https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

Systematic deviation in wavelet-based covariance estimation and implications for eddy-covariance flux calculations

Jonathan Bitton[1], Bernard Heinesch[1], Samuel Nicolay[2], Catherine Charles[1]

[1]Biosystems Dynamics and Exchanges, TERRA Teaching and Research Center, Gembloux Agro-Bio Tech, University of Liege, Gembloux, Belgium
[2]Departement of Mathematics, University of Liege, Liege, Belgium

Correspondence to: Jonathan Bitton (jonathan.bitton@uliege.be)

Abstract. Accurate estimation of variance and covariance is central to the analysis of geophysical time series and underpins key diagnostics such as turbulent fluxes in eddy-covariance (EC) measurements. Continuous wavelet transform (CWT) methods are widely used for this purpose due to their ability to represent non-stationary and multi-scale processes. However, the consistency of commonly used formulations with the theoretical properties of the CWT has not been fully clarified. In this work, we show that the widely used formulation of Torrence and Compo can lead to systematic deviations in covariance estimation, which can be traced to normalization choices that are not fully consistent with the CWT reconstruction framework. These deviations, of the order of 10–15 % under typical conditions, are observed across a range of signal types, including deterministic, white-noise, and long-range dependent processes. We derive an alternative formulation directly from the CWT reconstruction identity, ensuring consistency with wavelet energy conservation. This approach reduces bias and improves convergence properties. Application to EC data shows that the proposed formulation recovers classical covariance estimates under stationary conditions while providing a more consistent framework for the analysis of non-stationary signals. These results emphasize the importance of scale-consistent estimation of second-order moments for interpreting geophysical variability and turbulent fluxes, and provide a methodological basis for multi-scale analysis of environmental time series.

1 Introduction

Quantifying variance and covariance is a fundamental task in the analysis of geophysical and environmental time series (Kirlin et al., 1999; Whitcher et al., 2000; Witt and Malamud, 2013). Variance provides a measure of the amplitude of fluctuations across scales, while covariance characterizes interactions and co-variability between physical processes. These quantities underpin key diagnostics in Earth system science, including variability assessments, uncertainty quantification, and the identification of scale-dependent couplings between atmospheric, hydrological, and biogeochemical components (Amiri-Simkooei, 2013; Baldocchi et al., 2018; Franzke et al., 2020).

Geophysical time series are, however, rarely well described by simple statistical assumptions. They typically exhibit strong autocorrelation, non-stationarity, and variability spanning multiple temporal scales (Franzke and O'Kane, 2017). Their power spectra are often dominated by low-frequency contributions and frequently follow power-law behavior characteristic

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

of long-range dependent (1/f-type) processes (Fraedrich et al., 2009; Yuan et al., 2022). In such contexts, classical time-domain estimators of variance and covariance become sensitive to finite-sample effects and to the relative contribution of different scales, potentially leading to biased or unstable estimates (Percival and Walden, 2000). These characteristics complicate the retrieval of reliable statistical quantities from observational data.

Wavelet analysis provides a natural framework to address these challenges by offering a joint time–scale representation of signals. Since the work of Torrence and Compo (1998), the continuous wavelet transform (CWT) has been widely used in Earth science, climate science, hydrology, turbulence studies, and land–atmosphere interaction research to characterize non-stationary and scale-dependent processes (Debnath and Shah, 2015; Lovejoy and Schertzer, 2013; Montillet and Bos, 2020). Beyond spectral diagnostics, wavelet coefficients have been used to reconstruct second-order statistical moments, such as variance and covariance, by invoking wavelet analogues of Parseval's identity (Metzger et al., 2013; Schaller et al., 2017; Torrence and Compo, 1998).

Despite this widespread use, the theoretical basis of wavelet-based variance reconstruction remains only partially examined. The classical formulation of Torrence and Compo is generally interpreted as a practical discretization of energy conservation in the wavelet domain, yet it is not derived explicitly from the reconstruction properties of the CWT. As a result, it implicitly relies on assumptions regarding scale discretization, normalization, and statistical independence of wavelet coefficients. The implications of these assumptions for variance and covariance estimation, particularly for long-range dependent processes, remain insufficiently understood.

These issues are particularly relevant for high-frequency geophysical measurements, such as those obtained from eddy-covariance (EC) systems. In EC applications, covariance between turbulent fluctuations is used to derive fluxes of momentum, heat, and trace gases (Aubinet et al., 2012). Because covariance is the fundamental quantity underlying EC flux calculations, any systematic bias in its estimation directly propagates into flux estimates. While EC rely on time-averaged statistics under assumptions of stationarity, wavelet-based approaches have been proposed as alternatives capable of resolving scale-dependent flux contributions and handling non-stationary conditions (Dupont, 2020; Göckede et al., 2019; Rhif et al., 2019). However, the consistency between wavelet-based covariance reconstruction and classical EC estimates remains an open question, especially in regimes where stationarity assumptions are violated.

In this study, we revisit the problem of variance and covariance estimation from CWT coefficients, with the objective of establishing a mathematically consistent and practically robust framework for second-order moment reconstruction. Specifically, we aim to (i) derive variance and covariance estimators directly from the continuous reconstruction properties of the CWT, (ii) quantify the impact of discretization, wavelet choice, and spectral structure on estimation accuracy, and (iii) assess the practical implications of these differences for geophysical applications.

To this end, we first develop a formulation of second-order moments explicitly grounded in the CWT inner-product identity, ensuring theoretical consistency with wavelet energy conservation. We then compare this framework to the widely used Torrence and Compo formulation using synthetic signals spanning a range of spectral structures, including long-range dependent processes. Finally, we illustrate the relevance of these developments using eddy-covariance data, with a focus on

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

the interpretation of covariance under stationary and non-stationary conditions and on the extent to which wavelet-based estimates can reproduce or extend classical EC diagnostics.

2 Methods

2.1 Variance and covariance reconstruction from wavelet coefficients

2.1.1 Classical formulation

Let x and y be two zero-mean signals of length N, obtained by uniformly sampling two continuous functions $f(t)$ and $g(t)$ with a temporal resolution δt. Variance and covariance can be estimated from their CWT coefficients $W_x\left(\frac{a_j}{\delta t}, n\right)$ and $W_y\left(\frac{a_j}{\delta t}, n\right)$ using (Metzger et al., 2013; Torrence and Compo, 1998):

$$\sigma_x^2 = \frac{\delta t}{C_\delta}\frac{\delta j}{N}\sum_{n=1}^{N}\sum_{j=1}^{J}\frac{\left|W_x\left(\frac{a_j}{\delta t}, n\right)\right|^2}{a_j}, \qquad cov(x,y) = \frac{\delta t}{C_\delta}\frac{\delta j}{N}\sum_{n=1}^{N}\sum_{j=1}^{J}\frac{Re\left[W_x\left(\frac{a_j}{\delta t}, n\right)\overline{W_y\left(\frac{a_j}{\delta t}, n\right)}\right]}{a_j},$$

where $\overline{}$ denotes the complex conjugate, C_δ is a wavelet-dependent reconstruction factor, a_j are the set of analyzed scales and δj is the logarithmic scale resolution. The translation parameter is expressed as a discrete time index n and scales are written in dimensionless form $\frac{a_j}{\delta t}$.

