

ORIGINAL ARTICLE

TREND LOCALLY STATIONARY WAVELET PROCESSES

EUAN T. MCGONIGLE^{a,b}  REBECCA KILLICK^c  AND MATTHEW A. NUNES^d 

^a*STOR-i Centre for Doctoral Training, Lancaster University, Lancaster, UK*

^b*School of Mathematics, University of Bristol, Bristol, UK*

^c*Department of Mathematics and Statistics, Lancaster University, Lancaster, UK*

^d*Department of Mathematical Sciences, University of Bath, Bath, UK*

Most time series observed in practice exhibit first- as well as second-order non-stationarity. In this article we propose a novel framework for modelling series with simultaneous time-varying first- and second-order structure, removing the restrictive zero-mean assumption of locally stationary wavelet processes and extending the applicability of the locally stationary wavelet model to include trend components. We develop an associated estimation theory for both first- and second-order time series quantities and show that our estimators achieve good properties in isolation of each other by making appropriate assumptions on the series trend. We demonstrate the utility of the method by analysing the global mean sea temperature time series, highlighting the impact of the changing climate.

Received 17 February 2021; Revised 21 January 2022; Accepted 2 February 2022

Keywords: Climate data; locally stationary; non-stationary time series; trend estimation; wavelet spectrum.

MOS subject classification: *MSC 2020*: Primary 62M10; secondary 62G05.

1. INTRODUCTION

In many contexts, it is common to encounter time series for which the mean and autocovariance of the series vary over time. Examples of time series that exhibit such non-stationary first- and second-order behaviour span a broad range of applications, including financial time series (Fryzlewicz, 2005), climate data (Beaulieu and Killick, 2018), and meteorology (Shen, 2015). Consequently, methodology that can incorporate both first- and second-order non-stationarity within the same modelling framework is appealing in numerous practical scenarios. Often, however, analysis is only performed on either first- or second-order properties, rather than both. In this article, we propose a two-stage wavelet-based framework to estimate the trend – the smooth, long-term behaviour – and the time-varying second-order structure of a time series.

Due to their many advantageous properties, wavelet methods have enjoyed popularity for many years in time series analysis. In particular, wavelets' localised nature can provide sparse location-scale signal decompositions; for a recent overview of wavelet techniques, see for example Vidakovic (2009) or Nason (2008). Recently, there has been much focus in the area of locally stationary wavelet (LSW) modelling, introduced in the seminal work of Nason *et al.* (2000). The LSW framework provides a modelling approach for time series whose second-order structure evolves slowly over time. The model has been successfully applied for a number of tasks in time series analysis, including forecasting (Fryzlewicz *et al.*, 2003), changepoint analysis (Killick *et al.*, 2013), stationarity testing (Nason, 2013) and clustering/classification (Hargreaves *et al.*, 2018; Wilson *et al.*, 2019). This particular approach to modelling has shown to provide improved analytic insights in areas such as medicine (Sanderson *et al.*, 2010; Park *et al.*, 2014), biology (Hargreaves *et al.*, 2019) and environmental science (Chapman *et al.*, 2020). A comprehensive overview of second-order non-stationary time series modelling can be found in Dahlhaus (2012).

* Correspondence to: Euan T. McGonigle, School of Mathematics, University of Bristol, Bristol, UK.
 Email: euan.mcgonigle@bristol.ac.uk

Under the LSW framework, a time series is assumed to have zero mean, and so typically the time series is detrended prior to applying the model. This extra step will have implications for data analysis, for example, by introducing extra uncertainty, and potentially a bias in estimation. Furthermore, the task of estimating the mean function in the presence of non-stationary second-order structure is highly challenging. In principle, most standard techniques from non-parametric regression such as kernel or local polynomial estimates may be used. However, due to the non-stationarity of the error structure, theoretical and practical performance suffer.

There exists a body of work for mean estimation in the presence of time-varying dependent error structures. For example, von Sachs and MacGibbon (2000) use wavelet thresholding to estimate the first-order structure only. The threshold used is data-driven and depends on the unknown second-order structure, and hence the method can be difficult to implement in practice. In the semi-parametric setting, Dahlhaus (2000) and Dahlhaus and Neumann (2001) establish results where the mean function is time-varying and/or estimated. Zhang and Wu (2011) test whether the mean trend of a non-stationary time series is of certain parametric forms. For literature on trend estimation by wavelet thresholding, see, for example, Brillinger (1994), Donoho and Johnstone (1995) or Donoho *et al.* (1995). For a review of other non-parametric regression methods including kernel smoothing and splines, see Härdle (1990).

Some techniques do consider mean estimation alongside parameters relating to second-order non-stationarity, see, for example, Tunyavetchakit (2010) for time-varying autoregressive processes. Beran and Feng (2002) propose parallel first- and second-order estimation, but the second-order structure follows a parametric model and is assumed to be difference-stationary. However, for locally stationary series, many authors acknowledge that the topic ‘needs more investigation’ (Dahlhaus, 2012, p. 56), and thus a principled approach to this challenge remains open.

In this article, we propose a wavelet-based framework to model time series that exhibit time-varying first- and second-order structure, inspired by the LSW model of Nason *et al.* (2000). The trend is considered to be a deterministic smooth function representing long-term (low frequency) patterns or systematic variations in the time series. By making modest assumptions on the trend, we develop rigorous theory to achieve consistent estimation of the key time-varying second-order quantities of interest. Crucially this includes a proof that a key wavelet matrix has a bounded inverse for all Daubechies wavelets – previously only shown for the boundary cases of Haar and Shannon wavelets. Further, we propose a wavelet thresholding technique to consistently estimate the trend term in the presence of second-order non-stationarity. In practice, however, our estimators also work well for trends that depart from our assumptions. Finally, we provide new methodology for the boundary handling in the discrete wavelet transform employed during the estimation of first- and second-order quantities. This is of particular interest when analysing first and second-order non-stationary time series, which are not of dyadic length, as standard techniques can produce spurious artefacts at the boundaries of the time series.

The rest of the article is organised as follows. In Section 2 we give a brief overview of the LSW framework and propose the trend locally stationary wavelet process model, discussing the key assumptions. Sections 3 and 4 introduce the methodology for estimating the second-order structure and trend in the model. In Section 5 we assess the performance of the method via a simulation study. In Section 6, we examine the much-studied global mean sea temperature data set, providing new insights and highlighting the potential uses of our framework. Concluding remarks are given in Section 7, while proofs of results and additional simulations are contained in the Supporting Information.

2. THE TREND LOCALLY STATIONARY WAVELET PROCESS MODEL

2.1. Background to LSW Processes

We begin by giving a brief overview of the original LSW modelling paradigm of Nason *et al.* (2000), which has been widely used for the modelling of second-order non-stationary – but first-order stationary – time series. This will form the basis for our more general trend non-stationary time series model. A locally stationary wavelet (LSW) process is defined as follows.

Definition 2.1. A triangular stochastic array $\{X_{t,T}\}_{t=0}^{T-1}$ for $T = 2^J \geq 1$ is in a class of locally stationary wavelet (LSW) processes if there exists a mean square representation

$$X_{t,T} = \sum_{j=-\infty}^{-1} \sum_{k \in \mathbb{Z}} w_{j,k;T} \psi_{j,k-t} \xi_{j,k}, \tag{2.1}$$

where j and $k \in \mathbb{Z}$ are scale and location parameters respectively, $\{\xi_{j,k}\}$ is a random orthonormal increment sequence, $\{\psi_{j,k-t}\}_{j,k}$ is a set of discrete non-decimated wavelets, and $\{w_{j,k;T}\}$ is a set of amplitudes. The quantities in representation (2.1) possess the following properties:

1. $\mathbb{E}(\xi_{j,k}) = 0$ for all j, k .
2. $\text{Cov}(\xi_{jk}, \xi_{lm}) = \delta_{jl} \delta_{km}$.
3. There exists, for each $j \leq -1$, a Lipschitz continuous function $W_j(z)$ for $z \in (0, 1)$, which satisfies the following properties:

$$\sum_{j=-\infty}^{-1} |W_j(z)|^2 < \infty \text{ uniformly in } z \in (0, 1);$$

the Lipschitz constants L_j are uniformly bounded in j and $\sum_{j=-\infty}^{-1} 2^{-j} L_j < \infty$. There exists a sequence of constants C_j such that for each T ,

$$\sup_k \left| w_{j,k;T} - W_j\left(\frac{k}{T}\right) \right| \leq \frac{C_j}{T},$$

and where the sequence $\{C_j\}$ satisfies $\sum_{j=-\infty}^{-1} C_j < \infty$.

Since $\mathbb{E}(\xi_{j,k}) = 0$ for all j and k , the process mean $\mathbb{E}(X_{t,T}) = 0$ for all t . This means the LSW framework is only able to model time-varying *second-order* structure. As with classical time series theory, the second-order structure of an LSW process is encoded in the spectrum. The evolutionary wavelet spectrum (EWS) of an LSW process is defined as $S_j(z) := |W_j(z)|^2$ for rescaled time $z = k/T \in (0, 1)$ and measures the contribution to variance at a particular rescaled time z and scale j . Since the W_j are Lipschitz continuous, the spectrum at level j , S_j , is also Lipschitz continuous; however Fryzlewicz and Nason (2006) and Van Bellegem and von Sachs (2008) extend the LSW model to consider piecewise constant spectra and those of bounded total variation respectively.

The EWS is estimated via the empirical wavelet coefficients of the time series, given by $d_{j,k;T} := \langle X_{t,T}, \psi_{j,k-t} \rangle = \sum_t X_{t,T} \psi_{j,k-t}$. The raw wavelet periodogram, defined by $I_{k,T}^j := |d_{j,k;T}|^2$ is a biased, inconsistent estimator of the EWS (Nason *et al.*, 2000, proposition 4):

$$\mathbb{E}\left(I_{k,T}^j\right) = \sum_l A_{jl} S_l(k/T) + \mathcal{O}(T^{-1}), \tag{2.2}$$

$$\text{Var}(I_{k,T}^j) = 2 \left(\sum_l A_{jl} S_l(k/T) \right)^2 + \mathcal{O}(2^{-j} T^{-1}), \tag{2.3}$$

where the operator $A = (A_{jl})_{j,l < 0}$ is given by $A_{jl} = \sum_\tau \Psi_j(\tau) \Psi_l(\tau)$, and the autocorrelation wavelets $\Psi_j(\cdot)$ are defined by $\Psi_j(\tau) := \sum_{k \in \mathbb{Z}} \psi_{j,k} \psi_{j,k-\tau}$, $j < 0$, $\tau \in \mathbb{Z}$.

The vector of periodograms $\mathbf{I}(z) := \{I_{[z]T}^j\}_{j=-1, \dots, -J}$ can be bias-corrected using the J -dimensional matrix $A_J := (A_{jl})_{j,l=-1, \dots, -J}$ which is invertible. The vector of corrected periodograms $\mathbf{L}(z) := \{L_{[z]T}^j\}_{j=-1, \dots, -J}$ with

$\mathbf{L}(z) = A_J^{-1} \mathbf{I}(z)$ satisfies

$$\mathbb{E}(\mathbf{L}(z)) = \mathbb{E}(A_J^{-1} \mathbf{I}(z)) = \mathbf{S}(z) + \mathcal{O}(T^{-1}) \quad \forall z \in (0, 1), \quad (2.4)$$

where $\mathbf{S}(z) := \{S_j(z)\}_{j=-1, \dots, -J}$. The raw wavelet periodogram is first smoothed and then corrected by A_J^{-1} to produce an asymptotically unbiased, consistent estimator. Smoothing can be carried out using a number of techniques, for example, via a running mean as in Nason (2013) or using wavelet thresholding as in Nason *et al.* (2000).

