Article

Denoising Non-Stationary Signals via Dynamic Multivariate Complex Wavelet Thresholding

Kim C. Raath [1], Katherine B. Ensor [2], Alena Crivello [3] and David W. Scott [2,*]

1 Imperative Global; kcraath@gmail.com
2 Department of Statistics, Rice University, MS-138, Houston, TX 77005, USA; ensor@rice.edu
3 Chevron Co., San Ramon, CA 94583, USA; ixmocane@gmail.com
* Correspondence: scottdw@rice.edu; Tel.: +1-713-348-6037

Abstract: Over the past few years, we have seen an increased need to analyze the dynamically changing behaviors of economic and financial time series. These needs have led to significant demand for methods that denoise non-stationary time series across time and for specific investment horizons (scales) and localized windows (blocks) of time. Wavelets have long been known to decompose non-stationary time series into their different components or scale pieces. Recent methods satisfying this demand first decompose the non-stationary time series using wavelet techniques and then apply a thresholding method to separate and capture the signal and noise components of the series. Traditionally, wavelet thresholding methods rely on the discrete wavelet transform (DWT), which is a static thresholding technique that may not capture the time series of the estimated variance in the additive noise process. We introduce a novel continuous wavelet transform (CWT) dynamically optimized multivariate thresholding method ($WaveL_2E$). Applying this method, we are simultaneously able to separate and capture the signal and noise components while estimating the dynamic noise variance. Our method shows improved results when compared to well-known methods, especially for high-frequency signal-rich time series, typically observed in finance.

Keywords: continuous wavelet transform; data-driven and adaptive thresholding; partial density estimation; integrated squared error; $WaveL_2E$; nonparametric method

1. Introduction

Various time series in economics and finance are formed by non-stationary processes. For instance, Exchange Traded Funds (ETF's) consist of combinations of different equities functioning at different periodicities. In order to understand and identify dynamic associations, it is necessary to explore the characteristics within each individual time series and the changing dynamics across a series. Over the past decade, there has been a significant increase in the use of time-varying spectral representations to analyze dynamic time series behavior in finance. Recently, emphasis has been placed on using wavelets. Rostan and Rostan [1] mention that quantitative analysts (quants) "nibbling" with wavelets are heavily being recruited by major hedge funds. They specifically point out that the interest in wavelet analysis has spiked in recent years because new methods and techniques have come about to analyze physical phenomena, such as seismic and electrical signals, that propagate through time in waveform. These methods are being adapted and applied to economic and financial time series as these signals also propagate through time in waveform. The application of time-varying techniques to augment traditional portfolio management tools by distinguishing across multiple investment horizons or scales is becoming a growing field of interest as well [2–4]. Chaudhuri and Lo [5], who coined the term *spectral portfolio theory*, suggest that the flexibility of the wavelet transform could be used to overcome various difficulties of the Fourier transform for spectral portfolio analysis. Wavelet transforms, when specified correctly, are simple to interpret and can add tremendous value to the

quantitative finance models that financial engineers implement. Wavelets have also proven useful in characterizing two-dimensional fractional Brownian fields, often used in financial studies [6]. For an expanded discussion of wavelet theory and implementation examples in finance, see [7].

The decomposition of time series into their component pieces is frequently applied, but further inference such as optimal dynamic *thresholding* and *forecasting* has only increased in interest over the past few years. A recent paper by Reményi and Vidakovic [8] proposes adaptive wavelet denoising methodology using a fully Bayesian hierarchical model in the complex wavelet domain. They compare their results to two methods proposed by [9]. These authors also compare their methods to the well-known Empirical Bayesian Threshold (EBayesThresh) of [10] and various other superior universal hard and soft thresholding methods such as the SureShrink methods of [11]. Even though these two papers were written ten years apart and are over a decade old, they are still commonly referenced and can be seen as the gold standard in the domain of wavelet thresholding. For a survey of other wavelet denoising or thresholding techniques and their comparisons, see [12]. More recent thresholding (denoising) methods were developed by the authors of [9] who proposed two new complex-valued wavelet techniques, while the authors of [13] proposed a data-driven threshold called the SureBlock, which uses Stein's unbiased risk estimate (SURE) criterion. He et al. [14] designed a new threshold, considering interscale correlation by improving the basic universal hard threshold in order to address the issue of removing too many wavelet coefficients. Each method improved characteristics that are lacking in the existing "gold standard" methods.

In general, hard thresholding reduces all coefficients to zero that do not exceed the threshold. Soft thresholding pushes toward zero any coefficient whose magnitude exceeds the threshold, and zeros the coefficient otherwise. As a comparison to our proposed method, we specifically use the SureShrink methods for the hard and soft thresholds. We also compare our methods to the EBayesThresh. We choose these methods for our comparison as they show superior results to the more recent and well-cited methods pointed out earlier in this paper.

The choice of wavelet family is very important. Most solutions are not based on the continuous wavelet transform (CWT) due to the difficulty of having a tractable solution or an analytical solution [1]. The CWT is calculated by shifting, continuously, a scalable function over a signal and by calculating the correlation between the function and signal, resulting in potential redundancies. However, we mitigate these redundancies as well as address both the tractability and analytical issues by uniquely choosing the Morlet wavelet with the exact specifications for concise interpretation, see Section 2 and [15] for more details.

The novel wavelet thresholding method we propose in this paper builds upon work on minimum distance estimation, initiated by the authors of [16,17]. We propose decomposing each univariate time series into scale-specific wavelet coefficients using a CWT. After decomposition, we apply Scott's multivariate minimum distance partial density estimation (L_2E) to each time increment. In doing so, we separate each time point into a mixture model of signal and noise without having to explicitly assume a model for the signal. Upon analyzing these results, we are not only able to find the optimal signal, but we also estimate the variance of the noise component. In essence, we model the wavelet coefficients as a two-component mixture, where the signal component is unspecified and the noise component is Gaussian; see the Methods section for more details. The L_2E (distance criterion) can be used to exclude groups of outliers; therefore, by estimating the partial density parameters of the noise component, we do not have to make assumptions on the signal component. Ultimately, we are able to distinguish between signal and noise, by taking our signal to be "outliers" while assuming that the noise has a known structure which we can model.

Section 2 describes the $WaveL_2E$ method, starting with a mathematical exposition of wavelets in Section 2.1. We use [18–20] as examples to explain wavelet characteristics and introduce statistical methods that extend the capabilities of wavelets to visualize unexpected

structures in multidimensional data in Section 2.1.1. We then introduce the *WaveL₂E* in Section 2.2 and further identify how our method compares to other well-established wavelet thresholding methods, using simulated time series that have dynamically changing behaviors. In Section 3, we discuss and display the comparative analysis. In this section, we also note that our *WaveL₂E* results yield improved outcomes with added information and interpretations. Before concluding in Section 5 with the strengths of the new methods and a list of the next potential steps and applications, we introduce an example in Section 4. The selected example addresses the investment relationships in water and energy ETF's [19]. In this example, we highlight the post-thresholding improved coherence analysis of the water–energy nexus.

2. Methods

2.1. Wavelets

Wavelets are mathematical functions that decompose temporally localized frequency components, providing information of each component with a resolution matched to its scale. Wavelet analysis in economics and finance has demonstrated promising results. This has led to a resurgent of the literature focused on wavelet transforms and denoising.

2.1.1. CWT

A wavelet is usually derived by scaling and shifting a mother wavelet $\psi(t) \in L^2(\mathbb{R})$ into daughter wavelets $\psi_{a,b}(t) \in L^2(\mathbb{R})$:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right), \tag{1}$$

where $a > 0$ defines the scale and $b \in \mathbb{R}$ defines the shift. Given a time series $x(t) \in L^2(\mathbb{R})$, the continuous wavelet transform (CWT) provides a decomposition of $x(t)$ into time-scale components:

$$W_x(a,b) = \langle x, \psi_{a,b}\rangle = \int_{\mathbb{R}} x(t)\psi_{a,b}^*(t)\,dt, \tag{2}$$

where $*$ denotes the complex conjugate. Usually the CWT coefficients are presented in a power spectrum, where wavelet power is defined as $|W_x(a,b)|^2$ and visualized in the time-scale $\{b, a\}$ half plane with logarithmic scale a-axis (vertical) increasing upward and a linear time-scale b-axis (horizontal) (see Figure 1). More generally, the CWT can be seen as a set of continuous band-pass filters applied to a time series.

