SAND2014-4870P

Chapter Z for Joint Army/Navy/NASA/Air Force (JANNAF) *Guide to UQ, V&V, and Simulation Credibility for Continuum Physics Applications*

Case Study Example:

Application of UQ and V&V to Experiments and Simulations of Heated Pipes Pressurized to Failure[1]

Vicente Romero
Verification, Validation, Uncertainty Quantification, and Credibility Processes Dept.

J. Franklin Dempsey
Component Science and Mechanics Dept.

Bonnie Antoun
Mechanics of Materials Dept.

Sandia National Laboratories[2]
P.O. Box 5800
Albuquerque, New Mexico 87185-0825

Abstract

This chapter demonstrates versatile and practical model validation and uncertainty quantification techniques applied to the accuracy assessment of a computational model of heated steel pipes pressurized to failure. The "Real Space" validation methodology segregates aleatory and epistemic uncertainties to form straightforward model validation metrics especially suited for assessing models 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 (discrete temperaturedependent stress-strain curves);
- very limited (sparse) experimental data 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] This manuscript is a work of the United States Government and is not subject to copyright protection in the U.S.
[2] Sandia 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.

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 on the pipes. Figures 1.1 and 1.2 portray some of the associated "pipe bomb" (PB) hardware and experiments and coupled thermal-mechanical modeling. Only details of PB geometry and experimental conditions and results important to the model validation procedures and comparisons in this chapter are presented herein. Further details of the design of the hardware and experiments, and execution and results of the larger set of experiments, are given in [2].

Controlled nitrogen pressurization of the pipe is accomplished via pressure supply tanks. The pipe is heated by a hot inconel plate, creating a hot spot on the pipe. The pipe is pressurized until it bursts at the hot spot, tearing back along upper and lower thickness-transition shoulders in the pipe, leaving "butterfly wings" as shown in Figure 1.1. The 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 shoulder-region thicknesses of nominally 0.05 in. For safety reasons the inner slug (shiny silver) fills up most of inside volume of the pipe, lessening the explosive energy built up prior to pipe breach failure.

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 the temperature range of interest, nominally 25C to 800C. Information on the mathematical and algorithmic formulation of the temperature-dependent multilinear elastic-plastic (MLEP) constitutive model for material behavior is available from [3], [4]. The constitutive model is used within the Sierra solid mechanics code [5] (massively parallel 3-D implicit nonlinear quasi-statics) to model pipe response in the tests.

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

For each measured stress-strain curve, optimization techniques were used to solve the inverse problem of obtaining best-fit transformed stress-strain curves for the constitutive model. The inverse problem and solution procedures and results are briefly described in Section 2 and are more fully documented in [3].

For uncertainty quantification (UQ) purposes it is important to note that each stress-strain curve comprises a ***discrete*** random function that has no readily identifiable parametric relationship to other stress-strain curves at that temperature. A novel 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.

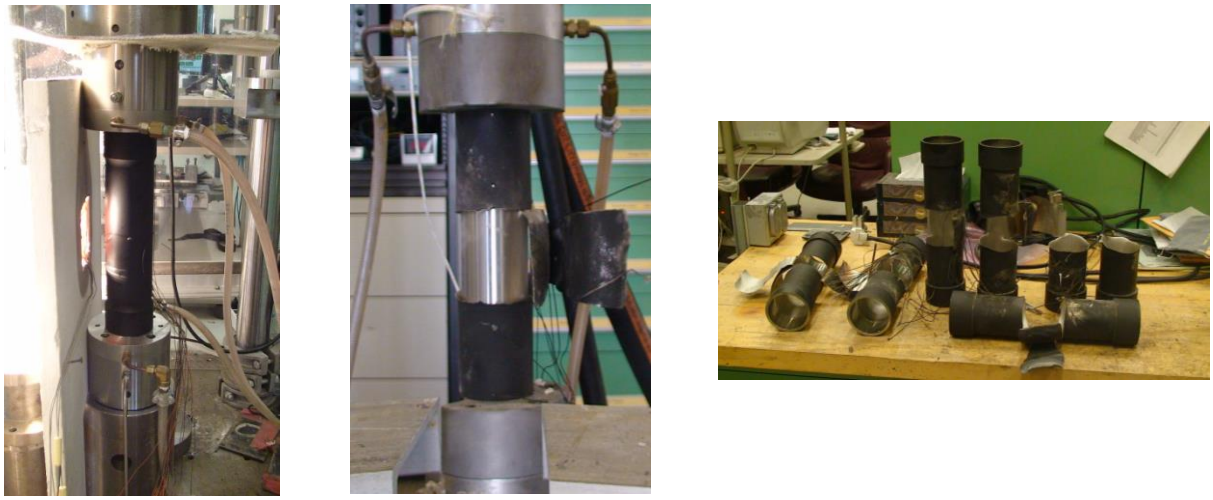


Figure 1.1 Model validation experiments.

coupled Sierra Thermal+Solid-Mechanics simulation model used to help design heating configuration and thermocouple locations

pipe coated black with Pyromark paint for repeatable control of emissivity (& thus heating)

quartz lamp boxes heat back side of shroud (glowing orange)

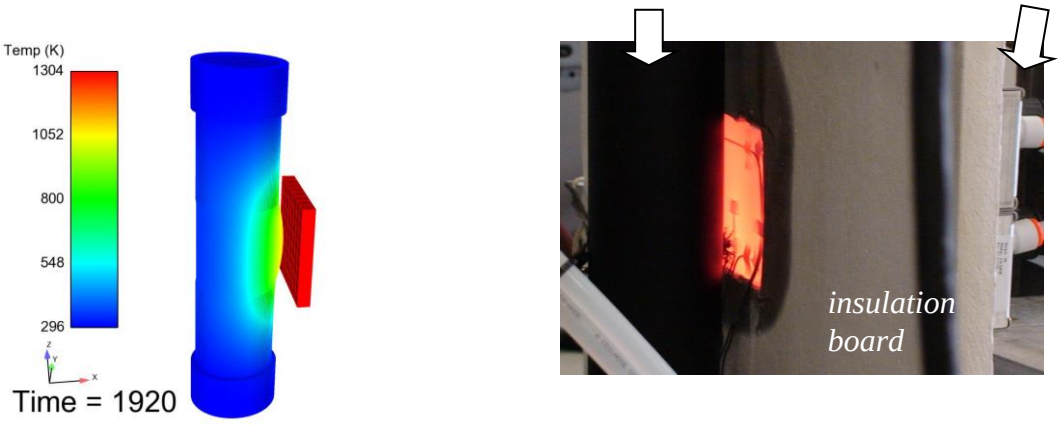


Figure 1.2 Thermal modeling helped design, instrumentation, and analysis of experiments. Thermocouples with wire leads can be seen on the inconel heating shroud (glowing orange).

Section 3 describes the project's finite-element (FE) models, geometries, mesh and solver choices, and calculation verification studies to characterize discretization related solution errors. The FE model and simulations employed in the constitutive model material characterization/ inversion procedure are described. These simulations emulate the cylinder test specimens' response, through deformation ("necking") and failure in the tension tests. Various versions of the heated pressurized pipe models and simulations employed in the study are also described.

Section 4 describes the use of modeling and simulation to help design the PB validation 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 5 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 interpretation of the comparisons are provided.

Section 6 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 rod tensile tests and results from which the stress-strain curves for the constitutive model were derived. Several replicate tests at each of the temperatures indicated were conducted, through necking and failure of the rod specimens.

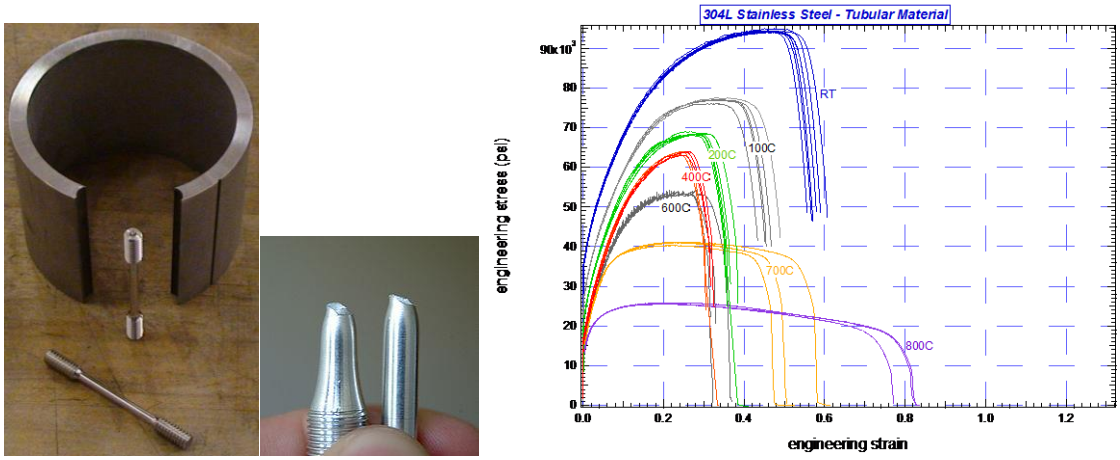


Figure 2.1 Rod tensile test material samples and measured stress-strain curves at the labeled temperatures. (Note: "RT" in the plot stands for "room temperature", nominally 20C).

The cylindrical material samples were cut and machined from the same 304L stainless steel tubular stock (3.5 in. dia., ¼ in. wall thickness) that the validation-experiment pipes were machined from. The rod specimens are 3 in. long and have a long thin "gage" section of diameter 1/8 inch (see [2] for detailed drawings). The specimens were vacuum annealed at 1000C for 30 minutes to produce the same anneal conditions present in the pipe vessels. The ends of the cylinders were displaced in axially opposing directions to produce axial tension at an engineering strain rate of 0.001/s, slow enough to be considered quasistatic and estimated from hand calculations to be representative of pipe wall membrane strain rates 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. More details of the testing apparatus, experimental conditions, measurement and control instrumentation and calibration, etc., are given in [2].

2.2 Inverse calculations to convert from measured stress-strain curves to constitutive model stress-strain curves

For each measured "engineering" stress-strain curve, optimization techniques were used to solve the inverse problem of determining the constitutive model's corresponding "true" stress-strain curve ([3]). This enables a FE model of the cylinder to reproduce the measured engineering stress-strain response in the tension tests. An example is presented in Figure 2.2. 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 inverse-problem solution procedures and results for this project are more fully documented in [18].

The solid mechanics code Adagio[5] was used for the FE model forward simulations in the inversion procedure. In section 3.1 a discussion of the cylinder model FE mesh and solver discretization choices indicates that model results are insensitive to significant perturbations from 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.

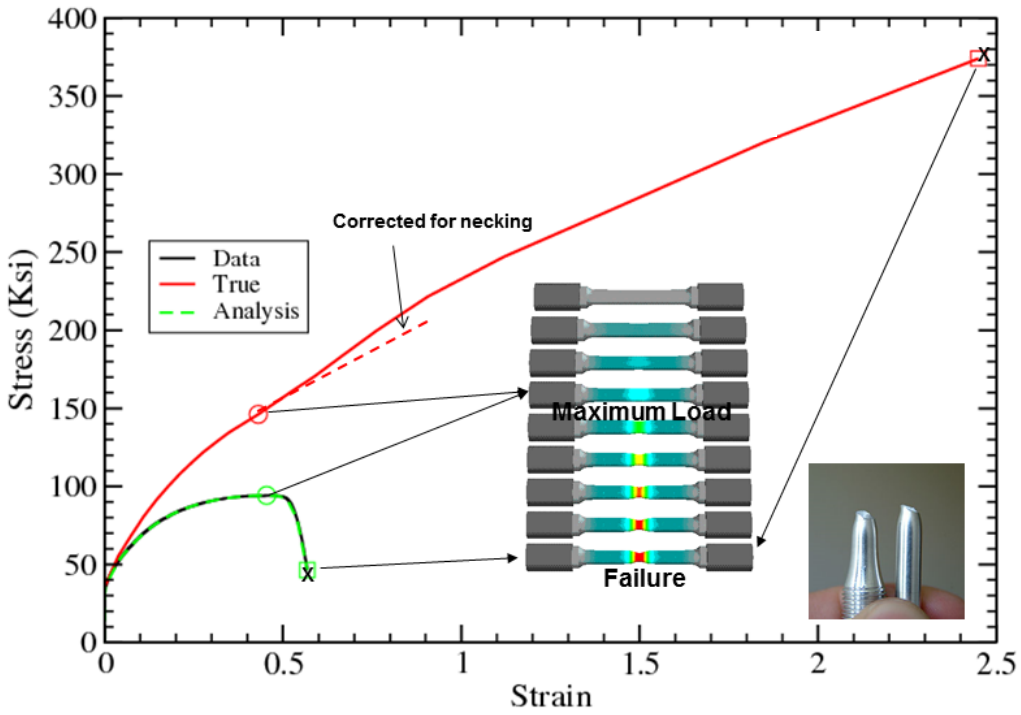


Figure 2.2 Relationship between experimental and constitutive-model stress strain curves obtained from a tensile test. The red curve is the Cauchy-Stress/Logarithmic-Strain "True-Stress/True-Strain" curve appropriate for the constitutive model. The black curve is the engineering-stress/engineering-strain curve from the test. The red curve is inversely calculated such that its use with the constitutive model in FE simulations of the cylinder tensile test yields the calculated green stress-strain curve which closely matches the experimental (black) stress-strain curve.

The material inversion procedure involves many steps and decision points where analyst judgment is used. Therefore the sensitivity of the inversion procedure results to different analysts applying the procedure was checked. We asked another seasoned solid mechanics analyst to see if they could reproduce one of the 700C True-Stress/True-Strain curves that was originally obtained by the author of [3] who retired a short time after generating the material curves for our project. The second analyst was familiar with the inversion procedure from applying it on other projects, but otherwise this was a relatively independent spot-check on our derived material curves. The differences were insignificant below ~230% strain (see [18]), which far exceeds the

point of pipe structural failure initiation in our calculations. We presume that the relevant portions of the other True-Stress/True-Strain curves used in our project are also adequately immune to analyst judgment particulars in the material inversion process.

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, substantial epistemic "sampling 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 underestimating the true full-population variance (at least for distribution types and combinations investigated in [9] – [11]). This is unconservative and therefore undesirable for many engineering purposes. If a structure or pressure-vessel model were perfect in every other way, use of the constitutive model would likely underestimate the (strength or displacement) response variance of the real system. In design and risk analysis one would normally want to avoid such underestimation.

Two approaches were tried to avoid underestimation of the larger-population variability. The first approach assumed that the variability of stress-strain curves at a given temperature can be parameterized. Then the parameter range corresponding to variations between the curves can be appropriately increased to correct for small-sample underestimation bias. The adjusted parameter range representing material curve variability could then be propagated through the pipe response model along with the other parametric uncertainties in the model. However, the attempted parametric representation of material curve variability failed an important "sanity test" for physical consistency as shown in [18].

Consequently, a different approach was devised and implemented. The uncertainty associated with material curve variability is decoupled from the other (parametric) uncertainties in the problem. It is represented and propagated alone, as described here. After propagation it is combined with the other (propagated) parametric uncertainties as described in section 5.2.

The approach treats the stress-strain curves at a given temperature as discrete random functions with no readily identifiable parametric relationship between them. Yet the approach recognizes that the stress-strain curves 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 mitigates chances of underestimating the full-population variability with relatively few data samples. Figure 2.3 conveys the approach for a random variable, but the idea also applies to random functions.

Employing the approach at right in Figure 2.3, 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. Tolerance Intervals 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. [12]). 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 2.3.

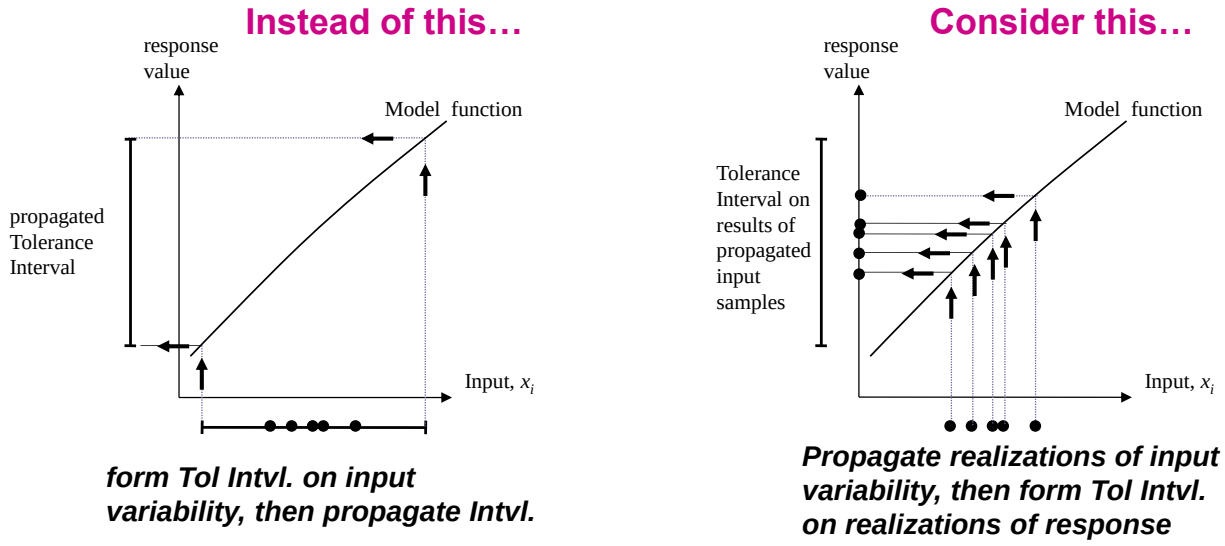


Figure 2.3 Two ways of propagating uncertainty from sparse samples of an input random variable. In the approach at left the sparse samples of the input quantity are fit with a tolerance interval (explained below) that is then propagated to an uncertainty of response. (Alternatively, an equivalent Normal PDF fitted to the tolerance interval is propagated.) The approach at right individually propagates each sample of the input quantity and then forms a tolerance interval (and equivalent Normal PDF, see Figure 2.6) from the propagated results.

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 [9]-[11]. It was found that a classical statistical Tolerance Interval approach (e.g. [13]) provides reliably conservative estimates of the combined epistemic and aleatory uncertainty associated with limited data. The approach is also very easy to use. The approach [14] 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 due to limited samples of data. 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 2.4 reveal that 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. (Selected results from tables in [13].)

# 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

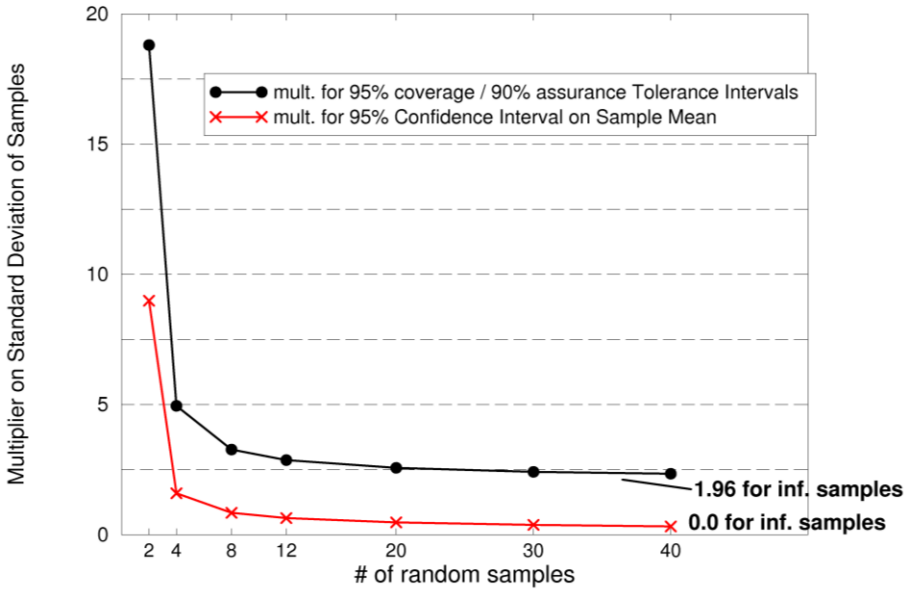


Figure 2.4 Multiplier on calculated standard deviation used to form 0.95/0.90 Tolerance Interval ranges vs. number of random samples. (Figure reproduced from [15], ignore confidence interval curve.)

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], [11] for uniform and right-triangular PDFs and for PDFs resulting from convolving various types of PDFs as depicted in Figure 2.5.

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 2.6). 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 [9]-[11] to be true for all tested PDF types and combinations in Figure 2.5.

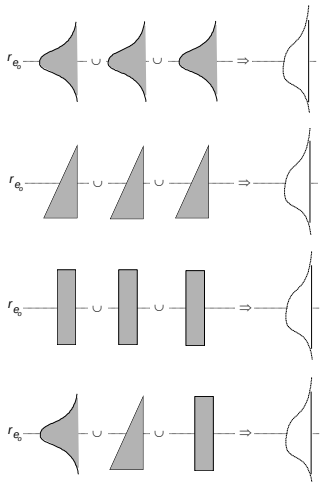


Figure 2.5 Test matrix for PDF representation study in [10], [11].

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 [9] – [11]). The Pradlwarter-Schueller approach [14] 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 2.5 (compared to an average of 92% reliability for TI).

2.4 Results for constant-temperature pipe bombs

At each material characterization temperature *T* in Figure 2.1 the derived constitutive model stress-strain curves are used in simulations of the pipe vessel held at uniform temperature *T* and pressurized at a linear ramp rate of 1psi/sec until pipe wall failure is indicated. (The rate of linear pressure rise is representative but does not impact these quasi-static failure pressures; creep is not an aspect of the current constitutive model and is assumed to be unimportant in the PB tests, based on a test for creep effects ([16]).) The pipe model, mesh and solver settings, pipe ends fixturing/loading conditions, simulations, and failure criteria are described in section 3.2. Table 2.2 gives the calculated failure pressures for the runs at the various characterization temperatures. The results at a given temperature are listed in order of predicted failure pressures and this orders the effective strengths 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.

For the six predicted failure pressures at 20C the mean is 1482 psi and the standard deviation σ is 3.97 psi. Table 2.3 lists these summary statistics and those for the other characterized temperatures. As indicated in the table, the upper and lower extents of the 0.95/0.9 tolerance intervals (TIs) are determined by adding/subtracting the quantity $f\sigma$ to/from the mean failure pressure, where the appropriate values of *f* are found in Table 2.1. The tolerance intervals are plotted in Figure 2.6.

Table 2.2 Predicted pipe failure pressures at listed temperatures where stress-strain curves were characterized (entries sorted from highest to lowest failure pressures).

Test and Temperature (degrees C)	Fail Press. (psi)	% Equiv. Plastic Strain	Tearing param.	hours runtime (192 CPUs)
try5-20	1485.2	57.50%	2.04	0.324
try6-20	1485.0	54.90%	1.54	0.348
try3-20	1484.5	60.10%	2.14	0.368
try39-20	1483.9	58.70%	2.09	0.402
try4-20	1482.8	57.10%	2.03	0.308
try40-20	1474.8	55.50%	1.96	0.309
try14-100	1227.1	58.60%	2.09	0.441
try36-100	1226.3	55.90%	1.98	0.335
try16-100	1225.3	56.10%	1.99	0.31
try37-100	1222.9	54.90%	1.95	0.284
try15-100	1208.7	52.80%	1.86	0.546
try11-200	1102.1	52.90%	1.66	0.335
try34-200	1089.9	44.20%	1.32	0.453
try13-200	1088.6	46.90%	1.43	2.26
try12-200	1085.8	42.60%	1.26	2.62
try35-200	1081.7	40.20%	1.17	0.342
try33-400	1014.0	38.40%	1.03	0.369
try17-400	1010.3	39.40%	1.06	0.393
try18-400	1007.2	38.60%	1.02	0.325
try19-400	1005.7	43.20%	1.2	0.312
try32-400	1001.9	37.30%	0.986	2.479
try24-600	884.7	52.30%	1.52	0.359
try23-600	880.1	49.00%	1.39	2.54
try22-600	869.2	40.90%	1.1	0.361
try25-700	714.0	61.70%	1.88	0.431
try27-700	704.2	60.60%	1.83	0.443
try26-700	703.7	60.50%	1.84	0.431
try31-800	448.8	64.50%	1.89	0.414
try29-800	448.0	50.10%	1.32	0.476
try30-800	440.8	63.20%	1.82	0.431

Table 2.3 Statistics of predicted failure pressures at temperatures where stress-strain curves were characterized.

Temperature	mean, μ fail press. (psi)	std.dev., σ (psi)	upper & lower extents of 0.95/0.9 TI	upper & lower extents of 0.95/0.9 TI (psi)
20C	1482.70	3.97 = 0.27% of mean	μ + 3.72σ = μ + 1% =	1497.5
			μ - 3.72σ = μ - 1% =	1467.9
100C	1222.06	7.63 = 0.63% of mean	μ + 4.15σ = μ + 2.6% =	1253.7
			μ - 4.15σ = μ - 2.6% =	1190.4
200C	1089.62	7.65 = 0.7% of mean	μ + 4.15σ = μ + 2.9% =	1121.4
			μ - 4.15σ = μ - 2.9% =	1057.9
400C	1007.82	4.59 = 0.46% of mean	μ + 4.15σ = μ + 1.9% =	1026.9
			μ - 4.15σ = μ - 1.9% =	988.8
600C	878.00	7.96 = 0.91% of mean	μ + 6.92σ = μ + 6.3% =	933.1
			μ - 6.92σ = μ - 6.3% =	822.9
700C	707.30	5.81 = 0.82% of mean	μ + 6.92σ = μ + 5.7% =	747.5
			μ - 6.92σ = μ - 5.7% =	667.1
800C	445.87	4.41 = 0.99% of mean	μ + 6.92σ = μ + 6.8% =	476.4
			μ - 6.92σ = μ - 6.8% =	415.4

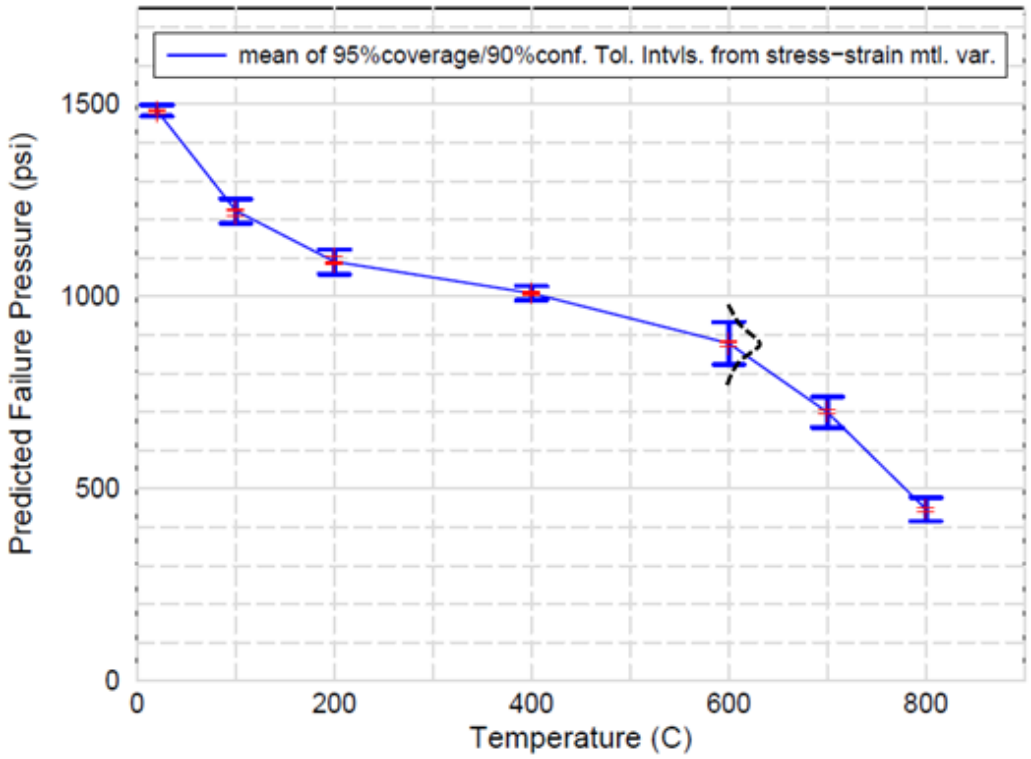


Figure 2.6 95% coverage / 90% confidence Tolerance Intervals from variation of pipe failure pressures at various pipe temperatures (for uniform temperature throughout pipe). Individual failure pressures predicted with the various stress-strain curves are plotted as red crosses. At 600C an illustrative Normal PDF constructed from the 0.95/0.9 TI is depicted.

