Data Mine

A Short-Period Surface-Wave Dispersion Dataset for Model Assessment of Africa's Crust: ADAMA

Tolulope Olugboji*[1] and Siyu Xue[1,2]

Abstract

We present the first in a series of dataset and model assessment products for investigating Africa's lithosphere (ADAMA). This is a comprehensive catalog of short-period interstation surface-wave dispersion measurements and uncertainties. It is derived from processing continuous recordings of all publicly available three-component seismograms, spanning four decades, from ∼1372 stations, across 62 seismic networks deployed in and around the African continent. It includes Love- and Rayleigh-wave dispersion derived from frequency-domain ambient noise cross-correlation functions (NCFs). Phase and group dispersion, as well as their uncertainties, are then obtained with an iterative nonlinear waveform fitting of the NCFs, using a spectral element representation of a path-average a priori Earth model. Our catalog represents the following advances: (1) a large distribution of short period dispersion measurements: ∼114,000 interstation pairs at periods between 5 s and 40 s, (2) inclusion of uncertainties useful for regularization in continent-wide model building, (3) preliminary model assessments for different tectonic domains on the continent, and (4) an exemplary Love-wave phase velocity map obtained by Bayesian inversion revealing detailed features not previously detected. ADAMA will be used to prepare short-period, high-resolution dispersion maps, and for assessment and updates of widely used seismic velocity models of the crust across a diversity of terranes on the continent.

Cite this article as Olugboji, T., and S. Xue (2022). A Short-Period Surface-Wave Dispersion Dataset for Model Assessment of Africa's Crust: ADAMA, *Seismol. Res. Lett.* **XX**, 1–17, doi: 10.1785/0220210355.

Supplemental Material

Introduction

The African continent hosts a remarkable diversity of geologic features and tectonic terranes spanning ∼80% of Earth's history (Burke, 1996; Begg *et al.*, 2009). These include large and small cratons, a superswell in the south, an active rift in the east, and a network of failed rifts in the west and central region (Van Hinsbergen, 2011; Van Hinsbergen *et al.*, 2011). A common approach to studying this rich tapestry of geological complexity is to turn to continent-wide (Pasyanos and Nyblade, 2007) or global seismic models of the crust (i.e., Crust1.0 and Litho1.0) (Nataf and Ricard, 1996; Pasyanos *et al.*, 2014). When combined with other complementary datasets, that is, gravity, geochemistry, geoneutrino, and rock physics, these seismic models have been useful for providing important constraints on the composition and secular evolution of the crust and mantle lithosphere (Christensen and Mooney, 1995; Keller and Schoene, 2012; Huang *et al.*, 2014; Rudnick and Gao, 2014; Globig *et al.*, 2016).

Continued investment in continent-wide (Nyblade *et al.*, 2008; Fishwick and Bastow, 2011; Durrheim and Nyblade, 2019) and regional passive-source experiments (Tokam *et al.*, 2010; Nyblade, 2015; Fadel *et al.*, 2018, 2020; Yu *et al.*, 2020) have recently broadened the seismic coverage across the continent, making it an opportune time to revise lithospheric models on the continent, particularly across major terrane boundaries where data coverage has now improved. We emphasize that while new approaches to passive source imaging of the lithosphere using ambient noise surface waves (Accardo *et al.*, 2017; Ojo *et al.*, 2017; Hopper *et al.*, 2020) are being used to investigate targeted regional questions, only a few of these seismic measurements have been incorporated

1. Department of Earth and Environmental Sciences, University of Rochester, Rochester, New York, U.S.A., https://orcid.org/0000-0002-2621-2752 (TO); https://orcid.org/0000-0003-3670-1493 (SX); 2. Goergen Institute for Data Science, University of Rochester, Rochester, New York, U.S.A.

*Corresponding author: tolugboj@ur.rochester.edu

into global lithospheric models like Litho1.0 (Pasyanos *et al.*, 2014; Emry *et al.*, 2018). In particular, a continent-wide ambient noise dispersion dataset, with associated uncertainties, comprising both Love and Rayleigh waves has yet to be published. How these measurements may affect updates to global lithospheric models, or should be used in joint inversion schemes, is not well understood (Olugboji *et al.*, 2017). Also, only a few measurements of Love-wave dispersion are available, despite being crucial to constraining radial anisotropy (Moschetti *et al.*, 2010; Ojo *et al.*, 2017). This is possibly due to the poor waveform and metadata quality of horizontal seismograms (Ojo *et al.*, 2019).

In this study, we address these limitations, following an approach similar to that of other continent-wide studies (Rawlinson *et al.*, 2016; Ekström, 2017). We describe our method for extracting short-period dispersion from ambient noise and how they extend other global (Ma and Masters, 2014; Ma *et al.*, 2014) and regional (Emry *et al.*, 2018) surface-wave catalogs. Our measurements include (1) a large waveform database spanning four decades resulting in high-density data coverage compared to the previous studies (Emry *et al.*, 2018), (2) new measurements of Love-wave response useful for constraining radial anisotropy (Lin *et al.*, 2010), (3) more short-wavelength (from temporary arrays: <100 km) and short-period (5–25 s) fundamental-mode measurements sensitive to the mid-upper crust (<30 km), (4) horizontal channels corrected for misorientations (Ojo *et al.*, 2019), and (5) the estimation of uncertainties using a nonlinear waveform fitting process (Hawkins and Sambridge, 2019) valuable for model regularization during model construction (Bodin *et al.*, 2012). The procedure described in this article is particularly useful for extracting dispersion with low-quality seismograms.

Data and Methods

Short-period surface-wave dispersion is estimated from ambient noise cross-correlation functions (NCFs) of seismic data between station pairs operating within the same time range. To ensure ~~a dataset with~~ the best resolution both in space and time, we begin by compiling a comprehensive dataset using all publicly available data (~40 yr: 1980–2021) on the continent ~~across~~. We download a total of ~1 million daily three-channel seismic records at 1372 stations from 62 networks. These include permanent and temporary networks with active station pairs within and across station arrays from different classes of networks in and around the continent (Figs. 1 and 2). Data availability varies between networks. We notice that improved connections between station pairs are provided by the continuously recording permanent networks (G, GT, G, II, IU), which span a wide range of tectonic regions and have been operating since the 1990s. With the advent of new temporary networks (i.e., XB, YB, XA, XW), and the continuously running AfricaArray (AF) networks, data recovery increased by a factor of 10 in the second epoch (2000–2010). In the last decade (2010–2020), large country-wide networks contribute significantly to data availability, improving spatial resolution especially in the south and southeastern regions of Africa as well as in Madagascar and Cameroon (see Fig. 2d).

Spatial sampling of central and western Africa, while poor at short wavelengths ($<2 \times 10^3$ km), is clearly improved by long-running networks like the GSN, Geoscope, Spanish, and Mediterranean networks (G, GT, ES, and MN: Figs. 1b and 2d). The illumination of these regions is improved by the interconnectivity of these long-running networks with short-aperture country-wide seismic arrays, especially at intermediate-to-long wavelengths (Fig. 2b,c). This improvement in resolution should be reflected in new details observed in phase velocity maps obtained from our new dataset and by extension lithospheric models constrained with them. We explore this in the Exemplary Phase Map for Love Waves section.

Signal processing: computing NCF ~~functions~~ (Love and Rayleigh)

Our measurement of short-period surface-wave phase velocity dispersion proceeds in two steps: (1) we compute the cross-spectra between individual station pairs, that is, the NCF (Ekström, 2014); (2) we use the NCFs to determine the phase velocity dispersion using the Aki formulation (Aki, 1957; Ekström *et al.*, 2009). The first step is a signal processing step, whereas the second is a parameter estimation problem. In computing the NCFs, we follow closely the approach described in Ekström (2014, 2017), with only slight modifications.

Data preprocessing begins with preparing daily three-channel displacement seismic records $u^x(t) : x = Z, R, T$, at all stations, downsampled to 1 Hz following instrument response removal. Each station is cross-referenced with every other station to determine the stations that they are connected to. Station pairs are connected when they have at least 60 days of overlapping operation and are spaced no less than 60 km apart ($r > 60$ km). Each NCF, $\bar{\rho}(\omega, r)$, is the stacked cross-spectra between connected station pairs separated by distance r following subdivision of daily records into 4 hr overlapping segments:

$$\bar{\rho}(\omega, r) = \sum_{k=1}^{k=N_{j\otimes l}} \rho_{jkl}(\omega, r), \tag{1}$$

$$\rho_{jkl}(\omega) = \frac{u_{jk}(\omega) u_{lk}^*(\omega)}{\sqrt{u_{jk}(\omega) u_{jk}^*(\omega)}\sqrt{u_{lk}(\omega) u_{lk}^*(\omega)}}, \tag{2}$$

in which k identifies each individual 4 hr seismic record, $N_{j\otimes l}$ is the total number of segments for each interconnected station pair, and $u_{jk}(\omega), u_{lk}(\omega)$ is the ground displacement spectra for stations j and l obtained by Fourier analysis ($N_{j\otimes l} > 360$: requiring that both the stations operate at the same time for

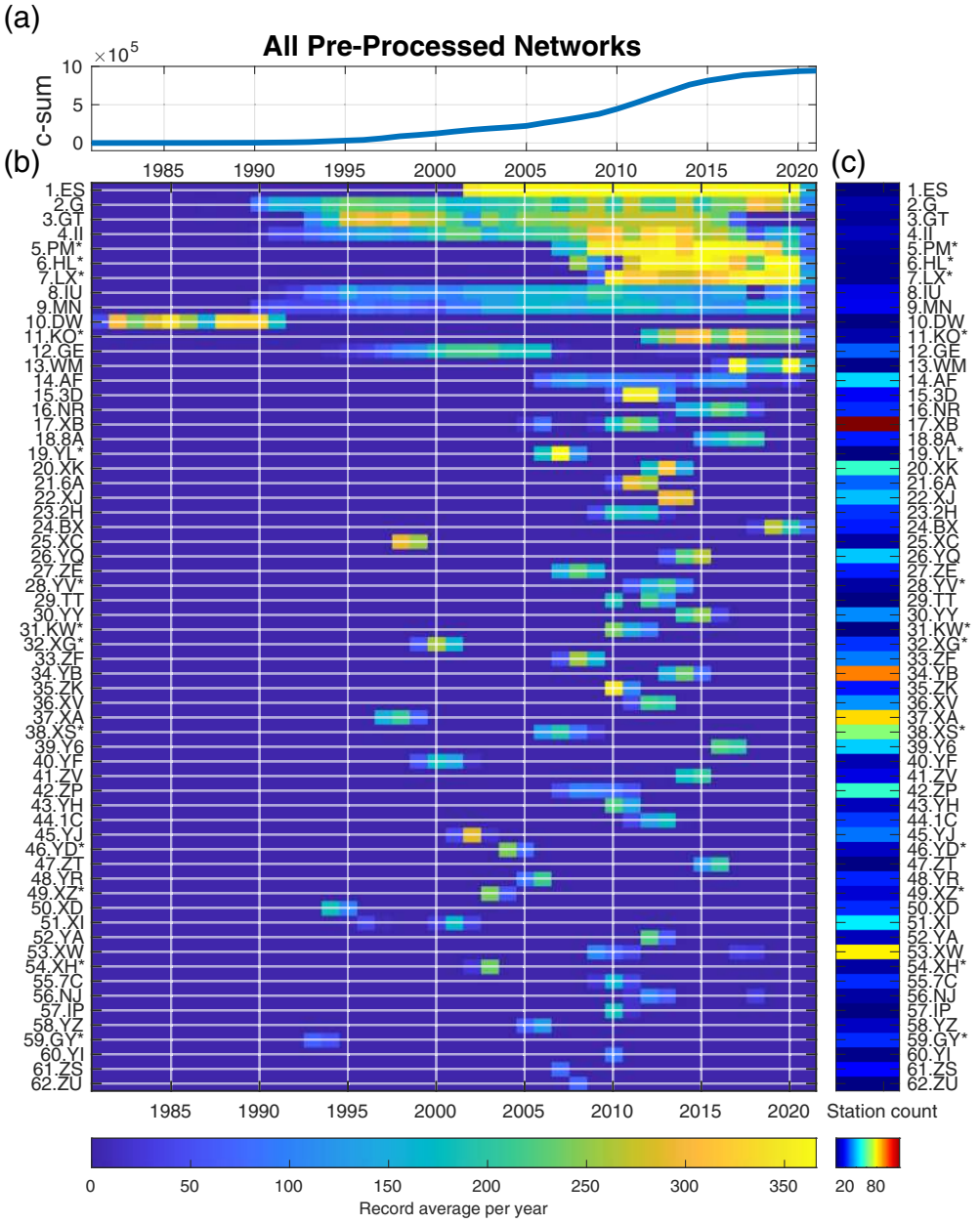

Figure 1. Waveform database showing distribution of daily records downloaded from 62 seismic networks, spanning four (4) decades: 1980–2020. (a) The cumulative count of the ~10^6 (1 million) total daily records highlights three epochs of network deployment: global networks (<1995), regional deployments (1995–2010), and country networks (>2010). (b) Histogram of annual data completeness for each network, allowing ranking based on total daily contribution to the entire database. (c) The station count for each network. Note that ranking is decided by temporal data availability, while spatial coverage is a function of the total number of stations concurrently operating within and across network arrays. So, the fact that multiple large-array networks are operating concurrently from 2005–2015 leads to improved spatial coverage of the continent. A connection is active when a station pair is recording concurrently for no less than 60 days (~2 months). The color version of this figure is available only in the electronic edition.

at least 60 days). We are able to obtain three sets of cross-spectra $\tilde{\rho}$, one each for vertical–vertical $\tilde{\rho}^R_{ZZ}$, radial–radial $\tilde{\rho}^R_{RR}$ (Rayleigh), and transverse–transverse $\tilde{\rho}^L_{TT}$, (Love) displacement cross-spectra representing a total of 114,487 interconnected station pairs. Our measurements span both long- and short-distance connections (see Fig. 2a–c).

