Response to Referee report 1 on
Improving subduction interface implementation in dynamic numerical models

Dan Sandiford Louis Moresi

April 25, 2019

We thank the referee, Dr. Duretz, for the extensive and detailed comments on our manuscript. These have helped us to define the scope of the study more clearly, and also to the express some of the technical details more succinctly. In this reply, the comments of the referee (R1) appear in `typewriter font`, with our responses following directly beneath.

General comments

`This study aims at providing a new approach to model long-term subduction processes, particularly at the plate interface. The authors investigate the widely-used weak layer (WL) approach and identifies some of its limitations. In this light, they propose an alternative approach, termed embedded fault (EF), and show that this approach remedies some limitations of the WL approach. This work is interesting because many geodynamic modelers studying subduction processes are facing these issues. The outcome of this study (i.e. the EF approach) might help subduction modelers designing their models. It will surely help informing the community about the caveats of subduction modeling. That said, I was personally not convinced of using the proposed EF approach. The essence of this approach is to remap the geometry of the plate interface at each time step of a simulation. The idea is thus to overrule the geometries predicted by the numerical simulation in order to facilitate interplate decoupling. I would personally not encourage code users to interfere within simulations by using ad hoc rules. I would rather expect an alternative solution that would not require interfering with an ongoing numerical simulation. This would necessitate an augmented modeling framework, by accounting for thermo-mechanical feedbacks (e.g. Thielmann and Kaus 2012) hydrological (e.g. Dymkova and Gerya 2013) or chemical processes) or an advanced rheological model (e.g. Bellas et al. 2018). This is likely beyond the scope of the current study, however this is a fundamental issue. and these aspects are so far not really discussed. Prior to publication, I would hence recommend the authors to enrich their discussion - potentially around the above-mentioned points. I found the discussion very short and mostly restricted to the difference between WL/EF approaches. I would also encourage the authors to make sure that the figures are cited in increasing order. Youll find below a list of comments and suggestions.`

The summary of R1 acknowledges that our study may help subduction modelers designing their models. We are encouraged by this reflection, as this was certainly one of our major intentions. The primary criticism of R1 here is somewhat philosophical - we would argue - and involves an alternative view on the best way forward for *improving subduction interface implementation in dynamic numerical models*. Referring to our proposed Embedded Fault strategy, R1 proposes

‘an alternative solution that would not require interfering with an ongoing numerical simulation. This would necessitate an augmented modeling framework, by accounting for thermo-mechanical feed-backs, hydrological or chemical processes) or an advanced rheological model...’

R1 advocates the development of a more self-consistent approach to modelling the subduction interface, in which a range of feed-back mechanisms, including hydrological processes are implemented. In response, we simply suggest that there are valid reasons for pursuing both strategies. The suggestion of R1 likely involves significant increase in the complexity of models. Moreover, it is not clear that such as strategy is presently capable of allowing long-term, stable, buoyancy-drive subduction to operate. In cases where the aim of the subduction model is to understand, for instance, the long-term dynamics of the slab, such complexity may be undesirable. In other words, where a deliberately simplified representation of the subduction interface is desirable.

We also point out that the WL approach generally represents an ad hoc strategy, rather than an attempt to directly model the complex (multi) physics of the subduction interface zone. Given the ad hoc nature of WL implementations, it seems sensible to at least develop a more thorough understanding of the behavior and characteristics of that approach. This is precisely what our study intends to do.

The points we have briefly covered here have been integrated into the revised manuscript, particularly in Section 3.

Specific comments

p.1 l. 8 - What is ‘fully dynamic’?

This term has been used in the literature to refer to a numerical subduction model driven only by internal buoyancy forces, rather than being additionally forced by velocity boundary conditions. In a number of places we have simplified to this phrase to ‘dynamic models’. The term fully dynamic is clarified in Section 4.1.

What are the requirements for a model to be ‘fully dynamic’?

See previous comment.

p.2 l.5 - ‘geodyanamics’

Typo fixed.

p.3 l. 15 - What is ‘full thermal modelling’? How can a thermal model be ‘full’?

The word ‘full’ was unintended here, now corrected.

p.5 l.9 - ‘an solution’

Typo fixed

p.5 l.11 & p.5 l.25 - brackets around citation,

Syntax fixed

p.5 l.26 - a more important issue is when the interface locally thins out and becomes unresolved

We agree that this is an important issue. We discuss this in Section 6.

p.5 l.31 - ?

Syntax fixed.

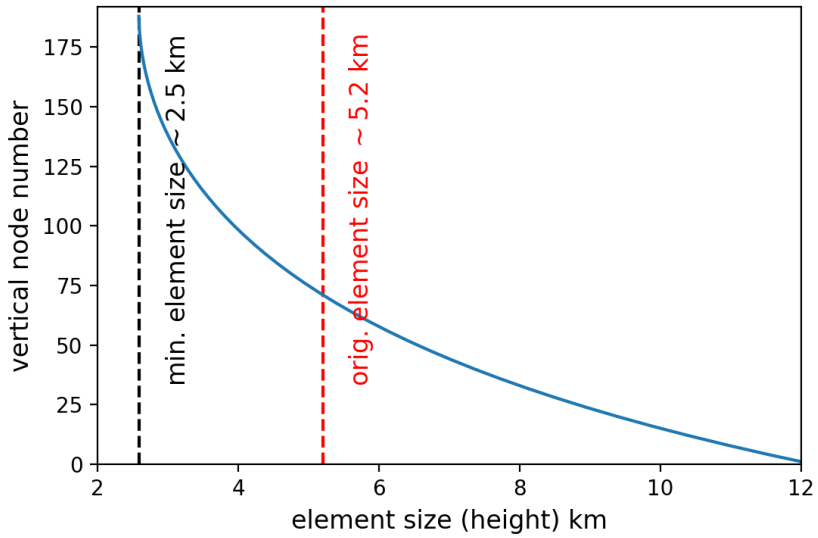

Figure 1: Vertical coordinate shows the vertical node number, for a 192 node mesh, with 192 being the surface node. Horizontal coordinate shows the element size. Blue line show the vertical element size in our highest resolution model.

`figure 1 caption: The layer representing the subduction interface appears to have viscosity variations (top-left) while it is described as having a constant linear viscosity.`

The rheology associated with the weak layer is merged with the background (plate/mantle) rheology at a significant distance (800 km) away from the trench on the subducting plate. The main reason for doing this is so that the weak layer does not extend all the way to the spreading ridge. This has been clarified in section 4.1. We think it is still appropriate to refer to the subduction interface as having a constant viscosity.

`Is figure 1 relevant at all, I dont see it called in the text?`

Figure 1 has been significantly changed and is now referred to in the main text. We think that it is relevant to provide context for the type of simulations we are discussing, as well as to clarify some of the SZ terminology that is used throughout the paper.

`p.6 l.15 - psuedo-brittle`

Typo fixed.

`p.6 l.24- what function you use to refine the mesh?`

Given a initial mesh with N uniformly spaced nodes (y_i), we shift the nodes to new locations (y'_i) using an normalised exponential function:

$$y'_i = (y_i - 1)e^{\lambda(\bar{Y}_i^2 - 1)} + 1$$

Where $\bar{Y}_i$ is the normalised position

$$\bar{Y}_i = \frac{-y_i - y_N}{y_N - y_0}$$

The constant $\lambda = 0.7$. Fig. 1 shows the element spacing based on this function.

p.6 l.24-what is the horizontal resolution?

All models have an aspect ratio of 5 which is reflected in the mesh resolution. So the highest-resolution models (192 elements in vertical dimension) have 960 elements in the horizontal dimension.

p.6 l.30 - on all side

Clarified.

p.6 l.33-34 - not clear whether left and right hand side wall are treated similarly.

The left and right hand side wall have the same boundary condition, this has been clarified.

figure 2 caption: Since you mention normal velocity to be zero at boundaries, you may also add that the tangential shear stress is zero.

We have added this information, and fixed some typos in the Figure 2 caption.

p.7 l.3 - 'Past studies...' I would add references or delete this statement.

Statement deleted.

p.7 l.8 - 'Throughout this study...' this sentence reads weird or incomplete.

This section has been rewritten.

p.8 l.1 - 'In this chapter' - is this part of a thesis or report?

Fixed this mistake.

p.8 l.4 - Fig. 8 is called right after Fig. 3

Removed this figure reference.

From Section 4.3 and Figure 3 caption, it is not entirely clear what the principle of EL is. Do you mean that: (1) you pre-define a channel geometry that will remains constant during model evolution (2) you remap, at each timestep, material types based on their relative position of the particles with regard to the channel? If yes, please state it clearly.

We mean (2). Section 4.3 and Figure 3 caption have been largely re-written to make this clearer.

p.8 l.7 'In a number of previous studies...' Do you mean variations in space or time? No sure whether the referencing is sound.

Primarily we mean spatially. We cannot say with certainly whether temporal variations are present in the studies we refer to, as this level of detail is not provided.

p. 10 l. 20 'effected'

Typo fixed.

p. 10 l. 21 'these feature'

Typo fixed.

p. 10 from l. 25 on. Better write $W_{\min}$ instead of W_min. Same for 'max' and 'init'

Fixed.

p. 10 l.30 please add a scale on this figure

This figure has been updated with a scale.

p. 10 l.30 ‘physically inconsistent’ do you mean geologically irrelevant? Is this so irrelevant by the way?

We mean, in a qualitative way, geologically irrelevant. The morphology shown in Figure 7. (to which p. 10 l.30 alludes) is dissimilar to current subduction zones on earth. One could make the argument that the extremely long forearc region ($\sim$ 500 km) is also seen in some flat slabs settings (although arc volcanoes may not actually be present in those regions). However flat slabs usually have a subduction hinge with relatively normal curvature, whereas at greater distance from the trench the slab dip reduces and the curvature changes sign. In contrast, the model shown in Fig 7b, the length of the forearc is due to the extremely low curvature, but there is no slab dip reduction (flattening).

p 10. l.34 - how would a free surface affect this behaviour?

While a free surface is likely to change some aspects of subduction dynamics, the variation in interface thickness - seen in the WL approach - would still be expected as the argument relating to variation in volume flux remains the same. We cannot guarantee that a model with a free surface would behave identically when the interface is not allowed to thicken (e.g. as in Fig 7 b).

figure 6 caption: what are red and black dashed lines? Which model is depicted here? The model presented in Fig. 7a or 7b?

Figure caption has been updated. The model depicted here is the same as in Fig.1a and Fig. 1b. (in the revised manuscript). This has been noted in the caption as well.

p.12 l.10 - reduce the amount the transient adjustment

We’re unsure what the comment is.

p.12 . l.16 - ‘One advantage
of the EF approach is that it offers improved precision in determining the thickness of the subduction interface.’ This is confusing, I thought the EL approach was aiming at imposing this thickness. How can it help to determine a thickness when it already imposes it?

Yes, by imposing a maximum and minimum thickness, we gain precision in the knowledge of its thickness, don’t we?

p.12 l.17 - ‘Such precision will be important for studying highly pressure- and temperature-sensiti processes, such as metamorphism and melting near the slab top.’ Well, sure. This new parameterization will pre-define everything related to it.

See previous comment.

figure 7: very difficult to appreciate the dimensions, a scale is missing. What is the grey line? I think the notion of ‘physically-consistent subduction morphology’ does not make sense.

Figure 7 has been substantially updated along with the caption. We have also incorporated R1’s suggestion on ‘physically-consistent subduction morphology’, both here and in the main text.

figure 8: How is this MDD monitored? Do you measure the stress differences across the plate interface?

The MDD is monitored by evaluating the vertical component of the velocity at a distance 5 km above the top of the slab (in a slab-normal sense). The vertical velocity clearly reveals the location of the onset of coupling.

figure 9: do you use the same scale in x and y? figure 9 caption: ‘effected’

Figure 9 has been updated and typo fixed.

figure 10: at what time do these snapshots correspond? Do you use any particle reseeding?

Snapshots are at 12.5 Myr, after the WL interface thickness has reached its quasi-equilibrium profile. This is now recorded in the figure caption. Particle reseeding is used to maintain regular particle numbers in the elements.

figure 10 - 11: Given the fact that the mesh is not distorted, adding element edges would help the reader to realise how well the interface is resolved.

True, but as the mesh is quite high resolution we feel it would also detract from the clarity of the image. We have tried to convey this information (i.e. the interface resolution) where necessary.

p.15 l.9 - you mentioned you ran simulations with '72, 96, 128, 160 elements'. You can also include 192 elements as I understand.

We did, and this has been corrected.

p. 16 l.5 - 'better' sounds very qualitative. You mean that low resolution models using a WL fails at capturing plate decoupling.

Agreed. We have expressed this differently.

figure 13: I dont understand, why do repeated models at the resolution of 192 produce any error? If the models were repeated you should obtain the exact same results, dont you?

The only difference in the models is the randomness in initial particle locations, as well as the repopulation of those particles. We would note that the error, or drift, is still quite small. We are not aware of a similar experiment previously being published, so we cannot comment on whether this value is unexpectedly high.

p. 20 l.10 - a mistake here, a mean stress is not lithostatic. It can be split into dynamic and lithostatic components.

This mistake has been corrected.

p. 22 - Why using a lithostatic pressure field in the viscosity expression. Are obtained numerical solution of the actual pressure field not accurate enough to be used in the rheology?

This is not indicative of poor results in the pressure calculation, but rather the assumption that the dynamic contribution of the pressure is expected to have to little effect on the viscous flow law compared to the lithostatic pressure contribution. In contrast, the dynamic pressure component is used in the determining the yield surface (Drucker-Prager criteria). The reviewer is correct that it would be more consistent to use the full pressure, however we doubt that the results of this study would be changed.

References

Thielmann, Marcel and Boris JP Kaus (2012). "Shear heating induced lithospheric-scale localization: Does it result in subduction?" In: *Earth and Planetary Science Letters* 359, pp. 1–13.

Dymkova, D and T Gerya (2013). "Porous fluid flow enables oceanic subduction initiation on Earth". In: *Geophysical Research Letters* 40.21, pp. 5671–5676.

Bellas, Ashley et al. (2018). "Dynamic weakening with grain-damage and implications for slab detachment". In: *Physics of the Earth and Planetary Interiors* 285, pp. 76–90.

Response to Referee report 2 on
Improving subduction interface implementation in dynamic numerical models

Dan Sandiford Louis Moresi

April 26, 2019

We thank the referee for their critical reading of the manuscript, which raised a number of important issues and has helped us to clarify the central arguments in the study. In this reply, the comments of the referee (R2) appear in `typewriter font`, with our responses following directly beneath. Note that in our updated manuscript the figure order has changed slightly. Unless otherwise stated, figure numbers here refer to the first submitted manuscript version.

General comments

`This paper explores different parameterizations of a weak layer that is used to model one-sided subduction. It particularly explores the controls on the width of this weak layer and finds that, within the context of the models presented, there is a preferential width. In general I think this is a nice, albeit technical, contribution to the literature describing fully dynamical modeling of subduction zones. I have a few somewhat broader comments and a fair bit of more detailed comments.`

A continuously-entrained weak layer (WL) is a common strategy used in numerical models to provide decoupling between the slab and upper plate. In the models presented we show that significant thickness variations develop when a uniform thickness WL is initially prescribed. We argue that changes in the volumetric flux in the weak layer lead to these spatial and temporal variations in thickness. The comment that a preferential width is found 'in the context of the models presented', seems to imply that our results are model specific and not generalizable. In later parts of the review, R2 suggests that mesh resolution might be a contributing factor to the observed thickness variations and that such effects would be mitigated with increasing resolution.

We feel that these suggestions are made without any real justification, and do not critically engage with the argument we make. Our explanation for the evolution of the interface thickness invokes one main assumption, which is a first-order characterization of the kinematics of the WL, namely that uniform flow of material on the incoming subducting plate transitions to simple shear (Couette flow) within the decoupling region. We show that the interface evolves to an equilibrium thickness profile close to that predicted under this assumption. To help further clarify these processes, we introduced a simple boundary-driven example (e.g. fig. 5) which shows analogous thickness changes in a weak horizon due to a changing boundary condition and concomitant gradients in downstream volume flux. Hence, we have already gone to some effort to show that our analysis reveals an intrinsic behaviour of this kind of flow.

Of course, depending on other aspects of the model setup (e.g. the subduction interface rheology, use of adaptive meshing), width variations of the WL will effect the evolution of the simulation differently. However, the width variations themselves are unlikely to be specific to our models,

nor are they an artifact of low resolution. We feel that we may not have communicated these arguments as clearly as we could have, and have therefore re-written substantial parts of Section 5.1.

Specific comments

`Section 2 contains a fair bit of information that was already mentioned in the introduction or other superfluous. A bit of careful editing can shorten this section and make it clearer.`

Based on comments from all reviewers, we made minor adjustments to this section, with some simplifications.

`The citations in the text dont seem to follow any chronological or alphabetical order (see example in P4.l9-l11).`

This has been remedied.

`The weak fault seems rather wide (10 km as mentioned on page 7). If you have a 2 km grid resolution then that it is probably a good idea to smear it out over such a distance but what are your thoughts about FEM grid refinement around this fault to be able to bring it to more realistic thicknesses?`

We chose a WL initial thickness that was representative of recent long-term numerical subduction models, e.g. (e.g. Čížková, Hunen, and Berg 2007; Garel et al. 2014; Agrusta, Goes, and Hunen 2017). Again, as our paper is methodological, we are content with using a model setup that is appropriate for the problem. Of course, in the highest-resolution models we showed (e.g. element width of $\sim 2.5.$ km) we could resolve a thinner interface. But we doubt that doing so would alter any of the conclusions of the study.

Based on our analysis of the resolution convergence, we would suggest that the ratio of prescribed WL thickness to element size should be greater than three. This is not to say that a model with less resolution won't produce physically consistent subduction, however it may not show satisfactory resolution convergence. Testing resolution convergence is obviously the best way to ascertain this for a specific setup. We argue that in the WL approach, there is a component of 'wasted' resolution that is required only to properly resolve the transient phase of interface adjustment. This is why, we argue, the error accumulation tends to decrease significantly at around the time the interface thickness reaches quasi-equilibrium (see Figure 13.).

