Referee#1

The paper presents a solid and carefully executed study on coupling a POD-based reduced order transport model with an iterative ensemble smoother for joint estimation of K and well attributes. Results are convincing, and the work is publishable after some focused improvements.

Answer: We are grateful for the positive and constructive comments. We have carefully revised the manuscript and provided our responses to each of the comments emerged. The line numbers provided correspond to the clean (without tracked changes) version of the revised manuscript.

Improve the novelty statement in the Introduction and Conclusions; what is new relative to existing ROM+DA studies: (i) joint estimation of heterogeneous K and hidden pumping well attributes, (ii) reduction of only the transport equation while keeping flow full-order, and (iii) the systematic multi-factor analysis (ROM size, ensemble size, prior stats, noise, snapshot size).

Answer: It is worth to note that only a few studies are devoted to performing uncertainty quantification relying on a reduced order model (ROM) approach for solute transport. To support this statement, we write: (lines: 171-176) “Although conceptual insights can be drawn from ROMC studies addressing groundwater flow (e.g., Pasetto et al., 2014; Xia et al., 2020, 2025), influence of key factors (such as, e.g., dimensionality of the reduced concentration space and strength of hydraulic conductivity heterogeneity) on accuracy and robustness of ROMC-based UQ still remains poorly characterized”. We then state that: (lines: 177-179) “Building upon these works, the present study introduces a novel framework that integrates the iES with a ROM for solute transport (hereafter referred to as iES_ROM)”.

We further emphasize that: (lines: 60-62) "Despite the relevance of these issues, only limited research has been devoted to the identification and quantification of pumping rates and spatial locations of such hidden wells.".

Additionally, we point out that the scenarios we consider include the key factors that can have an impact on ROM performance in these contexts and are designed to examine the potential of the novel framework we propose.

Provide concrete numbers and protocol: how many realizations and time levels are used, whether snapshots come from prior draws or a single reference field, and whether their statistics match those used in the DA experiments. Add a short discussion of how ROM performance might change if the prior used for snapshot generation differs from that used in assimilation. Add also a short discussion (no need for new runs) on how sensitive the ROM is if the prior used for snapshot generation differs from the prior used for DA.

Answer: We have listed the detailed settings for each test case in Table 1. We now write that: (lines: 353-355) "A uniform time step of 1 day is considered, our analyses encompassing a total simulation time of 10 days (i.e., T_s = 10 days and N_t = 10).".

Snapshots are taken as FSM solutions corresponding to m randomly selected realizations of the initial ensemble of long-conductivity (Y) fields. Our revised manuscript then states that: (lines: 267-273) "The basis functions forming the entries of **P** are computed as the leading eigenvectors (corresponding to the highest eigenvalues) of the covariance of solute concentration evaluated through N_{sn} numerical solutions (i.e., $\mathbf{c}^1$, $\mathbf{c}^2$, …, and $\mathbf{c}^{N_{sn}}$) of the FSM. Here, N_{sn} = $m \times N_t$, where m is the number of MC realizations of hydraulic conductivity that are randomly sampled from the initial ensemble of Y fields, each yielding $N_t = T_s / \Delta t$ (Δt corresponding to a uniform time step) numerical solutions of Equation (2)."

Concerning the discussion for the influence of snapshots on the accuracy of ROM solution, we now write that: (lines: 686-694) "Additionally, we emphasize that relying on realizations of *Y* associated with (spatial) statistics different from their theoretical counterparts linked to the initial ensemble of *Y* fields can contribute to deteriorate the quality of the selected snapshots. Low quality snapshots yield low quality basis functions and low accuracy of ROM outcomes (see our results in Section 4.1; Pasetto et al., 2014; Xia et al., 2020). The latter deteriorate the accuracy of conductivity estimates and pumping well attributes. Additional studies should be devoted to assess the potential of techniques (including, e.g., greedy algorithms) that might contribute to increase the quality of snapshots.".

State whether snapshots are mean-centered, and discuss briefly how omitting a separate mean field affects accuracy. Add a short justification of why a "mean + anomalies" representation is less convenient in your iES implementation, and whether it might reduce ROM error.

Answer: As we state in the original manuscript (lines: 300-306), "The degree of compatibility of ROM to iES is reduced when considering a typical Karhunen-Loève expansion of $\mathbf{c}^i$ (i.e., $\mathbf{c}^i \approx \langle \mathbf{c} \rangle + \sum_{j=1}^{n} \alpha_j^i \mathbf{p}_j = \langle \mathbf{c} \rangle + \mathbf{P}\boldsymbol{\alpha}^i$). This is related to the observation that $\langle \mathbf{c} \rangle$ evolves with time and needs to be evaluated at each time step. This, in turn, implies that *m* numerical solutions of solute concentration through FSM need to be obtained to evaluate $\langle \mathbf{c} \rangle$ at every outer iteration of iES. Hence, computational advantages of employing ROM are reduced while coding complexity increases.".

Make clear which ensemble sizes are realistic for applied hydrogeological problems (e.g. a few hundred to 1000), and present N_MC = 10000 explicitly as a reference benchmark. Emphasize results and cost accuracy trade offs for the practically relevant range.