The local autocovariance (LACV) function for an LSW process provides information about the covariance at a rescaled location $z = k/T \in (0, 1)$. The LACV, $c(z, \tau)$, of an LSW process with EWS $\{S_j(z)\}$ is defined as $c(z, \tau) = \sum_{j=-\infty}^{-1} S_j(z) \Psi_j(\tau)$, $\tau \in \mathbb{Z}, z \in (0, 1)$. The LACV can be thought of as a decomposition of the autocovariance of a process over scale and rescaled time. The process autocovariance arises as the asymptotic limit of the LACV. In practice, the LACV is estimated by plugging in the smoothed, corrected estimate for the EWS into the equation for the LACV, which results in a consistent estimator.

2.2. Modelling Locally Stationary Time Series with Trend

We introduce our proposed model, which relaxes the zero mean assumption inherited from Assumption 1 in Definition 2.1 above. We will show that it is possible to incorporate a non-stationary polynomial trend term within the context of a second-order non-stationary time series. Furthermore, we will show that this polynomial assumption can be relaxed, facilitating modelling of time series with more complex first-order structure within our paradigm. This allows for self-contained modelling of both first- and second-order non-stationarity of a time series. Hence, it is not necessary to remove the trend before estimating second-order structure, and vice versa, we can obtain an estimate of first-order structure in the presence of the non-stationary second-order behaviour. To this end, we define the polynomial trend locally stationary wavelet process as follows.

Definition 2.2. A polynomial trend LSW process $\{X_{t,T}\}$, $t = 0, \dots, T-1$, and $T = 2^J$ is a doubly-indexed stochastic process with the following representation in the mean square sense:

$$X_{t,T} = \mu\left(\frac{t}{T}\right) + \sum_{j=-\infty}^{-1} \sum_k w_{j,k;T} \Psi_{j,k-t} \xi_{j,k}, \quad (2.5)$$

where the quantities in representation (2.5) possess the same properties as in Definition 2.1, and in addition the function μ is a polynomial of degree p , that is,

$$\mu\left(\frac{t}{T}\right) = P_p\left(\frac{t}{T}\right) = \sum_{i=0}^p a_i \frac{t^i}{T^i}, \quad a_i \in \mathbb{R}.$$

Our model imposes the same assumptions on the LSW component as in Nason *et al.* (2000), allowing for locally stationary second-order structure, while also permitting non-stationary first-order behaviour by incorporating a polynomial mean function. Using the model defined in Equation (2.5), we build on the rigorous theory of Nason *et al.* (2000) to estimate the non-stationary first- and second-order structure of the time series. Note that the assumption that $T = 2^J$ is a theoretical one; in practice, using the boundary handling method described in Section 4.4, we can analyse time series of any length.

3. SPECTRAL ESTIMATION THEORY

Given the process model definition in Equation (2.5), we now turn to the question of estimating the appropriate quantities of the model. We discuss estimation of the evolutionary wavelet spectrum $S_j(z) = |W_j(z)|^2$ and local autocovariance function. In Nason *et al.* (2000), rigorous estimation theory is provided for the case where the

time series does not display a trend. We extend this work to allow for the inclusion of a trend by examining three separate scenarios: first, we examine the case where the trend is a polynomial of degree less than the order of the discrete wavelet used in the LSW process (2.5); second, the case where the trend is a higher degree than the discrete wavelet; and lastly, the case of a general Hölder continuous trend.

Using discrete wavelets to model the second-order non-stationarity allows the incorporation of smooth trends due to the ability of the wavelet to ‘cancel out’ this trend. This means we circumvent the problem of detrending the series first, which is challenging due to the second-order non-stationarity. In all three scenarios, we provide appropriate theoretical results to enable consistent estimation of both the EWS and local autocovariance of the time series. Given that trends are usually assumed to be slowly-varying, smooth functions, such as low-order polynomials (see, e.g. Priestley, 1983; Craigmile *et al.*, 2005; Kallache *et al.*, 2005), our model in (2.5) is applicable to a wide range of data scenarios. In practice, this can serve as a good approximation even when we depart from the given assumptions.

3.1. Spectral Estimation with Low-Order Polynomial Trends

First, we examine the case of a low-order polynomial trend. We show that we can use the same estimation procedure as in Nason *et al.* (2000) to obtain an asymptotically unbiased, consistent estimator of the EWS. This can then be used to obtain a consistent estimator of the LACV.

Recall that a function $\psi \in L^2(\mathbb{R})$ is said to have m vanishing moments if it satisfies $\langle x^l, \psi(x) \rangle = \int x^l \psi(x) dx = 0$, for $l = 0, 1, \dots, m-1$. Denote by $v(\psi)$ the number of vanishing moments of the discrete wavelet ψ . If $v(\psi) = m$, then all wavelet coefficients of any polynomial of degree $m-1$ or less will be zero. The commonly used extremal phase (EP) and least asymmetric (LA) Daubechies wavelet families have k vanishing moments for a filter of length $2k$. For example, the Haar wavelet has one vanishing moment and can only annihilate constant behaviour, while the Daubechies EP wavelet with 10 vanishing moments can annihilate polynomial behaviour up to degree 9.

We henceforth refer to the discrete wavelet ψ defined in the LSW process representation in Equation (2.5) as the *generating wavelet*. Since second-order estimation is performed using the generating wavelet coefficients, estimation will be unaffected by the trend, provided that $p < v(\psi)$. Hence, we can obtain an unbiased estimate of the spectrum in the usual way.

Lemma 3.1. Suppose that $p < v(\psi)$. Then, the expectation and variance of the non-boundary raw wavelet periodogram is given by Equations (2.2) and (2.3) respectively.

We can correct the raw wavelet periodogram estimator in the usual way with A_J^{-1} as in Equation (2.4) to obtain an asymptotically unbiased estimator, just as in the original LSW model. Hence, as long as the trend of the time series can be represented by a low-order polynomial and we use a high enough order discrete wavelet, we can still apply the standard EWS estimation procedure without incurring any additional bias.

Remark 1. In practice, only the non-boundary wavelet coefficients will be free from trend, while the boundary wavelet coefficients will still contain the trend. This is due to the way in which the non-decimated discrete wavelet transform (DWT) is computed at the boundary values of the time series, see, for example, section 4.11 of Percival and Walden (2006). In Section 4.4, we address this issue by modifying standard boundary handling methods for estimation at the boundaries of the time series.

As in the original LSW model, resulting spectral estimates are inconsistent and must be smoothed. To achieve consistency, we can simply follow the smoothing strategy taken in Nason *et al.* (2000): for each fixed scale j , the periodogram $I_{k,T}^j$ (which is scaled χ^2 -distributed) is smoothed as a function of $z = k/T$ using, for example, DWT shrinkage or translation invariant (TI) denoising of Coifman and Donoho (1995). The following two results, which hold for Gaussian LSW processes (and in turn polynomial trend Gaussian LSW processes), describe how consistency is achieved through DWT shrinkage.

Work in von Sachs *et al.* (1997) explains how to smooth using an orthonormal second stage wavelet basis of $L_2([0, 1])$, denoted $\{\tilde{\psi}_{rs}\}$. Smoothing is achieved using nonlinear thresholding of the empirical wavelet coefficients,

\hat{u}_{rs} , of $I^j(z)$ and then inverting using A^{-1} to give the estimate $\hat{S}_j(z)$. This leads to an extension of theorem 4 and proposition 5 of Nason *et al.* (2000) – which are only applicable for Haar and Shannon wavelet zero-mean LSW processes – to trend LSW processes generated by any discrete wavelet in the Daubechies compactly supported family.

Proposition 3.2. Let $\tilde{\psi}$ be a discrete wavelet of bounded variation, with $2^r = o(T)$ for wavelet coefficients \hat{u}_{rs} . Suppose that $S_j(k/T) \leq D2^{5j/6}$ for some constant D . For a polynomial trend Gaussian LSW process generated by any Daubechies compactly supported discrete wavelet, with threshold given by $\lambda^2(j, r, s, T) = \text{Var}(\hat{u}_{rs})\log^2(T)$, for each fixed j ,

$$\mathbb{E} \left[\int_0^1 (\hat{S}_j(z) - S_j(z))^2 dz \right] = \mathcal{O} (2^j \log^2(T) T^{-2/3}). \tag{3.1}$$

Lastly, we can consistently estimate the LACV by using the EWS estimate.

Proposition 3.3. Suppose the assumptions of Proposition 3.2 hold, and define $\hat{c}(z, \tau)$ by replacing $S_j(z)$ with $\hat{S}_j(z)$ in the equation for the LACV, that is,

$$\hat{c}(z, \tau) = \sum_{j=-J}^{-1} \hat{S}_j(z) \Psi_j(\tau). \tag{3.2}$$

Then, for each fixed $\tau \in \mathbb{Z}$, $\hat{c}(z, \tau)$ satisfies

$$\mathbb{E} \left[\int_0^1 (\hat{c}(z, \tau) - c(z, \tau))^2 dz \right] = \mathcal{O} (\log^2(T) T^{-2/3}).$$

Propositions 3.2 and 3.3 show that in the case of a low-order polynomial trend, we can still consistently estimate the evolutionary wavelet spectrum and local autocovariance on the non-boundary values of the wavelet periodogram. The assumption on the decay rate of the EWS, $S_j \leq D2^{5j/6}$, is necessary for the proof of Proposition 3.2, and hence Proposition 3.3 as well. This assumption is, for example, satisfied for the white noise process, where $S_j = 2^j$. The assumption is required to control the mean squared error in estimating the EWS using the J available scales, as we do not have access to all (infinite) scales in practice; it is chosen to balance the mean squared error rate (up to the log factor) of $\mathcal{O}(\log^2(T) T^{-2/3})$ that arises from the wavelet thresholding procedure. This is a slightly weaker assumption than that of Fryzlewicz and Nason (2006) and Sanderson *et al.* (2010), who assume that $S_j \leq D2^j$. Note that no explicit assumption on S_j is made in the original work of Nason *et al.* (2000). However, the modification to the original LSW model made in Fryzlewicz (2003), the accepted definition in the LSW literature, would necessitate the same assumption that we have here.

The extension of consistency results to all Daubechies compactly supported discrete wavelets is in itself an important contribution to the area of LSW analysis, and is due to a new result on properties of the autocorrelation wavelet inner product matrix A , given in Proposition 3.2 in the Supporting Information. This proof can be used to extend much of the current theory of LSW processes – which previously was only valid for Haar and Shannon wavelets – to all compactly supported Daubechies wavelets. This includes, for example, the work on bivariate LSW processes of Sanderson *et al.* (2010), and the extension of the LSW model to spectra of bounded total variation of Van Bellegem and von Sachs (2008).

3.2. Spectral Estimation with High-Order Polynomial Trends

Next, we consider the case of higher-order polynomial trends with respect to the generating wavelet, which allows for the modelling of less smooth trends. The key idea is to use a smoother discrete wavelet to analyse the series, which ensures that the trend is removed from the wavelet coefficients. By calculating the bias that this causes, we can correct the raw wavelet periodogram analogously to Equation (2.4) to obtain an unbiased estimate.

