Mem. S.A.It. Suppl. Vol. 13, 46

Memorie della
Supplementi

Testing planetary transit detection methods with grid-based Monte-Carlo simulations

A. S. Bonomo[1,2] and A. F. Lanza[2]

[1] Dipartimento di Fisica e Astronomia, Università degli Studi di Catania, Via S. Sofia, 78, 95123, Catania, Italy
[2] INAF - Osservatorio Astrofisico di Catania, Via S. Sofia, 78, 95123, Catania, Italy
e-mail: aldo.bonomo@oact.inaf.it

Abstract. The detection of extrasolar planets by means of the transit method is a rapidly growing field of modern astrophysics. The periodic light dips produced by the passage of a planet in front of its parent star can be used to reveal the presence of the planet itself, to measure its orbital period and relative radius, as well as to perform studies on the outer layers of the planet by analysing the light of the star passing through the planet's atmosphere. We have developed a new method to detect transits of Earth-sized planets in front of solar-like stars that allows us to reduce the impact of stellar microvariability on transit detection. A large Monte Carlo numerical experiment has been designed to test the performance of our approach in comparison with other transit detection methods for stars of different magnitudes and planets of different radius and orbital period, as will be observed by the space experiments CoRoT and Kepler. The large computational load of this experiment has been managed by means of the Grid infrastructure of the COMETA consortium.

Key words. planetary systems, methods: data analysis, techniques: photometric, stars: activity, stars: late-type

1. Introduction

The search for extrasolar planets is one of the most interesting and active fields of modern astronomy with more than 300 planets discovered to date. They are mostly Jupiter-mass objects, 25 per cent of which is orbiting at distances closer than 0.05 AU from their stars. The next frontier is the search for Earth-sized objects which is possible thanks to space-borne telescopes such as those of the missions CoRoT (launched on 27 December 2006) and Kepler (scheduled in 2009). They will detect planets in an indirect way, by looking for the light dips produced by the transits of the planets themselves in front of their parent stars (see Fig. 1). The probability that a transit occurs is $R_\star/a$, where $R_\star$ is the radius of the star and a the semi-major axis of the orbit assumed to be circular. Since $R_\star/a$ < 0.003-0.05 and the transit duration is of the order of a few hours or tens of hours for a solar-like star accompanied by a planet with an orbital period between 10 and 360 days, a large number of stars must be monitored for transits. CoRoT will observe about 60,000 stars, while Kepler aims at 100,000, for time intervals ranging from 150 days to 4 years. The light dips produced by the

Send offprint requests to: A. S. Bonomo

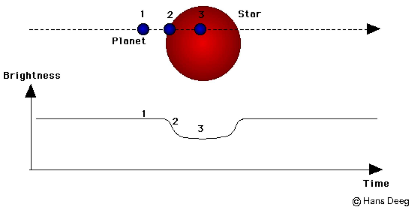

Fig. 1. A planet transiting in front of the disk of its parent star and the corresponding variation of the optical flux of the star which allows us to detect the transit (courtesy by Hans Deeg).

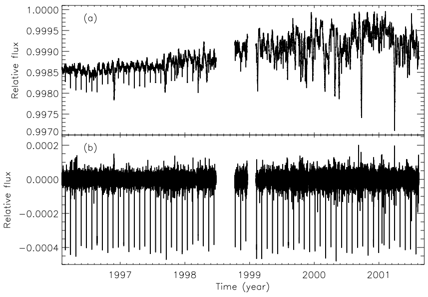

Fig. 2. Upper panel (a): the time series of the Total Solar Irradiance with superposed a planetary transit of an Earth-like planet of radius 2.2 times the radius of the Earth and an orbital period of 30 days. Lower panel (b): the residuals obtained after subtracting the model for stellar microvariability by Lanza et al. (2003).

transits of an Earth-sized planet in front of a solar-like star are smaller than 100 parts per million (ppm), so that intrinsic stellar variability must be accounted for at 30-50 ppm level in order to be able to detect its transits. Since the transit of a large sunspot group across the disk of the Sun produces a dip of about 3000 ppm, it is necessary to develop techniques to model the effects of stellar magnetic activity and use them to remove stellar microvariability before performing any successful search for transits by Earth-like planets.

