Determination of Destructive Strength and Chemical Composition Uncertainties for Material Property Verification

Janille Maragh[1], Emily Brady[1], Jeffrey Kornuta[1], Owen Lopez-Oneal[2], Peter Veloo[2]

[1] Exponent, Inc., [2]Pacific Gas & Electric Company

Organized by

https://doi.org/10.52202/078572-0042

https://doi.org/10.52202/078572-0042

Abstract

Federal regulations governing gas transmission pipelines allow non-destructive (NDT) or destructive testing (DT) techniques to verify the material properties of pipe joints. As part of its material property verification program, the Pacific Gas & Electric Company (PG&E) routinely performs pipe grade determination using NDT strength and chemical composition measurements with associated uncertainty values. Over the course of several years, PG&E has also assembled a database of DT data for pipe joints, which includes data for pipe joints both with and without known pipe grade. For the pipe joints without known grade, it is desired to determine the pipe grade to update PG&E's system of record. Any DT measurements have some degree of uncertainty associated with them that should be accounted for in the pipe grade determination process. However, DT data obtained for pipe joints may not include replicate measurements that could be used to estimate uncertainty (i.e., standard deviation), usually due to the cost and labor required to obtain that data. Therefore, we developed an alternative process to determine reasonably conservative uncertainty values for DT strength and chemical composition measurements.

For this process, we reviewed scientific and engineering literature, testing standards, and instrument manufacturer technical documents. We then obtained or calculated uncertainty values based on the literature and chose reasonably conservative uncertainty values within the element concentration ranges typically observed in low carbon pipeline steels. Additionally, we analyzed PG&E's historic database of DT measurements and used a subset of the database, which consisted of DT chemical composition and strength data for 633 pipe joints with known grades, to validate our proposed uncertainty values. We did so by comparing the predicted pipe grades calculated using the DT measurements and the proposed uncertainty values with the known pipe grades. This validation process was also used to assess the sensitivity of pipe grade predictions to uncertainty values.

Background

PG&E is interested in generating and validating a methodology for pipe grade determination using destructive (laboratory) measurement data. PG&E has asked Exponent to assist in developing this process and provide supporting technical analysis. As part of this effort, PG&E has requested proposed estimated uncertainties for destructive measurements for yield strength (YS), ultimate tensile strength (UTS), carbon (C), manganese (Mn), and sulfur (S) that could be used in their pipe joint grade determination process.

PG&E currently has a methodology for pipe grade determination; however, this methodology utilizes non-destructive testing (NDT) data for pipe grade predictions. The pipe grade prediction analyses, which are performed using supervised classification machine learning models, also known as pipe

https://doi.org/10.52202/078572-0042

grade calculators (PGCs)[1] provided by PG&E vendors, can utilize two different sets of inputs: the first requires chemical composition data (C, Mn, and S), and the second requires both chemical composition and strength (YS and UTS) data. Both sets of inputs also require outer diameter (OD) and nominal wall thickness (NWT) data for each pipe joint.

Although the destructive testing of pipe joints is considered the "gold standard," serving as both an API 5L requirement for pipe manufacturers to ensure their produced pipe meets desired specifications[2] and the basis of comparison with NDT methods for validation purposes,[3] all measurements have some degree of uncertainty associated with them. PG&E currently collects NDT data for materials verification, and the NDT tools used by PG&E have validation datasets with associated uncertainties quantified for YS, UTS, C, Mn, and S for each NDT method. As part of the grade determination process, these measurement uncertainties are propagated downstream into the PGCs.

PG&E would like to extend this grade prediction analysis approach to historic destructive test data stored in PG&E's Met database,[4] most of which was collected before PG&E's grade determination process was developed. By applying the grade determination process to this historic data, PG&E seeks to gain potentially valuable information regarding populations of pipe requiring materials verification according to 49 CFR 192.607. However, because a considerable portion of the destructive data collected and stored in PG&E's Met database does not include replicate measurements for chemical composition or tensile testing on a per-pipe joint basis, the uncertainties in these material property measurements are difficult to estimate.

In order to better align with PG&E's current grade determination process, Exponent was asked to propose reasonable uncertainty estimates to associate with the destructive chemical composition testing and destructive tensile testing data used for performing pipe grade predictions.

It is important to note that uncertainty in data collection, or measurement uncertainty, is an amalgamation of various sources of uncertainty, including but not limited to lab equipment, test procedure, technician, and material inhomogeneity. Due to experimental limitations, it can be challenging—if not impossible—to establish the specific contributions from every source of uncertainty. However, instead of attempting to estimate specific contributions from each potential source of uncertainty, an estimation of total measurement uncertainty for each material property listed above from destructive testing is pursued.

[1] These PGCs were trained, tested, and validated using destructive test data.
[2] API Specification 5L, Line Pipe, 46th edition, April 2018.
[3] 49 CFR 192.607(d)
[4] The Met database contains chemical composition and strength values collected using destructive measurements.

https://doi.org/10.52202/078572-0042

Literature Review and Analysis - Composition

Scope of Literature Review

Relevant literature was reviewed to obtain a range of uncertainty values for carbon, sulfur, and manganese as measured using cutout analysis. The cutout analysis techniques considered were combustion (LECO), for the measurement of carbon and sulfur, and laboratory spark optical emission spectroscopy (OES), for the measurement of manganese. The following references were reviewed. Note that not all sources contained information that was relevant to this study. The literature review that follows this list therefore only includes excerpts from the sources that contained relevant information.

- ASTM standards:
  - ASTM A751-2014a Chemical Analysis of Steel Products
  - ASTM E173-1993 Interlab Studies - Chem Analysis of Metals WITHDRAWN
  - ASTM E177-2014 Precision & Bias in ASTM Test Methods
  - ASTM E350-1995 (R2000) Chemical Analysis of Steel, Iron
  - ASTM E415-2017 Spark Atomic Emission Spectrometry Carbon & Low-Alloy Steel
  - ASTM E415-2015 Spark Atomic Emission Spectrometry Carbon & Low-Alloy Steel
  - E1601-1998 Standard Practice for Conducting an Interlaboratory Study to Evaluate the Performance of an Analytical Method
  - ASTM E691-2015 Interlab Study - Precision of Test Method
  - ASTM E691-2005 Interlab Study - Precision of Test Method
  - ASTM E1019-2011 Carbon, Sulfur, Nitrogen & Oxygen by Combustion & Fusion
  - ASTM E1806-2018 Sampling Steel & Iron for Chemical Composition
- Literature from the original equipment manufacturers (OEMs) of the destructive tools:
  - Brochure for the CS744 Series LECO combustion analyzer
  - Brochure for the Spectromaxx Lab Spark OES analyzer
- The following scientific and engineering literature:
  - Lab Spark OES (Mn):
    - Zhang, Yong, et al. "Comparison of the analytical performances of laser-induced breakdown spectroscopy and spark-OES." ISIJ International 54.1 (2014): 136-140.
    - Grünberger, Stefan, et al. "Analysis of minor elements in steel and chemical imaging of micro-patterned polymer by laser ablation-spark discharge-optical emission spectroscopy and laser-induced breakdown spectroscopy." Spectrochimica Acta Part B: Atomic Spectroscopy 169 (2020): 105884.
  - Combustion (C):
    - Hemmerlin, M., L. Paulard, and G. Schotter. "Determination of ultra-low carbon and nitrogen contents in steel: combustion versus electrical spark source optical emission spectrometry for steelmaking process control." Journal of Analytical Atomic Spectrometry 18.3 (2003): 282-286.

https://doi.org/10.52202/078572-0042

- Murray, Jr, W. M., and SE Q. Ashley. "Determination of Carbon by Low-Pressure Combustion Method." Industrial & Engineering Chemistry Analytical Edition 16.4 (1944): 242-248.

- Combustion (S):
  - Takada, Kunio, et al. "Determination of trace amounts of sulfur in high-purity iron by infrared absorption after combustion: removal of sulfur blank." Materials Transactions, JIM 41.1 (2000): 53-56.
  - Isham, Helen, and Joseph Aumer. "DIRECT COMBUSTION OF STEEL FOR CARBON AND SULPHUR." Journal of the American Chemical Society 30.8 (1908): 1236-1239.
  - Fulton, J. W., and R. E. Fryxell. "Combustion Method for Determination of Sulfur in Ferrous Alloys." Analytical Chemistry 31.3 (1959): 401-405.

Review of ASTM Standards

ASTM E173-1993 Interlab Studies - Chem Analysis of Metals WITHDRAWN

ASTM E173-1993 contained conceptual information about ensuring good accuracy and reproducibility when using multiple labs for analysis. This standard also defines repeatability (R_1) and reproducibility (R_2) intervals and includes equations for calculating those intervals using the number of determinations, m,[5] the between lab standard deviation, σ_L, and the within lab standard deviation, σ_w.

The repeatability interval, R_1, is defined as:

$$RR_1 = 2\sqrt{2}\frac{\sigma_{WW}}{\sqrt{mm}} \quad (1)$$

The reproducibility interval, R_2, is defined as:

$$RR_2 = 2\sqrt{2}\sqrt{\sigma_{LL}^2 + \sigma_{WW}^2/mm} \quad (2)$$

This standard, specifically Eqn. 1 above, was used to estimate the within lab standard deviation using reported repeatability values from other ASTM standards, as described in the following sections.

ASTM E177-2014 Precision & Bias in ASTM Test Methods

ASTM E177-2014 provides definitions of terms commonly used when discussing precision and bias. Selected definitions are provided below:

- *accepted reference value*, n—a value that serves as an agreed-upon reference for comparison, and which is derived as: (1) a theoretical or established value, based on scientific principles, (2) an assigned or certified value, based on experimental work of some national or international organization, or (3) a consensus or certified

[5] Note 1 of ASTM E173-1993 describes that the number of determinations, m, is distinct from the number of replicate test results collected by a lab. A test result is the average of mm determinations, which may be specified by the method. If it is not specified, then one may specify $mm = 1$, as was done for this analysis.

https://doi.org/10.52202/078572-0042

value, based on collaborative experimental work under the auspices of a scientific or engineering group.

- *accuracy*, n—the closeness of agreement between a test result and an accepted reference value.
- *bias*, n—the difference between the expectation of the test results and an accepted reference value.
- *coefficient of variation*, CV, n—for a nonnegative characteristic, the ratio of the standard deviation to the mean for a population or sample.
- *precision*, n—the closeness of agreement between independent test results obtained under stipulated conditions.

ASTM E350-1995 (R2000) Chemical Analysis of Steel, Iron

ASTM E350-1995 reported repeatability and reproducibility values for sulfur and manganese. Using the equations provided in ASTM E173-1993, within lab and between lab standard deviation estimates could be back-calculated using the reported repeatability and reproducibility values.

TABLE 1 Statistical Information—Manganese—Metaperiodate Photometric Method

Test Material	Manganese Found, %	Repeatability (R_1, E 173)	Reproducibility (R_2, E 173)
1. Alloy steel (BCS 252, 0.016 Mn)	0.022	0.004	0.006
2. Alloy steel (BCS 255/1 0.16 Mn)	0.161	0.004	0.010
3. Low-alloy steel (NBS 72f, 0.545 Mn)	0.551	0.010	0.020
4. Low-alloy steel (NBS 139a, 0.780 Mn)	0.780	0.009	0.030
5. Alloy steel (NBS, 159, 0.807 Mn)	0.819	0.010	0.034
6. Carbon steel (NBS 13f, 0.889 Mn)	0.892	0.015	0.027
7. Low-alloy steel (NBS 100b, 1.89 Mn)	1.91	0.02	0.04

Figure 1. Reported data copied from ASTM E350-1995, which shows the measured amount of manganese, the repeatability interval, and the reproducibility interval for seven test materials.

https://doi.org/10.52202/078572-0042

TABLE 3 Statistical Information—Sulfur—Combustion-Iodate Titration Method

Test Material	Sulfur Found, %	Repeatability (R_1, E 173)	Reproducibility (R_2, E 173)
Induction Furnace			
1. No. 1, E 352	0.006[A]	0.002	0.003
2. No. 2, E 352	0.008[A]	0.001	0.004
3. Low-alloy steel (NBS 11 lb, 0.015S)	0.014[A]	0.003	0.003
4. Carbon steel (NBS 13f, 0.016S)	0.016[A]	0.002	0.002
5. Carbon steel (NBS 152, 0.027S)	0.026[B]	0.003	0.004
6. Carbon steel (NBS 16d, 0.033 S)	0.032[B]	0.003	0.005
7. Carbon steel (NBS 129b + 8i (Mixed), 0.144 S)	0.141[C]	0.007	0.013
8. No. 7, E 353	0.286[D]	0.014	0.020
Resistance Furnace			
1. No. 1, E 352	0.006[A]	0.001	0.002
2. No. 2, E 352	0.009[A]	0.001	0.002
3. Low-alloy steel (NBS 11 lb, 0.015S)	0.014[A]	0.001	0.003
4. Carbon steel (NBS 13f, 0.016S)	0.015[A]	0.002	0.003
5. Carbon steel (NBS 152, 0.027S)	0.027[B]	0.004	0.004
6. Carbon steel (NBS 16d, 0.033S)	0.032[B]	0.003	0.004
7. Carbon steel (NBS 129b + 8i (Mixed), 0.144S)	0.140[C]	0.007	0.011
8. No. 7, E 353	0.288[D]	0.012	0.021

[A] Calibration Standards: NBS 169, Ni-Base Alloy, *002S*; NBS 125a, *0.013S*; NBS 32e, 1.2 Ni-0.7 Cr, *0.021S*.

[B] Calibration Standards: NBS 32e, 1.2 Ni-0.7 Cr, *0.021S*; NBS 8i, Low-Alloy Steel, *0.064S*; NBS 10g, Low-Alloy Steel, *0.109S*.

[C] Calibration Standards: NBS 10g, Carbon Steel, *0.109S*, NBS 32e, 1.2 Ni-0.7 Cr, 0.021S + NBS 133a, 13 Cr-0.3 Mo, 0.326S: *0.174S*; NBS 129b, Low-Alloy Steel, *0.221S*.

[D] Calibration Standards: NBS 129b, Low-Alloy Steel, *0.221S*; NBS 129b, Low-Alloy Steel, 0.221S + NBS 133a, 13 Cr-0.3 Mo, 0.329S: *0.273S*; NBS 133a, 13 Cr-0.3 Mo, *0.329S*.

Figure 2. Reported data copied from ASTM E350-1995, which shows the measured amount of sulfur, the repeatability interval, and the reproducibility interval for eight test materials. The sulfur was measured using both an induction furnace and a resistance furnace for each sample, and the reported means and statistics were comparable for both methods.

E1601-1998 Standard Practice for Conducting an Interlaboratory Study to Evaluate the Performance of an Analytical Method

E1601-1998 defines equations for various statistical parameters, including the repeatability index, r, and reproducibility index, R. The repeatability index, r, can be computed using the within-laboratory standard deviation, s_r, which "reflects all variability that may occur from day-to-day within a laboratory," as follows:

$$rr = 2.8(ss_{rr}) \tag{3}$$

https://doi.org/10.52202/078572-0042

ASTM E415-2017 Spark Atomic Emission Spectrometry Carbon & Low-Alloy Steel[6]

The data shown in Figure 3, copied from ASTM E415-2017, presents the mean value, Xbar, the repeatability index, *r*, and the reproducibility index, *R*, for the measurement of carbon, manganese, and sulfur in 13 samples, as measured using spark atomic emission vacuum spectrometry. The repeatability indices presented in Figure 3 can be used to compute the within-laboratory standard deviation, s_r, for each sample-element combination using Eqn. (3).

Material	Number of Laboratories	Certified Value, %	Xbar	r	R	Bias
			Carbon			
Sample 1	7	0.211	0.2169	0.0073	0.0252	0.0059
Sample 2	7	0.142	0.1525	0.0084	0.0230	0.0105
Sample 3	8	0.13	0.1384	0.0072	0.0167	0.0084
Sample 4	8	0.658	0.6605	0.0075	0.0163	0.0025
Sample 5	8	0.483	0.4892	0.0092	0.0124	0.0062
Sample 6	8	0.457	0.4687	0.0110	0.0156	0.0117
Sample 7	8	0.332	0.3251	0.0202	0.0279	-0.0069
Sample 8	7	0.128	0.1305	0.0045	0.0076	0.0025
Sample 9	7	0.12	0.1196	0.0039	0.0142	-0.0004
Sample 10	8	1.03	1.024	0.0170	0.0227	-0.006
Sample 11	8	0.255	0.2530	0.0072	0.0137	-0.0020
Sample 12	8	0.107	0.1114	0.0040	0.0115	0.0044
Sample 13	7	0.376	0.3593	0.0280	0.0280	-0.0167
			Manganese			
Sample 1	7	0.316	0.3153	0.0033	0.0130	-0.0007
Sample 2	7	1.12	1.148	0.0200	0.0373	0.028
Sample 3	8	0.44	0.4549	0.0058	0.0161	0.0149
Sample 4	8	0.82	0.8319	0.0176	0.0336	0.0119
Sample 5	8	0.72	0.7330	0.0081	0.0265	0.0130
Sample 6	8	0.772	0.7825	0.0116	0.0298	0.0105
Sample 7	8	0.169	0.1713	0.0033	0.0091	0.0023
Sample 8	7	0.441	0.4437	0.0066	0.0168	0.0027
Sample 9	7	0.55	0.5584	0.0075	0.0226	0.0084
Sample 10	8	0.33	0.3340	0.0075	0.0182	0.0040
Sample 11	8	1.42	1.445	0.0132	0.0551	0.025
Sample 12	8	0.333	0.3374	0.0028	0.0131	0.0044
Sample 13	7	0.8	0.8070	0.0251	0.0402	0.0070
			Sulfur			
Sample 1	7	0.005	0.0460	0.0024	0.3116	0.0410
Sample 2	7	0.008	0.0076	0.0008	0.0031	-0.0004
Sample 3	8	0.015	0.0146	0.0010	0.0021	-0.0004
Sample 4	8	0.012	0.0135	0.0018	0.0044	0.0015
Sample 5	8	0.025	0.0232	0.0039	0.0064	-0.0018
Sample 6	8	0.0234	0.0221	0.0035	0.0054	-0.0013
Sample 7	8	0.033	0.0321	0.0038	0.0063	-0.0009
Sample 8	7	0.026	0.0241	0.0026	0.0056	-0.0019
Sample 9	7	0.003	0.0013	0.0005	0.0014	-0.0017
Sample 10	8	0.014	0.0144	0.0032	0.0046	0.0004
Sample 11	8	0.004	0.0046	0.0007	0.0007	0.0006
Sample 12	8	0.008	0.0076	0.0005	0.0023	-0.0004
Sample 13	7	0.047	0.0454	0.0082	0.0112	-0.0016

Figure 3. Reported data copied from ASTM E415-2017 that shows the mean value (Xbar), the repeatability index (r), and the reproducibility index (R) for the measurement of carbon, manganese, and sulfur in 13 samples, as measured using spark atomic emission vacuum spectrometry.

