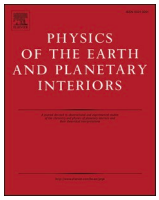
Physics of the Earth and Planetary Interiors 361 (2025) 107317

Contents lists available at ScienceDirect

Physics of the Earth and Planetary Interiors

journal homepage: www.elsevier.com/locate/pepi



SeisTeC: A neural network tool to constrain mantle thermal and chemical properties from seismic observables

Ashim Rijal *, Laura Cobden, Jeannot Trampert

Department of Earth Sciences, Utrecht University, Princetonlaan 8a, 3584 CB Utrecht, the Netherlands

ARTICLE INFO

Keywords:
Lower mantle
Bulk composition
Seismology
Wave speeds
Neural networks
Probability density

ABSTRACT

Three-dimensional variations of wave speeds and density have identified the presence of seismically distinct structures in the Earth's mantle. To determine the thermochemical properties and dynamic relevance of these structures, it is crucial to understand the relationship between seismic properties and temperature and composition. However, multiple thermochemical parameters influence seismic wave speeds simultaneously. A given wave speed pair (compressional and shear) and density can be generated by many possible combinations of thermochemical parameters, which makes the inversion of wave speeds and density for thermochemical parameters a non-unique problem. We have developed a tool which efficiently captures the mapping between seismic wave speeds (and density) and thermochemical properties, with the capacity to represent both the inherent trade-offs between parameters as well as data uncertainties. These trade-offs and uncertainties are represented by the posterior probability density function provided by a neural network. We demonstrate the concept for seismic wave speeds and density, but the same tool can also be adapted for other parameters such as attenuation or properties of seismic discontinuities. SeisTeC is available to the wider community and is intended to facilitate interpretations of seismic structures inside the Earth, or in general, any planetary bodies.

Our tool is based on a neural network, which implicitly learns the non-linear mapping between temperature and bulk composition. We chose the example of the lower mantle and expressed composition in terms of six end-member oxides ($SiO_2, MgO, Al_2O_3, FeO, Na_2O, CaO$) and modelled seismic wave speeds and density at appropriate temperature and pressure conditions. Wave speeds and density are calculated for 750,000 thermochemical models, whose temperature and composition are selected at random from pre-defined ranges, using thermodynamic modelling. We train neural networks with wave speeds plus or minus density as the input, and temperature and bulk composition as target outputs. The networks then approximate a probability density function for each output, which allows us to interpret seismic observables in terms of physical parameters, crucially, with uncertainties. When working with wave speeds (V_P and V_S) only, we find trade-offs between pairs of parameters such as temperature - FeO, SiO_2 - MgO, SiO_2 - Na_2O, and SiO_2 - Al_2O_3 which limits the constraints one can place on mantle temperature and chemistry using these observables. We also emphasise the importance of combining V_P and V_S for constraining SiO_2 content. The main advantage of including density with wave speeds is that it helps to better constrain the temperature and the most abundant and dynamically relevant compositional end-members, namely, the SiO_2, MgO and FeO by breaking down the trade-offs between them. Some trade-offs between pairs of parameters involving minor compositional end-members still remain, namely temperature - CaO, SiO_2 - Na_2O, SiO_2 - Al_2O_3. In general, except mid-ocean ridge basalt, most rocks only have a small fraction of Na_2O and Al_2O_3. By excluding mid-ocean ridge basalt in the training data most of the apparent trade-offs will disappear when considering more average mantle compositional ranges.

1. Introduction

Seismic data provide us with vital information about the physical properties inside the Earth or a planetary interior (e.g. Jeffreys and Bullen, 1940; Dziewonski and Anderson, 1981; Kennett et al., 1995; Garcia et al., 2011; Khan et al., 2021; Stähler et al., 2021). As a result, seismic models of the Earth have identified various seismic features at different length scales in the mantle (e.g. Lay and Helmberger, 1983; van

* Corresponding author.
E-mail address: a.rijal@uu.nl (A. Rijal).

https://doi.org/10.1016/j.pepi.2025.107317
Received 3 July 2023; Received in revised form 23 May 2024; Accepted 31 January 2025
Available online 13 February 2025

der Hilst et al., 1997; Garnero and Helmberger, 1998; Romanowicz, 2008; Fichtner et al., 2009; Thomas et al., 2009; Tauzin et al., 2010; Ritsema et al., 2011; Debayle and Ricard, 2012; Sun et al., 2013; Frost and Rost, 2014; French and Romanowicz, 2014; Waszek et al., 2015; Garnero et al., 2016; Lei et al., 2020). Seismology mainly provides us with seismic wave speeds, and although difficult, sometimes with density (e.g. Ishii and Tromp, 1999; Trampert et al., 2004; Mosca et al., 2012; Moulik and Ekström, 2016; Koelemeijer et al., 2017) or attenuation structure (Dziewonski and Anderson, 1981; Durek and Ekström, 1996; Lawrence and Wysession, 2006; Dalton et al., 2008; Hwang and Ritsema, 2011; Zhu et al., 2013; Karaoğlu and Romanowicz, 2018; Konishi et al., 2020; Talavera-Soza et al., 2025). However, to decipher the structure and dynamics of the Earth's mantle, one needs parameters such as temperature and composition as they are closely related to mantle dynamics. Hence, it is crucial to understand the temperature and compositional dependence of seismic wave speeds (and density).

It is not feasible to directly sample the Earth's lower mantle. Therefore, mineral elastic properties (bulk modulus K, density ρ and shear modulus G) for an assumed composition and at relevant mantle pressure (P) and temperature (T) are inferred from experimental measurements and theoretical calculations and/or by extrapolation thereof (e.g. Duffy and Ahrens, 1995; Sinogeikin and Bass, 2000; Speziale et al., 2001; Murakami et al., 2004; Stixrude and Lithgow-Bertelloni, 2005; Marquardt et al., 2009; Dorfman et al., 2013; Wolf et al., 2015; Thomson et al., 2019; Oganov and Dorogokupets, 2003; Wentzcovitch et al., 2004; Tsuchiya et al., 2004; Mao et al., 2005; Wang et al., 2015; Muir and Brodholt, 2015). In this forward procedure (Fig. 1), the elastic properties of individual minerals are used to compute seismic properties of the bulk assemblage. The seismic properties of the assemblage are then compared with those obtained from observed seismic data (e.g. Jackson, 1998; Deschamps and Trampert, 2004; Cammarano et al., 2005b; Matas et al., 2007; Xu et al., 2008; Cobden et al., 2009; Khan et al., 2009; Simmons et al., 2010; Cobden et al., 2012; Deschamps et al., 2019; Vilella et al., 2021). The inversion of seismic wave speeds for a thermochemical structure is a non-unique problem, i.e. for a given compressional and shear wave speed there are many possible combinations of temperature and composition that will fit.

Using only two or three (wave speeds $\pm$ density) observables, determining temperature and bulk composition (here, six oxide end-members) is an underdetermined problem. One way to simplify the problem is to assume the mantle has a fixed composition, in most cases pyrolite (Ringwood, 1962), and interpret wave speed variations in terms of temperature only (e.g. Goes et al., 2004; Cammarano et al., 2005a; Cammarano et al., 2005b; Cammarano and Romanowicz, 2007; Ritsema et al., 2009; Simmons et al., 2009; Schuberth et al., 2012). Some studies suggest a part or the whole lower mantle is composed of pyrolite (e.g. Wentzcovitch et al., 2004; Li and Zhang, 2005; Irifune et al., 2010; Chantel et al., 2012; Zhang et al., 2013; Tsuchiya and Kawai, 2013; Cottaar et al., 2014; Wang et al., 2015; Wu, 2016; Hyung et al., 2016), while other studies have indicated an average composition more enriched in silica (i.e. bridgmanite), and/or having characteristics of chondritic materials (e.g. Stixrude et al., 1992; Murakami et al., 2012; Ricolleau et al., 2009). While fixing a composition a priori eliminates possible trade-offs between temperature and composition, it potentially provides an unrealistic model of mantle thermal structure, if the wrong chemical model is chosen. (e.g. Cobden et al., 2024). Additionally, some studies suggest that the mantle has a depth dependent composition or lateral deviations from a uniform composition (e.g. Trampert et al., 2004; Matas et al., 2007; Khan et al., 2008; Cobden et al., 2009; Deschamps et al., 2012; Ballmer et al., 2017; Cobden et al., 2024). Another option is to increase the number of observables. For instance, Afonso et al., 2022 inverted multiple datasets for the thermochemical structure of cratons. This is more easily feasible for the uppermost lithosphere than the deep mantle, where additional datasets are difficult to generate and fraught with high uncertainites.

A number of studies have attempted to invert for mantle temperature and composition. Mattern et al., 2005 and Matas et al., 2007 applied a least square criterion, whereas Simmons et al., 2010 adapted a very fast simulated annealing approach. As the inverse problem of inferring thermochemical variations is a strongly non-linear problem (Khan et al., 2008), some random sampling and Monte Carlo type methods also have been invoked (e.g. Trampert et al., 2004; Deschamps and Trampert, 2004; Cammarano et al., 2005b; Khan et al., 2011; Mosca et al., 2012; Cobden et al., 2012; Deschamps et al., 2012; Afonso et al., 2022; Cobden et al., 2018). In order to make thermochemical interpretations (i.e. to get the posterior probability density of temperature or composition), these methods require re-sampling of the posterior every time we wish to interpret a different dataset or oberservation. This is labour intensive, and the advantage of using neural networks, which is the approach taken in this study, is that once trained on priors (Käufl et al., 2016) they are extremely efficient at predicting posterior probability density functions without retraining unless the prior changes. In addition, the neural networks are very flexible, i.e. readily adapted for different inputs and outputs, as well as different uncertainties.

In this paper, we investigate the feasibility of mapping the compositional variations at the same time as the temperature variation in the Earth's lower mantle with neural networks at fixed depth/pressure. Owing to the non-uniqueness of the inverse problem, we consider a wide variety of prior compositions ranging from harzburgite to mid-ocean ridge basalt (MORB). We calculate the seismic properties (wave speed and density) corresponding to this range of compositions via thermodynamic modelling. We then create a neural network which is able to map these seismic properties into probability density functions of bulk composition and temperature. This allows us to quantify the ranges of

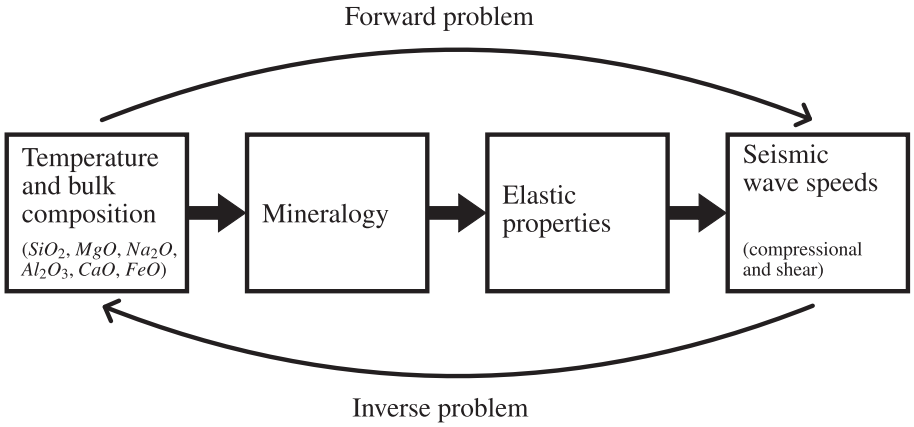


Fig. 1. The forward problem is the computation of the mineral seismic properties (e.g. compressional and shear wave speeds) from temperature and composition (plus pressure), whereas the inverse problem is going from seismic properties to thermochemical parameters.

temperature and composition that fit a given seismic wave speed (and density) of the lower mantle. Moreover, the approach accounts for all trade-offs among pairs of compositional parameters, as well as between temperature and composition. This in turn enables us to understand the relative contributions of thermal versus compositional effects on wave speeds and to what extent we can constrain these using seismology. We empasise that we train the neural networks with theoretical prior data without ever using actual seismic data. The advantage of prior sampling and convergence to classical Monte Carlo type methods is discussed in Käufl et al., 2016.

2. Methodology

2.1. Thermodynamic modelling

We define the bulk composition of the Earth's mantle in terms of six end-member oxides, namely: SiO_2, MgO, FeO, CaO, Al_2O_3 and Na_2O, which form over 98 % of the mantle mass (e.g. Irifune, 1994). For a given bulk composition, pressure and temperature, the mineral assemblages and elastic properties are calculated self-consistently using Perple_X thermodynamic modelling software (Connolly, 1990, 2005) together with the equation of state in Stixrude and Lithgow-Bertelloni, 2005 and mineral parameters in Stixrude and Lithgow-Bertelloni, 2011. Average elastic properties, namely, the Voigt average for ρ and Voigt-Reuss-Hill for K and G, are then taken to compute the average compressional (V_P) and shear wave (V_S) speeds of the mineral assemblage as

$$V_P(P,T,C) = \sqrt{\frac{K(P,T,C) + \frac{4}{3}G(P,T,C)}{\rho(P,T,C)}} \quad \text{and} \quad V_S(P,T,C) = \sqrt{\frac{G(P,T,C)}{\rho(P,T,C)}}, \tag{1}$$

respectively, where we have denoted the composition by C. To allow for variations in the mantle composition, we compute mineral assemblages for 2500 randomly chosen bulk compositions and for each composition, at a given depth (pressure), 300 temperatures are selected at random, giving a maximum of 750,000 different thermochemical models. These models cover all compositions from harzburgite to MORB, as shown by the range of each individual oxide in Table 1. Some of the low temperature models are thermodynamically unstable and therefore discarded, leaving a total of 749,471 models at 1000 km. While the temperature is drawn uniformly from the range shown in the Table 1, prior distributions for oxides are not uniform (see Appendix A) because the weight percentages, by definition, have to sum to 100 % (see Cobden et al., 2018 for details). In this study, we show results at 1000 and 2800 km depths corresponding to pressures of 38.61 GPa and 130.40 GPa, i.e. the top and bottom of the lower mantle, respectively. We use PREM (Dziewonski and Anderson, 1981) to convert depth to pressure, and we assume that the pressure is known and defines depth. We pick those two depths to avoid complexity from the iron spin transition (e.g. Sturhahn et al., 2005; Fei et al., 2007; Wentzcovitch et al., 2009; Marquardt et al., 2018) in the mid mantle, which is beyond the scope of this paper, but can readily be incorporated. In the following, we present results mainly for the anharmonic wave speeds. Introducing intrinsic anelasticity adds an extra set of parameters or assumptions to the modelling while - as we shall show - having a minor impact on the wave speeds themselves in the lower mantle.

2.2. Mixture density network (MDN)

Once the seismic properties of the thermochemical models are calculated, we try to answer the following: can we infer the temperature and composition for a given value of V_P and V_S (±density)? Because of the non-uniqueness of this inverse problem, we treat the problem within the probabilistic framework. The solution to our probabilistic inverse problem is the posterior probability density function (pdf) for temperature and composition, which is denoted as

$$\sigma(x|V_P, V_S), \tag{2}$$

where x can be one of the following: temperature, a component of bulk composition. An arbitrary posterior probability density function can be approximated by a combination of a feed-forward neural network and Gaussian kernels. The resulting combination is called a Mixture Density Network (MDN) (Bishop, 1994, 1995). For a description of this type of neural network and references to some examples of its application in Earth Sciences, readers are referred to our previous work (Rijal et al., 2021). Based on the samples generated using the forward procedure in Sub-section 2.1, we train an ensemble of independent MDNs to approximate the true posterior (of Eq. 2), which can be written as

$$p(x|V_P, V_S; \widehat{\boldsymbol{\alpha}}) = \sum_{n=1}^{M} w_n(x|V_P, V_S; \boldsymbol{\alpha}_n) q_n(x|V_P, V_S; \boldsymbol{\alpha}_n) \approx \sigma(x|V_P, V_S), \tag{3}$$

where $\widehat{\boldsymbol{\alpha}} = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ represents weights and biases, also called parameters, of the feed-forward part of ensemble members. In Eq. 3, we have performed a weighted sum of a total of M ensemble members, where q_n and w_n represent the output posterior pdf and the mixing coefficient (or weight), respectively, of the n_{th} ensemble member. $\boldsymbol{\alpha}_n$ and w_n are tuned by the ensemble during its training process providing a reasonable approximation of the underlying generator of the training data (see Section 2.3). For the details of this prior sampling based approach, and its comparison with MCMC, readers are referred to Käufl et al., 2016, for example. Often a range of network parameters fit the data equally well, depending on the initialisation and network architecture. To account for this, we train an ensemble of fifteen, i.e. M = 15, independent MDNs (e.g. de Wit and Trampert, 2015; Käufl et al., 2016; Rijal et al., 2021), with each having 3–5 Gaussian kernels. The output is a parameterised posterior pdf $p(x|V_P, V_S; \boldsymbol{\alpha})$ on temperature or a component of bulk composition for a pair of V_P and V_S. The architecture of an MDN is shown in Appendix B, along with the range of hidden nodes used to approximate the thermochemical posteriors. The hidden node range is determined based on a network complexity test. The test provides the range which has the lowest error on a dataset that is not used for training the networks. The MDNs are built, trained and deployed for prediction using TensorFlow 1.13.1 (Abadi et al., 2015). Sometimes, it is possible to get the density information along with the wave speeds from seismology. One can simply condition Eq. 3 (as we will see) on density as well, if such an additional constraint is available.

