DOCKET: EPA-HQ-OAR-2001-0012-0777

ONSITE TECHNICAL REVIEW REPORT OF SANDIA NATIONAL LABORATORIES CARLSBAD

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

July 2023

TABLE OF CONTENTS

List of Acronyms, Abbreviations and Initialisms iii
1.0 Executive Summary 1
2.0 Scope 1
3.0 EPA Participants 2
4.0 Technical Review 2
4.1 Document Review Process 2
4.1.1 Overview of EPA Findings 3
4.2 FM4 Database 4
4.2.1 Prioritization and Performance of Database Updates 5
4.2.2 Database Verification Exercises 8
4.3 Corrosion 9
4.3.1 Documentation 9
4.3.2 SNL Technical Review 10
4.4 Mineral Fragment Colloids 11
4.4.1 Issues of Scientific Rigor 11
4.5 TAUFAIL 14
4.5.1 Test Plan 14
4.5.2 Scientific Notebooks 16
4.5.3 Measuring and Test Equipment 18
4.5.4 Control of Samples and Standards 19
4.5.5 Document Review 19
4.5.6 Corrective Action 20
4.5.7 TAUFAIL Test Conclusions 21
4.6 Boreholes 21
4.7 BRAGFLO 22
4.7.1 Permeability Changes and Capillary Pressure 22
4.7.2 Convergence Criteria and Mass Balance 23
4.7.3 Summary and Potential Resolution 23
5.0 Conclusions 24
5.1 Observations 24
5.2 Observations and Recommendations 24
6.0 References 27
Attachment Correspondence from R. Patterson (DOE) to K. Economy (EPA) dated July 11, 2016 ..... A-1

LIST OF TABLES

Table 1 EPA Participants 2
Table 2 FM4 Database-Related Documents Reviewed 6
Table 3 Iron Corrosion-Related Documents Reviewed 9
Table 4 Mineral Fragment Colloids Documents Reviewed 11
Table 5 TAUFAIL Documents Reviewed 15
Table 6 Key Controlling Procedures 16
Table 7 Drilling Rate and Plugging Pattern Parameters for the CRA-2019 PA 22

LIST OF FIGURES

Figure 1 Correspondence from R. Patterson (DOE) to K. Economy (EPA) dated July 11, 2016, page 1 of 2 ..... A-1

Figure 2 Correspondence from R. Patterson (DOE) to K. Economy (EPA) dated July 11, 2016, page 2 of 2 ..... A-2

LIST OF ACRONYMS, ABBREVIATIONS AND INITIALISMS

Am	Americium
AP	Analysis Plan
BRAGFLO	Brine and Gas Flow Code
C	Carbon
CAQ	Condition(s) Adverse to Quality
CCA	Compliance Certification Application
CFR	*Code of Federal Regulations*
Cl	Chlorine
CRA	Compliance Recertification Application
CRA-2014	2014 Compliance Recertification Application
CRA-2019	2019 Compliance Recertification Application
DAS	Data Acquisition System
FM4	Data0.FM4 geochemistry database
DOE	U.S. Department of Energy
DRC	Document Review and Comment
EDTA	Ethylenediaminetetraacetic acid
EPA or the Agency	U.S. Environmental Protection Agency
ERMS	Electronic Records Management System
EQ3/6	Geochemistry modeling program
Fe	Iron
FR	*Federal Register*
H	Hydrogen
ICONVTEST	Convergence criteria test flag in BRAGFLO
Log K	Solubility constant
Mg	Magnesium
Na	Sodium
NP	Nuclear Waste Management Procedure
O	Oxygen
PA	Performance Assessment
PADB	WIPP PA Database
Pb	Lead
PFLOTRAN	Subsurface flow and reactive transport code
PI	Principal Investigator
QA	Quality Assurance
QC	Quality Control
SN	Scientific Notebook
SNL	Sandia National Laboratories
SNS	Scientific Notebook Supplement
SP	Specific Procedure
Sr	Strontium
TAUFAIL	Waste Shear Strength Parameter
TP	Test Plan
WIPP	Waste Isolation Pilot Plant

1.0 EXECUTIVE SUMMARY

During the week of December 5, 2022, U.S. Environmental Protection Agency (EPA) staff conducted an onsite review of Sandia National Laboratories (SNL) technical investigation practices in Carlsbad, NM. Staff visited the SNL Records Center to review work and adherence to SNL procedures on topics related to the Waste Isolation Pilot Plant (WIPP) performance assessment (PA) issues from the 2019 Compliance Recertification Application (CRA-2019). The review included the geochemical thermodynamic database Data0.FM4 (FM4), iron corrosion, mineral fragment colloids, waste shear strength (TAUFAIL), the borehole database and the BRAGFLO (Brine and Gas Flow Code) model. EPA staff also interviewed SNL staff. Results of the review indicated that SNL demonstrated many instances of good documentation and review practices in its iron corrosion, TAUFAIL and boreholes work. While this was not a quality assurance (QA) audit, EPA identified a major concern in SNL's lack of adherence to its document review for Nuclear Waste Management Procedure (NP) 6-1. Because of this deficiency, preventable technical issues emerged, such as problematic documentation (e.g., milestone reports providing insufficient technical depth and lack of justifications for database updates), instances where procedures were not followed (e.g., geochemistry database parameter fitting exercises were not performed per analysis plans (APs) and conditions adverse to quality (CAQ) were not identified per SNL NPs), and the quality of some data was degraded (e.g., there were instances where inappropriate conclusions were made using data that were published despite being below detection limits, such as in Xiong (2015)). In some cases, such as with the FM4 database, these technical issues resulted in a prolonged completeness process during EPA's review of CRA-2019 as EPA required SNL to perform additional calculations to determine the impacts of the errors in the database. EPA is concerned that SNL is not catching mistakes or procedural deviations in its current review process and believes that stronger adherence to the minimum criteria for a technical reviewer listed in Section 2.3 of NP 6-1 can improve the quality of the analyses and data that are published. EPA's overall conclusion is that SNL Carlsbad needs to reassess and strengthen its review process, as well as its current methodologies, to ensure data quality and the integrity of the WIPP PA. The U.S. Department of Energy (DOE) should also increase its oversight of SNL work to confirm it follows its procedures and produces data of acceptable quality.

2.0 SCOPE

The EPA WIPP Compliance Criteria allow the Agency to be afforded "unfettered and unannounced access to inspect any area of the WIPP, and any locations performing activities that provide information relevant to compliance application(s)," which includes access to records pertaining to the WIPP (40 CFR 194.21(a)(1)). During the week of December 5, 2022, EPA staff visited the Records Center at SNL in Carlsbad, NM, for an onsite review of SNL's technical document review process, which is documented according to SNL's QA processes. The Agency's goal for the visit was to learn how SNL's internal review processes were applied to various technical topics that ultimately inform parameters used in the WIPP PA. EPA's visit was an extension of its CRA-2019 review, as the topics the Agency investigated were related to many of the technical issues identified during the CRA-2019 review. The May 3, 2022, recertification decision (87 FR 26126) contains more discussion of those technical issues. Of concern to EPA is the increasing length of time EPA has spent reviewing Compliance Recertification Application (CRA) documentation due to the number of technical issues encountered in the submitted work.

Part of EPA's objective in this visit was also to assess SNL's document review process to determine whether SNL can identify and correct technical issues in advance before submitting documentation, thereby streamlining EPA's review. This, along with more proactive technical discussions between DOE and SNL technical staff, should allow the recertification and similar review process to be completed faster and more efficiently in the future.

EPA examined a diverse array of topics related to SNL's experimental and modeling work, including the FM4 thermodynamic database, iron (Fe) corrosion, mineral fragment colloids, TAUFAIL, boreholes, and the BRAGFLO model. Agency personnel examined test plans (TPs), APs, analysis reports, spreadsheet calculations, laboratory notebooks, and QA forms such as Document Review and Comment (DRC) forms per SNL procedure NP 6-1 (e.g., Nielsen 2022) within the Records Library. Additionally, EPA staff interviewed SNL personnel to gain a further understanding of their processes. A list of EPA participants is provided in Table 1.

3.0 EPA PARTICIPANTS

Table 1 EPA Participants

Participant	Affiliation
Jay Santillan	EPA
Xinyue Tong	EPA
David Back	SC&A (EPA contractor)
Charles Wilson	SC&A (EPA contractor)
Jonathan Major (virtual)	EPA
Joseph Rustick (virtual)	EPA
Janet Schramke (virtual)	SC&A (EPA contractor)

4.0 TECHNICAL REVIEW

4.1 Document Review Process

SNL reports are reviewed internally using procedures outlined in NP 6-1 and are documented on form NP 6-1-1. Per discussions with SNL QA staff member S. Nielsen, reviewers are SNL staff who have not been involved in the work for a given technical project. Because SNL Carlsbad is small, supervisors are easily able to identify reviewers within SNL Carlsbad who have not been directly involved in the work. If additional qualified reviewers are needed, reviewers are recruited from SNL in Albuquerque, NM. A reviewer's qualifications are documented in form NP 2-1-1, per NP 2-1 (Davis 2006), in which an employee lists technical proficiencies and other educational experiences. Generally, three sets of reviews are performed in sequential order: a technical review, followed by a QA review and then a management review. Although this sequence of reviewers is not outlined specifically in NP 6-1, per discussions with QA staff, SNL personnel generally follow this practice.

Section 2.3 in Revision 10 of procedure NP 6-1 lists the following minimum criteria for a technical reviewer of a document to address:

- *Are the objectives clearly stated and fulfilled?*
- *Is the technical activity clearly described?*
- *Are equations/calculations accurate?*
- *Does logic lead to reasonable conclusions?*
- *Are the results drawn from the data supported by the data presented?*
- *Are data/tables/figures easily understood? Are legends complete?*
- *Are assumptions stated clearly and do they lead logically to the conclusions presented?*

4.1.1 Overview of EPA Findings

Each of the following sections addresses a specific topical area of EPA's review. The process EPA followed and the documents reviewed are described in detail. Findings and recommendations for each topical area are provided, followed by detailed discussions of the review process.

As a general observation, throughout its review of the DRC forms associated with the documents selected for this onsite technical review, EPA noted multiple instances where reviews did not follow the documented processes or meet the minimum criteria listed above. For the development of the FM4 geochemistry database, for example, much of the technical, QA and management reviews occurred simultaneously due to time constraints to complete CRA-2019, which, although it is not listed as part of the NP 6-1 procedures, counters the current approach. SNL should evaluate and consider whether documenting the sequential order of reviewers to NP 6-1 procedures may improve data quality. In many cases, reviews were editorial and effectively identified transcription errors. One such review noted a potential discrepancy for the Na^{+} and $Mg(Oxalate)^{2-}$ system in the technical DRC form for Domski (2019) (Electronic Records Management System (ERMS) 571055). However, some of the database DRCs did not provide any in-depth questions about the validity and rationale of the database updates, nor did they critically examine their underlying assumptions, reasons for the additions of new thermodynamic data to the database, or the downstream PA impacts of these additions. In terms of NP 6-1, much of the DRC form content did not fulfill the criteria: "*Does logic lead to reasonable conclusions?*" and "*Are assumptions stated clearly and do they lead logically to the conclusions presented?*" Section 4.2 provides a more detailed discussion of the FM4 database review.

