COMPARISON OF METHODOLOGIES FOR FINITE ELEMENT MODEL VALIDATION OF RAILROAD TANK CAR SIDE IMPACT TESTS

Shaun Eshraghi, Michael Carolan, Benjamin Perlman, Francisco González III
Volpe National Transportation Systems Center U.S. Department of Transportation Cambridge, Massachusetts, USA

ABSTRACT

The U.S. Department of Transportation’s Federal Railroad Administration (FRA) has sponsored a series of full-scale dynamic shell impact tests on railroad tank cars. For each shell impact test a pre-test finite element (FE) model is created to predict the overall force-time or force-displacement histories of the impactor, puncture/non-puncture outcomes of the impacted tank shell, global motions of the tank car, internal pressures within the tank, and the energy absorbed by the tank during the impact. While qualitative comparisons (e.g. the shapes of the indentation) and quantitative comparisons (e.g. peak impact forces) have been made between tests and simulations, there are currently no standards or guidelines on how to compare the simulation results with the test results, or what measurable level of agreement would be an acceptable demonstration of model validation.

It is desirable that a framework for model validation, including well-defined criteria for comparison, be developed or adopted if FE analysis is to be used without companion full-scale shell impact testing for future tank car development. One of the challenges to developing model validation criteria and procedures for tank car shell puncture is the number of complex behaviors encountered in this problem, and the variety of approaches that could be used in simulating these behaviors. The FE models used to simulate tank car shell impacts include several complex behaviors, which increase the level of uncertainty in simulation results, including dynamic impacts, non-linear steel material behavior, two-phase (water and air) fluid-structure interaction, and contact between rigid and deformable bodies.

Approaches to model validation employed in other areas of transportation where validation procedures have been documented are applied to railroad tank car dynamic shell impact FE simulation results. This work compares and contrasts two model validation programs: Roadside Safety Verification and Validation Program (RSVVP) and Correlation and Analysis Plus (CORA). RSVVP and CORA are used to apply validation metrics and ratings specified by the National Cooperative Highway Research Program Project 22-24 (NCHRP 22-24) and ISO/TS 18571:2014 respectively. The validation methods are applied to recently-completed shell impact tests on two different types of railroad tank cars sponsored by the FRA. Additionally, this paper includes discussion on model validation difficulties unique to dynamic impacts involving puncture.

1. INTRODUCTION

In recent years, significant research has been conducted to analyze and improve the impact behavior and puncture resistance of railroad tank cars used in the transportation of hazardous materials. Ultimately, the results of this research can be used by the FRA in the US and Transport Canada (TC) in Canada to establish performance-based testing requirements and to develop methods to evaluate the crashworthiness and structural integrity of different tank car designs when subjected to a standardized shell (i.e., the side of the tank) impact scenario. A performance-based requirement for tank car head (i.e., the end of the tank) impact protection has already been defined within the current U.S. regulations [1].

The FRA has an ongoing research program to provide the technical basis for enhanced and alternative performance standards for tank cars. As a part of this program, new and innovative designs that are developed by the industry and other countries are also reviewed. In support of this research program, full-scale shell impact tests are necessary to provide the technical information to validate modeling efforts and to inform technology transfer and industry interaction activities. These tests evaluate the crashworthiness performance of tank cars used in the transportation of various categories of hazardous materials. The Volpe National Transportation Systems Center (Volpe Center) supports the FRA in this research effort, and has performed pre- and post-test finite element (FE) analyses
alongside several of the full-scale shell impact tests. As shell impact testing of tank cars is destructive, expensive, and challenging to perform, this type of problem is an attractive candidate for FE analysis.

Since 2007, the FRA has sponsored a series of shell impact tests of tank cars of various designs, using a standardized impact test setup. The standardized shell impact test setup constrains the tank car’s motion by supporting the car against a rigid barrier. This creates a severe impact condition for the tank car’s shell, as the kinetic energy (KE) of the initially-moving ram car must be dissipated solely through deformation of the struck tank car. Since the mass of the impacting ram car was approximately the same in all of the tests, impact speed and initial KE are interchangeable measures of the impact conditions in each test.

One of the key outcomes of each test is whether, under the defined impact conditions, the tank car was punctured or resisted the impact without puncturing. If puncture occurs, a key measurement from the test is the puncture energy calculated by integrating impactor force over impactor travel up to the point of puncture. A variety of impact heads have been used in the tests conducted to-date, but typically a smaller (sharper) impactor results in a much lower puncture energy because the impactor acts like a needle.

A table summarizing the shell impact tests is shown in Table 1. These shell impact tests have involved testing of tank cars constructed to various specifications, and included a mix of cars designed to carry pressurized gases (DOT-105) and flammable liquids (DOT-111 and DOT-117). Some of these tests were performed as a part of a government-industry collaborative program referred to as the Next Generation Railroad Tank Car Project (denoted with a †).

