Boeing Comments On The Radiological Background Study Report
Santa Susana Field Laboratory, Ventura, California

Technical review or clarification comments on the United States Environmental Protection Agency (USEPA) report *Radiological Background Study Report, Santa Susana Field Laboratory, Ventura, California* are provided below. General Comments are provided in Table 1 and Specific Comments in Table 2.

Table 1: Radiological Background Study Report General Comments

General Comment Number	Comment
1	Boeing appreciates the significant effort that has gone into the Radiological Background Study (RBS) and the "Statistical Methods" paper and commends EPA, HGL, Anita Singh and all others who have contributed to the study. Overall, the final report reflects the attention to detail that went into the project planning and the field work.
2	The use of full terms (e.g., Distance test locations) and acronyms/abbreviations (e.g., DTL) is mixed throughout the document and can be confusing. Suggest defining them the first time they are used in each section and then just use the acronyms/abbreviations thereafter.
3	If risk-based standards are included for comparison to background, then the suburban resident land use scenario should also be included.
4	An objective of the study is to minimize both false positives and false negatives when applying BTVs for onsite data comparisons. Since there are cases when the background study BTV is less than the measured maximum in the background dataset, false positive errors will occur when compared to onsite data (i.e., contamination will be identified when it is really background). In the case when the BTV is less than the maximum value, we suggest that EPA and risk managers consider either revision of the BTV to the maximum measured value, or inclusion of an additional step that allows consideration of the maximum detected value for cleanup decisions.
5	Boeing believes that prior comments by Tom Rucker ("Comments on SSFL Radionuclide Background Data Sets and their Statistical Treatment" 6/20/2011), Abe Weitzberg and others are still pertinent. These included, • Possible false detects for many radioisotopes including, Nb-94, Cs-134, Sb-125, Eu-155, Ho-166m, Na-22 and Te-125m, are based on misidentification due to interference with gamma peaks from naturally occurring radionuclides. Since EPA radiochemists acknowledge these results are not real, and detection limits vary between laboratories, the use of the same library for on-site measurements will not eliminate the possibility of similar false detects for these radioisotopes during the Area IV sampling program. • Rational for eliminating many of the U-238 and Th-232 daughter products from the AOC look-up table (EPA concurs with this position in Section 9.5) • Including both Cs-137 and its daughter Ba-137m in the look-up table should be avoided. EPA specifies a BTV for Cs-137+D (Cs-137 plus

General Comment Number	Comment
	Ba-137m) in Table 8.4 and a separate BTV for Ba-137m in Table 8.2. • Problem with applying the Kaplan-Meier process to uncensored data (see comments on Appendix B below, relating to application of the Kaplan-Meier process to uncensored data) • Lack of any background data for sediments, drainage channels and evaporative concentration areas

Table 2: Radiological Background Study Report Specific Comments

Page (pg) # and Paragraph (pp) #	Comment
MAIN TEXT	
Pg 2-1, pp6	Suggest description of the geology in Area IV at SSFL include additional detail. While 80% of Area IV is underlain by the Chatsworth Formation, the western end and northern edge of Area IV, especially the drainages to the north, are within or receive drainage from the Santa Susana Formation.
Pg 2-4, pp4	Suggest that the text describing the RBRAs (Chatsworth and Santa Susana formations) acknowledge that there is some uncertainty in the native concentrations of naturally-occurring radionuclides since geological formations also have variable sub-formation strata (shale, versus siltstone, versus conglomerate), and the concentrations may vary between them.
Pg 2-5, 2-6	For completeness, the descriptions of the DTLs should include the underlying geologic formations.
Pg 3-1, pp6	Since some locations were modified during field work, suggest documentation of change in a table that gives the sample number and the reason for the location movement. This text should identify how many samples in each of the three RBRAs were moved from the original randomly located position.
Pg 3-2, pp1	Clarify the criteria for a gamma measurement being classified as an anomaly.
Pg 3-2, Section 3.3	It is unclear how surface samples were collected, so additional description should be added to the text. Were these discrete samples, collected at <6" below ground surface, and/or multiple sleeves collected and composited across an area? Please describe how the surface sample collection differed from the subsurface composite sample collection.
Pg 3-3, Section 3.4	Subsurface samples should be clearly defined, based on the sampling methodology, as 'composite' samples collected over the entire subsurface sampling interval.
Pg 4-2, Section 4.2.4	Sampling equipment decontamination is generally followed by some type of quality control sampling (i.e., equipment rinsate blanks) to confirm the quality of the decontamination process. The report should describe whether these types of quality control samples were collected.

