Frequency-domain correlation of nonuniform time series: Coherence across distant wind stations

Bill D. Caraway,[1] Ryan Bogucki,[2] Ideal Rafuna,[1] Manuel Reyna,[3] Shyuan Cheng,[4] and Leonardo P. Chamorro[1, 3, 5, 6, a)]

[1)]*Department of Mechanical Science and Engineering, University of Illinois, Urbana, IL, 61801, USA.*

[2)]*Department of Physics, University of Illinois, Urbana, IL, 61801, USA.*

[3)]*Department of Civil and Environmental Engineering, University of Illinois, Urbana, IL, 61801, USA.*

[4)]*Department of Mechanical Engineering, University of South Carolina, Columbia, SC 29201, USA.*

[5)]*Department of Aerospace Engineering, University of Illinois, Urbana, IL, 61801, USA.*

[6)]*Department of Geology, University of Illinois, Urbana, IL, 61801, USA.*

(Dated: 12 November 2025)

We analyze multi-year wind velocity records from six U.S. airports to examine spatial and temporal correlations across the national airspace system. The velocity time series, characterized by uneven sampling and differing time bases, are evaluated using a frequency-domain approach based on the nonuniform discrete Fourier transform. This Fourier-domain based approach enables direct estimation of cross-covariance and cross-spectral relationships without interpolation or resampling in the time-domain, preserving variance and phase integrity. Applied to airport wind data, the analysis reveals scale-dependent and spatially organized correlations. We demonstrate that coherence in the inertial subrange decays rapidly with distance, while diurnal and annual components persist over much broader separations. The results show the scale-dependent organization of atmospheric variability across geographically distributed sites and highlight the utility of frequency-domain analysis for irregular environmental datasets where uniform sampling cannot be assumed.

[a] Author to whom correspondence should be addressed:lpchamo@illinois.edu

I.   INTRODUCTION

The statistical characterization of near-surface winds is needed for several engineering applications, including wind power[1,2], particle dispersion and airport traffic management. Modeling such flows is an evolving problem, where establishing interconnections between flow behaviors across distant locations may offer key insight. This phenomenon may be instrumental for improving wind velocity predictions, requiring a deeper exploration into the multiscale turbulence processes.

Velocity autocorrelation and cross-covariance and the associated counterparts in frequency domain, the power spectrum and cross-spectrum provide a baseline for assessing the basic dynamics of turbulent flows[3]. These quantities offer unique perspectives on the temporal scales of turbulence. When applied across disparate spatial locations, these techniques may provide information into the spatial features of turbulent flows and inter-dependence of velocity data. Particularly in the context of wind energy, smoothing effects achieved by geographically distributed wind turbines are of particular relevance[4]. This phenomenon relates to the collective capacity of dispersed turbines to mitigate the overall variability in power output from wind speed fluctuations. The primary metric for assessing this capability is correlation[5], based on the concept that integrating statistically independent turbines reduces joint power variability, a relationship quantitatively described by the square root of the turbine count as per the central limit theorem. Moreover, grouping turbines that are negatively correlated can speed up variability reduction.

The cross-covariance $C_{\alpha\beta}(\tau)$, or its nondimensional counterpart the cross-correlation $\rho_{\alpha\beta}$, measures the relation between two signals $\alpha(t)$ and $\beta(t+\tau)$ as a function of the time lag τ. This approach sheds light on the time-dependent interactions within a turbulent flow, offering a way to assess features of signal fluctuations based on the other's patterns. The cross-spectrum $E_{\alpha\beta}(f)$, representing the frequency domain counterpart of cross-covariance, shows the degree and temporal dynamics of signal correlation at specific frequencies. This is instrumental in identifying coherent structures within turbulent flows and understanding energy transfer across various scales and locations[6]. The application of the cross-spectrum in evaluating the smoothing effect of distributed turbines illustrates its utility in wind power studies[7,8]. The power spectrum $\phi(f)$ is central for analyzing a turbulent signal's frequency content and decomposing turbulent kinetic energy across frequencies. It is determined by

summing $|E_{u_i u_i}(f)|$ for $i = 1, 2, 3$, where $u_1(t)$, $u_2(t)$, and $u_3(t)$ represent the orthogonal velocity components. Though sometimes limited by the availability of measurements, the power spectrum facilitates the identification of dominant frequencies and describes energy distribution across turbulence scales.

A common challenge in these analyses is collecting the required data. Longer sampling durations and higher sampling rates tend to improve precision of the analysis, but there are often practical limits on these parameters. In some cases, analysis techniques have been developed to exploit known features of the data to simplify data collection. For example, compressed sensing exploits the sparsity or compressibility of many natural and physical signals in an appropriate transform domain (e.g., Fourier, wavelet, or empirical mode decomposition) to reconstruct full signals from a limited number of nonuniform or incomplete measurements, often below the Nyquist-Shannon sampling rate[9,10]. For environmental flows, compressed sensing may offer significant promise as complementary tools to traditional time-series analysis, enabling the analysis of nonuniform datasets without excessive interpolation or data loss, while still capturing essential structures and dynamic behaviors[11]. For data with strong but extremely narrow frequency features, such as in quasi-periodic phenomena, several types of techniques exist to identify these peaks with greater precision than would normally be possible for a given time series. These techniques refine the peak frequencies by a) using nonorthogonal temporal basis frunctions[12], b) collocation[13], c) autoregressive moving-average techniques[14,15], or d) averaging-extrapolation methods[16]. However, many phenomena, including most turbulent flows, do not have significant sparsity in the power spectra and have broad spectrum frequency content of significant power.

A particular concern with traditional analyses is their dependency on uniformly sampled data, a criterion not always feasible due to diverse sampling intervals or maintenance-induced data collection interruptions. This issue is particularly pronounced in long-term studies, or datasets aggregated from various sources without synchronized sampling efforts. For instance, we examine wind data collated over three years from six U.S. airports, highlighting the challenges in measurement times and sampling rates within and across these datasets. Standard calculation methods for cross-correlations, cross-spectra, and power spectra would require interpolation or substantial downsampling to standardize the timestep size, inevitably increasing result uncertainty.

We introduce an approach based on shifted Discrete Fourier Transform (DFT) to fulfill the

Nonuniform Discrete Fouroer Tramsform (NUDFT)[17,18] for computing the cross-covariance of nonuniformly sampled data through a frequency domain transformation. This method calculates the total cross-correlation and directly determines the cross-spectrum without requiring additional computational overhead. Furthermore, it facilitates a novel representation of cross-correlation within the frequency domain, the cumulative spectral cross-correlation based on shifted DFT that enables spectral analysis of signals with arbitrary shifts in time or frequency domains. While the standard DFT assumes a fixed sampling grid and computes frequency components based on equally spaced samples, the shifted DFT introduces flexibility by allowing non-integer or fractional shifts. This is especially useful in applications where signal alignment, frequency interpolation are required[19]. The methodology and its application are detailed in Section II, followed by a discussion in Section III, with final conclusions presented in Section IV.

