Determining Mid-Ocean Ridge Geography from Upper Mantle Temperature

Xiyuan Bao[1], Tushar Mittal[2], and Carolina Lithgow-Bertelloni[1]

[1]University of California, Los Angeles
[2]Pennsylvania State University

May 01, 2024

Highlights

Determining Mid-Ocean Ridge Geography from Upper Mantle Temperature

Xiyuan Bao, Tushar Mittal, Carolina R. Lithgow-Bertelloni

- Mantle temperatures beneath global mid-ocean ridges exhibit inter-basin differences
- We use machine learning to predict the geographic location of ridge segments based on the sub-ridge upper mantle temperature
- The integrated history of convection and tectonics is recorded in the large-scale patterns observed at mid-ocean ridges

Determining Mid-Ocean Ridge Geography from Upper Mantle Temperature

Xiyuan Bao[a,*], Tushar Mittal[b], Carolina R. Lithgow-Bertelloni[a]

[a]*Department of Earth, Planetary, and Space Sciences, University of California Los Angeles, Los Angeles, CA, USA*
[b]*Department of Geosciences, Pennsylvania State University, University Park, PA, USA*

Abstract

In this study, we examine the influence of the mantle and large-scale tectonics on the global mid-ocean ridge (MOR) system. Using solely seismically-inferred upper mantle temperatures below the melting zone (260-600 km) and an interpretable machine learning model (Random Forest and Principal Component Analysis), we can predict *a priori* the location (ocean basin and ridge system) of global MOR with up to 90% accuracy. Two features provide >50% of the discriminative power: the temperature difference between the mid-layer (340-500 km) and other depths, and the depth-averaged temperature of the upper mantle. We suggest long-term (100s Myr) tectonic and convective processes left ample, distinct but hidden fingerprints in the mantle that allow us to separate regions at $\sim$ 1000 km scale. Our result implies that the large-scale geophysical and geochemical differences observed along the MOR system are reflective, not primarily of shallow processes associated with melting, but of the integrated long-term tectonic and convective history

*Corresponding author
Email address: xiyuanbao@g.ucla.edu (Xiyuan Bao)

that determines the present-day upper mantle temperature structure.

Keywords: Mid-ocean Ridge, Potential Temperature, Mantle Convection, Random Forest

2020 MSC: 86-00

1. Introduction

The 60,000 km-long chain of mid-ocean ridges (MOR) is the most visible surface manifestation of plate tectonics and mantle flow. Deep (>250 km depth) mantle material is fed to ridges by largely passive convective currents, resulting in decompression melting at depths < 150 km, and the generation of new oceanic lithospheric plates. The plate tectonic factory is hence directly connected not only to the present-day structure of the mantle under ridges but also to the integrated convective and tectonic history of each ocean basin. For instance, since the breakup of Pangea, the circum-Pacific subduction girdle has produced an influx of cold downwelling slabs towards the mantle beneath the Atlantic and Indian Ocean basins and a relative absence in the Pacific basin (Supplementary Movie S1, Richards and Engebretson, 1992; Lithgow-Bertelloni and Richards, 1998; Mao and Zhong, 2018; Müller et al., 2019; Doucet et al., 2020)., even when considering intra-oceanic subduction in the Pacific basin (e.g., van de Lagemaat and van Hinsbergen, 2023). The downgoing Pacific slabs cool the mantle, and the downwelling flow generates a passive return upwelling flow at ridges.

The mantle convective and tectonic history, integrated over the last few hundred million years, will determine the average temperature of the upper mantle today (e.g., Conrad et al., 2013; Karlsen et al., 2021). Superposed on these average temperatures are additional dynamical processes. For instance, upwelling flow in the lowermost mantle above the Pacific and African Large Low Shear Velocity Provinces (LLSVPs) may contribute to the overall thermal structure of the upper mantle in some regions (MacLeod et al., 2023). More importantly, as LLSVPs are directly associated with most present-day hotspots and mantle plumes, the interaction of mid-ocean ridges with nearby (within a few 1000 km) mantle plumes can significantly affect the ridge thermal structure, volatile content (CO_2, H_2O content), and basalt geochemistry (especially radiogenic isotopes, e.g., Ito et al., 2003; Gibson and Richards, 2018).

We thus hypothesize that the tectonic and convective histories over the past few 100 Myr are reflected in the geophysical and geochemical characteristics of the ridge systems of individual ocean basins. For example, the Pacific ridges have a systematically higher spreading rate and are shallower (Fig. 1 a, b) than the mid-Atlantic ridge system, with the Indian Ocean ridge segments having intermediate values (e.g., Gale et al., 2014). Similar differences exist in the major element composition of mid-ocean ridge basalts (MORBs, Gale et al., 2014). Previous work on the origin of these inter-basin geochemical and geophysical differences has focused on the correlation amongst spreading rate, ridge depth, and MORB major and trace element chemistry (e.g., Klein and Langmuir, 1987; Sinton and Detrick, 1992; Brandl et al., 2013;Gale et al., 2014; Niu, 2016), inferences on the mantle temperature (Klein and Langmuir, 1987; Brown Krein et al., 2021), composition of the mantle source region (Niu and O'Hara, 2008), and melt-rock interaction during magma transport (Kimura and Sano, 2012). The premise of these studies is that the degree of partial melting and the nature of melt transport in the melting column is the primary control of the observed variability.

Instead, taking a geodynamics perspective, we posit that the differences in ridge characteristics at the ocean basin scale are a consequence not only of shallow melting but of deep mantle structure integrating over the ocean basin's convective history. Focusing only on shallow processes obscures the large-scale integrative role of mantle convection and tectonic history in shaping the source of mantle melting at the ridge on multiple spatio-temporal scales. However, it is very challenging to analyze the critical role of deep processes from existing approaches since the inferences regarding MORB geochemistry and mantle source potential temperature (T_P) are strongly affected by the poorly constrained details of the melting process at shallow depths (Stracke, 2021), such as the extent of melt channelization (Spiegelman and Kelemen, 2003; Keller et al., 2017; Brown Krein et al., 2021). For instance, current petrological estimates of the ridge potential temperature (T_P) disagree both in absolute value and inferred spatial patterns. Brandl et al. (2013) and Dalton et al. (2014) see a hotter Pacific compared to the Atlantic and Indian Ocean basins, while Brown Krein et al. (2021) see no distinct hemispheric difference. *In this study, we take an alternative, data-driven approach to search for unique fingerprints of the mantle flow processes in the deep upper mantle (260 - 600 km depth) temperature under MORs.* These variations would serve as inputs for the shallow melting processes that eventually give rise to the observed geochemical variations in MORB lavas.

Specifically, our work addresses the following question: *Can the temperature of the entire upper mantle below the melting zone accurately classify a* *priori the oceanic basin where ridge segments are found?* Within the context of this study, we will be using ocean basins interchangeably with ridge systems in those basins since the ridges are representative of the integrated convective history of the basin and are fed dominantly by passive mantle flow. This is true even if some ridge segments have an active component as suggested by Rowley et al., 2016. By focusing on ridges instead of geometric centers of ocean basins, we avoid contamination from smearing near complex features like subduction zones, which have very large velocity and thermal anomalies. Finally, as we directly sample the passive upwelling part of the mantle underneath the ridges, we avoid the potential effects of mixing at varying length and time scales, that the geometric center of the ocean basin may suffer from.

Given the significant overlap in mantle potential temperatures across basins (Fig. 1c, using results from Bao et al., 2022, see section 2.2), the answer to our primary question is not obvious. Furthermore, this question may seem superfluous for the present day, given that we already have the geographical data for each ridge segment. However, if we can identify unique sub-ridge mantle temperature fingerprints associated with each basin and even sub-basin-scale ridge systems, we can analyze them with respect to the basin's tectonic and convective history. This would, in turn, aid our understanding of how different mantle flow processes contribute to sub-ridge thermal structure. Overall, our study helps in addressing a fundamental geodynamics question: What is the dominant reason for the observed geophysical and geochemical differences in ridge systems - shallow melting or deep mantle processes?

To address our primary science question, we construct a model for predicting the basin in which ridge segments are located, using only the temperature of the upper mantle inferred from a full waveform seismic tomography model and self-consistent thermodynamics (Bao et al., 2022), combined with the power of an interpretable supervised classification method – the random forest (RF) algorithm (Breiman, 2001). Our end goal is not to develop a model for extrapolation to other geographic locations or time periods. Instead, we use the model to assess if the thermal structure beneath the ridges contains enough information to distinguish among ocean basins (or at spatial scales of at least $\sim$ 1000 km) without any prior geographic information. If we can find a high accuracy and interpretable model, this would clearly indicate that different ocean basins have unique upper mantle thermal fingerprints associated with specific mantle dynamics processes.

Our work builds upon earlier attempts to understand the deep mantle contribution to the global ridge system. Early studies, e.g., Ray and Anderson (1994) explored the connection between mantle seismic velocity and ridges, as shear wave speeds are particularly sensitive to temperature. However, Ray and Anderson (1994) were limited by the resolution of the global tomography and sparse mineral physics data and thermodynamic modeling available at the time. They could not infer temperatures directly from the seismic velocities. Dalton et al. (2014) provided a big step forward by using thermodynamic models of the physical properties of mantle rocks to infer mantle temperature at 300 km depth below the ridges from global seismic tomography. They found Pacific ridges to be hotter than those in the Indian and Atlantic oceans. Rowley et al. (2016) also found a possible contribution from active, hotter mantle upwellings to the faster-spreading rates at the East Pacific Rise. While these studies provide important clues regarding the role of convective and tectonic processes on seafloor spreading and MORB geochemistry, they lack predictive power (uniqueness of the mantle fingerprint) or a direct connection to convective and tectonic processes. Furthermore, full waveform seismic tomography models (as used here) from the past decade (e.g., French and Romanowicz, 2014) provide a more robust and faithful estimate of the amplitude of seismic anomalies, which is crucial for inferences of temperatures.

Thematically, our work is a counterpart to the recent study by Stracke et al. (2022), who used non-linear dimension reduction and clustering analysis on multiple isotopic data for global MORBs and Ocean Island Basalts. They showed that ridges and hotspots potentially sample distinct sub-basin-scale isotopic heterogeneities, thus highlighting the role of deep mantle processes in controlling ridge composition. Section 2 describes the datasets and analysis methods we use in this study, followed by the results of the random forest analysis in Section 3. Section 4 uses these results to discuss the main implications of our results in the context of the importance of shallow vs. deep mantle processes for ridges.

