

Cross-Wavelet Bias Corrected by Normalizing Scales

DORIS VELEDA

Centro de Estudos e Ensaios em Risco e Modelagem Ambiental, CEERMA–UFPE, and Laboratório de Oceanografia Física Estuarina e Costeira, Departamento de Oceanografia da Universidade Federal de Pernambuco LOFEC–DOCEAN–UFPE, Recife, Pernambuco, Brazil

RAUL MONTAGNE

Centro de Estudos e Ensaios em Risco e Modelagem Ambiental, CEERMA–UFPE, and Departamento de Física, Universidade Federal Rural de Pernambuco, Recife, Pernambuco, Brazil

MOACYR ARAUJO

Centro de Estudos e Ensaios em Risco e Modelagem Ambiental, CEERMA–UFPE, and Laboratório de Oceanografia Física Estuarina e Costeira, Departamento de Oceanografia da Universidade Federal de Pernambuco LOFEC–DOCEAN–UFPE, Recife, Pernambuco, Brazil

(Manuscript received 10 August 2011, in final form 13 February 2012)

ABSTRACT

The cross-wavelet transform (XWT) is a powerful tool for testing the proposed connections between two time series. Because of XWT's skeletal structure, which is based on the wavelet transform, it is suitable for the analysis of nonstationary periodic signals. Recent work has shown that the power spectrum based on the wavelet transform can produce a deviation, which can be corrected by choosing a proper rectification scale. In this study, it is shown that the standard application of the XWT can also lead to a biased result. A corrected version of the standard XWT was constructed using the scale of each series as normalizing factors. This correction was first tested with an artificial example involving two series built from combinations of two harmonic series with different amplitudes and frequencies. The standard XWT applied to this example produces a biased result, whereas the correct result is obtained with the use of the proposed normalization. This analysis was then applied to a real geophysical situation with important implications to climate modulation on the northwestern Brazilian coast. The linkage between the relative humidity and the shortwave radiation measurements, obtained from the 8°S, 30°W Autonomous Temperature Line Acquisition System (ATLAS) buoy of the Southwestern Extension of the Prediction and Research Moored Array in the Tropical Atlantic (PIRATA-SWE), was explored. The analysis revealed the importance of including the correction in order to not overlook any possible connections. The requirements of incorporating this correction in the XWT calculations are emphasized.

1. Introduction

Wavelet analysis is a widely used tool. In contrast to Fourier analysis, wavelet analysis analyzes signals with changing spectra. Thus, the estimation of the spectral characteristics of a time series is a function of time. The evolution of the periodic components over time can reveal a link between two signals, allowing the quantification of a nonstationary association for a specific time range. For

example, for time series composed of sine waves with the same amplitude and different frequencies, the conventionally adopted wavelet method does not produce a spectrum with identical peaks, which is in contrast to a Fourier analysis. The wavelet power spectrum, defined as the transform coefficient squared (to within a constant factor), is equivalent to the integration of energy over the influence period that the series spans (Torrence and Compo 1998, hereafter TC98; Kumar and Foufoula-Georgiou 1997). Thus, a physically consistent definition of energy for the wavelet power spectrum should be the squared transform coefficient divided by the scale with which it is associated (Farge 1992; Liu et al. 2007, hereafter L07;

Corresponding author address: Doris Veleda, LOFEC/DOCEAN/UFPE, Av. Arquitetura s/n, 50740-550, Recife, PE, Brazil.
E-mail: doris.veleda@ufpe.br

Zunino et al. 2007). Such a rectified wavelet power spectrum results in a substantial improvement in the spectral estimate, allowing for a comparison of spectral peaks across different scales.

While the analysis of a time series is very important, it is often desirable to examine two time series together that may be expected to be linked in some way. Specifically, it may be of interest whether nonstationary frequencies with a large common power have a consistent phase relationship, which also indicates causality between the time series. Cross-wavelet analysis can reveal transient phenomena synchrony and long-term relationships. Although the cross-wavelet transform (XWT) is the recommended tool for this purpose, this approach has the same inadequate outcome as the standard power spectrum; that is, it can screen out the peaks of resonant frequencies. As in the case of the wavelet power spectrum, the solution is to apply the appropriate normalizing factor.

TC98 introduced in their practical guide to wavelet analysis with the software that is currently widely adopted by the mainstream atmospheric–ocean science community. The power spectra calculated with this software induce a bias, already shown in L07, and a similar problem emerges in the XWT, as we show in this work.

This paper addresses the bias problem in the estimate of the cross wavelet for time series datasets. We built, based on the rectification presented by L07, the required correction for this bias problem of the XWT. In section 2, we present the basic concepts of wavelet theory and the correction on the wavelet power spectrum introduced by L07. Additionally in that section, we show the presence of the same problem in the XWT and how to correct it. The proposed approach for correction is validated in section 3 with an artificial example of two time series constructed as the combination of harmonic functions of different frequencies. The two sinusoidal time series were constructed so that they are resonant in just one frequency. The standard XWT instead reveals two resonant frequencies. However, when we apply the correction introduced in section 3, the XWT correctly shows only one resonance. In section 4, we analyze two weakly coupled geophysical signals: the relative humidity and the shortwave radiation, collected from the Autonomous Temperature Line Acquisition System (ATLAS) buoy moored at 8°S, 30°W, as part of the Southwestern Extension of the Prediction and Research Moored Array in the Tropical Atlantic (PIRATA-SWE; Servain et al. 1998; Bourlés et al. 2008). These are key-state variables for the prediction of easterly atmospheric wave propagation and extreme rainfall events at the eastern boundary of Northeast Brazil (Nobre et al. 2004; Rao et al. 1993). The main conclusions of the paper are presented in section 5.

