UNIVERSITAT DE BARCELONA

TFAW: Noise filtering Through the use of the Wavelet Transform in Astronomy Photometric Data

Daniel del Ser Badia

TFAW: NOISE FILTERING THROUGH THE USE OF THE WAVELET TRANSFORM IN ASTRONOMY PHOTOMETRIC DATA

DANIEL DEL SER BADIA

UNIVERSITAT DE
BARCELONA

Application to Exoplanetary Transit Surveys

September 2018

Programa de Doctorat en Física
Línia de Recerca en Astronomia i Astrofísica

Tesis doctoral presentada por Daniel del Ser Badia
para optar al título de doctor por la Universitat de Barcelona

DIRECTORES:
Dr. Octavi Fors Aldrich
Dr. Jorge Núñez de Murga
TUTOR:
Dr. Alberto Manrique Oliva

Daniel del Ser Badia: *TFAW: Noise filtering Through the use of the Wavelet Transform in Astronomy Photometric Data,* Application to Exoplanetary Transit Surveys, PhD Thesis, © September 2018

Of course there are worlds. Millions of them!
Every star you see has worlds, and most of those you don't see.

— Isaac Asimov, Pebble in the Sky

Go then, there are other worlds than these.

— Stephen King, The Gunslinger

I would rather spend one lifetime with you,
than face all the ages of this world alone.

— J.R.R. Tolkien, The Lord of the Rings

ACKNOWLEDGMENTS

En primer lugar me gustaría agradecer a mis directores de tesis, los doctores Octavi Fors Aldrich y Jorge Núñez de Murga su dedicación, entrega y apoyo a lo largo de la realización de esta tesis. En especial me gustaría agradecerles que me acogieran en el proyecto del TFRM que me ha permitido, junto con su experiencia y conocimientos científicos, crecer como persona y como astrónomo.

No puedo dejar de agradecer al resto de miembros actuales del TFRM por su acogida y por su enseñanza. A Maria Teresa Merino por las horas de terapia durante los viajes al Montsec y por introducirme en el apasionante mundo del hardware, software y de la seguridad informática; a Lluis Canals por su ayuda, comentarios y apoyo; a Ricardo Morcillo por enseñarme un poco de sus inconmensurables conocimientos tecnológicos y técnicos; y a Josep Tous, el último en llegar pero a quien le agradezco mucho que me haya liberado de unos cuantos viajes al Montsec. También mi agradecimiento para aquellas personas que están o estuvieron involucradas en el proyecto: a Javier Montojo, a Jaime Boloix, a Mike Mazur, a Elwood C. Downey, a Albert Rosich y a Martí Rosselló.

I would also like to thank Dr. Nicholas Law for inviting me twice to collaborate with The Evryscope group at the University of North Carolina at Chapel Hill and for his teachings, support and advice. Also thanks to the rest of The Evryscope group, Jeff Ratzloff, Hank Corbett, Ward Howard and Carl Ziegler for their warm welcome and their help during my research stays.

Igualmente, me gustaría agradecer a los doctores Marc Ribó Gomis y Xavier Paredes Fortuny que, a través de sus observaciones con el TFRM me permitieran aprender más sobre las fuentes gamma. Así mismo, me gustaría agradecer al doctor Benito Marcote Martín por introducirme en el fascinante mundo de los Fast Radio Bursts. También agradecer al doctor Josep Mª Solanes su ayuda facilitándome el acceso a su servidor para realizar parte de los cálculos presentados en este trabajo. No puedo dejar de agradecer también todo el apoyo, aprecio y comprensión de mi compañero de despacho el doctor Daniel Galindo. También agradecer toda la ayuda y trabajo realizado

por el personal de Secretaría del Departamento y al personal informático, en especial a JR por su buen hacer y por tener siempre una sonrisa amable para todos.

A mis padres y mi hermana, a quienes dedico esta tesis, no puedo acabar de agradecerles su cariño, su paciencia y su apoyo que en todo momento me han acompañado y me han dado fuerzas para terminar esta tesis. Y les pido perdón por todos aquellos momentos en los que se lo he hecho pasar mal y en los que el mal humor ha sacado lo peor de mi. Sólo puedo deciros: os quiero.

También agradecerles todo su apoyo, su ayuda y su compañía a mis amigos de La Vellés. No puedo expresar con palabras lo que contar con vuestra amistad ha supuesto para mí. A Ale y Trole, mis hermanos; a Rosi por todo y más; a Ana por su cariño y sus abrazos; a Tere por iluminar el mundo con su risa y alegría; a Chechu por todos los años compartidos desde el cole; a Javi por todas las horas discutiendo como dos físicos locos; a Kike por ponerme mis motes; a Quillo por hacer mi estancia en Barcelona más vellesina; a Héctor por su amistad y por esos algodones de azúcar en fiestas; a Nanú por sus "Nanunadas" que nos alegran la vida a todos; a Efrén por su desparpajo; a Manolillo, te queremos de vuelta; a Ezty, Sali, Imanol y Nacho por poner color, cada uno a vuestra manera, a nuestras vidas. A los Exponjas: Marta, Charlie, Marina y Jorge por ser como mis hermanos pequeños, por vuestra preocupación y vuestro apoyo, gracias por todo. A Chopo, Rober y Rubén por esas vueltas al pueblo hablando de lo divino y lo humano. A todos en general muchas gracias por estar ahí cuando os necesitaba. Alber, esta tesis también va por tí; te echo de menos.

Así mismo, darle las gracias a todos los socios de la Asociación Juvenil La Armuña de La Vellés. A los viejos por ayudarme a comenzar esta singladura hace ya 11 años y, a los nuevos, por insuflar aires frescos a esta nuestra Asociación.

A la Unión CastelloMurciana, por hacer más pasable la semana esperando a que llegase el finde para disfrutar de ese desparpajo del sur.

Darle las gracias también a mis compañeros de carrera en especial, a Lluís Hernández por su risa, sus chistes malos, por sacar tiempo para mí en su apretada agenda social y por presentarme el "Firefly" y así hacerme más llevadera la semana. A Carla Marín por todos estos años compartiendo prácticas de Astronomía, clases de carrera y máster y, sobretodo por su amistad y afecto. A Gonzalo Vilella, por estar ahí cuando lo necesitaba, por su amistad y por sus visitas a Salamanca, gracias.

También gracias a Joan, a Anna, a Marina, a Guillem, a Pau y a Quique por los años compartidos.

Por último, agradecerles todo su apoyo a mis compañeros de departamento. A Ignasi Pérez, a Laia Casamiquela y a Víctor Moreno por todos estos años de experiencias compartidas, por vuestra amistad y por vuestra ayuda cuando la he necesitado. A Nico, por demostrarme que una gran amistad puede surgir en muy poco tiempo, por estar siempre ahí. A Núria, porque pese a que Lluís diga que eres la "hater" oficial, tu risa y afecto nos hacen más llevadera la vida en el departamento. A Alfred por su alegría isleña, por hacerme reír incluso cuando menos ganas tenía y por enseñarme que El Señor de los Anillos en mallorquín tiene su gracia. A José Luís y a Pau por nuestras discusiones frikis: May the Force be with you. A Roger por nuestras charlas de cine fantástico y su risa contagiosa. A Edgar por proveernos de una válvula de escape para nuestras artes de canto. A Pacheco por exhortarnos a luchar por nuestros derechos como estudiantes y trabajadores y a que mantengamos una visión crítica. A Juan por demostrarle a alguien de Salamanca que puede existir algún pucelano bueno. A Héctor, por su reciente amistad. Also thanks to our international friends, David, Katie, Giulio, Hamish and Nicolas for expanding the horizons of our little, windowless department. También un recuerdo para Santi, Benito, Xavi, Pere, Rosa, Pol, Núria Miret, Núria Jordana y Guillem.

RESUMEN

Desde la primera detección confirmada de varios exoplanetas de masa terrestre orbitando el púlsar PSR B125712 (Wolszczan y Frail, 1992; Wolszczan, 1994), el interés por estos objetos no ha hecho más que crecer. El descubrimiento del exoplaneta gigante 51 Pegasi b (Mayor y Queloz, 1995), mediante el método de las velocidades radiales, promovió el desarrollo de una nueva técnica de detección basada en la alta probabilidad de que un planeta similar pudiera cruzar su estrella anfitriona en la línea de visión del observador. Esta técnica, conocida como el método del tránsito, detecta planetas extrasolares midiendo el pequeño oscurecimiento del flujo estelar cuando el planeta pasa entre la estrella y el observador. El método del tránsito es, actualmente, el modo más eficiente y sensible para detectar planetas extrasolares. Además, dada la posibilidad de escanear simultáneamente amplias regiones de cielo que contienen millares de estrellas, las observaciones dedicadas a tránsitos puede detectar más exoplanetas que otros métodos. Muchas misiones han seguido este modo de observación, aquellas basadas en tierra, como MEarth (Irwin y col., 2009), SuperWASP (Pollacco y col., 2006), KELT (Pepper y col., 2007), HAT-South (Bakos y col., 2008), TFRM-PSES (Fors y col., 2013), NGTS (Wheatley y col., 2013) o el Evryscope (Law y col., 2014; Law y col., 2015); así como aquellas misiones espaciales como COnvection ROtation and planetary Transits (CoRoT) (Fridlund y col., 2006), Kepler (Borucki y col., 2010), la recientemente lanzada, TESS (Ricker, 2014) y la futura misión PLATO (Rauer, 2013).

Los tránsitos pueden aportar información sobre las características de los planetas que orbitan su estrella anfitriona. Dado que la caída en el flujo proveniente de la estrella es proporcional al tamaño del planeta y, dado que el tamaño de las estrellas puede conocerse con suficiente precisión, el radio del planeta puede determinarse directamente de la profundidad del tránsito. Combinado con observaciones en velocidades radiales, también puede ofrecer una buena estimación de la masa del planeta y, por lo tanto, también de la densidad. Además, el método del tránsito también permite estudiar la atmósfera del planeta mediante espectrografía del alta resolución, así como su albedo, su temperatura (Charbonneau y col., 2005; Deming y col., 2005) e incluso, detectar la presencia (Kreidberg y col., 2014) o ausencia (Nikolov y col., 2018) de nubes en su atmósfera.

El número total de exoplanetas confirmados hasta la fecha (2 de agosto, 2018) es de 3,774 de los cuales, 2,951 han sido detectados mediante el método del tránsito. En particular, la misión Kepler de la NASA ha detectado 2,327 planetas confirmados y más de 2,000 candidatos. Su misión extendida K2, actualmente ha confirmado 323 planetas y ha detectado más de 400 candidatos.

A parte de los tránsitos exoplanetarios, hay otras fuentes de variabilidad estelar que son detectadas por estas misiones. Estas variaciones pueden ser debidas a cambios físicos en la estrella (como es el caso de las estrellas pulsantes, eruptivas, cataclísmicas o emisoras de rayos X) o, debidas al eclipse de la estrella por otra estrella en el sistema o por efectos de rotación estelar (estrellas eclipsantes, rotantes o eventos de microlensing). La primera estrella variable fue identificada en 1638 cuando Johannes Holwarda (1640) detectó que Omicron Ceti pulsaba cíclicamente cada 11 meses. Desde entonces, el número de estrellas variables ha ido creciendo rápidamente, especialmente desde 1890 cuando se hizo posible identificarlas a partir de placas fotográficas y, más adelante, a partir de instrumentos electrónicos como las CCDs. Actualmente, el General Catalogue of Variable Stars (GCVS) cuenta con 52,011 objectos variables descubiertos y catalogados (Samus' y col., 2017). De igual manera, el Variable Star Index (VSX) cuenta con un catálogo de 541,955 estrellas variables.

La precisión fotométrica y la exactitud conseguida por una misión es un factor clave en la detección y caracterización de una señal correspondiente a un tránsito o a cualquier otro tipo de variabilidad. Muchas de los efectos sistemáticos que afectan a una curva de luz de una estrella son compartidos por otras estrellas en el mismo conjunto de datos. Para eliminarlos, uno puede identificar aquellos objetos que comparten las mismas variaciones sistemáticas que la curva de luz objetivo y, a partir de ellos, construir algún tipo de filtro para corregirla.

Las wavelets tienen unas propiedades que las hacen ideales para analizar señales de naturaleza no estacionaria. A diferencia de la transformada de Fourier, la transformada wavelet descompone una señal en sus wavelets, es decir, impulsos altamente localizados obtenidos de escalar y desplazar la función wavelet madre. Estas operaciones de escalado y desplazamiento permiten calcular los coeficientes wavelet, que representan la correlación entre la wavelet y una sección localizada de la señal. La forma irregular de las wavelets permite analizar señales con discontinuidades, singularidades o cambios bruscos mientras

que su naturaleza compacta permite la localización temporal de las características de la señal.

A lo largo del texto desarrollamos el marco de trabajo a partir del cual presentar la principal contribución de esta tesis, el Wavelet-based Trend Filtering Algorithm (TFAW). TFAW es un algoritmo totalmente genérico, desarrollado y paralelizado en Python basado en el Trend Filtering Algorithm (TFA) desarrollado por Kovács, Bakos y Noyes (2005). El algoritmo es útil para mejorar el rendimiento en la detección, reconstrucción y caracterización de señales astrofísicas. TFAW difiere de otros métodos de filtrado de ruido basados en wavelets en que no requiere ningún modelo de ajuste paramétrico o cualquier otro método computacional. TFAW estima la contribución de ruido de la señal a partir de su Stationary Wavelet Transform (SWT) y el filtrado se realiza eliminando esta contribución de la señal. Además, TFAW es capaz de hacer el filtrado de la señal sin modificar ninguna de sus características intrínsecas a diferencia de otros métodos como los umbrales calibrados de los coeficientes wavelet que pueden dar lugar a distorsiones de la señal o introducir oscilaciones o perturbaciones artificiales alrededor de discontinuidades.

En el Capítulo 2 hacemos una breve descripción de cómo se desarrolló y en qué consiste la teoría wavelet. En el Capítulo 3 explicamos cómo pueden usarse las wavelets para filtrar ruido en curvas de luz fotométricas. También presentamos un método hecho a medida para este trabajo para la eliminación de valores atípicos en series temporales. De igual modo, también introducimos uno de los puntos clave en la aplicación del algoritmo TFAW: el criterio de selección de los denominados *signal* y *noise level*. En el Capítulo 4 hacemos una descripción extensiva del algoritmo y estudiamos su rendimiento sobre curvas de luz simuladas usando una batería de tests. El Capítulo 5 describe las misiones de tierra y espaciales usadas para validar el algoritmo y, el Capítulo 6 presenta los resultados de aplicar TFAW a curvas reales provenientes de dichas misiones. Finalmente el Capítulo 7 presenta las conclusiones de esta tesis.

CONTENTS

1 INTRODUCTION

Since the times of the ancient Greek atomists (Leucippus (5th cent. BCE), Democritus (circa 460-370 B.C.) or Epicurus (circa 341-270 B.C.)), humanity has theorized on the existence of worlds other than our own. Though the work of Aristotle did much to overshadow the advancement on this subject, the introduction of the heliocentric theory by Copernicus (Copernicus, 1543), supposed a paradigm shift. Philosophers, like Giordano Bruno (Bruno, 1584), or scientists, like Christiaan Huygens (Huygens, 1698), speculated that planets orbiting stars other than the Sun existed. However, there was no way of knowing how common they were or if they had any similarities to the planets of our Solar System. Some of the first detection claims, dating back as far as the nineteenth century involving the binary star 70 Ophiuchi (Jacob, 1855) or, in the mid-twentieth century, with Barnard's Star (van de Kamp, 1963), were all eventually rejected by astronomers. Some time had to pass until, in 1988, claims of an exoplanet around Gamma Cephei were presented based on Radial Velocity (RV) observations (Campbell, Walker, and Yang, 1988) (although the exoplanet was not confirmed until 2003 (Hatzes et al., 2003)). In 1989, the first claim of a sub-stellar companion orbiting the star HD114762 was published (Latham et al., 1989). The first confirmed detection came in 1992, with the discovery of several terrestrial-mass planets orbiting the pulsar PSR B125712 (Wolszczan and Frail (1992) & Wolszczan (1994)).

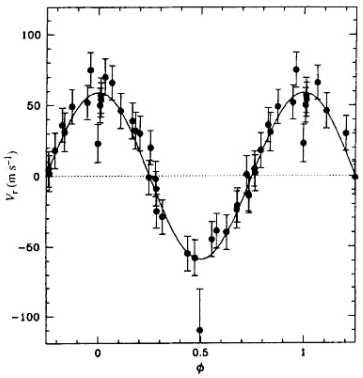


Figure 1: Orbital motion of the first discovered exoplanet 51 Peg b by Mayor and Queloz, 1995.

The first confirmed detection of an exoplanet orbiting a main-sequence star was made in 1995, when a giant planet was found by RV measurements in a four-day orbit around the nearby star 51 Pegasi (Mayor and Queloz, 1995). This detection, together with others of giant exoplanets with orbital periods of less than 10 days, defined a new class of planets called Hot Jupiters. The finding of the giant exoplanet 51 Pegasi b, encouraged the development of another search method based on the high probability that such planet would cross its host star in the line of sight of the observer. This method, known as transit method detects distant planets by measuring the small darkening of a star light curve as an orbiting planet passes between it and the Earth. The first to mention the possibility of observing exoplanetary transits was astronomer Struve (1952). He assumed the possibility of a giant planet (bigger than Jupiter) orbiting in a very small orbit (~1/5 AU) around its host star. This, and taking a mean planetary density of 5× that of the star, resulted in an observed loss of 0.02 magnitudes in the stellar light curve. The first detailed methodology was devised by Rosenblatt (1971). He calculated the color changes that happen during the transit and proposed a 3-wide field telescopes system located at three different sites to monitor stars for brightness variations. Major refinements to the method were done later by Borucki et al. (2010) and Hale and Doyle (1994). Starting in 1994, the Transits of Extrasolar Planets (TEP) program observed the eclipsing binary CM Draconis until 2002 with a network of 10 telescopes (with sizes ranging from 0.6m to 1.2m) placed in 6 countries (Deeg et al., 1998). Six small-amplitude candidate events were recorded, but all of them were ruled out by follow-up observations. The first detection of a transiting exoplanet, HD 209458 b, in 1999 by (Charbonneau et al., 2000) and (Henry et al., 1999) and the discoveries obtained for this planet during follow-up observations (first planet with a detectable atmosphere containing oxygen and carbon, first detection of an evaporating hydrogen atmosphere and being one of the first two exoplanets to be directly observed spectroscopically) demonstrated the high scientific potential of planets discovered with this method.

The first discovered exoplanet based on photometry (Udalski et al., 2002) and confirmed by RV measurements (Konacki et al., 2003) was OGLE-TR-56 b; observed during the 2001 observational campaign of the 1.2m OGLE-III telescope (Udalski et al., 2002). Its orbital period of 1.21 days was the shortest until the discovery of WASP-12 b (Hebb et al., 2009). Since then, thousands of exoplanet detections have been confirmed (see Figure 3) presenting a richer scenario than what could be guessed

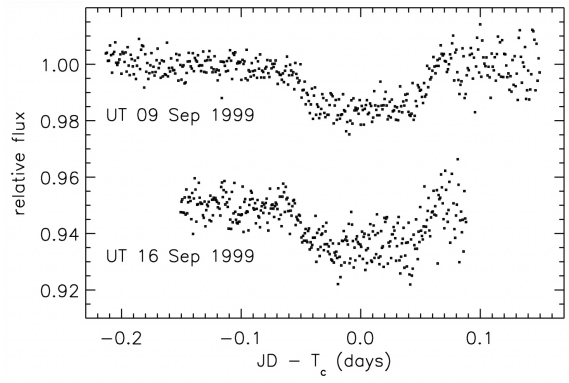


Figure 2: Photometric time series, corrected for gray and color-dependent extinction, of the planet HD 209458 b as shown in (Charbonneau et al., 2000).

from our Solar System, some examples are: the detection of the first circumbinary and also, the oldest to date, exoplanet (PSR B1620-26 b (Sigurdsson et al., 2003)); TrES-2 b the darkest (with a geometric albedo of 0.04%) known exoplanet has been detected (O'Donovan et al. (2006); Kipping and Spiegel (2011)); Formalhault b became the first exoplanet directly imaged by an optical telescope (Kalas et al., 2008); several multiplanetary systems have been detected (625 confirmed[1]) and the first one has been directly imaged (Marois et al., 2008); super-Earths like GJ1214b, have been detected (Charbonneau et al., 2009); a number of potential ocean planets, like Kepler-22 b, the first Kepler mission transiting planet orbiting the Habitable Zone (HZ) of a Sun-like star (Borucki et al., 2012), have been observed; in 2016 the Pale Red Dot[2] campaign announced the discovery of Proxima b, an Earth-sized exoplanet in the HZ of the closest star to the Sun, Proxima Centauri, was detected (Anglada-Escudé et al., 2016); in 2017, the discovery of seven potentially habitable exoplanets orbiting TRAPPIST-1 was announced (Gillon et al., 2017) and planets like K2-141 b (Malavolta et al., 2018) form part of the Ultra-short Period class (planets with orbital periods shorter than one day).

Transit photometry is currently the most effective and sensitive method for detecting extrasolar planets. In addition, by scanning large areas of the sky containing thousands of stars, transit surveys can find more extrasolar planets than other methods (like the RV one). Several surveys have taken this wide-FoV approach, such as the ground-based MEarth (Irwin et al., 2009), SuperWASP (Pollacco et al., 2006), KELT (Pepper et al., 2007), HAT-South (Bakos et al., 2008), TFRM-PSES (Fors et al., 2013),

1 https://exoplanetarchive.ipac.caltech.edu
2 https://palereddot.org/es/

NGTS (Wheatley et al., 2013) or the Evryscope (Law et al., 2014), as well as the space-based CoRoT (Fridlund et al., 2006), Kepler (Borucki et al., 2010), the recently commissioned, TESS (Ricker, 2014) and the future PLATO (Rauer, 2013) missions.

Transits can also provide information about the planets orbiting their hosts stars. The main advantage of the transit method comes from the fact that the drop in a star's flux during transit is proportional to the size of the planet. Due to the fact that the star's size can be known with some accuracy, the planet's size can thus be directly deduced from the transit depth. Combined with RV, it can also provide a good estimate of the planet's mass as the transiting planet necessarily has to be in the line-of-sight of the observer. Taken together, the planet's size and mass provide the planet's density. The second advantage of the transit method is that it also allows to study the atmosphere of the transiting planet. When the planet transits the host star, light from the star passes through the upper atmosphere of the planet. Using high-resolution stellar spectrum, the elements present in the planet's atmosphere can be detected by recreating the absorption spectrum. Another way of detecting exoplanetary atmospheres consist on measuring the polarization of the host star's light as it passes through or is reflected by the planet's atmosphere. In addition, the secondary eclipse provides a direct measurement of the planet's radiation. It allows to measure the planet's temperature (Charbonneau et al., 2005; Deming et al., 2005) and even to detect the presence of clouds (Kreidberg et al., 2014) or their absence (Nikolov et al., 2018).

The total number of confirmed exoplanets to date (August 2, 2018) is 3,774 of which 2,951 have been detected by the transit method. In particular, the Kepler mission, so far the most successful one, has detected 2,327 confirmed exoplanets and other 2,244 candidates, its extended K2 mission currently holds 323 confirmed exoplanets and 479 candidates[3].

Apart from exoplanetary transits, there are other sources of stellar variability. This variation may be caused by a change in emitted light or by something partly blocking the light. Thus, variable stars are classified in *intrinsic variables* if their variation is due to physical changes in the star or stellar system (pulsating, eruptive, cataclysmic or X-ray stars) and *extrinsic variables* if their variability is due to eclipse of one star by another or the effect of stellar rotation (eclipsing, rotating stars and microlensing events). The first variable star was identified in 1638 when Johannes Holwarda (1640) noticed that Omicron Ceti pulsated in a cycle taking 11 months; the star had previously been de-

3 https://exoplanetarchive.ipac.caltech.edu

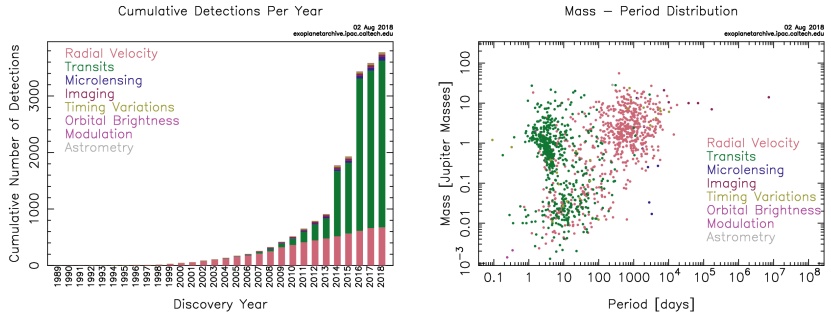


Figure 3: Confirmed exoplanets statistics updated to August 2, 2018. **Left**: Cumulative exoplanet detection per year and detection method. **Right**: Mass-orbital period diagram for all confirmed exoplanets. Credit: NASA Exoplanet Archive

scribed as a nova by David Fabricius in 1596 (Bolt et al., 2007). This discovery, combined with supernovae observed in 1572 and 1604, proved that the starry sky was not eternally invariable as Aristotle and other ancient philosophers had taught. In this way, the discovery of variable stars contributed to the astronomical revolution of the sixteenth and early seventeenth centuries. The second variable star to be described was the eclipsing variable Algol, by Geminiano Montanari in 1669 (Bolt et al., 2007); John Goodricke gave the correct explanation of its variability in 1784 (Bolt et al., 2007). Chi Cygni was identified in 1686 by Gottfried Kirch, then R Hydrae in 1704 by Giovanni Domenico Maraldi (Bolt et al., 2007). By 1786 ten variable stars were known. John Goodricke himself discovered Delta Cephei and Beta Lyrae (Bolt et al., 2007). Since 1850 the number of known variable stars has increased rapidly, especially after 1890 when it became possible to identify variable stars by means of photography. Currently, the General Catalogue of Variable Stars (GCVS) contains data for 52,011 individual variable objects discovered and cataloged as variable stars by 2015 and located mainly in the Milky Way galaxy (Samus' et al., 2017). Similarly, The International VSX[4] currently holds a catalog of 541,955 variable stars.

1.0.1 *Motivation and overview of the thesis*

The photometric precision and accuracy achieved by an astronomical survey is a key factor in detecting a transiting signal or any other kind of variability. Many of the systematic variations in a given light curve are shared by light curves of other stars in the same data set. In order to remove those sys-

4 https://www.aavso.org/vsx/index.php

tematics, one can identify the objects in the field that suffer from the same kind of variations as the target (correlated noise) and then build and apply a filter based on the light curves of these comparison stars. The Trend Filtering Algorithm (TFA) (Kovács, Bakos, and Noyes, 2005) and SysRem (Tamuz, Mazeh, and Zucker, 2005) are often applied to remove systematic variations in time-domain surveys, in particular for wide FoV ones.

Wavelets have unique properties that make them an ideal tool for analyzing signals of non-stationary nature. Unlike the Fourier transform, the wavelet transform decomposes a given time series into its *wavelets*, i.e., highly localized impulses obtained from scaling and shifting the *mother wavelet* function. These scaling and shifting operations allow us to calculate the wavelet coefficients, which represent the correlation between the wavelet and a localized section of the signal. The wavelet coefficients are calculated for each wavelet segment, giving a time-series function measuring the wavelets' correlation to the signal. In comparison to the sine wave used in the Fourier transform, which is smooth and of infinite length, the wavelet is irregular in shape and compactly supported. Their irregular shape allows to analyze signals with discontinuities, transients, singularities and sharp changes, while their compactly supported nature allows temporal localization of the signal's features.

Along this thesis we will lay out the framework from which the main goal of this thesis, the Wavelet-based Trend Filtering Algorithm (TFAW), will be built from. In Chapter 2 we will do a brief review of how wavelet theory was developed as well as give a summarized introduction to wavelet theory itself. In Chapter 3 we will explain how wavelets can be used to filter noise and present a custom-made method for this thesis to remove signal outliers. Also, we will introduce one of the key points in the application of TFAW, the *signal* and *noise level* selection criteria. In Chapter 4 we fully describe the TFAW algorithm and assess its performance over simulated light curves using a battery of tests. Chapter 5 describes the ground- and space-based surveys used to validate the algorithm and, Chapter 6 presents the results of applying TFAW to real light curves coming from these surveys. Finally, in Chapter 7 we present the overall conclusions of this thesis and briefly comment on some future perspectives.

Part I

WAVELET THEORY

2

WAVELETS

Fourier Transforms (FTs) have been long used for signal analysis. Although Fourier representations have perfect compact support in the frequency domain, they lack the capability of preserving time information and, as a result, they cannot be used to approximate signals of a non-stationary nature. In the late 1970's, Jean Morlet, a geophysical engineer working at the Elf Acquitaine oil company developed an alternative for the FT. Morlet wanted to analyze seismic signals with different features in time and frequency which he wanted to decouple. In order to gain time resolution for the high frequency transients and good frequency resolution for the low frequency components, Morlet generated transform functions taking a Gaussian-windowed cosine wave and compressed or spread it to obtain different frequency functions. These functions were shifted also in time to study what happened at different times. He then would take the inner product of the target signal with all these transform functions. The word *wavelet* was already being used in other contexts in geophysics; Morlet decided to name his transform functions *wavelets of constant shape*. Morlet looked for help to give a mathematically rigorous basis to his wavelet transform. He contacted Alex Grossman, a theoretical physicist working on quantum mechanics. Together, they constructed an exact inversion formula and explored several applications of the wavelet transform leading to the formalization of the wavelet transform (Grossmann and Morlet, 1984).

In 1985, Yves Meyer, a mathematician based at the Ecole Polytechnique of Paris, heard about the work of Grossmann and Morlet. He realized that their wavelet analysis was a rediscovery of a formula that Alberto Calderón had introduced in numerical analysis (Calderón, 1964). He also noticed that there was redundancy in the choice of the wavelet families done by Morlet. Meyer started working on developing wavelets with better time-frequency localization and ended up building an orthogonal wavelet basis. In 1986, Stephane Mallat, at that time a graduate student at the University of Pennsylvania, conceived a layered structure for wavelet expansions. Together with Meyer, they developed all the mathematical details of the Multi-Resolution Analysis (MRA). It led to a simple and recur-

sive filtering algorithm to compute the wavelet decomposition of a function from its finest scale approximation (Mallat, 1989).

Also in 1986, during her guest-research stay at the Courant Institute of Mathematical Sciences, Ingrid Daubechies constructed compactly supported continuous wavelets based on quadrature mirror filter-technology. These wavelets would require only a finite amount of processing, enabling wavelet theory to enter the digital signal processing field. In addition, in 1988 she developed the orthonormal bases of compactly supported wavelets (Daubechies, 1988). The works of Mallat and Daubechies set the foundations of the Discrete Wavelet Transform (DWT) analysis. Some years later, Cohen, Daubechies, and Feauveau (1992) developed the compactly supported biorthogonal wavelets which allowed for symmetric wavelets to be built thus increasing the number of applications of the wavelet theory. Holschneider et al. (1989) and later on Shensa (1992), developed the *à trous* algorithm also know as Stationary Wavelet Transform (SWT) in order to overcome the lack of translation invariance of the DWT.

Since then, wavelets have been used in several fields of knowledge with different applications. They have been very successful in image and data compression. Some implementations include the JPEG2000 (Taubman and Marcellin, 2002) for still images, or the Wavelet Scalar Quantization, a wavelet-based compression algorithm developed by the FBI used for gray-scale fingerprint images (Jain, Pankanti, and Archives, 2001). They can also be used for data denoising through the removal of the wavelet coefficients associated to noise through some kind of smart thresholding (Donoho and Johnstone (1994b), Donoho and Johnstone (1994a), Donoho (1995), and Starck and Murtagh (1994) and others). There has been also applications of wavelets to biomedical data and imaging, financial analysis, data mining, remote sensing, image fusion, time series classification or anomaly and singularity detection.

In astronomy, they have been used for astronomical signal processing (Starck and Murtagh, 1994; Núñez and Otazu, 1996; Starck, Murtagh, and Bijaoui, 1998); redshift spectra study (Machado et al., 2013); Cosmic Microwave Background (Moudden et al., 2005), Baryon Acoustic Oscillation analysis (Arnalte-Mur et al., 2012) and galaxy distribution morphology (Martinez, Paredes, and Saar, 1993; Martínez et al., 2005; Antoja et al., 2008; Antoja et al., 2012; Kushniruk, Schirmer, and Bensby, 2017); solar activity (Aschwanden et al., 1998; Giménez de Castro et al., 2001); stellar activity, pulsation and rotation analysis (de Freitas et al., 2010; Bravo et al., 2014); signal detection (Szatmary, Vinko, and Gal, 1994; Otazu et al., 2002; Fors et al., 2008) and light curve

noise analysis (Cubillos et al., 2017) and filtering (Carter and Winn, 2009; Waldmann, 2014; Grziwa, Korth, and Pätzold, 2014; Grziwa et al., 2016) or del Ser, Fors, and Núñez (2018).

2.1 INTRODUCTION TO WAVELET THEORY

2.1.1 *Fourier Transform (FT)*

A Fourier series consists in the expansion of a periodic function $f(t)$ in terms of an infinite sum of sines and cosines:

$$f(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos(nt) + \sum_{n=1}^{\infty} b_n \sin(nt) \tag{1}$$

where:

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(nt)dt \\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin(nt)dt \end{aligned} \tag{2}$$

We can also extend the Fourier series to complex coefficients:

$$f(t) = \sum_{n=-\infty}^{\infty} A_n e^{int} \tag{3}$$

where:

$$A_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}dt \tag{4}$$

which for a periodic function in [-T/2,T/2] becomes:

$$\begin{aligned} f(t) &= \sum_{n=-\infty}^{\infty} A_n e^{i(2\pi nt/T)} \\ A_n &= \frac{1}{T}\int_{-T/2}^{T/2} f(t)e^{-i(2\pi nt/T)}dt \end{aligned} \tag{5}$$

The FT is a generalization of the complex Fourier series. The FT and its inverse are defined by:

$$\begin{aligned} F(\omega) &= \int_{-\infty}^{\infty} f(t)e^{-2\pi i\omega t}dt \\ f(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{2\pi i\omega t}d\omega \end{aligned} \tag{6}$$

A discretized version of the FT can also be defined. Consider a discrete function $f(t_k)$ where $t_k = k\Delta$ with $k = 0, ..., N-1$. The Discrete Fourier Transform (DFT) can be then defined as:

$$F_n = \sum_{k=0}^{N-1} f(t_k) e^{-2\pi i n k/N} \tag{7}$$

and its inverse:

$$f(t_k) = \frac{1}{N} \sum_{k=0}^{N-1} F_n e^{2\pi i k n/N} \tag{8}$$

DFT are used because they can reveal periodicities in the input signal as well as the strength of any periodic component. However, while the FT has perfect frequency resolution, it retains no time information. This limits the usefulness of the FT for analyzing signals that are localized in time, notably transients, or any signal of finite extent. Cooley and Tukey (1965) developed a general version of the Fast Fourier Transform (FFT) allowing to reduce the computational cost of the DFT by reducing the number of operations involved from $O(N^2)$ to $O(N \log N)$.

2.1.2 *Wavelet Definition*

A wavelet is a highly localized impulse obtained from shifting and scaling a function, ψ, called *mother wavelet*. Unlike Fourier bases, which only preserve the frequency information of a signal, wavelet bases are able to preserve both the frequency information and their temporal distribution.

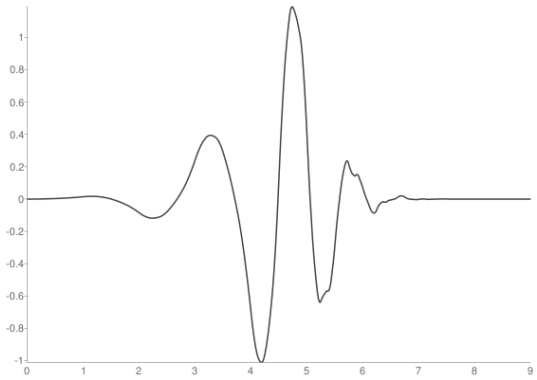


Figure 4: Mother Wavelet for the Daubechies 5 basis

Alfred Haar (Haar, 1910) introduced a piecewise constant function known as the Haar sequence. Haar used these functions to give an example of an orthonormal system for the

space of square-integrable functions on the unit interval [0, 1]. Though the name wavelet came much later, the Haar sequence is also the simplest possible wavelet:

$$\psi(t) = \begin{cases} 1 & \text{if } 0 \leqslant t < 1/2 \\ -1 & \text{if } 1/2 \leqslant t < 1 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

the dilations and translations of which, generate an orthonormal basis:

$$\left\{ \psi_{j,k}(t) = \frac{1}{\sqrt{2^{-j}}} \psi \left(\frac{t - 2^{-j}k}{2^{-j}} \right) \right\}_{(j,k)\in\mathbb{Z}^2} \tag{10}$$

In a more general form, Equation 10 can be written as:

$$\left\{ \psi_{u,v}(t) = \frac{1}{\sqrt{v}} \psi \left(\frac{t - u}{v} \right) \right\}_{(u,v)\in\mathbb{Z}^2} \tag{11}$$

with v and u representing the scale and translation parameters respectively and ψ, the *mother wavelet*. The scale parameter is related to the frequency information and either dilates or compresses the signal f. Small scales (related to high frequencies), compress the signal and provide detailed information while large scales (low frequencies) dilate the signal and provide global information. The collection $\{\psi_{u,v}\}$ forms an orthogonal basis of the Hilbert space $\mathbb{L}^2(\mathbb{R})$ of signals with a finite energy:

$$\| f \|^2 = \int_{-\infty}^{\infty} | f(t) |^2 \, dt < +\infty \tag{12}$$

A wavelet dictionary can be built from a *mother wavelet*. A function $\psi(t)$ is called a *mother wavelet* if it satisfies the following properties (Meyer, 1992):

- $\psi(t)$ and all its derivatives decrease rapidly as $x \to \pm\infty$.
- $\psi(t)$ has zero average:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \tag{13}$$

- The collection $\{\psi_{u,v}\}$ forms an orthogonal basis of $\mathbb{L}^2(\mathbb{R})$ and is built through equation 10.

2.1.3 *Wavelet Transform*

Let us define the inner product in $\mathbb{L}^2(\mathbb{R})$ as:

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(t) g^{\star}(t) dt \tag{14}$$

The *wavelet transform* of a signal f can thus be defined as the inner product of f with the scaled and translated versions of the *mother wavelet*:

$$W_{u,v} = \langle f, \psi_{u,v} \rangle = \int_{-\infty}^{\infty} f(t) \psi^{\star}_{u,v}(t) dt \tag{15}$$

The coefficients $W_{u,v}$ are called *wavelet coefficients* of signal *f(t)*. The latter can be recovered by summing them in the wavelet orthonormal basis:

$$f = \sum_{u=-\infty}^{+\infty} \sum_{v=-\infty}^{+\infty} \langle f, \psi_{u,v} \rangle \psi_{u,v} \tag{16}$$

2.2 CONTINUOUS WAVELET TRANSFORM (CWT)

In 1946, physicist Dennis Gabor, applying ideas from quantum physics, introduced Gaussian-windowed sinusoids for time-frequency decomposition and to provide the best trade-off between spatial and frequency resolution (Gabor, 1946).

In the late 70s Jean Morlet, was trying to develop a method to study waveforms in reflection seismology. Looking to give a rigorous basis to his method, Morlet approached Alexander Grossmann, leading to the formalization of the Continuous Wavelet Transform (CWT) (Grossmann and Morlet, 1984).

The CWT of a function *f(t)* at any scale $v > 0$ and position u is the projection of f on the corresponding wavelet:

$$W_{u,v} = \langle f, \psi_{u,v} \rangle = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{v}} \psi^{\star} \left(\frac{t-u}{v} \right) dt \tag{17}$$

Grossmann and Morlet proof that the Inverse Continuous Wavelet Transform (ICWT) can be computed through:

$$f(t) = \frac{1}{C_{\psi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_{u,v} \frac{1}{\sqrt{v}} \tilde{\psi} \left(\frac{t-u}{v} \right) du \frac{dv}{v^2} \tag{18}$$

where $\tilde{\psi}$ is the dual function of ψ and C_ψ is the admissible constant which for wavelets becomes:

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\,\hat{\psi}(\omega)\,|^2}{|\,\omega\,|} d\omega \tag{19}$$

where $\hat{\psi}(\omega)$ represents the Fourier spectrum of signal $\psi(t)$. A wavelet whose admissible constant satisfies that $0 < C_\psi < \infty$ is called an admissible wavelet. Also, given the presence of ω at the denominator, an admissible wavelet implies that $\hat{\psi}(0) = 0$ thus, and admissible wavelet must integrate to zero:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \tag{20}$$

These conditions imply that wavelets must have a band-pass like spectrum and be of oscillatory nature. In addition, wavelets should decrease quickly with decreasing scale. The regularity conditions (given by the vanishing moments theory) state that a wavelet should be smooth and localized in frequency and time domains (Daubechies, 1992). In summary, admissibility gives us the wave while regularity gives us the decay rate at decreasing scales and, together, give us the wavelet.

