We are grateful to the reviewer for the constructive suggestions which help to improve the quality of the submitted paper. We will consider all comments/concerns during revision of the manuscript. Below please find a point-by-point reply to each comment. Reviewer comments: bold, replies: plain text, changes in the text: highlighted yellow.

This manuscript presents a comprehensive field and laboratory study examining how waterlogging-induced oxygen limitation alters soil–atmosphere exchange of volatile organic compounds in a temperate mixed forest. By combining seasonal forest-floor chamber measurements with controlled anoxic incubations of soil cores, the authors show that waterlogged/anoxic soil patches, mainly associated with skid trails, can switch from being a sink to a strong source of aromatic VOCs, particularly toluene and p-cymene. The observed emission rates are unexpectedly high for a temperate forest system and are supported consistently across in situ observations and laboratory experiments. Overall, the manuscript addresses an underexplored source of biogenic aromatic VOCs and provides novel insights that are relevant for both forest biogeochemistry and atmospheric chemistry.

Overall, I would support the publication of the manuscript, but would like the authors to address or note the following comments/concerns:

Line 45-47, BVOC/AVOC ratio needs to be updated.

► We agree that the given ratio of Guenther et al. (1995) might not be up-to-date and underestimates the significance of BVOC contribution. GBeforestein and Galbally (2007) state that global BVOC emissions are up to 8 times higher than AVOC emissions. More recent estimates assume somewhat higher BVOC/AVOC ratios of up to 10 (Guenther et al. 2012). We therefore will change the text stating that BVOCs exceed AVOCs up to 10-fBefore and add the references of Guenther et al. (2012) and GBeforestein and Galbally (2007).

Before (L45-47): Both, anthropogenic and biogenic sources contribute to the global budget of VOCs in the atmosphere with biogenic sources exceeding anthropogenic sources by a factor of 7 (Guenther et al., 1995).

Revised: Both, anthropogenic and biogenic sources contribute to the global budget of VOCs in the atmosphere with biogenic sources exceeding anthropogenic sources up to 10-fBefore (Guenther et al., 1995; 2012; GBeforestein and Galbally, 2007).

References:

GBeforestein, A.H., Galbally, I.E. (2007). Known and unexplored organic constituents in the earth's atmosphere. Environmental science & technology, 41, 1514-1521.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T. A., Emmons, L. K., & Wang, X. (2012). The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2. 1): an extended and updated framework for modeling biogenic emissions. *Geoscientific Model Development*, *5*, 1471-1492.

Line 72-73, ref missing

► We will add a suitable reference

Before (L72/73): Waterlogging reduces the diffusion of oxygen from the atmosphere into the soil by a factor of 10,000.

Revised: Waterlogging reduces the diffusion of oxygen from the atmosphere into the soil by a factor of 10,000 (Armstrong, 1979).

Reference: Armstrong W. (1979) Aeration in higher plants. In: H.W.W. Woolhouse (Ed), *Advances in botanical research*. Academic Press, Vol. **7**. London, UK, pp 225–332.)

Line 91, could you please be more specific with examples? What microbial activities?

► We agree and will add examples of the effect of waterlogging on soil microbial activities.

Before (L91): In temperate forests, such extreme conditions of prolonged periods of waterlogging usually only occur at spatially small scales, i.e. at hotspots where particular microbial activities are strongly enhanced compared to the average soil conditions (Kuzyakov and Blagodatskaya, 2015).

Revised: In temperate forests, such extreme conditions of prolonged periods of waterlogging usually only occur at spatially small scales, i.e. at hotspots where particular microbial activities are strongly enhanced compared to the average soil conditions (Kuzyakov and Blagodatskaya, 2015). As a consequence of waterlogging events, the nutrient availability in the soil might change (Zhang et al., 2025). Moreover, nitrogen and carbon cycling are affected as seen, for example, from enhanced denitrification rates and nitrous oxide emission (Tomasek et al., 2019), altered organic carbon decomposition (Liu et al., 2022) and enhanced methane emission (Shi et al., 2026), respectively.

