The CoRoT space mission: early results

Noise properties of the CoRoT data

A planet-finding perspective[⋆]

S. Aigrain[1], F. Pont[1], F. Fressin[2], A. Alapini[1], R. Alonso[3], M. Auvergne[4], M. Barbieri[3], P. Barge[3], P. Bordé[5],
F. Bouchy[6], H. Deeg[7], R. De la Reza[8], M. Deleuil[3], R. Dvorak[9], A. Erikson[10], M. Fridlund[11], P. Gondoin[11],
P. Guterman[3], L. Jorda[3], H. Lammer[12], A. Léger[5], A. Llebaria[3], P. Magain[13], T. Mazeh[14], C. Moutou[3], M. Ollivier[5],
M. Pätzold[15], D. Queloz[16], H. Rauer[10,17], D. Rouan[4], J. Schneider[18], G. Wuchter[19], and S. Zucker[20]

(Affiliations can be found after the references)

ABSTRACT

In this short paper, we study the photometric precision of stellar light curves obtained by the CoRoT satellite in its planet-finding channel, with a particular emphasis on the time scales characteristic of planetary transits. Together with other articles in the same issue of this journal, it forms an attempt to provide the building blocks for a statistical interpretation of the CoRoT planet and eclipsing binary catch to date. After pre-processing the light curves so as to minimise long-term variations and outliers, we measure the scatter of the light curves in the first three CoRoT runs lasting more than 1 month, using an iterative non-linear filter to isolate signal on the time scales of interest. The behaviour of the noise on 2 h time scales is described well by a power-law with index 0.25 in R-magnitude, ranging from 0.1 mmag at $R = 11.5$ to 1 mmag at $R = 16$, which is close to the pre-launch specification, though still a factor 2−3 above the photon noise due to residual jitter noise and hot pixel events. There is evidence of slight degradation in the performance over time. We find clear evidence of enhanced variability on hour time scales (at the level of 0.5 mmag) in stars identified as likely giants from their R magnitude and $B − V$ colour, which represent approximately 60 and 20% of the observed population in the directions of Aquila and Monoceros, respectively. On the other hand, median correlated noise levels over 2 h for dwarf stars are extremely low, reaching 0.05 mmag at the bright end.

Key words. methods: data analysis – techniques: photometric – stars: planetary systems

1. Introduction

This paper presents a global assessment of the noise properties of the CoRoT planet-finding channel data to date, from a transit detection perspective, based on data collected over the first 14 months of the satellite's operations.

A preliminary discussion of CoRoT's photometric performance from an instrumental point of view was given by Auvergne et al. (2009). In this work, we start from the science-grade data as released to the CoRoT community and, one year later, to the public. We focus specifically on the "white" (~300−1000 nm) light curves collected in the planet-finding channel of CoRoT.

We outline the pre-processing steps taken to minimise instrumental and astrophysical signals which can impede transit detection, and discuss the noise on transit time scales after pre-processing. The motivation for doing this is two-fold. First, it enables us to check whether the CoRoT planet-finding data meets its pre-launch specifications and provides the first estimate of correlated noise levels in space-based time series based on a large dataset. Second, it provides us with a means of estimating the detection threshold for planetary transits as a function of stellar magnitude.

This is the first in a series of three papers: based on the photometric performance described here and on the results of the transit detection and ground-based follow-up programme to date: Moutou et al. (2009) presents an assessment of the overall CoRoT detection threshold based on candidates from the initial run, which in turn serves as the input for Fressin et al. (in prep.)'s attempt at a preliminary statistical interpretation of the CoRoT planet catch to date.

2. Light curve pre-processing

The CoRoT N2 (science-grade) data contain a number of artifacts and signals of astrophysical origin other than transits, which must be removed or reduced before the transit search can proceed. In this section, we outline the steps taken to do this, since it is the properties of the pre-processed light curves that determine the detectability of transits. In practice, of course, the CoRoT data are being analysed by different teams employing a variety of methods, which will no doubt be improved upon in the future. However, the output of the method presented here is, to the best of our knowledge, representative of the quality of the light curves used to detect transits in CoRoT data to date.