`I wonder if the improvements you find when choosing the initial and maximum width of the weak zone to be 20 km is just a matter of the numerical resolution (mesh and particles) that you use. Maybe you need at least 10 elements in this zone for things to stabilize. You mention that a lot of small wavelength variations disappear compared to using 10 km. This might be a symptom of discretization issues...`

The improvements we mention (using the EF approach) are specifically improvements relative to a standard WL setup. We assessed both implementations at several resolutions. We showed that the EF approach has better resolution convergence. This means that as we increase the resolution, the error (relative to a high-resolution reference model) is always smaller in the EF approach than in an equivalent WL approach. We expect that further adjustments to the EF approach (or similar strategies) could improve this further. R2 is correct to point out that the EF method always starts out with a higher effective resolution in the decoupling region - hence the comparison between the EF and WL approaches does have a limitation.

The EF approach has two separate features, one is the imposition of a variable initial thickness, the other involves controlling the maximum and minimum thickness. Both of these are targeted

at different behaviours we observed in the WL model. The first tries to limit the amount of transient adjustment of the interface thickness, the second helps to dampen the short-wavelength instabilities at the boundary of the subduction interface and the upper plate. We could, of course, investigate these aspects individually. However, our main goal here was to design and communicate an approach that achieves better outcomes (e.g. better resolution convergence) for a given computational expense. We think that the EF approach achieves this goal.

`It would be very helpful if you could repeat your exercise with higher resolution. I suspect your maximum width will go down with increasing resolution. In 5.3 you do a divergence test but I am not sure that this is very meaningful given that you show that coarser resolution meshes lead to worse results. It is far more exciting to see what happens when you start converging.`

As we discuss in our reply to the preceding comment, there are some limitations in the convergence (divergence?) test, and therefore some ambiguity in how these tests are interpreted. We hope that we have already answered these criticisms satisfactorily. The second point made here by R2 is that the pattern of width variations in the WL approach might be resolution dependent, and that such processes would be mitigated with increasing resolution.

Our reply here essentially follows from our answer to the comment in paragraph 1. The width variation are not likely to be a dependent on resolution or 'context' (see paragraph 1), because the width variations are an intrinsic outcome of the fluid dynamics (see reply to paragraph 1).

The referee has provided no physical argument to back up the expectation that these processes should be resolution dependent. Hence, we are skeptical that it would be helpful to reproduce our results with higher-resolution models.

We accept that we may not have communicated these arguments as clearly as we could have, and have therefore re-written substantial parts of Section 5.1.

`Figure 11. I am not certain it is useful to show under-resolved results.`

In this case we feel it is useful, because the style in which the models degenerate actually reveals something about the way that the interface thickness is dependent on the kinematics of the flow. In particular, when the interface is under-resolved and does not adequately decouple the plates, it is the induced partial-coupling that drives further thinning of the interface. This feedback cycle is a limitation of the WL method, and may be very relevant when additional processes are included - such as subduction of buoyant crust / continents.

`P1 footnote 1. why footnote?`

We have removed the footnote.

`might be good to add a few original references as to why the slab-mantle wedge coupling appears to start at 80 km (that seems to be the case in most SZs; please just clarify why we think this is the case)`

We have cited the paper of Wada and Wang 2009. In the scope of this study we feel this sufficient.

`P2.l5. typo`

Typo fixed.

`P2.l25. typo`

Citation fixed.

`P2.l25-27. Full sentence here. You can probably rephrase this a bit better as in that sediments may be important amidst various other issues controlling plate velocities.`

We have reorganised this paragraph.

`P2.l33. 10s -> tens`

Changed.

P3.l12-14. I am confused about this full sentence here. The seismogenic zone is generally not characterized by ending in the mantle (maybe between slab crust and mantle, but not slab mantle and mantle).

We have simplified this section.

P3.l14. Average stresses: where?

The value we quote (from Lamb 2006) is based on an average stress taken along the entire subduction interface.

P3. line starting at l22. This seems to take a single point of an antigorite-out boundary as some extreme. Its a bit more variable/complex than that. You can probably delete this sentence.

We have reconfigured this paragraph substantially based on the comment.

P4. Figure 1. last line 'representative' of what?

Figure 1 has been changed and captions and labels redone.

P4.l14 missing article

Changed.

P4.l5 missing plural

This paragraph has been condensed.

P4.l9 typo

Changed.

P5.l11 typos

Fixed.

P5.l25. What is 'over-resolving?

We mean providing resolution larger than is necessary to resolve an interface of a given thickness. This would ensure that the interface remains adequately resolved when the interface thins (as occurs in the deeper part of interface under WL approach - e.g. Fig. 4).

P5.l30 just benchmarking

Changed

P5.l31. '1 km, 20K' Not sure what this means. The benchmark study referenced here showed that finite element models agreed on temperature predictions along the slab surface to within about 1K for 1 km resolution. There were some finite difference models that had larger differences.

We have shortened the paragraph and excised this sentence.

And here and below. It is just 20K not 20°K.

Fixed.

P7 three sentences starting at l20. Dont you say essentially the same thing here? You might be able to condense this into a single sentence.

We have simplified this paragraph.

P8.l1 I hope a cosine taper is easy to implement in many subduction models. Is it not?

Past studies have used different approaches to effectively inhibit the subduction interface from extending to arbitrary depths. In the context of the Lagrangian particle method, we could simply (arbitrarily) replace the weak layer material with background material at a given depth. We choose a rheological transition, based on a cosine taper, to allow a somewhat graduated change.

`P9.l12ff. Maybe use subscript ‘conv?`

We have simplified the mathematical symbology in this section.

`P9.bottom. MDD reference is repetitive`

Changed.

`P10.l6 You are citing Figure 8 way out of order`

Yes. We have omitted this reference.

`P10.l21 typo`

Fixed.

References

Lamb, Simon (2006). “Shear stresses on megathrusts: Implications for mountain building behind subduction zones”. In: *Journal of Geophysical Research: Solid Earth* 111.B7.

Čížková, Hana, Jeroen van Hunen, and Arie van den Berg (2007). “Stress distribution within subducting slabs and their deformation in the transition zone”. In: *Physics of the Earth and Planetary Interiors* 161.3-4, pp. 202–214.

Wada, Ikuko and Kelin Wang (2009). “Common depth of slab-mantle decoupling: Reconciling diversity and uniformity of subduction zones”. In: *Geochemistry, Geophysics, Geosystems* 10.10.

Garel, Fanny et al. (2014). “Interaction of subducted slabs with the mantle transition-zone: A regime diagram from 2-D thermo-mechanical models with a mobile trench and an overriding plate”. In: *Geochemistry, Geophysics, Geosystems* 15.5, pp. 1739–1765.

Agrusta, Roberto, Saskia Goes, and Jeroen van Hunen (2017). “Subducting-slab transition-zone interaction: Stagnation, penetration and mode switches”. In: *Earth and Planetary Science Letters* 464, pp. 10–23.

Response to Referee report 3 on
Improving subduction interface implementation in dynamic numerical models

Dan Sandiford Louis Moresi

April 25, 2019

We thank the referre, Dr. Čížková, for providing a number of helpful comments on our manuscript. In this reply, the comments of the referee (R3) appear in `typewriter font`, with our responses following directly beneath. Note that in our updated manuscript the figure order has changed slightly. Unless otherwise stated, figure numbers here refer to the first submitted manuscript version.

General comments

`Manuscript presents a detailed analysis of the properties of entrained weak layer that is often used in numerical models of subduction to decouple the plates. The authors demonstrate that this weak layer is changing its width both spatially and with time. After entering the subduction channel at the trench it increases width and then thins again as the subducting plate is coupled with the mantle flow below the overriding plate. This thinning, that may temporarily result in a decoupling weak layer actually yet thinner than the prescribed initial thickness, may result in locking of the subduction in case of lower resolution and certainly affects the subduction evolution. The authors explain this phenomenon using an analogy with a boundary-driven Stokes flow in a two layer material with boundary condition changing from free-slip to no-slip and back.`

`They then introduce an improvement to the standard weak layer approach (called here embedded fault) where the thickness of the weak layer is controlled and modified during the subduction evolution. This approach prevents transient thinning of the lower part of the decoupling layer and potential coupling of the subducting and overriding plates incase of lower resolution.`

`I find this paper very interesting. The weak layer approach is often used in subduction modelling while the numerical aspects of the implementation of this decoupling are seldom discussed or even mentioned. I therefore very much appreciate this systematic evaluation of the problem and suggested solution using ad hoc control of the layer thickness. The crutial point of the paper is made in fig. 12 where the authors illustrate that in their embedded fault approach low resolution has much smaller effect than in standard weak layer approach. This paper thus provides the reader with a recipe how to tackle the problem of decoupling the plates with an entrained weak layer one may either use the suggested embedded fault approach, or use resolution high enough to resolve the thinned interface in the transient stage at the beginning of subduction. Alternatively, as just briefly mentioned in the discussion, this problem may be suppressed by nonlinear rheology, but that of course brings other complexities into the play`

The manuscript is nicely and clearly written and the topic fits the scope of Solid Earth,therefore I recommend it for publication. I only have couple of suggestions for mostly minor corrections.

As the general comments above mainly summarise our paper, and are largely supportive of the scope and conclusions, we thank the reviewer and simply address the specific points listed below.

Specific comments

You may perhaps explicitly mention that the weak layer has constant viscosity. The reader has to dig up this information by combining the sentence stating that except of subduction interface the rest of the domain is deforming according to composite rheology (at the end of paragraph 20 page 22 in Appendix) and the information in Table A2 (unless I overlooked this information somewhere else).

In the updated manuscript we have made additional reference to the subduction interface viscosity in Section 3 paragraph 2, as well as Section 4.2.

I dont see the logic in the order of figures they are sometimes ordered and referenced rather randomly. Figure 1 is not referenced in the text.

Figure 1 has been substantially changed and is now referenced in the text. Other figure references that were out of order have been rectified, while the figure order has been changed slightly.

Page 2, par. 25, line 4: remove of

Typo fixed.

Page 5, par. 5, line 5: an solution -> a solution

Typo fixed.

Page 5, par. 10, line 3: though -> thought

Typo fixed.

Page 6, par. 5, line 3: a initial -> an initial

Typo fixed.

Page 6, par. 20, line 5: a element -> an element

Typo fixed.

Caption Figure 11: Fig. a -> Fig. 12a

Fixed.

Page 22, par. 15, line 3: We -> we, mechanism -> mechanisms

Typo fixed.

Table A2: domain depth 100 -> 1000

Fixed.

Improving subduction interface implementation in dynamic numerical models

Dan Sandiford[1,2] and Louis Moresi[1]

[1]School of Earth Sciences, University of Melbourne, VIC, 3010, Australia
[2]Institute of Marine and Antarctic Studies, University of Tasmania, TAS, 7004, Australia

Correspondence: Dan Sandiford (dan.sandiford@utas.edu.au)

Abstract. ~~This study focuses on methodological issues related to dynamic subduction zone modelling. Numerical models often employ~~ Numerical subduction models often implement an entrained weak layer (WL~~approach~~) to facilitate decoupling ~~between the subducting and overriding plates. In such a setup, the kinematics of the flow lead to width variations in the subduction interface. When a uniform-width interface is prescribed, a transient evolution of the interface thickness occurs, during which the volumetric flux along the interface profile establishes equilibrium. Width variations can exceed 4× during this stage~~of the slab and upper plate. This approach is attractive in its simplicity, and can provide stable, asymmetric subduction systems that persist for many tens of Myr. In this study we undertake a methodological analysis of the WL approach, and use these insights to guide improvements to the implementation. The issue that primarily motivates the study is the emergence of significant spatial and temporal thickness variations within the WL. We show that these variations are mainly the response to volumetric flux gradients, caused by the change in boundary conditions as the WL material enters and exits the zone of decoupling. The time taken to reach a quasi-equilibrium thickness profile will depend on the total plate convergence, and is around 7 Myr for the models presented here. During the transient stage, width variations along the WL can exceed 4×, which may impact the effective strength of the interface, ~~both~~ through physical effects if the rheology is linear, ~~and numerical effects if the fault becomes poorly resolved. This transient process~~ or simply if the interface becomes inadequately numerically resolved. The transient stage also induces strong sensitivity to model resolution~~, and may present a significant challenge to reproducibility. Developing more robust ways to model the subduction interface will enable fully dynamic models to address sensitive subduction-zone processes, such as metamorphism near the slab top. In this study we discuss a simple strategy aimed at improving the standard WL approach~~. By prescribing a variable thickness ~~weak layer~~ WL at the outset of the model, and by controlling the limits of the layer thickness during the model evolution, we find improved stability and resolution convergence of the models.

1 Introduction

~~Subduction zone modelling, both through numerical and analogue experiments, has provided important insights into the dynamic and chemical processes associated with solid earth convection. Yet designing numerical subduction experiments~~

that are simple enough to allow a tractable study of these phenomena, yet rich enough to be quantitatively constrained by the observations remains challenging (Buiter and Ellis, 2016). The subduction interface is a critical and sensitive component in such models. It must simultaneously provide strong strain localisation, low stress, lateral translation, and long term stability (Gurnis and Hager, 1988; Arcay, 2012).

The process of stable asymmetric subduction requires that the down-going plate is substantially decoupled [1] from the overriding plate along the subduction interface (Gerya et al., 2008). In the majority of subduction zones, decoupling transitions rapidly to complete slab-mantle coupling at a depth of around 80 km (Wada and Wang, 2009). This means that the depth extent of decoupling is significantly greater than the limit of the locked megathrust zone ($<$ 50 km). The subduction interface zone is characterised by rheological and petrological complexity, low strength (relative to the slab), as well as the abundance of water and the critical role of fluid pressure (Bachmann et al., 2009; Bebout and Penniston-Dorland, 2016; Gao and Wang, 2014; Hardebeck, 2015 .

Determining the combination of processes that allow plate-bounding faults to develop is a longstanding problem in geodyanamics (Trompert and Hansen, 1998; Moresi and Solomatov, 1998; Tackley, 2000; Lenardic and Kaula, 1994; Bercovici, 2003; Bercovici and Ric geodynamics (Lenardic and Kaula, 1994; Trompert and Hansen, 1998; Moresi and Solomatov, 1998; Tackley, 2000; Bercovici, 2003; Ber . Even when the subduction interface is assumed *a priori*, the implementation within a continuum modelling framework is not trivial. Accordingly, a range of modelling approaches have been developed. Different approaches may reflect the style of model (e.g. instantaneous vs. long term), physics and boundary conditions the degree to which multiphyics are represented (e.g. hydrgeologic processes), the subduction driving forces (kinematic vs. dynamic), as well as numerical method (finite difference vs. finite element methods vs. boundary element methods). The most A common approach in long term, fully dynamic models is to implement the subduction interface as a continuously-entrained layer of weak material (WL). Fig. 1a shows an example of this approach, implemented using Underworld2. This kind of approach is relatively simple to implement, and can provide stable asymmetric, subduction regimes that persist for many tens of My, as shown in Fig. 1b. While the WL approach has been utilised in a broad range of studies, a detailed analysis of the technical implementation methodological analysis of this strategy is lacking (although valuable insights appear in Arcay (2017); Manea and Gurnis (2007); Arcay (2012); Čížková et al. (2002); Androvičová et al. (2013)). In this paper we focus on the dynamics and evolution of the subduction interface material under the WL approachThis study focuses on the behaviour of a constant viscosity WL implementation, within a 2d thermo-mechanical subduction setup. The results have relevance for model reproducibility and provide a basis for improving precisionwithin fully dynamic the precision, efficiency, and reproducibility of dynamic subduction models.

[1] In this study we use the term coupled in the long-term sense, specifically to the depths where the slab and mantle become continuously coupled ($\gtrapprox$ 80 km). While seismologists often refer to seismogenic megathrusts as coupled (or locked), from the long term perspective, the plates are decoupled at the megathrust.

2 The subduction interface

The subduction interface refers to the plate boundary fault at earth's convergent margins. Exhumed subduction interfaces are typified by melange zones, often 100's of meters in width, with coherent blocks embedded within a sedimentary and/or serpentinised matrix ~~(Kimura et al., 2012; Bebout and Penniston-Dorland, 2016; Lallemand, 1995; Vannucchi et al., 2008; Shreve and Cloos, 198~~ (Shreve and Cloos, 1986; Vannucchi et al., 2008; Kimura et al., 2012; Bebout and Penniston-Dorland, 2016). ~~Entrainment of sediment at subduction is highly variable (Shreve and Cloos, 1986; Huene and Scholl, 1991)~~The subduction interface zone is characterised by rheological and petrological complexity, low strength, as well as the abundance of water and the critical role of fluid pressure (Abers, 2005; Audet et al., 2009; Bachmann et al., 2009; Gao and Wang, 2014; Hardebeck, 2015; Duarte et al., 2015; Bebout and Pennist . Subducted sediments can reach depths of up to 80 km (Bayet et al., 2018), similar to the inferred depth at which the slab and mantle become coupled, and can be traced in the composition of arc magmas (Plank and Langmuir, 1993). The proportion of subducted sediment may lead to large variations in the mechanical properties in the deep subduction interface ~~Behr and Becker (2018). Whether these variations have an observable impact on subduction dynamics is debated, partly due to the complex set of controls on plate velocities (Duarte et al., 2015; Schellart and Rawlinson, 2013; Cloos and Shreve, 1988; Behr and Beck~~ (Behr and Becker, 2018). Entrainment of upper plate material in a process known as subduction erosion also occurs at a substantial number of ~~of~~ convergent margins (Huene and Scholl, 1991). Subduction interfaces therefore incorporate material from the subducting plate, the accretionary prism, and the upper plate (Vannucchi et al., 2008). ~~The deeper part of subduction interface where the slab is in contact with the serpentinised upper-plate mantle is thought to be controlled primarily by weak hydrous minerals which allow effective slab-mantle decoupling but also inhibits unstable seismic slip (Hirauchi and Katayama, 2013; Reyna . The thickness of subduction interfaces, particular in the deeper aseismic zone, is not well constrained. Estimates range from between 10's metres to a few kilometres (Vannucchi et al., 2012, 2008; Abers, 2005; Cloos and Shreve, 1988). Behr and Becker (2018) used structural relationships in exhumed subduction interfaces to infer the relative strengthof different lithologies. Metasediments and serpentinites which commonly form the matrix within the subduction melange appear to be weakest.~~ However, the degree to which variability in these influxes impacts long-term subduction dynamics is debated (Cloos and Shreve, 1988; Schellart and Rawlinson, 20 .

