
Optimizing differential photometry for stellar variability monitoring

Author:

Guifré Sánchez Serra

Director:

Dr. Enrique Herrero Casas

Tutor:

Dr. Emili Bagan Capella

Declaració d'autoria

Jo, Guifré Sánchez Serra, amb Document Nacional d'Identitat 25368879V, i estudiant del Grau en Física de la Universitat Autònoma de Barcelona, en relació amb la memòria del treball de final de Grau presentada per a la seva defensa i avaluació durant la convocatòria de setembre del curs 2019-2020, declaro que

- El document presentat és original i ha estat realitzat per la meva persona.
- El treball s'ha dut a terme principalment amb l'objectiu d'avaluar l'assignatura de treball de grau en física a la UAB, i no s'ha presentat prèviament per ser qualificat en l'avaluació de cap altra assignatura ni en aquesta ni en cap altra universitat.
- En el cas de continguts de treballs publicats per terceres persones, l'autoria està clarament atribuïda, citant les fonts degudament.
- En el casos en els que el meu treball s'ha realitzat en col·laboració amb altres investigadors i/o estudiants, es declara amb exactitud quines contribucions es deriven del treball de tercers i quines es deriven de la meva contribució.
- A l'excepció dels punts esmentats anteriorment, el treball presentat és de la meva autoria.

Signat:

Guifré Sánchez Serra

Arenys de Munt, 21 d'agost de 2020

Declaració d'extensió

Jo, Guifré Sánchez Serra, amb Document Nacional d'Identitat 25368879V, i estudiant del Grau en Física de la Universitat Autònoma de Barcelona, en relació amb la memòria del treball de final de Grau presentada per a la seva defensa i avaluació durant la convocatòria de setembre del curs 2019-2020, declaro que:

- El nombre total de paraules (segons comptatge proposat) incloses en les seccions des de la introducció a les conclusions és de 8986 paraules
- El nombre total de figures és de 5.

En total el document, comptabilitza:

$$8986 \text{ paraules} + 5 \times 200 \text{ paraules/figura} = 9986$$

Que compleix amb la normativa de ser inferior a 10000.

Signat:

Guifré Sánchez Serra

Arenys de Munt, 21 d'agost de 2020

Agraïments

En primer lloc, voldria agrair-li al Kike l'entusiasme, la paciència, la proximitat (malgrat les circumstàncies) i tots els “Endavant!” amb què sovint acabaven els correus i les trobades que han fet possible aquest treball de final de grau. També m'agradaria agrair-li haver-me permès tancar d'una forma tan bonica un cicle que va començar ara farà uns 7 anys, amb un programa que porto a la *motxilla* des d'aleshores. Gràcies també per haver-me redescobert l'estima per l'astronomia, ha estat un autèntic plaer.

Aquest treball és el darrer que faig com a estudiant del doble grau de física i matemàtiques a la UAB, i en cert sentit representa el final d'una etapa. Sobretot per aquests últims anys, i molt especialment pel suport que m'han donat al llarg dels darrers mesos, voldria agrair i dedicar als meus pares la memòria que segueix. Gràcies per ser-hi sempre.

Abstract

We present and implement different optimization and data filtering techniques to improve differential photometry light curves from 4 M dwarfs (Wolf 1069 [24], TOI-1266 [33], TZ Ari [37] and GJ 555 [37]) that show signs of low-amplitude periodicities. Images were obtained using the 80 cm Joan Oró telescope at Montsec Astronomical Observatory and extraction of differential photometry data was performed using *AstroImageJ*, [5]. Our main motivation is that stellar activity induced brightness variations (driven by starspots, stellar rotation, flares, etc.) represent a source of “noise” in both photometric and RV (radial velocity) measurements intended for exoplanet detection and characterization, e.g.: [3, 19]; precise modelling of such photometric modulations could then improve planet parameter estimates.

σ -clipping and SNR numerical clipping are adopted as the main light curve filtering procedures to reduce photon noise and remove outlier points. Moreover, two different techniques are proposed for an optimized (re)selection of comparison stars —and subsequent (potential) improvement of target signal—, based on target RSD (relative standard deviation) minimization and variability index performance (where 4 indices are considered, **IQR**, χ^2 , σ , η^{-1} , following indications in [32]).

Relative fluxes from each M dwarf are studied, applying the different techniques described above. Generalized Lomb-Scargle periodograms, [39], are used to seek for significant periodicities in the observed signals; in the cases of TZ Ari and GJ 555 a prewhitening scheme, [2, 7], is also considered. Our results show that the proposed procedures can improve periodogram peak powers up to $\sim 20\%$; however, further analysis with a larger number of light curves is advised, to allow for a systematic adjustment of optimization parameters. The obtained periods range from ~ 2 days (TZ Ari) to ~ 200 days (Wolf 1069).

Contents

1	Differential photometry. An overview.	1
2	Stellar variability detection	3
2.1	Light curve filtering	4
2.2	Variability indices	5
2.3	Period search	5
2.3.1	Practical aspects of the Lomb-Scargle periodogram	6
3	Light curve reduction methods	8
3.1	Comparison star selection criteria	8
3.1.1	Target RSD minimization	8
3.1.2	Variability index performance	9
3.2	Filtering procedures	9
4	Results and discussion	10
4.1	Wolf 1069	11
4.2	TOI-1266	13
4.3	TZ Ari	14
4.4	GJ 555	18
5	Conclusions	20
A	Complementary figures and tables	21
B	Sources of noise in photometric measurements	28
C	Photometry extraction with AstroImageJ	30
D	Basic principles of periodogram analysis	31
D.1	Fourier analysis and power spectrum	31
D.2	The Lomb-Scargle periodogram	32
D.3	Periodogram failure modes	34
E	Filtering procedures. Binning and implementation details.	35
E.1	Binning	35
E.2	Implementation details	35
F	Description of variability indices	37
F.1	Interquartile range (IQR)	37
F.2	χ^2 test	37
F.3	Standard deviation, σ	37
F.4	The von Neumann ratio, η	38

1 Differential photometry. An overview.

Photometry is a widely used technique in astronomy that allows for highly precise measurements of flux or intensity of light radiated by astronomical objects. Combined with telescopes, CCD¹ cameras have become in the past years the main device used to perform such measurements, both for their practicality and precision. Essentially, these are electronic devices consisting of a 2-dimensional grid of cells, that convert light into electric current by the photoelectric effect. Thus, each cell records, roughly speaking, the number of photons passing through it. The resulting set of data points is used to extract the photometry of our field of view, and infer valuable physical information of the observed objects. Photometry extraction can be performed on the recorded data in different ways, yielding mainly 3 types of photometry: relative, absolute and differential.

The basic principle behind the extraction of photometric information from raw data obtained using a CCD can be summarized as follows. In first place, we need to take into account the response of the camera (together with the telescope and possibly other optical devices) to a point source, which is described by the so called *point spread function* (PSF). This will make the observed signal cover many pixels (cells) of the CCD. The flux corresponding to the observed object will then be proportional to the sum of pixel counts (number of photons recorded per pixel) within a circular aperture centered on the object, minus the average sky flux within that aperture (which can be computed, for example, as the product of the number of pixels within the aperture and a nearby average sky count from the immediate surroundings of the object). This is known as *aperture photometry*. The true flux can then be recovered through calibration procedures.

The advance of technology, and particularly of electronics, has made possible achieving high-precision flux measurements using CCD photometry, which has proven to be a key technique in the discovery and study of a variety of astronomical objects. Thus, among the many applications of high-precision, wide-field, ground-based photometry, we may highlight the following ones (see [8]): searches for transiting extrasolar planets, low-level stellar variability, the study of microlensing events, and searches for nearby, low-mass binaries. Although much effort has been devoted since the birth of CCD photometry to improve precision of measurements, and many authors have contributed to it (see [8, 14, 18, 22]), there is still a lot of work to do, in order to attain the desired resolution for the systematic study of stellar light curves. One of the purposes of this work is to develop tools that help improve photometry data after its extraction from raw image fluxes obtained using finely tuned modern CCDs. The starting point for this will be what is known as *differential photometry*.

Flux variations in astronomical sources are due to both intrinsic and extrinsic causes. Clearly, only intrinsic variations provide us with specific information about the object of interest; consequently, we will need, in general, to get rid of external factors contaminating the light curve of the source. Differential photometry aims to solve this problem and is based in the idea that measuring objects near a certain target and using differencing techniques we can eliminate essentially all systematic causes of variation, which will leave the *true* signal of the target. A technique that has proven to dramatically improve the precision of raw flux measurements in the context of differential photometry (see [4, 5, 8]) is selecting several² *comparison stars* from the (relatively) immediate surroundings of the *target*, and computing the *relative* flux of the latter, ϕ_T , and each of the comparison stars, ϕ_i as follows³:

$$\phi_T = \frac{\Phi_T}{\sum_i \Phi_i}, \quad \phi_j = \frac{\Phi_j}{\sum_{i \neq j} \Phi_i} \quad (1)$$

where the index i in the sums goes over all comparison stars and Φ_T , Φ_i represent the raw fluxes of the target and the i -th comparison star.

¹Charge-coupled device.

²Typical numbers are in the range 15–30 comparison stars.

³Other differential photometry formulas may be used; however, expressions in equation (1) are used explicitly in **AstroImageJ**, [5], which is the program we employ to perform extraction of differential photometry from raw images.

Ground-based observations that use comparison stars in close proximity to the target have shown to reduce notably (compare [23, p. 33-35]): atmospheric variations, due to small angular distance between target and comparison stars (first-order atmospheric extinction); color-dependent effects, due to similar intrinsic color index (second-order extinction); or variations in background flux and sky brightness, among others. Understanding the different sources of noise that present in high-precision photometry measurements is therefore important, if we aim to increase their overall signal-to-noise ratio. A summary of the main sources of error in such measurements can be found in appendix B.

2 Stellar variability detection

Light curve variability of astronomical objects is one of the main sources of information we have for their own study and identification. Variable stars have proven to be particularly suited for the discovery and characterization of several phenomena, with the transiting of extrasolar planets being possibly one of the most celebrated contributions. The analysis of large ensembles of variable stars has also been key in the study of stellar populations and to the understanding of the structure of the Milky Way and the expansion of the Universe (see [30]). However, variable behaviour is not exclusively stellar, as it may be encountered in a variety of contexts, from variable reflecting nebulae, to active galactic nuclei. On the other hand, the amplitudes and time-scales of such changes can also vary greatly, from a fraction of a magnitude and minutes for stellar pulsations to tens of magnitudes and weeks in the case of supernovae explosions. These considerations make the task of classifying variable astronomical objects fairly difficult.

A typical approach would consist of the following 4 stages: detection, analysis in the time domain, analysis in the phase domain with period estimation, and classification of the object. In this work we are interested in the first 3 stages. Variability detection is notably conditioned by the nature of the data under study. Thus, detection algorithms are inevitably linked to the type of data we are given, which may vary in form, structure and content, depending on the variability type of the object of interest, the observation cadence, the systematic errors of the measuring device, etc. It is then understandable that a general and reliable procedure for detection of variable behaviour has not yet been conceived.

In the case of low-amplitude variability, the systematic errors of photometric measurements can be in some cases mistaken by real flux variations, and hide completely the true behaviour of the object under study. This is why minimizing such sources of noise is important; moreover, low-amplitude variable stars have shown to be key in the characterization of the structure and dynamics of stellar surface (e.g.: distribution of spots and microflares, coronal heating, etc.) and other aspects of stellar physics (see [25]). On the other hand, low-amplitude photometric modulations induced by stellar activity can also be viewed as a source of “noise” with regards to exoplanet detection through radial velocity⁴ (RV) measurements. Recent studies (see [3]) have shown that spectral analysis of stellar activity driven RV variations can be used to simultaneously fit activity and planetary signals to RV data from known host stars; this may help improving planetary and orbital parameter estimates.

Another approach to mitigate the effects of stellar activity in exoplanet detection is to perform a combined analysis using both RV measurements and time-series high-precision photometry. It has been recently pointed out (see [19]) that this methodology could increase precision of radial velocity surveys, which are currently unable to completely characterize true Earth analogues, due to significant host activity, as well as certain precision and monitoring requirements; the aforementioned investigation also suggests that stellar activity is a major noise component in RV measurements from Earth-like planetary systems, and may limit future exoplanet surveys unless it can be reduced.

Star rotation is essential to stellar activity (compare [31]) and, in particular, to the generation of starspots. These are produced in the photosphere, due to the coupling of convective motions with stellar rotation; a strong magnetic field originates from this process (the so called *dynamo effect*), inhibiting convection and effectively cooling the affected region. Large groups of small spots can gather in big spotted areas, whose main structure is maintained through long time periods, causing detectable coherent brightness variations (see [13]). These variations can introduce changes in stellar spectral line profiles, and are therefore an important part of the noise component attributed to stellar activity signals in radial velocity data. Hence, monitoring stellar rotation with high-precision photometry could be essential to mitigate the effects of stellar activity in RV variations driven by planetary motion. One of the objectives of the present work is to detect low-amplitude periodicities associated to stellar activity in differential photometry data from a number of target stars, through different light curve optimization procedures.

⁴Exoplanet detection can be performed through the study of radial velocity variations associated to the candidate host star (these would be attributed to the changing gravitational pull from the planet); they are usually measured using *Doppler spectroscopy*, which is based on the fact that the spectrum of the star will be blueshifted or redshifted if it moves towards us or away from us, respectively (regular inspection of the spectrum allows for detection of periodic signals).

Returning to the problem of classifying variable sources, there are primarily 3 different approaches regarding variability detection (see [32]): direct image comparison (*transient detection*), computation of variability indices and periodic search. Transient detection consists of seeking changes between two or more sets of sky images taken at different epochs. Variability indices result from the analysis of light curves and are aimed at describing the overall scatter of flux measurements and/or the degree of correlation between them (more precisely, between consecutive flux measurements). Finally, period search is applied when there are sufficient signs of flux variability and a form of periodicity is assumed a priori for the object of interest; this allows for time-series and/or spectral analysis, leading to certification of significant periodicity in the signal, and possible fitting of a best-model curve to the experimental data. In this work we will explore variability detection through variability indices and period search.

In general, when studying variability for large ensembles of candidate stars, of a certain variability type, we begin by preselecting some of them using variability indices, and then we proceed applying period search only to lightcurves that passed the previous stage. This is because period search and curve-fitting algorithms are computationally expensive, while variability indices are typically fast to compute.

2.1 Light curve filtering

As stated before, minimizing errors in flux measurements is important to distinguish between lightcurve scattering produced by the actual astronomical object and scattering associated to the different sources of noise contaminating such measurements. Improvement of high-precision photometric data can be accomplished after the measuring process by adequate filtering of the fluxes that make up the light curve. Among the many procedures aimed at enhancing data post-recording, the most widely used are: σ -clipping, numerical clipping and binning. Here we describe the first two, as they have been the most utilized throughout this work (see appendix E.1 for a brief description of flux binning).

σ -clipping. Given a set of measurements $\{y_i\}_{i=1}^n$, with y_i errors modelled by a 0-mean Gaussian distribution, and subject to possible outliers, σ -clipping provides us with a method to avoid/reduce the effects of such outliers in the given data. It is defined as the following iterative process:

- i. Compute/estimate the standard deviation, σ , and the mean (or median), m , of the data under study.
- ii. Remove all points, y_j , outside the range $m \pm \alpha\sigma$, with α a positive real number.
- iii. Go to step i., unless the selected exit criteria is reached.

The exit criteria in step iii. can be either stopping after a fixed number of iterations or forcing the standard deviation of the new data to be within a certain tolerance level of the old one. A possible definition for such tolerance level could be $(\sigma_{\text{old}} - \sigma_{\text{new}})/\sigma_{\text{new}}$, which represents the relative change in standard deviation between successive iterations (note that in each clipping, we are reducing the dispersion of the data, thus $\sigma_{\text{old}} \geq \sigma_{\text{new}}$).

In the case of (relative) flux measurements, as mentioned in appendix B, the associated noise follows a Poisson distribution (photon counting noise). Since $\text{Pois}(\mathcal{N})$ for large \mathcal{N} is very well approximated by a Gaussian distribution with mean $\mu = \mathcal{N}$ and standard deviation $\sigma = \sqrt{\mathcal{N}}$, it is reasonable to apply σ -clipping to photometric data.

For flux measurements m_i with associated errors σ_i , $i = 1, \dots, n$, one way to estimate the standard deviation is by using the classical (unbiased) estimator $\sqrt{\sum_{i=1}^n (m_i - \bar{m})^2 / (n-1)}$ with \bar{m} the arithmetic mean of the m_i s. However, if the estimated errors, σ_i , are assumed to reflect the relative accuracy of the measurements, one can use its weighted version (see [32]):

$$\sigma_w = \left[\frac{\sum_{i=1}^n w_i}{(\sum_{i=1}^n w_i)^2 - \sum_{i=1}^n w_i^2} \sum_{i=1}^n w_i (m_i - \bar{m}_w)^2 \right]^{1/2}, \text{ with } \bar{m}_w = \frac{\sum_{i=1}^n w_i m_i}{\sum_{i=1}^n w_i} \text{ the weighted mean} \quad (2)$$

with weights $w_i \equiv \sigma_i^{-2}$.

Numerical clipping. Values outside a specified range are removed or clipped. This range can be dependent on the data set that has been given. In the case of photometric measurements, numerical clipping is usually applied to select the flux measurements that have the best signal-to-noise ratios. Thus, given the different SNR values associated to a specific object in a certain sky image, $\{\text{SNR}_i\}_{i=1}^n$, a value $p \in (0, 1)$, and setting $M \equiv \max\{\text{SNR}_i\}_{i=1}^n$, only values larger than Mp are conserved, *i.e.* fluxes with SNRs lower than Mp are clipped. Appropriate selection of p after careful inspection of SNR values can lead to a considerable improvement of the data, since low SNR fluxes could introduce significant errors to the true signal of the object under study.

2.2 Variability indices

Any number quantifying “how variable” an astronomical object is can be referred to as a *variability index*. An important number of such indices has been presented, discussed and tested for variability detection in recent studies (e.g.: [30, 32, 34]). In [32], a total of 18 different variability indices are compared, testing their performance in 7 time series data sets of different nature, which include lightcurves of about 130 thousand objects (some simulated data is also tested). They study only scatter- and correlation-based indices, and conclude that the latter type performs better than the former for data sets containing at least hundreds of measurement epochs. They also suggest two particular indices ($1/\eta$ and IQR) as the best performing pair applicable to a wide variety of variability types. Finally, they highlight that the investigated variability indices are perfectly suitable to solve the inverse problem, that is: identifying well-measured constant stars.

In this work, 4 different variability indices are used to optimize the initial ensemble of comparison stars that are selected to perform differential photometry on target fluxes. A summary of index characteristics is presented in Table 1. For a more detailed description the reader is referred to appendix F.

Table 1: Formulas and normalization conventions for the 4 variability indices under consideration. Q_1 and Q_3 stand for the first (25%) and the third (75%) quartiles of the data. $\{m_i\}_{i=1}^n$ are relative flux values and \bar{m} , \bar{m}_w their arithmetic and weighted means, respectively.

<i>Index</i>	<i>Formula</i>	<i>Type</i>	<i>Normalization</i>
Interquartile range, IQR	$Q_3 - Q_1$	scatter-based	IQR $\simeq \text{IQR}/(1.349\sigma)$
χ^2 statistic, $\tilde{\chi}^2$	$\sum_{i=1}^n \frac{(m_i - \bar{m}_w)^2}{\sigma_i^2}$	scatter-based	$\chi^2 = \tilde{\chi}^2/n - 1$
Standard deviation, σ	$\sqrt{\frac{\sum_{i=1}^n (m_i - \bar{m})^2}{n-1}}$ or equation (2)	scatter-based	$\sigma = \sigma/\bar{m}$
Von Neumann ratio, η	$\frac{\sum_{i=1}^{n-1} (m_{i+1} - m_i)^2 / (n-1)}{\sum_{i=1}^n (m_i - \bar{m})^2 / (n-1)}$	correlation-based	\emptyset

2.3 Period search

Period search is the last step in the classification of variable objects. Since the birth of spectral analysis, different period searching techniques have been developed, from the classical Schuster periodogram to the correntropy kernel periodogram (see [29], [15] resp.). These techniques can be classified in a few broad categories: (1) Fourier methods, based on the Fourier transform and the study of signal power spectra; (2) phase folding methods, where a cost function associated to the phase-folded data is to be minimized; (3) least squares methods, based on fitting a periodic model to the data at each candidate frequency and selecting the most probable frequency through a certain statistical test; (4) Bayesian approaches, which apply Bayesian probability theory to the problem. Despite such diversity of procedures, the Lomb-Scargle periodogram (introduced in [20, 27]) remains one of the most widely used spectral density estimators, specially in the astronomical community. One of the reasons for this is that it falls in both the Fourier and least squares method categories, in addition to its simplicity and ease of use. For a brief introduction to the concept of power spectrum and the effects of window functions on real signals, the reader is referred to section D.1 in the appendix; for a more complete description of the Lomb-Scargle (LS) periodogram see appendix D.2.

