Exoplanets: Detection, Formation and Dynamics
Proceedings IAU Symposium No. 249, 2007
Y.-S. Sun, S. Ferraz-Mello and J.-L. Zhou, eds.

doi:10.1017/S1743921308016475

Selection effects in Doppler velocity planet searches

Simon O'Toole,[1] Chris Tinney[2] and Hugh Jones[3]

[1] Anglo-Australian Observatory, PO Box 296, Epping NSW 1710, Australia
email: otoole@aao.gov.au

[2] School of Physics, University of NSW, 2006 Australia
email: cgt@phys.unsw.edu.au

[3] Centre for Astrophysical Research, University of Hertfordshire, Hatfield, AL10 9AB, UK
email:hraj@star.herts.ac.uk

Abstract. The majority of extra-solar planets have been discovered by measuring the Doppler velocities of the host star. Like all exoplanet detection methods, the Doppler method is rife with observational biases. Before any robust comparison of mass, orbital period and eccentricity distributions can be made with theory, a detailed understanding of these selection effects is required, something which up to now is lacking. We present here a progress report on our analysis of the selection effects present in Anglo-Australian Planet Search data, including the methodology used and some preliminary results.

Keywords. Extra-solar planets, numerical simulations

1. Motivation

Planet searches - as with all surveys - necessarily suffer from selection effects and observational biases. Without characterising and properly accounting for these effects we cannot compare the catalogue of detected planets with models of planet formation. It has long been known that Doppler velocity planet searches preferentially detect short period planets with large masses - as exemplified by 51 Pegasi b - rather than solar system analogues or short period low mass planets. As surveys progress, longer temporal baselines are allowing us to find planets with longer periods and higher precision is lowering mass detection thresholds. The effects of data sampling and variable data quality have not been quantified, however, meaning a potentially significant selection effect has so far been ignored. Any studies designed to optimise the cadence of observations will inevitably introduce their own biases. Up to now, no detailed Monte-Carlo-like simulations designed to quantify these effects have been carried out.

Despite the lack of selection function analyses, statistical analyses of the sample are now routinely undertaken. Several of these have simply assumed that observed distribution of exoplanets is in fact the true distribution, while making overly simplified assumptions about the observational biases inherent in the sample (e.g. Lineweaver & Grether 2003). Others have either ignored the effects of eccentricity or used an inadequate treatment of it (e.g. Cumming 2004). No one has simulated observational data looking for the effects of data quality and sampling on a star-by-star basis. We have begun a project to investigate the latter effects, while examining the entire parameter space including period, eccentricity, and planet mass.

2. Methodology

To derive selection functions for our observations, we first require an automated toolbox. The Lomb-Scargle (LS) periodogram has been increasingly used in Doppler velocity planet searches where, however, circular orbits (giving rise to sinusoidal velocity curves) are apparently rare (Butler *et al.* 2006). It therefore makes more sense to fit Keplerians to data instead of sinusoids as discussed by Cumming (2004). As orbital eccentricity is an important parameter in these functions, we have expanded the traditional LS periodogram to two dimensions: period and eccentricity; we call this the 2D Keplerian LS (2DKLS) periodogram. The method we use to calculate the 2DKLS periodogram was introduced in O'Toole *et al.* (2007). Briefly, we use a grid of fixed periods and eccentricities to calculate the 2DKLS, with $e = 0 - 0.95$ in steps of 0.05, while the list of periods is on a logarithmic scale, with log P=–0.3 to 3.7. A Keplerian is then fitted to the data using a non-linear least squares fitting routine with Levenberg-Marquadt minimisation from Press *et al.* (1986).

3. Detection Criteria

As discussed above, automation is one of the important practical criteria of our simulations. This extends to the development of an adequate set of criteria to decide whether a planet has been detected. Given the number of simulations to be analysed, one cannot simply examine each power spectrum or velocity curve by eye as is often done (initially at least) with real observations.

There are several methods currently used to determine the reality of a planet detection. Marcy *et al.* (2005) present an excellent discussion of two different approaches based on the False Alarm Probability or FAP.

Because of the large number of simulations we plan to carry out during this project, it is necessary we have a simple set of criteria that can quickly test the reliability of a detection. This automatically rules out several approaches that are in themselves computationally intensive; in particular the determination of a FAP would add considerably to the time budget of our simulation analysis. The criteria must be able to detect as many planets as possible robustly, without introducing too many false positives. An individual criterion can be statistical or physical. There is a certain level of arbitrariness involved in selecting detection criteria, as many different combinations will lead to similar results. Below we list the criteria we have used in this work.

- Fit period must be less than twice the time-span of the data - periods beyond this are not constrained;
- RMS of data must be greater than twice median uncertainty - based on an often-used flag for real data;
- Fit period must be greater than twice fit error - the period must be reasonably constrained; and
- χ^2 must be within 1σ of median of all χ^2 values.

They are by no means the only criteria we could use; however, they allow a swift determination of planet detection and give false positives and false negatives of less than 5%. We will refine them when necessary during the project.