In Metzger et al. (2013) and subsequent studies (e.g. Foken and Harrison, 2026; Paleri et al., 2022; Schaller et al., 2017), the covariance is often expressed using the complex cross-product of wavelet coefficients without explicitly taking its real part. However, for real-valued signals, only the real component contributes to covariance. This follows directly from the Fourier-domain equivalence, where covariance is obtained by integrating the cospectrum, defined as the real part of the cross-spectrum, over all frequencies (Stull, 1988). For consistency with this spectral interpretation and to ensure that covariance remains a real-valued quantity, the real part is explicitly retained here.

These expressions, hereafter referred to as the TC formulation, are commonly used as a discrete wavelet analogue of Parseval's theorem. Since they are not derived explicitly from the CWT reconstruction framework, they can be interpreted as an empirical discretization, which may introduce deviations in variance and covariance estimation.

2.1.2 Formulation based on CWT theory

To address this limitation, we derive an alternative formulation that ensures consistency with wavelet energy conservation. This formulation originates from the CWT reconstruction identity (Daubechies, 1992) discretized for practical use in two cases: real-valued wavelets and analytic wavelets applied to real-valued signals. Full derivations are provided in Appendices A-B. Here, we present the final discretized expressions.

Both formulations rely on a modified reconstruction factor, defined as

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

$$C_\psi = \int_0^{+\infty} \frac{\left|\hat{\psi}(\omega)\right|^2}{\omega} d\omega < \infty,$$

where $\hat{\psi}(\omega)$ is the Fourier transform of the mother wavelet $\psi(t)$.

i. Real-valued wavelets

95 For real-valued wavelets, the covariance estimator becomes

$$cov(x,y) = \frac{\langle x,y \rangle}{N} = \frac{\delta t}{N}\frac{\delta j}{C_\psi}\ln(2)\sum_{n=1}^{N}\sum_{j=1}^{J}\frac{W_x\left(\frac{a_j}{\delta t},n\right)\overline{W_y\left(\frac{a_j}{\delta t},n\right)}}{a_j},$$

Since the wavelet is real-valued, the complex product is real for real-valued signals. The corresponding variance estimator, obtained by substituting the second function y with x, is

$$\sigma_x^2 = \frac{\langle x,x \rangle}{N} = \frac{\delta t}{N}\frac{\delta j}{C_\psi}\ln(2)\sum_{n=1}^{N}\sum_{j=1}^{J}\frac{\left|W_x\left(\frac{a_j}{\delta t},n\right)\right|^2}{a_j}.$$

100 **ii. Analytic wavelets and real-valued functions**

For analytic wavelets applied to real-valued signals, the estimators become:

$$cov(x,y) = 2\frac{\delta t}{N}\frac{\delta j}{C_\psi}\ln(2)\sum_{n=1}^{N}\sum_{j=1}^{J}\frac{Re\left[W_x\left(\frac{a_j}{\delta t},n\right)\overline{W_y\left(\frac{a_j}{\delta t},n\right)}\right]}{a_j},$$

$$\sigma_x^2 = 2\frac{\delta t}{N}\frac{\delta j}{C_\psi}\ln(2)\sum_{n=1}^{N}\sum_{j=1}^{J}\frac{\left|W_x\left(\frac{a_j}{\delta t},n\right)\right|^2}{a_j}.$$

The factor of 2 accounts for the exclusion of negative frequencies and ensures energy conservation.

105 **2.1.3 Alternative formulations**

The above expressions assume energy conservation under the standard L^2 normalization of the CWT. An alternative normalization, based on L^1 norm preservation, ensures amplitude invariance across scales and provides an unambiguous scale–frequency correspondence (Yi and Shu, 2012). These are particularly useful for visualization, since variance and covariance values are directly comparable across scales. Corresponding expressions are provided in Supplement S1.

110 Additionally, the discretization assumes logarithmic sampling of scales with base 2. Linear scale sampling may also be used to refine low frequency resolution. Corresponding formulations are provided in Supplement S2.

2.1.4 Practical considerations for CWT parametrization

Accurate variance reconstruction requires an appropriate choice of wavelet transform parameters, in particular the range and sampling of scales as well as the treatment of boundary effects.

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

Scales should be selected so that the corresponding temporal support spans the entire range of relevant signal periods, from the sampling interval δt to the total signal duration $N\delta t$. In practice, scales are usually sampled logarithmically, which ensures uniform coverage in log-frequency space and preserves the structure of the reconstruction integral. The scale resolution δj must be sufficiently small to avoid under-sampling; values $\delta j \leq 0.5$ are typically recommended (Torrence and Compo, 1998). Here, we use $\delta j = 0.05$.

Because wavelets extend beyond signal boundaries, edge effects arise for finite signals. While padding is often used to mitigate these effects, it introduces artificial information that may bias reconstruction. For variance reconstruction, analyses should be restricted to regions unaffected by boundary effects, as defined by the cone of influence. When available, extending the signal using adjacent data segments provides a more robust alternative.

2.2 Validation using synthetic signals

To assess the performance of the proposed variance reconstruction framework and to compare it with the classical TC formulation, we conducted a validation study based on synthetic signals with known statistical properties. The use of controlled signals allows direct comparison between the estimated variance and its analytical value independently of measurement noise or sampling artefacts and facilitates the identification of potential biases related to signal length, spectral content and wavelet choice.

2.2.1 Synthetic signals

Three types of zero-mean synthetic signals were considered:

1. Pure sinusoidal signal, defined as $x(t) = Asin(2\pi f_0 t)$, for which the theoretical variance is $\sigma_x^2 = \frac{A^2}{2}$. This signal concentrates all its energy at a single frequency and serves as a reference case to assess reconstruction accuracy for narrow-band signals.
2. White noise, generated as a Gaussian random process with variance σ^2. Its flat power spectrum (i.e. variance is uniformly distributed across frequencies) makes it well suited to evaluate scale-dependent biases and the robustness of the variance estimator under broadband conditions.
3. Pink noise, obtained by spectral shaping of Gaussian noise to impose a power spectral density inversely proportional to frequency. This signal emphasizes low-frequency variability and provides a stringent test for variance reconstruction, especially with respect to scale integration limits and finite-sample effects.

These three signals span a wide range of spectral structures, from fully deterministic to highly stochastic, and from narrowband to strongly scale-dependent processes.

2.2.2 Experimental design

For each signal type, ensembles of realizations were generated by varying both the signal length N from 2^{10} to 2^{19} and key constitutive parameters:

- amplitude A and frequency f_0 for the sinusoidal signal,

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

- variance σ^2 for white and pink noises.

Each parameter was sampled using ten distinct values, resulting in a pseudo-ensemble of realizations for each signal type. This design enables a systematic exploration of the sensitivity of variance estimates to signal length and signal parameters.

For each realization, variance was estimated using the TC (Sect. 2.1.1) and the CWT-based (Sect. 2.1.2) formulations for two common mother wavelets:

- the real-valued second derivative of the Gaussian (DOG2, also called Mexican Hat), for which $C_\psi \approx 2.1349$.
- the approximately analytic Morlet wavelet, with central angular frequency $\omega_0 = 6$, for which $C_\psi \approx 1.1988$.

