Interactive comment on “Sampling frequency trade-offs in the assessment of mean transit times of tropical montane catchment waters under semi-steady-state conditions” by E. Timbe et al.

E. Timbe et al.
edison_timbe@yahoo.com

Received and published: 24 December 2014

GENERAL COMMENTS

The paper "Sampling frequency trade-offs in the assessment of mean transit times of tropical montane catchment waters under semi-steady-state conditions" investigates the effects of different sampling intervals on Transit Time Distribution (TTD) estimates. Using weekly stream and soil water isotope samples together with event samples of precipitation, data is aggregated to different temporal resolutions for 2 scenarios: first, precipitation and stream/soil water is aggregated to the same temporal resolution; sec-
ond, stream/soil water is kept at weekly resolution while only precipitation is temporally aggregated. Based on NSE values and uncertainties estimated by the GLUE method differences in TTDs and parameter results were evaluated and the conclusion drawn that the temporal sampling resolution can add uncertainty to TTD estimates. Finer resolutions were more similar, while coarser resolutions (>1 month) differed to a greater extent.

=> We thank Referee 2 for the comments which were very useful to improve the paper and prepare an improved version of the manuscript. We answer below to each comment in a point-to-point reply.

1. A major part of the paper is redundant information when compared to the publication of almost the same authors: "Understanding uncertainties when inferring mean transit times of water trough tracer-based lumped-parameter models in Andean tropical montane cloud forest catchments", published 24 April 2014 in HESS. On closer inspection, the same (not quite; see Specific Comments) data set was used. It is only natural that one data set can suffice to conduct research into many different topics. However, when major parts of this submission can be deleted without loss of unique information as they are already sufficiently and well presented in Timbe et al. (2014) (the Sections: Study area, Hydrometric measurements, Sampling scheme and isotopic analyses, Isotopic gradient of rainfall, Lumped parameter equation to infer mean transit times of water, Transit time distributions of water, Model performance; additionally, part of Results and Discussions), it starts to feel uncomfortable. Comparison to Timbe et al. (2014) make Table 1, Table 2, Figure 1, Figure 2, parts of Figure 8 redundant, while parts of Table 3 can be reconstructed by Table 4 in Timbe et al. (2014). More detailed comparison is given in the Specific Comments.

=> We acknowledge that the mentioned Sub-Sections, Tables and Figures contain similar information as presented in the paper Timbe et al. (2014). According to the suggestions of Referee #2 the data described in detail in the former publication has been removed. In this regard: Figure 1, depicting the geographic location of the study area, has been removed.
area, has been modified (see attached), and the Figures 2 and 3 integrated into one simpler figure (now Figure 2, see attached). The revised version of these figures will be included in the revised version of our manuscript. To avoid redundancy between the two papers, references to the former published paper will be included throughout the Material and Methods Sections instead of the rephrasing of text material and the recopying of the Tables 1 and 2.

2. The Discussion section "4.2 Comparison of distribution functions" deals to a great part with a comparison of TTD models, their differences and their possible physical interpretation. Starting from the beginning of this subsection, only 7 lines out of 56 mention sampling frequencies, and that is in a subordinate clause, until this section returns on-topic. This is not and should not be the aim of the paper. In fact, this was a goal of Timbe et al. (2014), quote "(ii) characterization of the dominant TTD functions". However, the conclusions for dominant TTD even change in this submission when compared to Timbe et al. (2014), keeping in mind that it is the same data and this is nowhere discussed (see Specific comments).

=> Findings from the previous study were useful and in fact necessary to successfully complete the current research, and permitted the authors to limit the research to the three best performing models (TTDs) and the data of 14 out of the 32 sampling sites. Doing so permitted us to go more in detail of the behavior and physical interpretation of those distribution functions for each type of water source and to quantify the effect of the sampling frequencies on the resulting MTTs: variation ranges of behavioral solutions and similarities/differences between sites and models. As such, the present analysis complements the one published in the previous paper (Timbe et al. 2014), but has a distinct different focus. Separating the two components, it is the testing of the TTD model and the assessment of the effect of the sampling frequency, into two distinct manuscripts permitted to shed light on the effect of each component on the resulting MTTs in a high degree of detail, which is unprecedented. The results allowed us to clearly separate and quantify the individual effect of the TTD model and sampling
frequency. Additionally, the followed approach presents a suitable method which will also serve as a reference for future work on MTT analysis, beyond the scope of the presented results.

