

# Searching for periodicity in weighted time point series

L. Jetsu<sup>1,2,\*</sup> and J. Pelt<sup>3</sup>

<sup>1</sup> Observatory, P.O. Box 14, SF-00014 University of Helsinki, Finland

<sup>2</sup> NORDITA, Blegdamsvej 17, 2100 Copenhagen, Denmark

<sup>3</sup> Tartu Astrophysical Observatory, Tõravere 1–6, Tartu 202444, Estonia

Received July 17, 1995; accepted January 20, 1996

**Abstract.** — Consistent statistics for two methods of searching for periodicity in a series of weighted time points are formulated. An approach based on the bootstrap method to estimate the accuracy of detected periodicity is presented.

**Key words:** methods: statistical; analytical

## 1. Introduction

A random sample of single measurements representing directions in a plane or phases at different time values folded with a fixed period is called circular data. The excellent review by Batschelet (1981: especially Sect. 4.) of the eleven most frequently applied methods for studying such data is adequate, yet not too detailed, and contains the most important references related to the subject. A more detailed formulation of several of the above eleven, as well as other, methods can be found in, e.g., Mardia (1972).

For example, the phases of the photon arrival times from a radio pulsar with an unknown period are circular data, where information on the shape of the light curve and the period are both lacking. Thus an ideal method for searching for periodicity should be nonparametric (i.e. model independent). The phase distribution of data with the correct period may consist of one or several concentrations. Batschelet (1981) concludes that the following four methods are sensitive to both uni- and multimodal distributions: the Rao's spacing (Rao 1969, 1976), the Chi-squared (e.g. Rao 1972), the Kuiper (Kuiper 1960), and the Watson test (Watson 1961). Examples of circular data analysis in astronomy can be found in Protheroe et al. (1984), Bai (1987, 1988, 1990) and Swanepoel & De Beer (1990). The test by Swanepoel & De Beer (1990) is especially sensitive to multi-modal distributions.

But there are less examples of how to utilize the additional information that may be available in the form of known weights (or errors) associated with the circular data. We managed to find only one such case: the weighted modification of the Rayleigh test by Moore (1980). In this

paper we will present the weighted modifications of two nonparametric methods for studying circular data (Kuiper 1960; Swanepoel & De Beer 1990), which are subsequently applied to solar flare data in Jetsu et al. (1995, hereafter Paper II). We chose to formulate the weighted versions of these particular methods, because they are sensitive both to uni- and multimodal distributions, and the consistency of the statistics of these modified weighted versions could be verified.

## 2. Extension of the SD-method

A modified version of the method developed by Swanepoel & De Beer (1990: hereafter the SD-method) is presented in this section to search for periodicity in weighted time point series. The SD-method is first briefly discussed in Sect. 2.1. The main modification is the introduction of weights and the subsequent definition of remapped non-integer indices (Eqs. (8) and (9)). The modified version introduced in Sect. 2.2. is hereafter abbreviated as the WSD-method. Finally, the statistics for testing a number of periods with the WSD-method are discussed in Sect. 2.3.

### 2.1. The SD-method

Suppose we seek any periodicity present in a series of  $n$  time points  $t_1, \dots, t_n$ . The phase of each  $t_i$  for an arbitrarily chosen period ( $P$ ) and a zero-point in time ( $t_0$ ) is

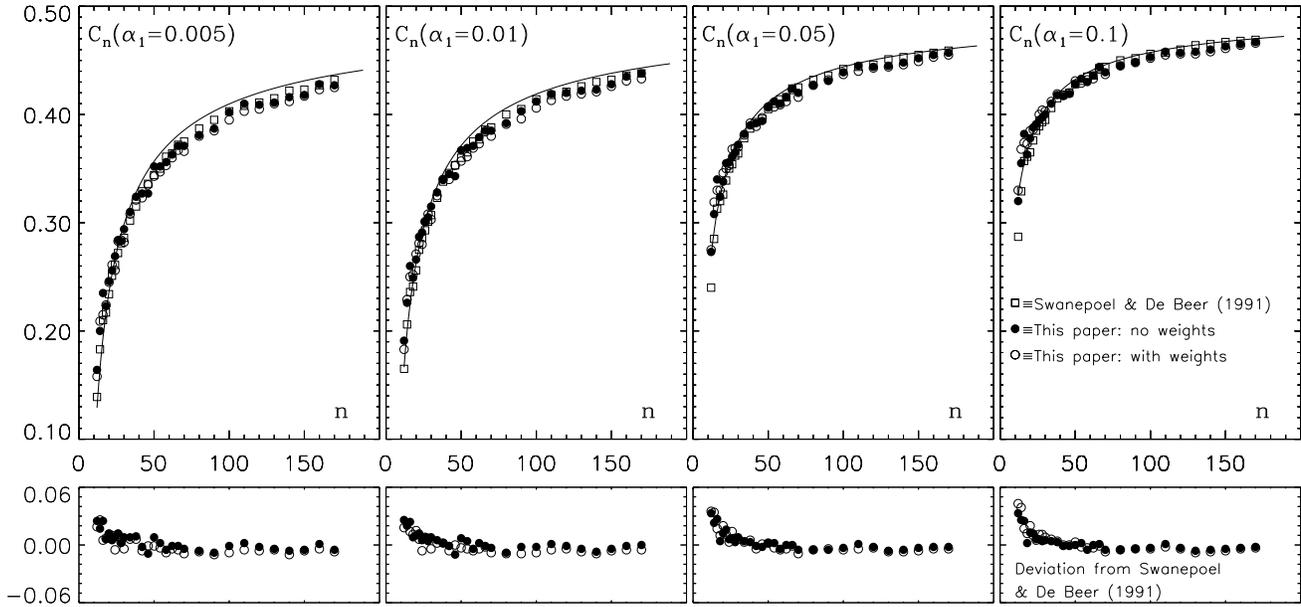
$$\phi_{i,P} = \text{FRAC} [(t_i - t_0)P^{-1}] \quad (1)$$

where the notation FRAC means that the integer part of  $\phi_{i,P}$  is removed. The phase differences are

$$\phi_{i,j,P} = |\text{FRAC} [(t_i - t_j)P^{-1}]| = |\phi_{i,P} - \phi_{j,P}|, \quad (2)$$

