Handling Bias and Uncertainty in Model Verification & Validation associated with Heated Pipes Pressurized to Failure[1]

Vicente Romero, J. Franklin Dempsey, Bonnie Antoun, Gerald Wellman, Martin Sherman
Sandia National Laboratories,[2] Albuquerque, NM

Abstract

This paper demonstrates versatile and practical model validation and uncertainty quantification (UQ) techniques applied to the accuracy assessment of a temperature dependent elastic-plastic constitutive model tested on heated steel pipes pressurized to failure. The "Real Space" validation methodology segregates aleatory and epistemic uncertainties in the validation procedure to form straightforward model validation metrics especially suited for assessing models and prediction quantities to be used in the analysis of performance and safety margins. The methodology handles difficulties associated with representing and propagating interval and/or probabilistic uncertainties from multiple correlated and uncorrelated sources in the experiments and simulations including:

- **material variability characterized by non-parametric random functions (temperature-dependent discrete stress-strain curves);**
- **very limited (sparse) experimental data both at the coupon testing level for material characterization and at the pipe-test validation level;**
- **boundary condition reconstruction uncertainties from spatially sparse sensor data;**
- **normalization of pipe experimental responses for measured input-condition differences among tests and for random and systematic uncertainties in measurement/processing/inference of experimental inputs and outputs;**
- **numerical solution uncertainty from model discretization and solver effects.**

1. Introduction

A current project at Sandia National Laboratories is the modeling of stainless-steel pressure vessel response at high pressures and temperatures, up to initiation of failure ([1]). The first vessels to be modeled are simple pipe geometries. It is desired to test a temperature-dependent constitutive model of stainless-steel response over large ranges of pressures and temperatures, ramp rates, and large temperature gradients and ranges of spatial temperature variation. Figures 1.1 and 1.2 portray some of the associated "pipe bomb" (PB) hardware and experiments and coupled thermal-mechanical modeling. Controlled pressurization of the pipe is accomplished via pressure supply tanks. Details of PB geometry and experimental conditions and results important to this paper are presented herein. Further details of the design of the hardware and experiments, and execution and results of the experiments, are given in [2].

[2] Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

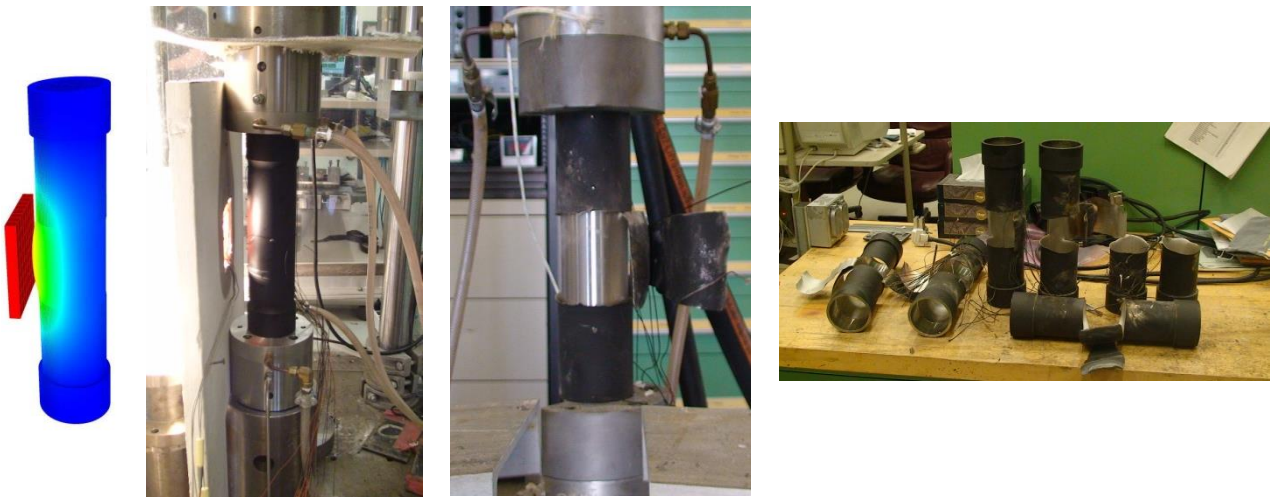

Figure 1.1 Model validation experiments: nitrogen-pressurized pipe is heated by hot inconel plate until pipe bursts at hot spot and tears back along upper and lower thickness-transition shoulders in pipe wall, leaving "butterfly wings" as shown. For safety reasons, inner slug shown (shiny silver) fills up most of inside volume of pipe, lessening explosive energy built up inside pipe prior to pipe breach failure. Pipe is approximately 14 inches high and 3.5-in. in diameter, with mid-region wall thickness of nominally 0.02 in. and upper & lower region thicknesses of nominally 0.05 in.

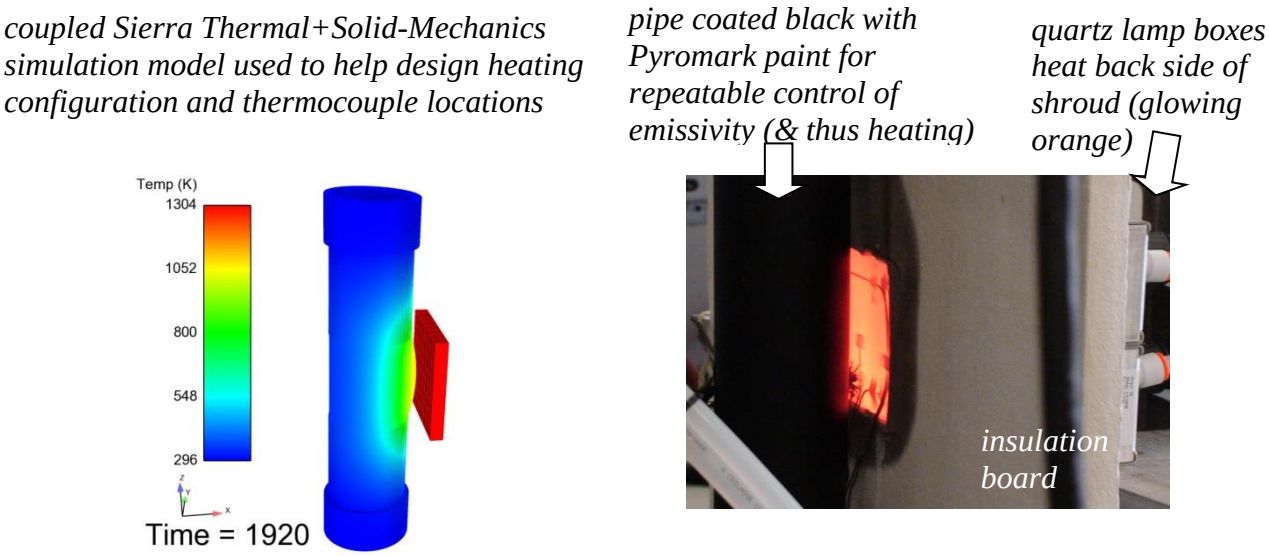

Figure 1.2 Thermal modeling helped with design, instrumentation, and analysis of experiments. Thermocouples with wire leads can be seen on the inconel heating shroud (glowing orange).

The project required the formulation and development of an approach for including the significant temperature dependence of strength (stress-strain response) of 304L stainless steel over temperature excursions of interest, nominally 25C to 800C. The mathematical and algorithmic formulation of the temperature-dependent multi-linear elastic-plastic (MLEP) constitutive model for material behavior is documented in [3], [4].

Development of the constitutive model required new experimental characterization of stress-strain behavior in material-coupon round-bar tension tests performed at set temperature levels spanning the range of interest (see [2]). Several nominally identical replicate tests were performed with new material samples each time to characterize the effects of material variability on exhibited strength at the set temperatures. The measured stress-strain curves for the investigated temperature levels are presented in Section 2 of this paper.

It is important to note that each stress-strain curve comprises a ***discrete*** random function that has no identifiable parametric relationship to other stress-strain curves at that temperature. A novel uncertainty quantification (UQ) approach based on just a few samples from a larger population of discrete random processes or functions is described in Section 2 that compensates for limited (sparse) numbers of material tests.

It should be emphasized that the temperature dependent constitutive model is the subject of the validation assessment. The pipe-vessel model depicted in Figures 1.1 and 1.2, and the solid mechanics code [5] (massively parallel 3-D implicit nonlinear quasi-statics), are the vehicles by which the constitutive model's validity is determined. But it is the constitutive model and not the pipe model that is intended to be used for further predictions subsequent to this model validation activity, and perhaps with other solid mechanics codes. Terminology developed to convey this distinction ([6] - [8]) is that the pipe-response model is the "experiment model" in this validation activity and the material constitutive model is the "traveling model" to be validated. It "travels" to subsequent uses beyond the present validation activity.

Section 3 describes the use of modeling and simulation to help design the experiments and thermocouple locations to minimize errors and uncertainty associated with the experiments and modeling of the boundary conditions from spatially sparse sensor information.

Section 4 presents the PB validation experiments and simulations, their uncertainties, and processing of results and uncertainties for comparison within a "Real Space" model validation framework. A major aspect of the processing of model predictions and experimental results involves accounting for small numbers of tests at the material characterization level which populates the constitutive model, and at the pipe validation testing level. Ultimately, uncertainty ranges of experimental and predicted 0.025 and 0.975 percentiles of response (failure pressure) are compared. Analysis and straightforward interpretation of the comparison results are provided.

Section 5 closes with some comments on the Real Space validation methodology features and capabilities for handling the challenging PB validation problem as compared to some other validation methodologies in the literature.

2. Quantification and Propagation of Material Behavior Variability and Epistemic Uncertainty associated with the Constitutive Model

2.1 Experimental Stress-Strain Variability of Tested Material Samples

Figure 2.1 shows elements of the cylinder tension-test experiments and results from which the stress-strain curves for the constitutive model were derived. Details of the tension-test apparatus, experimental setup, measurement and control instrumentation and calibration, etc. are given in [2].

The cylindrical material samples were cut and machined from the same 304L stainless steel tubular stock (3.5in. dia., ¼ in. wall thickness) that the validation-experiment pipe vessels were machined from. The coupon specimens were vacuum annealed at 1000C for 30 minutes to produce the same anneal conditions present in the pipe vessels. For each of the temperature levels indicated in Figure 2.1, several cylinder specimens were tested to failure. The ends of the cylinders were pulled in axially opposing directions with slowly increasing force to produce axial tension at a strain rate of 0.001/s, surmised from hand calculations to be representative of pipe wall strains in the PB validation tests. Measured axial displacement and resisting force were transformed to the experimental "engineering stress-strain" response curves plotted in Figure 2.1.

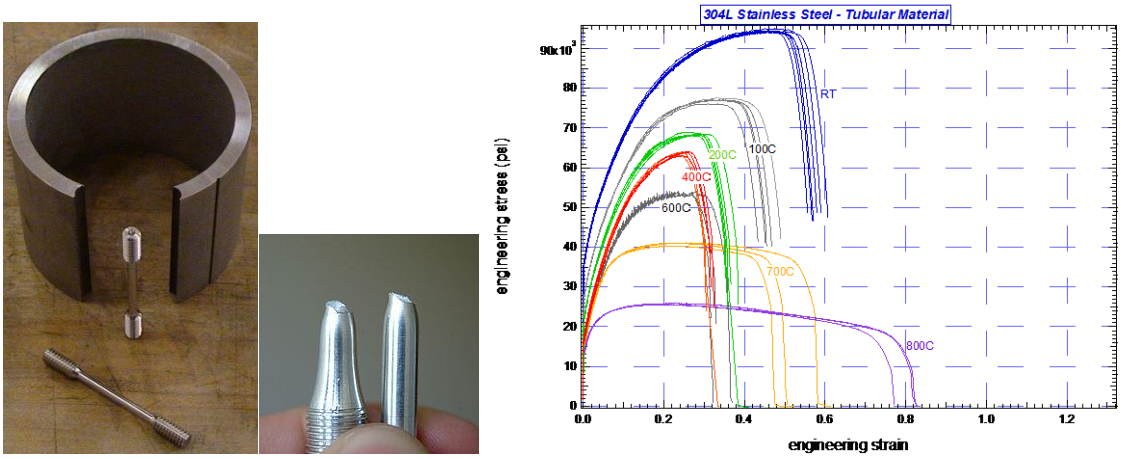

Figure 2.1 Cylinder tension-test material samples were cut and machined from the same 304L stainless steel tubular stock that the validation-experiment pipe vessels were machined from. Measured stress-strain response-to-failure curves plotted from cylinder tension-tests at a strain rate of 0.001/s for the labeled temperatures (note: "RT" in the plot stands for "room temperature", nominally 20C).

2.2 Inverse Calculations to Convert from Measured Stress-Strain Curves to Constitutive Model Stress-Strain Curves

For each measured "engineering" stress-strain curve, advanced optimization techniques were used to solve the inverse problem of determining the constitutive model's corresponding "true" stress-strain curve (see [4]). This enables the finite-element (FE) model of the cylinder to best reproduce the measured engineering stress-strain response in the tension tests. Thus, the modeled

cylinder deforms and necks when pulled to experimentally measured displacement and resisting force vs. time, in a manner that closely matches the test results.

The solid mechanics code Adagio[5] was used for the FE model forward simulations in the inversion procedure. In [9] a discussion of the cylinder model FE mesh and solver discretization choices indicates that model results are insensitive to significant perturbations about the model's discretization settings used in the inversion procedure. Therefore the cylinder model mesh and solver settings are presumed adequately refined for the inversion purposes here.

2.3 Incorporating multiple stress-strain curves of material variability and accompanying uncertainty from small numbers of material tests

Note from Figure 2.1 that the numbers of material variability tests (numbers of stress-strain curves at each temperature) are relatively small. When only a few samples of a random variable or function are available, these will usually significantly misrepresent the randomness properties of the source of variability that was sampled. The variability properties of the source (full population of random values or functions) generally cannot be accurately constructed from just a few samples of the population. Thus, when relatively few samples are available a substantial epistemic contribution of uncertainty exists in addition to the aleatory uncertainty due to stochastic variability in the source population.

The likely error that accompanies sparse sampling has a bias toward underestimation of the true full-population variability (see e.g. [10] - [12]). This is unconservative and therefore undesirable for many engineering purposes. If a system-level model were perfect in every other way, use of the constitutive model in system-level predictions of response variability (e.g. strength or displacement variability of support structures or pressure vessels) would underestimate the actual variability of the real systems. In design and risk analysis one would normally want to avoid underestimation of system variability. The following subsections summarize an approach intended to avoid underestimation of the larger-population variability. The procedure is more fully presented in [9] and [13] along with a failed attempt at a parametric representation of stress-strain curve variability and sparse-data uncertainty.

The following approach treats the stress-strain curves at a given temperature as discrete random functions with no readily identifiable parametric relationship between each other. Yet the approach recognizes that they issue from the same temperature-characteristic population of discrete random functions. Furthermore, because usually only a small number of experimental curves of behavior are available, the approach minimizes chances of underestimating the full-population variability with relatively few data samples. Figure 2.2 conveys the approach for a random variable, but the idea also applies to random functions as demonstrated next.

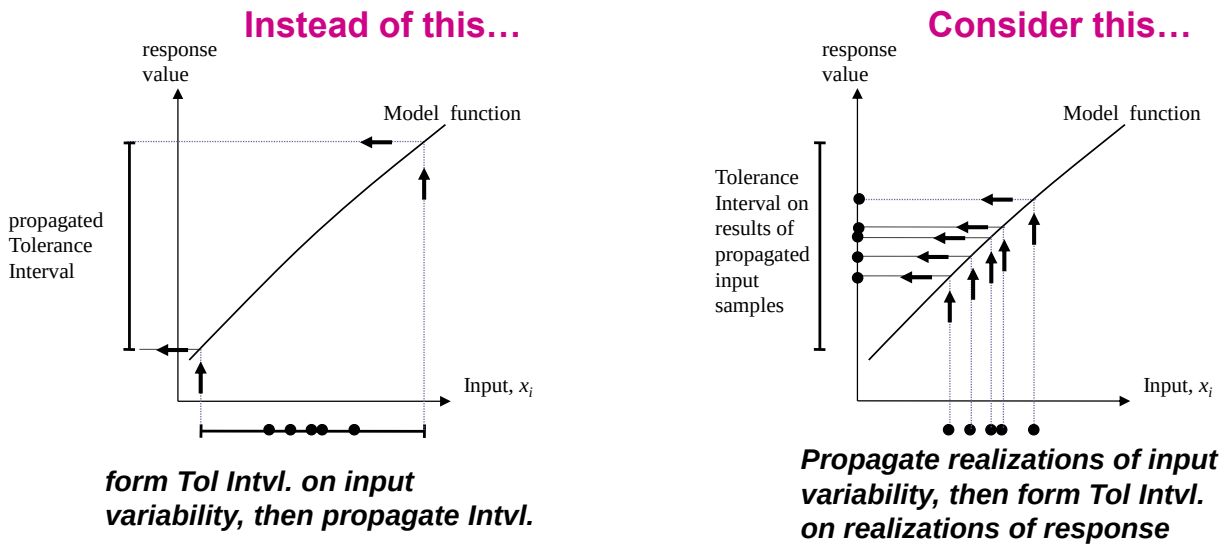

Figure 2.2 Two ways of propagating uncertainty from sparse samples of an input random variable. In the approach on the left the sparse samples of the input quantity are fit with a tolerance interval (explained below) that is then propagated to an uncertainty of response. (A more conventional manner of presentation is that an equivalent Normal PDF fitted to the tolerance interval is propagated.) The approach on the right individually propagates each sample of the input quantity and then forms a tolerance interval (and equivalent Normal PDF, see Figure 2.5) from the propagated results.

Employing the approach at right in Figure 2.2, the multiple stress-strain curves at a given temperature are individually propagated through the applicable system model (here the PB model) to yield corresponding samples of output quantities of response such as displacement, failure pressure, etc. PDFs of the response quantities are then constructed from the response samples for further analysis purposes as explained below.

This approach also accommodates propagation of model parameter sets that are discrete (are not parametrically continuous). For instance, in electronics modeling applications the Gummel-Poon (GP) model parameters (often 10 or more) are determined unique to each particular device tested. When multiple repeat tests on nominally identical devices are performed the resulting sets of GP parameters define different points in the parameter space. However, the parameter space is generally not considered to be continuous (e.g. [14]). The model is generally not "trusted" to yield suitable results when run with parameter values at other points in the space, e.g. at points interpolated on a line between any two established points/parameter-sets in the space. Hence device-to-device and other experimental variability resident in the discrete GP parameter sets can be treated by the paradigm illustrated at right in Figure 2.2.

Considering the samples of output response on the vertical axis at right in the figure, various approaches can be taken to compensate for small numbers of samples. Investigations were undertaken in [10]-[12]. It was found that a classical statistical Tolerance Interval approach provides a reliably conservative estimate of the combined epistemic and aleatory uncertainty associated with the limited data. The approach is also very easy to use. The approach [15] also worked well for sparse samples in many cases, but is somewhat more involved to implement and

its performance remains to be broadly tested and characterized because the method is very new. The rather common practice of simply fitting the random data with a normal distribution was found to be risky. It produces results skewed toward being non-conservative, especially if the sampled distribution is a Normal distribution.

For the purposes here the Tolerance Interval (TI) approach was used. The length of tolerance intervals accounts for both the epistemic and aleatory elements of uncertainty when limited data samples are available. Hence, TIs are characterized by two user-prescribed attainment levels: one for "coverage" of a subset of the variability, and one for statistical "confidence" in covering or bounding at least that subset of variability. For instance, a 0.95-coverage/0.90-confidence TI prescribes lower and upper values of a range of response that is said to have at least 90% odds that it covers or spans the 0.025 and 0.975 percentiles of the "true" probability distribution (or probability density function, PDF) from which the random samples were drawn (for a large array of PDF types).

A 0.95/0.90 TI is constructed by multiplying the calculated standard deviation $\tilde{\sigma}$ of the data samples by the following factors f in Table 2.1 to create an interval of total length 2f $\tilde{\sigma}$, where the interval is centered about the calculated mean $\tilde{\mu}$ of the samples. Table 2.1 and Figure 2.3 reveal that for a given PDF being sampled, the TI size decreases quickly with the number of data samples. For 0.95/0.90 TI a knee in the rate of uncertainty decrease per added sample occurs somewhere between 4 to 6 samples, with the rate of decrease being fairly small after 8 samples. The tolerance interval has an asymptotic standard-deviation multiplier of 1.96 for an infinite number of samples. This gives a TI that corresponds to the exact 0.95 central percentile range of a Normal PDF with μ and σ the same as $\tilde{\mu}$ and $\tilde{\sigma}$ from the ∞ samples. That is, the multipliers *f* for 0.95/0.90 TI are effectively constructed from randomly sampling a Normal PDF a number of times *M* in each of a large number of random trials and finding the multiplier f_M that gives TI which, in approximately 90% of the trials, span the true generating Normal PDF's 0.025 to 0.975 percentile range.

Table 2.1 0.95/0.90 Tolerance Interval Factors (standard deviation multipliers) vs. # of Samples of Random Quantity

# samples	$f_{0.95/0.90}$
2	18.80
3	6.92
4	4.94
5	4.15
6	3.72
8	3.26
12	2.86
20	2.56
30	2.41
40	2.33
∞	1.96

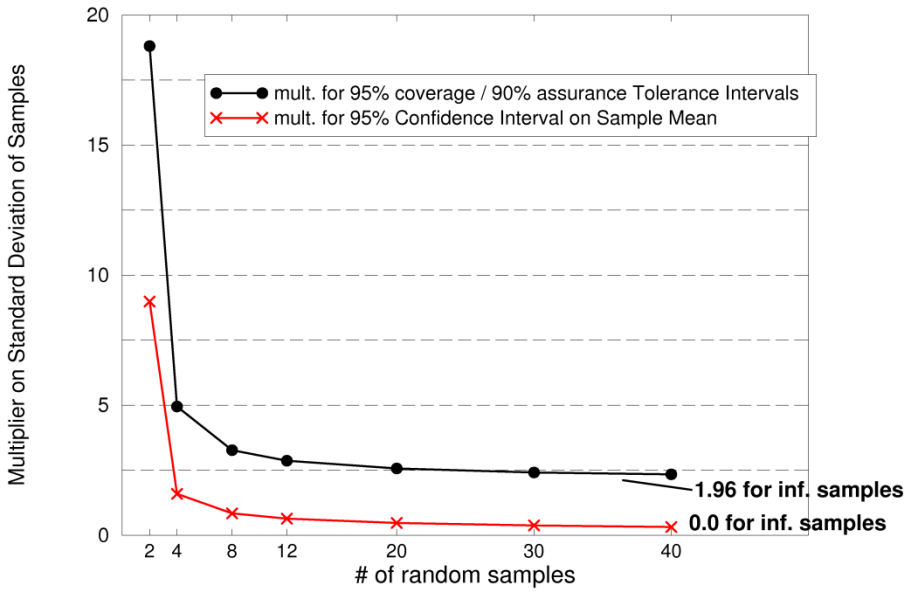

Figure 2.3 Multiplier on calculated standard deviation used to form uncertainty interval ranges for tolerance intervals associated with data samples. (Figure reproduced from [16], ignore confidence interval curve.)

Although constructed with respect to Normal PDFs, 0.95/0.90 TI will also span, with approximately 90% odds, the 0.025 to 0.975 percentile ranges of many other PDF types when sparsely sampled. This has been empirically established in [10]-[12] for uniform and right-triangular PDFs and for PDFs resulting from convolving various types of PDFs as depicted in Figure 2.4.

For subsequent uncertainty representation and analysis purposes a Normal PDF is constructed such that its 0.025 and 0.975 percentiles coincide with the end points of the established 0.95/0.90 Tolerance Interval (see Figure 2.5). The Normal PDF therefore has approximately 90% odds that its 0.025 and 0.975 percentiles contain the 0.025 and 0.975 percentiles of the "true" PDF from which the random samples come (for a large array of PDF types). Furthermore, because the constructed PDF is Normal, its characteristically long tails will have extended percentiles like 0.01 and 0.99 that in most cases extend beyond the same percentiles of the true PDF from which the data samples come. This was found in [11], [12] to be true for all tested PDF types and combinations in Figure 2.4.

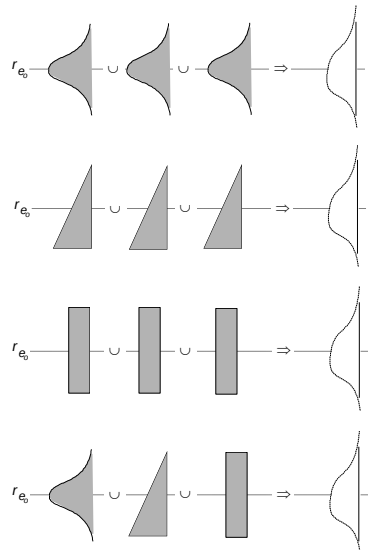

Figure 2.4. Test Matrix for PDF representation study in [11], [12].

The TI approach described here is presumed to likely (at the said odds) exaggerate the effects of the actual material variability. Such exaggeration can sometimes be egregious when very few samples are involved, depending on the particular samples obtained (see [10] – [12] for histograms of test results). The Pradlwarter-Schueller approach [15] has much smaller chances of egregious conservatism but averaged only 70% reliable in bracketing the true 0.025 to 0.975 percentile ranges of the PDF shapes and combinations shown in Figure 2.4 (compared to an average of 92% reliability for TI).

