Data-driven analysis of seismic phase using circular statistics

Akshika Rohatgi[1], Andrey Bakulin[1], and Sergey Fomel[1]

https://doi.org/10.1190/tle44090683.1

Abstract

Recognizing seismic phase as a primary attribute in seismic processing workflows, we apply circular statistics, a robust data-driven approach for correcting phase distortions in prestack seismic data. Unlike traditional linear methods that struggle with wrapped phase and often defer phase diagnostics to the final processing stages, the proposed approach treats phase as a circular variable. We compute the circular mean, variance, and von Mises concentration parameter directly from phase ensembles in the frequency domain. These parameters provide insights into phase stability and coherence without needing phase unwrapping or wavelet assumptions. Synthetic tests using additive and multiplicative noise models confirm that phase distributions follow the von Mises distribution, an analog of the normal distribution for circular variables, with circular statistics reliably tracking the true phase even in low-signal-quality scenarios. Field examples demonstrate how this framework can map phase behavior across frequency and offset, enabling the detection of coherence bands and assessing the impact of each processing step on phase fidelity. The proposed approach can be particularly valuable in land acquisition, where prestack data often exhibit a low signal-to-noise ratio. Circular statistics allow us to evaluate phase integrity at each frequency, facilitating novel data conditioning and acquisition design strategies.

Introduction and motivation

The phase component of seismic signals contains critical information for accurately imaging the subsurface and identifying structural and stratigraphic features (Oppenheim and Lim, 1981; Lichman, 1999; Ulrych et al., 2007; van der Baan and Fomel, 2009; Greer et al., 2019). In many seismic applications, particularly those involving migration, coherence analysis, or inversion, consistent phase alignment is as important, if not more so, than amplitude fidelity (Xie et al., 2016; Holt and Lubrano, 2020; Bakulin et al., 2022a, 2022b). As seismic acquisition and processing evolve toward higher spatial sampling and denser data sets, analyzing phase variability statistically becomes increasingly practical and necessary.

Conventional time processing approaches, including linear and random noise removal, surface-consistent scaling, deconvolution (Taner and Koehler, 1981; Cary and Lorentz, 1993; Chan and Stewart, 1994; Meunier, 1999; Liu et al., 2006), often overlook the true statistical nature of seismic phase, particularly under low signal-to-noise ratio (S/N) or significant wavelet perturbation. A central challenge is the circular nature of phase measurements: they are inherently bounded within $[-\pi, \pi]$, which violates assumptions behind standard linear statistical tools. This behavior can lead to misinterpretation if analyzed using conventional means or variances (Bakulin et al., 2024).

This paper introduces a robust analytical framework that leverages circular statistics (Mardia and Jupp, 1999) to describe and interpret seismic phase distributions across frequencies. By treating phase as a circular random variable, we quantify its average behavior (via the circular mean) and the spread or concentration (via the von Mises concentration parameter, κ), which is a frequency-dependent analog to S/N. This conceptual shift, from linear to conceptual statistical treatment of phase, allows for a more meaningful understanding of phase variability, particularly in challenging land data contaminated by noise.

Unlike traditional workflows that consider phase quality only in the final stacked or migrated images, the proposed framework enables phase diagnostics starting from the raw prestack data. This allows phase integrity to be monitored and optimized throughout each processing step. Because the analysis operates frequency by frequency, it provides a unique lens for evaluating how processing workflows impact phase fidelity, something amplitude- or stack-based S/N metrics cannot reliably detect. This capability is especially critical in land acquisition scenarios, where raw prestack data are often so degraded that even computing S/N for an entire frequency band becomes challenging (Bakulin et al., 2022a). Furthermore, by enabling direct assessment of phase coherence across frequency, this framework complements and strengthens frequency-dependent processing strategies that have long been advocated in the literature (e.g., Retailleau et al., 2014; Bakulin et al., 2019) and are increasingly adopted in practice. It also opens the door to a new class of algorithms, such as time-frequency phase masking (Bakulin et al., 2023), that manipulate phase by estimating the signal component through local stacking. In this study, we go a step further by analyzing the entire statistical distribution of phase to estimate the signal phase and assess the extent of perturbations caused by noise.

We propose to treat phase statistics as core diagnostic variables that, together with S/N, can be monitored continuously throughout the processing workflow to enable better informed decisions in acquisition design, evaluation of bandwidth-limited resolution, and the development of phase-aware processing strategies. We test the utility of this approach on both synthetic and field data sets, showing that it yields stable and interpretable estimates of phase behavior, even under significant noise contamination.

When phase misleads: A visual case for rethinking analysis

We start with a simple but illustrative synthetic example highlighting the challenges of phase analysis in noisy environments. Figure 1 shows a seismic ensemble with three flat reflectors,

Manuscript received 3 May 2025; revision received 27 June 2025; accepted 1 July 2025.
[1]The University of Texas at Austin, Bureau of Economic Geology, Austin, Texas, USA. E-mail: akshikarohatgi@utexas.edu; a_bakulin@yahoo.com; sergey.fomel@beg.utexas.edu.

each distorted by additive Gaussian noise. Two noise levels are considered: –2 and –10 dB S/N. The underlying signal is the same across all traces, so any deviation in phase is entirely due to noise. Figures 1c and 1d show histograms of estimated phase at 40 and 60 Hz on a linear scale. Without noise, phase would remain constant across traces, collapsing the distribution to a narrow spike, marked by the dashed red line. But even a modest amount of noise contaminates the phase. As the noise grows stronger, the contamination becomes worse.

At first glance, the distribution at 40 Hz (Figure 1c) appears reasonably symmetric, suggesting that the noise adds some random wobble around the true phase. However, a closer look at Figure 1k casts doubt on that idea. The shape is not as clean or centered as expected. At 60 Hz (Figures 1d and 1l), the

