## Anonymous Referee 1:

As an important component of the ecohydrological cycle, it is always difficult to analyze and quantitatively characterize Non-rainfall water (NRW). In this paper, the author designed, manufactured and tested a novel micro-lysimeter (ML) system, which has high precision and weighing range, and overcomes some defects of the existing lysimeters. Different types of NRW inputs, such as dew, hoar frost, fog, rime and the combinations among these, are well distinguished by auxiliary sensors. At the same time, it is also similar to the surrounding environment in terms of canopy and soil temperature, plant growth and soil humidity. The author applied the ML system at a field site in Switzerland. Through the monitoring of a hydrological year, different NRW events were effectively distinguished and the NRW inputs was quantified. In general, the paper has good innovation and practical value, but some parts need to be improved, and major revision are suggested. The specific opinions are as follows.

- 1. It is recommended to attach the location map of the study site in Switzerland and the real photos of the ML system;
- 2. How deep is the groundwater level? Should the impacts of groundwater be considered?
- 3. Line 215-220, it may be difficult to transfer the soil body from one ML to the second. How to ensure the stability of soil during transfer?
- 4. Do you think freezing-thawing would have impacts? Should the accuracy of the ML change at different temperatures?
- 5. Can the ML distinguish the influences from dust or other drifting materials falling on it?
- 6. According to the author's statement, the ML can only distinguish windless conditions, so it cannot work on windy days? NRW input over a year in Figure 8 might be underestimated?
- 7. The ML can quickly discharge the excess water after precipitation, but this would also cut off the evapotranspiration channel, which might misrepresent the conditions in the nearby soils. Will it affect NRW?
- 8. In figure 7 (b), it is described as WFPS in the text, but not in this figure.
- 9. Different types of NRW inputs in Figure 8 can be represented in different colors. The same is for Figure 9.

### Response to Referee's comments:

We thank the anonymous reviewer for his/her constructive feedback and inspiring questions. We did our best to improve our manuscript accordingly. Our response to reviewer's suggestions and critique is given in blue indented text point by point. Line numbers are referring to the clean revised manuscript version.

- 1. We added a location map and an aerial photograph of the site to Appendix A Fig. A1. Furthermore, we added photographs that show the installation process, the uninstalled and the installed ML system to Appendix B Fig. B1.
- 2. Please also have a look at our response in the final author comments. The disconnection of the ML pot to the underlying soil and in consequence to the groundwater body is a general lysimeter problem. We added to (lines 79–82) that there is also a disconnection to the groundwater body. "While limited rainfall retention capacity of ML is not a problem for NRW quantification, the potential prevention of upward direct water flow due to capillary rise from deeper soil layers or

the groundwater body cannot be neglected (Evett et al., 1995), because it replenishes plant available water in the rooting zone."

- 3. Indeed! This is quite a challenge. From our experience the monolith breaks most likely at the bottom, whereas breaking at the edges is not a big issue. At the time of installation, we were three people at the site and could manage to retrieve three undisturbed soil monoliths after one failure. We found that this work is best done when the soil is not completely dry or extremely wet to avoid any breaking of the monolith. We added: "The reversed soil monolith was carefully taken out from the ML pot and three people collaborated to transfer it to a second ML pot to be upright again. The ML pot was then ready for installation on the weighing platform." (lines 729–731). Furthermore, we added photographs of the installation process to Appendix B, Fig. B1 a-d.
- 4. Freezing-thawing could damage PVC parts of the ML, however this did not occur during the study period. We added to the text: "During or after freezing temperature conditions the ML system should be controlled, because expanding water in the reservoir or the ML pot could break PVC parts of the ML system. However, this did not occur during this study period." (lines 488–491). We added that the used load cells were temperature compensated. "Between the load plate and the base plate, a PW15AHY temperature-compensated load cell with 20 kg capacity (HBM, Darmstadt, Germany) was mounted." (lines 154–156).
- 5. Thank you for pointing this out. We added: "In deserts or arid regions (with low vegetation cover) additional sensors (e.g. infrared video cameras) would be needed to detect depositing materials like dust and sand that accumulate on the ML over time." (lines 556–558).
- 6. We added: "However, NRW inputs occur during conditions with low wind speed, the probability for dew formation decreases below 5% when wind speeds are smaller than 0.4 m s<sup>-1</sup> or bigger than 1.9 m s<sup>-1</sup> (Zhang et al., 2014). Thus, wind is not a big bias source for NRW quantification." (lines 477–479).
- 7. The water source of NRW inputs is atmospheric water vapor. The atmospheric water vapor can also stem from evaporated soil water, termed distillation. Distillation can contribute up to 42 % of the condensed water on foliage during dew nights (Li et al., 2020). Thus, a ML pot that is cut off from the evapotranspiration channel might contribute less to the distillation process. However, in the boundary layer atmospheric water vapor is distributed via turbulence and thus the ML is subject to the same atmospheric water vapor as the surrounding field. This cut-off might still lead to the same amount of condensation (e.g. dew formation on plants). However, prolonged altering of soil moisture could have an influence on plant growth and in consequence to NRW inputs (please have also a look at our response to a similar comment by reviewer 2).
- 8. Thanks for pointing this out. We used once "WFPS" instead of "Soil temperature" which created some confusion. We corrected it.
- 9. We changed the figures, each colour represents now a different type of NRW input.

