

Improved preservation of autocorrelative structure in surrogate data using an initial wavelet step

C. J. Keylock

School of Geography, University of Leeds, Leeds, UK

Received: 28 September 2007 – Revised: 17 April 2008 – Accepted: 18 April 2008 – Published: 2 June 2008

Abstract. Surrogate data generation algorithms are useful for hypothesis testing or for generating realisations of a process for data extension or modelling purposes. This paper tests a well known surrogate data generation method against a stochastic and also a hybrid wavelet-Fourier transform variant of the original algorithm. The data used for testing vary in their persistence and intermittency, and include synthetic and actual data. The hybrid wavelet-Fourier algorithm outperforms the others in its ability to match the autocorrelation function of the data, although the advantages decrease for high intermittencies and when attention is only directed towards the early part of the autocorrelation function. The improved performance is attributed to the wavelet step of the algorithm.

1 Introduction

Surrogate data techniques were originally developed for testing hypotheses regarding the nature of non-linearity in data (Theiler et al., 1992; Theiler and Prichard, 1996), and they have been used in this way in a variety of subdisciplines in the geosciences (Ashkenazy and Tziperman, 2004; Poggi et al., 2004). However, more recently there has been a move to apply such methods for producing realisations of a phenomenon for testing model performance (Venema et al., 2006c), or increasing the number of realisations for an experiment (Angelini et al., 2005), and it is the latter application that is of primary concern in this study. The most popular method for deriving such surrogates is a Fourier-based method due to Schreiber and Schmitz (1996) known as the Iterated Amplitude Adjusted Fourier Transform (IAAFT) technique.

With this technique, the surrogates are constrained to the values of the original series as well as their behaviour in the Fourier domain. One application of this type of test in geophysics is to use such surrogates for studying the intermittency properties of turbulence (Basu et al., 2007). Because these surrogates only preserve the histogram and Fourier spectrum, intermittency effects are removed, permitting a test for the presence of intermittency in the original data.

Venema et al. (2006b) tested this algorithm and several other surrogate-generating techniques against various geophysical datasets and concluded that the IAAFT method was more effective than methods that were solely based in the Fourier domain or fractal-based tools that gave a simplified model for the Fourier behaviour of the signal. However, there are convergence issues with the IAAFT algorithm and while this can be dealt with using a method such as simulated annealing (Schreiber, 1998), this is generally prohibitively slow. Venema et al. (2006a) proposed a stochastic variant of the IAAFT algorithm (the SIAAFT) where only a subset of the original values is matched to the data in an initial rank-ordering step. After convergence to a local minimum, the original IAAFT algorithm is applied to recover a signal with the same amplitude distribution as the original. The stochastic first stage helps the algorithm find a local minimum that is a more effective representation of the Fourier characteristics of the original signal by applying the amplitude constraint more gradually. The aim of this paper is to show how a recently proposed method for surrogate generation (Keylock, 2006) is of use for improving the convergence of the IAAFT method. As with the SIAAFT, this is a two step algorithm, but the first stage is based on a wavelet transform (hence, the WIAAFT). This method and the original IAAFT algorithm are described in Sect. 2, while the data and testing procedure are presented in Sect. 3. A comparison of the performance of the three methods is undertaken in Sect. 4.



Correspondence to: C. J. Keylock
(c.j.keylock@leeds.ac.uk)

2 The original and wavelet-based IAAFT algorithms

The original IAAFT algorithm is based on the following steps:

- (a) Take the Fourier transform of the original data series and store the amplitudes;
- (b) Make a random sort of the values in the original data series;
- (c) Fourier transform the random sort and replace the amplitudes with those from (a), while retaining the phases from the random sort;
- (d) Invert the Fourier transform to recover a surrogate data series and use rank-order matching to map the values of the original series onto the new series;

Step (d) will have degraded the Fourier representation of the surrogate series so stages (c) and (d) are repeated until no further re-ordering occurs, with (c) applied to the result from (d) at each iteration rather than the random sort. The SIAAFT algorithm of Venema et al. (2006a) modified step (d) in the first stage of their algorithm by only matching selected (perhaps randomly) values of the original data to the Fourier-transformed data. The alternative approach tested here supplements the IAAFT algorithm with an initial wavelet step.

The wavelet transform of a signal gives a decomposition both in terms of time/space as well as frequency (see Mallat, 1999, for a review of wavelet analysis methods). Hence, it was used by Keylock (2006) to generate surrogates that preserve the local mean and variance of a non-intermittent signal in the surrogates, while still randomising the intermittency or regularity of the signal (defined in Sect. 3). In general, in the geophysics literature, two types of wavelet transform are used: continuous, and discrete. The latter is extremely useful for data compression purposes because it is invertible, the dynamic range of the coefficients within a scale is less than for the whole signal and, for a signal of length N , it is defined over J scales, where $2^J = N$, and as the scale $j=1, \dots, J$ increases by 1, the number of coefficients decreases by ~ 2 (depending on the form of the mother wavelet). Wavelet-based surrogate algorithms such as Angelini et al. (2005) are based on the discrete transform. The continuous transform, while slower and more redundant, is often preferable for data analysis because N coefficients are retained at each scale. Furthermore, wavelet filters based on the Gaussian distribution and its derivatives, which have useful properties in wavelet analysis, are admissible in the continuous framework, allowing the development of, for example, modulus maxima based methods for multifractal analysis (Muzy et al., 1991) or estimates of intermittency (Nicolleau and Vassiliocos, 1999). However, the wavelet coefficients are not robust to translation of the data and may also lack exact invertibility due to aliasing errors (Simoncelli et al., 1992).

Consequently, a useful wavelet transform in this respect is the stationary wavelet transform or Maximal Overlap Discrete Wavelet Transform (MODWT) of Percival and Walden (2000). This transform is invertible while also producing N coefficients per scale (an undecimated transform). In addition, a circularly shifted version of a signal will have the same coefficients as the original signal with the MODWT, which is not the case for continuous methods, and analysis using this wavelet transform is less sensitive to the exact form of the wavelet used than continuous wavelet analysis. These properties mean that the MODWT is very useful for wavelet covariance and correlation analyses (Whitcher et al., 2000), as well as a tool for data visualisation (Keylock, 2007a). Its invertibility, coupled to its undecimated nature, the proportionality of the variance of the coefficients to the Fourier spectrum, and its robustness to when one begins to analyse the signal, make it a powerful analytical tool.

The WIAAFT algorithm has the following steps:

- (a) Take the MODWT using a wavelet with a high number of vanishing moments to deal with any potential non-stationarity in the series (Percival and Walden, 2000);
- (b) Apply the IAAFT algorithm to the wavelet detail coefficients at each scale j to generate a constrained realisation of the original detail coefficients, preserving the original values (hence, variance, hence, spectrum to a reasonable approximation) and their periodicity (hence, temporal structure);
- (c) Transpose this realisation so that the first detail coefficient in the transposed case is the last in the new variant;
- (d) Find the best match between the original detail coefficients at a particular scale and the two variants generated in (b) and (c) by circularly rotating until an error function (least-squares, penalised least-squares etc.) is minimised;
- (e) This will mean that the positions with high energies in the original data are mimicked in the surrogates, so we can invert the MODWT (using the original approximation coefficients) to yield a surrogate dataset;
- (f) Perform the rank-order matching described for the IAAFT method;
- (g) Use the output of (f) as initialisation of the standard IAAFT algorithm.

