to the force at a given instant obtained from the simulation using the parameter sets $\mathbf{X}_i$ and $\mathbf{X}_{true}$ respectively. The term $d\Delta_p$ refers to the incremental plastic displacement in that analysis step. The integrals in Equation (7) above are calculated numerically from the load and displacement data streams discussed previously. This error measure, developed by Smith et al. [27], is based on the energy (area) between responses derived from the true ($\mathbf{X}_{true}$) and the trial ($\mathbf{X}_i$) parameter sets. However, several modifications distinguish it from more traditional energy based error measures (e.g. Pal et al. [30]). Smith et al. [27] describes these modifications (and motivations for them) in comprehensive detail. Here, only the key features of this error measure are summarized, i.e., it is (1) accurate, implying that a perfect fit corresponds to zero error, with smoothly increasing error as the fit degrades, and (2) balanced, i.e. the error is not biased towards specific types of errors or specimens. Smith et al. [27] adopted a process of trial and error during which multiple error measures were examined and refined, ultimately resulting in the error measure proposed in Equation (7) above. Following this research, this error measure, i.e. $\epsilon_{F\Delta}^i$ is considered to be an objective indicator of the

agreement between load displacement, such that $\epsilon_{F\Delta}^i = 0$ implies that the load displacement curve obtained from parameter set $\mathbf{X}_i$ perfectly matches the curve obtained from parameter set $\mathbf{X}_{true}$. Similar to the other error measures, $\epsilon_{F\Delta}^i$ may be determined for specific specimen geometries and loading histories, and named accordingly, such as $\epsilon_{F\Delta}^{i,CT,C}$ for the CT specimen to which a cyclic loading history is applied. For illustrative purposes, Figures 7a and 7b provide a visual sense of the error magnitudes $\epsilon_{F\Delta}^i$ for monotonic and cyclic loading respectively. In the figures, load displacement data for CNT-1 specimens is plotted for parameter sets that correspond to $\epsilon_{F\Delta}^i = 0.02, 0.05$ and 0.10 . Also plotted on the figure is the load displacement curve corresponding to the true material properties, i.e. $\mathbf{X}_{true}$, for which $\epsilon_{F\Delta}^i = 0$. The figures suggest that $\epsilon_{F\Delta}^i = 0.02$ implies a fit that may be considered satisfactory within the traditional eyeball fitting approaches. On the other hand, the fit for $\epsilon_{F\Delta}^i = 0.05$ is less satisfactory, while the fit for $\epsilon_{F\Delta}^i = 0.10$ may be rejected with a high likelihood as per eyeball fitting methods.

The error measures $\epsilon_{\varepsilon_p}^i$, $\epsilon_{F_{max}}^i$, and $\epsilon_{F\Delta}^i$ are thus determined for all 2400 simulations. The first is determined as a time series, with the possibility of querying its value at selected instants in the loading history, whereas the two latter quantities are computed as discrete values for the entire simulation. The next section analyzes the relationships between these error measures and the implications for accurate calibration and simulation.

RESULTS AND DISCUSSION

This section presents a synthesis of the error measures determined as per the process outlined in the previous section. First, the potential for non-uniqueness of parameter fitting is examined by the development of “error contours” that plot the fitting error measure, i.e., $\epsilon_{F\Delta}^i$ within the

parameter space. The objective is to examine regions of the parameter space over which the fitting error is low, since large regions with low fitting error would suggest non-unique fits. For the purposes of illustration, Figure 8a shows superimposed contours of the error measures for two types of specimens, i.e., the CT and the CNT1 ($\epsilon_{F\Delta}^{i,CT,M}$ and $\epsilon_{F\Delta}^{i,CNT1,M}$) plotted in a two dimensional parameter space of Q_{inf} and γ_2 . Figure 8b shows similar error contours for the cyclic loading histories, i.e., $\epsilon_{F\Delta}^{i,CT,C}$ and $\epsilon_{F\Delta}^{i,CNT-1,C}$. Plotting the errors in a 2-d space allows for a visual interpretation of the results, and the variation of the error over the parameter space; a zero error implies a perfect fit. The plot is generated by selecting values of $\epsilon_{F\Delta}^{i,CT,M}$ and $\epsilon_{F\Delta}^{i,CNT-1,M}$ such that values of all other parameters (except Q_{inf} and γ_2) are held constant at their respective values within $\mathbf{X}_{true}$. In this manner, the parameter set $\mathbf{X}_{true}$ can be shown on the two-dimensional plot. In both Figures 8a and b, the contour maps (for both the monotonic and cyclic tests) are suggestive of valleys, with the lowest point (i.e. zero error) corresponding to the location of $\mathbf{X}_{true}$.

