We thank the reviewer for their detailed, timely and encouraging feedback. We have updated the manuscript based on the reviewer's suggestions.

The attached document contains a point-by-point response to each reviewer comment, with original comments in black and our response in blue. We have also attached updated versions of the manuscript and the supplement.

Summary and Recommendation

The paper describes a methodology for inferring coarse vs. fine grained fractions of the subsurface based on towed subsurface resistivity mapping (tTEM) calibrated to cone penetration test (CPT) data, in an almond orchard in the Central Valley California. Using various spatial and multi-variable statistics, this mapping is transformed to subsurface coarse-grained-fraction 3D realizations. Variably saturated flow model that simulates managed aquifer recharge by flooding part of the orchard in the spring is used to test hydrological (infiltration and recharge) and agricultural (root-zone saturation periods) farmers' interest. Results showed that coarse-grained structures accommodate the rapid recharge. Fine-grained sediments in the root zone are probable to cause long–term saturation affecting yield. Fine grained-blocks in the deeper unsaturated zone may retain significant volumes of the MAR operation water for years after the flooding. Infiltration, recharge and root-zone saturation were found sensitive only to the fine-grained fraction's hydraulic properties, therefore the authors conclude that better characterization of the fine-grain end member is needed to reduce uncertainty in similar Agricultural-MAR operations.

The innovative technical procedure including non-invasive and invasive geophysical methods, geostatistical, Monte-Carlo and multivariate statistics, and 3D variably saturated flow simulations, is of new and advanced nature and very well described, therefore a good fit for publication in HESS. I have some arguments on the overall understanding of recharge under thick unsaturated zone and the practical conclusion for AgMAR operation which I would like the authors to discuss and perhaps rethink. Therefore, I recommend moderate revisions (between minor and major).

Major Comments

1) As the authors describe nicely recharge is correlated good with surface input (precipitation, irrigation, flood-MAR) due to the pressure wave that propagates fast in the unsaturated zone and push deep-unsaturated older water to the water-table (sometimes called by old groundwater hydrologists as "the train-car model"). Today we can deal with the unsaturated-zone's retention and avoid the simplistic recharge coefficients used in many groundwater models. Nevertheless, in the life-time of an almond-grove (20-50 years) the result of retention of ~ 1/3 of the MAR water in the deep unsaturated zone after 2 years, is no news and hardly important and shouldn't be highlighted as a main result in the abstract. Higher storage in the deep unsaturated-zone will increase recharge in 1 of the following years of high input either due to a rainy year or MAR operation. The authors should discuss and maybe reconsider the important implication of their analysis in the scope of Ag–MAR in orchards.
To address this point, we have:

1) clarified that the timescale of recharge is an important consideration for water managers, if not for agricultural growers: "For example, in California, USA, the Sustainable Groundwater Management Act requires groundwater agencies to implement plans to mitigate groundwater overdraft and avoid undesirable consequences of overpumping by 2034 (Morrell, 2014). In addition, the negative effects of groundwater overdraft (e.g., subsidence) will continue to occur until the recharge

pressure pulse reaches the saturated zone. Accurately quantifying for saturated zone recharge rates will help water managers make an informed decision on whether to perform flood-MAR at a site or to use a different MAR technique with a more immediate impact, such as recharge through injection wells" (lines 589-594 in the updated manuscript without track changes).

2) explained that future work should investigate the long-term importance of this process at sites with repeated inundation over multiple years: "the extent of capillary trapping of recharge water in fine-grained sediments likely decreases with successive recharge events. Future work should strive to quantify the increase in saturated zone recharge efficiency with repeat recharge events over successive years" (lines 597-599).

More-ever, recent 3D simulations in heterogeneous variably saturated medium concerning transient flow from a drywell showed that fine-grained layers at the bottom of a dry well contribute to faster downward flow in the unsaturated zone and faster recharge - see Russo et al., 2022, WRR, https://doi.org/10.1029/2021WR031881. This phenomena is due to the turn over in unsaturated hydraulic conductivity during drying, where at increasing negative pressure head, fine-grain medium becomes more permeable than course-grain.
We have included this study in our discussion of capillary barriers in both the introduction (line 47) and the discussion (lines 520-522).

2) Although used often, the term "deep vadose-zone" is awkward, as vadose comes from Latin meaning shallow. I suggest to use deep unsaturated zone for the domain between the bottom of root zone and the water table (especially if it is of tens of meters thick).
We have replaced the phrase "deep vadose zone" with "thick unsaturated zone" in the abstract (line 3) to limit confusion, given the Latin etymology of the word "vadose".

Specific comments

1) L71 see also Rudnik et al., 2022, WRR for use of stochastic approaches in MAR
We have added this article to the examples referenced on line 71 in the revised manuscript.

2) L76-77 as discussed in major comment 1, in transient heterogeneous unsaturated flow fine-grained layers can increase flow in drying periods. Perhaps change "restrict flow" to impact flow.
We have changed "restrict flow" to "impact flow".

