Removing systematics from the CoRoT [*] light curves: I. Magnitude-Dependent Zero Point

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

(Affiliations can be found after the references)

ABSTRACT

This paper presents an analysis that searched for systematic effects within the CoRoT exoplanet field light curves. The analysis identified a systematic effect that modified the zero point of most CoRoT exposures as a function of stellar magnitude. We could find this effect only after preparing a set of learning light curves that were relatively free of stellar and instrumental noise. Correcting for this effect, rejecting outliers that appear in almost every exposure, and applying SysRem, reduced the stellar RMS by about 20%, without attenuating transit signals.

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

1. Introduction

CoRoT was launched on December 2006 to perform wide-field accurate stellar photometry (Rouan et al. 1998; Baglin et al. 2006), aiming to study the internal structure of stars using seismology and to detect small planets using the transit method. Both goals required ultra-high accuracy, interruption-free photometry, which the space platform, free from atmospheric disturbances, should afford. Nevertheless, the first six transiting planets discovered by CoRoT, CoRoT-1b to -6b (Barge et al. 2008; Alonso et al. 2008; Deleuil et al. 2008; Aigrain et al. 2008, Rauer et al. submitted; Fridlund et al., in preparation), all have a transit depth of 1–2%, easily detectable by ground-based photometry. Only at the first CoRoT international conference, held in February 2009, the CoRoT team announced the discovery of the super-Earth CoRoT-7b (Leger et al., submitted), with transit depth of 0.03%, which ground-based observations could never have detected.

The light curve of the bright star CoRoT-7 is of exceptional quality, with a RMS scatter over transit timescales of only 0.01%. However, the majority of CoRoT light curves contain systematics and correlated noise, which is probably associated with satellite jitter, stellar activity, cosmic ray impacts and possibly other effects. Albeit extremely low compared to ground-based surveys (Aigrain et al. 2009), this noise should nonetheless be removed before planets with shallow transits like the ones of CoRoT-7b could be detected.

In order to remove the systematic effects we have previously applied SysRem (Tamuz et al. 2005) to the CoRoT data, result-

ing in reduction of the noise level by 10–20%. However, SysRem was sensitive to *temporal* variability that was shared by many of the stars, and was not specifically tuned to detect collective effects that showed up when considering the measurements of each exposure separately. One such obvious effect was the zero point of the different exposures, which could have been modulated by the satellite motion in and out of the Earth shadow and going through the South Atlantic Anomaly, and could also depend on the stellar characteristics. When we searched for such an effect in the LRa01 CoRoT run we discovered that the zero points of most exposures depended on the stellar magnitude. This could be noticed only after we prepared a 'learning' set of light curves that were relatively free of stellar and instrumental noise. When this effect was removed, the noise level was reduced by about 20% for the faint stars.

Section 2 details our analysis, including the preparation of the learning set of light curves, and the procedure to remove the effect found, and Section 3 discusses in short our findings.

2. Data analysis

Our analysis was done on the white light curves obtained during the 150-day CoRoT LRa01 run (Auvergne et al. 2009). We first transformed the CoRoT fluxes into magnitudes, and then subtracted from each star its median magnitude. The residuals of each measurement relative to its stellar median — $\{r_{ij}\}$, where i was the star number and j was the exposure number, were the subject of this analysis. We concentrated on the residuals derived for any given exposure j_0: $\{r_{ij_0}; i = 1, M\}$, where M was the number of stars in this CoRoT run, searching for systematic effects in exposure j_0.

