

# Characterization of variable stars by robust wave variograms: An application to Hipparcos mission

L. Eyer<sup>1,2</sup> and M.G. Genton<sup>3</sup>

<sup>1</sup>Instituut voor Sterrenkunde, Katholieke Universiteit Leuven, Celestijnenlaan 200 B, B-3001 Leuven, Belgium

<sup>2</sup>Observatoire de Genève, CH-1290 Sauverny, Switzerland

<sup>3</sup>Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, U.S.A.

Received August 6, 1998; accepted January 21, 1999

**Abstract.** Variogram analysis is applied to time series of variable stars. This method allows to characterize time-scales of varying signals. It is well suited for periodic and pseudo-periodic signals. The aspects of robustness, bin positions and bin widths for the variogram estimator are assessed, evaluated with the help of simulations and adapted to the Hipparcos photometry. For instance variogram methodology is successfully applied for the detection of spurious periods, the estimation of measurement noise and the analysis of the pseudo-period of some typical Hipparcos variable stars.

**Key words:** methods: data analysis — methods: statistical — stars: variables: general — techniques: photometric

## 1. Introduction

Often subjects, like astronomy and statistics, are developed rather independently and new methods have some difficulties to break the wall between different disciplines (see Rousseeuw 1987). There is a need of improved “tools” for the astronomer for describing the observed objects. Nature can show a huge range of different behaviours and the measurement process can lead sometimes to erroneous interpretations and results. This makes the task difficult and implies the use of robust methods.

The goal of the proposed method is to analyse objects with pseudo-periodic light curves (i.e. approximately periodic signal) such as super giant, spotted, and semi-regular variable stars. The behaviour of these stars is difficult to describe, because it is not strictly periodic and there are different time-scales involved of different amplitudes. Even in Mira type stars, for example, very short time-scales are

present (de Laverny et al. 1998). We want to determine the general time-scales causing the dispersion in the data. The method is more qualitative than quantitative at the present time although procedures are being developed to make it quantitative.

The photometric data of Hipparcos have two origins: the Main mission and the Tycho mission. The latter used the star mappers and gave two magnitudes  $B_T$  and  $V_T$ , whereas the former used the image dissector tube and produced the so called *Hp* magnitude (van Leeuwen 1997). The Hipparcos satellite provided around 13 000 000 measurements of fluxes for 118 204 stars during a time interval of 3.3 years. A systematic search for variability and periodic stars was carried out (cf. Eyer 1998) and the results were published (see Grenon et al. 1997 and van Leeuwen et al. 1997). On this occasion, the variogram methodology was applied. However, pseudo-periodic signals were published with only an information of amplitude, and no information about time-scales. Because time-scales of magnitude variations are also a valuable information, a refined method was developed in order to be able to describe pseudo-periodic signals.

The structure of the paper is now briefly described. We begin by reviewing the methods of statistical analysis of periodicity to be found in the literature. The variogram approach is introduced and closely analysed in Sect. 2. The third part contains some interesting results of our method on simulated data examples. The use of the variogram for spurious period discrimination is explained on a real case in Sect. 4. The variogram allows also to estimate the measurement noise, and is applied to the whole sample of Hipparcos stars. Finally, the robust variogram approach is applied to individual stars from Hipparcos photometry.

**2. Statistical analysis of periodicity for irregular sampling**

There are many ways to describe periodic signals (Cuypers 1997). Most of the proposed methods are valid for rather strictly periodic signals: they are not very efficient when there is not a full conservation of the phase or of the amplitude. We may cite the following methods: “string” (Renson 1978), analysis of variances (Schwarzenberg-Czerny 1989), Fourier (Deeming 1975; Ferraz-Mello 1981; Babu & Feigelson 1996). Fortunately there are other methods which are better adapted to nearly periodic signals. The wavelets (Foster 1996), the autocorrelation (Bartholdi 1988; Edelson & Krolik 1988) methods and finally the structure function (Hughes et al. 1992), or variogram. These methods are obviously linked to each other. The last one is appropriate for our situation, as will be shown.

*2.1. The variogram*

In order to describe a pseudo-periodic signal, we introduce the concept of variogram. Let us consider a time series  $\{m(t): t \in D\}$ , where  $D$  is a subset of  $\mathbb{R}_+$ . We can think of  $m(t)$  as the magnitude of a star at time  $t$ . Suppose that the signal  $m(t)$  can be decomposed into the sum of a deterministic part,  $\mu(t)$ , plus a stochastic part, the noise. Assume that this series satisfies the following hypothesis

$$E(m(t)) = \mu(t), \quad \forall t \in D, \tag{a}$$

$$\text{Var}(m(t+h) - m(t)) = 2\gamma(h), \quad \forall t, t+h \in D, \tag{b}$$

where  $2\gamma(h)$  is the variogram, which can be used in practice in order to study the periodicity of the time series.

