Response to the editorial comments for the Manuscript entitled "**On Parameter Bias in Earthquake Sequence Models using Data Assimilation**."

The authors are thankful to the editor for his insightful guidance in making us improve the quality of the manuscript and for considering us for possible publication in the esteemed journal of 'Nonlinear Processes in Geophysics'**.** The authors are thankful to the reviewers for their helpful comments. Following the comments of the reviewers, the necessary changes are made in the revised manuscript.

Comments from Reviewer 1

Comment 1:
Synthetic observations are produced by sampling from the synthetic truth and adding an observational error from a Gaussian distribution with standard deviation. However, the real observations could be affected by instrumental noise, missing data, spikes, etc, and a short time step of four time units may no longer be applicable. I understand that the authors lack real observations. But they should at least discuss this limitation in Section 5.

Response:
The authors thank the reviewer for the comment. Indeed, the assumption of data availability (once every four-time units) and the assumptions on the standard deviation and distribution of the observational error may not be valid when assimilating real data. We have mentioned this as a limitation in our discussion (page 21, line 348-356) as:

An additional point worth mentioning is the use of synthetic observations for fault displacement and velocities for data assimilation in this study. In realistic applications, the assumptions that we have considered with respect to data assimilation frequency and the standard deviation and distribution of the observational errors may not be valid. However, if we know the distribution of the measurement errors, we can use that information in choosing the relevant likelihood function that can greatly affect our fault estimates. Fault shear stress observations are usually not available, and if they are, they are subject to large errors. In contrast, fault velocities can be observed fairly accurately using GPS, as discussed by Van Dinther et al. (2019), who demonstrate that stress measurements are useful despite of their large errors. Following van Dinther et al (2019), we emphasize the need for additional sensitivity studies to understand the implications of data gaps, outliers, and instrumental noise before our proposed methods can be used on real data.

Comment No. 2:

In introduction, please review some previous studies where either the frictional parameters have been estimated as part of the data assimilation or assumed to be perfectly known.

Response:
The authors have included some references which estimates frictional parameters using data assimilation and are mentioned in the manuscript as (page 2, line 35-37):

‘On the other hand, Van Dinther et al. (2019) and Diab-Montero et al. (2022) assumed the frictional parameters to be known and used an ensemble Kalman filter to estimate the fault states.’

Comment No. 3:

Section 2.1: Please specify how to quantitively determine the observation noise error. Please review some data assimilation frameworks and explain the reason why this framework is selected.

Response:
Observational errors can be determined by comparing observations of velocity or displacement with independent observations of these variables. This text is added to the manuscript. In the introduction, we motivate the choice for using ensemble-based data assimilation methods especially particle filters (page 2, line 38-44). For a further review of data assimilation frameworks, we refer the reader to the recent book by Evensen et al (2022). This reference has been added to the respective text in the introduction as (line 40-41, page 2):

For further discussion of available data-assimilation methods, we refer the reader to Evensen et al. (2022).

Comment No. 4:

Equation 4: What does j mean? Is it a typo?

Response:
The authors appreciate the reviewer’s comments. It is a typo and it should be i which is the number of realization. It has been corrected in the manuscript.

Comment No. 5:
Line 89: In the presence of filter degeneracy, how to guarantee that one or few particles with high weight are sufficiently representative as the input?

Response:
In degeneracy, the weight of one particle is close to one while the weight of all the other particles is close to zero. In this case, a single particle represents the filtered distribution, which results in an extremely poor approximation. Hence it is important to avoid filter degeneracy by (i) adding jitter in the prior distribution and (ii) using resampling step in particle filter. In the present work, we have included both to avoid filter degeneracy.

Comment No. 6:
Line 96: The sequential importance resampling process duplicates particles with high weight. Please explain its physical meaning in data assimilation.

Response:
The authors appreciate the reviewers' comments. In this implementation of the particle filter, the sequential importance resampling attributes higher weight to particles that are closer to the observations. This is done by multiplying the prior by the likelihood, which can be considered a weight function (in this study, a Lorentz function). Then, in the resampling step, the

importance resampling process removes those particles which have low weight in the distribution and thereby retains only those particles which have a higher weight. These are then duplicated according to their weight, in such a way that the number of particles remains constant. This ensures an approximation of the prior distribution that is less sensitive to particle degeneracy. We have mentioned this in text from line 102-105 in page 4.

Comment No. 7:
Section 2.2: The model of forwarding simulation is important to data assimilation. In this study, a zero-dimensional (0D) model is considered. However, 1D, 2/3D models are also available. Please specify the reason why 0D model is selected. More details of its pros and cons are expected.