II. METHODS

A. Cross-Covariance

For two infinite, continuous, and statistically stationary signals, $\alpha(t)$ and $\beta(t)$, the cross-covariance is defined as:

$$C_{\alpha\beta}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} [\alpha(t) - \overline{\alpha}][\beta(t + \tau) - \overline{\beta}] \, dt, \tag{1}$$

where $\overline{\alpha}$ and $\overline{\beta}$ represent the time averages of $\alpha(t)$ and $\beta(t)$, respectively. However, practical applications often involve uniformly sampled time series of finite duration T and a finite time series with I measurements; as opposed to infinite, continuous signals. Under these conditions, the cross-covariance of the underlying continuous signals can be approximated by:

$$C_{\alpha\beta}(\tau) \approx \frac{1}{T} \sum_{i=0}^{I-1} (\alpha_i - \overline{\alpha})(\beta_{i+\frac{\tau}{\Delta t}} - \overline{\beta}), \tag{2}$$

where $\Delta t = T/I$ denotes the sampling time step, with $\alpha_i = \alpha(i\Delta t)$ and $\beta_i = \beta(i\Delta t)$ representing the i-th measurement in the sampled time series. For indices i that exceed I, β_i is considered to be 0. The lag τ is restricted to nonnegative integer multiples of Δt. Negative multiples can be accounted for by using the fact that $C_{\alpha\beta}(-\tau) = C_{\beta\alpha}(\tau)$.

A significant limitation of this approximation is its assumption of uniform sampling. For nonuniform sampling, the $\tau = 0$ case might be manageable by incorporating weighting factors. However, for $\tau > 0$, it becomes problematic as mismatch may occur between the uniformly spaced $i-$th lag $\tau_i = i \times \Delta t$ and the sampling instant for the $i-$th measurement α_i and β_i. To overcome this issue, we shift our analysis from the time domain to the frequency domain, where the constraints of uniform sampling are less restrictive, and the impact of nonuniform sampling can be more easily mitigated.

B. Transition to the Frequency Domain

By applying the Discrete Fourier Transform (DFT) or, for irregular sampling, the Nonuniform Discrete Fourier Transform (NUDFT)[20,21] to the time series, we approximate the signals $\alpha(t)$ and $\beta(t)$ as

$$\alpha(t) \approx \overline{\alpha} + \sum_{n=1}^{N} \tilde{\alpha}_n \cos\left(2\pi \frac{n}{T} t + \theta_{\alpha,n}\right), \tag{3a}$$

$$\beta(t) \approx \overline{\beta} + \sum_{n=1}^{N} \tilde{\beta}_n \cos\left(2\pi \frac{n}{T} t + \theta_{\beta,n}\right), \tag{3b}$$

where N is the total measurement numbers in the time series, $\tilde{\alpha}_n$ and $\theta_{\alpha,n}$ are the amplitude and phase of the $n-$th mode, and N/T denotes the dominant ~~approximate~~ sampling frequency in the time series (e.g., based on the mean or dominant sampling interval); the corresponding effective Nyquist frequency is $f_N = (N/T)/2$. We use NUFFT implementation (a gridding-based, oversampled-FFT approach rather than an ℓ_1-sparse recovery) adopted from Potter, Gumerov, and Duraiswami[18], which efficiently estimates amplitudes and phases on a prescribed frequency grid for nonuniform samples under a band-limited signal assumption.

For time series that are uniformly sampled, Equations 3 provide an exact representation at the sampled points. The NUDFT generalizes this to nonuniformly sampled series, providing estimates of the Fourier amplitudes and phases over a prescribed frequency grid. A common frequency grid is adopted to ensure consistency across time series with different sampling intervals, enabling coherent spectral comparisons even though the basis functions are not strictly orthogonal under nonuniform sampling.

By substituting Equations 3 into Equation 2 and using the orthogonality of the continuous

trigonometric basis on $[0, T]$, we obtain

$$C_{\alpha\beta}(\tau) \approx \sum_{n=1}^{N} \frac{\tilde{\alpha}_n \tilde{\beta}_n}{2} \cos\left(-2\pi\frac{n}{T}\tau + \theta_{\alpha,n} - \theta_{\beta,n}\right), \tag{4}$$

The (one-sided) cross-spectrum amplitudes at $f_n = n/T$ are then

$$\left|E_{\alpha\beta}(f_n)\right| \approx \frac{\tilde{\alpha}_n \tilde{\beta}_n}{2}, \qquad 1 \leq n \leq N, \tag{5}$$

and the zero-lag cross-correlation coefficient, $\rho_{\alpha\beta} = C_{\alpha\beta}(0)/[C_{\alpha\alpha}(0)\, C_{\beta\beta}(0)]^{1/2}$, is

$$\rho_{\alpha\beta} \approx \left[\left(\sum_{p=1}^{N} \tilde{\alpha}_p^2\right)\left(\sum_{q=1}^{N} \tilde{\beta}_q^2\right)\right]^{-1/2} \sum_{n=1}^{N} \tilde{\alpha}_n \tilde{\beta}_n \cos(\theta_{\alpha,n} - \theta_{\beta,n}). \tag{6}$$

This transition to the frequency domain provides clarity on aspects of signal behavior that may not be apparent in the time domain. Equation 4 illustrates that two periodic signals with a phase difference of $\pi/2$ exhibit minimal correlation at zero lag, and their cross-correlation reaches maximum or minimum values at time lags corresponding to one quarter of the signal period. Furthermore, performing cross-covariance analysis in the frequency domain simplifies the computation of cross-correlation coefficients for nonuniformly sampled data or for signals with differing sampling rates. This approach avoids time-domain interpolation or data decimation, thereby minimizing distortion and preserving the integrity of the original datasets.

However, it is important to acknowledge a few considerations. Large gaps in the data, where the intervals between measurements greatly exceed the nominal sampling interval, reduce the accuracy of the NUDFT approximation in Equation 3. Such gaps increase spectral leakage and lower the effective spectral resolution. Although the frequency-domain approach mitigates many of the difficulties associated with nonuniform sampling, it remains sensitive to extended periods of missing data. Irregular sampling can also generate spurious high-frequency components in the cross-spectrum due to sampling jitter or uneven measurement spacing, resulting in artificial spectral amplification near the effective sampling rate that would not occur under uniform sampling.