2. Detecting transits in the presence of stellar variability

Several techniques have been recently proposed to reduce the impact of stellar variability due to magnetic activity (i.e., cool spots, bright faculae and network) on the detection of planetary transits across solar-like stars, see, e.g., Defaÿ et al. (2001), Jenkins (2002), Carpano et al. (2003), Lanza et al. (2003), Aigrain & Irwin (2004). They are based on a variety of methods that make use of filtering in Fourier space, wavelet decomposition or physical models of the surface brightness inhomogeneities in order to account for the effects of magnetic activity on stellar light curves. In Fig. 2, we show the application of the fitting method developed by Lanza et al. (2003) to the transits by a planet of 2.2 Earth radii across the disk of a Sun-like star, the variability of which has been assumed to be identical to the Total Solar Irradiance time series as observed by the experiment VIRGO on board of the SoHO satellite (Fröhlich & Lean 2004). The transit is not detectable by eye during phases of intermediate and maximum activity owing to the larger intrinsic stellar variability. It becomes immediately visible after subtracting the best fit to the stellar variability obtained with the model by Lanza et al. (2003). Different methods have different performance according to the orbital period of the planet, the ratio between the transit depth and the standard deviation of the photon shot noise and the activity level of the star.

3. Filtering methods

We have recently designed a large Monte Carlo numerical experiment to compare the performance of the method proposed by Lanza et al. (2003) with the two filtering techniques that, coupled with box-shaped transit finder algorithms, showed the best performance during the First CoRoT blind test carried out by Moutou et al. (2005) to test different strategies for the analysis of the CoRoT light curves. A brief description of the filtering methods whose performance we want to compare is as follows:

a) 3-spot model (Lanza et al. 2003, 2007): it is a simplified physical model of solar-like variability based on the rotational modulation of the flux produced by three active regions, containing both cool spots and warm faculae, plus a constant component to account for uniformly distributed active regions. In the case of the Sun, the model accounts for the flux variability up to a time scale of 14 days, after which the position and areas of the three regions and the uniform component have to be changed;

b) 200-harmonic fitting (Moutou et al. 2005, team 3): it fits stellar variability by means of a linear combination of 200 harmonic functions whose frequencies are multiples of the fundamental frequency $f_L = 1/2T$, where T is the whole duration of the time series, i. e. $T \sim 150$ days in the case of the CoRoT mission;

c) Iterative non-linear (INL) filter (Aigrain & Irwin 2004; Moutou et al. 2005, team 5): it is based on the computation of a continuum by applying a sliding median-boxcar filter. Points where the difference between the continuum and the original light curve is greater than 3 standard deviations are flagged and the continuum is recomputed without the flagged points, iterating the process up to convergence. The final continuum is then subtracted from the original light curve.

4. Light curve simulation and analysis

We apply a Monte Carlo approach by simulating a large number of light curves of duration 150 days (the extension of the CoRoT long runs) for different values of planetary radius R_p ranging from 1.0 to 2.0 Earth radii, orbital period P between 5 and 50 days, and standard deviation of the photon shot noise σ from 100 to 1000 parts per millions (ppm). A noise level $\sigma = 100$ ppm is obtained for a star of $V \sim 12$ observed in white light by CoRoT with 1 hr integration time, while $\sigma = 200$, 300 and 1000 ppm corrispond to stars of $V \sim 13$, 14 and 16, respectively, observed with the same instrument and 1 hr integration time. The phase of the first transit is taken from a uniform random distribution. The star is assumed to have the solar radius and mass. We add stellar variability, assumed in all the cases to be given by the Total Solar Irradiance variations as observed close to the maximum of solar cycle 23 (e.g., Fröhlich & Lean 2004). For each set of planetary parameters and noise level, we simulate 100 light curves with different noise and activity realizations, for a total of 8000 light curves.

After filtering solar variability with the three different filtering methods, transits are searched by means of the BLS algorithm. The ratio of the transit depth to the noise level is indicated by α, whose statistics determine the confidence level of a given transit detection (see Kovács et al. 2002).

Transitless light curves are analysed in the same way to establish the transit detection threshold for each filtering method, by requiring a maximum false-alarm rate of 1 percent.

The large computational load of our experiment is managed by running our analyses on the grid infrastructure of the project PI2S2, allowing us to use up to 560 CPUs in parallel (e.g., Becciani 2007). The CPU time for filtering stellar microvariability with the three methods is on the average about 15 minutes, while about 2.5 days of elapsed time have been necessary to analyse the complete set of 8,000 light curves.

