

Rapid shear zone weakening during subduction initiation

Yida Li and Michael Gurnis

Seismological Laboratory, California Institute of Technology, 1200 East California Blvd., Pasadena, 91125, CA, USA

This manuscript was compiled on September 1, 2024

Subduction zones play a pivotal role in the mechanics of plate tectonics by providing the driving force through slab pull and weak megathrusts that facilitate the relative motion between tectonic plates. The initiation of subduction zones is intricately linked to the accumulation of slab pull and development of weakness at plate boundaries and by consequence the largest changes in the energetics of mantle convection. However, the transient nature of subduction initiation accompanied by intense subsequent tectonic activity, leaves critical evidence poorly preserved and making subduction initiation difficult to constrain. We overcome these limitations through a comprehensive analysis focused on Puysegur, a well-constrained extant example of subduction initiation offshore South Island, New Zealand. Through time-dependent, three-dimensional thermo-mechanical computations and quantitative comparison to new geophysical and geological observations, including topography, stratigraphy, and seismicity, we demonstrate that subduction initiation develops with a fast strain weakening described with a small characteristic displacement ($\Delta_s \approx 4 - 8$ km). Potential physical mechanisms contributing to the strain weakening are explored and we find that the observed fast weakening may arise through a combination of grain-size reduction within the lower lithosphere and fluid pressurization at shallower depths. With the shared commonality in the underlying physics of tectonic processes, the rapid strain weakening constrained at Puysegur offers insights into the formation of the first subduction during early Earth and the onset of plate tectonics.

subduction | plate tectonics | geodynamics | New Zealand

Slab pull during subduction is arguably the dominant driving force for plate motion and mantle convection, and for mature subduction zones, this pull force is sufficiently large for the convective system to be self-sustaining. However, when a new subduction zone forms, slab pull is typically insufficient to overcome the resistance to plate motion that occurs within nascent plate boundaries. How subduction initiation can be triggered with an initially insufficient driving force remains a fundamental, open question in geodynamics and plate tectonics. Arguably, the key reason this uncertainty in the energetics and dynamics of plate tectonics remains is that the strain required to weaken plate boundaries during initiation has not been constrained.

Based on mechanical models and the geology and geophysics where subduction has initiated, a variety of initiation mechanisms have been proposed, including compression-induced (1), plume-induced (2), spontaneous initiation (3), and collapse of passive margins (4). Despite the variability and complex tectonics that often surrounds subduction initiation, the mechanisms share much in common through the balance of forces (5), such that a reduction of resistance or an increase of driving forces provide more favorable conditions for subduction initiation. Therefore, either external forces, which supply additions to those driving the system, or small yield stresses, which reduce the resistance, have been invoked in previous subduction initiation scenarios. Strain weakening is nearly universally advanced in models as a means to lower the strength of rocks within evolving plate boundaries (6–8); weakening is essential because it leads to the localization of deformation and formation of faults and shear zones. Despite proposals of weakening mechanisms like grain size reduction (9), shear heating (10, 11), fluid-pressure (12) and reaction induced (13) weakening, the rate at which a margin loses its strength with strain has been a crucial quantity that has remained elusive for known examples of subduction initiation. Here, we overcome this fundamental limitation with the well-observed Puysegur subduction zone within the context of four-dimensional dynamic models, place bounds on this critical weakening for the first time, and then compare this rate to values predicted by the principal physical models of weakening.

Significance Statement

The subduction zone megathrust is a critical weak interface facilitating the movement of adjacent tectonic plates. However, the mechanisms governing its initiation, notably the rate of weakening during subduction initiation, remain enigmatic. Focusing on the Puysegur subduction zone south of New Zealand, we employ advanced 3D-numerical models to study this process. Through comprehensive integration of these models with geophysical and geological observations, we find a rapid weakening is preferred during the formation of a new subduction zone. With the quantitatively-constrained strain weakening rate, we evaluate shear heating, grain size reduction, and fluid pressurization as three potential mechanisms, and find that only a combination of fluid pressurization at shallower depth and grain-size reduction at greater depth provides the observed rapid weakening.

Author affiliations: ¹Seismological Laboratory, California Institute of Technology, 1200 East California Blvd., Pasadena, 91125, CA, USA

Y.L. and M.G. designed the research. Y.L. performed the calculation. Y.L. and M.G. analyzed the data and wrote the paper.

The authors declare no conflict of interest.

¹To whom correspondence should be addressed. E-mail: yidali@caltech.edu

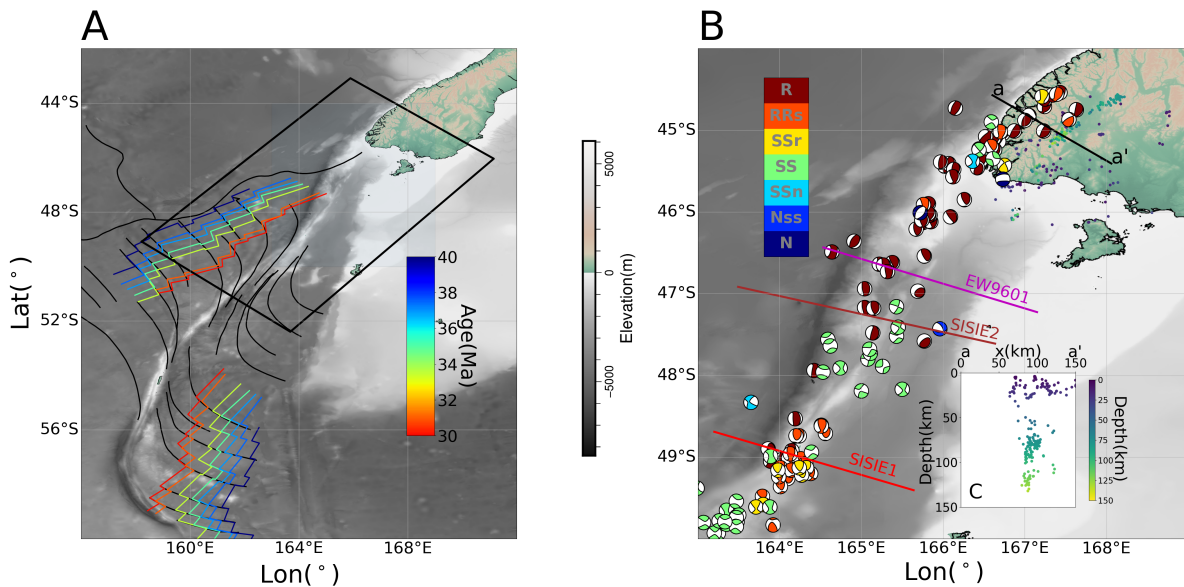


Fig. 1. Summary of Puysegur subduction zone. A. Fracture zones (black curves) and seafloor magnetic anomalies (color-coded with age). In the background is shown bathymetry (greyscale) from GMRT (23). Black box shows the computational model domain. B. Focal mechanisms from GCMT catalogue (24) with classified fault types (25) for shallow earthquake (depths < 30 km), and seismicity under Fiordland (colored dots near aa' line). Three colored, labeled lines are the locations of high-resolution seismic lines used to infer the evolving stress. C. (Inset) Cross-section of seismicity along aa' (26) under Fiordland. Seismicity in map view(B) and cross-section(C) are color-coded with depth using the colorbar of C.

The Puysegur Trench (Fig. 1), south of New Zealand, is an extant example of initiation, as the subduction zone is experiencing ongoing initiation since 15 Ma. Well defined by magnetic lineations and fracture zones (Fig. 1A), the kinematics is well known (14–16) with the plate boundary between the Australian and Pacific Plates experiencing a transition from oceanic spreading (40 to 25 Ma), to strike-slip motion (25 to 15 Ma), and eventually transpression and subduction (15 Ma to the present). Recently, targeted multi-channel seismic imaging with specific seismic-stratigraphic horizons dated using constraints from offshore drilling, were used to place bounds on a progressive transition of the in-plane stress at the nucleating boundary (17). The observations provide constraints on the relative state of stress, compressive versus tensional through the amount of strain (fault offsets) in time and space. Specifically, the northern Puysegur Trench displays fold-thrust features in the strata between 16 Ma and 8 Ma, with on-lapping passive strata, indicating the stress state started with compression between 15 Ma and 8 Ma, but switched to extension after 8 Ma. However, for southern Puysegur the compressive sequence began later, from 8 Ma to 5 Ma, and that a reversal in the state of stress (from compression to tension) has yet to occur. This observation of stress evolution is consistent with the earlier inference of uplift followed by ≈ 1.5 km of subsidence of the Puysegur Ridge in the northern section and only uplift in the southern (18) as well as the strong free-air gravity anomalies along the ridge (19). The vertical motions and compression-extension transition is a characteristic feature of induced subduction initiation, indicating a slab pull that builds and eventually dominates as the driving force, making the system self-sustaining. The 8 Myr phase lag between the northern and southern sections of Puysegur Trench may indicate the nucleation of Puysegur subduction starting in the north while propagating southward (17). Today, the northern

part of Puysegur is becoming a more mature subduction zone, with the seismicity of the slab reaching more than 130 km depth below Fiordland (20), the immediate onshore region of South Island, New Zealand (Fig. 1C). Meanwhile, the southern Puysegur exhibits strain partitioning, the process by which the relative plate motion between the Australian and Pacific plates is distributed across a series of strike-slip and oblique thrust faults, as a result of transpression (21). Although Puysegur subduction initiated at the margin of stretched continental crust, it is a sliver of oceanic crust that became trapped between the vertical strike-slip Puysegur fault and the dipping Puysegur megathrust (19, 21, 22).

Together, these observations provide a globally unique data set spanning the four-dimensional nature of subduction initiation from the nascent state with known antecedent tectonics, well-constrained plate kinematics during the entire period of initiation, state-of-stress in time and space along the plate boundary, and present-day structural controls from topography, gravity and seismology. Puysegur is a natural experiment to constrain a key unknown, ε_{P0} , in the mechanics of initiation. In the mechanical models, we approximate the process of strain-weakening by a linear reduction in yield stress as strain accumulates until it reaches $\varepsilon = \varepsilon_{P0}$. Although, numerical solutions of deformation with such strain weakening show that weak zone width depends on model resolution (6, 27, 28), a characteristic displacement, $\Delta_s \approx 2\varepsilon_f \delta$ (with ε_f strain in the shear zone and δ shear zone width) is approximately invariant with resolution (6). Consequently, ε_{P0} represents the strain (or Δ_s , the shear zone displacement) at which a plate boundary loses most of its strength. With the plate kinematics incorporated into the models, a smaller Δ_s indicates that rock weakening occurs more rapidly, requiring less strain, while a larger Δ_s results in a slower weakening process.