While it’s correct stating that the characterization of the dominant TTD functions was one of the aims of the first paper, a clear preference of just one model per water type was beyond the scope of that paper. For instance when referring to the best distribution functions for stream waters it was concluded that, quote "models such as EPM or GM which have a better performance in terms of considering the non-linearity, in most cases yielded better fits to the observed data and at the same time better identifiability of its variables ($\tau$, $\eta$ or $\alpha$)”. Similarly for soil waters, LPM showed shorter uncertainty ranges, but GM higher fitting efficiencies. Hence, for the present research topic these three models were selected to check their sensibility to sampling frequency, respectively for stream and soil waters.

The message from both studies - dealing with sources of uncertainties when estimating MTTs and its TTD functions - is clear: For the given study site more research is needed in order to find the prevailing distribution function that mimics in the best way the complex behavior of every analyzed water type. Furthermore, future research should be based on the integrated use of more sources of information, such as chemical tracers, conceptual models, and longer data series or high-frequency sampling campaigns of rainfall-runoff events. In this context the present and previous research are very valuable in terms of providing a solid base for more elaborated sampling campaigns.

To highlight the new insights from this research from the ones found in former related research a paragraph was added at the beginning of Subsection 4.2 (Discussion Section/Comparison of distribution functions) detailing the insights gained on the characterization of the dominant TTD functions (see also the reply to comment #13).

3. One of the messages of the paper is that coarser data sets (starting with monthly resolution) perform more different than higher resolutions (e.g., daily). I am mostly
familiar with studies using weekly data, not so much monthly data. The authors could discuss this issue more with examples of studies that use monthly tracer data and are affected by this.

=> We added the following paragraph at the end of Discussion sub-section 4.1:

"For studies dealing with coarse stable isotope data sets (e.g., monthly or bimonthly), considering the differences of the performances between data sets of diverse sampling resolutions, the uncertainties associated to the predictions should be acknowledged and considered at the moment of the evaluation of hypotheses associated to these results. Monthly sampling resolution and monthly data is still frequently used in stable water isotope studies when either the effort or the costs are too high to realize a higher sampling frequency (e.g., Goller et al., 2005; Rodgers et al., 2005; Viville et al., 2006; Liu et al., 2007; Rock and Mayer 2007; Chen et al., 2012), which goes in line with a large share of observation points of the Global Network of Isotopes in Precipitation and Rivers (GNIP) of the I.A.E.A.-W.M.O."

4a. Some sentences are really long and convoluted and the use of comma interrupt the reading flow in places where it is unnecessary, making the paper not as easy to read as it could be.

=> After the addition/deletion of sentences/paragraphs in the manuscript (according to the observations of the reviewers) the paper has been checked for this specific error type.

4b. I have read and thought about this paper several times and compared it to Timbe et al. (2014). Although both studies investigate unique topics (sampling frequency vs. uncertainties), it is my opinion that they could have been combined into one paper given the amount of overlap. This theoretical paper would have dealt with sampling frequencies. It could have used all TTD models that were used in Timbe et al. (2014) and checked their performance based on the quality rules detailed in Timbe et al. (2014). Although the topic of sampling frequency is very interesting and of scientific relevance,
based on these considerations I recommend to reject this submission for publication in HESS for reasons of self-plagiarism and scientific discrepancies between the two papers using the same data set (with just a few months in-between these two papers). Should the editor decide otherwise, I hope that my detailed explanations of changes that need to be done are helpful to the authors in improving this manuscript.

=> While writing the follow-up paper on the previous published paper it was quite a delicate issue to find the bare minimum of information necessary to understand the second article on its own without knowing the first manuscript. In response to the raised concerns of Referee #2 we further minimized overlaps by deleting or modifying sentences/paragraphs, figures and tables from the Material and Methods sections. To achieve this we increased the number of cross references with previous published paper (Timbe et al., 2014), as to enable readers to know in detail how we assessed the effect of different TTDs on inferred MTTs, and to reduce the impression of any kind of "self-plagiarism".

At this point we like to highlight the opinion of Referee #2, that he agrees on the very interesting topic and the general scientific relevance of the presented work and stress that all other concerns of the two reviewers have been thoroughly addressed in the revised version of the manuscript.