2.4 Results for Constant-Temperature Pipe Bombs

At each material characterization temperature in Figure 2.1 the derived constitutive model stress-strain curves are used in simulations of the pipe vessel held at 20C and pressurized at a linear ramp rate of 1psi/sec until pipe wall failure is indicated. (The rate of linear pressure rise is immaterial here because time does not impact these quasi-static failure pressures; creep is not an aspect of the current constitutive model and is not thought to be physically important here.) The pipe model, mesh and solver settings, pipe ends fixturing/loading conditions, simulations, and failure pressure results are described in [9]. The results at a given temperature are ordered by predicted failure pressures and therefore also by effective strength of the stress-strain curves at that temperature. Note that because structural response depends on the history or path that a given stress-strain curve entails, it is not clear a priori how the curves rank in effective strength in a given application until application-model simulations are run. The predicted failure pressures and derived tolerance intervals at the characterization temperatures are plotted in Figure 2.5.

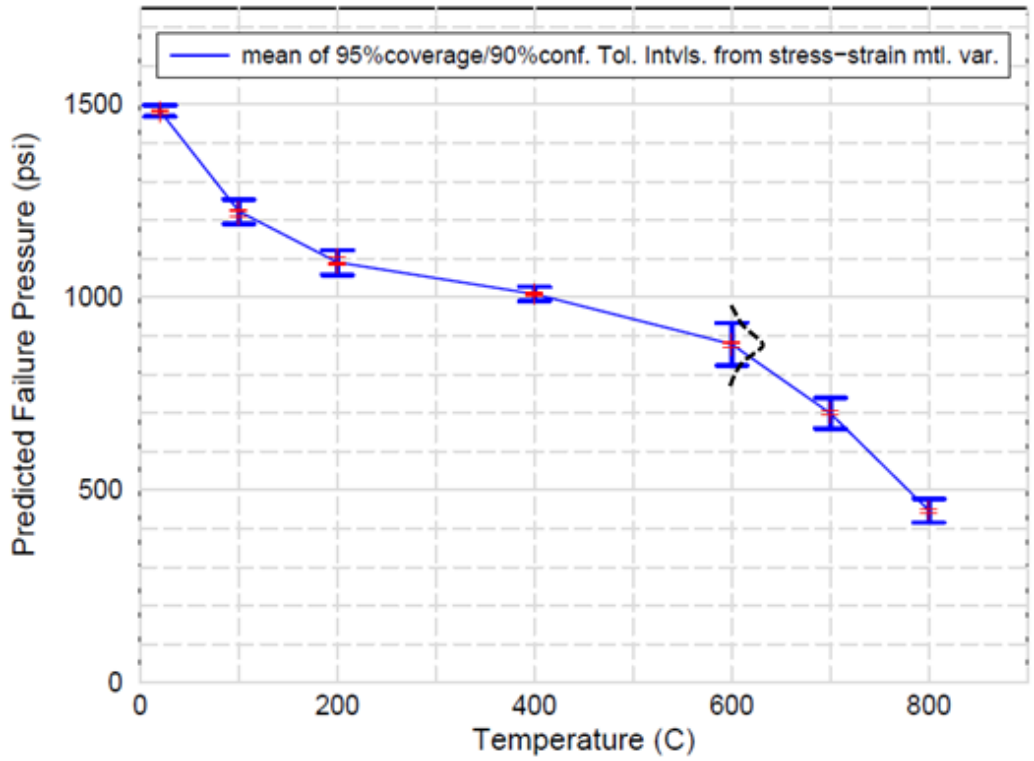

Figure 2.5 95% coverage / 90% confidence Tolerance Intervals from variation of pipe failure pressures at various pipe temperatures (for uniform temperature throughout pipe). Individual failure pressures predicted with the various stress-strain curves are plotted as red crosses. At 600C data an illustrative Normal PDF constructed from the 0.95/0.9 TI is depicted.

2.5 Extension to Other Pipe Bomb Application Conditions

In validation simulations with the PB model, differences exist vs. the initial and boundary conditions underlying the TIs in Figure 2.5, such as wall thickness, fixturing/loading conditions on the pipe ends, pressure loading history, and non-uniform pipe temperature. A two-part approach explained next is used to represent PB failure pressure variability under these different initial and boundary conditions.

Consider a pipe simulation with the same pressure and end-loading conditions underlying the results in Figure 2.5. But now let the pipe have a spatially uniform temperature that increases in time. Let the simulation start at the first characterization temperature (20C) in the material data set. When the pipe temperature rises from 20C to 100C, the stress-strain curves characterized at 20C gradually become less applicable and the stress-strain curves at the next characterization temperature of 100C gradually become more applicable until at 100C they are exclusively the applicable curves. This prompts a generalized question of how to “best” (best balance of effectiveness and economics) weight or transition between two sets of stress-strain curves from two adjacent characterization temperatures that bracket the local pipe temperature. The word ‘local’ is used here because the general problem involves alternatively or in combination a

temperature field that varies in space. In general, how does the simulation model represent material constitutive behavior for local pointwise pipe temperature (in time and space) that lies between adjacent characterization temperatures at which stress-strain curves exist?

If a single stress-strain curve exists at each bracketing temperature, then the procedure in [4] is employed to linearly interpolate the stress-strain state from the bracketing stress-strain curves, given the local temperature and the bracketing curve temperatures. However, when multiple stress-strain curves of material variability exist at each bracketing temperature, it must be decided how to handle this. The issues and our approach are discussed next.

Consider a PB simulation with a spatially uniform temperature that increases in time, from 20C to 200C. Figure 2.6 is an illustrative representation (not to scale) of the TIs and underlying PB failure-pressure data points for the first three material characterization temperatures in Figure 2.5. Each dot on each TI in Figure 2.6 corresponds to a particular stress-strain (s-s) curve. The temperature transient in the simulation can be negotiated with the interpolation scheme [4] after selecting one dot/s-s curve at each of the relevant characterization temperatures, 20C, 100C, and 200C. Several possible selection combinations are shown in Figure 2.6. Vastly more combinations exist. One could think about sampling all combinations ("exhaustive" sampling) and running a PB simulation for each combination. This would amount to 6x5x5 = 150 runs of the PB model. With the full set of material data there would be 20,250 combinations/runs for a simulation involving a uniform pipe temperature that increases from 20C to 800C. This is clearly unaffordable and is not necessary anyway in the present circumstances. Alternatively, a random sampling of say 30 to 50 random combinations could be performed. This might be affordable in some circumstances, but not for validation simulations to be described in Section 4.2. Hence the following two-run "bounding" approach was taken.

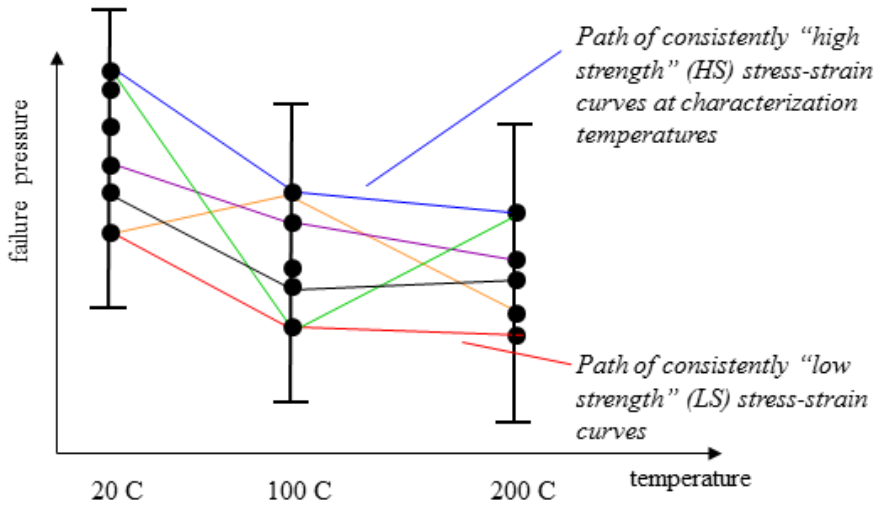

Figure 2.6 Illustration (not to scale) of the 0.95/0.9 Tolerance Intervals and underlying PB failure pressure data (for uniform pipe temperature) for the first three material characterization temperatures in Figure 2.5. Several possible combinations ("paths") of material stress-strain curves are shown that could be used in a computational simulation to negotiate a temperature transition over the depicted temperature range.

Consider a local region of material on the pipe. At a given characterization temperature, say 20C, this local region has particular stress-strain properties with an uncertainty of effective strength reflected by the six data samples at 20C in Figure 2.6. Consider a case where the effective strength of the local material lies near the highest dot at 20C. If the local material region undergoes a temperature transition from say 20C to 200C, it is physically plausible[3] that a local material region with a high relative effective strength at the starting temperature 20C will retain high relative effective strength as it transitions temperatures to 200C. Thus, a material realization with effective strength that starts in the neighborhood of the highest-strength dot at 20C will tend to correlate with the highest-strength dots at 100C and 200C. So the HS "high strength" combination path in Figure 2.6 is reasoned to be a highly physically reasonable combination. Similar reasoning is applied to the lowest dots at 20C, 100C, and 200C (LS "low strength" path).

Material realizations corresponding to mid-range dots at 20C in Figure 2.6 will also tend to correlate with the mid-range dots at 100C and 200C. But PB response in simulations using mid-range path combinations is assumed to be bounded by the consistently high strength (HS) and consistently low strength (LS) combinations. Therefore there is no need to expend computational resources investigating moderate paths.

We use similar reasoning explained in [9] for steady or transient spatially non-uniform temperature fields. But we acknowledge that a spatially stochastic uncertainty representation would be relevant here and may even be necessary for certain analysis requirements. We conclude similarly for spatially varying pipe wall thickness in section 4.3.

Hence, in view of the constraints in this project, material variability effects are parameterized in terms of just two simulation runs of the PB model. One simulation uses only the high strength stress-strain curves at each characterization temperature, using the methodology [4] to transition between the HS s-s curves. The other simulation uses only the low strength s-s curves at each characterization temperature. This parameterization is used in Section 4.2 to scale PB failure pressure variability TIs in Figure 2.5 to TIs for the validation application conditions.

[3] This assumes that material weakening mechanisms and % weakening are roughly similar with increasing temperature whether the material is initially of higher, medium, or lower relative strength.

3. Modeling and Design of the Pipe Bomb Validation Experiments to Minimize Error and Uncertainty in the Experiments and Validation Simulations

In accordance with model validation best practices (see e.g. [17], [18]) a model of the envisioned validation experiments was constructed and used to help refine the experiment conditions and the locations of measurement sensors to best support the model validation objectives. It was desired to test the temperature-dependent constitutive model at realistic pressure and temperature ramp rates and with large temperature gradients on the pipe and large ranges of spatial temperature variation. Coupled thermal-structural modeling (section 3.1) was used to help define the length of the pipe and the size of the heating shroud and its location relative to the pipe in order to achieve large spatial temperature variations. The model was also used to devise a thermocouple (TC) placement scheme suitable for the anticipated temperature field variations and mathematical interpolation scheme used to construct the temperature field boundary condition on the pipe surface (section 3.2). The temperature field reconstruction capability prompted an experimental adjustment to compensate for convection effects (section 3.4). Representative simulations were performed to test the temperature mapping and interpolation capability and characterize associated error and uncertainty (section 3.4).

Design of the pipe wall thickness and of the size of the internal slug for internal volume reduction (explosive energy reduction) were based on simplified handbook models and formulas. These latter design aspects were driven by experimental safety related to explosive rating limits of the test facility.

3.1 Coupled Thermal-Mechanical Modeling to help Design the Validation Experiments

The model included thermal radiation exchange between the heating shroud and the thermal-expansion affected pipe and the surroundings. Radiation exchange between the pipe and the internal slug (see Fig. 1.1) was also modeled, along with heat conduction along and through the pipe walls. Convection was ignored because of the difficulties mentioned in section 3.2. Nonetheless, the simplified thermal model was deemed sufficient to aid in the design and planning of the thermal aspects of the validation experiments. Figure 3.1 indicates that the temperature drops several hundred degrees with distance from the hot spot. Note the thermal expansion elongation and pressure-induced bulging of the pipe on the side facing the heating shroud. The temperature drop through the 0.07 inch pipe wall in this initial model was only about 1K at the hot spot. Therefore, when assigning temperature boundary conditions to the pipe walls as described in the next section it is sufficient to assign a constant temperature through the pipe wall thickness (to be that of the local surface temperature obtained from the interpolated thermocouple data).

Figure 3.1 provides a rough indication of the temperature pattern on the pipe in the planned experiments. This was useful in devising judicious locations of thermocouples (TCs) on the pipe surface for approximate reconstruction of the experimental temperature field. The thermocouple location and interpolation procedures are described next.

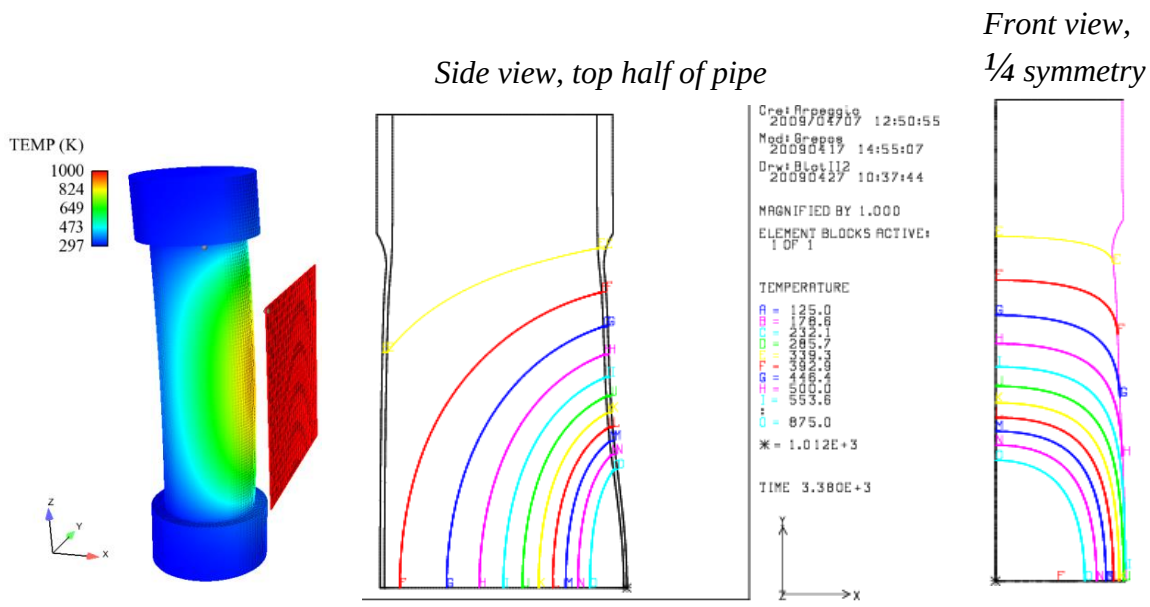

Figure 3.1 Coupled SIERRA/Thermal-Solid Mechanics simulation results for circumstances described in text.

3.2 Thermocouple Placement and Interpolation Scheme for Approximate Temperature Field Boundary Condition Reconstruction on Pipe Surface

Early in the design and planning of the validation activity it was judged from experience that modeling the thermal aspects of the experiments would contribute prohibitively large uncertainty to the validation study, thus unacceptably degrading the precision with which we could resolve how accurate the constitutive model is. In particular, for modeling purposes the emissivities of the radiating surfaces of the shroud, pipe, and slug inside it would not be known to within ±10% uncertainty even if coated with Pyromark paint for increased predictability of the emissivity. It was also estimated that modeling convective heat losses from the interior and exterior surfaces of the pipe could be in error by ±25% or more. The difficulty is compounded by the fact that pressurization of the pipe involved continually introducing new gas (mass) and its associated enthalpy into the pipe. These would not only affect the gas temperature inside the pipe in a complex way, but would also foreseeably affect convection inside the pipe. It was not even known if convection correlations that account for such introduced mass and/or the very high pressures in the experiment (several hundred atmospheres) even exist to be found.

Fortunately, error and uncertainty associated with the thermal aspects in the validation problem were mostly circumvented by measuring pipe surface temperatures at specifically designed thermocouple locations (based on model simulations) and then spatially interpolating to provide pipe wall temperature boundary conditions for the pipe structural response simulations. The procedures for thermocouple placement, temperature interpolation, and associated UQ are explained below. FE modeling of the planned experiments was central to these pursuits and also to analyze and help design the heating aspects of the experiments as already described.

An initial constraint in the strategy to reconstruct the experimental temperature field from temporally and spatially interpolated TC data was that at most 18 TCs were allowed per experiment because of limited data channel availability in the data acquisition system. Freedoms in the sensor placement and interpolation design-problem were that: 1) the 18 TCs could be placed anywhere on the pipe surface; and 2) any method of spatial interpolation could be used that was compatible with proposed sensor patterns and was expected to yield reasonably good interpolation accuracy and was within grasp of limited project time and resources.

Another consideration in the sensor placement and interpolation design was that the quantities we wanted to predict (failure time, pressure, temperature, etc.) were anticipated to be most affected by the hot spot region on the pipe. There the steel would experience greatest temperature-associated material strength loss, bulging, and tendency for failure. It was recognized, however, that the response prediction problem is a global one. Stress, strain, and ultimately failure at or near the hot spot is dependent on time-dependent material "give" everywhere else in the pipe. Therefore, temperature-dependent material relaxation needed to be modeled everywhere on the pipe.

Even though the global nature dictates that temperature-dependent material behavior effects should be modeled everywhere on the pipe, it is reasonable to assume that it is most important to model the effects most accurately in the hotter regions of the pipe, where most of the stress/strain/failure action is expected to occur. Accordingly, the temperature contour information in Fig. 3.1 was used as an indicator for relative concentration of TC sensor coverage over the pipe: greater concentration of TCs in the hotter regions, less in the cooler regions. Within this generalized objective, the locations of the individual TCs were determined as follows.

A Hermite 2D bi-cubic polynomial interpolation scheme was identified as something that could be quickly developed. The bi-cubic interpolating-shape freedoms of 2D Hermite polynomials could conform reasonably well to the temperature field over the surface of the pipe, when the field is subdivided into a suitable set of interpolation “patches” as shown in Fig. 3.2. The assemblage of all interpolation patches yields a C^0 interpolated temperature field that is continuous but not necessarily smooth across patch boundaries; i.e. temperature slope across patch boundaries is not necessarily continuous.

Generation of a Hermite bi-cubic interpolation function over a quadrilateral interpolation domain (patch) requires input of temperature and three slope-affiliated terms at each corner of the patch: T, $\partial T/\partial z$, $\partial T/\partial \theta$, $\partial^2 T/\partial z \partial \theta$. If the derivative information at the patch corners is from numerical approximation as employed here, the overall interpolation scheme is referred to as “Quasi”-Hermitian. The objective then becomes one of laying out a TC pattern that supports the best (most accurate) determination of temperature and the three derivatives at the corners of the patches in Fig. 3.2, assuming that it is important to get greatest accuracy in the vicinity of the hot spot and that accuracy importance drops as temperature drops with distance from the hot spot.

In Fig. 3.2 the θ-z "pipe-surface" coordinate system has its origin at the hot-spot at pipe mid-height. A preferential TC placement scheme increases the concentration of TCs as the hot spot location ($\theta = 0$, $z = 0$) is approached. This is shown to be roughly the case in the figure. The

location of the origin also coincides with an experimental pipe temperature distribution expected to be nominally symmetric about the vertical z-axis at $\theta = 0$, and to a lesser extent (because of convection) to be approximately symmetric about the horizontal θ-axis at $z = 0$. Hence, our test for interpolation accuracy and error (section 3.4) does not test non-symmetric conditions in either direction. Nonetheless, the TC placements were designed to enable temperature non-symmetry in either or both directions to be detected and to be approximately modeled by the interpolation scheme.

(***drawing and TC locations not to scale***)

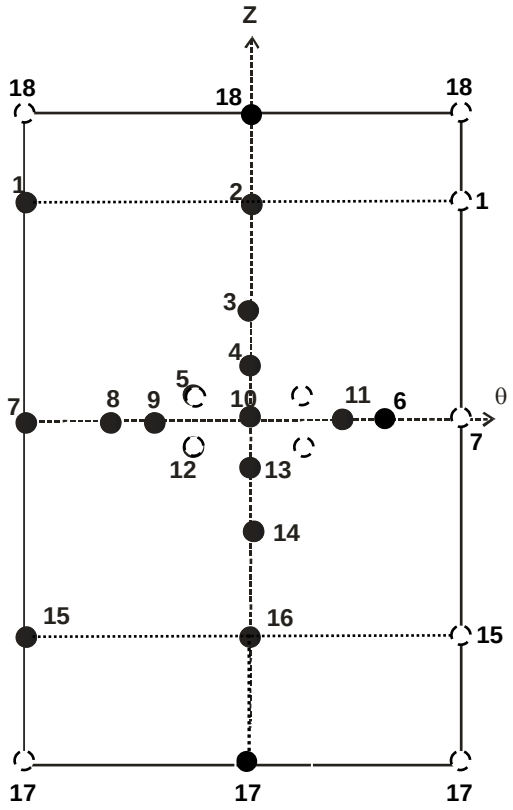

Figure 3.2. Division of pipe exterior surface into eight quadrilateral "interpolation patches". Surface of the pipe represented as though sliced at the back, unwrapped, and laid flat in the θ-z space shown. TC locations on pipe surface are shown (not drawn to scale). TC 10 marks the origin ($\theta = 0$, $z = 0$) of the θ-z "pipe-surface" coordinate system, where pipe circumference ranges from $-\pi < \theta \leq \pi$. Note that the TC numbering here is from convention used in the interpolation subroutine (see [9, Appendix C]) and <u>not</u> the ID numbering of TCs in the experiments (Fig. 3.3).

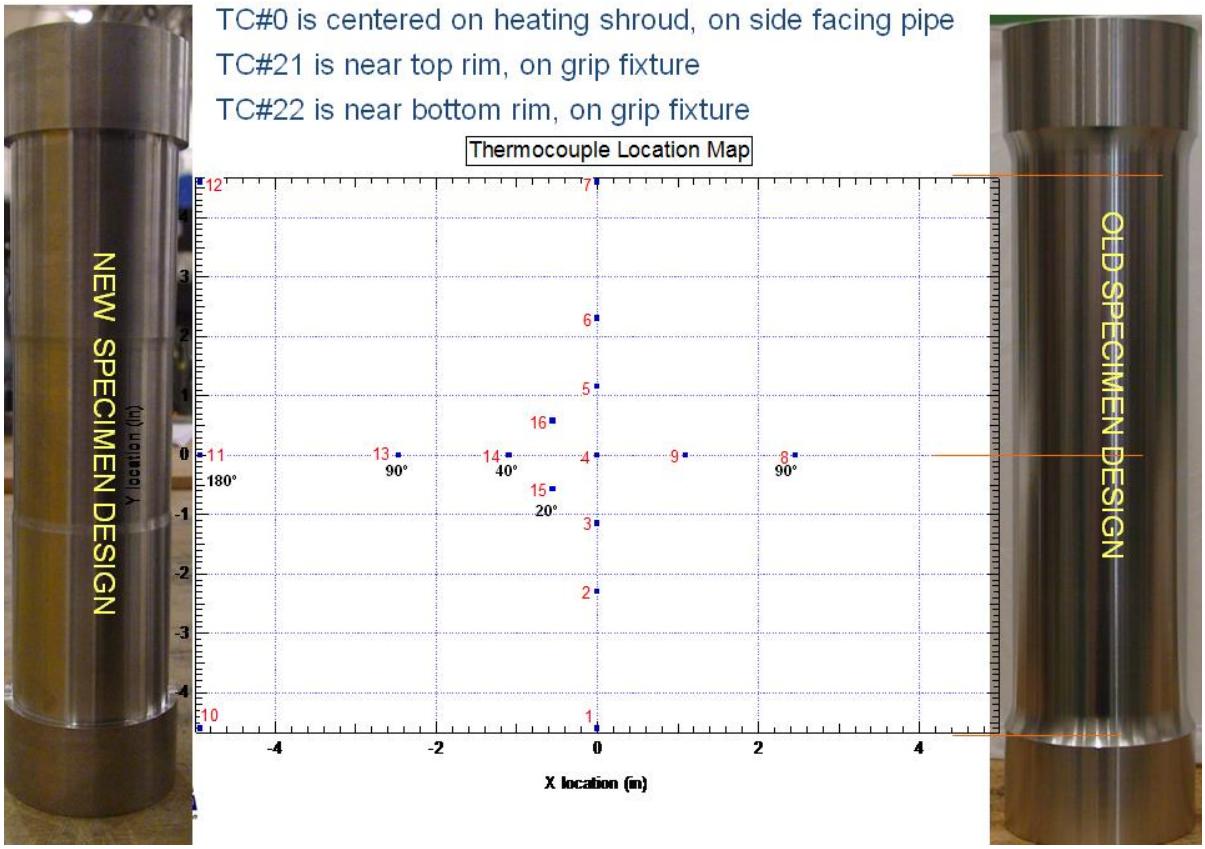

Figure 3.3 TC numbering used in the experiments, which is different from the interpolation code's numbering scheme shown in Fig. 3.2. The pipe shown on the right is an early specimen with uniform 0.05 in.-thick wall on reduced-diameter section of pipe. The pipe on the left is a later test specimen with a 3-inch long further-reduced section with 0.02 in. wall thickness. The thinner walls were adopted to reduce pressure levels at which the pipes failed (to meet test facility explosive hazard ratings).

The layout of TCs in Fig. 3.2 allows some of the derivatives to be evaluated non-uniquely, so any of several choices could be made. For example, $\partial T/\partial\theta$ at TC10 could be evaluated from a Lagrange cubic polynomial fitted to the temperatures of TCs 7,8,9,10 or TCs 8,9,10,11; or simply be set to zero according to a physical argument of C^1 smoothness (continuous slope across patch boundaries) + temperature field symmetry about the *z* axis. Testing each of these alternatives showed that interpolation error on the test problems to be described next was least for the first option. It was similarly found that setting $\partial T/\partial\theta$ at TC7 according to the option described in [9, Appendix C] gave less overall interpolation error than using a C^1 smoothness + symmetry condition $\partial T/\partial\theta = 0$ there.