This method is illustrated in Keylock (2007b). Similar to the SIAAFT, the first stage assists the IAAFT algorithm in choosing an appropriate local minimum. Applying the IAAFT algorithm to the coefficients of an undecimated wavelet transform has the advantage over block resampling methods used in other wavelet-based methods (Breakspear et al., 2003; Angelini et al., 2005) that the structure of the

wavelet coefficients is preserved more precisely. In addition, the degrees of freedom in the IAAFT step (b) do not change with scale because the number of coefficients is constant. While wavelet versions of the IAAFT algorithm that fix in place the largest absolute wavelet coefficients may be useful for replicating the major singularities between data and surrogate (Keylock, 2007b), the algorithm used here was designed to be able to detect changes in the regularity of a signal. Because the local mean and variance for a non-intermittent signal are aligned between data and surrogate when using the WIAAFT algorithm, intermittency can be detected either through a direct comparison of local Hurst exponents (for slightly intermittent signals) or through a degradation in the preservation of the local mean and variance (for highly intermittent signals).

The overall degrees of freedom of the differing surrogate algorithms is an issue that is difficult to assess in practice or even define because the results are dependent on the nature of the data. For example, one can construct extreme cases of a single value or an infinitely long series of the same value where surrogates and data are always identical. In practice, the main constraint is the length of the series and in this study there are sufficient degrees of freedom available for WIAAFT surrogates relative to IAAFT surrogates, if circularly rotated versions of the latter are not considered to be independent (because random rotations could also be applied to the WIAAFT surrogates if required). To demonstrate this, 100 WIAAFT surrogates for a realisation of the data described below (with a value for the parameter a , as defined in Eq. 3, of 0.21) were generated and there was a median value of 5 exact matches between data and surrogate for a data series of 512 values (with a maximum of 10 exact matches). That is, $\sim 1\%$ to 2% of values in the data series were replicated exactly in a surrogate. One hundred IAAFT surrogates were also generated and then circularly rotated to find the maximum number of matches of individual values in the data and surrogate. The median over all surrogates was also 5 exact matches, with a maximum of 8.

3 The data and testing procedures

The properties of stationary geophysical signals can vary from signals with high degrees of persistence and approximately constant regularity (annual maxima river discharges have a Hurst exponent of ~ 0.7) to signals that vary in their persistence, implying intermittency (such as the change from background seismicity to an earthquake or no rainfall to a thunderstorm). Here we define a persistent signal as one with a Hurst exponent H between 0.5 and 1.0 (with $H=0.5$ the value for Brownian motion) and anti-persistence for a value of H between 0.0 and 0.5. In order to define intermittency, it is useful to move away from consideration of H , which is derived from a rescaled range analysis of a whole (fractal) signal and, consequently, takes a single value (see Abaimov

et al., 2007, for a recent discussion). Instead, we need a notion of a “local value for H ” that can vary over a signal. This measure of signal regularity is the Hölder/Lipshitz exponent α_p (Mallat, 1999; Kolwankar and Lévy-Véhel, 2002).

The definition of α_p at a particular time of interest t_0 for some function f measured over time t , follows from consideration of a polynomial G of degree less than s (where $s > 0$) and a constant C . A search is undertaken for the cases where the following inequality holds:

$$|f(t) - G(t - t_0)| \leq C |t - t_0|^s \quad (1)$$

We can define the set $C_G^s(t_0)$ as all the cases where Eq. 1 holds and our value for α_p at t_0 is given by the largest value for s in this set. More formally,

$$\alpha_p(t_0) = \sup \{s : f \in C_G^s(t_0)\} \quad (2)$$

Hence, s represents the smoothest fit to the data about this point, as the higher the value for s , the greater its differentiability. A signal where α_p is constant is equivalent to one characterised by a single value for H . Variability in α_p introduces intermittency. In order to study signals with a range of intermittency and, thus, varying persistence, our initial analysis is based on data signals X where the Hölder characteristics $\alpha_p(t)$ are prescribed according to:

$$\alpha_p(a, t) = 0.5 + a \sin(4\pi t) \quad (3)$$

where $a \in \{0.00, 0.01, 0.05, 0.09, 0.13, 0.17, \dots, 0.41\}$. This choice results in a signal that is at the boundary of persistence/anti-persistence because when intermittency is removed ($a=0$), we recover a Brownian motion. However, we recognise that the degree of persistence in geophysical data varies, with, for example, turbulence data anti-persistent, and river discharge series persistent. Hence, we also study:

$$\alpha_p(b_{33}, t) = 0.33 + b_{33} \sin(4\pi t) \quad (4)$$

and

$$\alpha_p(b_{66}, t) = 0.66 + b_{66} \sin(4\pi t) \quad (5)$$

where both b_{33} and b_{66} are given by $\{0.00, 0.01, 0.05, 0.09, 0.13, 0.17, \dots, 0.29\}$. We also examine two geophysical datasets described in more detail below. The algorithm used to derive $\alpha_p(t)$ is based on that of Wood and Chan (1994) and Chan and Wood (1997) as implemented in the *FracLab* toolbox (see Acknowledgements).

We adopted two general approaches to testing the algorithms. Initially, we generated 100 surrogates for two different realisations of the 12 signals given by Eq. (3). Although this testing procedure is akin to the way in which real data are tested (where one only has a single realisation of the phenomenon) the interpretation of the results was somewhat dependent upon the realisations (see below). Hence, we also used an approach where 100 realisations were generated for

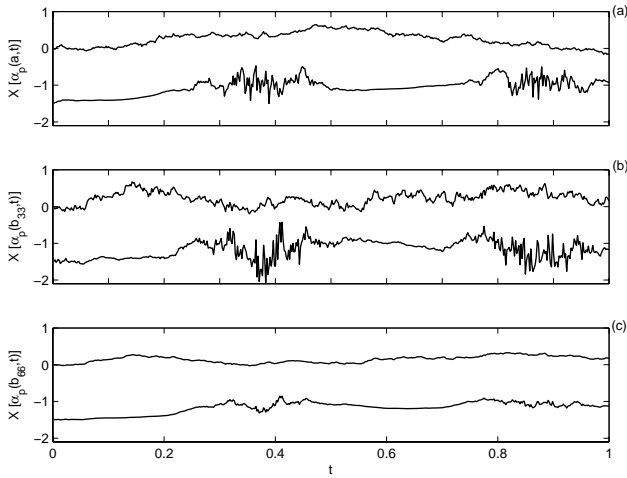


Fig. 1. Example signals with varying levels of persistence and intermittency. The signals in each plot are displaced vertically by -1.5 for clarity, with the lower signal the more intermittent. In 1a signals are generated using Eq. 3 with values for a of 0.00 (upper signal) and 0.41 (lower). In 1b signals are generated using Eq. 4 with values for b_{33} of 0.00 (upper signal) and 0.29 (lower). In 1c signals are generated using eq. 5 with values for b_{66} of 0.00 (upper signal) and 0.29 (lower).

