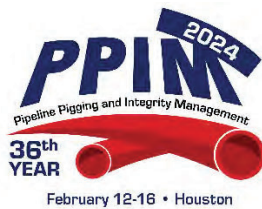# A Tool for the Identification of Anomalous NDT Chemical Composition Measurements During In Situ Testing

Janille Maragh[1], Peter Martin[2], Jonathan Gibbs[3]
Jeffrey Kornuta[1], Peter Veloo[3]
[1] Exponent, Inc., [2] RSI Pipeline Solutions, LLC
[3] Pacific Gas and Electric Company

Pipeline Pigging and Integrity Management Conference

February 12-16, 2024

Organized by
Clarion Technical Conferences

Abstract

Nondestructive testing (NDT) of chemical composition is a critical component of the Pacific Gas and Electric Company's (PG&E) materials verification program. Additionally, 49 CFR § 192.607 states that the operator must "conservatively account for measurement inaccuracy and uncertainty using reliable engineering tests and analyses." Accurate and precise NDT composition data can be used to determine or verify certain characteristics of a pipe joint feature, for example vintage, grade, or manufacturing process. However, it has been observed that composition measurements may at times be inconsistent across various field analytical tools, possibly due to the variability of experimental, environmental, and other statistical (random) factors. Anomalous field NDT measurements are problematic because they could lead to the mischaracterization of pipe features during the materials verification process.

In this paper, we present a tool for the in-field checking of chemical composition NDT data for potentially erroneous or inconsistent measurements. The presented tool, which is spreadsheet-based and therefore easily and independently implemented by operators, identifies sets of measurements that exhibit two commonly observed irregularities in field data: (1) large standard deviations, which usually indicate the presence of outliers or excessive scatter; and (2) greater-than-usual differences between the chemical compositions at two different analyzed positions on the same pipe joint. For a given set of measurements, the tool flags the standard deviation as too large if it is high relative to standard deviations that have been observed in a database of thousands of pipe joint field chemical composition measurements previously collected as part of PG&E's materials verification program for the tool/element combination of interest. Large differences between the chemical composition measurements at different test locations on the pipe joint, which may be the result of surface contamination, surface preparation issues, or environmental differences, are identified through comparison to compositional differences that have been observed in the PG&E materials verification database. Because this tool is designed to work with raw data from NDT chemical composition tools, it may be used to enable field technicians to identify potential errors in near real-time. The improved data reliability increases confidence in downstream analysis and reduces the need to return to a site to perform re-work, allowing for a more efficient and less costly materials verification process. The implementation and in-field use of this tool for live retesting to mitigate potential downstream issues are demonstrated for a set of portable laser-induced breakdown spectroscopy (LIBS) data.

Introduction

Materials Verification

The Department of Transportation's Pipeline and Hazardous Materials Safety Administration (PHMSA) in 49 CFR § 192.607 has required gas pipeline operators to perform materials verification for pipes and fittings for which traceable, verifiable, and complete (TVC) records do not exist. Verification by nondestructive testing (NDT) methods can be preferable to the use of destructive testing (DT) methods because NDT allows for materials testing to occur without service interruptions and without physically cutting into the pipe.

The Pacific Gas and Electric Company (PG&E) has developed a methodology for using NDT methods to obtain measurements for material properties, such as elemental composition and strength, which are used for pipe grade estimation.[1] Consistent with 49 CFR § 192.607, this framework accounts for measurement inaccuracy and uncertainty when using NDT methods. PG&E has used various NDT methodologies, such as filings analysis, portable spark optical emission spectroscopy (OES), X-ray fluorescence spectroscopy (XRF), and laser induced breakdown spectroscopy (LIBS).[2,3]

Improving the Accuracy of Field Data

Review and post-processing of NDT field data are critical components of the materials verification process since they help to ensure that the NDT measurements are representative of the pipe features being tested. This can include a review of NDT data by a subject matter expert (SME) to help the operator identify and discard erroneous data prior to subsequent analyses[4] and post-processing of NDT datasets that contain measurements below an instrument's limit of detection.[5] Data review is crucial since it can identify issues such as inconsistent tool behavior, excessive statistical variations in the data, and poor agreement between data collected at different regions of the same pipe. Since the stages of the materials verification process depend on upstream data collection and analyses, ensuring good data quality as early as possible leads to a more efficient workflow.