2.3 DISCRETE WAVELET TRANSFORM (DWT)

2.3.1 *Dyadic Wavelets*

In order to obtain the Discrete Wavelet Transform (DWT), the translation and scale parameters (u, v) have to be discretized. Daubechies 1992 showed that for a given ψ, the collection $\{\psi_{u,v}\}$ forms and orthogonal basis by means of dyadic translations and dilations of $u = k2^{-j}$ and $v = 2^{-j}$. This way, equation 11 becomes:

$$\left\{ \psi_{j,k}(t) = \frac{1}{\sqrt{2^{-j}}} \psi \left(\frac{t - 2^{-j}k}{2^{-j}} \right) \right\}_{(j,k)\in\mathbb{Z}^2} \rightarrow \psi_{j,k}(t) = 2^{j/2} \psi \left(2^j t - k \right) \tag{21}$$

The previous collection forms an orthonormal base first, if under the inner product (equation 14) it satisfies:

$$\langle \psi_{j,k}, \psi_{l,m} \rangle = \int_{-\infty}^{\infty} \psi_{j,k}(t) \psi^\star_{l,m}(t) dt = \delta_{j,l} \delta_{k,m} \tag{22}$$

where $\delta_{j,l}$, is the Kronecker delta and, second, if every function $f(t) \in \mathbb{L}^2(\mathbb{R})$ can be written as:

$$f = \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} W_{j,k}\psi_{j,k} \tag{23}$$

where the wavelet coefficients $W_{j,k}$ are given by the inner products:

$$W_{j,k} = \langle f, \psi_{j,k} \rangle = 2^{j/2} \int_{-\infty}^{\infty} f(t)\psi^{\star}\left(2^{j}t - k\right) dt \tag{24}$$

It is worth noting that equations 9 and 21 are the same. Thus, the Haar wavelet can be considered the first, compact, dyadic and orthonormal wavelet.

2.3.2 *Scaling Function and Multi Resolution Analysis*

We have already seen in Section 2.2 that one wavelet can be seen as a band-pass filter. Then, a series of dilated wavelets can be described as a *filter bank*. A *filter bank* is an array of band-pass filters that separates the input signal into multiple components, each one carrying a single frequency sub-band of the original signal.

We have also stated that the signal $f(t)$ should have finite energy (equation 12). This allows us to cover its frequency spectrum and its time duration using wavelets. However, given that dyadic wavelets stretch the frequency spectrum and shift the frequency components by a factor of two, we would need an infinite number of wavelets to cover the entire spectrum. In order to solve this, Mallat (1989) introduced the *scaling function* which basically works as a low-pass spectrum signal. Taking this property of being a signal with a low-pass spectrum, we can decompose it in wavelet components:

$$\phi(t) = \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} W_{j,k}\psi_{j,k} \tag{25}$$

Now, $\phi(t)$ can be used together with the wavelets to cover all the spectrum. It can be built in such a way that it covers all the frequency spectrum from $-\infty$ up to a given scale j while the rest is covered by the wavelets. The scaling function together with the mother wavelet form what is called a *Multi Resolution Analysis* (MRA) for the Hilbert space $\mathbb{L}^2(\mathbb{R})$.

The MRA, introduced by Mallat (1989) in the wavelet context, is the design method of most of the DWTs. If we consider the wavelet transform as a *filter bank*, then, transforming a signal $f(t)$ through wavelets, would consist on passing it through this filter bank. The wavelets would provide a band-pass filter while the scaling function would work as a low-pass filter. The outputs of the different filter stages would give us the wavelet and scaling transform coefficients.

2.3.3 *The Discrete Wavelet Transform (DWT)*

The DWT introduced by Mallat (1989) is an efficient method to calculate the wavelet transform of a signal. It uses MRA to compute the wavelet coefficients by successively passing the signal through a set of low-pass, $\{h_k\}$, and high-pass, $\{g_k\}$, filters.

In the previous section we built the scaling function in such a way that it would cover the frequency spectrum from $-\infty$ up to a given scale j. Taking into account that the scaling function can be described with wavelets (see equation 25), using multiresolution formulation (Burrus, Gopinath, and Guo, 1998), we can express the scaling function at scale j in terms of the scaling function at scale $j+1$:

$$\phi\left(2^j t\right) = \sum_k h_{j+1}(k)\phi(2^{j+1}t - k) \tag{26}$$

Given that the scaling function at scale j replaced a set of wavelets, we can also express the wavelets in this set in terms of the scaling function at scale $j+1$:

$$\psi\left(2^j t\right) = \sum_k g_{j+1}(k)\phi(2^{j+1}t - k) \tag{27}$$

Also, given that our signal $f(t)$ could be expressed as a combination of dilated and translated wavelets up to scale $j-1$, we can express it in terms of dilated and translated scaling functions at scale j:

$$f(t) = \sum_k \lambda_j(k)\phi(2^j t - k) \tag{28}$$

If in this equation we set up a scale $j-1$, in order to keep the same level of detail, we have to add wavelets so the signal $f(t)$ can be expressed as:

$$f(t) = \sum_k \lambda_{j-1}(k)\phi(2^{j-1}t-k) + \sum_k \gamma_{j-1}(k)\psi(2^{j-1}t-k) \quad (29)$$

Given that the scaling function $\phi_{j,k}$ and the wavelet $\psi_{j,k}$ are orthonormal, the coefficients $\lambda_{j-1}(k)$ and $\gamma_{j-1}(k)$ can be found using the inner products:

$$\begin{aligned} \lambda_{j-1}(k) &= \langle f(t), \phi_{j,k}(t) \rangle \\ \gamma_{j-1}(k) &= \langle f(t), \psi_{j,k}(t) \rangle \end{aligned} \quad (30)$$

By replacing $2^j t$ for t in equations 26 and 27 and substituting in the previous inner products we find:

$$\lambda_{j-1}(t) = \sum_k h(k-2t)\lambda_j(k) \quad (31)$$

$$\gamma_{j-1}(t) = \sum_k g(k-2t)\gamma_j(k) \quad (32)$$

These two equations establish that the wavelet and scaling coefficients on a certain scale can be found by calculating a weighted sum of the scaling coefficients of the previous scale. Since the $\lambda_j(t)$ coefficients come from the low-pass part of the signal spectrum the, $h(t)$ weights must form a low-pass filter. On the other hand, since the $\gamma_j(t)$ coefficients come from the high-pass part of the signal spectrum, then the $g(t)$ coefficients must form a high-pass filter. $\lambda_j(t)$ and $\gamma_j(t)$ are know as *approximation coefficients* and *detail coefficients* respectively (hereafter cA and cD). It is worth mentioning that our signal $f(t)$ would be equal to $\lambda(t)$ at the largest scale.

The low-pass filter is assumed to satisfy the internal orthogonality relation and to have sum of squares equal to one (Nason and Silverman, 1995). The high-pass filter can then be defined as a n-element sequence that satisfies:

$$g(k) = (-1)^k h(N-1-k), \ \forall n \quad (33)$$

with N being the total number of samples in $\lambda(k)$. Filters constructed this way are called *quadrature mirror filters* (Croisier, Esteban, and Galand, 1976).

An alternative formulation for the DWT is to construct an orthogonal matrix A associated to the filters of the wavelet basis of choice. This way, the DWT can be defined as the product of matrix A with the signal vector $\vec{f}(t)$:

$$\text{Matrix of coefficients} = A\vec{f}(t) \tag{34}$$

To summarize, the DWT of a signal is computed by successively passing it through a set of low- and high-pass filters ($h(t)$ and $g(t)$) that produce the *approximation* and *detail coefficients* at each decomposition level. These filters down-sample the signal by a factor two at each level due to the step-size of two in equations 31 and 32. This decomposition and filtering can be repeated to increase the frequency resolution. In the usual approach, the *approximation coefficients*, cA, are decomposed and downsampled while the *detail*, cD, ones remain untouched. This generates a tree known as *Mallat's decomposition tree* representing the DWT filter bank (see Figure 5). This decomposition scheme can be done up to a maximum decomposition level/scale given by $\log_2 N$ where N is the length of the signal.

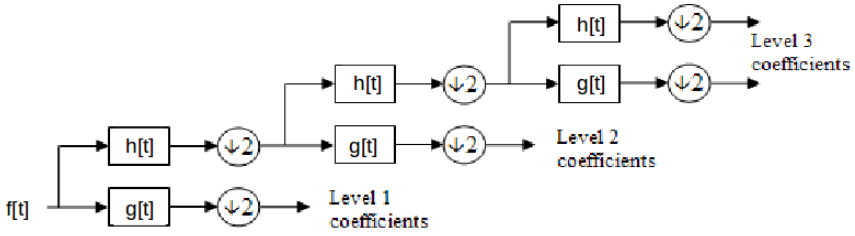


Figure 5: DWT filter bank example.

In this case, equations 31 and 32 would become:

$$cA_{j-1}(t) = \sum_k h(k-2t)cA_j(k) \tag{35}$$

$$cD_{j-1}(t) = \sum_k g(k-2t)cA_j(k) \tag{36}$$

The original signal can then be recovered through the Inverse Discrete Wavelet Transform (IDWT). Starting from the approximation and detail coefficients at scale j, the IDWT reconstructs $cA_{j-1}(t)$ inverting the decomposition step by inserting zeros and convolving the results with the reconstruction filters.

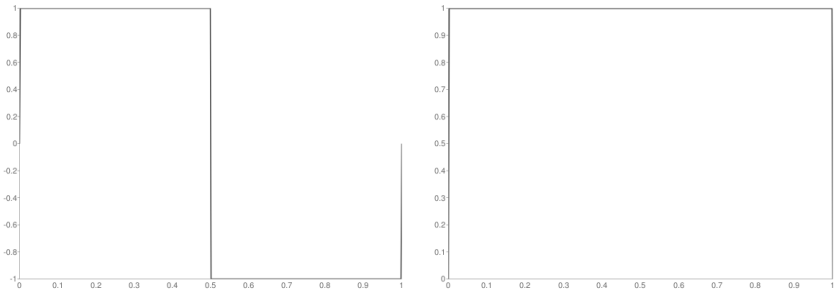
2.3.4 *Example of the DWT*

In Section 2.1.2 we have already introduced the first and simplest of all wavelet basis, the Haar wavelet (Haar, 1910):

$$\psi(t) = \begin{cases} 1 & \text{if } 0 \leqslant t < 1/2 \\ -1 & \text{if } 1/2 \leqslant t < 1 \\ 0 & \text{otherwise} \end{cases} \tag{37}$$

Its scaling function is given by:

$$\phi(t) = \begin{cases} 1 & \text{if } 0 \leqslant t < 1 \\ 0 & \text{otherwise} \end{cases} \tag{38}$$



(a) (b)

Figure 6: a) Haar Mother Wavelet and b) Scaling Function. Credit: PyWavelets

The family of filters for the Haar wavelet is:

$$\begin{aligned} \{h\} &= (h(0), h(1)) = \left(1/\sqrt{2}, 1/\sqrt{2}\right) \text{ and } h(i) = 0 \; \forall i \neq \{0, 1\} \\ \{g\} &= (g(0), g(1)) = \left(-1/\sqrt{2}, 1/\sqrt{2}\right) \text{ and } g(i) = 0 \; \forall i \neq \{0, 1\} \end{aligned} \tag{39}$$

Consider a time series of length N = 8:

$$S = (6, 8, 32, 6, 10, 4, 5, 21)$$

Taking into account that the number of maximum decomposition levels is given by $\log_2 N$, S can be decomposed up to level 3. Using equations 31, 32 and 39 the DWT of the first decomposition level is given by:

$$cA_2 = \begin{cases} cA_{2,0} = 6 \times 1/\sqrt{2} + 8 \times 1/\sqrt{2} = 7\sqrt{2} \\ cA_{2,1} = 32 \times 1/\sqrt{2} + 6 \times 1/\sqrt{2} = 19\sqrt{2} \\ cA_{2,2} = 10 \times 1/\sqrt{2} + 4 \times 1/\sqrt{2} = 7\sqrt{2} \\ cA_{2,3} = 5 \times 1/\sqrt{2} + 21 \times 1/\sqrt{2} = 13\sqrt{2} \end{cases}$$

$$cD_2 = \begin{cases} cD_{2,0} = -6 \times 1/\sqrt{2} + 8 \times 1/\sqrt{2} = -\sqrt{2} \\ cD_{2,1} = -32 \times 1/\sqrt{2} + 6 \times 1/\sqrt{2} = 13\sqrt{2} \\ cD_{2,2} = -10 \times 1/\sqrt{2} + 4 \times 1/\sqrt{2} = 3\sqrt{2} \\ cD_{2,3} = -5 \times 1/\sqrt{2} + 21 \times 1/\sqrt{2} = -8\sqrt{2} \end{cases}$$

Where $cA_{j-1,t} \equiv \lambda_{j-1}(t)$ and $cD_{j-1,t} \equiv \gamma_{j-1}(t)$. Just for clarification purposes, the matrix representation of the above is given by equation 34 as follows:

$$\begin{pmatrix} cA_{2,0} \\ cA_{2,1} \\ cA_{2,2} \\ cA_{2,3} \\ cD_{2,0} \\ cD_{2,1} \\ cD_{2,2} \\ cD_{2,3} \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \\ 32 \\ 6 \\ 10 \\ 4 \\ 5 \\ 21 \end{pmatrix}$$

Using the filter bank scheme shown in Figure 5, the *approximation* and *details coefficients* of the next decomposition level are given by the DWT of cA_2:

$$cA_1 = \begin{cases} cA_{1,0} &= 7\sqrt{2} \times 1/\sqrt{2} + 19\sqrt{2} \times 1/\sqrt{2} = 26 \\ cA_{1,1} &= 7\sqrt{2} \times 1/\sqrt{2} + 13\sqrt{2} \times 1/\sqrt{2} = 20 \end{cases}$$

$$cD_1 = \begin{cases} cD_{1,0} &= -7\sqrt{2} \times 1/\sqrt{2} + 19\sqrt{2} \times 1/\sqrt{2} = -12 \\ cD_{1,1} &= -7\sqrt{2} \times 1/\sqrt{2} + 13\sqrt{2} \times 1/\sqrt{2} = -6 \end{cases}$$

We can apply the Haar wavelet function one last time to cA_1 to obtain the coefficients for the last decomposition level:

$$cA_0 = \{cA_{0,0} = 26 \times 1/\sqrt{2} + 20 \times 1/\sqrt{2} = 23\sqrt{2}$$
$$cD_0 = \{cD_{0,1} = -26 \times 1/\sqrt{2} + 20 \times 1/\sqrt{2} = 6\sqrt{2}$$

Given the approximation coefficient of the last level, cA_2, the detail coefficients of every level and the wavelet function, the

original signal can be easily reconstructed using the inverse Haar wavelet transform through its reconstruction filters:

$$\begin{aligned}\{h_r\} &= (h_0, h_1) = \left(1/\sqrt{2}, 1/\sqrt{2}\right) \\ \{g_r\} &= (g_0, g_1) = \left(1/\sqrt{2}, -1/\sqrt{2}\right)\end{aligned} \tag{40}$$

2.4 THE STATIONARY WAVELET TRANSFORM (SWT)

The DWT is not a time-invariant transform and is therefore very sensitive to the alignment of the signal in time. Due to the fact that the DWT downsamples the signal at each decomposition level, a shift in the input signal might not necessarily manifest itself as an equivalent shift in the DWT coefficients at all levels. The SWT also know as Undecimated Wavelet Transform or *à trous* algorithm (Holschneider et al., 1989; Shensa, 1992) is a wavelet transform algorithm designed to overcome the lack of translation-invariance of the DWT. As with the DWT, high and low pass filters are applied to the signal at each level. However, in order to achieve translation-invariance, the downsamplers and upsamplers are removed and the filter coefficients are upsampled by a factor of $2^{(j-1)}$ in the jth decomposition level. By padding the filters at each level with zeros, the two new sequences at each level have the same length as the original time series. Following the notation of Nason and Silverman (1995), the padding is done in the following way.

- Given a sequence $S = (x_1, x_2, ...x_i, ..., x_{2^J})$, let $\mathcal{Z}$ be the operator that alternates that sequence with zeroes so that: $\forall i \in \mathbb{Z}, (\mathcal{Z}x)_{2i} = x_i$ and $(\mathcal{Z}x)_{2i+1} = 0$.
- Define a new set of filters $\mathcal{H}^{[j]}$ and $\mathcal{G}^{[j]}$ with weights $\mathcal{Z}^{[j]}h$ and $\mathcal{Z}^{[j]}g$ respectively.
- Thus the filter $\mathcal{H}^{[j]}$ has weights $h^{[j]}_{2^j i} = h_i$ and $h^{[j]}_k = 0$ if k is not a multiple of 2^j. The filter $\mathcal{H}^{[j]}$ is obtained by inserting a zero between each adjacent pair of elements of the filter $\mathcal{H}^{[j-1]}$, and similarly for $\mathcal{G}^{[j]}$.

This way, considering a filter bank as in Figure 5, equations 31 and 32 would become:

$$\lambda_{j-1} = \mathcal{H}^{[J-j]}\lambda_j \tag{41}$$

$$\gamma_{j-1} = \mathcal{G}^{[J-j]}\lambda_j \tag{42}$$

With this new set of filters we are able to not decimate the data resulting in the two new sequences at each level having the same length as the original signal. The $\mathcal{H}^{[j]}$ and $\mathcal{G}^{[j]}$ however, do not have an orthogonality property, so to obtain it, two extra decimation operators are needed:

$$\begin{aligned}(\mathcal{D}_0 x)_i &= x_{2i} \\ (\mathcal{D}_1 x)_i &= x_{2i+1}\end{aligned} \tag{43}$$

Applying $\left(\mathcal{D}_0\mathcal{H}^{(J-j)}, \mathcal{D}_0\mathcal{G}^{(J-j)}\right)$ and $\left(\mathcal{D}_1\mathcal{H}^{(J-j)}, \mathcal{D}_1\mathcal{G}^{(J-j)}\right)$ to the signal, we obtain the even and odd elements of the wavelet coefficients λ_{j-1} and γ_{j-1}, respectively through the following process (Nason and Silverman, 1995):

$$\begin{aligned}\left(\lambda_{j-1}\right)_{2i} &= \mathcal{D}_0\mathcal{H}^{(J-j)}\lambda_j \\ \left(\lambda_{j-1}\right)_{2i+1} &= \mathcal{D}_1\mathcal{H}^{(J-j)}\lambda_j \\ \left(\gamma_{j-1}\right)_{2i} &= \mathcal{D}_0\mathcal{G}^{(J-j)}\lambda_j \\ \left(\gamma_{j-1}\right)_{2i+1} &= \mathcal{D}_1\mathcal{G}^{(J-j)}\lambda_j\end{aligned} \tag{44}$$

By applying the filter twice at the even and odd position for each decomposition level, shifts do not affect the transformation results, meaning that the SWT is time-invariant.

2.4.1 *The Inverse Stationary Wavelet Transform (ISWT)*

Basically, the SWT can be understood as an overdetermined representation of the original signal compared to the DWT one (Nason and Silverman, 1995).

The first approach to compute the ISWT takes into account that the coefficients of the DWT are also present in the SWT ones. The idea is to build some sequence of zeros and ones that allows to select these coefficients and then use them to perform the reconstruction. This approach is called ϵ-*basis inverse* of the SWT (Nason and Silverman, 1995).

The second approach, called *average basis inverse*, consists on finding the ϵ-*basis inverse* for each ϵ and average the result. For a 2^J-length series, there are 2^J possible values of ϵ. For any particular scale j, we have seen in the previous section that the decimated transforms $\left(\mathcal{D}_0\mathcal{H}^{(j)}, \mathcal{D}_0\mathcal{G}^{(j)}\right)$ are each orthogonal

and that allow us to obtain the odd and even elements of the wavelet coefficients λ_{j-1} and γ_{j-1}. Following Nason and Silverman notation, we can denote the inverses of these two transformations by $\mathcal{R}_0^{(j)}$ and $\mathcal{R}_1^{(j)}$ respectively. Thus, we can define a reconstruction operator as:

$$\bar{\mathcal{R}}^{(j)} = 1/2\left(\mathcal{R}_0^{(j)} + \mathcal{R}_1^{(j)})\right) \tag{45}$$

The original data can be then obtained by recursively evaluating:

$$\lambda_j = \bar{\mathcal{R}}^{(J-j)}(\lambda_{j-1}, \gamma_{j-1}) \tag{46}$$

2.4.2 *An example of the SWT signal characterization capabilities*

An example of a MRA using the SWT is shown in Figure 7. Our test signal is a combination of two sinusoidal signals, one of high frequency and another of low frequency. They both have the same amplitude and are affected by Gaussian random noise. As can be seen in Figure 7, SWT is able to decouple each of these signal contributions at different decomposition levels. High-frequency noise has been separated from the sinusoidal signals, appearing as an extra contribution to the lower decomposition levels. In this example, noise can be fairly represented by the combination of the first two decomposition levels (j=1 and j=2). The high frequency sinusoidal signal can be recovered by summing the third and fourth decomposition levels while the low frequency one would be recovered through the combination of the remaining levels. Crucially, the wavelet decomposition allows the signal characteristics to change through the time series, while retrieving both time and frequency information from the original signal. As we have already mentioned, other techniques such as the discrete Fourier Transform convert data from time into frequency domain but in doing so, time information is lost.

2.5 BIORTHOGONAL WAVELET BASES

Except for the particular case of the Haar wavelet basis (Haar, 1910), it is impossible to construct symmetric orthogonal wavelets. However, biorthogonal formulation allows to build smooth symmetric or antisymmetric wavelet basis of compact support (Cohen, Daubechies, and Feauveau, 1992).

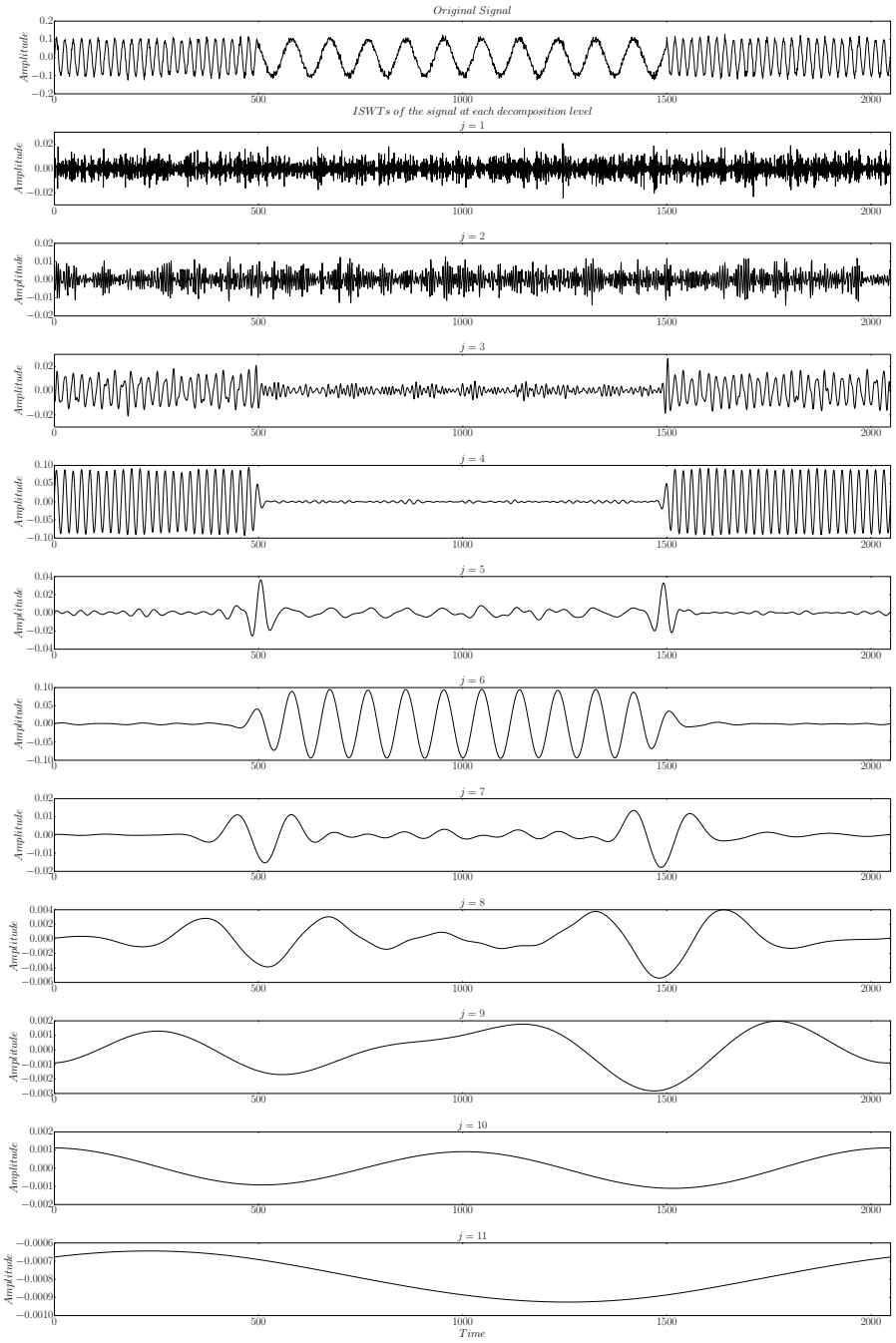


Figure 7: SWT decomposition of a test signal. The uppermost plot shows the test signal as described in Section 2.4.2, consisting of two sinusoidal signals and Gaussian random noise. The following plots represent the ISWT transform of the signal at each SWT decomposition level (j = 1,...,11).

2.5.1 *Biorthogonal Bases of* $\mathbb{L}^2(\mathbb{Z})$

The decomposition of a signal in a *filter bank* can be interpreted as an expansion in a basis of $\mathbb{L}^2(\mathbb{Z})$. Equations 35 and 36 can be rewritten as inner products in this basis:

$$cA_{j-1}(t) = \sum_k h(k-2t)cA_j(k) = \langle h(k-2t), cA_j(k)\rangle \quad (47)$$

$$cD_{j-1}(t) = \sum_k g(k-2t)cA_j(k) = \langle g(k-2t), cA_j(k)\rangle \quad (48)$$

The signal recovered using the reconstruction filters can be expressed as:

$$cA_j = \sum_t cA_{j-1}\tilde{h}(k-2t) + \sum_t cD_{j-1}\tilde{g}(k-2t) \quad (49)$$

Given the filters h, g, $\tilde{h}$ and $\tilde{g}$, if their Fourier transforms are bounded, then the families $\left\{\tilde{h}(k-2t), \tilde{g}(k-2t)\right\}_{t\in\mathbb{Z}}$ and $\{h(k-2t), g(k-2t)\}_{t\in\mathbb{Z}}$ are biorthogonal bases of $\mathbb{L}^2(\mathbb{Z})$ (Mallat, 2008).

2.5.2 *Construction of Biorthogonal Wavelet Bases*

A set of filters (h, g) and $(\tilde{h}, \tilde{g})$ yield two scaling and two wavelet functions that satisfy (Cohen, Daubechies, and Feauveau, 1992):

$$\phi(t) = \sqrt{2}\sum_k h(k)\phi(2t-k),\ \tilde{\phi}(t) = \sqrt{2}\sum_k \tilde{h}(k)\tilde{\phi}(2t-k) \quad (50)$$

$$\psi(t) = \sqrt{2}\sum_k g(k)\phi(2t-k),\ \tilde{\psi}(t) = \sqrt{2}\sum_k \tilde{g}(k)\tilde{\phi}(2t-k) \quad (51)$$

Where ϕ and $\tilde{\phi}$ satisfy biorthogonal relations:

$$\langle \phi(t), \tilde{\phi}(t-k)\rangle = \delta(k) \quad (52)$$

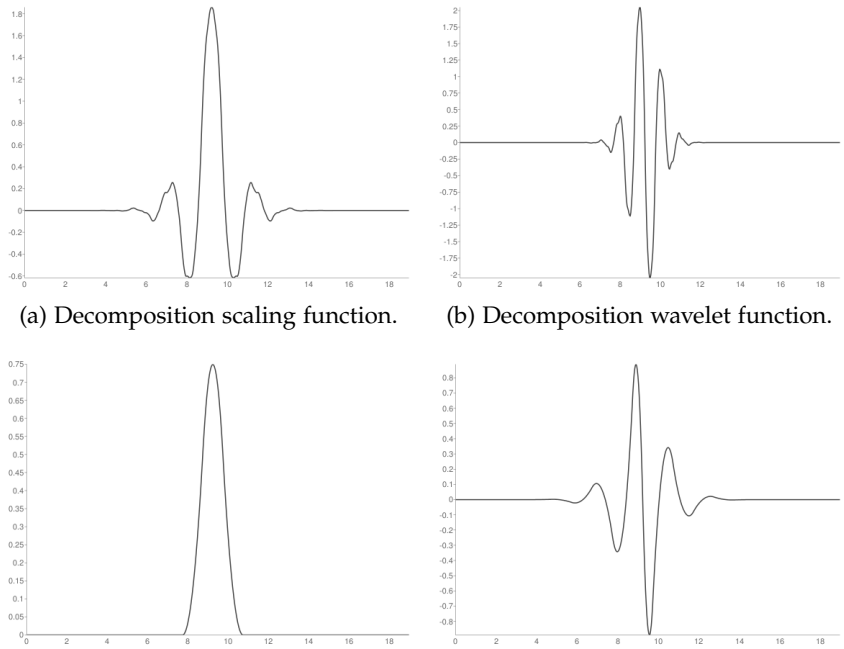
and the two wavelet families $\{\psi_{j,k}\}_{(j,k)\in\mathbb{Z}^2}$, $\{\tilde{\psi}_{j,k}\}_{(j,k)\in\mathbb{Z}^2}$ are biorthogonal bases of $\mathbb{L}^2(\mathbb{Z})$ that satisfy the biorthogonality condition:

$$\langle \psi_{j,k}, \tilde{\psi}_{j',k'}\rangle = \delta(k-k')\delta(j-j') \quad (53)$$

Any function $f \in \mathbb{L}^2(\mathbb{Z})$ has two possible decomposition in these bases:

$$f = \sum_k \sum_j \langle f, \psi_{j,k} \rangle \tilde{\psi}_{j',k'} = \sum_k \sum_j \langle f, \tilde{\psi}_{j,k} \rangle \psi_{j,k} \tag{54}$$

The number of vanishing moments, the regularity, the wavelet ordering and the symmetry of the biorthogonal wavelets is controlled with an appropriate design of the filters. Since ψ and $\tilde{\psi}$ might not have the same regularity and number of vanishing moments, the two formulas in equation 54 are not equivalent. The decomposition using the left hand side equation is obtained with filters (h, g) and the reconstruction with $(\tilde{h}, \tilde{g})$. The right hand side one corresponds to the opposite case. Usually the decomposition should be done using the wavelet with the maximum number of vanishing moments while the reconstruction should be done with the other one (Mallat, 2008). An example of a biorthogonal wavelet basis is shown in Figure 8.



(a) Decomposition scaling function.

(b) Decomposition wavelet function.

(c) Reconstruction scaling function.

(d) Reconstruction wavelet function.

Figure 8: Biorthogonal 3.9 mother wavelet and scaling functions. Credit: PyWavelets

2.6 WAVELET POWER SPECTRUM

The wavelet transform converts signal f into a series of wavelet coefficients $W_j(t)$. Each of these coefficients represents the am-

plitude of the wavelet function at a particular location within the signal at a particular wavelet scale. A useful way to determine the distribution of energy within the signal is to compute the Wavelet Power Spectrum (WPS). The local WPS at a particular decomposition level j is calculated by summing up the squares of wavelet coefficients for that level (Torrence and Compo, 1998):

$$\mathrm{WPS}(j) = \sum_{t} \mid W_j(t) \mid^2 \tag{55}$$

Following Liu, San Liang, and Weisberg, 2007 we can define a more robust power spectrum taking into account a correcting factor related to the scale j:

$$\mathrm{WPS}(j) = \sum_{t} 2^j \mid W_j(t) \mid^2 \tag{56}$$

Two sine waves, with 0.08 and 0.1 amplitudes and periods of 1 and 10 days, respectively, are summed together to form a simulated signal as shown in Figure 9 a). The sample interval is 512 seconds, and the length of the signal is 8192 data points (approximately 50 days). Figure 9 b) and c) show the power spectrum map (squared wavelet coefficients at each decomposition level) and local power spectrum, respectively, of the CWT of the signal using the order 6 Morlet mother wavelet. In the case of the CWT we can easily compute a relationship between the scale and the corresponding frequency/period. Following the method of Meyers, Kelly, and O'Brien (1993), the relationship between the equivalent Fourier period and the wavelet scale can be derived analytically for a particular wavelet function, ψ, by substituting a cosine wave of a known frequency into the convolution of the mother wavelet with the DFT (Torrence and Compo, 1998). As can be seen in Figure 9 c), the peaks of the power spectrum allow to recover the period of both sine waves composing our signal. The use of the CWT power spectrum as a way to detect periodicities and study other astrophysical properties such as pulsation and rotation has already been studied by de Freitas et al., 2010 and Bravo et al. (2014).

Figure 9 d) and c) show the power spectrum map and local power spectrum for the SWT of the signal using the Biorthogonal 3.9 mother wavelet. As in the case of the CWT we recover two peaks at different scales corresponding to the low and high frequency sine waves. In this case though, the relationship between the scale and the frequency cannot be so

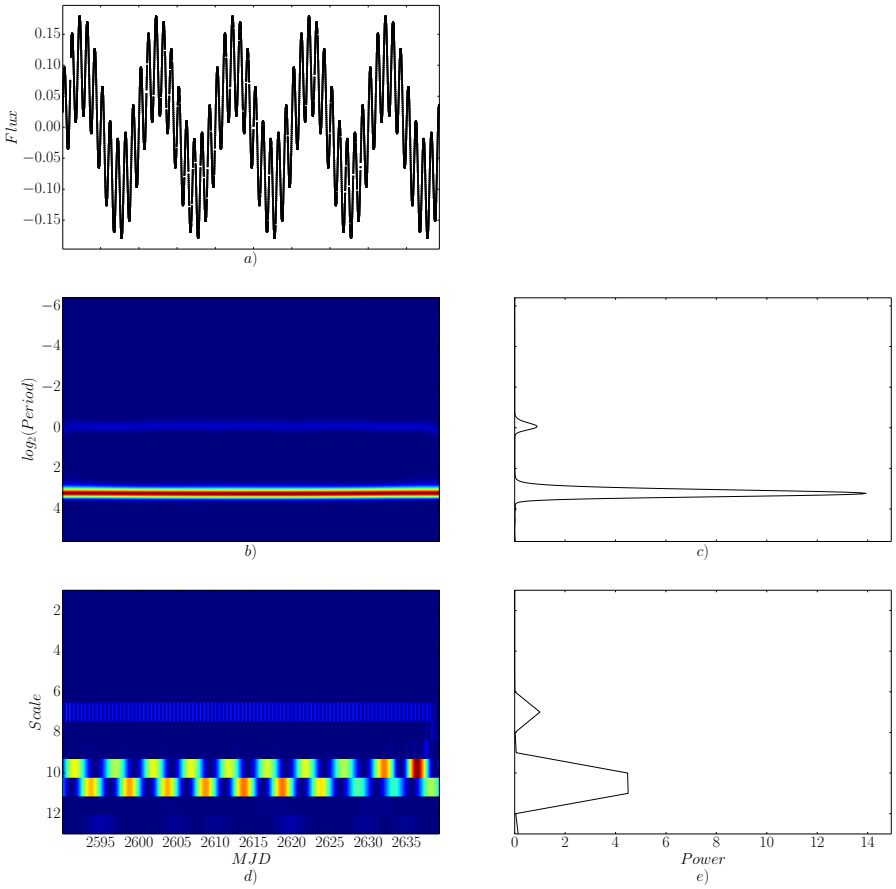


Figure 9: a) Original signal. b) CWT power spectrum map. c) CWT local power spectrum. d) SWT power spectrum map. e) SWT local power spectrum. Morlet wavelet of order 6 was used to compute the CWT power spectrum and Biorthogonal 3.9 wavelet to compute the SWT one.

easily computed as the signal energy can be distributed among different scales (as seen in Figure 9 d)). However, the fact that the SWT power spectra presents peaks will be part of the signal detection method used in Chapter 4.

Part II

WAVELETS AND NOISE: THE TFAW ALGORITHM

3 WAVELET-BASED DENOISING

3.1 INTRODUCTION

The limiting factor for most planetary transit surveys, specially for ground-based ones, is the presence of correlated noise (systematics) resulting from non-astrophysical sources (Pont, Zucker, and Queloz, 2006) more than pure Gaussian noise. The precision that can be obtained with a ground-based photometric survey depends on several factors such as the atmosphere conditions, the instrument, the number of non-distorted stellar PSFs that can be imaged on the CCD, the reduction pipeline and the photometric errors of the measurements. Detectors with wide Field-of-view (FoV) often suffer from additional issues such as under-sampled PSFs, atmospheric extinction, source blending and optical distortion at the edges of the detector. On the other hand, space-based surveys such as CoRoT or Kepler benefit from their stable and non-atmosphere affected environments. However, their on-board instruments are not perfect and their light curves are affected by effects that hinder transit detection. Some of these effects go from spacecraft jitter and maneuvers, CCD long-term aging or residuals coming from the orbital periods of the spacecraft or sudden changes in magnitude due to high energetic particle fluxes. These and other effects increase the noise of the light curves and also introduce systematic variations that decrease the detection probability of any periodic signal in the data.

3.2 WAVELET DENOISING BY THRESHOLDING

The idea of signal denoising using wavelets was first introduced by Donoho and Johnstone (1994b). If a signal is well represented by a small number of wavelet dimensions/scales, its associated coefficients will be relatively large compared to any other signal or noise that has its energy spread over several coefficients. This allows to threshold the wavelet transform coefficients removing the noise or any undesired signal in the wavelet domain. The inverse wavelet transform would then retrieve the original signal with minimal loss of detail.

Assume a signal y_i with additive noise of the form:

$$y_i = s_i + n_i, \quad i = 1, ..., n \tag{57}$$

where s_i is the signal of interest and n_i represents independent and identically distributed zero mean, white Gaussian noise with standard deviation σ. Using Donoho and Johnstone's notation, we can rewrite the previous equation in the wavelet transform domain as:

$$y_i[\mathcal{B}] = s_i[\mathcal{B}] + n_i[\mathcal{B}] \tag{58}$$

where the suffix $[\mathcal{B}]$ denotes the transform of the indicated vector into the wavelet orthonormal basis.

The signal $s_i[\mathcal{B}]$ can be estimated by transforming the noisy data $y_i[\mathcal{B}]$ with a given operator D:

$$\hat{s}[\mathcal{B}] = Dy_i[\mathcal{B}] \tag{59}$$

where the mean-squared error or *risk* of estimator $\hat{s}$ is given by:

$$R(\hat{s}, s) = E \parallel \hat{s} - s \parallel_2^2 \tag{60}$$

Consider a diagonal estimator in the basis $\mathcal{B}$:

$$D = diag(\delta_1, ..., \delta_N), \quad \delta_i \in \{0, 1\} \tag{61}$$

thus, the estimate $\hat{s}$ is obtained by keeping or zeroing the individual wavelet coefficients.

To summarize, the scheme Donoho and Johnstone proposes to denoise the signal basically consists on computing the wavelet transform of signal y_i, perform some kind of thresholding in the wavelet domain and then compute the inverse wavelet transform to retrieve the denoised estimate $\hat{s}$ of the signal.

3.2.1 *Thresholding Estimation*

In an orthogonal basis $\mathcal{B} = g_i$, equation 59 can be rewritten as (Mallat, 2008):

$$\hat{s} = Dy_i[\mathcal{B}] = \sum_{i=0}^{N-1} a_i(y_i[\mathcal{B}])y_i[\mathcal{B}]g_i \tag{62}$$

A **hard-thresholding** estimator can be implemented using:

$$\hat{s} = \sum_{i=0}^{N-1} \rho_T(y_i[\mathcal{B}])g_i \tag{63}$$

where the hard-thresholding function ρ_t is defined by:

$$\rho_T(x) = \begin{cases} x & \text{if } |x| > T \\ 0 & \text{if } |x| \leqslant T \end{cases} \tag{64}$$

where T is the threshold of choice.

In the case of **soft-thresholding**, the amplitude of all the coefficients is slightly decreased in order to reduce the added noise. The estimator can be written again as in equation 63 but, in the new thresholding function is given by:

$$\rho_T(x) = \begin{cases} x - T & \text{if } x \geqslant T \\ x + T & \text{if } x \leqslant -T \\ 0 & \text{if } |x| \leqslant T \end{cases} \tag{65}$$

Although the previously mentioned thresholding functions are the most commonly used, there have been several efforts to create improved ones such as the *Non-negative Garrote thresholding* (Breiman, 1995). Or the ones taking into account not only Gaussian noise but also correlated noise (Johnstone and Silverman, 1997).

3.2.2 *Threshold Selection Criteria*

The threshold T is generally chosen so that it is just above the maximum level of the wavelet coefficients associated to noise $n_i[\mathcal{B}]$. Given that a wavelet basis has been selected in which large coefficients are created by the signal s_i, reducing the amplitude of all noise coefficients by T ensures that the estimation restores a signal that is at least as regular as the original (Mallat, 2008).

Several threshold selection criteria have been developed, here we are going to briefly discuss some of the most widely used ones. For a more detailed review of these criteria see Mallat (2008) and Abramovich, Bailey, and Sapatinas (2000).