Using the frequency domain representation of the cross-correlation coefficient, we can evaluate the partial sums up to f_n to examine how frequency modes contribute to the correlation. The cumulative spectral cross-correlation for nonuniform sampling is defined as

$$\gamma_{\alpha\beta}(f_n) = k\left[\left(\sum_{p=1}^{N} \tilde{\alpha}_p^2\right)\left(\sum_{q=1}^{N} \tilde{\beta}_q^2\right)\right]^{-1/2} \sum_{m=1}^{n} \tilde{\alpha}_m \tilde{\beta}_m cos(\theta_{\alpha,m} - \theta_{\beta,m}), \tag{7}$$
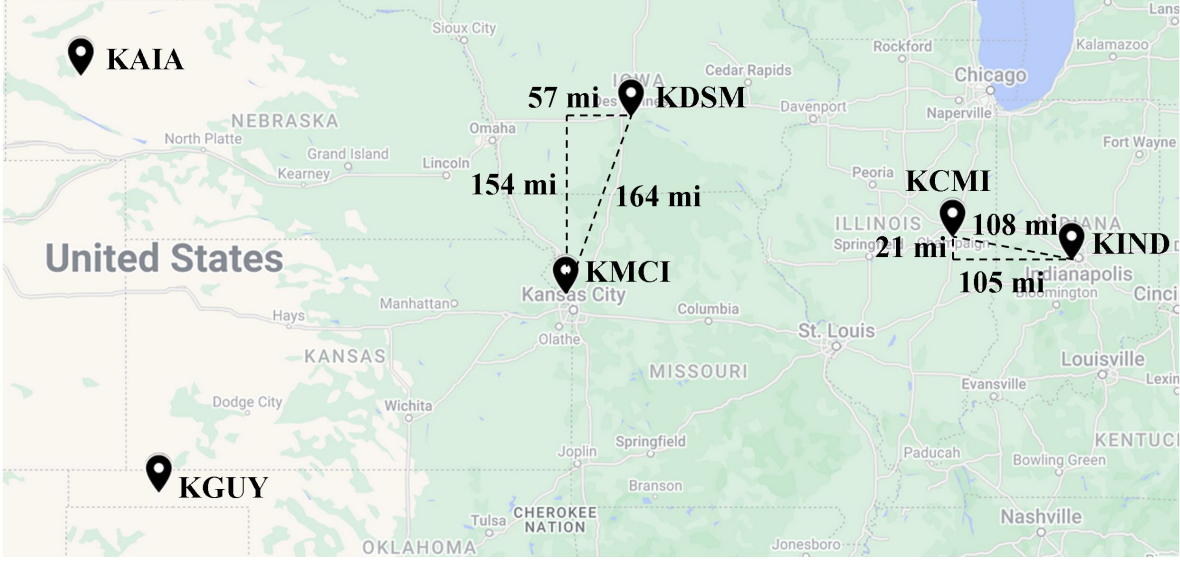
where $f_n = n/T$ and $1 \le n \le N$.

FIG. 1. Geographical locations of the selected Automated Surface Observing Systems (ASOS). Approximate great-circle distances and their longitudinal and latitudinal components are shown for the KCMI–KIND and KDSM–KMCI station pairs.

TABLE I. Approximate great-circle distances (in miles) between the selected Automated Surface Observing Systems (ASOS) stations shown in Figure 1.

	KIND	KDSM	KMCI	KAIA	KGUY
KCMI	108	300	347	768	752
KIND	–	406	451	876	851
KDSM	–	–	164	472	537
KMCI	–	–	–	463	410
KAIA	–	–	–	–	378

III. APPLICATION TO AIRPORTS WIND DATA

A. Airport Wind Data

The methodology given in Section II is used to analyze near-surface wind speed records (i.e., heights between 5 to 10 m) from six U.S. airports within the National Weather Service Automated Surface Observing System (ASOS) network[22]. The selected stations are KCMI
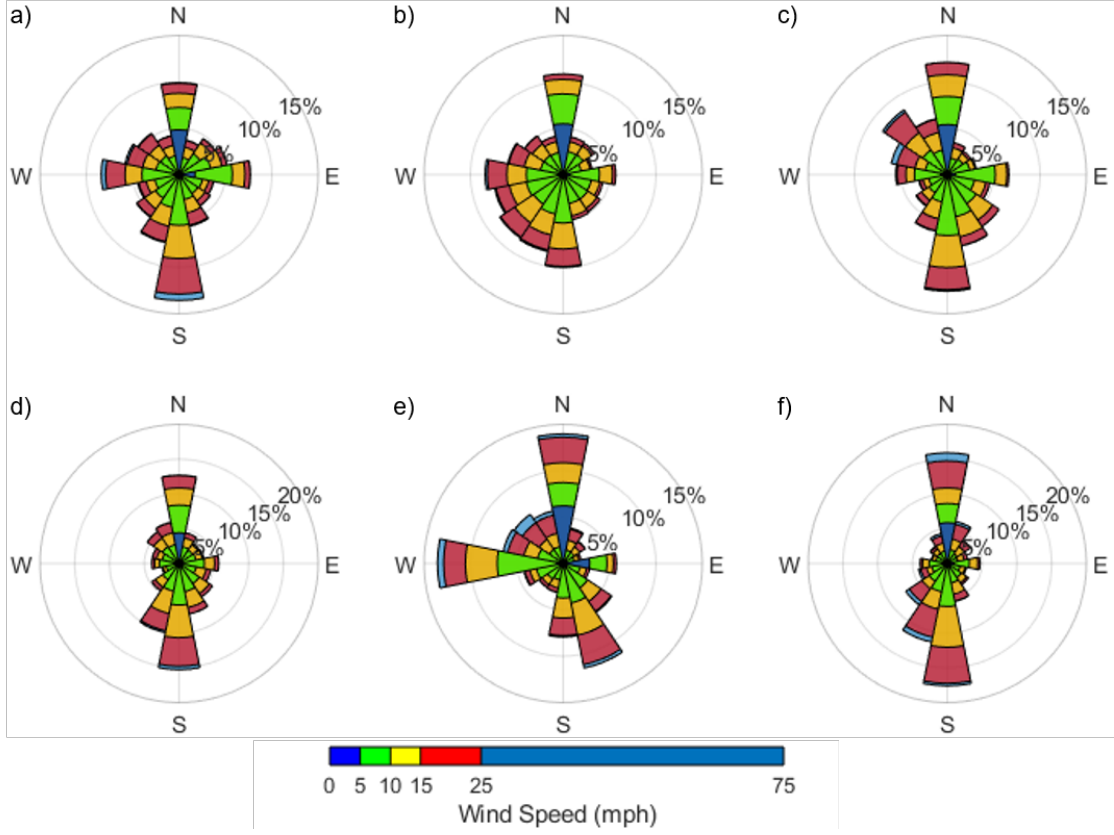
FIG. 2. Wind roses for the selected Automated Surface Observing Systems (ASOS) stations: (a) KCMI, (b) KIND, (c) KDSM, (d) KMCI, (e) KAIA, and (f) KGUY. Wind speeds between 0–5 mph are shown in dark blue, 5–10 mph in green, 10–15 mph in yellow, 15–25 mph in red, and greater than 25 mph in light blue. The maximum recorded wind speed was 66 mph.