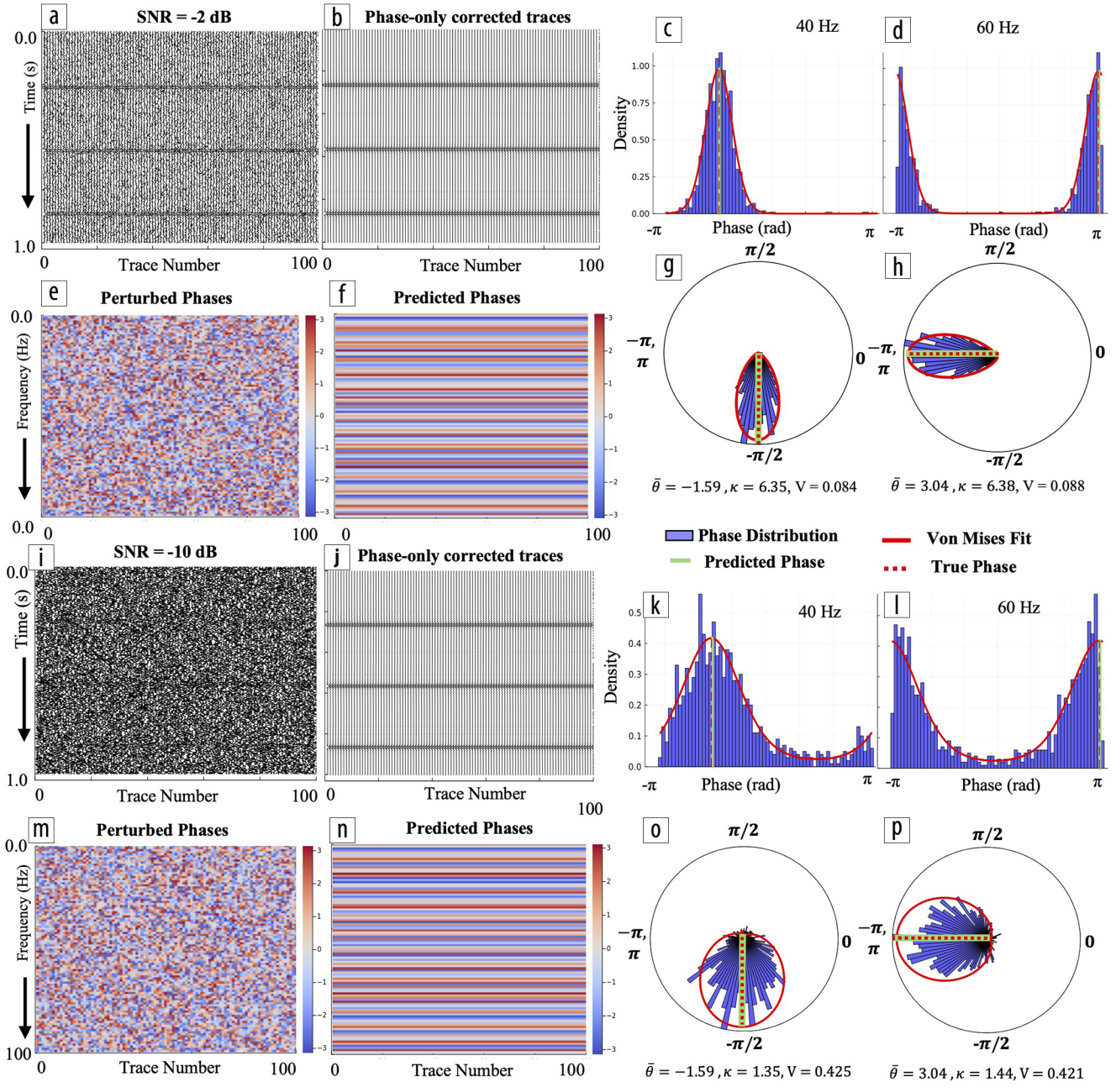

Figure 1. Seismic phase analysis under noisy conditions using an additive noise model with different S/N. (a) Seismic traces contaminated with additive noise at S/N = –2 dB. (b) Seismic traces after phase-only correction, where the noisy phase is replaced with the circular mean computed from a local window centered around the trace of interest, while retaining the original amplitude, demonstrating recovery of the original signal. (c) and (d) Linear representations of seismic phase at 40 and 60 Hz, respectively, highlighting distortions caused by noise. (e) Phase spectra of the noisy traces in the frequency domain, illustrating noise impact across frequencies. (f) Phase spectra after correction. The bimodal structure in (d) underscores the limitations of linear phase visualization for circular data. (g) and (h) Circular (rose plot) representations of seismic phase at 40 and 60 Hz, respectively. The calculated circular mean (red dashed line) remains aligned with the true phase, validating the robustness of circular statistics. (i)–(p) Corresponding results for traces with stronger noise (S/N = –10 dB), showing qualitatively similar observations but with significantly larger phase spread. Despite the increased noise, circular statistics still demonstrate improved phase coherence and signal recovery.

phase distribution appears bimodal, splitting into two peaks. This is not merely random noise; it is a breakdown of limitations of linear statistics, which assume that phase behaves like a regular number.

From confusion to clarity: Circular statistics explained visually

To make sense of the distorted phase patterns observed earlier, we must first reconsider how phase behaves. Phase is not a typical number; it is periodic and circular. When it exceeds π or drops below −π, it wraps around to the other side. This is not a minor detail — it fundamentally alters how we compute averages and assess variability using conventional linear tools. Figure 2 illustrates this visually: while the unwrapped phase evolves smoothly over time, its wrapped counterpart, confined to the interval [−π, π], exhibits abrupt jumps.

These discontinuities are not features of the underlying signal but artifacts introduced by how the data are represented. Yet they pose serious challenges when applying standard statistical measures like arithmetic means or variances.

Phase unwrapping is critical in fields like radar, InSAR, or optical interferometry where analysts need continuous phase maps to estimate displacements, topography, or deformation. A large body of work supports such methods (e.g., Ghiglia and Romero, 1994; Zebker and Lu, 1998). However, seismic analysis often deals with noisy, complex wavefields where unwrapping algorithms break down and add little value. For seismic data, the wrapped phase is not a nuisance but a proper, bounded quantity that naturally lends itself to a different analysis.

This is where circular statistics come in. Rather than treating phase as a scalar on a number line, we treat it as a direction on a unit circle. This framework respects its periodicity and avoids the distortions that plague traditional metrics. To compute the average direction of a set of N seismic phases ($\theta_1, \theta_2, \ldots, \theta_N$), we convert each into a unit vector and then average the cosine and sine components.

