Several years El Niño forecast using a wavelet-based mode decomposition

S. Nicolay & X. Fettweis

EGU General Assembly 2012

4/23/2012

From the Fourier series

A Fourier series decomposes a periodic signal into a possibly infinite sum of sines and cosines functions. From a practical point of view, a signal f is decomposed as a sum of K cosines:

$$f(t) \approx \sum_{k=1}^{K} c_k \cos(\omega_k t + \phi_k).$$

One thus only gets an approximation of the original signal.

From the Fourier series

A Fourier series decomposes a periodic signal into a possibly infinite sum of sines and cosines functions. From a practical point of view, a signal f is decomposed as a sum of K cosines:

$$f(t) \approx \sum_{k=1}^{K} c_k \cos(\omega_k t + \phi_k).$$

One thus only gets an approximation of the original signal. Such a decomposition often leads to a decomposition with too many terms, i.e. with a too large K.

...to a wavelet-based decomposition

The idea underlying the wavelet-based mode decomposition is to decrease this number of terms by considering the amplitudes c_k as functions of t.

...to a wavelet-based decomposition

The idea underlying the wavelet-based mode decomposition is to decrease this number of terms by considering the amplitudes c_k as functions of t.
One thus tries to have the following decomposition,

$$f(t) \approx \sum_{j=1}^{J} C_j(t) \cos(\omega_j t + \phi_j),$$

where J should be much smaller than K.

The wavelet spectrum

We use the progressive wavelet

$$\psi(t) = \frac{1}{2\sqrt{2\pi}} \exp(i\Omega t) \exp(-\frac{(2\Omega t + \pi)^2}{8\Omega^2})(\exp(\frac{\pi t}{\Omega}) + 1),$$

with $\Omega = \pi\sqrt{2/\ln 2}$.

The wavelet spectrum

We use the progressive wavelet

$$\psi(t) = \frac{1}{2\sqrt{2\pi}} \exp(i\Omega t) \exp\left(-\frac{(2\Omega t + \pi)^2}{8\Omega^2}\right)\left(\exp\left(\frac{\pi t}{\Omega}\right) + 1\right),$$

with $\Omega = \pi\sqrt{2/\ln 2}$.
The wavelet spectrum associated to the signal f is defined as

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

where Wf is the continuous wavelet transform of f and E denotes the mean over the time t.

The wavelet decomposition

The scales $a_1, \ldots, a_J$ for which the spectrum Λ of f reaches a maximum are kept (let us recall that a scale behaves as the inverse of a frequency).

The wavelet decomposition

The scales $a_1, \ldots, a_J$ for which the spectrum Λ of f reaches a maximum are kept (let us recall that a scale behaves as the inverse of a frequency).
The signal f is then decomposed in the following way,

$$f(t) = f_0(t) + \sum_{j=1}^{J} f_j(t),$$

with

$$f_j(t) = |Wf(t, a_j)| \cos\big(\arg Wf(t, a_j)\big),$$

if $j \geq 1$ and $f_0 = f - \sum_{j=1}^{J} f_j$.

Approximation of the signal

The signal $\sum_{j=1}^{J} f_j$ can be seen as an approximation of the original data in terms of oscillatory components.

Approximation of the signal

The signal $\sum_{j=1}^{J} f_j$ can be seen as an approximation of the original data in terms of oscillatory components.

If we assume that the mean frequency of a component f_j is constant, the reconstructed signal has the following form,

$$f(t) \approx \hat{f}(t) = \sum_{j=1}^{J} C_j(t) \cos(\omega_j + \phi_j),$$

where ω_j corresponds to the mean (angular) frequency.

Extrapolation

The amplitude C_j is supposed to vary smoothly; its values can therefore be extrapolated using Lagrange polynomials.

Extrapolation

The amplitude C_j is supposed to vary smoothly; its values can therefore be extrapolated using Lagrange polynomials.
The computation was based on the last 2 available data of the signal.

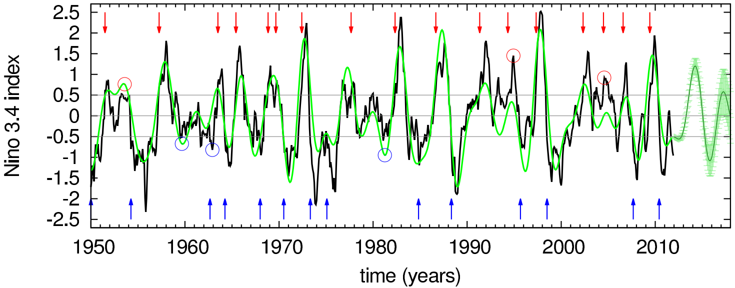
Extrapolation

The amplitude C_j is supposed to vary smoothly; its values can therefore be extrapolated using Lagrange polynomials.
The computation was based on the last 2 available data of the signal.
Summing the so-obtained extrapolated values related to each component defines the forecast; more precisely, it corresponds to a forecast of the reconstructed signal $\hat{f}$.

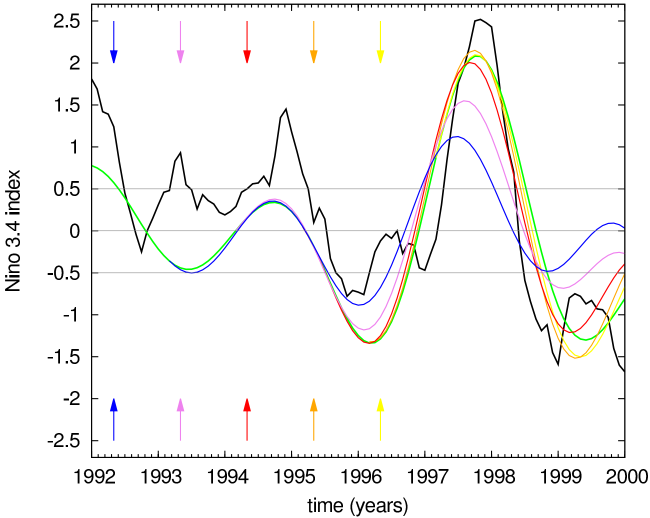
Extrapolation

The amplitude C_j is supposed to vary smoothly; its values can therefore be extrapolated using Lagrange polynomials.
The computation was based on the last 2 available data of the signal.
Summing the so-obtained extrapolated values related to each component defines the forecast; more precisely, it corresponds to a forecast of the reconstructed signal $\hat{f}$.
The efficiency of the predictions was tested with probing hindcasts.

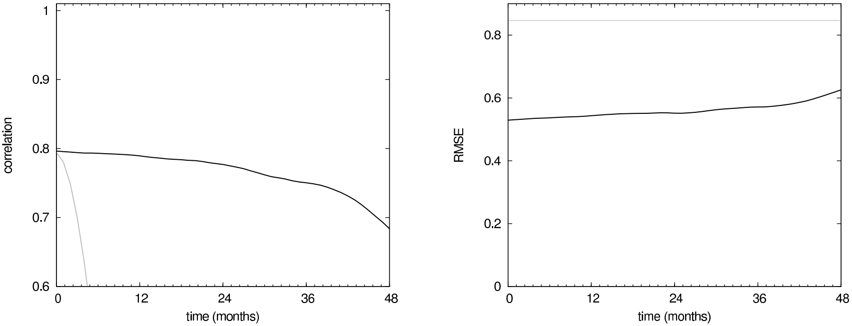
Application to the Nino 3.4 index



Forecast of the 1997–1998 El Niño event



Anomaly correlation and RMSE



The end

Thanks for your attention!