EPSC Abstracts
Vol. 8, EPSC2013-479, 2013
European Planetary Science Congress 2013

Removing the human bias from exoplanetary data de-trending

I. P. Waldmann
(1) University College London, Dept. Physics & Astronomy, Gower Street, WC1E 6BT, London (ingo@star.ucl.ac.uk)

Abstract

Independent Component Analysis (ICA) has recently been shown to be a promising new path in data analysis and de-trending of exoplanetary time series signals. Such approaches do not require or assume any prior or auxiliary knowledge on the data or instrument in order to de-convolve the astrophysical light curve signal from instrument or stellar systematic noise. These methods are often known as 'blind source separation' (BSS) algorithms. Unfortunately all BSS methods suffer from a amplitude and sign ambiguity of their de-convolved components which severely limits these methods in low signal-to-noise (S/N) observations where their scalings cannot be determined otherwise. Here we present a novel approach to calibrate ICA using sparse wavelet calibrators. The Amplitude Calibrated Independent Component Analysis (`ACICA`) allows for the direct retrieval of the independent components' scalings and the robust de-trending of low S/N data. Such an approach gives us an unique and unprecedented insight in the underlying morphology of a data set, making this method a powerful tool for exoplanetary data de-trending and signal diagnostics.

1. Introduction

As we explore smaller and smaller extrasolar planet around ever fainter stars, it is unsurprising that the need for ever more accurate data-calibration and de-trending techniques is a growing one. In the recent past, there has been a notable emergence of so called 'non-parametric' data de-trending algorithms in the fields of transiting extrasolar planet and time-resolved exoplanetary spectroscopy [1, 4, 2, 5, 6]. The use of such 'non-parametric' algorithms is a reactionary response to the difficulties of calibrating and de-trending time series observations when the instrument response function is not known at the precision of the science signal to be extracted. In [5] and [6], we have demonstrated independent component analysis (ICA) as novel de-correlation strategy for exoplanetary time series. ICA [3] belongs to the class of blind-source separation (BSS) algorithms, which attempt to de-correlate an observed mixture of signals into its individual source components without prior knowledge of the original signals nor the way they were mixed together. In this conference and [7], I will demonstrate a marked improvement over previous methods employing a sparse coding scheme for blind-source separation problems.

2. Using sparsity to break degeneracies

Previous blind-source separation algorithms were limited by two major factors: 1) the amount of Gaussian noise in the data, and 2) a sign and amplitude ambiguity of the de-correlated signals. For high signal to noise (S/N) observations, these limitations were circumvented by linearly regressing the de-correlated signals to the lightcurve's out of transit data to fit for the missing amplitude information. However for very low S/N data this methodology proves difficult. Here we present a novel approach featuring the introduction of sparse calibration signals in multi-resolution orthogonal wavelet space. In wavelet space, it becomes possible to 1) suppress Gaussian noise by selectively filtering for it and 2) to introduce artificial, time and frequency localised calibrator impulses which can be shown not to impact the original observed data. These calibrator pulses allow us to break the main degeneracies usually encountered in blind-source separation algorithms and to open up such techniques to a plethora of very low S/N observations.

3. Low Signal-to-noise Data

At this conference and in [7] I demonstrate this highly versatile and non-parametric machine learning algorithm using simulations and very low S/N Spitzer/IRS spectra. Thanks to the now absolute calibration of the

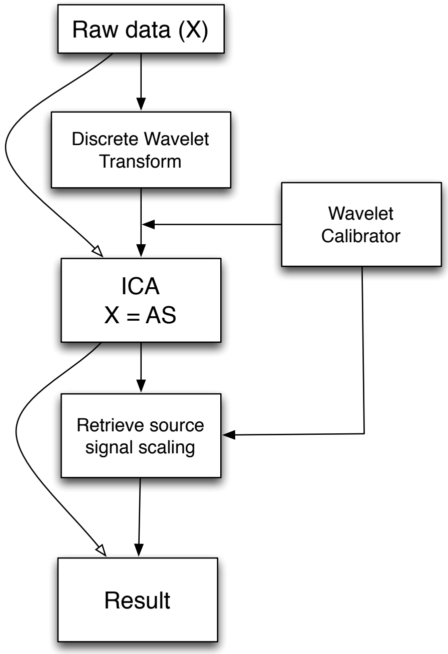

Figure 1: Flowchart of the ACICA algorithm. Beginning with the raw observed data, it is transformed into orthogonal wavelet space to which sparse calibrator impulses are added. The ICA blind-source detrending is then performed in wavelet space and the de-trended signals calibrated with the retrieved calibrator impulses. Curved arrows represent previous algorithms without wavelet support.

retrieved de-correlated signals, it is possible to systematically study the systematic (may these be stellar or instrumental) trends encountered in an observation of an exoplanetary atmosphere.

4. Summary and Conclusions

In this conference, I will discuss and present recent advances in the use of non-supervised machine learning algorithms in the de-trending of exoplanetary spectroscopic data. Given recent controversies in the use of parametric de-trending techniques, it is important to pursue non-parametric avenues that allow for an optimal correction of the data without human biases. `ACICA` allows us to de-trend otherwise in- accessible data sets non-parametrically. We demonstrated this using simulations and archival Spitzer/IRS data. It furthermore offers us an un- precedented and unique insight into the morphol- ogy of a data set by allowing us to directly map out temporal/wavelength dependent variations of instrumental or stellar noise in the data set. Together, these attributes make the algorithm proposed here a highly versatile and powerful tool for exoplanetary time series analysis.

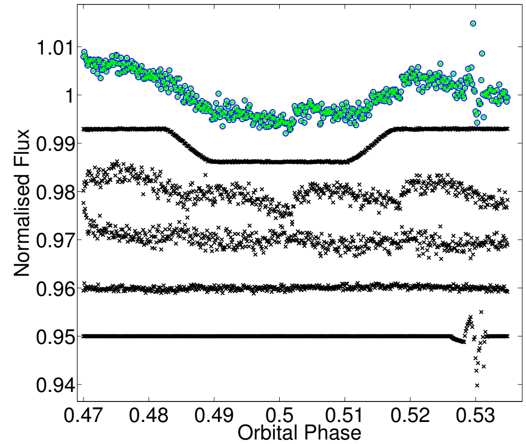

Figure 2: Simulation of a noisy time series observation. From top: simulated observed data, de-trended lightcurve component, systematic instrument noise components, bottom: wavelet calibration signal impulse. Signals are offset for clarity.

References

[1] Carter J.A., Winn J.N., 2009, ApJ, 704, 51

[2] Gibson, N. P., Aigrain, S., Roberts, S., et al., 2012, MNRAS, 419, 2683

[3] Hyvärinen A., Oja, E., 2000, Neural Networks, 13, 411

[4] Thatte, A. and Deroo, P. and Swain, M. R., 2010, A&A, 523, 35

[5] Waldmann, I. P., 2012, ApJ, 747, 12

[6] Waldmann, I P., Tinetti, G., Deroo, P., et al., 2013, ApJ, 766,7

[7] Waldmann, I. P., ApJ, submitted, arXiv: 1302.6714