References:

Liu, J., Zhen, B., Qiu, H., Zhou, X., & Zhang, H. (2023). Impact of waterlogging and heat stress on rice rhizosphere microbiome assembly and potential function in carbon and nitrogen transformation. *Archives of Agronomy and Soil Science*, *69*(10), 1920-1932.

Shi, B., Wang, M., Wang, G., Li, X., & Zhang, T. (2026). Effects of hummock–hollow microtopography on soil microbial communities and their links to soil carbon cycling in a sedge peatland of the Changbai mountains. *Applied Soil Ecology*, *220*, 106861.

Tomasek, A. A., Hondzo, M., Kozarek, J. L., Staley, C., Wang, P., Lurndahl, N., & Sadowsky, M. J. (2019). Intermittent flooding of organic-rich soil promotes the formation of denitrification hot moments and hot spots. *Ecosphere*, *10*(1), e02549.

Zhang, P., Li, F., Wang, D., Tian, Y., Rong, Y., Wu, Y., ... & Zhang, H. (2025). Soil bacteria and fungi diversity analysis reveals effects of sesame (Sesamum indicum L.) under waterlogging stress. *BMC microbiology*, *25*(1), 620.

Line 127, spell out OL and OF

► We will explain OL and OF:

Before (L127): The thin forest floor consists mainly of (OL) litter with a patchy thin OF-layer, pH (in KCl) throughout the profile is between 3.5 and 4 (further details see Kreuzwieser et al. 2025).

Revised: The forest floor consists mainly of a litter horizon (OL, organic layer composed of undegraded litter) with a patchy thin fermentation horizon (OF, organic fermentation layer with partially decomposed litter material), pH (in KCl) throughout the profile is between 3.5 and 4 (further details see Kreuzwieser et al. 2025).

Line 151, did you use any base collars?

► No, we didn't use any base collars. To point this out, we will change the text accordingly:

Before (L151): Per undisturbed plot (26 plots; plot size ca. 1.5 x 1.5 m) one self-constructed Plexiglass chamber (40 x 25 x 10 cm, length, width, height) was randomly placed on the forest floor.

Revised: Per undisturbed plot (26 plots; plot size ca. 1.5 x 1.5 m) one self-constructed Plexiglass chamber (40 x 25 x 10 cm, length, width, height) without collar was randomly placed on the forest floor.

Line 156, I am a bit concerned about the large volume of air samples by the adsorption tubes, which may have exceeded the breakthrough volume of the adsorption materials. Check here for the potential breakthrough volume of the compounds you are interested in – https://www.ingenieria-analitica.com/attachment/pdf/print/td20-confirming-sorbent-tube-retention-volumes-and-checking-for-analyte-breakthrough-pdf-417

► Thank you very much for this important comment. We checked for the breakthrough volumes in the mentioned table. Based on these values, the sampled air volume may indeed exceed the safe sampling volume (SSV) of our sorbent tubes, particularly for toluene, potentially leading to an underestimation of true toluene emissions.

However, it should be noted that the toluene concentrations in air were low (ppt-range) and sampling temperature during the campaigns continuously decreased after August, reaching close to 0°C in December (when toluene emission was high), which is substantially lower than the reference temperature (20°C) used for the shown SSV values. Lower sampling temperatures and low VOC concentrations are expected to increase sorbent retention and reduce breakthrough (i.e. enhance breakthrough volume). Nevertheless, even when accounting for this temperature effect, the sampled air volume (24 L) may still exceed the mass-adjusted SSV, and breakthrough for toluene might not be ruled out completely. Consequently, the reported toluene emission rates should be regarded as conservative estimates, implying that actual emissions may even be higher and further emphasizing the significance of the identified toluene hotspot in the forest ecosystem.

We propose to add some sentences on this in the revised version of the manuscript. We would add them to M&M (section 2.2.1.).