Li, Y., Aemisegger, F., Riedl, A., Buchmann, N. and Eugster, W.: The role of dew and radiation fog inputs in the local water cycling of a temperate grassland in Central Europe, Hydrol. Earth Syst. Sci. Discuss., 1–27, doi:10.5194/hess-2020-493, 2020.

Zhang, Q., Wang, S., Yang, F. L., Yue, P., Yao, T. and Wang, W. Y.: Characteristics of Dew Formation and Distribution, and Its Contribution to the Surface Water Budget in a Semi-arid Region in China, Boundary-Layer Meteorol., 154(2), 317–331, doi:10.1007/s10546-014-9971-x, 2014.

#### Anonymous Referee 2:

We thank the anonymous reviewer for the constructive and detailed feedback. We did our best to improve our manuscript accordingly. Our changes to reviewer's critique are given in blue indented text after each point addressed by the reviewer. Line numbers are referring to the clean revised manuscript version.

The paper introduces a microlysimeter for the specific purpose of measuring the deposition of water on the surfaces of the soil and the vegetation by means other than rain (dew, fog, etc.). The deposited amounts of water are quite small, yet the paper argues that they could be important for the vegetation during dry periods. The design of the set-up, with its high observation frequency and high-resolution mass measurements is explained in some detail. The instrument has been in operation in the field. Its performance is reported and evaluated.

#### General comments

Overall, the paper gets across the relevance of MLs and the improvements the authors have made to earlier designs. The material fits the HESS mission and its readership.

That being said, the paper is wordy and tedious at times. The authors go in so much detail that is hard to follow the line of thought. In contrast, Figure 1 leaves out many technical details, and photographs of the ML and its components are not given.

We have now modified Fig. 1 to show a photograph of the inside part of the ML as panel (a) and a sketch which shows the dimensions of the components at scale in panel (b). We also added an item q in panel (b) to more clearly show that an additional sensor is required if also drainage water rate should be quantified.

The Introduction is comprehensive but incoherent, and would benefit from a careful revision. It can be shortened a little perhaps.

Thanks for this suggestion. We shortened and completely rearranged the Introduction to increase coherence.

In Materials and Methods, there is no information at all given about the soil of the experimental site. The technical details of the ML setup and its installation and operating procedures need to be described better. The rest of the Methods are very detailed, sometimes about well-established techniques. You can shorten the text there.

We added information about the soil at the site: "The soil at the site is a silt loam mixture (56% silt, 37% sand, 7% clay), with a bulk density of  $1.12 \pm 0.03$  g cm<sup>-3</sup> and an organic C content of  $4.4 \pm 0.2\%$  (Stiehl-Braun et al., 2011)." (lines 115–117). We added photographs of the ML system and of the installation process to Appendix B, to better explain the installation and operating procedures. To shorten the Methods, we moved the paragraphs: "Soil monolith preparation" to Appendix B, "Data collection, storage and delivery" to Appendix C, "Load cell data low-pass filtering" to Appendix D. We relegate at the beginning of the Methods to the Appendix, in order that interested readers can find these information.

The authors make a point about separating the various NRW modes, but given the small total flux, I do not understand why this is so important. On the other hand, the differences in soil temperature between the ML and the surrounding soil is downplayed, even though its various effects may linger into the night, when NRW occurs.

In accordance with the Editor (please have a look at the Editor report) we retained the part about separating various NRW inputs and added more explanation why this is important for the readership with another background: "Differentiation among different types of NRW inputs is important for various research disciplines, e.g. the prediction of fog events poses a major challenge for numerical weather prediction for meteorologists (Westerhuis et al., 2020). Thus, it is important to measure the frequency and water inputs of fog events during the whole year." (lines 535–538). For the difference in soil temperature, please have a look at the response on the comment to Section 5.

The material on which the paper is based is solid, but the presentation is not so good. The paper would benefit from a thorough rewrite that increases its coherence and clarity, and reduces its size a little.

We tried to increase to coherence of the paper by rearranging and deleting some parts.

The paper mentions a supplement but I could not find that, other than the data set.

We apologize for this error, the supplement has been included as an Appendix, but we overlooked the need for a change in wording. We changed "Supplement" to "Appendix".

## Detailed comments

L34: The sentence introducing the new ML looks out of place here. You are still developing the argument for its necessity, only moving from general terms to a specific ecosystem.

We agree and rearranged the Introduction.

L39: If rainfall ('RW') is absent, NRW necessarily IS the only atmospheric source of water, because RW and NRW are mutually exclusive and complementary.

We deleted this sentence while rearranging the Introduction.

L43: 'another temperate site' You did not mention the first one.

We deleted this sentence while rearranging the Introduction.

#### L51: What do you mean by plant water status?

We tried to clarify this term by adding: "Plant water status is a widely used measure in plant physiology for assessing plant water stress. It incorporates the amount of water in plants and its energy status (Jones, 2006). NRW inputs can increase the amount of water in plants (Limm et al., 2009; Munné-Bosch and Alegre, 1999) and increase thereby the plant water status, which lowers plant water stress." (lines 39– 43).

## L56 and 320: You mention dew deposition on soil, but earlier you stated that does rarely occur.

It does rarely occur, but still there is the possibility that it occurs. Thus, we would prefer to keep this sentence.

You need to have a careful look at the Introduction. Although I agree with the arguments it presents, they are presented in a confusing order. Above I mentioned the Introduction of a lysimeter in the middle of a discussion about NRW. Elsewhere too you jump between general statements and location-specific arguments without a logical connection between the two. This compromises the coherence of the text and disrupts the flow of thought. All the elements the Introduction needs are there, but please present them in a more coherent order and use more paragraphs as compartments for different focal points of the Introduction.

