

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
1	PRP-Integral	Summary and Conclusions	EPA presents a final summary of the findings of its BERA in the Executive Summary and in the Summary and Conclusions. EPA also lists a number of conservative assumptions made where necessary to overcome uncertainty. However, EPA does not present a final conclusion about the ecological significance of estimated risk. EPA states that the purpose of the BERA, along with other relevant information, will be used by risk managers to decide whether remedial actions are needed to protect ecological receptors from adverse effects attributable to site-related environmental contaminants. Risk management decisions (Step 8 of the BERA process; EPA 1997a) will be informed by EPA's description of the magnitude and potential ecological significance of the estimated risk. In essence, the BERA should clearly identify a subset of the chemicals of potential ecological concern (COPECs) as "risk drivers" in site media that contribute most significantly to unacceptable risk for purposes of remedial planning in the feasibility study (e.g., development of risk-based remedial goals).	Risk drivers are currently identified in the Summary and Conclusions, but a summary table will be added as well that highlights key findings.
2	PRP-Integral	Summary and Conclusions	Risk estimates (hazard quotients [HQs]) account for one line of evidence. Other lines of evidence, quantitative or qualitative, and the uncertainty analysis should be used to draw risk conclusions. Although the conclusions of the BERA summarized the risk results, it was not always clear how the other lines of evidence or the uncertainties were considered in ultimately identifying risk drivers. For example, in Section 7, EPA reports barium and manganese as primary risk drivers in the ponds but also summarizes HQ results/exceedances from other metals without characterizing the significance of the risk. Specific examples are further discussed in this comment letter. We suggest that EPA present a synthesis of lines of evidence that indicates its overall degree of confidence in the risk estimates, and interprets the potential adversity to assessment endpoints (EPA 1998) and categorize the following: • COPECs that pose no unacceptable risk and warrant no further evaluation • COPECs that pose unacceptable risk to ecological communities/wildlife populations based on lines of evidence and require risk management decisions • COPECs for which risks are uncertain and confidence in the results is low and may require further evaluation in the remedial investigation. A synthesis of risk conclusions, in addition to an understanding of background risk, will meet the objectives of the BERA. This will help guide risk management decisions to advance from COPECs in the screening level assessment (Step 1 of the BERA process) to identification of risk drivers (Step 8 of the BERA process) that can be considered chemicals of concern (COCs) for the feasibility study.	Risk drivers are currently identified in the Summary and Conclusions, but a summary table will be added as well that highlights key findings.
3	PRP-Integral	Data Evaluation	EPA presents several 2-way statistical comparisons between chemical concentrations in environmental media upstream of the former Frenchtown Mill with locations adjacent to and downstream of the site. The underlying assumption or premise of this approach is that the former Frenchtown Mill influences the environmental conditions reflected in the abiotic media (i.e., sediment, surface water) in the Clark Fork River from the upstream end of the site to at least Alberton, which is about 18 river miles downstream. This	The scope of the risk assessment focuses on evaluating potential exposures and associated risks within the current study area which includes areas within the Clark Fork River (CFR) where hazardous substances

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			assumption is likely incorrect and is certainly not supported by the available data. The assumption is also not supported by any of EPA’s conceptual site models. EPA should reconsider this assumption in light of evidence that other sources of hazardous substances exist both upstream and downstream of the site. EPA should more closely evaluate the evidence downstream (see Appendix F of Integral’s Addendum 9 Data Interpretive Report [Integral 2020]), and note that Stations 57-CFR, 58-CFR, and 59-CFR, 3 to 4 miles downstream, are the locations of some of the highest concentrations of metals and dioxin and furan congeners in sediment sampled in 2018. Concentrations of various analytes here and downstream of these locations are often higher than anything measured adjacent to the former mill. The most obvious potential sources of chemicals to the river in this area are Frenchtown itself, the local golf course, and possibly other developments in the area of the sampling stations listed above. Regardless of the reason for the variation in sediment chemistry in that area, these data undermine the basic assumption underlying statistical comparisons that lump all stations adjacent to and downstream of the site together for comparisons to conditions at the group of upstream stations. Only those sampling stations within the Clark Fork River that are directly adjacent to the site should be considered potentially representative of the site.	from site activities may have come to be located through typical transport processes. Within a dynamic river system, it is unlikely that sediment transport downstream is not occurring. Thus, the underlying assumption of the comment that contamination from the Site entering the CFR through outfalls and groundwater movement results in deposition only adjacent to the site with no downstream movement is unlikely and has not been definitively demonstrated. Nonetheless, the risk presentation tables and figures in the BERA denote which sample locations are directly adjacent to the Site and which are downstream. Evaluation of full nature and extent and source attribution are outside the scope of the risk assessment, but will be included within the Remedial Investigation/Feasibility Study (RI/FS). Such information will be considered by risk managers in conjunction with the findings of the risk assessments during decision making. Based on the comments received on this issue, the term “site” will be replaced with “study area” within the text of the risk assessment.
4	PRP-Integral	Manganese in the Clark Fork River	EPA’s representation of manganese risks in the Clark Fork River contains conceptual flaws and overlooks important information. EPA erroneously equates all locations sampled adjacent to and downstream of the site as indicative of site conditions, which as described above, is not supportable. EPA also mixes data sets inappropriately in a key line of evidence and does not recognize important uncertainties about dissolved manganese in the Clark Fork River. The data used by EPA are not clear. EPA’s ProUCL output reports the use of 19 records for the downstream component of the upstream/downstream comparison. The database contains 18 records in EPA’s downstream group for dissolved manganese in the Clark Fork River, including four from a flood event in May 2018. Including results	As noted in the response to comment #3, the risk assessment evaluates the current study area defined to include locations in the CFR where hazardous substances from site activities may have come to be located. As noted in the BERA, the full nature and extent to which hazardous substances from site activities have come to be located in the CFR adjacent to and downstream of

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			from the May 2018 flood event in the upstream/downstream comparison is inappropriate, because it represents a functionally different condition from that of the upstream data set, which is entirely from sampling conducted during normal flows. - EPA states that dissolved manganese is statistically significantly higher adjacent to and downstream of the site than upstream. Details of the comparison on PDF p. 287 (Appendix C) show that EPA's comparison included the uncertain result for dissolved manganese at Station 52-CFR collected in August 2018, of 1,370 µg/L. This value is qualified as uncertain because the total concentration of manganese was reported by the laboratory at 89.3 µg/L. The second highest value for dissolved manganese in the Clark Fork River was 486 µg/L at Station 57-CFR about 4 miles downstream, also higher than the total concentration at the same location reported at 330 µg/L. While this sort of inconsistency, when the total metal is reported at a concentration lower than the dissolved concentration, does occur, the example at Station CFR-52 for manganese reflects a significant uncertainty. Not thoroughly addressing this uncertainty results in a misleading presentation. We recommend EPA revise its risk assessment for manganese in the Clark Fork River to: - Consider only those stations directly adjacent to the site as representing "site" influence or conditions - Remove flood-stage manganese data from statistical comparisons with upstream data that are not from flood stage conditions. - Thoroughly address the uncertainty associated with use of the dissolved manganese concentration and impacts to the risk conclusions. These changes may lead to small samples sizes and change the statistical power and ability to detect differences with confidence. Resulting uncertainties should be thoroughly presented.	the Site has not been fully established but will be considered within the RI. Appendix A presents which data are included in the evaluation of the study area. Surface water data collected in May 2018 will remain in the dataset used to evaluate exposure to surface water in the BERA. The BERA currently includes discussion of the uncertainty in the HQs based on these data. Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment. The BERA includes discussion of the uncertainties associated with the reported dissolved manganese concentration of 1,370 µg/L in the SW52-CFR sample. No changes will be made on this basis.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
5	PRP-Integral	Assessment Endpoints	Selecting appropriate assessment endpoints is an important step in a BERA because it is useful to risk managers in the remedial decision-making process (EPA 2003a). Assessment endpoints define both the valued ecological resource and a characteristic of the resource to protect (EPA 1997a, 1998, 2003a). As stated in Section 2.5 of the BERA, “Measures of exposure and effects (also called measurement endpoints) represent quantifiable ecological characteristics that can be measured, interpreted, and related to the valued ecological components chosen as the assessment endpoints” EPA used only the estimates of HQs to predict risk to ecological receptors. Although other data are available, they were not used as additional lines of evidence in the BERA. Using multiple lines of evidence, as discussed below, provides a more robust approach for interpreting the risk results and evaluating assessment endpoints.	Assessment and measurement endpoints for the BERA were described in the BERA workplan. Additional lines of evidence, such as toxicity tests or community surveys, would be beneficial to address uncertainties, but those data are not available for the Site.
6	PRP-Integral	Risk Characterization	Table 2-4 of the BERA establishes plant communities and invertebrate communities as the assessment endpoint, not individual organisms. EPA’s conclusions on risk to plants and soil invertebrates are based on exceedances of screening levels. Exceedances of generic soil screening levels on the site at a subset of locations sampled does not represent soil conditions nor reflect harm to the assessment endpoint. Soil screening levels for plants and invertebrates were selected from published sources (e.g., Oak Ridge National Laboratory [ORNL] and ecological soil screening levels [EcoSSLs]). As noted by the authors, these values are conservative estimates of toxicity for plant and invertebrate communities and do not necessarily indicate that adverse effects are occurring. In many cases, the screening levels are based on a limited data set and there is low confidence in the ability of these screening levels to predict risk to plant and soil invertebrate communities. An ORNL document (Efroymson et al. 1997a) states that “If chemical concentrations reported in field soils that support vigorous and diverse plant communities exceed one or more of the benchmarks presented in this report or if a benchmark is exceeded by background soil concentrations, it is generally safe to assume that the benchmark is a poor measure of risk to the plant community at that site.” In most cases, the available toxicity data are limited to only one or two studies for which the tested species, soil types, and exposure conditions are unlikely to be present at the site. For screening levels from Efroymson et al. (1997a,b), the authors caution that the values are conservative and uncertain, and recommend using observations of toxicity at the site to confirm risk predictions. Many of the inorganics are naturally occurring in this region and in the state (Hydrometrics 2013). Soil screening levels below background (Table 5-5; barium, copper, selenium, and zinc) also indicate substantial uncertainty in the screening values, as ecological communities have been present at the sites over evolutionary time and exposure to background concentrations of constituents is not expected to result in unacceptable risk or harm to the assessment endpoint.	EPA agrees that community level studies would provide additional information to address the uncertainties associated with the predicted toxicities line of evidence. In the absence of such information, the BERA evaluates the frequency and magnitude of HQs exceeding 1 and describes the associated uncertainties. Further, HQs are not inherently generic and conservative. Their level of specificity to a particular receptor and level of conservativeness are based on assumptions and values used in calculating receptor exposure and the value of the TRV selected for comparing the exposure levels. Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001).

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			Observations that support the health of plant communities at the site are considered a stronger line of evidence than the exceedances of low-confidence generic soil screening values, which are widely acknowledged to have low ability to predict toxicity to plant and invertebrate communities. Because community level studies are complex to conduct, the utility of the frequencies and magnitude of HQs exceeding 1 ($HQ > 1$) may be better served for risk management decisions. We suggest EPA plot the locations with plant and soil invertebrate HQs >1 on a map. A spatial analysis would demonstrate co-located risks by stressor (i.e., the COPEC) and the receptor (plants and soil invertebrates), which would aid in focusing risk management decisions.	Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment. To support communication of the risk results, EPA will map key risk results for plants and soil invertebrates in the Risk Characterization.
7	PRP-Integral	Risk Characterization	Risks to wildlife at the site were characterized based on HQs, which were estimated using exposure dose and toxicity reference values (TRVs) as discussed in Section 4.2 of the BERA. For small home range terrestrial receptors (robin, catbird, shrew, vole, and mouse), EPA estimated HQs using TRVs and concentrations from individual sampling locations instead of exposure point concentrations (EPCs) as defined by EPA (1989) as “the arithmetic average of the concentration that is contacted over the exposure period.” It is unclear why EPA used concentrations from individual sampling locations to estimate risk to small home range receptor populations as exposures. As noted by EPA in Section 4.1.4.4, EPA guidance recommends using the 95% upper confidence limit (UCL) on the mean (95UCL) as an estimate for the EPC (EPA 1989, 1992) so that the estimate of the average (or mean) is conservative (i.e., unlikely to be underestimated). In calculating the 95UCL for a contaminant, the underlying statistical theory assumes that each observation in the data set used to calculate the EPC is a random realization of all possible observations from the underlying distribution of chemicals in the subject area of interest. With a sufficient number of random observations, the data set and its statistics can more accurately represent the exposure to populations of ecological receptors, irrespective of the home ranges. For small home range receptors, EPA used the HQs estimated from concentrations individual sampling locations and TRVs to infer effects at the population level. Accordingly, these population-level effect extrapolations include uncertainty associated with the extrapolation between a study endpoint (e.g., number of offspring, reduced litter size) to a population-level effect (e.g., abundance, density, persistence, extinction). Also, any adverse effect on reproduction does not necessarily lead to a decrease in population stability or extinction. Clearly, the reproductive effect must be sufficient in magnitude to	An EPC is intended to conservatively approximate the exposure concentration within an exposure area. Small home range receptors are assumed to have a small exposure area. It is not feasible to collect samples to represent every potential small home range exposure area, so risk is evaluated on a sample-by-sample basis assuming such an approach adequately approximates a small home range receptor exposure area. This was not intended to be a thorough characterization of population-level risks. It provides a simple method for determining if risks may be posed to a large proportion of individuals which implies that population-levels effects may occur. Further, the BERA also calculated total hazard quotients (HQ _t) for small home range receptors in the same manner as was done for the medium home range receptors using the EPCs generated as 95 UCL values for the larger exposure areas.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			result in such a population-level effect. Section 4.2.1.3 of the BERA uses HQs > 1 on a sample by sample basis to interpret population level effects to wildlife. It is not clear what fraction of locations with HQs > 1 EPA uses to determine that population level effects are likely to occur. Because population level studies are complex to conduct, similar to the recommendation for plants and soil invertebrates, the utility of the frequencies and magnitudes of HQs > 1 may be better served as plots on a map to identify locations/localized areas that contribute to potential unacceptable risk to be considered for risk management decisions. A spatial analysis would demonstrate co-located risks by stressor (i.e., the COPEC) and the receptor (including multiple receptors), which would aid in focusing risk management decisions.	EPA will map key risk results in the Risk Characterization.
8	PRP-Integral	Uncertainty Analysis	In general, the uncertainty analysis in a BERA includes qualitative discussions and, in some cases, quantitative evaluations intended to convey the magnitude and direction of uncertainty in the risk estimates. Conservative assumptions are used in the risk models to avoid type II error (i.e., to conclude no unacceptable risks when in fact there is) and consequently there is a high level of confidence that COPECs with a HQ less than 1 do not pose unacceptable risk. Due to the conservative assumptions used, the probability of a type I error (i.e., to conclude unacceptable risk when in fact there is none) based on a HQ> 1 is more likely, and typically requires further evaluation to address uncertainties (Allard et al. 2009). Because there are inherent uncertainties in the exposure and effects assumptions (discussed in the uncertainty analysis of EPA's BERA), the resulting HQs are also uncertain and could over- or underestimate unacceptable risk. Although EPA has addressed several of the uncertainties in the BERA, how they were incorporated into the risk characterization to make ultimate risk conclusions was not made clear (see General Comment on Risk Conclusions on suggested revisions to Section 7).	The uncertainty assessment discusses the primary sources of uncertainty consistent with EPA guidance.
9	PRP-Integral	Data Evaluation	EPA does not provide the rationale for the delineation of samples within the Clark Fork River that are considered representative of the site. As such, it is unclear which samples EPA included in this classification. It would be helpful to have a table that specifies which sampling stations fall in which category cited in Section 3.1 for clarity. Such sample classification should also be made apparent in Attachment A-1. EPA explains that surface water quality criteria for metals require consideration of hardness on p. 19. The text refers to "site-specific" hardness. This language is repeated on p. 20 in explaining the approach to screening for aluminum. EPA should clarify in this section whether it used sample-specific hardness to develop sample-specific criteria values for water. If not, EPA should describe how CaCO₃ and magnesium values used in the equation on p. 19 were calculated to reflect the "site" (i.e., were they means or some other statistical representation of multiple sample results). The hardness of natural surface waters of the river and creeks will likely not be represented by data for ponds and vice versa.	Sample locations in the CFR are shown in Figure 3-2. Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
				Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment. The text will be revised to clarify that hardness values were calculated separately for the CFR, ponds, and on-site creeks (Lavalle and O'Keefe creeks). The hardness-specific water quality criteria used in the screening assessment were calculated using the 5th percentile values for each of the three water body categories (CFR, ponds, on-site creeks). This was considered an adequate approach to maintain the conservativeness of the COPEC screens.
10	PRP-Integral	Data Evaluation	On p. 21, EPA describes screening flood fringe sediments as soil in the screening level ecological risk assessment and as sediments in the BERA. Use of sediment quality guidelines or other screening levels for assessment of risk to flood fringe sediments, which are generally not wetted or submerged, is not appropriate. Flood fringe sediments are deposits on the upper extent of the floodplain. Even when they are in contact with the river, the contact is temporary and does not create conditions for sediment invertebrate communities to develop. We suggest that flood fringe sediments are more accurately understood to be part of the terrestrial ecosystem, and screened using ecological screening values for terrestrial soils.	For the SLERA, available flood fringe samples were defined as not being under water and as such were compared to soil screening benchmarks. Further consideration was given to these samples in the context of the exposure areas identified for the BERA. It was determined that the locations of the flood fringe samples along the CFR are within the exposure areas evaluated for the piscivorous wildlife receptors assumed to forage along the CFR. Thus, the flood fringe data were included in the assessment of these receptors, which was considered an appropriate approach for the purposes of the BERA. The flood fringe data are not being used in the BERA to evaluate toxicity to sediment invertebrate communities.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
11	PRP-Integral	Risk Characterization	EPA concluded that manganese and barium are primary risk drivers in surface water and pore water based on exceedances of surface water screening levels. However, EPA also summarized risk estimates from other metals (arsenic, cadmium, copper, mercury, selenium, and zinc), and it is not clear if these COPECs pose unacceptable risk and require a risk management decision, or risks are uncertain and require further evaluation, or there are sufficient lines of evidence to conclude risks are acceptable. Based on the lines of evidence discussed in the BERA and additional information discussed below, these COPECs should not be considered risk drivers in onsite ponds.	Clarification will be added to the BERA where needed, but this information is also included in the BERA.
12	PRP-Integral	Barium Risk in Ponds	EPA acknowledged the uncertainties associated with the Tier II water quality criteria (Suter and Tsao 1996) used for evaluating barium (Section 6.3.1 of the BERA) and included availability of more recent data and adoption of alternate screening levels by other states, including the Ohio EPA chronic criterion, which is 55 times higher (220 µg/L) than the one used for this site (4 µg/L). In addition, the form of barium commonly present in the environment is barium sulfate, which is not considered very toxic due to low bioavailability (EPA 1993). The toxicity values used by EPA in the BERA are based on more toxic and bioavailable forms, such as barium chloride. Based on these limitations, confidence in the benchmark to predict risk to aquatic organisms is low. More recent studies by Golding et al. (2018) have demonstrated chronic toxic effects of dissolved barium to freshwater aquatic invertebrates ranging from 40 to 240 mg/L. The pond surface concentrations (average of 69.3 µg/L and maximum of 172 µg/L) are below the toxicity values based on more recent studies. Although the maximum barium concentration in pore water (608 µg/L) exceeds the screening level, the average concentration (187 µg/L) is below the toxicity values based on more recent studies. Furthermore, using an updated study of the Great Lakes hardness-based benchmark approach (Great Lakes Basin 2015) results in a site-specific value of 3,182 µg/L for barium (average hardness in the ponds is 646 mg/L). Porewater data can be used to estimate risk to benthic invertebrates. However, using porewater concentrations to evaluate potential risk to aquatic organisms is a conservative approach, as it assumes exposure to organisms is solely occurring at the discharge point/sediment–water interface, without accounting for mixing of pore water with the surface water and impacts to the larger aquatic life community within the ponds. Detected concentrations of barium in pore water may be indicative of an exposure pathway and bioavailability; however, the weight of evidence should rely more on the measured concentrations of barium in surface water, the exposure medium for aquatic life. Based on the rationale provided above, we recommend EPA remove barium as a risk driver in onsite ponds.	The risk benchmark used for barium in the BERA is consistent with the selection hierarchy and the benchmark used at other EPA sites. The uncertainty section will be revised to include a sensitivity analysis presenting HQ results using 220 µg/L.
13	PRP-Integral	Manganese Risk in Ponds	EPA used two lines of evidence to conclude manganese (and barium) is a primary risk driver to benthic organisms in the ponds: HQs estimated for bulk sediment and for pore water. However, other lines of evidence, especially site-specific factors, should be employed to reduce the uncertainty associated with use of generic screening levels and	As concluded in the BERA, manganese is a risk driver to benthic organisms in onsite ponds primarily based on pore water concentrations. Additional

