UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
REGION 7
11201 Renner Boulevard
Lenexa, Kansas 66219

FEB 13 2020

Mr. Paul V. Rosasco
Project Coordinator
Engineering Management Support, Inc.
25923 Gateway Drive
Golden, Colorado 80401

Dear Mr. Rosasco:

On January 8, 2020, Parsons submitted the Preliminary Excavation Plan, or PEP, West Lake Landfill Superfund Site Operable Unit-1, or OU-1, to the U.S. Environmental Protection Agency, on behalf of the West Lake OU-1 Respondents. This document was prepared and submitted pursuant to Section 3.4 of the May 6, 2019, Remedial Design Statement of Work, or SOW, OU-1, West Lake Landfill Superfund Site.

The EPA has completed review of this document and is approving the PEP with conditions. The comments enclosed with this letter specify the conditions of this approval. Failure to satisfactorily address these conditions as directed in the enclosed comments may jeopardize the EPA's ultimate ability to approve the Final Remedial Design.

The enclosed comments reflect the EPA's conclusion that there are substantial deficiencies in the PEP. The PEP is a building block for subsequent remedial design work to be completed including the Design Investigation Work Plan, the 30% Remedial Design, the Revised Excavation Plan, the 90% and Final Remedial Design. The EPA recognizes that additional information obtained during the design investigation has the potential to help resolve some of the deficiencies. Other deficiencies identified in the PEP may be resolved in the subsequent design deliverables, as directed in the enclosed comments. Therefore, the EPA determined that approving the PEP with conditions is appropriate to identify significant deficiencies requiring resolution while continuing to move the remedial design forward.

In accordance with the SOW, the Design Investigation Work Plan shall be submitted to the EPA 45 days from receipt of this approval, or Monday, March 30, 2020.

Please feel free to contact me with any questions or concerns by phone at (913) 551-7141 or by email at jump.chris@epa.gov.

Sincerely,

Christine R. Jump
Remedial Project Manager
Site Remediation Branch
Superfund and Emergency Management Division

Enclosure

cc: Mr. Ryan Seabaugh, MDNR

Preliminary Excavation Plan Comments

1. **General.** The Preliminary Excavation Plan (PEP) conditions heavily on the geostatistical model developed by S.S. Papadopolous & Associates (2017). The 2017 model was developed and tested for the purpose of radiologically impacted material (RIM) volume estimation at several thresholds to support the final feasibility study; however, it is not without weaknesses. For example, the Remedial Investigation Addendum (RIA) states on page 180, "*...as described by SSP&A (2017), there is significant uncertainty associated with the estimated extent and volume of RIM in Areas 1 and 2. Furthermore, SSP&A (2017) indicate that these estimates are more likely to be biased low than be biased high and therefore underestimate the actual volume of RIM in Areas 1 and 2.*" The EPA acknowledges that this statement applies to the 7.9 picoCuries per gram (pCi/g) concentration threshold, however, information has not been provided to the EPA that would explain whether and to what extent these statements would apply to the 52.9 pCi/g threshold. However, it is with this general frame of reference in mind that the EPA has developed comments to this PEP that relate to the Parson's adaptation of the existing model presented in the SSP&A Report for use in the Remedial Design.

   Because the respondents have proposed to use the prior model, developing sampling objectives in the Design Investigation Work Plan (DIWP) requires an evaluation of the limitations of this model for designing the excavation required in the Record of Decision Amendment (RODA). At a minimum, the PEP should have acknowledged the weaknesses of the 2017 model discussed in the SSP&A Report and other pertinent documents this model was developed for, e.g the RIA and the Final Feasibility Study (FFS). The limited sensitivity analysis presented in the PEP focuses too much on changes in volume and does not provide sufficient evaluation of associated changes in the extent of RIM. For example, the analysis of volume sensitivities to various model parameters, such as cumulative distribution functions (CDFs) and range lengths, is not sufficient given the present goal of defining RIM extent at the 52.9 pCi/g threshold for the remedial design (RD). Predicted RIM extent is likely to be sensitive to model parameters such as variograms and CDFs for reasons discussed elsewhere in this comment letter, and additional sensitivity analyses geared towards RIM extent should have been provided. Such analysis is important to evaluating and potentially identifying model inputs that could be improved with data collected during the upcoming Design Investigation. Since this work was not performed and presented in the PEP, it must be completed and presented in the DIWP. The EPA has provided specific comments below that identify conditions upon which the approval of the PEP is based and when those conditions must be met, many of which include developing content for the DIWP.

   Substantial additional complexity for the development of the RD excavation is introduced with the inclusion of a second model for activity estimation. However, kriging of raw or transformed concentration values has the potential to address both questions of extent and activity in a simpler and more transparent manner. Though the process of indicator kriging is more directly geared towards the objective of defining extent, the activity model may be an adequate tool that would reap similar benefits from improving regressions and variogram estimation as the extent model. Though uncertainties related to soft data transformation aren't explicitly defined in the form of a probability in the activity model, comments on Appendix B discuss a potential method for incorporating uncertainty captured by the standard error of the soft data regression models.

The RD model should be defined by data, information, and conceptual site knowledge. The Design investigation should aim to collect additional data in regions where large uncertainties severely qualify model predictions. While checking model results against data and prior models is a useful exercise, changes to the model must be supported by data and site knowledge. The Revised Excavation Plan must provide more details on the final modeling approach and how the data (historical and new) drive the selection of the input parameters, rather than on proving that the current model adequately replicates the predictions made by the 2017 model.

2. **General.** The PEP provides only an abbreviated description of how the geostatistical modeling tools were used to identify the extent of RIM, the volumes and the activity for the Preliminary Excavation. Specifically, Section 2 on the geostatistical model and Section 3.0 on RIM activity calculations are very brief and refer the reader to the associated Appendices A and B for detailed support and information. However, these appendices do not contain the necessary information to adequately understand and review the proposal. There is almost no discussion or evaluation of the limitations of the existing model. The EPA notes that Section 5.1 of the Design Criteria Report (DCR) states, "*The technical memo will include sufficient detail to develop sampling objectives in the DIWP as discussed in the paragraph below*" and "*Additional radioactivity data will be collected during the design investigation according to sampling objectives specified in the DIWP aimed at improving the accuracy of the model and to fill in data gaps necessary to complete the remedial design.*" The EPA does not believe sufficient information was provided in the PEP to develop all the necessary sampling objectives.

   Approval Condition: Specific objectives for improving the accuracy, precision and confidence of the geostatistical modeling process must be identified in the DIWP, or associated Field Sampling Plan (FSP) or Quality Assurance Project Plan (QAPP), as well as sampling strategies to accomplish those objectives, as required in Section 5.1 of the DCR.

3. **Section 1.0 Introduction.** The second paragraph states, "*This Preliminary Excavation Plan addresses proposed excavation of landfilled materials in Areas 1 and 2 that will be determined based on a geostatistical model.*" This statement fails to acknowledge the role of confirmation sampling for the remedy. The RODA states on page 65, "*The final boundaries of excavation will be confirmed through a combination of field screening and sampling within survey units no larger than 2,000 square meters. Confirmation sampling procedures will be specified in a site-specific sampling and analysis plan to be developed during the RD.*"

   Approval Condition: The 90% RD must explain the role of confirmation sampling in determining the final excavation boundaries in accordance with the confirmation sampling and analysis plan described in the RODA.

4. **Section 2.0 Geostatistical Model.** Additional details must be provided to clarify "kriging confidence and uncertainty" as used in the PEP. At a minimum, an explicit definition, such as a formula or calculation methodology, is needed for all the criteria being utilized to select additional sampling locations or to develop sampling objectives in the DIWP. Specifically, the intervals used to define confidence (stated to be the predicted value plus or minus 0.2 during a

conference call held on January 24, 2020, with Respondents) need to be clearly defined. The specific details behind the weighting of confidence and magnitude need to be defined for calculating kriging uncertainty.