<table>
<thead>
<tr>
<th>Test #</th>
<th>Test Date MM/DD/YYYY</th>
<th>Tank Car Specification</th>
<th>Impact Speed mph</th>
<th>Impactor Size inches</th>
<th>Initial KE 10^6 ft-lbf</th>
<th>Puncture Energy 10^6 ft-lbf</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4/26/2007</td>
<td>DOT-105†</td>
<td>14.0</td>
<td>17 x 23</td>
<td>1.9</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>7/11/2007</td>
<td>DOT-105†</td>
<td>15.1</td>
<td>6 x 6</td>
<td>2.1</td>
<td>0.9</td>
</tr>
<tr>
<td>3</td>
<td>5/18/2011</td>
<td>DOT-105 (w/ panel)</td>
<td>17.8</td>
<td>12 x 12</td>
<td>3.1</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>5/18/2013</td>
<td>DOT-111</td>
<td>14.0</td>
<td>12 x 12</td>
<td>1.9</td>
<td>1.5</td>
</tr>
<tr>
<td>5</td>
<td>2/26/2014</td>
<td>DOT-112</td>
<td>14.7</td>
<td>12 x 12</td>
<td>2.1</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>4/27/2016</td>
<td>DOT-105</td>
<td>15.2</td>
<td>12 x 12</td>
<td>2.3</td>
<td>2.3</td>
</tr>
<tr>
<td>7</td>
<td>9/28/2016</td>
<td>DOT-117</td>
<td>13.9</td>
<td>12 x 12</td>
<td>1.9</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>8/1/2018</td>
<td>DOT-105</td>
<td>9.7</td>
<td>6 x 6</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>9</td>
<td>10/30/2018</td>
<td>DOT-111 (CPC-1232)</td>
<td>13.9</td>
<td>12 x 12</td>
<td>1.9</td>
<td>1.9</td>
</tr>
</tbody>
</table>

From this table, it is apparent that only a small number of tank car shell impact tests have been performed over the past decade. Additionally, no two tests have had exactly the same impact conditions. While each test produced useful data for both understanding the shell impact response of a tank car under those particular conditions and for performing comparisons with FE models, the test data do not span every conceivable shell impact scenario. FE modeling is used in conjunction with the tests to plan for the impact conditions, estimate the tank’s response under those conditions, evaluate alternative impact conditions, and extrapolate from the test conditions to other conditions of interest.

A primary purpose for pre-test modeling is to estimate the target impact speed for an upcoming test, and how that speed may relate to a threshold puncture speed. One of the goals in performing a test is to estimate the threshold puncture speed of the tank car being tested under the prescribed impact conditions. Puncture speed is an attractive metric to use in comparing the relative performance of different tank car designs under similar impact conditions as the goal of the research program is to improve the performance of tank cars involved in incidents, including minimizing the loss of product. The threshold puncture speed can be thought of as the maximum speed at which the tank car can be impacted under the prescribed conditions without resulting in a tear to its shell that would allow its lading to escape.

A test speed that is too high will result in excessive destruction of the tank car, while a test speed that is too low may not provide enough data to be useful. The threshold puncture speed of the tank car is the speed at which, under the test conditions, the initial kinetic energy of the ram is exactly equal to the energy necessary to puncture the tank shell. At this speed, an incrementally slower test would be a non-puncture test, and an incrementally faster test would exceed the capacity of the tank car to resist puncturing. The threshold puncture speed can be thought of as a range of speeds rather than a single numerical value due to manufacturing variability, variation in material properties, accuracy of measurement, and limitations on the test setup. This concept of uncertainty is shown schematically in Figure 1 where the red zone denotes the range of impact speeds that could possibly result in puncture and the circle denotes an example target speed for a given test where puncture is a possibility but not a certainty. In the context of this study, puncture is defined as the point where lading is able to escape from the tank; however, in other studies it may be defined differently.

This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States. Approved for public release; distribution is unlimited.
Ideally, the pre-test FE model is capable of predicting all of the responses that are measured or observed during the test. In practice, some difference between the FE results and the test measurements is expected. Additionally, based on the actual impact conditions (e.g., measured impact speed, post-test material characterization), it is usually necessary to make some adjustment to the pre-test model after the test to be able to simulate the actual test conditions, creating a post-test model. Depending on the nature of the changes made to the model, these changes may be considered model calibration, where the intent is to adjust the physical modeling parameters in the model to better match the test data, or may simply be adjustments to the pre-test model to better match the actual test conditions [2].

Given that there will be differences between the test measurements and corresponding results from an FE model, it is valuable to develop targets for comparisons to be made between the test measurements and FE results to be used to validate that the model is producing physically-realistic results for the system being modeled [2]. This is especially important if an FE model is intended to be used to simulate conditions beyond what was tested, as there will not be corresponding test data to serve as a check on the reasonableness of the model’s results.

While FE model results have been compared to test measurements for each of the tests summarized in Table 1, there are currently no requirements or formal guidelines on which specific behaviors should be compared, or what measurable level of agreement would be acceptable demonstration of model validation. This paper focuses its discussion on two recent shell impact tests and companion FE analyses: a DOT-105 tank car (April 27, 2016) [3] and a DOT-117 tank car (September 28, 2016) [4]. Further discussion of other tests and analyses can be found in the respective test reports (references [5][6][7]).

2. SHELL IMPACT SCENARIO

For the tank car shell impact test, a standardized, repeatable, controllable, and safe impact scenario was chosen. The tank car undergoing testing is removed from its trucks (bogies) and placed on two skids intended to limit the amount of roll that can occur after impact. The tank car is then placed perpendicular to a set of railroad tracks, with the area of the shell to be impacted centered between the rails. The tank car is placed against a stiff wall, limiting its ability to move away from the impacting car. A heavy ram car, equipped with the desired impact head, is pulled back up a track with a descending grade that ends at the rigid wall. Based on the desired test impact speed, the ram car is released from an appropriate distance up this track. The ram car accelerates under gravity, ideally reaching the desired impact speed at the instant of contact between the end of the impact head and the shell of the tank car being tested. The test setup is shown in Figure 2.

Table 1 provides a brief summary of each full-scale shell impact test, but does not explain all of the details that differed from test-to-test. The design of each specification tank car has been optimized based on the requirements and characteristics of the commodities they transport. Additionally, tank cars are not completely filled with lading, but have an intentional outage left...
at the top to allow for the commodity to expand en route. Different specification tank cars carrying different commodities will have different outage volumes. For example, the DOT-105 tank car discussed in this paper was designed to carry pressurized gases, while the DOT-117 tank car was designed to carry flammable liquids. The test setup used in each test reflected the operational conditions of the tank car, with the tested DOT-105 tank car having an internal pressure of 100 psig and approximately 10% outage, and the DOT-117 initially at atmospheric pressure with approximately 5% outage. As will be discussed later, the initial internal pressure affects the overall characteristic of the shell impact response.