2.3. MDN prediction performance

Each of the 750,000 randomly generated thermochemical models (in Section 2.1), has its own wave speeds, density, bulk composition and temperature. We remind the reader that we work at a fixed pressure or

Table 1
The prior range of T (in K) and six oxides (in wt%) considered in this study. T_S represents the maximum temperature which is assumed not to exceed the mantle solidus of $MgSiO_3$ perovskite (Stixrude et al., 2009). It is approximately 3347 K at 1000 km depth.

	T	SiO_2	MgO	FeO	CaO	Al_2O_3	Na_2O
Lower bound	900.00	31.127	4.720	1.951	0.017	0.195	0.002
Upper bound	T_S	79.855	54.840	19.169	15.083	21.007	4.465

depth here. A randomly chosen training set consisting of 70 % of these models is used to train the ensemble of MDNs, which learn the relationship between inputs (V_S and V_P) and an output (temperature or a component oxide). A subset of 10 % of the models, called a test set, is used to test the prediction performance of the trained MDN ensemble. Another subset with remaining 20 % of the models are used to restrict the overfitting during the training process. We use V_S and V_P of the test set to predict pdfs for temperature and composition. Since we know the actual values (also called targets) of the temperature and composition for the wave speeds of the test set, we can compare them with the network predictions.

We train a separate MDN ensemble for each target thermochemical parameter. In other words, the output of each trained MDN ensemble is a 1-D marginal (posterior) pdf. A simple measure of the prediction performance is a comparison between the mean of the posterior pdf and the actual target value from the test set. Fig. 2 (plots labelled as w/o density) shows the performance of trained MDN ensembles at 1000 km depth. Each circle represents the mean of the posterior pdf predicted by neural networks. Ideally, a unimodal pdf with pronounced peak would mean the predictions fall along the diagonal line. That is, the closer the circles plot to the dashed black line, the better the trained MDN ensemble has performed. The networks resolve variations in temperature and composition. Although the network predictions are scattered for some targets, they cluster around the dashed diagonal line, showing, for instance, the wave speeds are able to infer temperature and SiO_2 quite well. In comparison, the networks provide less constraints (circles plot more horizontally) on the relationship between wave speeds and *MgO*, *FeO*, *CaO*, Al_2O_3 and Na_2O, because the same wave speeds are generated by different combinations of these oxides. The predicted mean values of these end-member oxides closely follow a horizontal trend, indicating that the network is inferring values close to the mean prior in these cases.

When the density information is available, one can train neural networks with wave speeds and density as inputs, i.e. we condition the parameterised posterior of Eq. 3 on density as

$$p(\boldsymbol{x}|\rho, V_P, V_S; \boldsymbol{\alpha}). \tag{4}$$

The prediction performance of these neural networks are shown in Fig. 2 (plots labelled as w/ density). Now, the plots for temperature and SiO_2 are more closely clustered around the diagonal line (compared to when the network was trained using wave speeds only). More importantly, the networks now resolve variations in *MgO* and *FeO* significantly better. This is because the density breaks the trade-off between various thermochemical parameters which have simultaneous effects on wave speed and density. However, the remaining oxides, namely: *CaO*, Al_2O_3 and Na_2O, are still not constrained. Seismic wavespeeds are much more sensitive to the volumetrically abundant components (SiO_2, *MgO* and *FeO*) and therefore mask the true variations of these minor components via trade-off (see below).

To quantify the information gained about a thermochemical parameter, we can calculate the relative difference between prior and posterior distributions, using measures such as Kullback-Leibler (KL) divergence (e.g. Cover and Thomas, 1991) in logarithmic information units or nats. If two distributions are identical, the KL divergence will be zero (i.e. no information gain). Conversely, significantly different distributions give a higher KL divergence value. The divergence of the posterior pdf of thermochemical parameters from their prior distributions at depths 1000 km and 2800 km are shown in Figs. 3 and 4, respectively. The posterior pdfs are predicted using wave speeds (and density) of three rock types shown in Table 2. As expected, when including density along with wave speeds the information gain or KL divergence value is relatively large compared to excluding it. The relative increase in information is particularly large for temperature, *MgO*, SiO_2 and *FeO*, giving the impression that these parameters are well resolved, because the posterior is much narrower than the prior. However, it is prudent to examine the posterior pdfs more closely as we propose in the next paragraph.

3. Inferring lower mantle temperature and bulk composition

3.1. 1-D marginals of temperature and oxides

Comparing the target thermochemical model with only the mean predicted by neural networks ignores the full information provided by pdfs. Moreover, the comparison can be difficult if the posterior pdf has a complex shape. We consider three rock types which may be significant in the Earth's mantle: peridotite, harzburgite and MORB (Eggins et al., 1998), Irifune and Ringwood, 1987 and Perrillat et al., 2006, respectively; see Table 2). The three rock types have different bulk compositions but similar wave speeds. This allows us to understand the sensitivity of wave speeds to thermochemical parameters in the context of trade-offs between parameters and the network's capability to distinguish between different compositions. The wave speeds and density of these three compositions are calculated at 1000 km and 2800 km using Perple_X, at temperatures corresponding to a 1300 °C (1573 K) adiabat.

Using only the V_P and V_S as inputs to the trained neural networks we predict the pdfs on temperature and bulk composition. Then, we compare these pdfs with the corresponding target temperature and compositions given in Table 2. The results are labelled w/o density in Figs. 3 and 4. Although the target temperature and SiO_2 are relatively close to the mean of the posterior pdfs, the uncertainty on these parameters are large. For example, for peridotite at 1000 km the width of SiO_2 pdf is in the range of 38.36–69.09 wt%. All posterior ranges are defined as approximately 99.9 % of the area under the posterior probability density functions. One can compute a smaller percentile such as 68.26 % but it may not provide an accurate estimate of uncertainty on thermochemical parameters due to the complex shape of the probability density functions. For MORB, the temperature pdf at 2800 km and SiO_2 pdf at 1000 km show high probability density at lower values of temperature and SiO_2 than the respective targets. It is important to mention that our prior (thermochemical parameter) ranges are extremely broad (see Appendix A). If we had narrower priors then the apparent uncertainty would be correspondingly smaller (see below). For peridotite, the temperature and *FeO* pdfs predicted by the neural network have width in the range 1012–2600 K and 1–19.56 wt%, respectively. The posteriors for *MgO*, *CaO*, Al_2O_3 and Na_2O resemble their prior distributions. Similarly, the posteriors for *FeO* also look much like the prior without the density. Changing the *FeO* content changes V_P and V_S in (more or less) the same proportions as changing the temperature (e.g. Deschamps and Trampert, 2003). Thus, the effect of iron on wave speeds can easily hide within the changes in temperature leading to broad pdfs for *FeO* and temperature. For example, a change of 0.05 volume fraction of iron would correspond to roughly 360 $\pm$ 160 K temperature change at 1000 km (Deschamps et al., 2012). We next investigated the predicted pdfs when both wave speeds and density were given as inputs to the (trained) neural network (Figs. 3 and 4, labelled w/ density). Temperature uncertainties for all three rock types decreased significantly. For peridotite at 2800 km depth, temperature is in the range 1845–3430 K, and the volumetrically abundant chemical components - SiO_2, *FeO*, *MgO* - are within 34.67–53.2 wt%, 4.27–8.4 wt%, 19.87–47.39 wt%, respectively. The less abundant oxides *CaO*, Al_2O_3 and Na_2O predicted by neural networks have uncertainties in the range 0–14 wt%, 0.83–15 wt%, 0–3.8 wt%, respectively. The width of the pdf shows that temperature and SiO_2 uncertainties are reduced by a factor of approximately 1.86 and 1.52, respectively, for harzburgite at 2800 km. Furthermore, combining wave speeds with density allows the neural networks to constrain the *MgO* and *FeO* content significantly better as shown by the narrower pdfs compared to working with wave speeds alone. It is important to note, although the KL divergence indicates that these paraemters are better resolved, the mean of the posterior is offset with respect to the target. While the target is still within the posterior, a bias is present and is

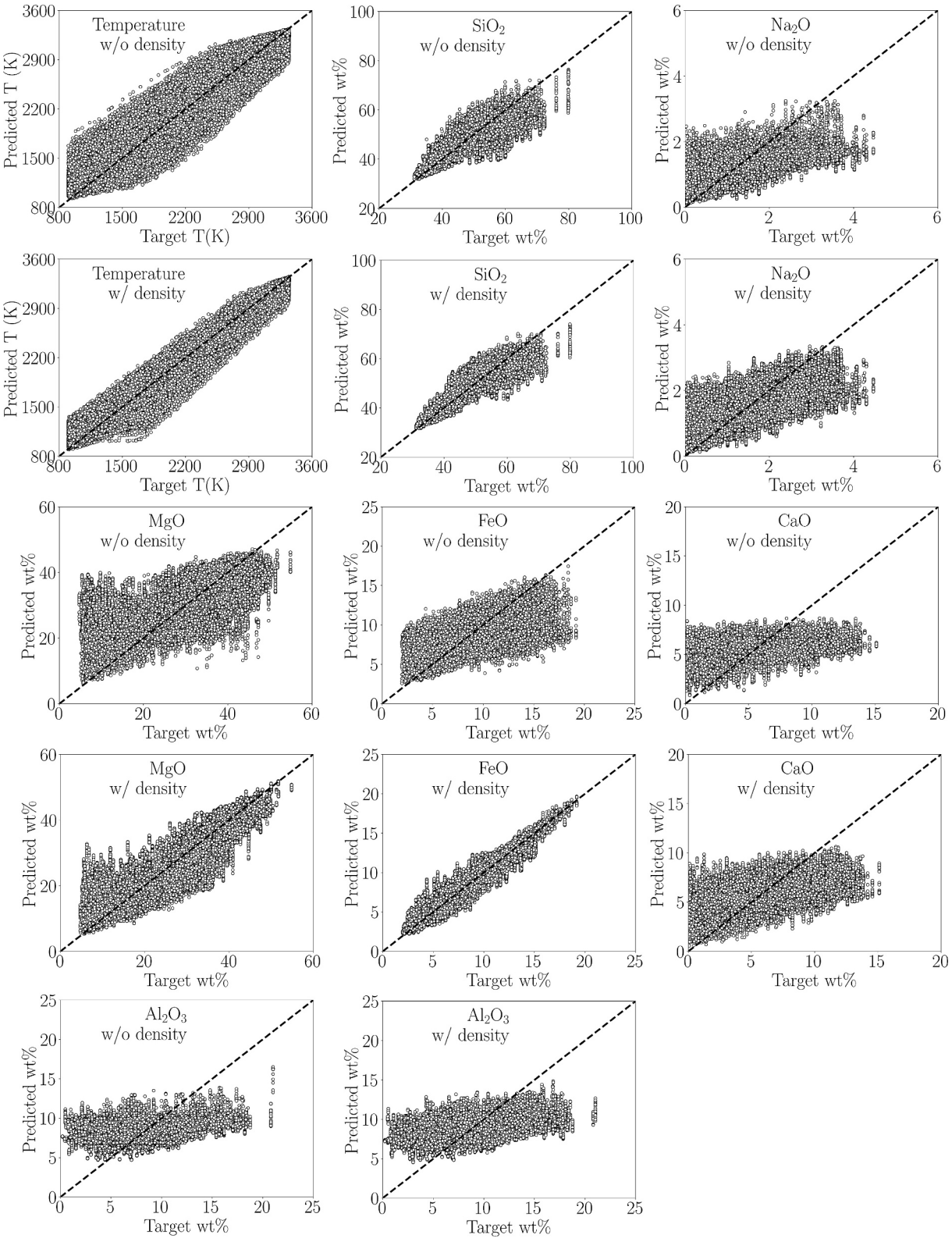


Fig. 2. MDN prediction performance plots at 1000 km depth. As a simple measure of the prediction performance, the mean of the posterior pdf is compared with the (actual or) target thermochemical value of the test set. A unimodal pdf with pronounced peak centred around the target value would plot along the diagonal dashed line. So, closer the predictions cluster around the diagonal line, the better the prediction performance of the trained network. Plots labelled with w/o and w/ density refer to predictions from neural network trained with wave speeds data only and with wave speeds plus density, respectively. When including density information, neural networks resolve variations in temperature, SiO_2, *MgO* and *FeO* significantly better (compared to w/o density).

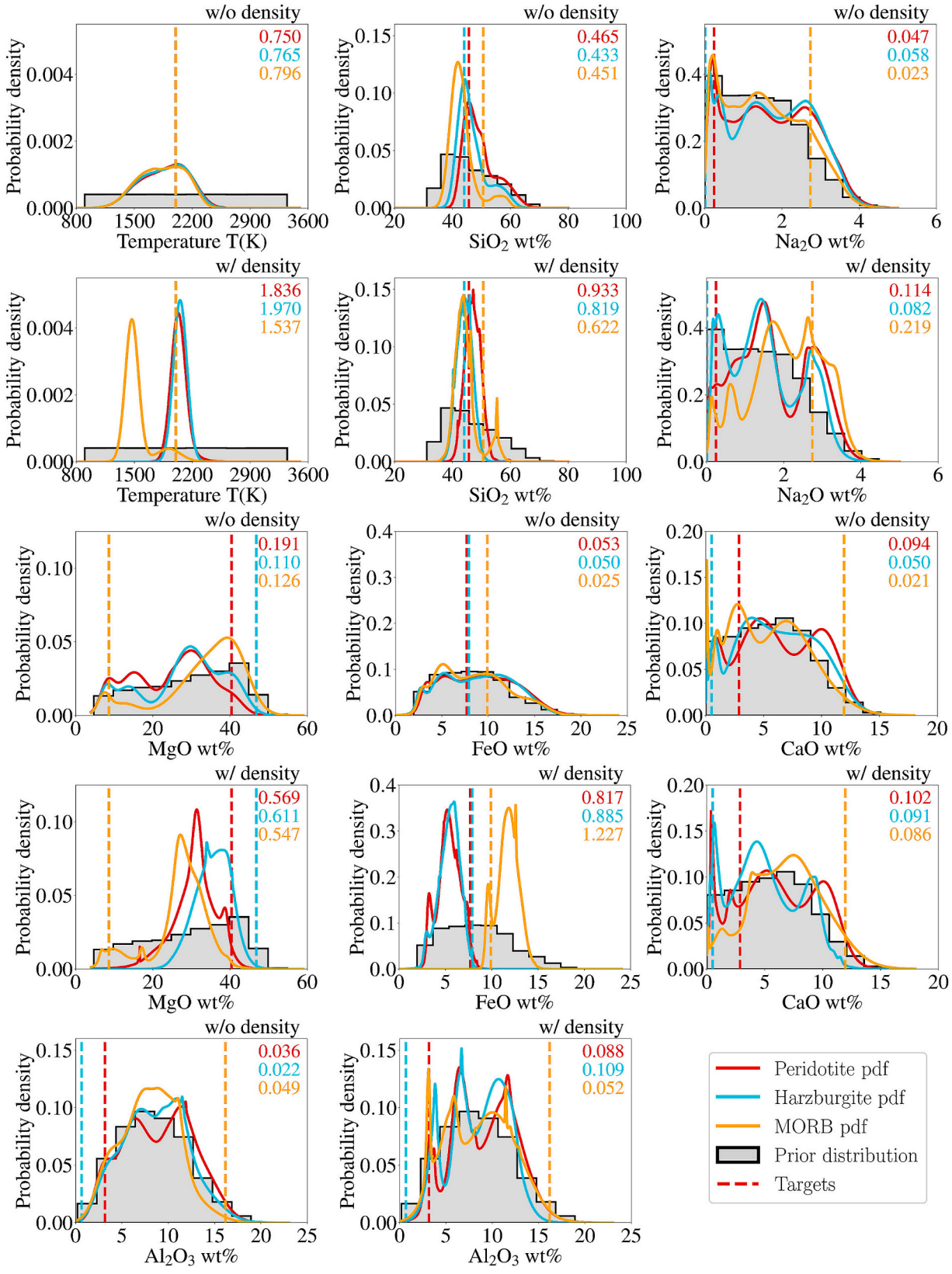


Fig. 3. Neural network predicted pdfs on temperature and composition at 1000 km depth for wave speeds (and density) shown in Table 2. The numbers on the top right corner represent KL divergence values or information gain (in nats). Again, plots labelled with w/o and w/ density refer to predictions from neural networks trained with wave speeds only and with wave speeds plus density, respectively. Temperature, SiO_2, *MgO* and *FeO* pdfs show significant reduction in uncertainties and increase in information gain when density is included in neural network trainings.