3) L113-126,Why not define the recharge of a 40 m deep aquifer as the downward flux at 39 m depth and stay with fluxes:1) it is straightforward and simpler; 2) the saturated zone part of a model may include sources and sinks (pumping wells) or transient head boundary conditions which have impact on the lateral flow not related to the MAR operation. Discuss.
While the technique the reviewer proposes is simple, it does not account for recharge due to mounding of the water table and would thus underestimate recharge rates. For example, if the water table is initially at 40 m and mounds up to 35 m, that water would not be counted as recharge until it moves below 39 m. This delay constitutes an underestimate in recharge rate. Moreover, in this case, we do not have any sources or sinks within the model domain that may impact lateral flow across each boundary.

4) Figure 3, caption last row, change "hydraulic conductivity" to saturated hydraulic conductivity
We have changed "hydraulic conductivity" to "saturated hydraulic conductivity".

5) L 209-210, perhaps better: Algebraically the resistivity of a tTEM cell is described by the harmonic mean of the fine and coarse grain resistivity's as:

We have updated this sentence to read: “Algebraically, the resistivity of a tTEM cell is described by the harmonic mean of the resistivity of the fine- and coarse-grained layers:” (line 210).

6) L249, May be better hydraulic functions than "water retention curve" (to include the unsaturated hydraulic conductivity function as well as the retention curve).
We have updated this sentence to clarify that the van Genuchten model is used to calculate both saturation and relative permeability.

7) L 360—370, only complete saturation? Or defined from some threshold of high saturation (e.g. 95% saturation)?
We have clarified that we use a 90% saturation threshold (section 3.4.1). Ganot and Dahlke (2021) estimated that a minimum volumetric air content of 0.09 m^3/m^3 is required to limit anoxia during Ag-MAR. However, since porosity varies both within and between simulations, for simplicity we use a constant saturation threshold instead.

8) L414 – 415 "or from the particular …simulation." not clear, explain or discard if not important for the sensitivity analysis description.
We have added an example to clarify this point: “For example, one stochastic realization of the model domain may have a higher percentage of fine-grained facies than another, limiting recharge" (line 414-415).

9) L 438 "0.15 +- 0.29" 0.29 standard deviation? define explicitly
We have defined this explicitly in the text (line 438).

10) L 461 discard "(7 acres)"
We have removed this parenthetical in the updated manuscript.

11) Figure 6 caption. Is it a single flooding of 0.8 m, or another scheme? Should be said in caption.
We have clarified that this simulation received a single, 0.8 m inundation event.

12) Figure 7: 1) What drives the Flux recharge after 6 years when the water table is back to its pre-MAR level, perhaps not a consequence of MAR (relates to specific comment # 3); 2) right hand vertical axis title - typo recharge.
We have edited the caption to explain that “even after the water table returns to its pre-MAR level at ~6 years, water content in the vadose zone remains higher than in simulations without recharge. As this water percolates to the saturated zone, it contributes to flux recharge.” We have also corrected the typo in the vertical axis.

13) L494-495, this is a trivial result no need for numbers and statistics.
The reviewer is correct that this result is intuitive to vadose zone hydrologists. The broader scientific community (along with water managers, consultants and other MAR practitioners) may be unfamiliar with vadose zone water retention processes, however, so we have chosen to keep these statistics in the manuscript.

14) L 540 "especially within the vadose zone", where else than the unsaturated zone?
We have edited this sentence to read: “However, over the long term, fine-grained sediments within the vadose zone absorb a large fraction of recharge water” (line 540).

15) L616-621 – conclusion 2 – As said in major comments, the significance of this result in Ag.-MAR in almond groves is minor. 63% in 2 years for free or cheap water is no good? And the rest 37% are not

lost forever they will recharge in the next rainy/MAR year (unless the pre-MAR unsaturated zone was in really low water contents)
As discussed in the major comments point 1 above, we have extended the discussion (section 5.3) to clarify the importance of the timescale over which saturated zone recharge occurs. See lines 589-594 and 597-599 in the updated manuscript.

16) L 629-634 conclusion 5 – Hard to belief that 1 flooding of 80 cm did not cause more saturation in rootzone than 16 inundations of 5 cm with 1 week between inundations. Check! A weekly 5 cm irrigation should not cause saturation of the entire root-zone unless very poor percolation in the soil.
We have edited this point to clarify that our results may be specific to this site. We have also edited section 5.2 to clarify that outcomes of interest are not sensitive to inundation frequency because "successive inundation events stack on top of one another and reach the water table as one cohesive pulse" (lines 566-567) and because "the large uncertainty on other parameters in the sensitivity analysis (Table 1) may dilute variability due to inundation frequency. For example, inundation frequency may impact root zone residence time, but uncertainty on other parameters (e.g., hydraulic conductivity of the fine-grained sediments) contribute more strongly to the observed variability in each outcome of interest at this site" (lines 571-574).

17) L633 saturated hydraulic conductivity rather than "hydraulic conductivity"
We have changed "hydraulic conductivity" to "saturated hydraulic conductivity".

References:

Ganot, Y. and Dahlke, H. E.: Natural and forced soil aeration during agricultural managed aquifer recharge, Vadose Zone Journal, 20, https://doi.org/10.1002/vzj2.20128, 2021.