Responses to Referees

Neng Lu

Contents

1 Response to RC1

G1, what the advantage of the ALE IB is over the conventional ALE? I would like to see some more discussion about ALE vs ALE IB, because they seem like the two real contenders.

Response: Generally, the ALE-IB scheme offers similar, though slightly less, accuracy compared to ALE and also encounters frequent remeshing challenges. However, it has several advantages: (1) It maintains stable and realistic surface evolution, even in challenging scenarios such as large asymmetric deformations (see Figure 7 in Experiment 3: Delamination). (2) By placing the free surface at the internal boundary of the computational domain, it effectively handles issues related to space constraints when moving particles. This is particularly beneficial when considering coupling with surface processes or other factors that reshape the surface, involving complex materials (such as air, rock, and sediments) exchanges. We have a related paper discussing the coupling framework within this ALE-IB approach (Lu et al., 2026).

G2, I do not think, however, that the performed experiments are sufficient to assess which method is most stable. Firstly, the issue is that numerical stability has not been properly defined. Secondly, if the authors really want to convince that their proposed method is more stable than conventional methods (i.e, than the ALE and Eulerian methods) then this should be motivated by either theoretical arguments (which I recon is quite challenging and may be outside the scope the paper), or by performing a proper numerical investigation where the time-step size is compared the norm of a evant prognostic output variable (e.g., the surface height). However, I only think such a study is interesting if the authors have reason to believe there are other sources of instability than overstepping the relaxation time (which is already taken care of by the FSSA), and which are ameliorated by the ALE IB method. If other sources of numerical instability are present, then this should be clearly communicated.

Response: We have included descriptions of numerical stability and the L2 norm test in Experiment 1: Topography Relaxation Models, as these allow for comparison with the analytical solution. Generally, the ALE-IB scheme is slightly less accurate than ALE but more accurate than the Eulerian method in this experiment. All methods experience the 'drunken sailor' instability, which can be addressed by implementing the Free Surface Stabilization Algorithm (FSSA) like the ones in Kaus et al. (2010). Both ALE and ALE-IB are able to use a simplified version of FSSA, as detailed by Andrés-Martínez et al. (2015), to apply stress boundary conditions only on the free surface (the top boundary in ALE and the internal boundary in ALE-IB). Since their paper provides an in-depth discussion of FSSA, we have not extended this discussion further here.

G3 Relating to General comment 2, I'm curious about Experiment 3, where you find ALE IB to be more stable than ALE. I'm wondering if you also experimented with different FSSA parameter, and if potentially a different choice control parameter would have influenced this conclusion? Also, it was not clear to me whether you also used the FSSA for the ALE and Eulerian methods, and in those cases did you use the Kaus, Mühlhaus, and May (2010) version?

Response: In our experiments, we did not use the FSSA, except for the test in Experiment 1 shown in Figure 5. Instead, we used a small time step to avoid the 'drunken sailor' issue. In Experiment 3, the ALE method exhibits strong asymmetry and fails to converge, even with a small time step, due to the presence of denser material on the left half of the model. Conversely, the ALE-IB method handles this instability effectively because the top boundary over the air layer is free-slip, fixing the vertical velocity, which stabilizes the layers below.

We conducted tests with different FSSA parameters and found that a parameter value of 0.5 was empirically better in our cases. However, as discussed in other papers, the optimal choice might be 2/3 or another value. Therefore, we did not include further tests regarding these parameters in our paper, as it is not the main discussion point. The FSSA used in Experiment 1 within ALE-IB demonstrates this scheme has the ability to handle a simplified FSSA version (similar like the one in Andrés-Martínez et al. (2015)). The Kaus2010 version of FSSA has been implemented in UW2 within an Eulerian scheme (but not suitable for methods ALE or ALEIB that involving mesh deformation), which are not used in the models of the paper.

G4, I'm missing a quantitative accuracy comparison between the proposed scheme and conventional methods. In particular, I would like to see Fig. 5 supplemented with plots of the error for the other methods (i.e., ALE and Eulerian).

Response: We added the plots of the results from other methods.