The revision of the manuscript allowed us at one hand to further narrow the focusing of the research topic and yet have a stand-alone manuscript. An explanation about the slight different analyzed period between the present and previous research were implemented in the paper (see reply to specific comment #6).

Furthermore, we like to state that the interest and need for the present research topic arose after the findings from previous research. Hence, the combination of both researches into one paper or the splitting into two papers was not pre-conceived. Additionally, the combination of both researches into one paper would have been extremely difficult given the large number of sampling sites, the associated computer simulations
and scenarios. For this new research and according to its main objective only the more representative sites of the previous research were selected, in order to account for trends of over or under estimations as compared to reference values.

SPECIFIC COMMENTS

Comparison of this paper to Timbe et al. (2014) and comments on Figures and Tables. Left-hand side figure number is this paper, right-hand side figure number is Timbe et al. (2014):

5. Figure 1 with Figure 1: Inevitably the same study site is shown. This is no big deal and only natural to occur. The amount of overlap regarding the rest of the paper however makes this figure problematic, as the Timbe et al. (2014) version shows more detail and could have easily been used in a composite paper.

=> Given that the main objective of the present paper is to account the similarities or discordances between results from diverse sampling frequencies, for this research selected and representative sampling sites were considered when compared to Timbe et al (2014). The site selection was based on the characterization of MTTs and TTDs performed in the previous paper were data of more sampling sites were used. Although Figure 1 in both manuscripts look similar, the previous version of the figure showed more sampling sites which were not considered in the present research and version of Figure 1 (excluded sampling sites were: spring waters SFS, PLS, QRS; micro-catchments TP, Q3, stream water sampling point SF). Similarly, although soil waters were collected at three depths (0.10, 0.25, 0.40 m.) per sampling site (6 sampling sites x 3 depths = 18 sampled waters) as it was described in the former paper, in order to check the resulting trends from diverse sampling frequencies, in the second manuscript we considered that it would be enough to take just the water sample at one depth (0.25 m.) per sampling site. This was justified by the fact that spatial differences between MTT of sampling sites were larger than differences between the water samples at different depths in a given sampling site.
Figure 1 has been redesigned (see attached) to clarify the differences in sampling sites between the previous published manuscript (Timbe et al., 2014) and the current one. Besides, an explanation was included in Subsection 2.2 of the water sites and types analyzed in this paper as compared to previous work:

"Concentrating on the effect of different sampling frequencies on the inferred MTTs, the number of sampled streams was limited to the seven main nested sub-catchments (QN, QP, QC, QM, QR, FH, QZ) and the main catchment outlet (PL) (Fig. 1). Based on the results of Timbe et al. (2014), that differences between soil water sampling sites were bigger than on site differences, we further limited the number of soil depths under consideration from three to one specific soil depth at 0.25 m."

6. Figure 2 with Figure 2: These are almost identical. Striking here is a shift in time series. While the submitted paper is from 9/1/10 to 9/1/12, Timbe et al. (2014) is from 8/1/10 to 8/1/12. This is very confusing and goes hand in hand with discrepancies in both texts about the start of the sampling campaign and discrepancies found in similar tables in both papers. Why is the time series shifted?

=> The first 1.5 month (mid-August 2010 until late September 2010) of the sampling campaign used in Timbe et al. (2014) was neglected to homogenize the different time series for the aggregation into different sampling frequencies (up to 3 months during tryouts). The starting period for rainfall data and all other datasets was therefore shifted to the beginning of the last quarter in 2010 (first of October 2010).

In order to clarify the shift in sampling campaign in the present research as compared to that of Timbe et al. (2014) the following explanation was included in Section 2.3: "It should be noticed that when compared the time series used for this research with those used in Timbe et al. (2014), the period mid-August 2010 – September was neglected. This decision was taken in order to homogenize the different time series for the aggregation into different sampling frequencies (up to 3 months during tryouts) and to assure that divergences among predictions are only due to the temporal resolutions
used. The starting period for rainfall data and all other data sets (e.g., stream waters) was therefore shifted to the beginning of the last quarter in 2010 (1st of October 2010)."