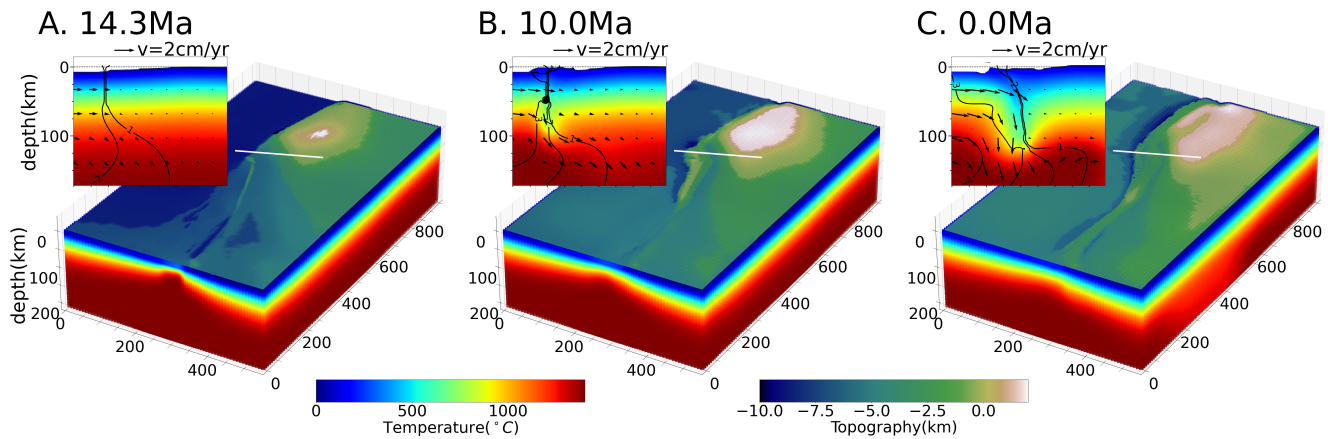


Fig. 2. Time evolution of a well-fitting case ($\varepsilon_{P0} = 0.5$, $\tau_u = 500$ MPa) at the starting (A. 14.3 Ma), middle (B. 10 Ma), and final stage (C. 0 Ma) of model evolution. The subpanels are the cross sections at the white line, where vectors are in-plane velocity, contours are velocity normal to the plane (strike-slip), and temperature is color-coded. The lower perspective views only show the upper 200 km of the computational domain.

Observationally-constrained dynamic models

The dynamics of Puysegur subduction initiation are studied in a model in which the mechanics are merged with the tight plate kinematic and structural controls while matching evolutionary outcomes (observed structure, stress and topography). The time-dependent 3D formulation starts at 15 Ma and is integrated to the present. The rheology is non-linear and uses realistic dislocation creep and plastic failure. The initial and boundary conditions are based on a Pacific-Australia reconstruction since 40 Ma using magnetic lineations and fracture zones (Fig. S1). The top of the domain is a free surface and the evolution of topography is tracked. We perform a parametric search on key rheological parameters, including the weakening strain, ε_{P0} , and maximum yield stress of the upper plate, τ_u , and evaluate the fit of models with respect to geophysical observations to find the best set of parameters. With ε_{P0} estimated from the Puysegur model, we subsequently bound the resolution invariant Δ_s .

Typically, with best-fitting models (Fig. 2, S2), the behavior is consistent with previous generic, 2-D models (29, 30), except now the outcomes are explicit in time and space and directly comparable with observations. Just after the model starts with the Australian Plate moving northward, the plate boundary is approximately a vertical strike-slip fault with some diffuse deformation in the lower lithosphere (Fig. 2, S3). With time, by 5 Myr of motion deformation continues to localize with a pronounced uplift on the Pacific Plate edge and some initial descent of the Australian Plate below the ridge in the northern section. Strain becomes distributed with some dipping thrust adjacent to the ridge. As time progresses, the slab extends downward below the ridge with a formerly uplifted ridge now subsiding, consistent with the vertical motion observed along the Puysegur Ridge (18). Starting at around 10 Myr from the beginning of the model, the stress transitions on the northern edge from growing more compressive to becoming less compressive, while the state of compression in the south shows more constant values.

By the present day, the model (Fig. 3) reproduces broad-scale topographic features characteristic of the region, including the Puysegur Trench (PT), Puysegur Ridge (PR), Puysegur Bank (PB), and Solander Basin (SB). Importantly,

a small-scale topographic low underlain by thicker crust is captured in some models (Fig. 3) and correlates with the Snares Zone (SZ), which has these same characteristics (19). The strain-partitioned fault system comprising the Puysegur Trench and the strike-slip Puysegur Ridge is a dynamic outcome of induced subduction initiation at a transpressional plate boundary (30), where a fragment of oceanic crust becomes trapped between the trench and the Puysegur fault. The computations demonstrate the transfer of strike-slip motion from the initial vertical fault onto the oblique subduction zone (Fig. 2).

We quantify how models fit the observations by either correlation (topography and stress), correctness (focal mechanisms and seismicity), or a combination of both. Although different observations give rise to different best-fitting outcomes (Fig. S4), the data fits are best for small values of the strain ($\varepsilon_{P0} < 1$) for most of the observations. One exception are the focal mechanisms, where the best fitting strain is $\varepsilon_{P0} = 1$, with some poor fits when $\varepsilon_{P0} < 1$. However, the smaller variation when fitting focal mechanisms shows that it is a weak constraint compared to the other three observations. Combining all of the measures, we obtain the best-fitting case with $[\varepsilon_{P0}, \tau_u] = [0.25, 500 \text{ MPa}]$ (Fig. 3), while case $[\varepsilon_{P0}, \tau_u] = [0.25, 300 \text{ MPa}]$ or $[0.5, 500 \text{ MPa}]$ fit the data nearly as well. The successful initiation of subduction requires that the upper plate's yield stress (τ_u) exceeds that of the subducting plate, 150 MPa. In instances where both the upper plate and the subducting plate share identical yield stresses (Fig. 3D, marked X), subduction initiation fails as the plate boundary is incapable of initiating the formation of a slab with a dipping shear zone above it. Instead, a vertical plate boundary persists, resulting in a subducting plate which is pushed horizontally beneath the upper crust (Fig. S5). This outcome is inconsistent with observed seismicity beneath Fiordland where a vertically dipping slab is observed (Fig. 1C).

Despite the variance among best-fitting models, there is a clear tendency from most of the observations to prefer a small ε_{P0} , typically $\varepsilon_{P0} < 1$. In terms of topography, a strong correlation between ε_{P0} and the width of an oceanic sliver is found, marked by the topography high between the

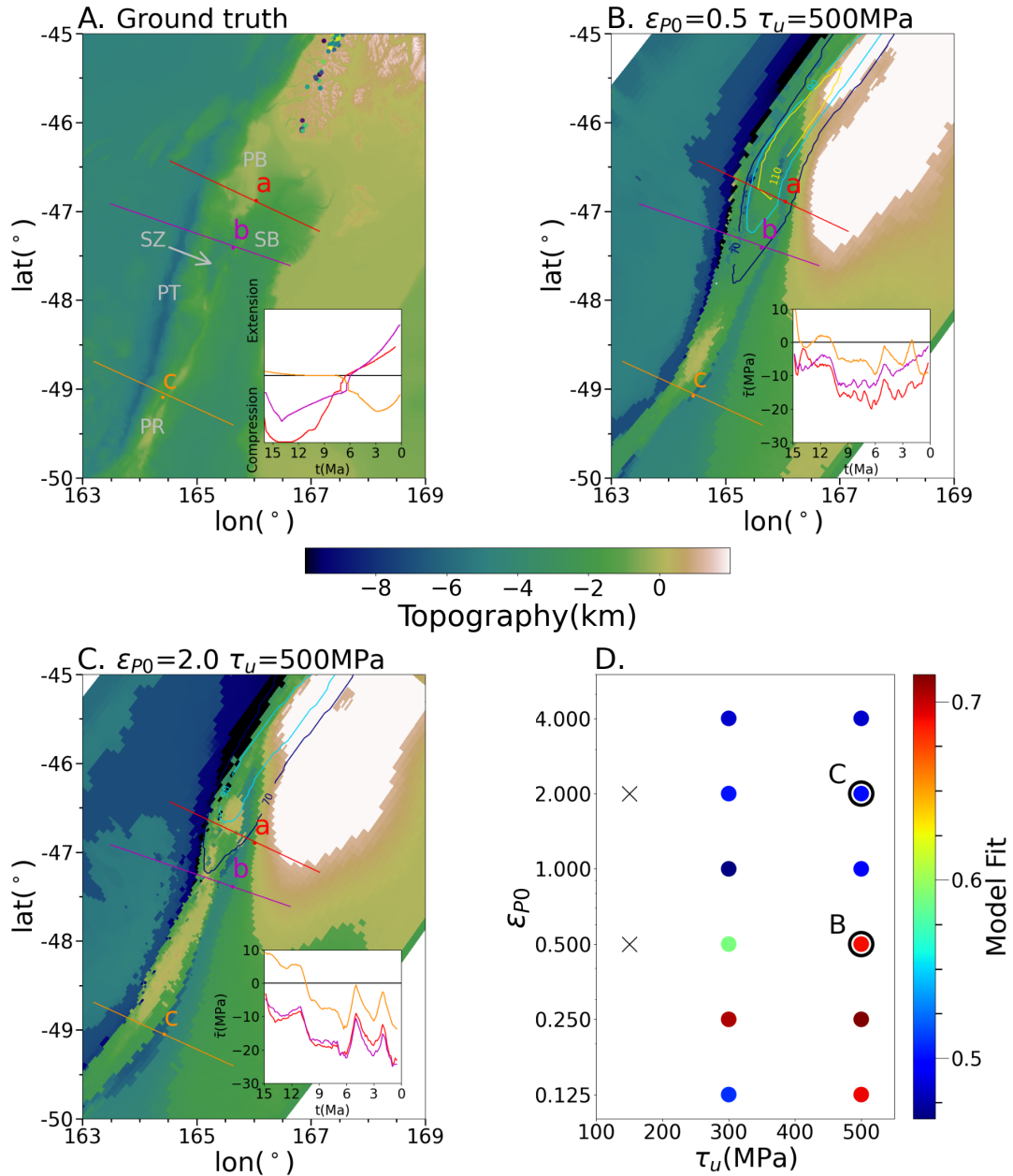


Fig. 3. Comparisons between observation and models at present day. A. Observed topography (background color), seismicity (dots color-coded with depth), and stress evolution (sub-panel) modified from (17). B. Modeled present-day bathymetry topography, depth of seismogenic zone (contours color-coded with depth using the same colormap as the seismicity in A), and measured stress evolution at a,b,c (sub-panel) from case $\epsilon_{P0} = 0.5$, $\tau_u = 500 \text{ MPa}$. C. Same as B but from case $\epsilon_{P0} = 2$, $\tau_u = 500 \text{ MPa}$. D. The total model correctness with varied ϵ_{P0} and τ_u . "X" indicates no subduction initiation. The correctness ranges from 0 to 1, with 1 being a perfect fit to observed data. PT, Puysegur Trench; PR, Puysegur Ridge; PB, Puysegur Bank; SZ, Snares zone; SB, Solander Basin. The seismicity (dots in A) and depth contours (B, C) are color-coded with depth, using the same color scale as in Fig. 1C