To illustrate the effects to be removed and the method employed to do so, three typical CoRoT N2 light curves for bright ($V < 12.5$) stars are shown in Fig. 1. All three are from the
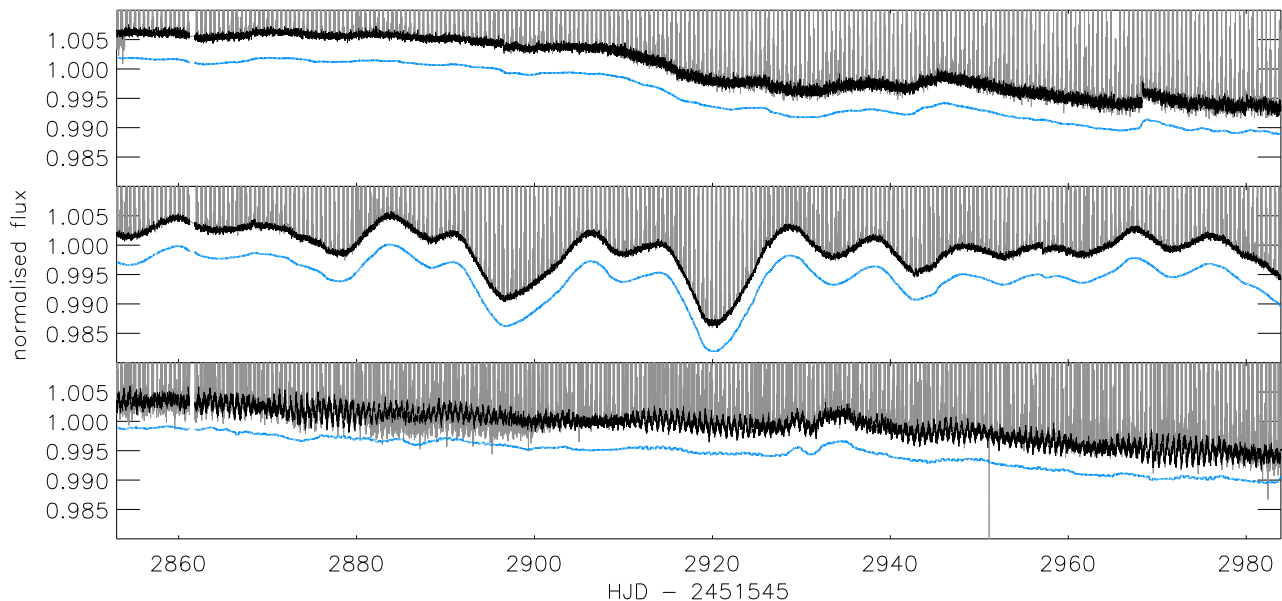

Fig. 1. Example light curves from run LRa01 before and after clipping (grey and black, respectively). The variations on time scales longer than a day, which are removed before searching for transits, are shown in blue, offset by 0.005 for clarity. See text for further details.

second CoRoT long run (150 d), LRa01, during which the satellite was pointed in the direction of the Monoceros constellation.

Most pro-eminent are upward outliers, due to the increased background during the satellite's quasi-periodic crossings of the South Atlantic Anomaly (SAA), and a long term decay assumed to have an instrumental origin. Although data points collected during the SAA crossings are flagged by the N2 pipeline, not all of these data points are outliers. To maximise the duty cycle of the final light curves, we apply a short baseline iterative non-linear filter with 3-σ outlier rejection to identify and reject outliers. The non-linear filter consists of a 5-point boxcar filter followed by a median filter whose baseline corresponds to one hour (see Aigrain & Irwin 2004, for further details of this type of filter).