Revised (to be added at line 156): Sampling occurred for two hours resulting in 24 L of chamber air channelled over the adsorbents. Considering this high sample volume, we might have slightly exceeded the breakthrough volume for toluene in our approach, which is 76 L for 200 mg Tenax TA at 20°C (www. ingenieria-analitica.com/attachment/pdf/print/td20-confirming-sorbent-tube-retention-volumes-and-checking-for-analyte-breakthrough-pdf-417). However, at times with high toluene abundance, temperatures were below 10°C increasing the breakthrough volume. Still, the reported toluene emission rates in autumn should be regarded as conservative estimates, implying that actual emissions may even be higher. To maintain ambient air pressure in the chambers, three inlet openings (diameter 4 mm) allowed ambient air to enter the chamber.

Section 2.4 I am also curious how many potential VOCs you have detected from the temperate forest soils. Were these compounds you focused on the most emitted ones?

► In this study, we focused on VOCs that exhibit distinct emission patterns between well-drained and waterlogged soils in order to investigate how soil compaction–induced changes in soil oxygen availability influence soil VOC emissions. Accordingly, we did not aim to provide a comprehensive characterization of all VOCs detected in the samples. For a broader overview of VOCs observed at this site, we refer readers to our recently published studies conducted at the same location. We fully agree that this information should be given in the paper and therefore will change the first paragraph of the discussion:

Before (358): We conducted the study in the ECOSENSE forest, a temperate mixed European beech and Douglas fir ecosystem, where seasonal measurements of terpenoid exchange have been performed previously (Kreuzwieser et al., 2025).

Revised: We conducted the study in the ECOSENSE forest, a temperate mixed European beech and Douglas fir ecosystem, where seasonal measurements of terpenoid exchange have been performed previously and exchange of over 20 compounds including monoterpenes, aldehydes and some aromatic compounds was observed (Kreuzwieser et al., 2025; Lee et al., 2025).

References:

Kreuzwieser, J., Lee, H., Köhler, A., Christen, A., Sulzer, M., Schack-Kirchner, H., Brzozon, J., Lang, F., Daber, L.E., BechtBefore, R. and Werner, C., 2025. Biotic and abiotic factors controlling terpenoid exchange from soil of a mixed temperate forest ecosystem. Soil Biology and Biochemistry, p.109991.

Lee, H., Katlewski, S., Weber, P.C., Wehlings-Schmitz, S., Brzozon, J., Schack-Kirchner, H., Werner, C. and Kreuzwieser, J., 2025. Soil terpenoid content and emissions are shaped by litter chemistry and soil depth. Plant and Soil, pp.1-18.

I noticed that the observed toluene emission rates were substantially higher from the waterlogged skid trail soils than from the undisturbed forest floor soils. Could longer-term accumulation of toluene in skid trail soils, for example through repeated deposition from vehicle exhaust, potentially contribute to the elevated emissions observed at these locations, in addition to the proposed biogeochemical mechanisms?

► We also had this idea when we saw the high emission of aromatics along the skid trails for the first time. However, there are reasons why this cannot be the case:

(i) The skid trail sites investigated in this study were not subject to repeated traffic by machinery/vehicle. Therefore, it is unlikely that the elevated toluene emissions observed at these locations resulted from long-term accumulation and subsequent release of toluene derived from vehicle exhaust deposition. Rather, we interpret the enhanced toluene emissions as being primarily associated with soil compaction caused changes in soil physical and chemical conditions.

(ii) Strong support for this interpretation is provided by our laboratory waterlogging experiments using undisturbed forest soil cores. These soil cores did not show detectable toluene emissions prior to treatment. However, toluene emissions gradually increased over time following waterlogging. This experimental result underlines that the observed toluene emissions arise from soil processes triggered by reduced oxygen availability, rather than from the remobilization of previously accumulated toluene.