2. Materials and Methods

2.1. Ridge Database

To test our hypothesis that the unique sub-ridge temperature fingerprints exist, we start by sampling mantle properties underlying MOR segments in the three major ocean basins (Pacific, Atlantic, Indian). We use the segment definitions from Gale et al. (2014) database with some filtering (choosing 655 out of 711 mid-ocean ridge segments, 771 in total if ridges in back-arc basins are included) to 1) avoid more complex tectonic settings (back-arc basins and ultra-slow ridges) and 2) simplify classification. The Gale et al. (2014) ridge segments are determined based on along-ridge axial depth variations, ridge offsets, transform faults, and non-transform offsets. Thus, using these segments is a reasonable choice for our question of interest rather than a uniform sampling per km of the ridge since each segment would correspond to a unique tectonic/convective regime. The average length of the mid-ocean ridge segments is 78 km. However, our analysis is not focused on interpreting the inferred temperature below any neighboring individual ridge segment at this ~ 100 km horizontal spacing. The inherent lateral averaging (smoothing and smearing) and resolution of global seismic tomography inversions preclude it and the results are only representative of the regional scale thermal structure ($>\sim 500$ km). Instead, we focus on ridge temperature variation with lateral scales of ~ 1000 km and ~ 100 km scale vertically (Fig. 2). We also test the robustness of our conclusions by doing ridge-basin classification for the entire mid-ocean ridge database with 711 segments (including smaller basins in the Arctic, Caribbean, and Red Sea), as well as finer sub-basin ridge system classification, but still at > 1000 km horizontal scale (discussed in section 3.2).

2.2. Temperature Inference

Following Bao et al. (2022), we extract shear wave seismic velocity from tomographic model SEMCUB-WM1 (French and Romanowicz, 2014) and convert it to temperature. We validate our results with 4 additional global tomographic models (Ritsema et al., 2011; Simmons et al., 2010; Schaeffer and Lebedev, 2013; Debayle et al., 2016). We extract velocity anomalies directly beneath each ridge segment, without any lateral averaging, from 260 to 600 km depth in 20 km intervals. This depth interval allows us to capture sufficient information given the radial spline basis functions used in recent global tomography models (e.g., French and Romanowicz, 2014). We focus on depths below 260 km to avoid the strongly reduced seismic velocities due to seismic attenuation, potentially caused by partial melt. Dry melting starts $<\sim$ 100 km depth beneath the ridge and at $<\sim$ 200 km in the presence of volatiles (Keller et al., 2017 and references therein). A depth $>$ 260 km is sufficient to avoid even the melting-influenced regions of intraplate volcanism, as seen seismically (Debayle et al., 2020) and geochemically (Ball et al., 2021). Because the velocity to temperature conversion is non-linear (Bao et al., 2022, Supplementary Fig. 1), we convert the shear-wave velocity anomalies to temperature using HeFESTo (Stixrude and Lithgow-Bertelloni, 2005, 2011, 2022, 2024). HeFESTo is a self-consistent thermodynamic model of the equilibrium phase assemblage of mantle minerals and their physical properties at a given pressure, temperature, and fixed bulk composition.

An obvious question here is whether this velocity to temperature conversion is necessary and if one could directly use the seismic velocity anomalies for ridge-basin classification from mantle tomography. However, we posit that analyzing the temperature structure is more physically informative because: (1) Seismic velocity of the sub-ridge mantle is dominantly thermal at the global and inter-basin spatial scales of interest in this study ($\sim$ 1000 km). At individual ridge segment compositional heterogeneities are more likely to play a role (Dalton et al., 2014); (2) Conversion of seismic velocity to temperature is not a constant or empirical relationship (Birch, 1969; Schubert et al., 1976), but is instead non-linear (and depth dependent) in a self-consistent thermodynamic model (Bao et al., 2022) (also see Supplement Figure), and; (3) Variations in temperature anomalies are more physically meaningful and interpretable in terms of mantle convective processes and tectonic histories in contrast to seismic velocity anomalies.

Since our focus is on ocean basin scale features, we use the conservative premise that the upper mantle is compositionally homogeneous to first order, consisting of Depleted MORB Mantle (DMM, Workman and Hart, 2005). Our final temperature data for the ridge segment catalog is high-dimensional (18 uniformally spaced depth layers between 260 to 600 km depth per each ridge segment, Fig. 2), which demands a strategy for dimensional reduction as discussed below (section 2.3).

2.3. Data Processing and Classification

2.3.1. Linear Classifier

We first use a linear classifier, i.e., multinomial logistic regression, to predict the basin where each ridge segment is located based on the MOR mantle temperature profiles . Specifically, we try to find curves in the space of each input pair (e.g., between temperature at 2 depths) to separate out the different basins. A softmax function (Bridle, 1989) is used to find the maximum probability of the particular class and to give a prediction. However, this yields low accuracy irrespective of whether we use dimensionality reduction (60% accuracy) or not (65% accuracy). This suggests that there is no clear, linear predictive separation between each ocean basin ridge segments (e.g.,

Fig. 3). The high-dimensional nature of the raw data (i.e., 18 depth layers) also makes the problem challenging. Thus, we need a higher-order machine learning model that can handle both linear and highly nonlinear relationships and remain interpretable. We further desire that the model features be physically meaningful quantities that can be related to dynamical processes, such as the average temperature of the upper mantle (related to long-term plate organization, e.g., Gurnis, 1988), and the difference in temperature between layers which can be linked to various convective length scales.

Dimensional reduction using Principal Component Analysis (PCA, Jolliffe, 2002) satisfies the requirements set above for optimal, interpretable classification. PCA is a commonly used method for high-dimensional datasets and calculates orthogonal principal components (PCs, Fig. 4). Each PC is a linear combination and weighted sum of the normalized T_{P} at the 18 distinct depths under each ridge segment. That is, $PC^i = \sum W_d^i \hat{T}_{p_d}$, $\hat{T}_{p_d} = (T_{p_d} - \mu_d)/\sigma_d$, where W_d^i is the weight for ith PC at depth d; $\hat{T}_{p_d}$ and T_{p_d} are the normalized and original potential temperature at depth d, respectively; μ_d and σ_d are the average potential temperature and standard deviation for all ridge segments at depth d, respectively. We normalize and rescale the original temperatures (from T_{p_d} to $\hat{T}_{p_d}$) for each depth before using it in the PCA calculation, to have zero mean and unit variance to achieve better performance (Duda et al., 1973). PCs are sorted from large to small values based on how much variance they can represent in the data. PC1 covers the largest variance of the data, PC2 the second largest, and so on for the remaining principal components. Mathematically, PCs are the eigenvectors of the covariance matrix C_X of the (normalized) original data matrix X (with dimensions of $18 \times N$ for 18 depths and $N = 655$ ridge segments accordingly). The eigenvectors are sorted by the absolute value of the corresponding eigenvalues. That is, if an eigenvalue is near zero, the corresponding eigenvector is not important. Because we have 18 depths, there will be 18 PCs in total. The covariance matrix C_X is symmetric, therefore PCs are mutually orthogonal. In other words, the PCs are just the result of rotation of the original space (18 depths) into a space where correlated depth slices are combined together into a new dimension. As a result, a positive weight (within a given PC) means the temperature at that depth contributes positively to that PC (or the rotation along that depth), and vice versa for negative values. Analyzing the PCs that capture the main variance ($\sim 99\%$) equates to projecting the data to a reduced dimensional space. We determine the optimal number of PCs to be used in our analysis based on their final performance in the subsequent machine-learning model, instead of just the variance of the data when projected to each PC. This ensures that we keep adequate information of the mantle temperature structure for the classification model to determine ridge location.

2.3.2. Non-Linear Classifier : Random Forest

Given the poor performance of a linear classifier even with PCs as inputs ($\sim 60\%$), we choose to use a nonlinear supervised classifier - Random Forest (Breiman, 2001) for our primary analysis here. Using the PCs as inputs, we train a Random Forest (RF) model to predict the ocean basin in which ridge segments are located. RF is a robust classification algorithm (reduced sensitivity to overfitting) and generates interpretable decision trees (Fig. 5a). RF consists of a decision tree generation algorithm, which chooses only one feature (i.e., PC) at each node and divides the data into two branches based on a cutoff value. To determine what PCs to use and their cutoff value for each tree branch, the tree algorithm calculates the entropy or Gini impurity G for each possible PC and cutoff combination. At each node, we have $G = \sum_k p_k(1 - p_k)$, where p_k is the proportion of each class (i.e., ocean basin) k. A low entropy or Gini impurity measure indicates that the sub-node/branch would be dominated by one class and it is thus a good choice for dividing the tree. This process is repeated until the whole dataset is classified by a tree consisting of many branches. Overall, the algorithm optimizes the PC selection and cutoffs at each branching point to match the input classification labels (here the ridge basins of origin). For each input datapoint consisting of a set of PC values, the final classification is the value of each end node (leaf node) that the datapoint reaches after traversing the trained tree model (e.g., Fig. 5b). A key feature of the tree-based classification algorithms is that they make it easier to understand the classification and the importance of each input feature in the final predictive classification model. RF generates a series of decision trees (here N = 20) as a forest and takes the predicted probability of the segment in a certain basin averaged from each tree. There are two built-in levels of randomness to avoid overfitting: 1) Random resampling of the dataset via bootstrapping when training each tree, and 2) PC selection from a randomly selected subset of PCs when growing the tree.

The nonlinear nature of the algorithm and its randomness enable RF to handle the complicated ridge database robustly. To further avoid overfitting and improve the robustness of the prediction, we also randomly split the input

PC data into training (80%) and testing (20%) sets. We repeat this 50 times to calculate the average classification accuracy. The modeling pipeline is constructed using Orange which enables visual programming for data mining (Demšar et al., 2013). Note that with PC as input of Random Forest, our model is similar to the Rotational Forest. In Rotational Forest, the raw feature is split into subsets randomly, and then PCA is performed for each subset. The result is then used as input for the RF algorithm (Rodriguez et al., 2006).