Solander Basin and Puysegur Trench. For instance, the width of the oceanic sliver at the latitude of seismic line EW9601 (Fig. 1B) is around 60 km (Fig. 3A), consistent with the case of $\epsilon_{P0} = 0.5$ (Fig. 3B) but exceeds the observed width when $\epsilon_{P0} = 2$ (Fig. 3C). The intermediate-depth seismicity indicates the depth to which the cold slab (defined in models as mantle $< 850^\circ \text{C}$ (31)) penetrates into the mantle. A large ϵ_{P0} (slower weakening) tends to produce a shallow slab (Fig. S6C), and only when ϵ_{P0} is as small as 0.25 do the computations give a slab deeper than 130 km (Fig. S6B). Predicted focal mechanisms (Fig. S7) and stress evolution

(subpanels of Fig. 3B,C) are both influenced by ϵ_{P0} through plate-coupling. Although models may not produce complete compression-extension reversals, cases with smaller ϵ_{P0} tend to predict a relaxation of horizontal compression since 6 Ma (Fig. 3B), while those with large ϵ_{P0} show strong compression persisting to the present day along the northern seismic lines (a and b). Fault type from seismic focal mechanisms (24) shows an overall compression-transpression stress state existing through the whole domain, while the fault type near the Snares Zone is dominantly strike-slip. Beneath Fiordland and the Puysegur Bank, there is normal faulting, potentially

indicative of a change from induced compression to pulling by the slab. Cases with larger ε_{P0} typically over-predict the extent of compression with thrust faulting throughout (Fig. S7C), while smaller ε_{P0} , i.e. faster weakening, produces transpressional to strike-slip faulting as well as an extensional stress state (normal faulting) beneath Fiordland (Fig. S7B,D). Quantitatively, with the evaluation of an angle defined by Kagan (32), we find that the best fit for earthquake focal mechanisms occurs at $\varepsilon_{P0} = 1$, in contrast to the combined observational data which suggest an optimal range of $\varepsilon_{P0} = 0.25 - 0.5$. This divergence in the focal mechanism fit can be attributed to complexities such as crustal heterogeneity and pre-existing faults that might influence shallow earthquakes. Nevertheless, the coherent pattern of fault types discussed above still demonstrates that the stress states in the models with smaller ε_{P0} align with the large-scale earthquake faulting patterns observed in the GCMT catalog, indicating that our models effectively capture the overarching background stress state.

The rheology of continental crust also influences model fit to observations. Two end-members for this rheology, one with a weak lower crust composed of wet quartzite (33) and a second with a strong lower crust composed of dry feldspar (34), are considered. All of the cases thus described have a weaker quartzitic lower crust, and we now test cases with a strong lower crust. The strong lower crust models yield equally good fits to the topography, seismicity, earthquake focal mechanisms, and stress evolution as their weak crustal counterparts (Fig. S8, S9). However, a notable discrepancy emerges: In the weaker crust models, the predicted slab position is beneath Fiordland, whereas with a stronger crust, the model slab shifts to the west of Fiordland. A weaker lower crust in the upper plate leads to more deformation of the upper plate, while a stronger lower crust in the upper plate favors subducting plate deformation. Given the observed seismicity beneath Fiordland, a weaker lower crust is preferred. Combined with the earlier discussion that a successful subduction initiation model prefers the overriding plate to be stronger than the subducting plate, we conclude that the overriding plate possesses a strong crust but with a slightly weaker lower crust. This aligns with the geology of Fiordland, where the upper crust consists of Mesozoic intrusive rocks and Cretaceous to Tertiary sediments, usually thought to be strong and brittle, while the lower crust is characterized by Precambrian granulites, typically exhibiting ductile behavior (35).

In the models, ε_{P0} influences outcomes by controlling the rate of weakening and localization within the shear zone. However, due to the finite resolution of the computations, there is a lower limit on the possible thickness of the shear zone, which may not accurately represent the true physical processes involved in shear zone weakening. Instead, the amount of slip, Δ_s defined above, proves to be a more resolution-independent measure of the weakening process (6, 27, 28). Given this, accurately determining δ becomes crucial. In the shallow part of the Puysegur subduction zone, we observe a strain-partitioned system with distributed deformation across several elements, where the shear zone thickness is about 10 km, about 4-5 elements (Fig. S34). At greater depths, strain becomes more localized onto a narrow, single fault system (Fig. S3), limited by model

resolution. Here, we estimate the thickness of the localized shear zone to be $\delta \approx 3\Delta_e$, consistent with generic models of subduction initiation with the same rheologies and weakening parameterization (6); with $\Delta_e = 2.5$ km being the element size, $\delta = 7.5$ km. Earlier generic models demonstrate that smaller ε_{P0} or smaller Δ_e lead to more rapid weakening of the plate boundary during subduction initiation (6) such that ε_{P0} represents an upper bound as meshes become more refined but the best fitting $\Delta_s \approx 2 \times 3\varepsilon_{P0}\Delta_e$ is $\approx 4 - 8$ km and approximately independent of resolution.

Bounding the physics of weakening

The observations indicate that the initiation of subduction at Puysegur is characterized by a relatively fast rate of weakening. However, the numerical models used thus far rely on the assumption that rock strength decreases linearly with accumulated strain, a simplification that may not fully capture the complexities within the lithosphere, as no specific physical mechanisms of strain weakening are assumed. Given the many cases which would need to be explored to find the optimal parameters, adding the detailed physical processes with many more uncertain parameters to the 4D mechanical/data model would make the approach computationally prohibitive. We now shift the focus to theoretical calculations to examine various physical mechanisms. We calculate their theoretical predictions of Δ_s , the total displacement for complete shear zone weakening, under Puysegur conditions and compare these predictions to the observationally constrained Δ_s , aiming to identify the mechanisms most likely responsible for the observed strain weakening.

A variety of mechanisms have been advanced to explain the progressive weakening of faults and shear zones, chief among them thermal weakening, grain size reduction, and fluid pressurization. Unfortunately, there has not been a sufficiently complete set of observations – spanning space and time – to distinguish between them in models of subduction initiation. Now, with an observationally-based bound on weakening, we determine if the parameterized strain required for weakening, predicts a consistent shear zone displacement Δ_s with the proposed processes. We test three possible physical mechanisms for strain weakening, shear heating, grain-size reduction, and pore fluid pressure, with reduced-dimensional systems as a function of time (see Materials and Methods section for details) in order to follow the reduction of strength with the underlying tectonic parameters.

Weakening from shear heating arises when viscous dissipation heats the rock and reduces its creep strength. We consider a simple shear system with an initial temperature, T_0 , that heats due to shearing. The dominant rheology experiences a transition from plastic-yielding with constant stress, to dislocation and diffusion creep with a constant strain rate, $\dot{\varepsilon}$. For a 7.5 km thick shear zone, the predicted Δ_s range from 600 to 1000 km under the range of T_0 and $\dot{\varepsilon}$ exhibited by the 4D numerical model (Fig. 4A), at least two orders of magnitude larger than the data-constrained value, $0.25 < \varepsilon_{P0} < 0.5$, corresponding to $4 \text{ km} < \Delta_s < 8 \text{ km}$. However, for a more localized shear zone with smaller thickness δ , the required displacement for weakening is substantially reduced (Fig. S10C, E), and a shear zone localized to a thickness of 100 m can produce a similar rate of weakening as we observed in Puysegur. This suggests that as subduction initiation

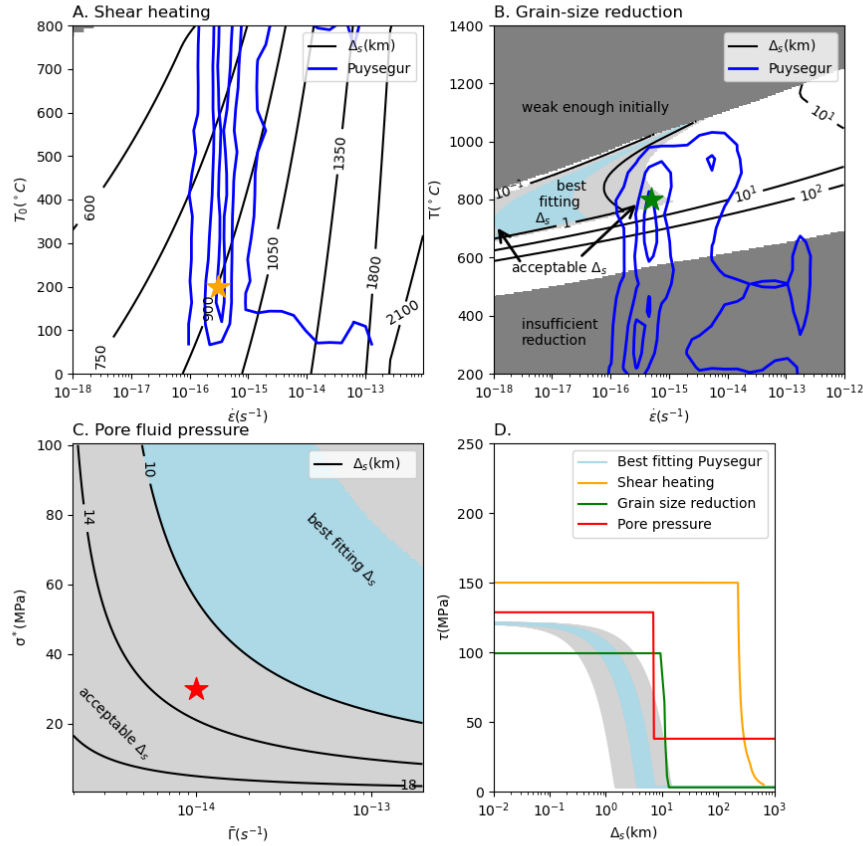


Fig. 4. Physical models for strain-weakening in the shear zone. A. Shear heating with the shear zone displacement $\Delta_s = 2\varepsilon_{P0}\delta$ (contoured in black). Blue contours represent the density distribution of T & $\dot{\varepsilon}$ in the lithosphere from the Puysegur model (beginning at 15 Ma for case $\varepsilon_{P0} = 0.5$, $\tau_u = 300$ MPa). T_0 is the initial temperature. B. Δ_s by grain-size reduction (black contours) with the initial condition $r_0 = 1$ mm. Blue contours are the same as A except from the end (0 Ma) of case $\varepsilon_{P0} = 0.5$, $\tau_u = 300$ MPa. Dark gray shaded regions in B represent the grain-size reduction is either unneeded or insufficient. C. The shear zone displacement Δ_s from pore pressure weakening; σ^* is the pressure-permeability exponential scaling factor and $\bar{\Gamma}$ the average fluid production rate in the top 10 km. D. Typical strain weakening paths for different weakening mechanisms compared with the best fitting Puysegur models. Three physical mechanisms' parameters are taken from 3 stars with corresponding colors in A, B, and C. The paths of pore pressure and best fitting numerical model sample a point at the depth of 5 km. The gray shaded regions (acceptable Δ_s) are for $0.1 \leq \varepsilon_{P0} \leq 0.5$ ($3.75\text{ km} \leq \Delta_s \leq 7.5\text{ km}$), and light blue shaded regions (best fitting Δ_s) are for $0.25 \leq \varepsilon_{P0} \leq 0.5$ ($3.75\text{ km} \leq \Delta_s \leq 7.5\text{ km}$). For shear heating and grain size reduction, we use a resolution limited shear zone thickness $\delta = 3\Delta_e = 7.5$ km. For pore fluid pressure, $\delta = 10$ km.

progresses shear heating could become more important. As the calculation assumes perfect efficiency and without heat loss, the weakening rate would only be slower (larger Δ_s) if thermal diffusion is considered. Consequently, shear heating alone is insufficient to explain the weakening needed to fit the Puysegur observations during the nucleation stage.