We rearranged the Introduction as recommended by the reviewer and deleted and shortened where necessary and meaningful.

Section 2: Some subsections are not directly important to understand the ML setup. Perhaps the very detailed material can be placed in the supplement.

We moved three rather detailed subsections to the Appendix: 2.2.3 Soil monolith preparation (now Appendix B); 2.2.4 Data collection, storage and delivery (now Appendix C); 2.2.5 Load cell data low-pass filtering (now Appendix D).

#### Section 2.1: Please give some information about the soil.

We added: "The soil at the site is a silt loam mixture (56% silt, 37% sand, 7% clay), with a bulk density of  $1.12 \pm 0.03$  g cm<sup>-3</sup> and an organic C content of  $4.4 \pm 0.2\%$  (Stiehl-Braun et al., 2011)." (lines 115–117).

## L124: What is FluxNet? Do we need to know?

This is part of the site information to tell readers in which network this site is embedded. Below is a statistics extracted from the Scopus database on the increasing importance of FluxNet in the field of eddy covariance flux measurements of  $H_2O$  and  $CO_2$  fluxes. It is thus not surprising that the name is new to some readers, but this might change rapidly if the spread of FluxNet data usage continues at its current pace. Thus, we prefer to keep this information and expect that readers not interested in it will simply ignore.



L154: Instead of the width and the thickness it is probably better to give the outer and inner diameter of the ML tube.

We changed it according to your suggestions. We use now outer diameter and inner diameter. Furthermore, we realized that our used abbreviations created some confusion, thus we have written out dimensions, e.g. we use now "diameter" instead of "d".

L173-174: How do you level the load cell and the ML during installation in the field? I cannot see the adjustment screws in the figure or how you can reach them during the installation process.

The adjustment screws (=adjustable support feet) are shown in Fig. 1b as item "h". We added a further component Fig. 1b:i counter nut for adjustable support feet to explain the "H" like structure better. We added labels for both visible support feet. Furthermore, we added a photograph of the weighing platform (Fig. 1a) and the adjustable support feet, which are basically machine screws (bolts) with a counter nut (otherwise the weighing platform would move). We could reach them with a prolonged hexagon socket wrench. We added to Appendix B: "The weighing platform was levelled out by adjusting the three adjustable standing feet with a prolonged hexagon socket wrench. The final position was fixed with the counter nut by using an open end wrench." (lines 731–733).

Fig. 1: What are the structures on either side of the load cell that look like an H on its side? The text mentions machine screws (bolts?) but I cannot find these in the figure. I also think it would be helpful to add a photograph of the instrument to the figure, as well as a scale or a set of dimensions within the drawing.

Please have a look at the response to L173-174.

Section 2.2.3: I estimate a filled ML weighs about 100 kg. How did you handle it when you excavated the monolith and when you installed the monolith in its field setup?

The ML mass was lower than 20 kg for all three ML (ML pot size of 25 cm diameter x 25 cm depth). The maximum capacity of the load cell was also 20 kg. We think that our used abbreviations in the Methods section (e.g. "d" instead of diameter) created some confusion, we hope that the dimensions are now better understandable. At the day of installation, we were three people at the site, and we were able to transfer the monoliths to the ML pots and afterwards the whole weighing platform to the outer part. For heavier weights one would need a crane, we agree, but with this weight it is not an issue. We added: "The reversed soil monolith was carefully taken out from the ML pot and three people collaborated to transfer it to a second ML pot to be upright again. The ML pot was then ready for installation on the weighing platform." (lines 729–731).

The ML pots were closed on one side. In order to transfer the monolith from the sampling pot to its ML pot you need to place both pots with the open sides against each other, so you end up with a cylinder that is closed on both sides. How did you transfer the monolith from one half of that cylinder to the other?

We added some photographs and a description to Appendix B Fig. B1, we think that helps to understand this process better. Basically, we took it out from one ML pot and transferred it by hand, in upright position, to another ML pot.

#### L240: What was the size of the averaging window?

We added that the size of the averaging window was 100 values (this was the maximum that this microcontroller could handle). Thus, with a sampling frequency of 3.3 Hz, the averaging window had a length of 30.3 seconds. We added: "The raw load cell data were then stored in an averaging window (ring memory) with a size of 100 values, where the oldest values were replaced by the newest ones." (lines 753–755).

# Section 2.2.6: it is helpful to mention here somewhere what the resolution of the load cell is converted to mm water layer.

This information is not available. The load cell itself outputs only a voltage, that's why also the company does not provide such a measure. The resolution depends on the parts of the system, e.g. if one uses a 8-Bit or a 16-Bit analog digital converter, the calibration range and many other factors. We provide information about the resolution of our system in Section 3.1, thus in the results section, not in the methods section, because this is a setup-dependent value. Here we specify the SE of the measurements of the three ML as  $\pm 0.31$ ,  $\pm 0.14$  and  $\pm 0.11$  g, respectively. We added the lacking units in Section 3.1 and Fig. 2.

# L279: Why the < sign in the sensitivity of the temperature sensor? Now we still do not know its sensitivity.

We changed it to: "To compare canopy temperatures of ML and the control (surrounding) during a NRW input period, we used a thermal camera (testo 882, Testo AG, Lenzkirch, Germany), with a thermal sensitivity of  $\pm 0.05$  °C." (lines 268–270).