When we visualize data in PC pair space (or input temperature variable space) with scatter plots in Orange (Demšar et al., 2013), it can compute the most informative projections (e.g., the 2D space that best separates different classes). For each point, Orange finds 10 nearest neighbors in the projected 2-D space, e.g., two PCs. It then checks the number of points out of 10, with the same ocean basin. The averaged number across the neighborhood of all points gives the final score, and we consider the PC (or temperature) pair with the highest score the most informative projection. In Figure 3, we show the results of this analysis for a pair of input temperature data variables. These results quantitatively illustrate that a 2D parameter space is insufficient to distinguish the ocean basin of origin of each ridge segment.

2.3.3. Summary: Classification analysis

In conclusion, we need a combination of PCA and RF for our classification analysis, since the separation of the ridge segments based on the depth-dependent temperature is challenging with a linear classifier alone. Due to inherent smoothing in seismic tomography and the effects of varying geodynamic processes, the temperatures in neighboring depth bins can be strongly correlated. Hence, using temperatures at 18 depths obscures the length-scale analysis and physical interpretation of the results. Instead, we use PCA to effectively determine the primary, uncorrelated information in the data and which length-scales are distinct.

We want to re-emphasize that the end goal of the analysis is not to just find a predictive model with the highest accuracy. Instead, we want to test if it is possible to find a model that accurately predicts the geographical location of ridge systems (and the segments within them to at least at $\sim$ 1000 km scale) using only the sub-ridge structure. While the mid ocean ridges are largely fed by passive mantle upwellings (c.f. section 1), our data-driven feature selection does not preclude the existence of temperature heterogeneities in the upper mantle along the ridge at the ocean-basin scale (Cammarano et al., 2003; Dalton et al., 2014). While our primary analysis is focused on inferred temperatures, we repeat the analysis with velocity anomalies (using the corresponding PCs and RF model) for completeness.

3. Results

3.1. Potential Temperature

Figure 2a shows the map of inferred T_{P} averaged over 260-600 km depths. The mean and median T_{P} of the Pacific are the hottest overall, while those of the Indian and Atlantic basins overlap (Fig. 1c), consistent with Dalton et al. (2014). The modal T_{P} for Pacific ridges is similar to that of Indian ridges but slightly hotter than that of Atlantic ridges. Overall, Indian ridges have T_{P} distribution intermediate between Pacific and Atlantic ridges. We see regional in-basin lateral temperature variations similar to Dalton et al.

(2014) and Bao et al. (2022). While the map (Fig. 2a) and overall statistics (Fig. 1c) already reveal some differences among basins, we observe additional multi-scale vertical variations, which we discuss in section 4.3 (Fig. 2b).

3.2. Principal Components and Random Forest

We find that the first 5 PCs cover $> 99\%$ of the variance in the temperature data (Fig. 6a). The proportion of variance explained by each PC decreases dramatically from more than 75% for PC1 to less than 1% for PC5. To understand what each PC represents physically, in Fig. 4a, we show the weighting coefficients of the linear combinations of PCs using the weight matrix of the first 5 PCs. For PC1, the weights are ~ 0.2 at all depths. Thus, PC1 corresponds to the scaled average T_{P} over all depths. Other PCs have an average weighting of 0, meaning they emphasize the T_{P} differences at depth for length scales smaller than the whole upper mantle. For example, the weighting coefficients for PC2 decrease from 0.3 to -0.3 from 260 km to 600 km, essentially giving the difference in T_{P} between the upper half of the upper mantle (260-420 km) and the transition zone (440-600 km). The coefficients for PC3 are positive around 400 km (340-500 km) and negative at the top (260-320 km) and bottom (520-600 km); thus PC3 quantifies the contrast between mid-upper mantle depths (340-500 km) and other depths (especially $<$300 km, where the weight is the most negative at about -0.5). Finally, PC4 and PC5 represent variations at smaller length scales ($\leq$80 km). The first 5 PC values for all ridge segments are shown in Fig. 2b.

Choice of PCs: PC1, or essentially the average upper mantle T_{P}, shows substantial overlap across basins around 1300-1500 °C (Fig. 1d), and it is insufficient for accurate basin classification. As PC1 is only the bulk temperature of the upper mantle, information at smaller length scales (through other PCs) is required to distinguish ridges from basins with similar bulk temperature from each other. To have a parsimonious model, we first try to predict the basin geography with just one other PC by finding the most informative 2-D projection, which gives the best classification accuracy among all PC pairs. We find that this is the PC1 vs. PC3 projection shown in Fig. 4b. The Pacific segments lie primarily on the right of the projection (PC1> -4), while the Atlantic can have extreme PC3 values (> 2 or < -2). Although one can approximately predict ocean basins based on this zoning, the PC1 and PC3 in each basin still overlap significantly. Thus, the predictive accuracy is less than 60% and we need more PCs and length scale information. The zoning in Fig. 4b also reinforces the need for non-linear classifiers since the boundary between different ocean basins is curved and complex.

To determine the best number of PCs in the RF model, we add one PC at a time, in the order of descending variance covered (e.g., PC1, PC1+PC2, PC1+PC2+PC3, and so forth), and calculate the classification accuracy as a function of the number of PCs (Fig. 6b). Not surprisingly, classification accuracy generally increases with more PCs. However, the increased accuracy gain generally reduces as the PC index increases. Three PCs are enough to achieve 70% classification accuracy. To reach 80% accuracy, we must include PC1 to PC5 (accuracy = 82%). Since adding more PCs does not significantly improve the accuracy, we will use the first 5 PCs for the subsequent analysis. We get prediction accuracies from 75% (Pacific) to 90% (Atlantic), shown in Table 1 as the confusion matrix.

Trained tree model : A typical example of how PCs work in RF is shown in Fig. 5b, which shows one decision tree of RF. At the root node where we have all samples (a random subset of all ridge segments), RF finds that PC1 can best split the data by bifurcating the samples at PC1 = 4.99 so that the child node with PC1 $>$ 4.99 (node A) is dominated by Pacific ridges. The other child node (node B) with PC1 $\leq$ 4.99 has fewer Pacific samples. In this way, the child nodes are more uniform and the entropy of the child nodes is minimized. Next, a random subset of PC candidates is generated at node A, and RF chooses to use PC4 = -0.22 to further bifurcate node A to A1 and A2. Consequently, the child node A1 has an even higher portion of Pacific segments than node A, while node A2 only has samples from the Atlantic Ocean. Similarly, node B is bifurcated at PC3 = 0.25 to B1 and B2 such that B1 has very few Pacific samples. A similar procedure is applied to A1, B1, and B2 with PC2, PC5, and PC1, respectively, and their child nodes repeatedly until the child node has four samples (or less) or samples in the child node are purely from one basin (like A2). We call these end nodes leaf nodes. Overall, as the decision tree grows from the root node to the leaf nodes, we gradually minimize the entropy at the next level and have one basin dominate each leaf node.

Classification robustness : We change the training/testing splitting over a wide range from 80%/20% to 50%/50% and even 25%/75%, and still achieve high classification accuracy (82%, 77% and 70%, respectively). We find that the classification accuracy is robust for all other tomographic models examined ranging from >83% (Debayle et al., 2016) to 90% (Ritsema et al., 2011; Simmons et al., 2010; Schaeffer and Lebedev, 2013). This higher accuracy may result from the global tomographic models explored here containing less heterogeneity at shorter wavelengths at depth (c.f., discussion in Meschede and Romanowicz, 2015). Consequently, these models suppress in-basin temperature variation and emphasize inter-basin differences. We also notice the weight matrix is reasonably consistent across models, i.e., PC1 always gives the average while each of the other 4 PCs gives the differences of the same layers. However, the sign of weights in certain PCs may flip (Fig. 7). These results are not unexpected as global tomographic models are broadly consistent with each other in the upper mantle. In addition, we can obtain a slightly improved classification accuracy (from 82 to 86%) if we average the inferred T_{P} at every depth in a disc, with radius $R = 500$ km centered at each ridge segment. The local average temperature beneath the ridge segment incorporates additional environmental information (i.e., cold and hot anomalies) and suppresses in-basin small-scale lateral variations. Using the seismic velocity anomalies for classification directly (without a temperature conversion), we obtain a similar accuracy - 86% instead of 82% - for the 80/20 training and testing data split. This illustrates that the velocity-temperature conversion does not, by itself, impart a bias in terms of our results. The results based on temperature are easier to physically interpret in terms of mantle dynamics.

Results with sub-basins : In our main results, we focus on the classification of three large main basins and exclude other small regions such as the Arctic, Red Sea, and Caribbean ridge systems. However, including them as additional ocean basins in our classification model only leads to negligible decrease (1%) in classification accuracy. Therefore, our primary conclusions do not change with the full mid-ocean ridge database of 711 segments. We further test our ability to predict smaller tectonic units within ocean basins (sub-basin ridge systems, e.g., East Pacific Rise). To do this, we slightly simplify the groups in the ridge database by merging the Chile Ridge with the Pacific-Antarctic Ridge and the Atlantic-Antarctic Ridge with the Mid-Atlantic Ridge given their spatial proximity and the small number of ridge segments in these ridge regions. We then obtain a sub-basin ridge system map as shown in Fig. 8. Using this for classification analysis (80/20 training and testing split), we obtain a good accuracy of 74%. Using the local temperature averaged inside a 500 km-radius disc surrounding each ridge segment, we get 80% accuracy because lateral variations within each ridge system are suppressed. These results suggest that the ocean-basin structure is distinct not only between Pacific, Atlantic, and Indian ocean, but also on the sub-basin scale.

4. Discussion

4.1. Robustness of our results and overall interpretation

Our results show that we can determine the ocean basin of origin with 80 to 90% accuracy. Note that with 25% of data for training and the rest for testing, we have an effective ridge segment scale closer to 500 km (instead of $\sim$ 100 km of the original data), so neighboring ridge segments should contain much more independent information. Nonetheless, we are still able to obtain 70% accuracy. The robustness of our results suggests that the sub-ridge mantle temperature is distinct across basins and could be an excellent indicator of large-scale convective contributions to surface differences in the MOR system. This directly addresses the primary motivating science question for our study affirmatively.

