EGUSPHERE-2025-2882, reply on RC1

Comments are copied in plain text.

Replies are blue.

Changes to the manuscript are in green.

General comments:

The authors made a comprehensive and interesting study on the size and applicability of TFA concentration on catchment scale. The study includes many data that may be used by others interested in similar catchment scale studies or in other study types. In general discussions and conclusions are both interesting and justified. However, I am skeptical about the justification of the calculations on agricultural input. While the higher TFA concentrations at the Dreisam Gauge is likely to be caused by an input from agricultural activity, as the authors suggest, I think the numbers used in the calculations in combination with the general understanding of agricultural sources are so uncertain that the calculated mass of TFA from agriculture makes no sense and should be removed from the abstract and probably from the article and more emphasis should be put on the need for more research into this important aspect of the study. Please see specific comments for more details on this.

We thank the reviewer for the comprehensive feedback and the time invested in the review. We agree that calculating an agricultural TFA excess is highly uncertain. We followed the reviewer's suggestion and removed the TFA surplus from the abstract and the article.

Specific comments:

Line 51 "the lack of transformation rates". Not only transformation rates are lacking but for most potential TFA-pesticides also whether the pesticides are transformed into TFA at all.

To take the potential of “non-transformation” into account, we changed the sentence to:

Moreover, it remains unknown whether, and to what extent, the degradation of PFAS PPPs releases TFA.

Line 93: How were HFCs (and which) used as tracers? Normally you would expect CFCs to be used as tracers. Are HFC/CFC data in Uhlenbrook et al, I cannot find them?

We assume the existence of two Uhlenbrook 2002 references has led to some confusion. There is a paper called “Hydrograph separations in a mesoscale mountainous basin at event and seasonal timescales” with two more co-authors cited as “Uhlenbrook et al. 2002” and another one called “Process-oriented catchment modelling and multiple-response validation” with one more coauthor cited as “Uhlenbrook and Leibundgut, 2002”. In short: The freons F-11, F-12 and F-113 were used and the full article can be found under doi: 1029/2001WR000938. And as the reviewer stated: Those molecules also contain chlorine atoms and are therefore called chlorofluorocarbons (CFCs), we changed the statement in brackets accordingly:

(major ions, silica, and clorofluorocarbons – CFCs, namely freons F-11, F-12 and F-113)

Line 95-96: “HOF" and "SOF" are these abbreviations necessary - they are not used any further?

We introduce the abbreviations in the new version of the draft and use them in the paragraph.

Hortonian Overland Flow (HOF) from impervious surfaces, such as roads, rock outcrops, or urban areas, and by Saturation Overland Flow (SOF) from saturated areas, including wetlands and riparian zones.

In terms of magnitude, SSF is the dominating process of storm discharge generation when compared to HOF and SOF (Steinbrich et al. 2016).

Line 143: "a polyethylene (PE) tank located below the funnel". This is not much information on the precipitation collection. What material was the funnel made of? Did you test that it did not leach or adsorb TFA? Was it assured that there was not evaporation from the tank? Was it a constantly open funnel, so that also some of the dry deposition would be sampled? In general, dry deposition is not considered in the study, but some studies claim that it is a substantial fraction of atmospheric TFA input (most extreme case by Zhuang et al (https://www.sciencedirect.com/science/article/abs/pii/S0304389424009622), and most studies estimate at least some dry deposition. I guess there will also be quite some fog/mist in the area, that can be captured by especially conifers, but will probably not be captured by the precipitation sampler, or?

We revised the section to address the reviewers' questions regarding blank testing of materials and evaporation protection. We follow the author's suggestion and added dry deposition as a potential source of errors to our discussion. We propose the following changes in the manuscript:

On the same day, we retrieved precipitation samples: rainwater was stored in a tank located below the funnel, which was emptied after sampling. The tank, funnel, and connecting tube (0.82 m) were composed of polyethylene (PE), and the material was blank tested to exclude any contamination of the material. Because of the strong hydrophily of the TFA molecule we did not consider sorption to the hydrophobic PE-material. The setup allowed the tank to be placed in the high grass to keep temperatures low. Furthermore, the tank was protected by a black rubber mat to minimize evaporation, shielding it from solar radiation. Evaporation protection was necessary because the funnel was left permanently open to allow for additional dry deposition sampling.

