Correlation analysis for unevenly spaced environmental time series

Josué. M. Polanco-Martínez[1], Dae-Jin Lee[2]

[1]josue.polanco@bc3research.org, Basque Centre for Climate Change - BC3
[2]dlee@bcamath.org, Basque Center for Applied Mathematics - BCAM

One of the main approaches used to estimate the correlation between two environmental time series are the Pearson's and Spearman's correlation coefficients or the estimation of the cross-correlation function. This task is straightforward when the time series are unevenly spaced, but it is not trivial for unevenly spaced time series. The most common way to face this problem is to interpolate in time the original unevenly spaced time series in order to obtain equidistance. Then, these series can be analysed using the existing standard correlation analysis techniques. However, a great deal of research has shown that interpolation has its drawbacks: interpolation is highly dependent of the kind of interpolation applied, the interpolated time series may show deviations in terms of variability or noise properties, and additional serial dependence can be introduced (Mudelsee 2014; Rehfeld et al., 2011). Thus, interpolation should be avoided as much as possible.

In order to face this challenge, in this work, we present a method to obtain directly the correlation between two irregularly sampled environmental time series. We will explore the computational implementation of this technique in R. Moreover, we will examine how this tool can help to obtain more information between unevenly spaced environmental time series with special emphasis in paleoclimate data.

Keywords: Environmental time series analysis, paleoclimate data, missing data interpolation

References

Mudelsee M (2014). Climate Time Series Analysis: Classical Statistical Bootstrap Methods, Springer, Dordrech, Germany.

Rehfeld et al (2011). Comparison of correlation analysis techniques for irregularly sampled time series. Nonlin. Processes Geophys. 18, 389-404.