2.2.3 Evaluation metrics and statistical representation

For each realization, the estimated variance $\hat{\sigma}^2$ was compared to the corresponding analytical variance σ^2 using the absolute relative error:

$$\varepsilon = \left|\frac{\hat{\sigma}^2}{\sigma^2} - 1\right|.$$

Distributions of ε obtained across all realizations of a given signal type were summarized using boxplots. This representation allows the bias (deviation of the median from unity), dispersion, and asymmetry of the variance estimates to be visually assessed, while explicitly accounting for variability induced by signal length and parameter modulation.

For conciseness, the benchmarking focuses on variance reconstruction. However, all tests are directly extensible to covariance estimation in a rigorous manner, since formulations derived in Sect. 2.1 share the same mathematical structure for both expressions. This extension was verified in practice but is not shown here to avoid redundancy.

2.3 Application to eddy-covariance

Wavelet-based covariance reconstruction was illustrated using eddy-covariance (EC) measurements from the ICOS Lonzée site (BE-Lon) for the year 2024.

Half-hour periods were selected among those exhibiting the largest absolute vertical CO_2 flux values. Stationary and non-stationary conditions were identified using the combined steady-state and integral turbulence characteristics test (SSITC), with SSITC = 0 indicating stationary conditions and SSITC = 2 indicating non-stationary conditions (Foken et al., 2004).

High-frequency EC time series of vertical wind speed and CO_2 concentration were pre-processed using standard despiking procedures and corrected for lag-time following conventional EC practice. To minimize artificial edge effects in the wavelet domain, each half-hour segment was extended using adjacent data on both sides to ensure all analyzed scales (up to an equivalent period of 30 min) were free of boundary effects. Due to its broader temporal support, Morlet required longer extensions (approx. 5 times that of DOG2), reducing the number of usable non-stationary periods.

After quality filtering, 49 stationary half-hours were retained for both wavelets, while 44 and 42 non-stationary half-hours were kept for DOG2 and Morlet analyses, respectively. Covariance estimates were computed in the wavelet domain using

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

both the TC formulation and the proposed CWT-based reconstruction. Due to the illustrative purpose of this analysis, covariance values were not converted into flux units.
All analyses were performed using Matlab R2022a (The MathWorks Inc, 2022).

3 Results

3.1 Variance estimation on synthetic signals

Across all signal types and wavelet configurations, the CWT-based estimator consistently yields smaller variance errors than the classical TC formulation (Fig. 1). This improvement reflects a reduction of systematic bias and enhanced robustness of variance estimation across signal configurations. In contrast, the TC formulation systematically overestimates or underestimates variance by approximately 10–15 % depending on the wavelet and signal characteristics. The sign of the bias reflects the combined effect of the wavelet used and the sign of the estimated statistic: for a given wavelet, the TC formulation consistently over- or underestimates variance (Fig. S1), whereas for covariance the bias sign additionally depends on the sign of the statistic.

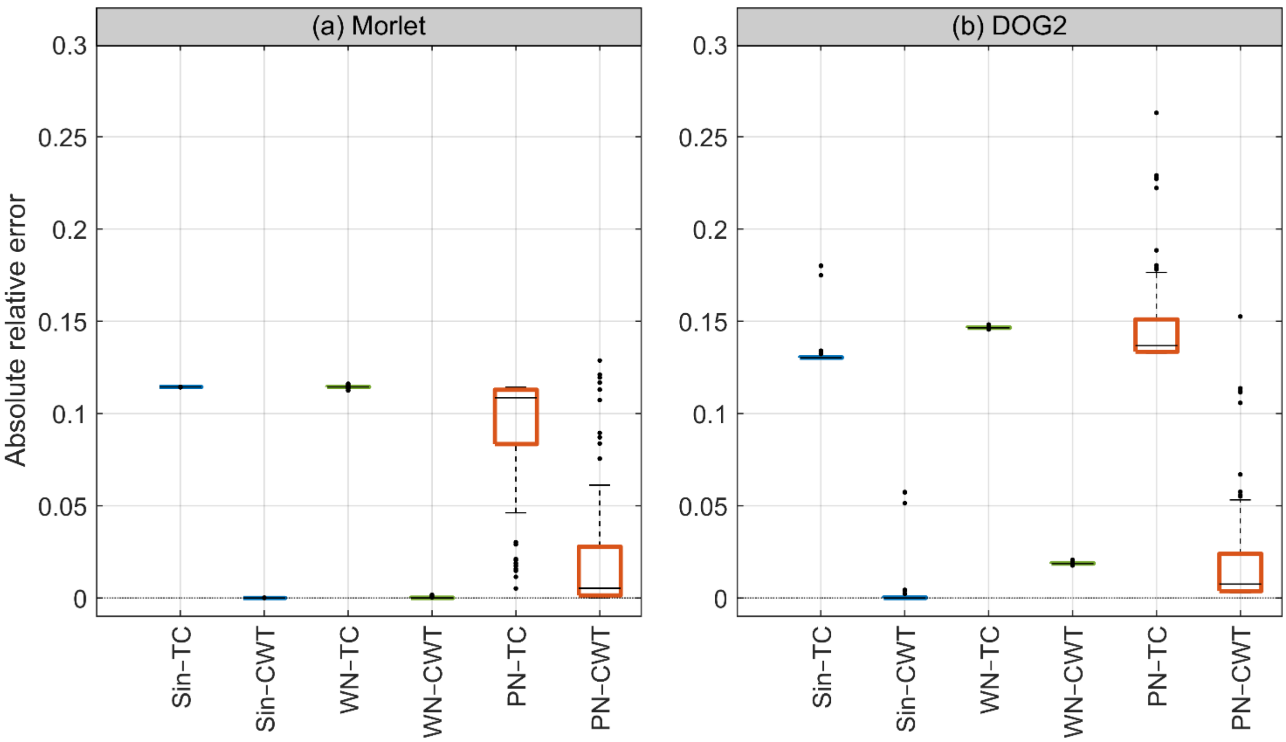


Figure 1: Absolute relative error of variance estimates obtained using the TC formulation and the proposed CWT-based estimator for sinusoidal (Sin), white-noise (WN) and pink-noise (PN) signals with (a) Morlet and (b) DOG2 wavelets. Boxplots summarize variability across signal lengths and parameter values. The horizontal line at zero absolute relative error denotes perfect variance reconstruction.

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

For stochastic signals, the contrast between the two approaches remains evident. While pink-noise signals exhibit substantial dispersion under both formulations, the TC-based approach yields systematically larger biases that increase with signal length when using the Morlet wavelet (Fig. S1). This behavior points to a lack of asymptotic consistency, with the estimator diverging from the true variance as more data is included. By contrast, the CWT-based estimator markedly improves accuracy, suggesting that the proposed formulation more effectively captures the scale-dependent contributions of broadband variability and long-range dependence to the overall signal variance.

