Dear editor,

We regret that one of the referees was disappointed with our first revision. We kept proposed changes to the text at times limited to maintain the manuscript's clarity. Finding the balance between providing more detail (as often requested by referees) vs. keeping a manuscript text fluid and comprehensible is always somewhat of a challenge. In hindsight, we apparently misjudged this and apologize that reviewer 2 ‘s concerns were insufficiently acknowledged. We think that our elaborate responses to each of the three referees demonstrate that we do take referee advice seriously. We were thus surprised by referee 2’s decision to advise ‘rejection’ of our manuscript and would have hoped to have gone further in dialogue. We are grateful the editor allowed us to do so and now propose to make further revisions to the manuscript and provide extra details on comments raised by reviewer 2 and two by reviewer 3.

Below, we re-list the *major remarks of the referees*, along with updated author responses, as well as explicit the revised sections of the manuscript (<u>in green, with specifically altered parts underlined</u>). Grammar-related remarks or remarks that have already been addressed elaborately during the first revision (including revised text in the manuscript) have been omitted from this list.

RC2:

114 with respect to the difficult-to-interpret results of this study, I wonder whether the lab room temperature may have had an effect. As stated, the two GWT tables were imposed for two consecutive time periods. The C turnover is highly sensitive to temperature. Let's assume a more or less standard Q10 of 1.8 for an Arrhenius-type equation. That means for a 10 degrees increase in soil temperature you would have 1.8 times the amount of C turnover. The lab temperature has a standard deviation of plusminus 0.5 degrees. For a 1 degree increase you end up with 18 % more respiration. Now, if the average temperature in the lab for the two consecutive incubation periods differs…

- ➔ Firstly, a Q10 of 1.8 implies that per 1°C increase in temperature, respiration rises by 8 %, not 18 %. Notwithstanding, we acknowledge that by the batch design of our experiment it is important to understand whether or not differences in temperature could have instead explained part of the observed contrast in respiration between the GWT treatments. Average room temperatures were identical between the consecutive GWT treatments, but we temperature did fluctuate. We propose to now **further specify the average temperature and standard deviation during both GWT treatments separately** in the M&M:

  L120-121: The setup was installed in a temperature controlled (<u>20.8 ± 0.7°C during GWT –165 cm and 20.8 ± 1.5°C during GWT –115 cm</u>) dark room.

- ➔ **We believe** that despite the fluctuations, **the moisture regimes still had a significant impact onto ryegrass mineralization**. Particularly so in the silt loam soils, where the GWT treatment had the greatest effect on topsoil moisture due to capillary rise. Accordingly, the mineralization rate of the more stable $C_{ryegrass}$ pool (k_s) was significantly lower under the –115 cm GWT treatment compared to the –165 cm GWT treatment in these silt loam soils. In contrast, for other soil textures where similar air temperature fluctuations occurred but capillary rise had less influence on topsoil moisture, the k_s values remained consistent across both GWT treatments. Furthermore, after the third wetting event, mineralization rates in silt loam soils differed substantially between the GWT treatments with mineralization rates of 51 vs. 13 µg $C_{ryegrass}$ kg^{-1} soil h^{-1} (on day 57), or a 292.31 % increase in respiration. This is a lot higher than the suggested increase the T fluctuations would cause. In the **results** section, we propose to **include data of the specific mineralization rate values** to provide readers with a **clearer understanding of the magnitude** of differences between the GWT treatments, which may not have been fully apparent in the original figure due to its broad scale:

L352-356: After the watering applications, mineralization rates in the drier, –165 cm GWT treatment, soil seemed to be more sensitive to the moisture input. Significant differences were observed only in comparison to the –115 cm GWT from the second watering application onwards in the loamy sand soil (72 vs 51 µg $C_{ryegrass}$ kg^{-1} soil h^{-1} at day 36), and after the third application for the sandy loam (59 vs 32 µg $C_{ryegrass}$ kg^{-1} soil h^{-1} at day 59) and silt loam (51 vs 13 µg $C_{ryegrass}$ kg^{-1} soil h^{-1} at day 57) soil.

➔ However, with imperfect thermostatic conditions we acknowledge that we cannot 100% guarantee the Birch effect was the sole factor influencing soil respiration. We propose to reword instances where the Birch effect is discussed **more conditionally** in the **abstract**:

L24-30: One possible explanation could be that rewetting may have triggered a stronger mineralization response, commonly known as the Birch effect, in drier topsoils compared to conditions where the soil remained consistently wetter with a shallower GWT level. Based on our findings, including the process of texture-specific capillary supply from the GWT is required to adequately simulate moisture in the topsoil during droughts as they occurred over the past summers in North-West Europe, depending on GWT and texture combination. However, the net effect on topsoil carbon mineralization is complex and warrants further investigation, including the integration of processes related to fluctuating soil moisture following rewetting.