2.5 Extension to other pipe bomb application conditions

In later validation simulations with the PB model, differences exist vs. the initial and boundary conditions underlying the TIs in Figure 2.6, such as wall thickness, fixturing/loading conditions on the pipe ends, pressure loading history, and non-uniform pipe temperature. A two-part approach explained in the following two sections is used to represent PB failure pressure variability under these different initial and boundary conditions.

2.5.1 Temperature dependence of material variability, parameterized in terms of high & low strength material curves

Consider a pipe simulation with the same pressure and end-loading conditions underlying the results in Figure 2.6. 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 [3] 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 pipe that increases from 20C to 200C over time. Figure 2.7 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 Table 2.2 and Figure 2.6. Each dot on each TI in Figure 2.7 corresponds to a particular stress-strain (s-s) curve. The temperature transient in the simulation can be negotiated with the interpolation scheme [3] 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 2.7. 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 5. Hence the following two-run "bounding" approach was taken.

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 effective strength variability reflected by the six data samples at 20C in Figure 2.7. 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 2.7 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).

[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.

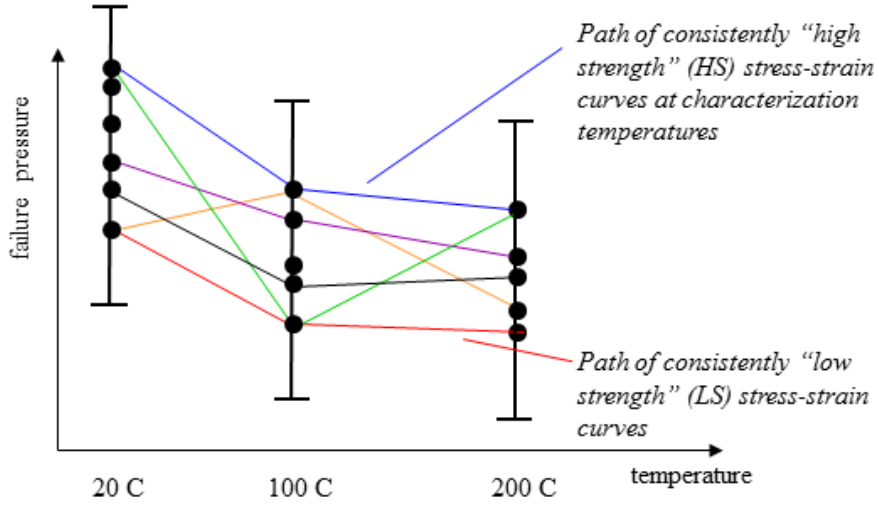


Figure 2.7 Illustrative (not to scale) 0.95/0.9 Tolerance Intervals and underlying PB failure pressure data (for uniform pipe temperature) for the first three material characterization temperatures in Table 2.2 and Figure 2.6. Several possible combinations ("paths") of material stress-strain curves are shown that could be used in a computational simulation to negotiate a temperature transition over the depicted temperature range.

Material realizations corresponding to mid-range dots at 20C in Figure 2.7 will also tend to correlate with the mid-range dots at 100C and 200C. 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 computational resources are not expended investigating moderate paths.

So far we have not discussed spatial material variability in the pipe. Figure 2.8 illustrates some possible combinations at three neighboring (> millimeters apart) locations on the pipe at uniform temperature 20C. Any of an exhaustive number of possible combinations would seem to be equally likely. To bound this problem we assume that PB simulation results from consistently high strength (HS) and consistently low strength (LS) combinations in Figure 2.8 would bound the simulation results from any other possible combinations like the others shown in Figure 2.8. Due to lack of time and resources we have not verified this assumption, but recommend that it be assessed in future projects that would seek to rely on this assumption.

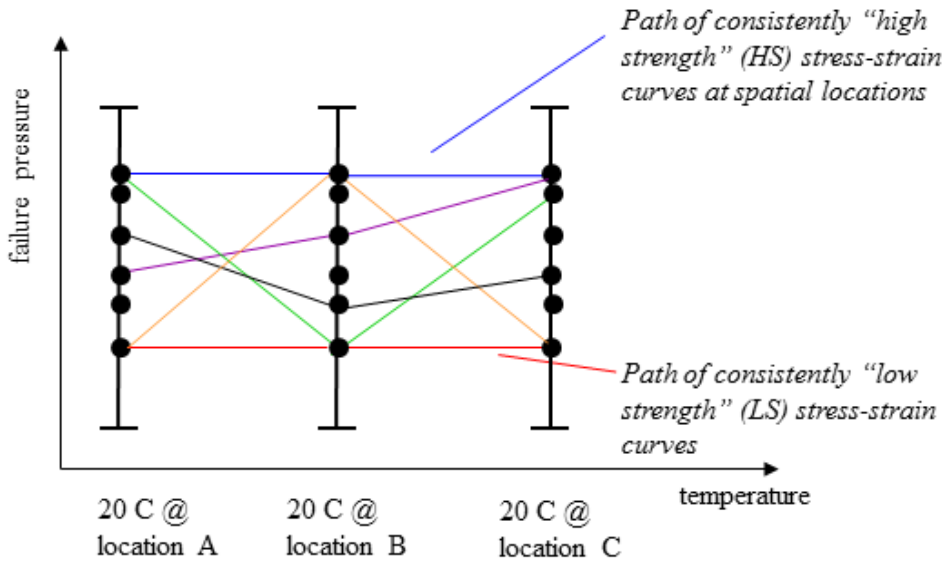


Figure 2.8 Illustrative (not to scale) 0.95/0.9 Tolerance Intervals and underlying PB failure pressure data for three neighboring locations on the pipe at uniform temperature 20C. Several possible stress-strain curves exist at each of three sample locations on the pipe. Some possible combinations of the stress-strain curves at these locations are shown that could be used in a computational simulation of a uniform-temperature 20C pipe. However, only the HS and LS combinations are used in the present work.

We use similar reasoning for steady or transient spatially non-uniform temperature fields. Consider a spatial temperature transition between 20C and 200C at two locations on the pipe. Figure 2.7 applies to this spatial temperature transition (and not just for the temporal transition previously considered with the figure). For spatial temperature transitions, strong correlation of ss-curve strengths is not expected between different locations because of physical material variations that can occur from point to point. Then any ss-curve computational path combinations are physically plausible such as all those shown in Figure 2.7. Again, we assume that the consistently high strength (HS) and consistently low strength (LS) combinations bound the simulation results from any other possible combinations. But we acknowledge that a spatially stochastic uncertainty representation would be relevant here and may even be necessary for certain analysis requirements. A follow-on investigation is recommended. We conclude similarly for spatially varying pipe wall thickness in section 5.3.

Thus, 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 [3] to transition between the HS s-s curves. The other simulation uses only the low strength s-s curves at each characterization temperature. It is next explained how this parameterization is used in Section 5 to scale PB failure pressure variability TIs in Figure 2.6 to TIs for the validation application conditions.

2.5.2 Scaling of PB failure pressure variability to new application conditions

Let the HS and LS simulations described above yield respective failure pressures P_fail-HS and P_fail-LS. These define a range of predicted failure pressure given by

$$\Delta = \text{P\_fail-HS} - \text{P\_fail-LS}. \qquad \text{Eqn. 2.1}$$

For the case of a pipe at uniform temperature 700C, the range $\Delta_{700Cunif} = \text{P\_fail-HS}_{700Cunif} - \text{P\_fail-LS}_{700Cunif}$ can be determined from the results in Table 2.2. This range is labeled in Figure 2.9 on the left TI, which corresponds to the 700C uniform-temperature PB tolerance interval defined in Table 2.3. The 700C-uniform TI is composed of three segments ($L_{\text{lower-}700Cunif}$, $\Delta_{700Cunif}$, $L_{\text{upper-}700Cunif}$) as shown in Figure 2.9.

The separation of the TI into three segments according to the HS and LS simulation results is handy for scaling the TI to other application conditions as illustrated in Figure 2.9. For example, let $TI_{app700C}$ at right in Figure 2.9 be for a non-uniform temperature pipe where failure occurs at the pipe hot spot with temperature 700C at its peak. In validation simulations with the PB model, several things will differ from the Figure 2.6 conditions, such as pipe wall thickness, pressure loading history, non-constant temperature, and end fixturing conditions on the pipe. Even if the failure point is at 700C at the time of failure, the calculated HS and LS failure pressures will differ (perhaps greatly) from the values $\text{P\_fail-HS}_{700Cunif}$ and $\text{P\_fail-LS}_{700Cunif}$ in Table 2.2 for the 700C uniform-temperature pipe. But the spans $\Delta_{700Cunif}$ and $\Delta_{app700C}$ in Figure 2.9 both issue from the same sets of low-strength and high-strength stress-strain curves. Therefore it is ventured that the spans $\Delta_{app700C}$ and $\Delta_{700Cunif}$ scale with each other. The scaling factor is

$$S = \Delta_{app700C} / \Delta_{700Cunif}. \qquad \text{Eqn. 2.2}$$

Everything on the right side of Equation 2.2 is obtainable from Table 2.2 and the HS and LS simulation results under the new application conditions. The lengths of the other segments that make up the TIs in Figure 2.9 are assumed to scale similarly. Mirroring Eqn. 2.2, $S = L_{\text{upper-app700C}} / L_{\text{upper-700Cunif}}$ or

$$L_{\text{upper-app700C}} = S \times L_{\text{upper-700Cunif}}. \qquad \text{Eqn. 2.3}$$

Analogously,

$$L_{\text{lower-app700C}} = S \times L_{\text{lower-700Cunif}}. \qquad \text{Eqn. 2.4}$$

Then $TI_{app700C}$ is fully defined by the three known segments ($L_{\text{lower- app700C}}$, $\Delta_{app700C}$, $L_{\text{upper-app700C}}$) as shown in Figure 2.9. Tolerance intervals constructed this way come from a loose but traceable basis of .95-coverage/.90-confidence TIs at "nearby" points in the problem parameter space. Because of several assumptions made in the construction, including interpolation and extrapolation (see next paragraph), it cannot be assured that the constructed

TI are likely be a conservative 95% range of failure pressure variation due to the underlying material property variability. We cannot assign a confidence level to the likelihood of conservatism, but we consider it to be a reasonable working estimate and use the lower and upper ends of the interval for .025 and .975 percentiles of a constructed Normal PDF of failure pressure variability in the new application, as indicated in Figure 2.9.

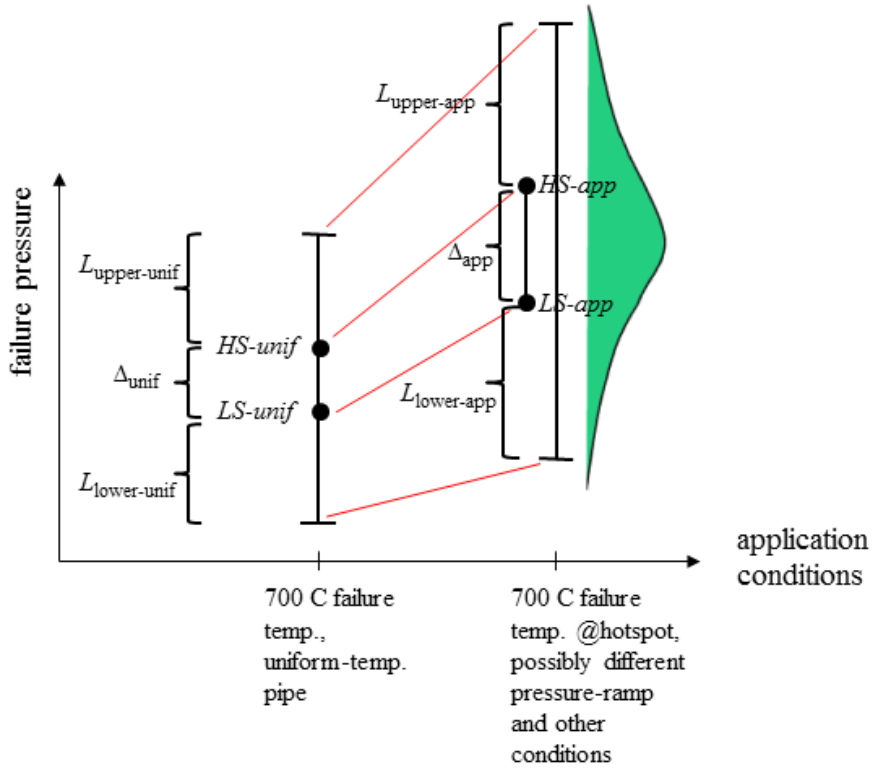


Figure 2.9 Proposed scaling between 0.95/0.9 Tolerance Intervals parameterized by failure pressures calculated with low and high strength material curves for (at left) a uniform-temperature pipe at 700C and (at right) a non-uniform temperature pipe where failure occurs at pipe hot spot with temperature 700C.

This procedure can be applied for the general problem where the failure temperature is some value *T* between the specific characterization temperatures in Figure 2.6. Then the subscript 700C in the above equations is replaced by the applicable temperature *T*. The span $\Delta_{T\text{unif}}$ in the new version of Eqn. 2.2 is linearly interpolated from the spans at the immediately surrounding characterization temperatures in Figure 2.6. The upper and lower segment lengths $L_{\text{upper-}T\text{unif}}$ and $L_{\text{lower-}T\text{unif}}$ are likewise interpolated.

Note that in the PB validation simulations, which involve both spatially and temporally varying temperature fields, the HS and LS predicted failures generally do not occur at exactly the same location or at exactly the same temperature. But failure locations and temperatures are within tenths of a percent for the HS and LS simulations so the average predicted failure temperature *T*avg-app is used in the above procedure, where the applicable subscripts are related by $T\text{unif} = T\text{avg-app}$.

3. FE Models, Geometries, Mesh and Solver Choices, and Calculation Verification

3.1 Cylinder tension test simulations

Here we consider the FE model and simulations of the cylindrical test specimens in the constitutive model material characterization/inversion procedure described in section 2.2. Given the geometry and testing conditions described in section 2.1 the simulations emulate cylinder response, through the necking and failure processes depicted in Figure 2.2.

The geometry and mesh of the corresponding FE model are shown in Figure 3.1. Only a 1-in. middle portion of the gage length is considered, where the loading and stress-strain response is considered to be axisymmetric. The FE model consists of only a ~2.77% portion of this 1-in. middle section of gage length, invoking symmetry to model a 20-degree portion and only the top half as shown. The model's radial dimension is 1/16 (0.0625) in. at the top cross-section, varying linearly to 0.0620 in. at the bottom. The bottom cross-section is modeled slightly less than actual so that computed failure will dependably occur at this minimum cross-section location.

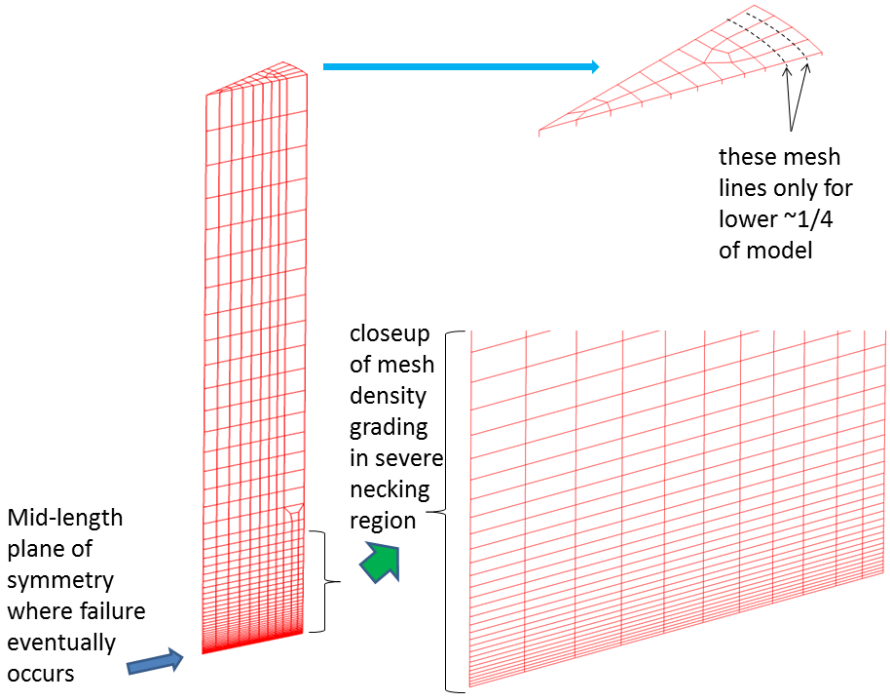


Figure 3.1 FE model geometry and mesh for modeled section of circular cylinder in tension tests.

By seeding the location of most deformation (thus failure) in this way, a mesh can be tailored for the simulations. Linear 8-node hexagonal finite elements are used. The mesh is most dense in the severe necking region. In this region the mesh is graded so that grid density increases substantially where approaching the mid-length plane of symmetry. Necking entails substantial

axial stretching and radial compression of elements. To compensate, the initial mesh that the simulations start from is made with axially compressed and radially stretched elements as shown. The mesh resolution appears to be adequate when compared to the much coarser meshes in a mesh sensitivity study in [18] for a similar material inversion problem that shows no mesh effects for the relevant portion of the derived stress-strain curves.

The solver and hourglass-control settings for the FE calculations in the material inversion procedure are also catalogued in [18]. These were chosen by the analyst ([3]) based on years of experience performing similar material inversion procedures for similar types of rod tensile tests. The settings were also deemed adequate by a second analyst with similar experience during a spot-check of reproducibility of the inversion procedure and results (see section 2.2).

3.2 Isothermal pipe bomb simulations at the material characterization temperatures

Isothermal PB simulations were performed to calculate the pipe failure pressures in Table 2.2. When the uniform-temperature pipes are pressurized to failure, a state is reached where the material can no longer resist the next increment in applied pressure. This is termed the 'structural instability' point ([17]). In the PB simulations this coincides with a mathematical instability, signified by the Adagio calculation going numerically unstable (unable to converge, [17]). The instability occurs because the Adagio quasi-statics governing equations have no inertial terms to balance the increment in pressure force by an acceleration of the material when its strength can no longer resist the internal pressure and the pipe bursts.

Identifying the burst pressure by arriving at the quasi-statics instability point is sufficient for our purposes here. If one wants to calculate the ensuing structural breakup of the pipe, a structural dynamics code can be used. It has the required inertial terms for continuation of the simulation from a starting point just prior to quasi-statics instability and continuing through structure breakup. Such continuation was demonstrated in [3].

Figure 3.2 shows representative Adagio calculation results associated with Table 2.2. A signature difficulty in these calculations is the large computational expense required to creep up to the structural instability point. As this point is neared, the solver must work harder to advance by taking smaller time-steps and performing more iterations per timestep to solve the increasingly nonlinear material response and force-balance problem.

The large majority of CPU time is spent in the final stages of the event. Alternate approaches are being investigated to more efficiently identify the structural/mathematical instability point or to identify alternative indicators of material damage and critical failure values. The latter approach would ideally suspend the calculation at a point simultaneously low on the CPU cost curve and near the final failure pressure. Such a point occurs, for example, at the sharp knee in the curve of the lower left CPU plot in Figure 3.2. This point is very near the ending (instability) time in the simulation, and because pressure is linearly ramped with time in these simulations, the knee also corresponds to a pressure that is very near the final pressure at calculation failure.

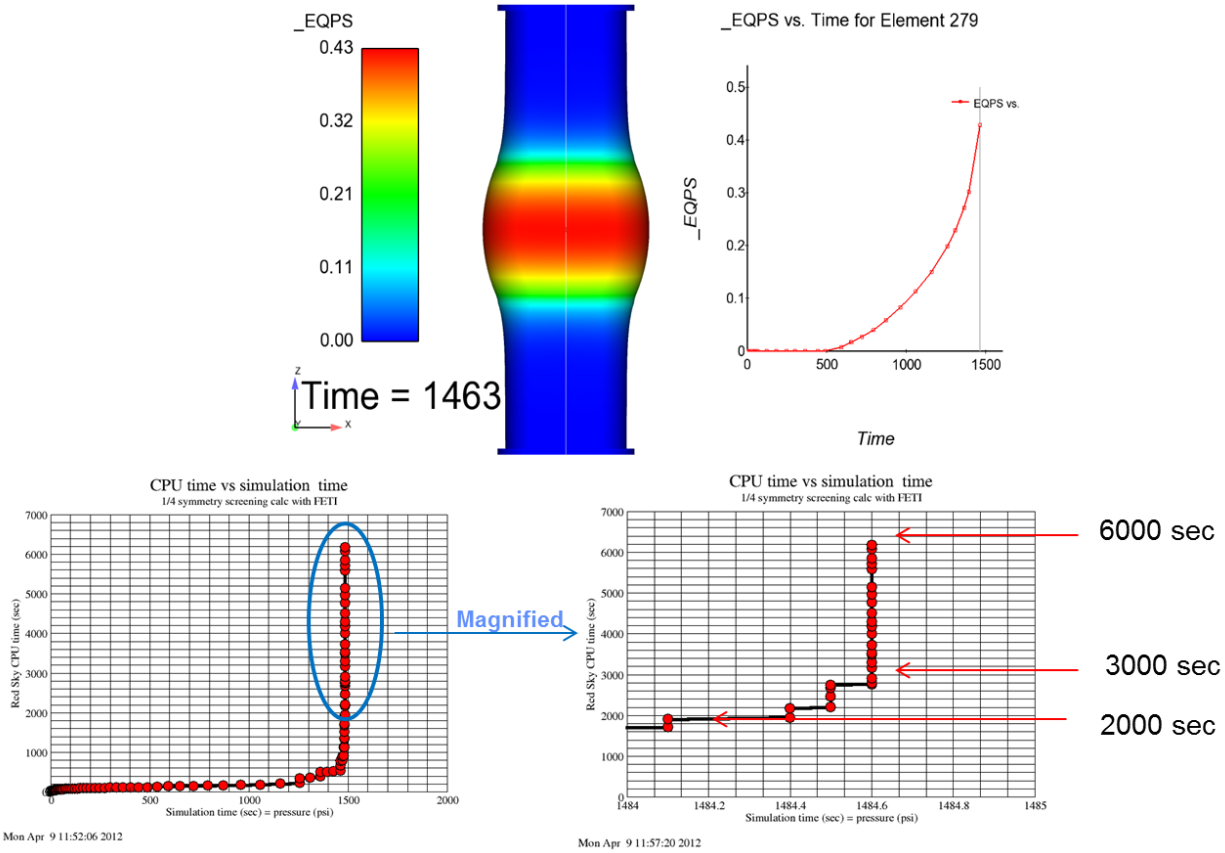


Figure 3.2 Representative Adagio calculation results associated with this section and Table 2.2. (EQPS is the 'equivalent plastic strain' of the material.)

3.2.1 Geometry and Mesh

The FE model has a pipe wall thickness of 0.02 in. and 4 finite-elements through the thickness of the wall. Mechanical end-loading conditions are fixed. The full-geometry model is shown at left in Figure 3.3 (see [2] for detailed drawings). For the calculations in section 3.3 a geometry change was made in the modeled length of the pipe. The full length of the PB cylinder is 14 inches. The truncated model at right in Figure 3.3 contains only the middle 10.8 inch section of the pipe as shown. Most of the thick-walled shoulder regions at either end are truncated because previous simulations found they are so strong relative to the thin-wall sections of the pipe that truncating them does not materially affect failure pressure results. This reduces the finite-element count by about 30%. To further save on computational resources a 1/8 section of the truncated geometry was used that takes advantage of the symmetry of geometry and uniform-temperature and fixed-end boundary conditions in the problem. See Figures 3.4 and 3.5 for details on the mesh structure and element aspect ratios.

In section 3.4 we investigate effects of different mesh densities and element aspect ratios for simulations of the spot-heated PB experiments. From that investigation we infer that the mesh in Figures 3.4 and 3.5 appears quite adequate for simulations of uniform-temperature pressurized

pipes described in this section. These simulations are used to calculate the failure pressures in Table 2.2.

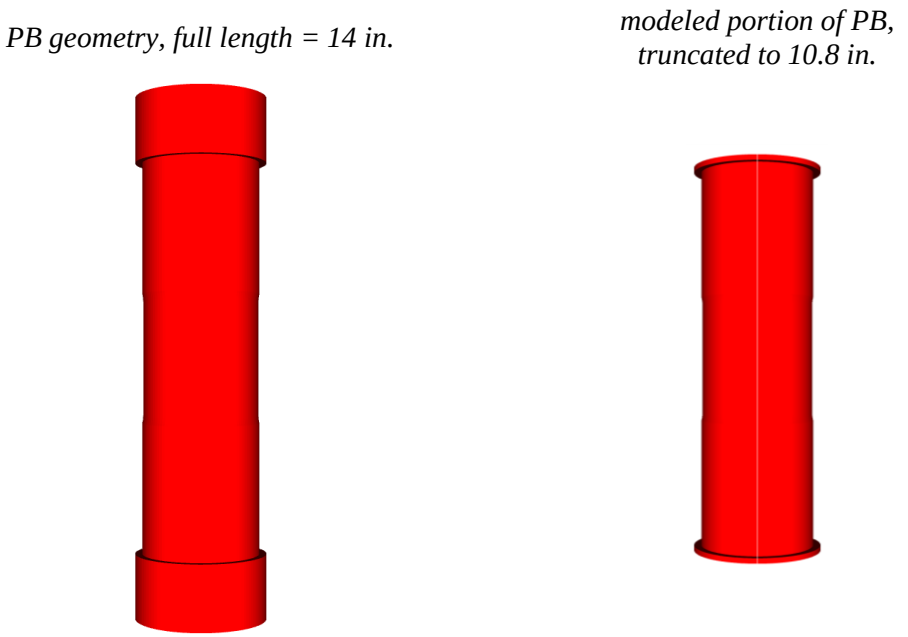


Figure 3.3. Actual and truncated-model pipe geometries. Full length of pipe is 14 inches. Modeled portion omits 1.6 inches from each end of pipe, leaving ¼ in. length of stiffening collar at pipe ends. This decreases the finite-element count by almost 30%.

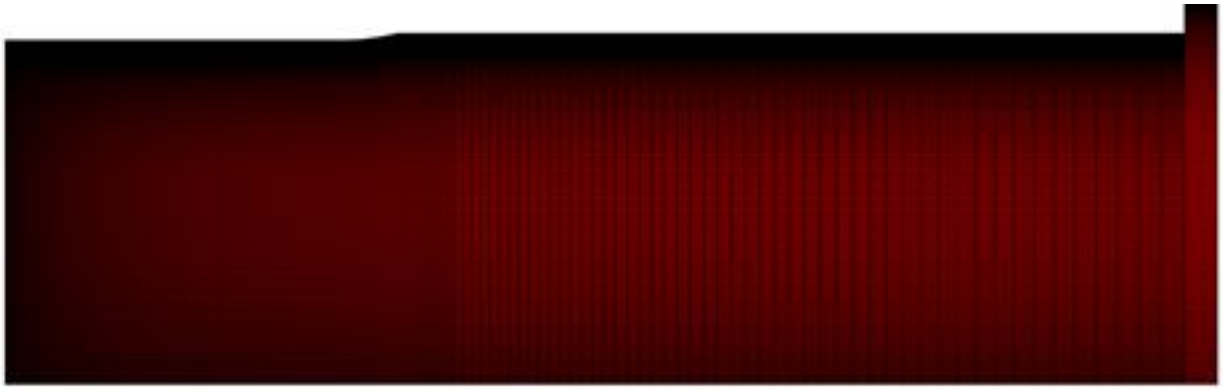


Figure 3.4 Finite element mesh of 1/8 symmetry model of truncated pipe section shown in Figure 3.3. The 1/8 symmetry model contains 200,322 8-node linear hex elements. Grid density increases significantly in going from truncated end of pipe to mid-length of pipe. Grid density also increases in going from back of pipe to front of pipe. (The front-to-back grid density variations were imposed in anticipation of simulating scenarios with the hot spot from non-uniform heating conditions in the PB experiments, but different meshes were ultimately used, see section 3.5).

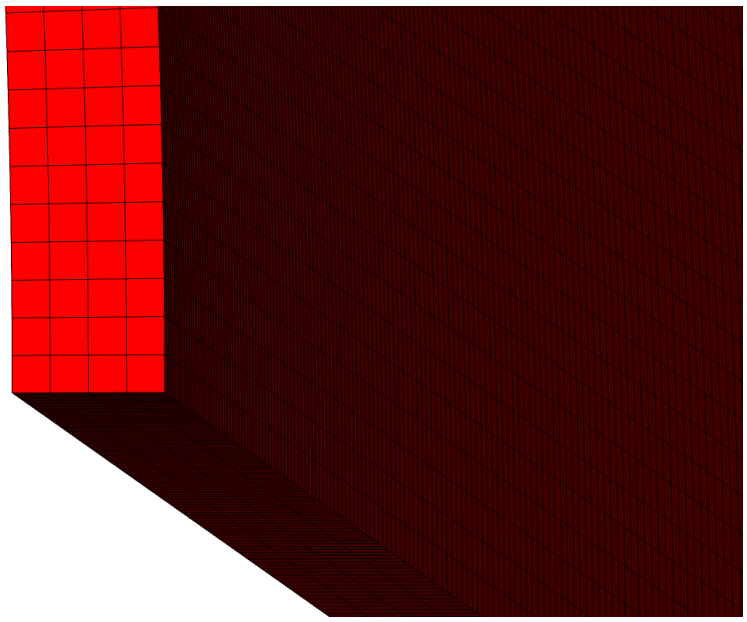


Figure 3.5 Mesh with four elements through the thickness of the pipe wall. The starting aspect ratios of the elements at pipe mid-length are 4:4:1 in the radial, axial, and circumferential directions respectively.