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			bring a higher level of confidence in the conclusions. Other lines of evidence were evaluated in Integral's report (Integral 2020) such as aquatic chemistry, properties, and bioavailability of manganese that are largely controlled by pH and redox potential (CCME 2018). In addition, site-specific observations of the pond ecology were considered in the evaluation of the severity of ecological risks associated with high concentrations of manganese in pore water. Macroinvertebrates were recovered from all locations, primarily as vegetation associated invertebrates, but often including benthic infauna. Robust aquatic invertebrate communities were also observed in several ponds during the July 2018 site reconnaissance visit with EPA. Although sample-specific toxicity thresholds could not be derived for each of the samples, the presence of high calcium concentrations or hardness, can be expected to impart a protective effect on invertebrates (Peters et al. 2011; CCME 2018). Because of uncertainties in the manganese toxicity levels predicted using generic screening levels or model estimates (e.g., BLM-Mn and the CCME models when site-specific samples were outside their calibration ranges), observations of invertebrates provide useful evidence to assess risks to the invertebrate assessment endpoints. Toxicity of manganese to benthic organisms and aquatic life would be reduced in Pond 5, CL-Pond, and IBJ, by high levels of hardness, which have been shown to mitigate manganese toxicity to invertebrates. In addition, manganese in surface water does not exceed screening levels. Pond P5 produced the second highest biomass of chironomids in the 2018 sampling effort. Based on the rationale provided above, we recommend EPA remove manganese as a risk driver in onsite ponds.	information from Integral (2020) may be summarized in the RI, but will not be included in the BERA. Discussion regarding hardness effects on toxicity will be added to the uncertainty assessment.
14	PRP-Integral	Risk Characterization	Site-specific measures of metals bioavailability may be used to further define the potential for sediment toxicity associated with metals. Integral (2020) used EPA's acid-volatile sulfide and simultaneously extracted metals (AVS-SEM) approach (EPA 2005), which accounts for the limited biological availability of the metals in some sediments. From the literature and their own tests, EPA scientists determined that AVS provides the means to quantify the solid-phase sulfides that result in dissolution of iron through a displacement reaction, which leads to the binding of metals other than iron and formation of less soluble sulfides. Those metals bound to AVS are in their solid sulfide form and are not bioavailable to sediment-dwelling organisms. For the Addendum 9 study, the molar concentrations of AVS and SEM were measured in each pond sediment sample and results were used to interpret pond sediment chemistry. As described by EPA (2005), values of $\Sigma(\text{SEM}-\text{AVS})/f_{\text{oc}}$ greater than $130 \mu\text{mol/g}_{\text{oc}}$ indicate that the SEM are likely present in the dissolved phase in sediment pore water (i.e., may be more bioavailable). Values of $\Sigma(\text{SEM}-\text{AVS})/f_{\text{oc}}$ less than $130 \mu\text{mol/g}_{\text{oc}}$ indicate that the SEM in sediment are not likely to be toxic to benthic invertebrates because they are not bioavailable (EPA 2005). EPA's approach to assessment of metals	AVS/SEM data are typically used in conjunction with toxicity tests, but none were performed at this Site. Furthermore, EPA has stated previously that AVS/SEM data will not be considered in the BERA as these data have been shown to be irreproducible among laboratories.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			toxicity in sediments and its application in this report are presented in Appendix B of Integral's 2020 BERA report. Using pond sediment data for several metals and the AVS-SEM approach, all values of $\Sigma(\text{SEM}-\text{AVS})/f_{oc}$ are below $130 \mu\text{mol/g}_{oc}$, indicating that the SEM are not present in pond sediments at concentrations expected to be toxic for cadmium, copper, and zinc. Notably, this line of evidence is consistent with results of screening porewater metals against EPA's surface water screening levels. We suggest including the AVS-SEM evaluation to further define the potential for sediment toxicity associated with metals to provide a more robust conclusion for risk to aquatic invertebrate communities in the ponds.	
15	PRP-Integral	Risk Characterization	Tables 3-8 and 3-9 of the BERA present the steps used to select soil COPECs including identifying chemicals that are bioaccumulative, the frequency of detections greater than 5 percent, availability of an ecological screening values (ESVs), and maximum detected concentrations exceeding the selected ESVs. In addition, EPA completed a statistical comparison of site data to background as discussed in Section 3.3. The results of the background comparison indicated several of the metals were statistically not greater than background including aluminum, arsenic, chromium, manganese, nickel, and vanadium (Table 3-16). However, in Section 3.4.2 and Table 3-16, EPA used no-observed-adverse-effect level (NOAEL)-based HQs > 1 based on maximum detected concentrations to select COPECs. It is not clear why the steps used to screen soil COPECs (Tables 3-8 and 3-9) were not used to screen wildlife COPECs (Table 3-16). Several of the metals listed as wildlife COPECs in EPA's Table 3-18 are not considered bioaccumulative as reported in Tables 3-8 and 3-9 of the BERA and other EPA publications (please see Table 4-2 of EPA 2000a). Thus, related risk conclusions are likely not a reflection of appropriately estimated doses and reflect likely overestimated doses via ingested tissue. We suggest that EPA reconsider the soil screening process for wildlife described in Section 3.4.2 and those metals that are both on the site in soils at concentrations greater than in background and are bioaccumulative be subjected to the screening approach as described in Section 3.3 for selecting soil COPECs.	The reasoning behind the approach taken for screening COPECs for wildlife is described in the BERA. Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment.
16	PRP-Integral	Aluminum as a Wildlife COPEC	Currently a soil ESV equivalent to a soil concentration is not available and EPA guidance (EPA 2003b) recommends a pH screening level of 5.5 to be environmentally protective. EPA reports soil pH data are not available for the site. However, Tables 3-8, 3-9, and 3-16 erroneously indicate exceedance of the aluminum ESV; we suggest EPA correct these tables.	The uncertainty text will be updated to reference Tooke (2006) with regard to the soil pH information to better characterize the uncertainties in risks from aluminum in soil.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			EPA correctly notes in Section 3.1.3 of the BERA that soil pH is an important consideration in evaluating the bioavailability for biological uptake and toxicity of aluminum in soils. Additional information to address the likely characteristics of soils on the site should be considered. Acidification of soil to pH below 5.5 increases the amount of aluminum in solution and depending on the presence of other inorganic molecules (e.g., sulfate, fluoride, phosphate, bicarbonates, or hydroxides), bioavailability and toxicity vary with pH (EPA 2003b). As a result, total aluminum is a poor predictor of ecological toxicity to plants or soil invertebrates. In the absence of site soil pH data, EPA considered aluminum in soils a source of uncertainty and retained it as a wildlife COPEC. However, previous investigations suggest site soil is alkaline (Tooke 2006). Soil pH was measured in surface soil (0–4 in.) at 12 locations over a 12-month period on the site. Average pH over the 12-month period at the various sampling locations on the site ranged from 7.9 to 11.9. EPA’s EcoSSL guidance suggests that aluminum should only be identified as a COPEC in soil at sites where the pH is less than 5.5 (EPA 2003b). In addition to the discussion above on the soil screening process for aluminum (i.e., not considered bioaccumulative and not detected at concentrations greater than background), the alkaline soil pH levels observed on the site indicate that aluminum should not be included as a COPEC for site soil in this assessment. While EPA acknowledges these uncertainties (Section 5.3.1.2), we propose EPA adjust its screening approach and its risk conclusions, removing risk conclusions associated with aluminum exposure, as the current presentation is confusing and places an overemphasis on aluminum, even though it passes all the necessary screening steps.	Aluminum was not retained as a wildlife COPEC solely on the basis of measured concentrations in soil. As shown in Table 3-17, maximum hazard quotient (HQ _{max}) values based on measured aluminum concentrations in dietary items exceeded 1.0. Following the COPEC screening process outlined in the BERA, aluminum is retained as a COPEC in the wildlife assessment.
17	PRP-Integral	Exposure Assessment	EPA discusses uncertainties associated with exposure parameter assumptions used in calculating exposure concentrations/dose. Several of these could have significant impact on the risk conclusions. EPA acknowledges the uncertainties associated with use of literature-based bioaccumulation models or simply assuming concentrations in biota are the same as in an abiotic medium (i.e., assuming uptake is linear; R₂ = 1). Bioaccumulation assumptions represent a large source of uncertainty, especially for terrestrial wildlife at this site, as prey concentrations estimated using literature-derived uptake models/factors do not account for assimilation, metabolism, or depuration of constituents, or site-specific factors that may influence uptake. Site-specific factors may include species-specific characteristics (e.g., feeding strategy, age, and gender) as well as abiotic factors (e.g., soil organic carbon and mineral content, COPEC weathering, climate). The soil at the site may not be representative of soils typically used to derive bioaccumulation factors in the published literature. However, published bioaccumulation models/factors are generally thought to be conservative even though the magnitude of this effect cannot be estimated without site-specific tissue residue data. Calculating risk to wildlife receptors based on uptake models or assumed factors, in addition to other conservative parameters used in the dose estimations, along with	A table summarizing the risk conclusions will be added to the Summary and Conclusions.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			conservative TRVs, would have a compounding conservative effect on HQs. However, EPA does not present a final conclusion about the ecological significance of estimated risk (see General Comment on Risk Conclusions on suggested revisions to Section 7).	
18	PRP-Integral	Exposure Assessment	In the BERA, dietary exposure for wildlife receptors to dioxin toxicity equivalence (TEQ) concentrations was calculated by estimating dioxin TEQ concentrations in their prey tissue. Uptake from soil-to-prey tissue was estimated using an uptake regression for a single congener (2,3,7,8-tetrachlorinated dibenzo-p-dioxin [TCDD]) published by Sample et al. (1998) for soil to earthworms (Table 4-4) to represent uptake of all dioxin/furan congeners included in dioxin TEQ. This is a critical uncertainty for dioxins/furans, as this approach ignores differential uptake of congeners based on differences in their structure and physico-chemical properties leading to potential and significant overestimation of risks to wildlife receptors. Uptake data available in the literature for earthworms (Fagervold et al. 2010) and published soil-to-invertebrate uptake factors (EPA 1999) indicate that 2,3,7,8-TCDD has among the highest uptake rates for the dioxin/furan congeners included in TEQ concentrations. Several dioxin/furan-impacted sites (e.g., Tittabawasee River [Galbraith 2004; Kay et al. 2005], Centredale Manor in Rhode Island [MACTEC 2004]) have demonstrated the overestimation of dioxin TEQ risk resulting from use of a single uptake factor compared with measured concentrations. Because TEQ uptake was estimated using uptake data for only 2,3,7,8-TCDD, TEQ uptake is likely overestimated in the BERA. Congener-specific uptake factors are readily available from EPA (1999) for soil-to-plant uptake and soil-to-terrestrial invertebrate uptake, and congener-specific soil-to-invertebrate uptake factors are also available from Fagervold et al. (2010). These congener-specific uptake factors can be used to model congener-specific tissue concentrations from soil concentrations detected in site soil to demonstrate how the use of the congener-specific uptake factors could impact exposure estimates. To demonstrate the magnitude of this effect, congener specific uptake factors were used to predict individual congener concentrations in diet (invertebrate) tissue, which were then used to calculate tissue TEQ concentrations and estimate HQs. Using the congener-specific uptake factors from EPA (1999), Integral calculated tissue concentrations for dioxin/furan TEQ for each soil sample used to estimate HQs for the shrew at OU3 (see Comment Table 1 [attached to original comment letter]). Integral's calculated tissue TEQ concentrations were compared to the tissue EPA's TEQ concentrations used in the BERA (estimated using the 2,3,7,8-TCDD uptake regressions from Sample et al. 1998; i.e., not congener-specific). As shown in Comment Table 1 (attached to original comment letter), the TEQ tissue concentrations based on congener-specific uptake factors are always less than tissue TEQ concentrations calculated in the BERA except for two instances (Grids 28 and 34), where	The uptake factors for dioxin/furan congeners presented in EPA (1999) represent bioconcentration factors (BCFs) measured in laboratory studies adjusted by a bioaccumulation equivalency factor (BEF). The BEFs used in EPA (1999) are based on sediment and fish residue data from Lake Ontario (EPA 1995) and may not be applicable to soil. Use of the soil-to-invertebrate uptake model from Sample et al. (1998) based on TCDD is considered a conservative approach given the data that suggests that TCDD has among the highest uptake rates for the dioxins/furan congeners. However, in the absence of measured data in soil invertebrates at the Site, this is considered an appropriate approach to ensure that risks are not being underestimated.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			they were similar. Similarly, the HQs for the shrew using congener-specific uptake factors are lower than those estimated in the BERA, up to 10 times in some cases. Most notably, the highest lowest-observed-adverse-effect level (LOAEL)-based HQ estimated (106) in the BERA was at location P17-BH1. Using the congener-specific uptake factors, the LOAEL-based HQ decreases to 10. In addition, using the congener-specific uptake factors, the LOAEL-based HQ for nine additional samples decrease to less than 1. We suggest EPA use the congener-specific uptake factors to estimate risk for TEQ or include discussion of uncertainties and likely overestimation of TEQ risk estimates for wildlife receptors based on use of a single congener uptake factor and influence on the risk conclusions. Quantifying the magnitude of exceedance of HQs using the approach used in the BERA will guide risk management decisions.	
19	PRP-Integral	Toxicity Values	In Section 4.2.7.1 of the BERA, EPA's TRVs for 2,3,7,8-TCDD TEQ for birds combine information from several studies on the toxicity of dioxins and relative sensitivity of various bird species to dioxins using a novel approach. EPA's presentation is missing important details and therefore, the process for deriving avian TRVs is not clear. The avian TRV is based on reproductive effects observed in pheasants in studies conducted by Nosek et al. (1992, 1993). We recommend EPA describe the details and assumptions of the TRV derivation in more detail, including: • We recommend EPA clarify how doses injected to birds reported by Nosek et al. (1992; 1993) were converted to daily ingested dose, and which specific result of Nosek et al.'s work is represented in Table 4-10. • EPA does not provide the reference to the Nosek et al. (1993) reference study. We recommend EPA add this reference to the references section. • EPA only references Nosek et al. 1992 in Section 4.2.7.1, and states that "data suggest that different bird species exhibit varying levels of sensitivity to TCDD toxicity" immediately after citing Nosek et al. 1992. This observation is made in Nosek et al. 1993, not in Nosek et al. 1992. EPA should correct this citation. Further, Table 4-9 references the dioxin TEQ TRV for pheasants as derived from Nosek et al. 1992. We recommend EPA clarify which study the TRV is based on and update references in Section 4.2.7.1 and Table 4-9. • EPA references "the available avian TRV derived from Nosek et al. 1993 (based on pheasants)," and states that it used the approach described below to extrapolate from the pheasant TRV to TRVs for other avian species using relative differences in TCDD or tetrachlorodibenzofuran (TCDF) observed in Cohen-Barnhouse et al. (2011). We recommend EPA provide additional information to clarify to which TRV derived from Nosek et al. (1993) it is referring, as Nosek et al. (1993) does not provide a TRV. • We recommend EPA provide more detail about the extrapolation of the avian TRVs for each species from the TRV for pheasants using data from Farmahin et al. (2013). EPA states that Farmahin et al. determined that sequencing the AhR ligand-binding domain can predict the sensitivity of individual avian species to dioxin-like compounds. Based on this study, EPA grouped the TCDD sensitivity of individual	Text will be revised to clarify that Nosek et al. (1992) is the basis for the Category 2 avian TRVs. Further, additional text discussion will be added to Appendix B and to Section 6 noting the uncertainties associated with the avian TRVs used in the BERA. Injected doses reported by Nosek et al. (1992) in units of µg/kg once per week were converted to daily ingested doses by dividing by 7 (1 week/7 days) and assuming an oral absorption fraction of 100%. In the absence of data, this is an appropriate and health-protective assumption. Following further consideration, ratios will be calculated using the LD50 values for TCDD exposure from Cohen-Barnhouse et al. (2002) study only. This study is the only study that presented LD50 values for all three avian sensitivity categories. Using only this study addresses uncertainties associated with estimating ratios from different types of studies using different methodologies. Using only the TCDD results address uncertainties raised by this comment in using TCDF

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			avian species into three categories. EPA then assigned each surrogate avian species included in this assessment to one of these three categories. We recommend EPA provide more detail about how this extrapolation was conducted, what information from Farmahin et al. (2013) this extrapolation was based on, and how surrogate species used in this assessment were assigned to the TCDD sensitivity categories. • In Table 4-10, EPA presents LC50 values from reference studies without characterizing the uncertainty associated with the derivation of those LC50 values. The uncertainty related to the derivation of the LC50 values should be translated to certainty around the derived TRVs. In addition to the variability in toxicity among individuals of a single species, the uncertainty in extrapolation of toxicity data for one species to derive TRVs for another species needs to be taken into account. We suggest EPA bring this information into the BERA. • In Table 4-10, EPA selected LC50 as its measurement endpoints for effects for calculating sensitivity ratios between categories of species and then applied these ratios to reproductive endpoints for deriving avian TRVs in Table 4-9. EPA assumed that sensitivity differences in mortality and acute toxicity can be equated to reproductive and chronic toxicity. We recommend EPA demonstrate if this assumption is valid or explain the associated uncertainty and how it impacts the risk estimates for birds. It is known that early life stages in vertebrates are the most susceptible to dioxin toxicity. It appears EPA overlooked substantial reporting in Cohen-Barnhouse et al. (2011) on early life stage and developmental effects and egg mortality. • EPA concludes low risk associated with dioxin TEQ to Clark's nutcracker in OU2, OU3 upland, and the OU3 floodplain (Table 7-5) and low risks to the American robin and gray catbird in OU2, the OU3 upland, and the OU3 floodplain (Table 7-6). While EPA does not conclude anything greater than low risks, we recommend EPA thoroughly characterize the uncertainty associated with the assumptions made in deriving the avian dioxin TRVs. It is possible that if this uncertainty and conservative assumptions were characterized, the estimated risk to birds would be de minimis. • As part of the TRV extrapolation process, EPA assumed that TCDD sensitivities of individual species that were not included in the Farmahin et al. (2013) evaluation (blue grouse and northern flicker) would be comparable to that of similar species evaluated in the study (ruffed grouse and common flicker). EPA assumed the individual species evaluated in this BERA would be in the same categories as the species included in Farmahin et al., where possible. Further, for species not included in Farmahin et al. (2013) with no directly comparable species, EPA assumed that these species would fall into the most sensitive category to TCDD. We recommend EPA provide rationale supporting this assumption, and/or explicitly include this assumption as a source of uncertainty in this assessment. • It is not clear how molar concentrations reported by Cohen-Barnhouse et al. (2011) were converted to mass concentrations necessary to derive an ingestion-based TRV and which specific results in this paper were used in Table 4-10. This is a potentially	results. The molecular weights were used to convert from the molar concentrations reported in the Cohen-Barnhouse study to mass concentrations: $\text{Mol/kg} * \text{g/Mol} * 1000 \text{ mg/g} = \text{mg/kg}$ This conversion is not a source of uncertainty. The resulting NOAEL and LOAEL TRV values are slightly higher for both Category 1 and Category 3 birds. However, the revised Category 1 NOAEL and LOAEL TRV values are supported by additional studies evaluating effects in chickens; discussion of these supporting data will be added to the BERA. As described in the BERA, the surrogate species evaluated in the risk assessment were assigned to category 1, 2, or 3 based on the species categorizations in the Farmahin et al. (2013) study (Table 1 subtypes 1, 2 and 3 correspond to the three sensitivity categories evaluated in the BERA). The assignments made for bird species not included in the Farmahin study are considered adequate and health protective in the absence of additional information as noted in the BERA. The uncertainty section will be revised to include discussion on the uncertainties associated with the assignment of avian sensitivity categories for dioxins/furans. The Nosek et al. (1992) study incubated fertilized eggs collected from hens exposed to TCDD by ip injection, and Cohen-Barnhouse (2011) incubated fertilized eggs following direct

