We gratefully thank the reviewer for the very constructive criticism. Implementing the suggestions helped to better visualize the simulation results, focus on the main findings and significantly improve the manuscript. During the review process, we have changed the structure of the manuscript significantly. In the results section of the revised manuscript, we present the evolution of the reference model and the wet olivine model separately. The results of the remaining models are presented in comparison to the results of the reference run for the distinct deformation stages. We then discuss the implications of our findings on several aspects, such as for example the impact of the viscosity structure on the convection, the onset of convection and the impact of convection on subduction, in the discussion section. The order of the figures and the style of visualization has been adapted accordingly.

Below we have answered to all the comments from the reviewer. Our answers are highlighted in blue and the original comments are highlighted in red.

- the fluid material is described as incompressible in section 2.1, but the thermal gradient is adiabatic, mantle densities vary greatly along the domain high due to phase transitions (Fig. A1) and the authors use the extended Boussinesq approximation for compressible fluid to solve the conservation equations (Appendix B line 579)

> The maximum value for the density time derivative is two orders of magnitude smaller compared to the velocity divergence. Also, (Bercovici, Schubert , & Glatzmaier, 1992) concluded that compressibility effects on the spatial mantle structure are minor when the superadiabatic temperature drop is close to the adiabatic temperature of the mantle, which is the case for the Earth. We therefore assume here that the Boussinesq approximation is still valid and suggest that density changes due to volumetric deformation are negligible. We consider only small density changes affecting the buoyancy stresses. Not considering adiabatic heating in the energy conservation equation leads to a significant deviation of the thermal structure from the initially imposed adiabatic temperature gradient over large time scales (>100 Myrs). The resulting vertical temperature profile (if adiabatic heating is neglected) is constant throughout the upper mantle and the newly equilibrated vertically-constant temperature is equal to the imposed temperature at the bottom boundary. In consequence, the density structure taken from the phase diagram table according to pressure and temperature values is wrong. By using the extended Boussinesq approximation, i.e., the adiabatic heating term is included in the energy conservation equation but not in the continuity equation, the initially imposed adiabatic (or isentropic) gradient is maintained over long time scales. The resulting density structure agrees well with the PREM model as shown in this study. A more detailed comparison between different approximations of the continuity equation is beyond the scope of this study.

- phase transitions are implemented through the variable density (Fig. A1), but latent heat associated to phase change is missing from equation (A14), and this approximation should be justified.

Indeed, latent heat released or consumed by a phase transition can perturb the thermal field by up to 100 K and induce a buoyancy force aiding or inhibiting the motion of cold subducting slabs (van Hunen, van den Berg, & Vlaar, 2001) or hot rising plumes. However, when the lateral differences in temperature do not vary much, the deflection of the phase transition by an ascending plume or a subducting slab has a much bigger impact on the buoyancy stresses than the latent heat released or consumed by the phase transition (Christensen , 1995). Also, most of the studies on the impact of latent heat rely on the assumption that density is temperature dependent only. In the models presented here, density is a function of temperature and pressure, which makes it difficult to estimate the impact of temperature changes due to latent heat near a phase transition on the buoyancy term a priori. Because a detailed parametric investigation of the impact of latent heat on buoyancy stresses for temperature and pressure dependent density is beyond the scope of study, we neglect latent heat for simplicity.

- the side velocity boundary conditions during the extension or convergence phase (Fig. 1a,d) are likely to induce a sheared weak zone near the side boundaries at the transition depth between lateral inflow and outflow (340 km depth). Also, for the extension set-up, the suction created by the divergence in uppermost mantle is likely to generate a bulk ascending mantle channel in the middle of the domain (X=0 km).
- more generally, the flow pattern over the entire model domain is never shown when in/out flows are imposed at the sides, and this is an issue when discussing application to Earth (section 4.4): what is the geological justification that the divergence or convergence was not only active at the plates 'surface but also across 300 km in the mantle below the plates? The justification of the side velocity boundary conditions should be developed, and the global flow pattern (velocity glyphs/arrow) should be shown during extension and convergence phases.

I also have other issues with the methods and results analysis:

As mentioned correctly by the reviewer, material inflow/outflow velocity boundary conditions may lead to a shear zone forming at the transition point between inflow and outflow. To avoid such boundary condition effects close to the mechanical lithosphere, we set the transition point deeper than the initially imposed lithospheric thickness. We chose z=-330 km as the point of transition, because values for deviatoric stresses at this depth are significantly smaller compared to those at the base of the lithosphere. The mechanical thickness of the lithosphere may vary greatly over time mainly due to temperature changes and is therefore difficult to constrain a priori. By setting the transition point of the velocity inflow/outflow boundary condition deeper than the initially imposed mechanical thickness of the lithosphere, we allow for self-consistent adjustment of the mechanical thickness of the lithosphere during the evolution of the model. We have added velocity arrows to many of our revised figures. These figures show that the thickness of the lithosphere, with respect to consistent horizontal velocities, is not controlled by the inflow/outflow boundary away from the model boundaries.

- the choice of which input parameter are varied is not explained: why choose to vary the minimum viscosity between models M1, M3 and M5, rather than for example the initial extension rate, the duration of the thermal relaxation phase or the inflow/outflow side velocity profiles? The discussion does not well explain why there is a single subduction in M1, but double subductions in M3/M5.

> One aim of this study is to quantify the impact of convection in the upper mantle on the long-term extension—cooling—convergence cycle of the lithosphere. Convection is controlled by the Rayleigh-number, which is a function of the viscosity of the convecting layer. Since the viscosity of the mantle is poorly constrained and estimated values vary by two orders of magnitude, this parameter is the most interesting for us to investigate. A frequently used technique to parameterize the impact of convection is the effective conductivity approach. This way the temperature field can be stabilized without having to calculate an enhanced convective velocity field in the mantle (which can be numerically challenging). We therefore compare this approach to the explicitly modelled convection, which calculates an enhanced convective velocity field in the mantle. The initial extension rate has been chosen to model the formation of an approximately 400 km wide basin containing exhumed mantle in an ultra-slow to slow spreading environment. The timing of the extension, cooling and convergence phases are motivated by data from the Alpine orogeny, which is now clearer explained in the revised manuscript. Testing the impact of different extension and compression rates as well as different cooling durations is beyond the scope of this study.

- a methodology study on the comparison of "explicitly modelled convection" and "effective conductivity mimicking a convective heat flow" is inserted in the middle of the main geodynamics study. This hinders the continuous read of the paper, and I suggest all analysis and related figures of models M2/M4 are moved to Appendix B, along with the heat flow profiles of Figure 10.

> The manuscript and result presentation have been re-structured based on the constructive comments of both reviewers. However, we keep the results of the simulations with an “effective conductivity” in the main manuscript, because these results are part of our main geodynamics study.

- the simulations have numerous features that do not seem relevant for the scientific question (erosion, sedimentations of alternating calcites and pelites), but add yet another set of free parameters that make the interpretation of the simulations more complex.

> We include erosion and sedimentation to avoid too high or low topography. To achieve this, we decided on one particular erosion/sedimentation model. A parametric study on the impact of different sedimentation/erosion processes or

sediment transport mechanisms on the deformation of the lithosphere is beyond the scope of this study.

- despite its central importance, the paper lacks a clear definition of "convection", that sometimes means "advection" or "drag" or "flow".

The word "convection" can be used to describe the motion of any fluid. Convection in a fluid can either develop freely, via thermal or compositional variations, or it can be induced by external forces (Ricard, 2007). In mantle convection simulations, free material motion in the mantle can be initiated by for example buoyancy contrasts due to variations in temperature or chemical composition. A geological example for induced mantle flow is a rigid plate that moves on top of the mantle. These statements have been incorporated to the introduction during the review process.

The manuscript could also maybe reference the following papers dealing with the plate-asthenosphere interactions or subduction initiation or various scales of mantle convection:
L. Husson. The dynamics of plate boundaries over a convecting mantle. Physics of the Earth and Planetary Interiors, Elsevier, 2012, 212-213, pp.32-43.
V. S. Solomatov. Initiation of subduction by small-scale convection. Journal of Geophysical Research: Solid Earth, 2004.
F. Lévy, C. Jaupart. The initiation of subduction by crustal extension at a continental margin. Geophysical Journal International, Volume 188, Issue 3, March 2012, Pages 779–797.
N. Coltice et al. Interactions of scales of convection in the Earth's mantle. Tectono-physics. Volume 746, 30 October 2018, Pages 669-677

We have added some of the references to the revised version of the manuscript.

Presentation quality: poor

This is a major flaw of the manuscript, which scientific contributions are hard to unearth because of confusing text and figure organization.
For example :
- showing vertical and horizontal velocity background colors (Fig. 4, Fig. 9) makes it difficult for the reader to visualize the flow pattern > could the authors show velocity glyphs or arrows, to better reveal e.g. the wavelength of small-scale convection in Fig. 4

We have added the velocity arrows to the figures. This was a very constructive comment for better visualization – thank you.

- Figure 5 should be referenced in the methods section since achieving realistic temperature, density and viscosity model output is rather a constrain on the input parameters than a surprising result

This has been done during the review process.

- Figure 8 and 10 are referred to very early in the text, whereas they belong in the discussion (or appendix?) rather than in the results section same for text on lines 187-199,156-164

The figure order has been changed in the review process.

- Appendix A belongs to the main text, otherwise the parameters of Table 2 are not defined

We have decided to shift the parameter table 2 to the appendix rather than describing all equations in the main text.

- a time-bar could be included in Fig. 1 showing to scale the 3 stages of boundary conditions with colours corresponding to the velocity profiles shown in Figure 1a,d (also please add a null-velocity profile fo the thermal relaxation). The same time bar could then be put on other figure to know at a glance which stage the figures belong to.

We have tried to add the time bar, but the figures become too busy. We have changed the order of the figures and implemented much of the constructive comments.

SPECIFIC COMMENTS

- in the introduction, the authors should define what they mean by "convection" and discuss the different scales

This suggestion has been implemented in the revised version of the manuscript.

- the exhumation of hot mantle (Fig. 2) is expected to lead to melting, please comment

This is explained in the introduction line 70 ff. and also in the methods section.

- you need to support some statements with results/data, i.e. "Alternating activity of the subduction zones is observed." (line 219)

This has been done in the review process.

- if you mention the importance of apply a force rather than a velocity BC, then you should also mention the importance of setting the lateral flow in/out of the domain

The timing of the deformation periods is chosen to allow for comparison with orogenies such as, f.e. the Alps. By choosing inflow/outflow we simulate the movement of the plates, decoupled from the rest of the domain, which we

consider here more realistic than extending or compressing the entire side walls of the model domain with a height of 660 km.

- I am not sure you can compare your model to the Atlantic (line 272) since old oceanic lithosphere there is much older than in your models

… Even that old oceanic crust did not undergo spontaneous subduction yet, which was the point to mention it as an example here.

- I disagree with the statement "the models are in a state of isostatic equilibrium at the onset of subduction initiation." (line 278): the convergence velocity and the topographic low above the new trench (Fig. 8) suggest a dynamic topography

Indeed, this was our mistake: we chose the wrong time step to show the topography at the end of the cooling period. This has been updated during the revision. In general, it is also true that the system cannot be in isostatic equilibrium by definition, since there are deviatoric stresses holding the topography of the passive margins. However, the difference between the height of margins and the depth of the basin is ca. 5 km, which is the calculated topographic difference for an idealised block of 30 km thick crust floating on top of the mantle (Turcotte & Schubert, 2014). We therefore argue that the topography across the passive margin system produced by our model is close to isostatic equilibrium at the end of the cooling period.

- line 294-307: the explanation of the control of single-sided subduction is not clear

The down-welling of two convecting cells meet directly below the margin at which subduction is going to be initiated later. The enhanced downward motion of upper-mantle material in this region likely exerts a suction force that assists in initiating and stabilising only one single-slab subduction.

- you cannot claim that "mantle convection seems active and largely confined to the upper region of the upper mantle. The convective patterns simulated in our study are in agreement with these observations." (line 399) since you impose the height of your simulation domain to be restricted to the upper mantle.

We want to highlight here that we model a convective pattern that is observed in nature, namely convection in the upper mantle, that is below the lithosphere and above 660 km. Other people might argue that such “upper-mantle convection” does not exist and only “one-layered convection” of the entire mantle, down to ca. 2900 km, exists in nature. We confined our model domain using the justification that in the Alps the convection seems to be two-layered. Besides that, all models are confined to the upper mantle, but the reach of the convection cells extends to different depth, depending on the Rayleigh-number (compare M1 to M3). This is now more obvious when visualising the flow pattern with velocity vectors.

- you should not boast that "the model has captured correctly the first order physics of the investigated processes." since model M6 shows the immense importance of rheology parameterisation - that is far from being constrained...

> As mentioned, we calibrated our initial model configurations in such a way that they match data from the PREM model and from GIA estimates. Starting from this point, we let the model freely evolve and do not change material parameters or geometries anymore. During rifting we generate margins of realistic first-order geometry, during cooling, the basin subsides to realistic depths and convection has realistic Rayleigh numbers, and finally subduction is initiated self-consistently (i.e. without imposing any major triangular weak zone cutting through the entire lithosphere at the OCT) via thermal softening. Therefore, we argue that model M1 captures the first order physics of the extension-cooling-subduction cycle correctly. M6 shows a scenario in which the initial viscosity profile is not backed-up by natural data. This model becomes unrealistic and is not applicable to the present-day Earth after the rifting phase. Hence, this model does not capture the first-order physics correctly, because Ra-numbers are too high and the lithospheric thickness becomes too thin. We therefore show end-member models that either capture the first order physics of the present-day mantle correctly, or not.

- what do you mean by "If convection in the mantle is suppressed by high effective thermal conductivities or high, lower viscosity limits" (line 453)? During the convergence phases, the mantle still flows in the domain (which is why you should show the velocity glyphs).

> We want to say that the vigor of convection is significantly reduced, for example, absolute magnitudes of convection velocities are significantly reduced. We modified the text accordingly.

- did you try to run models without shear heating to estimate the relative role of structural vs. thermal softening for localization? (lines460-462)

> We did not run models without shear heating. (Jaquet & Schmalholz, 2018) and (Kiss, Candioti, Duretz, & Schmalholz, 2020) investigated the importance of shear heating for shear zone formation and subduction initiation. They showed that in absence of any other active weakening mechanism (f.e. brittle-plastic strain weakening) the deactivation of thermal softening results in large scale folding of the crust without localisation of a major shear zone. The geometry at the onset of convergence in the models presented here is similar to the initial geometry of (Kiss, Candioti, Duretz, & Schmalholz, 2020). We, therefore, expect similar behaviour for our model. Since we do not employ any other strain weakening mechanism in the models presented here, we rely on thermal softening for localisation of shear zones.

- Figure 9 and Figure 2a,b: comment on the "slab-like" features between 100 and 200 km depth below the extended margins in M1 and M2

These features result from the convection cells that are already active.

- Fig.8d: what is the X-locations and the depth range for integration of the second invariant of deviatoric stress tensor?

The second invariant of the deviatoric stress tensor is integrated vertically over the entire domain. To estimate the plate driving forces we first average the vertically integrated second invariant of the deviatoric stress tensor horizontally over the left most and right most 100 km. Second, the estimated value for the plate driving force is computed as the average of the two average values obtained before. We have clarified this in the revised version of the manuscript.

- explain in caption of Fig. 8 "values for tau_II remain constant when no deformation is applied to the system" whereas Fig. 4 shows large convection cells un the mantle that may deform the plates above

We wanted to say: “when no far-field inflow/outflow deformation is applied”; we clarified the text. Values for tau_II are not equal to 0, because there are always deviatoric stresses, f.e. to sustain the margin geometry or shear forces induced by mantle flow at the bottom of the more-or-less rigid plates. However, there is no significant deviation from these“ background” stresses when no far-field deformation is applied to the system. Thus, the values for tau_II remain relatively constant during the cooling period.

- appendix B: it is not clear why D should be thickness of the whole upper mantle whereas Fig. 4a,f shows small convection cells

This paragraph has been rephrased for clarity during the review process.

- equation B1: how is effective viscosity average over the domain?

In our study, we calculate a local Rayleigh-Number on each grid point. The average Rayleigh number is calculated as the arithmetic average of local Rayleigh numbers >1000.

- the explanation on lines 554-561 is not convincing: what take a constant Rayleigh number that on D and on k and then claim that D and k can be adjusted?

The goal of this exercise is to match a realistic Nusselt number for the Earth’s mantle by assuming a conductive heat flow through the upper mantle. This means that q_LAB is parameterized via an enhanced, conductive heat flux. To match the Nusselt number of the Earth’s mantle, this artificial heat flow has to be

13x larger than the realistic conductive heat flow of the mantle. We therefore enhance the thermal conductivity in the upper mantle by a factor 13.

- the isentrop is Fig. A1 does not match the temperature profile in Fig. 1 TECHNICAL CORRECTIONS

This has been corrected in the revised version of the manuscript.

- "cooling" is more appropriate than "thermal relaxation" for stage 2

This has been changed in the revised version of the manuscript.

- the initial velocity condition is not given

This been corrected in Fig. 1 in the revised version of the manuscript

- do you have more references for the "common approach" to indirectly include the effects of thermal convection ? (line 48)

They are given line 52.

- line 97-98 : what does "free slip with constant material inflow/outflow velocities" mean?
- 7 units of 5 km each make a thickness of 35 km (not 33)(line 104)

This mistake has been corrected in the revised version of the manuscript. We have also clarified the description of the initial configuration.

- line 104 : why describe a 87-km thick mantle lithosphere if all parameters are the same (line 114)

To indicate the initial depth of the LAB. We have explained it in more detail during the review process.

- Table 1 : please highlight (bold ?) which parameters differ from model M1 for all models.

This suggestion has been implemented in the revised version of the manuscript.

- Table 2 : how are the column of the 2 sediments different? link with pelites/calcites or with sediments 1/2 of Figure 6?

This has been changed in the figure legends.

- Table 2 : why no diffusion creep in the crust ?

At low temperatures, the strain rate is a nonlinear function of the stress, which suggests that the active deformation mechanism of the crust is likely dominated by dislocation creep. Diffusion creep is usually active at high temperatures and low deviatoric stresses and is therefore more important in the upper mantle.

- Table 2 : which rock are analogue for strong and weak crust?

  The strength of crustal rocks depends on temperature, pressure and deformation rates. In the models presented here, the strength of the weak and strong layers should be regarded relative to each other. They represent a more heterogeneous crust, which is more realistic than a unified homogenous material for the entire crust. The weak layers represent for example silica-rich metasediments and the strong units represent for example mafic material (see also (Petri, et al., 2019) for more details).

- section 3.1.1: how do you define the length of the margin (threshold in crust thickness?)

  Thickness reduction from original thickness to <10 km following (Sutra & Manatschal, 2012).

- line 160: is the second invariant tensor of the deviatoric stress calculated for the whole lithosphere including the crust?]

  Yes, we clarified this in the revised version of the manuscript.

- why do you take the 10ˆ21 Pa.s contour as the base of the lithosphere? Why not take th 1350◦C isotherm?

  The viscosity contour remains horizontally straight, whereas the 1350 °C isotherm is deflected by the convection cells indicating mantle material flow rather than a rigid plate boundary. Above the viscosity contour the length of the velocity vectors is essentially zero, but below this contour line the convection cells are active. We have clarified this in the revised version of the manuscript.

- line 176: give X-location of special flow field at 120 km depth

  This has been clarified in the revised version of the manuscript.

- line 184: why do you claim that the lithosphere is delaminating whereas the isoviscous contour is almost flat?

  This has been addressed in the revised version of the manuscript.

- line 207: define GPE

A definition of GPE is given in line 197. A more detailed explanation is given in the appendix.

- line 223: Figure 8d rather than 8b?

  Yes, this has been changed in the revised version of the manuscript.

- rephrase "In our models, subduction is initiated self-consistently, without prescribing any major weak zone or an already existing slab." (line 286) since they are weak heterogeneities in the passive margin

  We argue that in our models, subduction is initiated “self-consistently”, because we do not ad-hoc prescribe any major weak zone cutting through the entire lithosphere at the passive margin. Also, in our initial model configuration, we do not impose any major weak zones, or seeds, to force mantle exhumation and separation of the continental crust. The heterogeneities at the passive margin have been modelled self-consistently within the same continuous numerical simulation. Also, the layered heterogeneities are only present in the crust and no major heterogeneities are imposed in the mantle lithosphere. We modified the text to clarify our statement.

- line 319 "if shear stresses are negligible" = is that really the case at subduction onset?

  Horizontally far away from the subduction zone, which is where we calculate the force, this assumption is valid.

- equation A3: define alpha and beta (which is different from the beta in Eq. B2 I guess...)

  This has been corrected in the revised version of the manuscript.

- Fig. 1a,d: the depth looks smaller than 680 km

  Indeed, this is not clear enough in the model configuration: depth is -660 km, surface level is 0 km and +20 km are left free to allow for topography. This has been addressed during the review process.

- Fig. 1c: initial random perturbations look denser between -20 and + 20 km, is that the case?

  This has also been clarified during the review process.

- Fig. 3d: issue with the bottom of the plit nera -200 km (vertical grey line?)

  I do not see the vertical grey line. The comment is not clear enough.

- Fig. 5: dashed lines for M4 and M5 are barely visible, I suggest you use thick lines with other colours

  We have changed the line style and colour for the figure in the revised version of the manuscript.

- Fig. 8: it would be helpful to have label on the topography such as "trench", and to mark the subduction initiation in the timeline of Figure 8d

  This suggestion has been implemented in the revised version of the manuscript.

- Fig. 10: what is the new information brought by this figure compared to Fig. 4?

  Figure 10 shows the conductive heat flow of the entire domain. Whenever convection is modelled properly, the conductive heat flow in the upper mantle must be close to 0. In models M4, M5 the conductive heat flow is still high due to the enhanced thermal conductivity. This illustrates, that this approach mimics the convective thermal structure but does not capture the physical process of convection in the upper mantle correctly. We have combined figure 10 with figure 4 in the review process.

Bibliography

Bercovici, D., Schubert , G., & Glatzmaier, G. A. (1992). Three-dimensional convection of an infinite-Prandtl-number compressible fluid in a basally heated spherical shell. *Journal of Fluid Mechanics, 239*, 683-719.

Christensen , U. (1995). Effects of phase transitions on mantle convection. *Annual Review of Earth and Planetary Sciences, 23*(1), 65-87.

Gülcher, A. J., Beaussier, S. J., & Gerya, T. V. (2019). On the formation of oceanic detachment faults and their influence on intra-oceanic subduction initiation: 3D thermomechanical modeling. *Earth and Planetary Science Letters, 506*, 195-208.

Jaquet, Y., & Schmalholz, S. M. (2018). Spontaneous ductile crustal shear zone formation by thermal softening and related stress, temperature and strain rate evolution. *Tectonophysics, 746*, 384-397.

Kiss, D., Candioti, L. G., Duretz, T., & Schmalholz, S. M. (2020). Thermal softening induced subduction initiation at a passive margin. *Geophysical Journal International, 220*(3), 2068-2073.

Petri, B., Duretz, T., Mohn, G., Schmalholz, S. M., Karner, G. D., & Müntener, O. (2019). Thinning mechanisms of heterogeneous continental lithosphere. *Earth and Planetary Science Letters, 512*, 147-162.

Ricard, Y. (2007). Physics of mantle convection. *Treatise on Geophysics*, 31-88.

Sutra, E., & Manatschal, G. (2012). How does the continental crust thin in a hyperextended rifted margin? Insights from the Iberia margin. *Geology, 40*(2), 139-142.

Tosi, N., Stein, C., Noack, L., Hüttig, C., Maierova, P., Samuel, H., . . . Glerum, A. (2015). A community benchmark for viscoplastic thermal convection in a 2-D square box. *Geochemistry, Geophysics, Geosystems, 16*(7), 2175-2196.

Turcotte, D. L., & Schubert, G. (2014). *Geodynamics.* Cambridge Unversity Press.

van Hunen, J., van den Berg, A. P., & Vlaar, N. J. (2001). Latent heat effects of the major mantle phase transitions on low-angle subduction. *Earth and Planetary Science Letters, 190*((3-4)), 125-135.

We gratefully thank the reviewer for the very constructive criticism. Implementing the suggestions helped to better visualize the simulation results, focus on the main findings and significantly improve the manuscript. During the review process, we have changed the structure of the manuscript significantly. In the results section of the revised manuscript, we present the evolution of the reference model and the wet olivine model separately. The results of the remaining models are presented in comparison to the results of the reference run for the distinct deformation stages. We then discuss the implications of our findings on several aspects, such as for example the impact of the viscosity structure on the convection, the onset of convection and the impact of convection on subduction, in the discussion section. The order of the figures and the style of visualization has been adapted accordingly.

Below we have answered to all the comments from the reviewer. Our answers are coloured in blue and are below the original comments coloured in red.

Major points:

1. Clarity of the manuscript. The model is very complex and it has lots of details, but the authors haven't always explained the concepts clearly or properly. I have written down some specific examples that the authors can fix easily. However, they should try to verify that their findings are backed by arguments that are explained in a logical way.

The abstract should be shortened to include the top 3 most important results, and be revised for clarity and shorter sentences. For example, what do the authors want the paper be known/cited for?

The Abstract has been shortened and reformulated during the review process in order to address the reviewer's comment.

In the introduction, the link between mantle convection and lithosphere deformation is quite abrupt (with a sentence about age of the Earth that is irrelevant to this study). The question 'why is convection important? 'is not satisfactorily introduced or linked to coupled lithosphere-mantle deformation.

Convection regulates the long-term temperature and mechanical structure of the lithosphere: the strength of the lithosphere is inter alia temperature dependent. Thus, convection may have a direct impact on the deformation of the lithosphere. Coupling convection to lithosphere deformation in numerical models can therefore improve our understanding of lithospheric scale processes, such as rifting and subduction.
Also, convection can generate forces due to up- and down-welling of mantle material, which can affect lithosphere deformation.
A paragraph has been added to the introduction for better explanation.

General suggestion: too many commas. Try to rephrase/split sentences with more than 2 commas or that are longer than 2 lines.

This has been implemented in the revised version of the manuscript.

2. Results section. I think the reference model (M1) should be described separately (evolution between extension, relation, convergence). Then compare models M2-M6 with M1 to highlight the effect of various factors. Figures should be adapted accordingly. The reason for this are the following: - in current form, the comparison is all over the place and it is confusing. It is not very clear which simulations the main text is referring to sometimes. - the current arrangement of figures is random. It starts with 2, 5, 8, 4, 10 etc. Their placement should follow a logical order of arguments.

We have changed the structure of the manuscript as follows:
Results section:

- The evolution of the reference run and M6 is now described separately (two standalone figures)
- M2-5 are then compared to M1 at each deformation stage.
- Figures have been modified and reordered accordingly

Discussion section:
We have restructured the discussion section and are now discussing implications of the models for several geodynamic problems including:

- Spontaneous vs. Induced subduction initiation
- Mantle convection stabilising single-slab subduction
- Onset of upper mantle convection and thermo-mechanical evolution of the lithospheric plates
- Impact of mantle viscosity structure and effective conductivity on passive margin formation

The comparison between M1-M6 should be done in terms of Ra. The k, viscosity cutoff, flow laws, they essentially affect the Ra.

We have largely implemented this suggestion in the revised version of the manuscript.

3. Thermal softening. A quick search in the manuscript finds 'thermal softening 'only in the abstract, very late discussion and conclusion, yet it is suggested as a key process that controls subduction initiation. I'm pointing out that it is incompletely described and linked to the hypothesis of the study and results.
For example, Line 425: thermal softening is introduced only now. not clear why bring it up here? Line 461: say that structural and thermal softening are important, but they were intro- duced late, without much context.
Moreover, the authors suggest in multiple places that it is the structural softening (in- heritance) and convection (slab suction) that help initiation. The authors need to clarify what are the main findings, and arguments need to be revised. One finding that I think is important: the required driving force to initiate subduction is much smaller, when convection and structural inheritance are considered.

We have discussed thermal and structural softening in more detail in the revised version of the manuscript.

4. Modelled vs parameterised convection. In Line 126, 3 types of simulations are introduced: 1) model convection with a weak asthenosphere, 2) parametrised convection, by scaling the thermal conductivity to the Nusselt number 3) impact of different viscosity structures
First, the treatment of the mantle convection is not clear in the main text (Lines 124- 130). What drives convection? How is the applied parametrised convection different? When is the onset of convection? Is convection only during the thermal relaxation stage? What controls the size of the convection cells? Also, explain how the Ra_avg is calculated.

This has been addressed during the review process. The Rayleigh number is calculated locally at each grid point. Ra_avg is the arithmetic average of all local Rayleigh numbers > 1000. We have clarified in the revised version of the manuscript.

In point 3) above which approach are you using: modelled/parameterized convection? While it is explained better in Appendix B, the differences between them are not clear in the main text. For example, 1) would be M1, while 2) is M, and 3) is M6?

This was explained in lines 130-135, but we changed the numbering of the models in the revised version of the manuscript, for clarity.