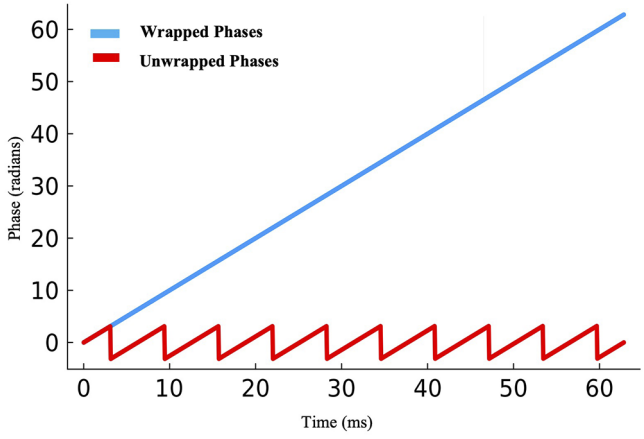

Figure 2. Comparison of unwrapped and wrapped phase as a function of time. The unwrapped phase (blue solid line) increases continuously, reflecting the natural progression of the phase over time. In contrast, the wrapped phase (red dashed line) is confined within the fixed interval [−π, π], illustrating the periodic nature of wrapped phases commonly encountered in signal processing and circular analysis.

$$C = \sum_{i=1}^{N} \cos\theta_i,\ S = \sum_{i=1}^{N} \sin\theta_i,\ R^2 = C^2 + S^2\ (R \geq 0). \quad (1)$$

The circular mean direction $\bar{\theta}$ is given by

$$\bar{\theta} = \begin{cases} \tan^{-1}\left(\frac{S}{C}\right),\ S > 0,\ C > 0 \\ \tan^{-1}\left(\frac{S}{C}\right) + \pi,\ C < 0 \\ \tan^{-1}\left(\frac{S}{C}\right) + 2\pi.\ S < 0,\ C > 0 \end{cases}. \quad (2)$$

This avoids the classic pitfall: if two angles are 1° and 359°, a linear average yields 180°, which is clearly wrong. The circular mean (equation 2) correctly returns 0°, which lies between the two and matches the proper average direction. Figure 3 illustrates this vividly. On the left, the linear mean misleads. On the right, circular averaging gets it right.

Beyond the average, we care about how tightly the phase values are clustered. The mean resultant length provides an appropriate measure:

$$\bar{R} = \frac{R}{N},\ 0 \leq \bar{R} \leq 1. \quad (3)$$

A value near 1 (Figure 4b) indicates strong phase alignment. A value near 0 (Figure 4a) means the phases are scattered, with no clear dominant direction. This gives us a compact, intuitive measure of phase coherence. An equally important metric is the circular variance, defined as:

$$V = 1 - \bar{R}. \quad (4)$$

Like traditional variance, it quantifies how spread out the values are. Unlike linear variance, circular variance is naturally bounded between 0 and 1. When all phase vectors point in the same direction (perfect alignment), $\bar{R} = 1$, so $V = 0$. When the phases are scattered uniformly across the circle,), $\bar{R} \approx 0$, leading to $V \approx 1$. This makes circular variance an intuitive and stable phase coherence measure, even in noisy or irregular data sets.

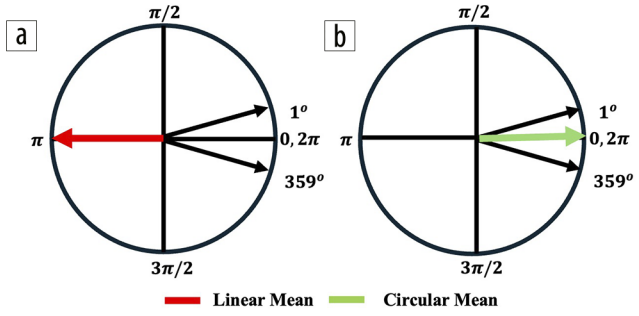

Figure 3. (a) Illustration of the linear mean, where opposing phase angles are averaged directly, leading to an incorrect mean direction after linear averaging of wrapped phases. (b) Illustration of the circular mean, where angles are treated as vectors and their average yields the correct mean direction, known as the circular mean.

Critically, all of these circular statistics — the circular mean, resultant length, and circular variance — can be computed directly from raw wrapped phase values, without requiring phase unwrapping or signal modeling. This is analogous to how mean and standard deviation are used in linear statistics: they offer simple, interpretable summaries of complex data distributions, but adapted to the geometry of angles. For geophysical applications, especially in land seismic or scattered wavefields, this offers a powerful and noise-resilient toolkit for phase diagnostics.

Table 1 summarizes the key attributes used throughout this study. While the circular mean, resultant length, and variance broadly apply to any wrapped phase distribution, the concentration parameter κ is specific to the von Mises distribution. We include it here because it can be directly computed from $\overline{R}$ (and thus from V), making it a derived but useful metric when the von Mises model is assumed. This connection becomes central in the next section.

Modeling phase distributions with von Mises: The circular Gaussian

The circular mean and variance (equations 2 and 4, respectively) give us powerful ways to summarize phase behavior, but what if we want a complete, interpretable model of how phase values are distributed? In linear statistics, the normal (Gaussian) distribution serves this role. For circular data such as phase, the equivalent is the von Mises distribution (Mardia and Jupp, 1999).

This distribution provides a natural model for how wrapped phase values behave in the presence of common types of seismic noise. Both additive Gaussian and multiplicative random noise, typical in land seismic data, lead to phase perturbations that closely follow von Mises distributions (Bakulin et al., 2022b, 2024; Rohatgi et al., 2024a).

The von Mises probability density function is:

$$f(\theta;\overline{\theta},\kappa) = \frac{1}{2\pi I_0(k)}\exp\left[k\cos(\theta-\overline{\theta})\right], \qquad (5)$$

where θ is the wrapped phase angle, $\overline{\theta}$ is the circular mean, representing the central direction, κ is the concentration parameter, analogous to the inverse of variance, and $I_0(\kappa)$ is the modified Bessel function of the first kind of order 0.

This distribution wraps a bell-shaped curve around the unit circle. For large κ, it peaks tightly around the mean. For small κ, it flattens out into a near-uniform circle.

The von Mises distribution fits real seismic data well because it captures the kind of phase variability induced by realistic random noise, not through theoretical assumptions, but directly measurable in ensembles of traces. This allows us to go beyond just detecting noise. We can statistically model it. The distribution provides two key parameters that matter most for analysis:

- The circular mean $\overline{\theta}$ represents the direction of dominant or average phase and estimates the underlying signal phase.
- The concentration parameter κ, specific to the von Mises distribution, quantifies how tightly the phase vectors are clustered around the mean. It serves as a phase-based analog of S/N.