An interesting variogram that exhibits negative correlations caused by periodicity of the series is the wave (or hole effect) variogram (Cressie 1993) given by:

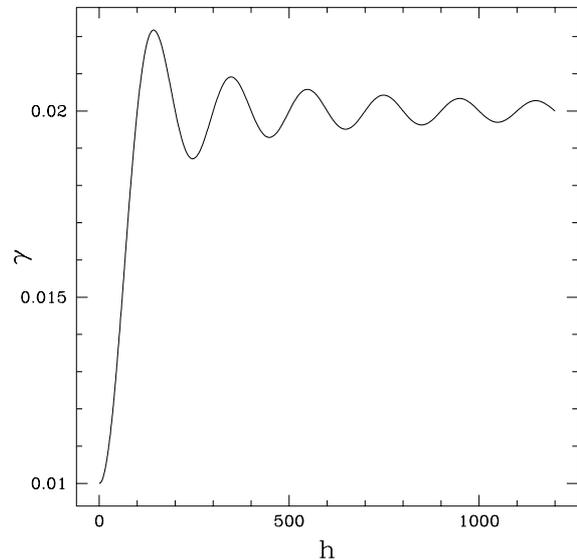
$$\gamma(h, a, b, c) = \begin{cases} 0 & \text{if } h = 0, \\ a + b\left(1 - \frac{c}{h} \sin\left(\frac{h}{c}\right)\right) & \text{otherwise,} \end{cases}$$

where  $a \geq 0, b \geq 0$  and  $c \geq 0$ . An example of a wave variogram is shown in Fig. 1. A hole effect shows oscillations of decreasing amplitude around the plateau called the sill,  $a + b = \sigma_{\text{signal}}^2$ , which is the total variance of the time series. The oscillation reflects periodicity in the data and the first minimum is considered as a period or pseudo-period of the time series. The parameter  $a$  is usually called nugget effect and represents the micro-scale variability  $\sigma_{\text{noise}}^2$  due to measurement noise.

Note that a strictly periodic signal would show a periodic variogram defined by Journel & Huijbregts (1978):

$$\gamma(h, a, b, c) = \begin{cases} 0 & \text{if } h = 0, \\ a + b\left(1 - \cos\left(\frac{h}{c}\right)\right) & \text{otherwise,} \end{cases}$$

where  $a \geq 0, b \geq 0$  and  $c \geq 0$ . However, in practice, the oscillations are often dampened.



**Fig. 1.** Wave variogram for a pseudo-periodic signal

*2.2. The Matheron variogram estimator*

Consider the series of differences at lag  $h$  defined by  $v(h) = m(t+h) - m(t)$ . Let  $\{m(t_1), \dots, m(t_n)\}$  be a sample of a time series, and  $\{v_1(h), \dots, v_{N_h}(h)\}$  be the corresponding sample of differences at lag  $h$ , where  $N(h) = \{(t_i, t_j): t_i - t_j = h\}$  and  $N_h$  is the cardinality of  $N(h)$ . The classical variogram estimator proposed by Matheron (1962), based on the method-of-moments, is

$$2\hat{\gamma}_{L^2}(h) = \frac{1}{N_h} \sum_{i=1}^{N_h} (v_i(h) - \bar{v}(h))^2, \quad h \in \mathbb{R}_+,$$

where  $\bar{v}(h) = \frac{1}{N_h} \sum_{i=1}^{N_h} v_i(h)$ . This estimator is based on the  $L^2$  scale estimator (Genton & Rousseeuw 1995), which is unbiased, but behaves poorly if there are outliers in the data. One single outlier can destroy this estimator completely. For that reason, Genton (1998) proposed a highly robust variogram estimator.

*2.3. The highly robust variogram estimator*

In the context of statistical analysis of periodicity, variogram estimation is a crucial stage, because it determines the period or pseudo-period. Therefore, it is important to have a variogram estimator which remains close to the true underlying variogram, even if outliers (faulty observations) are present in the data. Experience from a broad spectrum of applied sciences shows that measured data may contain between 10 to 15 percent of outlying values (Hampel 1973) due to gross errors, measurement mistakes, faulty recording, etc. One might argue that any reasonable exploratory data analysis would identify outliers in the data. However, this approach may be subjective

and informal. Furthermore, the existence of exploratory techniques does not supersede the usefulness of robust techniques. In this paper, we advocate the use of estimators which take account of all the available information in the data.

In the context of scale estimation, Rousseeuw & Croux (1992, 1993) have proposed a simple, explicit and highly robust estimator, called  $Q_{N_h}$ , which is defined by

$$Q_{N_h} = 2.2191 \left\{ |v_i(h) - v_j(h)|; i < j \right\}_{(k)},$$

where the factor 2.2191 is for consistency with the Gaussian distribution,  $k = \binom{[N_h/2]+1}{2} = ([N_h/2] + 1)[N_h/2]/2$ , and  $[N_h/2]$  denotes the integer part of  $N_h/2$ . This means that we sort the set of all absolute differences  $|v_i(h) - v_j(h)|$  for  $i < j$  and then compute its  $k$ -th order statistic (the first quartile for large  $N_h$ ). This value is multiplied by the factor 2.2191, thus yielding  $Q_{N_h}$ . Note that this estimator computes the  $k$ -th order statistic of the  $\binom{N_h}{2}$  interpoint distances. This estimator is shown to be very robust to outliers in the data. It is of interest to remark that  $Q_{N_h}$  does not rely on any location knowledge and is said to be location-free. The estimator  $Q_{N_h}$  can be computed using no more than  $O(N_h \log N_h)$  time and  $O(N_h)$  storage, by means of the fast algorithm described in Croux & Rousseeuw (1992).

Using the previous definitions, Genton (1998) defines a highly robust variogram estimator as

$$2\hat{\gamma}_Q(h) = (Q_{N_h})^2,$$

and discusses its use and properties.