ASTM E691-2015 Interlab Study - Precision of Test Method[7]

ASTM E691-2015 defines the within-laboratory and between-laboratory variabilities. The within-laboratory variability is defined as the "cell standard deviation." The square root of the within-laboratory variance is the repeatability standard deviation, s_r.

[6] See also ASTM E415-2015 Spark Atomic Emission Spectrometry Carbon & Low-Alloy Steel.
[7] See also ASTM E691-2005 Interlab Study - Precision of Test Method.

https://doi.org/10.52202/078572-0042

ASTM E1019-2011 Carbon, Sulfur, Nitrogen & Oxygen by Combustion & Fusion

ASTM E1019-2011 reports carbon and sulfur content measured using combustion, along with repeatability intervals and reproducibility intervals for several test materials. The reported values are shown below in Figure 4, Figure 5, and Figure 6. Estimates for the within-laboratory standard deviations were back-calculated using the repeatability indices and Eqn. (1).

TABLE 1 Statistical Information—Carbon, Range I

Test Specimen	Carbon Found, %	Repeatability (R_1, Practice E173)	Reproducibility (R_2, Practice E173)
1. Electrolytic iron (NIST 365, 0.0068 C)	0.007	0.002	0.003
2. Bessemer carbon steel (NIST 8j, 0.081 C)	0.080	0.003	0.006
3. Type 304L stainless steel 18Cr-8Ni (NIST 101f, 0.014 C)	0.014	0.002	0.004
4. Type 446 stainless steel 26Cr (NIST 367, 0.093 C)	0.094	0.003	0.004
5. Nickel steel 36Ni (NIST 126b, 0.090 C)	0.092	0.003	0.004
6. Waspaloy 57Ni-20Cr-14Co-4Mo (NIST 349, 0.080 C)	0.078	0.003	0.004
7. Silicon steel (NIST 131a, 0.004 C)	0.004	0.002	0.002
8. High temperature alloy A286 26Ni-15Cr (NIST 348, 0.044 C)	0.046	0.003	0.004

Figure 4. Reported data copied from ASTM E1019-2011 that shows the measured carbon content, the repeatability interval, and the reproducibility interval for eight test specimens.

TABLE 9 Statistical Information, Sulfur, Range I (0.002 % to 0.010 % S)

Test Material	Certified Value, % Sulfur	Sulfur Found, %	Repeatability (R_1, Practice E173)	Reproducibility (R_2, Practice E173)
JK NR24	0.0010	0.00103	0.00025	0.00078
NIST 132b	0.0030	0.0027	0.00033	0.00094
High Temperature Alloy	...	0.0043	0.00051	0.00118
NIST 50c	0.0064	0.0065	0.0005	0.0024

TABLE 10 Statistical Information, Sulfur, Range II (0.010 % to 0.10 % S)

Test Material	Certified Value, % Sulfur	Sulfur Found, %	Repeatability (R_1, Practice E173)	Reproducibility (R_2, Practice E173)
NIST 890	0.015	0.0149	0.0015	0.0039
NIST 163	0.027	0.0264	0.0015	0.0092
NIST 73c	0.036	0.0356	0.0032	0.0078

TABLE 11 Statistical Information, Sulfur, Range III (0.10 % to 0.35 % S)

Test Material	Certified Value, % Sulfur	Sulfur Found, %	Repeatability (R_1, Practice E173)	Reproducibility (R_2, Practice E173)
NIST 6g	0.124	0.1200	0.0076	0.0239
NIST 129c	0.245	0.2451	0.0074	0.0243
NIST 133	0.356	0.3683	0.0174	0.0373

Figure 5. Reported data copied from ASTM E1019-2011 that shows the measured sulfur content, the repeatability interval, and the reproducibility interval for 10 test materials with sulfur content in three different ranges: 0.002 – 0.010%, 0.010 – 0.10%, and 0.1 – 0.35%.

https://doi.org/10.52202/078572-0042

TABLE 4 Statistical Information—Sulfur

Test Specimen	Sulfur Found, %	Repeatability (R_1, Practice E173)	Reproducibility (R_2, Practice E173)
Low alloy steel (JK 24, 0.0010 S)	0.0010	0.00045	0.00051
Stainless steel (NIST 348, 0.0020 S)	0.00198	0.0005	0.00064
Silicon steel (IRSID 114-1, 0.0037 S)	0.00322	0.00051	0.0007
Plain carbon steel (JSS 240-8, 0.0060 S)	0.00549	0.00055	0.00099
Stainless steel (JSS 652-7, 0.0064 S)	0.00615	0.00084	0.00087

Figure 6. Reported data copied from ASTM E1019-2011 that shows the measured sulfur content, the repeatability interval, and the reproducibility interval for 5 steel samples.

Review of Information from Manufacturers

Application Note for the CS744 Series LECO Combustion Analyzer

Bryan Labs, one of the vendors PG&E uses for destructive testing, communicated that they perform combustion analysis using the 744 Series LECO instrument. In the application note published by LECO for the CS744 Series LECO combustion analyzer, example analysis results using two different methods were reported. The results for the two methods used are shown in Figure 7. In the application, the two methods are described as follows:

> *Different methods can be used for the analysis of carbon and sulfur in ferroalloy materials on the CS744. Method 1 utilizes LECOCEL II and iron chip accelerators to facilitate combustion, without the use of hazardous materials. Method 2 utilizes iron powder, vanadium pentoxide, and LECOCEL as accelerators to facilitate combustion. This accelerator combination works well for ferroalloys and may improve sulfur recovery and precision. Even though the carbon blank for this method is considered high, the blank is consistent enough to be properly removed from the analysis results. Vanadium pentoxide is considered a hazardous material.*

Although the details of Method 1 and Method 2 presented in the application note have not been studied further for this study at present, the standard deviations reported for both methods are similar. Both methods' results were included in the carbon and sulfur uncertainty data that was analyzed.

https://doi.org/10.52202/078572-0042

*Typical Results for Method 1**

Sample	Mass (g)	% Carbon	% Sulfur
Ferro-Vanadium	0.2489	0.090	0.0301
	0.2495	0.089	0.0300
	0.2492	0.088	0.0289
	0.2481	0.089	0.0303
	0.2508	0.090	0.0297
	X=	***0.089***	***0.0298***
	s=	***0.001***	***0.0005***
NIST 58a	0.2488	0.015	<0.0008
Ferrosilicon	0.2475	0.016	<0.0008
0.0143% ±0.0050% C	0.2501	0.015	<0.0008
	0.2500	0.015	<0.0008
	0.2481	0.017	<0.0008
	X=	***0.016***	***<0.0008***
	s=	***0.001***	-
Euro 578-1	0.2486	0.018	0.0639
Ferro-Molybdenum	0.2506	0.017	0.0648
0.016% ±0.002% C	0.2495	0.018	0.0650
0.065% ±0.003% S	0.2483	0.018	0.0645
	0.2499	0.018	0.0647
	X=	***0.018***	***0.0646***
	s=	***0.001***	***0.0004***

*Typical Results for Method 2**

Sample	Mass (g)	% Carbon	% Sulfur
Ferro-Vanadium	0.2482	0.090	0.0315
	0.2487	0.089	0.0321
	0.2499	0.088	0.0315
	0.2486	0.086	0.0311
	0.2494	0.086	0.0315
	X=	***0.088***	***0.0315***
	s=	***0.002***	***0.0003***
NIST 58a	0.2487	0.016	<0.0008
Ferrosilicon	0.2485	0.016	<0.0008
0.0143% ±0.0050% C	0.2503	0.016	<0.0008
	0.2479	0.015	<0.0008
	0.2508	0.015	<0.0008
	X=	***0.016***	***<0.0008***
	s=	***0.001***	-
Euro 578-1	0.2500	0.017	0.0658
Ferro-Molybdenum	0.2488	0.018	0.0655
0.016% ±0.002% C	0.2493	0.018	0.0651
0.065% ±0.003% S	0.2536	0.019	0.0656
	0.2490	0.018	0.0656
	X=	***0.018***	***0.0655***
	s=	***0.001***	***0.0003***

Figure 7. Typical results for Methods 1 and 2 provided for three ferroalloy materials in LECO's application note for the CS744 Series LECO combustion analyzer.

https://doi.org/10.52202/078572-0042

Application Report for the Spectromaxx Lab Spark OES Analyzer

Bryan Labs communicated that they perform lab spark OES using the Spectromaxx instrument. The Spectromaxx application report from Spectro Analytical Instruments reported example results for the analysis of various alloys, which are presented in Figure 8.

Table 2: RN19/... Low-Alloy Steel

Element	C	Si	Mn	P	S
1	0.916	1.08	1.57	0.109	0.123
2	0.905	1.07	1.58	0.108	0.114
3	0.905	1.08	1.57	0.108	0.121
4	0.918	1.07	1.57	0.108	0.113
5	0.901	1.07	1.57	0.109	0.119
6	0.907	1.07	1.57	0.108	0.112
7	0.917	1.08	1.57	0.108	0.118
<x>	0.91	1.07	1.57	0.108	0.117
SD	0.0067	0.0026	0.0026	0.00057	0.0041
RSD (%rel.)	0.741	0.24	0.167	0.53	3.52

Table 3: SUS5/... Cast Iron

Element	C	Si	Mn	P	S
1	3.78	2.04	0.618	0.0445	0.0131
2	3.79	2.04	0.621	0.0447	0.0164
3	3.78	2.03	0.619	0.0447	0.0169
4	3.82	2.04	0.621	0.0448	0.0157
5	3.78	2.03	0.620	0.0458	0.0180
6	3.76	2.03	0.619	0.0438	0.0141
7	3.79	2.03	0.620	0.0451	0.0148
<x>	3.79	2.03	0.620	0.0448	0.0156
SD	0.0156	0.0039	0.0012	0.00063	0.0017
RSD (%rel.)	0.411	0.194	0.191	1.41	10.79

Table 4: E/... Cr-Cr/Ni Steel

Element	C	Si	Mn	P	S
1	0.0675	0.636	1.52	0.0026	<0.0005
2	0.0698	0.643	1.52	0.0029	<0.0005
3	0.0689	0.635	1.52	0.0029	<0.0005
4	0.0638	0.637	1.52	0.0031	<0.0005
5	0.0666	0.638	1.53	0.0030	<0.0005
6	0.0651	0.637	1.52	0.0030	<0.0005
7	0.0679	0.638	1.52	0.0033	<0.0005
<x>	0.0671	0.638	1.52	0.0030	<0.0005
SD	0.0021	0.0023	0.0031	0.0002	
RSD (%rel.)	3.15	0.368	0.206		

Table 5: RH18/... Highspeed Steel

Element	C	Si	Mn	P	S
1	1.27	0.348	0.259	0.0136	0.0011
2	1.27	0.349	0.259	0.0138	0.0012
3	1.30	0.350	0.259	0.0147	0.0012
4	1.27	0.349	0.259	0.0139	0.0010
5	1.29	0.350	0.260	0.0140	0.0011
6	1.29	0.353	0.261	0.0141	0.0013
7	1.27	0.352	0.260	0.0138	0.0014
<x>	1.28	0.350	0.259	0.0140	0.0012
SD	0.0118	0.0019	0.00075	0.00035	0.00014
RSD (%rel.)	0.923	0.541	0.29	2.47	11.58

Table 6: BAS493/... Manganese Steel

Element	C	Si	Mn	P	S
1	0.846	0.840	11.14	0.121	0.0102
2	0.848	0.849	11.12	0.123	0.0104
3	0.858	0.848	11.14	0.122	0.0103
4	0.855	0.849	11.19	0.124	0.0104
5	0.858	0.855	11.21	0.125	0.0105
6	0.847	0.846	11.18	0.122	0.0100
7	0.853	0.854	11.20	0.125	0.0105
<x>	0.852	0.849	11.17	0.123	0.0103
SD	0.0052	0.005	0.0323	0.0016	0.00018
RSD (%rel.)	0.606	0.585	0.289	1.26	1.79

Figure 8. Excerpts from the Spectromaxx application report showing example lab spark OES analysis results for various types of steel.

https://doi.org/10.52202/078572-0042

Review of Scientific and Engineering Literature

Carbon

Two academic papers were reviewed to obtain example uncertainty values for carbon concentration as measured using combustion analysis.

In Hemmerlin, et al. (2003),[8] which compared spark-OES analysis carbon measurements for steel samples prepared using either milling or grinding, it was reported that the sample with the ground surface had a measured mean of 31.7 μg g^{-1} with a standard deviation of 2.2 μg g^{-1}, and the sample with the milled surface has a measured mean of 26.8 μg g^{-1} with a standard deviation of 0.7 μg g^{-1}.

In Murray, et al. (1944),[9] the reproducibility of carbon measurements using low-pressure combustion was studied. Data were reported for two 3% silicon steel samples. Sample A was analyzed by two sets of operators. First, eight units were analyzed by three operators at a site referred to as "Pittsfield," and four units were analyzed by three operators at "Brackenridge." The means were measured to be 0.0031 and 0.0032 with standard deviations of 0.0004 and 0.0005, respectively. Sample B was measured in two parts named "Portion A" and "Portion B," with measured means 0.0044 and 0.0040 and standard deviations of 0.0004 and 0.0004, respectively.

Manganese

Two academic papers were reviewed to obtain example uncertainty values for manganese concentration as measured using lab spark OES. Zhang, et al. (2014) [10] compared the performances of laser-induced breakdown spectroscopy and spark OES, and they presented values for the coefficients of variation observed when measuring carbon, silicon, manganese, phosphorus, sulfur, chromium, and nickel. The coefficient of variation of manganese was reported to be 0.74% for a measured manganese concentration of 0.800 wt.%.

[8] Hemmerlin, M., L. Paulard, and G. Schotter. "Determination of ultra-low carbon and nitrogen contents in steel: combustion versus electrical spark source optical emission spectrometry for steelmaking process control." *Journal of Analytical Atomic Spectrometry* 18.3 (2003): 282-286.

[9] Murray, Jr, W. M., and SE Q. Ashley. "Determination of Carbon by Low-Pressure Combustion Method." *Industrial & Engineering Chemistry Analytical Edition* 16.4 (1944): 242-248.

[10] Zhang, Yong, et al. "Comparison of the analytical performances of laser-induced breakdown spectroscopy and spark-OES." ISIJ International 54.1 (2014): 136-140.

https://doi.org/10.52202/078572-0042

Table 3. Comparison of precision and accuracy for LIBS and Spark-OES.

Elements	Measured by Spark-OES 750 (mass%)	RSD (Spark-OES) (%)	Measured by LIBS (mass%)	RSD (LIBS) (%)
C	0.427	0.81	0.410	2.37
Si	1.13	0.93	1.12	2.18
Mn	0.800	0.74	0.802	2.23
P	0.021	3.1	0.017	5.12
S	0.020	5.2	0.024	9.34
Cr	1.07	1.05	1.08	2.13
Ni	0.414	0.63	0.421	1.92

Figure 9. Comparison of precision and accuracy for spark OES and LIBS. Copied from Zhang, et al. (2014).

Grünberger, et al. (2020)[11] presented standard deviations for manganese measured using both laser ablation-spark discharge-optical emission spectroscopy (LA-SD-OES) and laser-induced breakdown spectroscopy (LIBS). The data were presented graphically as shown in Figure 10. Since numerical data were not reported in this paper, the results shown in Figure 10 were not included in the analysis for manganese. However, they are shown here for reference.

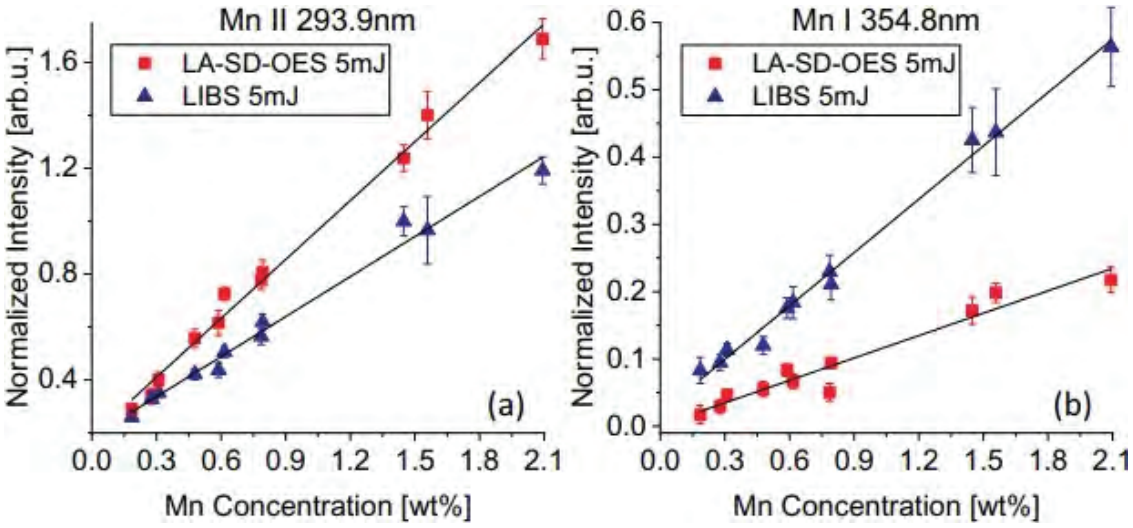

Figure 10. Excerpt from Figure 10 of Grünberger, et al. (2020), showing calibration curves for manganese in industrial steel samples measured with laser ablation-spark discharge-optical emission spectroscopy (LA-SD-OES) and laser-induced breakdown spectroscopy (LIBS).

