

Thank you for this opportunity to improve our paper, "New approaches to understand mercury in trees: radial and longitudinal patterns of mercury in tree rings and genetic control of mercury in maple sap," submitted to Water, Air, & Soil Pollution.

Our responses to these very helpful reviews (*copied in italics in black*) are detailed below in blue. Quoted material from the text is in black with additions in red. Our goal is to make your review of our revisions as easy as possible.

#### COMMENTS FOR THE AUTHOR:

*Reviewer #1: Yanai et al. describe Hg distribution in trees, providing further evidence for foliar uptake as the primary source of wood Hg. The paper is well-written and the methods and conclusions are scientifically sound. The two studies described in the paper are limited in scope, but appear to have just "enough" new data/information to make publication worthwhile.*

Thank you for your brief but positive evaluation.

*Reviewer #2: This short paper has two novel aspects: (1) it is the first to my knowledge to examine mercury in the tree bole at different heights above ground; and (2) it demonstrates that one source of variability in Hg concentrations within tree tissues is genetic. Many studies have demonstrated that Hg uptake by trees is dominated by foliar uptake of atmospheric Hg rather than root uptake of soil Hg, but the finding in this paper that Hg in bole wood decreases from the top of the canopy to the bottom, helps to reinforce that finding. A shortcoming of the paper is that this finding is based on only four trees, representing the three major Northern Hardwood deciduous species and a conifer. That they all display a similar pattern with bole height lends legitimacy to treating the four trees as replicates. In my view, the novelty and importance of these findings outweigh this weakness.*

Some of our co-authors doubted whether this would be publishable, so it's good that the reviewer supports its publication.

*The limited sample size is more problematic in interpreting the temporal Hg patterns in the rings of the individual trees, as the data are noisy. The three deciduous trees follow the general timeline of atmospheric Hg concentrations, albeit an early peak in 1960, while the spruce shows no apparent trend (see comment for lines 191-192 below). The sample size for the second point about genetic variability was adequate.*

Our description of the results was not clear on this point, but we have improved the text, as detailed below, in response to this reviewer's comment on lines 191-192. It's correct that spruce doesn't have a significant trend, but the four trees taken together as replicates do.

*The paper would benefit from adding some supporting references and topics, such as the issue of Hg translocation within the bole, to the Discussion. I have indicated these by line number in the comments below.*

Thank you. Our responses are detailed below.

*Line 66 (also line 230). This thought is incomplete, your argument hinges not on the mere fact that Hg is still being deposited to forest soils, but that the net Hg storage in soils continues to increase.*

We revised this sentence and added another. "Alternatively, if soil is the source of Hg to wood tissues, concentrations should be relatively stable or even increasing over time, due to the continued, albeit lessened, rate of Hg accretion in inputs to soils from throughfall and litterfall. The residence time of Hg in soil is very long, so we anticipate limited loss of Hg from soil pools (Yu et al. 2014)."

*82-89. The climate data for the two sites should be presented in parallel fashion for direct comparison.*

We were able to find archived climate data for Heiberg forest and compute comparable climate statistics.

The Sven O. Heiberg Memorial Forest is located in the Alleghany plateau in Tully, New York (42°45'53"N, 76°04'46"W). The average temperature is -7°C in January and 19°C in July and annual average precipitation is 115 cm (<http://climate.ncsu.edu/cronos/?station=308627>). Soils are classified in the Mongaup-Hawksnest Complex, formed from till of sandstone and siltstone.

*124. It states that all clonal trees were tapped, but only 8 of the 10 clonal groups are represented in Figure 3.*

We added a sentence: "Samples collected from Clone 3 and Clone 5 were not included in this study because of their limited sample volume (< 15 mL)." A volume of 15 mL was needed for Hg analysis using the direct Hg analyzer.

*132-133. Were the blocks adjacent to each other in a grid? If so sampling at the corners would result in a lot of the soil samples being close together.*

Our original description was vague; they weren't very close to the corners. "To collect soils, we divided each block into a 3x3 grid (forming 9 subplots) and sampled points at the center of each of the middle and four corner subplots. Samples of the organic horizon (Oea) and mineral soil (0-10 cm) samples were collected using soil corers at the center and the four corners of from these 5 locations in each of the six blocks and composited by block for analysis."