Experimental work, including mineral fragment colloid characterization experiments (Kirkes 2020), also lacked sufficient document review. The mineral fragment colloids milestone report (Kirkes 2020), for example, published results using conclusions about the colloid agglomeration process that were unsupported by the available data (see Section 4.4), despite DRC forms raising awareness of this process. The milestone report also provided no rationale for the use of an approach that seemingly repeated work done during the WIPP's initial certification. Although EPA did not review the DRC forms of lead solubility experiments (e.g., Kirkes et al. 2014, Xiong 2014), similar issues appeared in the SNL reports that made it through the DRC process, meaning that data were published that were below experimental detection limits (see Section 4.2 for more discussion of lead experimental work).

EPA also identified DRC issues in non-geochemistry-related experimental activities, though these issues did not result in the same data quality issues noted above. Although the TAUFAIL

data used to support the WIPP PA ultimately proved to be adequate (EPA Docket No. EPA-HQ-OAR-2014-0609 (EPA 2017a)), TAUFAIL DRCs also failed to identify instances in which data were measured despite calibrations being expired (see Section 4.5 for more discussion of calibration). Much of the work stated above did not meet the NP 6-1 review criteria: "*Are the objectives clearly stated and fulfilled?*", "*Are equations/calculations accurate?*" and "*Are the results drawn from the data supported by the data presented?*"

Based on this assessment, EPA concludes that, for the items reviewed, SNL reviewers did not follow the procedures outlined in NP 6-1. EPA believes that the SNL document review process can benefit significantly from a larger pool of qualified reviewers who are further removed from the work and able to critically examine it for both transcription errors and the underlying approaches and assumptions to the data. Furthermore, EPA recommends reemphasizing the current minimum criteria and adding other minimum criteria for technical reviewers to use, including assessment of the work's potential impact to the WIPP and the WIPP PA; assessment of the work's adherence to the related specific procedures (SPs), TPs and APs; and review of the quality of the data produced (e.g., whether calibrations were used and whether data fall within instrument detection limits). Because the DRC process was superficial in multiple instances, preventable inadequacies such as degraded data quality were allowed to emerge, as described in Sections 4.2, 4.4 and 4.5. In some cases, such as with the FM4 database, these inadequacires resulted in a prolonged completeness process during EPA's review of CRA-2019 as EPA required SNL to perform additional calculations to determine the impacts of the errors in the database. By contrast, the review of corrosion information in Section 4.3 indicated that the SNL review process was successfully followed and contributed to adequate documentation for this topic. This suggests that the identified deficiencies in documentation of other topical areas does not result from an inadequate process *per se*, but from staff not following the documented process. SNL needs to ensure that staff follow the existing processes. DOE should also increase its oversight of SNL work to ensure SNL follows its procedures and produces data of acceptable quality.

4.2 FM4 Database

EPA reviewed multiple documents related to the creation of the FM4 thermodynamic database used to calculate the dissolved actinide source term for CRA-2019 and presents the following observations:

- EPA concurs with the need to update the geochemistry database used in PA to reflect new information; however, SNL needs a documented and systematic approach for prioritizing database updates so that the updated information is appropriately incorporated. The approach should include clear justification for adding new systems or revisions to existing data, as well as an evaluation of the possible consequences to actinide solubility in PA results, especially if the addition of an incomplete system is prioritized. SNL should also document reasons for not updating or adding available data, showing whether such omissions are justified by a predicted lack of consequence to actinide solubilities or data quality uncertainties.

- SNL needs to ensure that it utilizes a clearly defined and consistent data assessment and selection methodology that includes documented verification tests and acceptance

criteria. This means SNL needs to perform parameter-fitting exercises that demonstrate that database updates are able to reasonably simulate available solubility data and include this information in the analysis report or technical memorandum documenting the database parameters. SNL has added parameter-fitting checks to recent AP procedures (AP-200) (Domski 2022), but this was largely absent for the CRA-2019. This is a step in addressing EPA's concern, and the Agency will verify that these procedures are followed in future reviews.

- SNL needs to improve its review of the scientific rigor of data before it is reported by ensuring instrument calibrations and detection limits are established; methods, data, and assumptions are assessed as they are collected; and that work is evaluated within the context of existing WIPP studies. EPA has noted instances in this technical review, such as lead experiments, where data quality was significantly compromised. Although data from the experiments EPA reviewed were not included in any PAs, EPA recommends SNL verify that other WIPP-related experiments do not have similar scientific rigor issues. Section 4.4 provides a full treatment of this issue.

4.2.1 Prioritization and Performance of Database Updates

The geochemistry database used in CRA-2019 has had several updates since the 2014 Compliance Recertification Application (CRA-2014), including updated solubility constants (log K) related to the magnesium mineral hydromagnesite and organic ligands, as well as the addition of iron and lead systems. EPA reviewed SNL's records to determine whether SNL had a process for prioritizing the selection of data. The Agency is aware that some of the updates, such as for hydromagnesite, updates were a result of EPA comments from the previous recertification. However, in its review of SNL's geochemistry database documentation, EPA found no clear rationale for why SNL chose to prioritize the addition of specific incomplete systems, such as lead, when updating the database. This is particularly important because adding incomplete systems may have downstream consequences on actinide solubilities that can adversely affect PA calculations.

EPA began its technical review of the geochemistry database by examining the procedures and APs related to the creation of FM4 for CRA-2019. AP-182 (Jang 2019) lists the steps used to update the database, AP-183 (Sisk-Scott 2019) lists the log K and Pitzer parameters to add to the database, Domski (2019) documents the creation of the database, and Domski (2018) documents the addition of $Ca_2EDTA•7H_2O$ to the database. Table 2 lists additional related documents, including the DRC forms associated with SNL's NP 6-1 procedure.

Table 2 FM4 Database-Related Documents Reviewed

Document Type	Title/Description	Citation	Associated DRC Forms
Analysis Plan (AP-182)	Analysis Plan for Derivation and Addition of Equilibrium Constants and Pitzer Interaction Parameters to the WIPP Geochemical Thermodynamic Database	Jang (2019)	Not examined by EPA
Analysis Plan (AP-183)	Analysis Plan to Update the WIPP Geochemical Thermodynamic Database (DATA0.FM1) to Data0.FM4 for CRA-2019, Revision 1	Sisk-Scott (2019)	Not examined by EPA
Memo	Memo on the Estimation of the Solubility Product (Log K_{sp}) for $Ca_2EDTA.7H_2O$, and Pitzer Parameters for the Na+ / $CaEDTA^{2-}$ Pair	Domski (2018)	ERMS 570203 (Technical) ERMS 570204 (QA) ERMS 570205 (Management)
Analysis Report	An Update to the EQ3/6 Pitzer Thermodynamic Database DATA0.FM1 With the Creation of Data0.FM4	Domski (2019)	ERMS 571053 (Signature authority for DRC) ERMS 571054 (Signature authority for DRC) ERMS 571055 (Technical) ERMS 571056 (QA) ERMS 571057 (Management)
Test Plan (TP 08-02)	Iron, Lead, Sulfide, and EDTA Solubilities Test Plan TP 08-02	Ismail et al. (2008)	Not examined by EPA
Analysis Plan (AP-176)	Analysis Plan for the Development of a Self-Consistent Extension of the WIPP Geochemical Thermodynamic Database that Includes Aqueous Ferrous Iron Chemistry	Jang and Domski (2016)	Not examined by EPA
Analysis Report	Derivation of Pitzer Interaction Parameters and Thermodynamic Properties for the Aqueous Species of Ferrous Iron and Their Pairs, Rev. 2	Jang and Kim (2016)	ERMS 567284 (Technical) ERMS 567285 (QA) ERMS 567286 (Management)
Memorandum	Revisit to the Pitzer Geochemical Thermodynamic Model for the $PbCl_2(s)$ – HCl – H_2O System	Jang (2021)	Not examined by EPA
Analysis Report (GEOC-21-03)	Analysis Report Documenting the Assessment of the Solubility of Lead, EDTA and Other Organic Ligands in Non-Sulfide Systems Performed Under TP 08-02 and Under TP 20-01, Rev. 0	Jang et al. (2021)	Not examined by EPA
Milestone Report	Third Milestone Report on Test Plan TP 08-02, "Iron, Lead, Sulfide, and EDTA Solubilities," Rev. 2	Kirkes et al. (2014)	Not examined by EPA
Analysis Report	Experimental and Thermodynamic Modeling of $PbEDTA^{2-}$ Interactions in NaCl and $MgCl_2$ Solutions	Xiong (2014)	Not examined by EPA

Although some rationale is provided for using the Hummel et al. (2005) log K values for updating organic ligands (i.e., the desire to align the database with the Nuclear Energy Agency's values), justification for the use of the Powell et al. (2009) data for lead is limited, except that

issues with SNL's own experimental work prevented the incorporation of its data in the FM4 database (Sisk-Scott 2019). As such, it appears SNL had the choice of including either data derived from its own experiments or data from Powell et al. (2009). However, even though SNL appropriately omitted its own lead data from the FM4 database, there is no documentation suggesting that SNL explored other options, including omitting lead from the database for CRA-2019 altogether, or performing and documenting a literature review that may have identified other potentially suitable lead thermodynamic data. It is unclear why SNL prioritized the inclusion of untested lead parameters that were inconsistent with the existing conceptual models because of the lack of Pitzer parameters. This addition impacted water balance during the dissolved actinide source term calculation, with downstream consequences for the WIPP PA and its defensibility.

EPA further investigated the experimental issues that resulted in the omission of SNL lead data. SNL's retrospective evaluation of these issues is documented in reports such as Jang et al. (2021) and Jang (2021). One of the issues examined with the lead experiments relates to uncertainty regarding the mineral phases of lead that were present in the experiments. Initial experiments were meant to examine the solubility of cerussite ($PbCO_3(s)$), though at the end of some experiments, the authors unexpectedly noted the presence of abellaite ($NaPb_2(CO_3)_2(OH)(s)$), suggesting the possible contamination of their solutions (Jang et al. 2021). Although Jang et al. (2021) re-performed the experiments under similar conditions and established the validity of the previous experiments, their examination of the previous data reveals several instances in which instrument detection limits were not established and lead concentrations below these limits were nonetheless reported and used in published data (e.g., Xiong 2015). SNL also identified significant flaws in the experimental evaluation of the lead (Pb)/ethylenediaminetetraacetic acid (EDTA) system in Kirkes et al. (2014) and Xiong (2014) that are described by Jang (2021). In these experiments, Jang (2021) points out that work was flawed in part because the brines used to measure solubility were inappropriate and should have been preceded by experiments in simpler systems to sequentially evaluate systems without including too many variables. EPA's review of these documents highlights the need to establish controls to ensure data quality. Such controls could include routine instrument detection limits in experimental methodology, evaluation of data quality at the time of collection, and a more thorough evaluation of work by technical reviewers. SNL has since taken steps to start correcting course in its lead experiments, and the Agency will monitor SNL's progress in future technical reviews. Moreover, milestone reports like Kirkes et al. (2014) and Xiong (2014) should have critically evaluated the data so that any experimental issues could have been corrected.