[11] Grünberger, Stefan, et al. "Analysis of minor elements in steel and chemical imaging of micro- patterned polymer by laser ablation-spark discharge-optical emission spectroscopy and laser- induced breakdown spectroscopy." *Spectrochimica Acta Part B: Atomic Spectroscopy* 169 (2020): 105884.

https://doi.org/10.52202/078572-0042

Sulfur

In Takada, et al. (2000),[12] the following sulfur measurements and statistics were reported for three kinds of iron: A-Iron, JSS 001-4, and LECO 501-078.

Table 1. Results of the analysis of sulfur in high-purity irons using infrared absorption after combustion, reported in Takada, et al. (2000).

Sample	No. of Runs	Sulfur Content [μg/g]	Coefficient of Variation [%]
A-Iron	10	0.071 ± 0.038	5.4
JSS 001-4	15	1.91 ± 0.054	2.8
LECO 501-078	12	1.02 ± 0.057	5.6

Analysis of Values Obtained from Literature

Carbon

The carbon values obtained in the literature review are summarized below in Table 3, where green cells show values directly taken from the listed source, and cells not shaded green show values that have been calculated using the values in other columns. Both the standard deviation and the coefficient of variation were explored as possible metrics for quantifying the uncertainty.

[12] Takada, Kunio, et al. "Determination of trace amounts of sulfur in high-purity iron by infrared absorption after combustion: removal of sulfur blank." *Materials Transactions, JIM 41.1* (2000): 53- 56.

https://doi.org/10.52202/078572-0042

Table 2. Summary of carbon measurements and statistics obtained from literature. Cells shaded green indicate values that were directly provided by a source, and values in unshaded cells were calculated[13] using the value(s) in the cells shaded green.

Source Type	Source	Sample	Mean	Standard Deviation	Coefficient of Variation	R1, Repeatability Interval	R2, Reproducibility Interval	r, Repeatability Index	R, Reproducibility Index
ASTM	E1019-2011, Table 1	1. Electrolytic iron (NIST 365, 0.0068 C)	0.0070	0.0007	0.1010	0.0020	0.0030		
ASTM	E1019-2011, Table 1	2. Bessemer carbon steel (NIST 8j, 0.081 C)	0.0800	0.0011	0.0133	0.0030	0.0060		
ASTM	E1019-2011, Table 1	3. Type 304L stainless steel 18Cr-8Ni (NIST 101f, 0.014 C)	0.0140	0.0007	0.0505	0.0020	0.0040		
ASTM	E1019-2011, Table 1	4. Type 446 stainless steel 26Cr (NIST 367, 0.093 C)	0.0940	0.0011	0.0113	0.0030	0.0040		
ASTM	E1019-2011, Table 1	5. Nickel steel 36Ni (NIST 126b, 0.090 C)	0.0920	0.0011	0.0115	0.0030	0.0040		
ASTM	E1019-2011, Table 1	6. Waspaloy 57Ni-20Cr-14Co-4Mo (NIST 349, 0.080 C)	0.0780	0.0011	0.0136	0.0030	0.0040		
ASTM	E1019-2011, Table 1	7. Silicon steel (NIST 131a, 0.004 C)	0.0040	0.0007	0.1768	0.0020	0.0020		
ASTM	E1019-2011, Table 1	8. High temperature alloy A286 26Ni-15Cr (NIST 348, 0.044 C)	0.0460	0.0011	0.0231	0.0030	0.0040		

[13] The (within lab) standard deviation was calculated using the repeatability interval, RR_1, as shown in Eqn. 1 in this report, which was taken from ASTM E173-1993. The coefficient of variation was calculated as the ratio of the standard deviation to the mean, as described in ASTM E177-2014.

https://doi.org/10.52202/078572-0042

Source Type	Source	Sample	Mean	Standard Deviation	Coefficient of Variation	R1, Repeatability Interval	R2, Reproducibility Interval	r, Repeatability Index	R, Reproducibility Index
ASTM	E415-2017	Sample 1	0.2169	0.0026	0.0120			0.0073	0.0252
ASTM	E415-2017	Sample 2	0.1525	0.0030	0.0197			0.0084	0.0230
ASTM	E415-2017	Sample 3	0.1384	0.0026	0.0186			0.0072	0.0167
ASTM	E415-2017	Sample 4	0.6605	0.0027	0.0041			0.0075	0.0163
ASTM	E415-2017	Sample 5	0.4892	0.0033	0.0067			0.0092	0.0124
ASTM	E415-2017	Sample 6	0.4687	0.0039	0.0084			0.0110	0.0156
ASTM	E415-2017	Sample 7	0.3251	0.0072	0.0222			0.0202	0.0279
ASTM	E415-2017	Sample 8	0.1305	0.0016	0.0123			0.0045	0.0076
ASTM	E415-2017	Sample 9	0.1196	0.0014	0.0116			0.0039	0.0142
ASTM	E415-2017	Sample 10	1.0240	0.0061	0.0059			0.0170	0.0227
ASTM	E415-2017	Sample 11	0.2530	0.0026	0.0102			0.0072	0.0137
ASTM	E415-2017	Sample 12	0.1114	0.0014	0.0128			0.0040	0.0115
ASTM	E415-2017	Sample 13	0.3593	0.0100	0.0278			0.0280	0.0280
Paper	Hemmerlin	Milling	0.0027	0.00007	0.0261				

Source Type	Source	Sample	Mean	Standard Deviation	Coefficient of Variation	R1, Repeatability Interval	R2, Reproducibility Interval	r, Repeatability Index	R, Reproducibility Index
Paper	Hemmerlin	Grinding	0.0032	0.00022	0.0694				
Spec Sheet	LECO 744 Series Brochure	Ferro-Vanadium (Method 1)	0.0890	0.0010	0.0112				
Spec Sheet	LECO 744 Series Brochure	NIST 58a Ferrosilicon (Method 1)	0.0160	0.0010	0.0625				
Spec Sheet	LECO 744 Series Brochure	Euro 578-1 Ferro-Molybdenum (Method 1)	0.0180	0.0010	0.0556				
Spec Sheet	LECO 744 Series Brochure	Ferro-Vanadium (Method 2)	0.0880	0.0020	0.0227				
Spec Sheet	LECO 744 Series Brochure	NIST 58a Ferrosilicon (Method 2)	0.0160	0.0010	0.0625				
Spec Sheet	LECO 744 Series Brochure	Euro 578-1 Ferro-Molybdenum (Method 2)	0.0180	0.0010	0.0556				
Paper	Murray	Table 3, Pittsfield	0.0031	0.0004	0.1290				
Paper	Murray	Table 3, Brackenridge	0.0032	0.0005	0.1563				
Paper	Murray	Table 4, Top	0.0044	0.0004	0.0909				
Paper	Murray	Table 4, Bottom	0.0040	0.0004	0.1000				
Paper	Zhang[14]	Table 3 Sample	0.4270	0.0035	0.0081				

[14] Note that this carbon measurement was reported by Zhang to have been measured using lab spark OES. Although this technique is not used by PG&E to analyze carbon content, the statistics are included in the analysis.

https://doi.org/10.52202/078572-0042

https://doi.org/10.52202/078572-0042

The data shown in Table 3 is shown in Figure 11 as a scatterplot displaying standard deviation vs. measured mean carbon value. The plot on the left shows all the data in Table 3, and the plot on the right shows the data after limiting it to mean values that lie between the expected limit values for carbon in low carbon pipeline steels, i.e., between 0.02 wt.% and 0.4 wt.%.[15]

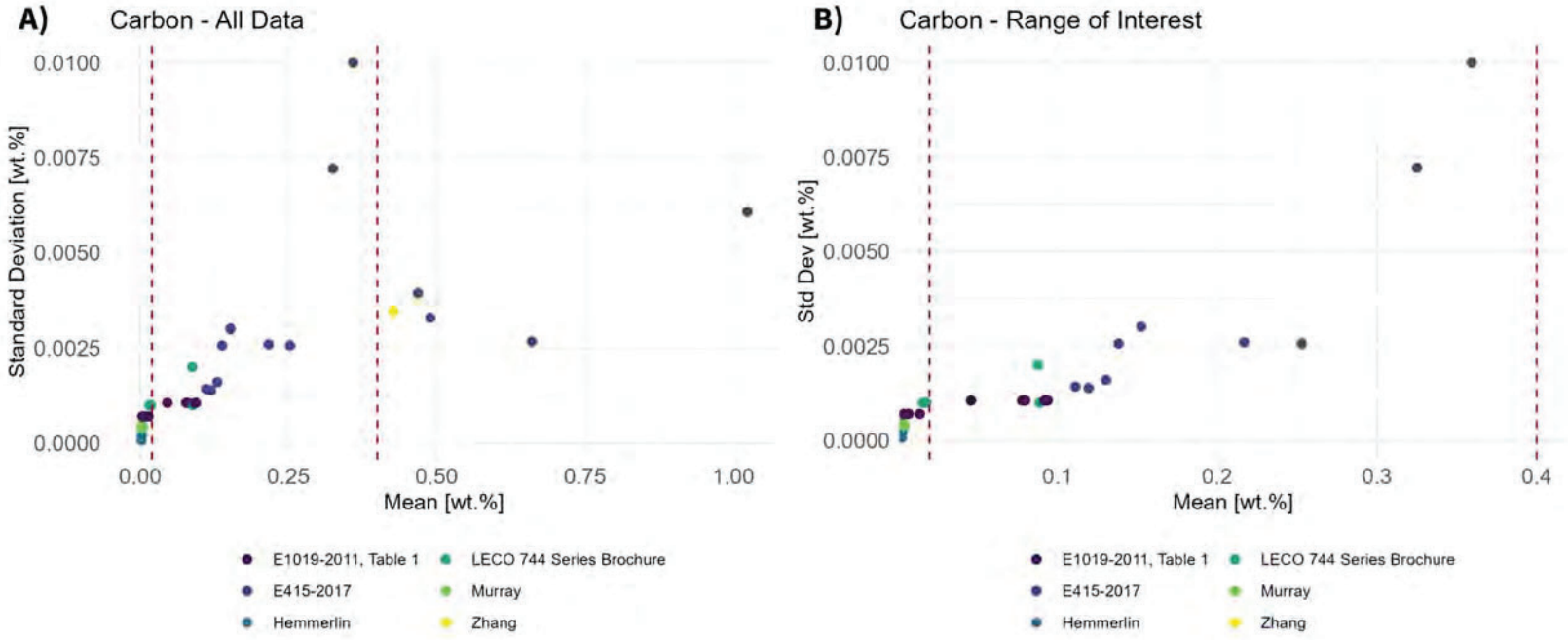

Figure 11. Scatterplots of standard deviation vs. mean value showing A) all data reported in Table 3, and B) data limited to mean carbon measurements between 0.02 wt.% and 0.4 wt.%.

Figure 12 below shows the same datapoints that are presented in Figure 11, but presented as the coefficient of variation vs. mean value. The coefficients of variation in the range of interest for carbon were more uniform in magnitude than for the standard deviations, which generally increased with increasing mean carbon content.

[15] These bounds were provided by RSI Pipeline Solutions LLC, one of PG&E's pipe grade prediction vendors, and they are based on an analysis of their internal database of low carbon pipeline steel composition measurements.

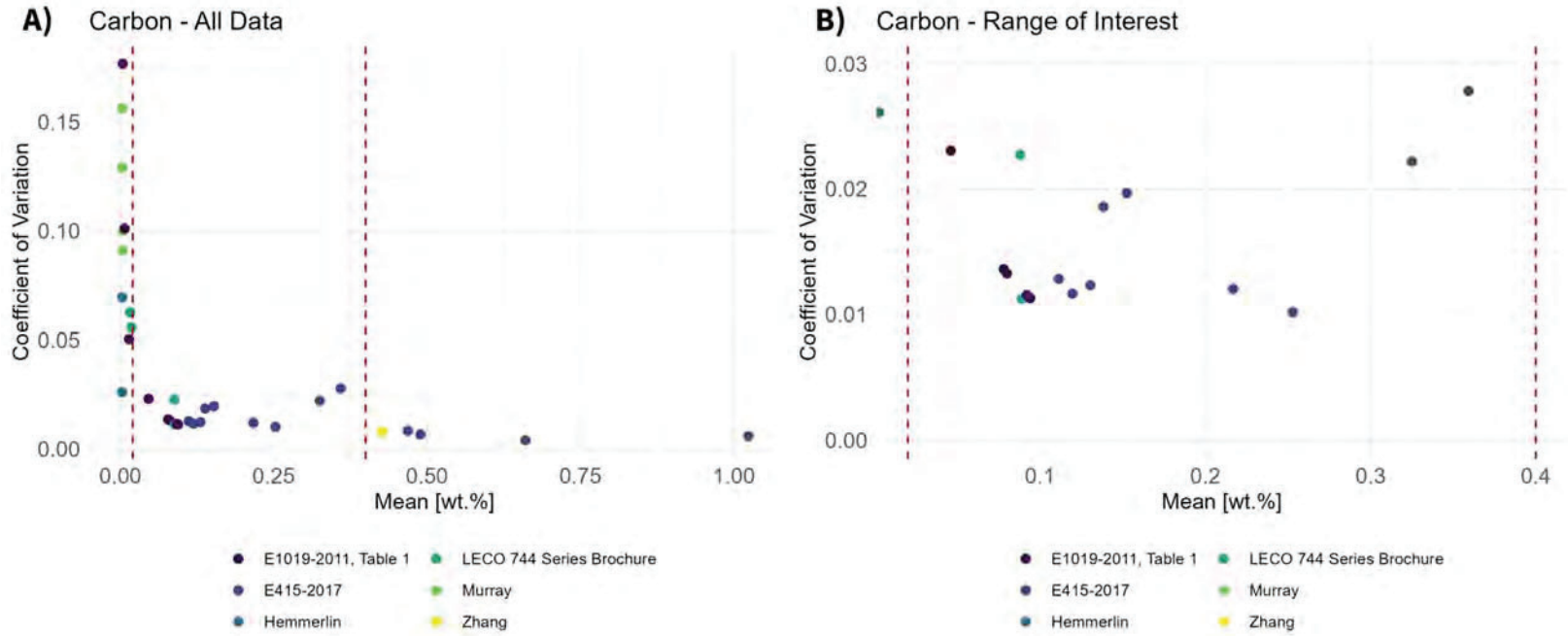

Figure 12. Scatterplots of coefficient of variation vs. mean value showing A) all data reported in Table 3, and B) data limited to mean carbon measurements between 0.02 wt.% and 0.4 wt.%.

https://doi.org/10.52202/078572-0042

https://doi.org/10.52202/078572-0042

Manganese

The manganese values obtained in the literature review are summarized below in Table 4, where green cells show values directly taken from the listed source, and cells not shaded green show values that have been calculated using the values in other columns. Both the standard deviation and the coefficient of variation were explored as possible metrics for quantifying the uncertainty.

Table 3. Summary of manganese measurements and statistics obtained from literature. Cells shaded green indicate values that were directly provided by a source, and values in unshaded cells were calculated16 using the value(s) in the cells shaded green.

Source Type	Source	Sample	Mean	Standard Deviation	Coefficient of Variation	R1, Repeatability Interval	R2, Reproducibility Interval	r, Repeatability Index	R, Reproducibility Index
ASTM	E350, Table 1	1. Alloy steel (BCS 252, 0.016 Mn)	0.0220	0.0014	0.0643	0.0040			
ASTM	E350, Table 1	2. Alloy steel (BCS 255/1 0.16 Mn)	0.1610	0.0014	0.0088	0.0040			
ASTM	E350, Table 1	3. Low-alloy steel (NBS 72f, 0.545 Mn)	0.5510	0.0035	0.0064	0.0100			
ASTM	E350, Table 1	4. Low-alloy steel (NBS 139a, 0.780 Mn)	0.7800	0.0032	0.0041	0.0090			
ASTM	E350, Table 1	5. Alloy steel (NBS, 159, 0.807 Mn)	0.8190	0.0035	0.0043	0.0100			
ASTM	E350, Table 1	6. Carbon steel (NBS 13f, 0.889 Mn)	0.8920	0.0053	0.0059	0.0150			
ASTM	E350, Table 1	7. Low-alloy steel (NBS 100b, 1.89 Mn)	1.9100	0.0071	0.0037	0.0200			

[16] The (within lab) standard deviation was calculated using the repeatability interval, RR_1, as shown in Eqn. 1 in this report, which was taken from ASTM E173-1993. The coefficient of variation was calculated as the ratio of the standard deviation to the mean, as described in ASTM E177-2014.