Grain-size reduction is invoked as a mechanism to weaken plate boundaries. When grain size decreases, diffusion creep increases and relaxes the stresses within the rock. Considering a two-phase peridotite system experiencing a grain-size evolution with Zener pinning (9), the governing rheology transitions from plastic failure/dislocation creep initially with large grain-size, to diffusion creep when grain size is reduced. In a 7.5 km thick shear zone, the predicted Δ_s varies from 10^{-1} to 10^2 km under different temperature and strain rate ($T - \dot{\varepsilon}$) conditions (Fig. 4B). As the grain size evolution is governed by two competing mechanisms,

grain growth and grain-size reduction, the grain sizes can eventually reach an equilibrium stage when the growth rate is equal to the reduction rate (Fig. S11, D-F). For a certain low-temperature region of the $T - \dot{\varepsilon}$ domain (Fig. 4B, labeled "insufficient reduction" region), the equilibrium grain size exceeds the desired grain size for the targeted stress levels, meaning that the desired grain size is unattainable in this temperature range. When the temperatures are higher, (800° to 1400°C, Fig. 4B, "weak enough initially" zone), the rock is already weak with its initial grain size, rendering any further weakening unnecessary. The choice of initial grain size affects the zone weak enough initially but has relatively little influence on the value of Δ_s (Fig. S11, A-C). The acceptable Puysegur-inferred weakening rate $\varepsilon_{P0} < 1$, equivalent to $\Delta_s < 15$ km assuming a 7.5 km thick shear zone, encompasses part of the grain-size reduction Δ_s diagram and this part corresponds to the condition of the mid-lower

lithosphere ($T > 600^\circ\text{C}$) for Puysegur. The best fitting range, however, falling within a narrower range of $4\text{ km} < \Delta_s < 8\text{ km}$, cannot be achieved with a shear zone thickness of $\delta = 7.5\text{ km}$. Reducing the thickness to $\delta = 1\text{ km}$ allows the best fit to be achieved (Fig. S10D), but a further reduction in thickness to $\delta = 100\text{ m}$ results in a weakening that is too fast (Fig. S10F). Nevertheless, regardless of the choice of δ , grain-size reduction is not ideal for achieving strain weakening within the shallow lithosphere as it is insufficient within the low-temperature domain, since under low-temperature conditions, the dislocation and diffusion creep are significantly reduced (Fig. S12).

Water is a third factor that can lead to rapid weakening. The subduction interface generally exhibits a lower yield stress compared to the crust away from the plate boundary, because fluid released from the sediment and crust along the slab interface increases the pore pressure and reduces the effective stress (36, 37). For subduction initiation, no sediment and crust pre-exist within the mantle, and the weakness introduced by the addition of fluid pore pressure can equivalently be regarded as a weakening process. Pore pressure is governed by Darcy's flow assuming permeability scales exponentially with effective pressure, where the scaling is determined by a parameter σ^* , with larger values of σ^* indicating reduced sensitivity of permeability to effective pressure (38). Additionally, a fluid production rate, Γ , significantly influences pore pressure. As such, we evaluate the shear zone displacement Δ_s , which replicates the weakening across the entire subduction interface resulting from excess pore fluid pressure due to crust and sediment dehydration, with varying σ^* and averaged production rate, $\bar{\Gamma}$ (Fig. 4C). The predicted Δ_s range from 5 to 20 km, and only with large σ^* and high fluid production rate does the weakening rate fall within the best fitting range ($\Delta_s < 10\text{ km}$ assuming $\delta = 10\text{ km}$). As the subducting Australian plate is young, about 20 Myr old at the start of initiation, the slab interface P - T path crosses the dehydration from Lawsonite breakdown at a shallow depth ($< 10\text{ km}$, Fig. S13), leading to large average dehydration rates (up to 10^{-14} s^{-1}), and in turn a Δ_s of around 12 km, close but larger than inferred for Puysegur. For Δ_s to be reduced even further, either an unrealistic low permeability ($k_0 = 10^{-20}\text{ m}^2$) or larger fluid production rate ($\bar{\Gamma} \approx 10^{-13}\text{ s}^{-1}$), or some combination of the two, are needed. In addition to the steady-state Darcy's flow previously discussed, transient mechanisms such as poroelastic compaction (39, 40), which occur on shorter time scales, might also contribute additional pore pressure and further weakening, potentially reconciling the underestimated weakening rates in our models.

Implications of the constrained weakening

The four-dimensional model has shown a significant ability to fit the different observations and provides a means to understand how plate boundaries like subduction zones nucleate and the stress, temperature, and strain rates existing at depth. Critically, we have constrained the key parameter, Δ_s , the displacement required to weaken the nascent plate boundary. From this approach, we find that a small displacement, or a small strain, essentially a fast strain weakening, is needed for Puysegur subduction initiation, with the shear zone displacement $\Delta_s \approx 4 - 8\text{ km}$. This estimation

aligns closely with findings from studies on transform fault stability, which suggest a $\varepsilon_{P0} = 1$ (at a model resolution of 0.5 km) and a $\Delta_s \approx 2 - 4\text{ km}$ (28).

Notably, although the total plate convergence during the initiation of subduction at Puysegur since 15 Ma exceeds 100 km (20), only a small fraction of that displacement, $\Delta_s \approx 4 - 8\text{ km}$, is required to weaken the nascent subduction interface. Theoretical and computational models suggest that the amount of overall convergence needed to transition to fully self-sustaining subduction is $100\text{--}150\text{ km}$ (5, 6). The $\approx 100\text{ km}$ slab length found in the best fitting models (Fig. S6) and the $\sim 1.8\text{ km}$ subsidence of the northern part of the Puysegur Ridge (Snare Zones) underlain by thickened crust (19) are all consistent with the transition to a self-sustaining subduction zone. The discrepancy between the total convergence and the small fraction needed to weaken the margin suggests that a substantial amount of the work the converging plate does at the plate boundary does not simply go into making a favorably orientated megathrust. Some of the work goes into distributed deformation across the nascent boundary and creating topography. Some work may also have gone into thickening the crust, as we do know the crust is thickened at the Snare zone (19) (although some of that thickened crust maybe due to preexisting blocks of continental crust associated with earlier Solander Basin rifting (22, 41)).

Three possible mechanisms for strain weakening, shear heating, grain size reduction, and fluid pore pressure, have been explored in light of this new constraint on weakening. We find shear heating to be too slow unless the shear zone is very localized which would be well after the nucleation stage of subduction initiation. Grain size reduction might be fast enough but only works within the lower lithosphere where the temperatures are sufficiently high. Fluid pore pressure can produce a Δ_s close to the estimated weakening rate, but a perfect match requires a faster fluid production rate. Puysegur Trench initiated at a vertical strike-slip boundary, and the subsequent development of weak zones varies with depth. At depths less than 10 km , pore pressure weakening could have played a significant role in the formation of the strain partitioning system. Deeper within the lithosphere, grain-size reduction may have been associated with the rotation of the presumed initially vertical strike-slip fault. However, it is crucial to note that the existence of a pre-existing fault that could be reused is also indispensable for the formation of the new Puysegur subduction zone. The interplay of these factors, including fast weakening mechanisms like pore pressure and grain-size reduction, and the availability of pre-existing fault structures (even those that are not optimally oriented), contributed to the complex process leading to the initiation and evolution of subduction zones.

The well-constrained fast strain weakening of Puysegur initiation offers new insight into a fundamental question surrounding the origin of plate tectonics: How subduction first initiates in the inception of plate tectonics on early Earth. Previously proposed triggering mechanisms that start the initial cycle of plate tectonics involve external processes, like plumes (2), meteorite impacts (42), the Moon-forming giant impact (43) or biologically-induced sedimentation (44), that drives deformation and localization of thermal and compositional variations – essential prerequisites for

subduction initiation. Now, our findings reveal that an intrinsic property of the lithosphere, the weakening rate, is fast during subduction initiation. We propose that a combination of two physical mechanisms, grain-size reduction and fluid pressurization, can explain the fast weakening, which opens intriguing possibilities for similar processes occurring on early Earth. Although the computations are tailored to Puysegur subduction, the fundamental principles underlying our theoretical calculations of the physical weakening models transcend specific geological contexts, including scenarios that may have been prevalent during the initiation of subduction on the early Earth. The grain-size reduction is an intrinsic weakening process under high-temperature conditions where diffusion creep dominates (9). During grain size reduction, an initial phase of grain growth occurs as interface roughness gradually decreases, followed by a sudden decrease in grain size when roughness reduces to a critical level (Fig. S11D-F). This pattern suggests that rocks might undergo a prolonged period of stability without significant rheological changes until the accumulated reduction in interface roughness triggers a rapid weakening. This process, typically observed under high-temperature conditions such as the lower lithosphere or under elevated surface temperatures, could also be relevant to the early Earth. Moreover, the long-term horizontal tectonic forces, like ridge push, might create some vertical weak bands through grain damage on a time scale of 100 Myr, leading to anisotropic weakness that favors vertical displacement, thereby facilitate the subduction initiation at the passive margin (45). Fluid pressurization remains the other potential mechanism for the formation of early subduction zones since oceans may have existed on early Earth (46). However, unlike present-day induced subduction initiation, the onset of the very first subduction zone may not have been driven by plate convergence. Consequently, the driving mechanism responsible for transporting water into the deep mantle remains unclear. An alternative mechanism is water percolation through weak zones caused by thermal cracking (47). Despite numerous uncertainties, the early Earth could share common characteristics with the present day: Fluid pressurization at shallow lithospheric depths and grain-size reduction at greater depths jointly governing a relatively fast weakening process that facilitates subduction initiation while shaping the foundations of plate tectonics.

Supporting Information Appendix (SI). Supplementary material, including supplementary figures, and tables, see the supplementary material document. The data and additional figures underlying this article are available in CaltechDATA, at <https://doi.org/10.22002/jjng3-qv546>.

Materials and Methods

A. 4D computational models. We perform computations in a 3D Cartesian domain as a function of time using the finite element method with the validated software *Underworld2* (48). The approach solves the continuity and momentum (the Stokes system) and energy equations. The computations start at 15 Ma in a Cartesian domain, with 540 km $\times$ 900 km horizontally (as black outline in Fig. 1A) and 450 km vertically, including 40 km sticky-air layer which mimics a free surface. The bottom of the domain is at a mantle depth of 410 km. The finest resolution near the trench is around 2.5 km/element (with linear elements for the velocity).

A particle-in-cell technique is implemented to trace the material composition and plastic strain.