L283-284: ,we considered standard deviation to account for spatial variability.' I do not understand this.

We changed it to: "To compare thermal images of the ML surface with the control, we compared the variance (F-test). Data were bootstrapped to reduce sample size from > 30k to 30 samples using the scikit-learn machine learning package of Python (Pedregosa et al., 2011)." (lines 273–276). Furthermore, we added to the results section: "The variance of canopy temperature between the ML vegetation and the control was not statistically significant different (F-test, p > 0.05, n = 30)." (lines 357–358).

L327: Accuracy of what? In the section that follows you use the term accuracy a lot, but I believe you sometimes mean precision (e.g., https://www.mccdaq.com/TechTips/TechTip-1.aspx).

We changed it to: "Accuracy of the ML system" (line 463). In the text we will carefully check the correct usage of accuracy vs. precision. We moved the definition of the terms accuracy, precision and resolution from the Introduction to the Methods section. We tried to better explain the term "precision" when using it: "Precision (repeatability of the measurements)..." (line 492).

L328: Please give more significant digits for the correlation coefficient.

We provide now four positions after decimal point: "Three replications showed an almost perfect linear correlation ( $R^2=0.9999$ ) between target mass and load cell mass." (lines 292–293).

L365: with your measurement frequency, individual eddies in the near-surface atmosphere can affect individual measurements. How did you use wind speed readings to correct for the effect of wind on the readings? Or do you mean you can only discard wind effects if the wind speed is low? In that case, do your data allow to place an upper limit in the wind speed below which its effects can be ignored?

Yes, we wanted to express that we can exclude wind effects at low wind speeds. We tried try to make this clearer in the text by simply stating in the Results section: "Wind speed remained low ( $< 1 \text{ m s}^{-1}$ ) during the whole potential water vapor adsorption event." (lines 329–330) and by stating in the discussion: "Ancillary wind measurements could be used to exclude periods with high wind speeds, because high wind could act as a force on ML and increase thereby mass. However, NRW inputs occur during conditions with low wind speed, the probability for dew formation decreases below 5% when wind speeds are smaller than 0.4 m s<sup>-1</sup> or bigger than 1.9 m s<sup>-1</sup> (Zhang et al., 2014). Thus, wind is not a big bias source for NRW quantification." (lines 475–479).

Table 1: In the two right-most columns, are the signs of the table entries reversed if the visibility exceeds 1000 m and the temperature is above freezing, respectively? The minus sign could be interpreted as leading to water loss from the ML, but it only signals an absence of the corresponding mode of water deposition. Perhaps explain this in the table heading.

We added this information according to your suggestions: "The sign '+' indicates the presence, whereas the sign '-' indicates the absence of a certain factor." (lines 350–351).

L393: Do you know what effect the closed lysimeter bottom has on the temperature profile inside the ML, compared to in situ values? Also, the ML was 4 degrees warmer at the end of the afternoon (Fig. 5), which you discuss in detail later on. Were you able to determine the cause of the temperature difference? Correction: I see that you discuss this later on.

### Thanks for having taken note of our discussion on this important aspect.

L426-436: The MLs had a lower soil water content than the surrounding soil. You state that this did not affect the NRW. However, it does affect the level of water stress experienced by the plants. In combination this leads to the conclusion that MLs can be used to measure NRW as long as the difference in water stress inside and outside the ML does not lead to changes in soil temperature, canopy architecture, plant height, etc., but cannot be used to study the effect of NRW on the water stress of the vegetation. For that you need deeper lysimeters. Is this correct?

Yes, this is correct, deeper and wider (greater diameter) lysimeters are necessary to minimize such artefacts, but also normal size lysimeters (taking on the order of one ton of soil) face the oasis problem that conditions inside a lysimeter are never perfectly equal to conditions in undisturbed soils. With respect to our ML system that aims at resolving small NRW inputs, we added: "Thus, the ML system can be used to reliably measure NRW inputs as long as the difference in soil moisture during prolonged drought periods does not influence plant height or canopy architecture." (lines 573–575). However, we think that a ML system can be used to study the effect of NRW on water stress. The ML system can be used to detect NRW inputs. Plant water stress measurements can be done next to the MLs. During destructive plant water stress measurements, e.g. water potential measurements with a pressure bomb, it is anyways considered to measure in the field in order to avoid manipulation of ML vegetation.

# Section 3.6: You present many numbers in the text, which is rather tedious. This information can better be organized in tables.

We collected these values in a new table "Table 2" to reduce numbers reported in the text (lines 425 onwards).

#### Section 4.1: I believe you mean resolution instead of accuracy.

In our understanding accuracy is "how close a reported measurement is to the true value being measured." (https://www.opto22.com/support/resourcestools/demos/accuracy-vs-resolution). The reported measurements are in our case the readings from the microcontroller and the true values stem from the calibration mass as they are used to calibrate commercial scales in grocery stores and elsewhere. We define the term "accuracy" in the Methods section: "In this study, weighing accuracy denotes the difference between the measured mass (determined with a ML) and the control (calibrated mass)." (lines 231–232) "Resolution is the smallest change that can be measured." (https://www.opto22.com/support/resources-tools/demos/accuracy-vs-resolution). So, resolution must be smaller than the accuracy, we report a resolution of 0.0002 mm. L505: stable decimal place: the meaning of this depends on the units you choose, which you specify elsewhere. I think it is better to rephrase and state the resolution you achieved, compared to that of earlier instruments.

