Response to Anonymous Reviewer #1

The text in italics refers to the literal reviewers' comments. The text in blue is the literal text edit suggestions for the revised manuscript.

Reviewer 1: *The manuscript by Hoffmeister et al. offers interesting new insights into the water dynamics of an agroforestry system. The study is methodologically very well implemented. The manuscript is complete, well structured and easy to read.*
I have few general comments:

- *The outcome of the study (the most important results) could be better elaborated. This could be implemented in the discussion (content headings for subchapters, summarizing opening or closing sentences) as well as it should be more strongly reflected in the abstract and conclusions.*
- *In my opinion, the abstract and conclusions would benefit from a stronger revision (see specific comments).*
- *The methods and results sections sometimes seem somewhat elaborate. The authors could shorten them by omitting some of the less relevant content, but above all by using shorter wording and sentence structure.*

The specific suggestions for improvement can be found directly in the pdf document.
I wish the authors much success and enjoyment in fine-tuning their manuscript!
Reply: Thank you very much for the positive and encouraging general feedback towards the manuscript. We enjoyed reading and discussing your constructive suggestions and interesting questions. Small text changes such as commas are not addressed specifically in the following but will be considered in the final version of the manuscript (will be visible in the tracked changes version).
We believe that the answers to the above-mentioned general comments are included in the following pages, where we address all specific comments.

L. 15: *The abstract needs more work. The starting motivation sentence is to weak and a conclusional last sentence is missing.*
Additionally, methods are completely missing. What was measured in situ, how was it evaluated? Which modeling tools were used?
Most of the presented results are very descriptive. Either they are not important enough to mention them in the abstract, or they lack interpretation. You could rethink: what are the really important findings from your study (and why)? Formulation tasks/aims with strong verbs (instead of explore, analyse) could additionally help to that (making more clear what the study really is about).
Reply: We agree that the abstract needs to provide more information about the methods and the synoptic analysis and will be revised it accordingly, especially regarding your follow-up comments.

L. 15: *"may" sounds somewhat weak and vague combined with "under forcasted scenarios". Maybe you can be more precise here, or is it at least "likely"?*
Reply: Thank you. We agree that this sounds very vague and we will reformulate the sentence: "The Western Cape in South Africa is a water scarce region which will likely receive less rainfall and higher air temperatures under projected climate change scenarios."

L. 20: *Describing you study, use either tasks or aims listing the elements. Mixing it up is irritating.*
Reply: Thank you! We will consider this and reformulate our objectives to avoid mixing elements:
"Our objectives are twofold; a) we want to explore whether the AFS positively impacts the water balance and b) we evaluate the proposed experimental design."

L. 23: *Here, I find it hard to imagine what exactly you analysed (or how). The step from rwu/event characteristics to infiltration felt quite large to me reading the sentence. Maybe you can link the elements better (prepare the reader better that this will be about infiltration) giving more details about the methods or rephrasing the sentence (changing the order, starting with infiltration).*

Reply: "This study combines measurement campaigns to characterize the spatial variability of various key system properties with continuous monitoring. The campaign measurements encompass extensive soil sampling to determine soil characteristics (nutrient concentrations, hydraulic conductivity, texture, water retention) in the laboratory as well as terrestrial scans of the field site, especially of the windbreaks. The continuous measurements cover meteorological, soil water content and soil water potential observations over a six-month period and were used to understand the state of the soil water and to estimate root water uptake as well as soil water storage dynamics during rain events."

L. 31: *In the last sentence, you repeat what you already said earlier, and you refer to hypothesis we not yet know about. It would be nice to have a conclusion sentence referring to the content of you study (and not if the study worked in general).*

Reply: We agree that the conclusions were formulated somewhat weak and suggest adding the following sentences to the abstract: "We could demonstrate positive effects of the windbreak on the water balance and dynamics in the blackberry field site, even though questions remain as to the extent these benefits and how they compared to less beneficial aspects of the trees (e.g. increased water usage). Our experimental approach was well-suited to examine the relevant processes."

L. 34: *Flexibility and adaptation measures of what?*
The sentence is quite long and complicated as a first sentence. I don't think "challenging" is necessary, the change is the challenge, the word elongates the sentence but does not add information. Maybe the rest of the sentence can be streamlined a bit, too.

Reply: Thank you, we will rephrase and streamline the sentence: "In a changing world, agricultural flexibility and adaptation measures are required to uphold and enhance global living standards, while protecting and restoring ecosystems, as well as to ensure agricultural productivity amid more frequent water shortages, particularly in the global south (Douville et al., 2021)."

L. 38: *This sentence is a bit irritating to me, as the listed elements overlap (enhance resilience and productivity is a benefit) or are unclear (which perspectives? which multiple benefits?). I think it could be rephrased to become more logical.*
(Adding a colon after the sentence, or something around "as will be described in the following", could also help.)

