

Proportions and metamorphic conditions of markers in numerical geodynamic models of subduction

Buchanan C. Kerswell ¹Matthew J. Kohn ¹Taras V. Gerya ²

¹Department of Geosciences, Boise State University, Boise, ID 83725

²Department of Earth Sciences, ETH-Zurich, Sonneggstrasse 5, Zurich 8092, Switzerland

Key Points:

-
-
-

Abstract

1 Introduction

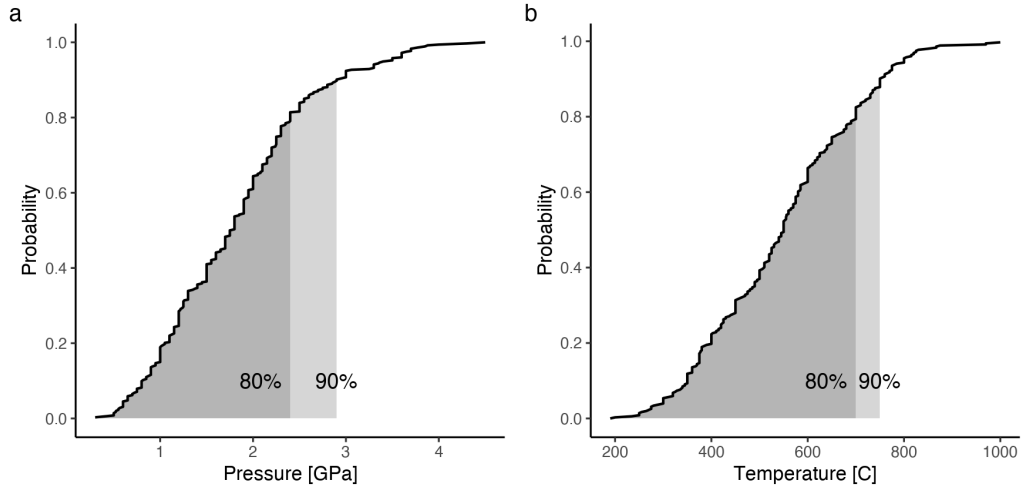


Figure 1: Cumulative probability of peak metamorphic pressures (a) and temperatures (b) for a global compilation of exhumed blueschists and eclogites. The rock record implies 80% of rocks are recovered from ≤ 2.4 GPa and ≤ 700 °C. Note the abrupt change in slope at $P=2.4$ GPa. Data from Penniston-Dorland et al. (2015).

2 Methods

This study presents a dataset of Lagrangian markers (described below) from the numerical experiments of Kerswell et al. (2020). The numerical experiments simulate 64 oceanic-continental subduction systems with variable oceanic plate age, convergence velocity, and upper plate (continental) lithospheric thickness. The range of oceanic plate ages and convergence velocities broadly represent the modern global subduction system (Syracuse & Abers, 2006; Wada & Wang, 2009). Initial conditions were modified from previous studies of active margins (Gorczyk et al., 2007; Sizova et al., 2010). The code, **I2VIS**, models visco-plastic flow of geologic materials by solving three conservative equations of mass, energy, and momentum on a fully-staggered finite difference grid with a *marker-in-cell* technique (Gerya & Yuen, 2003; Harlow & Welch, 1965). Further details about the initial setup and boundary conditions, rheologic model, metamorphic (de)hydration reactions, are in Kerswell et al. (2020). Details about the marker-in-cell technique are in Gerya & Yuen (2003) and Gerya (2019).

In this section we first define Lagrangian markers (now referred to as markers) and briefly elaborate on their usefulness in understanding fluid flow—including geodynamic problems like subduction. We then detail the maths and decisions involved in our marker classification algorithm, which we use to classify 1,214,757 markers from 64 numerical experiments of Kerswell et al. (2020) as either subducted or recovered based on characteristics of their PTt paths.

2.1 Lagrangian markers

Markers are mathematical objects representing discrete parcels of fluid flowing in a continuum (Harlow, 1962, 1964). Imagine tracking millions of parcels of air as they collectively move around an air foil. Each marker would experience a different flow path and pressure history relative to its neighbors. For example, some parcels of air may spin off and cause turbulence, while others move orderly with their neighbors around the foil. In this analogy, knowing which parcels experience turbulence, perhaps based on some characteristic of a parcel’s flow path, is useful information for aerodynamic testing.

