Dear referee #2:

General comments:

Thanks very much for taking your time to review this manuscript, and thanks for your comments and valuable suggestions. Below are our point-by-point responses to the referee's comments (in Italics). We hope you will find them to be comprehensive and satisfactory.

1. *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 maser 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, too.

   Answer: Thank for your comment. The time-variant TTD is a lumped representation of probability density of travel time in discharge. As is stated in this comment, the time-variant TTD encapsulates multiple physical properties such as the permeability architecture and preferential flow into a mathematical function, and the relations between time-variant TTD and the physical properties have been explored by several methods in previous studies [Ameli et al., 2016]. However, the age master equation model can only be applied to study these relations by experiment rather than by numerical simulation. This is because that these physical parameters are not included in the age master equation model but included in the governing equation of groundwater age distribution model. To clarify the aims of the study, we will rewrite and extend the statement (lines 49-53) as follows:

   "This approach of solving the age master equation mainly requires the information at the recharge and discharge boundaries and lacks the consideration of the intrinsic physical processes in the age master equation when calculating time-variant TTD. Therefore, the lumped model based on age master equation is not possible to study the controls of physical properties on the variation of TTD and storage selection preference by numerical simulation method, such as the influence of permeability architecture and preferential flow."

   Ameli, A. A., Amvrosiadi, N., Grabs, T., Laudon, H., Creed, I. F., McDonnell, J. J., & Bishop, K. (2016). Hillslope permeability architecture controls on subsurface transit time distribution and flow paths. Journal of Hydrology, 543, 17-30.

2. *Groundwater age distribution model*

   2.1 *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.*

Answer: We are sorry to give the impression that the derivation of the govern equation and its numerical implementation was originally done by the authors. This is indeed not purpose of section 2.1. As stated in line 90 of the manuscript, this section aims to brief the groundwater age distribution model and the approaches to solve the 5-D equation, rather than develop a new approach to solve this equation. The groundwater age distribution model in section is termed as the high-fidelity model, and is one of the components of multi-fidelity model. Therefore, section 2.1 is not the original work of this study but provides an overall description of theoretical foundation. The original work of this study starts from section 2.3 related to the multi-fidelity model.

2.2 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.

Answer: Thanks for your comment. To fully answer the four questions in this comment, we rewrite and reorganized the last paragraph of section 2.1 (line 135-140) as follows:

"The advantage of this approach is that it reduces the dimension of the problem by changing the age variable τ into the input time variable η, because the water from the same input event must have the same age. However, in real problems, the ADE model (Equation (8)) should be simulated for each input event to calculate the proportion of water parcels of all ages in storage and discharge. In this study, the ADE model (Equation (8)) for a certain input event is defined as the high-fidelity model. The high fidelity of this ADE model is compared with the fidelities of models introduced in the following subsections."

As the purpose of Section 2.1 is to introduce the groundwater age distribution model and high-fidelity model, the sentences in lines 138-140 related to multi-fidelity model are misleading and we will remove them in the revised manuscript. Meanwhile, we specifically answer each question in this comment as follows:

(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?*

The two questions are both related to the number of model runs, so we will answer them together. As is stated in the revised paragraph, the high-fidelity model should be run for each input event, i.e., be run multiple times to calculate the time-variant travel time distribution. The multi-fidelity model requires a limited number of runs of high-fidelity model. Since section 2.1 is to introduce high-fidelity model, the sentences related to the properties of multi-fidelity model are misleading and removed in the revised paragraph, and these issues are already fully discussed in line 242-250.

(3) *What does the "trend" of the variation in age distributions imply?*

(4) What does make the multi-fidelity model biased?

The two questions (3) and (4) are related to the multi-fidelity model. Building on the comments and our purpose of this section, we agree that it is misleading to present description on multi-fidelity model. We have deleted this misleading sentence in the revise manuscript. The answer to question (3) have been already stated in section 2.4. We notice that the "trend" in the original paragraph is a vague word, and a more accurate statement is presented in line 193-195, that is "the relation between the BTCs from low-fidelity model and high-fidelity model…" The answer to question (4) have been already fully described in the error analysis in Sections 3.2 and 3.3 and Figures 5 intensively.

