SF2A 2017
C. Reylé, P. Di Matteo, F. Herpin, É. Lagadec, A. Lançon, Z. Meliani, F. Royer (eds)

PERSPECTIVES FOR SHORT TIMESCALE VARIABILITY STUDIES WITH GAIA

M. Roelens[1], L. Eyer[1], N. Mowlavi[1], I. Lecoeur-Taïbi[2], L. Rimoldini[2], S. Blanco-Cuaresma[3], L. Palaversa[4], M. Süveges[5], J. Charnas[2] and T. Wevers[6]

Abstract. We assess the potential of *Gaia* for detecting and characterizing short timescale variables, i.e. at timescale from a few seconds to a dozen hours, through extensive light-curve simulations for various short timescale variable types, including both periodic and non-periodic variability. We evidence that the *variogram* analysis applied to *Gaia* photometry should enable to detect such fast variability phenomena, down to amplitudes of a few millimagnitudes, with limited contamination from longer timescale variables or constant sources. This approach also gives valuable information on the typical timescale(s) of the considered variation, which could complement results of classical period search methods, and help prepare ground-based follow-up of the *Gaia* short timescale candidates.

Keywords: Stars: variables: general, surveys, methods: data analysis, numerical, techniques: photometric

1 Introduction

Since the first reported discoveries of such astronomical objects, hundreds of thousands of variable stars have been identified and classified in different categories, revealing a great diversity in terms of timescales, amplitudes and phenomena at the origin of the variability.

In this work, we focus on *short timescale variability*, i.e. on sources showing variability at timescales from tens of seconds to a dozen hours. Various astronomical objects are known to exhibit such rapid variations in their light-curves, be it periodic or not, with amplitudes ranking from a few millimagnitudes (mmag) to a few magnitudes. Until a decade ago, only a relatively small number of these fast variables had been detected, due to the constraints in terms of time sampling and photometric precision when dealing with such objects. However, with the advent of Charged Coupled Devices (CCDs), and thanks to space and ground-based high cadence monitoring surveys (such as Kepler (Borucki et al. 2010) and the Optical Gravitational Lensing Experiment (Udalski et al. 1992) respectively), a deeper insight the domain of short timescale variability became accessible, and the number of known short timescale variables significantly increased.

In this context, the *Gaia* ESA cornerstone mission, launched in December 2013, offers a unique opportunity to drastically change the landscape. During its 5-year mission duration, *Gaia* will survey more than one billion objects over the entire sky, providing micro-arcsecond astrometry, photometry down to $G \approx 20.7$ mag (where G is the *Gaia* broad-band white light magnitude) with standard errors below the mmag level for bright sources, and medium resolution spectroscopy down to $G \approx 17$ mag (Gaia Collaboration et al. 2016). Moreover, the *Gaia* scanning law involves fast observing cadences, with groups of nine consecutive CCD observations separated by about 4.85 s from each other, followed by gaps of 1 h 46 min or 4 h 14 min between two successive groups, a group being referred to as a field-of-view (FoV) transit. Hence, *Gaia* will make a comprehensive variability search possible, enabling to investigate timescales as short as a few tens of seconds together with low amplitude variations.

[1] Observatoire de Genève, Département d'Astronomie, Université de Genève, Chemin des Maillettes 51, CH-1290 Versoix, Switzerland
[2] Observatoire de Genève, Département d'Astronomie, Université de Genève, Chemin d'Ecogia 16, CH-1290 Versoix, Switzerland
[3] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[4] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 OHA, United Kingdom
[5] Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany
[6] Department of Astrophysics / IMAPP, Radboud University, P.O. Box 9010, NL-6500GL, Nijmegen, The Netherlands