Send offprint requests to: L. Jetsu

\*E-mail address: jetsu@nordita.dk



**Fig. 1.** Monte Carlo simulations of  $C_n$  at four preassigned significance levels ( $\alpha_1 = 0.005, 0.01, 0.05$  and  $0.1$ ): open squares (Swanepoel & De Beer 1990, their Table 1), closed circles (WSD–method with equal weights), and open circles (WSD–method with unequal weights). The continuous line outlines the estimate of Eq. (16) with  $m = 1$ . The lower figures show the differences between the simulated values of  $C_n(\alpha_1)$  for the WSD–method minus those tabulated by Swanepoel & De Beer (1990)

where  $i = 1, \dots, n-1$  and  $j = i+1, \dots, n$ , i.e. the total number of phase differences is  $N = n(n-1)/2$ . It is this definition of  $\phi_{i,j,P}$ , which ensures that the SD–method is independent of the choice of the zero–point  $t_0$ . Note that the absolute value means here that all  $\phi_{i,j,P}$  greater than  $0.5$  must be converted to  $1 - \phi_{i,j,P}$ .

The next step is performed by arranging the phase differences  $\phi_{i,j,P}$  into ascending (i.e. rank) order and denoting them by  $V_1, \dots, V_N$ . The  $i$ :th smallest component is  $V_i$ . Two additional definitions are necessary:

$$\beta_n = 2^{\frac{1}{3}} n^{\frac{2}{3}} \quad (3)$$

$$a_n = \text{INT} [N\beta_n^{-1}] = \text{INT} \left[ \frac{n^{\frac{1}{3}}(n-1)}{2^{\frac{4}{3}}} \right], \quad (4)$$

where INT denotes the integer part of the value within the brackets.

The statistical “null hypothesis” ( $H_0$ ) is: “No periodicity is present in the time points  $t_i$  with the period  $P$ .” This  $H_0$  can be rejected only if the criterion

$$V_{a_n} \leq C_n(\alpha)\beta_n^{-1} \quad (5)$$

is fulfilled, where the critical parameter  $C_n$  is a function of the probability  $\alpha$  of falsely rejecting  $H_0$  when it is actually true. In other words, if this criterion is fulfilled for a given value of  $C_n(\alpha)$ , the probability of periodic variation is  $1 - \alpha$ . Thus  $V_{a_n}$  is the “test statistic” of the SD–method, and the periodogram is  $V_{a_n}$  derived as a function of  $P$ . A potential period candidate minimizes the  $a_n$  smallest  $V_i$ ,

while those with  $i > a_n$  have no effect to the test statistic. Swanepoel & De Beer (1990) applied Monte Carlo simulations to derive the critical value  $C_n(\alpha)$  for different numbers of time points ( $n$ ) and probabilities ( $\alpha$ ).

We will briefly discuss the reasons for the definitions of Eqs. (3) and (4) that were not given by Swanepoel & De Beer (1990). The total number of phase differences approaches  $2^{-1}n^2$  for larger samples, and combining Eqs. (3) and (4) gives  $a_n = (n/2)^{4/3}$ . The ratio  $a_n/n$  exceeds unity and approaches  $2^{-4/3}n^{1/3}$ . Hence the relations

$$n < a_n \propto n^{4/3} \quad (6)$$

$$a_n \ll N \propto n^2 \quad (7)$$

are valid for larger samples. The inequality of Eq. (6) is interesting, because only  $n-1$  independent  $V_i$  do in fact exist. For example, the phase difference  $\phi_{2,3,P}$  is already determined by  $\phi_{1,2,P}$  and  $\phi_{1,3,P}$ . Any period that minimizes the latter two will inevitably also minimize the first. The inequality of Eq. (6) implies that the increasing ratio  $a_n/n$  introduces more correlation into the test as the samples increase, but this correlation is balanced by the more rapid increase of the inequality of Eq. (7). The method is “free” of correlations only when  $a_n \leq n-1$ , and we conclude that the definitions of Eqs. (3) and (4) by Swanepoel & De Beer (1990) are a compromise between studying an adequate number of phase differences without introducing too much correlation into the test.

## 2.2. The WSD–method

One purpose of this study is to formulate a consistent weighted version of the SD–method, where the additional information of the known weight ( $w_i$ ) of each  $t_i$  is utilized. This is achieved by adopting weights  $w_i = \sigma_i^{-2}$ , where  $\sigma_i$  is the error connected to  $t_i$ . Hence the weight assigned for each  $\phi_{i,j,P}$  is

$$w_{i,j} = w_i w_j (w_i + w_j)^{-1}, \quad (8)$$

using the relation  $\sigma_{i,j}^2 = \sigma_i^2 + \sigma_j^2$ .