The analogy above highlights the representative advantage of markers. That is, markers save the physical state of the continuum at each timestep, thereby saving a historical record of physical changes from the perspective of the fluid (Harlow, 1962). “Trac-

ing” markers (saving marker states through time) is useful for understanding subduction dynamics like tracing air parcels is useful for understanding aerodynamics. But it is necessary to first accept implicit simplifications and uncertainties (known and unknown) about the geologic continuum. Uncertainties are especially rich in the petrologic model governing phase changes (Ito & Kennedy, 1971; Schmidt & Poli, 1998) and its effects on the highly non-linear rheologic model relating stress and strain by empirical flow laws (Hilalret et al., 2007; Karato & Wu, 1993; Ranalli, 1995; Turcotte & Schubert, 2002). See Kerswell et al. (2020) for details. Nonetheless, insofar as subducting crustal rocks on earth behave like markers deforming in an incompressible visco-plastic fluid (as parameterized by Gerya, 2019; Gerya & Yuen, 2003; Kerswell et al., 2020), principled comparisons between marker PTt paths and the rock record (Figure 1, Agard et al., 2018; Penniston-Dorland et al., 2015) may be made.

Markers also have a distinct numerical advantage. Unlike air parcels flowing in a relatively homogeneous atmosphere (air foil aside), markers are deforming in a partly layered, partly chaotic, visco-plastic continuum representing the interface between the top of subducting oceanic crust and the mantle it sinks into (often called the subduction interface). Current models of the subduction interface epitomize a geologic continuum with complex geometry, sharp thermal, chemical, and strain gradients, strong advection (high Peclet number), and abundant fluid flow (Agard et al., 2016, 2018; Bebout, 2007; Bebout & Barton, 2002; Gerya & Yuen, 2003; Penniston-Dorland et al., 2015; Syracuse et al., 2010). Solving the momentum and temperature equations on a finite-difference grid becomes unstable in such cases. Markers greatly improve solution accuracy and stability by interpolating and updating thermal, chemical, and velocity fields during each timestep (Gerya, 2019; Gerya & Yuen, 2003; Moresi et al., 2003).

2.2 Marker classification

On average, 18,981 markers are selected from within a 760 km wide and 8 km deep section of oceanic crust and seafloor sediments (Figure 2). Tracing then proceeds for a number of timesteps $t = \{1, 2, \dots, t_{dur}\}$, where t_{dur} depends on the subduction rate (see sec. A.1). From this set of markers, $x_i = \{x_1, x_2, \dots, x_n\}$. Only markers *recovered* from the subducting slab are relevant for comparison to PT estimates of natural rocks. The main challenge, therefore, is to first classify markers as either *subducted* or *recovered* without an inherited class label.

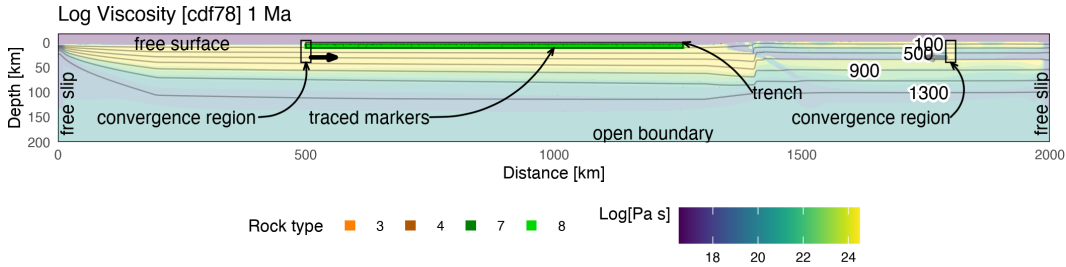


Figure 2: Initial conditions for marker tracing. Markers are selected from a 760 km wide and 8 km deep section representing 7 km of mafic oceanic crust (rock type 7 & 8) and 1 km of seafloor sediments (rock type 3 & 4). Two convergence regions far from the trench initiate and maintain subduction. On average, 18,981 markers are traced for each model. Numerical modelling details are in Kerswell et al. (2020).

At the heart of our marker classification algorithm is a finite Gaussian mixture model (GMM) fit by Expectation-Maximization (EM, Dempster et al., 1977). Please note that GMM fit by EM is a general purpose clustering algorithm broadly used in pattern recognition, anomaly detection, and estimating complex probability distribution functions (e.g., Banfield & Raftery, 1993; Celeux & Govaert, 1995; Figueiredo & Jain, 2002; Fraley & Raftery, 2002; Vermeesch, 2018). We derive GMM in sec. 2.2.1 and EM in sec. 2.2.2.

Before deriving the details of marker classification, we hypothesize that subducted and recovered markers may be distinguished by their PTt paths. If true, clustering algorithms like GMM may reliably classify markers by (dis)similarity along any number of dimensions computed from marker PTt paths (e.g., Dy & Brodley, 2004).

