

AUTHORS COMMENT: ANSWER TO REFEREE 1

'Drivers of the variability of the isotopic composition of water vapor in the surface boundary layer'

Referee comments: black,

Author comments: blue

Changes to the manuscript: green

General comments: This manuscript presents the isotope ratio data of atmospheric water vapor ($d_{18}O_v$ and dD_v) above a managed beech forest in central Germany. Together with EC measurements, values of $d_{18}O$ and dD associated with ET fluxes (dET) were also reported for a full growing season. The primary objectives of the study are to assess factors that are responsible for the observed variation in $d_{18}O_v$ and dD_v . The authors used a simple linear regression to seek for correlations between $d_{18}O_v$ and dD_v variability and an isolated variable and interpret their results on the basis of regression statistics (R^2 and p -values). As far as I can tell, the experiment was properly carried out and the data were carefully scrutinized and high quality. The topic is interesting to a broad audience especially to the stable isotope community and researchers who study ecohydrology. The measurements will contribute to a growing number of water vapor data collection, though a mechanistic interpretation of water vapor isotope data in the surface boundary layer remains challenging. The biggest issues I have with this manuscript are on the structure, assumptions made for the proposed problem, its statistical analysis and the interpretation of linear regression. The authors should consider addressing these major comments before its final publication.

Authors response: We thank the anonymous referee for the motivating, detailed and constructive feedback to our manuscript, below we answer the referee's comments in detail.

Specific comments on major issues:

1. This manuscript has two major purposes, 1) to demonstrate the ET fluxes do not dominate $d_{18}O_v$ and dD_v and 2) to evaluate potential factors that control $d_{18}O_v$ and dD_v variability on both diurnal and seasonal time scales. For the 1st objective, the authors treat the PBL as a box, and assume surface ET is the only flux component that contributes to volume of the box (or the diurnal evolution of boundary layer height) while neglecting horizontal advection and entrainment fluxes. They used 'isoforcing' associated with the ET fluxes, combined with PBL heights retrieved from ECMWF data product, to calculate dv/dt (for both ^{18}O and D) over the course of a day (Eq 2) and compared the results to the diurnal pattern from the time series measurements (Fig 2). Applying Eq 2 in this context is flawed as the height of the PBL cannot grow without entrainment (even if horizontal advection can be assume negligible under certain conditions). Assuming surface flux as the only flux component while applying a changing height (h) contradicts one another. This is a misinterpretation of the boundary layer budget, and it is no surprise that calculated and measured dv/dt diurnal pattern has nothing in common (Fig 2).

Authors response:

We thank the anonymous referee for pointing this out. We acknowledge, that the current version of the manuscript might not be sufficiently clear about the purpose, the assumptions, and the limits of this estimation.

We are well aware, that the calculation of $\frac{d\delta_v}{dt} |_{ET,est}$ based on equation (2) does not yield a real estimate of $\frac{d\delta_v}{dt}$ and in particular assuming a temporarily constant PBL height is only a theoretical assumption that is not justified for longer timescales. We want to clarify that temporarily changing PBL height changes the isotopic composition in two ways: 1) Entrainment of isotopically different

material from higher layers and 2) Changes in the relative fraction between the gas masses (i.e. dilution of the isoforcing signal over a larger volume). For our thought experiment we assume entrainment to not directly change δ_v by different material but only allow an influence of δ_v by dilution. The influence of local ET on δ_v is diluted by different PBL heights at different times of the day, thus throughout the day, IF/h reflects the influence of ET on δ_v in a boundary layer with a certain (slowly changing) height.

With the calculation based on equation 2, we quantify the influence of local ET on the isotopic composition of the boundary layer by making a quantitative thought experiment. More specifically, we do not aim at fully modelling δ_v (which is beyond the scope of our manuscript), but we want to answer the following question: How would local ET influence the delta value of the PBL (δ_v) if local ET would be the only process that (significantly) influences δ_v . More particular: We want to quantify/identify the influence of local ET for the theoretical case, that throughout the day entrainment would change PBL height as observed, but would not change δ_v . By doing so, we isolate the influence of local ET on δ_v . In particular, this reflects the influence of local ET better than assuming a constant PBL height and we find evidence that just using IF values alone is inappropriate to conclude about the influence of local ET on δ_v . This is discussed in Line 224 FF of the revised manuscript when we discuss possible interpretations of the diurnal cycle of isoforcing and in Line 267 FF, when we discuss that correlations between δ_v and δ_{ET} should not be over-interpreted.