, in the **discussion**:

L493-500: Accordingly, a potential explanation could be that the drier condition of topsoil with deeper GWT may have amplified the Birch effect, i.e. rewetting C mineralization pulses caused by the watering events. In the light of this, it is possible that the expected enhancement of C mineralization (as represented by k_s) due to increased moisture from capillary action under shallower GWT depths was offset in loamy sand and sandy loam soils, and perhaps even overruled in silt loam soils, by a stronger Birch effect in drier soils with less capillary moisture supply. In other words, the reduced mineralization rates observed in silt loam under a shallower GWT (k_s) might be explained, at least in part, by a larger capillary moisture supply compared to the other soil texture and GWT combinations, although other factors, such as an intertwined effects between moisture and temperature, could also be at play.

and in the **conclusion**:

L543-548: Moreover, we found that after rewetting, C mineralization pulses were larger for the deepest GWT treatment. We hypothesize that, as soil was drier for the deeper GWT owing to less capillary moisture supply, the imposed rewetting events caused stronger Birch effects. Hence, not just topsoil moisture state itself, but also the extent of its fluctuation over time, probably culminates into the overall observed net effect of GWT depth onto C mineralization. To use empirical data from experiments like this to improve soil models, further studies will be needed to deconvolute the effects of soil moisture regimes on soil C mineralization.

246 Are measured (saturated) hydraulic conductivities available? In combination with the measured soil water retention functions this would allow to estimate the infiltration depth, or even a soil water content profile over depth. To get an idea of how much the upper 20 cm with the labeled C were actually affected by the irrigation for the different soils…

315 well, now you end up with two factors affecting (or not affecting) the water content in the upper 20 cm with labelled C. The ground water table AND the irrigation. The irrigation will affect the respiration of the different soils differently. This might be difficult to interpret in the end.

500 but not only the GWTs were modified, simultaneously precipitation was simulated

- ➔ We here below address these three comments by referee 2 together, as they are closely related. We firstly want to clarify that the exact same amount of water was irrigated to each of the three soil textures, moreover following the same timing for both GWT batches. Hence irrigation dose and timing was kept identical across the GWT x texture treatments and should not be a factor in our experiment. To **evaluate** the impact of these **wetting events on soil moisture for the three differently textured soils**, we now conducted extra **simulations** using **HYDRUS-1D**. Because saturated hydraulic conductivities were not measured for our topsoil layer we estimated them using the Rosetta pedo-transfer function (Schaap et al., 2001), based on artificial neural network analysis, using texture data, bulk densities and water contents at field capacity and wilting point. We limited these simulations to the top 30 cm of the soil profiles, as the moisture sensors indicated that irrigation impacts would be negligible below this depth (i.e. there was no further fluctuation in the VWC at 30 cm after adding water). Simulated moisture content and hydraulic head along the soil profiles at various time points (0.5, 2, 6, 12, 24, 36 hours) after adding water (at time 0 h) are shown in the figures below. According to these simulations half an hour after water was added (blue line), it reached a depth of 15 cm in loamy sand, 11 cm in sandy loam, and 7 cm in silt loam soils. Over time, the water infiltrates to greater depths. Ryegrass-C mineralization measurements were taken no earlier than 24 hours after wetting, when the added water had ample time to infiltrate up to a depth of 20 cm for all three soil textures. In the loamy sand and sandy loam soils, by 24 hours the water was fairly evenly distributed within the 0-20 cm depth. In silt loam, a more uniform distribution was achieved by 36 hours after wetting. We propose to **add this figure of the simulations to the appendix**, and **add** the following six lines **to** the **M&M** section:

  L150-155: Moisture measurements at –10 cm depth were considered representative for the entire repacked topsoil layer (0 to –20 cm), and therefore suitable for assessing ^{13}C-labeled ryegrass mineralization within this layer. Simulations in HYDRUS-1D confirmed that moisture had infiltrated to about 20 – 30 cm and the water content had become nearly constant with depth across the 0 – 20 cm layer 24 – 36 h after water addition (Fig. A1). Since soil CO_2 efflux measurements (see 2.2.2) were taken then to the earliest after watering events, the –10 cm measurements provided a representative state of the topsoil moisture state.

