DOCKET: EPA-HQ-OAR-2014-0609

TECHNICAL SUPPORT DOCUMENT FOR SECTION 194.23
EPA REVIEW OF PROPOSED MODIFICATION TO THE
WASTE SHEAR STRENGTH PARAMETER TAUFAIL

Contract Number: EP-D-10-042
Work Assignment No. 5-12

Prepared for:

U.S. Environmental Protection Agency
Office of Radiation and Indoor Air
Center for Waste Management and Regulations
1200 Pennsylvania Ave., N.W. Washington, DC 20460

Kathleen Economy
Work Assignment Manager

Prepared by:

S. Cohen & Associates
1608 Spring Hill Road Suite 400
Vienna, Virginia 22182

July 2017

PREFACE

The U.S. Department of Energy (DOE) is required to submit a Compliance Recertification Application (CRA) to the U.S. Environmental Protection Agency (EPA) for the Waste Isolation Pilot Plant (WIPP) facility every five years including an updated assessment of future WIPP performance. During EPA's review of DOE's CRA-2014 performance assessment (PA), events associated with the February 2014 repository fire and radionuclide release have closed portions of the underground facility. This temporary closure has created a situation where certain parts of the underground facility could not be accessed for ground control. Panel 9 may be abandoned along with plans to modify the design change used for panel closures in panels 3, 4, 5 and 6.

Because the CRA PAs are predictions of post-closure repository performance and the EPA knows there will be modifications to the current repository design, modifying the CRA-2014 PA at this time to incorporate alternative parameter values would not add more reality to predictions of repository post-closure performance. Consequently, the EPA adopted the CRA-2014 PA as originally submitted by DOE as the baseline, rather than have DOE conduct a revised PA baseline calculation (PABC). In lieu of requesting a PABC-2014, the EPA requested that DOE conduct a set of sensitivity studies to address some of the significant technical concerns arising from the EPA's CRA-2014 review. The inputs to these sensitivity studies broadly address many of the EPA's technical concerns that could potentially impact long-term repository performance. The Agency has reviewed the results of these studies and determined that there exists an adequate level of confidence—that is, a reasonable expectation—that the repository will continue to comply with EPA regulations. These results are addressed in a companion Technical Support Document, "Review of EPA Sensitivity Studies of DOE CRA-2014 WIPP Compliance Recertification Performance Assessment."

Additionally, the EPA recommends further work that can be conducted to evaluate many of the technical concerns identified in the EPA's review of the CRA-2014 PA, as well as incorporate future repository design changes. The EPA will work with DOE to determine the best path forward for resolution of EPA's concerns, which could include additional data reviews, independent technical reviews, and possibly additional sensitivity analyses to reach a consensus for the next CRA. It is anticipated that the results of these efforts will be incorporated into the CRA-2019 PA or otherwise be made available during EPA's review of the CRA-2019 PA.

This Technical Support Document (TSD) addresses the Agency's technical review of the parameter value DOE adopted to represent the waste shear strength DOE's CRA-2014 PA.

EXECUTIVE SUMMARY

This report documents the U.S. Environmental Protection Agency's (EPA) review of the U.S. Department of Energy's (DOE) revised waste shear strength lower uncertainty boundary used in the CRA-2014 performance assessment (PA). The shear strength of the waste is used to calculate the amount of waste eroded by flowing drilling mud from the sidewall of a borehole that intersects the WIPP repository. Because the eroded waste reaches the ground surface, waste shear strength is an important parameter used in calculating potential repository releases. At the time of the initial WIPP certification it was acknowledged that the lower bound of the shear strength uncertainty range (0.05 Pa) was likely too low but was also likely a bounding value because it assumed complete waste degradation. More detailed knowledge now available from WIPP PA modeling results shows that complete waste degradation is expected to be rare in the anoxic WIPP subsurface environment during the 10,000-year regulatory time frame. A higher bounding value for the low end of the range would therefore be warranted and DOE authorized Sandia National Laboratories (SNL) to conduct a series of laboratory tests to provide a more appropriate value.

SNL designed and built a flume that specifically simulates the sidewall erosion resulting from vertical flow up a borehole. In SNL's flume, the eroding fluid enters a vertical channel from the bottom and flows up past a specimen of surrogate waste material held in a cylindrical sample holder in the sidewall of that channel. Selection of surrogate materials for degraded waste considered inventory, underground conditions, and theoretical and experimental results. The surrogate waste material was manually advanced into the vertical flume channel at a rate that corresponded to the rate of erosion. The shear strength is defined as the applied shear stress at the time that erosion begins.

Based on the result of the tests, DOE proposed increasing the lower bound of the waste shear strength uncertainty range from 0.05 Pa to 2.22 Pa and used this revised value in the CRA-2014 PA. The revised value was the arithmetic mean of a series of five tests on surrogate waste samples representing 50% degraded waste. EPA accepted these tests as appropriate for determining a lower bound value but questioned the use of the mean test result as representative of a lower bound when some samples in the series resulted in even lower values. EPA requested DOE to instead establish a lower bound equal to the lowest measured shear strength of 1.60 Pa for the five samples and DOE accepted this change. EPA does not expect this change to materially affect PA results because decreasing the lower bound from 2.22 Pa to 1.60 Pa does not significantly change the range of the shear strength distribution which has an upper bound of 77 Pa. EPA requested this change to provide consistency with the principle that in a bounding uncertainty analysis the lower bound should have a value that is reasonable and even lower values should not be expected. EPA disagreed with DOE's proposed use of the mean result of five tests because even lower values had been measured and could therefore be expected. EPA accepts the use of DOE's mean result in the CRA-2014 PA but expects DOE to use the lower bound of 1.60 Pa in the CRA-2019 PA.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

CCA	Compliance Certification Application
CFR	Code of Federal Regulations
CPR	Cellulosics, Plastics, and Rubber
CRA	Compliance Recertification Application
CTAC	Carlsbad Area Office Technical Assistance Contractor
DOE	U.S. Department of Energy
EPA	Environmental Protection Agency
PA	Performance Assessment
PABC	Performance Assessment Baseline Calculation
SEN	Sensitivity
SNL	Sandia National Laboratories
TRU	Transuranic
WIPP	Waste Isolation Pilot Plant

1.0 INTRODUCTION

The U.S. Department of Energy (DOE) developed the Waste Isolation Pilot Plant (WIPP) repository for the permanent disposal of transuranic (TRU) waste. The repository is located in deeply buried deposits of bedded salt in the Salado Formation in southeastern New Mexico. The U.S. Environmental Protection Agency (EPA or the Agency) regulates containment of TRU waste at WIPP in accordance with the radioactive waste disposal standards at Code of Federal Regulations (CFR) Title 40, Parts 191 and 194. EPA first certified the WIPP as complying with these standards and approved it for TRU waste disposal in 1998. The regulations require recertification of WIPP at five year intervals following the first waste shipment in 1999, with the most recent recertification occurring in 2014. EPA's decision to recertify WIPP is based in part on the results of an assessment of the projected ability of the facility to meet the Agency's waste isolation standards over the 10,000-year post-closure regulatory time frame. The ability to meet these standards is determined by the results of numerical modeling conducted for the DOE by Sandia National Laboratories (SNL). This modeling simulates the repository's future performance in a process called Performance Assessment (PA). The most recent assessment was included in DOE's 2014 Compliance Recertification Application (CRA) and is called the CRA-2014 PA.