3. *Particle-tracking model*

3.1 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

Answer: Thanks for the constrictive comment. The particle tracking model without random walk (low-fidelity model) does not consider molecular dispersion and mechanical diffusion, but the multi-fidelity model considers them. The aim of this study is to introduce the multi-fidelity model approach to calculate the time-variant TTD. We agree that the application of low-fidelity model in molecular diffusion and mechanical dispersion influenced cases can lead to inaccurate result. But this inaccurate result can be corrected by the multi-fidelity model, as illustrated in the error analysis sections 3.2, 3.3. It is a good strategy to include more comprehensive description to compare the computation efficiency of different particle tracking models. Thus, we will add the description on the comparison of the efficiency of different particle tracking models used in the provided literature with the low-fidelity model in this study (referring to line 152).

"Particle tracing model can be classified into two categories generally: particle tracking model with random walk, and particle tracking model without random walk. The particle tracking model without random walk will be taken as the low-fidelity model, and is set to be one of the components of the multi-fidelity model. Until now, there are several methods to improve the efficiency of particle tracking model (Yang et al, 2021a, b), but these are algorithm improvement, and the improved efficiency is very limited. The computation efficiency of particle tracking model without random walk can be several orders of magnitude lower than the efficiency of the particle tracking model with random walk by Yang et al, 2021a ,b. The improvement of the efficiency will be discussed after section 2.4 in detail. ”

The improvement of the efficiency is also discussed in the response to major comment 5.

[References]

YANG, C., Maxwell, R. M., & Valent, R. (2021). Accurate load balancing accelerates Lagrangian simulation of water ages on distributed, multi-GPU platforms.

Yang, C., Zhang, Y. K., Liang, X., Olschanowsky, C., Yang, X., & Maxwell, R. (2021). Accelerating the Lagrangian particle tracking of residence time distributions and source water mixing towards large scales. Computers & Geosciences, 151, 104760.

3.2 *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.*

Answer: Thanks for you notation in this comment. We agree that the injection with sufficient particle number could also reduce the bias in the tracked evapotranspired water. We have modified this sentence to make the statement more inclusive and complete as follows:
“As the number of particles is also closely related to the tracked evapotranspired water, the reduction of particle number may lead to bias of water flux in evapotranspiration. To avoid this problem, the evapotranspired water will not be tracked in the low-fidelity model, instead we employed infiltration rather evapotranspiration in the low fidelity model, and each particle represents a certain volume of infiltrated water.”

4. *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.

Answer: Thanks for your comment. This issue has been also raised by another referee. It seems that we gave a wrong impression for the purpose of this section with a misleading subtitle. We will change the section title to be “Connection between the results of low-fidelity model and high-fidelity model in one-dimensional scenario”. As is stated in the first paragraph, we actually introduce this section for the purpose of illustrating the mapping between BTCs of particle tracking model and ADE model, rather than to develop the analytical solution which can be applied to real world. According to the derived analytical solutions and equation (15), it can be found that the determination of the correlation is the diffusion parameter D. In other word, if the diffusion parameter D is determined, the mapping between the results of high-fidelity model and low-fidelity model is unique. Moreover, according to equation (15), it is difficult to find the mathematical form of mapping between the analytical solutions of the two models even in such a simple one-dimensional scenario. Undoubtedly, the mathematical form of the non-linear mapping for a much more complicated real-world case is infeasibly to be derived. To clearly clarify the conclusion of this section, the last paragraph is reorganized and extended as follows:

"Based on the analytical solutions, if the diffusion parameter D is given, the mapping between the results of high-fidelity model and low-fidelity model is unique. This is the theoretical foundation of the multi-fidelity model. However, according to equation (15), it is difficult to find the mathematical form of the mapping between the analytical solutions of the two models in such a simple one-dimensional scenario. Undoubtedly, the mathematical form of the non-linear mapping for a much more complicated real-world case is difficult to be derived."