(University of Illinois Willard Airport, Savoy, IL), KIND (Indianapolis International Airport, IN), KDSM (Des Moines International Airport, IA), KMCI (Kansas City International Airport, MO), KAIA (Alliance Municipal Airport, NE), and KGUY (Guymon Municipal Airport, OK), spanning the central United States, providing a range of intersite separations for examining spatial coherence. The station locations are shown in Figure 1, and their approximate straight-line distances, listed in Table I, serve as the spatial basis for correlation and spectral analyses[23].

For each location, wind speed and direction data were collected continuously over a three-year period, from 00:00 on January 1, 2020, to 23:59 on December 31, 2022. The corresponding wind roses are shown in Figure 2. Data were collected nominally every five minutes, with infrequent longer gaps producing minor nonuniformity among stations. The
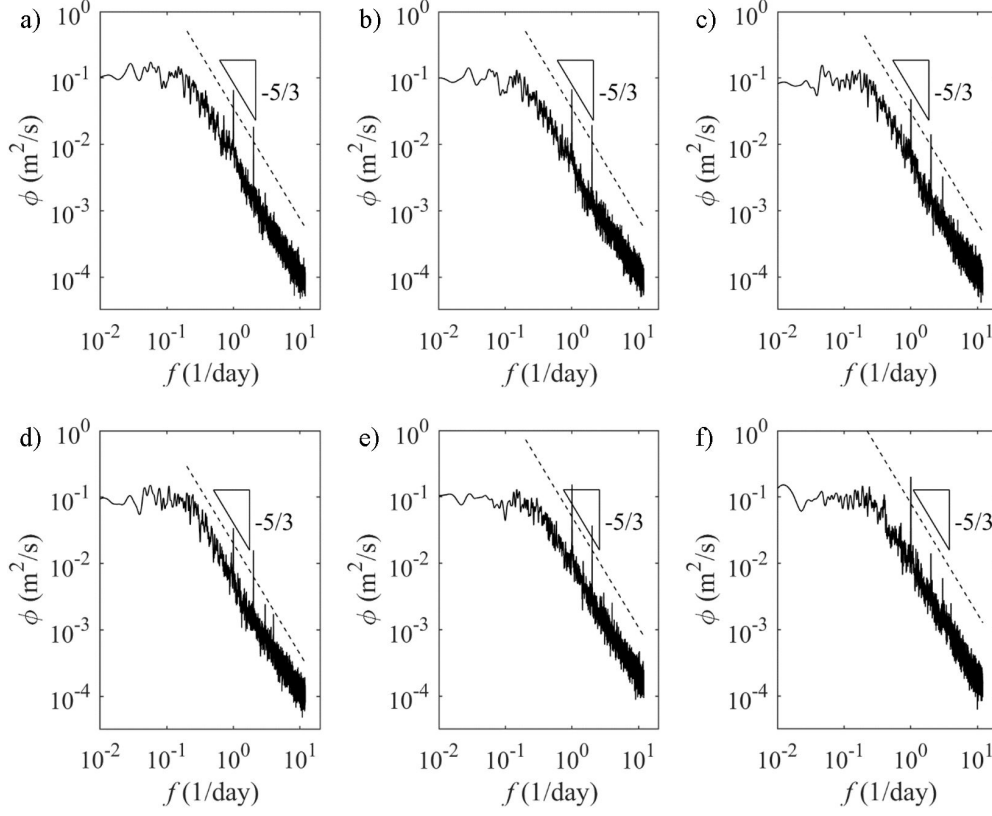
FIG. 3. Power spectra, $\phi(f)$, for the selected Automated Surface Observing Systems (ASOS) stations: (a) KCMI, (b) KIND, (c) KDSM, (d) KMCI, (e) KAIA, and (f) KGUY. The dotted lines indicate the $f^{-5/3}$ scaling. Each spectrum exhibits a pronounced peak at $f = 1$ day^{-1} and a weaker secondary peak near $f = 2$ day^{-1}.

six stations span three time zones, so the measurements were not recorded simultaneously in absolute time. These discrepancies were addressed in the frequency-domain analysis: irregular sampling intervals were handled using the NUDFT instead of the standard DFT, and temporal offsets between stations were corrected by applying a phase adjustment. This phase shift is naturally incorporated in the NUDFT formulation by aligning all measurement times to a common temporal reference, chosen here as 00:00 CST on January 1, 2020. Daylight saving time adjustments were also applied to ensure temporal consistency.

Because portions of the records were obtained at approximately hourly intervals, the effective Nyquist frequency was set to 12 day^{-1}. This threshold limits the analysis to frequencies below the range where aliasing and numerical noise become dominant, thereby suppressing spurious high-frequency modes associated with the sampling rate. This choice was supported by observations showing that the power spectra above 12 day^{-1} exhibited a
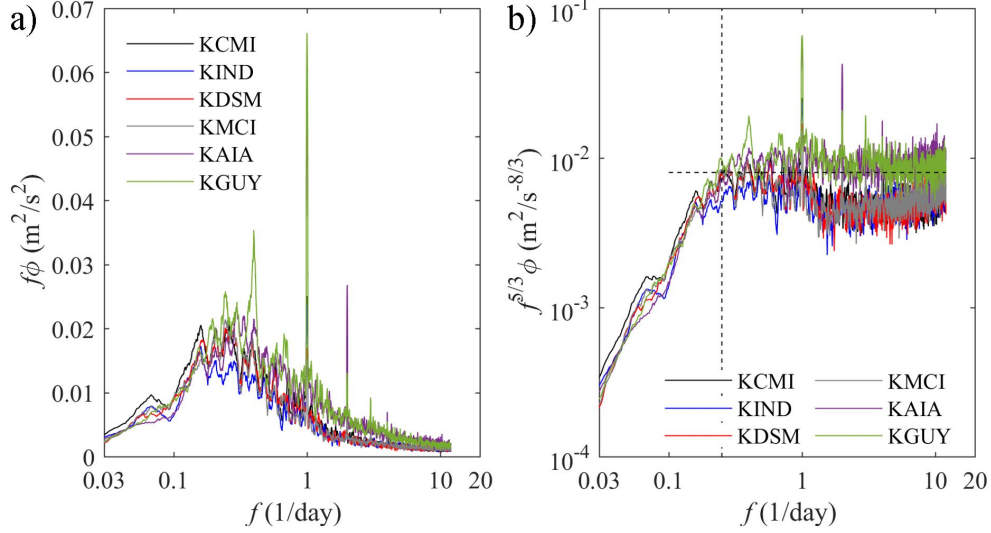
FIG. 4. Compensated power spectra: (a) $f\phi(f)$ and (b) $f^{5/3}\phi(f)$, with frequency f in day^{-1}, for the selected ASOS stations. In (b), the dotted line indicates the $f^{-5/3}$ scaling and the onset of the inertial subrange near $f \approx 0.25$ day^{-1}. In (a), the diurnal peak at $f = 1$ day^{-1} reaches $f\phi = 0.331$. The spectra for KCMI, KIND, KDSM, and KMCI exhibit comparable magnitudes and are slightly weaker at higher frequencies than those for KAIA and KGUY, which are themselves nearly identical.

nearly constant noise floor with small peaks at frequencies close to integer multiples of 24 day^{-1}, corresponding to diurnal harmonics.

