Author comment: response to referee n°2

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

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

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

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

We thank all referees for their careful and constructive evaluation of our manuscript. Their comments have helped us improve the clarity, methodological robustness and overall presentation of the study.

This document provides a detailed, point-by-point response to all comments. For transparency, responses are organized as follows: an overall assessment, followed by general comments and, where applicable, specific minor or line-by-line comments.

All modifications made in the manuscript are indicated at the end of each response and introduced by the symbol "⇨". References to sections, figures and tables in the revised manuscript are highlighted in blue.

Overall assessment

"This manuscript examines the reconstruction of variance and covariance from continuous wavelet transform (CWT) coefficients and argues that a commonly used formulation based on Torrence and Compo (1998) can introduce systematic deviations associated with normalization and discretization choices. The authors derive an alternative estimator from the CWT inner-product identity and compare the resulting formulation with the Torrence–Compo approach using sinusoidal, white-noise, and pink-noise signals, followed by an application to eddy-covariance surface turbulence observations.

The topic is relevant to Atmospheric Measurement Techniques. Accurate reconstruction of second-order moments is important in atmospheric signal analysis, and normalization issues in wavelet-based variance and covariance calculations deserve careful examination. The manuscript also has several positive attributes: the presentation is generally clear, the proposed formulation is explicitly connected to the CWT energy identity, and the numerical results suggest that differences between reconstruction conventions can be non-negligible. The reported deviations of approximately 10–15% are sufficiently large to merit attention in applications involving turbulent covariance and flux calculations since such level exceeds the instrumental noises and retrieval errors.

However, I believe that several aspects of the manuscript require substantial clarification and additional analysis before its broader conclusions can be supported. In particular, the manuscript currently moves too fast between four distinct issues: (1) an exact continuous CWT inner–product identity; (2) its numerical discretization; (3) the statistical properties of a finite–sample covariance estimator; and (4) the physical interpretation of covariance as turbulent transport under nonstationary atmospheric conditions. These levels are related but are not equivalent.

My principal concern is therefore not with the motivation for reconsidering CWT normalization, which I find worthwhile, but with the statistical and physical claims subsequently built upon the reconstruction result. I recommend Major Revisions."

We thank the referee for the detailed and constructive assessment of our manuscript. The comments led us to substantially revise both the numerical validation and the interpretation of the atmospheric application.

The principal revisions include:

- a clear distinction between the continuous CWT identity, its discrete approximation and the finite-sample statistical behavior of the reconstruction;
- genuine Monte Carlo experiments with 200 independent bivariate realizations per configuration;
- separate comparisons with finite-record and theoretical moments;
- direct validation of covariance for positive, negative and near-zero covariance;
- additional tests involving frequency-dependent coherence, frequency-dependent phase and cospectral cancellation;
- addition of fractional Gaussian noise with three Hurst exponents;
- sensitivity analyses for scale resolution and discrete filter partitioning;
- quantitative verification of the implemented Morlet wavelet's frequency support;

- scale-resolved covariance analysis and time–scale decomposition of representative stationary and non-stationary records;
- a more cautious distinction between covariance reconstruction and turbulent-flux interpretation.

General comments

<u>Comment n°1</u>: "*Mathematical reconstruction identity versus statistical estimator*.

The central derivation begins from the CWT inner-product identity and produces discrete expressions intended to reconstruct variance and covariance. However, the manuscript then describes the resulting quantity as a statistical estimator and makes statements concerning bias, robustness, convergence, and asymptotic consistency. The sequence continuous identity → discrete approximation → finite-sample estimator should not be treated as a single step. If "consistency" is intended in its statistical sense, the assumptions under which this result is claimed should be explicitly stated and preferably demonstrated analytically. Otherwise, the terminology should be modified."

We agree that the original manuscript did not adequately distinguish the three separate levels: the continuous CWT inner-product identity, its discrete numerical approximation and the statistical behavior of the reconstruction when applied to finite stochastic records.

The continuous identity establishes the analytical basis of the formulation. The implemented equations constitute a finite numerical quadrature of that identity. The revised Monte Carlo experiments (see response to Comment n°3) now address the third level by quantifying the bias, dispersion and RMSE of the numerical reconstructions over independent stochastic realizations. These experiments provide numerical evidence of convergence over the investigated processes, parameter ranges and record lengths. They do not constitute a general analytical proof of consistency. We have therefore removed or reformulated statements claiming "asymptotic consistency" in its formal statistical meaning.

We now distinguish two reference quantities:

- the finite-record reference, which is calculated directly from each simulated realization. It is used to assess whether the discretized wavelet integral reconstructs the moment of the record analyzed (finite-record reconstruction error) and isolates the effects of normalization, scale discretization, finite scale range and numerical implementation.
- the theoretical reference, which is the population moment prescribed by the generating process. It is used to assess the combined effects of finite-sample variability and wavelet-domain reconstruction relative to the population parameter.