Kriging variance or standard deviation must be provided in place of, or in addition to the kriging confidence metric discussed in the PEP. For the RIM extent model, kriging standard deviations are important for two reasons. First, standard deviations are common quantities that can be interpreted by a general audience without the additional context needed for deciphering kriging confidences. The interpretation of the kriging standard deviations is especially digestible because the predictions are probabilities between 0 and 1. A standard deviation of 0.1 and a prediction of 0.5, for example, suggests a 95% prediction interval from about 0.3 to 0.8. Second, lower standard deviations equate to greater confidence and are more consistent with interpretation of the subsequently provided kriging uncertainties. The endpoints of prediction intervals evaluated in concert with the predicted probabilities would provide some understanding of the efficacy of the model that is far more useful than a percentage representing kriging confidence. It is recommended that kriging standard deviation be presented in the same format that kriging confidence is presented in Figures A-19, A-20, A-23, A-24, and A-25.

If kriging uncertainty will be used to select boring locations, the following alterations to the calculation with respect to the predicted probability of non-exceedance are recommended:

> 1) Non-exceedance probabilities close to 0 or 1 should be down-weighted, and intermediate probabilities should be up-weighted, for two primary reasons:
>
> > i) the probability confers additional information about uncertainty that should not be ignored; and
> >
> > ii) using the concentration as a weight to sample locations will bias the data high and overall increase the extent and the estimate of total activity, when it is equally important to characterize areas where activity is relatively low but the probability of being above or below 52.9 is highly uncertain.
>
> 2) A simpler but conceptually similar approach would be to define a 95% confidence interval for the probability of non-exceedance and identify locations where 0.5 is within the confidence interval. These locations indicate a high probability of a decision error as far as inclusion in the extent and should be candidates for further sampling. If certain regions are of more interest than others, separate sampling campaigns can be developed to target specific objectives, though the data collected for these purposes should be carefully evaluated before use in the extent model due to inherent biases.

Approval Condition: Either kriging standard deviation or kriging variance must be provided in a similar manner to Figures A-19, A-20, A-23, A-24, and A-25 in the DIWP to develop Data Quality Objectives (DQOs). If Kriging confidence and uncertainty as used in the PEP are used to develop sampling objectives or select boring locations, an explicit definition must be provided in the DIWP.

5. **Section 3.0 RIM Activity Calculations.** The EPA has significant concerns about the methodology proposed in this PEP for Activity Calculations. Specific comments on activity calculations have been provided for Appendix B.

6. **Section 4.0 Optimization.** Although Section 4 contains the presentation of the optimized excavation, insufficient information and documentation are included to describe how the model was used to calculate the volumes of RIM and non-RIM material associated with the preliminary excavation. Therefore, the EPA is unable to fully consider whether the volumes presented are consistent with the requirements of the RODA and the output of the model. With the limited information submitted, the EPA was able to identify a potential issue with the volumes presented based on comparison with the prior volumes presented in the Focused Feasibility Study (FFS) and RODA. The Respondents confirmed that there were errors with the volumes and activities presented in the January 8, 2020, PEP and submitted replacement pages on January 28, 2020, with revised volumes and activities. The revised pages did not include an explanation of the prior error nor was additional documentation included describing how the model output was used to estimate the updated volumes. The EPA cannot evaluate whether the updated volumes presented are calculated accurately from the model output without additional information. See Section 5 of the S.S. Papadopulos & Associates, Inc., 2017, Estimated Three-Dimensional Extent of Radiologically Impacted Material, West Lake Landfill Operable Unit-1, Bridgeton, Missouri, December 22, 2017, (SSP&A Report) for an example of the information needed.

   Approval Condition : In order to reduce the likelihood that the EPA will require the Respondents to supplement the Design Investigation (DI) Evaluation Report and/or perform additional design studies, as described in Section in 3.6 (d) of the SOW, the use of the model to calculate volumes of RIM and non-RIM material for the excavation must be clearly presented in the 30 % RD.

7. **Section 4.0 Optimization and Figures 13 and 14.** The excavation optimization described in Section 4 and presented in Figures 13 and 14 appears to be incomplete, and therefore may not be consistent with the RODA. Page 4-1 states, "*Figures 13 and 14 highlight locations where RIM will not be excavated at depths of 8 to 16 feet B2005GS because there is little activity in those locations and significant additional excavation area and volume would be required to access it. Optimization in the RODA requires minimizing, to the extent practical, the volume of waste (both RIM and non-RIM) excavated in order to remove and dispose of the ROD specified amount of radioactivity by leaving isolated pockets in place between 8' and 12' and targeting higher concentrations of RIM between 12' and 20'.*" RIM between 12 and 16 feet is not required to be removed unless it is necessary to meet the required activity objective in the RODA. Therefore, leaving RIM in place between 12 and 16 feet does not require a demonstration that these RIM occurrences are "isolated." Further, the presented optimization does not appear to target only higher concentrations of RIM between 12 and 16 feet. Determining the volume of RIM and non-RIM (both overburden and set-back) between 8 and 12 feet, as well as the volume, concentration, and activity of the RIM occurrences between 12 and 20 feet is critical to performing a complete optimization of the excavation.

In general, the approach to visualizing the optimized excavation in Figures 13 and 14 is understandable and clear, however, additional information is needed as discussed in the previous paragraph to demonstrate the optimized excavation meets the RODA requirements. For example, in Figure 14 it appears that deeper excavation of the higher concentration area identified in Area 2 may result in a lower overall excavation volume than the multiple acre-sized excavations currently proposed in Area 2 within the 12 to 16-foot interval while still removing an equivalent amount of activity. In Figure 13, more specific information is needed about the depth of the RIM occurances for the first isolated pocket proposed for Area 1 so that the EPA can understand whether the RIM occurances identified from the 0 to 8-foot interval extend throughout the entire interval or are primarily near the surface. The EPA notes that the historical overland gamma survey identified a larger area of elevated RIM than that predicted by the model partially because this data could not be incorporated into the model. Therefore, the extent of RIM greater than 52.9 pCi/g between the 0 to 8-foot interval is likely larger than depicted for this portion of Area 1.

Approval Condition: Further optimization and supporting documentation for optimization must be presented in the Revised Excavation Plan in order to demonstrate compliance with the RODA requirements. Some of the other specific comments on the topic of optimization and activity estimates for Section 6.0 of Appendix A and Appendix B also include requirements for future submittals. In order to reduce the likelihood that the EPA will require the Respondents to supplement the DI Evaluation Report and/or perform additional design studies, as described in Section 3.6 (d) of the SOW, the preliminary optimized excavation proposed in the PEP should be revised to address these comments and be presented in the 30% RD Report to the extent feasible.

8. **Section 5.0 Data Gaps**. Section 5 presents a proposal to use the model to identify data gaps; however, the criteria proposed for selecting additional boring locations is limited. The EPA previously provided information in an attachment to an email from Tom Mahler to Paul Rosasco, October 7, 2019, which stated the EPA's expectation that, "*consideration must be given to any potential limitations inherent to the current data set (Hard and Soft) for the preliminary model. Sampling objectives must be developed for the design investigation to address these potential limitations.*" The PEP fails to provide sufficient evaluation of the preliminary model assumptions and limitations based on the existing data set and selected methodology, which is required in Section 5.1 of the DCR.

   Approval Condition: A further evaluation of the modeling methodologies must be performed to demonstrate that the model is reasonable for use in this RD. This must be demonstrated for the EPA to approve the Revised Excavation Plan and ultimately the Remedial Design. If the final model will be based on the prior FFS model methodologies, this evaluation must include an analysis of the limitations and assumptions in the SSP&A Report (citation to SOW if we can). In addition, identification and evaluation of other elements or issues throughout pre-processing, kriging, and post processing that could affect sensitivities and data needs should be further assessed and documented in the DIWP and associated documents. This includes topics such as the use of gamma soft data to represent the absence/presence of thorium (Th); sample collection at close intervals to evaluate the appropriateness of using a zero nugget; and sample collection within specific concentration ranges to improve regressions and evaluation of assumptions and parameters that could specifically affect extent of RIM determinations. In order to reduce the likelihood that the EPA will require the Respondents to supplement the DI Evaluation Report and/or perform additional design studies as described in Section 3.6 (d) of the SOW, these evaluations must be completed to the extent feasible so that sampling objectives can be established in the DIWP.