2.3.1 Practical aspects of the Lomb-Scargle periodogram

We will now discuss 3 important practical considerations regarding the LS periodogram: frequency limits and spacing, uncertainties in periodogram results and normalization conventions.

Frequency limits and spacing. For a set of observations with time span T , a signal with frequency $1/T$ will complete exactly one cycle. This can be used as the minimum frequency for our periodogram. For the higher frequency limit more options are available, from a Nyquist limit⁵ for non-uniformly sampled data (see [9]) to a frequency limit based on integration time of individual observations. However, a limit based on prior knowledge of the kind of signal under study can also be effective.

With regards to frequency interval, we should avoid too fine (computationally expensive) and too coarse grids (risk of missing narrow peaks), while keeping grid spacings smaller than the expected peak widths. For T -width rectangular windows, a prudent grid size would be $\Delta f = 1/(n_0 T)$. Common choices for n_0 range from 5 to 10, although higher values can also be considered.

Uncertainties in periodogram results. Uncertainties associated to periods derived from LS periodograms can not, in general, be meaningfully expressed as error bars. However, it is common to identify the precision of a peak's frequency with a certain function of its width; we often consider the half-width at half-maximum (HWHM), $\nu_{1/2}$, with some scaling depending on the number of samples, N , and their average SNR, Σ (see [16, §10]):

$$\sigma_{\nu_{\text{peak}}} \simeq \sqrt{\frac{2}{N\Sigma^2}} \nu_{1/2} \quad (3)$$

The latter may be viewed as an accurate characterization of the period precision only if the peak we are referencing to is the correct one. However, if the largest peak is induced by a failure mode (i.e. it is not the true main periodic component; see appendix D.3), the formula above would lose its validity. This is why error bars should be avoided when reporting uncertainties associated to periodogram results.

A more robust quantity for expressing uncertainty in periodogram results is the height of the peak and, in particular, the height compared to spurious background peaks —which depends on both the number of observations and their signal-to-noise ratio. False Alarm Probabilities (FAPs) are the common tool in this context, and allow quantifying peak significance. Scargle studied in [28] the statistical aspects of LS periodograms, and showed that, for a signal with pure Gaussian noise and no periodic components, the LS power at a preselected frequency, $\mathcal{P}^{LS}(\nu_0) \equiv Z$, follows an exponential distribution, that is, its CDF has the form: $P(Z < z) = 1 - \exp(-z)$. Thus, if we consider N frequency samples from a pure noise LS spectrum, ν_1, \dots, ν_N , and define $Z \equiv \max_n \mathcal{P}^{LS}(\nu_n)$, with each $Z_n \equiv \mathcal{P}^{LS}(\nu_n)$ following the previous distribution (and assuming independence), for $1 \leq n \leq N$, it is clear that:

$$\begin{aligned} P(Z > z) &= 1 - P(Z < z) = 1 - P(\max_n \mathcal{P}^{LS}(\nu_n) < z) = \\ &= 1 - \prod_{n=1}^N P(Z_n < z) = 1 - [1 - \exp(-z)]^N \end{aligned} \quad (4)$$

This is precisely the False Alarm Probability associated to a maximum height peak with power z : the probability of measuring a maximum peak with power z or higher conditioned on the assumption of non periodic data with pure Gaussian noise.

For practical purposes, we may invert the previous relation, to find a threshold power level, z_0 , such that only if the obtained maximum power, z , exceeds z_0 , we would claim the presence of a periodic component at the corresponding frequency in our data, with a certain confidence level (i.e. false alarm probability), p_0 . This leads to the following expression for z_0 :

$$z_0(p_0) = -\log[1 - (1 - p_0)^{1/N}] \quad (5)$$

⁵A definition can be found in appendix D.1.

Normalization conventions.

Power spectral density (PSD) normalization. This corresponds exactly to the expression presented in (19).

If the periodogram is constructed without taking measurement errors into account, periodogram units are $\text{units}(y)^2$, and periodogram values can be interpreted as squared-amplitudes of the Fourier power spectrum of the signal at each frequency; on the other hand, if measurement errors are included, periodogram values become unitless and can be interpreted as a measure of periodic content in signal-to-noise ratio instead. It allows direct comparison between Fourier power spectrum of different signals.

Least-squares normalization. In the least-squares interpretation of the LS periodogram (see appendix D.2), a perfect fit⁶ of a sinusoidal model, at frequency ν_0 , yields a maximum periodogram value $\mathcal{P}^{LS}(\nu_0) = \hat{\chi}_0^2/2$, according to equation (21). On the other hand, it is clear that $\hat{\chi}^2(\nu) \leq \hat{\chi}_0^2$ for all $\nu > 0$. Thus, the following normalization is proposed, so that the power of the normalized LS periodogram is dimensionless and ranges between 0 and 1:

$$\mathcal{P}^{LS}(\nu) = \frac{\mathcal{P}^{LS}(\nu)}{\hat{\chi}_0^2/2} = 1 - \frac{\hat{\chi}^2(\nu)}{\hat{\chi}_0^2} \quad (6)$$

This is the standard normalization choice when comparing periodograms from heterogeneous data sets.

We may highlight that, in the standard Lomb-Scargle periodogram, the data is assumed to be pre-centered around the mean value of the (unknown) signal. This is why direct comparison with the least-squares periodogram is possible, given that the same sinusoidal model is used in both cases. However, adding a frequency dependent offset term, y_ν^0 , to the least-squares model, can be advantageous in certain situations (e.g.: when a full phase coverage of the signal is not reflected in the data). This yields the so called *generalized Lomb-Scargle periodogram*, which was thoroughly investigated in [39], and has been proven to exhibit equivalent statistical properties to its standard analogue. In this work, the generalized approach is adopted, making extensive use of the `LombScargle` class from the Python library `Astropy` (see [36]).

⁶Equivalently, $\hat{\chi}^2(\nu_0) = 0$.

3 Light curve reduction methods

In this section, we detail the different methods that were implemented and/or used for light curve reduction of the different objects under study. Codes, as well as raw differential photometry files generated with **AstroImageJ**, can be found in the following GitHub repository: <https://github.com/GuifreSanchez/photometry-data-optimization>.

3.1 Comparison star selection criteria

As stated in 1, careful selection of comparison stars can significantly improve the light curve of the target. We extract differential photometry from raw images using **AstroImageJ** (see appendix C for details), [5], a Java-based visual interface with multiple research-level capabilities, that allows us to perform multi-aperture (differential) photometry in a very practical way, so that adequate comparison stars can be selected at first sight (avoiding bad pixels, stars with high aperture peak pixel count values, etc.). However, this initial selection can be improved substantially if we take into account the following criteria:

- (1) Comparison stars should be selected so that the signal-to-noise ratio of the target is enhanced. An estimate for this ratio is given by $\bar{\phi}/\sigma(\phi)$, where $\bar{\phi}$ is the mean of the target relative fluxes, and $\sigma(\phi)$ its (unbiased) standard deviation. If relative flux errors are to be accounted, $\bar{\phi}$ and $\sigma(\phi)$ should be substituted by their weighted counterparts, $\bar{\phi}_w$ and $\sigma_w(\phi)$ respectively. Thus, post-selection of comparison stars from this initial ensemble, can be driven by minimization of the relative standard deviation (RSD) of the target relative fluxes.
- (2) As pointed in appendix B, adding more comparison stars increases the target SNR by a factor proportional to $\sqrt{\# \text{ comp. stars}}$, while possibly incorporating variable comparison stars. It is then preferable to have an ensemble of non-variable comparison stars. Variability indices can be used to distinguish between constant and variable comparison stars, and determine those from the initial **AstroImageJ** selection that are to be neglected (e.g.: if a certain variability index exceeds a certain threshold).

Two different procedures were designed to optimize comparison star selection, according to the previous criteria, given an initial ensemble determined with **AstroImageJ**.

3.1.1 Target RSD minimization

Given an initial selection of C comparison stars, and a natural number $C' \leq C$, the following algorithm is applied:

- i.* Run over all possible configurations in which k comparison stars are discarded (start with $k = 1$; if $k > 1$ fix as discarded the $\Sigma_i \Gamma_{\star}^i$ comparison stars that are marked with 1s in configuration Γ_{\star} , from step *vi.*).
- ii.* Compute, for each of the configurations in *i.*, the updated relative fluxes (and their errors, if required) for both the target and the non-discarded comparison stars, following equation (8).
- iii.* Compute, again for each of the configurations in *i.*, using the results in *ii.*, the relative standard deviation (RSD) of the target updated relative fluxes.
- iv.* Among all RSD values obtained in *iii.*, keep the smallest $\rho(k)$, $\sigma_1 \leq \dots \leq \sigma_{\rho(k)}$, together with their corresponding configurations, $\Gamma_1, \dots, \Gamma_{\rho(k)}$ (where ρ is an integer-valued function yet to be determined). Configurations are defined as elements in $\{0, 1\}^C$, such that $\Gamma_j^i = 0$ ($\Gamma_j^i = 1$) means that the i th comparison star from the initial ensemble is still selected (discarded) in configuration j .
- v.* Check for comparison stars that are discarded in all the $\rho(k)$ configurations described in *iv.*, and construct a new configuration, with all comparison stars satisfying this property being discarded. The new configuration, Γ_{\star} , can be computed as the product: $\Gamma_{\star}^i = \prod_{j=1}^{\rho(k)} \Gamma_j^i$.
- vi.* If $k < C' - 1$, increase k by 1 and return to *i.*; else stop.

This can be a robust procedure to identify *bad* comparison stars. Great RSD changes between optimal configurations can be attributed to selection or rejection of certain comparison stars from the initial ensemble. We have implemented the previous algorithm with two different ρ functions: $\rho_1(k) = k$ and $\rho_2(k) = N_{\min} = \text{constant}$, for arbitrary $N_{\min} \geq 1$.

Updated relative fluxes for both target and selected comparison stars, ϕ_T^j, ϕ_i^j , corresponding to a certain configuration, Γ_j , can be computed from **AstroImageJ** initial relative fluxes, as follows. Let $\phi_{i_1}, \dots, \phi_{i_d}$ denote the relative fluxes of the discarded comparison stars in configuration Γ_j (i.e. $\Gamma_j^i = 1$ if and only if $i = i_k$ for some $1 \leq k \leq d$), by definition (see equation (1)), we have:

$$\phi_T^j = \frac{\Phi_T}{\sum_{i=1}^C \Phi_i - \sum_{k=1}^d \Phi_{i_k}}, \quad \phi_\ell^j = \frac{\Phi_\ell}{\sum_{i \neq i_k, \ell} \Phi_i} = \frac{\Phi_\ell}{\sum_{i=1}^C \Phi_i - \Phi_\ell - \sum_{k=1}^d \Phi_{i_k}}, \quad \ell \neq i_k \quad (7)$$

Denoting $\Phi_C \equiv \sum_{i=1}^C \Phi_i$, and observing that $\Phi_j/\Phi_C = \Phi_j/(\sum_{i \neq j} \Phi_i + \Phi_j) = \phi_j/(1 + \phi_j)$, we obtain:

$$\begin{aligned} \phi_T^j &= \frac{\Phi_T/\Phi_C}{1 - \sum_{k=1}^d \Phi_{i_k}/\Phi_C} = \frac{\phi_T}{1 - \sum_{k=1}^d \phi_{i_k}/(1 + \phi_{i_k})} \\ \phi_\ell^j &= \frac{\Phi_\ell/\Phi_C}{1 - \Phi_\ell/\Phi_C - \sum_{k=1}^d \Phi_{i_k}/\Phi_C} = \frac{\phi_\ell/(1 + \phi_\ell)}{1 - \phi_\ell/(1 + \phi_\ell) - \sum_{k=1}^d \phi_{i_k}/(1 + \phi_{i_k})} = \\ &= \frac{\phi_\ell}{1 + \phi_\ell - \phi_\ell - (1 + \phi_\ell) \sum_{k=1}^d \phi_{i_k}/(1 + \phi_{i_k})} = \frac{\phi_\ell}{1 - (1 + \phi_\ell) \sum_{k=1}^d \phi_{i_k}/(1 + \phi_{i_k})}, \quad \ell \neq i_k \end{aligned} \quad (8)$$

Updated relative flux errors associated to ϕ_T^j and ϕ_ℓ^j are computed using classical error propagation.

3.1.2 Variability index performance

As described in 2.2, variability indices help quantifying how variable an astronomical object is. We have already presented 4 classical examples of such indices, that have shown to perform adequately in a variety of situations, namely: the interquartile range (IQR); the χ^2 statistic, $\tilde{\chi}^2$, from the χ^2 variability test; the standard deviation, σ ; and the inverse of the von Neumann ratio, $1/\eta$.

For each of the C comparison stars from an initial **AstroImageJ** selection, we can compute the normalized value of each of the aforementioned indices: $\mathbf{IQR}_i, \chi_i^2, \sigma_i, \eta_i^{-1}$, $1 \leq i \leq C$. Two different strategies have been considered to determine which of the previous comparison stars should be discarded based on variability index values:

- (1) Given integers $k_I \geq 1$ and $b_I \in \{0, 1\}$, with $I \in \{\mathbf{IQR}, \chi^2, \sigma, \eta^{-1}\} \equiv \mathcal{I}$, compute normalized index values for indices I s.t. $b_I = 1$, and mark the k_I comparison stars with highest index values (i.e. the k_I worst performing comparison stars); we may denote indices satisfying $b_I = 1$, I_1, \dots, I_p , with $p = \sum_{I \in \mathcal{I}} b_I \leq 4$. Construct then p configurations (same definition as in 3.1.1), $\Gamma_1, \dots, \Gamma_p$, such that $\Gamma_j^i = 1$ if the i th comparison star has been previously marked based on its I_j value. As in point *v.* of the algorithm presented in 3.1.1, check for comparison stars discarded in all configurations, and build the corresponding new configuration, Γ_\star , namely: $\Gamma_\star^i = \prod_{j=1}^p \Gamma_j^i$, $1 \leq i \leq C$.
- (2) Given thresholds $\omega_I \in \mathbb{R}$, and integers $b_I \in \{0, 1\}$ for each index $I \in \mathcal{I}$, compute normalized index values for indices I , I_1, \dots, I_q , s.t. $b_I = 1$. For each of these indices, construct configurations, $\Gamma_1, \dots, \Gamma_q$, s.t. $\Gamma_j^i = 1$ if and only if $I_j^i \geq \omega_{I_j}$, where I_j^i is the value of the (normalized) index $I_j \in \mathcal{I}$ for comparison star i . Finally, also as in 3.1.1.v., construct the *product configuration*, Γ_\star , as $\Gamma_\star^i = \prod_{j=1}^q \Gamma_j^i$, to discard only comparison stars that have been marked in all the configurations, Γ_j , $1 \leq j \leq q$.

These strategies may help identifying variable or noisy comparison stars. Thus, rejecting them as part of the initial ensemble should improve the quality of the target light curve, increasing its overall signal-to-noise ratio, but also smoothing relative flux trends that can lead to the detection of periodicities.

3.2 Filtering procedures

In 2.1, 3 different procedures were exposed that serve as filtering techniques to reduce the presence of outliers in our data (σ -clipping), improve its overall signal-to-noise ratio (SNR numerical clipping) and/or lower its dispersion (binning). All 3 have been implemented and applied, when required, to photometric data. Implementation details can be found in appendix E.2.

4 Results and discussion

We present results from 4 different objects of interest, namely: Wolf 1069, TOI-1266, TZ Ari and GJ 555. Observations were conducted with the 80 cm Joan Oró telescope at Montsec Astronomical Observatory (Catalonia), using its main imaging camera LAIA (a $4k \times 4k$ back illuminated CCD with a pixel scale of 0.4 arcsec and a 30×30 arcmin² field of view) and employing a Johnson R filter. Several slots of images were obtained per night during the visibility periods of each object (observation dates, as well as the number of recorded data points, are given in the corresponding subsections). The obtained images were calibrated with darks, bias and flat-fields (see appendix B) following the ICAT pipeline (see [6]).

Figure 2 shows the results of applying target RSD minimization using the function $\rho_1(k) = k$, for the 4 objects under study. We note there is a similar qualitative behaviour for minimal σ values as a function of the number of selected (or discarded) comparison stars from the original `AstroImageJ` ensemble: we start initially with a decreasing trend, with typically highest σ_{\min} for the first selected comparison stars; then a minimum σ_{\min} is reached, usually for a low number of selected comparison stars (relative to the total number of stars in the initial ensemble); and finally a modest increasing trend is observed, that gradually approaches the *first* RSD value.

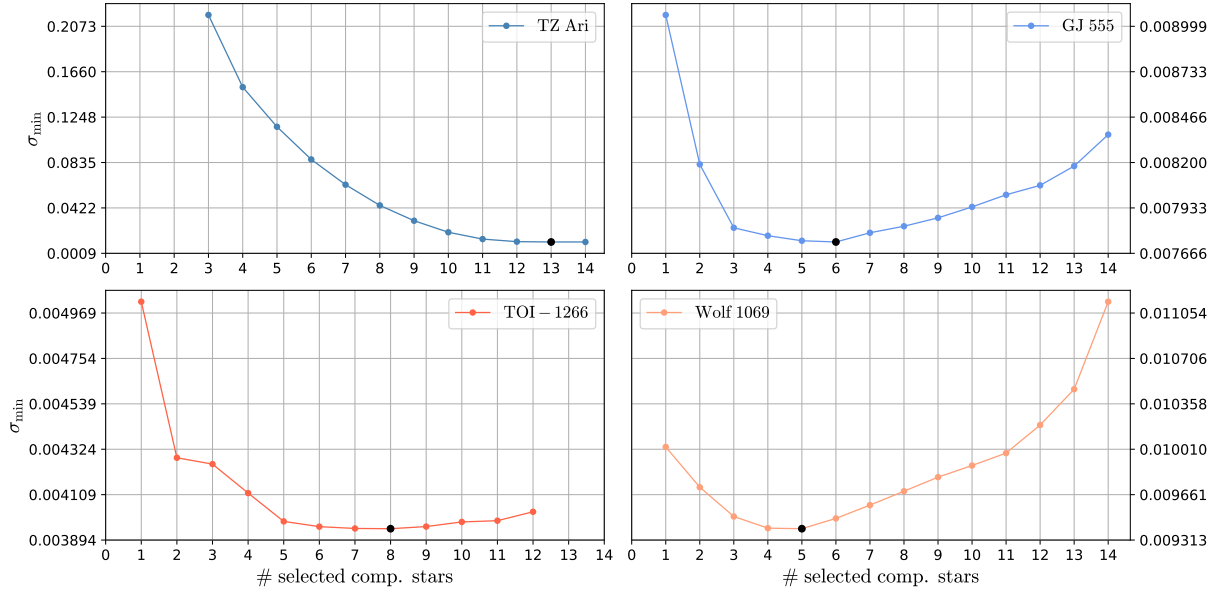


Figure 2: Minimal RSD values for optimal configurations vs. number of selected comparison stars, for the 4 objects under study: TZ Ari, GJ 555, TOI-1266 and Wolf 1069.