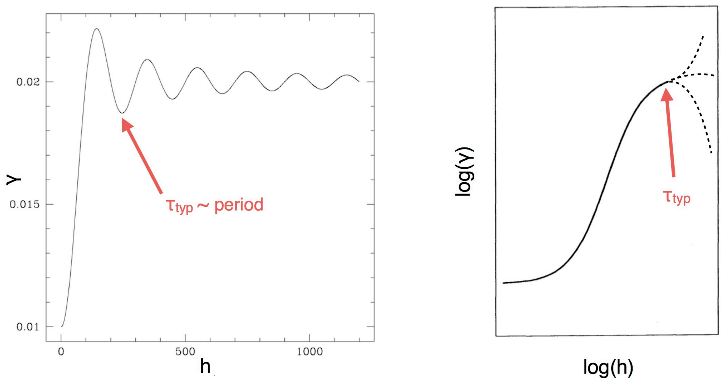

Fig. 1. Typical variogram plots. **Left:** for a periodic/pseudo-periodic variable (Eyer & Genton 1999). **Right:** for a transient variable, only exploring lags up to the first structure characteristic of variability (derived from Hughes et al. 1992). In each case, the feature used to estimate the typical timescale is pointed by a colored arrow.

In this proceeding, we estimate the *Gaia* capabilities for detecting and characterizing short timescale variability detection, using *Gaia* per-CCD photometry in the G band. Our study is based on light-curve simulations for various short timescale variable types, and relies on the *variogram method*, also known as the *structure function method*, which is extensively used in the fields of quasar and AGN studies, and in high-energy astrophysics (see e.g. Simonetti et al. 1985; MacLeod et al. 2012; Kozłowski 2016), but can also be applied to optical stellar variability (see e.g. Eyer & Genton 1999; Sumi et al. 2005). This variogram technique and its application to short timescale variability detection are detailed in Sect. 2. In Sect. 3, we describe the light-curve data sets we generated for our analysis, and summarize the expectations in the *Gaia* context. Section 4 recapitulates our conclusions.

2 Principle of the variogram analysis

The key idea of the variogram method is to investigate variability by quantifying the magnitude difference between two measurements as function of the time lag h between them. If the light-curve is defined by magnitudes $(m_i)_{i=1..n}$ observed at times $(t_i)_{i=1..n}$, then the variogram value for a time lag h is defined as (Hughes et al. 1992):

$$\gamma(h) = \sum_{i>j} \frac{(m_j - m_i)^2}{N_h} \tag{2.1}$$

and is computed on all pairs (i, j) such that $|t_j - t_i| = h \pm \epsilon_h$ where ϵ_h is the tolerance accepted for grouping the pairs by time lag (N_h is the number of such pairs).

The *variogram plot* (hereafter referred to as variogram) associated with this time series is built by exploring different lag values h_k and calculating the associated variogram values $\gamma(h_k)$. It gives information on how variable the considered source is, and on its variability characteristics if appropriate. Figure 1 shows the typical variograms for a periodic or pseudo-periodic variable (top), and for a transient variable (bottom). If the analyzed time series exhibits some variability, the expected features in its variogram are:

1. a plateau at the shortest lags,

2. towards longer lags, an increase in the variogram values, followed by a flattening phase.

When the underlying variation is periodic or pseudo-periodic, this flattening is followed by a succession of dips. In the case of a transient variation, the flattening can be followed by complex structures, e.g. other plateaus or a decrease in the variogram values, depending on the origin of variability. The lags at which those features occur allow us to estimate the characteristic variation timescales. For transient variability, the lag(s) at which the variogram flattens correspond to the typical timescale(s) τ_{typ} (see Fig. 1). For periodic variability, τ_{typ} corresponds to the lag of the first dip after the plateauing, and gives a rough estimate of the period of the variability.