Referring to Figures 7a and b (for a visual sense of the error magnitudes) and Figures 8a and b, the following points may be made –

1. The darkened lines on Figures 8a and b enclose the regions within which the errors are less than the threshold values (i.e. 0.02, 0.05 or 0.10) for both the tests considered. Thus, for the region enclosed by the line marked 0.02, the fit for both the CT and the CNT-1 specimens may be considered satisfactory (refer Figure 7 and previous discussion), with the implication that any trial parameter set within this region may be selected as a fit as per the eyeball fitting approach.
2. For the monotonic tests (i.e., Figure 8a), the contours indicate larger and longer error valleys as compared to the cyclic tests (i.e. Figure 8b), with the implication that a larger portion of

the parameter space produces good agreement with the load-displacement curves for the monotonic tests, than for the cyclic tests. This is not surprising, because it suggests that parameters that are otherwise used to improve characteristics of the cyclic aspects of response (e.g., isotropic vs kinematic hardening) act as additional degrees-of-freedom for model calibration when only monotonic response must be matched. This results in a large combination of parameter sets that can successfully replicate the load-deformation response of the tests. This is further corroborated by the observation that the shapes of the contours for the monotonic and cyclic calibrations are entirely different, because for the two situations (i.e., monotonic vs cyclic), the contributions of the different parameters towards the error are dissimilar as well.

3. Following the above point, Figure 8b (with smaller valleys and tighter contours), when compared to Figure 8a also suggests that cyclic tests provide a more rigorous and accurate way of calibrating a parameter set that is nearer to $\mathbf{X}_{true}$. On the other hand, the elongated valleys of Figure 8a indicate that if the application scenario is to simulate the load displacement curve of a monotonically loaded component, then any $\mathbf{X}_i$ value within the valley (i.e., not necessarily $\mathbf{X}_{true}$) may be appropriate for this purpose. This is an important observation, which emphasizes the pragmatic consideration that the constitutive model calibration need not characterize the “true” material properties, as long as the intended objective (e.g., characterizing peak force) is adequately achieved.

In summary, Figures 8a and b, and the associated discussion confirm the susceptibility of the calibration procedure to non-unique fits. The issue then, is whether this non-uniqueness impacts the accuracy of structural simulation. In particular, it is of interest whether multiple acceptable

trial parameter sets (e.g., within the $\epsilon_{F\Delta}^i = 0.02$ error valley) result in significantly different estimates of structural response metrics (such as peak loads or internal strains), when applied to an application scenario (i.e., a structural component). To address this issue, two other error measures (that measured inaccuracy in response prediction) were analyzed for the simulation set. As discussed previously, these include the error in the predicted peak force, and the equivalent plastic strain at the center of the specimen.

Recall that for any parameter set $\mathbf{X}_i$, the measure $\epsilon_{F_{max}}^i$ (which is a discrete value) represents the error in the estimate of peak force, relative to the true parameter set $\mathbf{X}_{true}$. Figures 9a-d plot the error $\epsilon_{F_{max}}^i$ for four different types of specimens against the corresponding calibration error. More specifically, the plots are constructed in the following way –

1. In each of the Figures 9a-d, a total of 300 points (i.e., markers on the scatter plot) are plotted; each point corresponds to one of the parameter sets $\mathbf{X}_i$, and only one type of loading history (either cyclic or monotonic).
2. The horizontal coordinate of each marker on the scatter plot corresponds to the calibration error $\epsilon_{F\Delta}^{i,ensemble}$ for the specific parameter set, which is evaluated as the error in the load displacement curve $\epsilon_{F\Delta}^i$ averaged over the CT, CNT1, and CNT2 specimen types for the parameter set $\mathbf{X}_i$. This error measure represents one obtained from a typical calibration procedure where multiple types of specimens are used for calibration, and an overall agreement with all of them is used as a criterion for acceptance of the fit. As a result, the horizontal coordinate of a marker represents an average measure of fit for the load displacement curves, given the parameter set corresponding to the marker. The GP specimen is purposefully left out of the calibration ensemble to examine the effect of the calibration

process when applied to an “application scenario” which may differ in terms of stress state (or other conditions) from the calibration specimens. This is discussed later.

3. The vertical coordinate of the scatter points is the relative error in the estimate of peak force $\epsilon_{F_{max}}^i$ for each parameter set $\mathbf{X}_i$ as determined through Equation 6. This is computed for all the four specimen types, including the GP specimen.