There are also downward outliers, though at a much lower level (these are most clearly visible in the third example in Fig. 1 around date 2885). These outliers occur when the satellite enters or exits the Earth's shadow. This causes a temporary loss of pointing accuracy, which induces a drop in the flux measured in each star's aperture. Each entry or exit event affects a single data point, which is flagged by the N2 pipeline. We use these flags to clip out the affected data.

In light curves that were binned on board to a sampling of 512 s[1], the pipeline did not flag all the downward outliers. For short (21 d) runs, usually taken around equinoxes when Earth eclipses entry and exit events are most pro-eminent, we folded the light curves at the satellite orbital period and automatically identified and clipped the remaining data points affected. We did not do this for the longer runs, as the magnitude of these events and their timing within the satellite's orbit evolve over time scales of months. Therefore, some minor effects of the Earth eclipses remain in those light curves.

All light curves share a long term downward trend, presumably due to some kind of instrumental decay or gradual pointing drift. In addition, most light curves show variations on time scales of days to weeks, which are generally interpreted as resulting from rotational modulation and evolution of active regions on the surface of the star. We estimate the long-term component of the light curve variations using a 1-day iterative non-linear filter (where the median filter has a 1-day baseline and the other parameters are as before, see Fig. 1). This estimate is subtracted from the clipped light curve, and the resulting de-trended light curve is used to perform the transit search.

Most light curves also contain one or more sudden jumps, which are often followed by an exponential decay, though in some cases the decay is delayed and sudden rather than exponential. These so-called "hot pixel events" are caused by high-energy particles impacts on the detector, which temporarily alter the sensitivity of one or more pixels (see Pinheiro da Silva et al. 2008). Hot pixels leave local high-frequency residuals in the detrended light curves, which can lead to false alarms in the transit search, though they do not obviously affect the global detrended light curve scatter.

In some cases, shorter time scale variations are also visible. These can be periodic or quasi-periodic, as expected for pulsations or residual satellite pointing jitter, or stochastic, as expected for e.g. surface granulation. This type of variability, occurring on typical transit time scales, cannot be filtered out without affecting a potential transit signal.

3. Noise estimates

As shown by Pont et al. (2006), what determines the detectability of transits is the noise level over a typical transit duration, which is generally not equal to what one might infer from the point-to-point scatter of the light curve, due to the presence of correlated noise on hour time scales. We therefore evaluated the scatter of each light curve after detrending over both 512 s (point-to-point) and 2 h. (CoRoT is sensitive to transits with periods up to a few weeks, which last approximately 2 h.) The latter estimate is