Updates to the iron database stemmed from TP 08-02 (Ismail et al. 2008) and AP-176 (Jang and Domski 2016). A 2016 analysis report from this work (Jang and Kim 2016) was prepared originally for a previous version of the iron database (FM2). This report appropriately provided the calculations and parameter fitting checks for many of the iron updates. Jang and Kim (2016) also assessed the potential impact of iron on EDTA in the repository (and ultimately Am(III) solubility) by calculating how complexation with sulfide may impact available iron. Although SNL ultimately chose to include only the Fe-Na-Cl-H_2O and Fe-Na-Cl-CO_3-H_2O systems for FM4, future database updates may include Fe-EDTA updates. Jang and Kim (2016) made it clear that SNL is aware of the potential downstream consequences of this addition to solubility. If SNL chooses to update the database to prioritize Fe-EDTA without including sulfide and other species that will preferentially complex with iron, it would be important to

provide detailed and clear justification given the documented understanding of the impacts described in Jang and Kim (2016). In general, SNL must make sure that incomplete systems added into the database have been evaluated for their downstream effects.

EPA had further discussions with SNL staff C. Miller, P. Domski, J. Jang, and M. Nemer on Wednesday, December 7, 2022, to clarify the database update process. Staff confirmed that the Hummel et al. (2005) and Powell et al. (2009) lead data was added to align the database with internationally recognized programs, including the NEA and the International Union of Pure and Applied Chemists, respectively. For the iron data, SNL chose to utilize data from previous experiments associated with TP 08-02 (Ismail et al. 2008) and AP-176 (Jang and Domski 2016). As noted above, these values appeared to be appropriate. Staff also explained that literature reviews were conducted to determine suitable parameter values to include in the database, but these literature reviews were not documented. Staff did not discuss how the literature reviews supported the chosen values. SNL also did not provide any additional information on how database updates are chosen and prioritized, nor did it explain why SNL decided to incorporate untested lead parameters.

4.2.2 Database Verification Exercises

SNL has previously demonstrated the ability to simulate experimental data, such as with Jang and Kim (2016), where the geochemistry modeling program EQ3/6 calculations adequately simulated experimental data for iron. EPA noted during its review of CRA-2019 that many of the FM4 database updates lacked the parameter fitting checks necessary to ensure the database could model available experimental data (EPA Docket No. EPA-HQ-OAR-2019-0534). This is especially important because +III solubility was underestimated as a result of the database updates that included inappropriate EDTA stability constants.

The Agency investigated how database verification exercises were incorporated into SNL's methodology. AP-182 provides some steps for database verification exercises in its acceptance criteria: "During the testing phases of the preliminary and interim databases comparisons between fitted parameters and EQ3/6 output will be made in Microsoft Excel" (Jang 2019). Domski (2019) does not provide any documentation demonstrating verification exercises according to this AP, while Domski (2018) provides only an incomplete assessment, that is, $Ca_2EDTA•7H2O$ solubility was simulated in NaCl solutions but not $MgCl_2$ solutions. DRCs associated with both documents do not provide any additional assessments of the parameter additions and are mainly editorial in nature. SNL document reviews need to confirm that testing and analysis follow the steps outlined in existing APs.

A new analysis plan, AP-200 (Domski 2022), appears to now provide a clear step for testing and verifying new and/or updated parameters. EPA is uncertain whether SNL intends to apply these added steps to literature values that have been added to the database, but EPA believes this will be necessary. EPA believes that this revised AP may address many of the database testing issues it identified during its technical review. Future Agency reviews will confirm that verification steps have been incorporated into SNL procedures and conducted as needed.

4.3 Corrosion

EPA's review of SNL's iron corrosion studies found adequate documentation that justified the work, evaluated the data produced, and compared the experiments to the broader WIPP project. Table 3 provides the list of documents reviewed for this topic. It also demonstrated where SNL successfully utilized the document review process.

Table 3 Iron Corrosion-Related Documents Reviewed

Document Type	Title/Description	Citation	DRC Forms
Test Plan (TP 06-02)	Iron and Lead Corrosion in WIPP-Relevant Conditions TP Rev. 3	Sisk-Scott and Icenhower (2016)	ERMS 565449 (Technical) ERMS 565450 (QA) ERMS 565451 (Management)
Milestone Report	Milestone Report for Iron and Lead Corrosion in WIPP-Relevant Conditions: Static and Solution Flow-Through Corrosion Experiments (Progress Report for the Period between September 2017 through September 2019)	Sisk-Scott and Icenhower (2019)	Not examined by EPA
Milestone Report	FY20 Milestone Report on TP 06-02 Rev. 3	Zhang (2020)	ERMS 574286 (Technical) ERMS 574287 (QA) ERMS 574288 (Management) ERMS 575339 (Memo for correcting typographical error on NP 6-1-1) ERMS 575340 (Memo for correcting typographical error on NP 6-1-1)

4.3.1 Documentation

Iron corrosion studies fall under TP 06-02 (Sisk-Scott and Icenhower 2016). The TP gives a comprehensive review of previous work on corrosion and makes thermodynamic calculations to identify the potential compounds that can exist under WIPP conditions. It also justifies why experiments are being performed (i.e., experiments "are designed to more effectively determine the dissolution rate of low carbon steel in aqueous solution over an ionic strength and a_{HS^-} [hydrogen sulfide activity] interval" and that a "greater emphasis on characterization of the starting steel coupons will also be enacted").

The first milestone report under this TP is Sisk-Scott and Icenhower (2019). The results of the study appear to be well documented:

- The text explicitly distinguishes the study from previous WIPP work (i.e., "to quantify the steel corrosion rates and $H_2(g)$ production rates as a function of chloride and sulfide concentration and temperature"), thereby justifying the work and its impact to the WIPP PA.

- The report evaluates its own data through the use of calculations. For example, the authors use the data they present in Table 5 (Sisk-Scott and Icenhower 2019) to calculate average surface retreat rates and are able to conclude that, as a result of the long amount of time elapsed between sample collection and measurement, gas loss occurs and that "the full inventory of generated H_2 was not captured in our sampling."

- Similarly, Sisk-Scott and Icenhower (2019) compare some of their own corrosion rates to previous experiments, such as experiments evaluating steel pipelines that are impacted by sulfide produced by sulfate-reducing bacteria. The authors also compare their calculated corrosion rates to WIPP values, allowing for a synthesis of the data, further understanding of WIPP processes, and a contribution to the general knowledge of the WIPP.

4.3.2 SNL Technical Review

The next milestone report in this TP is Zhang (2020), which presents data collected from electrochemical methods that were described briefly in TP 06-02:

- Data reported from electrochemical methods appear very preliminary, especially since the report notes that the tests have "an unknown degree of error associated with them and may result in some concerns on the Tafel slopes."[1]

- An examination of the technical DRC forms for Zhang (2020) (ERMS 574286) indicates that a thorough, in-depth review of the document was performed in which comments that addressed uncertainties in the work were incorporated in the milestone report:

  - For example, Comment 19 of ERMS 574286 asks a fundamental question regarding issues with the experimental setup of the electrochemical experiments.

  - The resultant response was that "since the Tafel slopes from PDP [potentiodynamic polarization, i.e., the electrochemical test performed] test are questionable, no corrosion rate will present [sic] in this report."

  - The author of the milestone report also subsequently took out a section of the report and several figures. Comments from this DRC form appear to have appropriately kept portions of the report from reaching unsupported conclusions with the given data and helped establish the preliminary nature of the work.

ERMS 574286 highlights the importance of an independent technical reviewer who will critically assess the data so that incorrect conclusions are not made.

Based on the review of the documents related to iron corrosion studies, EPA observes that much of the iron corrosion experiments and analyses were adequately documented in accordance with SNL procedures. TP 06-02 provides a thorough analysis of previous work on the topic, thermodynamic calculations and clear experimental goals. In their milestone report,

[1] A Tafel plot provides a method of determining corrosion rates from electrochemical impedance spectroscopy. It is one of a number of methods SNL uses for corrosion tests.

Sisk-Scott and Icenhower (2019) explicitly justify the experiments they included, review the work presented in the context of previous WIPP-related work, present uncertainties and statistical treatments of the data, and analyze the quality of the measurements and data. Zhang's (2020) technical review provides an instance in which a critical assessment of the work kept the milestone report from including any dubious findings.

4.4 Mineral Fragment Colloids

EPA's review of the mineral fragment colloids experimental work under TP 14-04 (Kirkes 2018) highlights SNL's need to establish robust scientific standards that ensure data quality. It also highlights the importance of technical reviewers able to critically assess the work. Although the data from these experiments were not used for any recent PAs, mineral fragment colloids are an active topic of investigation by SNL and may be incorporated into future PAs:

- Technical reviewers need to evaluate the validity of the experimental design and the assumptions, as well as the quality of the data.

- EPA identified multiple instances in this particular technical review (e.g., Kirkes 2020, Kirkes et al. 2014, Xiong 2014) in which the quality of data was severely degraded due to lack of scientific rigor. Notable issues include the absence of detection limits in methodology, flawed experimental designs and faulty assumptions. EPA recommends SNL take steps to improve its review of scientific rigor in geochemistry experimental and modeling work such that the quality and integrity of data are ensured before it is reported.

4.4.1 Issues of Scientific Rigor

TP 14-04 is broadly written, presumably to provide flexibility in experimental design. It provides the list of minerals to be synthesized and examined in experiments, brines to be used and possible instrumentation. Listed minerals include minerals that were previously characterized for the WIPP Compliance Certification Application (CCA). DRCs associated with the TP provided very few comments. Table 4 lists the documents EPA reviewed on this topic.

Table 4 Mineral Fragment Colloids Documents Reviewed

Document Type	Title/Description	Citation	DRC Forms
Test Plan (TP 14-04)	Experimental Investigation of Stability of Mineral Colloids Under WIPP Conditions, Rev. 3	Kirkes (2018)	ERMS 570241 (Technical) ERMS 570242 (QA) ERMS 570243 (Management
Milestone Report	Fourth Milestone Report on TP 14-04, Rev. 3, "Experimental Investigation of Stability of Mineral Colloids Under WIPP Conditions"	Kirkes (2020)	ERMS 574418 (Technical) ERMS 574419 (QA) ERMS 574420 (Management)
Analysis Report	Analysis Report on Mineral Fragment Colloids Under WIPP Relevant Conditions	Hora et al. (2021)	ERMS 576111 (Technical) ERMS 575112 (QA) ERMS 576113 (Management)

Document Type	Title/Description	Citation	DRC Forms
Lab Notebook 1	WIPP – Colloids – 1	ERMS 572222	ERMS 566670 (Technical) ERMS 566671 (QA) ERMS 568517 (Technical) ERMS 568518 (QA) ERMS 568519 (Technical) ERMS 568520 (QA) ERMS 569381 (Technical) ERMS 569382 (QA)
Lab Notebook 3	WIPP – Colloids – 3	ERMS 576247	ERMS 576248 (Technical) ERMS 576249 (QA) ERMS 576250 (Technical) ERMS 576251 (QA) ERMS 576252 (Technical) ERMS 576253 (QA) ERMS 576254 (Technical) ERMS 576255 (QA)

Kirkes (2020) is the milestone report associated with Revision 3 of TP 14-04 (this is labeled as the fourth milestone report, as there have been other milestone reports associated with previous revisions of TP 14-04). In these experiments, WIPP brines of various dilutions were pre-equilibrated with strontianite ($SrCO_3$) or calcite ($CaCO_3$). After equilibration, 0.2 micrometer aliquots of filtered brines were sampled for calcium or strontium concentrations as a "baseline" value before experimentation. Samples were then mixed in with colloidal versions of the minerals, which were created via grinding methods. Samples were jostled and then sampled and immediately acidified. Sampling time intervals ranged from minutes to days. Colloidal concentrations were measured using either aqueous Ca^{2+} or Sr^{2+} as surrogates, though particle sizes were also measured. EPA noted significant issues in the report, including a lack of critical evaluation of the data and problems with the experimental design and assumptions. For example—

- The documented experiments appear to repeat work previously done for the CCA on strontianite and calcite without clear justification (e.g., Papenguth 1996).