In the *Gaia* context, once we retrieve the variogram associated with a given light-curve, the first think we have to do is to decide whether the considered source is a true variable or not. To answer this question, one possibility is to fix a *detection threshold* γ_{det} such that: if, for at least one lag value h_k, $\gamma(h_k) \geq \gamma_{det}$, then the source is flagged as *variable*. Otherwise, the source is flagged as *constant*. Hence γ_{det} defines the variance level above which variability in the signal is significant enough not to be due only to noise, and can be more or less restrictive. If a source is detected as short timescale candidate with this criterion, then the corresponding *detection timescale* τ_{det} is defined as the smallest lag for which $\gamma(\tau_{det}) \geq \gamma_{det}$. It quantifies the average variation rate in the investigated light-curve and is characteristic of the underlying variability. So as to focus on fast variability, we complete our detection criterion by an upper limit on the detection timescale: a detected source is flagged as a *short timescale candidate* only if $\tau_{det} \leq 0.5$ d. Finally, we estimate the typical timescales for those flagged short timescale candidates as explained above.

All in all, the interest of the variogram method lies in the fact that it enables to detect and to characterize variable candidates, handling both periodic, pseudo-periodic and non-periodic variability, though it is a complement and absolutely not a substitute to more precise period search methods, e.g. the Fourier periodograms.

3 Short timescale variability detection: what we can expect with Gaia

To evaluate the efficiency of the variogram method for detecting short timescale variables from *Gaia* data, we simulate different light-curve data sets for various types of such astronomical objects. Our sample includes eight different periodic variable types, covering a wide range of periods (from 30 s to 12 h) and amplitudes (from a few mmags to a few mags), as well as transient events, i.e. M dwarf flares and supernovae (SNe). Note that not all short-period variable types are included in this work, and that we adopt a simplified approach, simulating each periodic light-cuve with one single period P and not treating multiperiodicity. Additionally, SNe are not short timescale variables per se, since their duration is much longer than 1 d. Nevertheless, SNe can experience quite fast and significant brightening, with a variation rate of the order of 0.1 mag/d. Given the precision of the *Gaia* G photometry, if the brightening phase of a supernova is sampled by *Gaia*, then we should be able to detect significant variation at the short timescale level.

For our analysis, we generate two different types of light-curves:

1. The *Gaia-like* light-curves, with a time sampling corresponding to the expected *Gaia* observation times for a random position in the sky, over a timespan $\Delta t \approx 5$ yrs (which is the nominal duration of the *Gaia* mission), and adding noise according to a magnitude-error distribution retrieved from real *Gaia* data, similar to the distribution presented in Fig. 6 of Eyer et al. (2017).

2. The *continuous* light-curves, corresponding to the same variables as in the continuous data set (same period or duration, amplitude and magnitude), but this time without noise and with a dense and regular time sampling, for comparison purposes.

For each simulated continuous light-curve, we calculate the associated *theoretical* variogram, for the appropriate lag values defined by the underlying time sampling (i.e. explored lags are multiple of the time interval δt). Similarly, we compute *observational* variograms associated with each simulated short timescale variable *Gaia*-like light-curve. This time, the explored lags are defined by the *Gaia* scanning law, i.e. the time intervals between CCD measurements (4.85 s, 9.7 s, 14.6 s, 19.4 s, 24.3 s, 29.2 s, 34 s and 38.8 s), and those between the different FoV transits (1 h 46 min, 4 h 14 min, 6 h, 7 h 46 min, etc), up to $h \approx 1.5$ d. Note that no lag can be explored from about 40 s to 1 h 46 min, which may have consequences on the detectability and timescales estimation of some sources.

Figures 2 and 3 show examples of *continuous* and *Gaia-like* light-curves and associated variograms, for a δ Scuti star and an M dwarf flare respectively. For those two cases, the source would be detected as short timescale candidate with $\gamma_{det} = 10^{-3}$ mag^2, not only in an ideal situation but also in the *Gaia-like* context, though the detection timescale is pushed from an ideal τ_{det} of a few minutes to 1 h 46 min due to the *Gaia* lag gap mentioned above. Moreover, the estimated τ_{typ} from the observational variogram matches the simulation input period as expected. For the transient source, we can identify visualy three different timescales from the theoretical variogram, which roughly correspond to the increase, decrease and total duration of the transient event. But when turning to the observational variogram, because of the specific lag values we can explore, all we can say is that this transient has a typical timescale which is between 40 s and 1 h 46 min. For now, for the simulated transients, we retrieve the lag of the maximum observational variogram value as an estimate of τ_{typ}, which should correspond to the decrease duration of the event.

