The aim of this study was to propose a multi-fidelity model is to quantify the time-variant travel time distribution (TTD). This multi-fidelity model undergoes four main steps of (i) decomposing the advection-diffusion equation (ADE) into a series of ADE’s, (ii) solving these ADEs through particle tracking without random walk (low-fidelity model), (iii) solving several ADE’s by numerical simulation (high-fidelity model), (iv) establishing the relationship relation between the low fidelity and high-fidelity models, and (v) calculating the time-variant TTD from the breakthrough curves of the multi-fidelity model. The manuscript is well written and one can follow the proposed theoretical framework smoothly. However, the superiority of the proposed multi-fidelity model over the very well-known approaches like the storage section function framework is not clear. Also, the impact on model accuracy of some major simplifications in both the low- and high-fidelity models has not been demonstrated. As such, I recommend major revision at this stage of review. Further major and minor comments are presented below.

Major comments

- Introduction

Line 49-53: This statement of is not correct because the influence of permeability architecture and preferential flow is already projected in the flow velocities and residence times, which in turn are reflected in the time-variant TTD. Therefore, the age mase equation can indeed be employed to explore the above issues by studying hillslopes with a wide range of physical conditions. Some other former studies have already explored them,
Groundwater age distribution model

The two main parts of this section is (i) the groundwater age distribution model and (ii) the approach to solve the so-called 5D equation (1). Both of them have already been contextualized in the work of Ginn (1999) and Gomez and Wilson (2013), respectively. The tone of material presentation in the section (e.g., in Line 95) looks like the authors are proposing a novel approach of treating the original equations by Ginn (1999), but it does not seem so.

Line 137-140: This part is written very concisely and is not clear to a reader. In particular, four questions are unknown: (1) Why should the high-fidelity model be run multiple times and in a limited number of runs? (2) Why should not the high-fidelity model required be run for all input pulses, but only some of them? (3) What does the “trend” of the variation in age distributions imply? And (4) what does make the multi-fidelity model biased? To clarify the above issue, the authors must first describe the elements of the multi-fidelity models and then explain how those elements work together to build the multi-fidelity model with sufficient elaboration on its properties/advantages/caveats.

Particle-tracking model

The claimed lower computational cost of the proposed low-fidelity model seems to be obtained at the cost of ignoring the molecular diffusion and mechanical dispersion of solutes. As such, the superiority/accuracy of the low-fidelity model as compared to the other competing particle-tracking models like EcoSILM is not clear. Also, please note the following studies in which some approaches to diminish the computational of particle-tracking are discussed:

Yang et al. (2021), Accurate load balancing accelerates Lagrangian simulation of water ages on distributed, multi-GPU platforms, https://doi.org/10.1002/essoar.10507899.1.


It is also noted that injecting a sufficient number of particles into the particle-tracking models is not required to only reduce the random noise of diffusion and mechanical dispersion, but also to reduce bias in the tracked evapotranspired water. Since the present study highlights its superiority in terms of employing a much smaller number of particles, the model accuracy on tracking fluxes like evapotranspiration in not clear and discussed.
- **Analytical theorem for one dimensional scenario**

The presented theoretical foundation of the multi-fidelity model in 1D seems to hold upon the assumption of constant advection velocity in the entire domain. But this is not certainty true in real-world cases and is against the goal of this study.

- **Unclear advantage of the multi-fidelity model and the limited experiments**

According to the motivation of this study, the proposed multi-fidelity model is expected to ease the estimation of TTD by skipping the complex process of SAS function quantification proposed by former studies (e.g., Botter et al., 2011). However, determining the mathematical form of the nonlinear mapping of equation (17) is said to be strongly dependent on the properties of a hydrogeological system under study. Therefore, it seems that the challenge of SAS function calibration in the formerly developed framework has only switched to quantifying another function of nonlinear type, i.e., g in equation (17), in the present study. From the designed experiments, the relative efficiency of the multi-fidelity model performance as compared to the SAS function calibration approach is not clear and appears to be exaggerated.

**Minor comments**

- Abstract, line 1: TTD is not necessarily representative of groundwater, but also surface runoff and any the processes that compose streamflow. If the focus of this work is to consider the travel time of flow pathways in groundwater, the authors should clarify it first in the manuscript.
- Abstract, line 5: The non-stationarity of TTD is not only because of non-stationarity n rainfall, but all the processes closing the water balance in a hydrological system.
- The abstract is a little bit confusing and the study storyline and objectives are not clear.
- Line 20: mixing process and also the heterogeneity of flow pathways and transient pore water velocity.
- Line 36: RTD was not defined earlier in the manuscript.
- Line 54 and 58: Is “periodic” the best word for this condition?
- Line 86: In comparison to... Please fix.
- Line 92: groundwater flow pathways... Please fix.