The second stage following NCF computation is parameter estimation, and it involves determining the phase dispersion curve, $c(\omega_0)$, using the Aki formulas (Aki, 1957; Ekström, 2014):

$$\tilde{\rho}(r,\omega_0) = J_0\left(\frac{\omega_0}{c(\omega_0)}r\right). \qquad (3)$$

When horizontal NCFs are used ($\tilde{\rho}^R_{RR}$ or $\tilde{\rho}^L_{TT}$), slight modifications to this expression have been proposed (Haney *et al.*, 2012), but Ekström (2017) argues that, in practice, the form presented earlier should be indistinguishable from new expressions proposed. Here, we describe the process of estimating phase dispersion using the classical Aki equation, and later we revisit the modified expressions and show how they can help determine biases that are present when the horizontal displacement channels are misoriented (see Correcting Love for Horizontal Misorientations and Estimating Uncertainties section).

There are two approaches to parameter estimation using the Aki formulas: (1) an automated method that fits the zero-crossing of the NCF (Ekström *et al.*, 2009) or (2) a linearized (Menke and Jin, 2015) or nonlinear waveform-fitting approach (Hawkins and Sambridge, 2019) that uses the entire amplitude information of the NCF. Here, we adopt the latter approach described in Hawkins and Sambridge (2019; hereafter, HS19). This method uses a priori information of a path-average model to effectively guide the nonlinear waveform fitting (Hawkins, 2018). In practice, it integrates all the best features of the zero crossing and waveform fitting methods. For a

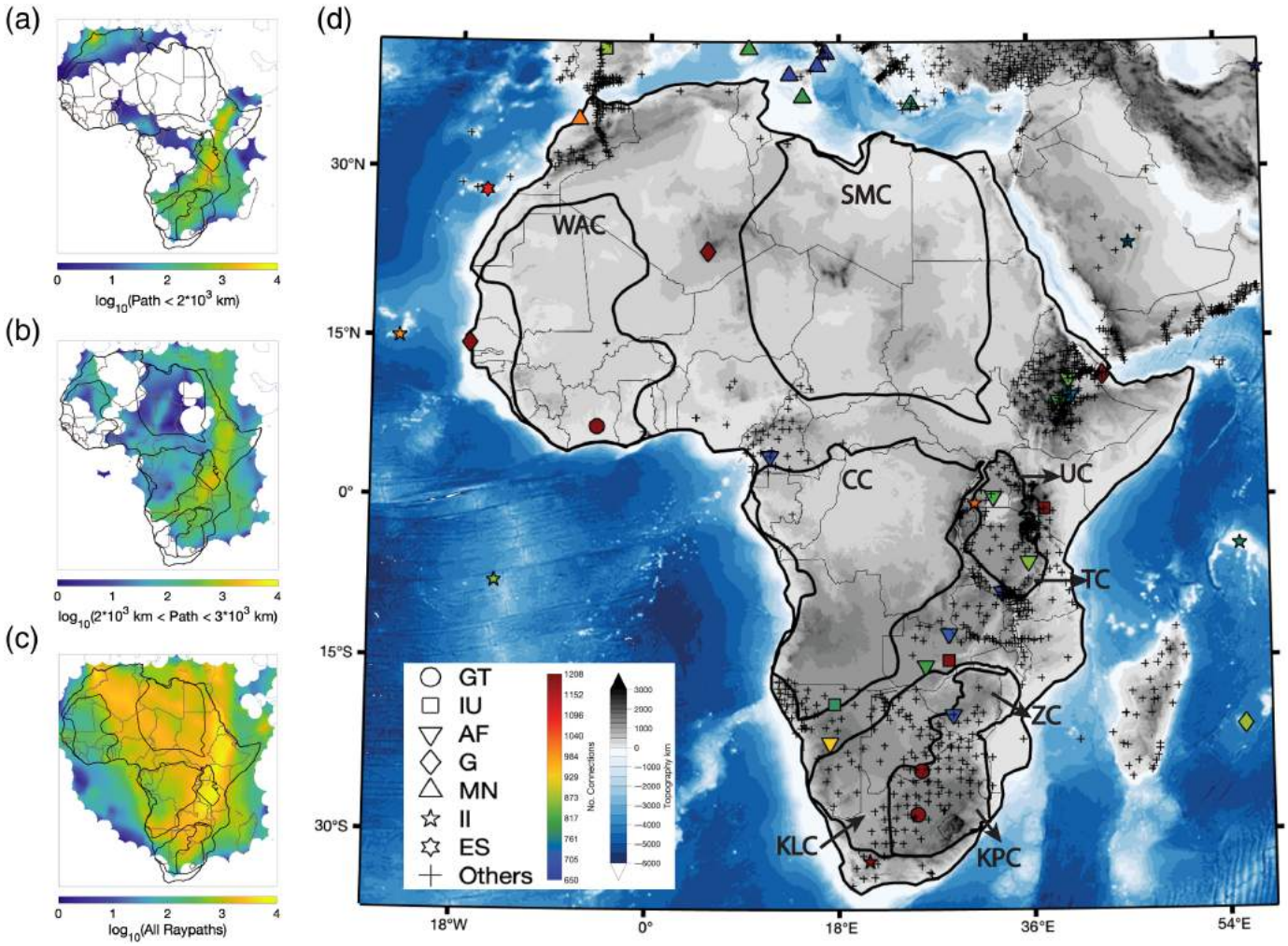

Figure 2. Station locations showing spatial distribution of pairwise connections (ray-path density) used in computing the ambient noise cross-correlation functions (NCFs) across the entire African continent. (a–c) Ray-path density (number of connected station pairs for every 0.25°) showing spatial sampling at (a) short interstation distances, $r_s < 2 \times 10^3$ km; (b) mid-range interstation distances 2×10^3 km $< r_m < 3 \times 10^3$ km; and (c) all interstation. (d) Station classification based on count of connected stations (red-green color bar). The best stations (connected to more than 650 other stations) are from long-running global and regional networks (GT, IU, AF, G, MN, II, and ES). These stations provide long-wavelength coverage while the short-aperture seismic networks (crosses) provide short-wavelength coverage (compare panel d with c). The broad craton outlines (black lines) are taken from Globig *et al.* (2016). CC, Congo Craton; KLC, Kalahari Craton; KPC, Kaapvaal Craton; SMC, Sahara Meta Craton; TC, Tanzanian Craton; UC, Uganda Craton; WAC, West African Craton; and ZC, Zimbabwe Craton. The color version of this figure is available only in the electronic edition.

full description of the technique, we refer the reader to the paper by Hawkins and Sambridge (2019). Here, we summarize the key steps, highlighting its suitability for extracting phase dispersion from poor-quality NCF measurements, that is, cross-spectra with low signal-to-noise ratio (SNR), either due to poor installations or shorter deployment times (i.e., Fig. 3b,d,f).

Phase dispersion estimation using nonlinear waveform fitting

The HS19 method is a two-stage, nonlinear iterative scheme to extract high-resolution Love- and Rayleigh-wave phase dispersion, and its uncertainties, using the Aki approach (Aki, 1957; Ekström *et al.*, 2009). It is based on the nonlinear waveform fitting of the modified Bessel function described in equation (1). In a preliminary step, an initial low-resolution phase dispersion curve c_0 is derived using a method similar to Ekström *et al.* (2009), which fits the zero-crossing phase information of the NCF. In constructing c_0, the zero crossing method of Ekström *et al.* (2009) is supplemented by including the peaks and troughs of the zeroth-order Bessel functions (J_0), which is obtained by finding the zeros of the first-order Bessel function (J_1), because:

$$\frac{d}{d\omega_0}\bar{\rho}(r,\omega_0) \propto -J_1\left(\frac{\omega_0}{c(\omega_0)}r\right). \tag{4}$$

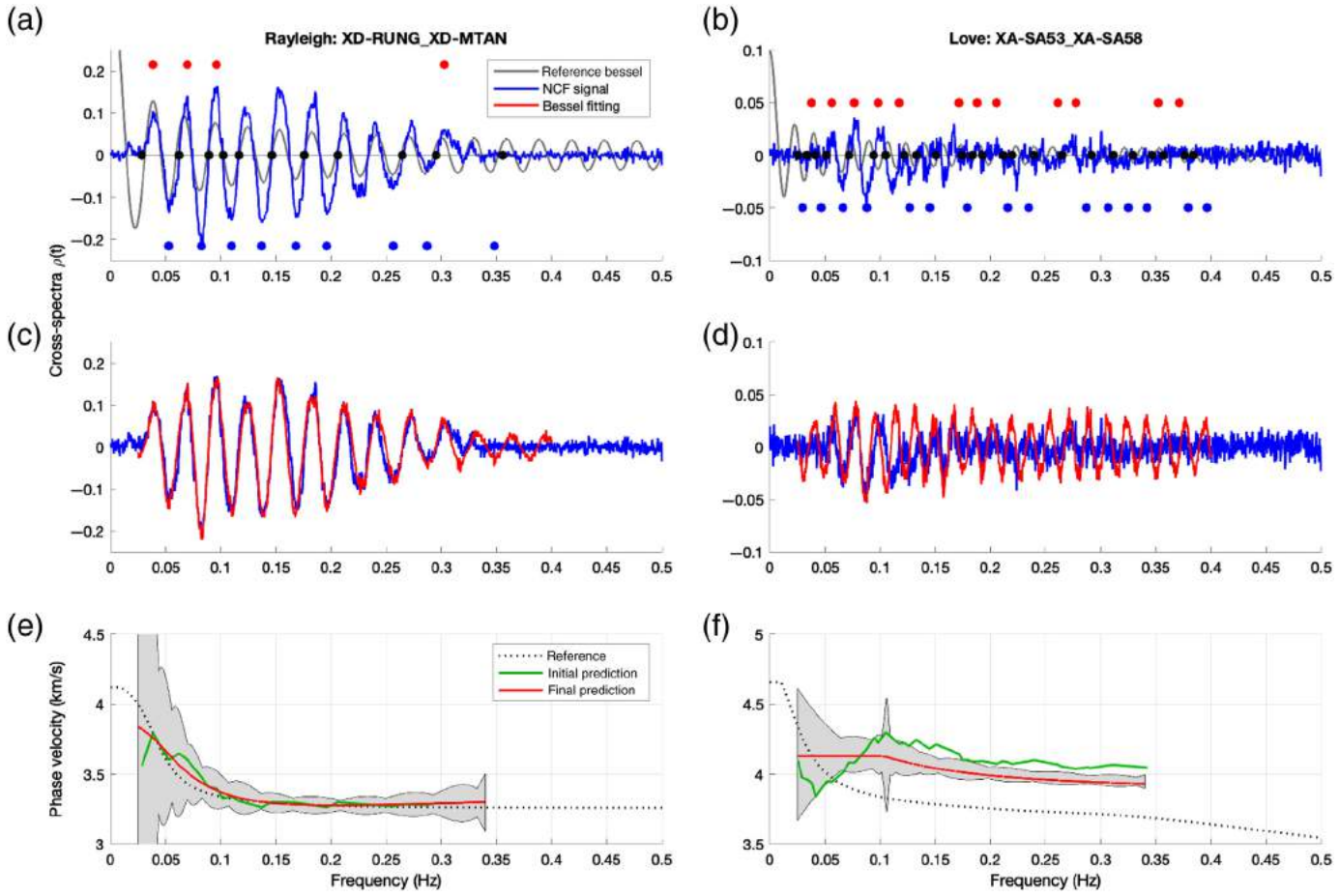

Figure 3. The estimation of Rayleigh- and Love-phase dispersion and uncertainties on two different networks. (a,b) Extracting reference dispersion c_0 using the zeros, peaks, and troughs (black, red, and blue dots) of the stacked NCFs: $\bar{\rho}$ (blue curves). (c,d) Nonlinear waveform fitting of $\bar{\rho}$ using the corrected Bessel function (red curve: ρ^{corr}, see equations 6–8). (e,f) A final high-resolution dispersion curve c_f (red line) and associated uncertainties $\widetilde{\mathbf{C}}$ (gray envelope) obtained from nonlinear waveform fitting using a preoptimized model-dispersion pair $\{\mathbf{m}_1, c_1\}$, in which c_1 (not shown) is itself derived from c_0 (green line, see equation 5). A reference dispersion curve (dotted lines) from Litho1 $c(\omega, \mathbf{m}_0)$ acts as a low-frequency anchor when tracking the position of the zeros, peaks, and troughs (gray curve in a,b). (b,d,f) Love response, although noisier (low signal-to-noise ratio [SNR] in $\bar{\rho}$ and a poor starting $\mathbf{m}_0$) than (a,c,e) Rayleigh, can still be used to extract a reliable c_f^L robust to missed zero crossings. Rayleigh response is from the XD network on the Tanzanian Craton (r = 109.5 km, N = 1749 records), and Love is from XA network on the Kalahari craton (r = 220.4 km, N = 1760 records). See Figure S1 for Love and Rayleigh of XD and XA station pairs. The color version of this figure is available only in the electronic edition.