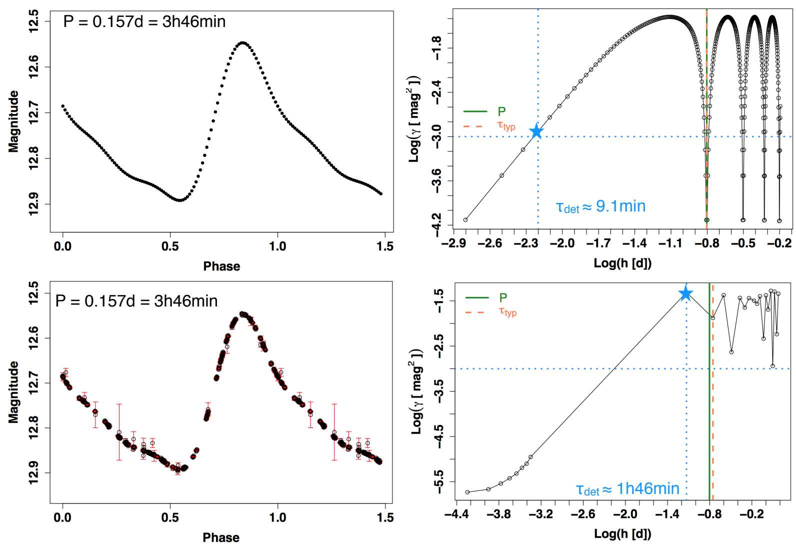

Fig. 2. Example of simulated δ Scuti light-curves. **Left:** *continuous* (top) and corresponding *Gaia-like* (bottom) light-curves, phase-folded with the input period of the simulation. **Right:** theoretical (top) and observational (bottom) variograms derived from the simulated light-curves. The blue dotted lines evidence the detection threshold (here 10^{-3} mag^2) and associated detection timescale. The green continuous line marks the simulation period, and the orange dashed line corresponds to the typical timescale estimated from the variograms.

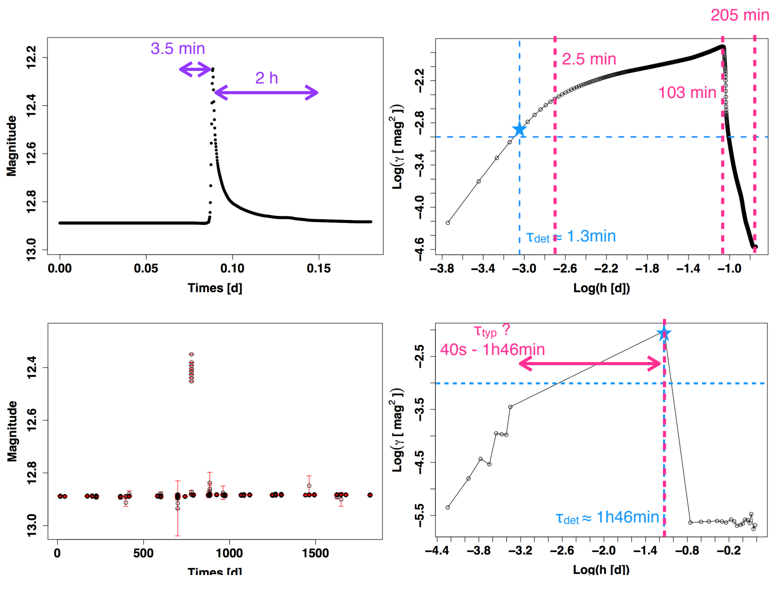

Fig. 3. Example of simulated M dwarf flare light-curves. **Left:** *continuous* (top) and corresponding *Gaia-like* (bottom) light-curves. The purple arrows indicate the approximate duration of the brightening and fading phases of the considered flare. **Right:** theoretical (top) and observational (bottom) variograms derived from the simulated light-curves. The blue dotted lines evidence the detection threshold (here 10^{-3} mag^2) and associated detection timescale. The pink dashed lines correspond to the typical timescale(s) estimated from the variograms.