Fluid pressure plays a key role in subduction zones interfaces as it determines the effective frictional strength, controls the seismic/aseismic character of slip (Audet and Schwartz, 2013) and dictates the extent of hydrous mineral formation along the subduction interface (Reynard, 2013). In the shallow megathrust, pore fluid water and clay minerals are thought to exert a major control on the rheology (Vrolijk, 1990). ~~At greater depths, fluids are generated by devolatilization reactions (Bebout and Penniston-Dorland, 2016).~~ Seismic imaging demonstrates the presence of fluids along the plate boundary ~~(Audet et al., 2009)~~ (Audet et al., 2009). In some cases pore pressures are inferred to be near lithostatic values (Audet et al., 2009). ~~Fluid pressure plays a key role in subduction zones interfaces as it determines the effective frictional strength, controls the seismic/aseismic character of slip (Audet and Schwartz, 2013) and dictates the extent of hydrous mineral formation along the subduction interface (Reynard, 2013).~~

While many subduction zone megathrusts are capable of hosting great earthquakes, there is a spectrum of behaviour from locked to creeping. Ongoing debate surrounds the strength implications of this divergent behavior (Gao and Wang, 2014; Hardebeck, 2015; Hardebeck and Loveless, 2018). The limit of the seismogenic zone can range from 5 to 50 kilometres in depth ~~, and is usually found to reside in the forearc mantle, rather than crust (Wang, 2010; Wada and Wang, 2009)~~(Tichelaar and Ruff, 1993) . The topography of most forearcs is consistent with average shear stresses of 15 MPa over the long-term (Lamb, 2006). ~~Heat flow studies, combined with full thermal modelling of particular subduction zones are generally consistent with megathrust~~ Megathrust shear stresses in the range 1-100 MPa, with mean shear stresses also around 15 MPa~~(Peacock, 1996; Gao and Wang, 2014) .~~, are consistent with results from heat flow studies (Peacock, 1996; Gao and Wang, 2014).

At greater depths, fluids are generated by devolatilization reactions (Bebout and Penniston-Dorland, 2016). The deeper part of subduction interface where the slab is in contact with the serpentinised upper plate mantle is thought to be controlled primarily by weak hydrous minerals which allow effective slab-mantle decoupling but also inhibits unstable seismic slip (Hirauchi and Katayama, 2013; Reynard, 2013). It remains unclear whether the rheology of the deep subduction interface is dominated by viscous or (stable) plastic behavior ~~(Li and Ghosh, 2017; Proctor and Hirth, 2016).~~

~~Phyllosilicates such as serpentine, chlorite and talc are inferred to have a strong influence on subduction zone dynamics because of their high water content and low strength at shallow and intermediate depths (Hilairet and Reynard, 2009; Schmidt and Poli, 1998; Hirau . Serpentine minerals form by hydration of olivine at the boundary of the slab and mantle wedge, where fluids from the subducting slab accumulate. Their stability range extends to 550°C and 200 km depth (Reynard, 2013). Chlorite forms primarily through hydration of orthopyroxene, another abundant mineral in dry peridotite. Talc typically forms by metasomatic alteration of the wedge through addition of silica-rich hydrous fluids derived from the slab. At the subduction interface the most stable serpentine polymorph, antigorite, is limited to depths between 70 to 120 km, depending on the thermal regime. This serpentinised (plus chlorite, talc) layer at the top of the slab is thought to be primarily responsible for accommodating decoupling to ~ 80 km (Hilairet and Reynard, 2009; Ulmer et al., 1995)~~(Proctor and Hirth, 2016; Li and Ghosh, 2017). The thickness of subduction interfaces, particular in the deeper aseismic zone, is not well constrained. Estimates range from between tens of metres to a few kilometres (Abers, 2005; Cloos and Shreve, 1988; Vannucchi et al., 2008, 2012). Behr and Becker (2018) used structural relationships in exhumed subduction interfaces to infer the relative strength of different lithologies. Metasediments and serpentinites which commonly form the matrix within the subduction melange appear to be the weakest components.

3 Past modelling approaches

In efforts to study subduction zone dynamics, a range of modelling approaches have been developed. The subduction interface is a necessary model component whenever both subducting and upper plates are included. One approach is to incorporate the interface directly as a kinematic constraint into the simulation, i.e. by specifying continuous normal and discontinuous tangential velocities in the model solution surrounding the fault (Billen et al., 2003; Christensen, 1996). This idea was developed into a ~~fully~~ dynamic framework with finite element models that incorporate internal stress boundaries; the result is zero

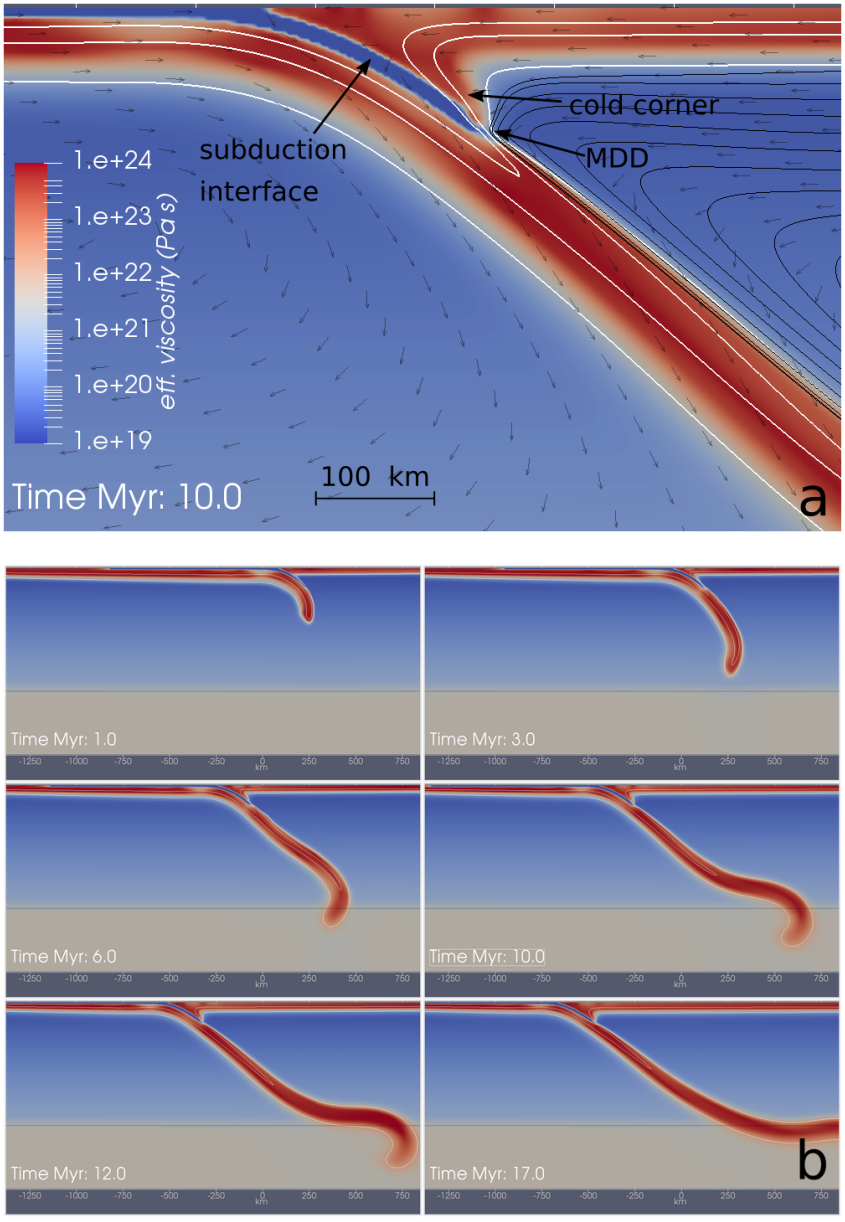

Figure 1. Example of ~~numerical~~ dynamic subduction simulation. ~~Closeup of the slab in the upper mantle from a fully dynamic subduction model.~~ Colormap shows the effective viscosity on a logarithmic scale, and reveals the weak finite-width subduction interface which enables decoupling of the slab and the ~~upper-plate. In this model the subduction interface has a simple constant linear viscosity~~upper plate. The implementation of the interface follows ~~the~~ a modified WL approach which we refer to as an embedded fault (EF)~~strategy developed in this study~~. The subduction interface extends to 100 km depth, beyond which the slab and mantle become coupled ~~. Dynamically, this is known as the Maximum Depth of Decoupling~~ (MDD). Above the MDD part of the mantle wedge tends to become stagnant due to the strong temperature dependence of the rheology. This is often referred to as a ~~'~~cold corner~~'~~. a) closeup of slab in the upper mantle. White contour lines show 400, ~~depicted here with~~ 700 and 1200°C isotherms. Black lines are streamlines of the velocity in a ~~green polygon~~fixed upper plate reference frame. ~~The region labelled melt transport~~is ~~representative~~Arrows show model velocity field. b) Time evolution of the same model, showing slab rollback and interaction with the transition zone.

width faults within the continuum mechanical representation of the lithosphere (Zhong and Gurnis, 1992, 1995; Aagaard et al., 2008). Because moving mesh nodes are needed to capture proper fault advection, accurate tracking of large scale deformation is challenging. A more common approach is to apply finite regions of weak constitutive behavior within a static mesh. The velocity field naturally develops strain localization around the weak zone, although the 'faults' are usually very much broader than real plate boundaries. This approach was first applied as a spatially-fixed low viscosity zone that could decouple the plates but would not allow trench motion (e.g. Gurnis and Hager, 1988). This type of implementation was developed so that narrow low-strength weak zone 'stencils' could also be advected to allow trench motion (Kincaid and Sacks, 1997; Billen and Hirth, 2007).

Over the past decade the use of an entrained weak layer (WL) has become an increasingly common strategy ~~(e.g. Capitanio et al., 2010; Garc~~ (e.g. Babeyko and Sobolev, 2008; Capitanio et al., 2010; Magni et al., 2012; Chertova et al., 2012; Čížková and Bina, 2013; Garel et al., 2 . Rather than using a fixed weak zone, the subduction interface is typically implemented by imposing a material layer at the top of the subducting plate that is advected with the flow and continuously entrained into the decoupling region. The WL approach enables a mobile trench while also helping to ~~facilitate plate bendingas it enters the subduction zone (even~~ accommodate plate bending, which is important in the absence of ~~free surface)~~a free surface. Additionally, it isolates the extremely weak parts of the system (the plate boundary) from the plates, enabling long-term asymmetric subduction to occur. ~~The WL implementation will often include decisions about where to create and remove the weak material after it has passed through the zone of decoupling. Similarly, choices are often made about where the weak material will be created, for instance across the entire upper boundary of the model, just the subducting plate, or part of the subducting plate only.~~ In this study, we ~~do not attempt to examine all the permutations of the WL approach that appear in previous model. Rather, we focus on the most basic and common aspects the WL implementation . The analysis considers both the physical evolution of the layer, as well the implications from a numerical and modelling standpoint~~focus on a simple constant-viscosity implementation of the WL using the Underworld2 code, as described in Section 2.

~~There is degree of~~ Although the WL is commonly implemented in numerical subduction models, there is some ambiguity as to whether ~~the WL~~ this approach represents an attempt to explicitly model subduction interface dynamics, or should instead be conceived as an ~~solution to the sub-grid scale physics of faulting. This also reflects complexity in the physical system, as the shallow subduction interface is probably more fault-like, while the deeper interface may be more channel-like Gerya et al. (2002); Behr and Becker (2018).~~ ad hoc solution to modelling the plate boundary. WL implementations do mimic several critical processes that are ~~though~~ thought to contribute to the ~~weakness~~ long-term weakness and stability of the subduction interface. One of these is the process of entrainment or 'self-lubrication' (Lenardic and Kaula, 1994). The entrainment of ~~a range of materials~~ both sediments and fluids along the slab interface likely plays a key role in maintaining subduction interface weakness. Additionally, the WL approach might be linked with deformation-localizing processes such as damage, grain size reduction ~~,~~ and fabric development. Indeed, WL models have been conceived as a limiting case in which comprehensive damage is assumed within the interface material and hence can be prescribed at the outset (Tagawa et al., 2007). ~~Whichever way the WL is conceptualized, the primary requirement is to establish a sufficient decoupling between the plates. If the imposed interface~~

is too strong, subduction will be unrealistically slow, or may entrain the upper plate in a mode of catastrophic buckling and failureDespite the fact that we can draw analogies to relevant physical processes, our view is that the WL approach constitutes an ad hoc solution to the sub-grid phenomena of plate boundary shear localisation.

The behavior of subduction interface in numerical models the WL is discussed is some detail in Arcay (2012, 2017). These studies highlight the tendency for the subduction interface to develop spontaeous spontaneous thickness variation as the models progressevolve. Typically the interface widens near the trench, building a prism-like complex, and thins at depths beyond the brittle-ductile transition at around (50 kmsdepth). This pattern was also noted in the boundary element models of (Gerardi and Ribe, 2018), who attributed it a downdip thickness variation to lubrication layer dynamics. Because of this tendency , Arcay (2017) proposes that to achieve proper resolution, one must preemptively over-resolve the interface. A related issue is that many dynamic The tendency for thinning will impact the level of mesh resolution required in a model (Arcay, 2017). Additionally, WL formulations are likely to evolve towards maximum subduction interface thickness around twice the imposed thickness, a phenomena we explore in this study. For instance, a model with a prescribed weak layer of 5 km is likely to develop an maximum thickness throughout the decoupling region of 10 kmpotentially impacting effective stress along the interface, as well as the accuracy of predicted thermal structure. Understanding and controlling these unintended spontaneous width variations will be vital important in terms of using dynamic models to explore sensitive subduction-zone processes, such as metamorphism and melting near the slab top. Currently, these type of questions are mainly addressed using thermo-kinematic models. Bench-marking of community codes indicates that precision to about 1 km, & 20°K can be expected in well resolved thermo-kinematic models (van Keken et al., 2008). This provides a useful, if ambitious target for dynamic numerical subduction models.

4 Methods

4.1 Numerical model setup

The numerical subduction models developed in this study represent the time evolution of simplified conservation equations for mass, momentum and energy within a 2D Cartesian domain. Solutions to the equations are derived using a Galerkin finite-element approximation. Fig. 2 provides an overview of the model domain, as well as initial and boundary conditions. The depth of the domain is 1000 km, and the aspect ratio is 5. Initial temperature conditions define two plates which meet at the centre of the domain, including a small asymmetric slab following a circular arc to a depth of 150 km. The subducting plate has a an initial age of 50 Myr at the trench, while the upper plate age is 10 Myr. Both plates have a linear age profile with an initial age of zero at the sidewalls. The initial conditions allow for self-sustaining (fully-dynamic) subduction to proceed. There This setup allows the model to evolve under the driving force of internal density anomalies (sometimes referred to as a fully dynamic model). Apart from the presence of a WL, there is no compositional difference between the subducting (oceanic) and upper (continental) platenor and upper plate, nor do we include any compositional differentiation within the oceanic lithosphere.

The only aspects of the model setup that are varied are the details of subduction interface implementation (described in the following section) and the model resolution.

The mantle is treated as an incompressible, highly viscous fluid in which inertial forces and elastic stresses can be neglected. The mechanical behaviour of mantle (including the thermal lithosphere) is prescribed by a composite rheological model that includes a linear high-temperature creep law, as well as a scalar visco-plastic flow law, sufficient for capturing ~~psuedo-brittle~~ pseudo-brittle as well as distributed plastic deformation within the slab. Thermal buoyancy is the only source of density variation in the model. The thermal variations are coupled to the momentum equation through their effect on density, which follows the Boussinesq approximation. A detailed description of the governing equations, constitutive laws, and physical parameters are given in Appendix A.

Approximate solutions to the incompressible momentum and energy conservation equations are derived using the finite element code ~~Underworld 2~~Underworld2. Underworld2 is a Python API ~~(Application Programming Interface)~~ which provides functionality for the modelling of geodynamic processes. Underworld2 solves the discrete Stokes system through the standard mixed Galerkin finite element formulation. The domain is partitioned into quadrilateral elements, with linear elements for velocity and constant elements for pressure (Q_1/dP_0)(Arnold and Logg, 2014). ~~The typical models discussed in this chapter~~ Material properties are advected on Lagrangian tracer particles. Unless otherwise stated, models have a mesh resolution of 160 elements in the vertical direction, refined to provide ~~a~~ an element width of ~~2~~ 3 km at the surface~~. Material properties are advected on Lagrangian tracer particles, with~~ , and a particle density of 30 tracers per element. Particles are added and removed to maintain density near this value. During quadrature, material properties are mapped to quadrature points using nearest-neighbour interpolation. The Lagrangian tracer particles are used to distinguish the subduction interface material from the rest of the system (lithosphere and mantle). Underworld2 solves the energy conservation equation using an explicit Streamline Upwind Petrov Galerkin (SUPG) method (Brooks and Hughes, 1982). In this approach, a Petrov-Galerkin formulation is obtained by using a modified weighting function which affects upwinding-type behaviour. The Stokes system has free-slip conditions on all ~~side and bottom~~ boundaries. The energy equation has constant (Dirichlet) and zero-flux (Neumann), on the top and bottom boundary respectively. The ~~right hand sidewall has a constant (potential) temperature which enforces a ridge at that boundary. The~~ left and right sidewalls have a constant ~~potential temperature of the mantle is~~ temperature equal to the mantle potential temperature (1673 ~~°K, and the~~ K). The surface temperature is 273 ~~°~~K.