In the next analysis, the wind velocity is represented by its magnitude U, the east–west component u, and the north–south component v. The two-dimensional kinetic energy spectrum is defined as

$$\phi(f) = E_{uu}(f) + E_{vv}(f), \tag{8}$$

where $E_{uu}(f)$ and $E_{vv}(f)$ denote the power spectral densities of the respective velocity components.

B. Power Spectra

The power spectra for each measurement location are shown in Figure 3. Each spectrum exhibits a well-defined inertial subrange, together with distinct peaks at $f = 1$ day^{-1} and less pronounced peaks at $f = 2$ day^{-1}. These peaks likely reflect the strong influence of

11

solar heating and other diurnal processes[24–26]. Figure 4 presents the compensated power spectra, highlighting the consistent spectral behavior across all stations. The onset of the inertial subrange occurs at approximately $f \approx 0.25$ day^{-1} for all six sites, corresponding to large-scale motions with characteristic timescales of one to ten days.

The overall spectral magnitudes are comparable across all stations, although those from KCMI, KIND, KDSM, and KMCI are slightly weaker at higher frequencies than those from KAIA and KGUY. Within each of these two groups, the spectral shapes and magnitudes are highly similar. Although wind records may exhibit some non-stationarity and variability in flow direction due to changing meteorological conditions, the frequency-domain representation remains valuable for assessing the general spectral structure of the wind field. It provides insight into the characteristic energy distribution, dominant temporal scales, and relative similarity among sites, even when the underlying dynamics are not perfectly stationary.

TABLE II. Cross-correlation coefficients, ρ_{uv}, between the east–west (u) and north–south (v) wind velocity components at each ASOS station.

KCMI	KIND	KDSM	KMCI	KAIA	KGUY
-0.04	0.04	-0.25	-0.04	-0.31	0.25

C.   Cross-correlation Coefficients

The cross correlation coefficients with zero lag, ρ_{uv}, were computed between the east and west (u) and north and south (v) components of the wind velocity at each station over the entire three year period. The results, summarized in Table II, indicate that these two horizontal components are generally weakly correlated. This behavior is consistent with expectations, since the u and v components often vary almost independently under changing synoptic and mesoscale conditions.

The magnitudes of the correlation coefficients for total wind speed $U = \sqrt{u^2 + v^2}$) closely match those observed between the individual components (refer to Table III and IV). In addition, the cross-correlation coefficients decrease with increasing distance between stations, indicating a clear spatial dependence in wind velocity coherence. Certain station pairs, such as KCMI and KIND, which are separated by more than 100 miles, exhibit relatively strong correlations even though the measurements were taken at modest heights of 5 to 10 meters.

This suggests that large-scale atmospheric structures or regional forcing can impose coherent variability over considerable distances, even within the surface layer.

TABLE III. Cross-correlation coefficients of wind speed between station pairs, $\rho_{U_A U_B}$.

	KIND	KDSM	KMCI	KAIA	KGUY
KCMI	0.71	0.45	0.39	0.21	0.18
KIND		0.33	0.22	0.18	0.09
KDSM			0.58	0.23	0.30
KMCI				0.22	0.38
KAIA					0.26

TABLE IV. Cross-correlation coefficients of the east–west component of wind velocity between station pairs, $\rho_{u_A u_B}$, and $\rho_{v_A v_B}$ provided in parenthesis.

	KIND	KDSM	KMCI	KAIA	KGUY
KCMI	0.83 (0.78)	0.62 (0.32)	0.43 (0.43)	0.10 (-0.28)	-0.08 (0.01)
KIND		0.52 (0.12)	0.35 (0.24)	0.07 (-0.27)	-0.07 (-0.09)
KDSM			0.65 (0.77)	0.24 (-0.03)	-0.11 (0.46)
KMCI				0.33 (-0.08)	0.07 (0.46)
KAIA					0.25 (0.26)

By examining these two-point correlations in greater detail, particularly their dependence on the lag τ (Equation 4) and identifying the lag at which the correlation reaches its maximum, we gain additional physical insight. Figure 5 shows this analysis of wind-speed correlations between KCMI and KIND, and between KDSM and KMCI. The maximum values of the normalized cross covariance function, $C_{U_A U_B}(\tau)/[C_{U_A U_A}(0) C_{U_B U_B}(0)]^{1/2}$, together with the corresponding lag times τ, are listed in Tables V.

In cases of strong correlation, the lag time at the peak provides a characteristic timescale that, when combined with the spatial distance between the stations, yields a characteristic correlation speed. For KCMI and KIND this speed is approximately 76 miles per hour, and for KDSM and KMCI it is about 341 miles per hour. Since these derived speeds exceed the maximum observed wind velocities, the total observed correlation cannot be explained solely by advection. It is likely that correlations at lower frequencies are associated with advective processes, whereas those at higher frequencies arise from shared local effects such as solar

heating or regional influences, including the presence of Lake Michigan for the KCMI and KIND pair.
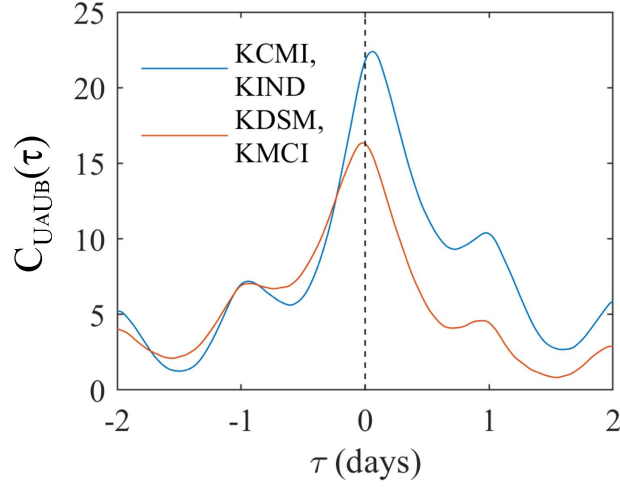

FIG. 5. Cross-covariance functions of wind speed, $C_{U_A U_B}(\tau)$, for the station pairs KCMI–KIND and KDSM–KMCI, where the first station is denoted as A and the second as B. The global maxima occur close to, but not exactly at, $|\tau| = 0$ days.