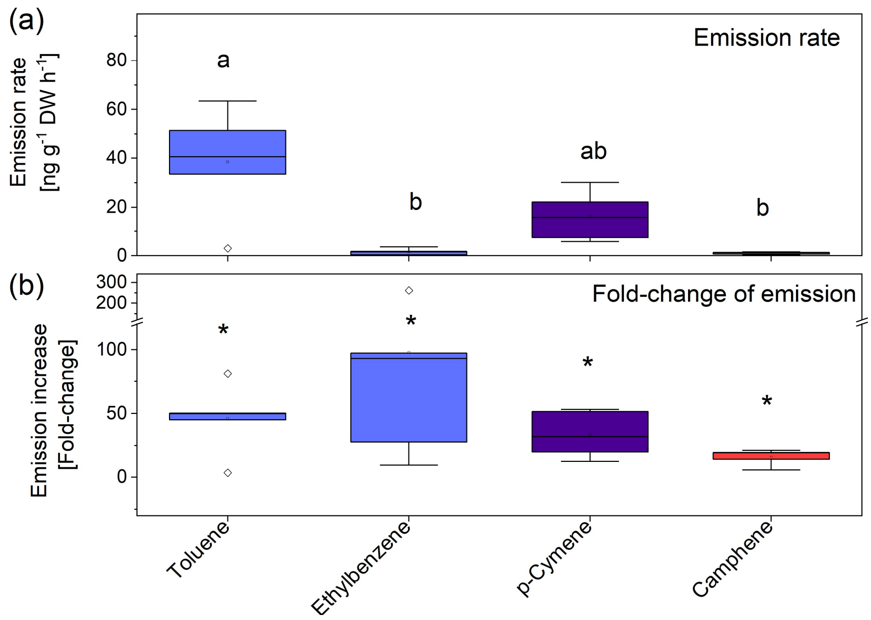
To make this point clearer, we propose to add the following sentence at line 368:

"…. to 5000 ng m^{-2} h^{-1} and 3000 ng m^{-2} h^{-1}, respectively. Since these high emissions were observed on skid trails, it might be argued that this is due to repeated deposition from vehicle exhaust and subsequent release of such aromatic compounds. However, there was no regular use of the skid trails before the time of our measurements. Moreover, also soil from undisturbed areas showed strong toluene emission after exposure to anoxic conditions, clearly indicating de novo formation. To the best of our knowledge…."

Figures 3 and 4: It may be clearer to present the emission increases on the y-axis as fBefore changes rather than percentages, as this would make the values easier to read and compare.

► We agree and designed new figures 3 and 4. We will replace the % increase by fBefore-changes of treatment vs control. The figure captions will be changed accordingly.