Formally, $\psi(t)$ is a mother wavelet if the following three conditions are met: (1) $\int_{\mathbb{R}} \psi(t)dt = 0$, (2) $\int_{\mathbb{R}} \psi^2(t)dt = 1$, and (3) it satisfies an admissibility condition. The admissibility condition ensures the reconstruction of a time series from its wavelet transform:

$$x(t) = \frac{1}{a^2 C_\psi} \int_0^\infty \left[\int_{\mathbb{R}} W_x(a,b)\psi_{a,b}(t)db\right] da, \tag{3}$$

where $C_\psi = \int_{\mathbb{R}} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega$ and $\hat{\psi}(\omega)$ is the Fourier transform of $\psi(t)$ in the CWT. By the convolution theorem, the CWT, $W_x(a,b)$, of a discrete time series $x(t)$ can be approximated by averaging n (the length of the time series) convolutions for each scale using the discrete Fourier transform (DFT) of $x(t)$ and $\hat{\psi}(\omega)$ [21]. Traditionally, the constant (C_ψ) is called the wavelet admissible constant and a wavelet whose admissible constant satisfies $0 < C_\psi < \infty$ is called an admissible wavelet. The most commonly used mother wavelet in finance and economic research is the Morlet wavelet:

$$\Psi^M(t) = \pi^{-\frac{1}{4}} e^{-\frac{1}{2}t^2} e^{i\omega_0 t}, \tag{4}$$

where $e^{i\omega_0 t}$ is the complex sinusoid, $e^{-\frac{1}{2}t^2}$ is a Gaussian envelope with a standard deviation of one, and $\pi^{-\frac{1}{4}}$ normalizes the wavelet to have unit energy, i.e., satisfying the admissibility conditions. As suggested in [22,23], we set the central frequency $\omega_0 = 6$. This value

provides a reasonable trade-off between frequency localization and temporal evolution. Choosing a Morlet wavelet with the above specifications provides us with unique properties described in [15]. First, by choosing a complex wavelet, we have a separation of information between the phase and amplitude. This allows us to examine leading–lagging relationships depicted in Section 4 (see [19] for more details). Furthermore, our method is able to recover the original series without using the inverse transform. However, the Morlet wavelet is also analytic, allowing us to recover the original time series through an inverse transform. Second, ref. [24] describes three different ways to convert scales into frequency. These special frequencies are all equal to the central frequency, ω_0, for the Morlet wavelet. Using the usual definition of the relation between scale and Fourier frequency, defined as $f(s) = \frac{\omega_0}{2\pi s}$, as well as selecting $\omega_0 = 6$, results in $f(s) \approx \frac{1}{s}$, which provides better interpretability. Further, with this parameterization, the standard deviation in both time and frequency are equal, proving an optimal relationship between time and frequency accuracy. The current literature indicates that the choice of Morlet wavelet with $\omega_0 = 6$ proves to be best when using wavelets for feature extraction purposes in economic research. Once again, for a further breakdown of the above summarized specifications, see [15]. A good source for different wavelets and their inherent characteristics can be found in [25].
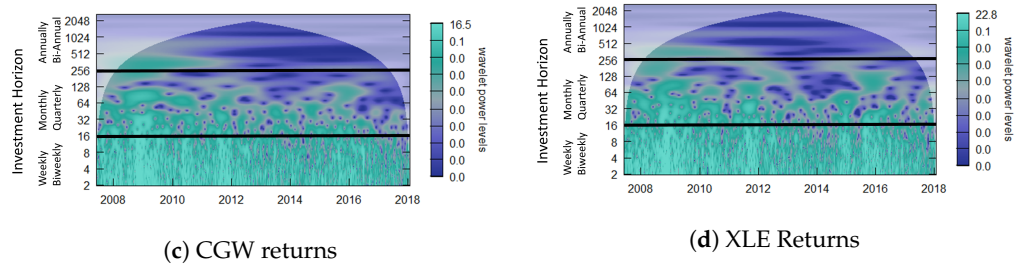

(**a**) CGW Prices

(**b**) XLE prices

(c) CGW returns
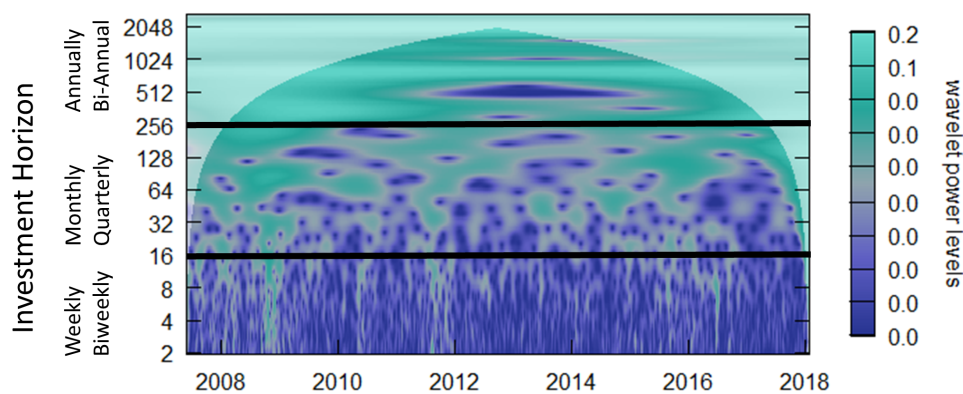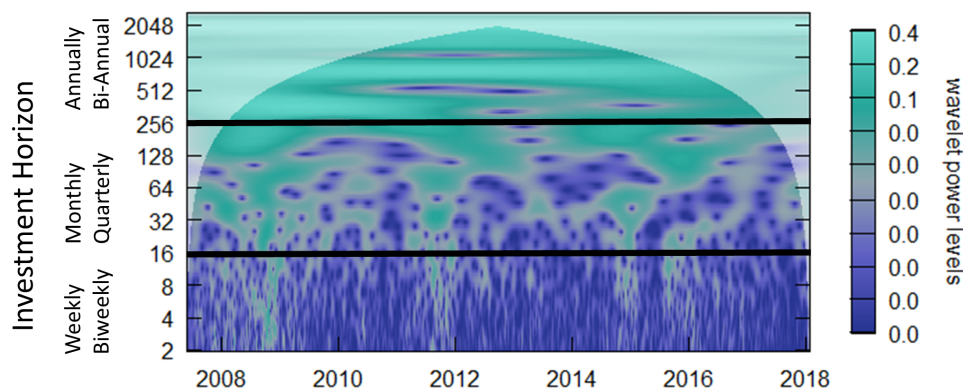
(d) XLE Returns

Figure 1. The continuous wavelet transform (CWT) power spectrum of water (**a**,**c**) and energy (**b**,**d**) commodity ETF prices (**a**,**b**) and returns (**c**,**d**). These plots are visual representations of the power spectrum of each individual series. The investment horizons (vertical axis) are such that the value one through fourteen represents weekly and biweekly investment horizons. Sixty-four days represent a quarterly investment horizon, whereas 250 and above represent annual and larger investment horizons. The horizontal axis indicates the 10 years of data. The white overlay defines the cone of influence.

