Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **464,** 274–292 (2017) doi:10.1093/mnras/stw2262
Advance Access publication 2016 September 8

Comparative performance of selected variability detection techniques in photometric time series data

K. V. Sokolovsky,[1,2,3★] P. Gavras,[1] A. Karampelas,[1] S. V. Antipin,[2,4] I. Bellas-Velidis,[1] P. Benni,[5] A. Z. Bonanos,[1] A. Y. Burdanov,[6,7] S. Derlopa,[8] D. Hatzidimitriou,[1,9] A. D. Khokhryakova,[6] D. M. Kolesnikova,[4] S. A. Korotkiy,[10] E. G. Lapukhin,[11] M. I. Moretti,[1] A. A. Popov,[6] E. Pouliasis,[1,9] N. N. Samus,[2,4] Z. Spetsieri,[1,9] S. A. Veselkov,[11] K. V. Volkov,[6] M. Yang[1] and A. M. Zubareva[2,4]

[1]*IAASARS, National Observatory of Athens, 15236 Penteli, Greece*
[2]*Sternberg Astronomical Institute, Moscow State University, Universitetskii pr. 13, 119992 Moscow, Russia*
[3]*Astro Space Center of Lebedev Physical Institute, Profsoyuznaya Str. 84/32, 117997 Moscow, Russia*
[4]*Institute of Astronomy (Russian Academy of Sciences), Pyatnitskaya Str. 48, 119017 Moscow, Russia*
[5]*Acton Sky Portal, 3 Concetta Circle, Acton, MA 01720, USA*
[6]*Kourovka Astronomical Observatory of Ural Federal University, Mira Str. 19, 620002 Ekaterinburg, Russia*
[7]*Institut d'Astrophysique et Géophysique, Université de Liège, allée du 6 Août 17, B-4000 Liège, Belgium*
[8]*Department of Physics, University of Patras, 26500 Patra, Greece*
[9]*Department of Astrophysics, Astronomy & Mechanics, Faculty of Physics, University of Athens, 15783 Athens, Greece*
[10]*Ka-Dar astronomy foundation, Kuzminki, PO Box 82, 142717 Moscow region, Russia*
[11]*Reshetnev Siberian State Aerospace University, Krasnoyarsky Rabochy Av. 31, 660037 Krasnoyarsk, Russia*

Accepted 2016 September 7. Received 2016 September 5; in original form 2016 May 2

ABSTRACT

Photometric measurements are prone to systematic errors presenting a challenge to low-amplitude variability detection. In search for a general-purpose variability detection technique able to recover a broad range of variability types including currently unknown ones, we test 18 statistical characteristics quantifying scatter and/or correlation between brightness measurements. We compare their performance in identifying variable objects in seven time series data sets obtained with telescopes ranging in size from a telephoto lens to 1 m-class and probing variability on time-scales from minutes to decades. The test data sets together include light curves of 127 539 objects, among them 1251 variable stars of various types and represent a range of observing conditions often found in ground-based variability surveys. The real data are complemented by simulations. We propose a combination of two indices that together recover a broad range of variability types from photometric data characterized by a wide variety of sampling patterns, photometric accuracies and percentages of outlier measurements. The first index is the interquartile range (IQR) of magnitude measurements, sensitive to variability irrespective of a time-scale and resistant to outliers. It can be complemented by the ratio of the light-curve variance to the mean square successive difference, $1/\eta$, which is efficient in detecting variability on time-scales longer than the typical time interval between observations. Variable objects have larger $1/\eta$ and/or IQR values than non-variable objects of similar brightness. Another approach to variability detection is to combine many variability indices using principal component analysis. We present 124 previously unknown variable stars found in the test data.

Key words: methods: data analysis – methods: statistical – stars: variables: general.

1 INTRODUCTION

A variety of phenomena manifest themselves as changes in apparent brightness of astronomical objects. The amplitudes and time-scales of these changes vary from tens of magnitudes and weeks for supernova explosions to a fraction of a magnitude and minutes for stellar pulsations. With the notable exceptions of light echoes (e.g. Bond et al. 2003), variable reflecting nebulae (e.g. Close et al. 1997) and the M87 jet (e.g. Perlman et al. 2011), variable objects are unresolved by single-dish telescopes.[1] Variable point-like objects are

★E-mail: kirx@noa.gr

[1] The light travel time argument implies that an object varying on a time-scale t cannot be larger than ct light seconds, otherwise its variability would be smeared.

often embedded in light of a resolved non-variable source (active nucleus or a supernova in a galaxy; young stellar object embedded in a nebula) that complicate measurements of the variable object's brightness. The variations may be associated with a single catastrophic event (supernova), may be approximately (dwarf novae) or strictly periodic (eclipsing binaries) or aperiodic (active galactic nuclei) in nature. Our understanding of these events depends on the efficient and reliable detection of brightness variations.

Photometric measurements are prone to systematic errors that are difficult to characterize. This makes it challenging to distinguish true low-amplitude variability from the apparent one caused by systematic effects and measurement errors. Imaging artefacts such as cosmetic defects of a CCD, diffraction spikes from bright objects and cosmic ray hits as well as blending between images of nearby objects can corrupt photometry and mimic high-amplitude variability. Three different lines of attack on the problem of variable object detection are described in the literature: direct image comparison ('transient detection'), light-curve analysis using variability indices and periodicity search.

Transient detection techniques seek to identify changes between two sets of sky images taken at different times (epochs). The changes may be found by subtracting the images pixel by pixel after resampling them to a common coordinate grid and accounting for seeing changes (difference image analysis – DIA; Tomaney & Crotts 1996; Alard & Lupton 1998; Alard 2000; Bramich 2008; Becker et al. 2012; Bramich et al. 2016; Zackay & Ofek 2015; applications of the method include Bonanos et al. 2003; Zheleznyak & Kravtsov 2003; Arellano Ferro et al. 2013; Sahay, Lebzelter & Wood 2014; Zhang et al. 2015). Large surveys such as Optical Gravitational Lensing Experiment (OGLE; Udalski, Szymański & Szymański 2015), PTF (Law et al. 2009), Pan-STARRS (Rest et al. 2014) and DES (Kessler et al. 2015) implement the image subtraction technique. Alternatively, one may extract astronomical objects (sources) from each image independently and compare the resulting source lists (Contreras Peña et al. 2014, CRTS – Drake et al. 2009). The second-epoch images are often taken in pairs, triplets or even longer series with dithering to eliminate image artefacts that are usually associated with a given position on the image detector, not in the sky.

More sophisticated detection strategies may be applied if measurements are obtained at more than two epochs. Their obvious advantage over the simple two-epoch data comparison is the potential to average out individual measurement errors and thus detect variability with a lower amplitude. One class of methods employs various 'variability indices' characterizing the overall scatter of measurements in a light curve and/or degree of correlation between consecutive flux measurements (some recent examples, Munari et al. 2014; Javadi et al. 2015; Yao et al. 2015b, see the detailed discussion in Section 2). The other class of methods search for significant periodicity in flux variations (e.g. Drake et al. 2014; Kaluzny et al. 2014; McCormac et al. 2014; Chakrabarti et al. 2015; Nardiello et al. 2015, 2016; Soszyński et al. 2015). While many types of variable stars show periodic or semi-periodic light variations, flux measurement errors are expected to be aperiodic, or associated with a known periodic process inherent to the observations (diurnal or seasonal cycle, synodic month, periodic guiding errors, orbital period of a space-borne telescope, etc.).

If a search is aimed at a specific variability type for which a light-curve shape is generally known in advance (e.g. exoplanet transits or eclipsing binaries in general, Cepheids, RR Lyrae stars, novae), template fitting (e.g. Jenkins, Doyle & Cullers 1996; Macri et al. 1999; Prša et al. 2011; Sesar et al. 2013; Angeloni et al. 2014; Hoffmann et al. 2016) with various trial periods/flare development time-scales can be performed. Simple cuts on light-curve parameters (Henze, Meusinger & Pietsch 2008; Graczyk & Eyer 2010) as well as advanced machine learning techniques (Feeney et al. 2005) can be used to select light curves of a known shape from a large photometric data set. A pre-selection based on colour can be used to reduce the number of candidates when searching for variables of a specific type (Kinemuchi et al. 2006; Tisserand et al. 2013; Zinn et al. 2014; Moretti et al. 2016; Ordoñez & Sarajedini 2016).

Since period search and template-fitting algorithms are computationally expensive, a two-step approach can be applied. Candidate variable stars are pre-selected using a fast-to-compute variability index (and/or colour) and only the light curves that passed this selection are subjected to period search (e.g. Akerlof et al. 2000; Drake et al. 2013; Kourniotis et al. 2014; Fernández-Trincado et al. 2015; Ferreira Lopes et al. 2015; Gran et al. 2015; Vivas et al. 2016) or template fitting (e.g. Shappee & Stanek 2011; Hoffmann & Macri 2015). If the total number of observed objects is low, both period search and light-curve scatter-based selection criteria are applied independently of each other to conduct exhaustive search for both periodic and non-periodic variables (Sitek & Pojmański 2014; Rebull et al. 2015). Selection based on period search may be followed by even more computationally intensive steps like binary system modelling (Devor 2005). Kim et al. (2014) used the period along with other variability features as an input for the random forest algorithm to select periodic variable star candidates in the EROS-2 data base and simultaneously classify them.

The methods described above may efficiently select variable object candidates from a large set of photometric data. However, the final decision to designate an object as 'variable star' rather than a 'candidate' is usually made only after visual inspection of the object's light curve by a human expert (e.g. Pojmanski, Pilecki & Szczygiel 2005; Graczyk et al. 2011; Palaversa et al. 2013; Pawlak et al. 2013; Pietrukowicz et al. 2013; Cusano et al. 2013; Klagyivik et al. 2016; Song et al. 2016). If the number of observations is small, the original images are checked for the presence of obvious problems [image artefacts, cosmic ray hits, point spread function (PSF) wings of a bright nearby object] affecting photometry of a candidate variable (e.g. Dolphin et al. 2003; Bernard et al. 2010; Denisenko & Sokolovsky 2011; Ramsay et al. 2014). While advanced image artefact rejection procedures exist (Fruchter & Hook 2002; Desai et al. 2016), visual image inspection remains an important data quality control tool as it may uncover unexpected problems (Melchior et al. 2016).

Variable star detection may be considered in the framework of classical hypothesis testing (e.g. Wall & Jenkins 2003): to establish that an object is variable, one needs to rule out the null hypothesis that it is constant given the observations (Eyer 2006). One may compare a value of variability detection statistic (Section 2) derived from the light curve to the distribution of this value expected for non-variable objects. The problem is that objects with corrupted measurements produce long tails in the aforementioned distributions. In the presence of badly measured objects, one is forced to set a low threshold for accepting candidate variables (Section 5.6) and rely on additional information not captured by the variability detection statistic to distinguish true variables from badly measured objects in the distribution tail.

Alternatively, one may view the search for variable stars as a classification problem that may be approached with machine learning techniques. The task is to classify a set of objects characterized by their light curves, images associated with each light-curve point and possibly additional pieces of information associated with each brightness measurement (object's position on the CCD frame,

airmass, seeing, temperature, etc.). One needs to distinguish various classes of variable stars from the class of well-measured constant stars and classes of stars affected by various types of measurement errors (bad pixels, diffraction spikes, blending). Objects that do not belong to one of the known classes should also be identified. While considerable progress has been made in light-curve-based automated classification of stars already known to be variable (Debosscher et al. 2007; Paegert, Stassun & Burger 2014; Kim & Bailer-Jones 2016), an automated system that could reliably identify variable stars among non-variables remains to be developed.

In practice, the following approach to variable star detection is often adopted. (i) Objects affected by blending and image artefacts are flagged at source extraction stage. (ii) The light curves of the detected objects are constructed and may be refined using the available additional information (Section 3.7). (iii) The techniques described in the previous paragraphs are used to select promising variable star candidates based on their light curves. (iv) The list of candidates is examined by a human expert who performs the final classification and removes false variables from the list. In this work, we explore the limits of the traditional approach outlined above and identify the best ways to select candidate variables.

We compare the performance of popular variability detection techniques on various real and simulated photometric data sets. We refer to any value that quantifies 'how variable' a given object is as a variability index'. The discussion is limited to variability indices based on light-curve scatter (Sections 2.1–2.7) and correlation (Sections 2.8–2.17) while the period search-based techniques will be discussed elsewhere. We attempt to find a general-purpose variability detection technique able to recover a broad range of variability types including currently unknown ones (Shin, Sekora & Byun 2009). Such a technique would also be useful for solving the opposite problem: reliable selection of non-variable objects that can be used as photometric standards (e.g. Ofek et al. 2012) or targets for searches of variations not intrinsic or not typical to the objects such as microlensing events (Udalski et al. 1994), occultations of stars by distant Solar system bodies (Zhang et al. 2013), tidal disruption events in nuclei of non-active galaxies (van Velzen et al. 2011) and failed supernovae (Kochanek et al. 2008).

Publications focused on comparing performance of variability search techniques include Enoch et al. (2012) who compared planetary-transit detection algorithms, while de Diego (2010) and Villforth, Koekemoer & Grogin (2010) discussed a number of variability detection tests in the context of active galactic nuclei. Ferreira Lopes & Cross (2016) compared performance of some multi-band correlation-based variability indices. Vaughan et al. (2003) and Allevato et al. (2013) discussed properties of 'excess variance' (Section 2.6) and 'fractional variability amplitude', the variability measures often used in X-ray astronomy. Graham et al. (2013) compared the accuracy and performance of period finding algorithms. Findeisen, Cody & Hillenbrand (2015) compare various methods of extracting a characteristic time-scale from an irregular light curve. Nun et al. (2015) provide an extensive list of features useful for light-curve characterization and classification. Preliminary results of our variability index comparison based solely on photographic light curves are presented by Sokolovsky et al. (2016).

This paper is structured as follows. Section 2 defines the variability indices we investigate. Section 3 describes the test data. Section 4 presents the technique for comparison of effectiveness of variability indices in selecting variable objects. Section 5 discusses the results of the comparison and Section 6 summarizes our findings.

Table 1. Information included in variability indices. For references, see the footnote in Table 5.

Index	Errors	Order	Time	Section	Ref.
		Scatter-based indices			
$\chi^2_{\rm red}$	√			2.1	(a)
σ				2.2	(b)
σ_w	√			2.2	(b)
MAD				2.3	(c)
IQR				2.4	(d)
RoMS	√			2.5	(e)
$\sigma^2_{\rm NXS}$	√			2.6	(f)
v	√			2.7	(g)
		Correlation-based indices			
l_1		√		2.8	(h)
I	√	√	√	2.9	(i)
J	√	√	√	2.10	(j)
J(time)	√	√	√	2.11	(k)
J(clip)	√	√	√	2.12	(d)
L	√	√	√	2.10	(j)
CSSD		√		2.13	(l)
E_x	√	√	√	2.14	(m)
$1/\eta$		√		2.15	(n)
$\mathcal{E}_{\mathcal{A}}$		√	√	2.16	(o)
S_B	√	√		2.17	(p)

2 VARIABILITY SEARCH METHODS

In this section, we define the numerical parameters characterizing the 'degree of variability' of an object – the variability indices, discussed in detail in the following paragraphs. The scatter-based indices (Sections 2.2–2.7) consider only the distribution of measured magnitudes ignoring the time information available in a light curve. Some also take into account the estimated errors. The correlation-based indices (Sections 2.8–2.17) in addition to the measured magnitudes themselves consider the order in which the measurements were taken and some indices also take into account the time difference between measurements. The use of this additional information makes correlation-based indices more sensitive to low-amplitude variability, but on the downside, correlation-based indices are insensitive to variability on time-scales shorter than the sampling time (Kim et al. 2011b). Table 1 summarizes the information used by each index. In the following sections, we compare the effectiveness of these variability indices in selecting variable stars.