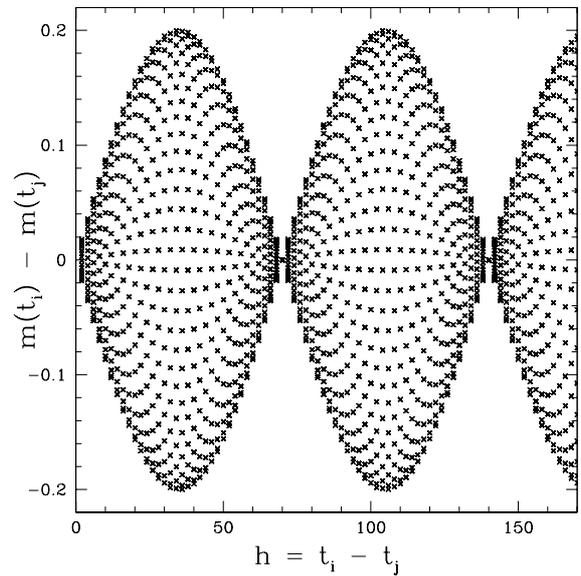
#### 2.4. Variogram estimation for irregularly spaced data

Structure functions are used in AGN studies (Hughes et al. 1992). Apparently this goes well with the signals they want to study, because they are mainly interested in the slope of the rising curve of the structure function which gives information in the frequency domain (Paltani 1996). However for periodic signals, the oscillations may be diminished in their analysis. We are left with the difficulty of the choice of the binning and of the width of the bins. A too narrow width would have the disastrous effect to have too little measurements in a bin, whereas a constant binning can be too large for short periods and smears out the periodic behaviour. This is the reason why we choose to have a binning linked with the period investigated. Denote by

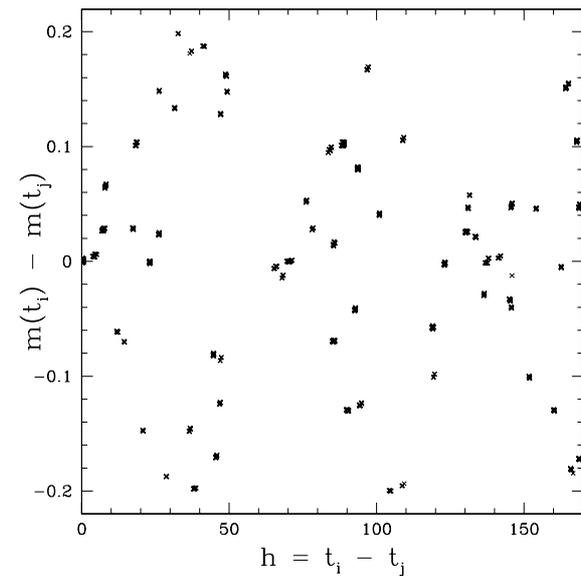
$$2\gamma_b(h) = \frac{1}{2f(h)} \int_{h-f(h)}^{h+f(h)} 2\gamma(u) du,$$

the binned variogram based on a lag function  $f$ . It is natural to require for a time-scale invariant variogram, i.e.  $2\gamma_1(h) = 2\gamma_2(ah)$ ,  $a > 0$ , that the binned variogram be invariant too. This is achieved by choosing the linear lag function  $f(h) = \delta h$ , as is shown by

$$2\gamma_{b2}(ah) = \frac{1}{2\delta ah} \int_{ah-\delta ah}^{ah+\delta ah} 2\gamma_2(u) du$$



(a)

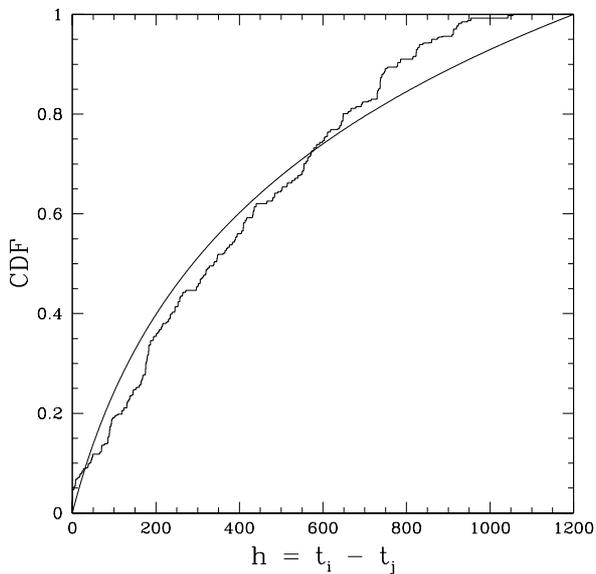


(b)

**Fig. 2.** Diagram of pairwise differences. **a)** regular time sampling, **b)** irregular sampling for the star HIP 111771

$$\begin{aligned} &= \frac{1}{2\delta ah} \int_{h-\delta h}^{h+\delta h} 2\gamma_2(av)adv \\ &= \frac{1}{2\delta h} \int_{h-\delta h}^{h+\delta h} 2\gamma_1(v)dv \\ &= 2\gamma_{b1}(h). \end{aligned}$$

One has to determine the fraction  $\delta$ , in order to get an optimal binning of the variogram. If data would be regularly