TABLE V. Maximum values of the normalized cross-covariance function, $C_{U_A U_B}(\tau)/[C_{U_A U_A}(0)C_{U_B U_B}(0)]^{1/2}$, and their corresponding lag τ (in days), where location A is indicated by the row and location B by the column.

	KIND	KDSM	KMCI	KAIA	KGUY
KCMI	0.74 (0.06)	0.46 (-0.05)	0.44 (-0.32)	0.27 (-0.93)	0.30 (-0.94)
KIND		0.36 (-0.95)	0.36 (-0.96)	0.26 (-0.96)	0.25 (-0.98)
KDSM			0.58 (-0.02)	0.31 (-0.94)	0.30 (-0.02)
KMCI				0.22 (0.02)	0.38 (-0.03)
KAIA					0.26 (0.00)

D.   Cumulative Spectral Cross-Correlations

After identifying the station pairs with relatively strong correlations, it is useful to examine how different frequency modes contribute to them. The correlations among pairs of measurement locations are examined and presented as the cumulative spectral cross correlation of wind speed at KCMI relative to each of the other stations in Figure 6a. The
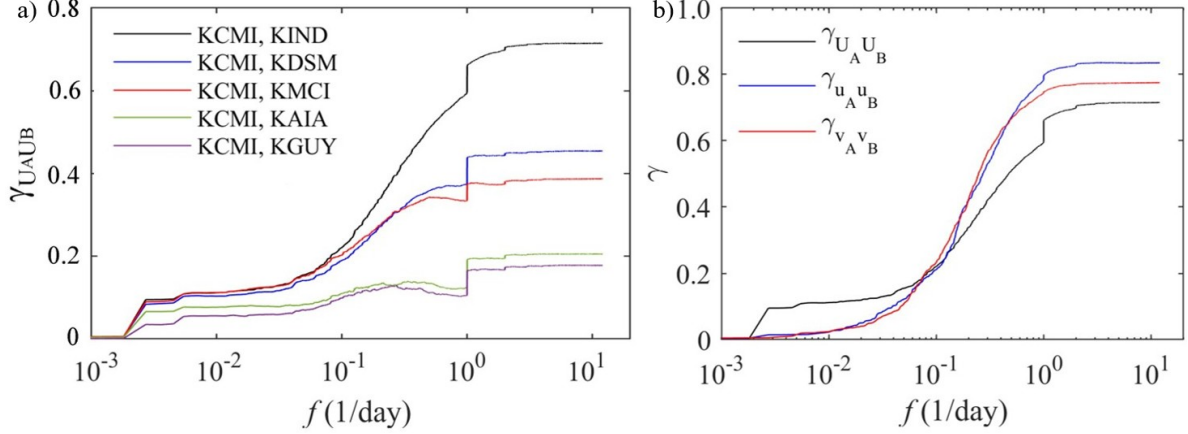
FIG. 6. Cumulative spectral cross correlations of wind speed for a) airports between KCMI and each of KIND, KDSM, KMCI, KAIA, and KGUY, and b) for velocity components between KCMI and KIND.

analysis shows that the lowest- and highest-frequency modes contribute only marginally to the overall correlation. Most of the correlation arises from intermediate frequencies within and near the inertial subrange. This distribution produces an S-shaped trend in the cumulative spectral cross correlations, with a steep slope observed in the intermediate frequency range.

The detailed examination of different velocity components ($\gamma_{U_A U_B}$, $\gamma_{u_A u_B}$, and $\gamma_{v_A v_B}$) is also carried out in Figure 6b for correlation between KCMI and KIND airport. General trend observed in Figure 6a holds, while the mode at $f = 1$ day^{-1} has a pronounced influence on $\gamma_{U_A U_B}$ and a more moderate influence on $\gamma_{u_A u_B}$, while its effect on $\gamma_{v_A v_B}$ is minimal. The mode near $f \approx 1/365$ day^{-1} also contributes noticeably to $\gamma_{U_A U_B}$ but has little impact on $\gamma_{u_A u_B}$ or $\gamma_{v_A v_B}$, reflecting the inclusion of a leap year in the analysis period. These results highlight the distinct roles that different frequency bands play in shaping the cumulative spectral cross correlations between wind components across the two measurement sites.

The contributions from the diurnal frequency modes remain relatively consistent across all pairs, while the correlations associated with the lower frequency modes within the inertial subrange decrease progressively with increasing distance between the locations. This decline eventually leads to near-zero correlation, with the overall coherence between sites weakening. The influence of the annual frequency scales also weakens with increasing spatial separation, though this reduction occurs more gradually than the attenuation observed for

inertial subrange frequencies, specifically from approximately $0.7 < f < 1$ day^{-1} for $\gamma_{U_A U_B}$, and from $0.5 < f < 0.7$ day^{-1} for $\gamma_{u_A u_B}$ and $\gamma_{v_A v_B}$.

To ensure reliable interpretation of these results, it is important to note that the NUDFT approximates the continuous Fourier transform when applied to nonuniformly sampled data. In the present analysis, the time series are sufficiently dense and approximately uniform for the NUDFT assumptions to hold (approximately 80% of the data points are separated by 5 minutes, 16 data points out of more than 1.9 million have a separation of 5 to 19 hours from the preceding data point with 19 hours being the longest separation), and previous studies have shown that the method reproduces spectral amplitudes and phases with high accuracy for band-limited signals[17]. Therefore, the cumulative spectral cross correlations in Figure 6 can be interpreted with confidence. Although minor numerical errors may occur at the highest frequencies due to variations in sampling intervals or short data gaps, these effects are negligible relative to the dominant spectral features and do not affect the overall conclusions.

E. Comparison with Wind Duration Curves

Analyzing correlations in the frequency domain provides insight into the scale dependence of wind variability across spatially separated sites. To examine the role of advection between locations, we construct wind duration curves for each station pair. This procedure involves averaging the probabilities that the component of the wind velocity parallel to the rhumb line connecting the two locations exceeds a given speed. The rhumb line represents the path that intersects all meridians at a constant angle. The resulting duration curve is then expressed in the frequency domain by dividing the wind speed by the distance between the locations, which provides a measure of the advective frequency.