Conceptually, our classification model can be regarded as a non-linear function that takes the present sub-ridge mantle structure as input, decodes the hidden fingerprint of the integrated records of past tectonic and convective history, and converts this fingerprint into location information of the ridges in terms of the basin of origin or smaller tectonic units, such as sub-basin ridge systems. *The hidden fingerprint from the deep mantle is sufficient to provide robust long-wavelength spatial information without the need for any shallow mantle or surface observations such as MORB chemistry, spreading rate, and ridge depth.* In the mantle tomography models, we are averaging laterally at at least the $\sim$ 500 km scale given the inherent smoothing and smearing in seismic inversion methods. Consequently, the sub-ridge mantle is representative of a large section of the ocean basin where the ridge segments reside.

4.2. Random Forest Feature importance

In our analysis, each principal component represents the sub-ridge temperature heterogeneity at different length scales, ranging from the entire upper mantle (PC1) to half (PC2) to 1/3 (PC3) of the upper mantle, and even smaller depth intervals (PC4 and PC5). Our results thus reveal the length-scale of thermal heterogeneity subsisting in the mantle and contributing to the integrated convective record. To assess which features (i.e. PCs) contribute to classification accuracy the most, we use feature importance analysis methods. For non-linear classifiers such as RF, we can use the permutation feature importance method (Breiman, 2001) to compute feature importance. This approach randomly permutates the data of a given PC and computes the corresponding decrease in classification accuracy with respect to the default case (Fig. 5a). We find that PC3 is the most critical feature with >30% importance, while PC1 is the second most important (>20%). Thus, PC3 and PC1 together provide more than half of the discriminative power of the 5 PCs. The fact that feature importance of one PC can be as high as 30% suggests a random dataset (like the original data after the permutation) would not provide enough information for classification.

The fact that PC1, which represents average mantle temperature (see discussion below), is a key feature is not surprising given previous work. Multiple studies have suggested that hemispherical temperature differences in the upper mantle may result from past subduction zone locations (i.e., the Indo-Atlantic box and Pacific box, Masters et al., 1982; Richards and Engebretson, 1992;Ricard et al., 1993; Lithgow-Bertelloni and Richards, 1998; Davaille et al., 2005). The key novelty of in our study is three fold: 1) Using a purely data-driven approach, we see such hemispherical differences from the global sub-ridge temperatures inferred from modern seismic tomography in the absence of any *a priori* geographical or surface morphological information. We can therefore avoid the limited and sampling bias affected petrological data (e.g., Brandl et al., 2013), while still being able to compare our inferred ridge temperature to global petrological data (Bao et al., 2022); 2) By using the temperature as a function of depth, rather than only the average temperature over the upper mantle or selected at a particular depth (e.g., Dalton et al., 2014; Bao et al., 2022), we can analyze the overall large-scale thermal structure below the melting depth. Even sophisticated ML models would fail if we used only the depth-averaged sub-ridge temperatures, because the values in each basin overlap significantly (Fig. 1c); 3) by adopting the state-of-the-art RF model combined with PCA, we can perform multi-scale analysis not just laterally (inter-basin versus each ridge system), but vertically, using the reduced dimensionality provided by PCA.

4.3. Physical interpretation of PCs

4.3.1. Principal Component 1

PC1, the average T_{P} over all depths and the second most important feature, broadly represents the current convective vigor of the upper mantle column. The distinct hemispherical pattern (higher PC1 in the Pacific, Fig. 2, 1c) is consistent with previous studies (Brandl et al., 2013; Dalton et al., 2014) and can be linked to the history of subduction for the past 200 Myr (c.f. Supplementary Fig. 2).

For instance, the Pacific ocean evolved from the Panthalassic ocean. It was filled with in-basin spreading ridges and was also surrounded by an outward subduction girdle predating the formation of Pangea (~300 Ma, Doucet et al., 2020, c.f. Supplementary Text S1 for more information). Intra-oceanic subduction, if present, was not extensive and confined to the edges of the basin (e.g., van de Lagemaat and van Hinsbergen, 2023). Thus, the hemisphere dominated by the Pacific plate and its mid-ocean ridges and has not been cooled significantly by past subduction (Pacific domain). Since that period there was also significant intraplate hotspot volcanism in the Pacific resulting in large oceanic plateaus also potentially reflective of the higher basin temperature (Bryan and Ferrari, 2013).

In contrast, the Atlantic region developed from the rifting of Pangea ~180 Ma and the formation of the mid-Atlantic ridge system. The Indian Ocean has a more complex tectonic history – it has undergone in-basin subduction, ridge spreading, and the closure of the Tethys (Müller et al., 2019). Thus, the Indian and Atlantic oceans and associated continental plates define the other hemisphere (African domain) where we instead expect the mantle to have been cooled by Tethyan and Pacific basin slabs. It has been suggested that the residual topography and lithospheric thickness potentially present a similar hemispherical pattern (Stewart et al., 2023), which might be linked to the differences between the corresponding mantle domains (separated by the dashed line in Supplementary Fig. 2). Overall, the different tectonic histories, in particular, the presence or absence of in-basin subduction and the subduction of slabs away from one basin and towards another, can change the first-order thermal structure of the mantle under each basin and is reflected in the MOR temperature today (Fig. 1c).

This tectonic history forming a subduction girdle, gives rise to a degree-2 convection regime with limited material transfer between the Pacific and African mantle domains (Conrad et al., 2013). This degree-2 pattern is also reflected in the presence of the antipodal Pacific and African Large Low Shear Velocity Provinces (LLSVPs) above the core-mantle boundary (Supplementary Fig. 2). The origin and specific nature of the LLSVPs are beyond the scope of this discussion. We speculate that their presence and differences between the LLSVPs may have influenced the thermal structure of their mantle domains, especially given their association with plumes (Koppers et al., 2021; Weis et al., 2023; MacLeod et al., 2023).

Beyond recent (< 200 My) subduction history, the long-term convective and tectonic history, such as the presence of supercontinents, may also alter the thermal structure of the mantle under each basin (Gurnis, 1988; Jellinek and Lenardic, 2009; O'Neill et al., 2009; Lenardic et al., 2011). Karlsen et al. (2021) argue that Rodinia, a longer-lived (1.1-0.7 Ga) supercontinent, might have allowed more heat to accumulate under the Pacific mantle domain in contrast to the impact of the shorter-lived Pangea (300-180 Ma) on the African domain. The additional supercontinent insulation may be partially responsible for the present-day hemispherical temperature difference T_{P} at depth (Fig.2a), despite faster cooling in the Pacific due to higher spreading rates after the breakup of Pangea (Karlsen et al., 2021).

Besides the impact on basin-wide average temperature and PC1, past subduction may also explain regional low PC1 values. For instance, a coherent slab-like structure has been observed beneath the Southeast Indian Ridge in seismic tomography models (Simmons et al., 2015) with a part of this potential slab remnant still trapped in the transition zone (Gurnis et al., 1998). This subduction event dates back to the Mesozoic and terminated near the edge of East Gondwana $\sim$ 140Ma. The presence of a trapped slab in the transition zone may explain the low temperatures and PC1 value of the associated nearby ridge ($T_{\mathrm{P}} \sim 1250$ °C, PC1$\sim$ -10, green box in Fig. 2) and contribute to the Indian basin's ridge system intermediate nature. These observations suggest a potentially persistent effect of subduction on upper mantle structure and temperature for over 100 Myr.

4.3.2. Principal Component 3

Interpreting PC3 – the difference in temperature between the middle of the mantle (340-500 km) and other depths – is more challenging. PC3 is more distinct among basins (Fig. 1f), and consequently, PC3 dominates the classification as indicated by the feature importance. The confusion matrix of our model (Table 1) shows that the smallest portion of mislabeled samples is between the Atlantic and the Indian region (around 8%) which is less than those related to the Pacific (usually >10%). This result illustrates that the hemispherical, first-order differences from PC1 are insufficient to determine whether a ridge segment is inside the Pacific Ocean (Fig. 1d). The modal PC3 value is highest in the Atlantic, then the Indian, and lowest in the Pacific (Fig. 1f). What controls the different temperatures at the length scale of 1/3 of the upper mantle across ocean basins? We posit that PC3 variations are potentially related to mantle flow associated with plume-ridge interaction as well as the interaction of the ridge with large-scale mantle upwellings (e.g., Ribe et al., 1995; Sleep, 2002; Gassmöller et al., 2016; Gibson and Richards, 2018). The nature and style of plume ridge interaction is distinct across different ocean basins and it can also affect regional scale mantle flow (e.g., Reunion mantle plume and Indian ocean ridge system, Barruol et al., 2019). A detailed analysis of the physical interpretation of PC3 will be discussed in a future companion paper.

4.3.3. PC2, PC4, PC5

PC2, the difference between the transition zone and mantle above the transition zone, is a feature that describes a larger length scale than PC3, and far larger than PC4/PC5. However, its importance is less than 20%, only about half and 80% of that of PC3 and PC1, respectively (Fig. 6c). Interestingly, we find that while PC2 covers ∼15% variance in contrast to 1% or less for PC4 and PC5 (Fig. 6a), the three PCs have similar feature importance (Fig. 6c). We attribute this to the fact that no single dynamical process dominates the difference at the three scales globally. Consequently, we observe no obvious modal/median difference among basins for PC4 and PC5 and PC2. But there are still differences between basins in terms of the shape of the density distribution, especially the distribution edges (Fig. 1e, g, h), so that each of PC2, PC4, and PC5 provides around 15% classification accuracy. A deeper physical understanding of the origin of these variations, such as the potential role of transition zone phase transitions and discontinuity topography, will be the subject of future work. We note it is hard to further improve classification accuracy to near 100% even when including more PCs. This may indicate the role of neglected dynamics such as those related to the melting process or heterogeneities shallower than 260 km depth.