4.2 Subduction interface implementation

The focus of our study is a common approach to modelling the subduction interface, ~~referred to here as continuously entrained weak layer (WL). Past studies have also termed this a~~ sometimes referred to as a weak layer or weak crust. ~~Generally, this approach have two defining attributes~~This approach has two intrinsic features: 1) the material that provides decoupling within the interplate zone is a distinct material type ~~contrasting with the background material (Billen and Hirth, 2007; Garel et al., 2014; Čížková a~~ with rheology that contrasts with the plates/background material; and 2) the weak material layer is distributed along some or all

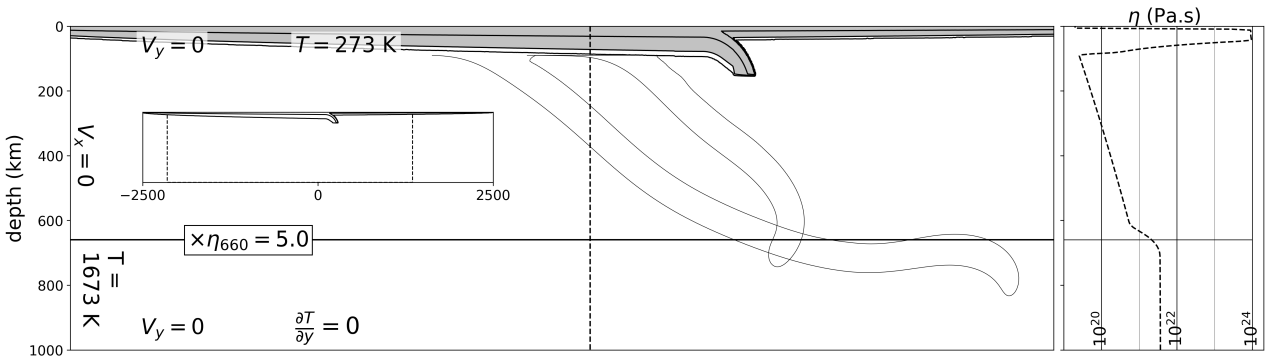

Figure 2. Initial conditions for 2D thermo-mechanical subduction models. All models in this study ~~models~~ have a 50 Myr initial slab age~~, and free-slip~~ . Velocity boundary conditions are free-slip on ~~upper boundary~~all walls of the domain (zero tangential stress). The dark gray region in the main panel shows parts of the model colder ~~that~~ than 1100 °C. Inset shows the full domain. Outlines ~~of~~ show the slab ~~is shown~~ position at ~~later times:~~ 10 ~~&~~ and 20 Myr. Right hand panel shows viscosity profile evaluated along the vertical dashed line in main panel.

of the subducting plate, so that the flow itself entrains new weak material into the decoupling region ~~. Throughout this study , we~~ (Billen and Hirth, 2007; Garel et al., 2014; Čížková and Bina, 2013; Arredondo and Billen, 2017; Agrusta et al., 2017). In this study we use the abbreviation WL to refer to a ~~standard WL approach as~~ typical implementation, namely one in which the distribution of weak material is fully self-evolving within the deforming subduction interface zone. We also demonstrate ~~an improved~~ a modified version of this approach, which we refer to as an embedded fault (EF). The EF implementation ~~includes a reference line advected along the base of the weak layer. The reference line provides a way to constrain the~~ differs from the WL in that the width of the interface is constrained in terms of its minimum and maximum thickness~~of the weak layer material , (i.e. to alter the material distribution within the model by setting subduction interface particles to background particles~~ [1] ~~and vice versa).~~ . While this manipulation of the interface material could be achieved in different ways, our implementation is based on a reference line of particles, which lie at the base of the weak interface and are advected along with the material swarm. At each timestep we remap material types based on the relative position of the material particles with regard to the reference line. The relationship between the EF ~~tracer~~ reference line, the ~~global~~ material swarm, and the mesh is shown in Fig. 3. In the EF approach, we also initialise the ~~weak material~~ interface with a non-uniform thickness. ~~Within the interplate zone, the weak material will typically have twice~~ For reasons discussed in the following sections, the thickness of the interface material within the decoupling zone should be close to double the thickness that is prescribed on the top of the subducting plate. ~~The choice emerges from the behavior we observe in the standard WL models, as discussed in the following section.~~

It is important to emphasise that differences between the WL and EF implementations relate only to the distribution of weak material within the model. All other aspects of the interface representation remain identical. In each case the interface has a constant viscosity. Weak material is continually prescribed along the upper-most part of the ~~subduction plate~~ subducting

[1] ~~In this study Lagrangian tracer particles are used to track material properties. Mesh/field based methods have also been used to implement a WL (e.g. Garel et al., 2014).~~

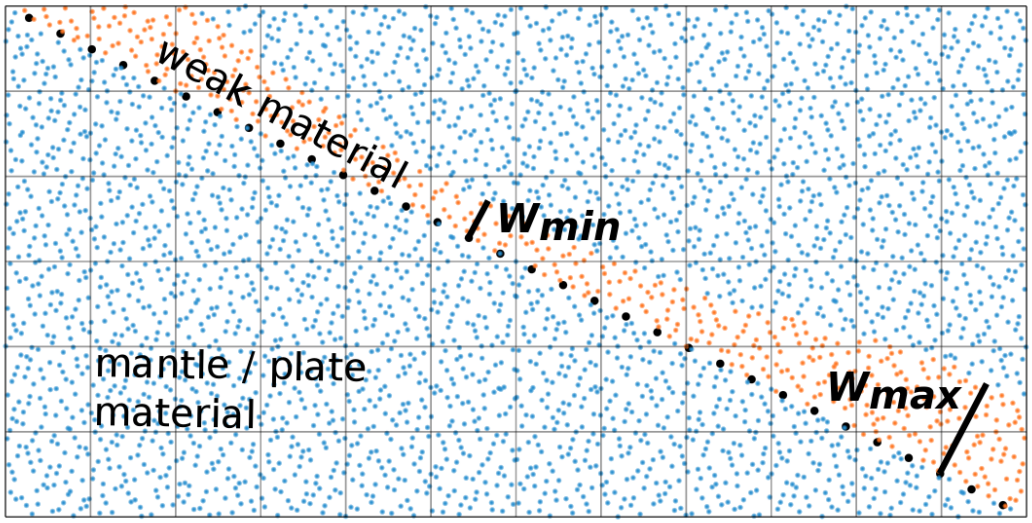

Figure 3. Schematic of the embedded fault (EF) implementation. The EF is a modification to the standard weak layer approach (WL)~~approach~~. As with the WL, shear ~~localization along the subduction interface is encouraged by~~ localizes due to a finite-width layer of weak material (represented by orange particles). ~~In the EF approach, a~~ A reference line of tracer particles (black points) is advected with the flow, at the base of the ~~interface~~weak material layer. This line provides a ~~set of~~ reference ~~points~~ to enforce width limits on the weak material, denoted by W_{min} and W_{max}. ~~This approach mitigates~~ At each timestep we remap material types based on the ~~strong thinning~~ relative position of the ~~layer that occurs in~~ material particles to the ~~standard WL approach (see Fig~~reference line. ~~4), and also reduces short-wavelength instabilities~~ Note that ~~develop at the boundary between~~ the ~~subduction interface and the upper plate~~element size here is not shown to scale.

plate as it moves away from the ridge, with a specified constant thickness (W_{init} = 10 km). ~~Meanwhile, the~~ The upper plate does not contain any weak/interface material. For simplicity, there is no rheological variation between the shallow (frictional) megathrust and the deeper, viscous interface. ~~A~~ To control the upper limit of the MDD, a depth-dependent cosine taper is used to transition the subduction interface rheology to the background mantle rheology. ~~The transition is imposed by a static depth dependent function. In this chapter the cosine~~ The taper for the transition (both WL and EF) begins at 100 km, and has a width of 30 km. ~~This type of behavior is very simple to implement using the Underworld2 function capability.~~ Decoupling is strongly inhibited at depths greater than the taper onset. Hence the depth of the taper onset (100 km) effectively ~~sets the maximum depth of decoupling, as shown in Fig. 9~~controls the upper limit of MDD. Likewise, at a distance of 800 km from the trench, along the subducting plate, the rheology of the interface material transitions to background mantle rheology. This avoids interaction of weak material with the spreading ridge.

5 ~~Results~~Analysis of modelling approaches

5.1 ~~Analysis of~~ The weak layer approach

~~In a number of previous studies it was observed that the WL approach tends to develop spontaneous variations in the interface thickness (Arcay, 2017; Duretz et al., 2012; Gerardi and Ribe, 2018). Here we identify two separate causes for these thickness variations. The first can be understood as an outcome of the evolution towards uniform volumetric flux along the interface. This phenomenon is likely to be present in any implementation of the WL approach. The second represents shorter-wavelength instabilities at the boundary between the subduction interface and the upper plate. Both of these characteristic behaviours can be seen in Fig. 4, where we plot the interface thickness as a function of the downdip distance during 20 million years of model evolution. During the first few million years of the model evolution, changes in the interface thickness have both spatial and temporal dependency. The interface thickens at distances of up to 150 km downdip from the trench, whereas beyond this it thins. The point of zero thickness change (10 km) migrates downdip with time.The interface reaches a minimum thickness of around 5 km at ∼ 2.0 Myr, while it reaches 20 km, close to its maximum thickness, at around 7 Myr. The thickened part of the interface evolves to a near-constant configuration, both in thickness (∼ 20 km) and downdip extent (∼ 200 km) . In contrast the thinning of the deeper interface is transient and by ∼ 15 Myr the deeper part of the interface has reestablished a thickness close to the initial value (10 km). In the following section we provide a simple explanation for these systematic width variations , based on a simplified kinematic description of flow within the WL.~~

~~In the WL approach, interface thickness variations appear to be a simple outcome of the evolution towards uniform volumetric flux along the interface. The flux is a function of both the interface thickness as well as the effective boundary conditions on the~~

Fig. 4 shows the evolution of the subduction interface thickness over a 20 Myr period, based on the WL setup described in section 4. This model reveals the spontaneous development of substantial WL thickness variations, not only near the trench (i.e. near the accretionary prism) but throughout the entirety of the decoupling zone. Previous studies have commented on the occurrence of similar thickness variations (Arcay, 2017; Gerardi and Ribe, 2018) although their origin has not been closely considered. We propose that such variations arise primarily from the way that the kinematics of the flow in the WL effect the downdip volumetric flux. In the following discussion we refer to a slab-based coordinate system in a fixed upper plate reference frame, where $\hat{y}$ is the direction orthogonal to the slab midplane, ~~and~~ $\hat{s}$ is ~~a vector~~ the direction parallel to the slab midplane, and the plate convergence velocity is equal to the subducting plate velocity (V_s). Before reaching the trench, the ~~weak interface~~ WL material travels with the subducting plate, with ~~effectively uniform velocity and zero shear~~ ($\frac{\partial V_s}{\partial \hat{y}} = 0, V_s = V^{\text{conv.}}$) near-uniform velocity and negligible shear ($\frac{\partial V_s}{\partial \hat{y}} \approx 0$). In the ~~interplate region~~subduction interface zone, however, the ~~weak interface~~ WL material decouples the slab and upper plate, and ~~the velocity gradient is~~ velocity gradients across the interface are finite. For a Newtonian rheology, the gradient is expected to be linear with $\frac{\partial V_s}{\partial \hat{y}} = \frac{V^{\text{conv.}}}{W} \frac{\partial V_s}{\partial \hat{y}} \approx \frac{V_s}{W}$ (i.e. Couette flow). This change in velocity ~~gradient means that the volumetric flux across the interfaceis reduced,~~ profile across the WL means that there is a

~~Interface thickness evolution in WL implementation. The initial layer thickness (W_{init}) is 10 km. Colored lines show the evolving thickness plotted as function of downdip distance from the trench.~~

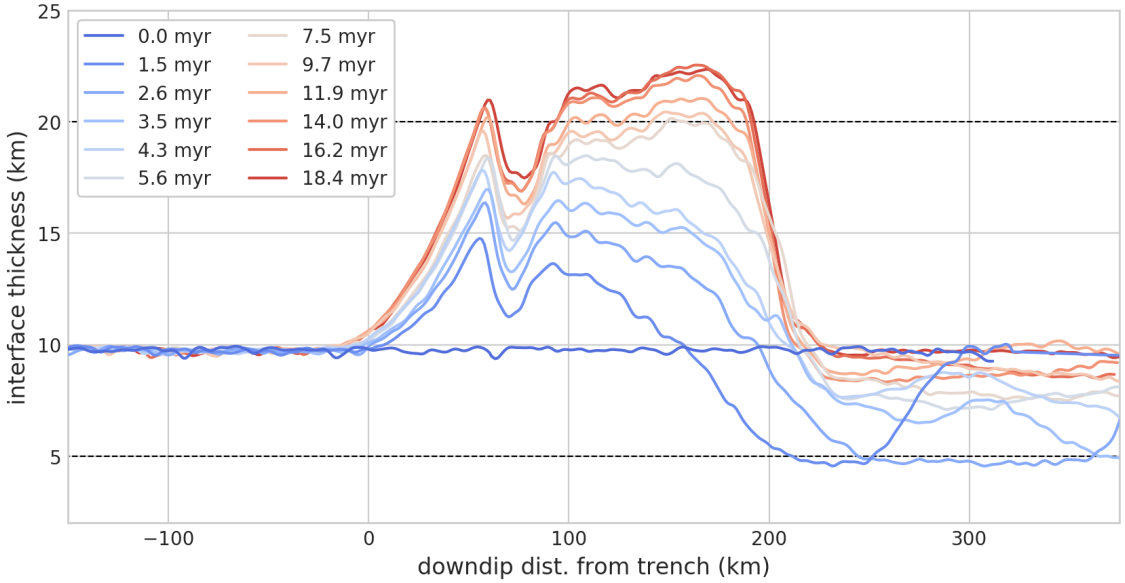

Figure 4. **Interface thickness evolution in WL implementation**. The initial layer thickness (W_{init}) is 10 km. Colored lines show the evolving thickness plotted as function of downdip distance from the trench. Dashed black lines show the expected maximum (equilibrium) and minimum (transient) thickness, based on a kinematic description of the flow evolving from a uniform thickness weak layer.

smaller volumetric flux (per unit length normal to the interface) compared to the flux of ~~interface material arriving~~ material being delivered on the incoming subducting plate. ~~This causes material to accumulate at the trench, progressively thickening the interface from the top down. Assuming a linear velocity profile, the volume flux will reach equilibrium when the thickness of the decoupling regionis twice the prescribed thickness (i. e. the thickness on the top of the subducting plate). This is~~ To illustrate this point we consider a simple boundary-driven Stokes flow, as shown in Fig. 5, which provides a useful analogy to the flow in the WL. In this model, flow is driven by a constant tangential velocity on the lower boundary, while the upper boundary of the model has a patch of frictional (no slip) nodes, surrounded on both sides by a free-slip boundary condition. The step change in boundary conditions imposes volumetric flux gradients causing the weak layer (shown in blue) to thicken near the start of the no slip region, and thin near the end. The width changes of the weak layer proceed until the overall volume flux reaches equilibrium.

Returning to our subduction model results, the validity of our description can be tested by considering the thickness profile required for volume flux equilibrium. Assuming that the decoupling zone maintains a Couette profile, independent of the local thickness, flux equilibrium requires that the thickness of the WL in decoupling zone will be twice that imposed on the incoming plate. This prediction, based on a simplified kinematic description of flow in the interface is broadly consistent with

the ~~equilibrium thickness of the interface that develops~~ long-term evolution observed in the WL models (e.g. Fig. 4). ~~At the base of the subduction interface~~

While our discussion so far has mainly emphasised the transition of the WL from the oceanic part of the subducting plate into the decoupling zone, similar processes take place where the slab and ~~mantle flow becomes coupled (sometimes referred to as the Maximum Depth of Decoupling - or~~ the mantle become coupled (the MDD). When a constant interface width is prescribed, the interface thickness ~~at~~ near the MDD initially decreases. Again, this is because the volumetric flux increases ~~at the MDD, as~~ as Couette flow in the decoupling region ~~(Couette flow)~~ transitions to the fully coupled flow. ~~In effect, we see a reversal of the process that occurs at the top of the subduction interface.~~ The location where the subduction interface thins~~is~~, provides a proxy for the MDD~~, again because of the relationship between the flow kinematics and volumetric flux. In typical thermo-mechanical subduction model setups, the MDD is unlikely to be~~. A complexity arise due to the fact that the MDD in our model is not constant. Instead the MDD tends to become deeper as the corner of the mantle wedge cools and stagnates~~(see Fig. 9).~~

~~A simple boundary-driven Stokes flow provides a useful analogy to the thickness evolution of the subduction interface in the WL approach. This boundary-driven model consists of an incompressible Newtonian fluid that transitions from uniform flow to Couette flow, due to a change in the mechanical boundary condition along one side of the model, from frictionless (or free slip) to frictional (no slip). The material distribution consists of a weaker and a stronger layer~~. This means that the effective boundary conditions on the flow within in the interface are also time varying. This accounts for why the location of interface thinning migrates downdip as time progresses, as shown in Fig. ~~5. The step change in boundary conditions imposes horizontal velocity gradients which must be balanced by vertical velocity gradients. The vertical velocity causes the weak layer to thicken near the start of the no slip region, and thin near the end. The width changes of the weak layer proceed until the overall volume flux reaches equilibrium~~4.