Importantly, κ can be estimated from $\overline{R}$ using the following approximations (Fisher et al., 1993):

$$\kappa = \begin{cases} 2\overline{R} + R^3 + \frac{5R^5}{6}, & \overline{R} < 0.53 \\ -0.4 + 1.39\overline{R} + \frac{0.43}{1-\overline{R}}, & 0.53 \le \overline{R} < 0.53 \\ \frac{1}{3\overline{R} - 4\overline{R}^3 + \overline{R}^4}, & \overline{R} \ge 0.85 \end{cases}. \qquad (6)$$

In short, low variance (V close to 0) corresponds to high concentration (large κ), indicating tightly clustered, coherent phases. High

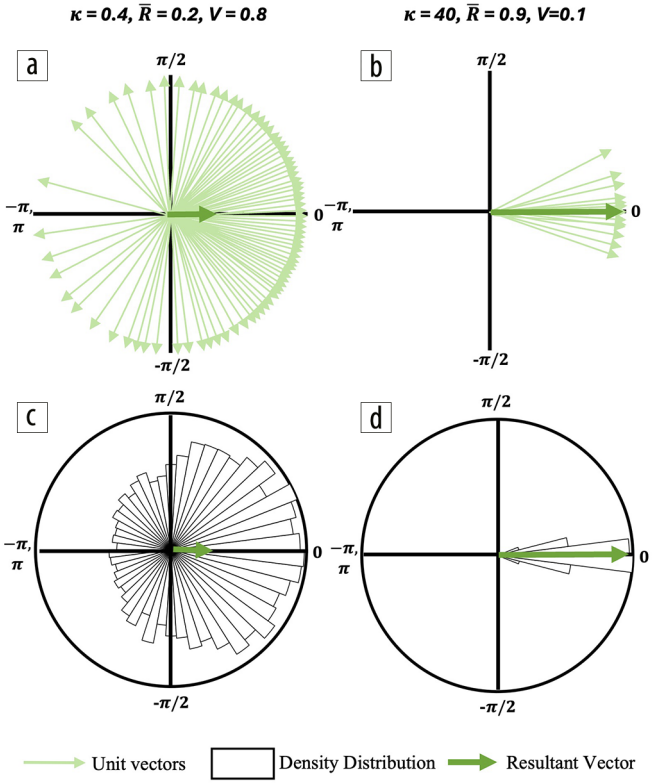

Figure 4. Polar representation of phase distributions modeled using the von Mises distribution, which is the circular analog of the normal distribution. The concentration parameter κ controls the spread of the distribution, with higher κ indicating stronger clustering around the mean direction. Panels (a) and (c) (left) correspond to a low κ value (0.4), resulting in high circular variance ($V = 0.8$) and low phase coherence. Panels (b) and (d) (right) show a high κ value (40), yielding low circular variance ($V = 0.1$) and tightly clustered phases. The top row (a) and (b) shows individual unit phase vectors (light green) and their resultant vector (dark green), while the bottom row (c) and (d) shows the corresponding circular histograms.

Table 1. Summary of key circular statistical attributes for seismic phase analysis.

Attribute	Symbol	Range	Interpretation
Circular mean	$\overline{\theta}$	$[-\pi, \pi]$	Central phase direction
Mean resultant length	$\overline{R}$	[0, 1]	Degree of phase alignment (clustering strength)
Circular variance	$V = 1 - \overline{R}$	[0, 1]	Phase dispersion or incoherence
Concentration (von Mises)	κ	$[0, \infty)$	Inverse dispersion (analogous to S/N in phase space)

variance (V near 1) implies low concentration (small κ), as in a uniform phase distribution. This relationship is visualized in Figure 5.

Theoretically, if phase unwrapping were reliably possible, one could fit a traditional Gaussian to the unwrapped phases and estimate standard deviation σ in radians or degrees, giving familiar uncertainty metrics. Bakulin et al. (2024) derived numerical correspondences between κ and σ for such cases. However, unwrapping is rarely feasible or stable for noisy seismic data in practice. Therefore, becoming fluent in interpreting wrapped-domain quantities like circular variance (V) and the von Mises concentration parameter (κ) is essential. Both describe phase dispersion, but differently. Circular variance is a general-purpose metric that can be computed for any phase distribution, whereas κ is specific to the von Mises model. V is often more robust and intuitive for practical diagnostics in field data due to its bounded range between 0 and 1. In contrast, κ ranges from 0 to ∞ and can exhibit large fluctuations between neighboring samples in noisy settings. Nonetheless, when the von Mises model is assumed, V and κ are tightly linked through the mean resultant length $\overline{R}$ and are analytically or graphically interconvertible (see Figure 5 and Table 1).

In short, the von Mises distribution gives us more than a fit — it offers a framework. It allows us to interpret phase coherence, model uncertainty, and detect reliable signal content across frequencies. With this model in hand, we are now ready to return to our earlier examples, this time through the lens of circular statistics, and see what was hidden in plain sight.

Application of circular statistics to the additive white noise example

Figures 1c–1h and 1k–1p show the application of circular statistics to the initial synthetic example with additive white noise. We accumulated phase ensembles from 1000 traces and analyzed them frequency by frequency. Suddenly, the patterns that previously appeared confusing now begin to make sense. The circular mean aligns precisely with the true signal phase, and the observed phase distributions fit well with the von Mises model (equation 5).

Traditionally, such phases are studied using linear histograms (Figures 1c and 1d). However, this approach fails to capture the true nature of phase behavior, especially in noisy environments. For instance, the apparent bimodality seen at 60 Hz (Figure 1d) arises not from a physical phenomenon but from an artifact of the linear representation. Seismic phase is fundamentally circular, naturally bounded within $[-\pi, \pi]$. Ignoring that structure leads to misleading interpretations.

The story becomes much clearer when the same distributions are displayed on a circular scale using rose diagrams (Figures 1g–1h and 1o–1p). What appeared bimodal on the line now forms a single, coherent cluster around the true phase direction. Figures 1h and 1p visually resemble Figures 1g and 1o, except for the mean rotation, which reflects natural changes in signal phase with frequency. These rose diagrams resolve visual ambiguity and allow for direct estimation of the signal phase using the circular mean. Each unit vector on the circle corresponds to a trace's phase at a given frequency. Their average direction provides a robust estimate of the dominant phase, so no unwrapping or modeling is needed.

As expected from white noise, the variance (quantified via the von Mises κ or circular variance V) remains constant across frequencies. What changes is the signal phase itself, which rotates naturally from frequency to frequency. This example demonstrates how circular statistics reveal structure and coherence hidden by traditional methods, clarifying what initially appeared as disordered phase chaos.

Phase analysis under multiplicative noise

To test the robustness of circular statistics under more realistic conditions, we apply a multiplicative noise model introduced by Bakulin et al. (2022b, 2023), specifically designed to induce phase perturbations representative of scattering and near-surface heterogeneity in land seismic data. Unlike linear models that define noise using standard deviation (σ), this approach uses a frequency-dependent concentration parameter (κ) to directly control the spread of phase distributions and better capture the non-uniform, frequency-dependent nature of seismic noise.

We invert κ from seismic trace ensembles using the mean resultant length $\overline{R}$, applying standard approximations from Fisher et al. (1993) and Mardia and Jupp (1999) (Figure 6a). The resulting $\kappa(f)$ curve quantifies phase alignment as a function of frequency, essentially mapping a phase-based analog to S/N (Figure 6b).