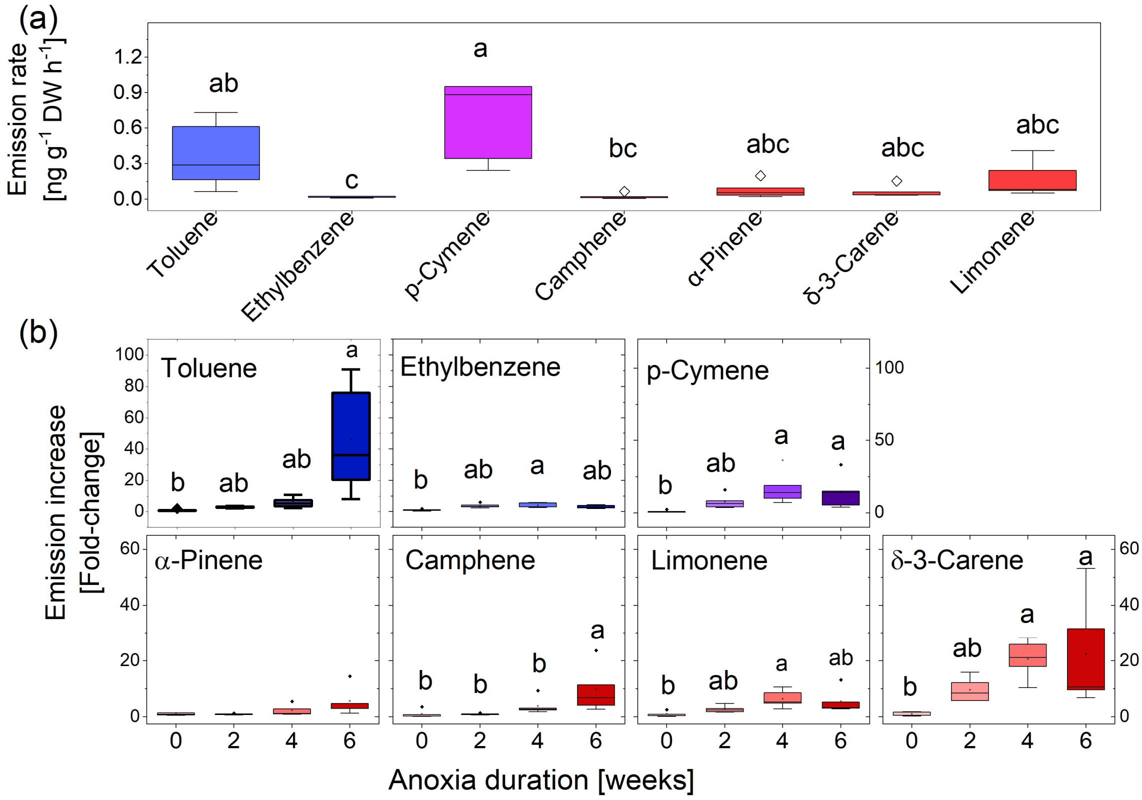
Revised figure 3



Before: Figure 3 Emission of main aromatics and terpenes under anoxic conditions (a) and percent increase of emission under anoxia compared to normoxic conditions (b). Five soil cores from waterlogged plots and five soil cores from well-drained control plots were collected in the ECOSENSE forest and transferred to the lab. (a) Emission of volatiles was immediately determined. Different letters indicate statistically significant differences at $p<0.05$ as calculated by Kruskal-Wallis test followed by Dunn's posthoc tests. (b) Percentages of increase of emissions between waterlogged soil compared to well-drained soil was calculated by setting emission rates of controls 100%. Asterisks indicate significant differences between waterlogged and well-drained soil as calculated by Mann-Whitney test at $p<0.05$.

Revised: Figure 3 Emission of main aromatics and terpenes under anoxic conditions (a) and fBefore-change increase of emission under anoxia compared to normoxic conditions (b). Five soil cores from waterlogged plots and five soil cores from well-drained control plots were collected in the ECOSENSE forest and transferred to the lab. (a) Emission of volatiles was immediately determined. Different letters indicate statistically significant differences at $p<0.05$ as calculated by Kruskal-Wallis test followed by Dunn's posthoc tests. (b) FBefore-change increase of emissions between waterlogged soil compared to well-drained soil was calculated by setting emission rates of controls 100%. Asterisks indicate significant differences between waterlogged and well-drained soil as calculated by Mann-Whitney test at $p<0.05$.

Revised figure 4



Before: Figure 4 Emission of main aromatics and terpenes after six weeks of anoxia under controlled conditions (a) and increase of emission compared to normoxic conditions (b). Volatile aromatics (blue), the aromatic monoterpene p-cymene (violet) and major terpenes (red). Thirty soil cores from well-drained plots were collected in the ECOSENSE forest and transferred to the laboratory. Fifteen soil cores were kept under normoxic conditions serving as controls and 15 were fumigated with nitrogen and covered with water to induce anoxia. Every second week, five samples each were used for VOC emission measurements. Data shown represent emissions rates after six weeks of anoxia (a) and the increase as percentages of exchange rates from anoxic compared to normoxic conditions, which were set 100%. Different letters indicate significant differences between emission rates of different compounds (a) or treatments (b) as calculated by Kruskal-Wallis test followed by Dunn's posthoc tests at $p<0.05$, DW, soil dry weight.

Revised: Figure 4 Emission of main aromatics and terpenes after six weeks of anoxia under controlled conditions (a) and fBefore-change increase of emission (anoxia vs. normoxia) (b). Volatile aromatics (blue), the aromatic monoterpene p-cymene (violet) and major terpenes (red). Thirty soil cores from well-drained plots were collected in the ECOSENSE forest and transferred to the laboratory. Fifteen soil cores were kept under normoxic conditions serving as controls and 15 were fumigated with nitrogen and covered with water to induce anoxia. Every second week, five samples each were used for VOC emission measurements. Data shown represent emissions rates after six weeks of anoxia (a) and the fBefore-change increase of exchange rates from anoxic compared to normoxic conditions. Different letters indicate significant differences between emission rates of different compounds (a) or treatments (b) as calculated by Kruskal-Wallis test followed by Dunn's posthoc tests at $p<0.05$, DW, soil dry weight.