Symmetry + C^1 smoothness conditions $\partial T/\partial\theta = 0$ at front and back of the pipe ($\theta = \pm\pi$) were also not used for patch corners at TCs 1, 2, 15, 16 although this would have been simpler for the interpolation code. Instead, the upper TCs 2,1 and lower TCs 16,15 were used to prescribe linear temperature drops with circumferential distance from front to back. It was not assessed whether this linear temperature decrease caused less overall interpolation error than would a nonlinear decrease from front to back that would accompany $\partial T/\partial\theta = 0$ specified at front and back. An

assessment was not performed because temperatures in this region vary only slightly around the pipe because its top and bottom rims are held to approximately 293K by active cooling from water jackets.

Two thermocouples (TCs 5 and 12 in Fig. 3.2) were placed to provide a check on interpolation accuracy by placing TCs that would not be used to construct the interpolated temperature field. It was figured that the accuracy check should best be conducted near where the hot spot was anticipated to be. Also, it was figured that one TC should be above the z=0 plane and one symmetrically below the z=0 plane so that information could be obtained on non-symmetric interpolation errors in the anticipated convection-induced non-symmetric temperature fields about the z=0 plane. It was envisioned that the information on interpolation error might be useful to inform uncertainty estimates in the model validation phase of the project. However, the very limited point information did not prove to be sufficient to make useful inferences about global interpolation error effects on computed failure pressures—the quantity of greatest validation interest. Therefore the strategy described in section 3.4 was used.

3.3 Use of Temperature Mapping/Interpolation Capability to Adjust Experiments to Compensate for Convection Effects

Figure 3.4 shows a tested pipe (0.05 in. uniform wall thickness) heated and pressurized to a bulging state and then depressurized before failure. The thermocouple leads can be seen on the tested pipe. The accompanying simulation results show a significantly non-symmetric temperature field with greater temperatures above pipe mid-height than below. The skewed temperature distribution was attributed to external and internal heat convection. The hottest point on the pipe was significantly above the center TC (#4 in Fig. 3.3). The experiment was therefore reconfigured to lower the heating plate by about ½ inch to move the hotspot approximately to the center TC. Thus, the temperature mapping and interpolation capability was directly useful in fine-tuning the heating configuration in the experiments.

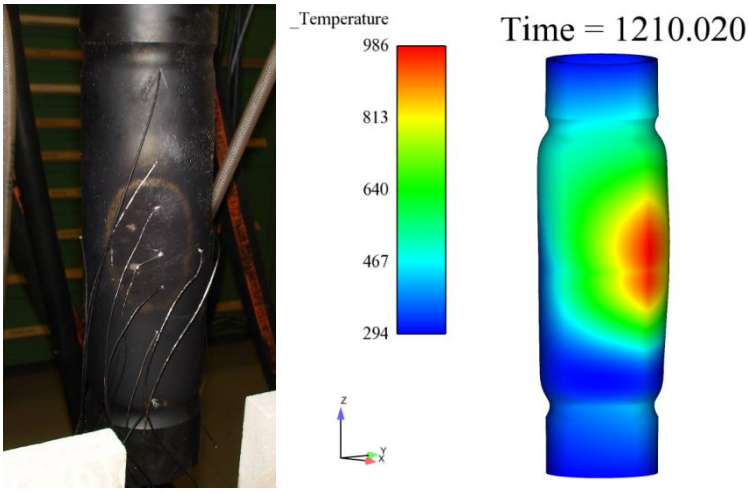

Figure 3.4 Convection-caused non-symmetric temperature field in early test setup. Heating shroud was subsequently lowered to move hot spot toward center thermocouple.

3.4 Simulations and Analysis to Test Temperature Mapping and Interpolation Procedures and Characterize associated Error and Uncertainty

Here we describe some relevant tests for the accuracy of the temperature interpolation and mapping procedure (TC data ⇨ temperature field BCs for model simulations) and the impact of interpolation error caused by sparseness of TC spatial coverage. We then describe how the results are leveraged to estimate a correction (with uncertainty) for interpolation-induced error in predicted pipe failure pressures in the validation simulations.

Such quantification of induced error can also be used to actively adjust/optimize the number of TCs, their locations on the pipe surface, and the interpolation scheme (Quasi-Hermite or others) to minimize the temperature BC reconstruction error and its effects on calculated failure pressure. This was done only at an informal level using expert judgment in the present project because of time and resource limitations.

3.4.1 Synthetic "Nearby Problem" representative of PB Validation Experiments

A synthetic "nearby problem" was constructed to closely emulate the tests that the model will be compared to (validated against) in Section 4. These tests, PB1 and PB4, undergo heating to produce a hot spot with a steady target temperature of 700C at the hottest point. Once this temperature is reached, the pipe is pressurized until failure. These will be referred to as the "700C-hold" PB1 and PB4 tests. See section 4.1 and [2] for further description of the test conditions.

To emulate the temperature field on the pipe from the 700C-hold experiments, a coupled thermal-mechanical model of the experiments was used (see Fig. 1.2). During pressurization the hot spot bulges outward because of temperature related weakening. The bulging brings the hot spot closer to the heating plate. During this process the viewfactors for radiative heat exchange on the inside and outside of the pipe must be updated after each timestep or a prescribed sufficiently small number of timesteps. This brings a challenging and expensive computational aspect to the coupled simulations. Fortunately the bi-symmetric problem allows a reduced 1/4 geometry to be used for the calculations in this section.

Convection is ignored in the model, so it predicts artificially high temperatures for a given heater plate temperature. Therefore the plate temperature from a different experiment (PB2) was used because it had a lower plate temperature which maintained the pipe hot-spot at 672C in the test. It was ventured that use of this lower plate temperature in a simulation would nominally offset the lack of convective cooling in the simulation such that the pipe temperature field in the simulation would be close to the temperature field in the 700C-hold experiments. We later show that this was indeed the case.

The pipe pressurization and heating-plate temperature inputs to the coupled thermal-mechanical simulations are plotted in Figure 3.5. The refined 4tt mesh and solver settings described in [9] were used in the following simulations.

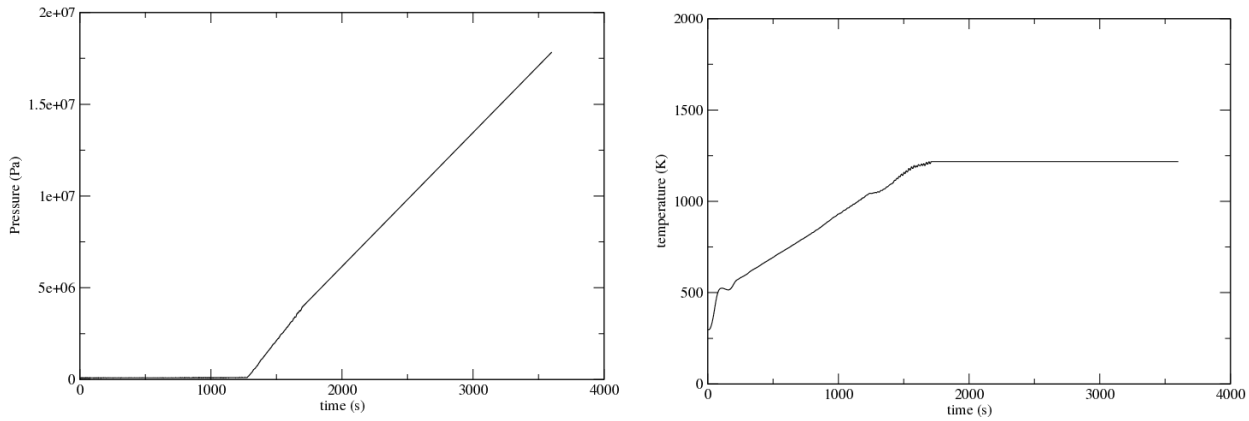

Figure 3.5 Test PB2 pressure and heating-plate temperature inputs to coupled thermal-mechanical simulations for testing and evaluating temperature mapping and interpolation procedure.

Two variants of the stainless-steel constitutive model were used in the simulations: stress-strain curves of high and low strength as explained in section 2.5. With these and with the radiative emissivity values described next, the coupled simulations yielded a representative set of time-developing spatial temperature fields on the pipe.

The following emissivities were nominal estimates from consulting various references in the literature. The emissivities are not used in the eventual validation simulations so their accuracy is not critical from this stand point. The values only need be representative enough to obtain temperature fields on the pipe that are close to the fields in the PB1 and PB4 experiments.

The outside surface of the solid slug inside the pipe (Figure 1.1) was assigned an emissivity of 0.4. The painted-black Inconel heating plate was assigned an emissivity of 0.7. Except for the heating plate, the surroundings that exchange radiation with the outside surface of the pipe were modeled with an emissivity of 1.0 and ambient temperature of 296 K. Emissivity of the outside and inside surfaces of the pipe is assigned a value stated in the simulation case description column in Table 3.1. The lower and upper emissivities of 0.7 and 0.84 in the table were deemed a reasonable range to represent the radiative heating uncertainty in the problem as it affects the temperature distribution on the pipe.

Table 3.1 Failure related quantities in coupled-simulation test cases.

Simulation case	Time @failure (sec.)	Temperature @failure (C)	Pressure @failure (psi)	Tear Param. @failure
Coupled-high-0.84	2277.1	759.0	809.9	6.23
Coupled-low-0.84	2254.9	759.0	791.6	4.40
Coupled-high-0.7	2386.0	706.9	903.7	3.9
Coupled-low-0.7	2350.4	706.8	873.0	3.50

The Table 3.1 combinations of high and low material strengths and high and low radiative heating give a representative range of temperature fields in the simulations to assess temperature

mapping and interpolation error effects on calculated failure pressures and other quantities. The hot-spot peak temperatures (at TC 10 location in Fig. 3.2) are essentially the same for high and low strength variants when emissivity = 0.84. Peak temperatures are also the same for high and low strength when emissivity = 0.7. Thus, peak temperatures (and the whole temperature fields, as discussed next) are essentially indifferent to changes between high and low strength material curves. But the temperature fields are significantly different when emissivity is changed from 0.7 to 0.84. Hot-spot temperatures in Table 3.1 are 52 C higher when emissivity = 0.84 than when emissivity = 0.7, whether high or low strength material is involved. Figure 3.6 shows the computed temperatures at other TC locations. Thus, the entire temperature field is hotter when emissivity = 0.84 than when emissivity = 0.7, as expected. Despite the linear visualization lines between TC temperatures in Figure 3.6, actual temperature variations between TCs are not linear. The actual variations are shown in the next subsection.

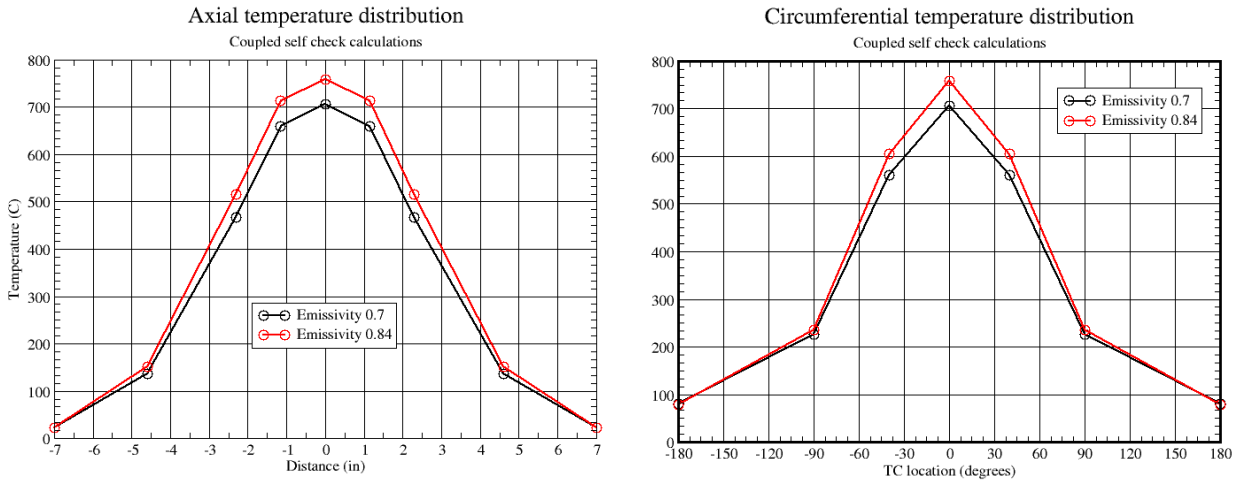

Figure 3.6. Coupled simulations temperature results (degrees C) plotted at thermocouple locations up and down front of pipe (left plot) and around pipe at mid-height (right plot). These locations correspond to TCs on *z* and θ axes in Figure 3.2 that the temperature interpolation scheme is built on. Temperatures are at time of failure in the indicated computations (the results for emissivity=0.84 essentially plot over each other regardless of whether high or low strength material curves are used, and likewise for emissivity=0.7). Temperatures at TC locations are spanned here by linear connecting segments for visualization purposes only and are not actual (see Figure 3.7).

Although changing from high to low strength negligibly impacts the temperature field, Table 3.1 reveals that failure pressure decreases by an average of ~3% (2.3% when emissivity=0.84 and 3.5% when emissivity=0.7). But the material strength variations affect failure pressure far less than changing emissivity from 0.7 to 0.84. This increase raises the temperature field, thereby reducing failure pressure by an average of ~11% (11.5% for high material strength and 10.3% for low strength).

3.4.2 Temperature Mapping and Interpolation Errors and Impact on Calculated Failure Pressures

Figure 3.7 shows, for a representative simulation case, the front and back views of the temperature fields of coupled and temperature-mapped/interpolated ("interp.") simulations at a common time, the time of failure in the coupled simulation. The interp. results are obtained from a mechanics-only Adagio simulation with prescribed temperature field BCs on the pipe surface. The temperature field is obtained using the mapping/interpolation methodology outlined in section 3.2 and synthetic temperature data (at the 18 TC locations) from the coupled simulations.

Noticeable differences exist between some areas of the coupled and interp. temperature fields in Figure 3.7. Figure 3.8 better reveals the spatial variation of error in the interpolated temperature field. Because the interpolated field is constructed from data at the TC locations, interpolation error at these locations is zero by default. Substantial error exists in many regions, especially where the temperature field varies quickly in space and simultaneously TC coverage is scarce.

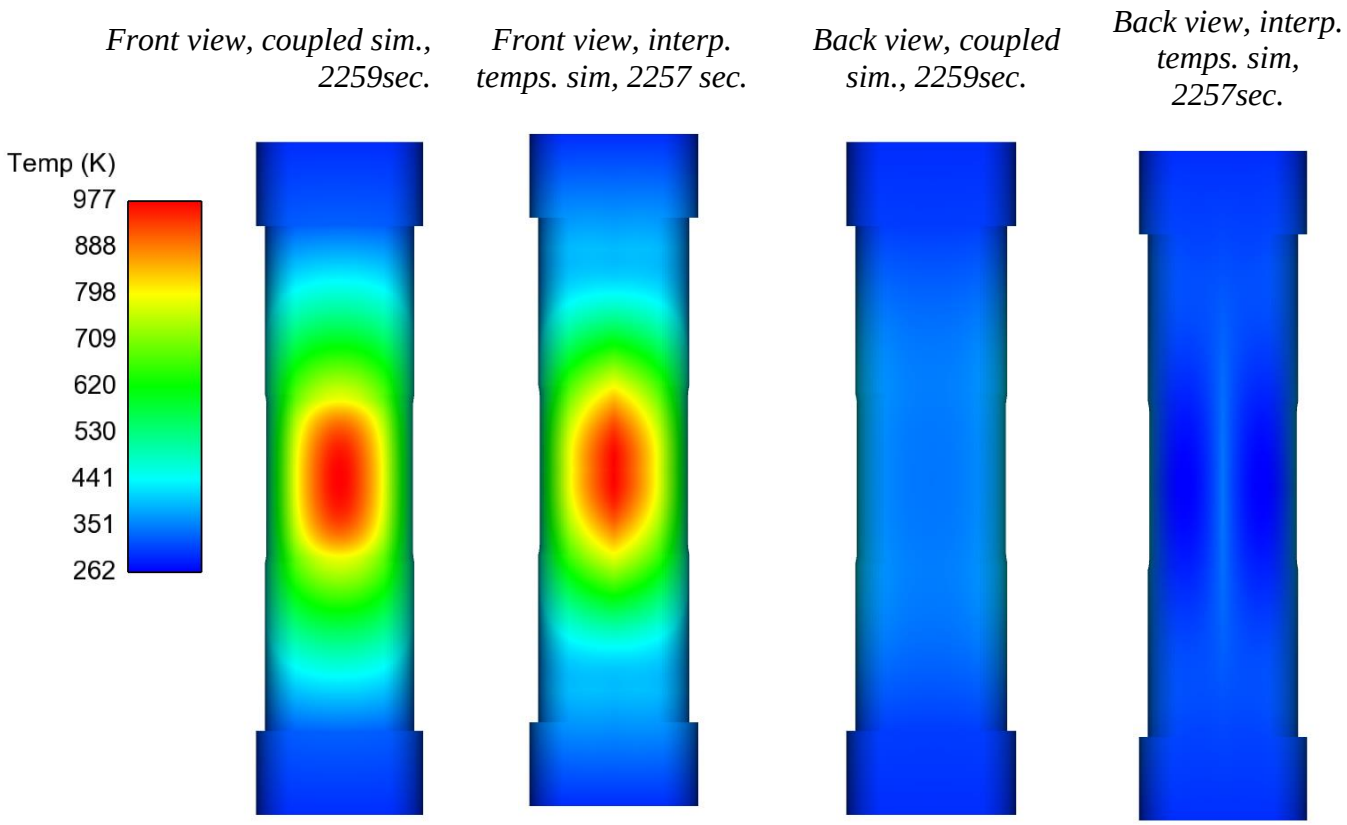

Figure 3.7 Reference and interpolated temperature fields for emissivity=0.7 and low strength.

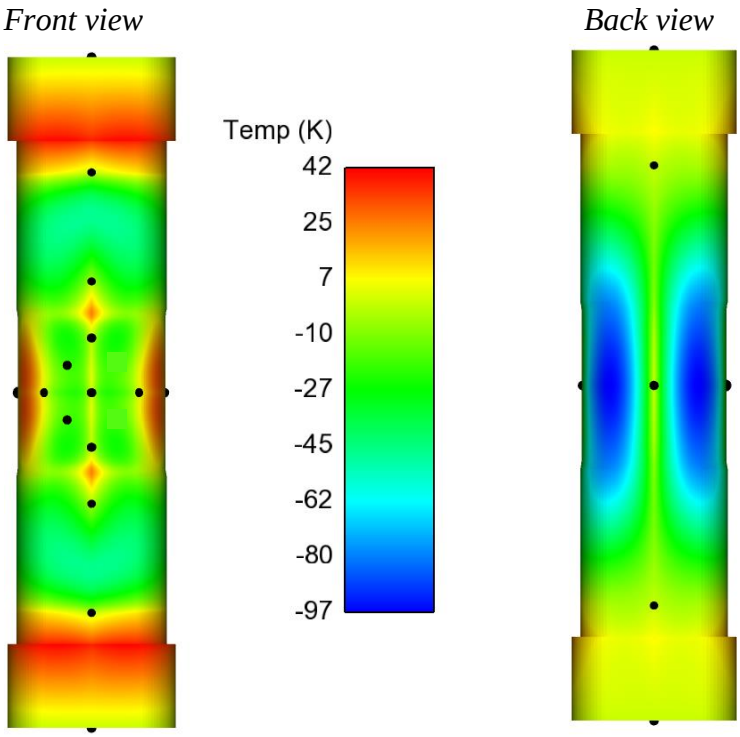

Figure 3.8 This difference plot of the interpolated minus coupled-sim. temperatures in Figure 3.7 quantities error in the interpolated temperatures at failure for emissivity=0.7 low-strength simulations. Zero interpolation error exists where yellow fades to green, in particular at the indicated TC locations (by construction of the interpolation scheme).

The interpolation error also varies over time. The integrated effect of the spatial-temporal interpolation error can be assessed by comparing the coupled sim. results in Table 3.1 to the interp. sim. results in Table 3.2. Various comparisons are discussed next.

Table 3.2 Failure related quantities in coupled simulations using mapped/interpolated temperature field BCs.

Simulation case	Time @failure (sec.)	Temperature @failure (C)	Pressure @failure (psi)	Tear Param. @failure
Interp-high-0.84	2332.8	758.9	857.9	7.27
Interp-low-0.84	2299.3	759.0	829.0	7.46
Interp-high-0.7	2400.9	706.9	916.2	4.43
Interp-low-0.7	2353.6	706.8	874.8	3.86

The hot-spot peak temperatures in Table 3.2 are effectively the same as in Table 3.1. The interpolation scheme ensures that the reference and interpolated temperature fields have the same temperature at all TC locations, including at the TC at the pipe hot spot. Therefore the trends

cited in the paragraph below Table 3.1 for hot-spot temperature vs. emissivity value and material strength also apply to the mapped temperatures here.

Figure 3.9 shows that the magnitude of material damage tracks with elevated temperature and failure occurs at the hot spot where the peak temperature exists. Damage contours and location of failure are shown to be essentially the same for the coupled simulation and the interp. simulation.

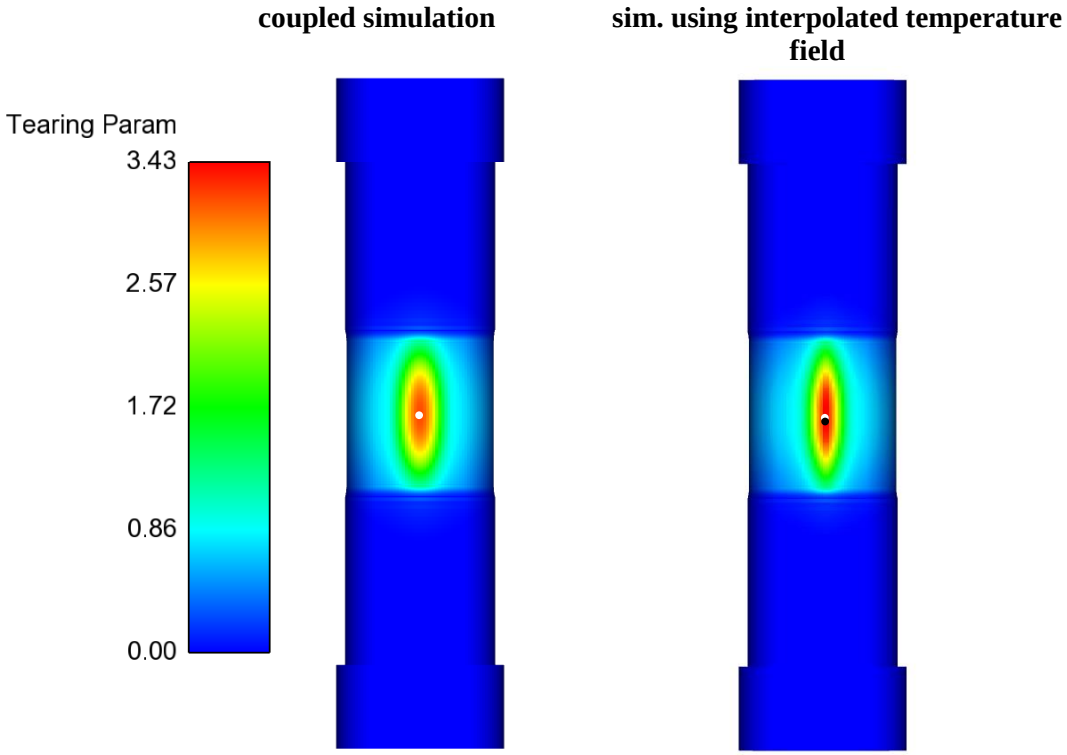

Figure 3.9 Spatial plots of computed tearing parameter (TP) indicating material damage at time of failure in emissivity=0.7 low-strength simulations. Damage is effectively confined to thin-wall section of pipe where the wall is 0.02 in. thick (wall is 0.05-in. thick elsewhere except at thicker shoulders). White dots locate maximum damage in coupled simulation. These are negligibly offset from pipe front-center where hottest point on pipe is. Black dot locates point of maximum damage in interp. simulations and is essentially the same location as failure in the coupled simulation.

Hence, both the coupled and the interp. simulations predict failure at the same location and at the same hot-spot temperature. Yet the failure times, pressures, etc. are different. Table 3.2 shows that failure occurs earlier and at lower pressures and tearing parameter levels in the coupled simulations than in the interp. simulations. The only difference between the coupled and interp. models is the temperature boundary condition. Even though the two models have the same temperatures at the 18 TC locations, spatial interpolation errors in the rest of the temperature field causes the interp. sim. output values to be different from those of the coupled simulations.

Nevertheless, the *trends* in output responses vs. changes in emissivity and material strength are in the same *directions* as those for the coupled calculations cited in the paragraph below Figure 3.6).

- Changing from high to low strength negligibly impacts the interpolated temperature field.
- Changing from high to low strength reduces failure pressures in Table 3.2 by an average of ~4% (3.5% when emissivity=0.84 and 4.6% when emissivity=0.7). These impacts are in the same direction as in the coupled simulations, but somewhat larger.
- Increasing emissivity from 0.7 to 0.84 raises the temperature field thereby reducing reduces failure pressure by an average of ~6% (6.8% at high strength and 4.6% at low strength). These decreases are in the same direction as in the coupled simulations, but significantly less.
- These emissivity changes affect failure pressure significantly more than these material strength variations.

