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.

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.

> The commenter is right that our system aims at quantifying the NRW inputs during periods without rain. It is however a misunderstanding, if the commenter understood the wording "more universal" to represent "longer term". The proposed ML system is well suitable for long term measurements, and of course a long-term sum obtained from such a system is a sum in the domain of conditional statistics (i.e. in this case: during dry periods when NRW input can be expected to become a relevant contribution to the hydrological budget). The Covid-19 pandemy has brought about many conditional statistical sums, so we think readers are able to deal with conditional statistics. However, to become even more explicit and clear, we will adopt our text to clarify that our estimate should be considered a conservative NRW input estimate if rainfall periods are included in the averaging, because when doing so (and as specified in Eq. 1), one would simply have to assume zero NRW input during periods that do not qualify for selection as detailed in Section 2.2.8. The critical points outside the scope of our paper given in these lines would then also be the transition periods, shortly after rainfall inputs, e.g. during nights when the sky clears after a rainfall event, NRW could then be underestimated, hence we will clarify in the text that our longer-term estimates of NRW inputs are conservative estimates. We will add more examples to explain what we mean by "more universal" to avoid any confusion. And we will add "during rainfree periods" in the Abstract and Introduction to clarify that the goal is not to obtain better estimates during rainfall periods.
>
> Our goal was to quantify NRW inputs during drought periods (lines 36–39). We now learn from this comment that this was not clearly enough stated, but by explicitly mentioning this in the Abstract and again in the Introduction (at the bottom where our main objectives are presented), this misunderstanding can be solved.
>
> 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 fogwater 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 chose 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.

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. We will add the information that this is a "conservative estimate" and the reason of course is still that we aimed at quantifying the NRW input during drought periods when there is a potential for this component to become relevant in the (dry period) hydrological budget. This is of course a conditional total, as if you sum up rainfall of days where the intensity is exceeding a certain threshold. That's sound statistics. We described this in the Methods section (2.2.8): "If rainfall occurred during an analysed 24-hour period, that period was excluded, except the rain event occurred directly after the NRW input event." (lines 308-309). Thus, the ML system was used the determine NRW input during periods with no rainfall. NRW inputs could be underestimated shortly after rainfall events. "Nevertheless, under conditions when drainage water flow persists for a longer time, the ML system

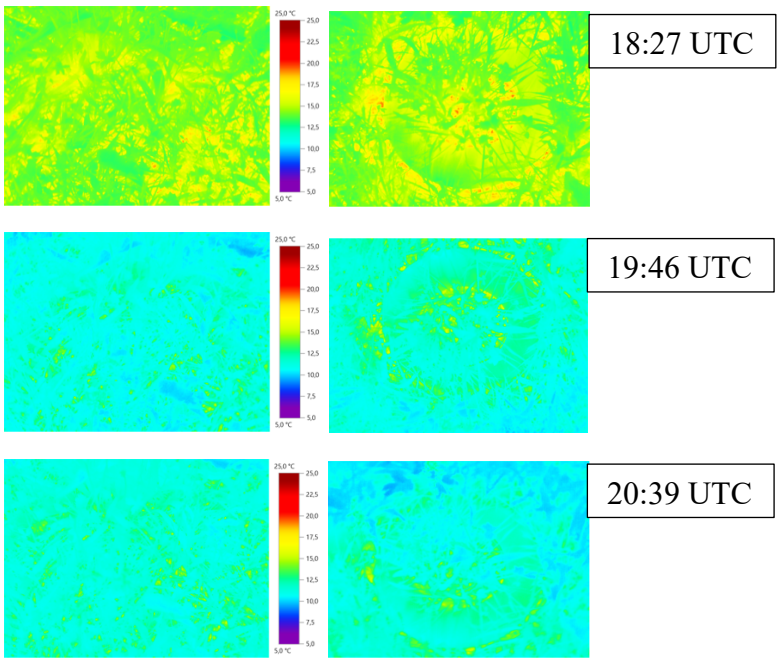
provides conservative estimates of NRW inputs. A possible modification of the ML system to also accurately quantify such drainage flow is suggested in the appendix." (lines 447-449). Please note that our suggestion added to the appendix is to show how this could be done, but in our view we would not get accurate estimates of NRW inputs during rainfall periods as our response above details. Since this aspect is out of scope of our manuscript, we offer this extension as a thought input for anyone who is interested in developing an accurate ML device to quantify NRW inputs during rainfall events. We will make it clearer in the manuscript text that rainfall periods were excluded, and that our ML system might give conservative estimates if drainage water flow persists for longer time. We were of the opinion that the term "non-rainfall water inputs" was clear enough to the reader that this is not including rainfall periods, but we accept this commenter's view that a reader expecting non-rainfall water inputs to be important during rainfall events could be misunderstanding our message, hence we will modify Abstract and Introduction to be more explicit and clear on this aspect that indeed is out of scope of our manuscript.