156-162. This paragraph is one long sentence, it could be broken up.

We revised the text into three sentences (changes not tracked.) "The method detection limit (MDL) was calculated using the U.S. Environmental Protection Agency Method Detection Limit procedure (40 CFR 136, Appendix B, revision 1.11). To determine the MDL, the critical Student's *t* value was multiplied by the standard deviation of the mean concentration of seven replicate 0.5 g samples of NIST 1515 (apple leaves). The calculation yielded the MDL of 0.03 ng in units of mass (corresponding to a concentration of 1.27 ng g<sup>-1</sup>). This value was lower than our measured Hg concentrations, which ranged from 0.05 to 0.45 ng for sap, 0.10 to 0.43 ng for wood, and > 0.7 ng for foliage and soils."

176-178. Soils were composited by block (line 132-133). So you have 6 soil Hg values (apiece for organic and mineral) by block which you are comparing to 4 foliar concentrations averaged by clone, but the clones are randomly distributed across the blocks. I don't see how you can do this. Have I misinterpreted something in the method?

Thank you for pointing this out. We did not compare foliar Hg to soil Hg. We revised the sentence describing the data analysis: "Pearson correlation was used to compare foliar Hg to sap Hg concentrations based on the 12 trees, and soil Hg concentrations (organic and mineral), based on the average of replicate trees for 4 clones. Pearson correlation was also used to compare foliar and sap Hg concentrations to organic and mineral soil Hg concentrations, averaging sap Hg for all the two trees in each of the 6 blocks."

In response to this comment, we also added a sentence about the sampling methods, because our analysis depended on the sampling design, which we had not described in sufficient detail: "We used pole pruners to collect sun-lit leaves from three replicate trees of each of Clones 1, 4, 2 and 9, which had the highest and lowest Hg concentrations in the sap. The 12 trees were selected to include two trees (of 2 different clones) in each of the six blocks. Organic horizon (Oea) and mineral soil (0-10 cm) samples were collected using soil corers at the center and the four corners of each of the six blocks and composited for analysis."

177. Does "average" here refer to foliar AND sap Hg concentrations? Or just foliar?

This section was confusing but has been edited as described above. We are now using Pearson correlation for each of 12 trees, which gives us more statistical power, but the relationship is still not significant. This result is reported on line 206: "Foliar Hg concentration was not related to concentration in the sap ( $p = 0.5329$ )."

178-180. This is a clear, apples to apples comparison that the preceding sentence lacks.

Thank you.

181. This is a one sentence paragraph that should fold into the previous paragraph.

The sentence is "Statistical tests were performed using SAS 9.4 (SAS Institute Inc. 2013)." We decided to make this a separate paragraph, even though it's only one sentence, to avoid the suggestion that it pertains only to the statistical tests in the preceding paragraph. It applies to all the statistical tests in the section (3 paragraphs).

191-192. I have a very hard time believing from Figure 1 that spruce Hg declined significantly since 1960, especially at the  $p < 0.001$  level. It looks quite flat to me, please check this.

This decline at  $p < 0.001$  level was based on a linear regression treating the four trees as replicates, for the data from 1955 to 2015. We changed the sentence to: "For all Treating the four species as replicates, Hg concentrations declined in wood formed after 1960 ( $p < 0.001$ , Figure 1)."

It's true that there is not a significant decline for spruce alone.

208. This statement is not supported by an existing figure, but it references figure 5. There are only 3 figures.

Thank you for catching this error. We decided not to show the insignificant patterns in foliar Hg by clone (this was once Figure 5) nor the insignificant pattern is sap Hg by block (formerly Figure 4). The reference to the non-existent Figure 5 has been replaced with "data not shown."

220. Shouldn't this just read "phloem transport"?

We deleted the reference to xylem transport: "Translocation of Hg from foliage to wood likely requires additional transformation during phloem and xylem transport (O'Connor et al. 2019)."

221. The paper alternately uses stomata and stomates.

We changed "stomates" to "stomata."

236-238. Does this statement pertain to the bole or to the tree in general?