The figures in the main manuscript use the finite-record reference because reconstruction of the analyzed finite signal is the primary methodological question addressed by the proposed normalization. Furthermore, for the sinusoidal experiment, the theoretical reference represents the infinite-duration or complete-cycle moment, whereas the finite-record reference accounts for phase and incomplete-cycle effects in each analyzed record. Convergence results relative to the theoretical process moments are provided in the Supplement. A quantitative table reports error metrics, including bias and RMSE, relative to both references.

For clarity, we now use "finite-record error" and "error relative to the theoretical process moment" to avoid ambiguity. We no longer claim general analytical consistency and instead refer to numerical convergence over the investigated models and record lengths.

⇨ Changes in Manuscript: In the revised manuscript, we have:

i. replaced the pseudo-ensemble design with genuine Monte Carlo experiments (see response to Comment 3) in Sect. 2.2.2;
ii. adapted the terminology used in the whole manuscript to distinguish analytical reconstruction from finite-sample performance;
iii. defined finite-record and theoretical process references and reported corresponding error metrics in Sect. 2.2.3;
iv. replaced formulations in Sects 3.1–3.2, 4.1–4.2 and 5 on “consistency” by statements concerning observed numerical convergence over the tested record lengths

---

Comment n°2: "***Relationship between covariance in the time, Fourier, and wavelet domains should be made explicit.***

The manuscript should explicitly present time-domain covariance at zero lag, the integrated cospectrum and the integrated CWT cross-product as representations of the same global second-order quantity. Total covariance reconstruction should be distinguished from time-scale decomposition of covariance and identification of turbulent transport."

As mentioned by the referee, for real-valued signals, the zero-lag covariance calculated from the finite-record fluctuations, the integrated Fourier cospectrum and the integrated real CWT cross-product after normalization are alternative representations of the same global second-order quantity. Their theoretical equivalence assumes that the three representations target the same signal components and frequency range. In numerical implementations, their equivalence is additionally subject to the approximations introduced by finite record duration and discretized implementation. The proposed CWT formula was not intended to define a new form of covariance.

The principal subject of this work is indeed the reconstruction of total covariance. The CWT reconstruction additionally provides a localized decomposition of the covariance by time and scale, but this decomposition does not change the definition of the globally integrated covariance and is not itself the focus of the reconstruction benchmark.

The observational application was intended to compare global covariance reconstructions and to examine how their scale composition changes when the assumptions underlying block-averaged Reynolds decomposition are not satisfied. Although the CWT is better suited to resolving non-stationary covariance contributions, we agree that conservation of the total inner product does not identify which scale-time contributions represent turbulent transport. Physical attribution of covariance to turbulent transport requires additional choices concerning averaging interval, detrending, frequency range, stationarity and separation of turbulent and lower-frequency contributions, which exceed the intended scope of the article.

⇨ Changes in Manuscript: In the revised manuscript, we have:

i. clarified the equivalence of covariance expressions by stating explicitly that the integral of the cospectrum and the real part of the integrated cross-wavelet product yields the covariance at zero lag in Sect. 2.1.1.
ii. separated the reconstruction of the integrated covariance, the localization of the covariance and the physical selection of turbulent contributions in Sects. 2.3, 3.4, 4.2 and 4.4;
iii. clarified that the method improves reconstruction of total covariance but does not by itself define a turbulence–mesoscale separation criterion in Sect. 4.4.

Comment n°3: "***The synthetic validation should use genuine Monte Carlo ensembles.***

Varying amplitude, frequency, variance, and sample length is not equivalent to generating independent realizations from the same stochastic process. Independent realizations should be generated for every combination of process parameters and N, and bias, variance, RMSE and confidence intervals should be reported as functions of N."

We agree and have replaced the pseudo-ensemble validation with genuine Monte Carlo experiments. For each combination of signal type, wavelet, record length and process parameters, we generated 200 independent bivariate realizations. For the mean reconstruction error, this corresponds to a Monte Carlo standard error of $\frac{1}{\sqrt{200}} \approx 0.071$ times the between-realization standard deviation.

The revised experiment includes:

- sinusoidal pairs with controlled amplitude (1, 2, 5, 10), period (1, 2, 5, 10, 20, 50, 100), random initial phase and relative phase shift (see response to Comment n°5);
- Gaussian white-noise pairs with controlled marginal variance (0.5, 1, 2, 5, 10, 25, 50, 100, 150) and correlation (see response to Comment n°5);
- pink-noise pairs retained as a finite-record stress test;
- fractional Gaussian noise pairs with controlled variance, correlation and Hurst exponent (see response to Comment n°4).

For each simulated pair, both marginal variances and the covariance were reconstructed using the new formulation and the original TC formulation. The simulations were paired across wavelets: Morlet and DOG2 were applied to the same realizations for each parameter combination.