When arranging the  $V_i$  in rank order, the order of the weights, now denoted as  $W_1, \dots, W_N$ , is arranged in such a way that  $w_{i,j}$  remains the weight of  $\phi_{i,j,P}$ . The weights are then used to derive a new set of indices with a remapping function

$$j(i) = N \left[ \sum_{k=1}^i W_k \right] \left[ \sum_{k=1}^N W_k \right]^{-1}. \quad (9)$$

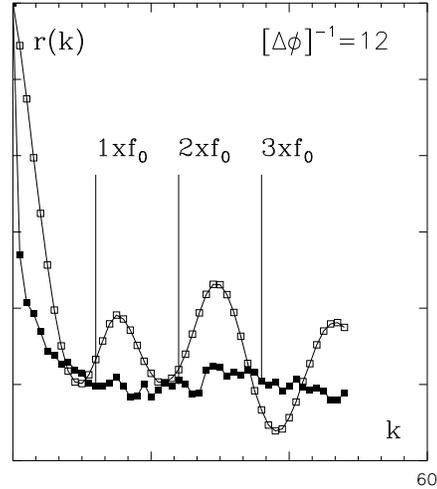
Note that this remapping does not change the number of  $V_i$  and  $W_i$ , nor the rank order because  $j(i) < j(i+1)$ . The definition of Eq. (9) ensures that the remapped index of the largest phase difference is  $j(N) = N$ , while the other indices  $j(i)$  are not necessarily integers. The value of the test statistic  $V_{a_n}$  is determined by these remapped indices. The  $a_n$ :th smallest value of  $V_i$  is the one with the particular remapped index

$$\max\{j(i)\} \leq a_n. \quad (10)$$

Monte Carlo simulations were applied to confirm that the statistics of the WSD– and SD–methods are identical. Random samples of  $n$  number of uniformly distributed  $\phi_{i,P}$  between 0 and 1 were first derived. The set of random weights ( $w_i$ ) was the same for every simulation with a fixed  $n$ . However, these weights were assigned in random order for each  $\phi_{i,P}$  during every simulation. The selected range of the weights  $w_i$  was between 3 and 17, i.e. the same as for the data to be analysed in Paper II. Then the corresponding  $N$  number of  $V_i$  and  $W_i$  were derived. Each simulation gave one estimate of  $C_n$ . The results of these simulations of  $C_n$  for the WSD–method with these unequal weights are shown in Fig. 1 (open circles). The statistics of the WSD–method seem to be the same as for the SD–method (Fig. 1: open squares). As a final check, we performed Monte Carlo simulations with equal weights (Fig. 1: closed circles). Although trivial, we note that this corresponds to a case with “no weights”, because in this case the WSD–method is identical with the SD–method due to the fact the remapping of Eq. (9) gives  $j(i) = i$  for all  $V_i$ . All simulations yield compatible results, except for a weak tendency of our values for  $n \leq 20$  to exceed those tabulated by Swanepoel & De Beer (1990).

The reason for the similarity of the statistics of the WSD– and SD–methods deserves a short discussion. The relation of Eq. (7) shows that only a minor fraction of all

phase differences, i.e. only those  $V_i$  with  $j(i) \leq a_n$ , contribute to the periodogram with an arbitrary trial period. The main reason for the similarity of the statistics of the WSD– and SD–methods is that the same applies to the weights. Any particular weight may, or may not, influence the periodogram.



**Fig. 2.** The variations of the “empirical” parameter  $r(k)$  of Eq. (14) as determined for the periodograms of Figs. 1a and 4a of Paper II are shown as closed and open squares, respectively. The “theoretical” positions of integer multiplets of  $f_0$  predicted by Eq. (12) are shown by vertical lines

## 2.3. Statistics of WSD–method for many tested periods

The statistics of the WSD–method presented in Sect. 2.2. are valid for *one* arbitrarily chosen period. This section is devoted to the case of many tested periods. Let us assume that a period interval between  $P_{\min}$  and  $P_{\max}$  is tested. Since we choose to perform the test in the frequency domain, the corresponding frequency limits are  $f_{\min} = P_{\max}^{-1}$  and  $f_{\max} = P_{\min}^{-1}$ . If the observing time interval is  $\Delta T = t_n - t_1$ , a suitable step for the tested frequencies  $f$  is

$$\Delta f = [GM\Delta T]^{-1}, \quad (11)$$

where  $M$  and  $G$  are integers. The reason for this definition is that the relation

$$|\phi_{i,j,f-1} - \phi_{i,j,(f \pm \Delta f)^{-1}}| \leq (MG)^{-1}$$

is valid for any  $i, j$  and  $f$ . If the distribution of time points is displayed with a histogram ( $N_{\text{bin}}$ ) of  $M$  bins in  $\phi_{i,P}$ , the phase shift of any  $t_i$  will not exceed  $(MG)^{-1}$  when the frequency changes from  $f$  to  $f \pm \Delta f$ . If  $G = 1$ , this means that an arbitrary  $t_i$  may shift from any of the  $M$  bins only to the next bin on the left or right, when the tested frequency changes by  $\pm \Delta f$ . Increasing the value of  $G$  can be used to further “overfill” the grid of tested frequencies.

The periodogram  $V_{a_n}$  is derived for all integer multiplets of  $\Delta f$  between  $f_{\min}$  and  $f_{\max}$ .