- **Universal Threshold**: proposed by Donoho and Johnstone (1994b). Given that $n_i[\mathcal{B}]$ is a vector of N independent

Gaussian random variables of variance σ^2, it can be proved that the maximum amplitude of the noise has a high probability of being below:

$$T = \sigma\sqrt{2\log_e N} \tag{66}$$

To improve the finite sample properties of the universal threshold, Donoho and Johnstone (1994b) suggested that the wavelet approximation coefficients of the first levels should be kept as they are even if they do not pass the threshold. For example, Hall and Patil (1996) and Efromovich (1999) suggest to start thresholding from the scale given by $\log_2(n)/(2r+1)$ where r is the regularity of the mother wavelet.

- **SURE**: The thresholding risk is often reduced by using a threshold smaller than the one given by the *Universal threshold*. A threshold adapted to the data can be calculated by minimizing an estimation of the risk. The *SURE* threshold chooser was proposed by Donoho and Johnstone (1995a) based on Stein's Unbiased Risk Estimate (Stein, 1981) and called *SureShrink* (Donoho and Johnstone, 1995b). The SURE threshold is the one that minimizes the SURE estimator:

$$Sure(y_i[\mathcal{B}], T) = \sum_{i=0}^{N-1} C(y_i[\mathcal{B}]) \tag{67}$$

with:

$$C(x) = \begin{cases} x^2 - \sigma^2 & \text{if } x \leqslant T \\ \sigma^2 + T^2 & \text{if } x > T \end{cases} \tag{68}$$

In order to efficiently remove all the noise in the cases when the signal energy is small relative to the noise energy, one must impose the *Universal threshold*. The resulting SURE threshold is given by:

$$T = \begin{cases} \sigma\sqrt{2\log_e N} & \text{if } \| y_i[\mathcal{B}] \|^2 - N\sigma^2 \leqslant \epsilon_N \\ argminSure(y_i[\mathcal{B}], T) & \text{if } \| y_i[\mathcal{B}] \|^2 - N\sigma^2 > \epsilon_N \end{cases}$$

(69)

Where ϵ_N is given by:

$$\epsilon_N = \sigma^2 N^{1/2} (\log_e N)^{3/2} \tag{70}$$

At each decomposition scale j, *SureShrink* computes a SURE threshold T_j using equation 69. A soft-thresholding is then applied at each scale using the previously computed threshold.

- **Cross-validation approaches**: these methods consist on minimizing the prediction error generated by comparing a prediction based on a subset of the data and compare it with the rest of the data in order to find an optimal threshold value (see Nason (1996) and Weyrich and Warhola (1995)).
- **Bayesian approaches**: various Bayesian thresholding approaches have been proposed. In the Bayesian approach, a prior distribution is imposed on the wavelet coefficients. Then the function is estimated by applying a suitable Bayesian rule to the resulting wavelet coefficients. Different choices of loss function lead to different Bayesian rules and thus to different thresholding solutions. See Abramovich, Bailey, and Sapatinas (2000) for a comprehensive review and reference list of different Bayesian approaches.

3.2.3 *Noise Variance Estimation*

The noise level σ needed by all the previously listed thresholding methods is not usually a priori known and has to be estimated from the data. To estimate the variance of the noise n_i one needs to suppress the influence of the signal s_i. Donoho and Johnstone (1994b) state that a robust estimator can be calculated from the Median Absolute Deviation (MAD) of the J finest-scale wavelet coefficients divided by 0.6745:

$$MAD = median(|\, cD_{J,i} - median(cD_J)) \,|)/0.6475 \tag{71}$$

the MAD is a robust statistic, more resilient to outliers than the standard deviation in which the distances from the mean are squared leading to outliers being weighted more heavily. The constant factor standardizes the MAD in order to make it consistent with the standard deviation of a normal distribution.

3.2.4 *Performance of Wavelet Thresholding*

Although simple, the thresholding of the wavelet coefficients can introduce distortions in the filtered signal. For example, thresholding DWT coefficients can create small ripples or oscillations around discontinuities due to the decimation related to orthogonal wavelet transforms (Mallat, 2008). One way of diminishing these artifacts is to use the *à trous* algorithm or SWT which does not decimate the data (Mallat, 2008; Starck, Murtagh, and Bijaoui, 1998). Moreover, in order to avoid the modification of the signal of interest, both the threshold value and the thresholding function have to be carefully selected. Neglecting this could cause the signal to retain some noise features or have some shape distortion or discontinuities introduced. As stated before, Donoho and Johnstone suggest that not all the wavelet transform coefficients should be thresholded and that some decomposition levels should be kept as they are. The correct selection of the wavelet family could also diminish the chance of perturbing the signal once the thresholding has been applied.

Figure 10 shows an example of the effects caused by a poorly selected threshold when applying soft-thresholding. Using `batman`[1] (Kreidberg, 2015) we simulate a Mandel & Agol planetary transit model (Mandel and Agol, 2002) to which a combination of Gaussian and one point to the next correlated noise has been added. Figure 10 b) is obtained by a soft-thresholding of the SWT coefficients (Biorthogonal 3.9 mother wavelet) using a global threshold value computed through equation 66 (following Mallat (2008) criteria) estimating σ at the finest scale. As can be seen, the noise contribution has been diminished and the shape of the signal has been kept almost undisturbed. However, the depth of the signal has slightly decreased and the wings coming from the convolution of the reconstruction wavelet function with the signal start to appear in the ingress and egress of the transit. Figure 10 c) also uses a soft-thresholding but, with a threshold two times the one given by equation 66. In this case, the noise contribution has completely disappeared but the shape and depth of the signal have been clearly modified. It has to be said, that it is possible to use more complex methods (i.e. SURE estimator, scale-dependent threshold, ...) to better compute an almost optimal threshold value. The example in Figure 10 just wants to show how a bad threshold estimation can lead to distorted signals and the appearance of unwanted artifacts.

1 https://www.cfa.harvard.edu/~lkreidberg/batman

In this work we follow a different approach. The algorithm presented in Chapter 4 combines the Trend Filtering Algorithm (TFA) (Kovács, Bakos, and Noyes, 2005) detrending and systematics removal capabilities with the wavelet transform's signal decoupling and denoising potential. However, in this case, instead of filtering the noise through the thresholding of the wavelet coefficients, the SWT is used to estimate the noise- and trend-free signal as well as the noise contribution. The former will be used to reconstruct the signal while the latter will be removed during TFA's signal reconstruction iterative process.

3.3 SIGNAL AND NOISE CHARACTERIZATION USING THE SWT

Section 2.4.2 showed an example of the SWT applied to a combination of a high-frequency sinusoidal plus a low-frequency one affected by random Gaussian noise (see Figure 7). At each decomposition level we set all the *detail coefficients*, cD, and *approximation coefficients*, cA, computed through the SWT of the original signal, to zero except the ones corresponding to that level. Then, by computing the ISWT at each decomposition level, one could separate the contributions of both sine waves from the high-frequency noise. Using the SWT decoupling capabilities, we could then define some kind of threshold level that allows to separate those scales corresponding to the noise from those representing the target signal.

Following the previous idea, we define two thresholds: *noise level* and *signal level*. The objective is to obtain the best noise estimation by summing the ISWTs from the lowest/finest decomposition level up to *noise level*. The same line of thought would allow to recover the underlying noise-free signal by summing the ISWTs from *signal level* to the highest decomposition level. The maximum number of allowed SWT decomposition levels for any signal is given by:

$$max\ level = \lfloor \frac{\log(\frac{n}{l-1})}{\log 2} \rfloor, \tag{72}$$

where n is the signal length and l is the wavelet filter length.

3.3.1 *Signal and noise level selection criteria*

The selection of the *signal level* and *noise level* thresholds must be carefully performed. As described above, high-frequency

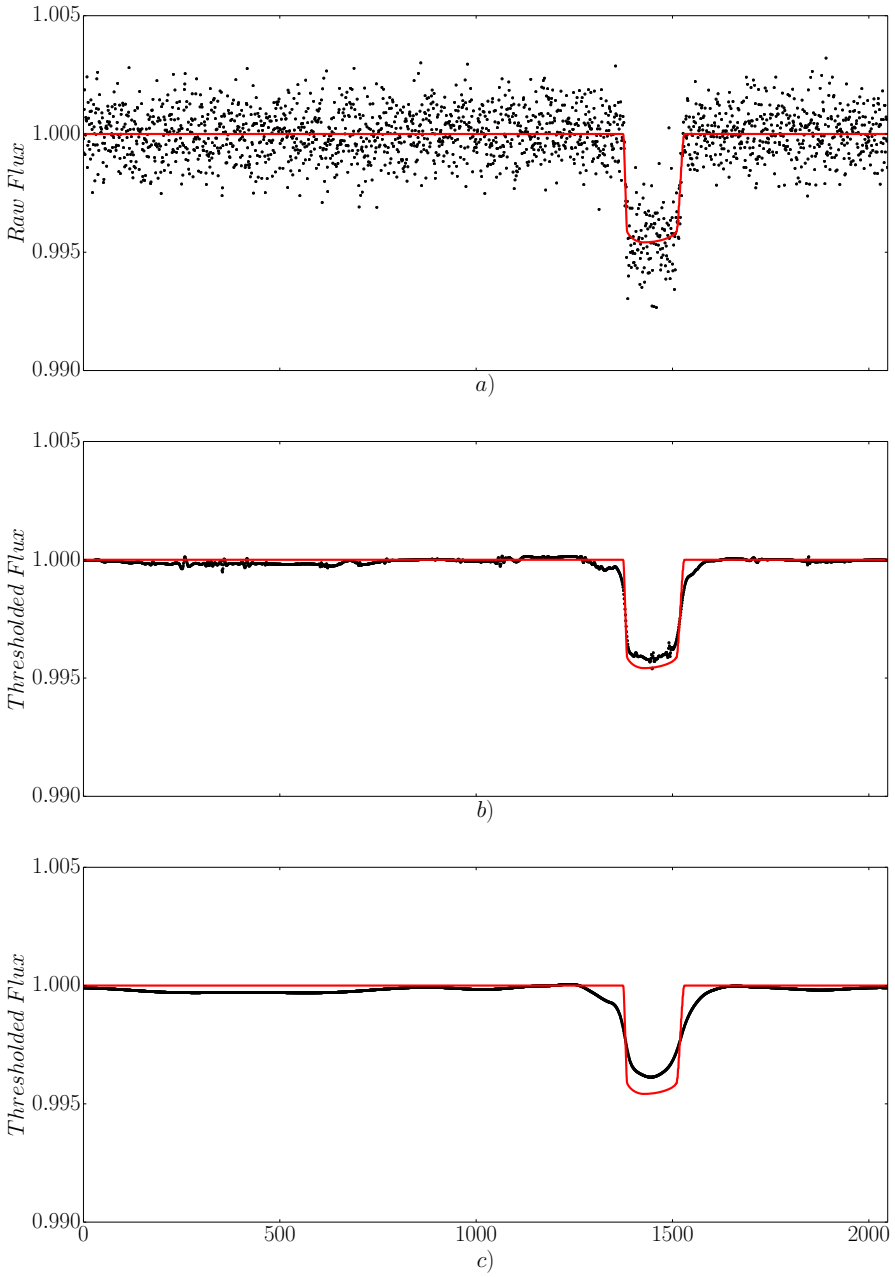


Figure 10: Wavelet soft-thresholding comparison for a simulated planetary transit (the red line in all plots corresponds to the noise-free simulated transit). a) Original signal. b) Soft-thresholded signal with fair threshold estimation. c) Soft-thresholded signal with bad threshold estimation.

noise appears as an extra contribution to the lower SWT decomposition levels. However, depending on the number of levels in which the original signal can be decomposed, some high-frequency signal contributions hidden within the data (i.e. second order sinusoidal modulations, etc...) could be total or partially present in the levels in which the noise is also present. One way of solving this issue would be to increase the number of data points to a higher order of two in order to increase the number of allowed decomposition levels. The greater the number of decomposition levels, the easier it is to separate the noise contribution from any high-frequency signals. In most of the cases, just setting the *noise level* to the finest decomposition level is enough to give a good estimation of the noise.

Regarding the *signal level*, it would suffice to take the sum of the ISWTs from *noise level* to the lowest one. However, as it has been stated before, depending on the selection of the *noise level* there might still be some noise contribution in the final signal. We could then set the *signal level* to, for example, a value around half of the maximum SWT decomposition level value in order to ensure the removal of most of the noise contributions but also to diminish both the chance of modifying the shape or removing part of the target signal.

The correct selection of the *signal level* will be critical when applying the wavelet-based modification of the TFA (see Chapter 4). Instead of using a global *signal level* for all the signals (light curves) in our sample, we can define it for each individual time series using the WPS defined in Section 2.6. We know that the WPS determines the distribution of energy within the signal and that, using the CWT, it allows to determine the periods of any significant signal within the time series (see Figure 9 b) and c)). In the case of the SWT, it allows to determine the decomposition levels at which the correlation between the signal and the wavelet transform at that scale is higher. We can use this to our advantage in order to determine the *signal level*.

Figure 11 shows the way to find the *signal level* for a non-multiperiodic signal. A sine wave of 0.1 amplitude and 2 day period plus a random Gaussian and correlated noise contribution is simulated as shown in Figure 11 a). The length of the time series is 8192 points allowing 13 decomposition levels. In Figure 11 b) and c), the wavelet power spectrum map and local power spectrum are computed following the method described in Section 2.6. As can be seen from the local power spectrum there is a clear peak at scale 8 corresponding to the sine wave. This would define the *signal level* for this time series and the wavelet noise-free signal would be recovered by summing the

ISWTs from scale 8 up to scale 13. However, it has to be noticed that there is a wing of the local power spectrum at scale 7 and some signal contribution can be seen in the power spectrum map at that scale. The *signal level* is set to also consider this previous scale in order to avoid losing that signal contribution. With this *signal level* we can then reconstruct the noise-free signal as explained before by summing the corresponding ISWTs. Figure 11 d) shows the noise-free signal plotted over the noisy one. As can be seen, with these method the signal can be recovered without modifying its shape, phase or amplitude.

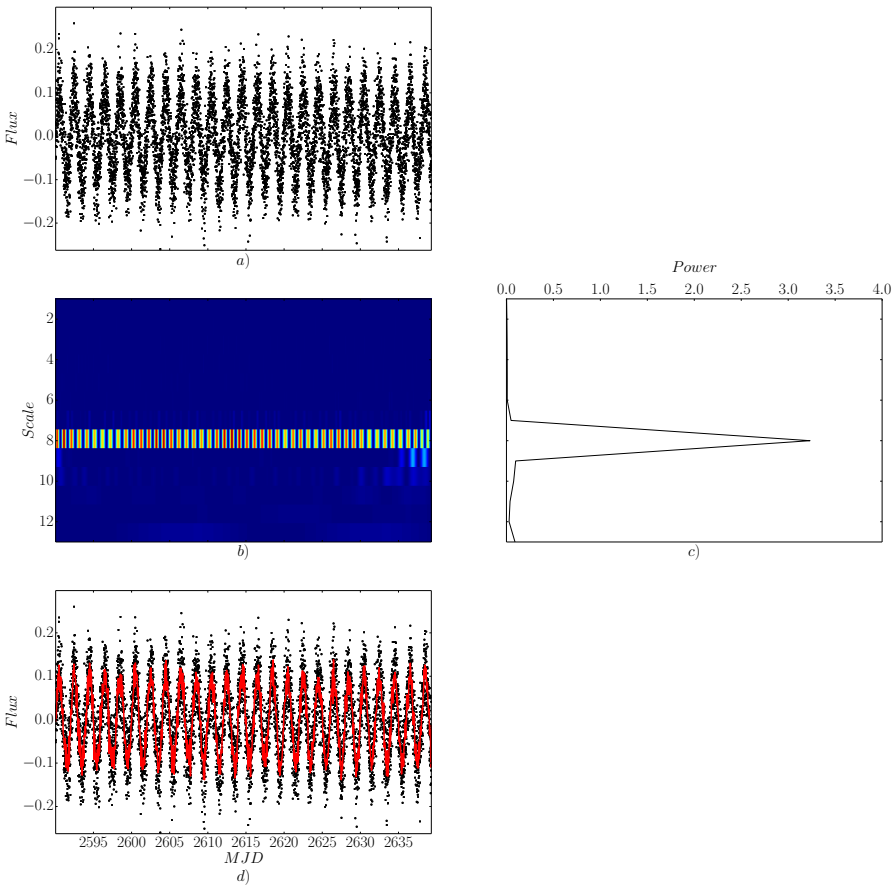


Figure 11: Signal level selection example using a simulated sinusoidal signal. a) Original signal. b) SWT wavelet power spectrum map. c) SWT local power spectrum. d) Original signal (black) and wavelet-recovered noise-free signal (red).

In a more general scenario, given that our signals could be of multi-periodic nature, resulting in WPS' peaks of diverse amplitudes at different scales, the **peak at the lowest scale** (i.e. highest frequency) is used to define the *signal level*. In addition, given the discrete nature of the SWT and, also depending on the number of decomposition levels, the energy of our signal can be distributed within more than one scale. Thus, as stated

before, the *signal level* is always set to also consider the previous scale in order to avoid losing as less signal contributions as possible. In general, the *signal level* should consider at most *max level - noise level* scales except when the scale corresponding to the maximum value of the WPS is the lowest decomposition level. This means that the signal (whether it is significant or not) has a very high frequency and is still partially or totally mixed with the noise for the available decomposition levels (given by equation 72). In this case, the *signal level* is set to be the lowest decomposition level (i.e. returning the original light curve).

3.4 OUTLIERS REMOVAL USING WAVELETS

In statistics, an outlier can be defined as an observation point that is distant from the norm of the sample. Their origin can be diverse, but the presence of outliers in a data sample leads to effects in the mean which can display a bias toward the outlier value. Also, depending on their nature, they may impact the time series analysis with respect to modeling, estimation or forecasting. There are several methods for outlier detection: model-based ones such as the Peirce's criterion (Peirce, 1852) or Grubbs' test (Grubbs, 1950); interquartile range methods such as Tukey's test (Tukey, 1949); distance-based methods (Knorr, Ng, and Tucakov, 2000) and wavelet-based methods (Mallat and Hwang, 1992; Bilen and Huzurbazar, 2002; Grané and Veiga, 2010).

One way of detecting outliers is through wavelet thresholding (Bilen and Huzurbazar, 2002). First, the wavelet transform of the target signal is computed. Then, a threshold value is computed at each decomposition level by means of any of the methods presented in Section 3.2.2. Any wavelet coefficient that exceeds the threshold value corresponding to that scale can be considered a potential outlier. Once the outliers have been detected, their associated coefficients are set to zero and the inverse wavelet transform is computed. However, we have already stated in Section 3.2.4 that this approach can create ripples around outliers and discontinuities and also, if the threshold is not correctly computed, modifications to the signal shape can result.

The outlier detection and removal approach developed for this work is a distance-based method that relies on an *estimated signal* computed from the *signal level* of the target light curve. As stated in Section 3.3, a noise-free signal can be obtained using the sum of the ISWTs for those levels from the *signal level* to the last decomposition level. Instead of using a pre-set *signal*

level, we can define it using the WPS defined in Section 2.6. With this *signal level* we can then reconstruct the *estimated signal*. A threshold can then be built in such a way that, given a point in the time series, if its distance to the *estimated signal* is above the threshold, then, it is considered an outlier and removed (or replaced by any given value). In this work, the threshold value is computed using the *Universal Threshold* (see equation 66) where σ is computed by the MAD of the original signal. Any data point exceeding 5 times this value is then replaced by the mean value of the raw time series. Once the target light curve is free of outliers, the *estimated signal* is recomputed. Except in the case of a very high frequency signal (for which the WPS peak at the lowest scale could coincide with the scale characterizing the noise), this final *estimated signal* represents an outlier-free and de-noised approximation to our target signal.

Figure 12 shows two examples of the outlier removal capabilities of the method. Figure 12 a) corresponds to the detrended light curve of EPIC 201245637 a suspected variable (Armstrong et al., 2015) from the K2 mission (see Chapter 5 for details). The light curve presents large outliers that affect the estimated signal obtained using the *signal level* as seen in the peak around phase 0.65. However, we can use this signal estimation to threshold the data following our distance-based method. Once the threshold value has been estimated as explained above, the algorithm is able to effectively remove any outlier as seen in Figure 12 b). In addition, the new signal estimation is free of any distortion caused by the presence of outliers as becomes clear from the disappearance of the small peak mentioned before around phase 0.65. In order to test whether this outlier removal method could somehow affect other type of variables not as smooth as EPIC 201245637 one, we test it on the multi planetary system K2-3. We apply it by phase folding the light curve to the K2-3 b period (see Figure 12 b) and c)). As can be seen, if the *signal level* has been correctly estimated, one can remove any outlier without modifying the signal of interest. In the case of K2-3 b it is clear that neither the shape or the depth of the transit has been modified once the outlier removal method has been applied.

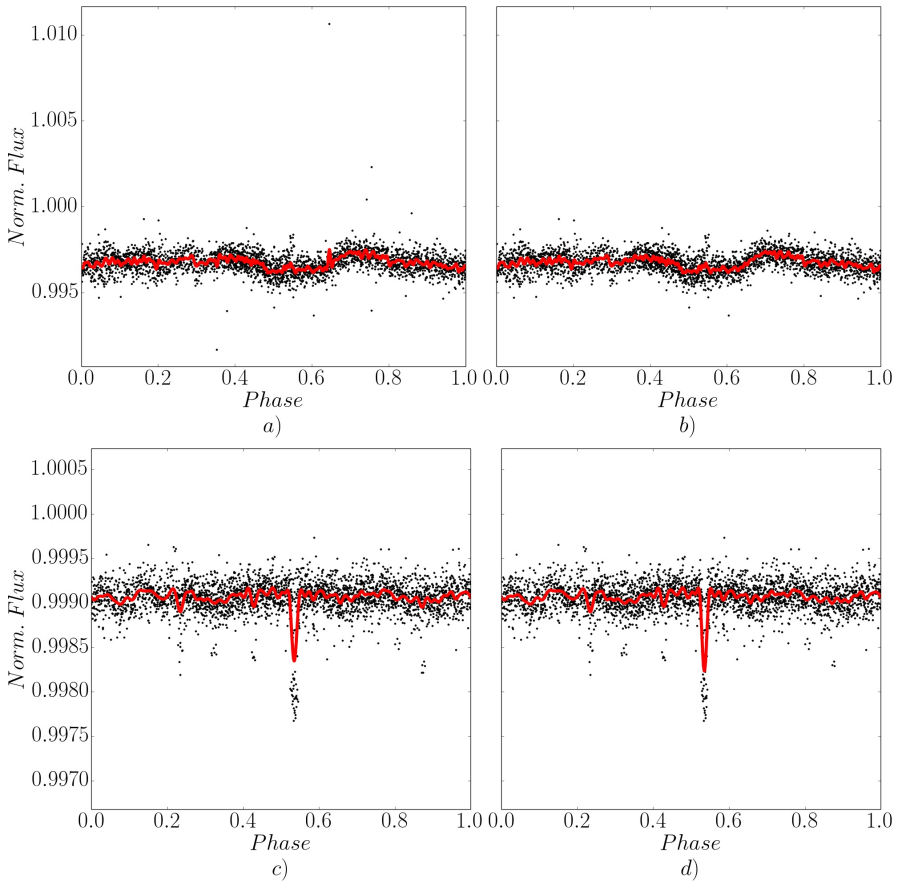


Figure 12: Two outlier removal examples. a) EPIC 201245637 detrended and phase folded light curve with outliers. Red line corresponds to the *estimated signal* obtained from the *signal level*. b) Same light curve without outliers. Red line corresponds to the new *estimated signal* after outliers have been removed. c) K2-3 detrended light curve phase folded to K2-3 b period of 10.054 days. Again, the red line corresponds to the *estimated signal* obtained from the *signal level*. d) Same light curve without outliers. Depth and shape of the transit has remained unchanged after outlier removal. As before, red line corresponds to the new *estimated signal*.

4

THE WAVELET-BASED TREND AND NOISE FILTERING ALGORITHM (TFAW)

4.1 INTRODUCTION

Algorithms such as the Trend Filtering Algorithm (TFA) (Kovács, Bakos, and Noyes, 2005) or SysRem (Tamuz, Mazeh, and Zucker, 2005) are usually applied to remove systematic variations to most time domain astronomy surveys. These comprise the ones sensitive to day-or-longer events timescales, like PTF (Law et al., 2009), Pan-STARRS (Kaiser et al., 2010), SkyMapper (Keller et al., 2007), CRTS (Djorgovski et al., 2011) or ATLAS (Tonry, 2011), that repeatedly observe few-degree-wide fields, using large apertures to achieve deep imaging, and tiling their observations across the sky; wide FoV and/or multi-telescope transiting exoplanetary surveys, such as SuperWASP (Pollacco et al., 2006), MEarth (Irwin et al., 2009), HAT-South (Bakos et al., 2008), TFRM-PSES (Fors et al., 2013) or NGTS (Wheatley et al., 2013), all-sky ones such as The Evryscope (Law et al., 2014; Law et al., 2015; Law et al., 2016) or even space-based ones such as CoRoT (using SARS (Ofir et al., 2010) a modified SysRem version).

4.2 THE TREND FILTERING ALGORITHM (TFA)

4.2.1 *Introduction*

The central idea behind the Trend Filtering Algorithm (TFA) is that many of the systematic effects present in a given photometric light curve are also present in the light curves of other stars in the same data set. To remove these systematics, one could select those objects in the field that suffer from the same kind of variations as the target star in order to build and apply some kind of filter. The filter is based on the light curves of this template of comparison stars under the assumption that they are representative of all the possible systematics. Given that no a priori knowledge on the functional form of the systematics is available, the algorithm uses a linear combination of the template stars to build the filter.

TFA can be applied to two types of situations. The first application, called *frequency analysis*, can be used to eliminate trends

from trend- and noise-dominated time series in order to enhance the probability of detecting weak signals (periodic or non-periodic). The second one, called *signal reconstruction*, can be used to iteratively reconstruct the shape of that signal if a significant period has been detected. The following subsections briefly summarize TFA's basics and formulation; for more details on the algorithm and its performance, we point to the original article (Kovács, Bakos, and Noyes, 2005).

4.2.2 *Mathematical Formulation*

Given a N-point target light curve and a set of M zero-averaged template stars $\{X_j(i); i = 1, 2, ..., N; j = 1, 2, ..., M\}$ sampled in the same moments and with the same length as the target, the filter is defined as:

$$F(i) = \sum_{j=1}^{M} c_j X_j(i) \tag{73}$$

The set of coefficients $\{c_j\}$ is determined by minimizing the following expression:

$$D = \sum_{i=1}^{N} [Y(i) - A(i) - F(i)]^2 \tag{74}$$

Where $\{Y(i)\}$ represents the target light curve and $\{A(i)\}$ is either constant in the case of the *frequency analysis* step of the algorithm or the best representation of the trend- and noise-free signal to be found in the iterative *signal reconstruction* step. Once the filter has been computed, the corrected light curve is defined by:

$$\hat{Y}(i) = Y(i) - \sum_{j=1}^{M} c_j X_j(i) \tag{75}$$

At a more technical level, TFA main steps are the following:

- Select M template stars distributed as uniformly as possible in the full field of the detector. Given that there is no a priori knowledge of which stars are variable, the selection is almost random, just those stars with high standard deviation or low brightness are discarded.

- Define the time base to be used by the filter and target time series. In our application we select the time-base corresponding to the target time series. If there are any missing points in the template light curves, they are replaced by the mean value of that light curve.
- Remove outliers and compute zero-averaged template light curves.
- Compute the normal matrix of the template:

$$g_{j,k} = \sum_{i=1}^{N} X_j(i) X_k(i); j, k = 1, 2, ..., M \tag{76}$$

  and compute its inverse $\{G_{j,k}\}$.
- For each template light curve, compute its scalar product with the target light curve:

$$h_j = \sum_{i=1}^{N} \tilde{Y}(i) X_j(i) \tag{77}$$

  where $\tilde{Y}(i) = Y(i) - A(i)$ is also assumed to be free of outliers.
- Compute the set of coefficients:

$$c_j = \sum_{k=1}^{M} G_{j,k} h_k \tag{78}$$

- Compute the corrected light curve using equation 75.

4.2.3 *Frequency analysis*

Given that there is no a priori knowledge of whether there is any periodic (or aperiodic) signal in our target light curve, TFA assumes that it is dominated by systematics and noise. Thus, $\{A(i)\}$ is set to be equal to the average of the target light curve. From this $\{A(i)\}$, the filter $F(i)$ is computed and the corrected light curve $\tilde{Y}(i)$ is obtained. The latter representing a systematic- and trend-free version of the original light curve. Now, with this corrected light curve variability detection algorithms, such as the Box Least Square (BLS) (Kovács, Zucker, and Mazeh, 2002) or Lomb-Scargle (LS) (Scargle, 1982), can be used to search for any periodicity. TFA is able to suppress most of the trends present in the light curve while simultaneously preserving the periodic signal component leading to significantly higher detection rates (Kovács, Bakos, and Noyes, 2005).

4.2.4 *Signal Reconstruction*

When TFA's *frequency analysis* step for signal detection is applied, the algorithm assumes that the light curve is trend- and noise-dominated, therefore the use of a constant $\{A(i)\}$ is justified. However, as a consequence, the detrended signal will suffer from some level of distortion. This is due to the requirement of minimum variance for the signal that is assumed to be constant while periodic signals by definition, are not. Given that for non-periodic signals no reasonable initial guess of the original shape can be done, TFA's *signal reconstruction* step can only be applied to periodic signals.

The objective of the *signal reconstruction* is to iteratively approximate the trend- and noise-free signal $\{A(i)\}$. Once a periodic signal has been found in the data, the filtered time series (equation 75) is phase folded and binned, then *re-mapped* to the original time base to give a new estimate of $\{A(i)\}$. This new estimate of the trend- and noise-free signal is then passed to equation 74 to compute a new set of filter coefficients c_j. This new filter leads to a better determination of $\{A(i)\}$, and the iteration continues until the relative difference between the standard deviations of the residuals (see equation 79) becomes under a certain limit.

$$\hat{\sigma}^2 = \mathcal{D}/(N - M) \tag{79}$$

It is important to find a good method to estimate $\{A(i)\}$ at each iteration step. Kovács, Bakos, and Noyes use the bin average method to derive the updated set of $A(i)$ through a fixed number of bins depending on the data length. This ensures statistical stability and allows to obtain a reasonable noise averaging.

4.2.5 *Application of TFA to Multi-periodic Data*

In Kovacs and Bakos (2008) they present the extension of TFA to multiperiodic signals. For signals for which the Fourier representation is adequate, the reconstruction can be done without iteration substituting equation 73 by:

$$F(i) = \sum_{j=1}^{M} c_j X_j(i) + \sum_{k=1}^{2L} a_k S_k(i) \tag{80}$$

where $S_k(i); k = 1, 2, ..., 2L; i = 1, 2, ..., N$ are the Fourier components with L different frequencies and a_k phase-dependent

amplitudes. The Fourier frequencies are determined from the analysis of the detrended signal obtained after the *frequency analysis* step has been applied (i.e. assuming $A(i)$ constant). If these frequencies approximate the ones representing the noise- and trend-free signal, equation 80 can yield an exact solution of signals with the form trend plus Fourier components plus noise. However, the authors point out that if the signal has extra components that cannot be well-represented by finite Fourier sums such as transients or transits, a more complicated model and the iterative reconstruction process should be used to obtain approximations for the signal components. In addition, the iterative reconstruction should also be employed even if the non-sinusoidal components are absent because, as stated before, we have assumed a constant noise- and trend-free signal as a starting model.

4.3 THE WAVELET-BASED TREND AND NOISE FILTERING ALGORITHM (TFAW)

4.3.1 *Preliminaries*

As explained before, in the original version of TFA, once a significant periodic signal has been found in the target time series during the frequency analysis step, the phase-folded light curve is used to iteratively estimate $\{A(i)\}$ (the noise-free signal). Kovács, Bakos, and Noyes (2005) employ the simple bin average method, using a fixed number of bins (100 for ~3000 data points) to ensure statistical stability and to obtain a reasonable noise averaging. They point out, however, that more accurate methods can be used to approximate $\{A(i)\}$. Once the signal shape is more accurately determined by the general method, it is possible to use this information to derive a more specific model. For example, in the case of planetary transits, one could use a Mandel and Agol (2002) model fit to get a more precise estimation of the noise- and trend-free signal. They emphasize that this second level of filtering can be used only if the general method supports the signal shape assumption otherwise, the output will be biased.

The Wavelet-based Trend Filtering Algorithm (TFAW) (del Ser, Fors, and Núñez, 2018) is a totally generic, Python-based, parallelized algorithm useful for any kind of survey which seeks to improve the performance of signal detection, reconstruction and characterization, leading to an overall SNR improvement without alteration of the signal's time sampling or astrophysical characteristics. TFAW introduces a wavelet filter in order to 1)

remove outliers using a wavelet-inferred estimation of the signal, 2) search for periods in this outlier-free and de-noised estimated signal during the frequency analysis step, 3) use the SWT to estimate the shape of the trend- and noise-free phase folded signal and 4) iteratively de-noise the trend-free light curve during the signal reconstruction process. TFAW combines TFA detrending and systematics removal capabilities with the wavelet transform's signal decoupling and de-noising potential. The filter is built using the SWT or *à trous* algorithm (Holschneider et al., 1989). TFAW differs from other wavelet-based noise-filtering algorithms in that it does not require any parametric model fitting as in Carter and Winn (2009) nor any extra computational method (Waldmann, 2014). Also, the noise contribution of the signal is estimated directly from its SWT at each iteration step and the de-noising is done through the subtraction of this contribution from the signal. This allows TFAW to de-noise the signal without modifying any of its intrinsic properties in contrast to wavelet coefficient thresholding (Grziwa, Korth, and Pätzold, 2014; Grziwa et al., 2016) that can lead to distortions of the signal and introduce artificial oscillations or ripples around discontinuities (Mallat, 2008).

4.3.2 *Mother wavelet selection criteria*

There are several mother wavelets with different analytic properties that can be used for signal decomposition. For our application of the SWT-modified version of TFA we use the non-orthogonal, symmetric base of the biorthogonal (bior) wavelet family, first constructed by Cohen, Daubechies, and Feauveau (1992). More specifically, we used the pre-computed values of the bior 3.9 wavelet included in the PyWavelets module[1]. We selected this wavelet family for TFAW as it allows the construction of symmetrical wavelet functions and because the shape of the reconstruction scaling function is very similar to the characteristic shape of a planetary transit (searches for other types of astrophysical variability may benefit from other wavelet shapes).

4.3.3 *TFAW Frequency Analysis*

In TFAW we follow a similar approach to the original TFA. First, an initial filter is computed using $\{A(i)\} = \langle Y \rangle$ and equations 73 and 74. The resulting filtered light curve (given by equation 75) is almost free of trends and systematics. Although,

1 https://pywavelets.readthedocs.io

as mentioned above (see Section 4.2.3), some perturbation might have been introduced by assuming a constant $A(i)$, TFA preserves the periodic signal components. In order to search for periodic signals in this first filtered TFAW light curve, the SWT and WPS of the filtered data $\hat{Y}(i) = Y(i) - \{A(i)\}$ are computed. Using the method explained in Section 3.4 outliers are removed and an *estimated signal* is obtained. The latter is then used to run Box Least Square (BLS) (Kovács, Zucker, and Mazeh, 2002) (or Lomb-Scargle (LS) (Scargle, 1982)) to search for any periodic signal in our target light curve.

4.3.4 *TFAW Signal Reconstruction*

Once BLS (or LS) has been applied during TFAW frequency analysis step, if any significant signal is found, the light curve is phase folded to its corresponding period. Instead of using a bin average, $\{A(i)\}$ is estimated using the SWT of the phase folded signal. We decompose the phase folded light curve up to a SWT decomposition level given by equation 72.

At each decomposition level, we set all the *detail coefficients* and *approximation coefficients* to zero except the ones corresponding to that level. Then, the ISWT is computed for each level separately (as seen in Figure 7). In order to estimate the shape of the noise- and trend-free signal we set a threshold decomposition level or *signal level* (see details in Section 3.3.1). $\{A(i)\}$ is then computed as the sum of the ISWTs of those levels from *signal level* to the last decomposition level. In this way we ensure that our signal's estimate is separated from the high frequency noise while preserving the high-frequency shape of the signal (unlike normal binning; see Section 4.4.4 for details in SWT signal approximation versus bin averaging): the high frequency noise components are better characterized by the lower decomposition levels as they have higher frequency resolution.

This new signal estimate, $\{A(i)\}$, can then be used to compute a new set of $\{c_j\}$ coefficients by means of equation 74. The new filter obtained from this set of coefficients gives us the new corrected light curve at each step of iteration through equation 75. We set an additional filter at each iteration given by the SWT of the phase folded light curve. In this case, instead of using those decomposition levels above *signal level* to reconstruct the noiseless signal, we set a *noise level* so that the noise is characterized

by the sum of the ISWTs of those levels below it (see Section 3.3.1). Thus equation 75 at each iteration step becomes:

$$\hat{Y}'(i) = \hat{Y}(i) - \mathrm{ISWT}(\hat{Y(i)}, \textit{noise level}), \tag{81}$$

The iteration of the algorithm stops when the relative difference between the standard deviations of the residuals in the successive iterations falls under a certain limit. As with the original TFA algorithm, we set the limit at 10^{-3}.

To summarize, the main steps of TFAW are the following:

1. An initial filter is computed using $\{A(i)\} = \langle Y \rangle$ as with the original TFA to remove trends and other systematics.
2. The SWT and WPS are computed from the filtered light curve obtained in 1.
3. Outliers are removed and an outlier-free and de-noised *estimated signal* is obtained.
4. The *estimated signal* is used to search for periodicities.
5. If a significant period is found, the signal is phase folded to run the iterative signal reconstruction. Otherwise, no reconstruction is performed.
6. Using the SWT a new signal estimation $\{A(i)\}$ is computed by means of the *signal level* of the phase folded light curve. The noise contribution of the light curve is estimated using the *noise level*.
7. The new $\{A(i)\}$ and noise contribution are used to compute the new filtered signal $\hat{Y}'(i)$.
8. Iteration continues until the convergence criterion is met.

4.4 TFAW PERFORMANCE

In Kovács, Bakos, and Noyes (2005) several tests are presented demonstrating the signal detection and reconstruction capabilities of TFA. We want to explore how the inclusion of the SWT filter in TFAW can improve the detection of variable signals, assess whether the use of the *estimated signal* has any impact on the frequency power spectra and study the effects on the Signal-to-noise Ratio (SNR) of the light curves. Also we compare the use of the SWT to compute the signal approximation against bin averaging, test the likeness of the final TFAW reconstructed light curve to the inserted variable signal and compare

the results with the ones obtained using the original TFA. For all the simulated light curves hereafter we have used 2048 data points and set the *noise* and *signal levels* to the lower decomposition level and the one given by the WPS peak as explained in Section 3.3.1 and 3.4.

4.4.1 *TFAW vs TFA transit detection efficiency*

In order to assess the transit detection capabilities of BLS during TFAW and TFA frequency analysis step, we generate a set of 5,000 time series with different random combinations of Gaussian and correlated noises. We also introduce linear and exponential trends (i.e. to simulate changes in airmass, changes in the CCD position of the object, etc), as well as gaps and jumps in the data. We use a set of 250 simulated template stars suffering from the same jumps, gaps and trends as the target time series, affected by random distributions of σ_w and σ_r combinations. To each of these light curves, we add the transit signal given by the parameters of planets 1 and 2 in Table 1 separately. We use `batman` (Kreidberg, 2015) package to simulate a Mandel and Agol (2002) planetary transit model. The total noise contribution for each light curve ranges from 0.005 to 0.2 mag for planet 1 and from 0.005 to 0.1 mag for planet 2.

Table 1: Planet parameters used for TFAW simulations

Planet	R_P (R_J)	M_P (M_J)	M_* ($M_\odot$)	R_* ($R_\odot$)	P (d)
1	1.98	1.40	1.36	1.23	0.8468
2	1.23	1.10	1.36	1.23	0.4842

We want to compare the detection rates obtained with TFA and TFAW light curves. First, we define the Signal Detection Efficiency (SDE) (Alcock et al., 2000; Kovács, Zucker, and Mazeh, 2002):

$$\mathrm{SDE} = \frac{\mathrm{SR}_{peak} - \langle \mathrm{SR} \rangle}{\mathrm{sd}\,(\mathrm{SR})} \tag{82}$$

where SR_{peak} is the power spectrum value at the highest peak, $langleSR\rangle$ is the average and $\mathrm{sd}\,(\mathrm{SR})$ is the standard deviation of the power spectrum over the frequency band being used. The SDE can be used to define a threshold above which, a signal can be considered *significant* or reliable. The selection criteria for this threshold is arbitrary and should be a compromise be-

tween losing too many true detected signals and too high false alarm rates.

We use the same two criteria as Kovács, Bakos, and Noyes (2005) and define a detection as the following:

- The highest peak in the BLS power spectrum must have a frequency between $[f_p - 0.001, f_p + 0.001]$, where f_p is the frequency of the simulated planetary transit.
- The SDE of the highest peak in the power spectrum must be greater than 6.