- The initial saturation steps to establish baseline concentrations did not provide any measurements to indicate whether the solutions were at equilibrium.

- The experimental design used solutions that were initially saturated with respect to the mineral being examined. To determine colloid formation, the experimenter measured colloid particle size and constituent concentrations (e.g., Ca^{2+}). However, these specific measurements did not account for any mineral precipitation that may have occurred once colloidal particles were introduced into solutions.

  – The report makes several assumptions based on the data, for example surmising that increases in particle size after experiments may be "due to the agglomeration

of the colloids in an unstable colloidal suspension." However, these assumptions do not consider the impact of experimental artifact, namely the impact of mineral precipitation in a supersaturated solution increasing colloid sizes.

- EPA reviewed the lab notebooks and DRCs associated with this study and found that the reviewers of the lab notebooks did not call attention to the items listed above.

The technical DRC for the milestone report (ERMS 574418) was able to identify and question some of the underlying assumptions in the report. The reviewer in the technical DRC questions the potential impacts from equilibrating brines with mineral solids. The author's response to the review dismisses the concern and provides no suggested way to address this issue beyond that the equilibration step is crucial to experiments to prevent colloids from dissolving. The reviewer accepted this response and the milestone report was published.

Following the milestone report, an analysis report was written (Hora et al. 2021). This report is significantly more critical of the data collected under this TP and outlines the implications of the data and some of the limitations of the analysis performed, and it suggested new approaches for future experiments. For example, the analysis report notes that the data found in Kirkes (2020) have error bars that extend below zero (i.e., data are physically meaningless), indicating that the experiments did not recognize the detection limits of their instrumentation. These issues were not identified in any of the lab notebook DRCs. Although the analysis report was able to identify crucial problems with the data to prevent their usage in the WIPP PA, these measurements should have been examined at the time of data collection and in the milestone report so that the appropriate adjustments could have been made to the study.

Based on its review of Jang (2021) on lead solubility experiments, EPA has identified other instances (e.g., Kirkes et al. 2014, Xiong 2014) in which data were not critically assessed and instrumentation detection limits were not established before data collection (see Prioritization and Performance of Database Updates in Section 4.2). Documentation suggests SNL has become aware of these problems and is taking steps to address them. However, some of the data have already been published, such as in Xiong (2015), which used flawed lead data to derive Pitzer parameters. EPA is concerned that SNL is not catching mistakes or procedural deviations in its current review process and believes that stronger adherence to the minimum criteria for a technical reviewer listed in Section 2.3 of NP 6-1 can improve the quality of the analyses and data that are published. Although these particular experiments were not included in the WIPP PA, EPA recommends SNL verify that its current methodologies and past work include processes that ensure data quality. EPA also recommends that the Xiong (2015) paper be reevaluated to determine whether it should be formally withdrawn from the published literature. EPA also reaffirms its recommendation that a broad pool of reviewers be available who can critically assess experiments, including data collection.

The result of experiments under TP 14-04 are recommendations for updates for more robust, higher quality analyses. The new TP 22-02 examines the mineral fragment colloids related to the magnesium oxide engineered barrier and incorporates several of the recommendations from Hora et al. (2021), such as uncertainty analyses for each sampling timestep. In future reviews, EPA will continue to examine and confirm whether these updated procedures have created improvements on data collection and quality.

4.5 TAUFAIL

TAUFAIL is a parameter used in WIPP PA to represent the shear strength of degraded repository waste. Because little is known about the strength of degraded waste, the value of TAUFAIL in a given performance analysis is sampled from a uniform distribution because only the upper and lower bounds need to be specified. Repository performance results are most sensitive to the lower bound, and the shear strength of a very weak material, San Francisco Bay mud, was conservatively selected as the lower bound for the initial 1996 WIPP PA. In 2009, DOE authorized SNL to conduct a series of laboratory flume tests to provide a more appropriate value for the lower bound. Testing began on August 5, 2011, and ended on October 2, 2012.

EPA selected the TAUFAIL tests for an in-depth review as an example of the application of SNL/DOE quality procedures to a geomechanical rather than a geochemical laboratory testing program. Despite their age, the TAUFAIL tests were selected for review because of their importance to the cavings release pathway in the WIPP PA and because SNL has completed little other geomechanical laboratory testing for the WIPP program since the CCA. Geomechanical laboratory testing programs are currently underway at SNL focused on creep closure of empty WIPP repository drifts and the endpoint properties of healed Salado halite. The updated procedural controls governing these current programs are little changed from the controls in effect at the time of the TAUFAIL tests. This allows the current programs to employ lessons learned from any concerns identified in the Agency's detailed review of the TAUFAIL tests.

Table 5 lists the documents EPA reviewed for the TAUFAIL tests, and Table 6 lists key SNL NPs in effect at the time the tests were conducted. A summary of the Agency's technical review of the TAUFAIL tests is presented in EPA (2017). EPA's review of the TAUFAIL experiments shows that issues with the experiments were well documented in scientific notebooks and that implementation of the test equipment was adequately performed. However, the Agency also identified instances in which experiments proceeded despite lapsed calibrations that were not identified by document reviewers. EPA also noted that SNL did not identify a CAQ regarding inattentiveness to equipment handling, which could have allowed for corrective action.

4.5.1 Test Plan

The TAUFAIL laboratory study was performed under SNL TP 09-01 (Roberts and Herrick, 2009). The plan was approved in July 2009 and became effective the following month. It was prepared following the requirements of SNL NP-20-1, Revision 5 (Johnsen 2008a), which became effective in September 2008. The procedure applies to the planning and documentation of field and laboratory experiments, as well as to testing and experimental activities that produce data. In summary, the procedure describes the responsibilities of the principal investigator (PI) and prescribes the format, review and approval process, change controls, and records keeping requirements. Appendix A of the procedure provides a detailed description of TP content and includes requirements for identifying sample control and data quality control (QC) measures, training requirements, and health and safety requirements.

EPA found that the TP provided generalized descriptions of the test procedures, some of which were modified during the testing program. EPA found this process to be acceptable because the testing equipment and procedures were new and innovative. Procedural refinements were to be

expected and were appropriately documented in the scientific notebooks (SNs) (SNL Records Center 2013) and the project report (Herrick et al. 2012).

Upon completing a detailed review, the Agency found that TP-09-01 met the requirements of the procedure.

Table 5 TAUFAIL Documents Reviewed

Document Type	Title/Description	Citation	DRC Forms
Test Program Recommendation	Recommendation for the Lower Limit of the Waste Shear Strength (Parameter BOREHOLE: TUFAIL), Rev. 0. ERMS 546033	Herrick et al. (2007)	
Test Plan TP 09-01	Waste Erodibility with Vertical and Horizontal Erosion Flumes, Rev. 0. ERMS 557149	Roberts and Herrick (2009)	
Surrogate Degraded Waste Sample Preparation	Description and Evaluation of a Mechanistically Based Conceptual Model for Spall. SAND97-1369	Hansen et al. (1997)	
Electronic Data Acquisition System	Data Report for Analysis Plan for Demonstration Test Process: Soil Flume Sixnet Data Acquisition System. ERMS 555892	Schuhen (2011a)	
Scientific Notebook	Scientific Notebook FLM-1. In Section 5.3, *Records Keeping*, TAUFAIL Records Package. ERMS 556992	SNL Records Center (2013)	ERMS 556994 (Technical) ERMS 556995 (QA) ERMS 558730 (Technical) ERMS 558731 (QA)
Scientific Notebook	Scientific Notebook FLM-2. In Section 5.3 *Records Keeping*, TAUFAIL Records Package. ERMS 556992	SNL Records Center (2013)	ERMS 558722 (Technical) ERMS 558723 (QA)
Scientific Notebook Supplement	Scientific Notebook Supplement. In Section 5.3 *Records Keeping*, TAUFAIL Records Package. ERMS 556992	SNL Records Center (2013)	ERMS 558730 (Technical) ERMS 558731 (QA)
Project Report	Determining the Hydrodynamic Shear Strength of Surrogate Degraded TRU Waste Materials as an Estimate for the Lower Limit of the Performance Assessment Parameter TAUFAIL, Rev. 0. ERMS 558479	Herrick et al. (2012)	ERMS 558480 (Signature authority for document and DRC) ERMS 558481 (Technical) ERMS 558482 (Signature authority for document and DRC) ERMS 558483 (Technical) ERMS 558484 (QA) ERMS 558485 (Management)
Follow-up Review of Test Results	Follow-up to Questions Concerning TAUFAIL Flume Testing Raised during the November 14–15, 2012 Technical Exchange between the DOE and EPA. ERMS 559081	Herrick and Kirchner (2013)	

Table 6 Key Controlling Procedures

Procedure Topic	Title/Description	Citation
Document Review	NP 6-1 Document Review Process, Rev. 8. ERMS 552261	Chavez (2009)
Measuring and Test Equipment	NP 12-1 Control of Measuring and Test Equipment, Rev. 8. ERMS 554113	Johnsen (2010)
Samples and Standards	NP 13-1 Control of Samples and Standards, Rev. 5. ERMS 550369	Johnsen (2008b)
Corrective Action	NP 16-1 Corrective Action, Rev. 6. ERMS 546325.	Davis (2007)
Test Plans	NP 20-1 Test Plans, Rev. 5. ERMS 549936	Johnsen (2008a)
Scientific Notebooks	NP 20-2 Scientific Notebooks, Rev. 9. ERMS 555311	Nielsen (2011)
Data Acquisition System Calibration	SP 12-4: Sixnet DAS Calibration, Rev. 2. ERMS 555064	Schuhen (2011b)

4.5.2 Scientific Notebooks

SNs gave the Agency insight into the day-to-day documentation of test implementation and results. EPA also reviewed scientific notebook supplements (SNSs), which are a collection of documents such as computer output, magnetic media and large drawings that may not be appropriate for archival in an SN.

The SN for the TAUFAIL flume tests was prepared following the requirements of SNL NP-20-2, Revision 9 (Nielsen 2011). In summary, the procedure identifies the initial qualification and training requirements that must be completed before recording observations or data in the SNs; the process to be followed for initiating an SN; the requirements for entering the date of work performed and the date of entry; the process for making corrections, supplements or changes; and the identification of supporting documentation, such as magnetic media, that may not be suitable for display in a notebook. The procedure also describes responsibilities for the security of the SNs, technical and QA reviews, comment resolution, and closure. Appendix A of the procedure provides a comprehensive list of requirements for the SNs themselves (bound with consecutively numbered pages, acceptable inks, etc.), required introductory information (unique identifying number, initiation date, PI name, statement of work objectives, etc.), and requirements for the entry of technical information in the body of the SN (equipment identification and calibration, description of work performed, signed and dated daily entries, etc.). Appendices B and C of the procedure provide detailed checklists for technical reviews of laboratory and field notebooks.

