



HAL
open science

Modelling lithospheric deformation using a compressible visco-elasto-viscoplastic rheology and the effective viscosity approach

Thibault Duretz, René Borst, Philippe Yamato

► To cite this version:

Thibault Duretz, René Borst, Philippe Yamato. Modelling lithospheric deformation using a compressible visco-elasto-viscoplastic rheology and the effective viscosity approach. *Geochemistry, Geophysics, Geosystems*, 2021, 22 (8), pp.e2021GC009675. 10.1029/2021GC009675 . insu-03283712

HAL Id: insu-03283712

<https://insu.hal.science/insu-03283712>

Submitted on 12 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modelling lithospheric deformation using a compressible visco-elasto-viscoplastic rheology and the effective viscosity approach

Thibault Duretz^{1,2}, René de Borst³, Philippe Yamato^{1,4}

¹Univ Rennes, CNRS, Géosciences Rennes UMR 6118, F-35000 Rennes, France

²Institut des Sciences de la Terre, University of Lausanne, Lausanne, Switzerland

³University of Sheffield, Department of Civil and Structural Engineering, Sheffield S1 3JD, UK

⁴Institut Universitaire de France (IUF)

Key Points:

- The effective viscosity approach is extended to include compressible elasto-plasticity, plastic dilatancy, power-law viscoplasticity and softening.
- This extended effective viscosity approach gives results as accurate as algorithms in engineering which use return mapping and consistent linearisation.
- Applications to crustal-scale shear banding and long-term lithospheric deformations are provided.

Corresponding author: T. Duretz, thibault.duretz@univ-rennes1.fr

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1029/2021GC009675](https://doi.org/10.1029/2021GC009675).

This article is protected by copyright. All rights reserved.

Abstract

Deformations of the colder regions of the lithosphere mainly occur in the frictional regime. In geodynamic models, frictional plastic deformations are often highly localised (shear bands) and are used as proxies for faults. However, capturing the generation and evolution of shear bands in geodynamic models is troublesome. Indeed, mesh dependency and lack of convergence affect, to some extent, the results of geodynamic models. Here we extend the most common plasticity implementation used in geodynamic codes (effective viscosity approach) to include the combined effects of elasto-plastic compressibility, plastic dilatancy, strain softening and viscoplasticity. The latter acts as a regularisation that cures most of the known issues of geodynamic models related to frictional plasticity. Using regularised models based on the M2Di MATLAB routines, we show that volumetric elasto-plastic deformations can significantly impact crustal-scale shear banding. We also show that the artificial overstress caused by viscoplasticity can be mitigated by employing power-law models. Furthermore, we demonstrate that plasticity algorithms common in geodynamics (based on the effective viscosity approach) can be as accurate as those obtained with algorithms typically used in engineering (return mapping with a consistent tangent operator). Finally, we show examples of long-term tectonic deformations using the state-of-the-art geodynamic code MDoodz. They indicate that viscoplastic regularisation can be used efficiently to obtain reliable simulations in geodynamics.

1 Introduction

In most geodynamic settings a large volume of the crust exhibits frictional-plastic deformations. Together with the equilibrium condition this rheological model can explain the occurrence of high-angle faults at the onset of extensional faulting and low-angle faults in compression. Moreover, this rheological model has been well calibrated by laboratory experiments of rock deformations [e.g. *Byerlee, 1978; Rutter and Glover, 2012*] and has further been validated in the context of deep continental drilling [*Zoback et al., 1993*]. In order to properly capture geologically relevant stress states and structures, geodynamic models must therefore include frictional plasticity. However, accounting for such a rheological model is not trivial and remains challenging. In particular, the modelling of localised shear bands, which serve as proxies for faults, causes issues. The frictional plastic rheology becomes unstable when softening or non-associated plastic flow are included

48 and this can lead to the formation of localised shear bands [Rudnicki and Rice, 1975].
49 Yet, most current implementations do not incorporate an internal length scale, which
50 is necessary to constrain the shear band to a finite width. The absence of such a length
51 scale causes patterns of the modelled faults to be fractal [Poliakov *et al.*, 1993] and their
52 properties (dimensions, number, stresses and strains) to depend on the numerical res-
53 olution [e.g. Duretz *et al.*, 2019]. Another consequence is that the models are numeri-
54 cally unstable in the sense that divergence of the equilibrium-finding iterative process
55 is often observed [Spiegelman *et al.*, 2016; Duretz *et al.*, 2018].

56 To address these issues numerous regularisation schemes have been proposed. All
57 of them introduce an internal length scale. Non-local plasticity involves a typical dimen-
58 sion of the area over which the plastic strain is averaged [Bažant and Lin, 1988]. Gra-
59 dient plasticity includes the spatial gradients of the plastic strain in the yield function
60 [de Borst and Mühlhaus, 1992]. Cosserat models include micro-rotations to reflect the
61 micro-structure of the material. The bending modulus which sets the stiffness between
62 the ensuing micro-curvatures and the couple stresses then introduces a length scale as
63 the quotient of this bending modulus and Young’s modulus has the dimension of length
64 [Mühlhaus and Vardoulakis, 1987; Stefanou *et al.*, 2017; Sabet and de Borst, 2019]. While
65 all of these approaches have been applied successfully, they typically require more com-
66 putational power than continuum models which are not enriched, since they either re-
67 quire additional degrees of freedom (rotational degrees of freedom in Cosserat media)
68 or complicated, time-consuming averaging procedures. Herein we focus on viscoplastic
69 regularisation, which has recently been applied to problems in geodynamics [Duretz *et al.*,
70 2019; Jacquey and Cacace, 2020; Duretz *et al.*, 2020]. This regularisation approach is based
71 on the inclusion of rate dependence of the yield function [Wang *et al.*, 1997; de Borst
72 and Duretz, 2020]. Viscoplasticity thus introduces explicitly a time scale rather than a
73 length scale [e.g. Wang, 2019]. The methodology is purely local and hence does not re-
74 quire the introduction of additional degrees of freedom at the global level. Likely, vis-
75 coplasticity is the simplest possible regularisation technique and its implementation in
76 existing codes is fairly trivial. While it is not the most rigorous regularisation technique,
77 especially for quasi-static (slow) process, the benefits of viscoplasticity are that a diver-
78 gence of the equilibrium-finding iterative procedure is usually avoided and that mesh de-
79 pendence is vastly reduced. Nevertheless, the method introduces an artificial overstress,
80 which may affect the solution – an issue that we will thoroughly address in this paper.

81 In this study we provide a detailed description of the implementation of viscoplas-
 82 ticity for geodynamic codes that are based on the velocity-pressure formulation and on
 83 the effective viscosity approach (EVA) for the plastic rheology [e.g. *Willett, 1992; Moresi*
 84 *et al., 2003; Gerya and Yuen, 2007; Lemiale et al., 2008; May et al., 2014; Kaus et al.,*
 85 *2016; Spiegelman et al., 2016; Glerum et al., 2018; Naliboff et al., 2020*]. We consider a
 86 compressible visco-elasto-viscoplastic (V-E-VP) formulation, power-law viscoplasticity
 87 and various softening laws. Both local rheological computations (local iterations, return
 88 mapping) and global computations (Newton-Raphson iterations) are explained in detail.
 89 Numerical implementation is based on the Finite Difference Method. A description of
 90 an equivalent formulation in the context of the Finite Element Method and displacement-
 91 pressure formulation can be found in *Commend et al. [2004]*. All the results shown can
 92 be reproduced using the open source MATLAB routines based on M2Di [*Räss et al., 2017*]
 93 (https://bitbucket.org/lraess/m2di/src/master/M2Di2-VEVP_Compressible) and the geo-
 94 dynamic modelling code MDoodz (<https://github.com/tduretz/MDOODZ6.0>).

95 2 Model formulation

We consider steady-state deformations of a compressible V-E-VP medium, so that
 the balance of momentum takes the form

$$\frac{\partial \sigma_{ij}}{\partial x_j} + \rho g_i = 0, \quad (1)$$

96 The body force acting on the medium is due to gravity acceleration g . x contains the
 97 spatial coordinates, ρ corresponds to the density and σ is the total stress tensor. The
 98 latter relates to the deviatoric stress τ and to the pressure p via $\sigma_{ij} = p\delta_{ij} + \tau_{ij}$, with
 99 δ_{ij} the Kronecker delta. The total strain rate $\dot{\epsilon}$ relates to the deviatoric and the volu-
 100 metric strain rates, $\dot{\epsilon}$ and $v_{k,k}$, respectively, as follows: $\dot{\epsilon}_{ij} = \dot{\epsilon}_{ij} + \frac{1}{3}v_{k,k}\delta_{ij}$, with v the
 101 velocity vector. The Einstein summation convention applies and the notation $_{,k}$ implies
 102 differentiation with respect to x_k .

The rheological model is based on the additive decomposition of the deviatoric strain
 rate tensor:

$$\dot{\epsilon}_{ij} = \dot{\epsilon}_{ij}^v + \dot{\epsilon}_{ij}^e + \dot{\epsilon}_{ij}^{vp}, \quad (2)$$

where the superscripts v, e and vp stand for elastic, viscous and viscoplastic, respectively.
 We consider isotropic power-law creep, linear isotropic elasticity and viscoplasticity such

that the additive decomposition can be expressed as:

$$\dot{\varepsilon}_{ij} = \frac{\tau_{ij}}{2\eta^v} + \frac{\dot{\tau}_{ij}}{2G} + \dot{\lambda} \frac{\partial Q}{\partial \tau_{ij}}, \quad (3)$$

where η^v is the effective creep viscosity, G is the shear modulus, $\dot{\lambda}$ is the rate of the plastic multiplier and Q is the plastic flow potential. The effective power-law creep viscosity is formulated as:

$$\eta^v = \eta_0^v \dot{\varepsilon}_{\text{II}}^v \frac{1}{n^v - 1}, \quad (4)$$

where η_0^v and n^v are material parameters which can be calibrated using laboratory experiments. $\dot{\varepsilon}_{\text{II}}^v$ is the second invariant of the viscous part of the deviatoric strain rate tensor defined as $\dot{\varepsilon}_{\text{II}}^v = \sqrt{\frac{1}{2}(\dot{\varepsilon}_{xx}^v{}^2 + \dot{\varepsilon}_{yy}^v{}^2 + \dot{\varepsilon}_{zz}^v{}^2) + \dot{\varepsilon}_{xy}^v{}^2}$. Here we consider the case of plane strain deformation, hence the out-of-plane deviatoric strain rate and deviatoric stress components do not vanish. Henceforth, we will consider a viscoplastic Drucker-Prager model with the yield function

$$F = \tau_{\text{II}} - p \sin \phi - c \cos \phi - \dot{\lambda} \eta^{\text{VP}}, \quad (5)$$

where τ_{II} is the second invariant of the deviatoric stress tensor defined as $\tau_{\text{II}} = \sqrt{\frac{1}{2}(\tau_{xx}^2 + \tau_{yy}^2 + \tau_{zz}^2) + \tau_{xy}^2}$, c is the cohesion, ϕ is the friction angle and η^{VP} is the viscoplastic viscosity. The parameters of the Drucker-Prager plasticity model are typically calibrated using experimental data [e.g. *Byerlee, 1978; Rutter and Glover, 2012*]. However classical local and rate-independent non-associated plasticity models cause mesh-dependence of numerical solutions. We employ a viscoplastic model to reduce the effects of mesh dependence during non-associated plastic flow. For generality, we express the viscoplastic viscosity using a power-law relation, so that:

$$\eta^{\text{VP}} = \eta_0^{\text{VP}} \dot{\lambda} \frac{1}{n^{\text{VP}} - 1}, \quad (6)$$

103 where η_0^{VP} and n^{VP} are material parameters. In this contribution, we do not aim to re-
104 late the parameters of the viscoplastic model to experimental rock deformation data. Vis-
105 coplasticity is here only used for its time regularisation effect.

Compared to inviscid plasticity models viscoplastic models introduce a rate-dependent overstress, $\bar{\sigma}$ [e.g. *Heeres et al., 2002; Niazi et al., 2013; de Borst and Duretz, 2020*]. The use of a power-law exponent in excess of 1 can significantly reduce overstress variations induced by changes in the magnitude of viscoplastic strain rate, see Figure (1A). Hardening/softening laws are defined for the variables that control the evolution of plastic-

ity:

$$\xi = \xi_0 - \frac{\Delta\xi}{2} \operatorname{erfc}\left(-\frac{\varepsilon^{\text{VP}} - \mu_\xi}{\sigma_\xi}\right), \quad (7)$$

where ξ represents either the cohesion, friction or dilatancy angle. ξ_0 indicates the initial value of ξ , while $\Delta\xi$, μ_ξ and σ_ξ correspond to the amplitude, the mean and the standard deviation of the prescribed variation of ξ , respectively. σ_ξ controls the rate of softening. Its effect on the variation of ξ with viscoplastic strain is depicted in Figure (1B). Such a non-linear hardening/softening law differs from the piecewise-linear laws which are commonly used in geodynamics. The advantage is that it does not include singularities and is continuously differentiable, which is very suitable in implicit rheological modelling using Newton-Raphson linearisation.

The accumulated viscoplastic strain tensor is expressed as

$$\varepsilon^{\text{VP}} = \int \sqrt{\frac{2}{3} \frac{\partial Q}{\partial \tau_{ij}}^T \frac{\partial Q}{\partial \tau_{ij}}} d\varepsilon^{\text{VP}} \equiv \int h d\varepsilon^{\text{VP}}. \quad (8)$$

The plastic flow potential is defined as:

$$Q = \tau_{\text{II}} - p \sin \psi, \quad (9)$$

where ψ is the dilatancy angle. This plasticity model is non-associated ($\frac{\partial F}{\partial p} \neq \frac{\partial Q}{\partial p}$) unless ψ equals ϕ . Plastic flow only occurs if $F \geq 0$. Then, the rate of the plastic multiplier is positive.

The volumetric rheological model is based on an additive decomposition of the divergence of velocity:

$$v_{k,k} = v_{k,k}^e + v_{k,k}^{\text{VP}}. \quad (10)$$

Thus, we assume that volumetric deformation is caused by either elasticity or viscoplasticity, and that volumetric viscous creep is excluded. This leads to:

$$v_{k,k} = -\frac{\dot{p}}{K} - \dot{\lambda} \frac{\partial Q}{\partial p}, \quad (11)$$

where K stands for the bulk modulus. When $\psi = 0$, the term $\frac{\partial Q}{\partial p} = -\sin \psi$ vanishes and the model reaches viscoplastic incompressibility. The fully incompressible limit $v_{k,k} = 0$ is obtained by assuming elastic incompressibility and setting the dilatancy angle to zero.

3 Numerical implementation

The momentum balance has been discretised using the staggered grid finite-difference method. The components of the velocity vector (v_i) and the pressure (p) are considered

123 as the primitive variables (i.e. a velocity-pressure formulation). The determination of
 124 velocity and pressure fields that satisfy momentum balance equation is a global proce-
 125 dure that requires the solution of a system of equations. Moreover, due to non-linear rhe-
 126 ological models, this procedure is iterative and requires successive global non-linear it-
 127 erations. The deviatoric rheological implementation is based on the effective viscosity
 128 approach (EVA), which is used in most geodynamic codes. We use a predictor-corrector
 129 procedure whereby the trial visco-elastic stress is computed at each global iteration and
 130 corrected in case of viscoplastic flow. Both the predictor and corrector steps involves non-
 131 linearities. The considered rheological model is local as, e.g., the yield function does not
 132 dependent on the gradients of the plastic strain. Hence all rheological computations are
 133 local procedures in the sense that there are applied to each cell/vertex or integration points
 134 independently.

135 3.1 Local rheological procedures: non-linear visco-elastic predictor

The stress rate is integrated using a backward Euler scheme, $\tau_{ij} = \tau_{ij}^0 + \dot{\tau}_{ij}\Delta t$,
 with τ^0 the deviatoric stress tensor at the previous time step and Δt is the time step.
 We now define $\eta^e = G\Delta t$. Then, Eq. (3) can be integrated to yield:

$$\tau_{ij} = 2\eta^{\text{ve}} \left(\dot{\epsilon}'_{ij} - \dot{\lambda} \frac{\partial Q}{\partial \tau_{ij}} \right) \quad (12)$$

136 where $\eta^{\text{ve}} = \left(\frac{1}{\eta^v} + \frac{1}{\eta^e} \right)^{-1}$ and $\dot{\epsilon}'_{ij} = \dot{\epsilon}_{ij} + \frac{\tau_{ij}^0}{2\eta^e}$ is an effective deviatoric strain rate ten-
 137 sor that accounts for the time discretisation of the stress rate [e.g. *Moresi et al.*, 2003;
 138 *Kaus et al.*, 2016; *Bauville et al.*, 2020]. Similarly, we define the effective divergence $v'_{k,k} =$
 139 $v_{k,k} - \frac{p^0}{K\Delta t}$, where p^0 is the pressure from the previous time step.