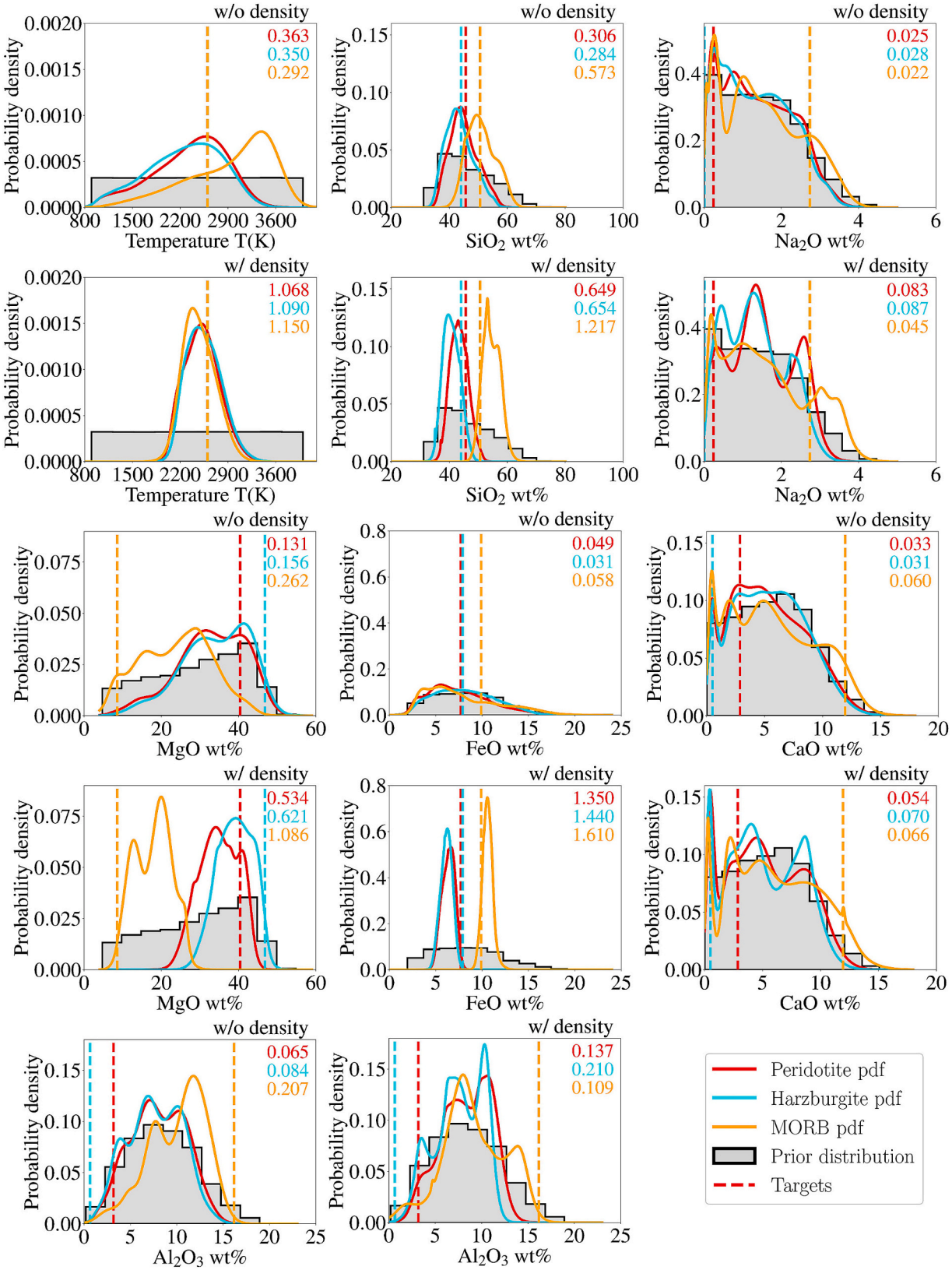


Fig. 4. Same as Fig. 3, but for wave speeds at 2800 km depth. In contrast to 1000 km depth (Fig. 3), the temperature pdfs at 2800 km are wider for all three rock types because the prior distribution for temperature is broader at 2800 km. In general, the same trends apply between 1000 and 2800 km depths when including density.

Table 2
Wave speeds and densities of peridotite (Eggins et al., 1998), harzburgite (Irifune and Ringwood, 1987) and MORB (Perrillat et al., 2006) calculated using Perple_X at 1000 km and 2800 km depth, with temperatures corresponding to a 1300 °C adiabat. The six oxide proportions have been normalised to 100 % (i.e. other components e.g. TiO_2, K_2O which account for <2 wt% were not included.)

			Depth		1000 km					
Rock	V_S	V_P	ρ	T	SiO_2	MgO	FeO	CaO	Al_2O_3	Na_2O
Peridotite	6.318	11.458	4553.31	2000	45.652	40.379	7.683	2.875	3.168	0.243
Harzburgite	6.309	11.391	4527.33	2000	44.085	46.833	7.910	0.505	0.657	0.010
MORB	6.304	11.430	4654.92	2000	50.626	8.611	9.903	11.949	16.173	2.738

			Depth		2800 km					
Rock	V_S	V_P	ρ	T	SiO_2	MgO	FeO	CaO	Al_2O_3	Na_2O
Peridotite	7.395	13.765	5531.56	2600	45.652	40.379	7.683	2.875	3.168	0.243
Harzburgite	7.389	13.744	5507.08	2600	44.085	46.833	7.910	0.505	0.657	0.010
MORB	7.280	13.710	5697.34	2600	50.626	8.611	9.903	11.949	16.173	2.738

another manifestation of the trade-offs between parameters. The temperature pdfs at 1000 km are narrower than at 2800 km, because the prior distribution for temperature is broader at 2800 km than at 1000 km. Our prior temperature distribution is depth dependent and the maximum temperature of the prior is fixed at the melting point of $MgSiO_3$ perovskite (Stixrude et al., 2009). In essence, pdfs of temperature and bulk compositions are a little narrower at one depth than another, but the same general trend applies, i.e. combining density with V_P and V_S reduces uncertainties.

Wide 1-D marginals predominantly represent the trade-off between different thermochemical parameters. If one assumes a fixed composition, obviously, there will be no parameter trade-offs. To demonstrate this more clearly, we used three fixed compositional models shown in Table 3 as priors and computed wave speeds for 1000 different temperatures selected at random between 900 K and 3347 K. Models 1 and 2 are selected to understand the difference in predicted temperature, for the same PREM wave speeds, when there is a relatively small change in the assumed composition. Model 3 is chosen to avoid overlapping its wave speeds with PREM, and the neural network is required to extrapolate. Next, we trained three separate neural networks, i.e. one for each fixed composition. We then put wave speeds taken from PREM at 1000 km depth as inputs to the three trained networks, and obtained a pdf on temperature as output (Fig. 5, right). The pdfs obtained from neural networks trained with wave speeds from models 1 and 2 are very narrow, indicating that the temperature is tightly constrained. But each neural network gives a different temperature as output and the mean temperatures differ by about 400 K. The temperature pdf obtained from the third neural network is multi-modal, and it shows relatively large uncertainty in temperature compared to other two networks. This is because the PREM wave speeds are not in the prior used to train the third neural network (see Fig. 5, left). This in turn forced the network to extrapolate away from the prior range, leading to uncertain temperature prediction. In essence, predictions can be unrealistic if observed seismic wave speeds are not inside the prior data (i.e. data used to train the neural networks) range. Moreover, assuming a unique composition may not be appropriate to constrain absolute temperatures in the Earth's mantle because the temperature depends on the choice of compositions that one assumes.

3.2. Parameter trade-offs

As mentioned earlier, we need to be cautious comparing target values with the posterior mean. A peak in the 1-D marginal doesn't necessarily mean a high probability in the (actual) higher dimensional thermochemical space, because a 1-D marginal only represents integrated information. The advantage of representing the inference of a thermochemical parameter as a 1-D marginal is that they are fast to implement and give a conservative estimate of the uncertainties. The disadvantage is that they cannot capture the correlations between parameters. To get a 2-D marginal, we need to train an additional neural network. For example, to get the 2-D marginal pdf of temperature and *FeO*, we take the 1-D conditional pdf for *FeO* conditioned on temperature (i.e. the extra network) and the 1-D marginal for temperature as

$$p(T, FeO|V_P, V_S) = p(FeO|T, V_P, V_S)\cdot p(T|V_P, V_S). \quad (5)$$

Correlations between temperature and composition are shown in Fig. 6, labelled as w/o density for results with only peridotite wave speeds as input. There are six possible combinations of temperature with oxides. We see a strong linear correlation between temperature and *FeO*. The relations between temperature and other oxides are more complex. 2-D marginal pdfs are an effective way of visualising the dependency between two parameters, however, there might still be interdependencies between three or more thermochemical parameters which are not seen in the 2-D pdfs. In a similar way, the linear correlation between *FeO* and temperature is not apparent by looking into their individual 1-D marginals. Of course, it is difficult to visualise such high-dimensional probability density spaces. Hence, we restrict ourselves to 2-D marginal posteriors in this study, but consider all possible pairs of thermochemical parameters.

We show here an example corresponding to the seismic properties of peridotite, but Eq. 5 is valid for any compositional model. Trade-offs inferred from wave speeds (and density) corresponding to other compositional models (of Section 3) are similar to those inferred for peridotite wave speeds (and density). Once we include the density, the trade-off between *FeO* content and temperature is much less, and we can place a much tighter constraint on both of them. Interestingly, combining density with wave speeds also reveals a linear correlation between temperature and CaO, as well as linear trade-offs between

Table 3
Three fixed isochemical models selected to train neural networks with only temperature variation at 1000 km. T_S is the maximum temperature, and it is about 3347 K as before.

Model	SiO_2	MgO	FeO	CaO	Al_2O_3	Na_2O	T_{min}	T_{max}
model #1	45.008	38.248	8.900	3.056	3.650	1.139	900	T_S
model #2	37.003	41.551	10.013	5.391	4.897	1.146	900	T_S
model #3	63.345	7.096	5.712	7.231	16.558	0.058	900	T_S

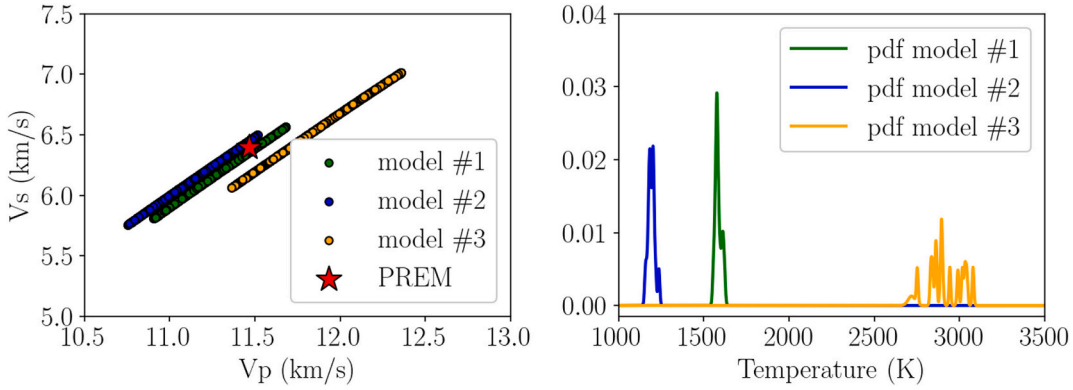


Fig. 5. Left: wave speeds generated from three fixed compositional models of Table 3 for temperatures uniformly drawn from the range 900 and 3347 K. We used the three sets of wave speeds to train three separate neural networks to predict temperature. Red star shows PREM wave speeds at 1000 km depth. The wave speeds of prior model 3 (orange) do not overlap with PREM at all. The neural network trained on these wave speeds is forced to extrapolate when PREM is given as an input. Right: All three neural networks take the same PREM wave speeds as inputs to predict a temperature pdf. The pdfs obtained from the neural networks trained with wave speeds from models 1 (green) and 2 (blue) resolve the temperature extremely well, but the output temperature depends on the choice of assumed compositions and they differ from each other by about 400 K. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

temperature and SiO_2 and between temperature and *MgO*.

Some trade-offs among component oxides are shown in Fig. 7. Although the joint pdfs are complicated, a correlation between pairs of parameters such as SiO_2 - *MgO*, SiO_2 - Al_2O_3 and SiO_2 - Na_2O is revealed without the density information. When including density, oxide pairs *FeO* and Al_2O_3, *FeO* and Na_2O, *FeO* and *CaO*, *CaO* and *MgO*, *FeO* and *MgO*, and *FeO* and SiO_2 (also see Appendix C) show correlation although some oxides show multiple peaks. Furthermore, density limits the trade-off between *MgO* and SiO_2 to a smaller range compared to results with wave speeds only. However, correlation between oxide pairs SiO_2 - Na_2O and SiO_2 - Al_2O_3 remain. *Al* and *Na* are minor components and in most rock types (i.e. other than MORB-type rocks) the abundances do not exceed more than a few percent (Table 2). Hence the effect of these parameters on the inferred SiO_2 content should be small for rocks similar to pyrolite/peridotite/harzburgite. The main advantage of including density is that it helps to better constrain the most abundant and dynamically relevant compositional end-members, namely, the SiO_2, *MgO* and *FeO*. The trade-offs identified here have no mineral physics origin, but are entirely due to the chosen prior and the under-determined nature of the problem.

4. Lower mantle temperature and compositions from PREM

One dimensional reference models of the Earth (e.g. PREM - Dziewonski and Anderson, 1981, or ak135- Kennett et al., 1995) are often used to infer the average thermal and chemical properties of the mantle (e.g. Jackson, 1998; Deschamps and Trampert, 2004; Matas et al., 2007; Murakami et al., 2012; Cottaar et al., 2014; Houser et al., 2020). Although the spherically symmetric reference model is a good first order approximation for wave speeds, it only provides an averaged Earth that is a biased representation of a potentially complex three dimensional interior (Cobden et al., 2009). Nevertheless, let's predict the temperature and composition for the PREM wave speeds and density using the neural networks trained with broad priors of Sub-section 2.1. The results at 1000 km depth are shown in Fig. 8. The combination of density with wave speeds reduces the uncertainties (compared to working with wave speeds only) in temperature and SiO_2, *MgO* and *FeO* which are the most abundant parameters. The temperature, *MgO*, *FeO* and SiO_2 predicted by neural networks are in the range 1260–1850 K, 22.86–51.34 wt%, 2.86–10.22 wt% and 31.33–48.16 wt%, respectively (when including density). For comparison, the bulk compositions of peridotite, harzburgite and MORB are plotted together with the pdfs. MORB falls on the edge of the pdfs. The *FeO* pdf shows a peak almost at the *FeO* wt% of peridotite and harzburgite rocks. Whilst both harzburgite and peridotite SiO_2 plot on the upper edge of the pdf, the *MgO* pdf favours peridotite. The pdfs of minor components, namely, Al_2O_3, Na_2O and *CaO* resemble their prior distributions. In this study, we have been working with a very broad prior which includes many intermediate compositions between MORB and peridotite that are unlikely to be abundant (or exist at all) in the mantle. One can restrict the prior temperature and/or composition ranges with the help of additional constraints (e.g. from geochemistry and petrology) and/or include an additional observable (e.g. attenuation) to further constrain the temperature and composition pdfs.