In order to be able to notice relatively small systematic effects we prepared 'clean' learning light curves by the following measures:

– Divided the LRa01 run into 13 blocks, each block contained data taken over only ten days.
– Removed invalid data marked by the CoRoT regular N2 pipeline and rejected outliers from each light curve on a temporal identification basis.
– In each block, rejected the stars that showed high variability in that block, either because of stellar variability or because of obvious instrumental effects, such as hot pixel features. Thus, some stars appeared only in some of the blocks.
– Removed long-term variability by subtracting a running median, taken over 3 satellite orbits. Small hot-pixel features were also attenuated by the median filtering.
– Considered the light curves from the two CCDs separately, but combined the two halves of each CCD together. This was done as we suspected that the two CCDs will have different systematics.
– Some of the light curves were sampled with a rate of 32 sec, while most stars with a rate of 512 sec. To make the analysis even more difficult, some stars had a light curve that was first sampled with the 512 sec rate, which was then switched into the 32 sec. We therefore had to put all the 512 sec and 32 sec measurements of CoRoT together in a common synchronous grid. An accurate common 512 sec exposure time was emulated as CoRoT would have done on board, rather than obtained by interpolation. This enabled us to associate each of the stellar measurements with the correct exposure.

As a result of our selection, we were left, in the first block on the E1 CCD for example, with 4111 learning light curves out of 5704 ones. Only after the preparation of the 'learning' sets of light curves the collective features of the data appeared in our analysis with relatively small spread, so we could accurately estimate and subtract these features.

We found that the zero points of most exposures depended on stellar magnitude, an effect that we dubbed MagZeP (Magnitude Zero Point). To show this effect we plotted in Figure 1 the residuals of two exposures, no. 12 and 55, from our first block. The figure showed two prominent features of the MagZeP effect:

– The scatter around the general trend was larger for fainter stars
– The residuals took a parabolic shape, with different curvature sign and slope for different exposures. The typical amplitudes of this effect were 0.005 mag for the faint objects.

While the first feature was expected, as the S/N is smaller, the second effect took us by surprise.

To remove the MagZeP effect we fitted the residuals of each exposure j by a parabola that depended on the stellar magnitude m_i:

$$p_{ij} = a_{0,j} + a_{1,j}m_i + a_{2,j}m_i^2 \,, \tag{1}$$

and then derived new residuals

$$\tilde{r}_{ij} = r_{ij} - p_{ij} \,, \tag{2}$$

where $\tilde{r}_{ij}$ presented our best estimate of the stellar magnitude, relative to its median.

We also noticed that for almost every exposure some outliers clearly stood out, as could be seen in Figure 1. To identify these outliers we assigned to each measurement an error, σ_{ij}, based on the collective scatter at the corresponding exposure and magnitude. We then rejected residuals that were *smaller* than zero

(after the removal of the parabola) by more than η times their error

$$\tilde{r}_{ij} \leq \eta\sigma_{ij} \,, \tag{3}$$

where η was a parameter, taken in our present implementation of the algorithm to be equal to 2. Keeping the positive residuals ensured us that we did not attenuate any transit signal. Removing these exposure outliers further improved the stellar scatter. The parabola was derived by a robust regression based on iteratively reweighed least-squares fit, which was relatively insensitive to the presence of outliers. In this way, a robust result was obtained even without a need for additional fitting-clipping cycle.

Finally, *all* CoRoT light curves were cleaned with the parameters determined by the selected learning set of light curves. This step of the analysis produced a homogeneous set of light curves in each block.

3. Discussion

We propose here a statistical algorithm to deal with the CoRoT data, as a complementary process of the regular N2 CoRoT pipeline. The latter includes only model-based corrections of identified physical effects, while ours, which is a generalization of a zero-point removal of each exposure (Tamuz et al. 2005; Collier Cameron et al. 2006), relies only on the collective effects identified in the data. We find that the zero point of each exposure depend on the stellar magnitude. Obviously, other effects can still be present in the data, and thus in our implementation we apply SysRem after the MagZeP removal.

The results of applying MagZeP and outlier removal and then applying SysRem are depicted in Figure 2, which presents the ratio between the RMS before and after applying our approach. We can see that the improvement is a strong function of the stellar magnitude, and the *averaged* improvement ratio is almost linear with the R magnitude, and can reach 25% for the faint stars. Most of the improvement is due to MagZeP and outlier removal, and adding SysRem has a minor impact on our approach. Applying SysRem *alone* does achieve less effective clean up. We find that the improvement of the light curve is less pronounced at the middle of the LRa01 run in both CCDs. This is because the CoRoT Earth-shadow crossing (see below) occurs only at the beginning and the end of the run, generating more pronounced systematic effects.

