Key Points:

- Bayesian inversion used to calibrate anelasticity parameterisations, allowing conversion of seismic velocity into thermomechanical structure
- Probabilistic approach enables uncertainty quantification of thermomechanical outputs
- Evidence for significant lateral heterogeneity in Antarctic viscosity, lithosphere-asthenosphere boundary depth, and geothermal heat flow, corroborated by data from the geological record

Author Contributions:
Conceptualization: James A. N. Hazzard, Fred D. Richards
Data curation: James A. N. Hazzard, Fred D. Richards
Formal analysis: James A. N. Hazzard, Fred D. Richards

Probabilistic Assessment of Antarctic Thermomechanical Structure: Impacts on Ice Sheet Stability

James A. N. Hazzard[1] ⓘ, Fred D. Richards[1] ⓘ, Saskia D. B. Goes[1] ⓘ, and Gareth G. Roberts[1] ⓘ

[1]Department of Earth Science and Engineering, Royal School of Mines, Imperial College London, London, UK

Abstract Uncertainty in present-day glacial isostatic adjustment (GIA) rates represents at least 44% of the total gravity-based ice mass balance signal over Antarctica. Meanwhile, physical couplings between solid Earth, sea level and ice dynamics enhance the dependency of the spatiotemporally varying GIA signal on three-dimensional variations in mantle rheology. Improved knowledge of thermomechanical mantle structure is therefore required to refine estimates of current and projected ice mass balance. Here, we present a Bayesian inverse method for self-consistently mapping shear-wave velocities from high-resolution adjoint tomography into thermomechanical structure using calibrated parameterisations of anelasticity at seismic frequency. We constrain the model using regional geophysical data sets containing information on upper mantle temperature, attenuation and viscosity structure. Our treatment allows formal quantification of parameter covariances, and naturally permits propagation of material parameter uncertainties into thermomechanical structure estimates. We find that uncertainty in steady-state viscosity structure at 150 km depth can be reduced by 4–5 orders of magnitude compared with a forward-modeling approach neglecting covariance between viscoelastic parameters. By accounting for the dependence of apparent viscosity on loading timescale, we find good agreement between our estimates of mantle viscosity beneath West Antarctica, and those derived from satellite GPS. Direct access to temperature structure allows us to estimate lateral variations in lithosphere-asthenosphere boundary (LAB) depth, geothermal heat flow (GHF), and associated uncertainties. We find evidence for shallow LAB depths (63 ± 13 km), and high GHF (76 ± 7 mW m^{-2}) beneath West Antarctica that, combined with low asthenospheric viscosities, indicate a highly dynamic response to ice mass loss.

Plain Language Summary The viscosity (i.e., "runniness") and temperature of Earth's interior exert a major influence on ice sheet stability and sea level change. Viscosity controls how the shape of Earth's surface and gravity field distorts when ice melts. Temperature controls the flow of heat to the base of ice sheets, determining how rapidly they slide and deform. Both parameters are expected to vary significantly with position inside Earth's mantle, but are poorly constrained. Improved information about mantle structure can be derived from recent models telling us about spatial variations in the speed at which earthquake-generated waves travel through Earth. In this study, we present a statistical method allowing us to convert from such models into estimates of viscosity and temperature. This method enables us to reduce uncertainty on such estimates, by feeding in regional geophysical data to help refine the range of plausible structures. Our estimates of viscosity beneath West Antarctica are in close agreement with observations from satellite GPS. In addition, our models of temperature structure allow us to estimate variations in Antarctic tectonic plate thickness, geothermal heat flow, and their associated uncertainties. We find evidence for significant disparity in each of these structures between West and East Antarctica.

1. Introduction

Antarctica is host to a volume of ice equivalent to 57.9 ± 0.9 m of global mean sea level (GMSL) rise, or roughly 90% of the global cryosphere (Bamber et al., 2018; Morlighem et al., 2020). The mantle structure, topography, and glacial stability of this continent expresses a dichotomy in tectonic setting between East and West Antarctica. The two regions are separated by the Transantarctic Mountain Range (TAM), which spans the continental interior from the Weddell Sea (WS) to the Ross Sea (RS). As a result, Antarctica's grounded ice volume is divided into an East and West Antarctic Ice Sheet (EAIS and WAIS, respectively). The EAIS is underlain by thick, cratonic lithosphere owing to minimal tectonic activity in this region since the Mesozoic Era (Noble et al., 2020). The WAIS is underlain by an active rift system, which has given rise to upwelling of low viscosity asthenosphere, and dynamically thinned lithosphere (Noble et al., 2020). Bedrock elevation is predominantly above GMSL in

the east, and below GMSL in the west (Figure 1a). This exerts a major influence on ice dynamics, due to the increased vulnerability of marine-grounded ice, especially when positioned on a reverse bed slope, as is the case in West Antarctica (Fretwell et al., 2013). The WAIS is therefore considered much more prone to short-term ice mass loss (Coulon et al., 2021). Indeed, it is declining by ∼200 Gt per year, while it is unknown whether the EAIS is gaining or losing mass (Shepherd et al., 2018).

To predict the contribution of the Antarctic Ice Sheet (AIS) to future patterns of spatiotemporally variable sea level, we require a reliable assessment of its stability. This involves detailed insight on past ice volumes from the geological record (DeConto & Pollard, 2016), quantification of present-day ice mass balance (Caron et al., 2018), and physically accurate models for the future evolution of the cryosphere (Slangen et al., 2017). A vital consideration in each of these pursuits is solid Earth structure and dynamics. Time-dependent lithospheric deflections caused by evolving surface loads and mantle flow alter the elevation of palaeo sea level indicators, the shape of the oceans and gravity field, and the stability of grounded ice (Austermann et al., 2015; Gomez et al., 2018; Mitrovica et al., 2020). These coupled interactions between solid Earth, ocean and cryosphere operate on physical timescales ranging from decadal to geological, and depend heavily on upper mantle thermomechanical structure.

For example, contemporary estimates of ice mass balance typically rely on satellite missions recording either altimetric or gravimetric data (Shepherd et al., 2018, 2020; Zwally & Giovinetto, 2011). The Gravity Recovery and Climate Experiment (GRACE) and its successor GRACE Follow-On (GRACE-FO) offer indirect regional scale insight into ice mass balance via the tracking of temporal changes in Earth's gravity field (King et al., 2012). However, glacial isostatic adjustment (GIA; the viscoelastic response of the solid Earth to changes in the distribution of ice and water over its surface) influences the gravity field significantly, even on decadal timescales. Since the Earth is still responding today to deglaciation following the Last Glacial Maximum (LGM; 21 ka), with elastically adjusted GPS uplift rates ranging from −5 to 5 mm a^{-1} across Antarctica (Thomas et al., 2011; Figure 1a), this contaminating GIA signal must be removed to accurately determine contemporary ice mass balance. However, calculation of the GIA signal relies upon two main inputs that remain weakly constrained; the first being a reconstruction of ice sheet history, and the second a viscoelastic Earth model. Caron et al. (2018) estimated that the Antarctic GIA uncertainty was around 44% of the total amplitude of the GRACE gravity signal itself. The true uncertainty is likely to be even larger, since the aforementioned study does not account for significant lateral heterogeneities in upper mantle viscosity inferred from GPS observations across Antarctica (Barletta et al., 2018).

The quality of projections of future sea level change is also heavily reliant on our ability to model GIA as accurately and precisely as possible. GIA models that incorporate solid Earth feedbacks will be particularly sensitive to the underlying mantle rheology, since this acts as a direct control on the evolution of the ice sheet (Whitehouse, 2018). For example, the presence of low-viscosity mantle beneath melting marine-based ice sheet sectors such as the Amundsen Sea Embayment (ASE) may delay or even prevent unstable grounding line retreat (Barletta & Bordoni, 2013). This enhanced dependence of sea level projections on our knowledge about mantle viscosity and temperature structure points to the requirement for coupled ice sheet-sea level modeling, incorporating reliable estimates of three-dimensional mantle structure (Gomez et al., 2018).

Seismic shear-wave velocity (V_S) can be used to gain insight into upper mantle structure beneath the ice sheets due to its strong sensitivity to temperature (Faul & Jackson, 2005). Laboratory experiments show sub-solidus temperature changes can induce up to 20% variations in V_S (Priestley & McKenzie, 2013; and references therein). Although volatiles and composition may also influence V_S (Karato & Jung, 1998; Lee, 2003), recent studies show close agreement between xenolith-derived temperature profiles and those inferred from seismic tomography models using anelasticity parameterisations that ignore the potential impact of compositional heterogeneity in the asthenosphere and lower lithosphere, indicating that temperature is indeed the dominant control on shallow mantle V_S variation (Hoggard et al., 2020; Klöcking et al., 2020).

Until recently, Antarctica has suffered from a significant shortage of seismic data due in part to difficulties operating polar seismic stations and the lack of proximal (latitudinally) land masses (Lloyd et al., 2020). However, ANT-20, a wave-equation traveltime adjoint tomography model, has recently been developed utilizing data from 323 seismic stations, the majority (297) of which reside on the Antarctic continent (Lloyd et al., 2020). ANT-20 is the first continental model to image Antarctica at regional-scale resolution (∼100 km), and thus serves as a suitable starting point for mapping temperature and viscosity with unprecedented fidelity. Promisingly, this tomographic model contains many features that are consistent with independent constraints. For example, lateral
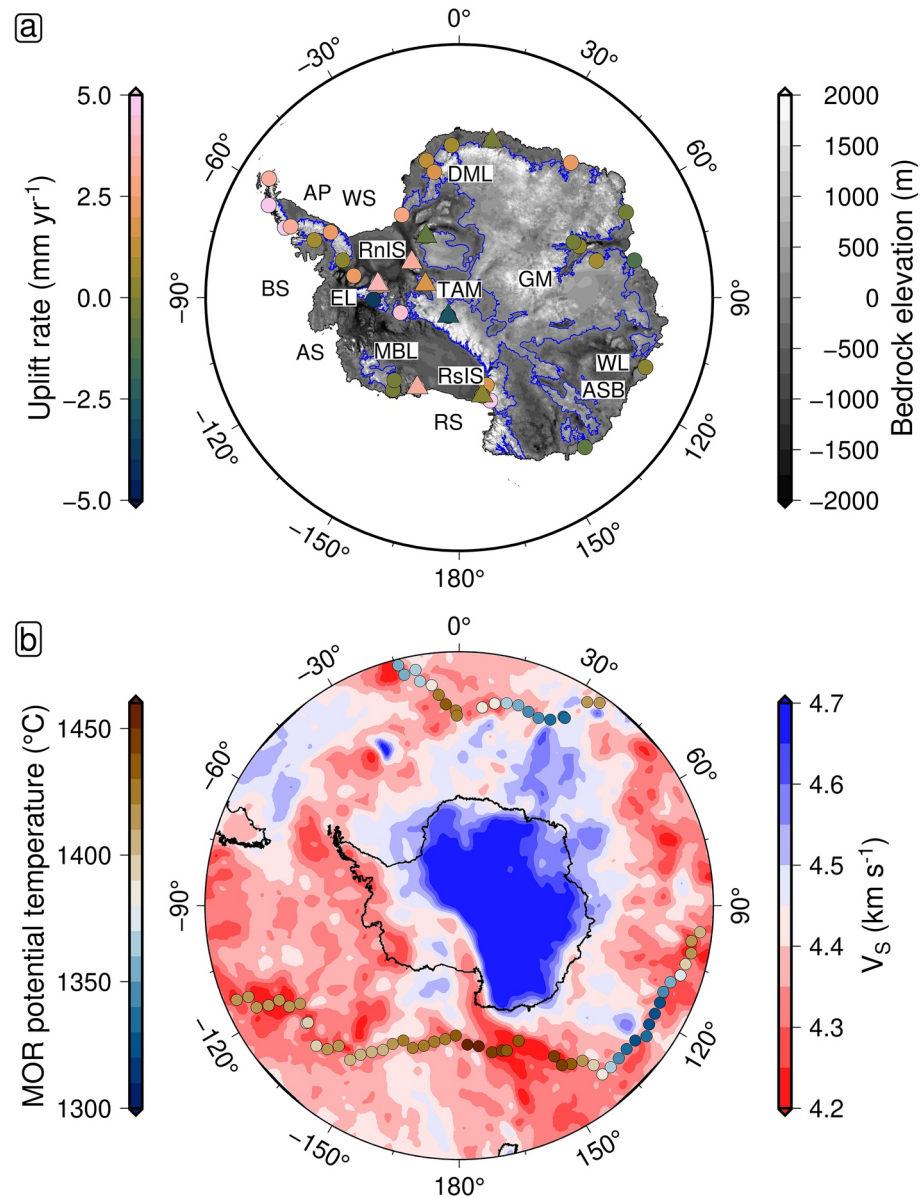
Figure 1. Geophysical and geochemical constraints on Antarctic mantle dynamics and structure. (a) Antarctic bedrock elevation taken from BEDMAP2 data (Fretwell et al., 2013), with elastically corrected GPS uplift rate overlain (Thomas et al., 2011, circles: individual measurement sites; triangles: averages over local sites). Blue contours delineate the transition between positive and negative bedrock elevation. Regions of negative elevation around the periphery of the continent indicate where the AIS is marine-grounded. Text labels indicate reference points within Antarctica and the surrounding ocean (AP, Antarctic Peninsula; RnIS, Ronne Ice Shelf; WS, Weddell Sea; DML, Dronning Maud Land; GM, Gamburtsev Mountain Range; WL, Wilkes Land; ASB, Aurora Subglacial Basin; RS, Ross Sea; RsIS, Ross Ice Shelf; MBL, Marie Byrd Land; TAM, Transantarctic Mountain Range; AS, Amundsen Sea; EL, Ellsworth Land; BS, Bellingshausen Sea). (b) V_S at 150 km depth from ANT-20 tomographic model (Lloyd et al., 2020), with mid-ocean ridge potential temperature overlain (Dalton et al., 2014).

variations in V_S beneath the Antarctic mid-ocean ridge system correlate well with point estimates of potential temperature (Dalton et al., 2014; Figure 1b). Fast shear-wave velocities below East Antarctica are indicative of high viscosity lithosphere and slow velocities in the West point to low viscosities and thus short viscoelastic response timescales (Coulon et al., 2021).

Here we present a novel Bayesian inverse framework for self-consistent quantification of upper mantle thermo-mechanical structure from seismic data via the calibration of experimental parameterisations of anelasticity. To

achieve this, the deterministic approach set out by Richards, Hoggard, Crosby, et al. (2020) is recast in terms of statistical methods. First, the details of the modeling approach are discussed, as well as the geophysical data used to constrain the inversion, and its algorithmic implementation. Second, the inversion is applied to the ANT-20 model of Antarctic shear-wave velocity structure and trade-offs between viscoelastic parameters are quantified. Third, seismologically derived estimates of viscosity, temperature, lithosphere-asthenosphere boundary (LAB) depth and geothermal heat flow (GHF) are presented. For the first time, uncertainties in each of these outputs are constrained using stochastic methods. Finally, the physical outputs presented herein are evaluated in the context of other studies, and potential implications and remaining challenges are discussed. Our principal goal is to show how disparate geophysical constraints can be integrated within a probabilistic inverse framework to develop a quantitative understanding of Antarctic upper mantle thermomechanical structure and its associated uncertainties.

2. Converting Seismic Velocities Into Thermomechanical Parameters

In order to use V_S data to gain insight into upper mantle structure, a physical model must be employed. Most studies have taken an empirical approach to converting between V_S and viscosity, making use of a constant or depth-dependent scaling (e.g., Austermann et al., 2013; Davies et al., 2019; Milne et al., 2018; Steinberger et al., 2019). Such conversions fail to appropriately capture non-linear viscosity reductions observed near the solidus in laboratory experiments (Faul & Jackson, 2007; McCarthy & Takei, 2011; Takei, 2017; Yabe & Hiraga, 2020).

To better account for observed non-linearities, we use the anelasticity parameterization of Yamauchi and Takei (2016), hereafter YT16. YT16 incorporates the effect of pre-melting, a process which enhances diffusionally accommodated grain boundary sliding and high-frequency seismic attenuation. By doing so, the model provides a physical basis for significant non-linearity in the $V_S(T)$ relationship near and beyond the solidus temperature (homologous temperature, $T/T_S \sim 1$). Since YT16 made use of a polycrystal analogue of the olivine-basalt system with a much smaller melting temperature, their forced oscillation experiments conducted near room temperature can tap into the same normalized frequency range as seismic waves at the near-solidus conditions relevant to the upper mantle.

YT16 make use of the complex compliance $J^*(\omega) = J_1(\omega) + iJ_2(\omega)$ in their description of anelasticity, which relates the complex strain response $\varepsilon^*(t)$ of a linear viscoelastic material to an applied complex stress $\sigma^*(t)$.

$$\sigma^*(t) = \int \sigma_0 \exp(-i\omega t)\, d\omega, \tag{1}$$

$$\varepsilon^*(t) = \int J^*(\omega)\sigma_0 \exp(-i\omega t)\, d\omega. \tag{2}$$

The in-phase term of the complex compliance J_1 is known as the storage compliance, as it is energy conserving. The out-of-phase term J_2 is known as the loss compliance, as it is responsible for dissipation. This relationship can be verified by considering the work done by the system during a complete oscillation cycle (Appendix A).

Consider a plane wave propagating in a linear viscoelastic medium, as an approximation of seismic wave propagation in the mantle. The dependence of the phase velocity $V(\omega)$ and attenuation $Q^{-1}(\omega)$ on the complex compliance is given by

$$V = \frac{1}{\sqrt{\rho J_1}}, \tag{3}$$

$$Q^{-1} = \frac{J_2}{J_1}, \tag{4}$$

where ρ is the density of the medium (McCarthy & Takei, 2011). These equations hold for seismological studies of the mantle, where the approximation $Q^{-1} \ll 1$ is valid (Takei, 2017). The complex compliance terms can be determined given knowledge of temperature, pressure, seismic frequency, and a set of viscoelastic parameters. This allows conversion from temperature to V_S, as well as vice versa given some optimisation procedure.

Of the various parameters involved in YT16, seven are material properties that depend upon the makeup of the mantle, and control its viscoelastic behavior. Hereafter, this group of seven viscoelastic parameters are referred to as the model space. The first three parameters regulate the elastic component of the material response, namely the unrelaxed shear modulus at reference temperature and pressure, μ_0, and the temperature and pressure derivatives of the shear modulus, $\partial\mu/\partial T$ and $\partial\mu/\partial P$, respectively. The last four parameters control the transient component of the response, namely reference viscosity, η_0, activation energy, E_A, activation volume, V_A and the depth gradient of the solidus temperature, $\partial T_S/\partial z$.

A forward-modeling approach is commonly used to account for anelasticity in the conversion of V_S into thermo-mechanical parameters (Cammarano et al., 2003). The elastic component of $V_S(P, T)$ is modeled by combining an assumed mantle composition with a computational Gibbs free energy minimization to estimate μ_0, $\partial\mu/\partial T$ and $\partial\mu/\partial P$. A correction for anelastic behavior is then applied using values of η_0, E_A, V_A, and $\partial T_S/\partial z$ compiled from laboratory-based experiments on mantle minerals. There are two key drawbacks to this method. The first is that applying experimentally determined parameter values to mantle conditions requires extrapolation of grain size-dependent behavior across several orders of magnitude, the validity of which remains unclear. The second is that discrepancies between tomography models are introduced by subjective choices such as regularization, model parameterization, and choice of reference model (Richards, Hoggard, White, et al., 2020). The forward approach then becomes problematic as, for a constant choice of viscoelastic parameters, highly discrepant physical predictions are generated depending on the chosen velocity model.

To tackle these issues and ensure a conversion consistent with Antarctic geophysical data, we instead calibrate the seven-dimensional YT16 model space against a suite of regional temperature, attenuation and viscosity constraints (Priestley & McKenzie, 2006, 2013; Richards, Hoggard, White, et al., 2020). A regional calibration is preferred to using viscoelastic parameters obtained from a global study, since the former approach ensures consistency with the chosen Antarctica-specific tomographic model (see Section 4.1 of Austermann et al. (2021) for further detail on intermodel seismic velocity variation). Calibration is achieved within the framework of a Bayesian inversion, incorporating stochastic sampling to characterize the model space. These samples can then be used to propagate uncertainties in the viscoelastic parameters into formal uncertainties in the resulting rheological model. Two additional sources of uncertainty are not accounted for during the propagation. The first is tomographic uncertainty, which we ignore in converting velocities into thermomechanical parameter estimates. This uncertainty is partially mitigated by the regional calibration procedure. The second is a phenomenological source of uncertainty, deriving from the assumption that YT16 is an accurate representation of upper mantle rheology. Although investigating this assumption further is beyond the scope of this study, we note that our inverse modeling framework is designed to be equally applicable to other choices of anelasticity parameterization (Havlin et al., 2021). Readers interested in the extent to which different rheological parameterisations agree within the context of Antarctic upper mantle structure are invited to view the work of Ivins et al. (2021).