2.2. WaveL₂E Threshold

In this section, we combine the CWT and the L_2E method. Scott [16] proposed minimizing the Integrated Square Error (ISE) with the aim of minimizing the squared distance between two probability density curves using the L_2 estimation criterion (referred to as the L_2E). In Appendix A, we show that for a general multivariate mixture distribution:

$$h(\boldsymbol{x}) = wf(\boldsymbol{x}|\theta) + (1-w)g(\boldsymbol{x}), \tag{5}$$

where $g(\boldsymbol{x})$ is an unspecified density, and $w \in (0,1)$ is a mixing parameter, then the L_2E for $\theta = (w, \sigma)$ is found by (numerically) minimizing

$$L_2E(w, \sigma) = \frac{w^2}{(2\sigma\sqrt{\pi})^d} - \frac{2w}{n}\sum_{i=1}^{n}\phi(\boldsymbol{x}_i|\mathbf{0}, \sigma^2 I_d). \tag{6}$$

To formulate the dynamic selection of the coefficients of the multivariate continuous wavelet transform using the L_2E, we first describe our time series $\boldsymbol{y}$ with the simple signal plus noise model:

$$\boldsymbol{y} = q(\boldsymbol{t}) + \boldsymbol{\epsilon}, \tag{7}$$

where $\boldsymbol{y} = (y_1, \ldots, y_n)$ denotes the observations at our n time points, the signal is defined by $q(\boldsymbol{t}) = (q(t_1), \ldots, q(t_n))$, $q(t) \in L^2(\mathbb{R})$, and $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n) \overset{iid}{\sim} N(\mathbf{0}, \sigma^2 I_n)$, with unknown variance σ^2. Taking the continuous wavelet transform (W) of the time series, we can further express the additive noise model as:

$$W\boldsymbol{y} = Wq(\boldsymbol{t}) + W\boldsymbol{\epsilon}. \tag{8}$$

Framing the additive noise model in Equation (8) as a general multivariate mixture distribution, we equate $W\boldsymbol{y}$ to $h(\boldsymbol{x})$ of Equation (5), with the strong signal components of the wavelet coefficients, $Wq(\boldsymbol{t})$, assigned to $(1-w)g(\boldsymbol{x})$. The remaining wavelet coefficients are categorized as noise, and are combined with $W\boldsymbol{\epsilon}$, which is equated to $wf(\boldsymbol{x}|\theta)$ of Equation (5). With this specification, minimizing (6) allows us to estimate the noise component parameters without making assumptions of the number or distribution of the 'significant' wavelet coefficients embodied in $g(\boldsymbol{x})$. Since we assume i.i.d Gaussian noise with variance σ^2, we also assume that the pure noise wavelet coefficients are i.i.d with covariance matrix $\sigma^2 I_d$. Thus, the wavelet coefficients with the covariance matrix, $\sigma^2 I_d$, are assumed to be independent. This assumption of independence means that we only have two parameters to estimate, namely σ and w, regardless of the dimension.

After minimizing the criteria and recovering our parameters $(\hat{\sigma}, \hat{w})$, we apply two thresholding methods to the wavelet coefficients based on these estimates. We apply the first of these thresholds to the distribution of the squared wavelet coefficients, referred to

as the $WaveL_2E$. The reason we apply the minimization to the squared wavelet coefficients is due to our assumptions that we have standard Gaussian noise, the square of the wavelet coefficients representing the Gaussian noise is therefore assumed to be distributed χ^2. The second threshold is determined using the 95% χ^2 critical value of the the squared wavelet coefficients distribution. This threshold is referred to as the $WaveL_2E_{\chi^2}$. The squared wavelet coefficients are also the power of the signal; in Section 2, we present these in a CWT power spectrum. After applying the threshold, the power spectrum can be used to visualize the applied thresholds. For a simple simulated demonstration of these thresholding methods, see Figure 2. This figure visualizes an additive noise model, with three stationary signals, before and after thresholding, using both the $WaveL_2E$ and the $WaveL_2E_{\chi^2}$ thresholding methods. We are able to recover the pure signal for the three periods in the original additive noise model by removing the induced noise across all wavelet coefficients. These methods work well for signals that are constant or stationary. However, when there are inter-scale dependencies, the static version of these thresholds do not necessarily extract all the signals. Therefore, we introduce a dynamic version of the static L_2E criterion.
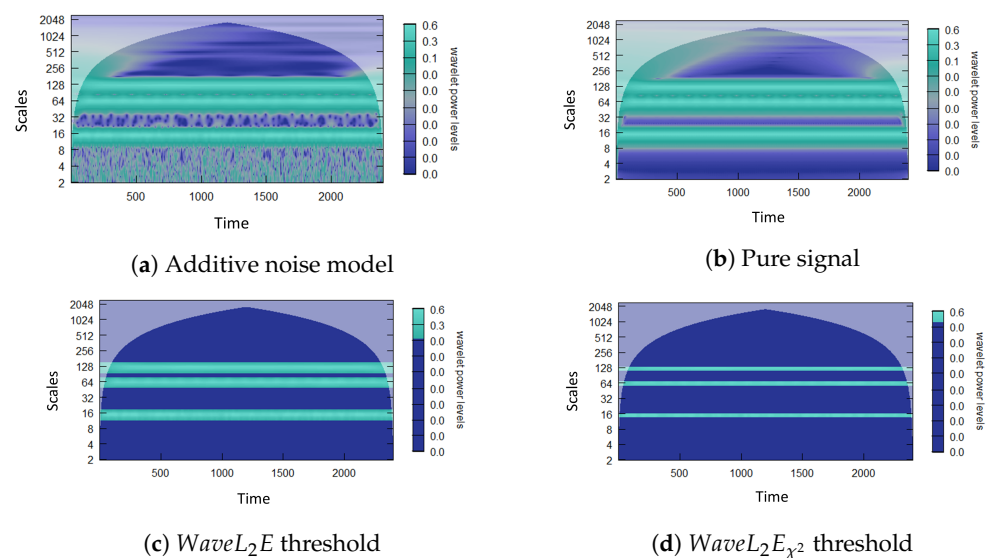

(**a**) Additive noise model

(**b**) Pure signal

(**c**) $WaveL_2E$ threshold

(**d**) $WaveL_2E_{\chi^2}$ threshold

Figure 2. The CWT power spectrum (**a**) of a simple additive noise model with three stationary signals. Biweekly (15 days), quarterly (64 days), and biannually (125 days). Hence, we have that: $y_t = sin(\frac{2\pi t}{15}) + sin(\frac{2\pi t}{64}) + sin(\frac{2\pi t}{125}) + \epsilon$, where $\epsilon \sim N(0, 0.1)$ and $t = (1, \ldots, 2400)$. Also, this figure shows the CWT power spectrum (**b**), which is the pure signal without noise, the CWT power spectrum (**c**) after the $WaveL_2E$ threshold, and the CWT power spectrum (**d**) after the $WaveL_2E_{\chi^2}$ threshold. For the minimization process, we use constrained optimization using PORT routines with $0 \le \sigma < \infty$ and $0 \le w < 1$ [26].

2.3. Dynamic WaveL$_2$E Threshold

Minimum distance estimators, like the L_2E, identify the portion of the data that matches the model and separates the data that do not match. Scott [16] showed that the inefficiencies (asymptotically) of the parameters estimated using the maximum likelihood estimator versus the minimization of the ISE or the L_2E criterion are roughly that of the the mean versus the median in a general statistical analysis. The L_2E criterion specified in [16] for mixture distributions is therefore very good at identifying groups of outliers in large datasets. This minimum distance estimator is also robust, without requiring additional specifications (meta-parameters) that are needed in other likelihood estimators, making it an ideal candidate for thresholding.

In order to evaluate the changing weight, w_t, and the changing noise variance, σ_t, across time, the implementation of the L_2E minimization needs to be adapted to optimally estimate $(\hat{w}_t, \hat{\sigma}_t)$ at each time-localized observation.

Therefore, by adapting Equation (6) for time-dependency, we change the L_2E criterion to obtain unique estimates $\hat{\sigma}_t$ and $\hat{w}_t$. The result is the dynamic optimization by minimizing the following L_2E criterion:

$$L_2E(w_t, \sigma_t) = \frac{w_t^2}{(2\sigma_t\sqrt{\pi})^d} - \frac{2w_t}{n}\sum_{i=1}^{n}\phi(\mathbf{x}_i|\mathbf{0}, \sigma_t^2 I_d). \tag{9}$$

We apply the same two thresholding methods mentioned previously, based on these estimates. The only difference in the adjusted implementation is that we apply these thresholds to the distribution of the squared wavelet coefficients at each time point. Figure 3 is a simple demonstration of the steps followed to implement the $WaveL_2E$ method and the resulting estimates recovered. The steps for implementation are as follows (see our R Package referenced in Data Availability Statement for more detailed implementation instructions):