Page (pg) # and Paragraph (pp) #	Comment
Pg 5-1, pp3 And Table 5.1	The gamma anomaly detected at TP-16 needs further description, including 1) whether TP-16 is a DTL or RBRA location and 2) how the +/-30% readings was selected as the criterion for an anomaly. Table 5.1 suggests that the TP-16 anomaly is not either a high or low reading but a range which is larger than the other DTLs. The highest value is still consistent with the gamma measurements for other DTLs in this quadrant. The rationale for elimination of TP-16 should be further described.
Pg 6-3, Section 6.2.1	Suggest that the description of “additional uncertainty” include potential sources, magnitude (in comparison to both expected uncertainties and detection limits) and consequences (in terms of data evaluation). In addition, the third paragraph in this section seems to limit the data use of the data to developing an overall range of background radionuclide concentrations and not to determining location-specific background. Suggest further discussion/explanation of this as it is important to understand the ways in which the data should and should not be used.
Pg 6-4, Section 6.2.3	The acceptable difference between primary and duplicate samples has been increased by 10% to account for under-estimated variability in background concentrations. The discussion is based on sigma (σ) and Z-values. It would be helpful to also include the percentage range of acceptable differences in the text since this is also a common measure of duplicate samples. It seems that the discussion has increased the range from +/- 20% (2σ or Z=1.96) to +/- 30% (3σ or Z=2.58), however this is not clear in the text.
Text Section 7.2, Appendix A, Appendix B	The univariate outlier tests available in Scout 2008 Version 1.00.01 includes Dixon’s Test and Rosner’s Test, which were used to identify outliers as indicated in Appendix A. Both Dixon’s Test and Rosner’s Test assume the data are normally distributed. Were the data checked for normality prior to applying these outlier tests? Were there datasets that are not normally distributed? Are there applicable outlier tests for data that were not normally distributed? Suggest additional text description to clarify this process.
Page 7-1 to 7-3, Section 7.3	The level of detail describing each of the DTL comparisons is not the same for each radionuclide. Suggest that presentation regarding the levels of significance of the tests be presented.

Page (pg) # and Paragraph (pp) #	Comment
Section 8.0	Suggest additional text to describe rationale for outlier exclusion given the amount of EPA's research for RBRA selection and the conclusion from the DTL study samples that the RBRAs were not affected by SSFL operations. Given the solid foundation for the background sample locations and the DTL conclusion, please carefully consider exclusion of any data from the dataset and provide rationale as to why the data were excluded, As described in EPA's 2006 document entitled: Data Quality Assessment: Statistical Methods for Practitioners, EPA QA/G-9S. EPA/240/B-06/003), statistical test identification of outliers is not recommended. The EPA document states the following: *"One should never discard an outlier based solely on a statistical test. Instead, the decision to discard an outlier should be based on some scientific or quality assurance basis. Discarding an outlier from a data set should be done with extreme caution, particularly for environmental data sets, which often contain legitimate extreme values. If an outlier is discarded from the data set, all statistical analysis of the data should be applied to both the full and truncated data set so that the effect of discarding observations may be assessed. If scientific reasoning does not explain the outlier, it should not be discarded from the data set."* If exclusion is solely based on statistical test results, these 'outlier' data may likely be part of the background. See comments on pp. B-2 below. Suggest each identified outlier be listed in a table and rationale provided for exclusion, and consideration of these outliers be included in cleanup planning.
Pg 8-1 to 8-3, Section 8-1	There are a few instances where the USL95 is lower than the maximum in the dataset, which may lead to increased Type I error rates if applied for onsite data comparisons. Suggest EPA consider other statistical parameters, including the maximum detection, for the BTV or adding a second comparison step (see General Comment 4).
Pg 8-3 to 8-4, Section 8.2.1	See General Comment 4 regarding false positives. Suggest including a discussion regarding how the selected uncensored ND values compare to the detected concentrations and how the selection of the maximum uncensored ND will affect the objective of minimizing false positives when the BTV is used for onsite comparisons. Also, see comments for Appendix B below regarding use of uncensored non-detect data.