Based on analysis of typical WL setup, we ~~see~~ argue that the primary ~~control on the interface thickness variation is the condition of uniform~~ thickness variations reflect a simple response to changes in volumetric flux along the interface. ~~The duration of the transient thickness evolution is driven both by the timescale of accumulation of material as well as by changing kinematics of flow due to deepening of the MDD.~~ In addition to these flux-controlled changes, Fig. 4 shows that the subduction interface also develops a persistent short-wavelength width perturbation just beneath the tip of the forearc at a downdip distance of $\sim$ 60 - 80 km. Here the interface thins by $\sim$ 3 - 4 km. This occurs in combination with strong, localised plastic deformation in the upper plate. It seems likely that ~~these~~ such short-wavelength anomalies may be ~~effected~~ affected by a range of factors, including the interface and lithosphere rheology, mesh resolution, and material advection scheme. If so, these ~~feature~~ features are likely to be ~~rather model /code~~ somewhat model dependent, in contrast to the to ~~longer-wavelength~~long-wavelength, flux-related thickness variations which reflect intrinsic kinematics of flow in the WL setup.

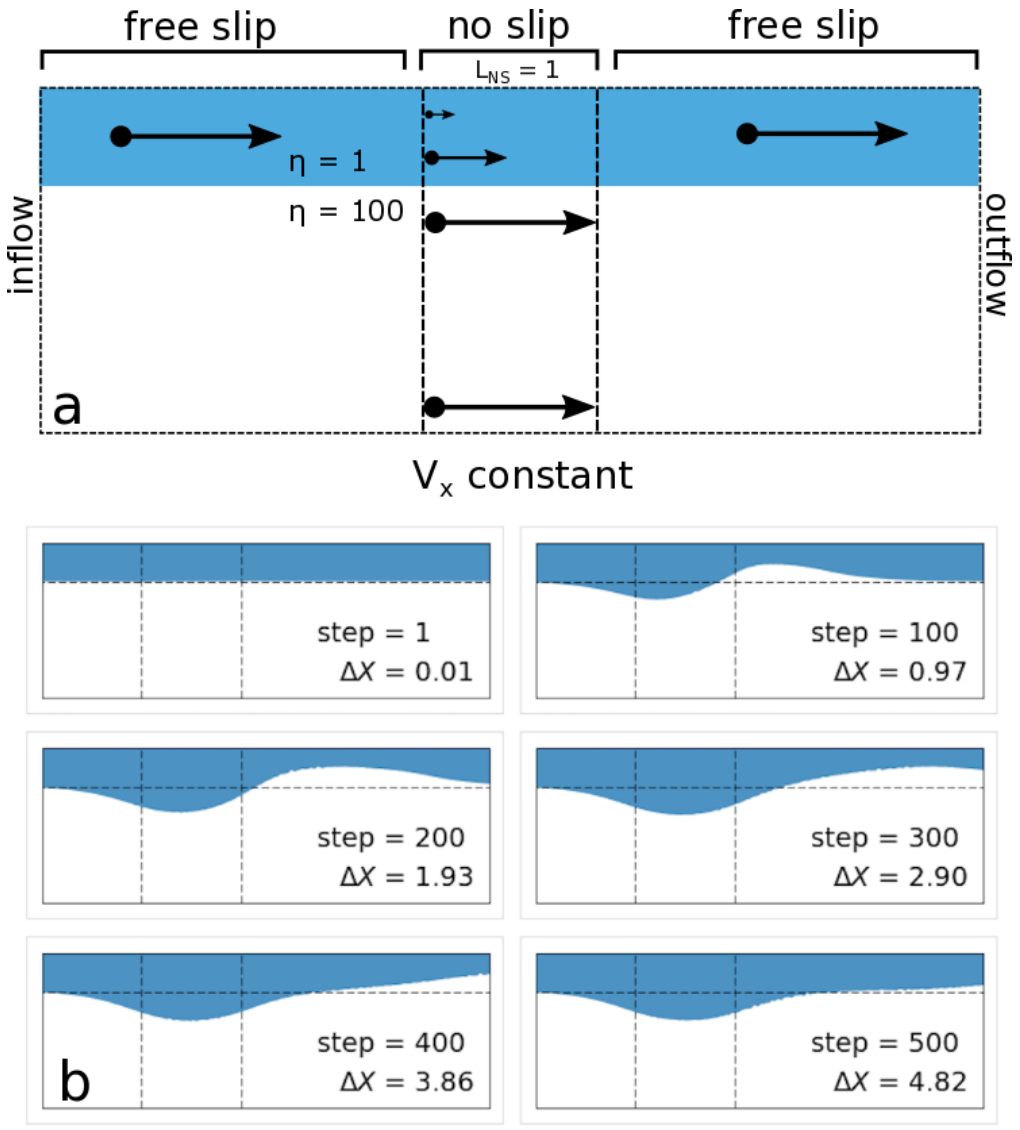

Figure 5. ~~Simple analogue~~ Analogue **for WL** ~~dynamics~~thickness evolution. a) model setup for a boundary-driven flow where a weak layer (shown in blue) interacts with a stronger layer (white). The top surface of the model is free slip, except for a patch of no-slip nodes ($V_x = 0$). The bottom surface has a constant horizontal velocity component. b) evolution of material in the boundary-driven model. ΔX represents the accumulated displacement along the bottom boundary, normalised by the width of the no-slip patch.

5.2 Improving the weak layer approach

The embedded fault (EF) implementation, described in Section 4, ~~primarily enables us to control the subduction interface thickness (i.e. the weak layer thickness). In general, this may consist of controlling the minimum thickness(W_{min}), the maximum thickness (W_{max}) , or both. Perhaps~~ consists of two complimentary strategies. Firstly, we may initiate the WL with a variable thickness. Secondly we can control the WL thickness throughout the simulation, by remapping WL material (particles) to background material, and vice versa. Given these controls, perhaps an obvious first issue to address is what happens if we simply enforce a constant width interface at all times. While this would in some ways be a desirable approach, doing so results in the development of a very spurious subduction morphology. Fig. 6~~a~~ b shows results from such a case (~~$W_{max} = W_{min} = W_{init}$~~ $W_{\max} = W_{\min} = W_{\text{init}} = 10$ km). After 15 million years of model evolution a very atypical subduction morphology has developed, with an extremely low angle megathrust beneath a forearc region with a width of greater than 600 km. While ~~physically inconsistent, this constant width EF example~~ geologically irrelevant, the example shown in Fig. 6b provides a useful insight into the way in which the ~~interface deformation~~ flow in the interface can influence model dynamics. When the subduction interface is forced to remain at constant width, the interface is unable to evolve towards flux equilibrium. Persistent interface-normal velocity components result, and the compounding effect eventually distorts the morphology of the entire subduction hinge region.

Fig. 6~~b~~ a shows the slab morphology developed by instead using ~~$W_{max} =$~~ $W_{\max} = 1.9 \times W_{\text{init}} =$ 19 km ~~$= 1.9 \times W_{init}$~~ (we discuss the choice of 1.9 in the next section). The subduction morphology in this simulation is ~~much~~ more realistic, and consistent with ~~the behavior we would expect using~~ outcomes from a standard WL approach. The evolution of the interface thickness from the same simulation is shown in Fig. 5.2. In addition to controlling the width of the interface throughout the EF simulation, we have also prescribed the initial interface thickness in the decoupling zone (from the trench to a depth of 100 km depth) to have value equal to ~~W_{max}~~ $W_{\max}$ (19 km). In this way, we have tried to preemptively impose a thickness profile closer to the flux equilibrium. Fig. 5.2 shows that this strategy reduces, but does not fully eliminate, the transient stage of interface adjustment. It is ~~challenging~~ difficult to fully eliminate the transient stage because the equilibrium interface thickness profile ~~depends critically on the depth of slab mantle coupling (MDD). In this thermo-mechanical subduction setup, the~~ is party determined by the MDD, which is not constant. The evolution of the MDD is a response to the cooling of the mantle wedge and the development of a stagnant 'cold corner' (see Fig. 9). Prescribing initial temperature conditions that include the cold corner, would be one way to further reduce the amount the transient adjustment of the interface. Fig. 7 shows the state of the material swarm in the EF model, compared to the equivalent WL model. This reveals a typical distribution of interface material once a quasi-equilibrium thickness has been established.

In addition to the thickness variations related to volumetric flux, ~~there also tends to be~~ the WL approach can develop short wavelength thickness variations, as seen in Fig. 4. In the EF implementation, we found that using a value of $W_{\max}$ slightly less than $2.0 \times W_{\text{init}}$ helps to suppress short wavelength thickness variations, without significantly effecting the overall model evolution~~relative to a larger value~~. In other words, while we need to allow the interface to develop some amount of thickness

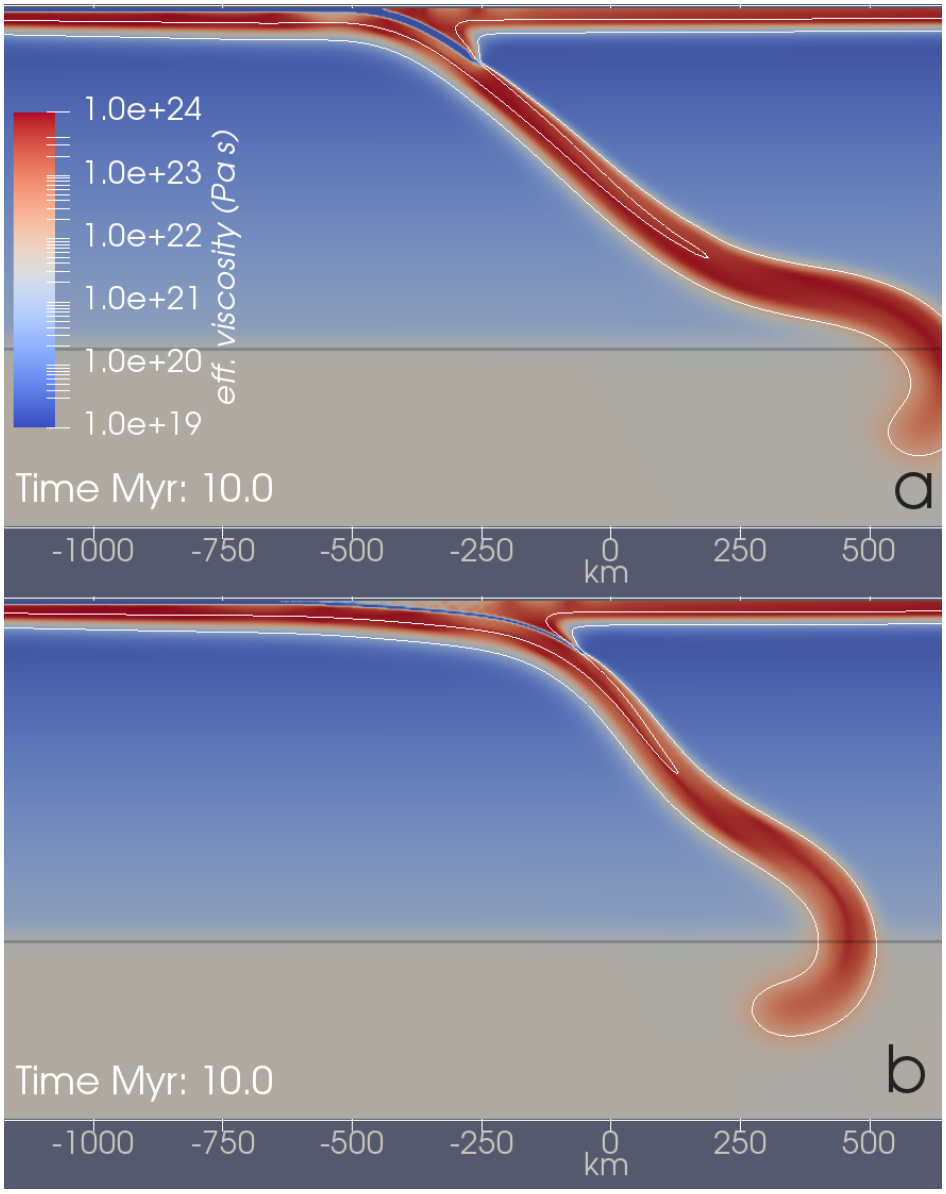

Figure 6. **Effect of variable maximum W_{max} in EF models.** The colormap shows the effective viscosity at a model time of 10 Myr. White lines show isotherms at 700 and 1200 °C. a) shows EF model with $W_{max} = 1.9 \times W_{init}$ = 19 km. b) shows EF model with $W_{max} = W_{init} = 10$ km. In this case, due to the fact that the subduction interface cannot adjust its thickness, an extremely long, low-angle interface develops, with forearc distances of $\sim$ 500 km.

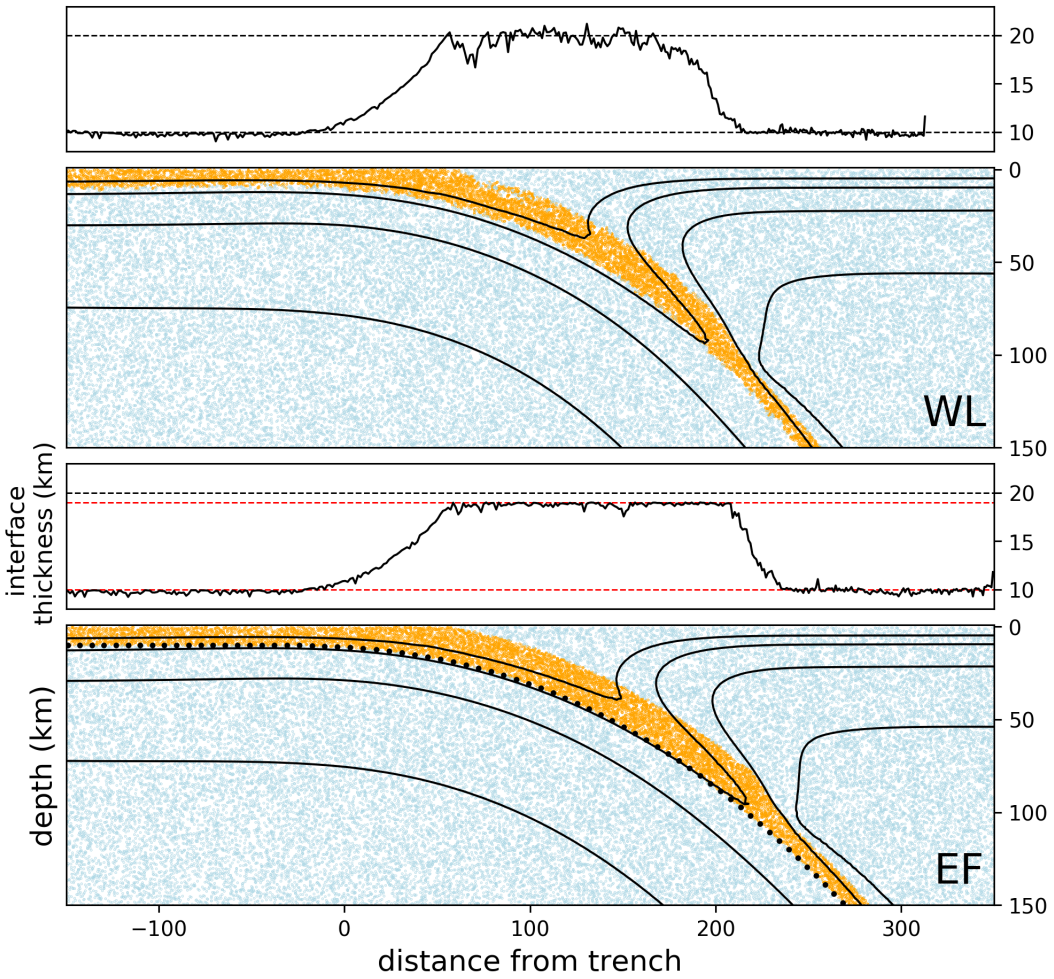

Figure 7. **Distribution of subduction interface material**. Results from a standard WL model are shown in the top panel, EF model shown in bottom panel. Both models have a constant viscosity interface rheology. Model time is 12.5 Myr in both cases. In the lower panels, orange points are the subduction interface material, blue points are the background material (mantle/lithosphere). Points in the material swarm and along the EF reference line have been down-sampled for clarity. Solid black lines over the material points are isotherms. Smaller panels show the measured interface thicknesses in each case; red horizontal lines show the thickness constraints $W_{\min}$, $W_{\max}$.

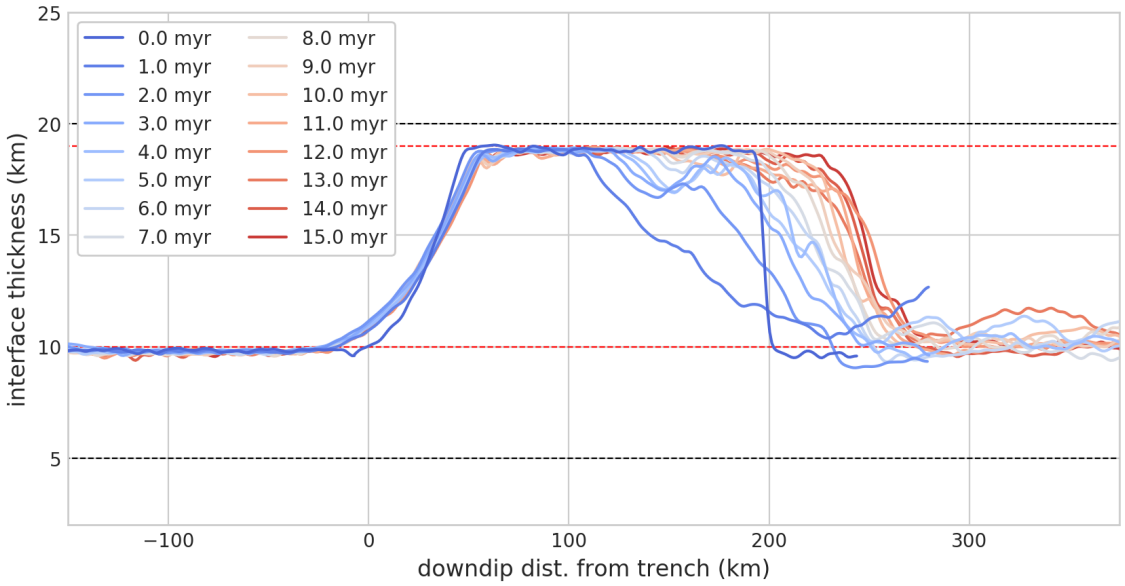

Figure 8. Interface thickness evolution, EF approach. In the EF implementation limits are imposed in the maximum and minimum thickness, as shown by the red dashed lines. Unlike the standard WL approach, the initial thickness of the interface is variable. The initial thickness of the interface within the decoupling region is equal to $W_{\max}$. The black lines show typical range in which a WL model will vary, assuming a constant initial thickness of 10 km (see also Fig. 4). The results shown here refer to the same model as shown in Fig. 1 and Fig. 6.

variation, it may be advantageous to use a value slightly less than $2 \times W_{\text{init}}$. Fig. ~~7 shows the material and interface thickness distribution using WL and EF implementation. Horizontal red lines show the location~~ 10 shows results from a number of experiments where the value of $W_{\max}$ ~~and $W_{\min}$. One advantage of the EF approach is that it offers improved precision in determining the thickness of the subduction interface. Such precision will be important for studying highly pressure- and temperature-sensitive processes, such as metamorphism and melting near the slab top~~was changed. Qualitatively, we see that that the models have very similar long term evolution once $W_{\max}$ is greater that $1.75 \times W_{\text{init}}$.