2. Biased cross-wavelet analysis and correction

The problem of data analysis is of central importance for numerous studies in the atmospheric and oceanic sciences, and in physics, biology, and economics, among others. Wavelet analysis has become a useful approach for time series analysis. Wavelet analysis provides a suitable tool for nonstationary series once intermittent periodicities can be detected with the wavelet analysis resulting from its time–frequency localization property. A complete description of the geophysical applications of wavelet analysis can be found in Kumar and Foufoula-Georgiou (1997); a theoretical treatment of wavelet analysis is given in Daubechies (1992, 1999), while a broad perspective on the principles and applications of transient signal processing with wavelet is presented in Mallat (1998).

Any time series can be decomposed using the wavelet family functions $\psi_n^j = 2^{-j/2}\psi(2^j t - n)$, with j, n integers, translations, and scaling indexes, respectively, where ψ is a function called the mother wavelet. If this family forms an orthonormal basis for the Hilbert space $L^2(\mathbb{R})$, then the space is consistent with finite-energy signals. The concept of the energy and power spectrum is linked with the usual notions derived from the classical Fourier theory.

The continuous wavelet transform of a function $f(t)$, $f(u, s)$, is defined as the inner product of the family of wavelets $\psi_{u,s}(t)$ with $f(t)$. This is given by

$$f(u, s) = \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{+\infty} f(t) \psi_{u,s}^* dt, \quad (1)$$

where ψ^* denotes the conjugate of $\psi_{u,s}$. To implement the wavelet transform on sampled signals, we need to discretize the parameters u and s . The most common choice is $s = 2^j$ and $u = n2^j$, with j, n integers. For a discrete data array $f[t]$, the discrete wavelet transform f_n^j is performed by

$$f_n^j = \int_{-\infty}^{+\infty} f[t] \frac{1}{\sqrt{2^j}} \psi\left(\frac{t - n2^j}{2^j}\right) dt = \int_{-\infty}^{+\infty} f[t] \psi_n^j dt, \quad (2)$$

where $n = 1, 2, \dots$ is the index of the data array. Thus, the function $f(t)$ can be approximated by

$$f(t) = \sum_j \sum_n f_n^j \psi_n^j(t). \quad (3)$$

The discrete wavelet transform measures the contribution to $f[t]$ of scale 2^j at location $n2^j$. These conditions are fairly broad and cover a wide range of situations.

The wavelet coefficients from that basis are given by $f_n^j = \langle f, \psi_n^j \rangle$ and, as in Fourier analysis, the corresponding associated energy is given by the squares $|f_n^j|^2$, up to some constant factor. In a recent article, L07 have shown