If the polynomial trend of the time series is a higher degree than the discrete wavelet generating the LSW process, then the standard non-decimated discrete wavelet transform will not zero out the trend, causing bias in our estimation. For example, if the Haar wavelet is the generating wavelet, and the time series exhibits a linear trend, then additional bias will be incurred in the computation of the raw wavelet periodogram, because the Haar wavelet is only capable of annihilating constants. If we instead use the Daubechies EP wavelet with two vanishing moments, the linear trend would be removed by the discrete wavelet. Then, the bias of the raw wavelet periodogram will be quantified in terms of the mismatch between the generating wavelet and analysing wavelet. The next result shows we can correct for this bias, and extends the work of Gott and Eckley (2013), which examined the effect of wavelet misspecification on EWS estimation.

Lemma 3.4. Suppose that the polynomial LSW process is generated by discrete wavelet ψ^0 , and let ψ^1 be a Daubechies compactly supported discrete wavelet with $p < \nu(\psi^1)$. Define the non-decimated discrete wavelet coefficients to be $d_{j,k} = \sum_t X_t \psi_{j,k-t}^1$. Let the matrix $C^{1,0}$ be the inner product matrix of autocorrelation wavelets whose j, l th entry is given by

$$C_{jl}^{1,0} = \sum_{\tau} \Psi_j^1(\tau) \Psi_l^0(\tau),$$

where the superscript denotes the fact that the underlying discrete wavelet is different. Then,

$$\begin{aligned} \mathbb{E}(d_{j,k}^2) &= \sum_l C_{jl}^{1,0} S_l(k/T) + \mathcal{O}(T^{-1}), \\ \text{Var}(d_{j,k}^2) &= 2 \left(\sum_l C_{jl}^{1,0} S_l(k/T) \right)^2 + \mathcal{O}(2^{-j} T^{-1}). \end{aligned}$$

Provided that $C_j^{1,0}$ is invertible, Lemma 3.4 implies that, for the vector of periodograms $\mathbf{I}(z) := \{I_{[zT,T]}^l\}_{l=-1, \dots, -J}$, and the vector of corrected periodograms $\mathbf{L}(z) := \{L_{[zT,T]}^j\}_{j=-1, \dots, -J}$ with $\mathbf{L}(z) = (C_j^{1,0})^{-1} \mathbf{I}(z)$,

$$\mathbb{E}(\mathbf{L}(z)) = \mathbb{E}((C_j^{1,0})^{-1} \mathbf{I}(z)) = \mathbf{S}(z) + \mathcal{O}(T^{-1}) \quad \forall z \in (0, 1), \quad (3.3)$$

where $\mathbf{S}(z) = \{S_j(z)\}_{j=-1, \dots, -J}$. Hence, we do not have to restrict to assuming the trend has a lower order than the discrete wavelet that generates the LSW process. In this scenario, we can perform the following altered estimation of the EWS. First, we apply the non-decimated discrete wavelet transform using a *higher order wavelet*, removing the trend. Next, we form the periodogram and perform smoothing. Finally, we correct the estimate by premultiplying the appropriate inverse matrix $(C^{1,0})^{-1}$.

Correcting in this fashion can be beneficial even when the time series contains no trend. Smoother wavelets have a shorter support in the Fourier domain and therefore experience less spectral leakage, in which power can leak into nearby scales. For example, a Haar wavelet estimate will experience more leakage than using a smoother discrete wavelet like the EP10 wavelet. Therefore, the off-diagonal entries of the Haar A matrix exhibit slower decay than that of the EP10 A matrix. Hence, the off-diagonal entries of the corresponding C matrix will be smaller than those of the less smooth discrete wavelet.

If the LSW process is generated by a Haar wavelet and we use a Haar wavelet to analyse the series, then the spectral estimate may be prone to leakage and the accuracy of the estimate will suffer as a result. However, if the process is generated by an EP10 wavelet and we use the EP10 wavelet to analyse, the estimate will exhibit less leakage. Therefore, it can be beneficial, in the case where the generating wavelet is not very smooth, to use a smoother wavelet to analyse, and then correct the estimate with the corresponding C -inverse matrix, rather than using the generating wavelet to analyse and the usual A -inverse matrix. Using the smoother wavelet to analyse helps to reduce leakage, while still achieving an unbiased estimate by correcting in the appropriate way. This leads

to a more accurate spectral estimate: in the Supporting Information, we provide numerical experiments to verify this observation.

Intuitively, there is strong evidence to suggest that the matrix $C^{1,0}$ is invertible. By Proposition 3.2 in the Supporting Information, we show that the inverse of the operator A is bounded for all Daubechies compactly supported discrete wavelets. The A -matrix is the Gram matrix of the set $\{\Psi_j(\tau)\}_{j \leq -1}$ of linearly independent autocorrelation wavelets. In our case, the matrix $C^{1,0}$ is the cross-Gram matrix of two sets of autocorrelation wavelets $\{\Psi_j^0\}_{j \leq -1}$ and $\{\Psi_j^1\}_{j \leq -1}$. Unfortunately, the properties of Gram matrices do not necessarily carry over to cross-Gram matrices. This means we cannot directly adapt the proof of the invertibility of A to the invertibility of C . Intuitively, forming the cross-Gram matrix uses a mix of two discrete wavelet families, and we might expect that the properties of the cross-Gram matrix are a ‘mixture’ or ‘average’ of the properties of the two Gram matrices associated to the two wavelet families.

For example, when $J = 10$, the numerically calculated condition number of the Haar A matrix is 860.47, while it is equal to 393.12 for the Daubechies extremal phase discrete wavelet with 10 vanishing moments. For the two C matrices associated with these two families, the condition numbers are 596.97 and 610.97. These numbers are close to the average of the two A matrix condition numbers. This suggests that the C matrices should also be invertible, with bounded inverse.

3.3. Spectral Estimation with General Smooth Trends

A polynomial trend may not be an appropriate assumption for a given time series. We investigate how we can modify the spectral estimation procedure when we place a Hölder continuous smoothness assumption on the trend. Often, wavelet methods for curve estimation (distinct from trend estimation where more general mean functions are considered) assume that the mean function lies in some Besov space (see e.g. von Sachs and MacGibbon, 2000; Neumann and von Sachs, 1995), of which Hölder continuous functions are a special case.

Intuitively, the discrete wavelet coefficients are calculated using difference operations, and so if the trend function is sufficiently smooth, we expect that these discrete wavelet coefficients will be reasonably small. When using these coefficients to estimate the EWS – in some of the finer scales at least – where we use less of the data to calculate the wavelet coefficients, and where local changes are small, we could expect the estimate to remain reasonably accurate.

Formally, consider the class of Hölder continuous functions with exponent β , where $0 < \beta \leq 1$, in place of the polynomial trend assumption in Definition 2.2. For example, $\mu(z) = z^\beta$ is Hölder continuous with exponent β , and an exponent of $\beta = 1$ corresponds to Lipschitz continuity. Again, this seems a reasonable assumption to make, given that trend functions are generally assumed to be smooth and slowly evolving – for example, sinusoids are Lipschitz continuous. As shown in Lemma 3.5, we accumulate bias in the raw wavelet periodogram proportional to the level of the discrete wavelet transform and the smoothness of the trend, which is governed by β .

Lemma 3.5. For any Hölder continuous trend LSW process with exponent β , the expectation of the raw wavelet periodogram is given by

$$\mathbb{E} \left(I_{k,T}^j \right) = \sum_l A_{jl} S_l(k/T) + \mathcal{O}(2^{-j(2\beta+1)} T^{-2\beta}) + \mathcal{O}(T^{-1}).$$

Thus we introduce extra bias due to the trend, which we must account for in our estimation procedure, that is, when smoothing and correcting. At coarser scales, the bias caused by the trend accumulates in the squared wavelet coefficients. Analogous behaviour occurs when performing Fourier-based spectral analysis. This is characterised by the phenomenon of a peak in the Fourier periodogram near the zero frequency, as discussed in Dalla *et al.* (2020). To avoid the accumulation of bias, one could apply differencing to the time series first to remove the trend. However, differencing can vastly alter the spectral structure of the time series. A standard LSW analysis performed on the differenced series may not be appropriate, and the differencing operation should be taken

into account in the estimation procedure. A thorough investigation of a difference-based approach can be found in McGonigle *et al.* (2021).

To smooth the wavelet periodogram, we propose to use a simple running mean, as in Nason (2013) and Sanderson *et al.* (2010), with bin width size $2n + 1$. The smoothed wavelet periodogram is thus defined as

$$\tilde{I}_k^j = \frac{1}{2n+1} \sum_{m=-n}^n I_{k+m}^j. \quad (3.4)$$

For simplicity, we use a running mean, but one could just as easily use wavelet thresholding, and achieve similarly consistent estimation. We must now correct the smoothed wavelet periodogram using the inverse matrix A^{-1} . When we do this, we alter the standard estimation procedure by only correcting across a certain number of scales – in relation to T – to control for the bias caused by the trend and the smoothing step. One can view this as akin to a tapered estimator, which we utilise for both EWS and LACV estimation. Based on the above observations, we can now state our results related to the consistency of the smoothed estimators of the EWS and local autocovariance in the presence of a Hölder continuous trend.

Theorem 3.6. Suppose that $S_j(k/T) \leq D2^j$ for some constant D , and μ is Hölder continuous with exponent β , and let $J_0 = \alpha \log_2(T)$ for $\alpha \in (0, 1)$. Then, uniformly in k , the EWS estimator $\hat{S}_j(k/T)$, defined by

$$\hat{S}_j(k/T) = \sum_{l=-J_0}^{-1} \tilde{I}_k^l A_{lj}^{-1}, \quad (3.5)$$

is mean square consistent for each fixed scale j , provided that $n^{-1}T^\alpha \rightarrow 0$, $nT^{-1} \rightarrow 0$ and $T^{4\beta(\alpha-1)+\alpha} \rightarrow 0$ as $T \rightarrow \infty$ and $n \rightarrow \infty$. The mean squared error of $\hat{S}_j(k/T)$ is given by

$$\mathbb{E}(\hat{S}_j(k/T) - S_j(k/T))^2 = \mathcal{O}(2^j n^{-1} T^\alpha) + \mathcal{O}(2^j n T^{-1}) + \mathcal{O}(2^j T^{4\beta(\alpha-1)+\alpha}).$$

Theorem 3.7. Suppose the assumptions of Theorem 3.6 hold. Let $\hat{S}_j(k/T)$ be the estimator of the EWS defined in Equation (3.5). Then, for each fixed $\tau \in \mathbb{Z}$, the estimator of the local autocovariance $c(k/T, \tau)$, defined by

$$\hat{c}(k/T, \tau) = \sum_{j=-J_0}^{-1} \hat{S}_j(k/T) \Psi_j(\tau) \quad (3.6)$$

is mean square consistent, provided that $n^{-1}T^\alpha \rightarrow 0$, $nT^{-1} \rightarrow 0$ and $T^{4\beta(\alpha-1)+\alpha} \rightarrow 0$ as $T \rightarrow \infty$ and $n \rightarrow \infty$. The mean squared error of $\hat{c}(k/T, \tau)$ is given by

$$\mathbb{E}(\hat{c}(k/T, \tau) - c(k/T, \tau))^2 = \mathcal{O}(n^{-1}T^\alpha) + \mathcal{O}(nT^{-1}) + \mathcal{O}(T^{4\beta(\alpha-1)+\alpha}).$$

Note that these results hold for all Daubechies compactly supported discrete wavelets, just as in Propositions 3.2 and 3.3. In the same manner as Propositions 3.2 and 3.3, we require an assumption on the decay rate of the EWS, which is necessary to ensure consistent estimation. For this weaker assumption of general smooth trends, the running mean smoother requires the slightly stronger assumption that $S_j \leq D2^j$, as in Sanderson *et al.* (2010). The third error term, corresponding to the error caused by the trend, will be asymptotically dominated by the first term, irrespective of the bin width n , for $\beta = 1$ and $\alpha < 3/4$. Therefore, the trend term can be seen to have minimal impact on the spectral estimation, provided it is smooth enough. In the presence of a smooth trend, we can still recover a consistent estimator of the EWS and LACV. We simply alter the standard estimation procedure by only using the finest J_0 scales. Further, one could just as easily use a different analysing wavelet to the generating wavelet as described in Section 3.2, and use the appropriate $(C^{1,0})^{-1}$ matrix for bias correction.