---

[1] On-board exposure times were 32 s, but most data were rebinned on board to 512 s by averaging sets of 16 exposures. Where this was not the case, we performed the rebinning ourselves so as to work with a data set with approximately uniform sampling, though we used medians rather than averages to minimise sensitivity to outliers.
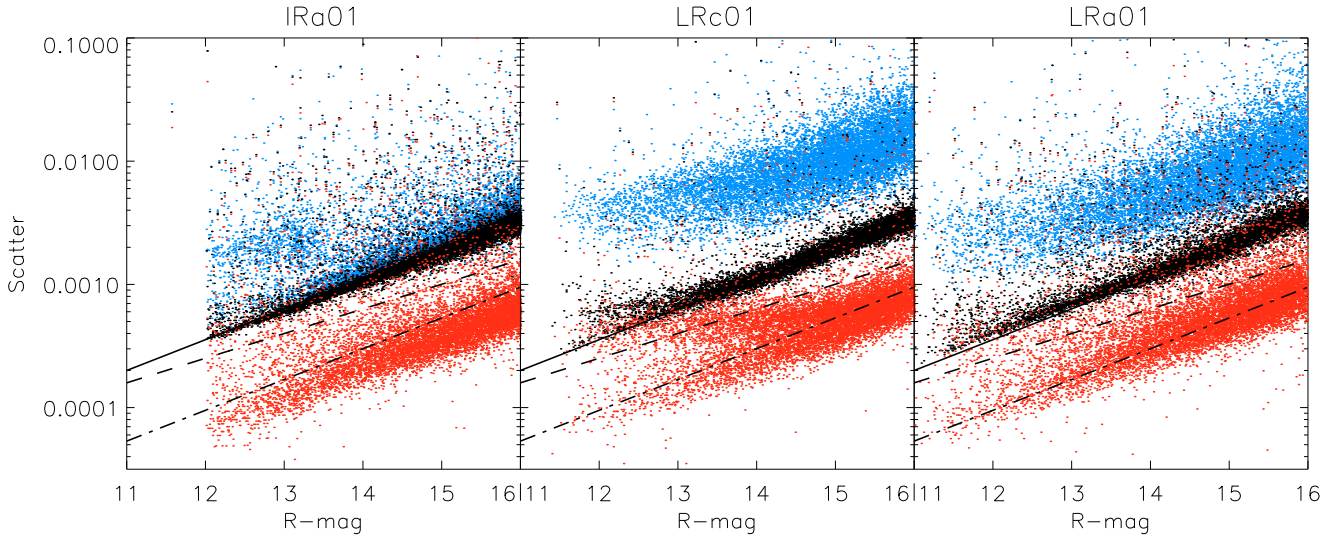
Fig. 2. Noise level as a function of *R*-magnitude for IRa01, LRc01 and LRa01. The black points represent the point-to-point scatter for each light curve. The solid line marks the lower envelope of these points, empirically approximated as a power law with slope 0.25 and value 0.2 mmag at *R* = 11. This can be directly compared with the theoretical (source) photon noise level, shown as the dashed line. The red points show the correlated noise level on 2 h time scales. This can be directly compared to the uncorrelated (white) noise on the same time scales, shown as the dash-dotted line (which is obtained by scaling the solid line down by a factor $\sqrt{7200/512} \approx \sqrt{14}$). Also shown in blue is the light curve scatter on time scales of a day and longer, including the effect of stellar variability, instrumental decay/drift and hot pixels.
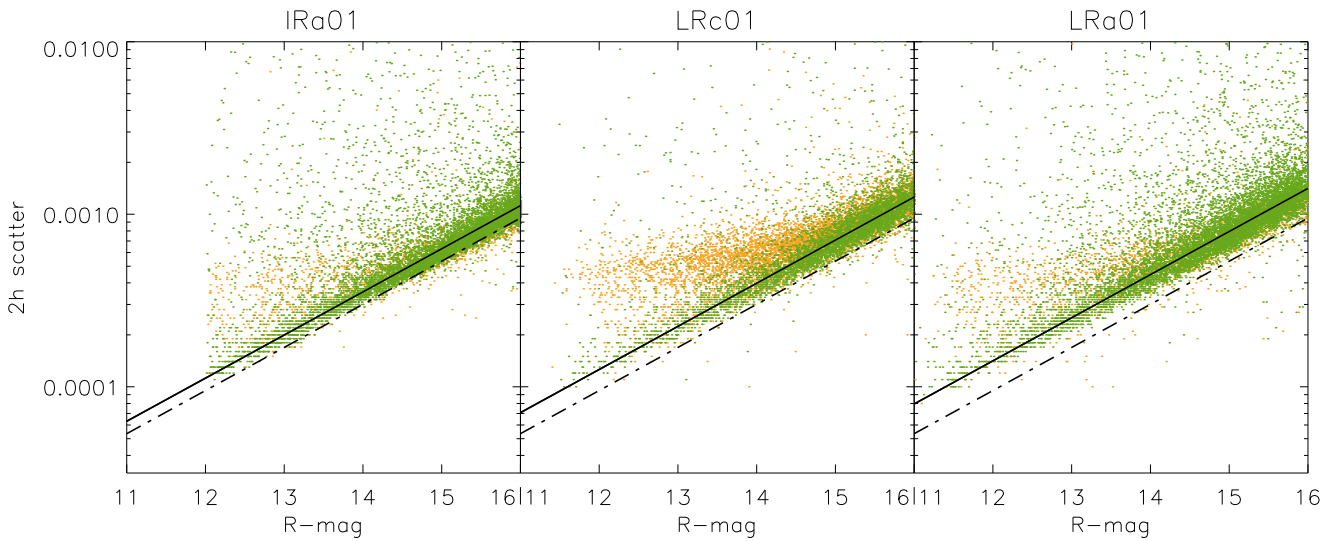