Source Type	Source	Sample	Mean	Standard Deviation	Coefficient of Variation	R1, Repeatability Interval	R2, Reproducibility Interval	r, Repeatability Index	R, Reproducibility Index
ASTM	E350, Table 14	BCS 255/1(0.16) - Peroxydisulfate-Arsenite Titrimetric Method	0.1620	0.0032	0.0196	0.0090			
ASTM	E350, Table 14	NBS 72F(0.595) - Peroxydisulfate-Arsenite Titrimetric Method	0.5430	0.0057	0.0104	0.0160			
ASTM	E350, Table 14	NBS 13f(0.889) - Peroxydisulfate-Arsenite Titrimetric Method	0.8900	0.0057	0.0064	0.0160			
ASTM	E350, Table 14	NBS 100b(1.89) - Peroxydisulfate-Arsenite Titrimetric Method	1.8900	0.0064	0.0034	0.0180			
ASTM	E415-2017	Sample 1	0.3153	0.0012	0.0037			0.0033	0.0130
ASTM	E415-2017	Sample 2	1.1480	0.0071	0.0062			0.0200	0.0373
ASTM	E415-2017	Sample 3	0.4549	0.0021	0.0046			0.0058	0.0161
ASTM	E415-2017	Sample 4	0.8319	0.0063	0.0076			0.0176	0.0336
ASTM	E415-2017	Sample 5	0.7330	0.0029	0.0039			0.0081	0.0265
ASTM	E415-2017	Sample 6	0.7825	0.0041	0.0053			0.0116	0.0298
ASTM	E415-2017	Sample 7	0.1713	0.0012	0.0069			0.0033	0.0091
ASTM	E415-2017	Sample 8	0.4437	0.0024	0.0053			0.0066	0.0168
ASTM	E415-2017	Sample 9	0.5584	0.0027	0.0048			0.0075	0.0226
ASTM	E415-2017	Sample 10	0.3340	0.0027	0.0080			0.0075	0.0182
ASTM	E415-2017	Sample 11	1.4450	0.0047	0.0033			0.0132	0.0551
ASTM	E415-2017	Sample 12	0.3374	0.0010	0.0030			0.0028	0.0131
ASTM	E415-2017	Sample 13	0.8070	0.0090	0.0111			0.0251	0.0402

https://doi.org/10.52202/078572-0042

https://doi.org/10.52202/078572-0042

Source Type	Source	Sample	Mean	Standard Deviation	Coefficient of Variation	R1, Repeatability Interval	R2, Reproducibility Interval	r, Repeatability Index	R, Reproducibility Index
Spec Sheet	Spectromaxx Brochure	Low-Alloy Steel	1.5700	0.0026	0.0017				
Spec Sheet	Spectromaxx Brochure	Cast Iron	0.6200	0.0012	0.0019				
Spec Sheet	Spectromaxx Brochure	Cr-Cr/Ni Steel	1.5200	0.0031	0.0020				
Spec Sheet	Spectromaxx Brochure	Highspeed Steel	0.2590	0.0008	0.0029				
Spec Sheet	Spectromaxx Brochure	Manganese Steel	11.1700	0.0323	0.0029				
Paper	Zhang	Table 3 Sample	0.8000	0.0059	0.0074				

The data shown in Table 4 is shown in Figure 13 as a scatterplot displaying standard deviation vs. measured mean manganese value. The plot on the left shows all the data in Table 4, and the plot on the right shows the data after limiting it to mean values that lie between the expected limit values for manganese in low carbon pipeline steels, i.e., between 0.23 wt.% and 1.8 wt.%.[17]

[17] These bounds were provided by RSI Pipeline Solutions LLC, one of PG&E's pipe grade prediction vendors, and they are based on an analysis of their internal database of low carbon pipeline steel composition measurements.

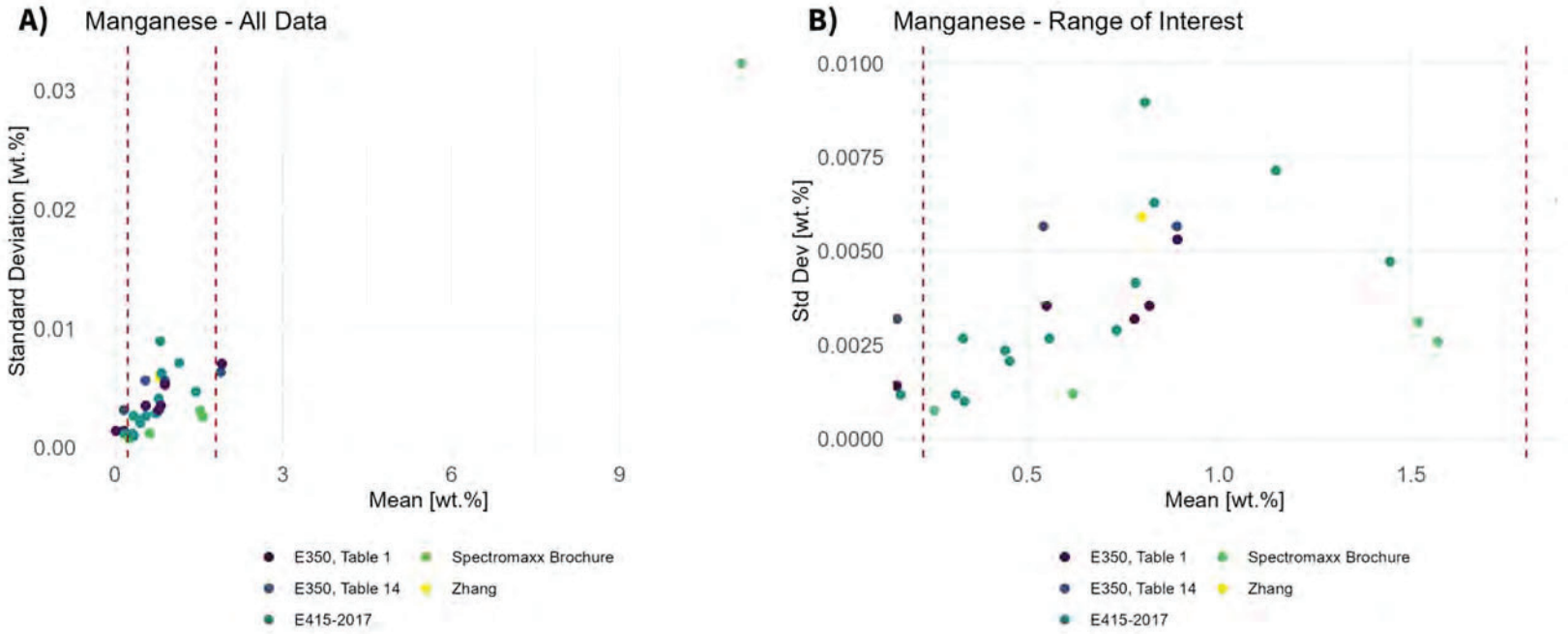

Figure 13. Scatterplots of standard deviation vs. mean value showing A) all data reported in Table 4, and B) data limited to mean manganese measurements between 0.23 wt.% and 1.8 wt.%.

Figure 14 below shows the same datapoints that are presented in Figure 13, but presented as the coefficient of variation vs. mean value. The magnitudes of both the standard deviations and coefficients of variation did not appear to be well-correlated with the mean manganese value in the range of interest for manganese.

https://doi.org/10.52202/078572-0042

https://doi.org/10.52202/078572-0042

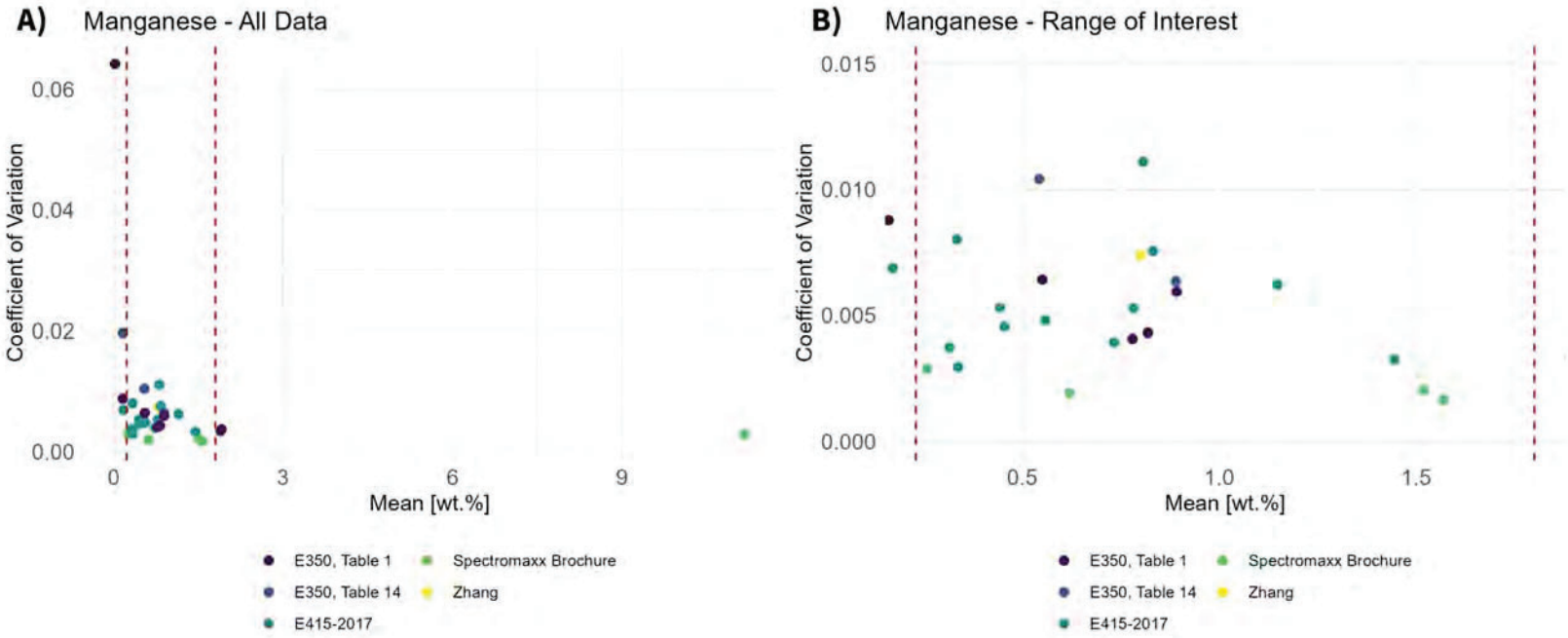

Figure 14. Scatterplots of the coefficient of variation vs. mean value showing A) all data reported in Table 4, and B) data limited to mean manganese measurements between 0.23 wt.% and 1.8 wt.%.

Sulfur

The sulfur values obtained in the literature review are summarized below in Table 5, where green cells show values directly taken from the listed source, and cells not shaded green show values that have been calculated using the values in other columns. Both the standard deviation and the coefficient of variation were explored as possible metrics for quantifying the uncertainty.

Table 4. Summary of sulfur measurements and statistics obtained from literature. Cells shaded green indicate values that were directly provided by a source, and values in unshaded cells were calculated[18] using the value(s) in the cells shaded green.

Source Type	Source	Sample	Mean	Standard Deviation	Coefficient of Variation	R1, Repeatability Interval	R2, Reproducibility Interval	r, Repeatability Index	R, Reproducibility Index
ASTM	E1019-2011, Table 10	NIST 890	0.0149	0.0005	0.0356	0.0015	0.0039		
ASTM	E1019-2011, Table 10	NIST 163	0.0264	0.0005	0.0201	0.0015	0.0092		
ASTM	E1019-2011, Table 10	NIST 73c	0.0356	0.0011	0.0318	0.0032	0.0078		
ASTM	E1019-2011, Table 11	NIST 6g	0.1200	0.0027	0.0224	0.0076	0.0239		
ASTM	E1019-2011, Table 11	NIST 129c	0.2451	0.0026	0.0107	0.0074	0.0243		
ASTM	E1019-2011, Table 11	NIST 133	0.3683	0.0062	0.0167	0.0174	0.0373		
ASTM	E1019-2011, Table 4	Low alloy steel (JK 24, 0.0010 S)	0.0010	0.0002	0.1591	0.0005	0.0005		
ASTM	E1019-2011, Table 4	Stainless steel (NIST 348, 0.0020 S)	0.0020	0.0002	0.0893	0.0005	0.0006		
ASTM	E1019-2011, Table 4	Silicon steel (IRSID 114-1, 0.0037 S)	0.0032	0.0002	0.0560	0.0005	0.0007		
ASTM	E1019-2011, Table 4	Plain carbon steel (JSS 240-8, 0.0060 S)	0.0055	0.0002	0.0354	0.0006	0.0010		
ASTM	E1019-2011, Table 4	Stainless steel (JSS 652-7, 0.0064 S)	0.0062	0.0003	0.0483	0.0008	0.0009		

[18] The (within lab) standard deviation was calculated using the repeatability interval, RR_1, as shown in Eqn. 1 in this report, which was taken from ASTM E173-1993. The coefficient of variation was calculated as the ratio of the standard deviation to the mean, as described in ASTM E177-2014.

https://doi.org/10.52202/078572-0042

Source Type	Source	Sample	Mean	Standard Deviation	Coefficient of Variation	r1, Repeatability Interval	R2, Reproducibility Interval	r, Repeatability Index	R, Reproducibility Index
ASTM	E1019-2011, Table 9	JK NR24	0.0010	0.0001	0.0858	0.0003	0.0008		
ASTM	E1019-2011, Table 9	NIST 132b	0.0027	0.0001	0.0432	0.0003	0.0009		
ASTM	E1019-2011, Table 9	High Temperature Alloy	0.0043	0.0002	0.0419	0.0005	0.0012		
ASTM	E1019-2011, Table 9	NIST 50c	0.0065	0.0002	0.0272	0.0005	0.0024		
ASTM	E350, Table 3	1. No. 1, E 352	0.0060	0.0007	0.1179	0.0020			
ASTM	E350, Table 3	2. No. 2, E 352	0.0080	0.0004	0.0442	0.0010			
ASTM	E350, Table 3	3. Low-alloy steel (NBS 11 lb, 0.015S)	0.0140	0.0011	0.0758	0.0030			
ASTM	E350, Table 3	4. Carbon steel (NBS 13f, 0.016S)	0.0160	0.0007	0.0442	0.0020			
ASTM	E350, Table 3	5. Carbon steel (NBS 152, 0.027S)	0.0260	0.0011	0.0408	0.0030			
ASTM	E350, Table 3	6. Carbon steel (NBS 16d, 0.033 S)	0.0320	0.0011	0.0331	0.0030			
ASTM	E350, Table 3	7. Carbon steel (NBS 129b + 8i (Mixed), 0.144 S)	0.1410	0.0025	0.0176	0.0070			
ASTM	E350, Table 3	8. No. 7, E 353	0.2860	0.0049	0.0173	0.0140			
ASTM	E350, Table 3	1. No. 1, E 352	0.0060	0.0004	0.0589	0.0010			
ASTM	E350, Table 3	2. No. 2, E 352	0.0090	0.0004	0.0393	0.0010			
ASTM	E350, Table 3	3. Low-alloy steel (NBS 11 lb, 0.015S)	0.0140	0.0004	0.0253	0.0010			
ASTM	E350, Table 3	4. Carbon steel (NBS 13f, 0.016S)	0.0150	0.0007	0.0471	0.0020			
ASTM	E350, Table 3	5. Carbon steel (NBS 152, 0.027S)	0.0270	0.0014	0.0524	0.0040			
ASTM	E350, Table 3	6. Carbon steel (NBS 16d, 0.033S)	0.0320	0.0011	0.0331	0.0030			
ASTM	E350, Table 3	7. Carbon steel (NBS 129b + 8i (Mixed), 0.144S)	0.1400	0.0025	0.0177	0.0070			

https://doi.org/10.52202/078572-0042

Source Type	Source	Sample	Mean	Standard Deviation	Coefficient of Variation	R1, Repeatability Interval	R2, Reproducibility Interval	r, Repeatability Index	R, Reproducibility Index
ASTM	E350, Table 3	8. No. 7, E 353	0.2880	0.0042	0.0147	0.0120			
ASTM	E415-2017	Sample 1	0.0460	0.0009	0.0186			0.0024	0.3116
ASTM	E415-2017	Sample 2	0.0076	0.0003	0.0376			0.0008	0.0031
ASTM	E415-2017	Sample 3	0.0146	0.0004	0.0245			0.0010	0.0021
ASTM	E415-2017	Sample 4	0.0135	0.0006	0.0476			0.0018	0.0044
ASTM	E415-2017	Sample 5	0.0232	0.0014	0.0600			0.0039	0.0064
ASTM	E415-2017	Sample 6	0.0221	0.0013	0.0566			0.0035	0.0054
ASTM	E415-2017	Sample 7	0.0321	0.0014	0.0423			0.0038	0.0063
ASTM	E415-2017	Sample 8	0.0241	0.0009	0.0385			0.0026	0.0056
ASTM	E415-2017	Sample 9	0.0013	0.0002	0.1374			0.0005	0.0014
ASTM	E415-2017	Sample 10	0.0144	0.0011	0.0794			0.0032	0.0046
ASTM	E415-2017	Sample 11	0.0046	0.0003	0.0543			0.0007	0.0007
ASTM	E415-2017	Sample 12	0.0076	0.0002	0.0235			0.0005	0.0023
ASTM	E415-2017	Sample 13	0.0454	0.0029	0.0645			0.0082	0.0112
Spec Sheet	LECO 744 Series Brochure	Ferro-Vanadium (Method 1)	0.0298	0.0005	0.0168				
Spec Sheet	LECO 744 Series Brochure	NIST 58a Ferrosilicon (Method 1)	< 0.0008						
Spec Sheet	LECO 744 Series Brochure	Euro 578-1 Ferro-Molybdenum (Method 1)	0.0646	0.0004	0.0062				
Spec Sheet	LECO 744 Series Brochure	Ferro-Vanadium (Method 2)	0.0315	0.0003	0.0095				
Spec Sheet	LECO 744 Series Brochure	NIST 58a Ferrosilicon (Method 2)	< 0.0008						

https://doi.org/10.52202/078572-0042

https://doi.org/10.52202/078572-0042

Source Type	Source	Sample	Mean	Standard Deviation	Coefficient of Variation	R1, Repeatability Interval	R2, Reproducibility Interval	r, Repeatability Index	R, Reproducibility Index
Spec Sheet	LECO 744 Series Brochure	Euro 578-1 Ferro-Molybdenum (Method 2)	0.0655	0.0003	0.0046				
Paper	Takada	A-Iron	0.000071	0.0000	0.0540				
Paper	Takada	JSS 001-4	0.000191	0.0000	0.0280				
Paper	Takada	LECO 501-078	0.000102	0.0000	0.0560				
Paper	Zhang	Table 3 Sample	0.0200	0.0010	0.0520				

The data shown in Table 5 is shown in Figure 15 as a scatterplot displaying standard deviation vs. measured mean sulfur value. The plot on the left shows all the data in Table 5, and the plot on the right shows the data after limiting it to mean values that lie between the expected limit values for sulfur in low carbon pipeline steels, i.e., between 0 wt.% and 0.1 wt.%.[19]

[19] These bounds were provided by RSI Pipeline Solutions LLC, one of PG&E's pipe grade prediction vendors, and they are based on an analysis of their internal database of low carbon pipeline steel composition measurements.