9. **Section 5.0 Data Gaps.** An important consideration for both the limitations of the existing data set/model and developing sampling objectives for the DIWP, is the extent to which spatial bias and/or spatial clustering occur within a data set. Pyrcz and Deutsch (2003) define spatial bias as "*sampling that does not characterize the full range of the distribution.*" For example, sampling that only consists of regions with high concentrations and completely neglects regions with low to median concentrations would present spatial bias. Construction of a simple histogram or other plot displaying the sampled concentrations in Areas 1 and 2 would provide a tool to consider spatial bias. This tool can then be compared to what concentrations are expected based on the conceptual site model and knowledge of the history of the site. If concentrations in a certain range are missing without prior expectation of this occurring, an objective of further sampling should be to target regions with those suspected concentrations in order to minimize spatial bias.

   Spatial clustering, by contrast, is the preponderance of samples in particular areas over other areas, and is easily identified at this site, as data density is clearly greater within certain portions of Area 1 and 2. Although spatial clustering could cause spatial bias if regions with high concentrations were particularly over-sampled compared with low-concentration regions, spatial clustering could also occur without spatial bias. Fortunately, declustering methods exist to properly account for the effects of spatial clustering (EPA 1996).

   For example, one such polygonal declustering method involves drawing a Voronoi polygon around each sample location that is representative of the areas where the sample in question is the closest available sample. Samples can then be weighted by polygon area. The grid-based cell declustering method is also common and is simpler to apply on a three-dimensional scale (Pyrcz et al. 2006). This method assigns weights to each datum inversely proportional to the number of samples in the same cell times the total number of occupied cells (Pyrcz et al. 2006). Another common method is the kriging weight method (Pyrcz and Deutsch 2003).

   Though clustered data can simply be discarded, it is an unnecessary loss of information (Isaaks and Srivastava 1989). To ensure all the data of sufficient quality collected for the site can be consider, established declustering methods can be used to develop weights for fitting a variogram model which will properly account for regions of higher data density. Earth Volumetric Studio (EVS) does not include declustering method options, but data can be exported to the GeoEAS (EPA) / GSLib (Stanford) file format. Declustering can then be performed in GSLib and results passed back to EVS. The EVS library external_kriging may be of useful in this process (Earth Volumetric Studio 2019). Additionally, custom code may be needed to fit the weighted variogram models.

   Approval Condition: Because spatial bias and spatial clustering exist to some extent in the existing data set, the DIWP must consider sampling objectives aimed at addressing these issues.

Furthermore, because certain sampling objectives may create additional bias, the revised excavation plan must address spatial bias and spatial clustering.

10. **Section 5.0 Data Gaps, page 5-1.** The last paragraph of the section discusses additional evaluations related to the geometric quality of the data that will be performed "*separately as part of the DIWP development.*" Specifically, the paragraph states, "*The available data set has a variety of "hard" and "soft" data that was obtained from borings and other methods, such as surface gamma scans.*" The EPA could not locate any other mention of the overland gamma survey in the PEP nor any discussion about its significance to the model predictions. It is not clear whether the Respondents intend to further evaluate this data. Because the overland gamma survey identifies areas of elevated gamma that are not identified by the IK model, additional evaluation of this data is necessary to determine whether it can be incorporated into the model and/or how the excavation design will address areas of known elevated gamma that are not identified by the IK model.

    The EPA notes the SSP&A Report contains a specific section (Section 8.3 Review of Overland Gamma Data) and an entire appendix (Appendix J Evaluation of Overland Gamma Measurement Data) that discuss this data at length. For example, "*Although it is not recommended that the overland gamma data be incorporated into the 3D geostatistical analysis of RIM in a manner similar to the borehole and core gamma, the overland gamma data do identify some locations that – depending upon the remedial alternatives, that might be further considered or evaluated – should be prioritized for retrieving additional surface grab samples from which to obtain a refined understanding of the presence and activity concentration of RIM at (or very close to) the surface.*" In addition, the discussion included elsewhere in this section of the SSP&A Report provides a description of the data gaps associated with developing a reasonable regression between the overland gamma data and hard data for radium and thorium. This discussion must be considered when developing sampling objectives in the DIWP related to further characterization of RIM "*at (or very close to) the surface.*"

    Approval Condition: Additional discussion regarding further evaluation of the overland gamma survey data and associated data gaps for characterizing the extent of RIM greater than 52.9 pCi/g at or very close to the surface must be provided in the DIWP. This discussion must consider and reference the information in Section 8.3 and Appendix J of the SSP&A Report. Sampling objective(s) must be developed to address this data gap.

Appendix A – Geostatistical Analysis Technical Memorandum

11. **Appendix A, General Comment.** No formal uncertainty analysis is provided. That is, results are deterministic; one preliminary excavation option is provided and is subject to all model assumptions. While the plan requires a deterministic approach, the impact of model uncertainties on extent and activity should be provided.

    There is inherent uncertainty in model predictions which can be quantified by the kriging variances at each predicted location. As discussed in comment 4, kriging variances allow for calculation of prediction intervals, which provide understanding of uncertainties in model predicted probabilities at individual locations. The endpoints of prediction intervals can then be evaluated in concert with the predicted probabilities to provide understanding of the efficacy of the model and uncertainty in extent. In addition, polygons depicting extent at other probability selection criteria would provide further illustration of model uncertainty, combining both uncertainty in the predicted probability and the probability itself. (see comment 12)

The EPA also notes that consideration of uncertainty in the model predictions is mentioned only within the context of the RIM extent model (IK model), but this topic must be discussed within the context of the activity model (OK model) as well. Minimally, a confidence interval for total activity will need to be computed. (See comments on Appendix B)

12. **Appendix A Section 1.0 Introduction**. The section states three categorical uses of the geostatistical model to support the RD. Categorical Use (1) is, "*Identification of areas and depths of RIM within OU-1 Area 1 and Area 2 <u>that have greater than 50 percent (%) probability</u> [emphasis added] of being above 52.9 picoCuries per gram (pCi/g) in activity concentration.*" The EPA recognizes that for modeling purposes, RIM is interpreted as more-likely-than-not present at any grid cell within the model boundary if the cumulative probability of non-exceedance for the corresponding threshold is less than 0.5. The EPA previously provided comments to the SSP&A Report that an "*element of subjectivity involved in the selection of a probability of 0.5*" (Singh, 2018, pdf page 4). However, Figure A-4 indicates that total volume of the predicted RIM changes significantly with changes to the non-exceedance probability thresholds and therefore the EPA expects similar changes to the extent of the predicted RIM. Therefore, extent polygons and estimated relative activities for other probability selection criteria should be should considered when evaluating model uncertainties.

    Categorical use (3) is, "*Support the Remedial Design Investigation through identification of additional borings with the intent of further delineating areas of RIM >52.9 pCi/g, <u>improving accuracy of geostatistical modeling by improving correlations between hard data and soft data</u> [emphasis added], and better defining RIM concentrations to facilitate meeting total activity removal requirements.*" No additional discussion or information about weaknesses or data gaps associated with the "*correlations between hard data and soft data*" is presented in the PEP. The model relies upon a correlation between hard and soft data which is a source of model uncertainty. The significance of this uncertainty on predicted RIM extent should be considered further for the development of an RD model.

    Approval Condition: The volume and extent of predicted RIM at probability thresholds other than 0.5 must be considered and additional sampling locations proposed in the DIWP. The DIWP must include evaluation and discussion of the weaknesses, limitations, and data gaps associated with correlations between hard and soft data (for both radium and thorium). Along with presenting any corresponding sampling objectives developed to address them.