Fig. 3. Median noise over 2 h (including both correlated and uncorrelated components) as a function of *R* magnitude for IRa01, LRc01 and LRa01. Green dots correspond to stars identified as likely dwarfs from the *R*, *B* − *V* colour−magnitude diagram (see Fig. 4), orange dots to stars identified as likely giants. The solid line represents the empirical formula of Eq. (1), and the uncorrelated (white) component of the noise on 2 h time scales is shown for comparison as the dash-dotted line.

obtained by applying a 2 h baseline iterative non-linear filter to the detrended light curve. In both cases, we evaluated the scatter as $\sigma = 1.48 \times$ MAD, where MAD is using the median of absolute deviations from the median, an estimate of the standard deviation that is robust to outliers (Hoaglin et al. 1983).

If the noise were entirely white, the ratio between the two estimates (point-to-point and over 2 h) should correspond to the square root of the number of data points in a 2 h interval (in this case 14). The quadrature difference between the measured noise level on 2 h time scales and the value expected if the noise were white provides an estimate of the component of the noise, which is correlated on 2 h time scales. For comparison, we also evaluate

the scatter on long time scales from the light curve component that is subtracted in the detrending process.

This was done for the CoRoT initial run IRa01 (60d in the Monoceros direction) and the first two long runs (LRc01 in the Aquila direction, and LRa01). The results are shown in Figs. 2 and 3. Each data point in both figures is a MAD-estimated scatter measurement for one star over the entire run. Figure 2 shows the noise per 512 s exposure, compared to the noise on daily and longer time scales, and the correlated noise over 2 h. Figure 3 shows the total noise over 2 h, for stars identified as likely dwarfs and giants from *R* versus *B* − *V* colour magnitude diagrams, which are shown in Fig. 4. We note that the colour−magnitude
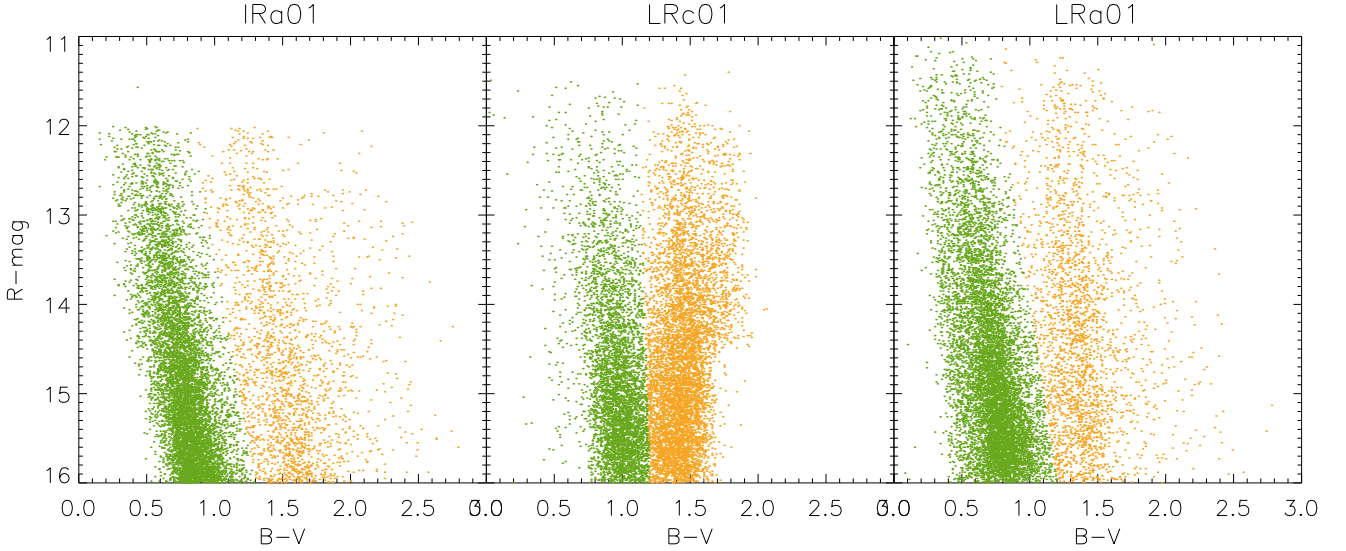
Fig. 4. *R*-magnitude versus *B* − *V* colour of all CoRoT targets in each field (data from Exodat, Deleuil et al., submitted). Green dots correspond to stars identified as likely dwarfs, orange dots to likely giants. The cut between dwarfs and giants was empirically set as a straight line running from *B* − *V* = 0.7, 1.1 and 0.7 at *R* = 11 to *B* − *V* = 1.3, 1.2 and 1.2 at *R* = 16 for IRa01, LRc01 and LRa01 respectively. The larger fraction of giants in LRc01 is clearly visible.