In the case of low-quality NCFs, including the peaks and troughs of the Bessel functions provides additional constraints that make for a better c_0 (blue and red dots in Fig. 3b). This preliminary step provides necessary a priori constraints to ensure convergence during the nonlinear waveform fitting of the NCF, which proceeds in two steps: first, an a priori path-averaged velocity model $\{\mathbf{m}_0\}$ obtained from the global lithospheric model, Litho1 (Pasyanos *et al.*, 2014), is used to compute a refined phase dispersion curve c_1 using c_0 as a target observational data. In this step, a new optimized model $\mathbf{m}_1$ is obtained in addition to a refined curve c_1. Then comes the actual waveform fitting of the NCF, which uses $\mathbf{m}_1$, alongside c_1 to obtain a final phase dispersion curve c_f. In this second step, the NCF is the target observational data, while $\{\mathbf{m}_1, c_1\}$ are a priori constraints.

We note that in updating the phase velocity curve $c_0 \Rightarrow c_1$, an updated model $\mathbf{m}_0 \Rightarrow \mathbf{m}_1$ is obtained using c_0 as the target constraint during the iterative minimization of the objective function:

$$\sum_i \frac{(c^{\text{pred}}(\omega_i, \mathbf{m}) - c_0(\omega_i))^2}{\sigma_0(\omega_i)} + (\mathbf{m} - \mathbf{m}_0)^T C_m^{-1} (\mathbf{m} - \mathbf{m}_0). \qquad (5)$$

The data misfit is calculated at the discrete frequencies, ω_i, representing the zeros, peaks, and troughs. The observational error $\sigma_0(\omega_i)$ is estimated from the variance of trial curves obtained using J_0 or J_1, and the final solution $\{c_1, \mathbf{m}_1\}$ is that which minimizes equation (5).

The last step in the procedure obtains a high-resolution phase velocity c_f by conducting a nonlinear waveform fitting of the NCFs, using $\{c_1, \mathbf{m}_1\}$ as starting constraints. That is, given a starting model $\mathbf{m}_1$, a predicted NCF can be calculated from equation (3):

$$\rho^{\text{pred}}(\omega, r, \mathbf{m}_1) = J_0\left(\frac{\omega r}{c(\omega, \mathbf{m}_1)}\right). \quad (6)$$

A final solution c_f is obtained by nonlinear optimization using the observed NCF $\bar{\rho}$ (equation 1) as the target constraint in the misfit function :

$$\sum_i \frac{(\rho^{\text{corr}}(\omega_i, r, \mathbf{m}) - \bar{\rho}(\omega_i))^2}{\sigma_i^2} + (\mathbf{m} - \mathbf{m}_1)^T C_m^{-1}(\mathbf{m} - \mathbf{m}_1). \quad (7)$$

In this expression, the predicted NCF is adjusted $\rho^{\text{pred}} \Rightarrow \rho^{\text{corr}}$ before a data misfit is calculated. The correction uses a ratio of amplitude envelopes:

$$\rho^{\text{corr}} = \frac{E(\bar{\rho})}{E(\rho^{\text{pred}})}\rho^{\text{pred}}. \quad (8)$$

This is done to account for the amplitude variation of $\bar{\rho}$ caused by the uneven distribution of noise sources (compare red and blue curves in Fig. 3c,d). The spectral element parameterization of $\mathbf{m}$, the gradients, and the misfits are all computed using adjoint methods described in Hawkins (2018). This technique also provides a final path-average model ($\mathbf{m}_f$) that best explains the Rayleigh and Love NCF response. A comparison of the misfit between c_f and c_0 (offset of red line from dashed line) provides a preliminary model assessment of the Litho1.0 model ($\mathbf{m}_0$). For example, it is clear from the data fit in Figure 3e that $\mathbf{m}_0$ is sufficient along the path described by station pairs RUNG-MTAN (network XD on the Tanzanian craton). In comparison, the station pair SA53 and SA58 on the Kalahari craton (XA network in South Africa) requires significant model updates, because c_f is offset from the reference curve c_0 generated from $\mathbf{m}_0$ (compare Fig. 3f to 3e). The additional information on uncertainties allows us to judge the statistical significance of this data-informed model assessment. We describe this next.

Correcting Love for horizontal misorientations and estimating uncertainties

The quality of our NCFs varies significantly across different station pairs in our database. Three key issues are expected to affect NCF quality and, by extension, the phase dispersions derived from them: (1) number of overlapping records, (2) waveform quality due to network installation or site properties, and (3) horizontal channel misorientations (Brink *et al.*, 2019; Ojo *et al.*, 2019). First, we address the concern of horizontal misorientations and how it might lead to biases in estimated Love response. In the original theoretical treatment of Aki (1957), and confirmed by Haney and Tsai (2020), when using horizontal components for Love-wave phase dispersion, a slight modification to equation (2) is required (in which ρ_{TT} is cross-spectra using tangential seismograms):

$$\bar{\rho}^L_{TT}(r, \omega_0) = 1/2[J_0(\omega r/c^L) - J_2(\omega r/c^L)]. \quad (9)$$

We show that this asymptotically approaches equation (3) and poses no problem for using the earlier expressions to estimate the Love-wave dispersion (Fig. 4). This is not the case when either of the horizontal sensors are not properly oriented; the most severe case is when either of the horizontal channels are orthogonal to the radial direction: $\bar{\rho}_{\text{RT}}$, $\bar{\rho}_{\text{TR}}$, or $\bar{\rho}_{\text{RR}}$. In this situation, the NCFs should rapidly approach zero, making phase estimation for Love waves difficult (Fig. 4). We note that this is not a problem for Rayleigh waves, because we can obtain its phase dispersion from the coherence of vertical channels or mixed-component (vertical–horizontal) displacement spectra. We ensure robustness in our Love measurements by correcting each station for orientation errors using a recently published metadata catalog (Ojo *et al.*, 2019). For stations in our database not in this catalog, we use the method of Doran and Laske (2017) to correct severe orientation biases. An exhaustive analysis shows that only 204 stations have >10° misalignment error and ~6% (85 stations) without orientation measurements.

Apart from horizontal channel misorientation, nonsystematic issues can lead to degradation of the NCFs, and, therefore, estimating uncertainties in our phase dispersion predictions is valuable for secondary users of our measurement. We use the local-linear approximation to estimate uncertainties in phase velocity measurements, that is the covariance matrices, $\widetilde{\mathbf{C}}$, (Tarantola, 2005; Hawkins and Sambridge, 2019):

$$\widetilde{\mathbf{C}} = \mathbf{J}^T\widetilde{\mathbf{C}}_M\mathbf{J}, \quad (10)$$

in which $\widetilde{\mathbf{C}}_M$ is the posterior covariance matrix of the final optimal model, and $\mathbf{J}$ is the Jacobian matrix of partial derivatives with respect to the model parameters. The reported uncertainties in the dispersion measurements can be investigated for influences from structural and nonstructural parameters: that is, duration of operation, metadata quality, and sediment or low-velocity zones (see Table 1).

Phase maps using a Bayesian auto-adaptive approach

We have compiled a large database of interstation phase dispersion measurements using the most comprehensive waveform database available in public archives. Our dataset is useful for model assessment of the African lithosphere. One way to do this is to generate new phase velocity maps across the entire continent. As an example, we discuss, very briefly, how we construct a phase velocity map for Love waves at a period of 40 s.

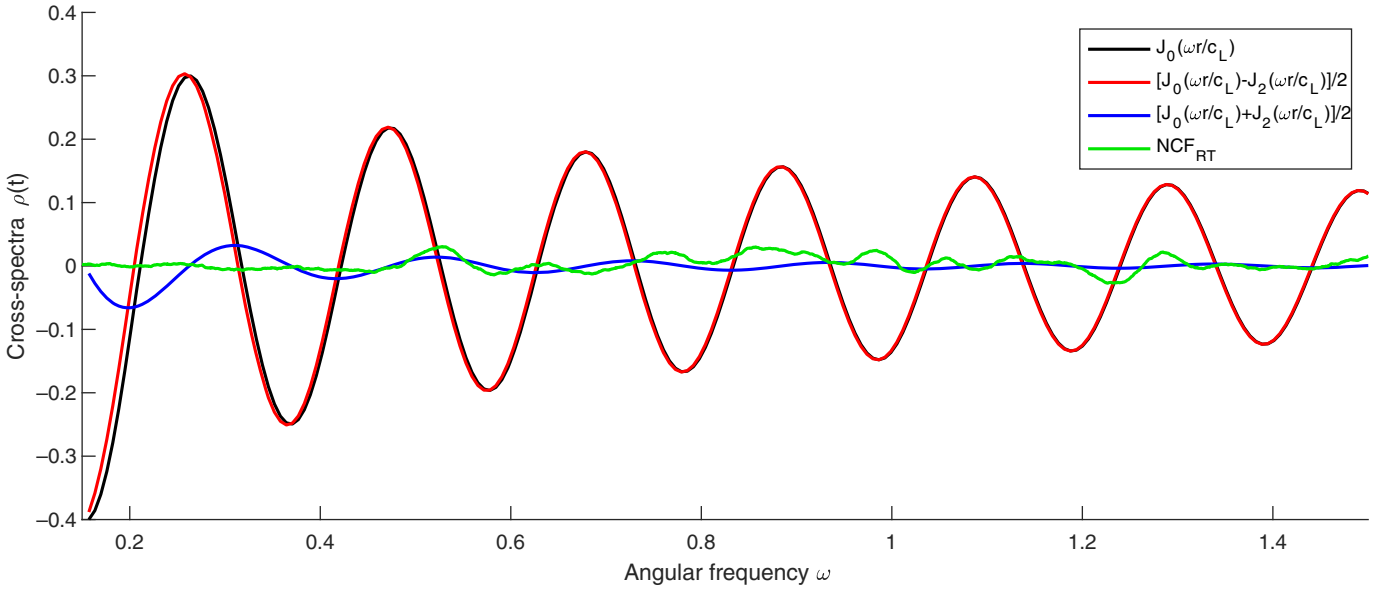

Figure 4. Effect of station misorientation on Love-wave cross-spectra. Orthogonal misalignment is the most severe. The color version of this figure is available only in the electronic edition.

This allows us to compare the ambient noise phase maps to those obtained using earthquake data (Ekström, 2011; Ma and Masters, 2014; Ma *et al.*, 2014). A continent-wide phase map with our new dataset should be broadly consistent with the earthquake maps at the longest wavelengths while providing new resolution in regions where new measurements have been acquired (see Fig. 2).