2.1 χ^2 test

A χ^2 test[2] is any statistical hypothesis test in which the sampling distribution of the test statistic is a χ^2 distribution when the null hypothesis is true. Given N magnitude measurements m_i (assumed to be independent of each other) and their associated errors σ_i (assumed to be Gaussian), the null hypothesis, H_0, that an object does not change its brightness can be tested by computing the value

$$\chi^2 = \sum_{i=1}^{N} \frac{(m_i - \bar{m})^2}{\sigma_i^2}, \quad (1)$$

[2] https://en.wikipedia.org/wiki/Chi-squared_test

where

$$\bar{m} = \sum_{i=1}^{N} \frac{m_i}{\sigma_i^2} \Bigg/ \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \tag{2}$$

is the weighted mean magnitude. χ^2 is compared to the critical value $\chi^2_{p,\nu}$ obtained from the χ^2-distribution with $\nu = N - 1$ degrees of freedom. The p-value indicates the statistical significance level at which H_0 can be rejected.

If measurement errors are estimated correctly, the majority of objects should have χ^2 values consistent with H_0, since the majority of stars are not variable. A notable exception from this rule is millimagnitude-precision photometric observations such as the ones obtained by *MOST* (Walker et al. 2003), *CoRoT* (Auvergne et al. 2009), *Kepler* (Borucki et al. 2010) and future photometric space missions (e.g. Rauer et al. 2014; Ricker et al. 2014), which are able to detect variability in the majority of field stars, including variability caused by transiting Solar system-like planets (Hippke & Angerhausen 2015).

In practice, poor knowledge of σ_i limits the applicability of the χ^2 test for variability detection in ground-based photometry. In this case, χ^2 may still be useful as a measure of scatter in a light curve compared to the expected measurement errors, but the cut-off value for discriminating variable objects from non-variable ones should be different from the one suggested by the χ^2 distribution. In the following, we use the reduced $\chi^2_{\rm red} = \chi^2/N - 1$ (e.g. Andrae, Schulze-Hartung & Melchior 2010) to compare its value for light curves with different N. Villforth et al. (2010) note that estimated photometric measurement errors are asymmetric and non-Gaussian when converted from flux to magnitude space. This violates the assumptions behind the critical value $\chi^2_{p,\nu}$ calculations. The χ^2 test, in its textbook form, should be performed in flux space and only when the contribution from all sources of photometric errors has been properly accounted for.

2.2 Standard deviation, σ_w

A detectable variable star, by definition, should have larger scatter of measurements in its light curve compared to non-variable stars that could be measured with the same accuracy. One way to characterize scatter of measurements is to compute the standard deviation,

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (m_i - \bar{m})^2} \tag{3}$$

or alternatively, if the estimated errors are assumed to reflect the relative accuracy of measurements, its weighted version

$$\sigma_w = \sqrt{\frac{\sum_{i=1}^{N} w_i}{\left(\sum_{i=1}^{N} w_i\right)^2 - \sum_{i=1}^{N} (w_i^2)} \sum_{i=1}^{N} w_i (m_i - \bar{m})^2}. \tag{4}$$

Assuming that m_i are drawn from Gaussian distributions having variances σ_i^2 and the same mean $\bar{m}$, the choice of weights $w_i = 1/\sigma_i^2$ maximizes the likelihood of obtaining the set of measurements (m_i). Therefore, given a set of measurements (m_i, σ_i), equation (2) is the best estimate of the mean under the above assumptions.

We define σ as a square root from an unbiased estimator of the population variance [the Bessel correction, i.e. $(N - 1)$ instead of N in the denominator of equation (3)] as this is the definition often adopted in statistical software, notably in the GNU Scientific Library.[3] For the purpose of variable star search, the use of Bessel's correction has minimal practical consequences.

Standard deviation is relatively sensitive to outlier points. In many cases, light-curve filtering (Section 3.7) might be needed before σ can serve as an efficient variable star selection tool. In the following paragraphs, we describe ways of characterizing light-curve scatter that are less sensitive to outliers.

2.3 Median absolute deviation (MAD)

The median absolute deviation,[4] MAD (Rousseeuw & Croux 1993; Richards et al. 2011), is a measure of scatter of observations m_i defined as

$$\mathrm{MAD} = \mathrm{median}(|m_i - \mathrm{median}(m_i)|). \tag{5}$$

For a Gaussian distribution

$$\sigma = \mathrm{MAD}/\Phi^{-1}(3/4) \simeq 1.4826 \times \mathrm{MAD}, \tag{6}$$

where $\Phi^{-1}(x)$ is the inverse of the cumulative distribution function for the Gaussian distribution. The MAD statistic is mostly insensitive to outliers (Zhang et al. 2016); its only disadvantage is that it is equally insensitive to real variations that occur only occasionally, like rare eclipses of an Algol-type binary that may show virtually constant brightness outside of the eclipses.

The use of MAD is computationally more expensive than σ as the sorting needed to compute the median is a relatively slow, $\mathcal{O}(n \log n)$, operation compared to calculating the average value, $\mathcal{O}(n)$. Here $\mathcal{O}(n \log n)$ ($\mathcal{O}(n)$) means that there is a constant $C > 0$ such that for any number of input measurements, n, the computation will be completed in less than $Cn \log n$ (Cn) steps. It should be noted that correlation-based indices discussed below in Sections 2.8–2.17 computationally depend on the order of data points and, therefore, require the input light curve to be sorted in time – an operation of $\mathcal{O}(n \log n)$ complexity.

2.4 Interquartile range (IQR)

Another robust measure of scatter is the interquartile range,[5] IQR (Kim et al. 2014), which includes the inner 50 per cent of measurement values (i.e. excludes 25 per cent of the brightest and 25 per cent of the faintest flux measurements). To compute the IQR, we first compute the median value that divides the set of flux measurements into upper and lower halves. The IQR is the difference between the median values computed for the upper and lower halves of the data set. For the normal distribution $\mathrm{IQR} = 2\Phi^{-1}(0.75)\sigma \simeq 1.349\sigma$, where $\Phi^{-1}(x)$ is the inverse of the cumulative distribution function for the Gaussian distribution. The IQR may be more appropriate than MAD (Section 2.3) for measuring the width of an asymmetric (skewed) distribution, such as the distribution of flux measurements of an eclipsing binary.

2.5 Robust median statistic (RoMS)

The robust median statistic, RoMS, was proposed by Enoch, Brown & Burgasser (2003) and successfully applied for variable star search

[3] https://www.gnu.org/software/gsl/; https://en.wikipedia.org/wiki/Bessel's_correction
[4] http://en.wikipedia.org/wiki/Median_absolute_deviation
[5] https://en.wikipedia.org/wiki/Interquartile_range

by Rose & Hintz (2007) and Burdanov, Krushinsky & Popov (2014). It is defined as

$$\mathrm{RoMS} = (N-1)^{-1} \sum_{i=1}^{N} \frac{|m_i - \mathrm{median}(m_i)|}{\sigma_i}. \quad (7)$$

For a non-variable object, the expected value of RoMS is around 1 as the majority of the measurements should be within 1σ of the median value (if σ is estimated correctly).

2.6 Normalized excess variance, σ^2_{NXS}

Normalized excess variance, σ^2_{NXS}, is used in X-ray (Nikołajuk, Czerny & Gurynowicz 2009; Ponti et al. 2012; Hernández-García et al. 2015; Yao et al. 2015a) and optical (Simm et al. 2015) astronomy to characterize variability amplitude in the presence of changing measurement errors. It is defined as

$$\sigma^2_{\mathrm{NXS}} = \frac{1}{N\bar{m}^2} \sum_{i=1}^{N} [(m_i - \bar{m})^2 - \sigma_i^2]. \quad (8)$$

Here we use the symbol σ^2_{NXS} for the normalized excess variance as this or similar symbols are widely used in the literature. Note that σ^2_{NXS} may be negative if the estimated errors σ_i are larger than the actual scatter of measurements, m_i. The fractional root-mean-square variability amplitude, F_{var}, another commonly used X-ray variability measure, is simply a square root of the normalized excess variance: $F_{\mathrm{var}} = \sqrt{\sigma_{\mathrm{NXS}}}$ (Vaughan et al. 2003) if σ^2_{NXS} is positive.

Lawrence & Papadakis (1993) note that in the presence of red noise, the expected value of σ^2_{NXS} depends on the length of a time series. The value of σ^2_{NXS} estimated from a light curve is related to the integral of the power spectral density (PSD) in the frequency range probed by the observations; however, this relation is complex (Allevato et al. 2013) and depends on the PSD slope and sampling (window function).

2.7 Peak-to-peak variability, v

The peak-to-peak variation, v, can be characterized as

$$v = \frac{(m_i - \sigma_i)_{\max} - (m_i + \sigma_i)_{\min}}{(m_i - \sigma_i)_{\max} + (m_i + \sigma_i)_{\min}}, \quad (9)$$

where $(m_i - \sigma_i)_{\max}$ and $(m_i + \sigma_i)_{\min}$ are the maximum and minimum values of the expressions $m_i - \sigma_i$ and $m_i + \sigma_i$ over the entire light curve. This variability index, with minor variations in its definition, is widely used in the radio astronomy community (e.g. Aller, Aller & Hughes 1992; Ciaramella et al. 2004; Hovatta et al. 2008; Fan et al. 2011; Gorshkov, Konnikova & Mingaliev 2012; Majorova & Zhelenkova 2012). It is of interest to compare v with variability characteristics traditionally used in optical and other bands. Here we use the definition of v adopted by Sokolovsky et al. (2009) and Mingaliev et al. (2014). The value of v may be negative if the measurement errors, σ_i, are overestimated (cf. σ^2_{NXS}, Section 2.6)

The peak-to-peak variation may be a sensitive variability indicator if we believe that a light curve is free from outliers (thanks to high data quality or successful filtering). While v can be computed for a light curve consisting of as few as two observations, the expected value of v for a non-variable source depends strongly on the number of measurements. Monte Carlo simulation is a practical way to estimate expected values of v for a non-variable object given a number of observations and their accuracy.

2.8 Lag-1 autocorrelation, l_1

Photometric observations are often planned so that the time span between consecutive flux measurements is smaller than the variability time-scale expected for the objects of interest. The simplest way to characterize the similarity of consecutive flux measurements is to compute the first-order autocorrelation coefficient (also known as 'serial correlation coefficient' or 'lag-1 autocorrelation') of a light curve (e.g. Kim et al. 2011a,b):

$$l_1 = \frac{\sum_{i=1}^{N-1} (m_i - \bar{m})(m_{i+1} - \bar{m})}{\sum_{i=1}^{N} (m_i - \bar{m})^2}. \quad (10)$$

It has been shown that, assuming that m_i are independent measurements subject to identically distributed measurement errors, l_1 follows an asymptotically normal distribution with the expected value of $=-1/N$ and the variance of $\simeq 1/N$, allowing one to assess if the obtained value of l_1 is consistent with the expected one under the above assumptions.

This simple method loses efficiency if a light curve is unevenly sampled since pairs of data points widely separated in time and weakly correlated or uncorrelated entirely contribute to the value of l_1 equally with the pairs of measurements taken close in time that may be well correlated.

2.9 Welch–Stetson variability index I

Welch & Stetson (1993) propose a variability index, I, characterizing the degree of correlation between n quasi-simultaneous pairs of measurements obtained in two filters b and v:

$$I = \sqrt{\frac{1}{n(n-1)}} \sum_{i=1}^{n} \left(\frac{b_i - \bar{b}}{\sigma_{b_i}} \right) \left(\frac{v_i - \bar{v}}{\sigma_{v_i}} \right), \quad (11)$$

where b_i (v_i) are the measured magnitudes, σ_{b_i} (σ_{v_i}) are the estimated errors and $\bar{b}$ ($\bar{v}$) is the mean magnitude in filter b (v).

Relying on the above assumption that a light curve contains pairs of measurements obtained close in time (compared to the expected variability time-scale), one can apply I to a single-band light curve by dividing it into two subsamples that would mimic measurements in two filters. One obvious way to accomplish this is to sort the light curve in time, number measurements (1, 2, 3. . .) and assign measurements having odd numbers to v subsample and even numbers to b subsample. In this case, $\bar{b} = \bar{v}$ may be taken to be the mean of all $N = 2n$ observations, rather than the means of two different samples each of size n.

If a single-filter light curve does not entirely consist of pairs of closely spaced points, one would like to avoid forming pairs from measurements taken far apart in time (cf. l_1 in Section 2.8). In that case, an additional parameter, $\Delta T_{\max}$, defines the maximum time difference between two observations that are considered to be taken sufficiently close in time for forming a pair. The performance of the algorithm on a given unevenly sampled data set depends strongly on the choice of $\Delta T_{\max}$. If $\Delta T_{\max}$ is too small, only few light-curve points will form a pair and contribute to I rendering the index unusable. An optimal value of $\Delta T_{\max}$ would be large enough to form many measurement pairs in an unevenly sampled light curve but small enough to remain sensitive to a wide range of variability time-scales as I is sensitive to variations on time-scales from $\Delta T_{\max}$ to the overall duration of the light curve. A histogram of the interval between observations may be useful in selecting an appropriate $\Delta T_{\max}$ value for a given data set (Ferreira Lopes & Cross 2016).

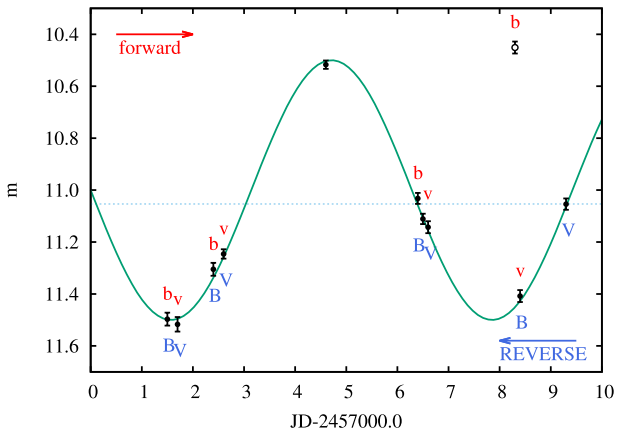

Figure 1. Single-band light curve simulated as $m = 11.0 + 0.5 \sin(\mathrm{JD}-245\,7000.0)$ + noise is divided into subsamples to calculate Stetson's variability indices (Sections 2.9 and 2.10). The arrows indicate the order in which light-curve points are considered: first to last (subsample names in lower case) or reverse (subsample names in upper case). $\Delta T_{\max} = 1\,d$. Dashed line is the weighted average. The weights are iteratively scaled by the factor f, see equation (14), Section 2.10. The open circle is an 'outlier' 1 mag brighter than it should be to follow the sine curve.

Table 2. Simulated light curve divided into subsamples.

JD (d)	m_i (mag)	σ_i (mag)	subsamples
245 7001.500 00	11.497	0.025	bB
245 7001.700 00	11.517	0.028	vV
245 7002.400 00	11.305	0.025	bB
245 7002.600 00	11.246	0.018	vV
245 7004.600 00	10.517	0.016	
245 7006.400 00	11.032	0.021	b
245 7006.500 00	11.111	0.020	vB
245 7006.600 00	11.143	0.023	V
245 7008.300 00	10.451	0.023	b
245 7008.400 00	11.408	0.023	vB
245 7009.300 00	11.054	0.022	V

In our tests, we use $\Delta T_{\max} = 2\,d$ for all the test data sets. Isolated data points that cannot be paired with others (for a given choice of $\Delta T_{\max}$) are omitted from the I computation.