For a non-linear viscous creep model there is no closed-form expression for η^{ve} and
 a local non-linear iteration is needed [*Popov and Sobolev*, 2008]. An exact additive par-
 titioning of the elastic and viscous deviatoric strain rates is ensured by the non-linear
 identity:

$$f(\eta^{\text{ve}}) = \dot{\epsilon}'_{\text{II}} - \frac{\tau_{\text{II}}}{2\eta^e} - \dot{\epsilon}_{\text{II}}^v = 0, \quad (13)$$

140 where $\dot{\epsilon}'_{\text{II}}$ is the second invariant of $\dot{\epsilon}'_{ij}$, $\tau_{\text{II}} = 2\eta^{\text{ve}}\dot{\epsilon}'_{\text{II}}$ is the second invariant of the de-
 141 viatoric stress tensor and $\dot{\epsilon}_{\text{II}}^v = C^v \tau_{\text{II}}^{n_v}$ is the viscous strain rate with $C^v = (2\eta_0^v)^{-n_v}$.

A local Newton-Raphson scheme can be used to solve $f(\eta^{\text{ve}}) = 0$. The effective
 visco-elastic viscosity is then determined iteratively by incrementing successive correc-

tions

$$\delta\eta^{\text{ve}} = - \left(\frac{\partial f}{\partial \eta^{\text{ve}}} \right)^{-1} f(\eta^{\text{ve}}), \quad (14)$$

where

$$\frac{\partial f}{\partial \eta^{\text{ve}}} = -2 \frac{\dot{\epsilon}'_{\text{II}}}{2\eta^{\text{e}}} - C^{\text{v}} n^{\text{v}} \frac{\tau_{\text{II}}}{\eta^{\text{ve}}}. \quad (15)$$

This procedure typically converges to machine precision in less than five iterations. The algorithm is strain-rate driven and $\dot{\epsilon}'_{\text{II}}$ is constant during these local iterations.

The viscous creep model does not include volumetric deformation. Hence, viscous creep is assumed to be purely deviatoric and does not feed back on the pressure field. The inclusion of a composite creep model based on different flow law expressions (exponential creep, grain-size evolution) is straightforward [e.g. *Popov and Sobolev, 2008; Schmalholz and Duretz, 2017; Bessat et al., 2020*]. The resulting deviatoric stress serves as a deviatoric trial stress for the subsequent plasticity computations ($\tau_{\text{II}}^{\text{trial}} = \tau_{\text{II}}$).

3.2 Local rheological procedures: non-linear viscoplastic corrector

Once the deviatoric trial stress has been determined, the condition for yielding,

$$F^{\text{trial}} = \tau_{\text{II}}^{\text{trial}} - p^{\text{trial}} \sin \phi - c \cos \phi \quad (16)$$

can be evaluated. As discussed, $\tau_{\text{II}}^{\text{trial}}$ is determined in the predictor stage, while p^{trial} is extracted from the global solution vector. The trial yield function assumes no viscoplastic flow, hence $\dot{\lambda} = 0$ at this stage. For viscoplastic yielding ($F^{\text{trial}} \geq 0$), the rate of plastic multiplier ($\dot{\lambda}$) needs to be evaluated, which will be used to determine the deviatoric viscoplastic strain rate tensor, $\dot{\epsilon}_{ij}^{\text{vp}} = \dot{\lambda} \frac{\partial Q}{\partial \tau_{ij}} = \dot{\lambda} \left[\frac{\tau_{xx}}{2\tau_{\text{II}}} \frac{\tau_{yy}}{2\tau_{\text{II}}} \frac{\tau_{zz}}{2\tau_{\text{II}}} \frac{\tau_{xy}}{\tau_{\text{II}}} \right]^T$ and the volumetric viscoplastic strain rate, $v_{k,k}^{\text{vp}} = -\dot{\lambda} \frac{\partial Q}{\partial p} = \dot{\lambda} \sin \psi$. Viscoplastic flow then implies a reduction of the deviatoric stress such that:

$$\tau_{\text{II}}^{\text{corr}} = \tau_{\text{II}}^{\text{trial}} - \eta^{\text{ve}} \dot{\lambda}. \quad (17)$$

The individual corrected deviatoric stress components may be evaluated as:

$$\tau_{ij}^{\text{corr}} = 2\eta^{\text{ve}} \left(\dot{\epsilon}'_{ij} - \dot{\epsilon}_{ij}^{\text{vp}} \right) = \tau_{ij}^{\text{trial}} - 2\eta^{\text{ve}} \dot{\lambda} \frac{\partial Q}{\partial \tau_{ij}}. \quad (18)$$

If elasto-plastic volume changes are considered, viscoplastic flow feeds back into the pressure. The pressure update then includes a correction proportional to the amount of viscoplastic volume change:

$$p^{\text{corr}} = p^{\text{trial}} + \dot{\lambda} K \Delta t \sin \psi \quad (19)$$

In the incompressible limit, the pressure can be interpreted as a Lagrange multiplier that enforces incompressibility. Hence, a pressure update may not be explicitly written and the rheological model does not require any local pressure corrections. The accumulated viscoplastic strain is updated with the following increment:

$$\Delta \varepsilon_{\text{II}}^{\text{VP}} = h \dot{\lambda} \Delta t \quad (20)$$

The corrected yield function is thus expressed as

$$F = \tau_{\text{II}}^{\text{trial}} - \eta^{\text{ve}} \dot{\lambda} - p^{\text{corr}} \sin \phi - c \cos \phi - \eta^{\text{VP}} \dot{\lambda} \quad (21)$$

For ideal, linear viscoplasticity (so without hardening or softening), setting $F = 0$ yields the following expression of plastic multiplier rate [*de Borst and Feenstra, 1990*]:

$$\dot{\lambda} = \frac{F^{\text{trial}}}{\eta^{\text{ve}} + \eta^{\text{VP}} + K \Delta t \sin \psi \sin \phi}. \quad (22)$$

However, as soon as either η^{VP} , c , ϕ or ψ involve non-linear expressions, the determination of $\dot{\lambda}$ necessitates non-linear iterations. This procedure is again achieved via local Newton-Raphson iterations, where corrections to the rate of the plastic multiplier are expressed as:

$$\delta \dot{\lambda} = - \left(\frac{\partial F}{\partial \dot{\lambda}} \right)^{-1} F(\dot{\lambda}), \quad (23)$$

with

$$\frac{\partial F}{\partial \dot{\lambda}} = -K \Delta t \sin \psi \sin \phi - \eta^{\text{ve}} - \frac{\eta^{\text{VP}}}{n^{\text{VP}}} - H, \quad (24)$$

where $H = \frac{\partial c}{\partial \dot{\lambda}} \cos \phi - \frac{\partial \phi}{\partial \dot{\lambda}} (c \sin \phi - p^{\text{corr}} \cos \phi) + K \Delta t \dot{\lambda} \cos \psi \sin \phi \frac{\partial \psi}{\partial \dot{\lambda}}$. The partial derivatives of c , ϕ or ψ are of the form:

$$\frac{\partial \xi}{\partial \dot{\lambda}} = - \frac{h \Delta t \Delta \xi}{\sqrt{\pi} \sigma_{\xi}} \exp \left[- \frac{(\mu_{\xi} - \varepsilon_{\text{II}}^{\text{pl}})^2}{\sigma_{\xi}^2} \right] \quad (25)$$

where ξ is either c , ϕ or ψ .

Upon convergence, the stress update is be formulated as:

$$\tau_{ij} = 2\eta^{\text{ve}} \left(\dot{\varepsilon}'_{ij} - \dot{\lambda} \frac{\partial Q}{\partial \tau_{ij}} \right) \equiv 2\eta^{\text{vep}} \dot{\varepsilon}'_{ij} \quad (26)$$

where

$$\eta^{\text{vep}} = \frac{p^{\text{corr}} \sin \phi + c \cos \phi + \eta^{\text{VP}} \dot{\lambda}}{2 \dot{\varepsilon}'_{\text{II}}}. \quad (27)$$

The effective visco-elasto-viscoplastic viscosity can then be used for the resolution of the global non-linear problem. Implementation details regarding local rheological computations (predictor and corrector phases) can be found in the Supporting Material (Code S1).

3.3 Global Newton-Raphson iterations

The activation of non-linear rheological elements (e.g. power-law creep, frictional plasticity) introduces a non-linearity at the global level. In order to reach global equilibrium, successive global Newton-Raphson iterations are applied. The corrections to the velocity and the pressure are given by:

$$\begin{bmatrix} \delta v \\ \delta p \end{bmatrix} = \mathbf{J}^{-1} \begin{bmatrix} f_v \\ f_p \end{bmatrix}. \quad (28)$$

where \mathbf{J} represents the Jacobian matrix and f_v , f_p are the residuals, defined as:

$$\begin{aligned} f_{v_i} &= \frac{\partial \tau_{ij}}{\partial x_j} - \frac{\partial p}{\partial x_i} + \rho g_i, \\ f_p &= -\frac{dp}{dt} - K \left(v_{k,k} - v_{k,k}^{\text{vp}} \right). \end{aligned} \quad (29)$$

It should be noted that the evaluation of the above residual equations requires knowledge of corrected deviatoric stress components and corrected pressure in case of viscoplastic deformation (see Sec 3.2). If p^{trial} is considered as the global pressure variable, the time discretized residuals are expressed as:

$$\begin{aligned} \tilde{f}_{v_i} &= \frac{\partial \tau_{ij}}{\partial x_j} - \frac{\partial p^{\text{trial}}}{\partial x_i} - \frac{\partial K \Delta t v_{k,k}^{\text{vp}}}{\partial x_i} + \rho g_i \\ \tilde{f}_p &= -\frac{p^{\text{trial}} - p^0}{\Delta t} - K v_{k,k}, \end{aligned} \quad (30)$$

where p^0 is the pressure from the preceding time increment and Δt is the time step. Alternatively one may define p^{corr} as the global pressure variable, in this case the time-discretized residuals are expressed as:

$$\begin{aligned} \tilde{f}_{v_i} &= \frac{\partial \tau_{ij}}{\partial x_j} - \frac{\partial p^{\text{corr}}}{\partial x_i} + \rho g_i, \\ \tilde{f}_p &= -\frac{p^{\text{corr}} - p^0}{\Delta t} - K \left(v_{k,k} - v_{k,k}^{\text{vp}} \right). \end{aligned} \quad (31)$$

The assembly of the Jacobian matrix requires the evaluation of the tangential operators, $\frac{\partial \tau_{ij}}{\partial \varepsilon_{kl}}$ and $\frac{\partial p}{\partial v_{k,k}}$, at each cell or integration point. The latter can be obtained either via a consistent linearisation of the finite-step visco-elasto-viscoplastic relation (as customary in computational engineering), or by explicitly evaluating the derivatives of the effective viscosity η^{vep} , which is commonly done in computational geodynamics. Herein, we have tested both approaches and the corresponding analytical expressions are given in the Appendices A and B, for the Newton linearisation of the effective viscosity approach and the finite-step consistent tangent linearisation, respectively. It is emphasised that, when the linearisation and differentiation are done correctly, both approaches result in

166 the same Jacobian, and hence, the convergence behaviour is exactly the same (see Fig.
167 2 and Supporting Material Figure S8).

168 At each global iteration solving for the correction vector requires the application
169 of the inverse of the Jacobian matrix to the current residual vector. To this end, we use
170 a sparse direct factorisation based on UMFPACK [Bates, 2007] in the M2Di examples.
171 For the MDOODZ applications, we rely on a direct-iterative scheme involving Powell and
172 Hestenes iterations, Generalised Conjugate Residual iterations and pre-conditioning us-
173 ing Cholesky factorisation of the symmetrised Jacobian [e.g. Räss *et al.*, 2017]. These
174 techniques require the explicit assembly of the Jacobian matrix and it is thus necessary
175 to compute the partial derivatives of the momentum and continuity equations with re-
176 spect to the velocities.

177 3.4 The choice of parameters for viscoplastic regularisation

The viscoplastic rheological model relies on the inclusion of a rate-dependent vis-
cous component in the yield function. The latter causes an extra stress, typically coined
the overstress $\bar{\sigma}$, compared to rate-independent models. The magnitude of the overstress
depends on the strain rate can be expressed as:

$$\bar{\sigma} = \eta^{\text{vp}} \dot{\lambda} = \eta_0^{\text{vp}} \dot{\lambda}^{\frac{1}{n^{\text{vp}}}}. \quad (32)$$

For a linear viscoplastic model, the overstress depends linearly on the rate of plas-
tic multiplier. Large variations of strain rate result in proportional variations of the over-
stress. This effect can be mitigated by using power-law models, which reduces the mag-
nitude of the overstress. For example, Fig. (1A) shows variations of the overstress for
different values of the stress exponent, n^{vp} . A linear model predicts a variation of over-
stress of an order of magnitude for a variation of the strain rate of an order of mag-
nitude. By contrast, a power law model using $n^{\text{vp}} = 2.0$ predicts only half of this over-
stress. In the following model we compute the reference viscosity factor ($\text{Pa} \cdot \text{s}^{n^{\text{vp}}}$) by
defining a reference overstress $\bar{\sigma}^{\text{ref}}$ for a given reference value of strain rate ($\dot{\lambda} = \dot{\epsilon}^{\text{ref}}$).
For example, we set $\bar{\sigma}^{\text{ref}} = 1 \text{ MPa}$ for $\dot{\epsilon}^{\text{ref}} = 1 \times 10^{-15} \text{ s}^{-1}$ as in *Duretz et al.* [2020].
The reference viscosity factor can thus be written as a function of the reference overstress,
the reference strain rate and the stress exponent:

$$\eta_0^{\text{vp}} = \bar{\sigma}^{\text{ref}} \dot{\epsilon}^{\text{ref} - \frac{1}{n^{\text{vp}}}}. \quad (33)$$

178 In the linear case ($n^{\text{VP}} = 1.0$), such as in *Duretz et al.* [2020], the viscosity coefficient
179 would be equal to 10^{21} Pa-s for the above stated reference overstress and strain rate. In
180 practice, we set the reference strain rate equal to that of the bulk strain rate applied to
181 the model boundaries. Variations of the reference overstress (or viscoplastic viscosity)
182 will influence the width of shear bands [*Duretz et al.*, 2019] and thus will have an im-
183 pact the evolution of geodynamic models. We provide some examples in the following
184 sections as well as in the Supporting Material (Figure S3).

185 4 Results

186 4.1 Comparison with previously published results

187 We first check whether the above model can successfully capture visco-elasto-(visco)plastic
188 shear banding. To this end, we compare the results of 3 shear banding simulations to
189 previously published results. The latter simulations were obtained using a different model
190 formulation (displacement-based) and linearisation technique (consistent tangent lineari-
191 sation), but with the same spatial and temporal discretisation (staggered grid finite dif-
192 ferences, backward Euler). The first simulation uses an elasto-plastic (E-P) rheology and
193 corresponds to *Test 1* of *Duretz et al.* [2018] (Fig. 2A). The second test accounts for de-
194 viatoric viscous creep (V-E-P rheology) and corresponds to *Test 4* of *Duretz et al.* [2018]
195 (Fig 2B). The third test accounts for an elasto-viscoplastic rheology (E-VP) and was pre-
196 sented in *Duretz et al.* [2019] as *Model 1* (Fig 2C). All tests include bulk elasticity and
197 plastic dilation. For each test, we report the evolution of the minimum, the maximum
198 and the mean value of the second deviatoric stress invariant in the domain. In all cases
199 excellent agreement between results obtained with the different formulations is obtained.
200 We conclude that models based on the effective viscosity approach with a correct dif-
201 ferentiation and a velocity-pressure formulation capture visco-elasto-(visco)plastic shear
202 banding as accurately as models based on the displacement-based formulations and us-
203 ing a consistent tangent linearisation. Moreover we have also derived a finite-step con-
204 sistent tangent linearisation suitable for the velocity-pressure formulation (Appendix A).
205 We could show that this approach delivers similar non-linear convergence than the New-
206 ton linearisation of the effective viscosity approach (see Supporting Material, Figures S7
207 and S8).

4.2 Crustal shear banding

We now apply the rheological model to model shear banding at the crustal scale (Fig. 3 to 6). The design of the experiments is similar to that in *Duretz et al.* [2020]. The model domain has a width of 100 km and a height of 30 km, and accounts for gravity acceleration. The crust is represented by a V-E-VP rheology and accounts for the West-erly granite flow parameters, constant shear and bulk modulus, and a viscoplastic pressure-dependent flow rule. The temperature varies linearly between 20°C and 466°C from top to bottom. The only difference with *Duretz et al.* [2020] is that we consider the top boundary to be a free surface. The simulations have been carried out in compression with a constant rate of 10^{-15} s^{-1} . We have assumed small strains and thus neither advection, nor rotation are taken into account. The models were run with a resolution of 404×124 cells up to a final time of 1.2811 My (101 time steps of $4 \times 10^{11} \text{ s}$). Non-linear iterations were performed until the norm of momentum residuals dropped below 5×10^{-12} . The results can be reproduced using the corresponding M2Di MATLAB routines.

4.2.1 Bulk elastic deformation and plastic dilation

The assumption of incompressible deformations is often made in geodynamic modelling. We now investigate the role of elastic compressibility as well as that of plastic dilation on the patterns of crustal scale shear banding. Fig. 3A shows shear banding patterns obtained with an incompressible V-E-VP formulation [*Duretz et al.*, 2020]. Including elastic bulk deformation (bulk modulus $K = 5 \times 10^{10} \text{ Pa}$) significantly changes the shear banding pattern. The strain rates tend to concentrate within the main shear band and secondary shear bands tend to disappear. The effect of shear banding on the pressure field is still noticeable, and strong pressure gradients across shear bands are preserved (Fig. 3B). The inclusion of a constant plastic dilatancy angle ($\psi = 10^\circ$) smears out most of the shear bands observed in the incompressible counterpart. The main shear bands are broader, hence both the intensity of deformation and pressure variations are attenuated (Fig. 3C).