For convenience of interpretation, Figures 9a-d also show average and maximum ratios between the response and calibration errors, and the 1:1 line. Referring to Figures 9a-d, the following observations may be made –

1. For each of the 4 specimen types, on average, the error in the performance metric $\epsilon_{F_{max}}^i$ is lower than the error in the fitting metric $\epsilon_{F\Delta}^{i,ensemble}$, such that for the four specimen types, the average value of $\epsilon_{F_{max}}^i / \epsilon_{F\Delta}^{i,ensemble}$ ranges between 0.73 (for the CNT specimens) to 0.82 (for the CT specimen) times the corresponding fitting error averaged over the entire ensemble of calibration data. This suggests that if the objective of the simulation (i.e., the application scenario) is to characterize peak load, then the fitting error based on load-displacement curve is an excellent metric, such that for a good fit based on load deformation curves (i.e., $\epsilon_{F_{max}}^i / \epsilon_{F\Delta}^{i,ensemble} < 0.02$), the relative error in the peak force will be less than 0.02 (on average). This is not surprising, because fitting based on $\epsilon_{F\Delta}^{i,ensemble}$ enforces a good fit over the entire load-deformation curve, of which F_{max} represents one point.
2. While the average values $\epsilon_{F_{max}}^i / \epsilon_{F\Delta}^{i,ensemble}$ are less than 1.0, the maximum values are higher, in the range of 1.12 (for CNT2) to 1.79 (for CT), suggesting that in some cases, the error in F_{max} may be significantly greater than the fitting error $\epsilon_{F\Delta}^{i,ensemble}$. This means that

potential for significant error in the response metric cannot be ruled out (although it can be mitigated) by selecting trial parameter sets with low fitting error.

3. Among the four specimen types, the largest mean value of $\epsilon_{F_{max}}^i / \epsilon_{F\Delta}^{i,ensemble} = 0.82$, with a maximum of 1.79 is determined for the CT specimens. This implies that given a calibration error, the prediction of the response metric is least accurate for the CT specimen. On one hand, this is somewhat surprising because the CT specimen was included in the calibration ensemble. On the other hand, it is also important to recall that the CT specimen shows the largest geometry change (deforming from a cylinder to a necked bar), and consequently the interactions between geometric nonlinearity and constitutive nonlinearity are the highest in the CT specimens. This exacerbates their proclivity to increased error in the response metric. The mean value of $\epsilon_{F_{max}}^i / \epsilon_{F\Delta}^{i,ensemble} = 0.80$ for the GP specimens. This indicates two things. First, the error ratio is greater as compared to the other two types of specimens in the calibration ensemble (i.e., the CNT1 and CNT2), suggesting that applying fitted parameters to a specimen not included in the calibration process decreases the accuracy of prediction. However, because the geometry change in the GP specimens is modest, the error ratio is less than that of the CT specimen. This indicates that the expected accuracy in an application scenario depends on a complex interplay of various factors.

The above discussion suggests that standard calibration procedures (based on eyeball or automated fitting of load displacement curves to a test ensemble) result in parameter fits that predict the peak component strength (in general) with a fair degree of accuracy. However, a key issue, not addressed by the preceding discussion, is the effect of non-unique calibration on the simulation of internal strains within the components. Recall that the non-uniqueness in parameter

fits is due to multiple admissible strain fields within the applied boundary conditions. As a result, it may be expected that different trial parameters that match load-displacement data will be associated with dissimilar strain fields. The following discussion examines this issue quantitatively.

As discussed previously, for each specimen type and parameter set $\mathbf{X}_i$, the equivalent plastic strain at the center of the specimen (see Figure 5) was determined, and the corresponding error $\epsilon_{\epsilon_p}^i$ was generated as a time series. Figures 10a,b, and c illustrate the dependence of the relative plastic strain error $\epsilon_{\epsilon_p}^i$ on the calibration ensemble error $\epsilon_{F\Delta}^{i,ensemble}$ for all specimen types, and for various stages in the loading. Before the plots are analyzed, some aspects of their development are discussed for clarity –