---

[1] M. Rosenfeld, J. Ma, "Probabilistic Determination of Pipe Grade", The 29th International Pipeline Pigging and Integrity Management Conference (PPIM), Houston, TX, USA 2017

[2] M.W. Louie, B. Amend, M. Liong, M. Gould, N. Switzner, T. Rovella, P. Martin, P. Veloo, "Nondestructive Testing of Pipeline Materials: Analysis of Chemical Composition from Metal Filings", The 31st International Pipeline Pigging and Integrity Management Conference (PPIM), Houston, TX, USA (February 2019)

[3] M. Liong, N. Switzner, P. Veloo, M. Gould, P. Martin, T. Rovella, "Evaluation of portable SOES, XRF, LIBS, and filings to quantify chemical composition of pipeline steel", The 32nd International Pipeline Pigging and Integrity Management Conference (PPIM), Houston, TX, USA (February 2020)

[4] J. Maragh, P. Martin, J. Anderson, J. Gibbs, J. Kornuta, P. Veloo, " Identifying irregular and erroneous chemical composition data from in situ nondestructive testing," The 35th International Pipeline Pigging and Integrity Management Conference (PPIM), Houston, TX, USA (February 2023)

[5] J. Maragh, C. Liu, N. Switzner, J. Gibbs, J. Kornuta, P. Veloo, "Reducing Bias in Chemical Composition Data with Measurements Below the Limit of Detection," The 34th International Pipeline Pigging and Integrity Management Conference (PPIM), Houston, TX, USA (February 2022)

Efficient and Accurate Field Data Collection

The tool presented in this paper improves the data collection and review process by enabling the more *efficient* collection of acceptable field data by allowing for the live review of data in the field by a technician. In a typical process, field technicians collect field composition data, send the data to an SME for review, and wait up to 24 hours for guidance on whether additional data must be collected. This process may, at times, occur yet again if issues are identified by the SME in the recollected data. This extended process increases both the time needed to collect acceptable field composition data and any labor costs (e.g., SME time for multiple cycles of data review and technician time for data recollection).

The field data checking tool presented in this paper is intended for use by field technicians during or shortly after data collection. The tool flags data if it appears to exhibit issues that have been commonly observed in previously collected field data that triggered re-testing. When the tool flags data, the technicians may anticipate possible retest requests by the SME and proceed with retesting any features with potentially problematic data prior to the SME review. By enabling the anticipation of retest requests, the use of this field data checking tool could greatly reduce the likelihood that the field technicians will have to wait to collect additional data following the SME review stage, thereby saving the operator both time and cost during the field data collection process.

Field Data Checker

Typical Issues Observed in Field Composition Data

In a typical data collection process during materials verification, field technicians are asked to prepare and analyze two regions on each pipe feature, which are referred to as "test locations." At each test location, approximately 10-12 individual measurements are collected using both XRF and LIBS. Through the collection of in-field XRF and LIBS data for dozens of stations, several chemical composition data issues have been observed during preliminary subject matter expert (SME) review, which determines whether the field data may be accepted as-is or contains deficiencies that require additional data collection. These chemical composition data issues may be broadly categorized as follows:

1. Excessive scatter is observed in the measurements collected at a given test location on a given pipe feature.
2. The mean composition data collected at one test location on a pipe feature is atypically different from the mean composition data collected at the second test location on the same pipe feature, as compared to previously observed chemical composition differences between test locations.

Flagging Scatter as Excessively Large

Excessive scatter may be indicative of several issues, such as the presence of outliers in the data, inconsistent behavior of the tool, or contamination occurring part way through data collection. The standard deviation represents the spread or scatter within a set of measurements collected at a single test location on a pipe feature. One method by which the standard deviation can be flagged as excessively large is by assessing whether it is large in comparison to previous standard deviations observed for prior "good" sets of measurements for this tool-element combination.

When selecting prior measurements to use for comparison, it may be necessary to use only measurements collected using the same tool model. Instruments from different manufacturers could vary in performance when measuring different elements, so typical scatter observed in the carbon measurements collected using one manufacturer's LIBS tool, for example, may be dissimilar from the scatter observed in carbon measurements collected using a different manufacturer's LIBS tool.