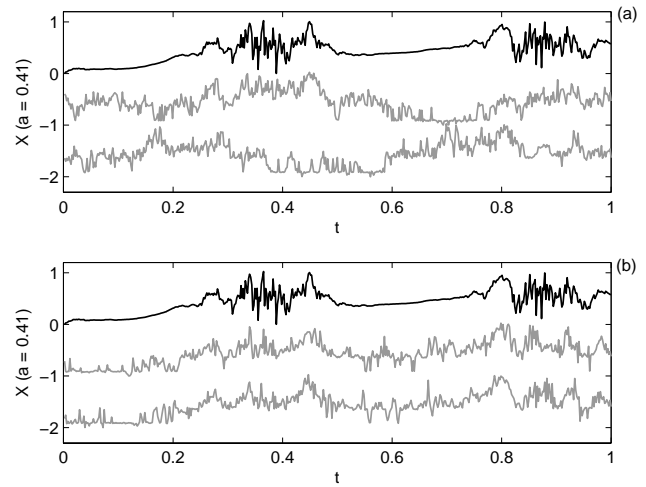


Fig. 3. Similar to Fig. 2 but for the bottom signal in Fig. 1a.

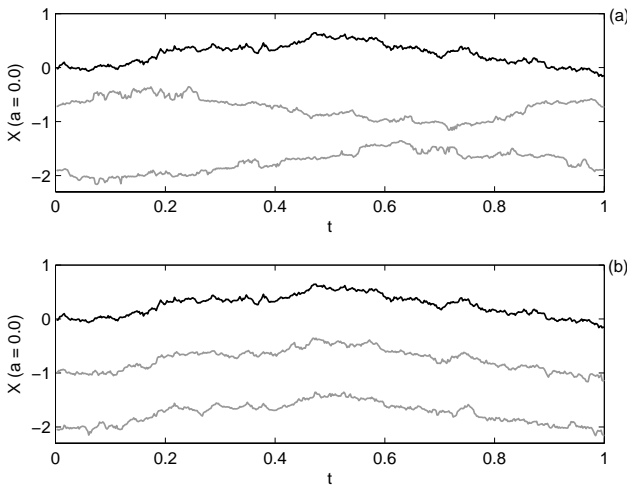


Fig. 2. Surrogates for the upper signal from Fig. 1a (black line in both panels). Figure 2a shows two IAAFT surrogates as grey lines, while 2b shows two WIAAFT surrogates as grey lines. Surrogates are displaced vertically for clarity.

each case, with one surrogate produced for each. This latter method was applied to the data generated using Eqs. (3), (4) and (5). Figure 1 shows realisations with values for a of 0.00 (1a, top) and 0.41 (1a, bottom), for b_{33} of 0.00 (1b, top) and 0.29 (1b, bottom) and for b_{66} of 0.00 (1c, top) and 0.29 (1c, bottom). A visual comparison of the difference between the IAAFT and WIAAFT algorithms for these signals is given in Figs. 2 and 3. While neither methods preserves the local

Hölder characteristics (by design), which makes representing the extremely intermittent signals in Fig. 3 problematic, the improved matching of the local mean and variance by the WIAAFT method is clear. However, there is a deterioration in the matching between Figs. 2 and 3, meaning that given a multifractal signal or one with sequential variation in the Hölder exponents, the WIAAFT method can be used to detect changes in α_p . This is because variation in the local mean or variance is similar between data and surrogates for the case of constant α_p , meaning that significant variation in α_p or the local mean/variance must be due to the changing roughness characteristics of the signal and not due to the lack of alignment of the two signals as could be the case for IAAFT methods (Keylock, 2006). This provides a suitable null hypothesis for testing for the presence of intermittency.

All three algorithms (the original IAAFT method, the SIAAFT and WIAAFT) preserve the actual values in the data and try to retain the Fourier amplitudes as precisely as possible. In order to test these methods, we use three related metrics. The autocovariance γ at lag τ for a mean-subtracted time series $X(t)$ is given by:

$$\gamma(\tau) = \overline{[X(t) - \overline{X(t)}][X(t + \hat{\tau}) - \overline{X(t)}]} \quad (6)$$

where an overbar indicates an average, τ ranges from 0 to $N-1$ and $\hat{\tau}$ is τ/N . In this study $N=512$ for all signals. The autocorrelation function $R(\tau)$ is then:

$$R(\tau) = \gamma(\tau)/\gamma(0) = \gamma(\tau)/\sigma^2 \quad (7)$$

which is the first measure used in this study. Using the Wiener-Khinchine theorem, the power spectral density function $f(\omega)$ is given by:

$$f(\omega) = \frac{1}{\pi} \sum_{\tau=-\infty}^{\infty} \gamma(\tau) e^{-i\omega\tau} \quad (8)$$

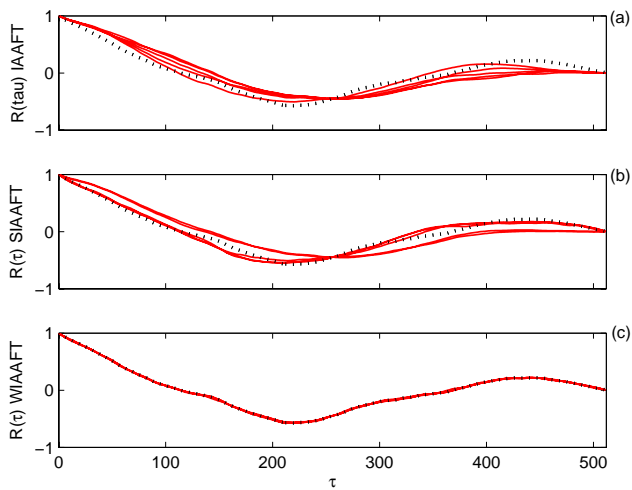


Fig. 4. A comparison of the autocorrelation functions $R(\tau)$ for the data in Fig. 2 (black dotted line) along with 5 surrogates (red lines). IAAFT surrogates are in (a), SIAAFT in (b) and WIAAFT in (c).

where ω is frequency over the range 0 to π . An estimate for $f(\omega)$ can be obtained directly from the data by taking the square of the Fourier transform to give the periodogram estimator $P(\omega)$:

$$P(\omega) = \frac{1}{N2\pi} \left| \sum_{\tau=1}^N X(\tau) e^{-i\omega\tau} \right|^2 \quad (9)$$

and this forms our second measure. Venema et al. (2006b) note that the periodogram is a noisy estimate and adopt a test statistic similar to the Kolmogorov-Smirnov test by expressing $P(\omega)$ in a cumulative form, normalised between 0 and 1 and then looking for the maximum distance, D , between these histograms for data and surrogates. We use this approach, but also study the standard deviation normalised Root-Mean-Squared Error (RMSE) between the original realisation, X , and a surrogate of that data, r , as an error measure, which we apply to both $R(\tau)$ and $P(\omega)$:

$$\text{RMSE} = \frac{1}{\sigma_X} \sqrt{\sum_{k=1}^N (M_k^X - M_k^r)^2} / hN \quad (10)$$

where M is the measure used [$R(\tau)$ or $P(\omega)$], $h=1$ for $R(\tau)$ and 0.5 for $P(\omega)$ as analysis of the latter is curtailed at the Nyquist frequency, and σ is the standard deviation.