3.2.2 Solver, control settings, and error tolerance refinement study

For the PB simulations the conjugate gradient (CG) solver in Adagio was used with the FETI (finite-element tearing and interconnecting) pre-conditioner. The solver settings in Table 3.1 were used to produce the uniform-temperature PB results in Table 2.2.

The settings in Table 3.1 were arrived at by performing the solution cost vs. accuracy study summarized in Table 3.2. Selected cases from Table 2.2 are considered in Table 3.2. The cases span a large temperature range from 20C to 700C and the stress-strain curves at each temperature yield closely coinciding failure pressures. This provides a tough test for the solver variants in Table 3.2: can the strength orderings of the selected stress-strain curves be accurately determined despite the small separations in the failure pressures, and is this successfully done at the low end 20C and high end 700C temperatures?

The error tolerances in column two of Table 3.2 are customary for Adagio analyses according to analyst experiences balancing calculation cost and accuracy in many projects prior to PB. Relatively small average differences of about 2.8 psi or 0.24% exist between results in the 2nd column and in the 3rd column where error tolerances are significantly stricter. But a difference in curve-strength orderings occurs for try3-20C and try6-20C. In further tightening the error tolerances from column three to column four, the curve-strength orderings do not change and the calculated failure pressures are virtually unchanged (only the results for try26-700C change, and only by 0.1 psi). Hence, the error tolerances in column 2 are not sufficient for curve strength

ranking purposes and the tolerances in columns 3 and 4 both appear adequate; results are stable when error tolerances are tightened to these levels (for this particular type of PB analysis).

Table 3.1 Solver error tolerances and algorithm controls in uniform-temperature PB calculations.

Solver Error Tolerances	CG
Target relative residual	1.00E-06
Acceptable relative residual	2.00E-04
Max iterations	10,000
Min iterations	3
Max cutback	10
cutback factor	0.5
Growth factor	1.1
Iteration window	100
Target iterations	400
Max multiplier	100
Min multiplier	1.00E-12
Max timestep	1.0
Hourglass Effective Moduli	Elastic
	FETI
Residual Norm Tolerance	1.00E-03
Iteration updates	125

We note that the CPU costs of the simulations in Table 3.2 do not necessarily increase with tightened error tolerances. Complex numerical interactions in adaptive solvers sometimes result in *slower* progress overall when looser iteration-convergence tolerances are used. This can allow larger time steps and disproportionately increased iterations to converge the nonlinear solution over the larger time steps. Furthermore, sometimes the time step has to be retried with a smaller step size if convergence does not occur within the specified iteration limits for the original time step. So larger steps can result in non-convergence and then much numerical rework as the size of the steps are cut down until convergence can be achieved. Because the third column only saves about 3.4% on average vs. the stricter fourth column, we used the more strict tolerances (see Table 3.1) in the simulations for Table 2.2.

To get a sense of the effectiveness of the FETI pre-conditioner we tried the CG solver without pre-conditioning. CG solver settings from Table 3.1 were used except for the following increased iteration targets and limits because of no preconditioning: *iteration window* = 10,000, *target iterations* = 35,000, *max iterations* = 50,000. On the five cases in Table 3.2 the non-pre-conditioned simulations took from 2 to 5 times longer than with FETI preconditioning and

results were insignificantly different, within 0.1%. Hence, on our PB problem the FETI pre-conditioner yields large computational savings with no appreciable drawbacks, so we used it in all that follows.

Table 3.2 Comparison of predicted failure pressures with uniform-temperature pipe model and FETI-CG solver using the listed error tolerances. Other settings not specified are the same as in Table 3.1. Cases are listed in order of increasing failure pressure per the last two columns which use the tightest error tolerances.

Test & temperature cases	Rel.Resid.=$1e^{-4}$ Accep.Resid.= $2x10^{-2}$	Rel.Resid.=10^{-5} Accep.Resid.=$2x10^{-4}$	Rel.Resid.=10^{-6} Accep.Resid.= $2x10^{-4}$
	Failure psi (CPU hrs.*)	Failure psi (CPU hrs.*)	Failure psi (CPU hrs.*)
try26-700C	702.0 (20.3)	703.8 (5.87)	703.7 (5.24)
try27-700C	704.1 (19.1)	704.2 (5.28)	704.2 (6.21)
try3-20C	1490.70 (12.1)	1484.5 (7.8)	1484.5 (9.78)
try6-20C	1487.20 (4.6)	1485.0 (2.9)	1485.0 (4.39)
try5-20C	1492.60 (41.3)	1485.2 (20.7)	1485.2 (8.26)

* CPU times reported in Adagio output file via global output variable *cpu_time*. Simulations were run on 192 processors of Red Sky.

3.3 Check of 8/8 truncated geometry model vs. 1/8 truncated geometry model

As a spot check, the 8/8 truncated geometry model at right in Figure 3.3 was run at 700C with the try26 and try27 stress-strain curves. (The experiments we will later validate the model against have a failure temperature of approximately 700C so the stress-strain curve rankings at 700C are of high relevance.) From Table 3.2 a very small difference of 0.5 psi separates the failure pressures for try26-700C and try27-700C. We wanted to assess how well this small separation was preserved when running the 8/8 geometry model. The calculations were run with the FETI solver and control settings listed in Table 3.1. The 8/8 model has 1.6 million elements and runs over an order of magnitude slower than the 200,322 element 1/8 geometry model. Table 3.3 reveals that for the try26 curve a shift of +3.1 psi increase in failure pressure to 706.8 psi occurred with the 8/8 model. For the try27 curve a shift of +3.7 psi to 707.9 psi occurred with the 8/8 model. These ~0.5% shifts are very small, supporting a conclusion that the mechanical loading symmetry boundary conditions were applied correctly in the 1/8 model.

Furthermore, because the 1/8 model is only used for relative curve-strength ranking purposes, and not where prediction accuracy on an absolute basis is sought, the shift between the 8/8 and 1/8 models is not a problem if the shifting is relatively uniform for the various stress-strain curves. Reasonably similar shifts of 3.1 psi and 3.7 psi occur for the try26 and try27 curves. The curve-strength rankings are the same with either model. From this spot check we assume that the

1/8 model gives accurate curve-strength rankings in Table 2.2 for all curves and all characterization temperatures.

Table 3.3. Comparison of predicted pipe failures with the 8/8 and 1/8 geometry models (all calculations with FETI/CG solver settings in Table 3.1)

Test cases	1/8 model Failure psi	8/8 model Failure psi	difference (8/8 – 1/8) Failure psi	% diff. Failure psi
try26-700C	703.7	706.8	+3.1	0.44%
try27-700C	704.2	707.9	+3.7	0.52%

3.4 Pipe bomb model variants for hot-spot heating and pressurization to failure

Four different categories of PB simulations with hot-spot heating and pressurization to failure are summarized below. Details of the calculation verification study (models, meshes, simulations, results, and analysis) are described in the remainder of this section. The findings are applicable to the other three categories of simulations as well. Further modeling and simulation details and results for the other three categories are given in the cited sections.

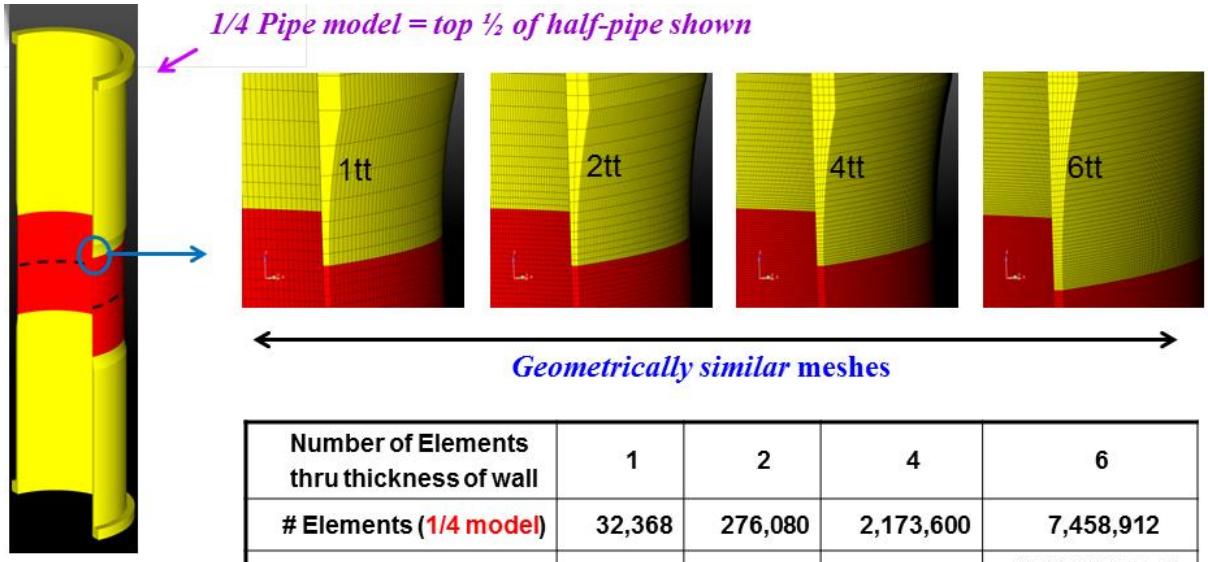
- **PB calculation verification study**: ¼ symmetry Adagio model with mapped/interpolated temperature field BCs from test PB1 thermocouples (see next section);
- **coupled thermo-mechanical simulations**: ¼ symmetry Adagio-Aria model with heating shroud and thermal radiation and conduction included in model—used to model the validation experiments for experiment design and improvement to minimize validation uncertainty (see section 4.1 and 4.2);
- **"self check" simulations to characterize temperature mapping/interpolation error**: ¼ symmetry Adagio model with mapped temperature field BCs from virtual TCs in coupled thermo-mechanical simulations —used to characterize BC temperature field errors from TC mapping/interpolation vs. known reference field from coupled thermo-mechanical simulations (see section 4.4);
- **Full-model validation/UQ simulations**: full pipe model Adagio simulations with TC mapped/interpolated temperature field BCs from PB tests (see section 5).

3.5 Calculation verification of simulations of hot-spot heating and pressurization to failure

3.5.1 Mesh refinements and solver settings in calculation verification study

Figure 3.6 shows the ¼ pipe geometry and meshes used in the mesh refinement study. The temperature field boundary condition on the pipe comes from test PB1 thermocouple temperature mapping and interpolation described in section 4.2. This BC is used to maximize the applicability of the mesh study and discretization error/uncertainty estimates in this section to the

PB validation simulations in section 5. Section 4.4 shows that the temperature field in the experiments varies markedly from front to back and axially upwards and downwards from mid-height. This variation is captured in the ¼ symmetry model except for minor circumferential and axial asymmetries characterized in section 4.4. A ¼ pipe model was required for affordability of the mesh refinement study, as will become evident.



Number of Elements thru thickness of wall	1	2	4	6
# Elements (1/4 model)	32,368	276,080	2,173,600	7,458,912
Pressure at Fail (psi)	1069	955	850	819.1* (*didn't finish, 36 days on 400cpu's)

Figure 3.6 ¼ truncated-length pipe geometry, meshes, linear HEX element counts, and failure pressure results in mesh refinement study.

The **1tt** (**1** finite element **t**hrough the **t**hickness of the wall) mesh in Figure 3.6 was created with approximately equal cell lengths in the axial, circumferential, and thru-wall dimensions in the 0.020 in. thin-wall extent of the pipe where the hot spot exists and hence large deformation and ultimately failure occur. Use of 1:1:1 aspect ratio elements is not necessarily optimal because pressurization-induced deformations cause the walls to thin in the radial direction because of wall stretching predominantly in the circumferential dimension but in the axial direction as well. Therefore, finite elements near the hot spot that start the simulation with equal cell lengths have highly non-uniform aspect ratios by the end of the simulation. Later in this section we study the effects of attempting to compensate for element thinning and stretching by using non-1:1:1 aspect elements in the starting meshes.

In the pipe wall-thickness transition region zoom-shots in Figure 3.6 the **1tt** mesh has axial and thru-wall cell dimensions that increase together according to wall thickness as it transitions from 0.02 in. to 0.05 in. The cell axial dimensions are coordinated to approximately match cell widths (thru-wall) so that a nearly 1:1 aspect ratio is maintained in these element dimensions in the wall transition region and hence throughout the model. However, to maintain an i-j-k structured hex mesh throughout the model, cell size in the circumferential dimension cannot likewise follow wall thickness. Cell size in the circumferential direction is constrained to be constant along the entire length of the pipe at its

inside-diameter (ID) surface. This (constant) arc-length cell dimension of approximately 0.02 in. at the pipe ID is set by the imposed target of 1:1:1 starting aspect ratios where the hot spot and deformation and failure occur.

Our mesh refinement study followed rules for constructing *geometrically similar* meshes (explained next) that are necessary for correctly estimating order or rate of convergence of the solution as the mesh is refined ([25]). Based on the rate of convergence, the asymptotic solution in the limit of infinite mesh refinement is estimated and assigned an associated uncertainty band. Then the solution error for any mesh in the sequence of refined meshed can also be estimated with assigned uncertainty.

The 2tt mesh in our refinement sequence doubles the number of elements across the wall. The two elements across the wall are equally spaced, having equal width in the thru-wall direction. The refinement rules require that a mesh refinement in one dimension be matched by similar refinements in the other dimensions of the mesh. Hence, nominally, each element of the 1tt mesh is subdivided into eight elements to get the 2tt mesh and each element in the 2tt mesh has essentially the same aspect ratio as the parent element in the 1tt mesh.

Another refinement rule is to maintain any grading of mesh density when subdividing elements. For example, in the 1tt mesh, the seven elements that vertically span the wall-thickness transition region have graded (non-constant) cell sizes as the region is traversed in the vertical/axial direction. This grading is required to meet the objective of 1:1 aspect between axial and thru-wall cell dimensions, as the wall thickness increases in the axial direction. To meet the mesh refinement rule that the vertical mesh-density grading function in this region be preserved under mesh refinements, the seven elements that vertically span the transition region are not halved when subdivided in the axial direction. Rather, they are vertically subdivided such that the mesh density grading in the vertical direction is nominally preserved.

The word ‘nominally’ in the two preceding paragraphs signifies that the mesh refinement objectives were met as well as reasonably possible by manual trial-and-error iteration with the present meshing facilities in CUBIT [26]. The 4tt and 6tt meshes in Figure 3.6 were constructed with similar considerations in mind**.** This experience spawned an initiative to develop automated mesh refinement capabilities in CUBIT that respect the rules for enabling estimation of solution convergence rate, asymptotic solution results, and solution error of the meshes in the refinement study.

The calculations were run with the solver settings in Table 3.4. The FETI *residual norm tolerance* was tightened to 10^{-4} vs. Table 3.1 and experimentation for improving solver efficiency (reducing CPU times) lead us to lower some of the iteration limits so that time-step cutbacks and re-tries would occur sooner instead of allowing high numbers of iterations that ultimately were often unsuccessful.

Table 3.4 Solver error tolerances and algorithm controls in mesh study simulations and subsequently.

Solver Error Tolerances	CG
Target relative residual	1.00E-06
Acceptable relative residual	2.00E-04
Max iterations	20
Min iterations	3
Max cutback	10
cutback factor	0.5
Growth factor	1.25
Iteration window	100
Target iterations	20
Max multiplier	500
Min multiplier	1.00E-12
Max timestep	10
Hourglass Effective Moduli	Elastic
	FETI
Residual Norm Tolerance	1.00E-04
Iteration updates	10

3.5.2 Mesh refinement results, convergence assessment, and solution error & uncertainty estimates

The failure pressures calculated with each mesh are plotted and listed in Figure 3.7. As noted in Figure 3.6 the calculation with the 6tt mesh did not finish within an allowed 36 days on 400 processors. (Adagio restarts were not working at the time of the mesh study so this was a continuous run for 36 days.) If the calculation was allowed to run until it reached instability failure, the associated failure pressure would be greater than the 837 psi level at the end of the 36 day calculation. Hence 837 psi is the lowest potential failure pressure plotted in Figure 3.7 for the 6tt mesh. The calculated material damage quantities of equivalent plastic strain and tearing parameter [3] at the point when the 6tt simulation timed-out were very low relative to values from 1tt, 2tt, and 4tt results (see Table 3.5). So the 6tt simulation was probably relatively far from reaching failure instability. Therefore 837 psi is certainly a lower bound, but not a plausible value of failure pressure that would be yielded by a finished 6tt mesh simulation.

For an upper bound, the trend of decreasing computed failure pressure with mesh refinement in Figure 3.7 indicates that an upper bound on failure pressure for the 6tt mesh can reasonably be assumed to be less than the 874 psi result for the 4tt mesh. Then representative potential failure

pressures within the lower (837 psi) and upper (874 psi) assumed limits are chosen as 845, 855, and 865 psi as plotted in Figure 3.7.

A standard method ([21], [22]) was used to determine the empirical rate of convergence of computed failure pressure as the mesh is refined. Results from the three coarsest meshes, 1tt, 2tt, and 4tt, yield an empirical rate of convergence of 0.26. This is much lower than the surmised ideal value of 1.0, which is the theoretical rate at which calculated stresses converge with mesh cell size for the second-order spatial discretization of displacement fields in Adagio. Although on thin ground from a theoretical standpoint, we relate computed failure pressure to computed stress state in the pipe. We venture that failure pressure might be expected to converge at the same rate as computed stresses, i.e., at an ideal rate of 1.0. Because failure pressure is not a nodal quantity or a quantity computed on the mesh, and because failure pressure is a value determined by the point of non-convergence of the calculation, we are not aware of any theory that would establish a theoretical rate of convergence for computed failure pressure. However, empirical results in Table 3.2 and Figures 3.7 and 5.8 provide a basis for treating failure pressure as a well-behaved quantity yielded by these computations.

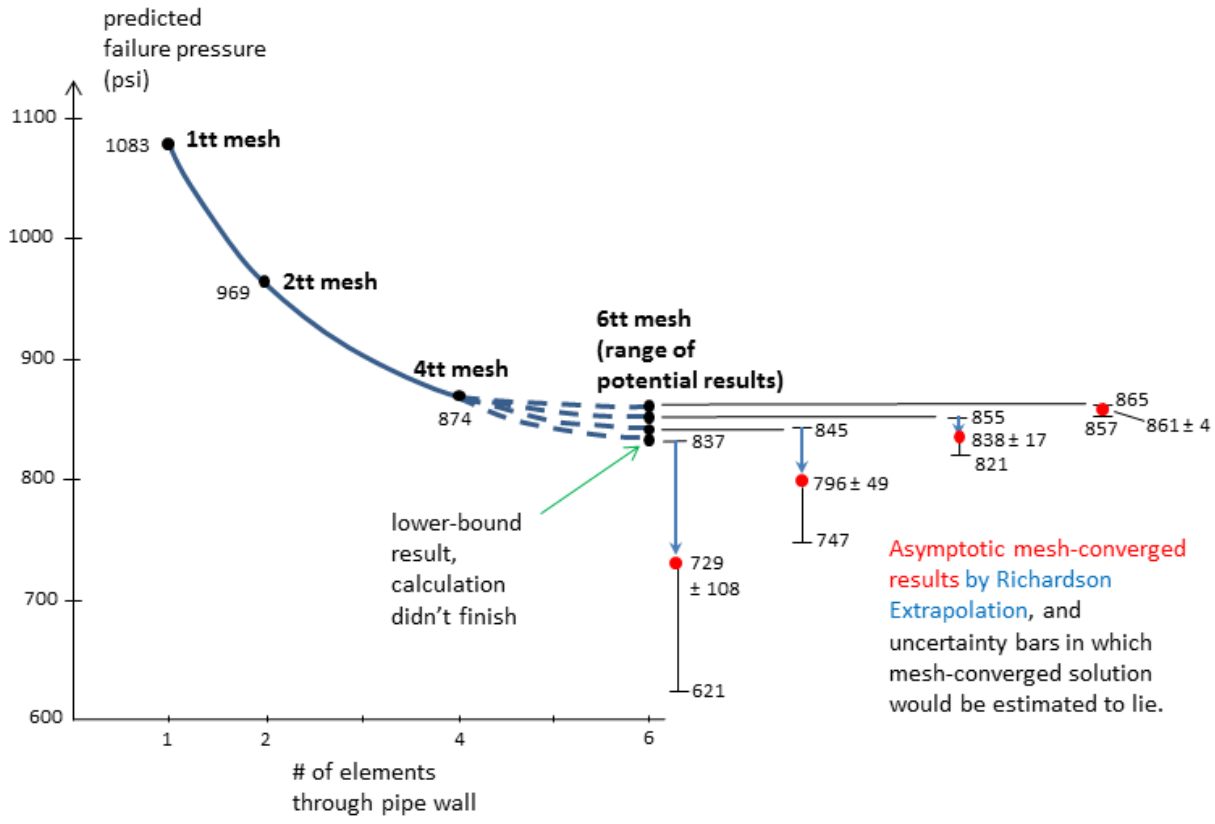


Figure 3.7 Calculated pipe failure pressures in mesh refinement study. Lowest result for 6tt mesh is from a calculation that did not finish within an allowed 36 days on 400 processors. Various plausible finishing results are plotted for the 6tt mesh, along with Richardson Extrapolation estimates of asymptotic mesh-converged values (red dots) and associated solution uncertainty bars (see text).

Table 3.5 Calculated material damage quantities Equivalent Plastic Strain and Tearing Parameter for simulations in mesh study (spatial maximums over pipe, see e.g. Figure 4.9).

mesh	maximum Equivalent Plastic Strain in simulation	maximum Tearing Parameter value in simulation
1tt	138%	5.5
2tt	208%	8.89
4tt	116%	3.53
6tt *sim. timed out	27%	0.71

Even more than the fact that the 0.26 convergence rate did not meet the surmised ideal rate, the consensus of the Sandia solid mechanics analysts consulted by the authors was that the 1tt-mesh results do not lie in the asymptotic regime of convergence for this problem, a requirement for Richardson Extrapolation to perform well. This is because structures modeled by single-element-thick linear hex elements have no resistance to bending moments. Therefore we pursued the more refined triplet of 2tt, 4tt, and 6tt meshes/results to apply Richardson Extrapolation (RE). As Table 3.6 presents, uncertainty in the finished 6tt mesh result yields uncertainty in the empirical rates of convergence and therefore uncertainty in the asymptotic failure pressure estimated by RE for the limit of mesh size decreasing to zero. The estimated potential asymptotic failure pressures are plotted in Figure 3.7.

In Figure 3.7 an uncertainty bar is shown about each estimated potential asymptotic value. Each uncertainty bar has a half-length equal to the distance between the asymptotic estimate and the corresponding 6tt mesh value. The half length of each estimated asymptotic solution's uncertainty bar is about 80% of that given by Roach's Grid Convergence Index ([21], [22]) when the recommended factor of safety 1.25 is used. This safety factor accompanies: 1) empirical rates of convergence used for RE; and 2) structured grids refined into geometrically similar meshes. The present study meets these criteria. Nonetheless, we used the simpler method of uncertainty bar sizing described at the start of this paragraph.

Table 3.6 Potential empirical orders of convergence based on 2tt, 4tt, 6tt meshes and Richardson Extrapolation estimates of asymptotic mesh-converged failure pressures.

6tt mesh potential finished result (failure pressure, psi)	potential empirical order of convergence	RE estimates of potential grid-converged failure pressure results (psi)
837	0.73	729
845	1.15	796
855	1.87	838
865	3.1	861

Recalling the prior arguments about the potential 6tt mesh results considered plausible, we further consider only the results in the bottom three rows of Table 3.6. The bottom row shows a potential order of convergence that seems implausibly high, whereas the middle two rows have much more reasonable values. Therefore we discount the last row's potential value of failure pressure (865 psi). We are left with the middle two rows in Table 3.6. We make no further judgment that one of these rows has greater or lesser plausibility or probability than the other. Hence we obtain a combined uncertainty range that spans the uncertainty bars in Figure 3.7 which correspond to the middle two rows in Table 3.6. That is, the asymptotic grid-converged failure pressure for the modeled conditions in this mesh study is estimated to lie within an uncertainty range between 747 psi and 855 psi. This uncertainty estimate will be used in the model validation uncertainty analysis in section 5.2. To prepare for this, we note that corrections of -228 and -336 psi added to the 1tt-mesh result yield the said uncertainty range [747, 855] psi within which the asymptotic grid-converged failure pressure is estimated to lie.

4. 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. [29], [30]) 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 over large ranges of pressures and temperatures, ramp rates, and large temperature gradients on the pipes. Coupled thermal-structural modeling (section 4.1) was used to help define the length of the pipe and the size and temperature of the heating shroud and its location relative to the pipe in order to achieve the spatial temperature variations sought. The model was also used to devise a thermocouple (TC) placement scheme suitable to capture the anticipated temperature field variations and support a mathematical interpolation scheme to reconstruct the temperature field boundary condition on the pipe surface (section 4.2). Temperature field reconstruction prompted an experimental adjustment to compensate for convection effects (section 4.3). Representative simulations were performed to test the temperature mapping and interpolation capability and characterize associated error and uncertainty (section 4.4).

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

4.1 Coupled thermal-mechanical modeling to help design the validation experiments

Coupled thermal-structural modeling was used to help define the length of the pipe and the size and temperature of the heating shroud and its location relative to the pipe in order to achieve the spatial temperature variations sought. The model included thermal radiation exchange between the heating shroud, the dynamically bulging 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 4.2. Nonetheless, the simplified thermal model was deemed sufficient to aid in the design and planning of the thermal aspects of the validation experiments.

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 temperature field in Figure 4.1 allows a reduced 1/4 geometry to be used for the calculations in all of section 4. But the 1/4 model used here has full shoulder extensions like the version at left in Figure 3.3.

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

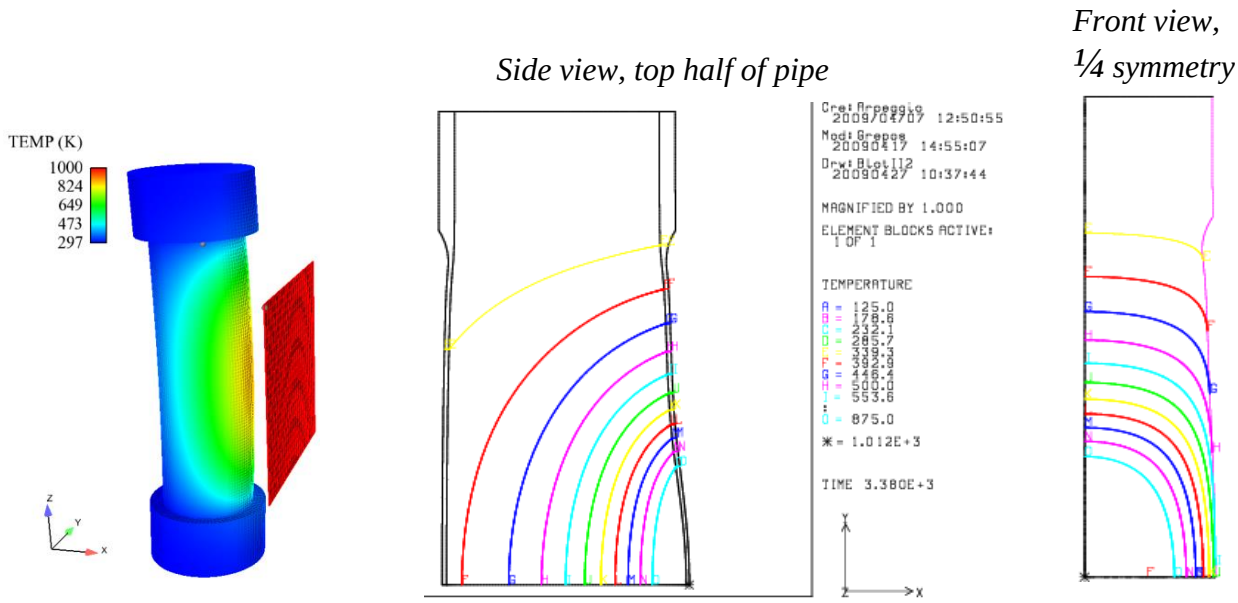


Figure 4.1 Coupled SIERRA/Thermal-Solid Mechanics simulation results for circumstances described in body of text.