5. Results and conclusions

We show in Fig. 3 a sample of our results for the case of a planet of 1.5 Earth radii with orbital periods of 5, 10, 25 and 50 days, as labelled, respectively. The noise value (standard deviation) corresponds to a star of magnitude 13 as will be observed by CoRoT with an integration time of one hour. The plots show the frequency distribution of the detection parameter α which measures the depth of a putative transit in a light curve with respect to the standard deviation of the noise (see Kovács et al. 2002 for details). In the present case, 400 light curves with transits have been analysed for each value of the parameters. The vertically dot-dashed lines indicate the threshold for α, i.e., the minimum level above which a transit is detected in the light curve with a false-alarm

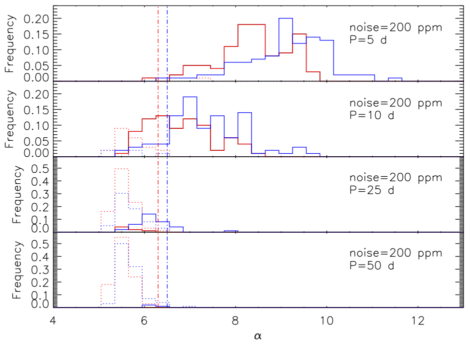

Fig. 3. Distributions of the values of α, the signal-to-noise ratio of a transit detection, obtained by analysing light curves with transits of a planet of 1.5 $R_{\oplus}$. Each set is characterized by a different orbital period of the planet (as labelled). The standard deviation of the white Gaussian noise is in all cases 200 ppm. The red and blue vertical dot-dashed lines indicate the 1 percent false-alarm threshold for the 200-harmonic fitting and the 3-spot model, respectively. Red solid histograms show the statistics of light curves where the period P was correctly identified after applying the 200-harmonic method, blue solid histograms show those after applying the 3-spot method. Dashed histograms refer to the statistics of light curves where the period P was incorrectly identified, with the same color coding.

probability lower than 0.01. The blue-solid histograms show the fraction of detections obtained with the 3-spot model, while the red-solid histograms show the detections with the 200-harmonic fitting. Dotted histograms show the frequency distribution of missed detections, i.e. those where the orbital period of the planet was incorrectly identified, with the same color coding. Missed detections increase with increasing orbital period, because the number of transits in the light curve becomes smaller giving a lower signal-to-noise ratio, and are concentrated below the threshold. Note that the method by Lanza et al. (2003) performs significantly better than the 200-harmonic fitting, giving systematically higher values of α. Such a conclusion is confirmed by the analysis of the complete set of our simulated light curves. Specifically, we find that the 3-spot model is always better than the 200-harmonic fitting when the standard deviation of the photon shot noise is at least 2-4 times larger than the depth of the transit, that is the case for stars with magnitude fainter than 13.0 in the fields of CoRoT (see Bonomo & Lanza 2008 for details).

Bonomo & Lanza (2008) show that the poor performance of the 200-harmonic fitting is due to the use of orthogonal functions to fit stellar variability, which makes it significantly affected by the Gibbs phenomenon (Morse & Feshbach 1954). This latter reduces the depth of the transits in the filtered light curves thus lowering the efficiency of detection in the presence of noise (see Fig. 4).

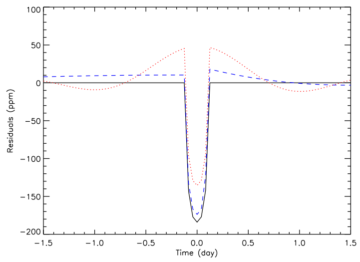

Fig. 4. The shape of a transit of an Earth-like planet as it appears in the ideal case (solid black line), in the residuals obtained with the 3-spot model (dashed blue line) and in those of the 200-harmonic fitting (dotted red line). Note the reduction of the transit depth and the overshooting at the edges of the transit dip due to the Gibbs phenomenon in the case of the 200-harmonic fitting.

Concerning the INL filter, it shows a performance comparable with that of the 3-spot model in most of the cases, even better in some instances, when we use an appropriate window of 2 days for the median boxcar filter (see Fig. 5).

However, the performance of the INL filter depends critically on the adopted extension of the filter window. An optimal window of 2 days has been chosen for our analysis. Shorter windows negatively affect the transit detection since they give rise to a reduction of the transit depth in the filtered light curve (see Fig. 6), in which case the 3-spot model would prove to be the best method for the cases with $\sigma \geq 200$ ppm.

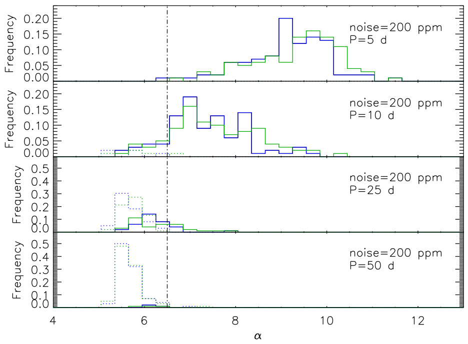

Fig. 5. As in Fig. 3, distributions of the values of α obtained analysing the simulated light curves after applying the 3-spot model (blue histograms) and the INL filter (green histograms).