The observed behaviour may be explained as follows. Initially, as the number of comparison stars decreases, minimum RSD values also decrease, due to an appropriate compromise between the number of selected stars and the optimization process. At a certain point, removing a comparison star (or several) leads to ensemble configurations where the minimum RSD can no longer decrease; this may be due to the fact that the resulting number of stars in the optimal ensemble is too small (which would minimize the impact differential photometry has on target flux SNRs) and/or to the previously discarded comparison stars, which could concentrate the scatter-induced noise we aim to eliminate. Subsequent rejection of other comparison stars, clearly aggravates the situation, resulting in increased σ_{\min} values, and a greater σ_{\min} is expected when only a few comparison stars (1–3) are selected.

4.1 Wolf 1069

Wolf 1069 is a bright M5 dwarf located in the northern constellation of Cygnus. It has been studied in a number of publications (e.g.: [10, 17, 35]) and its rotation period has been recently estimated to be of about 142.9 days (see [24]). Photometric data was obtained between 19/12/2018 and 09/03/2020 for this late-type star (653 images), and extraction of differential photometry with prior selection of 15 comparison stars was performed later.

Initially, σ -clipping⁷ and SNR numerical clipping were applied directly to raw target and comparison stars relative fluxes, with target and comparison stars σ - and SNR-clipping factors $\alpha_T = 2.25, p_T = 0.25$ and $\alpha_C = 4.00, p_C = 0.00$ (i.e. no SNR clipping⁸ is performed on comparison star fluxes). Setting minimum and maximum frequencies $\nu_{\min} = 0.002245 \text{ d}^{-1}$ and $\nu_{\max} = 0.1 \text{ d}^{-1}$ (following the presented criteria, with min. freq. equal to $(t_{\max} - t_{\min})^{-1}$ and assuming a maximum frequency of at most 10 days per cycle, since previous analysis with $\nu_{\max} = 0.5 \text{ d}^{-1}$ showed no significant peaks between frequencies 0.1 and 0.5), and grid factor $n_0 = 100$, LS periodogram⁹ results show peak frequencies¹⁰ at $\nu_r = 0.0051 \pm 0.0010 \text{ d}^{-1}$ and $\nu_f = 0.0052 \pm 0.0011 \text{ d}^{-1}$ for raw and filtered data, respectively, corresponding to first period estimates of $198 \pm 40 \text{ d}$ and $191 \pm 40 \text{ d}$. We note these results are, given the reported uncertainties, close to the 142.9 d period presented in [24]. LS least-squares normalized powers, associated to peak frequencies are $\mathcal{P}_r^{LS} = 0.5612$ and $\mathcal{P}_f^{LS} = 0.6193$. Having $\mathcal{P}_f^{LS} > \mathcal{P}_r^{LS}$ for similar peak frequency values favours the identification of periodic behaviour associated to target fluxes. Figure 3 helps visualizing these results, and shows clearly less dispersion with respect to LS best-model fit for peak frequencies in the case of filtered data. It also allows identifying the aforementioned periodic trend, particularly between times $\sim 2458800 \text{ d}$ and 2458900 d in BJD TDB¹¹, since more data points were collected then.

Optimized selection of comparison stars from the initial **AstroImageJ** ensemble is then performed. A configuration that minimizes the relative standard deviation of the target relative fluxes is found using target RSD minimization as presented in 3.1.1 (using $\rho = \rho_1$; see step *iv.*), with an optimal σ of 0.00940037. The number of discarded comparison stars for such configuration is 10. The ratio between maximum and minimum RSD values is ~ 1.2 . With same periodogram and filtering parameters as before, LS periodogram results for filtered data (using the obtained configuration) are: $\nu_{f,*} = 0.0051 \pm 0.0011 \text{ d}^{-1}$, $\mathcal{P}_{f,*}^{LS} = 0.5365$, with a corresponding period estimate of $196 \pm 43 \text{ days}$. We note that LS power at peak frequency with this minimal RSD configuration is slightly less than that obtained with raw data (without filtering and RSD minimization). This may be due to the small number (5) of comparison stars that were used in this case: RSD may be minimized excessively in this situation, effectively eliminating valuable flux variability information from the target light curve; the chosen comparison stars are not sufficient to improve the target photometry.

Power threshold (PT) values associated to 0.1%, 0.01% and 0.001% False Alarm Probabilities (FAPS) were computed for both raw and filtered data, with results in the range 0.04-0.08 (and filtered PTs slightly above raw PTs in the two described scenarios, i.e. with and without optimized selection of comparison stars). Taking into account that LS peak powers have been of order ~ 0.5 , we may claim the presence of a periodic component (with period around 195 days) in the observed signal, at a confidence level of at least 0.01%.

Other optimal configurations were tested, to validate previous observations. With 2 and 4 discarded comparison stars, denoting their configurations Γ^2 and Γ^4 resp. (using same filtering parameters as in the preceding situations), LS powers at peak frequencies improve slightly, with values of 0.5711 and 0.5539, respectively (peak frequencies are very similar, with corresponding periods of around 195 days). In the first case, we note the obtained power is greater than \mathcal{P}_r^{LS} . Using configuration Γ^2 and a less restrictive target filtering ($\alpha_T = 2.75, p_T = 0.20, \alpha_C = 3.5, p_C = 0.00$), the corresponding LS maximum power also increases to 0.6131, with a peak period of about 202 days, which is the most distanced from the obtained for raw data. Nonetheless, it remains below \mathcal{P}_f^{LS} , where no optimized selection of comparison stars is applied.

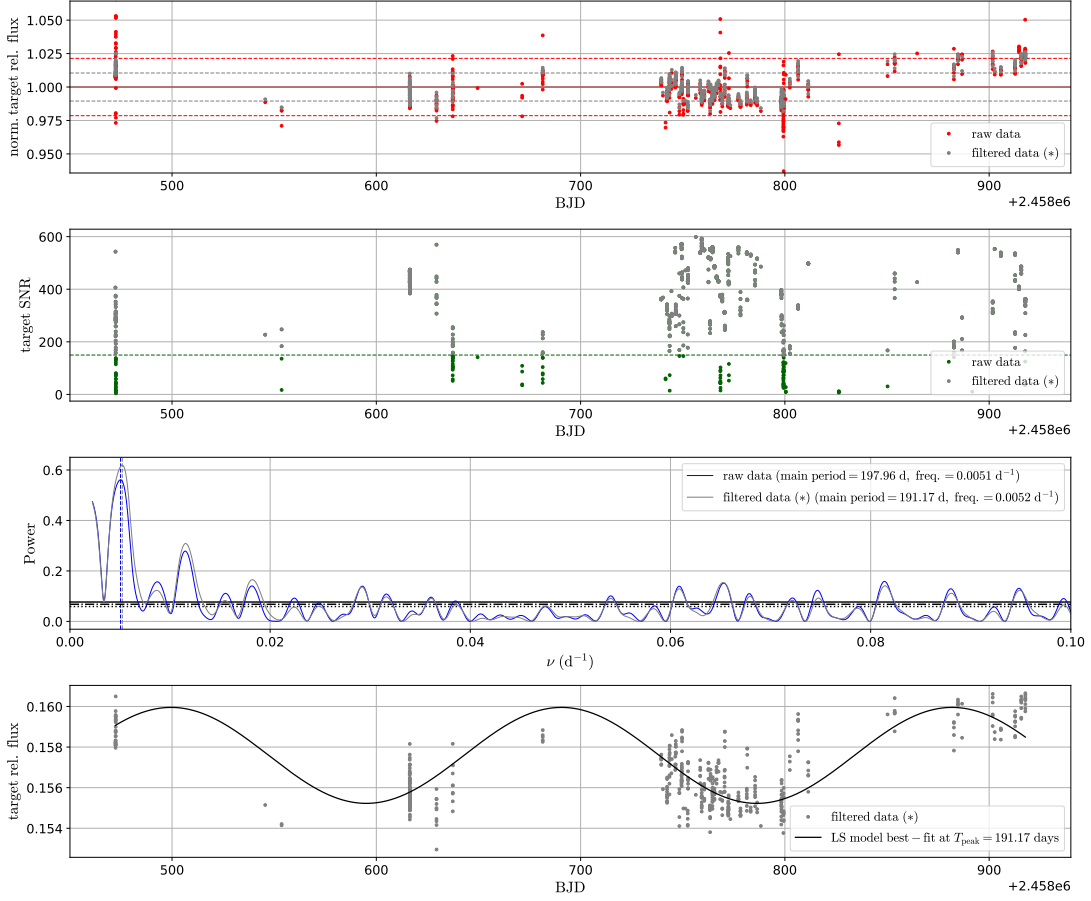
⁷All σ -clippings presented throughout the following discussion, were performed with respect to the arithmetic mean of the data, and a maximum of 4 iterations.

⁸We expect similar SNR levels in comparison star fluxes as in the target, since these are selected not far from it; thus, SNR clipping on target fluxes will effectively eliminate low SNR fluxes from comparison stars. This approach will be repeated for the remaining 3 astronomical objects under study.

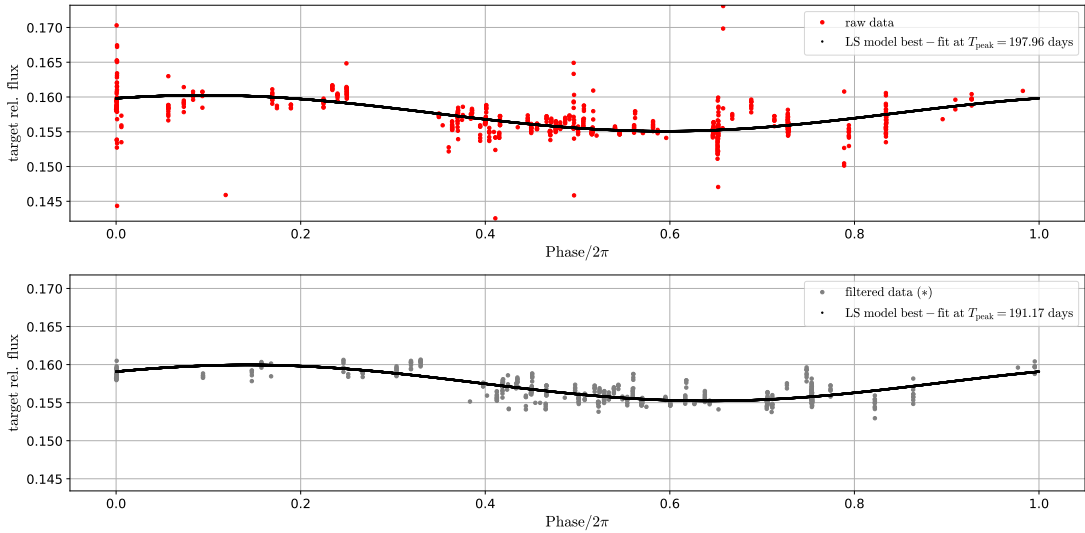
⁹As stated, throughout this section we consider the *generalized Lomb-Scargle periodogram*. All results are obtained using the **LombScargle** class from **Astropy**. A summary of its different features can be found here: <https://docs.astropy.org/en/stable/timeseries/lombscargle.html>.

¹⁰Frequency errors have been estimated using the HWHM, $\nu_{1/2}$, of the corresponding peak. Period error estimates are obtained using classical error propagation.

¹¹Barycentric Julian Date in Barycentric Dynamical Time (solar days).



(a) From top to bottom. (1) Normalized target relative flux (with respect to the mean) vs. BJD: raw (red) and filtered (gray) data are presented together; dashed lines represent (normalized) standard deviations above and below mean relative flux values for both raw and filtered data. (2) Target signal-to-noise ratio vs. BJD: raw (green) and filtered (gray) data are presented together, and separated by a dashed (green) line at $\max_i \{ \text{SNR}_i \} p_T$ (numerical clip). (3) LS periodogram for raw (blue) and filtered (gray) data; solid, dash-dotted and dotted horizontal lines represent power thresholds (filtered data) for 0.1%, 0.01% and 0.001% FAP values. (4) LS model best-fit at peak frequency ν_f (for filtered data), together with filtered (unnormalized) target relative fluxes.



(b) Phase-folded (true) target relative fluxes corresponding to raw (red, top) and filtered (gray, bottom) data, with max. power periods from LS periodogram results.

Figure 3: (a) Raw and filtered data comparison for Wolf 1069 photometry, with LS periodogram results and best-model fit for filtered data (used information 76.42%). (b) Phase-folding of raw and filtered relative fluxes with best-model fit from LS periodogram results.

4.2 TOI-1266

TOI-1266 has been recently confirmed in [33] as a host star with two orbiting exoplanets, the inner one (with a 10.9 d orbital period) being a sub-Neptune-sized planet, and the outer one (with a 18.8 d orbital period) residing in the transition region between rocky and gaseous planets (the so called *exoplanet radius valley*). As a TESS (Transit Exoplanet Survey Satellite, [26]) Object of Interest, some stellar parameters had already been studied prior to the aforementioned work, and are publicly available at <https://exofop.ipac.caltech.edu/tess/target.php?id=467179528>; it is classified as an M2 dwarf (i.e. cooler than Wolf 1069). TOI-1266 photometry was obtained between dates 26/03/2020 and 01/06/2020 (788 images). A total of 13 comparison stars were selected using **AstroImageJ** to perform differential photometry.

In first place, σ -clipping and SNR clipping were applied to both target and comparison stars relative fluxes. Filtering parameters were set as follows: (target) $\alpha_T = 2.5, p_T = 0.3$ and (comparison stars) $\alpha_C = 4.0, p_C = 0.0$. Minimum and maximum periodogram frequencies, as well as frequency grid spacing, were determined as in 4.1, with: $\nu_{\min} = 0.015085 \text{ d}^{-1}$, $\nu_{\max} = 0.5 \text{ d}^{-1}$ and grid factor $n_0 = 100$. LS periodogram results show raw and filtered peak frequencies at $\nu_r = 0.0259 \pm 0.0065 \text{ d}^{-1}$ and $\nu_f = 0.0255 \pm 0.0059 \text{ d}^{-1}$ with powers $\mathcal{P}_r^{LS} = 0.2618$ and $\mathcal{P}_f^{LS} = 0.3291$, respectively. As in the case of Wolf 1069, we observe that $\mathcal{P}_f^{LS} > \mathcal{P}_r^{LS}$, which reinforces the validity of the obtained results, given that peak frequencies are relatively close. Period estimates corresponding to ν_r and ν_f are (raw data) $38.5 \pm 9.7 \text{ d}$ and (filtered data) $39.2 \pm 9.1 \text{ d}$.

After optimized selection of comparison stars, through RSD minimization (with $\rho = \rho_1$), we obtain a minimal RSD value of 0.00394732 and a corresponding configuration with (only) 5 discarded stars. LS periodogram results with this new configuration, applying the same filtering procedures as before are: $\nu_{f,*} = 0.0255 \pm 0.0057 \text{ d}^{-1}$, $\mathcal{P}_{f,*}^{LS} = 0.3318$ (with a period estimate of $39.2 \pm 8.8 \text{ d}$). We note that $\mathcal{P}_{f,*}^{LS} > \mathcal{P}_f^{LS} > \mathcal{P}_r^{LS}$. These inequalities indicate that the proposed comparison star selection method can be combined with filtering techniques to improve periodogram results. Hence, this confirms the periodic nature of the recorded signal, with a robust LS peak period of ~ 40 days.

PT values associated to 0.1%, 0.01% and 0.001% FAPS were also computed for TOI-1266, with results ranging between 0.04 and 0.06 for both raw and filtered data, and with filtered PTs generally greater than raw PTs. Similar to what we encountered in the case of Wolf 1069, TOI-1266 LS peak powers are of order ~ 0.3 , which confirms the presence of a periodic component (with period around 40 days) in the observed signal, at a confidence level of at least 0.01%.

In Figure 9 a comparison between raw and filtered, RSD minimized data is shown. We observe that the dispersion in phase-folded filtered data with respect to the LS model best-fit at peak frequency is noticeably less than that of raw data (Figure 9b). On the other hand, we can see in Figure 9a that relative fluxes from filtered data do seem to follow an oscillatory trend, which reasonably fits the LS model at peak period. Finally, we note that the peaks in the Lomb-Scargle periodograms for both raw and filtered data are not excessively narrow, which may translate, according to equation (3), in higher uncertainty.

4.3 TZ Ari

TZ Ari is another bright M dwarf in the constellation of Aries (other usual designations are L 1159-16 and GJ 83.1). The same study that reported the detection of a planetary system in GJ 555, [37], also identified 3 candidate exoplanets orbiting this star, with mass uncertainties also particularly high, and mass values in the range 4–80 M_{\oplus} ; orbital periods were found to be ~ 773.4 d, 243.1 d and 1.93 d; however, it has not been confirmed whether the 1.93 d period can be attributed to an orbiting exoplanet. It is known to be a *flare star* (meaning that great changes in brightness during short periods of time can occur); however, long-term brightness variations have also been observed, and could possibly be attributed to starspots, [12]. TZ Ari photometry data was collected between dates 06/02/2019 and 12/03/2020, and differential photometry was computed using an initial ensemble of 15 comparison stars.

3646 relative flux values were computed for both target and comparison stars, with target SNR values generally above 200. This allows for a more restrictive light curve filtering. Thus, filtering parameters have been set in this case: (target) $\alpha_T = 3.00$, $p_T = 0.35$ and (comp. stars) $\alpha_C = 4.00$, $p_C = 0.00$. A general long-term trend is clearly observed in the target relative fluxes (both for raw and filtered data), which motivates the application of a prewhitening scheme. We proceed in a similar way as in [2, 7], with a technique that may be referred to as *iterative prewhitening*. In first place, we calculate the Lomb-Scargle periodogram¹² on filtered data, ϕ_i^0 , and select, as usual, the frequency of the peak with highest power, ν_{peak}^0 . We then subtract the LS model best-fit at ν_{peak}^0 to the previous relative fluxes, and obtain *shifted* (detrended) relative flux values, ϕ_i^1 . The same procedure is applied to these fluxes, which yields another peak frequency, ν_{peak}^1 , and new shifted relative fluxes, ϕ_i^2 . This process is iterated 2 more times. Iterative prewhitening, in the context of spectral analysis, helps removing periodic biases in our data, that may be induced by systematic errors (e.g. window function aliases associated to observation times), but also allows us to estimate the harmonic decomposition of the signal under study. Hence, within the first few iterations, given the long-term periodicity initially observed, we should identify other important periodic components in the relative fluxes of TZ Ari. Minimum and maximum frequencies used to construct LS periodograms were $\nu_{\text{min}} = 10^{-6} \text{ d}^{-1}$ and $\nu_{\text{max}} = 0.60 \text{ d}^{-1}$, with a grid factor of $n_0 = 100$. We chose $\nu_{\text{min}} \ll 1$ to account for large periods, as prewhitening was applied—a more direct inspection of the signal would have very likely required a greater value for ν_{min} .

Iterative prewhitening results can be visualized in Figure 4. Table 2 contains (filtered data column) numerical peak frequencies, periods and associated powers corresponding to LS results from prewhitening. We note that initially (iteration 0), the obtained peak frequency for filtered data performs better (i.e. has higher power) than that of raw data, while being relatively close to the latter. This is consistent with our previous analysis of Wolf 1069 and TOI-1266 photometry, and confirms the aforementioned long-term periodicity, with period estimates of around 320 days for both raw and filtered data. We note that subsequent prewhitening iterations produce much smaller power maxima, ~ 0.05 , compared to initial power peaks at ~ 0.2 . However, in Figure 4a, a narrow frequency peak centered near 0.51 d^{-1} can be clearly distinguished, while the main low-frequency component of the signal, has vanished almost completely, due to prewhitening—note that, as expected from the applied procedure, the obtained spectrum between frequencies 0.0–0.1 after detrending of the long-term variability resembles that of white noise, i.e. with no distinguished peaks. A period estimate of $\nu_f^1 = 1.96$ days is obtained in this case. Iterations 2 and 3 from prewhitening provide period estimates between 20 and 25 days; however, the obtained peaks are not as clear as the previous one, and have both slightly less power. Phase folded data corresponding to iterations 0 and 1 is represented in Figure 4b; we note that the LS model at peak frequency fits reasonably well with the original data (iteration 0), while the ~ 2 day trend assumed for iteration 1-fluxes is more subtle.

¹²In this case, relative flux errors are only taken into account to compute the initial LS periodogram. Periodogram results from iterations ≥ 1 do not make use of such errors (this corresponds to the usual approach; however, we tested both cases and appreciated no significant differences between the obtained periodograms).