Differences between wavelet types are modest but non-negligible. The divergence between TC and CWT is more pronounced for the DOG2 wavelet, likely reflecting differences in normalization factors between the two formulations. For the CWT-based estimator, both wavelets yield similar results in most cases, except for white-noise signals. In this case, the real-valued DOG2 wavelet exhibits a small residual bias, whereas the analytic Morlet wavelet produces nearly unbiased estimates. This behavior suggests a wavelet-dependent bias arising from differences in frequency response of the underlying wavelet filters.

3.2 Convergence properties of the CWT-based estimator

Variance estimates converge asymptotically toward the true value as signal length increases, demonstrating the consistency of the CWT-based estimator (Fig. 2). Convergence is almost immediate for the sinusoidal signal, where accurate variance reconstruction is achieved even for short sample lengths.

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

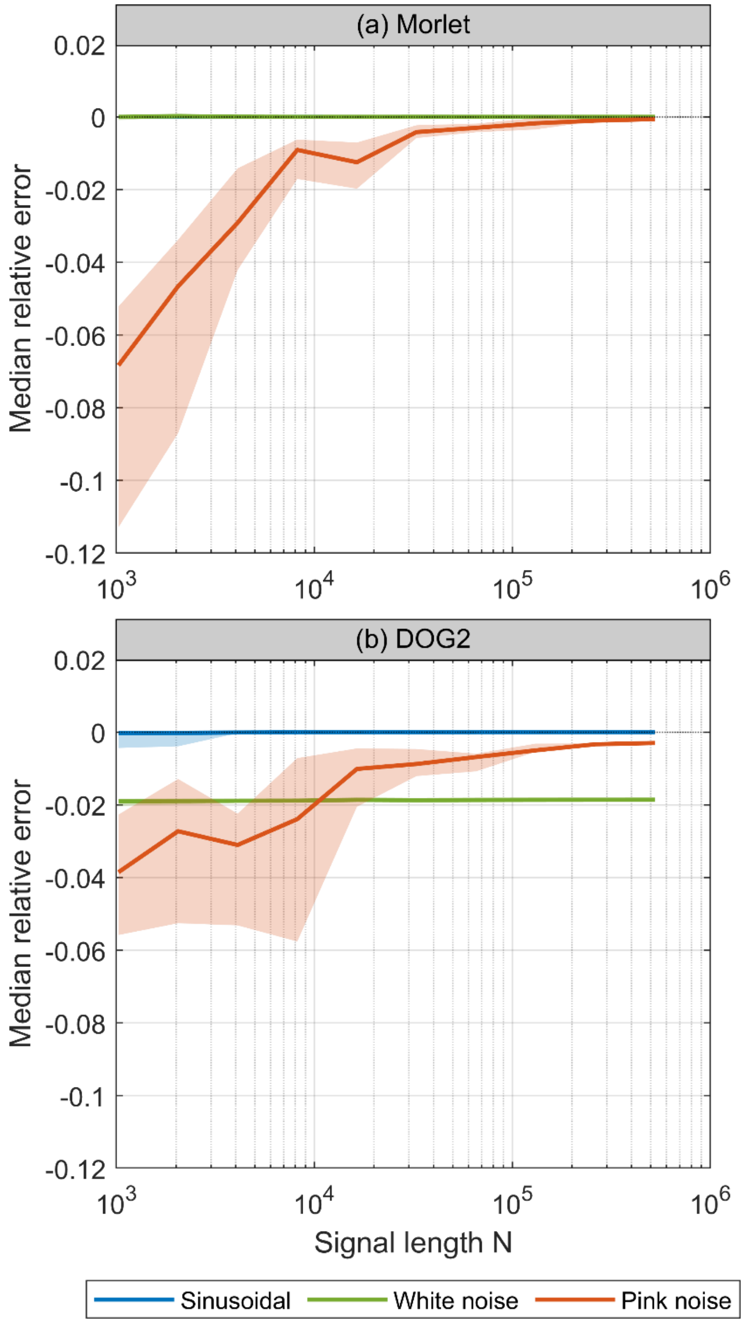


Figure 2: Convergence of the CWT-based variance estimator as a function of signal length for sinusoidal (blue), white-noise (green), and pink-noise (red) signals using (a) Morlet and (b) DOG2 wavelets. Solid lines denote median relative errors ($\hat{\sigma}^2/\sigma^2 - 1$) and shaded areas indicate interquartile ranges

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

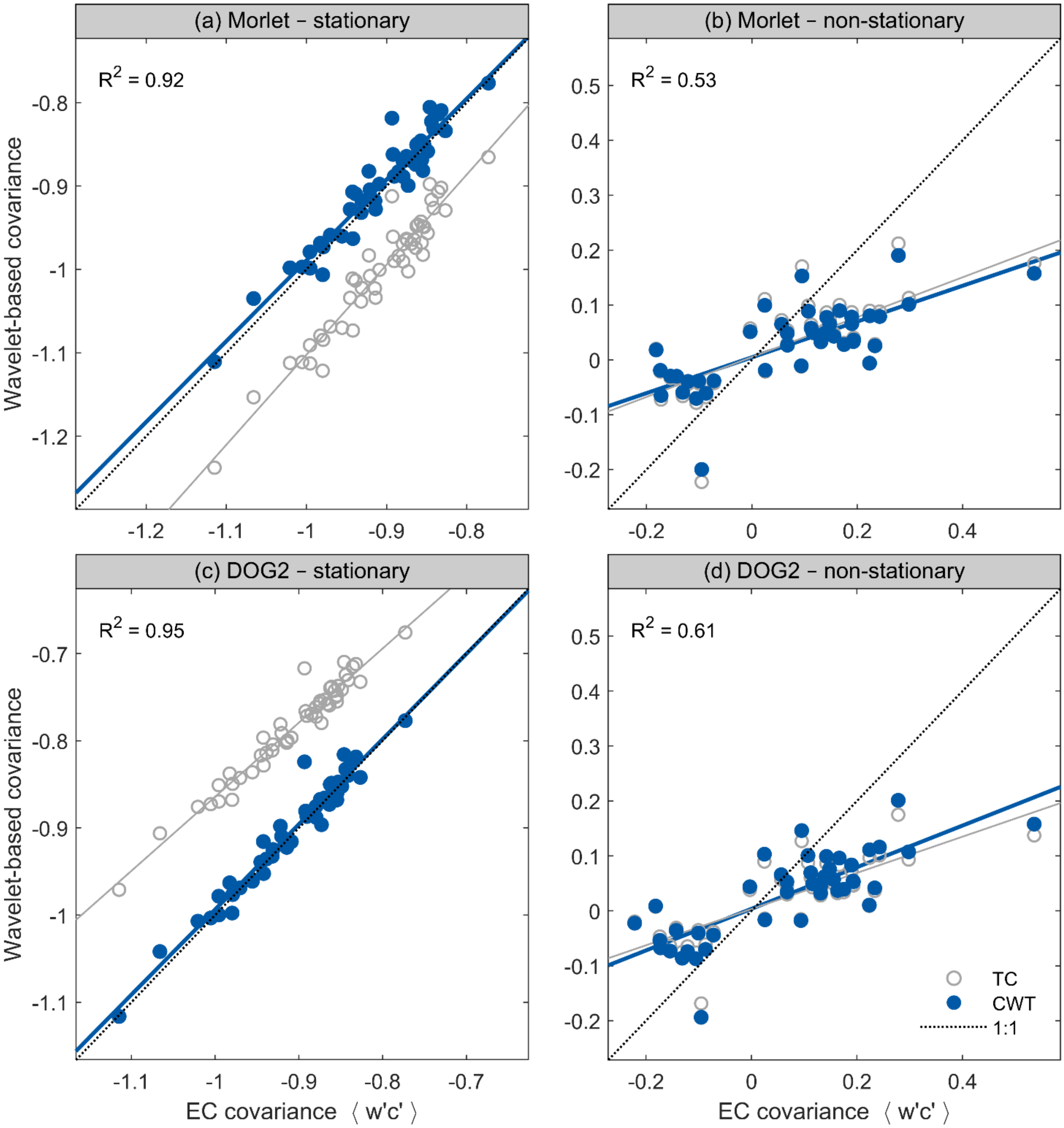
White-noise signals also exhibit rapid convergence. However, a small residual bias persists for the DOG2 wavelet, consistent with the filter-dependent effects identified in Fig. 1. In contrast, convergence using the Morlet wavelet is both faster and less dispersed, reflecting its improved spectral selectivity.