The initial and boundary conditions are based on a plate reconstruction starting from 40 Ma (Fig. S1) in *pyGPlates* (49). Since 40 Ma, the Australia-Pacific plate boundary experienced a transition from a spreading center, to strike-slip motion, and eventually to subduction, so that the majority of the subducting plate is younger than 40 Ma. We reconstruct the spreading center-transform fault plate boundary system according to the magnetic lineations and fracture zones (Fig. 1A), and calculate the explicit, evolving age of the plates using a tracer-based algorithm, *TracTec* (50) implemented using *pyGPlates*. Based on the evaluated plate age model at 15 Ma, we construct the initial thermal field of the geodynamic model following half-space cooling. Crustal thickness for the upper plate (Pacific) is interpolated from crust model *CRUST 1.0* (51). Transformation from geographical to Cartesian coordinates uses a Lambert equal area projection.

The boundary conditions of the geodynamic models are based on the AUS-PAC plate finite rotation (14–16, 52). The bottom boundary is no-slip, and the top boundary is free-slip so that with a sticky air layer forms a free surface on top of the plates and allows the tracking of surface topography. The plate kinematics from the plate reconstruction is incorporated into the model through the four vertical side boundaries, and the velocity field is composed of two terms. In the first term, the top 100 km mantle strictly follows the plate velocity of finite rotation, and between 100 km and 410 km depth the velocity drops linearly from the plate velocity to 0. In order to conserve mass, a second term is added as channel flow, $V_r(z) = S_0 \frac{(z-z_0)(z-z_1)}{r}$ where the depths z_0 and z_1 are 100 km and 410 km, r the radial distance from the sink (trench), S_0 is a constant that guarantees that the total flux out equals the influx.

In the numerical model, we implement a visco-plastic rheology with strain weakening to mimic the realistic mantle rheology (53). At shallow depths with low temperatures and pressures, the viscosity is governed by the Drucker-Prager yield stress. At greater depths with high temperatures, the dislocation creep is fast enough to dominate the flow law. This combination of rheologies leads to a profile where stress initially increases with depth, reaching a peak before decreasing at greater depths. Such combined rheology creates a scenario where the middle depths, transitioning from brittle to ductile behaviors, represent the strongest part of the plate. In the middle depths, we assign an upper limit of the yield stress which defines the strength of the core of the bending slab or the upper plate. For the continental plate, we treat the upper limit of the yield stress as a variable that defines the overall strength of the plate, while keeping the subducting plate relatively weak ($\tau_{max} = 150$ MPa) (54). The effective viscosity is governed by either dislocation creep or plastic-yielding

$$\eta_{eff} = \min \left(\frac{1}{2} A_{disl}^{\frac{1}{n}} e^{\frac{E_I}{nRT}} \dot{\epsilon}_{II}^{\frac{n}{n-1}}, \frac{\tau_y}{\dot{\epsilon}_{II}} \right) \quad [1]$$

Where E_I , A_{disl} , and n are activation energy, prefactor and non-Newtonian exponent of dislocation creep. $\dot{\epsilon}_{II}$ is the square root of the second invariant of strain rate, and τ_y the yield stress, defined by the Drucker-Prager failure criterion with strain weakening

$$\tau_y = \max \left(\min \left(\left(1 - \frac{\epsilon_P}{\epsilon_{P0}} \right) (\mu P + C_0), \tau_{max} \right), \tau_{min} \right) \quad [2]$$

Where μ and C_0 are the friction coefficient and cohesion. ϵ_P is the accumulated plastic strain and ϵ_{P0} the reference plastic strain. τ_{min} and τ_{max} the lower and upper bound of yield stress. This definition of τ_y approximates the strain weakening with a linear process until it saturates ($\epsilon_P = \epsilon_{P0}$). A larger ϵ_{P0} means a slower weakening and vice versa.

The rheology is also composition-dependent. We mainly test two end-members for the upper plate: A model whose crust is made out of weak quartzite, and one with strong dry feldspar (55). The two end-member rheology models are incorporated into the model with different crustal rheological parameters in Table S2. The two key parameters varied are weakening rate ϵ_{P0} and the τ_{max} of the upper plate τ_u . The subducted oceanic crust follows the metamorphism of mid-ocean basalt (MORB) and eventually transforms to dense eclogite (56). For key parameters, see Table S2

B. Model evaluation. Outcomes of the geodynamic model are compared against different geophysical and geological observations quantitatively. The four major observations are the bathymetry and topography, shallow earthquake focal mechanisms, intermediate depth seismicity as an inference of slab morphology, and stress evolution from stratigraphy.

In the model, the topography is derived from the interface between the mantle (crust) and sticky air, and we evaluate the similarity between model and measured topography with the correlation

$$C_{topo} = \frac{\int_{\Gamma_t} (P_m - \bar{P}_m)(P_o - \bar{P}_o) dxdy}{\left(\int_{\Gamma_t} (P_m - \bar{P}_m)^2 dxdy \right) \left(\int_{\Gamma_t} (P_o - \bar{P}_o)^2 dxdy \right)} \quad [3]$$

Where P_m and P_o are the topography from model and observation. Γ_t is the trench area, defined as the area within 100 km from the Puysegur Trench. $\bar{P}_m$ and $\bar{P}_o$ are the average model and observed topography in the area of Γ_t . The C_{topo} describes the similarity between model topography and observed topography in phase but not in amplitude. A perfect in-phase topography gives $C_{topo} = 1$ while a perfect anti-phase topography gives $C_{topo} = -1$.

The stress evolution is derived from seismic profiles across the Puysegur Trench (17), which describe the time evolution of stress polarity at different locations of the trench, $\tau_i^P(t)$ where $i = 1, 2, 3$ for line a, b, c and $\tau P_i(t) = -1$ for compression and $\tau_i^P(t) = 1$ for extension. In the model, we directly measure the stresses in the direction of the three seismic profiles $\tau_i(t)$. The similarity between modeled stress and observation is

$$C_{stress} = \frac{1}{3} \sum_1^3 \frac{\int_{15Ma}^0 \tau_i(t) \tau_i^P(t) dt}{\left(\int_{15Ma}^0 \tau_i(t)^2 dt \right) \left(\int_{15Ma}^0 \tau_i^P(t)^2 dt \right)} \quad [4]$$

Note the C_{stress} doesn't need to demean, as $\tau = 0$ has an absolute physical meaning of neutral stress.

The intermediate depth seismicity provides an indication of the spatial distribution of the cold slab (Benioff zone). We define C_{seis} as the proportion of observed seismicity that falls within the area with a temperature $\leq 850^\circ C$ (31). The evaluation of C_{seis} is weighted by the reciprocal of the spatial density of seismicity, given by the expression

$$C_{seis} = \frac{\sum_{T_i < 850} 1/\nu_i}{\sum_i 1/\nu_i} \quad [5]$$

Where ν_i denotes the number of earthquakes in the vicinity of earthquake i . This quantifies the proportion of the observed intermediate-depth earthquakes that correctly fall into the seismogenic zone predicted by the geodynamic models. The ν_i in eq 5 ensures the evaluation is independent of the spatial distribution of the earthquake in the catalog, such that the zone with sparser earthquakes is treated equally as the zone with denser earthquakes.

The shallow earthquake focal mechanism provides a direct inference of the stress state in the present day. We take all Mw > 5 earthquakes shallower than 30 km from GCMT catalog (24). In the geodynamic model, we evaluate the focal mechanism from the stress tensor measured at the same hypocenter location as the GCMT earthquakes. To assess the alignment of moment tensors between the GCMT catalog and our model prediction, we evaluate the Kagan angle (Φ) (32) between them. A Kagan angle ranges from 0 to 120° , with smaller values indicating a better match between observations (GCMT) and model predictions. Similar to the seismicity, we define the correctness of focal mechanism prediction C_{focal} as the volumetric average of Kagan angles, weighted by the spatial density distribution of earthquakes:

$$C_{focal} = (120^\circ - \sum_i \frac{\Phi_i \nu_i}{1/\nu_i}) / 120^\circ \quad [6]$$

Where Φ_i and ν_i are the Kagan angle and spatial density of earthquake i . The C_{focal} is normalized to the range [0,1], where a higher C_{focal} value indicates a more accurate model prediction.

As the C_{topo} and C_{stress} are defined as correlation ranging from -1 to 1, while C_{seis} and C_{focal} are defined as correctness

ranging from 0 to 1, we define the total model correctness C as the average of the four quantities all normalized to [0,1]

$$C = ((C_{topo} + 1)/2 + (C_{stress} + 1)/2 + C_{seis} + C_{focal}) / 4 \quad [7]$$

C. Physical models for plate boundary weakening.

C.1. Shear heating. Shear heating is one potential strain-weakening mechanism, with the rock becoming weaker as the temperature increases with deformation. The 0-dimensional heat equation ignoring transport yields

$$\frac{dT}{dt} = \chi \frac{\tau \dot{\epsilon}}{\rho_r C_P} \quad [8]$$

Where χ is the efficiency of shear heating ranging from 0 to 1 (57). We take $\chi = 1$, providing an upper limit of shear heating, such that

$$\frac{\rho_r C_P dT}{\tau} = \dot{\epsilon} dt \quad [9]$$

In eq 8, free variables include temperature T , strain rate $\dot{\epsilon}$, and stress τ , but $\dot{\epsilon}$ and τ are related through the constitutive relations. Therefore, in addition to the initial temperature, one more degree of freedom lies on either $\dot{\epsilon}$ or τ . In the following discussion, we will discuss the situation where either stress or strain is the free parameter(fixed), and finally combine the scenarios of fixed stress and fixed strain based on different rheology laws. Through integrating eq 9, the target is to obtain the total strain required for a system to reach the ultimate weakness, defined by the final stress τ_f . Here, we choose the final stress $\tau_f = 3\text{MPa}$, the same as the τ_f in the numerical models.