The effect of varying the bulk modulus while keeping the dilatancy angle constant is shown in Fig. 4. Panels A to C show a variation of K within the range $10^{10} - 10^{11} \text{ Pa}$. The progressive increase of K promotes strain localisation and shows that plastic strain localisation is favoured as elastic incompressibility is approached. On the other

239 hand, models with ψ ranging between 2.5 and 7.5 are shown in panels D to F. The re-
240 sults further confirm that plastic strain localisation is more likely to occur for low val-
241 ues of the dilatancy angle. The magnitude of the deviatoric stress (max. 500 MPa) and
242 the depth of the brittle-ductile transition (-17 km) are not much affected by variations
243 of the bulk modulus and the dilatancy angle. We conclude that shear banding is pro-
244 moted by either elastic or plastic incompressibility. Strain localisation is the most intense
245 in the limiting case for both elastic and plastic incompressibility (Fig. 3A).

246 In general we observe that simulations in which the effects of bulk elasticity and
247 plastic dilation are included exhibit a better global non-linear convergence than incom-
248 pressible simulations. This is also linked to the fact that both elastic and plastic incom-
249 pressibility promote shear banding, which renders the iteration procedure more challeng-
250 ing (see Supporting Information Figure S11).

251 **4.2.2 Power-law viscoplasticity**

252 We have also explored the impact of different values of the power-law viscoplas-
253 ticity exponent on the patterns of shear banding. All models are designed such that the
254 reference overstress is 1 MPa. In the reference model ($K = 5 \times 10^{10}$ Pa, $\psi = 10^\circ$), the
255 shear bands are wide and the overstress rises above 10 MPa (Fig. 3B, Fig. 5A). This is
256 because the strain rate in the shear band is one order larger in magnitude than the back-
257 ground value and the viscoplastic model is linear ($n^{\text{VP}} = 1.0$). Setting the power-law
258 exponent to 1.5 locally reduces the overstress to maximum values of about 10 MPa and
259 amplifies strain localisation (Fig. 5B). Increasing the exponent to 2.0 further decreases
260 the width of the shear bands and reduces the overstress to maximum values of about 5
261 MPa (Fig. 5C). By increasing the stress exponent, models tend towards the rate-independent
262 plastic limit. Moderate values of the stress exponent (≥ 3.0) are hence recommended to
263 benefit from the advantages of viscoplasticity (global non-linear convergence, regular-
264 isation) while keeping moderate values of overstress.

265 **4.2.3 Strain softening**

266 Material strain softening is often used to trigger strain localisation in the frictional
267 domain [Huismans and Beaumont, 2003; Naliboff *et al.*, 2017, e.g.]. We have tested the
268 effect of softening on the cohesion, the friction and the dilatancy angle for the V-E-VP

269 class of models. Implicit and explicit implementations were tested (Fig. 6). In the im-
270 plicit approach plastic variables are updated during the non-linear global iterations. In
271 the explicit approach, plastic variables are updated once per time step, i.e. after the global
272 non-linear iterations. Models with cohesion strain softening exhibit the largest sensitiv-
273 ity with regard to the type of implementation, Fig. 6A,B. The model with implicit strain
274 softening features numerous secondary shear bands, Fig. 6B. In both cases the minimum
275 cohesion (0 MPa) is reached within the shear bands. For friction softening, we have only
276 considered a small decrease of the friction angle (5°). This ensures a stable implicit time
277 integration at each time step without requiring any form of adaptative time stepping.
278 Both implementations deliver very similar shear banding patterns with most reduction
279 of the friction angle occurring in the shear bands, Fig. 6C,D. For dilation softening, we
280 also allowed for a 5° reduction of the dilatancy angle. Whatever the implementation, model
281 results are virtually similar (Fig. 6E,F) and very much resemble the reference model, Fig.
282 3B. A large part of the plastic region is not affected by shear localisation. Reduction of
283 the dilatancy angle is also observed where shear bands do not develop. In all cases we
284 observe that 'thick' shear bands (i.e. several elements/cells wide) develop. This indicates
285 that the regularising properties are preserved when strain softening is considered.

286 5 Application to lithosphere dynamics simulations

287 In order to demonstrate the practical use of the V-E-VP model, we have run long-
288 term lithospheric deformation simulations (Fig. 7 to 10). The V-E-VP model was im-
289 plemented in the thermo-mechanical code MDoodz that can handle tectonics deforma-
290 tions [Duretz *et al.*, 2016a; Kiss *et al.*, 2019; Poh *et al.*, 2020; Candiotti *et al.*, 2020] and
291 accounts for composite rheological models [Yamato *et al.*, 2019; Bessat *et al.*, 2020] and
292 a true free surface [Duretz *et al.*, 2016b]. The model domain has a 300 km width and 100
293 km height. The crust consists of three layers (15 km thick upper crust, 10 km thick mid-
294 dle crust and 10 km thick lower crust). To seed the deformation, a weak elliptical inclu-
295 sion is located within the lower crust. Its orientation will trigger either symmetric or asym-
296 metric deformation patterns (10 km long axis, 2 km short axis, with a 0° dip for the sym-
297 metric case and a 30° dip to the left for the asymmetric case). The underlying mantle
298 is represented by a single material phase. For the mechanical problem, we consider con-
299 stant normal velocities at the left (± 0.5 cm/y), right (± 0.5 cm/y) and bottom (± 0.31
300 cm/y) faces of the model, and a free surface at the top. The transient thermal problem

301 was initialised with an equilibrated temperature field (shear and adiabatic heating are
302 neglected in the initialisation step) and accounts for zero fluxes at the left and right sides,
303 and constant temperatures at the top (0° C) and at the bottom (1450° C). We have used
304 a linear viscoplastic model ($n^{\text{VP}} = 1.0$). Strain softening was applied for crustal ma-
305 terials only and was based on an explicit implementation (see the previous section). We
306 have used an explicit marker advection scheme (4th order Runger-Kutta in space) with
307 a variable time step (Courant number: 0.25). Governing equations and material prop-
308 erties are given in Appendix C and Table C.1.

309 **5.1 Evolution of reference models**

310 **5.1.1 The symmetric case**

311 In order to test the robustness of our algorithm, we first consider a case where both
312 initial and boundary conditions are symmetric. The viscoplastic viscosity was set to $2 \times$
313 10^{20} Pa·s ($n^{\text{VP}} = 1.0$), which corresponds to a reference overstress of 2×10^5 Pa. The
314 spatial resolution was set to 187 m (1600×592 cells). To enforce symmetry, the dip of
315 the weak elliptical inclusion is set to 0. After 1 My of extension, stress is built up and
316 frictional deformation occurs in the brittle regions of the whole crust (Fig. 7A). After
317 4.5 My, a symmetric neck has developed and the crust has thinned by a factor 3 in the
318 center of the model. With exhumation, the mantle cools down which leads to embrit-
319 tlement (Fig. 7B,C). At 7.6 My, the crust is hyperthinned and is about to break-up and
320 locally reaches 1/8 of its original thickness (Fig. 7C). In compression, the yield stress
321 is higher and more time is needed for fully building-up stresses. After 1 My, only the up-
322 per crust is affected by frictional plastic deformation (Fig. 7D). At 4.5 My, the lithosphere
323 has started to buckle and frictional deformation occurs throughout the entire crust (Fig.
324 7E). After 7.6 My, the crust was locally thickened by almost a factor 2. It is striking that,
325 either in compression or extension, all modelled structures are strictly symmetric despite
326 the use of frictional strain softening.

327 **5.1.2 The asymmetric case**

328 Now we introduce asymmetry by initially tilting the weak elliptical inclusion. Again,
329 models were run under either extension or compression for a total duration of 7.6 My.
330 The viscoplasticity parameters and model resolution were similar to that of the previ-

331 ous example (187 m). After 1.9 My of extension, a graben has developed at the centre
332 of the domain and is laterally limited by two conjugate crustal-scale shear bands. The
333 500°C isotherm is located at an average depth of 30 km. Asymmetry is inherited from
334 the initial tilt of the elliptical inclusion. As a result, the right shoulder rises higher than
335 the left one (Fig. 8A). After 4.4 My of stretching, the crust has necked down to a thick-
336 ness of about 10 km and the 500°C isotherm now coincides with the base of the thinned
337 crust. The isocontour of strain reveals accumulated strain in the necked crust and at the
338 base of the lithosphere (Fig. 8B). After 7.6 My, the crust reaches the stage of break-up
339 and two margins can be identified. An asymmetry can be observed as the right margin
340 is almost twice as wide as the left margin. The strain isocontour indicates that both plates
341 are weakly deformed above 90 km depth (Fig. 8C). After 1.9 My of compression, a crustal
342 scale pop-up structure develops in the centre of the domain. As in extension, asymme-
343 try is inherited from the orientation of the perturbation. Hence, the thrust that has the
344 closest orientation to that of the weak elliptical inclusion becomes the main thrust. The
345 right part of the domain thus becomes the lower plate (Fig. 8D). After 4.4 My, the Moho
346 of the lower plate reaches 60 km depth. The left plate starts to bend downwards and the
347 plateau is slightly tilted towards the hinterland. A secondary thrust develops in the fore-
348 land. The 500°C isocontour is characterised by an upward deflection below the plateau
349 (Fig. 8E). At 7.6 My, new thrusts have developed in the foreland and the tilt of the plateau
350 reaches 10°. The 500°C isocontour has started to diffuse laterally and the lower plate
351 Moho reaches a depth of 80 km. The accumulated strain contours reveal that foreland
352 deformation was mostly accommodated by the two deeply rooted thrusts (Fig. 8F).

353 5.2 Behaviour upon mesh refinement

354 We have next studied the behaviour of the models upon mesh refinement. In ex-
355 tension, the differences are minor when considering simulations for a low spatial reso-
356 lution (200×72 , Fig. 9A), a medium resolution (400×148 , Fig. 9B), a high resolu-
357 tion (800×296 , Fig. 9C), and the reference case (1600×592 , Fig. 8C). Invariably, two
358 plates that are weakly deformed above 90 km depth were individualised in response to
359 stretching. The margins exhibit similar morphologies and the stress distributions are in
360 broad agreement. The maximum stress values are located in the frictional plastic regions
361 of the lower crust of both plates (250-300 MPa).

362 In compression, the sensitivity of models to the spatial resolution is bigger. In par-
363 ticular, the low resolution model (Fig. 9D) does not capture some important character-
364 istics of the reference simulation (Fig. 8F). Instead, the upper plate is bent upwards, the
365 plateau is not developed and the upper plate deformation is limited. These differences
366 progressively diminish with increasing resolution. For example, the upper plate depicts
367 a slight downward bending for the medium resolution (Fig. 9E). For the high resolution,
368 the amplitude of bending reaches that of the reference model (Fig. 9F). The shape of
369 the 500°C isotherm also converges with increasing resolution. A stress of 500 MPa (and
370 above) is reached in the frictional portions on the middle crust, lower crust and the man-
371 tle lithosphere. The morphology of the frontal part of the wedge is the most critical re-
372 gion with regard to mesh convergence. Strain isocontours indicate that plateau growth
373 was mainly accommodated by two deeply rooted thrusts in both high and reference res-
374 olution models. However, there are still important differences in terms of the number of
375 secondary thrusts, which locally affect the topography.

376 5.3 Effect of the reference overstress

377 In V-E-VP models the rate-dependence of the plastic model is controlled by the
378 viscoplastic viscosity. The subsequent overstress is thus expected to influence the results
379 of numerical simulations. For example, setting a too small value of reference overstress
380 (or too small η^{VP}) is equivalent to running a model in the rate independent limit (V-E-
381 P). In this limit, the benefits of viscoplastic regularisation vanish and models will thus
382 likely fail at satisfying a global force equilibrium. For a too large value of overstress, the
383 computed stress will be well above the value predicted with the rate-independent Drucker-
384 Prager. This will impede shear banding and strain localisation in the frictional domain.
385 We have examined the impact of this parameter on the lithospheric extension and on
386 the compression models for the medium resolution (800×296 cells). For the reasons men-
387 tioned above, we have varied η^{VP} within a narrow range ($10^{20} - 10^{21}$ Pa·s). In extension,
388 the magnitude of the viscoplastic viscosity has little influence on the morphology of the
389 margins and the timing of break-up (Fig. 10,A,B,C). Small values of η^{VP} promote strain
390 localisation in the frictional plastic domain, enhances asymmetry as well as topographic
391 gradients. However, the overall stress distribution and the stress levels are not affected
392 strongly. The value of η^{VP} also has an influence on lithosphere compression models, in
393 particular on the distribution of strain around the plateau. For example, we observe that

394 deeply rooted foreland-dipping thrusts are better captured for $\eta^{\text{VP}} \leq 4 \times 10^{20}$ Pa·s (Fig.
395 10D,E,F). Interestingly, we notice that variations of η^{VP} affect the bending of the upper
396 plate in a similar way as the numerical resolution (see section above). However, the over-
397 all stress distribution and the stress level, and also the topography, are influenced rather
398 weakly by the value of η^{VP} .

399 6 Discussion

400 Geodynamic models generally rely on continuum mechanics and can, so far, at best
401 achieve resolutions of 100 m scale [Naliboff *et al.*, 2017; Petri *et al.*, 2019]. Such resolu-
402 tions are far larger than the actual width of fault zones that develop in the brittle part
403 of the lithosphere [Shipton *et al.*, 2006]. The latter could also be modelled using discon-
404 tinuous representations [e.g. de la Puente *et al.*, 2009]. However this approach does not
405 seem suitable for large deformation tectonic models that should capture the self-consistent
406 emergence and activation of frictional plastic shear bands. Moreover, such models will
407 not be able to resolve fault zones dimensions, in particular their width, when a sufficiently
408 high resolution will have come within reach. It is therefore essential to continue devel-
409 oping geodynamic models based on a continuum approach that can capture frictional
410 plastic strain localisation while providing stable and accurate numerical solutions.

411 To this end we have used viscoplastic regularisation, which offers a simple and di-
412 rect way to obtain convergence during the non-linear iterative process and vastly reduces
413 mesh dependence. However, it also clear that viscoplastic regularisation is a time reg-
414 ularisation and not a spatial regularisation approach [Wang, 2019], in contrast with non-
415 local plastic models [Bažant and Lin, 1988], gradient-based models [de Borst and Mühlhaus,
416 1992] or models based on Cosserat medium [Mühlhaus and Vardoulakis, 1987]. Viscoplas-
417 ticity is thus not expected to resolve all mesh convergence problems. Future models should
418 also consider spatial regularisation approach, which despite their algorithmic and com-
419 putational cost become more and more affordable in the context of fully iterative solv-
420 ing strategies (e.g. pseudo-transient schemes in Räss *et al.* [2019]). In particular, it will
421 be interesting to consider models that include both temporal and spatial regularisation
422 in the context of geodynamic modelling [Wang, 2019]. In general care has to be taken
423 when selecting the value of the viscoplastic viscosity which controls the overstress, since
424 the evolution of lithospheric models can be affected by this parameter (see Supporting
425 Information Figure S3). For the parameter range considered, these variations fortunately

426 appeared less important than variations caused by mesh refinement. In fact, even with
427 regularisation and for converged flow and pressure fields, secondary shear bands still de-
428 velop as the resolution increases. This is likely a limitation of viscoplastic regularisation.

429 With viscoplastic regularisation, the main parameter that allows for obtaining non-
430 linear convergence while preserving shear localisation is the reference overstress. We rec-
431 ommend to select values in the order of, or below, 1 MPa. Local variations of the over-
432 stress may become important in cases where strong localisation occurs. The above pre-
433 sented power-law model appears as a suitable way to limit this drawback. We have also
434 noticed that, in some cases, checkerboard-style shear banding patterns can arise (see Sup-
435 porting Information Figure S6). This was observed for very low values of the overstress
436 ($\bar{\sigma} < 10^5$ Pa), in the limit where viscoplastic regularisation becomes inefficient and where
437 shear band can not be captured by the model resolution. This effect occurred in com-
438 pressible models and is unrelated to the stability of the velocity-pressure discretisation.
439 Further work is needed to investigate what controls this behaviour and to further im-
440 prove the use of V-E-VP geodynamic models.

441 The effects of elastic bulk deformation can be included in geodynamic models [e.g
442 *Poliakov et al.*, 1993; *Gerbault et al.*, 1998; *Popov and Sobolev*, 2008; *Choi and Petersen*,
443 2015]. Herein we demonstrate that variations in the bulk modulus (K) or in Poisson ra-
444 tio (ν) within a realistic range can have a strong impact on the patterns and the inten-
445 sity of shear banding. More intense strain localisation was obtained for high values of
446 the elastic bulk modulus, in particular near the limit of elastic incompressibility. This
447 can be explained by the fact that an increase of bulk modulus decreases effective elasto-
448 plastic hardening, which can promote strain localisation. For a larger bulk modulus, the
449 strength contrast between shear bands and their less deformed host is increased, which
450 further contribution to localisation. A comparison of models involving or not the effect
451 of elastic compressibility was presented in *Choi and Petersen* [2015]. In comparison to
452 geoFLAC models ($\nu = 0.25$), those based on the code 2DPIC (elastically incompress-
453 ible, $\nu = 0.5$) displayed larger accumulated strain inside shear bands at high resolution
454 (their Fig. 2). This would also indicate that elastic incompressibility further contribute
455 to strain localisation. This is however to be taken with care since the models presented
456 in this study are based on fundamentally different algorithms. Models with a positive
457 dilatancy angle exhibit a more diffuse localisation pattern than plastic incompressible
458 models in agreement with the results of *Choi and Petersen* [2015]. The use of dilation

459 strain softening as suggested by *Choi and Petersen* [2015], for example, seems a reason-
460 able choice to capture the initial volume changes due to frictional plastic deformation,
461 while still promoting the localisation of strain. It is also in agreement with laboratory
462 measurements [*Zhao and Cai*, 2010, e.g.].