This report documents EPA's review of the revised lower boundary DOE has assigned for the range of uncertainty in waste shear strength in the CRA-2014 PA. Waste shear strength is an important parameter used in calculating potential releases of waste materials from the WIPP repository. Drilling mud flowing up a borehole applies a hydrodynamic shear stress to the borehole wall that could result in erosion of the wall material. If a future exploratory borehole intersects a WIPP waste panel and passes through consolidated waste, erosion of the waste at the borehole-waste interface could occur and may result in the transport of solid waste materials to the ground surface entrained in the drilling mud. As a consequence, radionuclides may be released to the accessible environment. The released waste materials are called *cavings* in WIPP PA and this erosion process is included as a potential radionuclide release mechanism in WIPP performance modeling. Cavings releases are important components of total releases in WIPP PA because they can occur in any intersecting borehole independent of gas pressure and brine saturation. Waste shear strength is a key parameter for calculating the quantity of waste released as cavings and is identified in WIPP PA as the parameter BOREHOLE:TAUFAIL. In this review, this parameter will more simply be called TAUFAIL.

Over time, competing processes within the WIPP repository will tend to both decrease and increase the waste shear strength. Degradation due to chemical and biological processes will tend to weaken the waste while other processes such as compaction and cementation will tend to strengthen it. As discussed below, the focus of DOE's shear strength analysis was on establishing a more realistic lower bound to the uncertainty boundaries of this parameter and therefore on the processes that would tend to weaken the waste. The strengthening effects of compaction were considered in DOE's analysis but other processes that could strengthen the waste were conservatively not included.

The rate of degradation and loss of strength depend on the type of waste and the nature of the degradation processes, both of which are variable. The amount of strength loss is therefore uncertain. To address this uncertainty and in view of its importance to WIPP performance, TAUFAIL was sampled in previous PAs as an uncertain parameter with a log-uniform distribution that ranged between 0.05 to 77 Pa. The range of this distribution was intended to be bounding. The upper end of this range was based on particle size distributions determined by an expert elicitation panel assuming that the degradation of waste is limited (CTAC 1997). The lower end, which is of particular interest in this review, assumed that waste degradation is complete and was based on the hydrodynamic shear strength of San Francisco bay mud (Partheniades and Paaswell 1970). Bay mud is a natural material with a low shear strength and was used as a surrogate to represent completely degraded WIPP waste. This distribution has remained unchanged from the time of the initial 1999 WIPP Compliance Certification Application (CCA) and was also used in the subsequent 2004 and 2009 Compliance Recertification Applications (CRAs).

2.0 BACKGROUND

At the time of the initial WIPP certification a peer review acknowledged that the low end of the shear strength range (0.05 Pa) was likely too low but was also likely a bounding value as intended (Wilson et al. 1997). More detailed knowledge now available from WIPP PA modeling results has shown that complete waste degradation is expected to be rare in the anoxic WIPP subsurface environment during the 10,000-year regulatory time frame. This is true for iron-based metallic wastes, which constitute a large fraction of the total metallic waste volume, as well as organic wastes consisting of cellulosics, plastics, and rubber (CPR). The predicted presence of undegraded iron and organic wastes at the end of the regulatory time frame indicates that complete degradation is unlikely to occur. Examples of relevant modeling results supporting this conclusion were provided by DOE in response to EPA completeness questions, and are discussed in Section 4.2 below. San Francisco bay mud may therefore not be an appropriate surrogate for establishing the low end of the uncertainty range for TAUFAIL.

To provide an updated lower bound for the range of values for TAUFAIL, the DOE requested Sandia National Laboratories (SNL) to perform experimental laboratory flume tests to better determine the hydrodynamic shear strength of surrogate WIPP wastes representing varying degrees of degradation. These tests, their results, and the proposed application of those results to the WIPP PA are described by Herrick et al. (2012). The supporting information for this review was taken principally from Herrick et al. (2012), the Agency's onsite review of SNL's laboratory test procedures and equipment, and the DOE's responses to EPA completeness questions on the CRA-2014 PA regarding this parameter.

3.0 SANDIA NATIONAL LABORATORIES SHEAR STRENGTH LABORATORY TESTS

3.1 Test Equipment and Conduct

This section provides a summary overview of the testing program performed by SNL to determine the hydrodynamic shear strength of degraded WIPP wastes. A comprehensive description of the program is presented in Herrick et al. (2012). Herrick et al. (2012, p. 1) state that hydrodynamic shear strength can only be measured in the laboratory by flume testing. SNL designed and built a flume for those tests that specifically simulates the sidewall erosion resulting from vertical flow up a borehole. In SNL's flume, the eroding fluid enters a vertical channel from the bottom and flows up past a specimen of surrogate waste material held in a cylindrical sample holder in the sidewall of that channel. The surrogate waste material was manually advanced into the vertical flume channel at a rate that corresponded to the rate of erosion. A photograph of SNL's test equipment is presented in Figure 1.

The surrogate samples of degraded WIPP waste used in SNL's flume tests were prepared following an approach developed by Hansen et al. (1997). The approach was based on consideration of the anticipated future state of the waste. Selection of surrogate materials for degraded waste considered inventory, underground conditions, and theoretical and experimental results (Hansen et al. 1997; Hansen 2005). The degraded surrogate for each waste constituent was individually considered. The potential for the waste to be strengthened due to cementation, mineral precipitation, and more durable packaging was not included when developing the surrogates. Consequently, Hansen et al. considered their surrogate degraded waste to represent the most highly degraded, plausible future state of the waste. Surrogate waste formulations were prepared to represent degradation states of 50% and 100% degraded waste by weight (Hansen 2005, Hansen et al. 2003). In the 50% degradation state half the waste was represented as fully degraded and the other half as completely undegraded. In the 100% degradation state all the waste was represented as fully degraded.

The materials used to create the surrogate WIPP waste for SNL's flume tests are described in Herrick et al. (2012, Table 4). The uncorroded metallic components consisted of strips of steel sheet metal, small nails (cut up), and scraps of steel or iron. The corroded metallic components included scrapings from rusted steel or iron, Fe(III)O.OH in the form of goethite or limonite rock samples, and crushed sand- to silt-sized particles. The undegraded CPR components included finely shredded paper, snipped cotton balls, sawdust, rubber bands, shredded plastic grocery bags, and peat. Other waste components included broken glassware representing other inorganic materials, broken concrete and mortar to represent the cement used to solidify the waste, natural soil to represent soils, and corrosion-induced salt precipitates. The degraded CPR components were represented by increased amounts of corrosion-induced salt precipitates and elimination of all undegraded CPR.

SNL's flume tests consisted of the following cases:

- A 50% case where half of the iron is corroded and half of the cellulosics, plastics, and rubber are degraded, per Hansen et al.'s (1997) approach;
- A 100% case where all of the iron is corroded and all cellulosics, plastics, and rubber are degraded, per Hansen et al.'s (1997) approach;
- An additional 75% case where three quarters of the iron is corroded and three quarters of the cellulosics, plastics, and rubber are degraded, per the methodology developed by Papenguth and Myers and cited in Hansen et al. (1997, Appendix A).