13. **Appendix A Section 1.0 Introduction**. The first unnumbered bullet on page 2 of 18 describing the objectives states: "*Review and evaluate the previously completed geostatistical approach and determine if there are potential intrinsic failures, oversimplifications, or areas for potential improvement.*" The information in the PEP is insufficient to demonstrate that this objective was met.

Approval Condition: Further analysis and description of the modeling processes will need to be performed and documented in the DIWP. This will identify and support sampling objectives, boring and sample locations.

14. **Appendix A Section 2.0 Previous Work.** The document states on page 4, *"Many aspects of these evaluations of inputs, assumptions, and procedures were documented in SSP&A's Evaluation of Uncertainty (Appendix I of SSP&A Report) which determined that the estimates of RIM volume and extent were quite stable over reasonable ranges of input values and assumptions."* The EPA does not agree with this statement because it mischaracterizes the evaluation of uncertainty presented in Appendix I and is overly broad (See page I-8 of the SSP&A Report). More generally, this section fails to acknowledge a number of statements in the SSP&A Report that caution against the use of the model as presented in that report for the design of an excavation remedy without performing additional evaluation. For example, there is no discussion of the nugget parameter in Appendix A, which is one of the primary parameters evaluated in Appendix I of the SSP&A Report. The analysis in Appendix I concluded that the nugget parameter has the potential to impact the design and cost of a partial excavation, which is one reason the attachment to the EPA's October 7, 2019, email to Paul Rosasco discussed the need for "modeling transparency" with regards to selecting model parameters, including the nugget.

    Approval Condition: All future summaries or descriptions of the SSP&A Report must accurately reflect the text and conclusions of that Report. Also, if the final RD model will be based on the prior FFS model methodologies presented in the SSP&A Report, a thorough evaluation and analysis of the limitations, assumptions, and potential concerns for use of that model for an excavation RD, which are noted in the SSP&A Report, must be performed and ultimately documented in the Revised Excavation Plan. In order to reduce the likelihood that the EPA will require the Respondents to supplement the DI Evaluation Report and/or to perform additional design studies as described in Section 3.6 (d) of the SOW, these evaluations must be completed to the extent feasible so that sampling objectives can be established in the DIWP.

15. **Appendix A Section 2.0 Previous Work.** It is unknown whether the use of soft data to predict thorium concentrations near 52.9 pCi/g as has been done in the proposed model is reasonable. The use of soft data for predicting thorium concentrations needs to be formally evaluated because Thorium-230 does not emit gamma radiation that can be detected at less than approximately 2,000 pCi/g in the field by the instruments used for borehole and core gamma measurements. The EPA recognizes that the substantial number of samples collected from Area 1 and Area 2 indicate that Radium-226 and Thorium-230 are collocated and generally related, (e.g. Thorium-230 is present in concentrations greater than Radium-226). Because Radium-226 emits gamma radiation that can generally be detected at concentrations near background, it is reasonable to utilize the soft data for predicting radium concentrations. Therefore, in theory soft data could be used as an indirect predictor of Thorium-230 because of the relationship between thorium and radium demonstrated by the existing data set. More evaluation is needed to determine whether this relationship is reliable for predicting thorium-230 concentrations near

52.9 pCi/g. This is particularly important, as the extent of RIM at the Site appears to be driven by thorium rather than radium.

Because the prior SSP&A model is being proposed for use in the remedial design, these relationships must be formally evaluated to estimate a reasonable detection limit for Radium-226 and the approximate thorium-230 concentration that corresponds to this detection limit. This exercise is necessary to determine the reasonableness of the model to predict the extent of thorium-230 at any concentration. These estimated detection limits must be understood, in order to evaluate the reasonableness of the model methodology for the remedial design.

Completing this exercise is not possible without further considering the regression between radium-226 and various soft data types, as well as, the regression between radium-226 and thorium-230. The uncertainty associated with these detection limits directly relates to the strength of these regressions. The PEP should have evaluated these regressions based on the existing data set to identify any limitations or data gaps so that sampling objectives could be developed in the DIWP. Therefore, this evaluation must be presented in the DIWP.

Ideally, the data collection for the design investigation would occur in stages. First, data should be collected to improve the geometric quality of the data or regressions. Assuming the sampling does improve confidence in sub-models, kriging uncertainties should be evaluated to identify further sampling locations after the sub-models have been updated. However, the EPA acknowledges that while additional sampling will improve the characterization of the extent of RIM greater 52.9 pCi/g, it may not improve the regressions. Therefore, alternative strategies for developing sampling objectives and additional boring locations may also be appropriate. For example, consideration of kriging uncertainties from modeled thorium predictions using only hard data may be useful to identify boring locations. Alternatively, the current radium-thorium regression could be evaluated to estimate the concentration of radium-226 that corresponds to 52.9 pCi/g of Thorium-230. Input could then be developed for an additional model run for radium-226 only at this concentration. This new model run could be compared to the extent of RIM presented in the PEP to develop sampling objectives and identify additional boring locations.

Approval Condition: In order to determine whether soft data is being used appropriately for modeling of thorium-230, the DIWP must include an estimation of the detection limit for Radium-226 for each type of soft data. In addition, the corresponding Thorium-230 concentration for each of the Radium-226 detection limits must be estimated from the radium-thorium regression and presented in the DIWP. An evaluation of the regressions to identify limitations or data gaps also must be presented in the DIWP so that a sampling objective can be developed aimed at improving these regressions. Finally, one or more sampling objectives must be included in the DIWP to address the uncertainty with the model's prediction of thorium.

16. **Appendix A Section 2.0 Previous Work.** The PEP fails to list and fully consider all aspects of the CDF development for the SSP&A model that may be significant for the remedial design. If CDFs will not be further developed in the Design Investigation, further proof is needed that BOTH extent and total activity estimates will not be affected by a change in CDFs. Page D-1-6

of the SSP&A report states, *"At the conclusion of this process, the standard errors used to describe the CDF were pragmatically adjusted to 0.35 and 0.50 for combined Ra and combined Th, respectively."* Because the CDF was manually adjusted beyond what was estimated from data regressions in development of the extent model, the model seems to be sensitive to CDFs, despite the claim in the PEP that it is not (Figure A-4). More information must be presented to resolve this discrepancy. Further, the PEP does not appear to include this step of the CFD development process in section 2.0, which should have been described in detail for increased transparency and so that sampling objectives could be considered in the DIWP. This needs to be resolved during the design investigation to reduce the likelihood of multiple investigation mobilizations and subsequent extension of the project schedule. The appropriateness of this approach to developing CDFs must be further demonstrated if this approach is utilized for the revised excavation plan in order for the EPA to consider approval. Ultimately, the process to determine CDFs for the final model must be described explicitly in the revised excavation plan if the data for the final model are processed in this manner.

Approval Condition: Additional summaries of the SSP&A Report presented in future deliverables must be accurate. The DIWP must present an evaluation and further supporting information related to the development of the CDFs. This evaluation must include further justification of any manual adjustments of the CDF due to the presence of *"small, isolated, unverifiable islands of RIM"* discussed in appendix D of the SSP&A Report to evaluate the need for an associated sampling objective. In addition, the CDF development process must be explicitly defined and fully supported in the revised excavation plan if the final model relies upon CDFs.

17. **Appendix A, Section 2.0 Prior Work, Page 6**: The document states: *"Thus far all sensitivity analyses conducted by SSP&A indicate that the final Base Case CDF is appropriate, and the results are insensitive to small differences in the CDFs when comparing the volumes and extent of RIM >52.9 pCi/g."* This statement appears to contradict statements from the PEP and SSP&A Report which reported that pragmatic adjustments to the CDF were made to adjust the extent of RIM greater than 52.9 pCi/g to eliminate occurrences beyond or within the convex hull as noted in Figure A-2 (See Section 1.2.1 Method 1 - Development of Base Case CDF, page D-1-6 in the SSP&A report). Because of this sensitivity, it is important to understand CDF development, use, and any impacts on the model outputs so that sampling objectives can be developed, if necessary. The PEP does not adequately describe and evaluate the prior CDF development process for appropriateness in the development of an RD model (e.g. increased emphasis on extent of RIM). Section D-1.2.2 of the SSP&A Report also states that the continuous CDF resulted in a different convex hull than the base case CDF.