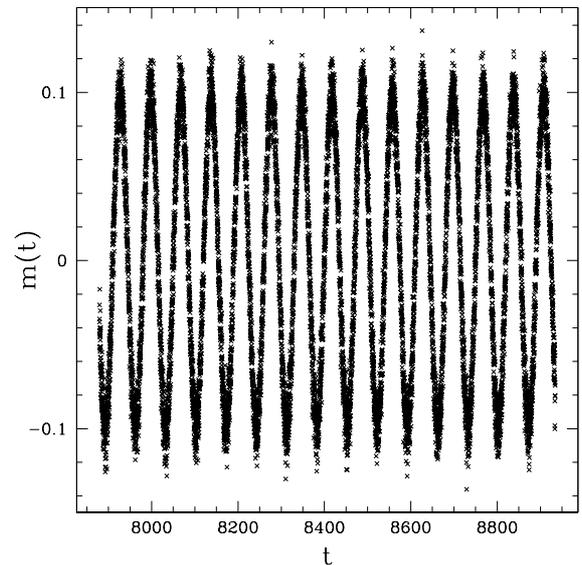
**Fig. 3.** Empirical cumulative distribution function (CDF) of the lag differences for the star HIP 111771, and logarithmic approximation  $F(h)$  (smooth line) of the empirical CDF

spaced, one could let  $\delta \rightarrow 0$ , and use the classical method  $2\gamma_b(h) = 2\gamma(h)$ . Figure 2 depicts the diagram of paired differences for regularly and irregularly (HIP 111771 time sampling) spaced data. A closer look at Fig. 2 indicates that if the period is  $h$ , we should set  $\delta < 1/2$ , otherwise too much information about the period is lost. Therefore, a good compromise between regularity and period information is to let  $\delta = 1/4$ .

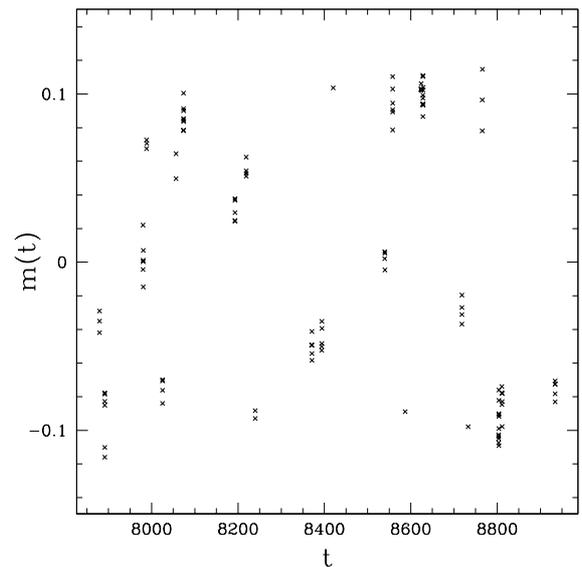
The remaining problem is the choice of the lags, i.e. the position of the bin location. If one chooses equidistant lags, there would be too many bins for large lags, and too few for small lags. Therefore, the distribution of the lags is a crucial question. Figure 3 represents the empirical cumulative distribution function (CDF) of the lags of the star HIP 111771, and indicates that a logarithmic scale of the lags

$$F(h) = \log\left(\frac{9}{1200}h + 1\right)$$

is a satisfactory approximation. Other stars have also been analysed and the same approximation can be used for most of them. Note that the greatest available lag is  $h = 1200$ , and we set  $F(0) = 0$  and  $F(1200) = 1$ . In consequence, we adopt the choice of lags  $h_0, \dots, h_k$  such that  $F(h_0), \dots, F(h_k)$  are equidistant on the interval  $[0, 1]$ . The number  $k$  of lags must be chosen such that a good temporal resolution is achieved (for instance,  $k = 90$  is a typical value with Hipparcos time sampling).



(a)



(b)

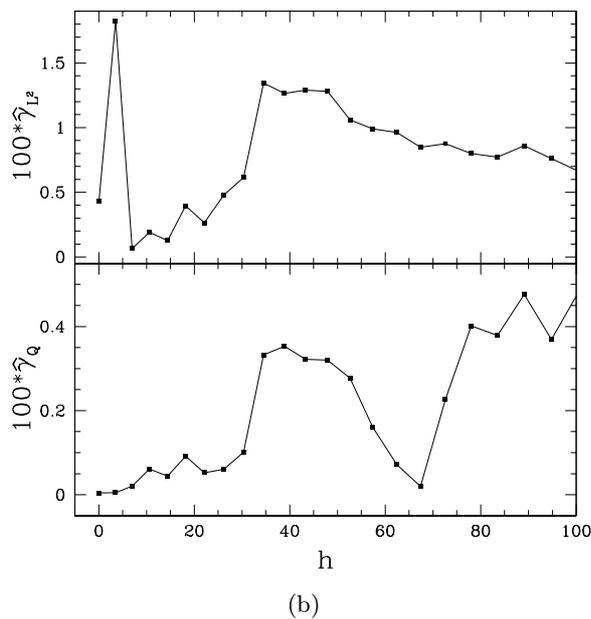
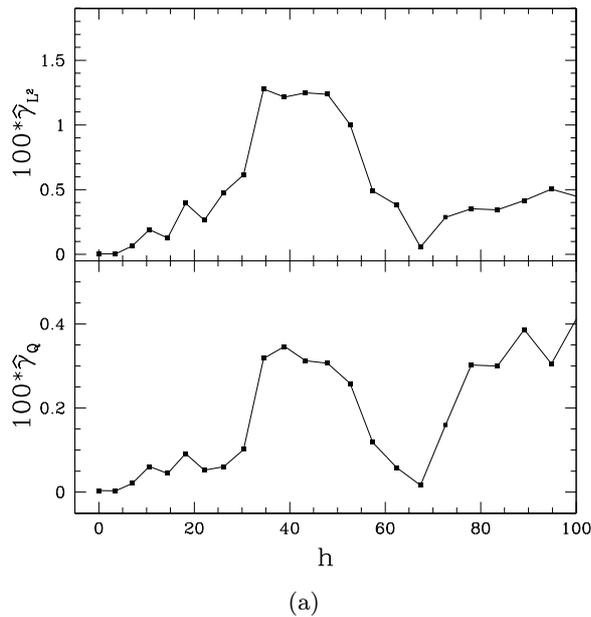
**Fig. 4.** Simulated sinusoidal process with noise. **a)** regular time sampling, **b)** irregular time sampling for the star HIP 111771

Another approach in the choice of the lags consists in computing quantiles of the empirical distribution of each star. The advantage of this method is its adaptability to the particular sampling of each star, whereas the former one was global. Moreover, lags are defined only when new data points are present. The drawback is that it provides clumped bin locations.