We deleted this sentence to avoid confusion. The resolution is described later on in the same paragraph.

L516: According to the dimension (L105), the ML pots have a volume of 67 liters. They cannot possibly weigh not even 20 kg at that size.

The dimension of the ML pots is 25 cm in diameter and 25 cm in depth (see lines 160–161). Converted to meters and calculating the volume yields  $(0.25/2)^2 \times \pi \times 0.25 = 0.0122719 \text{ m}^3$ , which corresponds to the 12.2 liters that we report. All three ML had a mass under 20 kg on the order of 15 kg. Maybe you calculated the 67 liters with the dimensions of the outer part (45 cm in diameter x 47.2 cm in depth). The outer part is to protect the inner part and is not the ML pot. We adapted the description of the dimensions in the methods Section 2.2.1 and hope thereby to have eliminated confusion about dimensions.

#### L521: I have the impression this confusion in terminology also appears in this paper.

We defined these terms in the Methods section and tried to stick to this definition throughout the manuscript. "In this study, weighing accuracy denotes the difference between the measured mass (determined with a ML) and the control (calibrated mass). Precision reflects the reliability of the measurements, and it specifies to what extent the experiment can be repeated. On the other hand, resolution is the smallest distinguishable unit for an observable change in mass and thus determines the upper limit of precision." (lines 231–234).

### L550: The grass, not the grasslands, grow.

We changed it according to your suggestion: "The highest NRW inputs occurred during the months of main grass growth (April–September), indicating a potential hydro ecological relevance." (lines 533–534).

L550: The claim that NRW is highly relevant is a bit too fast. To validate that claim you have to show that it can substantially reduce water stress and/or significantly increases actual transpiration.

We rephrased to: "...indicating a potential hydro ecological relevance" (line 534). In fact, local farmers concluded that the surprising October grass harvest (after a drought with very little rain only terminating the drought), so we probably somewhat overemphasised this aspect based on local understanding of farmers, and thus a more neutral wording is indeed a good idea.

L558-559: 'However, the NRW inputs of the potential water vapor adsorption events were with < 1 mm' I do not understand this, please rephrase.

We rephrased to: "However, the NRW inputs of the potential water vapor adsorption events were rather low (0.03 - 0.13 mm)." (lines 544–545).

L588: But you did not measure the plant mass (impossible to do non-destructively) or the leaf area index, so you may, in fact, have had reduced plant growth that you did not see.

We changed it to: "In this study, this had however no influence on plant standing height because measurements of plant height (before the drought period) and measurement of overall vegetation height (after the drought period) were not statistically different." (lines 569–571).

L625: ...on plants... Just above you limit yourself to grass. I believe you demonstrated that your system works for short vegetation. For high plants (Maize, shrubs) I am less sure. Also, for vegetation with interlocking leaves, or plants that can be flattened by wind (e.g., barley) and then get back up again, your very sensitive mass measurements may be compromised. Later you claim your system works for plants up to 120 cm. What is the rationale for this value?

We rephrased: "Thus, we conclude that our novel ML design is suitable for quantifying nocturnal NRW inputs on grasses and forbs reliably and accurately at high temporal resolution." (lines 608–610). The claim for up to 120 cm comes from another site, where the grass was that high. However, this other site was not part of this study, hence we will rephrase: "This ML size allowed natural plant growth and such a ML system can therefore be used in different ecosystems with most short to mid-size statured grasses and forbs or similar vegetation up to ca. 40 cm." (lines 681–683).

L689: You reported a diameter of 45 cm above, yet here you state that its area is comparable to 25 by 25 cm, which is 0.4 times your ML-area.

The dimension of the ML pot is 25 cm in diameter by 25 cm in depth. Our wording "25 x 25 cm" can easily be misunderstood, thus we changed it to: "25 cm diameter  $\times$  25 cm depth" (line 680).

L694: 'simulate' is not really the right word here. 'Represents', 'reproduce', or 'mimick' are all better, depending on what you want to convey.

We rephrased and used "represents" (line 684).

Section 5: You leave out the discrepancy in soil temperatures inside and outside the ML, but it worries me. The temperature difference affects the heat balance of the soil. Liquid water is less viscous in warmer soil, so the hydraulic conductivity increases for a given water content, which will have an effect on the vertical distribution of water and water uptake by roots. A change in the soil temperature affects the partitioning of the incoming energy between heating up the soil and generating evaporation. It also changes the microclimate near the soil surface. Even if the temperature difference vanishes at night, its effects on the soil hydrology may linger. I cannot offer a remedy, and I do not believe it invalidates your measurements, but it is an issue that deserves attention and hopefully can be improved if you continue your work.

To better account for this issue, we added: "Soil temperatures were higher in ML pots, especially during the day. This could influence the hydraulic characteristics of soil water, the heat balance of the soil and in consequence lead to biased latent and

sensible heat fluxes. Thus, further ML studies should primarily focus to get rid of soil temperature differences between ML pots and the surrounding soil." (lines 692–696).

Appendix A: It will be challenging to measure drainage with the accuracy necessary for reliable NRW quantification. A few droplets in the outlet tube may have a sizeable effect on the estimated NRW. This appendix is too detailed and wordy. Please condense it to get the message across better.