One danger of removing collective effects from a set of light curves is that the process might attenuate the signal of a possible transit. To show that this is not the case here we present in Figure 3 the light curve of CoRoT-7, after removing the stellar variability and then folding the residuals with the planetary orbital period of 0.854 days, binning the data such that each bin presents an average of 10 measurements. For comparison, the figure also depicts the light curve before applying our analysis. The signal is clearer after our analysis.

To look for the source of the MagZeP effect we considered the set of parameters of the fitted parabolas of all exposures — $\{a_{0,j}, j = 1, N\}$ for example, where N was the number of exposures included in the block being analysed. This set of parameters reflected the zero-order brightness removed by our algorithm. We folded this set of parameters with the orbital period of the satellite, the results were being plotted in Figure 4.

The two positive bumps seen in the figure occurred when the satellite was entering and exiting the Earth shadow, while the negative outliers, at phase 0.05 and 0.45, coincided with crossing the South Atlantic Anomaly (see for example Auvergne et al.
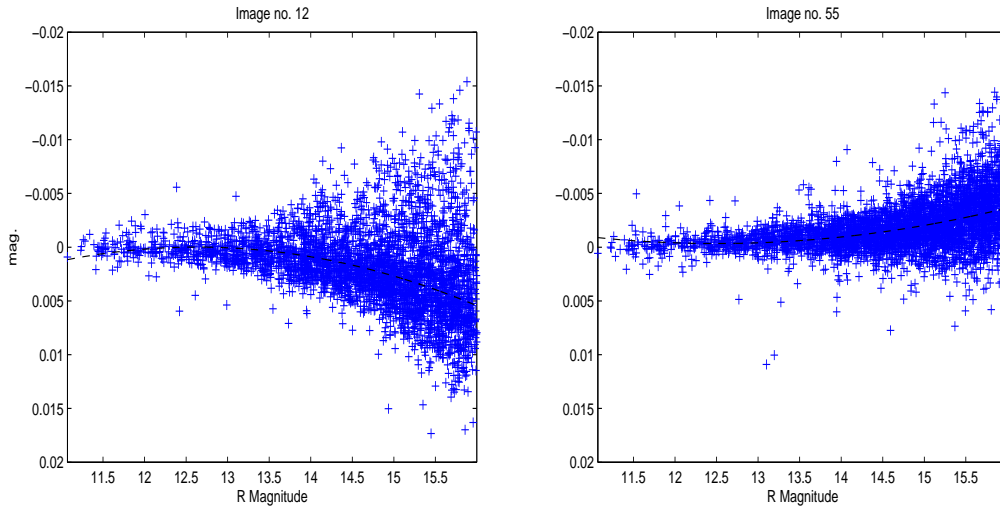
Fig. 1. The residuals of two exposures as a function of their stellar R magnitude. The continuous line is our parabolic fit. The left-hand-side panel shows the residuals of exposure 2007-10-24T14:09:15.000, while the right-hand-side panel comes from exposure 2007-10-24T20:16:11.000.
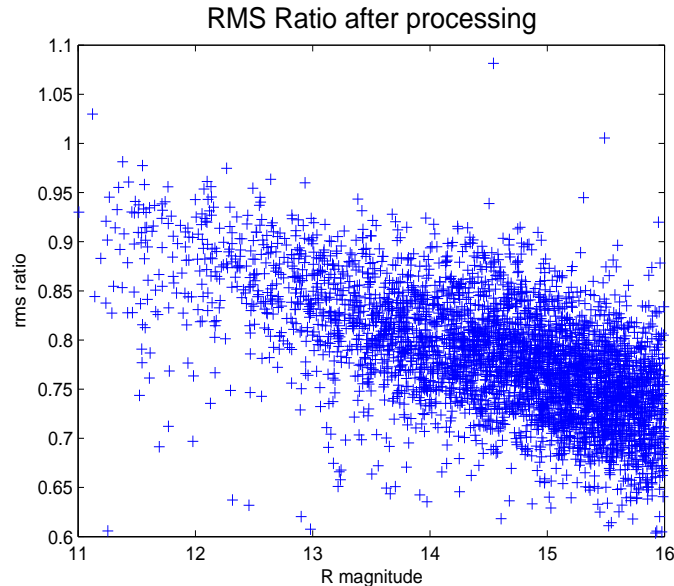

