

## Reply to reviewer 2

In what follows, we respond to the individual comments and recommendations of reviewer 2, R2. These responses are keyed to the specific comment by numbering, and are given in blue print, followed by indications of the changes made in the manuscript (in italics), and referring to the position in the original manuscript. Also, we revised again the entire manuscript for clarity, paying close attention to all of the reviewers' comments.

### R2 1

Hirl and coauthors present an impressive data set of seven years of isotopic observations in a grassland and an equally impressive modelling effort of the data.

We thank the reviewer for the encouraging and thought-provoking comments that helped us much to improve the presentation of our work.

### R2 2

The interpretation of the data is regrettably only discussing the isotopes and gives very little insight into the water fluxes of the ecosystem.

Yes, our results and interpretations centre on  $^{18}\text{O}$  of water in the different ecosystem components, although we do present model predictions of canopy conductance and transpiration as a function of soil water content and leaf-to-air vapour pressure deficit (Fig. 4c and 5c), we compare measurements and predictions of latent heat flux (Fig. S1), and we make predictions of the soil depth distribution of water contents and root water uptake (Fig. S12). This approach was dictated by the main focus of the work that consisted in systematically evaluating our (eco)system-scale understanding of the propagation of the  $\delta^{18}\text{O}$  signal of rainwater through soil water, root water uptake and  $^{18}\text{O}$  enrichment of leaf water (as specified in the Introduction, P3 L30ff), by comparison of model predictions and observations. In that sense, our work is 'restricted' to the  $^{18}\text{O}$ -ecohydrology of the system, as we explore how the different hydrological properties of the system (given by the parametrization of the model) dictate the dynamics of  $\delta^{18}\text{O}$  of water with depth in the soil, water taken up by the vegetation, and enrichment in the leaves. We believe that this is the most novel aspect of this work, and also the topic that we can support/validate/evaluate best with observations. Thus, our paper demonstrates how knowledge of  $\delta^{18}\text{O}$  of distinct water pools can help us to assess the ability of a locally-parameterized  $^{18}\text{O}$ -enabled mechanistic soil-plant-atmosphere model in predicting the hydrology of a system. For instance, the fact that the  $^{18}\text{O}$ -enabled hydrology inside MuSICA predicted well the observed  $^{18}\text{O}$ -dynamics at different depths in the soil and in the water taken up by the root system indicates strongly that the ensemble of model parameters also predicted well the spatio-temporal dynamics of soil water contents (including emptying and refilling dynamics) and depth distribution of root water uptake. We hope that this approach – when developed further – can also be helpful later on for the ecohydrological interpretation of  $\delta^{18}\text{O}$  in biological archives (e.g.  $\delta^{18}\text{O}_{\text{cellulose}}$  extracted from Park Grass Experiment hay samples originating from the last century).

To address the point in the revision, we

- added the definition of the term  $^{18}\text{O}$ -ecohydrology and its' objectives and potential (P2 L5). *“This science, that explores relationships between the spatio-temporal dynamics of water in the soil-vegetation-atmosphere system with help of the dynamics of  $\delta^{18}\text{O}$  of water in its different components, may be termed  $^{18}\text{O}$  ecohydrology.”*
- improved and expanded the discussion/interpretation of soil water dynamics and root water uptake (see below, and responses to reviewer 1, Matthias Beyer).

### R2 3

For example, if main water uptake is always at 7 cm depth even when this layer falls dry, then ecosystem transpires probably less than possible during this times because it would have access to more water in deeper soil. How is the ecosystem reacting? Is it shutting down the stomata? Is it

changing its carboxylation capacity and stomata close thence? Or both? And why would a grassland do this? I guess it is well established in trees that they would harvest deeper soil water.

Yes, these are important points, that we address in 4.2 (revised, see below). See also responses to MB and relevant changes made, above.

Being restricted to only 2 depths, the spatial resolution of our observations of  $\delta^{18}\text{O}_{\text{soil}}$  is limited, and there are methodical issues on the precision for estimation of the depth of root water uptake from such observations alone. Here, the (locally parameterized) hydrological model inside MuSICA does help. This predicted that root water uptake occurred over a broader zone (Fig. S12), with a mean (uptake-weighted) depth of root water uptake above a soil depth of 15 cm in 90% of all sampling dates (new Fig. S13).