We completely agree on this aspect of reduced accuracy (see our response to the commenter, Dr. Groh). We provided this information because some anonymous reviewer of an earlier manuscript version insisted on this. The sensor could be placed in an outlet tube with a very steep angle or a PTFE outlet tube (very low friction coefficient). This must be further developed, as we detail in our response to Dr. Groh. In the Appendix we provide information on water drainage rates of a ML pot of given size and provide potential solutions. The potential solutions must be tested and adapted if necessary. We shortened (now Appendix E) by moving numbers to a table and by deleting some sentences.

Jones, H. G.: Monitoring plant and soil water status: established and novel methods revisited and their relevance to studies of drought tolerance, J. Exp. Bot., 58(2), 119–130, doi:10.1093/jxb/erl118, 2006.

Limm, E., Simonin, K., Bothman, A. and Dawson, T.: Foliar water uptake: a common water acquisition strategy for plants of the redwood forest, Oecologia, 161(3), 449–459, doi:https://doi.org/10.1007/s00442-009-1400-3, 2009.

Munné-Bosch, S. and Alegre, L.: Role of Dew on the Recovery of Water-Stressed Melissa officinalis L. Plants, J. Plant Physiol., 154(5–6), 759–766, doi:10.1016/S0176-1617(99)80255-7, 1999.

Nolz, R., Kammerer, G. and Cepuder, P.: Interpretation of lysimeter weighing data affected by wind, J. Plant Nutr. Soil Sci., 176(2), 200–208, doi:10.1002/jpln.201200342, 2013. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B. and Grisel, O.: Scikitlearn: Machine learning in Python, J. Mach. Learn. Res., 12(Oct), J. Mach. Learn. Res., 2011.

Stiehl-Braun, P. A., Hartmann, A. A., Kandeler, E., Buchmann, N. and Niklaus, P. A.: Interactive effects of drought and N fertilization on the spatial distribution of methane assimilation in grassland soils, Glob. Chang. Biol., 17(8), 2629–2639, doi:10.1111/j.1365-2486.2011.02410.x, 2011.

Westerhuis, S., Fuhrer, O., Cermak, J. and Eugster, W.: Identifying the key challenges for fog and low stratus forecasting in complex terrain, Q. J. R. Meteorol. Soc., 146(732), 3347–3367, doi:10.1002/qj.3849, 2020.

Zhang, Q., Wang, S., Yang, F. L., Yue, P., Yao, T. and Wang, W. Y.: Characteristics of Dew Formation and Distribution, and Its Contribution to the Surface Water Budget in a Semi-arid Region in China, Boundary-Layer Meteorol., 154(2), 317–331, doi:10.1007/s10546-014-9971-x, 2014.

#### Community comment 1:

We thank the commenter Dr. Groh for the comments. Our response to commenter's critique is given in blue indented text after each point addressed by the commenter. Line numbers are referring to the clean revised manuscript version.

The authors present a new ML system and show how this ML system can be used to determine NRW for a grassland site. The manuscript describes partially very detailed the technical set-up of the ML system. The authors describe that the ML system doesn't include observations of the ML outflow/drainage and argue in line 207 the ML design was used here to quantify NRW inputs during dry spells and drought periods in summer. I agree that under this conditions the used assumption of no outflow from the ML system might be partially correct. It is also nice to see that the authors recommend in the same section that an ML system would need, when using it in a more general way e.g. to describe NRW inputs for longer time, and additional sensor to determine drainage outflow.

Thanks for the critical comment, but we respectfully have to disagree in aspects and hence decided to add a full-fledged Appendix F to simulate drainage water flow from the bottom of the soil monolith. We did **not** make the assumption that there is no drainage after rainfall events, but we emphasized that we are not quantifying NRW inputs during times when there might be drainage flow. Thus, we set NRW inputs = 0 mm during rainfall periods and the time thereafter until dry and drought conditions are observed on which our focus is. Thus, **our assumption is that NRW inputs are negligible during and shortly after rain events** given the orders of magnitude higher precipitation rates than NRW input rates during dry and drought conditions. We are sorry that our wording was not understandable to the commenter and hence have rephrased more explicitly.

In Appendix F with Fig. F1 and Table F1 we set up a model following Zhan et al. (2016) and simulated the duration t<sub>d</sub> until less than one drop of water per square meter and day (!) is draining at the bottom of a monolith. As the starting condition we chose fully saturated soils, which in reality requires more than 5 days of substantial aboveaverage precipitation (!). Fig. F1 then shows the range for different soil types. At our field site we have a silt loam, which yields a best estimate of  $t_d = 7.2$  hours; but when we tried to explore the full range of t<sub>d</sub> with parameter combinations within the ranges given by Rawls et al. (1991) that we present in Table F1, the maximum t<sub>d</sub> can be up to 33.8 hours with a two-sided 90% (80%, 50%) CI of 0.0–24.8 hours (0.0–20.3 hours, 0.0–12.2 hours). Because soils typically do not fully saturate during normal precipitation events, we also show simulated  $t_d$  where we measure the time when the average soil water content of the monolith passes 90%, 80%, and 70% of its saturation value. The corresponding  $t_d$  (see Fig. F1) are: 5.0 hours, 4.2 hours, and 3.5 hours, respectively. Thus, we are convinced that eliminating the 24 hours after rainfall in our analysis is a realistic approach to (a) really look at periods of dry or drought conditions, and (b) to support our approach that during dry and drought conditions we do not require a measurement of drainage flow at the bottom of the monolith. This modeling revealed to be quite a bit more challenging than expected, and thus we obtained the help of a Ph.D. student in mathematics, which helped to solve issues and supported WE with the model set-up and simulation. Hence, we added him to the coauthors, as it appears to be an essential question by the commenter and also by the editor, that went well beyond our original manuscript without a mathematician on board.