2.1. Independent Geophysical Data Sets

Independent constraints on mantle properties are collated and used as data sets in a joint inversion. These data are complementary in that they are collected over a range of depths (0–400 km) and temperatures (0–1500°C), and help to tackle the issue of non-uniqueness via their different sensitivities to a given change in the parameter space.

The first constraint used is the observed $V_S(T)$ relationship in conductively cooling oceanic lithosphere. V_S data from a tomographic model may be compared to thermal structure obtained via numerical modeling when binned by age and depth (Richards, Hoggard, White, et al., 2020). The 15 km maximum vertical resolution of ANT-20 informs our decision to sample V_S and T data points in 25 km bins over the range 50–125 km. This depth range is chosen to avoid non-negligible compositional effects at shallow depths due to mantle melting and the potential incorporation of spurious low velocity structure resulting from the bleeding of crustal velocities down into deeper depth ranges.

To construct a suitable thermal model for Antarctica, a Crank-Nicholson finite difference scheme with a predictor-corrector step is used to numerically integrate the heat diffusion equation. We follow the implementation set out by Richards et al. (2018) and Richards, Hoggard, Crosby, et al. (2020), in which the heat capacity, C_P, mantle density, ρ, and thermal conductivity, k, vary as a function of temperature, T, and composition, X. The latter two variables are also dependent on pressure, P.
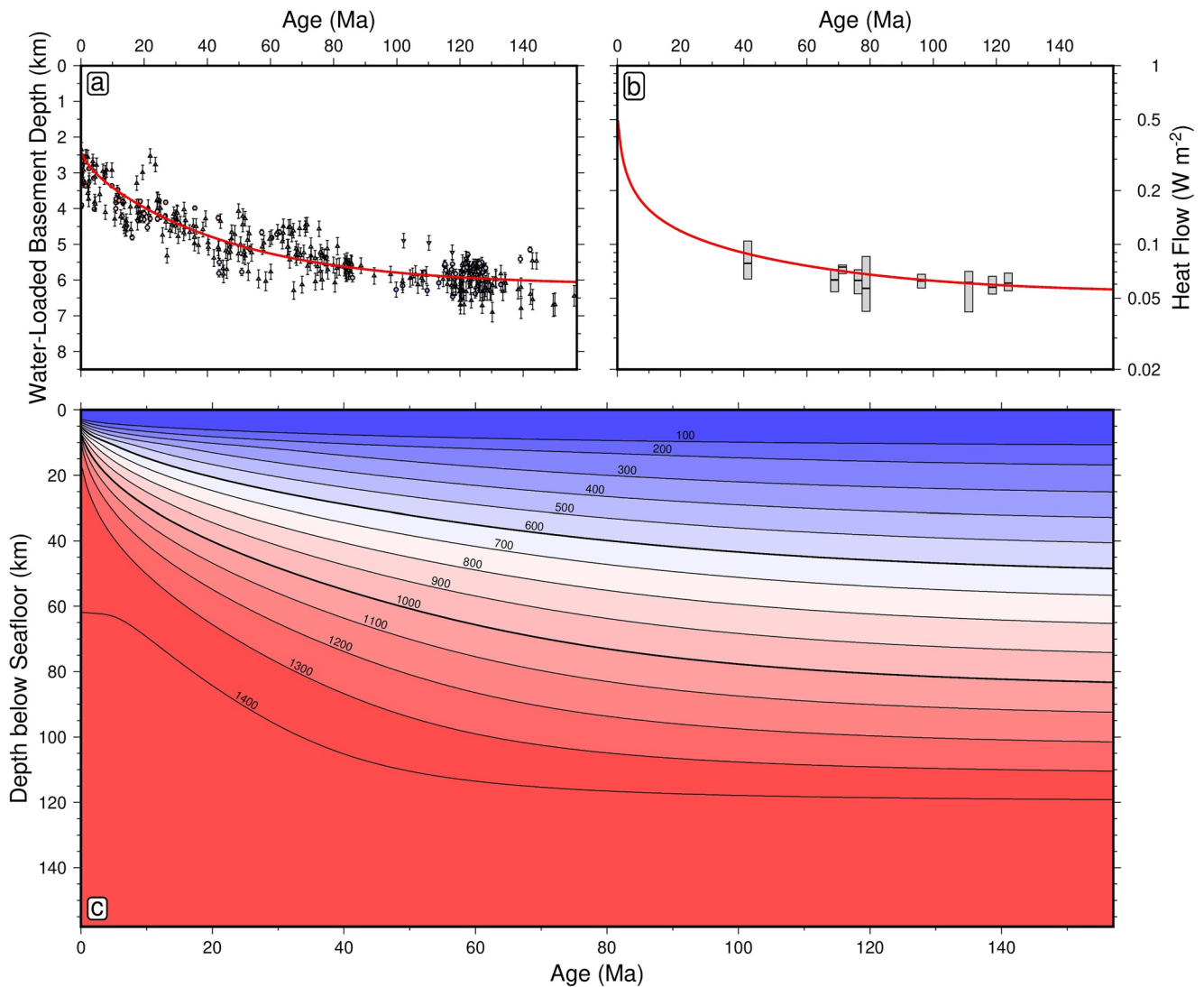
Figure 2. Thermal modeling of Antarctic oceanic lithosphere. (a) Thermal model fit to oceanic age-depth data from the Antarctic oceanic region placed into 2.5 Ma bins (Richards et al., 2018). (b) Same as (a) for age-heat flow data (Richards et al., 2018). (c) Plate cooling model solution constructed using a Crank-Nicholson finite-difference scheme to numerically solve the 1-D heat-diffusion equation (Richards et al., 2018, Richards, Hoggard, Crosby, et al., 2020). Antarctica-specific regional data are used to capture any local deviation from the globally averaged thermal trend. Model isotherms (°C) given by black curves in panel (c).

Mantle potential temperature, plate thickness and zero-age ridge depth are optimized by assessing the misfit to heat flux and subsidence data located within the footprint of the ANT-20 seismic tomographic model. The result is a best-fitting model describing the thermal structure beneath the oceans surrounding Antarctica (see Figure 2), suitable for comparison with V_S measurements over the same age-depth bins, such that the regional oceanic $V_S(T)$ relationship can be obtained for the lithosphere. We find a mantle potential temperature of $T_P = 1420 \pm 50$°C, approximately 5% hotter than the geochemically constrained global average $T_P = 1333$°C (Richards et al., 2018). Our regional best-fitting value of $T_P = 1420$°C is consistent with regional geochemical estimates of mantle potential temperature, which span the range 1314–1550°C, with an average 1385 ± 40°C (Figure 1). Although these constraints are only available along the circum-Antarctic ridge system, they are nevertheless indicative of anomalously hot mantle beneath the Southern Ocean (see Text S1 in Supporting Information S1). In the inversion, V_S measurements are compared to inferences of V_S from temperature at each age-depth bin (Figure 4a).

Since the lithospheric thermal model is only applicable at depths of $z \leq 125$ km, a $T_P = 1420$°C (1693 K) isentrope is used to characterize temperatures in the convective portion of the mantle, over the depth range $z = 225$–400 km. Here, temperature is calculated according to

$$T = T_P \exp\left(\frac{\alpha g z}{C_P}\right), \tag{5}$$

where α is thermal expansivity, g, acceleration due to gravity, and z, depth. This serves as the second constraint in the inversion, whereby V_S measurements are compared to inferences of V_S from temperature at each depth bin (Figure 4b). The third constraint is the QRFSI12 attenuation model of Dalton et al. (2008), which provides an average radial profile of seismic attenuation at depths $z = 150$–400 km beneath Antarctic ocean floor of age ≥ 100 Ma. Both the isentropic temperature and attenuation profiles are sampled at 25 km intervals to match the chosen V_S binning resolution (Figure 4c). To assess the misfit between data and model for these two constraints, tomographic V_S measurements are compared to V_S inferred from isentropic temperature, and attenuation measurements are compared to attenuation inferred from tomographic V_S, respectively. The final constraint used is a single estimate for the average steady state shear viscosity between 225 and 400 km depth, $\eta = 10^{20\pm1}$ Pa s (Lau et al., 2016, based on GIA modeling of relative sea-level and geodetic data assuming a Maxwell rheology, i.e., diffusion creep). Here, the viscosity constraint is compared to the corresponding depth-averaged viscosity inferred from V_S (Figure 4d).

2.2. Bayesian Modeling Framework

Formulating the inverse problem in a Bayesian framework entails treating each of the model parameters as random variables. There are several reasons why this is favorable to taking a deterministic approach. By incorporating hyperparameters, the reported uncertainties on each data set are scaled to more appropriately capture the misfit between data and model (see Eilon et al., 2018). This approach allows for integration of multiple constraints into a joint inversion without the need to use subjective weightings on each data set (Fukuda & Johnson, 2010). Secondly, prior information on the nature of the parameter space can be incorporated. Thirdly, the use of statistical sampling enables much more informative and rigorous treatment of uncertainty, and a natural way to propagate this into uncertainty in physical parameters of interest.

The objective of the inversion is to numerically characterize the a posteriori probability density function $p(\mathcal{X}|\mathcal{D})$. This function describes how the probability of an infinitesimal volume, $d\mathcal{X}$, of the model space, $\mathcal{X}$, varies as we traverse through it, given the observed data. In the following, we will refer to a particular choice of model using the superscript notation, $\mathcal{X}^i$. We will refer to a particular component of the model using the subscript notation, $\mathcal{X}_i$.

The model space $\mathcal{X} = \mathcal{X}(\boldsymbol{m}, \boldsymbol{\sigma})$ contains seven viscoelastic parameters

$$\boldsymbol{m} = \left[\mu_0, \partial\mu/\partial T, \partial\mu/\partial P, \eta_0, E_A, V_A, \partial T_S/\partial z\right]^T, \tag{6}$$

as well as three hyperparameters

$$\boldsymbol{\sigma} = [\sigma_1, \sigma_2, \sigma_3]^T. \tag{7}$$

An individual hyperparameter is used to tune the uncertainties for each data set. We omit a fourth hyperparameter associated with the viscosity constraint, due to the instability of this parameter when used to constrain a data set containing only a single data point. The posterior density, $p(\mathcal{X}|\mathcal{D})$, is dependent on the outcome of the experiments we undertake, via our data, $\mathcal{D}$. Since it is usually not possible to access $p(\mathcal{X}|\mathcal{D})$ analytically, we turn to stochastic methods.

Bayes' theorem states that the a posteriori density, $p(\mathcal{X}|\mathcal{D})$, is linked to the a priori information we have about the model space, as well as the likelihood of obtaining the observed data given a particular model, which are described by the density functions, $p(\mathcal{X})$ and $p(\mathcal{D}|\mathcal{X})$, respectively. The relationship is expressed mathematically as

$$p(\mathcal{X}|\mathcal{D}) = \frac{p(\mathcal{X})p(\mathcal{D}|\mathcal{X})}{p(\mathcal{D})}. \tag{8}$$

The a priori probability density on the data, $p(\mathcal{D})$, takes on a fixed value for a given set of observations and is thus treated as a normalization. This allows us to compare probability densities between two different models $\mathcal{X}^1$ and $\mathcal{X}^2$ by evaluating the ratio

Table 1
Prior and Posterior Estimates of the Inversion Parameters

Model sector	Model parameter i	Prior μ_i	Prior s_i	Posterior μ_i	Posterior s_i	MAP
Viscoelastic parameters $\boldsymbol{m}$	μ_0 [GPa]	81	8	74.8	0.4	74.8
	$\partial\mu/\partial T$ [GPa K^{-1}]	−0.014	0.003	−0.0129	0.0005	−0.0131
	$\partial\mu/\partial P$ [unitless]	1.6	0.2	2.04	0.06	2.09
	$\log_{10}(\eta_0[\text{Pa s}])$	22	3	23.2	0.7	22.9
	E_A [kJ mol^{-1}]	400	200	542	146	476
	V_A [cm^3 mol^{-1}]	6	4	5.35	0.32	5.02
	$\partial T_s/\partial z$ [K km^{-1}]	2.25	2.25	1.63	0.14	1.65
Hyperparameters $\boldsymbol{\sigma}$	$\log_{10}(\sigma_1\ [\text{unitless}])$	0	1	−0.317	0.024	−0.328
	$\log_{10}(\sigma_2\ [\text{unitless}])$	0	1	0.093	0.148	0.136
	$\log_{10}(\sigma_3\ [\text{unitless}])$	0	1	0.588	0.105	0.514

Note. The inversion parameters are made up of the seven material-dependent components of YT16, denoted by $\boldsymbol{m}$, as well as the three hyperparameters, denoted by $\boldsymbol{\sigma}$. Prior μ_i and s_i represent the mean and standard deviation of the Gaussian prior distribution for each parameter. For details of prior calculation, see Appendix B. Posterior μ_i and s_i are estimates of the mean and standard deviation of the posterior distribution for each parameter. MAP represents the maximum a posteriori model.

$$\frac{p(\mathcal{X} = \mathcal{X}^1|\boldsymbol{D})}{p(\mathcal{X} = \mathcal{X}^2|\boldsymbol{D})} = \frac{p(\mathcal{X} = \mathcal{X}^1)p(\boldsymbol{D}|\mathcal{X} = \mathcal{X}^1)}{p(\mathcal{X} = \mathcal{X}^2)p(\boldsymbol{D}|\mathcal{X} = \mathcal{X}^2)}. \tag{9}$$

Therefore, to estimate the variation in posterior density one needs a suitable method for calculating the prior and likelihood functions. The prior should be selected as a function which agnostically summarizes the knowledge one has about the model space before performing the inversion, usually in the form of a uniform or Gaussian distribution. Here we use a Gaussian distribution to summarize our prior knowledge of each parameter $\mathcal{X}_i$,

$$p(\mathcal{X}_i) = \frac{1}{\sqrt{2\pi}s_i}\exp\left(-\frac{(\mathcal{X}_i - \mu_i)^2}{2s_i^2}\right), \tag{10}$$

where μ_i and s_i represent our prior estimate and its uncertainty respectively. This distribution is useful as it enforces a non-zero probability density for any choice of model, $\mathcal{X}^i$, and enables us to use conservative uncertainty estimates for model parameters based on experimental studies (Table 1). The assumption that each model parameter is conditionally independent is taken, allowing the multiplication of the prior on each parameter to form an overall prior density given by

$$p(\mathcal{X}) = \prod_{i=1}^{i=N_p} \frac{1}{\sqrt{2\pi}s_i}\exp\left(-\frac{(\mathcal{X}_i - \mu_i)^2}{2s_i^2}\right), \tag{11}$$

where N_p is the number of parameters within the model.

We will also assume that the data points within each data set are independent, allowing us to use a Gaussian distribution to describe the likelihood function for each data set,

$$p(\boldsymbol{d}_k|\mathcal{X}(\boldsymbol{m},\boldsymbol{\sigma})) = \frac{1}{\left(2\pi\sigma_k^2\right)^{N_k/2}|\boldsymbol{\Sigma_k}|^{1/2}}\exp\left(-\frac{1}{2\sigma_k^2}\left(\boldsymbol{d}_k - \hat{\boldsymbol{d}}_k\right)^T\boldsymbol{\Sigma_k}^{-1}\left(\boldsymbol{d}_k - \hat{\boldsymbol{d}}_k\right)\right). \tag{12}$$

In this equation, $\boldsymbol{d}_k$ represents the kth data set containing N_k data points, $\hat{\boldsymbol{d}}_k = \hat{\boldsymbol{d}}_k(\mathcal{X})$ the corresponding model prediction, $\boldsymbol{\Sigma_k}$ the data covariance matrix containing the uncertainty on each data point, and σ_k the hyperparameter weighting applied to the data set.

If the data sets are independent of each other, the overall likelihood function can be constructed by simply multiplying together the likelihood function for each of the N_d data sets:

$$p(\boldsymbol{D}|\mathcal{X}(\boldsymbol{m},\boldsymbol{\sigma})) = \prod_{k=1}^{k=N_d} \frac{1}{\left(2\pi\sigma_k^2\right)^{N_k/2}|\boldsymbol{\Sigma_k}|^{1/2}}\exp\left(-\frac{1}{2\sigma_k^2}\left(\boldsymbol{d}_k - \hat{\boldsymbol{d}}_k\right)^T\boldsymbol{\Sigma_k}^{-1}\left(\boldsymbol{d}_k - \hat{\boldsymbol{d}}_k\right)\right). \tag{13}$$

Once a set of mathematical expressions for the prior and likelihood densities has been established as above, we may select a suitable algorithm to characterize the posterior space. The Metropolis-Hastings algorithm is one of the most common methods for doing so and involves generating a chain of models with associated posterior density values (Metropolis et al., 1953).

Given a current model $\mathcal{X}^n$, a proposal model $\mathcal{Y}^{n+1}$ is constructed according to the relationship

$$\mathcal{Y}^{n+1} = \mathcal{X}^n + \mathcal{P}, \tag{14}$$

where $\mathcal{P} \sim \mathcal{N}\left(\mathbf{0}, \mathbf{\Sigma}^{\text{proposal}}\right)$ and $\mathbf{\Sigma}^{\text{proposal}}$ is a suitable $N_p \times N_p$ proposal sampling covariance matrix. For simplicity, this matrix is typically chosen to be diagonal. The proposal model is accepted with probability

$$a^n = \min\left(1, \frac{p\left(\mathcal{Y}^{n+1}|\boldsymbol{D}\right)}{p\left(\mathcal{X}^n|\boldsymbol{D}\right)} \right), \tag{15}$$

where a^n is known as the acceptance ratio and is calculated using Equation 9. If the proposal model is accepted one sets $\mathcal{X}^{n+1} = \mathcal{Y}^{n+1}$. Otherwise the current model remains and one sets $\mathcal{X}^{n+1} = \mathcal{X}^n$. This process is repeated until the parameter space is suitably explored. Since the probability of a model being accepted is proportional to its posterior density, convergence toward optimal regions of the parameter space occurs. However, less probable models still have a finite acceptance probability, meaning the procedure is also capable of escaping local minima. To circumvent the issue that the evolution of samples is, at first, correlated with the initial starting point, the first 50% of trials are discarded as a so-called "burn-in" period. Only the post burn-in set of samples are used in the analysis.

While powerful, in our case, the Metropolis-Hastings algorithm in its original form is not sophisticated enough to perform the inversion efficiently, since strong trade-offs between model parameters invalidate the use of a diagonal proposal covariance matrix. The precise form of $\mathbf{\Sigma}^{\text{proposal}}$ has a strong impact on the average model acceptance rate a, which is optimized when $a \approx a^*$, where $a^* = 0.234$ (Gelman et al., 1997). When $\mathbf{\Sigma}^{\text{proposal}}$ is too small, a large proportion of models are accepted but only small steps around the model space are taken. When $\mathbf{\Sigma}^{\text{proposal}}$ is too large, only a small proportion of models are accepted and so the inversion algorithm tends to sample the same area of the model space for a prohibitively large number of trials, before wildly jumping elsewhere. This applies when any region of the multi-dimensional proposal covariance space is poorly estimated. Both situations lead to inefficient convergence toward the posterior distribution and so, for a finite number of trials, inhibit our ability to achieve a useful result. We therefore adopt the Global Adaptive Scaling Within Adaptive Metropolis (GASWAM) modification of Metropolis-Hastings (Andrieu & Thoms, 2008; Figure 3; see Appendix C for methodological details).