Failure pressure is the quantity of interest for the validation comparisons in Section 4. Accordingly, Table 3.3 lists the failure pressure errors and percent errors caused by temperature interpolation for the four test cases. The errors are always positive; failure pressures in the interp. simulations are higher than in the coupled simulations by 0.3% to 4.9%. The errors are much larger for the higher-temperature 0.86 emissivity cases than for the lower temperature 0.7 emissivity cases, regardless of whether high or low strength material curves are involved. Errors are slightly greater for high strength curves than for low strength curves, regardless of whether emissivity is 0.7 or 0.84.

Table 3.3 Over-estimation of failure pressures caused by interpolation of temperature field for synthetic test cases.

sim. case	△ = P_fail_interp - P_fail_coupled	△% = △/ P_fail_coupled
high-0.84	48.0 psi	4.9 %
low-0.84	37.4 psi	4.7 %
high-0.7	12.5 psi	1.4 %
low-0.7	2.8 psi	0.3 %

3.4.3 Application to Model Validation Simulations to Correct Calculated Failure Pressures (with uncertainty on the correction)

In the previous subsection, error in predicted failure pressures was characterized for pipe temperature boundary conditions constructed from limited temperature data at the discrete thermocouple locations. For actual temperature fields that closely resemble the synthetic fields in Figures 3.6 and 3.7, the characterized pressure prediction errors in Table 3.3 will approximately apply. Figure 3.10 compares the synthetic temperature fields against the actual temperature fields at time of failure in the validation tests PB1 and PB4. The temperature fields are similar enough that the results in Table 3.3 are used to estimate failure pressure prediction errors and associated corrections (with uncertainty) in Section 4.2 due to reconstruction of the pipe temperature BCs from the tests' TC data.

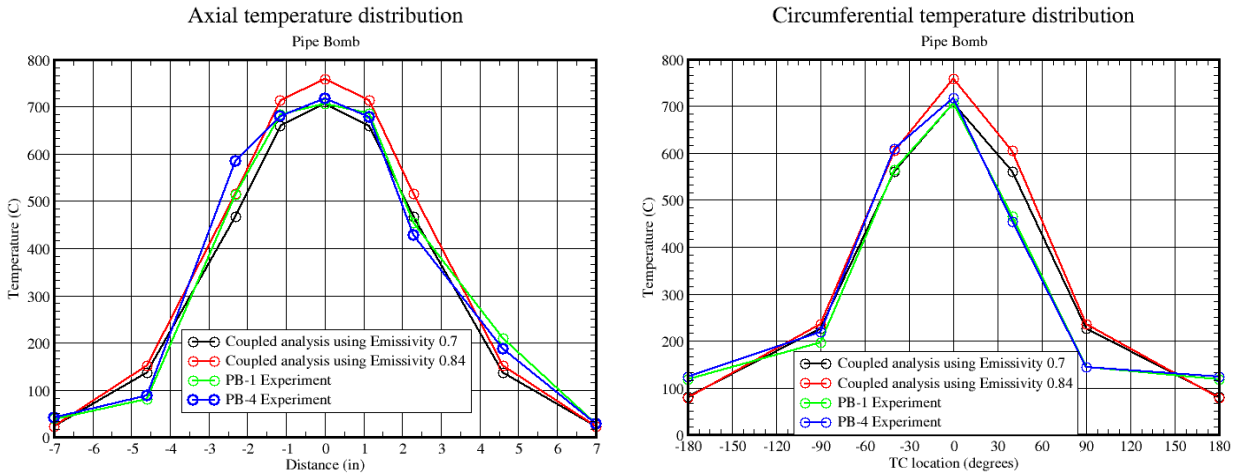

Figure 3.10. Pipe-bomb TC temperatures at time of failure in the 700C-hold experiments PB1 (green curves) and PB4 (red curves) plotted along with synthetic termperature fields (red and black curves) from Figure 3.6.

In the region where failure occurs (at or near the hot spot), the PB1 and PB4 experimental temperature profiles track closely with the synthetic temperature profiles (Figure 3.10). For the emissivity=0.7 temperature field, the failure pressure prediction error due to temperature field reconstruction from the sparse TC information is 2.8 psi for simulations with low-strength stress-strain curves and 12.5 psi for high-strength simulations (Table 3.3). These errors increase by about 35 psi to 37.4 (LS) or 48 psi (HS) for the emissivity=0.84 temperature field. The 35 psi change in the magnitude of interpolation-induced error apparently comes from the shape differences of the two temperature profiles in Figure 3.6. So the magnitude of interpolation-induced error is apparently fairly sensitive to the temperature distribution being interpolated. This is apparently a symptom of the spatial sparseness of the TC data available for intepolation. The interpolation error characterized here is used in Section 4.2 to estimate error and uncertainty associated with interpolating the actual TC data from the pipe temperature fields in the experiments.

4. Validation Experiments, Results, and Processing for Comparison to Model Predictions

In this section the constitutive model's performance is assessed by comparing experimental and predicted failure pressures of steel pipes heated to 700C and pressurized to failure. We present the experiments and simulations; their results and uncertainties; processing of these into a form suitable for the Real Space model validation methodology employed; and description of the comparisons along with their interpretation. The versatile and practical Real Space methodology has previously been applied to other complex calibration and validation problems in other physics realms: device thermal response and failure [19], [20]; modeled behavior of irradiated electronics [21], [22]; and combustion in fluids and solids [23], [24]. The PB model validation problem also has a large set of challenging features. Hence the methodology demonstrated here can be leveraged to a large set of model validation applications.

Various fidelities of uncertainty treatment can be applied with the Real Space validation framework. The particular UQ treatment applied here was driven by severe constraints in the number of simulations that could be completed in the time available. Only five simulations of the model were required in the linear decoupled UQ approach applied here. Each of the simulations took several weeks on 800 processors. That was at the limit of what was feasible under our resource constraints. Nevertheless, we judge that if allowed hundreds of simulations and higher-order UQ procedures, the main conclusions arrived at here would not be substantially different.

4.1 Experimental Inputs, Outputs, and Uncertainty in Failure Pressure Tests at 700C

In the following we consider the 700C-hold set of pipe pressurization experiments PB1 and PB4. The tests were planned replicates of each other. Figure 4.1 shows the transient internal absolute pressures in the pipes and the transient responses of temperature control TCs presumed to lie at the peaks of the hot spots in the two tests (i.e., TC#10 location in Figure 4.2). Later we attempt to account for uncertainties regarding TC accuracy and whether the TCs actually lie at the peak temperatures on the hot spots. The peak hot-spot temperatures (as indicated by the control TCs) were ramped at a rate of approximately 31 C/min.

When the control TCs reached nominally 700C, this temperature was maintained while the pipes were pressurized as shown in Figure 4.1. Controlled pressurization of the pipes was accomplished via pressure supply tanks as described in [2]. PB4 pressurization started about 1.5 minutes after 700C was reached, while PB1 pressurization started about 4 minutes after 700C was reached. Minor differences also exist about the nominal pressure ramp rate of 1.3 psi/sec. The differences in pressurization, were, at the outset of this project, anticipated to affect failure pressure levels negligibly. However, the experience of project participants concerning stainless steel behaviors at high pressures and especially high temperatures is very limited. Indeed, the PB1 pipe spent a few minutes longer at 700C before pressurization than PB4 did, and failed at a lower pressure. The lower failure pressure could conceivably be explained at least partially by the pipe soaking at 700C longer than the PB4 pipe. But the PB1 lower failure pressure appears to be at least partially explained by PB1 vs. PB4 wall-thickness differences and temperature distribution differences over the pipes, as will be establish later. Hence, neglibible effects of pressurization differences between the PB1 and PB4 tests are assumed in the following. But this

assumption should be reexamined as further experiences and model development contribute to better understanding of material response/failure dependencies in this regime.

In the following subsections, details of hardware geometry and experimental conditions and results are presented that are pertinent to subsequent model validation procedures and comparisons.

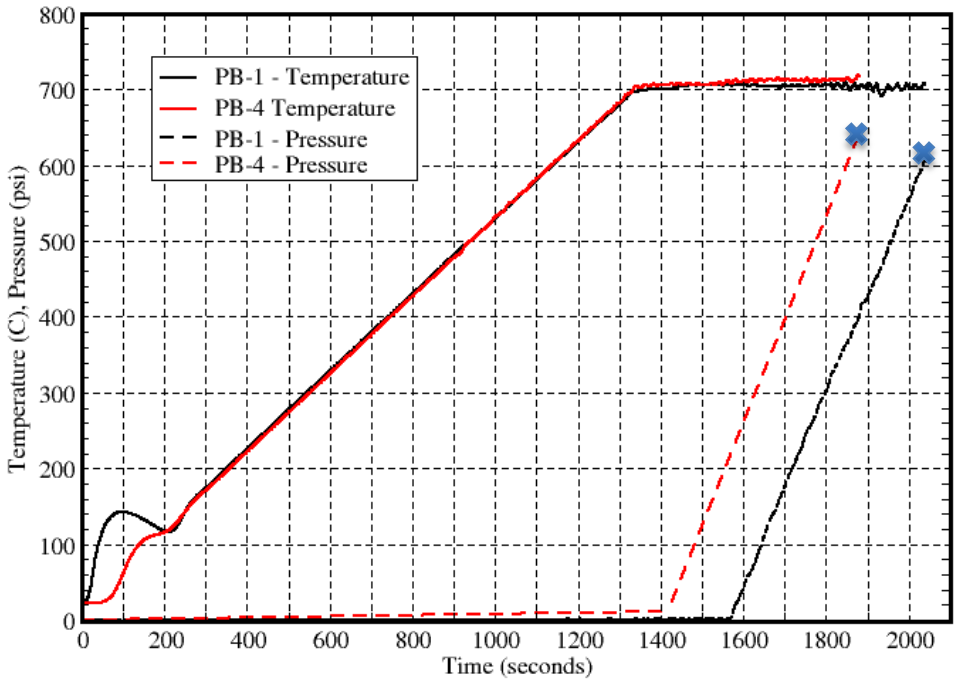

Figure 4.1 Test PB1 and PB4 transient internal pressures (absolute) and control TC temperatures at the hot spots. Pipes fail at the pressures indicated by the X's and are listed in Table 4.1.

Table 4.1 Failure pressures measured in 700C-hold experiments. Pressure measurement uncertainties are explained in Section 4.1.1.

Exper.	Measured pressure at failure	Uncertainty in pressure measurement
PB1	606 psi	± 10 psi
PB4	655 psi	± 10 psi

4.1.1 Experimental Pressure Loading and Fixture Loading BCs

Figure 4.2 replots the pressure loadings and shows the pipe axial loading profiles in the two experiments. The precipitous drops in the pressure readings indicate the violent-rupture depressurization in the tests. The manufacturer-specified uncertainty on the measured pressures is +/- 0.2% of the full-scale rating of the pressure gages used, which were 5000 psi max. gages. Therefore the uncertainty is within +/- 10 psi of the measured pressures in Figure 4.2. This amounts to about double the line thickness in the plots.

The uncertainty on the axial loading measurements is ±0.1% of reading, or < ±3 lb. for the highest axial load in the two tests. This is far less than the line thickness in the plots. Early investigations with the PB model indicated that this magnitude of uncertainty has negligible effect on failure pressures, so load measurement uncertainty is not considered further in this paper.

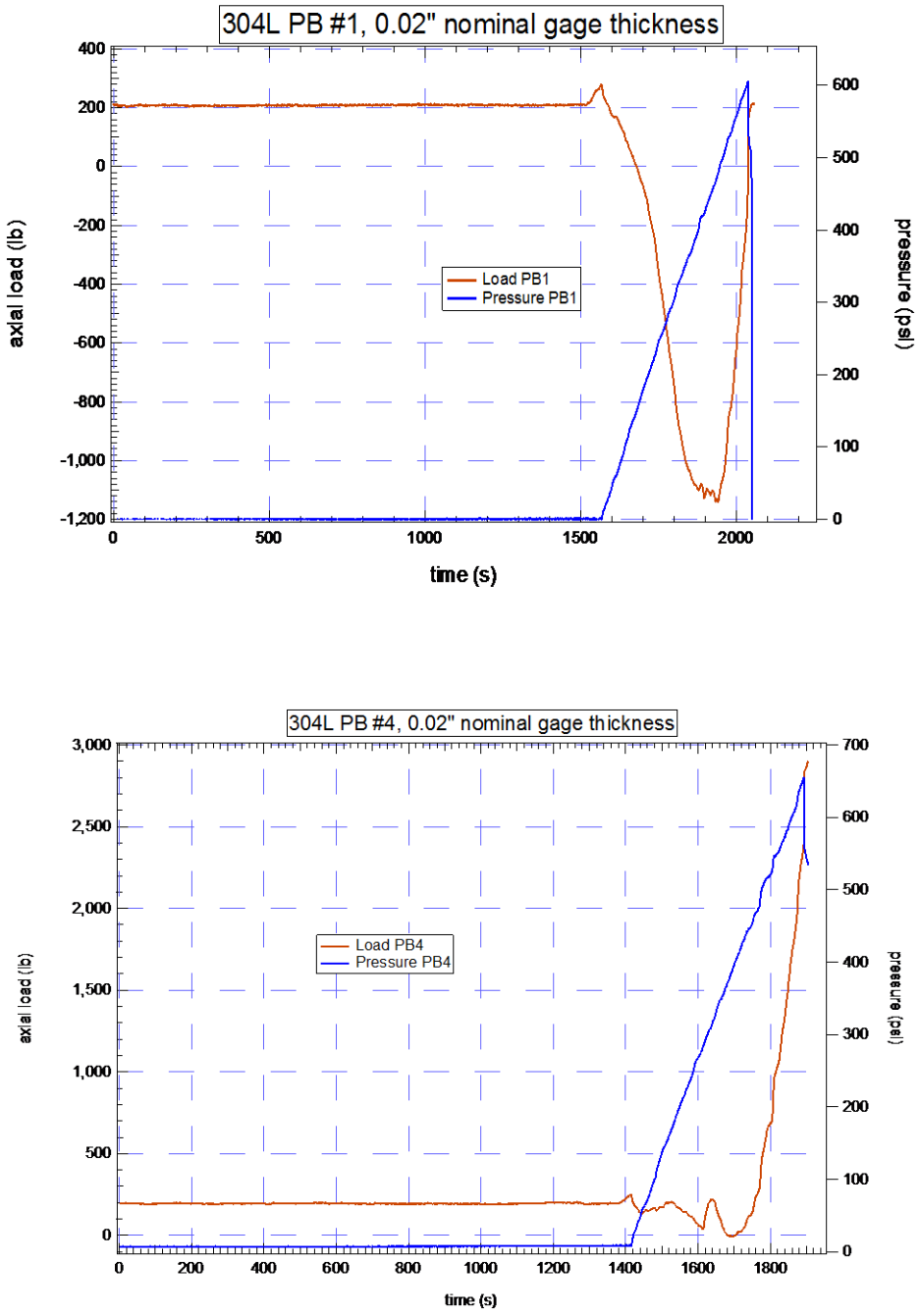

Figure 4.2 Measured pressure (referenced to scale on right axis), and axial fixture loading (referenced to scale on left axis) in 700C-hold pressurization experiments. Note similar pressure load scales in the two plots, but very different axial loading scales in the two plots.

4.1.2 Pipe Surface Temperature Measurements (TC Mapping/Interpolation Data)

Figures 4.3 and 4.4 plot the thermocouple temperature histories in tests PB1 and PB4. These TC data are supplied to the TC mapping/interpolation procedure described in section 4.2 for reconstructing the temperature fields on the pipe surface in the experiments. These reconstructed fields are the temperature BCs in simulations of the experiments (next section). The precipitous drops in the temperature data coincide with the pipe failures.

Table 4.2 lists linear regression values of the measured temperatures at the pipe front-center location, nominally at the pipe hot-spot temperature peaks (TC#4 in Figures 4.3, 4.3 and 4.4). Uncertainty of TC locations relative to the actual peak temperature locations will be addressed in section 4.3. The temperatures in Table 4.2 are obtained by linear regression of TC temperatures over the 60 seconds preceding failure in each experiment. Over the last several minutes in each test, electrical noise creates visible oscillations in TC temperatures in Figures 4.3 and 4.4. The oscillations over the last 60 seconds are on the order of ±10C, so citing a single temperature record at the time of failure could be in error by up to ±10C. Therefore we use linear regression mean-temperature curves evaluated at the times of failure. We do this for all TCs. The simulations must be carried out beyond the failure times in the tests because the simulations predict failures at higher pressures and longer times than the actual failures. No temperature data exists beyond the failure times in the tests, so at all TC locations we evaluate the linear regression mean curves at the times of failure. We then treat these temperatures as constant for the remainder of the simulations.

References [25], [26] discuss the various sources of thermocouple temperature measurement uncertainty applicable here. The PB experiments used intrinsic TCs of 0.005-inch diameter, which are the most accurate TCs commonly available. But small measurement uncertainties exist due to random and systematic sources of error described in Section 4.3 in relation to Equations 4.9 and 4.10. The largest uncertainties are indicated in Table 4.2, corresponding to the hottest TCs on the pipes. These uncertainties are less than the line thickness in the plots.

Table 4.2 Pipe front-center thermocouple temperature at failure in 700C-hold experiments

Exper.	Measured temperature at failure (after regression)	% uncertainty in temperature measurement	% uncertainty in temperature measurement
PB 1	707 C	[-0.25%, +0.5%] of reading in degrees C	[-1.8, +3.5] C
PB 4	711 C	[-0.25%, +0.5%] of reading in degrees C	[-1.8, +3.6] C

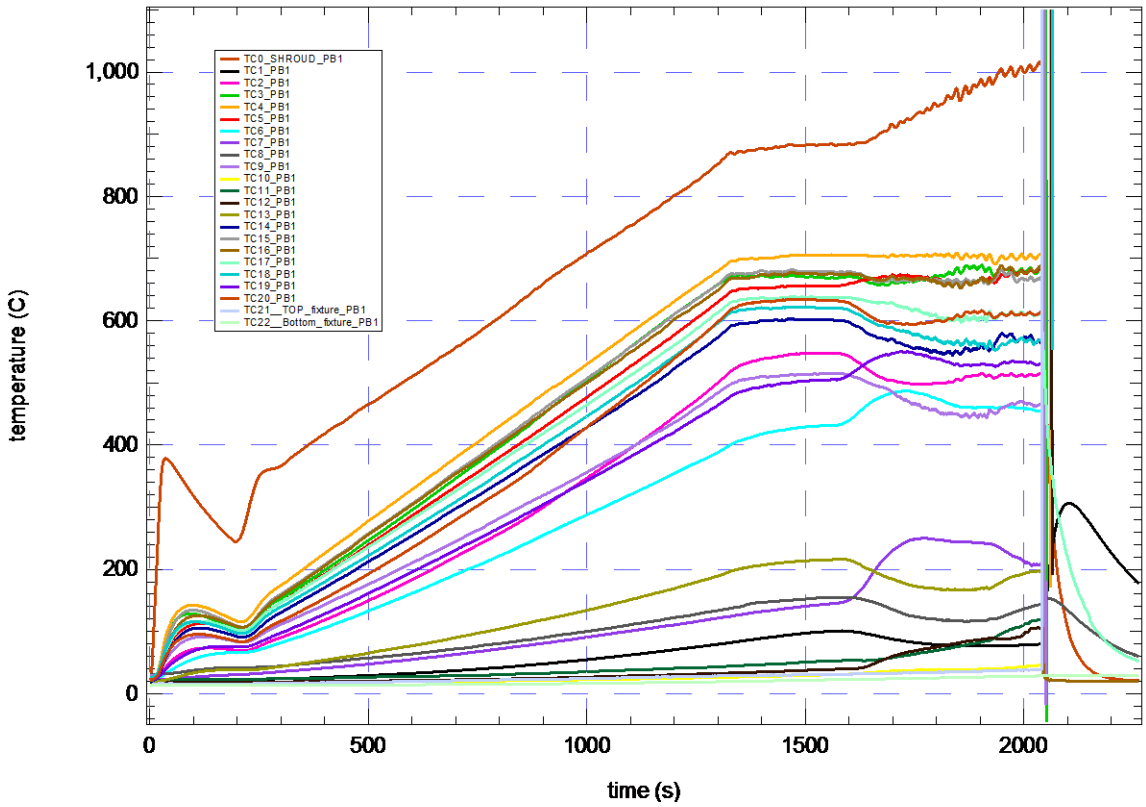

Figure 4.3 Thermocouple temperature measurements in experiment PB1. TC numbering in this figure corresponds to numbered locations in Figure 3.3.

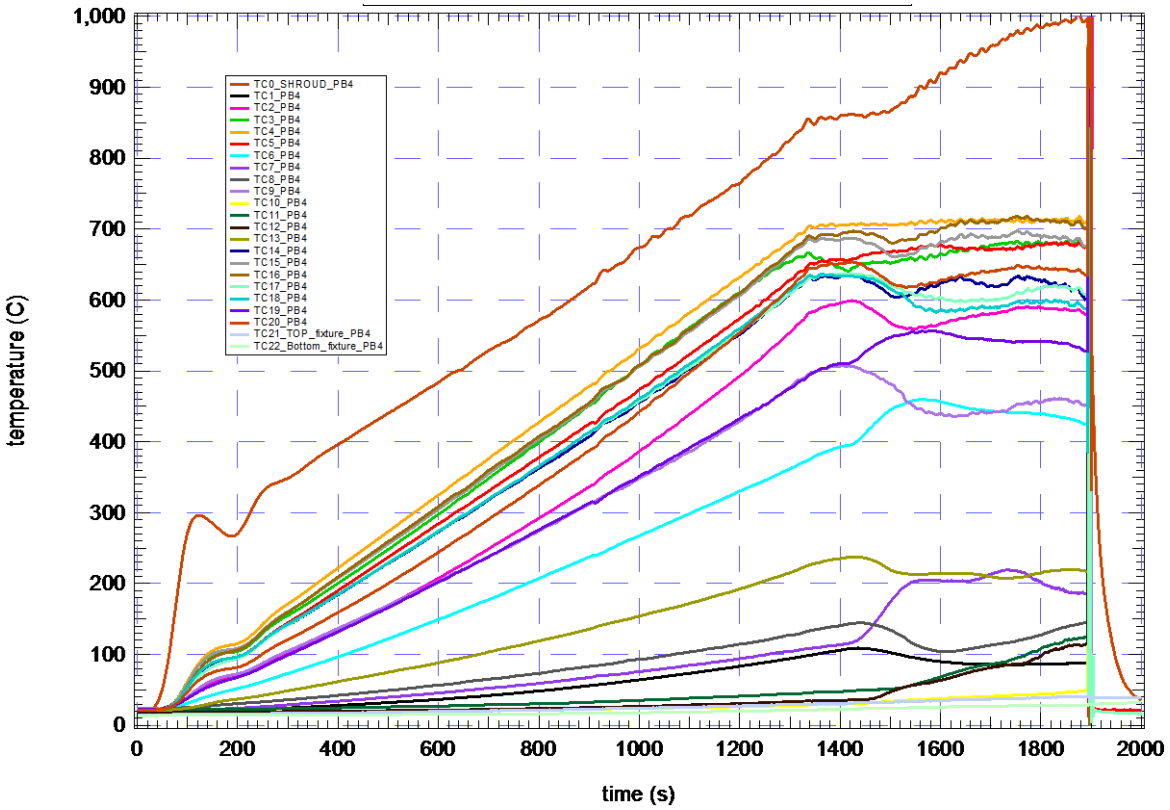

Figure 4.4 Thermocouple temperature measurements in experiment PB4. TC numbering in this figure corresponds to numbered locations in Figure 3.3.

4.1.3 Pipe Wall Thickness Variation

Table 4.3 shows measured minimum and maximum pipe-wall thicknesses for various PB tests, including PB1 and PB4. The measurement procedure was the following. Two V-blocks where set up to support a mandrel. A dial indicator was then used and zeroed on the mandrel. This would be the reference for the pipe inside diameter, ID. A pipe test unit was then placed on the mandrel and the dial indicator was placed on the outside diameter (OD) of the reduced section. The dial indicator was then moved along the pipe and the readings were taken. The specimen was then rotated on the mandrel and another set of measurements were taken. This was repeated 15-20 times and the minimum and maximum wall thicknesses were recorded in Table 4.3. This was done for each pipe unit.

The machined units held fairly tightly to the nominal 0.02-inch wall thickness specified for the reduced middle length of pipe. All tested pipe units were within the allowable tolerances specified on the drawings. PB4 seems slightly skewed to thicker walls than the other units. This might explain at least some of the reason why PB4 had a notably higher failure pressure than PB1, even though PB1's control TC reads slightly hotter than PB4's control TC. These issues are investigated in detail in section 4.3.

Table 4.3 Measured minimum and maximum pipe-wall thicknesses for various PB tests/'Parts' specified. Nominal machining spec. thickness was 0.02-inch.

Part No.	Wall Thickness
1	.019/.022"
2	.020/.022"
4	.020/.024"
5	.020/.022"
7	.019/.022"
8	.020/.021"
9	.019/.0225"
10	.019/.021"

4.2 Simulations of Validation Experiments and Uncertainty Processing & Rollup for Model Validation Comparisons

In this section we address random and systematic uncertainties intrinsically affiliated with the PB model and its computational solution. In the next section, 4.3, we address random and systematic uncertainties issuing from the PB experiments.

Here we use the PB1 experiment as the reference case for the model validation assessment at the 700C-hold conditions. The validation comparisons of experimental results vs. simulation results will be built around this reference case. We could alternatively use PB4 as the reference. Alternatively we could use averaged experimental conditions and inputs from the PB1 and PB4 replicate experiments, but such averaging is difficult here and in most real model validation applications. In quantifying the modeling and computational uncertainties in this subsection, we assume that these uncertainties would be effectively the same for simulation of the PB4 test. We reason that this a good assumption because the two tests are relatively small perturbations from each other. Unfortunately this assumption could not be checked because of the extreme computational time and cost of running more simulations.

