Seismic phase spectral analysis: Field-data insights from circular statistics

Akshika Rohatgi, Andrey Bakulin, and Sergey Fomel, Bureau of Economic Geology, University of Texas at Austin*

Summary

We introduce circular statistics as a robust tool for quantifying phase coherence in the spectral domain. Applied to real 3D prestack land seismic data from the continental US, this approach provides a data-driven framework for assessing phase stability across frequencies. While conventional seismic quality analysis emphasizes amplitude fidelity, our method shifts focus to phase fidelity by leveraging circular variance and von Mises distribution parameters, enabling direct analysis of frequency-dependent phase coherence.Our results show that phase correction via circular mean substitution significantly enhances seismic signal clarity, particularly in noise-dominated regions.

Introduction

Seismic phase is crucial in imaging workflows, influencing FWI, migration, and pre-stack inversion (van der Baan and Fomel, 2009; Fomel and van der Baan, 2014; Holt and Lubrano, 2020; Bakulin et al., 2022a, 2024). However, noise and scattering effects often compromise phase analysis, limiting its use in conventional processing.

Previous approaches aimed to enhance phase coherence through local stacking in the time domain, followed by phase substitution (Bakulin et al., 2021). While these methods improved phase alignment, they lacked a clear statistical framework to quantify phase variability across frequencies.

Here, we introduce a spectral-domain approach leveraging circular statistics (Fisher, 1993; Mardia and Jupp, 2009). This approach captures not just the estimated signal phase but the entire phase distribution at each frequency. It enables both targeted phase enhancement and quantitative assessment of phase coherence. Applied to field land prestack seismic data, the proposed approach improves phase fidelity and provides an objective measure of processing effectiveness.

Quantifying phase coherence in seismic data

Seismic phase is a crucial yet often overlooked aspect of data analysis, mainly because of the lack of rigorous metrics to track its variability. While seismic processing workflows indirectly improve phase behavior, they primarily focus on amplitude spectra and signal-to-noise ratio (SNR), with no widely accepted metric to evaluate phase stability in the frequency domain.

A key challenge is that the unwrapped phase is difficult to estimate accurately, while the wrapped phase is discontinuous and statistically challenging to analyze. Conventional methods lack a robust approach to quantifying phase coherence across frequencies.We propose circular statistics as a novel framework for assessing phase coherence. Unlike linear statistical methods, circular statistics effectively handle the periodic nature of phase data, allowing direct quantification of phase stability across frequencies. We avoid mixing different frequency components by analyzing phase frequency by frequency, thus preserving critical phase information.

Circular Mean: The mean of the circular data is computed differently from linear data, considering the periodic nature of the values. Given a set of phases, $(\theta_1, \theta_2, \ldots \theta_N)$, the circular mean is computed as:

$$\bar{\theta} = tan^{-1} \frac{\sum(\sin \theta_i)}{\sum(\cos \theta_i)}.$$

The circular mean provides a statistically robust estimate of seismic phase, even in noisy data, by averaging out random phase distortions at each frequency.

Circular variance: Circular variance measures the spread of angular data around the mean direction. It is defined as

$$V = 1 - R\ ,$$

where R is the mean resultant length, given by:

$$R = \frac{1}{N}\sqrt{(\sum cos\ \theta_i)^2 + (\sum sin\ \theta_i)^2}.$$

Circular variance provides a data-driven measure of phase coherence, ranging from 0 (stable, coherent phases) to 1 (random, noise-dominated phases). Intermediate values indicate varying levels of phase stability and noise contamination. By computing circular variance within localized time-frequency windows, we can objectively assess phase coherence at any stage of the analysis. Since seismic noise often results in symmetric phase distributions, the von Mises distribution serves as an appropriate statistical framework for modeling phase dispersion.

von Mises Distribution: The von Mises distribution (von Mises, 1918) is a fundamental probability distribution for circular data, analogous to the normal distribution in linear statistics. Seismic noise typically results in symmetric circular distributions well-modeled by the von Mises distribution with the probability density function given by

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

where $I_0(x)$ is the modified Bessel function of the first kind, $\overline{\theta}$ represents the mean direction, and κ is the concentration parameter controlling the clustering around the mean. Circular variance V and the concentration parameter κ are directly related and interconvertible, with the precise relationship given by Fisher (1993). The von Mises distribution offers a data-driven method for characterizing