Reply: We will rephrase the sentence to make it more logical and to improve the readability: "AFS describes the combination of woody perennial species with crops and/or livestock complements. It has the potential to deliver multiple benefits and new perspectives for existing agricultural systems including their greater resilience and productivity (Sheppard et al., 2020)."

L. 97: *I'm still missing the methods.*

Reply: We are not entirely sure what is meant here. We describe all the measurements undertaken and go into the details in the method section. If this is referring to the synoptical analyses, we suggest adding a sentence as follows to the manuscript text: "By merging the different methods we could infer process such as infiltration through the combination of nutrient analyses with soil water dynamics during rain events, or by reflecting on the energy budget through shade-cast simulations and evapotranspiration estimates."

L. 156: *the? each?*

Reply: The three samples come from one profile at three different depths. We will modify the sentence to make this clearer.

L. 158: *The order does not make sense here (taking the undisturbed samples, measuring textur, Ksat measurement)*

Reply: Indeed, that is correct. Thank you for noticing this. We listed the measurements in no specific order that they were undertaken in. We agree, it makes sense to list them in the order as carried out and will adapt it in the manuscript text.

L. 160: *I think there is no need to go into detail of how Ksat works. I consider it a standard method, everybody who does not know and is interested can look it up. This applies also to Hyprop and WP4C including the drying and weighing. The methods description should be as concise as possible.*

Reply: Thank you for the reminder of the audience addressed in HESS. We will remove detailed descriptions of the methods mentioned.

L. 180: *Maybe this makes better sense further down?*

Reply: Yes, we agree. Thank you for picking it up. We will move it into the respective data analysis section (2.3.3).

L. 223: *As above, I think this detail can be spared.*

Reply: Thank you. We will shorten the whole paragraph about root water uptake: "The volumetric water content time series were used to retrieve information on root water uptake (Guderle and Hildebrandt, 2015) and on changes in soil water storage during precipitation events. Daily RWU is derived after Jackisch et al. (2020) including a nocturnal correction from stepwise diurnal changes in soil water content between two consecutive days assuming that RWU is the decrease in soil water content between two nights. If the hourly soil water content time series of a sensor did not contain a stepwise decrease, RWU could not be calculated for that sensor on that day."

L. 293: *On average, events had a low intensity of … and a duration of …*

Reply: We will adapt the sentence accordingly: "On average, the events had a low intensity of 1.6 mm h^{-1} and a duration of 11 h 41 min."

L. 311: *I can't make sense of the three given here and the locations in the table.*

Reply: Thank you for pointing this out. We will adapt the caption of the table to clarify the abbreviations: "Abbreviations are in the first two digits according to the sampling row: WB = windbreak or BB = blackberries; followed by the location within the row: P = profile, E = east, M = middle, W = west."

L. 359: *I feel this paragraph can be shortened by rephrasing*

Reply: Thank you. We acknowledge the comment and will rephrase the paragraph to make it more concise.

L. 383: *?*

Reply: Is the "?" referring to the six- hour period between rain events? It is a common procedure to define the end of a rainfall event if it has not rained for six hours straight. In case, you are referring to the word choice (strongest), we will update the text to say: "The greatest precipitation event accumulated…".

L. 438: *You could give the precipitation/time (e.g. monthly average) of you study period in comparison to the long-term yearly precipitation/time, such that one has and understanding what your study is representative for. It is interesting to know how much winter and summer differ on the study site, and how typical your study period summer was.*

Reply: That is a great idea, thank you. For the duration of our measurement period (mid-September to mid-March), we summed the monthly averages of a 30-year time series and will add this rainfall information to the discussion: mid-September to mid-March: 206.5 mm (summer), rest of the year: 580.5 mm (winter), total: 787 mm. In the summer season, approximately half of the precipitation of the winter season occurs: "Total precipitation was 245 mm (30-year average for the same period is 206.5 mm) and partially covered the annual average of 787 mm for the region (Meadows, 2015; Veste et al., 2020; Climate-Data.org, 2024)."

L. 440: *Maybe name the monthly sum, for better comparison?*

Reply: That is a good point, especially after considering the previous comment. We will add the monthly sum to the text.

L. 455: *This transition to soil water retention is a bit surprising.*

Reply: This might be a misunderstanding due to a poorly chosen word on our part. We are not referring to soil water retention in this paragraph but to soil water storage following rain events, i. e. how much is the vol. water content increasing after a rain event and thus indicating how much water is stored in the soil. We will change the wording to avoid this misunderstanding: "Overall, we observed a higher proportion of rainwater stored in the soil at the blackberry location in contrast to the windbreak location, where on average 63 % and 54 % of the rainfall reached the soil column, respectively."