Next, we compute the cumulative mean spectral density, defined as:

$$\zeta_{\alpha\beta}(f_n) = \left[\left(\sum_{p=1}^{N} \tilde{\alpha}_p^2 \right) \left(\sum_{q=1}^{N} \tilde{\beta}_q^2 \right) \right]^{-1/2} \sum_{m=1}^{n} \tilde{\alpha}_m \tilde{\beta}_m, \tag{9}$$

where $\tilde{\alpha}_p$ and $\tilde{\beta}_q$ denote the spectral amplitudes of the wind speeds at the two locations. This formulation represents the maximum attainable cumulative spectral correlation under the assumption of perfect phase alignment between the two signals. Through this approach, the influence of spatial separation on wind velocity correlations can be more clearly interpreted,
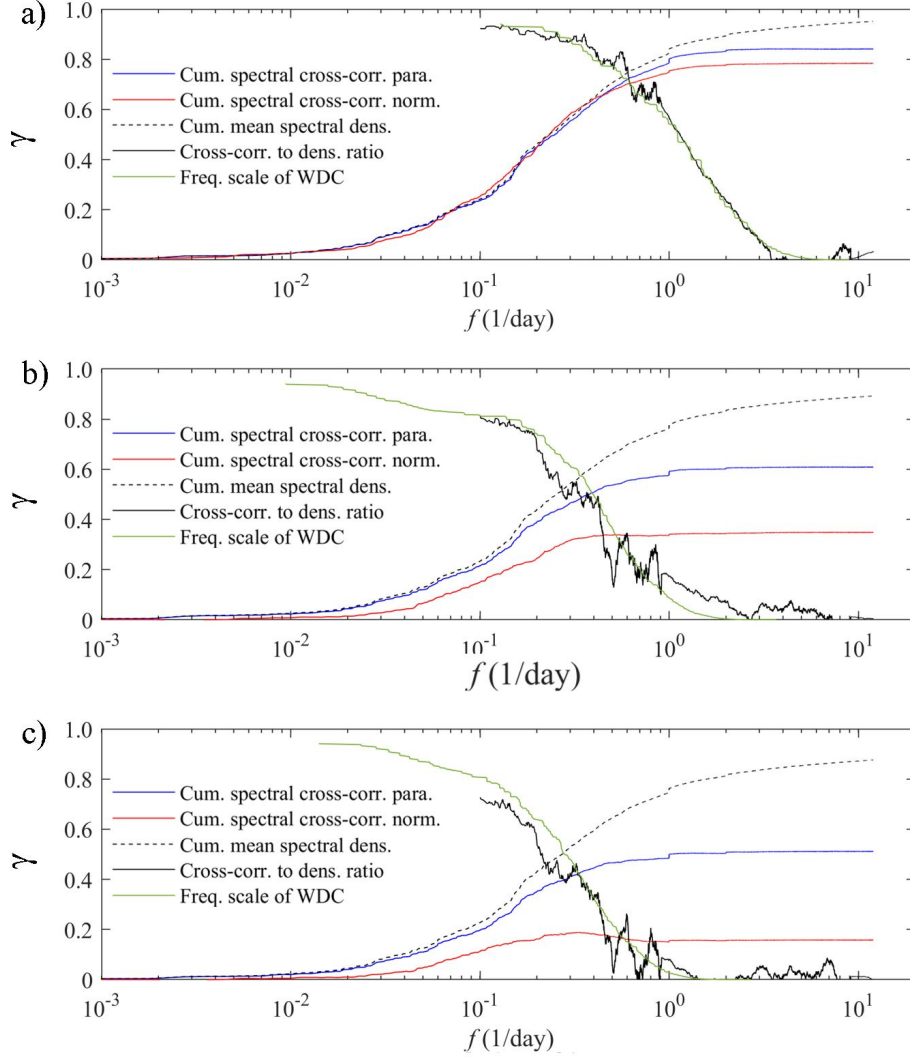
FIG. 7. Cumulative spectral cross correlation of the parallel component, cumulative spectral cross correlation of the normal component, cumulative mean spectral density of the parallel component, and the difference ratio between the cumulative spectral cross correlation and cumulative mean spectral density (i.e., the cosine of the phase difference at each frequency mode), shown with light smoothing for clarity. The wind duration curve (WDC) of the parallel component, transformed to a frequency scale, is also included. Results are shown for: (a) KCMI–KIND, (b) KCMI–KDSM, and (c) KIND–KDSM.

enabling the distinction between components of correlation arising from advective processes and those reflecting phase-coherent variability across sites.

For three selected pairs of locations, we analyze the cumulative spectral cross correlation of the wind speed components parallel and normal to the rhumb line, together with the

corresponding cumulative mean spectral density of these components, as shown in Figure 7. The results reveal a strong similarity between the cumulative spectral cross correlation and the cumulative mean spectral density when the wind duration curve exceeds a value of 0.7. This similarity indicates that the associated frequency modes are nearly in phase, suggesting that part of the observed correlation arises from advective processes.

Furthermore, Figure 7 shows that the cosine of the phase difference for each frequency mode, expressed as $\cos(\theta_{\alpha,n} - \theta_{\beta,n})$, follows closely the values defined by the wind duration curve at the corresponding frequencies. This relationship provides compelling evidence that the phase alignment of these modes is closely linked to the observed wind speeds, supporting the interpretation that advection contributes substantially to the correlation of frequency modes across spatially separated locations.

IV.   DISCUSSION

Frequency–domain analysis provides a scale–aware way to extract structure from irregularly sampled wind records without imposing time–domain interpolation or downsampling, both of which can bias variance, phase, and two–point statistics. Using the NUDFT on a common frequency grid preserves amplitude–phase information needed for cross-spectra, cross-covariance, and coherence across sites. NUDFT accuracy is governed by sampling density, the distribution of time stamps, and band–limitation. Extended gaps increase leakage and reduce effective resolution, and irregular sampling can introduce small spurious high–frequency energy. These effects are controlled here by long records and a conservative high–frequency cutoff; consequently, the dominant features are robust.

A central diagnostic is the cumulative spectral cross correlation, which quantifies how much of the two–point correlation is contributed by frequencies up to a given value. In our results, intermediate frequencies in and near the inertial subrange account for most intersite correlation, while very low and very high frequencies contribute little. Distinct bands emerge with clear physical meaning: the diurnal peak ($f = 1$ day^{-1}) strongly affects the magnitude correlation, the annual band ($f \approx 1/365$ day^{-1}) is weaker but detectable, and component–wise responses differ.

The scale–resolved view clarifies spatial organization. Contributions from inertial–subrange frequencies decay rapidly with distance, annual–scale contributions decay more slowly, and

18

diurnal contributions remain comparatively persistent over the separations considered. This pattern is consistent with finite–extent eddies, regional seasonal forcing, and local daily cycles driven by solar radiation, respectively. Comparing cumulative cross-correlation with a cumulative mean spectral density derived from wind–duration curves further distinguishes advection–driven coherence from phase alignment driven by shared external forcing.

The insights have practical implications. A single bulk correlation masks which processes set intersite coherence; models for forecasting, balancing, or risk assessment can target the bands that dominate shared variability rather than the full spectrum. Correlation length scales are frequency dependent, so spatial diversification in wind fleets will smooth power differently at diurnal versus inertial–subrange scales. The approach generalizes to other environmental records with irregular sampling (e.g., buoy networks, urban air quality, ocean currents). Future work may add uncertainty quantification for cumulative measures, explicit gap handling to bound leakage, and direction–conditional or stability–conditioned analyses to attribute scale–specific coherence to physical drivers.