10.1190/image2025-4313698.1

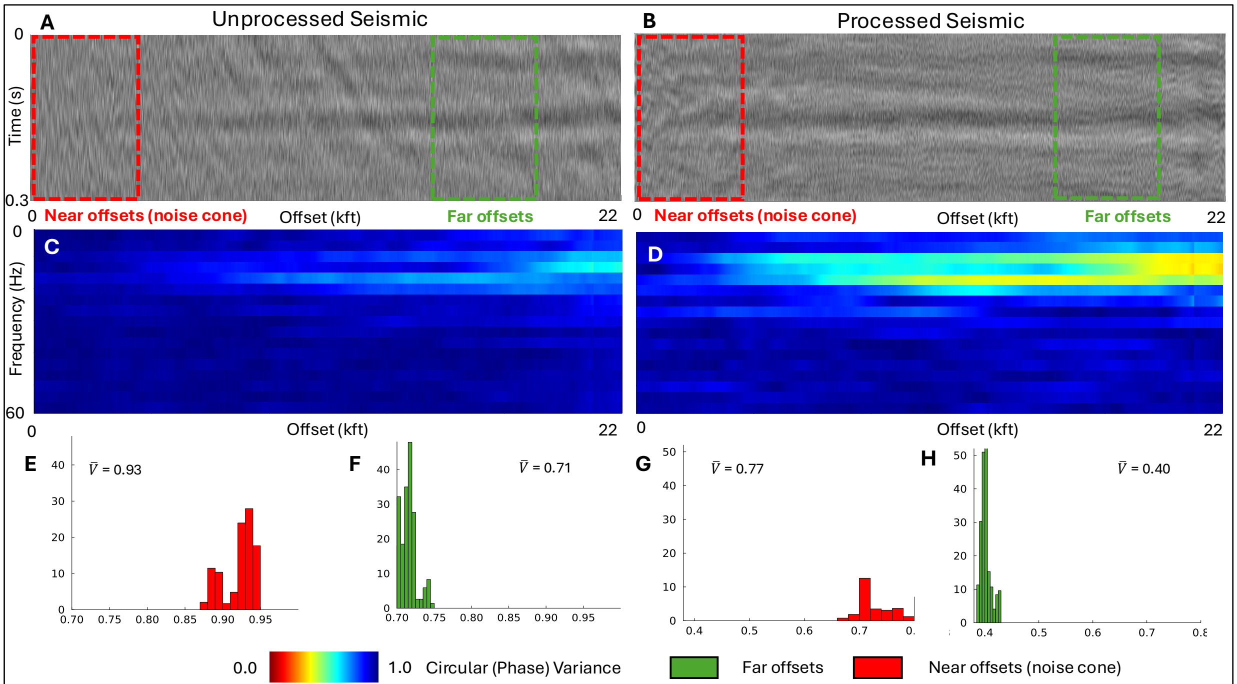

Figure 1: Comparison of unprocessed and processed prestack seismic land data: (A) and (B) show seismic sections before and after processing, with near-offset windows corresponding to data inside the noise cone and far-offset windows marked for comparison; (C) and (D) present frequency-offset heatmaps, illustrating how processing reduces phase variance and enhances coherence across different offsets; (E), (F), (G), and (H) quantify circular variance ($\bar{V}$) at 16 Hz, demonstrating a clear reduction in phase variability after processing, leading to improved signal consistency and continuity.

seismic phase coherence: large κ corresponds to high phase coherence (low variance), and small κ indicates random or noise-driven phase distributions (high variance).

Field example of land prestack seismic data

To illustrate our approach, we analyze a small, identical time window after NMO and statics corrections, comparing seismic phase statistics from the same prestack continental US land seismic data before (Figure 1A) and after final processing (Figure 1B). To ensure statistical robustness, we extract a prominent deep reflector from a 3D CMP supergather containing 10,000 traces sorted by offset.

The circular statistics framework allows for efficient and interpretable analysis of seismic phases directly in their wrapped form. Circular variance, computed frequency-by-frequency within sliding time windows, quantitatively captures phase variability as a function of offset and frequency (Figures 1C and 1D). The circular variance maps (Figures 1C and 1D) provide a direct, quantitative visualization of phase stability, revealing clear spatial trends: near offsets, corresponding to data inside the noise cone, suffer from higher variance (low coherence), whereas mid-to-far offsets exhibit significantly improved phase stability. Visual inspection aligns with these quantitative metrics: at 16 Hz, the near-offset window in unprocessed data has a variance of 0.93, close to pure noise, whereas mid-to-far offsets yield a lower variance of 0.71, indicating enhanced but still imperfect coherence (Figures 1E–F). The same trend persists after processing but with notably reduced variance levels (Figures 1G–H).