We had no observations of stomatal conductance and carboxylation capacity, that would allow us to address their responses to drying soil. However, the model did consider an effect of soil drying on stomatal conductance (dependent on predawn leaf water potential) (P7 L22-24). The predicted effect of that is displayed in Fig. 5c. The sensitivity analysis did show that predictions of  $\delta^{18}\text{O}_{\text{soil}}$  at the different depths was responsive to stomatal conductance. Therefore, the generally good agreement between observed and predicted  $\delta^{18}\text{O}_{\text{soil}}$  did suggest that the ensemble of (photosynthetic and hydrological) model parameterization predicted the spatio-temporal variation in SWC and root water uptake quite well.

Interestingly, the model also predicted that SWC were occasionally lower below 25 cm than above that depth, particularly when rainfall recharged the top soil, but was insufficient to recharge the soil at greater depths (Fig. S12). Such phenomena occurred relatively frequently in the second half of the growing season. That fact could contribute additionally to explain why root water uptake occurred mainly from shallow soil depths (i.e. <20 cm below soil surface).

Certainly, the shallow root distribution also dictated a shallow depth of root water uptake. That shallow root distribution probably resulted from morpho-physiological constraints, particularly in the grasses and white clover (which comprised about 90% of the total pasture vegetation): in these species, adventitious roots compose virtually the entire root system, and root turnover is rapid and connected with leaf turnover at phytomere level (Yang et al., 1998; Robin et al., 2010) and assimilate supply to roots is reduced when grazing pressure is high (e.g. Bazot et al., 2005). In addition, the extremely high nutrient demand of frequently-defoliated vegetation is another factor that contributes to explain the formation and maintenance of a very shallow root system, as virtually all nutrient returns (mainly excreta from the grazing cattle) occur superficially.

We revised the entire manuscript for clarity concerning the above issues and, particularly, revised rigorously the first part of section 4.2, which now reads:

*“The comparison of observed  $\delta^{18}\text{O}_{\text{stem}}$  and  $\delta^{18}\text{O}_{\text{soil}}$  (Fig. 3a) strongly suggested that root water uptake occurred mainly at shallow depths (<20 cm) throughout the vegetation periods, largely independently of changes in SWC. That interpretation of observed data was based on comparison of  $\delta^{18}\text{O}_{\text{stem}}$  and  $\delta^{18}\text{O}_{\text{soil}}$  at two depths (7 and 20 cm) only, which provides limited spatial resolution and cannot inform precisely on the depth of root water, if  $\delta^{18}\text{O}_{\text{soil}}$  does not change monotonously with soil depth (Rothfuss and Javaux, 2017; Brinkmann et al., 2018). Such information can be improved by a locally-parameterized, physically-based,  $^{18}\text{O}$ -enabled ecohydrological model, as shown here. For instance, the standard MuSICA runs (Fig. 3b) indicated near-monotonous increases of  $\delta^{18}\text{O}_{\text{soil}}$  between 20 and 7 cm depth, matching well the observations in the majority of sampling dates (Fig. S13). Further, the simulations predicted a mean (uptake-weighted) depth of root water uptake at <15 cm in 90% of all sampling dates, independently of SWC and observations of  $\delta^{18}\text{O}_{\text{soil}}$ . Support came also from the MuSICA sensitivity analysis (Fig. 6h) in showing that  $\delta^{18}\text{O}_{\text{stem}}$  was well predicted by the model only when root length density was maximum at shallow soil depth. The potential range of rooting depths is large in grassland, depending on site, species, climatic and management effects (Schenk and Jackson, 2002; Klapp, 1971). So, why was root water uptake constrained to shallow depths in this drought-prone permanent grassland system? Several factors likely contributed: (1) the shallow top-soil overlying calcareous gravel (Schnyder et al., 2006), (2) the rapid shoot and root biomass turnover, that is associated with high phytomer dynamics leading to short leaf and root lifespan in intensively managed grassland (Schleip et al., 2013; Yang et al., 1998; Auerswald and Schnyder, 2009; Robin et al., 2010), (3) the high rates of shoot tissue (mainly leaves) losses that elicit a priority for assimilate*