We run BLS using 100 bins and 99,000 frequency steps to ensure statistical stability (Kovács, Zucker, and Mazeh, 2002). The transit search is done in the (0.01, 12) days range. We count the number of times the signal is detected in the TFAW light curves but not detected in the TFA ones. By performing this and the opposite test using a template sample of 250 stars, we obtain the results shown in Table 2.

Table 2: Mutually exclusive detections and mean SDE values for simulated planetary transits 1 and 2 as shown in Table 1. N_{TFA}: not detected using TFAW light curves, but detected using TFA data. N_{TFAW}: detected using TFAW light curves, but not detected using TFA data. N_{mut}: simultaneous detections with TFA and TFAW. SDE_{TFA}: mean TFA SDE. SDE_{TFAW}: mean TFAW SDE. Percentage values in parenthesis are with respect to the 5,000 tested transits.

Planet	N_{TFA}	N_{TFAW}	N_{mut}	SDE_{TFA}	SDE_{TFAW}
1	265 (5.3%)	784 (15.7%)	561 (11.2%)	16.99	17.86
2	79 (1.58%)	503 (10.1%)	612 (12.2%)	14.61	15.92

Table 3: Detection distributions for simulated planetary transits 1 and 2 as shown in Table 1 for three bins of total noise contributions.

σ	< 0.01			$0.01 - 0.04$			> 0.04		
	N_{TFA}	N_{TFAW}	N_{mut}	N_{TFA}	N_{TFAW}	N_{mut}	N_{TFA}	N_{TFAW}	N_{mut}
Planet 1	-	-	75	76	131	318	189	653	167
	N_{TFA}	N_{TFAW}	N_{mut}	N_{TFA}	N_{TFAW}	N_{mut}	N_{TFA}	N_{TFAW}	N_{mut}
Planet 2	10	41	107	67	393	491	2	69	14

The example shown in Table 2 shows that the number of non-simultaneous detections (both in absolute and percentage) is always higher (a factor ~3× for planet 1 and ~6× for planet

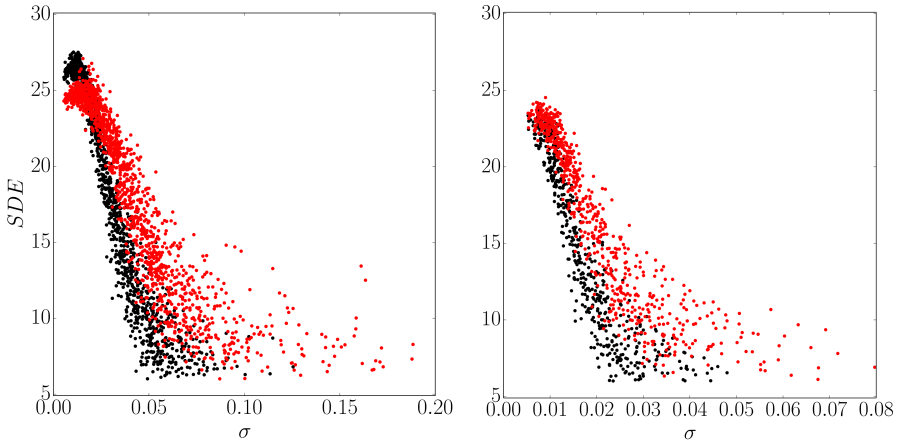


Figure 13: TFA vs TFAW detections. **Left:** SDEs of detections for planet 1 in Table 1 versus signal noise for TFA (black dots) and TFAW (red dots). **Right:** Same but for planet 2 in Table 1.

2) in the case of TFAW light curves for both planetary transits. Also the mean SDE values are higher for TFAW than for TFA. As can be seen in Figure 13 the SDE improvement for a given light curve can be up to a factor ~2.5× for low SNR signals. Table 3 shows the distribution of the TFA and TFAW detections in three total noise contribution bins. In the high-SNR regime both TFA and TFAW behave in a similar way. TFAW performs better than TFA in detecting the transit signal in the mid-SNR regime (σ=0.01-0.04) by a factor ~1.16× for planet 1 and a factor ~1.6× for planet 2. In the low-SNR case (σ >0.04), TFAW detects the transit in ~2.3× more light curves than TFA for Planet 1 and in ~5.2× for Planet 2. Taking into account the whole noise range, TFAW improves the detection rate by a factor ~1.6× for both planetary transits. Figure 14 shows two examples of the increase in the SDE of the BLS power spectrum peaks for Planet 1 in Table 1 in the low SNR regime. Top two panels correspond to a simultaneous detection for which TFAW yields an increased SDE for the true period of 11.2 versus 6.2 for TFA. The bottom two panels present the power spectrum for a TFAW mutually exclusive detection. In this case, TFA is not able to detect the transit as the peak corresponding to the true period is hidden within the noise peaks of the power spectrum with a SDE of ~3.9 (below the detection threshold defined above). For TFAW, the peak can easily be identified and has a SDE of ~8.3 (a factor ~2.1× higher than for TFA). In addition, TFAW is able to detect the planetary signals, with a SDE above 6, up to a total noise of ~0.19 mag for planet 1 and up to ~0.07 mag for planet 2 whereas TFA is able to detect them up to ~0.1 and ~0.045 mag respectively.

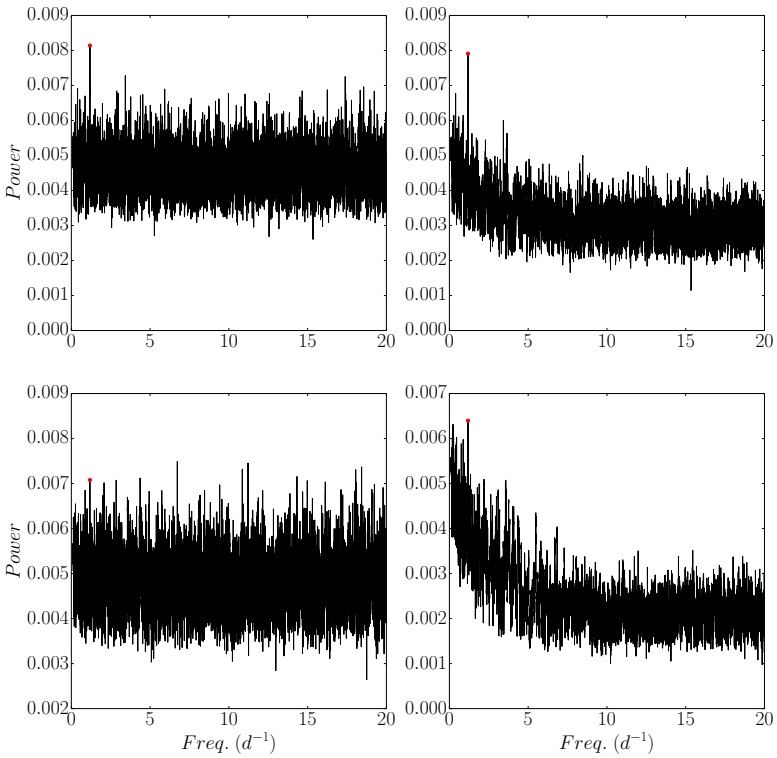


Figure 14: TFA vs TFAW BLS power spectrum examples for planet 1 in Table 1. **Top:** TFA BLS power spectrum (left) vs TFAW BLS power spectrum (right) for a simultaneous detection as defined in Section 4.4.1. Red dots mark the true period of the transit. **Bottom:** Same but for a TFAW mutually exclusive detection.

It is worth noting, that, the SDE as defined by Kovács, Zucker, and Mazeh (2002) may be biased depending on the frequency range used. BLS power spectra tend to have higher values towards low frequencies. Also, the SDE relies on the mean and standard deviation (which are not robust estimators in front of outliers). Thus, if the frequency range is cut at low frequencies, there is not enough floor to give a robust SDE value. This is something that is noted in Kovács, Zucker, and Mazeh (2002) where they state: "Because in the practical computation of SDE one uses all available spectral points, in the presence of periodic signal, the actual value of SDE also depends on the time spanned by the data and on the lengths and position of the frequency band of the analysis. If all other parameters are kept constant, increasing time span or frequency band leads to an increase in SDE for signals containing periodic component(s)." The use of the median and/or the MAD could yield more stable results. We have also done tests using the generalized SDE defined in Ofir (2014) that uses a median filter to remove the trend towards low frequencies in the BLS power spectrum. Again, in

these tests, TFAW performs better than TFA in giving increased SDE values and yields somehow more robust results regardless of the frequency range used.

4.4.2 *TFAW vs TFA signal reconstruction*

To illustrate the noise and trend filtering efficiency of the TFAW method during the iterative *signal reconstruction* step, we generate a sinusoidal signal affected by a combination of different levels of simulated Gaussian, σ_w, and correlated, σ_r noises. As in Section 4.4.1 we also introduce linear and exponential trends, as well as gaps and jumps in the data. The period of the simulated signal is 0.63 days and 0.03 mag amplitude. Again, we use a template of 250 simulated template stars suffering from the same trends and systematics as the target star, each of them affected by a random distribution of σ_w and σ_r combinations (see Table 4 for details).

Table 4: Noise parameters for the simulated light curves

Signal type	SNR	σ_w (mag)	σ_r (mag)
Sinusoidal	High	0.01	0.005
Sinusoidal	Low	0.1	0.005
Transit	High	0.01	0.005
Transit	Low	0.04	0.005

Figure 15 shows the noise filtering capabilities of TFAW compared to TFA for a simulated target star with a sinusoidal modulation. While the noise dispersion is clearly diminished, there is no modification to either the amplitude or the phase of the signal, and the sampling timescale of the signal remains unchanged (unlike in simple binning). We also show the LS power spectrum obtained in the frequency analysis step. As can be seen the height and amplitude of the peaks for the TFAW case have a higher power value compared to the TFA ones though the SDEs in both cases are very similar.

In Figure 16, the same sinusoidal signal is simulated but with lower SNR. As before, the noise dispersion is also significantly diminished while the period and amplitude of the signal remains unchanged. In this case, the LS power spectrum is improved by a greater factor than the one for the high SNR case. However, as in the previous case, although the amplitude of the power spectrum peaks is greater in the case of TFAW, the SDE

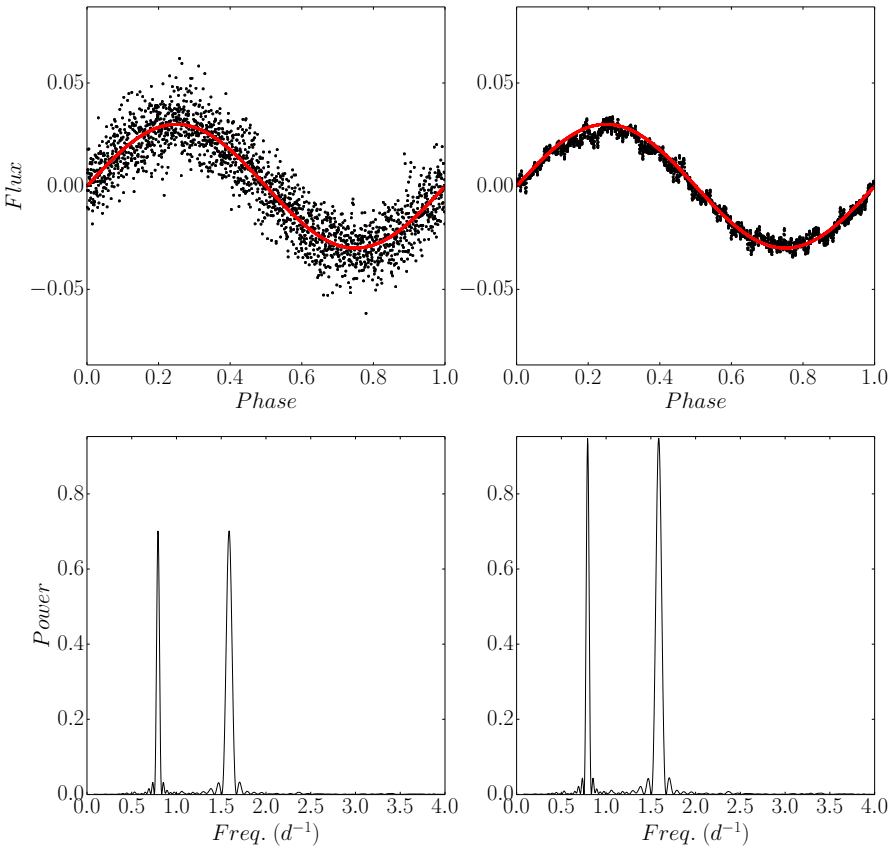


Figure 15: Noise filtering comparison of a simulated sinusoidal signal. The same number of template stars and LS parameters were used for TFA and TFAW. **Top left:** TFA-detrended and reconstructed phase folded signal. Red line corresponds to the simulated signal. **Top right:** The same phase folded signal but TFAW-detrended, reconstructed and de-noised. **Bottom left:** LS power spectrum of TFA frequency analysis step. **Bottom right:** LS power spectrum of TFAW frequency analysis step.

values for both TFA and TFAW are very similar (10.55 and 10.6, respectively) due to the low LS continuum.

Figure 17 simulates a planetary transit. We use `batman`, with the parameters of planet 1 in Table 1 to simulate a Mandel and Agol (2002) planetary transit model, and the noise parameters in Table 4. The high-frequency noise is filtered by TFAW during the iterative reconstruction step, resulting in a better defined transit without modifying its shape and depth. This result overcomes the artificial ripples and transit depth modification introduced in Figure 6 of Grziwa et al. (2016) due to wavelet coefficient thresholding.

The TFAW BLS power spectrum for this high SNR planetary transit remains almost unchanged with respect to the TFA version (with a lower noise floor), because the BLS power spectrum

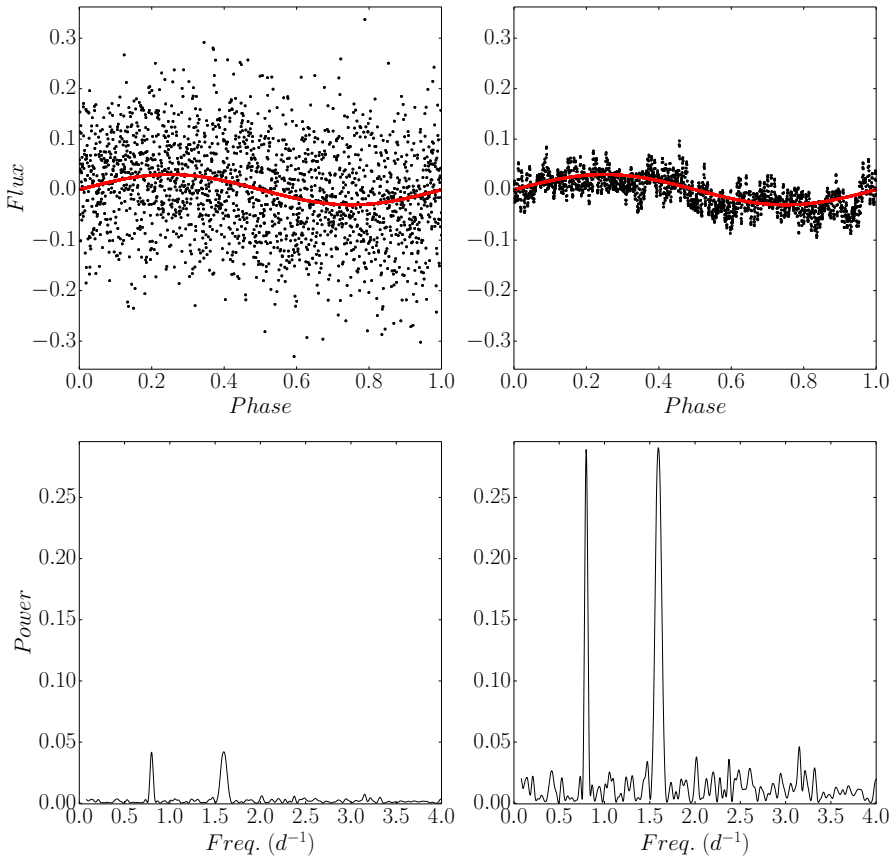


Figure 16: Noise filtering comparison of a simulated sinusoidal signal with lower SNR. Same notation, TFA and TFAW parameters as Figure 15.

is not primarily limited by high-frequency random noise. The SDE values for this case are also very similar; 23.6 for TFA and 24.5 for TFAW.

Figure 18 shows the same transit but with a much larger noise component (see Table 4). As in the case of the simulated sinusoidal with low SNR (see Figure 16), TFAW improves the characterization of the signal's shape while diminishing the noise, recovers the correct period and transit depth, and improves the detectability of the transit signal in the BLS periodogram (with and SDE of 17.8 for TFAW and 13.3 for TFA).

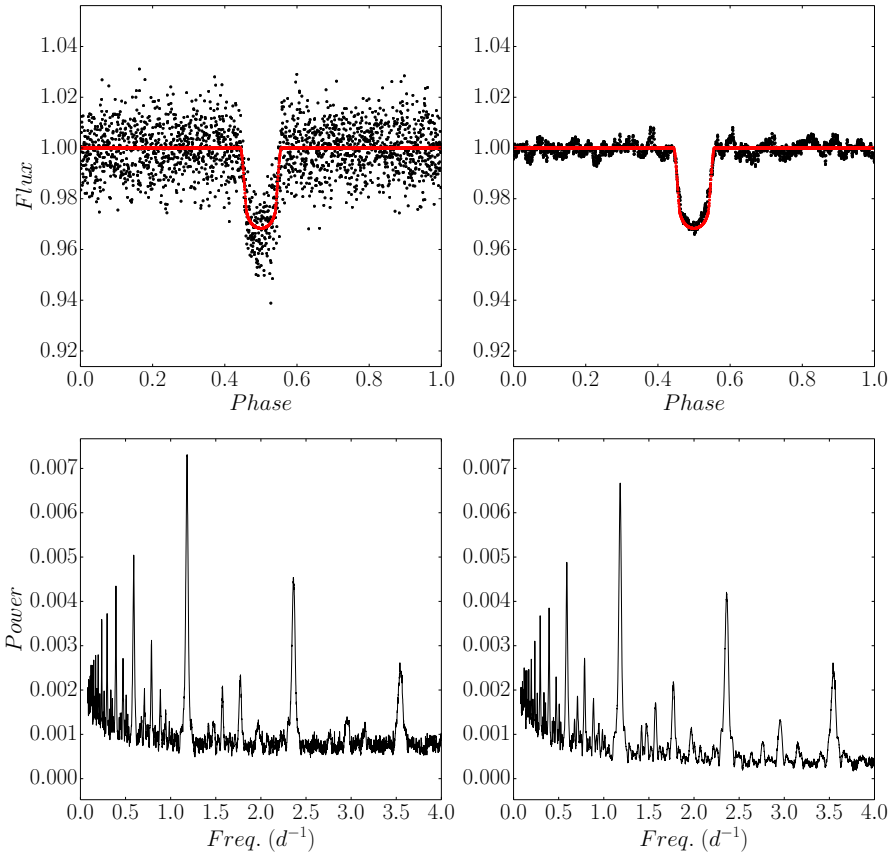


Figure 17: Noise filtering comparison of a simulated box-shaped transit (planet 1 in Table 1). Same notation, TFA and TFAW parameters as Figure 15.

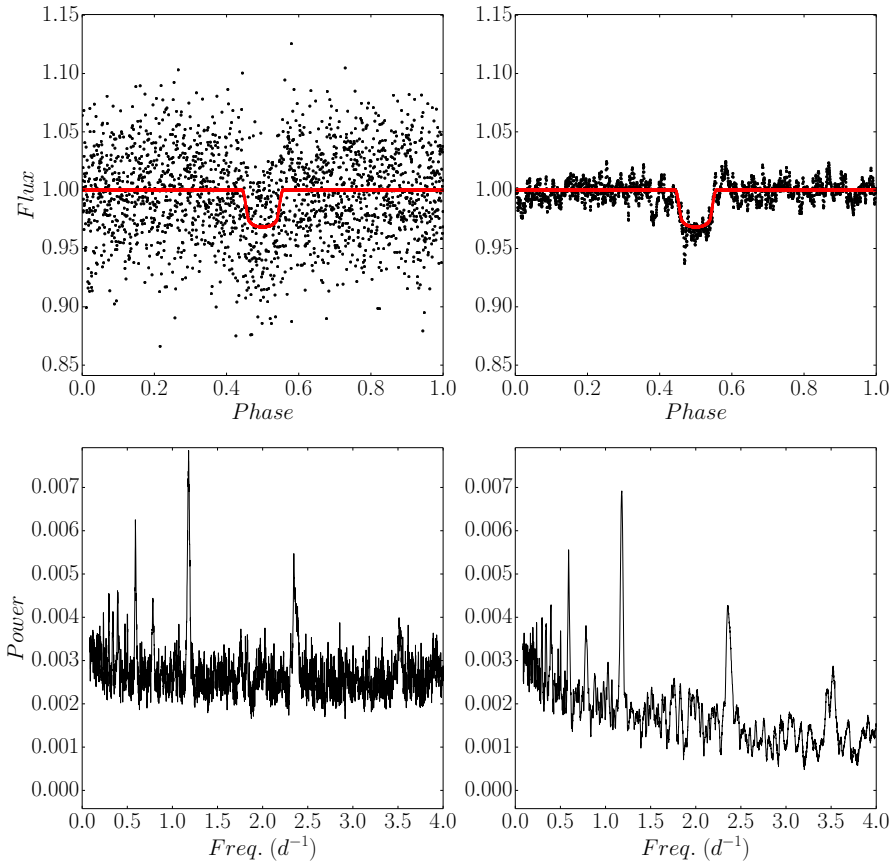


Figure 18: Noise filtering comparison of the same simulated box-shaped transit as in Figure 17 with lower SNR. Same notation, TFA and TFAW parameters as Figure 15.

4.4.3 *Application to multiperiodic signals*

Kovacs and Bakos (2008) presents a variation of TFA to extend its application to multiperiodic signals using their Fourier representation (for the cases in which such representation is adequate). However, they state that if the signal has additional components such as transients or planetary transits, a more complicated model should be used to approximate each of those extra signals present in the light curve. TFAW can be used to separate the different signal contributions directly from the signal after TFA's frequency analysis step (provided that the periods of each of the signals have been previously and correctly determined using BLS or another method). As a first example, we simulate the same high SNR-type sinusoidal signal as in Figure 15 and Figure 16 modulated with a 0.1438-day sinusoid with an amplitude of 0.006 mag, and with noise parameters according the first row of Table 4. As can be seen in Figure 19, if we are able to find the correct frequency of the secondary signal (notice the peak around 6.9 d^{-1} in the corresponding LS periodogram), we can fully recover it by applying TFAW directly (i.e. with no need of subtracting the primary signal) to the raw data without modifying the amplitude and shape of any of the underlying signals.

The second example in Figure 20 shows the results of applying TFAW to a high SNR-type light curve affected by two planetary transits simulated with `batman` using the parameters of planets 1 and 2 in Table 1, and noise parameters according the second-to-last row of Table 4. Again, if we are able to find the period of the secondary transit after the frequency analysis step, we can separate the secondary transit from the primary simply by phase folding the raw light curve and applying TFAW. As in the other examples, we are also able to improve the SNR of both the primary and secondary signals.

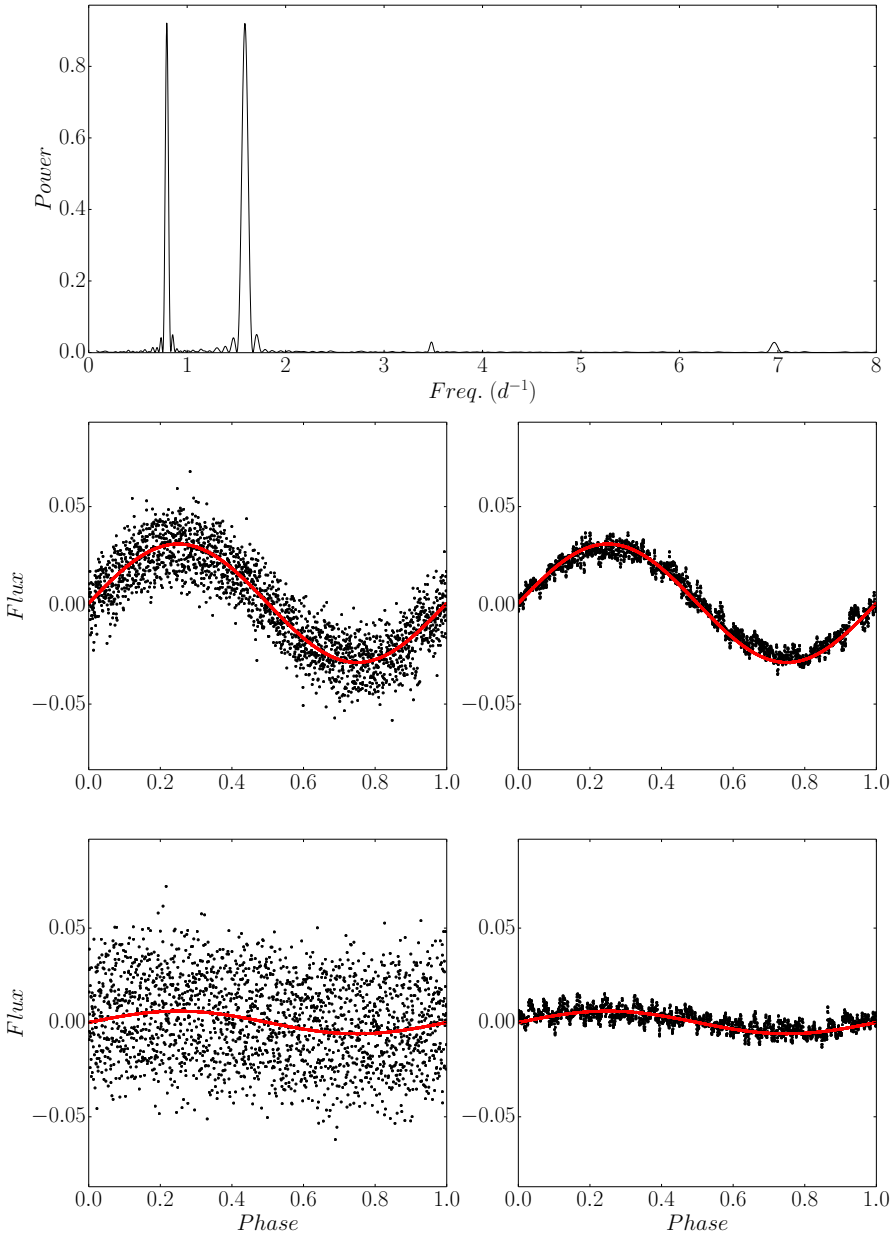


Figure 19: Example of the signal recovery for a multi-periodic sinusoidal signal. Red line corresponds to simulated signal. **Top:** LS power spectra of the signal after TFAW frequency analysis step (notice the small peak around 6.9 d^{-1} corresponding to the secondary signal). **Middle left:** TFA-detrended and reconstructed phase-folded low frequency signal. **Middle right:** Same TFAW-filtered phase folded low frequency signal. **Bottom left:** TFA-detrended and reconstructed secondary signal. **Bottom right:** TFAW phase-folded secondary signal.

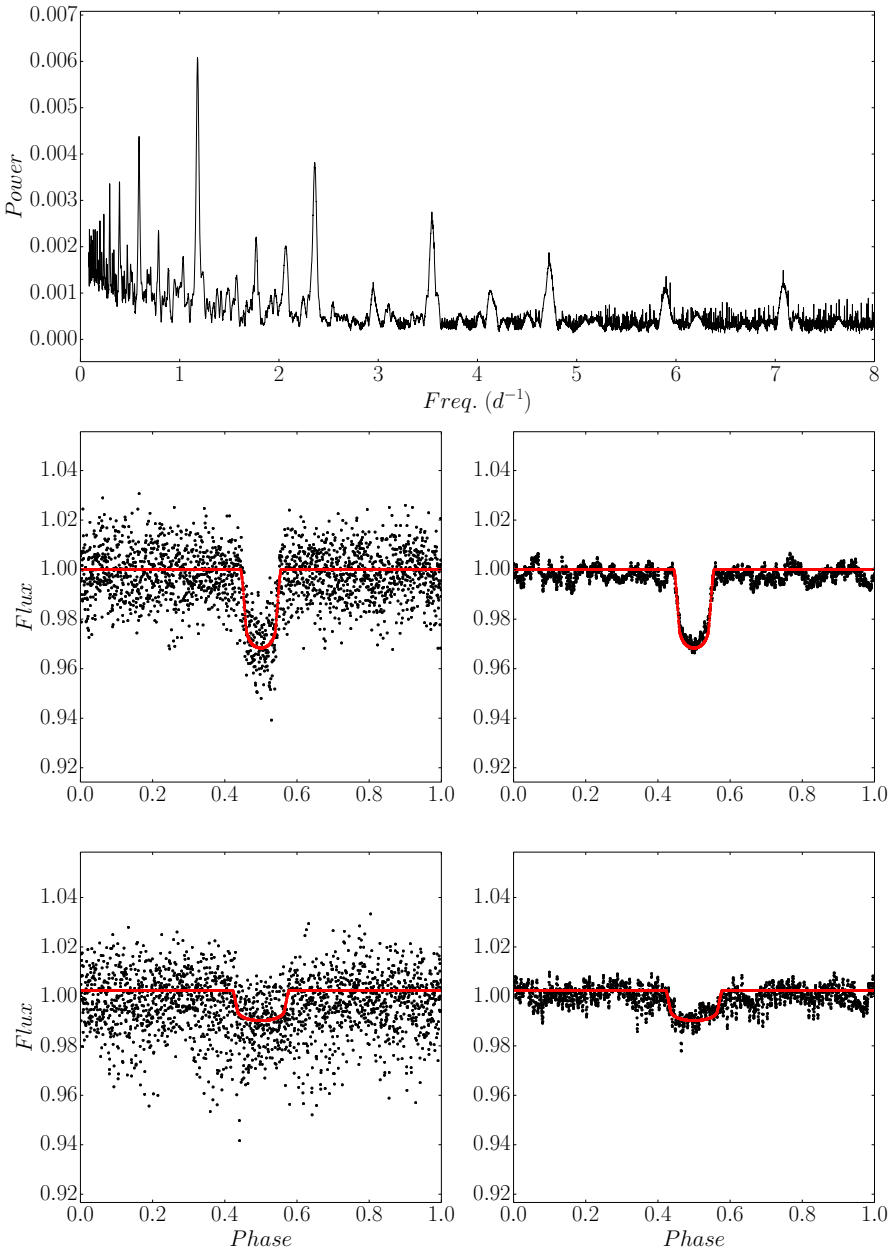


Figure 20: Example of the signal recovery in a multi-transit light curve. Red line corresponds to simulated signal. **Top:** BLS power spectra of the signal after TFAW frequency analysis step. Check the peak around 2 d^{-1} corresponding to the secondary transit. **Middle left:** TFA-detrended and reconstructed phase-folded planet 1 signal. **Middle right:** Same TFAW-filtered phase folded planet 1 signal. **Bottom left:** TFA-detrended and reconstructed planet 2 signal. **Bottom right:** TFAW phase-folded planet 2 signal.

4.4.4 *Wavelet versus bin average signal approximation*

In order to test the likeness of the signal approximation, $\{A(i)\}$, to the inserted signal TFAW compared to the bin averaging used by TFA, we use planet 1 in Table 1. We inject it in 500 light curves, using `batman` and a Mandel and Agol (2002) planetary transit model, for increasing values of noise contribution (i.e. lower transit depth). For each simulation, we compute the signal estimation given by the SWT and bin averaging with 100 bins and obtain their deviation with respect to the simulated input signal. Figure 21 shows that, in general, the estimation of the transit shape given by the sum of the ISWTs defined by *signal level* provides a better representation than the one given by the TFA bin average method. This is especially true for low σ_{signal}, with up to a factor of $\sim 2\times$ of improvement.

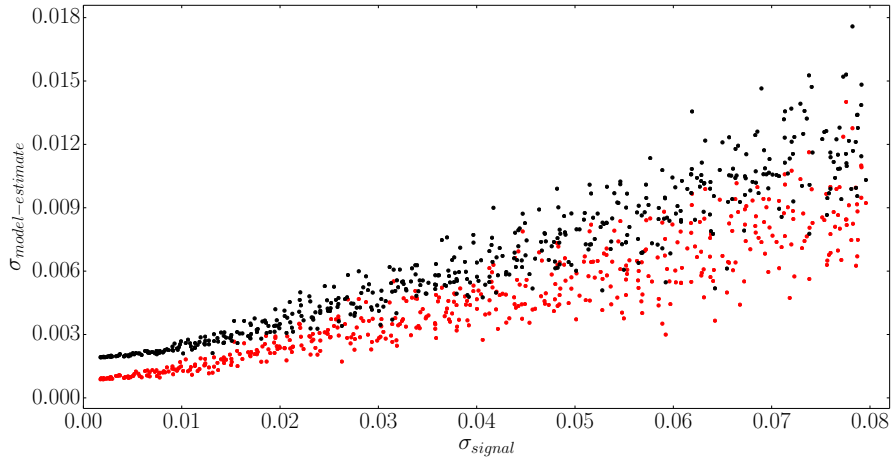


Figure 21: Wavelet signal approximation versus bin average comparison. Comparison of the standard deviations of the estimated signals obtained by the sum of the ISWTs given by *signal level* (red) and the one given by bin averaging (black) for Planet 1 in Table 1 for decreasing transit depth.

In the case of planetary transits, as can be seen in Figure 22, the improvement in the signal approximation is due to the fact that the wavelet reconstruction of the signal better fits the ingress and egress profiles of the transit even in the cases of low signal-to-noise ratios compared to the bin average method.

4.4.5 *Comparison of TFA and TFAW transit parameters fit values and uncertainties*

We want to compare TFA vs. TFAW performance in terms of assessing the bias of the fitted transit parameters values and their uncertainties. To quantify those, a low SNR-type plane-

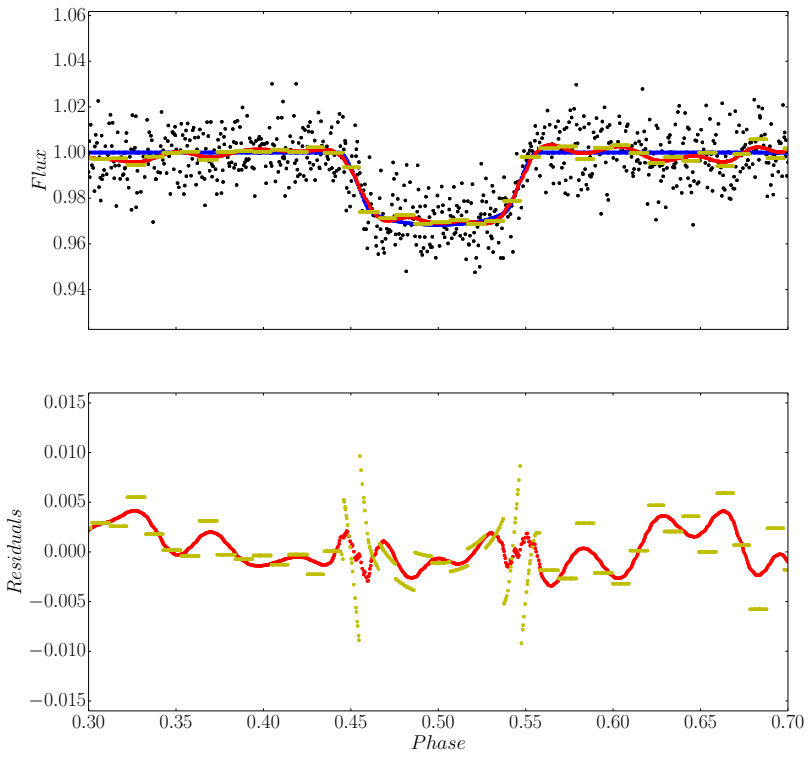


Figure 22: Wavelet signal approximation versus bin average comparison for Planet 1 in Table 1. **Top**: Planetary transit with high SNR. Blue line corresponds to the input Mandel and Agol (2002) model, red line represents the wavelet approximation of the signal, and yellow line is the bin average approximation. **Bottom**: Difference between the planetary model, and the wavelet and bin average approximations (same color notation).

tary transit of the planet 1 according Table 1 and last row of Table 4 was considered. This was modeled with `batman`, following a Mandel and Agol (2002) analytic transit model. We used the Markov chain Monte Carlo (MCMC) sampler provided by `emcee`[2] (Foreman-Mackey et al., 2013) to sample the posterior distribution of the 6 transit parameters (a, q, i, per, p, l). Keeping the eccentricity fixed, we consider a uniform distribution for the priors and run the sampler with 200 chains and 5,000 iterations with a burn-in phase of 1,000 iterations.

Figure 23 shows the 1-D and 2-D projections of the posterior probability distributions of the 6 MCMC fitted parameters for the TFA and TFAW detrended light curves. Similarly, at top panel of Table 5 we compare the injected transit parameters values with the ones obtained through MCMC for TFA and TFAW posterior probability distributions. MCMC parameter values correspond to the 50% quantile while the uncertainties are computed

2 http://dfm.io/emcee

from the 25% and 75% quantiles as the upper and lower errors. In the case of TFA, while some of the distributions are fairly behaved, for parameters a, q and p, they present wider features either characterized by larger uncertainties (a, q) and/or larger biases (a, p). On the other hand, for TFAW light curves, the biases with respect to the initial parameters values are strongly diminished compared to TFA ones. As for the uncertainties, they are, in general, smaller for TFAW than for TFA. In some parameters, such as i, l and a, they are largely decreased. At the bottom panel of Table 5, the 95% confidence highest probability density credibility intervals for both TFA and TFAW MCMC cases, are shown. The width of the credibility intervals is $\sim 10\times$ narrower (except for the period which is better determined) in the case of TFAW compared to the TFA ones.

Table 5: Top table: Actual parameters values used to simulate the transit of Planet 1. Posterior transit parameters values and their uncertainties (with the 25% and 75% quantile as the upper and lower errors) for TFA and TFAW MCMC fits. Bottom table: 95% confidence intervals of the highest probability density for Planet 1 transit parameters TFA and TFAW MCMC fits.

Parameters	a (R_*)	q	i (°)	per (days)	p	l
Simul. values	3.392862	0.214	88	0.8468	0.16542286	0.312
TFA MCMC	$3.53363^{+0.06460}_{-0.08075}$	$0.37952^{+0.10736}_{-0.09784}$	$87.7895^{+0.19976}_{-0.28489}$	$0.84628^{+0.0009}_{-0.00008}$	$0.19564^{+0.00347}_{-0.0032}$	$0.34398^{+0.09770}_{-0.08413}$
TFAW MCMC	$3.39438^{+0.01072}_{-0.00228}$	$0.21317^{+0.00206}_{-0.00591}$	$88.0075^{+0.00687}_{-0.00212}$	$0.84687^{+0.00008}_{-0.0001}$	$0.16656^{+0.00312}_{-0.00102}$	$0.31259^{+0.00603}_{-0.00189}$
95% confidence intervals of the highest posterior density						
TFA MCMC	3.368 - 3.672	0.186 - 0.577	87.31 - 88.21	0.8461 - 0.8465	0.186 - 0.204	0.161 - 0.564
TFAW MCMC	3.384 - 3.432	0.198 - 0.219	87.99 - 88.02	0.8467 - 0.8470	0.164 - 0.172	0.305 - 0.335

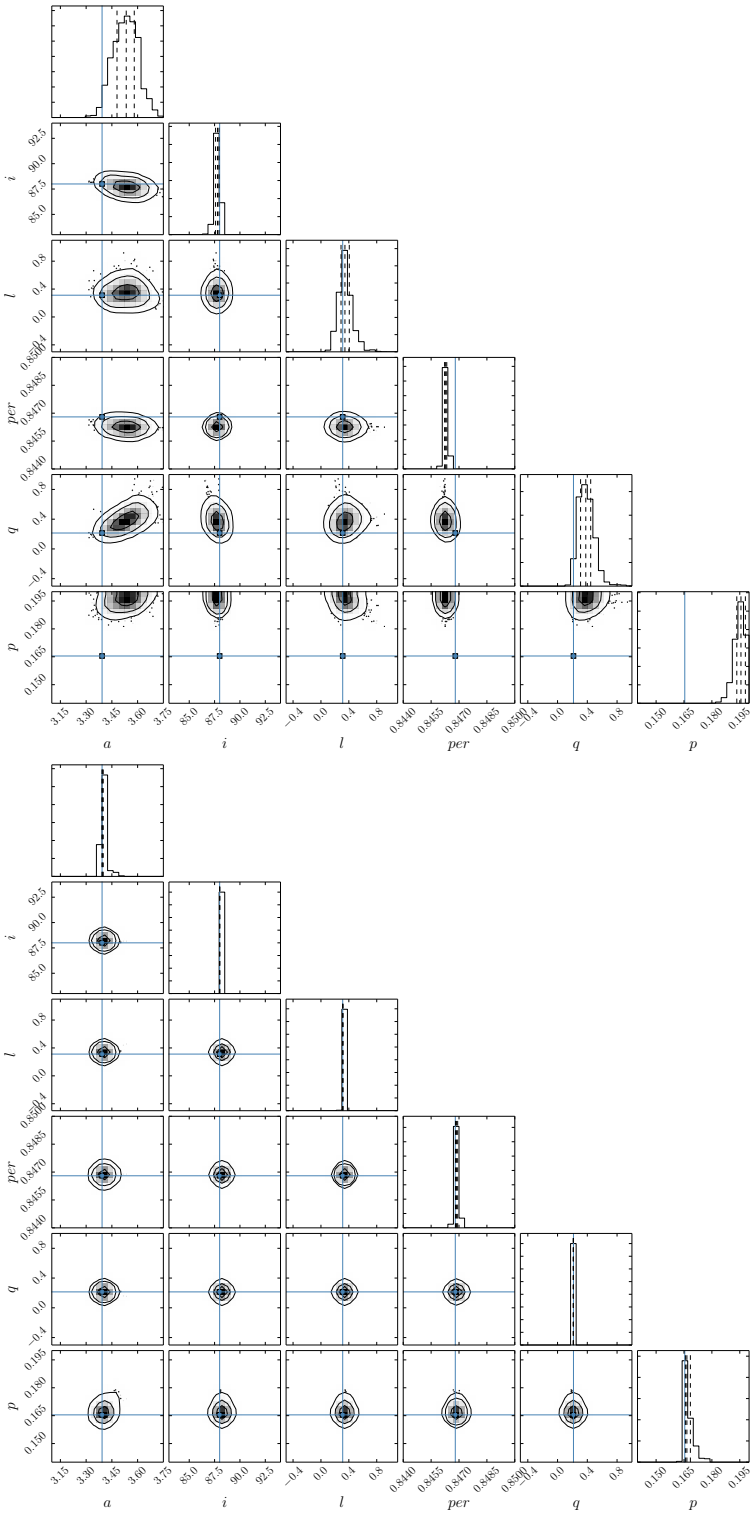


Figure 23: 1-D and 2-D projections of the posterior probability distributions of the 6 MCMC fitted parameters for the TFA (top) and TFAW (bottom) detrended light curves. The injected values for (a, q, i, P, p, l) are marked in solid blue. The 25%, 50%, 75% quantiles, are displayed in dash vertical lines on the 1-D histograms.