Time processing clearly reduces noise, a result visually evident when comparing Figures 1A and 1B. However, visual assessments alone can lead to subjective interpretations. In contrast, data-driven metrics like circular variance objectively quantify improvements, removing ambiguity. For instance, at 16 Hz, circular variance for mid-offsets decreases significantly from 0.71 (Figure 1F) to 0.40 (Figure 1H), quantitatively demonstrating substantial processing-driven phase enhancement, although still indicating a noise-dominated environment. Connecting this metric to the more familiar concept of signal-to-noise ratio (SNR), Figure 3A confirms this observation: initial prestack SNR is approximately -30 dB, improving only to around -18 dB after processing, underscoring persistent noise dominance in prestack seismic data.

In addition to assessing phase coherence across offsets, we also examine its behavior across the frequency spectrum. Our analysis shows that higher coherence is confined to a

10.1190/image2025-4313698.1

narrow band around 10–20 Hz, with a sharp decline below 10 Hz and above 25 Hz. This is visually evident in the frequency-variance heatmaps in Figures 1C and 1D, where phase coherence degrades outside this range. The loss of coherence at low frequencies is likely due to the vibrator's limited efficiency in exciting deep-penetrating low frequencies. As for the rapid degradation above 25 Hz, it may result from a combination of scattering noise and residual statics, both of which amplify phase distortions. The fact that coherence drops off precisely at 25 Hz suggests that only a narrow frequency band effectively contributes to resolution in this deep data window despite the broadband vibrator sweep. This trend is further highlighted in Figure 3C, where a single-location cross-section of variance as a function of frequency clearly illustrates that phase coherence shows no improvement above 25 Hz between processed and unprocessed data, despite the expectation that convolution-based processing significantly broadens the amplitude spectrum, as seen in Figure 3B.

Phase correction via circular mean

We numerically demonstrate the effectiveness of circular statistics in predicting seismic phase by computing the circular mean within a sliding window of 2,000 traces, frequency by frequency, and applying it in a phase substitution process (Bakulin et al., 2020a). Selecting an appropriate ensemble size is crucial for obtaining an accurate estimate of the signal phase (Rohatgi et al., 2024). This approach retains the original unprocessed raw data in amplitude while replacing its highly contaminated phase with the circular mean estimate, effectively recovering a clean signal phase.

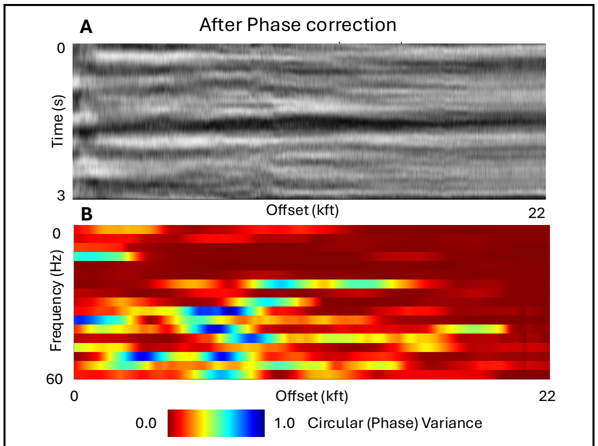

Figure 2: Phase Correction Using Circular Mean: (A) Unprocessed raw data after phase correction using the circular mean. (B) Corresponding circular phase variance map, showing a significant improvement in phase coherence across all frequencies and offsets compared to Figure 1C.

This approach parallels local stacking techniques (Bakulin et al., 2018, 2020b) but operates in the frequency domain using circular statistics. Despite the difference, it yields comparable signal quality improvements. We systematically correct phase distortions and enhance data clarity by substituting the estimated phase into the original seismic data. Figure 2A presents the seismic section after phase correction, showing clearer subsurface reflectors across both near and far offsets. Transforming the modified seismic traces back to the time domain reveals a significant improvement in structural continuity. Furthermore, an analysis of the circular variance in the frequency domain (Figure 2B) confirms improved phase coherence, reinforcing the effectiveness of this correction.