5. Conclusions

With thermodynamically inferred upper mantle temperature and a robust machine learning model, we show that we can predict the ocean basin where modern ridge segments are located with at least >80% accuracy (Fig. 6b) using only temperature information from the mantle column beneath the ridge below the melting zone and no geographic information. Unlike surface ridge characteristics (depth, geochemical signals, etc.) which can be altered by complex shallow melting processes, modern upper mantle temperature is a proxy that records 100s Myr of history of plate tectonics and mantle convection (Supplementary Fig. 2). Our results help reveal the significant contribution of the deep mantle to large-scale MOR geophysical signals and suggest distinct inter-basin and even sub-basin deep mantle variations. The cluster analysis of ridge isotope geochemistry in Stracke et al. (2022) highlighted similar spatial mantle compositional variations. These two results together reinforce the idea that the mantle is recording the integrated tectonic and convective history of the last few hundred million years, leading to inter-basin and sub-basin temperature and isotopic variations. We anticipate that future studies may be able to utilize our results and analyze the disentangled effect of shallow melting processes on various geophysical, geochemical, and petrological observations at MORs.

6. Data Availability

The machine learning pipeline was constructed using Orange Demšar et al. (2013), available at `https://orangedatamining.com/` licensed under GNU version 3.0 or later. The compiled ridge database, including the seismic velocity and inferred temperature, along with the Orange workflow file, are available at `https://figshare.com/s/1cc8a5bc0d6faa469fe1` (DOI:10.6084/m9.figshare.22256035). The thermodynamic package HeFESTo Stixrude and Lithgow-Bertelloni (2005, 2011) is available at `https://github.com/stixrude/HeFESToRepository`, and the parameter set is available at `https://github.com/stixrude/HeFESTo_Parameters_310516`. The Movie S1 was created with Gplates portal at `http://portal.gplates.org/` (Müller et al., 2016).

7. Acknowledgments

This project is supported by the National Science Foundation under grant EAR-1900633 to CLB. CLB was also supported by the Louis B. and Martha

B. Slichter Chair in Geosciences fund. TM was supported by the Crosby Fellowship at MIT.

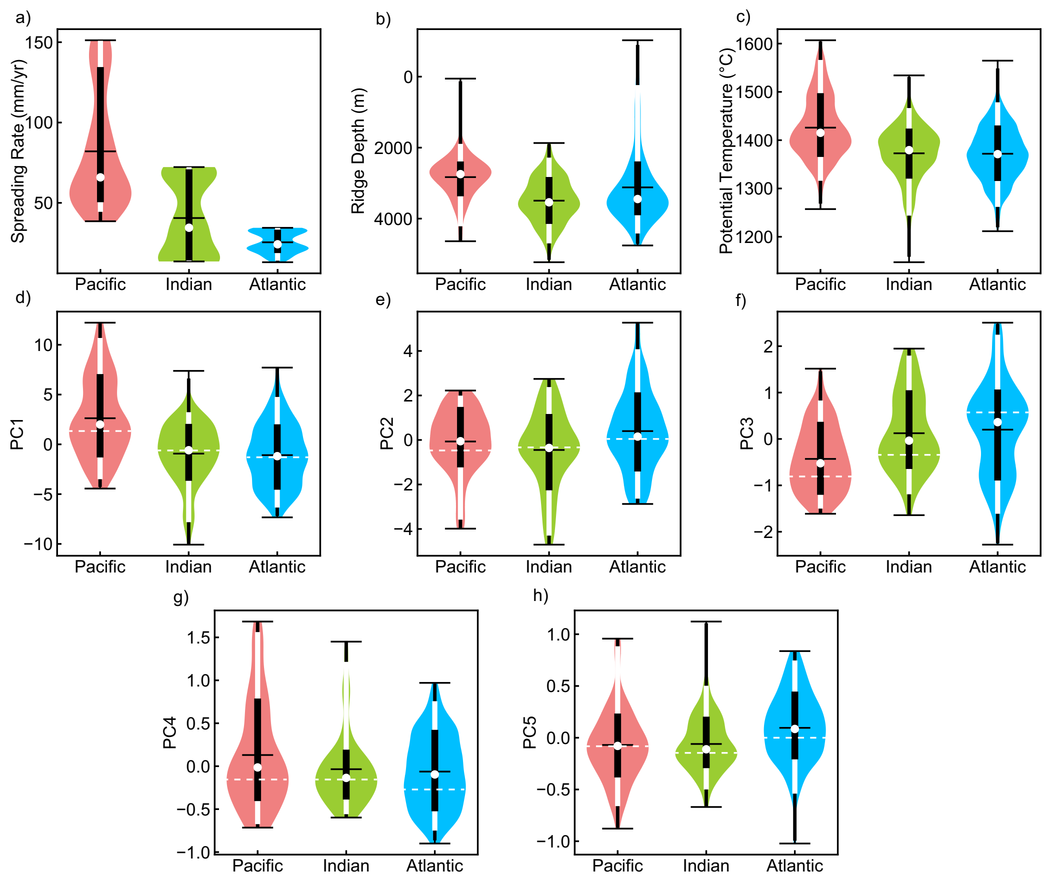


Figure 1: Violin plot of number density distribution of geophysical characteristics of each ocean basin. a) Spreading rate. b) Ridge Depth. c) Potential temperature stacked over all depths. d-h) PC1 to PC5. For each column, the horizontal bars are max, average, and min from top to bottom. The end points of vertical black and white bars are central 99, 95, 68 percentile from the median (white point). PC1 (d) and PC3 (f) have modal value position (dashed line) more distinct in the three basins, while PC2 (e), PC4 (g) and PC5 (h) have indistinguishable modal value positions in the three basins.

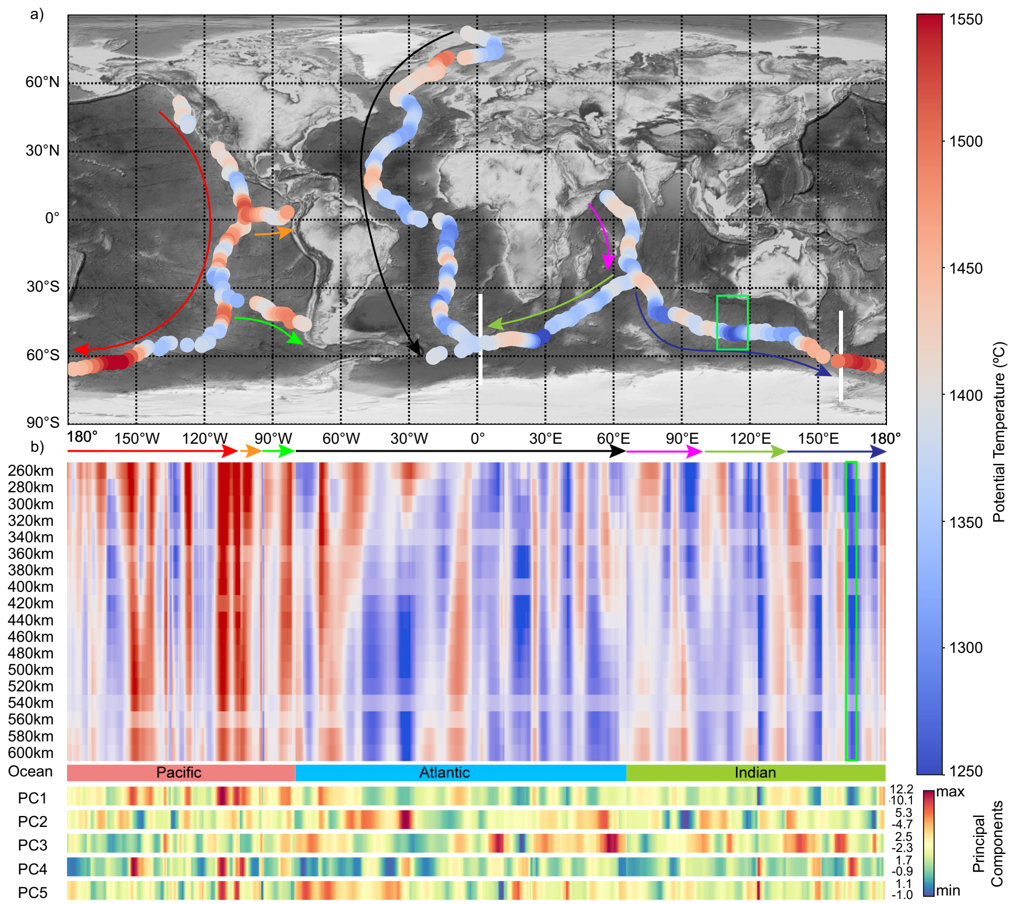


Figure 2: The inferred temperature $T_{\rm P}$ for MOR segments in the major ocean basins: the Pacific, the Atlantic, and the Indian. a) Map view of $T_{\rm P}$ averaged over all depths. White lines are ocean basin boundaries. b) $T_{\rm P}$ at depth. The order of ridge segments is shown with arrows in both panels. The ridge segments in the green box in both panels are possibly related to an ancient slab (Simmons et al., 2015). The bars on the bottom show the corresponding principal components for each segment.