5. Inferring temperature and bulk composition from V_S and ρ

So far we have inferred thermochemical parameters using compressional and shear wave speeds with or without density. In this section we predict thermochemical parameters from shear wave speed and density of peridotite (as an example). The combination of V_S and ρ constrains the temperature and *FeO* (Fig. 9) better than when V_S is combined with V_P. Pdfs of temperature and *MgO* show a width similar to working with V_P and V_S, but the prediction from V_S and density favours slightly higher temperature and *MgO* content. The *FeO* prediction from V_S and ρ shows only a slightly broader pdf compared to that from using both wave speeds (V_P and V_S) together with density. Combining V_P with density (not shown here) also predicts a similar *FeO* pdf (only narrower by just 1–2 wt%) to that obtained from combining shear wave speed and density. This shows that the density is an important parameter for constraining iron. The SiO_2 posterior pdf predicted from V_S and ρ of peridotite follows its prior distribution, in contrast to working with V_P and V_S. In a similar way, we found that the SiO_2 posterior predicted from V_P and ρ (also not shown here) also follows the prior. This highlights the significance of combining V_P and V_S in constraining SiO_2. Furthermore, in comparison to the pdf predicted by wave speeds alone, combining density with wave speeds further reduces the SiO_2 uncertainty.

The trade-offs between temperature and the most abundant parameters are shown in the 2-D marginals in Fig. 10. The region of *FeO* - temperature trade-off shows significant probability density at slightly higher *FeO* and lower temperature than working with both wave speeds and density (see Fig. 6 for comparison). However, the inherent trade-off is similar in both cases which is represented in the region of high probability density. In general, T - SiO_2, T - *MgO*, *FeO* - *MgO*, *FeO* - SiO_2 and SiO_2 - *MgO* all show trade-offs similar to working with both wave

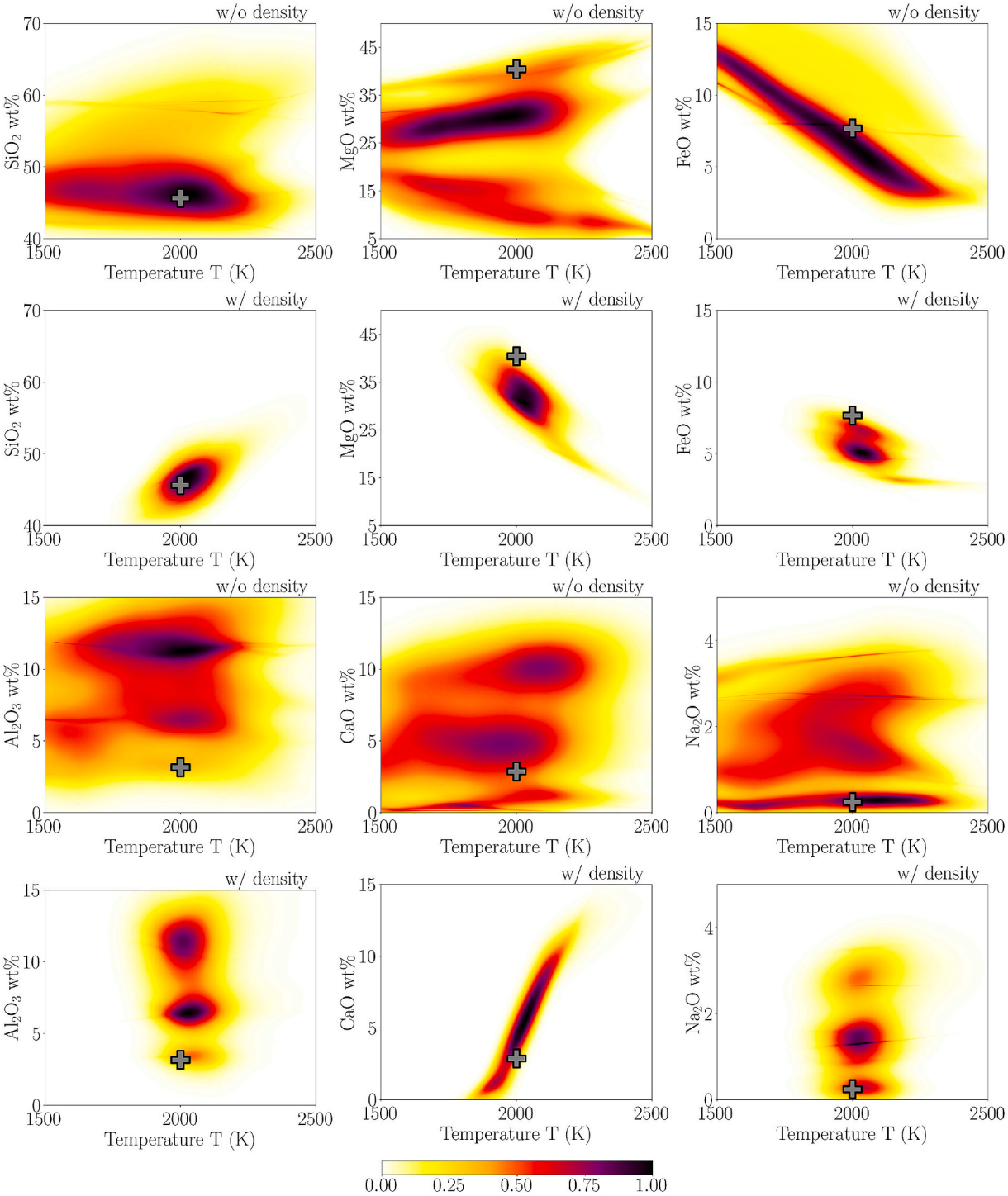


Fig. 6. 2-D marginal pdfs of temperature with all six oxides inferred from the wave speeds (and density) corresponding to peridotite (Table 2) at 1000 km depth. Plots labelled w/o and w/ density are results from neural networks trained using wave speeds only and with wave speeds plus density, respectively. Darker colour shows higher likelihood. We normalised the likelihood values to [0–1] for plotting purpose. Solid"+" symbol represents the target value, i.e. the actual value for peridotite (as in Table 2). A linear correlation between *FeO* and temperature is observed without the density information. Other oxides and temperature show complex dependencies. However, including density significantly reduces the joint probability space of trade-off by breaking down these complex dependencies.

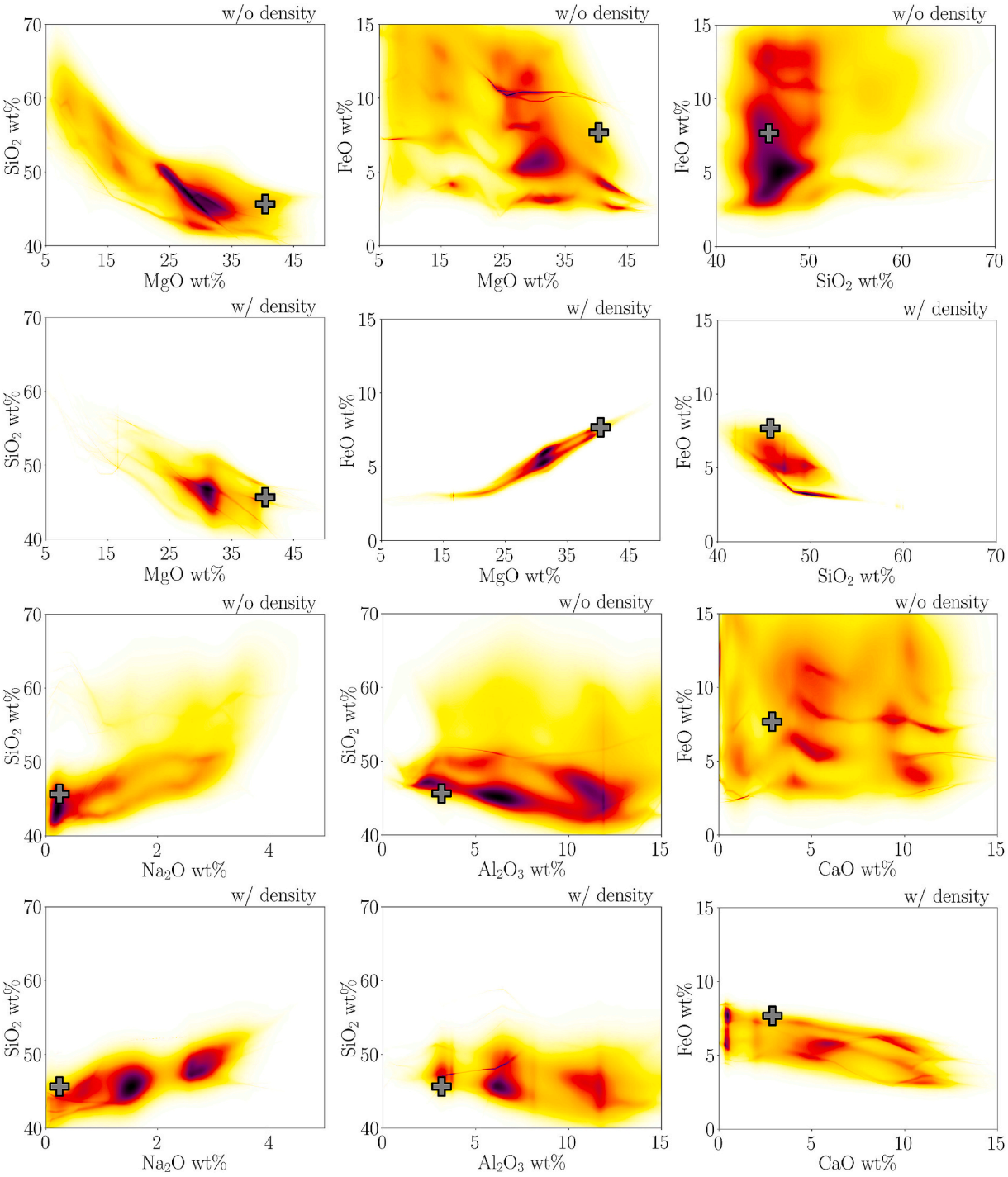


Fig. 7. 2-D marginal pdfs for different pairs of oxides inferred from the wave speeds corresponding to the peridotite model of Section 3 at 1000 km. Labels (w/o and w/ density) and the colour scheme are same as Fig. 6. Solid"+" symbol represents the target value.

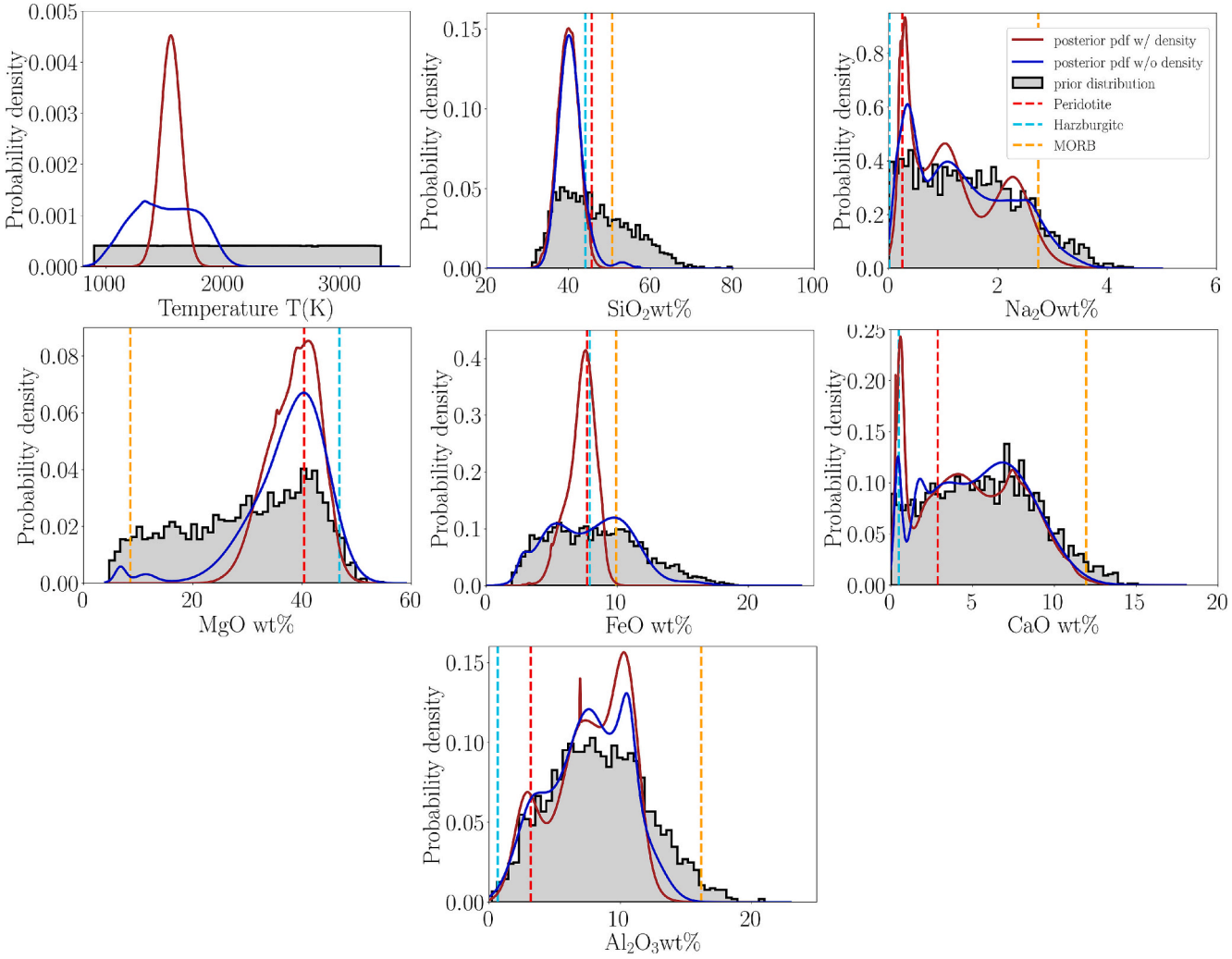


Fig. 8. 1-D marginal thermochemical pdfs at 1000 km for PREM wave speeds and density. The oxide pdfs are compared with oxide wt% of three rock types given in Table 2. The pdfs inferred using wave speeds data only are labelled as w/o density, and those with wave speeds plus density are labelled w/ density.

speeds and density, but in a wider region. Overall, the temperature and the most abundant oxide end-members (SiO_2, MgO and FeO) are better constrained when all three seismic observables are combined. In general, when working with V_S and ρ only, the target thermochemical parameters are on the edge or away from the high probability region of the 2-D marginal pdfs. If no additional observables are available, one would need to make the prior in composition narrower (using independent information) to get a tighter constraint on thermochemical parameters.

6. Narrower priors

The prior dataset used to train the neural networks is constantly evolving as new experimental and theoretical constraints on mineral thermoelastic properties become available. Defining the prior is also based on user-driven choices. With the broad priors we have used in this study, the temperature uncertainty, i.e. the width of the posterior pdf, at 1000 km for harzburgite V_P and V_S is 1542 K, whereas including ρ reduces the uncertainty to 555 K. Recently, Houser et al., 2020 examined the role of temperature uncertainties in discriminating between different compositions. Their approach shows that in order to discriminate between two compositions, perovskite and harzburgite, throughout the lower mantle the maximum temperature uncertainty should be ±150 K (±200 K to ±300 K in spin transition region). In an application to seismic observables with additional prior constraints (e.g. from petrology or geodynamics) one might choose to restrict the temperature and/or composition ranges. This in turn will reduce the widths of the pdfs in the output from the neural networks.