cuts we adopted are an over simplification, and in particular also select some K-dwarfs as giants, but that the global inferred fractions of giants in each run (20% in IRa01, 58% in LRc01 and 22% in LRa01) agree with more sophisticated estimates based on multi-colour photometry and, where available, spectroscopy (see Barbieri et al., in prep.; Gandolfi et al., in prep.).

4. Discussion

The lower envelope of the point-to-point scatter (Fig. 2) is a factor 2 above the photon noise at the bright end, and rises somewhat more steeply (slope 0.25 in log, versus 0.2 for photon noise). However, the correlated noise on transit time scales is comparable to or slightly smaller than the white noise on 2 h time scales. The empirical relations between total 2 h noise and *R*-magnitude for dwarfs, shown in Fig. 3, are given by:

$$\log \sigma_{2\,\mathrm{h}} = 0.25\,R + z \qquad (1)$$

where $z = -6.95$ for IRa01, -7.0 for LRc01 and -7.05 for IRa01, indicating a gradual degradation of the photometric performance over time, which may be associated with the increase in incidence of hot pixel events. Again, the slope of this relation is slightly above what is expected for source photon noise.

The pre-launch specification for CoRoT's photometric performance is 7.7×10^{-4} in 1 h down to $V = 15.5$ (Auvergne et al. 2003). This specification was designed to ensure that transits of planets down to approximately 1.5 R_{Earth} in short-period (<1 week) orbits are detectable, assuming that noise averages out in a white manner over time scales longer than 1 h. Thus it corresponds to 5.5×10^{-4} in 2 h, while $V = 15.5$ corresponds approximately to $R = 15$ for a typical CoRoT dwarf target. Equation (1) gives values of 6.3, 7.1 and 7.9×10^{-4} at $R = 15$, close to, but slowly receding from, the pre-launch specification.

Although this level of photometric performance is already a vast improvement over any previously available long-baseline photometry of large samples of stars, there clearly remains scope for improvement. The behaviour of the correlated noise with magnitude suggests it may dominated by effects such as satellite

Table 1. Percentage of variable stars as a function of magnitude (see text for details, numbers in brackets refer to dwarfs only).