Part III

WAVELETS AND PHOTOMETRY: APPLICATION OF THE TFAW ALGORITHM TO REAL SURVEY DATA

5 INSTRUMENTATION AND SURVEYS

5.1 GROUND-BASED TELESCOPES

5.1.1 *The Telescope Fabra-ROA at Montsec (TFRM)*

5.1.1.1 *Project Overview*

The launch of Sputnik I, the first artificial Earth satellite, on October 4, 1957 and of other satellites a few months later, marked the beginning of the early Space Age. As a solution for optically tracking these satellites, the Smithsonian Institution designed and built a new kind of telescope: the Baker-Nunn Camera (BNC) (Henize, 1957), named after James Baker and Joseph Nunn, its optical and mechanic designers respectively. These wide-field f/1, 0.5m aperture, 30° x 5° field-of-view (FoV), photographic telescopes were manufactured by Perkin-Elmer (optics) and Boller & Chivens (mechanics) with the highest quality specifications. As a result, the BNC was able to achieve satellite positional measurements with a typical accuracy of ~2" for one single station.

In order to maximize the satellite coverage and minimize the positional measurement error, a family of 21 BNCs were manufactured and placed all over the world spanning in longitude. In 1958 one of them was installed at the Real Instituto y Observatorio de la Armada (ROA), in San Fernando (Cádiz), Spain. Later on, during the 80s the BNC program became obsolete with the appearance of new technologies (such as laser, radar and CCD) and the BNC in San Fernando was donated to ROA where it remained inactive but in excellent state of conservation.

The TFRM project consists in the refurbishment of ROA's BNC for robotic CCD surveying purposes (Fors et al., 2013). It is a joint collaboration between the Reial Acadèmia de Ciències i Arts de Barcelona (RACAB) and ROA which counts with the participation of several members of the Dept. de Física Quàntica i Astrofísica, Institut de Ciències del Cosmos (ICCUB) at University of Barcelona. Currently, the BNC is installed at the summit of the Montsec d'Ares (Lleida, Spain) as part of the Observatori Astronòmic del Montsec (see `http://www.am.ub.edu/bnc` for more details).

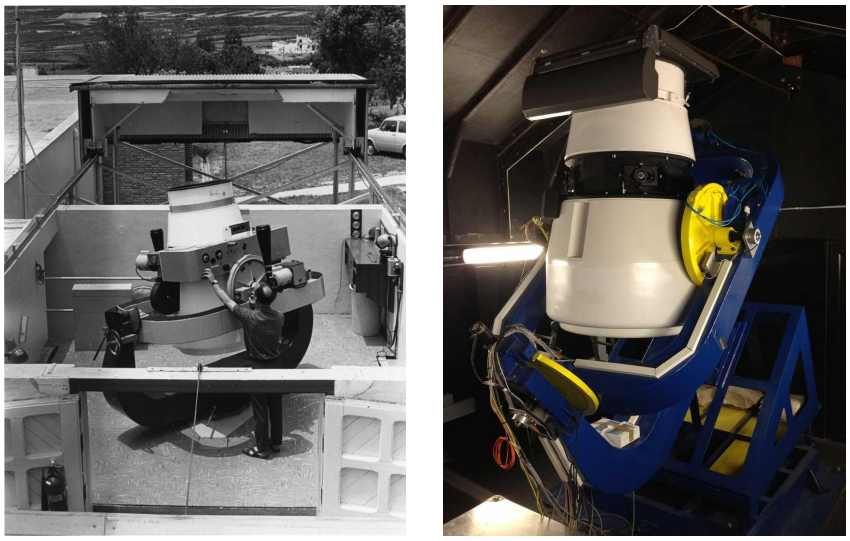


Figure 24: Baker-Nunn Camera (BNC) at ROA in 1958 (left) and at its present location at Observatori Astronòmic del Montsec (right).

5.1.1.2 *Optical Specifications*

Due to the original BNC's 30° × 5° FoV, a curved focal surface was necessary. However, commercial CCD detectors require a flat focal plane. During the refurbishment project, the manufacture of three new elements was mandatory to achieve it: a biconvex field flattening lens, a meniscus lens, and a plano-plano colour filter. Both the outermost surface of the telescope objective lens and the primary mirror had also to be repolished and recoated respectively to achieve maximum throughput of the system. As a result, this corrected design yielded an f/0.96 BNC system with a 4.4° × 4.4° FoV. In summary, the refurbished optical layout of the TFRM's BNC consists on a field flattener corrective lens close to the CCD and a meniscus lens (positioned slightly farther from the focus) that provides correction for the astigmatism introduced by the former and for barrel distortions. All technical specifications of the retrofitted BNC are summarized in Table 6.

Table 6: Optical Specifications of the TFRM refurbished BNC

Aperture	Focal Ratio	Mirror ϕ	Sensor	Scale	FoV
0.5 m	f/0.96	0.78 m	KAF-16803E	3.9"/pix	4.4 deg^2

Given the many observational programs carried out by the TFRM project, the use of a filter was found to be desirable. Johnson/Cousin filters were promptly discarded as the chromatic aberration caused by the great incidence angle of the f/0.96 beam was unavoidable. Since the BNC optics were not opti-

mized for blue wavelengths and due to the KAF-16803E lower quantum efficiency in this part of the visible spectrum, finally, a Schott GG475 yellow colored filter with a cut-off frequency of 475 nm was chosen.

5.1.1.3 *The TFRM-Preselected Super-Earth Survey (TFRM-PSES)*

The TFRM-PSES is an ongoing systematic search for super-Earth planets orbiting M-dwarf stars following a similar approach as the MEarth-North survey (Irwin et al., 2009). Due to the very low luminosities (relative to the ones of solar type stars) of M-dwarf stars they are very favorable targets to search for super-Earth planets in the HZ of their host stars (Charbonneau and Deming, 2007). Just as an example (Irwin et al., 2009), assume a M5-dwarf star (approximately $0.25M_{\odot}$ and $0.25R_{\odot}$) orbited by a planet of $7M_{\oplus}$ and $2R_{\oplus}$. In this scenario, the M-dwarf has a luminosity of approximately 1/200 that of the Sun and, consequently, a planet receiving the same stellar insolation as the Earth would lie at 0.074AU, corresponding to an orbital period of 14.8 days. Moreover, the geometric transit probability is 1.6% compared to that of the Earth-Sun system (~0.5%). In addition, the small radii of M-dwarf stars results in deeper transits (~0.5% for a M5-dwarf star, compared to the 0.03% for a solar type host) and, the combination of a smaller stellar mass and shorter orbital period, increases the semi-amplitude of the radial velocity measurements.

In order to maximize the probability of detecting a rocky super-Earth planet in the HZ, MEarth-North (Irwin et al., 2009) is photometrically monitoring a preselected sample of 2000 M-type stars compiled from the LSPM-North catalog (Lépine and Shara, 2005). MEarth-North operates 8, f/9, Ritchey-Chretien telescopes with a 25'x25' FoV each. Due to this limited FoV, this project can only monitor a single star per telescope at a time.

In comparison, the 19.4 deg^2 FoV of the TFRM combined with the fact that a 17-second exposure typically contains ~15,000 stars with a SNR greater than 5 (for stars with magnitude V<15.5 mag) and a photometric precision better than 10 milimagnitudes (3-4 milimagnitudes for V down to 13-13.5 mag), make the TFRM an excellent instrument to detect new exoplanets through the transit technique.

Since December 2011 the TFRM began to survey a selected set of fields from an input catalog of 556 M-type stars resulting from the crossmatch of the Palomar/Michigan State University Survey (Reid, Hawley, and Gizis, 1995) and a photometric catalog of bright M-dwarfs (Lépine and Gaidos, 2011). The TFRM-PSES is a able to monitor multiple fields per night,

each one containing typically around 20 M-type cataloged stars, mainly in the range of 9.0 mag < V < 15.5 mag. The cataloged TFRM fields and their current coverage status are plotted in Figure 25.

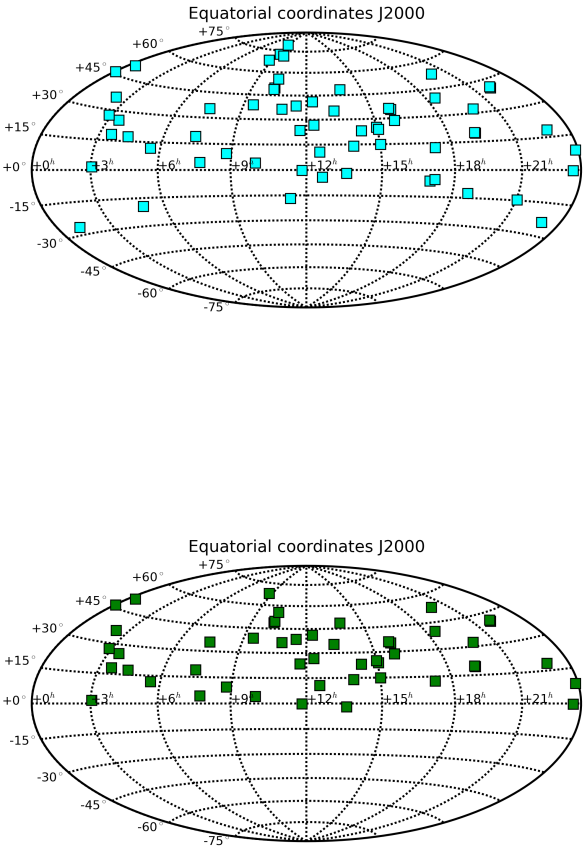


Figure 25: TFRM-PSES cataloged fields (top) and observed fields up to September 2018 (bottom).

5.1.1.4 *TFRM-PSES Pipeline and Data Reduction*

The astrometric and photometric data of the TFRM-PSES light curves are measured using the Apex II reduction pipeline (Devyatkin et al., 2010). The pipeline has been optimized to detect planetary transit signatures and consist on the following Python-based scripts:

1. *apex_superdark.py*: It generates combined master dark and bias frames from a set of individual frames. The script

divides all available dark frames into groups with identical characteristics (integration time, detector temperature and frame size) and processes each of them separately. For each group, a combined master dark is created and saved into a file containing all relevant frame parameters.

2. *apex_superflat.py*: It generates a combined master flat frame from a set of individual flat field frames. It also searches appropriate master dark frames for dark calibration. Given the big FoV of the TFRM, at the beginning of the TFRM-PSES survey sky flat frames were taken using the so-called "antisolar point" during the civil dawn and dusk (when the geometric center of the Sun is 6° below the horizon). The effects of the sky gradient present in the flat fields was minimized carefully combining a set of frames taken at dawn and dusk. Later on, in 2015, a flat field box was built to more accurately obtain gradient-free flat frames as well as to characterize the shutter map and the CCD pixel non-linearity.

3. *apex_epphot.py*: It is the main photometric reduction script of the pipeline. First, the script corrects all pixels from their intrinsic non-linearity using the following equation:

$$F_{corrected} = F + b \cdot F + c \cdot F^2 \tag{83}$$

where b and c are the matrices used for the non-linearity correction and F is the flux of the star in units of ADU.

Then, the script does bias, dark, shutter map and flat calibration of all input frames and computes the fluxes of all detections above a given SNR threshold for each frame. The photometric fluxes can be computed using the classical aperture photometry setting the *Kron aperture* and the inner and outer aperture radii. Otherwise, a background estimation can be used setting the *Kron aperture*, the default seeing and the *kernel factor*. The product of these last two denotes the characteristic non-uniformity scale of the background thus, features smaller than the size given by this product will belong to the background.

The script can discard some of the frames by a preliminary quality control that considers the minimum number of detections, the maximum allowed seeing, the minimum and maximum average intensity and the maximum ellipticity of the stars (defined as $\eta = \frac{\text{major axis}}{\text{minor axis}}$) to retain a frame.

Likewise, this script is also able to apply deblending correction to the frames. In crowded star fields, a large number of PSFs of neighboring stars can be overlapped and thus be detected as single objects. The deblending algorithm implemented in this script follows the same idea as the one in SExtractor (Bertin and Arnouts, 1996) introducing n intensity threshold levels to determine if the blended objects can be separated. This can be done using the *mask* option in which the flux from the brightest blended object is subtracted by means of an square mask centered at the source centroid. The flux subtraction on the surrounding blended objects is done using the same approach. Otherwise, a 2D-Gaussian fitting algorithm can be used instead of the square mask.

Finally, the script produces a catalog file for each input frame that contains the UTC epoch of mid-exposure and a list of detections with their corresponding (X,Y) pixel coordinates in the frame and their instrumental magnitudes.

4. *apex_epmatch.py*: This script deals with the source matching and astrometric reduction. A set of field stars is first crossmatched with a reference catalog (in our case, the UCAC4 catalog, see `http://www.usno.navy.mil/USNO/astrometry/optical-IR-prod/ucac` for more details). Then, their coordinates are first transformed and projected onto the frame and through their (X,Y) coordinates, the parameters of the chosen reduction model are determined. These parameters are then applied to the other objects in the frame and an inverse transformation is done to obtain their coordinates in the reference system of the catalog (Devyatkin et al., 2010).

   In addition, the script conducts a more exhaustive quality control by computing the magnitude zero point of each frame. This is done taking a set of reference stars and computing the median value of their magnitudes for each frame. This way, the script is able to discard those frames affected by bad atmospheric conditions, the passage of clouds or shutter malfunctions.

   On output the script produces a file which contains the raw magnitude light curves of all detections found by *apex_epphot.py* crossmatched and with their reference catalog coordinates.

5. *apex_epfind.py*: Last step of the calibration pipeline. The script produces a final file that contains the differential

magnitude (in instrumental system) light curves of all detections using the BESTRED algorithm (Voss, 2006) which, on a nightly basis, uses a set of the brightest, non-saturated and low RMS reference stars to compute the light curve correction. Once this algorithm has been applied, the magnitude residuals of each star are fit with a position-dependent polynomial and the magnitudes corrected accordingly.

The TFRM-PSES data calibration pipeline has been fully parallelized so that several frames can be calibrated, their sources detected and their raw instrumental magnitudes corrected at the same time thus reducing the CPU execution time to 60-120 minutes per night (depending on the number of frames and the crowding of the fields).

5.1.2 *The Evryscope*

5.1.2.1 *Project Overview*

The Evryscope ("wide-seer") (Law et al., 2014; Law et al., 2015; Law et al., 2016; Ratzloff et al., 2016) consists of a single hemisphere containing 24 (expandable to 27) 61mm-aperture telescopes simultaneously imaging the sky with a two minute cadence. The telescopes each use a rectangular 28.8 MPix KAI-29050 interline-transfer CCD and each telescope assembly rotates in a circular arc around the pole facing camera as the hemispherical housing tracks the sky. The Evryscope is designed to be sensitive to exoplanetary transits and other short timescale events not discernible from existing large-sky-area astronomical surveys. The Evryscope, tracks the sky on a standard German Equatorial mount, imaging a 691 MPix instantaneous 8,000 sq. deg. FoV, in two hour sections, ratcheting back after each section to cover a new sky area (with 75% overlap with the previous area). The system takes an exposure in g-band every 2 min exposure with 97% duty cycle efficiency, reaching a limiting magnitude of g~16, and co-adds over hours each night to monitor fainter objects. The Evryscope is located at Cerro Tololo Inter-American Observatory (CTIO), covering declinations between -90^{o} and $+10^{o}$.

The Evryscope operates fully robotically and it is streaming imaging data each night at 104MBit/sec, generating 50-100 TB/year of calibrated images. It is delivering multi-year 1%-precision 2-minute-cadence light curves for ~35,000 observations of every star brighter than g~16 and milimagnitude pre-

Figure 26: The Evryscope deployed at CTIO. The main dome is 6 feet in diameter. The system is mounted inside an AstroHaven dome maintained as part of the PROMPT telescope network.

cision 16-minute-cadence light curves for every object brighter than g~12.

Currently, the Evryscope has the largest ètendue of any ground-based survey. For bright-stars high-cadence regime, it is the only survey within 10% of the enormous planned Large Synoptic Survey Telescope (LSST) ètendue.

A Northern-Hemisphere Evryscope is to be deployed during 2018 in collaboration with Mount Laguna Observatory and San Diego State University. Together, the two Evryscope systems will provide true all-sky, high-cadence and multiband coverage of bright targets.

5.1.2.2 *Evryscope's Science Plans*

The Evryscope opens a new parameter space in optical astronomy providing short cadence and high precision coverage of the entire accessible sky. In addition, the telescope design provides limiting magnitudes that allow both galactic and extragalactic events to be detected. The Evryscope's dataset will thus enable real time transient detection, variable phenomena and wide-field transiting exoplanet characterization.

Given that the Evryscope has an order-of-magnitude more FoV than the next-largest exoplanet surveys, four transiting planet projects are currently being developed and conducted:

1. White-dwarf survey for transiting asteroid-size exoplanets.
2. Transiting rocky exoplanets search in the habitable zone of nearby M-dwarfs.
3. Transiting exoplanets search around nearby, bright stars.
4. Monitor large populations of eclipsing binaries and detection of eclipse timing variations induced by orbiting planets.
5. Monitoring nearby stars for microlensing events.
6. Transiting Exoplanet Survey Satellite (TESS) precursor observation providing long-term monitoring of TESS targets. The Evryscope will also increase TESS long period giant planet yield by recovering multiple transits from objects detected as single events in the TESS 27-day search period.
7. Monitoring of flares in active stars hosting planets.

In addition, the Evryscope will monitor the fluxes of millions of stars with a two-minute-cadence, building a multi-year database that will allow the detection and characterization on a huge number of variable stars and variability types.

Given that all the observed data is recorded into the database, the Evryscope will be able to provide pre-, in- and post-explosion optical counterpart data for rapid transients such as novae, supernovae, gamma-ray bursts, fast radio bursts, and gravitational waves counterparts.

5.1.2.3 *Evryscope Pipeline and Data Reduction*

The 0.8 GB/min of raw images generated by the Evryscope are reduced in real time using a fully parallelized framework written in Python backed by a PostgreSQL 9.5 database server (Law et al., 2016; Corbett et al., 2016). Each science frame is corrected from pixel non-linearity and a mask for bad-pixels is applied. Then they are masterdark and masterflat calibrated and then checked by a Data Quality Daemon (DQD) that takes into account the integrity of each 28.8 MPix FITS per camera file, the Sun elevation, the effective CCD calibration, the minimum and maximum median value of the pixels, the minimum number of detected stars in the frame, the median of the detected stars elongation, and others. Those frames that pass the DQD are astrometrically calibrated with a custom-made solver (Law et al., 2016). This consist on a third-order distortion polynomial solution (in x, y and radial terms) represented as a TPV

convention (Shupe et al., 2012). After careful background modeling and subtraction, forced-aperture undetrended photometry is extracted based on known source positions in a reference catalog, such as APASS-DR9 (Henden et al., 2009; Henden et al., 2010; Smith, Henden, and Terrell, 2010; Henden et al., 2016). Light curves are then generated for approximately 15 million sources across the Southern sky by differential photometry in small sky regions using carefully-selected reference stars and a range of apertures. These equal-area sky regions are DR9 indexed with PostgreSQL's Q3C (Koposov and Bartunov, 2006) and partitioned with HEALPix (Górski et al., 2005). In extremely crowded fields the pipeline is run for particular targets, optimizing the aperture sizes to avoid nearby stars.

The database can be queried to return a target light curve and a set of reference stars around it. The differential photometry calibration procedure for these light curves first rejects frames with very low number detections or with anomalous magnitude zero points. In order to correct atmospheric extinction and other systematics, all light curves have differential photometry applied using the BESTRED algorithm (Voss, 2006) in per two hour ratchet, per night, and per whole light curve timebase basis. Once the differential photometry has been applied, the magnitude residuals of the reference stars are fit with a second order polynomial to account for slowly-varying position-dependent effects.

5.2 SPACE TELESCOPES

5.2.1 *CoRoT Space Observatory*

5.2.1.1 *Project Overview*

The COnvection ROtation and planetary Transits (CoRoT) space mission (Fridlund et al., 2006) was the first space project designed for exoplanet research. The mission was led by the Centre National d'études Spatiales (CNES) together with the European Space Agency (ESA) and other international collaborators. Its two main objectives were the study of stellar seismology, detecting and measuring star vibrations, and the search for exoplanets (specially terrestrial ones). The spacecraft was launched on December 27, 2006 from the Baikonur Cosmodrome in Kazakhstan, Russia, atop a Soyuz 2.1b rocket. It began collecting science data on February 2007 beginning its 2.5 year mission. Although its operations were extended up to 2013, on November 2012 CoRoT suffered a computer failure that made it

impossible to retrieve data from the telescope. After some repair attempts it was announced that CoRoT would be retired and decommissioned. Finally, it was de-orbited on June 17, 2014.

During this period, CoRoT photometrically monitored 163,665 targets distributed over two opposite regions in the Galactic plane in 26 fields.

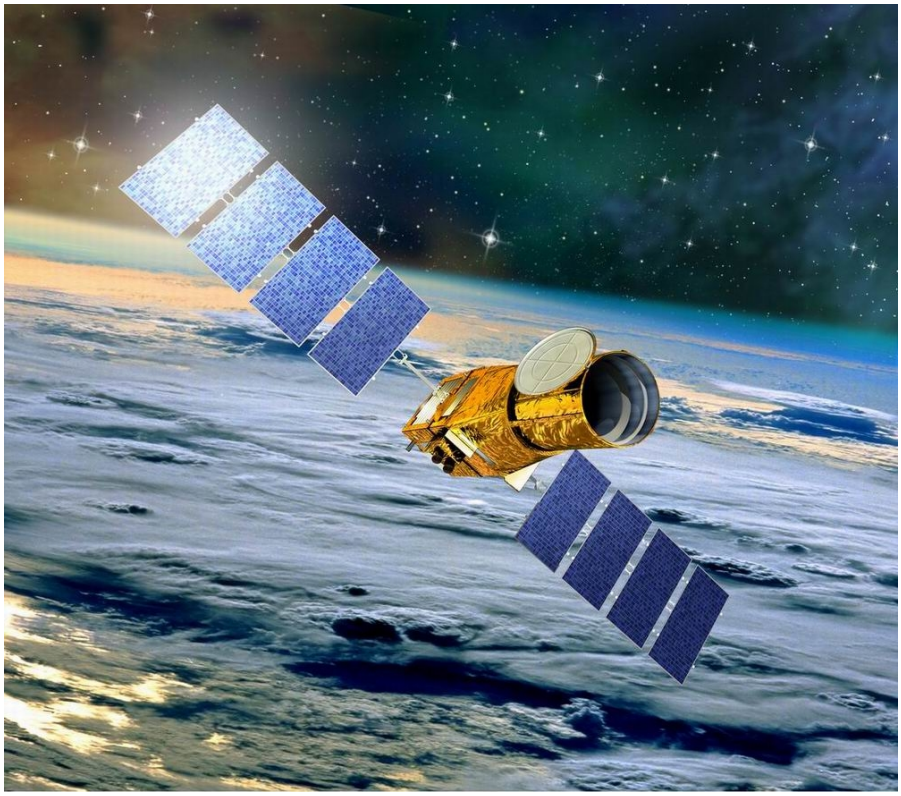


Figure 27: Artist's view of the CoRoT satellite. Credit: CNES - Mai 2004/Illus. D. Ducros

5.2.1.2 *The CoRoT exoplanet mission*

The CoRoT exoplanet program aimed at detecting the presence of extrasolar planets when they transit in front of their parent star. CoRoT was equipped with four CCDs with a $1.3 \times 1.3^{\circ}$ FoV. Given the nature of the CoRoT objectives, two of them were dedicated to the astroseismology observations and, the other two, to the exoplanet mission. In the case of the exoplanet channel, the satellite's processing capacities allowed for the observation of ~6,000 stars per CCD. For each star, a pixellised photometric aperture was assigned at the start of each run to optimize the SNR of the integrated flux (Llebaria and Guterman, 2006). For these stars, all the photometry was done on board the satellite and only the light curves were downloaded to Earth. The targets in the exoplanet channel have magnitudes

$11 \leqslant r \leqslant 16$ though brighter ones were also observed. In addition, a prism was located in the optical path of CoRoT's exoplanet channel to obtain a spectrum on the focal plane. The photometric apertures for stars brighter than r = 15 were divided into three regions that approximately coincided with the red, green and blue parts of the spectrum resulting in three-color light curves being transmitted to Earth. For fainter stars, only white light curves are available.

The fields accessible to CoRoT, known as the CoRoT continuous viewing zones (CVZs), were two circles of ~10° located in the ecliptic plane and centered at zero declination and RA 6^h50^m and 18^h50^m, respectively. The telescope switched from one zone to the other twice a year allowing the CVZs to be observed continuously during approximately six months. The position in the sky of all the 26 CoRoT fields observed for the exoplanet mission is shown in Figure 28. As can be seen, some targets were observed more than once due to the overlapping that existed between successive fields. The targets observed in these fields were selected to maximize the number of main-sequence stars with spectral types F or later. Thus, the spectral type and luminosity class of all CoRoT targets in the magnitude range $11 \leqslant r \leqslant 16$ needed to be estimated. The spectral classification used ground-based multi-colour photometric observations that were carried out prior to the instrument launch. The `Exo-Dat` database was built in order to provide as much stellar information as possible in the FoV of the CoRoT's exoplanetary science program (Deleuil et al., 2009).

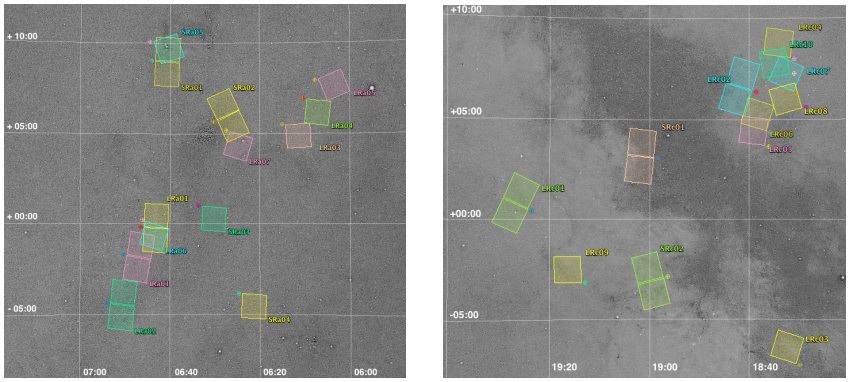


Figure 28: Position of the fields observed by CoRoT for the exoplanet mission in the direction of the Galactic anti-center (left) and center (right) as seen in Deleuil et al. (2018).

At first, the observations consisted on one, 140 days, long run (LR) and a shorter one (SR), lasting 20 to 30 days, every 6 months. The duration and number of pointings per year were flexible and evolved to take into account the scientific require-

ments of the exoplanet and astroseismology missions. In March 2009, one of the Data Processing Units broke down causing the loss of one CCD in each of the exoplanet and seismology channels. Since then, the observation strategy changed to a two intermediate-duration runs per half year. Apart from compensating the loss of star counts per pointing, it also allowed to observe again the same field after some time. The possibility of re-observing the fields allowed, for example, to more precisely determine the planetary parameters of some of the CoRoT confirmed exoplanets by scheduling simultaneous observations with other instruments like HARPS (Barros, S. C. C. et al., 2014) or Spitzer (Bonomo et al., 2017). For more information about the fields, runs and the CoRoT mission itself, the author points to Fridlund et al. (2006), the "CoRoT Legacy Book" (CoRot Team, 2016), Deleuil et al. (2018) and the CoRoT data archives[1] for a more extensive review.

5.2.1.3 *CoRoT data reduction pipeline*

In CoRoT, there was no official pipeline at mission level for transit detection and light curve analysis. Once the light curves of a given run were released, their analysis was done in parallel by different teams, using different algorithms. Given that the transit search was performed as soon as the runs were released, the transit detection algorithms evolved significantly from the first to the last run. Erikson, A. et al. (2012) summarizes all the transit search methods implemented for the CoRoT runs. In addition, several ways of improving the pre-filtering and detrending of the light curves have been proposed. For example, SARS (Ofir et al., 2010): a modification of the Sysrem algorithm (Tamuz, Mazeh, and Zucker, 2005); EXOTRANS (Grziwa, Pätzold, and Carone, 2012): a BLS-based algorithm; a detrending method based on a wavelet coefficient thresholding (Grziwa, Korth, and Pätzold, 2014), and the one from Bonomo et al. (2012).

5.2.2 *Kepler Space Observatory*

5.2.2.1 *Project Overview*

Kepler is a space observatory launched on March 7, 2009 into an Earth-trailing heliocentric orbit. The Kepler mission was designed to detect transits of Earth-size planets in the HZ of F through M type dwarf stars with magnitudes in the $9<m_v<15$

1 https://corot.cnes.fr/en/data-archives

range (E. and A, 2016). Kepler had a fixed FoV in the constellations Cygnus and Lyra (see Figure 29) centered on (α, δ) = (19:22:40, +44:30:00) (l = 76.32°, b = +13.5°). Kepler's sole scientific instrument is a photometer that continually monitors the brightness of ~150,000 stars. This photometer, a 0.95 m aperture Schmidt telescope feeds the 94.6 million pixel, 42 CCD, 21 modules detector array containing both science and Fine Guidance Sensor CCDs. Each 50x25 mm CCD has 2 outputs and 2200x1024 pixels and a pixel scale of 3.98 arcseconds. The half-maximum bandpass is 435 to 845 nm. Kepler's FoV has a diameter of 16.1°, of which 115.6 square degrees are covered with active pixels. The four Fine Guidance Sensors are mounted in the corners of the science array and were used to attain the <0.009 arcsec pointing stability on $\geqslant$15 min scales. The interval between reads of a given pixel of a CCD is composed of an exposure time of 6.02 seconds and a fixed readout time of 0.52 seconds. For each integration, all pixels were read-out and temporally coadded in the Science Data Accumulator (SDA). The data volume was reduced as target lists determined which target and calibration pixels were read-out of the SDA and transmitted to the Solid State Recorder (SSR) for later down link. Kepler acquired data at 29.424-minute intervals (270 integrations) for all target stars called long cadence (LC) targets. A set of 512 bright, non-saturated, relatively isolated stars were observed at 1-min intervals (9 integrations) and called short cadence (SC) targets. The resulting flux time series for both LC and SC targets were typically >90% complete. The observations were organized into seventeen 93-day "quarters" allowing the telescope to rotate by 90° to keep the Sun on the solar panels and the radiator pointed to deep space (Haas et al., 2010).

There were three main components of the Kepler observing program:

- Planetary transit search: measuring the flux of ~150,000 stars to identify candidates for planetary transits using precision photometry.
- Characterization of transiting systems: to confirm the nature of the transit, additional analysis of the candidate system and its neighborhood need to be conducted including: precise centroiding, high spatial resolution imaging from ground- and space-based telescopes, precise stellar radii determination or measurement of stellar variability.
- Guest observers: some targets in Kepler's FoV were selected for guest observers.

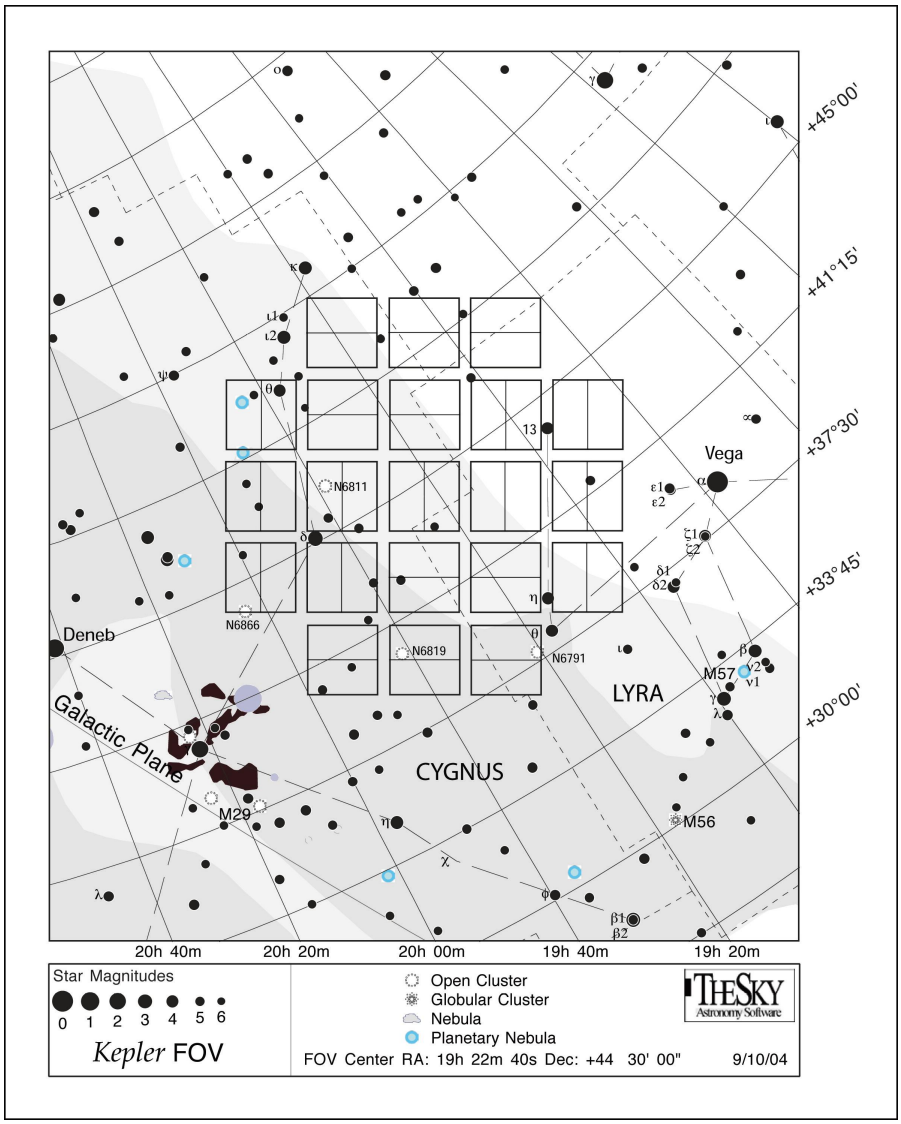


Figure 29: Kepler FoV in the constellations Cygnus and Lyra. Squares denote the outlines of individual CCD modules projected on the sky.

The Kepler mission lasted four years between May 2009 and May 2013. Light curves of target stars observed during this period have yielded 2,327 confirmed exoplanets so far and, more than 2,000 candidates remain to be confirmed([2]). Kepler discoveries comprises among others, the discovery of Kepler-10 b, the first confirmed terrestrial planet outside the Solar System (Batalha et al., 2011); Kepler-16 b, the first, unambiguous detection of a circumbinary planet (Doyle et al., 2011) or Kepler-37 b, the smallest, almost Moon-size exoplanet orbiting a main-sequence star (Barclay et al., 2013). For more information regarding the Kepler mission, the instrument itself and the Kepler archive, the author refers to Borucki et al. (2010), E. and A

2 https://exoplanetarchive.ipac.caltech.edu

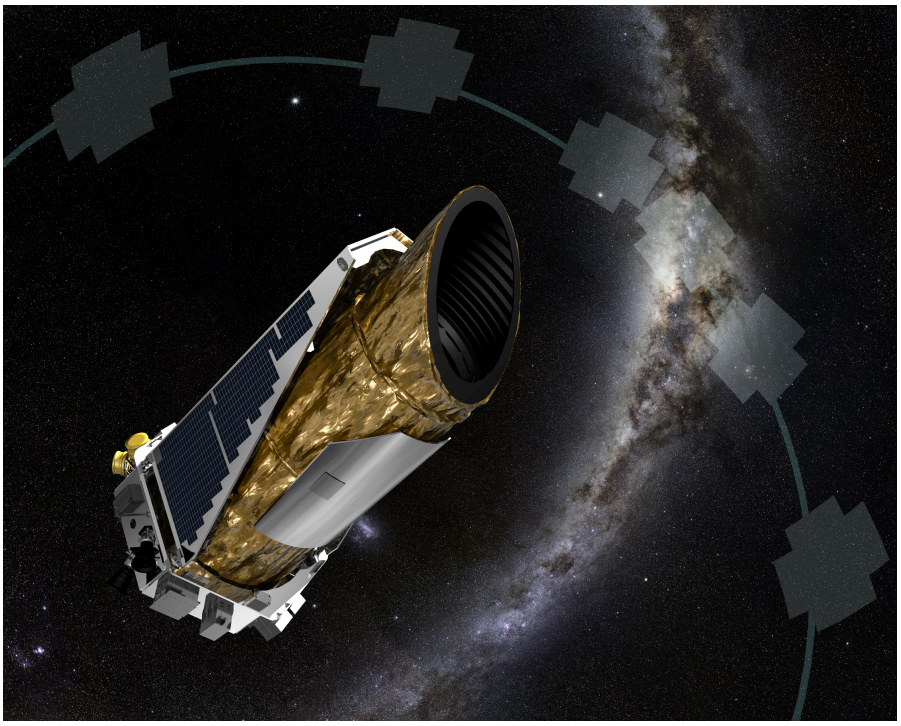


Figure 30: Artist's concept of NASA's Kepler space telescope. Image Credit: NASA.

(2016), J. E. Van Cleve and Zamudio (2016) and Thompson S. E. and A. (2016).

5.2.2.2 *The K2 Mission*

The initial planned lifetime for the Kepler mission was 3 years. In 2012, the mission was to be extended until 2016 but, on July 14, 2012, one of the spacecraft's four reaction wheels stopped turning. In order to ensure the pointing accuracy and thus, meeting the mission objectives, the remaining reaction wheels needed to keep working. On May 11, 2013, a second reaction wheel failed ending the primary mission data collection and threatening the continuation of the mission. Developed in the months following this failure, the K2 mission (Howell et al., 2014), represented a new concept for Kepler's operations given the spacecraft's ability to maintain pointing in all three axes with only two reaction wheels. This operation mode, that started in October 2013 and became fully operational in May 2014, provides enough fuel for an extended 2-3 year mission. K2 began a series of sequential campaigns observing a set of independent target fields distributed around the ecliptic plane. This way, K2 minimizes the torque exerted on the spacecraft by solar wind pressure, reducing the pointing drift and achieving a photometric precision close to the original for the Kepler mis-

sion (Vanderburg, 2014; Vanderburg and Johnson, 2014). The duration of each observing campaign is limited by solar illumination to a duration of about 75 days.

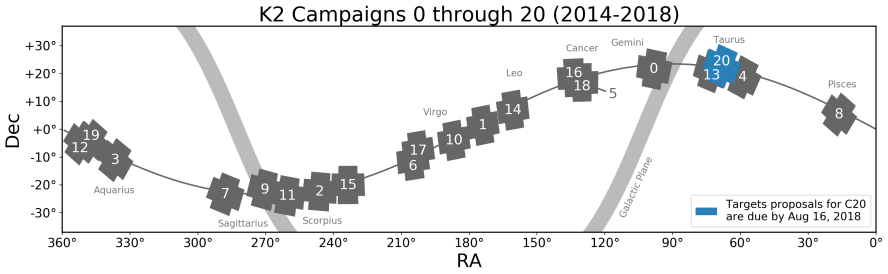


Figure 31: K2 campaigns 0 to 20 FoVs centered around the ecliptic plane. Image Credit: NASA.