4.2 Thermocouple placement and interpolation scheme for temperature field boundary condition reconstruction on pipe

Early in the design and planning of the validation activity it was judged 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 the accuracy of the constitutive model. In particular, for modeling purposes the emissivities of the radiating surfaces of the shroud, pipe, and internal slug would not be known to within ±10%. 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 not only affect the gas temperature inside the pipe in a complex way, but 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) exist to be found.

Fortunately, error and uncertainty associated with thermal aspects in the validation problem were largely circumvented by measuring pipe surface temperatures at specifically designed thermocouple locations (based on model simulations). Then spatial interpolation was used (with interpolation error/uncertainty factored in) 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.

An initial constraint was that 18 or fewer thermocouples (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 for sensor placement and interpolation 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 occurs. Accordingly, the temperature contour information in Fig. 4.1 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 quasi-Hermite 2D bi-cubic polynomial interpolation scheme was identified as something that could be quickly developed and otherwise met our needs. 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. 4.2. 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. 4.2, assuming it is important to get greatest accuracy in the vicinity of the hot spot and accuracy importance drops as temperature drops with distance from the hot spot.

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

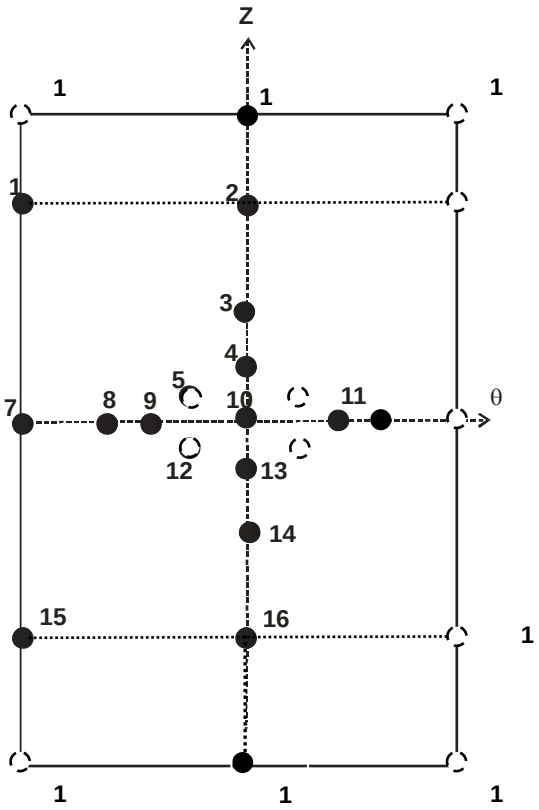


Figure 4.2 Division of pipe exterior surface into eight quadrilateral "interpolation patches". Surface of the pipe represented as though sliced at the back, unwrapped, and laid flat in θ-*z* space as shown. TC locations on pipe surface are shown (not drawn to scale, see Table 4.2). TC 10 marks the origin (θ = 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 the convention used in the interpolation subroutine (see [18]) and not the ID numbering of TCs in the experiments (Fig. 4.3).

In Fig. 4.2 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 (θ = 0, *z* = 0) is approached, as shown 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 θ = 0, and to a lesser extent (because of convection) to be approximately symmetric about the horizontal θ-axis at *z* = 0. But 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.

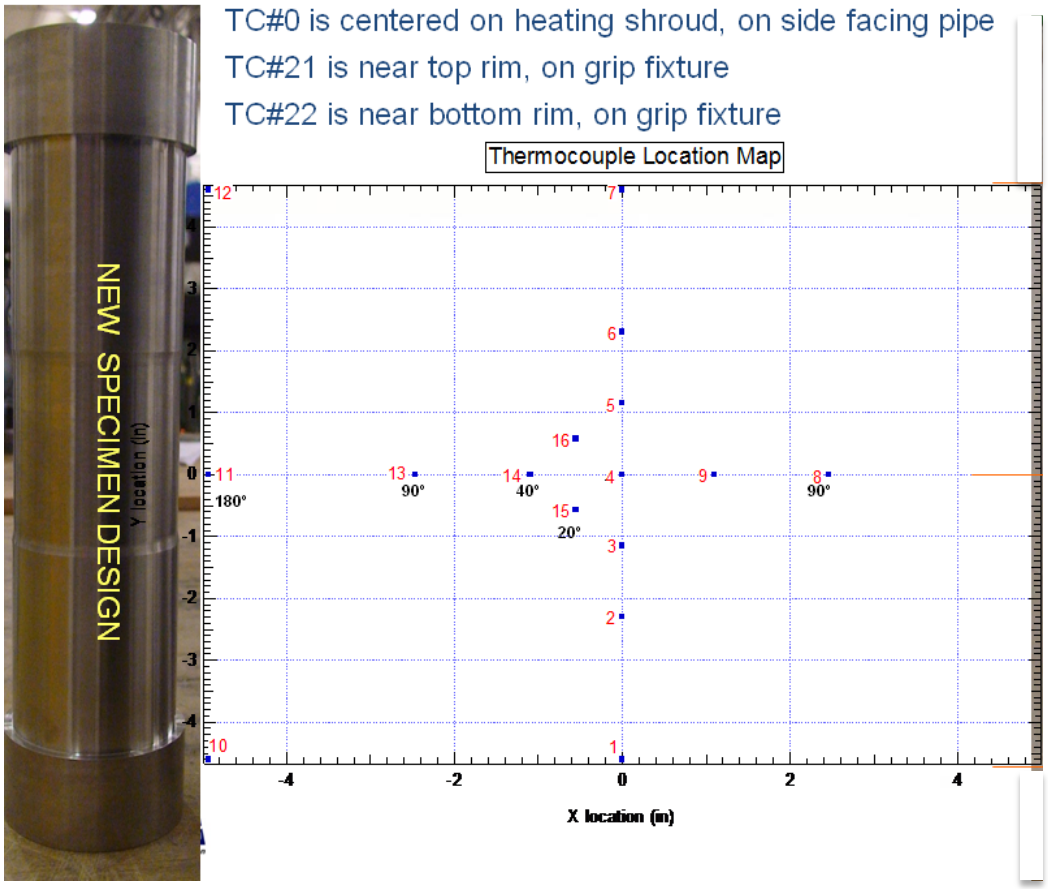


Figure 4.3 TC numbering used in the experiments, which is different from the interpolation code's numbering scheme shown in Fig. 4.2.

The layout of TCs in Fig. 4.2 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 [18, 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 for the following reasons. Temperatures in this region varied only

slightly around the pipe because its top and bottom rims were held to approximately 293K by active cooling from water jackets.

Two thermocouples (TCs 5 and 12 in Fig. 4.2) were located to provide for 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 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 4.4. was used.

4.3 Use of temperature mapping/interpolation to adjust experiments to compensate for convection effects

Figure 4.4 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. 4.3). 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.

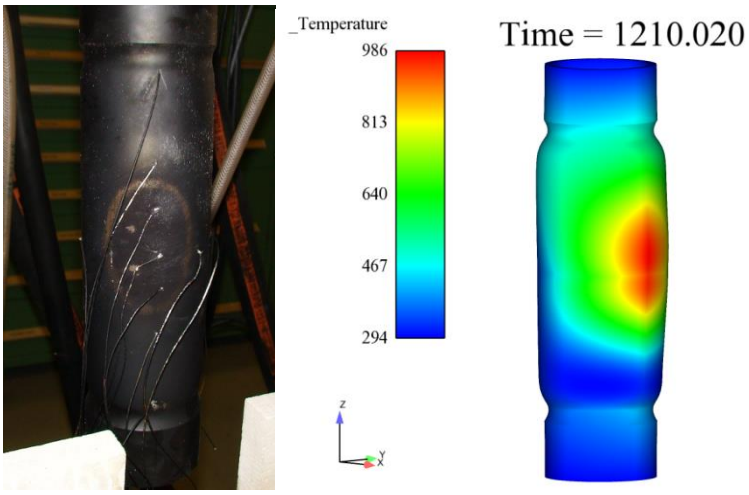


Figure 4.4 Convection-caused non-symmetric temperature field in early test setup. Heating shroud was subsequently lowered to move hot spot toward center thermocouple.

4.4 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 because of time and resource limitations in the project.

4.4.1 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 5. 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 5.1 and [2] for further description of the test conditions.

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 is indeed the case.

The pipe pressurization and heating-plate temperature inputs to the coupled thermal-mechanical simulations are plotted in Figure 4.5. The 4tt mesh (Figure 3.6) and solver settings in Table 3.4 were used in the following simulations.

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 2.5.1. 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.

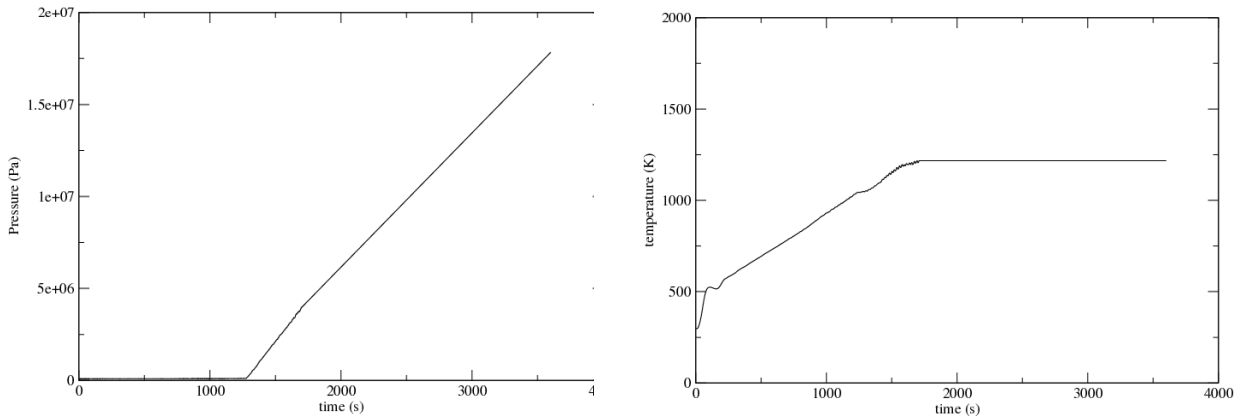


Figure 4.5 Test PB2 pressure and heating-plate temperature inputs to coupled thermal-mechanical simulations.

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 to 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 0.5. 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 4.1. 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 4.1 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	2255.9	759.0	791.6	5.40
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 4.1 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. The hot-spot peak temperatures (at TC 10 location in Fig. 4.2) 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 4.1 are 52 C higher when emissivity = 0.84 than when emissivity = 0.7, whether high or low strength material is involved. Figure 4.6 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 4.6, actual temperature variations between TCs are not linear. The actual variations are shown in the next section.

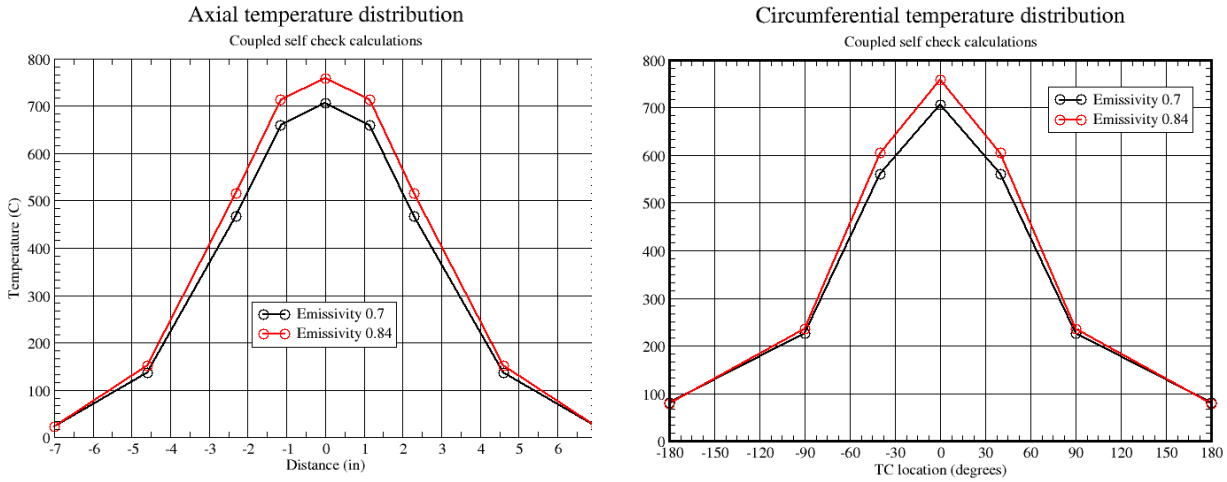


Figure 4.6. Coupled simulations temperature results (degrees C) plotted at thermocouple locations up and down front of pipe (left plot) and around pipe at mid-height (right plot). These locations correspond to TCs on *z* and θ axes in Figure 4.2 that the temperature interpolation scheme is built on. 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 4.7).

Although changing from high to low strength negligibly impacts the temperature field, Table 4.1 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).

4.4.2 Temperature mapping and interpolation errors and impact on calculated failure pressures

Figure 4.7 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 temperature field BCs on the pipe surface obtained from mapping/interpolation of synthetic temperature data (at the 18 TC locations) generated by the coupled simulations.

Noticeable differences exist between some areas of the coupled and interp. temperature fields in Figure 4.7. Figure 4.8 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.

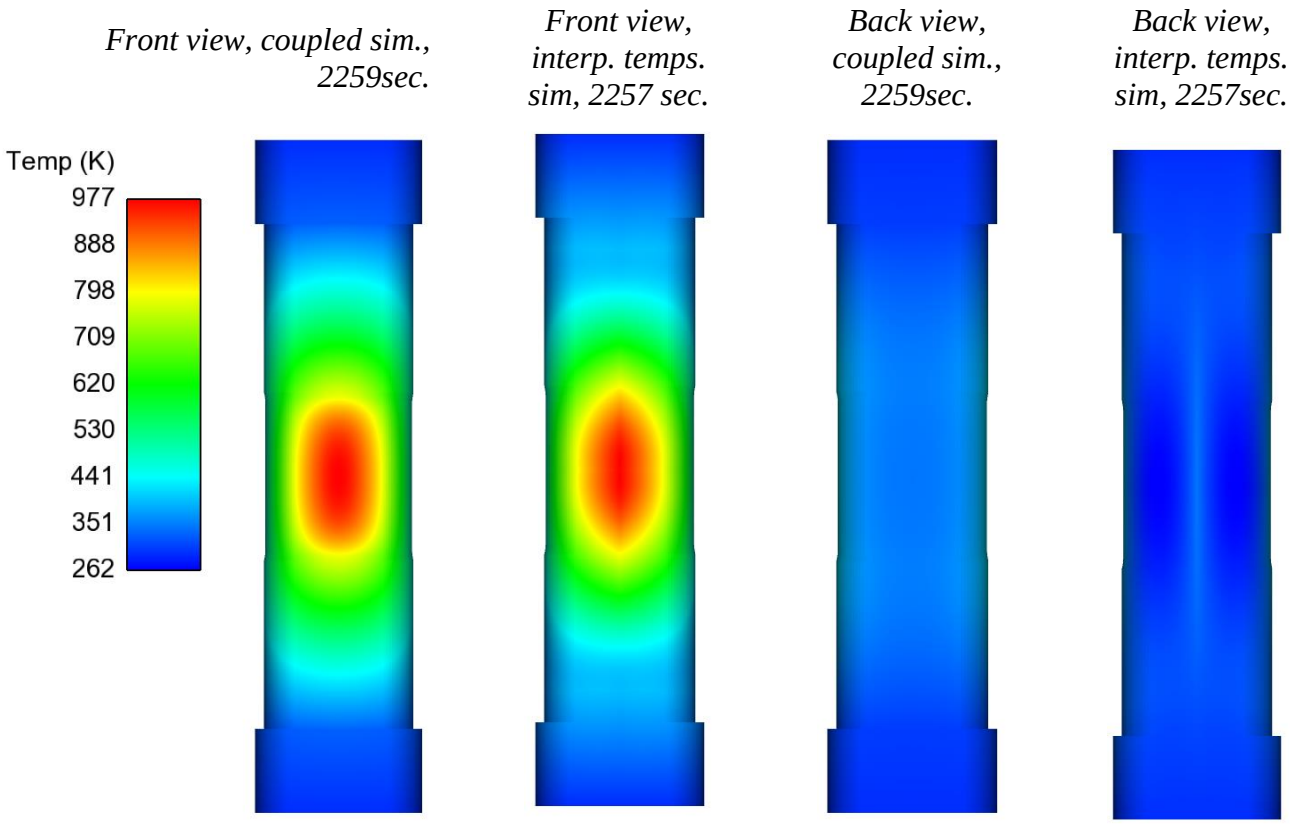


Figure 4.7 Reference and interpolated temperature fields for simulations with emissivity=0.7 and low-strength material curves.

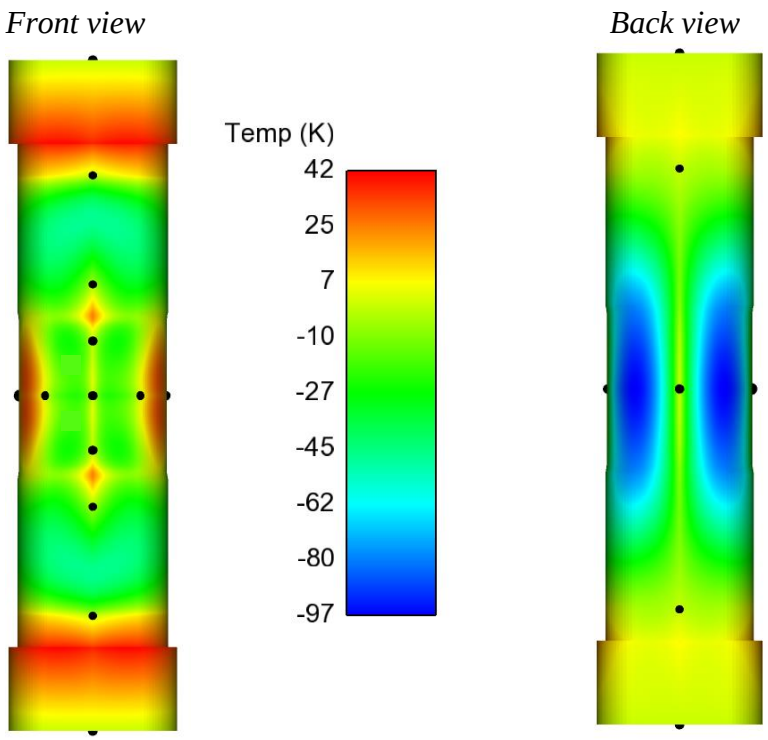


Figure 4.8 Difference plot of the interpolated minus coupled-sim. temperature fields in Figure 4.7 quantifying error in interpolated temperatures at failure for emissivity=0.7 and low-strength material curves. Zero interpolation error exists where yellow fades to green, in particular at the indicated 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 4.1 to the interp. sim. results in Table 4.2. Various comparisons are discussed next.

Table 4.2 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	875.8	3.86

The hot-spot peak temperatures in Table 4.2 are effectively the same as in Table 4.1. 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 4.1 for hot-spot temperature vs. emissivity value and material strength also apply to the mapped temperatures here.

Figure 4.9 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 very similar for the coupled and the interp. simulations.

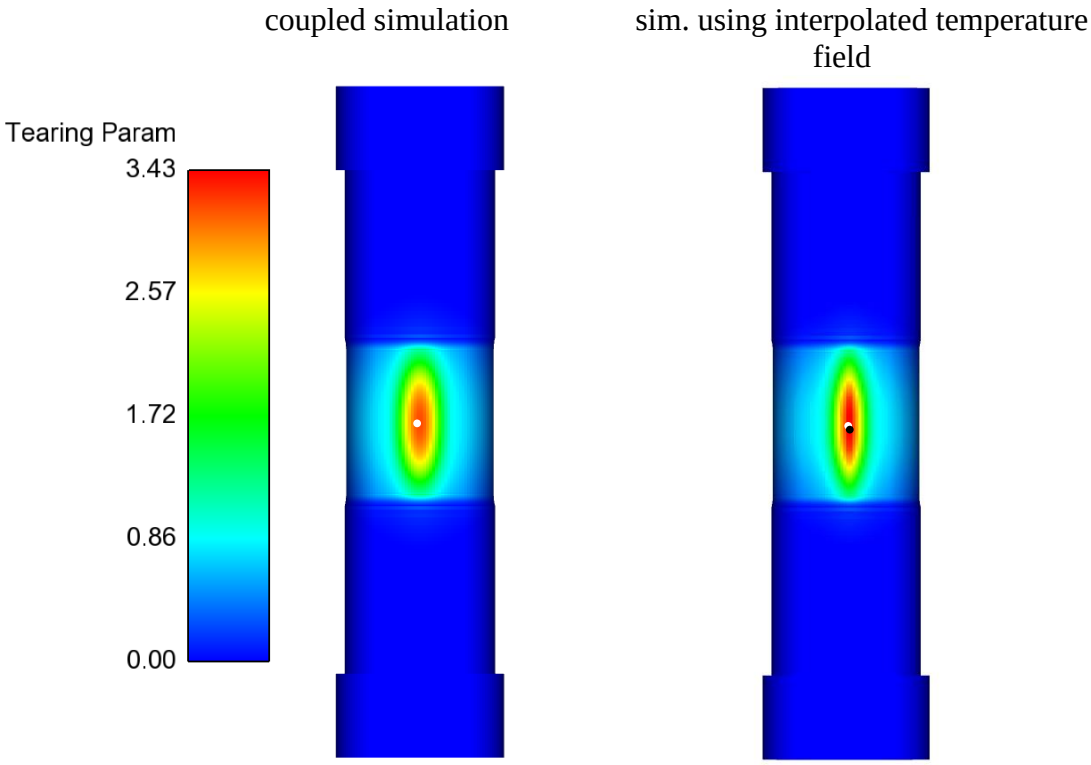


Figure 4.9 Spatial plots of computed tearing parameter indicating material damage at time of failure in emissivity=0.7 low-strength simulations. Damage is effectively confined to thin-wall section of pipe where the wall is 0.02 in. thick (wall is 0.05-in. thick elsewhere except at thicker shoulders). 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 interp. simulations predict failure at the same location and at the same hot-spot temperature. 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. simulation output values to be different from those of the coupled simulation. Tables 4.1 and 4.2 show that failure occurs earlier and at lower pressures and tearing parameter levels in the coupled simulation than in the interp. simulation.

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 4.6).

- Changing from high to low strength negligibly impacts the interpolated temperature field.

- Changing from high to low strength reduces failure pressures in Table 4.2 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 failure pressure by an average of ~6% (6.8% at high strength and 5.6% 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 5. Accordingly, Table 4.3 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 5.9%. 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 4.3 Overestimation 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	5.9 %
low-0.84	37.4 psi	4.7 %
high-0.7	12.5 psi	1.4 %
low-0.7	2.8 psi	0.3 %

4.4.3 Application to model validation simulations to correct calculated failure pressures, with uncertainty on the correction

In the previous section, 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 4.6 and 4.7, the characterized pressure prediction errors in Table 4.3 will approximately apply. Figure 4.10 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 4.3 are used to estimate failure pressure prediction errors and associated corrections (with uncertainty) in Section 5.2 due to reconstruction of the pipe temperature BCs from the tests' TC data.

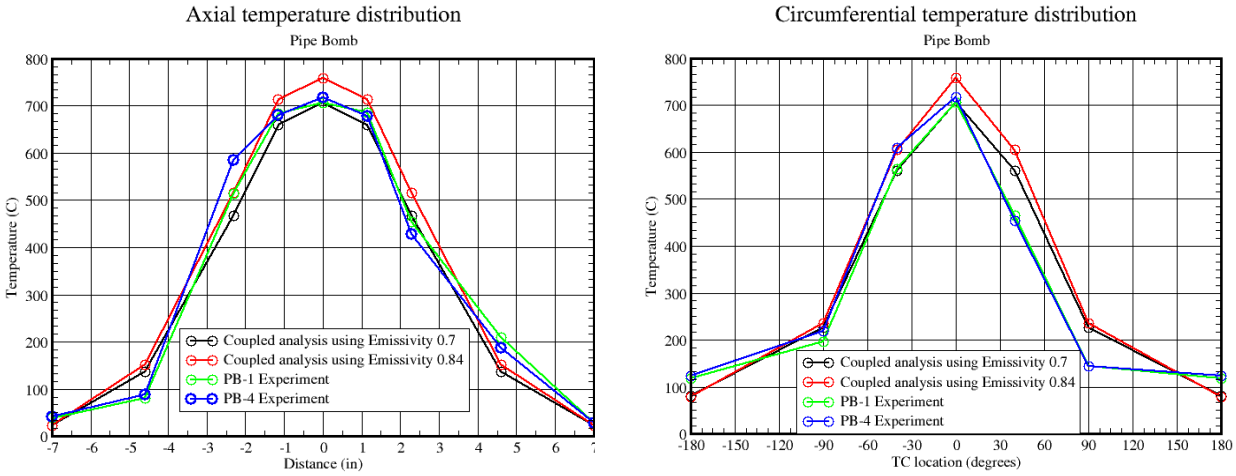


Figure 4.10 Pipe TC temperatures at time of failure in the 700C-hold experiments PB1 (green curves) and PB4 (red curves) plotted along with synthetic temperature fields (red and black curves) from Figure 4.6.

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 4.10). 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 4.3). 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 4.6. 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 5.2 to estimate error and uncertainty associated with interpolating the actual TC data from the pipe temperature fields in the experiments.

5. 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 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 [25], [26]; modeled behavior of irradiated electronics [27], [28]; and combustion in fluids and solids [29], [30]. 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. Appendix C briefly compares and constrasts the Real Space validation approach to other established model validation frameworks. See also [7] and [8] in this regard.

Various fidelities of uncertainty treatment can be applied in 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 on the order of a month on 800 processors. That was at the limit of what was feasible under the resource constraints. Nevertheless, we judge that if allowed considerably more simulations and higher-order UQ procedures, the main conclusions would not be significantly different than those arrived at in this chapter.

5.1 Experimental inputs, outputs, and uncertainty in failure pressure tests at 700C

Here we consider the 700C-hold set of pipe pressurization experiments PB1 and PB4. The tests were planned replicates of each other. Figure 5.1 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 pipes. 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 5.1. 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, from the outset of the 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, negligible 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 sections, details of hardware geometry and experimental conditions and results are presented that are pertinent to subsequent model validation procedures and comparisons.

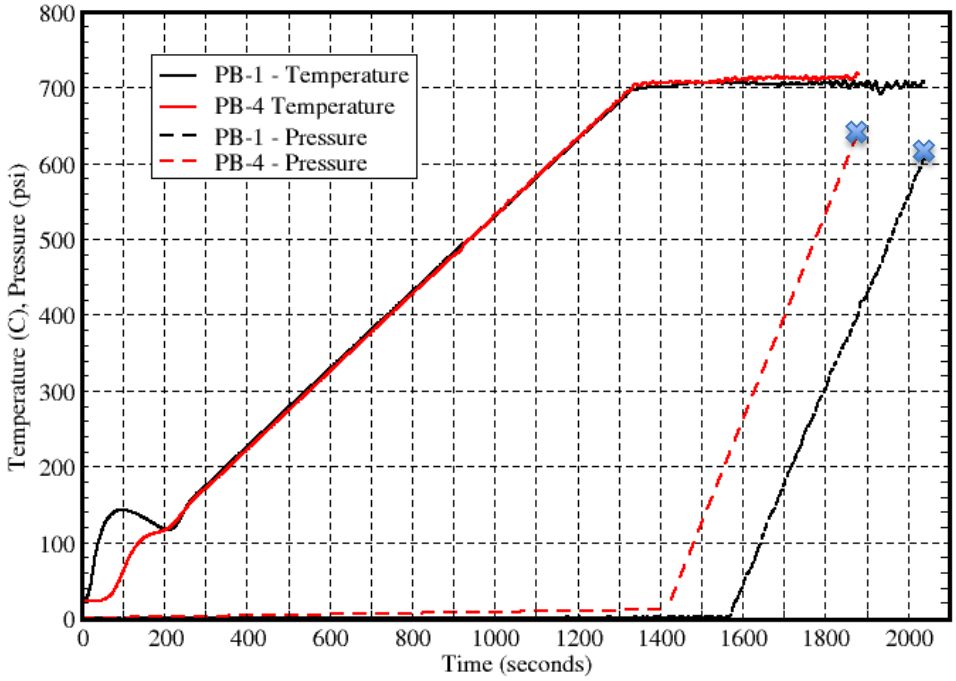


Figure 5.1 Tests PB1 and PB4 transient internal pressures (absolute) and control TC temperatures at the hot spots. Pipes fail at the pressures indicated by the X's and are listed in Table 5.1.

Table 5.1 Failure pressures measured in 700C-hold experiments. Pressure measurement uncertainties are explained in Section 5.1.1.