We adopt a Bayesian auto-adaptive approach (Bodin *et al.*, 2012; Olugboji *et al.*, 2017) to construct a high-resolution Love-wave phase dispersion map $c^L_{T=40}(\theta,\phi)$ from all N_o available interstation phase dispersion measurements c_{ij} at 40 s (Table 1). Following Bayes, we describe a posterior distribution on model parameters $\mathbf{X}$, given data $\mathbf{d}$:

$$P(\mathbf{X}|\mathbf{d}=t_{ij}) \propto P(\mathbf{d}=t_{ij}|\mathbf{X})p(\mathbf{X}), \tag{11}$$

in which the posterior distribution is obtained by a Markov chain Monte Carlo search on a set of $3N_i+2$ parameters every kth step: $\mathbf{X}_k = \{c_i(\theta_i,\phi_i),(\theta_i,\phi_i),N_i,\sigma_c^T\}$. The first set of $3N_i$ parameters allow for an adjustable parameterization, that is, the nearest-neighbor Voronoi tessellation prescribing N_i nodes at any location (θ_i,ϕ_i) on our domain with phase velocities at this node $C_i(\theta_i,\phi_i)$. The last two (2) model parameters (N_i,σ_c^T) are the hierarchical parameters that adjust resolution and prescribe the amount of observational or modeling uncertainty present in our forward problem, $g(\mathbf{X})$, using the likelihood:

$$P(\mathbf{d}=t_{ij}|\mathbf{X}) = \exp\left[-\frac{(g(\mathbf{X})-\mathbf{d})^2}{2\sigma_c^T}\right]. \tag{12}$$

Here $\mathbf{d}$ is our data vector of interstation travel-time observations from c_{ij} at a given period: $\mathbf{d}=t_{ij}=c_{ij}/r_{ij}$ for interstation distance r_{ij}. In this way, errors in our measurements and any simplifications in wave propagation can be incorporated into σ_c. A final map is obtained by averaging over all likely models:

$$c^L_{T=40}(\theta,\phi) = \sum_{k=b+\Delta_k}^{k=N_k} \mathbf{I}(C_{ik},\theta_{ik},\phi_{ik},\theta,\phi), \tag{13}$$

in which $\mathbf{I}$ is an interpolating function that uses the Voronoi nodes to generate a high-resolution pixel map over our entire spatial domain. Averaging is done after discarding b burning steps and resampling every Δ_k steps to decorrelate samples before model averaging. A full discussion of our transdimensional approach and a comprehensive compilation of all phase maps is provided in a companion article.

Results

The Love and Rayleigh dispersions are the largest compiled on the continent since the commencement of digital recordings (Fig. 1). A detailed analysis of this catalog can be best understood in broad terms, with some key targets. Therefore, we organize our results in the following way: (1) we present a qualitative description of continent-wide surface-wave response by highlighting the effect of interstation distance on signal quality; (2) we summarize the trends in our measurement using a statistical approach, illustrating the distribution of our dispersion curves (and their uncertainties) and highlighting new resolution at the shortest periods (>32 mHz); and (3) we present a comparison of our Love-wave phase maps with global earthquake maps showing long-wavelength similarities as well as clear improvements at shorter spatial wavelengths.

TABLE 1.
Summary of Distribution of Phase Dispersion Measurements Sorted by Distance

	Love (Transverse)											
	Short: $r<r_s$		Long: $r>r_s$		Short: $r<r_s$		Long: $r>r_s$		$C_f^L - C_0^L$		$2\sqrt{\sigma_f^2}$	
T^*	$\bar{C}_0^L$	N_o†	$\bar{C}_0^L$	N_o†	$\bar{C}_f^L$	N_o†	$\bar{C}_f^L$	N_o†	$r<r_s$	$r>r_s$	$r<r_s$	$r>r_s$
5	**3.2058**	**27,721**	3.4504	66,800	**3.1978**	**29,797**	3.4523	73,136	**0.0030**	−0.0019	**0.0769**	0.0120
6	**3.2915**	**28,669**	3.5170	66,800	**3.2812**	**29,797**	3.5198	73,136	**−0.0162**	−0.0004	**0.0215**	0.0027
8	**3.4239**	**28,870**	3.6112	66,800	**3.4159**	**29,797**	3.6134	73,136	**−0.0162**	−0.0004	**0.0320**	0.0036
10	**3.5343**	**28,875**	3.6921	66,800	**3.5268**	**29,797**	3.6938	73,136	**−0.0162**	−0.0004	**0.0455**	0.0049
12	**3.6440**	**28,861**	3.7715	66,800	**3.6359**	**29,797**	3.7729	73,136	**−0.0162**	−0.0004	**0.0648**	0.0062
15	**3.8026**	**28,782**	3.8929	66,800	**3.7931**	**29,797**	3.8932	73,136	**−0.0131**	−0.0003	**0.1100**	0.0092
20	**3.9978**	**28,305**	4.0581	66,800	**3.9877**	**29,797**	4.0569	73,136	**−0.0124**	−0.0002	**0.2254**	0.0166
25	**4.1148**	**27,186**	4.1665	66,800	**4.1042**	**29,797**	4.1655	73,136	**−0.0124**	−0.0002	**0.4026**	0.0285
30	**4.2063**	**25,377**	4.2535	66,796	**4.1940**	**29,797**	4.2531	73,136	**−0.0124**	−0.0002	**0.7002**	0.0475
35	**4.2819**	**20,469**	4.3251	66,650	**4.2659**	**29,797**	4.3255	73,136	**−0.0124**	−0.0002	**0.3139**‡	0.0730
40	**4.3277**	**14,405**	4.3725	64,775	**4.3198**	**29,797**	4.3812	73,136	**−0.0700**	0.0075	**0.7084**‡	0.1783

The number of observations: N_o, mean phase velocity for each wave type for Love using zero crossing C_0 or nonlinear waveform fitting C_f. The average standard deviation σ_f at eleven (11) distinct periods, T^*. We report the 95% confidence interval: $2\sqrt{\sigma_f^2}$. Short interstation distance ($r < r_s$) are shown in bold.

*Our shortest period measurements are 3.0 s; however, not all paths have reliable phase dispersion data at such short periods. So, here, we report measurements longer than 5 s. The comprehensive catalog will include some shorter period measurements for all interstation paths.

†The catalog contains $N = 114{,}487$ interstation connections representing ~73% of our initial database of 153,029 connections. Failure to obtain cross-spectra is a function of data and metadata errors. We were unable to process waveforms from ~6% (82 stations) of our total 1458 stations. See the supplemental material for further details.

‡Using the median instead of mean.

Empirical Green's function for Love and Rayleigh

The time-domain noise correlation functions show visible Rayleigh (ZZ) and Love (TT) waves across our entire arrays. The image is an approximate empirical Green's function, and it is obtained by stacking all the NCFs for similar interstation distances across the entire seismic array. We observe clear detection of both Rayleigh and Love waves for all interstation distances. Compared to Love waves, Rayleigh waves are more visible at longer distances (>3000 km). Love waves, though detectable, are weaker, especially when station separation exceeds ~3000 km. When station separation is less than 2000 km, we observe the best SNR for both the surface-wave types. Signal quality degrades between 2000 and 3000 km, with the lowest SNR being observed at distances >3000 km (Fig. 5).

Although we show time-domain NCFs here, we emphasize that the actual SNR and the dispersion measurements are computed in the frequency domain (e.g., Fig. 3). We observe very high SNR (>5) for all the frequency bands, even for short duration stacks (2–3 months, see Fig. S3, available in the supplemental material to this article). The detection of high-quality frequency-domain NCFs for short duration stacks has ~~previously~~ been documented by previous authors (e.g., Prieto *et al.*, 2009; Seats *et al.*, 2012) and justify our preference for frequency-domain estimation of surface-wave dispersion. In addition to fundamental-mode Rayleigh and Love waves, we also observe body-wave phases in our vertical cross-correlation image (Fig. 5a), although we do not explore them any further in this study.

Model assessment: path-average model

A comparison between the final path-averaged model $\mathbf{m}_f$ and the reference Litho1.0 model $\mathbf{m}_0$ provides a preliminary assessment of the information content present in our NCFs. We demonstrate this by showing two sets of exemplary path-averaged models obtained from station pairs that are located on the Tanzanian craton and the Kalahari craton (Fig. 6a,b). Although both stations are located on a similar geological setting, and their reference models are similar, we observe, however, that after fitting the NCF measurements, the models are distinctly different. For the station pair on the Tanzanian craton, $\mathbf{m}_f$ is similar to $\mathbf{m}_0$, with very little differences in the density, elasticity, and anisotropy across the entire crustal column (Fig. 6a).

For the station pair on the Kalahari craton, we observe a striking difference between $\mathbf{m}_f$ and $\mathbf{m}_0$. The ratio of compressional to shear wave velocity in the top crust differs from the value prescribed by the reference model. Similarly, a radially anisotropic model is strongly preferred over the isotropic reference model for the entire crustal column (Fig. 6b). Slight but noticeable differences are also recovered from the density model. These results are also evident in the data (compare

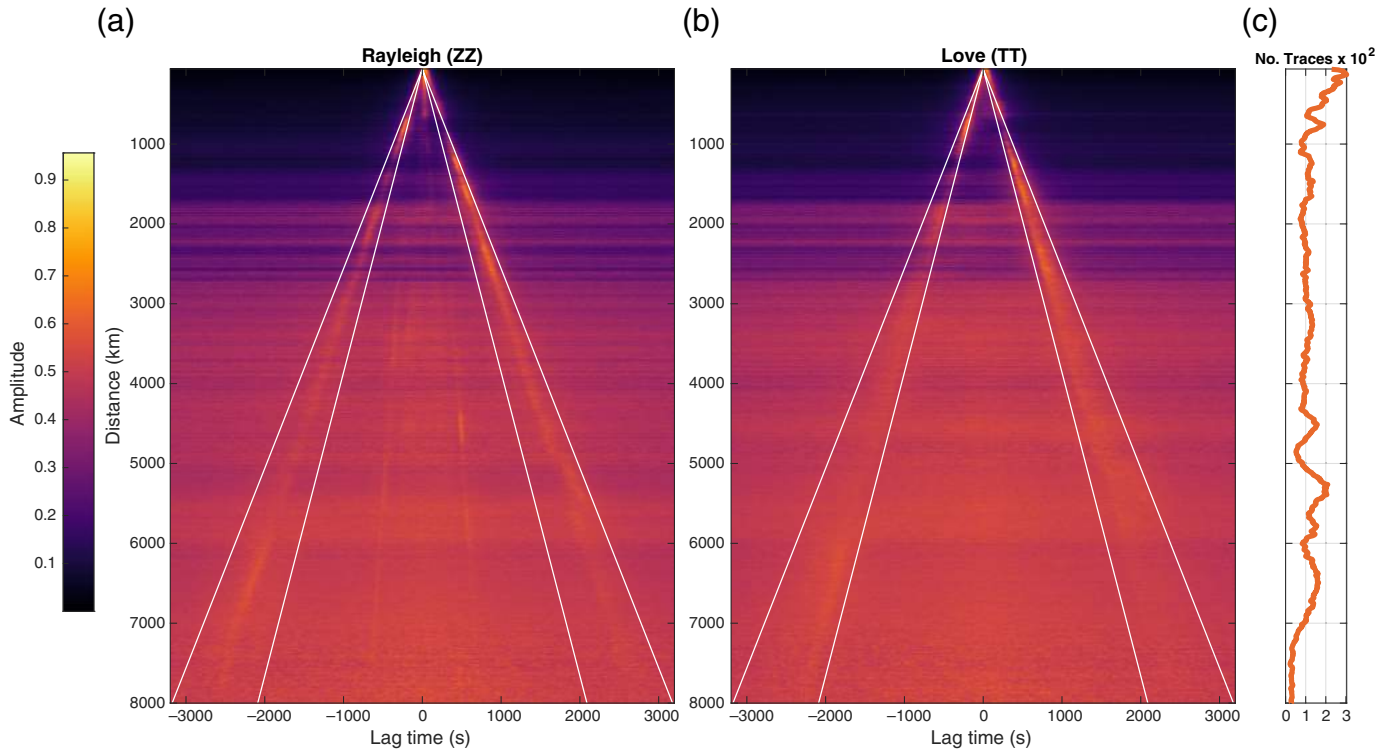

Figure 5. Short-period surface waves across the African continent obtained from filtered and stacked time-domain NCFs. (a) Rayleigh waves from vertical waveforms (ZZ) (b) Love waves from stacked transverse waveforms (TT). (c) Number of waveforms stacked at each interstation distance. For a close up view of the short-distance empirical Green's functions, see Figure S2. For a frequency-dependent estimation of SNR, see Figure S3. The color version of this figure is available only in the electronic edition.

Fig. 3e and 3f). This illustrates one example of how our dataset can be used for model assessment. The crucial improvement is the improved resolution, especially in the topmost crust.

Phase dispersion distribution

Our measurements include both Love- and Rayleigh-wave dispersion across different geological domains on the continent. The distribution of these measurements shows that we can extend ambient noise surface-wave measurements to very short periods, that is from 40 s all the way to 5 s ($\sim 10^4$ measurements on average), even in the presence of variable signal quality (Fig. 7). The dispersion curves vary smoothly, suggesting that they are consistent. Phase dispersion curves obtained from short interstation distances are less robust to outliers (Fig. 7a,b). This is evident by the unusually high velocities at the longest periods (>5 km/s at 40 s) and is clearly reflective of defective measurements. These outliers are not present ~~in~~ when the stations are separated farther apart. In our final catalog, we flag measurements as outliers when they exceed reasonable bounds of long-period measurements (Ekström, 2014) (Fig. 8c).

In addition to outliers, we do not observe systematic biases between the two methods of obtaining phase dispersion measurements. The phase velocities obtained by nonlinear waveform fitting are often not statistically distinguishable from those obtained by the zero crossing technique (Fig. 8). Clues as to why this is the case can be seen in the underlying NCF measurements on which the phase dispersions are based. When the NCF signal quality degrades, errors propagate into the estimates of phase velocity (see Fig. 3b,f). Fitting the zeros in this case becomes more problematic, and refinements from waveform fitting and constraints on continuity of the curve improve the estimates. When signal quality is good, as is the case for most of our measurements, such adjustments are not needed. Our final catalog report results from both techniques and can be used as complementary information for quantifying uncertainties in the reported phase dispersion measurements.