Conceptually, κ serves a role analogous to S/N. High κ values indicate stable, well-aligned phases and low phase noise. Low κ reflects scattered phases and high phase noise. This trend is evident in the phase histograms at 20, 40, and 60 Hz (Figures 6d–6F); however, the behavior becomes far more transparent and unambiguous in the corresponding rose plots (Figures 6g–6i). As frequency increases, phase vectors become more dispersed and κ decreases, consistent with the design of the multiplicative noise model.

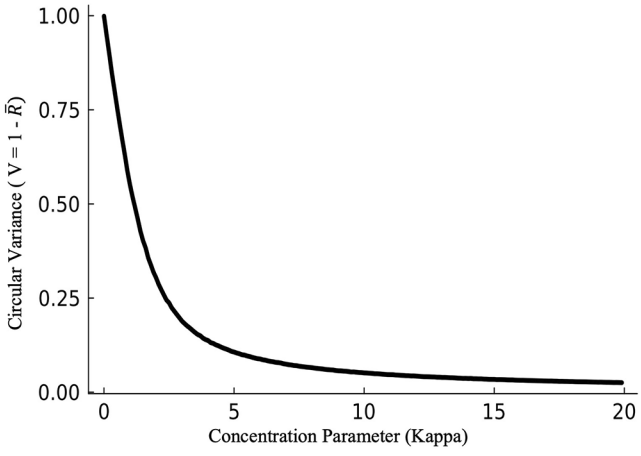

Figure 5. Relationship between circular variance (V) and the von Mises concentration parameter (κ). This plot illustrates the nonlinear mapping between the two, derived numerically via the mean resultant length $\overline{R}$. As κ increases, variance V decreases, indicating tighter clustering of phase angles. In noisy seismic data (where phase estimates fluctuate) variance tends to be high, making it a more stable and interpretable measure than κ. At low variance (i.e., high phase coherence), small changes in V can correspond to large variations in κ, which makes κ less robust for interpretation in those cases. This motivates our choice to primarily report circular variance when characterizing phase coherence.

At 20 Hz (Figures 6d and 6g), phase vectors are tightly clustered, and the circular mean closely aligns with the true phase. At 60 Hz (Figures 6f and 6i), phase scatter increases, and the circular mean drifts, signaling reduced reliability in phase estimation under higher perturbations.

This analysis also underscores the importance of ensemble size. In high-noise conditions, small ensembles may yield unstable phase estimates. Even if linear plots show significant deviations from the true phase, the circular representation (e.g., Figure 6i) often reveals that the mean direction is statistically meaningful. However, achieving convergence requires a sufficiently large number of traces. In an earlier paper (Rohatgi et al., 2024b), we proposed estimating the ensemble size using the standard deviation of the residual phase, which requires knowing the true phase. In contrast, the approach proposed here infers phase stability and confidence directly from raw phase distributions, without the need for phase unwrapping or prior wavelet information.

In summary, circular statistics quantify phase coherence and provide practical guidance for designing ensembles that yield robust phase estimates, even in extremely noisy seismic data.

Real-data example: Circular statistics reveal frequency-dependent phase coherence not captured by conventional measures

We now apply our circular statistical framework to a 3D prestack land data set from the continental United States. This field example illustrates how phase coherence varies with frequency and offset and how it can be quantified using circular variance and modeled using the von Mises distribution.

To generalize this method for field data, we adopt a sliding-window strategy similar to that of Bakulin et al. (2020b, 2023, 2024). At each time-frequency location, we extract a phase ensemble from a window moving horizontally across traces, compute the circular mean, circular variance, and κ, and assign these values to the center trace and time sample. Repeating this across the data set produces dense frequency- and offset-dependent phase behavior maps.

Unlike traditional QC methods that rely on amplitude or semblance, this approach provides a direct diagnostic of phase coherence at each frequency. This workflow can be conceptually summarized as shown in Figure 7.

Figures 8a and 8e present seismic gathers before and after conventional processing, respectively, along with circular variance maps across frequency and offset. These gathers, extracted from a 3D CDP supergather, are shown after normal moveout (NMO) and statics corrections. The conventional time processing flow included: linear noise removal, refraction statics, random noise attenuation, two passes of surface-consistent deconvolution, post-deconvolution noise attenuation (linear, random, and burst noise), merge phase and static matching, surface-consistent scaling, velocity analysis, and residual statics correction.

While reflectors appear visually clearer after conventional processing, the variance maps reveal deeper insights: high-frequency components remain dominated by noise, with circular variance approaching one. This indicates low phase coherence even after standard processing. The coherent signal is mostly confined to the 10–25 Hz range, revealing the true denoised effective bandwidth for phase-sensitive applications such as prestack inversion (Lichman, 1999). Furthermore, increased variance is clearly observed within the near-offset cone compared to the mid- and far-offset ranges.

To validate this quantitatively, we examine phase distributions at 16 Hz within the far-offset window highlighted by the red box (Figures 8a, 8e, and 8i), each fitted with a von Mises model (Figures 8c, 8g, and 8k). The