Ensuring that the variogram method properly detects short timescale variables is necessary, but not sufficient. We also have to make sure that the method limits the number of unexpected detections, be it constant sources or stars showing variability on longer timescales than 12 h. Hence, we complete our *Gaia-like* data set with supplementary light-curve simulations of constant stars and sources showing sinusoidal variations with periods greater than 10 d, and calculate the corresponding observational variograms.

The short timescale detection criterion we use can be summarized as follows: a source is flagged as short timescale variable candidate if its maximum variogram value is greater than the chosen detection threshold γ_{det}, and if the associated detection timescale τ_{det} is shorter than 0.5 d. But which value of γ_{det} should we choose? Figure 4 represents the maximum variogram value as function of the mean G magnitude of the source for our full *Gaia-like* data set. As one can see, a constant threshold does not seem to be an appropriate choice: with e.g. $\gamma_{det} = 10^{-3}\,\mathrm{mag}^2$ most of the bright low amplitude variables are missed, whereas many false positive arise at the faint end. Consequently, we adopt a detection threshold depending on the mean magnitude of the source (grey continuous line in Fig. 4).

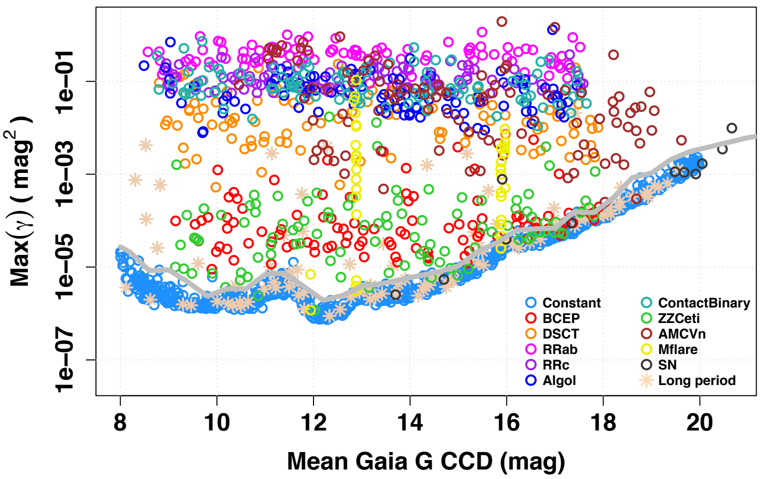

Fig. 4. Maximum variogram value as function of the mean *Gaia* G magnitude of the source, for the periodic variables and the constant sources of the *Gaia*-like data set. The grey continuous line corresponds to the definition of the detection threshold used.

With that short timescale detection criterion, about 94% of the periodic short timescale variables of the *Gaia-like* sample are recovered, as well as 30 M dwarf flares and 2 SNe. The false positive rate (i.e. contamination from constant sources) is around 0.1%, but contamination from longer period variables is as high as 16%. One way to limit that contamination could be to adopt a more restrictive definition of what short timescale variability is, e.g. with a lower detection timescale limit $\tau_{det} \leq 0.1$ d, focusing on the fastest phenomena detected. With this new upper limit on the detection timescale, the short period variable recovery rate drops from 94% to 91.9%, and we still detect 29 M dwarf flares and one supernova, whereas false positive and longer period contamination rates are significantly reduced, down to 0% and 2% respectively. Hence, with the variogram method we have a powerful criterion to identify short timescale variable candidates.