5. Other questions.

The geodynamic cycle modelled: 1) 30 Myrs extension at 2cm/yr 2) 70 Myrs thermal relaxation 3) 20 Myrs convergence at 3 cm/yr
What is the motivation behind these choices: 1) why thermal relaxation 2) why those time intervals 3) why those extension/convergence rates? Also, what are the boundary conditions during thermal relaxation?

The aim is to model the opening of a ca. 400 km wide oceanic basin without formation of a mature oceanic crust in an ultra-slow to slow spreading rift system. The durations of the periods and boundary conditions are chosen to allow for comparison of model results to orogens that formed from the collision of magma-poor hyper-extended margins, such as the European Western and Central Alps. We have clarified this in the revised version of the manuscript.

Why the choice of those parameters to change?

The viscosity structure of the mantle is poorly constrained and has a direct impact on the convective flow of the mantle. The effective conductivity approach is used to stabilise the thermal field in numerical simulations, but its impact on the deformation of the rigid plates and on self-consistent subduction initiation has not been tested yet. These statements have been added as a motivation in the introduction.

Are the surface processes important? Have you run models without? Do they introduce further heterogeneities in the model that affect the outcome?

Testing in more detail coupled surface processes to the deformation in the lithosphere and convection in the upper mantle is beyond the scope of this study. We have included only a simple parameterisation of surface process into the model to avoid unrealistically high and low topography. This has been clarified in the revised version of the manuscript.

6. Subduction initiation. It seems like symmetric vs asymmetric spreading also controls to a large extent subduction initiation, whether it is single/double subduction. I feel very little discussion is about that, and more on structural and thermal softening.

Subduction is always initiated during convergence. Most inheritance from the spreading is restricted to the margin geometry and is thus structural. Localisation occurs in the lithospheric mantle beneath the margins due stress concentration at the beginning of the necking zone. The heterogeneity in the upper mantle introduced by the convection and, therefore, the suction force of downward directed material flow seems to influence whether there is single or double-sided subduction. This has been clarified during the review process.

Also, there are other previous efforts to model extension/compression to obtain structural inheritance and subduction initiation (i.e. Gulcher et al 2019). The authors discuss simpler treatments of subduction initiation in paragraph 280, but do not relate to newer efforts to avoid the use of artificial features. So, are these newer models better for studying subduction initiation?

(Gülcher, Beaussier, & Gerya, 2019) investigate detachment faults as potential weak zones for intra-oceanic subduction initiation, not at a passive margin as wee do. Nevertheless, such detachment faults might be a geologically observable weak zone at which subduction could be potentially initiated. This topic has been addressed in the introduction of the revised manuscript.

Minor points:
Line 8-10: revise sentence

We have incorporated this suggestion in the revised version of the manuscript.

Line 10: only from the abstract it is not clear what the parameters were used, so saying that a viscosity of 5e20 Pa.s was used (as compared to what?) is not very meaningful. Rephrase

We have rephrased the abstract in the revised version of the manuscript

Line 20: multiple use of 'geodynamic 'in the same sentence Line 29-30: rephrase.

We have rephrased this sentence in the revised version of the manuscript.

Line 31: while it is an interesting fact - the calculation of the age of the earth - is not very relevant to the manuscript.

We have deleted this sentence in the revised version of the manuscript.

Line 35: rephrase

We have rephrased this sentence in the revised version of the manuscript.

Line 40: unlikely to be problematic

We have changed this accordingly in the revised version of the manuscript.

Line 41: delete likely

We have changed this accordingly in the revised version of the manuscript.

Line 55: authors relate to numerical aspects such as time step size, without mentioning why? The context was on physical aspects of convection.

This has been moved to the appendix in the revised version of the manuscript.

Line 64-65: should be in the first paragraph of introduction

This suggestion has been implemented in the revised version of the manuscript.

Line 68: Why only upper mantle? This is discussed late in discussion (section 4.4, paragraph 395)

This has been moved to the introduction section in the revised version of the manuscript.

Line 69: delete 'of applying'

This has been changed in the revised version of the manuscript.

Line 74: revise sentence - its meaning is not clear to someone who hasn't read the methods/results section.

We have restructured this part of the introduction in the revised version of the manuscript.

Line 88: reference to the code how it was benchmarked? (Info in appendix A, but should be in the main text too)

The benchmarks have been moved to the main text in the revised version of the manuscript.

Line 89: rephrase

We have rephrased this sentence in the revised version of the manuscript.

Line 96: repeats with Line 92, also Duretz et al 2016/2016a?

This has been changed in the revised version of the manuscript.

Line 105: based on the sentence the crust should be: 3*5+4*5 = 35 km thick. But a sentence earlier it is 33km

This mistake has been corrected and the description of the initial configuration has been clarified in the revised version of the manuscript.

Line 103: what is the mathematical expression for the perturbation? in case the model needs to be reproduced?

We have added the mathematical expression for the marker field perturbation to the algorithm description in the appendix of the revised version of the manuscript.

Line 111: more details on the rheology? Indicate appendix A for reference

We refer to the appendix in the revised version of the manuscript.

Line 113: reference to "corresponding laboratory flow law estimates"?

References are given in the footnote of table A1.

Line 113-114: rephrase. i.e. The mantle lithosphere is rheologically stronger than the mantle asthenosphere due to the temperature gradient.

With this sentence, we want to emphasize that we used the same material parameters for both the lithospheric and the upper mantle. We therefore keep the phrase as it is.

Line 120: what is the motivation for alternating between calcites and pelites for sedi- mentation algorithm?

To account for changes in sediment strength due to changing sedimentary environments. A detailed investigation on the impact of different implementations of surface processes is beyond the scope of this study. We have clarified this in the revised version of the manuscript.

Line 130 - give reference to Table 1.

This has been adapted in the revised version of the manuscript.

Line 133: viscosity cutoff for M1 is not provided to understand the difference.

The cut-off value is given in table 1.

Line 134: realistic value? Are the other values not realistic?

A thermal conductivity value of 36 is not realistic for a peridotite at upper mantle temperature and pressure conditions.

Line 148: Figure 2
-> can define a variable F = 2xtau_II

We have introduced a variable for the plate driving forces in the revised version of the manuscript.

Line 160: introduce the horizontal driving force per unit length, but what is it proxy for?

It is a proxy for the strength of the lithosphere and it indicates how much force is needed to localise deformation. Has been clarified during the review process.

Line 168: you can't see to a depth of 660 km as indicated

Figures have been adapted accordingly.

Line 195: values

This sentence has been rephrased during the review process.

Line 197: what is the delta GPE showing? (Info given in appendix)

The gravitational potential energy (GPE) has been explained in the appendix. Delta means that it is the difference of the GPE compared to a reference value, commonly the value close to one of the boundaries.

Line 224: reference to fig 9a, yet that figure is for extension stage.
Paragraph 220-225: confusing.

This has been addressed during the review process.

Line 228: what is mechanical heterogeneity? increases the strength of the weak layers

We have clarified this in the revised version of the manuscript.

Line 233: breaks later than the continental. after gives the impression of location.

This suggestion has been implemented into the revised version of the manuscript.

Line 234: what do you mean 'Mantle convection does not establish as early as rifting and crustal separation.'?

This sentence has been clarified in the revised version of the manuscript.

Lines 243, 245: use of realistic. close to the Ra estimated for the Earth. Line 250: which modulates mantle velocities.

We have implemented this suggestion in the revised version of the manuscript.

Line 250-254: why the discussion on time step size (a numerical feature) here? Lines 255-257: which simulation results are the authors referring here?

This has been addressed in the revised version of the manuscript

Paragraph 258: reference figure 5e,j in this paragraph. Also, maybe plot density aver- ages in passive margins/exhumed mantle separately?

We have tried to modify this figure, but the figure becomes too busy.

Line 272: you jump from density differences to values of tectonic forces. An additional sentence needs to connect them (i.e. estimate the buoyancy force due to modelled density differences). How much is needed to initiate subduction? (a similar calculation is done in line 315)

This depends on the strength of the lithosphere, which is still subject to debate.

Paragraph 280: this should come before the Cloos 1993 paragraph

We have restructured the discussion accordingly during the review process.

Line 286: yes, but under convergence

When a major weak zone is imposed, one also has to push from the sides to initiated subduction. Only when you skip the process of subduction initiation and already assume the presence of an inclined slab, subduction continues freely (given that the initial slab is long enough and that boundary conditions have been chosen correctly).

Line 294-295: total convergence is double sided, while in M1 is single-sided (asym- metric). Not clear why subduction initiation is stable only in M1. Convection cell size important? how about thickness of lithosphere at the point? M2-M5 are quite symmet- ric and they all have Ra_avg ~1e5, while M1 has Ra_avg~1e6. That should have an effect.

Likely the distribution of cells is important: the more asymmetric the mantle flow the more the model tends to produce single-slab subduction rather than double-slab subduction. Asymmetry decreases with decreasing Ra. This has been explained in more detail during the review process.

Line 312-315: - suction force induced by down-welling in the convection cell in M1. What is the similar force in the other simulations?

It is likely similar in M2-3 compared to M1, because of similar density distributions and vertical velocities, but it has not been computed here. Due to higher temperatures induced by the effective conductivity, values for densities are lower in M4-5 compared to M1-3, which explains the reduced absolute speed of material in the convecting cells. Most important seems to be the asymmetric distribution of the cells. We have clarified this during the review process.

Line 330: not sure what the reference is for. The double-subduction term was not coined by those workers.

We cited this reference, because they also modelled double-sided subduction.

Line 332: sentence not clear. Which simulation are you referring? would say M2-5 are more or less symmetric double subduction

Has been reformulated during the review process.

Line 343: onset of convergence - unclear when this happens?

This happens at 100 Myrs. We have clarified this in the revised version of the manuscript.

Line 357-361: use of 'realistic'

This has been rephrased in the revised version of the manuscript.

Paragraph 395: this paragraph should be in the methods, as it motivates/explains your model domain until 660 km. The sentence 'The convective patterns simulated in our study are in agreement with these observations. 'is irrelevant because you don't model the lower mantle.

This paragraph has been moved to the introduction in the revised version of the manuscript.

Paragraph 410: this should come earlier - I had questions about it earlier. on previous work on subduction initiation.

This paragraph has been moved to the introduction in the revised version of the manuscript.

Line 418: most definitely will have an impact

We have implemented this suggestion in the revised version of the manuscript.

Paragraph 430: and melting

We have accounted for this suggestion in the revised version of the manuscript.

Line 444-446: rephrase/simplify.

We have rephrased the conclusion section in the revised version of the manuscript.

Figures and Tables: Table 1: thermal conductivity should be 'k 'without the 'th 'sub- script. The authors can also provide the formula for the Ra number in the main text. How was the Ra_avg calculated?

This suggestion has been implemented in the revised version of the manuscript.

Table 2: there should be a column 'Description 'to describe the meaning of each pa- rameter i.e. 'rho0 '- reference density. Use k instead of k_th for thermal conductivity. What is dry/wet mantle? I assume wet mantle applies only to M6? Plastic and elastic parameters are also listed. Not very clear in the main text.

Description of the parameters is given in the appendix section. We also moved the table to the appendix for better comprehension. Dry and wet refer to the rheological parameters for dry and wet olivine. In the references given below the (now) appendix table 1 it is explained in which model we use either wet or dry rheologies.

Figure 4: why plot the vertical velocity field separate from the horizontal? should plot arrow/streamlines field to see the convection cells.

The figure has been adapted accordingly in the revised version of the manuscript.

Figure 10: should be merged with Figure 4. One column velocity, one column temperature.

These two figures have been merged in the revised version of the manuscript.

Figure 5: What if you plot the profiles at the rift axis (within a distance) vs off-axis on either flanks of the rifts? Caption: g-j show enlarged areas.

We have tested this version of the figure, but it becomes too busy, unfortunately.

Figure 6: legend: temp contours are red.

We have corrected this mistake in the revised version of the manuscript.

Figure 8: the line plots are not entirely clear. Maybe use a dotted line instead of dashed line? and same thickness.

We have changed line style and colour for the figures in the revised version of the manuscript.

Appendix A
Line 481: that's a strange notation of i,j indices (Einstein notation).
Eq A2: if written in Einstein notation, then vectors are written in terms of scalar com- ponents (a_i should not be bold). Same in Line 482 a=[0,g]. -> revise this appendix for completeness of sentences, and explanation of all parameters. For example, what is Ap, tauP etc. Gamma value? in eq A10

We have clarified these points in the revised version of the manuscript.

Appendix B
Paragraph 531: rephrase

This suggestion has been implemented in the revised version of the manuscript.

Line 560: not clear

This has been clarified in the revised version of the manuscript.

Appendix C
Line 595: gamma_T=1? Line 601: g=10^4?

To perform this benchmark one has to apply a local Rayleigh number of 100 at the top and using the Frank-Kamenetskii approximation a local Rayleigh number of 10^7 at the bottom. Choosing the dimensionless values as it is done here matches those numbers and reproduces the desired pattern with diagnostic quantities that are in the range of values reproduced by other algorithms as tested in (Tosi, et al., 2015).

Bibliography

Gülcher, A. J., Beaussier, S. J., & Gerya, T. V. (2019). On the formation of oceanic detachment faults and their influence on intra-oceanic subduction initiation: 3D thermomechanical modeling. *Earth and Planetary Science Letters, 506*, 195-208.

Tosi, N., Stein, C., Noack, L., Hüttig, C., Maierova, P., Samuel, H., . . . Glerum, A. (2015). A community benchmark for viscoplastic thermal convection in a 2-D square box. *Geochemistry, Geophysics, Geosystems, 16*(7), 2175-2196.

List of most relevant changes

Text:

- Abstract has been shortened and reformulated highlighting the most important findings
- Introduction has been rewritten and restructured:
    - Definition of convection and description of convection in the Earth's mantle
    - Motivation for studying coupled lithosphere-mantle deformation
- Method section:
    - More detailed description of the initial configuration and assumptions made
    - Table 2 has been moved to the appendix
- Results section:
    - Evolution of reference model and Wet Olivine model are described separately
    - Results of other models are presented in comparison to the reference model for the different deformation stages
- Discussion section has been completely restructured discussing:
    - Spontaneous vs. induced subduction initiation and estimates for plate driving forces
    - Mantle convection stabilising single-slab subduction
    - Onset of upper mantle convection and thermo-mechanical evolution of the lithospheric plates
    - Impact of mantle viscosity structure and effective conductivity on passive margin formation
    - Mantle convection, thermal erosion and tectonics in the Archean
    - Comparison with estimates of Earth's mantle viscosity and thermal structure
    - Formation and reactivation of magma-poor rifted margins: potential applications
- Conclusion has been shortened and reformulated
- Appendix: more detailed explanation on the assumptions made:
    - Justification for extended Boussinesq approximation
    - Justification for neglecting latent heat
- Mistakes have been corrected and paragraphs have been rephrased for clarity

Figures:

- Velocity glyphs have been added to figures for better visualisation of the material flow field
- Reference model and the Wet Olivine model are shown in separate figures
- Figures have been rearranged in a logical order following the structure of the revised version of the manuscript
- Figure 10 has been merged with figure 4
- Colour and style of lines have been changed for better visualisation

Impact of upper mantle convection on lithosphere hyper-extension and subsequent convergence-induced subduction

Lorenzo G. Candioti[1], Stefan M. Schmalholz[1], and Thibault Duretz[1,2]

[1]Institut des sciences de la Terre, Bâtiment Géopolis, Quartier UNIL-Mouline, Université de Lausanne, 1015 Lausanne (VD), Switzerland
[2]Univ.-Rennes, Rennes, France

Correspondence: Lorenzo G. Candioti (Lorenzo.Candioti@unil.ch)

Abstract.

~~We present~~ Many plate tectonic processes, such as subduction initiation, are embedded in long-term (>100 Myrs) geodynamic cycles often involving subsequent phases of extension, cooling without plate deformation and convergence. However, the impact of upper mantle convection on lithosphere dynamics during such long-term cycles is incompletely understood. We have designed two-dimensional ~~thermo-mechanical numerical models of coupled lithosphere-mantle deformation , considering the upper mantle~~ (2D) upper mantle-scale (down to a depth of 660 km) thermo-mechanical numerical models of coupled lithosphere-mantle deformation. We consider visco-elasto-plastic deformation and for the lithospheric and upper mantle a combination of diffusion, dislocation and Peierls creep. Mantle densities are calculated from petrological phase diagrams (Perple_X) for a Hawaiian pyrolite. ~~The~~ We tested the impact of the viscosity structure in the asthenosphere on upper mantle convection and lithosphere dynamics. Further, we quantified the plate driving forces necessary for subduction initiation in 2D thermo-mechanical models of coupled lithosphere-mantle deformation. Our model generates a 120 Myrs long geodynamic cycle of subsequent extension (30 Myrs), cooling (70 Myrs) and convergence (20 Myrs) coupled to upper mantle convection in a single and continuous simulation~~with explicitly modelling convection in the upper mantle. During lithosphere extension, the models generate an approximately~~ . Fundamental features such as the formation of hyper-extended margins, upper mantle convective flow and subduction initiation are captured by the simulations presented here. Vertically forced subduction initiation due to negative buoyancy of the 400 km ~~wide basin of exhumed mantle bounded by hyper-extended passive margins. The models show that considering only the thermal effects of upper mantle convection by using an effective thermal conductivity generates results of lithosphere hyper-extension that are similar to the ones of models that explicitly model the convective flow . Applying a lower viscosity limit of 5×10^{20} suppresses convection and generates results different to the ones for simulations with a low viscosity asthenosphere having minimal viscosity of approximately 10^{19} . During cooling without far-field deformation, no subduction of the~~ wide cooled exhumed mantle is ~~spontaneously initiated. Density differences between lithosphere and mantle are too small to generate a buoyancy force exceeding the mechanical strength of the lithosphere. The extension and cooling stages generate self-consistently a structural and thermal inheritance for the subsequent convergence stage. Convergence initiates subduction of the exhumed mantle at the transition~~ not observed after 100 Myrs. Instead, subduction initiation is horizontally forced and occurs at the transition from the exhumed mantle to the

hyper-extended ~~margins~~passive margin. The main ~~mechanism of subduction initiation~~ localization mechanism is thermal softening~~for a plate driving force (per unit length) of approximately~~ . Compared to a strong asthenosphere, a weak asthenosphere leads to (1) more asymmetric passive margins during rifting, (2) locally larger suction forces due to more vigorous convective flow assisting in establishing single-slab subduction rather than double-slab subduction and (3) horizontally forced subduction initiation for a smaller value of plate driving forces (15 TN m^{-1} ~~. If convection in the mantle is suppressed by high effective thermal conductivities or high, lower viscosity limits, then subduction initiates at both margins leading to divergent double-slab subduction. Convection in the mantle assists to generate a single-slab subduction at only one margin, likely due to mantle flow which exerts an additional suction force on the lithosphere.~~ instead of 22 TN m^{-1}). The first-order geodynamic processes simulated in ~~the geodynamic cycle of subsequent extension, cooling and convergence~~ this geodynamic cycle are applicable to orogenies that resulted from the opening and closure of embryonic oceans bounded by magma-poor hyper-extended ~~passive~~ rifted margins, which might have been the case for the Alpine orogeny.

1 Introduction

~~Heat transfer~~

1.1 Convection in the Earth's mantle

In general, the term convection can be used to describe any motion of a fluid driven by external or internal forces (Ricard et al., 1989). Prout (1834) derived this term from the Latin word "convectio" (to carry, or to convey) to distinguish between advection dominated heat transfer and conduction, radiation dominated heat transfer.

On Earth, heat transfer through the lithosphere is dominated by thermal ~~diffusion~~ conduction while heat transfer through the underlying mantle is dominated by ~~convection (e.g. Turcotte and Schubert, 2014). Over long-term (>≈80 Myrs) geodynamic time scales, convection below the lithosphere maintains a more or less constant thermal thickness and heat flow across the lithosphere, and hence regulates and stabilises the thermal structure of the lithosphere (Richter, 1973; Parsons and McKenzie, 1978). Consequently, convection in the mantleexerts a strong control on the thermal history, and consequently on the strength, of the lithosphere. Historically, ignoring convection was the main reason why early estimates of the age of the Earth, based on thermal diffusion only, provided estimates that were two orders of magnitude too young; it was not because of ignoring radioactive heat production (England et al., 2007).~~

~~Many coupled lithosphere-mantle deformation processes are still incompletely understood, such as the formation of hyper-extended passive margins (e.g. Peron-Pinvidic et al., 2019) or the initiation of subduction (e.g. Stern and Gerya, 2018). In addition to field and laboratory studies, theoretical studies, often employing numerical simulations, have proven useful to understand and quantify these processes. Many numerical studies modelling the deformation of~~ advection of material (e.g. Turcotte and Schubert, 2014). Convection can occur over the whole mantle, down to the core-mantle boundary, or it can be layered. At temperature and pressure conditions corresponding to a depth of about 660 km, the ~~lithosphere and~~ mineralogy of peridotite changes from γ-spinel-olivine to perovskite + magnesiowüstite. This phase transition is endothermic, which means it has a negative pressure-temperature,

so-called Clapeyron slope. This phase transition likely impedes the penetration of cold slabs subducting into the lower mantle and hot plumes rising into the upper mantle (Schubert et al., 2001). The 660-km phase transition can therefore represent a natural impermeable boundary, that separates two convecting layers. Laboratory experiments, tomographic images and calculations on the ~~underlying mantle , or thinner asthenosphere, do not directly model convective flows below the lithosphere (e.g. Jaquet and Schmalholz, 2018; Beaussier et al., 2019; Erdös et al., 2019; Li et al., 2019). Ignoring convection below the lithosphere in numerical simulations is likely not problematic if the duration of the simulated deformation is not exceeding a few tens of millions of years, because in such short time interval the diffusive cooling of the lithosphere is likely negligible. However, many studies show that geodynamic processes are embedded in longer term cycles, such as the Wilson cycle (Wilson, 1966; Wilson et al., 2019), and that tectonic inheritance of earlier extension and cooling events had presumably a major impact on subsequent convergence and subduction evolution (Chenin et al., 2019). To investigate and quantify the coupled lithosphere-mantle deformation processes during such long-term geodynamic cycles with numerical models, requires, in principle, that these models consider~~ Earth's heat budget deliver evidence that a mixed mode of both types best explains convection in the ~~mantle. However, modelling convection in the mantle directly in numerical models focusing on lithosphere deformation is computationally challenging since the convection in the mantle must be spatially and temporally resolved. A common approach to indirectly include the thermal effects of mantle convection into numerical simulations, without actually modelling the velocity field of the convection, is to use an effective thermal conductivity parameterization (Zahnle et al., 1988)~~ present-day Earth's mantle (Li et al., 2008; Chen, 2016).

Any convecting system can be described by a dimensionless number, the so-called Rayleigh number. It is defined as the ratio of the diffusive and the advective time scale of heat transfer (see also~~appendix B). This approach has been used, for example, in mantle convection studies for planetary bodies when the convection of the mantle is too vigorous to be modelled explicitly (e.g. Zahnle et al., 1988; Tackley et al., 2001; Golabek et al., 2011), in models of back-arc lithospheric thinning through mantle flow that is induced by subduction of an oceanic plate (e.g. Currie et al., 2008), or in models of lithosphere extension and subsequent compression (e.g. Jammes and Huismans, 2012). A main parameter controlling the vigour of upper mantle convection is~~ , eq. B1). The critical value of the Rayleigh number necessary for the ~~effective viscosity. The larger the viscosity the smaller the convectionvelocities and, with respect to numerical simulations, the larger is also the admissible time step and the smaller the number of total time steps~~onset of convection is typically in the order of 1000 (Schubert et al., 2001). Convection in the Earth's mantle can occur at Rayleigh numbers in the range of 10^6-10^9 depending on the heating mode of the system and whether convection is layered or it includes the whole mantle (Schubert et al., 2001). The higher the Rayleigh number, the more vigorous is the convection, i.e. advection of material occurs at a higher speed. Vigour of both whole and layered mantle convection is inter alia controlled by the effective viscosity of the mantle. However, compared to the density and thermal structures, the viscosity structure of the ~~upper~~ mantle is subject to large uncertainty~~, because it is not directly~~ . Viscosity is not a direct observable and can only be inferred by inverting observable geophysical data such as data for glacial isostatic adjustment (e.g. Mitrovica and Forte, 2004)) or seismic anisotropy data (e.g. Behn et al., 2004). Especially at depths of ca. 100-300 km, ~~within the~~ in the so-called asthenosphere, the inferred value for viscosity varies greatly (see fig. 2 in Forte et al. (2010))~~and~~ . Values for effective viscosity in this region can be up to two orders of magnitude lower than estimates for the average upper

mantle viscosity of $\approx 10^{21}$ Pa s (Hirth and Kohlstedt, 2003; Becker, 2017). ~~Furthermore, mantle convection does not only impact the~~

1.2 Why to couple mantle convection to the long-term deformation cycle of the lithosphere?

Many coupled lithosphere-mantle deformation processes, such as the formation of hyper-extended passive margins and the mechanisms leading to the initiation of subduction (e.g. Peron-Pinvidic et al., 2019; Stern and Gerya, 2018), are still elusive. Crameri et al. (2020) compiled a database from recent subduction zones to investigate whether subduction initiation was vertically (spontaneous) or horizontally forced (induced, see also Stern (2004) for terminology). They concluded that, during the last ca. 100 Myrs, the majority of subduction initiation events were likely horizontally forced. Recent numerical studies have investigated thermal softening as a feasible mechanism for horizontally forced subduction initiation (Thielmann and Kaus, 2012; Jaque In these models horizontally forced subduction was initiated without prescribing a major weak zone cross-cutting through the lithosphere. These models do not require further assumptions on other softening mechanisms, such as micro-scale grain growth or fluid- and reaction-induced softening. Therefore, these models are likely the simplest to study horizontally forced subduction initiation.

Geodynamic processes, such as lithosphere extension or convergence, are frequently studied separately. In fact, many studies show that these processes are embedded in longer term cycles, such as the Wilson cycle (Wilson, 1966; Wilson et al., 2019). Over large time scales (>≈80 Myrs), tectonic inheritance of earlier extension and cooling events (Chenin et al., 2019) together with mantle convection (Solomatov, 2004) had presumably a major impact on subsequent convergence and subduction. Certainly, subduction initiation at passive margins during convergence can be studied without a previous extension and cooling stage (Kiss et al., 2020). An initial passive margin geometry and thermal field must be then constructed ad-hoc for the model configuration. However, it is then uncertain whether the applied model would have generated such a margin geometry and thermal structure during an extension simulation. In other words, it is unclear whether the initial margin configuration is consistent with the applied model.

Plenty of numerical studies modelling the deformation of the lithosphere and the underlying mantle do not directly model convective flows below the lithosphere (e.g. Jaquet and Schmalholz, 2018; Gülcher et al., 2019; Beaussier et al., 2019; Erdös et al., 2019; I Ignoring convection below the lithosphere in numerical simulations is unlikely problematic, if the duration of the simulated deformation is not exceeding a few tens of millions of years. In such short time intervals, the diffusive cooling of the lithosphere is likely negligible. However, convection in the Earth's mantle has a fundamental impact on the long-term dynamics in the lithosphere: regulating the thermal structure of the lithosphere ~~but can also~~ (Richter, 1973; Parsons and McKenzie, 1978) mantle convection has a strong control on the lithospheric strength. Furthermore, mantle convection can exert suction forces on the lithosphere due to mantle flow ~~on the lithosphere (e.g. Conrad and Lithgow-Bertelloni, 2002). For example, numerical simulations show that suction forces due to mantle flow~~ (e.g. Conrad and Lithgow-Bertelloni, 2002). Numerical studies show that these suction forces can assist in the initiation of subduction (Baes et al., 2018). ~~Currently, the impact of convection in the upper mantle on the long-term deformation~~ Therefore, coupling mantle convection to lithospheric scale deformation

can potentially improve our understanding of processes acting on long-term geodynamic cycles of the lithosphere~~is not well understood~~.

Here, we ~~simulate numerically a long-term (~~present two-dimensional (2D) thermo-mechanical numerical simulations modelling the long-term cycle of coupled lithosphere-mantle deformation. The modelled geodynamic cycle comprises a 120 Myrs history of extension–cooling–convergence leading to horizontally forced subduction. We include the mantle down to a depth of 660 km assuming that convection is layered. Timings and deformation velocities for the distinct periods have been chosen to allow for comparison of the model results to the Alpine orogeny. With these models, we investigate and quantify the impact of (1) the viscosity structure of the upper mantle and (2) an effective conductivity parameterization on upper mantle convection and lithospheric deformation. Applying this parameterization diminishes the vigour of convection, but maintains a characteristic thermal field for high Rayleigh number convection (explained in more detail in the next section and in appendix B). (3) We also test creep law parameters for wet and dry olivine rheology. Finally, we investigate whether forces induced by upper mantle convection have an impact on horizontally forced subduction.