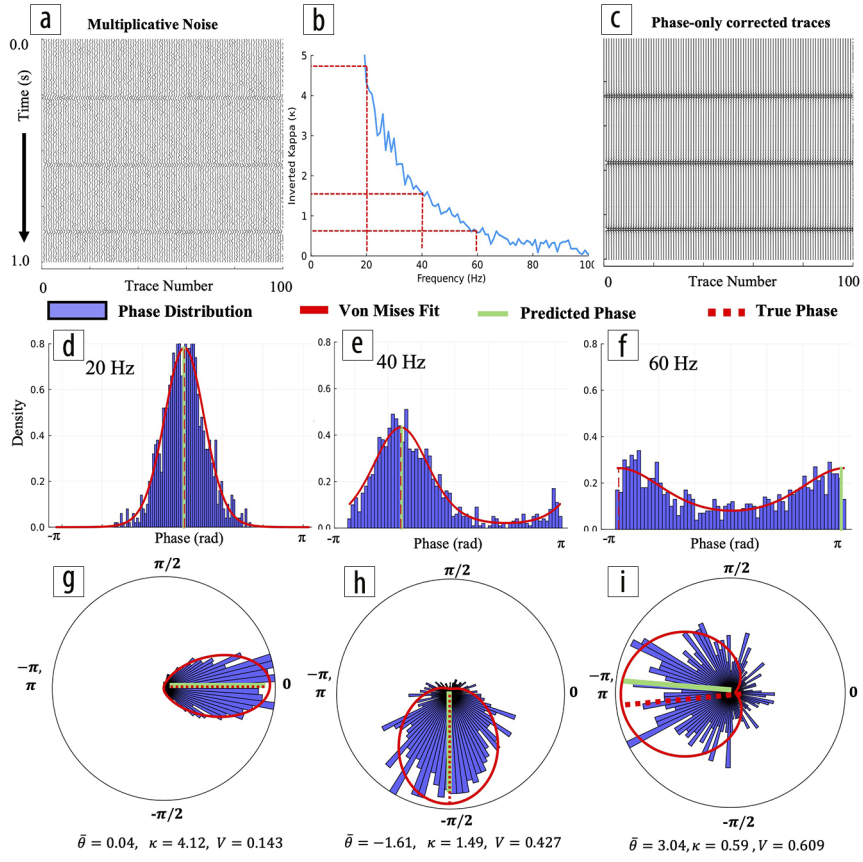

Figure 6. Analysis of seismic phases using a multiplicative noise model and a frequency-dependent κ parameter. (a) Seismic traces perturbed by multiplicative noise, illustrating the random nature of the phase perturbations. (b) Frequency-dependent inverted κ values, representing the concentration of phases around the true phase. (c) Reconstructed seismic traces after phase correction, demonstrating effective noise mitigation. (d)–(f) Linear phase histograms at 20, 40, and 60 Hz, respectively. (g)–(i) Corresponding rose plots (circular histograms) at the same frequencies, with κ values of 4.85, 1.53, and 0.76. The red dashed lines indicate the calculated circular mean, which aligns closely with the true phase at low frequencies but diverges at higher frequencies due to increased noise. Note how rose plots clearly depict phase clustering around the mean, providing a more interpretable visualization than the wrapped linear histograms, which appear ambiguous and even bimodal under strong noise.

DOI:10.1190/tle44090683.1

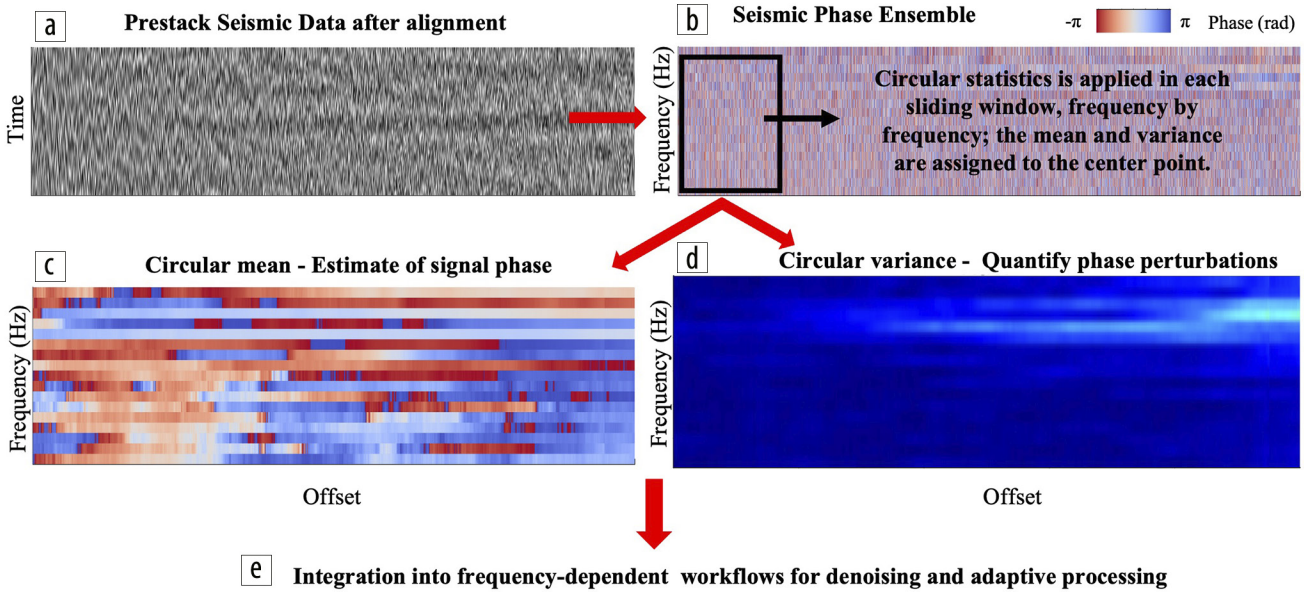

Figure 7. Circular statistics workflow for frequency-dependent phase analysis and correction of prestack seismic data. (a) Raw prestack seismic section after signal alignment using NMO, focused on a window around a strong reflector (time versus offset). (b) Phase ensemble extracted via Fourier transform and displayed as phase spectra ($-\pi$ to π); local sliding window applied for statistical analysis. (c) Frequency-dependent circular mean providing the best estimate of the coherent signal. (c) Circular variance map quantifying phase perturbations across frequency and offset. (e) Integration into frequency-dependent workflows for denoising and adaptive processing. This workflow enables seamless diagnostics and correction at the frequency-by-frequency level, supporting integration into advanced frequency-dependent processing workflows.

distribution is broad and noisy in the unprocessed data, with only a faint central peak, corresponding to a high circular variance of 0.77 and low κ (≈ 0). After standard processing, the distribution tightens modestly, and variance decreases to 0.49, indicating partial improvement in phase alignment. Only circular mean substitution produces a sharply peaked distribution (variance $\approx$ 0.01, $\kappa \gg 1$), with phases tightly clustered around the true direction. This progression demonstrates how circular statistics not only quantify phase coherence but also enable recovery of meaningful signal phase estimates that can be used for phase substitution (Bakulin et al., 2020b), time-frequency phase masking (Bakulin et al., 2023), or other frequency-dependent denoising strategies aimed at restoring stable phase structure even in low-S/N conditions.

Figure 8i shows the result of the phase substitution with circular mean. This result is similar to locally stacked phase substitution (Bakulin et al., 2020b). We retain the original amplitude but replace the contaminated phase with its circular mean, computed across a large ensemble. The reflectors become better aligned, and coherence improves across frequencies and offsets, as seen in the updated variance map (Figure 8j). Importantly, this was achieved without unwrapping, wavelet assumptions, or time-domain stacking, just raw phase statistics.

This method is most appropriate when applied to windows or frequency slices dominated by a single coherent event. Flatness is not required, as long as the moveout of the dominant event is estimated, the window can be aligned accordingly, as is typically done in local stacking (Bakulin et al., 2020a, 2020b, 2023). For windows with multiple dipping or interfering events, a single phase estimate may be inadequate, and more localized or adaptive strategies may be necessary to avoid misrepresentation. The method also assumes densely sampled seismic data, and the selection of ensemble size remains a critical consideration. Including many traces within each window helps ensure statistical stability of the calculated circular mean and variance. However, even limited phase averaging or local stacking can be extremely beneficial in reducing phase variability in the data, especially for prestack enhancement (Bakulin et al., 2024).