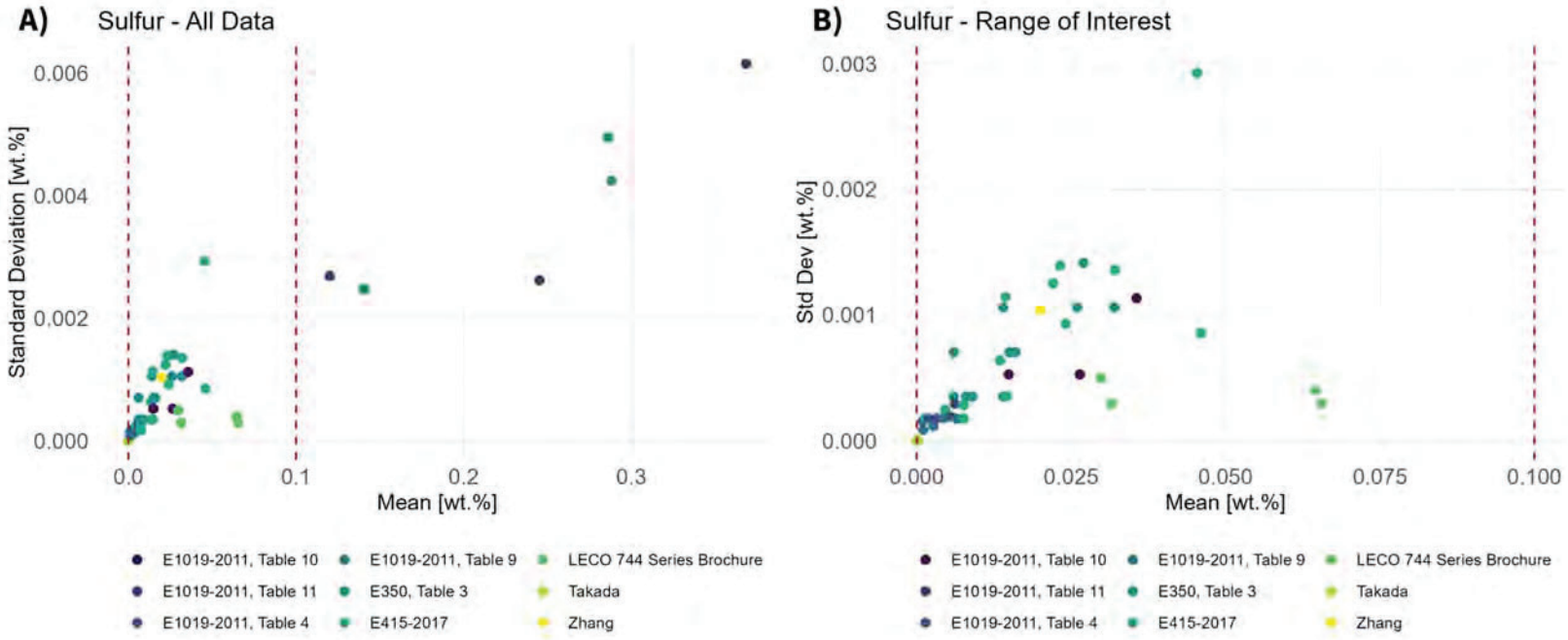

Figure 15. Scatterplots of standard deviation vs. mean value showing A) all data reported in Table 5, and B) data limited to mean sulfur measurements between 0 wt.% and 0.1 wt.%.

Figure 16 below shows the same datapoints that are presented in Figure 15, but presented as the coefficient of variation vs. mean value. In the range of interest for sulfur, the standard deviation appears to be roughly correlated with the mean, but the coefficient of variation appears to be anticorrelated with the mean. The observed anticorrelation is due to sulfur usually being present in very small amounts, since it is not intentionally added to the low carbon pipeline steel during manufacturing.

https://doi.org/10.52202/078572-0042

https://doi.org/10.52202/078572-0042

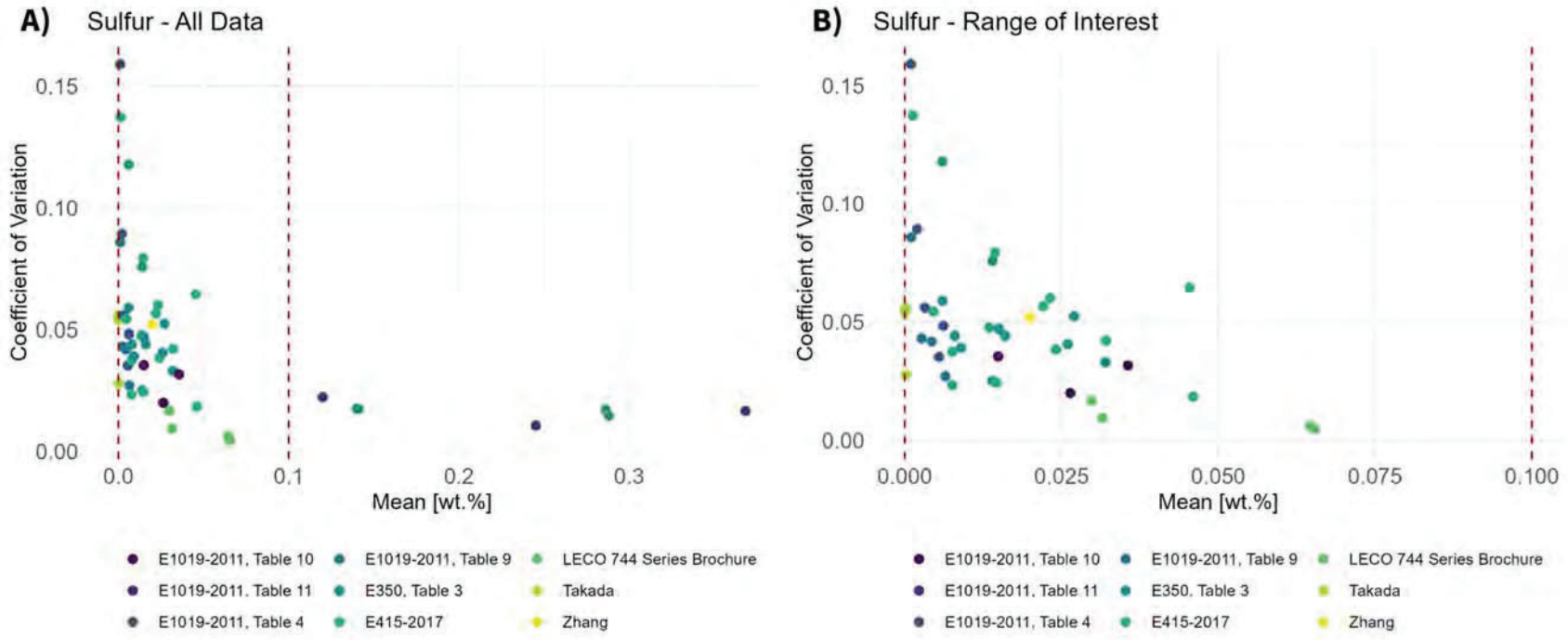

Figure 16. Scatterplots of coefficient of variation vs. mean value showing A) all data reported in Table 5, and B) data limited to mean sulfur measurements between 0 wt.% and 0.1 wt.%.

Discussion

The goal of the literature search was to obtain representative uncertainty values for the destructive testing of carbon, manganese, and sulfur in pipeline steels. For each element, ASTM standards, instrument application notes, and academic literature were reviewed to obtain example uncertainty values for a range of mean values. The assembled data were presented as both standard deviation values vs. mean values and coefficient of variation values vs. mean values for each element to determine which metric was more appropriate for quantifying the uncertainty generally.

Although the standard deviation was usually somewhat correlated with the mean value, the coefficient of variation was in some cases more anticorrelated with the mean, particularly for sulfur. Because sulfur is not an intentional additive to low carbon pipeline steel, it can be present in very small amounts, which means that small absolute standard deviations, σ, can result in overly large coefficients of variations, CV.[20] Therefore, the decision was made to not select the coefficient of variation as the metric for defining the uncertainty value. Instead, representative and reasonably conservative values for standard deviation were chosen for each element. The selected uncertainty values are summarized below in Table 6. For all three elements, the number of degrees of freedom for the t-distribution to be generated was conservatively selected to be 1, which corresponds to 2 sample measurements, i.e., the minimum number of measurements for which a standard deviation can be calculated.

Table 5. Summary table showing the recommended uncertainties for each element and the rationale for selecting the chosen estimated uncertainty values.

Element	Recommended Uncertainty (wt%)	Notes
C	0.003	This standard deviation was selected because it was the maximum standard deviation observed in the carbon range of interest after excluding two potential outlier values that were much larger than the others and corresponded to carbon mean values over 0.3 wt.%, which are rarely observed for low carbon pipeline steels.
Mn	0.009	This standard deviation was selected because it was the maximum standard deviation observed in the manganese range of interest.
S	0.0014	This standard deviation was selected because it was the maximum standard deviation observed in the sulfur range of interest after excluding a single measurement that was much larger than the other standard deviations (specifically, it was potentially an outlier that was over twice as large as the selected second largest standard deviation).

[20] The coefficient of variation, CV, is calculated using the standard deviation, σ, and the mean, μ, as $CV = \frac{\sigma}{\mu}$.

https://doi.org/10.52202/078572-0042

Literature Review and Database Analysis – Strength

Literature Review

Relevant literature was reviewed to obtain uncertainty values for YS and UTS from tensile testing.[21] Literature to be discussed in detail below includes:

- API standards
    - API RP 1176 Recommended Practice for Assessment and Management of Cracking in Pipelines
- ASTM standards
    - ASTM E-8-2015a Tension Testing of Metallic Materials
- ISO standards
    - ISO 6892-1 Tensile Testing at Room Temperature
- Engineering and scientific literature
    - A Historical Review and Analysis of the Effect of Tensile Test Sample Orientation on Pipeline Yield Strength (PRCI REX2023)

The following literature was also reviewed:

- ASTM standards:
    - ASTM E4-2015 Force Verification of Testing Machines
    - ASTM E74-2013a Calibration of Force-Measuring Instruments
    - ASTM E83-2010a Extensometer Verification and Classification
    - ASTM E177-2014 Precision and Bias in ASTM Test Methods

While these studies provide good background information on tensile testing, they do not provide sufficient information relevant for estimating tensile test uncertainty and therefore will not be discussed in detail below.

API RP 1176 Recommended Practice for Assessment and Management of Cracking in Pipelines (2016)

API RP 1176 provides YS and UTS data based on several hundred tests on pipeline samples covering broad ranges of grades, vintages, and sources. These data are shown in Table 7 below. From these data the highest standard deviation for YS is for X56 at 8.7 ksi, and the highest standard deviation for UTS is for X52 at 8.7 ksi.

It is important to note that these tensile test results (YS and UTS) are reported based on pipeline grade and not on a per pipe joint basis. It is recognized that there can be wide distributions of strength properties within a grade category whereas the strength distribution for a single pipe joint will likely be much narrower. Therefore, these uncertainty values could be viewed as an upper bound for YS and UTS variability for specific pipe joints.

[21] The transition between elastic and plastic behavior is called the proportional limit. Because this transition point is difficult to measure, and can be dependent on the precision of the testing equipment used, YS is used as a surrogate to mark the stress at which the material starts plastic deformation. API 5L references the 0.5% extension under load (EUL) definition of YS, which is the portion of the stress-strain curve corresponding to a strain of 0.005.

https://doi.org/10.52202/078572-0042

Additionally, API RP 1176 mentions that an "operator may also use sample testing to confirm properties; however, a limited sample size might not be adequate to establish the full range of values."[22]

Table 6. Database YS and UTS Properties by Grade from API RP 1176 Annex D.[23]

	YS (ksi)		UTS (ksi)	
Grade	Mean	STD	Mean	STD
A/Bsmr/OH	40.0	4.5	55.7	5.3
B	48.6	7.8	68.5	7.1
X42	52.2	6.4	70.5	5.1
X46	43.7	5.7	73.1	6.3
X52	59.2	6.0	78.7	8.7
X56	62.8	8.7	85.1	7.3
X60	68.7	5.4	86.6	6.6
X65	72.0	2.9	89.4	5.9
X70	80.4	5.0	91.1	5.8

ASTM E-8-2015a Tension Testing of Metallic Materials

ASTM E-8 provides YS and UTS data for an interlaboratory study where six different labs tested six samples each for six different bulk materials. This study sought to investigate the standard deviations seen within a lab compared to standard deviations seen between-labs for these six bulk materials. These data are shown in Table 8 below. From these data, the interlaboratory difference for YS is between approximately 0.5 to 2.8 ksi, while the interlaboratory difference for UTS is between approximately 0.6 and 1.3 ksi.

It is important to note that these data are for bulk materials and not specifically pipeline steel samples. It is expected that these bulk materials will likely have less material inhomogeneity than pipe joints, and therefore should have lower standard deviations for YS and UTS. These uncertainty values could be viewed as a lower bound for YS and UTS variability for pipeline materials. While none of the presented results are from pipeline steels, those shown in ASTM A105 could represent a reasonable minimum expected standard deviation for pipeline steels.

Additionally, ASTM E-8 recommends a suite of tests be run to determine typical lab uncertainties because tensile results are "easily affected by operators, measuring devices, and test equipment."[24]

[22] API RP 1176 Recommended Practice for Assessment and Management of Cracking in Pipelines (2016), Annex D, p. 93.

[23] API RP 1176 Annex D reports these values in MPa. The MPa values have been converted to ksi for ease of comparison between standards and the PG&E ECA2 and Met database analyses.

[24] ASTM E8/E8m – 15a, Standard Test Methods for Tension Testing of Metallic Materials, X3.7.4, p. 26.

https://doi.org/10.52202/078572-0042

Table 7. Interlaboratory study YS and UTS values from ASTM E8-2015a for bulk materials.

	YS (0.2%) (ksi)			UTS (ksi)		
Material	Mean	Within-lab STD	Between-lab STD	Mean	Within-lab STD	Between-lab STD
EC-H19	22.98	0.47	0.48	25.66	0.63	0.63
2024-T351	52.64	0.74	0.79	71.26	0.88	0.96
ASTM A105	58.36	0.83	1.44	86.57	0.60	1.27
AISI 316	69.78	0.95	2.83	100.75	0.39	1.22
Inconel 600	38.91	0.36	0.85	99.48	0.42	0.72
SAE 51410	140.33	1.29	2.30	181.73	0.46	1.14

ISO 6892-1 Tensile Testing at Room Temperature (2019)

ISO 6892-1 provides the mean and reproducibility for YS and UTS data for an interlaboratory study for twelve different bulk steel materials. The reproducibility values were converted to standard deviations using the equation mentioned in ISO6892-1 Annex L. These data are shown below in Table 9.

Similar to those materials presented in ASTM E-8, the bulk materials presented here would be expected to have less material inhomogeneity than pipe joints. While none of the presented results are from pipeline steels, similar to ASTM E-8, the results for the low carbon plate sample in ISO 6892-1 could represent a reasonable lower bound for anticipated strength variability for pipeline steels.

Table 8. Interlaboratory study YS and UTS values from ISO 6892-1 Annex L. Standard deviations calculated using formula with reproducibility and mean value described in Annex L.[25]

		YS (ksi)		UTS (ksi)	
Material	Code	Mean	STD	Mean	STD
Sheet	DX56	23.5	0.5	43.7	1.1
Low C plate	HR3	33.2	1.4	48.6	1.2
Sheet	ZStE 180	38.7	1.9	45.7	1.1
AISI 105	P245GH	53.3	1.3	80.1	0.8
Plate	C22	58.4	1.4	86.6	1.2
Austenitic SS	S355	62.0	1.9	81.9	1.0
Austenitic SS	SS316L	33.5	1.2	82.5	1.7
Austenitic SS	X2CrNi18-10	44.1	1.4	86.1	1.3
AISI 316	X2CrNiMo18-10	51.2	2.0	90.3	1.4
Martensitic SS	X5CrNiMo17-12-2	69.6	2.8	100.7	1.2
High Strength	X12Cr13	140.3	2.2	181.7	1.2
	30NiCrMo16	150.8	1.5	169.4	1.3

[25] ISO 6892-1 Annex L reports YS and UTS values in MPa. The MPa values have been converted to ksi for ease of comparison between standards and the PG&E ECA2 and Met database analyses. The reproducibility values were converted to standard deviations using the equation mentioned in ISO6892-1 Annex L.

https://doi.org/10.52202/078572-0042

A Historical Review and Analysis of the Effect of Tensile Test Sample Orientation on Pipeline Yield Strength (PRCI REX2023)

This PRCI REX2023 paper compares tensile results between different tensile sample types to evaluate the effect that sample type has on YS. The tensile sample types compared are transverse flat strap, longitudinal flat strap, and transverse round bar. In order to compare sample type, different sample types were taken from the same unique pipe joint.

The data comparing transverse and longitudinal flat strap are shown in Figure 17 below. The flat strap transverse and longitudinal samples have similar YS results. It is possible that transverse tests result in a slightly higher YS overall compared to longitudinal tests; however, the PG&E data appear to show comparable results between transverse and longitudinal samples.

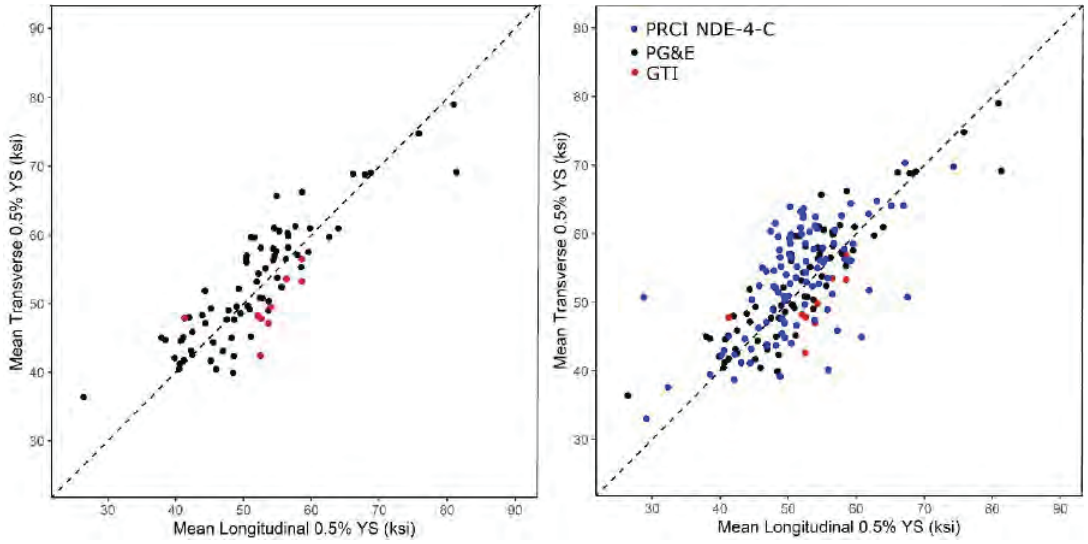

Figure 17. Transverse vs Longitudinal Strap originally published in PRCI REX 2023 proceedings.[26]

The data comparing transverse flat strap and round bar specimens are shown in Figure 18 below. The transverse flat strap and round bar samples have fairly similar YS results; however, the round bar specimens appear to exhibit a slight trend for a higher YS compared to flat strap specimens.