- Calculate the wavelet transform of the observed time series and recover the squared wavelet coefficients.
- Apply the $WaveL_2E$ and $WaveL_2E_{\chi^2}$ at each time point, solving for the estimates $(\hat{w}_t, \hat{\sigma}_t)$ of the L_2E criterion.
- Threshold the wavelet coefficients based on the estimates.
- Calculate the inverse transform of the denoised time series or solve for the pure signal using the recovered estimates.
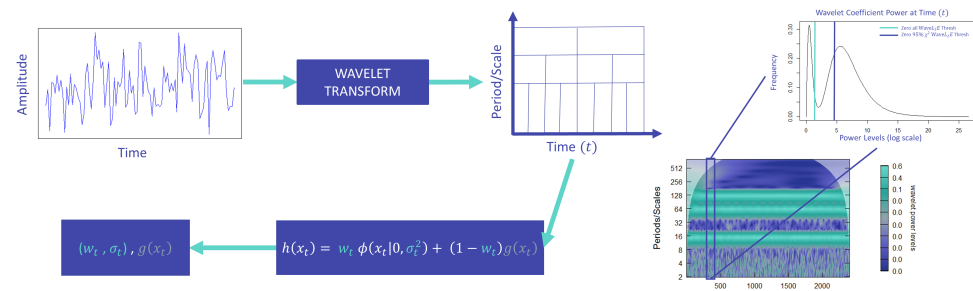

Figure 3. Demonstration of the process applied to a time series. Starting with the observed time series, we apply the Morlet wavelet, a continuous wavelet transform (CWT), and then after decomposition, we minimize the L_2E criterion. This minimization provides us with the optimal estimates for σ_t and w_t. After estimation, we can remove the noise component and recover the signal component of the original observed series.

Further extending and generalizing Equation (9), we include intra-scale dependence (within scale) by evaluation blocks (windows) of time across scales. We perform this by enabling block sizes of length h. The first block is labeled as block zero, with estimates $(\hat{w}_0, \hat{\sigma}_0)$. The $WaveL_2E$ and $WaveL_2E_{\chi^2}$ are applied to each of these blocks, across the scales; see Figure 3. For example, if we identify a time window (block) of 22 days, an average fiscal month, we would have estimates for $(\hat{w}_t, \hat{\sigma}_t)$, at $t = h$, where $(h = 0, 22, 44, 66, \ldots, T)$, with T being the length of the series. When $h = T$, we have the static L_2E criterion described in Equation (6). More specifically, the generalization of the L_2E criterion then becomes:

$$L_2E(w_h, \sigma_h) = \frac{w_h^2}{(2\sigma_h\sqrt{\pi})^d} - \frac{2w_h}{n}\sum_{i=1}^{n}\phi(\mathbf{x}_i|\mathbf{0}, \sigma_h^2 I_d). \tag{10}$$

3. Results and Comparative Analysis

3.1. Inter-Scale and Intra-Scale

In this section, our main goal is to show comparative results of our methods against the EBayesthresh and SureShrink referenced in Section 1. We refer to these methods as the gold standard for wavelet thresholding. To further understand our findings, we examine the percentage of significance area (PSA) and the percentage of total volume (PTV) given in [27]. We use these two measures to estimate the statistically significant area maintained

after our threshold implementation and the portion of volume in the CWT power spectrum, respectively. Thus, to simplify, PSA measures the percentage of the wavelet coefficients removed by our method and PTV measures the resulting statistical significance captured by our method.

Current thresholding methods have two limiting assumptions. First, most methods assume wavelet coefficients, representative of the signal components, which follow a specified parametric model. Unlike Bayesian methods with hyperparameters that need to be tuned, our multivariate wavelet thresholding method is adaptive and data-driven, so no specifications are needed. Second, traditional wavelet thresholding methods assume independence of all wavelet coefficients. However, wavelet correlation theory suggests that signal coefficients are not independent and only the wavelet coefficients of noise are uncorrelated (see [28] for more details). These methods generally assume that the smallest level of wavelet coefficients are likely to contain pure noise. Both SureShrink and EbayesThresh methods use adaptions of the median absolute deviation (MAD) of the finest level of coefficients to estimate the noise variance of the signal. However, these MAD estimates of the smallest scale wavelet coefficients overestimate the variance in high-frequency data when compared to our noise variance estimate. Overestimation results in too large a threshold. Visualizations of these comparative results for the different estimates of the noise variance can be found in [17]. Therefore, noise variance has to be estimated in a data-driven, dynamic manner as we suggest in our proposed wavelet thresholding method, the $WaveL_2E$.

The assumption that all levels of wavelet coefficients are independent leads to thresholding methods that do not capture both the inter-scale (between) and intra-scale (within) dependencies. Figure 4 summarizes these differences by pointing out how the $WaveL_2E$ thresholding method mitigates these within and between scale dependencies. Figure 3 provides a simple demonstration of how the $WaveL_2E$ method incorporates between scale dependencies, by evaluating the additive noise model at each increment, and within scale dependency, by allowing blocks of time to be evaluated simultaneously. Our method has the flexibility to separate the noise and signal components by minimizing the L_2E criterion across scales and for specific time windows. In the next subsection, we will show how this flexibility results in better, simplified thresholding and the interpretation of commonly simulated signals. This subsection then leads to an example of how our method can be implemented to analyze real- world applications.
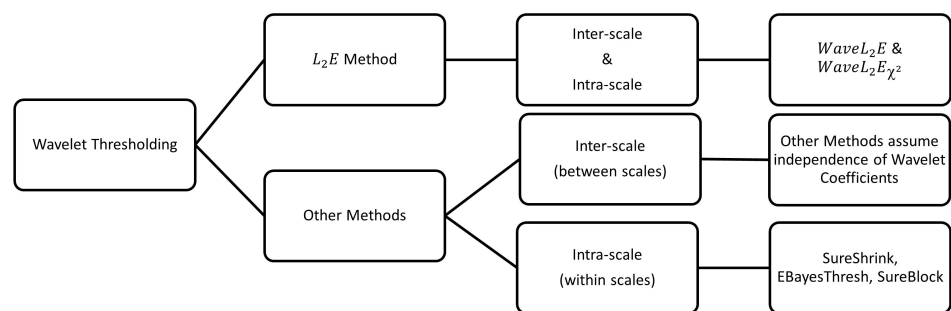

Figure 4. Summary of the difference between the wavelet L_2E thresholding method we introduce in this paper and the other gold standard methods. The fundamental difference is that our novel method incorporates both inter-scale and intra-scale dependencies with no limitations on the estimates of the noise variance or parametric specifications.

3.2. Comparative Analysis

We start our comparative analysis by analyzing the root-mean-squared error (RMSE) of both the gold standard methods and our method for the simulated periodic time series in Figure 2. Figure 5 extends the previous static analysis and includes the dynamic analysis from Equation (9) for both the $WaveL_2E$ and $WaveL_2E_{\chi^2}$ thresholds. The static implementation removed a significant amount of the coefficients, whereas the dynamic implementation

was more selective. This property becomes more relevant when the signals are not as simple as what was simulated in this example. In Table 1, we point out the percentage total volume (PTV) and the percentage significance area (PSA). These two measures refer to the percentage of the total wavelet coefficients that represent the estimated true signal after threshold implementation and the percentage of the total statistical significance represented by these thresholded wavelet coefficients. The static implementation thresholds about 80% of the coefficients for the $WaveL_2E$ and about 92% of the coefficients for the $WaveL_2E_{\chi^2}$. The remaining wavelet coefficients represent about 94% and about 55% of the statistical significance for the respective thresholding methods. On the other hand, the dynamic implementation thresholded about 42% and about 50% of the wavelet coefficients for the respective methods, but this implementation was able to maintain about 99% of the statistical significant wavelet coefficients for both methods.

Table 1. The percentage total volume (PTV) and percentage significance area (PSA) of the static and dynamic $WaveL_2E$ and $WaveL_2E_{\chi^2}$ thresholding methods, for the simulated series defined in Figure 2.