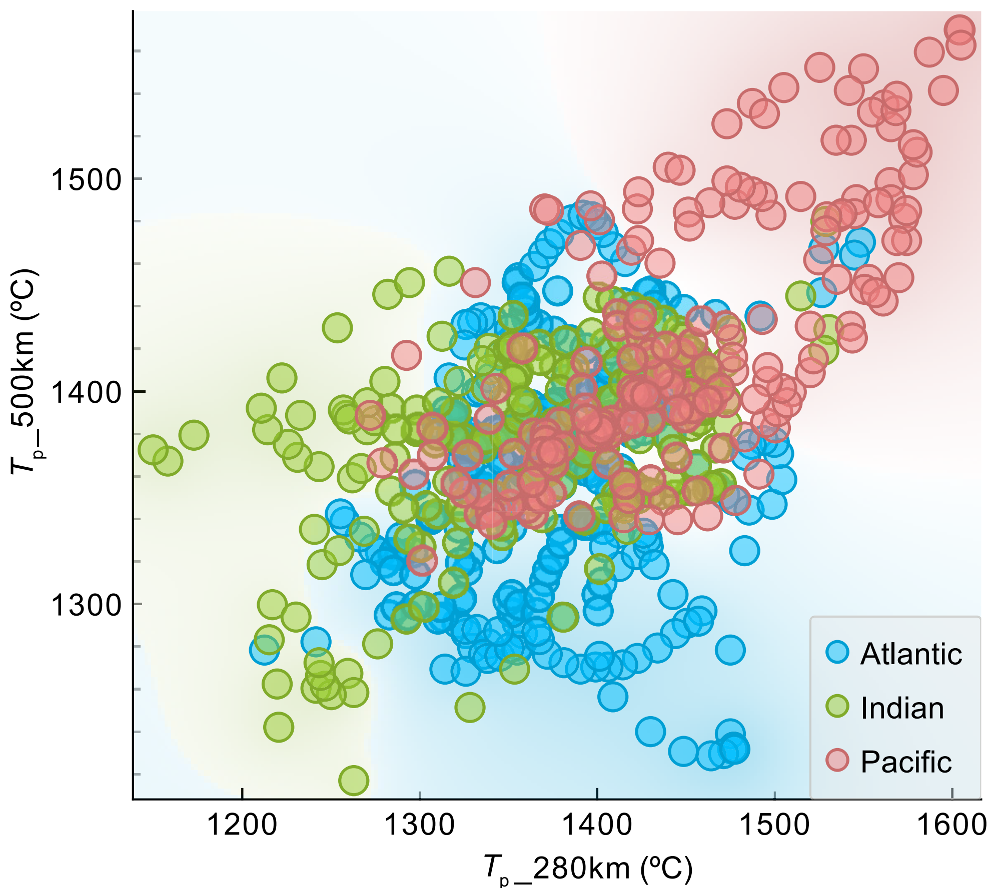
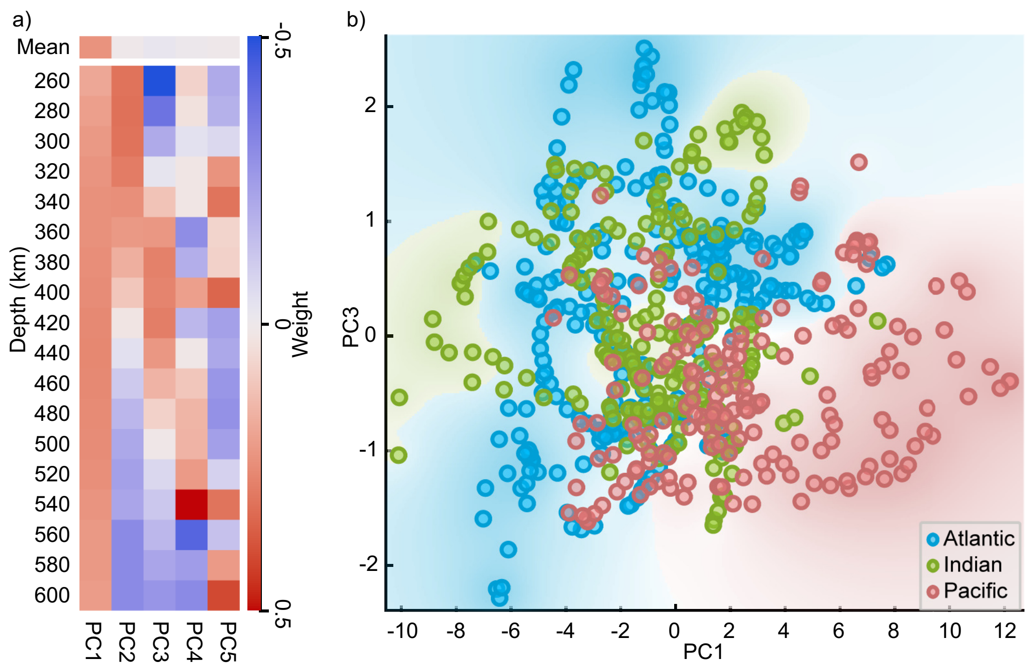


Figure 3: Scatter plot of potential temperature T_P at 280 km versus 500 km. This is the most informative projection among all pairs of temperatures, showing the best basin zoning, shown by the background colors.. Background colors are based on the density of points from each ocean basin in that space. Note that data from different basins are not easily separable with this linear classifier.



$PC^i = \sum_d W_d^i \hat{T}_{\mathrm{p}_d}$ W_d^i: weight for i th PC at depth d

$\hat{T_{\mathrm{p}_d}} = (T_{\mathrm{p}_d} - \mu_d)/(\sigma_d)$ at depth d T_{p_d}: potential temperature μ_d: average σ_d: standard deviation

Figure 4: Principal Component Analysis (PCA). a) Each PC is a weighted sum of normalized T_{P} at depth using the equations shown at the bottom. Individual weights (W_d^i) are shown as a heatmap. The top row shows the average weight of each column (over all depths). b) Ridge segment data is shown in the most informative space PC1 versus PC3 among all PC pairs. Background colors as in Fig. 3

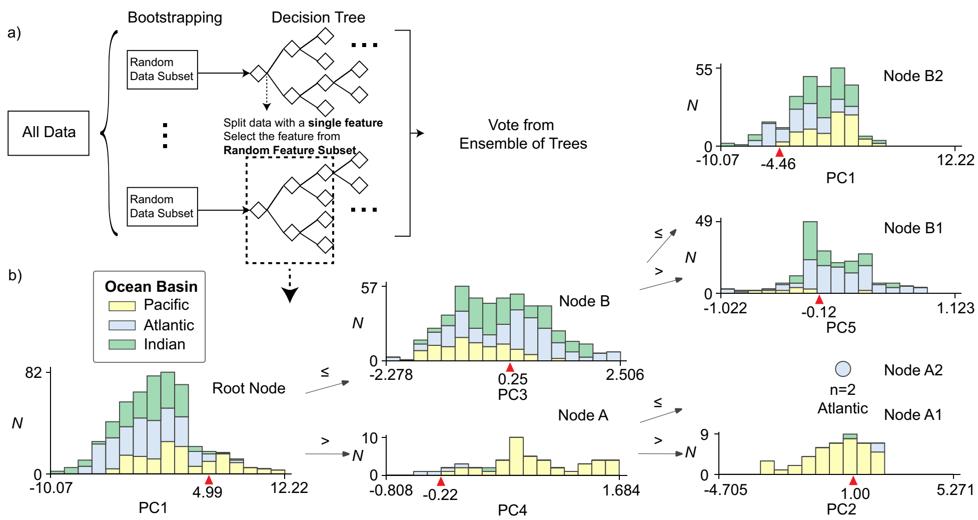


Figure 5: The Random Forest (RF) model. a) Schematic of RF. Data are randomly sampled as subsets with replacement (Boostrapping), and each subset is fed to a different decision tree. In each tree the data are bifurcated multiple times. For every bifurcation, the tree chooses a best PC from a random subset of PCs. Compared with the parent node, the child nodes are purified, i.e., they are gradually dominated by an ocean basin after bifurcation. The end node (leaf node) can predict probability of the ocean basin based on its basin fraction. The ensemble of trees then vote for the classification. b) The top 3 levels of one decision tree in the RF (dashed box in panel a). Each node bifurcates based on the PC shown (x axis) at the point indicated by the red triange. The y axis is the number of data points. The upper child node has data no larger than the point indicated by the red triange in its parent node, and vice versa. The tree stops at leaf nodes like A2, when all the points belong to one basin only, or with no more than 4 data points.

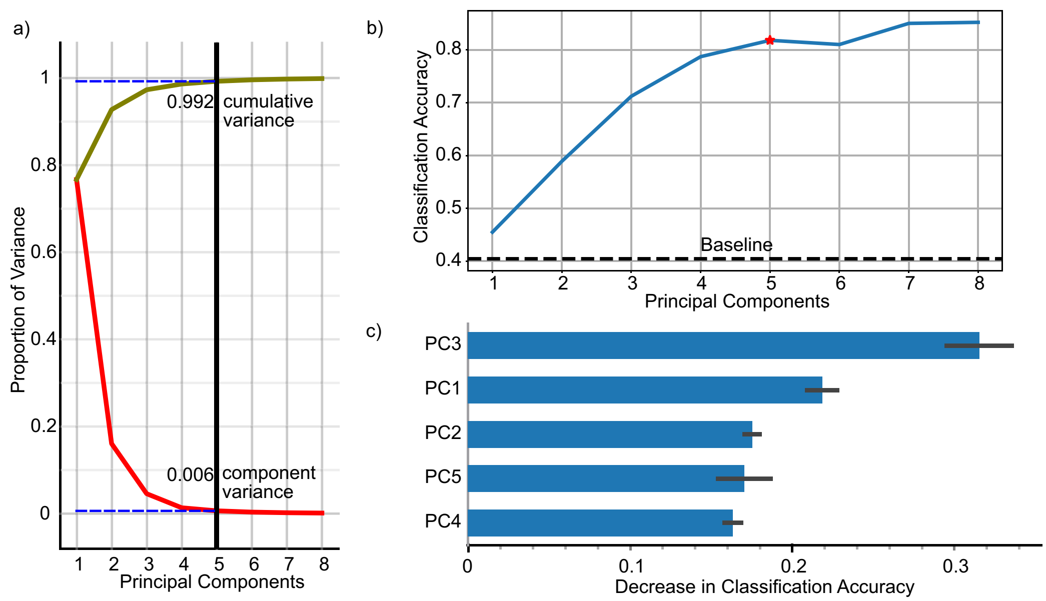


Figure 6: Effect of the first few PCs. a) The proportion of variance covered by each PC (red) and cumulative proportion (dark olive green). b)The cumulative Classification Accuracy with PC1 to PC8. The star denotes our final choice: PC1 to PC5, when we reach 82% accuracy. The baseline is to predict all ridge segments to be in the Atlantic basin, which has the most data. c) Feature importance is calculated from the decrease in classification accuracy by permuting data in each PC. Black bars show the standard deviation among all trees.