Fig. 1 and Table 2 show how an unevenly sampled single-band light curve can be divided into subsamples to calculate I or J. A point is assigned to subsample b, v or counted as 'isolated' depending on the value of $\Delta T_{\max}$ and the order in which one considers the light curve: from the first point to the last one (direction indicated by the top arrow in Fig. 1, the corresponding samples are named as lower case b and v) or the reverse direction (bottom arrow, capital B and V). Depending on the order, one may compute the 'forward' and 'reverse' values of an index that might differ from each other because the points are divided into pairs (assigned to b and v subsamples) in a different way (as illustrated by Fig. 1). In our implementation of the index, the 'forward' and 'reverse' values are averaged to have a single value describing a light curve. In the case of I (but not J), the 'forward' and 'reverse' values are equal if one allows a point to be counted in multiple pairs (enter two subsamples simultaneously).

The I and J indices are designed to detect variability on time-scales much longer than the typical time difference between observations forming pairs. If, however, the variability time-scale is comparable to the sampling rate of observations, the measurements in pairs may appear anticorrelated (correlation coefficient $l_1 \sim -1$, Section 2.8) rather than uncorrelated ($l_1 \sim 0$), resulting in near-zero or negative value of I (J) and rendering the index insensitive to the variations. The actual value of detectable variability time-scale is determined by the data and will be very different for data sets including observations taken minutes apart and data sets that include only observations taken on different nights.

2.10 Stetson's *J*, *K* and *L* variability indices

A more robust variability index proposed by Stetson (1996) is

$$J = \frac{\sum_{k=1}^{n} w_k \,\mathrm{sgn}(P_k)\sqrt{|P_k|}}{\sum_{k=1}^{n} w_k}, \tag{12}$$

where sgn is the sign function. Here the photometric data set is divided into n groups each consisting of two or more quasi-simultaneous observations (in one or more filters) or a single, isolated measurement. A single-filter light curve can be divided into subsamples to mimic multi-band data in the same way as for the I index (Section 2.9), with the difference that isolated points can be kept in the analysis. Each group consisting of one or more points is assigned a weight w_k. P_k is defined as

$$P_k = \begin{cases} \left(\sqrt{\frac{n_v}{n_v-1}}\,\frac{v_i-\bar{v}}{\sigma_{v_i}}\right)\left(\sqrt{\frac{n_b}{n_b-1}}\,\frac{b_i-\bar{b}}{\sigma_{b_i}}\right) & \text{pair} \\ \frac{n_v}{n_v-1}\left(\frac{v_i-\bar{v}}{\sigma_{v_i}}\right)^2 - 1 & \text{single observation.} \end{cases} \tag{13}$$

The definition of P_k can be generalized for groups containing more than two measurements by multiplying P_k (for a pair) by an additional factor of $\left(\sqrt{\frac{n}{n-1}}\,\frac{r_i-\bar{r}}{\sigma_{r_i}}\right)$, where r_i are the observations in the third filter or subsample. For simplicity, in the implementation of the Stetson indices used throughout this paper, we do not consider groups containing more than two points and do not allow a point to be counted as part of more than one group (see Fig. 1 and Table 2).

Instead of using the weighted arithmetic mean to derive $\bar{v}$, Stetson (1996) suggests to use an iterative re-weighting procedure to down-weight potential outlier points. After computing $\bar{v}$ as the weighted mean at the first step, weights of all points are multiplied by a factor

$$f = \left(1 + \left(\frac{\left|\sqrt{\frac{n_v}{n_v-1}}\,\frac{v_i-\bar{v}}{\sigma_{v_i}}\right|}{a}\right)^b\right)^{-1} \tag{14}$$

and $\bar{v}$ is re-computed with these new weights. The procedure is repeated until it converges.

Many types of variable stars show continuous brightness variations (with notable exceptions, the Algol-type eclipsing binaries and flare stars, which remain at about constant brightness most of the time only occasionally showing large variations). Stetson (1996) suggests to supplement J, which is a measure of the degree of correlation between consecutive brightness measurements, with a robust measure of the kurtosis ('peakedness') of the magnitude histogram:

$$K = \frac{1/N\sum_{i=1}^{N}\left|\sqrt{\frac{n_v}{n_v-1}}\,\frac{v_i-\bar{v}}{\sigma_{v_i}}\right|}{\sqrt{1/N\sum_{i=1}^{N}\left(\sqrt{\frac{n_v}{n_v-1}}\,\frac{v_i-\bar{v}}{\sigma_{v_i}}\right)^2}}. \tag{15}$$

For a Gaussian magnitude distribution, K tends to $K \underset{N\to\infty}{\longrightarrow} \sqrt{2/\pi}$ or will be less if there is an outlier point in the light curve making the overall magnitude distribution more 'peaked'.

The two indices J and K can be combined to the index L (Stetson 1996):

$$L = \sqrt{\pi/2} J K \left(\sum w/w_{\rm all} \right), \tag{16}$$

where $(\sum w/w_{\rm all})$ is the ratio of the weights of all of the light-curve points to a total weight that the star would have if it had been successfully measured on all images. This ratio is designed to reduce the combined variability index L value for stars with a small number of measurements for the reasons outlined in Section 3.7. The combined index is designed to maximize chances of detection for well-measured continuously variable stars. It is less effective for objects that show brightness variations only occasionally (Algol-type binaries, flare stars, transient events).

2.11 Stetson's variability indices with time-based weighting: *J*(time), *L*(time)

Zhang et al. (2003) and Fruth et al. (2012) suggested to weigh the pairs used to compute Stetson's J index (Section 2.10) according to the time difference between the observations used to form a pair:

$$w_i = \exp\left(-\frac{t_{i+1} - t_i}{\Delta t} \right), \tag{17}$$

where t_i is the time of observation i and Δt is the median of all pair time spans $(t_{i+1} - t_i)$. This weighting scheme eliminates the need to choose a specific maximum allowed time difference ($\Delta T_{\rm max}$, Section 2.9) for forming a pair.

2.12 Stetson's variability indices with a limit on the magnitude difference in a pair: *J*(clip), *L*(clip)

The example presented in Fig. 1 shows that it is undesirable to form a pair that would include an outlier point. Considering the assumption that a light curve contains pairs of observations taken close in time (compared to the expected variability time-scale), one can discard from the calculation of I (Section 2.9) or J index (Section 2.10) pairs with magnitude difference between the points greater than a few times the measurement uncertainty. In our tests, we do not form pairs from measurements that differ by more than five times their combined uncertainty, no matter how close in time the two measurements are taken.

2.13 Consecutive same-sign deviations from the mean magnitude (CSSD)

Wozniak (2000) and Shin et al. (2009) suggested to use the number of groups, CSSD, containing three consecutive measurements that are brighter or fainter than the mean (or median) magnitude by at least a factor of $c\sigma$ as a variability indicator. Typically, the value of c is set to 2 or 3. In the algorithm implementation tested in this work, we choose $c = 3$, replace σ with the MAD value scaled to σ (Section 2.3) and use the median as the baseline magnitude level, in order to make the index more robust against outliers. Following Shin et al. (2009), we normalize the number of groups by $(N - 2)$, where N is the number of points in a light curve.

2.14 Excursions, E_x

Plavchan et al. (2008) and Parks et al. (2014) point out that ground-based photometric time series can often be naturally divided into groups (scans) – dense series of observations separated by long gaps. If the variability time-scale is longer than the duration of an individual scan, average magnitudes will differ from scan to scan. Combining observations within a scan to form a single estimate of brightness has an obvious advantage of giving a more accurate estimate (compared to an individual measurement) at the expense of degraded time resolution.

To compare mean magnitudes of the scans, one could perform the analysis of variance (ANOVA, e.g. Kenney & Keeping 1956). However, a light curve obtained with a ground-based telescope is likely to violate the assumptions behind the parametric form of this test. The variance of measurements may differ between the scans (if the observations combined in different scans were performed at different elevations or weather conditions). The distribution of measurements may be non-Gaussian due to outliers. It is tempting to use a non-parametric test (like Mood's median test) to compare scans without having a pre-conception about the measurement error distribution. However, when applied to a typical ground-based photometric data set, such a test would give the (mathematically correct) answer that the majority of stars are variable due to night-to-night photometric zero-point variations.

In the algorithm implementation tested here, we use the absolute difference between the median magnitudes of scans normalized by their combined MADs (Section 2.3) and averaged over all pairs of scans in a light curve to form the variability index E_x. In practice, the exact way a light curve is split into scans has a strong impact on the usefulness of this variability test for a given data set. We divide the light curve into scans according to a pre-defined maximum time difference. This implies that each scan may have a different number of points. For each scan, we compute the median and MAD scaled to σ (Section 2.3) of the observed magnitudes during this scan. The index E_x is computed according to the equation

$$E_x = \frac{2}{N_{\rm scan}(N_{\rm scan}-1)} \sum_{i=1}^{N_{\rm scan}-1} \sum_{j>i,}^{N_{\rm scan}} \frac{\left|{\rm median}_i - {\rm median}_j\right|}{\sqrt{\sigma_i^2 + \sigma_j^2}}, \tag{18}$$

where $N_{\rm scan}$ is the number of scans, $N_{\rm scan}(N_{\rm scan}-1)/2 = C^2_{N_{\rm scan}}$ is the number of two-scan combinations in the data set, ${\rm median}_i$ and σ_i corresponds to median and MAD scaled to σ of the ith scan and the same notation is used for the jth scan, respectively.

2.15 The von Neumann ratio η

The ratio of the mean square successive difference to the distribution variance was discussed by von Neumann (1941, 1942) as an indicator of independence for a series of observations. It is defined as

$$\eta = \frac{\delta^2}{\sigma^2} = \frac{\sum_{i=1}^{N-1} (m_{i+1} - m_i)^2/(N-1)}{\sum_{i=1}^{N} (m_i - \bar{m})^2/(N-1)}. \tag{19}$$

It remains useful even if the observations are drawn from a non-Gaussian distribution as long as it is nearly symmetric (Lemeshko 2006; Strunov 2006).

The ratio η quantifies the smoothness of a time series. Shin et al. (2009) employed η as a variability indicator, noting that since photometric time series measurements do not follow a Gaussian

distribution, in practice, the cut-off value for selecting variable objects cannot be determined a priori (as in the case of χ^2, Section 2.1). One may use $1/\eta$ as a variability indicator to have larger values of the index corresponding to a greater likelihood of an object being variable as is the case with the other variability indices discussed here.

2.16 Excess Abbe value $\mathcal{E}_\mathcal{A}$

Mowlavi (2014) discussed the Abbe value $\mathcal{A} = \eta/2$ and the excess Abbe value

$$\mathcal{E}_\mathcal{A} \equiv \overline{\mathcal{A}_{\rm subi}} - \mathcal{A}, \tag{20}$$

where $\overline{\mathcal{A}_{\rm subi}}$ is the mean of $\mathcal{A}_{\rm subi}$ values computed for all measurements m_i obtained at times t_i. Each $\mathcal{A}_{\rm subi}$ is computed over the sub-interval $[t_i - \frac{1}{2}\Delta T_{\rm sub}, t_i + \frac{1}{2}\Delta T_{\rm sub}]$ ($\Delta T_{\rm sub} < \Delta T$, the overall duration of time series). The choice of $\Delta T_{\rm sub}$ determines the minimum time-scale of variability that may be detected by comparing $\mathcal{A}_{\rm subi}$ to $\mathcal{A}$. $\mathcal{E}_\mathcal{A}$ may be useful to identify unusual behaviour in well-sampled light curves. A large number of measurements (>5 in our implementation) should be taken within the time interval $\Delta T_{\rm sub}$ from each point to accurately determine $\mathcal{A}_{\rm subi}$.

2.17 S_B variability detection statistic

The χ^2 statistic applied to photometric time series data considers only the distribution of the measured magnitudes ignoring the information on when these measurements were obtained. Thus, the χ^2 statistic cannot distinguish between the cases where small-scale deviations in one direction from the mean value are randomly distributed across the light curve from the cases where many of the same-sign deviations are concentrated around a specific time (the second case is less likely to occur by chance).

Figuera Jaimes et al. (2013) suggested a variability detection statistic that combines the advantages of scatter-based and correlation-based variability indices. It is based on the 'alarm' statistic used by Tamuz, Mazeh & North (2006) to assess the quality of fitting binary light-curve models to observational data. Arellano Ferro et al. (2012) applied a similar statistic to detect the Blazhko effect in light curves of RR Lyrae stars. The variability detection statistic is defined as

$$S_B = \left(\frac{1}{NM}\right)\sum_{i=1}^{M}\left(\frac{r_{i,1}}{\sigma_{i,1}} + \frac{r_{i,2}}{\sigma_{i,2}} + \cdots + \frac{r_{i,k_i}}{\sigma_{i,k_i}}\right)^2, \tag{21}$$

where N represents the total number of data points in the light curve and M is the number of groups of consecutive residuals of the same sign from a constant-brightness light-curve model, $r_{i,j} = |m_i - \bar{m}|$ (j is the running number in the group containing k_i same-sign deviations from the mean, $\bar{m}$) and $\sigma_{i,j}$ are the uncertainties corresponding to $r_{i,j}$.

3 TEST DATA SETS

To compare the relative power of the indices (Section 2) in identifying variable objects, we use seven photometric data sets containing a large number of known variable stars (Table 3). The data sets represent a range of sampling patterns and measurement accuracies. Due to the diversity of instruments and reduction strategies, the data sets are characterized by a variety of numbers of badly measured objects that contaminate the lists of candidate variables. Overall, the selected data sets should represent a range of observing conditions typically found in ground-based variability surveys.

Table 3. Test data sets. $N_{\rm var}$ denotes number of variable stars identified in the data set, $N_{\rm stars}$ total number of stars and N maximum number of light-curve points.

Data set	$N_{\rm var}/N_{\rm stars}$	N	Time range	$m_{\rm lim}$	Section
TF1	271/21543	3900	2012-05-14 to 2013-08-19	18 R	3.1
TF2	51/ 8438	8004	2014-09-05 to 2014-11-22	16 R	3.1
Kr	235/29298	1171	2012-08-13 to 2012-10-18	17 v^a	3.2
W1	80/ 2615	242	2006-06-14 to 2006-07-24	19 I	3.3
And 1	124/29043	132	2011-10-31 to 2013-05-23	14 v^a	3.4
SC20	465/30265	268	1997-10-05 to 2000-11-24	21 I	3.5
66 Oph	26/ 6337	227	1976-02-04 to 1995-08-19	17 B^b	3.6

Notes. [a]Unfiltered magnitude calibrated against v zero-point.
[b]Photographic magnitudes calibrated against B zero-point.

The data sets used for our variability indices test were previously searched for variability and contain 1097 known variable objects. While preparing this publication, we manually checked the light curves of all stars standing out in any of the variability indices plotted against the mean magnitude (Fig. 3). We were able to identify 124 variable stars that were missed in the original searches. New variable stars[6] were found in each one of the test data sets. This highlights the fact that variability detection techniques used in previous searches can be improved by adding (a combination of) the variability indices considered here (Section 2).