463 We have also demonstrated that for frictional plasticity a velocity-based Effective
464 Viscosity Approach (EVA) can deliver the same results as algorithms more common in
465 the engineering literature, which are based on a displacement formulation and which lin-
466 earise a return-mapping scheme to compute the stresses in order to derive a consistently
467 linearised tangent operator (see also *Lemiale et al.* [2008]). This is of course subject to
468 the use of the same spatio-temporal discretisation. Moreover, it is also conditional on
469 a proper derivation of the visco-elastic-viscoplastic viscosity η^{vep} and a proper differen-
470 tiation of this quantity. This includes that (i) elastic deformations are taken into account,
471 (ii) plastic incompressibility is not assumed a priori, and (iii) a proper loading-unloading
472 criterion is utilised. It is noted that, at variance with these conditions, the original for-
473 mulation [*Willett*, 1992] assumes the absence of elastic deformations, plastic incompress-
474 ibility, and continued plastic loading. For the original formulation, identical results are
475 therefore obtained only for a small subset of constitutive models.

476 The cause for the generation of non-symmetric structures in geodynamic models
477 is typically attributed to plastic strain softening [e.g. *Huismans and Beaumont*, 2003].
478 In this study, strictly symmetric deformation structures arise from models with symmet-
479 ric initial and boundary conditions, despite the use plastic strain softening (Fig. 7). In
480 fact, it was necessary to introduce asymmetry in either the initial or boundary condi-
481 tions to model the formation of asymmetric rifts and mountain belts (Fig. 8). Here, asym-
482 metry is thus induced by either inheritance or far-field kinematics. This is provided that
483 all discretisation elements (mesh, markers, stencils) are symmetric and that frictional
484 plasticity is solved to reasonable accuracy (here, machine precision) at each time incre-
485 ment. Given these conditions, we argue that monitoring the symmetry of structures can
486 demonstrate the robustness of geodynamic models that encapsulate non-linear rheolo-
487 gies upon large strain. A comparison with the models presented in *Huismans and Beau-*
488 *mont* [2007] is provided in the Supporting Information. It shows that both symmetric
489 and asymmetric deformation can be modelled using a either a linear or power-law V-E-
490 VP model (Supporting Material Figure S1 to S3) and further confirms that the occur-
491 rence of asymmetric deformation is controlled by the amplification of initial perturba-

492 tions (Supporting Material Figure S1). Our result hence confirm those of *Huismans and*
493 *Beaumont* [2007] but also clarifies the role of strain softening which acts as a catalyser
494 rather than a source of asymmetry.

495 Numerical geodynamic models often exhibit lack of global non-linear convergence
496 [*Spiegelman et al.*, 2016] and the consequences of this is debatable. Here we have used
497 viscoplastic regularisation to obtain global convergence in geodynamic models. We have
498 observed that when the internal dynamics of the model is not entirely controlled by fric-
499 tional plastic strain localisation, the evolution of models is not very sensitive to the con-
500 vergence of global equilibrium (see Supporting information Figure S9). This is a posi-
501 tive outcome which is particularly true when the plastic layer (i.e. the crust) is consid-
502 ered as a passive stress limiter or when the internal dynamics is controlled by a mesh-
503 convergent strain localisation phenomenon (e.g., shear heating in [*Schmalholz et al.*, 2014]).
504 However, the propagation of frictional plastic shear bands is affected by the convergence
505 of global equilibrium (see Supporting information Figure S9). This can become prob-
506 lematic in cases where frictional plastic shear banding is the essential ingredient of a model.
507 It is therefore important to keep on developing aspects related to frictional plasticity and
508 to continue improving the reliability and robustness of future geodynamic models.

509 **7 Conclusions**

510 We have designed and tested new aspects of a frictional viscoplastic rheology for
511 geodynamic modelling, with emphasis on viscoplastic regularisation. Elastic compress-
512 ibility, power-law viscoplasticity, plastic dilatancy, non-associated plastic flow and strain
513 softening can be all combined, and incorporated in a numerical approach that is com-
514 mon in geodynamic modelling, i.e. a velocity-pressure formulation with an Effective Vis-
515 cosity Approach. Moreover, since we accounted for elastic deformation and compress-
516 ibility, the approach can deliver results which are as accurate as those in computational
517 engineering. We have shown that elastic volumetric deformations, usually neglected in
518 geodynamic modelling, can have a noticeable impact on patterns and intensity of shear
519 banding. Power-law viscoplasticity can be used to limit the overstress inherent in vis-
520 coplastic modelling. Finally, we have shown successful applications of the visco-elastic-
521 viscoplastic (V-E-VP) model in the context of state-of-the geodynamic simulations. Hence,
522 viscoplastic regularisation may be used for practical purposes when modelling long-term
523 tectonic deformations.

8 Figures

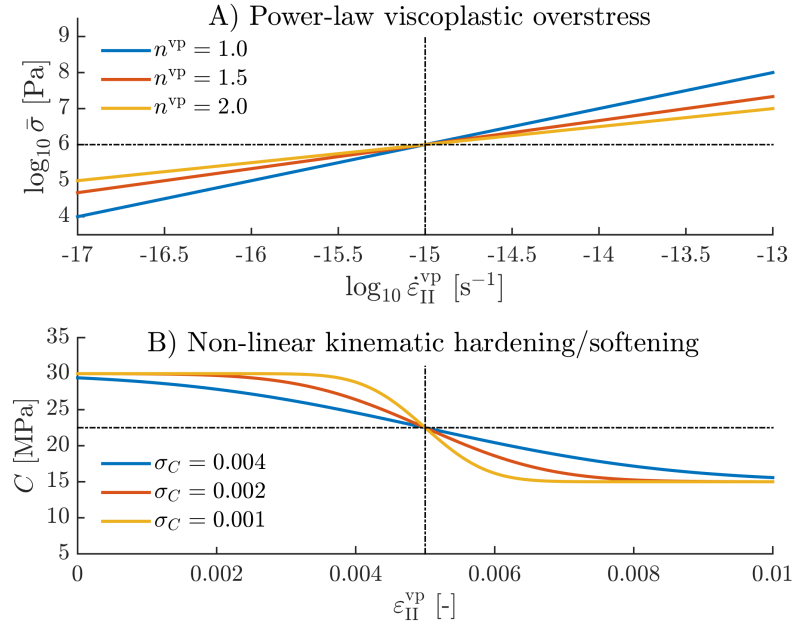


Figure 1. A) Effect of power-law viscoplasticity exponent n^{VP} on the magnitude of viscoplastic overstress $\bar{\sigma}$ for the variable viscoplastic strain rate $\dot{\epsilon}_{II}^{VP}$. B) Example of a non-linear hardening/softening law. Effect of variable standard deviation σ_c on the evolution of cohesion c with the viscoplastic strain ϵ_{II}^{VP} .

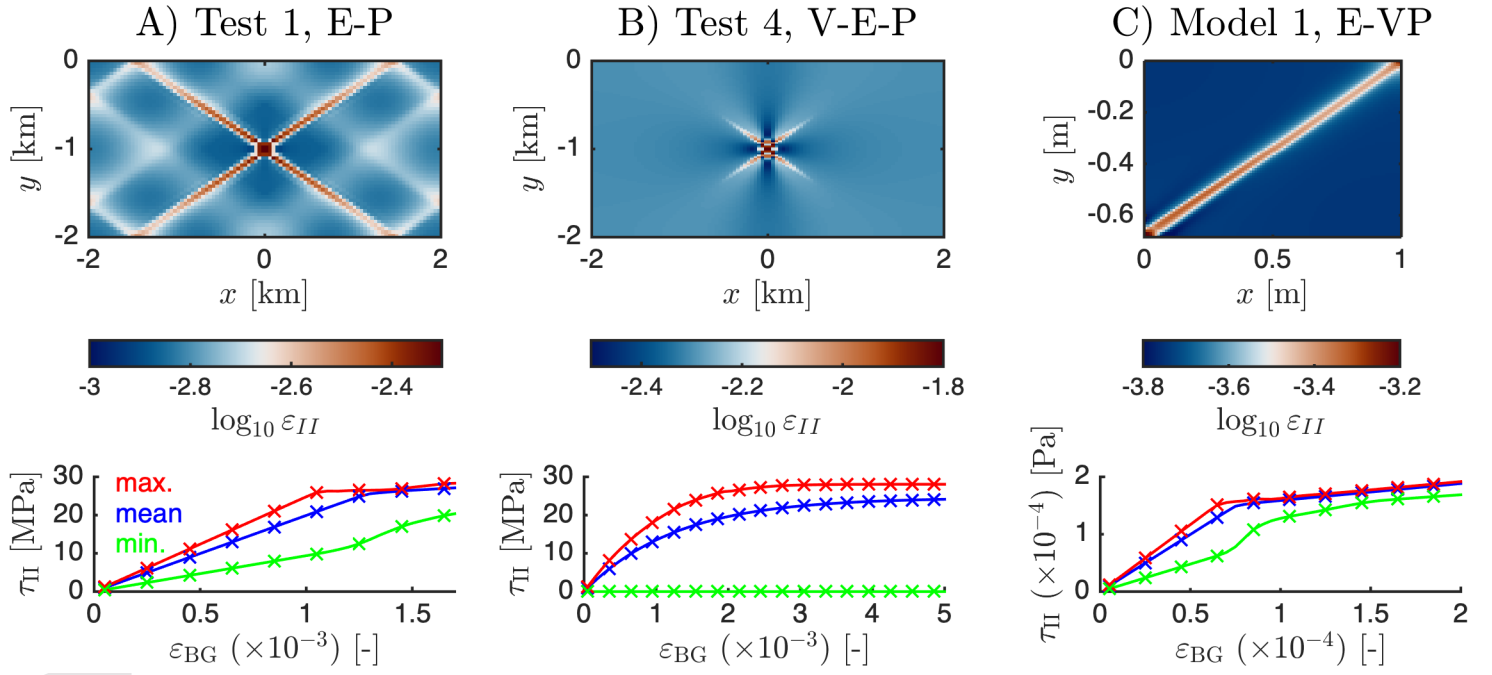


Figure 2. Simulation of shear banding using compressible elastic and plastic rheology. A) Test 1 from *Duretz et al.* [2018]: shear banding in the elasto-plastic regime. B) Test 4 from *Duretz et al.* [2018]: shear banding in the visco-elasto-plastic regime. C) Reference model from *Duretz et al.* [2019]: shear banding in the elasto-viscoplastic regime. Lower panel plots show comparison between measurements of the second deviatoric stress invariant (minimum, mean, maximum). Solid lines are from published studies [*Duretz et al.*, 2018, 2019] and are based on incremental displacement-based consistent tangent formulation. The crosses correspond to this study and are based on velocity-pressure effective viscosity formulation. The colour maps are *vik* from *Cramer* [2018].

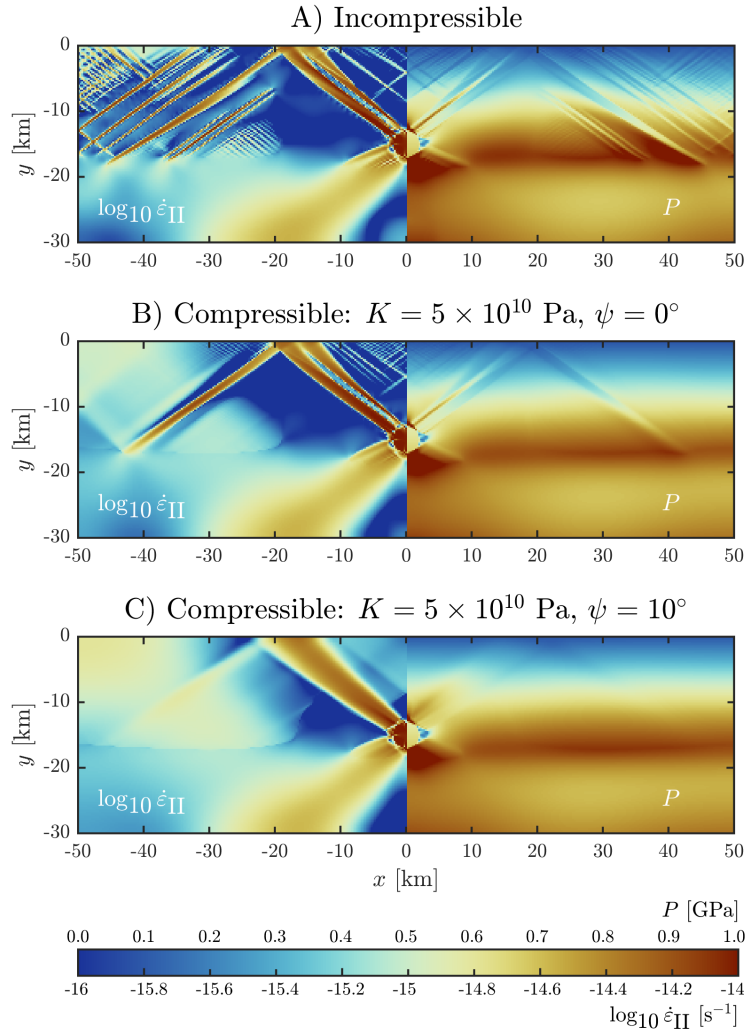


Figure 3. Role of compressibility on crustal shear banding. A) Strictly incompressible model. B) Model incorporating elastic volumetric deformations. C) Model incorporating both elastic volumetric deformation and plastic dilatancy. For each model, the second invariant of deviatoric strain rate tensor (logarithmic scale, left) and the pressure (right) are depicted. The colour map is *roma* (inverted) from *Crameri* [2018]. Initial and boundary conditions are similar to those of *Duretz et al.* [2020] excepted that we consider the top boundary to be a free surface (see text for details)

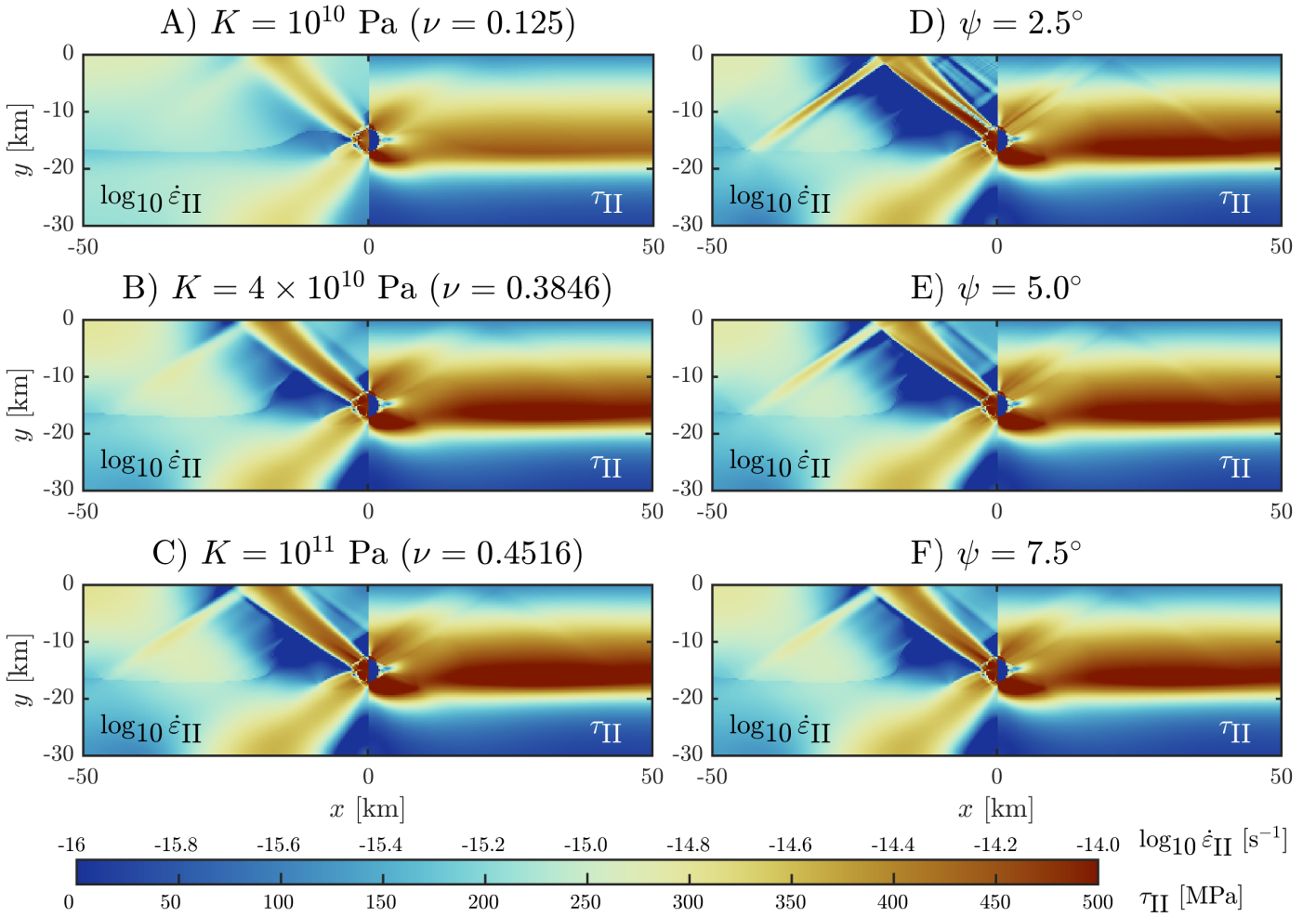


Figure 4. Effect of varying elastic compressibility and plastic dilation. A), B) and C) shows models with different elastic compressibilities. E), F) and G) correspond to runs with different dilatancy angles (constant). The colour map is *roma* (inverted) from *Cramer* [2018].