    Approval Condition: Summaries of the SSP&A Report presented in future deliverables must be accurate. CDF development, use, and potential impacts on the model outputs must be further evaluated in the DIWP to evaluate the need for an associated sampling objective.

18. **Appendix A Section 2.0 Prior Work. Appendix A, Page 6**: The document states, *"It is understood that as further data are collected and the geostatistical model is updated, additional CDFs can be evaluated as appropriate."* On page 13 in Appendix A it is stated that one of the

main objectives for identification of potential additional boring locations which the geostatistical model will support includes further evaluating and improving upon previously determined correlations between soft data and hard data. This implies that additional CDF development will be consider following additional data collection as these correlations were only used to develop the CDF.

Approval Condition: The explicit process for developing sampling objectives based on the identified weaknesses in the existing correlations must be stated in the DIWP. Likewise, the analysis of this data must be presented in the DI Evaluation Report and CDFs should be re-evaluated and modified with the DI data as necessary for presentation in the Revised Excavation Plan.

19. **Appendix A Section 3.0 Transition from IK3D to C-Tech EVS.** Page 8 of the document states: *"The determination of these settings was accomplished by starting with SSP&A parameters, and then, where subtle differences were noticed, conducting iterative response testing to provide acceptable matches of: (1) volumes, and (2) spatial extent of RIM >52.9 pCi/g."* There is no requirement for the Parsons RD model to be equal to the prior model in the SSP&A Report that was constructed for Feasibility Study purposes. The primary focus on model evaluation and revisions should be on assuring the model is appropriate and capable of accomplishing the new objectives for RD purposes.

20. **Appendix A Section 4.0 Sensitivity Analysis.** Page 9 of 18 states, *"This section summarizes the work done by SSP&A (as detailed in Appendix I of the Geostatistics Report, December 2017) which evaluated the sensitivity of RIM volume and extent estimates and demonstrated that within a reasonable range of variogram setting and CDFs there is no appreciable change in RIM extent or volume."* Neither this language nor similar language is present in Appendix I of the SSP&A Report.

    Similarly, page 11 of 18 states, *"The 52.9 pCi/g portion of Figure A-4 demonstrates that, across a range of settings, the indicator kriging produces remarkably similar estimates of the volume and extent of RIM >52.9 pCi/g for each Area, indicating the model is relatively insensitive to the CDF, variograms, and search criteria."* No description or discussion is provided in the PEP as to why the specific sensitivity analyses previously conducted in the SSP&A Report represents a "reasonable range" of variogram settings and CDFs. Further, no supporting information is provided to demonstrate the model-predicted extent is "relatively insensitive" to the CDF, variograms, and search criteria. Kriging models are always sensitive to the variogram model used, as the variogram model exactly describes the spatial interpolation that takes place. While volume may not be obviously sensitive within the range of 240 to 500 feet (see Figure A-80), indicator probabilities over space are certainly affected, and additional details need to be provided with regard to the variogram models for both the extent (indicator kriging, or IK) and the activity (ordinary kriging, or OK) models. For both the extent and activity models, the fitted model needs to be shown alongside the empirical variogram, similar to Figure 4-1 of SSP&A Report. Though the anisotropy approach may be justified, the implied z-directional variogram should be plotted alongside the vertical empirical variogram. The parameterization of the variogram model should also be defined (exponential or spherical, for example).

Spatial interpolations and kriging weights depend entirely on the variogram model, which is defined by the range and sill parameters as well as the shape of the fitted model. Empirical variograms are calculated by binning observations into groups based on distance to other sample locations. Therefore, a variety of observation pairs, some closer and some farther in proximity, are needed for accurate variogram estimation. The empirical variograms should have been evaluated in the PEP to reveal any potential data gaps. This evaluation must be presented in the DIWP so that a sampling objective can be developed if data gaps are identified. The PEP does consider kriging uncertainty to some extent for use as a tool when developing the DIWP, however, this does not address improvement of variogram estimation because spatial interpolations do not address uncertainty in the variogram model itself. Kriging uncertainty at each location is based on the variogram models and distance to sampled locations.

The PEP neither lists nor discusses the nugget parameter which is a critical defining parameter of the variogram model. The rationale for the selection of a zero nugget, if that is what was done, should have been described in the PEP for each model, so that it is clear whether the basis is related to the data or the desire to over-smooth the predictions for the purpose of excavation design. Additionally, the PEP should have evaluated the appropriateness of the selected nugget criteria and discussed the impact that the nugget parameter has on kriging confidence and variance. This evaluation must be presented in the DIWP so that consideration can be given to the extent to which the model uncertainty is relied upon for selecting additional sampling locations.

Approval Condition: The modeled variogram must be shown alongside the empirical variograms for both the extent and activity models (see Figure 4-1 of SSP&A Report); the implied z-directional variogram must be plotted alongside the vertical empirical variogram; and the parameterization of the variogram model must be defined (e.g., exponential or spherical) in the DIWP. A potential sampling objective for improvement of variograms must be evaluated in the DIWP with consideration of identifying or creating a variety of observation pairs at various proximities to improve variogram estimation accuracy. The nugget parameter must be identified in the DIWP with a discussion as to why that selected value of nugget was chosen. The DIWP must also include a discussion of the potential effects that selection (either zero or non-zero) has on kriging confidence and variance for each model so that those effects can be taken into consideration during selection of additional boring and sample locations.

21. **Appendix A Section 4.0 Sensitivity Analysis.** In general, demonstration of a lack of volume sensitivity is not adequate for the purposes of this document. Figures A-7 and A-8 do not provide an evaluation of the sensitivity to RIM extent. Illustrations depicting RIM extent under various variogram range length and CDF scenarios need to be provided for the indicator kriging model (such as those in A-6).

The activity model should also be evaluated separately for sensitivity of the relative total activities to the variogram parameters and CDFs. Even if extent is shown to be insensitive to a given parameter, the same parameter in the activity model may be sensitive with influences on activity estimates and isolated pocket delineations.

Approval Condition: Sensitivity analyses for both models depicting changes in extent and relative activities must be provided in the DIWP to the extent feasible in order to evaluate the need for an associated sampling objective.

22. **Appendix A Section 4.2, pages 11 and 12.** This portion of the section discusses the grid cell size used in the SSP&A model and include the following statement, *"The grid size was suitable for the data density in both Area 1 and Area 2 for initial geostatistical model development; however, it was computationally time consuming and provided higher resolution than will be necessary for the excavation design."* No further discussion about the resolution that is necessary for the excavation design is provided in this section as it relates to the selection of an alternate grid cell size.

    The following paragraph then discusses the nature and results of a sensitivity analysis performed for grid cell size on volume and extent. The EPA considers the evaluation of the sensitivity to extent with grid cell size to be somewhat subjective and simplistic. However, the EPA notes that the document states, *"For both Areas, visual inspection shows that the general locations of RIM are maintained, but <u>the shape and extent are altered</u> when using the different grid sizes... While <u>the spatial extent appears to be visually altered</u>, the estimated volumes stayed within 15% of the original estimate for Area 1, and 10% for Area 2."* Additional supporting information is needed to demonstrate the conclusion that, *"225 square meters (15 by 15-meter grid cells) was the most appropriate for adequately estimating RIM extent and volume while providing realistic design basis calculations."*

    Approval Condition: Additional supporting information must be provided in the revised excavation plan to demonstrate the appropriateness of the final grid cell size. At a minimum, must include a discussion of what resolution is considered necessary for the excavation design, a visual depiction of the RIM extent associated with the selected grid cell size like what is presented on figures A-5 and A-6, and further description of the observed changes in extent of RIM, if any, and why these observed changes are insignificant for the excavation design.