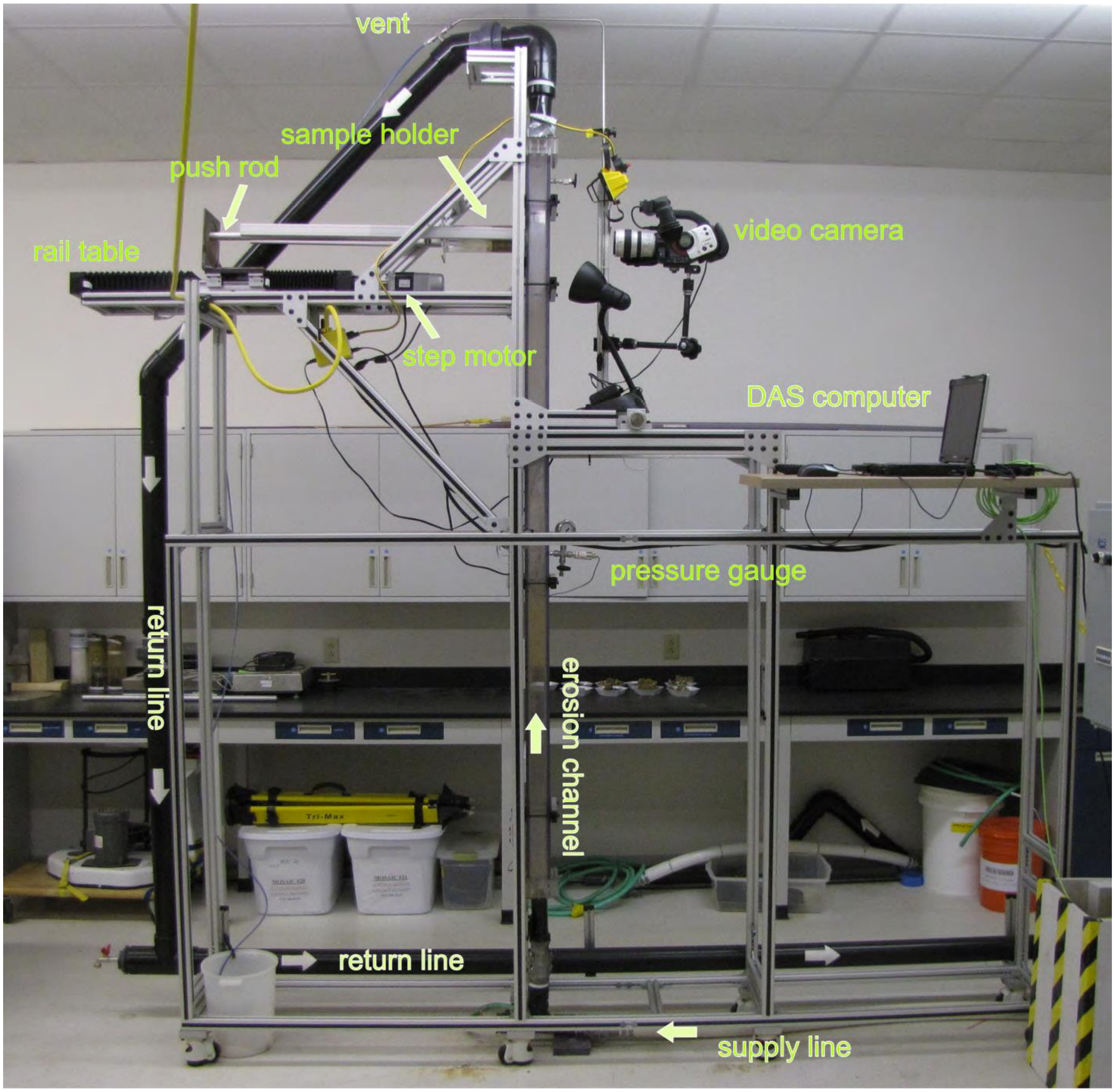

Figure 1. Photograph of SNL's test flume equipment (from Herrick et al. 2012, Figure 2).

Surrogate waste materials were compacted to 2.3 and 5.0 MPa prior to testing. A stress of 2.3 MPa was identified in modeling performed by Herrick et al. (2007) as the minimum pressure the degraded waste would be subjected to, which corresponds to the highest gas generation rate. Alternatively, a stress of 5.0 MPa was identified by Hansen et al. (1997, 2003) as a conservative estimate for the lower bound of pressure that the waste will undergo in the underground.

The flume tests were performed using stepwise increases in the shear stress on the sample. These increases were accomplished by increasing the flow rate in the vertical flume and measuring the applied shear stress and associated erosion rate at each step. The erosion rate was determined by manually advancing the test specimen into the vertical flume such that the sample face remained flush with the wall of the flume. This rate was governed by visual inspection of the eroding sample face in the flume. The results were analyzed using three different methods to estimate the shear stress required to initiate erosion (Herrick et al. 2012, p. 27). This stress is called the critical shear stress and is equal to the shear strength of the test specimen. Although the results of the three methods were often similar, Herrick et al. (2012, p. 79) concluded that the bilinear method originally proposed by Parchure and Mehta (1985) was most applicable to their data. Herrick et al. refer to the bilinear method as the University of Florida or Mehta method.

In the bilinear method, the critical shear stress is the applied stress at the inflection point where erosion begins and is determined by projecting the erosion rate data down to the zero rate axis. Although statistical curve fitting techniques were applied by SNL in this process, judgement is still required to interpret the scatter and requires a knowledge of the characteristics of the surrogate waste and the testing methodology. Examples of SNL's application of the bilinear method to test data are provided in Figures 2a and 2b. The Agency's concerns regarding test interpretation and the role of judgement are discussed in Section 4.1 below.

3.2 Test Results and SNL Recommendations

The flume tests were performed on three different waste surrogates at two different pre-compaction pressures. Between three and seven tests were performed on each combination of surrogate and pre-compaction pressure. The first tests were performed on the weakest, 100% degraded samples; however, problems were encountered with the testing apparatus that invalidated the results of all but one of the tests. Using the bilinear method, the single successful test on a 100% degraded sample yielded a shear strength of 0.17 Pa for a sample compacted to 2.3 MPa (Herrick et al. 2012, Figure 20). However, this sample may have been disturbed during testing and Herrick et al. (2012, p. 37) warn that the result may not be representative. Testing the 75% degraded samples benefited from the refinements employed in the previous tests and all such tests yielded useful results. The 50% degraded samples were tested at the same time as the 100% degraded samples and suffered some of the same problems. However, based on visual inspection SNL concluded that they did not exhibit the same damage as the 100% samples and accepted all results (Herrick et al. 2012, Section 6.4). The critical shear stress was greater than the maximum applied stress of SNL's testing apparatus for two of the 50% samples pre-compacted to 5.0 MPa and bulk erosion did not occur. For those two samples SNL assigned a shear strength equal to the maximum shear stress of 5.69 Pa that their testing apparatus could apply. The results of SNL's flume tests are summarized in Table 1.

Table 1. SNL Flume Test Results Using the UF Bilinear Fit Method

Degree of Degradation	Shear Strength (Pa)			
	Pre-compaction to 2.3 MPa		Pre-compaction to 5.0 MPa	
	Mean	Range	Mean	Range
100%	0.17	n/a	n/a	n/a
75%	1.53	1.06 - 2.00	2.17	1.46 - 2.80
50%	2.22	1.60 - 3.09	5.05	3.79 - 5.69

Source: Herrick et al. 2012

Herrick et al. (2012, p. 80) initially recommended using the arithmetic mean result of 5.05 Pa for the 50% degraded waste surrogate and a pre-compaction pressure of 5.0 MPa as the lower bound of the TAUFAIL distribution in WIPP PA. This recommendation was based on the following reasoning: Hansen et al. (1997, 2003) showed that for most performance assessment calculations, half or more of the initial iron and CPR inventory remains after 10,000 years; 5.0 MPa was identified by Hansen et al. (1997, 2003) as a conservative estimate for the lower bound of pressure that the waste will undergo in the underground; and the 50% degraded surrogate waste material pre-compacted to 5.0 MPa was accepted by the Spallings Conceptual Model Peer Review Panel (Yew et al. 2003) and by the EPA for use in establishing parameters for the spallings model.