If we systematically reduce the prior range of thermochemical parameters by 20 % and 40 %, the resulting uncertainities of the main parameters reduce as well, albeit by somewhat smaller amounts (see Table 4). In this test, we reduced the upper bound of the range of chemical parameters shown in Table 1, except for MgO for which we removed lower MgO models. Temperature range is symmetrically narrowed from both ends. As shown in Table 4, reducing the priors by 40 % constrains the SiO_2 and FeO contents within 14 wt% and 5.4 wt%, respectively. In essence, restricting the priors to a narrower range reduces the possibilities for trade-offs between thermochemical parameters, leading to narrower 1-D marginals.

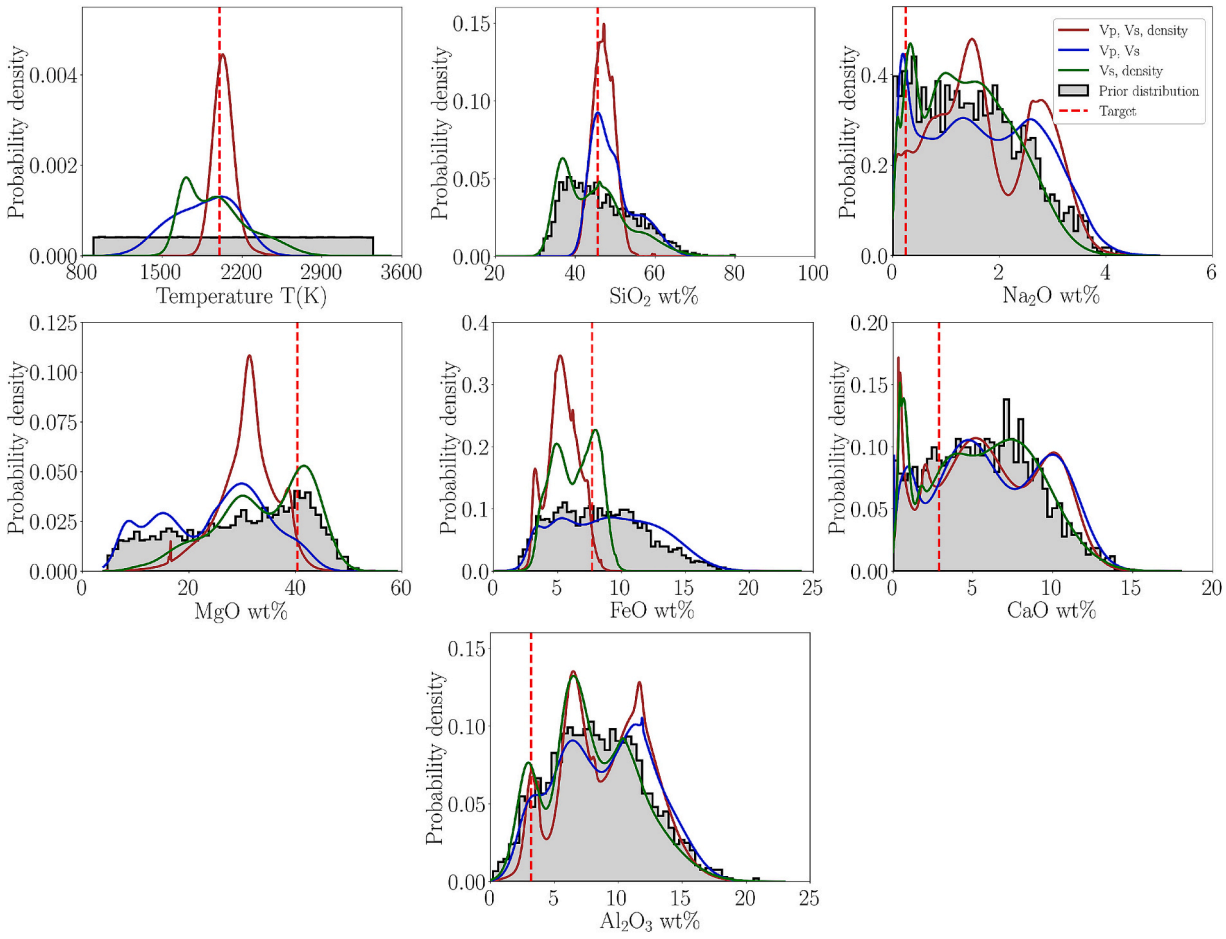


Fig. 9. 1-D marginals of temperature and oxide end-members at 1000 km inferred from V_S and ρ of peridotite rock from Table 2. The marginals are compared with predictions from V_P, V_S only and V_P, V_S together with ρ. The V_S and ρ pair constrains the temperature and *FeO* contents. The pair V_P and V_S constrain the SiO_2 content.

7. Discussion

Using an ensemble of neural networks, in particular, Mixture Density Networks (MDNs), we can infer the posterior probability density functions of temperature and bulk composition for a given value of V_P and V_S (±density). The computational demand for obtaining prior samples - i.e. calculating seismic properties for thermochemical models (Perple_X in our case) - and training the neural networks depends on several factors (details in Käufl et al., 2016): number of networks in an ensemble, number of parameters of each ensemble member, size of the prior samples, choice of error minimisation algorithm, code efficiency, etc. Roughly, to train an independent ensemble member - i.e. one MDN - with 3 input nodes, 45 hidden nodes, 5 Gaussian kernels (see architecture in Appendix B) and approximately 5.3×10^5 training data, it takes about 9.6 min of wall-clock time (about 14.5 min of CPU-time) in a compute node with two AMD EPYC 7451 24-Core Processors running at 2.3 GHz. Most time is spent generating the thermochemical dataset, lucklily it is trivial to parallelise Perple_X and run each Gibbs minimisation on a separate core.Once trained, the same network takes just a fraction of a second to evaluate a pdf at 2701 points. In comparison to a Monte Carlo type method, which is a posterior sampling method, our neural networks use samples drawn from prior ranges of temperature and composition. In principle, once trained the same trained neural networks can be re-used to interpret wave speeds (and density) obtained from different locations or different seismic studies, which makes the inference step extremely efficient. In contrast, in order to get the temperature and composition as well as trade-offs between thermochemical parameters, Monte Carlo type methods require re-sampling of the posterior every time we need to interpret new wave speeds which is inevitably less efficient.

From the point of view of seismic interpretation, the most important compositional end-members are the most volumetrically abundant oxides (*MgO*, SiO_2 and *FeO*), and we have seen that one can get a significantly better constraint on them by including density. In addition, including a fourth observable, such as attenuation (e.g. Hwang and Ritsema, 2011; Karaoğlu and Romanowicz, 2018; Konishi et al., 2020), could help to reduce the remaining trade-offs between thermochemical parameters. Although attenuation may be less sensitive to composition, it can help to constrain the temperature and grain size (Talavera-Soza et al., 2025).- However, the intrinsic attenuation depends on a number of parameters including activation energy, activation volume, frequency dependence and grain size, all of which are poorly constrained for the lower mantle. Hence, while attenuation could theoretically provide an extra constraint with which to reduce the trade offs between thermal and chemical parameters, in practise the results would only be meaningful if uncertainties in the anelasticity parameters were also taken into account.

Although seismic wave speeds depend on material properties such as the size and orientation of mineral grains, for simplicity our modelling is with isotropic mineral parameters. This is a good first order approximation (Marquardt and Thomson, 2020), and we do not attempt to estimate modelling errors due to unaccounted anisotropy. In addition, the results shown so far are based on anharmonic wave speeds. We trained two neural networks (including and excluding density) in which the

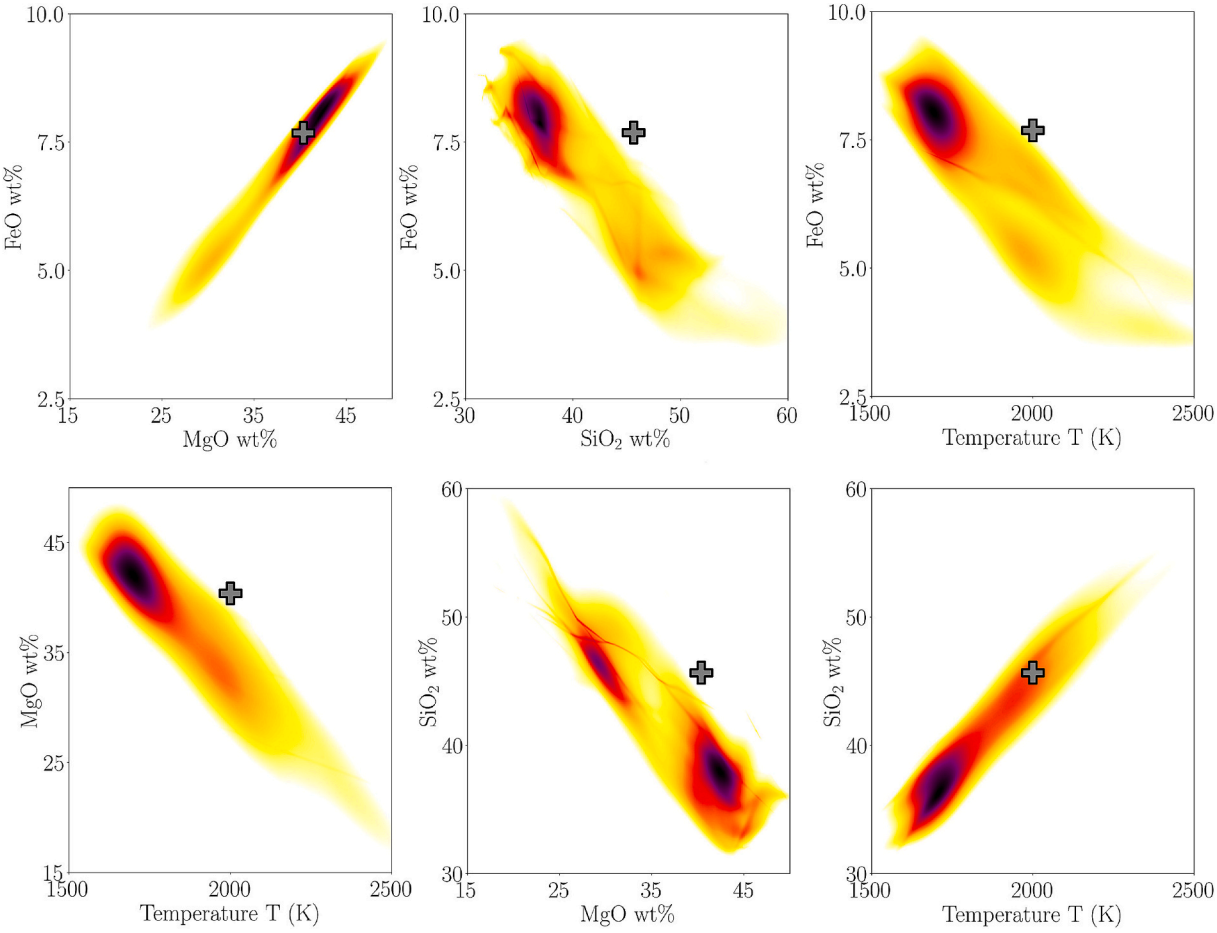


Fig. 10. 2-D marginal pdfs for pairs of different thermochemical parameters inferred from the shear wave speed and density corresponding to peridotite rock (Table 2). Colour scheme is according to Fig. 6. Solid"+" symbol represents the target value.

Table 4
The width (upper minus lower) of the posterior probability density functions of temperature and the most volumetrically abundant oxides inferred using broad and systematically narrowed priors (by 20 % and 40 % of broad priors). The thermochemical parameters are inferred using harzburgite wave speeds (± density) at 1000 km (see Table 2). T is in K and oxides in wt%.

		Broad prior		20 % narrower prior		40 % narrower prior	
		Lower	Upper	Lower	Upper	Lower	Upper
Without density	T	1018	2560	1120	2463	1397	2470
	SiO_2	36.44	64.91	36.33	55.13	36.91	50.80
	MgO	4.00	50.85	13.07	50.95	22.13	50.67
	FeO	0.70	18.98	1.25	16.49	0.98	13.38
With density	T	1775	2330	1790	2270	1820	2223
	SiO_2	35.70	52.70	36.27	51.27	36.38	50.22
	MgO	20.52	49.89	22.97	47.82	27.18	48.12
	FeO	1.70	8.70	1.79	8.62	3.13	8.53

anharmonic wave speeds were corrected for temperature-dependent anelasticity, following the model of Deschamps et al., 2019. We assumed fixed values for the various anelasticity parameters (see Appendix D). Temperature pdfs inferred from anelastic wave speeds are only slightly narrower (Fig. 11) than those inferred from anharmonic wave speeds, showing negligible influence of anelasticity (Brodholt et al., 2007). In this study, we used ρ, V_P and V_S. However, the inputs to the neural networks can readily be modified by conditioning the Eq. 4 for other parameters, including attenuation, anisotropy, or the depth and sharpness of seismic discontinuities. Seismic discontinuities are particularly relevant for transition zone and D″ studies. It is important to mention that any uncertainty in the wave speeds and density, if available, from observed seismic data or models can be propagated through the MDNs. We chose to illustrate our tool at fixed depth/pressure. Often a continuous description in pressure-space is desired (e.g. Afonso et al., 2015), it is trivial to condition our posteriors on pressure as well.

Recently, we used neural networks to characterise the elastic properties of MgO as a function of pressure and temperature (Rijal et al., 2021; Rijal et al., 2023). We found that the experimental uncertainties in density and seismic wave speeds may be comparable to or larger than reported variations in these properties in the lower mantle. These uncertainties would contribute further to the ranges of temperature and composition that can fit a given set of wave speeds and density. In the current study we chose to use an explicit EOS (Stixrude and Lithgow-Bertelloni, 2005, 2011) because this is currently one of the most comprehensive and self-consistent datasets for the elastic properties of

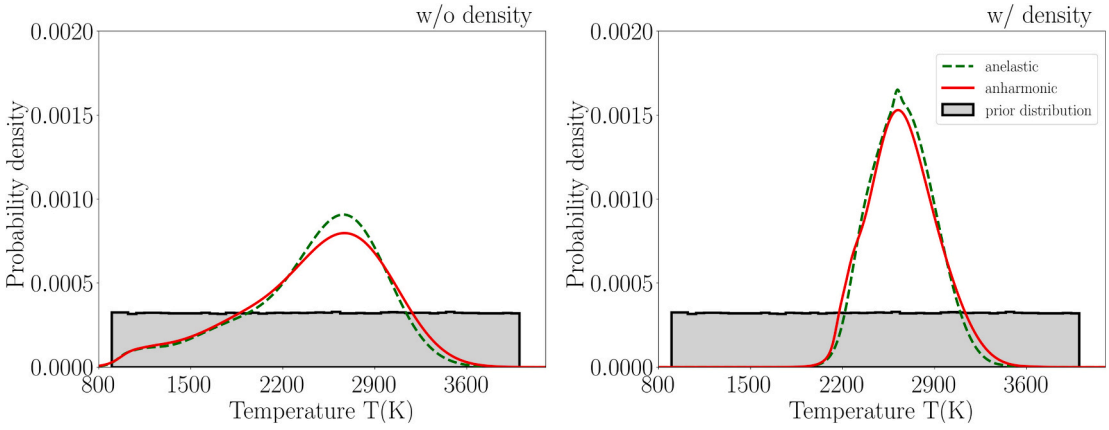


Fig. 11. 1-D marginal pdfs for temperature at 2800 km inferred from wave speeds corrected for anelasticity and anharmonic wave speeds. Left: inputs are wave speeds only. Right: inputs are wave speeds and density. Temperature pdfs inferred from anelastic wave speeds are only slightly narrower than those inferred from anharmonic wave speeds.

mantle minerals. We did not include uncertainties on the mineral elastic or seismic properties during the modelling or training, but it would be straightforward to do so. This would give even more conservative uncertainty ranges than the wide ranges seen in this study, although the ranges could in turn be reduced by narrowing the priors as discussed earlier. Trade-offs indentified here depend on the prior and non-uniqueness of the problem. Often systematics in composition-space are known in mineral physics, but not communicated. Including these systematics in the prior for generating the dataset could help defining a more realistic prior. The effect of such restructions on the prior would need further investigations and are beyond the scope of the current paper.