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

5.1.1 Experimental pressure loading and end fixturing BCs

Figure 5.2 replots the pressure loadings and shows the pipe axial loading profiles in the two experiments. The precipitous drops in the pressure readings indicate the 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 5.2. This amounts to about double the line thickness in the plots.

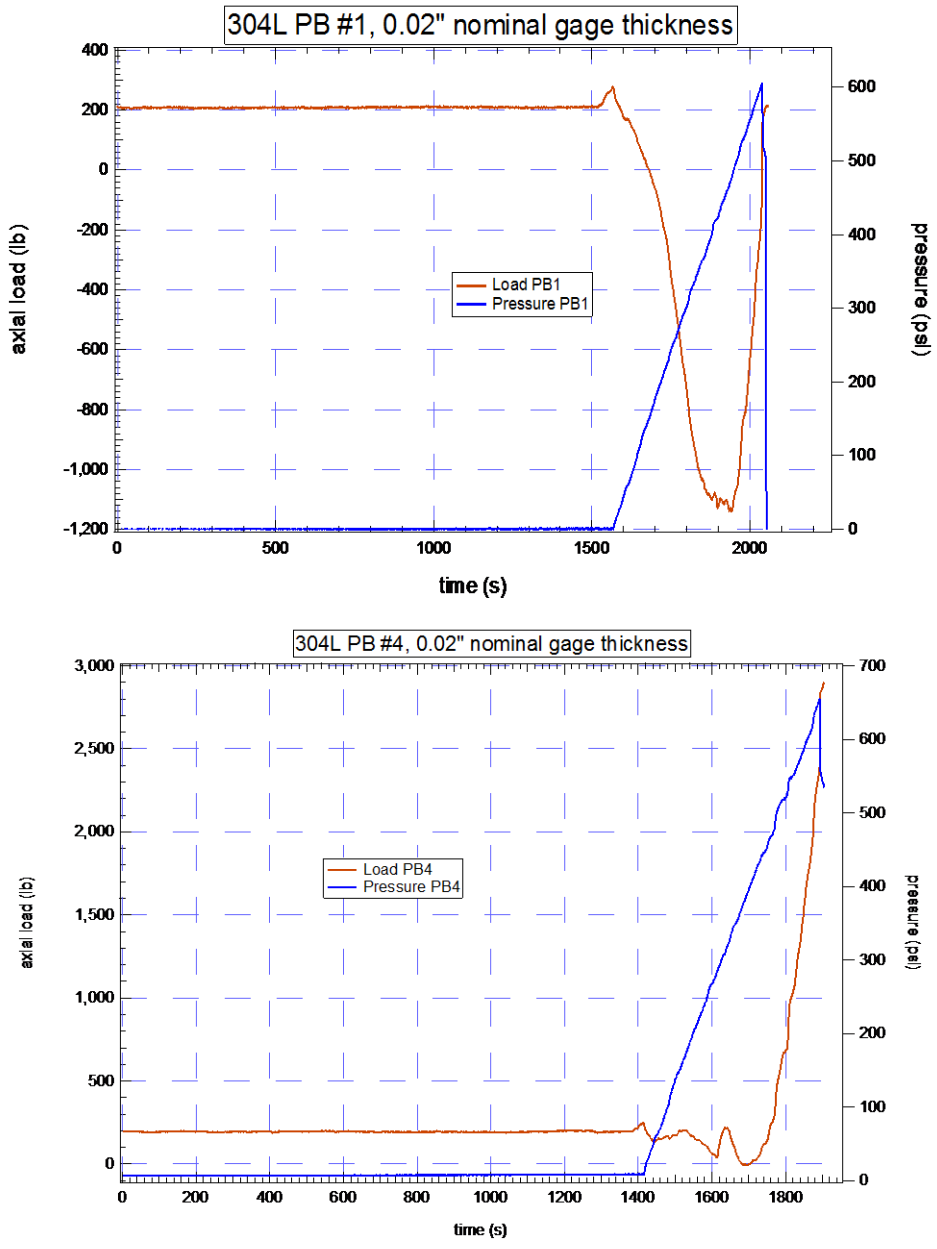


Figure 5.2 Measured pressure (referenced to scale on right axis), and axial fixture loading (referenced to scale on left axis) in 700C-hold pressurization experiments. Note similar pressure load scales in the two plots, but very different axial loading scales in the two plots.

5.1.2 Pipe surface temperature measurements (TC mapping/interpolation data)

Figures 5.3 and 5.4 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 5.2 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, 5.3 and 5.4). The temperatures in Table 5.2 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 5.3 and 5.4. 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 hold these temperatures as constant for the remainder of the simulations.

References [31] and [32] discuss 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 5.3. The largest uncertainties are indicated in Table 5.2, corresponding to the hottest TCs on the pipes. These uncertainties are less than the line thickness in figures 5.3 and 5.4.

Table 5.2 Pipe front-center control TC4 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

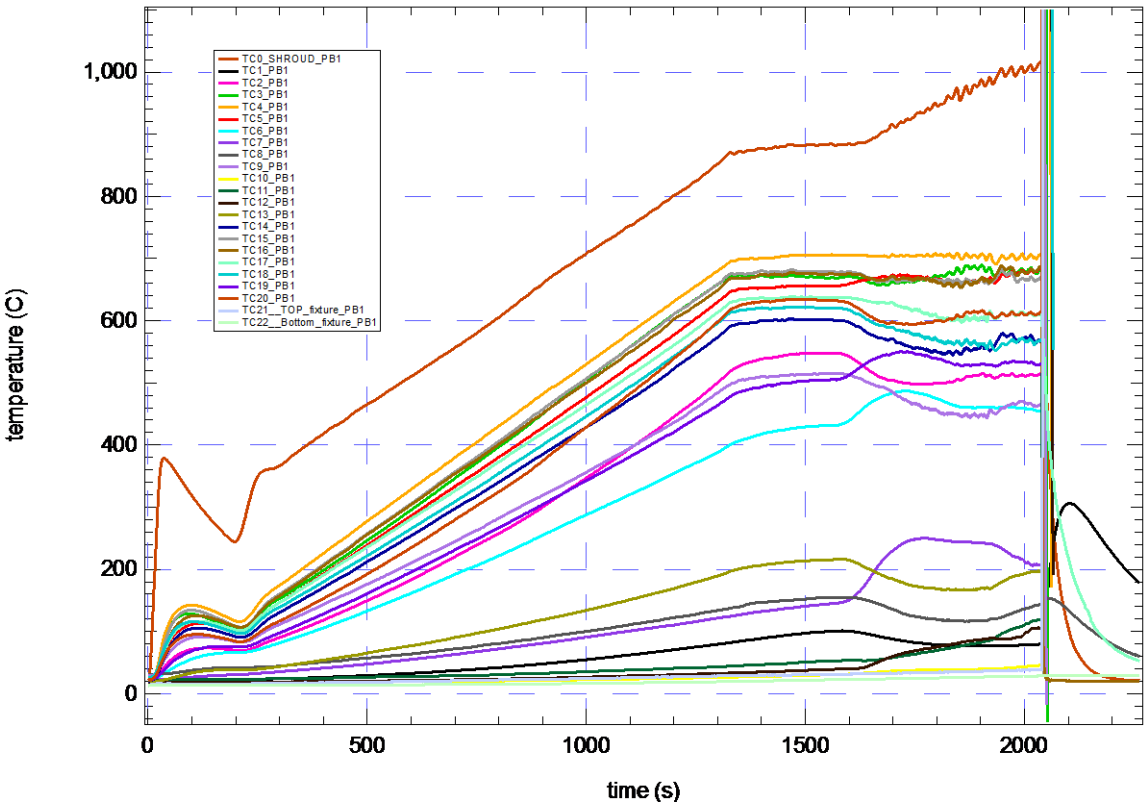


Figure 5.3 Thermocouple temperature measurements in experiment PB1. TC numbering in this figure corresponds to numbered locations in Figure 4.3.

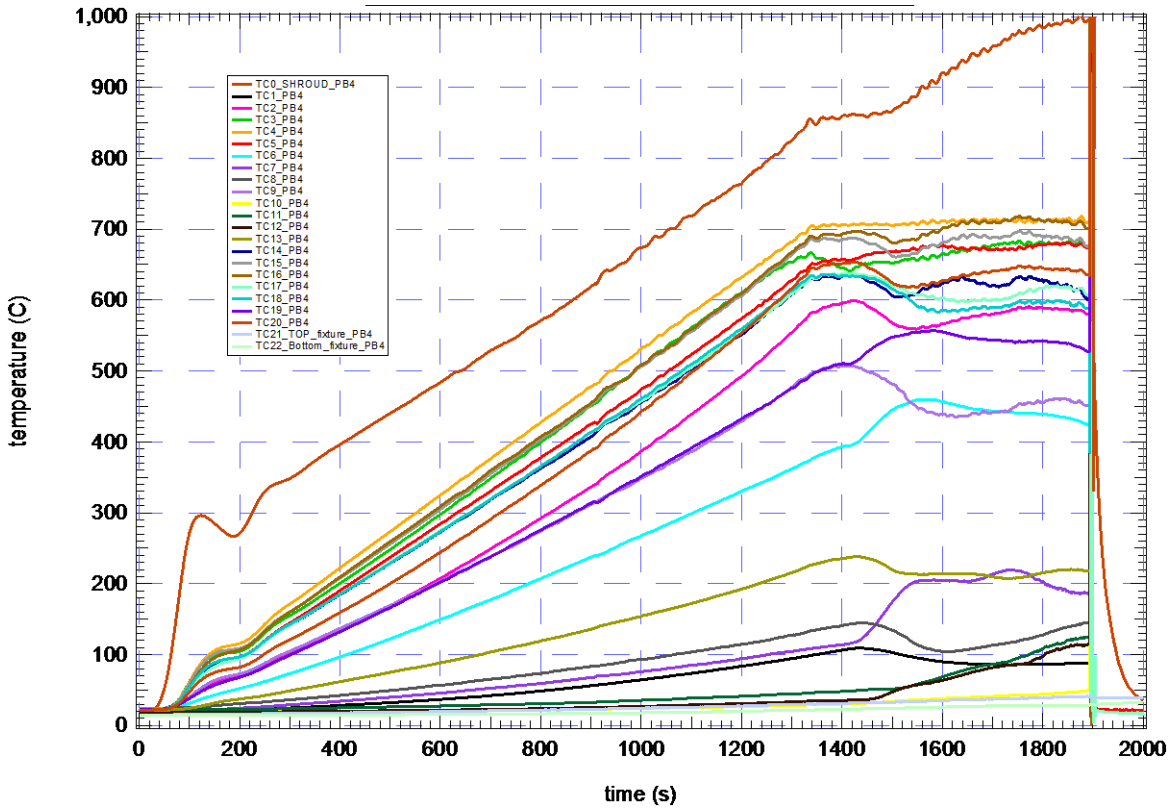


Figure 5.4 Thermocouple temperature measurements in experiment PB4. TC numbering in this figure corresponds to numbered locations in Figure 4.3.

5.1.3 Pipe wall thickness variation

Table 5.3 shows measured minimum and maximum pipe-wall thicknesses for various PB tests. 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 5.3. 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 pipes 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 5.3.

Table 5.3 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"

5.2 Simulations of validation experiments and uncertainty processing & rollup for comparisons to experimental results

In this section we address random and systematic uncertainties affiliated with the PB model and its computational solution. In the next section, 5.3, we address random and systematic uncertainties issuing from the PB experiments.

Here we use the PB1 experiment as the reference case for 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 instead 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 model validation applications.

The time-varying PB1 temperature, end loading, and pressure boundary conditions plotted in Figures 5.3 and 5.2 (top plot) are applied to a full-geometry PB model. The spec. wall thickness of 0.02" in the thin-wall section of the pipe is used, along with dimensions of the thicker sections as described in section 3. 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 section, working from the mesh refinement study in section 3.5.2. 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 in Table 3.4 were used. These settings were also used in

the mesh refinement study and in the TC temperature mapping/interpolation error characterization in section 4.4.

The full geometry model was necessary because of non-negligible effects of using the non-symmetric temperature distribution in PB1 (see Figure 4.10) vs. the ¼ symmetric mirrored PB1 temperature BCs used in the mesh study. Figure 3.6 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-section region, but the ¼ model assumes symmetry of the temperature distribution about the ¼ model's bottom and side boundaries. So its solution emulates a full model with ¼ symmetric temperature BCs. 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 PB1 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 5.5. Explanatory notes:

1. The top left corner in Figure 5.5 shows the result from a simulation with the high strength (HS) stress-strain curves.

2. A correction for use of a 1tt mesh is made using the results from section 3.5.2. From the last sentence in section 3.5.2, 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. These correction limits are applied here because the modeled PB1 experimental conditions in section 3.5.2 only differ from those in this section by the non-symmetry of the temperature field as explained earlier. Although the calculated failure pressures differ non-negligibly for the symmetric vs. non-symmetric temperature fields, it is not anticipated that convergence rates with mesh refinement will differ appreciably. We could not check this assumption.

   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 5.5. 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 5.5. Another way of expressing what we have done is to correct the 1tt-mesh 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 stress-strain curves instead of high strength curves reduces the calculated failure pressure 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 left in Figure 5.5. 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 spatially sparse TC temperature data in test PB1. A "global" element of temperature reconstruction error is addressed here. Local elements are addressed in section 5.3. Here we consider the global reconstruction errors characterized in section 4.4.2 for the synthetic temperature distributions in Figures 4.6 and 4.7. These distributions are similar to test PB1's temperature distribution (see Figure 4.10).

   At the times of failure in the experiments and similations, Figure 4.10 shows that the temperature BCs for simulations of test PB1 are bounded between the $\varepsilon = 0.7$ and $\varepsilon = 0.84$ synthetic temperature fields in the critical hot-spot region where failure occurs. From Table 4.3 the interpolation-induced 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. 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 that is the same as the global interpolation error for the actual PB1 temperature field. Then the limiting cases are:

   if $\varepsilon\_PB1 =$ **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 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 5.5 are indicated by the green line segments that end at the tops of the yellow and teal uncertainty bars respectively. These corrected pressure predictions are 827 psi and 793 psi respectively. Figure 5.6 shows the corresponding tolerance interval and Normal PDF of failure pressure variability due to material strength variability scaled from Figure 2.6.

For the limiting case **ε_PB1 = 0.84,** corrections to the HS and LS simulation results in Figure 5.5 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 5.7 shows the corresponding tolerance interval and Normal PDF scaled from Figure 2.6.

The Normal PDFs for the two bounding cases ε_PB1 = 0.7 and ε_PB1 = 0.84 are shown in Figure 5.5. The intermediate value of actual effective emissivity ε_PB1 corresponds to a Normal PDF 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 dictated by the (uncertain) value of the effective emissivity ε_PB1.

6. Last, the uncertainty of the correction for mesh effects is incorporated. In Figure 5.5 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 the starting reference value of 839 psi marked by the dot on the left-most uncertainty bar, would translate up or down as a fixed-shape pair if the reference value (the dot) lie at some other place 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 thought to be a good approximation, but we could not afford to check it.

   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. (existing reference of 839 psi + 17 psi upward translation = 855 psi = top of uncertainty bar per the last sentence in section 3.5.2.) 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 5.5.

At the other extreme, the existing reference value of 839 psi could be translated downwards by 91 psi to reach the bottom of the uncertainty bar = 747 per the last sentence in section 3.5.2. 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 5.5. The pair of upper and lower PDFs comprise a "probability box" equivalent to the CDF (cumulative density function) form introduced in [44]. Their separation signifies epistemic lack-of-knowledge uncertainty regarding where the actual PDF of stochastic results or values in the population lies.

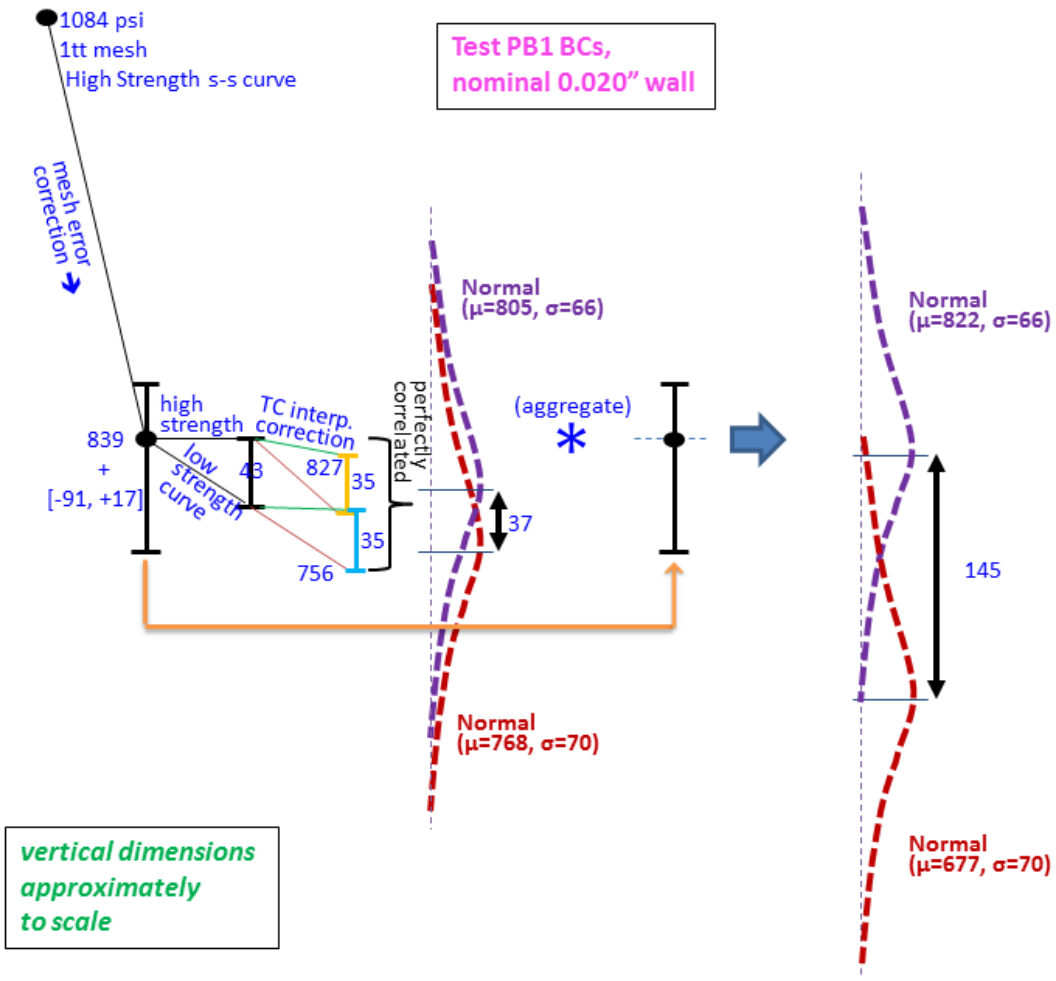


Figure 5.5 Uncertainty rollup for simulation results processed for Real Space validation comparisons.

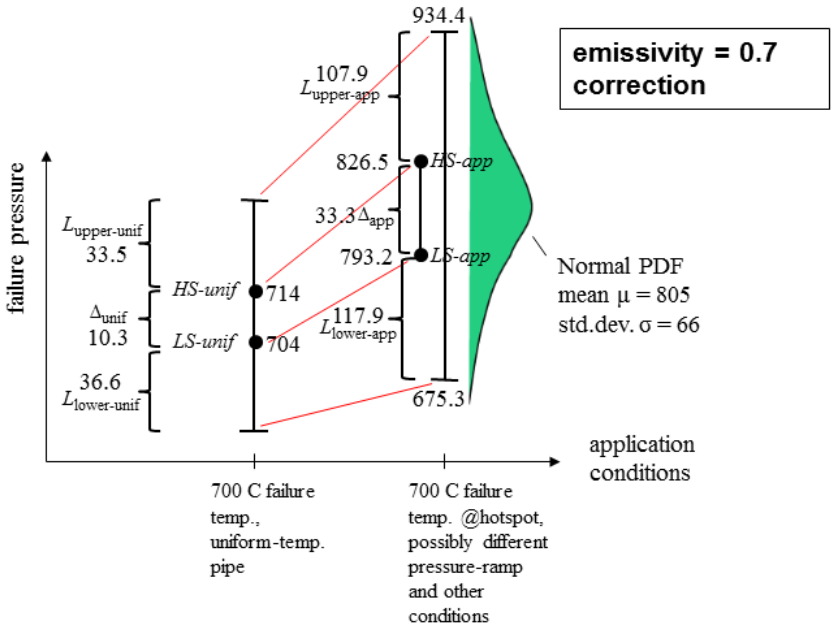


Figure 5.6 Section 2.5.2 scaling of failure pressure 95/90 TI for 700C uniform-temperature pipe to 95/90 TI and associated Normal PDF for PB1 test conditions and upper-bound (**ε = 0.7**) correction of temperature field reconstruction error.

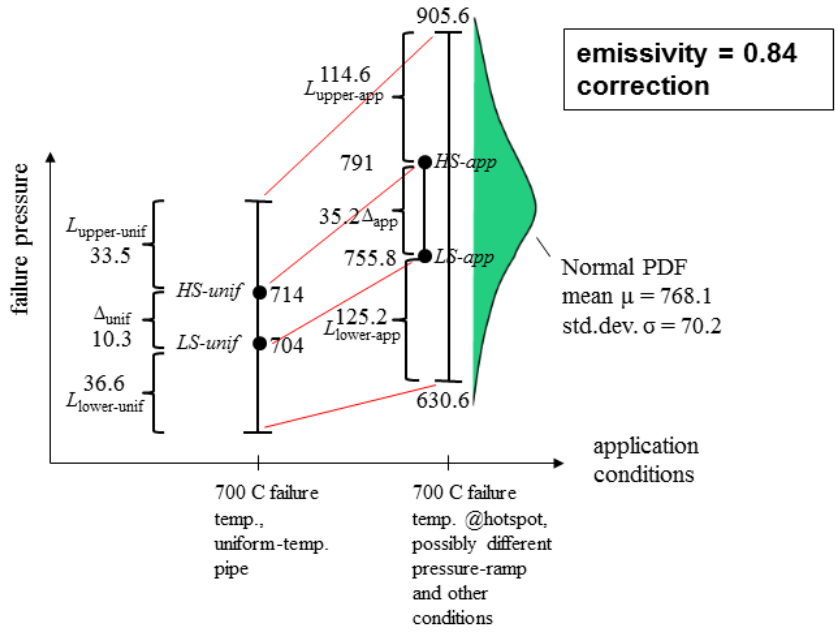


Figure 5.7 Section 2.5.2 scaling of failure pressure 95/90 TI for 700C uniform-temperature pipe to 95/90 TI and associated Normal PDF for PB1 test conditions and upper-bound (**ε = 0.84**) correction of temperature field reconstruction error.

5.3 Processing and rollup of experimental uncertainties for model validation comparisons

Here the experimental data is processed in a specific manner for comparison to the simulation results in Figure 5.5 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. To reduce UQ method complexity and cost we employ a linearized data normalization process in the following. Examples of higher-order UQ treatment are available in [28].

Let w_{act} be the actual pipe wall thickness in test PB1. From Table 5.1 the failure pressure is $Pfail_{\text{exper},w\_\text{act}}$ = 606psi. We will account for pressure measurement uncertainty later. Thus we have the input-out correspondence (w_{act}, $Pfail_{\text{PB1exper},w\_\text{act}}$). This can be thought of as particular point of a function $Pfail_{\text{exper}}(w)$ that describes experimental failure pressure as a function of wall thickness.

$$Pfail_{\text{exper},w\_\text{act}} = Pfail_{\text{exper}}(w=w_{\text{act}}) \qquad \text{Eqn. 5.1}$$

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 5.2:

$$\begin{aligned} Pfail_{\text{exper},w\_\text{ref}} = Pfail_{\text{exper}}(w_{\text{ref}}) &= Pfail_{\text{exper}}(w_{\text{act}}+[w_{\text{ref}}-w_{\text{act}}]) \\ &= Pfail_{\text{exper}}(w_{\text{act}}) + \frac{d(Pfail_{exper})}{d(w)}\bullet(w_{\text{ref}}-w_{\text{act}}) + \text{HOT.} \end{aligned} \qquad \text{Eqn. 5.2}$$

[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], [26], [33]). 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 5.2 to form the variability PDFs in Figure 5.5. 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 reflects 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 here we retain just the 1st-order term:

$$Pfail_{exper}(w_{ref}) \approx Pfail_{exper}(w_{act}) + \frac{d(Pfail_{exper})}{d(w)} \bullet (w_{ref} - w_{act}). \qquad \text{Eqn. 5.3}$$

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 5.3. Any uncertainty U[$\frac{d(Pfail_{exper})}{d(w)}$] associated with $\frac{d(Pfail_{exper})}{d(w)}$ can also be substituted. Then Equation 5.3 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 (e.g. 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 5.3 to multiple experimental inputs x_i, the normalization adjustments with respect to each of the individual inputs superpose (add linearly):

$$Pfail_{exper}(\vec{x}_{ref}) \approx Pfail_{exper}(\vec{x}_{act}) + \sum \frac{\partial(Pfail_{exper})}{\partial(x_i)} \bullet (x_{i\_ref} - x_{i\_act}). \qquad \text{Eqn. 5.4}$$

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

Any significant measurement and processing uncertainties associated with the experimental results of interest (here $Pfail_{exper}(\vec{x}_{act})$) are combined with any significant uncertainties in the partial derivatives and in the actual input values x_{i_act} on the right hand side (RHS) of Equation 5.4. 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. 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 5.4. But including 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 5.4 is demonstrated in [28]. Equation 5.4 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.

Equation 5.4 with linear UQ handles non-traveling systematic uncertainty of experimental inputs x_i in the same way as to two very differently derived linearization approaches [6] and [33]. This is reassuring.

When only systematic uncertainties are significant in the replicate experiments, the manner of uncertainty treatment and presentation of results in the RS approach can often be simplified, see section 5.5. The simplification allows a streamlined approach (e.g. [28]) vs. the spreadsheet methodology presented below. An earlier method of treatment in [29] handles all that the method below and in [28] do, but the interval approach for aggregating interval uncertainties in [29] almost certainly grossly exaggerates uncertainty when more than a couple of interval uncertainties are present. But the interval aggregation approach in [29] is simpler. It can be easily accomplished with hand calculations; it does not require statistical sampling and therefore a random number generator.

5.3.1 Normalization of reference test results

The next step in the procedure is to write Equation 5.4 for the reference test that we are normalizing all the test results to. The reference test is PB1 and we normalize it to the PB1 *nominal* input conditions used in the simulations in section 5.2. We will later write normalization equations for the other replicate experiments.

$$Pfail_{\text{PB1}}(\vec{x}_{\text{nomPB1}}) \approx Pfail_{\text{PB1}}(\vec{x}_{\text{actPB1}}) + \sum \frac{\partial(Pfail_{\text{PB1}})}{\partial(x_i)} \bullet (x_{\text{i\_nomPB1}} - x_{\text{i\_actPB1}}) \qquad \text{Eqn. 5.5}$$

We only know the failure pressure in PB1 to within the uncertainty of the pressure measurement in the test (see Table 5.1). Hence $Pfail_{\text{PB1}}(\vec{x}_{\text{actPB1}})$ on the RHS of Equation 5.5 has an uncertainty range U[$Pfail_{\text{PB1}}(\vec{x}_{\text{actPB1}})$] = 606 psi ± 10 psi = [596, 616] psi.

The two most substantial terms (by far) in the summation in Equation 5.5 involve differences between nominal and actual wall thickness, and nominal and actual temperature, at the location of failure initiation in PB1. Hence Equation 5.5 is rewritten as

$$Pfail_{\text{PB1}}(\vec{x}_{\text{nomPB1}}) \approx Pfail_{\text{PB1}}(\vec{x}_{\text{actPB1}}) + \frac{\partial(Pfail_{\text{PB1}})}{\partial(w)} \bullet (w_{\text{nomPB1}} - w_{\text{actPB1}}) + \frac{\partial(Pfail_{\text{PB1}})}{\partial(T@fail\_point)} \bullet (T_{\text{nomTC4-PB1}} - T_{\text{act@fail\_point-PB1}}). \qquad \text{Eqn. 5.6}$$

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 5.2. However, the nominal TC data are 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;
2) the actual temperature distribution may have a peak temperature that is hotter than at the TC4 location (the hottest TC location in all tests).

It is presumed that failure initiates at the location (point) where the combination of local temperature, wall thickness, and material strength determines 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 vicinity of the hottest TC, TC4. An investigation in [16] 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.

(Absent large temperature measurement uncertainties like in Appendix A which applies for much larger-diameter TCs than in the PB tests, the item 1) TC measurement errors/uncertainties in PB1and PB4 are much more important for the hot-spot TC4 than for the other TCs. Propagation of all the TCs' uncertainties (if large) to evaluate their impact involves a spatially coupled problem in the spatially distributed TCs, so evaluating each individual TC's error effect via a separate term in the summation in Equation 5.5 would not be adequate or cost effective. See Appendix A for an example evaluation of the effect of spatially correlated errors of the TC measurements and how the effect is folded into the larger validation analysis.)