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			large source of uncertainty and it should be recognized and discussed by EPA. We suggest EPA account for this when making risk management decisions related to the avian dioxin assessment. It is not clear why results for TCDF are referenced by EPA in Section 4.2.7.1 and shown in Table 4-10, when EPA's measure of exposure is TCDD TEQ values. We recommend EPA provide additional context supporting the inclusion of TCDF in these components of the report or compare all results as TCDD concentrations or as calculated TEQs.	injection of TCDD into the egg. Egg mortality was evaluated in both studies. Cohen-Barnhouse et al. (2011) notes that embryo mortality has been shown to be related to dose within species and can be used to evaluate species sensitivity. Using ratios (as was done in the BERA) further allows for evaluation of the relative differences in sensitivities among different bird species.
20	PRP-Integral	Toxicity Values	We found several subjects related to risk assessment for mercury that EPA should consider in preparing the final BERA: - EPA's selected inorganic mercury TRV does not follow the hierarchy of TRV selection described in Appendix B. Inorganic mercury TRVs for mammals are available from ORNL (Sample et al. 1996); however, an alternative study was selected, Khan et al. (2004) from primary literature, identified in Table 4-9 although not cited in the references or appendix. While there are in some cases more recent and relevant toxicity studies for mercury and wildlife, there is no discussion in the report about this deviation from the TRV hierarchy. We recommend EPA provide additional discussion of the derivation of this TRV. - As previously stated, more recent and relevant mercury toxicity studies are available in the primary literature; this is particularly true for methylmercury toxicity for birds. A significant amount of more recent literature is available since the publication by Heinz (1979) on the effects of methylmercury on avian reproduction. A more recent review of relevant studies was conducted by Fuchsman et al. (2016), which includes review of dietary toxicity studies from which an avian TRV may be derived. We recommend EPA consider the breadth of recent toxicological data relevant to avian mercury exposures and toxicity. - EPA's critical tissue residue (CTR) for fish is not clear. No information is provided in Appendix B on the derivation of the CTR for mercury in fish. We suggest EPA provide a description on the derivation of the fish CTR values.	Text will be added to Appendix B regarding the derivation of mercury TRVs and the CTR for mercury in fish.
21	PRP-Integral	Toxicity Values	It would be useful to provide a discussion of the use of uncertainty factors in deriving toxicity values. The following are two specific examples of why additional discussion of uncertainty factors is needed. The main text of the report indicates the relevant endpoints for evaluating effects to ecological receptors are growth, reproduction, and survival. It is unclear then, why in the case of surface soil toxicity values, that an additional uncertainty factor is being applied to estimate a toxicity value for reproduction/growth from a study that reports a toxicity value based on survival.	The use of uncertainty factors in deriving TRVs is an EPA-accepted approach. The application of UFs described in the derivation of the wildlife TRVs is consistent with the approach being used for other regional EPA risk assessments.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			While the use of an uncertainty factor of 10 between NOAEL and LOAEL values is commonly used, there are more recent recommendations that support chemical-specific uncertainty factor estimates other than 10.	
22	PRP-Integral	Toxicity Values	Additional detail on study selection criteria is recommended. This is particularly true for methylmercury TRVs for birds and CTRs for fish, where a significant amount of more recent literature is available.	Text will be added to Appendix B regarding the derivation of the CTR for mercury in fish.
23	PRP-Integral	Toxicity Values	We suggest adding additional details to discuss the methods used in selecting the final toxicity values (dose response modeling, geometric means, etc.). In some cases, a toxicity value is selected using a geometric mean of several studies and in other cases, the lowest value is selected. If data are available and of sufficient quality (as may be the case with more recent toxicological data on mercury in birds and fish), it may also be possible to derive toxicity values from dose-response relationship instead of from point-estimate values such as no-effect and low-effect values.	The selection of no-effect and low-effect values is consistent with EPA-approved approaches for deriving toxicity values.
24	PRP-Integral	Toxicity Values	While threshold effect concentrations (TECs) and probable effect concentrations (PECs) are commonly used in ecological risk assessments, it would be useful to include a discussion of the uncertainties in these values. Although the development of TECs and PECs involved the consideration of other previously derived sediment screening levels, TECs and PECs are based on associations observed between measures of adverse biological effects and the concentrations of potential COCs in sediments (Long and Morgan 2001). TEC and PEC values are derived for individual chemicals, yet the measured effects in invertebrates are based on sites where multiple contaminants are present. The consensus-based approach derives unified threshold and probable values by calculating the geometric mean of several published guidelines; however, the limitations posed by the individual guidelines are not necessarily eliminated in the process. The guidelines included in the consensus-based approach are based on empirical approaches that derive their respective levels from large databases where sediment chemistry data are paired with biological effects and do not take into account cause-and-effect relationships between a single chemical or a mixture of chemicals and the adverse biological effects on organisms. Therefore, the adequacy of these approaches is directly related to the size and variability of these databases. This results in an overestimation of the potential risk to benthic invertebrates, especially when conclusions of risk are based solely on sediment chemistry concentrations (in the case of organics).	The consensus-based sediment benchmark values are limited by the database that was evaluated by McDonald et al. (2000). However, the selected values are commonly used in ecological risk assessments. Use of these benchmark values is not considered a significant source of uncertainty, but additional discussion regarding the uncertainties in the sediment benchmarks will be added to Section 6.3.
25	PRP-Integral	Toxicity Values	Information in Appendix B on the CTR in fish is not sufficient for independent verification. Additional details are required.	Information regarding the dietary exposure studies with methylmercury in five species of fish used to derive the NOEC and LOEC critical fish tissue concentrations for mercury will be added to Appendix B.
26	PRP-Integral	Editorial	We suggest that EPA adopt notation that allows the reader to understand a HQ as based on a NOAEL as HQ _N , or based on a LOAEL as HQ _L throughout the document. The	NOAEL- and LOAEL-based HQs are noted in the text, and to improve

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			distinction is important and the current lack of specificity makes results difficult to interpret. The first sentence in the last paragraph on p. ES-4 illustrates the problem.	clarity, the BERA will be updated with this notation change.
27	PRP-Integral	Editorial	On p. ES-2, EPA states that the “management goal at the Site is to ensure adequate protection of ecological receptors that may utilize” the site. We ask that the following be added to the end of the sentence: “...within a mixed-use development for human use.” Because the statement does not acknowledge that the future uses of the site include human uses, it is incomplete as written.	The sentence in question will remain as is in the BERA. There is no set zoning plan, and a mixed-use development for human use is not a certainty at the Site. As such, it may remain open space and ecological habitat.
28	PRP-Integral	Editorial	On p. 9, EPA states that “areas in OU2 including the former bleach plant and chemical recovery systems are potential sources for polychlorinated biphenyls (PCBs).” This is incorrect and should be deleted.	This statement will be deleted.
29	PRP-Integral	Editorial	On p. 10, EPA states that “the presence of other chemical pollutants in contaminated soils may dissolve dioxins, making it easier for dioxins to move through soil.” This statement needs to be deleted; if true in a general sense, the statement should not be applied to any specific site without describing which chemicals are believed to be mobilizing dioxins and the evidence that there is mobilization at the site. There is no evidence at this site that chemicals that could “dissolve dioxins” are present.	This statement will be deleted.
30	PRP-Integral	Editorial	On p. 12, the BERA states: “...oral toxicity factors for fish are not generally available.” There are TRVs expressed as concentrations in the foods of fish for many metals.	EPA is aware of studies that evaluate toxicity of metals to fish and present exposure concentrations. However, EPA has not currently identified readily available and appropriate sources for exposure factors for fish. A food ingestion rate is needed to convert the exposure concentration to a dose.
31	PRP-Integral	Editorial	On p. 38, in the formula for estimating whole body fish tissue concentrations, it should be specified whether fillet weight (W_f) represents the mean of the fillet weights within a composite or the cumulative weight of all fillet in the composite, or something else. W_c should be clarified as well.	W_f (fillet weight) is the average weight (in grams) across fish fillets included in a composite sample. W_c (carcass weight) is the average weight (in grams) across fish carcasses (does not include fillets) included in a composite sample. This detail will be added to Section 4.1.2.2 of the BERA.
32	PRP-Integral	Editorial	The presentation of media-specific doses (e.g., for bat and fox in Appendix H) are presented in a manner that is difficult to follow. We propose EPA clarify that the media-specific doses need to be multiplied by the area use factor to yield the total dose for the receptor presented in these tables.	The equations detailing the dose calculations are in Section 4.1.4 of the BERA and include multiplication by the area use factor.
33	CFC- Abt	Exposure Assessment	EPA has attempted to assess ecological risk using a paucity of data and many modeling substitutions and assumptions to address data gaps. Many of the sediment and tissue samples are composite samples, and those composites at times are further aggregated to create a single estimate of dose or exposure over hundreds of acres of waste ponds. Some	In the absence of empirical data, it is considered acceptable to use established uptake models for estimating food web exposures. As

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			data gaps were not addressed at all, which led to methods that strain credulity, such as estimating risk to bats based on contaminant concentrations in earthworms. Other assessments were done with convoluted calculations and assumptions, such as comparing dioxin data from composite rainbow trout fillets and carcasses to tissue residue benchmarks (TRBs) from 30-year-old studies of dioxins injected into fish eggs. The lack of sufficient data and the attempts to estimate risk based on assumptions and models instead of actual data do not allow for a reasonable conclusion that this site poses little or no risk to biota. However, as will be discussed in the comments below, the existing data show that the site is a source of dioxins and dioxin-like compounds, there is a potential pathway for the compounds to reach the Clark Fork River, and the concentrations of those compounds in rainbow trout are higher downstream of the site than they are upstream. Concluding that the site is not a source of contaminants that pose a risk to biota in the Clark Fork River is potentially premature.	described in the uncertainty assessment, the EPA models used in the BERA are considered to conservatively estimate exposures whereby risks are likely overestimated. Composite sampling reduces data variability and provides a reasonably unbiased estimate of mean contaminant concentrations in a specified decision unit (ITRC 2021).
34	CFC- Abt	Data Evaluation	BERA Section 2.1, p. 3: <i>“Most of the pulp was used to produce unbleached linerboard, but a small fraction (about 6 percent) of the total pulp produced from 1960-1999 was used to create white linerboard or sold as bleached pulp.”</i> EPA and site contractors discount the amount of bleached pulp (a source of dioxins) produced at this site. If this “small fraction” of overall production was 6%, approximately 1,600,000 tons of bleached pulp were produced. This is a substantial quantity of a product that is a known source of dioxins and furans. Chlorinated dioxins and furans, including 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and 2,3,7,8-tetrachlorodibenzofuran (TCDF), are persistent organic pollutants (POPs) that do not readily degrade in the environment and are highly toxic. In the late 1980s, EPA and the paper industry conducted a study incorporating sampling data from all 104 bleaching kraft paper mills in the United States (U.S. EPA, 1990), including the Frenchtown mill. In 1988, when the mill was sampled for the 104-Mill Study, TCDD was found in pulp, sludge, and effluent; and TCDF was found in pulp and sludge at the mill. The effluent discharged to the Clark Fork River and the sludges were placed in unlined waste ponds. In the problem formulation, EPA could have prioritized dioxins and furans as likely drivers of ecological risk, with more sensitive sampling methods and a different sampling design focused on assessing these contaminants.	As noted in the recently released Air Deposition Fact Sheet (EPA, 2021a), OU1 and OU2 surface soil samples reported dioxin levels below MDEQ’s background level. Available data (MDEQ, 2011) indicate that dioxins and furans in Missoula soils are typical of rural areas elsewhere in the U.S., and well below urban areas (Urban et al., 2013). MDEQ (1995) concluded that burning dewatered primary sludge from the wastewater treatment system would not result in increased dioxin emissions. Finally, available dioxin/furan data were evaluated in the BERA and were not found to result in unacceptable risk to ecological receptors.
35	CFC- Abt	Site Characterization	BERA Section 2.1, p. 4: <i>“OU3 includes locations in the CFR where hazardous substances from Site activities have come to be located, if any... For the purposes of this assessment, sampling conducted to date within the CFR adjacent to the Site and downstream of the Site is considered to represent potential areas where Site-related contamination may have deposited.”</i> The 104-Mill Study report (U.S. EPA, 1990) confirmed that dioxins were in the effluent in 1988. The mill released dioxins and furans into the Clark Fork River. These POPs may have come to be located in sediment deposition areas well downstream of the site, and	The study area evaluated in the BERA was based on input from EPA hydrologists to identify an area within the CFR downstream of the site where sedimentation was likely to occur to be representative of where site-related contamination may have come to be located. The full nature and extent of

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			they may still be biologically available at concentrations that harm biota. EPA has not assessed potential ecological risk in the Clark Fork River beyond the immediate vicinity of the site. POPs released when the mill was operating may have come to be located in depositional areas well downstream of the site. At other contaminated sites, EPA requires the responsible parties to identify potential ecological risk downstream of the original release site. For example, the Clark Fork River Superfund complex upstream of Missoula extends from the original source areas to Milltown Dam, where contaminants from Butte and Anaconda came to be located, in a sediment deposition area over 100 miles downstream. Most of the contamination behind the dam was likely released decades earlier, but was still biologically available and required remediation. At the Frenchtown Mill site, EPA has not assessed downstream depositional areas where dioxin/furan releases from the mill may have come to be located.	contamination within the CFR is being evaluated within the RI. EPA is not planning to do additional sampling for the purposes of risk assessment of the current study area at this time.
36	CFC- Abt	PCBs	BERA Section 2.4.2, p. 9: “PCBs were only detected in OU2 soils at specific locations, the high-density pulp tank (HDPT) and the transformer storage building (TSB) foundation areas. Aroclor 1260 was detected in the HDPT area, and Aroclors 1254 and 1260 were detected in the TSB area.” The conceptual model for polychlorinated biphenyls (PCBs) focused on Aroclors, such as those in oil leaking from transformers. Coplanar (dioxin-like) PCBs can also be generated from incomplete combustion in the presence of chlorine, similar to dioxins (Lemieux et al., 1999). These dioxin-like congeners have a similar mode of toxicity as dioxins and furans, and therefore are often included in a calculation of TCDD toxicity equivalence (TEQ). In addition, the detection limits for analyses of PCB congeners are orders of magnitude lower than the detection limits for PCBs in Aroclors. The only samples of any media from the site that included PCB congeners were from the 2017 OU2 supplemental soil sampling. This study included PCB congeners from nine composite samples, where each composite covered from 10 to 20 acres. Coplanar PCBs, including PCB-126 (the most toxic), were detected in all nine samples. In four of the samples, PCB congeners comprised more than 50% of the TCDD TEQ, and in two samples, PCBs comprised nearly 80% of the TCDD TEQ. The data suggest that the site is a potential source of PCBs other than just Aroclors, and these coplanar PCBs should have been evaluated in other locations and other media.	PCBs were not utilized in the main mill operations, but were included in the initial site investigation studies. Aroclors were only detected in soils at a level of concern in isolated locations on the site where transformers were previously stored and in the high-density pulp tank foundation. Removal activities have already occurred, and ongoing groundwater monitoring has continued to evaluate for the presence of PCBs. EPA recently published a PCB fact sheet (EPA, 2021c) that provides additional information and will be cited in the BERA. PCBs were evaluated as a COPEC in the BERA for soil receptors and for wildlife, and were not found to result in significant risk to ecological receptors.
37	CFC- Abt	Dioxins/Furans	BERA Section 2.4.3, Fate and Transport, p. 10: “Soil erosion and surface runoff can transport dioxins into surface water. Dioxins do not easily dissolve in water and therefore, most of the dioxins that enter surface water become strongly attached to particles that eventually settle in the sediment.” Despite the hydrophobic properties of dioxins and furans, several studies have found that these chemicals can be mobilized, most likely via colloidal transport (e.g., Hofmann and Wendelborn, 2007). The BERA does not fully evaluate a potential groundwater pathway to the Clark Fork River, despite detectable dioxins in groundwater samples. Figure 1 (in	The BERA CSM is intended to identify significant and complete exposure pathways to potential ecological receptors. This includes evaluation of potential exposures within the CFR based on the transport of contamination from the site into the CFR. The groundwater CSM (NewFields 2020)

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			the original comment letter) shows shallow groundwater wells with elevated TCDD TEQ, calculated with non-detectable (ND) dioxin and furan congeners set to a value of 0 pg/L. The wells with red circles have a maximum TCDD TEQ > 2 pg/L (which exceeds the Montana Circular DEQ-7 groundwater standard). Wells with a maximum TEQ > 1 pg/L have yellow circles, and wells with a maximum TEQ > 0.5 pg/L have orange circles. The spatial pattern suggests a groundwater pathway that extends from waste disposal sites in OU2 to the Clark Fork River. The TEQ calculations in Figure 1 (in the original comment letter) include only detectable dioxin and furan congeners. To date, there has been no assessment of coplanar PCBs in groundwater. As noted elsewhere, the site may be a source of coplanar PCBs in addition to dioxins and furans. Dioxins and furans can be toxic at extremely minute concentrations and difficult to assess in water using standard methods because of their hydrophobicity. Groundwater standards are generally based on practical quantitation limits (PQLs) rather than low or no adverse effects levels. There is no TCDD TEQ surface water standard for the protection of aquatic life, but it is likely that it would be lower than the detection limits in the water samples collected at the site. For comparison, the surface water aquatic life criterion for TCDD TEQ based on risk to human consumption of fish and water is 0.005 pg/L, which is two orders of magnitude lower than typically achieved using standard methods. As noted in Figure 1 (in the original comment letter), even with standard sampling methods, the data suggest possible groundwater transport of dioxins/furans to the Clark Fork River, although the data are variable. At one site, a shallow groundwater sample contained 2.4 pg/L TCDD and a duplicate sample had no detectable TCDD at a detection level of 0.18 pg/L. This highlights the difficulty of capturing highly hydrophobic POPs using these standard sampling methods. To reduce this variability and greatly reduce the detection limit, specialized sampling methods have been developed to assess highly toxic hydrophobic chemicals such as dioxins. The alternate methods generally require a large volume of potentially contaminated water to flow-through materials that attract lipophilic compounds, allowing those compounds to become concentrated to levels that are easier to detect. This can be accomplished by actively pumping large volumes of water through a device such as a resin column that will attract the lipophilic compounds, or passively placing a device in flowing water and leaving it for a prolonged period of time. These methods have been used for many years at other sites with known dioxin issues. For example, Dinkins and Heath (1998) used high-volume water sampling to achieve detection limits below 0.001 pg/L in a dioxin-contaminated river in Ohio. For each water sample, 1,000 liters of water were pumped through a resin column. Charlestra et al. (2008) deployed semipermeable membrane devices into the Androscoggin River in Maine for 36 days to evaluate dioxin releases from a paper mill. Although there is more uncertainty in	being developed separately from the risk assessments will be cited in the BERA. The OU3 CSM includes an arrow showing potential transport from groundwater to the CFR. Groundwater discharge to the CFR is captured in surface water and sediment sampling. TEQ was evaluated as a COPEC in surface water, sediment, and soil in the BERA and was not determined to be a major risk driver to ecological receptors. Additional data collection to support the risk assessments is not planned at this time, but EPA will consider these recommendations for any potential sampling in the future to support the FS.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			the estimate of the volume of water sampled using this method, Charlestra et al. (2008) also achieved detection limits close to 0.001 pg/L. Both the Department of Defense (Lohmann, 2016) and U.S. EPA (2012) have developed guidance for assessing organic contaminants using passive sampling. The techniques have been used at numerous Superfund sites. Burgess et al. (2015) deployed three different passive sampling devices at three different Superfund sites, concluding that the devices generated scientifically accurate data that were critical for making informed management decisions at contaminated sites. Passive sampling can also be used to assess the transport of hydrophobic compounds in groundwater. The Interstate Technology & Regulatory Council (ITRC, 2007) published protocols for using five different passive samplers in groundwater wells. The U.S. Geological Survey recently published additional guidance for assessing groundwater chemistry using passive sampling methods (Imbrigiotta and Harte, 2020). Despite the hydrophobic nature of dioxin-like compounds, the existing data suggest a potential pathway to biota that has not been entirely evaluated. Prior to concluding that this site poses no risk from releases of POPs, EPA should consider using these existing technologies to assess potential offsite transport of highly toxic hydrophobic chemicals.	
38	CFC- Abt	Data Evaluation	EPA evaluated dioxins, furans, and coplanar PCBs in northern pike and rainbow trout tissues, and dioxins and furans in longnose dace. We believe the data are inconclusive. The 2018 northern pike and rainbow trout data collection had data quality issues that relegated the data to the discussion of uncertainty in the BERA. The 2019 rainbow trout data included three sample sites upstream of the mill, one site in Frenchtown, and only one site downstream of the site at St. Regis. Northern pike were collected at two sites upstream of the site and in Frenchtown; no pike were collected downstream in St. Regis. This is a very limited dataset. Dioxins may have been transported upstream from the site via stack emissions or biota moving upstream, but the vast majority of dioxins released to the Clark Fork River would have been transported downstream. This BERA does not include data to assess potential exposure and risk to POPs that have come to be located in downstream depositional areas. Longnose dace were collected at several locations adjacent to and immediately downstream of the site. This spatial coverage likely does not cover the entirety of areas where dioxins have come to be located. In addition, coplanar PCBs were not analyzed in longnose dace. This is a data gap and likely underestimates TCDD TEQ in longnose dace.	Fish tissue data were collected and analyzed for dioxins, furans and PCBs in separate sampling events, as noted. Efforts were made to collect fish substantially upstream and downstream of the site, and data collection efforts were consistent with DQOs. Additional data collection efforts are not planned at this time. The uncertainty assessment will be expanded to include discussion of uncertainties associated with the lack of PCB data in longnose dace. Uncertainties associated with the available database are described in the uncertainty assessment.
39	CFC- Abt	Data Evaluation	In the BERA, TCDD TEQ is calculated using one-half of the detection limit of undetected congeners. This is a standard practice and would typically be considered “conservative” (i.e., potentially overestimating risk) because it assumes that all congeners are actually present at concentrations below the detection limit, even if they are not. However, when many of the congeners are undetected and data quality issues result in inconsistent and at times elevated detection limits, this method can cause samples with almost no dioxins to appear to have TEQs that are as high as or higher than samples with lower detection limits	EPA evaluates non-detects assuming ½ the sample-specific detection limit which accounts for sample-specific volume and matrix interferences. The same approach was used to evaluate exposures within the study area and upstream of the site. The summary