Only to the bole. We revised the sentence: "In a study of hardwood and conifer forests across four sites in the northeastern USA, yellow birch was significantly higher in wood Hg concentration than the other five species studied (Yang et al. 2018)."

238-239. Is this "soil warming" from an experiment? There is a reference but it should be clear from context.  
We added "experimental" to "soil warming."

260-264. Another relevant reference for this section is:

Navrátil et al. (2018): Navrátil, T., Nováková, T., Shanley, J.B., Rohovec, J., Matoušková, S., Vaňková, M. and Norton, S.A., 2018. Larch tree rings as a tool for reconstructing 20th century Central European atmospheric mercury trends. *Environmental Science & Technology*, 52(19), pp.11060-11068.

Thanks. We added this paper: "...Hg contamination and industrial activity for European beech, Norway spruce (Hojdova et al. 2011), and Scots pine (Navrátil et al. 2017), and European Larch (*Larix decidua*) (Navrátil et al. 2018) in the central Czech Republic, Japanese cedar (*Cryptomeria japonica*) in southern Korea...."

269-271. This wording is confusing as to whether a correlation or lack thereof was expected.

"However, sap Hg concentrations were also not correlated with foliar Hg concentrations; as might be we would expect a positive correlation if the differences in clones were due to differences in stomatal uptake of Hg. "Pearson correlation was used to compare foliar Hg to sap Hg concentrations based on the 12 trees, and soil Hg concentrations (organic and mineral), based on the average of replicate trees for 4 clones. Pearson correlation was also used to compare foliar and sap Hg concentrations to organic and mineral soil Hg concentrations, averaging sap Hg for all the two trees in each of the 6 blocks."

276. I think you mean "poplar"?

Very funny, sorry about that! "popular" was corrected to "poplar."

284. Not a sentence.

This was fixed by moving a misplaced "that."

293. I think the relevant point is not that the atmospheric sources were similar, but that the atmospheric Hg concentrations were similar.

Good point. "atmospheric sources" was changed to "atmospheric deposition rates."

305-309. The two sentences in this paragraph are disjointed. Mentioning Heiberg also in the first sentence, which would be appropriate, would help tie it together.

We decided to delete the paragraph. Unfortunately, the two data sets presented in this paper are disjointed! The first sentence was about wood Hg at Hubbard Brook and the second is about sap Hg at Heiberg. "Our results suggest that the atmosphere is the dominant source of Hg in wood for several species, including sugar maple, at the Hubbard Brook Experimental Forest in New Hampshire. It is not clear what mechanism controls variation in sap Hg by clone in sugar maple at Heiberg Forest in New York State." These ideas are presented elsewhere in the paper.

312. This number (7.3 ng/g) is quite precise given the range of sap Hg concentrations. Is it based on the average sap Hg across sites? Median? Please state. Now as to the accuracy: The average Hg concentration in sap is 4.8 ng/L (line 198). So going to 7.3 ng/g is a 3 order-of-magnitude concentration increase. But making syrup from sap takes a less than 2 order-of-magnitude concentration of the liquid. I also wonder how much Hg might be driven off (evaded) during boiling? The syrup is even less harmful than you figured.

Thank you for checking our calculations. We corrected this result: "However, even after concentrating maple sap (with a sugar concentration of 2.5%) into maple syrup (with a sugar concentration of 67%), the Hg concentration would be only 7.3 0.1 ng g<sup>-1</sup>, much which is orders of magnitude lower than the safe level in food of 500 ng g<sup>-1</sup> (Choi 2011)."

313-314. Yes, it is generally safe to consume plants, but rice can be a notable exception. A couple papers that would be good to cite:

Meng B., Feng X., Qiu G., Anderson C.W., Wang J., Zhao L. 2014a. Localization and speciation of mercury in brown rice with implications for pan-Asian public health. *Environ. Sci. Technol.* 48:(14), 7974-81, <https://doi.org/10.1021/es502000d>.

Brombach C.C., Manorut P., Kolambage-Dona P.P.P., Ezzeldin M.F., Chen B., Corns W.T., et al. 2017. Methylmercury varies more than one order of magnitude in commercial European rice. *Food Chem.* 214, 360-365, <https://doi.org/10.1016/j.foodchem.2016.07.064>.