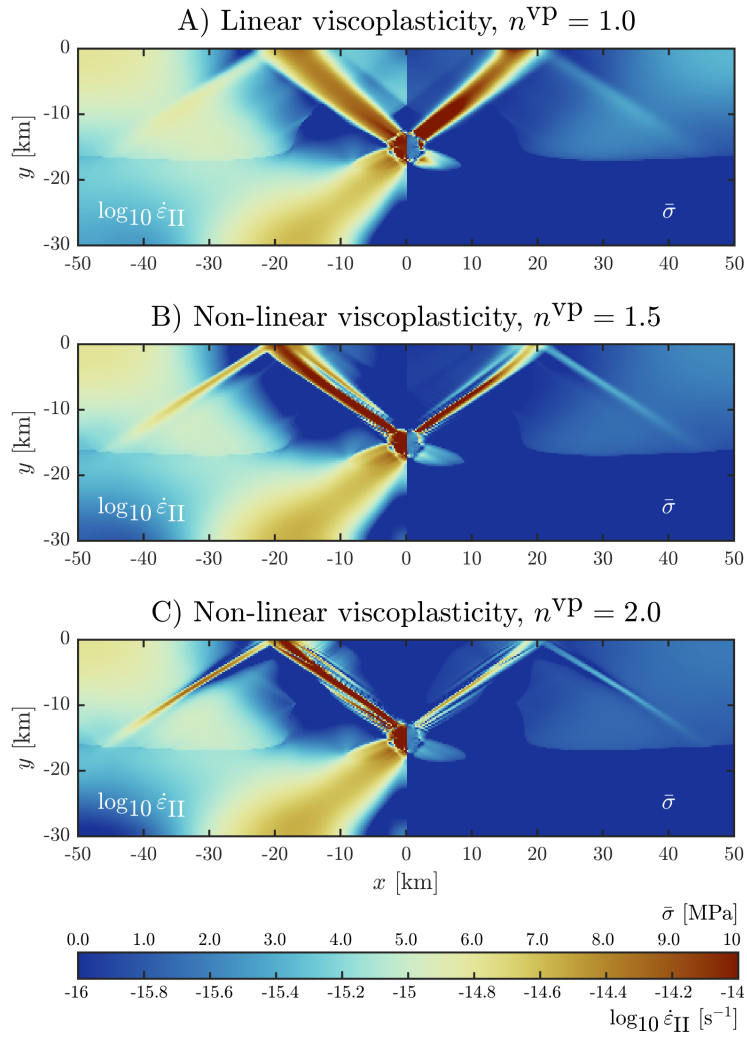


Figure 5. Effect of power-law viscoplasticity. A) Compressible model ($K = 5 \times 10^{10}$ Pa, $\psi = 10^\circ$) with linear viscoplasticity. B) Similar model as in A) but with power-law viscoplastic exponent set to 1.5. C) Similar model as in A) but with power-law viscoplastic exponent set to 2.0. For each model, both the second invariant of deviatoric strain rate tensor (logarithmic scale, left) and the viscoplastic overstress (right) are depicted. The colour map is *roma* (inverted) from *Cramer* [2018].

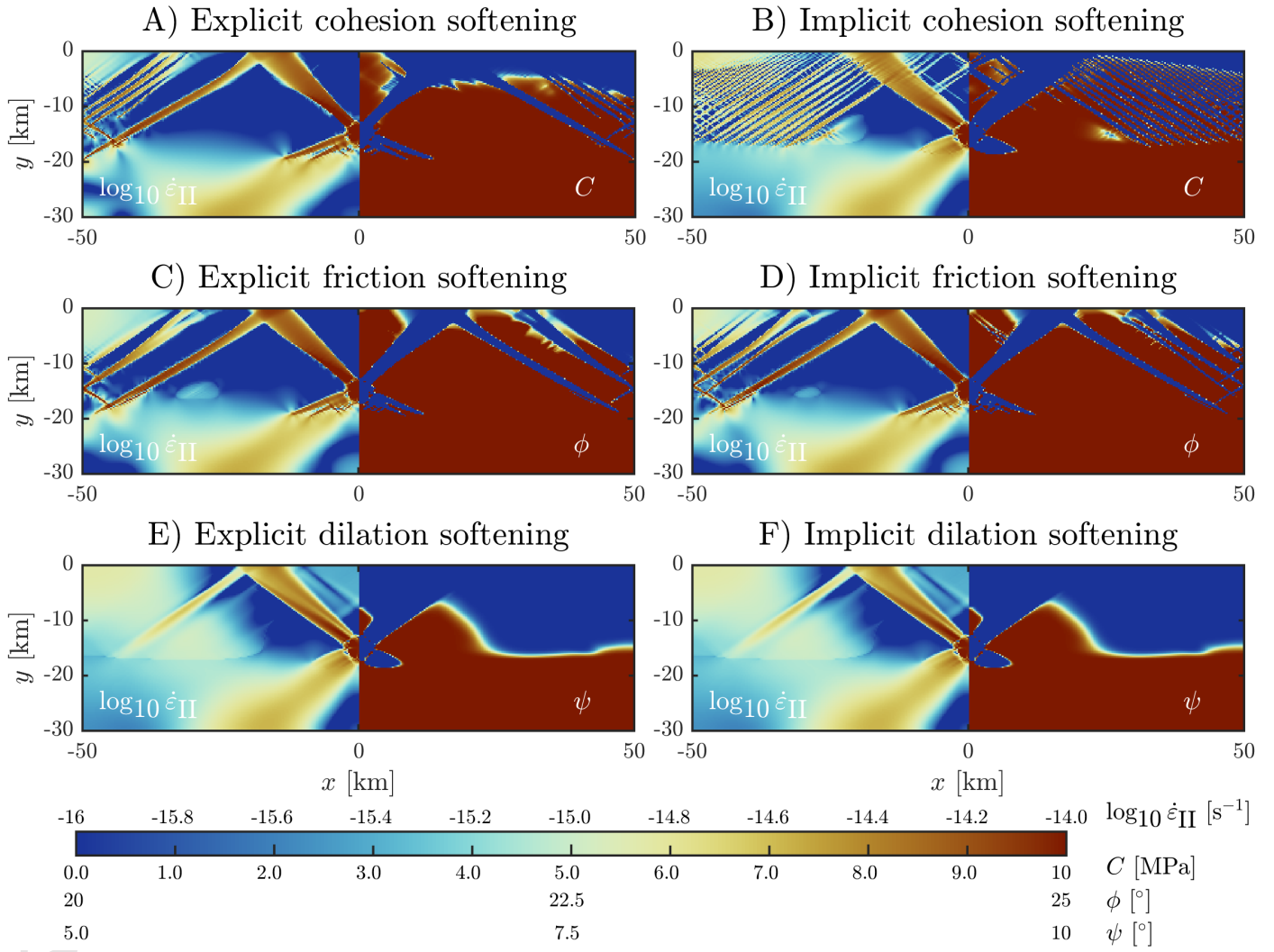


Figure 6. Effect of strain softening on crustal shear banding. The left column shows results obtained with an explicit strain softening implementation. The right column correspond to an implicit implementation of strain softening. A) and B) correspond to cohesion softening from 50 MPa to 25 MPa. C) and D) correspond to a friction softening from 25 to 20°. E) and F) correspond to a dilation softening from 10 to 5°. The colour map is *roma* (inverted) from *Crameri* [2018].

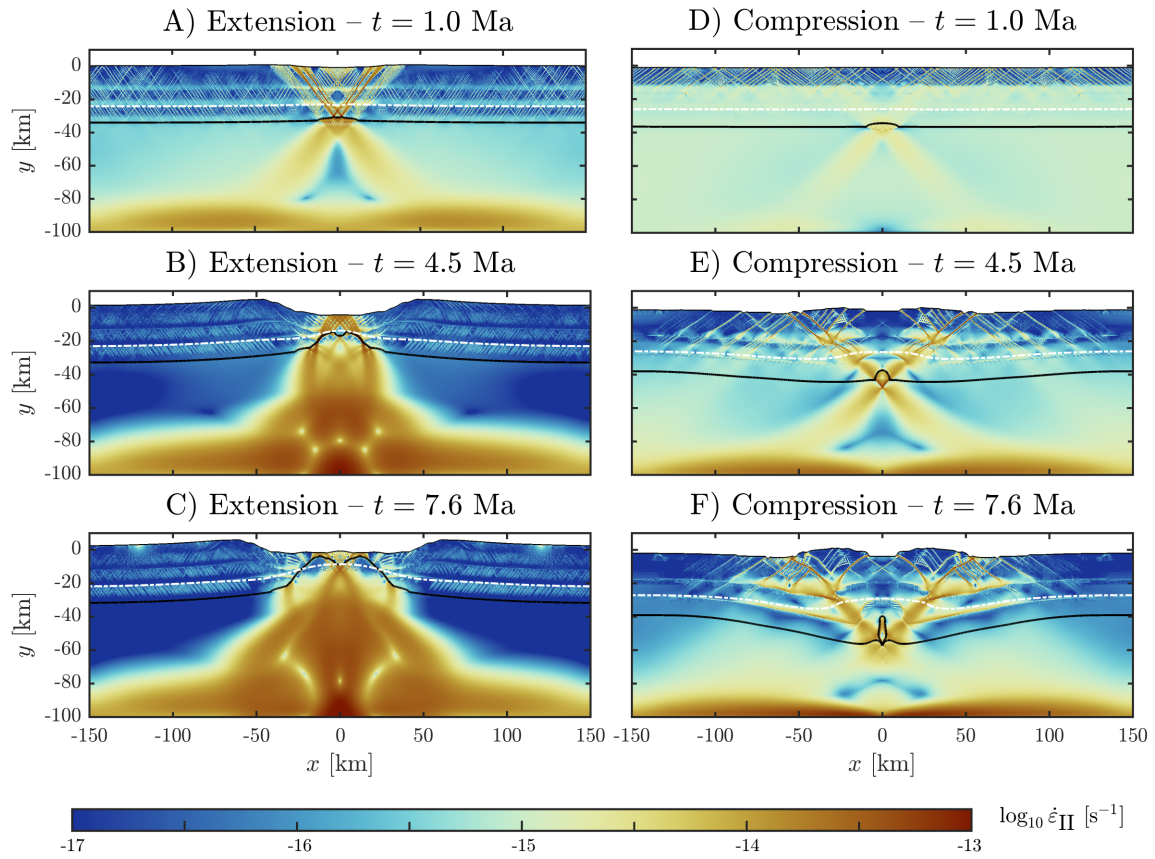


Figure 7. Time evolution of the lithospheric models in extension and compression with symmetric initial and boundary conditions. The colour map corresponds to the magnitude of the second deviatoric strain-rate invariant (logarithmic scale). Panels A, B, and C depict a model undergoing extension, panels D, E, F correspond to compression. The black lines indicate the location of the Moho, the dash-dotted line is the 500° isotherm.

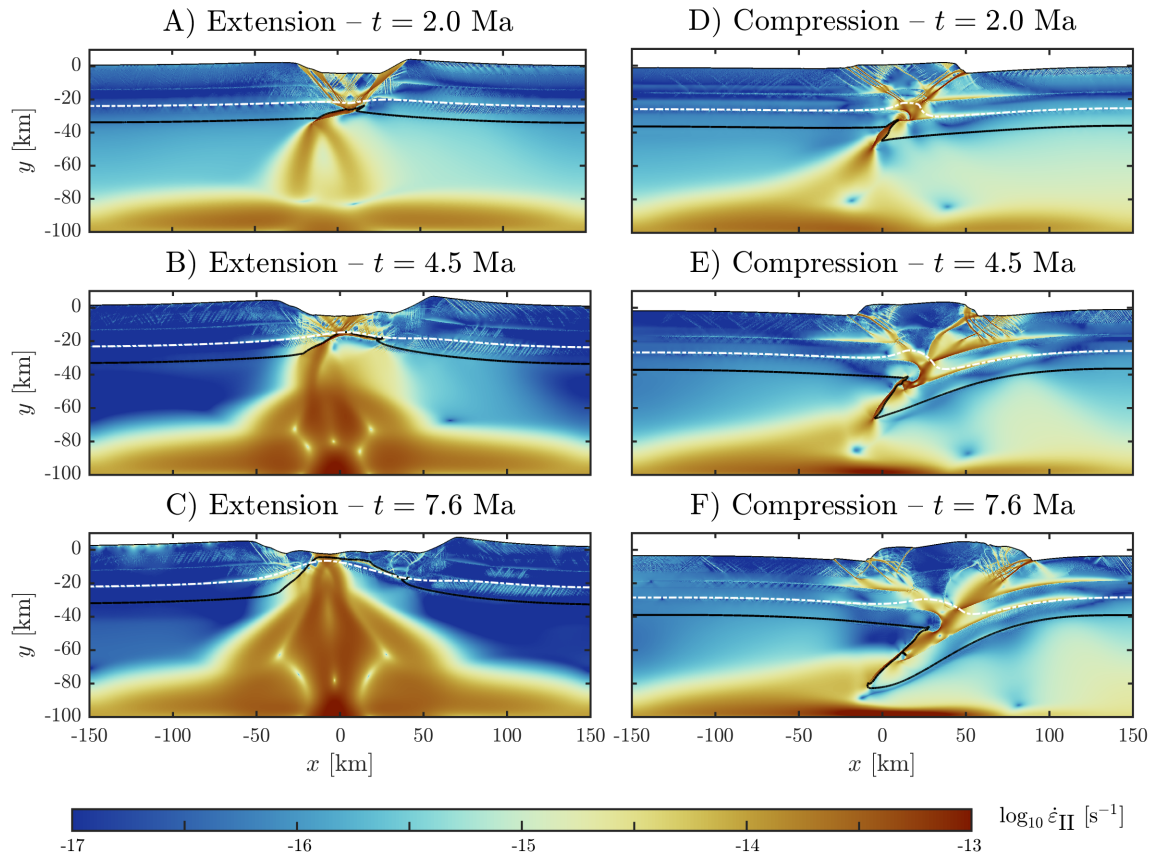


Figure 8. Time evolution of the reference lithospheric models in extension and compression with a non-symmetric initial condition. The colour map corresponds to the magnitude of the second deviatoric strain-rate invariant (logarithmic scale). Panels A, B, and C depict a model undergoing extension, panels D, E, F correspond to compression. The black lines indicate the location of the Moho, and the dash-dotted line is the 500° isotherm.

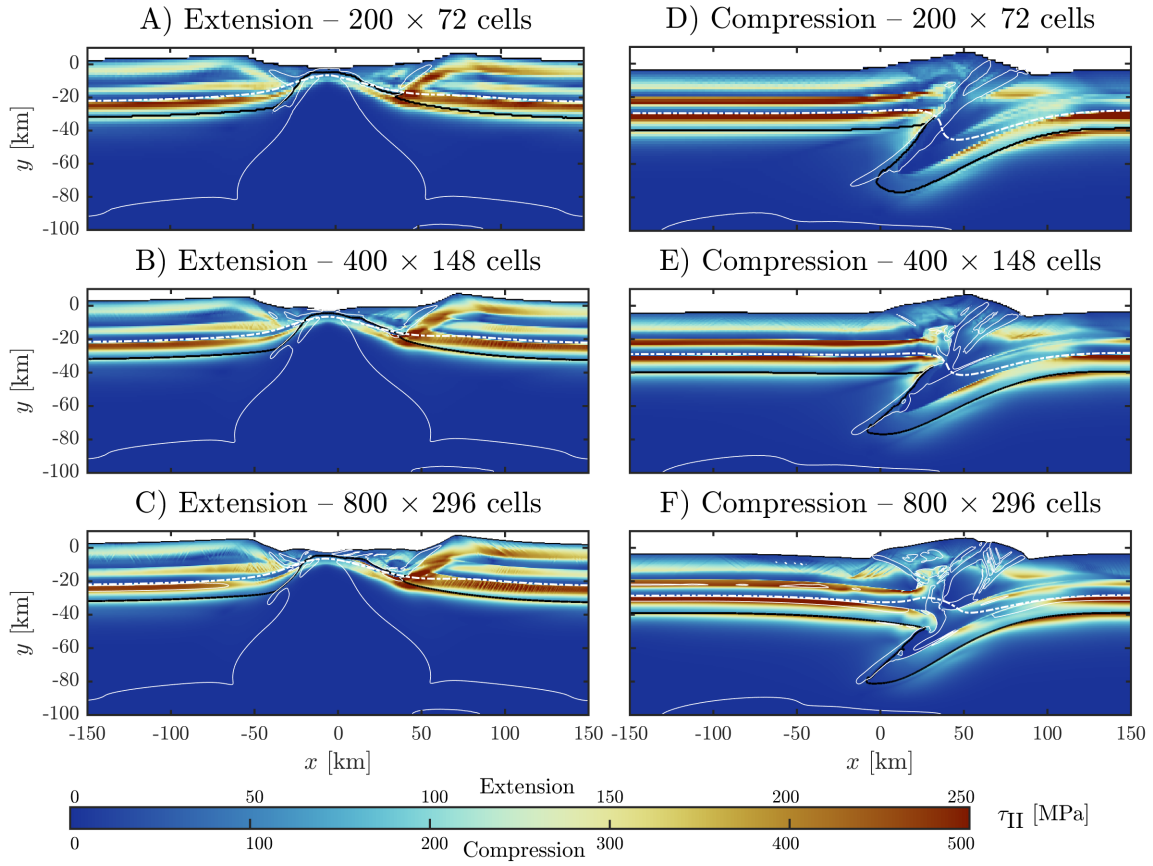


Figure 9. Sensitivity of lithospheric models to numerical resolution. Extension to the left, compression to the right. Three different resolutions are depicted (200×72 , 400×148 and 800×296 cells). The model time is about 7.6 My for all six simulations. The colormap corresponds to the second deviatoric stress invariant (τ_{II}). The black lines indicate the location of the Moho, the dash-dotted line is the 500° isotherm and the thin solid white is the 1.8 accumulated strain contour.