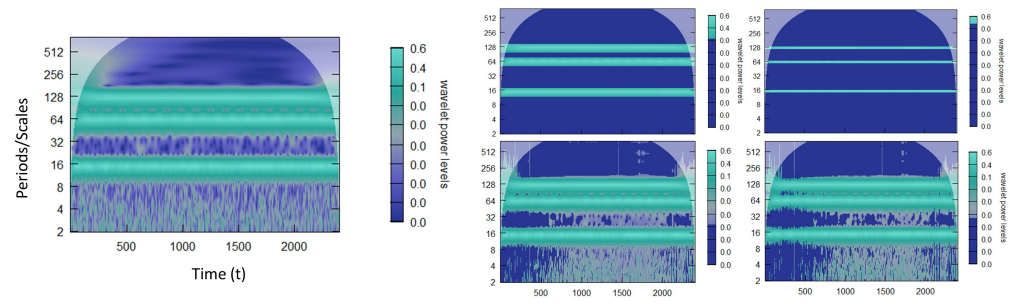| Analysis	$WaveL_2E$	$WaveL_2E_{\chi^2}$
Static PTV	20.48%	7.80%
Static PSA	94.10%	54.59%
Dynamic PTV	57.77%	50.79%
Dynamic PSA	99.84%	99.64%

Figure 5. Two different comparative analyses: The first plot (left) is the base CWT power spectrum for the signal $y_t = sin(\frac{2\pi t}{15}) + sin(\frac{2\pi t}{64}) + sin(\frac{2\pi t}{125}) + \epsilon_t$, where $\epsilon_t \sim N(0, 0.1)$ and $t = (1, \ldots, 2400)$. The top analyses is the $WaveL_2E$ (left) and $WaveL_2E_{\chi^2}$ (right) thresholds from Equation (6) and the bottom analysis is the dynamic $WaveL_2E$ (**left**) and $WaveL_2E_{\chi^2}$ (**right**) thresholds from Equation (9).

In Table 2, we include the root-mean-squared error (RMSE) for the wavelet SureShrink (WavShrink) universal hard threshold (RMSE: 0.1120), the WavShrink universal soft threshold (RMSE: 0.2205), and the empirical Bayes threshold (RMSE: 0.0734). Our static implementation for the $WaveL_2E$ threshold has an RMSE of 0.0775 and for the $WaveL_2E_{\chi^2}$ threshold has an RMSE of 0.1271. The generalization and implementation of the dynamic method results in an RMSE of 0.0931 and an RMSE of 0.0963 for the $WaveL_2E$ and $WaveL_2E_{\chi^2}$ thresholds, respectively.

Table 2. RMSE comparison of our method using the static L_2E criterion and the dynamic L_2E criterion to the gold standard (GS) methods. The RMSE of the WaveShrink (WShrink) hard (H) and WShrink soft (S) thresholds as well as the EBayes threshold, for the simulated series defined in Figure 2.

Analysis	WShrink (H)	WShrink (S)	EBayes	$WaveL_2E$	$WaveL_2E_{\chi^2}$
Static RMSE	0.1120	0.2205	0.0734	0.0775	0.1271
Dynamic RMSE	0.1120	0.2205	0.0734	0.0931	0.0963

As we pointed out earlier in this section and in Figure 4, most traditional methods assume independence between scales (inter-scale). When we only have stationary signals,

as we see in Figure 5, the EBayesThresh method slightly outperforms the new method we introduce in this paper. However, the moment high-frequency data and inter-scale dependency are introduced, see Figure 6 and Table 3, our method outperforms the other methods when implementing the static L_2E criterion, and the results are equivalent in RMSE when applying the dynamic L_2E criterion. We also see similar results for the PTV and PSA in Table 4, where the dynamic implementation maintains more of the statistical significant wavelet coefficients. The signal we identified to analyze with the aforementioned specifications consists of a ten-year cycle (as the first periodic component) and the second periodic component is a one-year cycle that changes to a three-year cycle between 10 and 20 years and a five-year cycle between 50 and 60 years with variance that changes five times, i.e., every 20 years. The simulated series represents 100 years of monthly values, or 1200 observations. Specifically,

$$y_t = cos\left(\frac{2\pi t}{p_1 12}\right) + cos\left(\frac{2\pi t}{p_2 12}\right) + \epsilon_t, \quad t = (1, 2, \ldots, 1200), \tag{11}$$

where $\epsilon_t \sim N(0, p_3)$ with $p_3 = (0.5, 0.05, 0.25, 0.15, 0.05)$ with $p_1 = 10$, $p_2 = 3$ when $10 \leq \frac{t}{12} \leq 20$, $p_2 = 5$ when $50 \leq \frac{t}{12} \leq 60$, and $p_2 = 1$ otherwise. Figure 6 shows two different analyses, the first using the static L_2E criterion and the second using the dynamic L_2E criterion.
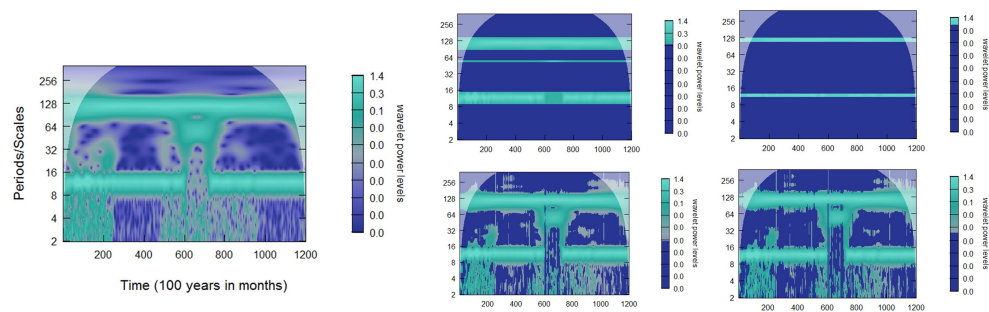

Figure 6. Two different comparative analyses. The first plot (left) is the base CWT power spectrum for the signal from Equation (11). The top analyses is the $WaveL_2E$ and $WaveL_2E_{\chi^2}$ thresholds from Equation (6) and the bottom analysis is the dynamic $WaveL_2E$ and $WaveL_2E_{\chi^2}$ thresholds from Equation (9).

Table 3. RMSE comparison of our method using the static L_2E criterion and the dynamic L_2E criterion to the gold standard (GS) methods. The RMSE of the WaveShrink (WShrink) hard (H) and WShrink soft (S) thresholds as well as the EBayes threshold, for the simulated series from model (11).

Analysis	WShrink (H)	WShrink (S)	EBayes	$WaveL_2E$	$WaveL_2E_{\chi^2}$
Static RMSE	0.2740	0.2939	0.2430	0.1844	0.2682
Dynamic RMSE	0.2740	0.2939	0.2430	0.2535	0.2547

Table 4. The percentage total volume (PTV) and percentage significance area (PSA) of the static and dynamic $WaveL_2E$ and $WaveL_2E_{\chi^2}$ thresholding methods, for the simulated series from model 11.

Analysis	$WaveL_2E$	$WaveL_2E_{\chi^2}$
Static PTV	20.54%	7.03%
Static PSA	87.44%	55.81%
Dynamic PTV	75.17%	70.81%
Dynamic PSA	99.93%	99.86%

Introducing more inter-scale dependency, we analyze eight different signals by implementing the dynamic L_2E criterion. We also add to the analysis the empirical $WaveL_2E$ calculation ($EWaveL_2E$). After recovering both $(\hat{w}_t, \hat{\sigma}_t)$ and since we are assuming independence, we can recover the denoised signal from the original equation by solving for the estimated true signal, $\hat{g}(\boldsymbol{x})$:

$$\hat{g}(x) = \frac{h(x) - \hat{w}_h \phi(x|\mathbf{0}, \hat{\sigma}_h^2 I_d)}{(1 - \hat{w}_h)}. \tag{12}$$

Equation (12) provides us with the flexibility to create a potential forecasting model; however, this task is left for future research. In Table 5, the last column contains the RMSE of an adjusted $EWaveL_2E$ for all the signals. We made an adjustment to ensure that we can include the results of various different signals. We use the median absolute deviation (MAD) of each of the estimates $(\hat{\sigma}_h, \hat{w}_h)$ to recover the estimated $\hat{g}(x)$. The time series and CWT power spectrum of each of these signals are visualized in Figure 7. These spectra clearly show the inter-scale dependencies.

Tables 5 and 6 are the RMSE results of the eight different signals with two different signal-to-noise ratio's (SNR), 2 and 5, respectively. As soon as there is more inter-scale dependency and the signals are not only noise, our model outperforms. Another interesting observation is the difference between the empirical calculation and the inverse wavelet transform. It seems that the inversion does yield better results. However, the adjusted empirical results ($EWaveL_2E$) compared to the $WaveL_2E$ and $WaveL_2E_{\chi^2}$ for some of the signals are very similar. To show the value in our new method, we expand the example of the water–energy nexus in the next Section (see [19]).