Thus, this approach, despite its limitations, yields a quantitative estimate for isoforcing-related changes in δ_v which is more appropriate than directly using isoforcing values or using isoforcing values in combination with assuming a PBL height that is constant throughout the day. For this reason, we would like to keep this approach in the manuscript, but we revised the manuscript to be clearer about the purpose and the limitations of this approach, in particular, we rewrote section 2.4 of the manuscript. We changed the manuscript to explicitly mention that the quantity, which we now call $\frac{d\delta_v}{dt}|_{ET,est}$ is not the real change in delta value, but only a theoretical estimate of the influence of ET, which is only one out of many. In the revised manuscript we changed section 2.4 about the **Calculation of evapotranspiration-related change in δ_v** , to be clearer about the purpose, the assumptions and the limits of this estimation. Further, we differentiate strictly between $d\delta_v/dt_{meas}$ and $\frac{d\delta_v}{dt}|_{ET,est}$ to make clear that these two quantities are not the same.

2.4 Calculation of evapotranspiration-related change in δ_v

We quantify the influence of local ET on the isotopic composition of the boundary layer by making a quantitative thought experiment. How would local ET influence the delta value of the PBL (δ_v) if local ET would be the only process that (significantly) influences δ_v ? To answer this question, we use isoforcing values, that are based on EC measurements of the magnitude of ET F_{ET} and its isotopic composition δ_{ET} (see Braden-Behrens2019). We further assume a simple isotopic mass balance model (see e.g. Lai2006) with only one flux component (ET) from the surface and no influence of horizontal advection or entrainment on δ_v (see also Sturm2012, Braden-Behrens2019). If this assumption would be fulfilled, isoforcing IF can be interpreted as the rate of change of the atmospheric delta value multiplied by the temporarily constant boundary layer height h (see e.g. Lai2006).

$$IF = \frac{F_{ET}}{C_a \rho_a} (\delta_{ET} - \delta_v) = h \frac{d\delta_v}{dt}|_{ET,est}$$

$$\Rightarrow \frac{d\delta_v}{dt}|_{ET,est} = \frac{IF}{h} \quad (\text{Eq. 2 and 3})$$

With the evaporative flux F_{ET} , its isotopic composition δ_{ET} the atmospheric mole fraction C_a , the molar density of atmospheric air ρ_a , the atmosphere's isotopic composition δ_v and the height h of the planetary boundary layer (PBL).

We use Eq. \ref{eq:isoforcing2} to calculate $\frac{d\delta_v}{dt}|_{ET,est}$ for our measurements at different times of the day with a simultaneous estimation of the PBL height for each data point. As evident from Eq. 3 the influence of local ET on δ_v is diluted by different PBL heights h . Thus in particular as h changes throughout the day, $\frac{d\delta_v}{dt}|_{ET,est} = IF/h$ reflects the influence of ET on δ_v in a boundary layer with a certain (slowly changing) height. The resulting quantity $\frac{d\delta_v}{dt}|_{ET,est}$ yields a theoretical estimate for the influence of local ET on δ_v . However, the real change of δ_v is composed of changes related to many different drivers such as entrainment or horizontal advection see e.g.Griffis2007.

Secondly, why would the authors even bother to do this exercise? As later stated by the authors (In 189-190) “A discussion of the influence of local ET that is purely based on isoforcing IF overlooks the influence of boundary layer mixing processes.”