We have also performed an optimized selection of comparison stars through variability indices¹³, an carried out a similar analysis using the same prewhitening scheme. More specifically, we have followed method (1) in 3.1.2, computing only σ and χ^2 index values, and comparing the worst 10 performing¹⁴ comparison stars in each case, i.e. $k_\sigma = k_{\chi^2} = 10$, $b_\sigma = b_{\chi^2} = 1$ and $b_I = 0$ if $I \neq \sigma, \chi^2$; a total of 7 stars were discarded. LS periodogram iterative prewhitening results associated to such an ensemble, with same filtering parameters as before, are summarized also in Table 2. Figure 8 contains relative flux light curves from the 7 discarded comparison stars, together with individual variability index charts that help visualizing their overall *index performance*. Iteration 1 in this situation yields (almost) the same peak frequency as before, i.e. $\nu_f^1 \simeq \nu_{f,*}^1 = 0.5110 \text{ d}^{-1}$, with slightly higher power; more specifically, with a $\sim 23\%$ gain. This is the highest relative difference between powers from filtered data with star selection optimization and powers from filtered (only) data. Figure 5 provides the corresponding LS periodograms, and allows making the aforementioned difference visible. Phase-folded (shifted) relative fluxes from iteration 1 are represented also in Figure 5, to allow for direct comparison with phase-folded data in Figure 4b.

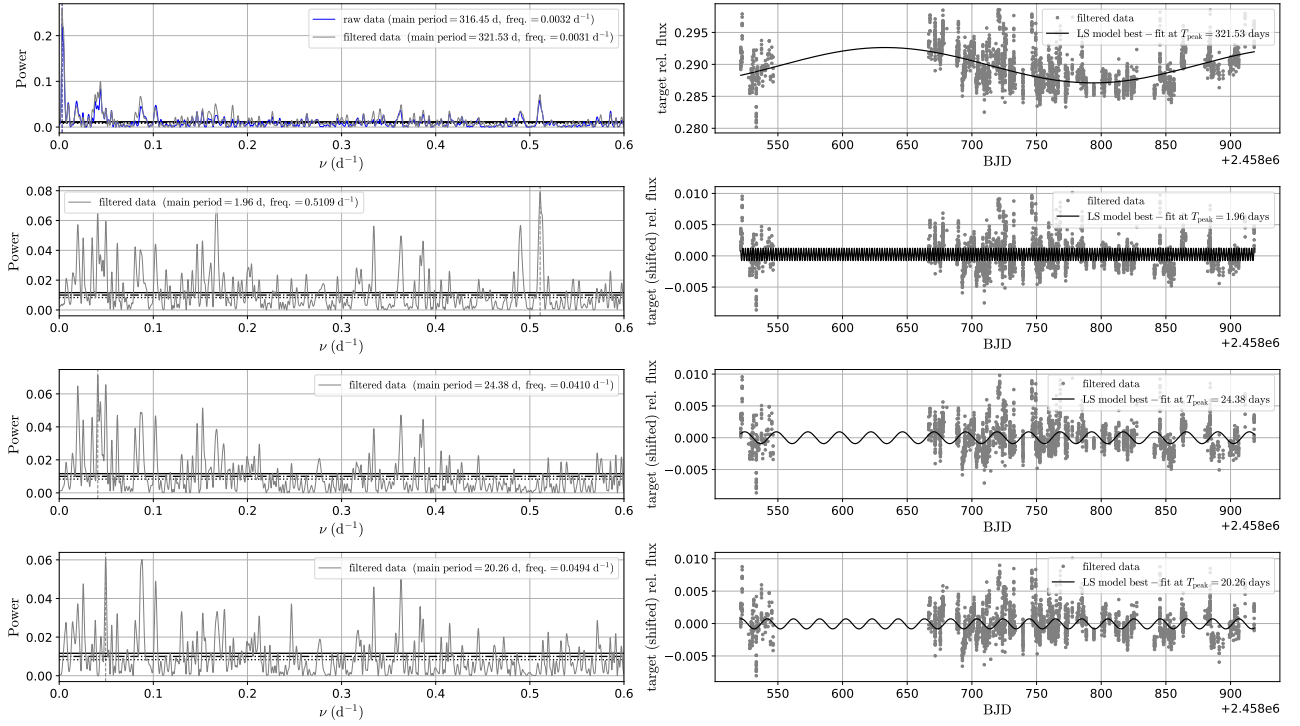
The obtained results reinforce the initial hypothesis of an intrinsic periodic component in the signal with a period estimate of 1.96 days. The observed long-term variations in the original data may be due to seasonal biases induced by the observational time span, which is considerably large (February 2019–March 2020). Moreover, while close to 2.00 days, a period of 1.96 days is very unlikely to correspond to a 2.00 (or 1.00) days window function, given the notable number of measured fluxes. On the other hand, as a young active red dwarf, TZ Ari may exhibit a small rotation period, which could be in direct relation with the found spectral peak. This period estimate has also been observed¹⁵ in other photometric time-series and RV measurements from the next-generation CARMENES spectrograph¹⁶.

¹³Computation of variability indices is performed always after relative fluxes have been filtered, as recommended in 3.2, due to the presence of outliers that can be eliminated using σ -clipping.

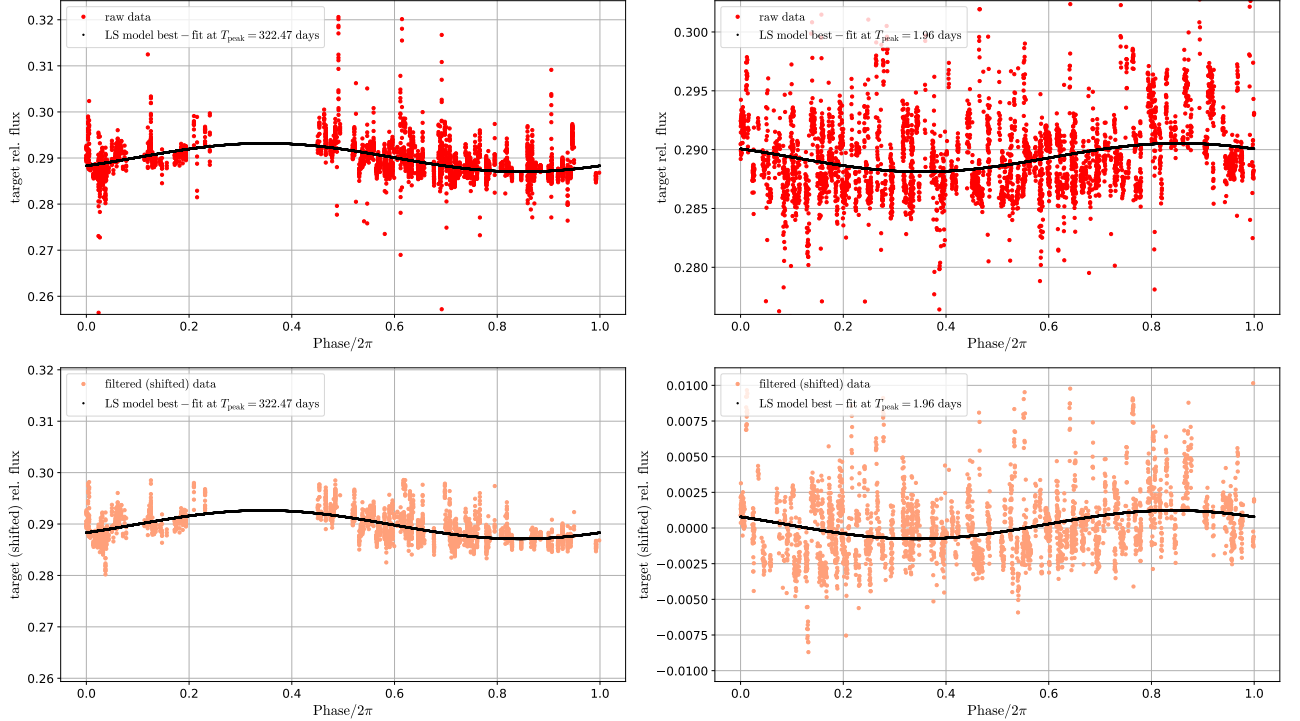
¹⁴Other values for k_σ and k_{χ^2} were tested but no improvement was observed w.r.t. peak power values.

¹⁵This information refers to a private communication from A. Quirrenbach (Zentrum für Astronomie der Universität Heidelberg), with IEEC members, who are currently collaborating in the analysis of photometric and RV measurements from TZ Ari.

¹⁶Calar Alto high-Resolution search for M dwarfs with Exoearths with Near-infrared and optical Échelle Spectrographs (for more information see: <https://carmenes.caha.es/ext/instrument/index.html>).



(a) TZ Ari successive LS periodograms from iterative prewhitening, together with LS model best-fits for filtered, shifted fluxes at peak frequencies. In the first row (initial step), both raw (blue) and (gray) filtered data LS periodograms are represented. Solid, dash-dotted and dotted horizontal lines indicate filtered data power thresholds at 10%, 1% and 0.01% False Alarm Probabilities, respectively.



(b) Phase-folded raw (top) and filtered, shifted (bottom) target relative fluxes, together with LS model best-fits at peak frequencies, corresponding to iterative prewhitening stages 0 (left, $T_{\text{peak}} = 322.47$ days) and 1 (right, $T_{\text{peak}} = 1.96$ days).

Figure 4: (a) LS periodograms and best-fits from iterative prewhitening applied to TZ Ari (used information 79.81%). (b) Phase-folding results from initial raw and filtered data and first prewhitening iteration.

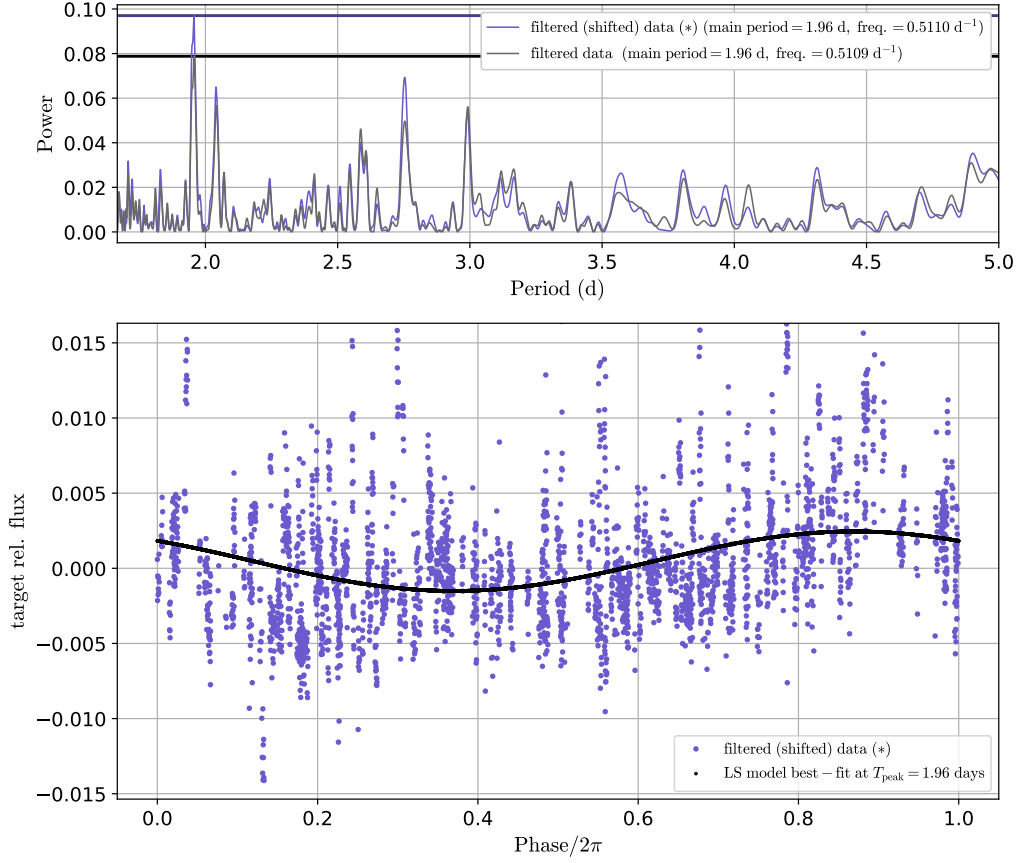


Figure 5: (Top) TZ Ari Lomb-Scargle periodograms for filtered (shifted) data from first prewhitening iteration with optimized comp. star selection through σ - and χ^2 -performance (blue) and without it (gray). Horizontal lines represent highest (peak) powers, respectively (used information 79.81%). (Bottom) Phase-folded relative fluxes and LS model best-fit at peak frequency for filtered (shifted, with optimized star selection) data.

Table 2: Iterative prewhitening results from Lomb-Scargle periodograms on raw, filtered (shifted), and filtered (shifted, with optimized star selection through variability indices, here symbolized with *) TZ Ari differential photometry data.

TZ Ari				
Iteration	Peak variable	Raw data	Filtered (shifted) data	Filtered (shifted, *) data
0	Frequency (d^{-1})	0.0032 ± 0.0013	0.0031 ± 0.0015	0.0028 ± 0.0018
	Period (d)	$(3.2 \pm 1.3) \cdot 10^2$	$(3.2 \pm 1.6) \cdot 10^2$	$(3.6 \pm 1.6) \cdot 10^2$
	Power, \mathcal{P}^{LS}	0.2260	0.2569	0.2313
1	Frequency (d^{-1})		0.5109 ± 0.0022	0.5110 ± 0.0026
	Period (d)	\emptyset	1.9573 ± 0.0086	1.957 ± 0.010
	Power, \mathcal{P}^{LS}		0.0787	0.0970
2	Frequency (d^{-1})		0.0410 ± 0.0022	0.0198 ± 0.0024
	Period (d)	\emptyset	24.4 ± 1.6	50.4 ± 6.7
	Power, \mathcal{P}^{LS}		0.0712	0.0736
3	Frequency (d^{-1})		0.0494 ± 0.0013	0.1527 ± 0.0013
	Period (d)	\emptyset	20.26 ± 0.55	6.550 ± 0.056
	Power, \mathcal{P}^{LS}		0.0612	0.0514

4.4 GJ 555

GJ 555 is a red dwarf of spectral type between TOI-1266 and Wolf 1069 (M3.5), located in the constellation of Libra (it is also designated HN Lib). A planetary system was recently detected (see [37]), and several orbital parameters, together with its mass, were estimated (the reported orbital period was $449.6^{+6.8}_{-5.9}$ d). However, mass uncertainties remain high, with a reported mass estimate of $30.1^{+16.3}_{-14.6} M_{\oplus}$. Monitoring photometric variability induced by stellar activity could help reducing noise triggering such uncertainties. Thus, photometric measurements were performed between dates 08/01/2019 and 20/05/2020 (2739 images), with later extraction of differential photometry by careful selection of 15 comparison stars using *AstroImageJ*.

We conduct a similar analysis of photometric data to the case of TZ Ari. Filtering parameters are set to usual (less restrictive) values: $\alpha_T = 3.00$, $p_T = 0.15$, $\alpha_C = 4.25$, $p_C = 0.00$. Iterative prewhitening is then applied to filtered data, with a grid factor of $n_0 = 100$, and minimum and maximum periodogram frequencies of $\nu_{\min} = 10^{-4} \text{ d}^{-1}$ and $\nu_{\max} = 0.25 \text{ d}^{-1}$ (results are summarized in Table 3 and Figure 10). Optimized selection of comparison stars through variability indices is performed later, with indices **IQR** (interquartile range) and η^{-1} (inverse of the von Neumann ratio); an optimal 3 star configuration is obtained by direct comparison between the worst 4 performing¹⁷ stars in each case (see 3.1.2 (1)), i.e. $k_{\text{IQR}} = k_{\eta^{-1}} = 4$, $b_{\text{IQR}} = b_{\eta^{-1}} = 1$, and $b_I = 0$ if $I \neq \text{IQR}, \eta^{-1}$. After filtering of the updated target fluxes, iterative prewhitening is again executed, with same parameters as before (results can be found in the last column of Table 3).

In first place, we notice that periodogram results associated to raw data differ significantly from their filtered counterparts. This may be due to certain biases dominating raw data trends, that are effectively suppressed after application of the corresponding clipping procedures. On the other hand, Figure 10a shows very clearly the periodic trend of (iteration 1) filtered data, that is notably close to the LS model best-fit at peak frequency (which can also be appreciated in the corresponding phase-fold diagram in Figure 10b). The obtained period in this case is 109.5 days, which has a clear peak in the LS periodogram (well above the peak threshold at FAP 0.1%, and clearly distinguished among the majority of peaks within the studied frequency range). Prewhitening iterations 2 and 3 lead to LS periodograms with increased noise. However, iteration 2 produces a peak period of around 350 days, with power ~ 0.1 (i.e. of the same order of powers found in iterations 0 and 1), which may be taken into account (filtered data could be initially detrended with a sinusoidal fit at this frequency; this could improve results from iterations 0 and 1, but we do not expect radically different periodogram estimates, given the visible oscillatory nature of the measurements).

With optimized selection of comparison stars through **IQR**- and η^{-1} -performance, similar peak period and power values are obtained for prewhitening iterations 0 and 1, while iterations 2 and 3 yield much smaller periods. As in the previous situation, noise levels in the last two LS periodograms are noticeably high, and peak frequencies are obtained at considerably smaller power values. This indicates that the main periodic components of our signal may be concentrated in filtered (shifted) data from the previous two stages of the prewhitening. Periods found in iterations 0 and 1 have relative differences¹⁸ of about $\sim 1.4\%$ in both cases and, in the case of iteration 1, a $\sim 20\%$ power gain is observed (this can be readily seen in Figure 6 where both periodograms are presented at once). These considerations certify peak frequencies encountered through the first two prewhitening stages as the main periodic components in our photometric data, particularly in the case of the ~ 110 days component from iteration 1, given the obtained power gain, but also the *goodness* of fit depicted in Figures 10b and 6. Finally, the fact that the corresponding periods (~ 65 days and ~ 110 days resp.) are not excessively far from each other, suggests that the true signal may be dominated by a single periodic component, with period in the range 70–100 days.

¹⁷Other values for k_{IQR} and $k_{\eta^{-1}}$ were tested but no improvement was observed w.r.t. peak power values.

¹⁸With respect to filtered data results without “reselection” of comparison stars (i.e. column 4 in Table 3).

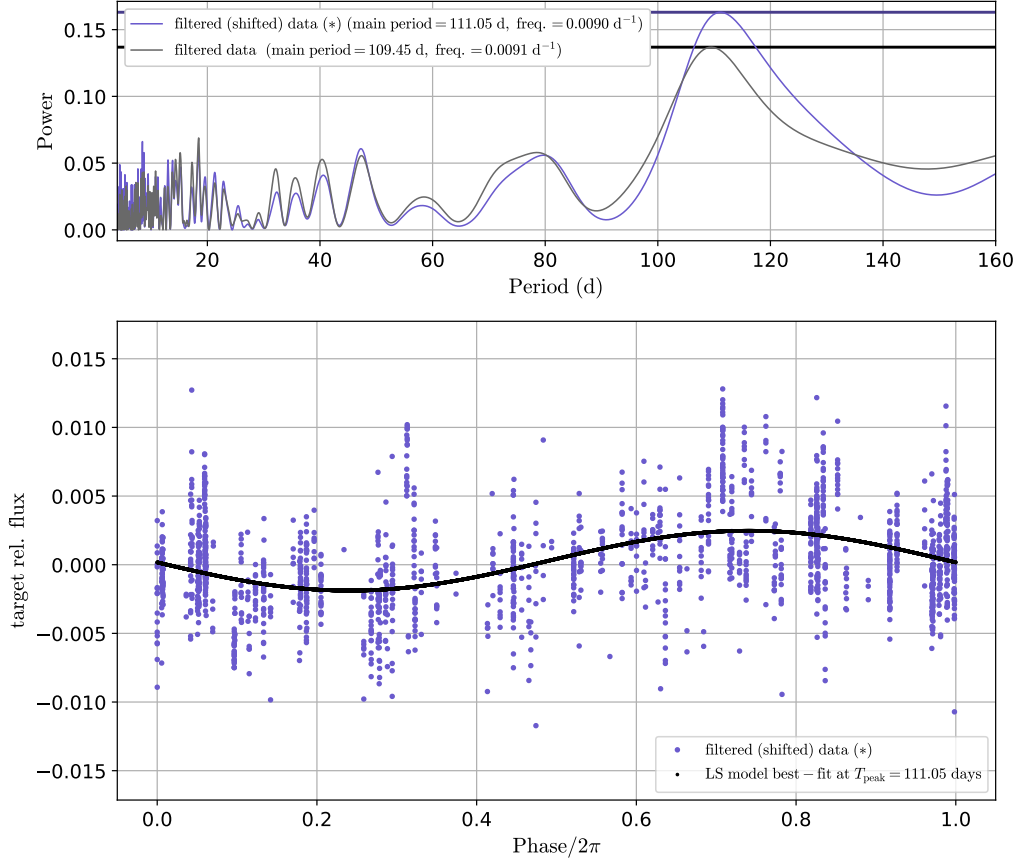


Figure 6: (Top) GJ 555 Lomb-Scargle periodograms for filtered (shifted) data from first prewhitening iteration with optimized comp. star selection through **IQR**- and η^{-1} -performance (blue) and without it (gray). Horizontal lines represent highest (peak) powers, respectively (used information 89.46%). (Bottom) Phase-folded relative fluxes and LS model best-fit at peak frequency for filtered (shifted, optimized star selection) data.