It is important to emphasize that phase substitution is not necessarily intended as a final processing solution. Instead, it serves as a diagnostic test to demonstrate that the estimated mean phase is physically meaningful and capable of producing a coherent, trackable signal. Once verified, the circular mean can be used directly or serve as input to more advanced workflows, such as time-frequency phase masking (Bakulin et al., 2023) where the goal is to decontaminate the signal phase and address amplitude noise through separate, targeted manipulation of phase and amplitude, unlike traditional local stacking which handles both simultaneously.

Together, these results demonstrate that wrapped seismic phase can be a powerful diagnostic and correction tool when analyzed through circular statistics. Circular variance maps reveal where the signal is coherent, von Mises modeling explains phase behavior, and circular mean substitution provides a robust correction path, all in a unified, interpretable framework.

Conclusions

We have presented circular statistics as a practical and statistically rigorous framework for seismic phase analysis. By treating phase as a directional variable confined to the interval $[-\pi, \pi]$, we avoid the distortions and ambiguities inherent in conventional linear analysis, particularly under low signal-to-noise conditions.

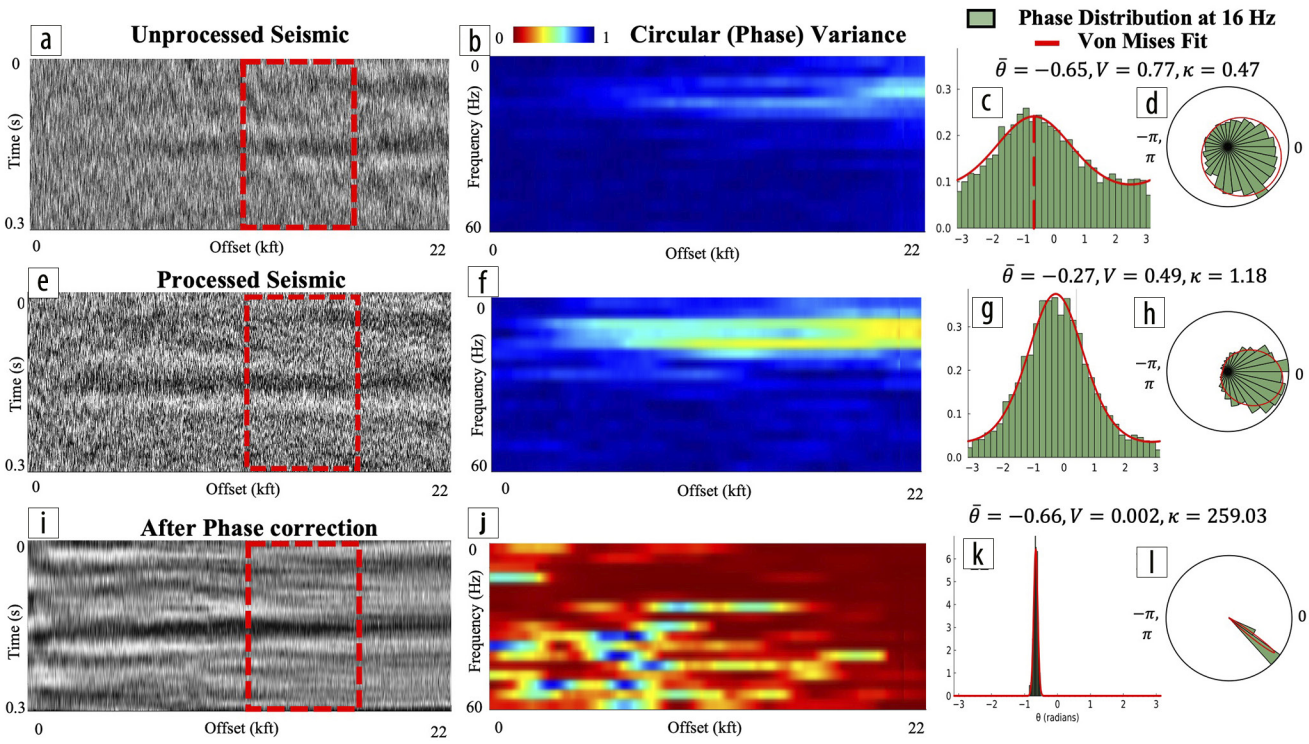

Figure 8. Comparison of unprocessed, processed, and phase-corrected prestack seismic land data. Panels (a), (e), and (i) show seismic sections focused on a window around a strong deep reflector after NMO —before processing, after conventional time processing, and after phase correction using the circular mean, respectively. Panels (b), (f), and (j) present corresponding frequency-offset heatmaps of circular variance, illustrating how processing and phase correction improve phase coherence across offsets and frequencies. Panel (j) shows a significant improvement in phase coherence across the band compared to (b) and (f). The red box in (a), (e), and (i) indicates the far-offset window used to compute the phase distributions shown in the right-hand panels. Panels (c), (g), and (k) display phase histograms at 16 Hz fitted with a von Mises distribution (red curve) for the unprocessed, processed, and phase-corrected data, respectively, demonstrating circular symmetry and tightening phase concentration with processing. Panels (d), (h), and (l) show the corresponding circular rose plots. The von Mises model consistently captures phase behavior, with tighter distributions indicating improved coherence after phase correction.

The von Mises distribution provides a natural model for phase variability in noisy seismic data, enabling direct quantification of central tendency (via the circular mean) and coherence (via circular variance or κ) from raw phase values without the need for unwrapping or wavelet assumptions. The circular mean can reliably track the underlying signal phase for practically important additive and multiplicative random noise cases, even in highly scattered or noise-dominated environments. While similar phase estimates could be obtained through local stacking, the circular statistics approach offers a deeper advantage: it recovers and visualizes the full absolute phase distribution, frequency by frequency, delivering both the mean phase and the associated phase spread. This richer quantitative insight into phase stability allows direct assessment of phase quality, which is critical for designing data-driven workflows, optimizing ensemble size, and enhancing frequency-dependent processing strategies.

In comparison to traditional linear phase diagnostics, circular statistics, as visualized through rose plots, provide clearer and more interpretable insights into phase behavior across frequency, offset, and processing parameters. All key quantities — circular mean, variance, and concentration — can be computed directly from the phase ensembles, eliminating the need for unwrapping, wavelet assumptions, or manual tuning. This reduces subjective bias and enables reproducible, objective data-driven analysis.

Phase coherence varies with frequency across the seismic spectrum, presenting opportunities for frequency-dependent processing. These strategies are increasingly critical in modern broadband workflows to enhance resolution and recover subtle structural details. Whether used for quality control, preprocessing, or as a foundation for advanced imaging techniques, circular statistics can serve as a robust and interpretable basis for phase-aware seismic analysis. TLE

Acknowledgments

We acknowledge Fairfield Geotechnologies for granting permission to use the data presented in this study.