3.1 The Kourovka Planet Search (TF1, TF2)

As our test data we used observations of two dense sky fields in the Galactic plane conducted within the framework of the Kourovka Planet Search (Burdanov et al. 2016). The first field (TF1) was observed with the MASTER-II-Ural telescope at the Kourovka Astronomical Observatory of the Ural Federal University ($\varphi = 57^\circ$ N, $\lambda = 59^\circ$ E). The mean full width at half-maximum (FWHM) seeing at the site is 3 arcsec. The telescope consists of a pair of Hamilton catadioptric tubes (400 mm *f*/2.5) on a single equatorial mount Astelco NTM-500 without autoguiding. Each tube is equipped with 4098 × 4098 pixels Apogee Alta U16M CCD giving an image scale of 1.85 arcsec pixel^{-1} in a 2 × 2 deg^2 field. The field TF1 is centred at $\alpha_{\rm J2000}$ = 20:30:00 $\delta_{\rm J2000}$ = +50:30:00 (Cygnus). The main observing set of TF1 was completed during short and bright nights from 2012 May to August. We obtained 3900 frames with an exposure time of 50 s in the R filter. The time interval between consequent frames was about 1.5 min. TF1 was observed for 90 h in the R band (36 nights) with an average duration of 2.5 h per observing run.

The second field (TF2) was observed with the Rowe-Ackermann Schmidt Astrograph (RASA) telescope (279 mm *f*/2.2) at the Acton Sky Portal private observatory ($\varphi = 43^\circ$ N, $\lambda = 71^\circ$ W). The telescope is equipped with a 3352 × 2532 pixels SBIG STF-8300M CCD which provides an image scale of 1.79 arcsec pixel^{-1} in a 1.2 × 1.6 deg^2 field. The typical seeing at the site is 2 arcsec. TF2 is centred at $\alpha_{\rm J2000}$ = 02:47:00 $\delta_{\rm J2000}$ = +63:00:00 (Cassiopeia). The RASA telescope obtained about 8000 frames of TF2 in 2014 September–November during all available clear nights. Observations were performed in the R filter with 50 s exposure time. The time interval between the consequent frames is 1 min. The field

[6] Information about known variable stars was extracted from the AAVSO International Variable Star Index (VSX; Watson 2006; https://www.aavso.org/vsx) and VizieR service (http://vizier.u-strasbg.fr/). Variable stars were considered 'new' if no information about their variability could be found in these services.

Table 4. New variable stars found in the test data (the full table is available online).

Name	Alias	α_{J2000} δ_{J2000}	Mag. range	Type	Period (d)	Epoch
ogle_17707	LMC_SC20_17707	05:44:59.79 −70:53:45.1	17.85–17.95 I	SR	86	max 2451164.798
ogle_33977	LMC_SC20_33977	05:45:07.10 −70:38:57.3	19.00–19.3 I	E:/CEP:	8.220	min 2450842.770
ogle_14141	LMC_SC20_14141	05:45:07.95 −70:56:55.9	17.65–17.75 I	SR	34.5	max 2450726.854

was observed for 130 h (18 nights) with an average duration of 7.2 h per night.

Before processing the data, we had to filter out some of the images because not all of them were obtained in optimal weather conditions. We use the standard deviation of image pixel counts $\sigma_{\rm pix}$ as an indicator of weather conditions. The value of $\sigma_{\rm pix}$ varies smoothly from image to image in photometric nights. In the presence of clouds, $\sigma_{\rm pix}$ value of a particular image noticeably decreases (or increases if the clouds are lit by the moonlight).

We used the console version of the ASTROMETRY.NET application (Lang et al. 2010) to set the correct World Coordinate System parameters in the FITS header of each image. The IRAF package (Tody 1986) is then used to perform dark frame subtraction and division by the flat-field. Dark frames are taken before each observing night. Flat-field images are taken during the dawn. The PHOT/APPHOT task is used to perform aperture photometry in each frame with aperture size and sky background level adjusted for each image. The aperture radius is set to 0.8 FWHM of the stellar PSF in the frame. A total of 21 500 and 8500 stars were measured in TF1 and TF2 fields, respectively. The ASTROKIT software (Burdanov et al. 2014) is used to correct for the star brightness variations caused by changing atmospheric transparency. The program selects for each star an individual ensemble of reference stars having similar magnitudes and positions in the frame. We achieved photometric accuracy of 0.005–0.05 mag in the interval 11–16 mag for data from the MASTER-II-Ural telescope. For the RASA telescope data, we achieved precision of 0.006–0.08 mag in the magnitude interval 11–16 mag for the TF2 field. These light-curve data were originally searched for variability by Popov et al. (2015).

3.2 Krasnoyarsk SibSAU 400 mm telescope (Kr)

A 2.3 × 2.3 deg^2 field centred at α_{J2000} = 22:50:00 δ_{J2000} = +52:00:00 (Lacerta) was observed with the 400 mm *f*/2.3 telescope of the Siberian State Aerospace University using the 3056 × 3056 pixels (2.7 arcsec pixel^{-1}) unfiltered CCD camera FLI ML09000. The telescope is installed on top of the University building in the city of Krasnoyarsk. The turbulence caused by the building results in 7–8 arcsec seeing. The observing site is affected by light pollution. A total of 1171 30-s exposures of the field were obtained in 2012 August–October. After applying bias, dark and flat-field corrections using the MAXIM DL software, the images were loaded into VAST[7] (Sokolovsky & Lebedev 2005) for photometric analysis.

After comparing results of aperture and PSF-fitting photometry performed using SEXTRACTOR (Bertin & Arnouts 1996) with PSFEX (Bertin 2011), we discovered that for the brightest stars in the field, the aperture photometry is about a factor of 2 more accurate than PSF photometry probably due to the insufficient accuracy of the reconstruction of PSF variations across the field. We applied six iterations of SYSREM (Tamuz, Mazeh & Zucker 2005; Macfarlane et al. 2015) to remove effects of these variations and bring scatter of PSF-photometry light curves for the bright stars to the level of scatter obtained with the aperture photometry. For the final analysis, we used SYSREM-corrected PSF-photometry light curves as they provide better measurement accuracy for the faint stars compared to fixed-aperture photometry. Only isolated objects with SEXTRACTOR flag = 0 and measured on at least 200 images were considered. The instrumental magnitude scale is calibrated to Cousins $R = v - 1.09^{\star}(r - i) - 0.22$ (Jester et al. 2005) computed from UCAC4/APASS v, r and i magnitudes (Zacharias et al. 2013; Henden et al. 2016) of 2644 stars in the field. These images were originally investigated by Lapukhin, Veselkov & Zubareva (2013, 2016) who used VAST with SEXTRACTOR in the aperture photometry mode and identified variable objects using the σ–mag plot.

3.3 LCO 1 m Swope telescope (W1)

Observations of the Galactic super star cluster Westerlund 1 were obtained during 17 nights between 2006 June 14 and July 24 using the 1 m *f*/7 Henrietta Swope telescope at Las Campanas Observatory, Chile by Bonanos (2007) who identified 129 new variable stars in the field using image subtraction. A 1200 × 1200 pixels section of the 2048 × 3150 SITe CCD (0.435 arcsec pixel^{-1}) corresponding to 8.7 arcmin field of view was read to increase cadence. The initial image processing steps including overscan correction, linearity correction and flat-fielding were performed in IRAF. We reprocessed 242 *I*-band images (including some rejected from the original study due to poor seeing) with VAST, performing PSF-fitting photometry using SEXTRACTOR and PSFEX. The magnitude scale was calibrated using *I*-band magnitudes of 1276 stars in the field measured by Bonanos (2007). We considered only isolated objects (SEXTRACTOR flag = 0) detected on ≥100 images to minimize the effects of crowding. Three cycles of SYSREM are applied to the data. From the list of Bonanos (2007), we select 78 objects showing detectable variability in the *I* band and pass our selection criteria. We add two previously unknown variable objects found during our tests (Table 4, Fig. 2).

3.4 The New Milky Way survey (And 1)

The New Milky Way survey[8] (Sokolovsky, Korotkiy & Lebedev 2014a) hunted for bright (v < 13.5 mag) transients near the Galactic plane using a Canon EF f = 135 mm (*f*/2) telephoto lens attached to an unfiltered 3352 × 2532 SBIG ST-8300M CCD camera (8.4 arcsec pixel^{-1}, 8 × 6 deg^2 field). The observations were conducted in 2011–2013. We used 132 images of the field centred at α_{J2000} = 23:00:00 δ_{J2000} = +50:00:00 (And 1)[9] reprocessed with VAST and SEXTRACTOR in the aperture photometry mode accepting blended stars for the analysis (SEXTRACTOR flag ≤ 3). Since the CCD chip is blue-sensitive, APASS v-band magnitudes of 1200 UCAC4 stars within the field of view are used for magnitude calibration.

[7] http://scan.sai.msu.ru/vast/

[8] http://scan.sai.msu.ru/nmw/

[9] The And 1 field fully includes the deeper SibSAU 400 mm Lacerta field described in Section 3.2.

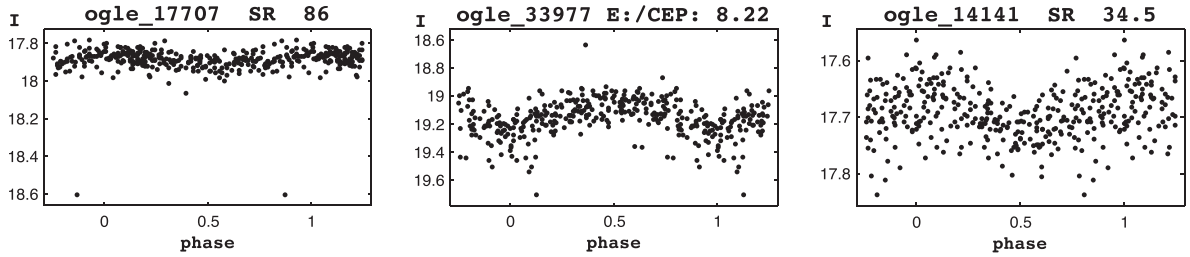

Figure 2. Light curves of new variable stars found in the test data sets (the complete figure is available online). The magnitudes measured in a band indicated in the top-left corner of each panel are plotted as a function of time (Julian day) for irregular or phase for periodic variables. The title of each panel indicates the object identifier in Table 4, its variability type and period in days (if applicable).

Three cycles of SYSREM were applied to the data in order to mitigate systematic effects caused by chromatic aberration of the lens and changing atmospheric extinction across the large field of view. We used three SYSREM cycles as adding more cycles did not further improve (reduce) light-curve scatter for the majority of objects in this data set.

Light curves and images of all objects that stand out in index versus mag plots were visually inspected for variability. We identified 91 known and 33 previously unknown variable stars (Table 4, Fig. 2). The list of detectable variable stars in the field may be incomplete as we accepted only those red objects showing slow irregular variability that are either matched with a known variable star or their variability can be confirmed from ROTSE-I/NSVS (Woźniak et al. 2004) or SuperWASP (Butters et al. 2010) archival data. This should safeguard us from mistaking for real variability any residual colour-related systematics not removed by SYSREM. An example of such residual systematic variation is the dip around JD245 6000 visible in the light curves of many red (SR- and LB-type) variables in the field (Fig. 2). As the final check, we repeat the processing using elliptical aperture of size and orientation that are individually tuned for each object (SExtractor parameter MAG_AUTO). This allows us to recover the flux of defocused red stars at the cost of reducing photometric accuracy for the well-focused point sources and make sure that for the selected variable star candidates the MAG_AUTO light-curve shape is not contradicting the one obtained with a circular aperture of a size fixed for all objects in a given image.

3.5 OGLE-II (LMC_SC20)

The OGLE utilizes the dedicated 1.3 m Warsaw telescope at the Las Campanas Observatory, Chile, to conduct a photometric survey of dense stellar fields in the Magellanic Clouds and Galactic bulge (Udalski, Kubiak & Szymanski 1997). We extract data from the second phase of the experiment OGLE-II PSF *I*-band photometry data base (Szymanski 2005). For the variability index tests, we select one field in the Large Magellanic Cloud, LMC_SC20, which is least affected by crowding. To keep the number of selected sources below the limit of 50 000 imposed by the data base's web-interface and retrieve only high-quality light curves, we selected sources having the percentage of good measurements Pgood $\geq$ 98. In total, 30 265 sources in this field satisfy the selection criteria each having from 262 to 268 photometric measurements. The data set contains 168 variable stars 20 of which (see Table 4, Fig. 2) were not previously known. The new variable stars were identified by visual inspection of the light curves standing out in variability index versus mag plots. To make sure the detected variability is not caused by nearby bright variable stars, we visually checked PSF-fitting light curves of stars located within 20 arcsec of each of the new variables. Only stars brighter than the variable were considered and no limit on the percentage of good measurements was applied.

The use of a fixed centroid position when conducting photometry may introduce spurious long-term variability if the measured star has a detectable proper motion. If the DIA is used, the moving star will have a characteristic dipole shape in the residual image, resulting in detection of two spurious variable sources apparently changing brightness in opposite directions (Eyer & Woźniak 2001). To make sure the variability of ogle_43681, ogle_63585 and ogle_72706 is not caused by the proper motion, we (i) check that there are no records in the OGLE-II DIA catalogue by Zebrun et al. (2001) within 3 arcsec of the new variables and (ii) manually check OGLE-II DIA light curves of nearby sources to make sure none of them show brightness trends mirroring the new variables.

3.6 Digitized photographic plates (66 Oph)

Photographic images of the sky obtained in late 19th and 20th centuries contain a wealth of information about historical positions (e.g. Laycock et al. 2010; Vicente et al. 2010; Berezhnoi 2013; Robert et al. 2014) and brightness (e.g. Kolesnikova et al. 2008; Tang et al. 2013; Sokolovsky et al. 2014c) of celestial objects. The efficient use of this information requires it to be converted to a digital form using purpose-built digitizing machines (Simcoe et al. 2006; De Cuyper et al. 2012) or a commercially available flatbed scanners capable of working with transparent materials (Vicente, Abad & Garzón 2007; Simcoe 2009; Tuvikene et al. 2014).

We use an Epson Expression 11000XL flatbed scanner operating at 2400 dpi resolution (1.4 arcsec pixel^{-1}, 16 bits per pixel colour depth) to digitize a 1.26 $\times$ 1.26 deg^2 area centred at α_{J2000} = 17:57:44.7 δ_{J2000} = +04:59:54 (66 Oph field; Kolesnikova et al. 2010) on 227 photographic plates obtained in 1976–1995 with the 40 cm astrograph. The digitized images were processed with VAST following the procedure described by Sokolovsky et al. (2014b). APASS *B*-band photometry of 1600 UCAC4 stars in the magnitude range B = 10–16 is used to calibrate the instrumental magnitude scale using the relation between aperture photographic and photoelectric magnitudes proposed by Bacher, Kimeswenger & Teutsch (2005). We identify 23 variable stars including five not previously known (Table 4, Fig. 2) by means of period search and visual inspection of light curves standing out in the magnitude versus σ_w plot.

3.7 Light-curve filtering

Often raw photometric data have to be pre-processed before computing the variability indices discussed in Section 2. This may include (i) removing outliers from a light curve (possibly by applying

iterative σ-clipping or median filtering); (ii) removing systematic effects from a set of light curves by applying local zero-point corrections (e.g. Nascimbeni et al. 2014) and/or the SYSREM algorithm, decorrelating each light curve with external parameters such as airmass, seeing, object position on a CCD, detector temperature (e.g. Pál 2009; Bakos et al. 2010; López-Morales et al. 2010; Hartman et al. 2011; Burton et al. 2012; Guterman, Mazeh & Faigler 2015; Baade et al. 2016) or de-trending the light curves if one is interested only in fast variability (e.g. Kovács, Bakos & Noyes 2005; Weingrill 2015).