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			and many detected congeners. The substitution of a single value (such as half the detection limit) for censored data can introduce false or misleading trends (Helsel, 2005). At this site, calculating TEQs using half the detection limit for undetected congeners can lead to artificially high TEQ calculations in samples from upstream of the site, which could mask a trend of increasing TEQ downstream of the site. We have evaluated TEQs with undetected congeners set to a value of 0 rather than half the detection limit in our analyses presented in these comments.	statistics tables in the BERA present TEQ calculations using non-detects equal to one-half the detection limit, and non-detects equal to zero.
40	CFC- Abt	Exposure Assessment	The method used to assess risk to higher trophic-level biota from exposure to TCDD TEQ required many calculations and assumptions, with little underlying data to support any conclusions. Many of the samples collected at this site were composited, where multiple samples were mixed together, which implicitly assumes that concentrations are uniform in the individual samples, or that organisms will be exposed to different concentrations across a home range and the composite sample provides a better point estimate of exposure. Large areas of the site are represented by a single data point, which may not be representative of the potential exposure. All of the data from 11 ponds in OU3 covering hundreds of acres were digested down to a single exposure area using one assumed value for each contaminant. The existing data may not capture the areas of contamination and potential risk when such broad areas are represented by composited samples and aggregated data. To assess risk to osprey from dioxins, the BERA relies on longnose dace data from adjacent to and immediately downstream of the site, and rainbow trout data from Frenchtown and St. Regis. All of the fish data were aggregated to create a single estimate of ingested dose. The data are sparse, and each step of aggregation increases the uncertainty. Rather than relying on these calculations and literature values, EPA could assess risk to osprey by collecting data from Clark Fork River ospreys.	Composite sampling reduces data variability and provides a reasonably unbiased estimate of mean contaminant concentrations in a specified decision unit (ITRC 2021). The BERA identified the belted kingfisher as a surrogate receptor for assessing risk to piscivorous birds that use the site. Although belted kingfisher is a more sensitive receptor than osprey due to their relatively higher food ingestion rate, EPA also evaluated exposures to osprey. Exposure point concentrations (EPCs) were calculated as the upper 95th percentile confidence limit (95UCL) to account for uncertainties in estimating the mean concentration from a limited number of measurements. It is considered acceptable to use established uptake models for estimating food web exposures. As described in the uncertainty assessment, the models used in the BERA are considered to conservatively estimate exposures whereby risks are likely overestimated.
41	CFC- Abt	Toxicity Values	To assess risk to fish, the BERA relies on TRBs from Steevens et al. (2005). Much of the underlying data in Steevens et al. (2005) are 30-year-old studies where adverse effects were measured at different concentrations of TCDD in fish eggs, including multiple studies where TCDD was injected into the eggs, and studies where reported exposure concentrations were nominal (i.e., estimated based on dilutions of a stock solution) rather than measured. In the BERA, TCDD TEQ concentrations in rainbow trout and northern pike were calculated from composite fillet and carcass samples of adult fish (one sample	EPA has identified Steevens et al. (2005) as an appropriate basis for fish tissue benchmarks to use in the BERA. Russell et al. (1999) demonstrated that lipid mobilization has little effect on the relationship between maternal tissue and egg concentrations of

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			site for northern pike, and two for rainbow trout), and compared to benchmarks developed to show harm in eggs. This requires assuming that benchmarks in eggs and adult fish tissues are the same when normalized to lipid weight. Lipid concentrations for rainbow trout (fillet with skin on) collected for the BERA range from 5.7% to 43.5%; 5 of the 20 reported values exceed 35% lipid. This is considerably higher than literature ranges of lipid values reported for multiple salmonids. We examined data from several studies (Niimi, 1983; Miller, 1993; Russell et al., 1999; Kandemir and Polat, 2007) and found lipids ranged from 0.7% to 26%, with an average of 8.1% lipid (in whole fish, skin-on fillet, and fillet). Eighteen of the 20 Clark Fork River rainbow trout fillet lipid values exceed 8.1%, which could confound the assumption that the TCDD TEQ concentrations in these fish can be compared to literature TRBs. Furthermore, the Steevens et al. (2005) benchmarks are based on tissue residue effects levels in fish eggs, which are converted to “whole-fish” by Steevens et al. (2005) based on the 1:1 lipid-normalized chlorinated organic compound relationship reported by Russell et al. (1999). Though Steevens et al. (2005) describes these as adult “whole fish” concentrations, Russell et al. (1999) used adult dorsal muscle in their analyses, and they cited studies where “whole fish” included fish with eggs and gut contents removed (Niimi, 1983) or just skin-on fillets (Miller, 1993). Therefore, comparisons of lipid-normalized TEQ values derived from fish carcass samples may not represent a direct comparison to the Russell et al. (1999) analysis and subsequent benchmarks derived by Steevens et al. (2005). Additional risk thresholds should be evaluated and additional data collected to achieve a lower level of uncertainty. There are other issues with the calculation of risks to fish in the BERA. Without belaboring the point, the assumptions and calculations required to compare the trout and pike data to fish egg residue benchmarks creates a level of uncertainty too high to allow for any conclusions about potential risk.	contaminants in fish. The uncertainty section will be expanded to include a specific discussion of the uncertainties associated with the fish tissue-based TRVs. In addition, to address the uncertainty associated with Russell et al. (1999) using adult dorsal muscle in their analysis, the uncertainty analysis will be revised to include an evaluation of lipid-weight adjusted fillet concentrations similar to the evaluation of the whole-body concentrations presented in the risk characterization.
42	CFC- Abt	Site Characterization	We have not attempted to calculate risk to fish, but we have examined the spatial trend in fish tissues, using simpler analysis methods. For example, we examined average TCDD TEQ in rainbow trout (“whole body,” using the method of adding fillet and carcass as described in the BERA) at each 2019 sampling location (Figure 2 [in the original comment letter]). While the data show potential sources of dioxins/furans and coplanar PCBs upstream of the site, the highest concentrations of both dioxins/furans and coplanar PCBs were in fish collected at St. Regis, downstream of the site. As noted in a previous comment, coplanar PCBs can be a byproduct of combustion, and they are present in OU2 soils in the only samples collected at the site that included PCB congeners. Given that the site is a known source of these POPs, and the concentrations are increasing downstream, additional data should be collected to properly assess potential risk from both past and ongoing releases into the Clark Fork River.	Source attribution will be conducted as part of the RI and is not included in the BERA. Further, using biotic data to evaluate source attribution is highly uncertain given the many variables that influence tissue concentrations. Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
				should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment. See EPA's recently published PCB fact sheet for more information (EPA, 2021c).
43	CFC- Abt	Risk Characterization	The assessment of risks to fish is based on a very limited dataset, with many assumptions and calculations required to compare the existing data to TRBs derived from fish eggs in Steevens et al. (2005). The existing data for rainbow trout suggest that TCDD TEQ increases downstream of the site. Additional research is required to properly assess exposure to fish and piscivores downstream.	Additional sampling is not planned prior to finalizing the BERA. Uncertainties regarding fish tissue data and toxicity values are discussed in the BERA.
44	CFC	Aerial Deposition	Much of the area upstream of the Smurfit site was subject to decades of aerial deposition from emissions at the mill site, and EPA's own research shows Kraft liner board facilities were common sources of airborne emissions of dioxins and furans. For these reasons, we ask that all reference conditions used in the ongoing investigations be located outside of Smurfit's aerial deposition zone.	Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
				to better characterize the results presented in the risk assessment. See EPA's recently published aerial deposition fact sheet (EPA, 2021a) and MDEQ's state-wide dioxin background investigation report (MDEQ 2011) for more information.
45	CFC	Flooding Concerns	High flows during the spring of 2018 provided a stark reminder that berm failure at Smurfit is an inevitability given enough time. No assessment of risks to human health and the environment at Smurfit can be complete without consideration of the scour and transport of toxic material from the site that would occur during a high flow berm failure event. We request that berm failure be added to the analysis in the BERA and HHRA.	Risk assessment does not/cannot account for potential future catastrophic events such as berm failure. However, the PRPs, with EPA and MDEQ oversight, conducted a berm stability study to address these concerns. This study will be cited in the BERA.
46	MDEQ	Toxicity Values	Please be aware that any compounds present above DEQ's Water Quality Standards for human health and aquatic life will need to be addressed when making risk management decisions in the future for both surface water or groundwater.	This will be discussed in the uncertainty section of the BERA.
47	MDEQ	COPEC Screen	DEQ has concerns about the process to determine COPCs for the following reasons: no consideration of risk-based water quality standards (federal or state), the use of a 5% rule with no analysis for hot spots, the statistical analyses of background and site data to determine COPSs. In some cases, there does not appear to be sufficient data to ensure we are not prematurely screening out COPCs (e.g. surface water PAHs with DLs greater than screening levels). Please clearly describe all locations where background and site data are compared and for each location, reference a table where the data is presented (including background concentrations used with a reference to the source), along with the statistical analyses.	The approach used to determine COPECs in the BERA that does not evaluate hot spots is consistent with RAGS (EPA, 1989). Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
				to better characterize the results presented in the risk assessment.
48	MDEQ	Background Evaluation	Appendix A, Section 2.1.2.2: DEQ does not agree with the use of the bulk fraction data for the background evaluation. Fine particles are known to be readily transported, ingested, or inhaled and typically are the fraction of soil inorganics readily sorb to as well. Therefore, bulk concentrations may not accurately capture exposure concentrations. Please conduct a background evaluation using the fine fraction soil.	Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment.
49	MDEQ	Editorial	In the Introduction of the Executive Summary (page ES-1), please make the following edit: <i>“...This information, along with other relevant information, will be used by risk managers to decide whether remedial actions are needed to protect ecological receptors from adverse effects attributable to Site-related environmental contaminants.”</i> This parenthetical should be deleted as there has as yet been no RA implemented at the Facility.	Soils have been removed from the high-density pulp tank (HPDT) and transformer storage building (TSB) based on PCBs. As such, no change will be made to this sentence.
50	MDEQ	Editorial	The second paragraph of Section 2.2 (Summary of Site Investigations) ends with the following sentence: <i>“The EPA 2018 and 2019 fish data were not uploaded into the Scribe database as of the writing of this report, but were available in electronic format on the Site webpage.”</i> Has this changed?	The 2018 and 2019 fish tissue data have been uploaded to the Scribe database. This text will be deleted.
51	MDEQ	Site Characterization	In Section 2.3.2.2 (O’Keefe and Lavalley Creeks), the report states, <i>“O’Keefe Creek is a ditch for much of its length.”</i> Please clarify what is meant by, or what factors were considered, in characterizing this stretch as a “ditch.” The state waters within O’Keefe and Lavalley Creeks would still be considered a B-1 classification surface water.	This was meant as a qualitative descriptor and not a strict surface water classification. This sentence will be deleted.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
52	MDEQ	Exposure Assessment	Section 2.4.4 lists the major exposure routes considered in the BERA. DEQ suggests including the ingestion of soil by soil invertebrates and wildlife receptors exposure route to the list below as this exposure route was considered in the CSM.	Incidental ingestion of soil by wildlife receptors is currently included in the list in Section 2.4.4. Direct contact of plants and soil invertebrates with soil will be added to the bullet list.
53	MDEQ	COPEC Screen	In the third paragraph of Section 3.1.1 (Updated Surface Water COPEC Screen), please note that Montana's DEQ-7 Water Quality Standards cannot be exceeded therefore, please ensure the ESVs used are below DEQ-7 standards and include a statement that no screening values exceeding DEQ-7 Water Quality Standards were used in this process. Also, please note the DEQ-7 Standards were updated in June 2019.	As described in the BERA, the hierarchy for selected ESVs prioritizes NRWQC values. A footnote has been added to note that Montana DEQ-7 Water Quality Standards will be considered within the Feasibility Study.
54	MDEQ	COPEC Screen	Section 3.1.1, 4th paragraph: It should be noted that DEQ-7 standards for metals (except aluminum) in surface water are based upon the analysis of samples following a "total recoverable" digestion procedure. Therefore, please ensure the appropriate data is compared to DEQ-7 aquatic life standards.	Montana DEQ-7 Water Quality Standards will be considered within the Feasibility Study.
55	MDEQ	Background Evaluation	In Section 3.3.3.3 (Soil), the text states that " <i>there are no background data for PCBs in soils.</i> " There are EPA nationwide studies of PCB background levels in soils. Could it be compared to that?	Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment.
56	MDEQ	Background Evaluation	In Section 3.3.3.3 (Soil), please clarify whether or not the following statement is referring to OU2, OU3, or both: " <i>Site concentrations of copper, mercury, selenium, and zinc in Site soils were statistically greater than background concentrations.</i> "	Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
				existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment.
57	WQAC	Site Characterization	The data upon which the risk assessment has been made are insufficient. Problems include: a. too few samples to adequately characterize risks on and beyond the site and no recent samples in important areas such as the 140 acres of sludge ponds. b. combining samples together, which has the effect of losing important detail associated with the site given that contamination is not evenly distributed. c. need for additional dioxin testing of groundwater and surface water of the Clark Fork using high-volume sampling for dioxin-like PCBs that have detection limits above the risk-based levels. The lack of site and river sampling is hiding the potential connection of the site to the river, which limits the geographical extent of the “facility.” CERCLA’s definition of a facility is “any site or area where a hazardous substance has been deposited, stored, disposed of, placed, or otherwise come to be located,” so this type of testing is important for defining the facility site boundary to inform site remediation. For reference, we include text from WQAC letter sent to Sara Sparks and Keith Large on July 9, 2018: The WQAC discussed the May 23, 2018 Water Sample Results (Dioxin/Furans) Former Smurfit Stone/Frenchtown Mill Site at our June meeting. It was clear from looking at the results that the detection limits and reporting limits of the analytical method are not useful for resolving concentrations of dioxin near surface water standards. We understand this is a common situation at dioxin contaminated sites in the nation, because the standard method (EPA 8290) commonly results in detection limits of 1 pg/L at best, whereas surface water standards are 200 times or more below this detectable concentration. The human health standard for dioxin is very low, 0.005 pg/L, owing to its incredible high bioaccumulation and toxicity when accumulated in fish tissue.	Composite sampling reduces data variability and provides a reasonably unbiased estimate of mean contaminant concentrations in a specified decision unit (ITRC 2021). Additional data collection efforts to support the risk assessments are not planned at this time. Uncertainties associated with the available database are described in the uncertainty assessment.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			Dioxin is extremely expensive to sample for and we do not believe additional sampling at using the past method is likely to provide useful information about the contaminants at the Smurfit site or Clark Fork River water quality. Newer dioxin sampling methods are able to detect the presence of dioxin in surface water at lower concentrations. The Ohio River Valley Water Sanitation Commission (ORSANCO) uses a high-volume water sampling method to obtain detection limits and reporting limits 1000 times lower than the sampling method currently being used (DL below 0.001 pg/L (parts per quadrillion, ppq) for dioxin). This actually resolves pollution at and near the level of Montana DEQ-7 human health surface water of 0.005 pg/L. I (Ian Magruder, Chair of the WQAC) have personally talked with Sam Dinkins at ORSANCO who described their use of the method. He reports it has been useful for them to characterize ambient dioxin concentrations in the Ohio River at very low levels. The WQAC believes that using this method may help to resolve lingering questions about dioxin impacts to the Clark Fork from Smurfit. Given the expense of sampling dioxin we believe it would be more worthwhile to collect fewer samples using the high-volume method which provides meaningful results, than the current sampling plan which does not add much to site characterization.	
58	WQAC	Flooding Concerns	The assessment fails to include the scope of risks known to have the potential to affect the site and surrounding environment. Examples include risks associated with channel avulsion and berm failure.	Risk assessment does not/cannot account for potential future catastrophic events such as berm failure. However, the PRPs, with EPA and MDEQ oversight, conducted a berm stability study to address these concerns. This study will be cited in the BERA.
59	WQAC	Site Characterization	The risk characterization in the documents we comment on here represent a gross under-sampling of the area and (given how data were combined) do not sufficiently identify hot spots or the extent of contamination (facility boundary). Consequently, the assessment very likely represents a lower level of risks than what actually exist. Further, the document completely omits potential known risks such as channel avulsion and berm failure. We believe that understanding of the site is therefore still subject to an unacceptably high level of uncertainty.	EPA believes that soil sampling conducted to date adequately characterizes surface soil conditions at the site for risk assessment purposes.
60	MVWQD	Conceptual Site Model (CSM)	Conceptual Site Models (site-wide, groundwater and air) need further development before finalizing risk assessments. Risk cannot be adequately assessed, nor pathways fully understood, until a sitewide CSM is complete, including an understanding of aerial deposition. We believe the comments that have been submitted regarding the Groundwater CSM should be addressed prior to evaluation of the risk assessments. We are fundamentally concerned that, because we don't understand the idiosyncrasies of the contaminant pathways that the risk assessments may be missing COPCs and evaluation of the most-contaminated areas (that would still be routes of exposure). Specifically: •Heterogeneity in well logs provides evidence of preferential flow paths not incorporated in any CSMs.	The CSM presented in the BERA is specific to identifying ecological receptors and exposure pathways. This objective differs slightly from the fate and transport CSMs being developed in other site documents. EPA recently published an aerial deposition fact sheet (EPA 2021a), which will be cited in the BERA. In addition, the groundwater CSM (NewFields 2020) is