There are two practical ideas underpinning the GASWAM algorithm. The first is that the most efficient choice of proposal covariance matrix, $\mathbf{\Sigma}^{\text{proposal}}$, is a scalar function of the model covariance matrix $\mathbf{\Sigma}^{\mathcal{X}}$. The second is that $\mathbf{\Sigma}^{\mathcal{X}}$ can be estimated for a given trial, n, of the inversion from the complete history of all preceding trials, $\{0, 1, \ldots, n-1\}$, using the formula

$$\mathbf{\Sigma}^{\mathcal{X}} \approx \frac{1}{n-2} \sum_{i=1}^{n-1}\left(\mathcal{X}^i - \overline{\mathcal{X}}\right)\left(\mathcal{X}^i - \overline{\mathcal{X}}\right)^T, \tag{16}$$

where $\overline{\mathcal{X}}$ is the iteratively updated average model

$$\overline{\mathcal{X}} = \frac{1}{n-1} \sum_{i=1}^{n-1} \mathcal{X}^i. \tag{17}$$

The GASWAM algorithm makes use of this empirically calculated covariance structure and an associated scale factor to progressively update the proposal covariance matrix. By simultaneously updating the shape and size of the proposal covariance matrix, stabilization of the inversion procedure can be achieved by enforcing the optimal acceptance ratio, such that $a \approx 0.234$. This stability is ensured by looking at a suite of convergence diagnostics including the running mean of each parameter as the trial proceeds, frequency density plots of each parameter, the potential scale reduction factor (Gelman et al., 1997; Roy, 2020), and the fit of the models to the data (Figure 4). The performance of the inversion algorithm was also tested against synthetic data, verifying that it behaved as expected (see Text S2 in Supporting Information S1 for details).
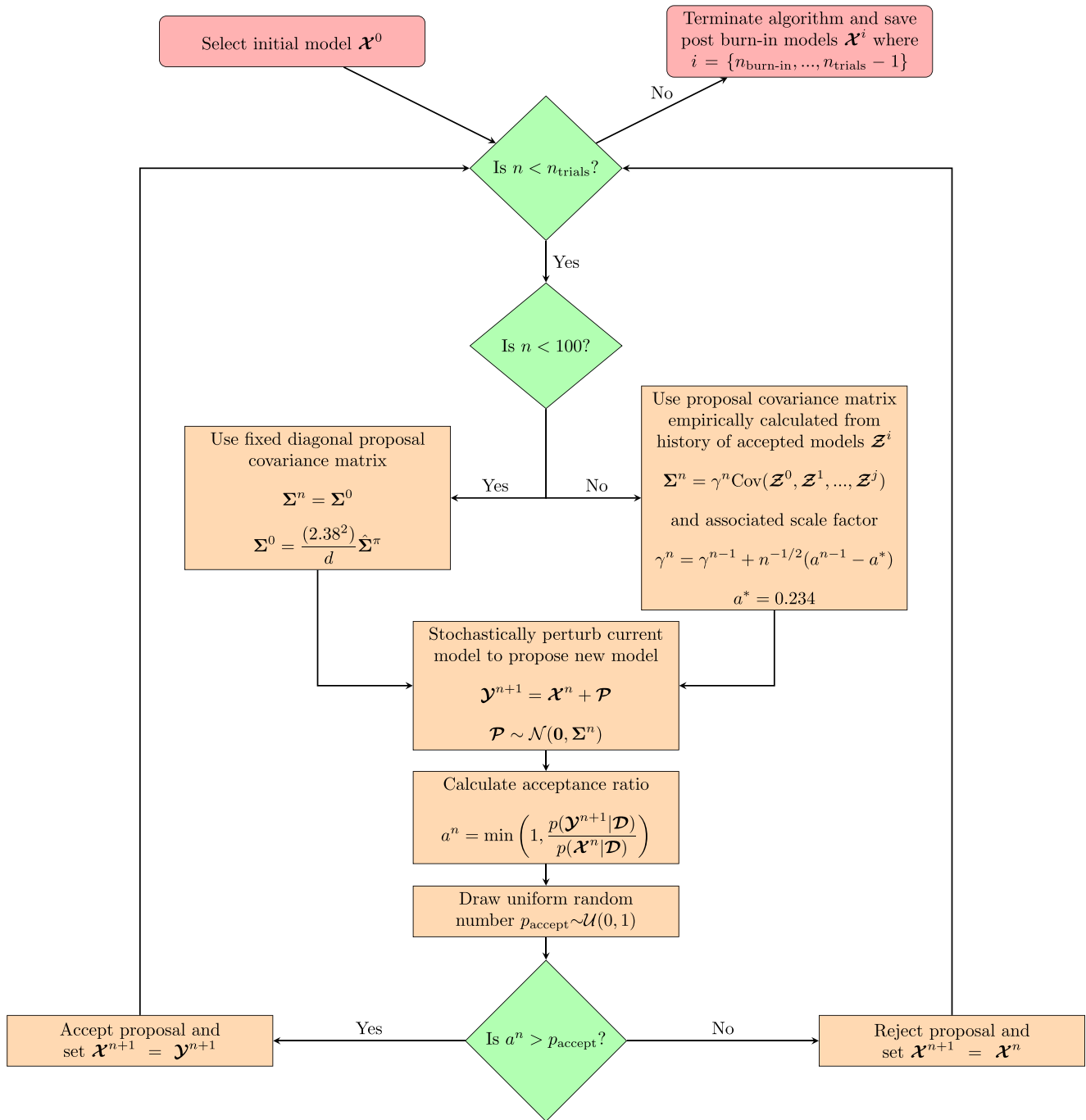
Select initial model $\boldsymbol{\mathcal{X}}^0$

Terminate algorithm and save post burn-in models $\boldsymbol{\mathcal{X}}^i$ where $i = \{n_{\text{burn-in}}, ..., n_{\text{trials}} - 1\}$

No

Is $n < n_{\text{trials}}$?

Yes

Is $n < 100$?

Yes / No

Use fixed diagonal proposal covariance matrix

$$\boldsymbol{\Sigma}^n = \boldsymbol{\Sigma}^0$$

$$\boldsymbol{\Sigma}^0 = \frac{(2.38^2)}{d}\hat{\boldsymbol{\Sigma}}^\pi$$

Use proposal covariance matrix empirically calculated from history of accepted models $\boldsymbol{\mathcal{Z}}^i$

$$\boldsymbol{\Sigma}^n = \gamma^n \text{Cov}(\boldsymbol{\mathcal{Z}}^0, \boldsymbol{\mathcal{Z}}^1, ..., \boldsymbol{\mathcal{Z}}^j)$$

and associated scale factor

$$\gamma^n = \gamma^{n-1} + n^{-1/2}(a^{n-1} - a^*)$$

$$a^* = 0.234$$

Stochastically perturb current model to propose new model

$$\boldsymbol{\mathcal{Y}}^{n+1} = \boldsymbol{\mathcal{X}}^n + \boldsymbol{\mathcal{P}}$$

$$\boldsymbol{\mathcal{P}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}^n)$$

Calculate acceptance ratio

$$a^n = \min\left(1, \frac{p(\boldsymbol{\mathcal{Y}}^{n+1}|\boldsymbol{\mathcal{D}})}{p(\boldsymbol{\mathcal{X}}^n|\boldsymbol{\mathcal{D}})}\right)$$

Draw uniform random number $p_{\text{accept}} \sim \mathcal{U}(0,1)$

Accept proposal and set $\boldsymbol{\mathcal{X}}^{n+1} = \boldsymbol{\mathcal{Y}}^{n+1}$

Yes

Is $a^n > p_{\text{accept}}$?

No

Reject proposal and set $\boldsymbol{\mathcal{X}}^{n+1} = \boldsymbol{\mathcal{X}}^n$

Figure 3. Flow chart representation of the Globally Adaptive Scaling Within Adaptive Metropolis adaptation (Andrieu & Thoms, 2008) of the Metropolis-Hastings algorithm (Metropolis et al., 1953). Optimal acceptance ratio, $a^* = 0.234$, from Gelman et al. (1997). Initial proposal sampling covariance matrix based on the proposition of Haario et al. (2001). Algebraic superscripts refer to a particular choice of model.

The result of the inversion is a set of post burn-in models, $\boldsymbol{\mathcal{X}}^{\text{posterior}}$. This serves as a discrete set of samples over the continuous posterior density function, $p(\boldsymbol{\mathcal{X}}|\boldsymbol{\mathcal{D}})$. A greater sampling density is indicative of a more probable region of the model space. Since the sampled posterior distribution (ignoring hyperparameters) is seven-dimensional, it cannot be visualized as a whole. Instead, we calculate the sampling density for each combination of model parameters, $\mathcal{X}_i$ and $\mathcal{X}_j$. To achieve this, the posterior space of each parameter is discretized into 1,000 blocks, spanning the range of values over which this parameter was sampled. This results in a step-size given by
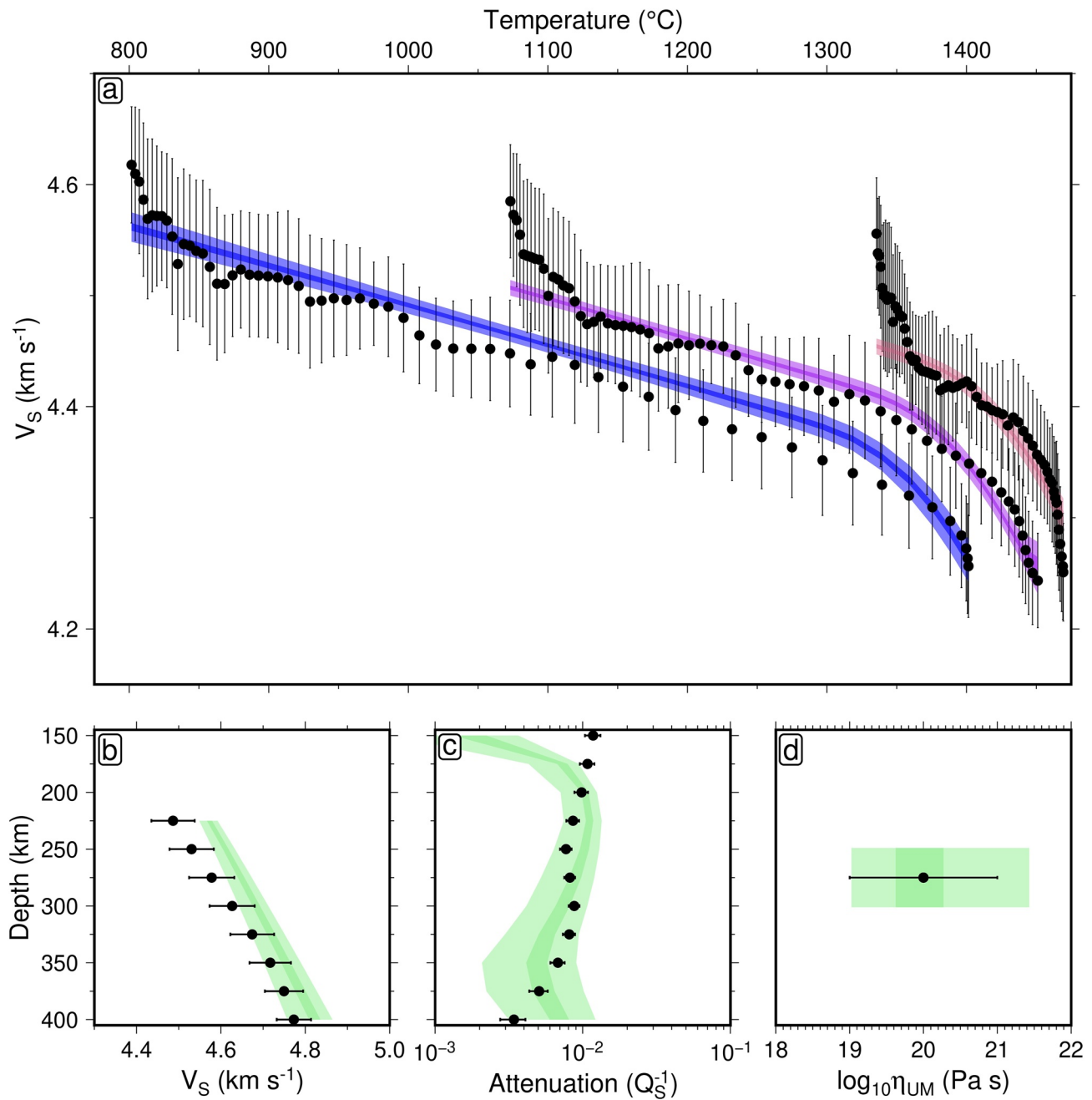
Figure 4. Assessing fit of inverted viscoelastic parameters. Fit of post burn-in models to the four geophysical data sets used to constrain the inversion procedure (circles/error bars; see Section 2.1). Pale shaded regions represent the 99% credible interval, and dark shaded regions represent the 50% credible interval. (a) Plate cooling model fit for depth ranges 50–75 km (blue), 75–100 km (purple), and 100–125 km (red). (b) Adiabatic model fit for depth range 225–400 km. (c) QRFSI12 seismic attenuation model fit at depths 150–400 km beneath ocean floor for ages ≥ 100 Ma. (d) Average viscosity between 225 and 400 km compared to $\eta = 10^{20 \pm 1}$ Pa s estimate.

$$h_i = \frac{\text{maximum}(\mathcal{X}_i) - \text{minimum}(\mathcal{X}_i)}{1,000}. \tag{18}$$

The sampling density is then calculated as

$$\rho_{ij}(x, y) = n_{ij}(x, y)/A_{ij}, \tag{19}$$

where (x, y) is the grid reference pertaining to each of the $1{,}000 \times 1{,}000$ discrete areas in which density values are calculated, $n_{ij}(x, y)$ is the corresponding number of posterior samples, and $A_{ij} = h_i h_j$ is the corresponding area. To yield further information from the inversion output, the model samples and their corresponding physical predictions must be summarized mathematically. The expectation value of each parameter can be estimated using the discrete summation

$$\hat{E}(\mathcal{X}_i) = \frac{1}{N_s} \sum_{j=1}^{N_s} \mathcal{X}_i^j, \tag{20}$$

where N_s is the number of discrete model samples (Gallagher et al., 2009). The corresponding variance of each parameter may be estimated using the formula

$$\hat{V}(\mathcal{X}_i) = \frac{1}{N_s - 1} \sum_{j=1}^{N_s} \left(\mathcal{X}_i^j - \hat{E}(\mathcal{X}_i) \right)^2. \tag{21}$$

A summary of the posterior parameters we obtain is shown in Table 1, and compared to the results of Richards, Hoggard, White, et al. (2020) in Text S3 in Supporting Information S1. However, the vectors $\hat{E}(\mathcal{X})$ and $\hat{V}(\mathcal{X})$ do not tell the full story. The anelasticity model $\mathcal{X}$ serves as a means for converting V_S into physical predictions of temperature, T, viscosity, η, and density, ρ. We are therefore interested in estimating the expectation value and variance of functions of the model $f(\mathcal{X})$, rather than the model itself. This can be achieved easily, by constructing a vector, $\mathcal{F}$, where each component is calculated according to the formula

$$\mathcal{F}^i = f\left(\mathcal{X}^i\right). \tag{22}$$

The expectation value and variance of the physical prediction can be estimated analogously to Equations 20 and 21, resulting in the equations

$$\hat{E}(\mathcal{F}) = \frac{1}{N_s} \sum_{i=1}^{N_s} \mathcal{F}^i, \tag{23}$$

and

$$\hat{V}(\mathcal{F}) = \frac{1}{N_s - 1} \sum_{i=1}^{N_s} \left(\mathcal{F}^i - \hat{E}(\mathcal{F}) \right)^2. \tag{24}$$

The estimates for the expectation value, $\hat{E}(\mathcal{F})$, and variance, $\hat{V}(\mathcal{F})$, are referred to as the average and uncertainty, respectively. If it is not practical to calculate all N_S values of $\mathcal{F}^i$, due to computational expense, a subset N_U of the overall set of post burn-in models may be used (see Section 4). The relationship between the uncertainty on a physical prediction, $\hat{V}(f(\mathcal{X}))$, and the uncertainty on the underlying model parameters, $\hat{V}(\mathcal{X})$, is dependent on the sensitivity of $f(\mathcal{X})$ to each parameter, $\mathcal{X}_i$ (i.e., the gradient, $\partial f(\mathcal{X})/\partial \mathcal{X}_i$), and the covariance structure of the model, $\Sigma^{\mathcal{X}}$ (Champac & Garcia Gervacio, 2018). In the case of the anelasticity parameterization, $T(\mathcal{X})$ and $\eta(\mathcal{X})$ are non-linear functions of V_S, complicating the analytical calculation of their expectation value and variance. This highlights one of the key benefits of taking a Bayesian approach, as it provides a simple way of propagating uncertainties, using the discrete summaries of Equations 23 and 24.

3. Anelasticity Model Covariance Structure

To investigate how dependent a particular model parameter is on the choice of another, we plot the posterior sampling density for each parameter combination (Figure 5). This highlights the presence of clear trade-offs, as expected given our need to adapt the proposal sampling scheme to handle non-diagonal model covariance structure. We find that the anelasticity model $\boldsymbol{m}$ can be approximately separated into two independent components, $\boldsymbol{A} = \{\mu_0, \partial\mu/\partial T, \partial\mu/\partial P\}$ and $\boldsymbol{B} = \{\eta_0, E_A, V_A, \partial T_S/\partial z\}$, such that $\boldsymbol{m} = \{\boldsymbol{A}, \boldsymbol{B}\}$. A reasonable approximation for the model covariance structure therefore takes the form

$$\Sigma^{\mathcal{X}} \approx \begin{bmatrix} \Sigma^{\mathbf{A}} & \mathbf{0} \\ \mathbf{0} & \Sigma^{\mathbf{B}} \end{bmatrix}. \tag{25}$$
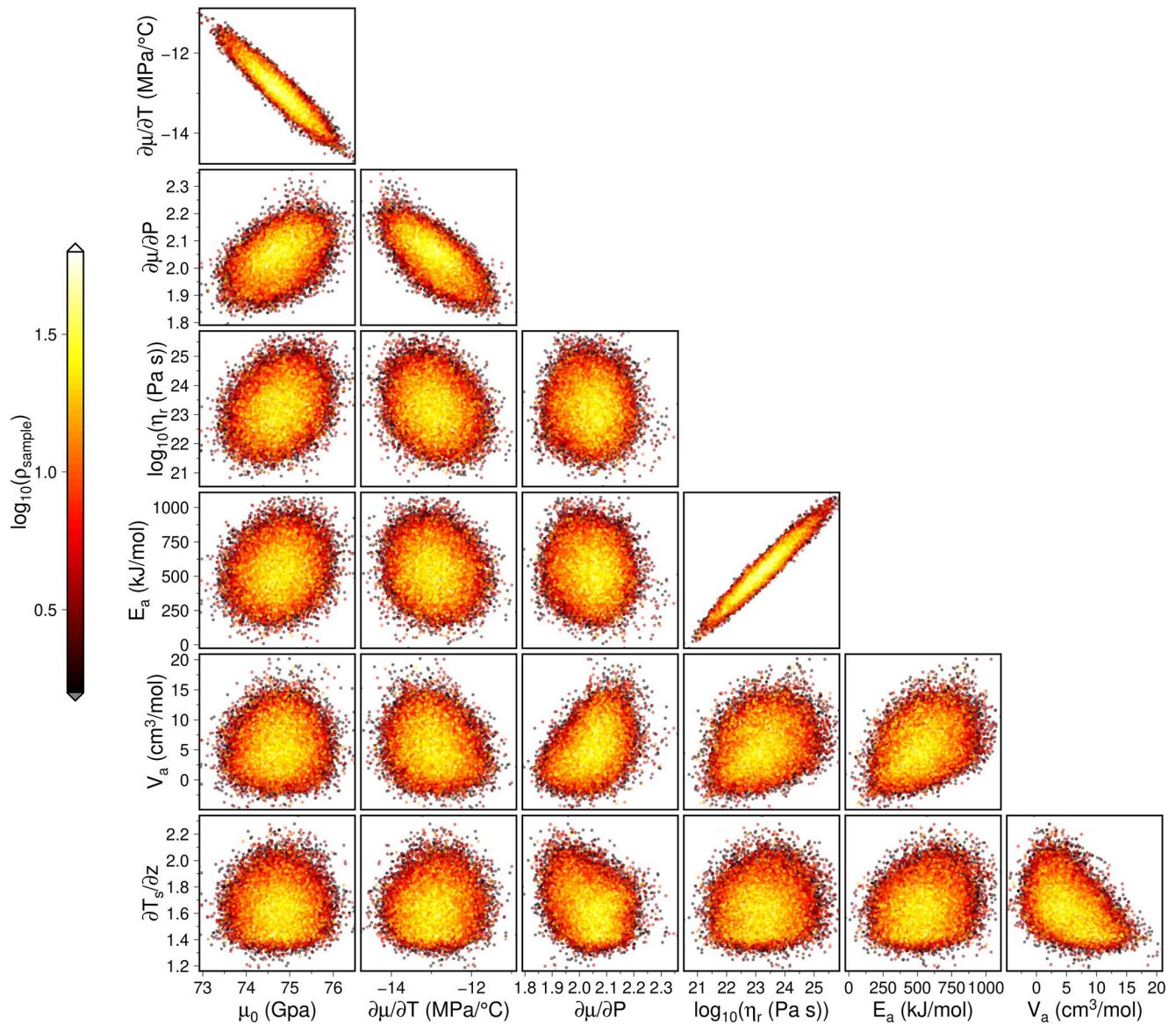
Figure 5. Posterior distributions of viscoelastic parameters. Posterior sampling density (ρ, arbitrary units) of each combination of anelasticity models, highlighting physical trade-offs between parameters.

There exist strong parameter trade-offs within **A** and **B** separately, but only weak trade-offs between **A** and **B**. This is in accordance with what we expect physically, whereby **A** regulates the elastic component of the physical response, and **B** the transient component.