For a fixed stress, τ , the reference strain ε_{P0} describes the total strain for complete weakening. For shear heating, we define ε_{P0} as the point when rock heats to T_1

$$\varepsilon_{P0}(T_0, T_1) = \int_0^{T_1} \dot{\epsilon} dt = \int_{T_0}^{T_1} \frac{\rho_r C_P}{\tau} dT = \frac{\rho_r C_P (T_1 - T_0)}{\tau} \quad [10]$$

For the fixed strain rate $\dot{\epsilon}$, we follow the definition of ε_{P0} that rock is heated to T_1 through shear heating, but dislocation and diffusion creep need to be considered. Diffusion creep is:

$$\dot{\epsilon}_{diff} = A_{diff} e^{-\frac{E_f}{RT}} h^{-m} \tau \quad [11]$$

Where E_f , R , m and A_{diff} are the activation energy, ideal gas constant, grain-size exponent, and prefactor of diffusion creep. Dislocation creep is

$$\dot{\epsilon}_{disl} = A_{disl} e^{-\frac{E_l}{RT}} \tau^n \quad [12]$$

For the composite rheology, the maximum of these two mechanisms is taken as the dominant creep mechanism

$$\dot{\epsilon} = \max(\dot{\epsilon}_{disl}, \dot{\epsilon}_{diff}) \quad [13]$$

As the $\dot{\epsilon}$ is fixed, eq. 13 is equivalent to

$$\tau = \min \left(\frac{\dot{\epsilon} h^m}{A_{diff} e^{-\frac{E_f}{RT}}}, \left(\frac{\dot{\epsilon}}{A_{disl} e^{-\frac{E_l}{RT}}} \right)^{\frac{1}{n}} \right) \quad [14]$$

Therefore, we obtain the dislocation-diffusion transition temperature T_T

$$T_T = -\frac{nE_f - E_l}{R \ln \frac{\dot{\epsilon}^{n-1} h^{mn} A_{disl}}{A_{diff}^n}} \quad [15]$$

When $T < T_T$, $\dot{\epsilon} = \dot{\epsilon}_{disl}$ and when $T > T_T$, $\dot{\epsilon} = \dot{\epsilon}_{diff}$, such that

$$\begin{aligned}
\dot{\varepsilon}_{P0}(T_0, T_1) &= \int_0^{t_1} \dot{\varepsilon} dt = \int_{T_0}^{T_1} \frac{\rho_r C_P}{\tau} dT \\
&= \int_{T_0}^{T_C} \rho_r C_P \left(\frac{A_{disl} e^{-\frac{E_l}{RT}}}{\dot{\varepsilon}} \right)^{1/n} dT + \int_{T_C}^{T_1} \rho_r C_P \frac{A_{diff} e^{-\frac{E_f}{RT}}}{\dot{\varepsilon} h^m} dT \\
&= \frac{\rho_r C_P E_l}{nR} \left(\frac{A_{disl}}{\dot{\varepsilon}} \right)^{1/n} \left(x e^{-\frac{1}{x}} + \text{Ei} \left(-\frac{1}{x} \right) \right) \Big|_{nRT_0/E_l}^{nRT_C/E_l} \\
&\quad + \frac{\rho_r C_P A_{diff} E_f}{\dot{\varepsilon} R h^m} \left(x e^{-\frac{1}{x}} + \text{Ei} \left(-\frac{1}{x} \right) \right) \Big|_{RT_C/E_f}^{RT_1/E_f}
\end{aligned} \tag{16}$$

Where $T_C = \max(T_0, \min(T_T, T_1))$. Ei is the exponential integral $\text{Ei}(x) = \int_{-\infty}^x \frac{e^t}{t} dt$.

The realistic mantle rheology follows neither fixed stress nor fixed strain rate but can be represented by a combination of fixed stress (low-temperature plasticity) and fixed strain rate (dislocation/diffusion creep). With that, eq. 14 is modified as

$$\tau = \min \left(\frac{\dot{\varepsilon} h^m}{A_{diff} e^{-\frac{E_f}{RT}}}, \left(\frac{\dot{\varepsilon}}{A_{disl} e^{-\frac{E_l}{RT}}} \right)^{\frac{1}{n}}, \tau_y \right) \tag{17}$$

Similar to eq. 15, we define T_{C1} and T_{C2} the transition temperature of $\dot{\varepsilon}_{disl} = \dot{\varepsilon}_y$ and $\dot{\varepsilon}_{diff} = \dot{\varepsilon}_y$.

$$T_{C1} = \max \left(T_0, \min \left(\frac{E_l}{R \ln \frac{\tau_y A_{disl}}{\dot{\varepsilon}}}, T_1 \right) \right) \tag{18}$$

$$T_{C2} = \max \left(T_0, \min \left(\frac{E_f}{R \ln \frac{\tau_y A_{diff}}{\dot{\varepsilon} h^m}}, T_1 \right) \right) \tag{19}$$

$$\begin{aligned}
\text{When } T_{C1} < T_{C2}, T_0 &\xrightarrow{\text{plas}} T_{C1} \xrightarrow{\text{disl}} T_C \xrightarrow{\text{diff}} T_1 \\
\varepsilon_{P0} &= \varepsilon_{P0}^{\tau_y}(T_0, T_{C1}) + \varepsilon_{P0}^{\dot{\varepsilon}}(T_{C1}, T_1)
\end{aligned} \tag{20}$$

$$\begin{aligned}
\text{When } T_{C1} > T_{C2}, T_0 &\xrightarrow{\text{plas}} T_{C2} \xrightarrow{\text{diff}} T_1. \\
\varepsilon_{P0} &= \varepsilon_{P0}^{\tau_y}(T_0, T_{C2}) + \varepsilon_{P0}^{\dot{\varepsilon}}(T_{C2}, T_1)
\end{aligned} \tag{21}$$

We define T_1 as the temperature when stress reaches the targeted final stress $\tau_f = 3$ MPa in diffusion creep regime, therefore $T_1 = \frac{E_f}{R \ln \frac{A_{diff} \tau_f}{\dot{\varepsilon} h^m}}$.

The inclusion of heat transport (such as diffusive cooling) would only slow the weakening, making the estimate conservative.

In a shear zone, it's easy to transfer the strain ε into the shear displacement Δ_s with the relationship $\Delta_s = 2\varepsilon\delta$, with the δ being the thickness of the shear zone. Therefore, the total displacement for shear heating weakening is $\Delta_s = 2\varepsilon_{P0}\delta$.

C.2. Grain-size reduction. Grain-size reduction is a physical process for strain weakening. The grain size, R , influences the rheology through diffusion creep (eq. 11). We follow a detailed treatment (9, 58, 59) in which the grain size evolution of peridotite (40% Pyroxene, $\phi_1 = 0.4$, and 60% Olivine, $\phi_2 = 0.6$, where ϕ_i is the partitioning of phase i) is governed by two-phase Zener pinning (a model of surface tension-like forces at the boundary between two phases). Three equations govern the evolution of the mean grain size for the two phases, R_i , and the interface roughness r :

$$\frac{dR_i}{dt} = \frac{G_i}{pR_i^{p-1}} Z_i - \lambda \frac{R_i^2}{3\gamma_i} f_G \Psi_i Z_i^{-1} \tag{22}$$

$$\frac{dr}{dt} = \frac{\eta G_I}{qr^{q-1}} - \frac{f_I r^2}{\gamma_I \eta} \Psi \tag{23}$$

In the equations, the first and second terms govern the coarsening (grain growth) and grain damage (grain size reduction), respectively. Under earth-like conditions, the grain growth exponent $p=2$, $q=4$,

and the grain growth rate G_i and interface coarsening coefficient G_I follows (60, 61)

$$G_I = 2(\mu\text{m})^2 G_i / 250 = 2(\mu\text{m})^2 (k_0 e^{-\frac{E_g}{RT}}) / 250 \tag{24}$$

Where $k_0 = 2 \times 10^4 (\mu\text{m})^p \text{ s}^{-1}$ is the kinetic factor and E_g is an activation energy for the grain growth.

The grain-size reduction terms are associated with the work done by dislocation creep and diffusion creep in each phase:

$$\Psi_i = \tau_i : (\dot{\varepsilon}_{disl,i} + \dot{\varepsilon}_{diff,i}) \tag{25}$$

The reduction rate of roughness, r , is related to the total work

$$\Psi = \sum_i \Psi_i \tag{26}$$

f_I is the partitioning factor that evaluate the fraction of damage energy that turns into the work creating interface area (60, 62), approximated by

$$f_I \approx f = f_0 e^{-2((T+273)/1000)^{2.9}} \tag{27}$$

Where $f_0 = 10^{-3}$. f_G accounts for the energy partition into creating new grain boundaries, driven solely by dislocation creep (62), such that

$$f_G \Psi_i = f \tau_i : \dot{\varepsilon}_{disl,i} \tag{28}$$

Finally, the grain-size evolution involves the Zenor pinning factor Z_i in both grain growth and grain size reduction terms, with

$$Z_i = 1 - c(1 - \phi_i) \frac{R_i^2}{r^2} \tag{29}$$

Where $c = 0.87$.

Other parameters include surface tension $\gamma_i \approx \gamma \approx 1 \text{ J} \cdot \text{m}^{-1}$, $\eta \approx 3\phi_1\phi_2$, $\lambda = 4.95$ (62).

Similar to shear heating, we consider a rheology combining dislocation creep, diffusion creep and plasticity. With an assigned strain rate $\dot{\varepsilon}$, the stress in each phase is determined by the weakest rheology (eq. 17), and the diffusion creep and dislocation creep following the same relation as eq. 12 and 11.

For simplicity, we assume the two phases share similar rheological parameters (A_{disl} , E_{disl} , A_{diff} , E_{diff}), but the different grain sizes R_i create distinct diffusion creep rate in the different phases. The total stress is evaluated

$$\bar{\tau} = \min \left(\frac{\dot{\varepsilon} \bar{R}^m}{A_{diff} e^{-\frac{E_f}{RT}}}, \left(\frac{\dot{\varepsilon}}{A_{disl} e^{-\frac{E_l}{RT}}} \right)^{\frac{1}{n}}, \tau_y \right) \tag{30}$$

With $\bar{R} = \sum_i R_i \phi_i$ being the mean grain size.

Beginning with an initial grain size of $R_1 = R_2 = 2r = 2$ mm, we integrate eq. 22 and 23 in time using the Radau method until the system reaches the targeted weakness, i.e. $\bar{\tau} = \tau_f$, under different temperatures, T , and strain rate $\dot{\varepsilon}$. For each strain rates, $\dot{\varepsilon}$, the total strain for a desired weakness is simply

$$\varepsilon_{P0} = \int_{t=0}^{t_1} \dot{\varepsilon} dt = \dot{\varepsilon} t_1 \tag{31}$$

Where t_1 is time when the condition $\bar{\tau} = \tau_f$ is achieved due to grain size reduction.

In eq. 25 and 26, there is an implicit assumption that only diffusion creep and dislocation creep are involved in the grain damage process. However, eq. 30 informs us that strain can be partitioned into three components: dislocation creep, diffusion creep, and plasticity. While dislocation and diffusion creep are much more active under high-temperature conditions (eq. 11, 12), plasticity predominates in regimes with lower temperatures. Plasticity is an indispensable process in the subduction zone dynamics, which accounts for the failure of the rock through processes like faulting when the stress in the cold lithosphere reaches the yielding limit. Consequently, plastic rheology primarily reflects the sporadic yet significant events, such as earthquakes, rather than the gradual mechanisms like creep. Thus, it is unlikely that plasticity significantly influences the overall grain size in the lithosphere. In light of this, along with Eq. 31, we propose an

alternative definition for strain, namely the effective weakening strain, denoted as $\varepsilon_{P0,eff}$:

$$\varepsilon_{P0,eff} = \int_{t=0}^{t_1} \dot{\varepsilon}_{eff} dt \quad [32]$$

Where $\dot{\varepsilon}_{eff}$ is the effective strain rate composite of dislocation creep and dislocation creep, but not the plastic strain rate:

$$\dot{\varepsilon}_{eff} = \max \left(A_{dist} e^{-\frac{E_l}{RT}} \bar{\tau}^n, A_{diff} e^{-\frac{E_f}{RT}} R_i^{-m} \bar{\tau} \right) \quad [33]$$

The geodynamic models always account for the total strain rate, therefore better reflecting the definition of eq. 31, but we also provide the evaluation of $\varepsilon_{P0,eff}$ (Fig. S12), which better reflects the rate of weakening from the microscopic process. The difference between ε_{P0} and $\varepsilon_{P0,eff}$ is prominent in the low temperature domain ($T < 600^\circ \text{C}$), where plasticity becomes dominant. In both definition, only the lower lithosphere condition (600 to 1000°C) predict a sufficiently fast weakening comparable to the Puysegur-inferred weakening ($\varepsilon_{P0} < 1$).