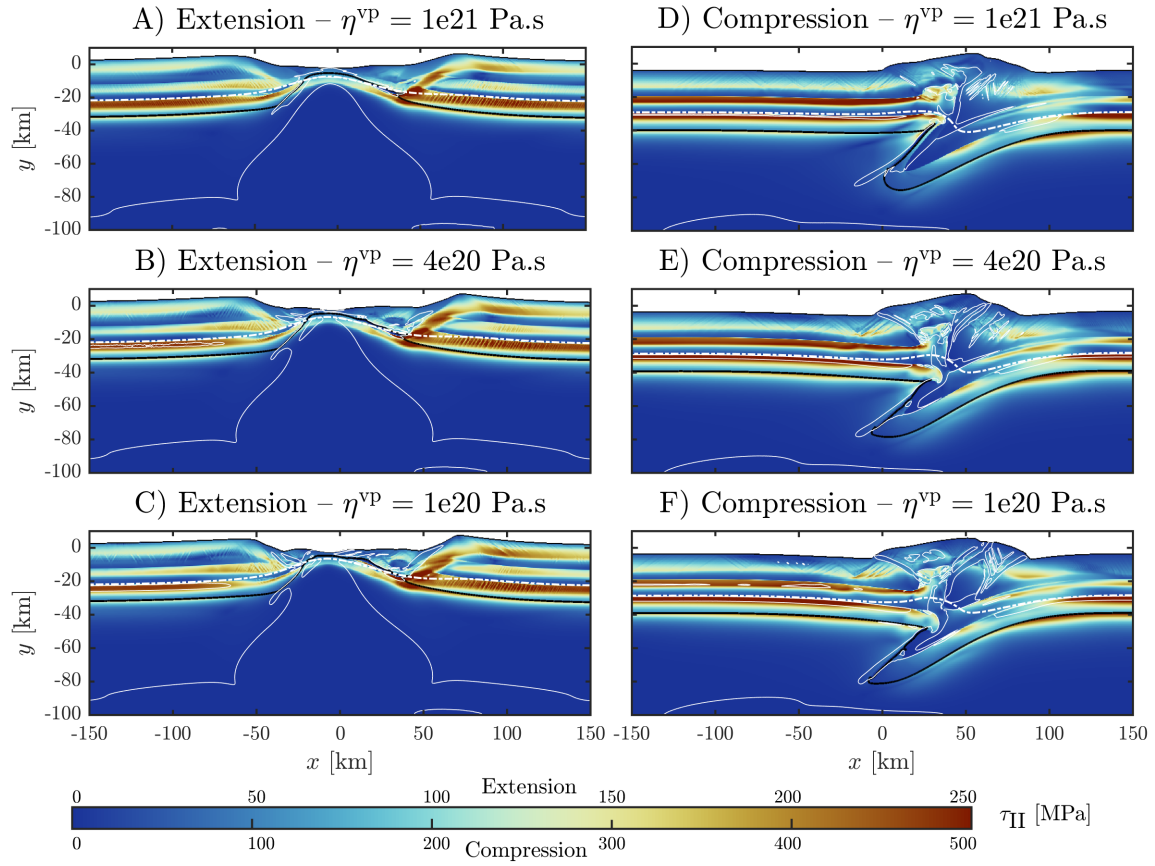


Figure 10. Influence of the viscoplastic viscosity on the evolution of lithospheric models in both extension (left panels) and compression (right panels). Three values of η^{VP} were investigated (10^{21} , 4×10^{20} and 10^{20} Pa·s). Models were run up to about 7.6 My. The colormap corresponds the second deviatoric stress invariant (τ_{II}). The black lines indicate the location of the Moho, the dash-dotted line is the 500° isotherm and the thin solid white is the 1.8 accumulated strain contour.

Table 1. Model parameters used for Fig. 2. The friction and dilatancy angles were set to 30 and 10 degrees, respectively.

Parameter	<i>Test 1</i>	<i>Test 4</i>	<i>Model 1</i>
L_x [m]	4000	4000	1
L_y [m]	2000	2000	0.685
$n_x \times n_y$	100×50	100×50	100×69
K [Pa]	2×10^{10}	2×10^{10}	2.0
G_{mat} [Pa]	10^{10}	10^{10}	1.0
G_{inc} [Pa]	2.5×10^9	10^{10}	0.25
η_{mat} [Pa.s]	-	2.5×10^{21}	-
η_{inc} [Pa.s]	-	10^{17}	-
C [Pa]	3×10^7	3×10^7	1.75×10^{-4}
η^{vp} [Pa.s]	-	-	2.5×10^2
$\dot{\epsilon}_{BG}$ [s $^{-1}$]	10^{-15}	10^{-15}	5×10^{-10}
Δt [s]	10^{10}	10^{10}	10^4

A: The tangent operator: Newton linearisation of the effective viscosity approach (EVA Newton)

The tangent operator needed for the global Newton-Raphson iterations is expressed as: $\mathbf{D} \equiv \frac{\delta \boldsymbol{\tau}}{\delta \dot{\boldsymbol{\epsilon}}}$, where $\boldsymbol{\tau} = [\tau_{xx} \ \tau_{yy} \ \tau_{zz} \ \tau_{xy} \ p^{corr}]^T$, $\dot{\boldsymbol{\epsilon}} = [\epsilon_{xx} \ \epsilon_{yy} \ \epsilon_{zz} \ \epsilon_{xy} \ v'_{k,k}]^T$ and may be explicitly written as:

$$\mathbf{D}^{eva} = \begin{bmatrix} \frac{\partial \tau_{xx}}{\partial \dot{\epsilon}'_{xx}} & \frac{\partial \tau_{xx}}{\partial \dot{\epsilon}'_{yy}} & \frac{\partial \tau_{xx}}{\partial \dot{\epsilon}'_{zz}} & \frac{\partial \tau_{xx}}{\partial \dot{\epsilon}'_{xy}} & \frac{\partial \tau_{xx}}{\partial v'_{k,k}} \\ \frac{\partial \tau_{yy}}{\partial \dot{\epsilon}'_{xx}} & \frac{\partial \tau_{yy}}{\partial \dot{\epsilon}'_{yy}} & \frac{\partial \tau_{yy}}{\partial \dot{\epsilon}'_{zz}} & \frac{\partial \tau_{yy}}{\partial \dot{\epsilon}'_{xy}} & \frac{\partial \tau_{yy}}{\partial v'_{k,k}} \\ \frac{\partial \tau_{zz}}{\partial \dot{\epsilon}'_{xx}} & \frac{\partial \tau_{zz}}{\partial \dot{\epsilon}'_{yy}} & \frac{\partial \tau_{zz}}{\partial \dot{\epsilon}'_{zz}} & \frac{\partial \tau_{zz}}{\partial \dot{\epsilon}'_{xy}} & \frac{\partial \tau_{zz}}{\partial v'_{k,k}} \\ \frac{\partial \tau_{xy}}{\partial \dot{\epsilon}'_{xx}} & \frac{\partial \tau_{xy}}{\partial \dot{\epsilon}'_{yy}} & \frac{\partial \tau_{xy}}{\partial \dot{\epsilon}'_{zz}} & \frac{\partial \tau_{xy}}{\partial \dot{\epsilon}'_{xy}} & \frac{\partial \tau_{xy}}{\partial v'_{k,k}} \\ \frac{\partial p^{corr}}{\partial \dot{\epsilon}'_{xx}} & \frac{\partial p^{corr}}{\partial \dot{\epsilon}'_{yy}} & \frac{\partial p^{corr}}{\partial \dot{\epsilon}'_{zz}} & \frac{\partial p^{corr}}{\partial \dot{\epsilon}'_{xy}} & \frac{\partial p^{corr}}{\partial v'_{k,k}} \end{bmatrix} \quad (\text{A.1})$$

To construct the tangent operator matrix needed, we first reformulate the rheology as:

$$\begin{aligned} \tau_{ij} &= 2\eta \left(\dot{\epsilon}'_{ij}, p^{trial} \right) \dot{\epsilon}'_{ij} \\ p^{corr} &= p^{trial} + K \Delta t v_{k,k}^{vp} \left(\dot{\epsilon}'_{ij}, p^{trial} \right) \end{aligned} \quad (\text{A.2})$$

where $\eta = \eta^{\text{ve}}$ below the plastic yield and $\eta = \eta^{\text{vep}}$ at the yield. Then, the components of \mathbf{D}^{eva} can be computed as follows:

$$\begin{aligned}\frac{\partial \tau_{ij}}{\partial \dot{\epsilon}'_{kl}} &= 2\eta \left(\delta_{ik}\delta_{jl} + \dot{\epsilon}'_{ij} \frac{\partial \eta}{\partial \dot{\epsilon}'_{kl}} \right) \\ \frac{\partial \tau_{ij}}{\partial v_{k,k}} &= 2\dot{\epsilon}'_{ij} \frac{\partial \eta}{\partial v_{k,k}} \\ \frac{\partial p^{\text{corr}}}{\partial \dot{\epsilon}'_{kl}} &= K \Delta t \frac{\partial v_{i,i}^{\text{vp}}}{\partial \dot{\epsilon}'_{kl}} \\ \frac{\partial p^{\text{corr}}}{\partial v_{k,k}} &= -K \Delta t \left(1 - \frac{\partial v_{i,i}^{\text{vp}}}{\partial v_{k,k}} \right).\end{aligned}\tag{A.3}$$

530 In case of visco-elastic creep or incompressible plastic flow, the partial derivatives of p^{corr}
531 vanish. Since we consider p^{trial} as our global variable for the pressure, it is necessary to
532 express partial derivatives with regard to the trial pressure. The latter can be further
533 computed using the chain rule as, $\frac{\partial \tau_{ij}}{\partial p^{\text{trial}}} = \frac{\partial \tau_{ij}}{\partial v_{k,k}} \frac{\partial v_{k,k}}{\partial p^{\text{trial}}}$ and $\frac{\partial p^{\text{corr}}}{\partial p^{\text{trial}}} = \frac{\partial p^{\text{corr}}}{\partial v_{k,k}} \frac{\partial v_{k,k}}{\partial p^{\text{trial}}}$, where
534 $\frac{\partial v_{k,k}}{\partial p^{\text{trial}}} = -\frac{1}{K \Delta t}$.

For completeness we also provide the partial derivatives of η^{ve} and η^{vep} . Upon completion of the local iterations that determine the visco-elastic trial stress, the effective viscosity is expressed as:

$$\eta^{\text{ve}} = \left(\frac{1}{\eta^{\text{v}}} + \frac{1}{\eta^{\text{e}}} \right)^{-1} \text{ with } \eta^{\text{v}} = C^{\text{v}} \tau_{\text{II}}^{n^{\text{v}}-1}.\tag{A.4}$$

The partial derivatives of η^{ve} take the following form:

$$\frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{ij}} = \frac{\partial \eta^{\text{ve}}}{\partial \tau_{ij}} \frac{2\eta^{\text{ve}}}{1 - 2 \left(\frac{\partial \eta^{\text{ve}}}{\partial \tau_{xx}} \dot{\epsilon}'_{xx} + \frac{\partial \eta^{\text{ve}}}{\partial \tau_{yy}} \dot{\epsilon}'_{yy} + \frac{\partial \eta^{\text{ve}}}{\partial \tau_{zz}} \dot{\epsilon}'_{zz} + \frac{\partial \eta^{\text{ve}}}{\partial \tau_{xy}} \dot{\epsilon}'_{xy} \right)}\tag{A.5}$$

Differentiation with regard to τ_{ij} yields:

$$\frac{\partial \eta^{\text{ve}}}{\partial \tau_{ij}} = -b(n^{\text{v}} - 1) (\eta^{\text{ve}})^2 C^{\text{v}} \tau_{\text{II}}^{n^{\text{v}}-3} \tau_{ij},\tag{A.6}$$

535 with $b = 1$ for normal components and $b = 2$ for the shear components. Note that the
536 derivative of η^{ve} with regard to p^{trial} is 0 since we consider that visco-elastic creep in pressure-
537 independent in this study.

In case of plastic flow, the partial derivatives of η^{vep} and $v_{i,i}^{\text{vp}}$ need to be determined. The latter may be evaluated upon completion of the local iteration that determines the values of $\dot{\lambda}$ may be formulated as:

$$\begin{aligned}\frac{\partial \eta^{\text{vep}}}{\partial \dot{\epsilon}'_{ij}} &= -\frac{\dot{\epsilon}'_{ij} \tau_{\text{II}}}{4 (\dot{\epsilon}'_{\text{II}})^3} + \frac{1}{2\dot{\epsilon}'_{\text{II}}} \left(a \frac{\partial \dot{\lambda}}{\partial \dot{\epsilon}'_{ij}} + \dot{\lambda} \frac{\partial \eta^{\text{vp}}}{\partial \dot{\epsilon}'_{ij}} \right) \\ \frac{\partial \eta^{\text{vep}}}{\partial p^{\text{trial}}} &= \frac{1}{2\dot{\epsilon}'_{\text{II}}} \left(\sin \phi + a \frac{\partial \dot{\lambda}}{\partial p^{\text{trial}}} + \dot{\lambda} \frac{\partial \eta^{\text{vp}}}{\partial p^{\text{trial}}} \right)\end{aligned}\tag{A.7}$$

where $a = H + \eta^{\text{VP}} + K \Delta t \sin \psi \sin \phi$. The partial derivatives of η^{VP} are non-zero only when $n^{\text{VP}} > 1$. They are then expressed as:

$$\begin{aligned}\frac{\partial \eta^{\text{VP}}}{\partial \dot{\varepsilon}'_{ij}} &= \left(\frac{1}{n^{\text{VP}}} - 1 \right) \frac{\eta^{\text{VP}}}{\dot{\lambda}} \frac{\partial \dot{\lambda}}{\partial \dot{\varepsilon}'_{ij}} \\ \frac{\partial \eta^{\text{VP}}}{\partial p^{\text{trial}}} &= \left(\frac{1}{n^{\text{VP}}} - 1 \right) \frac{\eta^{\text{VP}}}{\dot{\lambda}} \frac{\partial \dot{\lambda}}{\partial p^{\text{trial}}}\end{aligned}\quad (\text{A.8})$$

and the derivatives of the plastic multiplier rate read:

$$\begin{aligned}\frac{\partial \dot{\lambda}}{\partial \dot{\varepsilon}'_{ij}} &= -\frac{\partial F^{-1}}{\partial \dot{\lambda}} \left(\frac{\partial F}{\partial \dot{\varepsilon}'_{ij}} - \dot{\lambda} \frac{\partial \eta^{\text{ve}}}{\partial \dot{\varepsilon}'_{ij}} \right) \\ \frac{\partial \dot{\lambda}}{\partial p^{\text{trial}}} &= -\frac{\partial F^{-1}}{\partial \dot{\lambda}} \frac{\partial F}{\partial p^{\text{trial}}}\end{aligned}\quad (\text{A.9})$$

Finally, the partial derivatives of the yield function can be expressed as:

$$\begin{aligned}\frac{\partial F}{\partial \dot{\varepsilon}'_{ij}} &= b \eta^{\text{ve}} \frac{\dot{\varepsilon}'_{ij}}{\dot{\varepsilon}'_{\text{II}}} + 2 \dot{\varepsilon}'_{\text{II}} \frac{\partial \eta^{\text{ve}}}{\partial \dot{\varepsilon}'_{ij}} \\ \frac{\partial F}{\partial p^{\text{trial}}} &= -\sin \phi.\end{aligned}\quad (\text{A.10})$$

The derivatives of plastic strain rate are expressed as:

$$\begin{aligned}\frac{\partial v_{i,i}^{\text{VP}}}{\partial \dot{\varepsilon}'_{ij}} &= -\left(\frac{\partial Q}{\partial P} \frac{\partial \dot{\lambda}}{\partial \dot{\varepsilon}'_{ij}} + \dot{\lambda} \frac{\partial^2 Q}{\partial \dot{\varepsilon}'_{ij} \partial P} \right) \\ \frac{\partial v_{i,i}^{\text{VP}}}{\partial p^{\text{trial}}} &= -\left(\frac{\partial Q}{\partial P} \frac{\partial \dot{\lambda}}{\partial p^{\text{trial}}} + \dot{\lambda} \frac{\partial^2 Q}{\partial p^{\text{trial}^2}} \right)\end{aligned}\quad (\text{A.11})$$

where the second order partial derivatives of Q are:

$$\begin{aligned}\frac{\partial^2 Q}{\partial \dot{\varepsilon}'_{ij} \partial p^{\text{trial}}} &= \cos(\psi) \frac{\partial \psi}{\partial \dot{\lambda}} \frac{\partial \dot{\lambda}}{\partial \dot{\varepsilon}'_{ij}} \\ \frac{\partial^2 Q}{\partial^2 p^{\text{trial}}} &= \cos(\psi) \frac{\partial \psi}{\partial \dot{\lambda}} \frac{\partial \dot{\lambda}}{\partial p^{\text{trial}}}\end{aligned}\quad (\text{A.12})$$

In the case of Picard iterations, partial derivatives of effective viscosity and corrected pressure with regard to $\dot{\varepsilon}$ are neglected, the tangent operator then simply reads:

$$\mathbf{D}^{\text{pic}} = \begin{bmatrix} 2\eta & 0 & 0 & 0 & 0 \\ 0 & 2\eta & 0 & 0 & 0 \\ 0 & 0 & 2\eta & 0 & 0 \\ 0 & 0 & 0 & 2\eta & 0 \\ 0 & 0 & 0 & 0 & -K \Delta t \end{bmatrix}, \quad (\text{A.13})$$

538 where $\eta = \eta^{\text{ve}}$ below the plastic yield and $\eta = \eta^{\text{vep}}$ at the yield.