2 The applied numerical model

The applied numerical algorithm solves the partial differential equations for conservation of mass and momentum coupled to conservation of energy. We consider the deformation of incompressible visco-elasto-plastic slowly (no inertia) flowing fluids under gravity. The equations are discretized on a 2D finite difference staggered grid in the Cartesian coordinate system. Material properties are advected using a marker-in-cell method (Gerya and Yuen, 2003). A 4th order Runge-Kutta scheme is employed for marker advection and a true free surface is applied (Duretz et al., 2016a). A detailed description of the algorithm is given in the appendix A. A convection benchmark of the applied algorithm is presented in appendix C. The applied algorithm has already been used to model processes at different scales, such as deformation of eclogites on the centimetre-scale (Yamato et al., 2019), crystal-melt segregation of magma during its ascent in a meter-scale conduit (Yamato et al., 2015), rifting of continental lithosphere (Duretz et al., 2016b; Petri et al., 2019), stress calculations around the Tibetan Plateau (Schmalholz et al., 2019) a within and around the subduction of an oceanic plate (Bessat et al., 2020), as well as modelling Precambrian orogenic processes (Poh et al., 2020).

2.1 An approach to model coupled lithosphere-mantle deformation

We test the algorithm for capability of reproducing first order features of a long-term (>100 ~~Myrs) geodynamic cycle of subsequent lithosphere extension, cooling without far-field deformationand convergence causing subductioninitiation, and consider convection in the~~ Myrs) geodynamic cycle namely: (i) formation of hyper-extended magma-poor rifted margins during a 30 Myrs extension period applying an absolute extension velocity of 2 cm yr^{-1}. (ii) Separation of the continental crust and opening of a ca. 400 km wide marine basin floored by exhumed mantle material. (iii) Generation of upper mantle ~~. We refer here to the region between the lithosphere and the density discontinuity at 660 km depth as upper mantle. We quantify the impact of different viscosity structures and of applying an effective thermal conductivity on upper mantle~~ convection

and coupled lithospheredeformation. during a 70 Myrs cooling period without significant plate deformation. (iv) Subsequent convergence for a period of 20 Myrs and horizontally forced subduction initiation in a self-consistent way, that means without modifying the simulation by, for example, adding ad-hoc a prominent weak zone across the lithosphere. During convergence, the self-consistently evolved passive margin system is shortened applying an absolute convergence velocity of 3 $\mathrm{cm\ yr^{-1}}$.

For simplicity, we consider lithosphere extension that generates magma-poor hyper-extended margins and crustal separation leading to mantle exhumationso . This means that we do not need to model melting, lithosphere break-up, mid-ocean ridge formation and generation of new oceanic crust and lithosphere. Such Wilson-type cycles, involving only embryonic oceans, presumably formed orogens such as the Pyrenees, the Western and Central Alps, and most of the Variscides of Western Europe (e.g. Chenin et al., 2019). We present two-dimensional (Values for deformation periods and rates in the models presented here are chosen to allow for comparison of the model evolution to the Alpine orogeny.

Further, tomographic images from the Mediterranean show large p-wave anomalies in the transition zone (Piromallo and Morelli, 2003) i that the 660-km phase transition inhibits the sinking material to penetrate further into the lower mantle. Therefore, we do not include the lower mantle into the model domain and assume that mantle convection is layered. Figure 1 shows the initial configuration of the models.

We present 2D) thermo-mechanical numerical simulationsin which convection in the upper mantle is . We model a long-term (>100 Myrs) geodynamic cycle of subsequent lithospheric extension, cooling and horizontally forced subduction coupled to upper mantle convection in a single continuous simulation.

As mentioned above, the viscosity structure of the mantle is far from being well constrained. Therefore, we test (1) directly modelled and numerically resolved, (2) slowed down by a viscosity models with different upper mantle viscosity structures. In model M1, the reference model, the asthenosphere is assumed to be weak with values for viscosity in the order of $\approx 10^{19}$ Pa s resulting from the applied flow laws. In models M2 and M3 the asthenosphere is assumed to be stronger. Values for viscosities in the asthenosphere are limited by a numerical cut-off and (3) indirectly considered by applying a high effective thermal conductivity. In addition to testing the impact of convection, another aim of modelling an entire extension-cooling-convergence-subduction cycle is to test whether the applied numerical algorithm can generate the first order features of such a cycle within a single and continuous numerical simulation. These first order features are : (value of 1) formation of hyper-extended passive margins, $\times 10^{20}$ Pa s in M2 and 5×10^{20} Pa s in M3. Coupling of lithosphere-mantle deformation is achieved by resolving numerically both lithospheric deformation and upper mantle convection in M1-3. (2) separation of the continental crust and exhumation of mantle material, (3)upper mantle convection stabilizing the thermal thickness of the lithosphere and heat flow at its base and (4) subduction initiation during convergence. A further aim of modelling such cycle is to test whether convergence-induced subduction at passive margins can occur in a self-consistent way without prescribing ad-hoc a prominent weak zone across the lithosphere We further test the impact of parameterizing convection on the cycle by scaling the thermal conductivity to the Nusselt number of upper mantle convection (see app. B). The effective conductivity approach has been used, for example, in mantle convection studies for planetary bodies when convection in the mantle is too vigorous to be modelled explicitly (e.g. Zahnle et al., 1988; Tackley et al., 2001; Golabek et al., 2011). Also, it has been used in models of back-arc lithospheric thinning through mantle flow that is induced by subduction of an oceanic plate (e.g. Currie et al., 2008) and in models of

lithosphere extension and subsequent compression (e.g. Jammes and Huismans, 2012). In these models, we also assume both a weak asthenosphere (as in M1) in model M4 and a strong asthenosphere (as in M2) in model M5. (3) We finally investigate the role of the olivine rheology. To this end, we perform an additional model M6 in which the material parameters of the dislocation and diffusion creep mechanism of a dry olivine rheology are replaced by the parameters of a wet olivine rheology (table A1). ~~We also investigate whether velocities and forces in the upper mantle associated to convection have an impact on convergence-induced subduction.~~

3 ~~Model~~

In M6, values for all the other parameters, both physical and numerical, are initially equal to those set in M1. Within the error range of values for activation volume and energy of the wet olivine rheology, the viscosity is calibrated to the data obtained by Ricard et al. (1989). However, using the highest possible values for the wet olivine flow law parameters, the maximum viscosity in the upper mantle is initially one order of magnitude lower compared to models M1-5. A summary of all simulations is given in table 1 and all material parameters are summarised in table A1.

2.1 ~~Mathematical model~~

~~We use 2D numerical thermo-mechanical simulations to investigate coupled lithosphere and mantle deformation during a geodynamic cycle of extension, cooling and subsequent convergence. The applied code solves the partial differential equations for deformation of incompressible visco-elasto-plastic slowly (no inertia)flowing fluids under gravity coupled with energy conservation. The equations are discretized on a 2D finite difference staggered grid in the Cartesian coordinate system. Material properties are advected using a marker-in-cell method (Gerya and Yuen, 2003). A 4th order Runge-Kutta scheme is employed for marker advection and a true free surface is applied (Duretz et al. , 2016). A detailed description of the algorithm is given in the appendix A. A convection benchmark of the applied algorithm is presented in appendix C.~~

2.1 Model configuration

The model domain is 1600 km wide and 680 km high and the applied model resolution is 801×681 grid points (fig. 1). Minimum z-coordinate is set to -660 km and the top +20 km are left free to allow for build-up of topography. The top surface (initially at z = 0 km) is stress free~~(free surface, see Duretz et al. (2016a)) and thus evolves dynamically throughout the model evolution~~. Thus, its position evolves dynamically as topography develops. Mechanical boundary conditions on the remaining boundaries are set to free slip at the bottom ~~, free slip with~~ and constant material inflow/outflow velocities at the left and right boundary. The boundary velocity is calculated such that the total volume of material flowing through the lateral boundary is conserved. The transition between inflow and outflow occurs at ~~half of the model height.~~ z = -330 km and not at the initially imposed lithosphere-asthenosphere boundary (LAB). Values for deviatoric stresses at this depth are significantly lower compared to those at the base of the lithosphere. This choice avoids boundary effects close to the mechanical lithosphere and the LAB can develop freely away from the lateral model boundaries. We use material flow velocity boundary condition

rather than bulk extension rates to deform the model units. Applying bulk extension rates and deforming the model domain would change the height of the model domain which has strong control on the Rayleigh number of the system (eq. B1).

Also, evolution of passive margin geometries becomes dependent on the model width when using bulk extension rates as mechanical boundary condition (Chenin et al., 2018). It is therefore more practical to use material flow velocity boundary conditions in the type of models presented here.

Initial temperature at the surface is set to 15 °C and temperatures at the crust-mantle (Moho) and at the ~~lithosphere-asthenosphere boundaries (LAB)~~ LAB are 600 °C and 1350 °C, respectively. Assuming an adiabatic gradient of 0.49 °C ~~/~~km^{-1} (see appendix B), the temperature at the model bottom is 1612 °C. Thermal boundary conditions are set to isothermal at the bottom and at the top of the domain and the left and right ~~boundary~~ boundaries are assumed to be insulating ~~, (~~i.e. no heat flows through lateral boundaries). Model units include a 33 km thick, mechanically layered crust which overlies an 87 km thick mantle lithosphere on top of the upper mantle. The resulting initial thickness of the lithosphere is thus 120 km. The crust includes three mechanically strong and four mechanically weak layers~~of~~ . The thickness of the weak layers is set to 5 km ~~thickness each.~~ each, the thickness of the uppermost and lowermost strong layer is set to 4 km, whereas the strong layer in the middle is 5 km thick. This thickness variation allows to match the total 33 km thickness of the crust without introducing an additional vertical asymmetry. Mechanical layering of the crust was chosen, because it is a simple way of considering mechanical heterogeneities in the crust~~that lead to more realistic~~ . The layering leads to the formation of numerous structural features observed in natural hyper-extended passive ~~margin geometries during rifting (Duretz et al., 2016b) . To localize the deformation in the horizontal center of the model, we perturbed the initially flat layer interfaces with a random vertical displacement, within the x-position -75 and 75 . The magnitude of the vertical displacement is in average approximately 1 . The viscosity of all model units is~~ margins (Duretz et al., 2016b) without relying on pre-defined strain softening.

We consider viscous, elastic and brittle-plastic deformation of material in all models presented here. Viscous flow of material is described as a combination of several flow laws~~including, ~~. We use dislocation creep for the crustal units and dislocation, diffusion and Peierls creep for ~~lithospheric mantle and the upper mantle~~mantle units (see also appendix A). The initial viscosity profile through the upper mantle is calibrated to match viscosity data obtained by Ricard et al. (1989). The applied flow law parameters lie within the error range of the corresponding laboratory flow law estimates ~~.~~ (see tab. A1).

The difference between mantle lithosphere and upper mantle is temperature only, i.e. all material parameters are the same. Density of the crustal phases is computed with a simple equation of state (eq. A3), whereas density of the mantle phase is precomputed using Perple_X (Connolly, 2005) for the bulk rock composition of a pyrolite (Workman and Hart (2005), fig. A1). A detailed description of the phase transitions and how the initial thermal field is calibrated is given in appendix B. Surface processes (e.g., erosion and sedimentation) are taken into account by a kinematic approach: if the topography falls below a level of 5 km depth or rises above 2 km height, it undergoes either sedimentation or erosion with a constant velocity of 0.5 mm yr^{-1} ~~within one numerical time step~~. In case of sedimentation, the generated cavity between the old and corrected topographic level is filled with sediments, alternating between calcites and pelites every 2 Myrs. This simple parameterization is employed to avoid exaggeratedly high and low topography.

Table 1. Parameters ~~used~~ varied in models M1-6.

Parameter	Unit	M1	M2	M3	M4	M5	M6
k ~~$_{\text{th}}$~~	$\mathrm{W\ m^{-1}\ K^{-1}}$	2.75	~~36~~ 2.75	2.75	~~36~~ **36**	~~2.75~~ **36**	2.75
η_{cutoff}	Pa s	1×10^{18}	~~1×10^{18}~~ **1×10^{20}**	~~1×10^{20}~~ **5×10^{20}**	1~~$\times10^{20}$~~ $\times10^{18}$	~~5×10^{20}~~ **1×10^{20}**	1×10^{18}
Ra_{avg}	-	~~2.2~~9.95$\times10^{6}$	~~4.4~~3.92$\times10^{5}$~~^{6}~~	~~8.8~~1.18$\times10^{5}$~~^{6}~~	~~1.2~~1.97$\times10^{5}$~~^{6}~~	~~2.75~~5.48$\times10^{5}$	~~5.0~~9.26$\times10$
Rheology mantle	-	Dry Olivine	Dry Olivine	Dry Olivine	Dry Olivine	Dry Olivine	~~Wet Olivine~~ **Wet**

k is thermal conductivity and η_{cutoff} is the lower viscosity limiter. Ra_{avg} is the arithmetic average of Rayleigh numbers >1000. Rayleigh numbers are computed locally at each cell center according to after 99 Myrs in model history for models M1-5 and after 26 Myrs for model M6. Bold font highlights the parameters varied compared to the reference model M1.

~~Three subsequent stages of a geodynamic cycle are modelled in one continuous simulation: (1) Opening of a ca. 360 wide marine basin floored by exhumed mantle and bounded by hyper-extended passive margins during a 30 extension period applying an absolute extension velocity of 2 $^{-1}$, (2) a 70 thermal relaxation period applying no far-field deformation and (3) a 20 convergence period in which the self-consistently evolved passive margin system is shortened applying an absolute convergence velocity of 3 $^{-1}$. Figure 1 shows the initial configuration of the models.~~

~~We perform three types of simulations to investigate the impact of upper mantle convection on lithosphere deformation: (1) explicitly modelling and numerically resolving the convection for a viscosity structure with a weak asthenosphere, (2) parameterizing convection, without explicitly modelling convective flow, by scaling the thermal conductivity to the Nusselt number of upper mantle convection (see app. B) and (3) quantifying the impact of different viscosity structures of the upper mantle. We perform a set of 6 simulations, which will be named M1-6 hereafter. M1 is defined as the reference model in which the process of convection is explicitly modelled and numerically resolved for a weak asthenosphere. M2 is equivalent to M1, except that the impact of convection on the thermal field is investigated using an effective conductivity. M3 and M4 are equivalent to M1 and M2, but the viscosity is limited to a minimum value of 10^{20} . In M5, the viscosity is limited to a minimum value of 5×10^{20} using a realistic value for thermal conductivity in the upper mantle. Mantle material in M1-5 is described using flow law parameters for a dry olivine rheology. In M6, the material parameters of the dislocation and diffusion creep mechanism of a dry olivine rheology are replaced by the parameters for a wet olivine rheology (table A1). Within the error range of values for activation volume and energy of the wet olivine rheology, the viscosity is calibrated to the data obtained by Ricard et al. (1989). However, using the highest possible values for the wet olivine flow law parameters, the maximum viscosity in the upper mantle is one order of magnitude lower compared to models M1-5. In M1,2 and M6, we apply a lower viscosity cut-off at 10^{18} . A summary of all simulations is given in table 1 and all material parameters are summarised in table A1.~~

~~Physical parameters used in the numerical simulations M1-6. Parameter Unit Strong Crust[1] Weak Crust[2] Sediments[3] Sediments[4] Dry Mantle[5,6] Wet Mantle[5,6] ρ_0 $^{-3}$ 2800 2800 2800 2800 - - G 2×10^{10} 2×10^{10} 2×10^{10} 2×10^{10} 2×10^{10} 2×10^{10} c_P $^{-1}$ $^{-1}$ 1050 1050 1050 1050 1050 1050 k $^{-1}$ $^{-1}$ 2.25 2.25 2.37 2.55 2.75 2.75 H_R $^{-3}$ 0.9×10^{-6} 0.9×10^{-6} 0.56×10^{-6} 2.9×10^{-6} 2.1139×10^{-8} 2.1139×10^{-8} C 10^7 10^6 10^7 10^6 10^7 10^7 φ 30 5 30 5 30 30 α $^{-1}$ 3×10^{-5} 3×10^{-5} 3×10^{-5} 3×10^{-5} 3×10^{-5} β $^{-1}$ 1×10^{-11} 1×10^{-11} 1×10^{-11} 1×10^{-11} 1×10^{-11} 1×10^{-11} Dislocation A $^{-n-r}$ $^{-1}$ 5.0477×10^{-28} 5.0717×10^{-18} 1.5849×10^{-25} 10^{-138} 1.1×10^{-16} 5.6786×10^{-27} n - 4.7 2.3 4.7 18 3.5 3.5 Q $^{-1}$ 485×10^3 154×10^3 297×10^3 51×10^3 530×10^3 460×10^3 V ^{3} $^{-1}$ 0 0 0 0 14×10^{-6} 11×10^{-6} r - 0 0 0 0 0 1.2 f_{H_2O} 0 0 0 0 0 10^9 Diffusion A^* $^{-n-r}$ ^{m} $^{-1}$ - - - - 1.5×10^{-15} 2.5×10^{-23} n - - - - - 1 1 Q $^{-1}$ - - - - 370×10^3 375×10^3 V ^{3} $^{-1}$ - - - - 7.5×10^{-6} 20×10^{-6} m - - - - - 3 3 r - - - - - 0 1 f_{H_2O} - - - - - 0 10^9 d - - - - 10^{-3} 10^{-3} Peierls A_P $^{-1}$ - - - - 5.7×10^{11} 5.7×10^{11} Q $^{-1}$ - - - - 540×10^3 540×10^3 σ_P - - - - 8.5×10^9 8.5×10^9 γ - - - - - 0.1 0.1~~

3 Results

We ~~describe here the results of M1-5 for the subsequent extension, thermal relaxation and convergence stages. We describe results~~ first describe the evolution of the reference model M1 and of M6~~for the extension stageonly~~. Model M6 is stopped after the extensional stage, because its later evolution is not applicable to ~~Earth due to unrealistic Moho temperatures and mantle viscosities~~present-day Earth. Thereafter, we compare the results of M2-5 to the results of M1 for the individual deformation stages. Finally, we compare the evolution of the plate driving forces in models M1-5 during the entire cycle. The arithmetic average Rayleigh number in all models is significantly larger compared to the critical Rayleigh number Ra_{crit} = 1000. The rifting and cooling period laterally perturb the thermal field sufficiently to initiate and drive the convection over large time scales in all models presented here.

3.1 ~~Extension~~Dry Olivine rheology: Model M1 - Reference run

3.1.1 ~~Dry Olivine rheology: Models M1-5~~

~~After 13 Myrs of extension M1-4 produce two conjugate asymmetric hyper-extended margins (c. f. 5~~Crustal break-up during the rifting phase in M1 occurs after ca. 8 Myrs (fig. 3(a) ~~-~~& (e)). ~~In M1 and M2, the~~ The left continental margin has a length of ca. 200 km and the right margin has a length of ca. 150 km (fig. ~~5(a)& (b)). Both margins of M3 and M4 are approximately 150 long (fig. 5(c) & (d)). The two margins of M5 are connected by a long zone of hyper-extended crust (i.e., its thickness is less than 10) and crustal break-up does not occur (fig. 5~~3(e)). Velocity arrows indicate upward motion of hot material in the centre of the domain (fig. 3(e))~~up to this stage.~~

~~In none of the simulations the 1350 isotherm comes closer than 30 km to the surface.~~ . Two convection cells begin to establish at this stage below each margin. The viscosity of the upper mantle decreases to minimal values ~~around~~ in the order of 10^{19} Pa s and increases again up to values in the order of 10^{21} Pa s at the bottom of the model domain ~~in~~ (fig. 3(a) and fig. 2(c)). Towards the end of the cooling period (at 97 Myrs), M1 ~~and M2. In M3, M4 and M5 the minimal value for effective~~

viscosity has developed circular shaped convection cells in the upper mantle is at the applied cut-off value of 1region of the upper mantle (above $z \approx$ -400 km, fig 3(b)). The average Rayleigh-number (see footnote of tab. 1) of the system computed at this late stage of the cooling period is ca. 9.95×10^{20} and 5×10^{20} and increases up to values in the order of 10^{21} at a depth of 660 ^{6} and the size of the cells varies between ca. 50 km and ca. 300 km in diameter below the left and right margin respectively (fig.3(b)). Below the right margin at $z \approx$ -150 km (see fig. 2(c)& 5(c)-(eand $x \approx$ +300 km the downward directed mantle flow of two neighbouring convection cells unify (fig. 3(f)). The top ca. 100 km of the modelled domain remain undeformed; no material is flowing in this region (area without velocity arrows in fig. 3(f)). Horizontally averaged vertical temperature profiles are similar in all the models (fig. 2(aConvergence starts at 100 Myrs and at ca. 102 Myrs, a major shear zone forms breaking the lithosphere below the right margin (inclined zone of reduced effective viscosity in fig. 3(c) & (g)). The level of the 1350 isotherm remains at its initial depth in M2 and M4, whereas it subsides by ca. 20 in M1, M3 and M5. Horizontally averaged, density profiles (fig. 2(b)) show density differences $<10^{-3}$ between ca. 35 and ca.120 depth. Velocity arrows in the lithosphere indicate the far-field convergence, whereas velocity arrows in the upper mantle show that convection cells are still active. The exhumed mantle is subducted in one stable subduction zone below the right continental margin. Several convection cells are active in the upper mantle during subduction (see velocity arrows fig. 3(d) & (h)). A trench forms in which sediments are deposited (fig.3(d) & (h)). Folding of the crustal layers in the overriding plate indicates significant deformation of the crust. The crustal layers of the subducting plate remain relatively undeformed. The viscosity in the asthenosphere remains stable at values of 10^{19} Pa s during the entire model history.

The vertically integrated second invariant of the deviatoric stress tensor, $\bar{\tau}_{II}$, is a measure for the strength of the lithosphere and twice its value is representative for the horizontal driving force (per unit length) during lithosphere extension and compression (appendix D). During the pure shear thinning phase in the first

3.2 Wet Olivine rheology: Model M6

Crustal necking in M6 starts at ca. 2 Myrs of extension, values for $2 \times \bar{\tau}_{II}$ in model history (fig. 4(a) & (e)). Two convecting cells develop in the horizontal centre of the domain transporting material from $z \approx$-200 km to $z \approx$-100 km (fig. 7(d) 4(b) & (f))reach 14^{-1}. At ca. 2-3 13 Myrs , values for $2 \times \bar{\tau}_{II}$ decrease below ca. convection cells are active in the upper 500 km of the domain (fig. 4(c)). Crustal thickness varies laterally between ca. 20 km and $<5^{-1}$. At the end of the extension period the values for $2 \times \bar{\tau}_{II}$ stabilise between ca. $2\text{-}3^{-1}$ for all models.

3.2.1 Wet Olivine rheology: Model M6

km (fig. 4(g)). The mantle lithosphere is thermally eroded, indicated by a rising level of the 10^{21} Pa s contour in fig. 4(f)-(h) after 26 Myrs in model history. In contrast to M1, M6 does not reach the stage of crustal break-up (fig. 4(a) -(dd) & (h)) within 30 Myrs. Both margins are connected by an approximately 100 wide zone of hyper-extended crust (Two large convection cells are active: in the left half of the domain convection occurs at relatively enhanced flow speeds, whereas in the right half of the domain flow speeds are relatively lower (compare relative length of velocity arrows in fig. 4(ed)). Values for The average Rayleigh number of the system at this stage is 9.26×10^{7}. Values for temperature at the Moho reach ca. 1000 °C locally (fig.

4(h) and fig. 2(a)). The horizontally-averaged density profile (fig. 2(b)) shows that values for density in the lithosphere are in average 100 kg m^{-3} lower in M6 compared to M1. Values for effective viscosity in the ~~upper mantle~~ asthenosphere decrease to minimal values ~~around~~ in the order of 10^{18} Pa s and increase up to values of 10^{19} Pa s at a depth of 660 km ~~in the early stage of the model evolution~~ at the end of the extension period (fig. ~~4(a~~2(c)). ~~The mantle lithosphere delaminates, indicated by a rising level of~~

3.3 Comparison of reference run with models M2-5: Extension phase

In contrast to M1, M2 produces two conjugate passive margins that are both approximately 150 km long (fig. 5(b)) after 13 Myrs. Crustal separation has not occurred up to this stage in M3: the two passive margins are still connected by a crustal bridge of ca. 10 km thickness. Mantle material rises below the centre of the domain and then diverges below the plates in M2 and M3, but no convection cells have formed yet (fig. 5(b) & (c)), which is different compared to M1. The minimal value for effective viscosity in the upper mantle is at the applied cut-off value of 1×10^{21} ~~^{20}~~ Pa s ~~contour in fig. 4(a) -(d). Values for viscosity~~ in ~~the entire upper mantle and transition zone fall to the applied lower viscosity limit of~~ M2 and 5×10^{18} ~~^{20}~~ Pa s ~~. After 26 in model history, the majority of the mantle lithosphere has been delaminated (see level of~~ in M3 and increases up to ca. 1×10^{21} Pa s ~~contour in fig. 4(d)). Values for temperature at the Moho reach ca. 1000 .~~

3.4 ~~Thermal relaxation~~

~~At the end of the thermal relaxation period (ca. 99), M1 develops convection cells in the upper mantle. The size of the cells varies between ca. 50~~ at a depth of 660 km ~~and~~ (fig. 2(c))in both models. Similar to the reference model (M1), in M4, the left continental margin has a length of ca. ~~300~~ 200 km ~~in diameter below the left and right margin respectively (fig.6(a~~and the right margin has a length of ca. 150 km (fig. 5(d)). Like in M1, two convection cells have formed below the two passive margins in this model (see arrows in fig. 5(c)) and the minimal value for effective viscosity below the lithosphere is in the order of 1×10^{19} Pa s and increases to approximately 1×10^{21} Pa s at a depth of 660 km (fig. 2(c)). ~~Vertical velocities are ca. 1-1.5 $^{-1}$. Below the right margin a downward directed material flow with ca. 2-3 $^{-1}$ develops (see dark blue cell at ca. +300~~ Both margins in M5 are approximately equally long (ca. 150 km~~in horizontal x-direction in fig.6~~, fig. 5(e)) and values for viscosity are at the lower cut-off value of 1×10^{20} Pa s in the upper mantle and increase up to ca. 1×10^{21} Pa s at a depth of 660 km (fig. 2(c)). The overall evolution of the extension period in M5 is more similar to M2 than to M1. In M1-5, the 1350 °C isotherm does not come closer than 30 km to the surface. Horizontally-averaged vertical temperature profiles are similar in M1-5 (fig. 2(a) ~~)~~& (g)). The level of the 1350 °C isotherm remains at its initial depth in M4 and M5, whereas it subsides by ca. 20 km in M1-3. ~~At a depth of~~ Horizontally-averaged density profiles (fig. 2(b)) show density differences of <10 kg m^{-3} between ca. 35 km and ca. 120 km ~~a laterally heterogeneous horizontal mantle flow field develops (~~depth.

3.4 Comparison of reference run with models M2-5: Cooling phase without plate deformation

Models M1-5 maintain a stable lithospheric thickness of 90-100 km over 100 Myrs (top magenta viscosity contour in fig. ~~6(f))~~ ~~showing a region of convergent flow directions below the right margin. Absolute values for velocities of mantle flow are ca. 2~~ ~~$^{-1}$. Material transport in M2 and M4 occurs with velocities below ca. 0.5 $^{-1}$ (fig.~~ 6(~~b),(d),(g) ,(i)). Essentially no convection cells develop in these models (see undeflected 1350 and 1450 isotherms~~ a)-(e)) and no thermal erosion of the lithosphere occurs. Below, the upper mantle convects at decreasing Rayleigh numbers from M1-5. In M1, the vertical mantle flow speed within the convecting cell at $x \approx +350$ km is elevated (indicated by darker blue coloured region in fig. 6(~~b)& (d)~~ a)) compared to the average flow speed of ≈1-2 cm yr^{-1} in neighbouring cells). The size of the convection cells in ~~M3 is~~ M2 is larger compared to M1 and in the order of ca. 100-300 km in diameter. ~~The cell size and the~~ Characteristic is the more elliptical shape of the cells compared to the circular cells in M1. The magnitude of material flow velocity ~~are distributed~~ is similar but distributed more horizontally symmetric below both margins ~~(~~compared to M1 (compare arrows and colour field of fig.6(~~e~~a) & (~~h~~b)). A zone of strong downward directed movement ~~with > 2 $^{-1}$ develops~~ develops below both margins (dark blue cells at ~~ca.~~ $x \approx$ -300 km and ~~ca.~~ $x \approx$ +300 km in ~~horizontal x-direction in~~ fig.6(~~c))below both margins. M5 develops two zones where the lithosphere is delaminating.~~ b)). Magnitude of material flow speed is in the order of 1.5 cm yr^{-1}. The average Rayleigh-number of the convecting system in M2 is approximately 3.92×10^6 and is about a factor 2.5 lower compared to M1. M3 develops four large convection cells, two below each margin, that are active up to depths of approximately 600 km (see arrows fig. 6(c)). Downward directed movement of material occurs with ca. 2 cm yr^{-1} (darker blue regions at ~~ca.~~ $x \approx$ -500 km and ~~ca.~~ $x \approx$ 500 km in ~~horizontal x-direction in~~ fig.6(~~d~~c)).