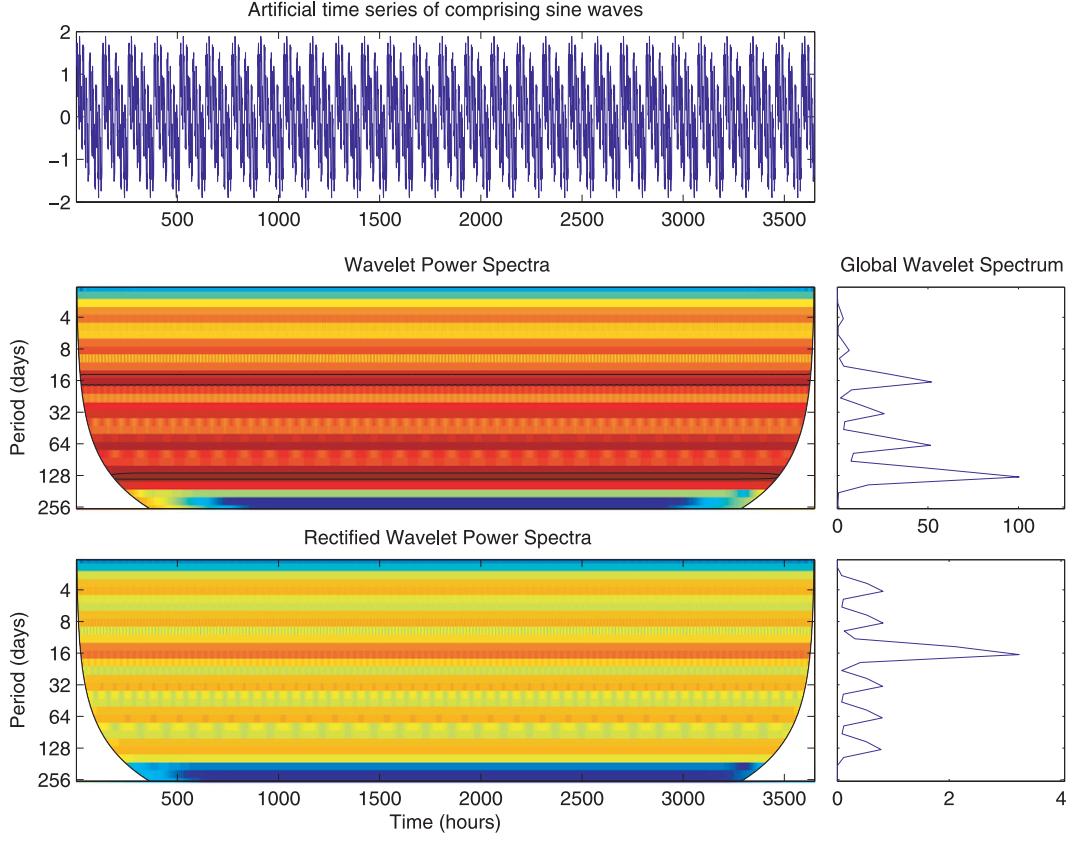


FIG. 1. (top) The artificial time series comprising sine waves of eight different periods (1, 2, 4, 8, 16, 32, 64, and 128 days). Also shown are the wavelet power spectra in logarithm (base 2) (left middle) without rectifying and (left bottom) rectified, and (right bottom) the global wavelet power spectra of the artificial time series. High (red) and low (blue) wavelet power spectrum values are indicated. The COI, where zero padding has reduced the variance, is shown (white shaded regions).

that the wavelet spectrum is biased, accentuating the lower frequencies, if the scale factor is not taken into account. A physically consistent definition of energy at location n (corresponding to the time $t = 2^j n$) and a scale level j (corresponding to scale 2^j) in terms of the wavelet transform coefficient is

$$E_n^j \equiv \frac{1}{\Delta} |f_n^j|^2 = 2^j \times |f_n^j|^2 = |2^{j/2} f_n|^2. \quad (4)$$

L07 showed that if this factor is not considered, then the wavelet spectral analysis produces a “biased” power spectra, where the high-frequency peaks tend to be underestimated. Actually, software that is widely adopted by the mainstream atmosphere–ocean science community, based on TC98’s practical guide to wavelet analysis, has this analogous problem when the power spectra are calculated.

The problem with the power spectra reported by L07 can also be found when addressing the cross-wavelet spectrum. Given two time series $f(t)$ and $g(t)$, with (discrete)

wavelet transforms coefficients from f_n^j to $f(t)$ and from g_n^k to $g(t)$, the cross wavelet at location n is conventionally (TC98; Grinsted et al. 2004) defined as

$$W_n^{j,k} \equiv (f_n^j) \times (g_n^k)^*, \quad (5)$$

where the asterisk (*) denotes complex conjugation. The cross-wavelet spectrum is intended to identify a peak for the common frequency (period). Instead, when one of the two signals exhibits an important contribution in a high frequency, the conventional cross wavelet will indicate that there was a resonance in this high frequency, incorrectly indicating that the two signals had important contributions at that same frequency. The problem has its origin in the same scale factor issue; however, because two time series are involved, the solution is somewhat different. It was shown in Eq. (4) that the total power spectrum coefficient should be corrected by the scale factor 2^j . The cross wavelet involves two different signals, so the scale factor may differ for each time series (the energy distribution may