L. 461: *is it "such as" (an example) or "in other words"? Also, it would be helpful to keep the same order when opening the paralell (surface, infiltration, subsurface). But I actually think you just need one half of the sentence, it is redundant.*

Reply: Thank you for pointing this out. We will remove half of the sentence to avoid redundancy.

L. 463: *What about surface/crust effects? Maybe you can give an estimation how representative topsoil values are for the soil surface from your visual experience.*

Reply: Thanks for mentioning this interesting point. It is a valid question as we have regular irrigation and high clay and silt contents in the topsoil under semi-arid conditions. From our visual experience of the observation plot, we did not notice any physical soil crusts and no apparent biocrusts. The reason for this might be the frequent walking on the work paths preventing physical incrustation, while a high cover of accompanying plants leaves little room for biocrusts. However, we do not expect spatial homogeneity in the topsoil Ksat values. There is a substantial difference between the berry rows themselves (rather porous, lightly packed soil, flat) and the space between the rows (compacted, rock fragments, steeper parts), that are frequently used for working on the plants. In the six topsoil samples from March 2022, Ksat is on average 302.3 mm h^{-1} (± 191.3 mm hr^{-1}), thus showing that the sample from 2019 is in a similar order of magnitude. However, the data from the samples demonstrate inhomogeneity between the windbreak and berry sample location, where Ksat is on average 163.2 mm h^{-1} and 441.4 mm h^{-1}, respectively.

Furthermore, the maximum precipitation intensity is quite an outstanding value. The average precipitation rate (> 0) of the whole measurement period is 7.1 mm h^{-1}, confirming what is already observed in the manuscript with Ksat exceeding max. precipitation intensity. We will add this information to the main text of the manuscript.

L. 477: *Can you remind your readers about your texture, if it is easily transported by water, and how much of a textural change you would expect - such, is it still possible that the hypothesis is true, because differences could be minimal, or do you think it becomes unlikely because of the missing textural differences? As transport by surface runoff would take place mainly during extreme events, the kinetic energy is high.*

Reply: The texture of the topsoil was classified as silty clay loam with clay fractions up to 40 %. We agree that surface runoff would mainly occur during extreme events with high kinetic energy.

As mentioned before, we do have events occurring with high precipitation intensities (max. observed 82.6 mm h^{-1}), which definitely have the potential to produce a splash erosion even in cohesive soils which become then erodible. We will address this suggestion for the revised manuscript by including the content of this comment and reply into the revised manuscript: "We observed very high precipitation intensities (max. observed 82.6 mm h^{-1}), which probably produce surface runoff with high kinetic energy and therefore have the potential to produce splash or sheet erosion even in cohesive soils. Possibly, the windbreak may not be apparent in the soil water content changes but downslope erosion of fine soil could explain the unexpected observed lower Ksat values near the windbreak, which is underpinned by larger bulk density and lower porosity at the windbreak. We did not find considerable texture differences between the two locations, but fine particles could be masked through the formation of aggregated particles (Jackisch et al., 2017). Carbon addition may moreover increase and stabilize aggregates in fine-grained soils."

L. 483: *or heterogeneous preferential flow out of the system*

Reply: We consider redistribution to include preferential flow and, naturally, this may not only occur in the lateral but also in the vertical direction. We will add this detail to the text.

L. 488: *thus, not allowing for an absolute, but a relative (between locations) interpretation of the results?*

Reply: Yes, this is correct. The absolute values might be biased by the irrigation. As the irrigation influence should be the same at both locations, relative comparison between the two locations is still possible. We will add half a sentence to clarify this.

L. 510: *I would be careful here, to my knowledge, water can be limiting before, I mean this is the permanent wiliting point - the point of no return (for sunflowers). I have seen matric potentials much before PWP beeing considered as the point where transpiration is reduced in literature.*

Reply: Thank you for referring to the critical context of the permanent wilting point (PWP). Yes, it is true that the permanent wilting point is only based on sunflowers and that water limitation of other plans might be reached much earlier. The PWP mainly describes the soil water status at which water flux are depleted and the water's mobility is strongly limited, but as mentioned water fluxes already slow down earlier than that. We suggest that we reformulate the text so that it is rather in the frame of "before reaching complete water limitation" at the permanent wilting point: "This is confirmed by the matric potential sensors, which show the plant does not reach the PWP (fig. 2), i.e. the point at which water fluxes are nearly immobile. Water becomes a limiting resource for many plants already at lower absolute matric potential values."