~~Distribution of subduction interface materials. Results from a standard WL model are shown in the top panel, EF model shown in bottom panel. Both models have a constant viscosity interface rheology. In the lower panels, orange points are the subduction interface material, blue points are the background material (mantle/lithosphere). Points in the material swarm and along the EF reference line have been down-sampled for clarity. Solid black lines over the material points are isotherms. Smaller panels show the measured interface thicknesses in each case; red horizontal lines show the thickness constraints $W_{\min}$, $W_{\max}$.~~

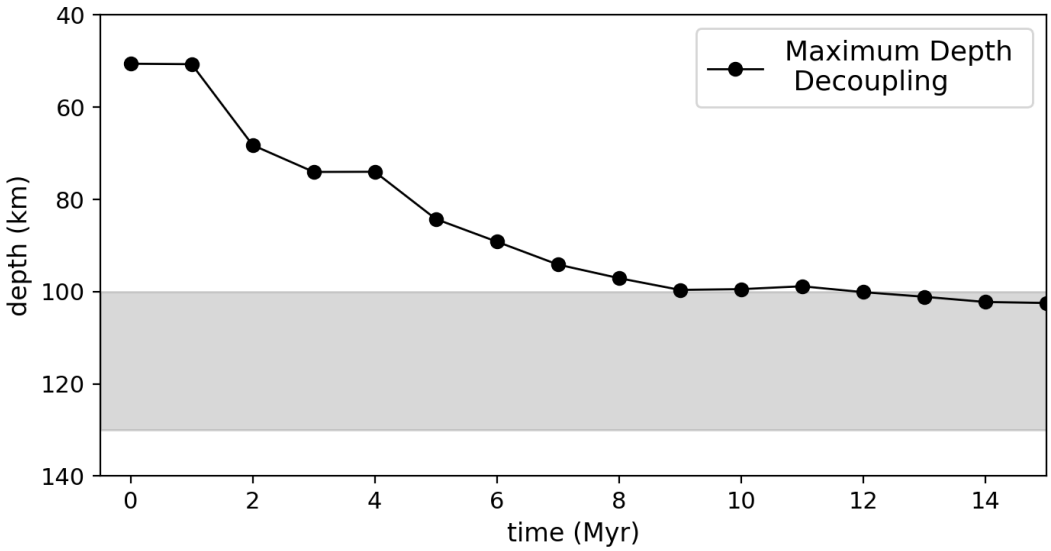

Figure 9. ~~**Embedded fault models (EF) with variable $W_{\max}$. The colormap shows the effective viscosity at a model time of 14 Myr. Parts E**~~**volution of the ~~model with temperature above 1250 °C are masked~~ maximum depth of decoupling (~~white). White lines show isotherms at 700 and 1100 °C. a~~MDD)** ~~$W_{\max} = W_{\text{init}} = 10$ km~~. ~~Note~~ As the ~~extremely long~~mantle wedge cools, ~~low-angle morphology of~~ it progressively stagnates, driving deeper decoupling within the subduction interface~~that develops when~~. The grey region in the figure shows the depth interval over which the subduction interface ~~cannot adjust its thickness. b~~transitions to the background mantle rheology as prescribed with a cosine taper (see Section 4)~~$W_{\max} = 1.9 \times W_{\text{init}}$~~. ~~This choice results~~ Results are from the same model as shown in Fig. 6a~~more physically-consistent subduction morphology~~, i.e. an EF model with $W_{\max} = 1.9 \times W_{\text{init}}$

~~Evolution of the maximum depth of decoupling (MDD). As the mantle wedge cools, it progressively stagnates, requiring deeper decoupling within the subduction interface. The grey region in the figure shows the depth interval over which the subduction interface transitions to the background mantle rheology as prescribed with a cosine taper (see main text and Appendix A for further information on the interface rheology).~~

5.3 Stability and convergence

So far we have discussed results based on well-resolved models, with 160 elements across the 1000 km vertical domain. At this resolution, the vertically-refined mesh provides 3.2 elements within the subduction interface (at ~~$W_{init} =$~~ $W_{init} =$ 10 km). Note that this increases to more than 6 elements in the decoupling zone, once the interface has thickened to $\sim$ 20 km. We now look at the convergence of models under variable resolution, based on a standard WL approach as well as the EF implementation. Fig. 11 shows the slab temperature field ~~,~~ at 10 Myr ~~after the initiation of the model, for different~~ for a set of models with varying resolution: 72, 96, 128, 160, 192 elements in the vertical dimension. The ratio of the initial interface width (~~W_{init}~~W_{init}) to the element width has values of 1.4, 1.9, 2.6, 3.2, 3.8 for the respective mesh resolutions. The dashed slab outline in Fig. 11 shows the morphology for a simulation with 192 elements in the vertical dimension; this is used as a reference model in

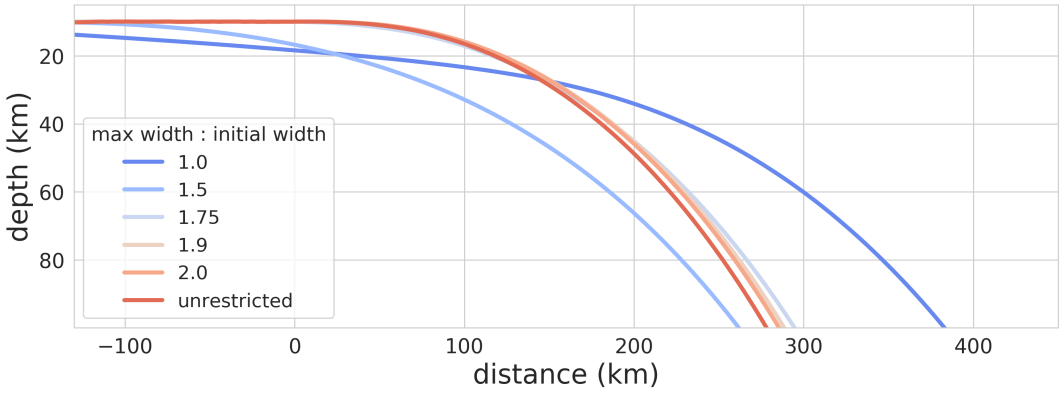

Figure 10. Embedded fault models (EF) with variable $W_{\max}$. Lines ~~represent~~ show the morphology of the base of the subduction interface (the EF reference line) at 10 Myr for different models with varying $W_{\max}$. ~~If~~ When the interface thickness variation is strongly constrained, the evolution of the model is significantly effected. See also Fig. 6.

the following analysis. Note that while we have varied the mesh resolution, all models have the same spatial particle density. Fig. 11a shows results using a standard WL implementation. At the lowest resolution (72 elements) the simulation stalls and the slab undergoes runaway thermal decay. At 96 elements, the model is still strongly impacted by under-resolution of the subduction interface. Fig. 11b shows equivalent results using the EF implementation. Qualitatively, we see that EF models are more stable at lower resolution. For instance, the EF model at lowest resolution (72 elements) ~~produces better results (closer to~~ more closely reproduces the evolution of the reference model ~~)~~ than does the WL model with 96 elements.

~~Interface thickness in under-resolved WL model.In the lower panel orange points show the subduction interface material, blue points are the background material (mantle/lithosphere). Upper panel shows interface thickness. Figure shows the model state 10 Myr after model initiation. This model is also shown in the upper-left panel of Fig. a.~~

Fig. 11 suggests that the EF models converge more closely with increasing resolution. We quantify this by tracking the relative error (L_2) of the temperature field in the lower resolution models with respect to the reference model (192 elements). The relative error with respect to the reference model (T_{ref}) is:

$$E = \left[\frac{\int_\Omega (T - T_{\text{ref}}) \cdot (T - T_{\text{ref}}) dV}{\int_\Omega T_{\text{ref}} \cdot T_{\text{ref}} dV} \right]^{1/2} \tag{1}$$

Fig. 13 shows relative error results for the same set of models shown in Fig. 11, confirming ~~that the EF has better resolution convergence than~~ more rapid resolution convergence in the EF models relative to the standard WL approach. For most of the models shown in Fig. 13 the relative error accumulates rapidly in the first 7-10 million years of the simulation, while the error rate flattens after this. This is similar to the time taken for the WL interface to reach its equilibrium thickness (e.g. Fig. 4). This suggests that models are ~~particularly~~ strongly resolution-sensitive during the transient phase of the interface ~~thickness~~

adjustment. At low resolution (72, 96, 128 elements), errors in both WL or EF models express this sensitivity. Interestingly, at 160 elements, the EF case exhibits nearly-constant error accumulation during the model evolution. This suggests that we have succeeded in reducing the sensitivity of the subduction interface implementation, relative to the overall error accumulation rate. The latter may be influenced by additional factors which we have not controlled for here, such as the timestep size in the advection-diffusion implementation which is controlled by the mesh resolution).

Fig. 12 shows the distribution of interface material in the lowest-resolution WL model (shown in the top left panel in Fig. 11a). In this case, under-resolution of the weak layer induces strong coupling between the slab and the upper plate at relatively shallow depths and begins to thin the WL, ~~according to the principles outlined~~ due to the development of strong flux gradients (as described in Section 5). This induces further coupling and yet more thinning. This proves to be catastrophic feedback process, causing the simulation to stall and enter runaway thermal decay. The EF approach provides stability in this context, by inhibiting the feedback cycle. While this behavior is mainly relevant for models run at low resolution, the increased stability of the EF approach is a useful property, particularly from a model development perspective.

6 Discussion and Conclusions

The entrained weak layer (WL) is a common approach for implementing the subduction interface in long-term dynamic simulations. We have discussed aspects of WL implementation that can have an unintended impact on model evolution. The first involves the transient evolution of a uniform thickness interface to a variable thickness - uniform flux, configuration. If not properly accounted for, thinning of the deeper part of the WL could lead to numerical under-resolution, as previously suggested (Arcay, 2017). If the WL has a viscous rheology, the thinning will lead to higher stresses. This can induce a positive feedback when higher stresses increase the amount of partial coupling, inducing further thinning. Even for seemingly well-resolved models, the transient behavior of the subduction interface appears to be responsible for strong mesh sensitivity and poor resolution convergence. In general, models with plastic/frictional rheologies should be less sensitive to these transient adjustments, as the stress should not depend on the width of the interface.

Another tendency of the WL models is to develop persistent short wavelength thickness variations. These may represent interface instabilities, as are observed in Couette flows past a deformable boundary (Shankar and Kumar, 2004). These tend to dominate on the shallow part of the boundary with the upper plate. While flux-related thickness variation will be expected for any model implementation of WL, boundary instabilities (short wavelength) are likely to be more variable across different codes. They may depend on additional details of implementation, such as the rheology of the plates, material advection and interpolation schemes. Together, these issues are likely to hinder efforts to produce reproducible results between codes.

These unintended behaviors of the WL approach can be partly mitigated by controlling the thickness of the interface. We demonstrate a simple implementation of this concept, which utilises a line of reference points at the base of the WL, allowing us to remap material types based on proximity to this line. We call this approach an embedded fault (EF). The ability to

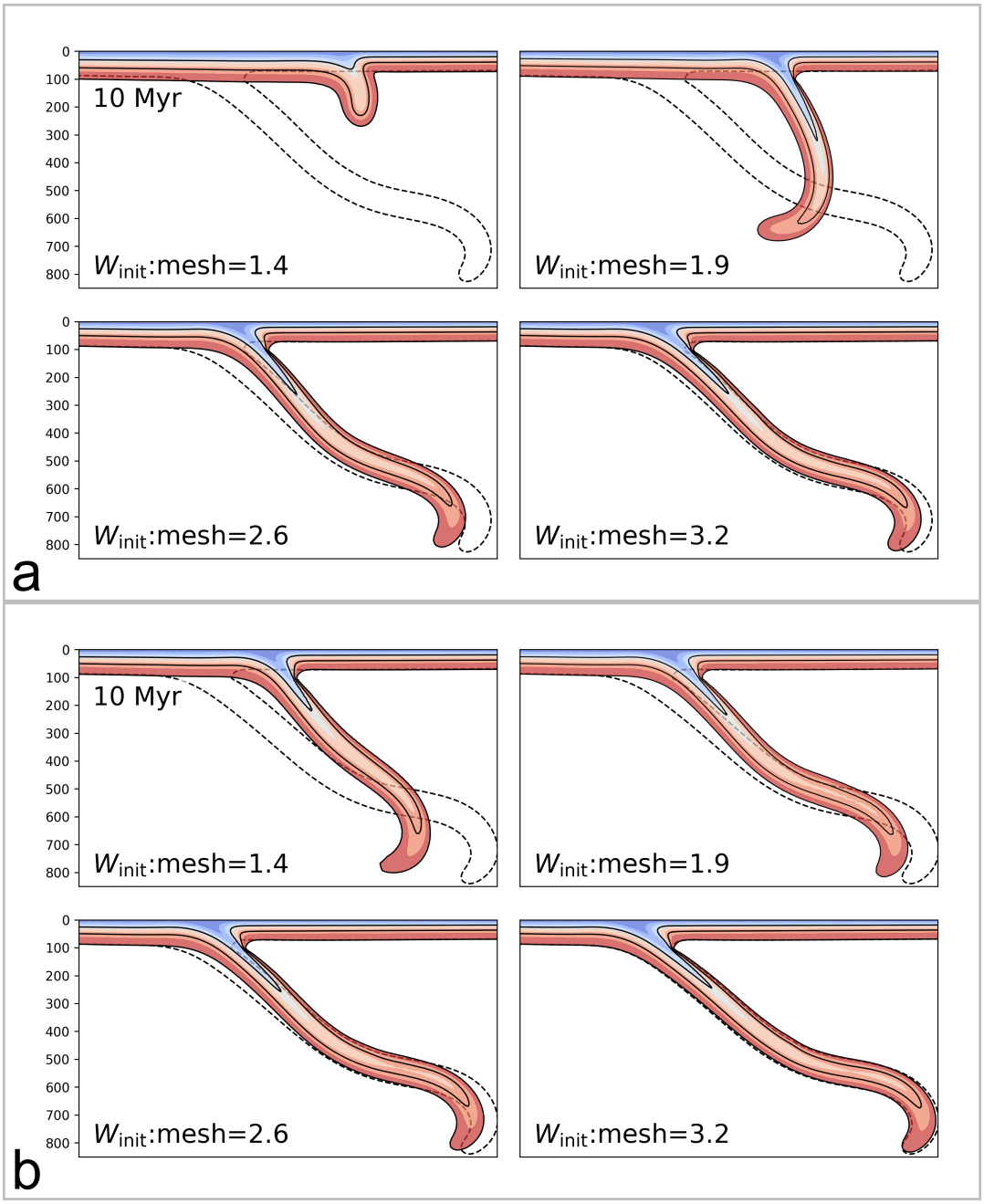

Figure 11. ~~Model behaviour~~ Implementation comparison **with variable resolution.** Colormap shows slab temperature field at 10 Myr, masked above 1250 °C. Each figure shows a series of models run at different resolution. The total number of elements in the vertical coordinate was 96, 127, 160, 192. Corresponding subduction interface resolution is displayed in the figure, representing the initial fault thickness (W_{init}) divided by the local element size. a) models using standard WL implementation. b) models using the EF implementation.