Fig. 2. The ratio between the RMS before and after our algorithm (MagZeP and exposure-outlier removal and SysRem) was applied. The figure presents 3893 stars observed by the E2 CCD in the first block of LRa01, with 1693 exposures

2009). We also noticed an overall slight curvature over the whole phase, with a small displacement between the two bumps. The reason for this feature was not clear.

We wish we had a complete model that accounts for the MagZeP effect, which is probably associated with an additive and multiplicative factor of stellar flux. The parabolic fit we use is only an approximation to the exact function of the MagZeP effect. However, a detailed analysis of the nature of the effect is out of the scope of this paper and is therefore deferred for future work. The goal of this short communication is to point out to the effect and the impact of its removal.

While trying to improve the algorithm and searching for other collective effects, we have processed the whole CoRoT N2 white colour dataset. The cleaned data are now available to the whole CoRoT community.

References

Aigrain, S., et al. 2008, A&A, 488, L43
Aigrain, S., et al. 2009, arXiv:0903.1829
Alonso, R., et al. 2008, A&A, 482, L21
Auvergne, M., Bodin, P., Boisnard, L., Buey, J. -., Chaintreuil, S., & CoRoT team 2009, arXiv:0901.2206
Baglin, A., et al. 2006, 36th COSPAR Scientific Assembly, 36, 3749
Barge, P., et al. 2008, A&A, 482, L17
Collier Cameron, A., et al. 2006, MNRAS, 373, 799
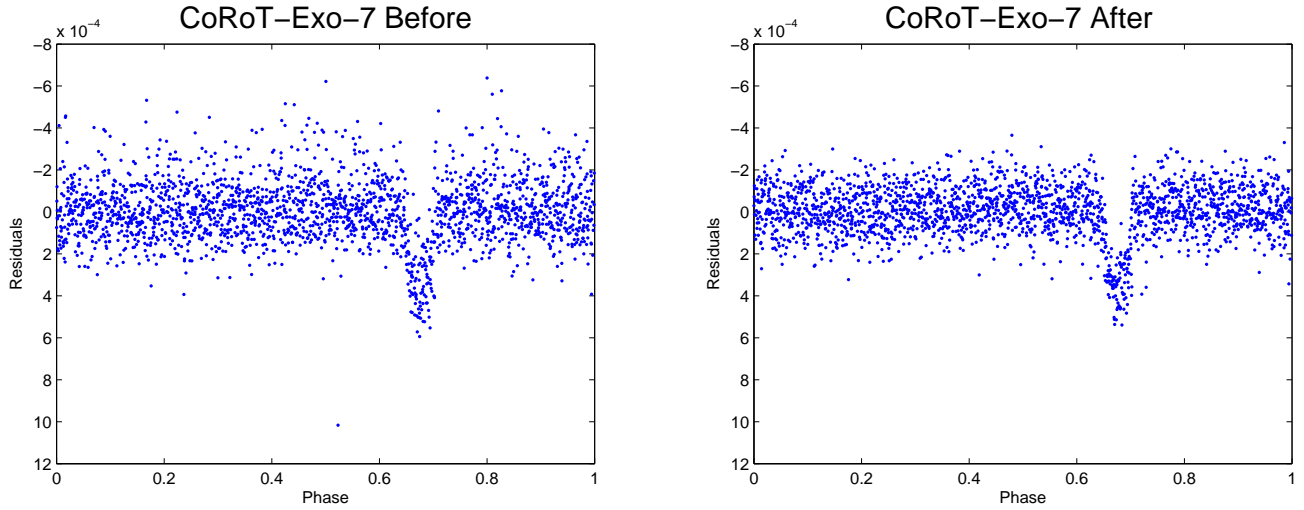Deleuil, M., et al. 2008, A&A, 491, 889