A smaller-than-expected number of detections is an indirect indication of many photometry problems including the object being close to an image edge, a cosmetic defect, a bright star, a detection or saturation limit. Objects systematically affected by any of these problems can be removed from the analysis by discarding light curves having less than a given number of points. The obvious disadvantage is that together with problematic objects, one may discard a transient object that appears only on a small number of images. The power of discarding light curves with a small number of measurements to improve the overall quality of a photometric data set might be the reason why 'variable star detection' and 'optical transient detection' are traditionally considered as two separate technical problems.

From all the data sets considered in this work, we discard light curves having fewer than 40 points, unless indicated otherwise. We apply no σ-clipping to the test data; however, we note that it can considerably improve performance of variability indices that are not robust to outliers. The SYSREM algorithm is applied to the data sets described in Sections 3.2, 3.3 and 3.4. For the other data sets, it does not lead to a noticeable reduction in light-curve scatter.

3.8 Simulated data sets

The data sets described above (Section 3) include in total 1251 variable stars of various types, but this list still provides us limited coverage of a possible range of variability amplitudes and timescales. We overcome this limitation by adding simulated variability to the test data. Following Enoch et al. (2012), we use light curves of non-variable stars as realistic photometric noise models. This approach has an advantage over simple bootstrapping[10] in that it preserves the correlated nature of the noise. It naturally requires a set of constant stars to have multiple realizations of the noise process while the bootstrapping can be applied to an individual light curve.

From each set of light curves described above, we remove the known variable stars and introduce artificial variability to the remaining stars that are presumed to be constant. Among these constant stars, there are both well-measured ones and some affected by blending or other sources of large photometric errors. According to the simulation parameters, each star has a 1 per cent chance to be variable with a random peak-to-peak amplitude uniformly distributed between 0 and 1.0 mag. The simulation is done in two versions: in version 1 all variables are assumed to be periodic while in the second version they are all assumed to be aperiodic.

We model periodic variability by adding a simple sine signal (e.g. Mislis et al. 2016) to the observed light curve of a constant star. The signal phase is chosen randomly for each simulated variable star. The frequency of the sine signal is drawn from a uniform random distribution in the range 0.05–20 d^{-1}. This results in a large fraction of variables with periods <1 d approximately resembling the period distribution typically found in the Galactic field.

To simulate non-periodic variability, we sum up 10 000 sine waves with logarithmically spaced frequencies in the range 0.0001–1000 d^{-1} and having random phases. The amplitude of each sine wave is the square root of the power spectrum value. If the real and imaginary parts of the Fourier transform of the light curve are independent and vary according to the Gaussian distribution (Campante 2012), the resulting power of the sine waves is varying according to the χ^2_2 distribution with 2 degrees of freedom (Timmer & Koenig 1995; Emmanoulopoulos, McHardy & Papadakis 2013) around the expected values. The expected values in our simulation are defined by a power law with the slope of -1 (e.g. Max-Moerbeck et al. 2014). The exact choice of the power-law slope in the range -0.5 to -3.0 has minimal effect on the following discussion. The simulations are repeated 1000 times for each data set and the averaged results are reported.

4 COMPARISON TECHNIQUE

To select the variability index that is the most efficient in identifying variable stars, we compute the indices defined in Section 2 for all light curves in the test data sets (Section 3). The variable objects have to be distinguished from two broad types of interlopers: non-variable objects and objects with corrupted photometry. To quantify the performance of each index following Kim et al. (2011b) and Graham et al. (2014), we compute the completeness C and purity P:

$$C = \frac{\text{Number of selected variables}}{\text{Total number of confirmed variables}} \quad (22)$$

$$P = \frac{\text{Number of selected variables}}{\text{Total number of selected candidates}} \quad (23)$$

as well as the fidelity F_1-score[11] which is the harmonic mean of the two parameters:

$$F_1 = 2(C \times P)/(C + P). \quad (24)$$

F_1 reaches a maximum of 1.0 for a perfect selection when all confirmed variables and no false candidates pass the selection criteria while $F_1 = 0$ if no confirmed variables are selected.

For each variability index, A, described in Section 2, we compute its expected value $\bar{A}$ and dispersion σ_A as functions of magnitude. The operation is performed for each data set that includes real (Section 3) and simulated (Section 3.8) variable objects. For each point in the magnitude versus index (mag–A) plot, we use points within ± 0.25 mag from it to compute $\bar{A}$ as a median of indices within the magnitude bin. If the bin contains <40 points, its width is increased to include at least 40 points. The expected dispersion, σ_A, is computed as the MAD scaled to σ (Section 2.3) for the points in the bin. After completing these computations for all the points in the magnitude–index plot, the estimated values of $\bar{A}$ and σ_A are smoothed with a simple running average. The robust estimators of $\bar{A}$ and σ_A are necessary considering that a bin is likely to contain

[10] Here by bootstrapping we mean shuffling JD–magnitude pairs in a light curve to eliminate any correlated variability. The method is often used to assess the significance of a periodogram peak (e.g. Barclay et al. 2011).

[11] The three parameters are often referred to as 'recall' or 'sensitivity' or 'true positive rate', 'precision' and 'F-factor' for C, P and F, respectively. See https://en.wikipedia.org/wiki/Precision_and_recall, https://en.wikipedia.org/wiki/F1_score and https://en.wikipedia.org/wiki/Information_retrieval

variable or badly measured objects that have variability index values not typical for constant stars.

Variable star candidates are selected as objects having a variability index value deviating by more than $a\sigma_A$ from the value $\bar{A}$ expected at this magnitude, where a is a factor defining the variability detection threshold (Fig. 3). This approach is similar to the one employed by Barclay et al. (2011) who selected periodic variable stars using a cut in false alarm probability (FAP)–period space. The authors used the median and MAD as robust estimators of the expected FAP value and its scatter as a function of a period. Unlike Villforth et al. (2010), we compare the variability indices not at some specific cut-off level a common for all indices, but instead choose the optimal value of a individually for each index as described below.

For each index and data set, we compute C, P and F_1 parameters as functions of a (Fig. 4). For some optimal value of a, F_1 reaches the maximum, $F_{1\max}$, corresponding to a trade-off between the completeness and purity of the selected list of candidates. We consider the index with the highest value of $F_{1\max}$ as the most efficient in selecting true variable stars in a given data set. By comparing results for various data sets (Sections 3 and 3.8), we draw general conclusions about which indices perform better under a wide range of observing conditions (Section 5). Since F_1 characterizes only the list of selected candidates and does not consider the rejected ones, we also report a fraction of objects that do not pass the selection (at the cut-off value corresponding to $F_{1\max}$), R, as a supplementary measure of variability index performance.

5 RESULTS AND DISCUSSION

5.1 Overall performance comparison

Fig. 3 presents the variability index–magnitude plots. The completeness, purity and F_1-score as a function of the cut-off limit, $a\sigma_A$, are presented in Fig. 4. Table 5 lists the highest F_1-score, $F_{1\max}$, and the corresponding fraction of rejected objects, R, for each index and data set described in Section 3. Tables 6 and 7 (available online) present this information for the simulated data sets discussed in Section 3.8.

While performance of each individual index varies considerably between the data sets, the correlation-based indices I (Section 2.9), J, L (including their time-weighted and clipped versions; Sections 2.10, 2.11, 2.12) and $1/\eta$ (Section 2.15) typically provide higher $F_{1\max}$ values than scatter-based indices. Among the scatter-based indices, the IQR (Section 2.4) and MAD (Section 2.3) show the highest $F_{1\max}$ values with RoMS (Section 2.5), σ_w (Section 2.2) and $\chi^2_{\rm red}$ (Section 2.1) falling slightly behind due to their sensitivity to individual outlier measurements. The l_1 (Section 2.8), S_B (Section 2.17), E_x (Section 2.14) and E_A (Section 2.16) perform well in some data sets, but not in the others and, therefore, cannot be recommended as general-purpose variability detection statistics. The indices $\sigma^2_{\rm NXS}$ (Section 2.6) and υ (Section 2.7) typically reach smaller $F_{1\max}$ values compared to the other scatter-based indices.

The CSSD index (Section 2.13) in our implementation appears practically useless for variable objects detection. The requirement for three consecutive data points to be $2\sigma_{\rm MAD}$ brighter or fainter than the median brightness where $\sigma_{\rm MAD}$ is the σ scaled from the light-curve MAD (Section 2.3) appears to be too strict. Indeed, Wozniak (2000) used individual measurement errors to compute CSSD while Shin et al. (2009) used the light-curve σ to compute CSSD (similar to our implementation), but it was only one of the many light-curve features used simultaneously for variability detection in that work.

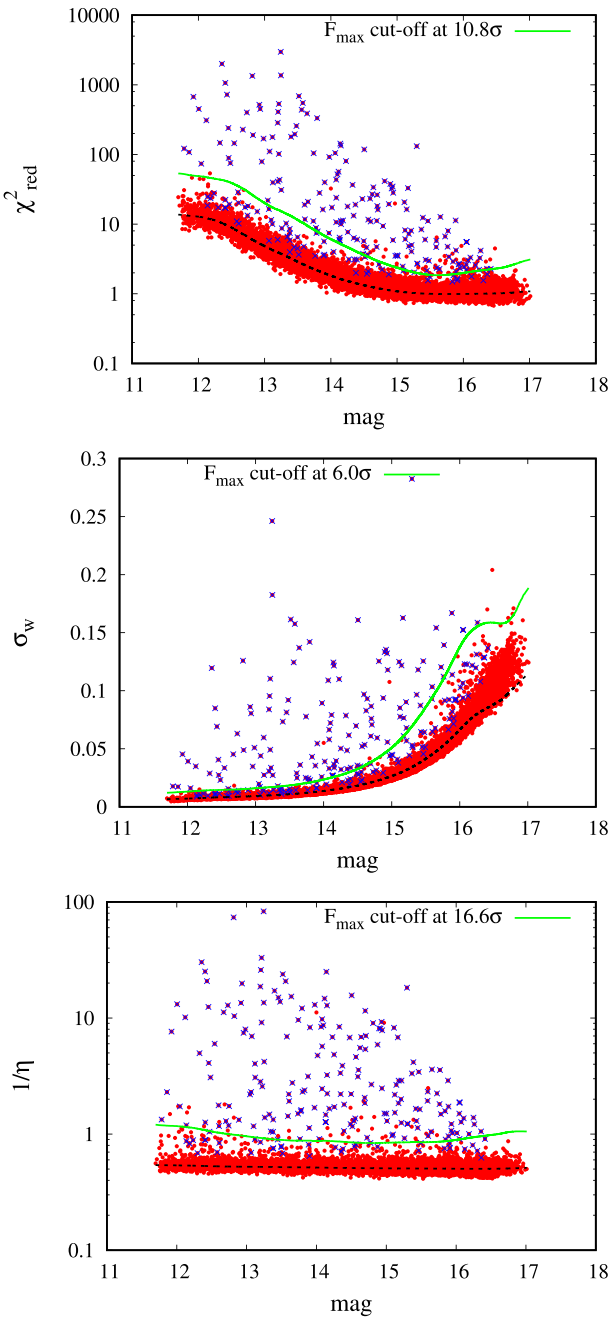

Figure 3. Variability indices $\chi^2_{\rm red}$ (Section 2.1), σ_w (Section 2.2) and $1/\eta$ (Section 2.15) plotted as a function of magnitude for the Krasnoyarsk data set (Section 3.2). Variable stars are marked with 'x'. The curves represent the expected value of $\chi^2_{\rm red}$, σ_w and $1/\eta$ for a given magnitude and the selection threshold corresponding to the best trade-off between the completeness and purity of the candidates list ($F_{\max}$; see Section 4, Fig. 4). The index–magnitude plots for the other indices and data sets may be found online at http://scan.sai.msu.ru/kirx/var_idx_paper/.

The $1/\eta$ appears to be the best compromise index as it performs better than most of the other discussed indices in all tested data sets (real and simulated) judging both from $F_{1\max}$ and R values. The $1/\eta$ index is sensitive only to variability on time-scales longer than the sampling time which causes it to miss fast variables in sparsely sampled data sets like LMC_SC20 (Section 3.5) and 66 Oph (Section 3.6). If the data set has no measurements taken very close in

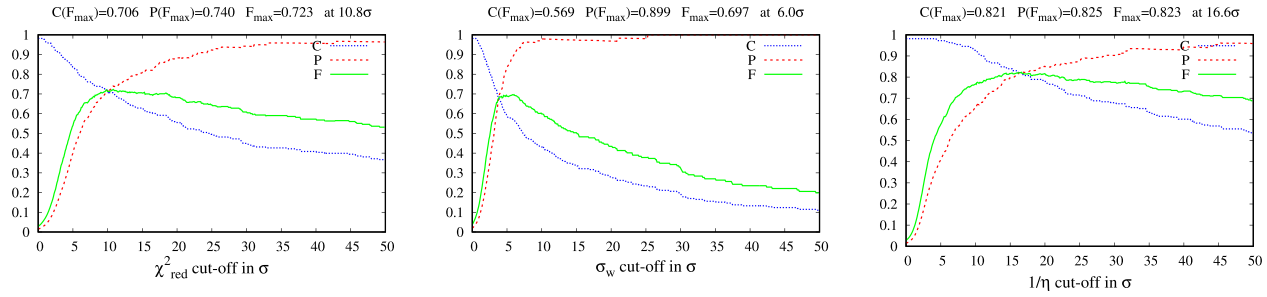

Figure 4. Variable star selection completeness (C), purity (P) and F_1-score (F; see Section 4) as a function of selection threshold for the variability indices $\chi^2_{\rm red}$ (Section 2.1), σ_w (Section 2.2) and $1/\eta$ (Section 2.15) computed for the Krasnoyarsk data set (Section 3.2). C-, P- and F_1-score plots for the other indices and data sets may be found online at http://scan.sai.msu.ru/kirx/var_idx_paper/.

Table 5. Performance of variability indices in selecting real variable stars.