Table 5. Using a signal-to-noise ratio (SNR) of two, we analyze the RMSE comparison of our method using the dynamic L_2E criterion to the gold standard (GS) methods for eight different signals: the RMSE of the WaveShrink (WShrink) hard (H) and WShrink soft (S) thresholds, as well as the EBayes threshold. We identify the smallest RMSE with a star (*).

Signal	WShrink (H)	WShrink (S)	EBayes	$WaveL_2E$	$WaveL_2E_{\chi^2}$	$EWaveL_2E$
hisine	0.7069	0.7069	0.7116	0.3443	0.3425 *	0.3832
losine	0.1457	0.2813	0.1301 *	0.3196	0.3201	0.3642
linchirp	0.3145	0.4520	0.2203 *	0.3296	0.3257	0.4026
twochirp	0.9662	0.9728	0.9315	0.4897	0.4835 *	0.5543
quadchirp	0.5694	0.6232	0.4812	0.3416 *	0.3543	0.4018
mishmash1	1.1166	1.1635	1.0778	0.6493	0.6478 *	0.7096
mishmash2	1.3406	1.3620	1.2559	0.8034 *	0.9486	0.9420
mishmash3	0.9354	1.0594	0.6603	0.5919	0.5899 *	0.6702

Table 6. Using a signal-to-noise ratio (SNR) of five, we analyze the RMSE comparison of our method using the dynamic L_2E criterion to the gold standard (GS) methods for eight different signals: the RMSE of the WaveShrink (WShrink) hard (H) and WShrink soft (S) thresholds, as well as the EBayes threshold. We identify the smallest RMSE with a star (*).

Signal	WShrink (H)	WShrink (S)	EBayes	$WaveL_2E$	$WaveL_2E_{\chi^2}$	$EWaveL_2E$
hisine	0.7068	0.7068	0.7076	0.1371	0.1359 *	0.1619
losine	0.0810	0.1339	0.0559 *	0.1248	0.1207	0.1830
linchirp	0.1378	0.2498	0.0925 *	0.1295	0.1929	0.1804
twochirp	0.9668	0.9727	0.9127	0.2145 *	0.2389	0.2901
quadchirp	0.3507	0.5088	0.1744	0.1541 *	0.2338	0.2791
mishmash1	1.1156	1.1637	1.0772	0.2693 *	0.3241	0.3078
mishmash2	1.3386	1.3618	1.2360	0.5202 *	0.6562	0.5630
mishmash3	0.7803	0.9463	0.4679	0.2525 *	0.2567	0.3467
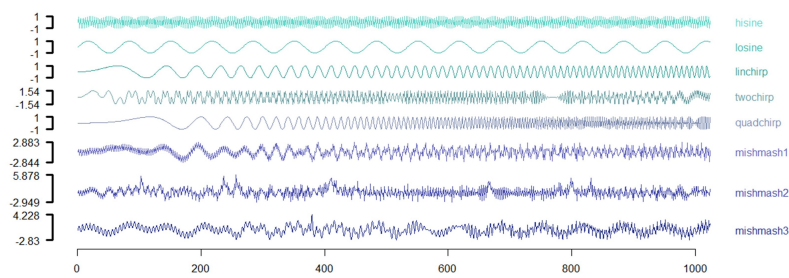
(**a**) Signal time series plots.
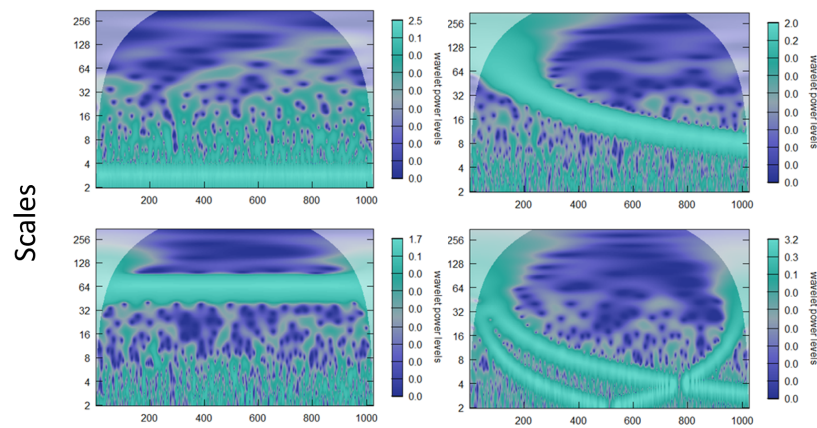
(**b**) CWT power spectrum of hisine, losine, linchirp, and twochirp.
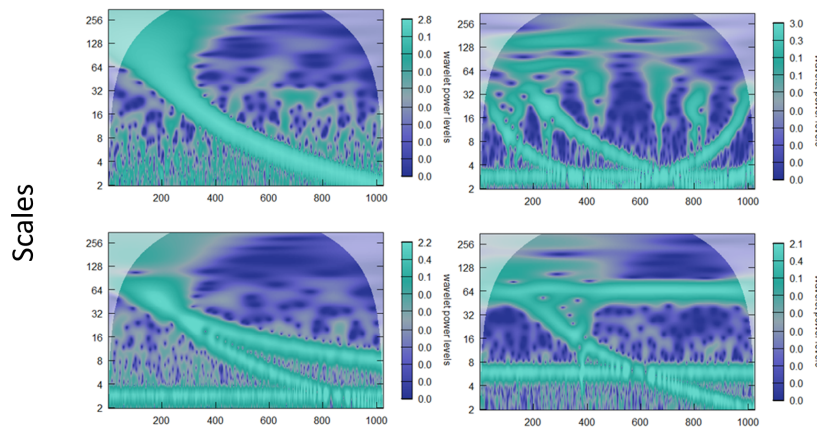
(**c**) CWT power spectrum of quadchirp and mishmash 1 through 3.

Figure 7. Comparative analysis of eight traditional time series depicted in the top panel (7a) before a wavelet transform. The series are labeled hisine, losine, linchirp, twochirp, quadchirp, mishmash1, mishmash2 and mismash3. The CWT for hisine, losine, linchirp and two chirp is given in panel 7b reading from left to right and top to bottom, whereas the CWT for the remaining four series is provided in panel 7c, namely quadchirp and mishmash1 through mishmash3.

4. Motivating Example: Water–Energy Nexus

Raath and Ensor [19] quantify the dynamic relationship between energy and water commodities by applying wavelet techniques to better understand the dynamic relationship that exists in the water–energy nexus. Using daily water and energy commodity ETF price data from 2007 to 2017, they evaluate the respective wavelet transforms of each time series. We use this water–energy nexus quantification as a motivating example in explaining each component of our novel thresholding method.

After thresholding both the water (CGW ETF) and energy (XLE ETF) prices and returns, as seen in Figure 1, we further analyze by comparing the wavelet-squared coherence (WSC) before and after thresholding, using the *WaveL₂E*. The WSC analyzes local linear correlations in regions of statistical significant co-movement and can be used to define investment horizon-specific behavior. In Figure 8, the top two plots represent the before (left) and after (right) thresholding results. It can clearly be seen that there are less statistically significant co-moving components in the WSC plot after thresholding (top right). We also note that during the financial crisis, from 2008 to 2012, there were no statistically significant behaviors at the weekly and biweekly investment horizons (scale); the relationship was only noise.
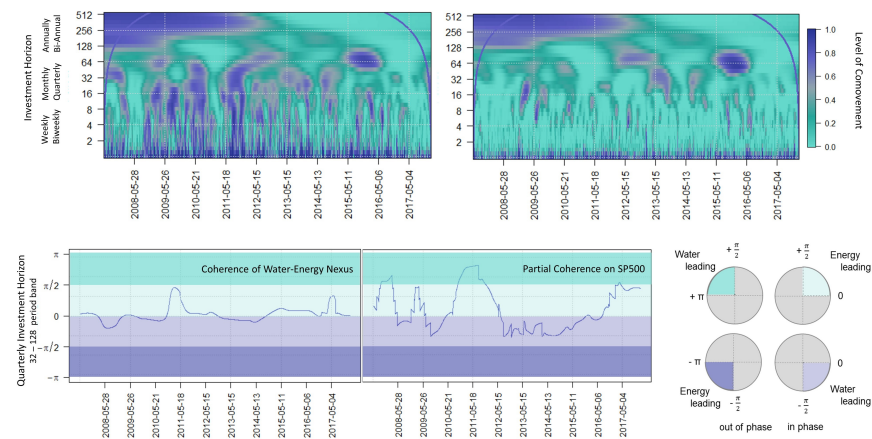
Figure 8. Analyzing the wavelet-squared coherence (WSC) of the water–energy nexus before (**top left**) and after (**top right**) *WaveL₂E* thresholding. We evaluate both the WSC (**bottom left**) and the partial coherence (**bottom middle**) results of the quarterly investment horizon by analyzing the leading–lagging relationship (**bottom right**) of the nexus.