On the other hand the final date of study period is the same for both studies (present research and Timbe et al., 2014): August 2012. Although in both referred figures data are shown until this date a possible cause of confusion can be attributable to that while for the current Figure 2 (where event data sampling is shown) the last tick mark for the x-axis and its respective label is showed for 9/01/2012, for the previous study (where weekly data was shown) the last labeled tick mark was 8/01/2012.

To avoid confusion with the similar figure published in the former paper, the Figures 2 and 3 in the current manuscript have been integrated. The redesigned figure (now Figure 2) clearly shows the seasonality of the event and monthly based Oxygen-18 isotopic signals together with daily rainfall. Reference to the former paper has been included in the text. Also the total samples of rainfall and stream water for the present study were checked and corrected.

7. Figure 3: There is no similar figure in Timbe et al. (2014), but this figure could be removed without drastic loss of information. It is nice to see how the aggregated time series look, but not absolutely necessary to understand this paper.

=> As requested by Referee #2 we deleted Figure 3, and Figure 2 was redesigned such that it shows now a high (event based) and coarse (monthly) time series (see also reply to previous comment).

8. Figure 8 with Figure 8: the left panels of soil water site C in this paper is the same as Timbe et al. (2014) panels a and b, with the Timbe et al. (2014) version showing more information due to more TTD models. This is another example of how the composite paper could have looked and that redundant results are presented.

=> We agree. Figure 8 has been deleted, and in the Discussion Section of the current paper reference to this picture in Timbe et al. (2014) has been included.
9. Table 1 with Table 1: This is a direct copy, with slight changes in formatting, with Timbe et al. (2014) showing more information.

=> We agree. Table 1 has been deleted and a reference to this table in the former published paper has been added in Section 2.1: "A detailed description of the physical, hydrological and land cover characteristics of the catchment and main tributaries are given in Timbe et al. (2014)".

10. Table 2 with Table 2: Not a direct copy, but very similar with lot of redundant and even mismatching information. While rainfall collection started in one case in August 2010 and was done by Collector, in the other case it started in October 2010 and was done Manually. This coincides with the text information given in the respective papers. Main Rivers and Tributaries collection dates also differ. Which is true? Why is there even a difference? Where is the exact difference between Manually and Collector for rainfall samples? This goes hand-in-hand with the question of the 1-month shifted time series and does raise some concerns.

=> We used “Manually” in order to distinguish from an automatic collector since during rainfall events, samples were taken manually using 1 L bottles (i.e., collector). As to prevent confusion, and according to the suggestions from Referee #2, the terminology regarding the collection of rainfall samples between both papers have been standardized to the terminology used in the first published paper (i.e., "Collector" instead of "Manually") and further Table 2 has been deleted. Besides, a reference to Timbe et al. (2014) has been placed together with an explanation of the shift between the time series involved in the current research as compared to the previous related study (see also reply to comment # 6).

11. Table 3 with Table 4: These tables deal both with soil water simulation results using the LPM model. As the time series is only shifted a month, they must be comparable, which they are to a certain extent. The Sampling frequency of “1 week” in this paper is the same setup as in Timbe et al. (2014). In most cases, the soil sites A to F of this
paper are almost (?) the mean value of the split-up version in Timbe et al. (2014), e.g., for parameter tau: B = 4.53 (4.50), C = 3.79 (3.93), D = 6.29 (6.20), E = 6.11 (6.60), F= 4.49 (4.87). Only soil site A is different with 5.69 (4.56). Why is A suddenly different in mean transit time using weekly data?

=> Explanation: In Timbe et al. (2014), at each sampling site (3 sites in pastures and 3 in forest), the analyzed soil waters were taken at three soil depths (0.10, 0.25 and 0.40 m) yielding a total of 18 sampled waters. For the present study we only considered the sampled waters collected at a depth of 0.25 m (one third of all soil waters sampled) since the spatial variability of predicted MTTs was larger than the variability between those of soil waters collected at different depths (Timbe et al., 2014). Thereby the intercomparison of the current weekly predictions should be performed directly with the results for the 0.25 m depths of the former published paper (instead, Reviewer #2 performed a comparison of the current results with the averaged results from the three depths for every site). In this sense, considering the weekly based predictions for MTT, the difference between the current prediction for soil site A at 0.25 m is 5.69 weeks vs 5.30 weeks of the former predicted result (representing a difference of 6.8% instead of 20% as it would be for the case of 5.69 vs 4.56). Keeping in mind that those results only represent the best predictive results, when compared to the uncertainty range for the particular soil water site and used model ($\Delta \tau = 1.75$ weeks, representing a 30.8% of variability, see Table 3 of the present research or Table 4 of former published paper), we consider that the difference between both referred simulations is small and attributable to less rainfall data used as input function for the models (for the present work one and a half months of the time series data of rainfall were not considered as they were for the previous published paper, as explained in the reply to the comment #6).