Choice of contributing scales and bin width. When using the tapered estimator, we must choose a value of α , the proportion of scales over which we correct. Higher values of α ensure a decomposition over a larger number of scales. In practice, not all scales will be informative, and the number of informative scales will depend on both the data and choice of discrete wavelet, as noted in Sanderson *et al.* (2010). In our experiments, we have observed no significant bias incurred by using most scales, and propose using the proportion $\alpha = 2/3$. This is motivated by results in the next section on trend estimation and is in alignment with Sanderson *et al.* (2010), who also utilise running mean smoothing for LSW spectral estimation. We recommend the choice $\alpha = 2/3$ as a balance between boundary effects, bias in coarse scales occurring due to trend, and capturing the coarse-scale behaviour of the time series. If the practitioner is interested in very coarse-scale behaviour, then provided the trend is not too pronounced, a higher value of α can be used along with a wavelet with shorter filter length. Different values of α can also be compared visually to assess the fit.

The choice of the bin width parameter n in Equation (3.4) will also affect the quality of the estimate. We can choose to use level-dependent smoothing of the periodogram, and so a larger bin width is used in coarser scales. Choosing a bin width in this way will still ensure consistent estimation, and further is natural as coarser scales will experience stronger autocorrelation.

The results in this section establish that we can still consistently estimate the non-stationary second-order structure in the presence of a trend, provided that the trend function is smooth. In practice, using a smoother discrete wavelet to analyse the series is recommended, as it is able to better remove the trend of the series, while also helping to reduce leakage. Smoothing of the wavelet periodogram can be carried out using wavelet thresholding or a running mean smoother. We advocate the latter due to its simplicity and because only one parameter needs to be specified.

Finally, we note that many existing methods that utilise the LSW framework immediately extend to our new model. For example, the extension of the LSW model to spectra of bounded total variation in Van Bellegem and von Sachs (2008) is valid within our model. Furthermore, the test of second-order stationarity of Nason (2013) will still be applicable, provided the discrete wavelet used is smooth enough to remove the trend of the time series. Similarly, locally stationary wavelet models for multi-variate time series, such as Sanderson *et al.* (2010) and Park *et al.* (2014), will be valid for any trend scenario described above.

4. TREND ESTIMATION

Having described the necessary methodology for estimating the non-stationary second-order structure of the time series, we now address the question of estimating the trend of the series. If we assume that the trend function is smooth, a natural method for estimating it is to use wavelets, given that polynomial functions can be exactly represented using wavelet scaling functions (Vidakovic, 2009). Using a discrete wavelet-based approach enables us to obtain a simple intuitive estimator from which we can derive appropriate confidence intervals.

4.1. Trend Estimation Theory

We follow the same strategy as Craigmile *et al.* (2004) and Craigmile *et al.* (2005), who apply wavelet thresholding in the long memory time series setting. Our problem can be posed as the following. The observed data vector $\mathbf{X} = (X_0, \dots, X_{T-1})^\top$ is given by

$$\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\varepsilon},$$

where $\boldsymbol{\mu}$ is the trend component and $\boldsymbol{\varepsilon}$ is the LSW process term. If we use a discrete wavelet that has a high enough order with respect to the trend, the discrete wavelet transform of the time series will decompose it into a trend and noise component. This is because the non-boundary wavelet coefficients of a polynomial are zero due to the vanishing moments property, and so since the increments in an LSW process are zero mean, these coefficients have expectation zero. The boundary wavelet coefficients and the scaling coefficients will contain a contribution due to the trend alone.

Motivated by this observation, we can estimate the trend using a type of wavelet thresholding estimator. We perform the discrete wavelet transform on the time series, choosing the coarsest scale j_0 to which we analyse the series. Then, we set the non-boundary wavelet coefficients to zero. Then, we perform the inverse discrete wavelet transform to obtain the estimator of the trend. This estimator captures the long-term behaviour of the time series (trend), but will also contain short-scale variability at the boundaries due to the inclusion of boundary wavelet coefficients.

Formally, let us write the discrete wavelet transform of the time series in matrix form, from which we can derive the matrix form of the estimator. Given the time series is of length $T = 2^j$, the total number of (decimated) wavelet coefficients at each level j is given by $N_j = T2^j$. Then, the $B_j = [(2m - 2)(1 - 2^j)]$ boundary coefficients are given by the outer coefficients, where m is the number of vanishing moments of the wavelet. If p is the order of the polynomial trend, then we can zero out this trend so long as $p < m$. The number of non-boundary coefficients is denoted by $M_j = N_j - B_j$. Denote by $\text{diag}(\mathbf{x})$ the diagonal matrix with \mathbf{x} along the main diagonal, and let $\mathbf{0}_n$ and $\mathbf{1}_n$ denote a vector of n zeros or ones respectively. Let $A = \text{diag}(\mathbf{1}_{B_1}, \mathbf{0}_{M_1}, \mathbf{1}_{B_2}, \mathbf{0}_{M_2}, \dots, \mathbf{1}_{B_j}, \mathbf{0}_{M_j}, \mathbf{1}_{N_j})$, and let I_T denote the $T \times T$ identity matrix. Using Daubechies wavelets ensures the DWT is orthogonal, and so we can partition the data vector \mathbf{x} into a trend component and noise component, as

$$\mathbf{X} = W^T W \mathbf{X} = W^T A W \mathbf{X} + W^T (I_T - A) W \mathbf{X} := \hat{\boldsymbol{\mu}} + \hat{\boldsymbol{\varepsilon}}.$$

Analogously to Proposition 1 of Craigmile *et al.* (2004), the wavelet-based trend estimator is unbiased, and further we can show that the estimator is mean square consistent, in the case of both polynomial and Hölder continuous trend functions.

Proposition 4.1. For a polynomial trend of degree p , and using a discrete wavelet with $m > p$ vanishing moments, $\mathbb{E}(\hat{\boldsymbol{\mu}}) = \boldsymbol{\mu}$.

Proposition 4.2. Let $j_0 = \alpha \log_2 T$, for some $\alpha \in (0, 1)$, be the coarsest scale of the decomposition using discrete wavelet ψ . Suppose that $p < \nu(\psi)$, and assume that $S_j(z) \leq D2^j$ for all $z \in (0, 1)$ and some constant D . Ignoring boundary effects, the trend estimator is consistent in the mean square sense, that is:

$$\frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}((\mu(t/T) - \hat{\mu}(t/T))^2) = \mathcal{O}(T^{-\alpha}).$$

Proposition 4.3. Suppose μ is Hölder continuous with exponent β . Then, under the same conditions on α and S_j as Proposition 4.2, the following holds:

$$\frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}((\mu(t/T) - \hat{\mu}(t/T))^2) = \mathcal{O}(T^{-\alpha}) + \mathcal{O}(T^{2\beta(\alpha-1)}).$$

Choice of the parameter α can be viewed as akin to selecting the bandwidth in a kernel-based estimation procedure. Larger values of α loosely correspond to a larger bandwidth, and ensure more wavelet coefficients are thresholded, giving a smoother fit. For a Lipschitz continuous trend, balancing the rates in Proposition 4.3 leads to taking $\alpha = 2/3$. We use this for both trend and second-order estimation in practice for simplicity. Under the conditions of Proposition 4.2 and subject to some mild regularity conditions, the trend estimate is multi-variate Gaussian, with mean $\boldsymbol{\mu}$ and covariance $RCov(\boldsymbol{\varepsilon})R^T := R\Sigma R^T$, where $R = W^T A W$, following from asymptotic normality of the wavelet coefficients (Stevens, 2013, Theorem 3.1.1). Full discussion of the appropriate assumptions can be found in Stevens (2013). LSW processes with random innovations from the exponential, gamma, inverse-Gaussian, and F-distribution families all satisfy these assumptions.

4.2. Pointwise Confidence Intervals

Using the results of the previous section, we can derive pointwise confidence intervals (CIs) for a polynomial trend. Let q_a denote the a th quantile of the standard normal distribution, then a pointwise $100(1 - a)\%$ CI for the

trend estimate is given by

$$\hat{\mu}\left(\frac{t}{T}\right) \pm q_{1-a/2} \sqrt{\text{Var}\left(\hat{\mu}\left(\frac{t}{T}\right)\right)}, \quad t = 0, \dots, T-1.$$

In practice we do not know the true variance. Instead, we have access to a consistent estimate of it on the non-boundary values of the trend estimate, because we have a consistent estimate of the local autocovariance of the time series. Thus, we can construct confidence intervals for the trend using this estimate in a rather straightforward manner. This is in contrast to nonlinear wavelet thresholding techniques, for which construction of confidence intervals is far from straightforward. In practice, we may need to regularise the covariance matrix to ensure positive definiteness: this can be achieved straightforwardly using the method of Rothman (2012), for example.

4.3. Simultaneous Confidence Intervals

We can also compute simultaneous $100(1-a)\%$ confidence intervals for the trend estimate. Denote the vector of standard errors of the trend estimate by \mathbf{v} , that is,

$$\mathbf{v} = \left(\sqrt{\text{Var}(\hat{\mu}_0)}, \sqrt{\text{Var}(\hat{\mu}_1)}, \dots, \sqrt{\text{Var}(\hat{\mu}_{T-1})} \right)^\top.$$

Recall that the critical value is calculated as the value δ that satisfies

$$1 - a = \Pr(\hat{\boldsymbol{\mu}} - \delta \mathbf{v} \leq \boldsymbol{\mu} \leq \hat{\boldsymbol{\mu}} + \delta \mathbf{v}) = 1 - 2 \Pr(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}} > \delta \mathbf{v}),$$

using the symmetry and continuity properties of the multi-variate normal distribution. Therefore, δ is calculated to satisfy

$$\Pr(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}} > \delta \mathbf{v}) = a/2.$$

Assuming that the true spectral structure is known (in practice we have an asymptotically unbiased, consistent estimate), then

$$\boldsymbol{\mu} - \hat{\boldsymbol{\mu}} \sim \mathcal{N}(\mathbf{0}, R\Sigma R^\top).$$

Hence we calculate δ as the value such that the probability of this multi-variate normal random variable exceeding $\delta \mathbf{v}$ is equal to $a/2$. Thus, within our framework, we can straightforwardly calculate an estimate and confidence intervals for the trend. In turn this could be used for hypothesis testing of the presence of trend. Although we have assumed a low-order polynomial for the trend, the simulation results demonstrate that even if we depart from this assumption, the trend estimator is still able to perform well. This is due to the fact that a polynomial can serve as an accurate approximation to a more complicated trend function, and the ability of smooth wavelets to zero out the trend, as explained in Section 3.3.