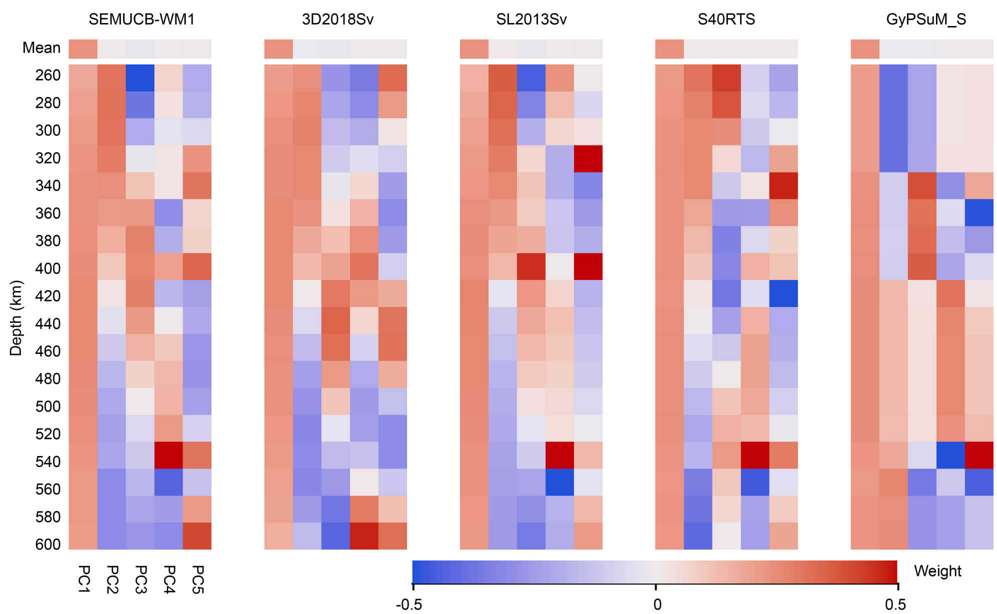


Figure 7: The PCA weight matrix of potential temperature at depth inferred from all tomographic models considered in this study.

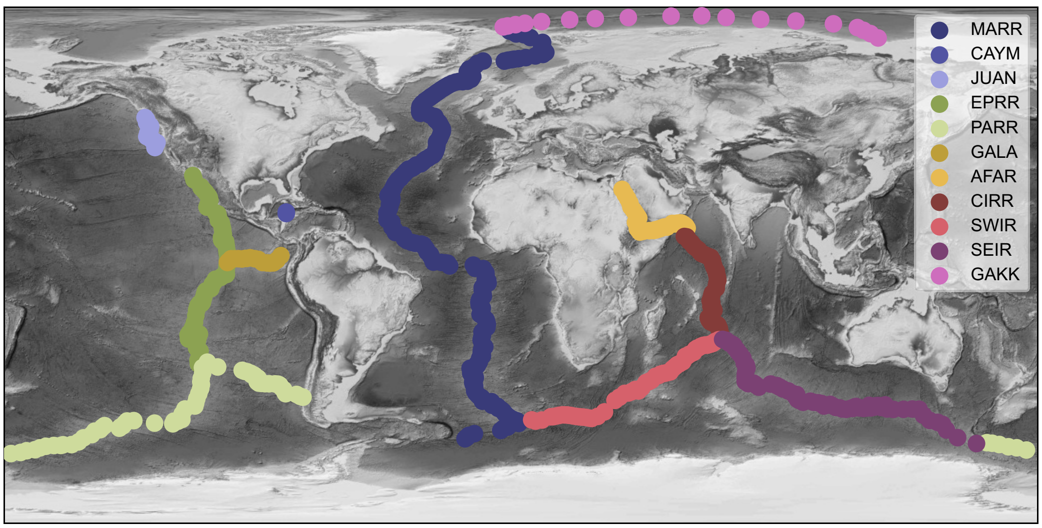


Figure 8: Sub-basin ridge systems as classified by our model. MARR: Mid-Atlantic Rise Ridge. CAYM: Cayman Ridge. JUAN: Juan De Fuca Ridge. EPRR: East Pacific Rise Ridge. PARR: Pacific-Antarctic Rise Ridge. GALA: Galapagos Ridge. AFAR: Red Sea Rift. CIRR: Central Indian Rise Ridge. SWIR: Southwest Indian Ridge. SEIR: Southeast Indian Ridge. GAKK: Gakkel Ridge.

Table 1: The confusion matrix from our classification models.

		Predicted			
		Atlantic	Indian	Pacific	Σ_{samples} [b]
Actual	Atlantic	88.0%[a]	6.8%	5.1%	2650
	Indian	8.7%	79.6%	11.7%	2050
	Pacific	10.5%	14.3%	75.1%	1850
	Σ_{samples} [b]	2707	2078	1765	6550

[a]Each row with percentages shows the fraction of all segments actually from a basin predicted to be in a different basin. The diagonal parts are the correct predicted fractions.

[b]The last row and last column show the numbers of bootstrapped test samples summed over all 50 trained random forest models.

References

Ball, P.W., White, N.J., Maclennan, J., Stephenson, S.N., 2021. Global influence of mantle temperature and plate thickness on intraplate volcanism. Nature Communications 12, 1–13. doi:10.1038/s41467-021-22323-9.

Bao, X., Lithgow-Bertelloni, C.R., Jackson, M.G., Romanowicz, B., 2022. On the relative temperatures of Earth's volcanic hotspots and mid-ocean ridges. Science 375, 57–61. doi:10.1126/science.abj8944.

Barruol, G., Sigloch, K., Scholz, J.R., Mazzullo, A., Stutzmann, E., Montagner, J.P., Kiselev, S., Fontaine, F.R., Michon, L., Deplus, C., Dyment, J., 2019. Large-scale flow of Indian Ocean asthenosphere driven by Réunion plume. Nature Geoscience 12, 1043–1049. doi:10.1038/s41561-019-0479-3.

Birch, F., 1969. Density and composition of the upper mantle: first approximation as an olivine layer. Geophysical Monograph Series 13, 18–36. doi:10.1029/GM013p0018.

Brandl, P.A., Regelous, M., Beier, C., Haase, K.M., 2013. High mantle temperatures following rifting caused by continental insulation. Nature Geoscience 6, 391–394. doi:10.1038/ngeo1758.

Breiman, L., 2001. Random forests. Machine learning 45, 5–32.

Bridle, J., 1989. Training stochastic model recognition algorithms as networks can lead to maximum mutual information estimation of parameters. Advances in neural information processing systems 2.

Brown Krein, S., Molitor, Z.J., Grove, T.L., 2021. ReversePetrogen : A Multiphase Dry Reverse Fractional Crystallization-Mantle Melting Thermobarometer Applied to 13,589 Mid-Ocean Ridge Basalt Glasses. Journal of Geophysical Research: Solid Earth 126, 1–20. doi:10.1029/2020JB021292.

Bryan, S.E., Ferrari, L., 2013. Large igneous provinces and silicic large igneous provinces: Progress in our understanding over the last 25 years. GSA Bulletin 125, 1053–1078. doi:10.1130/B30820.1.

Cammarano, F., Goes, S., Vacher, P., Giardini, D., 2003. Inferring upper-mantle temperatures from seismic velocities. Physics of the Earth and Planetary Interiors 138, 197–222. doi:10.1016/S0031-9201(03)00156-0.

Conrad, C.P., Steinberger, B., Torsvik, T.H., 2013. Stability of active mantle upwelling revealed by net characteristics of plate tectonics. Nature 498, 479–482. doi:10.1038/nature12203.

Dalton, C.A., Langmuir, C.H., Gale, A., 2014. Geophysical and Geochemical Evidence for Deep Temperature Variations Beneath Mid-Ocean Ridges. Science 344, 80–83. doi:10.1126/science.1249466.

Davaille, A., Stutzmann, E., Silveira, G., Besse, J., Courtillot, V., 2005. Convective patterns under the indo-atlantic box. Earth and Planetary Science Letters 239, 233–252. doi:10.1016/j.epsl.2005.07.024.

Debayle, E., Bodin, T., Durand, S., Ricard, Y., Durand, S., 2020. Seismic evidence for partial melt below tectonic plates. Nature 586, 555–559. doi:10.1038/s41586-020-2809-4.

Debayle, E., Dubuffet, F., Durand, S., 2016. An automatically updated S-wave model of the upper mantle and the depth extent of azimuthal anisotropy. Geophysical Research Letters 43, 674–682. doi:10.1002/2015GL067329.

Demšar, J., Curk, T., Erjavec, A., Črt Gorup, Hočevar, T., Milutinovič, M., Možina, M., Polajnar, M., Toplak, M., Starič, A., Štajdohar, M., Umek, L., Žagar, L., Žbontar, J., Žitnik, M., Zupan, B., 2013. Orange: Data mining toolbox in python. Journal of Machine Learning Research 14, 2349–2353.

Doucet, L.S., Li, Z.X., Gamal El Dien, H., Pourteau, A., Murphy, J.B., Collins, W.J., Mattielli, N., Olierook, H.K., Spencer, C.J., Mitchell, R.N., 2020. Distinct formation history for deep-mantle domains reflected in geochemical differences. Nature Geoscience 13, 511–515. doi:10.1038/s41561-020-0599-9.

Duda, R.O., Hart, P.E., et al., 1973. Pattern classification and scene analysis. volume 3. Wiley New York.

French, S.W., Romanowicz, B.A., 2014. Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography. Geophysical Journal International 199, 1303–1327. doi:10.1093/gji/ggu334.

Gale, A., Langmuir, C.H., Dalton, C.A., 2014. The Global Systematics of Ocean Ridge Basalts and their Origin. Journal of Petrology 55, 1051–1082. doi:10.1093/petrology/egu017.

Gassmöller, R., Dannberg, J., Bredow, E., Steinberger, B., Torsvik, T.H., 2016. Major influence of plume-ridge interaction, lithosphere thickness variations, and global mantle flow on hotspot volcanism—the example of tristan. Geochemistry, Geophysics, Geosystems 17, 1454–1479. doi:10.1002/2015GC006177.

Gibson, S.A., Richards, M.A., 2018. Delivery of deep-sourced, volatile-rich plume material to the global ridge system. Earth and Planetary Science Letters 499, 205–218. doi:10.1016/j.epsl.2018.07.028.

Gurnis, M., 1988. Large-scale mantle convection and the aggregation and dispersal of supercontinents. Nature 332, 695–699. URL: https://www.nature.com/articles/332695a0, doi:10.1038/332695a0.

Gurnis, M., Müller, R.D., Moresi, L., 1998. Cretaceous Vertical Motion of

Australia and the AustralianAntarctic Discordance. Science 279, 1499–1504. doi:10.1126/science.279.5356.1499.

Ito, G., Lin, J., Graham, D., 2003. Observational and theoretical studies of the dynamics of mantle plume–mid-ocean ridge interaction. Reviews of Geophysics 41. doi:10.1029/2002RG000117.