The SNs for the TAUFAIL flume tests were identified as FLM-1 and FLM-2 and are archived in the TAUFAIL Records Package (SNL Records Center 2013). The SNs appropriately document early problems encountered due to the uniqueness of the testing program and the learning curve for handling test equipment. SNL conducted technical and QA reviews. The first 116 pages of FLM-1 were reviewed in January and February 2012 (Chapin 2012; Nielsen 2012a); pages 119 to 157 (the last page) of FLM-1 were reviewed in November 2012 (Kicker 2012a; Nielsen

2012b), as were all 24 pages of FLM-2 (Kicker 2012b; Nielsen 2012c). The intervening pages 117 and 118 of FLM-1 document the technical and QA reviews. The Kicker (2012a) and Nielsen (2012b) technical and QA reviews also included the SNS, the data analysis Excel spreadsheets and the archival CD. In all instances, the reviewers accepted the originator's responses.

The following problems were encountered during the test program and were documented in the SNs, SNS, the Excel spreadsheets, and/or the project report:

- The TP (Roberts and Herrick 2009) describes the original intent to measure the shear strength of 50% and 100% surrogate degraded waste samples. However, the SNs describe difficulties encountered in testing the weak 100% degraded surrogate test samples (a reliable test result was obtained for only one 100% sample), and a supplemental series of 75% degraded samples was introduced to provide an intermediate sample set. This deviation from the original TP did not constitute a change in test methodology. It was explained in the project report where the test program and results are documented (Herrick et al. 2012). The Agency found this documentation to be acceptable.

- The weaker samples did not move smoothly into the flume. The investigators noted this as a "stick-slip" phenomenon that could have disrupted and weakened the samples and made some test results unacceptable.

- Air bubbles in the flume were observed to impact the weaker samples and may have affected the calculated shear stress. This contributed to making some test results unacceptable.

- The electronic Data Acquisition System (DAS) buffer was filled and no data were recorded for four tests of 75% degraded samples. Test results for these samples were evaluated using SN entries. This problem is further discussed below under EPA's review of corrective actions.

- Two 75% degraded samples continued to extrude into the flume during lunch breaks, resulting in the loss of data.

- Sample integrity may have been affected by a vacuum created when removing the shipping platens. The sample holder design was changed to correct this problem.

- The DAS calibration was found to have expired before the test program was completed. This problem is discussed further below.

EPA observes that problems are not unexpected with innovative testing equipment, that the problems that were encountered appear to have been frankly documented, and that their potential effects on test results were qualitatively explained. As previously noted, some test results were discarded by the investigators because of these problems. While these problems were unfortunate and potentially affected the quality of the results, the Agency found no fault in the documentation.

4.5.3 Measuring and Test Equipment

The measuring and test equipment used in the TAUFAIL tests were controlled under SNL NP 12-1, Revision 8 (Johnsen 2010). This procedure applies to equipment used to monitor, measure, test and collect data. The requirements include equipment use, care and maintenance, including unique identification, calibration, and records handling.

An electronic DAS using Sixnet Corporation SixTRAK software was used for the TAUFAIL tests. Schuhen (2011a) describes the test procedures, equipment, DAS, and qualification of the DAS for acquiring the data. The Sixnet software was used to configure system hardware and facilitate communication between the DAS hardware and the computer. Because Sixnet is commercial software used without modification, it was not necessary to meet the software requirements of NP 19-1 (Long 2020). Instead, the software was subjected to a series of test cases, including, for example, the system's ability to accurately measure an analog input signal and convert the value to engineering units. SNL SP 12-4, Revision 2 (Schuhen 2011b), describes the calibration requirements for the Sixnet DAS. Section 2.4 of SP 12-4, Revision 2, addresses calibration frequency. It states the following:

> *The Sixnet Controller and connected analog modules will be calibrated semi-annually (six month intervals). The results of the calibration will be documented on the applicable SN or SNS. The principal investigator can elect to lengthen or shorten the calibration interval based on the results of previous calibrations and the stability of the equipment. Any deviations from the calibration interval will be justified and noted in the applicable scientific notebook or SNS.*

DAS calibration reports are documented in Section 3.1 of SNS #1 (SNL Records Center 2013). EPA's review concluded that the DAS system was implemented appropriately, except for two concerns related to system calibration:

- An SNL QA review dated November 8, 2012 (Nielsen 2012b), found that the calibration of the DAS had expired on June 5, 2012, and that it was not recalibrated until July 10, 2012. However, the reviewer noted that "because none of the 100% degraded samples were used in the final report no action is necessary." In addition, the person requesting the review (identified on the SNL DRC form as the review requester) responded that an as-found calibration check demonstrated that the system had remained "well within tolerance." The reviewer accepted this response.

- Section 3.1 of the SNS has three calibration reports with calibration dates of March 15, 2011, December 5, 2011, and July 10, 2012. A final as-found calibration check was performed on September 11, 2013, and is documented on page 25 of the FLM-2 SN (SNL Records Center 2013). Although this final check demonstrated that the system had remained within tolerance, EPA found that the calibration intervals were all greater than 6 months and did not find justification or notation of these deviations in the SN or SNS, as required by procedure.

In the first case noted above, the SNL reviewer concluded that the missed calibrations had no effect on the test results used in the final report because the potentially affected tests had been

discarded for other reasons, and therefore no remedial action was necessary. In the second case noted above, the Agency found no evidence that the missed calibration dates had been detected. Although the Agency agrees that the test system had fortuitously remained in calibration and the missed calibrations apparently did not impact the final data set, the inattention to calibration requirements throughout the testing indicates a systemic lapse in measuring and test equipment maintenance and a deviation from the requirements of NP 12-1 and SP 12-4. This issue is discussed further under Corrective Action below.

4.5.4 Control of Samples and Standards

The surrogate degraded waste samples used in the TAUFAIL tests were controlled under NP 13-1, Revision 5 (Johnsen 2008). This procedure is intended to ensure that samples and standards are identified and controlled in a manner consistent with their intended use. The preparation and handling of the surrogate waste samples are described in the TP (Roberts and Herrick 2009), the SNs (SNL Records Center 2013) and the project report (Herrick et al. 2012). In summary, the samples were prepared at SNL in Albuquerque following the methods developed by Hansen et al. (1997) and accepted by EPA for CRA-2014. The samples were hand-carried to Carlsbad in sealed containers. EPA found the control of samples for the TAUFAIL tests to be acceptable.

4.5.5 Document Review

SNL's document review process applicable to the TAUFAIL tests is described in NP 6-1, Revision 8 (Chavez 2009). This procedure establishes requirements for conducting all types of reviews and documenting the resolution of comments. Examples of the implementation of this procedure were reviewed by the Agency as applied to the SNs and project report for the TAUFAIL tests and are described above. Because of the apparent inability of SNL's QA/QC review protocols to flag the early calibration problems identified by the Agency, EPA found the application of this procedure to the TAUFAIL tests to be insufficient in frequency and depth.

In comparing the aforementioned calibration dates with the testing dates, EPA found that the test start date of August 5, 2011, was approximately 5 months after the first recorded calibration date of March 15, 2011, and that the first recalibration was therefore due by September 15, 2011, 1 month after testing began. However, the first recalibration apparently did not occur until December 5, 2011, about 4 months after testing began and nearly 3 months after the recalibration was due. The next recalibration was due by June 5, 2012, but this recalibration was also missed and was performed 1 month late, on July 10, 2012. EPA observes that the first missed recalibration was not identified and was apparently not detected in the first SNL QA/QC reviews in January and February 2012, and the second missed recalibration was not identified until November 2012, 1 month after testing ended. Whether the missed recalibrations were not detected or were identified only after testing ended, the frequency and depth of the review cycle were insufficient for a timely identification of potentially serious problems with test results. This issue is discussed further under Corrective Action below.

4.5.6 Corrective Action

SNL's corrective action procedure applicable to the TAUFAIL tests is described in NP 16-1, Revision 6 (Davis 2007). This procedure establishes requirements for identifying, documenting, evaluating, preventing, controlling and correcting CAQs. According to the procedure, "[a] CAQ is a deviation from a requirement, a deficiency, or some other condition that adversely impacts the quality of a process or product including failures, malfunctions and technical inadequacies." EPA found two instances in which CAQs could have been but were not identified by SNL. The first involved filling of the DAS buffer and loss of detailed test data, and the second involved expiration of DAS calibrations. In each case, the investigators continued testing for several weeks or more, apparently unaware of the problems:

- According to SN FLM-1 entries beginning on page 131, on August 28, 2012, the investigators became aware that the DAS buffer had been filled and no data were being recorded. This lapse affected four tests of 75% degraded samples. Test results for these samples were instead evaluated using manual data from the SN. Given that the SNs functioned as intended by providing a redundant data source, EPA considers the CAQ in this instance to be the investigators' lack of awareness of the problem (i.e., a deficiency or technical inadequacy per NP 16-1) rather than an impact on test results. EPA believes that the investigators should have had an ongoing awareness of DAS buffer capacity limitations. It might have been possible to avoid this issue if an increased emphasis on systematic checks of DAS integrity and archival had been included in the training module for NP 20-2 on SN and/or other appropriate training modules.

- Per Section 2.4 of SP 12-4, Revision 2, the issue of missed calibrations could have been addressed procedurally if the PI had justified lengthening the calibration interval in advance of need. However, this was not done because the PI was apparently unaware that the calibration interval had been or was about to be exceeded and the SNL QA/QC reviews did not identify the missed calibrations in a timely manner. The SNL QA review of November 8, 2012, found that the calibration of the DAS for the TAUFAIL tests had expired on June 5, 2012, some 5 months earlier (Nielsen 2012b). In addition, EPA's review found that earlier calibration intervals seemed to have exceeded the prescribed 6 months without being detected and without the justification or notation required by the governing procedure (Schuhen 2011b).

- Agency personnel discussed the effects of these lapses in calibration with the SNL reviewer during the Agency's December 2022 in-depth technical review and were told that a CAQ was not issued because the lapses did not impact the quality of the data used in the final report. However, EPA considers that the investigators' lack of attention to system calibration requirements, and the failure to identify the early missed calibration in QA/QC reviews, represent CAQs under the requirements of NP 16-1, Revision 6. This procedure addresses preventing CAQs as well as identifying, documenting and evaluating them. EPA believes that the lack of impact on test results appears to have been fortuitous and that maintaining a quality testing program should not have to rely on luck. The investigators should have been aware of the DAS calibration schedule. While SNL did not consider this event to be a CAQ, the situation could potentially have been avoided by an increased frequency and depth of QA/QC reviews, and an increased emphasis on

systematic checks of DAS integrity in the training modules for NP 12-1 and NP 20-2 and/or other appropriate training modules.

4.5.7 TAUFAIL Test Conclusions

EPA's in-depth review of SNL's TAUFAIL testing program found that over 99% of the hundreds of requirements in the six principal governing procedures were fully met. Although the Agency's concerns are not numerous, they address lapses of attention to the functioning of the DAS and improvements in the frequency and depth of QA/QC reviews. These concerns are especially important in cases where experiments continue despite expired calibrations. The proper functioning of the test system is critical to documenting the test results and therefore also critical to the success of the test itself. Should similar situations arise in the future, EPA's review identified the need for improvements in the following areas:

- Increased attention to the data logging buffer capacity;
- Increased attention to system calibration schedules;
- Increased frequency of QA/QC reviews; and
- More stringent, in-depth QA/QC reviews of test equipment, conduct and data.