539 **B: The tangent operator: consistent tangent finite-step approach (Cons. Tangent)**

540

541 The consistent tangent operator can also be derived by differentiating a finite-step
542 visco-elasto-viscoplastic relation, which is a customary approach in computational en-

543 gineering. However, this is generally done for a displacement-based approach. Geody-
 544 namic codes are usually based on a velocity-pressure formulation, and it is therefore use-
 545 ful to express an appropriate form of the consistent tangent operator.

To this end we first formulate the rheology as:

$$\begin{aligned}\tau_{ij} &= 2\eta^{\text{ve}} \left(\dot{\epsilon}'_{ij} \right) \left[\dot{\epsilon}'_{ij} - \dot{\epsilon}'_{ij}{}^{\text{VP}} \left(\dot{\epsilon}'_{ij}, p^{\text{trial}} \right) \right] \\ p^{\text{corr}} &= p^{\text{trial}} + K\Delta t v_{k,k}^{\text{VP}} \left(\dot{\epsilon}'_{ij}, p^{\text{trial}} \right) = -K\Delta t \left[v'_{k,k} - v_{k,k}^{\text{VP}} \left(\dot{\epsilon}'_{ij}, p^{\text{trial}} \right) \right],\end{aligned}\quad (\text{B.1})$$

where $v'_{k,k} = v_{k,k} - \frac{p^0}{K\Delta t}$. Different from the preceding section, the viscoplastic strain rates and divergence appear explicitly in the deviatoric stress update. There is no need to define an effective V-E-VP viscosity as the trial visco-elastic viscosity will be used.

The constitutive relation can be recast further as:

$$\boldsymbol{\tau} = \mathbf{D}^{\text{ve}} (\dot{\boldsymbol{\epsilon}} - \dot{\boldsymbol{\epsilon}}^{\text{P}}) = \mathbf{D}^{\text{ve}} \dot{\boldsymbol{\epsilon}} - \dot{\lambda} \mathbf{D}^{\text{ve}} \mathbf{m} \quad (\text{B.2})$$

546 where $\mathbf{m} = \left[\frac{\partial Q}{\partial \tau_{xx}} \frac{\partial Q}{\partial \tau_{yy}} \frac{\partial Q}{\partial \tau_{zz}} \frac{\partial Q}{\partial \tau_{xy}} - \frac{\partial Q}{\partial p^{\text{trial}}} \right]^T$, and thus contains derivatives with regard
 547 to the stress deviators as well as the pressure. Due to power-law viscous rheology, \mathbf{D}^{ve}
 548 depends on the deviatoric strain rate, thus on $\dot{\boldsymbol{\epsilon}}$, where $\boldsymbol{\tau} = [\tau_{xx} \ \tau_{yy} \ \tau_{zz} \ \tau_{xy} \ p^{\text{corr}}]^T$, $\dot{\boldsymbol{\epsilon}} =$
 549 $[\epsilon_{xx} \ \epsilon_{yy} \ \epsilon_{zz} \ \epsilon_{xy} \ v'_{k,k}]^T$. The operator \mathbf{D}^{ve} is expressed as:

$$\mathbf{D}^{\text{ve}} = \begin{bmatrix} 2\eta^{\text{ve}} & 0 & 0 & 0 & 0 \\ 0 & 2\eta^{\text{ve}} & 0 & 0 & 0 \\ 0 & 0 & 2\eta^{\text{ve}} & 0 & 0 \\ 0 & 0 & 0 & 2\eta^{\text{ve}} & 0 \\ 0 & 0 & 0 & 0 & -K\Delta t \end{bmatrix}, \quad (\text{B.3})$$

550 For generality, we will consider the case of power-law visco-elastic flow, so that η^{ve}
 551 is a non-linear function of the strain rate.

We consider an infinitesimal perturbation of stress-strain rate relationship such that:

$$\delta \boldsymbol{\tau} = \mathbf{D}^{\text{ve}} \delta \dot{\boldsymbol{\epsilon}} + \delta \mathbf{D}^{\text{ve}} \dot{\boldsymbol{\epsilon}} - \delta \dot{\lambda} \mathbf{D}^{\text{ve}} \mathbf{m} - \delta \mathbf{D}^{\text{ve}} \dot{\lambda} \mathbf{m} - \dot{\lambda} \mathbf{D}^{\text{ve}} \frac{\partial \mathbf{m}}{\partial \boldsymbol{\tau}} \delta \boldsymbol{\tau}. \quad (\text{B.4})$$

The above expression can be recast as:

$$\delta \boldsymbol{\tau} = \mathbf{E}^{-1} \mathbf{D}^{\text{ve}} \delta \dot{\boldsymbol{\epsilon}} + \mathbf{E}^{-1} \delta \mathbf{D}^{\text{ve}} \dot{\boldsymbol{\epsilon}}^{\text{ve}} - \mathbf{E}^{-1} \mathbf{D}^{\text{ve}} \mathbf{m} \delta \dot{\lambda} \quad (\text{B.5})$$

where $\mathbf{E} = \left(\mathbf{I} + \mathbf{D}^{\text{ve}} \dot{\lambda} \frac{\partial \mathbf{m}}{\partial \boldsymbol{\tau}} \right)$ and $\dot{\boldsymbol{\epsilon}}^{\text{ve}} = \dot{\boldsymbol{\epsilon}} - \dot{\boldsymbol{\epsilon}}^{\text{P}}$. The consistency condition implies that $\delta F = 0$ during plastic loading, hence:

$$\left(\frac{\partial F}{\partial \boldsymbol{\tau}} \right)^T \delta \boldsymbol{\tau} + \frac{\partial F}{\partial \dot{\lambda}} \delta \dot{\lambda} + \frac{\partial F}{\partial \lambda} \delta \lambda = 0. \quad (\text{B.6})$$

The infinitesimal perturbation of plastic multiplier rate, $\delta\dot{\lambda}$, can be obtained by introducing infinitesimal perturbation of stress-strain rate relationship into the consistency condition:

$$\delta\dot{\lambda} = \frac{\left(\frac{\partial F}{\partial \boldsymbol{\tau}}\right)^T \mathbf{E}^{-1} \mathbf{D}^{\text{ve}} \delta \dot{\boldsymbol{\epsilon}} + \left(\frac{\partial F}{\partial \boldsymbol{\tau}}\right)^T \mathbf{E}^{-1} \delta \mathbf{D}^{\text{ve}} \dot{\boldsymbol{\epsilon}}^{\text{ve}}}{\left(\frac{\partial F}{\partial \boldsymbol{\tau}}\right)^T \mathbf{E}^{-1} \mathbf{D}^{\text{ve}} \mathbf{m} - \frac{\partial F}{\partial \lambda} - \Delta t \frac{\partial F}{\partial \lambda}}, \quad (\text{B.7})$$

where we have assumed that $\delta\lambda = \delta\dot{\lambda}\Delta t$.

Finally, the consistent tangent operator may be expressed by substituting $\delta\dot{\lambda}$ into expression B.5 and by letting $\mathbf{D}^{\text{ctl}} \equiv \frac{\partial \boldsymbol{\tau}}{\partial \dot{\boldsymbol{\epsilon}}}$:

$$\mathbf{D}^{\text{ctl}} = \mathbf{E}^{-1} \mathbf{B} + \mathbf{E}^{-1} \mathbf{D}^{\text{ve}} \left(\mathbf{I} - \frac{\mathbf{m} \left(\frac{\partial F}{\partial \boldsymbol{\tau}}\right)^T \mathbf{E}^{-1} (\mathbf{B} + \mathbf{D}^{\text{ve}})}{\left(\frac{\partial F}{\partial \boldsymbol{\tau}}\right)^T \mathbf{E}^{-1} \mathbf{D}^{\text{ve}} \mathbf{m} - \frac{\partial F}{\partial \lambda} - \Delta t \frac{\partial F}{\partial \lambda}} \right) \quad (\text{B.8})$$

where $\mathbf{B} = \frac{\partial \mathbf{D}^{\text{ve}} \dot{\boldsymbol{\epsilon}}^{\text{ve}}}{\partial \boldsymbol{\tau}}$, which can be written explicitly written as:

$$\mathbf{B} = \begin{bmatrix} 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{xx}} \dot{\epsilon}^{\text{ve}}_{xx} & 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{yy}} \dot{\epsilon}^{\text{ve}}_{xx} & 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{zz}} \dot{\epsilon}^{\text{ve}}_{xx} & 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{xy}} \dot{\epsilon}^{\text{ve}}_{xx} & 0 \\ 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{xx}} \dot{\epsilon}^{\text{ve}}_{yy} & 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{yy}} \dot{\epsilon}^{\text{ve}}_{yy} & 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{zz}} \dot{\epsilon}^{\text{ve}}_{yy} & 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{xy}} \dot{\epsilon}^{\text{ve}}_{yy} & 0 \\ 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{zz}} \dot{\epsilon}^{\text{ve}}_{zz} & 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{yy}} \dot{\epsilon}^{\text{ve}}_{zz} & 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{zz}} \dot{\epsilon}^{\text{ve}}_{zz} & 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{xy}} \dot{\epsilon}^{\text{ve}}_{zz} & 0 \\ 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{xx}} \dot{\epsilon}^{\text{ve}}_{xy} & 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{yy}} \dot{\epsilon}^{\text{ve}}_{xy} & 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{zz}} \dot{\epsilon}^{\text{ve}}_{xy} & 2 \frac{\partial \eta^{\text{ve}}}{\partial \dot{\epsilon}'_{xy}} \dot{\epsilon}^{\text{ve}}_{xy} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (\text{B.9})$$

The partial derivatives of the trial viscosity with regard to effective deviatoric strain rates are given in the previous appendix section. In practice, one may further reduce the size of systems to 3×3 by introducing constraint on out-of-plane deformation. The inverse of the matrix \mathbf{E} can then be computed analytically.

The continuum tangent operator (Cont. Tangent) can be formulated by neglecting the second order derivatives of the plastic flow potential ($\frac{\partial \mathbf{m}}{\partial \boldsymbol{\tau}} = \mathbf{0}$) and thus setting $\mathbf{E} = \mathbf{I}$.

C: Governing equations for lithospheric models (MDoodz6.0)

The models presented in Sec 5.1 were intended to show the evolution of a V-E-VP in a state-of-the-art geodynamic simulation, thus involving large deformations and com-

posite rheological modelling. The following set of thermo-mechanical equations were solved:

$$\begin{aligned}\frac{\partial \tau_{ij}}{\partial x_j} - \frac{\partial p}{\partial x_i} &= -\rho g_i, \\ \frac{1}{K} \frac{dp}{dt} &= -\left(v_{i,i} - v_{i,i}^{\text{VP}} - v_{i,i}^{\text{th}}\right) \\ \rho c_P \frac{dT}{dt} &= -\frac{\partial q_i}{\partial x_i} + H_R + H_S + H_A,\end{aligned}\tag{C.1}$$

where T is the temperature and ρ , c_P , k and α are the density, the heat capacity at constant pressure, the thermal conductivity, and the thermal expansivity, respectively.

The term $q_i = -k \frac{\partial T}{\partial x_i}$ is the heat flux vector and $v_{i,i}^{\text{th}} = \alpha \frac{dT}{dt}$ is the thermal divergence rate. H_R is the radioactive heating, $H_S = \tau_{ij} (\dot{\epsilon}_{ij} - \dot{\epsilon}_{ij}^e)$ is shear heating and $H_A = \alpha T \frac{dp}{dt}$ is adiabatic heating.

The deviatoric stress rate includes term arising from advection, rotation and stretching:

$$\dot{\boldsymbol{\tau}} = \dot{\boldsymbol{\tau}} - \mathbf{v} \nabla \boldsymbol{\tau} - \nabla \mathbf{v}^T \cdot \boldsymbol{\tau} - \boldsymbol{\tau} \cdot \nabla \mathbf{v}\tag{C.2}$$

In the current implementation, this operation is split such that the update is semi-implicit. At each time step, the deviatoric stress tensor is advected and rotated using the velocity and stress field obtained after convergence of the non-linear solver. The rheological model is composite and is based on an additive decomposition the different creep strain rates. Such rheological model is used to represent the creep mechanisms in the lithospheric mantle (dislocation, diffusion and Peierls). The different mechanisms are combined within local rheological iterations and allow for determination of trial visco-elastic stress (See *Popov and Sobolev* [2008] and Sec. 13). Diffusion and dislocation creep are expressed as power-laws, such that:

$$\eta^{\text{Dif,Dis}} = C^{\text{Dif,Dis}} \dot{\epsilon}_{\text{II}}^{\frac{1}{n^{\text{Dif,Dis}}} - 1} d^{\frac{m^{\text{Dif,Dis}}}{n^{\text{Dif,Dis}}}}\tag{C.3}$$

where d is the grain size, m is the grain size exponent, n is the stress exponent. The factor C is expressed as:

$$C^{\text{Dif,Dis}} = \left[2F^{\text{Dif,Dis}} \left(A^{\text{Dif,Dis}} \right)^{-\frac{1}{n}} \exp \left(\frac{Q^{\text{Dif,Dis}} + PV^{\text{Dif,Dis}}}{n^{\text{Dif,Dis}} RT} \right) \right]^{-n^{\text{Dif,Dis}}}\tag{C.4}$$

where Q is activation energy, V is activation volume, F is the correction factor for conversion of experimental data to invariant formulation (see *Schmalholz and Fletcher* [2011]) and R is the gas constant. Peierls creep is implemented using the effective power-law formulation of *Kameyama et al.* [1999], the subsequent Peierls effective viscosity is thus expressed as:

$$\eta^{\text{Peierls}} = C^{\text{Peierls}} \dot{\epsilon}_{\text{II}}^{\frac{1}{n^{\text{Peierls}}} - 1}\tag{C.5}$$

Table C.1. Material parameters used in the lithospheric simulations. Additional parameters for Peierls creep are $\gamma = 0.1$ and $\sigma^{\text{Peierls}} = 8.5 \times 10^9$ Pa. Thermal expansivity, heat capacity, bulk modulus, shear modulus are assumed constant for each phases ($\alpha = 3.2 \times 10^{-5} \text{ K}^{-1}$, $c_P = 1050 \text{ J.kg}^{-1}.\text{K}^{-1}$, $K = 2 \times 10^{10} \text{ Pa}$, $G = 3 \times 10^{10} \text{ Pa}$). Softening of the friction angle (upper and middle crust) occurs within an accumulated plastic strain of 0.5. Flow law parameter are taken from *Hansen and Carter* [1983] for the upper crust, *Rybacki Mackwell et al.* [1998] for the lower crust, *Hirth and Kohlstedt* [2004] and *Goetze and Evans* [1979] for the lithospheric mantle.