We observe a clear signature of very slow velocities, at the shortest periods, for both the surface-wave types (Fig. 7). This is most likely due to wave propagation through shallow sedimentary basins across the continent. The detection of the slow phase velocities, regardless of station separation, is a strong indication that this signal is structural and not stochastic. Spatial distribution of slow sedimentary basins will be explored in a companion article and should be clearly discernible in the phase velocity maps at the shortest periods (<13 s). We point out that the short-period measurements are a new contribution of our catalog, and improved spatial resolution of phase velocity maps at the same periods as earthquakes provides new insights for regions on the continent that have not been well sampled in the past. A comparison of the Love phase velocity

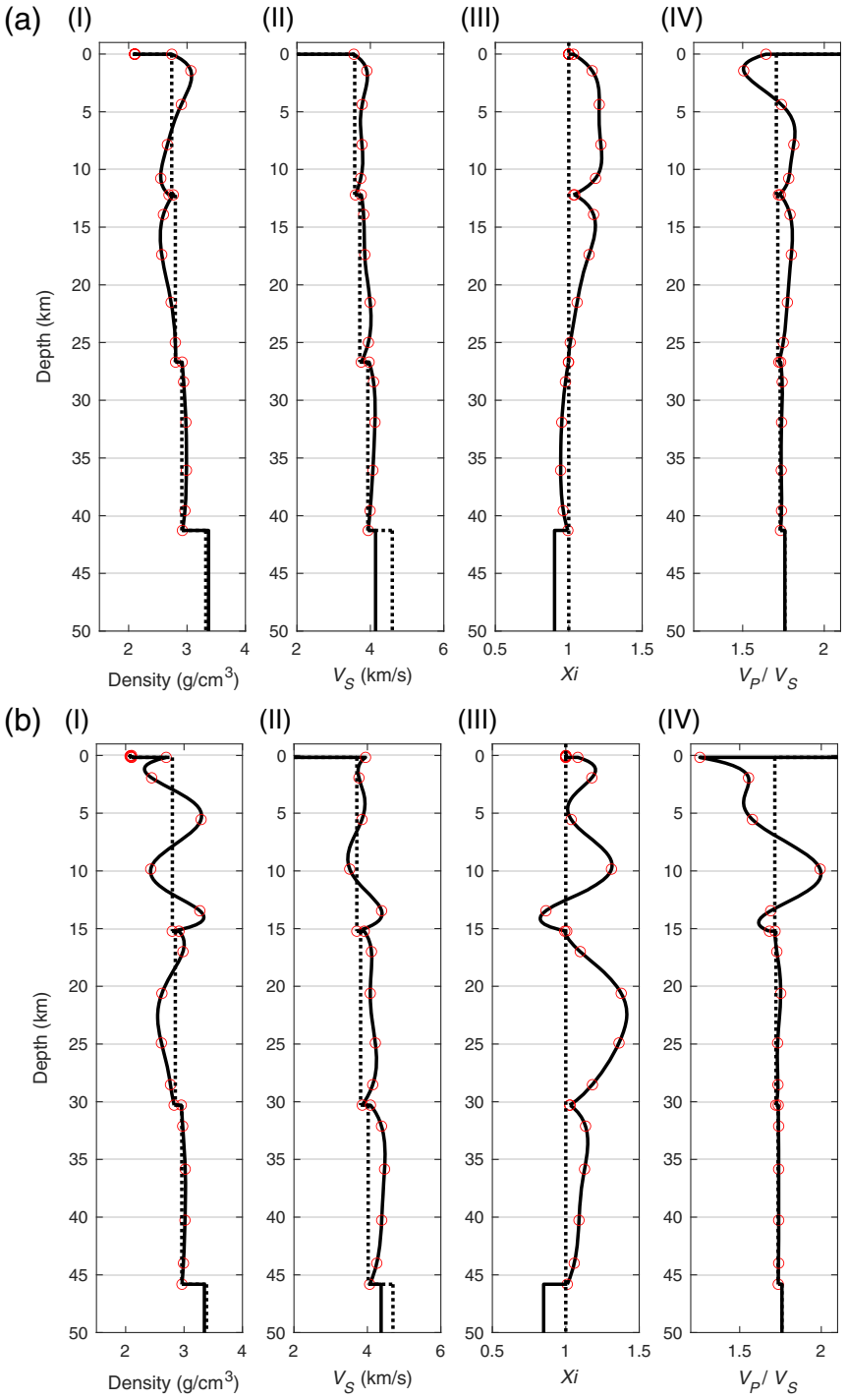

Figure 6. (a) Reference and final crust models for the station pair RUNG and MTAN (network XD on the Tanzanian craton). (I) Density, (II) shear velocity, V_S, (III) radial anisotropy, Xi, and (IV) ratio of shear to compressional velocity, V_P/V_S. The reference model $\mathbf{m}_0$ (dashed line) is the layered Litho1.0 model, while the final models $\mathbf{m}_f$ (solid line in ~~a–d~~) are derived from the joint waveform fitting of the Love and Rayleigh NCF data (Fig. 3c and equation 7) showing the nodes (red dots) used in the spectral elements description. (b) Same as Figure 6a, but for station pairs SA53 and SA58 (network XA on the Kalahari Craton). ~~Discernible~~ differences between $\mathbf{m}_0$ and $\mathbf{m}_f$ are detectable in all of the model parameters suggesting a unique crustal structure underneath this location not previously captured by Litho1.0. The color version of this figure is available only in the electronic edition.

map obtained at the same period as a globally sampled dispersion model obtained from earthquakes illustrates this point.

Exemplary phase map for love waves

We verify the consistency of our phase dispersion measurements by constructing a 40 s Love-wave phase velocity map using the Bayesian approach, described in the Phase Maps using a Bayesian Auto-Adaptive Approach section, and compare it with global dispersion maps derived from earthquake measurements. We chose two representative dispersion models for our comparison: the GDM52 model developed by Ekström (2011) and the Scripps model (Ma *et al.*, 2014) used to construct the Litho1 global lithospheric velocity model (Pasyanos *et al.*, 2014).

The long-wavelength patterns in our Love map are broadly consistent with those obtained from global earthquake studies (Fig. 9). The range and distribution of phase velocities are broadly consistent between our ambient noise and the earthquake measurements—faster than average phase velocities in Madagascar, south of the Kaapvaal-Zimbabwe craton (KPC) and the West African Craton (WAC), compared to the rest of the continent, and slower than average phase velocities along the East African and Ethiopian Plateaus, and in the Atlas mountain ranges of Northern Africa. Our map also shows distinct patterns of short-wavelength features that are spatially coherent with geological and topographic features.

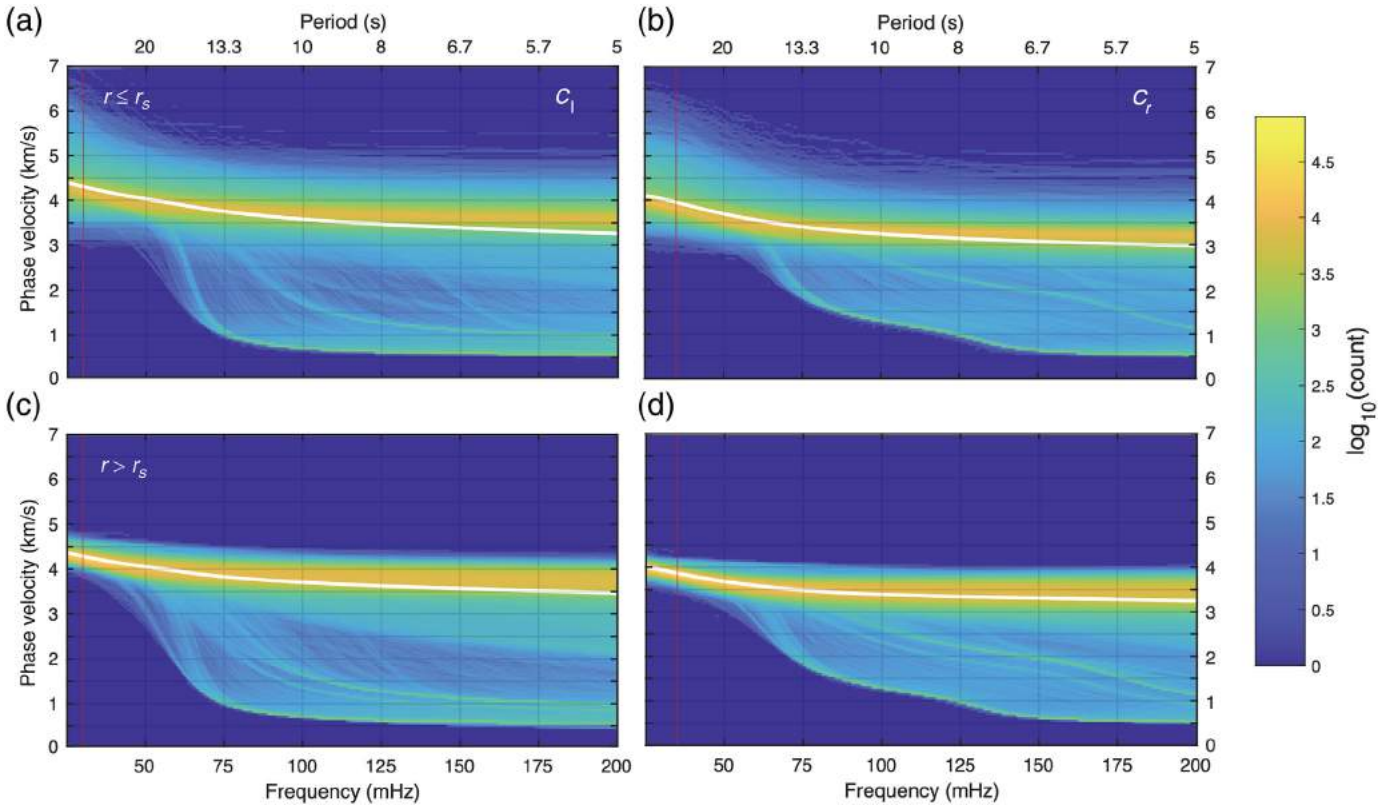

Figure 7. Distribution of phase dispersion curves, c_f, sorted by interstation distance r. (a,b) Distribution for short interstation distance: $r < r_s$ (c,d) Distribution for longer interstation distances $r > r_s$. (a,c) Love and (b,d) Rayleigh waves (100,000 measurements). The color version of this figure is available only in the electronic edition.

The most clearly resolved is the low-velocity feature in the Congo craton, which is spatially coherent with the Congo basin (visible as a low topography and geoid anomaly: Raveloson *et al.*, 2015, 2021). Other striking features are the slow velocities within the Sahara metacraton and the fast and slow velocity features on the West African craton. The value of the Bayesian approach is clearly evident in the fact that we can evaluate, using the uncertainties, which of these features are well resolved (Fig. 9d). We observe clearly that all, except for the West African craton, is well resolved. Regardless, their spatial extents, and the confidence limits, are clearly constrained by the data and have previously been undetected. We point out that this is only a preliminary presentation of the value of our new measurements for improved resolution of geological and tectonic features across the entire African continent. In a companion article, we will present the complete set of dispersion maps and their utility for investigating the crustal architecture of the African continent.