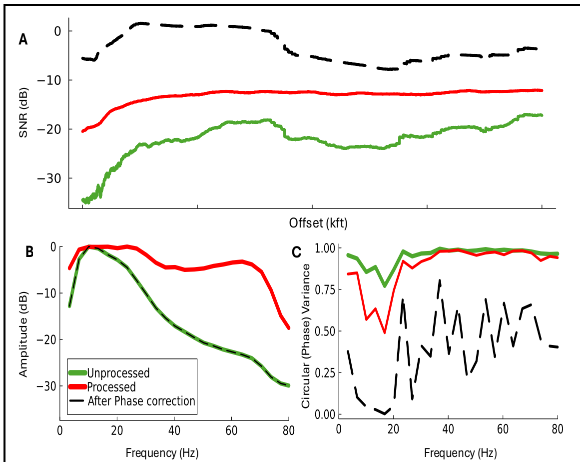

Figure 3: Quantitative assessment of three data versions (raw, processed, phase-substituted): (A) Stack-based SNR vs. offset, showing improvements with processing and further enhancement after phase correction. (B) Amplitude spectra for the three datasets, highlighting spectral preservation in raw and enhanced data, while processed data shows a flattened band; (C) Circular phase variance vs. frequency, showing no improvement in coherence above 25 Hz with processing, while phase correction significantly enhances coherence at higher frequencies.

To quantify the improvement, we evaluate signal-to-noise ratio (SNR) and phase coherence across different processing stages (Figure 3). The SNR, computed using the stack-based semblance method (Bakulin et al., 2022b), shows that conventional processing increases SNR from approximately -25 dB to -15 dB. However, applying phase substitution alone to raw data, without altering amplitude content, improves SNR dramatically to ~ -5 dB (Figure 3A).

Amplitude spectra analysis (Figure 3B) further illustrates the impact of different processing approaches. By design, phase-corrected data preserves the original amplitude spectrum. In contrast, conventional processing artificially boosts high frequencies due to various deconvolution steps applied in the time domain (red curve in Figure 3B). This creates a misleading impression of broadening the bandwidth and improving resolution.

10.1190/image2025-4313698.1

However, a different picture emerges when we examine phase coherence across frequencies using circular variance as a function of frequency (Figure 3C). The variance of conventionally processed data remains nearly identical to unprocessed data above 25 Hz, with values close to 1- indicating that this frequency range is dominated by pure noise. This suggests that surface-consistent scaling or deconvolution fails to improve signal quality when applied to low-SNR data (Cary and Nagarajappa, 2013). Instead, conventional processing amplifies noise energy at high frequencies without enhancing or recovering the true signal.

The phase-corrected data (black dashed line in Figure 3C) exhibits significantly lower variance between 25 and 80 Hz, confirming a recovery of clean signal even in previously noise-dominated frequency bands.

Phase distributions and their statistical interpretation

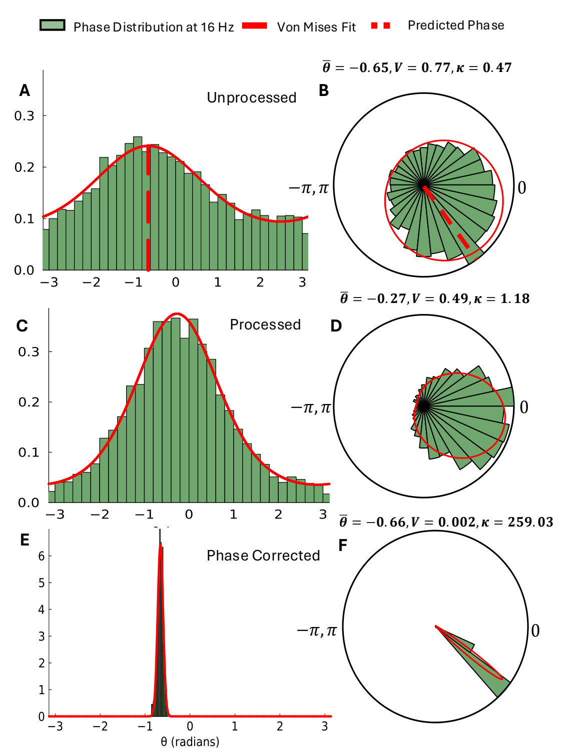

Figure 4: Phase distribution at 16 Hz for unprocessed, processed, and phase-corrected seismic data (A, C, E) phase histograms fitted with a von Mises distribution (red curve), demonstrating that phase distributions exhibit circular symmetry. (B, D, F) circular phase distributions, showing how processing and phase correction influence phase concentration. The von Mises model accurately represents phase behavior, with tighter distributions indicating improved coherence.