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			•Background levels for metals within the Missoula aquifer are lower than those used in Smurfit Stone Assessments. Studies indicate highly variable interaction between groundwater and surface water. Variable water levels will drive changes in groundwater pathways and need to be accounted for.	currently being revised and will be cited in the BERA as well.
61	MVWQD	Flooding Concerns	River channel migration is a risk that warrants assessment. Smurfit Stone berms impede natural river function by constricting movement of the Clark Fork River. Side channels that were active in 1955 can be identified aerially. These antecedent channels are indication of areas where there is likely preferential flow common to alluvial deposition. These preferential flows present risk of undermining of the berm through winnowing. Ten repairs between 1974 and 2007 were identified in the Missoula Conservation District 310 permit records. While most of the repairs involved adding significant armor to the bank, some of these repairs were not directly related to flood damage but rather animal burrows resulting in partial berm failure, or barbs to redirect the river's energy away from the berm that were causing erosion at normal flow. Flood analyses do not take into account the impact that ice jams, frequent on the Clark Fork River, may have in damaging the berms. Failure of internal berms, which have not been evaluated recently, and do not require permitting, could lead to increased risk of river avulsion and unpredictable erosion of sludge ponds, dumps, and contaminated soils. Considering the likelihood of channel migration, current constriction of the river, previous examples of berm failure unrelated to high flows, and highly variable state of berm maintenance, we believe the risk assessments to be inadequate by not addressing the risk of avulsion and channel migration.	Risk assessment does not/cannot account for potential future catastrophic events such as berm failure or river channel migration. However, the PRPs, with EPA and MDEQ oversight, conducted a berm stability study to address these concerns. This study will be cited in the BERA.
62	MVWQD	Data Evaluation	Percent of fine sediments (% Fines) need to be accounted for and normalized when assessing metals and dioxin concentrations. Dioxins and heavy metals sorb to fine materials more than larger grain sediments. The surface area of smaller particles also results in greater concentrations of contaminants by volume. Fines percentages and effect on samples were not discussed in analysis of the 2016 Clark Fork River sampling or 2017 soil sampling. Nor is percent fines data readily available through SCRIBE. Fine sediment is where contaminants are most likely to be found and therefore should be targeted to identify levels of contamination across geographic areas – otherwise comparison is not valid. This is commonly understood science that should have been incorporated into sampling design and quality control thresholds. Initial review of data from the 2016 Clark Fork River sediment sampling shows the samples that had the most fines also tended to have the highest TEQ values. There is also a linear relationship between fines and aluminum (graph included in original comment letter). Inclusion of the holding pond data results in more variability, likely due to the range of 93% to 3.4% fines. It is likely more contaminants of concern would have been found in sampling if fines were targeted. Within the Clark Fork River, samples cannot be compared to understand upstream, site, and downstream levels of contaminants without further standardization by further	Sediment sampling conducted in 2015 and 2018 targeted fine grain sediments from depositional areas as described in NewFields (2015, 2018). These sampling locations were also selected to be spatially representative of the site to support risk assessment evaluations. Within a risk assessment it is assumed that a receptor will be randomly exposed within an exposure area. Thus, by selecting spatially representative locations, data were collected to inform potential chronic exposures over time, while at the same time targeting fine grain sediments from depositional areas where high proportions of fines may provide a greater surface area (compared to coarse-grained

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			sampling of fines. Data is insufficient to determine whether there are risks in either the BERA or HHRA until adequate fine sediment sampling can occur. 2017 soil sampling data includes surface soil samples and subsurface soil samples. The subsurface samples range in fines from 97% to 0.4% fines. Surface samples range from 87% to 3.7%. This range of variability means that contaminant results from data cannot be compared. Given the toxicity of even the tiniest amount of TEQ, data cannot be normalized to account for this degree of variability. Although it was stated in a presentation regarding the BERA and HHRA that samples were reflective of environments on the site, variability does not reflect onsite use. The primary use for OU3 was for fines to settle in order to meet permit requirements. Fines variability of this degree is likely a function of sampling rather than site characteristics. Further review of data is necessary to determine what data can be compared and is reflective of contaminants on site. Without this analysis, much of the data used within the BERA and HHRA is misleading.	sediments) for sorption of organic carbon and contaminants. Additional sampling at the site is not planned prior to finalization of the BERA.
63	MVWQD	Site Characterization	Smurfit-Stone Facility Boundary. CERCLA's definition of a facility is "any site or area where a hazardous substance has been deposited, stored, disposed of, or placed, or otherwise come to be located." As more data is collected onsite, it is becoming clear that barriers are not in place to prevent groundwater and surface water interactions or to prevent flood waters from impacting the site and carrying contaminants downstream. The release and contamination of soils and water by pulp and paper mills has been well-documented by the EPA. Specifically, manganese is credited as the metal with the highest toxic-weighted pound equivalents (TWPE) reported by the Toxics Release Inventory (TRI) database for mills (EPA, 2006). Given that high levels of manganese have been detected in groundwater onsite, and at least four samples adjacent to or downstream of the site exceed the chronic toxicity benchmark for manganese, does the EPA believe that this contaminant is migrating offsite? If so, what are the next steps to further sample the Clark Fork River for manganese and other contaminants and assess risk?	Additional sampling is not planned by EPA prior to finalization of the BERA. Further nature and extent evaluation will occur in the RI. Manganese is evaluated as a COPEC in surface water (including the CFR), pore water, pond sediment, and in wildlife in the BERA.
64	MVWQD	Assessment Endpoints	Section 2.5 Assessment Endpoints, Table 2-4 and Section 4.1.3 Exposure Assessment: In the Exposure Assessment, terrestrial invertebrates were chosen as an ecological resource worth protecting. We support this goal. However, no site-specific toxicity tests or community surveys have been conducted. Instead, levels of contaminants within the top 0-7 inches of soil were used to model toxicity to terrestrial invertebrates. In the 1980s EPA tested sludge from pulp and paper mills and found that dioxins were highest in sludges from bleached kraft mills (Fed. Reg., Vol. 58, p. 66092). The Smurfit-Stone site is estimated to contain 138 acres of sludge ponds and dump sites. We understand that the risk to invertebrates from soil concentrations of dioxins onsite was not determined due to the absence of a soil ESV for invertebrates and general assumption of low dioxin/furan susceptibility (invertebrates generally lack the AhR receptor). However, isn't it also possible that other contaminants in sub-surface soils, specifically metals in dumps/landfills, are high enough to cause adverse effects to soil invertebrate (e.g.	EPA does not discount the use of site-specific toxicity testing and benthic community surveys, and agrees that these data can be useful in evaluating ecological risks. EPA characterized risks based on the HQ approach as the existing site database supports this line of evidence. If data derived from site-specific toxicity tests or community surveys is available in the future, the results of those tests will be considered by EPA.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			earthworms) populations, reproduction, or survival? And because one of the primary goals of the BERA is to determine if soil levels of contaminants are high enough to adversely affect survival, growth or reproduction of invertebrates, why not survey the existing population? If chronic contamination has already adversely altered the population, how would we know? Since the sludge ponds and dumps contain contaminants of concern aside from dioxin, including high levels of metals (the most significant risk exceedances (of HQ values > 1) for soil invertebrates are with barium, mercury, and zinc), why have there been no site-specific toxicity tests conducted for soil invertebrates? Due to the historic uses in the OU3 exposure area, including sludges and wastes buried at depths between 7” and > 6’, soil exposure assumptions do not adequately reflect onsite risk to terrestrial invertebrates. The District requests further testing be conducted to complete the Exposure Assessment (community surveys and site-specific toxicity tests).	
65	MVWQD	Site Characterization	Overall, the risk assessments need improved spatial representation of soils in both operable units. Risk assessments inform risk management decisions. We fear that inadequate sampling now will underestimate risk and carry through to the feasibility study. We disagree with taking composite samples in exposure units being represented as a 20-acre grid for upland in OU3 and 100-acre grid for the OU3 floodplain. If these areas were comprised of homogenous strata and no knowledge of potential historic use differences existed between them, then it would be a more appropriate sampling design. However, at Smurfit, even within an exposure unit, the historic uses vary. Within each type of historic use, the waste types and stratification vary. Within a single grid there could be 10 ft deep solid waste dumps, settling ponds, aeration basins, conveyances, sludge impoundments, etc. If the goal of the sampling method is to identify hot spots or delineate extent of plumes, compositing on a heterogenous site and media is not appropriate when other sampling methods exist. One such method would involve dividing the site by strata (waste layers, use types, etc.), taking grid or random samples of those strata, and then employing an adaptive cluster method, would provide more meaningful insight. When a sample comes back elevated or a new contaminant is identified, this method allows those “hot spots” to be mapped as adjacent samples are taken instead of losing the spatial information and variability in the data common with compositing large areas.	EPA believes that soil sampling conducted to date adequately characterizes surface soil conditions at the site for risk assessment purposes. The focus of a risk assessment is on estimating exposures that a receptor may reasonably be exposed to within a given exposure area over their entire exposure duration. Composite sampling reduces data variability and provides a reasonably unbiased estimate of mean contaminant concentrations in a specified decision unit (ITRC 2021). The 20-point/30-point composite samples were collected to provide an estimate of the mean concentration over an identified exposure area (e.g., grid). Discrete samples (5-point composites collected over a 1-m2 area) were collected under a biased sampling strategy to target areas of most concern. Due to low variability within the current data set, EPA is not currently recommending additional soil sampling to support the risk assessments beyond that which was carried out in the RI/FS Addendum 7 (2017) in the fall of 2017.
66	MVWQD	COPEC Screen	Despite new sampling efforts, the creek sampling is insufficient for assessing risk questions. Because of the statistical limitation of the low number of surface water	Although conducting a comparison to background concentrations was

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			samples, the COPEC data for the Lavalley and O’Keefe Creek was combined instead of data from each creek being compared to the background concentrations. Taken alone, dioxins and metals in O’Keefe Cr sediments were higher than background for arsenic, copper, and zinc and thus should have been retained as COPECs. This averaging of data obscures the differences between the creeks (surface water levels of manganese were 15x higher in LaValley Cr compared to O’Keefe), and led to the conclusion that the creek contaminants are statistically equal to upstream values, and eliminated COPECs that otherwise would have been retained. More surface water samples should be taken to achieve a statistically justifiable assessment of each creek independently.	included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment.
67	MVWQD	Background Evaluation	Section 3.3.3.2 Results of the Background Comparisons – Sediment, HHRA 3.3.5, 3.1.2 Updated Sediment Screen, Fig 3.2. Bleaching pulp creates specific substituted CDDs and CDFs. In the 1980s in response to a lawsuit to regulate dioxins, the EPA tested wastewater and sludge from 104 mills that bleached pulp. No mills were free of dioxins and furans and these contaminants were highest in the sludges of bleached kraft pulp mills (EPA, 1990). As a result of this study and analysis of effluent and sludge from kraft pulp mills since then, the EPA has identified congeners and developed default emissions factors for these congeners deemed to be representative of the pulp wastewater and sludge categories (EPA, 2019). The Smurfit-Stone mill produced bleached pulp and paper from 1960 through 1999. It was estimated that mill operations discharged an average of 4.8 million gallons of wastewater from the bleach plant and generated 54 tons of sludge each day. Sludge at Smurfit was pumped from the clarifier into four sludge ponds numbered 3, 4, 5, and 17. An elevated risk to residents exist for TEQ in surface soils in P17. But a clarifier wasn’t built until 1969 so in earlier years this sludge was part of wastewater and wastewater ponds as well. As described in the initial site investigation in 2011, Tables 3-1 to 3-11 of the BERA, HHRA, Table 3-3 of HHRA, dioxins are COPECs onsite have been detected in OU2 and OU3 soils, floodplain soils, pond sediment, pond water, and deeper soils. Despite the clear impact to the site, the BERA concludes that presence of dioxins in the CFR and bed sediments appears indistinguishable from background and poses no significant risk to receptors in these matrices (HQ < 1). Given the long history of	Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			discharge of dioxins and furans from the site, the conclusion about risk to the river and its receptors is worth consideration and could be interpreted in a number of ways. First, isn't it possible the "background" is not truly background? How do we know the areas selected roughly < 10 miles upstream of the site were not historically influenced by air-deposition? Aerial deposition modeling suggests at much as 20% of total emissions can be deposited within hundreds of kilometers surrounding a source (Cohen et al. 2012). Montana DEQs state-wide dioxin soil study showed many high concentration outliers in the Missoula area suggesting the mill or other industrial sources have impacted areas upstream of the mill. The same study also reports statistically significant differences between the rural and urban areas of the state (urban TEQ was 7.5 ng/kg and the rural TEQ was 1.8 ng/kg) casting doubt on relevance of comparisons of the mill site, one that is historically rural and agricultural, to areas that are urbanized or influenced by urban areas. Also, even though the differences between upstream and downstream samples for dioxins and furans are nuanced once normalized as TEQ, individual congener profiles up and downstream are different. This suggests that site conditions may have contributed and still could be contributing to impacts downstream. Second, the risk may be low simply because dioxins and furans are not present in the samples. Due to annual runoff washing away bed sediments and ongoing photodegradation this is reasonable. Alternatively, dioxins and furans may be present but not in the areas chosen for sampling. Only 3 bed sediment samples were taken in areas adjacent to site. An inside bend spanning 1.25 miles between outfall 2 and outfall 3 was omitted from any sediment or surface water sampling. There are a number of inside bends omitted downstream as well (e.g., between 60-CFR and 61-CFR, 61-CFR and 62-CFR). And even though the number of total sediment samples analyzed for the BERA (20 total) was a big improvement over the small sample size (n= 6) used in the SLERA, the District feels more sampling of bed sediment at appropriate slack water areas are needed. Unlike surface water and some sediments, fish have longer residence times in the area. The BERA and data from FWP demonstrate that dioxins and furans consistent with pulp and paper mills (FWP study) are found in fish (i.e. 5 of the 10 compounds in TRI report in pike: 1,2,3,4,6,7,8 HpCDD, 1,2,3,4,6,7,8,9 OCDD, 2,3,7,8 TCDF, 2,3,4,7,8 PeCDF, and 1,2,3,6,7,8 HxCDD. The presence of pulp-mill related dioxin/furan congeners onsite, in the river sediment, and in fish, is a concerning pattern. As is the fact that these contaminants persist despite flooding, erosion and over 10 years passing since the mill's last documented discharge. How do we know the site is not continuing to contribute to the levels offsite? More sampling is needed in areas shown to have high levels in the BERA and HHRA, namely sludge ponds (P3, P4, P5, P17) to assess the potential risk to receptors exposed to downstream surface water and sediment as these sediments and soils leave the site.	EPA recently published an aerial deposition fact sheet (EPA, 2021a) that will be cited in the BERA and addresses these concerns.
68	MVWQD	Conceptual Site Model	Landfill cells and sludge ponds should be evaluated as independent exposure units for human health and ecological risk. The BERA states that "the EPA has developed this	The CSM shown in the BERA is intended to highlight the complete

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			BERA to quantitatively describe the potential excess ecological risk posed by the Site in the absence of remediation”. The BERA also identifies the leaching of contaminants into groundwater and migration of contaminants from groundwater to soils, sediment, and surface as important “Fate and Transport” mechanisms for the site. In Section 3.1.1 Primary Sources of Contamination of the draft HHRA the historic use of OU3 for “storage of wastewater and solid wastes” is identified as the “primary source of contamination” of OU3 soils. It is therefore surprising that the largest repository of waste on the site, namely the buried sludge and waste, amounting to 91 acres of sludge ponds that received an estimated 20,000 tons of sludge per year, and 47 acres of other solid waste dump sites, are omitted as a source of direct contamination of soil in the CSM. Groundwater data is sparse in characterizing the groundwater under individual sludge ponds and landfill cells. The CSM described in the BERA does consider solid waste leaching to groundwater as a source. It would be more accurate to sample the interface of this waste with groundwater. The BERA CSM also does not show groundwater connected to the CFR from a solid waste/surface soil path. This pathway should be evaluated thoroughly.	exposure pathways from potentially contaminated media present at the site owing to site operations or waste disposal activities. The groundwater CSM (NewFields 2020) will be cited in the BERA as well. Discrete samples (5-point composites collected over a 1-m² area) were collected under a biased sampling strategy to target areas of most concern. Due to low variability within the current data set, EPA is not currently recommending additional soil sampling to support the risk assessments. Documentation regarding the waste management area (WMA) is located in DEQ, EPA and Missoula County files; the site was well regulated beginning in the early 1980s. The PA/SI (and the appendices) conducted by EPA in 2011/2012 references documents about the older materials generated during operation including concentrations and quantities. Additional work to better understand the WMA conducted during EPA's PA/SI work effort included soil boring within the WMA. The groundwater sampling data collected at the site are being used to evaluate this potential transport pathway within the RI.
69	MVWQD	Data Evaluation	Section 6.1.3: Earthworms (<i>E. foetida</i>) provide a food base for many higher order receptors including shrews, voles, and a variety of birds. In fact, due to their exposures to various soil depths (0-6 feet), feeding habits, and distributions, they have been selected as representatives of terrestrial invertebrates in other superfund risk assessments to evaluate not only site-specific toxicity but also bioaccumulation impacts to the higher trophic consumers (e.g. American Woodcock Earthworms as relevant to food-web toxicity) (Powell, 1997; EPA, 2004). However, as stated in the Uncertainty Assessment, no data has been collected for contaminants in terrestrial invertebrate tissue. The BERA states that	While site-specific soil invertebrate data would reduce uncertainties associated with modeling uptake of COPECs from soil to earthworms, conservatism is built into the risk calculations to account for this uncertainty as discussed in the uncertainty assessment.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			equating the concentration in the top 0-7" soil with invertebrate tissue concentration ($C_{\text{soil}} = C_{\text{tissue}}$) conservatively assesses risk to higher trophic receptors. We feel this inadequately quantifies risk. Isn't it possible that earthworms are bioaccumulating soil contaminants such that $C_{\text{soil}} < C_{\text{tissue}}$? This is especially relevant considering that earthworm habitat extends far beyond 7" of soil and are expected to reside in deeper sub-surface habitats comprised of historic dumps and sludge basins. In lieu of direct sampling, why haven't earthworms been exposed to site soils and analyzed for site-specific bioaccumulation of soil contaminants in sludge (dioxins, furans, metals)? We would like to see risks to higher trophic level receptors (robins, shrew, etc.) modeled based on these site-specific contaminant levels in soil invertebrates.	
70	Missoula Co Commissioners	Site Characterization	The structure of the sampling that has been used to evaluate the site is fundamentally structured in such a way that risks are muted.	EPA believes that soil sampling conducted to date adequately characterizes surface soil conditions at the site for risk assessment purposes. The focus of a risk assessment is on estimating exposures that a receptor may reasonably be exposed to within a given exposure area over their entire exposure duration. Composite sampling reduces data variability and provides a reasonably unbiased estimate of mean contaminant concentrations in a specified decision unit (ITRC 2021). The 20-point/30-point composite samples were collected to provide an estimate of the mean concentration over an identified exposure area (e.g., grid). Discrete samples (5-point composites collected over a 1-m ² area) were collected under a biased sampling strategy to target areas of most concern. Due to low variability within the current data set, EPA is not currently recommending additional soil sampling to support the risk assessments.
71	Missoula Co Commissioners	Flooding Concerns	The risk of a catastrophic channel avulsion into the site from the Clark Fork River are not evaluated.	Risk assessments are performed using available site data and presumed exposure scenarios to estimate the chronic risks posed by the contaminants present at the site. CERCLA risk assessments cannot and

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
				should not be used to estimate short-term risks from hypothetical future events such as flooding, since the extent of soil disturbance, the location of such erosion, and the lateral extent of deposition cannot be known ahead of time.
72	Missoula Co Commissioners	Conceptual Site Model	Conceptual Site Model should include updated Groundwater Conceptual Site Model and an understanding of aerial deposition.	The groundwater CSM (NewFields 2020) and aerial deposition fact sheet (EPA 2021a) will be cited in the BERA.
73	Missoula Co Commissioners	Site Characterization	Regarding sampling structure: Onsite sampling of soils and groundwater appear to be weighted heavily on grid-based sampling on 20-, 30- and 100-acre grids. We do not understand how this accurately describes the risks of the particularly toxic operations that occurred in discrete areas of the site (specifically the dumps and landfills). As stated in previous correspondence, along with provided references and photos, we are concerned that there are buried materials on this site that are contaminating groundwater or will contaminate groundwater (leaking drums and sludge from the days of bleaching pulp). We believe the risk assessment exposure units should be broken up by discrete units for each individual landfill and sludge pond. Some sites are difficult to ascertain where dumping took place and thus, require a grid-based approach. This site has very discernable boundaries of cells due to the tremendous amount of earth moving that took place (even artificially altering floodplain boundaries by filling the site with the waste that's being evaluated).	Risk assessment evaluates potential exposure over areas not necessarily defined by past operations. Regardless, in addition to the grid sampling, discrete samples (5-point composites collected over a 1-m ² area) were collected under a biased sampling strategy to target areas of most concern. EPA sampled and evaluated all storage ponds and cells throughout the entire treatment chain between the plant and the river, along with select locations in O'Keefe and LaValle Creeks.
74	Missoula Co Commissioners	Flooding Concerns	Regarding channel movement: The Clark Fork River meanders. Over time, our community has experienced the implications of this meandering in the loss of bridges, roads and structures. We know the best intentions to mitigate channel migration through engineering are not without their own consequences, shifting the problem upstream or downstream. Missoula's own floodplain regulations prohibit new use of riprap because of the consequences of shifting energy downstream. Plainly stated, the berms surrounding the Smurfit Stone site are masking a huge liability and risk to the ecosystem and to people. Catastrophic failure is a real threat, and that risk should be evaluated. What would the ecological and human health effects be of the Clark Fork's reoccupation of its flood plain or river channel?	Risk assessment does not/cannot account for potential future catastrophic events such as berm failure or river channel migration. However, the PRPs, with EPA and MDEQ oversight, conducted a berm stability study to address these concerns. This study will be cited in the BERA.
75	Missoula City-County Board of Health	Site Characterization	The inherent ability to miss-characterize the presence of contaminants by grid-based composite sampling is concerning. We disagree with taking composite samples in exposure units being represented as a 20-acre grid for upland in OU3 and 100-acre grid for the OU3 floodplain. If these areas were comprised of homogenous strata and no knowledge of potential historic use differences existed between them, then it would be a more appropriate sampling design. However, at Smurfit, even within an exposure unit, the historic uses vary. Within each type of historic use, the waste types and stratification vary.	EPA believes that soil sampling conducted to date adequately characterizes surface soil conditions at the site for risk assessment purposes. The focus of a risk assessment is on estimating exposures that a receptor may reasonably be exposed to within a