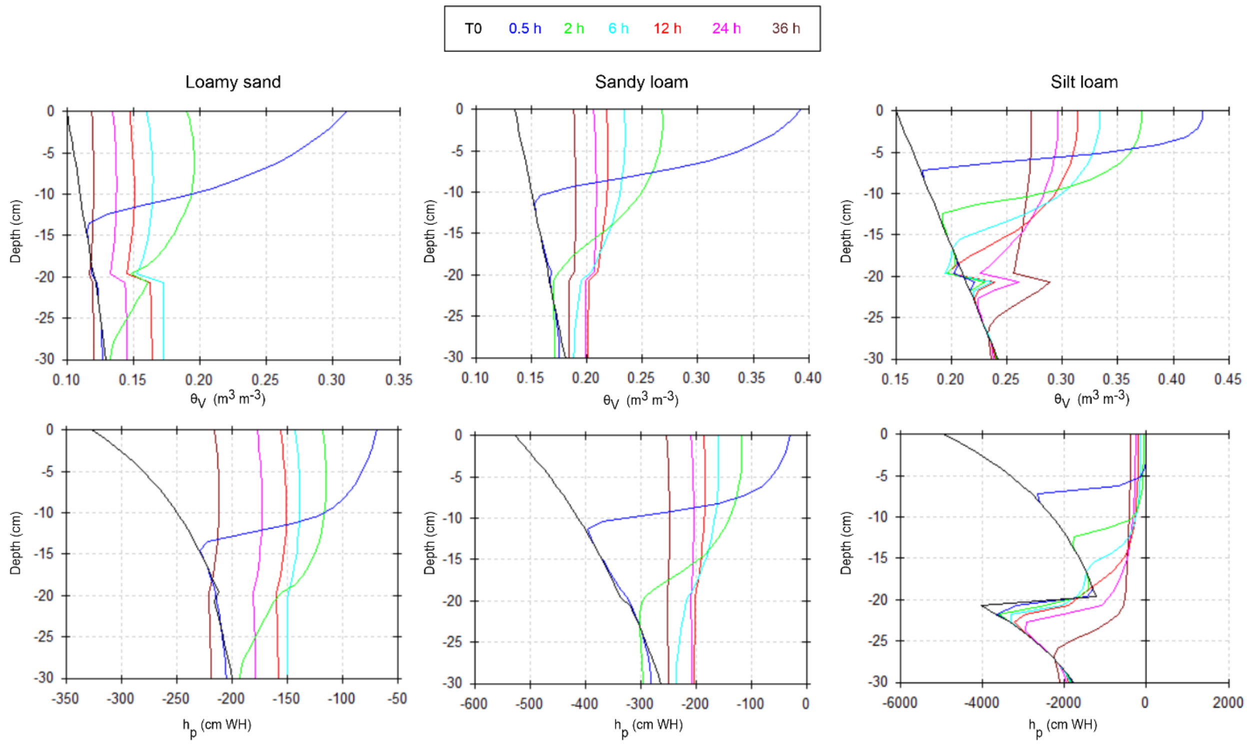

Figure A1: Simulations using HYDRUS-1D to determine when the water, added at T0, reached the lower boundary of the repacked topsoil containing ^{13}C-labeled ryegrass (at –20 cm depth). A 24-hour period after water addition was found sufficient for all three soil textures (loamy sand, sandy loam, and silt loam) to reach an approximately constant moisture level across the topsoil layer.

406 No, that is not unlikely. This could be a result of macropores.

- ➔ We recognize that macropores can serve as preferential flow paths, facilitating a downward water flux through the soil pore network. However, in combination with the volumetric soil moisture data, we think it could not have been a dominant avenue for moisture transport. Higher measured water contents for the –115 cm GWT treatment compared to the –165 cm GWT treatment demonstrate that the net transport of moisture must have predominantly been upwards. **We propose to include this remark the following way in the discussion text**:

  L423-429 These observations imply a downward moisture transport near the GWT, which could be the case if macropores served as preferential flow pathways. However, as we observed an increase in water content during the shallower GWT treatments above this soil segment, upward moisture movement would have been the dominant water flow, with the GWT as starting point of the flow path. We believe these unexpected negative ΔH values are therefore more likely a result of cumulative errors when converting measured θ_V into H using soil water retention curves obtained from drying phases. Due to hysteresis, such curves can be inaccurate for soil wetting (Hillel, 2003), which was the main expected process at the considered deeper depths near the GWT.

409 well, there are very small lab tensiometers on the market.

- ➔ We **propose to add four lines to the discussion text** regarding this matter:

  L429-433: In retrospect, a better experimental approach would have been to directly measure hydraulic head using compact tensiometers, especially in the undisturbed parts of the soil columns, where following moisture transport was the main objective. However, even for the smallest lab-scale models, their installation would likely have caused significant disturbance to the soil columns, in contrast to the volumetric sensors used in this study, which had very sharp fine rods able to pierce the waterproof foil surrounding the soil columns with minimal disruption.