In a well-known paper on surrogate data generation, Schreiber and Schmitz (2000) note that the most important parts of the autocorrelation function to be replicated are those at short lags. Furthermore, from a statistical perspective, as the lag increases, the significance of a particular value for R declines due to the reduced number of samples at this lag. Hence, an alternative means of implementing Eq. (10) is to curtail the testing at either the first zero-crossing or where

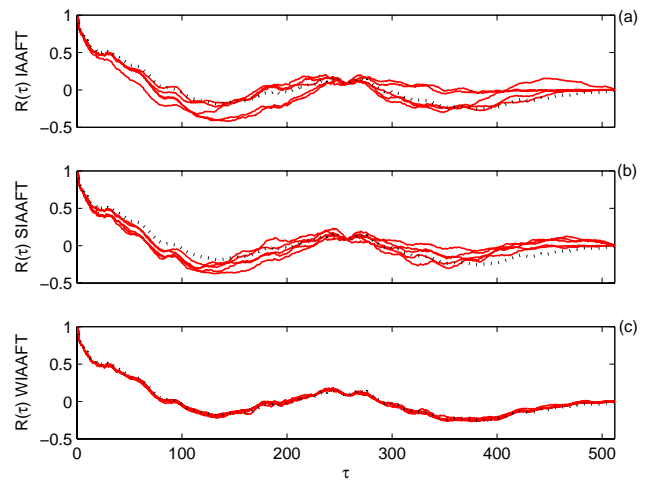


Fig. 5. Similar to Fig. 4 but for the data series used in Fig. 3.

a t -test based on the null hypothesis that R is not significantly to zero first changes from significant to insignificant. In this study, we test our geophysical data using the first zero-crossing of R approach, but because of the long-range correlations that exist by definition in fractional Brownian motions, zero is not necessarily crossed. Hence, for these data, the lag at which R is first insignificant at the 5% level is also used.

Venema et al. (2006a) suggest a variety of ways of implementing the SIAAFT. Their favoured approach is the partially stochastic variant, where the algorithm chooses at random one vector from five to update, where the vectors are given by: {1, 6, 11, ...}, {2, 7, 12, ...}, {3, 8, 13, ...}, {4, 9, 14, ...}, {5, 10, 15, ...}, and the values in each correspond to the positions of the rank-ordered data. Their convergence criterion for the first part of their algorithm is to accept a minimum if it is not superseded in 1000 additional iterations. We adopt these methods in this study and for each algorithm, the convergence criterion for the IAAFT stage of the algorithm (i.e. the second steps of the WIAAFT and SIAAFT methods and the whole of the IAAFT method) is identical. No additional adjustments were made such as the end-point matching that is often applied to improve convergence of the IAAFT algorithm. This is because this technique can be costly in terms of information loss for relatively short geophysical data series. To test the effect of this on our results, we generated 100 surrogates for sine waves sampled at 256 points with periods of 8π and 8.117π , the mean RMSE values for the R statistic were 0.071 (IAAFT) and 0.017 (WIAAFT) for the 8.1π data, and 0.033 (IAAFT), and 0.002 (WIAAFT) for the 8π data. Hence, end-matching improves convergence for the IAAFT algorithm, but not relative to the WIAAFT method, which supports the hypothesis that better methods than the IAAFT approach are needed for working with (pseudo)periodic data (Keylock, 2007b).

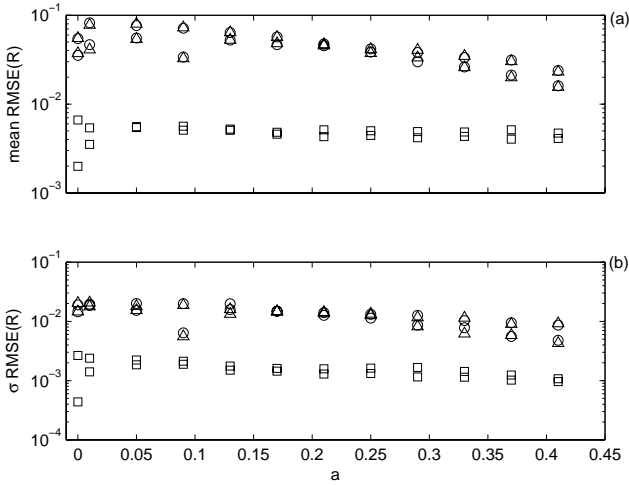


Fig. 6. The mean (a) and standard deviation (b) over 100 surrogates of the RMSE statistic for $R(\tau)$ over all lags obtained for two different signals with the same value for a . Circles are for IAAFT surrogates, triangles for SIAAFT and squares for WIAAFT surrogates.

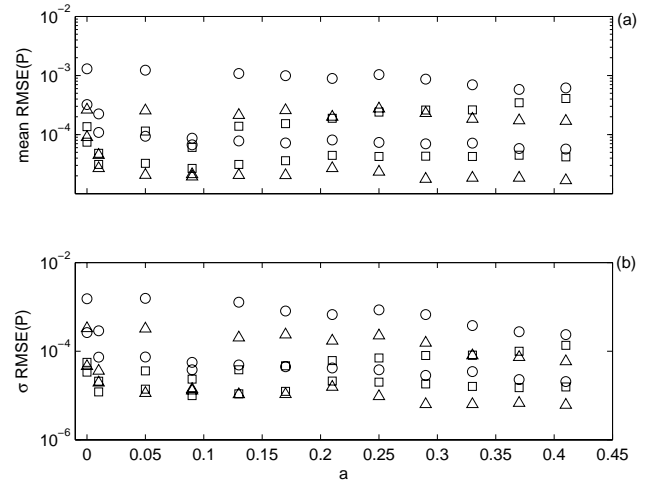


Fig. 8. The mean and standard deviation for the RMSE statistic applied to the $P(\omega)$ measure. The two signals for each value of a are those used in Fig. 6. Circles are for IAAFT surrogates, triangles for SIAAFT and squares for WIAAFT surrogates.