Discussion

Improved spatial and depth resolution

The phase dispersion catalog we describe in this study improves upon previous work in the following key respects: (1) it is a continent-wide study with the most comprehensive dataset available on public archives, (2) it extends to short periods (20–5 s), (3) it provides a drastic improvement in ray-path density and geometry with measurements totaling $\sim 10^5$, (4) it includes the Love-wave response at short periods, and (5) it provides estimates of uncertainties useful in follow-on studies of crust and upper-mantle structure. In comparison, the most recent continent-wide ambient noise study of Africa provides ~1500 interstation Rayleigh-wave dispersion measurements focused on periods >40 s (see fig. S4 of Emry *et al.*, 2018). Shorter-period measurements and improved ray-path geometry provide better spatial and depth resolution of the bulk crust. In particular, improvements in ray coverage across gaps on the West African craton, Congo craton, and the Sahara metacraton enable better illumination of these regions. As our phase maps show, even at the longest periods (~40 s), our results are not only broadly consistent with those obtained from earthquakes, but they also provide better spatial resolution. This is expected because the spacing, geometry, and density of our ray-path coverage result in better spatial sampling compared to earthquake station geometries used in global studies (see Fig. 1). In other words, even for the best global dispersion models, the ray paths are often long wavelengths, with ray-path density two orders of magnitude less than provided in our catalog (Ekström, 2011; Ma and

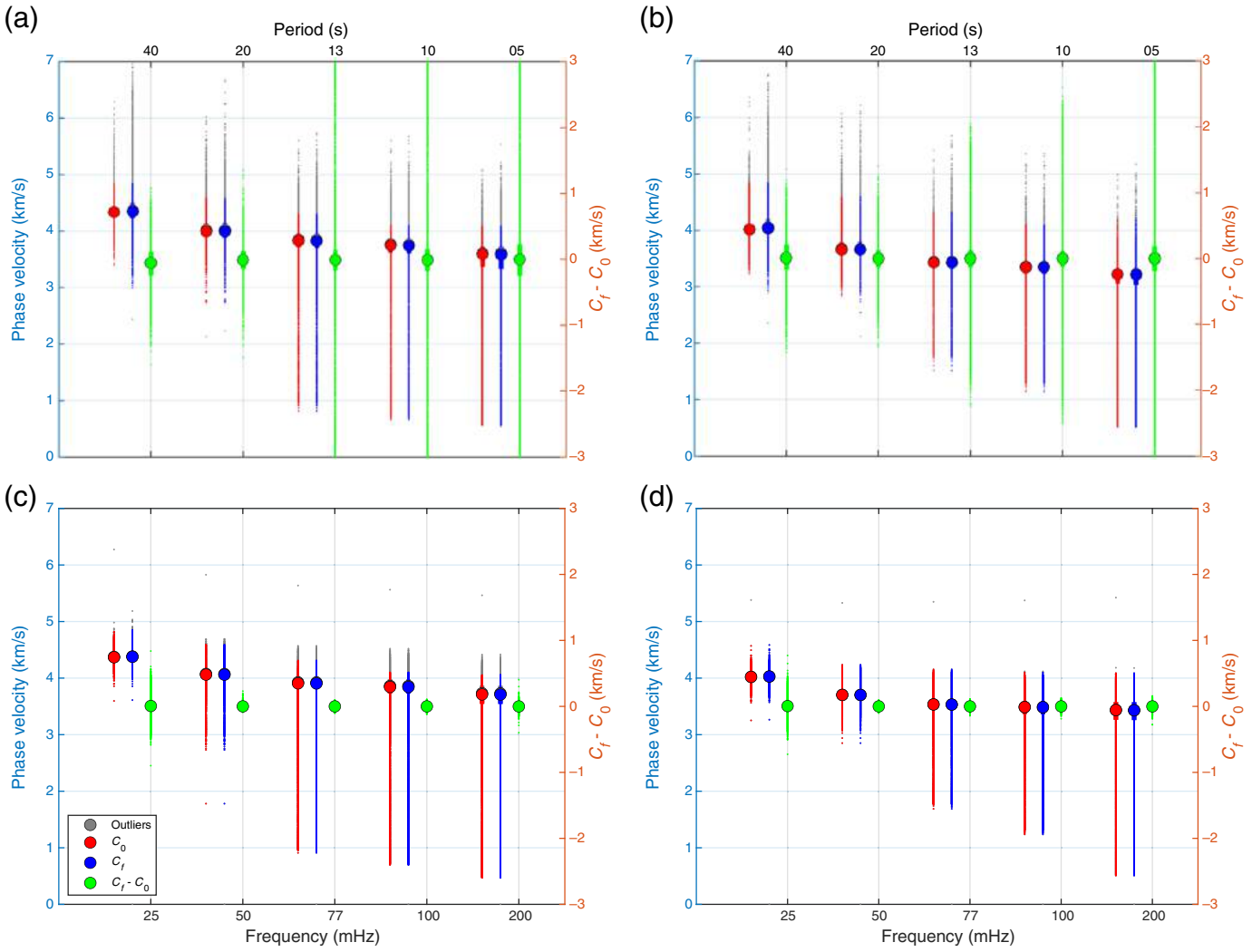

Figure 8. Detection of outliers and bias of different measurement techniques. (a,b) Distribution of phase dispersion curves using the technique of zero crossing, c_0 (red) and the nonlinear fitting c_f (blue) for short interstation distance: $r < r_s$. (c,d) Distribution for longer interstation distances $r > r_s$. The row organization is similar to Figure 7—(a,c) Love and (b,d) Rayleigh. Outliers are flagged as high phase velocities at the longest periods (gray dots) especially visible for Love waves on small aperture station pairs. Little bias in phase velocities depending on technique is visible in the distribution of $c_f - c_0$ (green) for station pairs. See Table 1 for a full description of our measurement distributions. The color version of this figure is available only in the electronic edition.

Masters, 2014; Moulik *et al.*, 2021). Taken together, these new measurements will produce improvements to velocity models of the crust and upper mantle, with new insights into the role radial anisotropy may play in shaping the rheology and strength of the crust.

Sources of error: love waves, signal quality and installation

The source of error in our phase dispersion measurements are primarily controlled by two key factors: (1) the station separation, and (2) the SNR of the frequency-domain NCFs affected either by (a) a short-duration network deployment or (b) a low-quality seismic recording especially on the horizontal channel. Inspection of the time-domain NCFs (Fig. 5) shows that, on average, the SNR is higher when the stations are spaced closer, which may initially suggest that the standard deviation should be lower for shorter station spacing. However, the opposite behavior is observed. In general, larger standard deviations, and more outliers, are detected when the station separation is small compared to when they are further apart (see Fig. 8 and Tables 1–3). This is a well-documented observation where phase velocity uncertainties are lower for longer station separation distances (Ekström *et al.*, 2009; Luo *et al.*, 2015, 2020). An inspection of the Aki formulation (equations 3 and 4) provides a simple explanation for why this is the case. When the station separation is only within a few wavelengths, the Bessel function (and the NCFs) has only a few cycles, and tracking the zero crossings or fitting

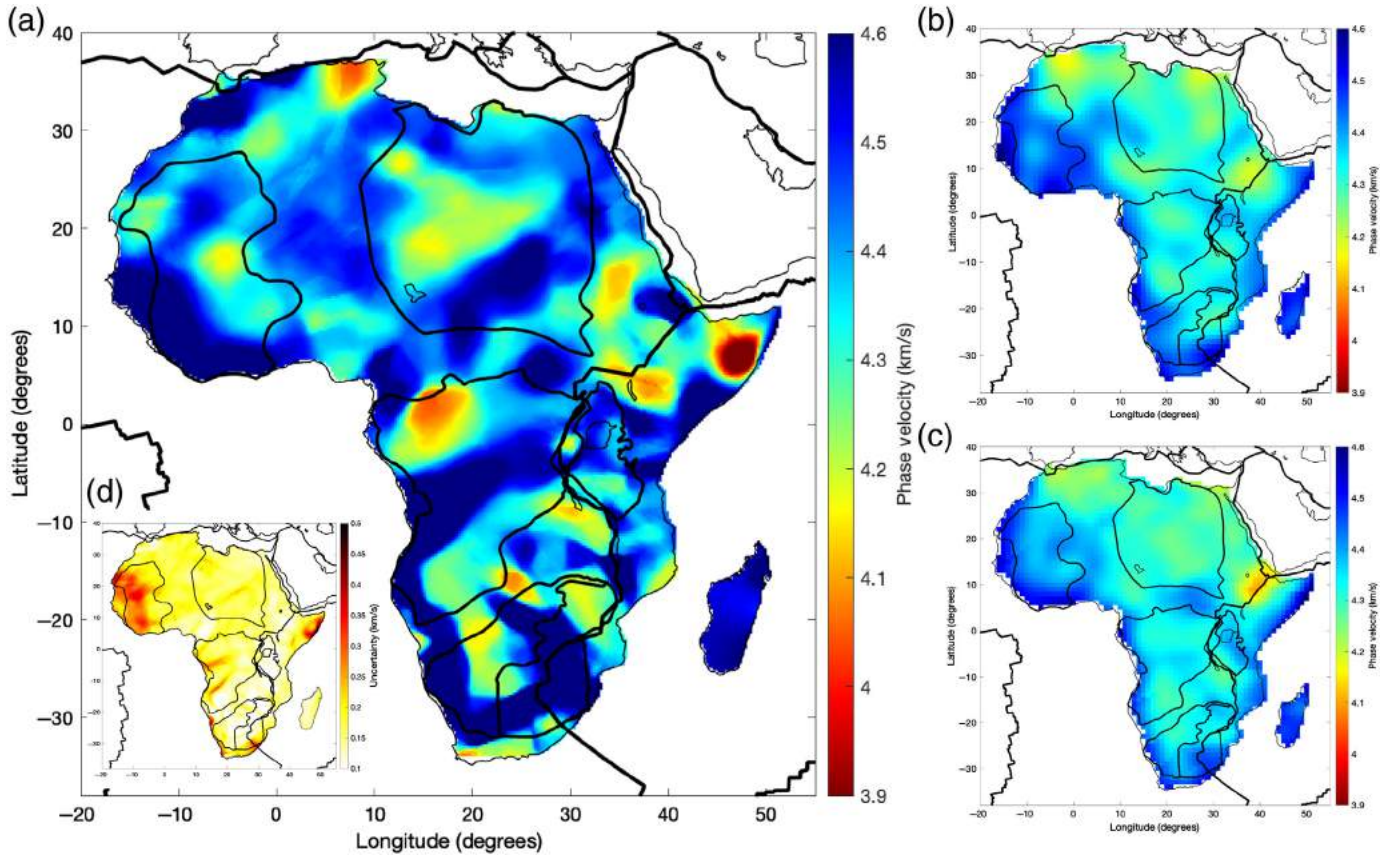

Figure 9. Bayesian construction of Love-wave phase map and uncertainties at 40 s using ADAMA dataset. (a) Most likely solution obtained from average of a large McMC ensemble (2.5×10^6 models using $\sim 10^5$ measurements). (b) Love-wave map derived from global earthquake measurements (GDM52 from Ekström, 2011). (c) Predictions from Litho1 model constructed using earthquake measurements of Ma *et al.* (2014). (d) Error map estimated from McMC ensemble. See Figure S4 for maps without removing outliers and using only 30,000 measurements. The color version of this figure is available only in the electronic edition.

the waveform amplitude becomes more difficult, leading to larger uncertainties in the final dispersion measurements (Ekström, 2014, 2017). Similarly, at the longest periods, it may be harder to determine the locations of the first few zeros (or peaks or troughs), especially when the amplitude of the NCFs are small, due to weak excitation at these periods. This explains why uncertainties are greater in these cases.

The second reason for large errors, especially for Love waves, is low SNR, caused by noisy seismograms from poor installations. It is evident from the way the NCFs are computed that the SNR of the final stack is strongly influenced by the quality of the seismograms present in the stack or the duration of the stack (number of records used to compute $\bar{\rho}$: equations 1 and 2). When station installation is poor, as is often the case for short-term deployments, this can lead to very low SNR and higher uncertainties. This explains the larger likelihood of detecting outliers in the Love-wave measurements. We observe that most of the defective measurements are from a few stations all from short-duration network deployments (see Fig. 10, compare to Fig. 1). For example, for two network installations with the same duration of $\sim$ 1 yr (XD and XA), they provide different SNRs, because the horizontal seismograms are noisier at XA compared to XD (Fig. 3 and Fig. S1). This means that for stations that have noisier seismograms in general, phase dispersion uncertainties will be larger, even when paired with seismograms obtained from good stations.