4.4. Boundary Handling

Using the results from Section 3, we can obtain an unbiased, consistent spectral estimator for the non-boundary coefficients when the time series exhibits a smooth trend. However, this will not be true for the boundary coefficients as they will still contain the trend of the time series. This behaviour is not bias but simply a consequence of the way the boundary coefficients are handled when performing the non-decimated wavelet transform. The two most common methods for boundary handling are to reflect the data at the boundaries or to periodise it – see section

2.8 of Nason (2008) or section 4.11 of Percival and Walden (2006) for more discussion. These two methods can be used effectively in a zero mean time series, but are less suitable when the time series displays a trend.

We can adapt the reflective boundary handling to produce better estimates of the EWS in the coarser scales. We use reflection of the time series as this ensures that the spectral properties of the time series in the reflected part evolve slowly as we move from the end of the time series past into the reflected, extended part. This is more appealing than periodising the time series, as spectral properties at one boundary of the series may not be as similar to those at the other boundary, due to the non-stationarity of the time series. Starting with a time series of length T , we form a time series of length $4T$, whereby the original data are centred within the extended time series. That is, we add $3T/2$ data points to the beginning, and end, of the original series.

We note that in most applications, the standard practice of boundary handling involves repeating the series half of the length in each direction, giving a series of length $2T$. However, we have found in our empirical analysis that using a series of length $4T$ offers stronger practical performance. One of the reasons for this is that in the coarser scales of the wavelet transform, the wavelet filter is often already longer than the length of the time series. Hence, in these coarse scales, the wavelet coefficients are formed only using boundary handling, which causes issues due to the trend of the time series. Therefore, extending the time series to length $4T$ ensures that a larger number of wavelet coefficients are not affected by the trend. This enables more accurate estimation in the coarser scales for the evolutionary wavelet spectrum.

We illustrate the boundary handling procedure for the right-hand boundary of the data with Figure 1. An analogous procedure is used to add data to the beginning of the series, and so here we focus on an illustration of adding data to the end of the series. In (a), we see a polynomial trend of length $T = 1024$ (in the example, we use a deterministic series for illustrative purposes). We wish to extend the data at the right boundary in such a way that we continuously extend the trend to minimise the boundary effect caused when calculating the boundary non-decimated wavelet coefficients. To achieve this, we first (b) reflect the entire time series at the boundary, given by the dashed vertical line, motivated by the reasoning above, obtaining a series of length $2T = 2048$. Next, we multiply the boundary series (data points 1025–2048) by -1 , to ensure that the gradient of the reflected trend part is similar to that at the edge of the time series (c). Finally in (d), we must add some value, c , to this reflected part, to align the edge of the time series with the boundary part. The value c used to align the boundary part with the original series is calculated using a simple pre-estimate of the trend function, obtained via a polynomial regression. From this we obtain a pre-estimate for the value of the trend at the boundary. The value c is given by two times this value, as this reverses the effect caused by multiplying the boundary series by -1 , realigning the series.

Having performed this step, we can repeat this again to add another $T/2 = 512$ data points on the right side, giving a series of length $5T/2 = 2560$. We then repeat the whole procedure for the left side boundary, adding $3T/2 = 1536$ data points to the start of the series, giving the final series of length $4T = 4096$. In practice, we can apply the procedure when performing trend estimation, to produce a more reliable estimate at the boundaries of the time series. In this case, we no longer have strictly valid confidence intervals at the boundary locations. However, we have found that performing the boundary handling step significantly improves estimation at the boundaries by reducing the variability of the estimate.

Finally, we observe that this procedure can be easily adapted to time series whose length is non-dyadic. A time series of general length n is extended to one of dyadic length d , where $d = 4 \times 2^{\lceil \log_2 n \rceil}$. Indeed, the data example used in Section 6 is non-dyadic.

5. SIMULATION STUDY

To illustrate the ability to incorporate a smooth trend within the LSW framework, we perform a simulation study. In the simulation, we define three different evolutionary wavelet spectra, shown in Figure 2 and explicitly defined in the Supporting Information. Spectrum S^1 , studied in Nason (2008), displays coarse-scale, slowly-evolving sinusoidal behaviour with a fine-scale burst in power at time point 800; spectrum S^2 is a concatenation of moving average processes and contains power moving from fine to coarser scales, and was examined in Nason *et al.* (2000); spectrum S^3 contains slowly-evolving power at fine scales.

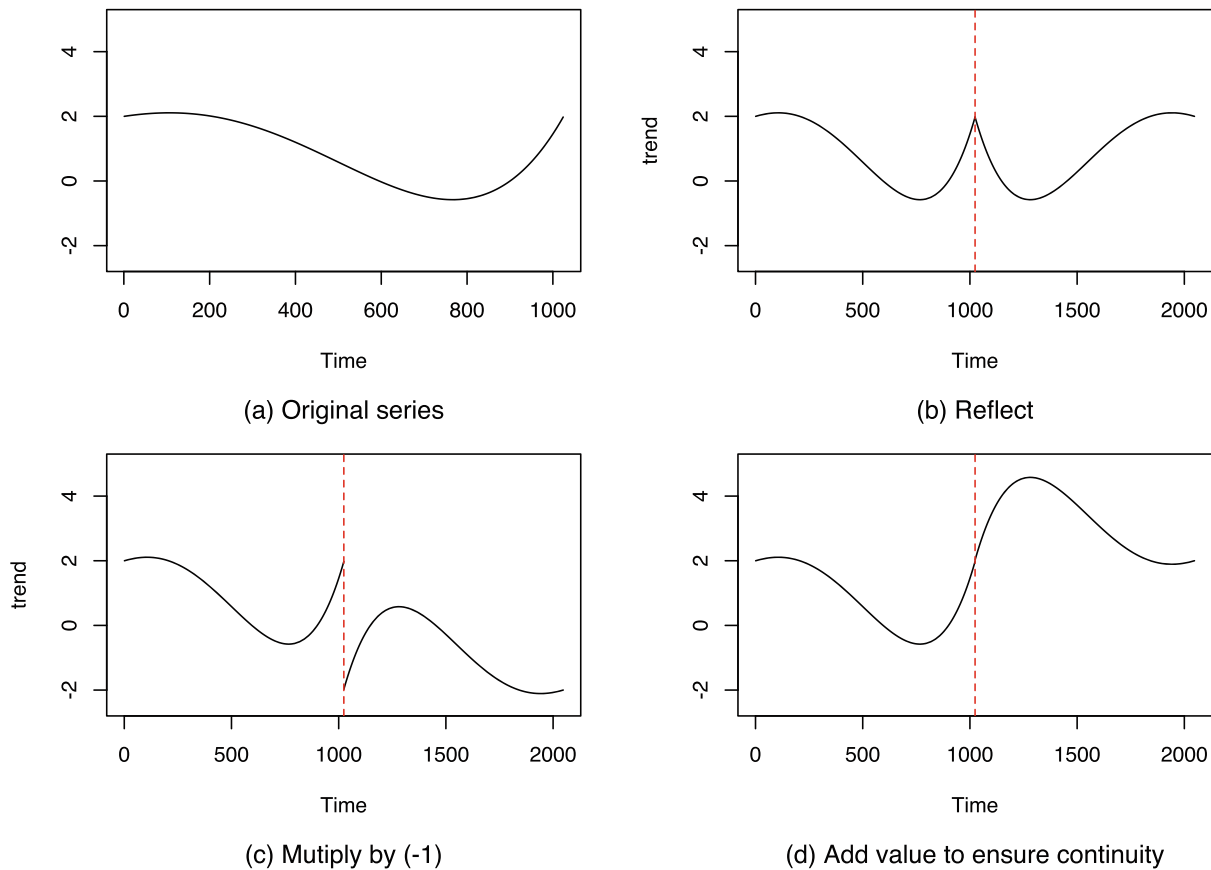


Figure 1. Boundary handling process for an example trend

We simulate time series $\{X_t\}_{t=0}^{T-1}$ of length $T = 2^{10} = 1024$ from LSW processes with those spectra, using Gaussian innovations. Depending on the simulation, different wavelets were used, to highlight the observations from Section 3. The simulations were performed in R (R Core Team (2019)) using the `wavethresh` package of Nason (2016). For each spectrum, 100 LSW processes with different trends added on were simulated.

We examine the effects of including a trend across three scenarios, corresponding to Sections 3.1, 3.2, and 3.3 of this article. In the first scenario, we add low-order trends to LSW processes simulated using a Daubechies EP wavelet with four vanishing moments, which corresponds to the setting of Section 3.1. In the second, contained within the Supporting Information, we add low-order trends to LSW processes simulated using Haar wavelets, which corresponds to Section 3.2. In the final set of simulations, we add non-polynomial trends to LSW processes simulated using the Daubechies EP4 wavelet, corresponding to Section 3.3.

To assess the spectral estimation performance, for each realisation, the un-smoothed estimate of the EWS was calculated, which was then used to obtain an averaged estimate for the EWS across the 100 realisations. In alignment with the discussion in Sections 3 and 4, we correct the EWS across the finest seven scales and use the boundary handling procedure, which ensures that the boundary effects of the trend are minimised. Furthermore, in the Supporting Information, we also provide numerical evidence to show that using a smoother wavelet for spectral estimation is beneficial.

We also assess the trend estimation performance by reporting the averaged mean squared error of the trend estimate across the 100 realisations. The trend estimate is calculated using the Daubechies LA4 wavelet, with

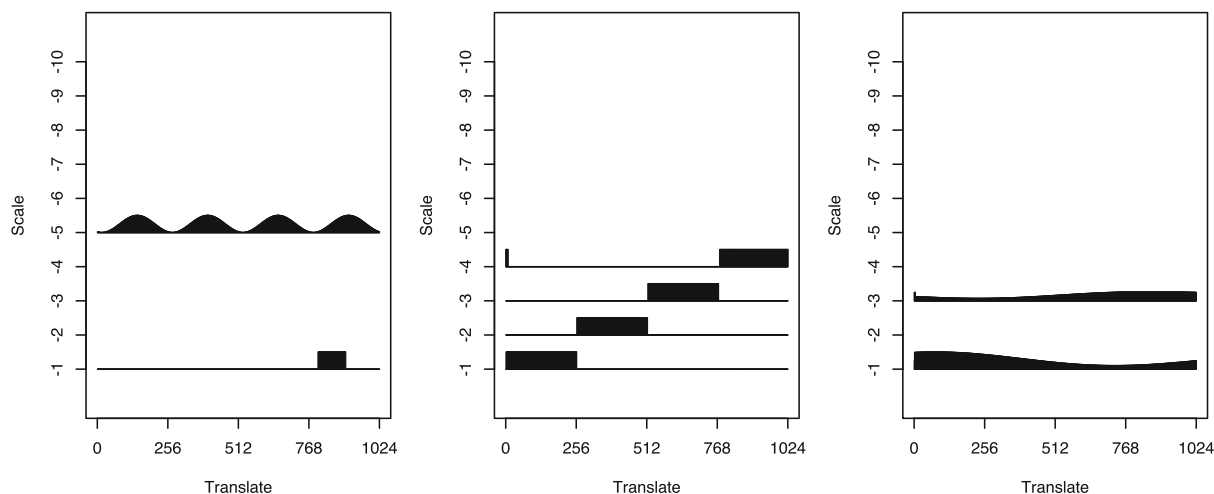


Figure 2. Spectra used in the simulation study. Left: S^1 , sinusoid with ‘burst’; centre: S^2 , concatenated moving average process; right: S^3 , slowly-evolving fine-scale power

boundary handling applied. Our linear wavelet thresholding method is referred to as LSWT in the comparison tables. We compare our method to three other trend estimation methods. We first compare to the approach of von Sachs and MacGibbon (2000), which utilises a local MAD estimator in a wavelet thresholding procedure. No code is publicly available for this method, and so we have implemented the method utilising `wavethresh` and following the description of the computation of the threshold in section 2.5 of von Sachs and MacGibbon (2000). This method is referred to as VSMWT in the comparison tables. Secondly, we also compare to the spline-based method using the R function `smooth.spline`, where the smoothing parameter is chosen via cross validation. Lastly, we compare to a local polynomial regression estimator, calculated using the `loess` function in R. We use the default settings and a local quadratic polynomial for fitting the trend.