5. *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.

   Answer: Thanks for your comment. We will reply to this comment in the two aspects: (1) the motivation of this study; (2) the advantage of the multi-fidelity model

   (1) The motivation of this study

       The initial motivation of this study is to calculate time-variant TTD by the direct numerical modeling approach, that is, to quantify time-variant TTD given hillslope physical properties (e.g., permeability architecture, stratification etc.) and corresponding boundary conditions. In the introduction section, several approaches to estimate time-variant TTD are reviewed. The age master equation model based on SAS function is an inverse modeling approach. i.e., calculating the parameters of SAS function, TTD and RTD based on long-term observed tracers which are the modeling results. As the long-term tracer data are not available for all catchments, the inverse modeling can be limitedly applied to intensive instrumented catchments.

       Then several direct modeling approaches are reviewed. The direct modeling approach can overcome the disadvantage of inverse modeling approach. The work flows of direct modeling approach are: (1) simulating water and solute transport given internal properties and boundary conditions; (2) calculate SAS function, TTD and RTD based on the simulated results. As the internal properties and boundary conditions of concerned catchment can be easily obtained for a certain catchment, the direct modeling approach has wider applications than the inverse modeling approach. The direct modeling approach mainly includes particle tracking model and groundwater age distribution. However, the computation cost including the computation time, computation memory, etc. will increase with the spatial scale or increase as the Péclet number decrease. The multi-fidelity model proposed in this study is a direct modeling approach to overcome this computational problem.

(2) The advantage of multi-fidelity model

Except the age master equation approach which is inverse modeling approaches, the other approaches reviewed in the introduction section are direct modeling approaches, including the particle tracking modeling approach, and the groundwater water age distribution modeling approach.. The particle tracking model with random walk and groundwater age distribution model are the two approaches with high accuracy, but the computation cost is sometime too high to differentiate the water parcels and solute particles of different ages (or different entering time). When considering molecular diffusion and mechanical dispersion, the computational cost can increase largely as the spatial scale of the model increases. However, the computation cost of particle tracking model without random walk can be reduced dramatically by sacrificing the accuracy of the model, i.e., by ignoring the influence molecular diffusion and hydrodynamic dispersion. Therefore, the multi-fidelity model is proposed in this study, and it combines the accuracy of high-fidelity model and the efficiency of low-fidelity model. Specifically, in the model with low Péclet number and large spatial scale, the number of particles required for multi-fidelity model can be reduced significantly, compared the particle tracking model with random walk. As for the usage of computation memory, if the transport is strongly influenced by molecular diffusion and mechanical dispersion (low Péclet number scenario), or if the model has a large spatial scale, the multi-fidelity model can reduce the number of particles significantly compared to the particle tracking model with random walk (How significantly the number of particles can be reduced will be discussed in the following paragraph in detail). This reduction of the usage of computation memory can be far more than the reduction by simply using a coarser grid as stated in this comment. The reduction of the usage of the computation memory allows the multi-fidelity model to be run at most personal computers, rather than to be run parallelly in a specific cluster system.

To clarify and highlight the advantage of the multi-fidelity and show the case ‘*where we really need the multi-fidelity model*’, line 244-250 will be fully reorganized and extended as a separate section as follows:

“Section 2.5 The advantage of the multi-fidelity model

The advantage of multi-fidelity model is manifested in two aspects: (1) advantage compared to the inverse modeling approach; (2) advantage compared to the approaches in direct modeling approach.

Firstly, the multi-fidelity model is a direct modeling approach which calculates time-variant TTD, RTD and SAS function by simulating water and solute transport in the system with the given physical properties and boundary conditions. The input data of direct modeling approach can be readily obtained for most of the catchment compared to the input data of inverse modeling approach which requires long-term observational data i.e., gauge data (flow rates, rainfall etc.) and tracer data in both rainfall and discharge.