2.2.1 Gaussian mixture model

Let the traced markers represent a d -dimensional array of n random independent variables $x_i \in \mathbb{R}$. Assume markers x_i were drawn from k discrete probability distribu-

tions with parameters Φ . The probability distribution of markers x_i can be modeled with a mixture of k components:

$$p(x_i|\Phi) = \sum_{j=1}^k \pi_j p(x_i|\Theta_j) \quad (1)$$

where $p(x_i|\Theta_j)$ is the probability of x_i under the j^{th} mixture component and π_j is the mixture proportion representing the probability that x_i belongs to the j^{th} component ($\pi_j \geq 0; \sum_{j=1}^k \pi_j = 1$).

Assuming Θ_j describes a Gaussian probability distributions with mean μ_j and covariance Σ_j , Equation 1 becomes:

$$p(x_i|\Phi) = \sum_{j=1}^k \pi_j \mathcal{N}(x_i|\mu_j, \Sigma_j) \quad (2)$$

where

$$\mathcal{N}(x_i|\mu_j, \Sigma_j) = \frac{\exp\{-\frac{1}{2}(x_i - \mu_j)(x_i - \mu_j)^T \Sigma_j^{-1}\}}{\sqrt{\det(2\pi \Sigma_j)}} \quad (3)$$

The parameters μ_j and Σ_j , representing the center and shape of each cluster, are estimated by maximizing the log of the likelihood function, $L(x_i|\Phi) = \prod_{i=1}^n p(x_i|\Phi)$:

$$\log L(x_i|\Phi) = \log \prod_{i=1}^n p(x_i|\Phi) = \sum_{i=1}^n \log \left[\sum_{j=1}^k \pi_j p(x_i|\Theta_j) \right] \quad (4)$$

Taking the derivative of Equation 4 with respect to each parameter, π , μ , Σ , setting the equation to zero, and solving for each parameter gives the Maximum Likelihood Estimators (MLE):

$$\begin{aligned} N_j &= \sum_{i=1}^n \omega_{ij} \\ \pi_j &= \frac{N_j}{n} \\ \mu_j &= \frac{1}{N_j} \sum_{i=1}^n \omega_{ij} x_i \\ \Sigma_j &= \frac{1}{N_j} \sum_{i=1}^n \omega_{ij} (x_i - \mu_j)(x_i - \mu_j)^T \end{aligned} \quad (5)$$

where ω_{ij} ($\omega_{ij} \geq 0$; $\sum_{j=1}^k \omega_{ij} = 1$) are membership weights representing the probability of an observation x_i belonging to the j^{th} Gaussian and N_j represents the number of observations belonging to the j^{th} Gaussian. Please note that ω_{ij} is unknown for unlabelled datasets, like the unclassified markers, so MLE cannot be computed with Equation 5. The solution to this problem is derived in sec. 2.2.2.

General purpose functions in the R package `Mclust` (Scrucca et al., 2016) are used to fit Gaussian mixture models. After Banfield & Raftery (1993), covariance matrices Σ in `Mclust` are parameterized to be flexible in their shape, volume, and orientation (Scrucca et al., 2016):

$$\Sigma_j = \lambda_j D_j A_j D_j^T \quad (6)$$

where D_j is the orthogonal eigenvector matrix, A_j and λ_j are diagonal matrices of values proportional to the eigenvalues. This implementation allows fixing one, two, or three geometric elements of the covariance matrices. That is, the volume λ_j , shape A_j , and orientation D_j of Gaussian clusters can change or be fixed among all k clusters (e.g., Celeux & Govaert, 1995; Fraley & Raftery, 2002). Fourteen parameterizations of Equation 6 are tried, representing different geometric combinations of the covariance matrices Σ (see Scrucca et al., 2016) and the Bayesian Information Criterion (BIC, Schwarz & others, 1978) is computed. The best parameterization for Equation 6 is chosen by BIC.

2.2.2 Expectation-Maximization fitting of Gaussian Mixtures

The EM algorithm estimates GMM parameters by initializing k Gaussians with parameters (π_j, μ_j, Σ_j) , then iteratively computing membership weights with Equation 7 (E-step) and updating Gaussian parameters with Equation 5 (M-step) until convergence (Dempster et al., 1977).

The *expectation* (E)-step involves a “latent” multinomial variable $z_{ij} \in \{1, 2, \dots, k\}$ representing the unknown classifications of x_i with a joint distribution $p(x_i, z_{ij}) = p(x_i|z_{ij})p(z_j)$. Membership weights ω_{ij} are equivalent to the conditional probability $p(z_{ij}|x_i)$, which represents the probability of observation x_i belonging to the j^{th} Gaussian. Given initial guesses for k sets of Gaussian parameters π_j, μ_j, Σ_j , membership weights are computed using Bayes Theorem (E-step):

$$p(z_{ij}|x_i) = \frac{p(x_i|z_{ij})p(z_{ij})}{p(x_i)} = \frac{\pi_j \mathcal{N}(\mu_j, \Sigma_j)}{\sum_{j=1}^k \pi_j \mathcal{N}(\mu_j, \Sigma_j)} = \omega_{ij} \quad (7)$$

and new Gaussian estimates are computed with ω_{ij} (Equation 5) during the *maximization* (M-)step.