R-mag	% variable stars					
	IRa01		LRc01		LRa01	
11.0	–		–		60	(63)
11.5	–		67	(35)	62	(47)
12.0	56	(38)	68	(36)	46	(42)
12.5	42	(33)	58	(27)	39	(31)
13.0	35	(30)	54	(26)	31	(33)
13.5	28	(24)	37	(17)	22	(25)
14.0	20	(21)	20	(14)	17	(22)
14.5	17	(19)	9	(9)	14	(19)
15.0	11	(13)	5	(8)	11	(15)
15.5	9	(11)	5	(6)	11	(13)
16.0	8	(9)	5	(8)	–	

No data is shown for magnitude bins containing fewer than 20 objects.

pointing jitter (as the aperture masks are smaller for fainter stars) and imperfect background correction. Improvements to the light curve generation have recently been tested on CoRoT data, including a new jitter correction based on improved PSF modelling (De Oliveira Fiahlo et al., in prep.), or post-light curve generation ensemble analysis (Mazeh et al. 2009). These will hopefully lead to improved performance in the short to medium term. The importance of developing an effective method for treating light curve discontinuities, more sophisticated than the detrending applied here, is rapidly increasing, as visual examination of light curves from LRa01 indicates that almost every light curve is affected by at least one discernible discontinuity, compared to approximately one in four in IRa01. Such techniques are under development, though it is challenging to apply them in an automated fashion without unwittingly affecting other signals in the light curves.

Although we now have a simple relationship between noise on transit time scales and magnitude for dwarf stars, which can be used to estimate transit detection thresholds (Moutou et al. 2009) for a statistical interpretation of the CoRoT planet catch

(Fressin et al., in prep.), it is also important to know what fraction of the stars deviate significantly from this relation. Table 1 lists the percentage of light curves, in each run, whose scatter on 2 h time scales is more than 1.5 times above the relation given by Eq. (1). At the bright end, 50% or more of the light curves show significant variability on transit time scales, rising to 70% in LRc01, while the fraction of drops to a few percent at the faint end.

Interestingly, because it is closer to the direction of the Galactic centre, LRc01 has a significantly larger fraction of giants (see Fig. 4). Figure 3 clearly shows that a large fraction of stars identified as likely K-dwarfs or giants from the R versus $B - V$ colour−magnitude diagram vary at the 0.5 mmag level on 2 h time scales, whereas no such trend is seen in stars identified as likely F or G dwarfs. In a detailed Fourier domain study of the light curves of a sample of red giants observed by CoRoT, Kallinger et al. (2009) find significantly enhanced variability on time scales of a few hours compared to the Sun, but we confirm here that this is a systematic effect, clearly visible over thousands of stars spanning a range of colours. We note that this enhanced variability was expected from 3D simulations of granulation (Freytag 2001; Svensson & Ludwig 2005), which indicate that granulation cell size – hence the strength and time scale of the photometric signature of granulation – increase significantly as surface gravity decreases.

5. Conclusions

We have evaluated the noise per exposure and on 2 h time scales for the first three CoRoT observing runs lasting more than a month. We find that the photometric performance on transit time scales is close to the pre-launch specification, with a level of correlated noise at least an order of magnitude below that obtained from the ground. The observed performance is good enough to detect transits of planets of a few Earth radii in short period (<3 d) orbits around bright stars.

However, there is scope for improvement, since the noise level per exposure exceeds the photon noise by a factor 2 to 3 and the level of correlated noise, although low, is non-zero. The slight degradation of the performance from IRa01 to LRa01 indicates that the impact of hot pixels is gradually increasing, so that improving the treatment of these artifacts is becoming increasingly important.

We also show that giants, tentatively identified from the R versus $B - V$ colour−magnitude diagram, constitute a significant fraction of the CoRoT targets (from 20% for Monoceros runs to 50% for Aquila runs). Giants are in any case less than ideal targets for transit surveys due to their large radii, but we also find that they tend to be variable at the 0.5 mmag level on transit time scales, which we interpret as a granulation effect.