Africa's crustal architecture

The short-period surface-wave catalog presented here is useful for assessment and updates of regional and global seismic models (Olugboji *et al.*, 2017). We have demonstrated better illumination of geological domains not previously well sampled (i.e., Congo basin, Sahara Metacraton, West African craton). Improved lateral and depth resolution will advance our understanding of the tectonic and geodynamic evolution of the African lithosphere, revealing the architecture of the longest-lived cratons on our planet, and the structure of their deep lithospheric roots. This can be done using our catalog in integrated modeling, alongside complementary seismological and geophysical datasets (Afonso *et al.*, 2008; Young *et al.*, 2013; Calò *et al.*, 2016; Raveloson *et al.*, 2021). Data-driven geophysical models will advance our

TABLE 2
Same as Table 1 but for Rayleigh Waves, $\bar{C}^R_{0,f}$, Obtained from Cross-Spectra of Vertical Seismograms (ρ^R_{ZZ})

	Rayleigh (Vertical)											
	Short: $r<r_s$		Long:$r>r_s$		Short: $r<r_s$		Long:$r>r_s$		$C^R_f-C^R_0$		$2\sqrt{\sigma^2_f}$	
T	$\bar{C}^R_0$	N_o	$\bar{C}^R_0$	N_o	$\bar{C}^R_f$	N_o	$\bar{C}^R_f$	N_o	$r<r_s$	$r>r_s$	$r<r_s$	$r>r_s$
5	**2.9714**	**29,405**	3.2721	72,820	**2.9718**	**32,135**	3.2713	78,878	**0.0122**	–0.0011	**0.0767**	0.0097
6	**3.0427**	**30,733**	3.3097	72,820	**3.0366**	**32,135**	3.3097	78,878	**–0.0092**	–0.0004	**0.0237**	0.0034
8	**3.1529**	**31,006**	3.3667	72,820	**3.1468**	**32,135**	3.3670	78,878	**–0.0092**	–0.0004	**0.0321**	0.0046
10	**3.2472**	**31,008**	3.4159	72,820	**3.2416**	**32,135**	3.4162	78,878	**–0.0092**	–0.0004	**0.0409**	0.0058
12	**3.3398**	**30,995**	3.4690	72,820	**3.3339**	**32,135**	3.4693	78,878	**–0.0092**	–0.0004	**0.0516**	0.0071
15	**3.4816**	**30,954**	3.5610	72,820	**3.4750**	**32,135**	3.5608	78,878	**–0.0075**	–0.0002	**0.0730**	0.0098
20	**3.6807**	**30,749**	3.7026	72,820	**3.6742**	**32,135**	3.7022	78,878	**–0.0066**	–0.0002	**0.1430**	0.0166
25	**3.8394**	**30,054**	3.8228	72,820	**3.8337**	**32,135**	3.8229	78,878	**–0.0066**	–0.0002	**0.2397**	0.0209
30	**3.9456**	**27,633**	3.9157	72,817	**3.9486**	**32,135**	3.9162	78,878	**–0.0066**	–0.0002	**0.4081**	0.0313
35	**4.0153**	**22,428**	3.9826	72,677	**4.0255**	**32,135**	3.9834	78,878	**–0.0066**	–0.0002	**0.2902***	0.0517
40	**4.0499**	**15,971**	4.0215	70,647	**4.0736**	**32,135**	4.0283	78,878	**0.0298**	0.0053	**0.7594***	0.1472

*Using the median instead of mean.

TABLE 3
Same as Table 1 but for Rayleigh Waves Obtained from Cross-Spectra of Radial Seismograms (ρ^R_{RR})

	Rayleigh (Radial)											
	Short: $r<r_s$		Long: $r>r_s$		Short: $r<r_s$		Long: $r>r_s$		$C^R_f-C^R_0$		$2\sqrt{\sigma^2_f}$	
	$\bar{C}^R_0$	N_o	$\bar{C}^R_0$	N_o	$\bar{C}^R_f$	N_o	$\bar{C}^R_f$	N_o	$r<r_s$	$r>r_s$	$r<r_s$	$r>r_s$
5	**2.9449**	**32,703**	3.2756	79,450	**2.9468**	**34,285**	3.2745	79,447	**0.0149**	-0.0011	**0.0767**	0.0097
6	**3.0105**	**34,074**	3.3135	79,450	**3.0079**	**34,285**	3.3129	79,447	**-0.0021**	-0.0003	**0.0240**	0.0034
8	**3.1140**	**34,337**	3.3709	79,450	**3.1113**	**34,285**	3.3704	79,447	**-0.0021**	-0.0003	**0.0322**	0.0046
10	**3.2018**	**34,340**	3.4203	79,450	**3.1995**	**34,285**	3.4199	79,447	**-0.0021**	-0.0003	**0.0405**	0.0058
12	**3.2875**	**34,338**	3.4734	79,450	**3.2852**	**34,285**	3.4732	79,447	**-0.0021**	-0.0003	**0.0508**	0.0071
15	**3.4190**	**34,333**	3.5652	79,450	**3.4164**	**34,285**	3.5651	79,447	**-0.0022**	-0.0002	**0.0719**	0.0098
20	**3.6049**	**34,260**	3.7075	79,450	**3.6011**	**34,285**	3.7076	79,447	**-0.0040**	0.00005	**0.1416**	0.0167
25	**3.7549**	**33,760**	3.8293	79,450	**3.7498**	**34,285**	3.8294	79,447	**-0.0040**	0.00005	**0.2402**	0.0209
30	**3.8639**	**31,869**	3.9236	79,444	**3.8597**	**34,285**	3.9237	79,447	**-0.0040**	0.00005	**0.4120**	0.0313
35	**3.9425**	**26,538**	3.9916	79,289	**3.9352**	**34,285**	3.9917	79,447	**-0.0040**	0.00005	**0.2988***	0.0518
40	**3.9830**	**19,107**	4.0313	77,043	**3.9833**	**34,285**	4.0371	79,447	**-0.0240**	0.0052	**0.7793***	0.1475

*Using the median instead of mean.

understanding of (1) the origin of Africa's elevated topography, (2) the broad-scale evolution of continental rifting and formation of new ocean basins, and (3) the multistage assembly and break up of Gondwanaland. Future work will include a compilation of phase and group velocity maps with associated uncertainty using the Bayesian inverse approach introduced in this study (Tilmann *et al.*, 2019; Zhang *et al.*, 2019).

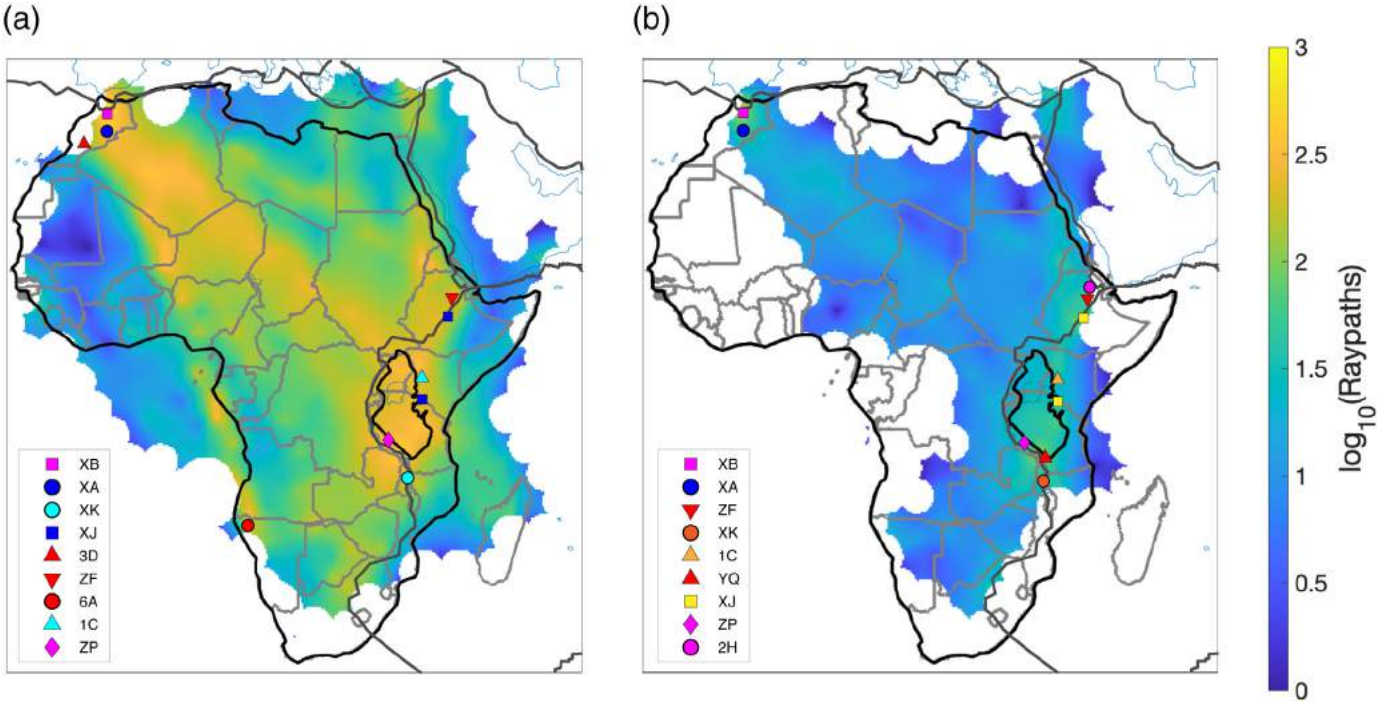

Figure 10. Spatial distribution and network location of outlier measurements (Love and Rayleigh). (a) Ray-path density of outlier Love measurements. (b) Ray-path density for outlier Rayleigh measurements. 80% of the outliers are from stations from eight networks (symbols on map). More outliers are detected on Love compared to Rayleigh even in cases when the noisy stations are similar (XB, XA, and XK). The color version of this figure is available only in the electronic edition.

Conclusions

We have compiled a large catalog of short-period Love- and Rayleigh-wave dispersion measurements with associated uncertainties. We extracted phase and group dispersion measurements from stacked interstation cross-spectra from ~1 million seismograms derived from all publicly available three-component seismograms recorded on seismic networks in and around Africa. This involved using both the phase and amplitude information present in our NCF cross-spectra: zero crossings, peak, troughs, and nonlinear waveform fitting of the amplitudes using a starting path-averaged model. This approach is robust to low-quality cross-spectra estimates, especially Love waves, and low-quality waveforms obtained for short-duration network deployments. An exemplary phase velocity map constructed using Bayesian inversion of our dispersion measurements shows statistically significant features with improved spatial and depth resolution. We expect that our dataset and model assessment product, ADAMA.D1, will be used to investigate important geophysical questions across the diversity of geological terranes on the African continent.

Data and Resources

All seismic data used in this study can be obtained from the Incorporated Research Institutions for Seismology (IRIS) Data Management Center (https://ds.iris.edu/ds, last accessed December 2021). Horizontal orientations for individual stations are obtained from Ojo *et al.* (2019) or using the DLOpy (Doran and Laske, 2017) code available on GitHub. The software, Aki Estimate, developed by Hawkins and Sambridge (2019) was used for extracting phase and group dispersions from noise cross-spectra. The supplemental material for this article includes Figures S1, S2, S3 and S4, as well as a description of the entire catalog that is published on the open source repository at doi: 10.5281/zenodo.5753071.

Declaration of Competing Interests

The authors acknowledge that there are no conflicts of interest recorded.

Acknowledgments

This work was made possible by a National Science Foundation Grant 2102495. The authors Siyu Xue and Tolulope Olugboji acknowledge financial support from the Furth Award at University of Rochester. Computational resources of the BlueHive Linux cluster were indispensable to the results obtained in this work. The authors also acknowledge many helpful discussions with Baowei Liu, Yuri Tamama, Zhijie Ji, Ziqi Zhang, Jean-Joel Legre, and Steve Carr. The authors thank two anonymous reviewers and the editors whose recommendation helped improve this work.

References

Accardo, N. J., J. B. Gaherty, D. J. Shillington, C. J. Ebinger, A. A. Nyblade, G. J. Mbogoni, P. R. N. Chindandali, R. W. Ferdinand, G. D. Mulibo, G. Kamihanda, *et al.* (2017). Surface wave imaging of the weakly extended Malawi Rift from ambient-noise and teleseismic Rayleigh waves from onshore and lake-bottom seismometers, *Geophys. J. Int.* **209,** no. 3, 1892–1905, doi: 10.1093/gji/ggx133.

Afonso, J. C., M. Fernàndez, G. Ranalli, W. L. Griffin, and J. A. D. Connolly (2008). Integrated geophysical-petrological modeling of the lithosphere and sublithospheric upper mantle: Methodology and applications, *Geochem. Geophys. Geosys.* **9,** no. 5, doi: 10.1029/2007gc001834.

Aki, K. (1957). Space and time spectra of stationary stochastic waves, with special reference to microtremors, *Bull. Earthq. Res. Inst.* **35,** 415–456.

Begg, G. C., W. L. Griffin, L. M. Natapov, S. Y. O'Reilly, S. P. Grand, C. J. O'Neill, J. M. A. Hronsky, Y. Poudjom Djomani, C. J. Swain, and T. Deen (2009). The lithospheric architecture of Africa: Seismic tomography, mantle petrology, and tectonic evolution, *Geosphere* **5,** no. 1, 23–50, doi: 10.1130/GES00179.1.

Bodin, T., M. Sambridge, N. Rawlinson, and P. Arroucau (2012). Transdimensional tomography with unknown data noise, *Geophys. J. Int.* **189,** no. 3, 1536–1556, doi: 10.1111/j.1365-246X.2012.05414.x.

Brink, T. J., L. P. Moser, and T. M. Olugboji (2019). Seismic signature of highly resistive south-eastern North American lithosphere, *Fall Meeting 2019, AGU*, Abstract T21G-0427, available at https://ui.adsabs.harvard.edu/abs/2019AGUFM.T21G0427B.

Burke, K. (1996). The African plate, *S. Afr. J. Geol.* **99,** no. 4, 341–409.

Calò, M., T. Bodin, and B. Romanowicz (2016). Layered structure in the upper mantle across North America from joint inversion of long and short period seismic data, *Earth Planet. Sci. Lett.* **449,** 164–175, doi: 10.1016/j.epsl.2016.05.054.

Christensen, N. I., and W. D. Mooney (1995). Seismic velocity structure and composition of the continental crust: A global view, *J. Geophys. Res.* **100,** 9761–9788, doi: 10.1029/95JB00259.