1. Each figure has four curves, each corresponding to one type of specimen as indicated in the legend. For each curve, the horizontal axis shows the stage of deformation (or loading); this is expressed in a normalized manner through the variable $\Delta_{cum}/\Delta_{F-max}$, where the numerator is the cumulative cyclic displacement at any stage in the loading, whereas the denominator (discussed previously), is the displacement associated with peak load or necking in a monotonic simulation of the specimen conducted with the parameter set $\mathbf{X}_{true}$. Normalizing the deformations in this way enables a meaningful comparison in trends between the various specimens, since Δ_{F-max} is indicative of the onset of localization in each specimen.
2. All the data shown (i.e., vertical coordinate) in each plot corresponds to a consistent level of ensemble error $\epsilon_{F\Delta}^{i,ensemble} \approx 0.02, 0.05, \text{ or } 0.10$. For example, in Figure 10a, the vertical coordinate of each curve is the average error $\epsilon_{\epsilon_p}^{avg}$, which is calculated as the average value

of the relative strain error $\epsilon_{\varepsilon_p}^i$ (as per Equation 5) over all simulations (i.e., parameter sets $\mathbf{X}_i$) that have $\epsilon_{F\Delta}^{i,ensemble}$ within a tolerance of 0.02 for a given specimen type. It is important to note here that the errors $\epsilon_{F\Delta}^{i,ensemble}$ are calculated after the generation of the parameter sets $\mathbf{X}_i$. Consequently, it is not possible to find $\mathbf{X}_i$ which have an error $\epsilon_{F\Delta}^{i,ensemble}$ precisely equal to 0.02. As a result, for the average calculation of $\epsilon_{\varepsilon_p}^i$, the parameter sets are selected such the resulting error $\epsilon_{F\Delta}^{i,ensemble}$ is within a tolerance ± 0.01 of 0.02. Figures 10b and 10c are generated similarly. As a result, the curves in these figures represent the evolution of the relative error in the plastic strain given a level of fit (in terms of force displacement curve) to the ensemble calibration set.

3. In Figures 10a-c, results only from cyclic simulations are plotted.

Referring to Figures 10a-c, several observations may be made –

1. In each of the figures, i.e., at each of the error levels, the CT specimen appears to have the maximum error, which is consistent with intuition because the CT specimen involves greater interaction between geometry change and material response, enabling the admissibility of highly dissimilar deformation fields given the applied boundary conditions or loading. The curves for the other specimens are significantly lower across all the error levels.
2. Figure 10a represents $\epsilon_{F\Delta}^{i,ensemble} \approx 0.02$, implying that any of the $\mathbf{X}_i$ that are represented in this figure will likely be acceptable as high-quality fits based on matching of the load displacement curve. Given this, it is disconcerting that as deformations increase (i.e., $\Delta_{cum}/\Delta_{F-max} > 0.5$), the average errors in the strain begin to reach unacceptable values. For example, at the end of the simulations (i.e., $\Delta_{cum}/\Delta_{F-max} \approx 1.5$), the error $\epsilon_{\varepsilon_p}^{i,AVG}$ for the CT specimen is approximately 0.3 (30%), whereas for the other specimens it is in the range of

0.05 to 0.09. It may be argued that if the strain at the center of the specimen is used for damage/fracture analyses, then these errors are unacceptably large. As expected, Figures 10b and c indicate a similar phenomenon, except that the error magnitudes are significantly larger. For example, referring to Figure 10b (which represents a good, although not excellent load-displacement fit), the $\epsilon_{\epsilon_p}^{i,AVG}$ for CT specimens is 0.4, whereas for GP it is 0.15. In any case, the main observation here is that fitting based on the load displacement curve provides no guarantee that the internal strain fields are accurately reproduced. Recall that the curves shown in Figure 10 represent the average value of $\epsilon_{\epsilon_p}^i$ for the parameter sets selected; the maximum errors for these are up to 25% higher than the average, depending on the specimen type. Moreover, it is also important to emphasize that the maximum displacements applied in the simulations are consistent with typical fracture deformations, implying that the large errors observed at the conclusion of the simulations correspond to physically realistic situations, and not some arbitrary applied deformation.

3. It is informative to examine the evolution of the error $\epsilon_{\epsilon_p}^{i,avg}$ with respect to deformation.

Note that the deformation is normalized by Δ_{F-max} which represents the onset of unstable necking. It is interesting that for the CT specimen, the error $\epsilon_{\epsilon_p}^{i,AVG}$ begins to rise sharply at approximately $\Delta_{cum}/\Delta_{F-max} = 0.5$. Although unstable necking has not been triggered until this point, note that the stress-state is still non-homogenous due to lateral constraints (and associated stable necking) at the ends of the CT specimen. This suggests that at this point in the loading, the geometric nonlinearity in the simulations is large enough to admit multiple dissimilar deformation fields for given boundary conditions. The error again rises sharply around $\Delta_{cum}/\Delta_{F-max} = 1.0$; this is not unexpected since this marks the onset of unstable necking and an even higher degree of geometric nonlinearity. For the GP specimens, a

similar increase in $\epsilon_{\varepsilon_p}^{i,AVG}$ is observed around $\Delta_{cum}/\Delta_{F-max}= 1.0$. The error increase in both the CNT specimens is more uniform.