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			External review by Abt Associates found “Large areas of the site are represented by a single data point, which may not be representative of the potential exposure. All of the data from 11 ponds in OU3 covering hundreds of acres were digested down to a single exposure area using one assumed value for each contaminant. The existing data may not capture the areas of contamination and potential risk when such broad areas are represented by composited samples and aggregated data.” The Board is concerned that the dumps and sludge ponds are not adequately characterized. The groundwater flow is not uniform because of the disruption of the site geology and the remnant river channels that run through the site. Risk from groundwater and soils should be characterized based on historic use (evaluated by landfill cell/sludge pond cell) rather than in grids that composite and run across multiple historic uses.	given exposure area over their entire exposure duration. Composite sampling reduces data variability and provides a reasonably unbiased estimate of mean contaminant concentrations in a specified decision unit (ITRC 2021). The 20-point/30-point composite samples were collected to provide an estimate of the mean concentration over an identified exposure area (e.g., grid). Discrete samples (5-point composites collected over a 1-m² area) were collected under a biased sampling strategy to target areas of most concern. Due to low variability within the current data set, EPA is not currently recommending additional soil sampling to support the risk assessments. Documentation regarding the waste management area (WMA) is located in DEQ, EPA and Missoula County files; the site was well regulated beginning in the early 1980s. The PA/SI (and the appendices) conducted by EPA in 2011/2012 references documents about the older materials generated during operation including concentrations and quantities. Additional work to better understand the WMA conducted during EPA's PA/SI work effort included soil boring within the WMA. The groundwater sampling data collected at the site are being used to evaluate this potential transport pathway within the RI.
76	Missoula City-County Board of Health	Conceptual Site Model	Risk cannot be adequately assessed, nor pathways fully understood, until a sitewide Conceptual Site Model is complete. Determining risk, biological or human health, is premature when issues brought forward in the Groundwater Conceptual Site Model have not been addressed or responded to. Further, air deposition of contaminants is widely acknowledged, yet aerial deposition of contaminants is not fully discussed or included in	EPA is already considering aerial deposition within the risk assessment CSM and the broader RI. The groundwater CSM (NewFields 2020) and aerial deposition fact sheet (EPA

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			identification of background levels. Even though current investigations are identifying contaminants at present, contamination from the site has extended beyond the site boundaries and should be included in risk assessments. Although CSMs can be updated, incorporating additional COPCs at a later stage puts the onus on the public as opposed to the PRPs to provide evidence of contamination.	2021a) will be cited in the BERA. Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment.
77	Missoula City-County Board of Health	Site Characterization	The risk assessments remain incomplete in their failure to include what may be the most important exposure pathway: erosion and discharge of contaminated material into the Clark Fork River during high water events. Channel migration and avulsion could wash large volumes of sediment into the river and impact this resource into the future.	Risk assessments are performed using available site data and presumed exposure scenarios to estimate the chronic risks posed by the contaminants present at the site. CERCLA risk assessments cannot and should not be used to estimate short-term risks from hypothetical future events such as flooding, since the extent of soil disturbance, the location of such erosion, and the lateral extent of deposition cannot be known ahead of time.
78	CSKT	Exposure Assessment	The BERA does not assess the ecological risk posed to the endangered bull trout or to a surrogate species with similar habitat requirements and feeding habits. Bull trout is one of the most important species to the CSKT. EPA's own guidance states that "When conducting an ecological risk assessment, it often is necessary to determine or estimate the effects of site contaminants to federal threatened or endangered species." (EPA 1997a). As CSKT explained to EPA at the October 2018 meeting, and in other occasions, CSKT's	Appropriate surrogates are chosen for the assessment endpoints that represent site receptors and are inclusive of threatened and endangered species, if present. The BERA included evaluation of northern pike as a surrogate for a

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			vision for this stretch of the CFR is one of full restoration of the site such that it could once again support a thriving bull trout population. Particularly for an endangered species that is also of such importance to the CSKT, the failure to analyze the risk to bull trout or a surrogate species is a significant omission.	high trophic level piscivore (i.e., bull trout); risks to northern pike within the study area appear to be below a level of concern as reported in the BERA.
79	Trustees	Data Evaluation	Request to Add Data to SCRIBES: As part of the remedial investigation /feasibility study process, the National Contingency Plan (NCP [40 CFR 300.430(b)(1)]) specifically requires the U.S. Environmental Protection Agency (EPA) to “(1) Assemble and evaluate existing data on the site, including the results of any removal actions, remedial preliminary assessment and site inspections, and the NPL listing process.” This specifically includes documents such as the Hazard Ranking System (HRS) evaluation for the Smurfit-Stone Mill Site (Site). This investigation revealed the presence of dioxin/furans and dioxin-like compounds in waste piles on the former mill property and sediment in unlined storage ponds on the former mill property. Furthermore, these compounds were also found in fish tissues near the Site and downstream of the Site. Accordingly, there is a clear pathway from the Site to the Clark Fork river and the presence of these highly toxic compounds in the environment both on the former mill property and off of it (including in fish tissue) requires further investigation. The Trustees specifically request this data be added to the Site database (SCRIBES) and used in the BERA and HHRA. If EPA is unwilling to include this data, the Trustees request a formal response regarding why this data should not be used in evaluating Site risks.	Data collected prior to 2014 to support the Preliminary Assessment/Site Investigation are not used in the BERA due to issues raised with data quality. EPA used those data to inform the sampling plan and focused collection on those areas identified as having potentially elevated contamination. The sampling conducted as part of RI activities are more recent, were collected in accordance with the EPA approved QAPP, and have been evaluated and determined to meet the needs of the risk assessment.
80	Trustees	Site Characterization	The Trustees request that additional Site characterization, including potential locations in the Clark Fork River where hazardous substances have become located, be performed prior to finalization of the HHRA and BERA for the Site. Sampling completed to date is insufficient to adequately characterize risk to human health and the environment. Dioxins and furans are persistent highly toxic organic pollutants even at low level at or near analytic detection limits. Because of their hydrophobic nature and extremely low probability of detection, sampling methods such as surface water grab samples and groundwater samples are not adequate to characterize contamination and associated risk to human health and biota. Existing data show that there are coplanar PCBs in OU2. Further characterization is warranted to define the extent of coplanar PCBs and to identify potential pathway (both current and historic) to the river. The Trustees request that the overall Site Conceptual Model and the groundwater Site Conceptual Model be further developed and finalized to address all potential exposure pathways from the Site to the river prior to finalization of the BERA and HHRA. The NCP provides direction on collecting enough data to characterize the nature and extent of the contamination. The data collected must allow EPA to assess, “(iii) The general characteristics of the waste, including quantities, state, concentration, toxicity, propensity to bioaccumulate, persistence, and mobility, and (v) Actual and potential exposure pathways through environmental media.” 40 CFR 300.430(d)(2). The data	A large amount of data has been collected and evaluated at the site since 2014. Additional data evaluations will occur to support the Feasibility Study. The full nature and extent to which hazardous substances from site activities have come to be located in the Clark Fork River adjacent to and downstream of the Site will be considered within the RI. Based on this, the term “site” will be replaced with “study area” within the text of the risk assessment.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			collected then must be used to conduct a “site-specific baseline risk assessment to characterize the current and potential threats to human health and the environment that may be posed by contaminants migrating to ground water or surface water, releasing to air, leaching through soil, remaining in the soil, and bioaccumulating in the food chain.” 40 CFR 300.430(d)(4). Therefore, the NCP requires enough data to identify quantities of contamination that have been released from the original mill property into the Clark Fork River and to determine the exposure pathways in the Clark Fork River downstream of the Site. As noted elsewhere in this comment letter, we do not believe that this data has been fully collected as required by the NCP. Further, the risk assessments are not complete until all the pathways for migration to and through surface water and bioaccumulating in the food chain have been identified, so we would request that EPA reevaluate the risk assessments as additional information is gathered as part of the process of finalizing the two Site Conceptual Models and as additional data is gathered to fully characterize the nature and extent of contamination in the Clark Fork River. The 104 Mill Study (EPA, 1990) included samples of effluent, sludge, and pulp at the Site. 2,3,7,8-tetrachlorodibenzofuran (TCDF) was detected in samples of effluent, sludge, and pulp. Note the study had a relatively high detection limit. This study provides evidence that dioxin and furans were produced at the Site and have the potential to come to be located in the Clark Fork River sediments downstream of the Site. The Trustees strongly request that additional sampling of sediment depositional areas downstream of the Site within the Clark fork River be performed to further define the OU3 boundary and evaluate if contamination from the Site has come to be located in these areas. This type of paper mill has been shown (based on EPA studies) be a source of dioxins and furans and these depositional areas within the Clark Fork River (and likely many miles downstream from the Site, and even in the reservoirs downstream) may pose an ongoing unacceptable risk to the environment. In addition to sampling of sediment depositional areas downstream of the Site, the Trustees request high resolution analytical methods specifically targeted to quantify hydrophobic chemicals that pose a risk at low levels, such as dioxins, be used. These methods have been used for many years at other sites with known dioxin issues and are critical for informing risk management at sites where persistent organic pollutants are present.	
81	Trustees	Background Evaluation	Request for Additional Background Analysis: The Trustees understand that EPA is preparing a technical memorandum addressing air emission from the Site. The Trustees look forward to reviewing and commenting on this document. However, the Trustees suggest that airborne emissions from the Site did contain dioxins and furans and the air deposition was widespread throughout the Missoula Valley. EPA has prepared technical guidance that dioxins and furans are aerially emitted from the type of kraft paper processes used at the Site (EPA, 1997b). Therefore, the Trustees disagree with using locations within the aerial emissions area to compare to background concentrations for all matrices including soil, sediment, water, and tissue. Further sampling and characterization	Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			should be performed to determine the aerial influence of the Site and to select background locations that are outside the area where aerial contamination from the Site came to be located.	evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment. EPA's recently published fact sheets on aerial deposition and dioxins/furans (EPA, 2021a,b) provides further information on these topics and will be cited in the BERA.
82	Trustees	Data Evaluation	Composite Fish Tissue Samples: The use of composite samples is not appropriate for aquatic-dependent receptors or to assess risk to individual fish based on body burdens. Additionally, the BERA only lists a single data point (Figure 5-13) for a species at a sampling location. Stark differences in Toxic Equivalent (TEQ) concentrations exists between fish sizes. Since aquatic-dependent receptors target specific sizes of prey, it follows that average size group concentration and not average total species concentrations should be assessed at each location. In most cases this greatly increases the concentrations and associated risk at each location. Given the variability found between composites of same sized fish species at a location (i.e., as high as 2632% for fillets and 342% for whole body) it is reasonable to assume an even higher degree of variability occurs within a composite sample. Based on the observed variation of TEQ in fish tissues from composites (which are designed to provide average conditions), a statistical calculation should be conducted to predict the highest likely concentrations within the composite. The majority of fish tissue concentrations sampled at a location, within a species and size, were more than two standard deviations apart. If one fish within each 5-fish composite has elevated concentrations of TEQ, then 20% of the population could potentially be at risk. The BERA states <i>"if all or nearly all of the HQs for individuals in a population of receptors are below 1.0, it is very unlikely that adverse, population-level effects will occur in the exposed population."</i> The BERA did not assess risk to individual fish based on body burdens but rather used composites, and then averaged composites of different sized fish. The Trustees request the risk evaluation be performed based on individual samples and not composite samples.	Composite sampling reduces data variability and provides a reasonably unbiased estimate of mean contaminant concentrations, particularly when general homogeneity among composited materials are understood (e.g., compositing fish by species).

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
83	Trustees	Exposure Assessment	The BERA assumes risk through direct contact with surface water to be a major route of exposure. Dioxins and PCBs are hydrophobic, and aquatic exposure comes through their diet of contaminated prey items, not respiration. The BERA lists ingestion of contaminated dietary items by aquatic invertebrates and fish to be minor. This assumption is not supported and inappropriate. The BERA later states ingestion may be more important for contaminants that bioaccumulate (i.e., dioxins and PCBs), but specific ingestion exposure routes are not evaluated because “oral toxicity values for fish are not generally available.” The same applies for benthic macro-invertebrates (BMIs), but the BERA only considers “some” contribution from ingestion. The highest contaminant concentrations in fish tissue occurred at the St. Regis location. However, Clark Fork River (CFR) sediment samples were collected adjacent to or immediately downstream of the Site, within an area that is only 1/6th of the distance toward where the highest contaminant concentrations were found in fish tissue. At a minimum, sediment sampling, or preferably sediment sampling in addition to sampling methods specifically targeted to quantify hydrophobic chemicals that pose a risk at low levels (see general comments) should be used. Moreover, since elevated concentrations of Site-related contaminants were found at St. Regis, it follows that additional sampling should occur downstream of the Site and potentially downstream of St. Regis. Currently, resident downstream of St. Regis are unaware of any potential risk from the Site.	The characterization of the exposure routes in the BERA will be redefined as the exposure routes quantified in the BERA as opposed to the “major” exposure routes. Further, ingestion of contaminated dietary items by aquatic invertebrates and fish is not quantified due to limited toxicity information. The BERA evaluated potential risks within the current study area. Nature and extent of contamination is outside the scope of the risk assessment, but is being considered within the RI.
84	Trustees	Background Evaluation	The BERA used locations immediately upstream of the Site as background data sets for surface water, sediment, and soils. This is not appropriate since there is likely airborne deposition from the former mill (see general comments). Additionally, benthic macro-invertebrates (BMIs) move upstream to complete life cycle, and fish migrate upstream. These locations are then combined with highly variable fish samples to inaccurately conclude (with statistical analyses) that upstream locations are no different than downstream locations. The use of background data for making site comparisons for surface water, soil, and sediment should be clarified in section 3.3 and Appendix A of the BERA. The BERA (page 16) indicates that refinements to the screening assessment were made using additional data to compare Site concentrations to background concentrations, and Section 3.3 on page 24 describes the comparison to background data. The section, as well as section 2.1 to 2.3 in Appendix A, indicates surface water and sediment data were compared to data collected by the U.S. Geological Survey (USGS) in 2015-2020, and to background soil data collected by the Montana Department of Environmental Quality. However, on-site reference samples such as upstream samples were also collected, and it is unclear if both background and reference data were used for the comparison to Site data, or if the upstream data listed in the tables on pages 27 and 29 include the USGS data. More information should be added regarding how the on-site reference samples were used to compare Site concentrations, and data quality assessment following guidance such as EPA (2000b) should be conducted to determine if the data variation was sufficient to	Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			characterize background. If data do not meet data quality assessment rules, then additional sampling efforts should be conducted to make comparisons and decisions for the Site.	
85	Trustees	Background Evaluation	For surface water, the creek data as listed in the table on page 27 of the BERA show results from a Site vs upstream comparison, but there is no comparison to USGS background data presented for these creeks. Also, though too few samples were collected to make a statistical comparison on the individual locations as noted in the text, the section indicates that “When the data from the two creeks are combined and compared in ProUCL using the approach described above, the concentrations of all of the COPECs were determined to be statistically equal within the site boundary as compared to the upstream of the Site boundary.” This statistical evaluation appears to be in direct conflict with ProUCL Technical Guidance (EPA 2015), which specifically states on page 17 that “Decisions based upon statistics computed using discrete data sets of small sizes (e.g., < 6) cannot be considered reliable enough to make decisions that affect human health and the environment. For example, a background data set of size < 6 is not large enough to characterize a background population, compute BTV [background threshold value] estimates, or to perform background versus site comparisons.” In fact, the ProUCL guidance (pages 24 to 27) also indicates that for comparing two populations or calculating confidence limits, at least 10 samples (and preferably much more) should generally be collected from each population. We recommend the data for the creeks be just compared graphically, or with box plots or quantile-quantile plots if data are sufficient for plotting.	Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment.
86	Trustees	Background Evaluation	Similar to the creek surface water, sediment from the Site appears to be compared only to reference upstream data (see table in BERA on page 29). Again, very few samples were collected for statistical comparisons, so data were combined from the two creeks. However, even after combining samples, the Site creek data only met minimum preferred sample size requirements and the upstream reference data lacked sufficient number of samples (≤ 3) for comparison. For the same reasons stated above for surface water, we recommend the sediment data for the creeks be just compared graphically, or with box plots or quantile-quantile plots if data are sufficient for plotting.	Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
				additional information to risk managers to better characterize the results presented in the risk assessment.
87	Trustees	Background Evaluation	For surface soil, comparisons of Site to background data were made using a variety of soil collection efforts from 2014 to 2016. These collection efforts had different objectives, different surface soil profiles characterized (such as from 0 to 2 inches in some cases and 0 to 6 inches in others), and different sampling approaches (e.g., using five-point composites in earlier years and 20- to 30-point composites in 2016). In addition, samples collected in 2014 and 2015 were biased samples, as stated in Newfields (2017) on page 6 “A biased sampling approach was used whereby discrete samples were collected from locations most likely to be contaminated from former operations.” EPA (2002) notes the limitations of a biased (or judgmental) sampling approaches and states that “Judgmental sampling does not allow the level of confidence (uncertainty) of the investigation to be accurately quantified. In addition, judgmental sampling limits the statistical inferences that can be made to the units actually analyzed, and extrapolation from those units to the overall population from which the units were collected is subject to unknown selection bias.” Comparisons should not be made from data combined over these years for comparison to background, nor should statistical comparisons be used with biased or judgmental data. For soil collected in 2016, a non-biased, systematic grid sampling approach was used in the collection of 20- to 30-point composite samples (Newfields 2017). This approach appears similar to samples collected using incremental sampling methodology (ISM), although ISM incorporates a minimum of 30 increments collected per sample based on a grid, has strict mass and sample processing and subsampling requirements (typically involving grinding of soil), and strict quality control measures (ITRC 2021). These requirements help control sample heterogeneity, reduce variability, and provide more reproducible results that can be applied to background calculations and for making other comparisons. The sampling process in 2016 did not follow ISM protocols and it remains unknown if the results are reproducible and defensible. For example, the 2016 sampling did not include grinding of samples to reduce sampling error, but rather the Newfields (2017) Work Plan reports on page 15 that “The material will be mixed in alternating clockwise and counterclockwise directions with a decontaminated stainless steel trowel for at least 2 minutes.” Mixing in this fashion separates particles by grain size and can increase sampling error. Also, subsampling by the laboratory for the analysis is not documented, and the subsampling process is critical to reducing sampling error (Gerlach and Nocerino 2003). In addition, field triplicate samples are necessary for collection in at least one decision unit in ISM to ensure that sampling has controlled variation and data are replicable, yet it is unclear how duplicate samples were collected in the 2016 sampling and if they truly replicate the field collection process. For the surface soil sampled in the BERA, a data assessment should be conducted to ensure variability is controlled and data	Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA 2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment. The 20-point/30-point composite samples were collected to provide an estimate of the mean concentration over an identified exposure area (e.g., grid). Discrete samples (5-point composites collected over a 1-m2 area) were collected under a biased sampling strategy to target areas of most concern. Due to low variability within the current data set, EPA is not currently recommending additional soil sampling to support the risk assessments beyond that which was carried out in the RI/FS Addendum 7 (2017) in the fall of 2017.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			are sufficient to make statistical comparisons (although statistical comparison should not be made for biased samples collected in 2014 and 2015). For 2016 samples, duplicate field results should be evaluated and displayed prominently in the BERA alongside the field data so readers can discern if data variability is within sufficient limits to base decisions with confidence, and a more complete description of the sampling and subsampling process is needed to better understand the limitations and uncertainty of any comparisons made using these data. If data do not meet data quality assessment rules, then additional sampling efforts should be conducted to make comparisons and decisions for the Site.	
88	J Harrington	Site Characterization	I grew up here in Western Montana, surrounded by industry who have profited-handsomely from our resources but are not held responsible for the environmental damage they leave behind. From the first meeting with the EPA in the Frenchtown firehall in 2017, the public and Natural Resource Trustees have asked that action be taken that are protective of human health and the environment, the genesis of which is a thorough and complete characterization of the site. I do not feel confident that the actions taken so far have captured the decades of environmental impact that mill activities left behind. Transparency and accountability are lacking in this process, which leads to social and environmental injustice. EPA has been sharing limited and redundant information for years, without addressing data gaps and unanswered questions. It has been asked multiple times over the years why the 2011 URS and 2013 HRN reports (EPA, 2013) have not been incorporated into the sampling plans, nor in the draft HHRA or BERA, without an adequate answer. The presentations of the HHRA, BERA, and the long-awaited Ground Water Site Conceptual Model were superficial, and additionally, the continued use of statements like, “only 6% of the pulp was bleached” is concerning. It has been asked again and again that the Agency report actual figures, like, 1,600,000 tons of bleached pulp was produced, but reports and presentations by the EPA on the HHRA and BERA regurgitate the, “only 6%” figure without providing context. If a stratified soil sampling plan was in place, the deeper soil horizons may reveal different concentrations than the samples taken so far, especially in the ponds holding wastewater sludge from the bleaching plant operations. On page 4 of the EIS report it states that at the time, 14 years after the mill began bleaching pulp, there were still no EPA regulations for the bleached pulp industry. The volume of wastewater from the bleaching plant alone, especially before the EPA issued any regulations for Bleaching plants in the 1960’s and 1970’s, should trigger a concentrated sampling event of the lower horizons in the settling and sludge ponds associated with the Bleach plant discharge. On page 27 of the 1974 EIS: <i>An average of 159 tons (based on 350 operating days) of bleached product was produced in 1973 by HW. HW reports 30,000 gallons of water are required per</i>	Comment noted.. Data collected prior to 2014 to support the Preliminary Assessment/Site Investigation are not used in the BERA due to issues raised with data quality. EPA used those data to inform the sampling plan and focused collection on those areas identified as having potentially elevated contamination. The sampling conducted as part of RI activities are more recent, were collected in accordance with the EPA approved QAPP, and have been evaluated and determined to meet the needs of the risk assessment. The RI is still underway and additional sampling to support RI/FS activities will occur as deemed appropriate. Risk assessments evaluate current and hypothetical future exposure scenarios, and sampling to support the risk assessments has been determined to be adequate. The objective of the multimedia sampling conducted as part of the Remedial Investigation is to capture the likely contaminants in the data set and determine whether they pose an unacceptable risk to human health or the environment, regardless of the quantities historically used or produced. Remedial alternatives will be