In the anoxic incubation experiments, nitrogen was used to establish anoxic conditions. I was wondering how quickly the soils became anoxic after the onset of nitrogen flushing, and whether this transition can be considered effectively instantaneous. If the establishment of anoxia was rapid, I find it somewhat surprising that VOC emission fluxes could increase by several tens of times over such a short timescale, as microbial processes are not generally expected to respond that quickly.

► We would like to clarify that anoxic conditions in our laboratory experiments were not established instantaneously by nitrogen flushing. As described in Section 2.3, (1) soil cores collected from waterlogged plots for the first experiment and (2) soil cores collected from well-drained plots but used for anoxic treatment in the second experiment were immediately filled with water after sampling in order to minimize oxygen diffusion into the soil. Accordingly, (1) soil cores from waterlogged plots were already under anoxic conditions in the field, and (2) the transition toward low-oxygen or anoxic conditions in soil cores from well-drained plots began after water filling in the laboratory, rather than being induced by nitrogen flushing.

To make this point clearer, we will change the text:

Before (lines 185/186): Glass tubes containing soil cores from waterlogged sites and soil cores dedicated for anoxic treatment, were immediately filled with water to minimize oxygen diffusion into the soil.

Revised: Glass tubes containing soil cores from waterlogged sites and soil cores dedicated for anoxic treatment, were immediately filled with water to minimize oxygen diffusion into the soil and gradually establish anoxic conditions comparable to the situation at the field site.

Section 3.3 Ambient toluene mixing ratios show a pronounced seasonal increase, while soil exchange rates indicate net uptake during the same period. How would you reconcile this decoupling between atmospheric concentrations and net soil fluxes? To what extent might changes in canopy structure, phenology, or boundary-layer dynamics during the seasonal transition influence the observed ambient VOC mixing ratios? Elevated ambient toluene concentrations kind of coincide with the period of leaf senescence. Could there be any potential contribution of senescing leaves or fresh litter to the observed atmospheric signal?

► You are fully right, and we agree that we have to be more cautious with our statements. The idea that the elevated toluene concentrations in autumn might be related to enhanced emissions from soil hotspots is discussed in lines 410-420. Here we state that the "data provide circumstantial evidence" for a connection between increasing toluene concentrations and high toluene emission from waterlogged soil in late autumn. We propose here that toluene emissions from hotspots overcompensate for toluene uptake by well-drained soil.

In line 424/425 we briefly mention that also leaf litter and senescing leaves might act as source of toluene. However, we agree that this statement is too short and the role of leaf litter and leaves should be further considered. Therefore, we propose to strengthen their potential role as toluene source:

Before (line 424/425): Still, another potential source of atmospheric toluene might be leaf litter (Gray et al., 2010; Isidorov and Zaitsev, 2022; Rocco et al., 2025) or senescing European beech leaves. This is because the time of highest atmospheric toluene abundance coincided with leaf senescence, when leaves were yellowing but still not shed as seen from available LAI data.

Revised: Still, also other potential sources of atmospheric toluene should be considered. Several studies demonstrated that leaf litter emits toluene (Gray et al., 2010; Isidorov and Zaitsev, 2022; Rocco et al., 2025). However, this assumption was not supported by our data because well-drained soil including fresh leaf litter rather took up toluene than emitting it into the atmosphere. Another potential toluene source might be senescing European beech leaves. This is because elevated toluene levels in the atmosphere were detected at the same time when beech leaves started senescing. Future studies should include seasonal leaf level measurements to elucidate if toluene is emitted by European beech leaves at all, and if emissions rise in parallel with leaf senescence in autumn.

.

Section 3.3 The chamber-based exchange rates integrate fluxes over a two-hour period, whereas ambient mixing ratios reflect cumulative atmospheric conditions. Could differences in temporal integration affect the interpretation of the soil–atmosphere coupling?