~~All models maintain a stable lithospheric thickness of 90-100 over 100 (magenta viscosity contour~~ The average Rayleigh number of the system is ca. 1.18×10^6, which is about a factor 8.4 smaller compared to M1. In M4 and M5, material transport occurs at absolute values for vertical velocity <0.5 cm yr^{-1} (see coloured velocity field in fig. 6~~) and no thermal erosion of the lithosphere occurs. Horizontally averaged temperature profiles remain similar for M1-5 (fig. 2~~(d) & (e)) which is one order of magnitude lower compared to M1-3. Two horizontally symmetric convection cells develop in these models that are active between $z \approx$ -150 km and $z \approx$ -400 km (arrows in fig. 6(d) ~~). Figure ?? &~~ (e)). The average Rayleigh number is ca. 1.97×10^6 and 5.48×10^5 in M4 and M5, respectively.

Figure 6(f)-(j) shows the conductive heat flow of the entire domain in absolute values. M1-5 reproduce a heat flux of 20-30 mW m^{-2} through the base of the lithosphere (indicated by the 10^{21} Pa s isopleth at a depth between 100-110 km). The conductive heat flow below the lithosphere is close to 0 mW m^{-2} in ~~M1, M3 and M5. In M2 and~~ M1-3. In M4 and M5, values for conductive heat flow remain at ca. 20 mW m^{-2} ~~in the~~ through the entire upper mantle. Density differences in the upper part of the mantle lithosphere reach ca. 20 kg m^{-3} between ca. 35-120 km in depth (fig. 2(h) & (j)). Values for effective viscosity range in the order of 10^{19} Pa s for M1 and M5 and in the order 10^{20} Pa s in ~~the upper mantle directly~~ M2-4 directly below the lithosphere and 10^{21} Pa s at the bottom of the upper mantle (fig. 2(f)).

~~Values of $2 \times \bar{\tau}_{II}$ remain constant at values between 2.5-3.0 $^{-1}$ for all models during the entire thermal relaxation period (fig. 7(d)). Because no far-field deformation is applied during the thermal relaxation, the stresses result only from convective flow and from lateral variations of gravitational potential energy (GPE) between elevated regions with continental crust and lower ones with exhumed mantle.~~

3.5 ~~Convergence and subduction~~

Figure 7(a) shows the topography at ca. ~~102~~ 99 Myrs in model history, that is ~~2~~ 1 Myrs ~~after~~ before the start of convergence. Topography does not exceed 1.5 km and the average depth of the basin is ca. ~~4~~ 3.75 km. ~~In~~

3.5 Comparison of reference run with models M2-5: Convergence and subduction phase

In contrast to M1~~and M2, a single, major shear zone develops in the lithosphere at the right continental margin (fig. 8(a) & (b)). The strain rate distribution is asymmetric (fig.8(a) & (b)). In M3-5~~ , two major symmetric shear zones develop in the lithosphere in M2, M3 and M5, one shear zone at each of the continental margins (fig.8(~~c)~~ b), (c) & (e)).

Like in M1, one shear zone forms below the right margin in M4. At this early stage of subduction initiation, the strain rate in the shear zone is in the order 10^{-14}-10^{-13} s^{-1}. In the region of the shear zones, the temperature is increased, which is indicated by the deflection of the isotherms (red contour lines in fig.8(a)-(e)). ~~Convection is still active in the upper mantle in M1 (strain rate distribution below viscosity contour in fig. 8(a)) . Horizontal profiles of GPE~~ At 103 Myrs, profiles for vertical velocities at $z \approx$-175 km (fig. 7(b)) show increasing symmetry ~~from M1 to M5 (fig. 7(c)), but the profiles are in general similar. Figure 7(b) shows a horizontal profile of the vertical velocity at 175 depth indicating a significant lateral asymmetry in M1~~with decreasing Rayleigh number (M1-5). Magnitudes of vertical velocity reach ca. -3 cm yr^{-1} at ~~x-position of ca. $x \approx$~~+300 km ~~.~~ in M1 (thick black line in fig. 7(b)). The vertical velocities are three times lower below the other margin. The distribution of vertical velocity in M2-5 is laterally ~~symmetric~~more symmetric compared to M1. Minimal values of approximately -2 cm yr^{-1} are reached in M2, M3 and M4 (~~red and black dashed lines in~~ fig. 7(b)), whereas in M5 minimal values of vertical velocity are ca. -1 cm yr^{-1}. ~~Values of $2\times\bar{\varepsilon}_{II}$ reach maximum values of ca. 22 $^{-1}$ in M5 and range between 12-20 $^{-1}$ in M1-4~~ Horizontal profiles of gravitational potential energy (*GPE*, see appendix D) are in general similar (fig. 7(~~d))~~ c)) for M1-5.

In ~~M1, one stable subduction zone develops below one of the continental margins and a trench forms in which sediments are deposited (fig.9(a)). Folding of the crustal layers in the overriding plate indicates significant deformation of the crust. The crustal layers of the subducting plate remain relatively undeformed. The viscosity in the upper mantle remains stable at values of 10^{19} during the entire model history. Also, the thermal thickness of the lithosphere remains stable and relatively constant over time. In M2-5, subduction is initiated~~ contrast to the single-slab subduction evolving in M1, double-slab subduction is observed below both margins in M2-5 and sediments are deposited in two trenches as the subduction evolves (fig.9(b)-(e)). ~~Alternating activity of the subduction zones is observed.~~ Folding of the crustal layers indicates deformation in both margins of M2-5. Values for viscosity in the upper mantle remain stable at values of 10^{19} Pa s in ~~M2~~M4, whereas the viscosity values are at the applied lower cut-off value in ~~M3-5~~ M2, M3 and M5 throughout the entire simulation history. ~~Subduction initiation and strain localization at ca. 102 is associated with a rapid decrease of $2\times\bar{\varepsilon}_{II}$ whereby values of $2\times\bar{\varepsilon}_{II}$ decrease by ca. 2-5 $^{-1}$ in all models. At ca. 105 , values of $2\times\bar{\varepsilon}_{II}$ increase again until the end of the simulation as shown in fig. 7(b). At this stage the vertical velocity field of~~ Convection cells remain active during subduction. Similar to M1~~shows only one active convection cell (fig. ??(a)). Convection is not active in M4 (cf. fig. ??(b~~, a stable convection cell is active below one subducting slab in M4. The distribution of convection cells is more symmetric in front of the slabs in M2, M3 and M5 compared to M1. Defined by

the applied boundary condition, material flows into the model domain up to z = -330 km at the lateral boundaries in all models (see section 2.1). Though, far away from the boundary in the centre of the domain in M1 and M2 the horizontal material inflow from the lateral boundaries is limited up to a depth of ca. 100-150 km (see horizontally directed arrows in fig. 9(a) & (eb)). The 10^{21} viscosity isopleth remains at a constant depth away from the subduction zone. In M3-5 the lateral inflow of material reaches a depth of about 200 km in the centre of the domain.

4 Discussion

3.1 Extension

Higher values for the lower viscosity cut-off (M3 compared to M1) decrease the mechanical heterogeneity in the crust. In consequence, the multi-layered crust necks effectively as a single layer. The resulting passive margins are shorter and more symmetric (Chenin et al., 2017, see also M3). The highest cut-off value of 5×10^{20} (M5) leads to a two-stage necking as investigated by Huismans and Beaumont (2011): first, the lithosphere is necking while the crust deforms by more or less homogeneous thinning, leading to the development of a large continuous zone of hyper-extended crust, below which the mantle lithosphere has been removed. Second, the hyper-extended crust is breaking up after the continental mantle lithosphere. Mantle convection does not establish as early as rifting and crustal separation. Therefore, the evolution of M1-4 is similar and the effective conductivity approach (M2, M4) leads to similar margin geometries as predicted by the realistic conductivity model (M1, M3).

In our models, we drive extension with horizontal boundary velocities that are constant during extension. Another end-member boundary condition is to assume a lateral constant extensional force (e.g. Brune et al., 2016). These different boundary conditions likely have a significant impact on the extensional evolution , but quantifying the impact of these boundary conditions is beyond the scope of our study.

3.1 Thermal relaxation

Vigour of the convecting system is described by its Rayleigh number. The average Rayleigh number (see appendix eq.B1) in M1 is

3.1 Estimates for plate driving forces

The vertically-integrated second invariant of the deviatoric stress tensor, $\bar{\tau}_{\mathrm{II}}$, is a measure for the strength of the lithosphere and twice its value is representative for the horizontal driving force (per unit length, F_{D} hereafter) during lithosphere extension and compression (appendix D). Figure 7(d) shows the evolution of F_{D} during the entire cycle. During the pure shear thinning phase in the first ca. 2 $\times10^{6}$ (see tab. 1) which is in the range of realistic values for upper mantle convection (Torrance and Turcotte, 1971). Heat flow through the lithosphere at a depth slightly above the convecting cells (ca. 100-110) is between 20-30 Myrs of extension, values for F_{D} reach 14 TN m^{-2} (fig. **??**(a))which is in the range of realistic values for heating at the base of the

lithosphere confirmed by other numerical studies (Turcotte and Schubert, 2014; Petersen et al., 2010). Below the lithosphere the conductive heat flux is essentially zero, because heat transport is mainly due to advection of material in the convecting cells. M3 and M5 produce similar values for $^{-1}$. At ca. 2-3 Myrs (1 in fig. 7(d)), this value drops below ca. 5 TN m^{-1}. At the end of the conductive heat flux as M1. The main difference between M1, extension period F_D is stabilised at values between ca. 2-3 TN m^{-1} for all models (2 in fig. 7(d)). This value remains relatively constant for all models during the entire thermal relaxation period. The maximum value for F_D necessary to initiate subduction in all models is observed in M3 and M5 is the speed at which material is transported: values for the Rayleigh number in M3 and M5 are ca. one order of magnitude lower compared to M1 (see tab. 1), because the higher viscosity limit raises values for viscosity in the upper mantle. The limited velocities in the upper mantle reduce the total number of time steps needed for the computation by ca. a factor 1.3 for M3 and is ca. 23 TN m^{-1} (thin blue curve in fig. 7(d)). The minimum value necessary for subduction initiation of ca. 13 TN m^{-1} is observed in M4. The reference run M1 initiates subduction with a value of $F_D \approx 17$ TN m^{-1}. Strain localization at ca. 102 Myrs is associated with a rapid decrease of F_D by ca. a factor 3.5 for M5 compared to the total number of time steps computed for M1. In M2 and M4, the Rayleigh number is also ca. one order of magnitude lower compared to M1 (see table 1) , due to the high effective thermal conductivity (included in κ in eq. B1) , which reduces the total number of time steps by a factor of 1.8 and 3.0 for M2 and M4, respectively, compared to M1. Although the effective conductivity approach maintains a reasonable heat flux at the base of the lithosphere (figs. **??**(b) & (d)) all processes in the upper mantle are conduction dominated. This results in values for the Rayleigh and Nusselt number of the system which are not representative for the Earth's mantle convection. 2-5 TN m^{-1} in all models (3 in fig. 7(d)). At ca. 105 Myrs, values for F_D increase again until the end of the simulation.

4 Discussion

4.1 Spontaneous vs. induced subduction initiation and estimates for plate driving forces

Currently, the processes and tectonic settings leading to subduction initiation remain unclear (Stern, 2004; Stern and Gerya, 2018; Crameri
Stern (2004) proposed two fundamental mechanisms for subduction initiation:

(1) Spontaneous (or vertically forced, Crameri et al. (2020)) subduction initiation occurs, for example, due to densification of the oceanic lithosphere during secular cooling. Cloos (1993) proposed that the density increase of cooling oceanic lithosphere compared to the underlying asthenosphere is in the order of 80 kg m^{-3}. According to Cloos (1993), this difference is sufficient to initiate subduction spontaneously by negative buoyancy of the oceanic lithosphere (see Stern (2004); Stern and Gerya (2018, & references therein) for detailed explanation). However, McKenzie (1977) and Mueller and Phillips (1991) showed using analytical solutions that the forces acting on the lithosphere due to buoyancy contrasts are not high enough to overcome the strength of the cold oceanic lithosphere. Observations of old plate ages (>100 MyrsMyrs) around passive margins in the South Atlantic indicate long-term stability of old oceanic lithosphere.

In the models presented here, the applied thermodynamic density of the Hawaiian pyrolite leads to density differences between the exhumed mantle in the basin and the underlying asthenosphere of ca. 10-20 kg m^{-3} across a vertical distance of 85 after 99 Myrs in model history. Compared to the buoyancy contrast proposed by Cloos (1993) the density differences

presented here are by a factor These buoyancy contrasts are ca. 4smaller, which is, however, consistent with more recent estimates presented by Afonso et al. (2007)× smaller than the contrast proposed by Cloos (1993). Boonma et al. (2019) report a density difference of +19 $\mathrm{kg\ m^{-3}}$ for an 80 km thick continental lithosphere and +17 $\mathrm{kg\ m^{-3}}$ for an oceanic lithosphere of 120 Myrs of age, which is . These values are in agreement with the values we report in our study. In the 2D models presented here, these buoyancy contrasts are insufficient to overcome the internal strength of the cooled oceanic lithosphere and initiate subduction spontaneously. Our results are, hence, in agreement with the stability of old oceanic lithosphere in, for example, the South Atlantic, where even an additional ridge push of approximately 3.9 $\mathrm{TN\ m^{-1}}$ (Turcotte and Schubert, 2014) was not sufficient to initiate subduction.

4.2 Subduction

Figure 7(a) shows the topographic elevation at the stage of subduction initiation. In all modelsend of the cooling period. In M1-5, the difference between the lowest point in the trenches basin and the highest point of the margins is ca. 5 km which is predicted by the principle of isostasy for a ca. 30 km thick crust with an average density of 2800 $\mathrm{kg\ m^{-3}}$ (Turcotte and Schubert, 2014). Therefore, the models are in a state of isostatic equilibrium at the onset of subduction initiation.

Currently, the processes leading to subduction initiation remain elusive (Stern, 2004; Stern and Gerya, 2018; Crameri et al., 2019). Stern (2004) proposed two fundamental mechanisms for subduction initiation: (1) spontaneous and (topography of the models at the end of cooling is close to the topography predicted by isostatic equilibrium between continent and exhumed mantle.

(2) induced subduction initiation . Spontaneous subduction initiation Induced (or horizontally forced, Crameri et al. (2020)) subduction initiation occurs, for example, due to densification of the oceanic lithosphere during secular cooling. Induced subduction initiation occurs, for example, due to far-field plate motion. In fact, many numerical studies that investigate subduction processes do not model the process of subduction initiation. In these studies, either a major weak zone across the lithosphere is usually prescribed ad-hoc in the initial model configuration to enable subduction (Ruh et al., 2015; Zhou et al., 2020)or the initial model configuration includes already . Another possibility to model subduction is to include a prescribed slab so in the initial configuration. This means that subduction has already initiated at the onset of the simulation (Kaus et al., 2009; Garel et al., 2014; Dal Zilio et al., 2018)(Kaus et al., 2009; Garel et al., 2014; Holt et al., 2017; Dal Zilio et al., 2018). In our models, subduction is initiated self-consistently, without prescribing which means here: (i) we do not prescribe any major weak zone or an already existing slab. The , and (ii) the model geometry and temperature field at the onset of convergence were also simulated in a self-consistent way by the previous modelling of extension and thermal relaxation. During convergence(100 Myrs) were simulated during a previous extension and cooling phase with the same numerical model and parameters. Subduction is initiated during the initial stages of convergence and subduction initiation is, hence, induced by horizontal shortening. Notably, Crameri et al. (2020) analysed more than a dozen documented subduction zone initiation events from the last hundred million years and found that horizontally forced subduction zone initiation is dominant over the last 100 Myrs. During convergence in our models, shear heating together with the temperature and strain rate dependent viscosity formulation (dislocation creep flow law, eq. A6) causes the spontaneous generation of a lithosphere-scale shear zone that evolves into a subduction zone (Thielmann and Kaus, 2012; Kiss et al., 2020). However, shear heating is a transient process

which means the increase in temperature is immediately counterbalanced by thermal diffusion. Efficiency of shear heating is restricted to the first ca. 2-3 Myrs after shear zone formation in the presented models. After this time span, heat generated by mechanical work is diffused away~~and single-slab subduction may not be stable anymore (see M2-5) . This means that~~ . Therefore, thermal softening is unlikely the mechanism responsible for stabilisation of long-term ~~occurrence of a single-slab subductionin~~ subduction zones, but likely important for initiating and triggering subduction zones.

The value of F_D (see appendix D) represents the plate driving force. In M1~~is likely caused by a different mechanism than shear heating.~~ , the maximal value of F_D just before stress drop caused by subduction initiation at ≈103 Myrs is ≈17 TN m^{-1} (3 in fig. 7(d)). However, $F_D \approx 2$ TN m^{-1} at the end of the cooling stage, resulting from stresses due to mantle convection and lateral variation of *GPE* between continent and basin. Therefore, this value can be subtracted from the ≈17 TN m^{-1}. The required plate tectonic driving force for convergence-induced subduction initiation is then ≈15 TN m^{-1} for M1. Kiss et al. (2020) modelled thermal softening induced subduction initiation during convergence of a passive margin, whose geometry and thermal structure was generated ad-hoc as initial model configuration. Their initial passive margin structure was significantly less heterogeneous than ours. To initiate subduction, they needed a driving force of ≈37 TN m^{-1} which is significantly larger than the ≈15 TN m^{-1} required in our model M1. We suggest that mechanical and geometrical heterogeneity inherited by previous deformation periods and thermal heterogeneity due to mantle convection reduces the required driving force necessary for subduction initiation by thermal softening. We suppose that even more heterogeneous models with higher numerical resolution than ours would require even less driving force than ≈15 TN m^{-1} to initiate subduction by thermal softening. This is relevant, because the main argument against thermal softening as an important localization mechanism during lithosphere strain localization and subduction initiation is commonly that the required stresses, and hence driving forces, are too high. In nature, more softening mechanisms act in concert with thermal softening, such as grain damage (e.g. Bercovici and Ricard, 2012; Thielmann and Schmalholz, 2020),

fabric and anisotropy evolution (e.g. Montési, 2013) or reaction-induced softening (e.g. White and Knipe, 1978), likely further reducing forces required for subduction initiation. Furthermore, Mallard et al. (2016) show with 3D spherical full-mantle convection models that a constant yield stress between 150 and 200 MPa in their outer boundary layer, representing the lithosphere, provides the most realistic distribution of plate sizes in their models. The yield stress in their models corresponds to a deviatoric, von Mises, stress which is comparable to the value of τ_{II} calculated for our models. If we assume a 100 km thick lithosphere, then F_D = 15 TN m^{-1} yields a vertically-averaged deviatoric stress, τ_{II}, of 75 MPa for our model lithosphere. Therefore, vertically-averaged deviatoric stresses for our model lithosphere are even smaller than deviatoric, or shear, stresses employed in global mantle convection models. Based on the above-mentioned arguments, we propose that ≈15 TN m^{-1} is a feasible value for the horizontal driving force per unit length.

4.2 Mantle convection stabilising single-slab subduction

Figure 7(c) shows the difference in ~~GPE~~*GPE* at 103 Myrs. The horizontal profiles do not reveal significant differences between the models. The ~~GPE~~ *GPE* is sensitive to topography and density distribution throughout the model domain. ~~The convection induced heterogeneity in the density field of the upper mantle are confined to a thin layer and the magnitude of these differences~~

~~is only~~ Values for density differences are ca. 10-20 $\mathrm{kg\ m^{-3}}$ across the mantle lithosphere. When integrated over the entire depth, such variations do not significantly impact the ~~GPE~~ *GPE* profiles. Instead, the signal reflects the topography, which is similar in all the models due to the similar margin geometry (see fig. 7(a) and fig. 5(a)-(e)). ~~However, the horizontal profile at a depth of 175 of the vertical velocity (fig.7(b)) reveals a possible explanation for why single-slab subduction is stable~~

The average Rayleigh number (see appendix eq.B1) in M1 ~~. The increased lateral heterogeneity~~ is ca. 9.95×10^6 (see tab. 1) which is close to estimated values for upper mantle convection (Torrance and Turcotte, 1971; Schubert et al., 2001). Potential lateral heterogeneity caused by the convecting cells in the upper mantle is inherited from the ~~convection active during the thermal relaxation~~ cooling period. The vertical velocity field at the end of the ~~thermal relaxation~~ cooling period of M1 (fig. 6(a)) reveals ~~a zone of convection with velocities~~ one convection cell at $x\approx+350$ km with flow speeds in the order of the convergence velocity applied later~~in the upper mantle at ca. +350 in horizontal x-direction. Horizontal velocities~~ . This convection cell is active directly below the ~~mechanical lithosphere (magenta isopleth in fig. 6(f)) indicate that mantle drag in this region forms a zone of locally convergent mantle flow in proximity of the future subduction zone~~passive margin at which subduction will be initiated later in the model and probably provides an additional suction force on the lithospheric plate. Conrad and Lithgow-Bertelloni (2002) quantified the importance of slab-pull vs. slab-suction force and showed that the slab pull force and the suction force of a detached slab sinking into the mantle induces similar mantle flow fields. They argued that a detached fraction of a slab sinking into the mantle can exert shear traction forces at the base of the plate and drive the plate. Baes et al. (2018) showed with numerical simulations that sinking of a detached slab below a passive margin can contribute significantly to the initiation of subduction. We ~~argue that the down-welling mantle region~~ suggest that the circular convection cell observed in M1 and the associated mantle flow also generates a force similar to a slab suction force. To quantify the suction force induced by the down-welling in the convection cell of M1, we calculated the difference of the density in this region with respect to the corresponding initial model density and integrated the density difference spatially over this region. The resulting buoyancy force per unit length is ca. 40 $\mathrm{TN\ m^{-1}}$, which is in the order of the slab pull force (Turcotte and Schubert, 2014). It is therefore likely that the suction force induced by the convection assists to initiate the subduction and to stabilise the single-slab subduction in simulation M1.

~~The value of $2\times\bar{\tau}_{\mathrm{II}}$ is identical to the vertical integral of the difference between the horizontal total stress and the lithostatic pressure, if shear stresses are negligible (Schmalholz et al., 2019). During onset of convergence, when the lithosphere is more or less homogeneously shortened, shear stresses are most likely negligible compared to normal deviatoric stresses. The vertical integral of the difference between the horizontal total stress and the lithostatic pressure represents one of the horizontal driving forces (per unit length)of tectonic plates (Molnar and Lyon-Caen, 1988; Schmalholz et al., 2019). Hence, during convergence before subduction initiation the value of $2\times\bar{\tau}_{\mathrm{II}}$ represents the plate driving force. In model M1, the maximal value of $2\times\bar{\tau}_{\mathrm{II}}$ just before stress drop caused by subduction initiation at ≈103 Myrs is ≈17 $^{-1}$. However, $2\times\bar{\tau}_{\mathrm{II}}$ =≈ 2 $^{-1}$ at the end of the cooling stage, resulting from stresses due to mantle convection and lateral variation of GPE between continent and basin. Therefore, the ≈2 $^{-1}$ can be subtracted from the ≈17 $^{-1}$ so that the required plate tectonic driving force for subduction initation is ≈15 $^{-1}$ in model M1. Although this value could be further reduced if additional softening mechanisms, such as grain damage or serpentinization during extension, would be considered in the model, we argue that ≈15 $^{-1}$ is a feasible value.~~

~~Convection is a physical mechanism that can provide additional forces locally to develop and stabilise a single-slab subduction. Double-slab subduction (Holt et al., 2017) is initiated below both of the margins (fig~~Values for the Rayleigh number in M2 and M3 are lower (3.92×10^6 and 1.18×10^6 respectively) compared to M1 (see tab. 1), because of the relatively higher viscosity (see eq. B1). ~~8(c)-(e)~~With decreasing Rayleigh-number the asymmetry of convection also decreases (Schubert et al., 2001). In consequence, the size of the convecting cells becomes larger and more elliptic (M2 compared to M1) and the number of active cells decreases (M3 compared to M1). ~~These subduction zones are stable only for a short period and the subduction alternates from one margin to the other with ongoing convergence~~Enhancing the thermal conductivity (included in the denominator in eq. B1) by ca. one order of magnitude in models M4 and M5 also decreases the Rayleigh number by one order of magnitude compared to M1 (see table 1). Decreasing the Rayleigh-number (M2-5) leads to more laterally symmetric convective flow patterns and decreases the speed and distribution of material flow (see fig. 7(b)). In turn, this probably leads to more equally distributed suction forces induced by the cells. Hence, we argue that decreasing the Rayleigh-number, assuming a relatively strong asthenosphere or applying an enhanced thermal conductivity likely favours the initiation of divergent double-slab subduction. The resulting slab geometries resemble a symmetric push-down (M2-5), rather than a stable asymmetric subduction ~~in M2-5~~(M1). This impacts the deformation in the lithosphere, especially in the crust: the crustal layers of both plates are strongly folded (see fig. 9(b)-(e)). In M1, the deformation of the crustal layers only occurs in the overriding plate. ~~Folding of the crustal layers might be the reason why values of $2\times\bar{\tau}_{II}$ increases again after subduction initiation. Although folding is a structural softening mechanism (e.g. Schmalholz et al., 2005; Schmalholz and Mancktelow, 2016), when the folds become more and more isoclinal, the stresses in the crust can be transmitted and are built up again. We, therefore, also stopped the simulations once the crustal folds were essentially isoclinal, because crustal deformation becomes unrealistic. The slightly higher temperatures induced by the effective conductivity reduce the stresses which~~

4.3 Onset of upper mantle convection and thermo-mechanical evolution of the lithospheric plates

Rifting in M1-5 causes up-welling of hot asthenospheric material in the horizontal centre of the domain. The resulting lateral thermal gradients are high enough to induce small-scale convection (Buck, 1986). Huang et al. (2003) derived scaling laws to predict the onset time for small-scale convection in 2D and 3D numerical simulations. They investigated the impact of plate motion, layered viscosity, temperature perturbations and surface fracture zones on the onset time. The observed onset time in layered viscosity systems becomes larger when the thickness of the weak asthenosphere decreases. Also, plate motion can delay the onset of small-scale convection. In contrast, fracture zones at the surface may lead to earlier onset of convection depending on the thermal structure of the fracture zones. Using their scaling law and the parameters used in our reference model M1, onset time of convection is predicted for ca. 43 Myrs. However, onset of convection in M1 is observed as early as ca. 8 Myrs (see fig. 3) which is consistent with onset times observed in numerical simulations conducted by Van Wijk et al. (2008). There may be several reasons for the discrepancy between the prediction and observation. First, the models presented here are likely a combination of the configurations tested by Huang et al. (2003). Second, the choice of boundary conditions and initial configuration is different which is probably important when testing the impact of plate motion on the onset time. They located the rift centre at the left lateral boundary, whereas the rift centre in our models is located far away from the lateral boundaries in

the horizontal centre of the domain. This likely impacts the flow direction of hot material ascending beneath the rift centre and, consequently, alters lateral thermal gradients which are important for the onset of convection. The onset time for convection is delayed and at ca. 20 Myrs in ~~turn might explain the lower values of $2\times\bar{\tau}_{II}$~~ M2 and at 30 Myrs in M3. This delay is likely due to the increased viscosity in the asthenosphere in M2 & M3 compared to ~~values obtained for M1. Also, the higher viscosity limits of M3 and~~ M1 which decreases the Rayleigh number of the system. This observation is in agreement with the general inverse proportionality of the onset time of convection to the Rayleigh number as predicted by Huang et al. (2003).

Once convection has started, it stabilises the temperature field over large time scales and controls the thickness of the lithosphere (Richter, 1973; Parsons and McKenzie, 1978). Although we allow for material inflow up to z = -330 km, the arrows in fig. 9(a) & (b) indicate that lateral inflow of material far away from the boundary is limited up to z $>\approx$ -150 km. Below this coordinate, the convection cells transport material even towards the lateral boundary, in the opposite direction of material inflow. This observation suggests, that the thermal thickness of the lithosphere adjusts self-consistently and far away from the boundary during the geodynamic cycle. The thermal thickness of the lithosphere seems to vary with decreasing Rayleigh number: in M5 ~~increase the stresses which results in higher values of $2\times\bar{\tau}_{II}$ compared to M1. These observations indicate that the~~ the value for the thermal thickness of the lithosphere is $\approx$ 200 km (see arrows in fig. 9(e)). This observation suggests that the thermal thickness of the moving lithosphere is presumable regulated by the vigour of convection.

For realistic values for thermal conductivity in the upper mantle (M1-3), heat flow through the base of the mechanical lithosphere is between 20-30 mW m^{-2} (fig. 6(f)-(h)). These values are in the range of heat flow estimates for heating at the base of the lithosphere confirmed by other numerical studies (Turcotte and Schubert, 2014; Petersen et al., 2010). Below the lithosphere, the conductive heat flux is essentially zero, because heat transport is mainly due to advection of material in the convecting cells. The effective conductivity approach ~~and high viscosity limits~~ maintains a reasonable heat flux directly at the base of the lithosphere (figs. 6(i) & (j)) and convection cells develop. However, all processes in the upper mantle are conduction dominated (see elevated values for q_z in fig. 6(i) & (j)). This implies that the characteristic physics of mantle convection are not captured correctly in these models. This becomes evident when comparing fig. 6(a) to fig. 6(d): although the physical parameters - except the thermal conductivity in the upper mantle - are the same in both simulations and temperature profiles are similar (see fig. 2(d) & (i)), the Rayleigh-numbers of the systems differ by one order of magnitude (see tab. 1) and the convective patterns are entirely different.