Authors response: Our sentence was not clear enough in the original manuscript. We referred to a discussion of the impact of ET based on isoforcing values IF versus a discussion based on $\frac{d\delta_v}{dt}|_{ET,est}$ (which includes a changing PBL-height). We do not refer to a general discussion of drivers of δ_v , but more specifically on the role of h when calculating the impact of local ET on measured δ_v . This has sometimes been discussed and estimated with assuming a PBL height that is constant on longer timescales. We change this sentence and add a discussion of the diurnal cycle of isoforcing: ‘Our data shows that a discussion of the influence of local ET that is purely based on isoforcing IF and does not include PBL height yields an over/underestimation of $\frac{d\delta_v}{dt}|_{ET,est}$. If we simply would assume a constant PBL height of eg. 1km, we would underestimate the influence of local ET for most of the times except around midday in spring and autumn. Further, if we would have used the diurnal cycles of isoforcing (see Fig. 1) as an indication for the influence of ET on δ_v throughout the day, we would have concluded that ET has the strongest influence on ET around midday. Our estimation of $\frac{d\delta_v}{dt}|_{ET,est}$ on the other hand shows a comparable magnitude in the mornings and in the evenings, while the comparison to $\frac{d\delta_v}{dt}$ shows that δ_v is dominantly driven by other processes such as entrainment around midday. Thus, we further conclude that due to the large variability of the boundary layer height h , it is essential to account for h when estimating the influence of local ET on ambient water vapor.’

The authors later (LINES 190 FF) stated that “the concurrent trends in the diurnal cycles of CH₂O and δ_v indicate, that entrainment dominantly influences δ_v from the forenoon to the afternoon:” and “: : we observe this indication for a dominant influence of entrainment from the forenoon to the afternoon

also in summer.” If you can make these conclusions from the observation (which the authors did), why trying to prove (and did it incorrectly) something that the data have already shown? This whole section should be scratched in my view.

Authors response: We still would like to keep the quantification of $\frac{d\delta_v}{dt}|_{ET,est}$, because we think it is more convincing to show both: a) a direct but only qualitative **indication** for the influence of entrainment by simply interpreting the shape of the diurnal cycles of C and δ_v and b) the quantitative estimation of $\frac{d\delta_v}{dt}|_{ET,est}$ in comparison with the measured $\frac{d\delta_v}{dt}$. This way, we can identify the magnitude of isoforcing related change in δ_v .

2. For the second objective, the authors identify 4 potential factors that influence seasonal availability of δ_v : local ET, Rayleigh distillation, selective water use by plants and temperature. The author applied a simple linear regression between δ_v and each of these factors to look for correlations. There are several problems in the statistical analysis used by the authors. First, the authors should distinguish processes from state variables. Secondly, these factors are not independent from one another, for example, ET and Rayleigh distillation are both temperature dependent. A simple linear regression ignores the interactive

effect between processes and state variables. Ideally, one should carry out a full BL budget calculation with a numeric model that considers thermodynamic isotopic fractionation. At the very least, the authors need to consider a multivariate regression that considers the interactive effect among variables. A simple linear regression is inappropriate.

Authors response: Thanks for this remark. We agree that the statistical analysis benefits from a multivariate regression. A full BL budget calculation that includes thermodynamic isotopic fractionation would be beyond the scope of this work. In the revised manuscript, we present a multivariate regression of the dataset.

3. The authors use sloppy statistics. This manuscript reports incredibly small p-values (10^{-35}) that are simply not meaningful. The p-value is calculated from the data and depends on the sample size (number of data points). It is possible to get p values to the -35 decimal points but that is simply because of we have the computing power to do this. More data points give you smaller p values. The bigger issue is, is the p value reliable? The p values shown in Table 3 are simply not meaningful. The difference in the p values between all times and period of green leaves is likely an artifact of sample size. Some statisticians have urged not to use p values but to use other alternative statistical matrix because it is too often misinterpreted (see Halsey 2019 <http://dx.doi.org/10.1098/rsbl.2019.0174>). This study is another example of why. The authors should limit reporting p values to a more reliable estimate.

Authors response: Thanks for this comment and for pointing out the interesting paper by Halsey et al. 2019. We fully agree that reporting such small p-values is not helpful. We will correct this and will use a p-value notation marking only ($p < 10^{-5}$) with a * and also give AIC numbers for multivariate regressions.

4. After redo the statistical analysis, the authors must re-evaluate their interpretation of the results and draw proper conclusion accordingly.

Authors response: We added a multivariate regression and changes the discussion accordingly.