G5, Section 2.1 is incomplete and should be supplemented with a brief explanation of the physical meaning behind each equation (including boundary conditions). I also think the part of Sect. 4.5, should be included in Sect. 2.1, because in the current form it requires skipping back-and-forth between the sections to disentangle how Eq. (1a) re duces to an equation for the velocity and the pressure. Furthermore, I find it unnecessarily complicated to introduce three different tensors.

Response: We have extended Section 2.1 to include more explanations about the physical meaning. We retained Section 4.5, as the subduction experiment is the only model using non-linear viscosity and temperature-dependent density. All other experiments use linear viscosity and density. We also made adjustments to Equation 1a regarding the tensor definitions.

Specific comments:

P2:L34, removed the equation.

P3:Eq. (1c) Maybe this is standard notation, but I find it slightly confusing to refer to the heat production as H as this could be mistaken for a geometric parameter. Perhaps Q is more appropriate/standard?

Response: H is more commonly used, so we kept it for consistency with other references Moresi et al. (2007); Crameri et al. (2017). And we prefer using h to denote height.

P3:L74, corrected to 'identity matrix'.

P4:L106–109, add the explanation of the physical significance of the isostatic compensation factor. "C_{isost} is a nondimensional combination of geometric and material parameters that quantifies the ratio of dynamic stresses to the static pressure scale set by the system."

P4:L106–109, You state there are multiple conditions that should be satisfied, however, I only see the single condition $C_{\text{isost}} \ll 1$ mentioned? Did you also at some point verify that this is the case? I think this could be interesting for the reader to know.

Response: We reference Crameri et al. (2012) to emphasize that our chosen thickness and viscosity for the sticky air layer comply with these fundamental requirements, ensuring the accuracy of our results. There are other conditions such as $C_{\text{stokes}} \ll 1$, as thoroughly discussed in Crameri et al. (2012).We verified that when these conditions are met, the results align with the increased accuracy mentioned in their paper.

P6:Eq. (9) What is Γ (it has not been introduced), or do you mean Γ_{fs} ? Also since you are using the FEM I expect there to be a test function?

Response: Γ denotes the boundary surface, and fs is short for the free surface. The test function N_i can be find in details in Moresi et al. (2003).

P6:L153–154 ...the optimal values is 0.5 : This needs to be motivated and/or backed by a reference. And in what sense is it optimal?

Response: We added the relevant reference Kaus et al. (2010). The optimal value can be complex, as discussed in Andrés-Martínez et al. (2015), so we do not elaborate further. However, using a value of 0.5 yields results that are closer to the analytical solution compared to a value of 1, same as mentioned in Kaus et al. (2010).

P7:L175 When λ ¡= D, t $\approx$ t0 . Can you please clarify the interpretation of t0 and why it is interesting to note?

Response: t0 is the simplified version of t in mathematical terms when λ ¡= D. In many studies, t0 is primarily used as the relaxation time.

P8:L215 If the sides are not subject to periodic boundary conditions, then what are they subject to?

Response: The boundary condition is set to free slip on the left and right walls to ensure consistency with other experiments in our paper.

P9:L229 It seems like you are using the same symbol to denote both the effective strain-rate and the strain rate tensor? Response: corrected.

P9:Eq. (16) Equation for what? And could you please clarify in what sense it is nonlinear? Maybe I'm missing something, but you state that the viscosity is strain-rate dependent, at the same time from Table 1 it says that n = 1. To me this seems contradictory. Could please clarify what type of rheology you are using, i.e., is it linear (n = 1) or nonlinear (n ¿ 1) and in the latter case, what is the value of n?

Response: For temperature-dependent rheology. Here, we use n=1 and have made adjustments to the equation for clarity.

P9:L233–234 The effective strain-rate was already defined in line 229. Response: corrected and removed it.

P9:L244 If the age-law is standard, I expect there to be a reference. Response: added the reference, and remove the 'standard' description.

P10:L283 How do you define the Courant criterion?

Response: The Courant–Friedrichs–Lewy (CFL) condition is a necessary stability criterion for solving partial differential equations numerically, requiring that the simulation's time step dt is small enough for information to travel less than one grid cell dx per step. It is defined by the Courant number: $C = \frac{u\Delta t}{\Delta x} \leq 1$.