The K2 mission observes ~10000 targets per FoV collecting data at 30-minute and 1-minute cadences (Howell et al., 2014). Due to its multi-field, ecliptic-pointed operation mode, it expanded Kepler's opportunities for new discoveries in Galactic regions not previously observed. K2 is a community-driven mission were all targets are peer-reviewed proposals done by the community through the Guest Observer program. Given that there is no predefined set or type of targets, the mission science goals are diverse; some of them are listed below[3] (Howell et al., 2014):

- Observation of extrasolar planets around bright stars: K2 inherits Kepler's mains objective to measure the occurrence rate of planets around Sun-like stars. More specifically, the detection of transits of hot planets around bright stars to enable precise Doppler spectroscopy to provide planetary masses and densities and to facilitate improvements in the characterization of exoplanet atmospheres.
- Observation of extrasolar planets around low-mass stars: identify characteristics of potentially-habitable planets around bright M-dwarfs in the solar neighborhood for approximately 4,000 M dwarfs brighter than m_v=16 per K2 field.
- Observation of open clusters: study of the incidence of large and small planets and to discover and characterize binary stars within open clusters and stellar associations.
- Asteroseismology to advance in the understanding of stellar evolution, stellar interiors and stellar populations.

3 https://keplerscience.arc.nasa.gov/objectives.html

- Observation of variable extragalactic sources: observe in statistically significant numbers bright AGNs to provide robust testing of current models or to identify the progenitors of Type Ia supernovae.
- Micro-lensing observations: combined with ground-based observations to obtain host and planet masses from the measured parallaxes.

The difference between an exoplanet candidate and a confirmed exoplanet is very important. During the Kepler mission, an exoplanet candidate was a any transit signal that had passed a set of astrophysical false-positive and instrumental false-alarm tests. In K2, a candidate is any signal that a given team has identified as a possible planet. A validated planet is a candidate that has been vetted with follow-up observations and determined quantitatively to be more likely an exoplanet than a false positive. Validated planets, are more promising targets than planet candidates for follow-up observations, characterization, and eventual confirmation. Confirmation is more rigorous than validation and is usually attained through a reliable mass determination using RV observations (Mayo et al., 2018).

On December 18, 2014, the K2 mission detected its first confirmed exoplanet, a super-Earth named HIP 116454 b or K2-2 b (Vanderburg et al., 2015). Since then, the K2 mission has found 325 exoplanets and 493 candidates are yet to be confirmed.

5.2.2.3 *K2 data reduction pipeline*

The data recorded by Kepler on orbit is downlinked, archived and ends up at he Space Science Division of NASA's Ames Research Center. All science data is processed and calibrated by means of the Kepler Science Pipeline. The main elements of the pipeline are summarized below (Jenkins et al., 2010):

- Pixel calibration (CAL): the first step in the Kepler pipeline is performed by the CAL module (Jenkins et al., 2010; Quintana et al., 2010). It corrects the raw Kepler photometric data into calibrated pixels prior to the extraction of photometry and astrometry. CAL module is responsible of bias level, dark current, smear (caused by the lack of shutter in Kepler's photometer), nonlinearity, undershooting and flat field corrections as well as cosmic ray removal in the bias and smear measurements.
- Photometric Analysis (PA): this module (Bryson et al., 2010) delivers photometric light curves from the calibrated pix-

els defined to contain the optimal aperture and its associated background pixels. The optimal apertures criteria tries to maximize the signal-to-noise ratio of the light curves and, also takes into account the pixel response function across the focal plane. PA module performs barycentric time correction, "Argabrightening" event (transient increases in the background flux due to an unknown source, probably due to small dust particles (Jenkins et al., 2010)) detection, cosmic ray cleaning and background signal removal. It is also responsible of the aperture photometry, detecting the photocenter of each source as well as providing the astrometric solution for each pixel.

- Pre-search Data Conditioning (PDC): the PDC's task (Twicken et al., 2010) is to examine the light curves produced by PA and remove systematic errors from them. These systematics include pointing errors, focus changes and thermal effects on the instruments. The "Pre-search" keyword refers to conditioning the light curves before executing a transit search. PDC conducts data anomaly flagging, identifies and corrects of discontinuities, identifies variable stars and astrophysical events, corrects target excess flux coming from nearby sources and performs systematic error correction for quiet and variable stars.

Once the data is processed, it is formated into FITS files and exported to the MAST archive[4] at the Space Telescope Science Institute. The archived data includes the raw and calibrated pixel values for all sources, background pixels, calibrated and corrected light curves, and related engineering data.

In the case of K2, a part from the PDC module, several detrending and systematic removal algorithms have been developed: Gaussian process-based ones like `K2SC` (Aigrain, Parviainen, and Pope, 2016); correlation-based like `K2SFF` (Vanderburg and Johnson, 2014), and `EVEREST 2.0`[5] (Luger et al., 2016; Luger et al., 2017). The latter is the most accurate algorithm for removing instrumental systematics from K2 light curves. `EVEREST 2.0` is an open-source pipeline that uses a combination of pixel-level decorrelations to remove spacecraft pointing errors and Gaussian processes to capture astrophysical variability. The corrected light curves have a precision comparable to the original Kepler mission for targets brighter than Kp = 13, and within a factor of 2 for fainter targets .

4 https://archive.stsci.edu/index.html
5 https://github.com/rodluger/everest

6 APPLICATION OF TFAW TO ASTRONOMICAL DATA

6.1 GROUND-BASED DATA

6.1.1 *TFRM Data*

6.1.1.1 *Data description*

The data from the TFRM-PSES survey used for the TFAW performance assessment comprises 2,048 data points from 30 nights observed during 2013, 2014 and 2015 for a field centered at $(\alpha,\ \delta)$ = (10:14:44, +48:30:00). All light curves in the data set have been observed with a 475 nm cutoff frequency glass filter (Schott GG475), 17 sec exposure time and an altitude above 50°. The light curves were generated with the APEX-2 reduction pipeline (Devyatkin et al., 2010). The pipeline includes aperture photometry, atmospheric extinction correction, differential photometry (using the BESTRED algorithm from Voss (2006)) and astrometric calibration (as explained in Section 5.1.1.3).

6.1.1.2 *Results*

In Figure 32 and Figure 33 we present two examples obtained from the TFRM-PSES data. Figure 32, 2MASS J10144313+5018191, is a newly-discovered variable star, detected both with TFA and TFAW, probably a δ-Scuti star with a 0.1592-day period. As can be seen, after TFAW is applied, the noise contribution decreases significantly and outliers are efficiently removed. The LS power spectrum during the *frequency analysis* step greatly improves with respect the TFA one presenting higher power of the peaks and an increase in the SDE (10.1 for TFAW in front of 9.8 for TFA). The one in Figure 33, a W Ursae Majoris-type variable star with a 0.371018-day period (the maximum peak in the LS power spectrum corresponds half the cataloged period) is cataloged as NSVS 4921994 (Woźniak et al., 2004). As in the previous case, the SNR of the signal is improved and the correct period is recovered from the TFAW LS power spectrum. Also, the power spectrum itself presents higher peaks and an increased SDE (10.31 for TFAW in front of 10.1 for TFA).

The small increase in the SDEs of both example light curves for TFAW compared to TFA is to be expected from Figure 13 as they are in the high SNR regime.

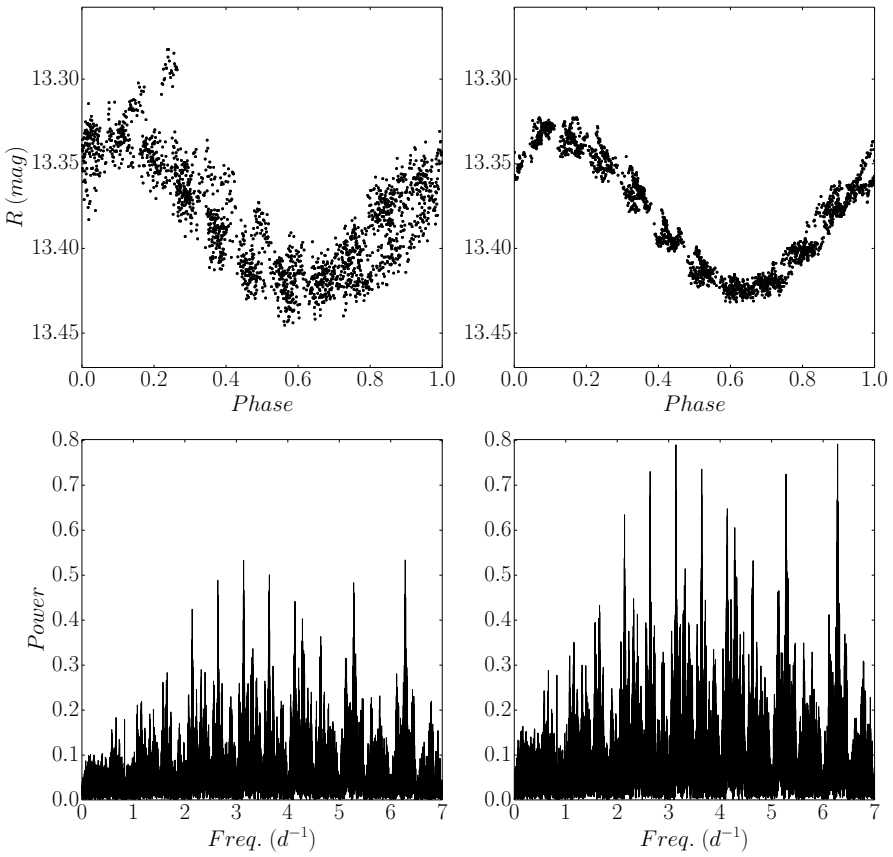


Figure 32: Example of the TFAW filtering capabilities on the observed TFRM-PSES light curve of 2MASS J10144313+5018191, a newly-discovered variable star. Same notation and TFAW parameters as Figure 15.

6.1.1.3 *TFAW quantitative performance over TFRM light curves*

Tests made on real light curves show that after TFAW is applied, signals affected by systematics and noise are recovered with no shape and period distortion. Top panel of Figure 34 shows the TFRM-PSES photometric precision (standard deviation σ) vs. R magnitude for both TFA and TFAW light curves. The set of 6,485 displayed light curves correspond to the same field, observed over three years. TFAW photometric performance is not only consistently better than TFA over all the R magnitude range, but also shows a closer fit with the expected blue-lined trend-free stochastic noises performance, especially the faint end. Figure 34 bottom panel shows, similarly to Figure 5 in Kovács, Bakos, and Noyes (2005), the decrease in the standard devia-

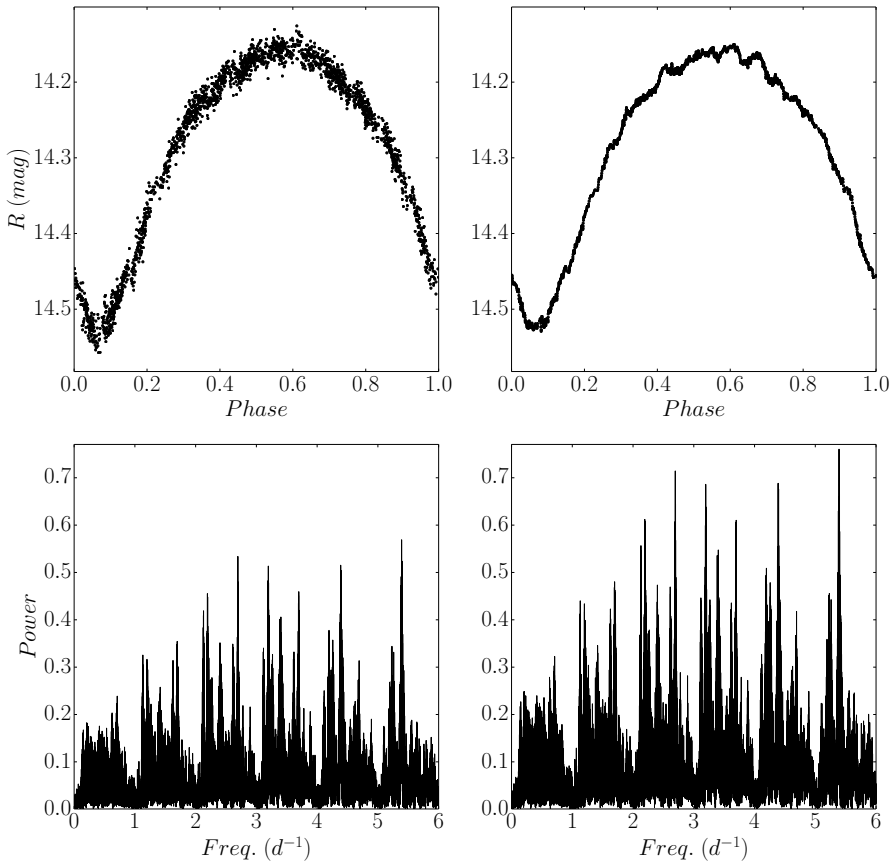


Figure 33: Example of the TFAW filtering capabilities on the observed TFRM-PSES light curve of NSVS 4921994. Same notation and TFAW parameters as Figure 15.

tions of the same sample of 6485 TFAW-filtered TFRM-PSES light curves compared to the original TFA ones. It can be seen, that for almost all the TFA standard deviation range (conversely brighter to fainter magnitudes), the standard deviations of TFAW light curves is ~40% better than that for TFA light curves. In the case of low standard deviations (i.e. bright stars), the improvement is smaller because the noise contribution is also smaller. The presence of real variable stars in the sample could also explain a fraction of such TFA and TFAW with similar standard deviations, as their intrinsic variabilities dominate the noise contribution (see Figure 33 as an example).

The results presented in Section 6.1.1.2 in the LS power spectra demonstrate that TFAW does not introduce false periodicities or eliminate any of the signal peaks during the frequency analysis step. For the examples presented before as well as for the rest of TFRM-PSES light curves analyzed, TFAW outlier removal and frequency analysis step results in an overall improvement of the power spectrum and of the SDEs.

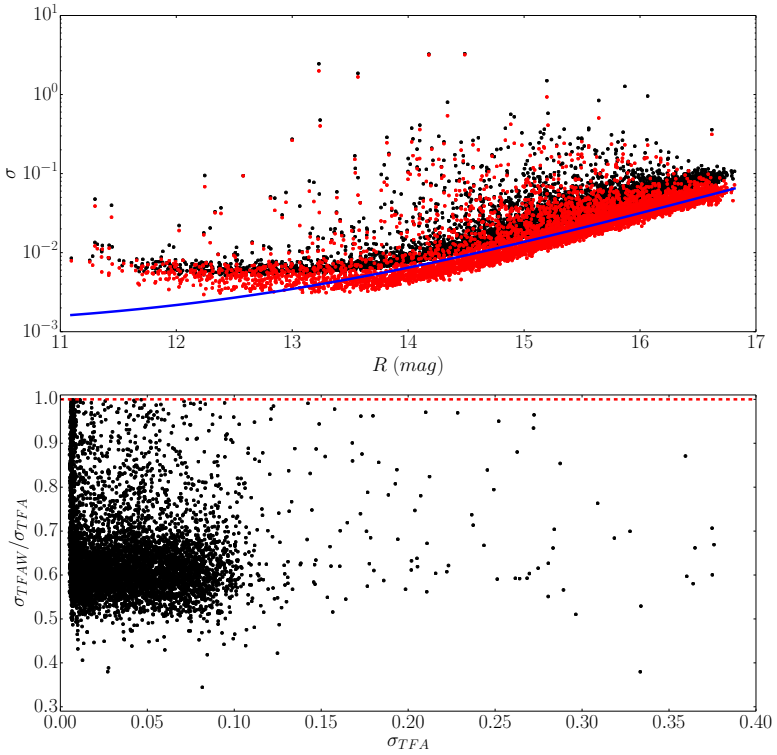


Figure 34: **Top:** Standard deviation vs. R magnitude for TFA (black dots) and TFAW (red dots) of a set of 6,485 TFRM-PSES light curves. Blue line corresponds to the sum of the scintillation, photon, background and read-out noises. **Bottom:** Decrease of the standard deviations for the same set of light curves due to the application of TFAW compared to TFA. Both TFA and TFAW results were obtained using 250 template stars. Red line corresponds to the zero correction level.

Regarding the number of template stars, tests run both with simulated and real light curves show that using 250 template stars gives the best results. TFAW can also be run with less stars (<10 stars). However, the noise filtering is less efficient and some trends and systematics could still be present in the filtered data. As with the original TFA, it is not necessary to compute the template matrix for each target star, only in the cases in which the latter accidentally coincides with one of the template ones.

6.1.1.4 *TFAW computing performance for TFRM data*

The algorithm has been fully implemented and parallelized in Python for the TFRM-PSES dataset. The CPU time needed for a 2,048 data points, single light curve TFAW analysis using 250 template stars is typically around 60 seconds and about 100 iterations. A full run of the TFAW on 6,485 TFRM-PSES light curves

with 2,048 points per time series, 250 template stars and 90,000 BLS frequency steps takes 7.5 hours of 24 × CPUs @ 2.00 GHz. As a comparison, TFA runs the same set of light curves with the same parameters and hardware capabilities in 0.5 hours, i.e., 15× less than TFAW. The main reason for this performance loss is the extra time in the computation of the SWT and the ISWT at each iteration step. The current `PyWavelets` implementation we are using for these transforms is coded in `Cython`. Future effort is planned in order to significantly speed up TFAW.

6.1.2 *Evryscope Data*

6.1.2.1 *Data description*

The Evryscope data used for TFAW performance assessment comprises 10,240 data points observed from January 10, 2016 to February 26, 2016 with one of The Evryscope telescopes centered around the South Celestial Pole. Prior to the TFA and TFAW analysis of the data, all target light curves have been processed using the pipeline and data reduction method described in Section 5.1.2.3. To build the set of 250 stars template stars needed by TFA and TFAW we selected those light curves with low standard deviations and small Stetson's L (Stetson, 1996) indexes. In addition, although the pipeline fills any missing value from the light curves to ensure that all of them have the same time-base, only those stars with less than 10% initial missing values are selected as possible template stars.

6.1.2.2 *Results*

2MASS J13190996-8347115, located at $(\alpha,\ \delta)$ = (13:19:09.9669, -83:47:11.5332) (Gaia Collaboration, 2018), is a semi-detached contact binary of the W Ursae Majoris-type with a 0.658-day period. It was detected during visual inspection of the pipeline-processed Evryscope light curves in 2017. At that moment, the target was not classified as variable in the VSX or GCVS. Its period and variability type was later confirmed by the ASAS-SN survey in Jayasinghe et al. (2018). TFA and TFAW recover the variability period reported by Jayasinghe et al. (2018) during the *frequency analysis* step. After the signal has been iteratively reconstructed, the TFAW-filtered light curve is significantly less noisy than the TFA one and has its outliers efficiently removed. The shape and depth of the signal has not been modified after applying TFAW's SWT-based iterative filtering.

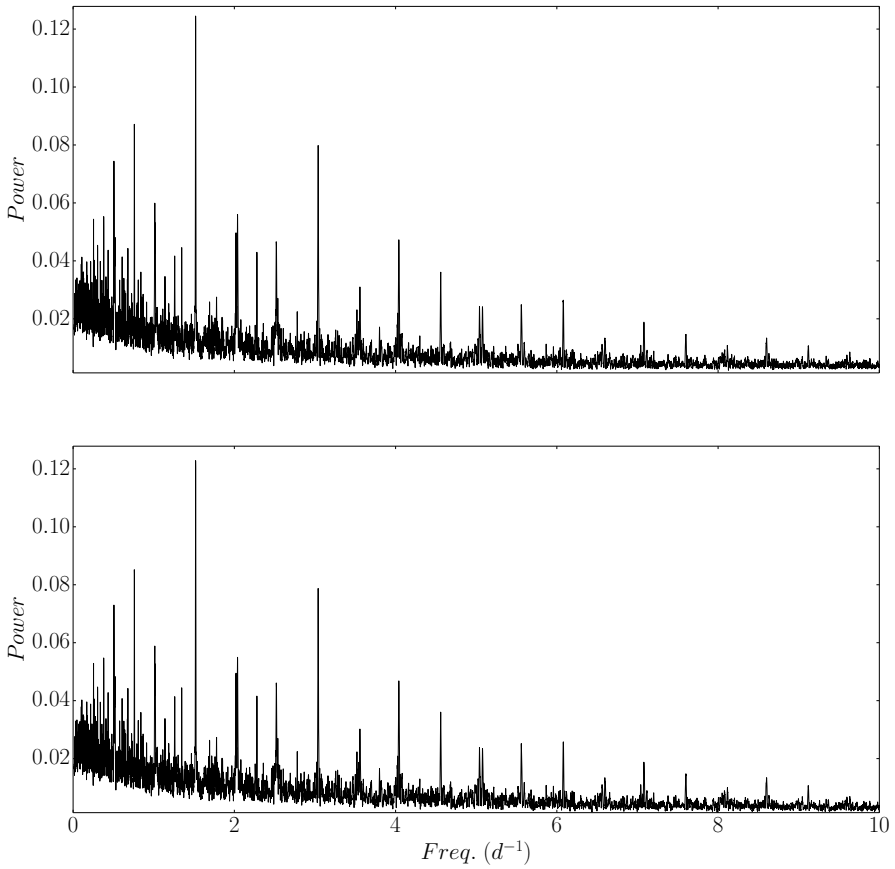


Figure 35: **Top**: TFA's BLS power spectrum for eclipsing binary 2MASS J13190996-8347115 observed with The Evryscope. **Bottom**: TFAW's BLS power spectrum. The peak at ~1.52 d^{-1} in both plots corresponds to the spin period reported by Barstow et al. (1995).

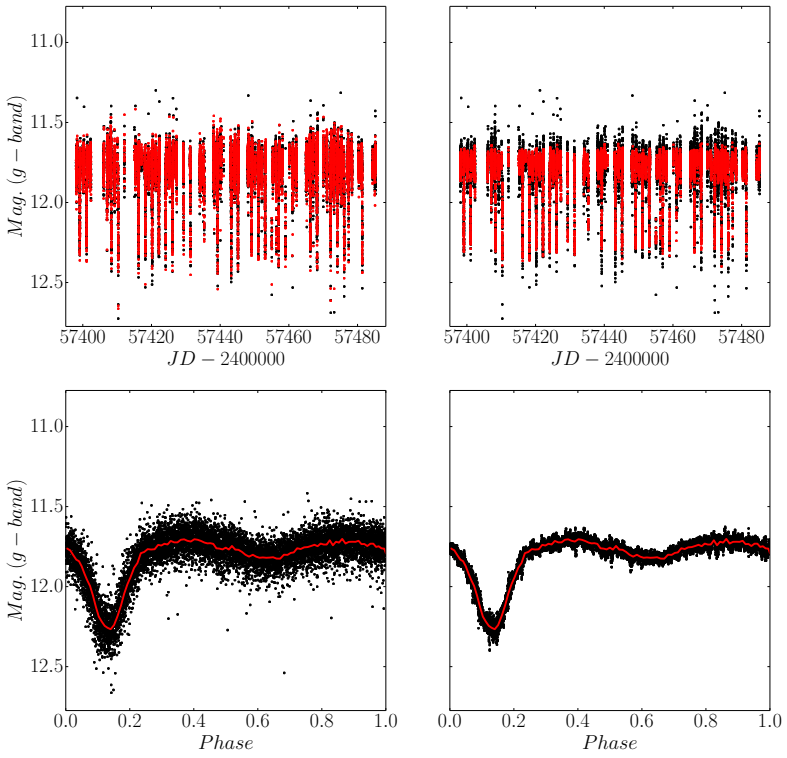


Figure 36: TFAW noise and systematics filtering capabilities for eclipsing binary 2MASS J13190996-8347115. **Top left**: Normalized raw (black dots) and TFA-filtered (red dots) light curves of 2MASS J13190996-8347115. **Top right**: Normalized raw (black dots) and TFAW-filtered (red dots) light curves of 2MASS J13190996-8347115. **Bottom left**: Phase folded, to twice the period, TFA-filtered light curve of 2MASS J13190996-8347115. Red line corresponds to the 300 binned phase folded light curve. **Bottom right**: Phase folded, to twice the period, TFAW-filtered light curve of 2MASS J13190996-8347115. Red line corresponds to the 300 binned phase folded light curve.

RE J0317-853, located at (α, δ) = (03:17:15.8452, -85:32:25.5611) (Gaia Collaboration, 2018), is one of the four spotted white dwarfs with rotation periods shorter than 1 hr (Hermes et al., 2017). This star spins every 725 sec, with 0.2 mag peak-to-peak photometric variations (Barstow et al. (1995), Ferrario et al. (1997), Vennes et al. (2003)). Along with spots, this target exhibits a strong magnetic field (340-450 MG) (Kilic et al., 2015), the latter suggesting a binary merger origin (García-Berro et al., 2012).

The brightness of this white dwarf is close to The Evryscope two-minute cadence limiting magnitude, with expected photon-limited photometric around 10%. In any case, both TFA and TFAW find the spin period reported by Barstow et al. (1995) during their *frequency analysis* steps (see Figure 37). Given that we had an a priori knowledge of the sinusoidal nature of RE J0317-853's variability, we have employed LS (Scargle, 1982) to run

the period search. For the TFAW's *iterative reconstruction* step, we set the *noise level* to the lowest SWT decomposition scale of the phase folded light curve and, compute the *signal level* using the method explained in Section 3.3.1. Figure 38 shows a comparison between TFA and TFAW reconstructed signals. TFAW has been able to greatly improve (by a factor ~3×) the SNR of the sinusoidal variability by filtering the high frequency noise and removing most of the outliers present in the light curve.

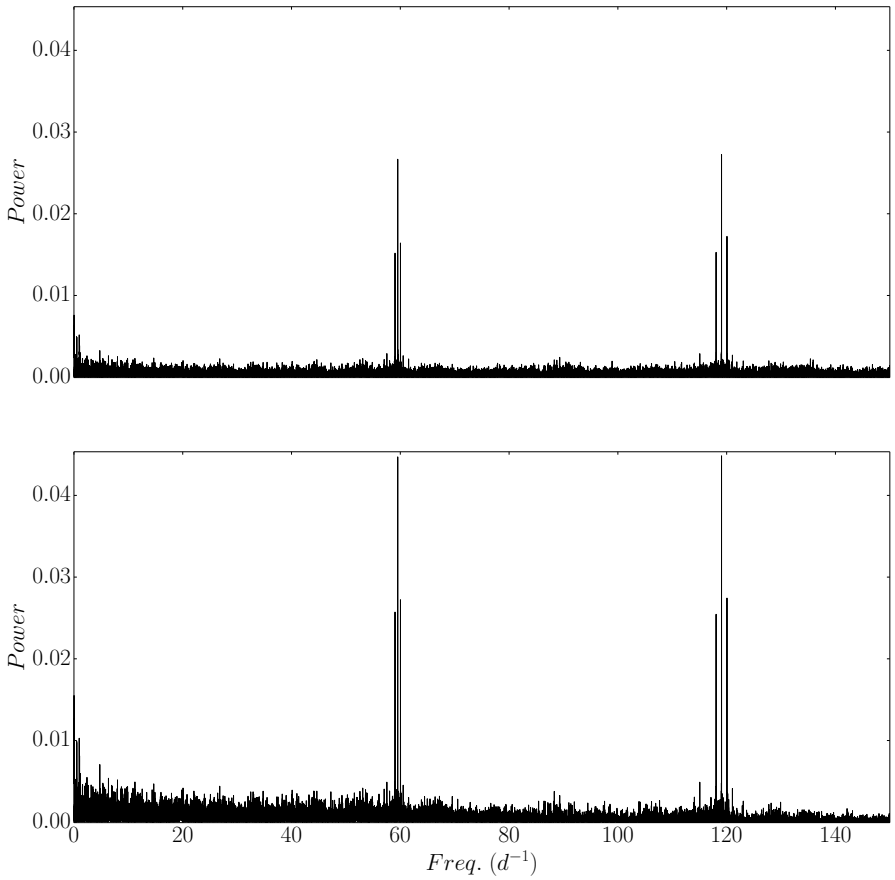


Figure 37: **Top**: TFA's LS power spectrum for white dwarf RE J0317-853 observed with the Evryscope. **Bottom**: TFAW's LS power spectrum. The peak at ~120 d^{-1} in both plots corresponds to the spin period reported by Barstow et al. (1995).

A large fraction of magnetic white dwarfs studied with Kepler (Hermes et al., 2017) have shown optical variations on the hour-to-day timescales, at least in the range of (60-2,000 ppm). A larger sample of these ultra short period spotted white dwarfs could be detected with The Evryscope thanks to its larger footprint, using a longer cadence survey (16-32 minutes), and taking advantage of the improvement in the SNR of the light curves obtained by TFAW.

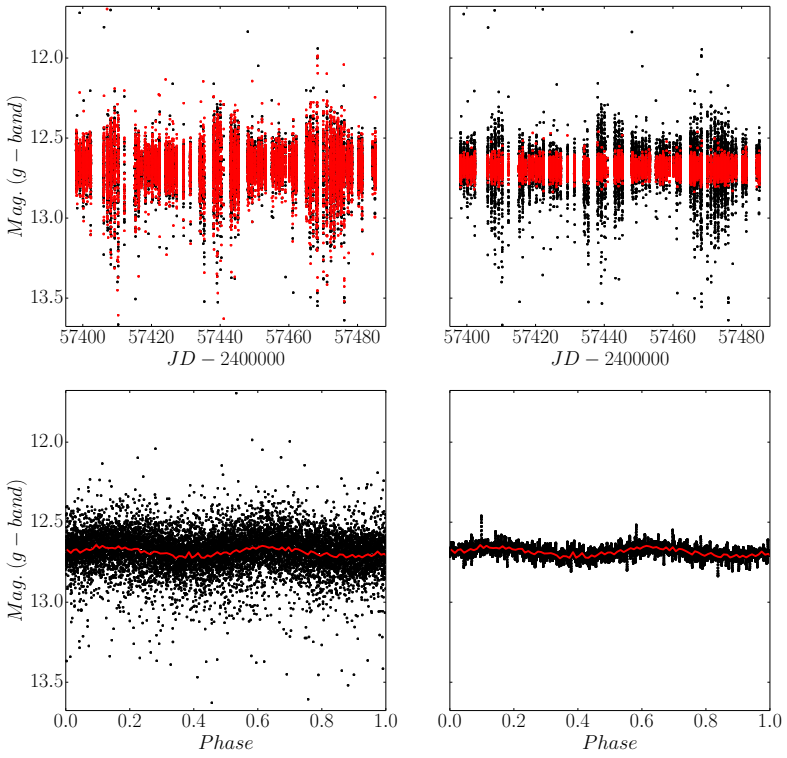


Figure 38: TFAW noise and systematics filtering capabilities for white dwarf RE J0317-853. **Top left**: Normalized raw (black dots) and TFA-filtered (red dots) light curves of RE J0317-853. **Top right**: Normalized raw (black dots) and TFAW-filtered (red dots) light curves of RE J0317-853. **Bottom left**: TFA-filtered light curve of RE J0317-853 phase folded to twice the period reported by Barstow et al. (1995). Red line corresponds to the 300 binned phase folded light curve. **Bottom right**: TFAW-filtered light curve of RE J0317-853 phase folded to twice the period reported by Barstow et al. (1995). Red line corresponds to the 300 binned phase folded light curve.

6.1.2.3 *TFAW quantitative performance over Evryscope light curves*

As in the case of the TFRM data (see Section 6.1.1.3), we wanted to check the photometric performance of TFAW applied to the Evryscope dataset. Top panel of Figure 39 shows The Evryscope's photometric precision (standard deviation σ) vs. The Evryscope magnitude for 965 TFA- and TFAW-corrected light curves with SDEs>15 (i.e. above the median value of the significances for the Evryscope's data sample and that can be considered bona fide variables) observed in the field considered for this work.

Again, TFAW photometric performance is much better than the one provided by TFA. Bottom panel of Figure 39 shows the decrease in the standard deviations of the 965 sample of TFAW-filtered light curves compared to the TFA ones. It can be seen, that for almost all the standard deviation range, TFAW yields an ~67% better final photometric precision than TFA.

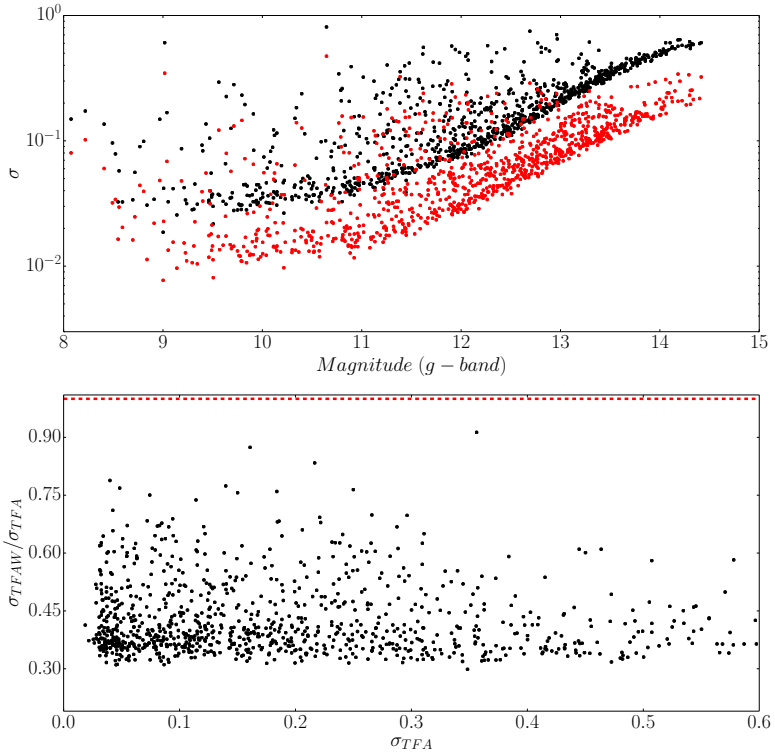


Figure 39: **Top:** Standard deviation vs. R magnitude for TFA (black dots) and TFAW (red dots) of a set of 965 The Evryscope light curves. **Bottom:** Decrease of the standard deviations for the same set of light curves due to the application of TFAW compared to TFA. Both TFA and TFAW results were obtained using 250 template stars. Red line corresponds to the zero correction level.

Regarding the power spectra, we have shown that TFAW does not introduce any fake periodicity in The Evryscope light curves. Even more, we have demonstrated that a careful selection of the *signal level* during TFAW's *frequency analysis* step, can recover high frequency signals like the one for white dwarf RE J0317-853. In addition, a correct *signal level* for the phase folded light curve preserves the high frequency signal's shape while at the same time, efficiently de-noising and removing most of the outliers.

6.2 SPACE TELESCOPE DATA

6.2.1 *CoRoT Data*

6.2.1.1 *Data description*

We downloaded the IRa01 and LRa01 runs from the N2 legacy data official IAS CoRoT Public Archive[1]. We consider the faint star (exoplanet) channel **EN2_STAR**, which comprises stars with magnitude between 10 and 16. The two data sets, classified as level 2 and versions 4.5 and 4.4, respectively, both include chromatic and monochromatic light curves. For this work we focus on those white flux light curves which have a uniform sampling of 512s of cadence. The observation windows, pointing coordinates and the number of light curves of these runs are shown in Table 7.

6.2.1.2 *CoRoT 102886012 and CoRoT 102881832: a binary candidate and a false positive planetary candidate*

In any extrasolar planet survey, once a transit feature has been identified, the possibility of it being caused by a binary system has to be ruled out. In the case of CoRoT (and also Kepler), transit features in the target light curve could be caused by a real exoplanet, a real eclipsing binary or, a Background Eclipsing Binary (BEB) (fainter background eclipsing binary contained in the target photometric aperture). In addition, some configurations of diluted eclipsing binaries or transiting planets can mimic the photometric transit of an real exoplanet (Santerne et al., 2013). For the CoRoT data set, Almenara, J. M. et al. (2009) found that the main types of false positives are eclipsing binaries that are observed directly ("undiluted binaries"), and those whose light is diluted by a nearby third star (which might be physically related or not to the system). Typically, the third star is brighter than the binary and corresponds to the target, while the binary is a faint background system. In some low-amplitude candidates, a transiting system consisting of a star and a giant planet has also to be considered (Léger, A. et al., 2009). Among the undiluted binaries, planet-like eclipses may be caused by grazing eclipsing binaries and the central eclipses of two stellar components with large ratios in area or surface brightness. Clear eclipsing binaries can be detected directly from light curves that exhibit secondary eclipses, out-of-transit photometric modulations or transits too deep to be caused by

1 http://idoc-corot.ias.u-psud.fr/sitools/client-user/COROT_N2_PUBLIC_DATA/project-index.html

Table 7: Observation windows, pointing coordinates, and number of light curves of the CoRoT runs used in this work.

Run	Start time (BSF)	Duration in the Bright star field (days h:m:s)	Pointing direction (J2000)		Number of light curves	
			α (°)	δ (°)	Chromatic	Monochromatic
IRa01	31/01/2007 11:06:34	60 20:05:41.0	102.60	−1.70	3,907	5,973
LRa01	18/10/2007 08:57:24	137 00:52:11.0	101.66	−0.20	7,470	3,938

planets. Also, planetary transits usually have a u-shape while binaries are more v-shaped. (Carpano, S. et al., 2009). However, this rule-out criteria can only be used for light curves with high SNRs. CoRoT has an extra tool that can help decide whether a transit is caused by a planet or eclipsing binary. For stars brighter than r=15, CoRoT provides photometry in 3 different colors that can be used to reject planet candidates (see Section 5.2.1.2 for more details). The chromatic information is helpful to distinguish between achromatic planetary transits and chromatic eclipsing binaries and also to identify false alarms from diluted background binaries. If a transit happens in all three filters, then the chances of it being a real planetary transit increase. If it is only observed in one or two filters at a time, then the transit might be due to a contaminating eclipsing binary (Carone, L. et al., 2012). For the rest of scenarios, complementary observations are needed to discard or not the eclipsing binary nature of the transits. This observations include the search for photometric variations on nearby stars during the target transit; deep, high-angular resolution observations for fainter and close contaminating stars (Deeg, H. J. et al., 2009; Law et al., 2014; Ziegler et al., 2016) ; centroid measurement (Bryson et al., 2013); high-resolution spectroscopy; infrared spectroscopy (Guenther, E. W. and Tal-Or, L., 2010), to search for low-mass companions; and, precise RV measurements to estimate the mass of the transiting object (e.g. Léger, A. et al. (2009), Pätzold, M. et al. (2012), and Santerne et al. (2013)).

We want to check if TFAW is able to help in ruling-out transits caused by binary systems by improving their characterization or by directly detecting secondary transits. TFAW does so increasing the SNR of the target light curves by means of the SWT-based signal estimation and noise removal (see Chapter 4 for more details). As test light curves we have selected CoRoT's IRa01 binary candidate CoRoT 102886012 and the false positive planetary candidate CoRoT 102881832 (Carpano, S. et al., 2009).

CoRoT 102886012 is a star of magnitude V=16.34 located at $(\alpha,\ \delta)$ = (06:48:13.0, -01:36:54) listed as an eclipsing binary candidate by Carpano, S. et al. (2009) with a period of 1.58466$\pm$0.00025 days. It was observed by CoRoT during the Initial Run (IRa01) from January to April 2007. We use 8,192 data points in our light curve extracted from the monochromatic white flux of the N2, version 4.5 legacy data for CoRoT 102886012. The target was observed in the E1R CCD with CoRoT's long cadence of 512s. To run TFA and TFAW we build or sample of template stars by selecting 200 stars randomly distributed in the target's CCD. In or-

der to minimize any bias or introducing correlated periodicities in the target light curve coming from the template light curves, only those stars with low standard deviation (<0.006 mags) and with a small (<-0.6) L variability index are selected as template stars. During TFA and TFAW's *frequency analysis* steps (see Sections 4.2.3 and 4.3.3), the period of the primary transit is found as seen in Figure 40 (though the half period has slightly higher significance). Once the period has been found, we run the *iterative reconstruction* step (see Sections 4.4.2 and 4.3.4) with TFA and TFAW. For the latter, the *signal level* has been chosen following the method explained in Section 3.3.1 and the *noise level* has been set to the lowest (i.e. highest frequency) SWT's decomposition scale.

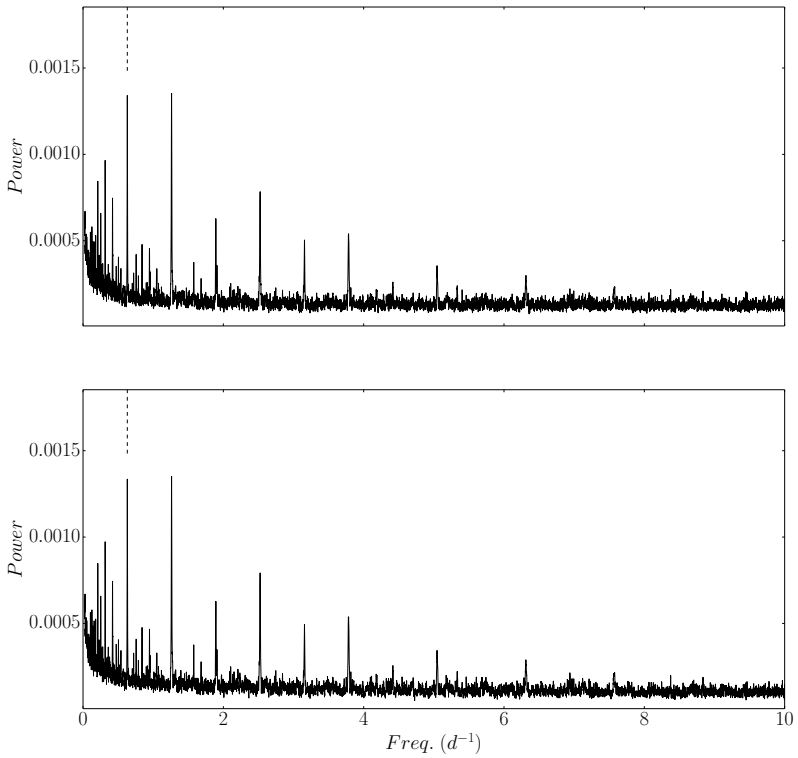


Figure 40: **Top**: TFA's BLS power spectrum for CoRoT's IRa01 eclipsing binary candidate CoRoT 102886102. **Bottom**: TFAW's BLS power spectrum. Dashed lines denote the 1.58466-day period reported by Carpano, S. et al. (2009).