We expect our  $t_d$  estimates to be high estimates that help with the planning of the design of a ML that also quantifies drainage losses for more general applications than we aimed at (dry and drought conditions).

Going out of the topic of the paper under discussion we have to remind the commenter that we were unable to suggest a ML system that would also resolve the NRW inputs that could occur during rainfall events with an accuracy that is comparable to that what the proposed ML system can provide under dry conditions. There are four reason why this is a challenge of itself and would have to be solved by someone interested in these inputs at times when there are non-dry conditions in a separate study:

During rainfall events there is the issue of the splashing of droplets off the ground; a clear definition would be required of how to separate splash droplet water gains from other water gains; should this then be counted as NRW input (most likely not!), and how could it be distinguished from ground fog water inputs, condensation and adsorption?

Fog droplet inputs under rainfall conditions are also a difficult aspect that would require a special definition. Often during rainfall a near-surface layer of foggy air establishes. Via stable isotopes it can clearly be shown that the fog water comes from the concurrent rainfall, but once the rainfall stops the fog layer may still be present; should one now count fog droplet input during rainfall and after it has stopped both as NRW inputs? Or only the amount gained after rainfall has stopped? Or should both be considered a secondary pathway of normal rainfall input? In any case a more detailed assessment of the terminology would be required – which is not necessary if one focuses on drought conditions (knowing that by definition we do not count the first minutes and hours after a rainfall event a "drought").

During rainfall events the wetting of the outer side of the ML weighing pot becomes important. While one could quantify the water amount adsorbed to the outer side of the weighing pot in the laboratory, we would expect an increasingly higher weight of this water pool under real-world outdoor conditions because of e.g. algae growing on the outer face of the ML pot, and consequently also an increasing share of accumulated hydrophibic soil particles, etc.

No matter how one would choose the definitions for separating the components, the experimental errors to quantify "NRW" inputs during and (shortly) after rainfall events would explode. Currently we could only think of a system that would be 5–10 times less accurate than what we present for drought conditions.

This is not to say that it is impossible to achieve high-quality NRW input data during rainfall events, but it is to explain why we did not include this aspect in our project. By focusing on drought periods we circumvented all these technical challenges, well knowing that during rainfall events NRW inputs may not be perfectly zero (as our

approach assumes, see Eq. 1), hence our suggestion to explicitly mention that our estimates are conservative estimates if rainfall periods are included in the total. We added and changed in the discussion section to: "Under conditions with water lost via drainage, NRW inputs would be underestimated. Especially during and shortly after intensive rainfall periods, when drainage water flow is more likely (see Appendix F, Fig. F1 and Table F1), the application of the ML system is limited. During transition periods, shortly after rainfall, e.g. during nights when the sky clears after rainfall, NRW inputs may be underestimated. Therefore, we excluded such periods (see Eq. 1) from the analysis and limited our analysis for dry periods. Our longer-term NRW estimates might thus be conservative estimates if rainfall periods are included in the total hydrological input. At our site, drainage water flow from the ML pots reached low levels rather quickly after rainfall events (see the Appendix E and F for more details). Nevertheless, depending on soil characteristics and conditions, drainage water flow could persist for longer time (Fig. F1 and Table F1)." (lines 641–650).

We added some examples to explain the before used term "more universal" (the term was deleted in the new version): "However, to further develop this ML system and use it during and shortly after rainfall periods or to improve the measurements during other periods when the soil cannot hold excessive water universally, it is recommended to quantify drainage water flow." (lines 763–765).

Furthermore we added to the Abstract and to the Introduction: "Here we present a novel micro-lysimeter (ML) system and its application which allows to quantify very small water inputs from NRW **during rainfree periods** with an unprecedented high accuracy of  $\pm$  0.25 g, which corresponds to  $\pm$  0.005 mm water input." (lines 9–12) and "The goal of this study was to design and test an automated long-term ML system for NRW quantification to grasslands **during rainfree periods** in the field, that overcomes drawbacks of existing small ML systems in terms of hampered plant growth and altered canopy and soil temperatures compared to the control (surrounding area)." (lines 92–95).

Thus I don't understand why the authors showing in section 3.6 NRW inputs over one year here, when the know that the ML system cannot provide such data? I recommend to delete the section 3.6 as the shown total amounts are strongly biased by the inability of the ML system to correctly quantify NRW input during time were also drainage occurs.