(including reserve) allocation to shoot regeneration at the expense of the root system (e.g. Bazot et al., 2005), and (4) predominant placement of the root system near the soil surface dictated by the high need for nutrient interception and uptake (e.g. from excreta deposits), to compensate the high rates of nutrient losses due to grazing (Lemaire et al., 2000). Importantly, (5) in a relatively high number of cases, the model predicted situations in which rainfall recharged mainly the top soil, while SWC at depths >20 cm remained low (e.g. June-end of year 2006, April-October 2007, or May-end of year 2008; Fig. S12; see also below). Principally, however, factors (2)-(4) alone can explain why shallow rooting depth is a typical feature of intensively grazed grasslands (Troughton, 1957; Klapp, 1971). Also, Prechsl et al. (2015) did not find an ...”

## R2 4

Are any of the other variables telling me something about the ecophysiology of the plants or the ecohydrology of the ecosystem? Are leaf water isotopes telling me something? They tell me at least that there is nighttime conductance. Is there also nighttime transpiration? Anything else?

In the main, the ecophysiology of the plants and the ecohydrology of the ecosystem is reflected in the parameterization of vegetation and soil in MuSICA (Methods S2, Table S1, Figure S5, S6, S8), with many parameter values obtained from local measurement. The spatio-temporal dynamics of root water uptake (Fig. S12), and canopy conductance (Fig. 5c) and transpiration rate (Fig. 4c) at midday predicted by MuSICA are a result of that parameterization.

And yes, the diurnal  $\delta^{18}\text{O}_{\text{leaf}}$  data indicate that stomates were not completely closed during the night (P7 L18-19), a factor that was reflected in the parameterization of MuSICA (Table S1). Yet, predicted night-time transpiration (estimated by latent energy flux) was always very low, in agreement with the eddy flux data (Fig. S1) and the generally high nocturnal relative humidity.

We did not have the detailed ecophysiological and ecohydrological observations to validate those specific predictions. However, we did validate MuSICA for the evapotranspiration (i.e. latent heat flux) predictions, and estimations of plant-available soil water in the entire top-soil (see also changes made in response to reviewer 1).

Most importantly, the good agreement between observed and predicted  $\delta^{18}\text{O}$  in soil (at 7 and 20 cm depth), stem and leaf water does indicate that the model described the ecohydrology of the grassland system well.

In the revision, we added several sentences and phrases, clarifying those points (see also responses to reviewer 1):

P5 L27ff: “The model was validated with latent energy flux (LE) data obtained from an eddy covariance station (EC) at the site. According to that comparison (Fig. S1), MuSICA estimates were unbiased ( $LE_{\text{MuSICA}} = 0.997 LE_{\text{EC}}$ ;  $R^2 = 0.59$ ). Further, we compared MuSICA predictions of total plant-available soil water (PAW, mm) in the entire top soil with PAW modelling and data for the same site presented in Schnyder et al. (2006). For the 2007-2012 data, this yielded the relationship  $PAW_{\text{MuSICA}} = 0.99 PAW_{\text{Schnyder et al. 2006}} + 7.8$  ( $R^2 0.83$ ).”

P7 L20ff: “Although the diurnal pattern of  $\delta^{18}\text{O}_{\text{leaf}}$  (Fig. S7) indicated some nocturnal stomatal conductance, the model generally predicted very low nighttime transpiration, in agreement with the eddy flux data (Fig. S1) and the generally high nocturnal relative humidity.”

P 14 L12ff: These ecohydrological processes are described explicitly in MuSICA, and agreement between observations and predictions of  $\delta^{18}\text{O}_{\text{stem}}$  and  $\delta^{18}\text{O}_{\text{soil}}$  at 7 and 20 cm depth indicates that MuSICA is capable of simulating these ecohydrological processes including  $^{18}\text{O}$  of the different water pools.