[26] E. Brady, M. Gould, J. Kornuta, N. Switzner, P. Veloo, "A Historical Review and Analysis of the Effect of Tensile Test Sample Orientation on Pipeline Yield Strength." REX 2023 PRCI Research Exchange, March 2023.

https://doi.org/10.52202/078572-0042

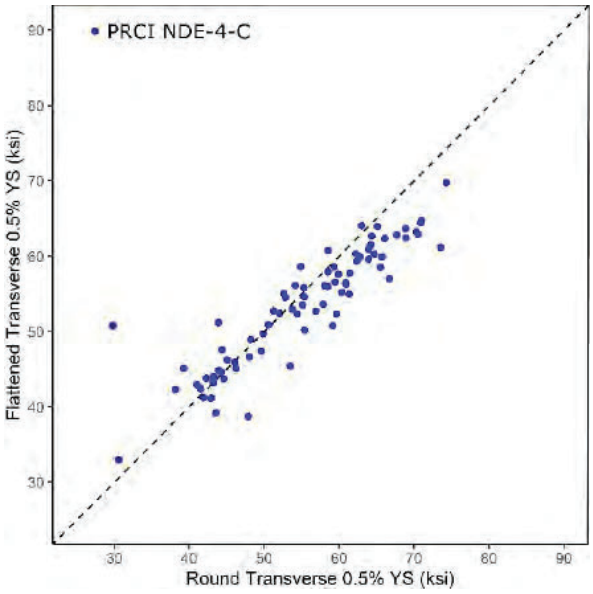

Figure 18. Flattened Transverse Strap vs Round Transverse Bar originally published in PRCI REX 2023 proceedings.[27]

Since these tensile sample types reviewed above have similar YS results for the same pipe joint, grouping these tensile sample types for the same pipe joint to calculate mean YS appears to be a reasonable practice. This observation is of importance as in the following section, the tensile test sample types (transverse and longitudinal, and flat strap and round bar) are grouped together for features in the PG&E databases to calculate mean and standard deviations for YS and UTS.

Database Analysis

PG&E maintains databases (such as the ECA2 and Met databases) that store the results of non-destructive and destructive testing for a number of unique pipe joints. The uncertainty estimates for destructive YS and UTS measurements presented in this section were determined by an analysis of the tensile testing results in the ECA2 and Met databases. The ECA2 database is a relatively newer database whose destructive tensile testing data were collected using a more controlled procedure. This more rigorous and detailed tensile testing procedure[28] would be expected to lead to more consistent results since each experimental variable would be more closely controlled.[29] In contrast, the Met database contains a significant amount of relatively older data collected by PG&E prior to this tensile testing procedure being implemented. Therefore, the Met database might be expected to exhibit greater uncertainty in the destructive tensile results, and it is necessary to analyze these databases separately.

[27] E. Brady, M. Gould, J. Kornuta, N. Switzner, P. Veloo, "A Historical Review and Analysis of the Effect of Tensile Test Sample Orientation on Pipeline Yield Strength." REX 2023 PRCI Research Exchange, March 2023.

[28] ECA2: Tensile Testing Procedure, effective date: 11/05/2020, Rev: 0.0 (DRAFT)

[29] Gould, M., Amend, B., Johnson, D., Rutherford, L., and Rapp, S., "Tensile Testing of Flattened Strap Pipe Specimens," Proceedings of the 28th Pipeline Pigging and Integrity Management (PPIM) Conference, Houston, TX, February 2016.

https://doi.org/10.52202/078572-0042

Both databases were analyzed to determine the mean and median standard deviation values for YS and UTS on a per-joint basis where multiple tensile tests were performed. To determine standard deviations for each unique pipe joint, the database was filtered to exclude: unique pipe joints with fewer than 3 tensile tests, calibration blocks, and weld samples. There was no filtering on type of tensile sample (i.e., round bar samples and strap samples were grouped together for each unique joint). After the above filters were applied, there were 113 unique pipe joints in the Met database and 101 in the ECA2 database. The distribution of standard deviation for the unique pipe joints for both databases are shown in Figure 19 while the mean and median standard deviations are shown in Table 10 below. From these results, the mean standard deviation values calculated from the Met database were observed to be more conservative.

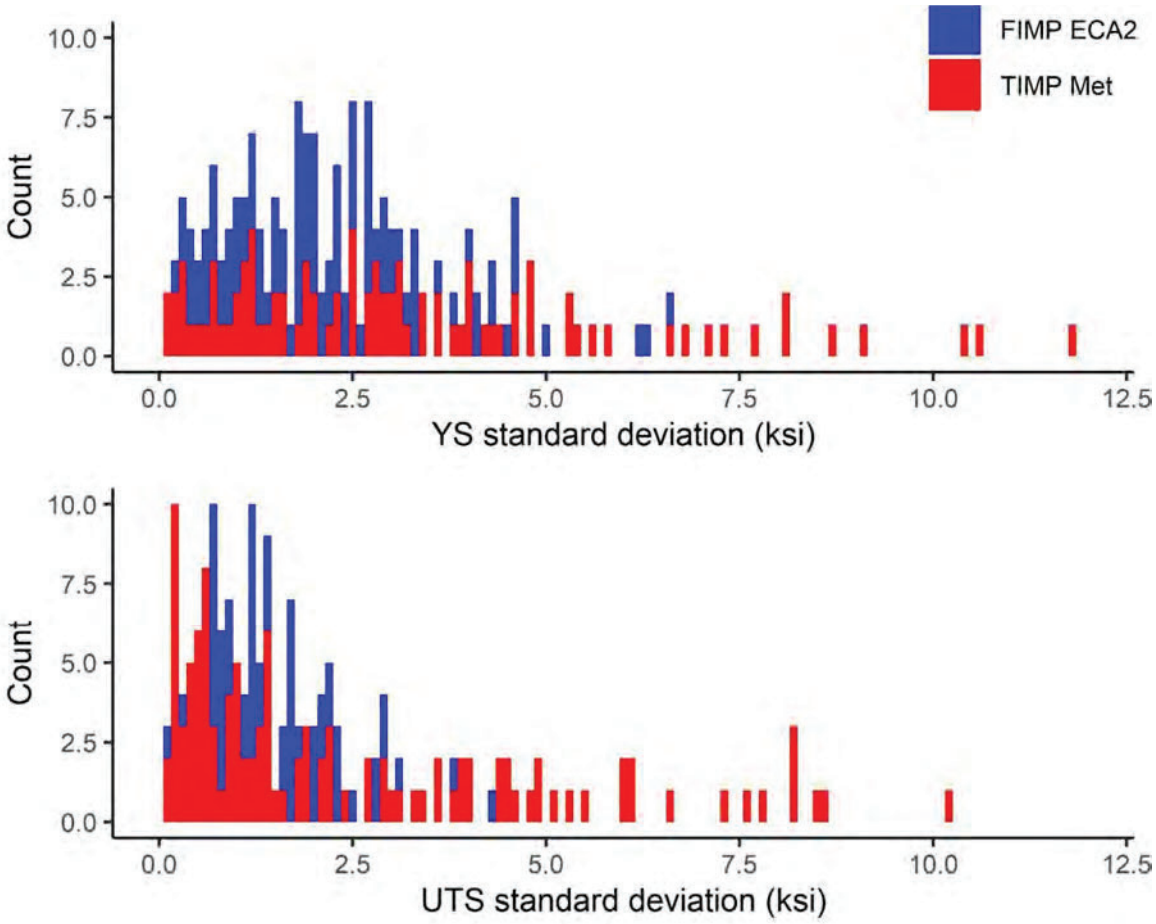

Figure 19. The distribution of calculated standard deviations for unique features in the ECA2 and Met databases.

https://doi.org/10.52202/078572-0042

Table 9. YS and UTS standard deviation values calculated for destructive measurements in the ECA2 and Met databases.

	YS standard deviation (ksi)		UTS standard deviation (ksi)	
Database	Mean	Median	Mean	Median
PG&E ECA2	2.22	2.03	1.17	0.96
PG&E Met	3.24	2.74	2.44	1.36

Discussion

Standard deviation values for YS and UTS from the relevant literature, PG&E ECA2 and Met database analyses, and the NDT strength technique currently employed by PG&E (instrumented indentation testing, IIT) are shown in Table 11 below.

The relevant literature provides reasonable lower and upper bounds to expect for tensile test uncertainty for pipe joints. The literature for bulk materials (ASTM E8-2015a and ISO 6892-1) have the smallest standard deviations for YS and UTS for materials most comparable to pipeline steels from the materials presented. It is expected that these bulk materials would have a lower standard deviation compared to pipeline materials, as they would likely have less material inhomogeneity. API RP 1176, which provides standard deviations for YS and UTS on pipeline steels by API grade category, has the largest standard deviations out of the data presented below. This result is expected since the measured variability in strength among a relatively large population of pipe joints based on grade would likely be larger than the exhibited strength variability among replicate tensile test specimens from a single pipe joint.

The mean YS and UTS uncertainties calculated from the PG&E ECA2 and Met databases lie within the lower and upper bounds from relevant literature. As a point of comparison, these values are also below the reported tensile prediction uncertainty for IIT. Hence, it is recommended to use the mean standard deviation values calculated from PG&E's Met database as assumed tensile uncertainties as these are more conservative than the values calculated from the ECA2 database. These estimates are also higher than the uncertainties for bulk material and lower than the uncertainties from API RP 1176 and for IIT. Moreover, these uncertainty values for YS and UTS align with expectations for where the destructive uncertainties for pipeline materials for consistent tensile testing would likely lie. For both YS and UTS, the number of degrees of freedom for the t-distribution to be generated was selected to be 3, which corresponds to 4 sample measurements, which was the median number of replicate measurements in the PG&E Met database.

https://doi.org/10.52202/078572-0042

Table 10. Summary of uncertainty values for YS and UTS from literature, PG&E database analyses, and IIT (for reference).

Source	YS standard deviation (ksi)	UTS standard deviation (ksi)	Number of Replicate Measurements
ASTM E8-2015a (ASTM A105) with-in lab[1]	0.83	0.60	6
ASTM E8-2015a (ASTM A105) between-labs[1]	1.44	1.27	6
ISO 6892-1 (Low carbon, plate)[2]	1.4	1.2	—
PG&E ECA2 (mean)	2.22	1.17	4[3]
PG&E Met (mean)	3.24	2.44	4[3]
IIT Tensile Prediction[4] (approximate)	6	4	85
API RP 1176	8.7[5]	8.7[6]	100's[7]

[1] Bulk material.

[2] Bulk material from interlaboratory testing.

[3] Is the median number of tensile samples per pipe joint for the database analyzed. The average number of tensile samples per pipe joint was 4.8 and 5.5 for the ECA2 database and Met database, respectively.

[4] Tompkins Hill Meter & Regulator Station NDT Strength Results and Analysis, March 22, 2024.

[5] This is the largest standard deviation which corresponds to X56. The average standard deviation for Grades A to X70 is 5.8 ksi, with the lowest being 2.9 ksi.

[6] This is the largest standard deviation which corresponds to X52. The average standard deviation for Grades A to X70 is 6.5 ksi, with the lowest being 5.1 ksi.

[7] From API RP 1176, "These values are based on several hundred tests performed on material samples covering a broad range of grades, vintages, and sources." Based on this text, it is unclear how many samples were used for each API pipe grade presented in API RP 1176.

https://doi.org/10.52202/078572-0042

Validation of Proposed Uncertainties

PG&E provided Exponent with destructively obtained chemical composition and strength values for 633 pipe joints with known grades. This dataset spans across grades LTB[30] to X70, OD values from 2.375 to 42 inches, NWT values from 0.141 to 0.812 and known vintages from 1928 to 2016.[31] In order to evaluate the proposed assumed uncertainties, three different grade prediction analyses were run on this dataset using the chemical composition and strength (CDS) grade prediction model developed by RSI Pipeline Solutions, LLC (RSI), one of PG&E's pipe grade prediction vendors. The three analyses are listed below.

1. No uncertainties for chemical composition or strength (633 pipe samples). The destructive measurements for each of the 633 pipe joints were used to generate one predicted grade per pipe joint as a control.
2. Proposed assumed uncertainties for both chemical composition and strength (633 pipe samples). The uncertainties presented in Table 6 and Table 10 were used along with the destructive measurements for each of the 633 pipe joints to generate 5000 Monte Carlo inputs per pipe joint to analyze the predicted grade of each joint.
3. Proposed assumed uncertainties for chemical composition and calculated strength uncertainties (32 pipe samples). Of the 633 pipe joints, 32 had at least three destructive tensile tests performed, so YS and UTS standard deviations could be calculated for these joints. For these 32 pipe joints, the C, Mn, and S uncertainties from Table 6 and the calculated standard deviations for YS and UTS were used to generate 5000 Monte Carlo inputs per pipe joint to analyze the predicted grade of each joint.

The grade prediction case with calculated strength uncertainties applied to cases in the dataset where three or more tensile tests were available from a single pipe joint to calculate standard deviations for YS and UTS. It is noted that PG&E's process to generate inputs used in the PGCs uses a t-distribution, which accounts for smaller sample sizes. For this study, the standard deviation calculated from pipe joints with three or more samples is a sample standard deviation and cannot be guaranteed to be representative of the population standard deviation due to the small sample size.

Grade Predictions for Proposed Chemical Composition and Tensile Test Uncertainties (633 Samples)

The confusion matrices shown below in Figure 20 compare the accuracy of the grade predictions for each set of pipe grade calculator inputs. For the confusion matrix on the left, the predicted grades were generated using destructive testing measurements without uncertainties as a control, i.e., one input per pipe joint in the set of 633 validation samples.

For the confusion matrix on the right in Figure 20, the destructive testing measurements all had assigned uncertainties for C, Mn, S, YS, and UTS. The assigned uncertainties were used to generate a set of 5000 Monte Carlo inputs following a t-distribution for each pipe joint. For this confusion matrix, the predicted grade for each pipe joint was determined by selecting the grade with the highest

[30] LTB stands for Less than Grade B. This includes all grades with a specified minimum yield strength (SMYS) less than 35 ksi.

[31] P. Veloo et al., "Comparing Machine Learning Model Predictions to SME Determinations of Pipe Grade," 2024 American Gas Association (AGA) Operations Conference, Seattle, WA, April 28 - May 2, 2024.

https://doi.org/10.52202/078572-0042

mean probability after averaging the probabilities per grade for all Monte Carlo inputs for that pipe joint.

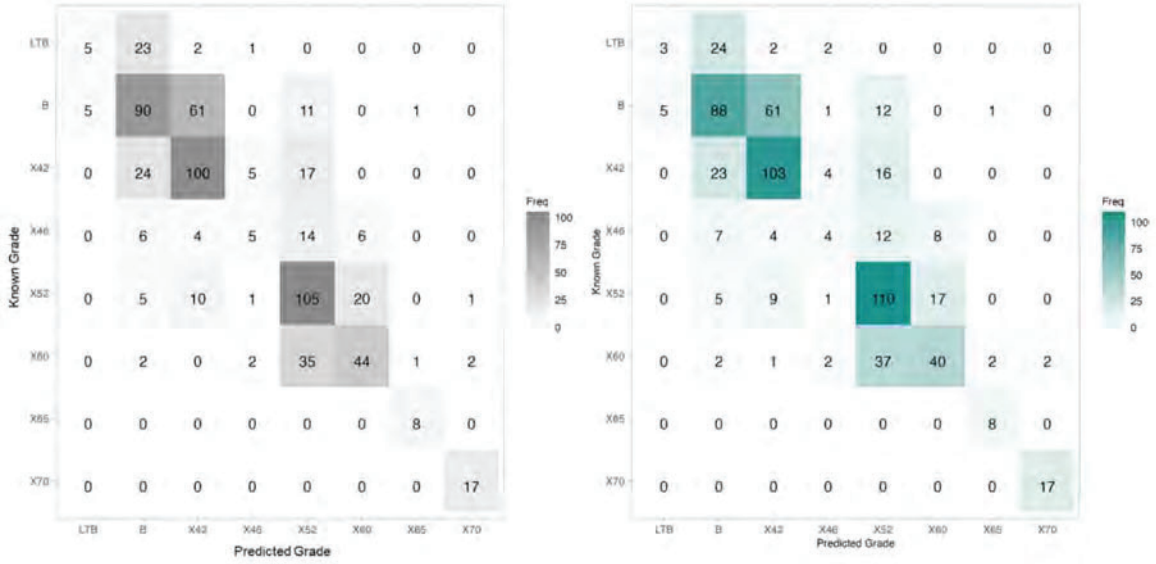

Figure 20. Confusion matrices without uncertainty (left) and with uncertainty (right) using the RSI CDS grade prediction model. Confusion matrix on the right for the highest mean probability.

Figure 21 reports the data shown in Figure 20, but the predicted grade for each pipe joint used to generate the confusion matrix on the right was determined by selecting the grade that was identified for the highest percentage of all of the Monte Carlo inputs generated for that pipe joint.