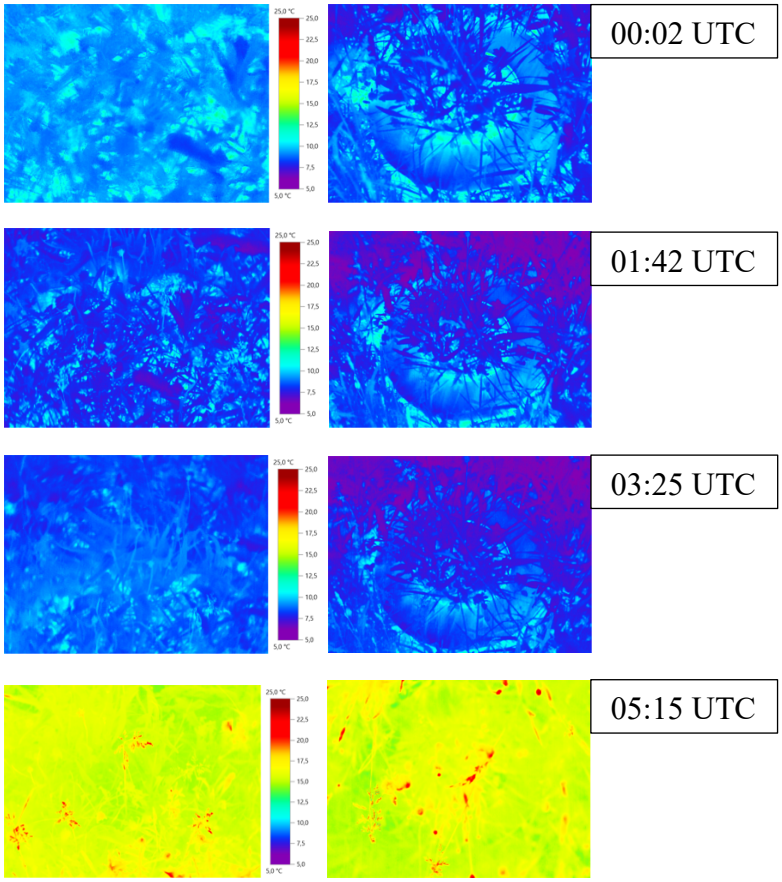
please show the ML system installed in the field somewhere in the Material and Method section

We will add photographs of the ML system and photographs taken during installation of the ML system.

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.". 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**." 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 (line 286) 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 micro-lysimeter (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.

> On lines 433–436 we quantify this effect of temperature differences: "Over the period from May till October 2019 (Fig. 7c), 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. Thus, during most of the night when NRW input occurs, the temperature differences between the soil of ML pots and the control are typically less than 1 °C.". We will change the last sentence to: "…90% of nocturnal temperature values were below 2.90 °C and 50% of the values were below 0.69 °C."

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 will thus add this aspect to the discussion section 4.4 and clarify 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 the appendix A 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 will present this argument clearer in the revised manuscript. Furthermore, we will add that the ML system was tested at the Früebüel study site and that soil characteristics at other sites might differ, and thus different patterns of drainage water outflow might occur.