4. Preliminary Results

Selection functions (SF) for eccentricity and period have been derived for each data set using the detection criteria described above. The functions we show here are integrated

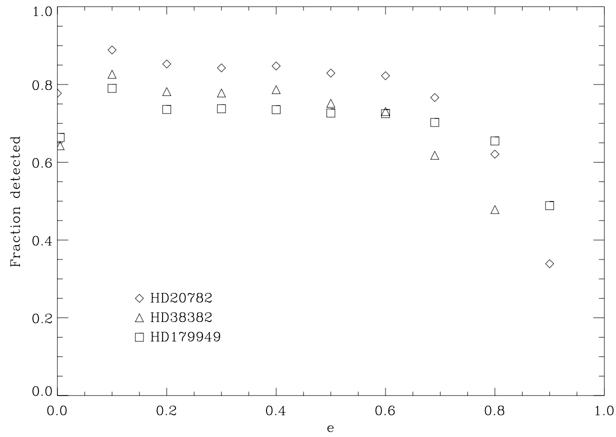

Figure 1. Eccentricity selection functions as a function of eccentricity for HD 20782, HD 38382 and HD 179949.

over semi-amplitude and period (eccentricity SF) and semi-amplitude and eccentricity (period SF). Currently the ordinate is simply the fraction of trials with a detection; this will be converted into a detection probability.

Figure 1 shows the eccentricity SFs for three of the objects we have simulated. There is a clear difference between each star, especially at high eccentricities. This difference does not appear to be dependent on the number of data points, as the HD 38382 simulations with 17 observations have only around 25-30% the number of observations as the HD 179949 simulations (56 observations). The sensitivity of our 35 observations of HD 20782 is greatest up to $e \sim 0.7$, but then drops off to below HD 179949. We suggest that data quality plays a significant role in determining the shape of the eccentricity SF. Below around $e = 0.5$ the sensitivity is approximately constant down to $e \sim 0.1$, where there appears to be a jump and then a drop at $e = 0.0$. The cause of this may be simply an artifact, and we are currently testing this hypothesis. The effects described above suggest that there is no appropriate parameterisation that can model the eccentricity SF, and that simulations such as these *must* be carried out on a star-by-star basis.

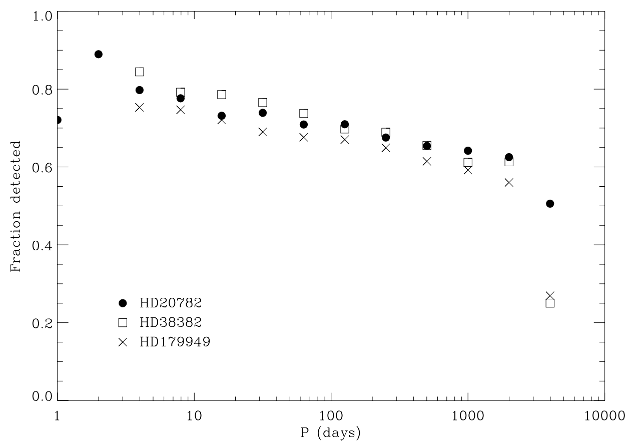

Figure 2. Period selection functions as a function of period for the three objects simulated.

https://doi.org/10.1017/S1743921308016475 Published online by Cambridge University Press

The period SFs shown in the bottom left panel of Figure 2 are what we might have expected: a steady decrease in sensitivity as the input period increases. Two interesting points arise though. Firstly, the sensitivity at short periods (only shown for HD38382 in the figure) is high, even for periods around 2 days; it only drops off at very short periods, where data sampling becomes an issue. Secondly, the drop-off in sensitivity at very long periods is simply caused by the input period being longer than the length of the data string.

5. Conclusions and Future Work

One of the key conclusions of the preliminary selection function analysis presented in this poster is that carrying out the simulations to quantify selection effects on a star-by-star basis is of tremendous importance. We also find that Doppler velocity planet search observations are biased against finding planets in highly eccentric orbits. This does not mean that there are necessarily more of these planets to be found, just that we are not very sensitive to them. Finally, we find that our observations should be able to detect planets with orbital periods around 2 days, but we become less sensitive at shorter periods.

The ongoing and future work on this project includes:

• Continue simulations using simple noise model using Keter at UCL and Swinburne University supercomputer;

• Determine mass selection functions for stars already investigated; and

• Develop a more sophisticated noise model, incorporating stellar "jitter": magnetic activity $>\sim 2$ m/s (Wright 2005); solar-like oscillations ~ 0.5 m/s (O'Toole *et al.* 2008); stellar convection and granulation noise, expected to be $\sim$ 1-2 m/s, and possibly activity cycles.

References

Butler, R. P., *et al.* 2006, *ApJ*, 646, 505
Cumming, A. 2004, *MNRAS*, 354, 1165
Jones, H. R. A., *et al.* 2006, *MNRAS*, 369, 249
Lineweaver, C. & Grether, D. 2003, *ApJ*, 598, 1350
Marcy, G., *et al.* 2005, *ApJ*, 619, 570
O'Toole, S. J., *et al.* 2007, *ApJ*, 660, 1636
O'Toole, S. J., *et al.* 2008, *MNRAS*, submitted
Press, W., *et al.* 1986, *Numerical Recipes*, Cambridge University Press
Santos, N. C.. *et al.* 2001, *A&A*, 379, 999
Tinney, C. G., *et al.* 2001, *ApJ*, 551, 507
Valenti, J. & Fischer, D. 2005, *ApJS*, 159, 141
Wright, J. 2005, *PASP*, 117, 657

https://doi.org/10.1017/S1743921308016475 Published online by Cambridge University Press