The revised analysis no longer pools all record lengths into a single distribution when evaluating convergence. We now present:

- a summary comparison at the largest record length, $N = 2^{19}$;
- separate convergence curves as functions of N, showing the median and interquartile range;
- tabulated Monte Carlo bias, variance and RMSE for each method, signal type, wavelet and record length.

For marginal variance, we verified that summaries based on X, Y or the within-pair mean of their errors were indistinguishable. To avoid treating the two correlated marginal reconstructions as independent observations, the main results report variance results for X only.

⇨ Changes in Manuscript: In the revised manuscript, we have:

i. replaced the pseudo-ensemble experimental design by Monte Carlo simulations in Sect. 2.2.2 and expanded parameter ranges tested;
ii. described error metrics in Sect. 2.2.3;
iii. adapted Figs. 1–2 with the Monte Carlo results and added a summary Table S1 indicating error metrics for all signals;

Comment n°4: "*The pink-noise experiment requires reconsideration.*

A nominal 1/f process depends critically on the low-frequency cutoff. As N changes, the low-frequency content and potentially the target variance may change. The authors should explain exactly how pink noise is generated and should supplement or replace this experiment with a stationary long-memory process such as fractional Gaussian noise, using several Hurst exponents."

We agree with the referee that the original interpretation of the pink-noise experiment was too strong.

The pink-noise generator produces finite-length colored-noise records by passing a uniformly distributed white-noise input through a fixed cascade of six second-order filter sections with randomized initial states. The generator reproduces the MATLAB pinknoise implementation, including its six second-order filter sections, gain normalization and randomized initial filter states. The applied scale parameter controls the signal amplitude but is not interpreted as an exact ensemble variance. Because the same filter is used for all record lengths, the underlying discrete filtering mechanism is fixed; however, the lowest sampled Fourier frequency and the realized contribution of the longest timescales depend on record duration. Consequently, the pink-noise sequence is not used in the revised manuscript as the principal basis for claims concerning long-memory asymptotic behavior.

Nevertheless, we retain pink noise as a practical finite-record stress test because it produces strong low-frequency spectral weighting and remains relevant to the type of broadband spectra motivating the study. However, we no longer describe its target scaling parameter as an exact theoretical ensemble variance. Reconstruction errors for pink noise are evaluated relative to the variance and covariance calculated directly from each finite realization.

More importantly, as suggested by the referee, we added a controlled stationary long-memory experiment using fractional Gaussian noise (fGn). The fGn simulations use a circulant-embedding generator with prescribed marginal variance and three Hurst exponents: 0.6, 0.75 and 0.9. All three values correspond to persistent fGn, with progressively stronger long-range dependence. This provides an explicit progression from weak to strong long-range dependence while retaining a well-defined stationary process and controlled marginal variance.

The revised convergence analysis reports the three Hurst scenarios separately. This allows the effect of increasingly strong low-frequency persistence to be assessed without relying on the record-length dependence of the lowest sampled frequency in the finite pink-noise records.

⇨ Changes in Manuscript: In the revised manuscript, we have:

i. presented the three fractional Gaussian noise signals corresponding to the three selected Hurst parameter values in Sect. 2.2.1;
ii. added a dedicated fGn panel in Figs. 1–2 showing the three values of H;
iii. explained the precise definition of the pink noise in Sect. 2.2.1;
iv. used fGn as the principal controlled long-memory result and present pink noise only as a finite-record stress test.

Comment n°5: "*Variance validation cannot simply substitute for covariance validation.*

Covariance introduces phase, coherence, sign changes in the cospectrum, and cancellation between positive and negative contributions. The authors should construct bivariate stochastic signals with a known cross-spectrum and test positive and negative covariance, broadband and narrowband covariance, frequency-dependent coherence, phase structure, and cospectral cancellation."

We agree that the original variance-only benchmark should be supplemented by a covariance benchmark given the title and the atmospheric application. We have therefore extended the entire Monte Carlo experiment to bivariate signals and directly evaluated covariance reconstruction.

For every simulated realization, we generated a pair (x,y) and reconstructed $Cov(x,y)$. The revised experiments include both narrowband and broadband covariance structures:

1. For the sinusoidal experiment:
$$x(t) = A\sin(\omega t + \phi_0),$$
$$y(t) = A\sin(\omega t + \phi_0 + \Delta\phi),$$
with:
$$\Delta\phi \in \left\{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}, \pi\right\}.$$
and a randomly sampled initial phase ϕ_0. The theoretical covariance is
$$Cov(x,y) = \frac{A^2}{2}\cos(\Delta\phi).$$
This experiment directly tests positive/negative/zero covariance under a controlled narrowband phase relationship.
2. For white noise and fGn, independent unit-variance processes u and v were combined as:
$$x(t) = \sigma u(t),$$
$$y(t) = \sigma\left(\rho u(t) + \sqrt{1-\rho^2}v(t)\right),$$
with:
$$\rho \in \{-0.9, -0.5, 0, 0.5, 0.9\}.$$
The theoretical covariance is:
$$Cov(x,y) = \rho\sigma^2.$$
These experiments test positive/negative/zero broadband covariance under both short-memory and long-memory conditions.
Pink-noise pairs were generated using the same correlation construction, but their covariance was evaluated relative to the empirical covariance of each finite pair rather than an assumed theoretical target.