	TF1		TF2		Kr		Westerlund 1		And 1		LMC_SC20		66 Oph			
Index	$F_{1\max}$	R	$F_{1\max}$	R	$F_{1\max}$	R	$F_{1\max}$	R	$F_{1\max}$	R	$F_{1\max}$	R	$F_{1\max}$	R	Section	Ref.
Scatter-based indices																
$\chi^2_{\rm red}$	0.110	0.902	0.076	0.884	0.723	0.993	0.270	0.969	0.556	0.996	0.284	0.996	0.192	0.992	2.1	(a)
σ_w	0.114	0.899	0.076	0.879	0.697	0.995	0.264	0.950	0.544	0.996	0.254	0.994	0.155	0.988	2.2	(b)
MAD	0.161	0.927	0.086	0.940	0.710	0.994	0.287	0.940	0.582	0.996	0.483	0.994	0.375	0.996	2.3	(c)
IQR	0.162	0.927	0.086	0.951	0.726	0.994	0.298	0.945	0.608	0.996	0.470	0.992	0.383	0.997	2.4	(d)
RoMS	0.130	0.917	0.070	0.922	0.729	0.993	0.270	0.963	0.563	0.996	0.382	0.993	0.381	0.997	2.5	(e)
$\sigma^2_{\rm NXS}$	0.026	0.198	0.012	0.197	0.047	0.731	0.059	0.522	0.032	0.752	0.034	0.754	0.324	0.992	2.6	(f)
v	0.053	0.835	0.032	0.901	0.347	0.996	0.140	0.984	0.450	0.997	0.049	0.899	0.098	0.994	2.7	(g)
Correlation-based indices																
l_1	0.370	0.992	0.175	0.999	0.400	0.995	0.188	0.935	0.569	0.996	0.470	0.996	0.450	0.998	2.8	(h)
I	0.116	0.896	0.082	0.891	0.819	0.993	0.281	0.973	0.611	0.994	0.500	0.996	0.341	0.997	2.9	(i)
J	0.144	0.927	0.079	0.931	0.819	0.993	0.286	0.977	0.628	0.994	0.448	0.994	0.368	0.998	2.10	(j)
J(time)	0.152	0.931	0.081	0.932	0.819	0.992	0.291	0.975	0.659	0.995	0.519	0.996	0.410	0.998	2.11	(k)
J(clip)	0.134	0.922	0.074	0.917	0.788	0.993	0.267	0.977	0.587	0.995	0.375	0.991	0.364	0.997	2.12	(d)
L	0.169	0.923	0.092	0.942	0.821	0.992	0.283	0.979	0.706	0.996	0.470	0.994	0.571	0.997	2.10	(j)
CSSD	0.231	0.957	0.105	0.977	0.014	0.008	0.034	0.013	0.008	0.007	0.011	0.012	0.008	0.001	2.13	(l)
E_x	0.181	0.973	0.090	0.998	0.347	0.997	0.159	0.983	0.500	0.997	0.357	0.996	0.263	0.998	2.14	(m)
$1/\eta$	0.549	0.991	0.414	0.992	0.823	0.993	0.378	0.982	0.588	0.997	0.471	0.997	0.424	0.999	2.15	(n)
$\mathcal{E}_\mathcal{A}$	0.154	0.962	0.156	0.995	0.434	0.997	0.250	0.989	0.151	0.994	0.228	0.997	0.133	0.999	2.16	(o)
S_B	0.146	0.893	0.092	0.891	0.766	0.992	0.261	0.982	0.463	0.993	0.303	0.989	0.246	0.995	2.17	(p)
α_1	0.112	0.878	0.078	0.875	0.782	0.994	0.245	0.961	0.639	0.995	0.441	0.994	0.426	0.997	5.4	(d)

References: (a) de Diego (2010), (b) Kolesnikova et al. (2008), (c) Zhang et al. (2016), (d) this work, (e) Rose & Hintz (2007), (f) Nandra et al. (1997), (g) Brown et al. (1989), (h) Kim et al. (2011a), (i) Welch & Stetson (1993), (j) Stetson (1996), (k) Fruth et al. (2012), (l) Shin et al. (2009), (m) Parks et al. (2014), (n) Shin et al. (2009), (o) Mowlavi (2014), (p) Figuera Jaimes et al. (2013).

time (compared to the fastest expected variability time-scale), the $1/\eta$ index sensitive to slow variations should be complemented with a scatter-based index such as the IQR (Section 2.4) that would pick fast variables missed by $1/\eta$.

5.2 Performance based on the number of points in a light curve

The results presented in Tables 5–7 allow us to identify indices that perform well in all the test data sets (Section 5.1). All the test data sets are well sampled containing hundreds to thousands of observations. The question remains how well these indices perform on light curves containing a smaller number of points? This is especially interesting considering that the alternative period search-based methods of variability detection (not considered in this work) are ineffective for light curves having a small number of points. A rule of thumb is that $\gtrsim$100 points randomly sampling a light curve (Graham et al. 2013 suggest $\gtrsim$200 for CRTS sampling) are often sufficient to determine a variable star period. A smaller number of points may be sufficient if the sampling is favourable or the range of possible periods is constrained by prior knowledge of the variability type. If the number of observations is too small to attempt a periodicity search, variability indices are the best hope for identifying variable stars among such undersampled light curves.

To test this, we use the OGLE-II LMC_SC20 data set described in Section 3.5 that is characterized by quasi-random sampling (i.e. it includes a small number of measurements taken on the same night). We randomly select a subset of N observations from the LMC_SC20 data to construct an artificial data set and test how many known variables can be recovered using the same technique as applied to the full data set (Section 4). The results of index comparison are presented in Fig. 5. While σ_w does not show a strong dependence on the number of points, the $F_{1\max}$-score of $1/\eta$ and l_1 linearly increases with increasing number of points in a light curve. The IQR at $N \lesssim 15$ shows $F_{1\max}$ values similar to σ_w, but it shows larger $F_{1\max}$ values for a larger number of points. The reason for IQR being more efficient than σ_w for large N is that the IQR is insensitive to outlier

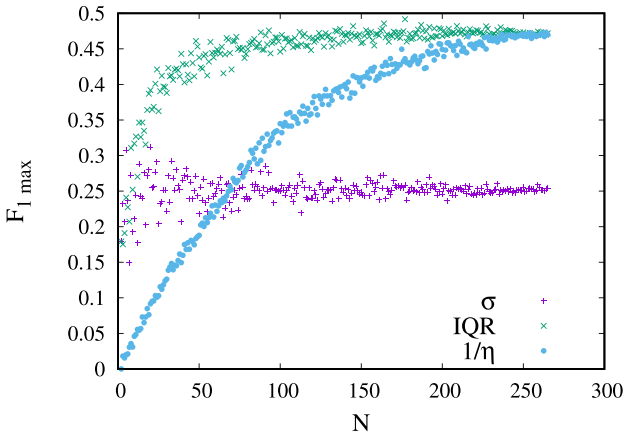

Figure 5. The $F_{1\max}$-score as a function of the number of light-curve points. For each N, the random selection of N points is repeated 10 times and the median $F_{1\max}$ value is plotted.

measurements. Stetson's J (and L) indices, MAD and RoMS also behave similarly to the IQR as these indices can characterize the light-curve scatter while remaining relatively insensitive to outliers. The Welch–Stetson I index becomes useful only for a large number of points because only in this way there are light-curve points obtained close enough in time to form pairs (unlike the J index, I cannot take into account the individual, unpaired measurements). S_B does not show a strong dependence of its $F_{1\max}$ values on the number of points, while $F_{1\max}$ values of E_A slowly increase with increasing N. Overall, we can conclude that the indices characterizing the light-curve scatter perform well even on undersampled light curves while the indices that are purely correlation-based linearly increase their effectiveness with increasing number of light-curve points.

5.3 Correlation between variability indices

Many of the variability indices considered above reflect the same information, just in a slightly different way. Consider, for example, the three versions of Stetson's J index described in Sections 2.10–2.12 which essentially differ from each other only in the relative weights assigned to various pairs of observations.

To quantify the degree of similarity between the indices, we computed the Pearson product-moment correlation coefficient, r, for all possible pairs of indices using the full data sets (i.e. indices computed for variable and non-variable objects were considered together). The linear Pearson correlation coefficient of two variables measures the degree of linear dependence between the variables. It is defined as the ratio of the covariance of the two variables to the product of their standard deviations. It is a direct measure of how well two sample populations vary jointly. It ranges from -1 (total anticorrelation) to 1 (total correlation). A zero value corresponds to a lack of linear correlation (however, non-linear correlations may exist).

The majority of variability indices considered in Section 2 are strongly ($r > 0.8$) correlated with each other. The exceptions are l_1, CSSD, $1/\eta$, $\mathcal{E}_{\mathcal{A}}$. This suggests that the correlated variability indices reflect mostly the same information. This is understandable considering that the indices quantifying the degree of correlation between consecutive brightness measurements are also sensitive to the overall light-curve scatter (with the exception of l_1).

5.4 Principal component analysis

To further quantify the relative importance of the variability indices and to search for a possible linear combination of indices that may be a better variability indicator than any individual index, we performed the principal component analysis (PCA; Pearson 1901).

PCA is an unsupervised, non-parametric method that provides a linear orthogonal transformation of a data set into a new base, where the data variance (assumed to represent the useful information) is highlighted. The new set of (uncorrelated) 'optimal' axes is called the principal components (PCs). The original data can be expressed as a linear combination of the PCs. Usually, very few of the PCs (even two to three of them) are capable of describing the data in terms of variance without a significant loss of information. This dimensionality reduction/data compression is the reason why PCA is very effective in extracting information from huge data sets. However, the results should be interpreted with caution, since the data may not reflect uncorrelated physical phenomena. PCA is extensively used in astronomy, e.g. in applications on stellar spectra (Bailer-Jones, Irwin & von Hippel 1998; Re Fiorentin et al. 2007), on galaxy spectra (Yip et al. 2004; Karampelas et al. 2012), on spectroscopic imaging (Steiner et al. 2009), etc. It was suggested as a variability detection tool for photometric data sets containing quasi-simultaneous multi-colour observations (Eyer 2006; Süveges et al. 2012).

The PCA implementation on an (n observations) $\times$ (m features) data set involves (i) the construction of either (usually) the data variance–covariance matrix or the correlation matrix, (ii) the calculation of the respective eigenvectors PCi (the principal components) and (iii) the calculation of the admixture coefficients α_i, which are the data coordinates on the new axes.

Thus, each original observation x is decomposed on to the new set of axes PCi as

$$x = \sum_{i=1}^{m} \alpha_i \cdot \mathrm{PCi}.$$

The first principal component PC1 summarizes the majority of the data variance (the most widespread information), PC2 summarizes the majority of the rest of the data variance, being uncorrelated to PC1, etc. It is expected that low-order PCs correspond to rare/weak processes, noise, etc. (Tso & Mather 2001).

PCA was applied to each of the test data sets (Section 3). The variability indices of the sample's stars were normalized by their expected value $\bar{A}$ and scatter σ_A as a function of magnitude as discussed in Section 4. Since the indices represent different, albeit often correlated, characteristics and PCA is data dependent, we performed a zero mean and unit variance standardization prior to the analysis. Additionally, the variance–covariance matrix of the data was used. PCA was implemented in IDL (PCOMP procedure).

We consider the first three PCs. For the Kr data set, we find that PC1 is responsible for 56.5 per cent of the data variance, PC2 for 8.2 per cent and PC3 for 7.1 per cent. The distribution of variance between the first principal components for the other test data sets is very similar. The admixture coefficients corresponding to the first principal components are presented in Fig. 6. Variable objects tend to have large positive values of α_1, while they may have any α_2 and α_3 values. This suggests that most of the information related to variability in general is encoded in PC1. The components PC2 and PC3 may encode light-curve characteristics that differ for different variability types. Fig. 7 presents the relative contribution of the variability indices to the first three PCs. While many scatter- and correlation-based indices provide comparable contribution to PC1,

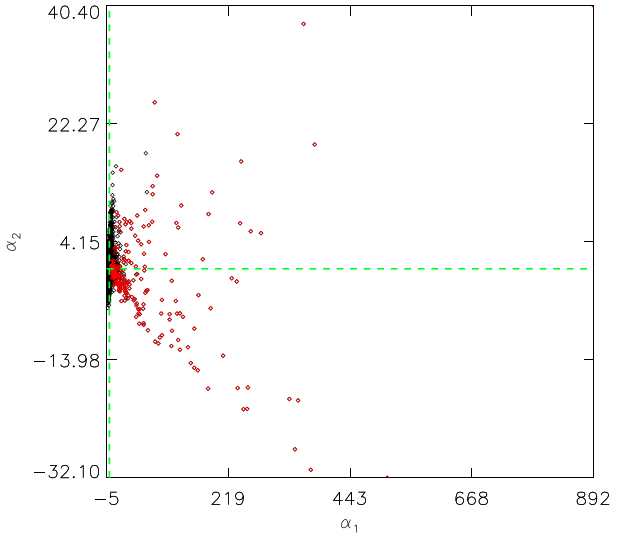

Figure 6. The admixture coefficients corresponding to PC1 (α_1), PC2 (α_2) for the Kr data set (Section 3.2). Variable stars are marked in red. Similar plots for PC3 (α_3) and the other data sets may be found at http://scan.sai.msu.ru/kirx/var_idx_paper/.

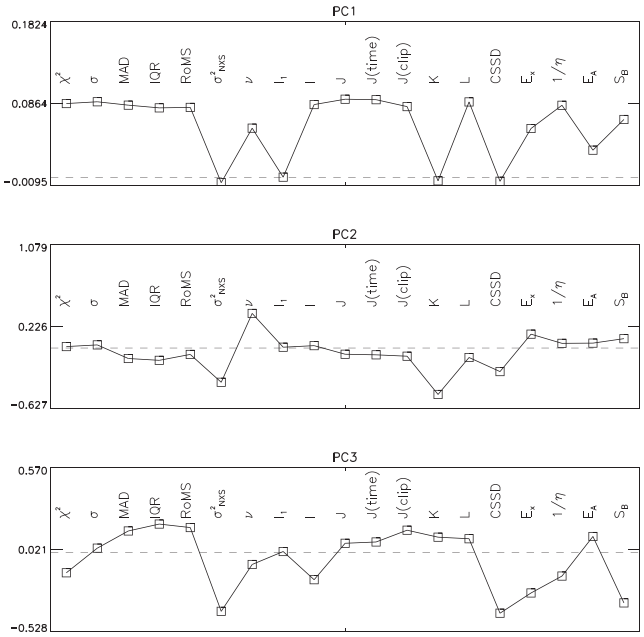

Figure 7. The first three PCs in the Kr data set (Section 3.2). The dashed line indicates zero contribution of an index to the PC. Similar plots for the other data sets may be found at http://scan.sai.msu.ru/kirx/var_idx_paper/.

the indices l_1 and $\mathcal{E}_\mathcal{A}$ contribute less, and the contribution of K, CSSD, $\sigma^2_{\rm NXS}$ is near zero. PC1 is dominated by the indices that generally perform better in identifying variable objects (Section 5). The indices l_1, S_B, $\chi^2_{\rm red}$ contribute the most to PC2 while K, v and $\sigma^2_{\rm NXS}$ dominate PC3.

The admixture coefficient α_1 may be used as a composite variability index since all variable objects tend to have large positive values of α_1 (Fig. 6). It reaches the value of $F_1 = 0.659$ (Fig. 8) and $R = 0.995$ on par with the best variability indices for this field (Table 5), but does not provide an improvement over them. One possible use of α_1 is to investigate a new data set for which it is not known a priori which variability indices are most suitable. In this

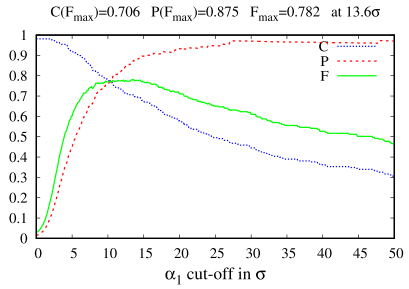

Figure 8. Variable star selection completeness (C), purity (P) and F_1-score (F; see Section 4) as a function of selection threshold for the admixture coefficient α_1 used as a composite variability index (Section 5.4) computed for the Kr data set (Section 3.2). Similar plots for the other data sets may be found online at http://scan.sai.msu.ru/kirx/var_idx_paper/.

case, one could compute multiple indices and perform the PCA of them. The coefficient α_1 is by construction one of the best variability indices (that captures most of variability-related information) for this particular data set.

5.5 Limitations of the indices as variability indicators

Besides the random errors (caused by the background and photon noise[12]) that are usually easy to estimate, photometric measurements are subject to systematic error (due to atmospheric and instrumental variation) that are hard to quantify. Since the overall measurement errors are not accurately known, it is not possible to apply the χ^2 test (Section 2.1) to select (non-)variable objects. The absence of accurate error estimates can be substituted with the assumptions that (i) the majority of field stars are non-variable and (ii) stars of similar brightness in a given field are measured with about the same photometric accuracy. If these assumptions hold, the field stars may be used to measure the actual accuracy of a given set of photometric observations. The variability indices (Section 2) can be used to select objects showing larger-than-expected brightness variations.