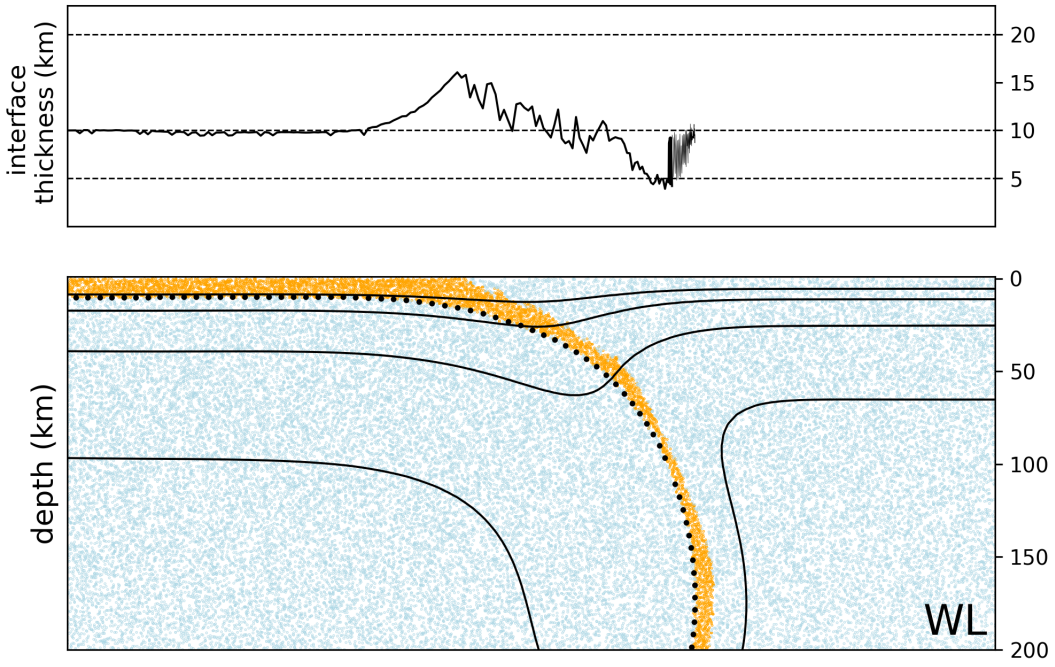

Figure 12. ~~**Convergence of models with varying resolution**~~**Interface thickness in under-resolved WL model.** ~~Vertical axis shows~~ In the ~~relative error~~ lower panel orange points show the subduction interface material, blue points are the background material (~~L_2~~mantle/lithosphere)~~of the temperature field~~. ~~We have truncated the vertical axis so as to focus on trends in the higher-resolution results~~Upper panel shows interface thickness. ~~The error value is based on the error between~~ Figure shows the ~~highest resolution~~ model ~~(192 elements~~ state 10 Myr after model initiation. This model is also shown in the ~~vertical axis) and each~~ upper-left panel of ~~the models with lower resolution, as labelled in the figure~~Fig. ~~Experiments were repeated at the highest resolution to provide~~ 11a~~baseline for model reproducibility (labelled 'repeat')~~ .

constrain the thickness of the interface improves the resolution convergence of numerical models, as well as the stability at lower resolutions. The EF does add complexity to models, both in the sense of implementation as well as the introduction of new parameters to control the specific details (e.g. $W_{\max}$). There are obviously other implementation strategies that could be developed in order to achieve similar outcomes. These should be explored in future studies. Overall, this study provides a better understanding of the behaviour of subduction models utilising WL approaches. These insights offer a basis for achieving better outcomes in terms of model reproducibility and precision.

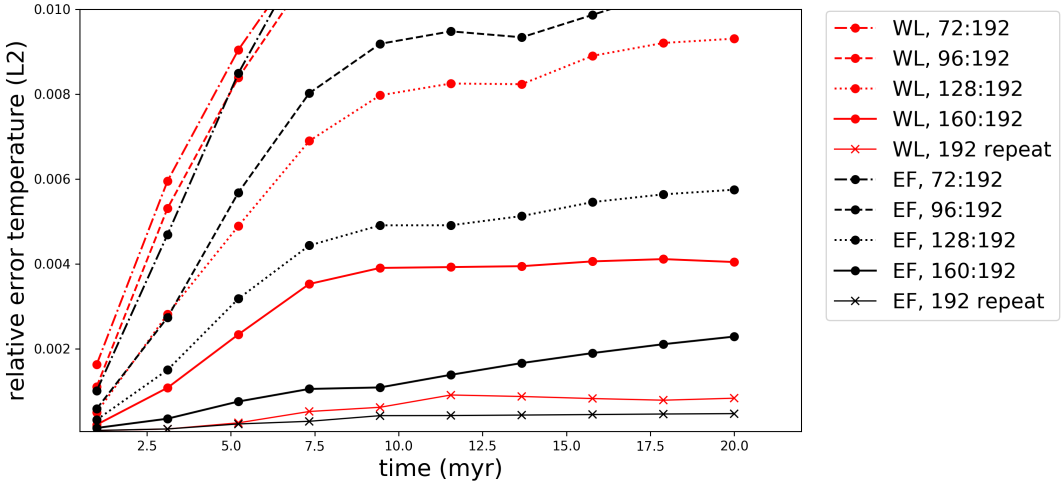

Figure 13. Convergence of models with varying resolution. Vertical axis shows the relative error (L_2) of the temperature field. We have truncated the vertical axis so as to focus on trends in the higher-resolution results. The error value is based on the error between the highest resolution model (192 elements in the vertical axis) and each of the models with lower resolution, as labelled in the figure. Experiments were repeated at the highest resolution to provide a baseline for model reproducibility (labelled 'repeat').

Appendix A: Governing equations, constitutive relationships and physical parameters

A1 Continuity, momentum and energy equations

On geological time scales the Earth's mantle behaves as a highly viscous, incompressible fluid, in which inertial forces can be neglected. The flow caused by internal buoyancy anomalies is described by the static force-balance (momentum conservation) and continuity equations:

$$\sigma_{ij,j} + \rho g_i = 0 \tag{A1}$$

$$u_{i,i} = 0 \tag{A2}$$

where u_i is the i^{th} component of the velocity. Repeated indices denote summation, and $u_{,i}$ represents partial derivative with respect to the spatial coordinate x_i. The full stress tensor appearing in Eq.A1 can be decomposed into deviatoric and mean ~~(lithostatic)~~ components:

$$\sigma_{ij} = \tau_{ij} + p\delta_{ij} \tag{A3}$$

It is noted that sign of the pressure (p) is opposite to the mean stress tensor, consistent with the convention that fluid flows from high to low pressure. The deviatoric stress tensor (τ) and the strain rate tensor (D_{ij}) are related according to the constitutive relationship:

$$\tau_{ij} = 2\eta D_{ij} = \eta(u_{i,j} + u_{j,i}) \tag{A4}$$

Substituting Eqs.A4 & A3 into Eq.A1 gives the Stokes equation, which involves two unknown variables: pressure, and velocity. The Stokes and continuity equation are sufficient to solve for the two unknowns, together with appropriate boundary conditions. An approximate solution to these equations is derived using a Galerkin Finite Element method, implemented in the ~~*Underworld 2*~~*Underworld2* code.

The thermal evolution of the system expresses the balance between heat transport by fluid motion, thermal diffusion and internal heat generation by the 1st Law of Thermodynamics, assuming incompressibility:

$$\rho C_p \frac{DT}{Dt} = q_{i,i} + \rho Q \tag{A5}$$

where T is the temperature and Q is the heat production rate (everywhere zero in this study). Diffusion rates are described by Fourier's Law, which satisfies the 2nd Law for positive conductivity (k):

$$q_i = -kT_{,i} \tag{A6}$$

Inserting Eq.A6 into Eq.A5, and using the definition of the material derivative gives:

$$\frac{\partial T}{\partial t} + u_i T_{,i} = (\kappa T_{,i})_{,i} + \frac{Q}{C_p} \tag{A7}$$

where $\kappa = \frac{k}{\rho C_p}$ is the thermal diffusivity.

The thermal variations are coupled to the momentum equation through their effect on density. At pressures in planetary interiors, silicate minerals are weakly compressible and this is generally considered as a perturbation to an incompressible flow. The Boussinesq approximation accounts for the buoyancy forces while neglecting the associated volume change allowing us to assume incompressibility (Eq. A2). In the case of density variations due to temperature, the equation of state is:

$$\rho = \rho_0(1 - \alpha(T - T_p)) \tag{A8}$$

where ρ_0 is the density at a reference temperature (here the mantle potential temperature T_p). α is the coefficient of thermal expansion. It is generally much smaller than one, making the Boussinesq approximation reasonable.

The equations and parameters that appear in the numerical models are based on equivalent dimensionless forms of the governing equations. We use the following characteristic scales (e.g Christensen, 1984):

$$\bar{x_i} = x_i \left[\frac{1}{d}\right],\ \bar{u_i} = u_i \left[\frac{d}{\kappa}\right],\ \bar{\eta} = \eta \left[\frac{1}{\eta_0}\right],\ \bar{\tau} = \tau \left[\frac{d^2}{\kappa \eta_0}\right],\ \bar{t} = t \left[\frac{\kappa}{d^2}\right],\ \bar{T} = T \left[\frac{1}{\Delta T}\right], \tag{A9}$$

where d is the mantle depth, t is time, η_0 is the reference viscosity and $\Delta T = (T_s - T_p)$, is the superadiabatic temperature difference across the fluid layer. Substituting dimensional terms for scaled dimensionless values (e.g $x \rightarrow \bar{x}d$), and rearranging allows us to write the Stokes equation as:

$$2\bar{\eta}\bar{D}_{ij,j} + \bar{p}_{,i} = Ra(1 - \bar{T})(-\delta_{iz}) \tag{A10}$$

Overbars in Eq.A10 represent dimensionless quantities, and all dimensional parameters are contained in the dimensionless ratio Ra, the Rayleigh number which can be interpreted as a ratio of advection and diffusion timescales:

$$Ra = \frac{\rho_0 g \alpha \Delta T D^3}{\eta_0 \kappa} \tag{A11}$$

The dimensionless viscosity, which has a functional dependence on the total pressure, the temperature and the second invariant of the stress tensor, are described below.

The dimensionless form of the heat conservation equation is:

$$\frac{\partial \bar{T}}{\partial \bar{t}} - \bar{u}_i \bar{T}_{,i} = (\bar{T}_{,i})_{,i} + \bar{Q} \tag{A12}$$

where the dimensionless internal heating is given by:

$$\bar{Q} = Q \left[\frac{d^2}{\kappa C_p \Delta T} \right] \tag{A13}$$

A2 Rheology

Mantle silicates deform through a range of mechanisms. The most important high-temperature creep mechanisms are diffusion creep (low stress), which results in a linear relationship between stress and strain that is strongly dependent on grain size; and dislocation creep (high stress), which leads to a power law relationship between stress and strain that is independent of grain size. In addition to high-temperature creep, some form of stress limiting behaviour is expected to occur at low temperature, and high-stress. Glide-controlled dislocation creep (or Peierls creep), which includes a stress dependence of the activation energy, is likely to play a role, particularly in the cold part of subducted slabs Karato (2012). Nevertheless, it remains unclear whether Peierls creep allows sufficient weakening, as geophysical constraints on slabs would imply (e.g. Jain et al., 2017; Krien and Fleitout, 2008; Alisic et al., 2010). In geodynamic calculations, the effect of the Peierls mechanism is similar to that of plastic, temperature-independent plasticity models (Agrusta et al., 2017; Garel et al., 2014; Čížková and Bina, 2013).

To keep the models as simple as possible ~~We~~ we include two deformation ~~mechanism only~~mechanisms: grainsize-independent diffusion creep and a plastic yielding based on a truncated Drucker-Prager plasticity model. The plastic strain rates are capable of representing brittle failure at low pressure (near the surface) and low-temperature plasticity at high pressure (deep within the slab). The value of the yield stress limit is chosen on the basis of previous studies, and is consistent with several lines of evidence suggesting that slabs to not support more than a few hundred MPa (Richards et al., 2001; Tackley, 2000; Watts and Zhong, 2000; Krien and Fleitout, 2008; Alisic et al., 2010). The viscosity associated with each deformation mechanism is

combined using a harmonic average, denoted by η_c. The entire computational domain, except for the subduction interface, is governed by the same composite rheology: there is no compositional distinction between the mantle and plates.

Ductile flow laws for silicates often have an Arrhenius temperature and pressure dependence, controlled by the activation energy E, and activation volume V (Hirth and Kohlstedt, 2004). Additional dependencies, such as grain size and melt fraction, are neglected in this study, resulting in the following diffusion viscosity:

$$\eta_d = A \exp\left(\frac{E + p_l V}{R T_a}\right) \tag{A14}$$

where p_l indicates the lithostatic component of the pressure, and A is a constant. A linearised adiabatic term is added to the dimensionless temperature field, whenever it appears in an Arrhenius law:

$$T_a = T + z \times T_{,z}$$
$$T_{,z} = \frac{-\alpha g T_p}{C_p}$$

The dimensionless form of the creep law applied in the models uses the following scalings:

$$\bar{E} = E\left[\frac{1}{R\Delta T}\right],\ \bar{W} = V\left[\frac{\rho_0 g d}{R\Delta T}\right],\ \bar{A} = A\left[\frac{1}{\eta_0}\right] \tag{A15}$$

Note that $V \rightarrow \bar{W}$ includes a change from pressure dependence (dimensional) to depth dependence (dimensionless): The dimensionless diffusion creep viscosity can be written:

$$\bar{\eta}_d = \bar{A} \exp\left(\frac{\bar{E} + \bar{z}\bar{W}}{\bar{T}_s + \bar{T}_a}\right) \tag{A16}$$

where $\bar{z}$ is the dimensionless depth and $\bar{T}_s$ is the dimensionless surface temperature. The dimensionless linearised adiabatic component is incorporated as follows:

$$\bar{T}_a = \bar{T} + \bar{z} \times \bar{T}_{,\bar{z}}$$
$$\bar{T}'_{,\bar{z}} = T_{,z}\left[\frac{d}{\Delta T}\right]$$

The parameters chosen for the diffusion creep law are consistent with those derived from experimental data on dry olivine ~~(Karato and Wu, 1993)~~(?), providing an average upper mantle viscosity close to 1×10^{20} Pas. The relatively high value of

the activation volume produces relatively low viscosity asthenosphere $\sim 0.3 \times 10^{20}$, relative to the transition zone $\sim 5 \times 10^{20}$. Additionally, a viscosity increase ($\times\eta_{660}$) is applied at the 660 km discontinuity, consistent with inferences based on the geoid (Hager and O'Connell, 1981). For ($\times\eta_{660}$ = 10), the lower mantle just beneath the 660 km is 50 times more viscous than the mean viscosity of the upper mantle. The parameters chosen produce radial viscosity profiles that are slightly higher than the 'Haskell constraint' ($\eta_{\text{mean}} = 1 \times 10^{21}$) over the upper 1400 km of the mantle (e.g. Becker, 2017).

A range of ~~psuedo-brittle~~ pseudo-brittle and plastic deformation mechanisms can be approximated in the fluid constitutive model by allowing non-linearity in the viscosity ($\eta = \eta(T, p, J_I, ...)$). The rheological model itself should be defined independently of the coordinate system, so it is necessary to define the constitutive model in terms of stress invariants (J_I). The standard viscoplastic approach (Spiegelman et al., 2016) defines an effective plastic viscosity η_p such that the deviatoric stress tensor is bounded by a yield stress τ_y:

$$\tau_y = 2\eta_p D_{ij} \tag{A17}$$

Assuming that η_p is isotropic and scalar (i.e. eigenvectors of the strain-rate tensor and deviatoric stress are identical), one can use the magnitude of both sides to define the scalar effective plastic viscosity as:

$$\eta_p = \frac{\tau_{y(\text{II})}}{2\epsilon_{\text{II}}} \tag{A18}$$

where the subscript II, denotes the square root of the tensor second invariant.

The yield stress function in the computational models is a truncated Drucker-Prager criterion:

$$\tau_y = \min(\tau_{\text{max}}, \mu p + C) \tag{A19}$$

where μ is the friction coefficient, and C is the cohesion. The Drucker-Prager yield surface is defined by the full pressure p. Because the pressure that appears in the dimensionless Stokes equation (Eq. A10) is a dynamic pressure ($\bar{p}$), due to density variations only, the lithostatic pressure (a function of vertical coordinate) needs to be accounted for. The dimensionless form of the yield stress is given by

$$\bar{\tau}_y = \min(\bar{\tau}_{\text{max}}, \bar{\mu}(\bar{p} + \bar{p}_l \bar{z}) + \bar{C}) \tag{A20}$$

where:

$$\begin{aligned} \bar{\mu} &= \mu, \\ \bar{C} &= C\left[\frac{d^2}{\kappa\eta_0}\right], \\ \bar{\tau}_{\max} &= \tau_{\max}\left[\frac{d^2}{\kappa\eta_0}\right], \\ \bar{P}_l &= \bar{z}\left[\frac{\rho_0 g d^3}{\kappa\eta_0}\right], \end{aligned} \tag{A21}$$

The effective plastic viscosity (dimensionless) is given by:

$$\bar{\eta}_p = \frac{\bar{\tau}_{y(II)}}{2\dot{\epsilon}_{II}} \tag{A22}$$

The final (composite) viscosity is the harmonic average of the viscosity associated with creep and plastic yielding:

$$\bar{\eta}_{\mathrm{c}} = \frac{\bar{\eta}_{\mathrm{d}}\bar{\eta}_p}{\bar{\eta}_{\mathrm{d}} + \bar{\eta}_p} \tag{A23}$$

A3 Model parameters and scaling values

This section provides a record of model parameters and reference values that are used in the models. The dimensional parameters quoted here are non-dimensionalised using the scaling system described in Appendix A and reference values provided in Table A1. This scaling system is identical for all models used within the thesis.

Reference value	**Value**	**Symbol**	**Units**
length	2900	d	km
viscosity	1×10^{20}	η_0	$\mathrm{Pa\,s}$
density	3300	ρ_0	$\mathrm{kg\,m^{-3}}$
thermal diffusivity	1×10^{-6}	κ	$\mathrm{m^2\,s^{-1}}$
gravity	9.8	g	$\mathrm{m\,s^{-2}}$
temperature	1400	ΔT	K
gas constant	8.314	R	$\mathrm{J\,mol^{-1}\,K^{-1}}$
Rayleigh Number	3.31×10^8	Ra	-

Table A1. Reference values used to non-dimensionalise the Stokes and Energy equations, as described in Appendix A

Parameter name	Value	Symbol	Units
domain depth	~~100~~ 1000	-	km
domain width	5000	-	km
potential temp	1673	T_p	K
surface temp	273	T_s	K
viscosity min.	1×10^{18}	-	Pa s
viscosity max.	1×10^{24}	-	Pa s
diffusion creep volume UM**	5.27×10^{-6}	V	$\mathrm{m^3\,mol^{-1}}$
diffusion creep energy UM	316	E	$\mathrm{kJ\,mol^{-1}}$
diffusion creep constant UM	1.87×10^{9}	A	$\mathrm{Pa^n\,s^1}$
diffusion creep volume LM***	1.58×10^{-6}	V	$\mathrm{m^3\,mol^{-1}}$
diffusion creep energy LM	210	E	$\mathrm{kJ\,mol^{-1}}$
diffusion creep constant LM	1.77×10^{14}	A	$\mathrm{Pa^n\,s^1}$
DP* friction coefficient	0.1	μ	-
DP cohesion	20	C	MPa
yield stress max.	200	$\tau_{\max}$	MPa
sub. interface thickness	10	W_{init}	km
sub. interface max. thickness	19	$W_{\max}$	km
sub. interface min. thickness	10	$W_{\min}$	km
sub. interface viscosity	5×10^{19}	-	Pa s
sub. interface depth taper start	100	-	km
sub. interface depth taper width	30	-	km
slab age at trench	50	-	Myr
slab radius of curv.	200	-	km
initial slab depth	150	-	km
upper plate age at trench	10	-	Myr
lower mantle viscosity increase	15	-	-
adiabatic temp. gradient	3.7×10^{-4}	-	-
internal heating	0.0	Q	$\mathrm{W.m^{-3}}$

Table A2. Dimensional model parameters: * Drucker-Prager, ** Upper Mantle, *** Lower mantle. Typical model element resolution was 800×160.