The problem of determining the number of statistically independent frequencies ( $m$ ) within an interval  $[f_{\min}, f_{\max}]$  is not trivial. No unique solution may exist, although several approaches have been suggested, e.g. by Scott (1991, his Sect. 4.1.). In this paper the frequency  $f_0 = \Delta T^{-1}$  is used to estimate the number of independent frequencies. The relation between  $\Delta f$  and  $f_0$  is

$$f_0[\Delta f]^{-1} = GM = [\Delta\phi]^{-1}, \quad (12)$$

where  $[\Delta\phi]^{-1}$  will be called the ‘‘overfilling factor’’. The estimate of the number of independent tested frequencies is

$$m = \text{INT} [(f_{\max} - f_{\min})f_0^{-1}]. \quad (13)$$

The validity of this estimate of  $m$  can be verified empirically by studying the correlation between subsequent points of an arbitrary periodogram ( $Y(f)$ ). The linear correlation coefficient  $r(k)$  for  $Y(f)$  is

$$r(k) = \frac{\sum_{i=1}^{m_1} (Y(f_i) - \langle Y \rangle)(Y(f_{i+k}) - \langle Y \rangle)}{\sqrt{[\sum_{i=1}^{m_1} (Y(f_i) - \langle Y \rangle)^2][\sum_{i=1}^{m_1} (Y(f_{i+k}) - \langle Y \rangle)^2]}} \quad (14)$$

where  $\langle Y \rangle$  is the mean of  $Y$  and  $m_1 = m[\Delta\phi]^{-1} - k$ . The relation  $r([\Delta\phi]^{-1}) = 0$  should be satisfied, if the values of  $Y$  at integer multiplets of  $f_0$  are independent. For example, this function  $r(k)$  for the periodograms of Figs. 1a and 4a of Paper II is shown in Fig. 2, where the correlation reaches zero at approximately  $f_0$  in both cases.

Several conclusions can be drawn from Fig. 2. Firstly, it is usually necessary to derive the periodogram for more than  $m$  frequencies ( $[\Delta\phi]^{-1} > 1$ ) to detect all potential period candidates. Fortunately, the number of independent frequencies does not exceed  $m$ , because the additional points of the periodogram derived at the ‘‘overfilled’’ frequencies correlate with those at independent frequencies. Therefore the significance estimates based on Eq. (15), to be discussed below, are reliable for any value of  $[\Delta\phi]^{-1} \geq 1$ . Secondly, any periodogram may contain regularities not only connected with the distribution of the time points, but also with  $\Delta T$ , as shown by the interference pattern of  $r(k)$  for larger values of  $k$  in Fig. 2 (open squares). Thirdly, while the chosen definition of  $f_0$  does not depend on the form of the periodogram, the parameter  $r(k)$  of a given distribution of time points may not necessarily vanish at exactly  $k = [\Delta\phi]^{-1}$ . A ‘‘pathological’’ example of such a distribution would be a case where one time point is separated by a large gap from the rest of the data. For example, tests performed for the large number of periodograms analysed in Paper II did show that the first value  $r(k) \sim 0$  occurred very close to  $k = [\Delta\phi]^{-1}$ . A similar approach, connected to our  $r(k)$  of Eq. (14), can

be found in, e.g., Buccheri & De Jager (1989: especially their Fig. 6).

It is straightforward to show that the probability  $\gamma$  of falsely rejecting  $H_0$  for  $m$  independent tested frequencies transforms to a condition that the probability of falsely rejecting  $H_0$  with *one* of these tested frequencies must be

$$\alpha_m = 1 - (1 - \gamma)^{1/m}. \quad (15)$$

This means that it is no longer possible to apply the criterion of Eq. (5) by using directly the values of  $C_n$  specified in Table 1 by Swanepoel & De Beer (1990). It is necessary first to choose a ‘‘preassigned level of significance’’, for example  $\gamma = 0.01$ , for falsely rejecting  $H_0$  for  $m$  independent tested frequencies and then to derive the corresponding  $\alpha_m$  of Eq. (15).

The analytical connection between  $C_n$  and  $\alpha_m$  is yet undetermined. Instead of determining  $C_n$  for a fixed  $\alpha_m$  by using Monte Carlo simulations for each particular data set, the following approximation by Swanepoel & De Beer (1990) is adopted in this paper:

$$C_n(\alpha_m) = 0.5 + [0.125 - s(\alpha_m)]/\beta_n. \quad (16)$$

The value of  $s(\alpha_m)$  is given by

$$U(s(\alpha_m)) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{s(\alpha_m)} e^{-t^2/2} dt = 1 - \alpha_m, \quad (17)$$

where  $U(s(\alpha_m))$  is the standard cumulative function of the normal distribution and the solution for  $s(\alpha_m)$  is performed by numerical inversion. Note that this ‘‘asymptotic’’ estimate of  $C_n$  with Eqs. (16) and (17) is not exact and introduces uncertainty to the significance estimates of the WSD-method for smaller samples. But, as illustrated in Fig. 1, where this estimate is compared to the simulated values of  $C_n$ , the accuracy is sufficient for larger samples (e.g. Paper II:  $n=44, 62$  or  $102$ ).

The statistical tests with the WSD-method in the case of  $m$  independent tested frequencies can now be summarized. If the criterion

$$V_{a_n} \leq C_n(\alpha_m)\beta_n^{-1} \quad (18)$$

is fulfilled, the probability of falsely rejecting  $H_0$  is  $\gamma$ . Combining Eqs. (15)–(18) offers an alternative approach to test  $H_0$ . The probability of falsely rejecting  $H_0$  is  $\gamma$ , if

$$Q_{\text{WSD}} \leq \gamma, \quad (19)$$

where  $Q_{\text{WSD}} = 1 - [U(s_0)]^m$  and  $s_0 = 0.125 - \beta_n(\beta_n V_{a_n} - 0.5)$ . Under the ‘‘null hypothesis’’ the parameter  $Q_{\text{WSD}}$  gives the probability for the occurrence of a particular value of  $V_{a_n}$ , or of a more extreme, smaller, value. Hence  $Q_{\text{WSD}}$  will be hereafter called the ‘‘critical level’’ of the WSD-method.