And P15 L4ff: “Predictions of  $\delta^{18}\text{O}_{\text{soil}}$ , particularly below the main zone of most water uptake, at 20 cm, were influenced markedly by estimates of LAI...”

## R2 5

I think, therefore, that the claim in the conclusions that the "work highlights the usefulness of mechanistic  $^{18}\text{O}$ -enabled modelling for explorations and quantitative analyses of the ecohydrology of ecosystems." is premature because only point (2) of the three points, i.e. root water uptake is actually ecohydrology of the ecosystem. The other points are about  $^{18}\text{O}$  ecohydrology, as the title of the paper suggests.

We understand the criticism, which is – we believe – partly due to our omission of a clear definition of  $^{18}\text{O}$ -ecohydrology, and the objectives of its application in the present context.

In the revision, we added the definition in the Introduction. Here, we employed the ecohydrological model implemented in MuSICA to predict the  $\delta^{18}\text{O}$  of water at different soil depths, the  $\delta^{18}\text{O}$  of water taken up from the soil, and the  $^{18}\text{O}$ -enrichment in leaves. In that we also evaluated several methodical/conceptual,  $^{18}\text{O}$ -ecohydrological uncertainties impacting on such predictions, such as (1) the choice of the water vapour effective diffusivity in the soil (Moldrup vs Penman), (2) the source of the rain water and atmospheric vapour  $\delta^{18}\text{O}$  (local data vs IsoGSM estimations), and (3) alternative leaf water- $^{18}\text{O}$ -enrichment models (two-pool vs Péclet). The capability of the model to predict the  $\delta^{18}\text{O}$  of the different water pools then indicates that the model is equally capable to predict the different ecohydrological processes (that underlie the  $\delta^{18}\text{O}$  predictions and observations).

Also, we revised all text carefully to eliminate any opportunity for misunderstandings. In particular, we highlight that a physically-based  $^{18}\text{O}$ -enabled ecohydrological model (as implemented in MuSICA) can provide insight in ecohydrological processes, such as the spatio-temporal dynamics of soil water and root water uptake, and transpiration or canopy/stomatal conductance. Concerning the latter, we find it interesting that dry soil conditions (under similar atmospheric conditions) led to increased  $^{18}\text{O}$ -enrichment (on average) in both the observed and predicted data, although evidence for a Péclet effect was missing in our data.

In the revision, we made the following main corrections, additions and deletions:

Abstract

P1 L16: *“Using the ecohydrology part of a physically-based,  $^{18}\text{O}$ -enabled soil-plant-atmosphere transfer model (MuSICA), we evaluated our ability to predict the dynamics ...”*

P1 L18: *“The model accurately predicted the  $\delta^{18}\text{O}$  dynamics of the different ecosystem water pools, suggesting that the model generated realistic predictions of the vertical distribution of soil water and root water uptake dynamics. Observations and model predictions indicated that water uptake occurred predominantly from shallow (<20 cm) soil depths ...”*

Introduction

P2 L5: *“This science, that explores relationships between the spatio-temporal dynamics of water in the soil-vegetation-atmosphere system with help of the temporal dynamics of  $\delta^{18}\text{O}$  of water in its different components, may be termed  $^{18}\text{O}$  ecohydrology”.*

Conclusion

We deleted the Conclusions section (see also response to reviewer1, MB 39)

**R2 6**

I have to admit that I had problems with the sensitivity analysis. Firstly, the mean difference is not a good measure. Differences can cancel out even when the model reacts strongly to a change. Most people use variance, standard deviation or root mean square error to avoid this. I guess that would be something like the error bars in Fig. 6. Secondly, one can of course use "arbitrary" ranges of model parameters to look at the output range, but then one cannot compare anymore the output ranges between the different parameters as done in Fig. 6. One wants to disturb each parameter similarly. So a derivative would probably be a good idea, or an elasticity.

We understand the point raised by the reviewer. We realize that our description of the sensitivity analysis and of Fig. 6 lacked some precision, and we improved that in the revision (see below).