Finally, one aspect that has not been touched upon here so far, but will be explored in more detail in a forthcoming paper, is that the global variability levels on time scales of a day and longer are relatively high (see Fig. 2): very few CoRoT light curves have dispersions below 1 mmag on those time scales. For comparison, we measured the scatter of the total irradiance variations of the Sun from SoHO/VIRGO/PMO6 at times of minimum and maximum activity in the same way (see Aigrain et al. 2004, and references therein for details on this dataset), finding 0.07 and 0.3 mmag respectively. However, to quantify this tentative but interesting result further requires differentiation between the effects of instrumental decay, hot pixels, satellite pointing jitter and true stellar variability, which is beyond the scope of this paper.

References

Aigrain, S., & Irwin, M. 2004, MNRAS, 350, 331
Aigrain, S., Favata, F., & Gilmore, G. 2004, A&A, 414, 1139
Auvergne, M., Boisnard, L., Buey, J.-T. M., et al. 2003, in Proc. SPIE 4854, ed. J. C. Blades, & O. H. W. Siegmund, 170
Auvergne, M., Bodin, P., Boisnard, L., et al. 2009, A&A, 506, 411
Freytag, B. 2001, in 11th Cambridge Workshop on Cool Stars, Stellar Systems and the Sun, ed. R. J. Garcia Lopez, R. Rebolo, & M. R. Zapaterio Osorio, ASP Conf. Ser., 223, 785
Hoaglin, D. C., Mostellar, F., & Tukey, J. W. 1983, Understanding Robust and exoploratory data analysis (New York: John Wiley)
Kallinger, T., Weiss, W. W., Barban, C., et al. 2009, A&A, submitted
Mazeh, T., Guterman, P., Aigrain, S., et al. 2009, A&A, 506, 431
Moutou, C., Pont, F., Bouchy, F., et al. 2009, A&A, 506, 321
Pinheiro da Silva, L., Rolland, G., Lapeyrere, V., & Auvergne, M. 2008, MNRAS, 384, 1337
Pont, F., Zucker, S., & Queloz, D. 2006, MNRAS, 373, 231
Svensson, F., & Ludwig, H.-G. 2005, in 13th Cambridge Workshop on Cool Stars, Stellar Systems and the Sun, ed. F. Favata, G. A. J. Hussain, & B. Battrick, ESA SP, 560, 979

[1] School of Physics, University of Exeter, Exeter, EX4 4QL, UK
e-mail: suz@astro.ex.ac.uk
[2] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, US
[3] LAM, Université de Provence, 13388 Marseille, France
[4] LESIA, Observatoire de Paris, 92195 Meudon, France
[5] IAS, Université Paris XI, 91405 Orsay, France
[6] Observatoire de Haute-Provence, 04870 St Michel l'Observatoire, France
[7] IAC, 38205 La Laguna, Spain
[8] ON/MCT, 20921-030, Rio de Janeiro, Brazil
[9] IfA, University of Vienna, 1180 Vienna, Austria
[10] Institute of Planetary Research, DLR, 12489 Berlin, Germany
[11] RSSD, ESA/ESTEC, 2200 Noordwijk, The Netherlands
[12] IWF, Austrian Academy of Sciences, 8042 Graz, Austria
[13] IAG, Université de Liège, Liège 1, Belgium
[14] Sch. Physics & Astronomy, Tel Aviv Univ., Tel Aviv 69978, Israel
[15] RIU, Universität zu Köln, 50931 Köln, Germany
[16] Observatoire de Genève, 1290 Sauverny, Switzerland
[17] ZAA, TU Berlin, 10623 Berlin, Germany
[18] LUTH, Observatoire de Paris, 92195 Meudon, France
[19] Thüringer Landessternwarte, 07778 Tautenburg, Germany
[20] Dept. Geophysics & Planetary Sciences, Tel Aviv Univ., Tel Aviv 69978, Israel