435 these 'bell-shaped' curves were usually determined for disturbed soil samples. And in the wet range the decrease of respiration is rather related to an oxygen deficit and is actually not related to water content. For undisturbed samples the moisture sensitivity function of respiration actually looks different.

- ➔ We believe we have clearly stated that the repacked topsoil layer from which we monitored ryegrass-C mineralization was in fact 'disturbed' , as already explained in L124-125 and L135-136 in the M&M section. We did not further modify the revised manuscript accordingly.

RC3:

L121 I was disappointed that the authors did not report their results for the mineralization of native SOC and only show the $^{13}CO_2$ results. This would have given a fuller picture of C dynamics in their systems. Also, because the added ^{13}C litter is likely to 'exist' as particulate organic matter and a large proportion of SOC 'exists' as mineral associated organic matter, quantifying their relative mineralization under different GWT treatments could have greatly improved your hypotheses and results, especially given the lack of difference in ^{13}C mineralization between soils. Since that data surely exists and does not require repeating or doing new experiments, I highly encourage the authors to include this data. The authors` concerns regarding soil properties effects on SOC mineralization can be partially addressed by normalizing CO2 to C content, and besides the mineralization of the 13C litter is likely also impacted by various properties.

- ➔ CO_2 efflux from native SOC was higher during the first GWT treatment batch (deeper GWT with lower topsoil θ_V) compared to the second one (shallower GWT with higher θ_V) for all three textures. This suggests that a substantial portion of the native SOC derived CO_2 efflux resulted from SOC from the soil below 20 cm, which had not been renewed between both GWT-treatments. If we would have intended to as suggested by the referee compare native SOC mineralization, we should have used completely fresh soil columns for both GWT-treatments. Aside from considerable practical constraints (collecting 0-2 m intact soil cores is not trivial) we prioritized assessing the effect of GWT depth on identical columns to eliminate the potential impact of soil heterogeneity on moisture dynamics. This setup trade-off was initially decided, and is also the very reason why we worked for each batch with new ^{13}C-labeled ryegrass and topsoil material. For clarification we propose to **add the following** 3 lines to the M&M:

  L130-135: We used a model OM substrate (in casu ^{13}C-labeled clipped ryegrass) and did not compare native soil OM mineralization, as its inherently different quality and quantity between the three soils would no longer allow studying the effect of GWT depth, soil texture and their interaction on soil OM. The native soil OM-derived CO_2 efflux in part originated from mineralization of native soil OM in the undisturbed (–20 to –200 cm) soil. As this soil column was not renewed between both GWT treatment batches, the quality of subsoil OM differed between both GWT treatments. For these reasons it is not meaningful to present native SOC mineralization data and results are kept restricted to ryegrass C-mineralization.

L180 delta13C of CO2 undergoes fractionation at different diffusivities (e.g. water contents). https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2008JG000766

Given that your labeled material wasn't very highly enriched, such differences can have an effect on your calculations. Were the parallel incubations of end members (L179-180) done at the same GWTs as the experimental incubations?

➔ We assessed the isotopic signature of CO_2 emitted from native SOC and ryegrass using parallel identical 20 cm packed soil columns, so not in the 2 m soil columns, as described in L180-186. Working with soil on top of undisturbed 180 cm soil columns in GWT barrels, would have been complicated as there would also be 'contamination' of CO_2 being emitted from the underlying soil, limiting to precisely assess the $\delta^{13}C$ of CO_2 derived from ryegrass mineralization. Nevertheless, in these ancillary soil core incubations care was taken to as closely as possible simulate the moisture regime as it occurred in the main experiment: viz. every few days the 20 cm soil columns were weighed, and (a small amount of) water was added if their θ_V became lower than certain thresholds (0.1, 0.12 and 0.15 m^3 m^{-3} for respectively loamy sand, sandy loam and silt loam soils), while also the larger wetting events were applied identically as in the 2 m columns. To now better clarify this, we propose to **add the following** 2 lines to the M&M:

L195-200: The isotopic signature of CO_2 emitted from either end member, i.e. $\delta^{13}C\cdot CO_{2\cdot ryegrass}$ and $\delta^{13}C\cdot CO_{2\cdot SOC}$, were analyzed in ancillary soil incubations as in Mendoza et al. (2022). The $\delta^{13}C\cdot CO_{2\cdot SOC}$ was determined in parallel 20 cm packed soil columns with no ryegrass added. For $\delta^{13}C\cdot CO_{2\cdot ryegrass}$, such soil columns were amended with a high dose of ryegrass (3 g C kg^{-1} soil; indicated as "high"). The rationale for using smaller columns compared to the main experimental setup was to exclude CO_2 emissions from the underlying soil OC, ensuring a more accurate assessment of $\delta^{13}C\cdot CO_{2\cdot ryegrass}$.