For the data checker tool, the prior values to use for comparison with the field data measurements' standard deviation are calculated using thousands of measurements for hundreds of pipe features previously analyzed as part of PG&E's materials verification program as follows. As an example, the field measurements being assessed are XRF sulfur measurements collected using Manufacturer A's XRF tool, and these field measurements have a standard deviation of σ_{field}.

1. Starting with all prior field measurements, filter the data to include only XRF sulfur measurements that were collected using Manufacturer A's XRF tool.
2. Compute the standard deviation for each of the N prior test locations, $\sigma_{TL,i}$
3. Compute the mean for all the standard deviations as:

$$\mu_{\sigma,prior} = \frac{\sum \sigma_{TL,i}}{N} \tag{1}$$

4. Compute the standard deviation of all the standard deviations as:

$$\sigma_{\sigma,prior} = \sqrt{\frac{\sum(\sigma_{TL,i} - \mu_{\sigma,prior})^2}{N-1}} \tag{2}$$

The standard deviation of the XRF sulfur measurements being assessed, σ_{field}, is then flagged as potentially too large if it is more than one standard deviation greater than the mean of the prior standard deviations, that is, if

$$\sigma_{field} > \mu_{\sigma,prior} + \sigma_{\sigma,prior} \tag{3}$$

Examples of the calculated LIBS and XRF values for $\mu_{\sigma,prior}$ and $\sigma_{\sigma,prior}$ are presented below in Table *1* for the elements of interest carbon, silicon, manganese, and sulfur. Note that the values presented in Table *1* are specific to one particular field equipment manufacturer, so these values may have to be recalculated for each field equipment manufacturer in use by the operator.

Table 1. Calculated means and standard deviations of the standard deviations of sets of measurements on pipe feature test locations as measured using LIBS (for carbon and silicon) and XRF (for manganese and sulfur).

Field Analysis Technique	Element	$\mu_{\sigma,prior}$	$\sigma_{\sigma,prior}$
LIBS	Carbon	0.019	0.013
LIBS	Silicon	0.018	0.012
XRF	Manganese	0.014	0.019
XRF	Sulfur	0.002	0.002

Flagging Measurements for a Pipe Feature as Inconsistent

Since the chemical composition of a pipe feature should be independent of the axial and circumferential position of the test location, the chemical composition measurements collected at two test locations ground to similar depths on the same pipe feature should also be similar. Large differences between the chemical composition measurements at two test locations could indicate instrument issues, surface preparation issues, or other problems with the data. Note that for this metric, it is assumed that the differences in test location chemical composition are independent of the instrument manufacturer. Although different instrument manufacturers may exhibit varying performance (e.g., degree of scatter, offsets), these differences are considered to be less important since the value of interest is the difference in composition observed at two arbitrary locations on a pipe feature.

One method by which the difference between the mean measurement of two test locations on the same pipe feature can be flagged as "excessive" is by assessing whether the difference is large relative to all differences that have been previously observed between two test locations on the same pipe feature. For example, if the difference between the means of the XRF sulfur measurements at two test locations on the same pipe feature is Δ_{field}, then Δ_{field} can be compared to previously observed values that are calculated as follows.

1. Starting with all prior field measurements, filter the data to include only XRF sulfur measurements for pipe features that have data collected on two test locations.

2. For each of the N pipe features with two test locations, compute the difference between the means at the two test locations on that pipe feature, $\Delta\mu_{feature,i}$

3. Compute the mean of all the test location differences as:

$$\mu_{\Delta\mu,prior} = \frac{\sum \Delta\mu_{feature,i}}{N} \tag{4}$$

4. Compute the standard deviation of all the test location differences as:

$$\sigma_{\Delta\mu,prior} = \sqrt{\frac{\sum(\Delta\mu_{feature,i}-\mu_{\Delta\mu,prior})^2}{N-1}} \tag{5}$$

The difference in test location XRF sulfur means for the pipe feature assessed is flagged as potentially too large if it is more than one standard deviation greater than the mean of the previously observed test location differences, that is, if

$$\Delta_{field} > \mu_{\Delta,prior} + \sigma_{\Delta,prior} \tag{6}$$

Examples of the calculated LIBS and XRF values for $\mu_{\Delta,prior}$ and $\sigma_{\Delta,prior}$ are presented below in Table 2 for the elements of interest carbon, silicon, manganese, and sulfur.