Pink-noise signals exhibit the strongest sensitivity to signal length, with large finite-sample variability and slower convergence. The Morlet wavelet yields faster convergence and smaller dispersion than the DOG2 wavelet. For both cases, substantially longer time series are required for accurate variance reconstruction, which is dominated by low-frequency components. Consequently, estimator dispersion prevails over residual filter-dependent biases in this regime.

3.3 Illustrative application to eddy-covariance data

Figure 3 compares classical EC estimates of the covariance between vertical wind speed and CO_2 concentration with covariance reconstructed from wavelet coefficients using the TC formulation and the proposed CWT-based approach.

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026



225

Figure 3: Comparison between classical EC covariance and wavelet-reconstructed covariance between vertical wind speed and CO_2 concentration under stationary (a–c) and non-stationary (b–d) conditions. Results are shown for Morlet (a,b) and DOG2 (c,d) wavelets using TC (gray circles) and CWT-based (blue circles) formulations. Each point corresponds to a half-hour period. Solid lines indicate linear regressions, and the dotted line indicates the 1:1 relationship. The coefficient of determination R^2 is identical
230 **for the CWT- and TC-based relationship with EC.**

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

Under stationary conditions (Fig. 3a,c), covariance reconstructed with the CWT-based formulation closely matches EC estimates for both wavelet types. The relationship is nearly linear, with coefficients of determination exceeding 0.9, which indicates that the proposed approach preserves classical EC covariance when stationarity assumptions are satisfied. TC covariance exhibits larger deviations from the one-to-one relationship, consistent with the finite-sample biases identified in synthetic experiments. The proposed formulation therefore reduces discrepancies with classical covariance estimates under stationary conditions. These results demonstrate that the differences between formulations are not limited to theoretical cases but directly affect flux estimates derived from covariance. Interestingly, and in contrast with the faster convergence and reduced bias observed for the Morlet wavelet in synthetic experiments, the real-valued DOG2 wavelet shows a slightly closer agreement with the one-to-one relationship.

Under non-stationary conditions (Fig. 3b,d), the agreement between wavelet-based and EC covariance decreases substantially for both wavelet types, with increased scatter and lower coefficients of determination. In this regime, covariance contributions become intermittent and scale dependent, violating the assumptions underlying classical EC averaging. The wavelet-based reconstruction therefore departs from a single bulk covariance value and highlights the limited representativeness of time-averaged covariance under non-stationary conditions. Differences between the TC and CWT-based formulations become less visible under non-stationary conditions because covariance values are concentrated within a narrower range around zero. As both formulations differ primarily by a multiplicative factor, their estimates naturally converge as covariance approaches zero, leading to reduced apparent discrepancies.

4 Discussion

4.1 Improved performance of the CWT-based formulation

The results obtained on synthetic signals demonstrate that variance and covariance estimation derived explicitly from CWT theory provides a consistent and robust alternative to the classical TC formulation. Across all tested configurations, the CWT-based estimator systematically reduces finite-sample bias while preserving asymptotic consistency, including for long-range dependent processes.

These improvements arise directly from the theoretical construction of the estimator. Unlike the TC formulation, which is introduced as a discrete analogue of Parseval's theorem, the proposed approach is derived from the continuous wavelet inner-product identity. Although both formulations rely on the same wavelet coefficients, differences in normalization lead to systematic deviations in the TC approach when integrating contributions across scales.

The eddy-covariance application confirms that these properties extend to geophysical signals. Under stationary turbulent conditions, the CWT-based reconstruction closely reproduces classical EC estimates. In contrast, the bias observed for the TC formulation persists in the EC analysis, indicating that it originates from the estimation framework itself rather than from data characteristics.

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

4.2 Role of signal spectral structure, long-range dependence, and non-stationarity

Differences between estimators become increasingly pronounced with signal complexity. The most demanding cases are pink-noise signals, representative of long-range dependent geophysical processes. For these signals, variance is dominated by contributions from the largest scales, where only a limited number of statistically independent wavelet coefficients are available in finite-length records (Percival and Walden, 2000). This results in a strong reduction in effective degrees of freedom and increased sensitivity to scale truncation, boundary effects, and integration limits. The slow convergence observed for pink-noise signals is therefore an intrinsic property of the process rather than a limitation of the estimator itself. In this regime, estimator dispersion dominates, but the TC formulation still exhibits persistent bias, whereas the CWT-based estimator remains consistent.

The EC application provides a real-world counterpart to this behavior. Under non-stationary conditions, covariance contributions become intermittent and strongly scale-dependent. In such cases, a single time-averaged EC covariance loses representativeness, and divergence between EC and wavelet-based estimates reflects the breakdown of the underlying stationarity assumptions. The wavelet-based formulation therefore acts not only as an estimator but also as a diagnostic tool, revealing the multi-scale structure of covariance and highlighting the limitations of bulk averaging approaches.

4.3 Influence of wavelet choice

Wavelet choice plays a secondary but non-negligible role in reconstruction accuracy, as magnitude of wavelet-dependent bias remains limited and both wavelet types preserve asymptotic consistency.

The Morlet wavelet yields faster convergence and reduced dispersion across signal types. This behavior can be explained by its improved frequency localization, which leads to a more consistent distribution of energy across scales. In addition, the complex nature of its coefficients provides two statistically independent components, such that their squared amplitude follows a χ^2 distribution with two degrees of freedom. In contrast, real-valued wavelets produce coefficients with a single degree of freedom (Ge, 2007; Torrence and Compo, 1998). This difference contributes to reduced estimator variance for the Morlet wavelet, although its effect remains secondary compared to dominant factors such as spectral structure and sample size.