Thank you for these references. We have added, "with the notable exception of rice" and cited these papers.

Figure 2. Although all 4 trees show a decrease in bole Hg from crown to base, the rate of decrease is small, and it is based on only three points per tree. Can you put error bars on these points, even if only analytical error?

We didn't want to add error bars because they would be assumed to represent sampling error, which is more commonly reported; readers would need to be attentive to the figure caption to understand the error bars. There is

also a practical difficulty in showing these values: the analytical error, based on duplicate samples, was 0.05 ng/g, which is the size of the symbols on the graph. It's good if measurement error is small relative to sampling error!

*Note: All 3 figures say "Figure 1" in the upper left-hand corner.*

I apologize, but I don't know how to correct this; it was added in the automated generation of the pdf from the files we uploaded.

*Reviewer #3: This manuscript describes mercury concentrations measured in tree rings, with the results addressing a wide range of questions. These include: How do mercury concentrations change from the inner to outer rings, and what does this tell us about foliar versus root uptake of mercury? What are the differences in concentrations between tree species? How do bole wood mercury concentrations vary with respect to height in the tree? What is the concentration of mercury in sap taken from sugar maples? Are sap concentrations related to soil or foliar concentrations of mercury? How do wood concentrations differ among ten sugar maple clones?*

*Correct, except that we don't really have the statistical basis to attribute differences among our four trees to species, because we don't have replicates within species. We tried to be careful how we described those differences, and we don't include this question in our introduction or objectives.*

*This is a lot of ground to cover in a single manuscript. The experimental design is fairly modest in terms of number of samples collected and analyzed (by the authors' own acknowledgment), so this results in more of an exploratory analysis. However, there are some interesting findings presented here. First and foremost, the longitudinal variability in mercury concentrations along the tree trunk (i.e. mercury concentrations increase with respect to height) is novel and will likely spur further investigation. The consistency of this effect across tree species provides convincing evidence that this phenomenon is real. This finding could have important implications for understanding how mercury is adsorbed and incorporated into the woody tissue. The investigation of mercury concentrations in clonal plants is also well conceived, and there is some opportunity here to explore tree-to-tree variability (as opposed to writing this off as being due to "genetic control", which is somewhat of a bail out in terms of explaining this variability).*

*Thank you. The issue of genetic control is discussed in more detail below.*

*There are some other portions of the manuscript in which the relatively small sample size limits the inferences that can be drawn from the study. Given the existence in the literature of carefully controlled experiments to discern of relative dominance of foliar versus root uptake of mercury, it is perhaps unreasonable to expect that mercury concentrations measured in the wood of four trees could shed much light on this particular question. The literature has converged on the relative dominance of foliar uptake (perhaps with the exception of settings with severely polluted soils), so the findings presented here relating the tree core mercury to presumed historical levels of atmospheric mercury do not contribute much to answering the "foliar versus soil uptake" question. They do, though, indicate that the mercury concentrations are consistent with historical levels of atmospheric mercury exposure, which provides some affirmation about the quality of the dataset.*

*We would be in a worse position, given the small sample sizes, had our results indicated a soil source!*

*This manuscript could benefit from being better constrained. An easy way to do this would be to include specific research questions at the end of the Introduction section (currently the manuscript lacks specific research questions or hypotheses). Given the novelty of the longitudinal patterns of mercury found in trees, this should feature prominently as a specific research question.*

*This was a good suggestion. We revised the Objectives paragraph and included these two questions. "First, are there vertical or radial patterns of Hg concentrations in bole wood that could provide insight on the source of Hg to trees?" "Second, does sap Hg concentration reflect soil or foliar sources?"*

*The manuscript should also do a better job of trying to explain differences in mercury concentrations between the cloned trees. The observed differences are currently being chalked up as "genetic control", but there are other factors that could contribute to tree-to-tree differences (e.g. DBH, tree height, tree nutrient status) that would be really interesting to explore. More should be done along these lines, leveraging the really nice experimental design of the cloned individuals.*