### 3. Extension of the K-method

In this paper the Kuiper (1960) method is abbreviated as the K-method and the statistics of this method are summarized in Sect. 3.1. To be precise, this method was originally devised to analyse a random sample of angular directions in a plane, while the version discussed in this paper is formulated for time points of a periodic phenomenon. The K-method is one modification of the well-known Kolmogorov–Smirnov statistics and the connection between these two tests is discussed, e.g. by Barr & Shudde (1973). The formulation of the two-sample method presented by Kuiper (1960) will be discussed in Sect 3.2. (hereafter the K<sup>2</sup>-method). Finally, the weighted version of the K-method is presented in Sect. 3.3. (hereafter the WK-method).

#### 3.1. The K-method

A summary of the more detailed formulation by Stephens (1965) for testing one arbitrarily chosen period  $P$  with the K-method is given in the beginning of this section. The procedure of testing  $m$  independent frequencies is formulated at the end of this section.

First, the  $n$  phases  $\phi_{1,P}, \dots, \phi_{n,P}$  are arranged into rank order. Then, a sample distribution function

$$F_n(\phi) = \begin{cases} 0, & \phi < \phi_1 \\ iN^{-1}, & \phi_{i,P} \leq \phi < \phi_{i+1,P}, 1 \leq i \leq n-1 \\ 1, & \phi \geq \phi_n \end{cases} \quad (20)$$

is derived. The “null hypothesis” ( $H_0$ ) is that the phases  $\phi_{i,P}$  represent a random sample drawn from a theoretical distribution determined by a cumulative distribution function  $F(\phi)$ .  $H_0$  in this paper is that the phases  $\phi_{i,P}$  are uniformly distributed between 0 and 1, i.e. no periodicity is present and  $F(\phi) = \phi$ . Note that this  $H_0$  is the same as that already adopted for the WSD-method in Sect. 2. The Kuiper test statistic is

$$V_n = D^+ + D^-, \quad (21)$$

where  $D^+$  and  $D^-$  denote the maximum values of  $F_n(\phi) - F(\phi)$  and  $F(\phi) - F_n(\phi)$ , respectively. Kuiper (1960) proved three important characteristics of this test statistic. Firstly, the distribution of  $V_n$  is independent of  $F(\phi)$ , i.e. the K-method has more general applications than the particular test of the “null hypothesis” of uniform distribution performed in this paper. Secondly, the test statistic is independent of the choice of the zero point of  $\phi_{i,P}$ . Thirdly, under  $H_0$ , the probability that  $n^{1/2}V_n$  does reach a value larger or equal to  $z$  is

$$P(n^{1/2}V_n \geq z) = P_n(z) = \sum_{k=1}^{\infty} 2(4k^2z^2 - 1)e^{-2k^2z^2} - (8z/3)n^{-1/2} \sum_{k=1}^{\infty} k^2(4k^2z^2 - 3)e^{-2k^2z^2} + O(n^{-1}), \quad (22)$$

where the contribution of the residual  $O(n^{-1})$  becomes insignificant for larger samples ( $n \gtrsim 20$ ). An exact analytical solution of  $P_n(z)$  for smaller samples has been derived by Stephens (1965: Theorem 2).

Thus the condition for reaching a preassigned significance level  $\gamma$  of falsely rejecting  $H_0$  for  $m$  independent tested frequencies transforms to the condition  $P_n(z_0) = \alpha_m$ , where  $\alpha_m$  is first found from Eq. (15) and then Eq. (22) is solved iteratively for  $z_0$ . A fixed preassigned significance level  $\gamma$  is reached, if

$$V_n \geq z_0 n^{-1/2} \quad (23)$$

or alternatively, if

$$Q_K \leq \gamma, \quad (24)$$

where  $Q_K = 1 - [1 - P(z)]^m$  is the probability that under  $H_0$  a particular value of  $z = n^{1/2}V_n$ , or an even larger value, can occur. This parameter  $Q_K$  can therefore be called the “critical level” of the K-method.

#### 3.2. The K<sup>2</sup>-method

The “null hypothesis” that two independent samples, with respectively  $n_1$  and  $n_2$  time points, have been drawn from the same distribution can be tested by the K<sup>2</sup>-method. The main idea is the definition of the test statistic

$$V_{n_1, n_2} = D_{n_1, n_2}^+ + D_{n_1, n_2}^-, \quad (25)$$

where

$$D_{n_1, n_2}^+ = \max\{F_{n_1}(\phi) - F_{n_2}(\phi)\} \\ D_{n_1, n_2}^- = \max\{F_{n_2}(\phi) - F_{n_1}(\phi)\}$$

and  $F_{n_1}$  and  $F_{n_2}$  are the sample distribution functions of Eq. (20). This test statistic is independent of the zero point of phase. Kuiper (1960) proved two important relations for the case  $n_1 < n_2$ :

$$\lim_{n_2 \rightarrow \infty} V_{n_1, n_2} = V_{n_1} \quad (26)$$

$$P(n_1^{1/2}V_{n_1} \geq z) < P(n_1^{1/2}V_{n_1, n_2} \geq z). \quad (27)$$

These relations show that the test statistic  $V_{n_1, n_2}$  approaches  $V_{n_1}$  as  $n_2$  increases, i.e. the statistics of Sect. 3.1. will apply for  $V_{n_1}$ . Note that the inequality of Eq. (27) implies that Eqs. (23) and (24) will over-estimate the significance for falsely rejecting the “null hypothesis” that the two samples have been drawn from the same distribution. In other words, the possibility that both distributions do in fact represent a random sample drawn from the same theoretical distribution is *under-estimated*.