23. **Appendix A, Section 4.3, page 13.** The sensitivity analysis presented for range length is not complete because only one value below the selected range length presented in Section 2.0 of the PEP was tested and the corresponding decrease in the RIM volume was significant. The Parson's sensitivity analysis concludes that, *"the estimates of RIM volume and extent are generally robust (insensitive) over reasonable values of the range length..."* The text in this section and the associated figures do not provide sufficient information to support this conclusion. Further, this conclusion does not seem consistent with some of the statements made elsewhere is this section. For example, the text also states on page 13, *"A range length of 100 feet produced the largest change in RIM volume, dropping it by 78%."* Figure A-8 indicates that no volumes were estimated for any range lengths between 240 feet and 100 feet even though the volume decreased significantly at 100 feet. While the text states that values of range length above 500 feet and lower than 100 feet are not reasonable, there is no discussion of what are considered *"reasonable values of range length"* or how this conclusion was reached. It is critical to compare the volume and concentrations of predicted RIM between portions of Area 2 where sample spacing tends to

be larger and portions of Area 2 where sample spacing tends to be smaller. The discussion also doesn't address sensitivity in extent of RIM which is critical for the design of the excavation.

Approval Condition: A discussion of the relationship between data density (e.g. boring/sampling spacing) and the variogram, specifically range length, must be included in the DIWP. In addition, further description and rationale for what is considered reasonable values of the range length must be described in the DIWP so that the conclusion provided in the PEP can be understood. Because the volume dropped signficntly with a range length of 100 feet, additional sensitivity analysis must be performed to understand the significance of this decrease in volume. This analysis must be provided in the DIWP to inform any sampling objectives that relate to further characterizing the predicted extent of RIM.

24. **Appendix A Section 5.0 Geostatistical Model for Supporting Preliminary Design and Additional Boring Locations.** The objectives for selecting additional boring locations must be stated in the DIWP or associated documents. While there does seem to be some focus on increasing confidence in predictions from the geostatistical models, it is not clear whether the focus is on increasing confidence in the predicted probabilities from the indicator kriging (extent) statistical model, on increasing confidence in the predicted concentrations from the ordinary kriging (activity) model, or both. In addition, other objectives such as improvements to the CDFs or evaluating sensitivities associated with extent of RIM appear to be overlooked. The methods used to evaluate the modeling process and generate sampling objectives should determine specific sampling locations as part of the DIWP rather than vague regions for sampling to take place. Each time a new boring is selected, this will change the uncertainty and the predicted concentrations/probabilities both nearby that location and elsewhere in the model. In some cases, it may be beneficial to consider an iterative approach to selecting locations in order to evaluate how additional borings effect model uncertainty. This can be done as a simulation by first selecting a location, drawing a random value for concentration, and re-running the model in order to select the next location.

    Approval Condition: Sampling objectives for data collection proposed in the DIWP must be presented in the DIWP or associated documents. Details regarding location and depth of each boring should be included along with the type of data collection that will occur at each location, with clear descriptions of methodology for determining sample depth and type.

25. **Appendix A, Section 5.0, page 14 of 18.** The quoted text from the C-Tech manual included on this page states, "*As we approach the location of our samples, the standard deviation will approach zero (0.0) since there should be no error or deviation at the measured locations.*" The EPA notes that for this Site, there is uncertainty (error) in the selection of the depth for hard data and certain types of soft data due to limitations in the drilling methods related to core recovery. Therefore, some assumptions that may be reasonable for typical applications of kriging models must be carefully considered for this remedial design.

26. **Appendix A Section 6.0, Pages 16-17 of 18.** The "*general overview processes for total activity calculations*" provided in section 6.0 is not clear. The reference to Figure A-29 - *Area 1 Uncertainty Cross Section Conceptualization* in this section appears as though it should be

referencing Figure A-31 – *Total Activity Calculation Conceptualization.* Some additional information is provided in Figure A-31, however, several of the statements on this figure are also not clear. In general, it appears as though the three steps described in Section 3 are meant to correspond to the flow diagram presented on Figure A-31. The EPA notes the following about these steps as presented:

a. The first step in the flow diagram is titled "*Volume*" yet step 1 in the text and the description beneath this title on the figure refer to "*3D extent of RIM.*"
b. Some of the descriptions underneath the "*3D Extent of RIM*" portion of figure A-31 are not presented so that the reader can understand the significance or are just not clear, including, "*Involves standard error*" and "*Groups activities.*"
c. Step 2 in the text clearly communicates that activity concentrations are being estimated from within the RIM extent, but it does not provide any context as to how. Further, it is unclear what the significance and meaning of "*Relative calculation*" and "*Limit to RIM Extent*" are on Figure A-31. The EPA acknowledges that in theory an ordinary kriging model of activity concentrations can result in an improved discretization of these concentrations compared to an indicator kriging model, as described in the third bullet in the center of figure A-31. However, it is unclear what is being referred to by "*avoids the grossly averaging data*" and how this statement supports this theory.
d. Step 3 does not indicate which model will be used to estimate volume. While not explicitly stated, the EPA assumes that the grid cell size utilized in the ordinary kriging model is the same as the grid cell size utilized in the indicator model. This process must be clarified and documented so that it can be demonstrated that the actual cells determined to be within the extent of RIM greater than 52.9 pCi/g from the indicator kriging model are identical to the cells used in the activity calculations from the ordinary kriging model. In addition, as described in previous comments regarding Figures 4 and 5, the depicted extent of the preliminary excavation appears to be smoothed. It is not clear whether this smoothing process is or is not being accounted for within EVS and/or the Volumetric model described on page 6 of appendix B.
e. Use of the phrase "*Combining the volume and activity concentration results*" is confusing, as it implies the values will be summed to estimate total activity. The EPA notes that similar terminology is depicted on Figure A-31. The EPA suggests that step 3 be written similarly to the first two steps by stating what is to be estimated first and then describing what information will be used to obtain this estimate, e.g. estimating total activity from the volume associated with the extent of RIM greater than 52.9 pCi/g and the predicted activity concentrations, as described in Appendix B.
f. Because step 3 involves a calculation that relies on information produced from both the indicator kriging and ordinary kriging models, the first bullet on the right column of figure A-31 which states, "*Utilizes same concepts on same data*" is not clear. Further, the second bullet in this list which states, "*Indicators are transformed from concentrations*" is also not clear as to how this relates to the total activity estimations. The third bullet is also not clear. The EPA cannot determine what is meant by "*Updated regression process*" and how this process has been "*improved beyond initial use*". Similarly, the EPA does not understand why the regressions were "*not revisited because of lack of need*". See prior comments related to the development of a sampling objective during the

design investigations aimed at improving the regressions utilized in any geostatistical model. Lastly, the final bullet states that these regressions "*can be improved further with loess line adjustments and error matrix*". This statement is not clear, nor does it indicate whether or when such improvements will be considered.

Approval Condition: The 30 % RD must provide an explicit and clear overview of the activity estimation process and which of the two proposed geostatistical models are being used for each step. The EPA notes that this section states that the details of the activity estimation process are, "*provided in West Lake Activity Calculations for Preliminary Excavation Plan Technical Memorandum (Appendix B of the Preliminary Excavation Plan)*" Additional EPA comments are provided for Appendix B below.

27. **Appendix A Section 6.0, Page 17 of 18.** The last paragraph of the section distinguishes significant depth intervals related to the excavation according to the following: "*isolated pockets of RIM >52.9 pCi/g between 8 and 16 feet B2005GS*" and "*removal of "make-up" RIM between 16 and 20 feet B2005GS.*" The specified depths associated with isolated pockets and deeper excavations are inconsistent with the RODA. In order to reduce confusion and maintain consistency with the RODA, isolated pockets of RIM greater than 52.9 pCi/g must be identified and considered between 8 and 12 feet below the 2005 ground surface (B2005GS) and deeper excavation of relatively higher concentration of RIM must be identified and considered between 12 and 20 feet B2005GS. Approaching the excavation design with these depths will help to ensure that the excavation is sufficiently optimized and in accordance with the requirements in the RODA.