### 3. Simulations

In order to point out the behaviour and robustness properties of the  $\hat{\gamma}_{L^2}$  and  $\hat{\gamma}_Q$  variogram estimators, we simulate a sinusoidal process

$$m(t_i) = 0.1 \sin\left(\frac{2\pi}{70}t_i\right) + \varepsilon_i,$$



**Fig. 5.** Estimated variograms (above  $\hat{\gamma}_{L^2}$ , below  $\hat{\gamma}_Q$ ) of the simulated sinusoidal process. **a)** without outlier, **b)** with one outlier

where  $\varepsilon_i$  is the noise, independently and identically distributed according to a Gaussian law  $N(0, \sigma_{\text{noise}}^2)$  with  $\sigma_{\text{noise}} = 0.01$ . The amplitude of the signal is set to 0.10 and the period to 70 days, a possible period for semi-regular red giant stars (Jorissen et al. 1997). This process is visualized in Fig. 4a. As we would like to come to a more realistic case, we generate the same behaviour but with the time sampling  $t_i$  taken from the star HIP 111771, which is taken as representative of the mission (although the sample is rather heterogeneous). This data set is shown in Fig. 4b and is used to compute the variogram with  $\hat{\gamma}_{L^2}$  (Fig. 5a above) and  $\hat{\gamma}_Q$  (Fig. 5a below) respectively. Both estimators are able to detect the period around 70 days, given by the first significant minimum.

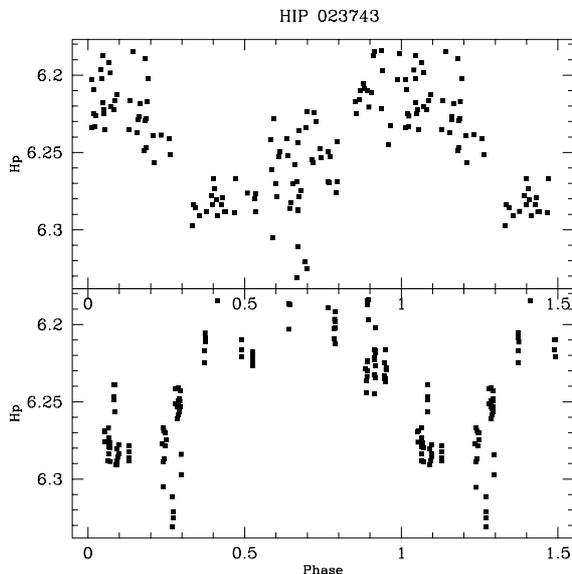
In order to show conspicuously the differences between the two estimators when the signal is perturbed by outliers, we took the previous simulated data and changed only one value. We put it at  $5\sigma_{\text{signal}}$  from the mean. Actually, this value can sometimes occur in real data from Hipparcos. The effect of the substitution of that single value can be seen in Fig. 5b, which should be compared with the Fig. 5a. Several remarks can be made. First, the  $\hat{\gamma}_{L^2}$  estimator shows a flat and slight declining curve around the period: the signature of the periodicity has totally disappeared. In comparison, the general behaviour of  $\hat{\gamma}_Q$  has not changed. Second, the estimation of the micro-scale variability  $\sigma_{\text{noise}}^2$  jumped with  $\hat{\gamma}_{L^2}$ . Furthermore, there is a higher jump at the second lag, which would suggest that there is some variability in the signal for extremely short time-scales. Again  $\hat{\gamma}_Q$  stays unchanged. With this example we see that  $\hat{\gamma}_{L^2}$  is not reliable when outlying values are present in the data, and that  $\hat{\gamma}_Q$  is more invariant to such values, i.e.  $\hat{\gamma}_Q$  is a highly robust estimate of the variogram.

### 4. Application to Hipparcos

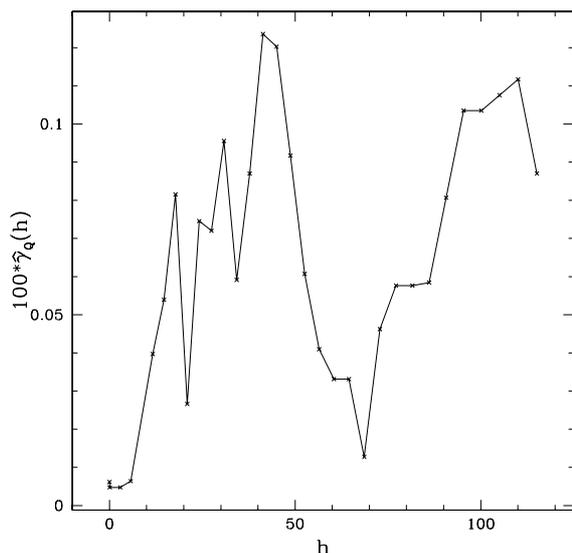
#### 4.1. Detection of spurious periods

When using usual period search methods, the well known problem of aliasing can produce spurious periods. On the basis of the phase diagram only, these periods can not be rejected. However, the variogram can be used as a complementary method. It can be very helpful to rule out certain periods, or to increase the level of confidence in others. The accurate determination of the period is left to the usual period search algorithm.