SUMMARY AND CONCLUSIONS

Material constitutive models are central to FE simulation of steel structures, especially when extreme limit states are simulated. However, calibrating the parameters of these constitutive models is challenging because standard coupon tests produce homogenous stress states only for relatively low strains. For calibrating the parameters over larger strain ranges, the common approach involves matching load-displacement curves from FE simulations to those obtained from tests. These load displacement curves are a product not only of the material constitutive response, but also of its interaction with geometric nonlinearity (due to necking or buckling). This raises the possibility of non-unique calibration, wherein multiple material definitions produce similar load-displacement curves that closely match the experimental response. The paper aims (1) to examine the susceptibility of common calibration specimens/procedures to non-unique fitting, (2) to examine the impact of non-uniqueness on performance prediction of structural components, and (3) to suggest mitigation strategies.

The main scientific basis for the study is a set of 2400 FE simulations. The simulations are conducted on four specimen types representing common calibration specimens or generic structural geometries. The methodology involves: (1) generation of a synthetic benchmark material, which is considered to have “true” constitutive properties, (2) generation of 299 other trial materials or parameter sets, (3) examining the patterns of non-uniqueness, i.e., the extent of the parameter space over which trial materials may be considered acceptable fits, and (4)

examining the effect of selecting such acceptable fits on the accuracy of response metrics such as peak load and internal strain, assessed relative to the true material.

A rigorous index (or calibration error measure) was used to quantify the load-displacement curve fit between any trial parameter set and the true parameter set. For any given trial parameter set, a value of this error measure in less than 0.02 implies an excellent visual fit with load displacement data obtained from the true material. Based on this index, it was determined that large regions (error valleys) of the trial parameter space resulted in an acceptable or excellent fit with the true material load-deformation response. This confirmed the fundamental premise of this study, i.e., non-unique parameter fits are likely for structural steels, if standard fitting procedures are adopted. In practice, the non-uniqueness in the fitted parameters arises because any point within the error valley shows good agreement, and within this valley, the final selection of the parameter set may be arbitrary, or based on noise in the test data. A corollary of the large error valleys is that if characterization of only the load-displacement curves is the ultimate application of the FE simulations, then it may not be necessary to determine the true material parameters, since several other trial parameter sets (within the valley) also produce excellent agreement with the load-displacement curve.

After the susceptibility to non-unique fitting was established, the impact of non-uniqueness was evaluated for two structural response metrics, which are: the peak force (induced by unstable necking), and the equivalent plastic strain at critical locations within the specimens. For the assessment, the error in the predicted value of the response metric was evaluated against the calibration error in the trial parameter sets used for the prediction. This enabled the

characterization of errors in the response metrics, for material parameters that would be acceptable as per standard calibration procedures.

The error in estimation of the peak force is generally small if the trial material parameters satisfy acceptable tolerances for standard calibration procedures based on load deformation curves. However, while true on average, this observation has several exceptions, especially for the CT specimens, for which the error in the predicted peak force was significantly larger than that for the calibration fit for some of the trial parameters. Notwithstanding this, the general observation is that calibrating the material parameters based on load-deformation curves (with a set fitting tolerance or eyeball fitting) does not produce unexpectedly large errors in the peak force.

However, the examination of error in the internal plastic strain reveals a more concerning trend. It is observed that even for trial parameter sets that show excellent fit with the “true” material data (fitting error < 0.02), the errors in the plastic strains are fairly large, and increase with the applied deformation. For the CT specimens, where the geometry change is the largest, the error in the plastic strain is as large as 0.3, whereas for the other specimens, it is in the range of 0.05 to 0.09. These errors may be considered unacceptable if the plastic strains are used for characterizing damage or fracture. The errors in plastic strain are (expectedly) even larger if trial parameter sets with fitting errors of 0.05 or 0.10 are examined. The main conclusions of the study are –

1. Constitutive model calibration for structural steels is highly prone to non-unique parameter fits if standard procedures using matching of load-displacement curves are adopted.