To further assess phase coherence improvements, we examine the actual phase distributions from field data, showing how they conform to the von Mises distribution and how processing, particularly phase substitution, tightens these distributions both visually and statistically. Phase values are inherently wrapped within [-π, π], making their statistical behavior distinct from conventional linear distributions. Circular polar plots provide a more intuitive way to interpret phase behavior, as shown in Figures 4B, D, and F. The unprocessed data's phase distribution appears broad, indicating low coherence (Figures 4A and 4B). The circular variance is measured as 0.77, suggesting strong phase perturbations and noise dominance. This distribution closely resembles a low-concentration von Mises model, confirming the statistical nature of random phase variations. Following conventional processing, phase variance decreases to 0.49, indicating partial improvement in phase coherence (Figure 4C and 4D). However, the phase distribution remains relatively broad, showing that standard processing does not fully restore coherent phase alignment. In contrast, phase substitution generates a much tighter phase distribution, with variance significantly lower (Figure 4E and 4F). This suggests that the circular mean effectively recovers the true signal phase, even in noise-dominated regions. The resulting phase distribution aligns closely with a von Mises model with a higher concentration parameter, demonstrating a clear improvement in statistical phase stability. These results confirm that circular statistics not only quantify phase coherence but also provide an objective framework for evaluating processing effectiveness.

Conclusions

Using a field data example, we have demonstrated that circular statistics can provide a robust, data-driven approach for quantifying phase coherence in seismic data. By leveraging circular variance and the von Mises distribution, we objectively assess phase stability, track processing effectiveness, and identify frequency bands requiring correction. Applied to prestack land seismic data from the continental US, our method reveals that conventional processing does little to improve phase coherence at high frequencies, where noise dominates. In contrast, phase correction via circular mean substitution effectively recovers the true signal phase, even in noisy regions. The strong agreement with the von Mises model further validates the proposed statistical framework. Beyond analyzing individual datasets, circular statistics offer an unbiased tool for evaluating acquisition and processing techniques, optimizing workflows, and improving seismic imaging.

Acknowledgments

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

REFERENCES

Bakulin, A., P. Golikov, M. Dmitriev, D. Neklyudov, P. Leger, and V. Dolgov, 2018, Application of supergrouping to enhance 3D prestack seismic data from a desert environment: The Leading Edge, **37**, no. 3, pp. 200–207.

Bakulin, A., I. Silvestrov, and D. Neklyudov, 2020a, Importance of phase guides from beamformed data for processing multi-channel data in highly scattering media: The Journal of the Acoustical Society of America, **147**, no. 6, pp. EL447–EL452.

Bakulin, A., I. Silvestrov, M. Dmitriev, D. Neklyudov, M. Protasov, K. Gadylshin, and V. Dolgov, 2020b, Nonlinear beamforming for enhancement of 3D prestack land seismic data, Geophysics, **85**, no. 3, pp. V283–V296.

Bakulin, A., D. Neklyudov, and I. Silvestrov, 2021, Prestack data enhancement with phase substitution and phase corrections guided by local multi-dimensional stacking: 82nd EAGE Annual Conference & Exhibition, EAGE, 1–5.

Bakulin, A., D. Neklyudov, and I. Silvestrov, 2022a. Multiplicative random seismic noise caused by small-scale near-surface scattering and its transformation during stacking: Geophysics, **87**, no. 5, pp. V419–V435.

Bakulin, A., I. Silvestrov, and M. Protasov, 2022b, Research note: Signal-to-noise ratio computation for challenging land single-sensor seismic data: Geophysical Prospecting, **70**, no. 3, pp. 629–638.

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, pp. V551–V561.

Cary, P., and N. Nagarajappa, 2013, Questioning the basics of surface-consistent scaling: 83rd Annual International Meeting, SEG.

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

Fomel, S., and M. van der Baan, 2014. Local skewness attribute as a seismic phase detector: Interpretation, **2**, no. 1, pp. SA49–SA56.

Mardia, K. V. and P. E. Jupp, 2009, Directional statistics: John Wiley & Sons.

Rohatgi, A., A. Bakulin, and S. Fomel, 2024. 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.

van der Baan, M. and S. Fomel, 2009, Nonstationary phase estimation using regularized local kurtosis maximization, Geophysics, **74**, no. 6, pp. A75–A80.

von Mises, R., 1918, “Über die “Ganzzahligkeit” der Atomgewichte und verwandete Fragen.”: Physikalische Zeitschrift, **19**, 490.

10.1190/image2025-4313698.1