Relative covariance errors were summarized only for configurations with a prescribed non-zero covariance. Nominally zero-covariance configurations, corresponding to $\Delta\phi = \frac{\pi}{2}$ for the sinusoids and $\rho = 0$ for the stochastic processes, were retained as a separate control using absolute covariance error. Including these cases in pooled summaries increased dispersion, which is expected for a target covariance near zero, but did not alter the relative ranking of the methods or their convergence with record length.

We further added a targeted frequency-domain experiment with prescribed spectra $S_{xx}(f)$ and $S_{yy}(f)$, coherence $\gamma_{xy}(f)$ and phase $\phi(f)$:

$$S_{xy}(f) = \gamma_{xy}(f)\sqrt{S_{xx}(f)S_{yy}(f)}e^{i\phi(f)}$$

Three structures were tested:

- broadband positive covariance;
- frequency-dependent coherence and phase;
- positive and negative cospectral bands divided into a low-frequency contribution C_{low} and a high-frequency contribution C_{high} whose integrated contributions nearly cancel.

For the cancellation case, the two bands had similar magnitudes and opposite signs, such that their sum approached zero. We therefore evaluated reconstruction error relative to the total magnitude involved $|C_{low}| + |C_{high}|$, rather than relative to the near-zero net covariance.

The proposed formulation more closely reproduced the prescribed integrated covariance in the broadband-positive and frequency-dependent coherence and phase experiments. In the cancellation experiment, the difference between the two formulations was smaller because the principal TC deviation is approximately multiplicative across both bands. Similar scaling of the positive and negative contributions preserved their cancellation in the integrated covariance, although their individual bandwise magnitudes were less accurately reconstructed.

⇨ Changes in Manuscript: In the revised manuscript, we have:

i. described the bivariate construction with theoretical covariances and the controlled cross-spectral experiment in Sects. 2.2.1 and 2.2.4;
ii. generalized error metrics used for both variance and covariance in Sect. 2.2.3;
iii. added the direct covariance-reconstruction results for the sinusoidal, white-noise, pink-noise and fractional-Gaussian-noise experiments in Sects. 3.1 and 3.2;
iv. added a Results section (Sect. 3.3.2) dedicated to covariance benchmarking for narrowband and broadband covariance structures;
v. added a Supplement section dedicated to the controlled cross-spectral experiment.

Comment n°6: "*Reconstruction of total covariance should not be equated with turbulent flux under nonstationarity.*

The manuscript argues that wavelet-based approaches can relax the stationarity requirement of conventional eddy-covariance analysis and allow otherwise rejected data to be retained. However, time-frequency localization and conservation of total covariance do not by themselves provide a physical criterion for separating turbulent transport from mesoscale variability, trends, transitions, coherent motions, or other nonstationary contributions. The manuscript should moderate the claim that wavelet reconstruction permits nonstationary periods to be incorporated directly into flux estimation."

We agree with the referee's central distinction, that time-scale localization and covariance reconstruction do not, by themselves, determine which contributions should be interpreted as turbulent transport.

Nevertheless, we would like to clarify the scope of the statement made in the original manuscript. The ability to analyze non-stationary intervals arises from the wavelet transform itself, rather than from the reconstruction normalization introduced in this study. Because the CWT resolves fluctuations jointly in time and scale, it can identify non-stationary events and integrate their covariance contributions over a prescribed range of periods without requiring the statistical properties of the complete 30 min interval to remain constant. This property has been used in previous wavelet-flux studies to investigate periods for which conventional EC steady-state tests indicate low data quality (notably in Schaller et al., 2017).

The role of the present study is different and more specific. Given an adopted wavelet, scale range and maximum period, we examine whether the selected wavelet cross-products are integrated with the correct reconstruction factor. The proposed normalization does not influence the turbulence–mesoscale cutoff and does not determine whether an individual low-frequency contribution should be regarded as turbulent transport. The question of whether these contributions should or should not be included in the total flux exceeds the scope of the manuscript. Likewise, equality or improved reconstruction of total covariance does not determine which part of that covariance should be interpreted as turbulent vertical transport.