8. Conclusion

We have developed a neural networks tool to map seismic observables into thermochemical parameters.We use a Bayesian formulation which infers posterior marginals conditioned on a number of input parameters. The networks are trained on syntheic models, which associate thermochemical parameters to the seismic observables. The models are generated using appropriate prior information from mineral physics databases. We chose to use PerpleX (Connolly, 1990; Connolly, 2005) to generate the models, but other packages can be used equally well, such as e.g. MAGEmin (Riel et al., 2022 or BurnMan (Myhill et al., 2023). Input observables and additional parameters as well as output targets are fully flexible.

To illustrate and test our tool, we quantify the full ranges of thermochemical parameters that fit V_P and V_S ($\pm$density) at a given depth in the lower mantle. It is challenging from only two or three observables to infer seven thermochemical parameters (temperature and six oxide end-members) because of the non-uniqueness of this inverse problem. The MDN is an efficient method for getting an overview of the uncertainties and trade-offs which are often neglected during seismic interpretation in the form of a probability density function. In this work, we have tested this approach at fixed lower mantle depths but the concept can be applied at any depth where data is available on the elastic properties of constituent minerals. For this, one could either include pressure as an additional input parameter to the neural networks or train separate networks at each depth of interest.

In our lower mantle example, we made a few noteworthy observations. With the neural networks trained on wave speed (V_P and V_S) data only, one can start to constrain the temperature and SiO_2 in the lower mantle. Without the density, we obtain limited sensitivity to other components of bulk composition and large uncertainties on the temperature. This is partly because we allowed a wide range of possible lower mantle temperature and compositions in our training dataset and the resulting trade-off between different thermochemical parameters gives us broad 1-D marginals. By using density together with V_P and V_S, we break down some complex dependencies between thermochemical parameters. As a result, one obtains a tighter constraint on the temperature as well as the *Fe*, *Mg* and *Si* content of the bulk composition. In all cases there is limited sensitivity to the minor components (Al_2O_3, CaO and Na_2O), regardless that the seismic properties of these end-members arepoorly constrained at lower mantle depths from a mineral physics point of view (Marquardt and Thomson, 2020).

This work highlights the need to have both V_S and ρ to constrain the iron content, while for constraining the SiO_2 (or MgO) content, it is important to have both V_P and V_S. We could incorporate additional parameters such as attenuation, anisotropy or properties of seismic discontinuities in the same tool with the potential to enhance the robustness of seismic interpretations and avoid the potential to over interpret seismic signals. The neural network tool is freely available for academic use (see Section Code availability).

Code availability

The code used in this study and the details on how to use it can be found in this bitbucket repository: https://bitbucket.org/ashim_rijal/mdns_tensorflow/src/master/ or by contacting the corresponding author of this literature.

CRediT authorship contribution statement

Ashim Rijal: Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. **Laura Cobden:** Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing. **Jeannot Trampert:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

AR and LC received funding from the Dutch Research Council (NWO) on grant number 016.Vidi.171.022. We thank JC Afonso for thoutful comments, which improved our manuscript.

Appendix A. Prior distributions

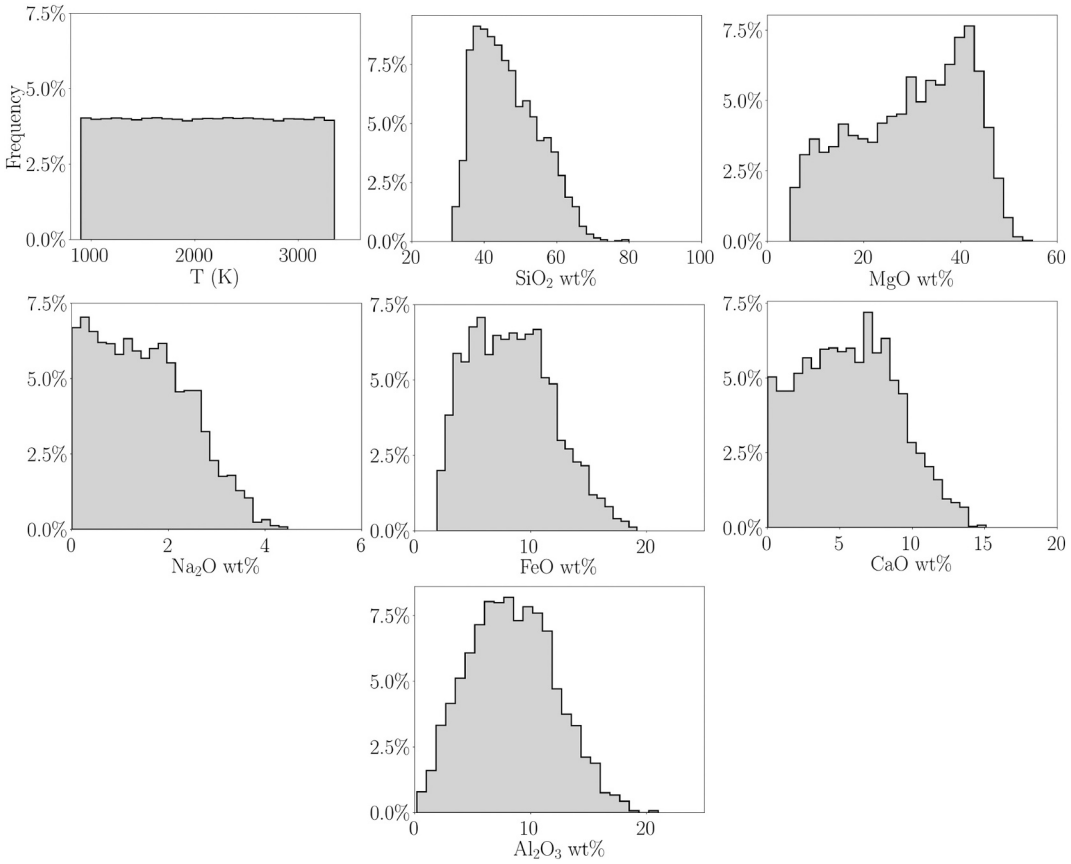


Fig. A.12. Prior distributions of temperature (T) in Kelvin (K) and oxides (SiO_2, MgO, FeO, CaO, Al_2O_3 and Na_2O) in wt%. at 1000 km depth.

Appendix B. MDN architecture

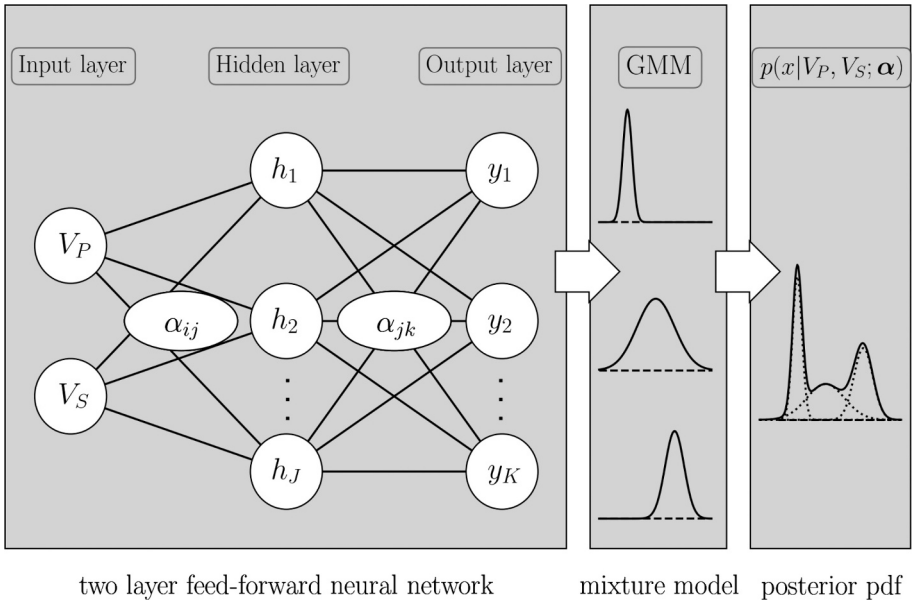


Fig. B.13. A Mixture Density Network (figure modified after Bishop, 1994 and Rijal et al., 2021) that approximates the posterior probability density function of a thermochemical parameter (x) for a given pair of wave speeds (V_P and V_S). In order to approximate the posterior, a conventional feed-forward network is combined with a Gaussian Mixture Model (GMM). The GMM consists of Gaussian functions, and the mean, standard deviation and weight of each Gaussian is computed from the output of the feed-forward network y_k. The hidden nodes and the weights and biases of the feed-forward network are denoted by h_j and $\boldsymbol{\alpha}$, respectively.

Table B.5
The number of hidden nodes used in neural networks to approximate the pdf of thermochemical parameters.

Thermochemical parameter	Hidden nodes without density	Hidden nodes with density
Temperature	20–40	20–60
FeO	25–50	10–25
MgO	15–35	60–105
SiO_2, *CaO*, Na_2O, Al_2O_3	15–65	15–65

Appendix C. Correlations between compositional parameters

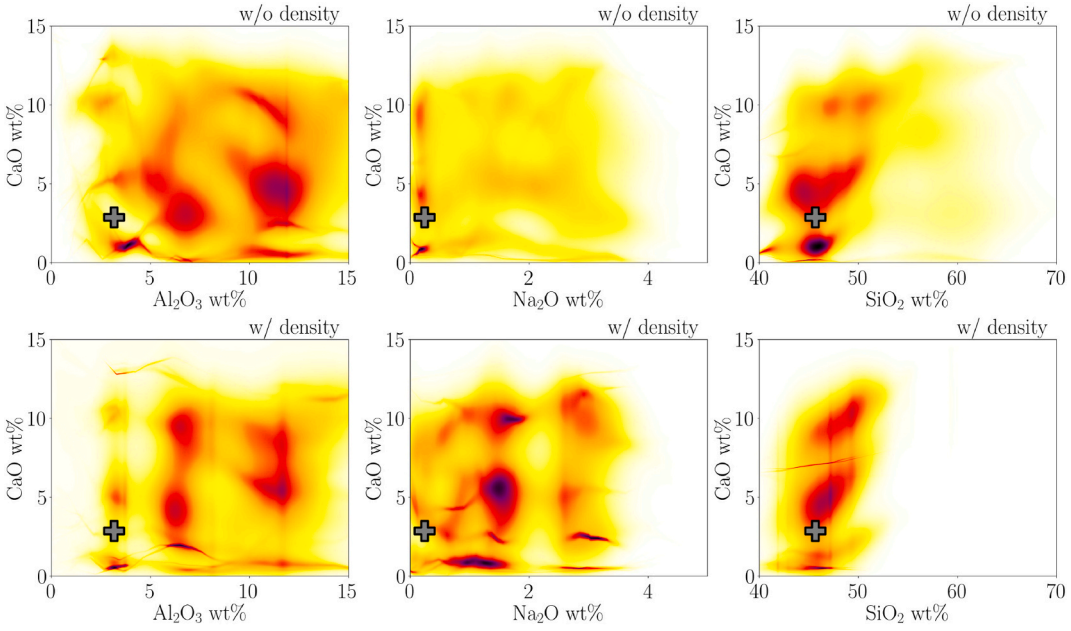


Fig. C.14. 2-D marginal pdfs for different pairs of oxides for the peridotite model of Section 3 at 1000 km. Results from neural networks trained without and with density information are labelled as w/o and w/ density, respectively. Colour scheme is according to Fig. 6. Solid"+" symbol represents the target value.

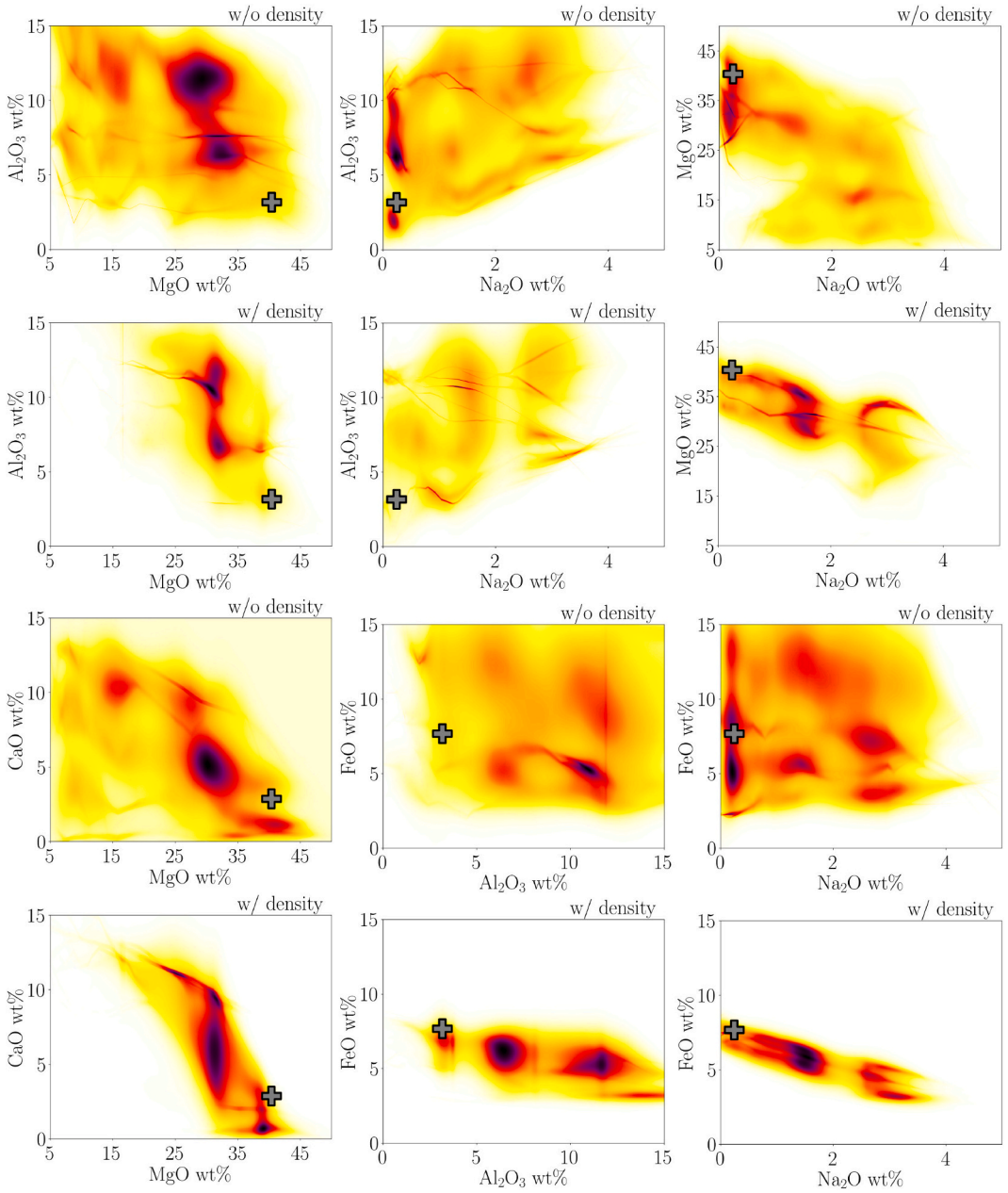


Fig. C.15. 2-D marginal pdfs for different pairs of oxides for the peridotite model of Section 3 at 1000 km. Results from neural networks trained without and with density information are labelled as w/o and w/ density, respectively. Colour scheme is according to Fig. 6. Solid"+" symbol represents the target value.

Appendix D. Parameters for temperature-dependent anelasticity

Table D.6
Anelasticity parameters taken form Deschamps et al., 2019 to correct anharmonic wave speeds for temperature-dependant anelasticity.

Parameter	Value
Seismic period	1 s
Frequency dependance	0.274
Activation energy	286 kJ/mol
Activation volume	1.2×10^{-6} m^3/mol
Reference quality factor	312
Core-mantle-boundary temperature	3500 K

Data availability

Data will be made available on request.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., 2015. TensorFlow: Large-scale machine learning on heterogeneous systems. URL. https://www.tensorflow.org/. Software available from. tensorflow.org.

Afonso, J.C., Zlotnik, S., Diaz, P., 2015. An efficient and general approach for implementing thermodynamic phase equilibria information in geophysical and geodynamic studies. Geochem. Geophys. Geosyst. 16, 3767–3777.