The next set of plots point out the partial wavelet coherence conditional on the market (SPY ETF); for more details, see [15]. We wanted to eliminate the market effects on the water–energy nexus and then evaluate the leading–lagging (bottom right) relationship for the quarterly investment horizon (bottom left and center). The quarterly investment horizon was identified because of the statistically significant co-movement during the 2014–2016 oil glut (top right). The phase difference (bottom left) of our thresholded series does not seem to have any significant leading–lagging relationship. The light purple and darker green indicate that water is leading and the darker purple and lighter green indicate that energy is leading (bottom right). However, as soon as we implement the partial wavelet coherence on our thresholded series, we see significant leading–lagging relationships. Raath and Ensor [19] pointed out the significance of water-leading energy during the 2014–2016 oil glut, which is very clearly confirmed here using our *WaveL₂E* thresholding method.

We point out in Figure 9 the plots of the estimates for $(\hat{\sigma}_h, \hat{w}_h)$, where $h = 22$, a fiscal month (top) on the XLE ETF. We added these here to show the flexibility of our method and to visually demonstrate that a static weight or variance would not work for the water–energy nexus.

In this figure, we also add the reconstructed price series and return series using our *WaveL₂E* threshold method. In the return plot (bottom right), we also include the *EWaveL2E* estimate of $\hat{g}(x)$ from Equation (12) using only the $h = 22$ estimates $(\hat{\sigma}_h, \hat{w}_h)$.
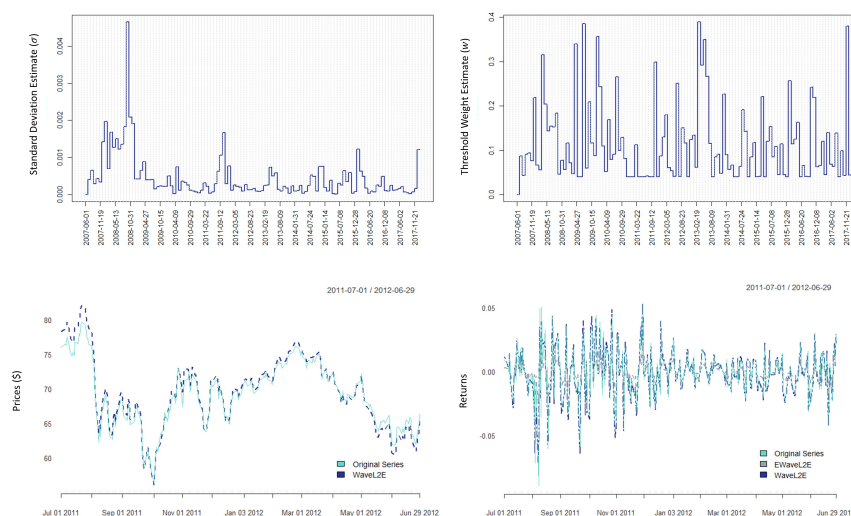
Figure 9. Analyzing and recovering the fiscal month, on the 22nd day; estimates for the $WaveL_2E$ and visualizing the results of both the $WaveL_2E$ and the $EWaveL_2E$ after inversion and recovery of the estimated true time series. Method executed on the XLE ETF.

5. Conclusions

In this paper, we present a method that denoises non-stationary time series by dynamic multivariate complex wavelet thresholding. We demonstrate that our method performs best among a broad class of methods for the type of series seen in finance. Our method performs especially well in environments when there are inter-scale dependencies. Further, we are able to capture not only the signal, but also an estimate of the variance for the additive noise of a time series. Other objectives met in this manuscript include: (1) choosing a continuous wavelet transform to increase interpretability while incorporating and adjusting the L_2E criterion to optimize a data-driven threshold; (2) increasing the dynamic adaptability by generalizing our $WaveL_2E$ method; (3) describing how to implement these thresholds in practice; and (4) developing a user-friendly R package, which implements the thresholds introduced in this paper.

Our methodology is applied to a financial series, addressing the water–energy nexus [19]. By denoising the price series, using our methodology, we form a deeper understanding of the relationship between the price series representing water and that representing energy. These results would be difficult to discern without implementing our threshold. We confirmed that the water–energy nexus depends on general market behavior. When the market behavior is removed, we find that water prices lead energy prices at the quarterly investment horizon during the 2014–2016 oil glut.

An added feature of our thresholding method is the estimation of the variance of the additive noise at specific localized windows of time. These estimates can be used to forecast the volatility series, work which will be further explored in future papers. Future investigations also include revisiting wavelet estimation for temporal random fields, as in [6].

Data Availability Statement: R-package for *WaveL2E* routine: The R-package *CoFESWave* contains code used to perform the diagnostic methods described in the article. Data and software used in this article may be obtained from https://github.com/kcraath/CoFESWaveResearch (accessed on 9 November 2023) and https://github.com/kcraath/CoFESWave (accessed on 9 November 2023), respectively.

Conflicts of Interest: Author Kim C. Raath was employed by the company Imperative Global and author Alena Crivello was employed by the company Chevron Corporation. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CGW	Global Water Fund
CWT	Continuous Wavelet Transform
ETF	Exchange Traded Fund
GS	Gold Standard
ISE	Integrated Squared Error
L_2E	Integrated Squared Error Estimation Criterion
PSA	Percentage of Significant Area
PTV	Percentage of Total Volume
RMSE	Root-Mean-Squared Error
SNR	Signal-to-Noise Ratio
WSC	Wavelet-Squared Coherence
XLE	Energy Select Sector

Appendix A

Scott [16] proposed using the Integrated Square Error (ISE) criterion to minimize the squared distance between two probability density curves using the L_2 estimation criterion (referred to as the L_2E). The ISE can be expressed in terms of two curves, namely the assumed parametric density $f(x|\theta)$ and the true but unknown density, $h(x)$, from which the i.i.d. sample is taken, and is given by:

$$ISE(f,h) \equiv \int_{-\infty}^{\infty} \left[f(x|\theta) - h(x) \right]^2 dx. \tag{A1}$$

In this formulation, we allow $f(x|\theta)$ and $h(x)$ to come from different parametric families. In fact, this is the general expectation.

The L_2E parameter estimate, $\hat{\theta}$, is found by minimizing an unbiased estimate of a shifted version of $ISE(f,h)$. Specifically, we consider the integrals in the curly brackets

$$
\begin{aligned}
\hat{\theta} &= \underset{\theta}{\operatorname{argmin}} \left\{ \int \left[f(x|\theta) - h(x) \right]^2 dx \right\}, \\
&= \underset{\theta}{\operatorname{argmin}} \left\{ \int \left[f(x|\theta)^2 - 2f(x|\theta)h(x) + h(x)^2 \right] dx \right\}, \\
&= \underset{\theta}{\operatorname{argmin}} \left\{ \int f(x|\theta)^2 dx - 2\, Ef(X|\theta) \right\},
\end{aligned}
\tag{A2}
$$

where $X \sim h(x)$ and the constant term $\int h(x)^2 dx$ are dropped. The first term is computable and does need not be estimated stochastically. For the second term, we substitute the unbiased average based on our random sample to estimate $Ef(X|\theta)$ as proposed by [16]. The resulting L_2E criterion is:

$$\hat{\theta}_{L_2E} = \underset{\theta}{\operatorname{argmin}} \left\{ \int f(x|\theta)^2 dx - \frac{2}{n} \sum_{i=1}^{n} f(x_i|\theta) \right\}, \tag{A3}$$

which may be obtained using nonlinear optimization software.