Herrick et al. (2012) presented their recommendations in an EPA/DOE technical exchange meeting in November, 2012. During that meeting EPA requested DOE to provide additional explanation for selecting a pre-compaction stress of 5.0 MPa as appropriate for determining a lower bound for the distribution of TAUFAIL. In a further evaluation, documented by Herrick and Kirchner (2013), SNL concluded that Hansen et al. (1997) had apparently taken the waste porosity results from the CCA BRAGFLO calculations and calculated the minimum vertical stress necessary to produce the deformation of a drum stack consistent with that porosity. When applying a similar approach to the waste porosity results from the more recent CRA-2009 BRAGFLO calculations, Herrick and Kirchner (2013, p. 5) estimated vertical stresses that were fairly consistent at 4.3 to 4.4 MPa but were somewhat lower than the pre-compaction pressure of 5.0 MPa recommended by Hansen et al. (1997). In view of the direct correlation between increasing pre-compaction pressure and increasing waste shear strength, and also of the need to establish a defensible lower bound for TAUFAIL, Herrick and Kirchner (2013, p. 5) modified their recommendation to instead use the lower average shear strength value of 2.22 Pa from the 50% degraded experimental samples compacted at 2.3 MPa as the lower bound for a uniform distribution of TAUFAIL in WIPP PA. These recommendations were adopted by DOE and used in the CRA-2014 PA.

4.0 EPA EVALUATION OF SNL's SHEAR STRENGTH TESTS

4.1 Evaluation of SNL Test Equipment and Conduct

EPA's evaluation of SNL's flume tests is based on detailed documentation of the test equipment, its design considerations, and the methodology adopted for conducting the tests (Herrick et al.

2012); onsite examination of the test equipment and testing environment; and discussions with SNL's lead investigator Courtney Herrick. EPA believes that the test equipment was carefully conceived and well documented. Given that the previous history of laboratory hydrodynamic shear strength measurements involved horizontal fluid movement over horizontal beds, SNL's adaptation to vertical fluid movement was innovative and it is not surprising that the need for refinements to the test equipment and procedures became evident in the early part of the program. SNL's success in this refinement is apparent in the differences between the results of the early tests on the 50% and 100% degraded samples and the later tests on the 75% degraded samples that were performed after the refinements were accomplished. Examples of test results on 50% and 75% degraded samples are provided for comparison in Figures 2a and 2b. The color-coded data points on the figures correspond to their use in the linear regression analyses used to interpret the test results.

Figure 2a shows the results of testing a 50% degraded, 2.3 MPa pre-compacted sample before the equipment and procedural refinements were completed. The results show a considerable scatter that increases uncertainty in how the results should be interpreted. These results were typical of three out of the five tests performed on 50% degraded, 2.3 MPa pre-compacted samples.

Figure 2b shows the results of testing a 75% degraded, 2.3 MPa pre-compacted sample after the equipment and procedural refinements were completed. The results in Figure 2b show very little scatter and the bilinear fit methodology was applied with textbook clarity. These results were typical of four out of the seven tests performed on 75% degraded, 2.3 MPa pre-compacted samples.

Based on test results of the type illustrated in Figure 2b, the Agency concluded that the equipment and procedural refinements made by SNL were successful. As previously noted, the tests on the 50% degraded samples were conducted before these refinements were completed. Those tests were accepted by SNL despite the difficulties in interpretation because the test specimens did not appear to exhibit the same damage as the 100% degraded samples. However, the differences in the quality of test results before and after the equipment and procedural refinements were completed prompted EPA to perform a detailed review of the 50% degradation tests. This concern is further addressed in the Agency's evaluation of SNL's test procedures, results and recommendations in Section 4.

Also as previously noted, the erosion rate (the vertical axis in Figures 2a and 2b) was determined by the manually controlled rate of advancement of the sample into the vertical flume based on visual inspection of the eroding sample face. In recognizing potential difficulties in deciding when to advance the sample based on a visual interpretation of the degree of erosion of a highly heterogeneous surrogate waste material during testing, EPA submitted the following completeness comment to DOE:

> **EPA Comment 1-23-5 Waste Shear Strength.** *Provide and justify the criteria used in advancing the surrogate waste samples during the shear strength tests when the eroded sample face was not smooth but irregular.*

DOE's response, documented in its 2nd Response Submittal to the EPA (DOE 2015, Enclosure 1), stated in summary that the surface of the sample is initially positioned flush with the flume channel wall. The eroding fluid is forced upward through the enclosed channel across the surface of the sample, producing a shear stress across that surface that may cause the sample to erode.

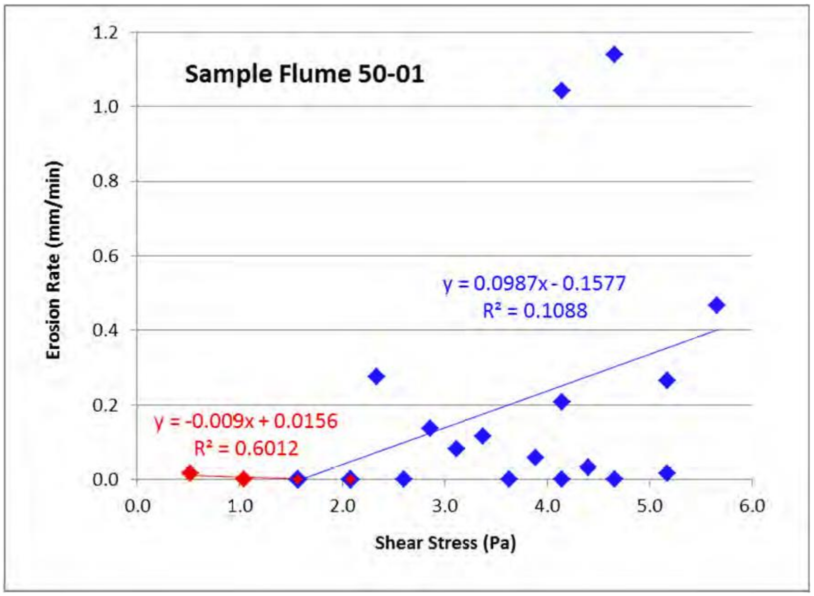

Figure 2a. Bilinear analysis of test results for Sample Flume 50-01 yielding a critical shear stress of 1.60 Pa. Test was performed on 50% degraded material with a pre-compaction pressure of 2.3 MPa. (Source: Herrick et al. 2012, Figure 37a).