Doran, A. K., and G. Laske (2017). Ocean-bottom seismometer instrument orientations via automated Rayleigh-wave arrival-angle measurementsocean-bottom seismometer instrument orientations via automated Rayleigh-wave arrival-angle measurements, *Bull. Seismol. Soc. Am.* **107,** no. 2, 691–708, doi: 10.1785/0120160165.

Durrheim, R. J., and A. A. Nyblade (2019). AfricaArray studies of the lithosphere in sub-Saharan Africa, *Acta Geol. Sinica* **93,** 22–24.

Ekström, G. (2011). A global model of Love and Rayleigh surface wave dispersion and anisotropy, 25-250s, *Geophys. J. Int.* **187,** 1668–1686, doi: 10.1111/j.1365-246X.2011.05225.x.

Ekström, G. (2014). Love and Rayleigh phase-velocity maps, 5–40 s, of the western and central USA from USArray data, *Earth Planet. Sci. Lett.* **402,** 42–49, doi: 10.1016/j.epsl.2013.11.022.

Ekström, G. (2017). Short-period surface-wave phase velocities across the conterminous United States, *Phys. Earth Planet. In.* **270,** 168–175, doi: 10.1016/j.pepi.2017.07.010.

Ekström, G., G. A. Abers, and S. C. Webb (2009). Determination of surface-wave phase velocities across USArray from noise and Aki's spectral formulation, *Geophys. Res. Lett.* **36,** no. 18, L18301 , doi: 10.1029/2009GL039131.

Emry, E. L., Y. Shen, A. A. Nyblade, A. Flinders, and X. Bao (2018). Upper mantle earth structure in Africa from full-wave ambient noise tomography, *Geochem. Geophys. Geosys.* doi: 10.1029/2018GC007804.

Fadel, I., H. Paulssen, M. van der Meijde, M. Kwadiba, O. Ntibinyane, A. Nyblade, and R. Durrheim (2020). Crustal and upper mantle shear wave velocity structure of Botswana: The 3 April 2017 Central Botswana earthquake linked to the East African rift system, *Geophys. Res. Lett.* **47,** no. 4, doi: 10.1029/2019GL085598.

Fadel, I., M. van der Meijde, and H. Paulssen (2018). Crustal structure and dynamics of Botswana, *J. Geophys. Res.* doi: 10.1029/2018JB016190.

Fishwick, S., and I. D. Bastow (2011). Towards a better understanding of African topography: A review of passive-source seismic studies of the African crust and upper mantle, *Geol. Soc. Lond. Spec. Publ.* **357,** no. 1, 343–371, doi: 10.1144/SP357.19.

Globig, J., M. Fernàndez, M. Torne, J. Vergés, A. Robert, and C. Faccenna (2016). New insights into the crust and lithospheric mantle structure of Africa from elevation, geoid, and thermal analysis: Lithosphere structure of Africa, *J. Geophys. Res.* **121,** no. 7, 5389–5424, doi: 10.1002/2016JB012972.

Haney, M. M., and V. C. Tsai (2020). Perturbational and nonperturbational inversion of Love-wave velocities, *Geophysics* **85**, no. 1, F19–F26, doi: 10.1190/geo2018-0882.1.

Haney, M. M., T. D. Mikesell, K. van Wijk, and H. Nakahara (2012). Extension of the spatial autocorrelation (SPAC) method to mixed-component correlations of surface waves, *Geophys. J. Int.* **191,** no. 1, 189–206, doi: 10.1111/j.1365-246X.2012.05597.x.

Hawkins, R. (2018). A spectral element method for surface wave dispersion and adjoints, *Geophys. J. Int.* **215,** no. 1, 267–302, doi: 10.1093/gji/ggy277.

Hawkins, R., and M. Sambridge (2019). An adjoint technique for estimation of interstation phase and group dispersion from ambient noise cross correlations, *Bull. Seismol. Soc. Am.* **109,** no. 5, 1716–1728.

Hopper, E., J. B. Gaherty, D. J. Shillington, N. J. Accardo, A. A. Nyblade, B. K. Holtzman, C. Havlin, C. A. Scholz, P. R. N. Chindandali, R. W. Ferdinand, *et al.* (2020). Preferential localized thinning of lithospheric mantle in the melt-poor Malawi Rift, *Nat. Geosci.* **13,** no. 8, 584–589, doi: 10.1038/s41561-020-0609-y.

Huang, Y., V. Strati, F. Mantovani, S. B. Shirey, and W. F. McDonough (2014). Regional study of the archean to proterozoic crust at the Sudbury Neutrino Observatory (SNO+), Ontario: Predicting the geoneutrino flux, *Geochem. Geophys. Geosys.* doi: 10.1002/2014GC005397.

Keller, C. B., and B. Schoene (2012). Statistical geochemistry reveals disruption in secular lithospheric evolution about 2.5 Gyr ago, *Nature* **485,** no. 7399, 490–493, doi: 10.1038/nature11024.

Lin, F.-C., M. H. Ritzwoller, Y. Yang, M. P. Moschetti, and M. J. Fouch (2010). Complex and variable crustal and uppermost mantle seismic anisotropy in the western United States, *Nat. Geosci.* **4,** no. 1, 55–61, doi: 10.1038/ngeo1036.

Luo, Y., Y. Yang, J. Xie, X. Yang, F. Ren, K. Zhao, and H. Xu (2020). Evaluating uncertainties of phase velocity measurements from cross-correlations of ambient seismic noise, *Seismol. Res. Lett.* **91,** no. 3, 1717–1729, doi: 10.1785/0220190308.

Luo, Y., Y. Yang, Y. Xu, H. Xu, K. Zhao, and K. Wang (2015). On the limitations of interstation distances in ambient noise tomography, *Geophys. J. Int.* **201,** no. 2, 652–661, doi: 10.1093/gji/ggv043.

Ma, Z., and G. Masters (2014). A new global Rayleigh-and Love-wave group velocity dataset for constraining lithosphere properties, *Bull. Seismol. Soc. Am.* **104,** no. 4, 2007–2026.

Ma, Z., G. Masters, G. Laske, and M. Pasyanos (2014). A comprehensive dispersion model of surface wave phase and group velocity for

the globe, *Geophys. J. Int.* **199,** no. 1, 113–135, doi: 10.1093/gji/ggu246.

Menke, W., and G. Jin (2015). Waveform fitting of cross spectra to determine phase velocity using Aki's formula, *Bull. Seismol. Soc. Am.* **105,** no. 3, 1–9, doi: 10.1785/0120140245.

Moschetti, M. P., M. H. Ritzwoller, F. Lin, and Y. Yang (2010). Seismic evidence for widespread western-US deep-crustal deformation caused by extension, *Nature* **464,** no. 7290, 885–889, doi: 10.1038/nature08951.

Moulik, P., V. Lekic, B. Romanowicz, Z. Ma, A. Schaeffer, T. Ho, E. Beucler, E. Debayle, A. Deuss, S. Durand, *et al.* (2021). Global reference seismological datasets: Multi-mode surface wave dispersion, *Geophys. J. Int.* doi: 10.1093/gji/ggab418.

Nataf, H.-C., and Y. Ricard (1996). 3SMAC: An a priori tomographic model of the upper mantle based on geophysical modeling, *Phys. Earth Planet. In.* **95,** nos. 1/2, 101–122, doi: 10.1016/0031-9201(95)03105-7.

Nyblade, A. (2015). Africaarray-Namibia [data Set] [Data set], doi: 10.7914/SN/8A_2015.

Nyblade, A., P. Dirks, R. Durrheim, S. Webb, M. Jones, G. Cooper, and G. Graham (2008). AfricaArray: Developing a geosciences workforce for Africa's natural resource sector, *Lead. Edge* **27,** no. 10, 1358–1361, doi: 10.1190/1.2996547.

Ojo, A. O., S. Ni, and Z. Li (2017). Crustal radial anisotropy beneath Cameroon from ambient noise tomography, *Tectonophysics* doi: 10.1016/j.tecto.2016.12.018.

Ojo, A. O., L. Zhao, and X. Wang (2019). Estimations of sensor misorientation for broadband seismic stations in and around Africa, *Seismol. Res. Lett.* doi: 10.1785/0220190103.

Olugboji, T. M., V. Lekic, and W. McDonough (2017). A statistical assessment of seismic models of the U.S. continental crust using Bayesian inversion of ambient noise surface wave dispersion data, *Tectonics* **36,** no. 7, 1232–1253, doi: 10.1002/2017TC004468.

Pasyanos, M. E., and A. A. Nyblade (2007). A top to bottom lithospheric study of Africa and Arabia, *Tectonophysics* **444,** no. 1, 27–44, doi: 10.1016/j.tecto.2007.07.008.

Pasyanos, M. E., T. G. Masters, G. Laske, and Z. Ma (2014). LITHO1.0: An updated crust and lithospheric model of the Earth, *J. Geophys. Res.* **119,** no. 3, 2153–2173, doi: 10.1002/2013JB010626.

Prieto, G. A., J. F. Lawrence, and G. C. Beroza (2009). Anelastic Earth structure from the coherency of the ambient seismic field, *J. Geophys. Res.* **114,** no. B7, doi: 10.1029/2008JB006067.

Raveloson, A., A. Nyblade, and R. Durrheim (2021). Joint inversion of surface wave and gravity data reveals subbasin architecture of the Congo Basin, *Geology* **49,** no. 7, 810–815, doi: 10.1130/G48408.1.

Raveloson, A., A. Nyblade, S. Fishwick, A. Mangongolo, and S. Master (2015). The upper mantle seismic velocity structure of south-central Africa and the seismic architecture of Precambrian lithosphere beneath the Congo Basin, in *Geology and Resource Potential of the Congo Basin*, M. J. de Wit, F. Guillocheau, and M. C. J. de Wit (Editors), Springer Berlin Heidelberg, Berlin, Heidelberg.

Rawlinson, N., S. Pilia, M. Young, M. Salmon, and Y. Yang (2016). Crust and upper mantle structure beneath southeast Australia from ambient noise and teleseismic tomography, *Tectonophysics* **689,** 143–156, doi: 10.1016/j.tecto.2015.11.034.

Rudnick, R. L., and S. Gao (2014). Composition of the continental crust, in *Treatise on Geochemistry*, Second Ed., Vol. 4, Elsevier, Amsterdam, doi: 10.1016/B978-0-08-095975-7.00301-6.

Seats, K. J., J. F. Lawrence, and G. A. Prieto (2012). Improved ambient noise correlation functions using Welch's method, *Geophys. J. Int.* **188,** no. 2, 513–523, doi: 10.1111/j.1365-246X.2011.05263.x.

Tarantola, A. (2005). Inverse problem theory and methods for model parameter estimation, *Soc. Ind. Appl. Math.* doi: 10.1137/1.9780898717921.bm.

Tilmann, F., H. Sadeghisorkhani, and A. Mauerberger (2019). Another look at the treatment of data uncertainty in Markov chain Monte Carlo inversion and other probabilistic methods, *Geophys. J. Int.* doi: 10.31223/osf.io/bcs4j.

Tokam, A.-P. K., C. T. Tabod, A. A. Nyblade, J. Julia, D. A. Wiens, and M. E. Pasyanos (2010). Structure of the crust beneath Cameroon, West Africa, from the joint inversion of Rayleigh wave group velocities and receiver functions, *Geophys. J. Int.* **183,** no. 2, 1061–1076.

Van Hinsbergen, D. J. J. (2011). *The Formation and Evolution of Africa: A Synopsis of 3.8 Ga of Earth History*, Geological Society of London, London, United Kingdom.

Van Hinsbergen, D. J. J., S. J. H. Buiter, T. H. Torsvik, C. Gaina, and S. J. Webb (2011). The formation and evolution of Africa from the Archaean to Present: introduction. *Geol. Soc. Lond. Spec. Publ.* **357,** no. 1, 1–8, doi: 10.1144/SP357.1.

Young, M. K., N. Rawlinson, and T. Bodin (2013). Transdimensional inversion of ambient seismic noise for 3D shear velocity structure of the Tasmanian crustTransdimensional shear velocity inversion, *Geophysics* **78,** no. 3, WB49–WB62, doi: 10.1190/geo2012-0356.1.

Yu, Y., S. S. Gao, and K. H. Liu (2020). Topography of the 410 and 660 km discontinuities beneath the Cenozoic Okavango rift zone and adjacent Precambrian Provinces, *J. Geophys. Res.* **125,** no. 9, doi: 10.1029/2019JB019290.

Zhang, H., K. Pankow, and W. Stephenson (2019). A Bayesian Monte Carlo inversion of spatial auto-correlation (SPAC) for near-surface Vs structure applied to both broad-band and geophone data, *Geophys. J. Int.* **217,** no. 3, 2056–2070, doi: 10.1093/gji/ggz136.

Manuscript received 14 December 2021