Data and materials availability

Data associated with this research are confidential and cannot be released.

Corresponding author: akshikarohatgi@utexas.edu

References

Bakulin, A., D. Neklyudov, and I. Silvestrov, 2024, The impact of receiver arrays on suppressing seismic speckle scattering noise caused by the meter-scale near-surface heterogeneity: Geophysics, **89**, no. 6, V551–V561, https://doi.org/10.1190/geo2023-0489.1.

DOI:10.1190/tle44090683.1

Bakulin, A., D. Neklyudov, and I. Silvestrov, 2023, Seismic time-frequency masking for suppression of seismic speckle noise: Geophysics, **88**, no. 5, V371–V385, https://doi.org/10.1190/geo2022-0779.1.

Bakulin, A., I. Silvestrov, and M. Protasov, 2022a, Signal-to-noise ratio computation for challenging land single-sensor seismic data: Geophysical Prospecting, **70**, no. 3, 629–638, https://doi.org/10.1111/1365-2478.13178.

Bakulin, A., D. Neklyudov, and I. Silvestrov, 2022b, Multiplicative random seismic noise caused by small-scale near-surface scattering and its transformation during stacking: Geophysics, **87**, no. 5, V419–V435, https://doi.org/10.1190/geo2021-0830.1.

Bakulin, A., I. Silvestrov, M. Dmitriev, D. Neklyudov, M. Protasov, K. Gadylshin, and A. Dolgov, 2020a, Nonlinear beamforming for enhancement of 3D prestack land seismic data: Geophysics, **85**, no. 3, V283-V296, https://doi.org/10.1190/geo2019-0341.1.

Bakulin, A., D. Neklyudov, and I. Silvestrov, 2020b, Prestack data enhancement with phase corrections in time-frequency domain guided by local multidimensional stacking: Geophysical Prospecting, **68**, no. 6, 1811–1818, https://doi.org/10.1111/1365-2478.12956.

Bakulin, A., I. Silvestrov, and M. Dmitriev, 2019, Adaptive multiscale processing of challenging 3D seismic data for first-break picking, FWI and imaging: 89th Annual International Meeting, SEG, Expanded Abstracts, 3979–3984, https://doi.org/10.1190/segam2019-3214616.1.

Cary, P. W., and G. A. Lorentz, 1993, Four-component surface-consistent deconvolution: Geophysics, **58**, no. 3, 383–392, https://doi.org/10.1190/1.1443421.

Chan, W. K., and R. R. Stewart, 1994, 3-D f-k filtering: CREWES Research Report, 6, 15.1–15.7.

Fisher, N. I., T. Lewis, and B. J. J. Embleton, 1993, Statistical analysis of spherical data: Cambridge University Press.

Ghiglia, D. C., and L. A. Romero, 1994, Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods: Journal of the Optical Society of America. A, Optics, Image Science, and Vision, **11**, no. 1, 107–117, https://doi.org/10.1364/JOSAA.11.000107.

Greer, S., S. Fomel, and M. Fry, 2019, Prestack phase corrections using local seismic attributes: 89th Annual International Meeting, SEG, 3949–3953, https://doi.org/10.1190/segam2019-3214515.1.

Holt, R., and A. Lubrano, 2020, Stabilizing the phase of onshore 3D seismic data: Geophysics, **85**, no. 6, V473–V479, https://doi.org/10.1190/geo2019-0695.1.

Lichman, E., 1999, Automated phase-based moveout correction: 69th Annual International Meeting, SEG, Expanded Abstracts, 1150–1153, https://doi.org/10.1190/1.1820706.

Liu, C., Y. Liu, B. Yang, D. Wang, and J. Sun, 2006, A 2D multistage median filter to reduce random seismic noise: Geophysics, **71**, no. 5, V105–V110, https://doi.org/10.1190/1.2236003.

Mardia, K. V., and P. E. Jupp, 1999, Directional statistics: John Wiley & Sons, https://doi.org/10.1002/9780470316979.

Meunier, J., 1999, 3D geometry, velocity filtering and scattered noise: 69th Annual International Meeting, SEG, Expanded Abstracts, 1216–1219, https://doi.org/10.1190/1.1820725.

Oppenheim, A. V., and J. S. Lim, 1981, The importance of phase in signals: Proceedings of the IEEE, **69**, no. 5, 529–541, https://doi.org/10.1109/PROC.1981.12022.

Retailleau, M., R. El Asrag, and J. Shorter, 2014, Processing land broadband data: Challenges that Oman surveys present and how they are addressed: EAGE/SPG Workshop on Broadband Seismic, https://doi.org/10.3997/2214-4609.20141699 l.

Rohatgi, A., A. Bakulin, and S. Fomel, 2024a, Analyzing the impact of additive and multiplicative noise on seismic data analysis: Fourth International Meeting for Applied Geoscience & Energy, SEG/APPG, Expanded Abstracts, 2132–2136, https://doi.org/10.1190/image2024-4086176.1.

Rohatgi, A., A. Bakulin, and S. Fomel, 2024b, Phase pilot recovery: A foundation for mitigating speckle scattering noise in seismic data: Fourth International Meeting for Applied Geoscience & Energy, SEG/AAPG, Expanded Abstracts, 2314–2318, https://doi.org/10.1190/image2024-4091774.1.

Taner, M. T., and F. Koehler, 1981, Surface consistent corrections: Geophysics, **46**, no. 1, 17–22, https://doi.org/10.1190/1.1441133.

Ulrych, T. J., S. T. Kaplan, M. D. Sacchi, and E. Galloway, 2007, The essence of phase in seismic data processing and inversion: 77th Annual International Meeting, SEG, Expanded Abstracts, 1765–1769, https://doi.org/10.1190/1.2792834.

van der Baan, M., and S. Fomel, 2009, Nonstationary phase estimation using regularized local kurtosis maximization: Geophysics, **74**, no. 6, A75–A80, https://doi.org/10.1190/1.3213533.

Xie, X.-B., H. Ning, and B. Chen, 2016, How scatterings from small-scale near-surface heterogeneities affecting seismic data and the quality of depth image, analysis based on seismic resolution: 76th Annual International Meeting, SEG, Expanded Abstracts, 4278–4282, https://doi.org/10.1190/segam2016-13780223.1.

Zebker, H. A., and Y. Lu, 1998, Phase unwrapping algorithms for radar interferometry: Residue-cut, least-squares, and synthesis algorithms: Journal of the Optical Society of America. A, Optics, Image Science, and Vision, **15**, no. 3, 586–598, https://doi.org/10.1364/JOSAA.15.000586.