Concerning ET as a potential driver, we find a negative dependency between $\frac{d\delta_v}{dt} |_{ET,est}$ and δ_v also in the multivariate regression. We discuss that this is physically not meaningful. Further, the multivariate regression did not yield a physically meaningful explanation with lower AIC than a simple correlation with temperature as the only driver. Thus, in the revised manuscript we focus more on this correlation.

Technical comments:

Ln 25. Do you mean a major driver of δ_v variability? Why the remove of precipitation only acts on seasonal time scales?

Authors response: Here we focus on Rayleigh distillation as a cumulative removal of rain from the atmosphere. We think this might not have been clear enough in the original manuscript. We changed this sentence to 'At seasonal time scales the cumulative rainout of an air mass as it ages from its origin (e.g by Rayleigh distillation) is a major driver of the variability of δ_v .'

Ln29. Your description of the amount effect is very crude and can cause confusion. Please be more elaborative on the amount effect.

Authors response: We rephrased the writing to be clearer about the complexity of the empirical amount effect, that can be a result of many different processes depending on the location. We further added some information on the influence of deep convection on the amount effect, as mentioned by Tharammal et al. 2017 JGR-A: 'These complex processes yield the 'temperature effect', a positive correlation between condensation temperatures and higher delta-values of precipitation (see e.g. Dansgaard1964) and the empirical 'amount effect', a negative correlation between the total amount and the mean isotopic composition of precipitation (see e.g. Dansgaard1964, Tharammal2017). However, the 'amount effect' can be a result of many different processes

depending on the location. For example the amount effect can be strongly moderated by deep convection (see e.g. Tharammal2017).'

Ln37-38. It's unclear what 'different importance' means based on R2 values; this sentence is hard to read. It's easier to see the effect by a state variable (such as temperature) but it becomes harder to visualize by a process (like Rayleigh rainout). Can you explain how Rayleigh process may differ seasonally that in turn affect seasonal variability of δv ?

Authors response:

With 'difference importance' we wanted to refer to the considerable differences in correlations between $\log(T)$ and δv that have been done by many other authors and that have been interpreted to reflect in how far the data could be explained by Rayleigh distillation. In the revised manuscript we explain this a bit more detailed: 'Thus, at different field sites, δv and $\log(T)$ are differently strong correlated. This indicates, that Rayleigh processes might play a dominant role in some cases (potentially explaining up to 78% of the variability) while in other cases other processes are more relevant (see also Huang2014 for details).'

We also add some more information and citations here to explain that Rayleigh distillation is only a very simple model for the cumulative removal of rain from the atmosphere: 'However, the removal of rain from the atmosphere by Rayleigh distillation is only a very simple model, while both, changes in the originating air masses and rainout processes are much more complex (see e.g. Noone2011).'

Ln51-52, a correlation does not suggest a causal effect; maybe that was not what you meant to suggest but the writing makes it seem that way.

Authors response: We agree and changed the writing to not imply that a correlation suggests a causal effect. New version: 'At seasonal time scales, some authors found evidence for a dominant role of Rayleigh processes (Lee2006, Wen2010).'

Ln61-62, consider revise this sentence to "Only one of these studies performed direct δ_{ET} measurements in a forest".

Authors response: We followed the suggestion and changed to 'Only one of these studies, the one by Huang2014, performed direct δ_{ET} measurements in a forest, however based on a flux-gradient approach, not eddy covariance.'

Ln99 pls provide more details on how exactly δ_{ET} was calculated. Did you perform a spectral analysis to examine potential loss of energy due to the differences in the sampling frequency between EC and isotope measurements?

Authors response: Thanks for this remark we agree that it is helpful to add some more details about the data evaluation to the manuscript – this might have been too short in the original manuscript.

Concerning the measurement frequencies of the different instruments, we add: 'We combined the 20Hz anemometer measurements with the 2Hz measurements of C_{H_2O} , $\delta^{18}O_v$ and δD_v yielding a 2Hz dataset of simultaneous measurements of isotopologue concentrations and 3D windspeed to calculate the magnitude and the isotopic composition of ET using the eddy covariance software EddyPro, version 6.2.0 LiCorBiosciences2016.'