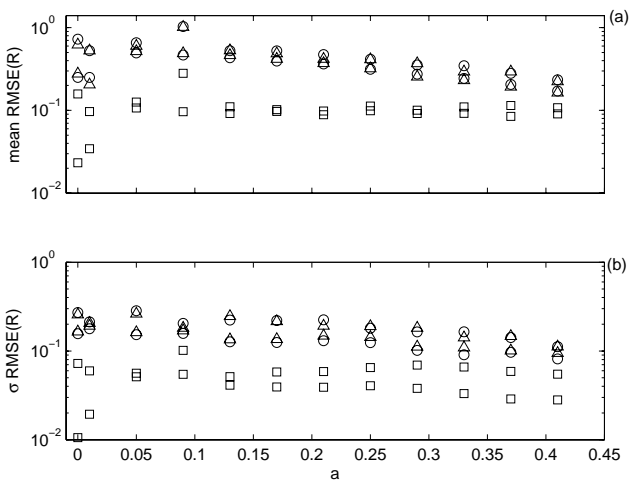


Fig. 7. The mean (a) and standard deviation (b) over 100 surrogates of the RMSE statistic for $R(\tau)$ curtailed at the lag where R was no longer significant at the 5% level. The two signals for each value of a are those used in Fig. 6. Circles are for IAAFT surrogates, triangles for SIAAFT and squares for WIAAFT surrogates.

4 Algorithm testing

Figures 4 and 5 compare values for $R(\tau)$ for five randomly selected surrogates of the data shown in Figs. 2 and 3. While the SIAAFT appears to give improved estimates over the original algorithm, in agreement with Venema et al. (2006a), the WIAAFT algorithm gives a better fit. This is assessed using the mean and standard deviation of the RMSE over 100 surrogates for 2 different realisations of the process at a chosen value for a in Fig. 6. The advantage of the WIAAFT

method is clear, especially for signals with low intermittency. However, even at higher intermittencies, the WIAAFT method is still four times more accurate in the average for $a=0.41$. Results when the RMSE statistic is only calculated over the statistically significant part of the autocorrelation function are shown in Fig. 7. While the differences between the algorithms is less marked, the WIAAFT algorithm still outperforms the other algorithms, and again, the contrast is greater for signals with lower intermittencies.

Owing to the noisy nature of the periodogram estimator, the RMSE statistic for $P(\omega)$ is not a particularly robust measure as noted by Venema et al. (2006b). Based on this statistic, there is less to choose between the SIAAFT and WIAAFT methods (Fig. 8), with the former performing better at high intermittency. However, both methods outperform the IAAFT technique across the range of values for a . Note that for $a > 0.1$, the results are dataset dependent with the statistics for the IAAFT algorithm for one dataset better than those for the SIAAFT and WIAAFT for a different dataset with the same a . Using the D statistic, the SIAAFT and WIAAFT methods again outperform the IAAFT (Fig. 9), which is consistent with the analysis of Venema et al. (2006b). For low intermittencies, the mean difference between the IAAFT and WIAAFT methods can be in excess of an order of magnitude. However, for $a > 0.25$, the SIAAFT appears to converge more successfully.

The alternative approach to testing was to generate 100 realisations of the process and one surrogate for each case and examine the distribution of the errors over all 100 realisations. Boxplots of the results for the RMSE of the R , P statistics and the D statistic are shown in Figs. 10 to 12. The central line of the boxplot indicates the median with upper and lower quartiles giving the upper and lower edges of

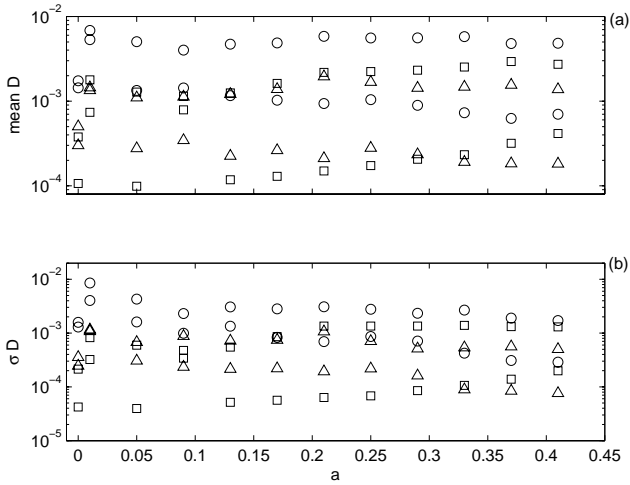


Fig. 9. The mean (a) and standard deviation (b) over 100 surrogates of the D error statistic for $P(\omega)$. The two signals for each value of a are those used in Fig. 6. Circles are for IAAFT surrogates, triangles for SIAAFT and squares for WIAAFT surrogates.

the box. The whiskers extend up to 1.5 times the interquartile range from the top and bottom of the boxes. Note the difference in the y-axis scaling in Fig. 10c and in Figs. 11a and 12a. These results are consistent with the earlier analysis, with WIAAFT surrogates out-performing the other techniques by an order of magnitude for R , and WIAAFT and SIAAFT surrogates out-performing the original Schreiber and Schmitz IAAFT algorithm by a factor of 4 for P and D , with SIAAFT surrogates offering a greater advantage relative to WIAAFT surrogates for P and D statistics as a increases. Figures 13 and 14 show the results for the R statistic for the anti-persistent and persistent data series given by Eqs. 4 and 5. Again, the WIAAFT algorithm outperforms the others over all values for a , while a comparison between Figs. 10, 13 and 14 shows that the algorithms perform less effectively on smoother data (higher mean value for α_p).

Our final two cases are geophysical datasets with contrasting characteristics. The first case is 256 years of annual sunspot data from 1700 to 1955 collected by the Solar Influences Data analysis Centre (SIDC, 2007) illustrated in Fig. 15a. The second set of data are 2^{14} daily river discharge measurements from 29 March, 1904 to 17 February 1959 on the San Pedro river at Charleston, Arizona (USGS site number 09471000) downloaded from the USGS National Water Information System at <http://nwis.waterdata.usgs.gov/nwis/> and shown in Fig. 15b. Thirty nine surrogates were generated for each case, which is the minimum number for hypothesis testing at the 5% level using a two-tailed test. The nature of these signals again varies with the sunspot data showing a relatively smooth, periodic behaviour and the discharge data showing a great deal of variability on top of the annual periodicity. In agreement with earlier results, Fig. 16 shows that all algorithms perform better on the rougher discharge data.

