

Brief communication

“A statistical validation for the cycles found in air temperature data using a Morlet wavelet-based method”

S. Nicolay¹, G. Mabille², X. Fettweis², and M. Erpicum²

¹Département de mathématique, University of Liège, Liège, Belgium

²Département de géographie, University of Liège, Liège, Belgium

Received: 23 February 2010 – Revised: 29 April 2010 – Accepted: 8 May 2010 – Published: 20 May 2010

Abstract. Recently, new cycles, associated with periods of 30 and 43 months, respectively, have been observed by the authors in surface air temperature time series, using a wavelet-based methodology. Although many evidences attest the validity of this method applied to climatic data, no systematic study of its efficiency has been carried out. Here, we estimate confidence levels for this approach and show that the observed cycles are significant. Taking these cycles into consideration should prove helpful in increasing the accuracy of the climate model projections of climate change and weather forecast.

1 Introduction

Cycles associated with periods longer than one year in surface air temperature data have been observed in several studies (see e.g. Paluš and Novotná, 2006; Nicolay et al., 2009; Mabille and Nicolay, 2009; Matyasovszky, 2010). However, there is no clear evidence concerning the efficiency of these methods: Can we trust the results or is there a high probability that these oscillations occurred by pure chance? To answer this question, we estimate such a probability for the approach proposed in Nicolay et al. (2009) and show that the cycles are indeed significant.

This therefore shows that the results previously obtained in this work are meaningful. In particular, it suggests that a cycle of about 30 months is coupled with the Arctic/North Atlantic Oscillation, while a cycle associated with a period of 43 months can be affiliated to the El Niño Southern Oscillation. In Nicolay et al. (2009), the temperature fluctuations induced by such cycles are estimated to be about one-tenth of the annual amplitude.

2 Method

We first briefly describe the wavelet-based methodology used in Nicolay et al. (2009) to study the existence of cycles in climatic data. We then explain how to build confidence levels for the observed cycles.

2.1 The wavelet spectrum

The wavelet spectrum is a tool designed for spectral studies; it is described in further details in Nicolay et al. (2009); Mabille and Nicolay (2009) for example.

A wavelet ψ is a function defined on the real line which

- is integrable,
- is square-integrable,
- satisfies the relation $\int |\hat{\psi}|^2 |\omega|^{-1} d\omega < \infty$, where $\hat{\psi}$ denotes the Fourier transform of ψ .

The wavelet transform of a square-integrable function f defined on the real line is the following function,

$$Wf(t, a) = \int \bar{\psi} \left(\frac{x-t}{a} \right) f(x) \frac{dx}{a},$$

with $t \in \mathbf{R}$ and $a > 0$, where $\bar{\psi}$ denotes the complex conjugate of ψ . The wavelet transform can be seen as a mathematical microscope for which position and magnification correspond to t and $1/a$ respectively, the performance of the optic being determined by the wavelet (see e.g. Arneodo et al., 1988; Freysz et al., 1990). The function f can be recovered, in some way, from its wavelet transform (see e.g. Daubechies, 1992 for more details).



Correspondence to: S. Nicolay
(s.nicolay@ulg.ac.be)

If ψ is a wavelet such that $\hat{\psi}(\omega) = 0$ whenever $\omega \leq 0$, the wavelet spectrum Λ of a square-integrable function f is defined as

$$\Lambda(a) = E|Wf(t, a)|,$$

where E denotes the mean over time t . Let us remark that our wavelet spectrum is not defined in terms of density, but rather as a marginal spectrum (see Huang et al., 1998). The use of the mean, rather than the root mean square allows a better balance between small and large values: small values are less taken into account when using the root mean square. This can be more rigorously stated, since the mean square is equal to the square of the arithmetic mean plus the variance. However, it can be shown (see Nicolay et al. (2009)) that such a spectrum leads to frequential informations, in a way analogous to that of the Fourier spectrum: if a_0 is a maximum of $\Lambda(a)$, there is a higher likelihood for a sine (or cosine) wave with period a_0 to have appeared locally (see Huang et al., 1998; Nicolay et al., 2009). This is therefore a “local approach”; such a technique is thus different from the Fourier spectrum, since this method is related to waves that persisted through the whole time span. Moreover, unlike the Fourier spectrum, the wavelet spectrum is not affected by trends.

