We thank the anonymous commenter for the careful reading of the manuscript and for the constructive suggestions provided during the interactive discussion. Below we address the points raised, with clarifications focused on the scope, assumptions, and objectives of the present study.

Reviewer's comment:

(1) The structural features (faults) are missing in the figure 1. It can also include another parameter i.e. focal depth by filling the varying color in circle depicting earthquake location and magnitude.

Reply to comment 1:

We thank the commenter for the suggestion. We would like to clarify that the present study does not investigate fault geometry or fault-controlled stress interactions. The analysis is intentionally formulated in a two-dimensional framework, focusing on stress evolution within depth-dependent horizontal slices rather than on fault-specific variables.

Figure 1 is designed to define the study domain and to illustrate the spatial distribution of seismicity used as initial conditions for the stress evolution model. Introducing mapped fault traces would imply an explicit role of fault geometry in the modeling, which is beyond the scope of this work and could potentially mislead readers regarding the variables considered.

Similarly, focal depth is treated in this study through separate depth-resolved analyses rather than as a continuous color-mapped parameter on a single map. Each depth slice is analyzed independently to isolate depth-dependent stress evolution behavior.

For these reasons, we retain Figure 1 as a domain and data-context figure, while depth effects are addressed explicitly in subsequent sections. We have clarified this scope distinction in the revised manuscript to avoid potential confusion.

Reviewer's comment:

(2) Terms like $n_{critical}$, i.e., the iteration where diffusion dominates, are mentioned but not mathematically derived (Section 2.3). Further, a brief formula or sensitivity equation could strengthen why STD peaks differ by depth. These explanations feel descriptive rather than fully predictive and sufficient for intuition, which assists a better understanding without digging into earlier equations.

Reply to comment 2:

We thank the reviewer for the detailed comments. To characterize the transition from reaction-dominated to diffusion-dominated regimes, we employ a scaling-based balance argument between the reaction and diffusion contributions in the iterative stress evolution.

Assuming a constant stress-memory parameter $\alpha \in (0,1)$, the stress amplitude at the n-th iteration is approximated by

$$\sigma_n \approx \sigma_0 \alpha^n$$

where σ_0 denotes the initial stress perturbation.

The reaction contribution at iteration n scales with the local stress amplitude and can be written as

$$R_n \sim A_0\sigma^n = A_0\sigma_0\alpha^n$$

where A_0 represents the effective reaction strength. The diffusion contribution is associated with the smoothing of the initial stress heterogeneity and is characterized by

$$D_n \sim D_0\Delta\sigma_0$$

where D_0 is the diffusion coefficient and $\Delta\sigma_0$ denotes a characteristic measure of the initial stress gradient.

We define the critical iteration index $n_{critical}$ as the iteration at which the reaction and diffusion contributions become comparable in magnitude, i.e.,

$$R_n \sim D_n$$

Substituting the above expressions yields

$$A_0\sigma_0\alpha^n = D_0\Delta\sigma_0$$

Solving for n gives

$$\alpha^n = (D_0\Delta\sigma_0)/(A_0\sigma_0)$$

and therefore

$$n_{critical} = \frac{\ln\left(\frac{A_0\Delta\sigma_0}{D_0\sigma_0}\right)}{-\ln(\alpha)}$$

Because $\ln(\alpha)<0$, this expression indicates that $n_{critical}$ increases with stronger initial stress levels or larger reaction efficiency, and decreases with stronger diffusion or weaker initial heterogeneity.

Reviewer's comment:

(3) In my understanding, the STD and MSE are used effectively for model internals and the persistence is inferred from simulations, not directly compared to observed aftershock decay (e.g., Omori's law rates from the Hualien CatLog). Like aftershock patterns (e.g., clustering at 5–15 km), but section 3.2 could link STD trends more explicitly to empirical metrics like aftershock productivity or temporal decay exponents. I think, it would make the heterogeneity explanation more robust, especially since the model aims for aftershock prediction.

Reply to comment 3:

We thank the commenter for the thoughtful suggestion. We would like to clarify the intended role of STD and MSE in the present study. These quantities are introduced as internal diagnostic measures of the stress-evolution model, designed to characterize the persistence and smoothing of stress heterogeneity under controlled variations of α, β and depth.

In this framework, STD quantifies the spatial variability of the continuous stress field and is used to compare relative evolution behaviors between shallow and deep layers, rather than to represent aftershock rates or event counts. As such, STD and MSE are not directly comparable to catalog-based statistical measures such as Omori–Utsu decay exponents, which describe temporal variations in discrete earthquake occurrence rates.