Concerning data evaluation steps for flux calculations, we added:

'The used method to correct for high-frequency dampening, was based on the work of Ibrom2007, as recommended for closed path analyzers with loge tubing (LiCorBiosciences2016).'

Concerning the influence of the reduces measurement frequency, we add the following to the revised manuscript: 'In particular, we analyzed the influence of technical limitation such as the comparably slow measurement frequency of 2 Hz on water vapor flux measurements by additionally using 20Hz measurements of C_{H_2O} using a standard closed path CO_2 and H_2O_v analyzer (LI-6262 LiCor Inc., Lincoln, USA). We mathematically reduced its measurement frequency down to 2Hz see Braden-Behrens2019 and found that the resulting 2Hz dataset captured more than 98% of the variability of the 20Hz dataset (see Braden-Behrens2019).'

However, for a detailed description of the different and complex data evaluation steps, we refer to

our technical manuscript about EC measurements of CO₂ (Braden-Behrens2019).

Ln123 More precisely speaking, VPD is calculated from temperature and RH data which were directly measured.

Authors response: Thanks, for pointing this out, in the revised manuscript, we removed 'VPD' from this list, because it is not directly measured.

Ln128. Can you give a brief description on how the rain sampler is designed to store its water to prevent evaporation?

Authors response: Yes, we include the following to the manuscript:

'In brief, these rain samplers, reduce evaporation by minimizing the water surface exposed to the atmosphere. This is achieved by using a thin tube from the funnel down to the bottom of the sampling bottle and additionally using a very long and thin tube to adjust the air pressure in the sampling bottle, (see Groning2012).'

Ln135. Avoid jargon; just say using ECMWF data product

Authors response: We changed the whole section about PBL height and avoid jargon.

Ln145 delete this sentence **DONE**; rework this paragraph. Rather than copying from the manual, it would be more useful to describe how you retrieve PBL h from IFS.

Authors response: The whole section on PBL height has been reworked, replacing the quotes from the manual, and describing how PBL h has been retrieved from the IFS/ERA5 product. It should be clearer now that both PBL height as well as the associated random uncertainty is a product readily delivered as part of ERA5 rather than being derived in the current study. In addition, we have added information on the uncertainty of the product relative to radiosonde measurements and on the representativity of the grid cell of ERA 5 relative to the study site.

Ln179 what is the time unit here? Is this 0.1 permil per second, per minute or per hour? Assuming 0.1 permil per hour, from 8am to 5pm, δ_v would've increased by 0.9 permil d18O and almost 10 permil dD. But Fig 1 shows a decrease in d18O by _ 1 permil while a decrease in dD by _ 5 permil. How do you reconcile the inconsistency between these results?

Authors response:

Yes, the unit is per hour. We added this missing unit to the manuscript. The inconsistency that you are referring to is the difference between $\frac{d\delta_v}{dt} |_{ET,est}$. This is exactly what we refer to in line 179ff – but instead of focusing on δ_v (Fig 1), we focus on its temporal derivative $\frac{d\delta_v}{dt} |_{meas}$ in Fig. 2. We address this discrepancy in the followingsentence: 'The directly measured diurnal cycles of $\frac{d\delta_v}{dt} |_{meas}$ do not agree with this isoforcing-related estimate $\frac{d\delta_v}{dt} |_{ET,est}$ (see Fig. 2). In particular in spring and summer, we measure negative values of $\frac{d\delta_v}{dt} |_{meas}$ around midday, associated with a depletion of ambient water vapor, while the isoforcing-related change δ_v always yields an enrichment.'

We think this is now clearer after we distinguished more consistently between $\frac{d\delta_v}{dt} |_{ET,est}$ and between $\frac{d\delta_v}{dt} |_{meas}$ and also changed the axis label in Fig 2 accordingly.

Ln180-185 As stated above, this conclusion is flawed as the calculation was based on an invalid assumption of no entrainment while BL h is allowed to grow. **Authors response:** Please see our comment above.