2.2.3 Feature selection

EM is sensitive to local optima and initialization (Figueiredo & Jain, 2002), so a number of features were computed from marker PTt paths and tested in combination. Redundant or useless features (e.g., Dy & Brodley, 2004) were filtered out. We settled on two features:

$$\begin{aligned} \max P &\leftarrow \max_{1 \leq t \leq t} P \\ \text{sum} dP &\leftarrow \sum_1^{t_{dur}} dP \end{aligned} \quad (8)$$

where $\max P$ and $\text{sum} dP$ represent the maximum pressure attained each marker's PTt path and the sum total of all pressure changes along each marker's PTt path, respectively. The bivariate mixture model described by Equation 2 and Equation 8 is fit with $k = 6$ Gaussian clusters with EM (Equations 7, 5). The results of this clustering step are arbitrary class labels $z_i \in \{1, \dots, k\}$ representing assignment of marker x_i to one of k clusters. Next, a decision must be made to determine which clusters (if any) represent recovered markers.

2.2.4 Subducted or recovered?

GMM assigns markers into up to 6 clusters, so a final decision to classify a cluster as containing either all *subducted* or *recovered* markers is made by comparing each cluster's centroid (μ_j , Equation 5) to the overall probability distribution of markers along $\max P$ and $\text{sum} dP$. Groups with centroids μ_j well above the median $\max P$ and $\text{sum} dP$ classify as *subducted*. The exact threshold is defined as three-quarters of an inter quartile range below the median (Figure 3).

Marker classification [cdf78]

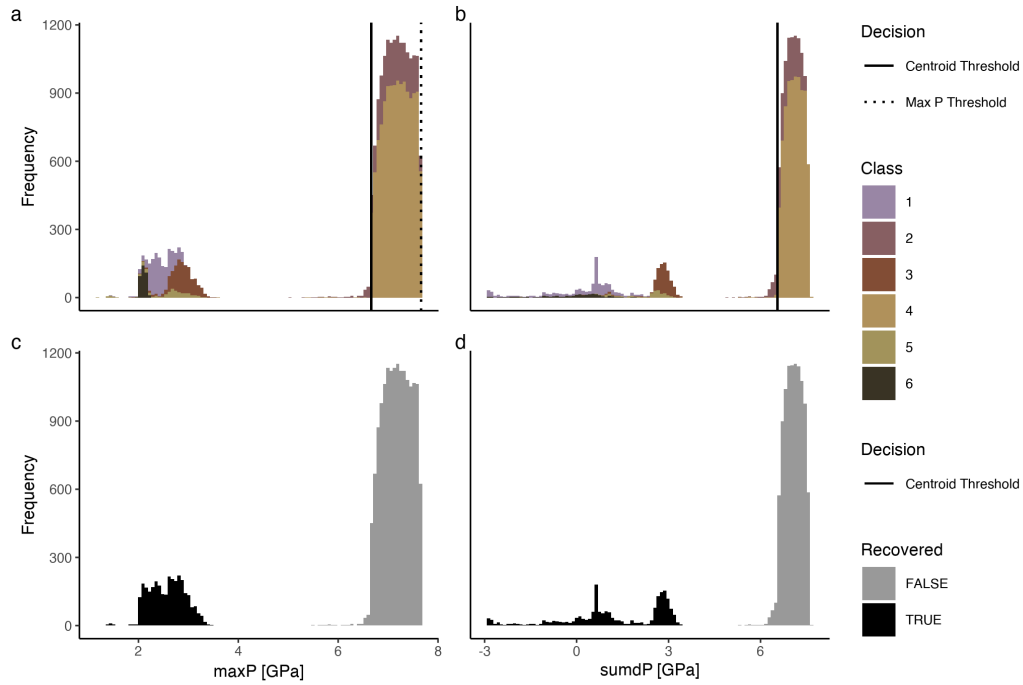
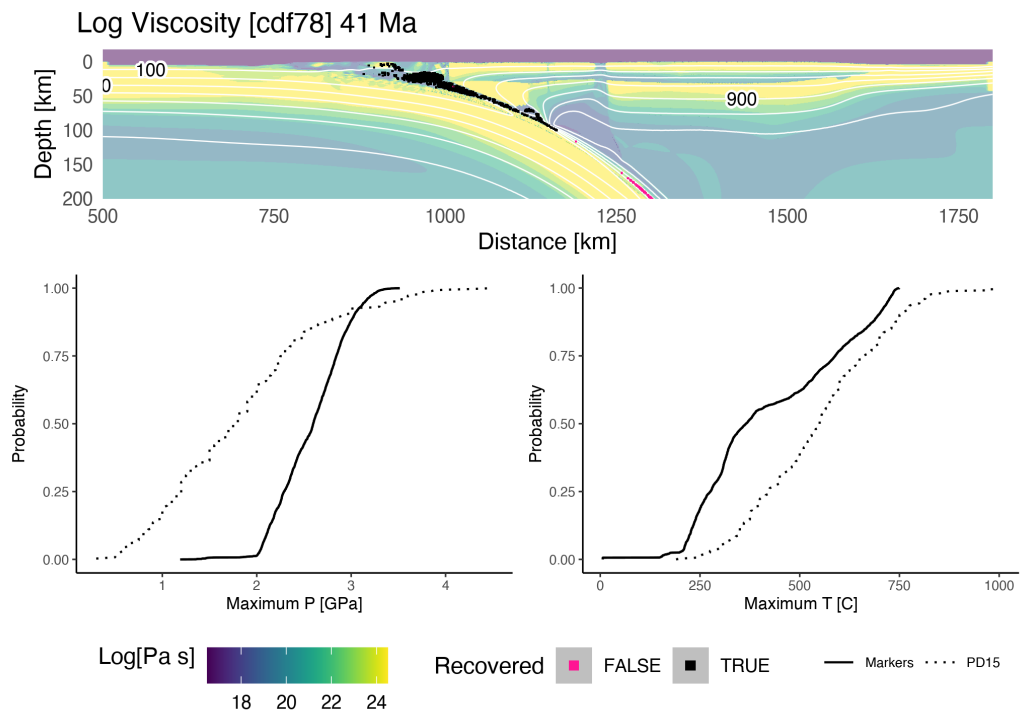