2. The implications of non-uniqueness, i.e., of inaccurate parameter fits, are especially adverse (resulting in unacceptable errors) if the ultimate objective of the simulation is to characterize internal strain fields. As the use of strain-based damage and fracture indices proliferates within steel research and practice, it is critical to consider the inaccuracy generated by non-unique parameter fits.
3. Inaccurate parameter fits, and the associated errors in performance characterization may be mitigated by considering the following points. First, if the calibration specimens, loading histories, and the calibration data streams, are similar to (or include) the application scenario or response quantities, then the error in response characterization may be significantly reduced. For example, if characterizing the ultimate load due to necking/buckling is the ultimate application scenario, then it is desirable to have calibration specimens that exhibit similar behavior, and perform fitting based on load displacement data. Second, if the ultimate purpose is to characterize internal strain distributions, then it is highly desirable to perform the calibration against test data streams that may include such strain data (if possible or economical to measure).

It is useful to discuss here the implications of using a simulated “true” material, and an associated parameter set, instead of using experimental data. It is important to recall that the purpose of the paper is not to develop or refine a constitutive model for steel response; that is the domain of the constitutive modeling community, and several studies (including the ones cited in this paper) have demonstrated that existing constitutive models such as the Armstrong-Frederick model have excellent capabilities of simulating the response of structural steel subjected to monotonic or cyclic loading. With this considered, the benefit of using a simulated “true”

material model are obvious: (1) it provides the opportunity to examine internal stress and strain fields, in a manner that experimental data cannot, and perhaps more importantly (2) it provides an exactly searchable solution within the parameter space, that corresponds to “true” response. The latter is especially relevant, because it assesses the efficacy of the method, where the optimal solution of the calibration process is known *a priori* – meaning that an optimal calibration process will have the ability to return a zero error. Using experimental data for the “true” material implies that the perfect fit solution does not exist within the search space, and as a result the inaccuracy of the calibration method cannot be isolated from the inaccuracy of the constitutive model. As a result, it is believed the using a simulated parameter set is the correct approach. In fact, the presence of the large errors even when the perfect fit is available within the search space provides even a stronger criticism of the current fitting approach.

Finally, the existence of non-unique parameter fits underscores that multiple parameter sets are able to predict certain types of response with accuracy, even if they do not represent the true material. This raises a broader issue about the overall philosophy of material calibration. Specifically, it suggests that material calibration is ultimately controlled by pragmatic considerations, wherein the aim is not to characterize true material response in an absolute sense, but rather to calibrate a parameter set that is acceptable for a given application scenario. If interpreted as such, calibration may be performed considering tradeoffs between the cost of calibration (for example instrumenting specimens to measure local strains, in addition to loads and displacement), against the cost of inaccurate performance prediction.

The study quantitatively demonstrates the pitfalls of common material calibration methods, and the associated loss of accuracy in performance assessment. The limitations of the study must be considered while interpreting the results. Mainly, these limitations are in the areas of (1) the size and scope, such that only a small number of specimen types were considered along with one constitutive material model, and (2) the methodology, wherein a synthetic (rather than true) material was used along with a limited number (300) trial parameter sets, and only a limited number of error measures. Finally, the observations presented in this study follow an empirical format, since the problem is essentially a practical one. As a result, any extrapolation of the observed quantitative trends to situations (i.e., material models, specimen types) significantly different than the ones examined may not be valid. Nevertheless, it may also be expected that for real structural components (in which the application geometries are complex and highly dissimilar to the calibration specimens), the impact of non-unique calibration will be greater than it is for the relatively small specimens examined in this study.

ACKNOWLEDGMENTS

Major funding for the study was provided by the National Science Foundation (CMMI 0825155). The authors are grateful to Christopher Smith of the National Institute of Standards and Technology, and to Gregory Deierlein of Stanford University for constructive discussions. The assistance of Kimberly Stillmaker, David Grilli, and Vincente Pericoli in reviewing the manuscript is greatly appreciated.