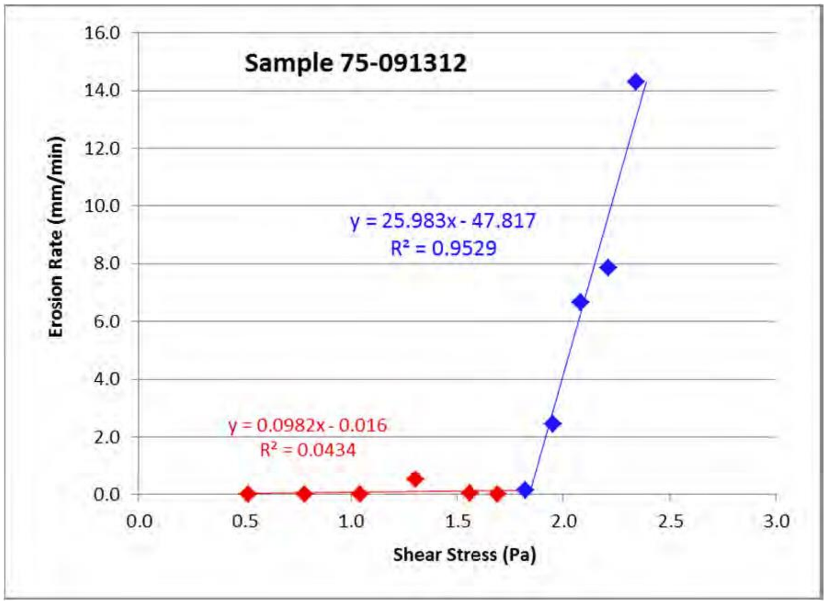

Figure 2b. Bilinear analysis of test results for Sample 75-091312 yielding a critical shear stress of 1.84 Pa. Test was performed on 75% degraded material with a pre-compaction pressure of 2.3 MPa. (Source: Herrick et al. 2012, Figure 27a).

Erosion is determined visually. If erosion is observed, additional sample material is advanced laterally by the operator until the surface of the sample is once again flush with the channel wall. Sample advancement is continued as needed to keep pace with erosion of the sample. The erosion rate is determined from the lateral movement of the sample in the sample container over time. During erosion, the surface of the sample may at times become irregular or pitted due to differential erosion rates that may not be consistent over the entire face of the sample. Because of this tendency, erosion of the leading edge is used to drive the operator's decision to move the sample forward. The leading edge is the lower edge of the sample that is first contacted by the fluid flowing up the channel. The erosion channel and sample holder are identified on Figure 1.

DOE justified this methodology by stating, in summary, that the leading edge is used to decide to advance the sample because the shear stresses acting on this edge are developed over the long, smooth-walled section of the flume channel below the sample. This allows the operator to analytically estimate the shear stresses in advance of the tests, with final corrections to the shear stress calculations being made post-test once all the testing conditions have been measured. An additional advantage of establishing the leading edge as the specific point used by the operator to decide whether to advance the specimen is that it establishes a consistent methodology. This consistent method of evaluation reduces the degree of human judgment that could influence the results.

Relying only on erosion of the small area of the leading edge to determine when to advance the test specimen is a limitation of the test methodology that can lead to increased scatter in test results due to the inherent inhomogeneity of the degraded and undegraded components of the degraded surrogate waste. More scatter would be expected for a 50% degraded surrogate waste than for a 75% or 100% degraded surrogate because of its greater degree of inhomogeneity and the test results support this expectation. The results are representative of only that small part of the test sample at the leading edge where the material can vary from undegraded (and zero erosion) to completely degraded (and rapid erosion). An example of such a wide range of results is illustrated in Figure 2a. Given the adopted leading edge test methodology, SNL has appropriately averaged the scatter in shear test results to approximate what would be seen in a release from a borehole that penetrates the waste area of the repository. Although individual waste containers will vary in type of waste and degree of degradation, the boreholes will typically penetrate stacks of three waste drums since this is the most common type of container in the repository. The subsequent release will therefore represent average degradation conditions across all three drums rather than the conditions in a small part of a single drum. Averaging the results of individual tests by applying a linear fit to the data is therefore appropriate.

Although the Agency believes that a better approach could have been to reduce the scatter by basing the advance rate on the average erosion of the entire specimen face, this would require an additional level of judgement that could be difficult to consistently apply. The Agency also considers that the scatter in test results could in part be due to sample damage that was not evident in SNL's visual examination but that the primary source of the scatter was the heterogeneity of the test samples. The Agency believes that the test results do provide a better basis than San Francisco bay mud for establishing a lower bound for TAUFAIL. The Agency concludes that, given the constraints involved, the testing methodology is acceptable.

4.2 Evaluation of SNL Test Results and Recommendations

As previously noted, DOE proposed using the average shear strength value of 2.22 Pa from the 50% degraded experimental samples pre-compacted at 2.3 MPa as the lower bound for a uniform distribution of TAUFAIL in WIPP PA. The Agency's evaluation of this proposal is described below.

Degraded Waste Surrogates. The materials used to prepare the surrogate waste samples are described in Section 3.1. The surrogate degraded waste formulations of the type used in SNL's flume tests were originally developed by Hansen et al. (1997) in support of a peer review of the tensile failure aspects of the WIPP spallings model used in DOE's 1996 CCA. The surrogate formulations and results of tensile tests on those surrogates were reviewed by the Conceptual Models Peer Review Panel and found to be reasonable (Wilson et al. 1997):

> "The Panel considered the adequacy of the values of waste tensile strength used in the tensile failure calculations. The Panel concluded that there is some uncertainty in the degree to which the values measured on waste surrogates truly represent the actual waste following extended exposure in WIPP, but that generally, the selected values are reasonable for the intended purposes."

The Agency accepted the Peer Panel's 1997 conclusions at the time of the CCA. There remains uncertainty in the degree to which the surrogate formulations represent the actual future state of WIPP waste, however the approaches taken to develop those surrogates were logical and reasonable. The Agency also considered the ramifications of the limited size of the test specimens. The surrogate waste test specimens were 8.25 cm in diameter and could not accommodate the larger pieces of metal or CPR that are actually present in the waste. These larger pieces with larger surface areas would degrade more slowly than the smaller pieces that would fit into the test specimens. Because the more highly degraded waste is weaker, SNL's capability to accommodate only the smaller pieces is more suitable for evaluating the low end of the TAUFAIL range than the high end. Based on these considerations, the Agency concludes that SNL's surrogate formulations are acceptable for use in evaluating the low end of the TAUFAIL range.

Pre-compaction Pressure. EPA accepts SNL's recommendation to use test results for surrogate waste samples pre-compacted to 2.3 MPa instead of 5.0 MPa for the reasons presented by SNL and described above in Section 3.2. The lower bound of the TAUFAIL distribution should be bounding and SNL's additional computations demonstrated that, while 5.0 MPa was not bounding, a 2.3 MPa pre-compaction pressure is considerably lower than the 15 MPa lithostatic pressure that may ultimately be applied to the waste upon creep closure of the waste panels. The applied pre-compaction pressure of 2.3 MPa therefore appears to be reasonably bounding for WIPP waste subjected to compaction by creep closure.

Degree of Degradation. EPA reviewed SNL's recommendation to base the lower bound of the TAUFAIL distribution on 50% degraded waste by requesting DOE to confirm that substantial fractions of metallic (iron) and organic (CPR) waste materials are predicted to remain undegraded in the WIPP repository after 10,000 years. This request took the form of two additional completeness comments.

> **EPA Comment 1-23-5 Waste Shear Strength.** *Please address the following:*
> *1. Provide horsetail plots of the remaining fraction of uncorroded iron in the repository throughout the 10,000-year regulatory time frame from the CRA-2009 PABC from each of the three replicates and each scenario.*
> *2. Provide horsetail plots of the remaining fraction of undegraded CPR in the repository throughout the 10,000-year regulatory time frame from the CRA-2009 PABC from each of the three replicates and each scenario*

EPA requested plots from the CRA-2009 PABC because that was the Agency-approved baseline assessment at the time of the Agency's review. PA predictions of the amount of uncorroded iron and undegraded CPR waste over the 10,000-year period would help support the selection of reasonable bounding values for TAUFAIL. DOE provided 36 plots in response to this request, of which 18 showed the fraction of uncorroded iron and 18 showed the fraction of undegraded CPR over the regulatory time frame. Examples of these plots for iron and CPR are presented in Figures 3a and 3b. Each figure shows 100 results for Scenario 2 of Replicate 2. Scenario 2 consists of a single borehole intrusion that penetrates both a waste panel and pressurized brine in the Castile Formation at 350 years after closure. This is called an E1 intrusion. This scenario was chosen because inundation of the waste panel with Castile brine generally provides the most severe corrosion/degradation conditions.