In the observational comparison, we restricted the CWT integration to periods comparable to the conventional 30 min EC averaging interval. This choice was made to compare quantities covering broadly similar temporal scales and to limit contributions from slower motions. In many boundary-layer conditions, a spectral gap separates turbulent and slower variability, although this separation can weaken during stable, intermittent or transitional conditions (Stull, 1988). We therefore agree that a 30 min cutoff reduces, but does not eliminate, the possibility that trends, transitions, coherent motions or mesoscale variability contribute to the integrated wavelet covariance.

We have revised the manuscript to distinguish more clearly among three points: (i) the applicability of the CWT to non-stationary records, (ii) the numerical reconstruction of covariance over a selected scale range and (iii) the physical interpretation of that covariance as turbulent transport. The first is an established property and application of wavelet analysis; the second is the methodological focus of the present study; and the third requires an independently justified scale-selection criterion and is not resolved by the normalization developed here.

Accordingly, we retain the statement that wavelet analysis provides a framework for examining non-stationary turbulent events, but we have moderated the broader suggestion that all non-stationary half-hours can be incorporated directly into operational flux estimates.

⇨ Changes in Manuscript: In the revised manuscript, we have:

i. moderated the atmospheric interpretation in Sect. 3.4 and in the Discussion, by presenting the proposed formulation as a reconstruction of total finite-record covariance and a framework for subsequent scale-resolved analyses.
ii. added an explicit distinction between capacity to analyze non-stationary motions and selection of physically relevant scales for separating turbulence and mesoscales components;
iii. emphasized the rationale behind the 30-min limit choice;
iv. avoided claiming direct operational flux improvement.

Comment n°7: "*The interpretation of the nonstationary eddy–covariance results is not unique.*

The disagreement between EC and CWT covariance under nonstationarity does not establish which estimate represents the appropriate turbulent flux. Additional low-frequency or transient contributions may enter the integrated wavelet covariance. Representative scale- and time-resolved diagnostics are recommended."

We thank the referee for this suggestion. We agree that the original manuscript did not sufficiently identify the time-scale origin of the increasing difference between conventional EC and wavelet covariance.

We would nevertheless clarify our interpretation of this difference. Conventional 30 min EC covariance relies on a block-averaged Reynolds decomposition, whose physical interpretation becomes problematic when the mean and higher-order properties evolve within the averaging interval. The CWT is better suited to such conditions because it localizes covariance contributions jointly in time and scale without requiring the complete half-hour to be stationary. However, this advantage does not by itself determine which upper period should define turbulent transport.

We therefore added a scale-resolved analysis of stationary and non-stationary records, following the referee's suggestion. The analysis examined signed and absolute covariance contributions as functions of equivalent Fourier period, together with the cumulative signed covariance obtained as progressively longer periods were included.

The non-stationary records exhibited stronger and more variable contributions at intermediate and long periods than the stationary records. Their cumulative covariance also continued to evolve more strongly as the maximum included period approached the 30 min cutoff. Representative cross-scalograms further showed that the additional covariance did not follow a single temporal pattern. Depending on the record, it appeared as a persistent long-period structure, a temporally localized event or an alternation of positive and negative contributions.

The comparison targets high-magnitude records representative of each quality class rather than two populations matched by time of day, covariance sign or atmospheric regime. Therefore, several selected non-stationary records occur at night or near diurnal transitions. We repeated the analysis after excluding every record located within ±2.5 h of a day/night transition. This exclusion extends approximately one hour beyond the maximum external temporal support used for the Morlet reconstruction at the 30 min upper-period limit. The main spectral pattern remained, with 23 Morlet and 24 DOG2 records retained (approx. half of the non-stationary dataset). Representative daytime records located 3.5–6.5 h from the nearest transition also displayed enhanced intermediate- and long-period covariance. The observed scale redistribution therefore cannot be attributed primarily to sunrise or sunset transitions.

We additionally tested whether this result depended on the use of observed neighboring data at the boundaries. Replacing the neighboring observations with reflected extension produced very similar relative contributions from periods between 10 and 30 min. For Morlet, the median non-stationary minus stationary contrast was 0.175 under reflected extension, compared with 0.167 using observed neighboring data. For DOG2, the corresponding contrasts were 0.209 and 0.195. Thus, the increased contribution of periods approaching the 30 min cutoff was not created solely by the external observations used to calculate the longest-period coefficients.

These analyses therefore support the referee's interpretation that the reduced agreement between EC and CWT covariance under non-stationarity coincides with a different time-scale composition of the covariance. In particular, the integrated covariance becomes more sensitive to contributions near the upper integration limit. However, we agree that the localization of these contributions does not determine whether all of them

should be interpreted as turbulent transport. The additional covariance may include turbulent structures, coherent motions, transient events or slower variability. Distinguishing among these physical processes requires an independently justified scale-selection criterion and remains separate from the reconstruction focus of the present study.