REFERENCES

- [1] Fell, BV, Kanvinde, AM, Deierlein, GG, Myers, AT. Experimental Investigation of Inelastic Cyclic Buckling and Fracture of Steel Braces. J Struct Eng ASCE 2009; 135(1):19-32.
- [2] Myers, AT, Kanvinde AM, Deierlein, GG, Fell, BV. Effect of Weld Details on the Ductility of Steel Column Baseplate Connections. Journal of constructional steel research, 65(6). Elsevier; 2009. 1366-73
- [3] Purba, RH, Bruneau, M. Finite Element Investigation and Design Recommendations for Perforated Steel Plate Shear Walls. J Struct Eng ASCE 2009; 135(11): 1367-76.
- [4] Berman, JW, Bruneau, M. Experimental and Analytical Investigation of Tubular Links for Eccentrically Braced Frames. Engineering Structures 2007; 29(8): 1929-38.
- [5] Chao, S, Khandelwal, K, El-Tawil, S. Ductile Web Fracture Initiation in Steel Shear Links. J Struct Eng ASCE 2006;
- [6] Lubliner, J. Plasticity Theory. Dover Publications 2008.
- [7] Dafalias, YF, Popov, EP. A Model of Nonlinearly Hardening Materials for Complex Loading. Acta Mechanica 1975; 21: 173-92
- [8] Armstrong, PJ, Frederick, CO. A Mathematical Representation of the Multiaxial Bauschinger Effect. CEGB Report, RD/BN731, Berkeley Nuclear Laboratories 1966.
- [9] ASTM Standard E8. Metal Tensile Strength Testing. ASTM International 2014.
- [10] Ling, Y. Uniaxial True Stress-Strain after Necking. AMP Journal of Technology 1996; 5: 37-48.
- [11] Bridgeman, PW. Studies in Large Plastic Flow and Fracture. McGraw-Hill Book Company, London 1952.

- [12] Rodriguez, ME, Botero, JC, Villa, J. Cyclic Stress-Strain Behavior of Reinforcing Steel Including Effect of Buckling. J Struct Eng ASCE 1999; 125(6): 605-12.
- [13] Rousellier, G. Ductile Fracture Models and their Potential in Local Approach to Fracture. Nuclear Engineering and Design 1987; 105: 97-111.
- [14] Takeuchi, T, Matsui, R. Cumulative Cyclic Deformation Capacity of Circular Tubular Braces under Local Buckling. J Struct Eng ASCE 2011; 137(11): 1311-18.
- [15] Brandonisio, G, De Luca, A, Mele, E. Finite Element Analysis of Panel Zone in Welded Beam to Column Connections. Proceedings of the 5th International Conference on Behavior of Steel Structures in Seismic Areas – STESSA 2006; 285-90.
- [16] Chi, W, Deierlein, GG, Ingraffea, AR. Fracture Analysis of Welded Beam-Column Connections. Fatigue and Fracture Mechanics ASTM 1998; 30th Volume ASTM STP 1360.
- [17] Kulkarni, SA, Bing, L, Woon, KY. Finite Element Analysis of Precast Hybrid-Steel Concrete Connections Under Cyclic Loading. Journal of Constructional Steel Research 2008; 64(2): 190-201.
- [18] Yoshida, F, Urabe, M, Hino, R, Toropov, VV. Inverse Approach to Identification of Material Parameters of Cyclic Elasto-Plasticity for Component Layers of a Bimetallic Sheet. International Journal of Plasticity 2003; 19(12): 2149-70.
- [19] Yun, GJ, Shang, S. A Self-Optimizing Inverse Analysis Method for Estimation of Cyclic Elasto-Plasticity Model Parameters. Int J Plast 2011; 27(4): 576-95.
- [20] Gelin, JC, Ghouati, O. An Inverse Solution Procedure for Material Parameters Identification in Large Plastic Deformations. Communication in Numerical Methods Engineering 1996; 12(3): 161-73.

- [21] Hancock, JW, Mackenzie, AC. On the Mechanics of Ductile Failure in High-strength Steel Subjected to Multi-axial Stress States. *J Mech Phys Solids* 1976; 24: 147-69.
- [22] Rice, JR, Tracey, DM. On the Ductile Enlargement of Voids in a Triaxial Stress Field. *J Mech Phys Solids* 1969; 17: 201-17.
- [23] Prinz, GS, Nussbaumer, A. On the Low-cycle Fatigue Capacity of Unacnhored Steel Liquid Storage Tank Shell-to-Base Connections. *Bulletin of Earthquake Engineering* 2012; 10(6): 1943-58.
- [24] ABAQUS. ABAQUS/Standard User's Manual, version 6.14. Providence (RI, USA): Desault Systems; 2014.
- [25] ANSYS. Modeling and Meshing Guide for ANSYS Release 15.0. ANSYS, Inc.; 2014.
- [26] Mahan, M, Dafalias, YF, Taiebat, M, Heo, YA, Kunnath, SK. SANISTEEL: Simple Anisotropic Steel Plasticity Model. *J Struct Eng ASCE* 2011; 137(2): 185-94.
- [27] Smith, CM, Kanvinde, AM, Deierlein, GG. A Stress-Weighted Damage Model for Ductile Fracture Initiation in Structural Steel under Cyclic Loading and Generalized Stress States. Blume Center Technical Report TR-187. Stanford, CA: Stanford University; 2014.
- [28] Kanvinde, AM, Deierlein, GG. The Void Growth Model and the Stress Modified Critical Strain Model to Predict Ductile Fracture in Structural Steels. *J Struct Eng ASCE* 2006; 132(12): 1907-18.
- [29] Panontin, TL, Sheppard, SD. The Relationship between Constraint and Ductile Fracture Initiation as Defined by Micromechanical Analyses. *Fracture Mechanics* 1995; 26th Volume ASTM STP 1256.
- [30] Pal, S, Wathugala, GW., and Kundu, S. Calibration of a constitutive model using genetic algorithms." *Computers and Geotechnics*, 1995, 19(4), 325–348.