In each case, the problems were overcome either by the use of redundant, manually entered backup data, by demonstrations that the system had remained in calibration, or because the potentially affected test results were discarded for other reasons. As a result, the consequences to the TAUFAIL test program were minimized, and SNL concluded that no further actions were required. SNL's candor in including and reporting test problems in the SNs in detail, rather than to omit or obscure them, is a strong example of good documentation practices in adherence to the experimental procedures. These experiments would have been strengthened had SNL performed a similar level of documentation for the early calibration lapses. EPA considers the investigators' repeated inattention to DAS system operation to represent a systematic CAQ that could negatively impact future testing programs. Identifying this inattention as a CAQ would have raised awareness of the problem and could be corrected in future test programs, for example, by an increased emphasis on systematic checks of DAS integrity in the training modules for NP 12-1 and NP 20-2, as well as by an increased frequency and depth of QA/QC reviews.

4.6 Boreholes

WIPP regulations require that current drilling practices be assumed for future inadvertent intrusions. DOE continues to survey drilling activity in the Delaware Basin in accordance with the criteria established in 40 CFR 194.33. The drilling rate and plugging pattern parameters for the CRA-2019 PA are developed from survey results in the Delaware Basin monitoring annual report (DOE/WIPP-18-2308, Revision 1). Table 7 presents the drilling parameters used in the CRA-2019 PA.

Table 7 Drilling Rate and Plugging Pattern Parameters for the CRA-2019 PA

Material	Property	Description	Units	Distribution	Value
GLOBAL	LAMBDAD	Drilling rate per unit area	$km^{-2}yr^{-1}$	Constant	0.0099
GLOBAL	ONEPLG	Probability of having Plug Pattern 1	-	Constant	0.403
GLOBAL	TWOPLG	Probability of having Plug Pattern 2	-	Constant	0.331
GLOBAL	THREEPLG	Probability of having Plug Pattern 3	-	Constant	0.266

Update of the drilling parameter values follows NP 9-2 (e.g., Long 2023). Documents EPA reviewed related to this topic included AP-181 (Zeitler 2019, ERMS #571150), the Parameter Data Entry Form (NP 9-2-1, ERMS #571305), and the WIPP PA Database (PADB) viewer pages for GLOBAL:LAMBDAD, GLOBAL:ONEPLG, GLOBAL:TWOPLG, and GLOBAL:THREEPLG. NP 9-2-1 provided the value and justification of these four global material parameter updates for the CRA-2019 PA. It also included the name, signature and date of data entry and QA reviewer. The attached PADB viewer pages provided a clear record of the property information, version information, unit, effective date, distribution, and value of these four updated parameters. Reference documents for this version and previous versions are also listed in the PADB viewer pages. EPA's review of SNL's update of drilling rate and plugging pattern parameters for the CRA-2019 PA found that SNL provided clear documentation and followed its procedure (NP 9-2).

4.7 BRAGFLO

As part of EPA's review of DOE's peer review of conceptual changes incorporated in the additional panels PA model, EPA observed that DOE performed three-dimensional PFLOTRAN (subsurface flow and reactive transport code) simulations to confirm the two-dimensional BRAGFLO results. This approach raises the possibility that a similar method could be undertaken in future PAs to address several of the limitations identified with BRAGFLO during past PAs. These limitations include (1) the difficulty that BRAGFLO has in simulating step changes in material properties in conjunction with the activation of capillary pressures and 2) the means by which convergence criteria are set in BRAGLO until convergence is achieved without the ability to compute mass-balance.

4.7.1 Permeability Changes and Capillary Pressure

As described in a 2017 email from R. Patterson (DOE) to K. Economy (EPA), DOE has been aware of potential convergence problems since 2012 (see the attachment). EPA has also discussed the issues related to BRAGFLO convergence issues in two EPA CRA-2014 technical support documents (TSDs) (EPA 2017b, EPA 2017c). EPA's Sensitivity Studies TSD (EPA 2017b) notes (p. 5) that the BRAGFLO model would only properly converge under the following conditions:

- If the permeability is varied over time, the capillary pressure model must be turned off.
- If the permeability is constant over time, the capillary pressure model can be turned on.

DOE's solution to address these issues is to either ignore capillary forces when porosity and permeability were changing (as in run-of-mine salt panel closure system consolidation), or

vice-versa (as in assuming constant halite permeability and porosity for closed rooms beginning at time zero).

4.7.2 Convergence Criteria and Mass Balance

For some vectors, BRAGFLO has difficulty converging within the number of steps provided based on multiple factors described below. DOE's current approach is to adjust the parameter ICONVTEST to allow for this convergence. As described in the BRAGFLO Analysis Package (Stein and Zelinski 2003, p. 29)—

> *The standard settings optimize calculations under most circumstances, but occasionally BRAGFLO does not complete the calculations for individual vectors, which are referred to as exception vectors. The most common failure is that BRAGFLO calculations do not reach 10,000 years within the maximum number of time steps prescribed (10,000 time steps). Exception vectors usually result from the combination of extreme conditions of coincident sampled variables and very small grid cells (e.g., the intersection of the borehole or shaft with a marker bed). These circumstances can lead to extreme spatial or temporal gradients within the model domain that exceed tolerances specified in the input control file. These conditions cause BRAGFLO to shorten its time step. For most vectors this automatic time-step control is sufficient to solve the short-lived numerical problem, however for exception vectors it is not and it is necessary to relax, tighten, or otherwise adjust BRAGFLO input numerical control parameters in order to complete the calculations. Changing the value of the input control parameter, ICONVTEST, is the most common adjustment to BRAGFLO to allow the completion of calculations for "exception" vectors. The following excerpt from the user's manual (WPP PA, 2003a) explains when the standard value, "1", should be changed to "0". ICONVTEST: Flag specifying whether either or both convergence criteria must be satisfied before a solution is considered to have converged.*

However, as noted in the BRAGFLO User's Manual (Day and Zeitler 2019, p. 119), adjusting convergence criteria can result in important mass balance issues: "Caution is advised when using this feature because large mass balance errors can be introduced which may not be apparent in a cursory examination of the results."

In EPA's review of DOE's BRAGFLO-related documents, the Agency has learned that in some simulations DOE has relaxed the convergence criteria in BRAGFLO to facilitate convergence. However, this has not been the case for the PAs submitted to EPA for CRA-2019 and CRA-2014. EPA believes this is an important technical issue to resolve to improve the defensibility of PA calculations.

4.7.3 Summary and Potential Resolution

Since BRAGFLO is known to have two-phase flow issues and PFLOTRAN may not, PFLOTRAN could provide a means to independently check whether the two-phase flow and convergence issues could lead to significant mass-balance errors in BRAGFLO. To perform this

comparison, PFLOTRAN could be set up in two dimensions to match the two-dimensional BRAGFLO formulation in a similar fashion to how the two codes were used to compare the two-dimensional BRAGFLO versus the PFLOTRAN three-dimensional model. EPA will work with DOE to develop a path forward to resolve these issues. Concurrently, SNL and DOE should also work towards the use of PFLOTRAN or another qualified replacement model for BRAGFLO, as the Agency has stipulated in its review of the CRA-2019 (EPA Docket No. EPA-HQ-OAR-2019-0534).

5.0 CONCLUSIONS

5.1 Observations

In its review of several select topics, EPA has identified examples where SNL has demonstrated good documentation and review practices:

- SNL's iron corrosion studies provide examples of well-documented work. The DRC for Zhang (2020) was highly critical and consequently prevented inappropriate conclusions from being reported. Earlier corrosion studies (Sisk-Scott and Icenhower 2019) also critically assessed the data generated, examined uncertainties in the experiments, and justified the work performed.

- Challenges in the TAUFAIL experiments were candidly and explicitly documented in SNs in adherence to their experimental procedures, increasing the transparency and traceability of the work.

- The presentation and explanation of SNL's update of drilling rate and plugging pattern parameters for the CRA-2019 PA provided for the in-depth technical review by SNL gave clear documentation and demonstrated that procedures were followed.

5.2 Observations and Recommendations

The Agency has some observations from its onsite technical review that ultimately stem from the lack of adherence to the procedures listed in NP 6-1. The requirements in NP 6-1 appear reasonable, but they could be supplemented as discussed below, and the technical reviewers need to adhere to them. EPA identified multiple instances in which SNL reviewers did not address the minimum criteria for a technical reviewer in Section 2.3 of NP 6-1, and the question arises on how broadly applicable this issue is within the entirety of the WIPP PA program. These minimum criteria include answering the questions:

- *Are the objectives clearly stated and fulfilled?*
- *Is the technical activity clearly described?*
- *Are equations/calculations accurate?*
- *Does logic lead to reasonable conclusions?*
- *Are the results drawn from the data supported by the data presented?*
- *Are data/tables/figures easily understood? Are legends complete?*
- *Are assumptions stated clearly and do they lead logically to the conclusions presented?*

Because SNL did not address the minimum criteria in the documents EPA reviewed, the Agency observed several problems in the documentation including milestone reports that provided insufficient technical depth, geochemistry database updates that lacked justifications and did not follow parameter fitting exercises discussed in APs, and published documents with degraded data quality (e.g., Kirkes (2020) and Xiong (2015)). Since the issues observed in what was examined by EPA staff could extend broadly across all of SNL's reviews on WIPP work, then the review issue needs to be addressed systematically by SNL. DOE should also increase its oversight of SNL work to ensure SNL follows its procedures and produces data of acceptable quality.

Additionally, SNL's document review process would benefit from increased frequency of reviews and a larger pool of qualified reviewers further removed from the work and the PIs so that reviewers can be more objective and critical of data quality and the conclusions and assumptions stemming from the work. Furthermore, EPA recommends adding other minimum criteria for technical reviewers, including assessment of the work's impact to the WIPP and the WIPP PA, assessment of the work's adherence to the related SP/TP/APs, and review of the quality of the data produced (e.g., whether measured data fall within calibrations and whether data fall within instrument detection limits):

- Reviewers for the FM4 database updates did not question the validity of the updates, nor did they question any underlying assumptions or downstream consequences to the PA (i.e., database reviews did not fulfill minimum criteria listed in NP 6-1 of: *Are objectives clearly stated and fulfilled?* and *Are assumptions stated clearly and do they lead logically to the conclusions presented?)*.

- Data quality was not closely assessed in multiple experiments, resulting in instances in which measurements below detection limits were reported and even published, such as with mineral fragment colloid and lead experiments (i.e., reviews did not fulfill minimum criteria listed in NP 6-1 of: *Are equations/calculations accurate?* and *Are the results drawn from the data supported by the data presented?*).

- In TAUFAIL experiments, reviewers missed multiple instances in which data were collected while calibrations had lapsed (i.e., results do not fulfill minimum criteria listed in NP 6-1 of: *Are the results drawn from the data supported by the data presented?*).