While empirical aftershock laws provide important observational constraints, establishing a direct mapping between stress-field diagnostics and event-based statistics would require an additional triggering or catalog-generation model, which is beyond the scope of the present work. The objective here is to demonstrate how stress heterogeneity evolves and stabilizes within the proposed PDE framework, rather than to fit observed aftershock decay curves.

We have clarified this distinction in the revised manuscript to avoid potential confusion between model-internal diagnostics and observational earthquake statistics.

Reviewer's comment:

(4) The STD captures spatial spread (heterogeneity), and MSE tracks convergence (persistence decay), but they do not fully address temporal aspects. For example, the paper notes iterations that are not directly related with time (shallow: hours/days per iteration; deep: months/years), i.e. it does not quantify time-scaling factors. This leaves persistence explanations a bit abstract—sufficient for parameter sensitivity, but could benefit from a time-calibrated example.

Reply to comment 4:

We thank the commenter for raising the point regarding temporal interpretation. We would like to clarify that, in the present study, STD and MSE are not intended as time-dependent observables, nor are they designed to quantify physical time scales. Instead, they are internal diagnostic metrics used to characterize the *state of the evolving stress field* under controlled numerical iterations.

Specifically, STD measures the spatial spread (heterogeneity) of the stress field after evolution, while MSE quantifies the degree of convergence toward a stable configuration (persistence decay**)**. Both quantities are evaluated in iteration space and are intentionally defined to be independent of physical time, allowing a clean comparison of stress evolution behavior across depths and parameter settings.

The coupling between iterations and physical time is addressed separately in the manuscript. For shallow depths (above ~18 km), temporal effects are incorporated only through the time-dependent memory parameter $\alpha(t)$, which is linked to Omori-type aftershock decay to represent the progressive loss of historical stress memory. The corresponding adjustment of $\beta(t) \sim 1/\alpha(t)$ ensures balanced evolution between diffusion and nonlinear reaction terms. This $\alpha(t)$–$\beta(t)$ coupling constitutes the sole time-control mechanism in the present framework.

At greater depths (>18 km), where seismicity becomes sparse and Omori-type decay is no longer applicable, no direct time calibration is imposed, and the evolution is interpreted purely in terms of internal stress redistribution rather than explicit temporal progression. In such cases, iterations represent abstract relaxation steps, and only averaged or qualitative time interpretations are appropriate.

Therefore, while a fully time-calibrated example could be considered in future applications, it is beyond the scope of the present study, whose primary objective is to establish and validate a stable, physically interpretable stress-evolution framework. The current treatment is sufficient for parameter sensitivity analysis and for comparing depth-dependent stress evolution behaviors without introducing additional assumptions on poorly constrained deep-crustal time scales.

Reviewer's comment:

(5) STD and MSE trends with β at fixed α=0.75 (Figure 3a, b) and with α at fixed β=10 (Figure 3c, d) is explored thoroughly in result section. However, by exploring β variations across multiple α values, can we expect the following?

[1] The peak STD occurs at different β (e.g., β=10 at 6 km vs. β=5 at 30 km with fix α=0.75). The higher α increases overall STD and potentially alters the optimum value of β at which diffusion dominates over reaction. Thus, under varying α, at shallower depths require higher β for smoothing, and could it highlight depth-specific couplings?

[2] A large value of α slows stress decay, limiting the β's effective range to avoid divergence. Can we identify optimum parameter pairs by varying both parameters together that could optimize convergence rate without overshoot?

[3] The α governs memory decay like R-S aging, while β scales diffusion/reaction like KPP terms (Table 1). Could the coupled variations better capture time-dependent behaviors (e.g., decreasing α(t) with increasing β(t) for aftershock decay) and provide a deeper understanding of stress redistribution in heterogeneous crust?

[4] Is it possible that the extensions of analysis of explicitly 2D plot of β-α variations plot, may provide the refined model calibration, especially for real-time PSHA integration?

Reply to comment 5:

[1] We thank the commenter for the detailed observation regarding the dependence of STD and MSE on α and β. We would like to clarify that, in the present framework, neither STD nor MSE is introduced as an optimization target, and therefore the concept of an "optimal" β (or α–β pair) does not have a physical meaning in this study. Each admissible combination of α and β represents a distinct stress-evolution regime, provided that the numerical stability and convergence conditions are satisfied. STD serves solely as a descriptive metric characterizing the degree of spatial heterogeneity of the evolved stress field under a given parameter setting, while MSE quantifies convergence behavior. Higher or lower STD values do not imply better or worse performance, nor do they define an optimal state.