The EC application highlights a complementary perspective. Under stationary conditions, both wavelet types reproduce EC estimates well using the CWT-based formulation, but the DOG2 wavelet shows slightly stronger agreement with the one-to-one relationship. This behavior, consistent with previous wavelet-based analyses of turbulent signals (Li et al., 2023; Schaller et al., 2017), reflects its stronger temporal localization and its closer conceptual proximity to time-averaged covariance estimates used in EC computations. In contrast, the broader temporal support of the Morlet wavelet increases sensitivity to low-frequency variability and transitional contributions, leading to earlier divergence from a single bulk covariance under non-stationary conditions. This divergence is therefore diagnostic rather than indicative of reduced methodological consistency. These results indicate that wavelet choice should be guided by the analysis objective: localized

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

real-valued wavelets for compatibility with classical estimates, and broader analytic wavelets for improved sensitivity to scale-dependent and non-stationary processes.

A persistent bias was observed for the DOG2 wavelet in the white-noise experiments. This bias can be attributed to its non-ideal frequency response and the resulting imperfect spectral partitioning across scales, consistent with wavelet variance leakage associated with low-order real-valued filters (Percival and Walden, 2000). In wavelet-based variance estimation, unbiasedness requires that the discretized wavelet filters provide an approximately energy-preserving partition of the spectrum, ensuring consistent integration of variance across scales. The DOG2 wavelet, characterized by a broad and strongly overlapping spectral response, deviates from this condition, leading to systematic over- or under-weighting of certain frequency bands. In contrast, the Morlet wavelet provides a more localized band-pass behavior, resulting in more uniform spectral coverage and reduced bias. While filter-induced bias is clearly detectable for white-noise signals, it becomes secondary for pink-noise processes, where estimation uncertainty is dominated by limited degrees of freedom at large scales.

4.4 Implications, limitations and perspectives

The proposed formulation provides a coherent and physically meaningful framework for estimating second-order statistics across scales. This property is particularly relevant for geophysical applications involving long, correlated, and multi-scale time series, where variance and covariance are used to characterize variability, quantify energy transfer, and diagnose coupling between processes (Aubinet et al., 2012; Seo et al., 2023; Stull, 1988).

For such applications, errors of the order of 10–15 %, as observed for the classical TC formulation even in simple cases, are non-negligeable. From a measurement perspective, these results indicate that commonly used wavelet-based methods may introduce intrinsic biases in covariance estimation. These biases are comparable in magnitude to typical uncertainties encountered in practical applications (e.g. EC measurements; Richardson et al., 2012), but are systematic in nature and may therefore lead to persistent errors in derived quantities, long-term analyses and inter-study comparisons. The reduced bias and improved robustness of the CWT-based estimator thus provide a clear methodological advantage.

In the context of eddy-covariance measurements, this work supports the use of wavelet-based covariance reconstruction as a complement to classical flux estimation. While EC methods provide robust estimates under stationarity (Aubinet et al., 2012), wavelet-based approaches offer additional insight into the scale-dependent structure of turbulent transport and enable a more consistent interpretation of intermittent and non-stationary fluxes (Göckede et al., 2019; Paleri et al., 2022). Beyond their diagnostic value, wavelet-based approaches relax the strict requirement of stationarity inherent to classical EC processing. This allows a larger fraction of high-frequency observations to be retained, including periods that would otherwise be discarded due to non-stationarity, thereby reducing reliance on gap-filling procedures and improving the representativeness of flux estimates (Coimbra et al., 2025).

Future work should extend the proposed framework to fully scale-resolved analyses of covariance under non-stationary conditions, including the explicit representation of intermittency and coherent structures. Extending the framework to

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

multivariate settings would further enable the characterization of scale-dependent interactions among multiple variables, providing new insight into coupled Earth system processes. Ultimately, integrating scale-resolved covariance diagnostics into operational flux processing frameworks would improve the robustness and interpretability of high-frequency environmental observations.

5 Conclusions

This study shows that wavelet-based covariance estimates derived from the widely used Torrence and Compo formulation can exhibit systematic deviations linked to normalization choices that are not fully consistent with the CWT reconstruction framework. These deviations are observed across signal types and can propagate to eddy-covariance flux estimates, with magnitudes comparable to typical measurement uncertainties.

By deriving estimators consistent with the CWT reconstruction identity, we provide an alternative formulation that reduces these deviations and improves estimation reliability across signal configurations. These results underline the importance of theoretically consistent formulations in multi-scale analysis and indicate that wavelet-based estimates of geophysical variability and turbulent fluxes may benefit from a careful reassessment within this framework.

Appendix A: Derivation of reconstruction formulas

Let $f(t)$ and $g(t)$ be square-integrable zero-mean functions, and $W_f(a,b)$ and $W_g(a,b)$ denote their wavelet coefficients at scale $a \in \mathbb{R}$ and translation $b \in \mathbb{R}$. Using the standard inner product in $L^2(\mathbb{R})$, $\langle f, g\rangle = \int_{-\infty}^{+\infty} f(t)\overline{g(t)}dt$, the CWT reconstruction identity (Daubechies, 1992) yields:

$$\langle f, g\rangle = \frac{1}{C_\psi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} W_f(a,b)\overline{W_g(a,b)}\,db\,\frac{da}{a^2},$$

where $C_\psi = \int_{-\infty}^{+\infty}\frac{|\hat{\psi}(\omega)|^2}{|\omega|}d\omega < \infty$ is the wavelet admissibility constant and $\hat{\psi}(\omega)$ denotes the Fourier transform of the mother wavelet $\psi(t)$.

In practice, we restrict the analysis to positive scales, as negative scales do not provide additional information for real-valued signals but introduce redundancy in the representation. This can be achieved by adopting the slightly more stringent admissibility condition (Daubechies, 1992)

$$C_\psi = \int_0^{+\infty}\frac{|\hat{\psi}(\omega)|^2}{\omega}d\omega < \infty.$$

Two common situations encountered in practice satisfy this condition: (i) real-valued wavelets for which $\hat{\psi}(-\omega) = \overline{\hat{\psi}(\omega)}$ and (ii) analytical wavelets for which support $\hat{\psi}(\omega) \in [0, +\infty)$. Theoretical background for both cases is provided in Mallat (2009).

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

In each case, the covariance between $f(t)$ and $g(t)$ can be reconstructed from (see Appendix B for proofs):

i. Real-valued wavelets

$$\langle f,g\rangle = \frac{1}{C_\psi}\int_0^{+\infty}\int_{-\infty}^{+\infty} W_f(a,b)\overline{W_g(a,b)}\,db\,\frac{da}{a^2}.$$

ii. Analytic wavelets applied to real-valued functions

$$\langle f,g\rangle = \frac{2}{C_\psi}\int_0^{+\infty}\int_{-\infty}^{+\infty} Re\left[W_f(a,b)\overline{W_g(a,b)}\right]db\,\frac{da}{a^2}.$$

In practical applications, continuous formulations must be discretized for sampled signals. The discretization relies on the following assumptions:

1. Scales are logarithmically spaced with base 2, such that two consecutive scales differ by a factor $2^{\delta j}$. Under this assumption, $\frac{da}{a} = d\big(\ln(a)\big) \simeq \delta j \ln(2)$.
2. The translation parameter b is replaced by the discrete time index n and scales a are expressed in dimensionless form as $\frac{a_j}{\delta t}$.

Applying these approximations leads to the discrete formulas presented in the main text.