Although our precipitation sampler was permanently open, dry TFA deposition may be enhanced in forested environments due to the larger surface area provided by tree canopy, which increases the potential for particle deposition. Indeed, studies have reported increased TFA levels in throughfall compared to bulk precipitation, supporting the likelihood of additional inputs via fog and dry deposition in forest ecosystems (Jordan und Frank 1999). This suggests that our atmospheric TFA input estimates based solely on open field bulk precipitation measurements might be too low, resulting in an apparent lack of retention when performing mass balance calculations.

Line 147: "Storage time was up to four months". Stored how?

We added information on sample storage to the sentence.

Samples were stored in the dark and at room-temperature for up to for month, but…

Line 275: "the highest TFA levels in the Dreisam River were observed during the 2023/2024 winter". To me this is not clear from the figure - it looks relatively constant, when considering the fluctuation from sample to sample...?

The intention was to emphasize the high TFA levels in the Dreisam River, despite the low precipitation levels. We rephrased accordingly.

2. Despite low TFA levels in precipitation, high TFA levels were observed in the Dreisam River during winter.

Line 277: "TFA concentrations increased with discharge". Is this statistically significant - it is not so easy to see from the figure, I think. Sometimes there seem to be a clear positive relationship between the two, sometimes not...

We added a sentence to link the information on significance levels and correlations to this observation.

We also refer to Table 2 where correlations and significance levels are listed. The discharge-TFA correlation is statistically significant at Dreisam and Brugga ($p \leq 0.001$), but not at the Spring Zipfeldobel.

Line 334: "The Dreisam River exported 48 ± 21 % more TFA." This seems like a very uncertain calculation, since according to the figure in 2023 there was no surplus and in 2024 the surplus was around 100%? Taking average of two such different numbers makes no sense to me. Furthermore, how can the uncertainty be so small when taking average of two so different years? There should be input from agriculture in both years unless there is a very good (hydrological) reason for this to be so different. Although quantifying the contribution from agriculture is very important, I am not convinced it is scientifically reasonable to do from these data as apparently uncertainty is very high (which is actually expected for a whole catchment exercise)

We share the concerns of the reviewer regarding the average years uncertainty. We removed the "average" year from the plot and replaced it with the cumulative input and output over the two years.

Line 341: "main Dreisam catchment" is this the same as what is denoted DRC throughout the text?

We forgot to use the abbreviation; we changed to DRC.

Figure 4: Why use standard error, not standard deviation, so that it is easier to compare statistical significance?

As the reviewer suggested, the SE is better suited to compare significance. If we want to make a statement about retention/excess, then we need a threshold when defining these differences. The SE provides error bands suited to do so.

We added this to the draft:

To be able to compare statistical significance, whiskers indicate standard errors, calculated from bootstrapping standard errors for mean concentrations and volumes, and consecutive Gaussian error propagation.

Table 5: "Eq. 9" Do you mean Eq.8? I see no eq. 9.

We agree with the reviewer: this was a spelling mistake; it should be Eq. 8. We have made the correction accordingly.

Line 382: In general, water isotopes seem to show minor seasonal fluctuation, compared to what you would expect in the precipitation (though precipitation isotopes for the present study seems not to be

shown). It guess this would suggest the rivers are mainly fed by groundwater? Does this fit with your understanding of the system? Does it fit with the fluctuations in TFA? Do you have stable isotope data for precipitation that could be included?

The rivers are mainly groundwater-fed. We calculated Kirchner's new water fraction (the proportion of water younger than two month) from water stable isotopes and ended up with a low single-digit percent value. Still under storm conditions, SSF can make up to 50 % of the discharge (Bachmair and Weiler, 2014). For precipitation stable isotope data we refer to Figure F2.

We included a paragraph in the draft to account for the reviewer's comment:

Generally, this aligns with our understanding of a groundwater-dominated system. We calculated Kirchner's new water fraction, i.e. the proportion of water younger than two months (Kirchner 2019) from water stable isotopes and ended up with a low single-digit percent value. Under dry conditions, groundwater influence stabilized TFA concentrations in streamflow and under storm conditions SSF gains importance, mobilizing TFA.