    Approval Condition: A revised approach to calculating activities including the identification of isolated pockets and deeper areas of relatively higher concentrations of RIM that is consistent with the RODA must be included in the 30% Design.

28. **Appendix A Section 6.0, Page 17 of 18.** The last paragraph of this section states, "*This is an exercise of accounting for and balancing of the RIM >52.9 pCi/g and therefore the calculations are of a relative nature. Therefore, this relative calculation is a comparison of activities between different depths (and areas) using the same mathematical process. Because the same mathematical process is applied to each depth interval and / or spatial extent, the relative estimated comparison is valid for the purposes of meeting the RODA requirements.*" The assertion made regarding the validity of comparing total activity estimates is not clear. One way to define the required activity removal associated with the RODA required excavation is as a ratio where the total activity removed with the optimized excavation to the total activity removed to 16 feet B2005GS is equal to one.

$$\frac{TAct(removed)}{TAct_{0-16}} = 1$$

Where:

TAct(removed) = Total activity removed for the RD optimized excavation

$TAct_{0-16}$ = Total activity associated with RIM greater than 52.9 pCi/g that occurs within 16 feet below the 2005 topographical surface

Because this is a ratio, criteria which are identical for both activity estimates, e.g. estimated density, become inconsequential because they cancel out in the ratio. For illustrative purposes:

$$TAct = \sum_{i=1}^{n} (v_i\, a_i)\, \rho_b$$

Where:

$TAct$ = Total activity
a_i = Activity concentration at grid cell i
v_i = Soil volume of grid cell i
n = Number of grid cells where RIM is greater than 52.9 pCi/g
ρ_B = Soil Bulk Density (weight of the dry soil/total soil volume)

So

$$\frac{TAct(removed)}{TAct_{0-16}} = \frac{\sum_{i=1}^{n}(v_i\, a_i)\, \rho_b}{\sum_{j=1}^{n}(v_j\, a_j)\, \rho_b} = \frac{\rho_b * \sum_{i=1}^{n}(v_i\, a_i)}{\rho_b * \sum_{j=1}^{n}(v_j\, a_j)} = \frac{\sum_{i=1}^{n}(v_i\, a_i)}{\sum_{j=1}^{n}(v_j\, a_j)}$$

However, the EPA does not agree that this relationship holds true for the geostatistical model's predictions for individual grid cells. This is because model prediction at each grid cell is dependent upon multiple factors, such as the variogram parameters, unique nearest neighbor hard and soft data locations, relative proximity to hard and soft data locations, and varying reliance on either hard or soft data. Thus, the kriging prediction of activity concentration at each cell is unique, and therefore the model uncertainty and parameter sensitivity must be evaluated similar to the extent model.

Because the proposed activity model is a new model (distinctly different from the prior indicator kriging model), paragraph 3.4 of the statement of work requires the EPA's approval of this model before it can used as a basis for the preliminary excavation presented in the PEP. The PEP does not provide an acceptable description of how this model was developed, nor does it provide any associated supporting information to describe the overall model uncertainty and sensitivity to model parameters. The EPA is therefore unable to fully evaluate the appropriateness of the model and thus cannot approve the use of the model at this time.

Approval Condition: The EPA cannot approve the use of the ordinary kriging model at this time. The development and methodologies for the ordinary kriging model must be explicitly defined and the model uncertainty and parameter sensitivity must be evaluated. These items must be presented to the EPA in order for the model to be considered for approval. This information must be presented in the Revised Excavation Plan in order for the EPA to consider approving this for the use of this model in the remedial design. The EPA suggests that this information be provided sooner to reduce the likelihood of delays to the RD schedule.

APPENDIX B – Total Activity Calculation Technical Memorandum.

29. Appendix B, page 1. The first paragraph provides a description of the intent of the appendix followed by four steps defining the approach to determining various activity estimates. The activity estimates are then compared to the RODA criterion to determine if the optimized excavation complies with this criterion. However, the language that precedes the steps is unclear and the referenced depths appear to be inconsistent with the RODA.

(1) The RODA states on page 65, "*The resulting estimate of radioactivity removed for Alternative 4 must then be achieved during the implementation of the Amended Remedy...*" Alternative 4 required excavation of RIM greater than 52.9 pCi/g to a depth of 16 feet below the 2005 topographical surface.

(2) The RODA also states on page 62, "*During the RD, an evaluation will be performed to identify isolated pockets of RIM to be left at depths between 8 and 12 feet that, if excavated, would require excavation of large volumes of non-RIM waste as overburden and setback.*"

(3) The RODA states on page 65 and 66 that

(a) "*If RIM greater than 52.9 pCi/g occurs between 12 and 20 feet below the surface, then evaluate and excavate where necessary...*";

(b) "*...EPA expects the areas between 12 and 16 feet will be excavated if they are greater than 1,000 pCi/g...*"; and

(c) "*EPA also expects to focus the excavation in the areas between 16 to 20 feet on the higher activity occurrences of RIM (greater than 1,000 pCi/g) if it doesn't add significant excavation of non-RIM waste.*"

As stated in comment 28, isolated pockets of RIM greater than 52.9 pCi/g must be identified and considered between 8 and 12 feet B2005GS and deeper excavation of relatively higher concentration of RIM must be identified and considered between 12 and 20 feet B2005GS to ensure that the excavation is sufficiently optimized and in accordance with the requirements in the RODA.

30. Appendix B, page 1. Step 2 of the logic for the total activity analysis lists the sub-intervals for defining the calculations of activity: 0 to 8 feet, 8 to 12 feet, 12 to 16 feet, and 16 to 20 feet. The EPA agrees that consideration of these sub-intervals for activity estimates is useful for developing the optimized excavation. The EPA requests that future sub-intervals at a minimum be evenly spaced, i.e. 0 to 4 feet and 4 to 8 feet, and if feasible, smaller, e.g., one or two-foot intervals.

Step 3 of the logic for the total activity analysis states, "*identification of isolated pockets of RIM >52.9 pCi/g that may potentially be left in place, while making up activity removal in intervals between 16 and 20 feet B2005GS where material is within the footprint of the known area to be excavated.*" This statement is not clear and appears to be inconsistent with the RODA (See comment 28.)

The RODA states on page 64, "*The EPA expects the deeper excavations to focus on removing higher concentrations of RIM that don't also require excavation of significant volumes of non-RIM waste. When selecting areas for excavation deeper than 12 feet, the EPA specifically prefers removal of RIM at concentrations greater than 1,000 pCi/g. In addition, any isolated pockets of RIM that are not excavated between <u>8 feet and 12 feet</u> will be offset by removing additional radioactivity between <u>12 feet and 20 feet</u> elsewhere in Areas 1 and 2.*" Therefore, the EPA expects that deeper excavations will generally be located underneath excavations that extend to or near 12 feet below the 2005 topographical surface because these locations should generally not require excavation of significant volumes of non-RIM wastes. Additional consideration must be given to the concentrations of RIM in these proposed deeper excavations as the RODA expects a focus on higher concentrations, preferably greater than 1,000 pCi/g.

Approval Condition: All the steps of the total activity logic must be revised and presented in the 30% RD. A complete optimized excavation proposal that fully considers this comment and the other referenced comments must be presented in the Revised Excavation Plan.

31. **Appendix B, page 1 and 2.** The last paragraph on page 1 of Appendix B that continues to the top of page 2 contains similar language to what is provided in Section 6.0 of Appendix A (see comment 28). Similarly, the language discussing the "*relative nature*" of the activity estimates is unclear. The kriging prediction of activity concentration at each cell is unique and therefore the model must be evaluated to the extent necessary to demonstrate it is reasonable for its intended use.