Appendix B: Proof of formulas

i. Real-valued wavelets

For a real-valued mother wavelet $\psi(t)$, the reconstruction identity reads

$$\langle f,g\rangle = \frac{1}{C_\psi}\int_0^{+\infty}\int_{-\infty}^{+\infty} W_f(a,b)\overline{W_g(a,b)}\,db\,\frac{da}{a^2}.$$

To demonstrate this identity, we start from the Fourier-domain Parseval relation:

$$\langle f,g\rangle = \int_{-\infty}^{+\infty} f(t)\overline{g(t)}\,dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{f}(\omega)\overline{\hat{g}(\omega)}\,d\omega.$$

Applying Parseval's relation to the wavelet coefficients and using the Fourier-domain expression of the wavelet transform, $\widehat{W}_f(a,\omega) = \hat{f}(\omega)\sqrt{a}\,\overline{\hat{\psi}(a\omega)}$, yields

$$\frac{1}{C_\psi}\int_0^{+\infty}\int_{-\infty}^{+\infty} W_f(a,b)\overline{W_g(a,b)}\,db\,\frac{da}{a^2} = \frac{1}{C_\psi}\int_0^{+\infty}\frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{f}(\omega)\overline{\hat{g}(\omega)}\left|\hat{\psi}(a\omega)\right|^2 d\omega\,\frac{da}{a}.$$

Introducing the change of variables $v = a\omega$ and applying Fubini's theorem allows isolation of $\int_0^{+\infty}\left|\hat{\psi}(v)\right|^2\frac{dv}{v} = C_\psi$. Admissibility constants cancel out, leading to

$$\langle f,g\rangle = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{f}(\omega)\overline{\hat{g}(\omega)}\,d\omega.$$

Replacing $g(t)$ by $f(t)$ in the covariance expression gives the variance formula described in Daubechies (1992) and Mallat (2009):

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

$$\langle f,f\rangle = \frac{1}{C_\psi}\int_0^{+\infty}\int_{-\infty}^{+\infty}\left|W_f(a,b)\right|^2 db\,\frac{da}{a^2}.$$

ii. Analytic wavelets applied to real-valued functions

Let $f(t)$ be a real-valued function and let $f_a(t)$ denote its analytic signal defined in the Fourier domain as

$$\hat{f}_a(\omega) = \begin{cases} 2\hat{f}(\omega) & \text{if } w \geq 0, \\ 0 & \text{if } w < 0. \end{cases}$$

An analogous definition holds for $g_a(t)$. Using the same reasoning as above, the energy conservation relation applied to the analytic signals reads

$$\langle f_a, g_a\rangle = \frac{1}{C_\psi}\int_0^{+\infty}\int_{-\infty}^{+\infty} W_{f_a}(a,b)\overline{W_{g_a}(a,b)}\,db\,\frac{da}{a^2}.$$

For real-valued functions, the analytic signal satisfies $f_a = f + iH[f]$ where $H[\cdot]$ denotes the Hilbert transform. This implies

$$Re[\langle f_a, g_a\rangle] = \langle f,g\rangle + \langle \mathrm{H}[f], \mathrm{H}[g]\rangle = 2\langle f,g\rangle,$$

since the Hilbert transform preserves inner products for real signals, i.e. $\langle \mathrm{H}[f], \mathrm{H}[g]\rangle = \langle f,g\rangle$.

For an analytic wavelet, whose Fourier support is restricted to positive frequencies, the wavelet coefficients satisfy $W_{f_a}(a,b) = 2W_f(a,b)$ since

$$\widehat{W}_{f_a}(a,\omega) = \hat{f}_a(\omega)\sqrt{a}\overline{\hat{\psi}(a\omega)} = 2\hat{f}(\omega)\sqrt{a}\overline{\hat{\psi}(a\omega)} = 2\widehat{W}_f(a,\omega).$$

Combining these results leads to

$$\langle f,g\rangle = \frac{1}{2}Re[\langle f_a, g_a\rangle] = \frac{2}{C_\psi}\int_0^{+\infty}\int_{-\infty}^{+\infty} Re\left[W_f(a,b)\overline{W_g(a,b)}\right] db\,\frac{da}{a^2}.$$

The corresponding variance expression provided in Daubechies (1992) and Mallat (2009) is obtained by setting $g = f$:

$$\langle f,f\rangle = \frac{2}{C_\psi}\int_0^{+\infty}\int_{-\infty}^{+\infty}\left|W_f(a,b)\right|^2 db\,\frac{da}{a^2}.$$

Code and data availability

The code is openly available on GitHub at https://github.com/jonathanbitton/Wavelet-WAI. A permanent archived version is available through Zenodo at https://doi.org/10.5281/zenodo.20187583.
Data from BE-Lon is publicly available through the ICOS Carbon Portal.

Author contributions

JB: Conceptualization, Methodology, Formal analysis, Software, Data Curation, Writing - Original Draft, Visualization. BH: Conceptualization, Formal analysis, Resources, Writing - Review & Editing, Supervision, Project administration. SN:

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

Conceptualization, Methodology, Formal analysis, Writing - Review & Editing. CC: Conceptualization, Methodology, Formal analysis, Writing - Review & Editing, Supervision, Project administration.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

The authors gratefully acknowledge the support of the Integrated Carbon Observation System (ICOS) research infrastructure for providing continuous high-quality observations essential to this study, and particularly the team from the ICOS-Wallonia phase 2 project, supported by the Public Service of Wallonia (grant no. 1217769), and responsible for the BE-Lon station.

Financial support

The researchers were funded by the Fédération Wallonie-Bruxelles.

References

Amiri-Simkooei, A. R.: Application of least squares variance component estimation to errors-in-variables models, J. Geod., 87, 935–944, https://doi.org/10.1007/s00190-013-0658-8, 2013.

Aubinet, M., Vesala, T., and Papale, D. (Eds.): Eddy Covariance: A Practical Guide to Measurement and Data Analysis, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-007-2351-1, 2012.

Baldocchi, D., Chu, H., and Reichstein, M.: Inter-annual variability of net and gross ecosystem carbon fluxes: A review, Agric. For. Meteorol., 249, 520–533, https://doi.org/10.1016/j.agrformet.2017.05.015, 2018.

Coimbra, P. H. H., Loubet, B., Laurent, O., Mauder, M., Heinesch, B., Bitton, J., Delpierre, N., Berveiller, D., Depuydt, J., and Buysse, P.: Evaluation of a novel approach to partitioning respiration and photosynthesis using eddy covariance, wavelets and conditional sampling, Agric. For. Meteorol., 372, 110684, https://doi.org/10.1016/j.agrformet.2025.110684, 2025.

Daubechies, I.: Ten Lectures on Wavelets, Society for Industrial and Applied Mathematics, 369 pp., https://doi.org/10.1137/1.9781611970104, 1992.

Debnath, L. and Shah, F. A.: Wavelet Transforms and Their Applications, Birkhäuser, Boston, MA, https://doi.org/10.1007/978-0-8176-8418-1, 2015.

Dupont, S.: Scaling of Dust Flux With Friction Velocity: Time Resolution Effects, J. Geophys. Res. Atmospheres, 125, e2019JD031192, https://doi.org/10.1029/2019JD031192, 2020.