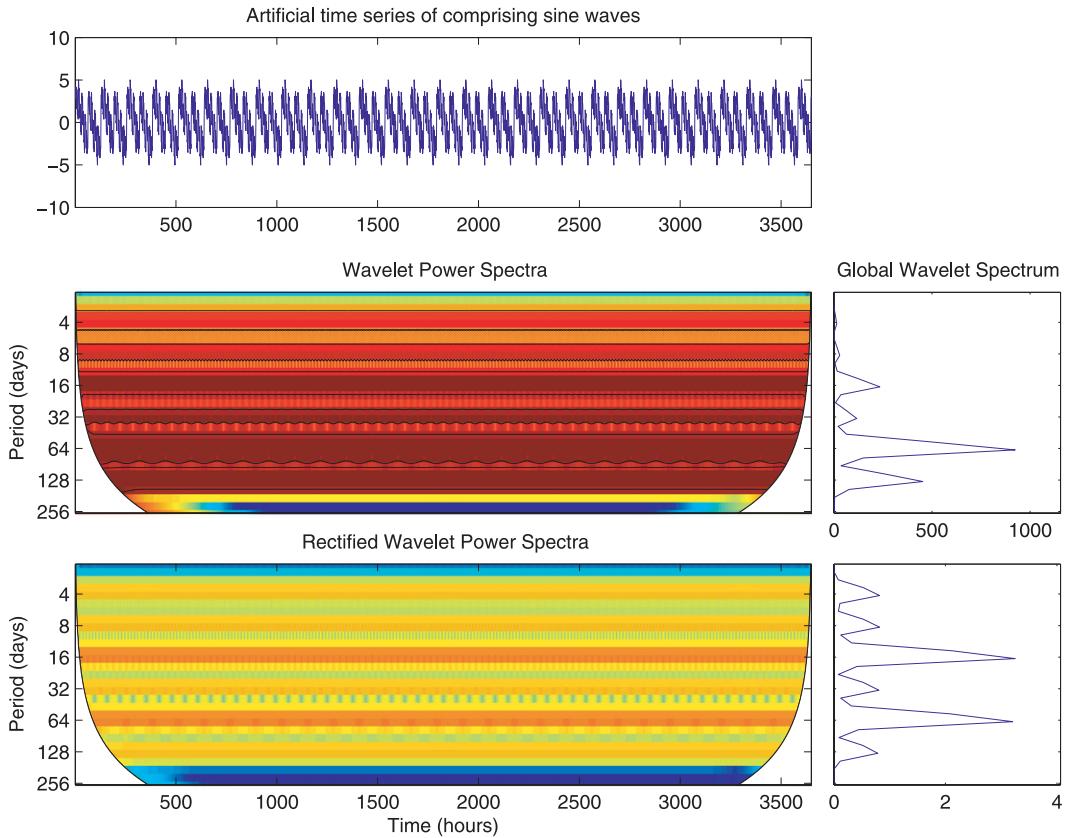


FIG. 2. As in Fig. 1, but for the other artificial time series comprising sine waves of eight different periods.

be different along the scales for the two signals). Thus, the square root of the scale factor correction has to be applied to each wavelet transform, namely, $2^{j/2}$ and $2^{k/2}$. The cross wavelet is then defined as the product of the two wavelet transforms with the corresponding scale factor,

$$W_n^{j,k} \equiv (2^{j/2}f_n^j) \times (2^{k/2}g_n^k)^*. \quad (6)$$

In this section we have introduced the necessary correction for the cross wavelet to avoid erroneous overestimations of resonant frequencies. In the next section we will show how this correction works in an artificial controlled example, and in section 4 we analyze a geophysical situation.

3. Cross-wavelet correction in an artificial example

To show the error produced in the calculation of the cross wavelet of two series when the wavelet is not normalized, we start with a simple and controlled example: two time series made up of harmonic waves. We consider an artificial signal that is 10 years long, sampled every hour, and composed of the sum of eight sine waves with

the same amplitude and periods of 1, 2, 4, 8, 16, 32, 64, and 128 days, except for the sine with a period of 16 days that has an amplitude threefold higher than the rest. The second signal is constructed in the same way, except that the amplitude for the sine waves corresponding to a period of 16 and 64 days are also threefold greater than the others. The TC98 MATLAB program is used, and the wavelet parameters are chosen as those for the geophysical data analyzed below in section 4.

The time series of the first artificial signal is shown in the upper panel of Fig. 1. The localized wavelet power spectrum and the time-averaged wavelet spectrum (called “global wavelet spectrum” in TC98) are shown in the middle panel of Fig. 1. The cone of influence (COI), that is, the region of the wavelet spectrum where the edge effects become important, is treated in the usual way (TC98; Grinsted et al. 2004; L07), and it is shown as a white shaded area. As was first pointed out by L07, the spectral peaks are distorted, with the low-frequency peaks at almost the same value as the high-frequency peaks. However, when the spectrum is divided by the scale s , this problem is rectified, as in L07, except that the relevant peak (the one corresponding to the period of 16 days) appears isolated, and the others recover the relative height. The same analysis

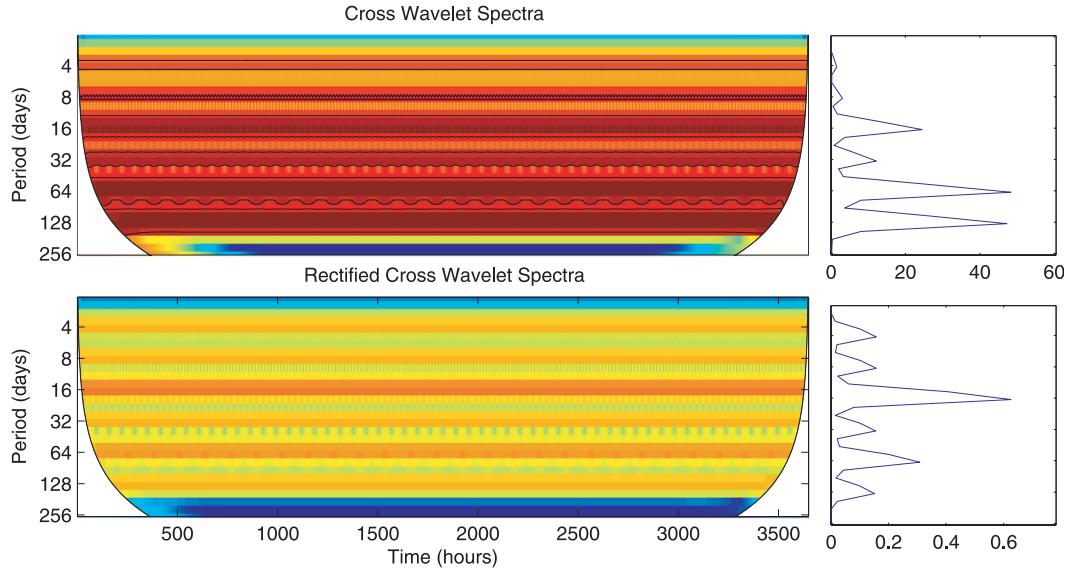


FIG. 3. (top) Standard (not rectified) and (bottom) rectified XWT from the wavelet analysis of the two artificial time series. (right) Global XWT. High (red contours) and low (blue contours) XWT values are indicated. The COI, where zero padding has reduced the variance, is shown (white shaded regions).