2.1 Test Instrumentation

The instrumentation setup can vary slightly from test-to-test, depending on the details of the test, the tank car being tested, and the desired measurements. In general, each test includes instrumentation on both the initially-moving ram car and the initially-standing tank car. Tape switches are installed on the surface of the impact head and in the contact zone on the tank itself to allow the data acquisition systems on the ram car and the struck tank car to be synchronized to the time of impact. The instrumentation used in the DOT-105 and DOT-117 tests are summarized in Table 2.

<table>
<thead>
<tr>
<th>Type of Instrumentation</th>
<th>Channel Count</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DOT-105 Test</td>
</tr>
<tr>
<td>Accelerometers</td>
<td>11</td>
</tr>
<tr>
<td>Speed Sensors</td>
<td>2</td>
</tr>
<tr>
<td>Pressure Transducers</td>
<td>11</td>
</tr>
<tr>
<td>String Potentiometers</td>
<td>10</td>
</tr>
<tr>
<td>Total Data Channels</td>
<td>34</td>
</tr>
</tbody>
</table>

Test data was acquired using GMH Engineering Data BRICK Model III units. The data was anti-alias filtered at 1,735 Hz then sampled and recorded at a frequency of 12,800 Hz.

In addition to the measurements from the test instrumentation, one of the most readily apparent results of the test is whether the tank car punctured or resisted the impact without puncturing. This behavior, as well as the measured test data, are all candidates for inclusion in a procedure for model validation.

Details of the accelerometer, string potentiometer, and pressure transducer instrumentation used in the impact tests has been summarized previously [8] and is described in detail in the tank car side impact test reports [3][4][5][6][7].

3. FINITE ELEMENT ANALYSIS

Previous work by Tang et al. [10] focused on verification of FEA on tank car shell impacts using Abaqus/Explicit and LS-DYNA. In this study, Abaqus/Explicit commercial FE program version 6.14 [9] was used to simulate both the DOT-105 and DOT-117 shell impact tests [9].

This impact problem presents several challenges to simulation, each of which will affect the ultimate performance of the FE model and its suitability to simulate further impact conditions. The shell impact problem involves a dynamic impact with contact that evolves over time. The tank car shell will undergo elastic and plastic deformations, necessitating a material response for the steel shell that can adequately capture both behaviors. The model must also be capable of determining if puncture is likely to occur and if so, implementing a physically-realistic numerical representation of material failure. The tested cars featured fluid-structure interactions between the tank shell and two different fluid species, lading (water) and outage (air). Depending on the test conditions, it may also be necessary to represent the tank car’s pressure relief valve within the model, should tank deformation result in an adequate rise in outage pressure to begin to vent.

For each shell impact test, the pre-test FE model includes a combination of deformable and rigid parts. For the DOT-105 and DOT-117 tank car impact tests, half-symmetric FE models were used to reduce computation time. Based on previous test experience and the use of a half-symmetric model, the impactor is typically modeled as a rigid body having a point mass equal to half the total mass of the ram car with impact head. The backing wall and skids are also modeled as rigid bodies. The tank car and its two-phase contents are represented as deformable bodies. The overall setup is shown in Figure 3 for the DOT-105 shell impact FE model. This model used shell elements for most of the tank and jacket, with a patch of solid elements in the tank’s impact zone. Shell-to-solid coupling constraints were defined at the interfaces between the two different types of elements.

Figure 3. Annotated FE Model for DOT-105 Simulations (Insulation Hidden)
Material behaviors must be defined within each pre-test model for the steel tank, steel external jacket, water (lading), and air (outage). Details on the approaches used to model the material behaviors of the various materials in the DOT-105 and DOT-117 tank cars can be found in the respective test reports [3][4].

Because both the strength and ductility of the steels can have an effect on the puncture resistance of the tank car, understanding the material properties of the actual material of construction is an important element to consider in assessing the validity of a particular tank car model. The variation in material behaviors across different cars made of material meeting the same specification poses a challenge to using a particular material response to extrapolate to a tank car manufactured at a different time, by a different manufacturer, or using steel from a different origin, as the actual material properties for the tank car of interest can only be known by cutting and testing that particular material.

3.1 Validation of Finite Element Models

In a previous paper presented at the ASME 2018 V&V Symposium titled Validation of Puncture Simulations of Railroad Tank Cars Using Full-scale Impact Test Data [8], the authors discussed FE model validation methodologies currently applied to tank car side impact simulations as well as other crashworthiness simulations including: (1) passenger rail vehicles [11][12][13], (2) roadside hardware [14][15][16], (3) automobile [17][18], and (4) aircraft seating [19].

One of the key findings from the previous ASME V&V paper was that the qualitative comparisons currently used on time-history results can be highly subjective, i.e. two different engineers could disagree on the level of agreement between FEA results and test data. A literature review identified two software packages, RSVVP [16] and CORA [17], as potential tools for quantifying the level of agreement between time-history results. The standardized procedures from NCHRP 22-24 [15] and ISO/TS 18571:2014 [18] utilize RSVVP and CORA respectively. The current work applies the standardized procedures to compare FE results with test data from 14 signal channels in both the DOT-105 and DOT-117 side impact tests.

3.1.1 Validation Methodology

In this paper, the true curve (test data) is $T$, the CAE curve (FEA result) is $C$. For pre-processing, the FE results and test data were resampled by linear interpolation so that they had the same time step ($\Delta t$) of 0.08 ms (12,500 Hz). No pre-processing was done to time-shift or scale the signals. The test data was truncated to match the length ($n$) of the FE result. Accelerometer and pressure data were filtered by a CFC-60 filter per SAE J211 [20].