4.4 Impact of mantle viscosity structure and effective conductivity on passive margin formation

Higher values for the lower viscosity cut-off (M2 & M3 compared to M1) do not only ~~impact the style of subduction~~change the viscosity structure of the mantle, but also ~~impact on the style of deformation in the crust during convergence~~

~~The onset of convergence in the simulations has a strong impact on the convection, which can be observed in the change of~~ increase the strength of the ~~velocity field between 99 and 118 in M1~~ weak layers. In consequence, the multi-layered crust necks effectively as a single layer. The resulting passive margins are shorter and more symmetric (Chenin et al., 2017, see also M3). The highest cut-off value of 5×10^{20} Pa s (M3) leads to a two-stage necking as investigated by Huismans and Beaumont (2011). First, the lithosphere is necking while the crust deforms by more or less homogeneous thinning, leading to the development

of a large continuous zone of hyper-extended crust, below which the mantle lithosphere has been removed. Second, the hyper-extended crust is breaking up later than the continental mantle lithosphere. During the rifting stage, the thermal field in all simulations is very similar (see fig. ~~??(b and c)). Not only the subducting slab may be responsible for changing the convection dynamics, but also the relative horizontal motion of the lithosphere~~ 2(a)), presumably because effects of heat loss due to diffusion are not significant over the relatively short time scale.

4.5 Mantle convection, thermal erosion and tectonics in the Archean

Due to significantly reduced viscosities (see fig. 6(c)), convection in the upper mantle occurs at an average Rayleigh number of ca. 9×10^7 in ~~general. The shear stress induced by the plate movement leads to the decay of the amplitude of these convection cells (e.g. Richter, 1973). Richter (1973) showed analytically that the temporal stability of transversal Rayleigh-Bénard convection cells underneath a moving plate is below~~ M6. Compared to estimates for present-day Earth's upper mantle convection (Torrance and Turcott Rayleigh-number is an order of magnitude higher.

The lithosphere is recycled rapidly and the resulting values for density directly below the lithosphere are ca. 100 kg m^{-3} lower in M6 compared to M1-5. The resulting density structure in the upper region of the mantle deviates significantly from the PREM model (Dziewonski and Anderson (1981), see fig. 2(b)). Convection at such high Rayleigh numbers leads to an enhanced temperature field. Resulting values for temperature at the Moho locally reach ca. 1000 °C. The scenario calculated in this model does not correspond to present-day observations. In the Archean eon, the mantle potential temperature was probably 200-300 °C higher (Herzberg et al., 2010) and therefore convection was more vigorous (Schubert et al., 2001). Agrusta et al. (2018) have investigated the impact of variations in mantle potential temperature compared to present-day's value on slab-dynamics. Assuming a 200 ~~when plate velocities are larger than 2 $^{-1}$. Three dimensional numerical simulations showed that Rayleigh-Bénard convection cells with axis parallel to the plate motion are stable below a moving plate (Ballmer et al., 2007), but do, of course, not exist in our 2D models~~°C higher value for mantle potential temperatures compared to present-day estimates leads to viscosity structures similar to the average viscosity structure we report for M6 (compare fig. 6c in Agrusta et al. (2018) to fig. 2(c) of this study). Hence, we argue that such vigorously convecting systems as presented in M6 may be applicable to the mantle earlier in Earth's history.

4.6 ~~Applications~~

4.6 Comparison with estimates of Earth's mantle viscosity and thermal structure

To apply our models to geodynamic processes on Earth, we compare several model quantities with measurements and indirect estimates of these quantities. ~~The~~ Even after 118 Myrs of model evolution, the mantle density structure of our model ~~is~~ remains in good agreement with the preliminary reference Earth model (PREM Dziewonski and Anderson (1981), see fig. 2(b),(e)). The geotherm of the conduction dominated regime ~~lies~~ remains well in the range of pressure-temperature (~~P-T~~*P-T*) estimates from mantle xenolith data~~and~~ . Also, the geotherm of the convection dominated regime ~~falls within the realistic~~ remains within the range for adiabatic gradients and potential ~~mantle temperatures~~ temperatures applicable to the Earth's mantle (Hasterok

and Chapman (2011), see fig. 2(a),(d)) during the entire long-term cycle. The heat flux at the base of the lithosphere is in a ~~realistic~~ range of 20-30 $\mathrm{mW\ m^{-2}}$. ~~The viscosity~~ Viscosity profiles lie within the range of estimates inferred by inversion of observable geophysical data and from experimentally determined flow law parameters of olivine rheology (Mitrovica and Forte (2004); Behn et al. (2004); Hirth and Kohlstedt (2003), see fig. 2(c),(f)). ~~The lithosphere~~ Lithosphere and mantle velocities are in a ~~realistic~~ range of several ~~. The first order lithosphere geometry of the modelled hyper-extended margins and the geometry of the subduction zone seem realistic~~$\mathrm{cm\ yr^{-1}}$ which is in agreement with predictions from boundary layer theory of layered mantle convection (Schubert et al., 2001) and plate velocity estimates from GPS measurements (Reilinger et al., 2006).

4.7 Formation and reactivation of magma-poor rifted margins: potential applications

Our model can be applied to some first-order geodynamic processes that were likely important for the orogeny of the Alps. Rifting in the Early to Middle Jurassic (Favre and Stampfli, 1992; Froitzheim and Manatschal, 1996; Handy et al., 2010) lead to the formation of the Piemont-Liguria ocean which was bounded by the hyper-extended magma-poor rifted margins of the Adriatic plate and the Briançonnais domain on the side of the European plate. We follow here the interpretation that the Piemont-Liguria ocean was an embryonic ocean which formed during ultra-slow spreading and was dominated by exhumed subcontinental mantle ~~(e.g. Picazo et al., 2016; McCarthy et al., 2018; Chenin et al., 2019).~~ (e.g. Picazo et al., 2016; McCarthy et al., 2018; Chenin et al., 201

If true, there was no stable mid-ocean ridge producing a several 100 ~~km~~ km wide ocean with a typically 8 ~~km~~ km thick oceanic crust and our model would be applicable to the formation of an embryonic ocean with exhumed mantle bounded by magma-poor hyper-extended ~~passive marings~~rifted margins. Our models show the formation of a basin with exhumed mantle bounded by hyper-extended margins above a convecting mantle. Hence, our model may describe the first-order thermo-mechanical processes during formation of an embryonic ocean. During closure of the Piemont-Liguria ocean, remnants of those magma-poor ocean-continent transitions escaped subduction and are preserved in the Eastern Alps (Manatschal and Müntener, 2009). We follow the interpretation that the initiation and at least the early stages of ocean closure were caused by far-field convergence between the African and European plates (e.g. Handy et al., 2010). We further assume that subduction was induced, or horizontally forced, by this convergence and was not initiated spontaneously due to buoyancy of a cold oceanic lithosphere (De Graciansky et al., 2010). This interpretation agrees with recent results of Crameri et al. (2020) who suggest that most subduction zones that formed during the last 100 Myrs were likely induced by horizontal shortening. Our models show that a cooling exhumed mantle does not subduct spontaneously because buoyancy forces are not significant enough to overcome the strength of the lithosphere. However, the models show that convergence of the basin generates a horizontally forced subduction initiation at the hyper-extended passive margin causing subduction of the exhumed mantle below the passive margin. Such subduction initiation at the passive margin agrees with geological reconstructions which suggest that the Alpine subduction initiated at the hyper-extended margin of Adria (e.g. Manzotti et al., 2014) and not, for example, within the ocean. Overall, the here modelled, more than 100 ~~Myrs~~ Myrs long, geodynamic cycle ~~of (1) lithosphere extension generating hyper-extended passive margins and a basin with exhumed mantle, (2) lithosphere cooling for 70 Myrs without spontaneous subduction initiation and (3) convergence leading to forced subduction initiation at one of the hyper-extended margins is~~ is thus

in agreement with several geological reconstructions of the Alpine orogeny. However, the Alpine orogeny exhibits a distinct three-dimensional evolution including major stages of strike-slip deformation and a considerably radial shortening direction so that any 2D model can always only address the fundamental aspects of the involved geodynamic processes.

During convergence, several of our models show the formation of a divergent double-slab subduction. Such divergent double-slab subduction likely applies to the ~~Eastward and Westward~~ eastward and westward dipping subduction of the Adriatic plate (Faccenna and Becker, 2010; Hua et al., 2017). Another example of divergent double-slab subduction is presumably the Paleo-Asian Ocean, which has been subducted beneath both the southern Siberian Craton in the north and the northern margin of the North China Craton in the south during the Paleozoic (Yang et al., 2017). Furthermore, a divergent double-slab subduction was also suggested between the North Qiangtang and South Qiangtang terrane (Li et al., 2020; Zhao et al., 2015). Our models show that a divergent double-slab subduction is a thermo-mechanically feasible process during convergence of tectonic plates.

Tomographic images from the Mediterranean show large p-wave anomalies in the transition zone (Piromallo and Morelli, 2003) indicating that the ~~660~~ 660-km phase transition inhibits the sinking material to penetrate further into the lower mantle. This observation suggests that convection in the Alpine-Mediterranean region could be two-layered ~~. Also, mantle convection seems active~~ and largely confined to the upper region of the upper mantle. The convective patterns resulting from the presence of a weak asthenosphere simulated in our study are in agreement with these observations. We speculate that upper mantle convection might have played a role in the formation of the Alpine orogeny in the form of inducing an additional suction force below the Adriatic margin and assisting in the onset of subduction.

4.8 ~~Modelling geodynamic cycles and structural inheritance~~

~~We modelled here a geodynamic cycle of subsequent lithosphere hyper-extension with mantle exhumation, thermal relaxation and convergence inducing subduction, all coupled to upper mantle convection, because we wanted to test whether a single and continuous numerical simulation can generate the first order features of these different processes. As we discuss above, our simulation captures the fundamental features of such cycle with model quantities that are applicable to Earth. We argue that the ability of a numerical algorithm with a specific model configuration to generate several geodynamic processes with results applicable to Earth provides further confidence that the model has captured correctly the first order physics of the investigated processes.~~

~~Geodynamic processes, such as lithosphere extension or convergence, are frequently studied separately. For example, subduction initation at a passive margin during convergence can be studied without a previous extension and cooling stage. An initial passive margin geometry and thermal field must be then constructed ad-hoc for the model configuration. However, it is then uncertain whether the applied model would have generated such a margin geometry and thermal structure during an extension simulation. In other words, it is uncertain whether the initial margin configuration is consistent with the applied model. We modelled here passive margin formation and thermal relaxation before modelling convergence, because we wanted to model convergence for a passive margin geometry and thermal structure that arises in a self-consistent manner from the applied algorithm and model parameters. Compared to ad-hoc passive margin geometries and temperatures, our modelled margin geometry is much more heterogeneous. This heterogeneity, or structural inheritance, may have an impact on the subduction~~

initiation. For example, Kiss et al. (2020) modelled thermal softening induced subduction initiation during convergence of a passive margin, whose geometry and thermal structure was generated ad-hoc as initial model configuration. Their initial passive margin structure was significantly less heterogeneous than ours. Their model is not directly comparable with ours, because they considered, amongst others, a homogeneous and not a layered crust and not identical model parameters. However, they applied the same numerical algorithm. To initiate subduction, they needed a driving force of $\approx 37^{-1}$ which is significantly larger than the $\approx 15^{-1}$ required in our model M1. We argue that the larger the mechanical, geometrical and thermal heterogeneities are, the smaller may be the required driving force to model subduction initiation by thermal softening. Structural inheritance, in the form of mechanical and geometrical heterogeneities, cause stress concentrations that can act as trigger for strain localization and shear zone formation. We speculate that even more heterogeneous models with higher numerical resolution than ours would require even less driving force than $\approx 15^{-1}$ to initiate subduction by thermal softening. This is relevant, because the main argument against thermal softening as an important localization mechanism during lithosphere strain localization and subduction initiation is commonly that the required stresses, and hence driving forces, are too high. In nature, more softening mechanisms act in concert with thermal softening, such as grain damage (e.g. Bercovici and Ricard, 2012), fabric and anisotropy evolution (e.g. Montési, 2013) or reaction-induced softening (e.g. White and Knipe, 1978). Including such additional processes in our mantle-scale numerical models means to include more parameterizations based on not well constrained parameters and to make more assumptions concerning the up-scaling of such pore- and mineral-scale processes to the applied numerical resolution adapted to the lithosphere and mantle scale. We, therefore, argue that our model is simple but applicable to first-order processes of lithosphere hyper-extension, upper mantle convection and induced subduction initiation; and is a starting algorithm for further elaboration with upscaled micro-scale processes.

5 Conclusions

Our 2D thermo-mechanical numerical model of coupled lithosphere-mantle deformation is able to generate a 120 Myrs long geodynamic cycle of subsequent extension (30 Myrs), cooling (70 Myrs) and convergence (20 Myrs) in a single and continuous simulation. During lithosphere extension, our models generate an approximately 400 wide basin of exhumed mantle bounded by The simulations capture the fundamental features of such cycles, such as formation of hyper-extended magma-poor passive margins. Physical fields, such as temperature, density and viscosity structure produced by our models are in good agreement with petrological data, data inferred from geophysical observable quantities and from laboratory experiments. The models show that considering only the thermal effects of upper mantle convection by using an effective thermal conductivity generates results of lithosphere hyper-extension that are similar to the ones of models that explicitly model the convective flow. The inferred viscosity structure in the mantle, parameterized by different effective viscosity cut-offs has a significant impact on the simulation results. After margins, upper mantle convective flow or subduction initiation, with model outputs that are applicable to Earth. We propose that the ability of a model to generate such long-term cycles in a continuous simulation with constant parameters provides further confidence that the model has captured correctly the first-order physics.

In our models, subduction at a hyper-extended passive margin is initiated during horizontal shortening and by shear localization due to mainly thermal softening. In contrast, after 70 Myrs of cooling without far-field deformation, subduction of ~~the~~ a 400 km wide exhumed and cold mantle ~~was~~ is not spontaneously initiated. The buoyancy force due to the density difference between lithosphere and ~~underlying mantle~~ asthenosphere is too small to ~~generate a downward pulling buoyancy force that can exceed~~ overcome the mechanical strength of the ~~cold lithosphere. The modelled extension and cooling stages generate a structural and thermal inheritance, in the form of mechanical, geometrical and thermal heterogeneities, in a self-consistent way for the subsequent convergence stage. Convergence of the basin initiates subduction of the exhumed mantle at the transition to the~~ lithosphere.

Our models show that the viscosity structure of the asthenosphere and the associated vigour of upper mantle convection has a significant impact on lithosphere dynamics during a long-term geodynamic cycle. In comparison to a strong asthenosphere with minimum viscosities of 5×10^{20} Pa s, a weak asthenosphere with minimum viscosities of ca. 10^{19} Pa s generates the following differences: (1) More asymmetric hyper-extended ~~margins. The main mechanism of strain localization causing subduction initiation is thermal softening. If convection in the mantle is suppressed by high effective thermal conductivities or high, lower viscosity limits, then subduction initiates at both margins leading to divergent double-slab subduction. Thermal convection in the mantle assists to generate~~ margins during extension, (2) locally larger suction forces due to convective flow, which are able to assist in establishing a single-slab subduction ~~at only one margin, likely due to mantle flow which exerts an additional suction force on the lithosphere. Subduction is initiated during convergence for plate driving forces of approximately~~ instead of a double-slab subduction, and (3) smaller horizontal driving forces to initiate convergence-induced subduction, namely ca. 15 TN m^{-1} instead of ca. 22 TN m^{-1}. Therefore, quantifying the viscosity structure of the asthenosphere is important for understanding the actual geodynamic processes acting in specific regions. ~~Overall, our models suggest that upper mantle convection can have a considerable impact on lithosphere deformation, mainly due to mantle flow and associated suction forces.~~

~~For conditions applicable to Earth, our models generate the first-order, large-scale features of hyper-extension and subduction initiation. On the crustal and lithospheric scales, structural softening, due to non-linear flow laws and mechanical heterogeneities, and thermal softening, due to energy conservation and temperature-dependent viscosities, are softening mechanisms that are sufficient to generate these first-order features. We, hence, argue that these two softening mechanisms are important for lithosphere deformation.~~

~~Our models are also useful to test mantle flow laws, or flow law parameters, because flow laws generating too low effective viscosities, in our case for wet olivine, generate too vigorous convection associated with an unrealistic thermal erosion of the lithosphere.~~

The first-order geodynamic processes simulated in the geodynamic cycle of subsequent extension, cooling and convergence are applicable to orogenies that resulted from the opening and closure of embryonic oceans, which might have been the case for the Alpine orogeny.

Data availability. The data presented in this study are available on request from Lorenzo G. Candioti.

Appendix A: Numerical algorithm

~~The applied algorithm has already been used to model processes at different scales, such as deformation of eclogites on the centimetre-scale (Yamato et al., 2019), crystal-melt segregation of magma during its ascent in a meter-scale conduit (Yamato et al., 2015), rifting of continental lithosphere (Duretz et al., 2016b; Petri et al., 2019), stress calculations around the Tibetan Plateau (Schmalholz et al., 2~~
~~within and around a subduction of an oceanic plate (Bessat et al., 2020), as well as modelling Precambrian orogenic processes (Poh et al., 2020).~~

Table A1. Physical parameters used in the numerical simulations M1-6.

Parameter	Unit	Strong Crust[1]	Weak Crust[2]	Calcite[3]	Mica[4]	Dry Mantle[5]	Wet Mantle[6]
ρ_0	kg m^{-3}	2800	2800	2800	2800	-	-
G	Pa	2×10^{10}	2×10^{10}	2×10^{10}	2×10^{10}	2×10^{10}	2×10^{10}
c_P	J kg^{-1} K^{-1}	1050	1050	1050	1050	1050	1050
k	W m^{-1} K^{-1}	2.25	2.25	2.37	2.55	2.75	2.75
H_R	W m^{-3}	0.9×10^{-6}	0.9×10^{-6}	0.56×10^{-6}	2.9×10^{-6}	2.1139×10^{-8}	2.1139×10^{-8}
C	Pa	10^7	10^6	10^7	10^6	10^7	10^7
φ	°	30	5	30	5	30	30
α	K^{-1}	3×10^{-5}	3×10^{-5}	3×10^{-5}	3×10^{-5}	3×10^{-5}	3×10^{-5}
β	Pa^{-1}	1×10^{-11}	1×10^{-11}	1×10^{-11}	1×10^{-11}	1×10^{-11}	1×10^{-11}
Dislocation							
A	Pa^{-n-r} s^{-1}	5.0477×10^{-28}	5.0717×10^{-18}	1.5849×10^{-25}	10^{-138}	1.1×10^{-16}	5.6786×10^{-27}
n	-	4.7	2.3	4.7	18	3.5	3.5
Q	J mol^{-1}	485×10^3	154×10^3	297×10^3	51×10^3	530×10^3	460×10^3
V	m^3 mol^{-1}	0	0	0	0	14×10^{-6}	11×10^{-6}
r	-	0	0	0	0	0	1.2
f_{H_2O}	Pa	0	0	0	0	0	10^9
Diffusion							
A^*	Pa^{-n-r} m^m s^{-1}	-	-	-	-	1.5×10^{-15}	2.5×10^{-23}
n	-	-	-	-	-	1	1
Q	J mol^{-1}	-	-	-	-	370×10^3	375×10^3
V	m^3 mol^{-1}	-	-	-	-	7.5×10^{-6}	20×10^{-6}
m	-	-	-	-	-	3	3
r	-	-	-	-	-	0	1
f_{H_2O}	Pa	-	-	-	-	0	10^9
d	m	-	-	-	-	10^{-3}	10^{-3}
Peierls							
A_P	s^{-1}	-	-	-	-	5.7×10^{11}	5.7×10^{11}
Q	J mol^{-1}	-	-	-	-	540×10^3	540×10^3
σ_P	Pa	-	-	-	-	8.5×10^9	8.5×10^9
γ	-	-	-	-	-	0.1	0.1

Flow law parameters: [1]Maryland Diabase (Mackwell et al., 1998), [2]Wet Quartzite (Ranalli, 1995), [3]Calcite (Schmid et al., 1977), [4]Mica (Kronenberg et al., 1990), [5]Dry Olivine (Hirth and Kohlstedt, 2003) and [6]Wet Olivine (Hirth and Kohlstedt, 2003). Peierls creep: (Goetze and Evans, 1979) regularised by Kameyama et al. (1999). *Converted to SI units from original units: $A = 2.5 \times 10^7$ $([\mathrm{MPa}])^{-n-r}$ $([\mu\mathrm{m}])^m$ $([\mathrm{s}])^{-1} = 2.5 \times 10^7 \times (10^{-6n-6r}\,[\mathrm{Pa}]^{-6n-6r}) \times (10^{-6m}\,[\mathrm{m}]^m) \times ([s]^{-1}) = 2.5 \times 10^{-23}$ $[\mathrm{Pa}^{-2}\mathrm{m}^3\mathrm{s}^{-1}]$.

Continuity and force balance equations for an incompressible slowly flowing (no inertial forces) fluid under gravity are given by

$$\frac{\partial v_i}{\partial x_i} = 0 \tag{A1}$$

$$\frac{\partial \sigma_{ij}}{\partial x_j} = -\rho\, a_i \,, \tag{A2}$$

where v_i denotes velocity vector components and x_i spatial coordinate components, where (i,j=1) indicates the horizontal direction and (i,j=2) the vertical direction, σ_{ij} are components of the total stress tensor, ρ is density and ~~$a_i = [0; g]$~~ $a_i = [0; g]$ is a vector with g being the gravitational acceleration. Density is a function of pressure P (negative mean stress) and temperature T computed as a simplified equation of state for the crustal phases like

$$\rho(P,T) = \rho_0\,(1 - \alpha\Delta T)\,(1 + \beta\Delta P) \,, \tag{A3}$$

where ρ_0 is the material density at the reference temperature T_0 and pressure P_0, α is the thermal expansion coefficient, β is the compressibility coefficient, $\Delta T = T - T_0$ and $\Delta P = P - P_0$. Effective density for the mantle phases is pre-computed using the software package Perple_X (Connolly, 2005) for the bulk rock composition of a Hawaiian pyrolite (Workman and Hart, 2005). Figure A1 shows the density distribution for the calculated pressure and temperature range.

The viscoelastic stress tensor components are defined using a backward-Euler scheme (e.g., Schmalholz et al., 2001) as

$$\sigma_{ij} = -P\delta_{ij} + 2\left(\frac{1}{\eta^{\text{eff}}} + \frac{1}{G\Delta t}\right)^{-1} \dot{\varepsilon}_{ij} + \left(1 + \frac{G\Delta t}{\eta^{\text{eff}}}\right)^{-1} \tau^o_{ij} + J_{ij} \,, \tag{A4}$$

where $\delta_{ij} = 0$ if $i \neq j$, or $\delta_{ij} = 1$ if $i = j$, η^{eff} is the effective viscosity, G is shear modulus, Δt is the time step, $\dot{\varepsilon}_{ij}$ are the components of the deviatoric strain rate tensor, τ^o_{ij} is the deviatoric stress tensor of the preceding time step and J_{ij} comprises components of the Jaumann stress rate as described in detail in Beuchert and Podladchikov (2010). A visco-elasto-plastic Maxwell model is used to describe the rheology, implying that the components of the total deviatoric strain rate tensor $\dot{\varepsilon}_{\text{ij}}$ are additively decomposed into contributions from the viscous (dislocation, diffusion and Peierls creep), plastic and elastic deformation as

$$\dot{\varepsilon}_{ij} = \dot{\varepsilon}^{\text{ela}}_{ij} + \dot{\varepsilon}^{\text{pla}}_{ij} + \dot{\varepsilon}^{\text{dis}}_{ij} + \dot{\varepsilon}^{\text{dif}}_{ij} + \dot{\varepsilon}^{\text{pei}}_{ij} \,, \tag{A5}$$

A local iteration cycle is performed on each cell/node until eq. A5 is satisfied (e.g., Popov and Sobolev, 2008). The effective viscosity for the dislocation and Peierls creep flow law is a function of the second invariant of the respective strain rate components $\dot{\varepsilon}^{\text{dis,pei}}_{\text{II}} = \tau_{\text{II}}/(2\eta^{\text{dis,pei}})$

$$\eta^{\text{dis}} = \frac{2^{\frac{1-n}{n}}}{3^{\frac{1+n}{2n}}}\, A^{-1}\, \left(\dot{\varepsilon}^{\text{dis}}_{\text{II}}\right)^{\frac{1}{n}-1} \exp\left(\frac{Q + PV}{nRT}\right) \left(f_{\text{H}_2\text{O}}\right)^{-\frac{r}{n}} \,, \tag{A6}$$

where the ratio in front of the pre-factor A results from conversion of the experimentally derived 1D flow law, obtained from laboratory experiments, to a flow law for tensor components (e.g., Schmalholz and Fletcher, 2011). For the mantle material diffusion creep is taken into account and its viscosity takes the following form

$$\eta^{\mathrm{dif}} = \frac{1}{3} A^{-1} d^{m} \exp\left(\frac{Q+PV}{RT}\right) \left(f_{\mathrm{H_2O}}\right)^{-r}, \tag{A7}$$

where d is grain size and m is a grain size exponent. Effective Peierls viscosity is calculated using the regularised form of Kameyama et al. (1999) for the experimentally derived flow law by Goetze and Evans (1979) as

$$\eta^{\mathrm{pei}} = \frac{2^{\frac{1-s}{s}}}{3^{\frac{1+s}{2s}}} \hat{A} \left(\dot{\varepsilon}_{\mathrm{II}}^{\mathrm{pei}}\right)^{\frac{1}{s}-1}, \tag{A8}$$

where s is an effective, temperature dependent stress exponent:

$$s = 2\,\gamma\,\frac{Q}{RT}(1-\gamma)\,. \tag{A9}$$

$\hat{A}$ in Eq. (A8) is

$$\hat{A} = \left[A_{\mathrm{P}} \exp\left(-\frac{Q(1-\gamma)^2}{RT}\right)\right]^{-\frac{1}{s}} \gamma\sigma_{\mathrm{P}}\,, \tag{A10}$$

where A_{P} is a pre-factor, γ is a fitting parameter and σ_{P} is a characteristic stress value. In the frictional domain, stresses are limited by ~~a yield stress τ_{y} obtained by a~~ the Drucker-Prager yield ~~criterion~~ function

$$\tau_{\mathrm{y}} F = \tau_{\mathrm{II}} - P\,\sin\phi + C\,\cos\phi\,, \tag{A11}$$

and the non-associated potential function:

$$Q = \tau_{\mathrm{II}}, \tag{A12}$$

where ϕ is the internal angle of friction and C is the cohesion. If the yield condition is met ($f = \tau_{\mathrm{II}} - \tau_{\mathrm{y}} > 0$) the equivalent plastic viscosity is computed as

$$\eta^{\mathrm{pla}} = \frac{\tau_{\mathrm{y}}}{2\dot{\varepsilon}_{\mathrm{II}}^{\mathrm{pla}}} \frac{P\,\sin\phi + C\,\cos\phi}{2\dot{\varepsilon}_{\mathrm{II}}^{\mathrm{pla}}}, \tag{A13}$$

At the end of the iteration cycle, the effective viscosity in eq. A4 is computed as the inverse of the quasi-harmonic average of the viscous and the plastic contributions for the corresponding invariants of the strain rate tensor components of the distinct deformation mechanisms as

$$\eta^{\mathrm{eff}} = \left(\frac{1}{\eta^{\mathrm{dis}}} + \frac{1}{\eta^{\mathrm{dif}}} + \frac{1}{\eta^{\mathrm{pei}}} + \frac{1}{\eta^{\mathrm{pla}}} \right)^{-1} . \tag{A14}$$

Thermal evolution of the model is calculated with the heat transfer equation

$$\rho \, c_{\mathrm{P}} \, \frac{\mathrm{D}T}{\mathrm{D}t} = \frac{\partial}{\partial x_i} \left(k_{\mathrm{th}} \, \frac{\partial T}{\partial x_i} \right) + H_{\mathrm{A}} + H_{\mathrm{D}} + H_{\mathrm{R}} \, , \tag{A15}$$

where c_P is the specific heat capacity at constant pressure, $\mathrm{D}/\mathrm{D}t$ is the material time derivative, ~~k_{th}~~ k is thermal conductivity, $H_{\mathrm{A}} = T\alpha v_z g \rho$ is a heat source or sink resulting from adiabatic processes assuming lithostatic pressure conditions, $H_{\mathrm{D}} = \tau_{ij}^2 / \left(2\eta^{\mathrm{eff}}\right)$ results from the conversion of dissipative work into heat (so-called shear heating) and H_{R} is a radiogenic heat source.

To initiate the deformation, we perturbed the initial marker field with a random amplitude vertical displacement like

$$z_{\mathrm{M}} = z_{\mathrm{M}} + A \, \exp\left(- \frac{x_{\mathrm{M}}}{\lambda} \right) , \tag{A16}$$

where z_{M} is the vertical marker coordinate in km, A is a random amplitude varying between -1.25 km and 1.25 km, x_{M} is the horizontal marker coordinate in km and $\lambda = 25 \times 10^3$ km is the half-width of the curve. The perturbation is applied to the horizontal centre of the domain between -75 km and 75 km.