Foken, T. and Harrison, R. G.: The Natural Laboratory of the Solar Eclipse for Investigating Switching of Surface Layer Turbulent Fluxes, Bound.-Layer Meteorol., 192, 21, https://doi.org/10.1007/s10546-026-00963-y, 2026.

Foken, T., Göockede, M., Mauder, M., Mahrt, L., Amiro, B., and Munger, W.: Post-Field Data Quality Control, in: Handbook of Micrometeorology: A Guide for Surface Flux Measurement and Analysis, edited by: Lee, X., Massman, W., and Law, B., Springer Netherlands, Dordrecht, 181–208, https://doi.org/10.1007/1-4020-2265-4_9, 2004.

Fraedrich, K., Blender, R., and Zhu, X.: Continuum climate variability: long-term memory, scaling, and 1/f-noise, Int. J. Mod. Phys. B, 23, 5403–5416, https://doi.org/10.1142/S0217979209063729, 2009.

Franzke, C. L. E. and O'Kane, T. J. (Eds.): Nonlinear and Stochastic Climate Dynamics, Cambridge University Press, Cambridge, https://doi.org/10.1017/9781316339251, 2017.

Franzke, C. L. E., Barbosa, S., Blender, R., Fredriksen, H.-B., Laepple, T., Lambert, F., Nilsen, T., Rypdal, K., Rypdal, M., Scotto, Manuel G., Vannitsem, S., Watkins, N. W., Yang, L., and Yuan, N.: The Structure of Climate Variability Across Scales, Rev. Geophys., 58, e2019RG000657, https://doi.org/10.1029/2019RG000657, 2020.

Ge, Z.: Significance tests for the wavelet power and the wavelet power spectrum, Ann. Geophys., 25, 2259–2269, https://doi.org/10.5194/angeo-25-2259-2007, 2007.

Göckede, M., Kittler, F., and Schaller, C.: Quantifying the impact of emission outbursts and non-stationary flow on eddy-covariance CH_4 flux measurements using wavelet techniques, Biogeosciences, 16, 3113–3131, https://doi.org/10.5194/bg-16-3113-2019, 2019.

Kirlin, R. L., Done, W. J., and Hill, S. J. (Eds.): Covariance Analysis for Seismic Signal Processing, Society of Exploration Geophysicists, https://doi.org/10.1190/1.9781560802037, 1999.

Li, Y., Wu, Y., Tang, J., Zhu, P., Gao, Z., and Yang, Y.: Quantitative Evaluation of Wavelet Analysis Method for Turbulent Flux Calculation of Non-Stationary Series, Geophys. Res. Lett., 50, e2022GL101591, https://doi.org/10.1029/2022GL101591, 2023.

Lovejoy, S. and Schertzer, D.: The Weather and Climate: Emergent Laws and Multifractal Cascades, Cambridge University Press, Cambridge, https://doi.org/10.1017/CBO9781139093811, 2013.

Mallat, S. (Ed.): A Wavelet Tour of Signal Processing, Third Edition., Academic Press, Boston, https://doi.org/10.1016/B978-0-12-374370-1.50001-9, 2009.

Metzger, S., Junkermann, W., Mauder, M., Butterbach-Bahl, K., Trancón y Widemann, B., Neidl, F., Schäfer, K., Wieneke, S., Zheng, X. H., Schmid, H. P., and Foken, T.: Spatially explicit regionalization of airborne flux measurements using environmental response functions, Biogeosciences, 10, 2193–2217, https://doi.org/10.5194/bg-10-2193-2013, 2013.

Montillet, J.-P. and Bos, M. S. (Eds.): Geodetic Time Series Analysis in Earth Sciences, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-030-21718-1, 2020.

Paleri, S., Desai, A. R., Metzger, S., Durden, D., Butterworth, B. J., Mauder, M., Kohnert, K., and Serafimovich, A.: Space-Scale Resolved Surface Fluxes Across a Heterogeneous, Mid-Latitude Forested Landscape, J. Geophys. Res. Atmospheres, 127, e2022JD037138, https://doi.org/10.1029/2022JD037138, 2022.

Percival, D. B. and Walden, A. T.: Wavelet Methods for Time Series Analysis, Cambridge University Press, Cambridge, https://doi.org/10.1017/CBO9780511841040, 2000.

https://doi.org/10.5194/egusphere-2026-2881
Preprint. Discussion started: 17 July 2026

Rhif, M., Ben Abbes, A., Farah, I. R., Martínez, B., and Sang, Y.: Wavelet Transform Application for/in Non-Stationary Time-Series Analysis: A Review, Appl. Sci., 9, 1345, https://doi.org/10.3390/app9071345, 2019.

Richardson, A. D., Aubinet, M., Barr, A. G., Hollinger, D. Y., Ibrom, A., Lasslop, G., and Reichstein, M.: Uncertainty Quantification, in: Eddy Covariance: A Practical Guide to Measurement and Data Analysis, edited by: Aubinet, M., Vesala, T., and Papale, D., Springer Netherlands, Dordrecht, 173–209, https://doi.org/10.1007/978-94-007-2351-1_7, 2012.

Schaller, C., Göckede, M., and Foken, T.: Flux calculation of short turbulent events; comparison of three methods, Atmospheric Meas. Tech., 10, 869–880, https://doi.org/10.5194/amt-10-869-2017, 2017.

Seo, E., Dirmeyer, P. A., Barlage, M., Wei, H., and Ek, M.: Evaluation of Land–Atmosphere Coupling Processes and Climatological Bias in the UFS Global Coupled Model, J. Hydrometeorol., 25, 161–175, https://doi.org/10.1175/JHM-D-23-0097.1, 2023.

Stull, R. B. (Ed.): An Introduction to Boundary Layer Meteorology, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-009-3027-8, 1988.

The MathWorks Inc: MATLAB version: 9.12.0 (R2022a), 2022.

Torrence, C. and Compo, G. P.: A Practical Guide to Wavelet Analysis, Bull. Am. Meteorol. Soc., 79, 61–78, https://doi.org/10.1175/1520-0477(1998)079%3C0061:APGTWA%3E2.0.CO;2, 1998.

Whitcher, B., Guttorp, P., and Percival, D. B.: Wavelet analysis of covariance with application to atmospheric time series, J. Geophys. Res. Atmospheres, 105, 14941–14962, https://doi.org/10.1029/2000JD900110, 2000.

Witt, A. and Malamud, B. D.: Quantification of Long-Range Persistence in Geophysical Time Series: Conventional and Benchmark-Based Improvement Techniques, Surv. Geophys., 34, 541–651, https://doi.org/10.1007/s10712-012-9217-8, 2013.

Yi, H. and Shu, H.: The improvement of the Morlet wavelet for multi-period analysis of climate data, Comptes Rendus Geosci., 344, 483–497, https://doi.org/10.1016/j.crte.2012.09.007, 2012.

Yuan, N., Franzke, C. L. E., Xiong, F., Fu, Z., and Dong, W.: The impact of long-term memory on the climate response to greenhouse gas emissions, Npj Clim. Atmospheric Sci., 5, 70, https://doi.org/10.1038/s41612-022-00298-8, 2022.