Competing interests. TEXT

The authors declare that no competing interests are present.

Acknowledgements. This research was partially funded by the Australian Government through the Australian Research Council Discovery grant DP150102887. Development of the Underworld2 code used in the simulations supported by AuScope. ~~Dan~~DS's postgraduate research at the University of Melbourne was supported by a Baragwanath Geology Research Scholarship. The study benefited from the authors' attendance of the CIDER Summer programs in 2016 & 2017 (funded by NSF grant EAR-1135452). This work was supported by resources provided by The Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. This research was supported by use of the Nectar Research Cloud, a collaborative Australian research platform supported by the National Collaborative Research Infrastructure Strategy (NCRIS).

References

Aagaard, B., Williams, C., and Knepley, M.: PyLith: A finite-element code for modeling quasi-static and dynamic crustal deformation, Eos Trans. AGU, 89, 2008.

Abers, G. A.: Seismic low-velocity layer at the top of subducting slabs: observations, predictions, and systematics, Physics of the Earth and Planetary Interiors, 149, 7–29, 2005.

Agrusta, R., Goes, S., and van Hunen, J.: Subducting-slab transition-zone interaction: Stagnation, penetration and mode switches, Earth and Planetary Science Letters, 464, 10–23, 2017.

Alisic, L., Gurnis, M., Stadler, G., Burstedde, C., Wilcox, L. C., and Ghattas, O.: Slab stress and strain rate as constraints on global mantle flow, Geophysical Research Letters, 37, https://doi.org/10.1029/2010gl045312, http://dx.doi.org/10.1029/2010gl045312, 2010.

Androvičová, A., Čížková, H., and van den Berg, A.: The effects of rheological decoupling on slab deformation in the Earths upper mantle, Studia Geophysica et Geodaetica, 57, 460–481, 2013.

Arcay, D.: Dynamics of interplate domain in subduction zones: influence of rheological parameters and subducting plate age, Solid Earth, 3, 467, 2012.

Arcay, D.: Modelling the interplate domain in thermo-mechanical simulations of subduction: Critical effects of resolution and rheology, and consequences on wet mantle melting, Physics of the Earth and Planetary Interiors, 269, 112–132, 2017.

Arnold, D. N. and Logg, A.: Periodic table of the finite elements, SIAM News, 47, 212, 2014.

Arredondo, K. M. and Billen, M. I.: Coupled Effects of Phase Transitions and Rheology in 2D Dynamical Models of Subduction, Journal of Geophysical Research: Solid Earth, 2017.

Audet, P. and Schwartz, S. Y.: Hydrologic control of forearc strength and seismicity in the Costa Rican subduction zone, Nature Geoscience, 6, 852, 2013.

Audet, P., Bostock, M. G., Christensen, N. I., and Peacock, S. M.: Seismic evidence for overpressured subducted oceanic crust and megathrust fault sealing, Nature, 457, 76–78, 2009.

Babeyko, A. and Sobolev, S.: High-resolution numerical modeling of stress distribution in visco-elasto-plastic subducting slabs, Lithos, 103, 205–216, 2008.

Bachmann, R., Oncken, O., Glodny, J., Seifert, W., Georgieva, V., and Sudo, M.: Exposed plate interface in the European Alps reveals fabric styles and gradients related to an ancient seismogenic coupling zone, Journal of Geophysical Research: Solid Earth, 114, 2009.

Bayet, L., John, T., Agard, P., Gao, J., and Li, J.-L.: Massive sediment accretion at˜ 80 km depth along the subduction interface: Evidence from the southern Chinese Tianshan, Geology, 2018.

Bebout, G. E. and Penniston-Dorland, S. C.: Fluid and mass transfer at subduction interfaces, The field metamorphic record, Lithos, 240, 228–258, 2016.

Becker, T. W.: Superweak asthenosphere in light of upper mantle seismic anisotropy, Geochemistry, Geophysics, Geosystems, 18, 1986–2003, 2017.

Behr, W. M. and Becker, T. W.: Sediment control on subduction plate speeds, Earth and Planetary Science Letters, 502, 166–173, https://doi.org/10.1016/j.epsl.2018.08.057, 2018.

Bercovici, D.: The generation of plate tectonics from mantle convection, Earth and Planetary Science Letters, 205, 107–121, 2003.

Bercovici, D. and Ricard, Y.: Plate tectonics damage and inheritance, Nature, 508, 513–516, https://doi.org/10.1038/nature13072, http://dx.doi.org/10.1038/nature13072, 2014.

Billen, M. I. and Hirth, G.: Rheologic controls on slab dynamics, Geochemistry Geophysics, Geosystems, 8, https://doi.org/10.1029/2007gc001597, http://dx.doi.org/10.1029/2007gc001597, 2007.

Billen, M. I., Gurnis, M., and Simons, M.: Multiscale dynamics of the Tonga-Kermadec subduction zone, Geophysical Journal International, 153, 359–388, 2003.

Brooks, A. N. and Hughes, T. J.: Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations, Computer methods in applied mechanics and engineering, 32, 199–259, 1982.

Buiter, S. and Ellis, S.: A discussion of numerical subduction initiation, in: EGU General Assembly Conference Abstracts, vol. 18, p. 4791, 2016.

Capitanio, F., Stegman, D., Moresi, L., and Sharples, W.: Upper plate controls on deep subduction, trench migrations and deformations at convergent margins, Tectonophysics, 483, 80–92, 2010.

Chertova, M., Geenen, T., Van Den Berg, A., and Spakman, W.: Using open sidewalls for modelling self-consistent lithosphere subduction dynamics, Solid Earth, 3, 313, 2012.

Christensen, U.: Convection with pressure-and temperature-dependent non-Newtonian rheology, Geophysical Journal International, 77, 343–384, 1984.

Christensen, U. R.: The influence of trench migration on slab penetration into the lower mantle, Earth and Planetary Science Letters, 140, 27–39, 1996.

Čížková, H. and Bina, C. R.: Effects of mantle and subduction-interface rheologies on slab stagnation and trench rollback, Earth and Planetary Science Letters, 379, 95–103, 2013.

Čížková, H., van Hunen, J., van den Berg, A. P., and Vlaar, N. J.: The influence of rheological weakening and yield stress on the interaction of slabs with the 670 km discontinuity, Earth and Planetary Science Letters, 199, 447–457, 2002.

Cloos, M. and Shreve, R. L.: Subduction-channel model of prism accretion, melange formation, sediment subduction, and subduction erosion at convergent plate margins: 1. Background and description, Pure and Applied Geophysics, 128, 455–500, 1988.

Duarte, J. C., Schellart, W. P., and Cruden, A. R.: How weak is the subduction zone interface?, Geophysical Research Letters, 42, 2664–2673, 2015.

Duretz, T., Schmalholz, S. M., and Gerya, T. V.: Dynamics of slab detachment, Geochemistry Geophysics, Geosystems, 13, https://doi.org/10.1029/2011gc004024, http://dx.doi.org/10.1029/2011gc004024, 2012.

Gao, X. and Wang, K.: Strength of stick-slip and creeping subduction megathrusts from heat flow observations, Science, 345, 1038–1041, 2014.

Garel, F., Goes, S., Davies, D. R., Davies, J. H., Kramer, S. C., and Wilson, C. R.: Interaction of subducted slabs with the mantle transition-zone: A regime diagram from 2-D thermo-mechanical models with a mobile trench and an overriding plate, Geochemistry Geophysics, Geosystems, 15, 1739–1765, https://doi.org/10.1002/2014gc005257, http://dx.doi.org/10.1002/2014gc005257, 2014.

Gerardi, G. and Ribe, N. M.: Boundary-element modeling of two-plate interaction at subduction zones: scaling laws and application to the Aleutian subduction zone, Journal of Geophysical Research: Solid Earth, 2018.

Gerya, T. V., Stöckhert, B., and Perchuk, A. L.: Exhumation of high-pressure metamorphic rocks in a subduction channel: A numerical simulation, Tectonics, 21, 2002.

Gerya, T. V., Connolly, J. A., and Yuen, D. A.: Why is terrestrial subduction one-sided?, Geology, 36, 43–46, 2008.

Glerum, A., Thieulot, C., Fraters, M., Blom, C., and Spakman, W.: Implementing nonlinear viscoplasticity in ASPECT: benchmarking and applications to 3D subduction modeling, Solid Earth Discussions, 2017, 1–47, https://doi.org/10.5194/se-2017-9, https://www.solid-earth-discuss.net/se-2017-9/, 2017.

Gurnis, M. and Hager, B. H.: Controls of the structure of subducted slabs, Nature, 335, 317–321, 1988.

Hager, B. H. and O'Connell, R. J.: A simple global model of plate dynamics and mantle convection, Journal of Geophysical Research, 86, 4843, https://doi.org/10.1029/jb086ib06p04843, http://dx.doi.org/10.1029/jb086ib06p04843, 1981.

Hardebeck, J. L.: Stress orientations in subduction zones and the strength of subduction megathrust faults, Science, 349, 1213–1216, 2015.

Hardebeck, J. L. and Loveless, J. P.: Creeping subduction zones are weaker than locked subduction zones, Nature Geoscience, 11, 60, 2018.

Hilairet, N. and Reynard, B.: Stability and dynamics of serpentinite layer in subduction zone, Tectonophysics, 465, 24–29, 2009.

Hirauchi, K.-i. and Katayama, I.: Rheological contrast between serpentine species and implications for slab–mantle wedge decoupling, Tectonophysics, 608, 545–551, 2013.

Hirth, G. and Kohlstedt, D.: Rheology of the upper mantle and the mantle wedge: A view from the experimentalists, Inside the subduction Factory, 138, 83–105, 2004.

Holt, A. F., Buffett, B. A., and Becker, T. W.: Overriding plate thickness control on subducting plate curvature, Geophysical Research Letters, 42, 3802–3810, 2015.

Huene, R. and Scholl, D. W.: Observations at convergent margins concerning sediment subduction, subduction erosion, and the growth of continental crust, Reviews of Geophysics, 29, 279–316, 1991.

Jain, C., Korenaga, J., and Karato, S.-i.: On the yield strength of oceanic lithosphere, Geophysical Research Letters, 44, 9716–9722, 2017.

Karato, S.-i.: Deformation of earth materials: an introduction to the rheology of solid Earth, 2012.

Karato, S.-i. and Wu, P.: Rheology of the upper mantle: A synthesis, Science, 260, 771–778, 1993.

Kimura, G., Yamaguchi, A., Hojo, M., Kitamura, Y., Kameda, J., Ujiie, K., Hamada, Y., Hamahashi, M., and Hina, S.: Tectonic mélange as fault rock of subduction plate boundary, Tectonophysics, 568, 25–38, 2012.

Kincaid, C. and Sacks, I. S.: Thermal and dynamical evolution of the upper mantle in subduction zones, Journal of Geophysical Research: Solid Earth, 102, 12 295–12 315, 1997.

Krien, Y. and Fleitout, L.: Gravity above subduction zones and forces controlling plate motions, Journal of Geophysical Research: Solid Earth, 113, 2008.

Lallemand, S.: High rates of arc consumption by subduction processes: Some consequences, Geology, 23, 551–554, 1995.

Lamb, S.: Shear stresses on megathrusts: Implications for mountain building behind subduction zones, Journal of Geophysical Research: Solid Earth, 111, 2006.

Lenardic, A. and Kaula, W.: Self-lubricated mantle convection: Two-dimensional models, Geophysical research letters, 21, 1707–1710, 1994.

Li, B. and Ghosh, A.: Near-continuous tremor and low frequency earthquake (LFE) activities in the Alaska-Aleutian subduction zone revealed by a mini seismic array, Geophysical Research Letters, 2017.

Magni, V. v., Van Hunen, J., Funiciello, F., and Faccenna, C.: Numerical models of slab migration in continental collision zones, Solid Earth, 3, 293, 2012.

Manea, V. and Gurnis, M.: Subduction zone evolution and low viscosity wedges and channels, Earth and Planetary Science Letters, 264, 22–45, 2007.

Moresi, L. and Solomatov, V.: Mantle convection with a brittle lithosphere: thoughts on the global tectonic styles of the Earth and Venus, Geophys. J. Int., 133, 669–682, https://doi.org/10.1046/j.1365-246x.1998.00521.x, http://dx.doi.org/10.1046/j.1365-246x.1998.00521.x, 1998.

Peacock, S. M.: Thermal and petrologic structure of subduction zones, Subduction top to bottom, pp. 119–133, 1996.

Plank, T. and Langmuir, C. H.: Tracing trace elements from sediment input to volcanic output at subduction zones, Nature, 362, 739–743, 1993.

Proctor, B. and Hirth, G.: Ductile to brittle transition in thermally stable antigorite gouge at mantle pressures, Journal of Geophysical Research: Solid Earth, 121, 1652–1663, 2016.

Reynard, B.: Serpentine in active subduction zones, Lithos, 178, 171–185, 2013.

Richards, M. A., Yang, W.-S., Baumgardner, J. R., and Bunge, H.-P.: Role of a low-viscosity zone in stabilizing plate tectonics: Implications for comparative terrestrial planetology, Geochemistry, Geophysics, Geosystems, 2, 2001.

Schellart, W. and Rawlinson, N.: Global correlations between maximum magnitudes of subduction zone interface thrust earthquakes and physical parameters of subduction zones, Physics of the Earth and Planetary Interiors, 225, 41–67, 2013.

Schmidt, M. W. and Poli, S.: Experimentally based water budgets for dehydrating slabs and consequences for arc magma generation, Earth and Planetary Science Letters, 163, 361–379, 1998.

Shankar, V. and Kumar, L.: Stability of two-layer Newtonian plane Couette flow past a deformable solid layer, Physics of Fluids, 16, 4426–4442, 2004.

Shreve, R. L. and Cloos, M.: Dynamics of sediment subduction, melange formation, and prism accretion, Journal of Geophysical Research: Solid Earth, 91, 10 229–10 245, 1986.

Spiegelman, M., May, D. A., and Wilson, C. R.: On the solvability of incompressible Stokes with viscoplastic rheologies in geodynamics, Geochemistry, Geophysics, Geosystems, 17, 2213–2238, 2016.

Tackley, P. J.: Self-consistent generation of tectonic plates in time-dependent, three-dimensional mantle convection simulations, Geochemistry, Geophysics, Geosystems, 1, 2000.

Tagawa, M., Nakakuki, T., Kameyama, M., and Tajima, F.: The role of history-dependent rheology in plate boundary lubrication for generating one-sided subduction, Pure and Applied Geophysics, 164, 879–907, 2007.

Tichelaar, B. W. and Ruff, L. J.: Depth of seismic coupling along subduction zones, Journal of Geophysical Research: Solid Earth, 98, 2017–2037, 1993.

Trompert, R. and Hansen, U.: Mantle convection simulations with rheologies that generate plate-like behaviour, Nature, 395, 686, 1998.

Ulmer, P., Trommsdorff, V., et al.: Serpentine stability to mantle depths and subduction-related magmatism, Science, pp. 858–858, 1995.

van Keken, P. E., Currie, C., King, S. D., Behn, M. D., Cagnioncle, A., He, J., Katz, R. F., Lin, S.-C., Parmentier, E. M., Spiegelman, M., et al.: A community benchmark for subduction zone modeling, Physics of the Earth and Planetary Interiors, 171, 187–197, 2008.

Vannucchi, P., Remitti, F., and Bettelli, G.: Geological record of fluid flow and seismogenesis along an erosive subducting plate boundary, Nature, 451, 699, 2008.

Vannucchi, P., Sage, F., Phipps Morgan, J., Remitti, F., and Collot, J.-Y.: Toward a dynamic concept of the subduction channel at erosive convergent margins with implications for interplate material transfer, Geochemistry, Geophysics, Geosystems, 13, 2012.

Vrolijk, P.: On the mechanical role of smectite in subduction zones, Geology, 18, 703–707, 1990.

Wada, I. and Wang, K.: Common depth of slab-mantle decoupling: Reconciling diversity and uniformity of subduction zones, Geochemistry, Geophysics, Geosystems, 10, 2009.

Wang, K.: Finding fault in fault zones, Science, 329, 152–153, 2010.

Watts, A. and Zhong, S.: Observations of flexure and the rheology of oceanic lithosphere, Geophysical Journal International, 142, 855–875, 2000.

Zhong, S. and Gurnis, M.: Viscous flow model of a subduction zone with a faulted lithosphere: Long and short wavelength topography gravity and geoid, Geophysical Research Letters, 19, 1891–1894, https://doi.org/10.1029/92gl02142, http://dx.doi.org/10.1029/92gl02142, 1992.

Zhong, S. and Gurnis, M.: Mantle Convection with Plates and Mobile Faulted Plate Margins, Science, 267, 838–843, https://doi.org/10.1126/science.267.5199.838, http://dx.doi.org/10.1126/science.267.5199.838, 1995.