Figure 41 shows the reconstructed signals for TFA and TFAW for CoRoT 102886012. As can be seen, TFAW efficiently removes the high frequency noise preserving the underlying signal's shape and depth for the primary transit. The high frequency noise removal achieved by TFAW leads to a better signal SNR than for TFA's light curve. Specially for the secondary transit whose profile is better defined after TFAW has been applied.

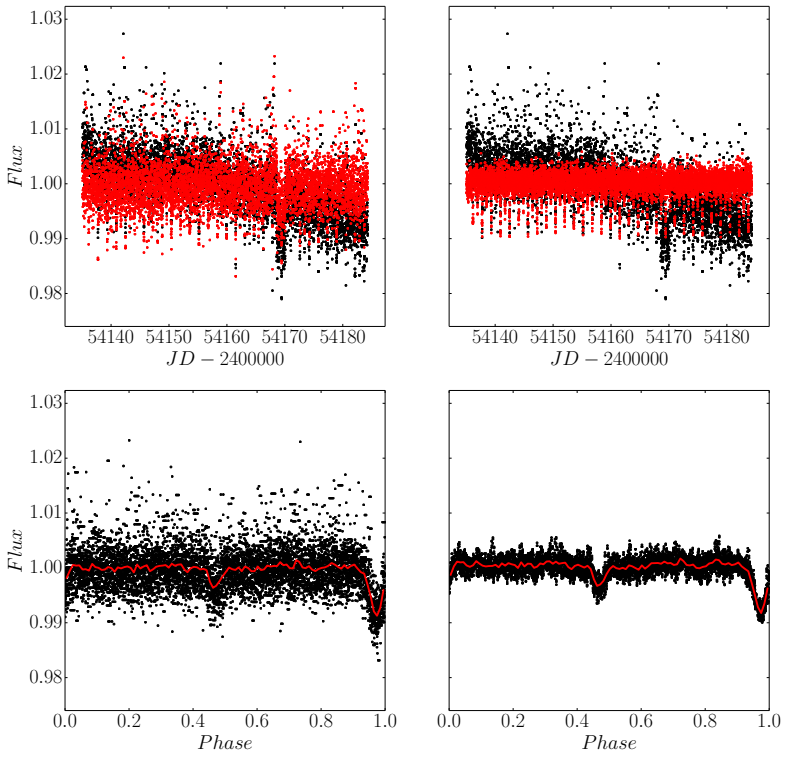


Figure 41: **Top left**: Normalized raw (black dots) and TFA-filtered (red dots) light curves of CoRoT 102886012. **Top right**: Normalized raw (black dots) and TFAW-filtered (red dots) light curves of CoRoT 102886012. **Bottom left**: Phase folded TFA-filtered light curve of CoRoT 102886012. Red line corresponds to the 300 binned phase folded light curve. **Bottom right**: Phase folded TFAW-filtered light curve of CoRoT 102886012. Red line corresponds to the 300 binned phase folded light curve. The arrows mark the phase position of the secondary transit.

CoRoT 102881832 is a V=15.273 (Zacharias et al., 2012) star located at $(\alpha,\ \delta)$ = (06:48:07.5462, -00:47:09.0765) (Gaia Collaboration, 2018). It is listed as one of the 50 transit candidates in Carpano, S. et al. (2009) but it is not mentioned in the follow-up paper of Moutou, C. et al. (2009). The transit period reported by Carpano, S. et al. (2009) is 2.16638±0.000087 days. As the previous example, CoRoT 102881832 was observed during CoRoT's Initial Run (IRa01) but, this time, by the E1L CCD. Again, we use 8,192 data points from the white flux light curve provided by CoRoT's legacy data and build the TFA and TFAW template samples following the same method as before. Both TFA and TFAW's *frequency analysis* steps recover the period reported by Carpano, S. et al. (2009) as can be seen in Figure 42.

As before, once the period has been found, we run TFA and TFAW's *iterative reconstruction* steps. For TFAW, we set the *noise level* to the lowest decomposition scale and the *signal level* has

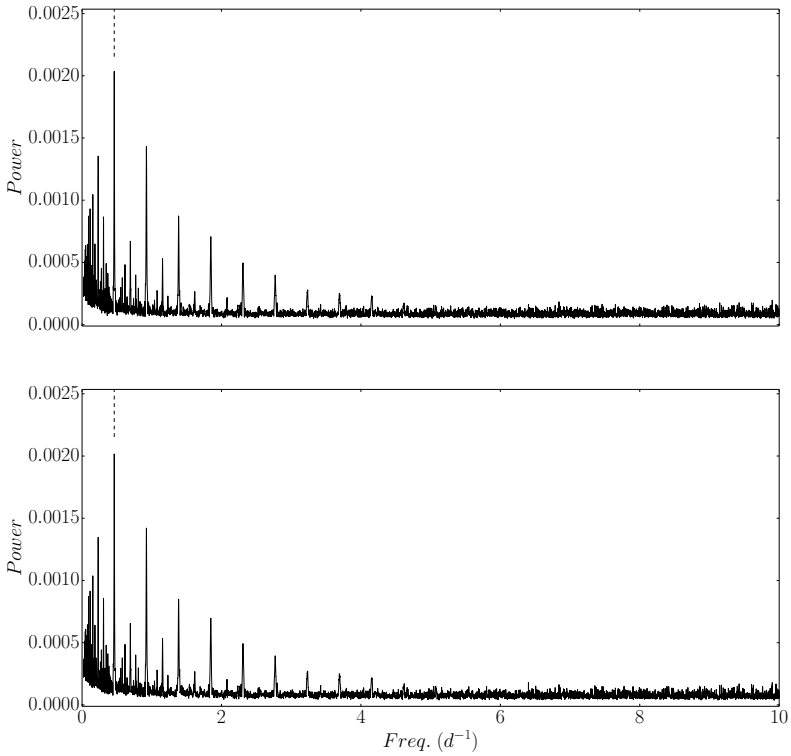


Figure 42: **Top**: TFA's BLS power spectrum for CoRoT's IRa01 planetary transit candidate CoRoT 102881832. **Bottom**: TFAW's BLS power spectrum. Dashed lines denote the 2.16638-day period reported by Carpano, S. et al. (2009).

been selected using the method in Section 3.3.1. Figure 43 shows TFA and TFAW's reconstructed signals for CoRoT 102881832. TFAW has adequately denoised the signal and removed the outliers present in the light curve. Moreover, in both light curves, but, clearer in TFAW's one, a shallow faint transit can be seen at phase ~0.9 (primary transit's phase plus 0.5) suggesting an eclipsing binary origin for the transit. Effectively, planetary transit candidate CoRoT 10288132 was later classified as an eclipsing binary by the "Automated supervised classification of variable stars" developed by Debosscher, J. et al. (2007) and applied to CoRoT's data in Debosscher, J. et al. (2009). This transit candidate was also given low chances of being of planetary origin by Nefs, S. V., Snellen, I. A. G., and de Mooij, E. J. W. (2012). Based on the fact that "the v-shaped light curves of near-grazing planet systems are strongly degenerate with blended eclipsing binary systems and can therefore not be distinguished from each other", Nefs, S. V., Snellen, I. A. G., and de Mooij, E. J. W. (2012) apply a cut in the impact parameter, b $<$0.85, to minimize the number of candidate systems that would be assigned follow-up observations. They find an impact parameter

of b=1.12 for CoRoT 102881832. It also appears an eclipsing binary system in the "Unofficial CoRoT Eclipsing Binary Catalog" of Jonathan Devor[2] with a radius ratio, p=0.660±0.053 and a mass ratio, M_2/M_1=0.31±0.23. CoRoT 102881832 spectral type is A5V according to EXODAT. Assuming that the primary is a main-sequence star and has an approximate radius of $1.8R_\odot$ and mass of $2M_\odot$ the fit parameters found by Jonathan Devor yield a radius of ~$1.18R_\odot$ and mass of ~$0.62M_\odot$ for the transiting object.

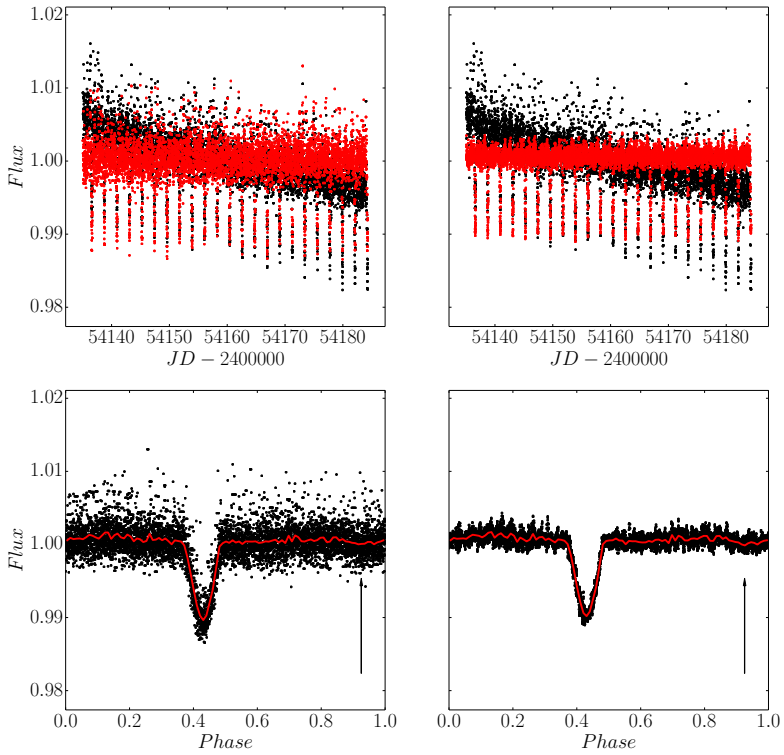


Figure 43: **Top left**: Normalized raw (black dots) and TFA-filtered (red dots) light curves of CoRoT 102881832. **Top right**: Normalized raw (black dots) and TFAW-filtered (red dots) light curves of CoRoT 102881832 for a *noise level* equal to the lowest decomposition scale. **Bottom left**: Phase folded TFA-filtered light curve of CoRoT 102886012. Red line corresponds to the 300 binned phase folded light curve. **Bottom right**: Phase folded TFAW-filtered light curve of CoRoT 102881832. Red line corresponds to the 300 binned phase folded light curve. Arrows point to the shallow secondary transit.

Given that there does not seem to be any significant high frequency signals neither in the BLS power spectrum (see Figure 42) or the phase folded light curve (see Figure 43), we can still improve the SNR of the secondary transit by increasing the *noise level*. We run again TFAW's iterative *signal reconstruction* step but,

2 http://www.astro.tau.ac.il/~jdevor/CoRoT_catalog/catalog.html

this time considering the **three** lowest decomposition scales of the phase folded light curve as *noise levels*. The *signal level* is computed as before through the method explained in Section 3.3.1. Figure 44 shows the new TFAW's reconstructed signal. As can be seen, the signal has been greatly denoised and the shallow secondary transit at phase ~0.9 has its SNR improved.

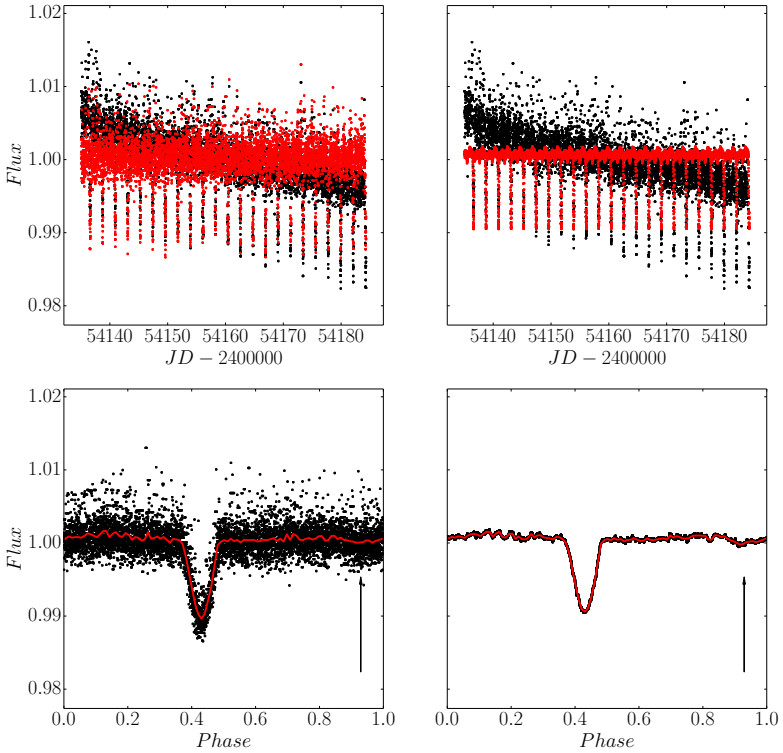


Figure 44: **Top left**: Normalized raw (black dots) and TFA-filtered (red dots) light curves of CoRoT 102881832. **Top right**: Normalized raw (black dots) and TFAW-filtered (red dots) light curves of CoRoT 102881832 for a *noise level* equal to the three lowest decomposition scales. **Bottom left**: Phase folded TFA-filtered light curve of CoRoT 102886012. Red line corresponds to the 300 binned phase folded light curve. **Bottom right**: Phase folded TFAW-filtered light curve of CoRoT 102881832. Red line corresponds to the 300 binned phase folded light curve. Arrows point to the shallow secondary transit.

6.2.1.3 *CoRoT-21 b: a test case*

CoRoT-21 is a F8IV star of magnitude V=15.84 mag (Zacharias et al., 2012) located in the Monoceros constellation at (α, δ) = (06:44:12.6302, -00:17:56.987) (Gaia Collaboration, 2018). It was observed by the CoRoT mission during the Long Run 01 (LRa01) from October 2007 to March 2008. CoRoT-21 b, CoRoT ID 102725122, was not discovered, in contrast to all other CoRoT planets up to that moment, by its alarm mode, but by six de-

tection teams that processed the raw light curves for the LRa01 run (Carone, L. et al., 2012). Candidate 102725122 received a low priority for follow-up observations due to the long transit duration (4.8 h) for a G2 star (CoRoT-21 was classified as such in the `Exo-Dat` database at that time (see Section 5.2.1.2 for more details on the `Exo-Dat`)). Later on, CoRoT-21 was given a higher priority and follow-up observations were performed with HARPS in November 2009, and with the HIRES instrument at Keck in December 2009, January 2010 and January 2011 (Pätzold, M. et al., 2012).

A transit with a period of 2.72482 days was detected using the `EXOTRANS` package (Grziwa, Pätzold, and Carone, 2012) and BLS. The possibility of the system being a binary was ruled out by means of the `DEBiL` package (Devor, 2005) and by later HARPS and HIRES follow-up observations. The raw, the `EXOTRANS`-corrected and the binned and phase folded light curves obtained by Pätzold, M. et al. (2012) are shown in Figure 45.

To characterize the transit Pätzold, M. et al. (2012), fitted a Mandel and Agol (2002) model to the phase-folded light curve using the Genetic Algorithm developed by Kim, Geem, and Kim (2001). Assuming a circular orbit (i.e. e=0), they fitted the transit epoch T0, the semi-major axis a (R_*), the planet to star radius ratio p, the impact parameter $b = a\cos i/R_*$ and the combinations $u_+ = q_1 + q_2$ and $u_- = q_1 - q_2$ where q_1 and q_2 are the quadratic limb-darkening coefficients. RV and spectroscopy follow-up observations of CoRoT-21 were performed with the HARPS and the HIRES spectrographs. This RV data was used to estimate the mass of the planet and the spectroscopy was used to estimate the star radius, mass and age through the mean stellar density, the effective temperature and metallicity. The stellar and planetary parameters obtained through the `EXOTRANS` light curve and the RV and spectroscopic observations are summarized in Table 8.

We want to compare TFA and TFAW performance for a light curve of a confirmed CoRoT exoplanet. We use 16,384 data points (allowing 14 SWT decomposition levels) to build our target light curve from the monochromatic white flux of the N2, version 4.4 legacy data provided by the CoRoT mission for CoRoT-21 b. Our target was observed with the E2L CCD with and exposure time of 512s. To build our sample of template stars, we use 200 stars randomly distributed in the E2L CCD FoV. To minimize any biasing of our target light curve after applying TFA and TFAW, only stars with low standard deviation (<0.006) and with a small (<-0.6) L Stetson's variability index are selected. The resulting TFA and TFAW corrected light curves are shown

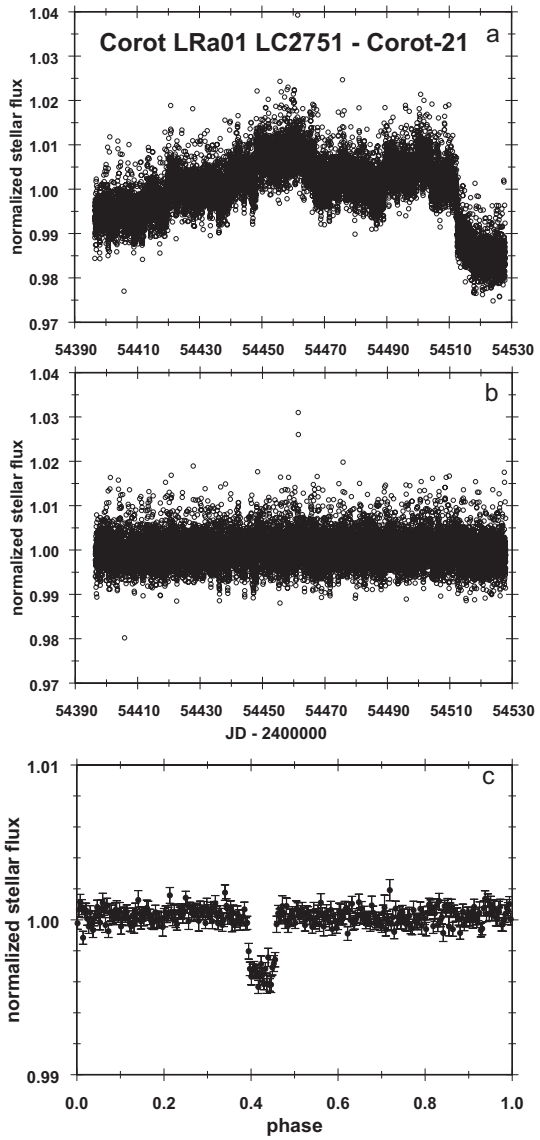


Figure 45: CoRoT-21 b light curves as seen in (Pätzold, M. et al., 2012). **Top**: Normalized raw light curve of CoRoT-21. **Middle**: EXOTRANS-corrected light curve. **Bottom**: Phase folded light curve. The phase is divided into 300 bins with mean flux and standard deviation displayed for each bin.

Table 8: Stellar and planetary parameters obtained for CoRoT-21 b by Pätzold, M. et al. (2012).

CoRoT ID	102725122
Transit parameters	
Period P (days)	2.72474 ± 0.00014
Transit epoch To (JD - 2450000)	4399.0282 ± 0.0009
Transit duration (hours)	4.76
Eccentricity e	0 (fixed)
Radius ratio p	0.067 ± 0.0018
u_+	0.66 ± 0.09
u_-	0.13 ± 0.11
Impact parameter b	0.25 ± 0.17
Scaled semi-major axis a (R_*)	4.60 ± 0.26
Inclination i (^{o})	86.8 ± 2.1
RV parameters	
Stellar rotational velocity (km s^{-1})	11 ± 1.0
System velocity Γ (m s^{-1})	101
Half-amplitude K (m s^{-1})	274 ± 35
Stellar parameters	
Stellar radius R_* ($R_\odot$)	1.95 ± 0.21
Stellar mass M_* ($M_\odot$)	1.29 ± 0.09
Age (Gyr)	$4.1^{+0.1}_{-0.5}$
Effective temperature (K)	6200 ± 100
Surface gravity log g (dex)	3.7 ± 0.1
Metallicity [Fe/H]	0.0 ± 0.1
Spectral Type	F8IV
Planetary parameters	
Planetary radius R_p (R_J)	1.30 ± 0.14
Planetary mass M_p (M_J)	2.26 ± 0.31
Planetary bulk density $\langle\rho\rangle$ (10^3 kg/m^3)	1.36 ± 0.48

in Figure 47 and the BLS power spectra for their corresponding *frequency analysis* steps (see Sections 4.2.3 and 4.3.3) are shown in Figure 46. As can be seen, after TFAW is applied, the noise contribution decreases significantly and outliers are efficiently removed. Regarding the power spectra, TFAW's *frequency analysis* step obtains a very similar result compared to TFA.

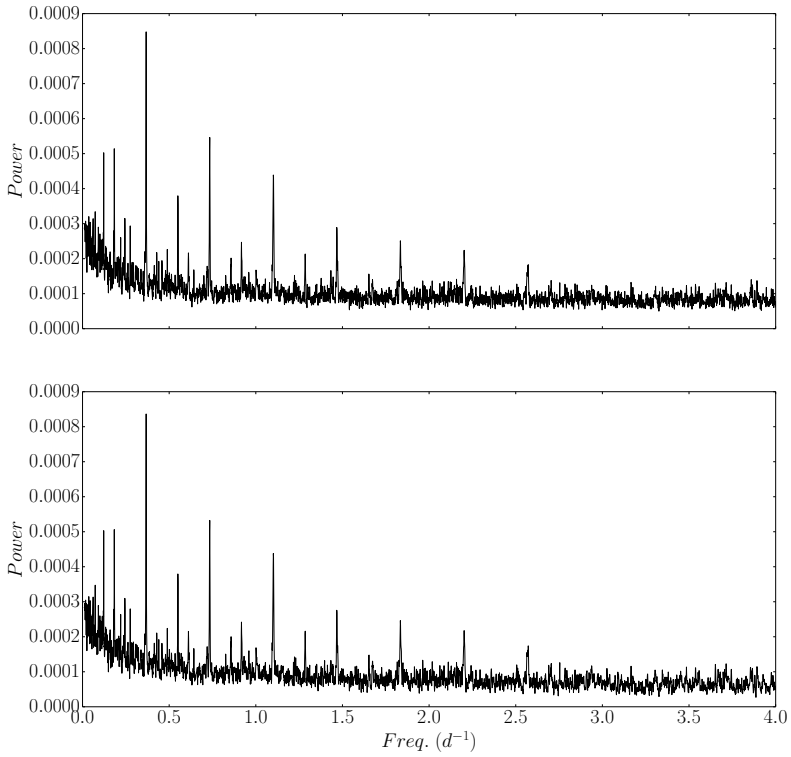


Figure 46: **Top**: TFA's BLS power spectrum for confirmed exoplanet CoRoT-21 b. **Bottom**: TFAW's BLS power spectrum for confirmed exoplanet CoRoT-21 b.

To characterize the transit, we model the fit with the `batman` package using a Mandel and Agol (2002) analytic transit model. We used the MCMC sampler provided by `emcee` (Foreman-Mackey et al., 2013) to solve for 5 transit parameters: the transit epoch `T0`, the inclination of the orbit `i`, the ratio between the planetary radius and the stellar radius `p`, the period `P` and the semi-major axis of the orbit `a`. Given the faint magnitude of CoRot-21, the correct determination of the limb-darkening coefficients is difficult to obtain. Thus, we have fixed the quadratic limb darkening coefficients to their theoretical values taken from the tables by Sing, D. K. (2010). For a T_{eff}=6200K, `log g`=3.7 dex, [Fe/H]=0.0 star this values are `l`=0.395 and `q`=0.265. As in Pätzold, M. et al. (2012), we have also considered a circular orbit (i.e. `e`=0). For the MCMC fit, we have considered a uniform distribution of the priors and run the sampler with 200 chains

and 10,000 iterations with a burn-in phase of 2,000 iterations. This way we ensure that the chains run for more than 50 autocorrelation times for each parameter and that the mean acceptance fraction is between 0.25 and 0.5 (Bernardo et al., 1996 and Foreman-Mackey et al., 2013). Also, we have started the walkers for each parameter in a tight ball around the values obtained by Pätzold, M. et al. (2012). In Table 9 we compare the transit parameters obtained by Pätzold, M. et al. (2012) with the ones obtained through the MCMC fit for TFA and TFAW posterior probability distributions. The fitted MCMC parameter values correspond to the 50% quantile while the upper and lower errors are computed from the 25% and 75% quantiles.

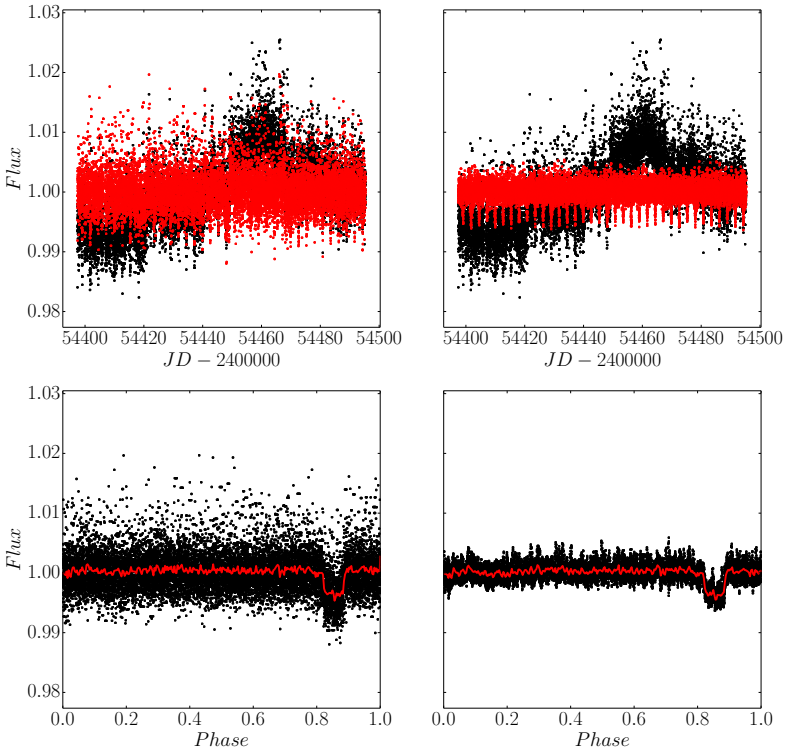


Figure 47: TFAW noise and systematics filtering capabilities for confirmed exoplanet CoRoT-21 b. **Top left**: Normalized raw (black dots) and TFA-filtered (red dots) light curves of CoRoT-21 b. **Top right**: Normalized raw (black dots) and TFAW-filtered (red dots) light curves of CoRoT-21 b. **Bottom left**: Phase folded TFA-filtered light curve of CoRoT-21 b. Red line corresponds to the 300 binned phase folded light curve. **Bottom right**: Phase folded TFAW-filtered light curve of CoRoT-21 b. Red line corresponds to the 300 binned phase folded light curve.

For the best fit, we find both for TFA and TFAW that the time of inferior conjunction, T0, lies somewhere around 4,399.03 (JD-2450000). This is approximately 8s later than the one reported by Pätzold, M. et al. (2012) but within the error of the tran-

sit epoch. Regarding the semi-major axis of the orbit, a, TFA yields a slightly lower value than the one reported by Pätzold, M. et al. (2012) and TFAW a higher one but both compatible with the measured value of $4.60\pm0.26 R_*$. For the impact parameter, b, and the orbital inclination, i, TFA returns $0.44^{+0.76}_{-0.64}$ and $84.24669^{+0.0751}_{-0.03465}{}^{o}$ while TFAW returns 0.0 ± 0.62 and $89.9979^{+0.0751}_{-0.03465}{}^{o}$ which are compatible within the errors with the values given by Pätzold, M. et al. (2012). The difference in the uncertainties for the impact parameter with respect the one in the literature is due to the fact that we fit the orbital inclination and from it derive b while Pätzold, M. et al. (2012) do it reversely. For the planetary to star radius ratio, p, both TFA and TFAW obtain smaller values ($\sim1.11\times$) than the one reported by Pätzold, M. et al. (2012). However, they would be compatible with a bigger star radius than the reference value of $R_*=1.95\pm0.21 R_\odot$ (around 2.2 $R_\odot$) assuming a planet radius of 1.30 ± 0.14 R_J. This difference in the transit depth for both TFA and TFAW with respect the literature values could probably be understood taking into account the main difference between Pätzold, M. et al. (2012) and this work. The TFA implementation used by `EXOTRANS` (Grziwa, Pätzold, and Carone, 2012) does not run the *signal reconstruction* step described in Section 4.4.2. As explained in Kovács, Bakos, and Noyes (2005), this can lead to a biased or distorted light curve as the signal is assumed to be constant during the *frequency analysis* step. With the best fit parameters, both TFA and TFAW obtain a mean planetary radius slightly above that of Jupiter (1.11 ± 0.12 for TFA and 1.15 ± 0.13 for TFAW). Using the same RV results for the half-amplitude, K, as in Pätzold, M. et al. (2012), the mass of CoRoT-21 b for TFA is 2.21 ± 0.31 and 2.26 ± 0.31 for TFAW; compatible with the literature value. Finally, the bulk density $\langle\rho\rangle$, is 1.97 ± 0.46 and 1.85 ± 0.43 for TFA and TFAW respectively. The larger values for this parameter compared to the one in (Pätzold, M. et al., 2012), is easily explained by the smaller radii obtained in this work. Figure 48 shows the best fit obtained for TFA and TFAW over-plotted to their corresponding phase folded light curves.

As seen in Section 4.4.5, TFAW yields a closer representation of the real transit than TFA and with less uncertainties. For CoRoT-21 b, although the planetary radii obtained with TFA and TFAW are smaller (which as explained above could be due to *signal reconstruction* step missing in `EXOTRANS`), all transit and derived parameters are compatible with the ones given by Pätzold, M. et al. (2012). Using CoRoT-21 b light curve, we have confirmed the correct performance of TFAW when applied to a real, confirmed exoplanet scenario.

Table 9: Top table: CoRoT-21 b parameters obtained by Pätzold, M. et al. (2012) using a Mandel and Agol (2002) model and the Genetic Algorithm (Kim, Geem, and Kim, 2001). Posterior transit parameters values and their uncertainties (with the 25% and 75% quantile as the upper and lower errors) for TFA and TFAW's MCMC fits. Middle table: 95% confidence intervals of the highest probability density for CoRoT-21 b transit parameters TFA and TFAW's MCMC fits. Bottom table: Derived planetary parameters for TFA and TFAW.

Parameters	To (JD-2450000)	P (days)	a (R_*)	p	i (°)
Pätzold, M. et al., 2012	4,399.0282±0.0009	2.72474±0.00014	4.60±0.26	0.067±0.0018	86.8±2.1
TFA MCMC	$4{,}399.03701^{+0.00012}_{-0.00012}$	$2.72486^{+0.00001}_{-0.00001}$	$4.8174^{+0.0039}_{-0.0042}$	$0.06037^{+0.00023}_{-0.00023}$	$90.0021^{+0.3099}_{-0.3128}$
TFAW MCMC	$4{,}399.03081^{+0.00006}_{-0.00006}$	$2.72485^{+0.000001}_{-0.000001}$	$4.7496^{+0.00125}_{-0.00128}$	$0.06077^{+0.00007}_{-0.00007}$	$89.9979^{+0.13239}_{-0.13235}$
	95% confidence intervals of the highest posterior density				
TFA MCMC	4,399.03676 - 4,399.03724	2.72485 - 2.72488	4.80886 - 4.82541	0.05991 - 0.06083	89.3851 - 90.6080
TFAW MCMC	4,399.03070 - 4,399.0392	2.72484 - 2.72485	4.74714 - 4.75216	0.06063 - 0.06091	89.7374 - 90.2615
System parameters	R_p (R_J)	M_p (M_J)	V_p (km/s)	b	$\langle\rho\rangle$ (10^3 kg/m^3)
Pätzold, M. et al., 2012	1.30±0.14	2.26±0.31	165.94±??	0.25±0.17	1.36±0.48
TFA	1.14±0.12	2.28±0.30	162.06±5.65	$0.0^{+1.493}_{-1.507}$	1.89±0.45
TFAW	1.15±0.13	2.26±0.30	163.21±5.69	$0.0^{+0.62}_{-0.62}$	1.85±0.43

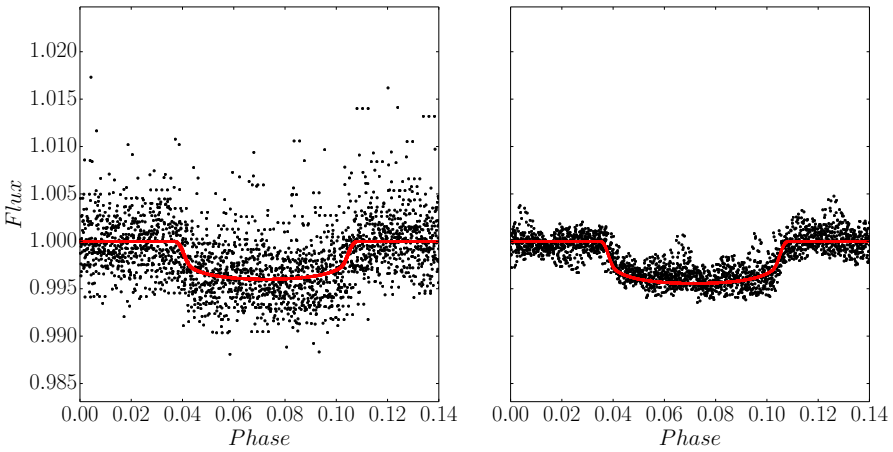


Figure 48: **Left**: Phase folded TFA light curve for CoRoT-21 b. The solid red line corresponds to the best fit given by the parameters in Table 9. **Right**: Phase folded TFAW light curve for CoRoT-21 b. As before, the solid red line corresponds to the best fit for TFAW following the parameters in Table 9.

6.2.1.4 *CoRoT 102588881: a planetary candidate*

CoRoT 102588881 is a K2V star (Guenther, E. W. et al., 2013) (originally classified as a G2IV star by `EXODAT`) of magnitude V=16.08 (Carone, L. et al., 2012) located at $(\alpha,\ \delta)$ = (06:41:07.805, +00:34:15.26) (Cutri et al., 2003). It was observed by the CoRoT mission during the Long Run 01 (LRa01) from October 2007 to March 2008. A 1.52% deep transit with a period of 27.29 days was detected by Carone, L. et al. (2012). Observations done with the IAC80 Telescope[3] excluded contaminating eclipsing binaries. In addition Guenther, E. W. et al. (2013) through high-resolution spectroscopy excluded a companion star with spectral type earlier than M3.5V and the presence of a companion of similar brightness as the target within 2 arcsec. The Gaia mission (Gaia Collaboration, 2018) does not detect any other source within this distance. RV observations with HARPS showed variations of 84 ms^{-1} that were comparable with the errors (~70 ms^{-1}). Carone, L. et al. (2012) assuming a stellar mass M_*=1$M_\odot$, discarded the presence of a Jupiter-mass planet in this 27.29 days orbit as it would produce peak-to-peak RV variations of ~120 ms^{-1}. Although Carone, L. et al. (2012) state that more HARPS measurements are required to assess the nature of the transit and that follow-up observations were ongoing, no more info is found on this target since (Guenther, E. W. et al., 2013).

3 http://vivaldi.ll.iac.es/OOCC/iac-managed-telescopes/iac80/

Table 10: Stellar and planetary parameters used to initialize the MCMC fit for transit candidate 102588881.

CoRoT ID	102588881
Transit parameters	
Period P (days)	$27.2881 \pm 0.0012^{\dagger}$
Transit epoch To (JD - 2450000)	$4409.3616 \pm 0.00340^{\dagger}$
Transit duration (hours)	$2.64^{\dagger}$
Eccentricity e	0.5
Radius ratio p	0.1148
u_+	0.5715
u_-	0.1368
Scaled semi-major axis a (R_*)	37.39
Inclination i (^{o})	87.7
RV parameters	
Half-amplitude K (m s^{-1})	$84 \pm 35^{\dagger}$
Stellar parameters	
Stellar radius R_* ($R_\odot$)	1.02
Stellar mass M_* ($M_\odot$)	0.7
Effective temperature (K)	5000
Surface gravity log g (dex)	4.3
Metallicity [Fe/H]	0.0
Spectral Type	K2V
Planetary parameters	
Planetary radius R_p (R_J)	1.14
Planetary mass M_p (M_J)	1.1

† denote values taken from Carone, L. et al. (2012).

As with CoRoT-21 b, we want to compare the detrending and de-noising performance of TFA and TFAW. We use 16,384 data points to build CoRoT 102588881 light curve obtained from the monochromatic white flux of the N2, version 4.4 legacy data. This CoRoT target was observed with the E1L CCD with a cadence of 512s. We use the same method employed with CoRoT-21 b (see Section 6.2.1.3 for more details) to build our sample of template stars. Although the transit is discernible from the raw light curve, and during TFA and TFAW's *frequency analysis* step the 27.29 days period is detected in the BLS power spectrum (see Figure 49), the corresponding peak is not the most significant one. Thus to run the *signal reconstruction* step, we had to manually force the period to phase fold the light curve. As in the previous examples, the *noise level* is set to the lowest SWT's decomposition scale and the *signal level* of the phase folded light curve is chosen using the method explained in Sect 3.3.1.

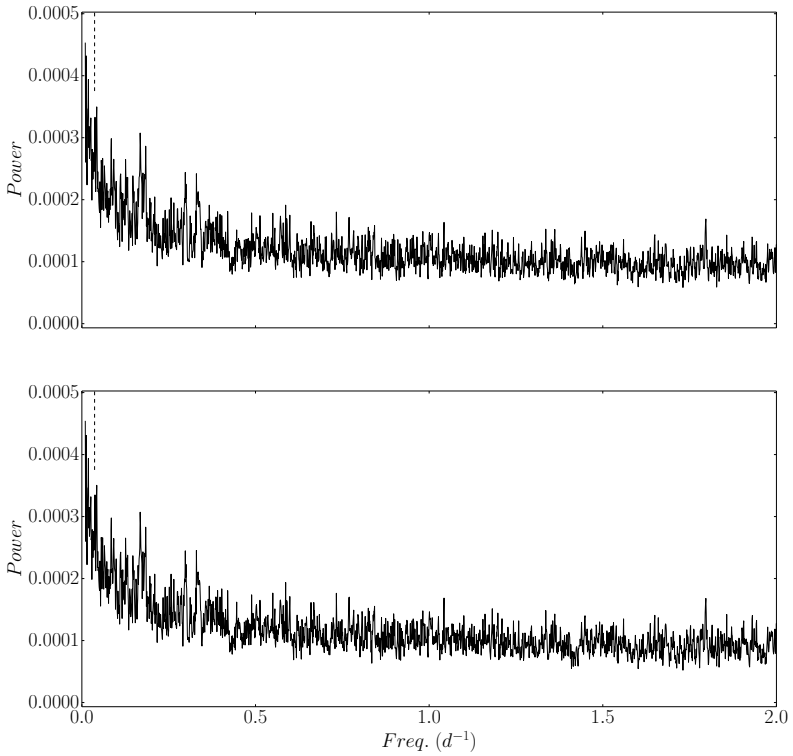


Figure 49: **Top**: TFA's BLS power spectrum for planetary candidate CoRoT 102588881. **Bottom**: TFAW's BLS power spectrum for planet candidate CoRoT 102588881. The dashed vertical line marks the transit's period as reported by Carone, L. et al. (2012).

Figure 50 shows CoRoT 102588881 reconstructed light curves for TFA and TFAW. As in the previous cases, the final TFAW-filtered light curve presents a higher SNR than the TFA one thanks to the SWT-based denoising and outlier removal.