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			<i>ton of bleached product. The discharge from the bleach plant would, therefore, average 4.8 million gallons per day (MGD). The total contaminated waste flow (not including the uncontaminated cooling water which is discharged separately) averaged 15.8 mgd (MDHES 1974).</i> The report also identifies that the mill's bleaching plant used 380 tons of chlorine monthly in 1973. It is important in attempting to understand the potential and real harm to human health and the environment that the study of the site include all areas that have been impacted by site activities.	evaluated in the FS, and damage is evaluated within the NRDA process.
89	J Harrington	Exposure Assessment	Before the mill broke ground, two Missoula-area scientists surveyed the river to understand the species that occupied the basin. Bottom samples were taken two miles downstream from the effluent outlet where there were bottom organisms identified. Roughly 50%, or 765 were types considered to be intolerant of pollution (Whitney and Spindler 1959). In July of 1958, seven months after the mill started operating, the percent of sensitive organisms had dropped to 17% at the same station (Spindler 1959). One of these scientists, John C. Spindler, warned in 1959 that: <i>It has frequently been found that the black liquor wastes from Kraft or sulphate process paper mills are toxic to fish and bottom invertebrates of streams (Whitney and Spindler 1959).</i> The BERA does not look at actual data about the site and its' impact to ecological receptors. The plant species that cover the grounds are invasive plants that can tolerate toxins in the soil and water. Where are the plants that are sensitive to pollutants in the environment? Why not sample the biota on site to see what is actually on site instead of using assumptions and models. It is important to know if there are species missing from the area because of potential impacts of pollution?	Comment noted. Risk assessments evaluate current and hypothetical future exposure scenarios. To that end, site-specific media (soil, surface water, sediment fish) were sampled and used to estimate potential risk. Damage is evaluated within the NRDA process.
90	J Harrington	Aerial Deposition	The argument that there has been no impact from aerial deposition during 53 years of operation is unsupported. Why would the mill need air pollution permits if what was being discharged over the years was not a hazard to human health and the environment? In the 1974 EIS for the proposed expansion of the mill, on page 1, it states: <i>HW has applied to the AQB for a permit to expand the capacity of its pulp and paper mill from 1150 tons per day (TPD) of paper products to 1850 TPD. Operation of the expanded plant would affect air pollution from the mill, and therefore a permit is required under regulations appurtenant to the 1967 Clean Air Act of Montana. The act and subsequent regulations say there shall be no "installation, alteration, or use of any machine, equipment, device or other article which . . . may cause or contribute to air pollution ..." unless a permit is issued (MDHES 1974).</i>	Comment noted. EPA is already considering aerial deposition within the risk assessment CSMs and the broader RI. EPA recently released fact sheets on aerial deposition, dioxins/furans, and PCBs at the site (EPA, 2021a, b, c) that speaks to these issues.
91	J Harrington	Background Evaluation	My comments on the HHRA in 2018 shared the research conducted by Paul Woods Bartlett, Barry Commoner, Kim Couchot, Brian Bush, Holger Eisl and Paul Cooney on established environmental impacts involving air pollution from industry (Bartlett et al. 2000). The fact that the background areas lie within reach of aerial deposition of the mill.	Although conducting a comparison to background concentrations was included in the BERA Work Plan as part of the COPEC refinement (EPA

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			The background sites are not far enough away from the potential impact area of the mill to be defensible. New background sites need to be chosen and sampled. On page 19 of the EIS the sources of pollution are described: <i>Nearly every item of process equipment in a pulp and paper mill can contribute to pollution of air and water. Air pollutants generated in the pulping process are many and varied. Gases and particulates are emitted as a variety of compounds. Many of the gaseous emissions are odorous, among them hydrogen sulfide (H₂S), methyl mercaptan, (CH₃SH), dimethyl sulfide (CH₃SCH₃) and dimethyl disulfide (CH₃SSCH₃). Other gaseous pollutants emitted are sulfur dioxide (SO₂) and carbon monoxide (CO). Particulate emissions include wood fiber, sodium compounds from the recovery furnaces, fly ash, and calcium compounds. Table 1 summarizes the various pollutants and sources (MDHES 1974).</i>	2018), upon further consideration of existing guidance it was determined that comparison with background levels should not be used to remove contaminants of concern from further evaluation owing to the need to fully characterize site risks (EPA 2001). Comparison to background is not a toxicological rationale for eliminating COPECs from further consideration. Comparisons of site concentrations to background concentrations should still be considered within the RI outside of the risk assessment to provide additional information to risk managers to better characterize the results presented in the risk assessment.
92	CAG	Site Characterization	Identification of the concentrations and distribution of contaminants across the Mill Site is most accurately done by on-site sampling. Rather than rigorously measuring for onsite contamination, the Draft Human Health Risk Assessment discusses Exposure Pathways of Chief Concern and indirect measurements. The soil sampling conducted is insufficient. Soil sampling density was far too low to adequately determine the presence or absence of contamination across the 3000-acre site (see figure 4.3 attached to original comment letter: “Surface Soil Sampling Locations in OU2: Smurfit-Stone/Frenchtown Mill Site, Montana”). Operational Unit 2 is approximately 265 acres in size. This to one sample taken for every 7.4 acres. Concerns have been expressed as to the adequacy of having such large areas represented by so few samples. Representing large areas with minimal sample locations may not represent the potential risks. No new soil samples since 2017.	EPA believes that soil sampling conducted to date adequately characterizes surface soil conditions at the site for risk assessment purposes. The focus of a risk assessment is on estimating exposures that a receptor may reasonably be exposed to within a given exposure area over their entire exposure duration. Composite sampling reduces data variability and provides a reasonably unbiased estimate of mean contaminant concentrations in a specified decision unit (ITRC 2021). The 20-point/30-point composite samples were collected to provide an estimate of the mean concentration over an identified exposure area (e.g., grid). Discrete samples (5-point composites collected over a 1-m² area) were collected under a biased sampling strategy to target areas of most concern. Due to low variability within the current data set, EPA is not currently

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
				recommending additional soil sampling to support the risk assessments.
93	CAG	Site Characterization	Soil sample data were collected in 2014, 2015 and 2016 as 5-point composite samples collected from locations across OU2 and OU3. Composite sampling would tend to mask samples with high concentrations and limit their use for statistical analyses and inference.	Composite sampling reduces data variability and provides a reasonably unbiased estimate of mean contaminant concentrations in a specified decision unit (ITRC 2021). The 20-point/30-point composite samples were collected to provide an estimate of the mean concentration over an identified exposure area (e.g., grid). Discrete samples (5-point composites collected over a 1-m ² area) were collected under a biased sampling strategy to target areas of most concern.
94	CAG	Site Characterization	A biased sampling approach (sampling locations selected based on site knowledge to identify areas of highest environmental concern) was used. Given that contamination may have been spread by post discharge surface flow as well as atmospheric deposition not using a sufficient number of randomized selections of sample sites across the Mill site limits any quantitative inference to identify the extent of contamination at the site. The number and spatial distribution of sites sampled should be related to the variability in the measurements.	Some site soil data was collected using biased sampling, but grid sampling was also used, which is randomized.
95	CAG	Site Characterization	PCBs, furans and dioxins strongly attach to soil particles, are very persistent, and can be transported via erosion and transport of contaminated soils. The BERA and HHRA soil samples were last taken in 2017, giving limited and potentially inaccurate assumptions about the current and future danger the Persistent Organic Pollutants pose to Human Health and the Environment.	EPA believes that soil sampling conducted to date adequately characterizes surface soil conditions at the site for risk assessment purposes. Site activities that could be expected to increase soil contamination have not occurred since the last round of soil sampling. EPA recently published fact sheets on dioxin/furans and PCBs (EPA, 2021a,b) that provide additional information and will be cited in the BERA.
96	CAG	Conceptual Site Model	GW Site Conceptual Model is flawed. It ignores the significant threat of erosion of contaminated materials into the Clark Fork River. It is incomplete, resulting in a questionably vetted HHRA and BERA. You are not getting a good picture of the risk without a good GWSCM.	The CSM presented in the BERA is specific to identifying ecological receptors and exposure pathways. This objective differs slightly from the fate and transport CSMs being developed in other site documents. EPA recently published an aerial deposition fact

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
				sheet (EPA 2021a), which will be cited in the BERA. In addition, the groundwater CSM (NewFields 2020) is currently being revised and will be cited in the BERA as well.
97	CAG	Site Characterization	The Channel Migration Zone and Avulsion hazards have not been accounted for.	Risk assessments are performed using available site data and presumed exposure scenarios to estimate the chronic risks posed by the contaminants present at the site. CERCLA risk assessments cannot and should not be used to estimate short-term risks from hypothetical future events such as flooding, since the extent of soil disturbance, the location of such erosion, and the lateral extent of deposition cannot be known ahead of time.
98	CAG	Site Characterization	The Landfills and ponds have not been sampled to groundwater depths. The Landfills do not have an adequate number of samples. HHRA and BERA for OU3 does not address the landfills and sludge ponds. Methane generation potential in the sludge ponds and landfills should be investigated in OU3.	Documentation regarding the waste management area (WMA) are located in DEQ, EPA and Missoula County files; the site was well regulated beginning in the early 1980s. Please review the PA/SI and the Missoula County letter of July 26, 2016. The PA/SI (and the appendices) conducted by EPA in 2011/2012 references documents about the older materials generated during Mill operation including concentrations and quantities. Additional work to better understand the WMA conducted during EPA's PA/SI work effort included soil boring within the WMA. One of the goals of the RI is to sample the groundwater to determine if the WMA is leaching contaminants to the GW. EPA is requiring additional monitoring in the future because some COPCs are impacting GW.

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
99	CAG	Site Characterization	The early years of production produced up to 30% bleached product. The total volume of bleached product is approximately 1,600,000 tons was produced during a period of 39 years. There is a need for dioxin data at lower horizons in the sludge ponds.	Additional data will not be collected prior to finalization of the BERA. Uncertainties associated with the available database are described in the uncertainty assessment.
100	CAG	Site Characterization	Exclusion of 104 Mill study (EPA 1990) and the 2011 URS data has not been justified and if the reasoning is not articulated, the data should be included in the HHRA and BERA.	Data collected prior to 2014 to support the Preliminary Assessment/Site Investigation are not used in the BERA due to issues raised with data quality. EPA used those data to inform the sampling plan and focused collection on those areas identified as having potentially elevated contamination. The sampling conducted as part of RI activities are more recent, were collected in accordance with the EPA approved QAPP, and have been evaluated and determined to meet the needs of the risk assessment.
101	CAG	Site Characterization	OU3 comprises of all potential locations in the Clark Fork River where hazardous substances may be located. Sampling at locations downriver from the former mill site is necessary to fully characterize the impact of the mill.	As noted in the BERA, the full nature and extent to which hazardous substances from site activities have come to be located in the CFR adjacent to and downstream of the Site has not been fully established but will be considered within the RI.
102	CAG	Site Characterization	Re: calculating risks to the berm and floodplain: Continued threat of high ground water influencing materials in the dumps and sludge ponds.	Risk assessments are performed using available site data and presumed exposure scenarios to estimate the chronic risks posed by the contaminants present at the site. CERCLA risk assessments cannot and should not be used to estimate short-term risks from hypothetical future events such as flooding or short-term rise in groundwater table.
103	CAG	Site Characterization	Groundwater is neither uniform nor homogenous and yet grab samples from surface water are being used to describe the health of the entire water system associated with the former mill.	The groundwater CSM (NewFields 2020) is currently being revised and will be cited in the BERA. Surface water data were collected consistent

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
				with DQOs for the site, and were evaluated in the BERA.
104	CAG	Site Characterization	It is unclear if the soil strata have been correctly profiled to make sure that each soil strata are represented.	The BERA evaluates potential exposures to ecological receptors which primarily occurs within the top 6 inches of soil. Additional soil characterization may occur as part of the RI.
105	CAG	Exposure Assessment	The HHRA does not address potential risks to receptors, nor does it address site-specific impacts to receptors, such as sensitive benthic macroinvertebrates. The HHRA has not considered site-specific species no longer present at the former mill or in the adjacent Clark Fork River. The use of modeling is not capturing the impact the mill has had on the ecosystem.	Risk assessments evaluate current and potential future risks to ecological receptors that are known to inhabit the site. They are not retroactive.
106	CAG	Flooding Concerns	Little attention was given to a low probability flood in excess of what is calculated to be a 100-year event would breach the berm and allow flood waters to scour large portions of the site. Climate change likely has and will continue to increase the risk of erratic weather and increase the probability of floods. What is estimated to be a 100-year flood at present may well become a much higher probability event in future years.	Risk assessments are performed using available site data and presumed exposure scenarios to estimate the chronic risks posed by the contaminants present at the site. CERCLA risk assessments cannot and should not be used to estimate short-term risks from hypothetical future events such as flooding, since the extent of soil disturbance, the location of such erosion, and the lateral extent of deposition cannot be known ahead of time.
107	CAG	General	It is impossible to quantify the risks without more intensive soil sampling. To date there is little data that would prove or disprove the sites impact to the Clark Fork River though State of Montana fish tissue data indicate an increase in Dioxins downstream of the site. These data were considered sufficient to issue a fish consumption advisory. It is always better to error on the side of caution when human health is at risk. Allocating resources to direct sampling of contaminants to determining the types, amounts, and locations of contaminants will give a much clearer and defensible picture of risks and identify areas needing treatment. In conclusion, the Frenchtown CAG is in support of actions that remove the sources of contamination instead of leaving contaminants for future generations to contend with. We are asking the EPA and MTDEQ for a commitment for a thorough Risk Assessment to protect human health and the environment now and for generations to come. A resident of the area has observed that the process is driving the data instead of the data driving the process. The desired result of the Risk Assessment is an accurate	Sampling conducted to support the risk assessment is not necessarily equivalent to sampling conducted to evaluate nature and extent. Evaluation of nature and extent is outside of the scope of the risk assessment which focuses on evaluation of the current study area. Further evaluation of nature and extent is being conducted as part of the RI. Risk assessments evaluate current and hypothetical future exposure scenarios, and sampling to support the risk assessments has been determined to be adequate. The objective of the multimedia sampling conducted as part

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			representation of the impacts that the site has had and continues to have on the environment. We do not feel that the process up to this point has done enough to characterize the influence and impact that 53 years of mill operations.	of the RI is to capture the likely contaminants in the data set and determine whether they pose an unacceptable risk to human health or the environment, regardless of the quantities historically used or produced. Remedial alternatives will be evaluated in the FS, and damage is evaluated within the NRDA process.
108	J Salazar	Site Characterization	...The studies EPA has conducted to date are inadequate to address fully and meaningfully the risks associated with this site. From inadequate stratification of sampling to small sample sizes to omission of some key chemicals of concern, no meaningful conclusions can be drawn concerning the levels of risk presented by this site. Upstream and downstream comparisons are invalid both because small sample size means there is no power of the test, and because aerial deposition and fish migration can move chemicals upstream.	EPA believes that sampling conducted to date adequately characterizes site conditions for risk assessment purposes to evaluate the current study area. The focus of a risk assessment is on estimating exposures that a receptor may reasonably be exposed to within a given exposure area over their entire exposure duration. EPA recently published fact sheets on aerial deposition, dioxins/furans, and PCBs at the Site (EPA, 2021a, b, c) that provides additional information and will be cited in the BERA.
109	J Salazar	Flooding Concerns	One of the greatest inadequacies of the study design is the omission of one of the largest sources of risk associated with this site – the risk of dike failure and downstream transport of the toxic chemicals in the historic floodplain. As long as the mill’s toxic waste remains in the Clark Fork River’s historic floodplain, these risk assessments must also address the potential impacts of a dike failure followed by the river washing that waste downstream. The river will eventually reclaim its floodplain. And climate scientists predict that climate change will continue to produce bigger and bigger floods, hence the chance of such a catastrophic failure is increasing and will be even greater than historic flooding would suggest. Obviously, it is difficult to predict where the waste will wind up once it is washed downstream – but that does not mean this catastrophic scenario should be ignored. These risk assessments must assess all the possible outcomes of leaving the waste where it is (where it is subject to being washed downstream), as well as assessing the much lower risks if those wastes are moved to a location less likely to be impacted by bigger floods	Risk assessments are performed using available site data and presumed exposure scenarios to estimate the chronic risks posed by the contaminants present at the site. CERCLA risk assessments cannot and should not be used to estimate short-term risks from hypothetical future events such as flooding, since the extent of soil disturbance, the location of such erosion, and the lateral extent of deposition cannot be known ahead of time. The on-site toxic waste, referred to as the ‘waste management area’ is not

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			and dike failure. Without such an assessment, these risk assessments are missing information critical to making decisions about needed remediation actions at this site. In my view, the EPA should move quickly to require the removal of the toxic waste in the historic floodplain at the Smurfit-Stone/Frenchtown Mill site, as it has in removing the contaminated ponds at the Columbia Falls Aluminum Company Superfund site. That EPA action was taken to protect the Flathead River from the eventual failure of those ponds. The much larger area of contaminated historic floodplain at the Smurfit Stone site presents an even greater risk to the Clark Fork River.	located within the CFR 100-year floodplain.
110	V Watson	Site Characterization	...The studies EPA has conducted to date are inadequate to address fully and meaningfully the risks associated with this site. From inadequate stratification of sampling to small sample sizes to omission of some key chemicals of concern, no meaningful conclusions can be drawn concerning the levels of risk presented by this site. Upstream and downstream comparisons are invalid both because small sample size means there is no power of the test, and because aerial deposition and fish migration can move chemicals upstream.	EPA believes that soil sampling conducted to date adequately characterizes surface soil conditions at the site for risk assessment purposes. The focus of a risk assessment is on estimating exposures that a receptor may reasonably be exposed to within a given exposure area over their entire exposure duration. Composite sampling reduces data variability and provides a reasonably unbiased estimate of mean contaminant concentrations in a specified decision unit (ITRC 2021). The 20-point/30-point composite samples were collected to provide an estimate of the mean concentration over an identified exposure area (e.g., grid). Discrete samples (5-point composites collected over a 1-m2 area) were collected under a biased sampling strategy to target areas of most concern. EPA recently published an aerial deposition fact sheet (EPA, 2021a) that provides additional information and will be cited in the BERA.
111	V Watson	Flooding Concerns	One of the greatest inadequacies of the study design is the omission of one of the largest sources of risk associated with this site – the risk of dike failure and downstream transport of the toxic chemicals in the historic floodplain. As long as the mill’s toxic waste remains in the Clark Fork River’s historic floodplain, these risk assessments must also address the potential impacts of a dike failure followed by the river washing that waste downstream. The river will eventually reclaim its	Risk assessments are performed using available site data and presumed exposure scenarios to estimate the chronic risks posed by the contaminants present at the site. CERCLA risk assessments cannot and should not be used to estimate short-

Response to Comments on the 2020 External Review Draft BERA for the Smurfit-Stone/Frenchtown Mill Site

Comment #	Commenter	Topic	Comment	Response
			floodplain. And climate scientists predict that climate change will continue to produce bigger and bigger floods, hence the chance of such a catastrophic failure is increasing and will be even greater than historic flooding would suggest. Given the evidence that water moves through this uncertified, simple gravel dike during 25 year flood events, the dike does not have to be overtopped to fail. The dike can simply liquefy and wash away. Obviously, it is difficult to predict where the waste will wind up once it is washed downstream – but that does not mean this catastrophic scenario should be ignored. These risk assessments must assess all the possible outcomes of leaving the waste where it is (where it is subject to being washed downstream), as well as assessing the much lower risks if those wastes are moved to a location less likely to be impacted by bigger floods and dike failure. Without such an assessment, these risk assessments are missing information critical to making decisions about needed remediation actions at this site. In my view, the EPA should move quickly to require the removal of the toxic waste in the historic floodplain at the Smurfit-Stone/Frenchtown Mill site, as it has in removing the contaminated ponds at the Columbia Falls Aluminum Company Superfund site. That EPA action was taken to protect the Flathead River from the eventual failure of those ponds. The much larger area of contaminated historic floodplain at the Smurfit Stone site presents an even greater risk to the Clark Fork River.	term risks from hypothetical future events such as flooding, since the extent of soil disturbance, the location of such erosion, and the lateral extent of deposition cannot be known ahead of time. The on-site toxic waste, referred to as the ‘waste management area’ is not located within the CFR 100-year floodplain.