Figure 3: Classification results for model cdf78. GMM clusters $n = 18,934$ markers into six classes according to each marker's maximum pressure (a) and sum of all pressure changes along its PTt path (b). If a class' centroid (mean) is below either centroid threshold in (a) or (b), it is classified as recovered (c, d). If a marker gets classified as recovered, but is above the maximum pressure threshold in (a), it is reclassified as subducted (c, d).

150

3 Results



151

Table 1: Summary of subduction parameters and marker tracing results by numerical experiment

model	z_{cpl} [km]	Δz_{lith} [km]	age [Ma]	\vec{v}_{conv} [$\frac{km}{Ma}$]	n_{max}	n_{rec}	σ_{rec}	n_{sub}	σ_{sub}	ratio	σ_{ratio}	P_{max} [GPa]	$\sigma_{P_{max}}$	T_{max} [C]	$\sigma_{T_{max}}$
cda46	66	46	33	40	18802	4777	0	14025	0	0.34	0	6	0	747	0
cdb46	74	46	33	66	18913	5676	0	13237	0	0.43	0	6	0	750	0
cdc46	69	46	33	80	18921	5243	0	13678	0	0.38	0	8	0	830	0
cdd46	67	46	33	100	18912	5246	9.19	13666	9.19	0.38	0	7	0	837	0
cde46	72	46	55	40	18760	4477	0	14283	0	0.31	0	7	0	752	0
cdf46	78	46	55	66	18782	6047	0	12735	0	0.47	0	3	0	745	0
cdg46	78	46	55	80	18794	5390	0	13404	0	0.4	0	3	0	732	0
cdh46	59	46	55	100	18812	5344	0.71	13468	0.71	0.4	0	7	0	835	0
cdi46	80	46	85	40	18719	2690	0	16029	0	0.17	0	3	0	740	0
cdj46	70	46	85	66	18750	5212	0	13538	0	0.38	0	3	0	732	0
cdk46	58	46	85	80	18760	4996	1240	13764	1240	0.37	0.12	7	0	814	10.79
cdl46	65	46	85	100	18779	4724	0	14055	0	0.34	0	6	0	793	0
cdm46	79	46	110	40	18739	4447	0	14292	0	0.31	0	7	0	826	0
cdn46	70	46	110	66	18769	5134	0	13635	0	0.38	0	3	0	735	0
cdo46	68	46	110	80	18764	4231	0	14533	0	0.29	0	3	0	729	0