The wavelet ψ used in this work is the accustomed Morlet wavelet, whose Fourier transform is approximated by

$$\hat{\psi}(\omega) = \exp\left(-\frac{(\omega - \Omega)^2}{2}\right),$$

where $\Omega = \pi\sqrt{2/\log 2}$. Such a tool is well designed for non-stationary signals; it has been successfully applied to climatic data, where it has led to the detection of cycles in air temperature time series associated with the periods of 30 and 43 months respectively (see Nicolay et al., 2009; Mabile and Nicolay, 2009).

2.2 Construction of confidence levels for the observed cycles

In order to test the significance of the detected cycles on the whole planet, the NCEP/NCAR reanalysis time series (Kalney et al., 1996) were selected as a gridded ($2.5^\circ \times 2.5^\circ$) data set. These signals present the state of the Earth’s atmosphere, incorporating observations and global climate model output. Let us notice that the data associated with the oceans must be carefully interpreted, since the number of observations for such grid points is rather small. The signals are monthly-sampled and start in 1948.

Let us first remark that since we are not using the accustomed definition of wavelet spectrum, the usual significance tests can not be applied here. Moreover, such methods have many pitfalls, especially when the model is not a Gaussian white noise (see e.g. Maraun and Kurths, 2004). Nevertheless, it could be interesting to adapt such methods to the definition given here.

To check if the cycles appearing in the time series did not occur by pure chance, the following methodology has been applied to each grid point:

1. the linear trends are first removed,
2. the climatological anomaly time series is computed: for each month, the mean temperature is computed from the whole signal and the so-obtained monthly-sampled signal M is then subtracted from the original one,
3. each anomaly time series is fitted separately using an autoregressive model of the first order (AR(1) model, see e.g. Percival and Walden, 1993):

$$x_n = \alpha x_{n-1} + \sigma \eta_n,$$

where η_n is a Gaussian white noise with zero mean and unit variance. Such processes are present in many climatic and geophysical data (see e.g. Percival and Walden, 1993; Allen and Robertson, 1996) and are well suited for the study of climatic time-series (Mann and Lees, 1996; Mann et al., 2007),

4. the mean M is then added to the simulated noise in order to obtain a simulation of the time series,
5. $N=10\,000$ such simulations are computed,
6. the distribution of the highest maximum y_M of the wavelet spectrum of the data in the range of 26 to 47 months (in order to compare these maxima with the ones detected in Nicolay et al., 2009) is estimated from these realizations, i.e. one computes the distribution of

$$y_M = \sup_{26 \leq a \leq 47} \tilde{\Lambda}(a),$$

where $\tilde{\Lambda}$ is the wavelet spectrum of a realization,

7. the probability P to obtain a maximum of higher amplitude than the one corresponding to the period of 30 months (or 43 months) observed in the wavelet spectrum of the grid point is finally computed, using the distribution obtained in 6.

Let us remark that the 30 and 43 months period cycles can also be detected through the Fourier transform; however, the signals have to be preprocessed first, since the lower amplitude cycles are hidden by the dominating cycle corresponding to 1 year. The two lower-frequency cycles are clearly observed if

- the linear trends are removed,
- only the climatological anomaly time series is kept.

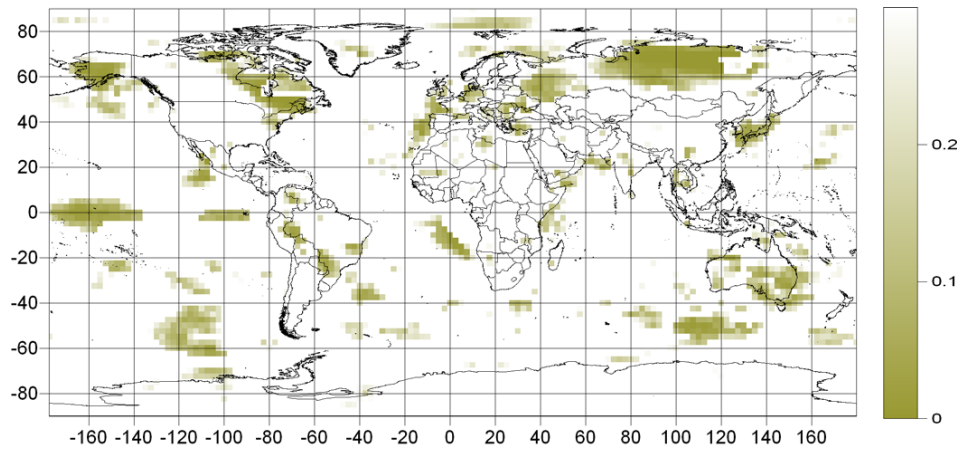


Fig. 1. The probability values associated with the cycle of 30 months (see Sect. 2.2). The cycles observed in a zone corresponding to the colour white are not significant.