Within **A**, we observe a very strong negative trade-off between the reference shear modulus and its temperature derivative. This implies that with respect to the maximum a posteriori estimate for this combination of parameters, a similar fit to the data can be obtained by co-varying μ_0 and $\partial\mu/\partial T$ in opposite directions. It is possible to verify that this makes sense in the context of the plate model $V_S(T)$ relationship (Figure 4a), which serves as the main data constraint on the inversion, as follows. The linear region of the $V_S(T)$ relationship in a given depth bin is well-approximated by assuming an elastic response at fixed pressure, and may therefore be expressed as (Appendix D)

$$V_S(T) \approx \sqrt{\frac{\mu_0^*}{\rho}} - \sqrt{\frac{|\partial\mu/\partial T|^2}{4\mu_0^*\rho}}(T - T_0). \tag{26}$$

Therefore, an increase in the reference shear modulus has the effect of increasing the V_S value at which the $V_S(T)$ trend is initialized, as well as reducing the absolute gradient of the trend. The temperature gradient of the shear modulus must assume a correspondingly more negative value to compensate, in order to preserve the squared distance between data and model. A symmetric argument can be used to interpret the positive trade-off between reference shear modulus and its pressure gradient. The relative weakness of this $(\mu_0, \partial\mu/\partial P)$ trade-off compared to that of $(\mu_0, \partial\mu/\partial T)$ may, in part, be down to the lower level of information that our data set contains on the variation of V_S with depth. In addition, a negative trade-off between the temperature and pressure derivatives of the shear modulus appears to be present.

By far the strongest parameter trade-off observed within the anelasticity model is contained within B, between activation energy and reference viscosity; parameters controlling the onset and strength of anelastic effects respectively. The non-linear relationship between parameters in the anelastic regime prevents an analytical derivation of the trade-off between E_A and η_0. However, it appears that while the individual uncertainties on E_A and η_0 are very large, the strength of trade-off between the two ensures only a small variation in the misfit between data and model. Importantly, this relationship reduces the extent to which uncertainty in the individual parameters propagates into uncertainty in upper mantle thermomechanical structure (see Section 4.1). Trade-offs between other parameters within B appear to be present, although relatively weak. In order to further constrain the model covariance, more data containing information about anelastic behavior, especially at high pressure, are required.

4. Predictions of Thermomechanical Structure Beneath Antarctica

Given a depth slice of ANT-20, it is possible to convert each $V_S(\theta, \phi)$ value—where θ and ϕ represent longitude and latitude, respectively—into an estimate of thermomechanical state (viscosity, η, temperature, T, and density, ρ) by assuming a choice of anelasticity model $\mathcal{X}^i$. To assess the improvement achieved by using the inversion procedure to refine this choice of model, we generate a series of three mean and standard deviation viscosity structures, each calculated based on the 150 km depth slice of ANT-20. In each case, we select N_U anelasticity models, summarizing the results by substituting $\log_{10}\eta$ into Equations 23 and 24. This results in a geometric mean and standard deviation of the viscosity at each location. In case I, each parameter is sampled independently from the prior distribution (Table 1). This represents, conservatively, the quality of Antarctic viscosity prediction that we can make based purely on experimental data pertaining to the mechanical behavior of the upper mantle. In case II, each parameter is sampled independently from the posterior distribution (Table 1). This represents the quality of prediction it is possible to make having calibrated the viscoelastic parameters with independent geophysical data, but ignoring any information on the covariance between parameters. Finally, in case III, the optimal approach laid out in Section 2.2 is taken, using a uniform random sample of posterior anelasticity models from the full set of $N_S = 200,000$ post burn-in models. This represents our best constraint on viscosity structure, including not only the refinement of individual parameters based on the data, but also information that the data provides about the model covariance structure. The use of a subset of the post burn-in models ensures computational viability. A suitable value for the sieving ratio N_U/N_S, representing the proportion of total post burn-in models used at the prediction stage, was found by investigating the additional information obtained by increasing N_U in integer steps, starting at 1 (Figure 6). It was ascertained that N_U as small as 100 was sufficient to bring deviations in the mean and standard deviation viscosity structure down to a fraction of a percentage upon the addition of an extra anelasticity model, and therefore a safe choice of $N_U = 1,000$ was taken.

A large reduction in uncertainty (4–5 orders of magnitude) is observed from case I–III (Figure 7), highlighting the benefit of the inversion as a whole. The most dramatic improvement occurs between case II and III, due to the effect of the highly non-diagonal covariance structure, which, due to compensation, results in muted variation in physical predictions for posterior models that encompass wide parameter ranges. Constraining the covariance structure of the physical model used to convert between shear-wave velocity and thermomechanical parameters is therefore central to the quality of the result we obtain. As a result, complementary data sets such as those used to calibrate the inversion here are hugely important. We may conclude from this assessment that the statistical inverse framework, as utilized optimally in Case III, provides the basis for improved predictions of thermomechanical structure. Therefore, this approach is taken to calculate a range of physical outputs in the results that follow.
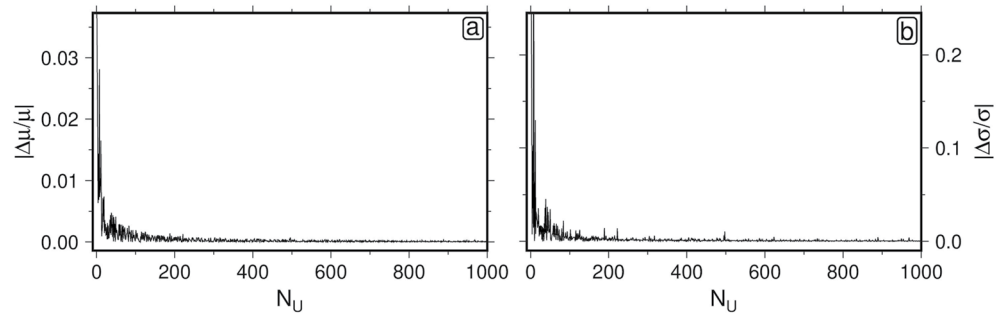
Figure 6. Determining the number of models required to accurately characterize posterior expectation and variance. Stability of the expectation value μ, and the uncertainty σ, of predicted viscosity at an arbitrary location ($z = 150$ km, longitude $\theta = 0.00°$, latitude $\phi = -90.00°$) of the ANT-20 V_S model, as a function of the number of randomly selected posterior anelasticity models used to construct them, denoted by N_U. Calculated by looking at the discrepancy in a physical prediction, X, before and after introducing an extra model, $\left(\hat{X}(n+1) - \hat{X}(n) \right) / \hat{X}(n+1)$. (a) $X = \mu$. (b) $X = \sigma$.

4.1. Viscosity Structure

The diffusion creep viscosity structure derived from the application of our conversion method to the ANT-20 V_S model contains significant lateral heterogeneity beneath the Antarctic continent and surrounding oceans. This is to be expected given the presence of shear-wave velocity anomalies up to 8% in relative amplitude observed in the underlying tomography (see Figure 11 in Lloyd et al., 2020). To show how this behavior manifests itself in terms of viscosity variation, we calculate geometric mean and standard deviation viscosity structures as a function of depth (Figures 8 and 9). Note that at low homologous temperatures, the anelastic contribution to V_S variation is negligible, meaning that viscosities cannot be reliably constrained when $\eta > 10^{22.5}$ Pa s (white contours in Figures 8 and 9). However, this is an issue of minor significance, since regions with viscosities above this threshold have Maxwell relaxation times exceeding 20 kyr and will behave elastically over the timescales relevant to GIA modeling. In the analysis that follows, we define the asthenosphere as the region in which $\eta < 10^{22.5}$ Pa s.

At 150 km depth, the thermomechanical dichotomy between East and West Antarctica is most obvious; a sharp viscosity boundary follows the path of the TAM across the continent from the Ross to the WS. The mantle at this depth is lithospheric beneath much of the EAIS, and asthenospheric beneath the WAIS. Within West Antarctica itself, viscosity varies within the range 10^{19-23} Pa s, and it is possible to identify two long-wavelength low-viscosity anomalies. The first arises at the Macquarie Triple Junction, extends to the Balleny Islands, and follows the TAM as it passes into West Antarctica through the western side of the Ross Embayment. The second passes from Marie Byrd Land (MBL), through the ASE, to the Antarctic Peninsula (AP). Both anomalies contain viscosities as low as $\eta \sim 10^{19}$ Pa s, and are also identifiable at 75 km, where they form a connected region which is the only portion of asthenosphere within the continental footprint at this depth. Model-based uncertainty in asthenospheric viscosity is very low (~0.3 orders of magnitude) at 75 and 150 km depth, and appears mostly homogeneous, albeit increasing appreciably within localized regions of very high viscosity. Given the small size of these regions (e.g., the higher viscosity patch beneath the Siple Coast at 150 km depth), it is difficult to rule out the possibility that they result from tomographic artifacts.

At deeper depths (250 and 350 km), average asthenospheric viscosities within the continent are higher ($\overline{\eta} = 10^{20.5 \pm 0.5}$ Pa s and $\overline{\eta} = 10^{21.4 \pm 0.6}$ Pa s, in terms of median and median absolute deviation, respectively) and the area of lithospheric coverage is reduced, leading to an overall more homogeneous structure. The low-viscosity anomaly observed at shallow depths beneath the AP has evolved into a high-viscosity anomaly that extends toward the South Scotia ridge by a depth of 350 km, possibly representing a fossil slab (An et al., 2015). Low viscosity regions present beneath the Ross and ASEs at 150 km persist at these depths, although the high viscosities that separate the two regions at shallower depths appear muted or absent. In addition, a large low viscosity anomaly can be seen in the Southern Ocean in the vicinity of MBL, consistent with the presence of a mantle plume (Seroussi et al., 2017). Average asthenospheric viscosity uncertainty increases with depth, likely reflecting the lack of deep geophysical data used to constrain the inversion for material properties. In particular, the inversion procedure is unable to constrain activation volume beyond an individual parameter precision of approximately 10%. Since this parameter governs the pressure-dependence of viscosity, deep viscosity uncertainty is highly
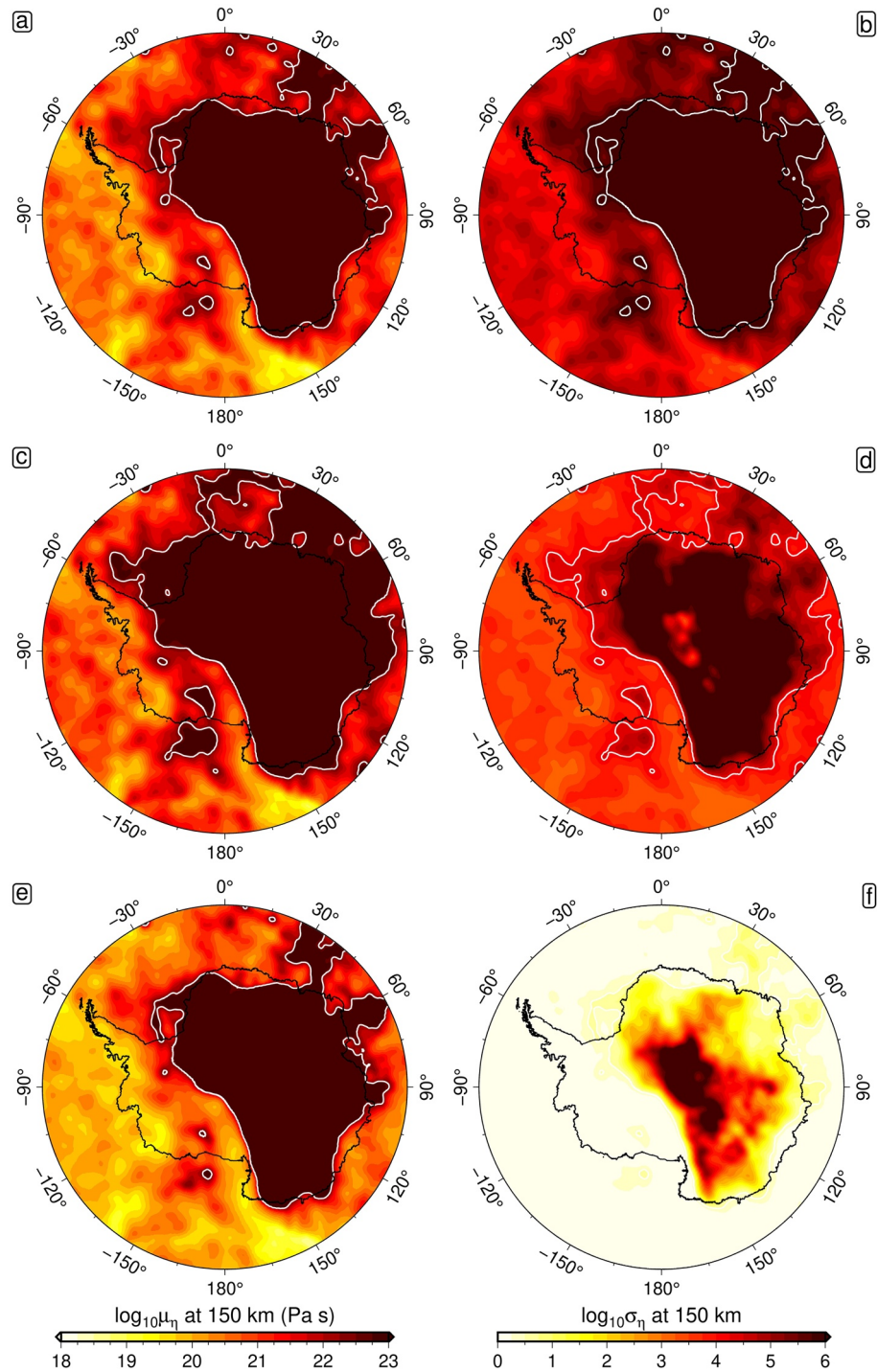
Figure 7. Diffusion creep viscosity and uncertainty based on forward and inverse modeling. Geometric mean (left-hand column; panels (a), (c), (e)) and standard deviation (right-hand column; panels (b), (d), (f)) viscosity structure at 150 km, calculated using three different methods. First (top row; panels (a), (b)), by sampling viscoelastic parameters independently from the prior distribution (see Table 1). Secondly (middle row; panels (c), (d)), by sampling viscoelastic parameters independently from the posterior distribution. Finally (bottom row; panels (e), (f)), by sampling sets of viscoelastic parameters from the posterior output. In each case, $N_U = 1,000$ models are used to generate the ensemble of viscosity predictions. White contours denote regions in which mean viscosity $\mu_\eta > 10^{22.5}$ Pa s.
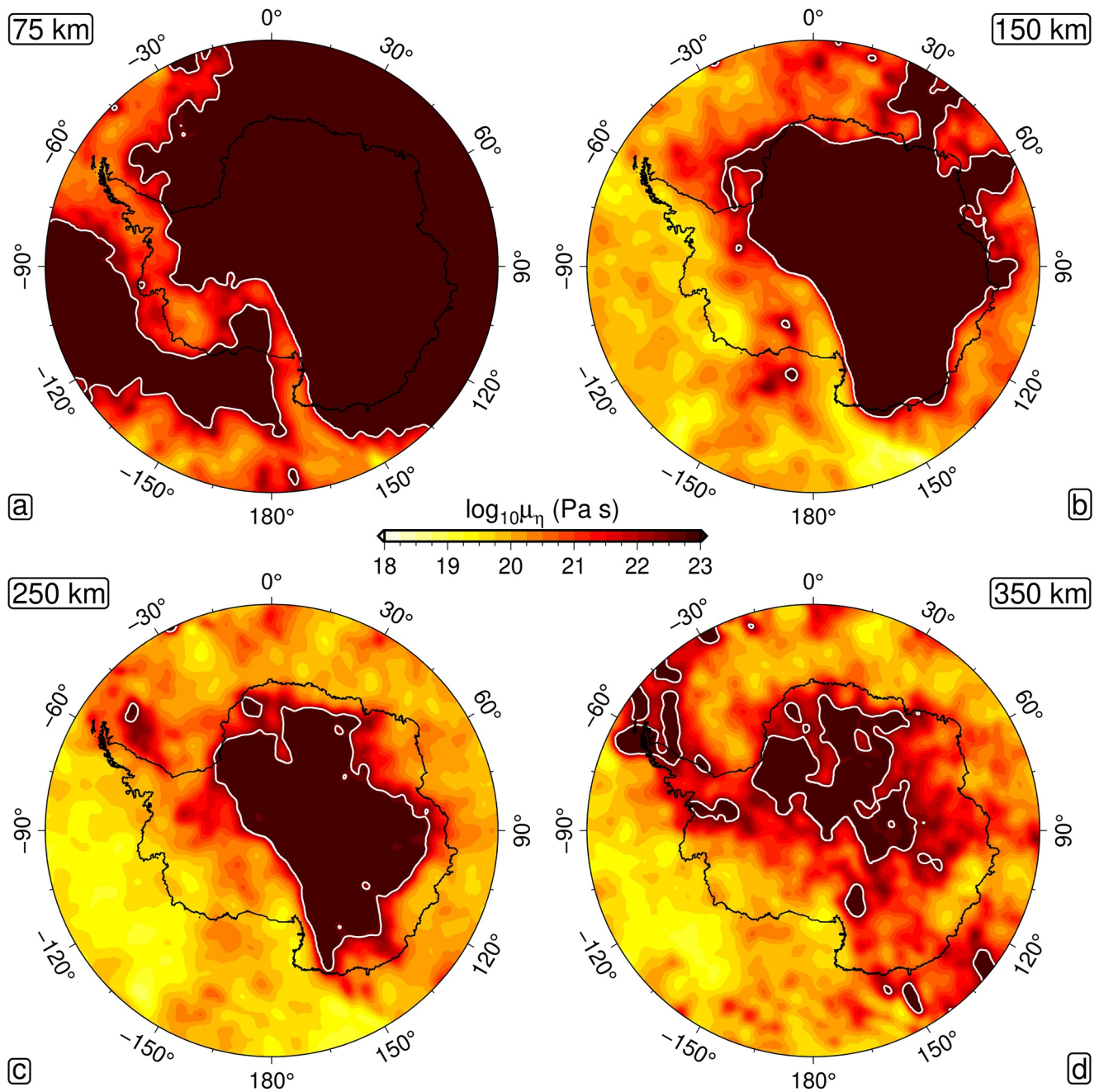
Figure 8. Diffusion creep viscosity beneath Antarctica. Geometric mean viscosity structure at 75, 150, 250, and 350 km depth ((a), (b), (c), and (d), respectively). Each structure is calculated by utilizing a uniform random sample of $N_U = 1,000$ posterior anelasticity models to convert ANT-20 shear-wave velocities into viscosity, and averaging the resulting ensemble.

correlated with activation volume uncertainty. However, lateral variations in uncertainty structure remain minimal, and even at 350 km depth do not exceed an order of magnitude.