If we assume that the univariate density $f(x|\theta)$ is Gaussian as in the additive noise model, then $h(x)$ in Equation (5) can be written as:

$$h(x) = w\phi(x|0, \sigma^2) + (1-w)g(x). \tag{A4}$$

By replacing $f(x|\theta)$ with $w\phi(x|0, \sigma^2)$, the partial L_2E criterion becomes:

$$L_2E(w, \sigma) = \frac{w^2}{2\sigma\sqrt{\pi}} - \frac{2w}{n} \sum_{i=1}^{n} \phi(x_i|0, \sigma^2). \tag{A5}$$

Here, we have used the result for Gaussian densities in [29]:

$$\int_{-\infty}^{\infty} \phi(x|\mu_1, \sigma_1^2)\, \phi(x|\mu_2, \sigma_2^2)\, dx = \phi(0|\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2),$$

which also holds for multivariate densities with mean vectors $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ and covariance matrices Σ_1 and Σ_2. Hence, Equation (A3) becomes

$$\begin{aligned} \hat{\theta}_{L_2E} &= \underset{\theta}{\operatorname{argmin}} \left\{ \int \left[w\phi(x|0, \sigma^2) \right]^2 dx - \frac{2}{n} \sum_{i=1}^{n} \left[w\phi(x_i|0, \sigma^2) \right] \right\} \\ &= \underset{\theta}{\operatorname{argmin}} \left\{ w^2 \phi(0|0, 2\sigma^2) dx - \frac{2w}{n} \sum_{i=1}^{n} \phi(x_i|0, \sigma^2) \right\} \\ &= \underset{\theta}{\operatorname{argmin}} \left\{ \frac{w^2}{2\sigma\sqrt{\pi}} - \frac{2w}{n} \sum_{i=1}^{n} \phi(x_i|0, \sigma^2) \right\}. \end{aligned}$$

The resulting L_2E criterion specified in Equation (A5) can be generalized to multivariate mixtures. Assume that $h(x)$ in Equation (5) can be written as:

$$h(\boldsymbol{x}) = w\phi(\boldsymbol{x}|\mathbf{0}, \sigma^2 I_d) + (1-w)g(\boldsymbol{x}); \tag{A6}$$

then

$$L_2E(w, \sigma) = \frac{w^2}{(2\sigma\sqrt{\pi})^d} - \frac{2w}{n} \sum_{i=1}^{n} \phi(\boldsymbol{x}_i|\mathbf{0}, \sigma^2 I_d), \tag{A7}$$

which demonstrates Equation (6).

References

1. Rostan, P.; Rostan, A. The versatility of spectrum analysis for forecasting financial time series. *J. Forecast.* **2018**, *37*, 327–339. [CrossRef]
2. Wang, G.J.; Xie, C.; Chen, S. Multiscale correlation networks analysis of the US stock market: A wavelet analysis. *J. Econ. Interact. Coord.* **2017**, *12*, 561–594. [CrossRef]
3. Ftiti, Z.; Jawadi, F.; Louhichi, W. Modelling the relationship between future energy intraday volatility and trading volume with wavelet. *Appl. Econ.* **2017**, *49*, 1981–1993. [CrossRef]
4. Kumar, S.; Pathak, R.; Tiwari, A.K.; Yoon, S.M. Are exchange rates interdependent? Evidence using wavelet analysis. *Appl. Econ.* **2017**, *49*, 3231–3245. [CrossRef]
5. Chaudhuri, S.; Lo, A. *Spectral Portfolio Theory*; Technical report; MIT Sloan School of Management: Cambridge, MA, USA, 2016. [CrossRef]
6. Nicolis, O.; Mateu, J.; Contreras-Reyes, J.E. Wavelet-Based Entropy Measures to Characterize Two-Dimensional Fractional Brownian Fields. *Entropy* **2020**, *22*, 196. [CrossRef]
7. In, F.; Kim, S. *An Introduction to Wavelet Theory in Finance: A Wavelet Multiscale Approach*; World Scientific: Singapore, 2012; p. 204.
8. Reményi, N.; Vidakovic, B. Bayesian nonparametric regression using complex wavelets. *arXiv* **2018**, arXiv:1803.02532.
9. Barber, S.; Nason, G.P. Real nonparametric regression using complex wavelets. *J. R. Statist. Soc. B* **2004**, *66*, 927–939. [CrossRef]
10. Johnstone, I.M.; Silverman, B.W. Empirical Bayes selection of wavelet thresholds. *Ann. Stat.* **2005**, *33*, 1700–1752. [CrossRef]

11. Donoho, D.L.; Johnstone, I.M. Adapting to Unknown Smoothness via Wavelet Shrinkage. *J. Am. Stat. Assoc.* **1995**, *90*, 1200–1224. [CrossRef]

12. Xiao, F.; Zhang, Y. A Comparative Study on Thresholding Methods in Wavelet-based Image Denoising. *Procedia Eng.* **2011**, *15*, 3998–4003. [CrossRef]

13. Cai, T.T.; Zhou, H.H. A data-driven block thresholding approach to wavelet estimation. *Ann. Stat.* **2009**, *37*, 569–595. [CrossRef]

14. He, C.; Xing, J.; Li, J.; Yang, Q.; Wang, R. A New Wavelet Threshold Determination Method Considering Interscale Correlation in Signal Denoising. *Math. Probl. Eng.* **2015**, *2015*,280251. [CrossRef]

15. Aguiar-Conraria, L.; Soares, M.J.; Aguiar-Conraria, L. The Continuous Wavelet Transform: Moving beyond uni- and bivariate analysis. *J. Econ. Surv.* **2014**, *28*, 344–375. [CrossRef]

16. Scott, D.W. Parametric statistical modeling by minimum integrated square error. *Technometrics* **2001**, *43*, 274–285. [CrossRef]

17. Scott, A. Denoising by Wavelet Thresholding Using Multivariate Minimum Distance Partial Density Estimation. Ph.D. Thesis, Rice University, Houston, TX, USA, 2006.

18. Raath, K. Dynamic Multivariate Wavelet Signal Extraction and Forecasting with Applications to Finance. Ph.D. Thesis, Rice University, Houston, TX, USA, 2020.

19. Raath, K.C.; Ensor, K.B. Time-varying wavelet-based applications for evaluating the Water-Energy Nexus. *Front. Energy Res.* **2020**, *8*, 118. [CrossRef]

20. Raath, K.C.; Ensor, K.B. Wavelet-L_2E Stochastic Volatility Models: An Application to the Water-Energy Nexus. *Sankhya B* **2020**, *85*, 150–176. [CrossRef]

21. Torrence, C.; Compo, G.P.; Torrence, C.; Compo, G.P. A Practical Guide to Wavelet Analysis. *Bull. Am. Meteorol. Soc.* **1998**, *79*, 61–78. [CrossRef]

22. Aguiar-Conraria, L.; Azevedo, N.; Soares, M.J. Using wavelets to decompose the time–frequency effects of monetary policy. *Physica A* **2008**, *387*, 2863–2878. [CrossRef]

23. Reboredo, J.C.; Rivera-Castro, M.A.; Ugolini, A. Wavelet-based test of co-movement and causality between oil and renewable energy stock prices. *Energy Econ.* **2017**, *61*, 241–252. [CrossRef]

24. Lilly, J.; Olhede, S. Higher-Order Properties of Analytic Wavelets. *IEEE Trans. Signal Process.* **2009**, *57*, 146–160. [CrossRef]

25. Addison, P.S. *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2017. [CrossRef]

26. Gay, D.M. *Usage Summary for Selected Optimization Routines*; Technical report; AT&T Bell Laboratories: Murray Hill, NJ, USA, 1990.

27. Ferrer, R.; Jammazi, R.; Bolós, V.J.; Benítez, R. Interactions between financial stress and economic activity for the U.S.: A time- and frequency-varying analysis using wavelets. *Phys. A Stat. Mech. Its Appl.* **2018**, *492*, 446–462. [CrossRef]

28. Xu, Y.; Weaver, J.B.; Healy, D.M.; Lu, J. Wavelet Transform Domain Filters: A Spatially Selective Noise Filtration Technique. *IEEE Trans. Image Process.* **1994**, *3*, 747–758. [CrossRef] [PubMed]

29. Wand, M.P.M.P.; Jones, M.C.M.C. *Kernel Smoothing*; Chapman & Hall: London, UK, 1995; p. 212.