Figure 3a depicts the remaining fraction of uncorroded iron. The mean of the individual vectors (the heavy red line) shows that an average of about 58% of the iron remains uncorroded at 10,000 years. The individual vectors (the light green lines) show that complete corrosion of all iron occurs in only one vector and this is predicted to happen at about 3,000 years after repository closure.

Figure 3b depicts the remaining fraction of undegraded CPR. Here the mean of the individual vectors (the heavy red line) shows that an average of about 75% of the CPR remains undegraded at 10,000 years. The individual vectors (the light green lines) show that complete degradation of all CPR occurs in four vectors beginning at about 7,000 years after repository closure.

For completeness, DOE also provided a similar set of plots from the CRA-2014 PA. The Agency's review confirmed that most of the iron in the waste was predicted to remain uncorroded for all realizations and all replicates in both PAs throughout the 10,000-year regulatory time frame. In the CRA-2009 PABC, an average of 58 to 64% of the iron remained uncorroded at 10,000 years in all replicates and scenarios, and there was at most only one vector in any given scenario where all of the iron was predicted to corrode. Some of the iron therefore remained uncorroded in at least 99% of the vectors. In the CRA-2014 PA, some of the iron remained uncorroded in at least 97% of the vectors.

A similar result was found for CPR degradation. The Agency's review confirmed that most of the CPR in the waste was predicted to remain undegraded for all realizations and replicates in both PAs throughout the 10,000-year regulatory time frame. In the CRA-2009 PABC, an average of 75 to 82% of the CPR remained undegraded at 10,000 years in all replicates and scenarios, and some of the CPR remained undegraded in at least 96% of the vectors. In the CRA-2014 PA, some of the CPR remained undegraded in at least 95% of the vectors.

Based on the foregoing results showing that more than half of the iron and CPR in the waste is predicted to remain undegraded after 10,000 years, the Agency accepts DOE's use of 50% degraded surrogate waste in determining the lower bound of the TAUFAIL distribution as reasonable and appropriate.

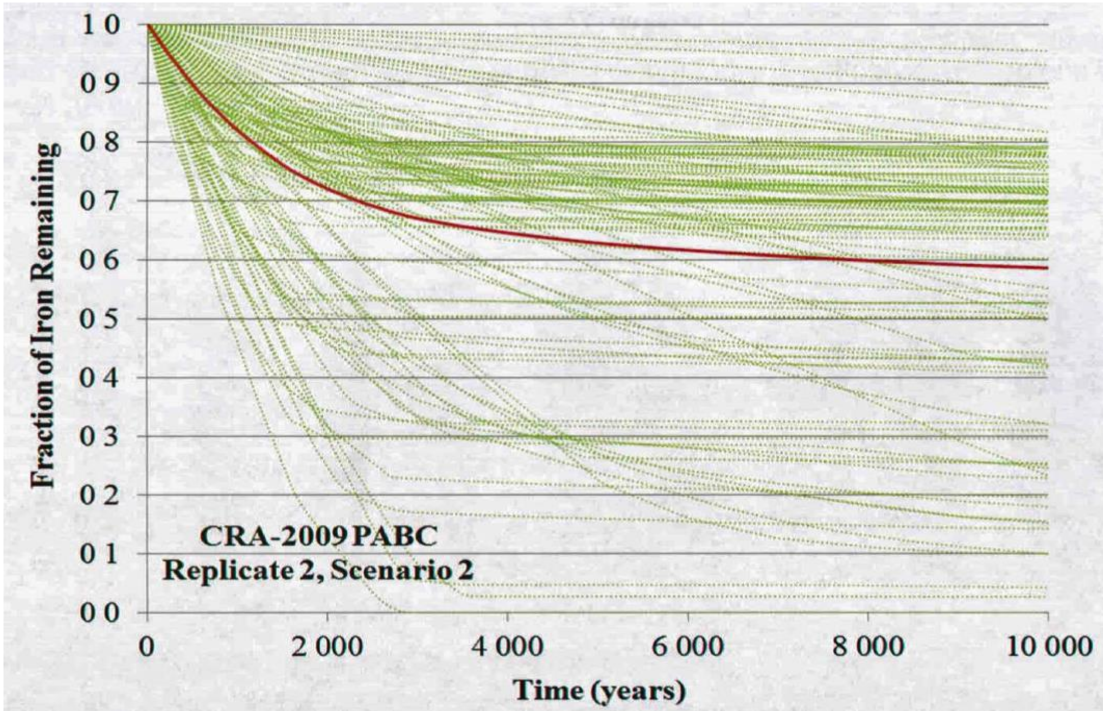

Figure 3a. Fraction of uncorroded iron remaining in the WIPP repository over the 10,000-year regulatory time frame. The red line is the mean of all 100 vectors (Source: Herrick 2015a).

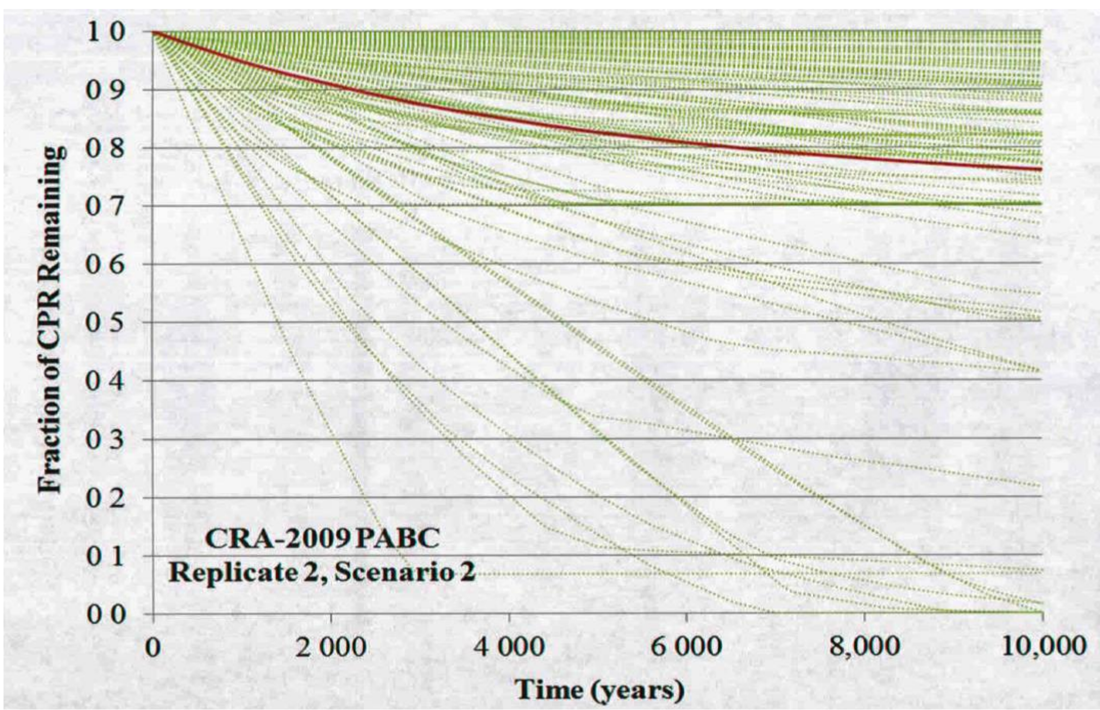

Figure 3b. Fraction of undegraded CPR remaining in the WIPP repository over the 10,000-year regulatory time frame. The red line is the mean of all 100 vectors (Source: Herrick 2015b).