We like to emphasize that our sensitivity analysis presents two types of (sensitivity) information/variables: (1) the mean sensitivity to a change of a parameter value (upper or lower value) on the metric of interest (e.g.  $\delta^{18}\text{O}_{\text{leaf}}$ ), in relation to the standard simulation, as shown on the x-axis as

‘mean sensitivity’, and (2) the standard deviation of the sensitivity (given by the error bar). The standard deviation captures the variability of the response to a parameter change between the individual sampling occasions. If cancelling effects result from the change of a parameter value, resulting in a mean sensitivity close to zero, that cancelling behavior is revealed by the (size of the) standard deviation of the sensitivity (e.g. the effect of the upper parameter value on  $\delta^{18}\text{O}_{\text{leaf}}$  in panel 6h). Panel 6a reports a very different type of behavior, as changing the parameter value caused no cancelling effects on  $\delta^{18}\text{O}_{\text{leaf}}$  (as was indicated by the small standard deviation of the sensitivity), but a strong change in the mean sensitivity for  $\delta^{18}\text{O}_{\text{leaf}}$ . So, there were instances where changes of parameter values caused a ‘general’ effect (causing a positive or negative mean sensitivity) and instances where there were strong cancelling effects (leading to a large standard deviation of the sensitivity). Both types of sensitivities can be gleaned from our presentation of parameter sensitivities.

Thus our sensitivity analysis revealed four different types of sensitivities: (a) strong mean sensitivities, with no or little cancelling (e.g.  $\delta^{18}\text{O}_{\text{leaf}}$  in panel 6a), (b) mean sensitivities combined with strong cancelling effects (e.g.  $\delta^{18}\text{O}_{\text{leaf}}$  in panel 6c), (c) no mean sensitivities resulting from strong positive and negative cancelling effects (e.g.  $\delta^{18}\text{O}_{\text{leaf}}$  in response to the high parameter value in panel 6h), and (d) absence of a mean sensitivity without cancelling effects (e.g.  $\delta^{18}\text{O}_{\text{stem}}$ ,  $\delta^{18}\text{O}_{\text{soil } 7}$  and  $\delta^{18}\text{O}_{\text{soil } 20}$  in panels 6a and 6b).

Although we like the idea of calculating elasticities, in principle, we did see some problems:

1) The  $\delta^{18}\text{O}$  values are not ratio-scaled (but interval-scaled) and the zero value (0‰) is not an absolute zero, resulting in problems when comparing parameter effects on the  $\delta^{18}\text{O}$  of the different water pools.

2) ‘Elasticity’ quantifies the percentage change of the output variable in response to a given percent change in the input parameter. This does not consider if a given percent change in the input parameter is hydrologically or physiologically plausible or relevant (particularly when model sensitivity is compared for different parameters).

3) It may not be possible to draw universally valid conclusions from the elasticity. In case of a non-linear response of the variable under study, elasticity depends on the extent of change of the parameter. Yet, varying parameters by the same percentages, e.g. by +50% and by -50%, in order to ‘disturb each parameter similarly’, would neglect morpho-physiological or system knowledge on the ‘realistic’ (or ‘plausible’) range of values for each parameter. So, changing a parameter by a certain percentage is likely a more arbitrary choice than the one that we have taken.

Point 3) is also valid for derivatives.

Regarding the second point of the reviewer “one can of course use "arbitrary" ranges of model parameters to look at the output range”:

This is a point that we had discussed extensively, during the work and preparation of the submitted manuscript. In effect, we did not use arbitrary values. Instead, we chose the upper and lower parameter values based on the range of values observed at the site (LAI, canopy height, mesophyll water content), ranges dictated by physical constraints of the system (root distribution), the origin of the  $\delta^{18}\text{O}_{\text{rain}}$  data (IsoGSM predictions as opposed to local measurements), or – where we did not have own measurements – based on the range found in the literature for grasses/grassland ( $\varphi$ ,  $m_{\text{gs}}$ ,  $g_0$ ,  $V_{\text{cmax}}$  and  $J_{\text{max}}$ ). In that way we ascertained realistic and physiologically meaningful upper and lower parameter values in the sensitivity analysis. In a way, this also dictated that we refrain from calculating elasticities.