Answer: Quantification of feasible and acceptable ensemble sizes is case-dependent. When considering a field scenario, one would need to find a trade-off that corresponds to relatively low computational costs while keeping acceptable accuracy of parameter estimates. We now write in our Conclusions that: (lines: 761-764) “Moreover, the values of N_{MC} that one should consider in a field application are case-dependent. In this context, localization techniques can be embedded in DA processes, as these can reduce negative influences of spurious correlation on parameter estimate arising from reliance on small ensemble sizes.”.

Use Tables 1–5 and possibly a small schematic/flowchart to clearly show what each group (A–E) varies and why. In the results, slightly condense repetitive descriptions and highlight cross-group patterns and any non-intuitive behaviors (e.g. non-monotonic trends).

Answer: We recall that we have clearly introduced the settings for each group in Section 3. The latter reads: (lines: 384-420) “To explore the potential of iES_ROM, several showcases are designed to highlight key features of interest. Five groups of test cases (TCs) are designed and organized as detailed in the following (see also Table 1).

- **Group A**. It includes twelve TCs (i.e., TC1-TC12), enabling us to compare performances of iES_FSM and iES_ROM associated with diverse values of n when the pumping rate and locations are either known (TC1-TC6) or unknown (TC7-TC12). The dimension of the ROM is considered equal to {5, 10, 15, 20, 25, 30}, these values being consistent with those most commonly analyzed in previous studies (Pasetto et al., 2014; Xia et al., 2020, 2025).
- **Group B**. It includes four TCs (i.e., TC6 and TC13-TC15), enabling us to compare the performances of iES_FSM and iES_ROM with the largest value of n analyzed (i.e, n = 30) and considering diverse values of N_{MC} corresponding to {30, 100, 500, 10,000}. The latter are values of N_{MC} commonly tested in previous studies (Chen and Zhang, 2006; Xia et al., 2021, 2024).

- **Group C**. It includes five TCs (i.e., TC6 and TC16-TC19), designed to analyze the ability of iES_ROM to cope with diverse quality and quantity of available measurements. Performances of iES_FSM and iES_ROM are also compared when σ_{obs} = {0.001, 0.01, 0.1} and the number of observation locations corresponds to a value selected from {9 (Fig. 1b), 18 (Fig. 1c), 55 (Fig. 1d)}.
- **Group D**. It includes five TCs (i.e., TC6 and TC_20-TC23), enabling us to study the effect of μ and σ_Y^2 of the initial ensemble of Y on the accuracies of estimates of conductivity and pumping rate and well location through iES_FSM and iES_ROM. Values of μ and σ_Y^2 of the initial ensemble of Y fields are selected from {-0.5, 1.2, 2.0} and {0.01, 1.0, 2.0}, respectively.
- **Group E**. It includes six TCs (i.e., TC6 and TC24-TC28), with the aim of investigating the effect of N_{sn} on the accuracies of the estimation of conductivity and well pumping rate and location through iES_ROM and on computation time requirements. Values of N_{sn} in TC24-TC28 and TC6 are equal to 30, 100, 300, 500, 1,000, and 10,000, respectively.

Note that, without specified otherwise, default settings for the above mentioned TCs correspond to TC6 which is designed with n = 30, N_{MC} = 10,000, N_{sn} = 10,000, N_m = 55, σ_{obs} = 0.01, and values of μ and σ_Y^2 of the initial ensemble of Y equal to 1.2 and 1.0, respectively. Except for TC8-TC12, the source/sink term is associated with uncertainty.”.

Prompted by the reviewer’s comment, we also slightly condense repetitive descriptions and highlight cross-group patterns and any non-intuitive behaviors in our revised manuscript.

In the Conclusions, clearly delimit the domain of validity: 2D confined aquifer, steady-state flow, single non-reactive solute, single well. Briefly comment on expected challenges and required modifications for transient flow, multiple wells, or reactive/density-dependent transport.

Answer: We state (at the beginning of our Conclusions) that: “This study addresses joint estimation of (uncertain, spatially heterogeneous) hydraulic conductivities and attributes (location and flow rate) of a pumping well in a two-dimensional confined aquifer in the presence of (non-reactive) solute transport taking place across a steady-state flow field.”.

We now add at the end of our Conclusions: “Additional elements of interest associated with future studies on coupling iES with ROM include the analysis of transient saturated/unsaturated flow, reactive transport, and density-dependent flow/transport scenarios. When considering nonlinear systems, reliance on discrete matrix interpolation schemes (Negri et al., 2015; Bonomi et al., 2017) constitutes a promising approach to enhance computational advantages of ROM.

References

Bonomi, D., Manzoni, A., Quarteroni, A., 2017. A matrix DEIM technique for model reduction of nonlinear parametrized problems in cardiac mechanics. Comput. Methods Appl. Mech. Eng. 324, 300-326.

Negri, F., Manzoni, A., Amsallem, D., 2015. Efficient model reduction of parametrized systems by matrix discrete empirical interpolation. J. Comput. Phys. 303, 431-454.

Overall, these changes are mostly clarifications and presentation refinements; the core methodology and results appear sound.

Answer: We thank the reviewer for the very positive appraisal of our work.