The wall thickness uncertainty in PB1 is modeled as the interval range given in Table 5.3:

$$U[w_{\text{actPB1}}] = [0.019, 0.022] \text{ inch.} \qquad \text{Eqn. 5.7}$$

The term $\frac{\partial(Pfail_{\text{PB1}})}{\partial(w)}$ in Equation 5.6 is approximated by considering the experimental and model simulation results in Figure 5.8. The figure shows predicted failure pressures (listed in Table 5.4) 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 5.3. The model used was the full-geometry 1tt-mesh model used for the predictions in Figure 5.5, except for changes to wall thickness. Thus, the value in Figure 5.8 for the nominal 0.02" wall is the same as in Figure 5.5. 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 5.5. The corrected values are listed in Table

5.4. The regression line through the simulation results in Figure 5.8 has a slope of 42,286 psi/inch-wall-thickness and a y-intercept of -16 psi when extrapolated to zero wall thickness.

Table 5.4 Predicted failure pressures 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

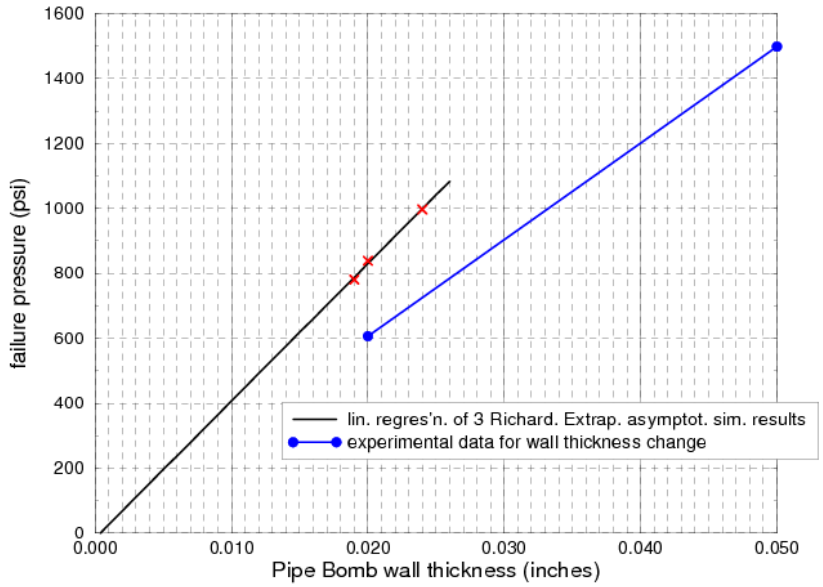


Figure 5.8 Failure pressures vs. wall thickness from simulations and experiments.

The experimental data line in Figure 5.8 connects failure pressures in the PB1 and PB3 tests (Table 5.5). 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 [16]) had a 0.05-inch thick wall and a hot-spot TC4 temperature 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 5.1 gives a glimpse of the magnitude of variation that exists among failure pressures even for nominal replicate tests.) It would 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 PB3 in Figure 5.8. Test PB4 at failure also has

close temperature conditions to PB3, but PB4 appears from Table 5.3 to have a wall thickness not as close to the 0.02" value used in Figure 5.8 and Table 5.5.

Table 5.5 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 failure pressures in the PB1 and PB4 replicate tests is 49 psi, from Table 5.1. This is used as a scale of experimental variability to work with. perturbing the experimental failure pressures in Table 5.5 by ±49 psi gives four combinations from which the lowest and highest slopes are 26,467 and 33,000 psi/inch-wall-thickness. If we liberally 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 5.8 from the model simulations. We choose to use the higher value of 42,286 psi/inch from the simulations as a liberal upper value in our UQ analysis. We use the experimentally based liberal lower value 23,067 psi/inch as a lower limit.

$$U\left[\frac{\partial(Pfail_{PB1})}{\partial(w)}\right] = [23{,}067,\ 42{,}286] \text{ psi/inch-wall-thickness.} \qquad \text{Eqn. 5.8}$$

The term ($T_{\text{nomTC4-PB1}}$ - $T_{\text{act@fail\_point-PB1}}$) in Equation 5.6 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 $\frac{\partial(Pfail_{PB1})}{\partial(T@fail\_point)}$ we normalize between the nominal temperature conditions modeled, and the actual temperature conditions in the test. For reasons mentioned earlier we normalize only for local temperature differences at the failure initiation point (close vicinity of TC4) and ignore measurement uncertainties of the other TCs. The difference ($T_{\text{nomTC4-PB1}}$ - $T_{\text{act@fail\_point-PB1}}$) is separated into two components:

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

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

The following identity is written in terms of components A and B.

$$T_{\text{nomTC4-PB1}} - T_{\text{act@fail\_point-PB1}} = (T_{\text{nomTC4-PB1}} - T_{\text{act@TC4-PB1}}) + (T_{\text{act@TC4-PB1}} - T_{\text{act@fail\_point-PB1}}).$$

= TC4 measurement error + TC4 location related error

Eqn. 5.9

Component A, TC4 measurement error, itself has several components:

$$\text{TC4 measurement error} = T_{\text{nomTC4-PB1}} - T_{\text{act@TC4-PB1}} = (T_{\text{nomTC4-PB1}} - T_{\text{ofTC4-PB1}}) + (T_{\text{ofTC4-PB1}} - T_{\text{act@TC4-PB1}}). \quad \text{Eqn. 5.10}$$

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 [31] as:

$$U[\Delta T_{\text{meas-TC/DAQ}}] = U[T_{\text{nomTC4-PB1}} - T_{\text{ofTC4-PB1}}] = [-0.25\%, 0.25\%] \text{ of TC4 reading in °C.} \quad \text{Eqn. 5.11}$$

$T_{\text{ofTC4-PB1}} - T_{\text{act@TC4-PB1}} = \Delta T_{\text{meas-contact}}$ 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 [32]:

$$U[\Delta T_{\text{meas-contact}}] = U[T_{\text{ofTC4-PB1}} - T_{\text{act@TC4-PB1}}] = [0., 0.25\%] \text{ of TC4 reading in °C.} \quad \text{Eqn. 5.12}$$

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 5.11 and 5.12 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 5.9 is due to TC4's location being potentially different from where failure occurs:

$$T_{\text{act@TC4-PB1}} - T_{\text{act@fail\_point-PB1}} = \Delta T_{\text{TC4location}} \quad \text{Eqn. 5.13}$$

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 5.13 yields $\Delta T_{\text{TC4location}} = -15\text{C}$ as the largest negative value we estimate can occur.

But in the real experiment, failure will not necessarily occur at the hottest point on the pipe due to wall thickness and material strength spatial variations. 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 5.13 yields a maximum value of 15C for $\Delta T_{\text{TC4location}}$. Thus,

$$U[\Delta T_{\text{TC4location}}] = U[T_{\text{act@TC4-PB1}} - T_{\text{act@fail\_point-PB1}}] = [-15, +15]°C. \qquad \text{Eqn. 5.14}$$

In all we have:

$$T_{\text{nomTC4-PB1}} - T_{\text{act@fail\_point-PB1}} = (T_{\text{nomTC4-PB1}} - T_{\text{act@TC4-PB1}}) + (T_{\text{act@TC4-PB1}} - T_{\text{act@fail\_point-PB1}}).$$
$$= \text{TC4 measurement error} + \text{TC4 location related error}$$
$$= (\Delta T_{\text{measTC/DAQ}} + \Delta T_{\text{meas-contact}}) + \Delta T_{\text{TC4location}}. \qquad \text{Eqn. 5.15}$$

Failure pressure difference due to different nominal and actual wall thicknesses in the PB1test is approximated in Equation 5.6 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 5.6 and 5.15):

$$\frac{\partial(Pfail_{\text{PB1}})}{\partial(T@fail\_point)} (\Delta T_{\text{meas-TC/DAQ}} + \Delta T_{\text{meas-contact}} + \Delta T_{\text{TC4location}}). \qquad \text{Eqn. 5.16}$$

The partial derivative in the above equation is approximated by the slope of the curve in Figure 2.9 evaluated for the temperature at the failure location. The slopes are calculated from the means of the data at the various temperature levels in Table 2.3. The slope in Figure 2.6 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 5.6 (and 5.16) are relative to the PB1 nominal 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,

$$U[\frac{\partial(Pfail_{\text{PB1}})}{\partial(T@fail\_point)}] = -[1.71, 2.61] \text{ psi/C}. \qquad \text{Eqn. 5.17}$$

Given the uncertainty descriptions of the applicable terms on the RHS of Equation 5.6 we are now ready to sample the uncertainties to estimate the uncertainty on the LHS term. The resulting uncertainty, $U[Pfail_{\text{PB1}}(\vec{x}_{\text{nomPB1}})]$, is the uncertainty of the PB1 failure pressure when normalized to the nominal experimental conditions input to the model simulations in Section 5.2. We obtain samples j=1,J of the uncertainty $U[Pfail_{\text{PB1}}(\vec{x}_{\text{nomPB1}})]$ by combining j=1,J realizations of the uncertainties on the RHS terms in Equation 5.6. In terms of the intervening developments, the realizations are written as:

$$Pfail_{\text{PB1}}(\vec{x}_{\text{nomPB1}})_{\text{j}} \approx Pfail_{\text{PB1}}(\vec{x}_{\text{actPB1}})_{\text{j}} + \{\frac{\partial(Pfail_{\text{PB1}})}{\partial(w)} \bullet \Delta w_{\text{PB1}}\}_{\text{j}}$$
$$+ \{\frac{\partial(Pfail_{\text{PB1}})}{\partial(T@fail\_point)} (\Delta T_{\text{meas-TC/DAQ}} + \Delta T_{\text{meas-contact}} + \Delta T_{\text{TC4location}})\}_{\text{j}} \qquad \text{Eqn. 5.18}$$

where the Δ terms have already been defined except for

$$\Delta w_{PB1} = w_{nomPB1} - w_{actPB1.} \qquad \text{Eqn. 5.19}$$

A spreadsheet is a convenient way to do the sampling and processing of the realizations. From Equation 5.18 we populate a spreadsheet as illustrated by Tables 5.6 and 5.7. The quantities in yellow highlighted columns 2, 5, 8, and 10 in Table 5.6 designate correlated uncertainties with the same-numbered columns for other tests to be described later. For all uncertainties in Tables 5.6 and 5.7 designated by interval ranges […] (e.g. from Equation 5.17), samples from uniform distributions are obtained from available sampling tools in the spreadsheet, or are imported 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 ([25], [28], [29]). A different method of treatment of interval uncertainties in [29] 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 gives a grossly exaggerated estimate 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 moderates a compounded interval treatment. But we do not interpret the resultant DF as a probability density. We use the DF to arrive at a moderated magnitude of interval uncertainty for the validation assessment (see section 5.4).

A similar moderating approach is used in [28] where both interval and probabilistic uncertainties are present and are propagated in a segregated fashion. In [25] the intervals were simply treated as uniform PDFs and comingled with the probabilistic uncertainties in propagation. The treatment in [28] 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 5.7 (and Table 5.9 etc.) are independent of each other unless correlation is intended and is purposely imposed. For the particular comparisons to be made in Section 5.4 it is recommended that at least 1000 realizations be taken ($J \geq 1000$ in Tables 5.6, 5.7, etc.). In the following we use J=1000.

Table 5.6 Spreadsheet formulas for uncertain experimental quantities in Equation 5.18 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 5.7 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 5.7 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 5.2. These realizations are histogrammed in Figure 5.9 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 5.7 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 5.2. 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 5.1. The resulting average normalized failure pressure is 586.2 psi listed in the pink box at the bottom of Table 5.7. Regarding sensitivity rankings, 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 failure 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 5.8 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 5.7. 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 ranking, 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 failure temperature normalization is contributed by the [-15, +15] failure temperature uncertainty due to uncertainty in the location of failure initiation.

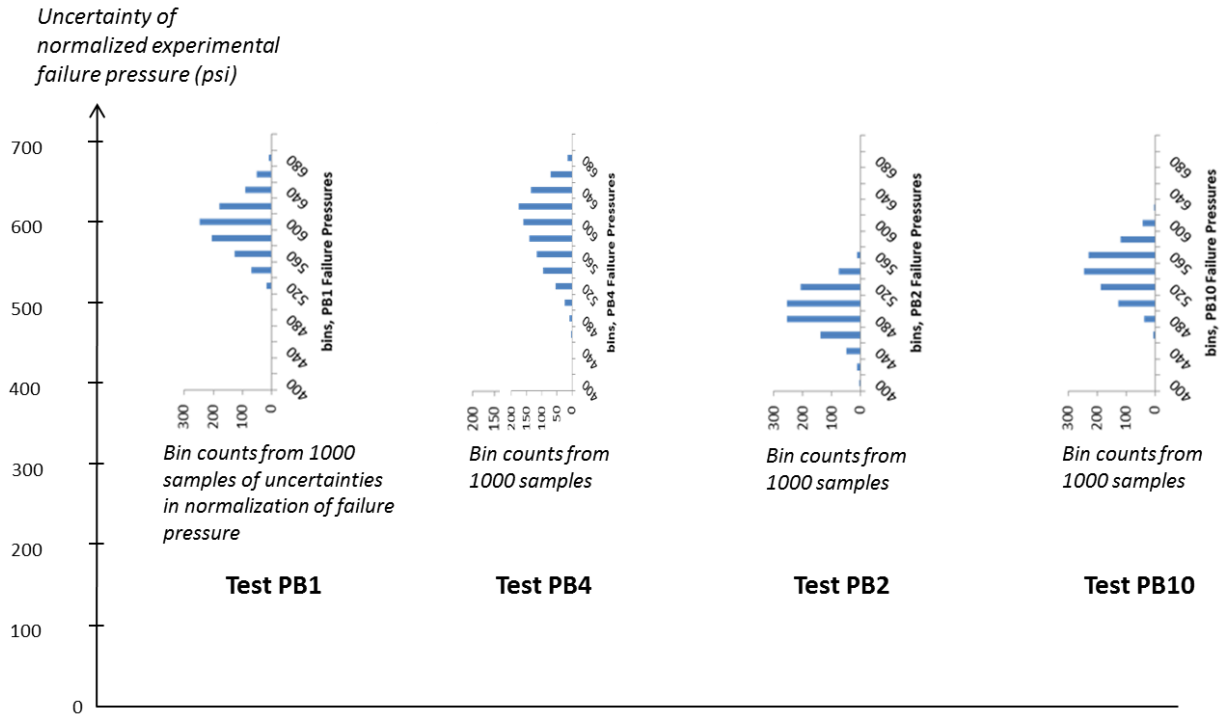


Figure 5.9 Uncertainty histograms of PB test failure pressures normalized to input conditions of model validation simulations.

5.3.2 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. For PB4:

$$\begin{aligned} Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{nomPB1}}) = {} & Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{actPB4}}) \\ & + [Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{nomPB4}}) - Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{actPB4}})] \\ & + [Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{nomPB1}}) - Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{nomPB4}})]. \end{aligned} \qquad \text{Eqn. 5.20}$$

The term in the second row of Equation 5.20 is evaluated from a PB4 analogue of Equation 5.5, to normalize test results between PB4 actual and nominal experimental conditions:

$$Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{nomPB4}}) - Pfail_{\mathrm{PB4}}(\vec{x}_{\mathrm{actPB4}}) \approx \sum \frac{\partial(PfailPB4)}{\partial(x_i)} \bullet (x_{\mathrm{i\_nomPB4}} - x_{\mathrm{i\_actPB4}}). \qquad \text{Eqn. 5.21}$$

The term in the third row of Equation 5.20 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:

$$Pfail_{\text{PB4}}(\vec{x}_{\text{nomPB1}}) - Pfail_{\text{PB4}}(\vec{x}_{\text{nomPB4}}) \approx Pfail_{\text{model}}(\vec{x}_{\text{nomPB1}}) - Pfail_{\text{model}}(\vec{x}_{\text{nomPB4}}). \quad \text{Eqn. 5.22}$$

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. But over the uncertainty ranges of the uncertain inputs $\vec{x}_{\text{i}}$ it must be sufficiently accurate in a relative sense of providing trend information (e.g. slope if linear UQ is used) such that propagating and accounting for these uncertainties improves the validation analysis vs. simply ignoring them. This is often a judgment call by the physics modelers and VVUQ analysts in the project.

Substituting equations 5.21 and 5.22 into Equation 5.20 yields

$$Pfail_{\text{PB4}}(\vec{x}_{\text{nomPB1}}) \approx [Pfail_{\text{model}}(\vec{x}_{\text{nomPB1}}) - Pfail_{\text{model}}(\vec{x}_{\text{nomPB4}})] + Pfail_{\text{PB4}}(\vec{x}_{\text{actPB4}}) + \sum \frac{\partial(PfailPB4)}{\partial(x_i)} \bullet (x_{\text{i\_nomPB4}} - x_{\text{i\_actPB4}}). \quad \text{Eqn. 5.23}$$

The term in row 1 of the RHS of Equation 5.23 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 5.5 and hence ultimately Equation 5.18. The new information required for the PB4 evaluation in Table 5.8 follows.

The wall thickness uncertainty in PB4 is modeled as the interval range given in Table 5.3:

$$\text{U}[w_{\text{actPB4}}] = [0.02, 0.024] \text{ inch.} \quad \text{Eqn. 5.24}$$

The failure pressure in PB4 has an uncertainty $\text{U}[Pfail_{\text{PB4}}(\vec{x}_{\text{actPB4}})]$ = 655 psi ± 10 psi (see Table 5.1). In Tables 5.8 and 5.9 this ±10 psi uncertainty is treated as being perfectly correlated to the pressure measurement uncertainty ±10 psi in PB1 (Tables 5.6 and 5.7) 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 to the systematic component, 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 5.8 and 5.9 are also highlighted yellow for similar reasons. The green-highlighted entries in Table 5.8 denote changes from PB1's Table 5.6. Column A12 of Table 5.6 is shifted rightward to become column B13 in Table 5.8, with other changes highlighted in green in the column B13 heading. The new column B12 represents row 1 of the RHS of Equation 5.23. 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]C in column B9 and in column A9 for PB1 are of the same magnitude but independent of each other; the realizations in a given row j of columns A9 and B9 are uncorrelated. Different initial seeds are used to sample all column quantities for PB1 and PB4 except for the yellow highlighted columns as explained earlier.

Table 5.8 Spreadsheet formulas for uncertain experimental quantities in Equation 5.23 for test PB4.

realization j	$\Delta Pfail_{\text{measPB4,j}}$ = A2j, systematic betw. PB1 & PB4	$w_{\text{actPB4,j}} = [0.02, 0.024]$ inches	$\Delta w_{\text{PB4,j}} = w_{\text{nomPB4}} - w_{\text{actPB4,j}} = 0.02'' - \text{B3}$	$\frac{\partial(Pfail_{\text{PB4}})}{\partial(w)}_{\text{j}}$ = 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	$\frac{\partial(Pfail_{\text{PB4}})}{\partial(T@fail\_point)}_{\text{j}}$ = 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 5.9 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 5.9 contains the realizations of uncertainty of PB4 failure pressure when normalized to the nominal PB1 reference experimental conditions input to the model simulations in Section 5.2. These realizations are histogrammed in Figure 5.9. The

realizations are processed with realizations from the other replicate tests as explained in the next section.

The statistics at the bottom of Table 5.9 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 5.22). Next in magnitude is the average normalization adjustment of -0.96 psi due to failure 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 5.1. The resulting average normalized failure pressure is 583.2 psi listed in the pink box at the bottom of Table 5.9.

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 failure 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 failure temperature normalization. Column B2's repeat of column A2 of course also here yields a standard deviation of 5.8 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 5.9 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 5.9.

The variance contributed by wall thickness normalization is here proportionately even greater (compared to the PB1 case) than the variances contributed by failure temperature normalization and by pressure measurement uncertainty. As in the PB1 case, the largest component of variance due to failure temperature normalization is contributed by the [-15, +15]C failure temperature uncertainty due to uncertainty in the location of failure initiation.

5.3.3 Combining possible values of normalized failure pressures to yield estimated ranges of failure population statistics

Columns A12 of Table 5.7 and B13 of Table 5.9 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 5.2). 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 5.5 and originating from material variability parameterized by differing stress-strain curves from the material characterization 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 5.5. (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 assuming 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 5.7 and B13 of Table 5.9. A convenient notional representation of the significance of the 1000 TIs is

portrayed by the associated Normal PDFs depicted notionally in Figure 5.10. The Normal PDFs in the figure only serve as a conceptualization aid; they are not constructed in the course of the analysis. The Figure Aepicts 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 5.5) under the common reference experimental conditions. Such comparisons are made and interpreted in Section 5.4.

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 5.10 (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 5.5. 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, and appear more relevant to the validation assessment of a model that is to be used for design or safety margin predictions.

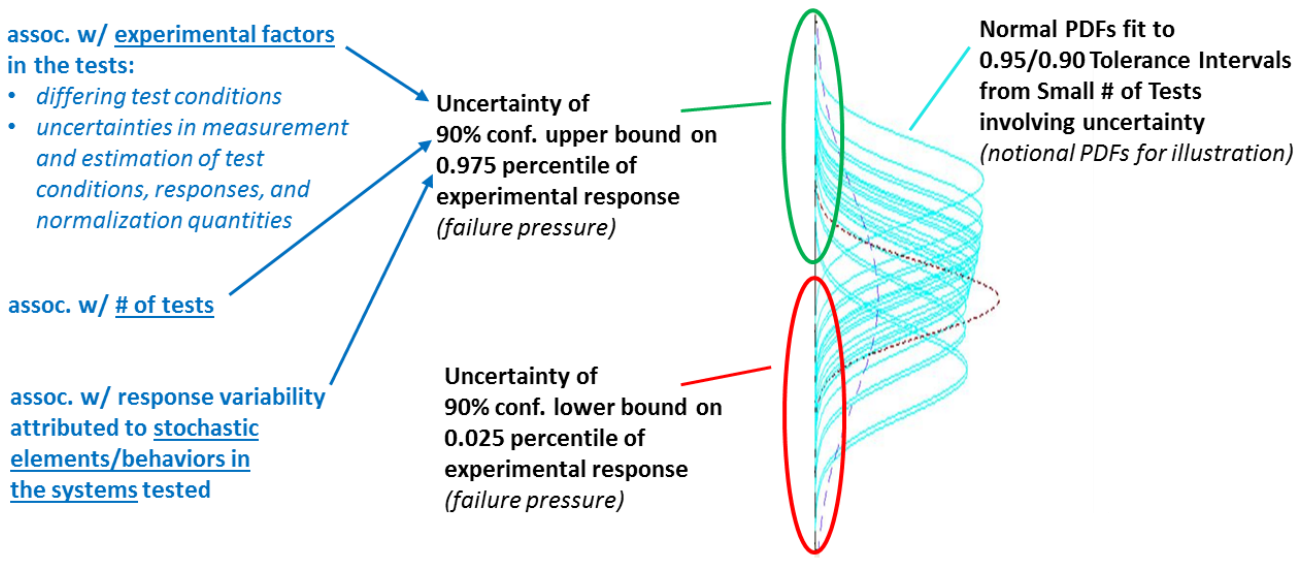


Figure 5.10 Uncertainty of statistical bounds on 0.025 and 0.975 failure pressure percentiles inferred from a small number of tests normalized to PB1 nominal experimental conditions.

5.3.4 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 B 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 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. Figure 5.11 shows the axial and circumferential temperature profiles at failure for PBs 1, 2, 4, 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 B 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 2.6 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 to normalize PB2 failure pressure for its differenent temperature profile vs. PB1 (see Equation 5.22 and column 12 in Table 5.8). 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. There was insufficient time to use the simulation model for PB 2 and 10 analogues of Equation 5.22 given the model's ~month-long run times on 800 processors.

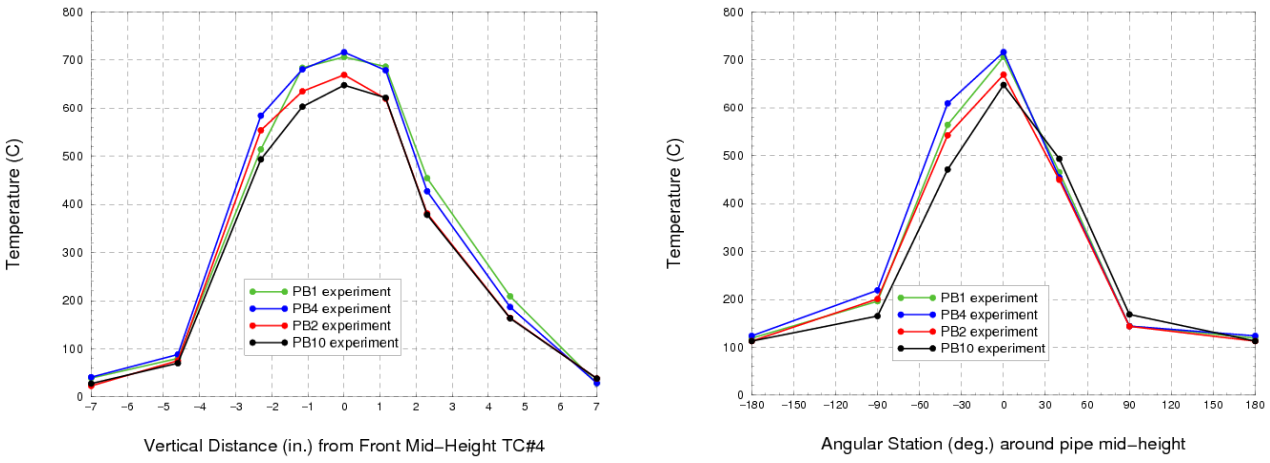


Figure 5.11. Pipe axial (left plot) and circumferential (right plot) TC temperatures at time of failures in the 650-hold and 700C-hold experiments.

Tables B.4 and B.6 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 5.9. 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 2.6 as described in Appendix B.

The J=1000 realizations for each of the four PB tests normalized to the reference PB1 experimental conditions are histogrammed in Figure 5.9. 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 5.7 and columns 13 in Tables 5.9, B.4, and B.6. Again, a convenient notional representation of the significance of the 1000 TIs is portrayed in Figure 5.10. 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 5.10. Corresponding histograms are shown at right in Figure 5.12. 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 5.5).

Table 5.10 Binned endpoints of 0.95/0.90 TIs constructed from realizations of normalized failure pressures of 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	95.10%
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%

5.4 Model validation comparisons and observations

In Figure 5.12 the histograms reflect Table 5.10's binned uncertain values of the upper and lower endpoints of 0.95/0.90 tolerance intervals from the four normalized tests. As explained earlier, the interval uncertainties input to the data normalization operations 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 5.12. 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 for 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 histrograms at right in Figure 5.12 were constructed from interval uncertainties input to the data normalization operations. But the methodology is readily extendable to non-interval uncertainties. Any of the random and systematic measurement/processing/inference uncertainties in the columns of the spreadsheets could alternatively have been PDFs or histograms or probability boxes [44]. The sampling and processing would be similar to what has been presented, but the sample realizations would be drawn from PDFs and/or histograms and/or probability boxes and/or interval uncertainties. In any of these cases, histograms like at right in Figure 5.12 would result. It is recommended that corresponding intervals like the ones at right in the figure be used for comparison to uncertainties of the model-predicted percentiles. But if all data uncertainties are represented as PDFs in the columns, the simplified treatment discussed in section 5.5 can be employed and the resulting histograms can be interpreted as a PDFs of probabilistic uncertainty. Then either the PDFs or representative spans, e.g. the 0.05 to 0.95 spans, can be compared to the uncertainty ranges of the model-predicted percentiles.