On the basis of these facts and considerations, we would like to retain the approach to sensitivity analyses presented in the original manuscript. However, we did take the reviewer’s comment/concerns very seriously and improved the presentation and description of the approach. This included: renaming the ‘mean difference’ by ‘*mean sensitivity*’ (which is more appropriate and illustrative) and standard deviation of the difference by ‘*standard deviation of the sensitivity*’, and explaining the rationale for the choice of this specific form of sensitivity analysis.

The legend to Fig. 6 now reads:

“Fig. 6: Sensitivity of modelled midday  $\delta^{18}\text{O}$  of leaf, stem and soil water at 7 and 20 cm depth to various parameters of the MuSICA model. The sensitivity was tested by varying one parameter while

keeping all other parameters the same as in the standard MuSICA parameter set (Table S1), as detailed in 2.5. Sensitivity (parameter effect) was quantified by two variables: the mean (or average) sensitivity (in ‰) resulting from the change of a parameter value relative to the reference run, and the standard deviation of the sensitivity which captures the variability of the response to a parameter-change for the different sampling times (displayed by error bars.) Strong averaging (cancelling) effects resulting from the change of a parameter value are revealed by large standard deviations of sensitivities. Note that the sensitivity analysis revealed four different combinations of parameter effects: (a) strong mean sensitivities, without cancelling effects, (b) strong mean sensitivities superposed with strong cancelling effects, (c) small mean sensitivities resulting from strong cancelling effects, or (d) absence of sensitivities unrelated to cancelling effects. Parameter identity is given in the upper left corner of each panel. In (a) to (h), blue down-pointing triangles refer to the low parameter value, red up-pointing triangles to the high parameter value of a sensitivity run, based on the range of values observed at the site or – where such values were missing – the range of reported values for grasses or grassland in literature (see Materials and Methods). In (i) the Moldrup submodel for the water vapour effective diffusivity in the soil was replaced by the Penman model. In (j) we used IsoGSM-predicted  $\delta^{18}\text{O}_{\text{rain}}$  and  $\delta^{18}\text{O}_{\text{vapour}}$  data instead of locally determined  $\delta^{18}\text{O}_{\text{rain}}$  and  $\delta^{18}\text{O}_{\text{vapour}}$  data for the isoforcing of MuSICA. Note that the low parameter value for Péclet number (a) predicted a far greater deviation of  $\delta^{18}\text{O}_{\text{leaf}}$  than any other parameter.

The relevant section of 2.5 was revised accordingly (P9 L21ff):

“Parameter effects (sensitivities) were quantified by two variables: (i) the mean sensitivity relative to the reference run, obtained as the mean differences from the reference run as  $(\sum_{i=1}^n (\delta_{\text{sens},i} - \delta_{\text{ref},i}))/n$ , with  $\delta_{\text{sens},i}$  the  $\delta^{18}\text{O}$  of a given water compartment (leaf, stem, or soil at 7 or 20 cm depth) in a sensitivity run and  $\delta_{\text{ref},i}$  that in the reference run, for a day  $i$ ; and (ii) the standard deviations of the sensitivity, obtained from the differences between  $\delta_{\text{sens},i}$  and  $\delta_{\text{ref},i}$ . The latter illustrated how strongly the effect of a parameter varied between sampling days, and hence how strongly it depended on the conditions encountered on one specific day. Thus, the sensitivity variables (mean and standard deviation of sensitivity) reported if changes in parameter values caused systematic/general effects (shown by the mean sensitivity), or cancelling effects (shown by the standard deviations of the sensitivity), or combinations, or lack of the two.”

Also, paragraph 3.6 and 4.2 were revised for consistency.

## R2 7

Lastly, the authors suggest that there is no Péclet effect but rather a second unenriched water pool. While the data seem to support this, I would have expected a much better discussion.

I cannot find any mentioning of the 2D formulation of Farquhar and Gan (2003) while this should probably be the correct model. For example, what would be the effect if the leaf followed exactly this 2D model but the leaves were sampled only partly, not sampling the least enriched part?