Jellinek, A.M., Lenardic, A., 2009. Effects of spatially varying roof cooling on thermal convection at high Rayleigh number in a fluid with a strongly temperature-dependent viscosity. Journal of Fluid Mechanics 629, 109–137. doi:10.1017/S0022112009006260.

Jolliffe, I.T., 2002. Principal component analysis for special types of data. Springer.

Karlsen, K.S., Conrad, C.P., Domeier, M., Trønnes, R.G., 2021. Spatiotemporal Variations in Surface Heat Loss Imply a Heterogeneous Mantle Cooling History. Geophysical Research Letters 48, 1–10. doi:10.1029/2020GL092119.

Keller, T., Katz, R.F., Hirschmann, M.M., 2017. Volatiles beneath mid-ocean ridges: Deep melting, channelised transport, focusing, and metasomatism. Earth and Planetary Science Letters 464, 55–68. doi:10.1016/j.epsl.2017.02.006, arXiv:1608.03841.

Kimura, J.I., Sano, S., 2012. Reactive melt flow as the origin of residual mantle lithologies and basalt chemistries in mid-ocean ridges: Implications from the red hills peridotite, New Zealand. Journal of Petrology 53, 1637–1671. doi:10.1093/petrology/egs028.

Klein, E.M., Langmuir, C.H., 1987. Global correlations of ocean ridge basalt chemistry with axial depth and crustal thickness. Journal of Geophysical Research 92, 8089–8115. doi:10.1029/JB092iB08p08089.

Koppers, A.A., Becker, T.W., Jackson, M.G., Konrad, K., Müller, R.D., Romanowicz, B., Steinberger, B., Whittaker, J.M., 2021. Mantle plumes and their role in earth processes. Nature Reviews Earth & Environment 2, 382–401. doi:10.1038/s43017-021-00168-6.

van de Lagemaat, S.H., van Hinsbergen, D.J., 2023. Plate tectonic cross-roads: Reconstructing the panthalassa-neotethys junction region from philippine sea plate and australasian oceans and orogens. Gondwana Research doi:10.1016/j.gr.2023.09.013.

Lenardic, A., Moresi, L., Jellinek, A.M., O'Neill, C.J., Cooper, C.M., Lee, C.T., 2011. Continents, supercontinents, mantle thermal mixing, and mantle thermal isolation: Theory, numerical simulations, and laboratory experiments. Geochemistry, Geophysics, Geosystems 12, 1–23. doi:10.1029/2011GC003663.

Lithgow-Bertelloni, C., Richards, M.A., 1998. The dynamics of cenozoic and mesozoic plate motions. Reviews of Geophysics 36, 27–78. doi:10.1029/97RG02282.

MacLeod, S.J., Müller, R.D., Hassan, R., Williams, S.E., 2023. Evolution of mantle plumes and lower mantle structure in numerical models using tectonic reconstructions as boundary conditions, in: Dynamics of Plate

Tectonics and Mantle Convection. Elsevier, pp. 427–458. doi:10.1016/B978-0-323-85733-8.00022-6.

Mao, W., Zhong, S., 2018. Slab stagnation due to a reduced viscosity layer beneath the mantle transition zone. Nature Geoscience 11, 876–881. doi:10.1038/s41561-018-0225-2.

Masters, G., Jordan, T.H., Silver, P.G., Gilbert, F., 1982. Aspherical earth structure from fundamental spheroidal-mode data. Nature 298, 609–613. doi:10.1038/298609a0.

Meschede, M., Romanowicz, B., 2015. Lateral heterogeneity scales in regional and global upper mantle shear velocity models. Geophysical Journal International 200, 1078–1095. doi:10.1093/gji/ggu424.

Müller, R.D., Qin, X., Sandwell, D.T., Dutkiewicz, A., Williams, S.E., Flament, N., Maus, S., Seton, M., 2016. The gplates portal: Cloud-based interactive 3d visualization of global geophysical and geological data in a web browser. PloS one 11, e0150883. doi:10.1371/journal.pone.0150883.

Müller, R.D., Zahirovic, S., Williams, S.E., Cannon, J., Seton, M., Bower, D.J., Tetley, M.G., Heine, C., Le Breton, E., Liu, S., et al., 2019. A global plate model including lithospheric deformation along major rifts and orogens since the triassic. Tectonics 38, 1884–1907. doi:10.1029/2018TC005462.

Niu, Y., 2016. The Meaning of Global Ocean Ridge Basalt Major Element Compositions. Journal of Petrology 57, 2081–2103. doi:10.1093/petrology/egw073.

Niu, Y., O'Hara, M.J., 2008. Global correlations of ocean ridge basalt chemistry with axial depth: A new perspective. Journal of Petrology 49, 633–664. doi:10.1093/petrology/egm051.

O'Neill, C., Lenardic, A., Jellinek, A.M., Moresi, L., 2009. Influence of supercontinents on deep mantle flow. Gondwana Research 15, 276–287. doi:10.1016/j.gr.2008.11.005.

Ray, T.W., Anderson, D.L., 1994. Spherical disharmonies in the Earth sciences and the spatial solution: Ridges, hotspots, slabs, geochemistry and tomography correlations. Journal of Geophysical Research: Solid Earth 99, 9605–9614. doi:10.1029/94JB00340.

Ribe, N., Christensen, U., Theissing, J., 1995. The dynamics of plume-ridge interaction, 1: Ridge-centered plumes. Earth and Planetary Science Letters 134, 155–168. doi:10.1016/0012-821X(95)00116-T.

Ricard, Y., Richards, M., Lithgow-Bertelloni, C., Le Stunff, Y., 1993. A geodynamic model of mantle density heterogeneity. Journal of Geophysical Research: Solid Earth 98, 21895–21909. doi:10.1029/93JB02216.

Richards, M.A., Engebretson, D.C., 1992. Large-scale mantle convection and the history of subduction. Nature 355, 437–440. doi:10.1038/355437a0.

Ritsema, J., Deuss, A., Van Heijst, H.J., Woodhouse, J.H., 2011. S40RTS: A degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements. Geophysical Journal International 184, 1223–1236. doi:10.1111/j.1365-246X.2010.04884.x.

Rodriguez, J., Kuncheva, L., Alonso, C., 2006. Rotation forest: A new classifier ensemble method. IEEE Transactions on Pattern Analysis and Machine Intelligence 28, 1619–1630. doi:10.1109/TPAMI.2006.211.

Rowley, D.B., Forte, A.M., Rowan, C.J., Glišović, P., Moucha, R., Grand, S.P., Simmons, N.A., 2016. Kinematics and dynamics of the East Pacific Rise linked to a stable, deep-mantle upwelling. Science Advances 2, e1601107. doi:10.1126/sciadv.1601107.

Schaeffer, A.J., Lebedev, S., 2013. Global shear speed structure of the upper mantle and transition zone. Geophysical Journal International 194, 417–449. doi:10.1093/gji/ggt095.

Schubert, G., Froidevaux, C., Yuen, D., 1976. Oceanic lithosphere and asthenosphere: thermal and mechanical structure. Journal of Geophysical Research 81, 3525–3540. doi:10.1029/JB081i020p03525.

Simmons, N., Myers, S., Johannesson, G., Matzel, E., Grand, S., 2015. Evidence for long-lived subduction of an ancient tectonic plate beneath the southern indian ocean. Geophysical Research Letters 42, 9270–9278. doi:10.1002/2015GL066237.

Simmons, N.A., Forte, A.M., Boschi, L., Grand, S.P., 2010. GyPSuM: A joint tomographic model of mantle density and seismic wave speeds. Journal of Geophysical Research: Solid Earth 115. doi:10.1029/2010JB007631.

Sinton, J.M., Detrick, R.S., 1992. Mid-ocean ridge magma chambers. Journal of Geophysical Research: Solid Earth 97, 197–216. doi:10.1029/91JB02508.

Sleep, N.H., 2002. Ridge-crossing mantle plumes and gaps in tracks. Geochemistry, Geophysics, Geosystems 3, 1–33. doi:10.1029/2001GC000290.

Spiegelman, M., Kelemen, P.B., 2003. Extreme chemical variability as a consequence of channelized melt transport. Geochemistry, Geophysics, Geosystems 4. doi:10.1029/2002GC000336.

Stewart, A.J., Li, Y., Cao, Z., Liu, L., 2023. Hemispheric dichotomy of mantle dynamics revealed by machine learning. arXiv preprint arXiv:2306.14312 .

Stixrude, L., Lithgow-Bertelloni, C., 2005. Thermodynamics of mantle minerals - I. Physical properties. Geophysical Journal International 162, 610–632. doi:10.1111/j.1365-246X.2005.02642.x.

Stixrude, L., Lithgow-Bertelloni, C., 2011. Thermodynamics of mantle minerals - II. Phase equilibria. Geophysical Journal International 184, 1180–1213. doi:10.1111/j.1365-246X.2010.04890.x.

Stixrude, L., Lithgow-Bertelloni, C., 2022. Thermal expansivity, heat capacity and bulk modulus of the mantle. Geophysical Journal International 228, 1119–1149. doi:10.1093/gji/ggab394.

Stixrude, L., Lithgow-Bertelloni, C., 2024. Thermodynamics of mantle minerals-iii. the role of iron. Geophysical Journal International , ggae126doi:10.1093/gji/ggae126.

Stracke, A., 2021. A process-oriented approach to mantle geochemistry. Chemical Geology 579, 120350.

Stracke, A., Willig, M., Genske, F., Béguelin, P., Todd, E., 2022. Chemical geodynamics insights from a machine learning approach. Geochemistry, Geophysics, Geosystems 23, e2022GC010606.

Weis, D., Harpp, K.S., Harrison, L.N., Boyet, M., Chauvel, C., Farnetani, C.G., Finlayson, V.A., Lee, K.K., Parai, R., Shahar, A., et al., 2023. Earth's mantle composition revealed by mantle plumes. Nature Reviews Earth & Environment 4, 604–625. doi:10.1038/s43017-023-00467-0.

Workman, R.K., Hart, S.R., 2005. Major and trace element composition of the depleted MORB mantle (DMM). Earth and Planetary Science Letters 231, 53–72. doi:10.1016/j.epsl.2004.12.005.