All physical parameters are summarised in table A1.

Appendix B: Nusselt number scaling laws and phase transitions

Modelling thermal convection beneath an actively deforming lithosphere can be numerically expensive, because the convection velocities can be as high as or even higher than the motion of the lithospheric plates, depending on the vigour of the convecting system. This significantly reduces the maximum time step necessary to ensure numerical stability. In consequence, it takes more time steps to run a simulation to the same physical time when convection is modelled together with deformation in the lithosphere. Hence, the computational time can be twice as long compared to models, where only the deforming lithosphere is modelled. However, it is possible to include the effect of ~~thermal~~ convection in the mantle on the thermal field and keep a constant vertical heat flux through the lithosphere-asthenosphere boundary (LAB) into a numerical model without explicitly modelling convection by using an effective thermal conductivity for the mantle material below the lithosphere. Two dimension-

less quantities have to be defined, namely the Rayleigh and Nusselt numbers. The Rayleigh number is the ratio of the thermal diffusion and advection time scale

$$Ra = \frac{t_{\mathrm{Dif}}}{t_{\mathrm{Adv}}} = \frac{\rho g \alpha \Delta T D^3}{\kappa \eta_{\mathrm{eff}}}, \tag{B1}$$

where ρ is density, g is gravitational acceleration, α is a coefficient of thermal expansion, ΔT is the temperature difference between the top and the bottom and D is the thickness of the convecting layer, ~~$\kappa = k_{\mathrm{th}}/\rho/c_P$~~ $\kappa = k/\rho/c_p$ is the thermal diffusivity and η_{eff} is the effective viscosity. The Nusselt number can be expressed in terms of the Rayleigh number as

$$Nu = \left(\frac{Ra}{Ra_{\mathrm{crit}}}\right)^{\beta}, \tag{B2}$$

where Ra_{crit} is the critical Rayleigh number ~~, typical~~ at which convection starts, typically in the order of 10^3, ~~at which convection starts~~ and β is a power-law exponent (Schubert et al., 2001). The Nusselt number is the ratio of advective heat flux, $q_{\mathrm{~~adv~~Adv}}$, which is the vertical heat flux through the base of the lithosphere, ~~q_{LAB},~~ imposed by the convecting upper mantle to the diffusive heat flux, $q_{\mathrm{~~dif~~Dif}}$, imposed by the lithosphere on top of the convecting upper mantle as

$$Nu = \frac{q_{\mathrm{Adv}}}{q_{\mathrm{Dif}}} = \frac{q_{\mathrm{LAB}}}{\frac{k_{\mathrm{th}}(T_{\mathrm{M}} - T_{\mathrm{LAB}})}{D}}. \tag{B3}$$

~~where k_{th} is thermal conductivity, T_{M} is the average temperature of the upper mantle and T_{LAB} is the temperature at the base of the lithosphere.~~ Using this relationship, it is possible to scale the thermal conductivity to the Nusselt number of the Earth's mantle and to maintain a constant heat flow through the base of the lithosphere via conduction when convection is absent. Assuming ~~a vertical heat flow of 30 $^{-2}$ through the base of the lithosphere,~~ $Ra = 2 \times 10^6$ and $\beta = 1/3$ for the Earth's upper mantle convection, eq. B2 predicts $Nu = 13$. This implies that the heat flow provided by advection is $13\times$ higher than the heat flow provided by conduction. ~~Equation B3 allows for three possibilities to balance the excess advective heat flow through the lithosphere: (1) D can be reduced by a factor 13, (2) (T_{M}-T_{LAB}) or (3) k_{th} can be increased by a factor 13. Since it is desired to mimic the convection~~ Using an effective conductivity approach, the heat flow provided by advection is mimicked using an enhanced conductive heat flow in the upper mantle~~down to a depth of 660 , neither D nor (T_{M}-T_{LAB}) can be changed. Hence, k_{th} of~~ . The effective conductivity can be determined by scaling the standard value of thermal conductivity of the upper mantle material to the ~~asthenosphere needs to be $13\times$ higher than the realistic value to balance the heat flow through the lithosphere by conduction only.~~ Nusselt number of the convecting system like

$$k_{\mathrm{eff}} = Nu\, k. \tag{B4}$$

For this study, the ~~realistic~~ standard value for $k_{\mathrm{~~th~~}} = 2.75$ ~~and $k_{\mathrm{eff}} = 36$. Calculating the effective conductive heat flow through the base of the lithosphere q_{eff}~~

$$q_{\text{eff}} = k_{\text{eff}} \frac{T_{\text{LAB}} - T_{\text{bot}}}{D}$$

using $T_{\text{LAB}} = 1350$ and $T_{\text{bot}} = 1612$ as temperatures at the LAB and at a depth of 660 and a total depth of the upper mantle and transition zone of 540 yields a vertical heat flow = 2.75 of the upper mantle material and $qNu_{\text{eff}} = 32^2$, which is consistent with the mean heat flow at the base of continental lithosphere (Turcotte and Schubert, 2014)= 13. The effective conductivity according to eq. B4 is $k_{\text{eff}} = 36$. To avoid a strong contrast of conductivities directly at the base of the lithosphere, we linearly increase the conductivity from 2.75-36 W m^{-1} K^{-1} over a temperature range of 1350-1376 °C. Applying this effective conductivity approach reduces the number of time steps necessary for computation up to the same physical time by ca. a factor 2 in M4 compared to M1.

As mentioned above, the vigour of convection is defined by the Rayleigh number (eq. B1). For $Ra >> Ra_{\text{crit}}$, the time scale for thermal diffusion is much larger than the time scale for advection of material. This means that the entropy of the system remains relatively constant in time. By definition, such a system is adiabatic (Kondepudi and Prigogine, 2014). In the presented simulationsmodels, the density and entropy for the mantle phases is pre-computed using Perple_X for a given bulk rock composition. Assuming that the temperature gradient in the upper mantle is adiabatic and stress conditions are close to lithostatic (i.e., deviatoric stresses are negligible), the temperature at any depth can be determined by following an isentrop from the Perple_X database. Starting coordinates in pressure-temperature space are the (lithostatic) conditions at the base of the lithosphere. From these pressure and temperature values one can follow the closest isentrop (black line Fig. A1) until the (lithostatic) pressure value at target depth (in this study 660 km, red diamond Fig. A1) is reached and extract the corresponding temperature value. Trubitsyn and Trubitsyna (2015) derived an analytical solution to calculate temperatures assuming an adiabatic gradient for given depths. We determined the temperature at the bottom of the model domain using both approaches and the obtained values that differ by only 0.01 °C. We use an extended Boussinesq approximation, which means that terms describing adiabatic processes only appear in eq. A15 (Petersen et al., 2010), density variations due to temperature gradients arenegligible on the long term, but temperature changes due to adiabatic processes can be as high as the contribution of radiogenic heat production in the mantle. Boundary conditions

Involving phase transitions necessitates mainly two major assumptions: (1) is compressibility of material due to large density variations important and (2) does latent heat released or consumed at a phase transition significantly change the convective pattern? Bercovici et al. (1992) concluded that compressibility effects on the spatial mantle structure are minor when the superadiabatic temperature drop is close to the adiabatic temperature of the mantle, which is the case for the Earth. Although the net density varies largely in *P-T* space of the phase diagram used in this study (fig. A1), the maximum value for the density time derivative computed from M1 is two orders of magnitude lower than the velocity divergence. We assume that density changes due to volumetric deformation are, hence, still negligible and density changes are accounted for in the buoyancy force only. This means the classical Boussinesq approximation is still valid. However, not considering adiabatic heating in the energy conservation equation leads to a significant deviation of the thermal structure from the initially imposed temperature gradient over large time scales (>100 Myrs). The resulting temperature profile is constant throughout the upper mantle and the newly

equilibrated constant temperature is equal to the imposed temperature at the bottom ~~of the domain are set to free slip, that is the vertical velocity is zero. This represents an impermeable mechanical boundary that mimics the phase transition of γ-spinel olivine to -Perovskite at pressure and temperature conditions at a depth of 660 and we here assume that convection in the mantle is two-layered~~boundary. In consequence, the density structure read in from the phase diagram table according to pressure and temperature values is significantly wrong. To avoid these problems, we use the extended Boussinesq approximation, i.e., the adiabatic heating term is included in the energy conservation equation. As a result, the initially imposed adiabatic (or isentropic) temperature gradient can be maintained over large time scales. The resulting density structure agrees well with the PREM model (Dziewonski and Anderson, 1981) as shown in this study. A detailed comparison between different approximations of the system of equations is clearly beyond the scope of this study. Latent heat that is released or consumed by a phase transition can perturb the thermal field by up to 100 K and induce a buoyancy force aiding or inhibiting the motion of cold, especially low-angle, subducting slabs (van Hunen et al., 2001) or hot rising plumes. However, when the lateral differences in temperature do not vary much, the deflection of the phase transition by an ascending plume or a subducting slab has a much bigger impact on the buoyancy stresses than the latent heat released or consumed by the phase transition (Christensen, 1995). Because a detailed parametric investigation of the impact of latent heat on buoyancy stresses is beyond the scope of study, we neglect latent heat for simplicity.

Appendix C: Convection benchmark

In this section, we present the results of a convection benchmark performed by the algorithm used in this study. Equations for continuity and force balance are solved as in eq. A1 and A2, density is a function of temperature only and calculated as

$$\rho(T) = \rho_0(1 - \alpha T) \tag{C1}$$

and the total stress tensor is decomposed into a pressure and a deviatoric part as

$$\sigma_{ij} = -P\delta_{ij} + \tau_{ij} \ . \tag{C2}$$

Transfer of heat is calculated as in eq. A15. Stresses and strain rates ($\dot{\varepsilon}_{ij}$) are related to each other via the viscosity η as

$$\tau_{ij} = 2\eta\dot{\varepsilon}_{ij} \ . \tag{C3}$$

Viscosity is computed via a linearized Arrhenius law, also called Frank-Kamenetskii approximation (Kamenetskii, 1969):

$$\eta(T,z) = \exp(-\gamma_T + \gamma_z) \ , \tag{C4}$$

with $\gamma_T = \log(\eta_T)$ and $\gamma_z = \log(\eta_z)$. By choosing $\eta_z = 1$, $\gamma_z = 0$ in eq. C4 and, therefore, the viscosity is only temperature dependent.

The dimensionless equations are discretized over a domain that extends from 0 to 1 in both horizontal and vertical directions and a small amplitude perturbation is applied to the initial temperature profile as

$$T(x,z) = (1-z) + A\cos(\pi x)\sin(\pi z)\ . \tag{C5}$$

As mentioned above, the vigour of the convecting system is described by its Rayleigh number (eq. B1). A local $Ra = 10^2$ is applied to the top boundary by setting $\alpha = 10^{-2}$, $g = 10^4$ and all other parameters of eq.B1 are set to 1. The applied viscosity decrease by choosing $\eta_T = 10^5$ in eq. C4 results in a global $Ra = 10^7$. All mechanical boundaries are set to free slip, the thermal boundary conditions are constant temperature at the top ($T = 0$) and bottom ($T = 1$) and insulating (i.e., zero flux) at the two vertical boundaries. Tosi et al. (2015) tested several algorithms, including finite element, finite differences, finite volume and spectral discretization, on their capability of modelling distinct rheologies of the mantle, from temperature dependent viscosity only up to visco-plastic rheologies. We have chosen the simplest test, case one in Tosi et al. (2015), and report the results of two distinct diagnostic quantities: the average temperature over the entire modelling domain

$$\langle T \rangle = \int_0^1 \int_0^1 T \,\mathrm{dxdz} \tag{C6}$$

and the root mean square velocity at the surface

$$u_{\mathrm{RMS}}^{\mathrm{surf}} = \left(\int_0^1 v_x^2 \Big|_{z=1} \,\mathrm{dx} \right)^{\frac{1}{2}}\ . \tag{C7}$$

The model develops one convection cell below a stagnant lid (fig.A2(a), (b)). Diagnostic quantities of average temperature (fig.A2(c)) and root-mean square velocity (fig.A2(d)) at the top of the domain agree with values that have been produced by the algorithms (grey areas in fig.A2(c) and (d)) tested in Tosi et al. (2015). We tested numerical resolutions of 50^2, 100^2, 150^2 and 300^2. Only for resolutions >100^2 the desired convective pattern developed. The numerical algorithms tested by Tosi et al. (2015) passed the benchmark already for lower resolutions. This is due to the fact that the algorithm presented here uses a uniform grid size across the domain. The algorithms tested by Tosi et al. (2015) used refined meshes. Sufficient resolution of the thermal boundary layers at the top and at the bottom is crucial to develop the desired pattern. Using a refined mesh in these regions, allows for lower total resolution, whereas using a regular mesh necessitates a much higher resolution in total. Nevertheless, values for the diagnostic quantities reproduced by the presented algorithm lie well within the minimum and maximum values calculated by the algorithms tested in Tosi et al. (2015) (grey areas in fig. A2(b) & (c)). This shows that the convection in the upper mantle, where the viscosity is essentially temperature dependent, in the ~~simulations~~ models presented in this study in which convection is not parameterized, is accurately modelled.

Appendix D: Gravitational potential energy and plate driving forces

We use the gravitational potential energy per unit surface (*GPE*) to quantify the impact of convection induced density variations in the upper mantle during the distinct stages of the simulations. The *GPE* varies along the horizontal x-direction and is computed as

$$GPE(x) = \int_{Sb}^{St(x)} P_{\mathrm{L}}(x,z)\ \mathrm{dz}, \tag{D1}$$

where P_{L} is the lithostatic pressure calculated as

$$P_{\mathrm{L}}(x,z) = \int_{z}^{St(x)} \rho(x,z')\ g\ \mathrm{dz'} \tag{D2}$$

and $St(x)$ is the stress-free surface and Sb is the model bottom. Horizontal variations in *GPE*, ΔGPE, are calculated by subtracting the leftmost value as a reference value from all other values. The *GPE* gives an estimate on the plate driving forces per unit length (Molnar and Lyon-Caen, 1988; Schmalholz et al., 2019; Bessat et al., 2020) acting in the system during the different stages of the hyper-extension and convergence cycle. We also calculate the vertical integral of the second invariant of the deviatoric stress tensor

$$\bar{\tau}_{\mathrm{II}}(x) = \int_{Sb}^{St(x)} \tau_{\mathrm{II}}(x,z)\ \mathrm{dz}\ . \tag{D3}$$

~~which also can~~ The value of $F_{\mathrm{D}} = 2 \times \bar{\tau}_{\mathrm{II}}^{\mathrm{avg}}(x)$, where $\bar{\tau}_{\mathrm{II}}^{\mathrm{avg}}(x)$ is calculated averaging the average of $\bar{\tau}_{\mathrm{II}}(x)$ both over the left and rightmost 100 km of the domain, is also identical to the vertical integral of the difference between the horizontal total stress and the lithostatic pressure, if shear stresses are negligible (Molnar and Lyon-Caen, 1988; Schmalholz et al., 2019). During onset of convergence (100 Myrs), when the lithosphere is more or less homogeneously shortened, shear stresses are most likely negligible compared to normal deviatoric stresses. F_{D} can therefore be used to estimate the plate driving force (see fig.7(d)). For calculation of the suction force per unit length (F_{S}) induced by the mantle flow in fig. 6(a) & (f) we used the following formula

$$F_{\mathrm{S}} = \int_{b}^{a} \int_{d}^{c} \Delta\rho g\ \mathrm{dz}\ \mathrm{dx}\ , \tag{D4}$$

where $a = 350$ km, $b = 150$ km, $c = -120$ km and $d = -450$ km are the integration bounds and $\Delta\rho$ is the difference in density between the initial density field and the density field at the end of the ~~thermal relaxation~~ cooling period.

Author contributions. Lorenzo G. Candioti configured and performed the numerical simulations and the convection benchmark, interpreted the numerical results, generated the figures and wrote the manuscript. Stefan M. Schmalholz designed the numerical study, helped in interpreting the results and designing the figures and contributed to writing the manuscript. Thibault Duretz developed the applied numerical algorithm and helped in configuring the model and the interpretation of the results.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We are grateful for two helpful and constructive anonymous reviews. This work is supported by SNF grant No. 200020 163169. ~~Scientific colour maps used in all figures are taken from Crameri (2018).~~

References

Afonso, J., Ranalli, G., and Fernandez, M.: Density structure and buoyancy of the oceanic lithosphere revisited, Geophysical Research Letters, 34, 2007.

Agrusta, R., van Hunen, J., and Goes, S.: Strong plates enhance mantle mixing in early Earth, Nature communications, 9, 1–10, 2018.

Baes, M., Sobolev, S. V., and Quinteros, J.: Subduction initiation in mid-ocean induced by mantle suction flow, Geophysical Journal International, 215, 1515–1522, 2018.

Ballmer, M., Van Hunen, J., Ito, G., Tackley, P., and Bianco, T.: Non-hotspot volcano chains originating from small-scale sublithospheric convection, Geophysical Research Letters, 34, 2007.

Beaussier, S. J., Gerya, T. V., and Burg, J.-P.: 3D numerical modelling of the Wilson cycle: structural inheritance of alternating subduction polarity, Geological Society, London, Special Publications, 470, 439–461, 2019.

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

Behn, M. D., Conrad, C. P., and Silver, P. G.: Detection of upper mantle flow associated with the African Superplume, Earth and Planetary Science Letters, 224, 259–274, 2004.

Bercovici, D. and Ricard, Y.: Mechanisms for the generation of plate tectonics by two-phase grain-damage and pinning, Physics of the Earth and Planetary Interiors, 202, 27–55, 2012.

Bercovici, D., Schubert, G., and Glatzmaier, G. A.: Three-dimensional convection of an infinite-Prandtl-number compressible fluid in a basally heated spherical shell, Journal of Fluid Mechanics, 239, 683–719, 1992.

Bessat, A., Duretz, T., Hetényi, G., Pilet, S., and Schmalholz, S. M.: Stress and deformation mechanisms at a subduction zone: insights from 2D thermo-mechanical numerical modelling, Geophysical Journal International, 2020.

Beuchert, M. J. and Podladchikov, Y. Y.: Viscoelastic mantle convection and lithospheric stresses, geophysical Journal international, 183, 35–63, 2010.

Boonma, K., Kumar, A., García-Castellanos, D., Jiménez-Munt, I., and Fernández, M.: Lithospheric mantle buoyancy: the role of tectonic convergence and mantle composition, Scientific Reports, 9, 1–8, 2019.

Brune, S., Williams, S. E., Butterworth, N. P., and Müller, R. D.: Abrupt plate accelerations shape rifted continental margins, Nature, 536, 201–204, 2016.

Buck, W. R.: Small-scale convection induced by passive rifting: the cause for uplift of rift shoulders, Earth and Planetary Science Letters, 77, 362–372, 1986.

Chen, J.: Lower-mantle materials under pressure, Science, 351, 122–123, 2016.

Chenin, P., Manatschal, G., Picazo, S., Müntener, O., Karner, G., Johnson, C., and Ulrich, M.: Influence of the architecture of magma-poor hyperextended rifted margins on orogens produced by the closure of narrow versus wide oceans, Geosphere, 13, 559–576, 2017.

Chenin, P., Schmalholz, S. M., Manatschal, G., and Karner, G. D.: Necking of the lithosphere: A reappraisal of basic concepts with thermo-mechanical numerical modeling, Journal of Geophysical Research: Solid Earth, 123, 5279–5299, 2018.

Chenin, P., Picazo, S., Jammes, S., Manatschal, G., Müntener, O., and Karner, G.: Potential role of lithospheric mantle composition in the Wilson cycle: a North Atlantic perspective, Geological Society, London, Special Publications, 470, 157–172, 2019.

Christensen, U.: Effects of phase transitions on mantle convection, Annual Review of Earth and Planetary Sciences, 23, 65–87, 1995.

Cloos, M.: Lithospheric buoyancy and collisional orogenesis: Subduction of oceanic plateaus, continental margins, island arcs, spreading ridges, and seamounts, Geological Society of America Bulletin, 105, 715–737, 1993.

Connolly, J. A.: Computation of phase equilibria by linear programming: a tool for geodynamic modeling and its application to subduction zone decarbonation, Earth and Planetary Science Letters, 236, 524–541, 2005.

Conrad, C. P. and Lithgow-Bertelloni, C.: How mantle slabs drive plate tectonics, Science, 298, 207–209, 2002.

Crameri, F.: Geodynamic diagnostics, scientific visualisation and StagLab 3.0, Geoscientific Model Development, 11, 2541–2562, 2018.

Crameri, F., Conrad, C. P., Montési, L., and Lithgow-Bertelloni, C. R.: The dynamic life of an oceanic plate, Tectonophysics, 760, 107–135, 2019.

Crameri, F., Magni, V., Domeier, M., Shephard, G. E., Chotalia, K., Cooper, G., Eakin, C. M., Grima, A. G., Gürer, D., Király, Á., et al.: A transdisciplinary and community-driven database to unravel subduction zone initiation, Nature Communications, 11, 1–14, 2020.

Currie, C. A., Huismans, R. S., and Beaumont, C.: Thinning of continental backarc lithosphere by flow-induced gravitational instability, Earth and Planetary Science Letters, 269, 436–447, 2008.

Dal Zilio, L., Faccenda, M., and Capitanio, F.: The role of deep subduction in supercontinent breakup, Tectonophysics, 746, 312–324, 2018.

De Graciansky, P.-C., Roberts, D. G., and Tricart, P.: The Western Alps, from rift to passive margin to orogenic belt: an integrated geoscience overview, Elsevier, 2010.

Duretz, T., May, D. A., and Yamato, P.: A free surface capturing discretization for the staggered grid finite difference scheme, Geophysical Journal International, 204, 1518–1530, 2016a.

Duretz, T., Petri, B., Mohn, G., Schmalholz, S., Schenker, F., and Müntener, O.: The importance of structural softening for the evolution and architecture of passive margins, Scientific reports, 6, 38 704, 2016b.

Dziewonski, A. M. and Anderson, D. L.: Preliminary reference Earth model, Physics of the earth and planetary interiors, 25, 297–356, 1981.

England, P., Molnar, P., and Richter, F.: John Perry's neglected critique of Kelvin's age for the Earth: A missed opportunity in geodynamics, GSA Today, 17, 2007.

Erdös, Z., Huismans, R., and van der Beek, P.: Control of increased sedimentation on orogenic fold-and-thrust belt structure-insights into the evolution of the Western Alps, 2019.

Faccenna, C. and Becker, T. W.: Shaping mobile belts by small-scale convection, Nature, 465, 602–605, 2010.

Favre, P. and Stampfli, G.: From rifting to passive margin: the examples of the Red Sea, Central Atlantic and Alpine Tethys, Tectonophysics, 215, 69–97, 1992.

Forte, A. M., Quéré, S., Moucha, R., Simmons, N. A., Grand, S. P., Mitrovica, J. X., and Rowley, D. B.: Joint seismic–geodynamic-mineral physical modelling of African geodynamics: A reconciliation of deep-mantle convection with surface geophysical constraints, Earth and Planetary Science Letters, 295, 329–341, 2010.

Froitzheim, N. and Manatschal, G.: Kinematics of Jurassic rifting, mantle exhumation, and passive-margin formation in the Austroalpine and Penninic nappes (eastern Switzerland), Geological society of America bulletin, 108, 1120–1133, 1996.

Garel, F., Goes, S., Davies, D., 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, 2014.

Gerya, T. V. and Yuen, D. A.: Characteristics-based marker-in-cell method with conservative finite-differences schemes for modeling geological flows with strongly variable transport properties, Physics of the Earth and Planetary Interiors, 140, 293–318, 2003.

Goetze, C. and Evans, B.: Stress and temperature in the bending lithosphere as constrained by experimental rock mechanics, Geophysical Journal International, 59, 463–478, 1979.

Golabek, G. J., Keller, T., Gerya, T. V., Zhu, G., Tackley, P. J., and Connolly, J. A.: Origin of the Martian dichotomy and Tharsis from a giant impact causing massive magmatism, Icarus, 215, 346–357, 2011.

Gülcher, A. J., Beaussier, S. J., and Gerya, T. V.: On the formation of oceanic detachment faults and their influence on intra-oceanic subduction initiation: 3D thermomechanical modeling, Earth and Planetary Science Letters, 506, 195–208, 2019.

Handy, M. R., Schmid, S. M., Bousquet, R., Kissling, E., and Bernoulli, D.: Reconciling plate-tectonic reconstructions of Alpine Tethys with the geological–geophysical record of spreading and subduction in the Alps, Earth-Science Reviews, 102, 121–158, 2010.

Hasterok, D. and Chapman, D.: Heat production and geotherms for the continental lithosphere, Earth and Planetary Science Letters, 307, 59–70, 2011.

Herzberg, C., Condie, K., and Korenaga, J.: Thermal history of the Earth and its petrological expression, Earth and Planetary Science Letters, 292, 79–88, 2010.

Hirth, G. and Kohlstedt, D.: Rheology of the upper mantle and the mantle wedge: A view from the experimentalists, GEOPHYSICAL MONOGRAPH-AMERICAN GEOPHYSICAL UNION, 138, 83–106, 2003.

Holt, A., Royden, L., and Becker, T.: The dynamics of double slab subduction, Geophysical Journal International, 209, 250–265, 2017.

Hua, Y., Zhao, D., and Xu, Y.: P wave anisotropic tomography of the Alps, Journal of Geophysical Research: Solid Earth, 122, 4509–4528, 2017.

Huang, J., Zhong, S., and van Hunen, J.: Controls on sublithospheric small-scale convection, Journal of Geophysical Research: Solid Earth, 108, 2003.

Huismans, R. and Beaumont, C.: Depth-dependent extension, two-stage breakup and cratonic underplating at rifted margins, Nature, 473, 74–78, 2011.

Jammes, S. and Huismans, R. S.: Structural styles of mountain building: Controls of lithospheric rheologic stratification and extensional inheritance, Journal of Geophysical Research: Solid Earth, 117, 2012.

Jaquet, Y. and Schmalholz, S. M.: Spontaneous ductile crustal shear zone formation by thermal softening and related stress, temperature and strain rate evolution, Tectonophysics, 746, 384–397, 2018.

Kamenetskii, D. F.: Diffusion and heat transfer in chemical kinetics, Plenum Press, New York, 1969.

Kameyama, M., Yuen, D. A., and Karato, S.-I.: Thermal-mechanical effects of low-temperature plasticity (the Peierls mechanism) on the deformation of a viscoelastic shear zone, Earth and Planetary Science Letters, 168, 159–172, 1999.

Kaus, B. J., Liu, Y., Becker, T., Yuen, D. A., and Shi, Y.: Lithospheric stress-states predicted from long-term tectonic models: Influence of rheology and possible application to Taiwan, Journal of Asian Earth Sciences, 36, 119–134, 2009.

Kiss, D., Candioti, L. G., Duretz, T., and Schmalholz, S. M.: Thermal softening induced subduction initiation at a passive margin, Geophysical Journal International, 220, 2068–2073, 2020.

Kondepudi, D. and Prigogine, I.: Modern thermodynamics: from heat engines to dissipative structures, John Wiley & Sons, 2014.

Kronenberg, A. K., Kirby, S. H., and Pinkston, J.: Basal slip and mechanical anisotropy of biotite, Journal of Geophysical Research: Solid Earth, 95, 19 257–19 278, 1990.

Li, C., van der Hilst, R. D., Engdahl, E. R., and Burdick, S.: A new global model for P wave speed variations in Earth's mantle, Geochemistry, Geophysics, Geosystems, 9, 2008.

Li, D., Wang, G., Bons, P., Zhao, Z., Du, J., Wang, S., Yuan, G., Liang, X., Zhang, L., Li, C., et al.: Subduction reversal in a divergent double subduction zone drives the exhumation of southern Qiangtang blueschist-bearing mélange, central Tibet, Tectonics, 2020.

Li, Z.-H., Gerya, T., and Connolly, J. A.: Variability of subducting slab morphologies in the mantle transition zone: Insight from petrological-thermomechanical modeling, Earth-Science Reviews, 2019.

Mackwell, S., Zimmerman, M., and Kohlstedt, D.: High-temperature deformation of dry diabase with application to tectonics on Venus, Journal of Geophysical Research: Solid Earth, 103, 975–984, 1998.

Mallard, C., Coltice, N., Seton, M., Müller, R. D., and Tackley, P. J.: Subduction controls the distribution and fragmentation of Earth's tectonic plates, Nature, 535, 140–143, 2016.

Manatschal, G. and Müntener, O.: A type sequence across an ancient magma-poor ocean–continent transition: the example of the western Alpine Tethys ophiolites, Tectonophysics, 473, 4–19, 2009.

Manzotti, P., Ballevre, M., Zucali, M., Robyr, M., and Engi, M.: The tectonometamorphic evolution of the Sesia–Dent Blanche nappes (internal Western Alps): review and synthesis, Swiss Journal of Geosciences, 107, 309–336, 2014.