The idea expressed in Eqs. (26) and (27) is relatively easy to explain. Whatever the theoretical cumulative distribution function  $F(\phi)$  “behind”  $F_{n_1}(\phi)$  and  $F_{n_2}(\phi)$ , the latter will inevitably approach  $F(\phi)$  as  $n_2$  increases, while the deviations of both from  $F(\phi)$  will contribute to the

test statistic  $V_{n_1, n_2}$ . One consistent interpretation for applications is that the sample distribution function  $F_{n_2}$  takes over the rôle of the cumulative distribution function  $F(\phi)$  as  $n_2$  increases. Thus the “null hypothesis” is that the smaller sample of  $n_1$  time points represents a random sample drawn from a distribution determined by  $F(\phi) \sim F_{n_2}(\phi)$ . In other words, the validity of the hypothesis that the larger sample ( $n_2$ ) can predict the distribution of the smaller sample ( $n_1$ ) is tested. The larger the test statistic  $V_{n_1, n_2} \sim V_{n_1}$ , the more probable the case that the prediction is false. An application of the  $K^2$ -method is presented in Paper II to test whether the results for previous data ( $n_2$ ) can predict the new data with  $n_1 < n_2$ . Finally, we wish to emphasize that the statistics of the  $K^2$ -method are valid only for *one* tested period.

### 3.3. The WK-method

The definition of the weighted sample distribution function of the WK-method for the ranked  $\phi_{i,P}$  is

$$F_{n,w}(\phi) = \begin{cases} 0, & \phi < \phi_1 \\ w_*^{-1} \sum_{k=1}^i w_i, & \phi_{i,P} \leq \phi < \phi_{i+1,P}, 1 \leq i \leq n-1, \\ 1, & \phi \geq \phi_n \end{cases} \quad (28)$$

where  $w_i$  is the weight associated with  $\phi_{i,P}$ , and  $w_* = \sum_{i=1}^n w_i$ . Because this definition of  $F_{n,w}(\phi)$  differs from that of  $F_n(\phi)$  in Eq. (20), the statistics are not necessarily the same as those determined by Eqs. (22)-(24). The definition of Eq. (28) is comparable to a case where  $w_i$  values of  $\phi_{i,P}$  are derived for each particular  $t_i$ . In this case, the total number of phases  $\phi_{i,P}$  would be  $w_*$ , and the shapes of  $F_n$  and  $F_{n,w}$  would coincide. But the average height of the steps of  $F_{n,w}(\phi)$  is  $n^{-1}$  and there are only  $n$  (not  $w_*$ ) steps, i.e.  $F_{n,w} = F_n$  is valid only for equal weights. The test statistic of the WK-method is derived by using  $F_{n,w}$  as the sample distribution function of Eq. (21).

We shall now study the consequences of using the weighted sample distribution function of Eq. (28). The exact analytical solution of the statistics of the WK-method is beyond the scope of this study. But an analytical proof of the more general case of an arbitrary distribution function  $F(\phi)$  is not necessary for the  $H_0$  chosen in this paper. Hence Monte Carlo simulations were applied to study whether the statistics of the K- and WK-methods are comparable for the special case of an even distribution (i.e.  $F(\phi) = \phi$ ). The results of these simulations are shown in Fig. 3. As could be expected the results for the WK-method with equal weights agree with those by Stephens (1965), but the case of unequal weights does not yield the same statistics. The results in Fig. 3 indicate that for unequal weights the criteria of Eqs. (23) and (24) will over-estimate the significance of any tested period for rejecting  $H_0$ . The reason for this deviation between the statistics of the K- and WK-method is that larger weights will contribute to the parameters  $D^+$  and  $D^-$  of Eq. (21) for

every tested period. The range of these unequal weights was the same as in the simulations performed in Sect. 2.2. with the WSD-method (i.e.  $3 \leq w_i \leq 17$ ). On the other hand, for a smaller scatter of weights and/or larger values of  $n$ , the statistics of the K- and WK-methods do approach each other. This tendency for larger samples is evident in Fig. 3, while the above mentioned influence of adopting weights with a smaller dispersion was checked with additional simulations. Furthermore, although Eqs. (23) and (24) do yield over-estimates of the significance of any value of the test statistic for the case of unequal weights, the order of the significance of the best periods for the data depends only on how large a value of  $V_n$  is reached.

The difference of the statistics of the K- and WK-methods in the case of smaller samples and/or larger scatter of weights can be expressed by the two inequalities given below. If  $V_n$  and  $Q_K$  are the parameters of the K-method and the notations  $V_{n,w}$  and  $Q_{WK}$  are used for the test statistic and the critical level of the WK-method respectively, the inequality

$$V_{n,w} \geq V_n \quad (29)$$

is fulfilled. When Eqs. (22) and (24) are applied to estimate  $Q_{WK}$ , the result is