Section 4.2: Although highly interesting, this section seems very speculative and I think it should be minimized and uncertainties put more forward.

Another reviewer also expressed concern about the lack of a thorough discussion of uncertainty. We rephrased the section to focus on adding a more thorough discussion of uncertainties. We refer to reviewer 2's comments on the discussion section.

Line 421-427: Taking an average of two very different years (in terms of hydrology and TFA balance) doesn´t make sense to me even if the average precipitation volume of the area lies between the two.

We share the concerns of the reviewer and have removed the calculation of a mass balance for the average year. Instead, we show the TFA import and export over the two years. Still, the average year seems necessary for calculating ET; therefore, we further stress the uncertainty, which is reported in the results coming with this method in Section 4.3.

Calculating *ET* based on weighted mean TFA concentrations resulted in higher error margins than using the deficit between water import and export. High fluctuations in TFA precipitation concentrations are the primary cause of elevated uncertainty.

Line 444: 100% molar yield is highly unlikely (pesticides are rarely degraded 100% to one compound, and a significant fraction of TFA and possibly the pesticides themselves will be removed with the crops). It is important to mention that this is highly unlikely, and also important to mention that so far, TFA from PPP is mainly hypothetical/potential (Joerss et al use the term "estimations of TFA

formation potentials "). For the few compounds, where TFA was demonstrated in the EFSA conclusions and elsewhere, (much) smaller fractions than 100% was found.

The reviewer is correct in stating the lack of data on PPP transformation to TFA; therefore, we have added a section that takes this into account.

Joerss et al. (2024) published a dataset estimating the potential agricultural TFA input based on a European dataset regarding emissions from plant protection products; however, it remains unknown whether and to what extent PPP degrade to TFA. The TFA excess in the DRC suggest at least some additional agricultural input.

Line 445: I doubt that manure will add much TFA compared to pesticides, but I guess you could calculate from published values for TFA in manure, mentioned in the Introduction?

We appreciate the idea and estimated liquid manure input, indeed it seems negligible:

Assuming a TFA concentration of 100 µg L^{-1} in liquid manure, as reported by the German Environment Agency (2023), an application rate of 15 tonnes per hectare (t ha^{-1}) would result in an annual input of less than 1 kg of TFA across the entire agricultural area of the DRC (5.2 km^2). Consequently, liquid manure can be considered a negligible source of TFA input in this context.

Line 454: “Eq. 9” Do you mean Eq.8? I see is no eq. 9.

This is a mistake, it is Eq. 8.

Line 458-465: What about the year-to-year variation that may be quite large, as you show in table 4 and which has also been shown by others (much higher than the 5% you mention as an average yearly increase). Maybe this should be discussed here as well?

We agree with the reviewer and have changed the section accordingly.

Adding to the uncertainty, the interannual variability of TFA concentrations in precipitation can be considerable (see Table 4 and Jordan und Frank (1999) and Henne et al. (2025)), as reflected in the relatively large error margins for *ET* values from TFA, ranging from 45% at Brugga to 70% at Zipfeldobel. This constitutes a methodological challenge, as the sampling period must be sufficiently long to capture interannual variation, while be short enough to minimize errors arising from the progressive increase in TFA concentrations over time.

Also see the change regarding the reviewer's comment on line 478.

Line 475: "reported for the degradation of precursor PPP". I would suggest to change into "reported for the potential degradation of precursor PPP"

We changed the sentence as the reviewer suggested.

Line 478: "mean residence times of 2-5 years". Why not shorter than 2 years? Is it not applicable for one year (11-16 months, line 460) as in the present manuscript?

We agree with the reviewer and changed the sentence to:

This approach is applicable in catchments with solely atmospheric TFA input. Mean residence times should remain below 5 years minimize the error from rising TFA concentrations in precipitation. Main uncertainties arise from difficulties in finding representative precipitation TFA concentrations, as overall TFA concentrations are increasing, and the interannual variability can be quite high.