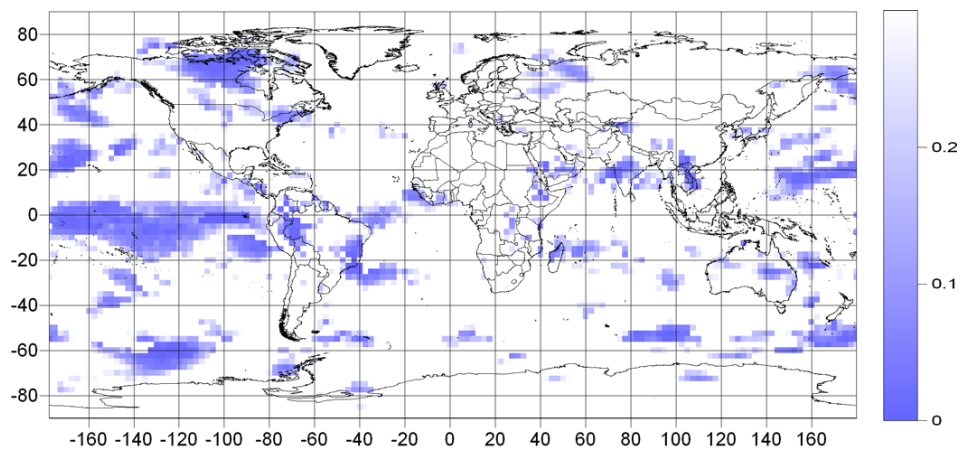


Fig. 2. The probability values associated with the cycle of 43 months (see Sect. 2.2). The cycles observed in a zone corresponding to the colour white are not significant.

The same methodology as above can then be applied to such signals, replacing the wavelet spectrum with the Fourier spectrum. The results are identical to the ones obtained with the wavelet spectrum (data not shown). However, since the Fourier spectrum is not well suited for non-stationary signals and since one could claim that the so-detected periods could result from the data preprocessing, the wavelet transform has been preferred in the present study (see also Nicolay et al., 2009).

3 Results

First of all, we have to ensure that the distributions associated to the highest maxima, obtained following the methodology described above, yield reliable information. All the probability density functions (PDF's) are unimodal and nearly mesokurtic. Moreover, the skewness is always lower than

0.5 and the test of D'Agostino leads to the conclusion that 10% of these PDF's are compatible with a Gaussian distribution ($\alpha=0.05$). Finally, when comparing one of the 10 000 AR(1) realizations with the 9999 others, the probability value P is almost always higher than 0.5; indeed 0.3% of the grid nodes only can be associated with a value P lower than 0.3.

The probability values concerning the 30 months cycle are displayed in Fig. 1; the probability values concerning the 43 months cycle are displayed in Fig. 2. The coloured area corresponds to regions where the cycle is significant (the more the area is coloured, the more the region is significant). These planispheres look very similar to the ones obtained concerning the existence of the corresponding cycles in Nicolay et al. (2009). In other words, the cycles associated with the period of 30 and 43 months previously observed are significant.

The cycle associated with 30 months is mainly seen in Europe, Northern Asia, Alaska and Eastern Canada, while the cycle associated with 43 months is principally observed in Northern America, Peru and Equatorial Pacific. The existence of a cycle of period close to 30 months in Central Europe has already been shown in Paluš and Novotná (2006). Moreover, these cycles seem to be related to climatic indices. Most of the regions affected by the 30 months period cycle correspond to the area under the influence of the North Atlantic Oscillation, while the ones affected by the 43 months period cycle coincide with the area affected by the Southern Oscillation. As a matter of fact, the wavelet spectra of the related indices display the corresponding period: a cycle associated with 30 months is significantly observed in the AO/NAO (CPC) indices ($P < 0.01$) and a cycle corresponding to 43 months is detected in the global-SST ENSO and ENSO MEI indices ($P < 0.01$). Let us also remark that a 30 months period cycle is observed in Australia and over some parts of the oceans; this could be a byproduct of the El Niño phenomenon, since a small maximum corresponding to 30 months is also observed in the Southern Oscillation (see Nicolay et al., 2009).