In addition to the temperature and strain rate, the choice of the initial grain size is the other dimension that impacts the weakening rate of the grain-size reduction system. The reference case uses $r_0 = 1 \text{ mm}$ for the mantle (Fig. 4.C), and we explore the cases of $r_0 = 0.1 \text{ mm}$ and $r_0 = 10 \text{ mm}$ cases (Fig. S11) and maintain $R_1 = R_2 = 2r_0$ initially. The initial grain size affect the boundary between the initially weak enough zone and the weakening regime, but within the regime of weakening, the choice of r_0 does not significantly influence the ε_{P0} .

Similar to shear heating, in a shear zone with the thickness of δ , the total displacement for grain size weakening is $\Delta_s = 2\varepsilon_{P0}\delta$.

C.3. Fluid pore pressure. Fluid is another potential source of plate boundary weakening, as the increased fluid pore pressure reduces the effective normal stress and yield stress (Fig. S14).

$$\tau_y = C + \mu(P - P_f) \quad [34]$$

Where τ_y the yield stress, C and μ are cohesion and coefficient of friction, P and P_f are rock and pore fluid pressure. The pore pressure follows steady state Darcy's law:

$$\nabla \left(\frac{k\rho}{\eta_w} (\nabla P_f - \rho \vec{g}) \right) + \rho \Gamma = 0 \quad [35]$$

Where k the permeability, ρ fluid density, η_w fluid viscosity, and Γ fluid production rate. For water, we can use constant density and viscosity $\rho = 1000 \text{ kg/m}^3$, $\eta_w = 10^{-4} \text{ Pa s}$, while the permeability is substantially dependent on pressure (38, 63–66)

$$k = k_0 e^{-\frac{\rho_r g z - P_f}{\sigma^*}} \quad [36]$$

where k_0 is the permeability at zero pressure, ρ_r the rock density and σ^* describes how permeability is sensitive to the change of pressure. The definition of k leads to

$$\nabla k = -\frac{k(\rho \vec{g} - \nabla P_f)}{\sigma^*} \quad [37]$$

$$\nabla^2 k = -\frac{1}{\sigma^*} (\nabla k \cdot \rho_r \vec{g} - \nabla k \cdot \nabla P_f - k \nabla^2 P_f) \quad [38]$$

With 35, 37, 38 yields a linear PDE of permeability k

$$\begin{aligned} \nabla(k(\nabla P_f - \rho \vec{g})) &= \nabla k \cdot \nabla P_f + k \nabla^2 P_f - \rho \nabla k \cdot \vec{g} \\ &= \sigma^* \nabla^2 k + (\rho_r - \rho) \nabla k \cdot \vec{g} = -\eta_w \Gamma \end{aligned} \quad [39]$$

As gravity $\vec{g}$ is only in vertical direction, z , horizontally k follows a diffusion equation, while vertically k is driven by the buoyancy term $(\rho_r - \rho)$, such that the vertical transport is much faster than the horizontal. Therefore, we collapse eq. 39 into an ODE in z direction, assuming a point source for each column

$$k'' + \frac{(\rho_r - \rho)g}{\sigma^*} k' = -\frac{\eta_w \Gamma \delta(z - z_0)}{\sigma^*} \quad [40]$$

Where δ is the Dirac delta function, z_0 is the location of the fluid source and Γ is the fluid production rate collapsing the total

volumetric fluid release rate of the crust and sediment layer. In our case, z_0 is the total depth of sediment and crust of the mega-thrust.

With the boundary condition $k|_{z=0} = k_0$, $k|_{z=\infty} = 0$, we can solve eq. 40 with Laplace transform.

$$\begin{aligned} k &= \frac{\eta_w \Gamma}{(\rho_r - \rho)g} - \frac{\eta_w \Gamma - (\rho_r - \rho)g k_0}{(\rho_r - \rho)g} e^{-\frac{(\rho_r - \rho)g}{\sigma^*} z} \\ &+ \frac{\eta_w \Gamma}{(\rho_r - \rho)g} u(z - z_0) \left(e^{\frac{(\rho_r - \rho)g}{\sigma^*} (z_0 - z)} - 1 \right) \end{aligned} \quad [41]$$

Where u is a heavy-side function.

With eq. 36 we can evaluated the pore pressure from the resulting k in eq. 41

$$P_f(z) = \sigma^* \ln(k/k_0) + \rho_r g z \quad [42]$$

Within the subduction zones interface where fluid is released from sediment and oceanic crust ($z = z_0$), the pore pressure P_f can be evaluated from eq. 36 and eq. 41:

$$P_f(z_0) = \rho_r g z_0 + \sigma^* \ln \left(\frac{\eta_w \Gamma}{(\rho_r - \rho)g K_0} (1 - e^{-\frac{(\rho_r - \rho)g z_0}{\sigma^*}}) + e^{-\frac{(\rho_r - \rho)g z_0}{\sigma^*}} \right) \quad [43]$$

With no fluid source ($\Gamma = 0$), $k = k_0 e^{-\frac{(\rho_r - \rho)g}{\sigma^*} z}$ is the background permeability, with eq. 36 we obtain background pore pressure $P_f = \rho_r g z$. Therefore the change of pore pressure at the subduction zone interface due to the fluid release is

$$\Delta P_f = P_f - \rho_r g z_0 = \sigma^* \ln \left(\frac{\eta_w \Gamma}{(\rho_r - \rho)g K_0} \left(e^{\frac{(\rho_r - \rho)g}{\sigma^*} z_0} - 1 \right) + 1 \right) \quad [44]$$

The yield stress reduction due to the change of fluid pore pressure is

$$\Delta \tau_y = -\mu \Delta P_f \quad [45]$$

In subduction zones, the fluid is released from the sedimentary layer and oceanic crust through porosity loss from compaction and clay dehydration at shallow depths (top 5 km), and hydrous mineral breakdown at deeper depth (37). We assign the fluid production rate following (36) assuming a 1 km thick sediment wedge, the value measured seismically (22), and 7km mid-ocean basalt.

Prior to subduction initiation, fluid production starts from $\Gamma = 0$, and turns non-zero ($\Gamma = \Gamma(z_0)$) as subduction develops once sediment reaches the depth of z_0 with the subduction channel. Within a simple shear channel flow, the total plastic strain for sediment to reach the depth of z_0 is $\varepsilon_P = \frac{z_0}{2\delta \sin \theta}$, with δ the channel width and θ the dip angle. In the Puysegur geodynamic models, the yielding stress of the shallow depth is governed by the Drucker–Prager failure criterion with a linear strain weakening

$$\tau_y = \left(1 - \frac{\varepsilon_P}{\varepsilon_{P0}} \right) (\mu P_r + C_0) \quad [46]$$

Where μ and C_0 are the initial friction coefficient and cohesion, and $\varepsilon_P < \varepsilon_{P0}$ as strain weakening isn't saturated at shallow depth. P_r is the rock pressure, and can be approximated with $P_r = \rho_r g z_0$.

Compared to the initial stress ($\varepsilon_P = 0$), the change of yielding stress in the model is

$$\Delta \tau_y = -\frac{\varepsilon_P}{\varepsilon_{P0}} (\mu P + C_0) = -\frac{\varepsilon_P}{\varepsilon_{P0}} (\mu \rho_r g z_0 + C_0) \quad [47]$$

Eq. 45 and eq. 47 link the weakening by pore pressure change with the linear strain weakening assumption in the model

$$\Delta P_f = \frac{z_0}{2\varepsilon_{P0}\delta \sin \theta} \left(\rho_r g z_0 + \frac{C_0}{\mu} \right) \quad [48]$$

This equation predicts the relationship between P_f/P_r and z_0 , which might not hold for every z_0 . Therefore, we evaluate averaged ε_{P0} by integrating eq. 48 with depth, representing the average weakening rate throughout the whole slab interface.

$$\bar{\varepsilon}_{P0} = \frac{1}{2\delta \sin \theta \int_0^{z_1} \Delta P_f dz_0} \left(\frac{\rho_r g z_1^3}{3} + \frac{C_0 z_1^2}{2\mu} \right) \quad [49]$$

We choose the integration upper limit to be $z_1 = 10 \text{ km}$, which is the depth of the strain partitioning zone.

In this section, we evaluate the correspondence between the

stress reduction at the subduction interface, driven by increased fluid pore pressure from dehydration processes (eq 45), and the strain weakening modeled in the numerical simulations (eq 46). Specifically, we quantify the extent of strain necessary within a shear zone in the model to replicate the weakening effect observed at the subduction zone interface. Given the assumed presence of a shear zone, displacement Δ_s , rather than strain ϵ_{P0} , might provide a more intuitive description of the weakening process:

$$\Delta_s = \frac{1}{\sin \theta \int_0^{z_1} \Delta P_f dz} \left(\frac{\rho_r g z_1^3}{3} + \frac{C_0 z_1^2}{2\mu} \right) \quad [50]$$

The shallowest part of the Puységur subduction zone is characterized by a strain-partitioning system with a well-defined thickness, prompting us to set the shear zone thickness at $\delta = 10$ km.