Table 3: Iterative prewhitening results from Lomb-Scargle periodograms on raw, filtered (shifted), and filtered (shifted, with optimized star selection through variability indices, here symbolized with *) GJ 555 differential photometry data.

GJ 555				
Iteration	Peak variable	Raw data	Filtered (shifted) data	Filtered (shifted, *) data
0	Frequency (d^{-1})	0.00340 ± 0.00056	0.011537 ± 0.00093	0.01559 ± 0.00074
	Period (d)	294 ± 48	65.0 ± 3.9	64.2 ± 3.0
	Power, \mathcal{P}^{LS}	0.0981	0.1153	0.1099
1	Frequency (d^{-1})		0.0091 ± 0.0011	0.00901 ± 0.00093
	Period (d)	\emptyset	109 ± 13	111 ± 11
	Power, \mathcal{P}^{LS}		0.1368	0.1630
2	Frequency (d^{-1})		0.00286 ± 0.00093	0.06615 ± 0.00013
	Period (d)	\emptyset	$(3.5 \pm 1.1) \cdot 10^2$	15.12 ± 0.30
	Power, \mathcal{P}^{LS}		0.1135	0.0773
3	Frequency (d^{-1})		0.02473 ± 0.00074	0.22571 ± 0.00093
	Period (d)	\emptyset	40.4 ± 1.2	4.430 ± 0.018
	Power, \mathcal{P}^{LS}		0.0613	0.0522

5 Conclusions

The main goal of this work was to present, study and implement a number of photometry data optimization techniques aimed at improving differential light curves from stars with low-amplitude flux variations. The leading motivation was the necessity for light curve reduction methods that allow for modelling stellar activity induced brightness variations, as these can be viewed as a source of noise with regards to exoplanet detection and characterization. Two different methods were designed to optimize comparison star selection (target RSD minimization and variability index performance) and three different filtering procedures were explored (σ -clipping, numerical clipping and binning). Photometric measurements from 4 M dwarfs (Wolf 1069, TOI-1266, TZ Ari and GJ 555) were studied using the aforementioned methodologies and Lomb-Scargle periodogram analysis was performed for each of them. From the obtained results we make the following general observations:

- Peak powers in LS periodogram results are generally higher for filtered data than for raw data, with gain ratios ranging from 10% to 25%; in the cases of Wolf 1069, TOI-1266 and TZ Ari, filtered peak frequencies are close to their raw analogues, with a maximum relative difference of 3.6% observed for Wolf 1069. This confirms the benefits of the applied filtering procedures with regards to the improvement of photometric data, as normalized LS powers are a measure of periodic content in signal-to-noise ratio.

On the other hand, star selection methods, in combination with filtering techniques, have also proven to slightly improve light curve periodogram results. This has been observed in TOI-1266 (RSD minimization), TZ Ari (σ -, χ^2 -performance) and GJ 555 (**IQR**-, η^{-1} -performance), where we obtained power gain ratios, with respect to filtered (only) data, of $\sim 20\%$ (at main periodic components, see Table 4). This validates the aforementioned methods as potential starting points for the development of more sophisticated selection tools, although investigation of their capabilities in larger data sets is clearly advised.

- Maximum power period estimates for the 4 studied objects are summarized in Table 4. These are to be understood as first analysis results, and relevant periodic components mentioned in the previous discussion (see Section 4) should be taken into consideration.

Table 4: Maximum power period estimates for Wolf 1069, TOI-1266, TZ Ari and GJ 555 and light curve reduction methods used in each case. σ and SNR stand for σ -clipping and numerical clipping of flux SNR values.

	WOLF 1069	TOI-1266	TZ ARI	GJ 555
Period (d)	191 ± 40	39.2 ± 8.8	1.957 ± 0.010	111 ± 11
Power, \mathcal{P}^{LS}	0.6193	0.3318	0.0970	0.1630
Filtering	σ , SNR	σ , SNR	σ , SNR	σ , SNR
Star selection	\emptyset	RSD minimization	σ, χ^2 performance	IQR , η^{-1} performance

We attribute these periodicities to starspots and stellar rotation, given the observations made in Section 2; moreover, the obtained results are, in general, far from orbital period estimates for the reported candidate exoplanets, which is consistent with our thesis. In the case of TZ Ari, as similar results have been obtained through radial velocity measurements and other photometric time-series, the reported period of around 1.96 days may be strongly correlated with the rotation period of the star (possibly disproving the presence of the third 1.93 d period exoplanet).

A Complementary figures and tables

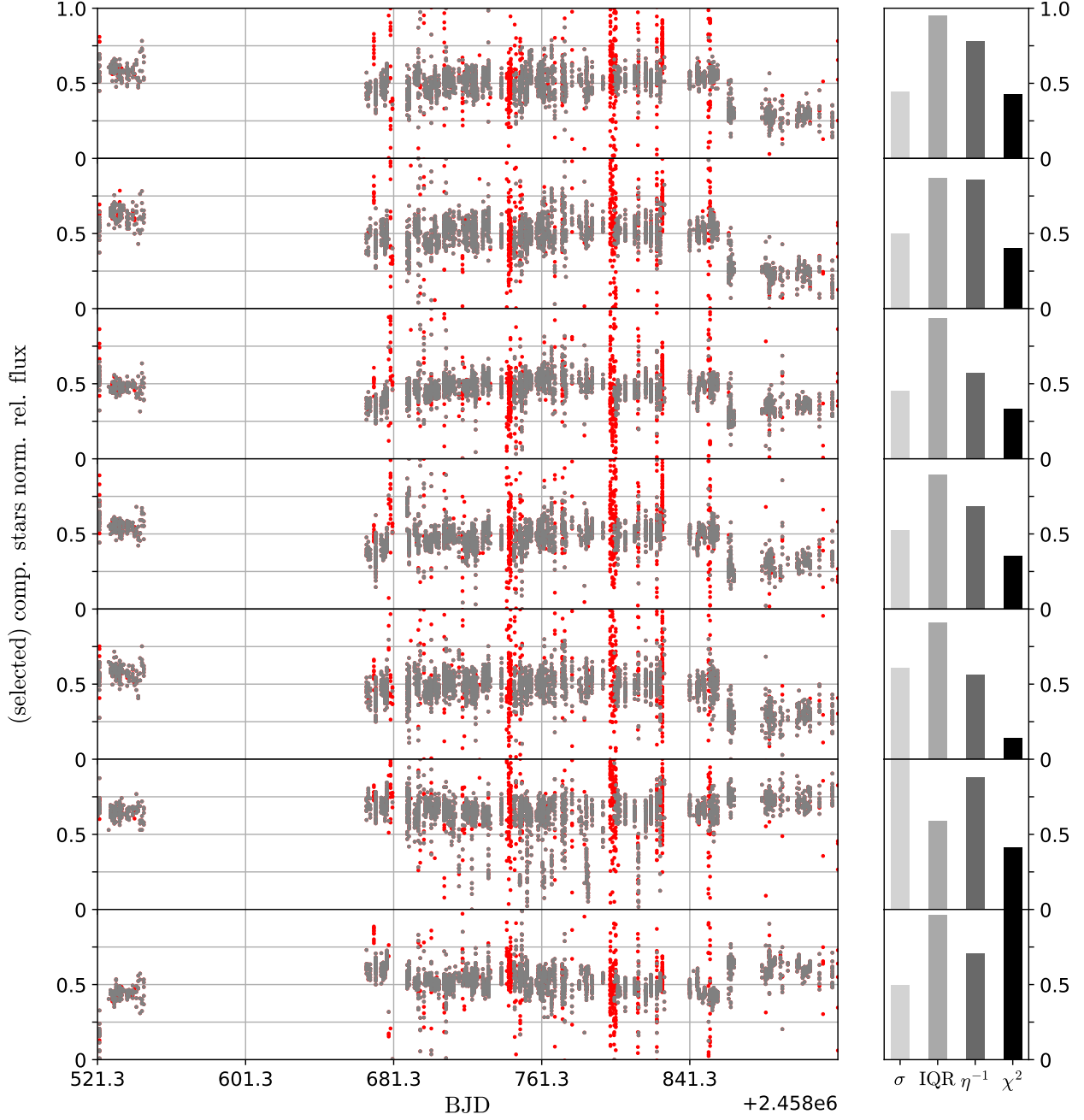


Figure 7: Relative light curves (TZ Ari) from the 7 discarded comparison stars through optimized star selection based on variability index performance (using indices σ and χ^2). Relative fluxes have been individually normalized so that each light curve fits within the interval $[0, 1]$. The second column contains index bar charts associated to each of the 7 discarded stars; each index value has been normalized by the corresponding maximum across all comparison stars.

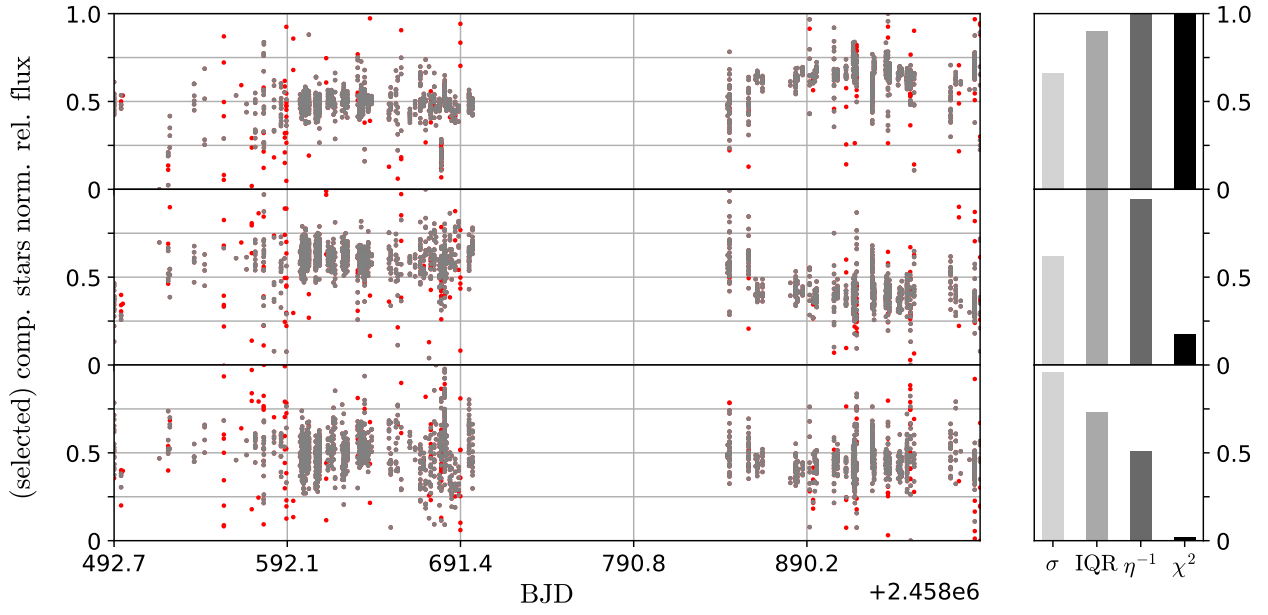
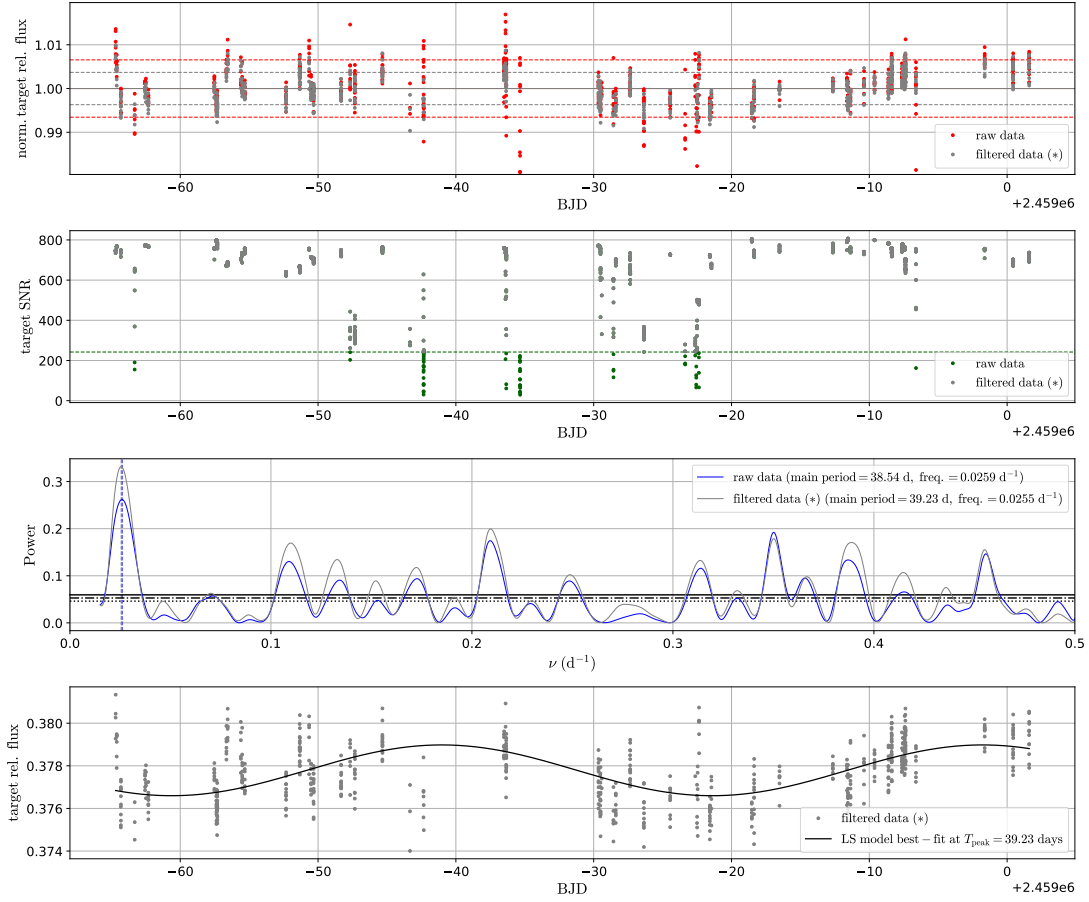
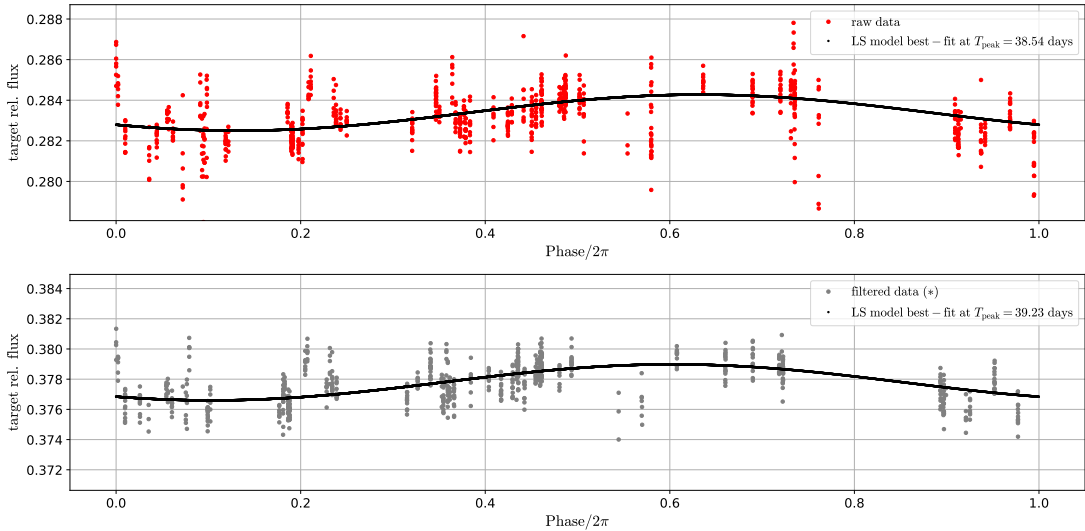


Figure 8: Relative light curves (GJ 555) from the 3 discarded comparison stars through optimized star selection based on variability index performance (using indices **IQR** and η^{-1}). Relative fluxes have been individually normalized so that each light curve fits within the interval $[0, 1]$. The second column contains index bar charts associated to each of the 3 discarded stars; each index value has been normalized by the corresponding maximum across all comparison stars.

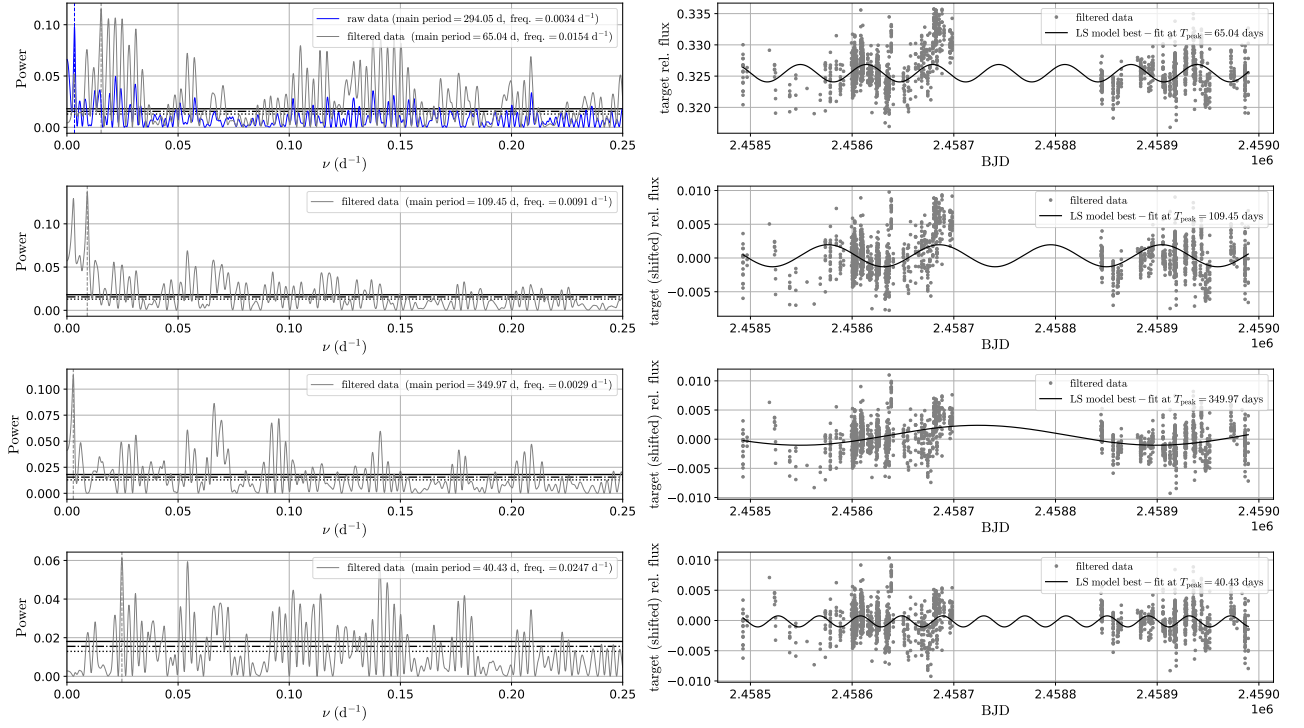


(a) From top to bottom. (1) Normalized target relative flux (with respect to the mean) vs. BJD: raw (red) and filtered data are presented together; dashed lines represent (normalized) standard deviations above and below mean relative flux values for both raw and filtered data. (2) Target signal-to-noise ratio vs. BJD: raw (green) and filtered (gray) data are presented together, and separated by a dashed (green) line at $\max_i \{SNR_i\} p_T$ (numerical clip). (3) LS periodogram for raw (blue) and filtered (gray) data; solid, dash-dotted and dotted horizontal lines represent power thresholds (filtered data) for 0.1%, 0.01% and 0.001% FAP values. (4) LS model best-fit at peak frequency ν_f (for filtered data), together with filtered (unnormalized) target relative fluxes.