	$A [\text{Pa}^{-n}.\text{s}^{-1}]$	$Q [\text{J.mol}^{-1}]$	$V [\text{m}^3.\text{mol}^{-1}]$	n	m	$C [\text{Pa}]$	$\phi [^\circ]$	$\psi [^\circ]$	$\rho_0 [\text{kg.m}^{-3}]$
Upper crust	3.1623×10^{-26}	186.5×10^3	0.0	3.3	0.0	2×10^7	$30 \rightarrow 10$	5	2700
Middle crust	3.9811×10^{-16}	356×10^3	0.0	3.0	0.0	2×10^7	$30 \rightarrow 20$	5	2750
Lower crust	5.0477×10^{-28}	485×10^3	0.0	4.7	0.0	2×10^7	30	5	2800
Inclusion	1.1×10^{-16}	530×10^3	11×10^{-6}	3.5	0.0	10^6	0	0	3260
Mantle Diffusion	1.5×10^{-15}	375×10^3	4×10^{-6}	1.0	3.0	10^7	30	5	3260
Mantle Dislocation	1.1×10^{-16}	530×10^3	11×10^{-6}	3.5	0.0	10^7	30	5	3260
Mantle Peierls	5.7×10^{11}	540×10^3	–	–	–	10^7	30	5	3260

where n^{Peierls} is formulated as

$$n^{\text{Peierls}} = \frac{Q^{\text{Peierls}}}{RT} q \gamma (1 - \gamma)^{q-1} \quad (\text{C.6})$$

and

$$C^{\text{Peierls}} = A^{\text{Peierls}} \left(\gamma \sigma^{\text{Peierls}} \right)^{-n^{\text{Peierls}}} \exp \left(-\frac{Q^{\text{Peierls}}}{RT} (1 - \gamma)^2 \right) \quad (\text{C.7})$$

567 where Q^{Peierls} , γ , σ^{Peierls} are material parameters (see table below).

The density is assumed to be pressure and temperature dependent and follows the equation of state:

$$\rho = \rho_0 \exp \left(-\alpha T + \frac{P}{K} \right) \quad (\text{C.8})$$

568 where ρ_0 is the reference density.

References

- 569
- 570 Bates, D. M. (2007), Direct Methods for Sparse Linear Systems by Timothy A.
- 571 Davis, *International Statistical Review*, *75*(2), 260–261.
- 572 Bauville, A., M. Furuichi, and M. (2020), Control of fault weakening on the struc-
- 573 tural styles of underthrusting-dominated non-cohesive accretionary wedges,
- 574 *Journal of Geophysical Research: Solid Earth*, *125*(3), e2019JB019,220, doi:
- 575 <https://doi.org/10.1029/2019JB019220>, e2019JB019220 10.1029/2019JB019220.
- 576 Bažant, Z. P., and F.-B. Lin (1988), Non-local yield limit degradation, *International*
- 577 *Journal for Numerical Methods in Engineering*, *26*(8), 1805–1823.
- 578 Bessat, A., T. Duretz, G. Hetényi, S. Pilet, and S. M. Schmalholz (2020), Stress and
- 579 deformation mechanisms at a subduction zone: Insights from 2-D thermomechanical
- 580 numerical modelling, *Geophysical Journal International*, *221*(3), 1605–1625.
- 581 Byerlee, J. (1978), Friction of rocks, *Pure and Applied Geophysics*, *116*, 615–626.
- 582 Candioti, L. G., S. M. Schmalholz, and T. Duretz (2020), Impact of upper mantle
- 583 convection on lithosphere hyper-extension and subsequent convergence-induced
- 584 subduction, *Solid Earth Discussions*, *2020*, 1–41, doi:10.5194/se-2020-88.
- 585 Choi, E., and K. D. Petersen (2015), Making Coulomb angle-oriented shear bands in
- 586 numerical tectonic models, *Tectonophysics*, *657*, 94–101.
- 587 Commend, S., A. Truty, and T. Zimmermann (2004), Stabilized finite elements
- 588 applied to elastoplasticity: I. mixed displacementpressure formulation, *Com-*
- 589 *puter Methods in Applied Mechanics and Engineering*, *193*(33), 3559–3586, doi:
- 590 <https://doi.org/10.1016/j.cma.2004.01.007>.
- 591 Crameri, F. (2018), Geodynamic diagnostics, scientific visualisation and staglab 3.0,
- 592 *Geoscientific Model Development*, *11*(6), 2541–2562.
- 593 de Borst, R., and T. Duretz (2020), On viscoplastic regularisation in strain-softening
- 594 rocks and soils, *International Journal for Numerical and Analytical Methods in*
- 595 *Geomechanics*, *44*, 890–903.
- 596 de Borst, R., and P. H. Feenstra (1990), Studies in anisotropic plasticity with refer-
- 597 ence to the Hill criterion, *International Journal for Numerical Methods in Engi-*
- 598 *neering*, *29*, 315–336.
- 599 de Borst, R., and H. B. Mühlhaus (1992), Gradient-dependent plasticity: Formu-
- 600 lation and algorithmic aspects, *International Journal for Numerical Methods in*

601 *Engineering*, 35, 521–539.

602 de la Puente, J., J.-P. Ampuero, and M. Käser (2009), Dynamic rupture model-
603 ing on unstructured meshes using a discontinuous Galerkin method, *Journal of*
604 *Geophysical Research: Solid Earth*, 114(B10302), 1–17.

605 Duretz, T., D. May, and P. Yamato (2016a), A free surface capturing discretization
606 for the staggered grid finite difference scheme, *Geophysical Journal International*,
607 204(3), 1518–1530.

608 Duretz, T., B. Petri, G. Mohn, S. M. Schmalholz, F. L. Schenker, and O. Müntener
609 (2016b), The importance of structural softening for the evolution and architecture
610 of passive margins, *Scientific Reports*, 6, 38,704.

611 Duretz, T., A. Souche, R. de Borst, and L. Le Pourhiet (2018), The benefits of using
612 a consistent tangent operator for viscoelastoplastic computations in geodynamics,
613 *Geochemistry, Geophysics, Geosystems*, 19(2018GC007877), 4904–4024.

614 Duretz, T., R. de Borst, and L. Le Pourhiet (2019), On finite thickness of shear
615 bands in frictional viscoplasticity, and implications for lithosphere dynamics, *Geo-*
616 *chemistry, Geophysics, Geosystems*, 20(2019GC008531), 5598–5616.

617 Duretz, T., R. de Borst, P. Yamato, and L. Le Pourhiet (2020), Toward robust and
618 predictive geodynamic modeling: The way forward in frictional plasticity, *Geo-*
619 *physical Research Letters*, 47(5), e2019GL086,027.

620 Gerbault, M., A. N. B. Poliakov, and M. Daignieres (1998), Prediction of faulting
621 from the theories of elasticity and plasticity: what are the limits?, *Journal of*
622 *Structural Geology*, 20(2), 301 – 320.

623 Gerya, T. V., and D. A. Yuen (2007), Robust characteristics method for modelling
624 multiphase visco-elasto-plastic thermo-mechanical problems, *Physics of the Earth*
625 *and Planetary Interiors*, 163, 83–105.

626 Glerum, A., C. Thieulot, M. Fraters, C. Blom, and W. Spakman (2018), Nonlinear
627 viscoplasticity in ASPECT: benchmarking and applications to subduction, *Solid*
628 *Earth*, 9(2), 267–294, doi:10.5194/se-9-267-2018.

629 Goetze, C., and B. Evans (1979), Stress and temperature in the bending lithosphere
630 as constrained by experimental rock mechanics, *Geophysical Journal International*,
631 59(3), 463–478.

632 Hansen, F., and N. Carter (1983), Semibrittle creep of dry and wet westerly granite
633 at 1000 mpa.

- 634 Heeres, O. M., A. S. J. Suiker, and R. de Borst (2002), A comparison between the
635 Perzyna viscoplastic model and the consistency viscoplastic model, *European*
636 *Journal of Mechanics: A/Solids*, *21*, 1 – 12.
- 637 Hirth, G., and D. Kohlstedt (2004), *Rheology of the Upper Mantle and the Mantle*
638 *Wedge: A View from the Experimentalists*, pp. 83–105, American Geophysical
639 Union (AGU).
- 640 Huismans, R. S., and C. Beaumont (2003), Symmetric and asymmetric lithospheric
641 extension: Relative effects of frictional-plastic and viscous strain softening, *Jour-*
642 *nal of Geophysical Research: Solid Earth*, *108*(B10), doi:10.1029/2002JB002026.
- 643 Huismans, R. S., and C. Beaumont (2007), Roles of lithospheric strain softening
644 and heterogeneity in determining the geometry of rifts and continental mar-
645 gins, *Geological Society, London, Special Publications*, *282*(1), 111–138, doi:
646 10.1144/SP282.6.
- 647 Jacquey, A. B., and M. Cacace (2020), Multiphysics modeling of a brittle-ductile
648 lithosphere: 1. explicit visco-elasto-plastic formulation and its numerical imple-
649 mentation, *Journal of Geophysical Research: Solid Earth*, *125*(1), e2019JB018474.
- 650 Kameyama, M., D. A. Yuen, and S.-I. Karato (1999), Thermal-mechanical effects
651 of low-temperature plasticity (the peierls mechanism) on the deformation of a
652 viscoelastic shear zone, *Earth and Planetary Science Letters*, *168*(1), 159 – 172.
- 653 Kaus, B. J. P., A. A. Popov, T. S. Baumann, A. E. Püsök, A. Bauville, N. Fer-
654 nandez, and M. Collignon (2016), Forward and inverse modelling of litho-
655 spheric deformation on geological timescales, in *NIC Symposium 2016*, vol. 48,
656 edited by K. Binder, M. Müller, A. Kremer, and A. Schnurpfeil, pp. 299–307,
657 Forschungszentrum Jülich, Jülich.
- 658 Kiss, D., L. G. Candiotti, T. Duretz, and S. M. Schmalholz (2019), Thermal softening
659 induced subduction initiation at a passive margin, *Geophysical Journal Interna-*
660 *tional*, *220*(3), 2068–2073.
- 661 Lemiale, V., H. B. Mühlhaus, L. Moresi, and J. Stafford (2008), Shear banding anal-
662 ysis of plastic models formulated for incompressible viscous flows, *Physics of the*
663 *Earth and Planetary Interiors*, *171*, 177–186.
- 664 Mackwell, S. J., M. E. Zimmerman, and D. L. Kohlstedt (1998), High-temperature
665 deformation of dry diabase with application to tectonics on venus, *Journal of*
666 *Geophysical Research: Solid Earth*, *103*(B1), 975–984.

- 667 May, D. A., J. Brown, and L. L. Pourhiet (2014), ptatin3d: High-performance meth-
668 ods for long-term lithospheric dynamics, in *SC '14: Proceedings of the Interna-*
669 *tional Conference for High Performance Computing, Networking, Storage and*
670 *Analysis*, pp. 274–284, doi:10.1109/SC.2014.28.
- 671 Moresi, L., F. Dufour, and H. Mühlhaus (2003), A Lagrangian integration point fi-
672 nite element method for large deformation modeling of viscoelastic geomaterials,
673 *Journal of Computational Physics*, *184*, 476–497.
- 674 Mühlhaus, H. B., and I. Vardoulakis (1987), The thickness of shear bands in granu-
675 lar materials, *Géotechnique*, *37*, 271–283.
- 676 Naliboff, J. B., S. J. H. Buiter, G. Péron-Pinvidic, P. T. Osmundsen, and
677 J. Tetreault (2017), Complex fault interaction controls continental rifting, *Na-*
678 *ture Communications*, *8*, 1179.
- 679 Naliboff, J. B., A. Glerum, S. Brune, G. Pron-Pinvidic, and T. Wrona (2020), De-
680 velopment of 3-d rift heterogeneity through fault network evolution, *Geophys-*
681 *ical Research Letters*, *47*(13), e2019GL086611, doi:https://doi.org/10.1029/
682 2019GL086611, e2019GL086611 2019GL086611.
- 683 Niazi, M. S., H. H. Wisselink, and T. Meinders (2013), Viscoplastic regularization of
684 local damage models: Revisited, *Comput. Mech.*, *51*(2), 203–216.
- 685 Petri, B., T. Duretz, G. Mohn, S. M. Schmalholz, G. D. Karner, and O. Müntener
686 (2019), Thinning mechanisms of heterogeneous continental lithosphere, *Earth and*
687 *Planetary Science Letters*, *512*, 147 – 162.
- 688 Poh, J., P. Yamato, T. Duretz, D. Gapais, and P. Ledru (2020), Precambrian defor-
689 mation belts in compressive tectonic regimes: A numerical perspective, *Tectono-*
690 *physics*, *777*, 228,350.
- 691 Poliakov, A., Y. Podladchikov, and C. Talbot (1993), Initiation of salt diapirs with
692 frictional overburdens: numerical experiments, *Tectonophysics*, *228*(3), 199 – 210.
- 693 Popov, A. A., and S. V. Sobolev (2008), SLIM3D: A tool for three-dimensional
694 thermomechanical modeling of lithospheric deformation with elasto-visco-plastic
695 rheology, *Physics of the Earth and Planetary Interiors*, *171*, 55–75.
- 696 Räss, L., T. Duretz, Y. Y. Podladchikov, and S. M. Schmalholz (2017), M2Di: Con-
697 cise and efficient MATLAB 2-D Stokes solvers using the Finite Difference Method,
698 *Geochemistry, Geophysics, Geosystems*, *18*(2), 755–768.

- 699 Räss, L., T. Duretz, and Y. Y. Podladchikov (2019), Resolving hydromechanical
700 coupling in two and three dimensions: spontaneous channelling of porous fluids
701 owing to decompaction weakening, *Geophysical Journal International*, *218*(3),
702 1591–1616, doi:10.1093/gji/ggz239.
- 703 Rudnicki, J. W., and J. R. Rice (1975), Conditions for the localization of deforma-
704 tion in pressure sensitive dilatant materials, *Journal of the Mechanics and Physics
705 of Solids*, *23*, 371–394.
- 706 Rutter, E. H., and C. T. Glover (2012), The deformation of porous sandstones:
707 are Byerlee friction and the critical state line equivalent?, *Journal of Structural
708 Geology*, *44*, 129 – 140.
- 709 Sabet, S. A., and R. de Borst (2019), Structural softening, mesh dependence, and
710 regularisation in non-associated plastic flow, *International Journal for Numerical
711 and Analytical Methods in Geomechanics*, *43*, 2170–2183.
- 712 Schmalholz, S., T. Duretz, F. Schenker, and Y. Podladchikov (2014), Kinematics and
713 dynamics of tectonic nappes: 2-d numerical modelling and implications for high
714 and ultra-high pressure tectonism in the western alps, *Tectonophysics*, *631*(C),
715 160–175, doi:10.1016/j.tecto.2014.05.018.
- 716 Schmalholz, S. M., and T. Duretz (2017), Impact of grain size evolution on necking
717 in calcite layers deforming by combined diffusion and dislocation creep, *Journal of
718 Structural Geology*, *103*, 37 – 56.
- 719 Schmalholz, S. M., and R. C. Fletcher (2011), The exponential flow law applied to
720 necking and folding of a ductile layer, *Geophysical Journal International*, *184*(1),
721 83–89.
- 722 Shipton, Z. K., A. M. Soden, J. D. Kirkpatrick, A. M. Bright, and R. J. Lunn
723 (2006), How thick is a fault? Fault displacement-thickness scaling revisited, in
724 *Earthquakes: Radiated Energy and the Physics of Faulting*, edited by R. Aber-
725 crombie, A. McGarr, G. Di Toro, and H. Kanamori, pp. 193–198, American Geo-
726 physical Union (AGU).
- 727 Spiegelman, M., D. A. May, and C. R. Wilson (2016), On the solvability of in-
728 compressible Stokes with viscoplastic rheologies in geodynamics, *Geochemistry,
729 Geophysics, Geosystems*, *17*, 2213–2238.
- 730 Stefanou, I., J. Sulem, and H. Rattetz (2017), *Cosserat Approach to Localiza-
731 tion in Geomaterials*, pp. 1–25, Springer International Publishing, Cham, doi:

732 10.1007/978-3-319-22977-5_10-1.

733 Wang, H. (2019), Viscous and second gradient regularization techniques for the
734 description of the behavior of geomaterials, Theses, École centrale de Nantes.

735 Wang, W. M., L. J. Sluys, and R. de Borst (1997), Viscoplasticity for instabilities
736 due to strain softening and strain-rate softening, *International Journal for Numerical*
737 *Methods in Engineering*, 40(20), 3839–3864.

738 Willett, S. D. (1992), Dynamic and kinematic growth and change of a Coulomb
739 wedge, in *Thrust Tectonics*, edited by K. R. McClay, pp. 19–31, Springer, Dor-
740 drecht.

741 Yamato, P., T. Duretz, and S. Angiboust (2019), Brittle/ductile deformation of
742 eclogites: Insights from numerical models, *Geochemistry, Geophysics, Geosystems*,
743 20(7), 3116–3133.

744 Zhao, X., and M. Cai (2010), A mobilized dilation angle model for rocks, *Internation-*
745 *al Journal of Rock Mechanics and Mining Sciences*, 47(3), 368 – 384.

746 Zoback, M. D., R. Apel, J. Baumgärtner, M. Brudy, R. Emmermann, B. Engeser,
747 K. Fuchs, W. Kessels, H. Rischmüller, F. Rummel, and L. Vernik (1993), Upper-
748 crustal strength inferred from stress measurements to 6 km depth in the KTB
749 borehole, *Nature*, 365(6447), 633–635.

750 **Acknowledgments**

751 The authors acknowledge Boris Kaus and John Naliboff for reviewing an earlier version
752 of this manuscript. The authors acknowledge Rennes Métropole and the French programme
753 NEEDS for their financial supports to the project. The authors thank Carolina Lithgow-
754 Bertelloni for editing the manuscript. The research reported in this article has been par-
755 tially supported by the European Research Council under grant 664734 "PoroFrac". Au-
756 thor TD thanks Prof. Taras Gerya for exciting discussions. This study is based on for-
757 ward numerical simulations and does not involve any new data.

Accepted Article