The optimal width of the median filter window depends on the magnetic activity level of the star and its rotation period. Specifically, the higher the activity level and the shorter the rotation period, the shorter the optimal window, because the time scales of the flux variations decrease with increasing activity. In other words, in the case of highly active stars, the window extension has to be shortened with respect to the solar case, otherwise some oscillations or transit-like features will appear in the residuals owing to a bad filtering of the variability. To fix automatically the window extension, we propose a method similar to that of Regulo et al. (2007), computing the power spectrum of the time series and choosing an extension corresponding to the frequency where the power spectral density goes below a fixed threshold, usually set at 10^{-6} of the maximum power level.

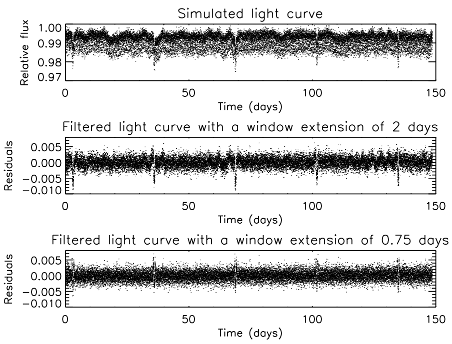

Fig. 6. *Upper panel*: one of the light curves with transits simulated for the First CoRoT Blind test (Moutou et al. 2005; ID=460). *Middle panel*: the light curve filtered by means of the INL filter with a window of 2 days. *Bottom panel*: the filtered light curve with a 0.75 day window. Note the disappearance of the transits when the window extension is reduced.

We conclude that the INL filter, when applied with a suitable choice of its window, has a better performance than more complicated and computationally intensive methods of fitting solar-like variability, like the 200-harmonic fitting or the 3-spot model.

Acknowledgements. The authors gratefully acknowledge Drs. U. Becciani, A. Costa and A. Grillo and the system managers of the Trigrid and COMETA Consortia for their technical advice and kind assistance during the implementation and running of our numerical experiments on the PI2S2 grid-based high performance computing system. The authors gratefully acknowledge support from the Italian Space Agency (ASI) under contract ASI/INAF I/015/07/0, work package 3170. This research has made use of results produced by the PI2S2 Project managed by the Consorzio COMETA, a project co-funded by the Italian Ministry of University and Research (MUR) within the *Piano Operativo Nazionale "Ricerca Scientifica, Sviluppo Tecnologico, Alta Formazione" (PON 2000-2006)*. More information is available at http://www.pi2s2.it and http://www.consorzio-cometa.it. Active star research and exoplanetary studies at INAF-Catania Astrophysical Observatory and the Department of Physics and Astronomy of Catania University is funded by MUR (*Ministero dell'Università e Ricerca*), and by *Regione Siciliana*, whose financial support is gratefully acknowledged. This research has made use of the ADS-CDS databases, operated at the CDS, Strasbourg, France.

References

Aigrain, S., & Irwin, M. 2004, MNRAS, 350, 331

Becciani, U. 2007, in Grid-enabled Astrophysics, L. Benacchio, F. Pasian, Eds., Edizioni Polimetrica; p. 179

Bonomo, A. S., & Lanza, A. F. 2008, A&A, 482, 341

Carpano, S., Aigrain, S., & Favata, F. 2003, A&A, 401, 743

Defaÿ C., Deleuil, M., & Barge, P. 2001, A&A, 365, 330

Fröhlich, C., & Lean, J. 2004, A&ARv, 12, 273

Jenkins, J. M. 2002, ApJ, 575, 493

Kovács, G., Zucker, S., & Mazeh, T. 2002, A&A, 391, 369

Lanza, A. F., Rodonò, M., Pagano, I., Barge, P., & Llebaria, A. 2003, A&A, 403, 1135

Lanza, A. F., Bonomo, A. S., & Rodonò, M. 2007, A&A, 464, 741

Morse, P. M., & Feshbach, H. 1954, Methods of Theoretical Physics, McGraw-Hill Book Co. Inc., New York

Moutou, C., Pont, F., Barge, P., Aigran, S., Auvergne, M., et al. 2005, A&A, 437, 355

Regulo, C., Almenara, J. M., Alonso, R., Deeg H., & Roca Cortes T. 2007, A&A, 467, 1345