Here we present an example of such a situation. The star HIP 023743 has in its power spectrum two possible periods:  $p_1 = 3.348$  days and  $p_2 = 80.9$  days. The associated phase diagrams are presented for these two possibilities in Fig. 6. Both seem to be reasonable. However, with the variogram, the short period can be eliminated. As we see in Fig. 7, there is no evidence of variability signature for the periods shorter than 3 days which could explain



**Fig. 6.** Phase diagram for the star HIP 023743. Above the data is folded with the period  $p_1 = 3.348$  days. Below, with the period  $p_2 = 80.9$  days

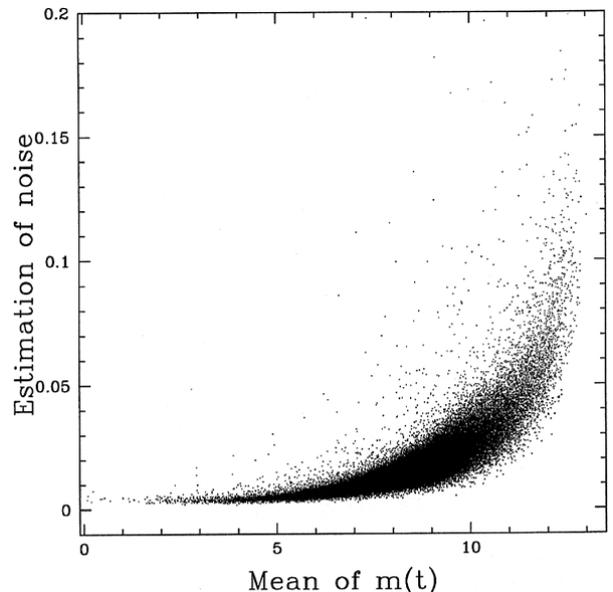


**Fig. 7.** Estimated variogram  $\hat{\gamma}_Q$  of the star HIP 023743

the dispersion of the data. Figure 6 shows that the phase diagrams both have the same amplitude of approximately 0.07. Figure 7 does not reveal any such amplitude around  $p_1 = 3.348$ , whereas it does around  $p_2 = 80.9$ . This case is also instructive because the minimum of the variogram does not exactly coincide with the period  $p_2 = 80.9$  days. This is due to the irregularity and scarcity of the pairwise differences near  $h = 80$ .

#### 4.2. Estimation of measurement noise

The precision on the Hipparcos magnitudes is strongly variable. It can be interesting to know its dependence



**Fig. 8.** Estimated measurement noise by short-time differences using  $\hat{\gamma}_Q$

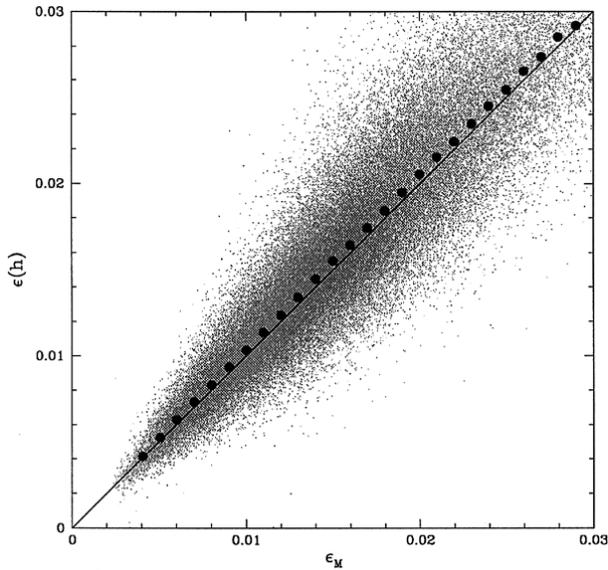
with respect to the magnitude. As already said, the nugget value can be viewed as a good estimator of the noise. Only the very short pairwise time differences are used ( $h = 20$  minutes or  $h = 108$  minutes). The underlying assumption is that very few variable stars have detectable variability at lags around  $h = 20$  minutes.

In fact the law (cf. Fig. 8) is very close to the one obtained by Eyer & Grenon (1997). For a given magnitude, the dispersion can be high and the degradation is very rapid (exponential). Furthermore there are some outliers which are caused by some problems of the satellite like light pollution from companions, mispointing, or by rapid oscillations of stars.

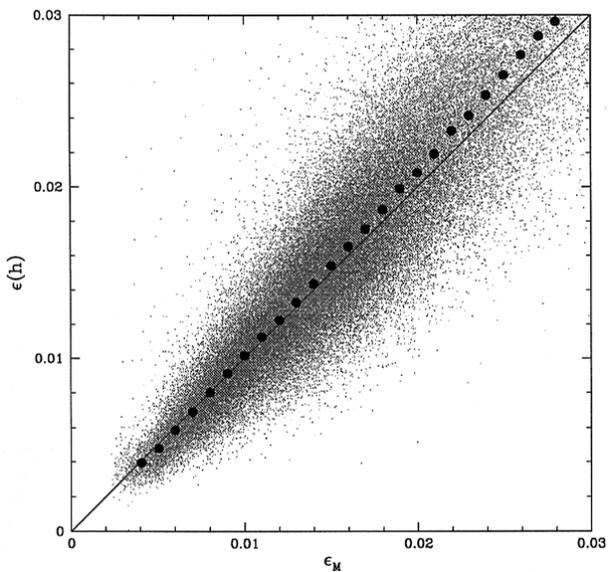
Based on Fig. 8, a set of potentially interesting stars was selected on which a high frequency analysis has been performed. As a result we found three periodic variable stars (HIP 071119, HIP 044025 and HIP 029055) which are not in the published catalogues (Grenon et al. 1997 and van Leeuwen et al. 1997).