McCarthy, A., Chelle-Michou, C., Müntener, O., Arculus, R., and Blundy, J.: Subduction initiation without magmatism: The case of the missing Alpine magmatic arc, Geology, 46, 1059–1062, 2018.

McCarthy, A., Tugend, J., Mohn, G., Candioti, L., Chelle-Michou, C., Arculus, R., Schmalholz, S. M., and Müntener, O.: A case of Ampferer-type subduction and consequences for the Alps and the Pyrenees, American Journal of Science, 320, 313–372, 2020.

McKenzie, D.: The initiation of trenches: a finite amplitude instability, Island arcs, deep sea trenches and back-arc basins, 1, 57–61, 1977.

Mitrovica, J. and Forte, A.: A new inference of mantle viscosity based upon joint inversion of convection and glacial isostatic adjustment data, Earth and Planetary Science Letters, 225, 177–189, 2004.

Molnar, P. and Lyon-Caen, H.: Some simple physical aspects of the support, structure, and evolution of mountain belts, Processes in continental lithospheric deformation, 218, 179–207, 1988.

Montési, L. G.: Fabric development as the key for forming ductile shear zones and enabling plate tectonics, Journal of Structural Geology, 50, 254–266, 2013.

Mueller, S. and Phillips, R. J.: On the initiation of subduction, Journal of Geophysical Research: Solid Earth, 96, 651–665, 1991.

Parsons, B. and McKenzie, D.: Mantle convection and the thermal structure of the plates, Journal of Geophysical Research: Solid Earth, 83, 4485–4496, 1978.

Peron-Pinvidic, G., Manatschal, G., et al.: Rifted margins: state of the art and future challenges, Frontiers in Earth Science, 7, 218, 2019.

Petersen, K. D., Nielsen, S., Clausen, O., Stephenson, R., and Gerya, T.: Small-scale mantle convection produces stratigraphic sequences in sedimentary basins, Science, 329, 827–830, 2010.

Petri, B., Duretz, T., Mohn, G., Schmalholz, S. M., Karner, G. D., and Müntener, O.: Thinning mechanisms of heterogeneous continental lithosphere, Earth and Planetary Science Letters, 512, 147–162, 2019.

Picazo, S., Müntener, O., Manatschal, G., Bauville, A., Karner, G., and Johnson, C.: Mapping the nature of mantle domains in Western and Central Europe based on clinopyroxene and spinel chemistry: Evidence for mantle modification during an extensional cycle, Lithos, 266, 233–263, 2016.

Piromallo, C. and Morelli, A.: P wave tomography of the mantle under the Alpine-Mediterranean area, Journal of Geophysical Research: Solid Earth, 108, 2003.

Poh, J., Yamato, P., Duretz, T., Gapais, D., and Ledru, P.: Precambrian deformation belts in compressive tectonic regimes: A numerical perspective, Tectonophysics, 777, 228 350, 2020.

Popov, A. and Sobolev, S.: SLIM3D: A tool for three-dimensional thermomechanical modeling of lithospheric deformation with elasto-visco-plastic rheology, Physics of the Earth and Planetary Interiors, 171, 55–75, 2008.

Prout, W.: Chemistry, meteorology and the function of digestion [Bridgewater Treatise No. 8], London: W. Pickering, 1834.

Ranalli, G.: Rheology of the Earth, Springer Science & Business Media, 1995.

Reilinger, R., McClusky, S., Vernant, P., Lawrence, S., Ergintav, S., Cakmak, R., Ozener, H., Kadirov, F., Guliev, I., Stepanyan, R., et al.: GPS constraints on continental deformation in the Africa-Arabia-Eurasia continental collision zone and implications for the dynamics of plate interactions, Journal of Geophysical Research: Solid Earth, 111, 2006.

Ricard, Y., Vigny, C., and Froidevaux, C.: Mantle heterogeneities, geoid, and plate motion: A Monte Carlo inversion, Journal of Geophysical Research: Solid Earth, 94, 13 739–13 754, 1989.

Richter, F. M.: Convection and the large-scale circulation of the mantle, Journal of Geophysical Research, 78, 8735–8745, 1973.

Ruh, J. B., Le Pourhiet, L., Agard, P., Burov, E., and Gerya, T.: Tectonic slicing of subducting oceanic crust along plate interfaces: Numerical modeling, Geochemistry, Geophysics, Geosystems, 16, 3505–3531, 2015.

Schmalholz, S., Podladchikov, Y., and Schmid, D.: A spectral/finite difference method for simulating large deformations of heterogeneous, viscoelastic materials, Geophysical Journal International, 145, 199–208, 2001.

Schmalholz, S. M. and Fletcher, R. C.: The exponential flow law applied to necking and folding of a ductile layer, Geophysical Journal International, 184, 83–89, 2011.

Schmalholz, S. M. and Mancktelow, N. S.: Folding and necking across the scales: a review of theoretical and experimental results and their applications., Solid Earth, 7, 2016.

Schmalholz, S. M., Podladchikov, Y. Y., and Jamtveit, B.: Structural softening of the lithosphere, Terra Nova, 17, 66–72, 2005.

Schmalholz, S. M., Duretz, T., Hetényi, G., and Medvedev, S.: Distribution and magnitude of stress due to lateral variation of gravitational potential energy between Indian lowland and Tibetan plateau, Geophysical Journal International, 216, 1313–1333, 2019.

Schmid, S., Boland, J., and Paterson, M.: Superplastic flow in finegrained limestone, Tectonophysics, 43, 257–291, 1977.

Schubert, G., Turcotte, D. L., and Olson, P.: Mantle convection in the Earth and planets, Cambridge University Press, 2001.

Solomatov, V.: Initiation of subduction by small-scale convection, Journal of Geophysical Research: Solid Earth, 109, 2004.

Stern, R. J.: Subduction initiation: spontaneous and induced, Earth and Planetary Science Letters, 226, 275–292, 2004.

Stern, R. J. and Gerya, T.: Subduction initiation in nature and models: A review, Tectonophysics, 746, 173–198, 2018.

Tackley, P. J., Schubert, G., Glatzmaier, G. A., Schenk, P., Ratcliff, J. T., and Matas, J.-P.: Three-dimensional simulations of mantle convection in Io, Icarus, 149, 79–93, 2001.

Thielmann, M. and Kaus, B. J.: Shear heating induced lithospheric-scale localization: Does it result in subduction?, Earth and Planetary Science Letters, 359, 1–13, 2012.

Thielmann, M. and Schmalholz, S. M.: Contributions of Grain Damage, Thermal Weakening, and Necking to Slab Detachment, Frontiers in Earth Science, 8, 254, 2020.

Torrance, K. and Turcotte, D.: Structure of convection cells in the mantle, Journal of Geophysical Research, 76, 1154–1161, 1971.

Tosi, N., Stein, C., Noack, L., Hüttig, C., Maierova, P., Samuel, H., Davies, D. R., Wilson, C. R., Kramer, S. C., Thieulot, C., et al.: A community benchmark for viscoplastic thermal convection in a 2-D square box, Geochemistry, Geophysics, Geosystems, 16, 2175–2196, 2015.

Trubitsyn, V. and Trubitsyna, A.: Effects of compressibility in the mantle convection equations, Izvestiya, Physics of the Solid Earth, 51, 801–813, 2015.

Turcotte, D. and Schubert, G.: Geodynamics, Cambridge University Press, 2014.

van Hunen, J., van den Berg, A. P., and Vlaar, N. J.: Latent heat effects of the major mantle phase transitions on low-angle subduction, Earth and Planetary Science Letters, 190, 125–135, 2001.

Van Wijk, J., Van Hunen, J., and Goes, S.: Small-scale convection during continental rifting: Evidence from the Rio Grande rift, Geology, 36, 575–578, 2008.

White, S. t. and Knipe, R.: Transformation-and reaction-enhanced ductility in rocks, Journal of the Geological Society, 135, 513–516, 1978.

Wilson, J. T.: Did the Atlantic close and then re-open?, Nature, 211, 676–681, 1966.

Wilson, R., Houseman, G., Buiter, S., McCaffrey, K., and Doré, A.: Fifty years of the Wilson Cycle concept in plate tectonics: an overview, 2019.

Workman, R. K. and Hart, S. R.: Major and trace element composition of the depleted MORB mantle (DMM), Earth and Planetary Science Letters, 231, 53–72, 2005.

Yamato, P., Duretz, T., May, D. A., and Tartese, R.: Quantifying magma segregation in dykes, Tectonophysics, 660, 132–147, 2015.

Yamato, P., Duretz, T., and Angiboust, S.: Brittle/ductile deformation of eclogites: insights from numerical models, Geochemistry, Geophysics, Geosystems, 2019.

Yang, F., Santosh, M., Tsunogae, T., Tang, L., and Teng, X.: Multiple magmatism in an evolving suprasubduction zone mantle wedge: the case of the composite mafic–ultramafic complex of Gaositai, North China Craton, Lithos, 284, 525–544, 2017.

Zahnle, K. J., Kasting, J. F., and Pollack, J. B.: Evolution of a steam atmosphere during Earth's accretion, Icarus, 74, 62–97, 1988.

Zhao, Z., Bons, P., Wang, G., Soesoo, A., and Liu, Y.: Tectonic evolution and high-pressure rock exhumation in the Qiangtang terrane, central Tibet, Solid Earth, 6, 457, 2015.

Zhou, X., Li, Z.-H., Gerya, T. V., and Stern, R. J.: Lateral propagation–induced subduction initiation at passive continental margins controlled by preexisting lithospheric weakness, Science Advances, 6, eaaz1048, 2020.

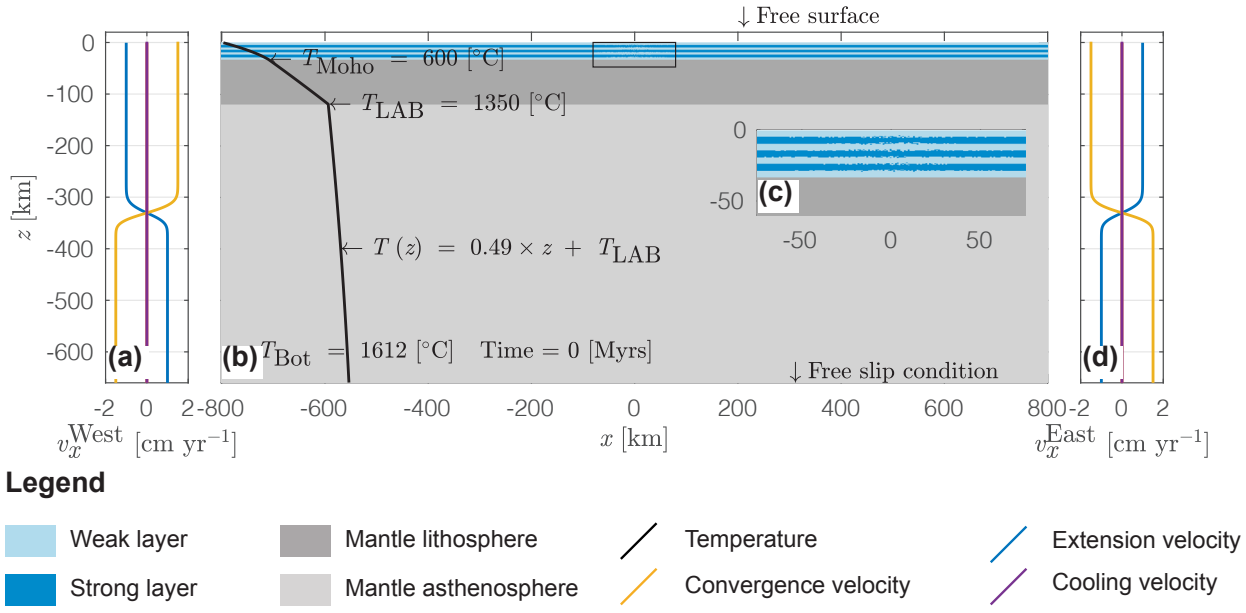

Figure 1. ~~Model~~ Initial model configuration and boundary conditions. Dark blue colours represent strong and light blue colours represent weak crustal units, dark grey colours represent the mantle lithosphere and light grey colours the upper mantle. (a) Profile of horizontal velocity for material inflow and outflow along the ~~left vertical~~ western model boundary. Blue ~~colour~~ line indicates the profile for the extension, purple line indicates the profile for the cooling and ~~red colour~~ the yellow line indicates the profile for the convergence~~, respectively,~~ . (b) ~~entire~~ Entire model domain including the material phases (colour-code as explained above) and initial vertical temperature profile (black line), (c) enlargement of the centre of the model domain showing the initial random perturbation on the ~~horizontal mechanical layer interfaces~~ marker field (see appendix A) used to localise deformation in the centre of the domain and (d) is the same profile as shown in (a), but along the ~~right vertical~~ eastern model boundary.

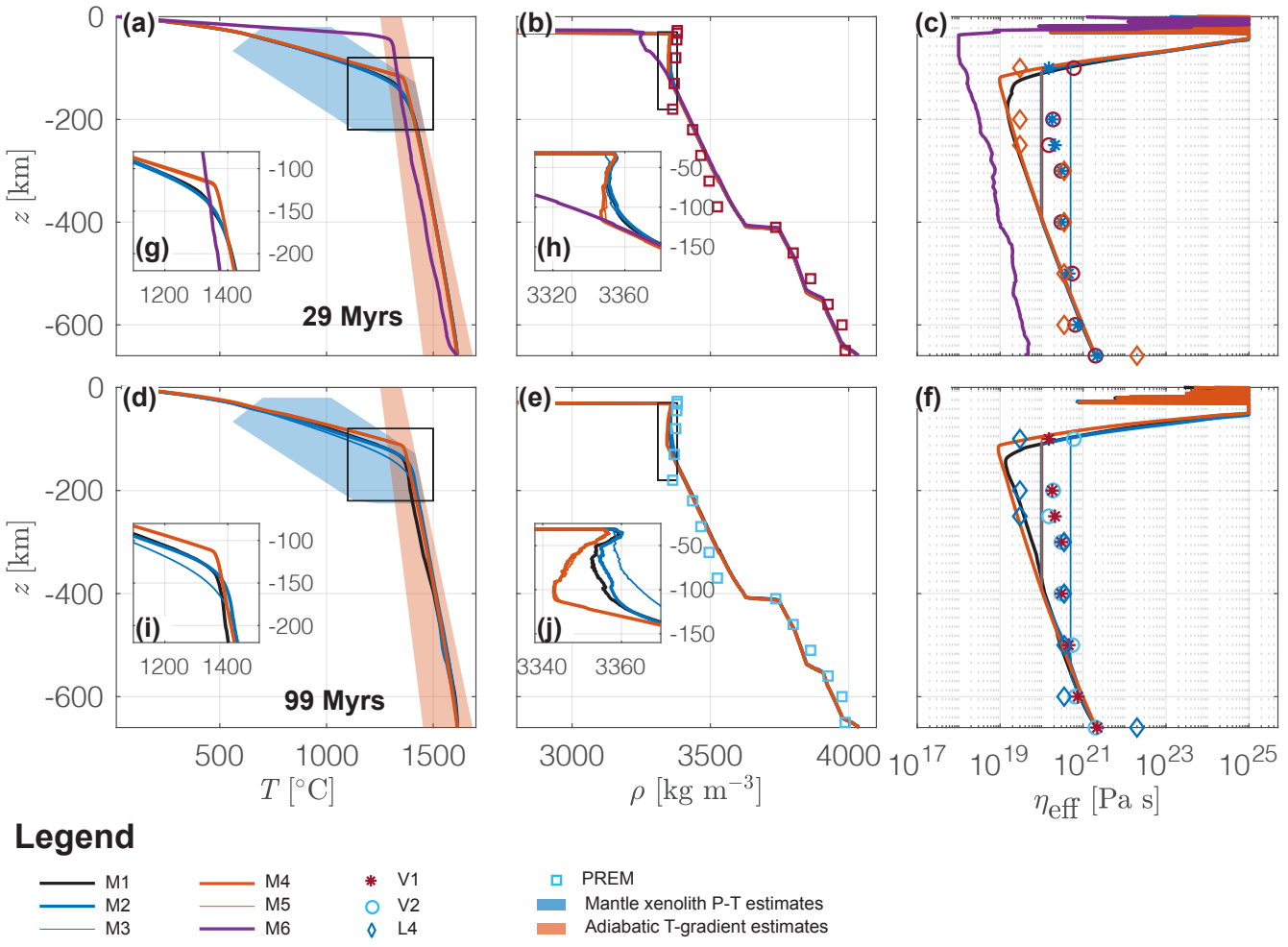

Figure 2. Horizontally averaged vertical profiles of temperature, T, density, ρ, and effective viscosity, η_{eff}. Top row: after 29 Myrs, bottom row: after 99 Myrs in model history. (a) ~~-~~and (d) show horizontally averaged temperature, (b) and (e) ~~Enlargements~~ show horizontally averaged density and (c) and (f) show horizontally averaged viscosity. (g)-(j) show an enlargement of the parental subfigure. Coloured lines show the results of models ~~M1-5 at 13~~ M1-6 as indicated in the legend. In (a) and (d): blue area indicates *P-T* condition estimates from mantle xenolith data and orange area indicates estimates for a range of ~~modelled time~~adiabatic gradients both taken from Hasterok and Chapman (2011) (fig. ~~Blue colours~~ 5). Estimates for adiabatic temperature gradients are extrapolated to 660 km depth. In (b) and (e): squares indicate ~~mechanically weak~~ density estimates from the preliminary reference Earth model (~~light~~PREM) (Dziewonski and Anderson, 1981). In (c) and ~~strong~~ (~~dark~~f)~~crustal units~~: circles and stars indicate viscosity profiles inferred from Occam-style inversion of glacial isostatic adjustment and convection related data originally by Mitrovica and Forte (2004), diamonds show a four-layer model fit of mantle flow to seismic anisotropy originally by Behn et al. (2004). All profiles taken from Forte et al. (2010) (fig. 2). The median is used as statistical quantity for averaging, because it is less sensitive to extreme values compared to the arithmetic mean.

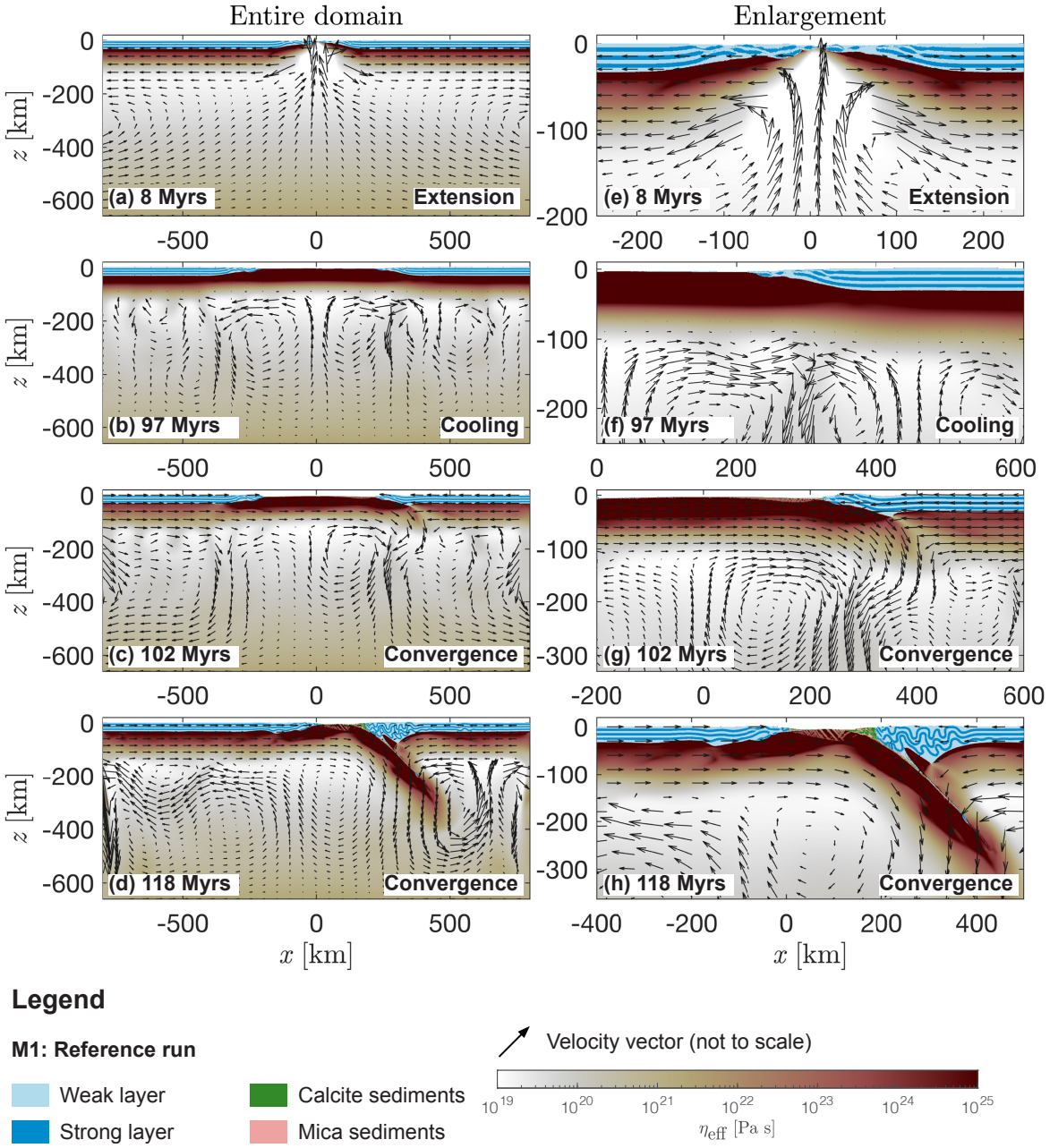

Figure 3. Model evolution of M1 (reference run). Left column: Entire domain. Right column: Enlargement. Dark and light blue colours, salmon and green colours indicate the material phase of weak and strong layers and mica and calcite sediments, respectively. White to red colours ~~show~~ indicate the effective viscosity field calculated by the ~~numerical~~ algorithm~~on a logarithmic scale for the lithospheric mantle~~. Arrows represent velocity vectors and the ~~upper mantle. Grey lines show the level~~ length of ~~550 and 1350 isotherm~~the arrows is not to scale. Scientific colour maps used in all figures are provided by Crameri (2018).

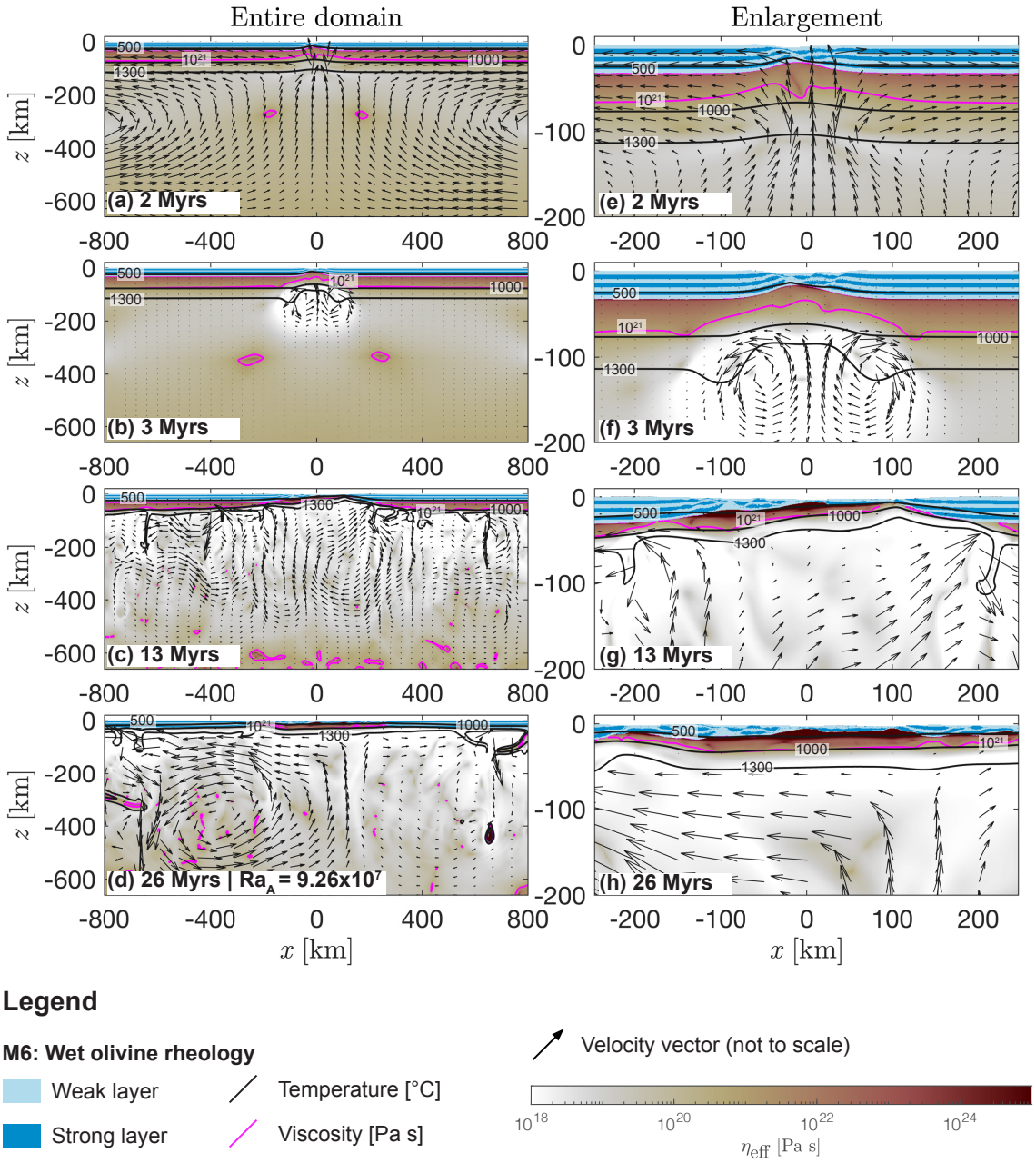

Figure 4. Evolution of model M6 with wet olivine flow law at (a) & (e) 2 Myrs, (b) & (f) 3 Myrs, (c) & (g) 13 Myrs and (d) &(h) 26 Myrs. Blue colours indicate mechanically weak (light) and strong (dark) crustal units. White to red colours show the effective viscosity field calculated by the numerical algorithm~~on a logarithmic scale for the lithospheric mantle and the upper mantle~~. Grey lines show the level of 500 °C, 1000 °C and 1350 °C isotherm and the magenta coloured line shows the level of the 10^{21} Pa s viscosity isopleth. This contour line represents the mechanical boundary between the mantle lithosphere and the convecting upper mantle. Arrows represent velocity vectors and the length of the arrows is not to scale.

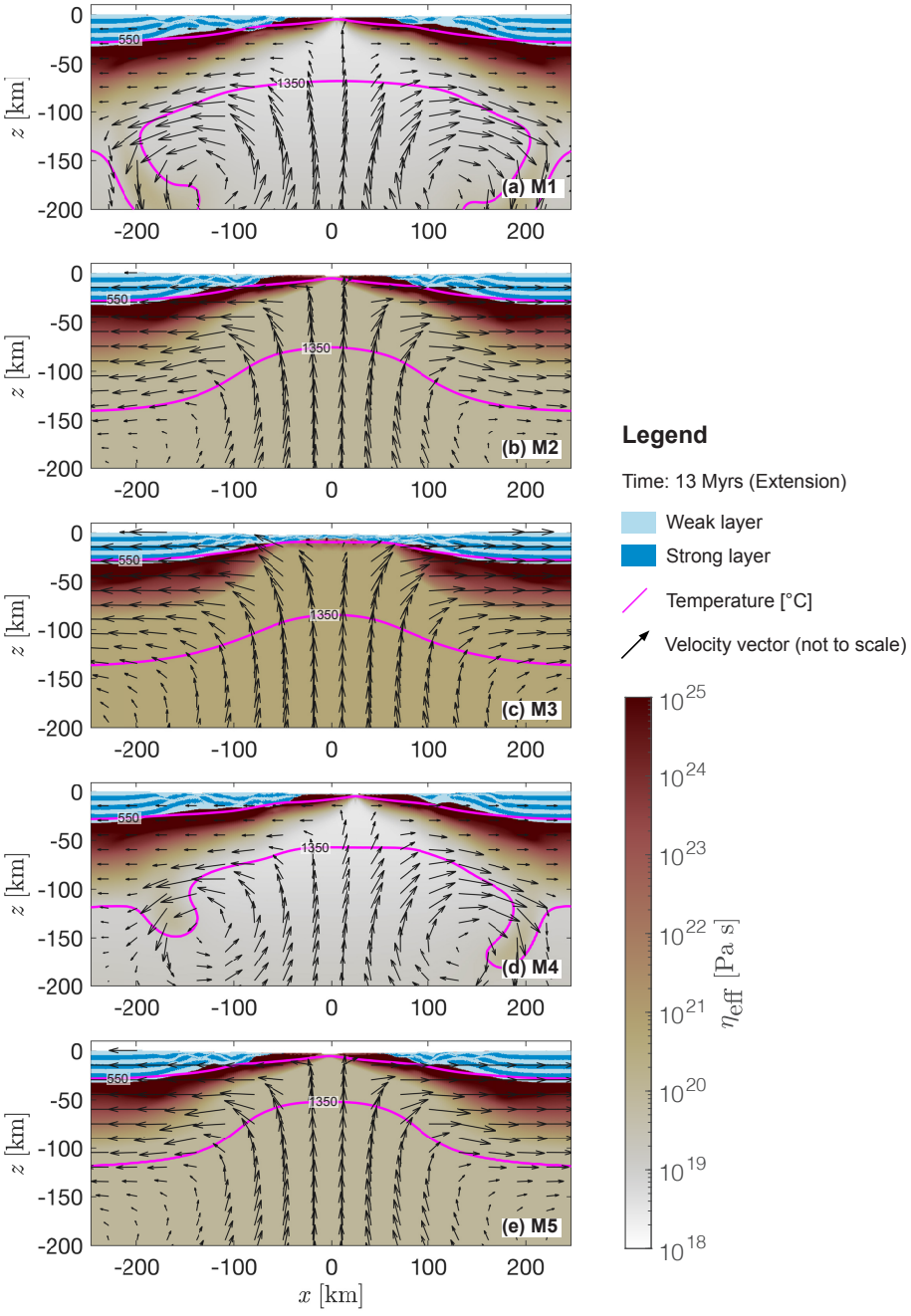

Figure 5. Comparison of M1-5 during the extension stage. (a)-(e) Enlargements of results of models M1-5 at 13 Myrs of modelled time. Blue colours indicate mechanically weak (light) and strong (dark) crustal units. White to red colours show the effective viscosity field calculated by the numerical algorithm. Magenta lines show the level of 550 °C and 1350 °C isotherm. Arrows represent velocity vectors and the length of the arrows is not to scale.