Page (pg) # and Paragraph (pp) #	Comment
Sections 9.1 thru 9.5	Boeing in general agrees with EPA's suggestions to utilize combined BTVs ("management decisions") in an effort to simplify comparison to onsite data and remedial decisions, and since the RBRAs were identified as un-impacted background locations. Further, use of a combined BTV would reflect actual site soil conditions. For example, much of the soil at SSFL has been excavated and mixed either during initial construction, operations, or during demolition. Therefore, for comparison to onsite concentrations, surface and subsurface background datasets would need to be combined in order to have an appropriate and representative BTV. It is also the case that locations at SSFL have mixed Chatsworth and Santa Susana formation soils and therefore the selection of a BTV that includes only one of these formations may increase the number of false positives when the BTVs are used onsite.
Section 9.1	Since PRGs are risk-based goals "incremental or in addition to background", it could be argued that the Lookup Table value should always be BTV + PRG. Depending on the relative sizes of the PRG and BTV, this summation would default to a Lookup Table value of PRG (if PRG >>> BTV) or BTV (if BTV >>> PRG).
Table 8-1	The selection of the highest uncensored ND as the BTV increases the probability of false positives when these values are used onsite since seven of the radionuclides were detected above the highest uncensored ND. While many of these reported detections and highest uncensored NDs appear within reasonable analytical variability, two radionuclides have reported detections approximately an order of magnitude (10-times) higher than the highest uncensored ND. The report concludes that these reported detects are not real, but they could occur onsite. See General Comment 4. Suggest BTVs for these two radionuclides be re-evaluated, and carefully considered for how they may be used for cleanup planning since they were detected in the background dataset.
APPENDIX A	
Appendix A	It appears the distribution test results were not summarized in the outputs in Appendix A for each step in which the distribution test was performed. It would be clear what tests were used if the normality test results were provided for each step.
Appendix A	The statistical comparisons between RBRAs are sometimes conducted between only the two Chatsworth RBRAs, and sometimes between all three RBRAs. Was the choice based on a visual inspection of the box plots? Please clarify this in the Appendix B text.
APPENDIX B	
Pg B-1, pp3	"However, the Project Team and the stakeholders decided to use univariate methods as described in this appendix." – A summary of the rationale/benefits of using univariate statistics instead of multivariate statistics would be beneficial, and, perhaps, an example provided.
Pg B-2, pp1 and Pg B-3, pp5	Please see Section 8.0 comment above regarding outlier analysis and exclusion. Suggest table of outliers be included and rationale provided.

Page (pg) # and Paragraph (pp) #	Comment
	Also, as stated on page 2 of Appendix A (regarding Cs-134 statistical analysis), some statistical analysis was performed using outliers as well as the truncated dataset. Please clarify where calculated statistical results with and without outliers are published.
Pg B-4, 3rd main bullet, 2nd sub-bullet	The text in this bullet is unclear, please clarify.
Pg B-3, Section 2.0	Statistical tests and examples of when the tests can be used are described. However, the tests are not listed in the order of when and what statistical tests should be conducted. It would be helpful if a flow chart was provided that describes the rationale for which statistical tests are used, and when and why they are used.
Section B2.0	For data that are normally distributed, the data for the three RBRAs were compared using a one-way ANOVA. Was the 2-way ANOVA considered to account for potential interactions between RBRAs and surface/subsurface soil?
Pg B-10	The discussion of the USL95 states that this statistic is expected to be above all measured background observations. However, for some of the radionuclides presented in Section 6 of the main report (Tables 8-3 to 8-7) there are measured observations that fall above the USL95. This seems like a contradiction with the statement above. See General Comment 4 for consideration of false positives if these BTVs are used for comparison to onsite data.
Page B-17, Section 4.0	The paper states *"Some technical stakeholders believe that radionuclide data consisting of NDs (positive as well as negative results) should be treated as detected data. They suggest that one should ignore the ND status of radionuclide concentrations and their detection limits/MDCs. All detected as well as ND values should be treated equally in the computation of various statistics of interest including BTV estimates. They do not acknowledge the fact that in practice concentrations cannot be negative." (Red text emphasis added)* Boeing believes this statement is incorrect as explained below:

Page (pg) # and Paragraph (pp) #	Comment
Page B-17, Section 4.0	**1. Censored vs. Non-censored Data** In the measurement of chemicals and in most of the literature on statistical treatment of “no-detects”, ND refers to a semi-quantitative value such as <5. <5 means the chemical laboratory cannot quantify the measurement other than to say it lies somewhere between 0 and 5 where 5 is a reporting limit. This data point is said to be censored or left-censored, meaning we have no knowledge of the “true” value to the “left” of 5. Indeed, one of the key references used by the paper and the source of the Kaplan-Meier (K-M) Method, is “Nondetects and Data Analysis – Statistics for Censored Environmental Data” by Dennis R. Helsel (underline added). Note the use of the term “censored” in the title, implying that these methods are to be used for data sets including <MDC data, but not for uncensored data. In contrast, radionuclide data is reported as quantitative numbers, that may be detects (above the MDC), positive non-detects (below the MDC) or even negative numbers (also less than then MDC). Therefore, a radionuclide ND is a quantitative number, e.g. 3, and is not reported as <5 even though the MDC may be 5. Measured, reported radionuclide results are therefore un-censored or non-censored, even if they are NDs or less than the MDC. The K-M method is used for treating chemical data sets that include some left-censored ND data such as <1, 5, <2, 6, 7, <3 using the methods discussed in the paper on pages B-14 through B-16. It should not be used to treat radionuclide data that includes some un-censored ND data less than the MDC of 5 (e.g., results such as 1, 5, 2, 6, 7, 3). All radionuclide data is based on measurement and is reported as uncensored data. As such, it should be treated statistically as uncensored data and included directly, as is, in the BTV calculations, and not censored. The classic statistical reference “Statistical Methods for Environmental Pollution Monitoring” by R. O. Gilbert (and also referenced in the paper), states on page 178, “ … *reporting of actual concentrations is the best procedure from both practical and statistical analysis points of view … It is strongly recommended here that, whenever the measurement technique permits, report the actual measurement, whatever it may be, even if it is negative.* ”

Page (pg) # and Paragraph (pp) #	Comment
Page B-17, Section 4.0	**2. Negative Concentrations?** The paper states *that "They [stakeholders] do not acknowledge the fact that in practice concentrations cannot be negative."* Although it is true that one cannot have a negative concentration, a negative value reported by the laboratory does have value and meaning. This is because a laboratory does not directly measure concentrations. It measures the number of radioactive particles detected during a fixed count period from a sample that exceeds the instrument background. The net count rate can be negative under certain conditions. This net count rate is then used to calculate a concentration using sample mass, count time, detection efficiencies, geometric factors, unit conversions etc. All radionuclide analysis involves counting a number of radioactive decays (either gammas, alphas or betas) emitted by the sample per unit time within a low-background laboratory counter. Even though counters are shielded to minimize any extraneous radiation entering from the outside or from within the equipment itself, there will always be a low level of radioactive particles detected even with no sample present. This is known as the instrument background, which is measured by counting a non-radioactive blank. For example, if the instrument background is measured at 10 counts per minute (cpm). The MDC expressed in cpm will be 2 x 1.645 x $(2 \times 10)^{1/2}$ = 14.7 cpm. If a sample that is <u>not radioactive</u> is counted 10 separate times, we would measure 10 cpm each time. However, since we are counting background <u>plus</u> the sample (gross count), and since instrument background is variable and will fluctuate during each of the counting periods, we may measure the following gross counts. 10, 11, 12, 9, 9, 10, 7, 13, 11, 8 Subtracting the single instrument background count of 10 cpm and ranking, we get the following net counts. -3, -2, -1, -1, 0, 0, 1, 1, 2, 3 Note that some are negative net counts, and all are less than the MDC of 14.7 cpm, therefore all are considered non-censored NDs. The simplest parametric statistic for this data set is the arithmetic mean which is calculated to be 0 cpm, which correctly confirms the prior statement that the sample is non-radioactive. However, if we were to dismiss the negative net counts as meaningless, the mean of the reduced data set of 0, 0, 1, 1, 2, 3 would be 1 cpm, which would incorrectly imply the sample exceeded background.

Page (pg) # and Paragraph (pp) #	Comment
Page B-17, Section 4.0	Likewise if we were to censor the data set and report all the data as <MDC, the data set would be <14.7, <14.7, <14.7, <14.7, <14.7, <14.7, <14.7, <14.7, <14.7, <14.7 By excluding negative measurements, valuable information is lost and parametric or non-parametric statistics calculated based on this censored data set do not give the correct conclusions. Typically the instrument background count is established once per batch of multiple samples. The instrument background count is therefore measured at a different earlier time than the subsequent batch of samples, which themselves are counted consecutively at different times. Thus, the contribution of instrument background to the gross count for each sample can and does vary between each sample in a batch. In this way, negative net counts and subsequently negative "concentrations" sometimes occur. In summary, dismissing negative radiochemical data is not recommended.
Pages B-21 and B-22, Section 4.0	The paper states *"It is not clear whether USL95 and UTL95-95 represent non-detects or detects."* USL95 and UTL95-96 are calculated test statistics - not measured data points. Therefore it is inappropriate to refer to them as detects or non-detects. They simply represent calculated upper level estimates of sets of uncensored measured data.