*We were excited to discover the clonal plantation of sugar maple; it allowed us to answer a unique question. It's disappointing that we can't be more specific as to what makes one clone different from another, except that they differ genetically. See more below, in response to the comment on line 29.*

*In summary, there are some interesting and novel elements to this study, but the focus of the manuscript needs to be better constrained to highlight these features. Plus, much less emphasis (if any at all) should be placed on the "foliar versus root uptake" debate. Reconfiguring the manuscript in this way will likely require major revisions. We improved the objectives, as described above. One of the references to "atmospheric vs. soil uptake" has been removed.*

Minor comments:

*-lines 26-28: The logic here in the Abstract is difficult to follow given the lack of background information.*

The sentence is: "Declining concentrations from the top to the bottom of the bole ( $p < 0.001$ ) and from early to recent wood ( $p = 0.001$ ) suggest that foliar uptake of Hg is more important than root uptake." The previous sentence gives the background information: "whether tree Hg is more influenced by soil Hg concentrations, which have been stable or increasing due to the cumulative retention of historical atmospheric Hg deposition, or by atmospheric Hg deposition, which has declined in recent decades." There is not much room in the abstract, but this logic has been made more explicit in the Introduction, as described above.

*-line 29: Can "genetic control" really be claimed here? Are there any other difference among individuals that could account for this?*

There is surely some difference associated with clones that explains the difference in sap Hg by clone, but it is not clear what it is. We did measure tree characteristics such as diameter at breast height and canopy position, and we described the competitive environment around each tree, based on the distance to nearby trees and their diameters. None of these characteristics, which will be reported in a future paper describing the sweetness of the maple sap, predicted the sap Hg concentrations.

*-line 38: enhanced by industrial emissions. There are natural sources of Hg too.*

We revised the sentence to: "Terrestrial ecosystems receive significant ~~amounts~~ **inputs** of broadly toxic mercury (Hg) from atmospheric deposition of ~~industrial~~ **direct natural and human** emissions **as well as re-emissions** (Grigal 2003; Driscoll et al. 2013)."

*-line 50: specify that this is vertical translocation*

We decided that "vertical" would not improve this phrase: "~~the~~ translocation of Hg from roots **into the bole** ~~to shoots~~ **was low**"

*-lines 62-63: Atmospheric Hg concentrations are declining in many, but not all, areas of the world. Plus, give timeframe for these trends.*

The sentence was revised to address this comment. "Atmospheric Hg concentrations ~~are~~ **have declined** declining in recent decades in the northeastern USA (Drevnick et al. 2012; Zhou et al. 2017) and worldwide (Slemr et al. 2011). ~~If this, atmospheric deposition to foliage is the source of Hg in tree tissues, radial trends of Hg in the bole would be lower in more wood, it should decrease in recently formed tree rings.~~"

*-lines 87-88: Be consistent and give annual average precipitation*

This was also requested by Reviewer #2 and has been achieved, as described above.

*-line 90: Is this both wet and dry deposition?*

We clarified this point. Total atmospheric Hg deposition is  $\sim 20 \mu\text{g m}^{-2} \text{yr}^{-1}$  at the Heiberg Forest and  $\sim 22 \mu\text{g m}^{-2} \text{yr}^{-1}$  at Hubbard Brook, based on **wet Hg deposition** data from the National Atmospheric Deposition Program Mercury Deposition Network **and model estimates of dry Hg deposition** (Yu et al. 2014).

*-line 107: Does "semi-decadal" mean every 5 years?*

We changed this to "5-year growth increment" to avoid confusion and to clarify that all 5 growth rings were sampled, not just one every 5 years.

*How done with a drill bit of constant diameter? Wouldn't the averaging times change if the rings are thicker or thinner?*

We added a sentence: "**Drill bits of different diameters were used to match the width of the 5-year growth increments.**"

*-lines 107-108: Was there not enough shaving mass for duplicate samples? Since duplicates are pretty standard QA, what is the reason for not having any duplicate samples?*

The reviewer is correct. One drilled sample did not provide enough mass for analysis. We improved this description; samples were collected "along ~~radii~~ **a radial vector** from the most recently formed annual ring to the pith." We added "In all cases, multiple samples ( $\sim 10$ ) were taken and combined to provide enough material for analysis." We did not analyze multiple samples from the same location, preferring to allocate our limited sample budget to describing vertical and radial patterns. Uncertainty in the composited samples is included in the random error term of our statistical models (with analytical error likely much smaller than sampling error, as described above).