► We are grateful to the reviewer's question as it indicates a misunderstanding. Actually, the ambient measurements and the chamber measurements were done at the same time and for exactly the same duration. Therefore, ambient mixing ratios of VOCs can be directly linked to soil VOC exchange rates. We propose to re-phrase the text to avoid any confusion

Before (line 157): During each chamber measurement, ambient forest air at 120 cm height above soil surface and in close vicinity of the soil chambers was sampled as well. VOC concentrations of these background measurements were subtracted from chamber values. Moreover, they were used to determine ambient mixing ratios of forest air. After sampling, air sampling tubes were transported to the laboratory and stored at 4°C until analysis which usually occurred within some days after sampling.

Revised: During each chamber measurement ambient mixing ratios of VOCs were quantified. For this purpose, ambient forest air at 120 cm height above soil surface and in close vicinity of the soil chambers was sampled with exactly the same method as described for chamber measurements, i.e. air was sucked for 120 min at a flow rate of 200 mL min^{-1} over air sampling tubes. VOC concentrations of these measurements were also considered background levels, which were subtracted from chamber values. ~~Moreover, they were used to determine ambient mixing ratios of forest air.~~ After sampling, air sampling tubes were transported to the laboratory and stored at 4°C until analysis which usually occurred within some days after sampling.

Section 3.3 Although waterlogging is proposed as a key driver of enhanced toluene emissions, ambient toluene mixing ratios do not show a clear positive correlation with measured soil moisture. Could this discrepancy be briefly discussed?

► We agree that this seems like a discrepancy at a first glance. However, also our laboratory-based experiments under controlled conditions demonstrated that it is not soil moisture alone that stimulates toluene production and emission. After adding water to the soil cores, it took two weeks until enhanced toluene emission was detected and even longer for very high toluene emission rates. We assume that the redox potential of the soil must decrease before the relevant microbial community forms toluene. This idea will be better presented in the revised discussion. We will add the following sentences at the end of section 4.1.

Revised: Considering that waterlogging stimulated toluene production and emission, it was astonishing that there was only a weak correlation between soil moisture and toluene emission from the forest floor. However, this finding is in good agreement with our experiments under controlled conditions; adding water to soil cores did not directly cause enhanced toluene emission. It took several days until toluene emission increased, most likely until the redox potential dropped and soil microorganisms started producing toluene.

For Section 4.3, I found parts of the mechanistic discussion to be somewhat speculative, although valuable and well grounded in the existing literature. Given that the proposed biochemical pathways are inferred rather than directly demonstrated in this study, it may help to more clearly distinguish between interpretations that are directly supported by the data and those that remain more hypothetical.

► We fully agree and propose to revise several sentences as follows:

Before (Line 435): “This finding provides evidence that emission of toluene and p-cymene is not due to release of trace gases abundant in soil pores or adsorbed at soil particles…”

Revised: “This finding indicates that emission of toluene and p-cymene is unlikely to result from the release of trace gases abundant in soil pores or adsorbed at soil particles…”

Before (Line 439): “enhanced toluene emission can only be explained by de novo formation in the soil.”

Revised: “the observed enhancement in toluene emission supports the interpretation of de novo formation in the soil.”

Before (Line 441): “There are two known pathways possibly contributing to formation of toluene in the present study.”

Revised: “Based on existing literature, two pathways have been proposed that could potentially contribute to toluene formation under the conditions observed in this study.”

The discussion suggests that emissions from waterlogged hotspots may influence ecosystem-scale toluene budgets. Given that such hotspots likely occupy only a small fraction of the forest area, might it be appropriate to slightly tone down the strength of this inference, or to more clearly state its limitations?

► We agree and try to be more cautious with our statements. According to the comments, we propose to revise as follows:

Before (Line 486): “toluene emission from anoxic hotspots has the potential to impact the whole ecosystem’s toluene budget, i.e., increasing atmospheric toluene mixing ratios.”

Revised: “toluene emission from anoxic hotspots might have the potential to contribute to the ecosystem-scale toluene budget, particularly under conditions where emissions from such hotspots are sustained over time.”