is performed on the second artificial signal (Fig. 2). As in the previous signal, the first panel shows the time series of the signal and the middle panel shows the localized wavelet power spectrum and the time-averaged wavelet spectrum. Despite the fact that the peaks of 16- and 64-day components have the same amplitude, they show different spectrum amplitudes. When the rectified wavelet power spectra are calculated the correct height of the peaks show up, that is, both peaks exhibit the same height.

The above results and corrections were clearly explained by L07. We move a step forward, calculating the XWT of the two signals, as defined in TC98, using the “standard” method and with the rectified wavelet. As was explained in section 2, each wavelet has to be rectified by the corresponding scale $2^{j/2}$, and after that the XWT is calculated.

Again, the XWT is meant to identify a peak for the common frequency (period). Instead, what we find with the biased XWT at both frequencies, corresponding to periods of 16 and 64 days, are equally relevant (upper panel of Fig. 3). This result would indicate that the two signals have a resonant frequency at a 16- and 64-day period. This result would incorrectly lead to the conclusion that the two signals are resonant for these two periods, and this inaccurate result can be fully corrected with the use of the rectified XWT. When the scale factor is considered, only the resonant period of 16 days appears, as can be seen in the second panel of Fig. 3, which is in agreement with the resonance induced by construction of the signals.

This artificial example clearly shows how an extremely biased XWT erroneously identified the resonant frequencies (periods) of the two signals as being two different frequencies, instead of just one, as is actually the case.

4. Correction applied to a specific geophysical problem

In the previous section, we showed that the cross-wavelet spectrum without the appropriate normalization disguises the real resonant frequencies. We now turn to an example of two weakly coupled geophysical signals: the relative humidity (%) and the shortwave radiation (W m^{-2}), collected from the ATLAS buoy moored at 8°S, 30°W, as part of PIRATA (Servain et al. 1998; Bourlés et al. 2008; NOAA/IRD/INPE/DHN 2011). As previously stated, these are important variables for the prediction of easterly atmospheric wave propagation and extreme rainfall events at the eastern boundary of Northeast Brazil (Nobre et al. 2004; Rao et al. 1993).

The effect of the clouds as regulators of the radiative heating of the surface is also well known (Ramanathan et al. 1989, 1995). In addition, the relative humidity directly effects cloudiness and, hence, the net radiation at the surface (Ododo 1997; Fotiadi et al. 2005; Lorenz et al. 2010). The low-level relative humidity affects low-level clouds; hence, the relative humidity affects the shortwave radiation (Lorenz et al. 2010). Therefore, the relative humidity and the shortwave radiation are coupled. We use the XWT to reveal this coupling. Without the