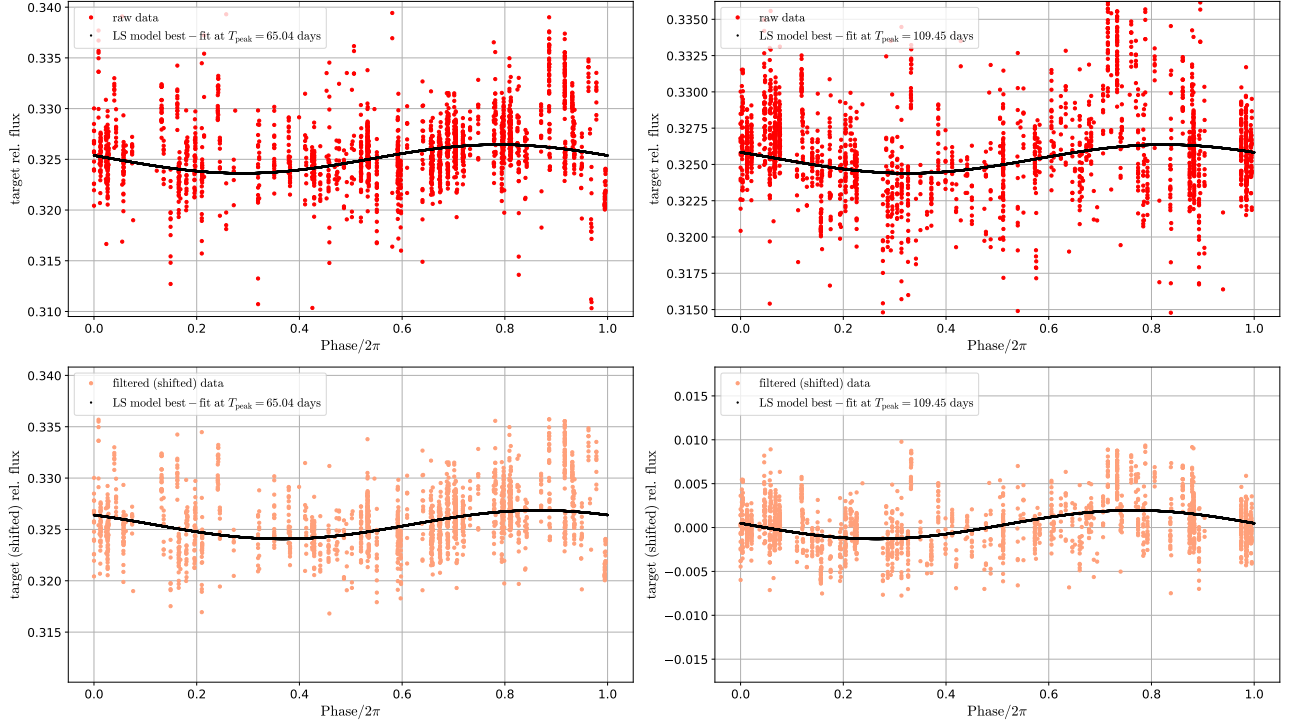


(b) Phase-folding of raw and filtered relative fluxes with best-model fit from LS periodogram results.

Figure 9: (a) Raw and filtered data comparison for TOI-1266 photometry, with LS periodogram results and best-model fit for filtered data; in this case, filtered data (*) indicates that only comparison stars from the optimal RSD ensemble were considered to compute (updated) relative fluxes (used information 85.66%). (b) Phase-folding of raw and filtered relative fluxes with best-model fit from LS periodogram results.



(a) GJ555 successive LS periodograms from iterative prewhitening, together with LS model best-fits for filtered, shifted fluxes at peak frequencies. In the first row (initial step), both raw (blue) and (gray) filtered data LS periodograms are represented. Solid, dash-dotted and dotted horizontal lines indicate filtered data power thresholds at 10%, 1% and 0.01% False Alarm Probabilities, respectively.



(b) Phase-folded raw (top) and filtered, shifted (bottom) target relative fluxes, together with LS model best-fits at peak frequencies, corresponding to iterative prewhitening stages 0 (left, $T_{\text{peak}} = 65.04$ days) and 1 (right, $T_{\text{peak}} = 109.45$ days).

Figure 10: (a) LS periodograms and best-fits from iterative prewhitening applied to GJ 555 (used information 89.46%). (b) Phase-folding results from initial raw and filtered data and first prewhitening iteration.

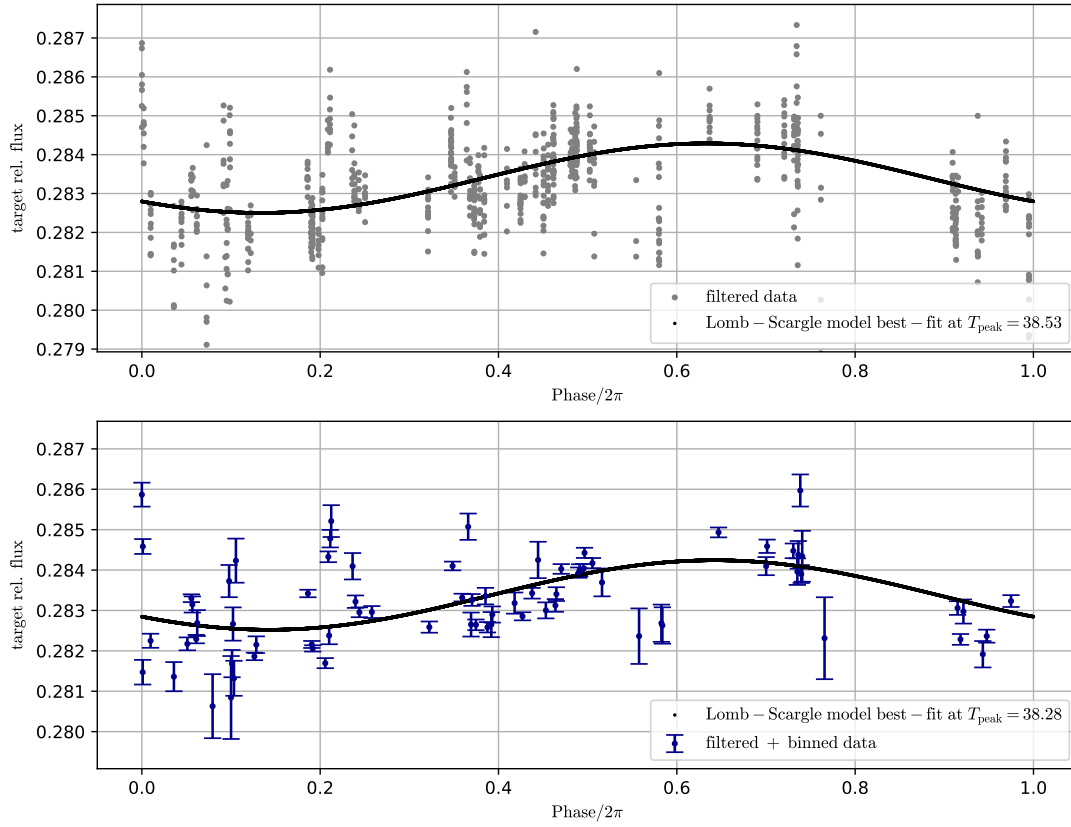


Figure 11: Comparison between phase-folded TOI-1266 filtered relative ($\alpha_T = 3.5, p_T = 0.15, \alpha_C = 4.0, p_C = 0.0$) fluxes with 1h-binning (bottom, blue) and without it (top, gray), together with LS model best-fits at the corresponding LS peak frequencies.

Table 5: Numerical values (with 6 decimal places) found for peak frequencies, ν_{peak} , and their corresponding FWHM (full width at half maximum) for each of the studied objects and the presented filtering/optimization techniques. PW i stands for prewhitening iteration i .

<i>Object</i>	Raw ν_{peak} (d^{-1})	FWHM	Filtered ν_{peak} (d^{-1})	FWHM	Filtered (*) ν_{peak} (d^{-1})	FWHM
WOLF 1069	0.005051	0.002029	0.005231	0.002174	0.005096	0.002247
TOI-1266	0.025946	0.012994	0.025493	0.011864	0.025494	0.011428
TZ ARI PW 0	0.003151	0.002631	0.003101	0.003070	0.002810	0.002586
TZ ARI PW 1			0.510918	0.004464	0.511018	0.005263
TZ ARI PW 2	\emptyset	\emptyset	0.041002	0.005357	0.019827	0.005263
TZ ARI PW 3			0.049353	0.002679	0.152681	0.002632
GJ 555 PW 0	0.003401	0.001111	0.115295	0.001852	0.015587	0.001481
GJ 555 PW 1			0.009137	0.002222	0.009005	0.001852
GJ 555 PW 2	\emptyset	\emptyset	0.002857	0.001852	0.066145	0.002593
GJ 555 PW 3			0.02474	0.001481	0.225710	0.001852

Table 6: Minimum RSD values, σ_{\min} , computed for the different objects under study, using the methods described in 3.1.1; these values are used in Figure 2.

# of discarded stars	σ_{\min}			
	WOLF 1069	TOI-1266	TZ ARI	GJ 555
1	0.01114083	0.00402744	0.01123498	0.00836340
2	0.01046982	0.00398526	0.01123102	0.00817861
3	0.01019530	0.00397937	0.01152197	0.00806548
4	0.00998004	0.00395742	0.01386154	0.00801013
5	0.00988470	0.00394732	0.02005772	0.00793891
6	0.00979557	0.00394842	0.03055165	0.00787460
7	0.00968783	0.00395710	0.04451001	0.00782547
8	0.00958116	0.00398187	0.06330981	0.00778703
9	0.00947923	0.00411618	0.08626858	0.00773306
10	0.00940037	0.00425339	0.11596963	0.00774017
11	0.00940452	0.00428346	0.15201317	0.00776987
12	0.00949379	0.00502312	0.21765441	0.00781633
13	0.00971858	—	—	0.00818897
14	0.01002703	—	—	0.00906581

B Sources of noise in photometric measurements

There are 3 fundamental sources of noise in photometric measurements, according to [11]: (1) Poisson noise in target data and comparison stars, (2) calibration and read noise, and (3) scintillation noise. Other sources of error that are generally less important are: flux variability in comparison stars or atmospheric conditions. For completion, we will give a brief description for each source mentioned above.

Poisson noise in target and comparison stars. Let us consider a source of constant brightness. As light comes in discrete packets, in any given period, there is a chance of receiving more or less photons than the mean photon number corresponding to that source. This is, by definition, *shot noise*, and will be present in both the target and comparison stars fluxes. It obeys a *Poisson distribution*. The reason for that is fairly simple if we take into account the quantum mechanical description of photons. In the classical realm, a constant intensity source of light can be described by a spatially and temporally coherent e.m. wave, *i.e.* $E(x, t) = E_0 \sin(kx - \omega t + \varphi)$. The quantum mechanical analogue for such source is represented by the so called coherent states:

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \quad (9)$$

with $\alpha \in \mathbb{C}$, and $|n\rangle$ the n -th Fock state. The probability of finding n photons, P_n , can then be computed as $|\langle n|\alpha\rangle|^2$ and yields:

$$P_n = \frac{|\alpha|^{2n}}{n!} e^{-|\alpha|^2} \quad (10)$$

The mean number of photons, \mathcal{N} , for $|\alpha\rangle$ can then be computed as follows:

$$\mathcal{N} = \sum_{n=0}^{\infty} n P_n = \sum_{n=1}^{\infty} n \frac{|\alpha|^{2n}}{n!} e^{-|\alpha|^2} = e^{-|\alpha|^2} |\alpha|^2 \sum_{n=1}^{\infty} \frac{(|\alpha|^2)^{n-1}}{(n-1)!} = |\alpha|^2 \quad (11)$$

In terms of \mathcal{N} , P_n reads:

$$P_n = \frac{\mathcal{N}^n}{n!} e^{-\mathcal{N}} \quad (12)$$

This is precisely the form of a Poisson distribution with *rate* \mathcal{N} , $\text{Pois}(\mathcal{N})$, and justifies the fact that shot noise follows this kind of distribution.

Taking now into account that a measure of signal-to-noise ratio (SNR) for a random variable following a certain distribution can be computed as the quotient μ/σ , and noting that $\text{Pois}(\mathcal{N})$ has variance \mathcal{N} , it is clear that the SNR for target and comparison stars scales by $\text{SNR} \propto \mathcal{N}/\sqrt{\mathcal{N}} = \sqrt{\mathcal{N}}$. It is then clear than one way to reduce target noise is by increasing signal (that is, increasing \mathcal{N}), which can be accomplished increasing exposure time. On the other hand, shot noise in comparison star fluxes may be addressed differently, since adding more of these stars will result in an overall decrease of this source of error across all stars of the ensemble, by a factor $\propto 1/\sqrt{\#\text{comp. stars}}$. However, proceeding in this way may incorporate bad pixels from the camera and/or variable comparison stars.

Calibration and read noise. Calibration noise is associated to readout, dark current and field illumination effects, that can be addressed by obtaining the bias, dark and flat-field frames. These are calibration frames that help increasing the signal-to-noise ratio of the received signal. The bias frame aims to remove the base level of readout of the CCD in the absence of light; the dark frame allows for reduction of thermal noise induced by long exposures; finally, the flat-field frame helps adding a correction to each pixel of the camera so that they all give the same value when exposed to the same quantity of light for a given optical path.

Read noise is primarily due to electronics, cables, e.m. interference in the observatory and the proper functioning of the CCD camera. It is the cumulative effect on the signal of errors produced during (and between) different operating stages of the CCD, from the conversion of electrons into voltage, to the actual digitalization of the signal into numbers representing the measured fluxes.

On the other hand, quantum efficiency (the ratio between electrons collected at each cell and the number of photons hitting the corresponding pixel surface, QE) or charge transfer efficiency (the fraction of electrons that are successfully moved from one pixel to another during readout, CTE) help quantifying sources of noise in this context. Finally, it is worth mentioning that exposing a CCD to a light source will produce a number of thermal electrons leaking into its pixels, which can become an important component of read noise. This *dark current* is known to increase linearly with exposure time, and is also sensitive to temperature (which justifies why CCDs should not be uncooled when operating). However, the advance of technology has allowed for an important decrease of read noise in modern CCDs and is currently not a major concern (see [23, p. 36]).

Scintillation noise. Scintillation is the change of apparent brightness of an astronomical object due to changes in the light path caused by atmospheric turbulence.

The uncertainty introduced by scintillation in flux measurements can be modelled analytically by the following formula ([11, p. 174]):

$$\sigma_{\text{sc}} = 0.004 D^{-\frac{2}{3}} X^{\frac{7}{4}} e^{-\frac{h}{H}} (2t_{\text{exp}})^{-\frac{1}{2}} \quad (13)$$

where D is diameter of the telescope, X the air mass at which we are operating, h the altitude of the observatory, H the scale height of the turbulence (which can be set ~ 8000 m), and t_{exp} the exposure time of the camera. The air mass is an important limiting factor regarding the optimization of target and comparison stars signal-to-noise ratios. It is defined as the integral of air density along the light ray; thus, to a certain extent, it quantifies the amount of air we are looking through to perform measurements on the object of interest. The simplest model for computation of air mass values assumes a plane-parallel atmosphere, and defines air mass X for an object with zenith angle¹⁹ z as $X = 1/\sin(z)$ (note that minimum and maximum air mass values are achieved at $z = 0$ and $z = \pi/2$, as expected).

For long-period observations, reduction of scintillation noise becomes more difficult, since we will have measurements over a range of air masses (and highest σ_{sc} at maximum air mass). From equation (13) we see that σ_{sc} increases with signal, since higher exposure times allow for more photons hitting the camera between recordings; on the other hand, increasing the diameter of the telescope will help reducing this source of noise. Another way of minimizing σ_{sc} is binning the data post calibration, which would yield a higher cadence as a drawback.

Other sources of noise. As mentioned before, variability of comparison stars when performing differential photometry and atmospheric conditions also contribute to lowering the signal-to-noise ratio of the object of interest. The case of variable comparison stars will be discussed more extensively in forthcoming sections. Clouds, sky glow, aurorae, among other events, can rapidly become catastrophic for our measurements. Sky noise, —that is, variation of sky signal due to changes in the emissivity of the atmosphere— can also affect aperture photometry measurements, when the local average sky counts are subtracted from the source count. It may be reduced by careful selection of photometric filter, but is nonetheless the main limiting factor in observing faint objects (see [11, p. 175]).

¹⁹Angular distance between the object and the zenith.

C Photometry extraction with AstroImageJ

AstroImageJ is a graphical Java-based user interface with several image processing tools specially aimed at the treatment and extraction of differential photometric data from .FITS image files, among other common formats. It provides research level analysis and allows for ultra-precise light curves from the observed objects. For an extensive survey of its capabilities the reader is referred to [5]. In this section, we will summarize the aspects of **AstroImageJ** photometry extraction that are central to subsequent discussion.

AstroImageJ allows performing aperture photometry of single astronomical objects. .FITS files can be visualized in real time, and placing of apertures can be done by sight with a mouse click near the star of interest, that will (if desired) place the center of the aperture at the centroid of the selected star. This option makes determination of star centers systematic and efficient. Both the radius of the aperture and the width of the annulus around its center (from which the local sky counts will be computed) can be set to equal a specific number of pixels (conversion between pixels and arcseconds is assumed to be known *a priori*). An optimal choice of these numbers is always desired: aperture radius should be close to the FWHM of the PSF of the studied object, and the annulus region should be close enough to the source, but contain as less from it or other sources as possible.

AstroImageJ can keep track of the selected object through the different images, and measure/compute, given an initial aperture, several variables: centroid position in .FITS coordinates, aperture net integrated counts, aperture peak pixel count, aperture mean pixel count, world coordinates (RA,DEC), etc. It can also take into account CCD parameters as gain, readout noise or dark current, which if included correctly, can contribute to noise reduction of measured fluxes.

Photometry extraction with **AstroImageJ** generates a file which contains, for all the selected images, the variables mentioned before (together with their associated errors if required), corresponding to the object of interest. Moreover, through loading specific plotting configurations, it allows to visualize in a clear way several of such variables on the same plot. This makes the task of studying the recorded data very efficient, which helps tuning important parameters, such as aperture radius, towards optimal values, very fluid.

One of the main advantages of **AstroImageJ** is its capability to perform multi-aperture differential photometry. This allows selecting visually both target and comparison stars and computing the variables of interest. Relative target and comparison stars fluxes are computed as in equation (1). Uncertainties associated to such measurements are also provided if requested (explicit expressions for errors and all other quantities produced by the program can be found in [5]). Among the measurements provided by **AstroImageJ**, the following are the ones that we will be using the most throughout this work, and for subsequent analysis: target and comparison stars relative fluxes (`rel_flux_T1`, `rel_flux_Ci`, with $i \geq 2$ a natural number) and their corresponding errors (`rel_flux_err_T1`, `rel_flux_err_Ci`), target and comparison stars relative fluxes signal-to-noise ratios (`rel_flux_SNR_T1`, `rel_flux_SNR_Ci`) and, finally, target and comparison stars aperture peak pixel count (`Peak_T1`, `Peak_Ci`).