- SNL should have identified the repeated inattention to the DAS system in the TAUFAIL experiments as a CAQ. This would have raised awareness of the problem and could be corrected, for example, by an increased emphasis on systematic checks of DAS integrity (i.e., this does not fulfill minimum criteria listed in NP 6-1 of: *Are the results drawn from the data supported by the data presented?*).

- Milestone reports, though a measure of progress, need to provide a more rigorous analysis and interpretation of the data and their implications, as EPA reviews many of these reports during recertifications. A more detailed analysis of the work in milestone reports can allow for corrections in experimental approaches before experiments are completed. Milestone reports for lead and mineral fragment colloids experiments, for example, did not critically assess any issues with the data.

Because of this deficiency in the review process, FM4 database work was published for which SNL did not follow its own protocols and provide proper documentation. This resulted in a prolonged completeness process during EPA's review of CRA-2019 as EPA required SNL to perform additional calculations to determine the impacts of the errors in the database:

- In reviewing updates to FM4, EPA observed that there was no systematic method for updating the database. Incomplete systems such as lead were included in the database without evaluation of their downstream effects. EPA recommends SNL document the justifications for additions to the database and evaluate their potential downstream impacts to the PA.

- SNL did not perform the necessary parameter-fitting exercises when data such as EDTA parameters were added to the FM4 database, as outlined in its procedures. SNL has updated procedures to ensure this will happen with database updates, and future Agency reviews will confirm that verification steps have been incorporated into SNL procedures and conducted as needed.

Moreover, the lack of stringent technical reviews has resulted in the publication of documents with questionable scientific rigor:

- EPA recommends that SNL reevaluate Xiong (2015) in particular to determine whether it should be formally withdrawn from the published literature.

- Although the experiments reviewed in this report were not included in any PAs, EPA is concerned at the lack of consistent scientific rigor in these specific studies in which instrument detection limits were not established; methods, data, and assumptions were not critically assessed; and work was not evaluated within the context of existing WIPP studies. EPA recommends SNL critically assess its experimental designs and doublecheck that these issues have not permeated into other WIPP-related experimental work. For example, SNL should evaluate its current methodologies to ensure instrument detection limits are established, stringent reviews of test equipment are performed, data are assessed as they are collected, and calibration schedules are adhered to. EPA will work with both DOE and SNL to verify that improvements have been made.

The Agency has also reached the following observation regarding the BRAGFLO calculations:

- SNL has addressed BRAGFLO mass balance issues by relaxing convergence criteria in some cases. EPA will continue to work with DOE to determine a path forward to continue addressing the issues of the Salado flow model for future PAs, noting that the Agency's acceptance of the approach of abandonment of panel closures in the south for CRA-2019 is contingent on DOE accepting an adequate three-dimensional model that can better address the Salado flow mass balance issues (EPA Docket No. EPA-HQ-OAR-2019-0534).

Lastly, EPA notes that its completeness process during CRA reviews will be shortened if SNL is able to identify and correct technical issues during its document review process before submitting documentation to EPA. This can partially be achieved through better adherence to

NP 6-1. The Agency will plan onsite technical reviews in the future to verify any progress SNL has made in addressing the concerns listed above.

6.0 REFERENCES

Chapin, D.M. 2012. Document Review and Comment, Flume Scientific Notebook #1, FLM-1, pages 1–116. Form Number NP-6-1-1. ERMS 556994. Records Center, Sandia National Laboratories, Carlsbad, NM. January.

Chavez, M. 2009. Nuclear Waste Management Procedure NP 6-1 Document Review Process, Rev. 8. ERMS 552261. Records Center, Sandia National Laboratories, Carlsbad, NM. October.

Davis, S. 2006. Nuclear Waste Management Procedure NP 2-1 Qualification and Training, Rev. 9. Sandia National Laboratories, Carlsbad, NM. September 28, 2006.

Davis, S. 2007. Nuclear Waste Management Procedure NP 16-1 Corrective Action, Rev. 6. ERMS 546325. Records Center, Sandia National Laboratories, Carlsbad, NM. June.

Day, B., and T. Zeitler. 2019. User's Manual for BRAGFLO Version 7.00. ERMS 570275. Sandia National Laboratories, Carlsbad, NM.

Domski, P.S. 2018. Memo on the Estimation of the Solubility Product (Log Ksp) for Ca2EDTA.7H2O, and Pitzer Parameters for the Na^+ / $CaEDTA^{2-}$ Pair. ERMS 570202. Sandia National Laboratories, Carlsbad, NM.

Domski, P.S. 2019. An Update to the EQ3/6 Pitzer Thermodynamic Database DATA0.FM1 With the Creation of DATA0.FM4. ERMS 571052. Sandia National Laboratories, Carlsbad, NM.

Domski, P.S. 2022. Analysis Plan: Procedure(s) for the Creation of Thermodynamic Databases for WIPP Compliance Calculations. ERMS 578657. Sandia National Laboratories, Carlsbad, NM.

EPA (U.S. Environmental Protection Agency). 2017a. Technical Support Document for Section 194.23, EPA Review of Proposed Modification to the Waste Shear Strength Parameter TAUFAIL. DOCKET: EPA-HQ-OAR-2014-0609. Prepared by S. Cohen & Associates, Vienna, Virginia, for Kathleen Economy, Work Assignment Manager, U.S. Environmental Protection Agency, Office of Radiation and Indoor Air, Center for Waste Management and Regulations, Washington, DC. July.

EPA. 2017b. Review of EPA Sensitivity Studies of the DOE CRA-2014 WIPP Compliance Recertification Performance Assessment. Docket: EPA-HQ-OAR-2014-0609. U.S. Environmental Protection Agency, Office of Radiation and Indoor Air, Center for Waste Management and Regulations, Washington, DC. July.

EPA. 2017c. Technical Support Document for Section 194.23 Technical Review of Salt Aggregate, Disturbed Rock Zone, and Open Drift Healing Characteristics. Docket: EPA-HQ-OAR-2014-0609-0053. Prepared by S. Cohen & Associates, Vienna, Virginia, for Kathleen

Economy, Work Assignment Manager, U.S. Environmental Protection Agency, Office of Radiation and Indoor Air, Center for Waste Management and Regulations, Washington, DC. July.

EPA 2022. Technical Support Document for Section 194.24 Evaluation of the Compliance Recertification Application (CRA-2019) Actinide Source Term, Gas Generation, Backfill Efficacy, Water Balance, and Culebra Dolomite Distribution Coefficient Values. Docket: EPA-HQ-OAR-2019-0534. U.S. Environmental Protection Agency, Office of Radiation and Indoor Air, Center for Waste Management and Regulations Washington, DC. April.

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, NM.

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, NM.

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, NM.

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, Rev. 0. ERMS 558479. Sandia National Laboratories, Carlsbad, NM.

Hora, P., C. Miller, J. Jang, L. Zhang, L. Kirkes. 2021. Analysis Report on Mineral Fragment Colloids Under WIPP Relevant Conditions. ERMS 576110. Sandia National Laboratories, Carlsbad, NM.

Hummel, W., G. Anderegg, I. Puigdomenech, L. Rao, and O. Tochiyama. 2005. Chemical Thermodynamics of Compounds and Complexes of U, Np, Pu, Am, Tc, Zr, Ni and Se with Selected Organic Ligands. OECD Nuclear Energy Agency, Data Bank, Issy-les-Moulineaux, France, Elsevier B.B. Amsterdam, The Netherlands. Chemical Thermodynamics, Vol. 9, p. 1088
*Ismail, A.E., M.B. Nemer, G.T. Roselle, and Y.L. Xiong. 2008. Iron, Lead, Sulfide, and EDTA Solubilities Test Plan TP 08-02. ERMS 548467. Sandia National Laboratories, Carlsbad, NM.

Jang, J. 2021. Revisit to the Pitzer Geochemical Thermodynamic Model for the $PbCl_2$(s) – HCl – H_2O System. Memorandum to P. Hora, L. Kirkes, L. Zhang, M. Nemer, P. Domski, C. Miller. ERMS 575309. Sandia National Laboratories, Carlsbad, NM.

Jang, J-H. 2019. Analysis Plan for Derivation and Addition of Equilibrium Constants and Pitzer Interaction Parameters to the WIPP Geochemical Thermodynamic Database. AP-182. ERMS 571074. Sandia National Laboratories, Carlsbad, NM.

Jang, J.-H., and P.S. Domski. 2016. Analysis Plan (AP) AP-176, Rev. 0, Analysis Plan for the Development of a Self-Consistent Extension of the WIPP Geochemical Thermodynamic Database that Includes Aqueous Ferrous Iron Chemistry. ERMS 566478. Sandia National Laboratories, Carlsbad, NM. June 27, 2016.

Jang, J., P. Hora, L. Kirkes, C. Miller, and L. Zhang. 2021. Analysis Report Documenting the Assessment of the Solubility of Lead, EDTA and Other Organic Ligands in Non-Sulfide Systems Performed Under TP 08-02 and Under TP 20-01, Rev. 0. GEOC-21-03. Sandia National Laboratories, Carlsbad, NM.

Jang, J.-H., and S. Kim. 2016. Derivation of Pitzer Interaction Parameters and Thermodynamic Properties for the Aqueous Species of Ferrous Iron and Their Pairs, Rev. 2. ERMS 567283. Sandia National Laboratories, Carlsbad, NM.

Johnsen, S. 2008a. Nuclear Waste Management Procedure NP 20-1 Test Plans, Rev. 5. ERMS 549936. Records Center, Sandia National Laboratories, Carlsbad, NM. September.

Johnsen, S. 2008b. Nuclear Waste Management Procedure NP 13-1 Control of Samples and Standards, Rev. 5. ERMS 550369. Records Center, Sandia National Laboratories, Carlsbad, NM. November.

Johnsen, S. 2010. Nuclear Waste Management Procedure NP 12-1 Control of Measuring and Test Equipment, Rev. 8. ERMS 554113. Records Center, Sandia National Laboratories, Carlsbad, NM. August.

Kicker, D. 2012a. Document Review and Comment, Flume Scientific Notebook #1, FLM-1, pages 119–157. Form Number NP-6-1-1. ERMS 558730. Records Center, Sandia National Laboratories, Carlsbad, NM. November.

Kicker, D. 2012b. Document Review and Comment, Flume Scientific Notebook #2, FLM-2, pages 1–24. Form Number NP-6-1-1. ERMS 558722. Records Center, Sandia National Laboratories, Carlsbad, NM. November.

Kirkes, L. 2018. Experimental Investigation of Stability of Mineral Colloids Under WIPP Conditions. TP 14-04, Rev. 3.ERMS 570240. Sandia National Laboratories, Carlsbad, NM.

Kirkes, L. 2020. Fourth Milestone Report on Test Plan TP 14-04, Rev. 3, “Experimental Investigation of Stability of Mineral Colloids Under WIPP Conditions. ERMS 574417. Sandia National Laboratories, Carlsbad, NM.

Kirkes, L., T. Olivas, J. Jang, S. Kim, and Y. Xiong. 2014. Third Milestone Report on Test Plan TP 08-02, “Iron, Lead, Sulfide, and EDTA Solubilities,” Rev. 2. ERMS 562947. Sandia National Laboratories, Carlsbad, NM. November 2014.

Long, J. 2020. Nuclear Waste Management Procedure NP 19-1 Software Requirements, Rev. 19. Sandia National Laboratories, Carlsbad, NM. July 28.