Afonso, J.C., Ben-Mansour, W., O'Reilly, S.Y., Griffin, W.L., Salajegheh, F., Foley, S., Begg, G., Selway, K., Macdonald, A., Januszczak, N., Fomin, I., Nyblade, A.A., Yang, Y., 2022. Thermochemical structure and evolution of cratonic lithosphere in central and southern africa. Nat. Geosci. 15, 405–410.

Ballmer, M., Houser, C., Hernlund, J., Wentzcovitch, R., Hirose, K., 2017. Persistence of strong silica-enriched domains in the earth's lower mantle. Nat. Geosci. 10, 02. https://doi.org/10.1038/ngeo2898.

Bishop, C.M., 1994. Mixture Density Networks. Technical report. Aston University, Birmingham.

Bishop, C., January 1995. Neural Networks for Pattern Recognition. Oxford University Press. URL. https://www.microsoft.com/en-us/research/publication/neural-networks-pattern-recognition-2/.

Brodholt, J., Helffrich, G., Trampert, J., 2007. Chemical versus thermal heterogeneity in the lower mantle: the most likely role of anelasticity. Earth Planet. Sci. Lett. 262, 10. https://doi.org/10.1016/j.epsl.2007.07.054.

Cammarano, F., Romanowicz, B., 2007. Insights into the nature of the transition zone from physically constrained inversion of long-period seismic data. Proc. Natl. Acad. Sci. 104 (22), 9139–9144.

Cammarano, F., Deuss, A., Goes, S., Giardini, D., 2005a. One-dimensional physical reference models for the upper mantle and transition zone: Combining seismic and mineral physics constraints. J. Geophys. Res. Solid Earth 110 (B1). https://doi.org/10.1029/2004JB003272. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JB003272.

Cammarano, F., Goes, S., Deuss, A., Giardini, D., 2005b. Is a pyrolitic adiabatic mantle compatible with seismic data? Earth Planet. Sci. Lett. 232 (3), 227–243. https://doi.org/10.1016/j.epsl.2005.01.031. ISSN 0012-821X. URL. http://www.sciencedirect.com/science/article/pii/S0012821X05000804.

Chantel, J., Frost, D.J., McCammon, C.A., Jing, Z., Wang, Y., 2012. Acoustic velocities of pure and iron-bearing magnesium silicate perovskite measured to 25 gpa and 1200 k. Geophys. Res. Lett. 39 (19). https://doi.org/10.1029/2012GL053075. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012GL053075.

Cobden, L., Goes, S., Ravenna, M., Styles, E., Cammarano, F., Gallagher, K., Connolly, J. A.D., 2009. Thermochemical interpretation of 1-D seismic data for the lower mantle: the significance of nonadiabatic thermal gradients and compositional heterogeneity. J. Geophys. Res. Solid Earth 114 (B11). https://doi.org/10.1029/2008JB006262. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JB006262.

Cobden, L., Mosca, I., Trampert, J., Ritsema, J., 2012. On the likelihood of post-perovskite near the core–mantle boundary: a statistical interpretation of seismic observations. Phys. Earth Planet. Inter. 210-211, 21–35. https://doi.org/10.1016/j.pepi.2012.08.007. ISSN 0031-9201. URL. https://www.sciencedirect.com/science/article/pii/S0031920112001537.

Cobden, L., Trampert, J., Fichtner, A., 2018. Insights on upper mantle melting, rheology, and anelastic behavior from seismic shear wave tomography. Geochem. Geophys. Geosyst. 19 (10), 3892–3916. https://doi.org/10.1029/2017GC007370. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2017GC007370.

Cobden, L., Zhuang, J., Lei, W., Wentzcovitch, R., Trampert, J., Tromp, J., 2024. Full-waveform tomography reveals iron spin crossover in earth's lower mantle. Nat. Commun. 15 (1961).

Connolly, J., 1990. Multivariable phase diagrams: an algorithm based on generalized thermodynamics. Am. J. Sci. 290, 666–718. https://doi.org/10.2475/ajs.290.6.666.

Connolly, J., 2005. Computation of phase equilibria by linear programming: a tool for geodynamic modeling and an application to subduction zone decarbonation. Earth Planet. Sci. Lett. 236, 524–541. https://doi.org/10.1016/j.epsl.2005.04.033.

Cottaar, S., Heister, T., Rose, I., Unterborn, C., 2014. Burnman: a lower mantle mineral physics toolkit. Geochem. Geophys. Geosyst. 15 (4), 1164–1179. https://doi.org/10.1002/2013GC005122. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013GC005122.

Cover, T.M., Thomas, J., 1991. Elements of Information Theory. John Wiley & Sons, New York.

Dalton, C.A., Ekström, G., Dziewoński, A.M., 2008. The global attenuation structure of the upper mantle. J. Geophys. Res. Solid Earth 113 (B9). https://doi.org/10.1029/2007JB005429. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JB005429.

de Wit, R., Trampert, J., 2015. Robust constraints on average radial lower mantle anisotropy and consequences for composition and texture. Earth Planet. Sci. Lett. 429, 101–109. https://doi.org/10.1016/j.epsl.2015.07.057. ISSN 0012-821X. URL. http://www.sciencedirect.com/science/article/pii/S0012821X15005002.

Debayle, E., Ricard, Y., 2012. A global shear velocity model of the upper mantle from fundamental and higher rayleigh mode measurements. J. Geophys. Res. Solid Earth 117 (B10). https://doi.org/10.1029/2012JB009288. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012JB009288.

Deschamps, F., Trampert, J., 2003. Mantle tomography and its relation to temperature and composition. Phys. Earth Planet. Inter. 140, 12. https://doi.org/10.1016/j.pepi.2003.09.004.

Deschamps, F., Trampert, J., 2004. Towards a lower mantle reference temperature and composition. Earth Planet. Sci. Lett. 222 (1), 161–175. https://doi.org/10.1016/j.epsl.2004.02.024. ISSN 0012-821X. URL. http://www.sciencedirect.com/science/article/pii/S0012821X04001426.

Deschamps, F., Cobden, L., Tackley, P.J., 2012. The primitive nature of large low shear-wave velocity provinces. Earth Planet. Sci. Lett. 349–350, 198–208. https://doi.org/10.1016/j.epsl.2012.07.012. ISSN 0012-821X. URL. http://www.sciencedirect.com/science/article/pii/S0012821X12003718.

Deschamps, F., Konishi, K., Fuji, N., Cobden, L., 2019. Radial thermo-chemical structure beneath western and northern pacific from seismic waveform inversion. Earth Planet. Sci. Lett. 520, 153–163. https://doi.org/10.1016/j.epsl.2019.05.040. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/S0012821X19303164.

Dorfman, S.M., Meng, Y., Prakapenka, V.B., Duffy, T.S., 2013. Effects of fe-enrichment on the equation of state and stability of (mg,fe)sio3 perovskite. Earth Planet. Sci. Lett. 361, 249–257. https://doi.org/10.1016/j.epsl.2012.10.033. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/S0012821X12006073.

Duffy, T.S., Ahrens, T.J., 1995. Compressional sound velocity, equation of state, and constitutive response of shock-compressed magnesium oxide. J. Geophys. Res. Solid Earth 100 (B1), 529–542. https://doi.org/10.1029/94JB02065.

Durek, J.J., Ekström, G., 1996. A radial model of anelasticity consistent with long-period surface-wave attenuation. Bull. Seismol. Soc. Am. 86 (1A), 144–158. https://doi.org/10.1785/BSSA08601A0144. ISSN 0037-1106.

Dziewonski, A.M., Anderson, D.L., 1981. Preliminary reference earth model. Phys. Earth Planet. Inter. 25 (4), 297–356. https://doi.org/10.1016/0031-9201(81)90046-7. ISSN 0031-9201. URL. http://www.sciencedirect.com/science/article/pii/0031920181900467.

Eggins, S., Rudnick, R., McDonough, W., 1998. The composition of peridotites and their minerals: a laser-ablation icp–ms study. Earth Planet. Sci. Lett. 154 (1), 53–71. https://doi.org/10.1016/S0012-821X(97)00195-7. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/S0012821X97001957.

Fei, Y., Zhang, L., Corgne, A., Watson, H., Ricolleau, A., Meng, Y., Prakapenka, V., 2007. Spin transition and equations of state of (mg, fe)o solid solutions. Geophys. Res. Lett. 34 (17). https://doi.org/10.1029/2007GL030712. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007GL030712.

Fichtner, A., Kennett, B.L.N., Igel, H., Bunge, H.-P., 2009. Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods. Geophys. J. Int. 179 (3), 1703–1725. https://doi.org/10.1111/j.1365-246X.2009.04368.x. ISSN 0956-540X.

French, S.W., Romanowicz, B.A., 2014. Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography. Geophys. J. Int. 199 (3), 1303–1327. https://doi.org/10.1093/gji/ggu334. ISSN 0956-540X.

Frost, D.A., Rost, S., 2014. The p-wave boundary of the large-low shear velocity province beneath the pacific. Earth Planet. Sci. Lett. 403, 380–392. https://doi.org/10.1016/j.epsl.2014.06.046. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/S0012821X14004373.

Garcia, R.F., Gagnepain-Beyneix, J., Chevrot, S., Lognonné, P., 2011. Very preliminary reference moon model. Phys. Earth Planet. Inter. 188 (1), 96–113. https://doi.org/10.1016/j.pepi.2011.06.015. ISSN 0031-9201. URL. https://www.sciencedirect.com/science/article/pii/S0031920111001373.

Garnero, E.J., Helmberger, D.V., 1998. Further structural constraints and uncertainties of a thin laterally varying ultralow-velocity layer at the base of the mantle. J. Geophys. Res. Solid Earth 103 (B6), 12495–12509. https://doi.org/10.1029/98JB00700. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98JB00700.

Garnero, E.J., McNamara, A.K., Shim, S.-H., 2016. Continent-sized anomalous zones with low seismic velocity at the base of earth's mantle. Nat. Geosci. 9, 481–489. https://doi.org/10.1038/ngeo2733.

Goes, S., Cammarano, F., Hansen, U., 2004. Synthetic seismic signature of thermal mantle plumes. Earth Planet. Sci. Lett. 218 (3), 403–419. https://doi.org/10.1016/S0012-821X(03)00680-0. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/S0012821X03006800.

Houser, C., Hernlund, J., Valencia-Cardona, J., Wentzcovitch, R., 2020. Discriminating lower mantle composition. Phys. Earth Planet. Inter. 308, 106552. https://doi.org/10.1016/j.pepi.2020.106552. ISSN 0031-9201. URL. http://www.sciencedirect.com/science/article/pii/S0031920118301845.

Hwang, Y.K., Ritsema, J., 2011. Radial Q_μ structure of the lower mantle from teleseismic body-wave spectra. Earth Planet. Sci. Lett. 303 (3), 369–375. https://doi.org/10.1016/j.epsl.2011.01.023. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/S0012821X11000471.

Hyung, E., Huang, S., Petaev, M.I., Jacobsen, S.B., 2016. Is the mantle chemically stratified? Insights from sound velocity modeling and isotope evolution of an early magma ocean. Earth Planet. Sci. Lett. 440, 158–168. https://doi.org/10.1016/j.epsl.2016.02.001. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/S0012821X16300127.

Irifune, T., 1994. Absence of an aluminous phase in the upper part of the earth's lower mantle. Nature 370, 131–133.

Irifune, T., Ringwood, A., 1987. Phase transformations in a harzburgite composition to 26 gpa: implications for dynamical behaviour of the subducting slab. Earth Planet. Sci. Lett. 86 (2), 365–376. https://doi.org/10.1016/0012-821X(87)90233-0. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/0012821X87902330.

Irifune, T., Shinmei, T., Mccammon, C., Miyajima, N., Rubie, D., Frost, D., 2010. Iron partitioning and density changes of pyrolite in earth's lower mantle. Science (New York, N.Y.) 327, 193–195. https://doi.org/10.1126/science.1181443.

Ishii, M., Tromp, J., 1999. Normal-mode and free-air gravity constraints on lateral variations in velocity and density of earth's mantle. Science 285 (5431), 1231–1236. https://doi.org/10.1126/science.285.5431.1231. ISSN 0036-8075. URL. https://science.sciencemag.org/content/285/5431/1231.

Jackson, I., 1998. Elasticity, composition and temperature of the Earth's lower mantle: a reappraisal. Geophys. J. Int. 134 (1), 291–311. https://doi.org/10.1046/j.1365-246x.1998.00560.x. ISSN 0956-540X.

Jeffreys, H., Bullen, K., 1940. Seismological Tables. British Assoc. Adv, Sci.

Karaoğlu, H., Romanowicz, B., 2018. Inferring global upper-mantle shear attenuation structure by waveform tomography using the spectral element method. Geophys. J. Int. 213 (3), 1536–1558. https://doi.org/10.1093/gji/ggy030. ISSN 0956-540X.

Käufl, P., Valentine, A., Wit, R., Trampert, J., 2016. Solving probabilistic inverse problems rapidly with prior samples. Geophys. J. Int. 205, ggw108. https://doi.org/10.1093/gji/ggw108.

Kennett, B.L.N., Engdahl, E.R., Buland, R., 1995. Constraints on seismic velocities in the earth from traveltimes. Geophys. J. Int. 122 (1), 108–124. https://doi.org/10.1111/j.1365-246X.1995.tb03540.x. URL. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-246X.1995.tb03540.x.

Khan, A., Connolly, J.A.D., Taylor, S.R., 2008. Inversion of seismic and geodetic data for the major element chemistry and temperature of the earth's mantle. J. Geophys. Res. Solid Earth 113 (B9). https://doi.org/10.1029/2007JB005239. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JB005239.

Khan, A., Boschi, L., Connolly, J.A.D., 2009. On mantle chemical and thermal heterogeneities and anisotropy as mapped by inversion of global surface wave data. J. Geophys. Res. Solid Earth 114 (B9). https://doi.org/10.1029/2009JB006399. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JB006399.

Khan, A., Boschi, L., Connolly, J.A.D., 2011. Mapping the earth's thermochemical and anisotropic structure using global surface wave data. J. Geophys. Res. Solid Earth 116 (B1). https://doi.org/10.1029/2010JB007828. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JB007828.

Khan, A., Ceylan, S., Driel, M., Giardini, D., Lognonné, P., Samuel, H., Schmerr, N., Stähler, S., Duran, A., Huang, Q., Kim, D., Broquet, A., Charalambous, C., Clinton, J., Davis, P., Drilleau, M., Karakostas, F., Lekic, V., McLennan, S., Banerdt, W., 2021. Upper mantle structure of mars from insight seismic data. Science 373, 434–438. https://doi.org/10.1126/science.abf2966.

Koelemeijer, P., Deuss, A., Ritsema, J., 2017. Density structure of earth's lowermost mantle from stoneley mode splitting observations. Nat. Commun. 8, 15241. https://doi.org/10.1038/ncomms15241.

Konishi, K., Fuji, N., Deschamps, F., 2020. Three-dimensional elastic and anelastic structure of the lowermost mantle beneath the western pacific from finite-frequency tomography. J. Geophys. Res. Solid Earth 125 (2), e2019JB018089. https://doi.org/10.1029/2019JB018089. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JB018089.

Lawrence, J.F., Wysession, M.E., 2006. Qlm9: a new radial quality factor (qμ) model for the lower mantle. Earth Planet. Sci. Lett. 241 (3), 962–971. https://doi.org/10.1016/j.epsl.2005.10.030. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/S0012821X05007090.

Lay, T., Helmberger, D.V., 1983. A lower mantle S-wave triplication and the shear velocity structure of D″. Geophys. J. Int. 75 (3), 799–837. https://doi.org/10.1111/j.1365-246X.1983.tb05010.x. ISSN 0956-540X.

Lei, W., Ruan, Y., Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N., Pugmire, D., 2020. Global adjoint tomography—model GLAD-M25. Geophys. J. Int. 223 (1), 1–21. https://doi.org/10.1093/gji/ggaa253. ISSN 0956-540X.

Li, B., Zhang, J., 2005. Pressure and temperature dependence of elastic wave velocity of mgsio3 perovskite and the composition of the lower mantle. Phys. Earth Planet. Inter. 151 (1), 143–154. https://doi.org/10.1016/j.pepi.2005.02.004. ISSN 0031-9201. URL. https://www.sciencedirect.com/science/article/pii/S0031920105000634.