P11:L298 Relating to General comment 2: What is the meaning of strong in stabilities, to me a numerical instability is the unbounded growth of perturbations. I would assume they are always strong. I suggest simply noting the presence of numerical oscillations, or something along that line.

Response: In this context, the instability indicates that, in the ALE method, the delamination model does not deform like the delamination itself. Instead, it exhibits a strong counter-clockwise deformation pattern. We have added adjustments accordingly.

P17:Fig. 1 This is a useful figure, but real and virtual interfaces needs explanation. Response: Added more explanations. "The real interface refers to the actual free surface, while the virtual interface represents the surface obtained from numerical modelling."

P19:Fig. 3 What is the meaning of the outer path? Response: It denotes the next step in the loop.

Technical corrections:

P8:L201–202, Corrected to 'centred at'

P8:L215, removed 'initial stabilization'

P19:Fig. 3, Corrected 'verticial' to 'vertical'

P19:Fig. 5, Corrected dt to Δt

P25:Fig. 9, written as (b–d).

Figures The figures look mostly good, but the font size of e.g., ticks and labels could be increased. Secondly, from an accessibility perspective, red, green, and blue may not be such a color-blind friendly choice.

Response: Increased the font size, and changes colours to colour-blind friendly choice.

2 Response to RC2

Abstract, free surface numerical fluctuation ("drunken sailor instability") is also characteristic for true Lagrangian free surface treatment (Kaus et al., 2010).

Response: The marker fluctuations in the tractional sticky air method are primarily caused when using the PIC methods. This issue can be addressed by employing the level set method or the ALEIB method we propose here. All free surface simulation methods, including ALE and ALEIB, can experience the "drunken sailor" issue when using relatively large time steps, which needs FSSA to fix. More details are provided in the introduction section, and additional context is added in the abstract.

Line 45: Corrected to 10^{22}-10^{24}.

Line 105: Add the water-load case into the text. "The density of this layer is set close to zero for ensuring it exerts no pressure on the actual free surface (the interface between the air and rock), or it is set to 1000 kg/m^3 to approximate a water-loaded free surface (Gerya and Yuen, 2003)."

Line 125: Yes, we consider this; that's why we perform the resampling process using algebraic techniques. Generally, the new coordinates of the topography are (x2,y2), where x2 = x0+Vx*dt, y2 = Vy*dt. We then interpolate the new topography (x2,y2) onto the fixed x0 to obtain (x1,y1), where x0=x1. This strategy is similar to ALE; for more details, please refer to Thieulot (2011).

Line 235, Please specify how η_0 is defined (from the dislocation creep parameters? Empirically?).

Response: The reference viscosity is defined based on the Rayleigh number, approximately 1e6, as outlined in Crameri et al. (2017). The temperature-based Rayleigh number (Ra) can be expressed in terms of density (q), gravitational acceleration (g), temperature scale (ΔT), mantle depth (D), thermal diffusivity (κ), and reference viscosity (η_0)as: $Ra = \frac{\rho g \alpha \Delta T D^3}{\eta_0 \kappa}$. As we are using different values for temperature scale and depth from their paper, the reference viscosity used here is 10^{21} (theirs is 10^{23}).

Line 265: We added this improved strategy into the context, the volume of fluid (VOF) method as suggested. We also cite the related thesis by Timothy Stephen Gray (Gray, 2025).

Line 285,the poor performance of the Eulerian MIC scheme is somewhat surprising and could be related to the inaccuracy of the applied marker advection scheme (4th order Runge-Kutta, ontinuity-based advection scheme should typically work better) and/or to the way or recovering the free surface position from markers (Eulearian material type isosurface interpolated from markers should typically work better, see above). Is this specific to the used FEM approach? Does a standard staggered Eulearian grid MIC produce similarly poor results?

Response: This issue is specifically caused by the FEM approach used in Underworld 2. In Underworld, material properties are assigned by applying the values of the closest particles to the Gauss points (9 points in each element in 2D) and then integrating these values to the mesh nodes. In Underworld 3, we address this issue by using the KD-tree method to assign values based on a distance-weighted average from the particles. And we believe that a standard staggered Eulerian grid MIC provides better results, though it is not as effective as ALE or ALE-IB unless additional interface tracing improvement strategies are used.