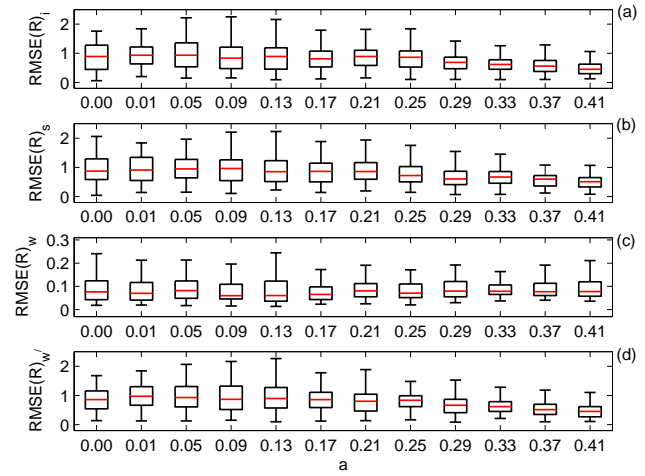


Fig. 10. Results for the R statistic based on generating one surrogate for 100 different realisations of the process given by Eq. 3. The subscripts i , s , and w in (a), (b), and (c), respectively, refer to the IAAFT, SIAAFT and WIAAFT methods. The subscript w' in (d) refers to the WIAAFT algorithm without the matching step that is discussed at the end of the paper. The central line of the boxplot indicates the median with upper and lower quartiles giving the upper and lower edges of the box. The whiskers extend up to 1.5 times the interquartile range from the top and bottom of the boxes. Points outside this range are not shown.

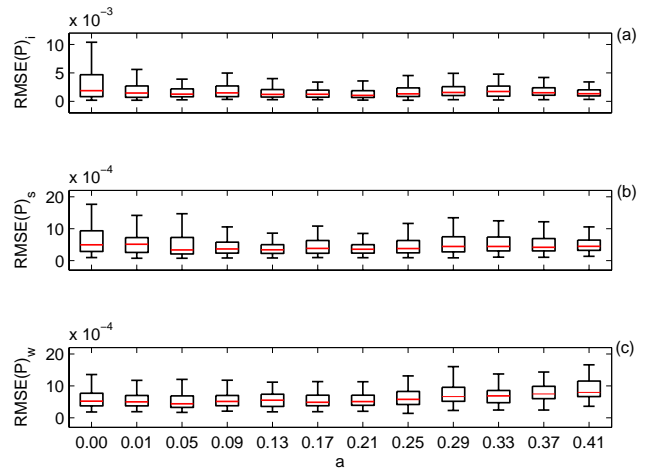


Fig. 11. Results for the P statistic based on generating one surrogate for 100 different realisations of the process given by Eq. 3. The definition for the boxplots is given in Fig. 10.

However, the difference in performance between the algorithms is also clearest in this case, with the WIAAFT method outperforming the others by more than a factor of 5 on average (medians of 6.9×10^{-4} , 5.7×10^{-4} , and 1.0×10^{-4} for the RMSE of the R statistic for IAAFT, SIAAFT and WIAAFT, respectively).

As a further check on our results, we calculated the RMSE for R over the number of lags that equate to the first zero

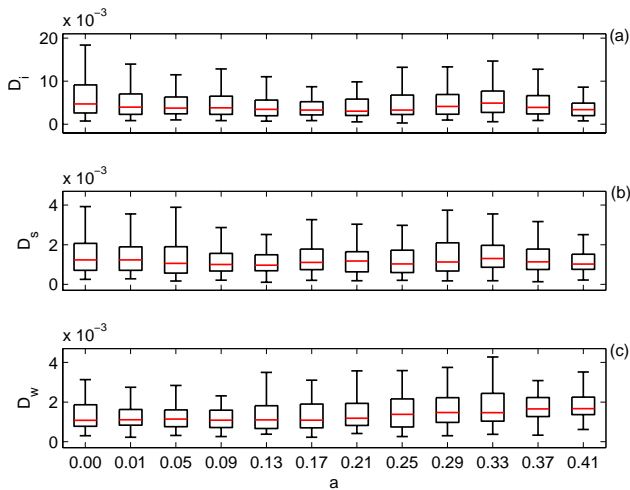


Fig. 12. Results for the D statistic based on generating one surrogate for 100 different realisations of the process given by Eq. 3. The definition for the boxplots is given in Fig. 10.

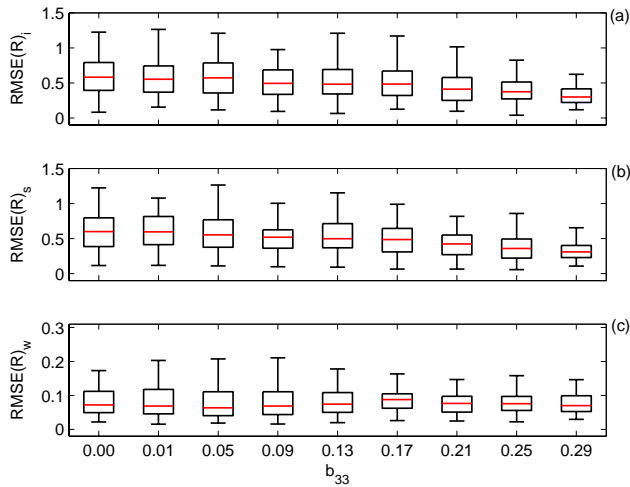


Fig. 13. Results for the R statistic based on generating one surrogate for 100 different realisations of the process given by Eq. 4 (the anti-persistent process). The subscripts i , s , and w in (a), (b), and (c), respectively, refer to the IAAFT, SIAAFT and WIAAFT methods. The definition for the boxplots is given in Fig. 10.

crossing of R (4 years for the sunspot data and 50 days for the discharge data). As shown in Fig. 17, this greatly improves the performance of the SIAAFT algorithm for the discharge data, although the WIAAFT still has the lowest average error for both sunspot and discharge datasets and also gives a low error more consistently.

5 Discussion of the testing and conclusion

A surrogate generating algorithm that was proposed by Keylock (2006) has been tested against the well known IAAFT

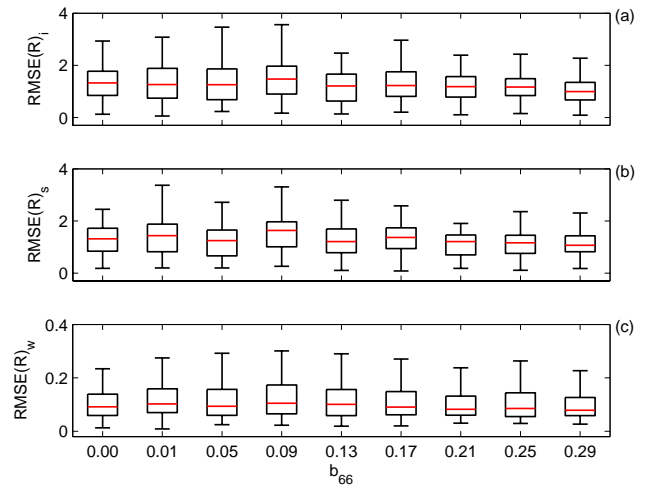


Fig. 14. Results for the R statistic based on generating one surrogate for 100 different realisations of the process given by eq. 5 (the persistent process). The subscripts i , s , and w in (a), (b), and (c), respectively, refer to the IAAFT, SIAAFT and WIAAFT methods. The definition for the boxplots is given in Fig. 10.