Additionally, with that approach we can further characterize our suspected variables. Figure 5 compares the typical timescale estimates for the detected short timescale candidates as function of their real characteristic timescale (i.e. input period P for the periodic sources, and decrease duration τ_{decr} for the transient ones). SNe are not treated here because their characteristic durations are longer than the maximum lag explored in the variogram analysis, thus we do not expect to retrieve a relevant timescale estimate for them. Outside the *Gaia* lag gap (brown arrows in Fig. 5) where no P or τ_{decr} good recovery is expected, we see that the typical timescale estimate, though not very precise, nevertheless gives an idea of the order of magnitude of the variation timescale. For 44% of the periodic flagged sources, τ_{typ} recovers the true period within a factor of 2. For 15% of them, our method fails to provide any typical timescale estimate. Regarding the detected M dwarf flares, for about half of them τ_{typ} recovers the decrease duration within a factor of 2. In the future, the idea would be to combine that not very precise but valuable information from the variogram with more accurate period search methods (e.g. Fourier periodograms), as well as with other variability studies performed in the *Gaia* DPAC context.

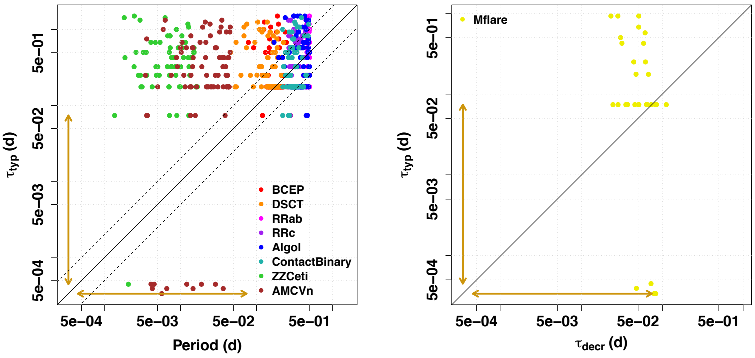

Fig. 5. Left: Typical timescale τ_{typ} as function of the input period, for the *Gaia-like* periodic simulations flagged as short timescale variables. **Right:** Typical timescale τ_{typ} as function of the decrease duration, for the *Gaia*-like transient simulations flagged as short timescale variables. The brown arrows indicate the *Gaia* lag gap.

4 Conclusions

In this proceeding, we evidence that, with an appropriate detection criterion, the variogram analysis is a very promising approach for identifying both periodic and transient short timescale variables observed by *Gaia*, from its per-CCD photometry, together with a limited contamination from non-variable and longer period variable objects. With the estimation of τ_{det} and τ_{typ}, we also retrieve valuable information on the timescale(s) and rapidity of the underlying variation. Among the perspectives it opens on the exploitation of *Gaia* data for variability analysis, in the case of periodic variability, the variogram results could be fruitfully combined with period search methods.

Nowadays, we are re-investing the knowledge and understanding we acquired on the variogram method through simulations, and are analyzing real *Gaia* data, searching for new short timescale variable candidates. We have obtained promising results, which we aim to include in the *Gaia* Data Release 2 planned for April 2018. In parallel, we have started a complementary ground-based follow-up campaign of some of our new short timescale candidates, so as to confirm the underlying supected fast variability. In the near future, we plan to further classifiy and characterize our candidates, combining all the *Gaia* data available (photometry in BP and RP, color, spectrum, parallaxes and proper motions), and exploring the performance of machine learning methods, to assess the variable type of the selected candidates.

References

Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977

Eyer, L. & Genton, M. G. 1999, A&AS, 136, 421

Eyer, L., Mowlavi, N., Evans, D. W., et al. 2017, arXiv:astro-ph/1702.03295

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2016, arXiv:astro-ph/1609.04172

Hughes, P. A., Aller, H. D., & Aller, M. F. 1992, ApJ, 396, 469

Kozłowski, S. 2016, The Astrophysical Journal, 826, 118

MacLeod, C. L., Ivezić, Ž., Sesar, B., et al. 2012, ApJ, 753, 106

Simonetti, J. H., Cordes, J. M., & Heeschen, D. S. 1985, ApJ, 296, 46

Sumi, T., Woźniak, P. R., Eyer, L., et al. 2005, MNRAS, 356, 331

Udalski, A., Szymanski, M., Kaluzny, J., Kubiak, M., & Mateo, M. 1992, Acta Astron., 42, 253