Mao, W.L., Meng, Y., Shen, G., Prakapenka, V.B., Campbell, A.J., Heinz, D.L., Shu, J., Caracas, R., Cohen, R.E., Fei, Y., Hemley, R.J., Mao, H.-k., 2005. Iron-rich silicates in the earth's d″ layer. Proc. Natl. Acad. Sci. 102 (28), 9751–9753. https://doi.org/10.1073/pnas.0503737102. ISSN 0027-8424. URL. https://www.pnas.org/content/102/28/9751.

Marquardt, H., Thomson, A.R., 2020. Experimental elasticity of earth's deep mantle. Nat. Rev. Earth Environ. 1, 455–469. https://doi.org/10.1038/s43017-020-0077-3.

Marquardt, H., Speziale, S., Reichmann, H.J., Frost, D.J., Schilling, F.R., Garnero, E.J., 2009. Elastic shear anisotropy of ferropericlase in earth's lower mantle. Science 324 (5924), 224–226.

Marquardt, H., Buchen, J., Mendez, A.S.J., Kurnosov, A., Wendt, M., Rothkirch, A., Pennicard, D., Liermann, H.-P., 2018. Elastic softening of $(Mg_{0.8}Fe_{0.2})O$ ferropericlase across the iron spin crossover measured at seismic frequencies. Geophys. Res. Lett. 45 (14), 6862–6868. https://doi.org/10.1029/2018GL077982. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL077982.

Matas, J., Bass, J., Ricard, Y., Mattern, E., Bukowinski, M.S.T., 2007. On the bulk composition of the lower mantle: predictions and limitations from generalized inversion of radial seismic profiles. Geophys. J. Int. 170 (2), 764–780. https://doi.org/10.1111/j.1365-246X.2007.03454.x. ISSN 0956-540X.

Mattern, E., Matas, J., Ricard, Y., Bass, J., 2005. Lower mantle composition and temperature from mineral physics and thermodynamic modelling. Geophys. J. Int. 160 (3), 973–990. https://doi.org/10.1111/j.1365-246X.2004.02549.x. ISSN 0956-540X.

Mosca, I., Cobden, L., Deuss, A., Ritsema, J., Trampert, J., 2012. Seismic and mineralogical structures of the lower mantle from probabilistic tomography. J. Geophys. Res. Solid Earth 117 (B6). https://doi.org/10.1029/2011JB008851. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JB008851.

Moulik, P., Ekström, G., 2016. The relationships between large-scale variations in shear velocity, density, and compressional velocity in the earth's mantle. J. Geophys. Res. Solid Earth 121 (4), 2737–2771. https://doi.org/10.1002/2015JB012679. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JB012679.

Muir, J.M., Brodholt, J.P., 2015. Elastic properties of ferrous bearing MgSiO3 and their relevance to ULVZs. Geophys. J. Int. 201 (1), 496–504. https://doi.org/10.1093/gji/ggv045. ISSN 0956-540X.

Murakami, M., Hirose, K., Kawamura, K., Sata, N., Ohishi, Y., 2004. Post-perovskite phase transition in mgsio3. Science (New York, N.Y.) 304, 855–858. https://doi.org/10.1126/science.1095932.

Murakami, M., Ohishi, Y., Hirao, N., Hirose, K., 2012. A perovskitic lower mantle inferred from high-pressure, high-temperature sound velocity data. Nature 485, 90–94. https://doi.org/10.1038/nature11004.

Myhill, R., Cottaar, S., Heister, T., Rose, I., Unterborn, C., Dannberg, J., Gassmoeller, R., 2023. Burnman – a python toolkit for planetary geophysics, geochemistry and thermodynamics. J. Open Source Softw. 8 (87), 5389. https://doi.org/10.21105/joss.05389.

Oganov, A., Dorogokupets, P.I., 2003. All-electron and pseudopotential study of MgO: equation of state, anharmonicity, and stability. Phys. Rev. B 67, 224110. https://doi.org/10.1103/PhysRevB.67.224110.

Perrillat, J.-P., Ricolleau, A., Daniel, I., Fiquet, G., Mezouar, M., Guignot, N., Cardon, H., 2006. Phase transformations of subducted basaltic crust in the upmost lower mantle. Phys. Earth Planet. Inter. 157 (1), 139–149. https://doi.org/10.1016/j.pepi.2006.04.001. ISSN 0031-9201. URL. https://www.sciencedirect.com/science/article/pii/S0031920106001063.

Ricolleau, A., Fei, Y., Cottrell, E., Watson, H., Deng, L., Zhang, L., Fiquet, G., Auzende, A.-L., Roskosz, M., Morard, G., Prakapenka, V., 2009. Density profile of pyrolite under the lower mantle conditions. Geophys. Res. Lett. 36 (6). https://doi.org/10.1029/2008GL036759. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008GL036759.

Riel, N., Kaus, B.J.P., Green, E.C.R., Berlie, N., 2022. Magemin, an efficient gibbs energy minimizer: application to igneous systems. Geochem. Geophys. Geosyst. 23 (7), e2022GC010427. https://doi.org/10.1029/2022GC010427. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022GC010427.

Rijal, A., Cobden, L., Trampert, J., Jackson, J.M., Valentine, A., 2021. Inferring material properties of the lower mantle minerals using mixture density networks. Phys. Earth Planet. Inter. 319, 106784. https://doi.org/10.1016/j.pepi.2021.106784. ISSN 0031-9201. URL. https://www.sciencedirect.com/science/article/pii/S0031920121001424.

Rijal, A., Cobden, L., Trampert, J., Marquardt, H., Jackson, J.M., 2023. Shear properties of mgo inferred using neural networks. Eur. J. Mineral. 35 (1), 45–58. https://doi.org/10.5194/ejm-35-45-2023. URL. https://ejm.copernicus.org/articles/35/45/2023/.

Ringwood, A.E., 1962. A model for the upper mantle. J. Geophys. Res. (1896-1977) 67 (2), 857–867. https://doi.org/10.1029/JZ067i002p00857. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JZ067i002p00857.

Ritsema, J., Xu, W., Stixrude, L., Lithgow-Bertelloni, C., 2009. Estimates of the transition zone temperature in a mechanically mixed upper mantle. Earth Planet. Sci. Lett. 277 (1), 244–252. https://doi.org/10.1016/j.epsl.2008.10.024. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/S0012821X08006845.

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. Geophys. J. Int. 184 (3), 1223–1236. https://doi.org/10.1111/j.1365-246X.2010.04884.x. URL. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-246X.2010.04884.x.

Romanowicz, B., 2008. Using seismic waves to image earth's internal structure. Nature 451, 266–268. https://doi.org/10.1038/nature06583.

Schuberth, B.S.A., Zaroli, C., Nolet, G., 2012. Synthetic seismograms for a synthetic earth: long-period P- and S-wave traveltime variations can be explained by temperature alone. Geophys. J. Int. 188 (3), 1393–1412. https://doi.org/10.1111/j.1365-246X.2011.05333.x. ISSN 0956-540X.

Simmons, N.A., Forte, A.M., Grand, S.P., 2009. Joint seismic, geodynamic and mineral physical constraints on three-dimensional mantle heterogeneity: implications for the relative importance of thermal versus compositional heterogeneity. Geophys. J. Int. 177 (3), 1284–1304. https://doi.org/10.1111/j.1365-246X.2009.04133.x. URL. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-246X.2009.04133.x.

Simmons, N.A., Forte, A.M., Boschi, L., Grand, S.P., 2010. Gypsum: a joint tomographic model of mantle density and seismic wave speeds. J. Geophys. Res. Solid Earth 115 (B12). https://doi.org/10.1029/2010JB007631. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JB007631.

Sinogeikin, S.V., Bass, J.D., 2000. Single-crystal elasticity of pyrope and MgO to 20 GPa by brillouin scattering in the diamond cell. Phys. Earth Planet. Inter. 120 (1), 43–62. https://doi.org/10.1016/S0031-9201(00)00143-6. ISSN 0031-9201.

Speziale, S., Zha, C.-S., Duffy, T.S., Hemley, R.J., Mao, H.-K., 2001. Quasi-hydrostatic compression of magnesium oxide to 52 GPa: implications for the pressure-volume-

temperature equation of state. J. Geophys. Res. Solid Earth 106 (B1), 515–528. https://doi.org/10.1029/2000JB900318.

Stähler, S.C., Khan, A., Banerdt, W.B., Lognonné, P., Giardini, D., Ceylan, S., Drilleau, M., Duran, A.C., Garcia, R.F., Huang, Q., Kim, D., Lekic, V., Samuel, H., Schimmel, M., Schmerr, N., Sollberger, D., Éléonore Stutzmann, Z. Xu, Antonangeli, D., Charalambous, C., Davis, P.M., Irving, J.C.E., Kawamura, T., Knapmeyer, M., Maguire, R., Marusiak, A.G., Panning, M.P., Perrin, C., Plesa, A.-C., Rivoldini, A., Schmelzbach, C., Zenhäusern, G., Éric Beucler, J., Clinton, N., Dahmen, M., Driel, T., Gudkova, A., Horleston, W.T., Pike, M. Plasman, Smrekar, S.E., 2021. Seismic detection of the martian core. Science 373 (6553), 443–448. https://doi.org/10.1126/science.abi7730. URL. https://www.science.org/doi/abs/10.1126/science.abi7730.

Stixrude, L., Lithgow-Bertelloni, C., 2005. Thermodynamics of mantle minerals – I. Physical properties. Geophys. J. Int. 162 (2), 610–632. https://doi.org/10.1111/j.1365-246X.2005.02642.x.

Stixrude, L., Lithgow-Bertelloni, C., 2011. Thermodynamics of mantle minerals- II. Phase equilibria. Geophys. J. Int. 184 (3), 1180–1213. https://doi.org/10.1111/j.1365-246X.2010.04890.x. URL. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-246X.2010.04890.x.

Stixrude, L., Hemley, R., Fei, Y., Mao, H., 1992. Thermoelasticity of silicate perovskite and magnesiowüstite and stratification of the earth's mantle. Science 257 (5073), 1099–1101.

Stixrude, L., de Koker, N., Sun, N., Mookherjee, M., Karki, B.B., 2009. Thermodynamics of silicate liquids in the deep earth. Earth Planet. Sci. Lett. 278 (3), 226–232. https://doi.org/10.1016/j.epsl.2008.12.006. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/S0012821X08007565.

Sturhahn, W., Jackson, J.M., Lin, J.-F., 2005. The spin state of iron in minerals of earth's lower mantle. Geophys. Res. Lett. 32 (12). https://doi.org/10.1029/2005GL022802. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005GL022802.

Sun, D., Helmberger, D.V., Jackson, J.M., Clayton, R.W., Bower, D.J., 2013. Rolling hills on the core–mantle boundary. Earth Planet. Sci. Lett. 361, 333–342. https://doi.org/10.1016/j.epsl.2012.10.027. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/S0012821X12005961.

Talavera-Soza, S., Cobden, L.J., Faul, U., Deuss, A.F., 2025. Global 3D model of mantle attenuation using normal modes. Nature. https://assets-eu.researchsquare.com/files/rs-1580818/v1_covered_eafee354-7fd4-4f73-bf90-df1196732f33.pdf?c=1737619627. Submitted.

Tauzin, B., Debayle, E., Wittlinger, G., 2010. Seismic evidence for a global low velocity layer in the earth's upper mantle. Nat. Geosci. 3, 09. https://doi.org/10.1038/ngeo969.

Thomas, C., Kendall, J.-M., Helffrich, G., 2009. Probing two low-velocity regions with PKP b-caustic amplitudes and scattering. Geophys. J. Int. 178 (1), 503–512. https://doi.org/10.1111/j.1365-246X.2009.04189.x. ISSN 0956-540X.

Thomson, A., Crichton, W., Brodholt, J., Wood, I., Siersch, N., Muir, J., Dobson, D., Hunt, S., 2019. Seismic velocities of casio3 perovskite can explain llsvps in earth's lower mantle. Nature 572, 1–5. https://doi.org/10.1038/s41586-019-1483-x.

Trampert, J., Deschamps, F., Resovsky, J., Yuen, D., 2004. Probabilistic tomography maps chemical heterogeneities throughout the lower mantle. Science 306 (5697), 853–856. https://doi.org/10.1126/science.1101996. URL. https://science.sciencemag.org/content/306/5697/853.

Tsuchiya, T., Kawai, K., 2013. Ab Initio Mineralogical Model of the Earth's Lower Mantle, Chapter 7. John Wiley & Sons, Ltd, pp. 213–243. https://doi.org/10.1002/9781118529492.ch7. ISBN 9781118529492.. URL. https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118529492.ch7.

Tsuchiya, T., Tsuchiya, J., Umemoto, K., Wentzcovitch, R.M., 2004. Phase transition in mgsio3 perovskite in the earth's lower mantle. Earth Planet. Sci. Lett. 224 (3), 241–248. https://doi.org/10.1016/j.epsl.2004.05.017. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/S0012821X04003383.

van der Hilst, R.D., Widiyantoro, S., Engdahl, E.R., 1997. Evidence for deep mantle circulation from global tomography. Nature 386, 578–584. https://doi.org/10.1038/386578a0.

Vilella, K., Bodin, T., Boukaré, C.-E., Deschamps, F., Badro, J., Ballmer, M.D., Li, Y., 2021. Constraints on the composition and temperature of llsvps from seismic properties of lower mantle minerals. Earth Planet. Sci. Lett. 554, 116685. https://doi.org/10.1016/j.epsl.2020.116685. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/S0012821X20306294.

Wang, X., Tsuchiya, T., Hase, A., 2015. Computational support for a pyrolitic lower mantle containing ferric iron. Nat. Geosci. 8, 556–559. https://doi.org/10.1038/ngeo2458.

Waszek, L., Thomas, C., Deuss, A., 2015. Pkp precursors: implications for global scatterers. Geophys. Res. Lett. 42 (10), 3829–3838. https://doi.org/10.1002/2015GL063869. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015GL063869.

Wentzcovitch, R.M., Karki, B.B., Cococcioni, M., de Gironcoli, S., Jan 2004. Thermoelastic properties of $mgsio_3$-perovskite: insights on the nature of the earth's lower mantle. Phys. Rev. Lett. 92, 018501. https://doi.org/10.1103/PhysRevLett.92.018501. URL. https://link.aps.org/doi/10.1103/PhysRevLett.92.018501.

Wentzcovitch, R.M., Justo, J.F., Wu, Z., da Silva, C.R.S., Yuen, D.A., Kohlstedt, D., 2009. Anomalous compressibility of ferropericlase throughout the iron spin cross-over. Proc. Natl. Acad. Sci. 106 (21), 8447–8452. https://doi.org/10.1073/pnas.0812150106. URL. https://www.pnas.org/doi/abs/10.1073/pnas.0812150106.

Wolf, A.S., Jackson, J.M., Dera, P., Prakapenka, V.B., 2015. The thermal equation of state of (mg, fe)sio3 bridgmanite (perovskite) and implications for lower mantle structures. J. Geophys. Res. Solid Earth 120 (11), 7460–7489. https://doi.org/10.1002/2015JB012108. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JB012108.

Wu, Z., 2016. Velocity structure and composition of the lower mantle with spin crossover in ferropericlase. J. Geophys. Res. Solid Earth 121 (4), 2304–2314. https://doi.org/10.1002/2015JB012667. URL. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JB012667.

Xu, W., Lithgow-Bertelloni, C., Stixrude, L., Ritsema, J., 2008. The effect of bulk composition and temperature on mantle seismic structure. Earth Planet. Sci. Lett. 275 (1), 70–79. https://doi.org/10.1016/j.epsl.2008.08.012. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/S0012821X08005281.

Zhang, Z., Stixrude, L., Brodholt, J., 2013. Elastic properties of mgsio3-perovskite under lower mantle conditions and the composition of the deep earth. Earth Planet. Sci. Lett. 379, 1–12. https://doi.org/10.1016/j.epsl.2013.07.034. ISSN 0012-821X. URL. https://www.sciencedirect.com/science/article/pii/S0012821X13004093.

Zhu, H., Bozdag, E., Duffy, T., Tromp, J., 2013. Seismic attenuation beneath europe and the North Atlantic: implications for water in the mantle. Earth Planet. Sci. Lett. 381, 1–11. https://doi.org/10.1016/j.epsl.2013.08.030.