5.1. Low-Order Polynomial Trend, High-Order Generating Wavelet

We add various low-order polynomial trends to LSW processes generated by the three different spectra. In particular, we use three different trends μ_t^1 , μ_t^2 , and μ_t^3 (where $\mu_t = \mu(t/T)$); polynomials of degree 1, 2, and 3 given by

$$\mu_t^1 = 4t, \quad \mu_t^2 = 12t^2 - 12t, \quad \mu_t^3 = 32t^3 - 48t^2 + 22t - 3. \quad (5.1)$$

A plot of example realisations from each trend and spectrum scenario is shown in Figure 3. We use the EP4 wavelet to analyse the series, the same wavelet that generates the LSW processes, which is a high enough order to remove the polynomial trends.

In each of the simulations, we compare the estimates of the EWS obtained when trends are added, to the estimate of the EWS obtained in the absence of any trend (‘None’ in Table I), obtained in the standard way using the `ewspec` command within `wavethresh`. The averaged estimates for each spectrum across the four trend scenarios are shown in the Supporting Information. Table I contains the mean squared error across the four trends and three spectra, which shows that the trend has negligible impact. In fact, for the third spectrum, the mean squared error is actually lower when a trend is added. From the simulation results, we see that there is no discernible difference in the estimation quality between the no trend case and the case when a trend is added.

In Table II, we report the average mean squared error for the trend estimates, calculated over 100 realisations, with standard deviation given in brackets. Bold values in the table indicate the lowest mean squared error for

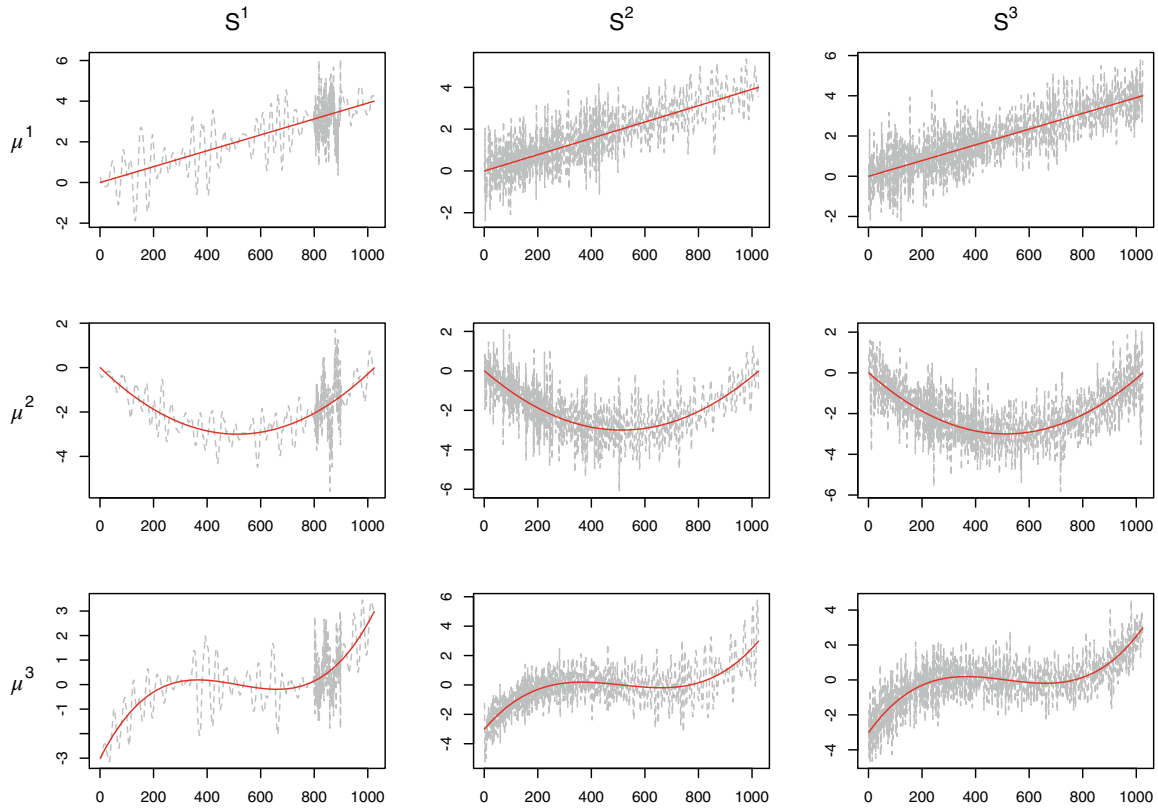


Figure 3. Example realisations from each trend and spectrum scenario. Dashed line shows time series with true underlying trend shown in solid line. Left: spectrum S^1 , sinusoid with ‘burst’; centre: S^2 , concatenated moving average process; right: S^3 , slowly-evolving fine-scale power

Table I. Mean squared error (in units $\times 10^{-3}$) of the averaged spectral estimate over 100 realisations

Trend	Spectrum 1	Spectrum 2	Spectrum 3
None	3.13	4.88	1.87
Linear	3.13	5.50	1.79
Quadratic	3.13	5.50	1.79
Cubic	3.13	5.50	1.79

each trend and spectrum combination. From the table, we see that our method performs well and gives a low mean squared error across the various scenarios. The method consistently outperforms the competing methods, with only one scenario in which it does not give the lowest mean squared error. In particular, we outperform the wavelet-based method of von Sachs and MacGibbon (2000), which is in part due to the lack of proper boundary handling. As we might expect, the local polynomial method also performs well. The spline-based method performs well for spectrum 3, which is likely due to the fact this process contains only weak autocorrelation.

5.2. Non-Polynomial Trend, High-Order Analysing Wavelet

Lastly, we show simulation results when the trend is less smooth. In this simulation, the spectra are simulated using the Daubechies EP4 wavelet, and we add three non-polynomial trends to the series. The three trends are a

Table II. Average mean squared error and standard deviation of trend estimate over 100 realisations

Trend	Spectrum	Method			
		LSWT	VSMWT	Spline	Loc. Poly.
Linear	1	7.10×10^{-4} (4.66×10^{-4})	0.273 (0.094)	0.438 (0.115)	1.59×10^{-3} (1.75×10^{-3})
	2	5.32×10^{-4} (7.71×10^{-4})	0.226 (0.064)	0.154 (0.087)	7.32×10^{-4} (6.45×10^{-4})
	3	1.01×10^{-4} (1.31×10^{-4})	0.161 (0.041)	2.26×10^{-5} (2.99×10^{-5})	4.63×10^{-4} (3.37×10^{-4})
Quadratic	1	7.05×10^{-4} (4.50×10^{-4})	0.284 (0.101)	0.453 (0.118)	1.51×10^{-3} (1.48×10^{-3})
	2	4.61×10^{-4} (6.91×10^{-4})	0.208 (0.061)	0.154 (0.086)	6.79×10^{-4} (7.77×10^{-4})
	3	9.69×10^{-5} (1.13×10^{-4})	0.157 (0.039)	1.55×10^{-3} (8.81×10^{-5})	4.87×10^{-4} (4.15×10^{-4})
Cubic	1	7.18×10^{-4} (4.49×10^{-4})	0.285 (0.098)	0.466 (0.122)	9.07×10^{-3} (3.14×10^{-3})
	2	6.77×10^{-4} (8.41×10^{-4})	0.228 (0.070)	0.161 (0.092)	6.79×10^{-3} (1.52×10^{-3})
	3	1.07×10^{-4} (1.41×10^{-4})	0.177 (0.044)	2.57×10^{-3} (8.17×10^{-4})	6.85×10^{-3} (1.25×10^{-3})

Table III. Mean squared error (in units $\times 10^{-3}$) of the averaged spectral estimate over 100 realisations, with trend scenarios in Equation (5.2)

Trend	Spectrum 1	Spectrum 2	Spectrum 3
None	2.60	5.00	1.80
Sine	2.93	4.34	1.53
Logistic	2.94	4.35	1.54
Exponential	2.94	4.35	1.54

sinusoidal, logistic, and exponential trend, defined as follows:

$$\mu_t^s = \sin(2\pi t) + \cos(\pi t), \quad \mu_t^l = \frac{4}{1 + \exp(4 - 7 \log 4t)}, \quad \mu_t^e = 2 \exp(3t). \quad (5.2)$$

The sinusoidal and logistic trends are Lipschitz continuous; however, the exponential trend is not, and further it is not Hölder continuous. Realisations from the simulated time series can be found in the Supporting Information. We use an EP10 wavelet to compute the spectral estimate, which is the smoothest available wavelet in this family within `wavethresh` and will therefore be the best at removing the effect of the trend.

The Supporting Information contains plots of the averaged spectrum again showing that then trend has not biased the spectral estimate. In Table III we compare the mean squared error for the averaged spectral estimates, and find similar results to that in Section 5.1. The error when including a trend is higher for spectrum S^1 , while lower for the other two spectra. These results confirm the finding in Section 3.3, that including a smooth non-polynomial trend will cause little bias to the estimate of the EWS. They also serve to highlight that using a smoother wavelet can achieve a more accurate spectral estimate.

