Proceedings of the 2nd International Conference on Statistics: Theory and Applications (ICSTA'20)
~~Prague, Czech Republic~~ Virtual Conference – August, 2020
Paper No. ICSTA 137
DOI: 10.11159/icsta20.137

Evaluation and Development of Time Series Interpolators

Sophie Castel[1] and Wesley S. Burr[3]
[1]Department of Mathematics, Trent University
1600 West Bank Drive, Peterborough, Ontario, Canada
wesleyburr@trentu.ca

Extended Abstract

The power spectral density, and by extension, the power spectrum, of a given time series is a sufficient statistic, and highly useful for describing the frequency-domain properties of series. Unfortunately, most spectrum estimation methods require time series data to be contiguous in order for robust estimators to retain their performance. This poses a fundamental challenge, especially when considering real-world scientific data that is often plagued by missing values, and/or irregularly recorded measurements.

One area of research devoted to this problem seeks to repair the original time series through interpolation. There are several computational algorithms that have proven successful for the interpolation of considerably large gaps of missing data, but most are only valid for use on stationary time series (processes whose statistical properties are time-invariant), which is not a common property of real-world data. In practice, organic processes of a scientific nature may be complex valued, have nonstationary trends, seasonal periodic variation, and other such idiosyncrasies. The Hybrid Wiener Interpolator (HWI) [1] is a recent method that was developed to remove the restriction of stationarity. It is a multi-step, iterative algorithm that involves the estimation and removal of mean and periodic trends such that classic Wiener interpolation [2] can be applied to the leftover noise component at the last step.

Using this formula, the HWI can be applied to a variety of different types of scientific time series – however, its statistical performance seems to be influenced by the characteristic features of the data, such as the number of and type of embedded periodic and polynomial trends, as well as the shape of the underlying noise structure. The ability of the HWI to repair time series data also seems to depend on gap structure parameters, like the overall proportion of data missing (p), and the lengths of the gap widths (g).

In this work, we present a comprehensive analysis of the performance of the HWI against a number of state-of-the-art interpolation algorithms with the assistance of a custom-designed R package called interpTools [3] that allows for systematic testing on the statistical performance of interpolation methods in light of changes to gap structure and departures from the stationarity assumption.

Most metrics used to quantify statistical accuracy are a function of the deviation between the original and interpolated point. However, the true value of a missing sample at a given index of a stochastic time-ordered process is generally an unknown quantity. To address this significant practical challenge, interpTools provides a means of simulating scientific test series, such that the user may then impose a particular gap structure, perform interpolation using a variety of available methods, and finally, evaluate statistical performance using a variety of different metrics and visualization tools. Among the interpolators tested, we found that the HWI was the least sensitive to changes in gap structure and had the most consistent performance. We did not find evidence of a clear relationship between performance and the number of sinusoidal trends, however, the algorithms seemed to perform better on data containing lower-frequency sinusoidal variation.

References

[1] W. S. Burr. Air Pollution and Health: Time Series Tools and Analysis. PhD thesis, 2012.
[2] N. Wiener. Extrapolation, Interpolation, and Smoothing of Stationary Time Series. The MIT press, 1949.
[3] S. Castel & W.S. Burr. interpTools: Tools for systematic testing and evaluation of interpolation algorithms. GitHub. https://github.com/wesleyburr/interpTools/. June, 2020.