Ln186-194 These remarks acknowledge the authors have known the answer from the observation but still decided to use a reverse logic to disapprove something they already knew could not be true. Hmmm interesting

Authors response: There are different aspects that might have caused unclarity here:

1. Here we distinguish between isoforcing IF and $\frac{d\delta_v}{dt}$. (which is based on Isoforcing, but includes dilution by the PBL). Please see our comment above, referring to lines 189-190.
2. As explained above (referring to line 190 ff), we draw our conclusions on the diurnal cycles shown in figures 1 and 2 because we think it is more convincing to show both: a) a direct but only qualitative **indication** for the influence of entrainment by simply interpreting the shape of the diurnal cycles of C and δ_v and b) the **quantitative estimation** of $\frac{d\delta_v}{dt}|_{ET,est}$ in comparison with the measured $\frac{d\delta_v}{dt}$. This way, we can identify the magnitude of isoforcing related change in δ_v . We think using both approaches is much more convincing than only discussing the shape of measured C and δ_v values.
3. The focus of our analysis was on quantifying the influence of ET for different timescales. We try to be clearer about that in the revised manuscript. Eg. We added the following to the introduction: 'We hypothesize, that at our measurement position, local ET is an important driver of δ_v at both, the seasonal and the diurnal time scale and use our direct measurements in combination with PBL height h to quantify the influence of ET on δ_v .'

We additionally restructured the paragraph (lines 186FF of the original manuscript) because we think this help to be clearer about the reasons for our conclusions.

Ln195-200 I found this section puzzling. TKE is a measure for the intensity of turbulence. h is most commonly defined by an inversion in potential temperature and dewpoint and is often estimated by radio sounding or lidar. It does not make sense to make direct comparison between TKE to PBL h (yes, they are both part of the boundary layer dynamics) because there is not a causal effect between the two. Simply presenting correlations without context is meaningless (if seeking covariation is the goal, why not presenting correlations with other meteorological variables? why do you choose to only present TKE? I would suggest removing TKE altogether.

Authors response: We removed this section/ the analysis of TKE from our manuscript.

Ln204 -12 permil for dDv ? Is this a typo?

Authors response: Yes, this was a typo. We changed it to to -88\permil

Ln206 Shouldn't selective water use by plants be included in ET?

Authors response: Thanks for pointing this out. We removed 'selective water use by plants' in this sentence.

Ln209 some would argue 7m above the top of the canopy is pretty far out; it is likely outside the subsurface BL near the forest canopy. Since you mention TKE, why don't you show a profile of vertical wind speed and momentum fluxes? It will give you an idea if your sensors are within the canopy subsurface BL.

Authors response: We agree that this would be interesting, but we do not have wind profile data available for this cite.

Ln222-223, do you have an explanation of why you found a correlation between dET and dDv but not with $d18Ov$?

Authors response: We think this might be related to the signal to noise ratio, that is better for dD than for $d18O$. We added this hypothesis to the manuscript.

Ln225-234 these interpretations are based on flawed stats

Authors response: Please see our comment above – we changed to multivariate regression.

Ln238. Bowling et al. 2017 is not an appropriate citation here. Remove. **Authors response:** We removed this citation.

Fig 7. Right panel: after leaf coloring - was that diamond or cross symbol?

Authors response: We changed the diamonds to crosses.

Ln243-248 & Fig7. Are GMWL and LMWL statistically different? I am skeptical that you can use GMWL and LMWL to contrast impacts by far-field v.s near-field factors.

Authors response: The LMWL is 7.4 ± 0.3 . Thus, we have a 2-sigma deviation from the GMWL. If we assume a standard distribution of errors, this yields $p < 0.05$. In the revised manuscript we explicitly mentioned the 2-sigma derivation to the interpretation: 'Thus the GMWL with a slope of 8 is at a 2-sigma difference away from the LMWL, yielding a p-value of $p < 0.05$.'

Ln254. entrainment is a diurnal process. Why would you expect entrainment be a factor on seasonal time scales in the first place?

Authors response: After reading all the referee reports, we removed the analysis of TKE and u^* from our manuscript. This involved also removing this line. However, originally, we wrote this sentence because entrainment integrated throughout the day can be differently strong on different days, this would yield seasonal variability.