The time-varying PB1 temperature, fixture loading, and pressure boundary conditions plotted in Figures 4.3 and 4.2 (top plot) are applied to a full-geometry PB model. A nominal wall thickness of 0.02" in the thin section of the pipe is used, along with the nominal 0.25" wall thickness of the upper and lower thick sections of the pipe. Both high-strength (HS) stress-strain curves and low-strength (LS) variants of the model were run, as described in Notes 1 and 3 below.

A coarse 1tt mesh was used because of the extreme computation demands of this problem. Corrections for mesh size (with associated uncertainty) are described later in this subsection, working from the mesh refinement study in [9]. The full model was first tried with a 4tt mesh and then a 2tt mesh, but even with 1200 processors the progress rates of these simulations indicated months of run time would be necessary to get through the multiple simulations required for uncertainty quantification (UQ). Because UQ/validation studies routinely take two or three iterations to refine and complete the analysis, the run times for 2tt and 4tt meshes were too long

for the project schedule. The 1tt-mesh runs took about four weeks (involving four or five restarts) on several hundred processors, so this was just within workable limits. The solver settings described in [9] were used. These settings were also used in the mesh refinement study and in the TC temperature mapping/interpolation error/uncertainty characterization in section 3.4.3.

The full geometry was necessary because of non-negligible effects of using the non-symmetric temperature distribution in PB1 (see Figure 4.6, PB1 experiment) vs. the ¼ symmetric mirrored PB1 temperature BCs used in the mesh study. Figure 3.5 shows the ¼ geometry section used for the ¼ model. The ¼ model and the full model have the same temperature distribution over their common quarter-region section, but the ¼ model assumes symmetry of the temperature distribution about the ¼ model's bottom and side boundaries such that it emulates a full model with ¼ symmetric temperature BCs. The failure pressure for the symmetric temperature distribution (¼ model, 1tt mesh) is 55 psi or ~ 5% higher than the result from the 1tt-mesh full model with non-symmetric temperature distribution. (The high-strength stress-strain curves were used in this investigation.)

The results and uncertainties associated with the PB1 reference prediction are summarized in Figure 4.4. Explanatory notes:

1. The top left corner in Figure 4.5 shows the result from a simulation with the high-strength (HS) stress-strain curves at each characterization temperature (see section 2.5).

2. A correction for use of a 1tt mesh is made from the mesh convergence study in [9]. From the study, corrections of -228 and -336 to the failure pressure calculated on the 1tt-mesh yield upper and lower bounds within which the asymptotic grid-converged failure pressure is estimated to lie.

   Hence we estimate that the asymptotic grid-converged failure pressure lies within the range [856, 748] psi = 1084 psi (the 1tt-mesh result) + [-228, -336] psi. These upper and lower values define the left-most uncertainty bar in Figure 4.4. For convenience this uncertainty will be incorporated in a later step. In the present step, only a nominal reference value within this range is selected. A value midway between the upper and lower extremes might be the most common choice, but we select 839 psi and demonstrate that this works just as well. Thus we parameterize the corrected range [856, 748] psi in terms of a nominal corrected value 839 psi plus an uncertainty range [+17, -91] psi about the nominal 839 psi. These numbers are written to the left of the left-most uncertainty bar in Figure 4.4. Another way of expressing what we have done is to correct the 1tt result (1084 psi) by a nominal correction of -245 psi plus an uncertainty of [+17, -91] psi about the nominal correction: -245 psi + [+17, -91] = [-228, -336] psi. The nominal correction of -245 psi yields the nominal corrected value of 839 psi.

3. Running the model with low-strength (LS) stress-strain curves instead of high-strength curves, the calculated failure pressure drops by 43 psi, from 1084 psi to 1041 psi. Applying the nominal correction (from step 2) of -245 psi to the LS result yields a

nominal mesh-corrected value of 1041 psi - 245 psi = 796 psi. This nominal correction assumes that the mesh convergence behavior is similar whether the model has LS or HS stress-strain curves. The nominal corrected LS value of 796 psi coincides with the bottom of the 2nd uncertainty bar from the left in Figure 4.4. This bar, spanning the LS and HS results, has a length of 43 psi as labeled in the figure. The LS and HS results parameterize the aleatory variability of predicted failure pressures due to stochastic variations in material strength as characterized from the cylinder tension tests in section 2.

4. Before the aleatory variability of failure pressure can be predicted, the LS and HS failure pressures must be corrected for errors caused by reconstruction of the pipe temperature field BCs from the PB1 temperature data. A "global" element of temperature reconstruction error is addressed here. Local elements are addressed in section 4.3. Here we consider the reconstruction errors characterized in section 3.4.2 for the symmetric temperature distributions shown in Figures 4.6 and 4.7. These distributions are from a simulated "nearby" problem that is close to the actual PB1 tests from the perspective of closeness of the PB1 temperature distribution to the synthetic temperature distributions (see Figure 4.10). If the experimental temperature distributions in Figure 4.10 were symmetric, and TC4 were located at the actual peak of the hot-spot in the experiments, there would be failure pressure prediction error from temperature field interpolation error due to spatial sparseness of the TCs. We refer to this as global error in the temperature field reconstruction. Table 3.3 lists potential global error magnitudes.

   We argue that this global component of interpolation-induced error is approximately bounded by the results in Table 3.3 because the actual PB1 temperature field being interpolated is very close in shape to, and is bounded by, the fields that produced the error results in Table 3.3. Specifically, at the times of failure in the experiments and similations, Figure 4.10 shows that the temperature BCs for the PB1 simulations are bounded between the $\varepsilon = 0.7$ and $\varepsilon = 0.84$ synthetic temperature fields in a critical region including at least 1-inch above and below the hot-spot TC and up to 40 degrees away in at least one circumferential direction. The interpolation-induced "bounding" errors for the $\varepsilon = 0.7$ and 0.84 "bounding" temperature distributions are respectively +12.5 psi and +48 psi if the simulations are performed with the high-strength stress-strain curves. The "bounding" errors are respectively +2.5 psi and +37.4 psi if the simulations are carried out with the low-strength curves.

5. Thus, we treat the actual PB1 temperature field as ***effectively*** lying somewhere within a parametric continuum between the $\varepsilon = 0.7$ and 0.84 "bounding" temperature distributions. That is, it is assumed that there is some emissivity value $0.7 < \varepsilon\_PB1 < 0.84$ that yields a global interpolation error equaling that for the actual PB1 temperature field. Then the limiting cases are:

if ε_PB1 = **0.7**, the correction for the global interpolation error is:
= -12.5 psi if the simulations are performed with the HS curves,
= - 2.5 psi if the simulations are performed with the LS curves;

if ε_PB1 = **0.84**, the correction for the global interpolation error is:
= -48 psi if the simulations are performed with the HS curves,
= - 37.4 psi and -48 psi if the simulations are performed with the LS curves.

For the limiting case **ε_PB1 = 0.7,** corrections to the HS and LS simulation results in Figure 4.5 are indicated by the green lines that end at the tops of the yellow and teal uncertainty bars, respectively. For these corrected pressure predictions of 827 psi and 793 psi, respectively, Figure 4.6 shows the corresponding tolerance interval and Normal PDF of aleatory variability due to material strength variability characterized in section 2.

For the limiting case **ε_PB1 = 0.84,** corrections to the HS and LS simulation results in Figure 4.5 are indicated by the brown lines that end at the bottoms of the yellow and teal uncertainty bars, respectively. The corrected pressure predictions are 791 psi and 756 psi, respectively. Figure 4.7 shows the constructed tolerance interval and Normal PDF for the limiting case ε_PB1 = 0.84.

The Normal PDFs for the two bounding cases ε_PB1 = 0.7 and ε_PB1 = 0.84 are shown in Figure 4.4. A intermediate value of the effective emissivity ε_PB1 corresponds to a Normal PDF commensurately interpolated between the upper and lower bounding PDFs in the figure. As noted in the figure, the yellow and teal uncertainty bars for corrected predicted failure pressures are perfectly correlated uncertainties both parametrically tied to the uncertain value of the effective emissivity ε_PB1.

6. Last, the uncertainty of the correction of mesh effects is incorporated. In Figure 4.5 the mesh correction uncertainty is shown translated to the right-most uncertainty bar in the figure. To aggregate this source of uncertainty with the uncertainty represented by the upper and lower PDFs defined in Note 5, linearity and independence are assumed. Essentially, the approximation is made that the upper and lower PDFs, which are built off of the starting reference value of 839 psi marked by the dot on the left-most uncertainty bar, would translate up or down as a grouped unit exactly with the reference value (the dot) if it was instead somewhere else on the uncertainty bar. Thus the mean separation and the standard deviations of the two PDFs would not change, but the PDF assembly would simply translate up and down with the reference value (dot). This is probably a relatively good approximation, but again could not be checked.

   Given the assumed linearity and independence, if the reference value (dot) was at the top of the mesh correction uncertainty bar, the PDF assembly would be translated upwards by 17 psi. Then the upper PDF in the upward translated PDF assembly would have a mean increased by 17 psi, making its mean (805psi + 17 psi) = 822 psi. This is shown at right in Figure 4.4.

At the other extreme, the existing reference value of 839 psi could be translated downwards by 91 psi. Then the lower PDF in the downward translated PDF assembly would have a mean decreased by 91 psi, making its mean (768 psi - 91 psi) = 677 psi. This is shown at right in Figure 4.4.

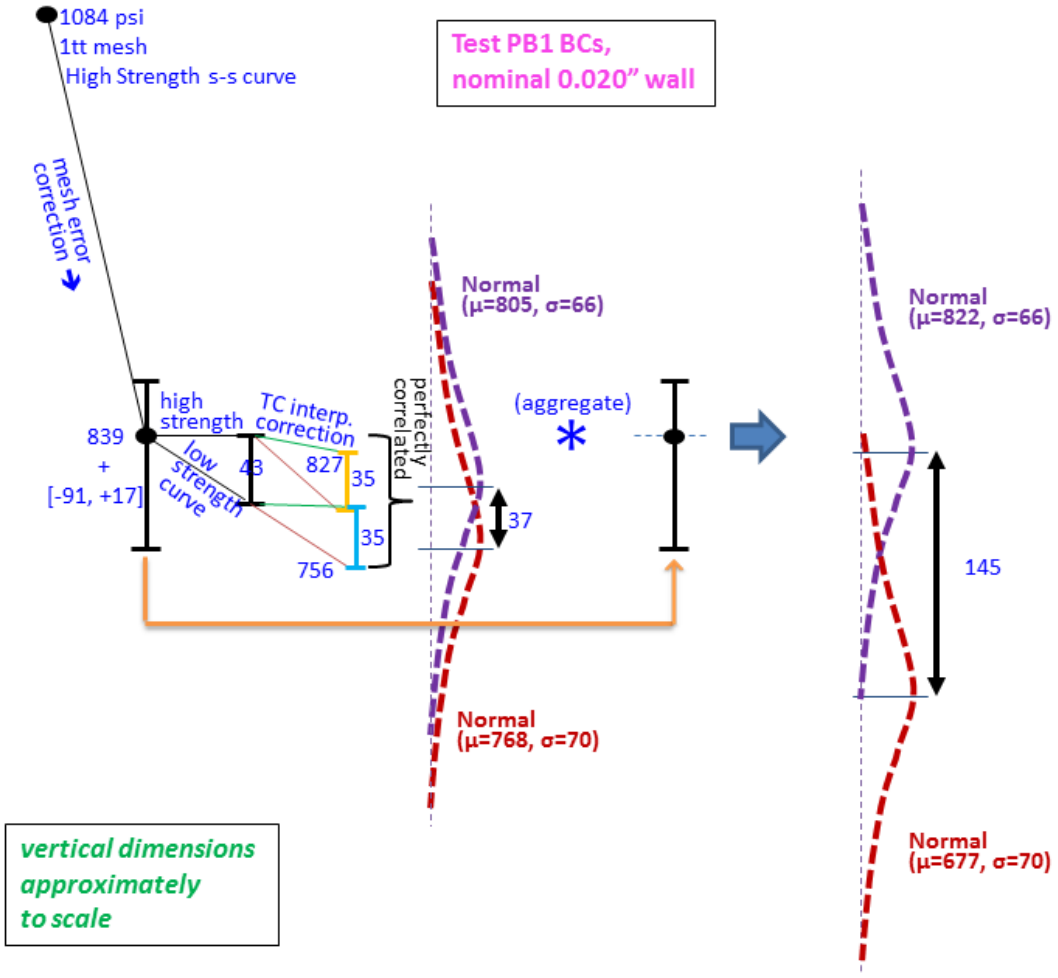

Figure 4.5 Uncertainty rollup for simulation results for validation comparisons.

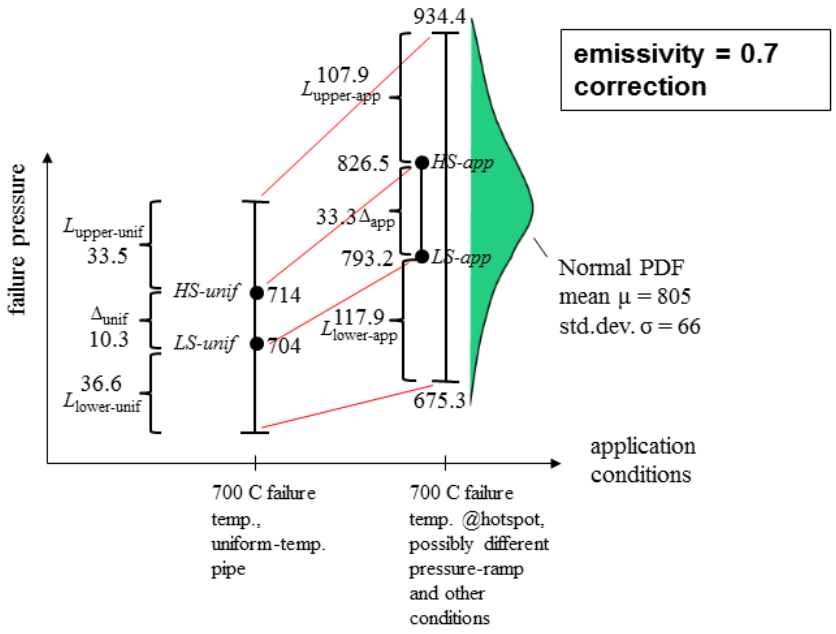

Figure 4.6 Scaling of 700C uniform-temperature 95/90 tolerance interval to approximate 95/90 TI and associated Normal PDF of predicted failure pressures. (For PB1experimental conditions and ε = 0.7 correction for temperature field reconstruction error.)

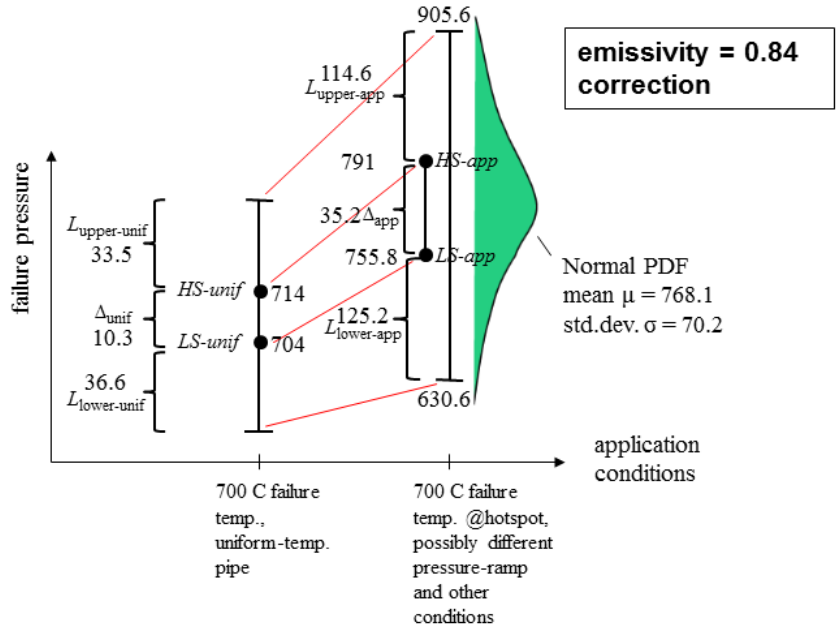

Figure 4.7 Scaling of 700C uniform-temperature 95/90 TI to approximate 95/90 TI and associated Normal PDF of predicted failure pressures. (For PB1experimental conditions and ε = 0.84 correction for temperature field reconstruction error.)

4.3 Processing and Rollup of Experimental Data Uncertainties for Model Validation Comparisons

Here the experimental data is processed in a specific manner for comparison to the simulation results in Figure 4.5 via Real Space validation metrics. An "apples-to-apples" comparison basis must first be established between predicted PDF percentiles of failure pressure and experimentally derived percentiles of failure pressure. Consequently, accounting for **non-traveling**[4] experiment-to-experiment variability and systematic uncertainty, we *normalize* the experimental data to the reference conditions input to the simulations:

- 0.02" pipe wall thickness
- PB1 nominal pressure, temperature, and mechanical end-loading boundary conditions.

We also account for inference uncertainty on estimated percentiles of response from small numbers of experiments.

An example of a linearized data normalization process is the following. Let w_{act} be the actual pipe wall thickness in test PB1. From Table 4.1 the failure pressure is $Pfail_{exper,w\_act}$ = 606psi. We will account for pressure measurement uncertainty later. Thus we have the input-out correspondence (w_{act}, $Pfail_{PB1exper,w\_act}$). This can be thought of as particular point of a function $Pfail_{exper}(w)$ that describes experimental failure pressure as a function of wall thickness.

$$Pfail_{exper,w\_act} = Pfail_{exper}(w=w_{act}) \qquad \text{Eqn. 4.1}$$

A Taylor Series is used to normalize the data to the reference wall thickness w_{ref} = 0.02" used in the failure pressure calculations in section 4.2:

$$Pfail_{exper,w\_ref} = Pfail_{exper}(w_{ref}) = Pfail_{exper}(w_{act}+[w_{ref} - w_{act}])$$
$$= Pfail_{exper}(w_{act}) + \frac{d(Pfail_{exper})}{d(w)} \bullet (w_{ref} - w_{act}) + \text{HOT.} \quad \text{Eqn. 4.2}$$

[4] *Non-traveling* uncertainties in the experiments and/or simulations in a validation activity do not transfer or "travel" consistently to application settings of intended model use that the validation is meant to inform ([6]-[8], [20], [27]). For example, non-traveling uncertainties in the PB validation activity include uncertainties on pressure and temperature measurement errors. The intended post-validation uses for the constitutive model will involve different pressure vessel geometries and different pressurization and temperature conditions, even though at similar levels as in the PB tests. Any uncertainties or parametric variations in post-validation model use are to be considered *scenario uncertainties* in the envisioned analyses, and will have no direct linkage to the measurement uncertainties in the PB validation activity. Other examples of non-traveling uncertainties are associated with the model and simulations in the validation activity, such as uncertainties associated with mesh discretization and temperature field reconstruction on the pipe (from the TC data). But an uncertainty that *is* proposed to travel consistently is the constitutive model's material-strength variability as a function of temperature, characterized in Section 2 and propagated to predictions in section 4.2 to form the variability distributions in Figure 4.4. Uncertainties are treated differently in the Real Space validation framework according to whether they are traveling or non-traveling between the validation and post-validation model use settings. This distinction governs their different significance and consequences to prediction, see [6], [8].

Here HOT stands for 'higher order terms'. The actual wall thickness w_{act} is a small perturbation from the machining spec. wall thickness of $w_{ref} = 0.02$", so the HOT contribution is relatively small and we retain just the 1st-order term:

$$Pfail_{exper}(w_{ref}) \approx Pfail_{exper}(w_{act}) + \frac{d(Pfail_{exper})}{d(w)} \bullet (w_{ref} - w_{act}). \qquad \text{Eqn. 4.3}$$

The above equation is used to approximately normalize the PB1 failure pressure to a reference wall thickness w_{ref} that differs from the actual wall thickness w_{act} at the point where failure onset occurs in the experiment. The derivative term $\frac{d(Pfail_{exper})}{d(w)}$ can be estimated either with the model or by using available test data from purposefully different wall-thickness pipes exposed to temperature and pressure conditions similar to PB1. This will be further discussed below. The more important issue for immediate discussion is that the actual wall thickness w_{act} is not known. But if an uncertainty description regarding the value of w_{act} can be reasonably determined, then this uncertainty description U[w_{act}] can be substituted into Equation 4.3. Any uncertainty U[$\frac{d(Pfail_{exper})}{d(w)}$] associated with $\frac{d(Pfail_{exper})}{d(w)}$ can also be substituted. Then Equation 4.3 becomes an equation for uncertainty of the normalized failure pressure for a reference wall thickness of 0.02 inches, U[$Pfail_{exper,w\_ref=0.02"}$].

If we normalize all replicate tests (here just two, PB1 and PB4) to the same reference wall thickness as in the failure pressure predictions, then we are on the same apples-to-apples basis of wall thickness to compare the predicted and experimental PDFs of failure pressure.
We must similarly normalize the experimental failure pressure results to the same basis of pressure loading and temperature boundary conditions before validation comparisons can be made. Generalizing the 1st-order Taylor Series approximation Equation 4.3 to multiple experimental inputs x_i, the normalization adjustments with respect to each of the individual inputs superpose (add linearly):

$$Pfail_{exper}(\vec{x}_{ref}) \approx Pfail_{exper}(\vec{x}_{act}) + \sum \frac{\partial(Pfail_{exper})}{\partial(x_i)} \bullet (x_{i\_ref} - x_{i\_act}). \qquad \text{Eqn. 4.4}$$

Here $Pfail_{exper}(\vec{x}_{act})$ is the measured failure pressure, which occurs under all the actual experimental input values $\vec{x}_{act}$.

Any measurement and processing uncertainties associated with the experimental output results of interest (here $Pfail_{exper}(\vec{x}_{act})$) are combined with any uncertainties in the partial derivatives and actual input values x_{i_act} on the right hand side (RHS) of Equation 4.4. All these uncertainties are propagated to output uncertainty on the normalized failure pressure on the left side of the equation. Any correlation between the uncertainties on the RHS of the equation must be accounted for in the propagation. A simple spread-sheet-based approach is demonstrated later in this subsection. This approach would be a 1st-order or linear UQ method if only incorporating uncertainties of the inputs x_{i_act} and of measurements of outputs (first term on RHS) in Equation 4.4. But the including the uncertainty of the partial derivatives makes this a nonlinear UQ method. It will be referred to as a 1st-order+ or linear+ UQ method.

A higher-order UQ approach to evaluation of the RHS of Equation 4.4 is presented in [22]. Equation 4.4 and any higher-fidelity UQ approaches are applicable to either or both of random and systematic non-traveling uncertainties in replicate experiments supporting model calibration or validation.

For problems involving only non-traveling *systematic* uncertainty of experimental inputs x_i, Equation 4.4 with linear UQ reduces to results that are equivalent to two very different methods of construction in [6] and [27]. For problems involving non-traveling input *and* output uncertainties, Equation 4.4 with linear UQ ultimately yields the combined "validation uncertainty" in [28] when its 1st-order linear UQ version is used and non-traveling uncertainties from the simulations are also taken into account in both approaches—provided the more restrictive conditions in [28] are met: A) uncertainty is assumed to be probabilistic and represented and propagated accordingly; B) only non-traveling uncertainties exist in the models and experiments; C) the system of interest (in the models and experiments) has no significant degree of stochastic behavior/response for the quantities of interest. This is a reassuring corroboration of both methods (RS and [28]) for the subset of conditions cited, given that their derivations come from very different conceptual approaches.

The manner of uncertainty treatment and presentation of results in the RS approach is simpler when only systematic uncertainties are present in the replicate experiments (as demonstrated in [22]) vs. the more involved treatment when both random and systematic (or only random) experimental uncertainties are present. The treatment below also handles probabilistic and/or interval descriptions of the uncertainties. An earlier method of treatment in [23] handles all that the method below does, but the method of aggregating interval uncertainties in [23] almost certainly grossly exaggerates uncertainty when more than a couple of interval uncertainties are present. But the simpler treatment in [23] has the advantage that it can be easily accomplished with hand calculations; it does not require a spreadsheet or random number generator.

4.3.1 Normalization of Reference Test results

The next step in the present procedure is to write Equation 4.4 for the reference test that we are normalizing all the test results to. In this paper the reference test is PB1, and normalization is to the PB1 *nominal* input conditions used in the simulations in subsection 4.2 and listed in the first paragraph of the present subsection. We will later write normalization equations for the other replicate experiments (here PB4).

$$Pfail_{\mathrm{PB1}}(\vec{x}_{\mathrm{nomPB1}}) \approx Pfail_{\mathrm{PB1}}(\vec{x}_{\mathrm{actPB1}}) + \sum \frac{\partial(Pfail_{\mathrm{PB1}})}{\partial(x_i)} \bullet (x_{\mathrm{i\_nomPB1}} - x_{\mathrm{i\_actPB1}}) \qquad \text{Eqn. 4.5}$$

We only know the failure pressure in PB1 to within the uncertainty of the pressure measurement in the test (see Table 4.1). Hence $Pfail_{\mathrm{PB1}}(\vec{x}_{\mathrm{actPB1}})$ on the RHS of Equation 4.5 has an uncertainty U[$Pfail_{\mathrm{PB1}}(\vec{x}_{\mathrm{actPB1}})$] = 606 psi ± 10 psi = [596, 616] psi.