Fig. 3. The folded and binned light curve of CoRoT-7, before and after our analysis.
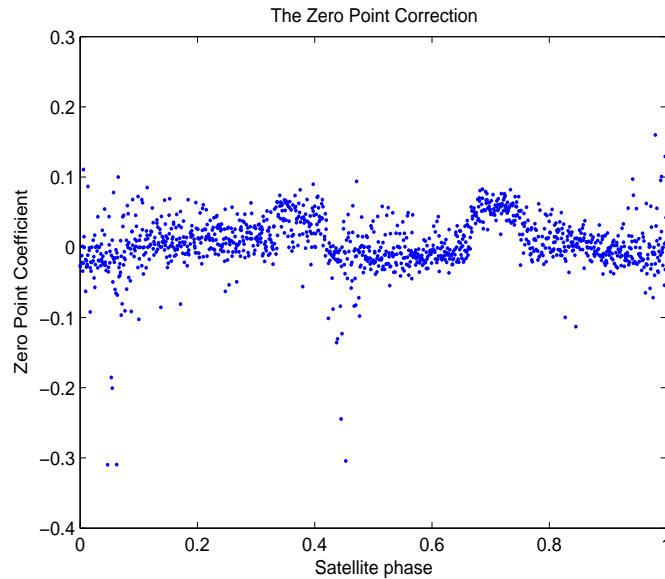

Fig. 4. The set of zero-order parameters of all parabolas of the first block, $\{a_{0,j}, j = 1, N\}$, folded with the orbital period of the satellite.

Rouan, D., Baglin, A., Copet, E., Schneider, J., Barge, P., Deleuil, M., Vuillemin, A., & Léger, A. 1998, Earth Moon and Planets, 81, 79

Tamuz, O., Mazeh, T., & Zucker, S. 2005, MNRAS, 356, 1466

[1] Radcliffe Institute for Advanced Studies at Harvard, 8 Garden St., Cambridge, MA 02138, USA

[2] On Sabbatical leave from School of Physics and Astronomy, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel

[3] Le Laboratoire d'Astrophysique de Marseille, UMR 6110 CNRS, Technopôle de Marseille-Etoile, F-13388 Marseille cedex 13, France

[4] School of Physics, University of Exeter, Physics Building, Stocker Road, Exeter, Ex4 4QL, UK

[5] Geophysics and Planetary Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel

[6] School of Physics and Astronomy, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel

[7] LESIA, Observatoire de Paris, 92195 Meudon, France

[8] IAS, Université Paris XI, 91405 Orsay, France

[9] Observatoire de Haute-Provence, 04870 St Michel lObservatoire,France

[10] IAC, E-38205 La Laguna, Spain

[11] ON/MCT, 20921-030, Rio de Janeiro, Brazil

[12] IfA, University of Vienna, 1180 Vienna, Austria

[13] Institute of Planetary Research, DLR, 12489 Berlin, Germany

[14] RSSD, ESA/ESTEC, 2200 Noordwijk, The Netherlands

[15] IWF, Austrian Academy of Sciences, A-8042 Graz, Austria

[16] IAG, Université de Liège, Liège 1, Belgium

[17] RIU, Universität zu Köln, 50931 Köln, Germany

[18] Observatoire de Genève, 1290 Sauverny, Switzerland

[19] ZAA, TU Berlin, D-10623 Berlin, Germany

[20] LUTH, Observatoire de Paris, 92195 Meudon, France

[21] Thüringer Landessternwarte, 07778 Tautenburg, Germany