model	z_{cpl} [km]	Δz_{lith} [km]	age [Ma]	\vec{v}_{conv} [$\frac{km}{Ma}$]	n_{max}	n_{rec}	σ_{rec}	n_{sub}	σ_{sub}	ratio	σ_{ratio}	P_{max} [GPa]	$\sigma_{P_{max}}$	T_{max} [C]	$\sigma_{T_{max}}$
cdp46	64	46	110	100	18771	3837	0	14934	0	0.26	0	3	0	721	0
cda62	80	62	33	40	18942	5861	0	13081	0	0.45	0	7	0	778	0
cdb62	79	62	33	66	19034	4839	0	14195	0	0.34	0	6	0	742	0
cdc62	78	62	33	80	19036	4580	0	14456	0	0.32	0	8	0	744	0
cdd62	77	62	33	100	19028	4010	0.71	15018	0.71	0.27	0	8	0	731	0
cde62	87	62	55	40	18770	5248	0	13522	0	0.39	0	6	0	778	0
cdf62	82	62	55	66	18836	3733	0	15103	0	0.25	0	3	0	736	0
cdg62	75	62	55	80	18843	3658	0	15185	0	0.24	0	3	0	735	0
cdh62	70	62	55	100	18852	3237	0	15615	0	0.21	0	3	0	736	0
cdi62	91	62	85	40	18775	4666	0	14109	0	0.33	0	4	0	773	0
cdj62	77	62	85	66	18839	3612	0	15227	0	0.24	0	3	0	749	0
cdk62	72	62	85	80	18850	3139	0	15711	0	0.2	0	7	0	759	0
cdl62	67	62	85	100	18866	3646	22.63	15220	22.63	0.24	0	7	0	770	0
cdm62	88	62	110	40	18829	4328	0	14501	0	0.3	0	4	0	778	0
cdn62	77	62	110	66	18889	2643	0	16246	0	0.16	0	4	0	773	0
cdo62	74	62	110	80	18908	3479	0	15429	0	0.23	0	7	0	750	0
cdp62	75	62	110	100	18908	3181	0	15727	0	0.2	0	3	0	745	0

model	z_{cpl} [km]	Δz_{lith} [km]	age [Ma]	\vec{v}_{conv} [$\frac{km}{Ma}$]	n_{max}	n_{rec}	σ_{rec}	n_{sub}	σ_{sub}	ratio	σ_{ratio}	P_{max} [GPa]	$\sigma_{P_{max}}$	T_{max} [C]	$\sigma_{T_{max}}$
cda78	87	78	33	40	19228	3729	0	15499	0	0.24	0	6	0	781	0
cdb78	94	78	33	66	19326	3256	0	16070	0	0.2	0	6	0	764	0
cdc78	97	78	33	80	19324	2452	2.12	16872	2.12	0.15	0	8	0	835	0
cdd78	97	78	33	100	19338	3315	1469	16023	1469	0.21	0.11	8	0	829	0.51
cde78	90	78	55	40	18942	4226	0	14716	0	0.29	0	7	0	765	0
cdf78	90	78	55	66	18934	3080	0	15854	0	0.19	0	4	0	750	0
cdg78	88	78	55	80	18945	3078	0	15867	0	0.19	0	3	0	743	0
cdh78	85	78	55	100	18960	2844	0.71	16116	0.71	0.18	0	3	0	740	0
cdi78	97	78	85	40	18860	4186	0	14674	0	0.29	0	4	0	756	0
cdj78	91	78	85	66	18911	2861	0	16050	0	0.18	0	4	0	758	0
cdk78	84	78	85	80	18909	2658	1.41	16251	1.41	0.16	0	3	0	744	0
cdl78	77	78	85	100	18910	2682	0	16228	0	0.17	0	7	0	795	0
cdm78	78	78	110	40	18885	3397	0	15488	0	0.22	0	4	0	757	0
cdn78	87	78	110	66	18935	3230	0.71	15704	0.71	0.21	0	4	0	753	0
cdo78	85	78	110	80	18944	2813	0	16131	0	0.17	0	4	0	753	0
cdp78	78	78	110	100	18940	2744	45.96	16196	45.96	0.17	0	7	0	789	0
cda94	95	94	33	40	19634	2635	0	16999	0	0.16	0	6	0	772	0