The fixed-α and fixed-β analyses presented in Figure 3 are intentionally designed to isolate the individual roles of memory decay (α) and diffusion–reaction balance (β), rather than to identify depth-specific optimal couplings. While varying β across multiple α values would indeed shift the location of STD peaks, such shifts simply reflect different parameter regimes and do not correspond to physically preferred or depth-dependent optimal values.

Importantly, the dominance of diffusion over reaction is governed by admissibility and stability constraints (e.g., CFL-type conditions), not by the maximization or minimization of STD. Consequently, extending the analysis toward multi-parameter optimization would introduce an artificial objective function that is not supported by the physics of the proposed model and is therefore beyond the scope of this study.

[2] We thank the reviewer for raising the question regarding the joint variation of α and β and the possibility of identifying an "optimal" parameter pair. We would like to clarify that within the

present framework, the roles of α and β are defined primarily by stability and convergence guarantees, rather than by optimization criteria. Specifically, for $\alpha<1$ and β satisfying the CFL-type stability condition, the system is mathematically guaranteed to converge toward a bounded attractor. Under these conditions, the admissible (α, β) space is continuous and infinite, and there is no unique or physically meaningful "optimal" pair in terms of convergence speed. Moreover, the iteration index in our model represents a dimensionless evolution measure rather than physical time. As a result, the notion of "fastest convergence" depends on numerical scaling choices and does not carry an unambiguous physical interpretation. For this reason, we intentionally avoid defining or optimizing a specific convergence rate and instead focus on demonstrating robust convergence behavior across a broad admissible parameter domain. We have clarified this modeling philosophy in the revised manuscript to avoid potential misinterpretation of the parameter roles as optimization targets.

[3] On the coupling between α(t) and β(t).

In the present framework, the parameters α(t) and β(t) are not treated as independent degrees of freedom. Instead, they represent complementary aspects of the same physical process governing post-seismic stress evolution. The parameter α(t) controls stress memory and is associated with the persistence of historical stress influence, analogous to aging effects in rate-and-state friction. As post-seismic evolution proceeds, the influence of past stress perturbations necessarily diminishes, leading to a monotonic decrease in α(t).

Conversely, β(t) characterizes diffusion and nonlinear smoothing effects, which become increasingly dominant as stress heterogeneity is redistributed and dissipated. This complementary behavior motivates a coupled representation in which decreasing α(t) is accompanied by an effectively increasing influence of β(t). For numerical stability and physical consistency, this coupling is implemented through a reciprocal scaling (e.g., $\alpha(t) \sim 1/\beta(t)$), ensuring bounded evolution toward an attractor while capturing the observed decay of aftershock activity. In the present study, fixed effective values are employed to isolate depth-dependent effects, while the coupled time-dependent formulation represents the underlying physical mechanism.

[4] We thank the reviewer for the suggestion regarding an explicit two-dimensional exploration of the (α,β) parameter space. We would like to clarify that in the present framework, α and β are not treated as calibration parameters in the sense of parameter inversion or optimization.

For $\alpha<1$ and β satisfying a CFL-type stability condition, convergence toward a bounded attractor is mathematically guaranteed. Under these conditions, the admissible (α,β) space is continuous and infinite. Consequently, there is no unique or physically meaningful mapping between a given α and a specific β, nor a well-defined optimal region in the α−β plane. A two-dimensional parameter scan would therefore only reproduce the known stability domain, without providing additional physical insight or refined calibration. In particular, such a scanwould not be suitable for real-time PSHA integration, which relies on seismicity rates, hazard functions, and uncertainty aggregation rather than on internal numerical control parameters of a stress-evolution model. The objective of this study is to demonstrate robust stress evolution behavior within a mathematically admissible parameter domain, rather than to perform parameter optimization or real-time calibration. We have clarified this modeling scope in the revised manuscript to avoid potential misinterpretation.

Reviewer's comment:

(6) The analysis assumes uniform parameter ranges (α=0.5–0.9, β=0.5–20) across depths, but real crust varies (e.g., rock rheology, fluids). The explanation is sufficient for the Hualien case, but generalizability to other tectonic settings (e.g., subduction zones) is not discussed, which might limit its perceived completeness.The abstract mentions time-dependent α(t)/β(t), but the provided sections use fixed values—clarify if/when time-dependence is implemented, or emphasize fixed as an approximation.

Reply to comment 6:

We thank the commenter for the remarks. We would like to clarify that the parameter ranges (α = 0.5–0.9, β = 0.5–20) are not assumed as uniform or fixed physical properties of the crust across depths, nor are they adopted as a single parameterization. Instead, these ranges are deliberately selected to conduct controlled parameter-sensitivity analyses, in which α and β are varied independently (i.e., fixed α with varying β, and fixed β with varying α) in order to examine how stress evolution responds to memory decay and diffusion–reaction balance at different depths.