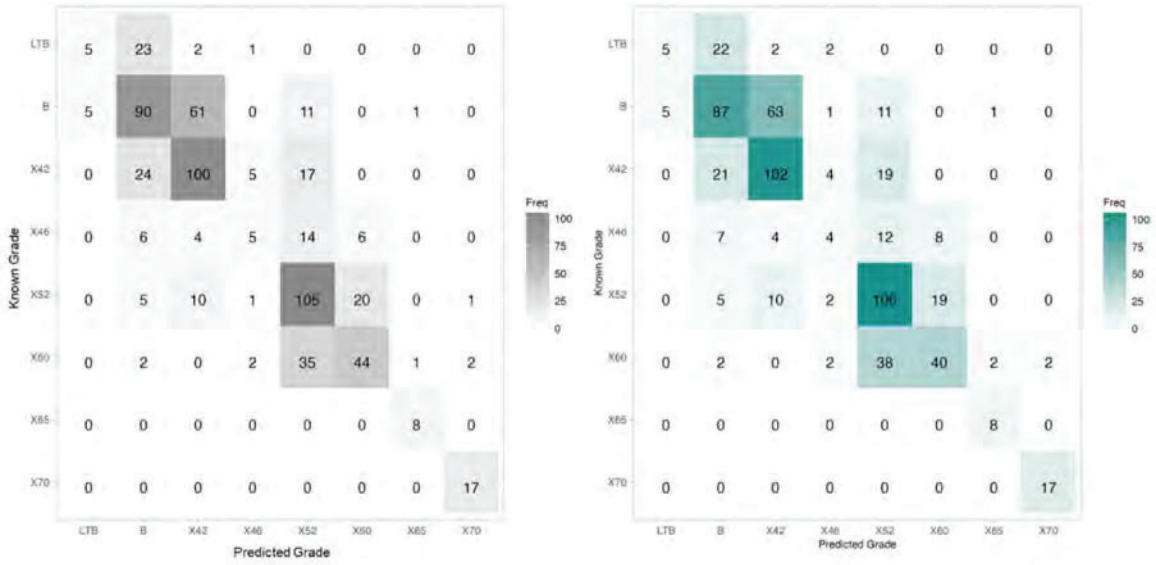

Figure 21. Confusion matrices without uncertainty (left) and with uncertainty (right) using the RSI CDS grade prediction model. Confusion matrix on the right for the highest percentage of cases.

https://doi.org/10.52202/078572-0042

The data presented in the confusion matrices in Figure 20 are summarized in the bar charts presented below in Figure 22. In each bar chart, the following is shown:

- Left: The number of pipe joints for which the predicted grade did not match the known grade but was conservative, i.e., predicted grade with a lower SMYS than the known grade.
- Center: The number of pipe joints for which the predicted grade matched the known grade.
- Right: The number of pipe joints for which the predicted grade did not match the known grade but was unconservative, i.e., predicted grade with a higher SMYS than the known grade.

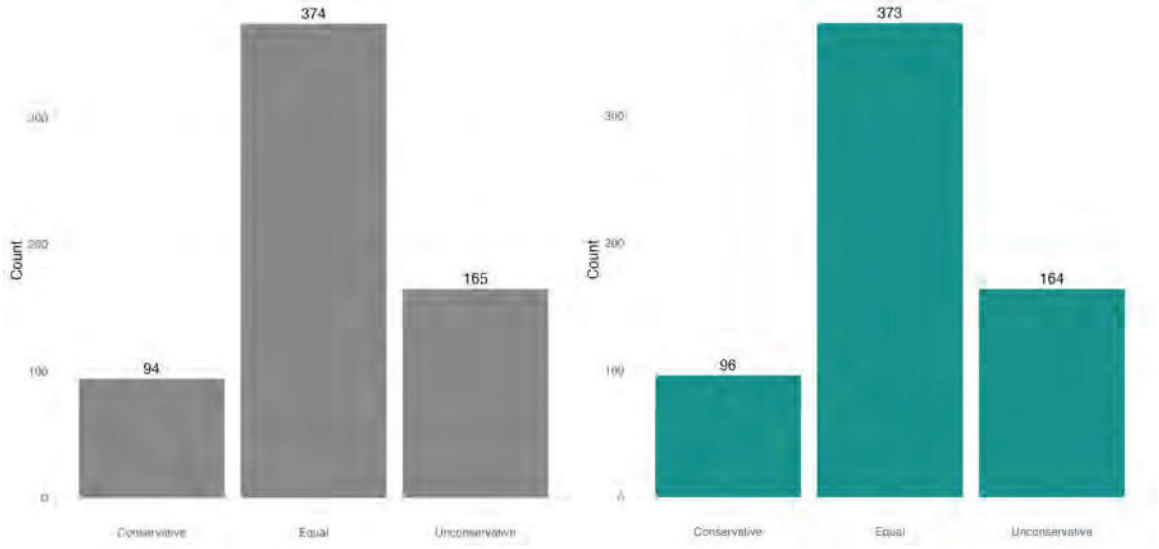

Figure 22. Bar charts without uncertainty (left) and with uncertainty (right) using the RSI CDS grade prediction model. Bar chart on the right is for the highest mean probability.

The data presented in the confusion matrices in Figure 21 are summarized in the bar charts presented below in Figure 23.

https://doi.org/10.52202/078572-0042

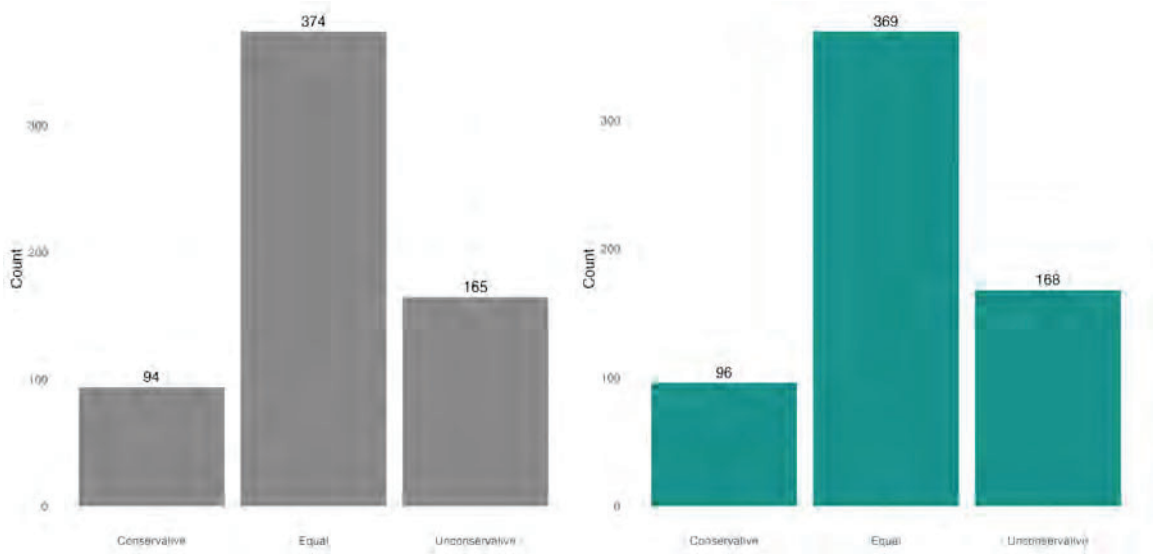

Figure 23. Bar charts without uncertainty (left) and with uncertainty (right) using the RSI CDS grade prediction model. Bar chart on the right is for the highest percentage of cases.

There were 19 instances (3%) where the grade with the highest mean probability was in disagreement with the grade corresponding to the highest percentage of cases. In these instances, the highest mean probability disagreed with the known grade 10/19 times while the highest percentage of cases disagreed with the known grade 14/19 times. These instances are shown in Table 12 below. Further analysis on the pipe joints listed in Table 12 may be useful in better understanding the sensitivity of the pipe grade prediction methodology to material properties. Notably, the highest mean probability was an unconservative grade prediction only 5 times whereas the highest percentage of cases was unconservative 9 times, making the highest mean probability metric both more accurate and more conservative for this dataset. Based on PG&E's experience, it is not uncommon for there to be discrepancies between predicted grade when comparing highest mean probability and highest percentage of cases. In these cases, it is often that two similar grades are "competing" with each other for the mostly likely predicted grade. That this dataset yielded a relatively small percentage of the total cases where highest mean probability and/or highest percentage of cases did not match known grade is a good indication that the assumed uncertainties selected are reasonable.

https://doi.org/10.52202/078572-0042

Table 11. Instances where the highest mean probability disagreed with the highest percentage of cases. Cells are shaded wherever the predicted grade does not match the known grade.

Pipe Joint	Highest Mean Probability	Avg. Probability	Highest Percentage of Cases	Perc. of Predictions	Known Grade
1091O	X42	0.28	X52	49%	X60
1093E	X52	0.30	X65	55%	X60
1128D	X60	0.51	X52	57%	X60
1159	B	0.46	LTB	50%	LTB
1168	X52	0.34	B	48%	B
1291	B	0.48	X42	50%	B
1347E	X42	0.52	B	56%	X42
136	B	0.44	X52	59%	X42
1422	B	0.46	LTB	52%	LTB
209	B	0.51	X42	51%	B
234	B	0.46	LTB	50%	LTB
330	X65	0.36	X60	47%	X60
338	X52	0.38	X46	49%	X52
362	X52	0.45	X60	63%	X52
36A	LTB	0.47	B	54%	LTB
403	X52	0.52	X60	51%	X52
422	B	0.44	X52	59%	X42
727	X52	0.42	X42	52%	X52
93A	B	0.44	X52	60%	X42

The bar charts and boxplots in Figure 24 below show the grade prediction distributions for three examples from the 19 instances where the highest mean probability disagreed with the highest percentage of cases. In each composite plot for each pipe joint, the following is shown:

- A: A bar chart of the mean probability of each grade over all the Monte Carlo cases for that pipe joint.
- B: A bar chart of the percentage of Monte Carlo cases where a predicted grade was selected as having the highest probability for that pipe joint.
- C: The boxplot of the probability of each grade over all the Monte Carlo cases for that pipe joint. The boxplot is important in that it helps visualize the distribution and overlap of the PGC results across grades.

https://doi.org/10.52202/078572-0042

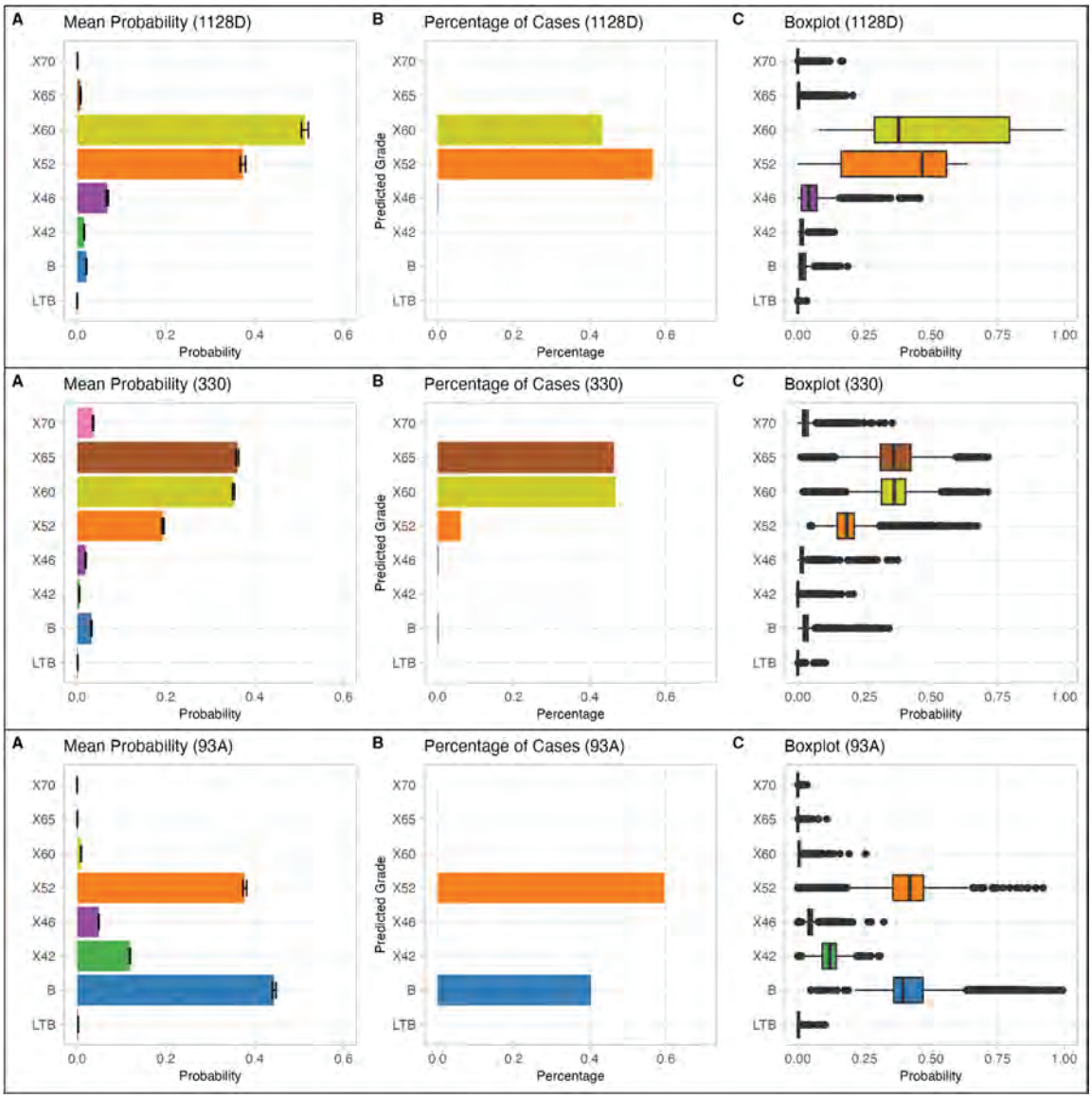

Figure 24. Examples of pipe joints where highest mean probability disagreed with the highest percentage of cases. Upper: The highest mean probability agreed with the known grade (X60). Middle: The highest percentage of cases agreed with the known grade (X60). Lower: Neither agreed with known grade (X42).

In all, the grade predictions with the proposed assumed uncertainties resulted in similar accuracy performance compared to the results that did not consider measurement uncertainty. Predicted grade was obtained using two different methodologies: (1) selection of the grade with the highest mean probability, or (2) selection of the grade with the highest percentage of cases. Both of these methodologies yielded similar accuracy performance.

https://doi.org/10.52202/078572-0042

Grade Predictions for Proposed Chemical Composition Uncertainties and Calculated Tensile Test Uncertainties (32 Samples)

The confusion matrices shown below in Figure 25 compare the accuracy of the grade predictions for each set of pipe grade calculator inputs. For the confusion matrix on the left, the predicted grades were generated using destructive testing measurements with the proposed assumed uncertainties. For the confusion matrix on the right in Figure 25, the destructive testing measurements for the chemical composition values used the proposed assumed uncertainties while the strength values utilized calculated uncertainties. These uncertainties were used to generate a set of 5000 Monte Carlo inputs following a t-distribution for each pipe joint. For this confusion matrix, the predicted grade for each pipe joint was determined by selecting the grade with the highest mean probability after averaging the probabilities per grade for all Monte Carlo inputs for that pipe joint.

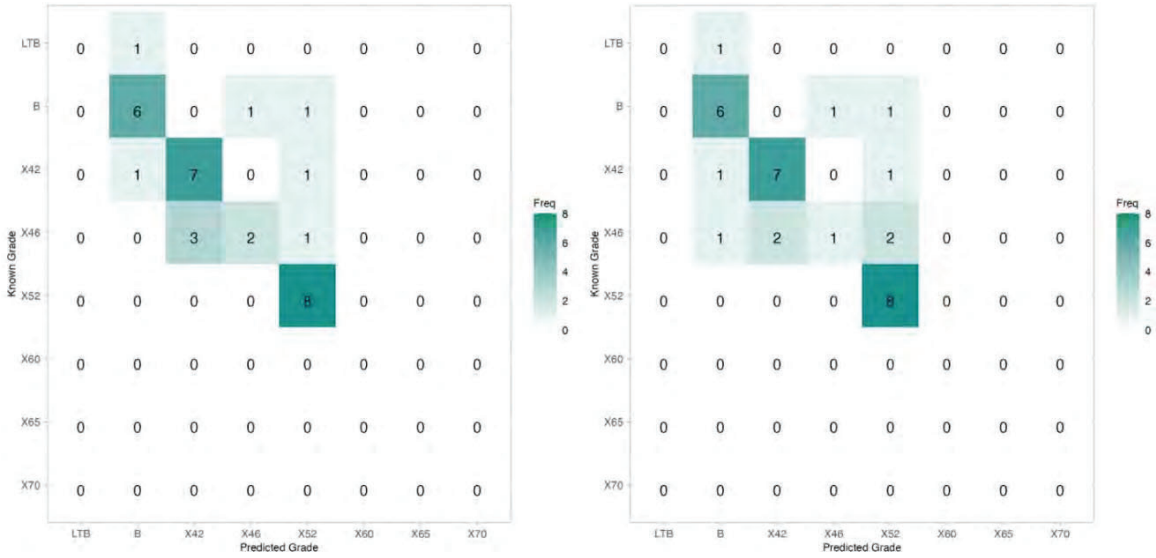

Figure 25. Confusion matrices with assumed tensile test uncertainty (left) and with calculated tensile test uncertainty (right) using the RSI CDS grade prediction model (highest mean probability).