model	z_{cpl} [km]	Δz_{lith} [km]	age [Ma]	\vec{v}_{conv} [$\frac{km}{Ma}$]	n_{max}	n_{rec}	σ_{rec}	n_{sub}	σ_{sub}	ratio	σ_{ratio}	P_{max} [GPa]	$\sigma_{P_{max}}$	T_{max} [C]	$\sigma_{T_{max}}$
cdb94	101	94	33	66	19701	2863	1.41	16838	1.41	0.17	0	8	0	836	0
cdc94	108	94	33	80	19722	2167	0	17555	0	0.12	0	8	0	754	0
cdd94	113	94	33	100	19732	3648	760.1	16084	760.1	0.23	0.06	8	0	831	22.04
cde94	100	94	55	40	19169	2936	3.54	16232	3.54	0.18	0	8	0	832	18.07
cdf94	104	94	55	66	19166	2530	0	16636	0	0.15	0	4	0	761	0
cdg94	104	94	55	80	19190	2622	2.12	16568	2.12	0.16	0	4	0	756	0
cdh94	104	94	55	100	19220	2350	6.36	16870	6.36	0.14	0	4	0	746	0
cdi94	101	94	85	40	18996	3108	0	15888	0	0.2	0	4	0	772	0
cdj94	102	94	85	66	19003	2580	0	16423	0	0.16	0	4	0	753	0
cdk94	101	94	85	80	19001	2584	0	16417	0	0.16	0	4	0	755	0
cdl94	107	94	85	100	19008	2704	0	16304	0	0.17	0	7	0	786	0
cdm94	106	94	110	40	19011	3098	0	15913	0	0.19	0	4	0	769	0
cdn94	102	94	110	66	18997	2382	0	16615	0	0.14	0	3	0	752	0
cdo94	98	94	110	80	18988	2558	0	16430	0	0.16	0	4	0	771	0
cdp94	108	94	110	100	18976	2944	0	16032	0	0.18	0	4	0	751	0

152 **4 Discussion**

153 **4.1 Diapirs**

154 **5 Conclusion**

155 **6 Open Research**

156 **Acknowledgments**

157 This work was supported by the National Science Foundation grant OIA1545903
158 to M. Kohn, S. Penniston-Dorland, and M. Feineman.

159 **7 References**

A Appendix

A.1 Marker tracing duration

Spontaneous sinking motion of the oceanic plate, as opposed to a fixed subduction rate (e.g., Syracuse et al., 2010; Wada & Wang, 2009), induces right-to-left plate motions as the sinking oceanic plate (the *slab*) provides a leftward horizontal force (known as *slab rollback*). Slab rollback eventually leads to mechanical interference (collision) between trench sediments and the stationary convergence region centered at 500 *km* from the left boundary (Figure 2). The fixed, high-viscosity, convergence region acts as a barrier to the incoming sediments, deforming the accretionary wedge into a rapidly thickening pile. The sudden change in accretionary wedge geometry flattens the slab, causing intense mantle circulation and crustal deformation in the forearc and backarc regions. We consider the dynamics after interference begins unrepresentative of natural buoyancy-driven slab motion. Therefore, marker PTt paths are also increasingly meaningless after mechanical interference begins.

A tracing duration t_{dur} must therefore be chosen for each model that immediately precedes mechanical interference. Models with rapid subduction rates, for example, interfere with the convergence region sooner than models with slower subduction rates. We define t_{dur} automatically for each model by computing the topographic surface profile for each timestep. Tracing stops when the sediment pile deforming against the convergence region becomes the overall topographic high, usually within one or two timesteps after interference. Markers PTt paths from different models, therefore, represent approximately the same amount of total convergence in *km*, but different subduction durations.

Agard, P., Yamato, P., Soret, M., Prigent, C., Guillot, S., Plunder, A., et al. (2016). Plate interface rheological switches during subduction infancy: Control on slab penetration and metamorphic sole formation. *Earth and Planetary Science Letters*, 451, 208–220.

Agard, P., Plunder, A., Angiboust, S., Bonnet, G., & Ruh, J. (2018). The subduction plate interface: Rock record and mechanical coupling (from long to short timescales). *Lithos*, 320, 537–566.

Banfield, J. D., & Raftery, A. E. (1993). Model-based gaussian and non-gaussian clustering. *Biometrics*, 803–821.

- 191 Bebout, G. E. (2007). Metamorphic chemical geodynamics of subduction zones. *Earth*
192 *and Planetary Science Letters*, 260(3-4), 373–393.
- 193 Bebout, G. E., & Barton, M. D. (2002). Tectonic and metasomatic mixing in a high-t,
194 subduction-zone mélange—insights into the geochemical evolution of the slab–mantle
195 interface. *Chemical Geology*, 187(1-2), 79–106.
- 196 Celeux, G., & Govaert, G. (1995). Gaussian parsimonious clustering models. *Pattern Recog-*
197 *niton*, 28(5), 781–793.
- 198 Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incom-
199 plete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B*
200 *(Methodological)*, 39(1), 1–22.
- 201 Dy, J. G., & Brodley, C. E. (2004). Feature selection for unsupervised learning. *Jour-*
202 *nal of Machine Learning Research*, 5(Aug), 845–889.
- 203 Figueiredo, M. A. T., & Jain, A. K. (2002). Unsupervised learning of finite mixture mod-
204 els. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3), 381–
205 396.
- 206 Fraley, C., & Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and
207 density estimation. *Journal of the American Statistical Association*, 97(458), 611–
208 631.
- 209 Gerya, T. V. (2019). *Introduction to numerical geodynamic modelling*. Cambridge Uni-
210 versity Press.
- 211 Gerya, T. V., & Yuen, D. A. (2003). Characteristics-based marker-in-cell method with
212 conservative finite-differences schemes for modeling geological flows with strongly vari-
213 able transport properties. *Physics of the Earth and Planetary Interiors*, 140(4), 293–
214 318.
- 215 Gorczyk, W., Willner, A. P., Gerya, T. V., Connolly, J. A., & Burg, J.-P. (2007). Phys-
216 ical controls of magmatic productivity at pacific-type convergent margins: Numer-
217 ical modelling. *Physics of the Earth and Planetary Interiors*, 163(1-4), 209–232.
- 218 Harlow, F. H. (1962). *The particle-in-cell method for numerical solution of problems in*
219 *fluid dynamics*. Los Alamos Scientific Lab., N. Mex.