4.2. Lithosphere-Asthenosphere Boundary Depth

The framework used to construct self-consistent predictions of thermomechanical structure beneath Antarctica can also be utilized to constrain other parameters important for GIA and ice-sheet modeling studies. First, we utilize the ensemble of three-dimensional temperature structures to infer LAB depth. For each temperature
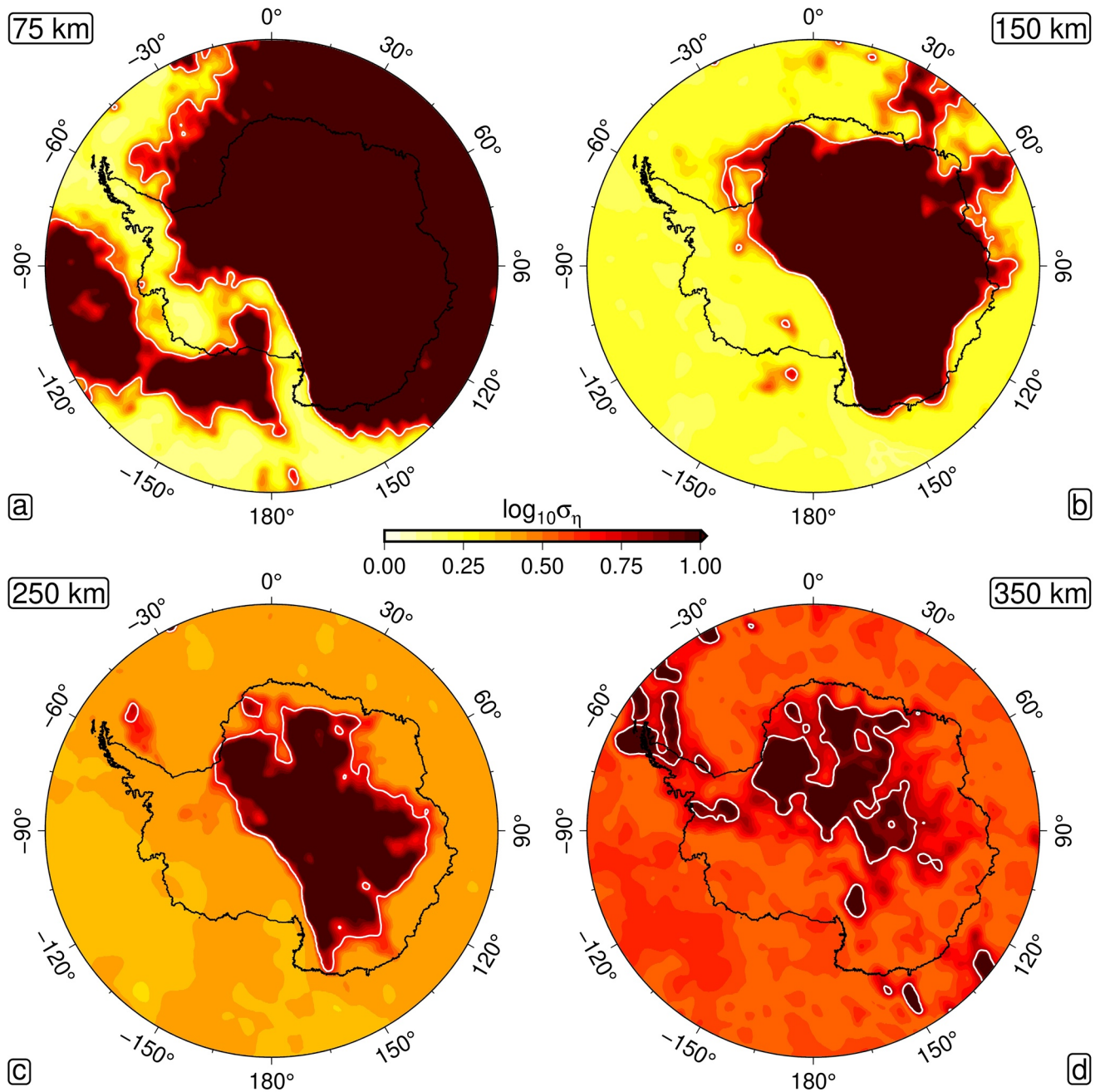
Figure 9. Uncertainty in diffusion creep viscosity beneath Antarctica. Geometric standard deviation viscosity structure at 75, 150, 250, and 350 depth ((a), (b), (c), and (d), respectively). Each structure is calculated by utilizing a uniform random sample of $N_U = 1,000$ posterior anelasticity models to convert ANT-20 shear-wave velocities into viscosity, and determining the variance of the resulting ensemble.

structure associated with a given choice of anelasticity model in the ensemble, the laterally varying geothermal profiles are interpolated to a 1 km depth interval. Prior to interpolation, anomalous temperatures associated with downward bleeding of crustal velocities in the underlying tomography are removed by identifying spurious reversals of the geothermal gradient and excising temperatures above these loci. In all cases, a temperature of 0°C is enforced at the basement depth, which can be estimated using the Moho depth and crustal thickness grids associated with the tomography. Following interpolation, we extract the depth at which the resulting profile intersects a temperature of 1200°C, a proxy for LAB depth (Figure S6 in Supporting Information S1; Burgos et al., 2014;

Richards et al., 2018). By summarizing the set of ensemble predictions of laterally varying LAB depth, according to Equations 23 and 24, we arrive at a mean and standard deviation LAB depth structure (Figure 10).

The resulting mean LAB depth displays a number of interesting features. We find good agreement with long-wavelength structure observed elsewhere in the literature (Priestley et al., 2018; Richards, Hoggard, White, et al., 2020), whereby LAB depth exceeds 150 km beneath the EAIS, and is much lower beneath the WAIS. We find spatially averaged LAB depths of 233 ± 41 and 63 ± 13 km beneath the respective ice sheets, calculated according to the median and median absolute deviation. The overall strength of this heterogeneity is high, involving LAB depths as shallow as 35 km in the West, and as deep as 365 km in the East (see Text S4 in Supporting Information S1). The full distribution of LAB depths represented by East and West Antarctica are shown in Figure 11, along with the classification used to distinguish between the two continental components, which is based on the satellite-mapped drainage network (Zwally et al., 2012).

The maximum ~15 km depth resolution and ~100 km lateral resolution of the underlying tomography is the dominant source of uncertainty on the calculated LAB structure over much of Antarctica, as a result of the low variance in LAB depth predictions provided by the ensemble of anelasticity models. However, this is not the case in certain areas of East Antarctica, where very large inferred LAB depths are also associated with large uncertainties, of order 30–40 km. The statistical uncertainty associated with the ensemble of anelasticity models is expected to rise with increasing LAB depth due to elevated temperature uncertainty with depth arising from the previously discussed uncertainty in activation volume.

By comparing the predictions made from our model of LAB depth to geological constraints not linked to the anelasticity calibration procedure, it is possible to verify that the temperature structures arrived at via the inversion method are realistic. We looked at the location and timing of Cenozoic magmatism, using a compilation of geochemical analyses on volcanic material (Ball et al., 2021; DIGIS Team, 2021). The data were spatially binned over a length scale of 100 km, in accordance with the seismological resolution, with minimum time since last eruption and its associated uncertainty extracted. The data were further processed to remove points with age uncertainties exceeding 10 Ma and the resulting data set mapped (Figure 10). Two key observations are immediately apparent when comparing magmatism and LAB depth. First, all sites containing a record of Pliocene or Quaternary (i.e., 5.33 Ma to present) eruptions lie above ANT-20 derived LAB depths in the range 35–70 km; the shallowest continental depths predicted by the present day seismic structure. This result is consistent with geodynamic expectations, since for a reasonable range of mantle temperature and hydration conditions, significant decompression melting is only expected in regions with LAB depth shallower than 80 km (Ball et al., 2021). Secondly, the minimum age since last eruption falls within the Miocene epoch for the remaining site, and here, LAB depth exceeds 70 km. The lack of more recent magmatism in this region indicates that the source of such magmatism has been removed over geological timescales. If this is the case, the LAB would have recovered to an equilibrium depth more representative of mean mantle conditions, thus further validating our model predictions.

To validate this hypothesis we use models of conductive cooling to determine the expected increase in LAB depth as a function of geological time. We adapt the thermal modeling approach of Richards, Hoggard, Crosby, et al. (2020), imposing an initial temperature condition that assumes a steady-state geotherm has been established by the time active magmatism ceases. We test a range of initial geotherms with 1358–1507°C mantle potential temperatures, 35–65 km initial LAB depths, and 10–40 km crustal thicknesses, based on seismically inferred values beneath the Antarctic magmatic provinces. In all models we assume an equilibrium plate thickness of 250 km and, based on calculated relationships between potential temperature and time since last eruption, we assume that initial thermal anomalies decay linearly to ambient temperatures (1333°C) over a 15 Myr period (see Text S5 in Supporting Information S1 for details). In order to compare the output of our conductive cooling models to the data, we tie the spatially binned eruption age values to a prediction of LAB depth and its uncertainty, calculated by taking the average and standard deviation of the depths within each bin (Figure 10e). The magmatic data are fully consistent with the post-magmatic lithospheric thickening models, suggesting that our seismically inferred LAB values are reliable.

To further investigate whether our data implies the existence of a monotonic relationship between LAB depth and minimum age since last eruption, we applied a statistical test. A Monte Carlo approach was employed to simulate the distribution of possible trends according to the uncertainty reported on each LAB depth-age data point, as follows. For each data point, d_i, initially located at $d_i = (a_i, z_i)$ in age-depth space, a random perturbation, $\Delta_i = (\alpha_i, \zeta_i)$, is added by drawing from a normal distribution with diagonal covariance scaled by the location-dependent
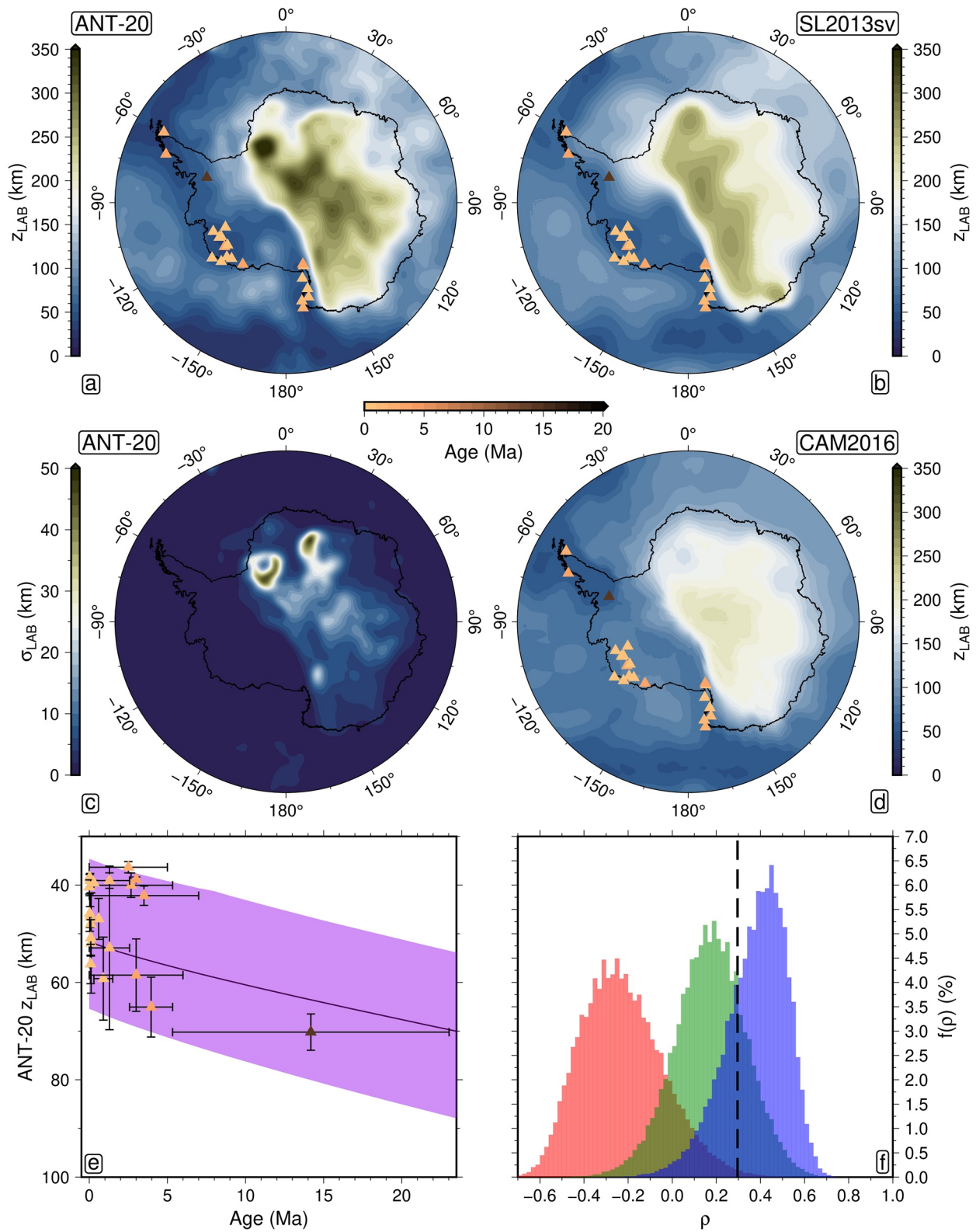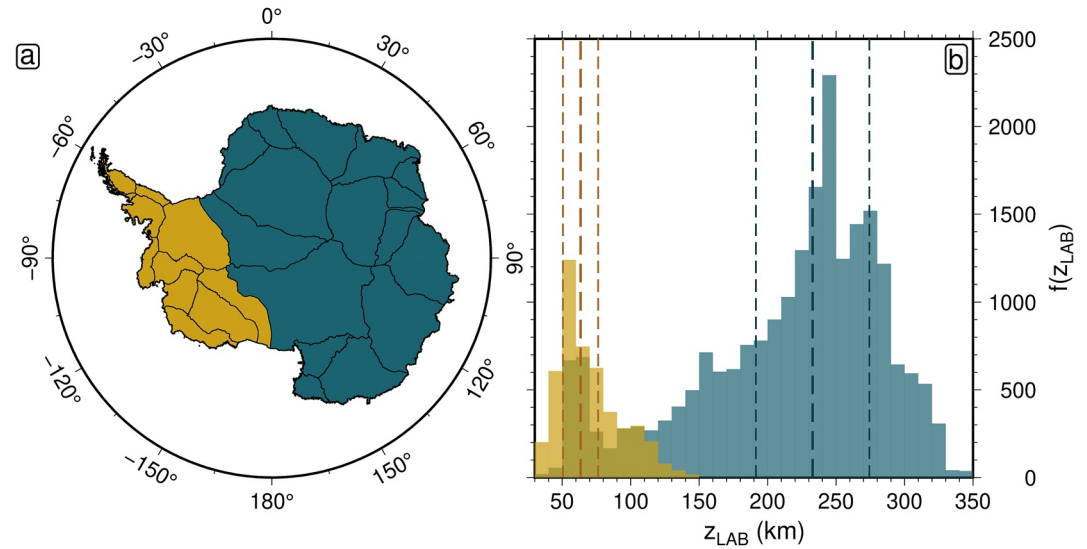
Figure 10.

Figure 11. Antarctic lithosphere-asthenosphere boundary (LAB) depth dichotomy. (a) Drainage network divides developed by the Goddard Ice Altimetry Group from ICESat data (Zwally et al., 2012). (b) Distribution of LAB depths beneath West Antarctic Ice Sheet and East Antarctic Ice Sheet (yellow and blue, respectively).

age and depth uncertainties. The resulting trend represents one possible combination of "true" age-depth values, and we calculate a corresponding Spearman's Rank correlation coefficient for this trend. This process is repeated until convergence, resulting in a distribution of possible correlation coefficients for each LAB depth model (Figure 10f). For us to associate a given coefficient with statistically significant evidence for the existence of a positive monotonic relationship between LAB depth and minimum age since last eruption at the 95% confidence level, it must exceed a value of $\rho = 0.296$.

We find that the ANT-20 and SL2013sv derived LAB models satisfy this test to at least the 1σ level, with coefficients $\rho = 0.17 \pm 0.16$ and $\rho = 0.38 \pm 0.14$ respectively. This result suggests that both models make reliable LAB depth predictions in the context of the geological record. It is unlikely however that the improved correlation offered by SL2013sv necessarily translates into this being a more realistic LAB model than ANT-20. This is because while, in each case, the spatial binning procedure was conducted over a 100 km radius, the true lateral resolution of SL2013sv is much poorer than ANT-20 over Antarctica. The resulting LAB model is therefore laterally smoother, reducing spatially binned LAB depth uncertainties and potentially improving the average trend observed in the age-depth data. It is unlikely that the LAB model derived from CAM2016 satisfies our test for statistical significance, owing to a coefficient, $\rho = -0.23 \pm 0.19$, such that less than 1% of possible age-depth trends contain a statistically significant positive gradient. This result may indicate that the LAB depth predictions of CAM2016 are less reliable in the Antarctic region than its counterparts. Nevertheless, there are several limitations on this analysis imposed by the small size of the magmatic data set, significant clustering of data points within age-depth space, and large age uncertainties on certain data points. These have a non-negligible impact on the calculated correlation coefficients, and so this analysis does not provide conclusive evidence for the reliability, or lack thereof, of any given seismologically derived LAB depth model.

4.3. Geothermal Heat Flow

In addition to calculating LAB depth, we constrain continental GHF by fitting a steady-state, laterally varying geotherm to our ensemble of three-dimensional temperature structures following the procedure laid out in

Figure 10. LAB depth variations beneath Antarctica. Mean (a) and standard deviation (b) lithosphere-asthenosphere boundary (LAB) depth derived from ANT-20, as estimated from depth to the 1200°C isotherm. ANT-20-derived LAB structure is compared to the predictions of Richards, Hoggard, White, et al. (2020) (c) and Priestley et al. (2018) (d), derived from the SL2013sv and CAM2016 seismic tomography models, respectively. The LAB depth models are overlain with the minimum age since last continental magmatic activity; the relationship between these two variables is shown for ANT-20 in (e). Histogram (f) displays the distribution of possible Spearman's Rank correlation coefficient values, ρ, between LAB depth and age for each LAB structure (CAM2016—red, ANT-20—green, SL2013sv—blue). Black dashed line = minimum value of ρ required for there to be a statistically significant increase in LAB depth with age at the 95% confidence level.

McKenzie et al. (2005). As in the construction of LAB depth estimates, anomalous temperatures associated with crustal bleeding were excised prior to interpolation of the geothermal profile back to 0°C at the basement depth. The Moho heat flux and mechanical boundary layer thickness are optimized based on the discrepancy between the modeled and V_S-derived geothermal profiles at each location, and the surface temperature gradient is utilized to calculate GHF (see Text S6 in Supporting Information S1). In constructing a modeled geothermal profile at a given location, it is necessary to account for lateral variations in crustal thickness, as well as depth variations in radiogenic heat production and conductivity. As previously, the crustal thickness grid associated with the tomography model was used, to ensure self-consistency. For crustal heat production, we assume a value of $H^*_{ocean} = 0.0$ µW m^{-3} within the ocean, distributed uniformly throughout the crustal layer. Within the continent, we divide the crust into two layers of equal depth. We assume values of $H^*_{cont} = 1.0$ µW m^{-3} and $H^*_{cont} = 0.3$ µW m^{-3} in the upper and lower crustal layers, respectively. This two-layer continental heat production parameterization is compatible with globally averaged values obtained from the comprehensive crustal geochemical analysis of Sammon et al. (2022), and is preferred for two main reasons. Firstly, the simplicity of the parameterization avoids assuming more detailed knowledge of the three-dimensional distribution of heat producing elements within the crust than is currently available. Secondly, it reduces the sensitivity of the crustal radiogenic heat content to regions of anomalously thick crust, as compared to assuming a single crustal layer of constant heat production (although this sensitivity remains non-negligible). Mantle and oceanic crust conductivity are calculated according to the temperature- and pressure-dependent parameterization of Korenaga and Korenaga (2016). In the continent, crustal conductivity is set to a constant value of $k_{crust} = 2.5$ W m^{-1} K^{-1}. These assumptions simplify the true lateral and depth dependence of heat production and conductivity within the continental crust, which are expected to vary within the range $H^*_{cont} \sim \{0.0, 3.0\}$ µW m^{-3} and $k_{crust} \sim \{1.0, 4.0\}$ W m^{-1} K^{-1} (Jennings et al., 2019). Investigating the effect of the variation of these two parameters on the resulting heat flow is beyond the scope of this study. However, a close fit between theoretically and V_S-derived geothermal profiles calculated using our simple parameter assumptions suggests the dominant control on GHF estimates is the seismically inferred thermal structure rather than the chosen crustal parameterization. With this in mind, we utilize a laterally varying mantle potential temperature during the fitting process, estimated according to the average V_S-derived temperature beneath the base of thermal boundary layer.

Resulting estimates of spatially averaged continental energy transfer rates are 44 ± 2 mW m^{-2} into the base of the EAIS, and 76 ± 7 mW m^{-2} into the base of the WAIS, where the 'uncertainty' in this case relates purely to lateral variations in mean heat flow (Figure 12). Both sides of the continent contain GHF variations in excess of 40 mW m^{-2}. However, East Antarctica is less heterogeneous, with over 75% of its area characterized by GHF in the region 40–50 mW m^{-2}. By contrast, the WAIS is underlain by bedrock feeding it anywhere from approximately 50–100 mW m^{-2} of geothermal energy, with heat flow unevenly distributed across the region. The connection of two long-wavelength (exceeding 10,000 km) thermal anomalies into a single anomaly, observed in the viscosity structure at 75 km depth (Figure 8a), can also be seen here to stretch from the RS through MBL and up to the AP, before stretching offshore toward the South Scotia ridge. Most of this anomaly is located within West Antarctica, with the exception of its eastern edge within MBL and Victoria Land. The presence of this anomaly, combined with shorter-wavelength (~1,000–10,000 km) cold anomalies observed in MBL and Ellsworth Land, together make up a highly heterogeneous West Antarctic GHF structure.

5. Discussion

In the text to follow, we show how the results presented in this study build upon existing evidence of strong lateral heterogeneity in Earth's internal thermomechanical structure beneath Antarctica, leading to spatially variable LAB depth and GHF. First, the dichotomy between West and East Antarctic thermomechanical structure is discussed, along with implications for ice sheet stability. Secondly, our GHF predictions are compared to those of a recent study by W. Shen et al., 2020. Thirdly, we discuss how a consideration of physical forcing timescale can be used to reconcile observations and model predictions of mantle viscosity. Finally, we summarize the most significant sources of remaining uncertainty in quantifying mantle structure.