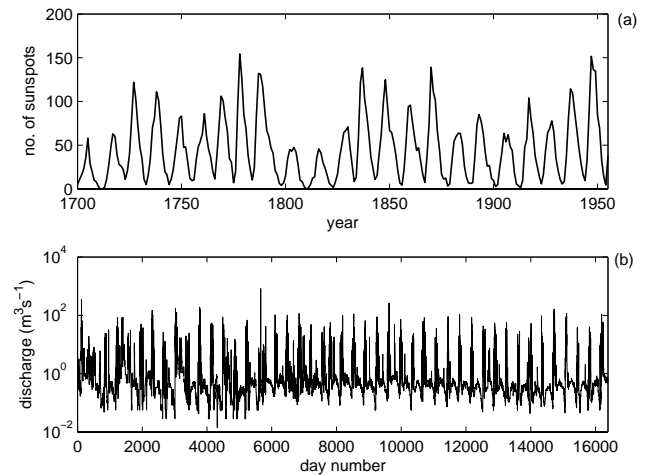


Fig. 15. The annual sunspot number (a) and daily river discharge data (b) examined in this study.

method and a stochastic variant of that method recently proposed in this journal by Venema et al. (2006a). Testing has been undertaken systematically using artificial data of varying persistence and intermittency, as well as with two geophysical datasets. Although the WIAAFT method was not developed with improved convergence of the IAAFT method in mind (in contrast to SIAAFT), it yields a better replication of the autocorrelation function and is at least as effective as the SIAAFT method using Fourier-based error criteria, especially for low intermittencies. The reason for this improved performance for the autocorrelation is that in the first stage of the WIAAFT algorithm, the approximate temporal behaviour of the wavelet coefficients is mimicked, while

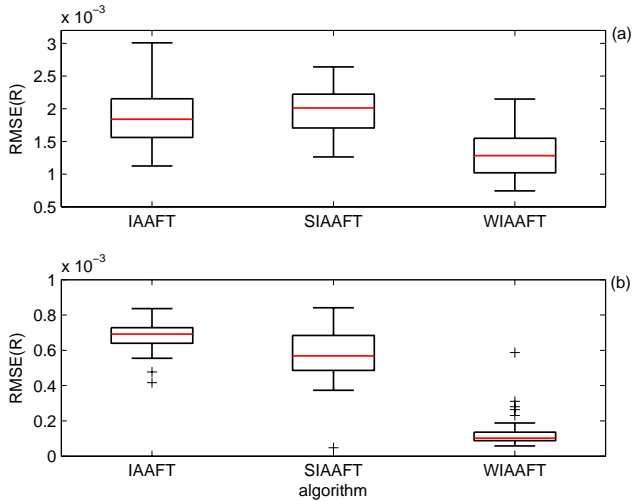


Fig. 16. Boxplots of results for tests using the RMSE for the R statistic over all lags for the sunspot (a) and discharge (b) data shown in Fig. 14 for the three algorithms. The definition for the boxplots is given in Fig. 10 but this figure also shows outliers that exceed the length of the whiskers as crosses.

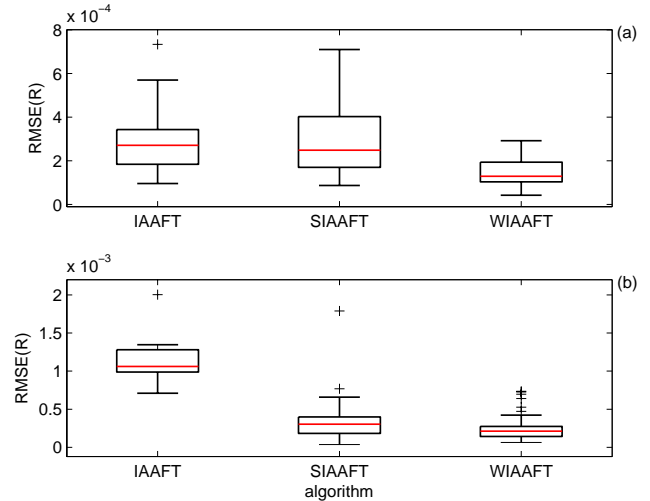


Fig. 17. Boxplots of results for tests using the RMSE for the R statistic defined until the first zero crossing for the sunspot (a) and discharge (b) data shown in Fig. 14 for all three algorithms. The definition for the boxplots is given in Fig. 10 but this figure also shows outliers that exceed the length of the whiskers as crosses.

the other algorithms have no similar constraint. To demonstrate unambiguously that the matching step explains the improved performance, Fig. 10d shows results for the R measure but for WIAAFT surrogates produced with no matching step (i.e. the wavelet coefficients are randomised using the IAAFT algorithm but are not rotated to optimally match the original coefficients). The results are virtually identical to those for the IAAFT and SIAAFT in Fig. 10a and b. Hence, the capability to introduce temporal alignment when using a time-frequency transform such as wavelets provides a means of improving algorithm convergence.

Greater insight into how the WIAAFT algorithm improves performance can be seen in Fig. 18 where the correlation between a realisation of a process (given by Eq. 3) and a surrogate $R(0)$ is plotted against the RMSE for the complete autocorrelation function (shown in Fig. 10). Results are shown for $a=0$ (black) and $a=0.41$ (red). Because neither of these algorithms is designed to preserve intermittency, the (unsigned) correlations are higher for $a=0$. Furthermore, the alignment steps of the WIAAFT algorithm ensure that $R(0)_w$ is generally of a high positive value. For IAAFT surrogates, there is a reduction in RMSE (R) as the correlation tends to ± 1 indicating that those surrogates that (by chance) are highly correlated to the original data are likely to yield a lower error. WIAAFT surrogates are highly correlated by design and there is a clear trend in the RMSE (R) relation with $R(0)_w$ when $a=0$. However, when $a=0.41$ not only does $R(0)_w$ decrease, but the trend is destroyed. Thus, the consequences of failing to preserve intermittency over-ride the RMSE(R) – $R(0)_w$ correlation and the results reflect the accuracy in representing the underlying process (Eq. 3).

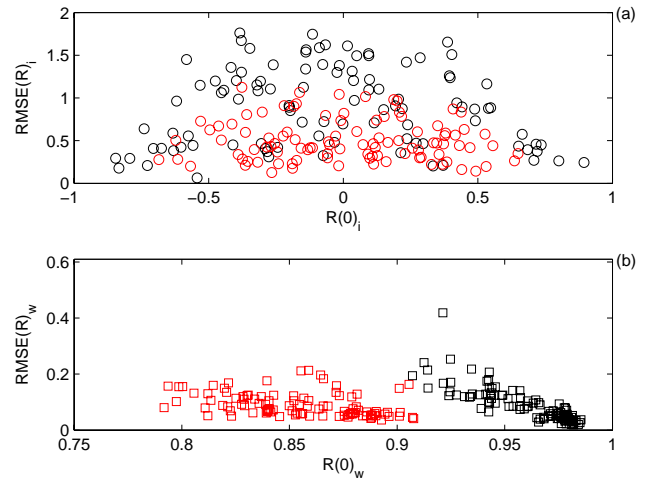


Fig. 18. The relation between RMSE (R) and the linear correlation $R(0)$ between a realisation and its surrogate for the 100 realisations generated to produce Fig. 10. Results are shown for IAAFT surrogates (i subscript and circles) and WIAAFT surrogates (w subscript and squares). Values in black are for $a=0$ (no intermittency) while values in red are for $a=0.41$ (high intermittency).