An explanation of the slight differences between predictions based on weekly data and those derived for the former paper has been included in the Discussion Section: "It should be noticed that, as a result of the use of a slightly shorter time series for the current research than those used for Timbe et al. (2014), slight differences for model
parameter predictions were found when weekly-based predictions are contrasted to the former published results.

12. Equations 3, 4 and 5 with Table 3: only natural that the TTD model equations are the same, but still, if you combined these to one paper, it would have saved space again.

=> Referee #2 is right, however in order to guarantee the reading of the paper as a stand-alone document we would prefer to keep the three referred equations.

13. Discrepancies of TTD models for soil and stream sites: Why is it that in Timbe et al. (2014) in the Conclusion it says, quote "In this sense, using the best predictions from models like LPM for soil waters and EPM for surface and spring waters yielded a more reliable range of MTT inferences through lowering the uncertainty associated in the predictions of certain models." while in this paper is reads, quote, "In this regard, for soil water the gamma distribution not only provides the highest goodness of fit but also more realistic and meaningful predictions; although for dampened isotopic signals (i.e. stream waters) a model preference is still not clear, [...].". This directly contradicts each other, and they were the findings of the same authors, in the same catchment, using the same data set and methods. This is not discussed in the paper.

Further, in this paper in the Discussion starting on p 12463 line 27: "For soil waters LPM yielded similar tau predictions than GM. However, a simple look at both distribution functions demonstrates that the gamma distribution function can provide more detailed information on how and when the tracer's signal increases/decreases and when the peak occurs". Yet, Timbe et al. (2014) prefers LPM for soil waters. This does not add up.

=> Please see reply to comment #2. Besides we will add the following paragraph to the Discussion section:

"The results of the current research provide new insights on the performance of the
analyzed TTD: 1) fitting efficiencies of model parameters can vary considerably depending on sampling resolution; 2) those efficiencies can be even greater for coarser than for higher model input resolution data; 3) at the same time, uncertainty ranges of model parameters can also vary greatly as a function of sampling resolution, being even possible that low resolution data sets display lower uncertainty ranges than higher resolution data sets. The mentioned findings lead us to consider that it is not possible to suggest the supremacy of one model over another based only on their performance (not necessarily the model with smaller uncertainties or neither the one with the highest efficiency); instead additional knowledge on the conceptual functioning of the studied system is necessary. For instance, in Timbe et al. (2014), where a weekly time step was considered, EPM and LPM predictions showed lesser uncertainty ranges (for stream and soil waters respectively) when compared to predictions provided by a GM model, which in counterpart provided better fitting efficiencies for most of the cases; however for knowing the best distribution function still further research is needed and meanwhile the use of a GM model cannot be discarded.

14. Definition of "tau" changes: 12452, line 12: tau is given as tracer transit time (agreed). On 12455, line 7 it is "tracer tau", 12457 onwards it is subsequently used as "MTT of the tracer (tau)" in the results discussion (e.g. for GM the median tau).

=> Has been checked and the same terminology and acronym is used throughout the manuscript.

15. p 12449, line 6: "The flashy reaction of the hydrograph [. . .]" where are those data coming from? From Timbe et al. (2014)? The source of the baseflow information should be cited here.

=> Baseflow was calculated in Timbe et al. (2014), and a citation to previous work has been included. The following sentence now reads as follows: "The flashy reaction of the hydrograph is due to rainstorms, while the slowly varying underlying trend corresponds to groundwater contribution (baseflow), which accounts for 85 % of the total runoff.
16. p 12452: Equation 1 is not necessary in this paper, as no radioactive tracer is used. It would save several lines of text if you only describe the equations you used.

=> We agree. The equation and text related to its description have been deleted.