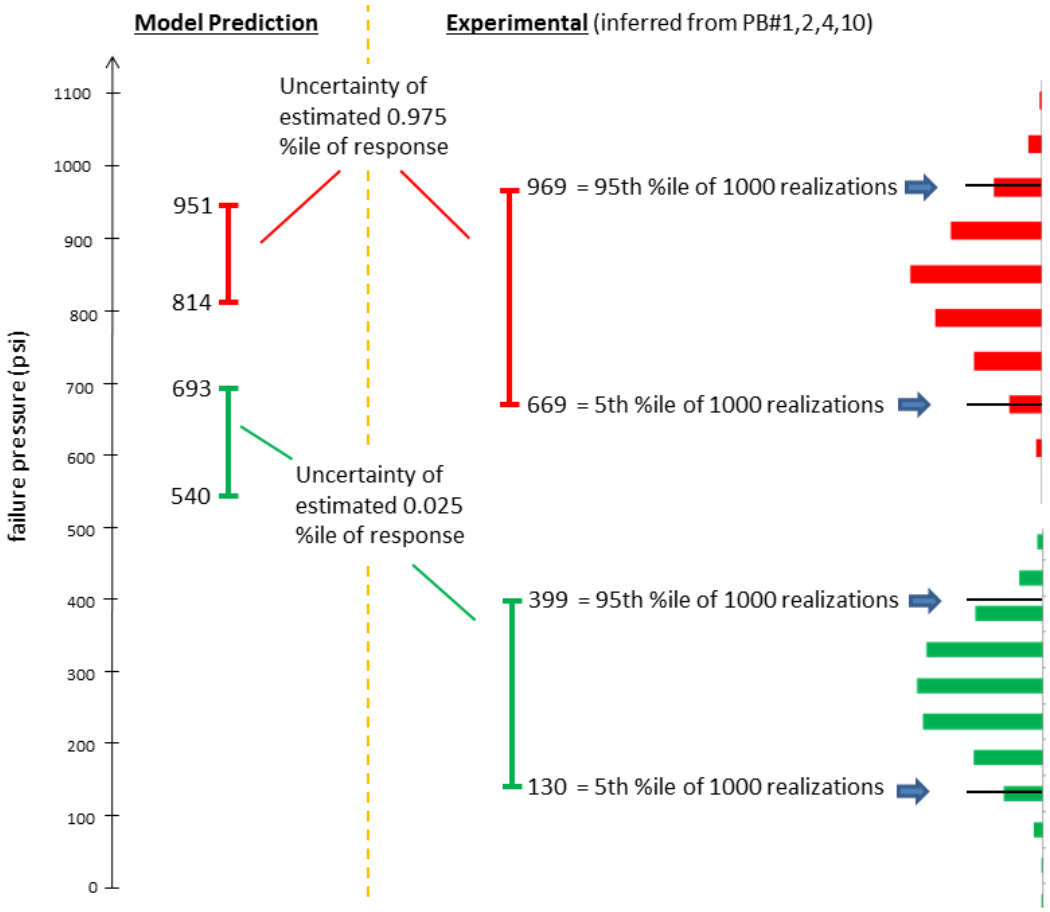


Figure 5.12 Model validation comparison of uncertainty ranges of simulation percentiles of response and normalized experimental percentiles of response.

The red and green intervals at left in Figure 5.12 are obtained from the 0.025 and 0.975 percentiles of the upper and lower Normal PDFs of predicted response in Figure 5.5. Consider the green intervals in Figure 5.12 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 5.12 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 for a more definitive assessment of prediction conservatism or not, and by how much, for specific individual percentiles of behavior. For example, the data can be reprocessed for two-sided confidence intervals on an individual percentile of interest like the 99th percentile of behavior, or can be reprocessed for one-sided confidence or tolerance *bounds* associated with a prescribed statistical confidence that a particular percentile of response meets, exceeds, or does not encroach upon a threshold response level, from below or above as specified.

Now consider the red intervals in Figure 5.12 for 0.975 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 0.975 percentile of response is higher than the predicted 0.975 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 0.975 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 0.975 percentile is up to 282 psi higher than the experimental percentile. The opposite limiting possibility is that the predicted 0.975 percentile is as low as 814 psi, while the experimental percentile is as high as 969 psi. Then the predicted 0.975 percentile is as much as 155 psi lower than the experimental percentile.

Thus the uncertainty U{predicted 0.975 percentile minus experimental 0.975 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. This combined non-traveling uncertainty defines the uncertainty resolution limit below which the model's exact bias cannot be determined (similar to the concept introduced in [34], and adopted in [22], but which does not distinguish between traveling and non-traveling uncertainties).

Because the non-traveling uncertainties of the simulations and experiments are combined, a given source of non-traveling uncertainty can be aggregated either with the simulation non-traveling uncertainties, or with the experimental non-traveling uncertainties, and the combined experimental and simulation uncertainty will be essentially the same—provided a few restrictions are met as discussed in Appendix A. Hence it is often a matter of convenience, constrained by convention and sensibility, where individual (negligibly correlated or interacting) non-traveling uncertainties are brought into the Real Space accounting ledger—either within the simulation UQ rollup or within the experimental UQ rollup. An example of such interchangeability is given in Appendix A.

How to best use a model's validation-characterized bias uncertainty to mitigate prediction risk in use of the model beyond the validation conditions is beyond the scope of this document. This is a very difficult question and an active area of research (see e.g. [6], [8], [26], [28]). See also [35] – [37] 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 5.12. Correlation (but weaker) also exists among the realizations that define the ranges of the red and green intervals on the experimental

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 firm conclusion.

5.5 Simpler treatment for problems involving suitably-random measurement errors or insignificantly small random measurement errors

Consider a situation where the exact locations of failure initiation in the four pipes were somehow known a-priori. Then the temperature and wall-thickness values at the failure locations could have potentially been measured. If measured, and if errors in the measurements are Normally distributed random perturbations about the actual values in the four tests, then the data normalization treatment in section 5.3 can be simplified and a significant decrease of overall experimental uncertainty in Figure 5.12 would result. Columns A9, B9, C9, D9 in Tables 4.6, 4.8, B.3, and B.5 would be eliminated because the uncertainty associated with location is eliminated. But more to the point of this section, a generic simplification is possible given Normality of random measurement errors over the tests. For example, let the (uncorrelated) temperature measurement-error uncertainty ranges in columns A7, B7, C7, D7 be replaced by uncorrelated Normal PDFs whose ± 3 standard-deviation extents lie at the ends of the stated ranges. Let analogous PDFs replace the (uncorrelated) wall-thickness uncertainty ranges in columns A3, B3, C3, D3.

If the normalization protocol in section 5.3 is followed, these Normal PDFs would be sampled for potential combinations of random measurement errors in the four tests. This is still legitimate under the new conditions of this section. But when linked with the 0.95/0.90 Tolerance Interval approach to account for limited #s of replicate tests (section 5.3.3), the processing although legitimate is unnecessary and perhaps yields overly conservative results. Numerical experiments for generic problems show this ([45]). Conservatism is shown in to increase as the number of replicate tests increases, as the number of measurement error sources increases (represented by the wall and temperature measurement sources here), and as the magnitude of potential measurement errors (spread of the error PDFs) increases.

Therefore, if the conditions described in the first paragraph of this section did exist, it is judged that the generic problems studied ([45]) are similar enough to support a recommendation to alter the normalization protocol from section 5.3. The random component of measurement error would be ignored. This is equivalent to having constant values in columns A3/B3/C3/D3 and A7/B7/C7/D7 instead of randomly sampling from the uncertainty ranges presently prescribed in the columns. The fixed values would be at the midpoints of the said uncertainty ranges, as these correspond to the means of the Normal PDFs of measurement errors postulated above. Based on the studies [11] and [45] it would not be surprising if this simplified treatment is also applicable for random measurement errors governed by reasonably symmetric central-tending PDF shapes

and even uniform PDFs over the stipulated uncertainty ranges. But numerical experiments have not yet been conducted to assess this.

If using a different approach such as the Pradlwarter-Schueller method [14] or bootstrapping to compensate for limited #s of replicate tests, it is presently not recommended that random measurement errors be ignored even if Normally distributed. More research is necessary. For example, over the test cases in [11] it was found that the Pradlwarter-Schueller approach is substantially less reliably conservative than the (~90% reliable) 0.95/0.90 TI approach. Sampling for the random measurement error possibilities, per the protocol in section 5.3, may serve to increase to acceptably high levels the reliability of using the Pradlwarter-Schueller approach. But this has not been investigated or established yet. Bootstrapping has also not been investigated in this context.

If random uncertainties among replicate tests are not present or are insignificantly small, or do not need to be sampled because they are Normal or otherwise suitably distributed and a TI approach is being used, then a simplified approach can be taken (see e.g. [28], calibration-data section) to process the *systematic* uncertainties in the problem if the following additional restriction is met:

- no scaling of the systematic uncertainties is present (*i.e.,* a value 1.0 exists in place of scale factors 1.07 and 0.95 in columns C5 and C8 (Table B.3) and 1.13 and 0.92 in columns D5 and D8 (Table B.5)).

The simplifications discussed in this section do not apply for the PB problem. The wall thicknesses and temperatures at the failure locations vary randomly from test to test. But the associated interval uncertainties in columns A3/B3/C3/D3, A7/B7/C7/D7, and A9/B9/C9/D9 are estimated by rather crude techniques—the variations are not actually measured (with known ~symmetric probability distributions for measurement error about the measured values). Hence the interval uncertainties reflect two types of uncertainty. A large element of epistemic uncertainty exists in the interval descriptions which attempt to characterize possible random variability in the tests. This brings into question potential similarities to the generic problems previously studied ([45]) and the scenario outlined at the top of this section, which have no error or epistemic uncertainty in their PDF characterizations of random variability in the replicate tests. So for the PB problem it is most prudent to use the more conservative interval-respecting UQ treatment in section 5.3.

6. 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.

Among the other established model validation frameworks discussed in Appendix C, the Real Space approach appears to uniquely have the required features to appropriately handle all the attributes of the PB validation problem. See [7] and [8] for further comparisons of the Real Space approach against other established model validation frameworks.

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 by experimental variations and uncertainties in the tests. Substantial simplifications, approximations, and assumptions have been made in representing and processing these uncertainties in a practical and affordable manner given the high computational expense of the model and the experimental challenges and constraints. It is judged that the large magnitude of the experimental and modeling/simulation uncertainties themselves—and not the propagation errors from uncertainty linearization and decoupling in processing the uncertainties—dominate the results. Given the uncertainties, it could not be determined whether the model is biased high or low (relative to the tests) in prediction of the 0.975 percentile of failure pressure. But at the lower end of response, model bias shows up unequivocally. The tests indicate a lower 0.025 percentile of failure pressure than the model predicts. So the model (without a factor of safety) may be unconservative for design or safety analysis purposes in that actual failures at the lower end of the failure-pressure spectrum are indicated to occur at lower pressures than the model predicts.

The experimental and simulation results could be reprocessed for similar validation comparisons of other statics of response such as mean, variance, and individual percentiles. By analogy with the presented example, the reader should be able to treat a large variety of model validation applications, issues, constraints, and purposes that arise in industrial practice.

Acknowledgments

The authors would like to recognize the contributions of retired Sandians Gerald Wellman and Martin Sherman to the work reported here. They were integral to many aspects of the project as referenced in the chapter. But we have not listed them as authors of this chapter because they were not aware of the final outcomes and analysis methodology and results reported here, and therefore could not be legitimately listed as signatories to this chapter.

The authors also thank the Advanced Simulation and Computing (ASC) program and the Weapons Systems Engineering Assessment Technology (WSEAT) experimental program of the National Nuclear Security Administration at Sandia National Laboratories for funding this work.

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] contact: Bill Scherzinger, Sandia National Laboratories, interpolation between temperature-dependent stress-strain curves in the Multilinear Elastic-Plastic (MLEP) material constitutive model

[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 Validation and Conditioning; •Hierarchical Modeling and Extrapolative Prediction," Sandia National Laboratories report SAND2011-7342, Nov. 2011.

[9] 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

[10] 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.

[11] 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.

[12] 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.

[13] Montgomery, D.C., and G.C. Runger, *Applied Statistics and Probability for Engineers*, Wiley & Sons, 1994..

[14] 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.

[15] 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.

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

[17] Scherzinger, W., "Modeling an Internally Pressurized Cylinder with Adagio," Sandia internal memo to Distribution, July 18, 2013.

[18] Romero, V., J.F. Dempsey, B. Antoun, "UQ and V&V Techniques applied to Experiments and Simulations of Heated Pipes Pressurized to Failure," Sandia National Laboratories report SAND2014-XXXX, printed April 2014.

[19] Eca, L., and M. Hoekstra, "Error Estimation based on Grid Refinement Studies: A Challenge for Grid Generation," Spanish Assoc. for Num. Mthds. in Engrg. 2009 Congress of Numerical Methods in Engineering, June 29 – July 2, 2009, Barcelona, Spain.

[20] The CUBIT Geometry and Mesh Generation Toolkit, Sandia National Laboratories, http://cubit.sandia.gov/

[21] Roache, P.J., *Verification and Validation in Computational Science and Engineering*, Hermosa Publishers, Albuquerque, NM, 1998.

[22] 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.

[23] 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.

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

[25] 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.

[26] 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.

[27] 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.

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

[29] 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.

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

[31] 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.

[32] 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.

[33] 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.

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

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

[36] Roy, C.J., and Oberkampf, W.L., "A Complete Framework for Verification, Validation, and Uncertainty Quantification in Scientific Computing," paper AIAA2010-124, 48th AIAA Aerospace Sciences Meeting, 4 - 7 Jan. 2010, Orlando, FL.

[37] Roy, C.J., and Oberkampf, W.L., "A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing," *Computer Methods in Applied Mechanics and Engineering*, 200 (2011).

[38] American Society of Mechanical Engineers PTC 19.1-2005, *Test Uncertainty*, 2006, available from ASME Codes & Standards website.

[39] Oberkampf, W.L., and Barone, M.F., (2006), "Measures of Agreement between Computation and Experiment: Validation Metrics," *Journal of Computational Physics*," Vol. 217, 5-36.

[40] 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.

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

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

[43] 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.

[44] Ferson, S., and L. Ginzburg, (1996) "Different Methods are Needed to Propagate Ignorance and Variability," *Reliability Engineering and System Safety*, Vol. 54, No. 11, pp. 133-144.

[45] contact: Vicente Romero and Conrad Woidyla, Sandia National Laboratories.

Appendix A: *Example of bias correction for spatially correlated temperature measurement errors*

Thermocouple (TC) temperature measurements for steel surfaces at temperatures very different from the effective temperature of the surroundings can have significant measurement error caused by TC contact resistance and fin effects ([Nakos et al.]). This "TC attachment" measurement bias can, for properly calibrated TCs, swamp errors associated with the data acquisition system, the calibration procedure, and TC-to-TC accuracy variations due to manufacturing variability ([Nakos]). The primary factors governing attachment bias error are: the TC wire/bead diameter; method of TC bonding or attachment to the surface; temperature differences between the measured surface and the effective radiative and convective temperatures of the surroundings; and the convective and radiative properties and conditions affecting heat transfer between the TC bead and its surroundings.

The experimental and physics modeling investigations and analysis in [Nakos et al.] provide a sufficient basis to estimate the attachment bias errors in the PB project. Early in the PB validation analysis a miscommunication occurred regarding the diameter of the TCs used in the project. It was initially understood that the TCs had a wire/bead diameter of 0.05". This lies between the 0.04" and 0.063" diameter data sets in [Nakos et al.]. Adjustments of the 0.04" and 0.063" diameter data were made for differences in geometry and the surrounding conditions in the PB experiments vs. the experiments in [Nakos et al.]. The adjusted data were then interpolated to get TC measurement biases and corresponding corrections for 0.05" diameter TCs as exemplified in Figure A.1. With the environment and geometry parameters fixed for a given PB test, TC measurement bias and associated corrections vary in sign and magnitude according to the TC's temperature and location (which dictates its radiative and convective environment). Therefore the % bias and correction varies in time and space. Figure A.1 shows that the TCs facing the heating plate register hotter temperatures than the surface they are attached to, and therefore require downward (negative) correction. The TCs on the unheated back side of the pipe are cooled by the environment they are exposed to, so they register cooler temperatures than the surface they are attached to, and therefore require upward correction.

A simulation was run with nominal corrections to the TC temperatures. The same solver settings and 1tt mesh full-model were used as for the calculations in sections 4.2 and 4.3.
The calculated failure pressure was 24 psi greater with TC correction. Lowering the hot-spot region temperatures as indicated in Figure A.1 has the effect of strengthening the pipe wall material there. So a higher pressure is required to initiate failure. Figure A.2 shows a corresponding +24 psi offset arising from TC measurement bias correction. Uncertainty on the nominal corrections was not determined because this would require more model simulations, which could not be afforded.

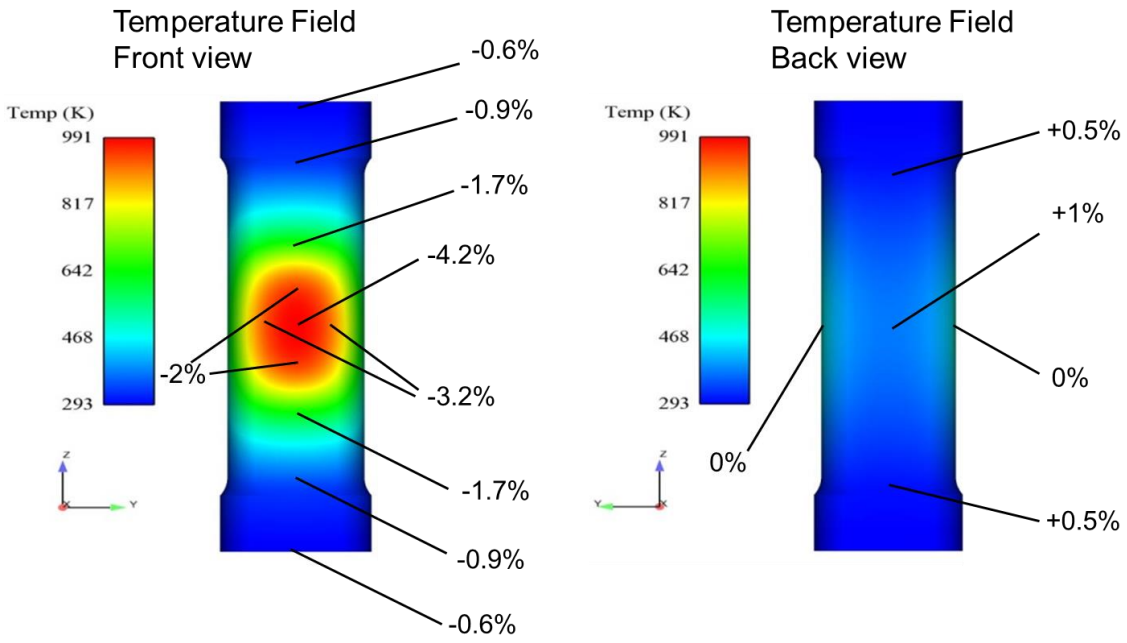


Figure A.1: Percent bias corrections to TC temperatures input to simulation of test PB1 with hypothetical large TC diameters of 0.05". Corrections are for the shown temperature field at the time of failure. Corrections are for temperature measurement errors caused by contact resistance between TC and the surface it is attached to, and by convective and radiative heat transfer between TC wire/bead and the surroundings.

It is judged that the true TC bias corrections for each TC are within +/- 50% of the nominal corrections used in the simulation. The methodology for estimating the nominal corrections arose from a single person's (the first author's) analysis and projection of the information in [Nakos et al.], so a high degree of systematic error may exist in the calculated biases and corrections. Therefore their uncertainties would be treated as highly correlated over the pipe surface. In any case, uncertainties could not be explored because of simulation budget constraints.

Later in the project it was determined that intrinsic TCs with an order of magnitude smaller diameter, 0.005", had actually been used. This made the TC-attachment component of measurement error much less important. Therefore the truncated approach explained in Section 5.3.1 was used for this component of error. A data set in [Nakos et al.] close to 0.005" diameter was used to get the TC attachment bias and uncertainty values in Equation 5.12.

Note that the approach in this appendix put the correction (and any uncertainty that would have been affordable to quantify) into the Simulation UQ Rollup. But the truncated approach in Section 5.3.1 adds the correction (and associated uncertainty) to the processing and rollup of experimental uncertainties. The correction is in the positive (+) direction in the simulation UQ rollup, whereas it is in the negative (-) direction in the experimental UQ rollup. This is consistent, and an important property of the Real Space validation method when treating non-traveling uncertainties, as we are dealing with here.

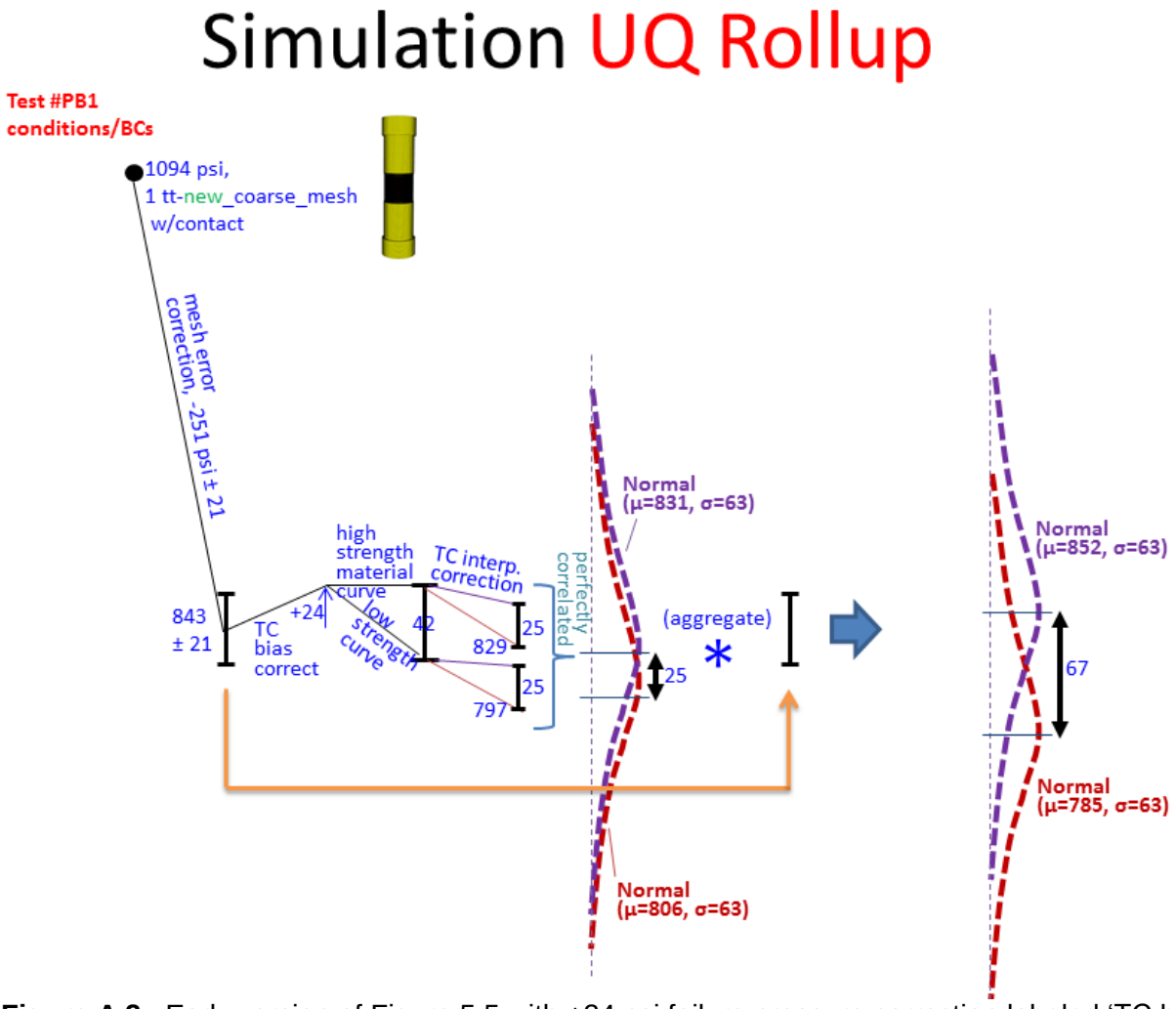


Figure A.2: Early version of Figure 5.5 with +24 psi failure pressure correction labeled 'TC bias correct' arising from nominal corrections to TC temperatures input to simulation of test PB1 (for hypothetical large TC diameters of 0.05").

For example, the indicated correction of +24 psi for 0.05" dia. TCs can be added to the simulation UQ rollup as we established for Figure A.2. This moves the simulation result up relative to the experimental failure pressures. The experimental data is not normalized with respect to this factor because the input TC temperatures in the simulation have been corrected (as close as we can reasonably come) to be the actual temperatures that occurred in the experiments. Then the simulations and experiments are on the same basis with respect to this factor.

Alternatively the bias correction can be applied to the experimental results. Then the rolled-up simulation results in Figure 5.5 apply (where no correction for TC measurement bias has been done). Then the experimental result is normalized downward by 24 psi to put it on the same basis of biased temperatures that the simulations were run with. That is, the experimental failure pressures correspond to the actual temperatures at the TC locations, not to the biased measured values input to the validation simulations. To normalize the experimental results to the temperature conditions input to the simulations (to put experimental and simulation results on the

same basis for comparison), we use the model to estimate how the experimental failure pressures would change if the pipe temperatures were perturbed to the biased temperatures in the simulations. We have already established that the model run with the biased temperatures gives a 24 psi lower failure pressure than the model run with the corrected temperatures. Thus we estimate that experimental failure pressures would change commensurately, decreasing by 24 psi and thereby moving downward relative to the simulated failure pressure.

So either the simulation results are corrected up by 24 psi relative to the experimental results, or the experimental results are corrected down by 24 psi relative to the simulation results. Either case gives essentially the same combined uncertainty range[5] for the model's prediction bias, as discussed in section 5.4. Thus it is often a matter of convenience, constrained by convention and sensibility, where individual (negligibly correlated or interacting) non-traveling uncertainties are bought into the RS accounting ledger—either within the simulation UQ rollup or within the experimental UQ rollup.

References

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

Nakos, J.T., "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 Chapter SAND 2004-1023, Released April 2004.

[5] This assumes that the same UQ propagation approaches are used for simulation uncertainty rollup and for normalization of the experimental data, something not done in the present chapter. Even if the same UQ propagation approaches are used, small differences in the combined uncertainty may exist if interval descriptions exist for some or all non-traveling uncertainties in the simulations and experiments.

Appendix B: *Description of PB2 and 10 tests and Normalization of their failure pressures to PB1 reference nominal conditions*

This appendix pertains to section 5.3.4, which introduces this appendix. Figures B.1 and B.2 plot the thermocouple temperature histories in tests PB2 and PB10. The peak hot-spot temperatures at the pipe front-center location, as indicated by TCs #4 in each plot, were ramped at a rate of approximately 31C/min., just as for PB1 and PB4.

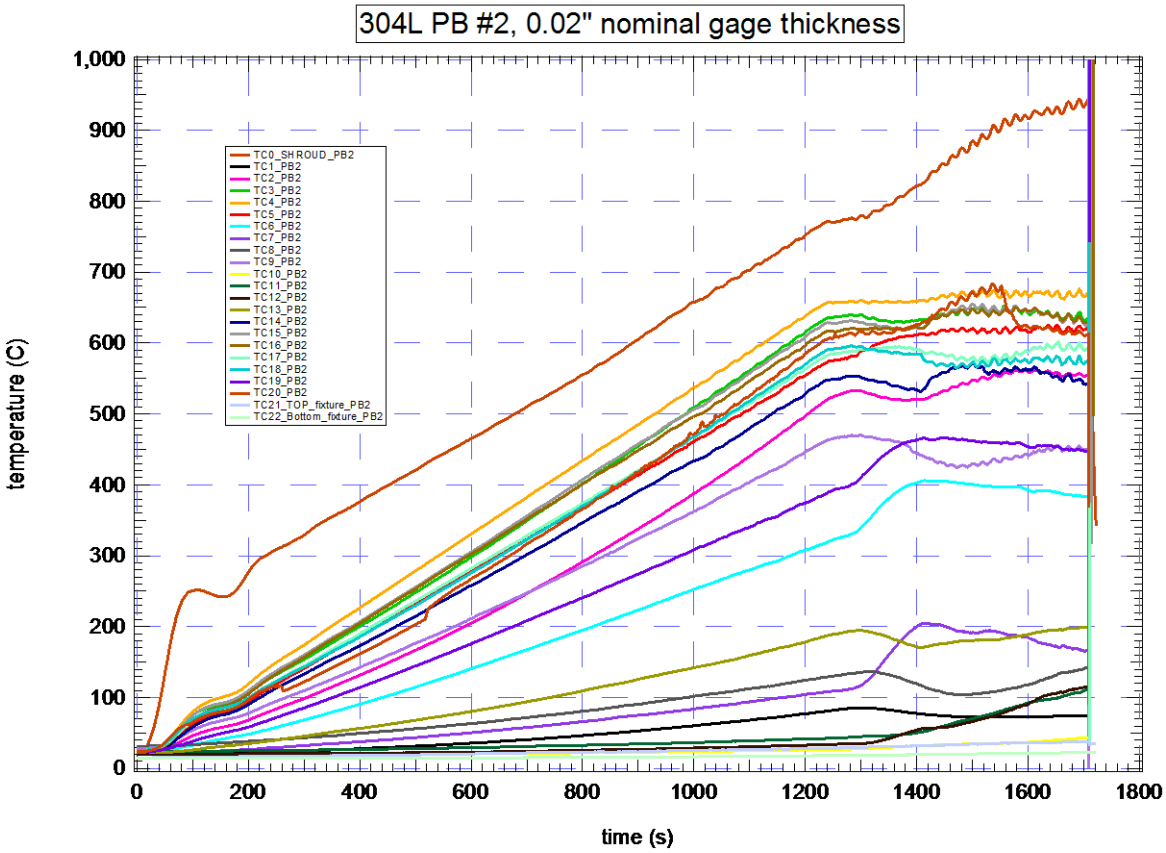


Figure B.1 Thermocouple temperature measurements in experiment PB2. TC numbering in this figure corresponds to numbered locations in Figure 4.3.