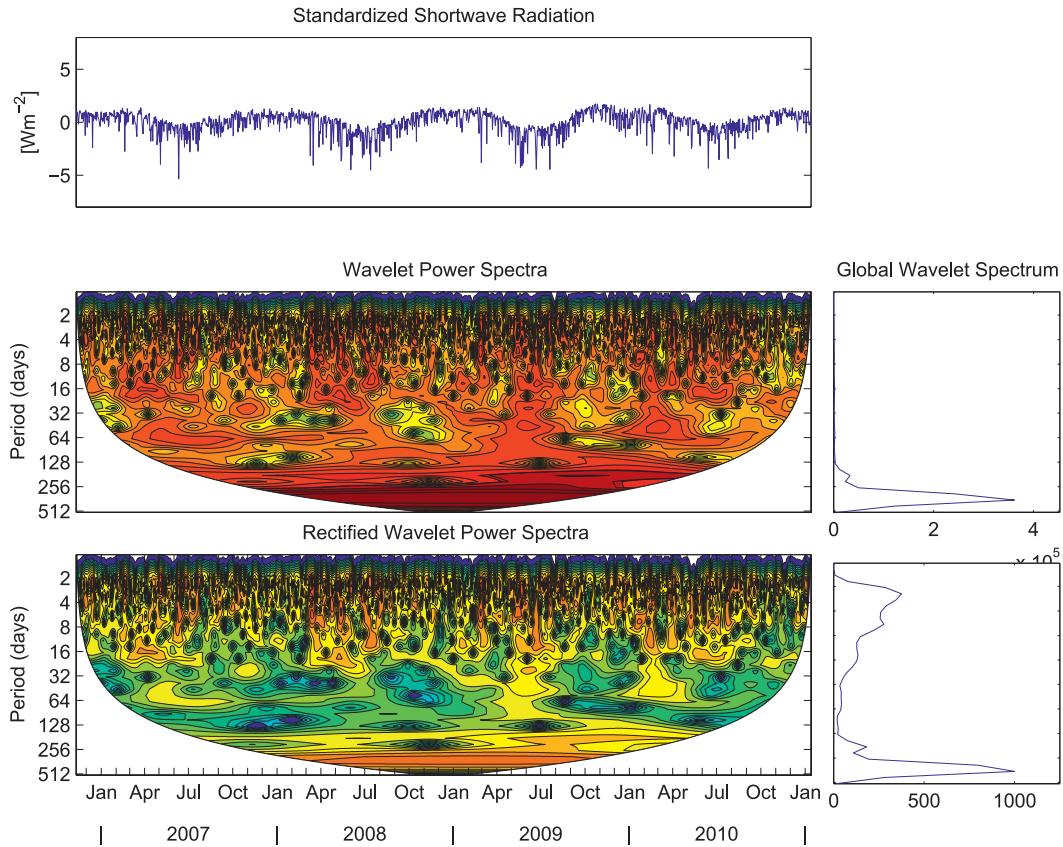


FIG. 4. (top) Signal of the daily shortwave radiation at the 8°S , 30°W PIRATA buoy. (middle) Wavelet power spectra of the shortwave radiation and the respective global wavelet spectrum. (bottom) Rectified wavelet power spectra of the shortwave radiation and (right) the respective global wavelet spectrum. The COI, where zero padding has reduced the variance, is shown (white shaded regions).

appropriate normalization, the real frequencies linking the two phenomena will be masked, as shown below.

We perform a wavelet analysis on both series, and then we quantify the relationship throughout the cross-wavelet spectrum. The results of this analysis are shown in Figs. 4–6.

The oscillations of the shortwave radiation time series are dominated by the annual mode. This result can clearly be seen in the upper panel of Fig. 4, which shows the standardized time series. The yearly modulation can be observed in the time series and in the wavelet spectrum in the middle panel of the same figure. In the same wavelet power spectrum, an intermittent oscillation mode at around the 60-day period broad peak can also be seen. Additionally, significant oscillations of $\approx 2\text{--}16$ days arise mainly from April to July every year. The time-averaged wavelet power spectrum shown in the right column (middle panel) clearly indicates that the low frequencies (365-day peak) are somehow overestimated. Once the correction is applied, the wavelet power spectrum reveals that the high frequencies ($\approx 2\text{--}16$ days) are also important and that the 60-day signal fades out as can be seen in the lower panel of Fig. 4

and in the corresponding time-averaged wavelet power spectrum in the right column.

The same analysis for the relative humidity, shown in Fig. 5, has a more dramatic outcome. The wavelet analysis reveals only a main annual signal (middle panel of Fig. 5), while the corrected wavelet power spectrum clearly signalize the ≈ 6 -day period oscillation mode as twice as important as the 365-day mode (lower panel of Fig. 5).

To investigate the connection between the shortwave radiation and the relative humidity, we make use of cross-wavelet analysis. The standard cross-wavelet analysis applied to these time series shows nothing but a significant common mode of oscillation of a 365-day period, as can be seen in the upper panel of Fig. 6. However, when the correction proposed in section 2 is applied to the XWT, the result is characterized by not only the 365-day period, but also by a broad band of $\approx 2\text{--}8$ -day mode oscillation (lower panel of Fig. 6).

This example clearly shows how a biased cross-wavelet spectrum can mask a significant common mode of oscillation hiding some relevant phenomena.

