A new investigation into noise injected by the Kepler Pre-search Data Conditioned pipeline

Simon J. Murphy

smurphy6@uclan.ac.uk
Jeremiah Horrocks Institute, University of Central Lancashire, Preston, PR1 2HE, UK

Introduction

Kepler time-series are readily available in three formats: "Raw Flux" describes those data on which no calibration is performed; "Simple Aperture Photometry" (SAP) describes those data which have undergone standard CCD calibration procedures; and "Pre-search Data Conditioned" (PDC) describes those data that have undergone the first step in preparation for planet searches. Upon Kepler Data Release 9, a new pipeline called PDC-MAP began to supersede the older PDC pipeline, which has now been renamed PDC-LS to alleviate confusion. Data are being reprocessed with the PDC-MAP pipeline with a completion target for long-cadence (LC) data of July 2012. The MAP pipeline does not currently treat the short-cadence (SC) data, for which PDC-LS is still used. This poster is concerned with the PDC-LS pipeline, and shows how it injects noise into the data.

For further reading on the pipelines the papers of Stumpe et al. (2012) and Smith et al. (2012) are recommended, along with the Kepler Data Characteristics Handbook. For how the characteristics affect asteroseismology, and more on PDC-LS, see Murphy (2012).

Prevalence of injected noise

The PDC-LS pipeline injects noise, which can be seen as a distortion of the light curve (Fig. 1). There is an associated sharp drop-off in power at ~24 d^{-1} (Fig. 2), which was seen in 15% of the stars analysed -- those with SC data available and Kepler Input Catalogue temperatures between 5500 and 9500 K.

Fig. 3 shows how for the star KIC 3429637 the PDC-LS pipeline injects noise into the SC Q7 data. The noise is restricted solely to PDC data and affects Q7 but not Q8 for this star. The grass level is higher in PDC Q7 LC than in SAP Q7 LC (after pre-whitening), so the effect is present in LC too, even if the low Nyquist frequency hides the power drop.

Sometimes the PDC fails to fit a light curve and passes the light-curve through without modification. In the lower panel of Fig. 2, it can be seen that although noise is not injected, low-frequency peaks, which are often caused by trends in the data, are not removed either. In this case, the light curve has perhaps emerged untouched by the pipeline.

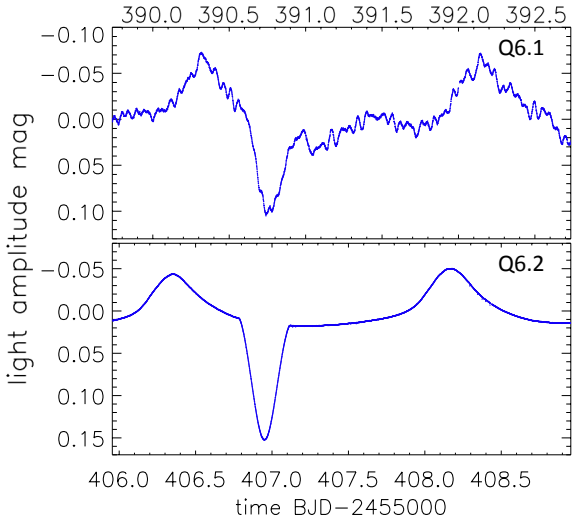

Fig. 1: 3-d segments of the light curves for the detached eclipsing binary system KIC 11285625 in SC PDC flux. In Q6.1 the pipeline is injecting noise, distorting the light-curve, whereas the Q6.2 light-curve looks comparatively very clean. The corresponding Fourier transforms are shown in Fig. 2. It is well-known that PDC-LS does not treat binaries well, but the noise injection is not limited to eclipsing binary systems.