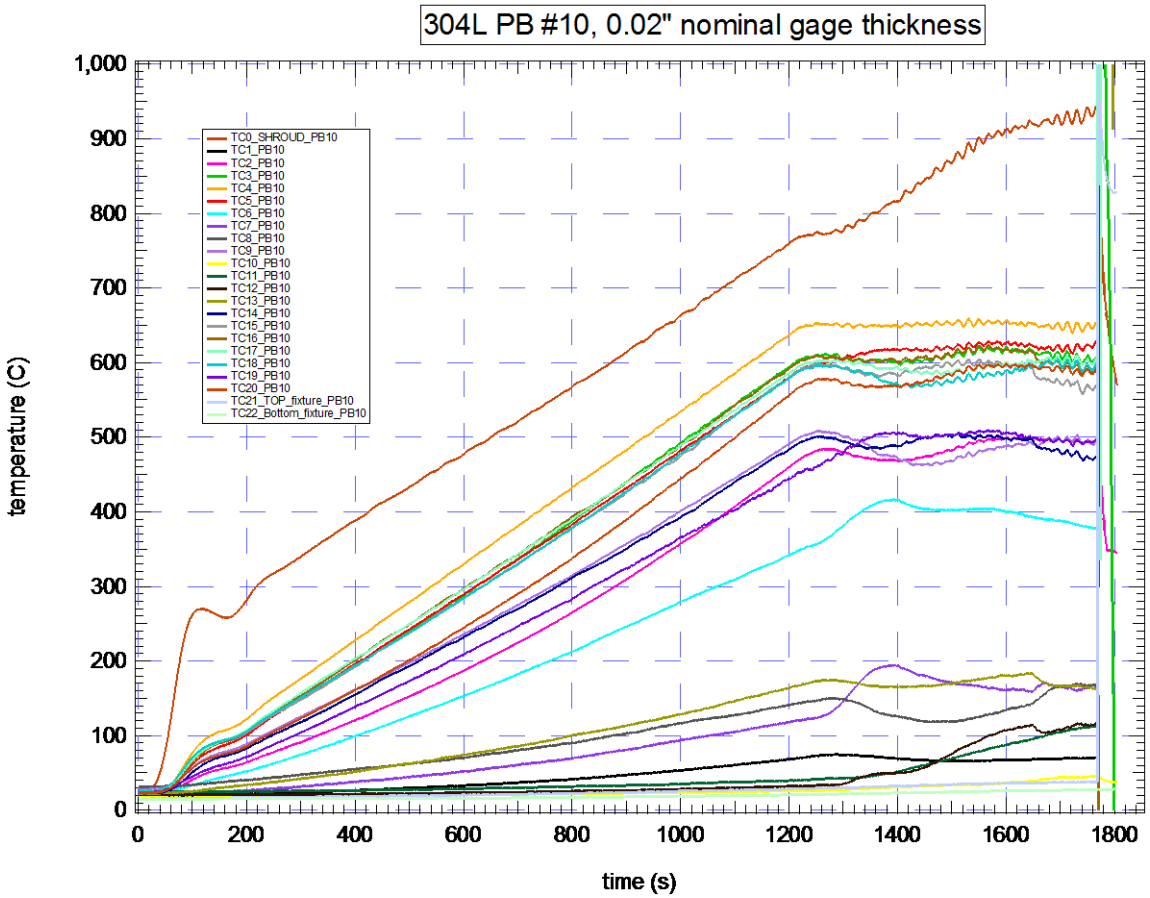


Figure B.2 Thermocouple temperature measurements in experiment PB10. TC numbering in this figure corresponds to numbered locations in Figure 4.3.

Table B.1 lists the linear regression values (over the last 60 seconds before failure) of the TC#4 temperatures. Tests PB 2 and 10 also used intrinsic TCs of 0.005-inch diameter, so very small measurement uncertainties exist due to random and systematic sources of error described in Section 5.3 (see Equations 5.9 and 5.10). The largest uncertainties are indicated in Table B.1, corresponding to the hottest TCs on the pipes. These uncertainties are less than the line thickness in the plots in Figures B.1 and B.2.

Table B.1 Pipe front-center thermocouple temperature at failure in 650C-hold experiments

Exper.	Measured temperature at failure (after regression)	% uncertainty in temperature measurement	uncertainty in temperature measurement
PB 2	672 C	[-0.25%, +0.5%] of reading in degrees C	[-1.7, +3.4] C
PB 10	648 C	[-0.25%, +0.5%] of reading in degrees C	[-1.6, +3.2] C

When the control TC4 in each test reached nominally 650C, this temperature was maintained while the pipes were pressurized until failure as shown in Figure B.3. Pressures were measured with the same pressure gage as in tests PB1 and 4. Therefore the uncertainty is within +/- 10 psi of the measured pressures in the figure. This amounts to about double the line thickness in the plots. Table B.2 lists the failure pressures and measurement uncertainties for the PB2 and PB10 tests. These measurement uncertainties are perfectly correlated with those in the PB1 and PB4 tests because the same pressure gage was used in all four tests.

Table B.2 Failure pressures measured in 650C-hold experiments. Pressure measurement uncertainties are explained in Section 5.1.1.

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

The target pressure ramp rate of 1.3 psi/sec. was met within a few % in both tests, same as in tests PB1 and PB4. PB2 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 delay effects.

Figure B.3 also shows the pipe axial loading profiles in the two experiments. The uncertainty on the axial loading measurements is ±0.1% of reading, or < ±3 lb. for the highest axial load in the tests. This is far less than the line thickness in the plots. Investigations with the PB model indicated that this magnitude of uncertainty has negligible effect on failure pressures.

Table 5.3 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.

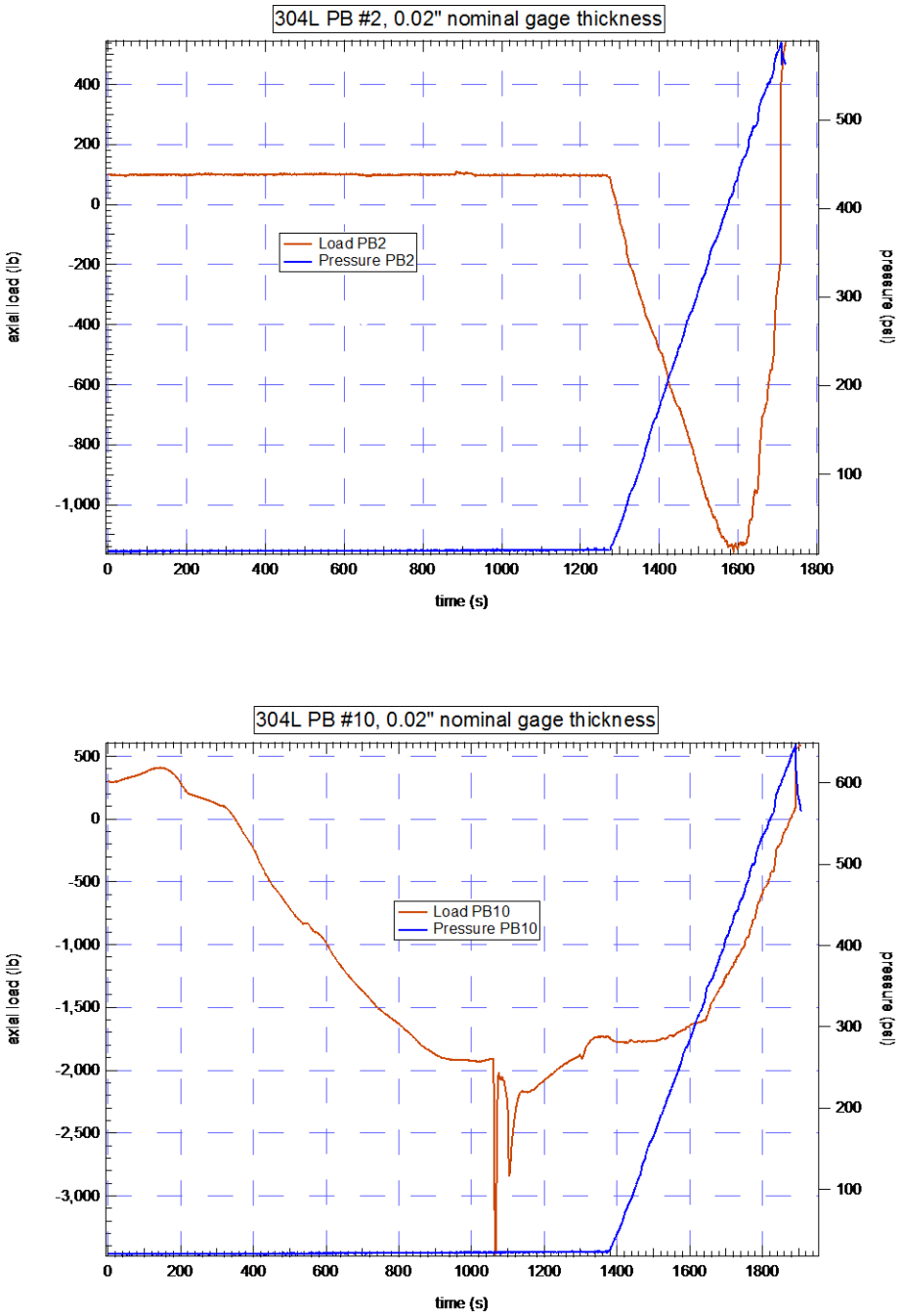


Figure B.3 Measured pressure (referenced to scale on right axis), and axial fixture loading (referenced to scale on left axis) in 650C-hold pressurization experiments. *Note that plotted PB10 results here were incorrectly shifted to zero reference time. These results must be shifted earlier by 123 sec. so that the failure time in this plot coincides with that in Figure B.2 at 1769 sec.*

PB2 normalization

Tables B.3 and B.4 list test PB2 normalization quantities and sample realizations. Again, the quantities in yellow highlighted columns C2, C5, C8, and C10 designate correlated results with the analogous columns for tests PB1 and PB4. Different initial seeds are used to sample all column quantities for PBs 1, 2, 4, and 10 except for the yellow-highlighted columns. The green-highlighted entries in Table B.3 denote changes from PB4's Table 5.8. Of particular note are the entry in column C12 and the multipliers 1.07 and 0.95 in columns C5 and C8. These are explained next.

The quantity in column C12 represents the third row of Equation 5.20, which here becomes the failure pressure adjustment for differences between the nominal input conditions for tests PB2 and PB1. It was determined late in the project to include the PB2 and 10 tests in the validation assessment, so there was insufficient time to use the simulation model to evaluate the PB2 analogue of Equation 5.22. Therefore we used the following strategy.

It is observed that the small differences in pressurization delay times and ramp rates in the tests will not yield differences in predicted failure pressures because the modeled physics are agnostic to such differences. But temperature field differences will affect predicted failure pressure. We note from Figure 5.11 that the spatial temperature field is approximately the same shape for PBs 1, 2, 4, and 10, but the fields are vertically shifted relative to each other in the hot spot region. This region, and in particular the mapped peak temperature at TC4 in the simulations, determines the calculated failure pressure far more than any of the other TCs. We make an approximation that the predicted failure pressure depends fully on the differences of the TC4 peak temperatures at failure. The applicable temperatures are then 707C @failure for TC4 in PB1 (Table 5.2) and 672C @failure for TC4 in PB2 (Table B.1). Therefore the entry in column C12 of Table B.3 serves as an approximate replacement for the term in row 1 of the RHS of Equation 5.23. The said temperature difference is proposed to affect the predicted failure pressure in the way that the other approximate temperature-related adjustments in columns C7-C9 do, via the temperature effect factor in column C10, obtained from the trends in Figure 2.6.

The multiplier 0.95 in column C8 adjusts the uncertainty of TC4 temperature measurement error $\Delta T_{\text{meas-contact}}$ caused by convective and radiative heat losses from the TC and by contact resistance between the pipe surface and the TC. An adjustment to the lower and upper uncertainty bounds in Equation 5.12 is required because the PB2 TC4 temperature at failure (672C) is significantly different from the PB1 and PB4 temperatures of 707C and 711C for which the values in Equation 5.12 apply. From [32] the uncertainty bounds in Equation 5.12 are reduced by 5% for PB2's TC #4 @ 672C. The implementation of this adjustment in column C8 preserves the correlation with realizations in columns A8, B8, and D8 for PBs 1, 4, and 10. This enforced correlation reflects the view that the $\Delta T_{\text{meas-contact}}$ error is consistent among TC4s in the PB tests rather than randomly differing. But epistemic uncertainty exists regarding the value of $\Delta T_{\text{meas-contact}}$ error at a given temperature (e.g. per Equation 5.12).

Table B.3 Spreadsheet formulas for uncertain experimental quantities in modified version of Equation 5.23 for test PB2

realization j	$\Delta Pfail_{\text{measPB2,}}$ j = A2j, systematic betw. PB1,4,2,10	$w_{\text{actPB2,j}}$ = [0.02, 0.022] inches	$\Delta w_{\text{PB2,j}} = w_{\text{nomPB4}} - w_{\text{actPB4,j}}$ = 0.02" – C3	$\frac{\partial(Pfail_{\text{PB2}})}{\partial(w)}$ j = 1.07*A5j, correlated betw. PB1,4,2,10	= (C4*C5) j psi	$\Delta T_{\text{meas-TC/DAQ,j}}$ = [–0.0025*672, +0.0025*672] C	$\Delta T_{\text{meas-contact,j}}$ = 0.95*A8j, correlated betw. PB1,4,2,10	$\Delta T_{\text{TC4location,j}}$ = [–15, 15] C	$\frac{\partial(Pfail_{\text{PB2}})}{\partial(T@fil\_point)}$ j = A10j, systematic betw. PB1,4,2,10	= (C7 + C8 + C9) j *C10j	= C10*(707C – 672C)	= 587 psi + (C2 + C6 + C11 + C12)j psi = $Pfail_{\text{PB2}}(\vec{x}_{\text{nomPB1}})$ j
C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
1												
2												
…	…	…	…	…	…	…	…	…	…	…	…	…
J												

Table B.4 Spreadsheet realizations for uncertain experimental quantities in test PB2

	C2=A2, ΔPm	C3=PB2wall	C4=Δ0.02"	C5=1.07*A5	C6 = C4*C5	C7=ΔTC_DA	C8=0.95*A8	C9=ΔTC_loc	C10=A10, ∂I	C11=C10*(C	C12=C10*(7	C13=587+C:
1	7.76	0.02026	-0.00026	32850.44	-8.39	0.71	0.49	-5.44	-2.07	8.79	-72.56	522.61
2	-5.88	0.02163	-0.00163	40485.73	-66.00	-1.57	1.12	13.93	-1.87	-25.18	-65.38	424.57
3	-0.69	0.02003	-0.00003	27944.54	-0.87	1.05	1.53	-14.44	-1.83	21.68	-63.97	543.15
999	-8.55	0.02013	-0.00013	40518.36	-5.24	-1.18	0.10	-11.53	-2.15	27.18	-75.41	524.98
1000	6.28	0.02118	-0.00118	31096.94	-36.67	0.74	0.06	9.48	-1.99	-20.50	-69.80	466.31
	avg	avg	avg	avg	avg	avg	avg	avg	avg	avg	avg	avg
	0.00	0.02099	-0.00099	34627.03	-34.15	-0.01	0.86	-0.10	-1.94	-1.46	-67.76	483.63
	stdev				stdev					stdev	stdev	stdev
	5.76				21.13					17.03	4.47	27.20
												max
												552.20
												min
												392.88

The multiplier 1.07 in column C5 adjusts the upper and lower bounding slopes $\frac{\partial(Pfail_{\text{PB1}})}{\partial(w)}$ in Equation 5.8 to PB2 temperature conditions. First it is noted that the calculated failure pressure at the nominal 0.02" wall thickness in Figure 5.8 corresponds to a PB1 hot-spot TC4 temperature of 707 C. This calculated failure pressure is then adjusted to a PB2 TC4 temperature of 672 C as follows. A temperature effect factor of $\frac{\partial(Pfail)}{\partial(T@fail\_point)}$ = -1.94 psi/C is calculated as the average of the upper and lower bounds in Equation 5.17. These bounds are derived from a consideration of Figure 2.6 over the temperature range 600C – 800C applicable here. We have

Eqn. E.1 $$Pfail_{\text{PB2-modeled}}(\vec{x}_{\text{nomPB2}}) \approx Pfail_{\text{PB1-modeled}}(\vec{x}_{\text{actPB1}}) + + \frac{\partial(Pfail_{\text{PB1}})}{\partial(T@fail\_point)} (\text{Temp\_TC4-PB2 - Temp\_TC4-PB1})$$

$$= 839 \text{ psi} + (-1.94 \text{ psi/C})(672 \text{ C} - 707 \text{ C})$$
$$= 907 \text{ psi.}$$

In Figure 5.8 we plot the temperature-adjusted failure pressure 907 psi at a wall thickness abscissa of 0.02". Through this point we create a line intersecting also the origin of the plot from physical reasoning that zero wall thickness must corresponding to failure pressure = 0. This line has a slope = 45,345 psi/in. Thus the original calculated slope 42,286 psi/inch for PB1 nominal conditions must be multiplied by an adjustment factor of 1.07 to get the PB2 temperature-adjusted slope of 45,345 psi/in. We use this same multiplier 1.07 to similarly adjust the lower-bound slope in Equation 5.8. Hence,

Eqn. E.2 $$\frac{\partial(Pfail_{\text{PB2}})}{\partial(T@fail\_point)} = 1.07 \cdot \frac{\partial(Pfail_{\text{PB1}})}{\partial(T@fail\_point)}$$

The implementation of this equation in column C5 preserves the correlation with realizations in columns A5, B5, and D5 for PBs 1, 4, and 10. This enforced correlation represents the view that failure pressure (adjusted for temperature differences as above) scales with wall thickness similarly in the tests—not randomly from test to test—but epistemic uncertainty exists regarding the scaling value as reflected in the different experimental and simulation-based estimates in Figure 5.8.

Column C13 of Table B.4 contains the realizations of uncertainty of PB2 failure pressure when normalized to the nominal PB1 reference experimental conditions input to the model simulations in Section 5.2. These realizations are histogrammed in Figure 5.9.

The statistics at the bottom of Table B.4 summarize the impacts of various uncertainties in the normalization of PB2 failure pressure. The green boxes reveal that the average adjustment in PB2 experimental failure pressure is greatest (-68 psi) due to normalization for the TC4 peak temperature difference between nominal PB1 and nominal PB2 conditions.

Next in magnitude is -34.2 psi for average normalization adjustment for possible differences from the nominal 0.02" wall thickness in the simulations. This is a much larger mean adjustment than the -17.8 psi mean adjustment for PB1 normalization, but much smaller than the -62.8 psi mean adjustment for PB4 normalization.

Next in magnitude is -1.5 psi average adjustment for the other temperature normalizations for columns C7-C9. The green box in column C2 shows a zero average bias adjustment for failure pressure measurements error. The green boxes show a combined bias shift of -103.4 psi from the nominal measured failure pressure of 587 psi in Table B.2. The resulting average normalized failure pressure is 483.6 psi listed in the pink box at the bottom of Table B.4.

The grey boxes in Table B.4 show the relative contributions to normalization uncertainty. Potential wall thickness differences are the largest uncertainty contributor (standard deviation of 21.1 psi), but this is significantly smaller than the values 28.1 psi for PB1 and 39.1 for PB4.

The square root of the sum of squares of the contributing standard deviations in the grey boxes is 28.1 psi. This closely agrees with the value 27.2 psi in the pink box calculated directly from the realizations in column C13. The normalization uncertainty for PB2 is smaller than for PBs 1, 4, and 10 as reflected by the relative widths of the distributions in Figure 5.9.

PB10 normalization

Tables B.5 and B.6 list test PB10 normalization quantities and sample realizations. The comments above concerning the yellow highlighting of columns apply here as well. The green-highlighted entries in Table B.5 denote changes from PB2's Table B.3. Changes that may need explanation are the following. **Column D8**: the uncertainty bounds of Equation 5.12 are reduced by 8% for PB10's TC4 failure temperature of 648C. **Column D5**: use Equation B.1 and change its PB2 failure temperature of 672 C to a value 648C for PB10. A result of 953 psi is obtained. Then perform the steps in the paragraph following Equation B.1, where 907 psi in that paragraph is replaced by 953 psi and PB2 is replaced by PB10. The result is the multiplier 1.13 in column D5.

Table B.5 Spreadsheet formulas for uncertain experimental quantities in modified version of Equation 5.23 for test PB10

realization j	$\Delta Pfail_{\text{measPB10,j}}$ = A2j, systematic betw. PB1,4,2,10	$w_{\text{actPB10,j}}$ = [0.019, 0.021] inches	$\Delta w_{\text{PB10,j}} = w_{\text{nomPB4}} - w_{\text{actPB4,j}}$ = 0.02" – D3	$\frac{\partial(Pfail_{\text{PB10}})}{\partial(w)}$ j = 1.13*A5j, correlated betw. PB1,4,2,10	= (D4*D5) j psi	$\Delta T_{\text{meas-TC/DAQ,j}}$ = [–0.0025*648, +0.0025*648] C	$\Delta T_{\text{meas-contact,j}}$ = 0.92*A8j, correlated betw. PB1,4,2,10	$\Delta T_{\text{TC4location,j}}$ = [–15, 15] C	$\frac{\partial(Pfail_{\text{PB10}})}{\partial(T@fail\_point)}$ j = A10j, systematic betw. PB1,4,2,10	= (D7 + D8 + D9) j *D10j	= D10*(707C – 648C)	= 647 psi + (D2 + D6 + D11 + D12)j psi = $Pfail_{\text{PB10}}(\vec{x}_{\text{nomPB1}})$ j
D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	D13
1												
2												
…	…	…	…	…	…	…	…	…	…	…	…	…
J												

Table B.6 Spreadsheet realizations for uncertain experimental quantities in test PB10

j	D2=A2, ΔPf	D3=PB10w	D4=Δ0.02"	D5=1.13*A	D6 = D4*D	D7=ΔTC_D	D8=0.92*A	D9=ΔTC_lc	D10=A10, s	D11=D10*	D12=D10*	D13=647+D2
1	7.76	0.0197	0.0003	34692.52	11.64	-1.26	0.45	14.93	-2.07	-29.28	-122.31	514.81
2	-5.88	0.0208	-0.0008	42755.96	-34.42	0.52	1.03	-13.17	-1.87	21.71	-110.20	518.21
3	-0.69	0.0204	-0.0004	29511.53	-12.22	-0.54	1.41	9.46	-1.83	-18.88	-107.84	507.38
999	-8.55	0.0203	-0.0003	42790.42	-12.74	-1.07	0.09	9.47	-2.15	-18.28	-127.12	480.33
1000	6.28	0.0208	-0.0008	32840.70	-27.41	-0.28	0.05	2.34	-1.99	-4.21	-117.66	503.99
	avg	avg	avg	avg	avg	avg	avg	avg	avg	avg	avg	avg
	0.00	0.0200	0.0000	36568.73	0.39	-0.04	0.79	0.65	-1.94	-2.69	-114.22	530.48
	stdev				stdev					stdev	stdev	stdev
	5.76				20.79					17.22	7.54	29.03
												max
												605.84
												min
												447.43

Column D13 of Table B.6 contains the realizations of uncertainty of PB10 failure pressure when normalized to the nominal PB1 reference conditions. These realizations are histogrammed in Figure 5.9.

The green boxes at the bottom of Table B.6 reveal that the average adjustment in PB10 experimental failure pressure is greatest (-114.2 psi) due to normalization for the TC4 peak temperature difference between nominal PB1 and nominal PB10 conditions. Next in magnitude is the average normalization adjustment of -2.7 psi due to the other temperature normalizations for columns D7-D9. Next in magnitude is the 0.4 psi mean adjustment due to possible differences from the nominal 0.02" wall thickness used in the simulations. This is so small (much smaller than for any of the other PBs 1, 2, and 4) because of the perfect symmetry (about the nominal 0.02") of the uncertainty range listed in column D3 of Table B.5.

The green boxes yield a combined normalization shift of -116.5 for PB10. This shift, due predominantly to PB10's much lower hot-spot temperature, is the largest combined normalization shift vs. any of the other tests (PB 1, 2, 4).

The grey boxes in Table B.6 show the relative contributions to normalization uncertainty. Potential wall thickness differences are the largest uncertainty contributor (standard deviation of 20.8 psi). This is similar to the value for PB2 but significantly smaller than the values 28.1 psi for PB1 and 39.1 for PB4.

The square root of the sum of squares of the contributing standard deviations in the grey boxes is 28.6 psi. This closely agrees with the value 29.0 psi in the pink box calculated directly from the realizations in column D13. The normalization uncertainty for PB2 is slightly larger than for PB10 but still significantly smaller than the values 33 psi and 43 psi for PBs 1 and 4, which are driven mostly by their larger uncertainties of wall thickness.

Appendix C: *Brief comparison of the Real Space model validation approach against several other established model validation approaches*

Some observations are made here concerning the applicability of several model validation methodologies to the pipe bomb validation problem. The Real Space approach appears to uniquely have the required features to appropriately handle all the attributes of the PB validation problem.

- It appears that the ASME VV20 methodology [22] and its companion document [38] on test uncertainty could be used to assess models of stochastic phenomena if the bias in mean response and uncertainty of the bias are the quantities of interest, although this is not discussed in [22]. Uncertainties due to limited #s of tests, due to stochastic variability in the systems tested, and due to random and systematic uncertainties contributed by the testing are all combined into a total "validation uncertainty". However, the Real Space method's percentile measures of stochastic behavior in Figure 5.12 reflect uncertainty of both mean **and** variance of stochastic response, and may be more relevant for assessing the accuracy of models for certain analysis purposes involving predictions of stochastic behavior/response. Another feature of the RS method is that it separates uncertainty contributions into the three categories at left in Figure 5.10. This is important primarily because the uncertainties associated with stochastic material behavior and the limited # of material characterization tests are intrinsic aspects of the constitutive model being validated for eventual use in other application settings. These traveling uncertainties are treated differently in the RS framework than the non-traveling uncertainties contributed by the execution of the tests because traveling and non-traveling uncertainties have different implications for model predictivity in post-validation use of the model (see e.g. Appendix A in [8]).

  For problems involving non-traveling experimental uncertainties, Equation 5.4 with linear UQ ultimately yields the combined "validation uncertainty" in [22] when its 1^{st}-order linear UQ version is used and the non-traveling uncertainties from the simulations are taken into account per the different approaches in the two methods—provided the conditions in the examples in [22] 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 affecting the quantities of interest. This is a reassuring corroboration of both methods (RS and [22]) for the subset of conditions cited, given that their derivations come from very different conceptual approaches.

- The Oberkampf & Barone approach [39] addresses the uncertainty elements at left in Figure 5.10 for stochastic system behavior/response and confidence levels from limited # of tests. But the measure of the stochastic behavior is limited to uncertainty of mean response. Furthermore, [39] does not address most of the uncertainties in the category 'experimental factors in the tests'. Systematic experimental uncertainties are ignored altogether. For model prediction a single deterministic model run is made for comparison

against the uncertainty PDF of the experimental mean (a very limited basis for assessing accuracy and adequacy of models of stochastic phenomena).

- The validation approaches in [40] and [41] focus on models of stochastic phenomena or systems. The "area" validation metric compares CDFs (cumulative density functions) of predicted and experimental responses. A numerical value for discrepancy between experimental and predicted CDFs is obtained. A value of zero indicates perfect agreement at all CDF 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 standard deviation of response, or error of predicted percentiles of response. Furthermore, the methodology does not directly address epistemic uncertainty (including accounting for the bias toward underestimating experimental variability) from limited #s of tests—although the uncertainty indicated by the area metric is generally found to be larger for smaller #s of tests. The remaining element at left in Figure 5.10 (experimental factors in the tests) is only partially addressed. Random variations of experimental inputs and of measurement errors in multiple replicate experiments are treated, but systematic experimental uncertainties are not. In this regard [40] and [41] incur somewhat more *Model User's Risk* concerning systematic 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.

- Two validation approaches are presented in the ASME VV10 methodology guide [42]. One is apparently an implementation of the approaches in [40] and [41], 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, the Real Space methodology recognizes that models can have traveling *epistemic* uncertainties that are an intrinsic aspect of the model. These are not present in the PB validation problem but occur e.g. in [28], [29], and [43] as parametric uncertainties in physics modeling parameters, and in [29] as multiple plausible discrete submodels for turbulence. The RS method treats these differently from the non-traveling epistemic uncertainties contributed by testing because they have different implications for model predictivity in post-validation use of the model (e.g. [6], and [8]-Appendix A and pp. 50).