Lower Bound Shear Strength. SNL recommended and DOE accepted use of the arithmetic mean result of 2.22 Pa for the lower bound shear strength for TAUFAIL. The Agency was concerned that use of the mean may not be appropriate for a value that is intended to be a lower bound. The results of the bilinear analyses for all five tests are provided in Table 2.

Table 2. Shear Strengths of 50% Degraded Samples Pre-compacted at 2.3 MPa

Sample Number	Shear Strength from Bilinear Analysis (Pa)
WF-50-02	2.54
Flume 50-01	1.60
WF-50-203-01	3.09
WF-50-203-02	1.78
WF-50-203-03	2.10
Average	**2.22**

Source: Herrick et al. 2012, Table 12

Although three out of the five tests provided highly scattered results, the scatter is likely due primarily to inherent inhomogeneity rather than sample damage. Because inhomogeneity would also be present in actual WIPP waste, the Agency concludes that use of test results that show considerable scatter is acceptable but questions the use of an overall average of all five tests as appropriate for a lower bound of TAUFAIL,

When evaluating the test results individually, two of the 50% degraded tests produced reasonably definitive trends. These tests yielded shear strengths of 2.54 and 3.09 Pa, and their results are shown in Figures 4a and 4b. The Agency notes that these shear strengths are both greater than the average of 2.22 Pa and were the highest results for all five tests. This suggests that these two samples may not have been as highly disturbed during testing as the others and may indicate that the typical shear strength for 50% degraded samples pre-compacted at 2.3 MPa may actually be higher than the average of 2.22 Pa. However, because of uncertainties associated with the source of the scatter, it cannot be concluded that the results of the two tests with reduced scatter would provide an acceptable lower bound. The three remaining tests produced shear strengths that were less than the overall average. Despite the Agency's consideration that sample damage may have more strongly affected the three lower value results, these results possibly represent the variable shear strength of 50% degraded samples pre-compacted at 2.3 MPa. The Agency believes that the lowest measured result of 1.60 Pa should be used in WIPP PA as that provides a lower bounding value for the parameter.

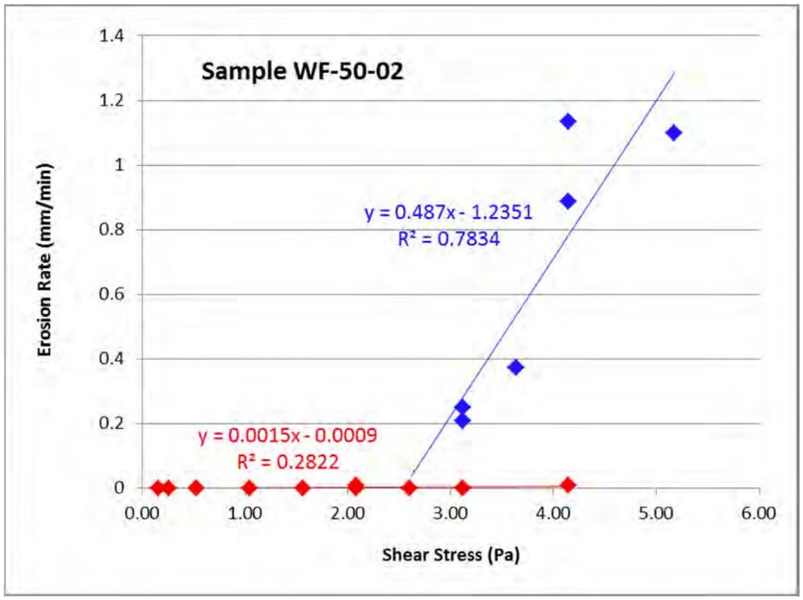

Figure 4a. Bilinear analysis of test results for Sample WF-50-02 yielding a critical shear stress of 2.54 Pa. Test was performed on 50% degraded material with a pre-compaction pressure of 2.3 MPa. (Source: Herrick et al. 2012, Figure 36a).

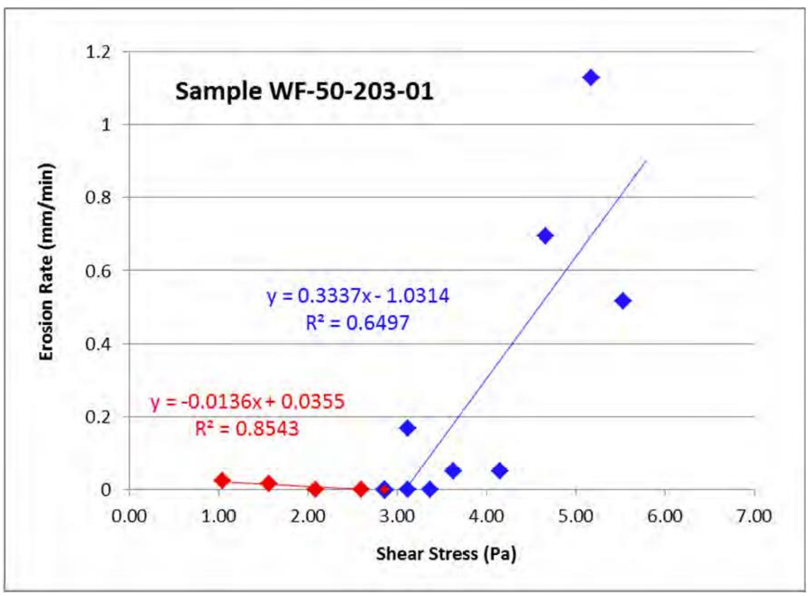

Figure 4b. Bilinear analysis of test results for Sample WF-50-203-01 yielding a critical shear stress of 3.09 Pa. Test was performed on 50% degraded material with a pre-compaction pressure of 2.3 MPa. (Source: Herrick et al. 2012, Figure 38a).

Uniform Distribution. DOE changed the proposed distribution of TAUFAIL from log-uniform to uniform in the CRA-2014 PA. A log-uniform distribution provides equal probability of sampling from each order of magnitude in an uncertainty range and is appropriate for distributions whose values span several orders of magnitude. A uniform distribution is appropriate for sampling a more limited range of values. This proposed change in type of distribution prompted the Agency to present the following request to DOE as a completeness comment.

> **EPA Comment 1-23-5 Waste Shear Strength.** *Identify and justify the consequences of using the proposed uniform distribution rather than the currently approved log-uniform distribution for TAUFAIL.*

In summary, DOE's response stated that if nothing is known about a distribution except the quantitative constraints defining its range, then the distribution with the largest entropy, the uniform distribution, should be chosen as the default. Further, by its nature, a log-uniform distribution assigns considerable negative skewedness to the distribution of possible values of the uncertain variable. Thus, its use implies that there exists more information about the distribution than just its minimum and maximum. DOE also points out a consensus view that use of a logarithmic distribution is only justified when the range of values spans more than two orders of magnitude, for which their proposed range of 2.22 to 77.0 Pa does not qualify.