17. p 12455: choosing of the second peak of the parameter that is not responsible for the MTT and writing that it will improve convergence of parameters needs a reference or written as assumption by the authors.

=> The referred text has been modified, and now reads as follows: "For cases with more than one solution peak, in order to improve the convergence of $\tau$, we restricted the behavioral solutions to the largest peak for the second model parameter (assumption made by the authors)."

18. p 12460, line 16: the value of 0.71 from bimonthly data actually looks to be the furthest away from the daily value of 0.63, when considering all sampling solutions. Why do you write it is close to it?

=> In fact they are not so different if the uncertainty ranges described for these particular parameter predictions ($\Delta \alpha = 0.14$ for each case) are considered. Although, we changed the sentence to consider the mentioned detail as follows: ". . . 0.71 (using bimonthly data) which is not far from 0.63, a value predicted using daily data, given that the range of behavioral solutions for this parameters is around 0.14 for every case."

19. p 12463, line 4: What are “mean conditions of soil water”?

=> Since soil water was collected on a weekly basis from containers that mixed all the water flowing towards them during that period of time, the isotopic signal of each sample represents a weekly mean condition of the soil water (i.e., the influence of particular rainfall-runoff events is not considered). To clarify we rephrased the related sentence. Instead of: "It should be noted that this finding is valid for groundwater systems or for mean conditions of soil water" now it reads: "It should be noted that this..."
finding is valid for semi-steady-state conditions of waters”.

20. p 12466, line 6-7: quote “For our study catchment, considering that NSE are high for all models and that TTD does not seem to influence their performance but greatly influences the predicted MTT, additional insights need be explored in order to unveil the correct TTD-function, as solely relying on model performances could lead to misleading results” Of course the TTD greatly influences the MTT, as the MTT is derived from the TTD. There is no more insight needed, as they are already results from Timbe et al. (2014). If they are not valid anymore, you need to give a reason.

=> The quoted sentence refers to the fact that conceptually different TTD (e.g. EPM vs GM, or LPM vs GM) could yield similar good performances in terms of fitting efficiencies, although they describe different flow characteristics. The mentioned insights needed are meant for additional information that could support or reject the characterization of the correct water processes through identifying the correct TTD. In order to clarify this, the related sentence was rephrased as:

"Considering that each TTD describe different flow characteristics, e.g. LPM vs. to GM, although they could yield similar performances, in terms of fitting efficiencies or uncertainties, for our studied catchment additional insights (e.g. tracer data associated with different flow paths) need be explored in order to finally unveil the prevailing TTD, as solely relying on model performances could lead to misleading results".

21. p 12467, line 3: What are micro-catchments?

=> We were referring to all the analyzed nested sub-catchments, however just two of them (QC and QM) have a small drainage area of around 1 sq km that could be considered as micro-catchments.

To avoid confusion we now rephrased the sentence as follows: "The study clearly demonstrates that estimations of the TTDs for catchments with similar characteristics in the same region using different frequencies of data sampling provides an additional
source of uncertainty" instead of "The study clearly demonstrates that estimations of the TTDs for micro-catchments in the same region using different frequencies of data sampling provides an additional source of uncertainty"

22. p 12467, line 11-13: Sampling frequency cannot account for time-variable conditions, how should it?

=> The suggestions was stated in a hypothetically way of speaking, it is the increase of the sampling frequency until a certain threshold could provide more and better insights of the behavior of a studied system (e.g., time-variable conditions). However, although we used a daily data resolution of rainfall as input function, the sampling frequencies for the outflows were conducted on a weekly basis and for steady-state conditions. Hence, we acknowledge that such affirmation is not supported by our present findings and to prevent confusion, and in agreement with the suggestions of Reviewer #1 and Reviewer #2 for this specific remark, we decided to delete this sentence.

TECHNICAL CORRECTIONS

We thank Referee #2 for all the technical corrections, which were carefully applied.

REFERENCES


Goller, R., Wilcke, W., Leng, M. J., Tobschall, H. J., Wagner, K., Valarezo, C. and Zech, W.: Tracing water paths through small catchments under a tropical montane rain forest in south Ecuador by an oxygen isotope approach, J. Hydrol., 308, 67–80,


Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 11, 12443, 2014.
Fig. 1. Figure 1
Fig. 2. Figure 2