Secondly, the multi-fidelity model consists of three component: low-fidelity model, high-fidelity model and non-linear Gaussian process regression model, and it takes advantage of both the accuracy of high-fidelity model and the low computation cost of low-fidelity model; and it also makes the evaluation of time-variant TTD to be easily implemented in real coordinate space compared with the complex approach that evaluates time-variant TTD numerically in Laplace space (Cornaton, 2012). The computation cost includes computation time and usage of computation memory. As for the computation time, the number of rainfall events considered in ADE model of the multi-fidelity model is reduced compared to the groundwater age distribution model; and the smaller number of particles in low-fidelity model also reduces the computation time. As for the usage of computation memory, when the spatial scale of model increases, the number of particles is reduced significantly compared with the random walk particle tracking model. Specifically, one particle is enough to simulate the transport of a certain water parcel injected at a certain time and a certain point in the model without molecular diffusion and mechanical dispersion. However, in the model considering molecular diffusion and mechanical dispersion, it is likely that more than 10 particles are required to simulate the transport of that water parcel to reduce random noise at laboratory scale; it is also likely that more than 100 particles are required to simulate the transport of that water parcel at hillslope scale; it is even likely more than 1000 particles are required to simulate the transport of that water parcel at large river basin scale. Given a certain Péclet number, the increase of particle number (10 - 100 - 1000) from laboratory scale to large basin scale depends on the increase of the spatial scales of the models. Besides, if the spatial scale of a model is given, the particle number also depends on the Péclet number, and the particle number will increase as the Péclet number decreases (i.e., the increase of the influence of diffusion). The significant reduction of the number of particles leads to the significant reduction of the usage of the computation memory and allows the multi-fidelity model to be run on most personal computers, rather than to be run parallelly in a specific cluster system.”

6. *Minor comments*
   6.1 *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.*

Answer: Well taken. Line 1 is a general definition of TTD and we should not only focused on groundwater, so we will change it as follows:

"The travel time distribution (TTD) is a lumped representation of water leaving the system responding to external forces such as rainfall."

6.2 *Abstract, line 5: The non-stationarity of TTD is not only because of non-stationarity rainfall, but all the processes closing the water balance in a hydrological system.*

Answer: Well taken. Line 5 introduces the general origin of time-variant TTD and we should not only focus on non-stationary rainfall input. We will change this sentence as:

"Under nonstationary hydrologic conditions such as nonstationary rainfall input, the TTD varies with transit groundwater flow system, leading to a time-variant TTD."

6.3 *The abstract is a little bit confusing and the study storyline and objectives are not clear.*

Answer: Thanks for your comment. We noticed the abstract is not concise and the objective is not clear. We have make effort to modify parts of abstract to make it clearer as follows::

"…Most methods for estimating time-variant TTD can be classified into two categories: (1) inverse modeling approach; (2) direct modeling approach. The inverse modeling approach such as the age master equation model requires long-term continuous tracer data which is only available for very limited number of catchments. The direct modeling approach always requires intensive computation to differentiate water parcels or solute particles of different ages…"

6.4 *Line 20: mixing process and also the heterogeneity of flow pathways and transient pore water velocity.*

Answer: Well taken, this sentence will be revised as follows:

"Distinct water parcels can have different travel time due to the mixing processes, the heterogeneity of flow pathways and transient pore water velocity."

6.5 *Line 36: RTD was not defined earlier in the manuscript.*

Answer: Thanks for your comment. We will define it before line 35 as follows:

"In the age master equation, the residence time distribution (RTD) is defined as age distribution of water residing in the system, and the backward TTD…"

6.6 *Line 54 and 58: Is "periodic" the best word for this condition?*

Answer: Yes, We would like to use it term as mentioned in previous studies (see line 53).

6.7 *Line 86: In comparison to… Please fix.*
6.8 *Line 92: groundwater flow pathways… Please fix.*

Answers: Well taken, we will revise them as suggested.