5.1. West and East Antarctic Mantle Structure

We find evidence that steady-state diffusion creep viscosities reach a lower threshold of $\eta \sim 10^{19}$ Pa s throughout the shallow mantle (150–350 km) beneath West Antarctica. Uncertainty in asthenospheric viscosity structure
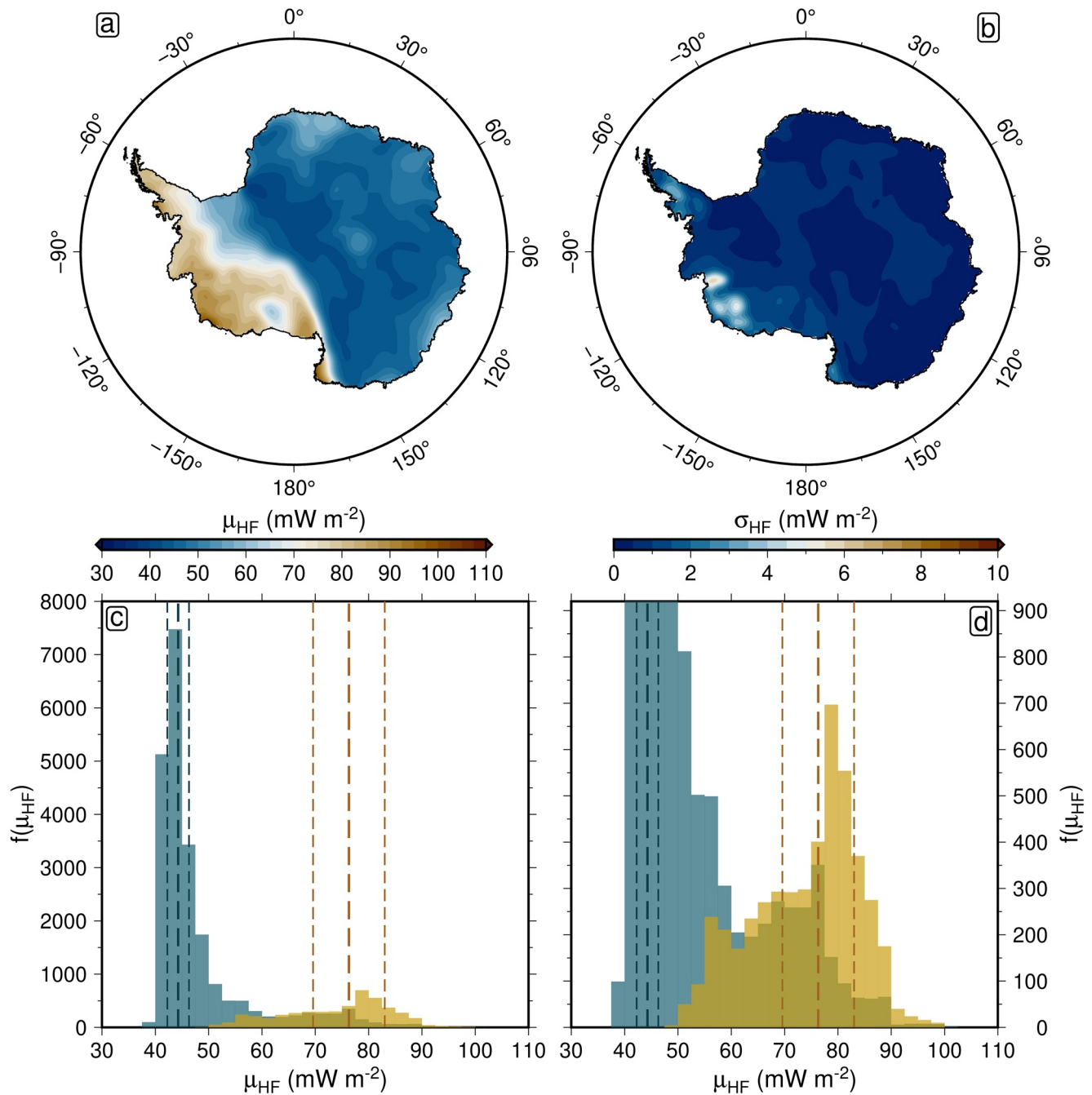
Figure 12. Geothermal heat flow (GHF) variations across Antarctica. Mean (a) and standard deviation (b) GHF derived from ANT-20. Estimated by fitting steady-state geotherms to temperatures inferred from an ensemble of anelasticity models. Distribution of GHF into base of West Antarctic and East Antarctic Ice Sheets ((c); yellow and blue, respectively). Zoom-in of distributions (d). Regional separation calculated according to the pattern of Antarctic drainage systems, see Figure 11. Thick dashed lines show the median of each distribution. Thin dashed lines are located one median absolute deviation away from the median of each distribution.

is found to be within one order of magnitude across the full depth range 0–400 km of study, and increases with depth. Low-viscosity anomalies observed within the mantle viscosity structure correspond with spatial patterns in LAB depth and GHF structure, whereby negative and positive anomalies are observed, respectively, which is to be expected given the self-consistent framework within which each of these parameters is estimated. For example, viscosities of $10^{19.5\pm0.3}$ Pa s present at 150 km depth in western MBL toward the ASE are associated with thin LAB depths (30–50 km) and elevated GHF (85–95 mW m^{-2}), where quoted ranges represent spatial variability within this region. The inference that such low viscosities beneath the WAIS are caused by a thermal

anomaly is consistent with the geological record of Cenozoic magmatism (Ball et al., 2021; DIGIS Team, 2021). The combination of high GHF, thin lithosphere and low viscosity points to a highly dynamic ice-sheet–solid-Earth interaction in regions including western MBL, the eastern Ross Embayment, and the AP. Large fluxes of thermal energy into the base of the ice sheet in these regions will likely enhance basal melting, reducing friction and increasing glacial sliding rates (Burton-Johnson et al., 2020; W. Shen et al., 2020). On the other hand, a thin lithosphere and low viscosity asthenosphere encourage rapid bedrock uplift and may help to stabilize and reduce grounding line retreat (Gomez et al., 2010).

We infer much higher viscosities beneath East Antarctica, with much of this side of the continent exhibiting LAB depths in excess of 150 km. LAB depth is on average 63 ± 13 km beneath West Antarctica, significantly lower than the 233 ± 41 km estimated beneath East Antarctica. The combination of thick lithosphere and moderate GHF suggest a less dynamic interaction between the ice sheet and the solid Earth. While there is evidence for a low viscosity anomaly of order $10^{19.5 \pm 0.7}$ Pa s at a depth of 350 km in Wilkes Land (WL), beneath the Aurora Subglacial Basin (ASB), it does not penetrate up to shallower depths of 150 km, and therefore the influence of this anomaly on GIA rates is expected to be greatly reduced compared with the shallow anomalies beneath much of the WAIS. Nonetheless, this region is associated with GHF of 45–60 mW m^{-2}. The upper end of this range is among the highest heat flow values estimated across East Antarctica (with the exception of the eastern Ross Embayment). Ice velocity and mass discharge rates across WL are accelerating in response to warming temperatures in the Southern Ocean (Noble et al., 2020). The marine-based ASB is positioned on a reverse bed slope, and may be susceptible to rapid ice mass loss (Q. Shen et al., 2018). Elevated GHF could therefore enhance this topographic instability by encouraging ice flow across the grounding line.

5.2. Comparing Predictions of Geothermal Heat Flow

Our estimates of GHF exhibit similar spatial structure to that estimated by W. Shen et al. (2020), albeit with less short-wavelength variation. The study conducted by W. Shen et al. (2020) calibrated an empirical mapping between GHF and V_S using the observed relationship across the continental United States between interpolated heat flow measurements and V_S at 80 km depth from a regional tomographic model. When the US-calibrated mapping is applied to their Antarctic velocity model, resulting GHF ranges from 40 to 90 mW m^{-2}. Our results agree on the presence of anomalously high heat flow (approximately 80 mW m^{-2}) stretching from the RS to the AP, avoiding the coast between the Ross and Amundsen Sea. The most obvious discrepancy between the two structures is the presence of a high heat flow anomaly in our study, situated within the footprint of the Gamburtsev Mountain Range. The amplitude of this anomaly is 15% above the East Antarctic average. The reliability of this particular prediction should be doubted, because while the geological origin of the Gamburtsev Mountains is not well known, the potential for it to be caused by a mantle plume would imply thin lithosphere in this region. This is not corroborated by our LAB depth model, or those of Priestley et al. (2018) or Richards, Hoggard, White, et al. (2020) (Figures 10a, 10b, and 10d). We find that the spatial pattern of elevated GHF coincides with anomalously thick (~60 km) crust found in ANT-20. Since the total crustal radiogenic heat content in a particular region is proportional to crustal thickness in our parameterization, thick crust steepens the geothermal temperature gradient and therefore increases the inferred GHF. We therefore hypothesize that the GHF anomaly arises from a combination of two factors. First, a discrepancy between the assumed and true crustal thickness in this region. Secondly, a discrepancy between the assumed and true radiogenic heat production. In addition, anomalously low seismic velocities in the uppermost mantle beneath the Gamburtsev Mountains suggest the presence of a compositional anomaly (W. Shen et al., 2018). Although our geotherm fitting methodology is designed to mitigate the impact of locally unphysical temperature estimates that would arise from such an anomaly, it may still reduce our ability to accurately infer heat flow from the seismic velocity structure of this region.

5.3. Reconciling Observations and Predictions of Mantle Viscosity

The AP and ASE are regions of special interest to the cryosphere and sea level communities, since they are currently experiencing significant ice mass loss and could hold important clues for determining the future stability of the wider WAIS. The northern AP is undergoing rapid atmospheric warming, which has increased surface melt rates and contributed to ice-shelf collapse (Davies et al., 2014). Recently, three major ice shelves along the AP were lost completely in the space of less than a decade, when Prince Gustav (1993–1995), Larsen A (1995), and Larsen B (2002) collapsed (Nield et al., 2014). As a result, tributary glaciers flowing

Table 2
Antarctic Upper Mantle Viscosity Estimates Derived From Geodetic Observations

Study	Location	Ice loading history	Observation period	$\log_{10}\eta$ estimate
B18—Barletta et al. (2018)	ASE	Retreat from 1900 to 2014	2002–2014	18.4–19.4
S21—Samrat et al. (2021)	AP	Retreat from 1999 to 2020	1999–2020	17.5–19.0
I11—Ivins et al. (2011)	AP	Overall retreat from LGM to present, modern phases of advance and retreat	2003–2009	19.3–20.0
W15—Wolstencroft et al. (2015)	AP	Retreat from LGM to present	2009–2013	20.0–20.5

Note. Each study assumed a particular ice loading history to estimate the reported viscosity values, a summary of which is reported here. The observation period represents the timeframe that best represents when data was collected. ASE, Amundsen Sea Embayment; AP, Antarctic Peninsula; LGM, Last Glacial Maximum.

from the AP plateau are accelerating and thinning (Cook & Vaughan, 2010). While the ASE accounts for less than 4% the area of the AIS, the marine-grounded portion of the WAIS in this region accounts for a quarter of the global present-day cryospheric contribution to GMSL rise (Barletta et al., 2018). Accelerating ice flow and rapidly retreating grounding lines have been observed at both the Pine Island and Thwaites glaciers. This recent change, combined with the reverse bed slope beneath both glaciers, suggests that they are vulnerable to catastrophic collapse (Barletta et al., 2018). Accurately capturing solid Earth structure beneath the AP and ASE is therefore of particular importance, since future ice retreat in these regions is especially sensitive to viscoelastic bedrock uplift rates, which—if sufficiently rapid—may help to stabilize grounding lines. Moreover, the present-day GIA rate is strongly affected by inferred thermomechanical structure beneath sites of recent ice loss, and must be accurately calculated in order to reliably estimate ongoing ice mass change from satellite gravity data.

Bedrock deformation rates observed by GPS can be used to shed light on solid Earth structure, since they depend on the rheology of the underlying mantle. Typically, observed deformation rates are combined with an estimate of the local ice sheet loading history and a Maxwell viscoelastic solid Earth model, to infer a viscosity consistent with the applied constraints. For example, the recent study of Barletta et al. (2018) provides a geodetic analysis of bedrock deformation rates across the ASE using six local GPS stations. Barletta et al. (2018) estimated upper mantle viscosities in the range $\log_{10}\eta = 18.4$ to $\log_{10}\eta = 19.4$. These extremely low viscosities imply Maxwell relaxation times of order 1–10 years, meaning that the topographic response to deglaciation following the LGM; ~21 ka, would have already decayed away in this region. Since models of GIA in response to modern-day ice mass loss typically assume upper mantle viscosities of $\eta \sim 10^{20}$ Pa s, a substantial upward revision of viscoelastic uplift rates would be required in the ASE, implying that local ice mass loss has been underestimated by ~10% in previous GRACE-based assessments. Conversely, the vulnerability of the WAIS in this region to catastrophic collapse would potentially be reduced by the faster GIA response rates, since rapid grounding line uplift might help to stabilize the ice sheet (Gomez et al., 2010).

A question which arises naturally is therefore: How well do geodetically constrained estimates of mantle viscosity beneath the AP and ASE agree with our inferred thermomechanical structure? To determine the answer, we have compiled the findings of four geodetic studies, summarized in Table 2. The viscosity range obtained from each study is based on the range of plausible upper mantle viscosities able to fit the geodetically observed horizontal and vertical deformation rates. We extracted our own corresponding viscosity estimates based on the ensemble of predicted diffusion creep viscosities beneath the AP and ASE. At each location a single upper mantle viscosity value was obtained for each anelasticity model by averaging predicted values over a three-dimensional region with a 50 km surface-equivalent lateral radius, and variable depth extent. For the AP, we used 125–175 and 125–250 km depth windows, with corresponding ranges of 150–175 and 150–250 km applied beneath the ASE. The shallow bounds of these depth windows are selected to coincide with local viscosity minima. Two different deep bounds are used in each region because stress changes associated with ice age loads will penetrate to greater depths than those induced by smaller-scale modern ice mass changes, affecting the appropriate depth window to use when comparing tomographic viscosity estimates with geodetic inferences (Blank et al., 2021; see Text S7 in Supporting Information S1 for further detail on viscosity sampling and averaging process). The resulting probability density distributions of inferred viscosity are shown in comparison to their geodetically derived counterparts in Figure 13. A reasonable agreement is found in the case of I11, when looking at the shallow depth window, and in the case of W15, when looking at the deeper depth window. A poor agreement is found for B18, with little
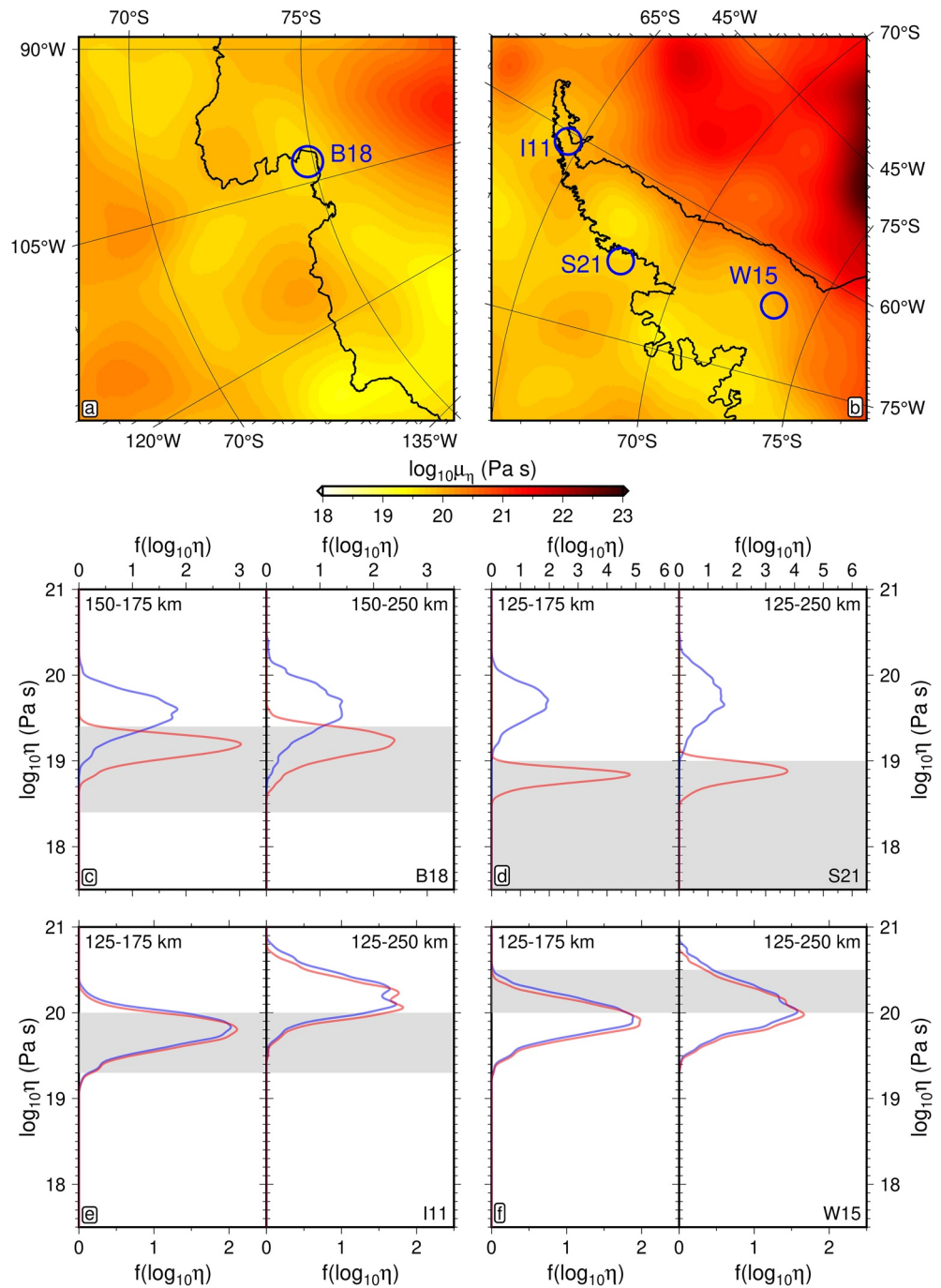
Figure 13. Comparison of upper mantle viscosity estimates based on GPS and tomography. Mean viscosity structure at 150 km depth beneath the Amundsen Sea Embayment and Antarctic Peninsula (panels (a) and (b), respectively), with lateral sampling regions overlain in blue. Probability density distribution of inferred diffusion creep steady-state (blue) and time-dependent (red) viscosity compared to geodetic estimates (gray) B18, S21, I11, and W15 (panels (c), (d), (e), and (f), respectively).

overlap between the geodetically and tomographically inferred viscosities. An extremely poor agreement is found for S21, where there is no overlap at all.

We note that the extent to which the geodetic analyses agree with our inferred steady-state viscosities appears to depend on the modeled ice loading timescale. The rheological response of a viscoelastic body to a given change in the stress field (i.e., loading) depends on the timescale over which it occurs (Lau & Holtzman, 2019).

Frequency-dependent anelasticity parameterisations can be used to calculate apparent viscosities at any forcing timescale, and the values we have reported so far in this study are for steady-state deformation, representing the theoretical limit of an infinite forcing timescale. On the other hand, geodetically derived viscosity estimates are relevant to the timescales corresponding with the forcing processes in operation. The shorter the forcing times-cale, the larger the discrepancy between steady state and apparent viscosity, as the deformation behavior tends toward the elastic regime. This relationship between forcing timescale and apparent viscosity could therefore be responsible for the discrepancy we found.

To quantify the potential impact of this time-dependent rheology on our tomographically inferred viscosity esti-mates, we modeled the deformation rate one would expect to observe given the ice loading histories assumed in each of the geodetic studies, combined with our estimates of steady-state viscosity and the transient rheological parameterization YT16. This approach was formulated based on the work of Lau et al. (2021). By performing a grid search over a range of Maxwell viscosity values, we were therefore able to invert for the apparent Maxwell viscosity value providing the best fit to the synthetic deformation rate observations. This apparent viscosity is dependent on the amplitude and timescale of forcing processes triggered by the ice loading history, or in other words, its frequency content. When we compare our tomographically inferred time-dependent viscosities to those derived from geodetic observations, we find a much better agreement than before (Figure 13). For example, our predicted time-dependent viscosity distributions lie almost entirely within the range of possible values predicted by B18 and S21, when looking at the shallow depth range. The observation that the shallow depth range provides the best fit to the geodetic observations for the short timescale ice loading histories lends further support to the hypothesis that time-dependent behavior is at play. This is because one would expect GPS observations to be sensitive to the viscosity within the portion of the mantle activated by the modeled loading history. In the case of short timescale and lower magnitude loading, dissipation of stress may only have occurred within the shallow upper mantle, thus making the observed deformation rates sensitive only to these depths.