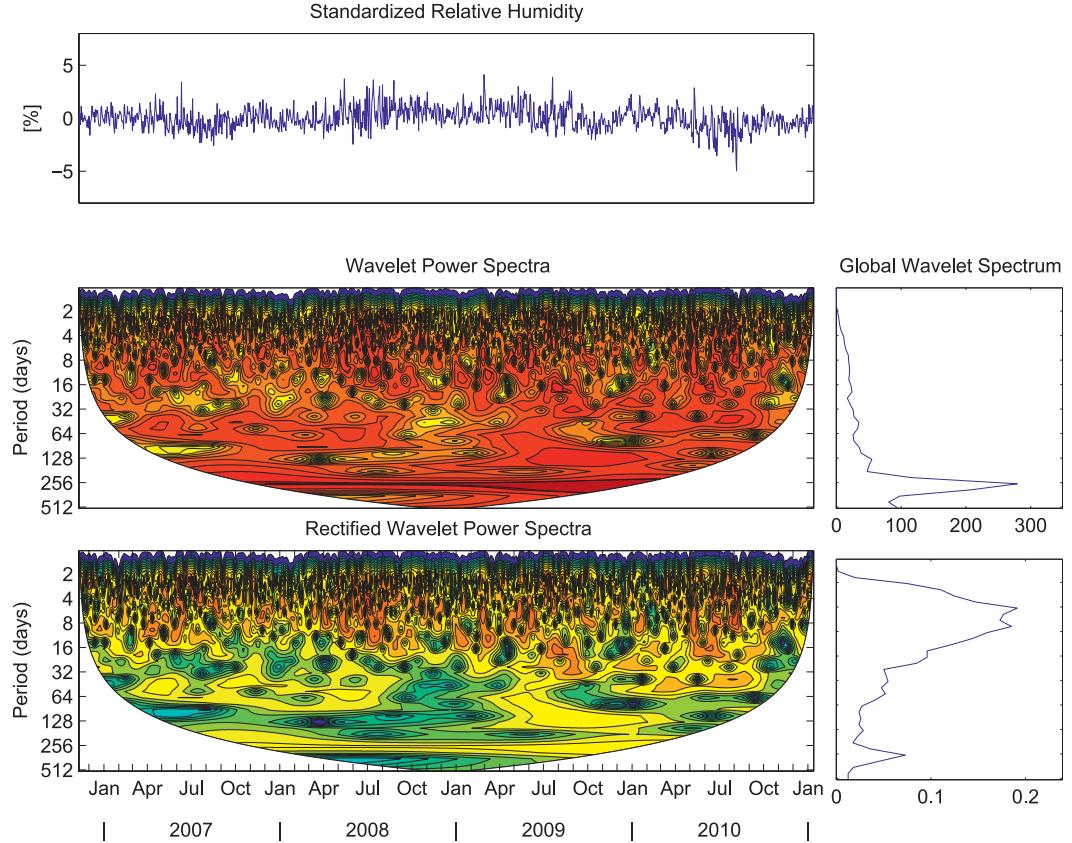


FIG. 5. (top) Signal of the daily relative humidity at the 8°S, 30°W PIRATA buoy. (middle) Wavelet power spectra of the relative humidity and the respective global wavelet spectrum. (bottom) Rectified wavelet power spectra of the relative humidity and (right) the respective global wavelet spectrum. The COI, where zero padding has reduced the variance, is shown (white shaded regions).

Namely, if the cross-wavelet spectrum is not rectified, the high frequencies (low periodicity signals) can be underestimated.

5. Summary and discussion

Cross-wavelet analysis is a useful instrument to investigate the likely relationship between two time series signals. A distorted or biased result can completely mask the connection between the involved phenomena. The cross-wavelet transform (XWT), as produced by the broadly used wavelet analysis, such as TC98, may generate incorrect distortions by giving more weight to the large-scale linking phenomena than to the small-scale phenomena. In this paper, we review the theoretical basis that leads to this bias in the power spectrum and how it was corrected by L07. We rewrite the correction to explicitly show the scale correction, making it self-evident how the XWT must be corrected. The scale correction, that is, the square root of the scale, is applied on the wavelet transform of each time series before taking the

inner product that yields the XWT. This correction not only avoids the bias, but is also physically consistent. The rectification of the cross wavelet was tested with two artificial time series composed of sine waves of known amplitudes and frequencies. The normalizing process worked as expected, showing the real resonant frequencies. The same rectification approach was also applied to the relative humidity and the shortwave radiation dataset collected by the buoy at 8°S, 30°W of the Southwestern Extension of the Prediction and Research Moored Array in the Tropical Atlantic (PIRATA; Servain et al. 1998; Bourlès et al. 2008; NOAA/IRD/INPE/DHN 2011). The rectified XWT confirmed in this case that the two phenomena are linked not only in the low frequencies, as is shown with the traditional biased XWT, but also in the high frequencies.

We trust that the insights into the XWT presented in this paper can produce more robust methods for researching for the analysis of linked phenomena. We also expect this correction minimize the risk of missing important coupled relationships.

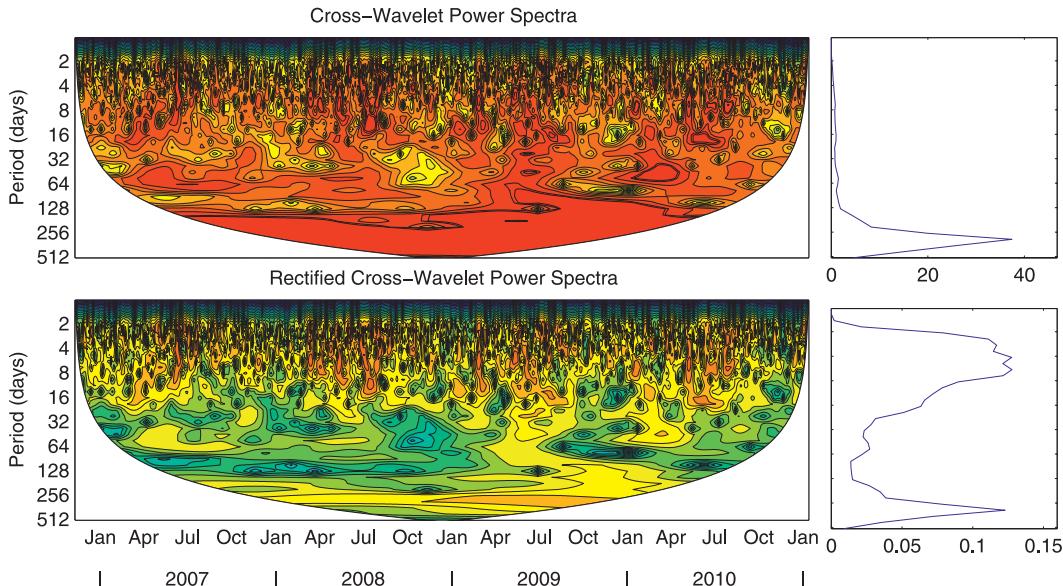


FIG. 6. (top) The XWT between the relative humidity and the shortwave radiation at the 8°S, 30°W PIRATA buoy. (bottom) Rectified XWT between the relative humidity and the shortwave radiation at the 8°S, 30°W PIRATA buoy. (right) XWT. The COI, where zero padding has reduced the variance, is shown (white shaded regions).