Table 2. Calculated means and standard deviations of the mean differences in test location measurements on pipe features as measured using LIBS (for carbon and silicon) and XRF (for manganese and sulfur).

Field Analysis Technique	Element	$\mu_{\Delta,prior}$	$\sigma_{\Delta,prior}$
LIBS	Carbon	0.022	0.020
LIBS	Silicon	0.013	0.013
XRF	Manganese	0.008	0.009
XRF	Sulfur	0.003	0.003

Note that the values presented in Table 1 and Table 2 were calculated using field measurements that are dependent on several factors, including field equipment performance, testing conditions, pipe joint material properties, and surface preparation. Therefore, the presented values are only provided as examples, and they are not intended to be universally applied in the field. Further, these values are not used to decide whether field data is or not acceptable. Rather, these calculated values are used to flag field data that are *potentially* inconsistent with prior observations during preliminary data review.

Tool Design and Implementation

The field checker tool was implemented as a workbook in Microsoft Excel[6] due to the wide availability and user-friendly interface of the software. The tool was also customized so that it could be used directly with the raw data as it is produced by the field tool's software. The user-friendliness and compatibility of the tool with raw data files are both crucial for minimizing the disruptiveness of the tool in the field technician's workflow.

Pre-loaded pivot tables are then used to convert the raw data table in one sheet of the workbook to a summary table in another sheet. This pivot table is constructed by:
1. Filtering on the unique measurement number and the feature name to display only the data for a single pipe feature in the pivot table.
2. Setting the rows to be the names of the test locations on the pipe feature.
3. Setting the values to be the averages and standard deviations of the elements of interest.

---

[6] Microsoft Corporation, 2023. *Microsoft Excel*, Available at: https://office.microsoft.com/excel.

Note that this procedure assumes a typical output table, in which data for each element, the pipe feature names, and the test location names are stored as single columns anywhere in the table. The resulting pivot table show the means and standard deviations of the measurements of the elements of interest (i.e., carbon, manganese, sulfur, and silicon) at each of the two test locations on the pipe feature.

The calculated means and averages are then compared to the historical field data as described in the previous two sections. If the measurements meet the criteria stated in Equations 3 and 6, they are highlighted in red using conditional formatting to flag to the technician that the measurements may trigger a retest request during SME review. The technician may then preemptively retest that location or pair of locations to clear the flag in the field data checker.