The two most substantial terms (by far) in the summation in Equation 4.5 involve differences between nominal and actual wall thickness, and nominal and actual temperature, at the location of failure initiation in PB1. Hence Equation 4.5 is rewritten as

$$Pfail_{\text{PB1}}(\vec{x}_{\text{nomPB1}}) \approx Pfail_{\text{PB1}}(\vec{x}_{\text{actPB1}}) + \frac{\partial(Pfail_{\text{PB1}})}{\partial(w)} \bullet (w_{\text{nomPB1}} - w_{\text{actPB1}}) + \frac{\partial(Pfail_{\text{PB1}})}{\partial(T@fail\_point)} \bullet (T_{\text{nomTC4-PB1}} - T_{\text{act@fail\_point-PB1}}). \quad \text{Eqn. 4.6}$$

Recall that the PB1 TC data provide the nominal temperature distribution on the pipe for the PB1 simulations, and the effect of any errors/uncertainties in reconstructing the temperature BCs from the *nominal* TC data have already been accounted for in Section 4.2. However, the nominal TC data is not representative of the actual PB1 temperature conditions in two other respects: 1) temperature measurement errors of the TCs still need to be accounted for; and 2) the actual temperature distribution may have a peak temperature that is hotter than at the TC4 location (the hottest TC location in all the tests).

It is presumed that failure initiates at the location (point) where the combination of local temperature and wall thickness provides the weakest point on the pipe wall, i.e., the lowest resistance to pressure loading. Yielding initiates there and proceeds until failure occurs there. Indeed, the model simulations, which have uniform wall thickness in the reduced-thickness 0.02" wall section, predict that initial yielding and then failure occurs at the hottest point on the pipe (see Figure 4.9). We assume that in the tests the failures initiate within a close viscinity of the hottest TC, TC4. An investigation in [29] suggests this. For the monitored specimen, failure initiates at a location indistinguishable from the TC4 location, where the pipe wall first splits upwards and downwards, and then circumferentially along the top and bottom thickness-transition shoulders as the "butterfly wings" are created (see Figure 1.1). The approximate symmetry of the final geometries in Figure 1.1 also implies that failure initiation is at the heated front center of the pipe. So we use TC4 as a reference temperature, above which we propose 15C as a reasonable maximum possible perturbation to address item 2) in the prior paragraph. Additionally, absent significant temperature measurement uncertainties like in Appendix D in [9] which applies to much larger-diameter TCs than in the PB tests, the actual TC errors/uncertainties in PB1and PB4 are much more pivotal for TC4 than for the other TCs. Furthermore, propagating the other TCs' uncertainties to evaluate similar terms in Equation 4.6 would require additional model simulations that the project could not afford. Therefore we address item 1) in the prior paragraph by considering only TC4's temperature measurement uncertainty.

The wall thickness uncertainty in PB1 is modeled as the interval range given in Table 4.3:

$$U[w_{\text{actPB1}}] = [0.019, 0.022] \text{ inch.} \quad \text{Eqn. 4.7}$$

The term $\frac{\partial(Pfail_{\text{PB1}})}{\partial(w)}$ in Equation 4.6 is approximated by considering the experimental and model simulation results in Figure 4.8. The figure shows predicted failure pressures (listed in Table 4.4) for test PB1 nominal boundary conditions and pipe wall thicknesses of: 0.02", the nominal value, and 0.019" and 0.024", the lowest and highest wall thickness measurements recorded in Table

4.3. The model used was the full-geometry 1tt-mesh model used for the predictions in Figure 4.5, except for changes to wall thickness. Thus, the value in Figure 4.8 for the nominal 0.02" wall is the same as in Figure 4.4. Failure pressures at the three thicknesses were all corrected downward to approximate mesh-converged results by the same nominal correction of -245 psi shown in Figure 4.4. The corrected values are listed in Table 4.4. The regression line through the simulation results in Figure 4.8 has a slope of 42,286 psi/inch-wall-thickness and a y-intercept of -16 psi when extrapolated to zero wall thickness.

Table 4.4 Failure pressures predicted for nominal PB1 conditions and pipe wall thicknesses listed.

Pipe wall thicknes (inches)	Predicted failure pressure (approximately corrected for mesh discretization effects)
0.019"	780 psi
0.020"	839 psi
0.024"	997 psi

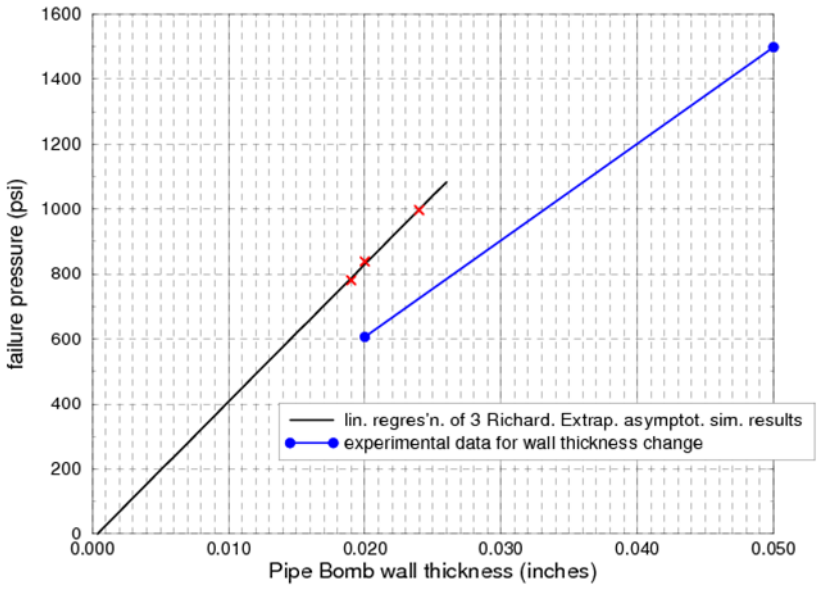

Figure 4.8 Failure pressures vs. wall thickness from simulations and experiments.

The experimental data line in Figure 4.8 connects failure pressures in the PB1 and PB3 tests (Table 4.5). The experimental line has a slope = 29,733 psi/inch-wall-thickness and a y-intercept of 11.3 psi when extrapolated to zero wall thickness. Test PB3 (see [29]) had a 0.05-inch thick wall and had a hot-spot TC4 temperature (Figure 4.3) of ~700C at failure, similar to PB1. But the pressure and temperature ramping conditions were somewhat different in the PB1 and PB3 tests. Furthermore, the experimental slope of failure pressure vs. wall thickness is subject to much uncertainty because it is based on only one test with 0.05" wall and one 0.02" wall result. (Table 4.1 gives a glimpse of the magnitude of variation that exists among failure pressures even for

nominal replicate tests.) It would of course be best to average over many tests at different wall thicknesses to obtain an experimental slope of failure pressure vs. wall thickness, but PB3 is the only available test with a wall thicknesses other than 0.02". Test PB1 had the closest temperature conditions at failure to test PB3, so is the best one to plot with PB1 in Figure 4.8. Test PB4 at failure also has close temperature conditions to PB3, but PB4 appears from Table 4.3 to have an actual wall thickness not as close to the 0.02" value used in Figure 4.8 and Table 4.4.

Table 4.5 Experimental failure pressures for nominal PB1 conditions and pipe wall thicknesses listed.

Pipe wall thicknes	measured failure pressure
0.02"	606 psi
0.05"	1498 psi

Given the potential errors in both the simulated and experimental slopes, the following course is taken. The differences in PB1 and PB4 failure pressures (for nominal repeat experiments) is 49 psi, from Table 4.1. Perturbing the experimental failure pressures in Table 4.5 by ±49 psi (as a plausible scale of uncertainty to work with) gives four combinations from which the lowest and highest slopes are 26,467 and 33,000 psi/inch-wall-thickness. If we double the perturbations to ±100 psi at both wall thicknesses, then the lowest and highest slopes are 23,067 and 36,400. The upper value is about 15% less than the slope in Figure 4.8 from the model simulations. We go ahead and use the higher value of 42,286 from the simulations as an upper value in our UQ analysis, and use as the lower limit the low of 23,067 from considering the experimental data.

$$\mathrm{U}\left[\frac{\partial(Pfail_{\mathrm{PB1}})}{\partial(w)}\right] = [23{,}067,\ 42{,}286] \text{ psi/inch-wall-thickness.} \qquad \text{Eqn. 4.8}$$

The term ($T_{\mathrm{nomTC4\text{-}PB1}}$ - $T_{\mathrm{act@fail\_point\text{-}PB1}}$) in Equation 4.6 captures the difference between the nominal hot-spot temperature (reading from TC4 in Figure 4.3), and the actual temperature at the location of failure initiation in the PB1 test. With this term and the partial derivative $\frac{\partial(Pfail_{\mathrm{PB1}})}{\partial(T@fail\_point)}$ we normalize between the nominal temperature conditions modeled, and the actual temperature conditions in the test. For the reasons mentioned earlier we normalize only for local temperature differences at the failure initiation point (close viscinity of TC4) and ignore measurement uncertainties of the other TCs.

The difference ($T_{\mathrm{nomTC4\text{-}PB1}}$ - $T_{\mathrm{act@fail\_point\text{-}PB1}}$) is separated into two components:

A) <u>TC4 measurement error</u> – the difference between the TC4 reading value $T_{\mathrm{nomTC4\text{-}PB1}}$ and the actual pipe temperature at that location, $T_{\mathrm{actTC4\text{-}PB1}}$;

B) <u>TC4 location related error</u> – any difference between the TC4 temperature after being corrected for any measurement error, $T_{\mathrm{actTC4\text{-}PB1}}$, and the temperature where failure occurs, $T_{\mathrm{act@fail\_point\text{-}PB1}}$.

Thus, the following identity is written in terms of components A and B.

$$T_{\text{nomTC4-PB1}} - T_{\text{act@fail\_point-PB1}} = (T_{\text{nomTC4-PB1}} - T_{\text{act@TC4-PB1}}) + (T_{\text{act@TC4-PB1}} - T_{\text{act@fail\_point-PB1}}).$$
$$= \text{TC4 measurement error} + \text{TC4 location related error} \quad \text{Eqn. 4.9}$$

TC4 measurement error itself has several components:

$$\text{TC4 measurement error} = T_{\text{nomTC4-PB1}} - T_{\text{act@TC4-PB1}} = (T_{\text{nomTC4-PB1}} - T_{\text{ofTC4-PB1}}) + (T_{\text{ofTC4-PB1}} - T_{\text{act@TC4-PB1}}). \quad \text{Eqn. 4.10}$$

Here, $\{T_{\text{nomTC4-PB1}} - T_{\text{ofTC4-PB1}}\} = \Delta T_{\text{meas-TC/DAQ}}$ is the difference or error between the TC's actual temperature, $T_{\text{ofTC4-PB1}}$, and the temperature $T_{\text{nomTC4-PB1}}$ reported by the data acquisition system (DAQ). This error is caused by inaccuracies associated with the TC transducer itself, the calibration standard and procedure used to calibrate the TCs if calibration is performed, and the DAQ. The combined TC/calibration/DAQ errors are typically very small. Uncertainty on the combined error is estimated from information in [25] as:

$$U[\Delta T_{\text{meas-TC/DAQ}}] = U[T_{\text{nomTC4-PB1}} - T_{\text{ofTC4-PB1}}] = [-0.25\%, 0.25\%] \text{ of TC4 reading in °C.} \quad \text{Eqn. 4.11}$$

$T_{\text{ofTC4-PB1}} - T_{\text{act@TC4-PB1}} = \Delta T_{\text{meas-contact}}$ is the difference or error between the TC's actual temperature, $T_{\text{ofTC4-PB1}}$, and the temperature $T_{\text{act@TC4-PB1}}$ of the pipe surface at the TC4 location. This error is caused by contact resistance between the surface and the attached TC, and by convective and radiative conditions affecting heat losses from the TC bead or attached wire tips if an intrinsic TC. The PB tests used intrinsic TCs with very small wire tip diameters of 0.005-in. Uncertainty on the temperature difference between the TC and the surface point it is attached to is estimated from information in [26]:

$$U[\Delta T_{\text{meas-contact}}] = U[T_{\text{ofTC4-PB1}} - T_{\text{act@TC4-PB1}}] = [0., 0.25\%] \text{ of TC4 reading in °C.} \quad \text{Eqn. 4.12}$$

The uncertainty ranges from zero to + 0.25% because the TC4 leads are between the surface being measured and the heating plate. Therefore, at the attachment location the leads will be hotter than the surface. Hence the error will be positive. In general, errors ΔT_{meas} in Equation 4.11 and 4.12 have positive values when, taken alone, they cause an overestimate of the true surface temperature at the TC4 location.

Component B of failure-temperature error in Equation 4.9 is due to TC4's location being potentially different from where failure occurs:

$$T_{\text{act@TC4-PB1}} - T_{\text{act@fail\_point-PB1}} = \Delta T_{\text{TC4location}} \quad \text{Eqn. 4.13}$$

We mentioned earlier that the actual peak temperature on the pipe could be a small distance away from TC4, and as much as 15C above the true pipe surface temperature at TC4. Then Equation 4.13 yields $\Delta T_{\text{TC4location}} = -15\text{C}$ as the largest negative value that can occur.

But in the real experiment, due to wall thickness and material strength spatial variations, failure will not necessarily occur at the hottest point on the pipe. Instead it occurs at the location of the weakest combination of thickness, temperature, and material strength. Therefore we must also consider possible failure temperatures lower than the temperature $T_{\text{act@TC4-PB1}}$ at TC4. It is more difficult to propose a reasonable lower bound for the possible true failure temperature $T_{\text{act@fail\_point-PB1}}$. We simply guess that failure could occur at a point that is as much as 15C below the pipe temperature at the TC4 location. Then Equation 4.13 yields a maximum value of 15C for $\Delta T_{\text{TC4location}}$. Thus,

$$\text{U}[\Delta T_{\text{TC4location}}] = \text{U}[T_{\text{act@TC4-PB1}} - T_{\text{act@fail\_point-PB1}}] = [-15, +15]^{\circ}\text{C}. \qquad \text{Eqn. 4.14}$$

In all we have:

$$\begin{aligned} T_{\text{nomTC4-PB1}} - T_{\text{act@fail\_point-PB1}} &= (T_{\text{nomTC4-PB1}} - T_{\text{act@TC4-PB1}}) + (T_{\text{act@TC4-PB1}} - T_{\text{act@fail\_point-PB1}}). \\ &= \text{TC4 measurement error} + \text{TC4 location related error} \\ &= (\Delta T_{\text{measTC/DAQ}} + \Delta T_{\text{meas-contact}}) + \Delta T_{\text{TC4location}}. \qquad \text{Eqn. 4.15} \end{aligned}$$

Failure pressure difference due to different nominal and actual wall thicknesses in the PB1test is approximated in Equation 4.6 by $\frac{\partial(Pfail_{\text{PB1}})}{\partial(w)} \bullet (w_{\text{nomPB1}} - w_{\text{actPB1}})$, whereas failure pressure difference due to different local material temperatures is approximated by (from equations 4.6 and 4.15):

$$\frac{\partial(Pfail_{\text{PB1}})}{\partial(T@fail\_point)} (\Delta T_{\text{meas-TC/DAQ}} + \Delta T_{\text{meas-contact}} + \Delta T_{\text{TC4location}}). \qquad \text{Eqn. 4.16}$$

The partial derivative in the above equation is approximated by the slope of the curve in Figure 2.5 evaluated for the temperature at the failure location. The slopes are calculated from the means of the data at the various temperature levels. The slope in Figure 2.5 between the data means at 600C and 700C is -1.71 psi/C. The slope between the data means at 700C and 800C is -2.61 psi/C. Averaging these two slopes gives -2.16 psi/C, which is the same value that would come from central-differencing for a second-order approximation to the slope at 700C. The temperature perturbations in Equations 4.6 (and 4.16) are relative to the PB1nominal failure temperature of 706C and applicable temperatures remain within about [-30, +80] of 700C. Given all this information the previously mentioned values of -1.71 psi/C and -2.61 psi/C are used as uncertainty extremes about the nominal 700C value of -2.16 psi/C. Thus,

$$\text{U}\left[\frac{\partial(Pfail_{\text{PB1}})}{\partial(T@fail\_point)}\right] = -[1.71, 2.16] \text{ psi/C}. \qquad \text{Eqn. 4.17}$$

Given the uncertainty descriptions of the applicable terms on the RHS of Equation 4.6 we are now ready to sample the uncertainties to estimate the uncertainty on the LHS term. The resulting LHS uncertainty, $\text{U}[Pfail_{\text{PB1}}(\vec{x}_{\text{nomPB1}})]$, is the uncertainty of the PB1 failure pressure when normalized to the nominal experimental conditions input to the model simulations in Section 4.2. We obtain samples j=1,J of the LHS uncertainty $\text{U}[Pfail_{\text{PB1}}(\vec{x}_{\text{nomPB1}})]$ by combining j=1,J realizations of the uncertainties on the RHS terms in Equation 4.6. In terms of the intervening developments, the realizations are written as:

$$Pfail_{\mathrm{PB1}}(\vec{x}_{\mathrm{nomPB1}})_{\,\mathrm{j}} \approx Pfail_{\mathrm{PB1}}(\vec{x}_{\mathrm{actPB1}})_{\,\mathrm{j}} + \{\frac{\partial(Pfail_{\mathrm{PB1}})}{\partial(w)} \bullet \Delta w_{\mathrm{PB1}}\}_{\mathrm{j}} + \{\frac{\partial(Pfail_{\mathrm{PB1}})}{\partial(T@fail\_point)}(\Delta T_{\text{meas-TC/DAQ}} + \Delta T_{\text{meas-contact}} + \Delta T_{\text{TC4location}})\}_{\mathrm{j}} \quad \text{Eqn. 4.18}$$

where the Δ terms have already been defined except for

$$\Delta w_{\mathrm{PB1}} = w_{\mathrm{nomPB1}} - w_{\mathrm{actPB1}}. \quad \text{Eqn. 4.19}$$

A spreadsheet is a convenient way to do the sampling and processing of the realizations. From Equation 4.18 we populate a spreadsheet as illustrated by Tables 4.6 and 4.7. The quantities in yellow highlighted columns 2, 5, 8, and 10 in Table 4.6 designate correlated uncertainties with the same-numbered columns for other tests to be described later. For all uncertainties in Tables 4.6 and 4.7 designated by interval ranges […] (e.g. from Equation 4.17), samples from uniform distributions are obtained from available sampling functions in the spreadsheet, or are imported into the spreadsheet from external sampling operations.

The uncertainties in the present application are defined as interval ranges. Both interval and probabilistic (PDF) uncertainties have been treated in other applications ([19], [22], [23]). A different method of treatment of interval uncertainties in [23] used interval propagation and aggregation techniques to deal with interval uncertainties over the monotonic uncertainty space in that problem. But this method of aggregation almost certainly grossly exaggerates representation of the uncertainty when more than a couple of interval uncertainties are present. This is because combinations of the extreme values of the interval uncertainties present an increasingly remote possibility as the number of interval uncertainties rises. We use the sensibility that the validation conclusions should not be driven by the remote possibilities of compounded extremes of the interval uncertainties. Rather, we here use uniform PDF representations of the interval uncertainties and propagate them probabilistically via the spreadsheet. If sufficient sampling is performed (and this is very fast and inexpensive inside or outside of the spreadsheet), then the resulting density function (DF) will have endpoints that coincide with the endpoints from an interval UQ treatment. But the DF will not weight the very remote extreme combinations as prominently. Thus, this manner of treatment takes the "edge" off a compounded interval treatment. But we do not interpret the resultant DF as a probability density. We simply use the DF to arrive at a moderated magnitude of interval uncertainty for the validation assessment (see section 4.4).

A similar moderating approach is used in [22] where both interval and probabilistic uncertainties are present and are propagated in a segregated fashion. In [19] the intervals were simply treated as uniform PDFs and comingled with the probabilistic uncertainties in propagation. The treatment in [22] is now preferred when both interval and probabilistic uncertainties are present.

Care should be exercised with random number generator seeds and sequences in the spreadsheet so that the columns of random samples in Table 4.7 (and Table 4.9 etc.) are independent of each other unless correlation is intended and is purposely imposed. For the particular comparisons to be made in Section 4.4 it is recommended that at least 1000 realizations be taken ($J \geq 1000$ in Tables 4.6, 4.7, etc.). In the following we use J=1000.

Table 4.6 Spreadsheet formulas for uncertain experimental quantities in Equation 4.18 for test PB1

realization j	$\Delta Pfail_{\text{measPB1, j}}$ = [-10, 10] psi	$w_{\text{actPB1,j}}$ = [0.019, 0.022] inches	$\Delta w_{\text{PB1,j}} = w_{\text{nomPB1}} - w_{\text{actPB1,j}} = 0.02'' - A3$	$\frac{\partial(Pfail_{\text{PB1}})}{\partial(w)}$ j = [23,067 , 42,286] psi/inch	= (A4*A5) j psi	$\Delta T_{\text{meas-TC/DAQ,j}}$ = [–0.0025*707, +0.0025*707] C	$\Delta T_{\text{meas-contact,j}}$ = [0., +0.0025*707] C	$\Delta T_{\text{TC4location,j}}$ = [–15, 15] C	$\frac{\partial(Pfail_{\text{PB1}})}{\partial(T@fail\_point)}$ j = [–2.16, –1.71] psi/C	= (A7 + A8 + A9) j *A10j psi	= 606 psi + (A2 + A6 + A11)j psi = $Pfail_{\text{PB1}}(\vec{x}_{\text{nomPB1}})$ j
A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
1											
2											
…	…	…	…	…	…	…	…	…	…	…	…
J											

Table 4.7 Spreadsheet realizations for uncertain experimental quantities in test PB1

	A2=ΔPmeas	A3=PB1wall	A4=Δ0.02"	A5=∂P/∂w	A6 = A4*A5	A7=ΔTC_DAC	A8=ΔTC_cntc	A9=ΔTC_loca	A10=∂P/∂T	A11=A10*(A7	A12=606+A2+
1	7.76	0.02135	-0.00135	30701.34	-41.57	1.00	0.52	-4.65	-2.07	6.49	578.68
2	-5.88	0.01928	0.00072	37837.13	27.38	0.23	1.18	-8.51	-1.87	13.27	640.77
3	-0.69	0.02135	-0.00135	26116.40	-35.35	0.39	1.61	-13.98	-1.83	21.89	591.86
999	-8.55	0.01958	0.00042	37867.63	15.98	1.58	0.10	-2.59	-2.15	1.97	615.41
1000	6.28	0.01943	0.00057	29062.56	16.56	-1.25	0.06	-9.80	-1.99	21.92	650.75
	avg	avg	avg	avg	avg	avg	avg	avg	avg	avg	avg
	0.00	0.02056	-0.00056	32361.71	-17.82	0.06	0.91	0.07	-1.94	-1.99	586.19
	stdev				stdev					stdev	stdev
	5.76				28.13					17.09	32.78
											max
											671.86
											min
											502.78

Column A12 of Table 4.7 contains the realizations of uncertainty of PB1 failure pressure, U[$Pfail_{\text{PB1}}(\vec{x}_{\text{nomPB1}})$], when normalized to the nominal PB1 experimental conditions input to the model simulations in Section 4.2. These realizations are histogrammed in Figure 4.9 along with

realizations from similar normalizations of other tests. The other tests and their normalized results will be discussed later.

Statistics at the bottom of Table 4.7 summarize the impacts of various uncertainties in the normalization of PB1 failure pressure. The green box in column A6 shows the average adjustment of PB1 experimental failure pressure normalized for possible differences from the nominal 0.02" wall thickness used in the simulations in Section 4.2. The green box in column A11 shows the average bias adjustment when PB1 failure pressure is normalized for possible differences from the nominal temperature field and failure location in the simulations. The green box in column A2 shows a zero average bias adjustment when PB1 failure pressure is normalized for possible measurement error affecting the nominal pressure values used in the simulations. The zero average adjustment occurs because the measurement uncertainty [-10, +10] psi is symmetric about a value of zero measurement error. The average adjustment of -17.8 psi for wall thickness adds with the average adjustment of -2 psi for temperature normalization to yield a combined adjustment of -19.8 psi from the nominal measured failure pressure of 606 psi in Table 4.1. The resulting average normalized failure pressure is 586.2 psi listed in the pink box at the bottom of Table 4.7. Regarding sensitivity, the average adjustment of -17.8 psi associated with wall thickness normalization is much larger than the average adjustment of -2 psi due to temperature normalization.

Possible wall thickness differences from nominal lead to variations in normalized failure pressure characterized by the standard deviation of 28.1 psi in the grey box in column A6. Possible temperature differences from nominal yield a standard deviation of 17.1 psi in normalized failure pressure per the grey box in column A11. Possible pressure measurement error yields a standard deviation of 4.8 psi in failure pressure per the grey box in column A2. These three sources of variance combine to yield a variance characterized by the standard deviation in the pink box at the bottom of Table 4.7. The individual variance effects add according to the sum of the squares of their standard deviations. The square root of this sum is 33.4 psi. This closely agrees with the standard deviation 32.8 psi in the pink box, calculated directly from the realizations in column A12. The 1.9% discrepancy between the two methods of calculation is attributed to sampling differences underlying the standard deviations calculated from the realizations down columns A2, A6, and A11 vs. the standard deviation of the realizations in column A12 calculated from row sums of uncorrelated random values across the rows of columns A2, A6, and A11. The discrepancy between standard deviation calculated these two ways is expected to diminish as the number of realizations $J \rightarrow \infty$. In terms of sensitivity, the variance contributed by wall thickness normalization is much larger than variance contributed by temperature normalization, which itself is much larger than variance contributed by pressure measurement uncertainty. Comparing the temperature uncertainties in columns A7, A8, and A9 shows that, by far, the largest component of variance due to temperature normalization is contributed by failure temperature uncertainty due to uncertainty in the actual location of failure initiation.