Figure 1

Figure 1 – Commonly used calibration coupons (a) CT – Cylindrical Tension (b) CNT – Circumferentially Notched Tension

Figure 2

Figure 2 – Research Methodology

Figure 3

Figure 3 – Constitutive model used in study (a) Cyclically loaded specimen (b) Stress-strain response (c) Yield surface motion and expansion in π – plane

Figure 4

Figure 4 – Specimen geometries (a) Cylindrical Tension – CT (b) Cylindrically Notched Tension – CNT, and (c) Grooved Plate – GP

Figure 5

Figure 5 – Finite element simulations showing contours of ϵ_p for (a) CNT-1 and (b) GP specimens

Figure 6

Figure 6 – Load displacement curves for CT specimens using the X_{true} parameter set for monotonic and cyclic loading histories

Figure 7

Figure 7 – Visual interpretations of various levels of fitting error (i.e., $\epsilon_{F\Delta}^i$) for (a) monotonic, and (b) cyclic simulations of CNT-1 specimens

Figure 8

Figure 8 – Contour plots of the fitting error (i.e., $\epsilon_{F\Delta}^i$) in the space of two of the constitutive parameters for the CT and CNT-1 specimens (a) monotonic, and (b) cyclic simulations

Figure 9

Figure 9 – Error in the estimated peak force F_{max} in the four types of specimens (a) CT (b) CNT-1 (c) CNT-2, and (d) GP, plotted against the corresponding fitting error $\epsilon_{F_{\Delta}}^{i, \text{ensemble}}$ for the entire simulation ensemble

Figure 10

Figure 10 – Evolution in the average error in equivalent plastic strain at center of specimens with increasing deformation for (a) $\epsilon_{F\Delta}^{i,ensemble} \approx 0.02$, (b) $\epsilon_{F\Delta}^{i,ensemble} \approx 0.05$, and (c) $\epsilon_{F\Delta}^{i,ensemble} \approx 0.10$

Figure Caption List

Figure 1 – Commonly used calibration coupons (a) CT – Cylindrical Tension (b) CNT – Circumferentially Notched Tension

Figure 2 – Research Methodology

Figure 3 – Constitutive model used in study (a) Cyclically loaded specimen (b) Stress-strain response (c) Yield surface motion and expansion in π – plane

Figure 4 – Specimen geometries (a) Cylindrical Tension – CT (b) Cylindrically Notched Tension – CNT, and (c) Grooved Plate – GP

Figure 5 – Finite element simulations showing contours of ε_p for (a) CNT-1 and (b) GP specimens

Figure 6 – Load displacement curves for CT specimens using the $\mathbf{X}_{true}$ parameter set for monotonic and cyclic loading histories

Figure 7 – Visual interpretations of various levels of fitting error (i.e., $\varepsilon_{F\Delta}^i$) for (a) monotonic, and (b) cyclic simulations of CNT-1 specimens

Figure 8 – Contour plots of the fitting error (i.e., $\varepsilon_{F\Delta}^i$) in the space of two of the constitutive parameters for the CT and CNT-1 specimens (a) monotonic, and (b) cyclic simulations

Figure 9 – Error in the estimated peak force F_{max} in the four types of specimens (a) CT (b) CNT-1 (c) CNT-2, and (d) GP, plotted against the corresponding fitting error $\varepsilon_{F\Delta}^{i,ensemble}$ for the entire simulation ensemble

Figure 10 – Evolution in the average error in equivalent plastic strain at center of specimens with increasing deformation for (a) $\varepsilon_{F\Delta}^{i,ensemble} \approx 0.02$, (b) $\varepsilon_{F\Delta}^{i,ensemble} \approx 0.05$, and (c) $\varepsilon_{F\Delta}^{i,ensemble} \approx 0.10$