*-lines 106-111: Was there any concern for ambient air exposure/contamination of samples during shaving? Many*



*papers indicate working under 'clean room conditions' so it would probably be worthwhile to at least state the conditions in more detail (type of air circulation, time samples exposed to ambient air). One way to avoid this question in the future would be to expose a standard material to ambient air for the same period of time as samples prior to analysis.*

We did a contamination test early on to prove that samples in the clean room had similar Hg concentrations compared to freeze-dry samples. We added this sentence: "This work was done in a clean room previously shown to be free of Hg contamination (Yang et al. 2017)." This paper was already cited in support of the analytical methods.

*-line 121: What types of materials were used in the collection of sap? Metal? Teflon?*

We added, "The sap passed through a metal spile and was collected in a plastic tube."

*-lines 136-137: Good job determining MDL up front to ensure sufficient sample mass*

Thanks.

*-line 153: Great recoveries! What were the blank concentrations? Below MDL?*

Yes. We added a sentence describing the blank concentrations: "The blanks had Hg concentrations of  $0.01 \pm 0.02 \text{ pg g}^{-1}$ ". This is 3 orders of magnitude less than the MDL. Thanks for asking.

*-line 198: Not clear what this p-value actually represents. Is this describing the level of significance for a test with a null hypothesis that all concentrations are the same (i.e. the low p-value could be due to just one tree being lower than the others)?*

That information was provided in the Methods: "An aligned rank test was used to examine the effects of block and clone on Hg concentration in sap, due to the small number of samples (54 in total)." The p value is telling us that it's not likely that the difference by clone are due to chance.

*-line 201: Need to report variation in soil Hg between sites. Not clear if the variation was significant.*

I see, this is important to add, since we eliminated the figures that showed the soil values. We added this sentence: "Soil Hg concentrations ranged from 102 to 127  $\text{ng g}^{-1}$  in the organic horizon and from 74 to 100  $\text{ng g}^{-1}$  in the mineral soil, across the six blocks."

*-line 208: Skipped figure 4?*

Thank you for catching this error. The reference to the non-existent Figure 5 has been replaced with "data not shown," as described above in response to Reviewer #2.

*-line 242-243: Evidence for this statement is somewhat lacking*

We revised the text: "The fact that yellow birch experienced the greatest decrease over time in wood Hg in our study (Figure 1) is consistent with these other observations of yellow birch having a higher capacity for Hg accumulation. Thus, accumulation of Hg by yellow birch appears to ~~may~~ be more sensitive to changes in the environment than other species. ~~in terms of accumulating Hg.~~"

*-line 244: "laid down" is not a very scientific term*

An internet search reveals plenty of occurrences of this usage, both in the scientific literature (Google Scholar) and in the popular press: "Young trees tend to lay down wider rings," "In winter, growth ceases and no new cells are laid down." As newer rings of sapwood are laid down, "Tree rings laid down over thousands of years may cut through the arguments surrounding global warming." "Every schoolchild learns that trees, as they grow, lay down new wood each year..."

Nonetheless, we changed the sentence to prevent other readers confusion. "Decreasing Hg concentrations in ~~wood laid down from~~ tree rings formed between 1960 to 2015 likely reflects declines in atmospheric Hg concentrations."

*-line 268: variation in soil Hg across blocks was ...?*

Sorry, this information is now provided in the Results (as detailed in response to the comment on line 201).

*-line 284: extra "that"*

Thanks, that "that" was out of place in the sentence and has been restored to its intended position.

*-line 322: "relatively free" not a very precise term. Again, there will always be some background Hg floating around.*

Thank you for pointing this out. We eliminated this term from the sentence regarding decreasing airborne sources of Hg: "we would predict that in the future, tree tissues ~~would be relatively~~ will become increasingly free of Hg in the future."