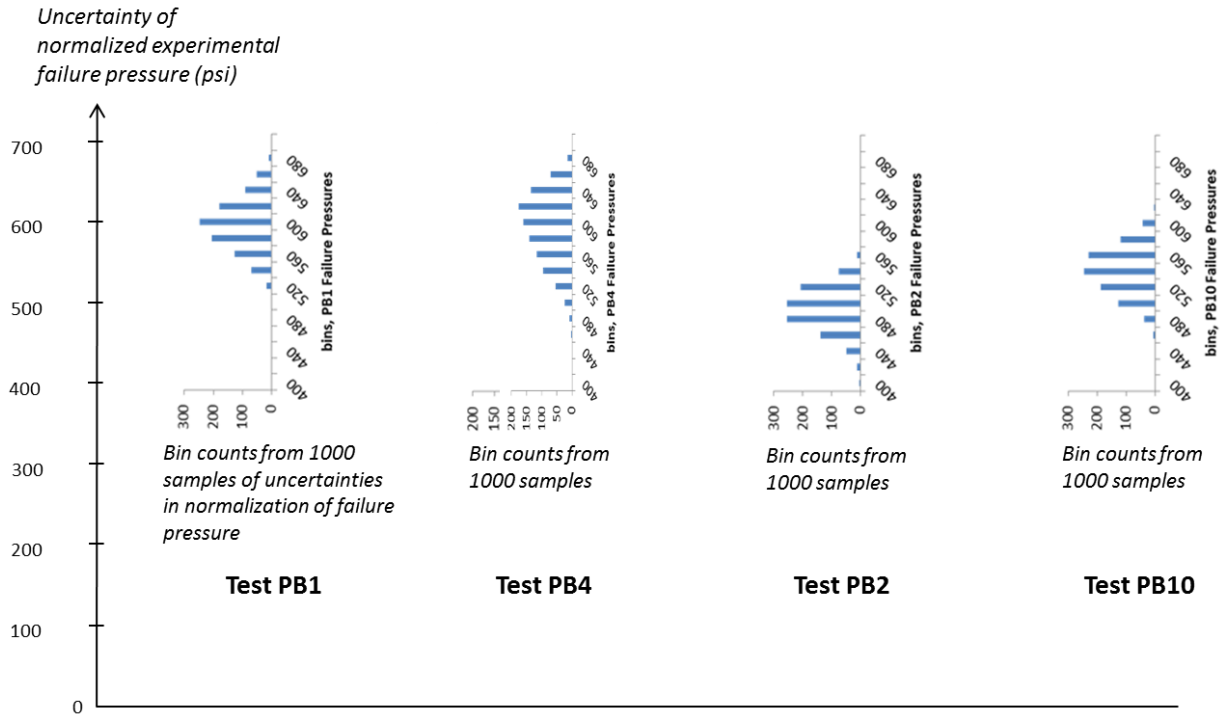

Figure 4.9 Uncertainty histograms of experimental failure pressures normalized to input conditions of model validation simulations.

4.3.2 Normalization for Replicates of the Reference Test

To write normalization equations for tests other than the one we are normalizing to (here PB1), we illustrate the procedure for test PB4. We start by writing a characteristic identity for such tests (here for PB4).

$$\begin{aligned} Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{nomPB1}}) = {} & Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{actPB4}}) \\ & + [Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{nomPB4}}) - Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{actPB4}})] \\ & + [Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{nomPB1}}) - Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{nomPB4}})]. \end{aligned} \qquad \text{Eqn. 4.20}$$

The term in the second row of Equation 4.20 is evaluated from a PB4 analogue of Equation 4.5, to normalize test results between PB4 actual and nominal experimental conditions:

$$Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{nomPB4}}) - Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{actPB4}}) \approx \sum \frac{\partial(PfailPB4)}{\partial(x_i)} \bullet (x_{\mathrm{i\_nomPB4}} - x_{\mathrm{i\_actPB4}}). \qquad \text{Eqn. 4.21}$$

The term in the third row of Equation 4.20 represents the difference that would exist if the PB4 pipe was subjected to the nominal input conditions from test PB1. We use the simulation model to approximate this difference by simulating to failure with the nominal PB1 inputs and then with the nominal PB4 inputs:

$$Pfail_{\text{PB4}}(\vec{x}_{\text{nomPB1}}) - Pfail_{\text{PB4}}(\vec{x}_{\text{nomPB4}}) \approx Pfail_{\text{model}}(\vec{x}_{\text{nomPB1}}) - Pfail_{\text{model}}(\vec{x}_{\text{nomPB4}}). \quad \text{Eqn. 4.22}$$

The model is used as the best available mechanism for estimating how things would change in reality under these perturbations to the input conditions. The model doesn't have to be accurate in an absolute sense, just sufficiently accurate in a relative sense to treat the uncertainties in a way that improves the uncertainty analysis vs. simply ignoring these uncertainties.

Substituting equations 4.21 and 4.22 into Equation 4.20 yields

$$\begin{aligned} Pfail_{\text{PB4}}(\vec{x}_{\text{nomPB1}}) \approx\ & [Pfail_{\text{model}}(\vec{x}_{\text{nomPB1}}) - Pfail_{\text{model}}(\vec{x}_{\text{nomPB4}})] \\ & + Pfail_{\text{PB4}}(\vec{x}_{\text{actPB4}}) + \sum \frac{\partial(PfailPB4)}{\partial(x_i)} \bullet (x_{\text{i\_nomPB4}} - x_{\text{i\_actPB4}}). \end{aligned} \quad \text{Eqn. 4.23}$$

The term in row 1 of the RHS of Equation 4.23 enacts an approximate adjustment for PB1 vs. PB4 nominal input conditions, $\vec{x}_{\text{nomPB1}}$ vs. $\vec{x}_{\text{nomPB4}}$. The term in row 2 of the RHS enacts an approximate adjustment for PB4 nominal vs. actual conditions, $\vec{x}_{\text{nomPB4}}$ vs. $\vec{x}_{\text{actPB4}}$. The treatment of row 2 is analogous to Equation 4.5 and hence ultimately Equation 4.18. The new information required for the PB4 evaluation in Table 4.7 follows.

The wall thickness uncertainty in PB4 is modeled as the interval range given in Table 4.3:

$$\text{U}[w_{\text{actPB4}}] = [0.02, 0.024] \text{ inch.} \quad \text{Eqn. 4.24}$$

The failure pressure in PB4 has an uncertainty $\text{U}[Pfail_{\text{PB4}}(\vec{x}_{\text{actPB4}})]$ = 655 psi ± 10 psi (see Table 4.1). In Tables 4.8 and 4.9 this ±10 psi uncertainty is treated as being perfectly correlated to the pressure measurement uncertainty ±10 psi in PB1 (Tables 4.6 and 4.7) because the same pressure gauge was used in the two tests so *systematic* measurement error exists among the two tests. Assuming that the random-error component is small relative, the gage's majority of error and the associated uncertainty are systematic over the replicate tests. Hence the realizations in column B2 are taken to be the same as in column A2. The heading in column B2 is highlighted yellow to signify commonality of this quantity with the quantity in column A2 for test PB1.

The quantities in columns B5, B8, and B10 of Tables 4.8 and 4.9 are also highlighted yellow, for similar reasons. The green-highlighted entries in Table 4.8 denote changes from PB1's Table 4.6. Column A12 of Table 4.6 is shifted rightward to become column B13 in Table 4.8, with other changes highlighted in green in the column B13 heading. The new column B12 represents row 1 of the RHS of Equation 4.23. This term enacts an approximate adjustment to PB4 failure pressure according to the difference in computational model results at PB1 nominal inputs and PB4 nominal inputs. Note that the uncertainties [-15, +15] in column B9 and in column A9 for PB1 are independent of each other; the realizations in a given row j of columns A9 and B9 are uncorrelated. Indeed, different initial seeds are used to sample all column quantities for PB1 and PB4 except for the yellow-highlighted columns as explained earlier.

Table 4.8 Spreadsheet formulas for uncertain experimental quantities in Equation 4.23 for test PB4

realization j	$\Delta Pfail_{\text{measPB4,}}$ j = A2j, systematic betw. PB1 & PB4	$w_{\text{actPB4,j}}$ = [0.02, 0.024] inches	$\Delta w_{\text{PB4,j}} = w_{\text{nomPB4}} - w_{\text{actPB4,j}} =$ 0.02" – B3	$\frac{\partial(Pfail_{\text{PB4}})}{\partial(w)}$ j = A5j, systematic betw. PB1 & PB4	= (B4*B5) j psi	$\Delta T_{\text{meas-TC/DAQ,j}}$ = [–0.0025*711, +0.0025*711] C	$\Delta T_{\text{meas-contact,j}}$ = A8j, systematic betw. PB1 & PB4	$\Delta T_{\text{TC4location,j}}$ = [–15, 15] C	$\frac{\partial(Pfail_{\text{PB4}})}{\partial(T@fail\_point)}$ j = A10j, systematic betw. PB1 & PB4	= (B7 + B8 + B9) j *B10j	$[Pfail_{\text{model}}(\vec{x}_{\text{nomPB1}}) - Pfail_{\text{model}}(\vec{x}_{\text{nomPB4}})]$	= 655 psi + (B2 + B6 + B11 + B12)j psi = $Pfail_{\text{PB4}}(\vec{x}_{\text{nomPB1}})$ j
B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13
1												
2												
…	…	…	…	…	…	…	…	…	…	…	…	…
J												

Table 4.9 Spreadsheet realizations for uncertain experimental quantities in test PB4

j	B2=A2, sys	B3=PB4wall	B4=Δ0.02"	B5=A5, sys	B6 = B4*B5	B7=ΔTC_DA	B8=A8, sys.	B9=ΔTC_loca	B10=A10, sy	B11=B10*(B7	B12=PfPB1-	B13=655+B2+
1	7.76	0.02266	-0.00266	30701.34	-81.65	-1.39	0.52	12.84	-2.07	-24.83	-8.00	548.29
2	-5.88	0.02327	-0.00327	37837.13	-123.76	0.91	1.18	-11.81	-1.87	18.17	-8.00	535.53
3	-0.69	0.02195	-0.00195	26116.40	-51.05	0.07	1.61	-6.42	-1.83	8.66	-8.00	603.93
999	-8.55	0.02399	-0.00399	37867.63	-150.97	0.67	0.10	-14.21	-2.15	28.95	-8.00	516.43
1000	6.28	0.02353	-0.00353	29062.56	-102.57	-0.85	0.06	2.18	-1.99	-2.78	-8.00	547.94
	avg	avg	avg	avg	avg	avg	avg	avg	avg	avg	avg	avg
	0.00	0.02195	-0.00195	32361.71	-62.83	-0.05	0.91	-0.39	-1.94	-0.96	-8.00	583.21
	stdev				stdev					stdev		stdev
	5.76				39.06					17.02		42.97
												max
												672.99
												min
												445.61

Column B13 of Table 4.8 contains the realizations of uncertainty U[$Pfail_{\text{PB4}}(\vec{x}_{\text{nomPB1}})$] of PB4 failure pressure when normalized to the nominal PB1 reference experimental conditions input to

the model simulations in Section 4.2. These realizations are histogrammed in Figure 4.9. The realizations are processed with realizations from the other replicate tests as explained in the next section.

The statistics at the bottom of Table 4.9 summarize the impacts of various uncertainties in the normalization of PB4 failure pressure. The green boxes reveal that the average normalization adjustment in PB4 experimental failure pressure is greatest (-62.8 psi) due to possible differences from the nominal 0.02" wall thickness used in the simulations. This is a much larger mean adjustment than the -17.8 psi mean adjustment for PB1 normalization because PB4's [0.02", 0.024"] range of wall thickness uncertainty is centered significantly further from the nominal 0.02" than is PB1's [0.019", 0.022"] uncertainty range. Next in magnitude of normalization adjustment is the -8 psi in column B12, which comes from simulations at nominal PB1 and PB4 test conditions (Equation 4.22). Next in magnitude is the average normalization adjustment of -0.96 psi due to temperature normalization. The green box in column B2 shows a zero average bias adjustment for failure pressure measurements error. The green boxes yield a combined bias shift of -71.8 psi from the nominal measured failure pressure of 655 psi in Table 4.1. The resulting average normalized failure pressure is 583.2 psi listed in the pink box at the bottom of Table 4.9.

Potential wall thickness differences from nominal lead to variations in normalized failure pressure characterized by a standard deviation of 39.1 psi. This is significantly larger than the value of 28.1 psi for PB1. This is because of PB4's larger [0.02", 0.024"] wall thickness uncertainty compared to PB1's [0.019", 0.022"] uncertainty. Potential temperature perturbations from PB4 nominal yield a standard deviation of 17.0 psi in normalized failure pressure. This very closely reflects the value of 17.1 psi for PB1 normalization. Column B2's repeat of column A2 of course also here yields a standard deviation of 4.8 psi due to possible pressure measurement error.

The square root of the sum of squares of the contributing standard deviations in the grey boxes in Table 4.9 is 42.99 psi. This closely agrees with the value 42.97 psi in the pink box, calculated directly from the realizations in column B13. This standard deviation is significantly greater than that for PB1 (32.8 psi). This is reflected in the relative widths of the distributions in Figure 4.9.

The variance contributed by wall thickness normalization is here proportionately even greater (compared to the PB1 case) than the variances contributed by temperature normalization and by pressure measurement uncertainty. As in the PB1 case, the largest component of variance due to temperature normalization is contributed by failure temperature uncertainty due to uncertainty in the actual location of failure initiation.

4.3.3 Combining Possible Values of Normalized Failure Pressures to Yield Estimated Ranges of Failure Population Statistics

Columns A12 of Table 4.7 and B13 of Table 4.9 yield, for each row j, a pair of potential failure pressures from replicate experiments normalized to the same experimental input conditions (and to the same input conditions used in the simulations in Section 4.2). Then, hypothetically, if the

realizations in a given row j happen to correspond to exact normalizations for the true input conditions in the tests, the disparities between the two values in that row are not attributable to differences in input conditions between tests or to measurement errors on the outputs. The disparities would then reflect differences between the two tests that were not accounted for in the normalization procedure. In the present case the only apparent element of major importance that was not explicitly normalized is material strength differences at the failure initiation locations in the tests. Thus we attribute any differences in exactly normalized failure pressures to material strength variability between tests.

The failure pressure disparities are a gauge of the material strength variability, just like in the simulated PDFs of failure pressure variability depicted in Figure 4.5 and originating from material variability parameterized by differing stress-strain curves from the material characterization tests (cylinder tension tests). For validation comparisons of experimental failure pressures against the predicted PDFs of failure pressure we form compatible PDFs of experimental failure pressure variability. From the row j (hypothetical) exactly normalized pair of PB1 and PB4 failure pressures we can form a n=2-sample 0.95/0.90 tolerance interval (TI) and associated Normal PDF to compare against the predicted range of 0.95/0.90 TIs and associated PDFs in Figure 4.4. (See Table 2.1 and surrounding text for details on how to construct 2-sample 0.95/0.90 TIs and associated Normal PDFs.)

In our procedure we do not expect that the J=1000 realizations will contain a row j that has perfectly normalized failure pressures for PB1 and PB4. But if the following conditions apply then one or more rows will come arbitrarily close to perfectly normalized failure pressures jointly (simultaneously) for PB1 and PB4.

Conditions

- Parameters explicitly normalized-for contain the actual experimental values within the stated uncertainty ranges.
- The modeling of response (here failure pressure) over the ranges of the normalization parameters is sufficiently accurate that sufficient sampling of the modeled response over the parameter values in the normalization procedure yields one or more rows of normalized PB1 and PB4 failure pressures that are simultaneously closely representative of exactly normalized PB1 and PB4 failure pressures.

In the present analysis it may be that the J=1000 realizations and/or the models used (the physics model, the experimental slope information, the estimated uncertainty ranges, and linear+ Taylor Series model) are not sufficient to meet the stated conditions. Because this is a relatively new methodology we have not yet looked into ways of determining or establishing the said sufficiencies. Nonetheless, we anticipate that several rows of realizations have values that are simultaneously closely representative of exactly normalized PB1 and PB4 failure pressures. We proceed as though this is the case. Then the sought results from exactly normalized PB1 and PB4 failure pressures will lie within the uncertainty ranges defined by the 1000 realizations.

The present analysis produces J=1000 2-sample 0.95/0.90 TIs from the 1000 rows of estimates of PB1 and PB4 normalized failure pressures in columns A12 of Table 4.7 and B13 of Table 4.9. A convenient notional representation of the significance of the 1000 TIs is portrayed by the

associated Normal PDFs depicted notionally in Figure 4.10. The Normal PDFs in the figure only serve as a conceptualization aid; they are not constructed in the course of the analysis. The figure depicts the uncertainty range of the 1000 TI estimates for lower bounds on the 0.025 percentile of response. An uncertainty range is also depicted for the estimated upper bounds on the 0.975 percentile of response. We can compare these uncertainty ranges against the uncertainty ranges for predicted 0.025 and 0.975 percentiles of response from model simulations (Figure 4.5) under the common reference experimental conditions. Such comparisons are made and interpreted in Section 4.4.

Note that the 0.025 and 0.975 percentiles of response are the only response statistics that can be addressed from the current construction. For example, the uncertainty of the means of the depicted Normal PDFs in Figure 4.10 (same as the means of the underlying 0.95/0.90 TIs) are not appropriate to compare against the range of means denoted by the upper and lower simulation PDFs in Figure 4.10. Instead, the 1000 rows of PB1 and PB4 normalized results would need to be processed to create 1000 *n*=2-sample t-distributions. Each such distribution characterizes the uncertainty of a population mean calculated from just two samples of data. So the processing of the PB1 and PB4 normalized failure pressures must be tailored to specific response statistics that are to be compared in the validation assessment. Experimental and simulated 0.025 and 0.975 percentiles of response were chosen as validation quantities in this project because these percentiles combine the effects of small-sample uncertainties in both the response mean and variance quantities, and appear more relevant to the validation assessment of a model that is to be used for design or safety margin predictions.

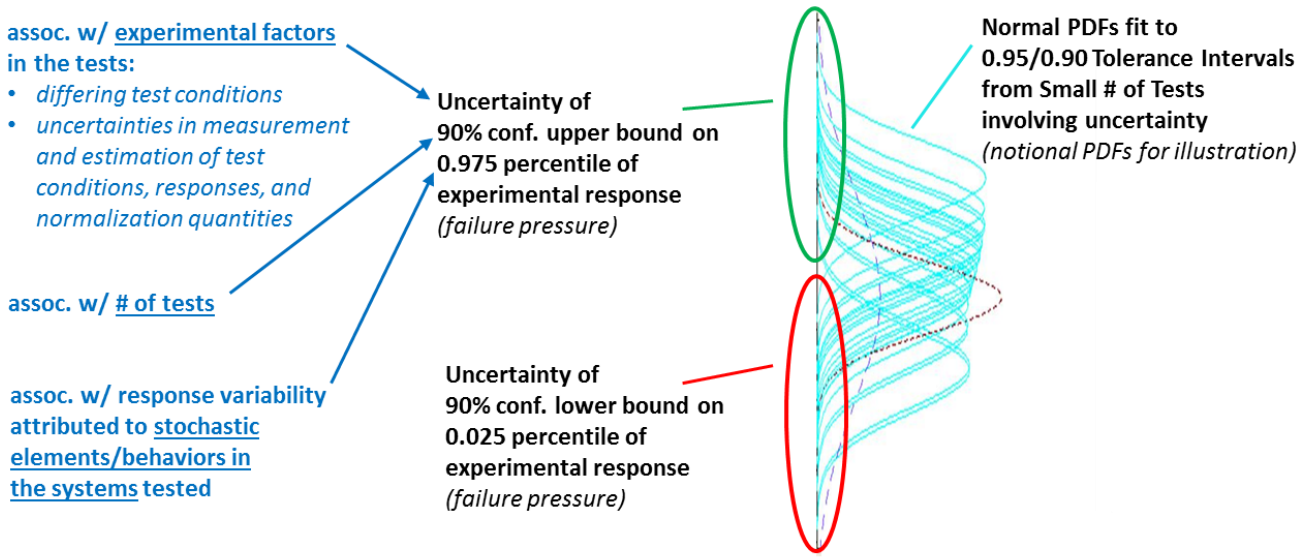

Figure 4.10 Uncertainty of statistical bounds on failure pressure percentiles inferred from a small number of tests normalized to PB1 nominal experimental conditions (accounting for differing conditions in the tests and for random and systematic uncertainties in measurement and estimation of the tests' conditions, responses, and normalization quantities).

4.3.4 Pooling more PB Tests/Results to Reduce Uncertainty of Population Statistics

The results from using the n=2 samples from the PB1 and PB4 tests yield unreasonable uncertainty ranges for the sought percentiles. In particular, a substantial portion of the uncertainty range for the lower percentile (0.025) of response reached into negative (impossible) values of failure pressure. A large contributor to these uncertainty ranges is the large value $f_{0.95/0.90}$ = 18.8 that multiplies the response standard deviation $\tilde{\sigma}$ from two tests to obtain 0.95/0.90 tolerance intervals of half-length = $18.8\tilde{\sigma}$ (see Table 2.1 and surrounding text).

To obtain more statistical precision we note that two similar tests, PB4 and PB10, were performed at 650C-hold conditions with similar pressurization profiles as in the PB1 and PB4 tests. The 650C-hold tests were planned replicates of each other. In Appendix E of [9] these test are described and their results are normalized to the PB1 nominal conditions so they can be pooled with the normalized PB1 and PB4 results to get a population of 4 samples, which decreases the multipler value to $f_{0.95/0.90}$ = 4.94 (Table 2.1), about ¼ the 2-sample magnitude of 18.8.

The 650C-hold tests PB2 and PB10 are similar enough to the 700C-hold tests PBs 1 and 4 that it is reasonable to expect that PB2 and PB10 failure pressures can be normalized to the nominal PB1 test conditions accurately enough that pooling of the four tests' results is justified. A brief summary follows. Extensive details are given in Appendix E of [9].

The peak hot-spot temperatures for PBs 2 and 10 (as indicated by TCs #4 in each test) were ramped at a rate of approximately 31C/min., just as for PBs 1 and 4. When the control TCs #4 in each test reached nominally 650C, this temperature was maintained while the pipes were pressurized until the failure (see Table 4.10). Pressures were measured with the same pressure gage as in tests PB1 and 4. Pressure ramp rates in all four tests 1, 2, 4, and 10 were within a few % of the 1.3 psi/sec target. In PB2 the pressurization started about 40 sec. after 650C was reached. PB10 pressurization started about 30 seconds after 650C was reached. In tests PB1 and PB4 pressurization started respectively about 4 minutes and 1.5 minutes after the target temperature of 700C was reached. These test-to-test differences in pressurization delay after target temperatures were reached are not expected to affect failure pressure levels significantly. Certainly they do not in the model simulations because the modeled physics are agnostic to any such time effects.

Table 4.10 Failure pressures measured in 650C-hold experiments. Pressure measurement uncertainties are explained in Section 4.1.1.

Exper.	Measured pressure at failure	Uncertainty in pressure measurement
PB2	587 psi	± 10 psi
PB10	647 psi	± 10 psi

Figure 4.11 shows the axial and circumferential temperature profiles at failure for PBs 1,2,4, and 10. The temperature profiles are approximately the same shape but are vertically shifted relative to each other in the hot spot region. The principal effects of the different hot-spot temperatures in the tests are approximately normalized out in Appendix E in [9] by assuming that the peak temperature at TC4 determines the failure pressure far more than temperatures away from this peak. Then the temperature vs. failure pressure relationship in Figure 2.5 is used to normalize for different peak (TC4) temperatures. Though this approach is deemed sufficient to support the ultimate conclusions of the validation analysis, it is less accurate than using the physics model as was done above to normalize PB2 failure pressure for its differenent temperature profile vs. PB1 (see Equation 4.22 and column 12 in Table 4.8). But it was determined late in the validation project that working with just PB1 and PB4 tests led to unacceptably large tolerance intervals for many of the realizations of normalized failure pressures. So there was insufficient time to use the simulation model for PB 2 and 10 analogues of Equation 4.22 given the model's ~month-long run times on 800 processors.

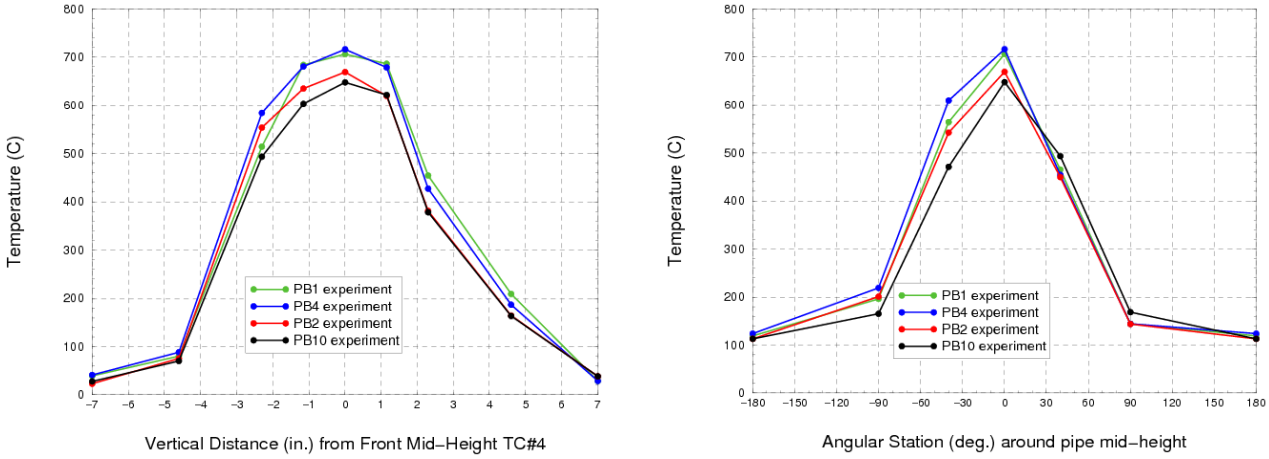

Figure 4.11 Pipe-bomb axial (left plot) and circumferential (right plot) TC temperatures at time of failures in the 650-hold and 700C-hold experiments.