**RSVVP (NCHRP 22-24)**

NCHRP 22-24 recommends utilizing 4 metrics proposed by Sprague and Geers [21][22] to compare FEA and test results.

1. Magnitude (Geers $M_G$)
2. Phase (Sprague-Geers $P_{SG}$)
3. Comprehensive (Sprague-Geers $C_{SG}$)
4. Normalized Residual Error (ANOVA $\epsilon^r$)
5. Variance of Residual Error (ANOVA $\sigma^r$)

Per NCHRP 22-24, the point-to-point comparison ANOVA scores are intended to be applied to acceleration signals while the magnitude, phase, and comprehensive (MPC) scores are intended to be applied to velocity signals.

For verification of the FEA solution, NCHRP 22-24 recommends checking the following criteria to ensure that a numerically stable result was produced, i.e. conservation of mass, energy, momentum, etc.:

- **Total energy** (kinetic, potential, contact, etc.) must not vary more than 10% from the beginning of the run to the end of the run.
- **Hourglass energy** of the analysis solution at the end of the run is less than 5% of the total initial energy at the beginning of the run.
- The part/material with the highest amount of **hourglass energy** at the end of the run is less than 10% of the total internal energy of the part/material at the end of the run.
- **Mass added** to the total model is less than 5% of the total model mass at the beginning of the run.
- The part/material with the most **mass added** had less than 10% of its initial mass added.
- The moving parts/materials in the model have less than 5% of **mass added** to the initial moving mass of the model.
- There are no **shooting nodes** in the solution.
- There are no solid elements with negative **volumes**


ISO/TS 18571:2014 specifies a procedure for comparing FEA and test results using an overall score ($R$) consisting of:

1. Corridor Score ($Z$)
2. Phase ($E_p$)
3. Magnitude ($E_M$)
4. Slope or Shape ($E_S$)

Table 3 shows the standardized constants that can be used in a configuration file when running the CORA software.

<table>
<thead>
<tr>
<th>Corridor</th>
<th>Phase</th>
<th>Magnitude</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0$</td>
<td>$b_0$</td>
<td>$k_z$</td>
<td>$k_p$</td>
</tr>
<tr>
<td>0.05</td>
<td>0.5</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

ISO/TS 18571:2014 limits the metric to non-ambiguous signals, e.g. time-history curves. The metric has been previously applied to time-history signals from channel types such as force, moment, acceleration, velocity, and displacement. While CORA, can average multiple test signals (repeated tests), the metric should only be applied to a single test-FEA pair.

3.1.2 Phase Score

The objective of a phase score is to quantify the phase shift between the test data and FEA result.

The RSVVP Sprague-Geers phase score $P_{SG}$ is calculated as a normalized cross-correlation (i.e. not zeroed) between the two signals as shown in (1).
\[
P_{SG} = \frac{1}{\pi} \cos^{-1}\left( \frac{\sum_{i=1}^{n} T_i C_i}{\sqrt{\sum_{i=1}^{n} T_i^2 \cdot \sum_{i=1}^{n} C_i^2}} \right) \tag{1}
\]

The resulting normalized cross-correlation is inverted using \(\frac{1}{\pi} \cos^{-1}(\cdot)\) so that a perfect score is 0 and no correlation results in a score of 1. The signal is not iteratively shifted when calculating \(P_{SG}\) so a direct calculation of the amount of phase shift is not calculated.

In CORA, the phase score \(E_p\) is calculated by iteratively time-shifting the FEA result to the left and right by the time step \(\Delta t\) up to the maximum allowable fraction \(\varepsilon_p\) of the total time length. At each time-shift increment, the zero-normalized cross-correlation (ZNCC) of the two signals is computed. The term zero means that the mean of each signal is subtracted, and the term normalized means that the cross-correlation is divided by the standard deviation of each signal.

The number of increments needed to shift the FEA result to maximize the ZNCC is termed \(n_e\). If \(n_e = 0\) then the phase score is perfect \((E_p = 1)\); however, if it is greater than or equal to the maximum allowable increment shift \((\varepsilon_p \cdot n)\) then the score is 0. In between 0 and the maximum allowable time-shift, the score scales linearly.

### 3.1.3 Magnitude Score

The objective of a magnitude score is to compare the relative amplitudes of the test data and FEA result.

In RSVVP, the Geers magnitude score \(M_G\) is calculated by dividing the root mean square (RMS) value of the test data by the FEA result as shown in Equation (2).

\[
M_G = \frac{\sum_{i=1}^{n} T_i^2}{\sum_{i=1}^{n} C_i^2} - 1 \tag{2}
\]

While it is true that \(M_G\) is not sensitive to phase, it does not compare the signals on a point-to-point basis. This means that two signals could have completely different shapes but would receive a perfect score if they had the same RMS value. The acceptance criterion of the RSVVP magnitude score is set at 0.4 which means that the RMS value of the FEA result must be within ±40% of the RMS value of the test data.

CORA’s magnitude score \(E_M\) is computed by comparing the signals after performing dynamic time warping (DTW) on the optimally time-shifted FEA result (as described in Section 3.2.1 Phase Score). DTW is an algorithm used to compare the amplitudes of temporal signals which might have varying rates or pauses. A well known application of DTW is in speech recognition, where different speakers typically have different speaking rates, pauses, etc. making it difficult to directly compare a spoken word with words in a database even when the word is clearly spoken. In CORA’s implementation, DTW is governed by a set of rules where:

- Every time point from the FEA result is matched with one or more time points from the test data, and vice versa;
- The first time point from the FEA result is matched to the first time point from the test result but it can also be matched with subsequent time points, and vice versa;
- The last time point from the FEA result is matched with the last time point from the truncated test data but it can also be matched with prior time points, and vice versa; and;
- The time points for both signals must be monotonically increasing but each time step can individually expand or contract, i.e. dynamic warping

Figure 4 shows a schematic example of DTW in CORA using the acceleration time-history signals from the DOT-117 test. The red curve is the FE result which has been first time shifted to maximize the phase score \(E_p\) and then time warped. The black curve is the test data which has been truncated to match the length of the FE result and then time warped. Time dilation is clearly visible as flat responses in the red and black curves and is annotated on the curves. Contraction of the time warped signals is not clearly visible in this example, but it would be visible as a vertical line if time were contracted while acceleration varied significantly. There are thousands of instances of dilation and contraction in the curves that are not large enough to be visually apparent.