D Basic principles of periodogram analysis

D.1 Fourier analysis and power spectrum

The Fourier transform of a continuous signal in time, $f(t)$, is defined as the integral

$$\hat{f}(\nu) = \int_{-\infty}^{\infty} f(t) e^{-2\pi i \nu t} dt$$

Under certain conditions²⁰, given the Fourier transform of a signal, $\hat{f}(\nu)$, we can recover the original signal, by means of the *inverse Fourier transform*:

$$f(t) = \int_{-\infty}^{\infty} \hat{f}(\nu) e^{2\pi i \nu t} d\nu$$

Sometimes we will denote the Fourier transform of a function f as $\mathcal{F}(f)$.

The most important property of the Fourier transform regarding period search is the fact that it transforms exponentials of the form $e^{2\pi i \nu_0 t}$ into Dirac deltas $\delta(\nu - \nu_0)$, i.e. $\mathcal{F}(e^{2\pi i \nu_0 t})(\nu) = \delta(\nu - \nu_0)$. An immediate consequence of this (using the Euler formula) is that sines and cosines of a certain frequency can also be expressed as a simple (finite) linear combination of deltas. More precisely:

$$\mathcal{F}\{\cos(2\pi \nu_0 t)\}(\nu) = \frac{1}{2}[\delta(\nu - \nu_0) + \delta(\nu + \nu_0)], \quad \mathcal{F}\{\sin(2\pi \nu_0 t)\}(\nu) = \frac{1}{2i}[\delta(\nu - \nu_0) - \delta(\nu + \nu_0)] \quad (14)$$

Another important feature of the Fourier transform is that it translates time shifts in the signal, $t \rightarrow t - t_0$, to multiplication by a phase, $e^{-2\pi i \nu t_0}$, in the frequency domain. Thus, applying \mathcal{F} to a real valued function, results, in general, into a complex valued function. To get rid of complex values and, in particular, the phase imparted by the choice of temporal baseline, we can take the amplitude of the Fourier transform, $|\mathcal{F}(f)|^2$. This is what is known as the *power spectral density* of f , or simply *power spectrum*. It is sometimes denoted \mathcal{P}_f , and, by the observations made, it quantifies the contribution of each frequency ν to the total signal.

The *convolution* of two functions f, g is defined as the integral:

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau$$

and is well behaved under the Fourier transform, by virtue of the so called *convolution theorem*:

$$\mathcal{F}(f * g)(\nu) = \mathcal{F}(f)(\nu) \cdot \mathcal{F}(g)(\nu), \quad \mathcal{F}(f \cdot g)(\nu) = [\mathcal{F}(f) * \mathcal{F}(g)](\nu) \quad (15)$$

which tells us that convolutions transform to point-wise products, and viceversa. However, real world measurements are given over a finite span of time, and at some finite sampling rate. The observed signal, $g_{\text{obs}}(t)$, is then generally described by a point-wise product of two functions: the true underlying continuous signal, $g(t)$, and a *window function*, $W(t)$, that characterizes the way in which the signal is being recorded. The Fourier transform of the measured data is then, by equation (15), a convolution: $\mathcal{F}(g_{\text{obs}}) = \mathcal{F}(g) * \mathcal{F}(W)$.

Rectangular- and Dirac comb-shaped windows are two important examples of window functions. The effect of the former, when applied to a given signal, is to replace, approximately, each frequency peak by a sinc function (with the standard sinc function being $\sin(\pi x)/(\pi x)$). This is because the Fourier transform of a rectangular window, $W_{\square}(t; A, T) = A \mathbb{1}_{[-T/2, T/2]}$, is given by $A \sin(\pi T \nu)/(\pi \nu) = AT \text{sinc}(\nu T)$. Hence, a wider observing window leads to proportionally less spread²¹ in the Fourier transform of g_{obs} (this is because of the so called *uncertainty principle*). A Dirac comb window function, on the other hand, is an infinite sequence of Dirac delta functions placed at even intervals of a certain size, T : $W_{\text{Dirac}}(t; A, T) = A \sum_{n \in \mathbb{Z}} \delta(t - nT)$.

²⁰For example, if both f and \hat{f} are continuous and integrable.

²¹Higher values of T translate in more narrower sinc functions.

Its Fourier transform is another Dirac comb: $\mathcal{F}\{W_{\text{Dirac}}(t; A, T)\}(\nu) = W_{\text{Dirac}}(\nu; AT^{-1}, T^{-1})(\nu)$. The effect of such an observing window is to create a long sequence of aliases (copies) of the underlying transform with a spacing of $1/T$, which makes evaluation of the observed transform in the range $0 \leq \nu < 1/T$ sufficient to capture all the available frequency information.

Increasing time between observations will decrease the spacing between peaks in the frequency domain. Let $g(t)$ be a Gaussian signal and $W_{\text{Dirac}}(t; A, T)$ a Dirac comb window function. Since the Fourier transform of a Gaussian is still a Gaussian, for large enough T , it is clear that \hat{g} will not *fit* between two deltas of the transformed Dirac comb. As a result, the Fourier transform of g_{obs} will be a mixed version of the different portions of the original signal, and its true transform could not be recovered entirely. This suggests there is a limiting sampling frequency above which there will always be some loss of periodic information from the original signal. This is known as the *Nyquist sampling limit*: a band-limited signal, with null Fourier transform outside $[-B, B]$ requires a sampling rate of at least $\nu_{\text{Ny}} = 2B$ to fully recover its frequency content.

Combining W_{\square} and W_{Dirac} results in a window function that perfectly represents a typical real world measurement. For a continuous signal $g(t)$, if the sampling frequency is $1/\Delta t$ (i.e. deltas in the Dirac comb window are regularly spaced with time interval Δt), and the number of measurements allowed by the rectangular window is N , the Fourier transform of $g_{\text{obs}}(t) = g(t)W_{\text{Dirac}}(t; 1, \Delta t)W_{\square}(t; 1, N\Delta t)$ reads:

$$\hat{g}_{\text{obs}}(\nu) = \sum_{n=0}^N g(n\Delta t)e^{-2\pi i\nu n\Delta t} \quad (16)$$

which follows directly from the definition of Fourier transform and the given expressions for both W_{Dirac} and W_{\square} . By previous arguments, since the distance between consecutive deltas in W_{Dirac} is Δt , the only relevant frequency range is $0 \leq \nu < 1/\Delta t$. Thus, we may define N evenly-spaced frequencies, with $\Delta\nu = 1/(N\Delta t)$, covering this range, to obtain the sampled Fourier transform of the observed signal:

$$\hat{g}_k = \sum_{n=0}^N g_n e^{-2\pi i k n / N} \quad (17)$$

where $\hat{g}_k = \hat{g}(k\Delta\nu)$ and $g_n = g(n\Delta t)$. This is the usual form of the *discrete Fourier transform*.

As discussed before, the effect of a rectangular window of size T over the Fourier transform of a certain signal, was to replace each frequency peak by a sinc function of width $\sim 1/T$. Hence, the window function $W_{\square}(t; 1, N\Delta t)$ replaces each peak in \hat{g} by a $1/(N\Delta t)$ -width sinc function. This implies that frequency values within $1/(N\Delta t)$ of each other in \hat{g}_{obs} will not be independent, which forces a spacing between transform samples of at least $1/(N\Delta t)$. Taking this into account, the choice of N regularly spaced frequencies to compute the sampled Fourier transform is clearly justified; subsequently, this indicates that the frequency spacing of the discrete Fourier transform is *optimal* in terms of both the Nyquist limit and the effect of a finite rectangular window.

D.2 The Lomb-Scargle periodogram

The classical or Schuster periodogram (first proposed in [29]) is essentially the power spectrum of the Fourier transform of the observed signal (see equation (16)):

$$\mathcal{P}_g^S(\nu) = \frac{1}{N} \left| \sum_{n=1}^N g_n e^{-2\pi i \nu t_n} \right|^2 \quad (18)$$

\mathcal{P}_g^S is an estimator of the real power spectrum, \mathcal{P}_g , of the signal, and, in the uniform sampling case, captures all the relevant frequency information present in the data (if a correct choice for N and the spacing of the frequency grid is made).

For non-uniform sampling, the task of recovering the Fourier transform of the original signal is even harder. In this case, the window function can be thought as a finite number of unevenly spaced Dirac deltas. As a consequence, its Fourier transform will be much more noisy and difficult to compute. When observations are not made instantaneously, as in the case of photometric measurements, another frequency limit arises in the context of non-uniform sampling (see [16]). In this case, the recorded photons are integrated over a specific period of time, δt , which allows for the observation process to be viewed as a convolution of the real signal with a δt -width rectangular window. The observed transform will then have width $\sim 1/\delta t$, which limits the maximum frequency to $\sim 1/(2\delta t)$ (if frequency information is not to be lost) —this is known as the *windowing limit*.

Photometric measurements are usually non-uniformly sampled, and can even be performed throughout different epochs, with important time intervals where no data is collected. Nonetheless, the necessity for period search algorithms is clear in the context of optical astronomy, and a generalization of the classical periodogram capable of fully recovering the frequency information —even in the case of non-uniform sampling— is at request. This is precisely what the Lomb-Scargle periodogram does. In 1982, Scargle considered a generalized form of the expression in (18), depending on a number of arbitrary functions, and showed that requiring the resulting periodogram to fulfill 3 properties, namely:

- (1) It should reduce to the classical form in the case of equally-spaced observations.
- (2) Its statistics should be analytically computable.
- (3) It must be insensitive to global time-shifts.

the aforementioned functions were completely determined. The obtained expression was:

$$\mathcal{P}_g^{LS}(\nu) = \frac{1}{2} \left\{ \frac{[\sum_n g_n \cos(2\pi\nu[t_n - \tau])]^2}{\sum_n \cos^2(2\pi\nu[t_n - \tau])} + \frac{[\sum_n g_n \sin(2\pi\nu[t_n - \tau])]^2}{\sum_n \sin^2(2\pi\nu[t_n - \tau])} \right\}, \quad (19)$$

$$\text{with } \tau = \frac{1}{4\pi\nu} \arctan \left[\frac{\sum_n \sin(4\pi\nu t_n)}{\sum_n \cos(4\pi\nu t_n)} \right]$$

The similarities between the Lomb-Scargle periodogram and the classical or Schuster periodogram grant that we can interpret the results given by the former in accordance to our previous discussion (window function effects, Nyquist and windowing limits, etc.), at least qualitatively. On the other hand, an important caveat of this periodogram is that its statistical properties are well behaved only when the observations have uncorrelated white noise; for data with more complicated noise characteristics, a different treatment is advised.

As stated, the Lomb-Scargle periodogram (LS periodogram) can be viewed as a least-squares method. To see why, we need to introduce the *least-squares periodogram* itself. This periodogram is constructed as follows: at each candidate frequency, ν , a sinusoidal model, and fit to the data in the standard least-squares sense, by minimizing the corresponding χ^2 statistic:

$$\tilde{\chi}_1^2(\nu) = \sum_n [y_n - y(t_n; \nu)]^2 \text{ or } \tilde{\chi}_2^2(\nu) = \sum_n \left[\frac{y_n - y(t_n; \nu)}{\sigma_n} \right]^2 \quad (20)$$

depending on if (Gaussian) measurement errors, σ_n , are included ($\tilde{\chi}_2^2$) or not ($\tilde{\chi}_1^2$). The optimized parameters, best-fit model and minimum χ^2 statistic value at frequency ν are denoted: \hat{A}_ν , $\hat{\phi}_\nu$, $\hat{y}(t; \nu)$ and $\hat{\chi}_i^2(\nu)$, respectively. Scargle showed, also in 1982, that the LS periodogram presented in equation (19) is equivalent to:

$$\mathcal{P}^{LS}(\nu) = \frac{1}{2} [\hat{\chi}_0^2 - \hat{\chi}^2(\nu)] \quad (21)$$

where $\hat{\chi}_0^2 \equiv \hat{\chi}^2(0)$, i.e. minimum χ^2 statistic corresponding to a best-fit of a constant signal to our data. The fundamental idea behind the equality of the LS and least-squares periodograms is that the former implicitly assumes a sinusoidal model for the data. Note that, the spectral power estimate given by $\hat{\chi}_2^2$, together with equation (21), for measurements with Gaussian error, will be closer to its real value, due to the noisy nature of the observations. The statistical properties of the periodogram in this case are virtually the same as in the case where errors are not included (see [40]).

D.3 Periodogram failure modes

Failure modes are an important cause for mischaracterizing the periodic behaviour of a real signal, when analyzing its periodogram (Lomb-Scargle, least-squares, etc.). The differences between the observed and the true signal, due to the measuring process, can cause the principal frequency of the signal to not coincide with the frequency corresponding to the largest peak in the periodogram. This problem is explored in [38], where Figure 12 is obtained, from 1000 simulated periodic (irregular) light curves. The results show that more than 50% of the points do not line on the line $y = x$, i.e. the period of the largest peak is the true period of the signal. The different patterns where points outside this line lie are known as *failure modes of the periodogram*, and are also present in real observations (see [21]). These modes can be classified in 3 broad categories:

Type I failure modes. For observations at night, the corresponding window functions have a strong diurnal component. This causes main frequencies of the signal, ν_p , to be partially aliased at $\nu_p + n\delta\nu$, where $\delta\nu$ is the major frequency component in the window (e.g.: 1 cycle / day). 36% of the data presented in Figure 12 is unsuccessfully characterized along failure modes of this kind.

Type II failure modes. The largest peak of an LS periodogram can be given at a multiple of the fundamental frequency of the signal. This may occur for periodic signals that are not strictly sinusoidal and thus present power spectrum peaks at higher (or lower) harmonics. This results in type II failure modes.

For positive multiples of the fundamental frequency, ν_0 , a partial aliasing is devised at frequencies of the form $m\nu_0 + n\delta\nu$, with $m \geq 0$.

The (even) symmetry of the LS periodogram allows for negative peak aliases to be reflected into the positive-frequency range, leading to possible largest peaks at frequency values of $|m\nu_0 + n\delta\nu|$, with n, m arbitrary integers.

Irregular failure modes. When an LS periodogram fails to characterize the main frequency of a variable (periodic) object, but the resulting point in a *Periodogram Peak vs. True Period* diagram can not be interpreted as part of a type I or a type II failure mode pattern, we say the corresponding failure mode is irregular.

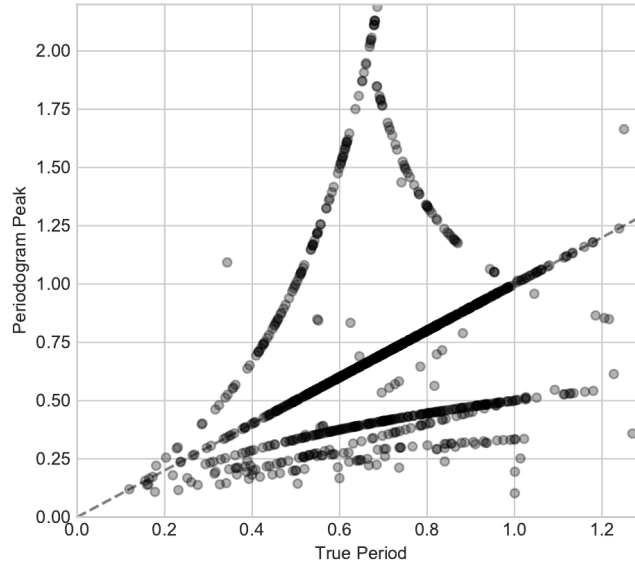


Figure 12: Comparison of the true period and peak LS period for 1000 simulated periodic light curves (with 60/180 irregular observations each).

Thus, it is important to take into account these effects when analysing periodogram results. If the observed peak is at frequency ν_{peak} , a general recommendation is to check for peaks at frequencies ν_{peak}/m with $m \geq 2$ an integer (it should suffice checking for peaks with $m = 2, 3$); if a significant peak is found, then ν_{peak} is probably an order m harmonic of the true frequency of the signal, ν_0 . On the other hand, if the observation window has strong power at known $\delta\nu$, one may also check for peaks at $|\nu_{\text{peak}} \pm n\delta\nu|$ for at least $n = 1, 2$; if such aliases are found, one of them may correspond also to the true frequency, ν_0 —fitting then a more complicated model should help clarifying whether or not the obtained frequency accurately describes the periodic behaviour of the signal.

E Filtering procedures. Binning and implementation details.

We begin by detailing some aspects of binning in the context of differential photometry, and continue by explaining the main implementation features of the 3 studied filtering procedures: σ -clipping, numerical clipping and binning itself.

E.1 Binning

Photon counting signal-to-noise ratio is proportional to $\sqrt{\mathcal{N}}$, with \mathcal{N} the number of photons. Higher exposure times imply then higher SNRs, but can also lower the resolution in the time-domain causing great losses of information regarding the variability of the signal. A good compromise between exposure time and cadence is therefore always requested. After data recording, the effect of increasing exposure times can be achieved through time-domain binning of the obtained fluxes.

Flux values associated to each bin can be estimated taking the mean (or median) of flux measurements within each of them. Since the standard error of the mean is $\propto 1/\sqrt{n}$ where n is the number of measurements considered, this has the desired effect: it reduces shot noise, resulting in an increase of flux SNR.

For flux measurements m_i taken at times t_i , $i = 1, \dots, n$, we can consider B equally spaced bins of size $\delta t = (t_{\max} - t_{\min})/B$, with t_{\min} and t_{\max} the minimum and maximum values of $\{t_i\}_{i=1}^n$. Bin j would then be defined by the interval $[t_{\min} + (j-1)\delta t, t_{\min} + j\delta t)$ for $j = 1, \dots, B-1$, the last bin, $[t_{\max} - \delta t, t_{\max}]$ being closed. Let $\{m_i^j\}_{i=1}^{n_j}$ be the fluxes within bin j , then, the error, $\bar{\sigma}_j$, associated to $\bar{m}_j = n_j^{-1} \sum_{i=1}^{n_j} m_i^j$ can be computed as:

$$\bar{\sigma}_j = \left[\frac{n_j^{-2} \sum_{i=1}^{n_j} (\sigma_i^j)^2 + (n_j - 1)^{-1} \sum_{i=1}^{n_j} (m_i^j - \bar{m}_j)^2}{n_j} \right]^{1/2} \quad (22)$$

where the first term in the numerator comes from error propagation using the expression for \bar{m}_j and the second is the classical standard deviation of fluxes within the bin. This last formula can be easily generalized for the case of weighted means and errors, using expressions in (2).

E.2 Implementation details

σ -clipping. Given a set of data points $\{y_i\}_{i=1}^n$, factors α_- , α_+ , and a maximum iteration number s , we can σ -clip the given data, removing, at each iteration, points y_j falling outside the interval $(m - \alpha_- \sigma, m + \alpha_+ \sigma)$, where σ is the usual estimate for the standard deviation and m can be set to be the mean, \bar{y} , or the median, $\text{median}(y_i)$, of y_i values considered in the current iteration. We will normally have $\alpha_- = \alpha_+ \equiv \alpha$.

σ -clipping can be used to remove outliers from both target and comparison stars relative fluxes; α_{\pm} values and center functions (mean or median) can be different in each case. For each of the stars under consideration, a (Boolean) mask is created, marking each relative flux value to be clipped with a **True** mask entry (valid values are associated to **False** mask entries). When all masks are computed, their logical sum provides a final mask that accounts for all σ -clipping applied to the different relative light curves.

Numerical clipping. Given a value $p \in (0, 1)$, the proposed implementation allows performing a numerical clipping of SNR values associated to relative fluxes as described in 2.1, generating a mask that is constructed as in the σ -clipping procedure. Again, SNR numerical clipping can be performed for both target and comparison stars, and p -values associated to each case can be different. Target and/or comparison stars masks obtained in each clipping are finally added to generate the corresponding total mask.