Table 4.3 lists minimum and maximum measured pipe-wall thicknesses for tests PB2 and PB10. These min to max ranges of thickness are less than the ranges for PBs 1 and 4. Therefore, significantly smaller normalization variance exists for wall-thickness normalizations of PBs 2 and 10 than for PBs 1 and 4.

Tables E.3 and E.5 in Appendix E of [9] list the PB2 and PB10 normalization quantities and their uncertainties. The structure of these normalization tables is the same as for PB4's Table 4.8. Columns 2, 5, 8, and 10 in tables E.3 and E.5 are highlighted yellow to signify their correlation with the same-numbered columns in Tables 4.6 and 4.8 for PBs 1 and 4. Different initial seeds are used to sample all column quantities for PBs 1, 2, 4, and 10 except for the yellow-highlighted columns.

The uncertainties in column 5 for the various tests and perfectly correlated but not necessarily the same; the uncertainty ranges in column 5 for PBs 2 and 10 are scaled versions of those in column 5 for PBs 1 and 4. The scaling reflects that the relation between failure pressure and wall thickness, for example, depends on wall temperature and PBs 2 and 10 occur at nontrivially different temperatures than PBs 1 and 4. For analogous reasons the thermocouple measurement uncertainties in columns 8 are also correlated but scaled. Appendix E in [9] describes the scaling procedures.

Tables E.4 and E.6 in Appendix E in [9] list sample realizations from the normalization of PBs 2 and 10. Columns 13 contain the realizations of the normalized failure pressures. The realizations are histogrammed in Figure 4.9. In the figure the normalized 650C-hold PB 2 and 10 results are lower by 50 – 100 psi on average than the normalized PB 1 and 4 results. It would not be surprising to get a systematic difference between the normalized 650C-hold PB 2 and 10 results vs. the normalized 700C-hold PB 1 and 4 results. The PB 2 and 10 tests involve pipe temperature fields that are more than a small perturbation away from the reference PB1 conditions. Moreover, normalization for this large difference was not as accurate as it could have been if project resources would have allowed running the physics model at the PB2 and PB10 input conditions. Instead, a less accurate approach considered only the hottest TC's temperature (TC4) and normalized using the relationship in Figure 2.5, as described above and in more detail in Appendix E in [9].

The normalization average shift in PB10 experimental failure pressure is -116.5 psi, due predominantly to adjustment for PB10's temperature differential vs. the reference PB1 conditions. The average adjustment for PB2 is -103.4 psi, also due mostly to temperature differential. In comparison, considerably smaller average normalization adjustments exist of -72 psi for PB4 and -20 psi for PB1, due mostly to their wall-thickness uncertainty ranges not being centered on the nominal 0.02" value.

PB2 normalization uncertainty has a standard deviation σPB2 = 27.2 psi. This is slightly less than σPB10 = 29.0 psi. PBs 1 and 4 have the next larger normalization uncertainties characterized by σPB1 = 32.8 psi and then σPB4 = 43.0 psi. For all PBs the normalization uncertainty is driven most by the combined uncertainty of wall thickness and its effect factor $\frac{\partial(Pfail)}{\partial(w)}$. In fact, the wall-thickness input uncertainties have the same ordering according to magnitude: smallest for PB2, next PB2, then PB1, and PB4 largest. For all PBs, the actual failure location's potential temperature differential from PB4 of [-15, 15]C in column 11, combined with the uncertain temperature effect factor $\frac{\partial(Pfail)}{\partial(T@fail\_point)}$, are also very significant and contribute 2nd most to the normalization uncertainty.

The J=1000 realizations for each of the four PB tests normalized to the reference PB1 experimental conditions are portrayed by the histograms in Figure 4.9. From these, J=1000 TIs are constructed: a 4-sample 0.95/0.90 TI is constructed for each row of estimates of normalized failure pressures in column A12 of Table 4.7 and columns 13 in Tables 4.9, E.2, and E.4. Again, a convenient notional representation of the significance of the 1000 TIs is portrayed in Figure 4.10. Uncertainty bands of the 1000 lower ends and 1000 upper ends of the constructed TIs are portrayed. Actual values defining such uncertainty bands for the present problem are binned in

Table 4.11. Corresponding histograms are shown at right in Figure 4.12. In the next section we compare and interpret these uncertainty ranges against the uncertainty ranges for predicted 0.025 and 0.975 percentiles of response from model simulations (Figure 4.5).

Table 4.11 Binned realizations of 0.025 and 0.975 percentiles bounds from 0.95/0.90 TIs constructed from realizations of normalized failure pressures from tests PB 1, 2, 4, 10

bins, 2.5%ile (μ - 4.94σ, 4-sample 95/90 Tol. Intvl.)	*Frequency*	*Cumulative %*
0	2	0.20%
50	2	0.40%
100	15	1.90%
150	67	8.61%
200	120	20.62%
250	208	41.44%
300	218	63.26%
350	202	83.48%
400	116	94.10%
450	40	99.10%
500	9	100.00%

bins, 97.5%ile (μ + 4.94σ, 4-sample 95/90 Tol. Intvl.)	*Frequency*	*Cumulative %*
570	0	0.00%
630	11	1.10%
690	65	7.61%
750	135	21.12%
810	214	42.54%
870	265	69.07%
930	183	87.39%
990	96	97.00%
1050	26	99.60%
1110	4	100.00%

4.4 Model Validation Comparisons and Observations

In Figure 4.12 the histograms reflect uncertainties in the upper and lower endpoints of 0.95/0.90 tolerance intervals from the four test results normalized to nominal PB1 input conditions. As explained earlier, the interval uncertainties input to the normalization operation are temporarily represented as uniform PDFs and sampled accordingly. If the interval uncertainties were instead propagated as intervals, the lower and upper interval limits on the 0.975 percentile of response would extend somewhat beyond the range of the red histogram in Figure 4.12. But the joint extremes of the uncertainty ranges that would produce these interval limits are considered exceedingly unlikely. Instead the 0.05 and 0.95 percentiles of the red histogram are chosen to represent a more reasonable range of uncertainty regarding the 0.975 percentile of response (failure pressure). The corresponding red interval in the figure represents the moderated interval uncertainty used in the following. Similar considerations underlie the rightmost green interval in the figure.

The red and green intervals at left in Figure 12 are obtained from the 0.025 and 0.975 percentiles of the upper and lower Normal PDFs of predicted response in Figure 4.4. Consider the green intervals in Figure 4.12 for 0.025 percentiles of predicted and experimental response. The green intervals do not overlap. Therefore it is straightforward that, for this lower percentile of response, the model predicts higher failure pressures than inferred from testing. If, for instance, this lower percentile of response is written into a design or safety spec that <2.5% of pipes of this design

are to fail under applied pressure and temperature conditions emulated in the tests, then the experiments are indicating a lower failure pressure for 2.5% of pipes than the model is predicting. The model therefore gives unconservative predictions for these circumstances.

How these results extrapolate to other applications of the constitutive model (different pressure vessel geometries, heating conditions, wall thicknesses, etc.) is a very difficult issue and beyond the scope of this project. But with some reasonable assumptions and a little more analysis one could take the results in Figure 4.12 and extend them to cases where the same pipe design and experimental conditions exist but the spec has lower allowable percentages of failure like 1% or 0.001%. We could tentatively conclude similarly that the model would be unconservative for those spec regimes as well. With less assumptions one could reprocess the experimental and simulation data to produce a figure like Fig. 4.12 for the different percentiles and confidence levels of interest. This would be a more definitive assessment of conservatism or not, and by how much, for the other percentiles of interest.

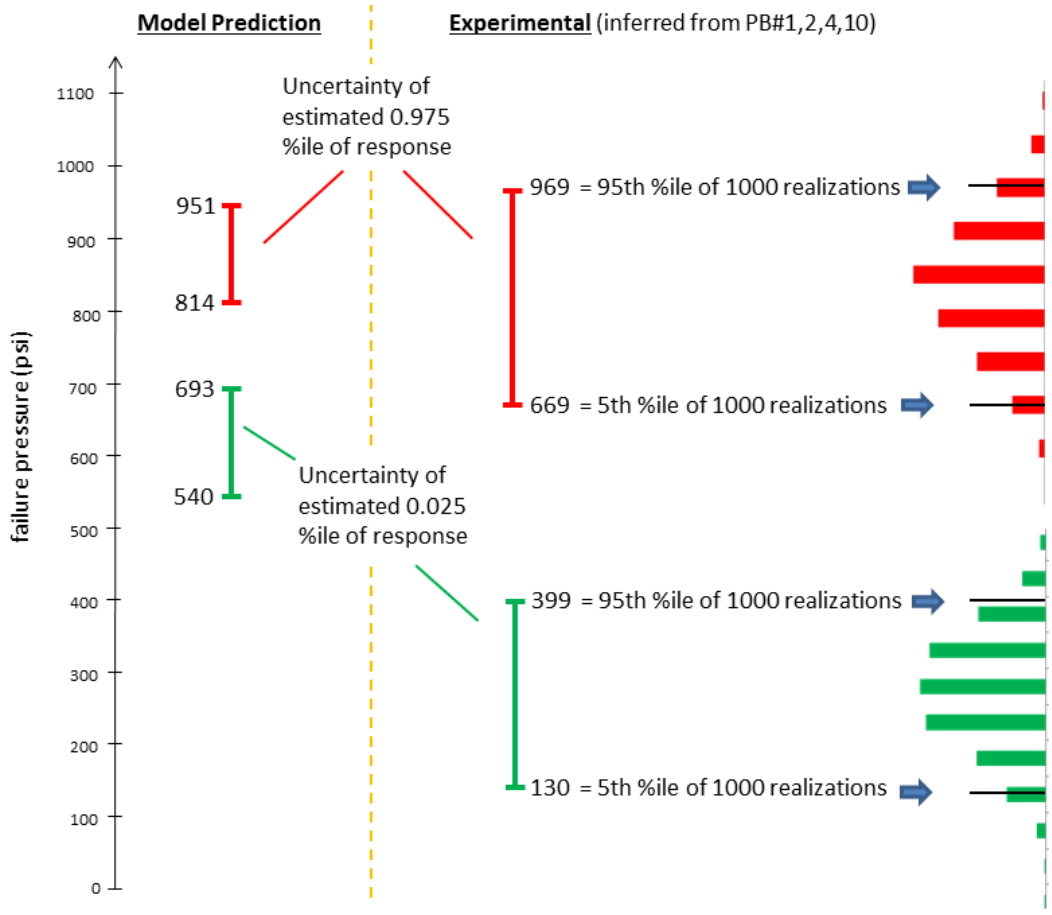

Figure 4.12 Model validation comparison of uncertainty ranges of simulation percentiles of response and normalized experimental percentiles of response.

Now consider the red intervals in Figure 4.12 for 0.925 percentiles of response. These intervals overlap and the experimental and simulation uncertainties they represent are statistically independent. Therefore there are numerous possibilities that the experimental 0.925 percentile of response is higher than the predicted 0.925 percentile, and vice versa. So we cannot conclude, as was done for the 0.025 percentile, that the predictions are unconservative (or alternatively that they are conservative). One limiting case for these ranges of uncertainty is that the predicted 0.925 percentile is as high as 951 psi as labeled in the figure, while the experimental percentile is as low as 669 psi, as labeled. In this limiting case the predicted 0.925 percentile is up to 282 psi higher than the experimental percentile. The opposite limiting possibility is that the predicted 0.925 percentile is as low as 814 psi, while the experimental percentile is as high as 969 psi. Then the predicted 0.925 percentile is as much as 155 psi lower than the experimental percentile.

Thus the uncertainty U{predicted 0.925 percentile minus experimental 0.925 percentile} = [-155, 282] psi = the model's range of potential prediction bias. Information in this form accounts for all the experimental and simulation non-traveling uncertainties combined, which defines the total uncertainty resolution limit below which the model's exact bias cannot be determined (similar to the concept introduced in [30] and adopted in [28] without the recognition that they consider only non-traveling uncertainties in their validation paradigm). How to best use the model's validation-characterized bias uncertainty to mitigate prediction risk in use of the model in settings other than the validation conditions is a very difficult question and an area of active research (see e.g. [6], [8], [20], [22]). See also [31] – [33] for different extrapolation approaches based on other representations of model bias information.

As a final observation, correlation exists among the realizations that define the ranges of the red and green intervals on the simulation side in Figure 4.12. Correlation (but weaker) also exists among the realizations that define the ranges of the red and green intervals on the computational side. Given the correlation structure there are many more possibilities (potential realizations) that the distance between the 0.025 and 0.975 predicted percentiles is less than the distance between the 0.025 and 0.975 experimental percentiles. This tends to indicate that the predicted variance of failure pressures is less than the experimental variance. However, arriving at a firmer quantitative conclusion requires processing the experimental and simulation data each for variance and uncertainty thereof, and then compare the results. Likewise, the relative positioning of uncertainty ranges tends to support a conclusion that mean experimental failure pressure is lower than the mean predicted failure pressure. But one would have to process for mean estimates and uncertainty thereof in order to form a firmer conclusion.

5. Closing Remarks

A pragmatic and novel Real Space model validation methodology has been presented that is geared for:

- very expensive computational models (minimal number of function evaluations);
- quantification and economical management of mesh and solver discretization effects;
- rollup of various types, sources, and representations of uncertainty;
- sparse experimental data;
- multiple replicate experiments;
- stochastic phenomena and models.

The validation approach and metrics:

- segregate aleatory and epistemic uncertainties in the validation activity;
- are relatively straightforward to interpret;
- are especially suited for assessing models and prediction quantities to be used in the analysis of performance and safety margins.

By analogy with the presented example the reader should be able to treat a large variety of model validation applications, issues, and constraints that arise in industrial practice.

In the PB validation problem the largest uncertainty contributors are the sparseness of repeat experiments at the pipe level, followed by solution uncertainty (discretization effects), and then experimental variations and uncertainties in the tests.

To close, some observations are made below concerning the applicability of several other validation methodologies to the pipe bomb validation problem. The Real Space approach appears to uniquely have the required features to handle all the attributes of the PB validation problem.

- The ASME VV20 methodology [28] is geared for validating deterministic models, so does not address the lower two uncertainty elements at left in Figure 10 (confidence levels associated with limited # of tests and percentiles or other measures of stochastic system behavior/response). The other uncertainty element in the figure (experimental factors in the tests) is addressed equivalently by the ASME VV20 and Real Space methods under the restricted problem conditions stated two paragraphs above the start of section 4.3.1. This is a reassuring corroboration of both the RS and ASME VV20 methods for the subset of conditions cited, given that their derivations come from very different conceptual approaches.

- The Oberkampf and Barone approach [32] addresses the uncertainty elements at left in Figure 10 for stochastic system behavior/response and confidence levels from limited # of tests. But the measure of the stochastic behavior in their validation metric is limited to uncertainty of mean response. The percentile measures of stochastic behavior in the RS method reflect uncertainty of both mean and variance of stochastic response, and seem more relevant for assessing the adequacy of models for engineering use purposes involving predictions of stochastic system behavior/response. Furthermore, [32] does not address most of the uncertainties in the third category (experimental factors in the tests).

Systematic experimental uncertainties are ignored altogether. Regarding model predictions, a single deterministic model run is made to compare the result against the uncertainty PDF of the experimental mean (an inadequate comparison basis for assessing accuracy and adequacy of models of stochastic phenomena).

- The validation approaches in [33] and its derivative [34] are geared for models of stochastic phenomena or systems. The "area" validation metric compares full PDFs of predicted and experimental responses. A numerical value for discrepancy between experimental and predicted PDFs is obtained. A value of zero indicates perfect agreement at all PDF percentiles. However, for non-perfect agreement it is not clear how non-zero values relate to more directly interpretable measures of prediction error like mean prediction error, error of predicted variance, or error of predicted percentiles of behavior. Furthermore, the methodology does not address uncertainty effects from limited data in a standard context of statistical confidence levels. The remaining element at left in Figure 4.10 (experimental factors in the tests) is only partially addressed. Systematic experimental uncertainties are not treated, leading to more *Model User's Risk* concerning these uncertainties than the RS and ASME VV20 approaches do ([8]). The latter take a more conservative slant, mitigating *Model User's Risk* by explicitly accounting for systematic measurement uncertainties. Also, the approaches in [33] and [34] treat random variations of input uncertainties over multiple replicate experiments but assume that random measurement errors are small and are not a significant source of the variations. In contrast, the RS approach explicitly estimates and propagates potential random and systematic measurement uncertainties of inputs and outputs, figuring that these are important to account for in the model validation assessment.

- Two validation approaches are presented in [35]. One is apparently an implementation of the approaches in [33] and [34], while the other does not address stochastic elements of behavior in the model predictions, so is very limited.

- Finally, in contrast to the other validation approaches mentioned here, the Real Space methodology recognizes that models can have *traveling* epistemic uncertainties that are an intrinsic aspect of the model. These are not in the present validation problem but occur e.g. in [22] and [36] as uncertainties in physics modeling parameters, and in [23] as different turbulence submodels. These are treated differently from the non-traveling systematic (epistemic) uncertainties contributed by testing because they have different implications for model predictivity in post-validation use of the model ([6], [8] pp. 50]).

References

[1] Dempsey, J.F., B. Antoun, V. Romero, G. Wellman, W. Scherzinger, S. Grange, "Temperature Dependent Ductile Material Failure Constitutive Modeling With Validation Experiments," SEM XII International Congress & Exposition on Experimental and Applied Mechanics, June 11-14, 2012, Costa Mesa, CA.

[2] Antoun, B., Sandia National Laboratories C6 L3 Milestone Report: "Material Characterization and Coupled Thermal-Mechanical Experiments for Pressurized, High Temperature Systems," September 11, 2009.

[3] Wellman, G. W., "A Simple Approach to Modeling Ductile Failure," Sandia National Laboratories report SAND2012-1343 printed June 2012.

[4] Scherzinger, B., and G. Wellman, "An Elastic-Plastic Constitutive Model Incorporating Temperature Dependence," Sandia National Laboratories draft document.

[5] Adagio 4.20 User Guide, Sandia National Laboratories report SAND2011-1825, printed March 2011.

[6] Romero, V.J., "Type X and Y Errors and Data & Model Conditioning for Systematic Uncertainty in Model Calibration, Validation, and Extrapolation," SAE paper 2008-01-1368 for Society of Automotive Engineers 2008 World Congress, April 14-17, 2008, Detroit, MI.

[7] Romero, V.J., "Comparison of Several Model Validation Conceptions against a "Real Space" End-to-End Approach," *Soc. Automotive Engrs. Intn'l. J. of Materials and Manufacturing*, June 2011.

[8] Romero, V.J., "Elements of a Pragmatic Approach for dealing with Bias and Uncertainty in Experiments through Predictions: •Data and Model Conditioning; •"Real Space" Model Verification and Conditioning; •Hierarchical Modeling and Extrapolative Prediction," Sandia National Laboratories report SAND2011-7342, Nov. 2011.

[9] Dempsey, F., G. Wellman , W. Scherzinger, V. Romero, B. Antoun, "UQ and V&V Techniques Applied to a Constitutive Model tested on Heated Pipes Pressurized to Failure," Sandia National Laboratories report in review.

[10] Romero, V., J. Mullins, L. Swiler, A. Urbina, "A Comparison of Methods for Representing and Aggregating Experimental Uncertainties involving Sparse Data—More Results," *Soc. Automot. Engrs. Int. J. of Materials and Manufacturing*, 6(3):2013, doi:10.4271/2013-01-0946

[11] Romero, V., L. Swiler, A. Urbina, J. Mullins, "A Comparison of Methods for Representing and Aggregating Uncertainties involving Sparsely Sampled Random Variables – Final Results," 15th AIAA Non-Deterministic Approaches Conference, April 8-11, 2013, Boston, MA.

[12] Romero, V., L. Swiler, A. Urbina, J. Mullins, "A Comparison of Methods for Representing Sparsely Sampled Random Quantities," Sandia National Laboratories report SAND2013-4561 printed September 2013.

[13] Romero, V., J.F. Dempsey, G. Wellman, B. Antoun, "A Method for Projecting Uncertainty from Sparse Samples of Discrete Random Functions ─ Example of Multiple Stress-Strain Curves," 14th AIAA Non-Deterministic Approaches Conference, April 23-26, 2012, Honolulu, HI.

[14] Kambour, K., C. Hembree, E. Keiter, "Calibration of Xyce Neutron Models for the QASPR Complex Prototype Circuit," SAND2010-7988 (Official Use Only/Export Controlled), printed December 2010.

[15] Pradlwarter, H.J., and G.I. Schuëller, "The use of kernel densities and confidence intervals to cope with insufficient data in validation experiments," *Computer Methods in Applied Mechanics and Engineering*. Vol. 197, Issues 29-32, May 2008, pp. 2550-2560.

[16] Romero, V., B. Rutherford, J. Newcomer, "Some Statistical Procedures to Refine Estimates of Uncertainty when Sparse Data are Available for Model Validation and Calibration," paper AIAA-2011-1709, 13th AIAA Non-Deterministic Approaches Conference, April 4-7, 2011, Denver, CO.

[17] Trucano, T.G., M. Pilch, W.L. Oberkampf, "General Concepts for Experimental Validation of ASCI Code Applications," Sandia National Laboratories Report SAND2002-0341, printed March 2002.

[18] American Society of Mechanical Engineers, V&V 10 – 2006, *Guide for Verification and Validation in Computational Solid Mechanics*, available from ASME Codes & Standards website.

[19] Romero, V.J., M.P. Sherman, J.F. Dempsey, J.D. Johnson, L.R. Edwards, K.C. Chen, R.V. Baron, C.F. King, "Development and Validation of a Component Failure Model," paper AIAA-2005-2141, 45th AIAA/ASME/ ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, April 18-21, 2005, Austin, TX.

[20] Romero, V.J., "Validated Model? Not So Fast. The Need for Model 'Conditioning' as an Essential Addendum to Model Validation," paper AIAA-2007-1953, 9th Non-Deterministic Approaches Conference, Honolulu, HI, April 23-26, 2007.

[21] Rutherford, R., V. Romero, J. Castro, R. Hoekstra, "Methods Used in Uncertainty Quantification, Calibration, Validation, and Prediction for the QASPR Silicon Device Prototype Demonstration," Sandia National Laboratories report SAND2011-7940 (Official Use Only) printed October 2011.

[22] Romero, V., B. Rutherford, J. Castro, "Model Calibration, Validation, and Extrapolative Prediction with Segregated Aleatory and Epistemic Uncertainties—Demonstration on Prototype Silicon Device," Sandia National Laboratories report in preparation.

[23] Romero, V.J., A. Luketa, M. Sherman, "Application of a Versatile "Real Space" Validation Methodology to a Fire Model" *AIAA J. of Thermophysics and Heat Transfer*, Vol. 24, No. 4, Oct. – Dec. 2010, pp. 730-744.

[24] Ricks, A., V. Nicolette, V. Romero, W. Erickson, "Fuego Solid-Propellant Fire Model Verification and Validation," Sandia National Laboratories report in review.

[25] J. T. Nakos, "Uncertainty Analysis of Thermocouple Measurements Used in Normal and Abnormal Thermal Environments Experiments at the Radiant Heat Facility and the Lurance Canyon Burn Site", Sandia National Laboratories Report SAND 2004-1023, Released April 2004.

[26] Nakos, J.T., J.M. Suo-Anttila, W. Gill, "Shroud Boundary Condition Characterization Experiments at the Radiant Heat Facility," Sandia National Laboratories report SAND2004-5080 printed Oct. 2004.

[27] Romero, V.J., "Data & Model Conditioning for Multivariate Systematic Uncertainty in Model Calibration, Validation, and Extrapolation," paper AIAA-2010-2511, 12th AIAA Non-Deterministic Approaches Conference, April 12-15, 2010, Orlando, FL.

[28] American Society of Mechanical Engineers, V&V 20 – 2009 *Standard for Verification and Validation in Computational Fluid Dynamics and Heat Transfer*, available from ASME Codes & Standards website.

[29] Antoun, B., K. Connelly, "Pipe Bomb Experiments for Abnormal Thermal-Mechanical Project" project review slides (Official Use Only), Sandia National Laboratories, Feb. 6, 2013.

[30] Coleman, H.W., and Stern, F., "Uncertainties in CFD Code Validation," *Journal of Fluids Engineering*, Dec. 1997, vol. 119, pp. 795-803.

[31] Hills, R.G., "Roll-up of Validation Results to a Target Application," Sandia National Laboratories report SAND2013-7424, printed Sept. 2013.

[32] Oberkampf, W.L., and Barone, M.F., 2004, "Measures of Agreement between Computation and Experiment: Validation Metrics," AIAA 34th Fluid Dynamics Conference, Portland, OR, June 2004.

[33] Ferson,S., W.L. Oberkampf, L. Ginzburg, "Model Validation and Predictive Capability for the Thermal Challenge Problem," *Comput. Methods in Applied Mechanics and Engrng.*, Vol. 197, 2009, pp. 2408 – 2430.

[34] Oberkampf, W.L., and Roy, C.J., *Verification and Validation in Scientific Computing*, Cambridge University Press, 2010.

[35] American Society of Mechanical Engineers, V&V 10.1-2012 – *An Illustration of the Concepts of Verification and Validation in Computational in Solid Mechanics*, issued April 16, 2012, available from ASME Codes & Standards website.

[36] Black, A., V. Romero, G. Wellman, A. Dodd, N. Breivik, J.F. Dempsey, "Predictive Capability Assessment Project: Abnormal Thermal-Mechanical Breach Plan," Sandia National Laboratories report SAND2012-9246 (Official Use Only), Oct. 2012.