Finally, we assess the trend estimation performance. Example trend estimates for each of the three trend scenarios for spectrum S^3 are shown in the Supporting Information. In Table IV, we report the average mean squared error for the methods calculated over 100 realisations, with standard deviation given in brackets. Despite the fact that the trends are non-polynomial, the wavelet-based estimator still performs well. Our method is consistently

Table IV. Average mean squared error and standard deviation of trend estimate over 100 realisations

Trend	Spectrum	Method			
		LSWT	VSMWT	Spline	Loc. Poly.
Sine	1	7.17×10^{-4} (4.93×10^{-4})	0.289 (0.108)	0.467 (0.127)	3.26×10^{-3} (2.49×10^{-3})
	2	6.35×10^{-4} (8.61×10^{-4})	0.201 (0.064)	0.152 (0.097)	1.88×10^{-3} (9.37×10^{-4})
	3	1.19×10^{-4} (1.57×10^{-4})	0.157 (0.0387)	1.49×10^{-3} (5.76×10^{-4})	1.65×10^{-3} (5.47×10^{-4})
Logistic	1	2.63×10^{-3} (1.08×10^{-3})	0.284 (0.110)	0.451 (0.124)	1.58×10^{-2} (3.65×10^{-3})
	2	2.52×10^{-3} (1.73×10^{-3})	0.214 (0.061)	0.148 (0.085)	1.41×10^{-2} (1.37×10^{-3})
	3	2.15×10^{-3} (7.56×10^{-4})	0.158 (0.039)	2.32×10^{-3} (1.41×10^{-4})	1.38×10^{-2} (7.94×10^{-4})
Exponential	1	9.45×10^{-4} (4.91×10^{-4})	0.386 (0.088)	0.446 (0.098)	2.99×10^{-2} (6.36×10^{-3})
	2	8.61×10^{-4} (9.94×10^{-4})	0.347 (0.072)	0.154 (0.084)	2.27×10^{-2} (3.62×10^{-3})
	3	2.54×10^{-4} (1.54×10^{-4})	0.294 (0.040)	3.61×10^{-3} (8.57×10^{-4})	2.29×10^{-2} (1.85×10^{-3})

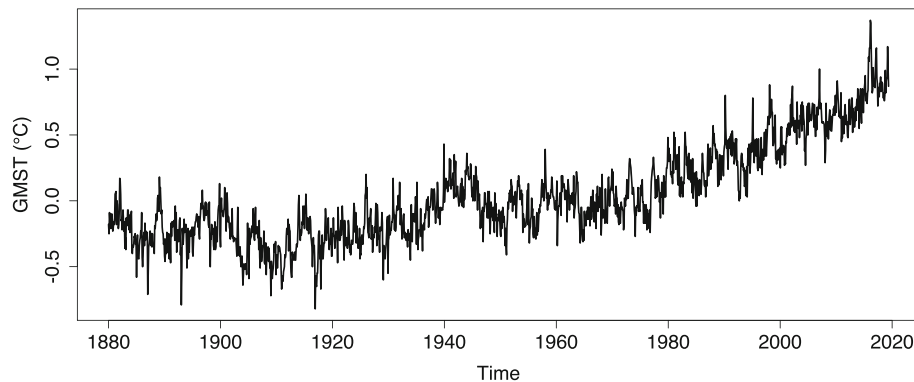


Figure 4. Global average monthly sea temperature data

the best performing across all of the trend and spectrum scenarios. Overall, the simulation results show that the linear wavelet thresholding estimator can perform well in a variety of trend settings in the presence of locally stationary errors. In particular, we observe that the linear wavelet thresholding estimator outperforms the nonlinear thresholding method of von Sachs and MacGibbon (2000).

6. GLOBAL MEAN SEA TEMPERATURE DATA APPLICATION

This section highlights the ability to model time series with time-varying first- and second-order structure on the monthly global mean sea temperature dataset (GMST) version 4.6 from the Goddard Institute for Space Studies (GISS) Surface Temperature Analysis (GISTEMP, NASA, 2020). The data are available at <http://data.giss.nasa.gov/gistemp>, and are shown in Figure 4. The GMST time series exhibits periods of warming separated by a long pause from approximately the mid-1940s to the mid-1970s (Kellogg, 1993) and possibly a second, shorter one, since the late 1990s/early 2000s, although this is highly debated (Trenberth, 2015; Beaulieu and Killick, 2018). The series should be well-suited to our methodology, as it appears to contain both long-term trends and short-term variability that may be non-stationary.

In general, the series exhibits an upward trend; therefore, using a high-order wavelet should remove the trend from the wavelet coefficients and allow us to obtain an estimate for the spectrum. In addition, we might reasonably expect there to be strong dependency within the data: as with many environmental time series, this may be a result of certain physical phenomena.

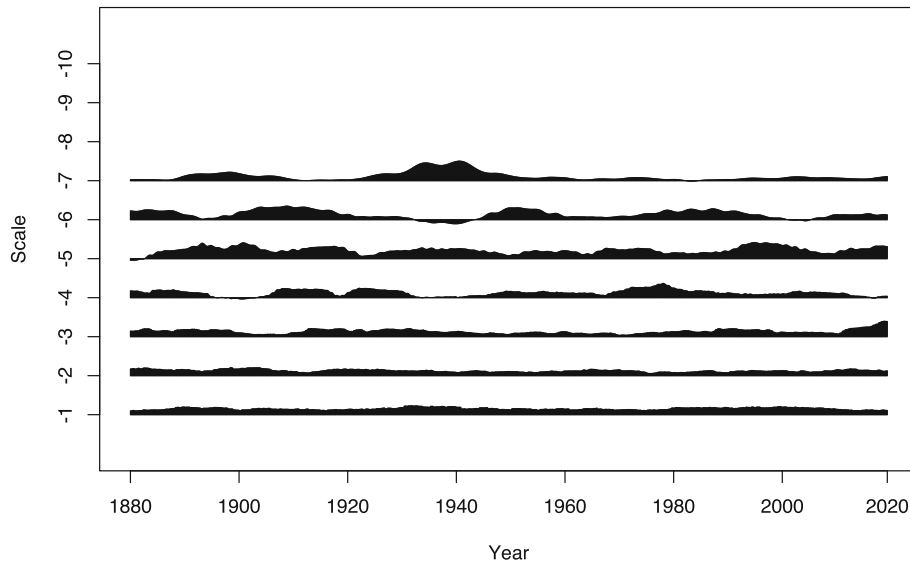


Figure 5. EWS estimate of the GMST series. Each level is scaled individually for clarity

In Figure 5, we see the spectral estimate for the GMST time series. This is estimated using the Daubechies LA10 wavelet, while the raw wavelet periodogram is smoothed using a running mean smoother with bin width 120, corresponding to 10 years. Each level is scaled individually for clarity. Additionally, we have used the boundary handling procedure outlined to properly handle the endpoints of the series. Motivated by the discussion in Section 3.3, we only calculate the EWS estimate across the finest seven scales, which corresponds to 70% of the scales. The choice of the LA10 wavelet is motivated by the discussion in Section 1 of the Supporting Information that suggests the benefit of using smoother wavelets. A more systematic approach to wavelet choice could be devised by adapting the work of Cardinali and Nason (2017). This method uses a penalised likelihood-based approach to select, for a given wavelet, an appropriate set of wavelet packets to use in the analysis of an LSW packet process.

We see that the spectral behaviour is generally slowly evolving with peaks at various locations: for example, note the burst of power that occurs in scale -4 around 1910. We also see an appearance of power in scale -7 at approximately 1940 and scale -6 slightly later, corresponding to the proposed period of pause in warming.

Next, we investigate how the autocorrelation structure of the GMST time series evolves over time. A common assertion in the Climate literature is that the autocorrelation within the Global Mean Sea Temperature Data is constant across time. Here we use the estimate of the spectrum to examine the autocorrelation function at different periods in the series. In Figure 6, we see a plot of the autocorrelation function of the time series at six different time-points across 25-year intervals, starting at 1890 and ending at 2015. Commonly, the error structure of climate change data is modelled using a low-order AR process, for example, Beaulieu and Killick (2018) and Karl *et al.* (2000). In similar analyses on gridded datasets, Lenton *et al.* (2017) find an increase in variance and autocorrelation in global surface sea temperature data, while Boulton and Lenton (2015) find a ‘reddening’ – a gradual increase in autocorrelation – in North Pacific surface sea temperature data.

Our analysis suggests that overall, the low-order AR assumption may be appropriate at certain time points; however, the degree of autocorrelation changes over time, and the assumption of a stationary AR model may not be appropriate. From the plot, the non-stationary nature of the second-order structure can be seen. In 1890, 1915, 1965, 1990, and 2015, the autocorrelation exhibits a similar decay behaviour, falling within the upper and lower confidence lines after approximately lag 60, corresponding to a time period of 5 years. However, at time 1940, given by the dark solid line, we see much stronger autocorrelation, exhibiting a shape that is not well represented

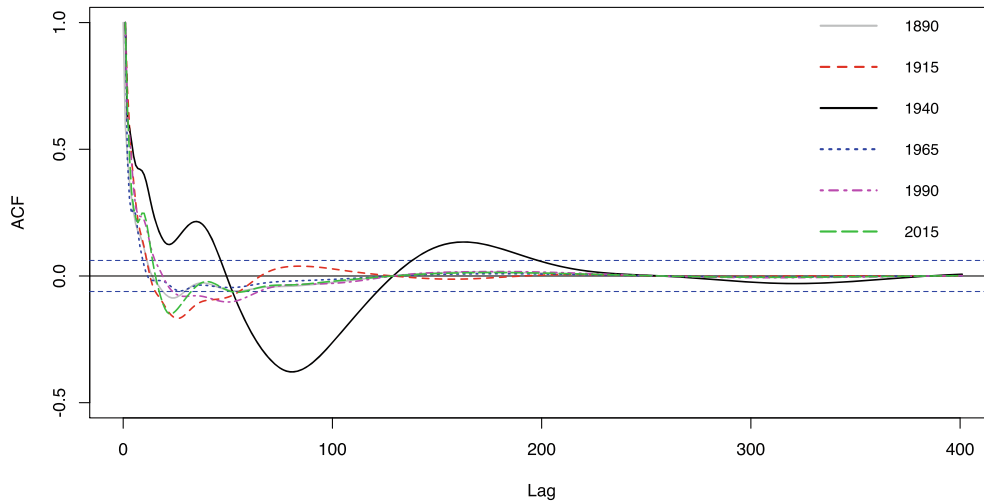


Figure 6. Autocorrelation of the GMST series across six time points, up to lag 400

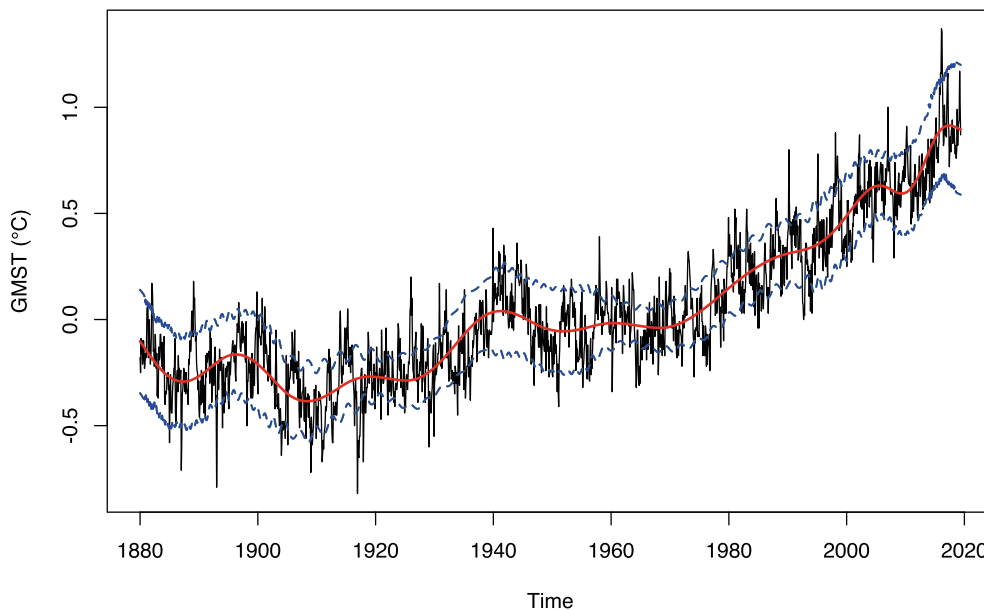


Figure 7. Global average monthly sea temperature with trend estimate (solid line) and 95% pointwise confidence intervals (dashed lines)

by a low-order AR model. This increase in autocorrelation can be attributed to the appearance of power at the coarsest scale of the spectrum, as seen in Figure 5. Moreover, our analysis shows that in the period between 2005 and 2020 – which is debated to be another warming pause – the variance of the series increases by roughly 63% from 0.0153 to 0.0240. This may explain the appearance of a pause due to increased variability, as many methods assume the series to be second-order stationary. Note that the trend estimate does suggest the appearance of a pause in this time period; however, because this occurs at the boundary, the estimate should be used with caution.