Line 315. "Our ALE-IB scheme results are more comparable to the free-surface case in the Eulerian scheme reported in Crameri et al. (2017), where a shape-function averaging method was employed in their modelling on all the uppermost rock tracers and the lowermost air tracers. This approach, combined with their sticky air method, yields more accurate surface representations." Would be good to show their results for comparison on your plots. Will your Eulerian results become more accurate with the use of a similar shape-function averaging method as Crameri et al. (2017) used?

Response: Thank you for your suggestion. Unfortunately, we cannot include their results directly in our plots due to slight modifications in parameter settings compared to their models. However, we have included a citation to their Figure 4 for reference. We agree that adopting a shape-function averaging method, as used by Crameri et al. (2017), could enhance accuracy. However, implementing this in Underworld 2 presents certain challenges.

Line 330, Does Underwold have these higher order elements? If so, presenting some additional tests would be very useful.

Response: In Underworld 2, the primary and secondary element types are offered as pairs (e.g., Q1/dQ0, Q2/dPc1, or Q2/dQ1). In Underworld 3, we provide a wider range of element type options. We have included a test case of topography relaxation models with different element types in the appendix.

Line 350. Advantages over the Eulerian scheme are only demonstrated for the specific scheme explored in Underworld and explored in this study. It performs notably worse than the Eulerian staggered grid MIC scheme of Cramery et. (2017) (see comments to line 315 above).

Response: We acknowledge that the advantages of the ALE-IB scheme, as presented in our study, are demonstrated specifically within the context of the Underworld framework (FEM-PIC). While our results show improvements over certain Eulerian schemes in FEM, we recognize that the Eulerian staggered grid MIC scheme using STAGYY by Crameri et al. (2017) may outperform in specific scenarios. We appreciate the opportunity to clarify these distinctions and will ensure that the limitations and specific contexts of our findings are clearly articulated.

References

Andrés-Martínez, M., Morgan, J. P., Pérez-Gussinyé, M., and Rüpke, L. (2015). A new free-surface stabilization algorithm for geodynamical modelling: Theory and numerical tests. *Physics of the Earth and Planetary Interiors*, 246:41–51.

Crameri, F., Lithgow-Bertelloni, C., and Tackley, P. J. (2017). The dynamical control of subduction parameters on surface topography. *Geochemistry, Geophysics, Geosystems*, 18(4):1661–1687.

Crameri, F., Schmeling, H., Golabek, G., Duretz, T., Orendt, R., Buiter, S., May, D., Kaus, B., Gerya, T., and Tackley, P. (2012). A comparison of numerical surface topography calculations in geodynamic modelling: an evaluation of the 'sticky air'method. *Geophysical Journal International*, 189(1):38–54.

Gerya, T. V. and Yuen, D. A. (2003). Rayleigh–taylor instabilities from hydration and melting propel 'cold plumes' at subduction zones. *Earth and Planetary Science Letters*, 212(1-2):47–62.

Gray, T. S. (2025). *Free surface methods applied to global scale numerical geodynamic models*. PhD thesis, ETH Zurich.

Kaus, B. J., Mühlhaus, H., and May, D. A. (2010). A stabilization algorithm for geodynamic numerical simulations with a free surface. *Physics of the Earth and Planetary Interiors*, 181(1-2):12–20.

Lu, N., Moresi, L., Giordani, J., and Knight, B. (2026). A novel ale scheme with the internal boundary for coupling tectonic and surface processes in geodynamic models. *EGUsphere*, 2026:1–35.

Moresi, L., Dufour, F., and Mühlhaus, H.-B. (2003). A lagrangian integration point finite element method for large deformation modeling of viscoelastic geomaterials. *Journal of computational physics*, 184(2):476–497.

Moresi, L., Quenette, S., Lemiale, V., Meriaux, C., Appelbe, B., and Mühlhaus, H.-B. (2007). Computational approaches to studying non-linear dynamics of the crust and mantle. *Physics of the Earth and Planetary Interiors*, 163(1-4):69–82.

Thieulot, C. (2011). Fantom: Two-and three-dimensional numerical modelling of creeping flows for the solution of geological problems. *Physics of the Earth and Planetary Interiors*, 188(1-2):47–68.