

**We thank referee #2 for the relevant and constructive comments made on our manuscript. These will certainly help to strengthen the manuscript.**

*“This manuscript describes a series of comparative monitoring techniques to determine annual fluxes of P leaving study catchments. The comparisons of grab and passive sampling are made against a benchmark of time-proportional composite sampling. Estimates of resource allocation are also made. It is an interesting comparison and adds to a body of work in this area of hydrological science. There is one very large and unqualified assumption in the work; that the time proportional composite sampling approach is a true benchmark from which to make comparisons of other methods using bias and precision as the metrics. Very late in the manuscript there is some discussion on 1) a better composite sampling method – the flow-proportional method – and 2) a comparison with a 7hr time integrated approach (not composite) to qualify the approach taken by the authors – pages 12-13. For 1) a comparison of time- and flow-proportional composite sampling methods, the authors should refer to the work by Ort et al., 2010 (ES&T, 44(16), 6024-6035) and references therein. Here the two methods, amongst others, are considered in varying sewage discharges. Time-proportional composite sampling is assumed to provide a poor estimate of annual chemical flux owing to a poor representation of higher discharges in the composite sample: “Conceptually it is clear that a time-proportional sampling mode will systematically under- or overestimate pollutant loads when the flow varies, and when flow and concentration are positively or negatively correlated.” p.6028. “Generally, . . .a time-proportional mode implies that low flows, with a higher proportion of less polluted (extraneous, infiltrating) water during the night, are over-represented in a composite sample; consequently, influent loads will generally be underestimated.” p.6030 Also, the work by Abtew and Powell, 2004 (Journal of the American Water Resources Association, 50(5), 1197-1204): “Weekly flow proportional composite auto-sampling resulted in the least bias in load estimation with competitive operational cost compared to daily grab, weekly grab sampling and time proportional auto-sampling.” Abstract. The authors need to provide some reassurance and evidence early in the manuscript that this isn’t a fatal flaw in the comparisons. From my understanding and brief review of the literature, the time-proportional method should be an additional method to compare with a ‘true load’ however this is (better) defined.”*

**We do agree with the referee, flow-proportional method is a better approach to measure the true load in comparison to time proportional method. The use of a flow-proportional method would have certainly given more assurance on the accuracy and bias of the results obtained by passive samplers and grab sampling. However, when comparing passive samplers, grab sampling and time-proportional**

sampling, we find it safe to rely on the time-proportional sampling as our best estimate of the true load. However, we will make sure to state and carefully word this assumption in the revised manuscript so that it is clearer for the readers. We do agree that we need to discuss this assumption more into details and furthermore, we will provide more details on the differences of performance between flow- and time-proportional methods from the international literature and own studies.

We thank the referee for the relevant references that were suggested — these will be added to the text.

*“For 2), above, the comparison between time-proportional composite sampling and time-integrated sampling (effectively a higher resolution grab sampling where discharge and concentration have discrete datapoints) cannot be made and so cannot justify time-proportional composite sampling. Without qualification to the assumption that time-proportional composite sampling is a ‘true’ method, then it would be an unsafe benchmark and the subsequent quantitative comparisons on bias, precision and resource allocation with other methods would also be unsafe. This qualification is therefore essential for the work to go forward.”*

The referee is correct, we cannot use the comparison between time-proportional composite sampling and time-integrated sampling to justify the use of time-proportional composite sampling. We apologize for this mistake and we will delete this part of the manuscript. As stated above, we will provide more insights into the assumptions and limitations of the time proportional method in the new revised manuscript.

*“General*

*Page 2 line 10 amend to: “Assuming hourly time-proportional”*

*Page 2 line 19 amend to: “a major transport route for”*

*Page 2 lines 25-26 amend to: “In recent decades, the transport of. . .has attracted particular attention”*

*Page 3 line 6 edit: Change to ‘pose’ and delete ‘the’*

*Page 3 line 9 amend to: “to establish at least”*

*Page 3 lines 11-12 delete: “And such mitigation is a costly affair”*

*Page 4 line 8 delete: “interesting” and amend to “as they do not”*

*Page 11 line 10 amend to: “using a minimum of equipment”*

*Page 12 and global check required – consistency in use of phosphorus abbreviations*

*e.g. line 19.*

*Page 12 and 13 not safe comparisons see specific comments”*

**All these corrections will be incorporated in the revised manuscript.**