Since source extraction is not perfect, in practice there are some objects measured with far worse accuracy than the majority, breaking the assumption (ii) above. The source extraction problems may be caused by blending and image artefacts. Neither scatter- nor correlation-based indices are effective in distinguishing true variable objects from the ones with corrupted photometry, which ultimately limits the usefulness of variability indices. The number of bad measurements in a photometric data set has a higher impact on the efficiency of variability search than the choice of a particular variability index. This is illustrated by comparison of variability search results in the data sets TF1/TF2 (Section 3.1) and Kr (Section 3) obtained with similar equipment. The Kr data set in which bad measurements are aggressively removed provides systematically higher $F_{1\rm max}$-scores than the TF1/TF2 data sets in which no flagging of bad measurements is applied (Table 5). The cost of removing 'suspicious' measurements that may be corrupted due to blending is that one may lose some variable stars that are blended, but have sufficiently high variability amplitude to be detected. The efficiency of variable star search with variability indices is determined by the ability to identify and discard bad measurements at

[12] Scintillation noise may also contribute significantly to random errors in ground-based photometric observations if conducted with short exposures and small telescopes (e.g. Kornilov et al. 2012).

the source extraction stage or assign appropriately high error bars to such measurements (and then use a variability index that takes error bars into account, see Table 1).

By computing the indices, one may pre-select candidate variables from a photometric data set reducing the initial number of considered objects typically by an order of magnitude. An index-based selection of candidates should be followed by a more sophisticated analysis such as period search and visual inspection of light curves and images to distinguish true variables from badly measured objects.

5.6 How to select a cut-off value?

The cut-off value, a, for variable objects selection, which provides a balance between the selection completeness and false positive rate (maximizing the F_1-score; Section 4), varies greatly between indices and data sets (Fig. 4). To select a for a new variability survey, one may use known variable stars covered by that survey. One would often tolerate a large number of false candidates in favour of a more complete variable objects selection, so a threshold set by maximizing the F_1-score (while being useful for comparing variability indices with each other) may be considered too high in practice. Instead, it is possible to search for the value of a maximizing

$$F_\beta = (1 + \beta^2)(C \times P)/(C + \beta^2 P), \quad (25)$$

where the parameter $\beta > 0$ determines how much importance we attach to completeness, C, relative of purity, P. For the test data sets described in Section 3, values of β as high as 50 are needed to have most of the known variables selected (with the majority of indices) above the cut-off limit that maximizes F_β.

For any variability index, the distributions of index values for variable and non-variable objects inherently intersect since (i) there is no lower limit on the possible amplitude of variability and (ii) there are often some objects with corrupted measurements resulting in elevated variability index values for them. The value of a should be chosen based on the false candidate rate that can be practically handled at the post-processing stage. For example, only a small number of false candidates are acceptable if selection based on variability indices is immediately followed by a visual inspection. A larger number of false candidates can be accepted if variability index-based selection is followed by a period search. If no list of known variables is available for the new survey data, one may start by setting, for example, $a = 3$ and gradually lowering the cut-off level until the number of false detections becomes unacceptable.

6 CONCLUSIONS

We compare 18 variability indices quantifying the overall scatter and/or degree of correlations between consecutive measurements in a light curve. The ability of these indices to distinguish variable stars from non-variable ones is tested on seven data sets collected with various ground-based telescopes and on simulated data incorporating actual light curves of non-variable objects as realistic models of photometric noise. We apply the PCA in search for an optimal combination of multiple variability indices.

We find that correlation-based indices are more efficient in selecting variable objects than the scatter-based indices for data sets containing hundreds of measurement epochs or more. The indices $1/\eta$, L, MAD and IQR perform better than others in selecting candidate variables from data sets affected by outliers. We suggest using the $1/\eta$ index together with the IQR as the pair of indices applicable to a wide variety of survey strategies and variability types. The indices $1/\eta$ and IQR provide stable high performance, albeit not always the highest one for each of the investigated data sets. However, the overall quality of a photometric data set including the percentage of outlier measurements and number of badly measured objects has a higher impact on the efficiency of variability search than the choice of a specific (set of) variability index(es).

Another efficient approach to variability detection is to compute many scatter- and correlation-based variability indices and perform the PCA over them. The admixture coefficient of the first principal component can be used as the composite index most suitable for the particular data set under investigation. This 'composite index' will perform on par with the best individual variability indices in this data set, but it requires no a priori knowledge of which indices are the best for the data set under investigation.

We also find that in practice, all the discussed variability indices as well as their combinations are not sufficient on their own to automatically select variable stars from a large set of light curves. The reason is that both variable and non-variable stars are diverse groups: variables may have various light-curve shapes, while non-variable stars include both the majority of objects displaying just noise and objects with photometry corrupted by nearby objects, cosmetic defects of a CCD, etc. The investigated indices cannot distinguish the badly measured objects from real variables because the corrupted measurements not only increase the light-curve scatter (compared to a non-variable object of similar brightness), but may also mimic correlated variability (due to night-to-night seeing variations, drift of the object's image across a cosmetic defect and so on). If all causes of measurement corruption in a particular data set can be identified and all such cases flagged at the source extraction stage, the discussed variability indices may efficiently distinguish variable objects standing out among the majority of non-variable stars.

At the same time, the variability indices are perfectly suitable to solve the inverse problem: identify well-measured constant stars in a photometric data set. The list of well-measured non-variable stars may be useful as photometric standards for calibration or targets for a search of variations not intrinsic to these objects such as microlensing events, occultations of stars by distant Solar system objects, etc.

The data sets used to test the variability indices were searched for variable objects previously. Despite that, we were able to identify 124 new variable stars during the tests. This highlights the fact that variability search techniques originally used to investigate the data sets can be improved by the application of the multiple variability indices tested here. The information about the new variables is summarized in Table 4, and their light curves are presented in Fig. 2. The variability types are assigned according to the GCVS system[13] (Samus et al. 2009) and high-amplitude δ Scuti/SX Phoenicis stars are indicated as HADS.

ACKNOWLEDGEMENTS

We thank Dr Laurent Eyer, Dr Ioannis Georgantopoulos and Dr Ilya Pashchenko for critically reading the manuscript, Nikolas Laskaris for the discussion of algorithms performance, Dr Valerio Nascimbeni and Dr Luigi Bedin for providing additional data for our technical tests and the entire Hubble Source Catalog (HSC; Whitmore et al. 2016) team for illuminating discussions and HSC data. KVS thanks Dr Boris Safonov and Dr Dmitry Chulkov for a discussion of scintillation-noise effects on photometry and Dmitry Litvinov for

[13] http://www.sai.msu.su/gcvs/gcvs/iii/vartype.txt

advice on how to improve the abstract. We acknowledge financial support by the European Space Agency (ESA) under the 'Hubble Catalog of Variables' programme, contract no. 4000112940. The work was supported by Act 211 Government of the Russian Federation, contract no. 02.A03.21.0006 and by the Ministry of Education and Science of the Russian Federation (the basic part of the state assignment, registration number 01201465056). KVS, SVA and NNS are supported by the Russian Foundation for Basic Research grant 13-02-00664. This publication makes use of data products from the AAVSO Photometric All Sky Survey (APASS). Funded by the Robert Martin Ayers Sciences Fund and the National Science Foundation. This research has made use of the International Variable Star Index (VSX) data base, operated at AAVSO, Cambridge, Massachusetts, USA. This research has made use of the VizieR catalogue access tool, CDS, Strasbourg, France. The original description of the VizieR service is presented by Ochsenbein, Bauer & Marcout (2000). This paper makes use of data from the DR1 of the WASP data (Butters et al. 2010) as provided by the WASP consortium, and the computing and storage facilities at the CERIT Scientific Cloud, reg. no. CZ.1.05/3.2.00/08.0144 which is operated by Masaryk University, Czech Republic. This research has made use of NASA's Astrophysics Data System.

REFERENCES

Akerlof C. et al., 2000, AJ, 119, 1901
Alard C., 2000, A&AS, 144, 363
Alard C., Lupton R. H., 1998, ApJ, 503, 325
Aller M. F., Aller H. D., Hughes P. A., 1992, ApJ, 399, 16
Allevato V., Paolillo M., Papadakis I., Pinto C., 2013, ApJ, 771, 9
Andrae R., Schulze-Hartung T., Melchior P., 2010, preprint (arXiv:1012.3754)
Angeloni R. et al., 2014, A&A, 567, A100
Arellano Ferro A., Bramich D. M., Figuera Jaimes R., Giridhar S., Kuppuswamy K., 2012, MNRAS, 420, 1333
Arellano Ferro A., Bramich D. M., Giridhar S., Figuera Jaimes R., Kains N., Kuppuswamy K., 2013, Acta Astron., 63, 429
Auvergne M. et al., 2009, A&A, 506, 411
Baade D. et al., 2016, A&A, 588, A56
Bacher A., Kimeswenger S., Teutsch P., 2005, MNRAS, 362, 542
Bailer-Jones C. A. L., Irwin M., von Hippel T., 1998, MNRAS, 298, 361
Bakos G. Á. et al., 2010, ApJ, 710, 1724
Barclay T., Ramsay G., Hakala P., Napiwotzki R., Nelemans G., Potter S., Todd I., 2011, MNRAS, 413, 2696
Becker A. C., Homrighausen D., Connolly A. J., Genovese C. R., Owen R., Bickerton S. J., Lupton R. H., 2012, MNRAS, 425, 1341
Berezhnoi A. A., 2013, Sol. Syst. Res., 47, 203
Bernard E. J. et al., 2010, ApJ, 712, 1259
Bertin E., 2011, in Evans I. N., Accomazzi A., Mink D. J., Rots A. H., eds, ASP Conf. Ser. Vol. 442, Astronomical Data Analysis Software and Systems XX. Astron. Soc. Pac., San Francisco, p. 435
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bonanos A. Z., 2007, AJ, 133, 2696
Bonanos A. Z., Stanek K. Z., Sasselov D. D., Mochejska B. J., Macri L. M., Kaluzny J., 2003, AJ, 126, 175
Bond H. E. et al., 2003, Nature, 422, 405
Borucki W. J. et al., 2010, Science, 327, 977
Bramich D. M., 2008, MNRAS, 386, L77
Bramich D. M., Horne K., Alsubai K. A., Bachelet E., Mislis D., Parley N., 2016, MNRAS, 457, 542
Brown L. M. J., Robson E. I., Gear W. K., Smith M. G., 1989, ApJ, 340, 150
Burdanov A. Y., Krushinsky V. V., Popov A. A., 2014, Astrophys. Bull., 69, 368
Burdanov A. Y. et al., 2016, MNRAS, 461, 3854
Burton J. R., Watson C. A., Littlefair S. P., Dhillon V. S., Gibson N. P., Marsh T. R., Pollacco D., 2012, ApJS, 201, 36
Butters O. W. et al., 2010, A&A, 520, L10
Campante T. L., 2012, PhD thesis, Universidade do Porto
Chakrabarti S., Saito R., Quillen A., Gran F., Klein C., Blitz L., 2015, ApJ, 802, L4
Ciaramella A. et al., 2004, A&A, 419, 485
Close L. M. et al., 1997, ApJ, 489, 210
Contreras Peña C. et al., 2014, MNRAS, 439, 1829
Cusano F. et al., 2013, ApJ, 779, 7
De Cuyper J., de Decker G., Laux U., Winter L., Zacharias N., 2012, in Ballester P., Egret D., Lorente N. P. F., eds, ASP Conf. Ser. Vol. 461, Astronomical Data Analysis Software and Systems XXI. Astron. Soc. Pac., San Francisco, p. 315
de Diego J. A., 2010, AJ, 139, 1269
Debosscher J., Sarro L. M., Aerts C., Cuypers J., Vandenbussche B., Garrido R., Solano E., 2007, A&A, 475, 1159
Denisenko D. V., Sokolovsky K. V., 2011, Astron. Lett., 37, 91
Desai S., Mohr J. J., Bertin E., Kümmel M., Wetzstein M., 2016, Astron. Comput., 16, 67
Devor J., 2005, ApJ, 628, 411
Dolphin A. E. et al., 2003, AJ, 125, 1261
Drake A. J. et al., 2009, ApJ, 696, 870
Drake A. J. et al., 2013, ApJ, 763, 32
Drake A. J. et al., 2014, ApJS, 213, 9
Emmanoulopoulos D., McHardy I. M., Papadakis I. E., 2013, MNRAS, 433, 907
Enoch M. L., Brown M. E., Burgasser A. J., 2003, AJ, 126, 1006
Enoch B., Haswell C. A., Norton A. J., Collier-Cameron A., West R. G., Smith A. M. S., Parley N. R., 2012, A&A, 548, A48
Eyer L., 2006, in Aerts C., Sterken C., eds, ASP Conf. Ser. Vol. 349, Astrophysics of Variable Stars. Astron. Soc. Pac., San Francisco, p. 15
Eyer L., Woźniak P. R., 2001, MNRAS, 327, 601
Fan J. H., Xu W., Pan J., Yuan Y. H., 2011, in Romero G. E., Sunyaev R. A., Belloni T., eds, Proc. IAU Symp. 275, Jets at All Scales. Cambridge Univ. Press, Cambridge, p. 164
Feeney S. M. et al., 2005, AJ, 130, 84
Fernández-Trincado J. G., Vivas A. K., Mateu C. E., Zinn R., Robin A. C., Valenzuela O., Moreno E., Pichardo B., 2015, A&A, 574, A15
Ferreira Lopes C. E., Cross N. J. G., 2016, A&A, 586, A36
Ferreira Lopes C. E., Dékány I., Catelan M., Cross N. J. G., Angeloni R., Leão I. C., De Medeiros J. R., 2015, A&A, 573, A100
Figuera Jaimes R., Arellano Ferro A., Bramich D. M., Giridhar S., Kuppuswamy K., 2013, A&A, 556, A20
Findeisen K., Cody A. M., Hillenbrand L., 2015, ApJ, 798, 89
Fruchter A. S., Hook R. N., 2002, PASP, 114, 144
Fruth T. et al., 2012, AJ, 143, 140
Gorshkov A. G., Konnikova V. K., Mingaliev M. G., 2012, Astron. Rep., 56, 345
Graczyk D., Eyer L., 2010, Acta Astron., 60, 109
Graczyk D. et al., 2011, Acta Astron., 61, 103
Graham M. J., Drake A. J., Djorgovski S. G., Mahabal A. A., Donalek C., Duan V., Maker A., 2013, MNRAS, 434, 3423
Graham M. J., Djorgovski S. G., Drake A. J., Mahabal A. A., Chang M., Stern D., Donalek C., Glikman E., 2014, MNRAS, 439, 703
Gran F., Minniti D., Saito R. K., Navarrete C., Dékány I., McDonald I., Contreras Ramos R., Catelan M., 2015, A&A, 575, A114
Guterman P., Mazeh T., Faigler S., 2015, in Martins F., Boissier S., Buat V., Cambrésy L., Petit P., eds, SF2A-2015: Proceedings of the Annual Meeting of the French Society of Astronomy and Astrophysics. p. 277
Hartman J. D., Bakos G. Á., Noyes R. W., Sipőcz B., Kovács G., Mazeh T., Shporer A., Pál A., 2011, AJ, 141, 166
Henden A. A., Templeton M., Terrell D., Smith T. C., Levine S., Welch D., 2016, VizieR Online Data Catalog, 2336
Henze M., Meusinger H., Pietsch W., 2008, A&A, 477, 67
Hernández-García L., Masegosa J., González-Martín O., Márquez I., 2015, A&A, 579, A90
Hippke M., Angerhausen D., 2015, ApJ, 810, 29