Commenter Key:

CAG: Frenchtown Community Advisory Group

CFC: Clark Fork Coalition

CSKT: Confederated Salish and Kootenai Tribes of the Flathead Reservation

J Harrington: Jen Harrington

J Salazar: Jordan Salazar

MDEQ: Montana Department of Environmental Quality

Missoula Co Commissioners: Missoula County Commissioners

MVWQD: Missoula Valley Water Quality District

PRP: Potentially Responsible Party

Trustees: Montana Department of Justice Natural Resource Damage Program, Montana Fish Wildlife & Parks, Confederated Salish and Kootenai Tribes of the Flathead Reservation, U.S. Fish and Wildlife Service, U.S. Forest Service, and Kalispel Tribe of Indians

V Watson: Dr. Vicki Watson

WQAC: Missoula Public Health City-County Health Department Water Quality Advisory Council

References:

Allard, P., A. Fairbrother, B.K. Hope, R.N. Hull, M.S. Johnson, L. Kapustka, G. Mann, B. McDonald, and B.E. Sample. 2009. Recommendations for the development and application of wildlife toxicity reference values. *Integ. Environ. Assess. Manag.* 6(1):28–37.

Burgess, R.M., R. Lohmann, J.P. Schubauer-Berigan, P. Reitsma, M.M. Perron, L. Lefkovitz, and M.G. Cantwell. 2015. Application of passive sampling for measuring dissolved concentrations of organic contaminants in the water column at three marine superfund sites. *Environmental Toxicology and Chemistry* 34(8):1720–1733.

CCME. 2018. Draft. Scientific Criteria Document for the development of the Canadian water quality guidelines for the protection of aquatic life. Manganese. Canadian Council of Ministers of the Environment. August.

Charlestra, L., D.L. Courtemanch, A. Amirbahman, and H. Patterson. 2008. Semipermeable membrane device (SPMD) for monitoring PCDD and PCDF levels from a paper mill effluent in the Androscoggin River, Maine, USA. *Chemosphere* 72:1171–1180.

Cohen, M.D. et al., 2002. Modeling the Atmospheric Transport and Deposition of PCDD/F to the Great Lakes. *Environmental Science & Technology*, 36 (22), pg 4831-4845.

Cohen-Barnhouse, A.M., M.J. Zwiernik, J.E. Link, S.D. Fitzgerald, S.W. Kennedy, J.P. Giesy, S. Wiseman, P.D. Jones, J.L. Newsted, D. Kay, and S.J. Bursian. 2011. Developmental and posthatch effects of in ovo exposure to 2,3,7,8-TCDD, 2,3,4,7,8-PECDF, and 2,3,7,8-TCDF in Japanese quail (*Coturnix japonica*), common pheasant (*Phasianus colchicus*), and white leghorn chicken (*Gallus gallus domesticus*) embryos. *Environ. Toxicol. Chem.* 30(7):1659–1668.

Dinkins, S.A. and J.P. Heath. 1998. Quantification of Dioxin Concentrations in the Ohio River Using High Volume Water Sampling. Presented at the Proceedings of the National Water Quality Monitoring Council National Monitoring Conference, Reno, NV. July 7–9.

Efroymson, R.A., M.E. Will, G.W. Suter II, and A.C. Wooten. 1997a. Toxicological Benchmarks for Screening Contaminants of Potential Concern for Effects on Terrestrial Plants: 1997 Revision. Prepared for Oak Ridge National Laboratory. November.

Efroymson, R.A., M.E. Will, G.W. Suter II, and A.C. Wooten. 1997b. Toxicological Benchmarks for Contaminants of Potential Concern for Effects on Soil and Litter Invertebrates and Heterotrophic Process: 1997 Revision. Prepared for Oak Ridge National Laboratory. November.

EPA. 1989. Risk Assessment Guidance for Superfund. Volume I:1, Human Health Evaluation Manual (Part A: Baseline Risk Assessment). Interim Final. EPA/540/1-89/002. Office of Emergency and Remedial Response. Washington, D.C. December.

EPA. 1990. U.S. EPA/Paper Industry Cooperative Dioxin Study: “The 104 Mill Study” Statistical Findings and Analyses. U.S. Environmental Protection Agency, Washington, DC. July 13.

EPA. 1992. Guidance for Data Usability in Risk Assessment (Part A)., PB92-963356. Office of Emergency and Remedial Response., Washington, DC.

EPA. 1993. Barium Sulfate; Toxic Chemical Release Reporting Community Right-to-Know. 40 CFR Part 372, Federal Register. 58, No. 111, 32622-32628. U.S. Environmental Protection Agency. June 10.

EPA. 1997a. Ecological Risk Assessment Guidance for Superfund, Process for Designing and Conducting Ecological Risk Assessments. EPA/540-R-97-006. U.S. Environmental Protection Agency, Office of Emergency and Remedial Response. June 5.

EPA. 1997b. Locating and Estimating Air Emissions from Source of Dioxins and Furans. U.S. Environmental Protection Agency, Research Triangle Park, North Carolina. May.

EPA. 1998. Guidelines for Ecological Risk Assessment. EPA/630/R-95/002F. U.S. Environmental Protection Agency. May.

EPA. 1999. Screening Level Ecological Risk Assessment Protocol for Hazardous Waste Combustion Facilities. Peer Review Draft. U.S. Environmental Protection Agency. August.

EPA. 2000a. Bioaccumulation Testing and Interpretation for the Purpose of Sediment Quality Assessment – Status and Needs. EPA 823-R-00-001. U.S. Environmental Protection Agency.

EPA. 2000b. Guidance for data quality assessment: Practical methods for data analysis. EPA QA/G-9, QA00 Update. Final Report EPA/600/R-96/084. U.S. Environmental Protection Agency, Washington, DC.

EPA. 2001. The Role of Screening-Level Risk Assessments and Refining Contaminants of Concern in Baseline Ecological Risk Assessments. ECO Update EPA540/F-01/014. June.

EPA. 2003a. Generic Ecological Assessment Endpoints (GEAE) for Ecological Risk Assessment. EPA/630/P-02/004F. U.S. Environmental Protection Agency, Risk Assessment Forum, Washington, DC. October.

EPA. 2003b. Ecological Soil Screening Levels for Aluminum. Interim Final. OSWER Directive 9285.7-60. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, DC. 808 pp. November.

- EPA, 2004. Final Study Design and DQO Process (Work Plan), Ecological Risk Assessment. Davis Timber Superfund Site, Lamar County, Mississippi.
- EPA. 2005. Procedures for the derivation of equilibrium partitioning sediment benchmarks (ESBs) for the protection of benthic organisms: Metal mixtures (cadmium, copper, lead, nickel, silver, and zinc). EPA-600-R-02-11. U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC.
- EPA. 2006. Final Report: Pulp, Paper, and Paperboard Detailed Study. EPA-821-R-06-016.
- EPA. 2012. Guidelines for Using Passive Samplers to Monitor Organic Contaminants at Superfund Sediment Sites. OSWER Directive 9200.1-110 FS. U.S. Environmental Protection Agency. December.
- EPA, 2013. Hazard Ranking Systems (HRS) Smurfit-Stone Mill. U.S. Environmental Protection Agency, Denver, CO. May.
- EPA. 2015. ProUCL version 5.1 technical guide: Statistical software for environmental applications for data sets with and without nondetect observations. Final Report EPA/600/R-07/041 prepared by A. Singh and A.K. Singh, Lockheed Martin/SERAS, Edison, New Jersey for F. Barnett, ORD Site Characterization and Monitoring Technical Support Center, U.S. Environmental Protection Agency, Atlanta, Georgia. U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC.
- EPA, 2018. Baseline Ecological Risk Assessment Work Plan for Operable Units 2 & 3 of the Smurfit-Stone/Frenchtown Mill Site Located in Missoula County, Montana. August.
- EPA, 2019. Toxics Release Inventory. Guidance for Reporting the Dioxin and Dioxin-Like Compounds Category EPA 745-B-19-004.
- EPA. 2021a. Smurfit Stone Mill Site Air Deposition Fact Sheet. United States Environmental Protection Agency. Region 8. March.
- EPA, 2021b. Smurfit Stone Site Dioxins/Furans Fact Sheet. United States Environmental Protection Agency, Region 8. June.
- EPA, 2021c. Smurfit Stone Site Polychlorinated Biphenyls (PCBs) Fact Sheet. United States Environmental Protection Agency, Region 8. August.
- Fagervold, S.K., Y. Chai, J.W. Davis, M. Wilken, G. Cornelissen, and U. Ghosh. 2010. Bioaccumulation of polychlorinated dibenzo-p-dioxins/dibenzofurans in *E. fetida* from floodplain soils and the effect of activated carbon amendment. *Environ. Sci. Technol.* 44(14):5546–5552.
- Farmahin, R., G.E. Manning, D. Crump, D. Wu, L.J. Mundy, S.P. Jones, M.E. Hahn, S.I. Karchner, J.P. Giesey, S.J. Bursian, M.J. Zwiernik, T.B. Fredericks, and S.W. Kennedy. 2013. Amino acid sequence of the ligand-binding domain of the aryl hydrocarbon receptor 1 predicts sensitivity of wild birds to effects of dioxin-like compounds. *Toxicol. Sci.* 131(1):139–152. <https://doi.org/10.1093/toxsci/kfs259>
- Fuchsman, P.C., L.E. Brown, M.H. Henning, M.J. Bock, and V.S. Magar. 2016. Toxicity reference values for methylmercury effects on avian reproduction: Critical review and analysis. *Environ. Toxicol. Chem.* 36(2):294–319. <https://doi.org/10.1002/etc.3606>
- Galbraith. 2004. Tittabawassee River Floodplain Screening-Level Ecological Risk Assessment- Polychlorinated Dibenzo-p-Dioxins, Polychlorinated Dibenzofurans. Submitted to the Michigan Department of Environmental Quality, Remediation and Redevelopment Division, Saginaw Bay District Office. Submitted by Galbraith Environmental Sciences LLC, Newfane, Vermont. April.

- Gerlach, R.W., and J.M. Nocerino. 2003. Guidance for obtaining representative laboratory analytical subsamples from particulate laboratory samples. U.S. Environmental Protection Agency, Office of Research and Development, Cincinnati, Ohio. 134 pp.
- Golding, L.A., K. McKnight, M. Binet, M. Adamn, and S.C. Apte. 2018. Toxicity of dissolved and precipitated forms of barium to a freshwater alga (*Chlorella* sp. 12) and water flea (*Ceriodaphnia dubia*). *Environ. Toxicol. Chem.* 37(6):1632–1642.
- Great Lakes Basin. 2015. Tier II values calculated using the Great Lakes Basin methodologies at 327 IAC 2-1.5-12-14. Available at: https://www.in.gov/idem/cleanwater/files/wqs_glb_wqs_tier_2.pdf.
- Heinz, G.H. 1979. Methyl mercury: reproductive and behavioral effects on three generations of mallard ducks. *J. Wildlife Manage.* 53:418–428.
- Helsel, D.R. 2005. More than obvious: Better methods for interpreting nondetect data. *Environmental Science and Technology* 39(20):419A–423A.
- Hofmann, T. and A. Wendelborn. 2007. Colloid facilitated transport of polychlorinated dibenzop-dioxins and dibenzofurans (PCDD/Fs) to the groundwater at Ma Da Area, Vietnam. *Environmental Science and Pollution Research* 14:223–224.
- Hydrometrics. 2013. Background Concentrations of Inorganic Constituents in Montana Surface Soils. Prepared for Montana Department of Environmental Quality. September.
- Imbrigiotta, T.E. and P.T. Harte. 2020. Passive Sampling of Groundwater Wells for Determination of Water Chemistry. Techniques and Methods. Chapter 8, Section D, Book 1. U.S. Geological Survey. Available: <https://doi.org/10.3133/tm1d8>. Accessed February 10, 2021.
- Integral. 2020. Smurfit/Frenchtown Mill, Baseline Ecological Risk Assessment. RI/FS Work Plan Addendum 9: Data Interpretive Report. January.
- ITRC. 2007. Protocol for Use of Five Passive Samplers to Sample for a Variety of Contaminants in Groundwater. DSP-5. Interstate Technology & Regulatory Council, Diffusion/Passive Sampler Team, Washington, DC.
- ITRC.2021. Incremental sampling methodology update. Interstate Technology & Regulatory Council, Incremental Sampling Methodology Team. Washington, D.C. <https://ism-2.itrcweb.org/>.
- Kandemir, Ş. and N. Polat. 2007. Seasonal variation of total lipid and total fatty acid in muscle and liver of rainbow trout (*Oncorhynchus mykiss* W., 1792) reared in Derbent Dam Lake. *Turkish Journal of Fisheries and Aquatic Sciences* 7:27–31.
- Kay, D., A. Blankenship, M. Zwiernik, J.L. Newsted, P.D. Jones, and J.P. Giesy. 2005. PCDDs and PCDFs in Aquatic and Terrestrial Food Webs of the Tittabawassee River, Michigan, USA. *Organohalogen Compounds* 67:2130-2133.
- Lemieux, P.M., C.W. Lee, J.D. Kilgroe, and J.V. Ryan. 1999. Emissions of Polychlorinated Biphenyls as Products of Incomplete Combustion from Incinerators. EPA 600/A-99/051. Presented at the International Conference on Incineration and Thermal Treatment Technologies.
- Lohmann, R. 2016. Development of a Passive Multisampling Method to Measure Dioxins/Furans and Other Contaminant Bioavailability in Aquatic Sediments. SERDP Project ER-2538. November.

- Long, E.R., and L.G. Morgan. 1991. The Potential for Biological Effects of Sediment-Sorbed Contaminants Tested in the National Status and Trends Program. NOAA Technical Memorandum NOS OMA 52. National Oceanic and Atmospheric Administration.
- Mactec. 2004. Interim Final Baseline Ecological Risk Assessment. Centredale Manor Restoration Project Superfund Site, Providence, Rhode Island. Submitted to U.S. Army Corps of Engineers, New England Division, under contract to Battelle. Contract No. DACW33-01-D-004. Mactec Engineering and Consulting, Inc. September.
- Nosek, J.A., S.R. Craven, J.R. Sullivan, S.S. Hurley, and R.E. Peterson. 1992. Toxicity and reproductive effects of 2,3,7,8-tetrachlorodibenzo-p-dioxin in ring-necked pheasant hens. *J. Toxicol. Environ. Health* 35:187–198.
- MDEQ. 1995. Montana Department of Environmental Quality. Permit Determination for Stone Container Permit #2589-05. Helena, MT.
- MDEQ. 2011. Montana Department of Environmental Quality. Montana Dioxin Background Investigation Report. Helena, MT.
- Miller, M.A. 1993. Maternal transfer of organochlorine compounds in salmonines to their eggs. *Canadian Journal of Fisheries and Aquatic Sciences* 50:1405–1413.
- NewFields. 2017. Addendum No. 7 to Remedial Investigation Work Plan. Supplemental Soil Sampling in OU2 and OU3. Former Frenchtown Mill, Missoula, County Montana. October.
- NewFields, 2020. Groundwater Conceptual Site Model. Former Smurfit-Stone/Frenchtown Mill Site. Missoula County, Montana. Draft Version 3 – June.
- Niimi, A. 1983. Biological and toxicological effects of environmental contaminants in fish and their eggs. *Canadian Journal of Fisheries and Aquatic Sciences* 40:306–312.
- Nosek et al. 1992. Toxicity and Reproductive Effects of 2,3,7,8-tetrachlorodibenzo-p-dioxin in ring-necked pheasant hens. *J. Toxicol. Environ. Health* 35(3): 187-198.
- Nosek, J.A., S.R. Craven, J.R. Sullivan, A. Gendron-Fitzpatrick, and R.E. Peterson. 1993. Embryotoxicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin in the ring-necked pheasant. *Environ. Toxicol. Chem.* 12(7):1215–1222.
- Peters, A., S. Lofts, G. Merrington, B. Brown, W. Stubblefield, and K. Harlow. 2011. Development of biotic ligand models for chronic manganese toxicity to fish, invertebrates, and algae. *Environ. Toxicol. Chem.* 30(11):2407–2415.
- Powell, M.R. 1997. Control of Dioxins (and other Organochlorines) from the Pulp and Paper Industry under the Clean Water Act and Lead in Soil at Superfund Mining Sites: Two Case Studies in EPA's Use of Science.
- Russell, R.W., F.A. Gobas, and G.D. Haffner. 1999. Maternal transfer and in ovo exposure of organochlorines in oviparous organisms: a model and field verification. *Environmental Science and Technology* 33:416–420.
- Sample, B.E., D.M. Opresko, and G.W. Suter II. 1996. Toxicological Benchmarks for Wildlife: 1996 Revision. ES/ER/TM-86-R3. U.S. Department of Energy, Office of Environmental Management.
- Sample, B.E., J.J. Beauchamp, R.A. Efroymson, G.W. Suter, II, and T.L. Ashwood. 1998. Development and Validation of Bioaccumulation Models for Earthworms. ES/ER/TM-220. ORNL, Oak Ridge TN. 93 pp.

- Steevens, J.A., M.R. Reiss, and A.V. Pawlisz. 2005. A methodology for deriving tissue residue benchmarks for aquatic biota: A case study for fish exposed to 2, 3, 7, 8-tetrachlorodibenzo-pdioxin and equivalents. *Integrated Environmental Assessment and Management: An International Journal* 1:142–151.
- Suter, G.W., and C.L. Tsao. 1996. Toxicological Benchmarks for Screening Potential Contaminants of Concern for Effects on Aquatic Biota: 1996 Revision. ES-ER-TM-96-R2 Risk Assessment Program, Health Science Research Division, Oak Ridge, TN. 151 pp.
- Tooke, D.L. 2006. Sulfur Transport and Fate in a Pulp Mill Wastewater Treatment System [Dissertation]. University of Montana.
<https://scholarworks.umt.edu/etd/9617/>
- Urban, J.D. et al. 2013. A Review of Background Dioxin concentrations in Urban/Suburban and Rural Soils across the United States: Implications for Site Assessment and the Establishment of Soil Cleanup Levels. *Science of the Total Environment*.
- Whitney and Spindler, 1959. Effects of Kraft Paper Wastes on a Montana Stream. (attached to original comment letter)