V. CONCLUSIONS

Environmental and geophysical observations are often characterized by nonuniform time sampling and inconsistent sampling practices across locations. Such irregularity complicates direct comparison, especially for two-point statistics such as covariance and correlation, which typically assume common time bases and uniform sampling intervals. The study presents a frequency-domain framework that addresses these challenges without interpolation or downsampling. By applying the nonuniform discrete Fourier transform, we obtain cross covariance, cross spectra, and power spectra on a common frequency grid, estimating the amplitude and phase of each frequency mode with sufficient fidelity to compare signals across locations even when the underlying time series are irregular.

A central diagnostic introduced here is the cumulative spectral cross correlation, which quantifies how different frequency bands contribute to intersite correlation. Applied to multi-year wind velocity records from airport stations, this diagnostic reveals that correlation is not uniform across scales: inertial-subrange frequencies dominate shared variability at moderate ranges, diurnal variability leaves a strong and spatially persistent signature, and annual variability decorrelates more slowly with distance. It further shows that some correlations

19

arise from coherent advection, while others reflect externally forced phase alignment (e.g., daily solar heating) rather than the transport of a common air mass.

These results demonstrate that a single bulk correlation coefficient conceals distinct physical mechanisms and spatial behaviors. In contrast, the frequency-domain approach resolves which temporal scales control the coherence between locations and how those scales decay with distance. This level of resolution is directly relevant for applications such as wind resource assessment, forecasting, load balancing across spatially distributed assets, and characterization of atmospheric variability.

Although demonstrated here for near-surface winds, the methodology is general. It can be applied to other observational networks in which sampling is irregular, records are long, and spatial coherence matters, including buoy arrays, air quality monitoring networks, and distributed structural or environmental sensing. More broadly, the combination of NUDFT-based spectral reconstruction with cumulative spectral cross correlation offers a physically interpretable path for separating advective coupling from shared external forcing in any system where strict stationarity and uniform sampling cannot be assumed.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]L. P. Chamorro and F. Porté-Agel, "A wind-tunnel investigation of wind-turbine wakes: Boundary-layer turbulence effects," Boundary-Layer Meteorology **132**, 129–149 (2009).

[2]L. P. Chamorro and F. Porté-Agel, "Turbulent flow inside and above a wind farm: A wind-tunnel study," Energies **4**, 1916–1936 (2011).

[3]W. K. George, *Lectures in Turbulence for the 21st Century* (Chalmers University of Technology, 2013).

[4]C. M. S. Martin, J. K. Lundquist, and M. A. Handschy, "Variability of interconnected wind plants: correlation length and its dependence on variability time scale," Environmental Research Letters **10**, 044004 (2015).

[5]C. M. S. Martin, J. K. Lundquist, and M. A. Handschy, en"Variability of interconnected wind plants: correlation length and its dependence on variability time scale," Environmental Research Letters **10**, 044004 (2015), publisher: IOP Publishing.

[6]W. J. Baars, N. Hutchins, and I. Marusic, "Self-similarity of wall-attached turbulence in boundary layers," Journal of Fluid Mechanics **823**, R2 (2017).

[7]T. Nanahara, M. Asari, T. Sato, K. Yamaguchi, M. Shibata, and T. Maejima, "Smoothing effects of distributed wind turbines. part 1. coherence and smoothing effects at a wind farm," **7**, 61–74 (), _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/we.109.

[8]T. Nanahara, M. Asari, T. Maejima, T. Sato, K. Yamaguchi, and M. Shibata, "Smoothing effects of distributed wind turbines. part 2. coherence among power output of distant wind turbines," **7**, 75–85 (), _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/we.108.

[9]D. L. Donoho, "Compressed sensing," IEEE Transactions on information theory **52**, 1289–1306 (2006).

[10]E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Transactions on information theory **52**, 489–509 (2006).

[11]S. L. Brunton and J. N. Kutz, *Data-driven science and engineering: Machine learning, dynamical systems, and control* (Cambridge University Press, 2022).

[12]J. Laskar, "Introduction to frequency map analysis," in *Hamiltonian Systems with Three or More Degrees of Freedom, NATO Adv. Sci Inst. Series C: Math. Phys. Sci.*, Vol. 533, edited by C. Simó (Kluwer, 1999) pp. 134–150.

[13]G. Gómez, J. Mondelo, and C. Simó, "A collocation method for the numerical fourier analysis of quasi-periodic functions. i: Numerical tests and examples," Discrete and Continuous Dynamical Systems Series B **14**, 41–74 (2010).

[14]A. Mitra and D. Kundu, "Consistent method for estimating sinusoidal frequencies: a non-iterative approach," Journal of Statistical Computation and Simulation **58**, 171–194 (1997).

[15]B. G. Quinn and E. J. Hannan, "The estimation and tracking of frequency," (Cambridge University Press, 2001) Chap. 4.

[16]A. Luque and J. Villanueva, "Quasi-periodic frequency analysis using averaging-extrapolation methods," SIAM Journal on Applied Dynamical Systems **13**, 1–46 (2014).

[17] A. Dutt and V. Rokhlin, "Fast fourier transforms for nonequispaced data," SIAM Journal on Scientific computing **14**, 1368–1393 (1993).

[18] S. F. Potter, N. A. Gumerov, and R. Duraiswami, "Fast interpolation of bandlimited functions," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, 2017) pp. 4516–4520.

[19] R. W. Schafer and A. V. Oppenheim, *Discrete-time signal processing* (Prentice Hall, 2010).

[20] A. Dutt and V. Rokhlin, "Fast fourier transforms for nonequispaced data," SIAM Journal on Scientific computing **14**, 1368–1393 (1993).

[21] S. F. Potter, N. A. Gumerov, and R. Duraiswami, "Fast interpolation of bandlimited functions," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, 2017) pp. 4516–4520.

[22] N. O. US Dept of Commerce, A. Administration, and N. W. Service, "National weather service western region headquarters,".

[23] N. O. US Dept of Commerce, A. Administration, and N. H. Center, "Latitude/longitude distance calculator,".

[24] T. Nitta, "On the Role of Transient Eddies in the Tropical Troposphere," Journal of the Meteorological Society of Japan. Ser. II **48**, 348–359 (1970).

[25] F. A. Gifford, S. Barr, R. C. Malone, and E. J. mroz, "Tropospheric relative diffusion to hemispheric scales," Atmospheric Environment (1967) **22**, 1871–1879 (1988).

[26] M. Huber, J. C. McWilliams, and M. Ghil, EN"A Climatology of Turbulent Dispersion in the Troposphere," Journal of the Atmospheric Sciences **58**, 2377–2394 (2001), publisher: American Meteorological Society Section: Journal of the Atmospheric Sciences.