The method permits also to compare our estimated noise with the quoted errors  $\epsilon_i$ , furnished in the Hipparcos photometric data base. For a given star we define a mean error by  $\epsilon_M = \sqrt{(\sum \epsilon_i^2/n)}$ , and a highly robust noise estimator  $\epsilon(h) = \sqrt{\hat{\gamma}_Q(h)}$ , with  $h = 20$  min. In order to confirm the adequacy of the method, a simulation is carried out on the 118204 stars. The magnitudes  $m(t_i)$  are drawn from a Gaussian distribution with constant mean and variance  $\epsilon_i^2$ . Then, the estimator  $\epsilon(h)$  is calculated on the simulated  $m(t_i)$  for each star.

We compare the two estimators  $\epsilon_M$  and  $\epsilon(h)$  in Fig. 9 for the simulated data; a moving median is computed and is represented by the filled circles, indicating a reasonable match between the two noise estimators. In Fig. 10, the estimated noise of the Hipparcos real data is

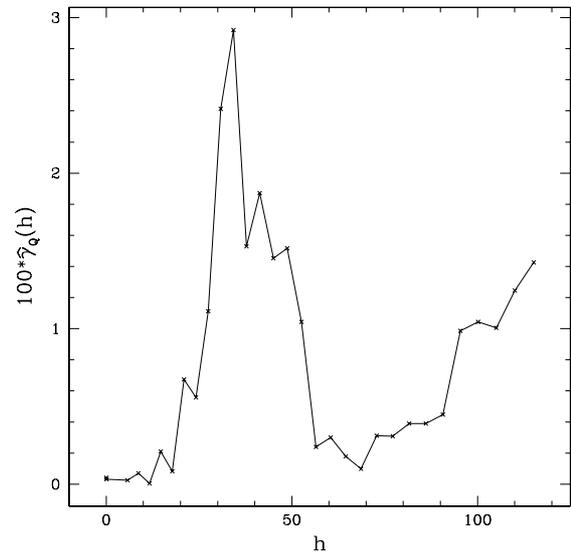


**Fig. 9.** Simulated measurement noise using the highly robust noise estimator  $\epsilon(h = 20)$  versus the mean error  $\epsilon_M$ . The grey data cloud represents the 118204 stars of Hipparcos, and the black filled circles denote the moving median. The graph shows that  $\epsilon_M$  and  $\epsilon(h = 20)$  are reasonable estimators for the noise

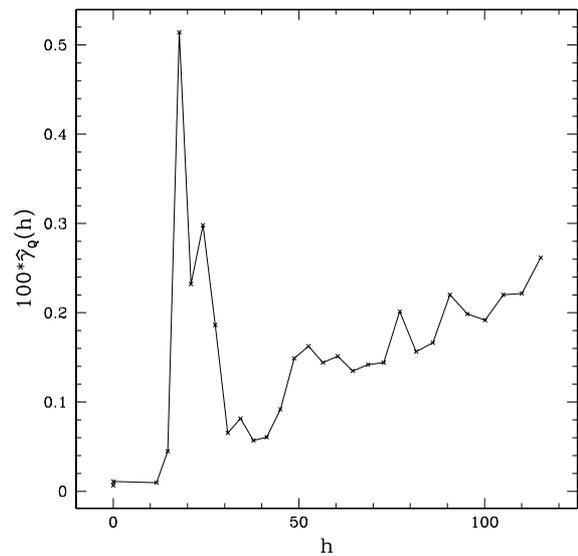


**Fig. 10.** Estimated measurement noise of the Hipparcos data using the highly robust noise estimator  $\epsilon(h = 20)$  versus the mean error  $\epsilon_M$ . The grey data cloud represents the 118204 stars of Hipparcos, and the black filled circles denote the moving median. The graph shows a bias of 9% for the slope

represented following the same method. The slope of the moving medians is 1.09 indicating a 9% bias probably due to  $\epsilon_i$ . It seems that for small  $\epsilon_M$  the  $\epsilon_i$  can be overestimated whereas for large  $\epsilon_M$  the  $\epsilon_i$  are underestimated in the Hipparcos data base. As a consequence, when doing a photometric variability analysis, some bright variable