5.4. Remaining Uncertainties

Despite making major progress in understanding the thermomechanical structure of the Antarctic upper mantle, our work highlights outstanding challenges that limit our ability to utilize seismological data to understand solid Earth structure and its relationship with cryospheric evolution. A lack of geophysical data sets containing infor-mation about the deep mantle restricts the precision with which we can estimate pressure-dependent behavior. For example, the uncertainty present in our estimate of activation volume remains high after calibrating the anelasticity parameterization, since the majority of our data relates to the shallowest 125 km of the mantle. This leads to increasing uncertainty in thermomechanical structure with depth. In addition, the microphysical process or processes responsible for the onset of anelasticity is subject to significant debate, and this translates into competing methods for modeling anelastic effects (Faul & Jackson, 2007; Yamauchi & Takei, 2016). As a result, heavily discrepant predictions may be made depending on the choice of physical model (Ivins et al., 2021). With this in mind, our inverse calibration has been designed structurally to work with any choice of anelasticity param-eterization. Further uncertainty relates to the particular viscous creep mechanism dominating Antarctic upper mantle rheology on timescales relevant to the modeling of geodynamic processes (Lau & Holtzman, 2019). If dislocation creep is the dominant mechanism, the diffusion creep viscosities predicted here will overestimate true steady-state values. Nevertheless, the temperatures we predict would remain robust (see Text S8 in Supporting Information S1 for temperature structures), being reliant only on the correct modeling of diffusionally controlled anelastic processes at seismic frequency. We also wish to emphasize that this independently constrained thermal structure significantly reduces uncertainty in dislocation creep viscosity, which, like its diffusional counterpart, is strongly temperature dependent.

With respect to secondary structures calculated using our estimates of three-dimensional temperature variations, namely LAB depth and GHF, a few specific challenges are yet to be addressed. First, vertical seismic resolution limits make it difficult to infer LAB depth variations smaller than ~15 km. Secondly, due largely to the downward bleeding of slow shear-wave velocities associated with discrepancies between the modeled and true crustal struc-ture, seismically inferred temperature structure becomes unreliable close to the Moho. This means that interpola-tion must be used to estimate shallow temperature structure. While this is not expected to influence our estimates of LAB depth, since the 1200°C isotherm is sufficiently deep, it will have an effect on our estimates of heat flow, which are proportional to the surface geothermal gradient. A lack of exposed outcrops where Antarctic GHF can

be measured makes it difficult to ground truth our geophysical predictions and refine the model. In addition, we currently have a poor understanding of the range, depth variation, and lateral variation in Antarctic crustal heat production and conductivity. Since both of these parameters must be assumed to fit a steady-state geothermal profile to our temperature-depth data, our GHF estimates are directly affected by this limitation. To address this, complementary geophysical methods should be used to gain insight into crustal heat production and conductivity structure, allowing for further refinement of GHF models.

6. Conclusions

A probabilistic approach to the calibration of experimental parameterisations of anelasticity has been developed to provide a self-consistent mapping between three-dimensional seismic tomography data and models of thermomechanical structure. By making use of a physical model designed to account for frequency dependence in the mantle stress-strain relationship, it is possible to translate experimentally constrained microphysical behavior into predictions of macroscopic variables including temperature, viscosity and density, as a function of shear-wave velocity. We calibrate the viscoelastic parameters with a suite of regional geophysical data constraints, reducing the discrepancy in physical predictions offered by different seismic tomography models, and ensuring a set of outputs compatible with well-constrained mantle properties. We provide an implementation of the inverse theory, using the Globally Adaptive Scaling Within Adaptive Metropolis (GASWAM) adaptation of the Metropolis-Hastings algorithm to allow ideal sampling efficiency and thus make the inverse problem tractable. We have shown it is possible to utilize a small subset (in our case, 0.5%) of the overall posterior data set to propagate shear-wave velocity into accurate estimates of thermomechanical structure and its uncertainty, which ensures computational viability. By probing the model covariance structure, this uncertainty is significantly reduced as compared to treating parameters independently (viscosity uncertainty reduced by 4–5 orders of magnitude at 150 km depth).

Dramatic differences in viscosity structure, LAB depth and GHF are predicted between East and West Antarctica, in accordance with other studies (Austermann et al., 2021; Barletta et al., 2018; Priestley et al., 2018; Richards, Hoggard, White, et al., 2020; W. Shen et al., 2020). We find evidence for mostly thick lithosphere, high viscosity asthenosphere, and uniformly low GHF beneath the EAIS. Shallow LAB depths and high GHF coincide with regions characterized by the presence of low viscosity anomalies, such as in western MBL where we observe values 30–50 km, 85–95 mW m^{-2}, and $\eta = 10^{19.5\pm0.3}$ Pa s, respectively. This combination of thermomechanical properties is consistent with the geological record of regional Plio-Pleistocene magmatism (Ball et al., 2021; DIGIS Team, 2021), and indicates that western MBL, along with the eastern Ross Embayment and AP, may be amongst the most dynamic in response to climate and ocean forcing. High GHF may significantly increase the flow of ice toward the continental perimeter, whereas the presence of low viscosities and thin lithosphere suggest much faster bedrock uplift rates than a one-dimensional average rheology, potentially providing a stabilizing effect on the grounding line (Gomez et al., 2010).

The outputs presented in this study may be used to refine our understanding of ice sheet stability in Antarctica. Our models of density structure can be used to improve time-dependent models of convectively supported surface topography, enabling correction of palaeo sea level markers used to inform ice sheet history. Our self-consistently determined viscosity and LAB depth structures, that also constrain time-dependent rheological variations, can be applied to three-dimensional GIA studies, where uplift rates are intimately tied to rheological structure. These high-resolution estimates of thermomechanical structure will be useful in constraining bedrock uplift rate across the continent, in turn altering corrections needed to produce gravimetric and altimetric estimates of present-day ice mass loss rates. Our seismically inferred maps of GHF can be incorporated in new ice-sheet modeling studies, where basal sliding rates are highly sensitive to the amount of thermal energy provided from below. As a result, we suggest that our new methodology for estimating solid Earth inputs and their associated uncertainties may enable accurate probabilistic assessment of ice sheet stability scenarios and projections of future sea level rise.

Appendix A: Work Done and Complex Compliance

In the following analysis, we make use of a complex representation of stress, σ^*, and strain, ε^*. The real stress, σ, and strain, ε, that would be measured can be calculated by projecting their complex representations onto a constant axis of choice. Here, we take the imaginary projection where $\sigma = \Im(\sigma^*)$ and $\varepsilon = \Im(\varepsilon^*)$.

If we consider a single Fourier component of the stress applied to the linear viscoelastic body, this can be written as

$$\sigma(t) = \mathfrak{I}(\sigma_0 \exp(-i\omega t)) = -\sigma_0 \sin(\omega t). \tag{A1}$$

The complex compliance, $J^*(\omega)$, provides us with the connection between applied stress and strain response

$$\varepsilon^*(t) = J^*(\omega)\sigma^*(t), \tag{A2}$$

$$\Rightarrow \varepsilon(t) = \mathfrak{I}(\sigma_0(J_1 + iJ_2)\exp(-i\omega t)), \tag{A3}$$

$$\Rightarrow \varepsilon(t) = -\sigma_0 J_1 \sin(\omega t) + \sigma_0 J_2 \cos(\omega t). \tag{A4}$$

The work done during each oscillatory cycle of applied stress is given by the integral

$$W = \oint \sigma \, d\varepsilon = \int_0^{2\pi/\omega} \sigma \frac{d\varepsilon}{dt} \, dt, \tag{A5}$$

and thus takes the form

$$W = \int_0^{2\pi/\omega} -\sigma_0 \sin(\omega t) \frac{d}{dt} (-\sigma_0 J_1 \sin(\omega t) + \sigma_0 J_2 \cos(\omega t)) \, dt, \tag{A6}$$

$$\Rightarrow W = \omega \sigma_0^2 \int_0^{2\pi/\omega} (J_1 \sin(\omega t)\cos(\omega t) + J_2 \sin(\omega t)\sin(\omega t)) \, dt, \tag{A7}$$

$$\Rightarrow W = \pi\sigma_0^2 (0 \cdot J_1 + 1 \cdot J_2) = \pi\sigma_0^2 J_2. \tag{A8}$$

We can therefore see that the out-of-phase compliance term, J_2, is responsible for energy dissipation, while the in-phase compliance term, J_1, is responsible for energy storage.

An equivalent representation for Equation A4 can be found using a double-angle trigonometric expansion as follows

$$\varepsilon(t) = -\varepsilon_0 \sin(\omega t + \phi), \tag{A9}$$

$$\Rightarrow \varepsilon(t) = -\varepsilon_0 \cos(\phi)\sin(\omega t) - \varepsilon_0 \sin(\phi)\cos(\omega t). \tag{A10}$$

By comparison with Equation A4, we can establish some useful relations between the complex compliance terms and the phase of the strain response as follows.

$$J_1 = \cos(\phi)\varepsilon_0/\sigma_0; \tag{A11}$$

$$J_2 = -\sin(\phi)\varepsilon_0/\sigma_0; \tag{A12}$$

$$\tan(\phi) = -J_2/J_1. \tag{A13}$$

This shows us that the superposition of the elastic and viscous response introduces a phase delay, ϕ, between stress and strain, and is controlled by the ratio between loss and storage terms of the complex compliance.

Appendix B: Prior Estimation

In order to specify the Gaussian prior distribution, a suitable estimate of each parameter, μ_i, and its uncertainty, s_i, is required. The priors on the elastic sector of the parameter space, $\mathcal{X}_{\text{elastic}} = \{\mu_0, \partial\mu/\partial T, \partial\mu/\partial P\}$, were calculated by sampling a range of thermochemical states, $S = \{X, P, T\}$, where X is pyrolitic composition defined in terms of the proportion of harzburgite to basalt. A database containing the dependence of elastic shear modulus on S was utilized to build a prior picture of $\mathcal{X}_{\text{elastic}}$. This database was constructed using the software Perple_X according to the method laid out by Cobden et al. (2008), using the compilation of thermodynamic parameters of Stixrude and Lithgow-Bertelloni (2011). Activation energy (E_A), activation volume (V_A) and the solidus gradient

$(\partial T_S/\partial z)$ were estimated by summarizing literature reported values (Hirth & Kohlstedt, 2004; Jain et al., 2019). Reference viscosity (η_0) was estimated using the following equation,

$$\eta_0 = \frac{d_0^p}{A}\exp\left(\frac{E_A + P_0 V_A}{RT_0}\right), \tag{B1}$$

where the reference thermodynamic state $(P_0, T_0) = (1.5\text{ GPa}, 1200°C)$, $d_0 = 1$ mm is the reference grain size, p its exponent, and A a scaling coefficient. By sampling A, p, E_A, and V_A over suitable ranges retrieved from the literature (Hirth & Kohlstedt, 2004; Jain et al., 2019), a summary of η_0 could be established.

Appendix C: Adaptive Metropolis Algorithms

The Metropolis-Hastings algorithm relies on a proposal distribution in order to generate transitions between a current state, $\mathcal{X}$, and a proposal state, $\mathcal{Y}$. The precise form and magnitude of the proposal distribution is a key component of the number of trials required to achieve ergodicity: the convergence of our discrete set of samples onto the underlying continuous posterior distribution.

Adaptive Metropolis algorithms are intended to improve the efficiency of this sampling process. Haario et al. (2001) serves as a good reference point for the implementation of such an algorithm. It utilizes the condition found by Gelman et al. (1997) that for a Metropolis algorithm on $\mathbb{R}^d$, the proposal is optimally scaled when the proposal state is generated according to

$$\mathcal{Y}^{n+1} \sim \mathcal{N}\left(\mathcal{X}^n, \mathbf{\Sigma}^{\text{proposal}}\right), \tag{C1}$$

$$\mathbf{\Sigma}^{\text{proposal}} = \gamma\mathbf{\Sigma}^{\mathcal{X}}, \tag{C2}$$

where $\mathbf{\Sigma}^{\mathcal{X}}$ is the posterior covariance matrix, and $\gamma = \frac{2.38^2}{d}$ is the scaling coefficient. The condition implies that the ideal proposal covariance matrix is a scalar multiple of the target posterior covariance. Since the posterior is the object we are attempting to access via our sampling procedure, employing a suitable proposal is challenging. To address this issue, Haario et al. (2001) use an unbiased estimate of the target posterior covariance, which can be calculated empirically based on the evolving chain of generated samples. On the nth trial, where $n - 1$ samples have been generated so far, the unbiased estimate of the posterior covariance is

$$\overline{\mathbf{\Sigma}}^{\mathcal{X}} = \frac{1}{n-2}\sum_{i=1}^{n-1}\left(\mathcal{X}^i - \overline{\mathcal{X}}\right)\left(\mathcal{X}^i - \overline{\mathcal{X}}\right)^T, \tag{C3}$$

where $\overline{\mathcal{X}} = \frac{1}{n-1}\sum_{i=1}^{n-1}\mathcal{X}_i$. The prefix "adaptive" therefore comes from the iterative adaptation of the proposal covariance matrix. For multi-dimensional parameter spaces ($d > 1$), $\overline{\mathbf{\Sigma}}^{\mathcal{X}}$ may take a considerable number of trials to resemble the true posterior covariance, $\mathbf{\Sigma}^{\mathcal{X}}$, however, it should provide better performance than a fixed proposal setup. Substituting $\mathcal{X}$ in Equation C3 for only the subset of trial models that were accepted, $\mathcal{Z}$, may offer more efficient convergence toward the posterior covariance. This approach is known as the "greedy start" procedure, and we make use of it in this study (Figure 3). It should also be noted that since Equation C3 relies on the history of all preceding trials, the chain of samples is no longer Markovian. However, it has been proven that ergodicity still holds for adaptive algorithms given some loose assumptions on the posterior (see Haario et al., 2001 for details).

Implementation of the Adaptive Metropolis algorithm shown above is theoretically easy, however the optimal scaling factor, $\gamma = \frac{2.38^3}{d}$, does not work in practice if there are significant correlations between the parameters in the model. In this case, the solution is to also update γ adaptively. The Global Adaptive Scaling Within Adaptive Metropolis (GASWAM) scheme employs this technique to estimate a suitable proposal covariance matrix

$$\mathbf{\Sigma}^{\text{proposal}} = \gamma^n\left(\overline{\mathbf{\Sigma}}^{\mathcal{X}} + \epsilon\mathbf{1}\right), \tag{C4}$$

$$\gamma^n = \gamma^{n-1} + \eta^n\left(a^{n-1} - a^*\right). \tag{C5}$$

Here, the scaling factor to be used for the nth trial, γ^n, is updated by a factor proportional to the difference between the current and ideal acceptance ratios, a^{n-1} and a^*, respectively. The function, $\eta^n = n^{-1/2}$, is used to ensure adaptation decays in size as the simulation progresses. The presence of the constant $\epsilon > 0$ ensures ergodicity, and is chosen to be negligibly small compared to the size of the proposal covariance matrix. This algorithm can be employed after some fixed number of trials — long enough to provide a suitable first estimate of Σ^{proposal} — and initiated with the traditional Adaptive Metropolis scaling factor $\gamma^0 = 2.38^2/d$.

Appendix D: Approximating the Relationship Between Elastic Shear-Wave Velocity and Temperature

The linear region of the $V_S(T)$ relationship in a given depth bin is well-approximated by assuming an elastic response at fixed pressure. Consider the 50–75 km depth bin (Figure 4a, blue circles), and let us define a reference shear modulus relevant to this depth slice as follows

$$\mu_0^* = \mu_0 + \partial\mu/\partial P(P - P_0), \tag{D1}$$

where $P = 2.1$ GPa is the corresponding pressure value. The $V_S(T)$ relationship can be expressed as

$$V_S(T) = \sqrt{\frac{\mu(T)}{\rho(T)}}, \tag{D2}$$

and therefore, in terms of our renormalized shear modulus,

$$V_S(T) = \sqrt{\frac{\mu_0^* + \partial\mu/\partial T(T - T_0)}{\rho(T)}}. \tag{D3}$$

Since density is only expected to vary by 2% over the temperature range covered by the 50–75 km depth bin of the plate V_S data, and even less so for the other two depth bins, we ignore its variation for the sake of simplicity here. Consider the numerator, $\zeta = \sqrt{\mu_0^* + \partial\mu/\partial T(T - T_0)}$, of Equation D3. This may be rewritten in the form

$$\frac{\zeta}{\sqrt{\mu_0^*}} = \left(1 + \frac{\partial\mu/\partial T}{\mu_0^*}\Delta T\right)^{\frac{1}{2}}, \tag{D4}$$

Assessing the magnitude of each term on the right hand side of this equation, $|\mu_0| \sim 10^2$ GPa and $|(\partial\mu/\partial T)\Delta T| \sim 10^1$ GPa, we find it is possible to perform a binomial expansion since the ratio

$$|\frac{\partial\mu/\partial T}{\mu_0^*}\Delta T| < 1. \tag{D5}$$

Therefore, expanding the square-root, we have

$$\zeta = \sqrt{\mu_0}\left(1 + \frac{1}{2}\frac{\partial\mu/\partial T}{\mu_0}\Delta T + \mathcal{O}\big((\Delta T)^2\big)\right). \tag{D6}$$

Ignoring terms of second-order and above, we can deduce the form of the linear relationship $\zeta(T)$ as follows

$$\zeta \approx \sqrt{\mu_0}\left(1 + \frac{1}{2}\frac{\partial\mu/\partial T}{\mu_0}\Delta T\right), \tag{D7}$$

and thus in terms of shear-wave velocity we have

$$V_S(T) = \sqrt{\frac{\mu_0}{\rho}} + \frac{1}{2}\frac{\partial\mu/\partial T}{\mu_0\rho}\Delta T. \tag{D8}$$

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

Figures were prepared using Generic Mapping Tools software. Model outputs are provided in an OSF online repository (https://doi.org/10.17605/OSF.IO/F4NTR). Code used to prepare data, perform the inversion, and analyse the inversion output is provided in a Zenodo repository (https://doi.org/10.5281/zenodo.7930907). Latest development version of the inversion code is provided in a GitHub repository (https://github.com/JamesHazzard/BANCAL22).

References

An, M., Wiens, D. A., Zhao, Y., Feng, M., Nyblade, A., Kanao, M., et al. (2015). Temperature, lithosphere-asthenosphere boundary, and heat flux beneath the Antarctic Plate inferred from seismic velocities. *Journal of Geophysical Research: Solid Earth*, *120*(12), 8720–8742. https://doi.org/10.1002/2015jb011917

Andrieu, C., & Thoms, J. (2008). A tutorial on adaptive MCMC. *Statistics and Computing*, *18*(4), 343–373. https://doi.org/10.1007/s11222-008-9110-y

Austermann, J., Hoggard, M. J., Latychev, K., Richards, F. D., & Mitrovica, J. X. (2021). The effect of lateral variations in Earth structure on Last Interglacial sea level. *Geophysical Journal International*, *227*(3), 1938–1960. https://doi.org/10.1093/gji/ggab289

Austermann, J., Mitrovica, J. X., Latychev, K., & Milne, G. A. (2013). Barbados-based estimate of ice volume at Last Glacial Maximum affected by subducted plate. *Nature Geoscience*, *6*(7), 553–557. https://doi.org/10.1038/ngeo1859

Austermann, J., Pollard, D., Mitrovica, J. X., Moucha, R., Forte, A. M., DeConto, R. M., et al. (2015). The impact of dynamic topography change on Antarctic ice sheet stability during the mid-Pliocene warm period. *Geology*, *43*(10), 927–930. https://doi.org/10.1130/g36988.1

Ball, P. W., White, N. J., Maclennan, J., & Stephenson, S. N. (2021). Global influence of mantle temperature and plate thickness on intraplate volcanism. *Nature Communications*, *12*(1), 2045. https://doi.org/10.1038/s41467-021-22323-9

Bamber, J. L., Westaway, R. M., Marzeion, B., & Wouters, B. (2018). The land ice contribution to sea level during the satellite era. *Environmental Research Letters*, *13*(6), 063008. https://doi.org/10.1088/1748-9326/aac2f0

Barletta, V. R., Bevis, M., Smith, B. E., Wilson, T., Brown, A., Bordoni, A., et al. (2018). Observed rapid bedrock uplift in Amundsen Sea Embayment promotes ice-sheet stability. *Science*, *360*(6395), 1335–1339. https://doi.org/10.1126/science.aao1447

Barletta, V. R., & Bordoni, A. (2013). Effect of different implementations of the same ice history in GIA modeling. *Journal of Geodynamics*, *71*, 65–73. https://doi.org/10.1016/j.jog.2013.07.002

Blank, B., Barletta, V., Hu, H., Pappa, F., & van der Wal, W. (2021). Effect of lateral and stress-dependent viscosity variations on GIA induced uplift rates in the Amundsen Sea Embayment. *Geochemistry, Geophysics, Geosystems*, *22*(9), e2021GC009807. https://doi.org/10.1029/2021gc009807

Burgos, G., Montagner, J. P., Beucler, E., Capdeville, Y., Mocquet, A., & Drilleau, M. (2014). Oceanic lithosphere-asthenosphere boundary from surface wave dispersion data. *Journal of Geophysical Research: Solid Earth*, *119*(2), 1079–1093. https://doi.org/10.1002/2013jb010528

Burton-Johnson, A., Dziadek, R., & Martin, C. (2020). Review article: Geothermal heat flow in Antarctica: Current and future directions. *The Cryosphere*, *14*(11), 3843–3873. https://doi.org/10.5194/tc-14-3843-2020

Cammarano, F., Goes, S., Vacher, P., & Giardini, D. (2003). Inferring upper-mantle temperatures from seismic velocities. *Physics of the Earth and Planetary Interiors*, *138*(3), 197–222. https://doi.org/10.1016/s0031-9201(03)00156-0

Caron, L., Ivins, E. R., Larour, E., Adhikari, S., Nilsson, J., & Blewitt, G. (2018). GIA model statistics for GRACE hydrology, cryosphere, and ocean science. *Geophysical Research Letters*, *45*(5), 2203–2212. https://doi.org/10.1002/2017gl076644

Champac, V., & Garcia Gervacio, J. (2018). *Timing performance of nanometer digital circuits under process variations* (1st ed., 204 pp.). Springer International Publishing.