Response:
The authors appreciate the comment of the reviewer. We have explained the reason of choosing a simplified model for this study in detail from line 357-366 (page 21) in the discussion of the manuscript as:

It is also very important to highlight the reason behind selecting a zero dimensional (0D) model for this study. Simplified fault slip models are computational efficient tools that help us to understand the physics behind the earthquake dynamics. A study by Li et al. (2021) compares the simulation of earthquakes in 0D, 1D, 2D, and 3D models and finds that lower-dimension models (0D and 1D) qualitatively represent the same dynamics as 2D and 3D models. Although 0D models cannot simulate the full complexity of the earthquake physics, they have the advantage that they are computationally inexpensive and provide the user with a tractable conceptual description of earthquake physics and the importance of the friction parameter. In our case, we were interested in investigating the effect of frictional parameter bias on the estimated fault states in earthquake cycle models. Hence in the present work, we investigated a simplified version of a Burridge-Knopoff spring-block slide model in a simplified 0D form. Eventually, to accurately simulate the behavior of real earthquake faults, 1D, 2D and 3D simulation models will be required (e.g., Li et al., 2021).

Comment No. 8:
Line 149: What if in the region a-b>0?

Response:

The parameter combination (a – b) < 0 corresponds to steady state slip rate-weakening properties causing an unstable rapid slip (frictionally unstable), while (a – b)> 0 corresponds to the steady state velocity- strengthening behaviour, causing a stable slip (frictionally stable). Since we are primarily interested in frictionally unstable earthquake cycles, we have focused on parameter combination for (a – b) < 0. According to Ruina (1983), if a velocity-strengthening system experiences a slip instability, the motion will be rapidly dampened down to a state of stability. A velocity-weakening system, on the other hand, will no matter how carefully driven, always exhibit growing oscillations and reach a state of regular stick slip (Scholz (2019)). The velocity-strengthening behaviour is thus intrinsically stable. For this reason, we have not investigated the case of (a-b) > 0 in this study.

Comment No. 9:
Section 3.2: The assimilation step may have an important effect on the results. In this study a very short time step is adopted. Please provide more discussions on its effect. If a longer time step is used, can a small parameter bias still be compensated?

Response:

Having large assimilation steps can also have a detrimental effect on the data assimilation process as it can miss characteristic variations of the earthquake cycle. A parameter bias can have a substantial effect on the evolution of the state variables, which may be difficult to correct if the assimilation step is large. Hence a short time step is to be chosen that allows the assimilation to capture the important characteristics of the earthquake cycle.

Comment 10:
Discussion: I appreciate the authors' efforts in stating the limitations of this study, but here I expect more discussion on their results and comparison with previous studies (without data assimilation).

Response:

We extended the discussion with a section that discusses our results in relation to previous studies without data assimilation (line 330-339, page 21) as:

Typically, earthquake forecasting is approached in a probabilistic manner (e.g., Marzocchi et al., 2017). Kinematic inversions of earthquake global positioning system (GPS) data have been used to estimate frictional properties in afterslip areas (Miyazaki et al., 2004; Hsu et al., 2006), but not for estimation of the earthquake dynamics themselves. As outlined by Van Dinther et al (2019), data assimilation for earthquake sequences has the advantage that it can take into account measurement and model errors, non-Gaussian probabilities and sequential updating as data becomes available. The results of this study demonstrate how, for a highly simplified representation of earthquake cycles, non-linear data assimilation provides a means to account for both measurement errors and parameter biases. It also highlights how observations can be included as they become available. While particle filters are not computationally efficient, they can propagate the full error distribution which makes them attractive for estimation and forecasting of highly nonlinear processes like earthquake generation.

References:

Evensen, G., Vossepoel, F. C., & van Leeuwen, P. J. (2022). Data Assimilation Fundamentals: A Unified Formulation of the State and Parameter Estimation Problem.

Hsu, Y. J., Simons, M., Avouac, J. P., Galetzka, J., Sieh, K., Chlieh, M., ... & Bock, Y. (2006). Frictional afterslip following the 2005 Nias-Simeulue earthquake, Sumatra. *Science*, *312*(5782), 1921-1926.

Li, M., Pranger, C., & van Dinther, Y. (2022). Characteristics of Earthquake Cycles: A Cross-Dimensional Comparison of 0D to 3D Numerical Models. *Journal of Geophysical Research: Solid Earth*, *127*(8), e2021JB023726.

Miyazaki, S. I., Segall, P., Fukuda, J., & Kato, T. (2004). Space time distribution of afterslip following the 2003 Tokachi-oki earthquake: Implications for variations in fault zone frictional properties. *Geophysical Research Letters*, *31*(6).

Ruina, A. (1983). Slip instability and state variable friction laws. *Journal of Geophysical Research: Solid Earth*, *88*(B12), 10359-10370.

Scholz, C. H. (2019). *The mechanics of earthquakes and faulting*. Cambridge university press.

van Dinther, Y., Künsch, H. R., & Fichtner, A. (2019). Ensemble data assimilation for earthquake sequences: probabilistic estimation and forecasting of fault stresses. *Geophysical Journal International*, *217*(3), 1453-1478.