4 Conclusions

We have built confidence levels for the cycles corresponding to the periods of 30 and 43 months previously observed in NCEP/NCAR reanalysis data, using a wavelet-based methodology (see Nicolay et al., 2009). To do so, we have simulated the background noise with an autoregressive model of the first order. From this point of view, the influence of these cycles is significant in each area where such an oscillation is observed.

Following the observations of Nicolay et al. (2009); Mabile and Nicolay (2009), these results suggest that the El Niño Southern Oscillation is related to a cycle of period of about 43 months in the near-surface air temperatures, whereas a 30 months period cycle in these data can be associated with the Arctic/North Atlantic Oscillation. Some regions of the globe (Equatorial Pacific, Canada, etc.) can be associated to both cycles (see Nicolay et al., 2009).

Recent results (Smith et al., 2007) have shown that modeling systems that predict both internal variability and externally forced changes forecast surface temperature with substantially improved skill. Climate models accounting for the cycles associated with 30 and 43 months should therefore predict surface temperature more accurately.

Acknowledgements. The authors would like to express their gratitude to P. Ditlevsen for providing fruitful discussions and thank the anonymous referees for useful comments.

Edited by: J. Kurths

Reviewed by: three anonymous referees

References

- Allen, M. R. and Robertson, A. W.: Distinguishing modulated oscillations from coloured noise in multivariate datasets, *Clim. Dynam.*, 12, 775–784, 1996.
- Arneodo, A., Argoul, F., Elezgaray, J., and Grasseau, G.: Wavelet transform analysis of fractals: application to nonequilibrium phase transitions, in: *Nonlinear Dynamics*, edited by: Turchetti, G., World Scientific, Singapore, 1988.
- Daubechies, I.: *Ten lectures on wavelets*, SIAM, Philadelphia, 1992.
- Freysz, E., Pouligny, B., Argoul, F., and Arneodo, A.: Optical wavelet transform of fractal aggregates, *Phys. Rev. Lett.*, 64, 745–748, 1990.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. L. C., Shih, H. H., Zheng, Q. N., Yen, N. C., Tung, C. C., and Liu, H. H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. R. Soc. Lon. Ser. A*, 454, 903–995, 1998.
- Kalney, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-Year reanalysis project, *B. Am. Meteorol. Soc.*, 104, 437–471, 1996.
- Mabile, G. and Nicolay, S.: Multi-year cycles observed in air temperature data and proxy series, *Eur. Phys. J.-Spec. Top.*, 174, 135–145, 2009.
- Mann, M. E. and Lees, J.: Robust estimation of background noise and signal detection, *Climatic Change*, 33, 409–445, 1996.
- Mann, M. E., Rutherford, S., Wahl, E., and Ammann, C.: Robustness of proxy-based climate field reconstruction methods, *J. Geophys. Res.*, 112, D12109, doi:10.1029/2006JD008272, 2007.
- Maraun, D. and Kurths, J.: Cross wavelet analysis: significance testing and pitfalls, *Nonlin. Processes Geophys.*, 11, 505–514, doi:10.5194/npg-11-505-2004, 2004.
- Matyasovszky, I.: Improving the methodology for spectral analysis of climatic time series, *Theor. Appl. Climatol.*, doi:10.1007/s00704-009-0212-z, in press, 2010.
- Nicolay, S., Mabile, G., Fettweis, X., and Erpicum, M.: 30 and 43 months period cycles found in air temperature time series using the Morlet wavelet method, *Clim. Dynam.*, 33, 1117–1129, 2009.
- Paluš, M. and Novotná, D.: Quasi-biennial oscillations extracted from the monthly NAO index and temperature records are phase-synchronized, *Nonlin. Processes Geophys.*, 13, 287–296, doi:10.5194/npg-13-287-2006, 2006.
- Percival, D. B. and Walden, A. T.: *Spectral analysis for physical applications*, Cambridge University Press, Cambridge, 1993.
- Smith, D., Cusack, S., Colman, A., Folland, C., Harris, G., and Murphy, J.: Improved surface temperature prediction for the coming decade from global climate model, *Science*, 317, 796–799, 2007.