- 220 Harlow, F. H. (1964). The particle-in-cell computing method for fluid dynamics. *Meth-*
221 *ods Comput. Phys.*, *3*, 319–343.
- 222 Harlow, F. H., & Welch, J. E. (1965). Numerical calculation of time-dependent viscous
223 incompressible flow of fluid with free surface. *The Physics of Fluids*, *8*(12), 2182–2189.
- 224 Hilaiet, N., Reynard, B., Wang, Y., Daniel, I., Merkel, S., Nishiyama, N., & Petitgirard,
225 S. (2007). High-pressure creep of serpentine, interseismic deformation, and initiation
226 of subduction. *Science*, *318*(5858), 1910–1913.
- 227 Ito, K., & Kennedy, G. C. (1971). An experimental study of the basalt-garnet granulite-
228 eclogite transition. *The Structure and Physical Properties of the Earth's Crust*, *14*,
229 303–314.
- 230 Karato, S., & Wu, P. (1993). Rheology of the upper mantle: A synthesis. *Science*, *260*(5109),
231 771–778.
- 232 Kerswell, B. C., Kohn, M. J., & Gerya, T. V. (2020). Backarc lithospheric thickness and
233 serpentine stability control slab-mantle coupling depths in subduction zones. *Earth*
234 *and Space Science Open Archive*, *34*. <https://doi.org/10.1002/essoar.10503710>
235 .1
- 236 Moresi, L., Dufour, F., & Mühlhaus, H.-B. (2003). A lagrangian integration point finite
237 element method for large deformation modeling of viscoelastic geomaterials. *Jour-*
238 *nal of Computational Physics*, *184*(2), 476–497.
- 239 Penniston-Dorland, S. C., Kohn, M. J., & Manning, C. E. (2015). The global range of
240 subduction zone thermal structures from exhumed blueschists and eclogites: Rocks
241 are hotter than models. *Earth and Planetary Science Letters*, *428*, 243–254.
- 242 Ranalli, G. (1995). *Rheology of the earth*. Springer Science & Business Media.
- 243 Schmidt, M. W., & Poli, S. (1998). Experimentally based water budgets for dehydrat-
244 ing slabs and consequences for arc magma generation. *Earth and Planetary Science*
245 *Letters*, *163*(1-4), 361–379.
- 246 Schwarz, G., & others. (1978). Estimating the dimension of a model. *Annals of Statis-*
247 *tics*, *6*(2), 461–464.

- 248 Scrucca, L., Fop, M., Murphy, T. B., & Raftery, A. E. (2016). Mclust 5: Clustering, clas-
 249 sification and density estimation using gaussian finite mixture models. *The R Jour-*
 250 *nal*, 8(1), 289.
- 251 Sizova, E., Gerya, T. V., Brown, M., & Perchuk, L. (2010). Subduction styles in the pre-
 252 cambrian: Insight from numerical experiments. *Lithos*, 116(3-4), 209–229.
- 253 Syracuse, E. M., & Abers, G. A. (2006). Global compilation of variations in slab depth
 254 beneath arc volcanoes and implications. *Geochemistry, Geophysics, Geosystems*, 7(5).
- 255 Syracuse, E. M., Keken, P. E. van, & Abers, G. A. (2010). The global range of subduc-
 256 tion zone thermal models. *Physics of the Earth and Planetary Interiors*, 183(1-2),
 257 73–90.
- 258 Turcotte, D. L., & Schubert, G. (2002). *Geodynamics*. Cambridge university press.
- 259 Vermeesch, P. (2018). IsoplotR: A free and open toolbox for geochronology. *Geoscience*
 260 *Frontiers*, 9(5), 1479–1493.
- 261 Wada, I., & Wang, K. (2009). Common depth of slab-mantle decoupling: Reconciling
 262 diversity and uniformity of subduction zones. *Geochemistry, Geophysics, Geosystems*,
 263 10(10).