![Figure 4. Example of Dynamic Time Warping using DOT-117 Acceleration Time-History (CFC-60)](image)

CORA’s magnitude error \(\varepsilon_{mag}\) is calculated on the signals after DTW (\(C_l^{DTW}\) and \(T_l^{DTW}\)) using Equation (3).

\[
\varepsilon_{mag} = \frac{\sum_{i=1}^{n} |C_i^{DTW} - T_i^{DTW}|}{\sum_{i=1}^{n} |T_i^{DTW}|} \tag{3}
\]

The magnitude score \(E_M\) is calculated by linearly normalizing \(\varepsilon_{mag}\) to the maximum allowable magnitude error \(\varepsilon_M = 0.5\). This means that if on average the FEA result is off by 50% or more after DTW then it gets a score of \(E_M = 0\). Conversely, if \(\varepsilon_{mag}\) is 0 then the magnitude score is a perfect \(E_M = 1\).

### 3.1.4 Slope (Shape) Score

The objective of the slope score is to quantify the level of agreement in the overall shapes of the test data and FEA result. RSVVP does not currently have a slope (i.e. derivative) comparison implemented.

In CORA, the slope is compared by decimating (down-sampling) both signals by a factor of 10 to remove high frequency noise. It is important to note that ISO/TS 18571:2014...
recommends a sampling rate of 10kHz with a CFC-60 filter for acceleration data; this results in a 1 ms time step after decimation and a 100 Hz 3-dB limit frequency. The slopes of each curve are then calculated using a forward difference approximation with the new time step. The slope percent error \( \varepsilon_{\text{slope}} \) is computed with respect to the slope of the test data at each new time point.

The slope percent error is then averaged and linearly scaled/scored, as previously discussed for phase and magnitude, according to the threshold \( \varepsilon_{s}^{*} \) of 200%.

### 3.1.5 Corridor and (ANOVA) Scores

The objective of the corridor and Analysis of Variance (ANOVA) scores is to compare the test data and FEA result on a point-to-point basis. The corridor and ANOVA scores compare the curves at each time step, and because of this, they are extremely sensitive to distortions in phase (timing) between the signals, i.e. a minor distortion in phase can result in a very poor score. In CORA, inner and outer corridors are defined by shifting the test data vertically by \( \pm a_{0} \cdot T_{\text{norm}} \) and \( \pm b_{0} \cdot T_{\text{norm}} \) respectively. At each time point, the signal receives a score of 1 if it is within the inner corridor and a score of 0 if it is outside the outer corridor. If it is between inner and outer corridors, then the score is calculated as a normalized quadratic function between the two corridors.

CORA’s corridor score and RSVVP’s variance score are analogous in that they essentially score the mean squared error (MSE) of the FE result at each time point when comparing it to the test data (true curve).

#### 3.1.6 Overall (Composite) Score

The recommended composite score for magnitude and phase is calculated using Equation (8) in RSVVP.

\[
C_{SG} = \sqrt{M_{G}^{2} + P_{SG}^{2}} \tag{8}
\]

RSVVP does not calculate an overall score for magnitude, phase, composite, error, and variance. Rather, the acceptance criteria (see Table 4) are applied individually to each score, i.e. if any one score is outside the threshold then the FEA result is not validated per NCHRP 22-24. However, NCHRP 22-24 only requires ANOVA of acceleration signals and recommends calculating magnitude, phase, and composite scores for the integrated velocity signals.

#### 3.2 CORA and RSVVP Validation Results

Table 6 and Table 7 give summaries of the validation results for the analyses of the DOT-105 and DOT-117 tests, respectively. For ease of comparison, the tables are color coded so that green corresponds to a perfect score (CORA=1, RSVVP=0) and red corresponds to a poor CORA rating and an unacceptable score in RSVVP.

While NCHRP 22-24 only recommends applying the MPC scores to velocity and ANOVA scores to acceleration when using RSVVP, both the MPC and ANOVA scores were calculated for all 14 signals in this study for the purpose of discussion. The absolute value of the RSVVP scores are shown, and the scores met the acceptance criteria for the recommended channels.

While ISO/TS 18571:2014 does not recommend applying the ratings in Table 5 to each component of the overall score, the ratings have been applied individually for the purpose of discussion.

---

This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States. Approved for public release; distribution is unlimited.
Table 6. DOT-105 Post-test FEA Validation using CORA and RSVVP