The main use for the WIAAFT algorithm in the context of this paper is for data simulation, similar to that of Angelini et al. (2005) or Venema et al. (2006b, 2006c). In the context of hypothesis testing for nonlinearity, the IAAFT and SIAAFT algorithms are designed to produce constrained, linear realisations of a process that can be contrasted with the actual data on some measure (and hence, test for nonlinearity). The WIAAFT algorithm restricts the possible class of realisations

to those that retain some aspect of the local mean and variance of a non-intermittent variant of the original data, which means that it cannot be used for the same type of testing. However, this algorithm can be used for testing for a change in signal intermittency, which is a specific type of nonlinearity. New algorithms are required to undertake simulation that preserves the underlying intermittency.

Acknowledgements. Data generation in this paper made use of the algorithms available in the FracLab toolbox available from <http://www.irccyn.ec-nantes.fr/hebergement/FracLab/>. We thank Ronald Van der Linden and the rest of the Solar Influences Data analysis Centre (SIDC), Belgium, for access to the sunspot data. The wavelet-based algorithms discussed in this paper may be downloaded from <http://www.geog.leeds.ac.uk/people/c.keylock/>.

Edited by: B. D. Malamud

Reviewed by: three anonymous referees

References

- Abaimov, S. G., Turcotte, D. L., Shcherbakov, R., and Rundle, J. B.: Recurrence and interoccurrence behavior of self-organized complex phenomena, *Nonlinear Proc. Geoph.*, 14, 455–464, 2007.
- Angelini, C., Cava, D., Katul, G., and Vidakovic, B.: Resampling hierarchical processes in the wavelet domain: A case study using atmospheric turbulence, *Physica D*, 207, 27–40, 2005.
- Ashkenazy, Y. and Tziperman, E.: Are the 41 kyr glacial oscillations a linear response to Milankovitch forcing? *Quaternary Sci. Rev.*, 23, 1879–1890, 2004.
- Basu, S., Foufoula-Georgiou, E., Lasheremes, B., and Arnéodo, A.: Estimating intermittency exponent in neutrally stratified atmospheric surface layer flows: A robust framework based on magnitude cumulant analysis and surrogate analyses, *Phys. Fluids*, 19, 115102, 2007.
- Breakspear, M., Brammer, M., and Robinson, P. A.: Construction of multivariate surrogate sets from nonlinear data using the wavelet transform, *Physica D*, 182, 1–22, 2003.
- Chan, G., and Wood, A. T. A.: AS 312 – An algorithm for simulating stationary Gaussian random fields, *Appl. Stat.-J. Roy. St. C*, 46, 171–181, 1997.
- Keylock, C. J.: Constrained surrogate time series with preservation of the mean and variance structure, *Phys. Rev. E*, 73, 036707, 2006.
- Keylock, C. J.: The visualisation of turbulence data using a wavelet-based method. *Earth Surf. Proc. Land.*, 32, 637–647, 2007a.
- Keylock, C. J.: A wavelet-based method for surrogate data generation, *Physica D*, 225, 219–228, 2007b.
- Kolwankar, K. M. and Lévy Véhel, J.: A time domain characterization of the fine local regularity of functions, *J. Fourier Anal. Appl.*, 8, 319–334, 2002.
- Kosugi, K., Sato, T., and Sato, A.: Dependence of drifting snow saltation lengths on snow surface hardness, *Cold. Reg. Sci. Tech.*, 39, 133–139, 2004.
- Mallat, S.: *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, 1999.
- Muzy, J. F., Bacry, E., and Arneodo, A.: Wavelets and multifractal formalism for singular signals – application to turbulence data, *Phys. Rev. Lett.*, 67, 3515–3518, 1991.
- Nemoto, M. and Nishimura, K.: Direct measurement of shear stress during snow saltation. *Boundary-Layer Meteorol.* 100, 149–170, 2001.
- Nicolleau, F., Vassilicos, J.C.: Wavelets for the study of intermittency and its topology, *Phil. T. R. Soc. A*, 357, 2439–2457, 1999.
- Percival, D. B. and Walden, A. T.: *Wavelet Methods for Times Series Analysis*. Cambridge University Press, Cambridge, 2000.
- Poggi, D., Porporato, A., Ridolfi, L., Albertson, J. D., and Katul, G. G.: Interaction between large and small scales in the canopy sublayer, *Geophys. Res. Lett.*, 31, L05102, doi:10.1029/2003GL018611, 2004.
- Schreiber, T.: Constrained randomization of time series data, *Phys. Rev. Lett.*, 80, 2105–2108, 1998.
- Schreiber, T., and Schmitz, A.: Improved surrogate data for nonlinearity tests, *Phys. Rev. Lett.*, 77, 635–638, 1996.
- Schreiber, T., and Schmitz, A.: Surrogate time series, *Physica D*, 142, 346–382, 2000.
- SIDC-team.: World Data Center for the Sunspot Index, Royal Observatory of Belgium, Monthly Report on the International Sunspot Number, online catalogue of the sunspot index: <http://www.sidc.be/sunspot-data/>, 2007.
- Simoncelli, E. P., Freeman, W. T., Adelson, E. H., and Heeger, D. J.: Shiftable multiscale transforms, *IEEE Trans. Inf. Theory*, 38, 587–607, 1992.
- Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Farmer, J. D.: Testing for nonlinearity in time-series – the method of surrogate data, *Physica D*, 58, 77–94, 1992.
- Theiler, J. and Prichard, D.: Constrained-realization Monte-Carlo method for hypothesis testing, *Physica D*, 94, 221–235, 1996.
- Venema, V., Ament, F., and Simmer, C.: A Stochastic iterative amplitude adjusted Fourier transform algorithm with improved accuracy, *Nonlinear Proc. Geoph.*, 13, 321–328, 2006a.
- Venema, V., Bachner, S., Rust, H. W., and Simmer, C.: Statistical characteristics of surrogate data based on geophysical measurements, *Nonlinear Proc. Geoph.*, 13, 449–466, 2006b.
- Venema, V., Meyer, S., Garcia, S. G., Kniffka, A., Simmer, C., Crewell, S., Lohnert, U., Trautmann, T., and Macke, A.: Surrogate cloud fields generated with the iterative amplitude adapted Fourier transform algorithm, *Tellus A*, 58, 104–120, 2006c.
- Whitcher, B., Guttorp, P., Percival, D.B.: Wavelet analysis of covariance with application to atmospheric time series, *J. Geophys. Res.*, 105, 14941–14962, 2000.
- Wood, A. T. A. and Chan, G.: Simulation of stationary Gaussian processes in $[0, 1]^d$, *J. Comput. Graph. Statist.*, 3, 409–432, 1994.