Cobden, L., Goes, S., Cammarano, F., & Connolly, J. A. D. (2008). Thermochemical interpretation of one-dimensional seismic reference models for the upper mantle: Evidence for bias due to heterogeneity. *Geophysical Journal International*, *175*(2), 627–648. https://doi.org/10.1111/j.1365-246x.2008.03903.x

Cook, A. J., & Vaughan, D. G. (2010). Overview of areal changes of the ice shelves on the Antarctic Peninsula over the past 50 years. *The Cryosphere*, *4*(1), 77–98. https://doi.org/10.5194/tc-4-77-2010

Coulon, V., Bulthuis, K., Whitehouse, P. L., Sun, S., Haubner, K., Zipf, L., & Pattyn, F. (2021). Contrasting response of West and East Antarctic ice sheets to glacial isostatic adjustment. *Journal of Geophysical Research: Earth Surface*, *126*(7), e2020JF006003. https://doi.org/10.1029/2020jf006003

Dalton, C. A., Ekström, G., & Dziewoński, A. M. (2008). The global attenuation structure of the upper mantle. *Journal of Geophysical Research*, *113*(B09303). https://doi.org/10.1029/2007JB005429

Dalton, C. A., Langmuir, C. H., & Gale, A. (2014). Geophysical and geochemical evidence for deep temperature variations beneath mid-ocean ridges. *Science*, *344*(6179), 80–83. https://doi.org/10.1126/science.1249466

Davies, B. J., Golledge, N. R., Glasser, N. F., Carrivick, J. L., Ligtenberg, S. R. M., Barrand, N. E., et al. (2014). Modelled glacier response to centennial temperature and precipitation trends on the Antarctic Peninsula. *Nature Climate Change*, *4*(11), 993–998. https://doi.org/10.1038/nclimate2369

Davies, D. R., Valentine, A. P., Kramer, S. C., Rawlinson, N., Hoggard, M. J., Eakin, C. M., & Wilson, C. R. (2019). Earth's multi-scale topographic response to global mantle flow. *Nature Geoscience*, *12*(10), 845–850. https://doi.org/10.1038/s41561-019-0441-4

DeConto, R. M., & Pollard, D. (2016). Contribution of Antarctica to past and future sea-level rise. *Nature*, *531*(7596), 591–597. https://doi.org/10.1038/nature17145

DIGIS Team. (2021). GEOROC compilation: Rift volcanics. Goettingen Research Online/Data, V6. https://doi.org/10.25625/KAIVCT

Eilon, Z., Fischer, K. M., & Dalton, C. A. (2018). An adaptive Bayesian inversion for upper-mantle structure using surface waves and scattered body waves. *Geophysical Journal International*, *214*(1), 232–253. https://doi.org/10.1093/gji/ggy137

Faul, U. H., & Jackson, I. (2005). The seismological signature of temperature and grain size variations in the upper mantle. *Earth and Planetary Science Letters*, *234*(1), 119–134. https://doi.org/10.1016/j.epsl.2005.02.008

Faul, U. H., & Jackson, I., (2007). Diffusion creep of dry, melt-free olivine. *Journal of Geophysical Research*, *112*(B4). https://doi.org/10.1029/2006jb004586

Fretwell, P., Pritchard, H. D., Vaughan, D. G., Bamber, J. L., Barrand, N. E., Bell, R., et al. (2013). Bedmap2: Improved ice bed, surface and thickness datasets for Antarctica. *The Cryosphere*, *7*(1), 375–393. https://doi.org/10.5194/tc-7-375-2013

Fukuda, J., & Johnson, K. M. (2010). Mixed linear–non-linear inversion of crustal deformation data: Bayesian inference of model, weighting and regularization parameters. *Geophysical Journal International*, *181*(3). https://doi.org/10.1111/j.1365-246x.2010.04564.x

Gallagher, K., Charvin, K., Nielsen, S., Sambridge, M., & Stephenson, J. (2009). Markov chain Monte Carlo (MCMC) sampling methods to determine optimal models, model resolution and model choice for Earth Science problems. *Marine and Petroleum Geology*, *26*(4), 525–535. https://doi.org/10.1016/j.marpetgeo.2009.01.003

Gelman, A., Gilks, W. R., & Roberts, G. O. (1997). Weak convergence and optimal scaling of random walk Metropolis algorithms. *Annals of Applied Probability*, *7*(1). https://doi.org/10.1214/aoap/1034625254

Gomez, N., Latychev, K., & Pollard, D. (2018). A coupled ice sheet–sea level model incorporating 3D Earth structure: Variations in Antarctica during the last deglacial retreat. *Journal of Climate*, *31*(10), 4041–4054. https://doi.org/10.1175/jcli-d-17-0352.1

Gomez, N., Mitrovica, J. X., Huybers, P., & Clark, P. U. (2010). Sea level as a stabilizing factor for marine-ice-sheet grounding lines. *Nature Geoscience*, *3*(12), 850–853. https://doi.org/10.1038/ngeo1012

Haario, H., Saksman, E., & Tamminen, J. (2001). An adaptive metropolis algorithm. *Bernoulli*, *7*(2), 223. https://doi.org/10.2307/3318737

Havlin, C., Holtzman, B. K., & Hopper, E. (2021). Inference of thermodynamic state in the asthenosphere from anelastic properties, with applications to North American upper mantle. *Physics of the Earth and Planetary Interiors*, *314*, 106639. https://doi.org/10.1016/j.pepi.2020.106639

Hirth, G., & Kohlstedt, D. (2004). Rheology of the upper mantle and the mantle wedge: A view from the experimentalists. In *Inside the subduction factory* (pp. 83–105). American Geophysical Union (AGU).

Hoggard, M. J., Czarnota, K., Richards, F. D., Huston, D. L., Jaques, A. L., & Ghelichkhan, S. (2020). Global distribution of sediment-hosted metals controlled by craton edge stability. *Nature Geoscience*, *13*(7), 504–510. https://doi.org/10.1038/s41561-020-0593-2

Ivins, E. R., van der Wal, W., Wiens, D. A., Lloyd, A. J., & Caron, L. (2021). Antarctic upper mantle rheology. *Geological Society, London, Memoirs*, *56*(1), 267–294. https://doi.org/10.1144/m56-2020-19

Ivins, E. R., Watkins, M. M., Yuan, D. N., Dietrich, R., Casassa, G., & Rülke, A. (2011). On-land ice loss and glacial isostatic adjustment at the Drake Passage: 2003–2009. *Journal of Geophysical Research*, *116*(B2), B02403. https://doi.org/10.1029/2010jb007607

Jain, C., Korenaga, J., & Karato, S.-I. (2019). Global analysis of experimental data on the rheology of olivine aggregates. *Journal of Geophysical Research: Solid Earth*, *124*(1), 310–334. https://doi.org/10.1029/2018jb016558

Jennings, S., Hasterok, D., & Payne, J. (2019). A new compositionally based thermal conductivity model for plutonic rocks. *Geophysical Journal International*, *219*(2), 1377–1394. https://doi.org/10.1093/gji/ggz376

Karato, S., & Jung, H. (1998). Water, partial melting and the origin of the seismic low velocity and high attenuation zone in the upper mantle. *Earth and Planetary Science Letters*, *157*(3), 193–207. https://doi.org/10.1016/s0012-821x(98)00034-x

King, M. A., Bingham, R. J., Moore, P., Whitehouse, P. L., Bentley, M. J., & Milne, G. A. (2012). Lower satellite-gravimetry estimates of Antarctic sea-level contribution. *Nature*, *491*(7425), 586–589. https://doi.org/10.1038/nature11621

Klöcking, M., Hoggard, M., Rodríguez Tribaldos, V., Richards, F., Guimarães, A., Maclennan, J., & White, N. (2020). A tale of two domes: Neogene to recent volcanism and dynamic uplift of northeast Brazil and southwest Africa. *Earth and Planetary Science Letters*, *547*, 116464. https://doi.org/10.1016/j.epsl.2020.116464

Korenaga, T., & Korenaga, J. (2016). Evolution of young oceanic lithosphere and the meaning of seafloor subsidence rate. *Journal of Geophysical Research: Solid Earth*, *121*(9), 6315–6332. https://doi.org/10.1002/2016jb013395

Lau, H. C. P., Austermann, J., Holtzman, B. K., Havlin, C., Lloyd, A. J., Book, C., & Hopper, E. (2021). Frequency ddependent mantle viscoelasticity via the complex viscosity: Cases from Antarctica. *Journal of Geophysical Research: Solid Earth*, *126*(11), e2021JB022622. https://doi.org/10.1029/2021jb022622

Lau, H. C. P., & Holtzman, B. K. (2019). "Measures of dissipation in viscoelastic media" extended: Toward continuous characterization across very broad geophysical time scales. *Geophysical Research Letters*, *46*(16), 9544–9553. https://doi.org/10.1029/2019gl083529

Lau, H. C. P., Mitrovica, J. X., Austermann, J., Crawford, O., Al-Attar, D., & Latychev, K. (2016). Inferences of mantle viscosity based on ice age data sets: Radial structure. *Journal of Geophysical Research: Solid Earth*, *121*(10), 6991–7012. https://doi.org/10.1002/2016jb013043

Lee, C. (2003). Compositional variation of density and seismic velocities in natural peridotites at STP conditions: Implications for seismic imaging of compositional heterogeneities in the upper mantle. *Journal of Geophysical Research*, *108*(B9). https://doi.org/10.1029/2003jb002413

Lloyd, A. J., Wiens, D. A., Zhu, H., Tromp, J., Nyblade, A. A., Aster, R. C., et al. (2020). Seismic structure of the Antarctic upper mantle imaged with adjoint tomography. *Journal of Geophysical Research: Solid Earth*, *125*(3). https://doi.org/10.1029/2019jb017823

McCarthy, C., & Takei, Y. (2011). Anelasticity and viscosity of partially molten rock analogue: Toward seismic detection of small quantities of melt. *Geophysical Research Letters*, *38*(18). https://doi.org/10.1029/2011gl048776

McKenzie, D., Jackson, J., & Priestley, K. (2005). Thermal structure of oceanic and continental lithosphere. *Earth and Planetary Science Letters*, *233*(3), 337–349. https://doi.org/10.1016/j.epsl.2005.02.005

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, *21*(6), 1087–1092. https://doi.org/10.1063/1.1699114

Milne, G. A., Latychev, K., Schaeffer, A., Crowley, J. W., Lecavalier, B. S., & Audette, A. (2018). The influence of lateral Earth structure on glacial isostatic adjustment in Greenland. *Geophysical Journal International*, *214*(2), 1252–1266. https://doi.org/10.1093/gji/ggy189

Mitrovica, J., Austermann, J., Coulson, S., Creveling, J., Hoggard, M., Jarvis, G., & Richards, F. (2020). Dynamic topography and ice age paleoclimate. *Annual Review of Earth and Planetary Sciences*, *48*(1), 585–621. https://doi.org/10.1146/annurev-earth-082517-010225

Morlighem, M., Rignot, E., Binder, T., Blankenship, D., Drews, R., Eagles, G., et al. (2020). Deep glacial troughs and stabilizing ridges unveiled beneath the margins of the Antarctic ice sheet. *Nature Geoscience*, *13*(2), 132–137. https://doi.org/10.1038/s41561-019-0510-8

Nield, G. A., Barletta, V. R., Bordoni, A., King, M. A., Whitehouse, P. L., Clarke, P. J., et al. (2014). Rapid bedrock uplift in the Antarctic Peninsula explained by viscoelastic response to recent ice unloading. *Earth and Planetary Science Letters*, *397*, 32–41. https://doi.org/10.1016/j.epsl.2014.04.019

Noble, T. L., Rohling, E. J., Aitken, A. R. A., Bostock, H. C., Chase, Z., Gomez, N., et al. (2020). The sensitivity of the Antarctic ice sheet to a changing climate: Past, present, and future. *Reviews of Geophysics*, *58*(4). https://doi.org/10.1029/2019rg000663

Priestley, K., & McKenzie, D. (2006). The thermal structure of the lithosphere from shear wave velocities. *Earth and Planetary Science Letters*, *244*(1), 285–301. https://doi.org/10.1016/j.epsl.2006.01.008

Priestley, K., & McKenzie, D. (2013). The relationship between shear wave velocity, temperature, attenuation and viscosity in the shallow part of the mantle. *Earth and Planetary Science Letters*, *381*, 78–91. https://doi.org/10.1016/j.epsl.2013.08.022

Priestley, K., McKenzie, D., & Ho, T. (2018). A lithosphere–asthenosphere boundary—A global model derived from multimode surface-wave tomography and petrology. In *Lithospheric discontinuities*, Chap. 6 (pp. 111–123). American Geophysical Union (AGU).

Richards, F. D., Hoggard, M. J., Crosby, A., Ghelichkhan, S., & White, N. (2020). Structure and dynamics of the oceanic lithosphere-asthenosphere system. *Physics of the Earth and Planetary Interiors*, *309*, 106559. https://doi.org/10.1016/j.pepi.2020.106559

Richards, F. D., Hoggard, M. J., Cowton, L. R., & White, N. J. (2018). Reassessing the thermal structure of oceanic lithosphere with revised global inventories of basement depths and heat flow measurements. *Journal of Geophysical Research: Solid Earth*, *123*(10), 9136–9161. https://doi.org/10.1029/2018jb015998

Richards, F. D., Hoggard, M. J., White, N., & Ghelichkhan, S. (2020). Quantifying the relationship between short-wavelength dynamic topography and thermomechanical structure of the upper mantle using calibrated parameterization of anelasticity. *Journal of Geophysical Research: Solid Earth*, *125*(9). https://doi.org/10.1029/2019jb019062

Roy, V. (2020). Convergence diagnostics for Markov Chain Monte Carlo. *Annual Review of Statistics and Its Application*, *7*(1), 387–412. https://doi.org/10.1146/annurev-statistics-031219-041300

Sammon, L. G., McDonough, W. F., & Mooney, W. D. (2022). Compositional attributes of the deep continental crust inferred from geochemical and geophysical data. *Journal of Geophysical Research: Solid Earth*, *127*(8). https://doi.org/10.1029/2022jb024041

Samrat, N. H., King, M. A., Watson, C., Hay, A., Barletta, V. R., & Bordoni, A. (2021). Upper mantle viscosity underneath Northern Marguerite Bay, Antarctic Peninsula constrained by bedrock uplift and ice mass variability. *Geophysical Research Letters*, *48*(24). https://doi.org/10.1029/2021GL097065

Seroussi, H., Ivins, E. R., Wiens, D. A., & Bondzio, J. (2017). Influence of a West Antarctic mantle plume on ice sheet basal conditions. *Journal of Geophysical Research: Solid Earth*, *122*(9), 7127–7155. https://doi.org/10.1002/2017jb014423

Shen, Q., Wang, H., Shum, C. K., Jiang, L., Hsu, H. T., & Dong, J. (2018). Recent high-resolution Antarctic ice velocity maps reveal increased mass loss in Wilkes Land, East Antarctica. *Scientific Reports*, *8*(1), 4477. https://doi.org/10.1038/s41598-018-22765-0

Shen, W., Wiens, D. A., Anandakrishnan, S., Aster, R. C., Gerstoft, P., Bromirski, P. D., et al. (2018). The crust and upper mantle structure of central and West Antarctica from Bayesian inversion of Rayleigh wave and receiver functions. *Journal of Geophysical Research: Solid Earth*, *123*(9), 7824–7849. https://doi.org/10.1029/2017jb015346

Shen, W., Wiens, D. A., Lloyd, A. J., & Nyblade, A. A. (2020). A geothermal heat flux map of Antarctica empirically constrained by seismic structure. *Geophysical Research Letters*, *47*(14). https://doi.org/10.1029/2020gl086955

Shepherd, A., Ivins, E., Rignot, E., Smith, B., van den Broeke, M., Velicogna, I., et al. (2018). Mass balance of the Antarctic Ice Sheet from 1992 to 2017. *Nature*, *558*(7709), 219–222. https://doi.org/10.1038/s41586-018-0179-y

Shepherd, A., Ivins, E., Rignot, E., Smith, B., van den Broeke, M., Velicogna, I., et al. (2020). Mass balance of the Greenland Ice Sheet from 1992 to 2018. *Nature*, *579*(7798), 233–239. https://doi.org/10.1038/s41586-019-1855-2

Slangen, A. B. A., Adloff, F., Jevrejeva, S., Leclercq, P. W., Marzeion, B., Wada, Y., & Winkelmann, R. (2017). A review of recent updates of sea-level projections at global and regional scales. *Surveys in Geophysics*, *38*(1), 385–406. https://doi.org/10.1007/s10712-016-9374-2

Steinberger, B., Conrad, C. P., Osei Tutu, A., & Hoggard, M. J. (2019). On the amplitude of dynamic topography at spherical harmonic degree two. *Tectonophysics*, *760*, 221–228. https://doi.org/10.1016/j.tecto.2017.11.032

Stixrude, L., & Lithgow-Bertelloni, C. (2011). Thermodynamics of mantle minerals—II. Phase equilibria. *Geophysical Journal International*, *184*(3), 1180–1213. https://doi.org/10.1111/j.1365-246x.2010.04890.x

Takei, Y. (2017). Effects of partial melting on seismic velocity and attenuation: A new insight from experiments. *Annual Review of Earth and Planetary Sciences*, *45*(1), 447–470. https://doi.org/10.1146/annurev-earth-063016-015820

Thomas, I. D., King, M. A., Bentley, M. J., Whitehouse, P. L., Penna, N. T., Williams, S. D. P., et al. (2011). Widespread low rates of Antarctic glacial isostatic adjustment revealed by GPS observations. *Geophysical Research Letters*, *38*(22). https://doi.org/10.1029/2011gl049277

Whitehouse, P. L. (2018). Glacial isostatic adjustment modelling: Historical perspectives, recent advances, and future directions. *Earth Surface Dynamics*, *6*(2), 401–429. https://doi.org/10.5194/esurf-6-401-2018

Wolstencroft, M., King, M. A., Whitehouse, P. L., Bentley, M. J., Nield, G. A., King, E. C., et al. (2015). Uplift rates from a new high-density GPS network in Palmer Land indicate significant late Holocene ice loss in the southwestern Weddell Sea. *Geophysical Journal International*, *203*(1), 737–754. https://doi.org/10.1093/gji/ggv327

Yabe, K., & Hiraga, T. (2020). Grain-boundary diffusion creep of olivine: 2. Solidus effects and consequences for the viscosity of the oceanic upper mantle. *Journal of Geophysical Research: Solid Earth*, *125*(8). https://doi.org/10.1029/2020jb019416

Yamauchi, H., & Takei, Y. (2016). Polycrystal anelasticity at near-solidus temperatures. *Journal of Geophysical Research: Solid Earth*, *121*(11), 7790–7820. https://doi.org/10.1002/2016jb013316

Zwally, H. J., & Giovinetto, M. B. (2011). Overview and assessment of Antarctic ice-sheet mass balance estimates: 1992-2009. *Surveys in Geophysics*, *32*(4), 351–376. https://doi.org/10.1007/s10712-011-9123-5

Zwally, J. H., Giovinetto, M. B., Beckley, M. A., & Saba, J. L. (2012). *Antarctic and Greenland drainage systems*. GSFC Cryospheric Sciences Laboratory.