Lastly, in Figure 7, we see the estimate for the trend shown in the solid line, with the 95% pointwise confidence interval shown in the dashed lines. This was calculated with the Daubechies EP10 wavelet, using the methodology

described in Section 4, using the boundary extension procedure. Note that the estimate near the endpoints is therefore dependent on this boundary handling and inference should be carefully considered. The EP10 wavelet was selected using advice from a domain expert, by comparing several choices and their resulting fitted trends. Overall, our estimate agrees with many previous analyses of the series: we see a long pause between approximately the 1940s and 1970s, as well as a sustained period of warming post-1970 to the early 2000s.

7. CONCLUDING REMARKS

In this article, we have addressed the commonly encountered problem of modelling time series that display first- and second-order non-stationarity. Our model extends the locally stationary wavelet model of Nason *et al.* (2000), used for modelling second-order non-stationarity, to include a non-stationary, smooth mean structure. Within our framework, we can model both smooth deterministic trends and slowly-evolving second-order behaviour, allowing for a flexible model that can be applicable to time series in many applications.

We have established results concerning the theoretical properties of the estimators of the first- and second-order quantities of interest, which are shown to have good consistency properties and strong practical performance. Furthermore, we have found using a smoother wavelet to be useful even in the case of a zero-mean time series. To ensure the applicability of our methodology, we propose a new method for handling the boundary of a time series when computing the discrete non-decimated wavelet transform, which incorporates data of any length and is suitable for data that display trend behaviour. In our framework, the results and methodology are suitable for polynomial and smooth trends; however, a polynomial or smooth function is a good approximation to many trends of practical interest. In practice, the techniques described here will work well when we depart from the assumptions of the model, as evidenced within the simulation study.

We demonstrated our methodology on a data example that shows that our model can be successfully applied to a variety of data sets that exhibit time-varying first- and second-order structure. In our analysis of the global mean sea temperature dataset, we find that the second-order structure of the series varies over time, while observing similar findings as previously described of a general increasing trend of global mean sea temperature.

ACKNOWLEDGEMENT

Euan T. McGonigle gratefully acknowledges financial support from EPSRC and Numerical Algorithms Group Ltd. via The Smith Institute i-CASE award No. EP/R511997/1.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are publicly available. The global mean sea temperature data set is available to download at <http://data.giss.nasa.gov/gistemp>.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

REFERENCES

- Beaulieu C, Killick R. 2018. Distinguishing trends and shifts from memory in climate data. *Journal of Climate* **31**(23):9519–9543.
- Beran J, Feng Y. 2002. SEMIFAR models: A semiparametric approach to modelling trends, long-range dependence and nonstationarity. *Computational Statistics Data Analysis* **40**(2):393–419.
- Boulton CA, Lenton TM. 2015. Slowing down of north pacific climate variability and its implications for abrupt ecosystem change. *Proceedings of the National Academy of Sciences* **112**(37):11496–11501.
- Brillinger DR. 1994. Some river wavelets. *Environmetrics* **5**(3):211–220.

- Cardinali A, Nason GP. 2017. Locally stationary wavelet packet processes: Basis selection and model fitting. *Journal of Time Series Analysis* **38**(2):151–174.
- Chapman J-L, Eckley IA, Killick R. 2020. A nonparametric approach to detecting changes in variance in locally stationary time series. *Environmetrics* **31**(1):e2576.
- Coifman RR, Donoho DL. 1995. *Translation-invariant de-noising, Wavelets and Statistics* Springer, New York; 125–150.
- Craigmile PF, Guttorp P, Percival DB. 2004. Trend assessment in a long memory dependence model using the discrete wavelet transform. *Environmetrics* **15**(4):313–335.
- Craigmile PF, Guttorp P, Percival DB. 2005. Wavelet-based parameter estimation for polynomial contaminated fractionally differenced processes. *IEEE Transactions on Signal Processing* **53**(8):3151–3161.
- Dahlhaus R. 2000. A likelihood approximation for locally stationary processes. *Annals of Statistics* **28**(6):1762–1794.
- Dahlhaus R. 2012. *Locally stationary processes*. In *Handbook of Statistics*, Vol. 30, Elsevier, Netherlands; 351–413.
- Dahlhaus R, Neumann MH. 2001. Locally adaptive fitting of semiparametric models to nonstationary time series. *Stochastic Processes and Their Applications* **91**(2):277–308.
- Dalla V, Giraitis L, Robinson PM. 2020. Asymptotic theory for time series with changing mean and variance. *The Journal of Econometrics* **219**(2):281–313.
- Donoho DL, Johnstone IM. 1995. Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association* **90**(432):1200–1224.
- Donoho DL, Johnstone IM, Kerkycharian G, Picard D. 1995. Wavelet shrinkage: Asymptopia? *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **57**(2):301–337.
- Fryzlewicz P. 2005. Modelling and forecasting financial log-returns as locally stationary wavelet processes. *Journal of Applied Statistics* **32**(5):503.
- Fryzlewicz P, Nason GP. 2006. Haar–Fisz estimation of evolutionary wavelet spectra. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **68**(4):611–634.
- Fryzlewicz P, Van Belleghem S, von Sachs R. 2003. Forecasting non-stationary time series by wavelet process modelling. *Annals of the Institute of Statistical Mathematics* **55**(4):737–764.
- Fryzlewicz PZ. 2003. *Wavelet techniques for time series and Poisson data*. PhD thesis, University of Bristol.
- Gott AN, Eckley IA. 2013. A note on the effect of wavelet choice on the estimation of the evolutionary wavelet spectrum. *Communications in Statistics Simulation and Computation* **42**(2):393–406.
- Härdle W. 1990. *Applied Nonparametric Regression* Cambridge University Press, Cambridge.
- Hargreaves JK, Knight MI, Pitchford JW, Oakenfull RJ, Chawla S, Munns J, Davis SJ. 2019. Wavelet spectral testing: Application to nonstationary circadian rhythms. *The Annals of Applied Statistics* **13**(3):1817–1846.
- Hargreaves JK, Knight MI, Pitchford JW, Oakenfull RJ, Davis SJ. 2018. Clustering nonstationary circadian rhythms using locally stationary wavelet representations. *Multiscale Modelling and Simulation* **16**(1):184–214.
- Kallache M, Rust HW, Kropp J. 2005. Trend assessment: Applications for hydrology and climate research. *Nonlinear Processes in Geophysics* **12**(2):201–210.
- Karl TR, Knight RW, Baker B. 2000. The record breaking global temperatures of 1997 and 1998: Evidence for an increase in the rate of global warming? *Geophysical Research Letters* **27**(5):719–722.
- Kellogg WW. 1993. An apparent moratorium on the greenhouse warming due to the deep ocean. *Climatic Change* **25**(1): 85–88.
- Killick R, Eckley IA, Jonathan P. 2013. A wavelet-based approach for detecting changes in second order structure within nonstationary time series. *Electronic Journal of Statistics* **7**:1167–1183.
- Lenton TM, Dakos V, Bathiany S, Scheffer M. 2017. Observed trends in the magnitude and persistence of monthly temperature variability. *Scientific Reports* **7**(1):5940.
- McGonigle ET, Killick R, Nunes MA. 2021. Modelling time-varying first and second-order structure of time series via wavelets and differencing. *arXiv preprint arXiv:2108.07550v1*.
- NASA. 2020. GISS Surface Temperature Analysis. <http://data.giss.nasa.gov/gistemp>
- Nason G. 2008. *Wavelet Methods in Statistics with R* Springer Science & Business Media, Berlin.
- Nason G. 2013. A test for second-order stationarity and approximate confidence intervals for localized autocovariances for locally stationary time series. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **75**(5):879–904.
- Nason G. 2016. *wavethresh: Wavelets Statistics and Transforms* (R package version 4.6.8) University of Bristol, Bristol.
- Nason GP, von Sachs R, Kroisandt G. 2000. Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **62**(2):271–292.
- Neumann MH, von Sachs R. 1995. *Wavelet thresholding: Beyond the Gaussian IID situation*. In *Wavelets and Statistics*, Springer, New York; 301–329.
- Park T, Eckley IA, Ombao HC. 2014. Estimating time-evolving partial coherence between signals via multivariate locally stationary wavelet processes. *IEEE Transactions on Signal Processing* **62**(20):5240–5250.
- Percival DB, Walden AT. 2006. *Wavelet Methods for Time Series Analysis* Cambridge University Press, Cambridge.
- Priestley M. 1983. *Spectral Analysis and Time Series* Academic Press, London.

- R Core Team. 2019. *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing, Vienna, Austria.
- Rothman AJ. 2012. Positive definite estimators of large covariance matrices. *Biometrika* **99**(3):733–740.
- Sanderson J, Fryzlewicz P, Jones MW. 2010. Estimating linear dependence between nonstationary time series using the locally stationary wavelet model. *Biometrika* **97**:435–446.
- Shen C. 2015. Analysis of detrended time-lagged cross-correlation between two nonstationary time series. *Physics Letters A* **379**(7):680–687.
- Stevens KN. 2013. *New Methods for Spectral Estimation with Confidence Intervals for Locally Stationary Wavelet Processes*. PhD thesis, University of Bristol.
- Trenberth KE. 2015. Has there been a hiatus? *Science* **349**(6249):691–692.
- Tunyavetchakit S. 2010. *On the optimal segment length for tapered Yule–Walker estimates for time-varying autoregressive processes*. PhD thesis, Heidelberg University.
- Van Belleghem S, von Sachs R. 2008. Locally adaptive estimation of evolutionary wavelet spectra. *The Annals of Statistics* **36**(4):1879–1924.
- Vidakovic B. 2009. *Statistical Modeling by Wavelets* John Wiley & Sons, Hoboken, NJ.
- von Sachs R, MacGibbon B. 2000. Non-parametric curve estimation by wavelet thresholding with locally stationary errors. *Scandinavian Journal of Statistics* **27**(3):475–499.
- von Sachs R, Nason GP, Kroisandt G. 1997. Adaptive estimation of the evolutionary wavelet spectrum. Technical report, University of Kaiserslautern and Bristol.
- Wilson RE, Eckley IA, Nunes MA, Park T. 2019. Dynamic detection of anomalous regions within distributed acoustic sensing data streams using locally stationary time series. *Data Mining and Knowledge Discovery* **33**(3):748–772.
- Zhang T, Wu WB. 2011. Testing parametric assumptions of trends of a nonstationary time series. *Biometrika* **98**(3):599–614.