We sampled the entire leaf blades and the entire exposed part of the growing leaf blade of grasses, (which was a minor component of the total sample), and trifoliolate leaves of white clover. In the case of *Taraxacum officinale*, we included half a leaf blade, severed along the length of the midrib. With that sampling protocol we integrated (but did not resolve) the entire gradients of evaporation-related  $^{18}\text{O}$ -enrichment that occurred within the individual leaf blades, permitting (and restricting us to) the use of the whole-leaf version of the  $^{18}\text{O}$ -enrichment model used to evaluate the occurrence of a Péclet effect. With that protocol, it was not possible to use the theory presented in the 2D formulation of Farquhar and Gan (2003); hence we used the non-steady-state version of the Péclet model, which is equivalent to that used by Gan, Wong, Yang and Farquhar (2003) for their experimental whole-leaf data.

In the revision we improved the respective paragraph, which now reads (P4 L27ff): “Each leaf sample included all leaf blades, including the exposed part of the growing leaf, but excluding senescing leaves (cf Fig. 1 of Liu et al., 2017) from each of two vegetative tillers of *D. glomerata* and 16 vegetative tillers of *L. perenne*, *P. pratensis* and *P. pratense*, one half of a leaf blade of *T. officinale* (with the latter severed along, but not including, the mid-vein) and two trifoliolate leaves of *T. repens*. This

*protocol ensured collection of the entire within-leaf evaporative  $^{18}\text{O}$ -gradient of all sampled leaf blade tissue of the different species.”*

## **R2 8**

The very small discussion starts with the possibility of xylem (or associated tissues) water and non-steady state but then only talks about the latter. I would have loved to see insights about grass blade anatomy, especially from this group who knows it that well.

We did not collect data on the anatomy of sampled leaves, as this was impractical (see also response to R2 9, below).

## **R2 9**

I also do not follow the argument that there is no non-steady-state effect in the missing correlation with transpiration because the model includes non-steady state. The model yes, the data no. Margaret Barbour’s group also claimed to see no Péclet effect but if they plotted their data against the isotopic composition of transpiration rather than xylem, the Péclet effect re-emerged.

Yes, correct, the model included non-steady-state. A significant fraction of the observations originated from non-steady-state conditions, others appeared to be close to steady-state (Figure S9).

In the revision, we looked at the subset of observations that exhibited seemingly near-steady-state  $^{18}\text{O}$ -enrichment (about half the data) to verify additionally if the relationship between the proportional difference between observed leaf water  $^{18}\text{O}$ -enrichment ( $\Delta^{18}\text{O}_{\text{leaf}}$ ) and evaporative site enrichment ( $\Delta^{18}\text{O}_e$ ) predicted by the Craig-Gordon model ( $\Delta^{18}\text{O}_{e,ss}$ ) would indicate the existence of a Péclet effect for that subset. Again, we did not observe evidence of such an effect.

In the revision, we deleted the sentence P12 L15-18, replacing it by: *“Also, the relationship between modelled transpiration rate and the proportional difference between the observed  $\Delta^{18}\text{O}_{\text{leaf}}$  and  $\Delta^{18}\text{O}_e$  predicted by the Craig-Gordon model (Fig. S11) was non-significant, revealing no evidence of a Péclet effect. This was also true, when investigating that relationship with a subset of the data that included only the leaves that exhibited near-steady-state  $^{18}\text{O}$ -enrichment. This subset was estimated using model output to identify the times when near-steady-state conditions were most likely, and included about half of the data (results not shown).”*

## **R2 10**

The data sampled 7 species while the model describes one mean species. What is the effect of this? Could an averaging of different leaf dynamics not lead to the observed missing correlation with transpiration?