The Agency considers that knowledge of the consequence of deemphasizing the low end of a parameter range by selecting a uniform distribution constitutes additional information about a parameter. This additional information can justify the use of a log-uniform distribution when the parameter range covers less than two orders of magnitude. However, in this case the Agency also considers that the revised range of 1.60 to 77.0 Pa is sufficiently small that the assignment of a uniform distribution is justified.

5.0 CONCLUSIONS

EPA accepts the results of DOE's flume tests and considers the results to provide an improved basis for determining the lower bounding value for the distribution of TAUFAIL in WIPP PA. With one exception, EPA also accepts SNL's recommendations for identifying the suite of flume tests from which the lower bounding value for that distribution was to be selected. The single exception is the lower bounding value itself, which SNL recommended be based on the average test results. This average result is demonstrably not bounding because test results that SNL considered viable were lower than this average. Therefore, the Agency considers the following determinations to be appropriate for inclusion in the next full WIPP PA that DOE submits.

Degraded Surrogate Waste: Surrogate formulations for degraded WIPP waste of the type documented in Hansen et al. (1997) are adequate for use in SNL's flume tests.
Pre-compaction Pressure: SNL's recommended pre-compaction pressure of 2.3 MPa is reasonable and appropriate based on the current understanding of creep closure and gas generation effects in a waste panel as modeled in WIPP PA.
Degree of Degradation. SNL's recommended use of 50% degraded surrogate waste as the basis for determining the lower bound of the TAUFAIL distribution is appropriate given PA results showing that large quantities of undegraded iron and CPR are present in the waste panels throughout the regulatory period.

Lower Bound Shear Strength. The Agency accepts the minimum value in this test suite of 1.60 Pa as the lower bounding value for the TAUFAIL distribution.
Uniform Distribution. SNL's recommendation of changing the distribution of TAUFAIL from log-uniform to uniform was accepted by EPA in view of the considerably smaller range between the upper and lower bounds in TAUFAIL values.

DOE accepted the replacement value of 1.60 Pa for the lower bound of the TAUFAIL distribution in an EPA/DOE technical exchange meeting held on September 1, 2016. In lieu of a revision to the CRA-2014 PA calculations (a PABC), the Agency requested DOE conduct the SEN4 sensitivity study incorporating this accepted lower bound for TAUFAIL to help determine its impact on WIPP PA (EPA 2017).

The SEN4 sensitivity study did not determine the sensitivity to TAUFAIL alone but instead determined the sensitivity to the combined results of several changes to the PA database. The sensitivity to the change in the lower bound for TAUFAIL could therefore not be separately determined and other database changes, such as those involving Castile brine releases and repository chemical conditions, were expected to have greater impacts. The change in the range of uncertainty in TAUFAIL was expected to have only a minor effect on overall PA results because decreasing the lower bound from 2.22 Pa to 1.60 Pa does not significantly change the range of a uniform distribution that has an upper bound of 77 Pa. EPA requested changing the lower bound of the range to provide consistency with the principle that in a bounding uncertainty analysis the lower bound should have a value that is reasonable and even lower values should not be expected. DOE's proposed use of the mean result of five tests was inappropriate because even lower values had been measured and would therefore be reasonably expected. Even though some of the parameters changed in the SEN4 analysis likely had a greater impact on results than the change in the lower bound of TAUFAIL, the mean total releases including all changes did not exceed EPA's WIPP release limits, nor the 95% confidence level in that mean, nor were the release limits exceeded by any individual vectors.

REFERENCES

CTAC (Carlsbad Area Office Technical Assistance Contractor) 1997. *Expert Elicitation on WIPP Waste Particle Diameter Size Distribution(s) during the 10,000-year Regulatory Postclosure Period: Final Report.* ERMS 541365. WIPP Records Center, Sandia National Laboratories, Carlsbad, New Mexico.

DOE (U.S. Department of Energy) 2015. *Response to Environmental Protection Agency Letter Dated December 17, 2014, Regarding the 2014 Compliance Recertification Application.* Letter from Jose Franco (DOE) to Jonathan Edwards (EPA), U.S. Department of Energy, Carlsbad Field Office, Carlsbad, New Mexico. March 18, 2015.

EPA (U.S. Environmental Protection Agency) 2017. *Review of EPA Sensitivity Studies of DOE CRA-2014 WIPP Compliance Recertification Performance Assessment. FDMS Docket ID No. EPA-HQ-OAR-2014-0609* Prepared by S. Cohen & Associates, Vienna, Virginia, for EPA Office of Radiation and Indoor Air, Washington, DC.

Hansen, F.D., M. K. Knowles, T. W. Thompson, M. Gross, J. D. McLennan, and J. F. Schatz 1997. *Description and Evaluation of a Mechanistically Based Conceptual Model for Spall*. SAND97–1369. Sandia National Laboratories, Albuquerque, New Mexico.

Hansen, F.D., T.W. Pfeifle, and D.L. Lord 2003. *Parameter Justification Report for DRSPALL*. SAND2003-2930. Sandia National Laboratories, Albuquerque, New Mexico.

Hansen, F.D. 2005. *A Revisit of Waste Shear Strength.* ERMS 541354. Sandia National Laboratories, Carlsbad, New Mexico.

Herrick, C.G., M. Riggins, and B.Y. Park 2007. *Recommendation for the Lower Limit of the Waste Shear Strength (Parameter BOREHOLE : TAUFAIL), Rev. 0*. ERMS 546033. Sandia National Laboratories, Carlsbad, New Mexico.

Herrick, C.G., M.D. Schuhen, D.M. Chapin, and D.C. Kicker 2012. *Determining the Hydrodynamic Shear Strength of Surrogate Degraded TRU Waste Materials as an Estimate for the Lower Limit of the Performance Assessment Parameter TAUFAIL, Revision 0.* ERMS 558479. Sandia National Laboratories, Carlsbad, New Mexico.

Herrick, C.G. and T. Kirchner 2013. *Follow-up to questions concerning TAUFAIL flume testing raised during the November 14-15, 2012 technical exchange between the DOE and EPA.* Memorandum to Chris Camphouse (SNL), January 24, 2013. ERMS 559081. Sandia National Laboratories, Carlsbad, New Mexico.

Herrick, C.G. 2015a. *CRA-2009 PABC and CRA-2014 plots of the fraction of iron remaining by replicate and scenario.* ERMS 563282. Sandia National Laboratories, Carlsbad, New Mexico. February.

Herrick, C.G. 2015b. *CRA-2009 PABC and CRA-2014 plots of the fraction of CPR remaining by replicate and scenario.* ERMS 563286. Sandia National Laboratories, Carlsbad, New Mexico. February.

Parchure, T.M. and A.J. Mehta 1985. “Erosion of soft sediment deposits,” *Journal of Hydraulic Engineering*, Vol. 111, No. 10, pp. 1308-1326.

Partheniades, E. and R.E. Paaswell 1970. “Erodibility of Channels with Cohesive Boundary.” *Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers*, Vol. 96, No. HY3, pp. 755-771.

Wilson, C., D. Porter, J. Gibbons, E. Oswald, G. Sjoblom, and F. Caporuscio 1997. *II-G-22: Conceptual Models Third Supplementary Peer Review Report, April 1997. A Peer Review Conducted for U.S. Department of Energy*. U.S. Department of Energy, Carlsbad Area Office, Carlsbad, New Mexico. April.

Yew, C., J. Hanson, and L. Teufel 2003. *Waste Isolation Pilot Plant Spallings Conceptual Model Peer Review Report*. ERMS 532520. Sandia National Laboratories, Carlsbad, New Mexico.