Field Implementation Example

In the screenshot shown below in Figure 1, the output summary table (the pivot table) previously described is shown for a set of LIBS field data. The LIBS raw data was copied to the specified sheet in the workbook, and the pivot table was updated using the "Refresh" button (yellow box) in the "Pivot Analyze" menu ribbon. The cells then auto-format to indicate potentially problematic sets of measurements. A red cell in the "Averages" section of the summary table (left) indicates that the mean values at the two test locations on the pipe feature may be too different compared to previously collected data. A red cell in the "Standard Deviations" section of the data checker table (right) indicates that there may be excessive scatter in a set of measurements at a single test location.
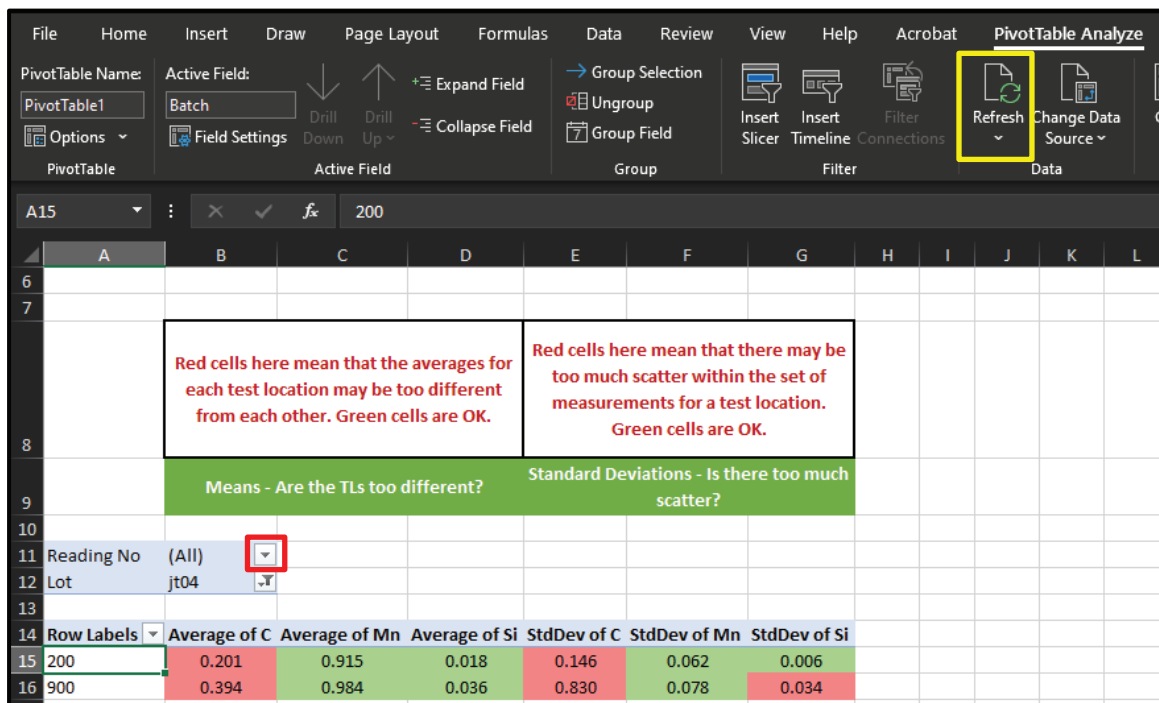

Figure 1. Screenshot from the field data checker showing the conditionally formatted summary (pivot) table of the LIBS raw data.

The technician may then inspect the data to determine whether there are obvious outlier measurements that, if removed from the analysis, would clear the flags from measurements sets shown in the field checker table. To do so while preserving the raw data file, those outlier measurements can be removed from the analysis by deselecting the individual measurements using the uppermost filter (in Figure *1* next to "Reading No", red box) and deselecting individual reading numbers (e.g., 363) as shown in Figure *2*.
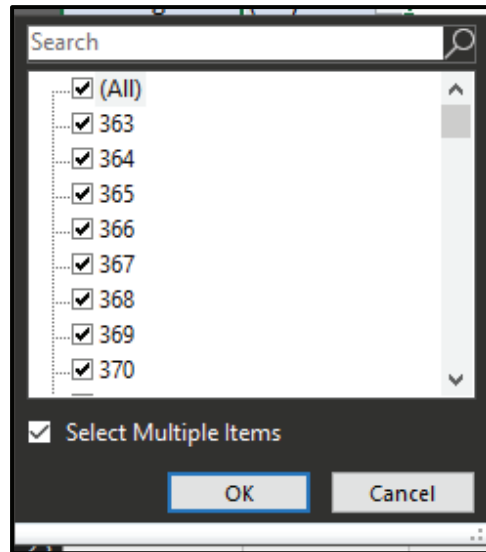

Figure 2. Screenshot showing default view after selecting the arrow next to Reading No in the data checker pivot table.

If, upon inspection of the data, the field technician determines that the inconsistencies flagged by the data checker are not due to obvious outlier measurements, they may then decide to proceed with retesting flagged features prior to SME review.

Conclusions

In this work, an Excel-based tool for performing preliminary checks of field data was presented. The tool uses pivot tables to convert the raw data produced by the tool to summary tables showing the mean measurements and standard deviations of the elements of interest. The mean measurements and standard deviations are then compared to values calculated from previously collected field data to determine whether the scatter of each measurement set and the differences between measurement sets on the same pipe feature are potentially problematic. Conditional formatting is then used to make it obvious to the technician which sets of measurements are potentially problematic and more likely to trigger a retest request. The implementation of this tool by field technicians as part of the field data collection and review process may improve the efficiency of the field work process by saving the operator both time in the field and labor costs.