Long, J. 2023. Nuclear Waste Management Procedure NP 9-2 Parameters, Rev. 3. Sandia National Laboratories, Carlsbad, NM. March 7.

Nielsen, S. 2011. Nuclear Waste Management Procedure NP 20-2 Scientific Notebooks, Rev. 9. ERMS 555311. Records Center, Sandia National Laboratories, Carlsbad, NM, April.

Nielsen, S. 2012a. Document Review and Comment, Flume Scientific Notebook #1, pages 1–116. Form Number NP-6-1-1. ERMS 556995. Records Center, Sandia National Laboratories, Carlsbad, NM. February.

Nielsen, S. 2012b. Document Review and Comment, Flume Scientific Notebook #1, pages 119–157. Form Number NP-6-1-1. ERMS 558731. Records Center, Sandia National Laboratories, Carlsbad, NM. November.

Nielsen, S. 2012c. Document Review and Comment, Flume Scientific Notebook #2, pages 1–24. Form Number NP-6-1-1. ERMS 558723. Records Center, Sandia National Laboratories, Carlsbad, NM. November.

Nielsen, S. 2022. Nuclear Waste Management Procedure NP 6-1 Document Review Process, Rev. 10. Sandia National Laboratories, Carlsbad, NM. September 9.

Papenguth, H.W. 1996. Parameter Record Package for Colloid Actinide Retardation Parameters. ERMS 238876. Sandia WIPP Central Files (SWCF). Albuquerque, NM.

Powell, K.J., Brown. P.L., Byrne, R.H., Gajda, T., Hefter. G., Luez, A.-K., Sjoberg, S. and Wanner, H., 2009. Chemical speciation of environmentally significant metals with inorganic ligands. Part 3: the Pb^{2+} + OH^-, Cl^-, CO_3^{2-}, SO_4^{2-}, and PO_4^{3-} systems (IUPAC Technical Report). *Pure and Applied Chemistry* 81 (12), 2425-2476.

Roberts, J.D., and C.G. Herrick. 2009. Waste Erodibility With Vertical and Horizontal Erosion Flumes. Test Plan TP 09-01, Rev. 0. ERMS 557149. Sandia National Laboratories, Carlsbad, NM. August.

Schuhen, M. 2011a. Data Report for Analysis Plan for Demonstration Test Process: Soil Flume Sixnet Data Acquisition System. AP 148, Rev. 0. ERMS 555892. Sandia National Laboratories, Carlsbad, NM. June.

Schuhen, M. 2011b. Special Procedure SP 12-4: Sixnet DAS Calibration, Rev. 2. ERMS 555064. Sandia National Laboratories, Carlsbad, NM. February.

Sisk-Sisk-Scott, C. 2019. Analysis Plan to Update the WIPP Geochemical Thermodynamic Database (DATA0.FM1) to Data0.FM4 for CRA-2019. AP-183, Rev. 1. ERMS 571001. Sandia National Laboratories, Carlsbad, NM. February 20, 2019.

Sisk-Scott, C., and J. Icenhower. 2016. Iron and Lead Corrosion in WIPP-Relevant Conditions Test Plan. TP 06-02. Sandia National Laboratories, Carlsbad, NM.

Sisk-Scott, C., and J. Icenhower. 2019. Milestone Report for Iron and Lead Corrosion in WIPP-Relevant Conditions: Static and Solution Flow-Through Corrosion Experiments (Progress Report for the Period Between September 2017 Through September 2019). ERMS 572066. Sandia National Laboratories, Carlsbad, NM.

SNL (Sandia National Laboratories) Records Center. 2013. TAUFAIL Records Package. ERMS 556992. Records Center, Sandia National Laboratories, Carlsbad, NM.

Stein, J.S., and W. Zelinski. 2003. Analysis Package for BRAGFLO: Compliance Recertification Application. ERMS 530163. Sandia National Laboratories, Carlsbad, NM.

Xiong, Y. 2014. Experimental and Thermodynamic Modeling of $PbEDTA^{2-}$ Interactions in NaCl and $MgCl_2$ Solutions. ERMS 562877. Sandia National Laboratories, Carlsbad, NM.

Xiong, Y. 2015. Experimental Determination of Lead Carbonate Solubility at High Ionic Strengths: A Pitzer Model Description. *Monatshefte für Chemie-Chemical Monthly*, 146: 1433–1443.

Zeitler, T.R. 2019. Analysis Plan for the 2019 WIPP Compliance Recertification Application Performance Assessment. ERMS 571150. Sandia National Laboratories, Carlsbad, NM.

Zhang, L. 2020. FY20 Milestone Report on TP 06-02 Rev. 3. ERMS 574285. Sandia National Laboratories, Carlsbad, NM.

ATTACHMENT
CORRESPONDENCE FROM R. PATTERSON (DOE) TO K. ECONOMY (EPA) DATED JULY 11, 2016

-----Original Message-----
From: Russ Patterson <russ.patterson@cbfo.doe.gov>
To: 'Economy, Kathleen' <Economy.Kathleen@epa.gov>; 'Chris Camphouse' <rccamph@sandia.gov>; 'Day, Brad A.' <baday@sandia.gov>; 'Zeitler, Todd (tzeitle@sandia.gov)' <tzeitle@sandia.gov>; 'Peake, Tom' <Peake.Tom@epa.gov>; betsy.forinash@em.doe.gov <betsy.forinash@em.doe.gov>
Cc: 'Dave Back' <davidback2@aol.com>; 'Charlie Wilson' <consultCW@aol.com>; 'Steve Marschke (smarschke@scainc.com)' <smarschke@scainc.com>; 'Shoemaker, Paul E' <peshoem@sandia.gov>; 'Gross, Mike - CTAC (mike_gross@earthlink.net)' <mike_gross@earthlink.net>; Anderson Ward <Anderson.Ward@cbfo.doe.gov>; 'Harris, Alton' <Alton.Harris@em.doe.gov>
Sent: Mon, Jul 11, 2016 5:48 pm
Subject: RE: BRAGFLO convergence issues

Kathy:

In response to your questions on BRAGFLO convergence issues last week please see the following. We would be happy to discuss this with you and your team if a discussion is desired.

Summary

Following discussions with EPA, DOE has reviewed the results of BRAGFLO calculations performed as part of performance assessment (PA) calculations for WIPP compliance and for sensitivity studies. The key results from this review are as follows:

- BRAGFLO calculations were reviewed for PCS-2012, and CRA-2014 for pressure fluctuations in the panel closures. The BRAGFLO results were well-behaved (i.e., stable and converged) for these compliance calculations.
- BRAGFLO calculations for sensitivity studies have defined the stability envelope for BRAGFLO. In particular, BRAGFLO results are well-behaved when:
    1. The capillary pressure model is turned on and there are no step changes in porosity/permeability, or
    2. The capillary pressure model is turned off for materials with step changes in porosity and permeability.

Discussion

In our teleconference on July 6, 2016 regarding the panel closure sensitivity study, we discussed our recent discovery of unrealistic instantaneous brine pressure fluctuations in the panel closures for a sensitivity study run by the EPA in 2012. In that study, the sensitivity of some panel closure properties was tested and compared to Sandia's PCS-2012 PA results. Of note, capillary pressure effects were activated when the step changes in porosity and permeability occurred. Also important to note is that those calculations were found to complete the 10,000 year simulations using solution residual tolerances previously utilized and within the maximum number of time steps allowed by BRAGFLO. The irregularities in brine pressure in the panel closures were only recently discovered because pressures in that area of the model have not been typically tracked/plotted in a Sandia compliance calculation analysis. Our experience with observed pressure fluctuations in the north-end sensitivity study led us to think that plotting pressures in the panel closures over time might reveal irregularities. In our early work on the north-end sensitivity study, we observed similar pressure irregularities in the Operations/Experimental (OPS/EXP) areas when (and only when) capillary pressure effects were active during a step change in material properties (porosity and permeability) in the OPS/EXP areas.

Following the teleconference, a question arose about the validity of converged BRAGFLO calculations (interpreted as those calculations that "completed" the full 10,000 year duration) in which "critical outputs" (taken to mean "typical outputs shown in analysis reports to-date") appear to be reasonable.

Immediately following our recent discovery of the irregularities with the 2012 sensitivity study results, we investigated the brine pressures in the panel closures for the PCS-2012 PA and CRA-2014 calculations and found absolutely no irregularities. Although step changes to material properties were implemented in those analyses, no unrealistic instantaneous brine pressure fluctuations were observed in the panel closures. This finding gave further support for the hypothesis (which was initially based on numerous tests performed during the north-end sensitivity study) that the irregularities stem from the combination of step changes in material properties and the activation of capillary pressure effects. The fact that our BRAGFLO model does not include any instantaneous property changes for materials that also have capillary pressure effects activated leads us to the conclusion that

Figure 1 Correspondence from R. Patterson (DOE) to K. Economy (EPA) dated July 11, 2016, page 1 of 2

there are no other regions of the BRAGFLO model for which these irregularities may arise. It is worth noting that convergence problems were anticipated prior to the PCS-2012 PA (Camphouse 2012, ERMS 557653) for cases in which capillary pressure effects were active and material properties changed instantaneously. Indeed, this is consistent with all of our observations to-date.

Thanks

Russ

From: Russ Patterson
Sent: Thursday, July 07, 2016 1:08 PM
To: 'Economy, Kathleen'; Chris Camphouse; Day, Brad A.; Zeitler, Todd (tzeitle@sandia.gov)
Cc: Dave Back; Charlie Wilson; Steve Marschke (smarschke@scainc.com); 'Shoemaker, Paul E'; Gross, Mike - CTAC (mike_gross@earthlink.net); Anderson Ward; 'Harris, Alton'
Subject: RE: BRAGFLO convergence issues

Kathy:

We are investigating past calculations and preparing a response to your questions. I believe we will have a more complete understanding and can discuss in more detail by early next week.

Thanks

Russ

From: Economy, Kathleen [mailto:Economy.Kathleen@epa.gov]
Sent: Wednesday, July 06, 2016 3:41 PM
To: Chris Camphouse; Day, Brad A.; Zeitler, Todd (tzeitle@sandia.gov); Russ Patterson
Cc: Dave Back; Charlie Wilson; Steve Marschke (smarschke@scainc.com)
Subject: BRAGFLO convergence issues

Greetings All,

I would like to continue the discussion brought up by Brad Day, that the BRAGFLO test runs were questionable that I performed with Chris Camphouse a couple years ago based on Brad finding questionable pressure fluctuations in the interior of the PCS. At the time, we looked at and plotted pressures and saturations in the waste panel, SROR, NROR, experimental and operations rooms with time for all vectors and found reasonable outputs.

Based on Brad's comments, this begs the question whether any BRAGFLO calculations are valid if they have 'converged' and critical outputs appear to be reasonable.

Regards,

Kathleen Economy

Radiation Protection Division-Center for Waste Management and Regulations
US EPA
phone: 202-343-9844
mobile: 202-650-7406

USPS Delivery:
1200 Pennsylvania Ave, NW (6608J)
Washington, DC 20460

(Note, I generally work from home on Mondays.)

"When one tugs at a single thing in Nature, he finds it attached to the rest of the world."

John Muir (1838-1914)

Figure 2 Correspondence from R. Patterson (DOE) to K. Economy (EPA) dated July 11, 2016, page 2 of 2