The purpose of this design is to compare depth-dependent stress evolution behaviors under admissible parameter regimes, using STD and MSE as internal diagnostic indicators of spatial heterogeneity and convergence, respectively. Importantly, no single α–β pair is treated as representative of a specific geological condition, and the analysis does not rely on fixed parameter values. Regarding time dependence, the coupling of α(t) and β(t) is explicitly implemented and demonstrated in Figures 4–7. In these sections, α(t) is prescribed to decay in accordance with Omori-type aftershock behavior to represent progressive loss of historical stress memory, while β(t) is adjusted correspondingly (β(t) ~ 1/α(t)) to regulate the balance between diffusion and nonlinear reaction terms. These figures therefore already illustrate the time-dependent formulation referenced in the abstract. The analyses using fixed α or β values are intentionally employed as diagnostic experiments to isolate individual parameter effects and should be interpreted as methodological probes rather than physical approximations.

Finally, the present study focuses on establishing and validating a general stress-evolution framework through a representative case study (the 2018 Hualien earthquake). While parameter domains may shift in other tectonic settings (e.g., subduction zones), the mathematical structure and admissibility principles of the framework remain unchanged. Detailed recalibration for different geological environments is therefore left for future applications and is beyond the scope of the current manuscript.

.

Reviewer's comment:

(7) In section 3.3, the gradients are more predictive in the case of early (<50 days, <12 km) and absolute changes and backed by AUC/Molchan. A table summarizing these metrics across

time/depth bins for: (i) static initial ΔCFF, (ii) dynamic ΔCFF(t), (iii) ∇σ(t), and (iv) combined. It may provide concrete evidence (e.g., AUC improvement of X% for gradients in early phase) and address why transitions occur.

Reply to comment 7:

We thank the commenter for the suggestion regarding a tabulated summary of predictive metrics. We would like to clarify that, in the present study, static ΔCFF, dynamic ΔCFF(t), and ∇σ(t) are continuous two-dimensional stress fields, rather than discrete scalar predictors. As such, they do not admit a meaningful representation in the form of a simple table across time–depth bins.

The predictive performance of these stress measures is evaluated through AUC and Molchan-type analyses, which are specifically designed to assess the statistical consistency between a continuous stress field and observed aftershock locations. Accordingly, the AUC is defined as a curve-based diagnostic, reflecting how well each stress metric discriminates aftershock locations relative to background space. For this reason, the temporal evolution of predictive performance is already presented in an appropriate and complete manner in Figure 10, where the AUC curves are shown as a function of time for different stress measures. This representation preserves the full statistical information of the evaluation, whereas a tabulated format would require arbitrary discretization and would risk obscuring the continuous nature of the underlying stress fields.

The observed transitions in predictive skill—such as the enhanced performance of stress gradients during the early post-seismic phase at shallow depths—are therefore interpreted directly from the AUC curve evolution, rather than from aggregated summary statistics. Introducing an additional table would not provide new physical insight beyond what is already captured by the curve-based analysis and is thus not pursued in the present manuscript.

Minor Query:

- Banach Fixed-Point Theorem can be cited in line number 157.

Reply to comment:

We will add the cited paper in line number 157 of the revised manuscript.

- The statement “Physically the contraction property reflects the Earth’s inherent damping mechanisms such as viscoelastic relaxation and afterslip” can be cited.

Reply to comment:

We thank the commenter for the suggestion regarding citation. We would like to clarify that the statement in question is intended as a physical interpretation of the mathematical contraction property established in the present framework, rather than as a previously demonstrated physical law. The contraction property itself follows rigorously from the mathematical structure of the proposed stress-evolution system. The reference to viscoelastic relaxation and afterslip is introduced to provide physical intuition, highlighting that long-term postseismic deformation in the Earth is widely understood to be governed by dissipative processes.

To avoid any possible misunderstanding, we have revised the text to explicitly frame this statement as an interpretative link and have added representative references documenting the existence of viscoelastic relaxation and afterslip as damping mechanisms in the crust (e.g., Dieterich, 1994; Chan and Stein, 2009). These references are cited as background physical context, not as direct derivations of the contraction property.

- Please mention the figure number properly in line numbers 312 and 315.

Reply to comment:

The figure references at Lines 312 and 315 have been corrected. The figures are now consistently referred to as Figure 3(a)–(d) in the revised manuscript.