<table>
<thead>
<tr>
<th>No.</th>
<th>Signals</th>
<th>Corridor</th>
<th>Cross-Correlation Rating</th>
<th>Overall</th>
<th>SPRAGUE-GEER</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Phase</td>
<td>Magnitude</td>
<td>Slope</td>
<td>STDEV</td>
</tr>
<tr>
<td>1</td>
<td>Impactor Acceleration</td>
<td>0.93</td>
<td>0.76</td>
<td>0.97</td>
<td>0.27</td>
<td>0.77</td>
</tr>
<tr>
<td>2</td>
<td>Impactor Change in Velocity</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.96</td>
<td>0.99</td>
</tr>
<tr>
<td>3</td>
<td>Impactor Displacement</td>
<td>1.00</td>
<td>0.90</td>
<td>1.00</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>4</td>
<td>Change in Air Pressure</td>
<td>0.87</td>
<td>0.68</td>
<td>0.89</td>
<td>0.53</td>
<td>0.81</td>
</tr>
<tr>
<td>5</td>
<td>String Pot 48” Offset A-End</td>
<td>0.87</td>
<td>0.68</td>
<td>0.89</td>
<td>0.53</td>
<td>0.77</td>
</tr>
<tr>
<td>6</td>
<td>String Pot 24” Offset A-End</td>
<td>1.00</td>
<td>0.90</td>
<td>1.00</td>
<td>0.81</td>
<td>0.94</td>
</tr>
<tr>
<td>7</td>
<td>String Pot 0” Offset</td>
<td>1.00</td>
<td>0.94</td>
<td>1.00</td>
<td>0.87</td>
<td>0.96</td>
</tr>
<tr>
<td>8</td>
<td>String Pot 24” Offset B-End</td>
<td>1.00</td>
<td>0.92</td>
<td>0.99</td>
<td>0.84</td>
<td>0.95</td>
</tr>
<tr>
<td>9</td>
<td>String Pot 48” Offset B-End</td>
<td>1.00</td>
<td>0.89</td>
<td>0.99</td>
<td>0.81</td>
<td>0.94</td>
</tr>
<tr>
<td>10</td>
<td>String Pot Vertical</td>
<td>0.99</td>
<td>0.86</td>
<td>0.99</td>
<td>0.68</td>
<td>0.90</td>
</tr>
<tr>
<td>11</td>
<td>String Pot Head A-End</td>
<td>0.87</td>
<td>0.62</td>
<td>0.97</td>
<td>0.93</td>
<td>0.85</td>
</tr>
<tr>
<td>12</td>
<td>String Pot Head B-End</td>
<td>0.83</td>
<td>0.63</td>
<td>1.00</td>
<td>0.93</td>
<td>0.85</td>
</tr>
<tr>
<td>13</td>
<td>String Pot Skid A-End</td>
<td>0.76</td>
<td>0.46</td>
<td>0.99</td>
<td>0.94</td>
<td>0.78</td>
</tr>
<tr>
<td>14</td>
<td>String Pot Skid B-End</td>
<td>0.70</td>
<td>0.44</td>
<td>0.99</td>
<td>0.75</td>
<td>0.71</td>
</tr>
</tbody>
</table>

Table 7. DOT-117 Post-test FEA Validation Using CORA and RSVVP

<table>
<thead>
<tr>
<th>No.</th>
<th>Signals</th>
<th>Corridor</th>
<th>Cross-Correlation Rating</th>
<th>Overall</th>
<th>SPRAGUE-GEER</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Phase</td>
<td>Magnitude</td>
<td>Slope</td>
<td>STDEV</td>
</tr>
<tr>
<td>1</td>
<td>Impactor Acceleration</td>
<td>0.91</td>
<td>0.83</td>
<td>0.97</td>
<td>0.51</td>
<td>0.83</td>
</tr>
<tr>
<td>2</td>
<td>Impactor Change in Velocity</td>
<td>1.00</td>
<td>0.91</td>
<td>1</td>
<td>0.95</td>
<td>0.97</td>
</tr>
<tr>
<td>3</td>
<td>Impactor Displacement</td>
<td>1.00</td>
<td>0.96</td>
<td>1</td>
<td>1</td>
<td>0.99</td>
</tr>
<tr>
<td>4</td>
<td>Change in Air Pressure</td>
<td>0.86</td>
<td>0.79</td>
<td>0.93</td>
<td>0.47</td>
<td>0.78</td>
</tr>
<tr>
<td>5</td>
<td>String Pot 48” Offset A-End</td>
<td>1.00</td>
<td>0.92</td>
<td>0.98</td>
<td>0.84</td>
<td>0.95</td>
</tr>
<tr>
<td>6</td>
<td>String Pot 24” Offset A-End</td>
<td>1.00</td>
<td>0.94</td>
<td>0.99</td>
<td>0.89</td>
<td>0.96</td>
</tr>
<tr>
<td>7</td>
<td>String Pot 0” Offset</td>
<td>1.00</td>
<td>0.9</td>
<td>0.98</td>
<td>0.89</td>
<td>0.96</td>
</tr>
<tr>
<td>8</td>
<td>String Pot 24” Offset B-End</td>
<td>1.00</td>
<td>0.97</td>
<td>0.98</td>
<td>0.89</td>
<td>0.97</td>
</tr>
<tr>
<td>9</td>
<td>String Pot 48” Offset B-End</td>
<td>0.99</td>
<td>0.96</td>
<td>0.97</td>
<td>0.87</td>
<td>0.96</td>
</tr>
<tr>
<td>10</td>
<td>String Pot Vertical</td>
<td>0.98</td>
<td>0.98</td>
<td>0.97</td>
<td>0.78</td>
<td>0.94</td>
</tr>
<tr>
<td>11</td>
<td>String Pot Head A-End</td>
<td>1.00</td>
<td>0.93</td>
<td>0.98</td>
<td>0.93</td>
<td>0.97</td>
</tr>
<tr>
<td>12</td>
<td>String Pot Head B-End</td>
<td>1.00</td>
<td>0.91</td>
<td>0.99</td>
<td>0.93</td>
<td>0.96</td>
</tr>
<tr>
<td>13</td>
<td>String Pot Skid A-End</td>
<td>0.99</td>
<td>0.8</td>
<td>0.97</td>
<td>0.92</td>
<td>0.93</td>
</tr>
<tr>
<td>14</td>
<td>String Pot Skid B-End</td>
<td>0.99</td>
<td>0.83</td>
<td>0.99</td>
<td>0.81</td>
<td>0.92</td>
</tr>
</tbody>
</table>
For all RSVVP and CORA scores in this paper, the calculations were done up to the point where the FEA terminated due to puncture. While the test data is shown beyond this point in Figure 5 through Figure 13, the test curve needed to be truncated for the calculations. Also, accelerations were converted to force using the estimated mass of the impactor (after verifying that it did not affect the scores) for ease of comparison.