⇨ Changes in Manuscript: In the revised manuscript, we have:

i. added a description of the scale-resolved analysis, including the calculation of covariance contributions as functions of equivalent Fourier period and cumulative covariance as progressively longer periods were included in Sect. 2.3;
ii. expanded Sect. 3.4 with results from the scale-resolved analysis;
iii. added a Supplement with cross-scalograms of representative cases;
iv. revised the Discussion and Conclusions accordingly.

Comment n°8: "*Morlet is only approximately analytic.*

The conventional Morlet wavelet with $\omega_0 = 6$ is approximately rather than strictly analytic. The exact implementation and negative-frequency leakage should be documented because the derivation uses a factor of two associated with positive-frequency support."

We thank the referee for requesting clarification of the exact Morlet implementation. The original manuscript described the Morlet wavelet as approximately analytic, which is correct for the Morlet kernel.

For the conventional Morlet kernel with $\omega_0 = 6$, we obtained:

$$\frac{\sum_{\omega_k<0}\left|\hat{\psi}(\omega_k)\right|^2}{\sum_k\left|\hat{\psi}(\omega_k)\right|^2} = 1.08\ 10^{-17}.$$

The zero-frequency amplitude was:

$$\left|\hat{\psi}(0)\right| = 1.22\ 10^{-8}.$$

These values confirm that the conventional $\omega_0 = 6$ Morlet kernel is admissible to numerical precision for the present calculations.

More importantly, the actual Fourier-domain implementation explicitly sets the Fourier-domain wavelet to zero at all negative discrete frequencies following the implementation of Torrence and Compo (1998). The factor of two used in the reconstruction consequently corresponds to a strictly one-sided numerical implementation.

⇨ Changes in Manuscript: In the revised manuscript, we have expanded the description of the Morlet wavelet in Sect. 2.1.6 to quantify the negligible leakage of the Morlet kernel.

Comment n°9: "*The statement concerning two independent Morlet coefficient components is too general.*

Complex Morlet coefficients do not automatically possess statistically independent real and imaginary components. Such a result requires additional assumptions that do not hold for arbitrary correlated or nonstationary geophysical records."

We thank the referee for this clarification. We agree that the original wording was too general. A complex Morlet coefficient always contains two quadrature components, but their statistical independence does not follow from the complex form of the wavelet alone.

For a stationary Gaussian process, the real and imaginary coefficients are jointly Gaussian. They can be treated as approximately uncorrelated and of similar variance when the Morlet wavelet is approximately analytic and the process spectrum varies only weakly over its effective bandwidth. Under these conditions, the squared coefficient amplitude follows the usual approximate two-degree-of-freedom interpretation. A real-valued wavelet such as DOG2 has one degree of freedom at each time–scale point (Torrence and Compo, 1998; Ge, 2007).

This assumption is most directly applicable to the Gaussian white-noise experiment and may also provide a useful approximation for the stationary fGn experiments, subject to their local spectral slope. It is not, however, a universal explanation across all signals considered here. For deterministic sinusoids, non-Gaussian pink noise or non-stationary atmospheric records, the two quadrature components cannot automatically be interpreted as statistically independent realizations. We have revised the discussion accordingly and now present the two-degree-of-freedom argument as a mechanism applicable to the Gaussian stochastic experiments rather than as a general property of all Morlet coefficients.

⇨ Changes in Manuscript: In the revised manuscript, we have modified Sect. 4.3 to distinguish the algebraic presence of two quadrature components from their statistical independence. The two-degree-of-freedom interpretation is retained for Gaussian stochastic processes satisfying the assumptions of Torrence and Compo (1998) and Ge (2007), but is no longer invoked as a universal explanation for deterministic or non-stationary records.

Comment n°10: "***Boundary effects and extension using neighboring observations deserve greater attention.***

Extending 30 min intervals with neighboring observations reduces mathematical edge effects but introduces data from outside the target interval, potentially from different meteorological regimes. The temporal extent should be reported in physical units and sensitivity to extension procedures should be examined."

In the original analysis, observed neighboring data were used to preserve the continuity of the atmospheric record across the arbitrary 30 min processing boundaries. The temporal integration of covariance remained restricted to the target half-hour, but coefficients at the longest periods had temporal support extending beyond its boundaries.

At the 30 min upper-period limit, the retained Morlet coefficients required approx. 81.0 min of data on either side of the target half-hour, whereas DOG2 required approx. 19.3 min on either side, due to its more compact temporal support. Because the input was loaded in complete half-hour blocks, three neighboring blocks (1.5 h) were provided on either side for Morlet and one (0.5 h) for DOG2.

We compared the original calculation, based on observed neighboring data (reference), with three alternative boundary treatments: no explicit extension, zero padding and reflected extension. Differences were normalized by the absolute covariance integrated below 30 min, which provides a stable reference even when positive and negative contributions produce a small net covariance.