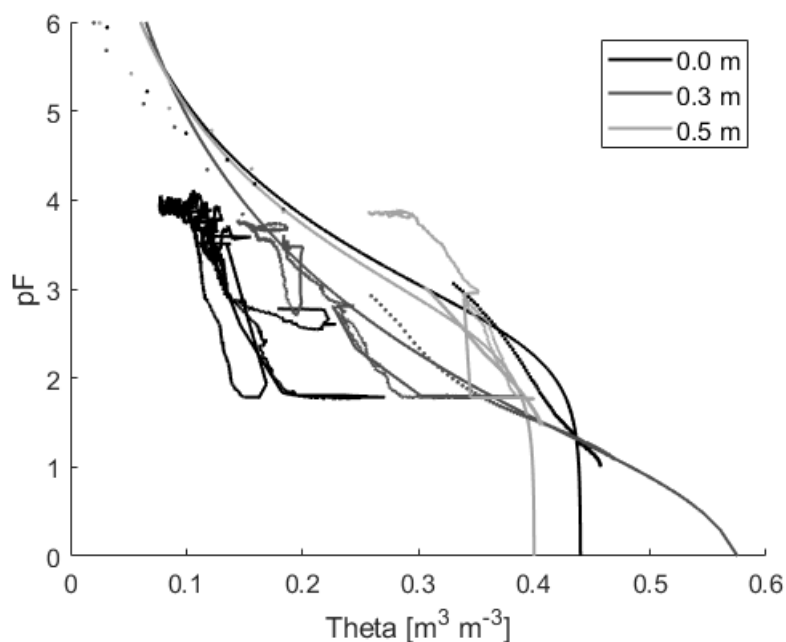
L. 532: *It sounds like a framing to me - not considering overall effects, but only benefits. An alternative would be to give a heading that already contains an outcome (Windbreaks induce benefits...). You could do this througout the discussion (giving the main outcome of the subsection in the heading), I always find it very helpful as a reader.*

Reply: Thank you. We like this suggestion and will consider renaming the subsections of the discussion. This could also help to address your general comment referring to the stronger highlighting of the main results of our study.

L. 536: *These are too many informations within one sentence (ideally, 1-2). Divide into two sentences?*
Reply: Thank you for this useful comment. We will split the sentence into three sentences to facilitate the information flow. "Carbon addition may increase and stabilize aggregates in fine-grained soils. Bogie et al. (2018) found significant differences in water retention at the PWP alongside changes in surface properties brought about by higher CEC of organic matter in coarse soils. This is in contrast to our samples, which had higher topsoil organic matter concentration but with similar PWP."

L. 545: *are the findings only at this location, or did you measure only here? it should be clear from the sentence.*
Reply: Thank you for pointing this out. We will modify the sentence to clarify the measurement location: "Both the volumetric water content (at both locations) and matric potential (measurements at the windbreak only) observations consistently show that the topsoil is drier than the soil at greater depths (fig. 2)."

L. 562: *...but matric potential sensors are more delayed. how big is the gap between the two sensors/the difference between lab and field retention curves?*
Reply: This is also an interesting point. We did not observe any substantial delay in the response of the sensors to the rain events (manuscript, fig. 2). Additionally, our water content sensors cover a rather big measurement volume, which might take some time to react but therefore also more representative. We also question the assumption of a general delay of matric potential sensors, as potential changes travel as waves at higher speeds (celerity) than soil water fluxes, which are controlled by filter velocities.

In the following figure, you can see the big difference between field (dots) and lab (lines) retention curves. Reasons for this are manifold; on the one hand, we can observe hysteresis effects. On the other hand, we need an equilibrium between the two measures, which does not happen so quickly during/after a rain event also the sensors might not be close enough to each other to be able to produce reliable retention curves.

L. 568: *choose which one is important here (September - March vs yearly) and use it for precipitation only as well as for the combination of the two*
Reply: Yes, we agree, one information is enough. We will modify the text accordingly. Thank you!

L. 569: *where is enery lost?*
Reply: Thank you for this question. We realized that our formulation was probably misleading. We are not referring to potential energy but to incoming solar radiation energy. The net radiation is partitioned into sensible and latent heat. We will change the wording in the manuscript.

L. 579: *repetition of "potential"*
Reply: Thank you for pointing this out. We will rephrase the sentence to avoid the repetition.

L. 580: *use unit symbols. and: per metre of what?*
Reply: Thank you. The unit must have slipped our attention. The number is referring to meter of windbreak. We will add this detail to the text.

L. 591: *windbreak themself or windbreaks themselves*
Reply: Thank you. We will correct this mistake in the text.

L. 592: *in what way? we want qualitative findings here, if there is an impact, in which direction is it?*
Reply: (see below, answered together with the next comment)

L. 593: *again*
Reply: Thank you. Yes, we understand and agree that this will improve the conclusions and elevate the main results of this study. We will include some qualitative statements into the revised manuscript.

L. 596: *I think the sentences I crossed out are unnecessary and they seem to not fit to or represent your work. If you want to make a kind of political statement, hide it more in the following sentences or make it shorter and put it at the end - would be my suggestion.*
Reply: Thanks for pointing this out. We will remove the sentence from the conclusions.