1. J Toth, M Gurnis, Dynamics of subduction initiation at preexisting fault zones. *J. Geophys. Res. Solid Earth* **103**, 18053–18067 (1998).
2. TV Gerya, RJ Stern, M Baes, SV Sobolev, SA Whattam, Plate tectonics on the Earth triggered by plume-induced subduction initiation. *Nature* **527**, 221–225 (2015).
3. R Stern, Subduction initiation: spontaneous and induced. *Earth Planet. Sci. Lett.* **226**, 275–292 (2004).
4. S Zhang, W Leng, L Chen, Continental tip with thinned lithosphere and thickened crust is a favorable location for subduction initiation. *J. Geophys. Res. Solid Earth* **128**, e2023JB027067 (2023).
5. Y Li, M Gurnis, A simple force balance model of subduction initiation. *Geophys. J. Int.* **232**, 128–146 (2022).
6. M Gurnis, C Hall, L Lavier, Evolving force balance during incipient subduction. *Geochem. Geophys. Geosystems* **5** (2004).
7. TV Gerya, D Bercovici, TW Becker, Dynamic slab segmentation due to brittle–ductile damage in the outer rise. *Nature* **599**, 245–250 (2021).
8. S Zhang, W Leng, Subduction polarity reversal: Induced or spontaneous? *Geophys. Res. Lett.* **48**, e2021GL093201 (2021).
9. D Bercovici, Y Ricard, Plate tectonics, damage and inheritance. *Nature* **508**, 513–516 (2014).
10. M Thielmann, BJ Kaus, Shear heating induced lithospheric-scale localization: Does it result in subduction? *Earth Planet. Sci. Lett.* **359–360**, 1–13 (2012).
11. D Kiss, LG Candiotti, T Duretz, SM Schmalholz, Thermal softening induced subduction initiation at a passive margin. *Geophys. J. Int.* **220**, 2068–2073 (2020).
12. D Dymkova, T Gerya, Porous fluid flow enables oceanic subduction initiation on earth. *Geophys. Res. Lett.* **40**, 5671–5676 (2013).
13. Ki Hiraiuchi, K Fukushima, M Kido, J Muto, A Okamoto, Reaction-induced rheological weakening enables oceanic plate subduction. *Nat. Commun.* **7**, 12550 (2016).
14. SC Cande, JM Stock, Pacific–Antarctic–Australia motion and the formation of the Macquarie Plate. *Geophys. J. Int.* **157**, 399–414 (2004).
15. WR Keller, “Cenozoic plate tectonic reconstructions and plate boundary processes in the Southwest Pacific,” PhD thesis, California Institute of Technology (2005). URL <https://resolver.caltech.edu/CaltechETD:etd-01102005-223039>. Medium: PDF Version Number: Final.
16. H Choi, et al., The kinematic evolution of the Macquarie Plate: A case study for the fragmentation of oceanic lithosphere. *Earth Planet. Sci. Lett.* **478**, 132–142 (2017).
17. B Shuck, et al., Stress transition from horizontal to vertical forces during subduction initiation. *Nat. Geosci.* **15**, 149–155 (2022).
18. JY Collot, et al., Morphostructure of an incipient subduction zone along a transform plate boundary: Puységur Fracture and Trench. *Geology* **23**, 519 (1995).
19. E Hightower, M Gurnis, H Van Avendonk, A Bayesian 3-D linear gravity inversion for complex density distributions: application to the Puységur subduction system. *Geophys. J. Int.* **223**, 1899–1918 (2020).
20. R Sutherland, P Barnes, C Uruski, Miocene–Recent deformation, surface elevation, and volcanic intrusion of the overriding plate during subduction initiation, offshore southern Fiordland, Puységur margin, southwest New Zealand. *New Zealand J. Geol. Geophys.* **49**, 131–149 (2006).
21. JF Lebrun, G Lamarche, JY Collot, Subduction initiation at a strike-slip plate boundary: The Cenozoic Pacific–Australian plate boundary, south of New Zealand. *J. Geophys. Res. Solid Earth* **108** (2003).
22. M Gurnis, et al., Incipient subduction at the contact with stretched continental crust: The Puységur Trench. *Earth Planet. Sci. Lett.* **520**, 212–219 (2019).
23. WBF Ryan, et al., Global Multi-Resolution Topography synthesis. *Geochem. Geophys. Geosystems* **10**, 2008GC002332 (2009).
24. G Ekström, M Nettles, A Dziwowski, The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. *Phys. Earth Planet. Interiors* **200–201**, 1–9 (2012).
25. JA Alvarez-Gomez, FMC—Earthquake focal mechanisms data management, cluster and classification. *SoftwareX* **9**, 299–307 (2019).
26. H Seebeck, et al., The New Zealand Community Fault Model – version 1.0: an improved geological foundation for seismic hazard modelling. *New Zealand J. Geol. Geophys.* pp. 1–21 (2023).
27. LL Lavier, WR Buck, AN Poliakov, Factors controlling normal fault offset in an ideal brittle layer. *J. Geophys. Res. Solid Earth* **105**, 23431–23442 (2000) Publisher: Wiley Online Library.
28. TV Gerya, Three-dimensional thermomechanical modeling of oceanic spreading initiation and evolution. *Phys. Earth Planet. Interiors* **214**, 35–52 (2013).
29. X Zhong, Z Li, Forced subduction initiation at passive continental margins: Velocity-driven versus stress-driven. *Geophys. Res. Lett.* **46**, 11054–11064 (2019).
30. Y Li, M Gurnis, Strike slip motion and the triggering of subduction initiation. *Front. Earth Sci.* **11**, 1168999 (2023).
31. PB Kelemen, G Hirth, A periodic shear-heating mechanism for intermediate-depth earthquakes in the mantle. *Nature* **446**, 787–790 (2007).
32. YY Kagan, Simplified algorithms for calculating double-couple rotation. *Geophys. J. Int.* **171**, 411–418 (2007).
33. WF Brace, DL Kohlstedt, Limits on lithospheric stress imposed by laboratory experiments. *J. Geophys. Res. Solid Earth* **85**, 6248–6252 (1980).
34. J Jackson, Strength of the continental lithosphere: Time to abandon the jelly sandwich? *GSA Today* **12**, 4 (2002).
35. G Oliver, J Coggon, Crustal structure of Fiordland, New Zealand. *Tectonophysics* **54**, 253–292 (1979).
36. DM Saffer, HJ Tobin, Hydrogeology and Mechanics of Subduction Zone Forearcs: Fluid Flow and Pore Pressure. *Annu. Rev. Earth Planet. Sci.* **39**, 157–186 (2011).
37. A Fagereng, JF Diener, S Ellis, F Remitti, Fluid-related deformation processes at the up- and down-dip limits of the subduction thrust seismogenic zone: What do the rocks tell us? in *Geology and Tectonics of Subduction Zones: A Tribute to Gaku Kimura*. (Geological Society of America), (2018).
38. W Zhu, KL Allison, EM Dunham, Y Yang, Fault valving and pore pressure evolution in simulations of earthquake sequences and aseismic slip. *Nat. Commun.* **11**, 4833 (2020).
39. C Petrin, et al., Seismo-hydro-mechanical modelling of the seismic cycle: Methodology and implications for subduction zone seismicity. *Tectonophysics* **791**, 228504 (2020).
40. L Dal Zilio, T Gerya, Subduction earthquake cycles controlled by episodic fluid pressure cycling. *Lithos* **426**, 106800 (2022).
41. B Shuck, et al., Strike-Slip Enables Subduction Initiation Beneath a Failed Rift: New Seismic Constraints From Puységur Margin, New Zealand. *Tectonics* **40** (2021).
42. VL Hansen, Subduction origin on early Earth: A hypothesis. *Geology* **35**, 1059 (2007).
43. Q Yuan, M Gurnis, PD Asimov, Y Li, A giant impact origin for the first subduction on earth. *Geophys. Res. Lett.* **51**, e2023GL106723 (2024).
44. S Zhang, Y Li, W Leng, M Gurnis, Photoferrothermophilic bacteria initiated plate tectonics in the Neoproterozoic. *Geophys. Res. Lett.* **50**, e2023GL103553 (2023).
45. D Bercovici, E Mulyukova, Evolution and demise of passive margins through grain mixing and damage. *Proc. Natl. Acad. Sci.* **118**, e2011247118 (2021).
46. SA Wilde, JW Valley, WH Peck, CM Graham, Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago. *Nature* **409**, 175–178 (2001).
47. J Korenaga, Thermal cracking and the deep hydration of oceanic lithosphere: A key to the generation of plate tectonics? *J. Geophys. Res. Solid Earth* **112**, 2006JB004502 (2007).
48. R Beucher, et al., Underworld2: Python Geodynamics Modelling for Desktop, HPC and Cloud (2022).
49. RD Müller, et al., GPlates: Building a Virtual Earth Through Deep Time. *Geochem. Geophys. Geosystems* **19**, 2243–2261 (2018).
50. KS Karlsen, M Domeier, C Gaina, CP Conrad, A tracer-based algorithm for automatic generation of seafloor age grids from plate tectonic reconstructions. *Comput. & Geosci.* **140**, 104508 (2020).
51. G Laske, G Masters, Z Ma, M Pasyanos, Update on CRUST1.0 - A 1-degree Global Model of Earth's Crust. (2013).
52. MB Croon, SC Cande, JM Stock, Revised Pacific–Antarctic plate motions and geophysics of the Menard Fracture Zone. *Geochem. Geophys. Geosystems* **9**, 2008GC002019 (2008).
53. Si Karato, P Wu, Rheology of the Upper Mantle: A Synthesis. *Science* **260**, 771–778 (1993).
54. D McKenzie, The initiation of trenches: a finite amplitude instability. *Island arcs, deep sea trenches back-arc basins* **1**, 57–61 (1977).
55. R Bürgmann, G Dresen, Rheology of the Lower Crust and Upper Mantle: Evidence from Rock Mechanics, Geodesy, and Field Observations. *Annu. Rev. Earth Planet. Sci.* **36**, 531–567 (2008).
56. BR Hacker, GA Abers, SM Peacock, Subduction factory 1. Theoretical mineralogy, densities, seismic wave speeds, and H₂O contents: SUBDUCTION ZONE MINERALOGY AND PHYSICAL PROPERTIES. *J. Geophys. Res. Solid Earth* **108** (2003).
57. BJP Kaus, YY Podladchikov, Initiation of localized shear zones in viscoelastoplastic rocks. *J. Geophys. Res.* **111**, B04412 (2006).
58. D Bercovici, Y Ricard, Mechanisms for the generation of plate tectonics by two-phase grain-damage and pinning. *Phys. Earth Planet. Interiors* **202–203**, 27–55 (2012).
59. D Bercovici, G Schubert, Y Ricard, Abrupt tectonics and rapid slab detachment with grain damage. *Proc. Natl. Acad. Sci.* **112**, 1287–1291 (2015).
60. A Rozel, Y Ricard, D Bercovici, A thermodynamically self-consistent damage equation for grain size evolution during dynamic recrystallization. *Geophys. J. Int.* **184**, 719–728 (2011).
61. D Bercovici, Y Ricard, Generation of plate tectonics with two-phase grain-damage and pinning: Source–sink model and toroidal flow. *Earth Planet. Sci. Lett.* **365**, 275–288 (2013).
62. D Bercovici, Y Ricard, Grain-damage hysteresis and plate tectonic states. *Phys. Earth Planet. Interiors* **253**, 31–47 (2016).
63. JP Evans, CB Forster, JV Goddard, Permeability of fault-related rocks, and implications for hydraulic structure of fault zones. *J. Struct. Geol.* **19**, 1393–1404 (1997).

DRAFT

1489	64. DR Faulkner, EH Rutter, Can the maintenance of overpressured fluids in large strike-slip	1551
1490	fault zones explain their apparent weakness? <i>Geology</i> 29 , 503 (2001).	1552
1491	65. JJ Dong, et al., Stress-dependence of the permeability and porosity of sandstone and shale	1553
1492	from TCDP Hole-A. <i>Int. J. Rock Mech. Min. Sci.</i> 47 , 1141–1157 (2010).	1554
1493	66. J Zheng, L Zheng, HH Liu, Y Ju, Relationships between permeability, porosity and effective	1555
1494	stress for low-permeability sedimentary rock. <i>Int. J. Rock Mech. Min. Sci.</i> 78 , 304–318	1556
1495	(2015).	1557
1496		1558
1497		1559
1498		1560
1499		1561
1500		1562
1501		1563
1502		1564
1503		1565
1504		1566
1505		1567
1506		1568
1507		1569
1508		1570
1509		1571
1510		1572
1511		1573
1512		1574
1513		1575
1514		1576
1515		1577
1516		1578
1517		1579
1518		1580
1519		1581
1520		1582
1521		1583
1522		1584
1523		1585
1524		1586
1525		1587
1526		1588
1527		1589
1528		1590
1529		1591
1530		1592
1531		1593
1532		1594
1533		1595
1534		1596
1535		1597
1536		1598
1537		1599
1538		1600
1539		1601
1540		1602
1541		1603
1542		1604
1543		1605
1544		1606
1545		1607
1546		1608
1547		1609
1548		1610
1549		1611
1550		1612