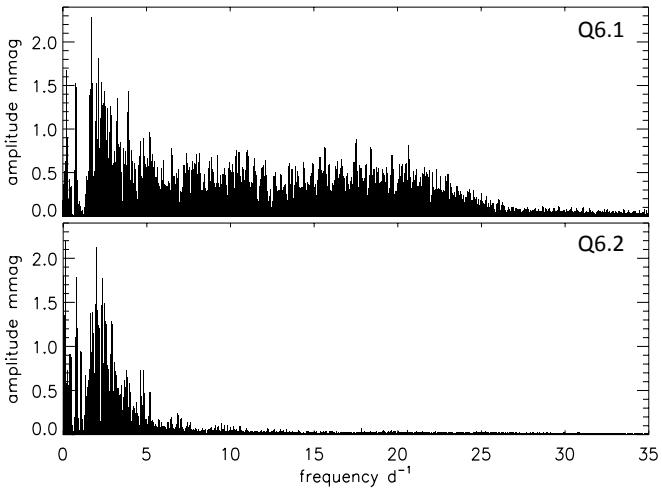

Fig. 2: The prewhitened Fourier transforms for Q6.1 and Q6.2 corresponding to the light curves in Fig. 1, and typical for stars in which the noise injection is seen. Pulsations and harmonics of the have been prewhitened. There is an elevated 'grass-level' up to a frequency of ~24 d^{-1} after which there is a sharp drop-off in power. The drop-off is not seen in LC data (not shown) because of the low Nyquist frequency of 24.4 d^{-1}.

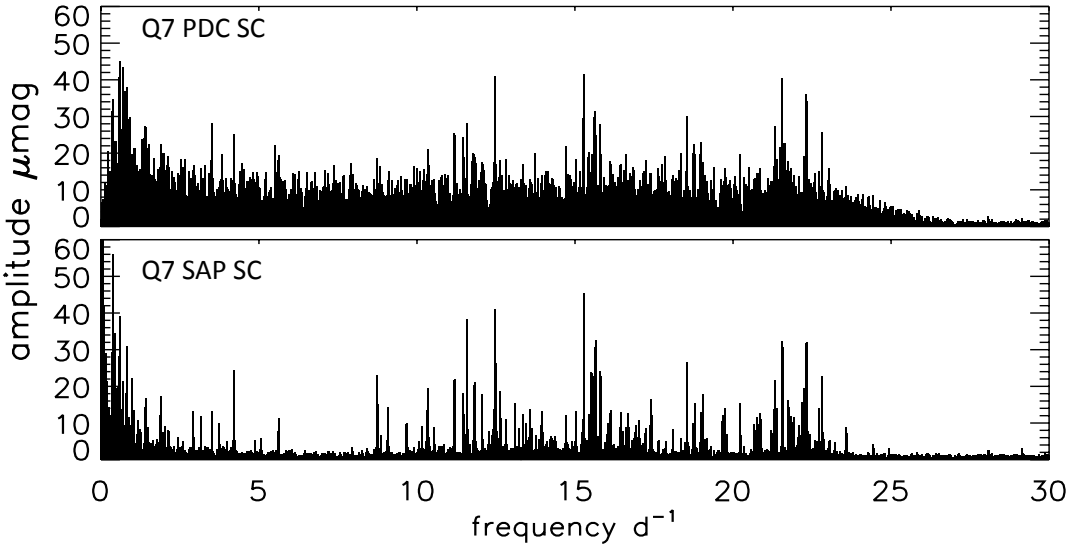

Fig. 3: Fourier transforms for KIC 3429637, showing the extent of the injected noise; the grass level is ~15 μmag in the PDC data, compared to just 2 μmag in the SAP data. The decrease in power at ~24 d^{-1} is clear. The effect is present in Q7 in both SC and LC PDC data for this star, but not present at all in Q8.

QR code for Murphy (2012) paper on Kepler Data Characteristics

References

Murphy S. J., 2012, MNRAS 422, 665

Smith J. C. et al., 2012, arXiv:1203.1383

Stumpe M. C. et al., 2012, arXiv:1203.1382