Yes, we also wondered if inability to detect a Péclet effect in the mixed-species leaf sample could have resulted from different leaf water and, hence,  $^{18}\text{O}$ -enrichment dynamics in the different species. As we could not answer that question with the data from our grassland ecosystem study, we included ancillary data obtained separately with *Lolium perenne* and *Dactylis glomerata* in different experiments in controlled conditions by Margaret Barbour. These species formed part of the mixed-species sample in our grassland ecosystem. The *L. perenne* data were based on destructive measurements of leaf water  $^{18}\text{O}$ -enrichment; conversely, the experiment with *D. glomerata* employed an online gas exchange and equilibrated leaf water method. In both cases, a Péclet effect was not apparent.

In the revision we expanded and improved the discussion of the putative causes for the absence of a Péclet effect or for our inability of detecting one (P15 L23ff):

*“...environmental conditions. We do not know if putative between-species differences in leaf water dynamics and associated  $^{18}\text{O}$ -enrichment, or any other morpho-physiological effects e.g. associated with leaf aging, could have led to a missing correlation between the proportional difference between measured leaf water  $^{18}\text{O}$ -enrichment and that predicted by the Craig-Gordon model ( $1 - \Delta^{18}\text{O}_{\text{leaf}}/\Delta^{18}\text{O}_e$ ) and transpiration rate. For these reasons, we explored this question with separate studies of *L.**

*perenne* and *D. glomerata*, two species that also formed part of the present grazed grassland ecosystem. Again, these studies found no evidence for a Pécelet effect, and supported the two-pool model, as there was no relationship between the proportional difference between measured leaf water enrichment and that predicted by the Craig-Gordon model ...”

## **R2 11**

Overall I compliment the authors on this very nice data set and the very careful modelling, and wish to see the paper published soon.

Thank you!

## **References not included in the Discussion paper**

Auerswald, K. and Schnyder, H.: Böden als Grünlandstandorte, in: Handbuch der Bodenkunde, edited by: Blume, H.-P., Felix-Henningsen, P., Frede, H.-G., Guggenberger, G., Horn, R., and Stahr, K., Wiley-VCH, 31, Erg.Lfg., 1-15, <https://doi.org/10.1002/9783527678495.hbbk2009003>, 2009.

Bazot, S., Mikola, J., Nguyen, C., and Robin, C.: Defoliation-induced changes in carbon allocation and root soluble carbon concentration in field-grown *Lolium perenne* plants: do they affect carbon availability, microbes and animal trophic groups in soil?, *Funct. Ecol.*, 19, 886-896, <https://doi.org/10.1111/j.1365-2435.2005.01037.x>, 2005.

Brinkmann, N., Seeger, S., Weiler, M., Buchmann, N., Eugster, W., and Kahmen, A.: Employing stable isotopes to determine the residence times of soil water and the temporal origin of water taken up by *Fagus sylvatica* and *Picea abies* in a temperate forest, *New Phytol.*, 219, 1300-1313, <https://doi.org/10.1111/nph.15255>, 2018.

Lemaire, G., Hodgson, J., de Moraes, A., and Nabinger, C.: *Grassland Ecophysiology and Grazing Ecology*, CABI Publishing, Wallingford, U.K., 2000.

Robin, A. H. K., Matthew, C., and Crush, J. R.: Time course of root initiation and development in perennial ryegrass – a new perspective, *Pr. N. Z. Grassl. Assoc.*, 72, 233-240, 2010.

Rothfuss, Y. and Javaux, M.: Review and syntheses: Isotopic approaches to quantify root water uptake: a review and comparison of methods, *Biogeosciences*, 14, 2199-2224, <https://doi.org/10.5194/bg-14-2199-2017>, 2017.

Schenk, H. J. and Jackson, R.B.: Rooting depths, lateral root spreads and below-ground/above-ground allometries of plants in water-limited ecosystems, *J. Ecol.*, 90, 480-494, <https://doi.org/10.1046/j.1365-2745.2002.00682.x>, 2002.

Yang, J. Z., Matthew, C., and Rowland, R. E.: Tiller axis observations for perennial ryegrass (*Lolium perenne*) and tall fescue (*Festuca arundinacea*): number of active phytomers, probability of tiller appearance, and frequency of root appearance per phytomere for three cutting heights, *New Zeal. J. Agr. Res.*, 41, 11-17, <https://doi.org/10.1080/00288233.1998.9513283>, 1998.