For the analysis of the NRW input for one year we excluded periods when rainfall occurred as mentioned above. It is not the case that the ML system "cannot provide such data", but we make a conservative estimate that when there is rainfall and during the time after, when drainage losses from the ML must be expected, we set NRW input to 0 mm, thus we obtain a conservative estimate of an annual total NRW gain. We added the information that this is a conservative estimate: "Under conditions with water lost via drainage, NRW inputs would be underestimated. Especially during and shortly after intensive rainfall periods, when drainage water flow is more likely (see Appendix F, Fig. F1 and Table F1), the application of the ML system is limited. During transition periods, shortly after rainfall, e.g. during nights when the sky clears after rainfall, NRW inputs may be underestimated. Therefore, we excluded such periods (see Eq. 1) from the analysis and limited our analysis for dry periods. Our

longer-term NRW estimates might thus be conservative estimates if rainfall periods are included in the total hydrological input." (lines 641–648). We now explicitly mention in our second main objective: "design a ML system that allows differentiating between different NRW inputs, here defined as dew, hoar frost, fog, rime as well as water vapor adsorption events **during dry and drought conditions**". (lines 99–101). At our site, drainage water flow from the ML pots reached low levels rather quickly after rainfall events (see the Appendix E and Appendix F for more details). "Nevertheless, depending on soil characteristics and conditions, drainage water flow could persist for longer time (Fig. F1 and Table F1)." (lines 649–650).

please show the ML system installed in the field somewhere in the Material and Method section

We added photographs of the ML system and photographs taken during installation of the ML system to Appendix B.

5 please show day and nighttime here and show in a) more the just the very close vicinity of the ML system in the pictures.

We interpret that with "5" Fig. 5 was meant and not line 5 or section 5. We already show more than the close vicinity of the ML system in Fig. 5b. "To compare ML pot temperatures to temperatures of the surrounding, separate images were taken in a distance of ca. 100 cm (images not shown here) with a size of ca. 75x75 cm, to exclude potential influence of the ML on its approximate surrounding." (lines 368–370). We added these separate images for the commenter. Thermal images were taken during 18:27 and 05:15, no daytime images were taken. This ML system was designed to measure NRW inputs during nights







The dry out periods in Fig.7 a) showing large difference from July until September between different pots and the control. The authors only show this for one period in July but it is also visible in august and September! The argument that nighttime difference are small is not correct as B) shows that only for a very small time window around 6 in the morning the difference is close to zero. However dew starts, as shown in a previous figure (Fig. 5) much earlier at around 7 pm where differences are still large (~2°C shown in Fig. 7 c). Thus the conclusion of the authors that that soil temperatures inside ML pots during the most relevant hours of day when dew forms (during the night before sunrise) from line 429-430 is partially not correct.

This might be a partial misunderstanding. On lines 429-430 we wrote: "From this we conclude that soil temperatures inside ML pots during the most relevant hours of day when dew forms (during the night before sunrise) were not strongly influenced by a lower water content and its resulting lower heat capacity." (lines 395–397). We only conclude that the cause of soil temperature differences were not resulting from a lower WFPS (given that WFPS varied quite strongly among pots). This is what you would expect: had the soil dried out too much, then its temperature measured at 15 cm depth would strongly increase, which was not the case (relative comparison), but this does not express that the ML pot soil temperatures are not a bit higher than the reference in the solid soil control plots (as is clearly shown in Fig. 7c). Such temperature differences can only be minimized by maximizing the size of the lysimeter beyond what would be termed a ML (and which is the trade-off for higher accuracy NRW quantification that we aimed at). Hence our information will help future investigators to further improve the system in this and other aspects. We changed and added: "Over the prolonged no-rainfall period, the hourly mean soil temperature deviations of ML pot 1 from the control ranged between -0.14 °C around sunrise and 2.57 °C in the later afternoon (Fig. 7c). Over the period from MayOctober 90 % of nocturnal one-minute soil temperature deviations (sunset–sunrise) were lower than 2.90 °C, 50 % were lower than 0.69 °C" (lines 399–402).

Section NRW inputs over one year showing strongly biased NRW data as the ML system fails to correctly quantify NRW under conditions were also drainage occurs and I was very surprise that this topic was even not picked up in the discussion section 4.4.

The reason for this impression by the commenter is that we thought we were clear enough to clarify that we are aiming at NRW inputs during drought periods. We thus added this aspect to the discussion section 4.4 and clarified that we excluded rainfall periods (as specified in Eq. 1) and that the ML system might give conservative NRW estimates when drainage water outflow occurs (see also first paragraph of the response).

The appendix A: drainage water flow of ML pots is too speculative from my perspective. Drainage occurs not only during rainfall and shortly after rainfall as mentioned in line 792. The outflow from soil depends on their soil characteristics and thus might differ when using ML system at other sites and different soils. The outflow from soils are typically low additional also bias ET during the day!

In Appendix A (now Appendix E) we used measured data from the ML system and the close-by rain bucket during and after a heavy rainfall event. If drainage persists for a longer time period, then the ML system would give conservative estimates of NRW inputs (as we then set NRW input to zero, see Eq. 1). We presented this argument clearer in the revised manuscript (please have a look at the responses above). Furthermore, we added: "Drainage flow was lower than 0.005 mm h-1 one hour before sunset during event 2, only 15 hours after the last rainfall. However, at other sites with different soil characteristics different drainage flow patterns might occur and our ML system might therefore provide conservative NRW inputs and accentuated evapotranspiration rates." (lines 841–844).

Rawls, W, J., Gish, T. J. and Brakensiek, D. L.: Estimating soil water retention from soil physical properties and characteristic, in: Stewart, B. A. (ed.) Advances in Soil Science, Springer, New York; pp. 213–234, doi:10.1007/978-1-4612-3144-8\_5, 1991.

Zhan, T. L. T., Qiu, Q. W., Xu, W. J.: Analytical solution for infiltration and deep percolation of rainwater into a monolithic cover subjected to different patterns of rainfall. Computers and Geotechnics 77: 1–10 DOI: 10.1016/j.compgeo.2016.03.008, 2016.

END OF DOCUMENT