- Reflected extension produced the closest alternative to the reference. Across the two wavelets and stationarity classes, the median normalized differences ranged from approximately −2.2 % to −0.6 %. For stationary records, the 95th percentile of the absolute difference remained below approximately 5.3 % for both wavelets. Median differences under non-stationary conditions were also close to zero, although individual records showed greater variability, particularly for Morlet.
- For stationary records, the no-extension calculation also remained close to the reference, with median normalized differences of −2.4 % for Morlet and −1.6 % for DOG2. Under non-stationary conditions, however, the no-extension calculation became more sensitive to the boundary treatment. The difference was negligible at short periods and accumulated primarily between 10 and 30 min. This sensitivity was stronger and more localized near the upper-period limit for Morlet.
- Zero padding produced large and highly record-dependent modifications over a broad range of periods. The imposed transition from the observed signal to zero generated artificial boundary contributions of both signs, particularly under non-stationary conditions. Median normalized differences were 29.4 % for Morlet and 8.8 % for DOG2, but their wide interquartile ranges showed substantial variation in both directions (from −406.2 % to 299.0 % and from −329.5 % to 287.8 %, respectively respectively). The 95th percentiles of the absolute differences reached 1224 % and 1134 %, indicating the presence of extreme deviations in the distribution tails. We therefore do not regard zero padding as an appropriate boundary treatment for these continuous atmospheric records.

The principal scale-resolved result described in response to Comment n°7 remained robust to the boundary treatment. Under reflected extension, the median non-stationary minus stationary contrast in the relative contribution of periods between 10 and 30 min was 0.175 for Morlet and 0.209 for DOG2. These values closely matched the contrasts of 0.167 and 0.195 obtained with observed neighboring data. The enhanced intermediate- and long-period contribution of the non-stationary records is therefore not an artefact caused solely by the use of adjacent observations.

We retained observed neighboring data as the primary treatment because the atmospheric record is continuous and the half-hour boundaries are imposed by the processing convention rather than by physical discontinuities. This treatment avoids the artificial boundary discontinuity introduced by zero padding and preserves events crossing the nominal block boundaries. We nevertheless now state explicitly that the longest-period coefficients have an effective temporal support extending beyond the target half-hour. Individual non-stationary records, particularly with Morlet, can therefore remain sensitive to changes in the surrounding atmospheric conditions.

⇨ Changes in Manuscript: In the revised manuscript, we have:

i. reported the temporal extension used for each wavelet in Sect. 2.3;
ii. clarified the distinction between the target integration interval and the wider support of the longest-period coefficients in Sect. 2.3;
iii. added the boundary-sensitivity analysis in Sects. 2.3 and 3.4.2;
iv. added period-resolved and cumulative comparisons to the Supplement;
v. expanded the Discussion and limitations regarding boundary treatment and temporal support.

Additional comments

"The manuscript would benefit from a clearer notation separating the continuous scale a, discrete scale a_j, Fourier frequency, and equivalent Fourier period. Because much of the paper concerns normalization, dimensional consistency should be explicitly checked and units associated with the wavelet coefficients should be stated."

We thank the referee for this suggestion. The manuscript already distinguishes continuous scale a from discrete scale a_j, but the conversion from scale to physical period using the wavelet-dependent Fourier factor F_ψ was not stated explicitly. We now define the wavelet-dependent equivalent Fourier period P_j and clarify that the equivalent 30 min cutoff refers to this specific period, not directly to scale.

The continuous formulations follow the standard CWT definitions. Both scale a and translation b have units of time, whereas the normalized wavelet argument $\frac{t-b}{a}$ is dimensionless. Therefore, under L^1 normalization, a wavelet coefficient has the units of the input signal, whereas under L^2 normalization it has the units of the input signal multiplied by the square root of time.

In the reconstruction formula, the corresponding L^1 and L^2 scale measures, $\frac{db\,da}{a}$ and $\frac{db\,da}{a^2}$, both reconstruct a time-domain inner product with units of signal product multiplied by time. Dividing by the analyzed duration therefore yields variance or covariance in the original physical units. In the discrete implementation, the explicit factors in δt preserve this correspondence.

⇨ Changes in Manuscript: In the revised manuscript, we have:

i. introduced the equivalent Fourier Factor in Sects. 2.1.5 and 2.1.6;
ii. added relevant information related to the dimensional analysis in Appendix B.

"The authors should also clarify whether the variance against which their reconstructed estimates are compared is the theoretical ensemble variance of the generating process or the conventional sample variance of each particular realization. These are different reference quantities for finite stochastic samples and lead to different interpretations of error."

We agree. This distinction is now central to the revised validation and is described in our response to Comment n°1. The main figures use finite-record references. Supplementary figures use theoretical process references for the sinusoidal, white-noise and fGn experiments; pink noise is evaluated only against finite-record moments. Quantitative tables contain both.