**Fig. 11.** Estimated variogram  $\hat{\gamma}_Q$  of the star HIP 111771

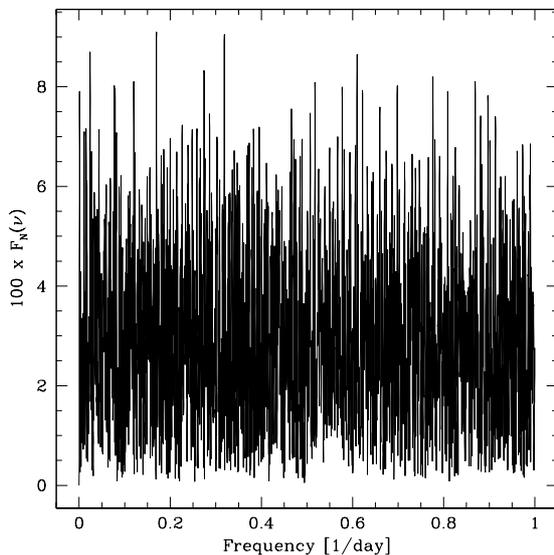


**Fig. 12.** Estimated variogram  $\hat{\gamma}_Q$  of the star HIP 060998

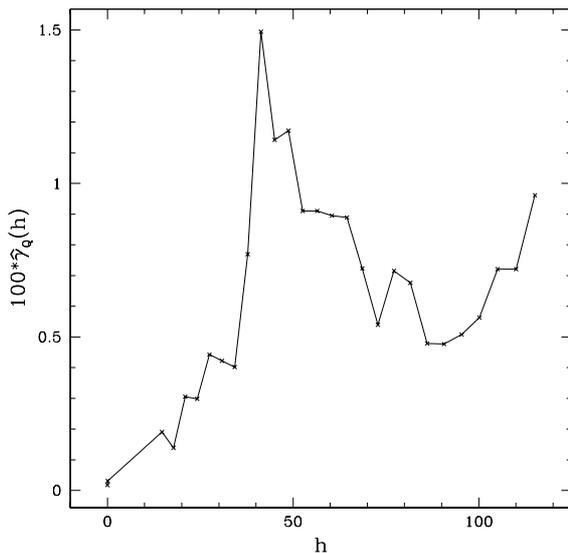
stars might be taken as constant, and some constant faint stars might be taken as variable.

#### 4.3. Example of three stars from the Main Mission

Three examples illustrate the use of the wave variogram. We have a sequence of giant stars, located in the same region of the HR diagram, but with different properties: HIP 111771 is the hottest star (M2III type) with a pseudo-period of approximately 70 days (cf. Fig. 11), and an amplitude (peak to peak) of 0.61 Hp mag; HIP 060998 (M4III type) has a small amplitude of 0.23 Hp mag, and a shorter period of about 35 days (cf. Fig. 12). For a comparison, the Fourier power spectrum of the latter star is presented



**Fig. 13.** Fourier power spectrum of the star HIP 060998. Although a characteristic time-scale can be given in the variogram (Fig. 12), the power spectrum is noisy



**Fig. 14.** Estimated variogram  $\hat{\gamma}_Q$  of the star HIP 052507

in Fig. 13, no clear information can be deduced from it. HIP 052507 (M5III type) has a very long time-scale variation (clearly seen in the light curve), with a shorter one superimposed. The latter is revealed by the variogram in Fig. 14, with a period of about 85 days, and an amplitude of 0.53 Hp mag.

## 5. Conclusion

In this paper, periodic structures in variable star light curves from the Hipparcos mission were studied with robust variogram estimators. Several problems were

investigated, such as bin positions, bin widths and robustness properties. Some specific applications to the Hipparcos mission were discussed, such as spurious periods detections and measurement noise. The method can be successful for detecting time-scales of semi-regular stars of the Hipparcos mission. Although bin positions are specific to the Hipparcos mission, the whole method can be easily applied to other irregular time series.

*Acknowledgements.* We would like to thank S. Paltani, M. Grenon and F. Mignard for their interesting discussions and comments. We also thank the referee X. Luri and D. Barthès for their remarks. The first author is grateful to the Swiss Society of Astronomy and Astrophysics for its financial support. The second author is grateful to the Swiss Federal Institute of Technology of Lausanne where the main part of this work was done.

## References

- Babu G.J., Feigelson E.D., 1996, *Astrostatistics*. Chapman & Hall
- Bartholdi P., 1988, *Comptes Rendus des Journées de Strasbourg* 10, 77
- Cressie N., 1993, *Statistics for Spatial Data*. Wiley
- Croux C., Rousseeuw P.J., 1992, *Computational Statistics* 1, 411
- Cuyper J., 1997, in *Application of Time Series Analysis in Astronomy and Meteorology*, Rao et al. (eds.). Chapman & Hall
- Deeming T.J., 1975, *Astrophys. Space Sci.* 36, 137
- Edelson R.A., Krolik J.H., 1988, *ApJ* 327, 693
- Eyer L., Grenon M., 1997, *HIPPARCOS Venice '97*, ESA SP-402, 467
- Eyer L., 1998, PhD Thesis, Geneva Observatory
- Ferraz-Mello S., 1981, *AJ* 86, 619
- Foster G., 1996, *AJ* 112, 1709
- Genton M.G., Rousseeuw P.J., 1995, *J. Comp. App. Math.* 64, 69-80
- Genton M.G., 1998, *Math. Geology* 30, 213
- Grenon M., et al., 1997, *ESA SP-1200*, Vol. 12
- Hampel F.R., 1973, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 27, 87
- Hughes P.A., Aller H.D., Aller M.F., 1992, *ApJ* 396, 469
- Jorissen A., Mowlavi N., Sterken C., Manfroid J., 1997, *A&A* 324, 578
- Journel A.G., Huijbregts C.J., 1978, *Mining Geostatistics*. Academic Press
- de Laverny P., Mennessier M.O., Mignard F., Mattei J.A., 1998, *A&A* 330, 169
- Matheron G., 1962, *Traité de Géostatistique Appliquée*, Tome I, Paris
- Paltani S., 1996, PhD Thesis, Geneva Observatory
- Renson P., 1978, *A&A* 63, 125
- Rousseeuw P.J., 1987, *Statistical Data Analysis Based on the  $L_1$ -Norm and Related Methods*. Elsevier Science Publishers B.V., North-Holland
- Rousseeuw P.J., Croux C., 1992,  *$L_1$  Statistical Analyses*, 77
- Rousseeuw P.J., Croux C., 1993, *J. Am. Stat. Assoc.* 88, 1273
- Schwarzenberg-Czerny A., 1989, *MNRAS* 241, 153
- van Leeuwen F., et al., 1997, *ESA SP-1200*, Vol. 11
- van Leeuwen F., 1997, *HIPPARCOS Venice '97*, ESA SP-402, 19