Figure 26 reports the data shown in Figure 25, but the predicted grade for each pipe joint used to generate the confusion matrix on the right was determined by selecting the most probable grade that was identified for the highest percentage of all of the Monte Carlo inputs generated for that pipe joint.

https://doi.org/10.52202/078572-0042

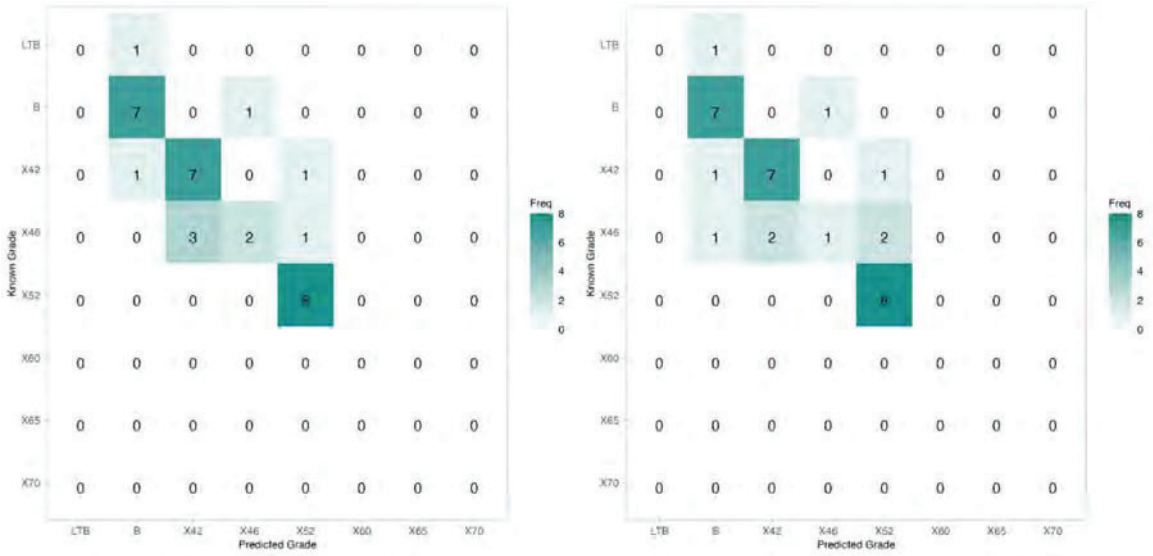

Figure 26. Confusion matrices with assumed tensile test uncertainty (left) and with calculated tensile test uncertainty (right) using the RSI CDS grade prediction model (highest percentage of cases).

The data presented in the confusion matrices in Figure 25 are summarized in the bar charts presented below in Figure 27. In each bar chart, the following is shown:

- Left: The number of pipe joints for which the predicted grade did not match the known grade but was conservative, i.e., grade prediction with a lower SMYS than the known grade.
- Center: The number of pipe joints for which the predicted grade matched the known grade.
- Right: The number of pipe joints for which the predicted grade did not match the known grade but was unconservative, i.e., grade prediction with a higher SMYS than the known grade.

https://doi.org/10.52202/078572-0042

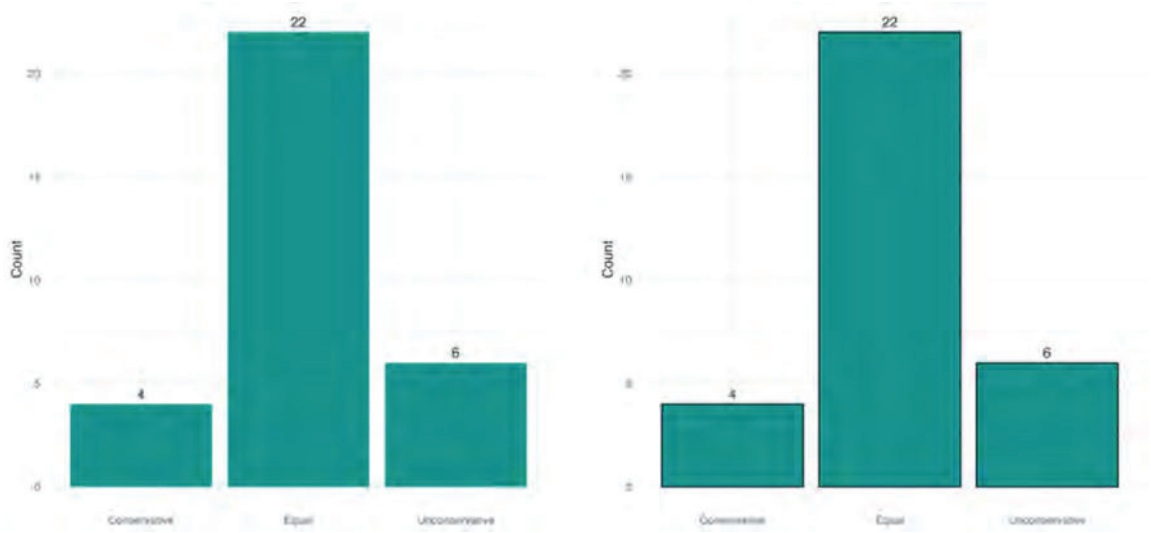

Figure 27. Bar charts indicating model performance with assumed tensile testing uncertainty (left) and with calculated tensile testing uncertainty (right) using the RSI CDS grade prediction model (highest mean probability).

The data presented in the confusion matrices in Figure 26 are summarized in the bar charts presented below in Figure 28.

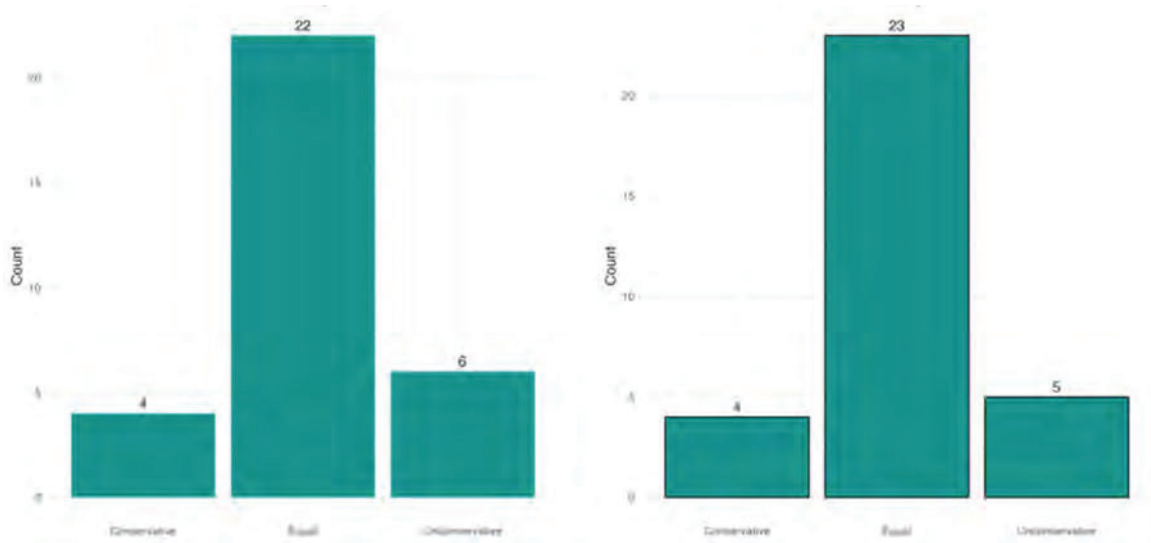

Figure 28. Bar charts indicating model performance with assumed tensile testing uncertainty (left) and with calculated tensile testing uncertainty (right) using the RSI CDS grade prediction model (highest percentage of cases).

Figure 29 through Figure 31 below show the grade prediction distributions for three examples where the calculated standard deviation for YS was either lower, approximately the same, or higher than the proposed assumed uncertainty. In each composite plot for each pipe joint, the following is shown:

https://doi.org/10.52202/078572-0042

- A: A bar chart of the mean probability of each grade over all the Monte Carlo cases for that pipe joint.
- B: A bar chart of the percentage of Monte Carlo cases where a predicted grade was selected as having the highest probability for that pipe joint.
- C: The boxplot of the probability of each grade over all the Monte Carlo cases for that pipe joint. The boxplot is important in that it helps visualize the distribution and overlap of the PGC results across grades.

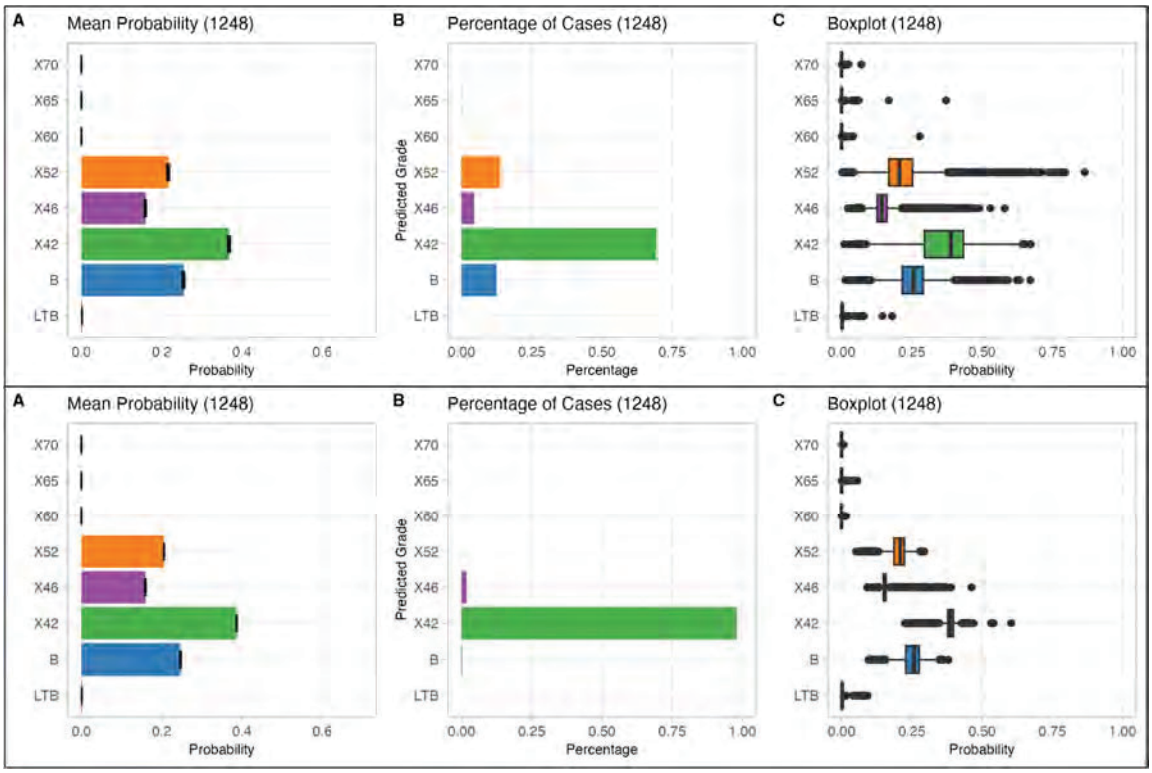

Figure 29. Example of pipe joint where calculated standard deviation for tensile results is lower than the assumed tensile standard deviation plot with (upper) assumed standard deviation (3.24 ksi), and (lower) calculated standard deviation (0.577 ksi).

https://doi.org/10.52202/078572-0042

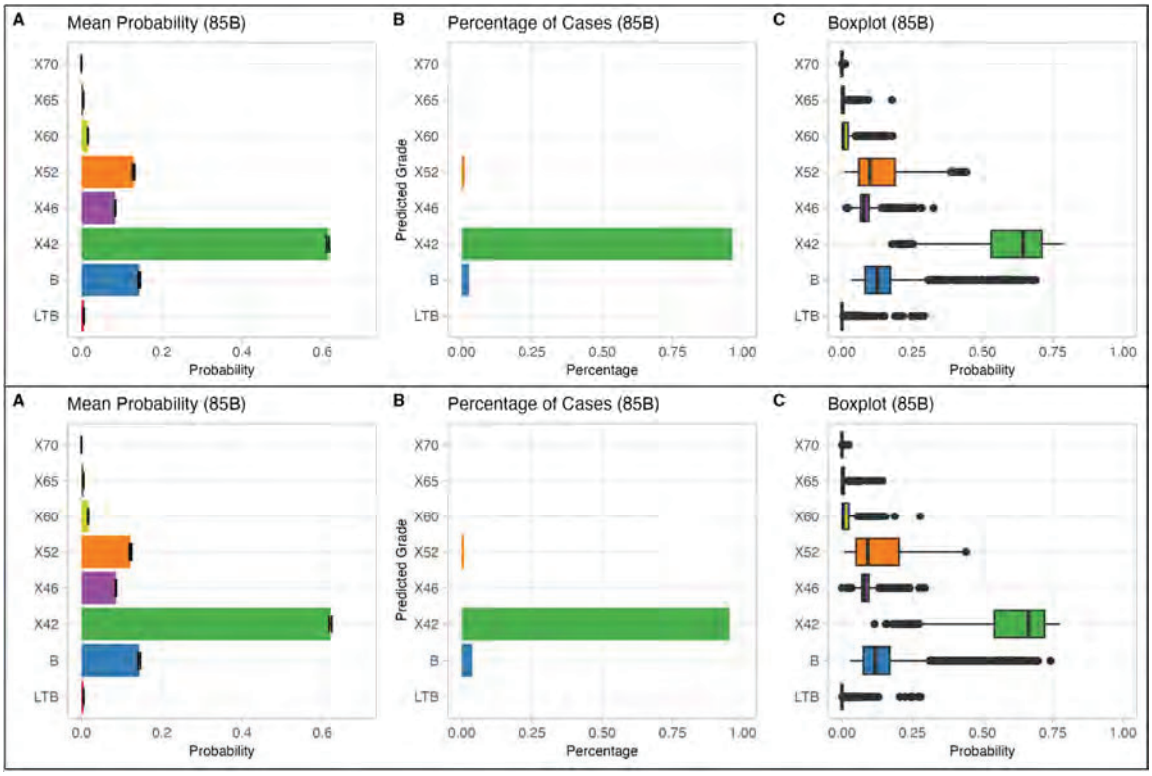

Figure 30. Example of pipe joint where calculated standard deviation for tensile results is around the same value as the assumed tensile standard deviation plot with (upper) assumed standard deviation (3.24 ksi), and (lower) calculated standard deviation (3.35 ksi).

https://doi.org/10.52202/078572-0042

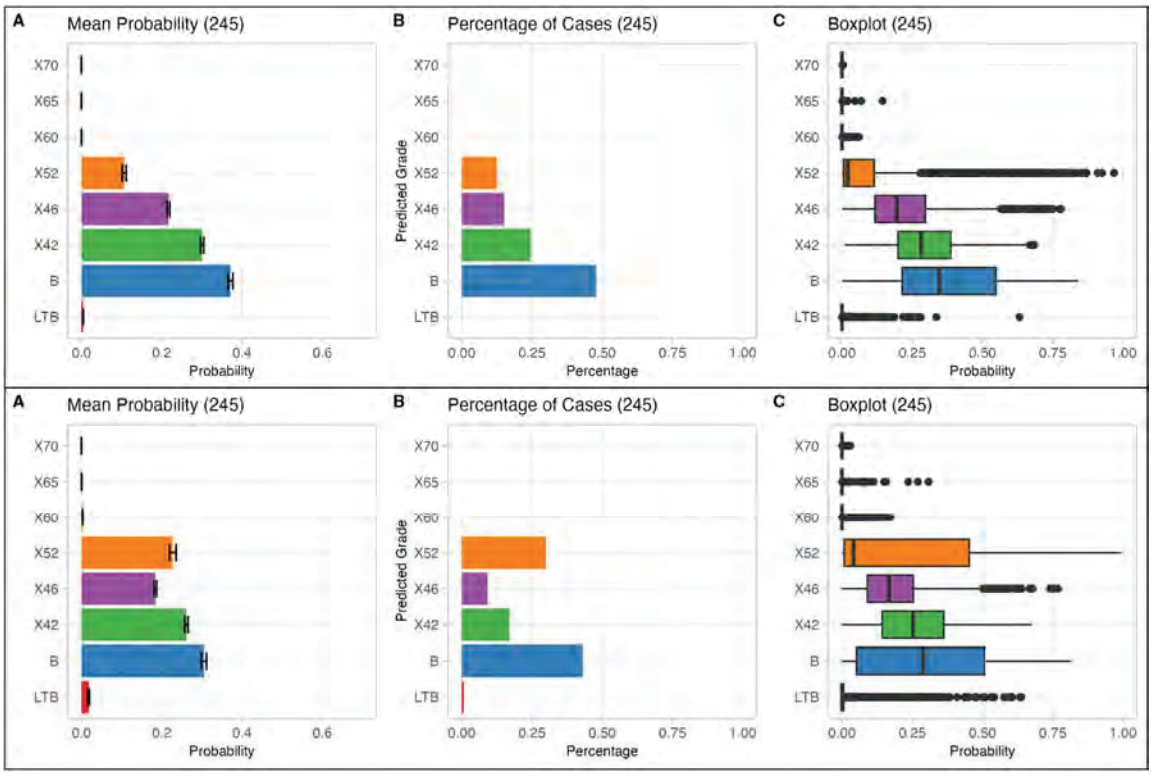

Figure 31. Example of pipe joint where calculated standard deviation for tensile results is higher than the assumed tensile standard deviation plot with (upper) assumed standard deviation (3.24 ksi), and (lower) calculated standard deviation (11.8 ksi).

Overall, the grade predictions with calculated strength uncertainties resulted in similar accuracy performance compared to the results that considered the proposed assumed strength uncertainties. Predicted grade was obtained using two different methodologies: (1) with the highest mean probability, or (2) with the highest percentage of cases. Both methodologies yielded similar accuracy performance.

Similar performance of accuracy between calculated and proposed assumed uncertainties for strength indicates that the proposed assumed tensile uncertainty should be representative for pipe joints where the estimated strength uncertainty cannot be calculated (i.e., where there are less than three tensile tests for a pipe joint).

While it did not appear to impact grade predictions, several of the 32 pipe joints had a calculated standard deviation for YS that were larger than the proposed assumed standard deviation. To achieve a robust grade determination process, it may be prudent to set a threshold flag to indicate that strength data should be reviewed for pipe joints where calculated standard deviations are notably larger than the proposed assumed standard deviation for YS.[32] In this way, a threshold flag for calculated YS uncertainty could help identify potential data issues when applied to older, previously recorded measurements.

[32] A reasonable proposed threshold would be 1.5 times the proposed assumed standard deviation for YS (4.86 ksi), but this value could be changed as PG&E continuously improves its grade determination process.

https://doi.org/10.52202/078572-0042