⇨ Changes in Manuscript: In the revised manuscript, we have:

i. defined both references explicitly in Sect. 2.2.3;
ii. distinguished between finite and theoretical references in Sect. 3.1 and 3.2;
iii. added a Supplementary section for theoretical references.

"For the sinusoidal experiment, it would be useful to indicate how the sinusoidal frequency is positioned relative to the discrete wavelet scales. Spectral leakage and scale-grid mismatch may contribute to numerical reconstruction errors independently of the normalization issue."

We agree that scale-grid mismatch and finite-record leakage can, in principle, affect sinusoidal reconstruction. In our implementation, the minimum scale was defined as $a_0 = \frac{\delta t}{F_\psi}$, where F_ψ is the Fourier scale-to-period conversion factor, such that the equivalent-period grid is $P_j = \delta t 2^{j\delta j} F_\psi$ for both wavelets. With $\delta j = 0.05$, the prescribed sinusoidal periods differed from the nearest equivalent grid period by at most 1.51 %. Moreover, reconstruction error was already small at the shortest investigated record lengths and did not show the behavior expected from substantial scale mismatch. We therefore clarified the grid construction and the maximum mismatch in the Methods, without adding a separate analysis figure.

⇨ Changes in Manuscript: In the revised manuscript, we have added in Sect. 2.2.2 a discussion on the scale-period matching.

"The selection of $\delta j = 0.05$ is very fine. A sensitivity analysis to scale discretization would be informative because one objective of the paper is specifically to separate normalization bias from numerical quadrature error."

We performed a targeted sensitivity analysis using $\delta j \in \{0.025, 0.05, 0.1, 0.25, 0.5\}$. The analysis included both wavelets and multiple record lengths for sinusoidal signals at several periods, white noise and fGn with $H = 0.9$.

The results show that the production value $\delta j = 0.05$ is already stable. For Morlet, the summed filter partition was nearly constant for $\delta j \leq 0.1$, whereas $\delta j = 0.5$ produced pronounced oscillations. The Monte Carlo reconstructions likewise showed no material improvement when refining the grid from 0.05 to 0.025. DOG2 remained less uniform than Morlet, but the production-grid results were stable and the residual white-noise bias was subsequently explained by the finite full-band partition rather than by insufficient scale density alone. This confirms that the main difference between the proposed and TC formulations is not caused by insufficient scale resolution at the production setting.

⇨ Changes in Manuscript: In the revised manuscript, we have:

i. added a sensitivity analysis on δj in Sect. 3.3.1 and in the Supplement;
ii. justified the choice $\delta j = 0.05$ using the sensitivity results.

"The manuscript states that the DOG2 residual white-noise bias arises from imperfect spectral partitioning and overlapping filter responses. This interpretation would be strengthened considerably by plotting the summed frequency responses of the discretized wavelet filters over the scales actually used in the analysis. Such a plot could directly demonstrate whether departures from an approximately constant partition of energy explain the observed residual bias."

We thank the referee for suggesting a direct examination of the summed discrete filter responses. This analysis clarified the origin of the residual DOG2 white-noise bias. At the scale step $\delta_j = 0.05$, the new Morlet formulation had a full-band white-noise multiplier of approximately 2, which was expected for an approx. analytic wavelet and a real-valued signal, corresponding to a predicted relative bias below 10^{-5}. The observed Monte Carlo bias was likewise negligible.

For DOG2, the summed discrete response had a full-band multiplier of 0.98137, predicting a white-noise reconstruction bias of −1.863 %. The independent Monte Carlo experiment yielded a mean relative error of −1.863 %, with very little variation among realizations (standard deviation of 3.9×10^{-5}). The near-exact agreement indicates that the residual DOG2 bias originates from the finite discrete scale partition and its non-uniform full-band spectral coverage, rather than from filter overlap itself.

⇨ Changes in Manuscript: In the revised manuscript, we have:

i. developed the discrete filter partition used in Sect. 3.3.1;
ii. added a Figure in the Supplement showing the absolute frequency partition and its frequency-dependent variation for Morlet and DOG2 across all tested values of δj;
iii. replaced the previous qualitative explanation based on overlapping responses in Sect. 4.3 by this quantitative partition analysis. We also distinguish this residual finite-grid reconstruction bias from the analytical reconstruction constant addressed by the proposed formulation.

“Finally, because the paper emphasizes implications for eddy-covariance flux calculations, I encourage the authors to distinguish consistently between covariance reconstruction and flux estimation. Their observational analysis evaluates covariance and explicitly does not convert the result to flux units. Claims concerning operational flux improvement should therefore remain appropriately qualified.”

The observational analysis compares covariance reconstructions and does not apply the complete sequence of spectral corrections, storage terms or conversions required for an operational flux product. We therefore use “covariance reconstruction” for the analyzed quantity and reserve “flux” for the broader physical application.

⇨ Changes in Manuscript: Terminology has been revised consistently throughout.