Acknowledgments. The authors want to acknowledge the scientific and crew members of the R/V *Antares* of the Brazilian Navy for their effort and dedication during the PIRATA-SWE cruises. This work was carried out under the Project BIO-NE, CNPq Process 558143/2009-1. DV acknowledges financial support from the National Council of Scientific and Technological Development (CNPq, Brazil). RM acknowledges support from FACEPE, Grant APQ-0871-1.05/08 and CNPq Process 303718/2010-2. Wavelet software was provided by C. Torrence and G. Compo, and is available at <http://atoc.colorado.edu/research/wavelets/>. The authors also acknowledge enlightening comments of Y. Liu.

REFERENCES

- Bourlés, B., and Coauthors, 2008: The PIRATA program: History, accomplishments, and future directions. *Bull. Amer. Meteor. Soc.*, **89**, 1111–1125.
- Daubechies, I., 1992: *Ten Lectures on Wavelets*. CBMS/NSF Series in Applied Mathematics, Vol. 61, SIAM, 377 pp.
- , 1999: The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inf. Theory*, **36**, 961–1005.
- Farge, M., 1992: Wavelet transforms and their applications to turbulence. *Annu. Rev. Fluid Mech.*, **24**, 395–458.
- Fotiadi, A., N. Hatzianastassiou, C. Matsoukas, K. G. Pavlakis, E. Drakakis, D. Hatzidimitriou, and I. Vardavas, 2005: Analysis of the decrease in the tropical mean outgoing shortwave radiation at the top of atmosphere for the period 1984–2000. *Atmos. Chem. Phys.*, **5**, 1721–1730.
- Grinsted, A., J. Moore, and S. Jevrejeva, 2004: Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes Geophys.*, **11**, 561–566.
- Kumar, P., and E. Foufoula-Georgiou, 1997: Wavelet analysis for geophysical applications. *Rev. Geophys.*, **35**, 385–412.
- Liu, Y., X. S. Liang, and R. H. Weisberg, 2007: Rectification of the bias in the wavelet power spectrum. *J. Atmos. Oceanic Technol.*, **24**, 2093–2102.
- Lorenz, D. J., E. T. DeWeaver, and D. J. Vimont, 2010: Evaporation change and global warming: The role of net radiation and relative humidity. *J. Geophys. Res.*, **115**, D20118, doi:10.1029/2010JD013949.
- Mallat, S., 1998: *A Wavelet Tour of Signal Processing*. Academic Press, 577 pp.
- NOAA/IRD/INPE/DHN, cited 2011: PIRATA. [Available online at <http://www.pmel.noaa.gov/pirata/>].
- Nobre, P., E. Campos, P. S. Polito, O. T. Sato, and J. A. Lorenzetti, 2004: PIRATA western extension scientific rational report. INPE/CPTEC Tech. Rep., 43 pp.
- Ododo, J. C., 1997: Prediction of solar radiation using only maximum temperature and relative humidity: South-east and north-east Nigeria. *Energy Convers. Manage.*, **38**, 1807–1814.
- Ramanathan, V., R. Cess, E. Harrison, P. Minnis, B. Barkstrom, E. Ahmad, and D. Hartmann, 1989: Cloud-radiative forcing and climate: Results from the earth radiation budget experiment. *Science*, **243**, 57–63.
- , B. Subasilar, G. Zhang, W. Conant, R. Cess, J. Kiehl, H. Grassi, and L. Shi, 1995: Warm pool heat budget and shortwave cloud forcing: A missing physics? *Science*, **267**, 499–503.
- Rao, V. B., M. C. Lima, and S. H. Franchito, 1993: Seasonal and interannual variations of rainfall over eastern northeast Brazil. *J. Climate*, **6**, 1754–1763.
- Servain, J., A. J. Busalacchi, M. J. McPhaden, A. D. Moura, G. Reverdin, M. Vianna, and S. E. Zebiak, 1998: A pilot research moored array in the Tropical Atlantic (PIRATA). *Bull. Amer. Meteor. Soc.*, **79**, 2019–2031.
- Torrence, C., and G. P. Compo, 1998: A practical guide to wavelet analysis. *Bull. Amer. Meteor. Soc.*, **79**, 61–78.
- Zunino, L., D. Pérez, M. Martín, A. Plastino, M. Garavaglia, and O. Rosso, 2007: Characterization of Gaussian self-similar stochastic processes using wavelet-based informational tools. *Phys. Rev. E*, **75**, 021115, doi:10.1103/PhysRevE.75.021115.