Figure 5 and Figure 6 show the force time-histories for the tests and FEA of the DOT-105 and DOT-117, respectively, as well as the overall CORA scores.

The DOT-105 force-time history shows that the model was premature in predicting puncture by approximately 0.025 seconds. The premature puncture resulted in a fair phase score and a poor slope score while the magnitude score remained excellent. Surprisingly, the point-to-point comparisons (CORA corridor and RSVVP ANOVA) remained acceptable even with the disagreement after 0.2 seconds.

The DOT-117, which was not punctured in the test, showed overall good and acceptable scores until the model terminated after the impactor started to rebound. In both tests, the RSVVP MPC Sprague-Geers scores were found to be acceptable even though NCHRP 22-24 does not recommend applying the metrics to acceleration or force data.

Figure 7 and Figure 8 are time-history plots of the change in velocity of the impactor (ram car) for the test and FEA of the DOT-105 and DOT-117, respectively. To calculate the change in velocity, the acceleration signals were integrated starting at 0 seconds. Note that using this approach, the total change in impactor velocity can exceed the initial impact speed as the impactor eventually rebounds from the tank. The DOT-105 and DOT-117 models received nearly perfect CORA and RSVVP scores for change in impactor velocity and for impactor displacement (not shown).

The DOT-105 change in pressure was calculated by subtracting the initial 100 psig. The DOT-117 pressure already started at 0 psig so the gauge pressure and change in pressure were equivalent.
The CORA and RSVVP scores for change in air pressure were similar to the scores for acceleration (force). In all cases, the CORA slope scores were poor while the overall CORA scores were fair to good. The RSVVP ANOVA error score for the DOT-105 change in air pressure was unacceptable; however, the same score for the gauge pressure signal was well under the limit. This result clearly shows the importance of subtracting out the starting pressure (100 psig).

Figure 11 shows the string potentiometer located 48” towards the A-END (TD1Y) in the DOT-105 test which received a poor CORA slope score because the test data had an unusual amount of noise. It was observed that signal noise contributed to a poor slope score as sudden spikes in data caused unrealistically high slopes (calculated using the forward difference approximation on every 10 increments). Attempts were made to filter the data, but the noise was not fully eliminated and the raw signal was used.

Figure 12 shows a displacement time-history from the DOT-105 string potentiometer positioned vertically at the center of the tank. From this plot it is clear that the simulation which punctured and subsequently terminated prematurely did not capture a sudden increase in displacement which corresponds to ovalization of the tank. CORA and RSVVP gave the signal good scores, but qualitatively the signal response does not appear to be valid since it did not capture the change in slope after 0.2 seconds. To assess whether the model is capable of capturing this sudden increase in ovalization, a non-puncture model could be used which would be capable of calculating the response for the full time duration of the impact.

Figure 13 shows the DOT-105 string potentiometer on the B-end skid, which received the worst overall score in CORA and RSVVP. CORA’s magnitude score is nearly perfect while RSVVP’s magnitude score is halfway to failing. Qualitatively, the signals appear to be in agreement in magnitude. CORA’s phase score for the signal is poor while RSVVP’s phase score is less than a quarter of the way to failing. Qualitatively, the signals appear to be in disagreement in phase so it is expected that RSVVP’s magnitude score would be worse than the corresponding phase score. For point-to-point comparisons, CORA gave the signal a fair corridor score while RSVVP gave the signal an unacceptable ANOVA error score and an acceptable ANOVA variance score.

Many of the other channels measuring displacement received much higher scores than the two examples presented in this paper; however, because the scores were good to excellent, there were few things to discuss about the signals so Table 6 and Table 7 are used to summarize them.

### 3.3 Peak Value Percent Difference Comparison

A more basic approach to comparing model results with test measurements is to identify key measurements for comparison and determine the percent difference of the peak value calculated in the model with respect to the test. This approach is straightforward, and allows the entity making the comparison to
assess the level of agreement by means of a simple percentage. This is the approach used by the railroad industry-government Engineering Task Force’s Technical Criteria and Procedures report [12], with different target thresholds for different types of measurements. However, those thresholds are for quasi-static modeling and testing.

Further refinement of a magnitude comparison could be obtained by setting a target number of channels that should be compared, and a second target on the number of those channels that must be within the allowable difference between test and analysis. As an illustration of this approach, the peak measurements from the FEA and test measurements from the DOT-105 and DOT-117 tank car shell impacts are shown in Table 8. A color scale has been applied with 0% difference corresponding to green and 29% difference corresponding to red. The upper limit was chosen to correspond to a poor CORA score of 0.58 for magnitude $E_M$. If the analogous RSVVP magnitude score acceptance criterion were used, the red color would have been at 40%.