Binning. Given a series of relative fluxes, with their corresponding errors and times, $\{(t_i, m_i, \sigma_i)\}_{i=1}^n$, and a bin size, δt , (relative) flux measurements are binned, as described in 2.1. For each bin, a triple $(\bar{t}_j, \bar{m}_j, \bar{\sigma}_j)$ is computed, so that, if $\{(m_i^j, \sigma_i^j)\}_{i=1}^{n_j}$ are fluxes (and errors) within bin j : $\bar{t}_j = t_{\min} + \delta t(j-1/2)$ (i.e. the midpoint of its time range); \bar{m}_j is its weighted mean, following equation (2); and $\bar{\sigma}_j$ is the weighted version of equation (22):

$$\bar{\sigma}_j = \left[\frac{1/\sum_{i=1}^{n_j} (\sigma_i^j)^{-2} + (\sigma_w^j)^2}{n_j} \right]^{1/2} \quad (23)$$

with σ_w^j the weighted standard deviation of relative fluxes in bin j (see equation (2)).

Normally, only target relative fluxes are binned, as our priority is to improve its signal-to-noise ratio, as well as to minimize its dispersion. However, if star selection is executed after binning target photometry measurements, it is necessary to also apply this filtering technique to all comparison stars from the initial **AstroImageJ** ensemble.

Besides σ -clipping, numerical clipping and binning, two more filtering procedures were implemented to ensure an optimal use of photometry data. In first place, due to possible spurious entries in the generated files from **AstroImageJ**, we only keep relative flux values in the range $[0, 1]$; for the same reason, relative fluxes with SNR values outside a specified interval $[0, M]$, with $M \gtrsim 1000$, are rejected. On the other hand, relative fluxes with associated aperture peak pixel count (**Peak_T1**, **Peak_Ci**) above a certain threshold (in our case 45000) are also discarded. These 2 procedures are always to be applied, to all stars under consideration, before performing σ -clipping, numerical clipping and/or binning (and also before optimized selection of comparison stars).

F Description of variability indices

F.1 Interquartile range (IQR)

The *interquartile range* index is a robust measure of scatter, meaning that is mostly insensitive to outliers. This is mainly because of its definition, which excludes, in the context of photometry, 25% of both the brightest and the faintest flux measurements. It is defined as the difference between the median values of the upper and lower halves of the data set, which is (previously) divided by its overall median. For a normal distribution with mean μ and variance σ^2 , the value of its IQR can be expressed analytically using the cumulative distribution function, $\Phi(z)$, of the standard normal $\text{Norm}(0, 1)$. The first and third quartiles, $Q_1(\mu, \sigma^2), Q_3(\mu, \sigma^2)$ can be easily computed: $Q_1(\mu, \sigma^2) = \mu + \Phi^{-1}(0.25)\sigma = \mu - \Phi^{-1}(0.75)\sigma$ and $Q_3(\mu, \sigma^2) = \mu + \Phi^{-1}(0.75)\sigma$. Hence, we have:

$$\text{IQR for Norm}(\mu, \sigma^2): 2\Phi^{-1}(0.75)\sigma \simeq 1.349\sigma \quad (24)$$

This result can be used to compare IQR values for data that is assumed to follow a Gaussian distribution, or any distribution sufficiently close to it. Thus, given an estimation for the standard deviation of the data, σ , the normalized interquartile range can be computed as the quotient $\mathbf{IQR} = \text{IQR}/[2\Phi^{-1}(0.75)\sigma] \simeq \text{IQR}/(1.349\sigma)$.

F.2 χ^2 test

Stellar variability detection can be framed in the context of classical hypothesis testing, for example, using a χ^2 statistic. A χ^2 test is any statistical hypothesis test in which the sampling distribution is a χ^2 distribution of a certain degree when the null hypothesis is true (recall that a χ^2 distribution with d degrees of freedom is equivalent to the distribution resulting from the sum $\sum_{i=1}^d Z_i^2$ with each Z_i a random variable following a standard normal distribution, $\text{Norm}(0, 1)$). Thus, given n magnitude measurements, m_i , together with their associated errors, σ_i , and assuming the measurements are independent and the errors follow a Gaussian distribution centered at 0, the null hypothesis, H_0 , that the object under study is non-variable can be tested by computing the so called χ^2 statistic, that we represent with $\tilde{\chi}^2$:

$$\tilde{\chi}^2 \equiv \sum_{i=1}^n \frac{(m_i - \overline{m}_w)^2}{\sigma_i^2} \quad (25)$$

and follows a χ^2 distribution with $\nu = n - 1$ degrees of freedom. \overline{m}_w in the previous equation is the weighted mean flux, computed as in (2).

The χ^2 test with statistical significance level $\alpha \in (0, 1)$ consists in comparing the value of $\tilde{\chi}^2$ to the critical value $\chi_{\alpha, \nu}^2 = \text{CDF}(\chi_\nu^2)^{-1}(1 - \alpha)$, where $\text{CDF}(\chi_\nu^2)$ is the cumulative distribution function of the χ^2 distribution with ν degrees of freedom. Thus, if $\tilde{\chi}^2 > \chi_{\alpha, \nu}^2$ H_0 can be rejected with a confidence of $100 \times (1 - \alpha)\%$ and the object would be classified as variable; else, H_0 cannot be rejected.

If flux errors are estimated correctly, non-variable objects should have $\tilde{\chi}^2$ values consistent with the null hypothesis. The value of $\tilde{\chi}^2$ can then clearly be used as a measure of scatter in a light-curve. To compare $\tilde{\chi}^2$ values between light curves with different number of flux measurements, a reduced χ^2 statistic is commonly used (see [1]): $\chi^2 = \tilde{\chi}^2/n - 1$.

F.3 Standard deviation, σ

A variable source should have a detectable amount of dispersion in its light curve, compared to non-variable objects measured with similar precision. The classical unbiased estimator for the standard deviation of a variable light curve should serve as an indicator of flux variability. Alternatively, as described in 2.1, we can use its weighted version (see equation (2)), σ_w , if flux errors are known. One of the main problems of standard deviation is that it is relatively sensitive to outlier points. Hence, as anticipated, light curve filtering is generally needed before σ can be used as an efficient variable selection tool. To compare scattering of light curves using standard deviation values, we may normalize them with the corresponding estimated flux mean, \overline{m} (\overline{m}_w): $\sigma = \sigma/\overline{m}$ (σ_w/\overline{m}_w). With such normalization, the resulting index is known as the *coefficient of variation* (CV) or the *relative standard deviation* (RSD).

F.4 The von Neumann ratio, η

Each of the previous indices is scatter-based, meaning that they take into account information relative to the distribution of measured fluxes but fail to incorporate most of the time information relative to the associated light curve. In [32] 11 correlation-based indices are studied. This type of variability indices, in addition to flux measurements, consider also the order in which the measurements were performed and, in some cases, the time difference between them, which makes correlation-based indices more sensitive to low-amplitude stellar variability, as the time information encoded in a light curve is not completely lost.

An efficient correlation-based index according to [32] is the inverse of the von Neumann ratio, $1/\eta$, which has shown to perform adequately in detecting variability on time-scales longer than the typical time interval between observations. The von Neumann ratio, η , is defined as the ratio of the mean square successive difference to the distribution variance:

$$\eta \equiv \frac{\delta^2}{\sigma^2} = \frac{\sum_{i=1}^{n-1} (m_{i+1} - m_i)^2 / (n-1)}{\sum_{i=1}^n (m_i - \bar{m})^2 / (n-1)} \quad (26)$$

and quantifies the smoothness of a time series. Hence, its reciprocal, $1/\eta$, can be used as an indicator of variability, with larger index values corresponding to greater likelihood of an object being variable. One of the good properties of this index is that it remains useful if the measurements follow a nearly-symmetric (not necessarily Gaussian) distribution.

References

- [1] Rene Andrae, Tim Schulze-Hartung, and Peter Melchior. *Dos and don'ts of reduced chi-squared*. 2010. arXiv: [1012.3754 \[astro-ph.IM\]](#) (cit. on p. 37).
- [2] Blomme, R. et al. “Variability in the CoRoT photometry of three hot O-type stars - HD 46223, HD 46150, and HD 46966”. In: *A&A* 533 (2011), A4. DOI: [10.1051/0004-6361/201116949](#). URL: <https://doi.org/10.1051/0004-6361/201116949> (cit. on pp. 5, 14).
- [3] Boisse, I. et al. “Disentangling between stellar activity and planetary signals”. In: *A&A* 528 (2011), A4. DOI: [10.1051/0004-6361/201014354](#). URL: <https://doi.org/10.1051/0004-6361/201014354> (cit. on pp. 5, 3).
- [4] D. Boyd. “Differential Photometry Using Multiple Comparison Stars”. In: *Society for Astronomical Sciences Annual Symposium* 26 (May 2007), p. 119 (cit. on p. 1).
- [5] Karen A. Collins et al. “AstroImageJ: Image Processing and Photometric Extraction for Ultra-Precise Astronomical Light Curves”. In: *The Astronomical Journal* 153.2 (Jan. 2017), p. 77. DOI: [10.3847/1538-3881/153/2/77](#) (cit. on pp. 5, 1, 8, 30).
- [6] J. Colome and I. Ribas. “ICAT: a General Purpose Image Reduction and Analysis Tool for Robotic Observatories”. In: *IAU Special Session* 6, 11 (Aug. 2006), p. 11 (cit. on p. 10).
- [7] Trevor Z. Dorn-Wallenstein, Emily M. Levesque, and James R. A. Davenport. “Short-term Variability of Evolved Massive Stars with TESS”. In: *The Astrophysical Journal* 878.2 (June 2019), p. 155. ISSN: 1538-4357. DOI: [10.3847/1538-4357/ab223f](#). URL: <http://dx.doi.org/10.3847/1538-4357/ab223f> (cit. on pp. 5, 14).
- [8] Mark E. Everett and Steve B. Howell. “A Technique for Ultrahigh-Precision CCD Photometry”. In: *Publications of the Astronomical Society of the Pacific* 113.789 (Nov. 2001), pp. 1428–1435. DOI: [10.1086/323387](#). URL: <https://doi.org/10.1086/323387> (cit. on p. 1).
- [9] L. Eyer and P. Bartholdi. “Variable stars: Which Nyquist frequency?” In: 135 (Feb. 1999), pp. 1–3. DOI: [10.1051/aas:1999102](#). arXiv: [astro-ph/9808176 \[astro-ph\]](#) (cit. on p. 6).
- [10] J. Frith et al. “A catalogue of bright (K < 9) M dwarfs”. In: *Monthly Notices of the Royal Astronomical Society* 435.3 (Aug. 2013), pp. 2161–2170. ISSN: 0035-8711. DOI: [10.1093/mnras/stt1436](#). URL: <http://dx.doi.org/10.1093/mnras/stt1436> (cit. on p. 11).
- [11] Mark Gallaway. *An Introduction to Observational Astrophysics*. Switzerland: Springer Nature Switzerland AG 2020, 2016. ISBN: 978-3-030-43550-9 (cit. on pp. 28, 29).
- [12] Gershberg, R. E. et al. “Catalogue and bibliography of the UV Cet-type flare stars and related objects in the solar vicinity*”. In: *Astron. Astrophys. Suppl. Ser.* 139.3 (1999), pp. 555–558. DOI: [10.1051/aas:1999407](#). URL: <https://doi.org/10.1051/aas:1999407> (cit. on p. 14).
- [13] D. S. Hall and G. W. Henry. “The Law of Starspot Lifetimes”. In: *International Amateur-Professional Photoelectric Photometry Communications* 55 (Mar. 1994), p. 51 (cit. on p. 3).
- [14] S. B. Howell, K. J. Mitchell, and III Warnock A. “Statistical error analysis in CCD time-resolved photometry with applications to variable stars and quasars.” In: 95 (Jan. 1988), pp. 247–256. DOI: [10.1086/114634](#) (cit. on p. 1).
- [15] Pablo Huijse et al. “An Information Theoretic Algorithm for Finding Periodicities in Stellar Light Curves”. In: *IEEE Transactions on Signal Processing* 60.10 (Oct. 2012), pp. 5135–5145. ISSN: 1941-0476. DOI: [10.1109/tsp.2012.2204260](#). URL: <http://dx.doi.org/10.1109/TSP.2012.2204260> (cit. on p. 5).
- [16] Zeljko Ivezic et al. *Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data*. Princeton Series in Modern Observational Astronomy (Princeton University Press), 2014 (cit. on pp. 6, 33).
- [17] J. S. Jenkins et al. “Rotational velocities for M dwarfs”. In: *The Astrophysical Journal* 704.2 (Sept. 2009), pp. 975–988. ISSN: 1538-4357. DOI: [10.1088/0004-637x/704/2/975](#). URL: <http://dx.doi.org/10.1088/0004-637x/704/2/975> (cit. on p. 11).
- [18] Hans Kjeldsen and Soren Frandsen. “High-precision time-resolved CCD photometry”. In: *Publications of the Astronomical Society of the Pacific* 104 (June 1992), p. 413. DOI: [10.1086/133014](#). URL: <https://doi.org/10.1086/133014> (cit. on p. 1).

- [19] Molly R. Kosiarek and Ian J. M. Crossfield. “Photometry as a Proxy for Stellar Activity in Radial Velocity Analyses”. In: *The Astronomical Journal* 159.6 (May 2020), p. 271. DOI: [10.3847/1538-3881/ab8d3a](https://doi.org/10.3847/1538-3881/ab8d3a). URL: <https://doi.org/10.3847/1538-3881/ab8d3a> (cit. on pp. 5, 3).
- [20] N. R. Lomb. “Least-squares frequency analysis of unequally spaced data”. In: *Astrophysics and Space Science* 39 (Feb. 1976), pp. 447–462. DOI: [10.1007/BF00648343](https://doi.org/10.1007/BF00648343). URL: <https://doi.org/10.1007/BF00648343> (cit. on p. 5).
- [21] James P. Long, Eric C. Chi, and Richard G. Baraniuk. “Estimating a common period for a set of irregularly sampled functions with applications to periodic variable star data”. In: *Ann. Appl. Stat.* 10.1 (Mar. 2016), pp. 165–197. DOI: [10.1214/15-AOAS885](https://doi.org/10.1214/15-AOAS885). URL: <https://doi.org/10.1214/15-AOAS885> (cit. on p. 34).
- [22] W. J. Merline. “A realistic model for point-sources imaged on array detectors: The model and initial results”. In: *Experimental Astronomy* 6 (Jan. 1995), pp. 163–210. DOI: [10.1007/BF00421131](https://doi.org/10.1007/BF00421131). URL: <https://doi.org/10.1007/BF00421131> (cit. on p. 1).
- [23] E.F. Milone and C. Sterken. *Astronomical Photometry: Past, Present and Future*. New York: Springer Science + Business Media, LLC, 2011. ISBN: 978-1-4419-8049-6 (cit. on pp. 2, 29).
- [24] Elisabeth R. Newton et al. “The rotation and galactic kinematics of mid M dwarfs in the solar neighborhood”. In: *The Astrophysical Journal* 821.2 (Apr. 2016), p. 93. ISSN: 1538-4357. DOI: [10.3847/0004-637x/821/2/93](https://doi.org/10.3847/0004-637x/821/2/93). URL: <http://dx.doi.org/10.3847/0004-637x/821/2/93> (cit. on pp. 5, 11).
- [25] G. Peres et al. “Low amplitude variability and transient periodicity in FF Andromedae and other active stars.” In: 278 (Oct. 1993), pp. 179–186 (cit. on p. 3).
- [26] George R. Ricker et al. “Transiting Exoplanet Survey Satellite”. In: *Journal of Astronomical Telescopes, Instruments, and Systems* 1.1 (2014), pp. 1–10. DOI: [10.1117/1.JATIS.1.1.014003](https://doi.org/10.1117/1.JATIS.1.1.014003). URL: <https://doi.org/10.1117/1.JATIS.1.1.014003> (cit. on p. 13).
- [27] J. D. Scargle. “Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data.” In: 263 (Dec. 1982), pp. 835–853. DOI: [10.1086/160554](https://doi.org/10.1086/160554) (cit. on p. 5).
- [28] J. D. Scargle. “Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data.” In: 263 (Dec. 1982), pp. 835–853. DOI: [10.1086/160554](https://doi.org/10.1086/160554) (cit. on p. 6).
- [29] Arthur Schuster. “On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena”. In: *Terrestrial Magnetism (Journal of Geophysical Research)* 3.1 (Jan. 1898), p. 13. DOI: [10.1029/TM003i001p00013](https://doi.org/10.1029/TM003i001p00013) (cit. on pp. 5, 32).
- [30] Min-Su Shin, Michael Sekora, and Yong-Ik Byun. “Detecting variability in massive astronomical time series data – I. Application of an infinite Gaussian mixture model”. In: *Monthly Notices of the Royal Astronomical Society* 400.4 (Dec. 2009), pp. 1897–1910. ISSN: 0035-8711. DOI: [10.1111/j.1365-2966.2009.15576.x](https://doi.org/10.1111/j.1365-2966.2009.15576.x). eprint: <https://academic.oup.com/mnras/article-pdf/400/4/1897/5653684/mnras0400-1897.pdf>. URL: <https://doi.org/10.1111/j.1365-2966.2009.15576.x> (cit. on pp. 3, 5).
- [31] A. Skumanich. “Time Scales for Ca II Emission Decay, Rotational Braking, and Lithium Depletion”. In: 171 (Feb. 1972), p. 565. DOI: [10.1086/151310](https://doi.org/10.1086/151310) (cit. on p. 3).
- [32] K. V. Sokolovsky et al. “Comparative performance of selected variability detection techniques in photometric time series data”. English. In: *Monthly Notices of the Royal Astronomical Society* 464.1 (2017), pp. 274–292. ISSN: 0035-8711. DOI: [10.1093/mnras/stw2262](https://doi.org/10.1093/mnras/stw2262) (cit. on pp. 5, 4, 38).
- [33] Gudmundur Stefansson et al. *A Mini-Neptune and a Venus-Zone Planet in the Radius Valley Orbiting the Nearby M2-dwarf TOI-1266: Validation with the Habitable-zone Planet Finder*. 2020. arXiv: [2006.11180](https://arxiv.org/abs/2006.11180) [astro-ph.EP] (cit. on pp. 5, 13).
- [34] Peter B. Stetson. “On the Automatic Determination of Light-Curve Parameters for Cepheid Variables”. In: 108 (Oct. 1996), p. 851. DOI: [10.1086/133808](https://doi.org/10.1086/133808) (cit. on p. 5).
- [35] Ryan C. Terrien et al. “A near-infrared spectroscopic survey of 886 nearby M dwarfs”. In: *The Astrophysical Journal Supplement Series* 220.1 (Sept. 2015), p. 16. ISSN: 1538-4365. DOI: [10.1088/0067-0049/220/1/16](https://doi.org/10.1088/0067-0049/220/1/16). URL: <http://dx.doi.org/10.1088/0067-0049/220/1/16> (cit. on p. 11).
- [36] The Astropy Collaboration et al. “Astropy: A community Python package for astronomy”. In: *A&A* 558 (2013), A33. DOI: [10.1051/0004-6361/201322068](https://doi.org/10.1051/0004-6361/201322068). URL: <https://doi.org/10.1051/0004-6361/201322068> (cit. on p. 7).

- [37] M. Tuomi et al. *Frequency of planets orbiting M dwarfs in the Solar neighbourhood*. 2019. arXiv: [1906.04644 \[astro-ph.EP\]](#) (cit. on pp. [5](#), [14](#), [18](#)).
- [38] Jacob T. VanderPlas. “Understanding the Lomb–Scargle Periodogram”. In: *The Astrophysical Journal Supplement Series* 236.1 (May 2018), p. 16. ISSN: 1538-4365. DOI: [10.3847/1538-4365/aab766](#). URL: <http://dx.doi.org/10.3847/1538-4365/aab766> (cit. on p. [34](#)).
- [39] Zechmeister, M. and Kürster, M. “The generalised Lomb-Scargle periodogram - A new formalism for the floating-mean and Keplerian periodograms”. In: *A&A* 496.2 (2009), pp. 577–584. DOI: [10.1051/0004-6361:200811296](#). URL: <https://doi.org/10.1051/0004-6361:200811296> (cit. on pp. [5](#), [7](#)).
- [40] Zechmeister, M. and Kürster, M. “The generalised Lomb-Scargle periodogram - A new formalism for the floating-mean and Keplerian periodograms”. In: *A&A* 496.2 (2009), pp. 577–584. DOI: [10.1051/0004-6361:200811296](#). URL: <https://doi.org/10.1051/0004-6361:200811296> (cit. on p. [33](#)).