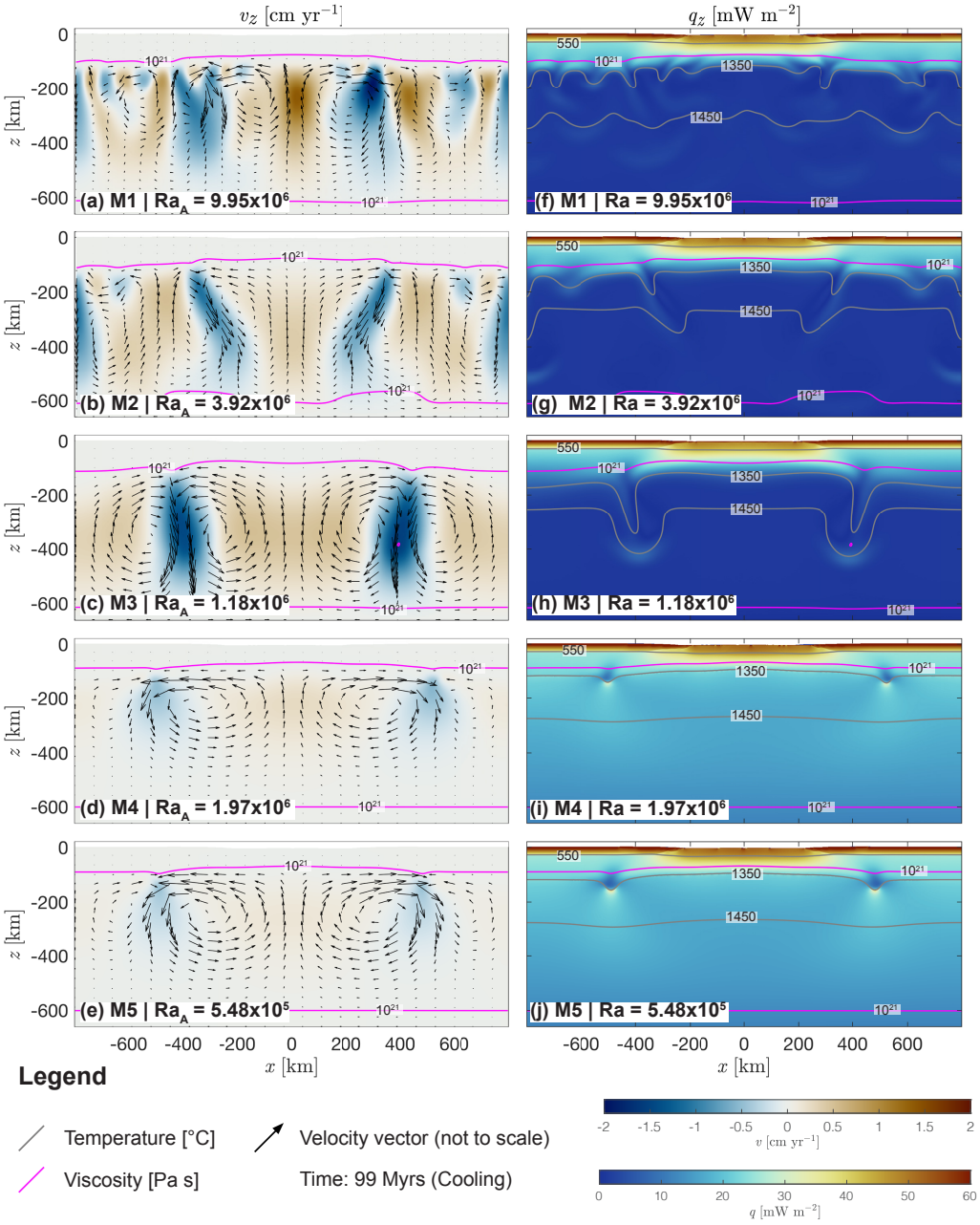

Figure 6. Results of models M1-5 at 99 Myrs (end of ~~thermal relaxation~~ cooling period). (a)-(e) ~~Vertical~~ Blue to red colours indicate the vertical velocity magnitude calculated by the numerical algorithm. Black arrows show not to scale velocity vectors to visualise the material flow field~~.~~, (f)-(j) ~~horizontal velocity field.~~ Blue to red ~~colour~~ colours indicate the ~~velocity magnitudes~~ conductive heat flow calculated by the numerical algorithm. Grey lines show the level of the 550 °C, 1350 °C and 1450 °C isotherms~~and the~~ . The magenta coloured line shows the level of the 10^{21} Pa s viscosity isopleth in all sub-figures. This contour line represents the mechanical boundary between the rigid mantle lithosphere (no velocity glyphs) and the convecting upper mantle.

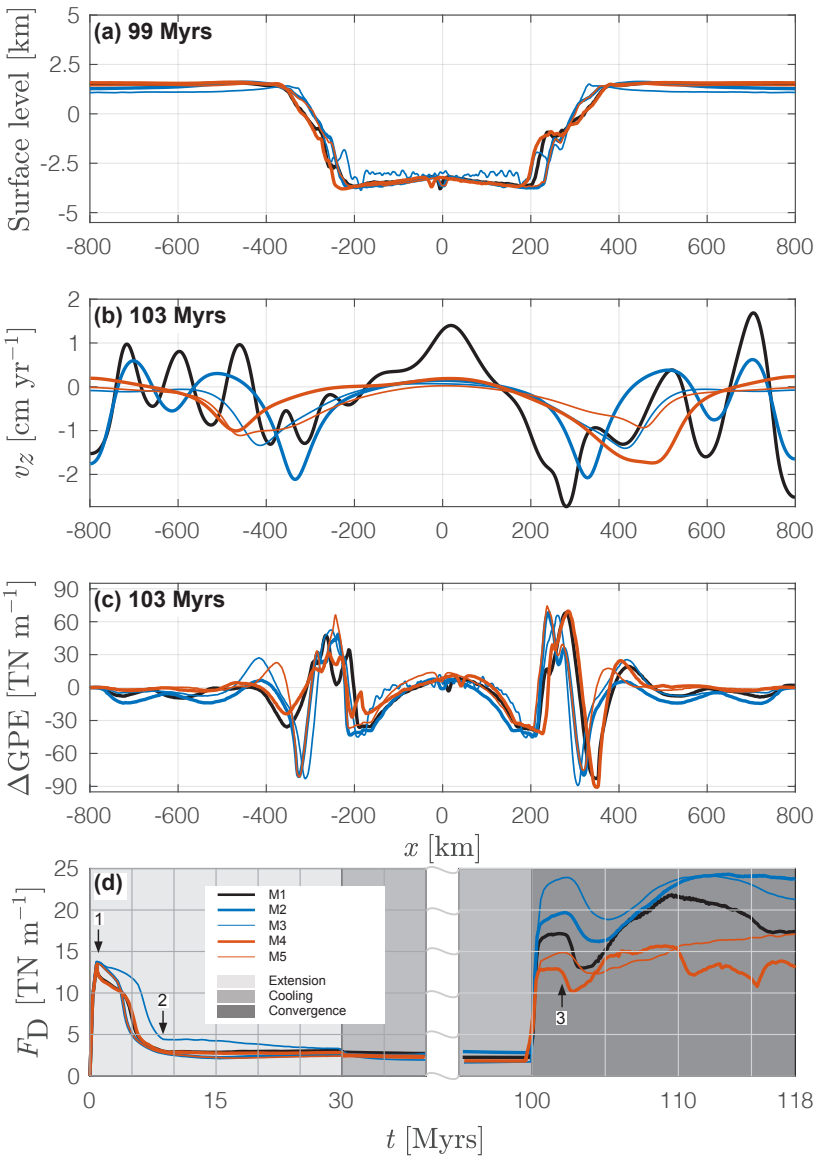

Figure 7. ~~Horizontally averaged vertical profiles of temperature, T, density, ρ, and effective viscosity, η_{eff}. Top row: after 29 , bottom row: after 99 in model history.~~ (a) ~~and (d) show horizontally averaged temperature~~Model topography at the end of the cooling period, (b) horizontal profile of vertical velocity at z≈-175 km and (~~e) show horizontally averaged density and (~~c) ~~and (f) show horizontally averaged viscosity~~difference in *GPE* of models M1-5 after 103 Myrs. (~~g~~d) ~~-~~Estimated plate driving forces (~~h~~F_D, see appendix D) ~~show an enlarged region of~~ through the ~~parental subfigure~~entire model time. ~~Black and red, solid and dashed lines show~~ Grey regions indicate the ~~results~~ stages of ~~models M1-5. In~~ extension (~~a~~light grey)~~and ~~, cooling without plate deformation (~~d~~grey) ~~: blue area indicates P-T condition estimates from mantle xenolith data~~ and ~~orange area indicates estimates for a range of adiabatic gradients both taken from Hasterok and Chapman (2011)~~ convergence (~~fig. 5~~dark grey). ~~Estimates~~ The cooling period is not entirely displayed, because values for ~~adiabatic temperature gradients are extrapolated~~ F_D remain constant when no deformation is applied to ~~660 depth~~the system. ~~In (b) and (e): squares~~ Numbers indicate ~~density estimates from~~ important events in the ~~preliminary reference Earth model (PREM) (Dziewonski and Anderson, 1981). In (c) and (f)~~evolution of the models: ~~circles and stars indicate viscosity profiles inferred from Occam-style inversion~~ 1 = Initiation of ~~glacial isostatic adjustment and convection related data originally by~~

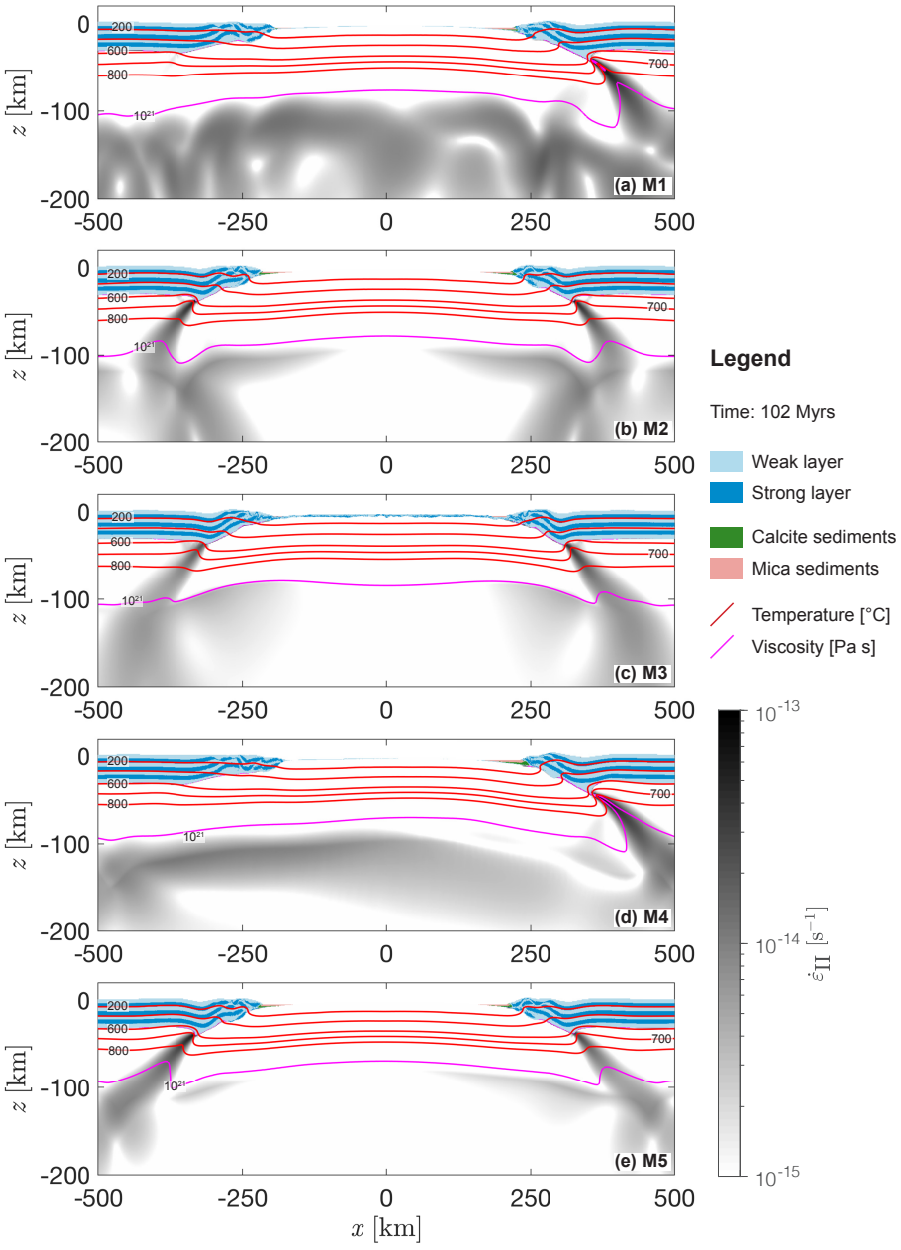

Figure 8. Stage of subduction initiation. (a)-(e) Results of models M1-5 after 102 Myrs of simulated deformation. Blue colours indicate mechanically weak (light) and strong (dark) crustal units, red and salmon colours indicate the sedimentary units (only minor volumes in trench regions). White to black colours indicate the second invariant of the strain rate tensor field calculated by the numerical algorithm ~~on a logarithmic scale~~. Red lines show levels of several isotherms and the magenta coloured line shows the level of the 10^{21} Pa s viscosity isopleth. This contour line represents approximately the mechanical transition between the mantle lithosphere and the convecting upper mantle.

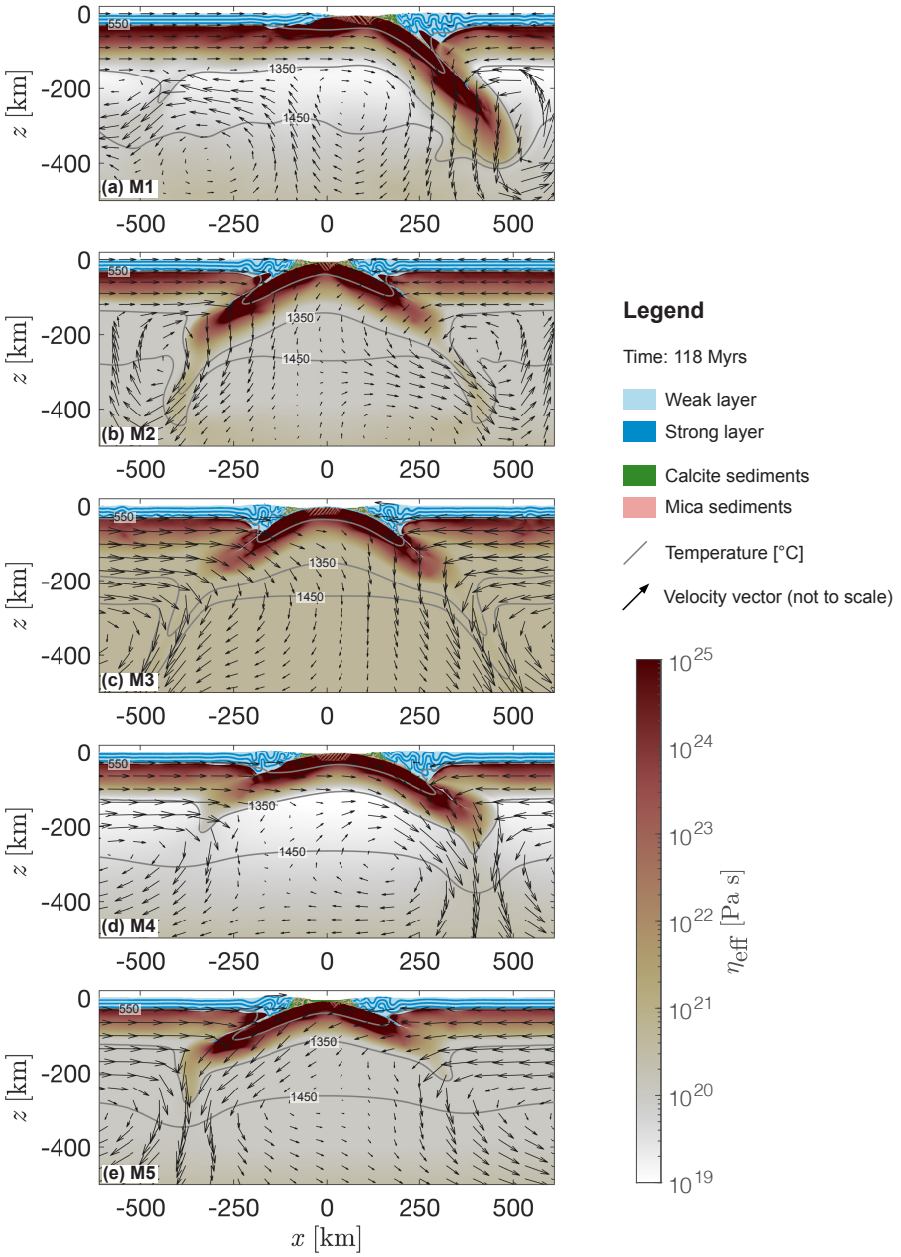

Figure 9. Evolution of subduction zones. (a)-(e) Results of models M1-5 after 118 Myrs of simulated deformation. Blue colours indicate mechanically weak (light) and strong (dark) crustal units, red and salmon colours indicate the sedimentary units. White to red colours show the effective viscosity field calculated by the numerical algorithm ~~on a logarithmic scale~~ for the mantle lithosphere and the upper mantle. ~~Grey lines show the level of the 550, 1350 and 1450 isotherm.(a) Model topography, (b) horizontal profile of vertical velocity at a depth of 175 and (c) difference in gravitational potential energy of models M1-5 after 103 . (d) Vertical integral of two times the second invariant of the deviatoric stress tensor ($\bar{\tau}_{II}$, see appendix D) through the entire model time. Grey regions indicate stage of extension (light grey), thermal relaxation (grey) and convergence (dark grey). The thermal relaxation period is not entirely displayed, because values for $\bar{\tau}_{II}$ remain constant when no deformation is applied to the system. The legend shown in (d) is valid for all panels above as well. Comparison of vertical~~ Arrows represent velocity ~~fields of models M1 (left column) vectors~~ and ~~M4 (right column) at: (a) and (d)~~ the ~~end~~ length of ~~rifting, (b) and (e) end of thermal relaxation period and (c) and (d) during subduction. Blue to red colours indicate velocity values and black lines indicate~~

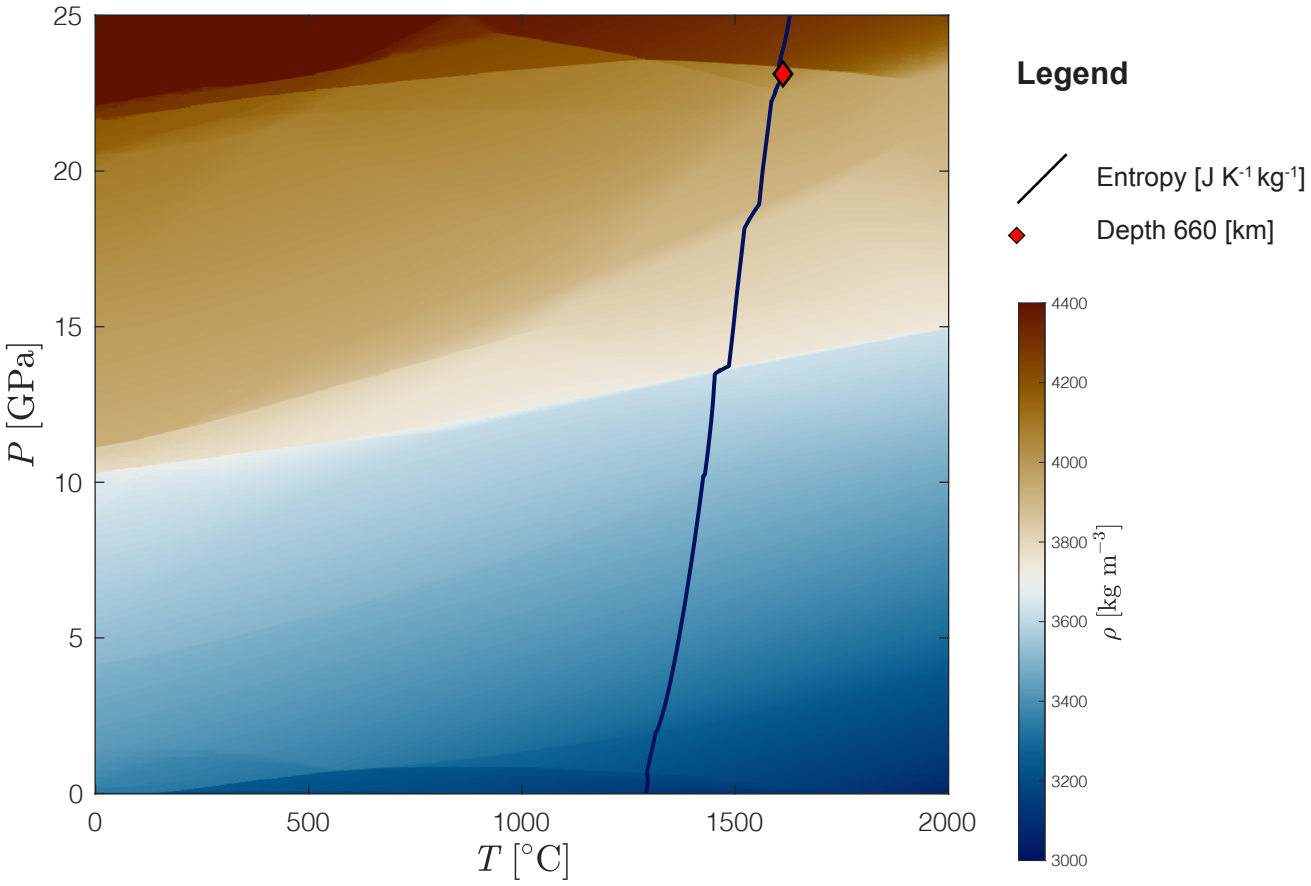

Figure A1. Hawaiian pyrolite phase diagram calculated with Perple_X (Connolly, 2005). Bulk rock composition in weight amount: 44.71 (SiO_2), 3.98 (Al_2O_3), 8.18 (FeO), 38.73 (MgO), 3.17 (CaO) and 0.13 (Na_2O). Bulk rock composition taken from Workman and Hart (2005). Blue to red colours indicate density calculated for the given pressure and temperature range, black line indicates the isentrop for a temperature of 1350 °C at the base of a 120 km thick lithosphere and the red diamond shows the pressure and temperature conditions at 660 km depth following this isentrop.

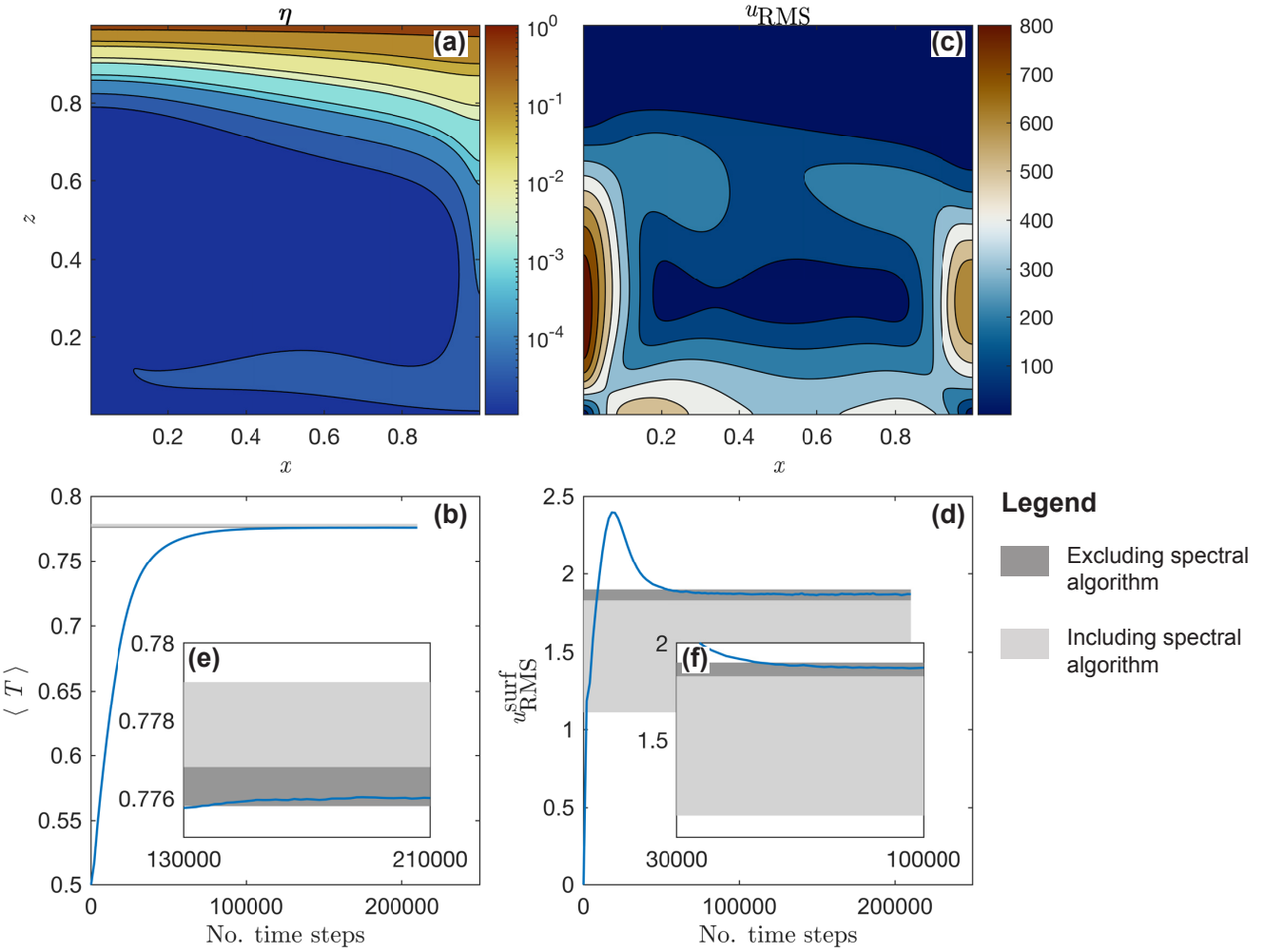

Figure A2. Results of 2D convection benchmark. (a) effective viscosity ~~on logarithmic scale~~ after convection reached a steady state, (b) normal average temperature calculated as in eq. C6 for the entire model history, (c) root mean square velocity field for the entire domain after convection reached a steady state, (d) root mean square velocity at the surface of the modelled domain of the entire model history calculated as in eq. C7, (e) normal average temperature calculated as in eq. C6 for the time period at which the convection reaches a steady state and (f) root mean square velocity at the surface of the modelled domain calculated as in eq. C7 for the time period at which the convection reaches a steady state. In fig. A2(b) and fig. A2(d)-(f), the dark grey area only shows the range of minimum maximum values for the given diagnostic quantity obtained by the algorithms tested by Tosi et al. (2015) excluding the spectral algorithm. Dark combined with light grey areas indicate the range obtained by all algorithms tested by Tosi et al. (2015) including the spectral algorithm.