Table 8. Summary of Peak Measurements and FE Results

<table>
<thead>
<tr>
<th>Measurement</th>
<th>DOT-105</th>
<th>DOT-117</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact Force</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>Displacement at Peak Force</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>Peak Energy Absorbed</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>Outage Pressure</td>
<td>0.08</td>
<td>0.01</td>
</tr>
<tr>
<td>String Pot 48” Offset A-End</td>
<td>0.10</td>
<td>0.03</td>
</tr>
<tr>
<td>String Pot 24” Offset A-End</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>String Pot 0” Offset</td>
<td>0.01</td>
<td>0.04</td>
</tr>
<tr>
<td>String Pot 24” Offset B-End</td>
<td>0.03</td>
<td>0.04</td>
</tr>
<tr>
<td>String Pot 48” Offset B-End</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td>String Pot Vertical</td>
<td>0.33</td>
<td>0.06</td>
</tr>
<tr>
<td>String Pot Head A-End</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>String Pot Head B-End</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>String Pot Skid A-End</td>
<td>0.02</td>
<td>0.08</td>
</tr>
<tr>
<td>String Pot Skid B-End</td>
<td>0.03</td>
<td>0.03</td>
</tr>
</tbody>
</table>

4. DISCUSSION

Softwares such as CORA and RSVVP automate the process of windowing (truncating) and phase-shifting the signals and do not necessarily compare the signals over the full duration of interest in the impact event. In the case of the displacement measurement from the vertical string potentiometer (TD3Z) in the DOT-105 test (Figure 12), the scores did not compare the FE result with the test data after the FEA terminated (punctured) so the signal did not receive a penalty when a peak was present in the test signal after the FEA terminated. Using a simple percent difference comparison over a larger time range, the signal had a high percent difference (33%) and that made it stand out.

While comparing the peak value for each measurement in a test with a corresponding FE result enables an assessment of the percentage difference between the two values to be made quickly and compared with an agreed-upon threshold, this approach also has several drawbacks. For a complicated dynamic impact, this approach to validation can be overly simple and runs the risk of providing a false sense of agreement between measurement and FEA results when such agreement may not exist. Specifically, comparing only the peak values does not take into account whether the time or indentation at which the peak values occur also agree with one another.

A combination of automated software (e.g. CORA or RSVVP) and simplified comparisons of peak measurements might be best at assessing the validity of analysis results.

4.1 Energy Metrics

At the instant of impact, the kinetic energy of the initially-moving ram can be readily calculated using the velocity of the impactor and its mass. The energy imparted to the tank can be determined by first calculating the force-indentation response from the acceleration- and indentation-time histories for the ram car, and then numerically integrating the force-displacement response. As described in EN 15227 [10], the energy absorbed during a dynamic impact is often a relevant result to be compared between test and analysis.

For the tank car shell impact problem, the energy absorbed by the tank will either be equal to the initial kinetic energy (in the event of a non-puncture outcome), or will be less than the initial kinetic energy (if the tank punctures). The difference between the initial energy and the absorbed energy gives some indication of how close the speed of the impactor was to the critical threshold puncture speed; however, the relationship is non-linear and is complicated by fluid effects, e.g. sloshing.

4.2 Fluid Response

As demonstrated by the test results presented earlier in this paper, the fluid behavior inside the tank car can have a significant effect on the overall impact response of the tank car. The DOT-105 and DOT-117 tests exhibited differently-shaped force-time histories owing at least partially to the DOT-105 tank car being initially pressurized and having a larger outage than the DOT-117 tank car.

One of the challenges associated with using test data to validate a model that will then be extrapolated to impact conditions beyond those tested is how valid the model remains as the conditions differ more significantly from what was tested. Because of the differences observed in pressurized and non-pressurized tank car responses, modeling techniques that are appropriate to model one state of pressure may not be equally suitable to model the other state of pressure. Significant changes to the impact setup, such as impactors of a different shape or size, may also have an effect on the relative significance of the fluid behavior. A validation framework may need to take into account the nature of the test and analysis initially used to validate the tank car shell impact model, and determine appropriate limits on the nature of the changes for which that model remains valid.

4.3 Puncture/Non-puncture Outcome

During the impact test itself, one of the readily apparent outcomes is whether the tank has punctured, or if the tank
resisted the impact without puncturing. If an FE model is used to simulate a test with the potential of a puncture outcome, the model must also be capable of simulating puncture. The quality of the puncture simulation will depend on such details as the availability of material coupons for testing, the ability of the failure/fracture behavior of the material to be characterized based upon those material tests, and the ability of the FE software to numerically implement the failure/fracture behavior.

Further, it may be appropriate to consider puncture not as a binary outcome, but to take into account the character of the puncture in assessing the performance of the model. The RSVVP and CORA frameworks were not developed to specifically evaluate puncture as a mode of failure between a simulation and a test. Future work may be appropriate to consider including a specific qualitative and/or quantitative score to account for the puncture response as a specific feature of the tank car shell impact scenario.

5. CONCLUDING REMARKS

FRA has sponsored a series of tests and corresponding FE analyses of shell impact tests of fluid-filled railroad tanks cars. While the test measurements have been compared with the results of corresponding FE models, specific model validation procedures have not yet been adopted for validating tank car shell impact models. As a starting point at choosing both the behaviors to be compared and the threshold for determining suitable agreement for validation, existing model validation criteria and procedures used by other segments of the transportation industry have been reviewed and applied to selected tank car models (DOT-105 and DOT-117). Several promising approaches to validation appear to be well-suited to the complexities of the tank car shell impact problem.

It should be noted that the simulations and test executions for both of the tank car tests discussed in this paper were successful in accomplishing their objectives, i.e. puncture was not observed for either of the tank car tests discussed in this paper. The testing described in this paper was performed by Nicholas Wilson and Przemek Rakoczy, Travis Gorhum, Shawn Trevithick, and Nicholas Wilson lead the testing effort. The authors also wish to acknowledge the contributions of Dr. David Jeong (retired) of the Volpe Center to the overall success of the tank car structural integrity research program.

ACKNOWLEDGEMENTS

The testing described in this paper was performed by Transportation Technology Center, Inc., in Pueblo, Colorado. Przemek Rakoczy, Travis Gorhum, Shawn Trevithick, and Nicholas Wilson lead the testing effort. The authors also wish to acknowledge the contributions of Dr. David Jeong (retired) of the Volpe Center to the overall success of the tank car structural integrity research program.

REFERENCES


This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States. Approved for public release; distribution is unlimited.