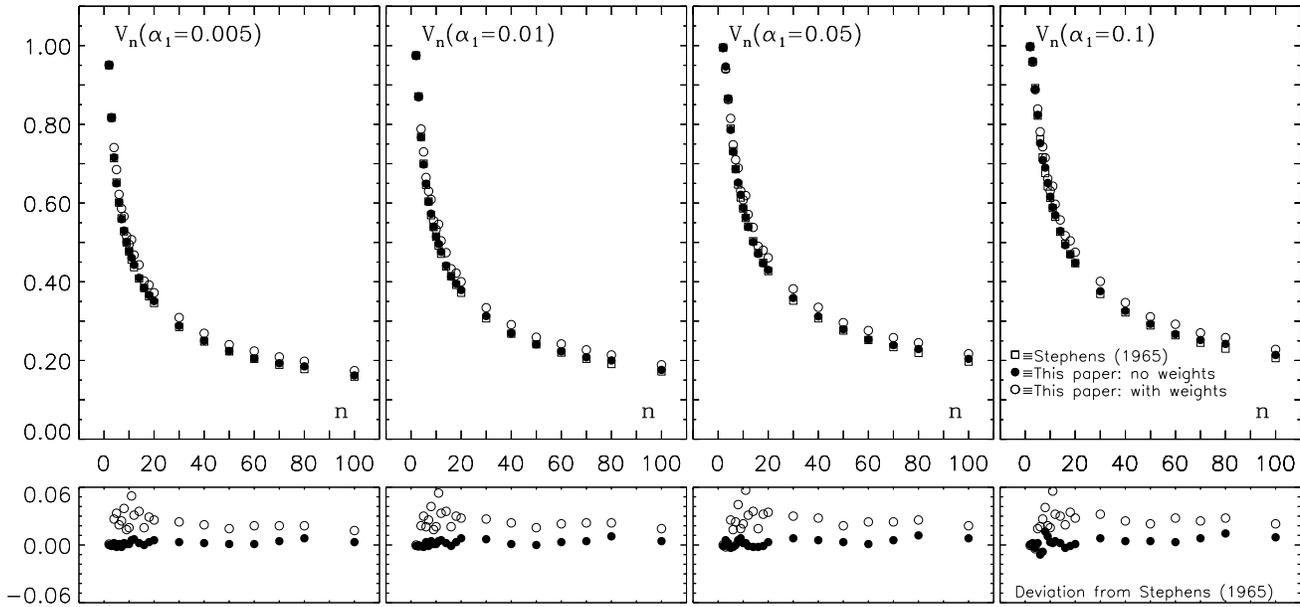
$$Q_{WK} \leq Q_K. \quad (30)$$

But it would be inconsistent to change the notation  $V_n$  for the test statistic to  $V_{n,w}$ , because the difference of these two parameters is solely due to the introduction of unequal weights into the sample distribution function of Eq. (28). However, to emphasize the case when the WK-method is applied or referred to, the notation  $Q_{WK}$  is used for the critical level, instead of  $Q_K$ . We conclude that Eqs. (22)-(24) can be used to estimate the statistics of the WK-method, if the limitations expressed by the inequalities of Eqs. (29) and (30) are taken into account.

Finally, the consequences of adopting  $F_{n_1, w_1}$  and  $F_{n_2, w_2}$  to determine the test statistic of Eq. (25) of the  $K^2$ -method, as e.g. in Paper II, can be inferred. The probability of *under-estimating* the validity of the prediction hypothesis will further increase due the stronger fluctuations of these weighted sample distribution functions.

## 4. Error estimates

Several sections have been devoted to the task of determining consistent statistics for different nonparametric methods for searching for periodicity in weighted time point series. But the significance estimates, although convenient for testing a given  $H_0$ , do not give an error for the period ( $\sigma_P$ ) when periodicity is clearly present and  $H_0$  should be rejected. A consistent determination of  $\sigma_P$  is not trivial, since no analytical model necessary for this purpose exists, because the methods are nonparametric.



**Fig. 3.** Monte Carlo simulations of  $V_n$  of Eq. (21) at four preassigned significance levels ( $\alpha = 0.005, 0.01, 0.05$  and  $0.1$ ) for the “null hypothesis”  $F(\phi) = \phi$ : closed circles (WK–method: with equal weights) and open circles (WK–method: with unequal weights). The analytical values of  $V_n$  given by Stephens (1965) are shown with open squares. Finally, the differences between the simulated values minus those tabulated by Stephens (1965) are shown in the lower figures

However, the bootstrap method offers a possibility for estimating the error of  $P$  (e.g. Efron & Tibshirani 1986). In the bootstrap applied below we will denote the *original* data simply by the vectors  $\bar{t}$  and  $\bar{w}$ . Referring to the discussion connected to Eq. (14), it is logical to restrict the study of the periodogram with the bootstrap method to the frequency interval  $[f_{\text{best}} - f_0, f_{\text{best}} + f_0]$ , where  $f_{\text{best}}$  is a particular value indicating that periodicity with  $P = [f_{\text{best}}]^{-1}$  is present. This frequency  $f_{\text{best}}$  will minimize ( $V_{a_n}$ ) or maximize ( $V_n$ ) the test statistic

$$Y(f_{\text{best}}) = Y(f_{\text{best}}, \bar{t}, \bar{w}), \tag{31}$$

of the WSD– or WK–method, respectively. The notations  $\bar{t}^*$  and  $\bar{w}^*$  are used for an arbitrary new sample of  $n$  components selected at random from the original data vectors  $\bar{t}$  and  $\bar{w}$ . Any component of  $\bar{t}$  and  $\bar{w}$  may enter these random samples  $\bar{t}^*$  and  $\bar{w}^*$  as many times as the random selection “favours” it. However, each  $w_i$  remains the weight of  $t_i$  through this random selection process. A large number ( $q_1$ ) of such random samples  $\bar{t}_1^*, \dots, \bar{t}_{q_1}^*$  and  $\bar{w}_1^*, \dots, \bar{w}_{q_1}^*$  are selected from the original data (e.g.  $q_1=300$  in Paper II). Each of these random samples is then used to derive the best frequency  $f_{\text{best},q}$  for the periodogram

$$Y(f_{\text{best},q}) = Y(f_{\text{best},q}, \bar{t}_q^*, \bar{w}_q^*), \tag{32}$$

in the frequency interval  $[f_{\text{best}} - f_0, f_{\text{best}} + f_0, ]$ , which yields a total number of  $q_1$  values  $f_{\text{best},1}, \dots, f_{\text{best},q_1}$ . The standard deviation of all values  $f_{\text{best},q}^{-1}$  is then adopted as an estimate of the error  $\sigma_P$  given by the bootstrap for  $P =$

$f_{\text{best}}^{-1}$ . We conclude that the bootstrap approach outlined in this section can be applied to any general method for searching for periodicity in weighted or nonweighted time point series.