32. **Appendix B, pages 2 and 3.** Equations 1, 2, and 3, bolded statements 1) and 2) and the associated discussion are adequate for generally describing approach to the activity calculations associated with these first two steps. However, some improvements are needed for clarity and must be addressed in the 30% RD:

    a. The text before equation one describes volume being considered in the activity calculation associated with an activity between two depths somewhat generally. It is difficult to determine whether the total volume of RIM greater than 52.9 pCi/g within that depth interval or the volume of an individual grid cell is being considered for the calculation. Further, "v_i" from equation 1 is defined as the "*Soil volume of grid cell i*". However, based on other statements provided in the PEP for the IK extent model, it appears that the grid cell volume is constant, i.e. 15 meters by 15 meters by 6 inches. If the volume of each cell is constant as was used for the IK model, then the equation could be simplified further by dropping the subscript "*i*" from the volume term, thereby allowing it to be treated similarly to bulk density from a mathematics perspective. The EPA acknowledges that this is discussed further in the last paragraph of page 6 and the equation at the top of page 7 of Appendix B. However, it is unclear how this discussion and equation relate to or can be applied to Equation 1 on page 2.

    Approval Condition: Revisions to Equation 1 and the associated text must be presented in the 30% RD to provide clarification.

b. As stated in comment 30, the EPA requests that future sub-intervals be evenly spaced, at a minimum. This will require revisions to bolded statement 2), the associated text, and Equation 3. The EPA believes that understanding and estimating the activity-versus-depth relationship is useful for developing an optimized excavation that will meet the RODA requirements. However, additional consideration should be given to the corresponding volume represented by the RIM in each interval.

   Approval Condition: Revisions to bolded step 2), the associated text, and equation 3 must be provided with the 30% RD. A table presenting the volume of RIM associated with these sub-intervals must be provided with the DIWP.

c. The first paragraph underneath the table on page 3 states, "*These values include small unverified anomalies that may be defined by an isolated sample and/or larger isolated pockets.*" This statement is unclear. All the model predictions for locations within Area 1 and Area 2 where there is no data (hard or soft) are "unverified" and thus, the EPA cannot understand the significance of the term "unverified" in this context. Further, additional information is needed to understand what is considered "small" and why these predicted RIM occurrences are considered anomalies. Very little information was provided about the details of the ordinary kriging model (activity) model such that the EPA cannot determine whether it is possible for a kriging prediction at an individual cell to be defined by "an isolated sample". The EPA notes however, the SSP&A Report states in Appendix D-2, on page 1702, "*The indicator value at the center of each node of the grid was calculated by weighting the indicator values of all hard and soft data identified within an anisotropic search radius of 280 ft in the horizontal direction and 20 ft in the vertical direction. A minimum of six (6) data points was required for interpolation to take place: no value was calculated for nodes where there were less than 6 data points (combined hard and soft) within the search radius.*" Similarly, it is unclear how "*small unverified anomalies*" could be defined by "*larger isolated pockets*".

   Approval Condition: This topic must be considered to the extent necessary to determine if a sampling objective is necessary in the DIWP and additional discussion to clarify the meaning of this statement must be provided in the 30% RD.

33. **Appendix B, pages 3 and 4.** Bolded statement 3 and the associated text is not consistent with the additional condition quoted from the DCR related to the optimized excavation described in the RODA. Revisions to this portion of the activity estimate process need to be revised consistent with the direction given in comment 7.

    Approval Condition: Bolded statement 3, the associated text, and Equation 4 must be revised consistent with Comments 7 to be presented in the 30% RD.

34. **Appendix B, Equation 4, Page 4:** The document states: "*$A_{IP\,z_1\text{-}z_2}$ = Area of isolated pockets of RIM over the depth interval z_1 to z_2.*" It appears that the word "*Area*" should be changed to "*Activity*".

Approval Condition: This discrepancy between "area" and "activity" must be clarified and presented in the 30% Design.

35. **Appendix B, page 5.** The text describes that Ra and Th concentrations were added together and the total concentrations were kriged for the activity model. This is inconsistent with the extent or indicator kriging model. It is recommended that the two radionuclides be kriged separately and then summed for the predicted concentrations at each location after kriging as part of the total activity calculation. This will confer two primary benefits:

    1) maintain consistency with the extent model and improve overall cohesiveness of the study, and
    2) allow for the comparison of the extent and activity models, which is an important piece of the validation of both models.

    Because the models are combined in the activity calculation by cropping the surface produced by the activity model using the shell of RIM produced by the indicator kriging model, the activity estimates depend on agreement between the two models. Without comparing the two models, reasonable confidence in the activity estimates cannot be achieved.

    Approval Condition: Any differences between the modeling approach or methodology utilized in the activity model and the extent model must be explicitly listed and supporting rationale provided as to why these differences are reasonable. Because the respondents have proposed to use both models together to design the excavation, the two models should be made as consistent as possible to increase the likelihood of obtaining the EPA approval for the model in the future.

36. **Appendix B, page 6.** The current data selection approach for the activity model described in the text uses hard data preferentially for those locations where it is available, and soft data otherwise. Where multiple estimates were available for either hard data or soft data, "*the lowest activity concentration was retained to limit anomaly-based irregularities*" (p. B-5). The defensibility of this approach is limited because the purpose of the geostatistical model is to identify pockets of anomalously high activity, and its capability to do this is limited if anomalous data are removed. If measurements vary drastically by location, this probably indicates either:

    1) significant measurement error, which should be evaluated during data validation;
    2) the sampling interval may need to be further discretized; or
    3) there is a great deal of heterogeneity that should be retained and modeled spatially.

    A more defensible approach would be taking the mean or maximum sampled result for locations with multiple estimates. The EPA also notes that the approach to selecting a data type when multiple types of hard and soft data are collocated appears to be different than the approach utilized in the extent model.

    Approval Condition: Any differences between the modeling approach or methodology utilized in the activity model and the extent model must be explicitly listed and supporting rationale provided as to why these differences are reasonable. Because the respondents have proposed to use both models together to design the excavation, the two models should be made as consistent as possible to increase the likelihood of obtaining the EPA approval for the model in the future.

37. **Appendix B, Page 5**: The document states, "*As no new data has been collected since the Final Feasibility Study (FFS), the SSP&A continuous cumulative distribution function (CDF) regression equations for the transformation of soft data (normalized gamma responses) to combined radium and combined thorium values were retained from SSP&A's geostatistical analysis.*" However, the document also states CDF Equations C-3 and C-4 from Section 1.2.2 of Appendix D of the SSP&A Report were selected. These equations were developed for the sensitivity analysis performed by SSP&A (method 2 CDFs) and were not retained for use in the final version of the SSP&A model. Therefore, the CDFs used for the activity model are not consistent with the CDFs used for the extent model.

    Approval Condition: Any differences between the modeling approach or methodology utilized in the activity model and the extent model must be explicitly listed and supporting rationale provided as to why these differences are reasonable. Because the respondents have proposed to use both models together to design the excavation, the two models should be made as consistent as possible to increase the likelihood of obtaining the EPA approval for the model in the future.

38. **Appendix B.** Adding a second model to the process of developing the optimized excavation adds significant complexity to the remedial design as discussed in comment 1. However, the EPA believes that an ordinary kriging model could be used in place of an indicator model and offers several benefits if applied appropriately for use in this remedial design. The EPA acknowledges that one disadvantage of the activity model, as presented, is that the variability in the inputs are currently underestimated due to the deterministic transformation of soft data into concentrations. However, there are methods that could simulate a dataset that reflects uncertainty in the regressions. For example, if random error were added to the transformations given on page B-5 to incorporate uncertainties in the regressions, a random draw could be taken from a normal distribution with a mean of zero and a standard deviation equal to the standard error of the regression for each soft data estimate.

    Approval Condition: Further consideration must be given to utilizing an ordinary kriging model for estimating the extent of RIM greater than 52.9 pCi/g and various activities for the optimized excavation. Discussion of the merits of utilizing ordinary kriging as the sole RD model must be included in the 30% RD.