Hoffmann S. L., Macri L. M., 2015, AJ, 149, 183
Hoffmann S. L. et al., 2016, preprint (arXiv:e-prints)
Hovatta T., Nieppola E., Tornikoski M., Valtaoja E., Aller M. F., Aller H. D., 2008, A&A, 485, 51
Javadi A., Saberi M., van Loon J. T., Khosroshahi H., Golabatooni N., Mirtorabi M. T., 2015, MNRAS, 447, 3973
Jenkins J. M., Doyle L. R., Cullers D. K., 1996, Icarus, 119, 244
Jester S. et al., 2005, AJ, 130, 873
Kaluzny J., Thompson I. B., Rozyczka M., Pych W., Narloch W., 2014, Acta Astron., 64, 309
Karampelas A., Kontizas M., Rocca-Volmerange B., Bellas-Velidis I., Kontizas E., Livanou E., Tsalmantza P., Dapergolas A., 2012, A&A, 538, A38
Kenney J. F., Keeping E. S., 1956, Mathematics of Statistics. Van Nostrand, Princeton, NJ
Kessler R. et al., 2015, AJ, 150, 172
Kim D.-W., Bailer-Jones C. A. L., 2016, A&A, 587, A18
Kim D.-W., Protopapas P., Alcock C., Byun Y.-I., Khardon R., 2011a, in Evans I. N., Accomazzi A., Mink D. J., Rots A. H., eds, ASP Conf. Ser. Vol. 442, Astronomical Data Analysis Software and Systems XX. Astron. Soc. Pac., San Francisco, p. 447
Kim D.-W., Protopapas P., Byun Y.-I., Alcock C., Khardon R., Trichas M., 2011b, ApJ, 735, 68
Kim D.-W., Protopapas P., Bailer-Jones C. A. L., Byun Y.-I., Chang S.-W., Marquette J.-B., Shin M.-S., 2014, A&A, 566, A43
Kinemuchi K., Smith H. A., Woźniak P. R., McKay T. A., ROTSE Collaboration, 2006, AJ, 132, 1202
Klagyivik P. et al., 2016, AJ, 151, 110
Kochanek C. S., Beacom J. F., Kistler M. D., Prieto J. L., Stanek K. Z., Thompson T. A., Yüksel H., 2008, ApJ, 684, 1336
Kolesnikova D. M., Sat L. A., Sokolovsky K. V., Antipin S. V., Samus N. N., 2008, Acta Astron., 58, 279
Kolesnikova D. M., Sat L. A., Sokolovsky K. V., Antipin S. V., Belinskii A. A., Samus' N. N., 2010, Astron. Rep., 54, 1000
Kornilov V., Sarazin M., Tokovinin A., Travouillon T., Voziakova O., 2012, A&A, 546, A41
Kourniotis M. et al., 2014, A&A, 562, A125
Kovács G., Bakos G., Noyes R. W., 2005, MNRAS, 356, 557
Lang D., Hogg D. W., Mierle K., Blanton M., Roweis S., 2010, AJ, 139, 1782
Lapukhin E. G., Veselkov S. A., Zubareva A. M., 2013, Perem. Zvezdy Prilozh., 13, 12
Lapukhin E. G., Veselkov S. A., Zubareva A. M., 2016, Perem. Zvezdy Prilozh., 16, 4
Law N. M. et al., 2009, PASP, 121, 1395
Lawrence A., Papadakis I., 1993, ApJ, 414, L85
Laycock S., Tang S., Grindlay J., Los E., Simcoe R., Mink D., 2010, AJ, 140, 1062
Lemeshko S. B., 2006, Meas. Tech., 49, 962
López-Morales M., Coughlin J. L., Sing D. K., Burrows A., Apai D., Rogers J. C., Spiegel D. S., Adams E. R., 2010, ApJ, 716, L36
McCormac J., Skillen I., Pollacco D., Faedi F., Ramsay G., Dhillon V. S., Todd I., Gonzalez A., 2014, MNRAS, 438, 3383
Macfarlane S. A., Toma R., Ramsay G., Groot P. J., Woudt P. A., Drew J. E., Barentsen G., Eislöffel J., 2015, MNRAS, 454, 507
Macri L. M. et al., 1999, ApJ, 521, 155
Majorova E. K., Zhelenkova O. P., 2012, Astrophys. Bull., 67, 318
Max-Moerbeck W., Richards J. L., Hovatta T., Pavlidou V., Pearson T. J., Readhead A. C. S., 2014, MNRAS, 445, 437
Melchior P. et al., 2016, Astron. Comput., 16, 99
Mingaliev M. G., Sotnikova Y. V., Udovitskiy R. Y., Mufakharov T. V., Nieppola E., Erkenov A. K., 2014, A&A, 572, A59
Mislis D., Bachelet E., Alsubai K. A., Bramich D. M., Parley N., 2016, MNRAS, 455, 626
Moretti M. I. et al., 2016, MNRAS, 459, 1687
Mowlavi N., 2014, A&A, 568, A78
Munari U., Henden A., Frigo A., Dallaporta S., 2014, J. Astron. Data, 20, 4
Nandra K., George I. M., Mushotzky R. F., Turner T. J., Yaqoob T., 1997, ApJ, 476, 70
Nardiello D. et al., 2015, MNRAS, 447, 3536
Nardiello D., Libralato M., Bedin L. R., Piotto G., Ochner P., Cunial A., Borsato L., Granata V., 2016, MNRAS, 455, 2337
Nascimbeni V. et al., 2014, MNRAS, 442, 2381
Nikołajuk M., Czerny B., Gurynowicz P., 2009, MNRAS, 394, 2141
Nun I., Protopapas P., Sim B., Zhu M., Dave R., Castro N., Pichara K., 2015, preprint (arXiv:e-prints)
Ochsenbein F., Bauer P., Marcout J., 2000, A&AS, 143, 23
Ofek E. O. et al., 2012, PASP, 124, 854
Ordoñez A. J., Sarajedini A., 2016, MNRAS, 455, 2163
Paegert M., Stassun K. G., Burger D. M., 2014, AJ, 148, 31
Pál A., 2009, PhD thesis, Eötvös Loránd University
Palaversa L. et al., 2013, AJ, 146, 101
Parks J. R., Plavchan P., White R. J., Gee A. H., 2014, ApJS, 211, 3
Pawlak M. et al., 2013, Acta Astron., 63, 323
Pearson K., 1901, Phil. Mag., 2, 559
Perlman E. S. et al., 2011, ApJ, 743, 119
Pietrukowicz P. et al., 2013, Acta Astron., 63, 115
Plavchan P., Jura M., Kirkpatrick J. D., Cutri R. M., Gallagher S. C., 2008, ApJS, 175, 191
Pojmanski G., Pilecki B., Szczygiel D., 2005, Acta Astron., 55, 275
Ponti G., Papadakis I., Bianchi S., Guainazzi M., Matt G., Uttley P., Bonilla N. F., 2012, A&A, 542, A83
Popov A. A., Burdanov A. Y., Zubareva A. M., Krushinsky V. V., Avvakumova E. A., Ivanov K., 2015, Perem. Zvezdy Prilozh., 15, 7
Prša A. et al., 2011, AJ, 141, 83
Ramsay G. et al., 2014, MNRAS, 437, 132
Rauer H. et al., 2014, Exp. Astron., 38, 249
Re Fiorentin P., Bailer-Jones C. A. L., Lee Y. S., Beers T. C., Sivarani T., Wilhelm R., Allende Prieto C., Norris J. E., 2007, A&A, 467, 1373
Rebull L. M. et al., 2015, AJ, 150, 175
Rest A. et al., 2014, ApJ, 795, 44
Richards J. W. et al., 2011, ApJ, 733, 10
Ricker G. R. et al., 2014, Proc. SPIE, 9143, 20
Robert V., Lainey V., Pascu D., Arlot J.-E., De Cuyper J.-P., Dehant V., Thuillot W., 2014, A&A, 572, A104
Rose M. B., Hintz E. G., 2007, AJ, 134, 2067
Rousseeuw P. J., Croux C., 1993, J. Am. Stat. Assoc., 88, 1273
Sahay A., Lebzelter T., Wood P. R., 2014, PASA, 31, 12
Samus N. N., Durlevich O. V., Goranskij V. P., Kazarovets E. V., Kireeva N. N., Pastukhova E. N., Zharova A. V., 2009, VizieR Online Catalog B/gcvs
Sesar B. et al., 2013, AJ, 146, 21
Shappee B. J., Stanek K. Z., 2011, ApJ, 733, 124
Shin M.-S., Sekora M., Byun Y.-I., 2009, MNRAS, 400, 1897
Simcoe R. J., 2009, in Osborn W., Robbins L., eds, ASP Conf. Ser. Vol. 410, Preserving Astronomy's Photographic Legacy: Current State and the Future of North American Astronomical Plates. Astron. Soc. Pac., San Francisco, p. 111
Simcoe R. J., Grindlay J. E., Los E. J., Doane A., Laycock S. G., Mink D. J., Champine G., Sliski A., 2006, Proc. SPIE, 6312, 631217
Simm T. et al., 2015, A&A, 584, A106
Sitek M., Pojmański G., 2014, Acta Astron., 64, 115
Sokolovsky K., Lebedev A., 2005, in Simon A., Golovin A., eds, 12th Young Scientists' Conference on Astronomy and Space Physics. Kyiv, Ukraine, p. 79
Sokolovsky K. V., Kovalev Y. Y., Kovalev Y. A., Nizhelskiy N. A., Zhekanis G. V., 2009, Astron. Nachr., 330, 199
Sokolovsky K., Korotkiy S., Lebedev A., 2014a, in Woudt P. A., Ribeiro V. A. R. M., eds, ASP Conf. Ser. Vol. 490, Stell Novae: Past and Future Decades. Astron. Soc. Pac., San Francisco, p. 395
Sokolovsky K., Antipin S., Kolesnikova D., Lebedev A., Samus N., Sat L., Zubareva A., 2014b, in Mišková L., Vítek S., eds, Astroplate 2014, p. 79
Sokolovsky K. V., Antipin S. V., Zubareva A. M., Kolesnikova D. M., Lebedev A. A., Samus' N. N., Sat L. A., 2014c, Astron. Rep., 58, 319

Sokolovsky K. V., Kolesnikova D. M., Zubareva A. M., Samus N. N., Antipin S. V., 2016, preprint (arXiv:1605.03571)
Song F.-F., Esamdin A., Ma L., Liu J.-Z., Zhang Y., Niu H.-B., Yang T.-Z., 2016, preprint (arXiv:1606.04792)
Soszyński I. et al., 2015, Acta Astron., 65, 297
Steiner J. E., Menezes R. B., Ricci T. V., Oliveira A. S., 2009, MNRAS, 395, 64
Stetson P. B., 1996, PASP, 108, 851
Strunov V. I., 2006, Meas. Tech., 49, 755
Süveges M. et al., 2012, MNRAS, 424, 2528
Szymanski M. K., 2005, Acta Astron., 55, 43
Tamuz O., Mazeh T., Zucker S., 2005, MNRAS, 356, 1466
Tamuz O., Mazeh T., North P., 2006, MNRAS, 367, 1521
Tang S., Grindlay J., Los E., Servillat M., 2013, PASP, 125, 857
Timmer J., Koenig M., 1995, A&A, 300, 707
Tisserand P., Clayton G. C., Welch D. L., Pilecki B., Wyrzykowski L., Kilkenny D., 2013, A&A, 551, A77
Tody D., 1986, in Crawford D. L., ed., Proc. SPIE, Vol. 627, Instrumentation in Astronomy VI. SPIE, Bellingham, p. 733
Tomaney A. B., Crotts A. P. S., 1996, AJ, 112, 2872
Tso B., Mather P. M., 2001, Classification Methods for Remotely Sensed Data. Taylor and Francis, London
Tuvikene T., Edelmann H., Groote D., Enke H., 2014, in Mišková L., Vítek S., eds, Astroplate 2014, p. 127
Udalski A., Szymanski M., Kaluzny J., Kubiak M., Mateo M., Krzeminski W., Paczynski B., 1994, Acta Astron., 44, 227
Udalski A., Kubiak M., Szymanski M., 1997, Acta Astron., 47, 319
Udalski A., Szymański M. K., Szymański G., 2015, Acta Astron., 65, 1
van Velzen S. et al., 2011, ApJ, 741, 73
Vaughan S., Edelson R., Warwick R. S., Uttley P., 2003, MNRAS, 345, 1271
Vicente B., Abad C., Garzón F., 2007, A&A, 471, 1077
Vicente B., Abad C., Garzón F., Girard T. M., 2010, A&A, 509, A62
Villforth C., Koekemoer A. M., Grogin N. A., 2010, ApJ, 723, 737
Vivas A. K. et al., 2016, AJ, 151, 118
von Neumann J., 1941, Ann. Math. Stat., 12, 367
von Neumann J., 1942, Ann. Math. Stat., 13, 86
Walker G. et al., 2003, PASP, 115, 1023
Wall J. V., Jenkins C. R., 2003, Practical Statistics for Astronomers. Cambridge Univ. Press, Cambridge
Watson C. L., 2006, Society for Astronomical Sciences 25th Annual Symposium. Society for Astronomical Sciences, Big Bear, CA, p. 47
Weingrill J., 2015, Astron. Nachr., 336, 125
Welch D. L., Stetson P. B., 1993, AJ, 105, 1813
Whitmore B. C. et al., 2016, AJ, 151, 134
Wozniak P. R., 2000, Acta Astron., 50, 421
Woźniak P. R. et al., 2004, AJ, 127, 2436
Yao S., Yuan W., Komossa S., Grupe D., Fuhrmann L., Liu B., 2015a, AJ, 150, 23
Yao X. et al., 2015b, AJ, 150, 107
Yip C. W. et al., 2004, AJ, 128, 585
Zacharias N., Finch C. T., Girard T. M., Henden A., Bartlett J. L., Monet D. G., Zacharias M. I., 2013, AJ, 145, 44
Zackay B., Ofek E. O., 2015, ApJ, preprint (arXiv:1512.06872)
Zebrun K. et al., 2001, Acta Astron., 51, 317
Zhang M., Bakos G. Á., Penev K., Csubry Z., Hartman J. D., Bhatti W., de Val-Borro M., 2016, PASP, 128, 035001
Zhang T.-M. et al., 2015, Res. Astron. Astrophys., 15, 215
Zhang X.-B., Deng L.-C., Xin Y., Zhou X., 2003, Chin. J. Astron. Astrophys., 3, 151
Zhang Z.-W. et al., 2013, AJ, 146, 14
Zheleznyak A. P., Kravtsov V. V., 2003, Astron. Lett., 29, 599
Zinn R., Horowitz B., Vivas A. K., Baltay C., Ellman N., Hadjiyska E., Rabinowitz D., Miller L., 2014, ApJ, 781, 22

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table 4. New variable stars found in the test data.
Table 6. Performance of variability indices on the data sets with simulated periodic variability.
Table 7. Performance of variability indices on the data sets with simulated non-periodic variability.
Figure 2. Light curves of new variable stars found in the test data sets. (http://www.mnras.oxfordjournals.org/lookup/suppl/doi:10.1093/mnras/stw2262/-/DC1).

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.