### 5. Conclusions

The WSD– and WK–methods for searching for periodicity in weighted time point series have been presented. They are extensions of the methods by Swanepoel & De Beer (1990) and Kuiper (1960). Two alternative criteria were presented for both methods (Eqs. (18), (19), (23) and (24)) to test a particular value of the test statistic ( $V_{a_n}$  and  $V_n$ ) against a preassigned significance level  $\gamma$  of falsely rejecting the “null hypothesis” ( $H_0$ ). The  $H_0$  for both methods was: “No periodicity is present”, or equivalently: “The phases  $\phi_{1,P}, \dots, \phi_{n,P}$  represent a random sample drawn from a uniform distribution between 0 and  $1''$ . Furthermore, the statistics for these methods when testing a number  $m$  statistically independent frequencies were derived, including an empirical approach to check  $m$  of any particular test within the chosen frequency interval. A consistent formulation, based on the bootstrap method, to estimate the error of the period detected when using the WSD– and WK–methods has been presented, which is applicable to any weighted or nonweighted method for searching for periodicity in time point series.

Six important limitations must be considered in applying the WSD– and WK–methods:

1. The estimate of the number  $m$  of independent frequencies within a tested frequency interval (Eq. 13) has the strongest influence on the significance estimates through Eq. (15). The “empirical” approach connected to  $r(k)$  of Eq. (14) should be applied after each test to verify whether the estimate of  $m$  is correct.
2. The WSD-method will weakly over-estimate the significance of any period for smaller samples of data, because Eq. (16) is an asymptotic estimate of  $C_n$  for larger samples.
3. In the case of smaller samples and/or larger scatter of weights, Eqs. (23) and (24) over-estimate the significance of any period detected with the WK-method. These estimates become more accurate when the sample size increases and/or the scatter of weights decreases. Note also that asymptotic estimate of Eq. (22) is valid only for larger samples.
4. The statistics of the  $K^2$ -method are exact only if  $n_2 \rightarrow \infty$ , and hence the criteria of Eqs. (23) and (24) will *under-estimate* the probability that  $F_{n_1}(\phi)$  and  $F_{n_2}(\phi)$  have a common  $F(\phi)$ , especially for the weighted sample distribution functions of Eq. (28).
5. The statistics of the WK-method presented in this paper are based on simulations with the “null hypothesis”  $F(\phi) = \phi$ , and do not necessarily apply for the more general case of an arbitrary  $F(\phi)$ .
6. Although the methods discussed in this paper are nonparametric, i.e. model independent, none of them can be equally sensitive to all types of distributions. A number of examples of these sensitivity effects will be given in Paper II.

Regardless of the above limitations for the application of the weighted tests presented in this paper, the most important conclusion is that the order of significance for the detected period candidates is *always* correct, as well as the error estimates derived with the bootstrap.

*Acknowledgements.* The work by L.J. was partly supported by the EC Human Capital and Mobility (Networks) project “Late type stars: activity, magnetism, turbulence” No. ERBCHRXCT940483. The work by J.P. was supported by a grant from the Centre for International Mobility (Nordic Scholarship Scheme for Baltic Countries and Northwest Russia). We would also like to thank Dr. David Moss for checking the English of the manuscript.

## References

- Bai T., 1987, ApJ 314, 795  
 Bai T., 1988, ApJ 328, 860  
 Bai T., 1990, ApJ 364, L17  
 Barr D.R., Shudde R.H., 1973, Biometrika 60, 663  
 Batschelet E., 1981, Circular Statistics in Biology. Academic Press, London  
 Buccheri R., De Jager O.C., 1989, in Timing Neutron Stars, NATO ASI Series C, Vol. 262. In: Ögelman H., van den Heuvel E.P.J. (eds.). Kluwer Academic Publishers, Dordrecht, p. 95  
 Efron B., Tibshirani R., 1986, Stat. Sci. 1, 54  
 Jetsu L., Pohjolainen S., Pelt J., Tuominen I., 1995, A&A (submitted) (Paper II)  
 Kuiper N.H., 1960, Proc. Koninkl. Nederl. Akad. Van Wetenschappen, Ser. A 63, 38  
 Mardia K.V., 1972, Statistics of Directional Data. Academic Press, London  
 Moore B.R., 1980, Biometrika 67, 175  
 Protheroe R.J., Clay R.W., Gerhardy P.R., 1984, ApJ 280, L47  
 Rao J.S., 1969, Thesis, Indian Statisc. Institute, Calcutta  
 Rao J.S., 1972, Z. Wahrscheinlichkeitstheorie verw. Geb. 22, 33  
 Rao J.S., 1976, Sankhyā, Ser. B 33, 1  
 Scott D., 1991, A&A 242, 1  
 Stephens M.A., 1965, Biometrika 52, 309  
 Swanepoel J.W.H, De Beer C.F., 1990, ApJ 350, 754  
 Watson G.S., 1961, Biometrika 48, 109