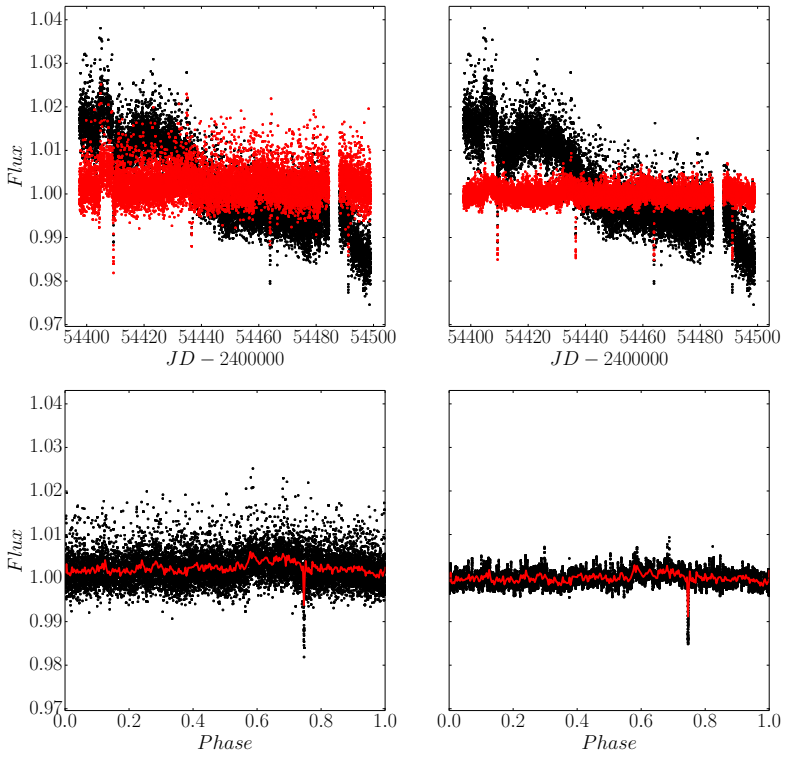


Figure 50: TFAW noise and systematics filtering capabilities for planetary candidate CoRoT 102588881. **Top left**: Normalized raw (black dots) and TFA-filtered (red dots) light curves of CoRoT 102588881. **Top right**: Normalized raw (black dots) and TFAW-filtered (red dots) light curves of CoRoT 102588881. **Bottom left**: Phase folded TFA-filtered light curve of CoRoT-21 b. Red line corresponds to the 300 binned phase folded light curve. **Bottom right**: Phase folded TFAW-filtered light curve of CoRoT 102588881. Red line corresponds to the 300 binned phase folded light curve.

Following the example of CoRoT-21 b, to characterize the transit we employ the `batman` package and a Mandel and Agol (2002) analytic model. To fit the transit we use the MCMC sampler provided by the `emcee` Python module. Initial tests run on the transit showed a great correlation between the transit's orbital inclination, i, and eccentricity, e. Thus, for CoRoT 102588881, we solve the fit for 6 transit parameters: the transit epoch, T0; the inclination and eccentricity of the orbit, i and e, respectively; the ration between the planetary and stellar radii, p; the period, per and the semi-major axis of the orbit, a. Also, as happened with CoRoT-21 b, given the faint magnitude of the stars, we have fixed the quadratic limb darkening coefficients to the theoretical values taken from the tables provided by Sing, D. K. (2010). Assuming a T_{eff}=5000K, $\log g$=4.3 dex and [Fe/H]=0.0 (i.e. similar to those for the K2V star Epsilon Eridani (Kovtyukh et al., 2003; Gonzalez, Carlson, and Tobin, 2010) this val-

ues are l=0.5715 and q=0.1368. Given that there are no precise measurements of the star's mass and radius, we have assumed M_*=0.7$M_\odot$ and R_*=1.02$R_\odot$ as the stellar parameters to initialize the transit model. In addition, the RV measurements indicate that if the transit is due to a planet, its mass should be smaller than 1M_J. However, given that Carone, L. et al. (2012) assume that the star is of spectral type G2IV (from EXODAT), whereas more recent observations (Guenther, E. W. et al., 2013) point towards a smaller K2V star, we assume as initial values for the planet mass and radius: 1.1M_J and 1.14R_J respectively. The mass value is compatible with the RV measurements done by Carone, L. et al. (2012) for an object in an orbit of 27.29 days around a K2V star of 0.7$M_\odot$. For the MCMC fit we have considered a uniform distribution of the priors and run the sampler with 200 chains and 10,000 iterations with a burn-in phase of 2000 iterations. As explained in Section 6.2.1.3, this way we try to ensure that the chains run for more than 50 autocorrelation times for each parameter and that the mean acceptance fraction lays between 0.25 and 0.5. The walkers start from a tight ball around the transit parameters in Table 10.

In Table 11 we compare the results obtained through the MCMC fit for TFA and TFAW's posterior probability distributions. The fitted values correspond to the 50% quantile while the upper and lower errors are computed from the 25% and 75% quantiles.

In this case, as can be seen from the posterior probability distributions in Figure 51, the best fit yields a high correlation between the semi-major axis, a, the orbit inclination, i, and its eccentricity, e. This leads to higher uncertainties for these parameters as shown in Table 11. Even though, TFAW obtains narrower probability distributions for these transit parameters (as well as for the other three) than TFA. In Figure 51, notice that for a, i and e the axis ranges of the probability distributions are the same for TFA and TFAW. For T0 and the period, though centered at different values, the axis ranges are also the same. For the semi-major axis of the orbit, a, both TFA and TFAW find compatible values around 38R_*; though the one for TFAW has lower uncertainties (~1.8-2×). For the orbit inclination, i, TFA returns $86.13844^{+0.99118^o}_{-1.31648}$ and, TFAW, $87.40814^{+0.36779^o}_{-0.46681}$. Again, both values could be compatible with each other within the errors, and TFAW yields lower (~3×) uncertainties than TFA. For the orbit eccentricity, e, TFA returns $0.50055^{+0.08575}_{-0.09559}$ and TFAW $0.47384^{+0.04982}_{-0.04997}$. As in the previous parameters, both values are compatible within the errors with TFAW returning lower uncertainties. Given the fitted eccentricity, it would mean that the

Table 11: Top table: CoRoT 102588881 transit parameters obtained by Carone, L. et al. (2012). Posterior transit parameters values and their uncertainties (with the 25% and 75% quantile as the upper and lower errors) for TFA and TFAW's MCMC fits. Middle table: 95% confidence intervals of the highest probability density for CoRoT 102588881 transit parameters TFA and TFAW's MCMC fits. Bottom table: Derived planetary parameters for TFA and TFAW.

Parameters	To (JD-2450000)	P (days)	a (R_*)	p	i (°)	e
Carone, L. et al., 2012	$4{,}409.3616^{+0.00340}_{-0.00340}$	$27.2881^{+0.0012}_{-0.0012}$	-	-	-	-
TFA MCMC	$4{,}409.35337^{+0.00028}_{-0.00028}$	$27.28551^{+0.00011}_{-0.00011}$	$38.0859^{+4.9011}_{-4.3265}$	$0.42482^{+0.13783}_{-0.09943}$	$86.13844^{+0.99118}_{-1.31648}$	$0.5005^{+0.0857}_{-0.0955}$
TFAW MCMC	$4{,}409.34736^{+0.00003}_{-0.00003}$	$27.28993^{+0.00002}_{-0.00002}$	$37.83114^{+2.43864}_{-2.41800}$	$0.13759^{+0.00062}_{-0.00059}$	$87.40814^{+0.36779}_{-0.46681}$	$0.4738^{+0.0498}_{-0.0499}$
	95% confidence intervals of the highest posterior density					
TFA MCMC	4,409.35281 - 4,409.35392	27.28529 - 27.28572	30.77704 - 46.68962	0.25809 - 0.67010	83.8229 - 87.8060	0.3368 - 0.6401
TFAW MCMC	4,409.34730 - 4,409.34742	27.28990 - 27.28997	33.84287 - 41.77990	0.13644 - 0.13883	86.65870 - 88.03138	0.3917 - 0.5543
System parameters	R_p (R_J)	M_p (M_J)	V_p (km/s)	b	$\langle\rho\rangle$ (10^3 kg/m^3)	
TFA	3.72 ±1.49	0.99±0.43	62.37±4.01	$2.56^{+37.72}_{-50.09}$	0.023±0.019	
TFAW	1.20±0.28	0.98±0.42	62.64±4.02	$1.71^{+13.89}_{-17.63}$	0.70±0.41	

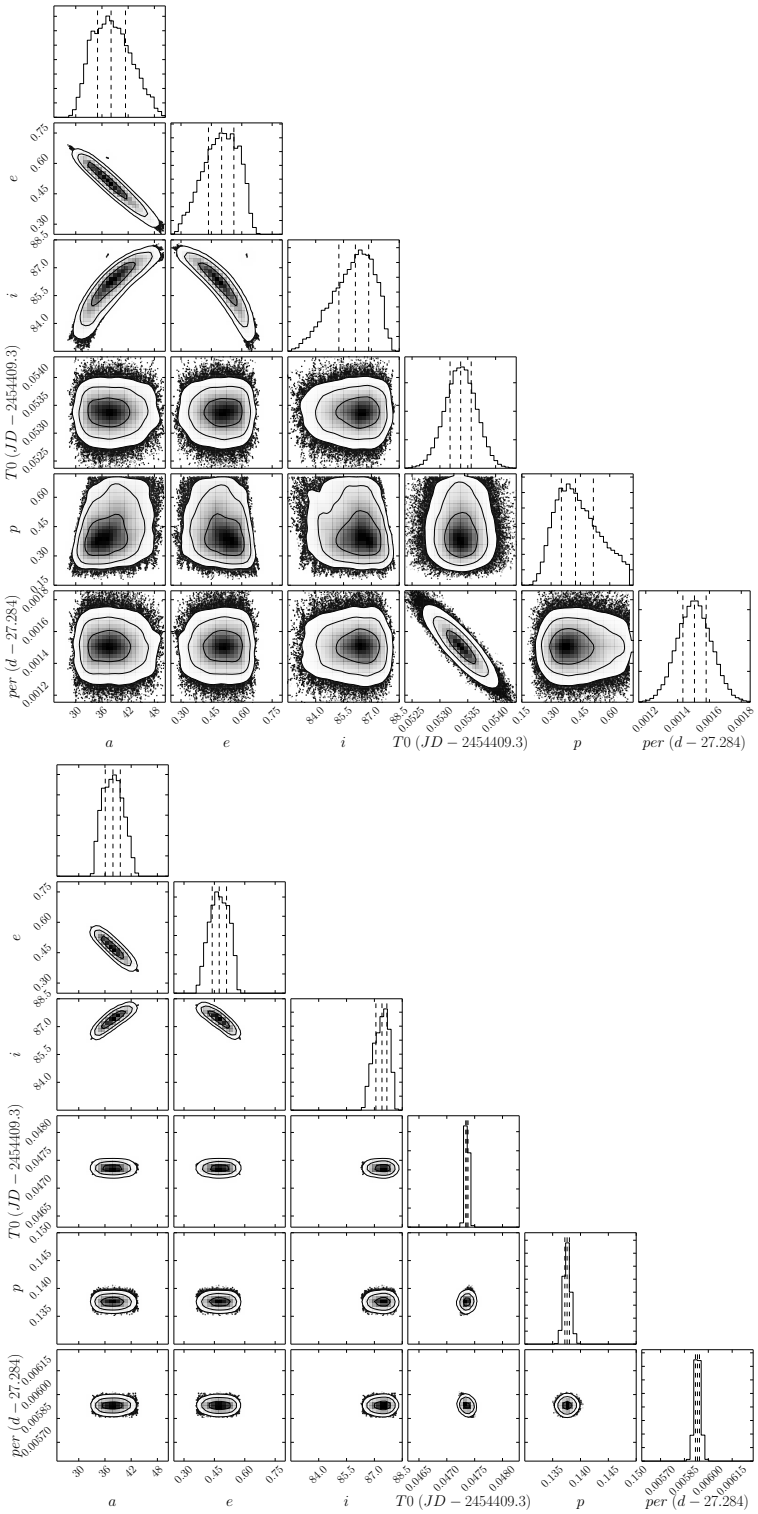


Figure 51: 1-D and 2-D projections of the posterior probability distributions of the 6 MCMC fitted parameters (a, e, i, To, p, P) for CoRoT 102588881 TFA (top) and TFAW (bottom) detrended light curves. The 25%, 50%, 75% quantiles, are displayed in dash vertical lines on the 1-D histograms.

candidate planet orbiting CoRoT 102588881 would have an eccentricity higher than that of Mercury (0.2056; https://nssdc.gsfc.nasa.gov/planetary/factsheet/mercuryfact.html) but still far from the highest eccentricity for a confirmed exoplanet (0.97±0.01 for HD 20782 (O'Toole et al., 2009; Stassun, Collins, and Gaudi, 2017). Regarding the time of inferior conjunction, T0, both TFA and TFAW find very similar values but, the transit happens slightly sooner (within two 512s exposures) than the value reported by Carone, L. et al. (2012). For the period, p, the fit values obtained for TFA and TFAW are compatible with the one in Carone, L. et al. (2012). Nevertheless, TFAW uncertainties for these parameters are lower than the ones reported by Carone, L. et al. (2012) and the ones returned by the fit of the TFA-filtered light curve. The main difference between TFA's MCMC fit and TFAW's one is the planet to stellar radii ratio, p. TFA finds a bigger (a factor ~3×) value than the one for TFAW's fit. This difference will be reflected in the planetary radii R_p obtained for both light curves. The planetary radius and mass obtained for TFA are 3.72±1.49R_J and 0.99±0.43M_J, respectively. Although the planetary mass is compatible with the RV measurements obtained by Carone, L. et al. (2012) and, considering a K2V star, the planet radius seems to be rather unrealistic when compared to the planet mass. In addition, the inferred bulk density of 0.023±0.019 seems to be too small. On the contrary, TFAW obtains a planetary mass of 0.98±0.42M_J, compatible with the RV, and a planet radius of 1.20±0.28R_J which seem to yield a more realistic scenario for the planetary transit candidate. The inferred bulk density of 0.70±0.41 would be compatible with those of other confirmed Jupiter-mass exoplanets. In total, for CoRoT 102588881, TFAW yields closer results to a real exoplanet candidate than TFA and provides smaller uncertainties for the Mandel and Agol (2002) fit parameters. However, due to the still high uncertainties in the semi-major axis, a, and the RV measurements, some derived parameters like the impact parameter, b, and the planet's bulk density, $\langle\rho\rangle$, still return high uncertainties. The latter could be improved, for example, with better RV measurements that would yield smaller errors in the planet mass. Figure 52 shows the best fit obtained for TFA and TFAW over-plotted to their corresponding phase folded light curves.

In summary, as in the test case of confirmed planet CoRoT-21 b, TFAW yields a closer, with less uncertainties and more realistic result than TFA for planet transit candidate CoRoT 102588881. Though in this case, some parameters (the impact parameter, b and the bulk density $\langle\rho\rangle$) still have high uncertainties, we have

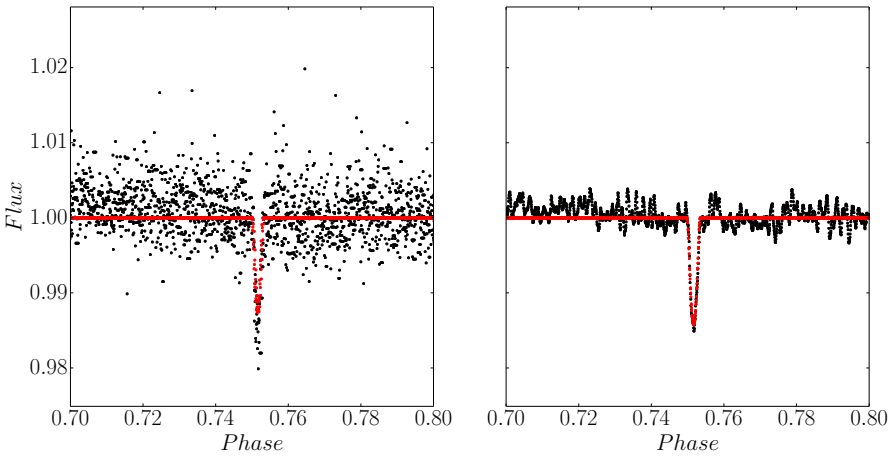


Figure 52: **Left**: Phase folded TFA light curve for CoRoT 102588881. The solid red line corresponds to the best fit given by the parameters in Table 11. **Right**: Phase folded TFAW light curve for CoRoT 102588881. As before, the solid red line corresponds to the best fit for TFAW following the parameters in Table 11.

demonstrated TFAW's potential to improve the characterization of potential planetary transit candidate signals.

6.2.1.5 *CoRoT 102850921: a multi-periodic example*

CoRoT 102850921 is a A6V star (Sebastian et al., 2012) of magnitude V=12.98 mag (Zacharias et al., 2012) located at (α, δ) = (06:47:23.8631, -03:08:32.3797) (Gaia Collaboration, 2018). It was observed by the CoRoT mission during the Initial Run 01 (IRa01) from January to April 2007. CoRoT 102850921 appears as a planet transit candidate with a period of 0.61161 days in Carpano, S. et al. (2009) but it is not mentioned in the follow-up paper of Moutou, C. et al. (2009). Nefs, S. V., Snellen, I. A. G., and de Mooij, E. J. W. (2012) classify this target as a blended eclipsing binary system. For CoRoT 102850921 they find a fitted impact parameter of b=1.18>0.85 thus minimizing the chances of the transit being of planetary origin.

We use 8,192 data points to build our target light curve from the monochromatic white flux of the N2, version 4.5 legacy data provided by the CoRoT mission for the IRa01 run. Our target was observed with the E2R CCD with an exposure time of 512s. As in the previous examples, we build our sample of template stars, using 200 stars randomly distributed in the CCD FoV. As before, to minimize any biasing of our target light curve after applying TFA and TFAW, only stars with low standard deviation (<0.006) and with a small (<-0.6) L Stetson's variability index are selected.

After running TFA and TFAW *frequency analysis* step, we analyzed the BLS power spectrum of the detrended light curve (see top panel in Figure 54). We find two clear signals at periods 0.1193 and 0.6116 days corresponding to a sinusoidal modulation and the transit, respectively. The 0.1193-day variability hints towards a δ-Scuti nature for this A-type star (Balona, 2011). In addition, we removed the contribution of these two signals and found two extra semi-sinusoidal contributions at periods 0.0951 and 0.1588 days (see second to fourth panels in Figure 54). Neither of them correspond to the orbital period of the satellite (1.7 h) (Affer et al., 2012). Running a LS search, two clear extra signals appear at periods of 17.95 and 16.25 mins (see Figure 53). These last two frequencies, not detected in other stars in the field, are not pure harmonics of the dominant mode of 0.1193 days (a factor ~10.56× and ~9.56×, respectively), or of the other semi-sinusoidal modes, and could hint towards CoRoT 102850921 being a roAp star (Kurtz, 1982; Balona, 2011; Balona et al., 2012) or being affected by other rotational variabilities. Also, the excess flux in the transit egress could be due to tidal distortions of the binary system (Kumar, Ao, and Quataert, 1995; Thompson et al., 2012; Hambleton et al., 2018) (determining the nature of the star is beyond the scope of this work).

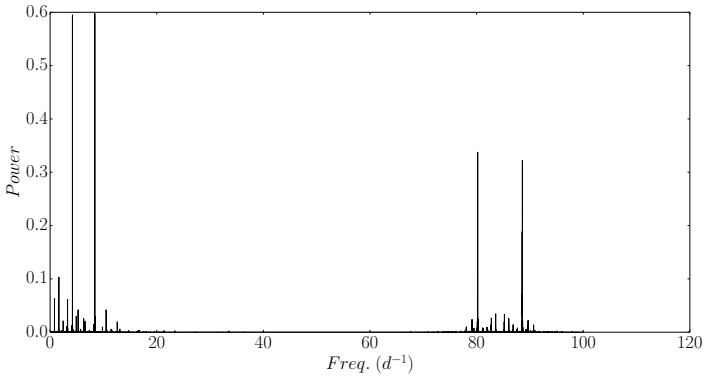


Figure 53: LS power spectrum for CoRoT 102850921 showing the two clear signals at 80.22 d^{-1} (17.95 min) and 88.61 d^{-1} (16.25 min).

Once the different signals have been detected during the *frequency analysis* steps, we run their *iterative reconstruction* with TFA and TFAW. Result are shown in Figure 54. As can be seen, if the signal period is correctly found, TFAW is able to directly (i.e. without pre-whitening) decouple the phase folded signal from the other signal contributions better than TFA and increase their SNR. Also, as demonstrated in Section 4.4.3, TFAW ensures that the signal shape is preserved.

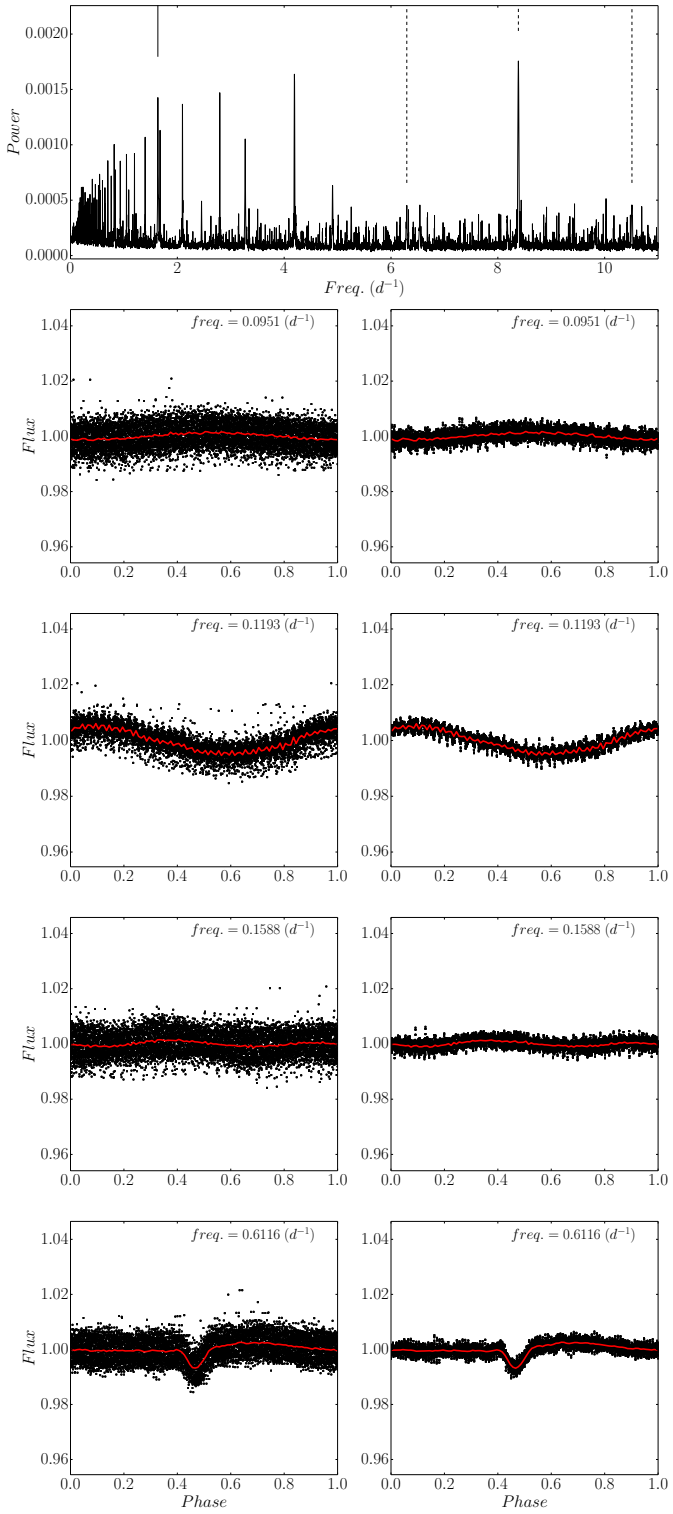


Figure 54: Top panel shows TFAW's BLS power spectrum for transit candidate CoRoT 102850921. The three dashed vertical lines mark the periods of the three sinusoidal signals described in Section 6.2.1.5. The solid vertical line corresponds to the transit period. The following plots show TFA's (left plots) and TFAW's signal reconstructions at the periods described in Section 6.2.1.5. Red lines correspond to the 100 binned phase folded light curve.

As explained in Section 3.3.1, the method developed to select the *signal level* and *noise level* has in mind the fact that all signal components (specially low frequency ones) present in the target light curve should be preserved. In the case of multiperiodic signals, if the *noise level* is selected to be the first decomposition scale (i.e. highest frequency), almost all the signal contributions (except those with the highest frequencies still coupled with the noise) will be conserved in the signal estimation and the *frequency analysis* step. If one is only interested in the low frequencies, one way of getting rid of the higher ones would be to increase the *noise level* to consider more of the lower decomposition scales. This way, during the *frequency analysis* step the low frequencies will no longer play a part and the signal estimation will be almost free of high frequency components. In order to preserve the shape of the low frequency signal, the *signal level* should still be selected following the method described in Section 3.3.1.

For CoRoT 102850921, we are interested in the low frequency transit. We can use the method explained before to remove the presence of the high frequency signals in order to: first, get a cleaner BLS power spectrum during the *frequency analysis* step and, second, to improve the signal characterization by diminishing the noise contribution and the effects of the high frequency components. Following Section 4.3.3, for CoRoT 102850921's *frequency analysis* step we have selected the *signal level* following the method described in Section 3.3.1 while the *noise level* has been set to consider the lowest **three** decomposition scales. Once the 0.6116 period has been recovered, the phase folded light curve is used to estimate a new *signal level* as per Section 4.3.4 and, again, the *noise level* has been selected to be the lowest **three** scales. Top panel in Figure 55 shows the BLS power spectrum once the high frequencies have been removed during the *frequency analysis* step. Comparing it with the top panel in Figure 54, it is clear that TFAW has efficiently removed all traces of the high frequency signal components while preserving the transit period. Bottom two plots in Figure 55 show TFAW light curves for *noise levels* equal to one (left) and three (right). As can be seen, the noise contribution is highly diminished for the second case compared to first one all the while, returning the correct transit shape thanks to the correct *signal estimation* (as explained in Section 4.3.4).

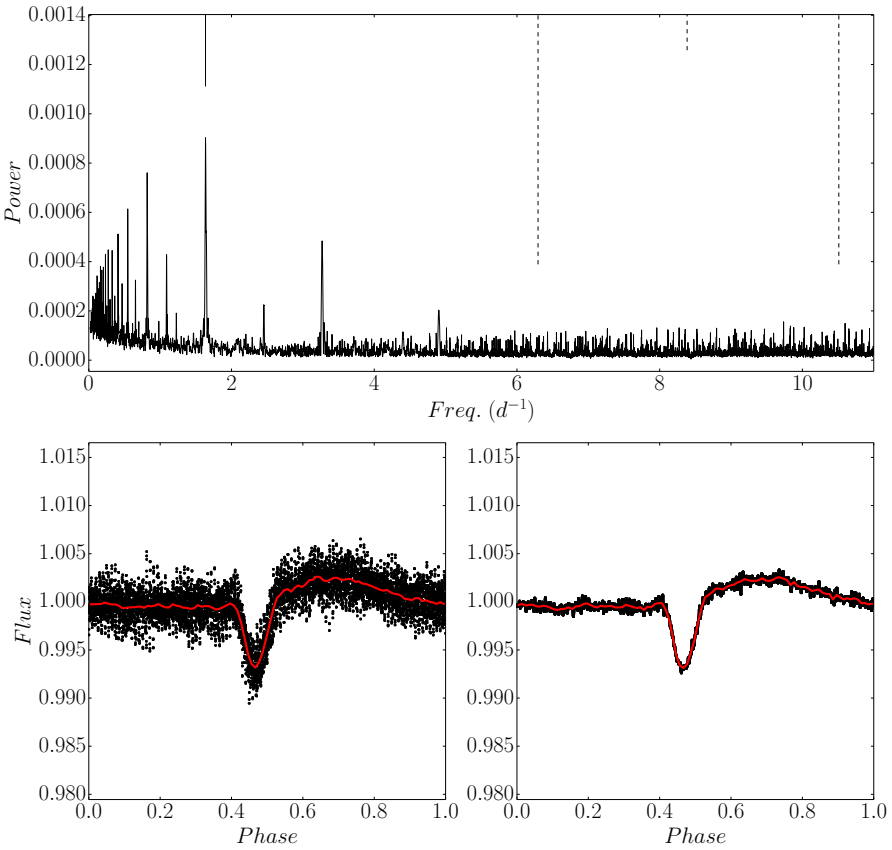


Figure 55: Top panel shows TFAW's BLS power spectrum for transit candidate CoRoT 102850921 after the high frequency signals have been removed as explained in 6.2.1.5. The three dashed vertical lines mark the periods of the three high frequency sinusoidal signals and the solid vertical one corresponds to the transit period. The bottom left plot shows TFAW's reconstructed signal for noise level equal to the lowest decomposition scale. Bottom right plot shows TFAW's reconstructed signal for noise level equal to the three lowest decomposition scales as described in 6.2.1.5. Red lines correspond to the 100 binned phase folded light curve.

6.2.2 *Kepler Data*

6.2.2.1 *Data description and objectives*

We downloaded the K2 mission monitoring campaign 1, K2-C1, Data Release 14 data from the MAST archive[4] (see Section 5.2.2.2 for more details). This field centered at (α, δ) = (11:35:45.51, 01:25:02.27) was observed between May 30 and Aug 21, 2014. During that time, Kepler observed 21,732 targets using the long cadence (LC) mode and 56 in short cadence (SC) (see Figure 56) as well as the trans-Neptunian Object 2002 GV31. Campaign C1 was the first full length observing cam-

4 https://archive.stsci.edu/pub/k2/lightcurves/

paign for the K2 mission where the targets were selected by peer reviewed proposals. For this work, we focus on the LC light curves of this campaign as the number of available template stars per CCD module is greater than for the SC data. We use the PDC-corrected (see Section 5.2.2.3 for more details) flux as starting point for our TFA and TFAW analysis.

K2-C1 comprises 41 confirmed planets in 32 planetary systems. We want to compare the photometric precision obtained with TFA and TFAW with that achieved by the `EVEREST` pipeline (Luger et al., 2016; Luger et al., 2017) for some of these objects.

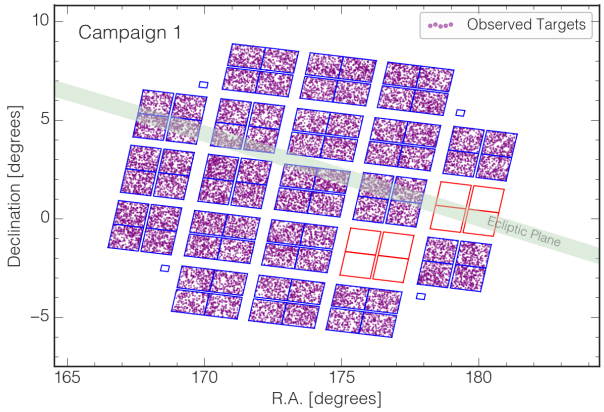


Figure 56: Schematic of K2's monitoring campaign C1 FoV with selected targets denoted by purple dots. Image credit: NASA.

6.2.2.2 *K2-44: a confirmed planet example*

K2-44 (EPIC201295312) is a V=12.19±0.12 mag (Zacharias et al., 2012) located at $(\alpha,\ \delta)$ = (11:36:02.79194, -02:31:15.16788) (Gaia Collaboration, 2018). It was observed by the K2 mission during the K2-C1 monitoring campaign from May 30 to Aug 21, 2014. K2-44 was first reported as a planetary hosting candidate by Montet et al. (2015). Doppler spectroscopy (Eylen et al., 2016) constrained the planet mass to $< 12 M_{\oplus}$. K2-35 was validated by Crossfield et al. (2016) as T_{eff}=5912±51 K, log g = 4.101±0.063, 1.58±0.15$R_{\odot}$ star orbited by a planet of 2.72±0.32$R_{\oplus}$ at 0.06511±0.011 AU with a period of 5.65688±0.00059 days. To characterize the host star they used Keck/HIRES optical spectra and applied the `SpecMatch` (Yee, Petigura, and von Braun, 2017) software (see Table 12 for more details). A companion star located 8" away from K2-44 was observed by Crossfield et al. (2016) with NIRC2[5] and PHARO (Hayward et al.,

5 https://www2.keck.hawaii.edu/inst/nirc2/ObserversManual.html

2001) on the Keck and Palomar 200 inch telescopes, respectively. Evans et al. (2018) observed the star in 2016 using the Two Colour Instrument on the Danish 1.54 m telescope (Skottfelt, J. et al., 2015). Combining their observations with measurements from 2MASS, Gaia DR1 and URAT1, they determine that the two stars are unassociated. The system was later reanalyzed by Mayo et al. (2018). They find a $2.392 \pm 0.262 R_{\oplus}$ planet orbiting the star at a distance of $9.3775^{+0.9217}_{-2.266163} R_{*}$ with a 5.6563 ± 0.0003 period.

Table 12: Stellar and planetary parameters obtained for K2-44 b by Crossfield et al. (2016).

K2 ID	EPIC201295312
Transit parameters	
Period P (days)	5.65688 ± 0.00059
Transit epoch To (BJD - 2454833)	1978.7176 ± 0.0044
Transit duration (hours)	4.36
Eccentricity e	0 (fixed)
Radius ratio p	0.0156 ± 0.0012
u_{+}	$0.3480 \pm 0.09^{\dagger}$
u_{-}	$0.2892 \pm 0.11^{\dagger}$
Scaled semi-major axis a (R_{*})	$8.856 \pm 1.49^{\star}$
Scaled semi-major axis a (AU)	0.0651 ± 0.011
Inclination i (^{o})	$87.3543^{+1.8561}_{-3.3003}$ *
Stellar parameters	
Stellar radius R_{*} ($R_{\odot}$)	1.58 ± 0.15
Stellar mass M_{*} ($M_{\odot}$)	1.15 ± 0.06
Effective temperature (K)	5912 ± 51
Surface gravity log g (dex)	4.101 ± 0.063
Metallicity [Fe/H]	0.0 (assumed)
Spectral Type	-
Planetary parameters	
Planetary radius R_{p} ($R_{\oplus}$)	2.72 ± 0.32
Planetary mass M_{p} (M_{J})	-
Planetary bulk density $\langle \rho \rangle$ (10^{3} kg/m^{3})	-

† denote values taken from Sing, D. K. (2010) assuming a 0.0 [Fe/H] metallicity.

* denote values taken from Mayo et al. (2018).

We want to compare the performance of TFA and TFAW applied to K2-44 b light curves with the EVEREST corrected one. We use 3,072 (i.e. yielding 10 decomposition scales) data points

obtained from the K2-C1, Data Release 14, PDC-corrected flux to build K2-44's light curve. K2-44 was observed with CCD channel 39 using K2's long cadence (LC) mode. To build TFA and TFAW's sample of template stars we compute the Stetson's L variability index for all the stars in the CCD channel. Looking at the distribution of L indexes for this CCD channel, we then select those (~300 stars) with a value <10 to avoid the inclusion of bona-fide variable stars in the sample. As can be seen in Figure 57, both TFA and TFAW find the period reported by Crossfield et al. (2016) for K2-44 b during their *frequency analysis* steps.

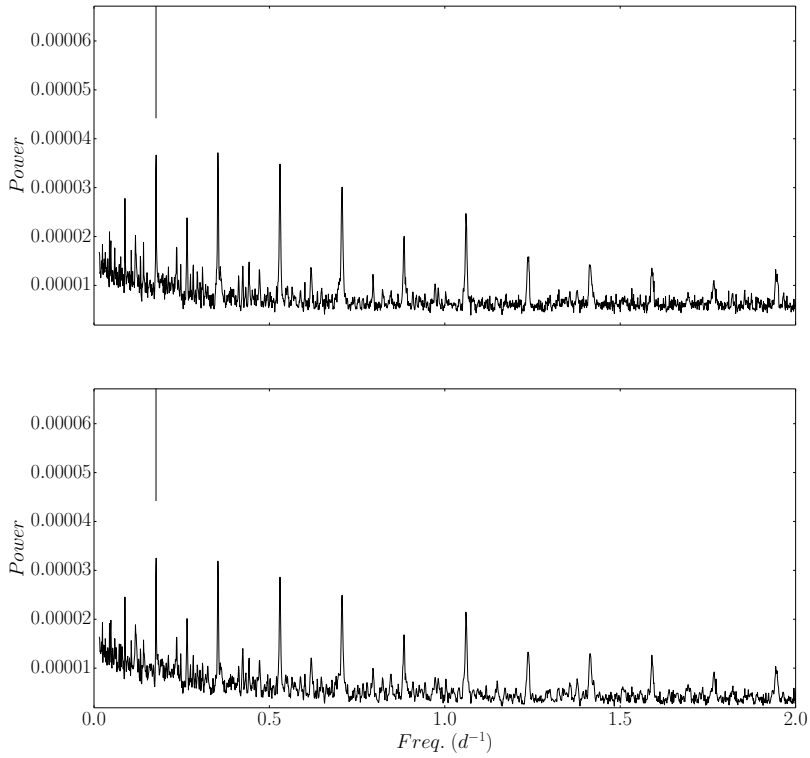


Figure 57: **Top**: TFA's BLS power spectrum for planet hosting star K2-44. **Bottom**: TFAW's BLS power spectrum for planet hosting star K2-44. The solid vertical line marks the planet period of 5.65688 as reported by Crossfield et al. (2016).

We then run the *signal reconstruction* setting the *noise level* to the lowest decomposition scale and the *signal level* to the scale determined by the method explained in Section 3.3.1 for the phase folded light curve of K2-44 b. Figure 60 shows the TFA- and TFAW-reconstructed light curves for K2-44 b compared to the EVEREST one. As can be seen, TFA is able to efficiently detrend and remove the star variability from the light curve but achieves a worse photometric precision (~1.4× worse) than EVEREST. As for TFAW, it is also able to detrend the light curve as TFA but, the SWT-based filter yields a better (~1.5×) photo-

metric precision than EVEREST. Also, the presence of outliers is highly diminished. Figure 58 shows a comparison of the phase folded light curves obtained for TFA and TFAW with the EVEREST one. For the latter, we have employed the method described in Luger et al. (2017) and have applied a Gaussian Process (GP) to detrend it before being phase folded (see Fig. 61 for a comparison between the GP-corrected EVEREST light curve and TFA and TFAW ones).

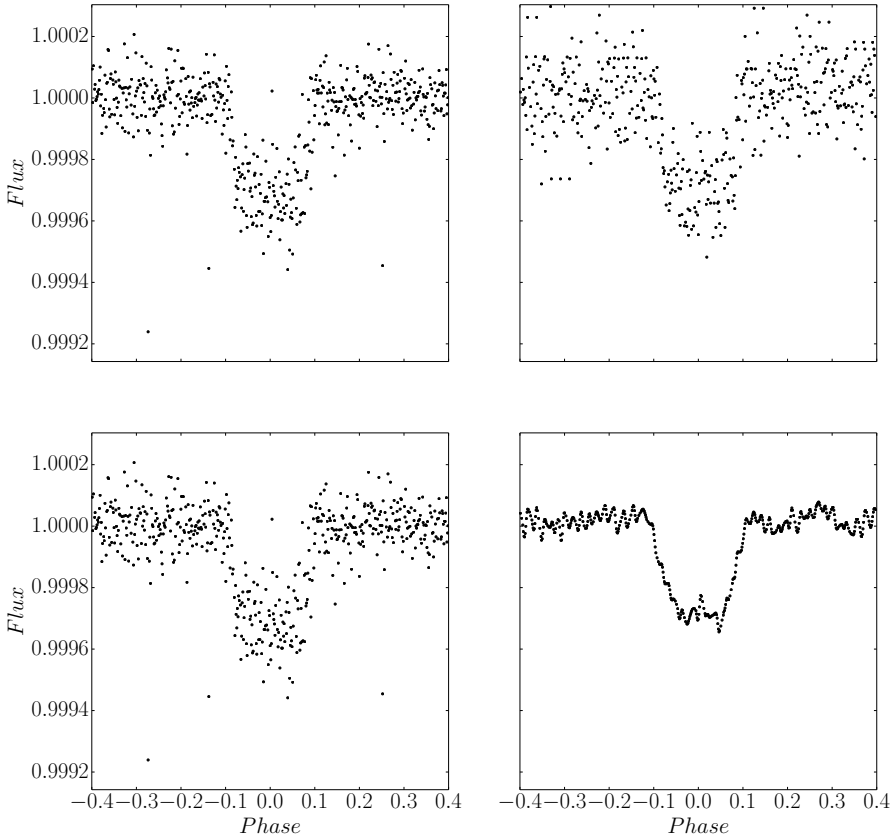


Figure 58: **Top left**: EVEREST light curve phase folded to the period of K2-44 b as reported by Crossfield et al. (2016). **Top right**: TFA-reconstructed and phase folded light curve for planet K2-44 b. **Bottom left**: EVEREST light curve phase folded for K2-44 b. **Bottom right** TFAW-reconstructed light curve for planet K2-44 b. The EVEREST light curve has been detrended applying the GP used by Luger et al. (2017) prior to the phase folding.

As in the case of the CoRoT light curves (see Sect. 6.2.1), we want to check whether the improved photometric precision yielded by TFAW results in a better characterization of the transiting signal or not. To do so, we analyze TFA and TFAW corrected light curves for confirmed exoplanet K2-44 b and compare the fitted parameters with the ones obtained by (Crossfield et al., 2016) (see Table 12). Again, to model the transit, we run a

MCMC fit using the sampler provided by `emcee` and the Mandel and Agol (2002) analytic model provided by the `batman` package. We solve for 5 transit parameters: the transit epoch, T0, the inclination of the orbit, i, the planetary radius to stellar radius ratio, p, the semi-major axis of the orbit, a and the period, P. In addition, we have fixed the quadratic limb darkening coefficients to their theoretical values taken from the tables by Sing, D. K. (2010). Using the stellar parameters provided by Crossfield et al. (2016) (see Table 12) and, assuming a metallicity of [Fe/H]=0.0, this values are l=0.3480 and q=0.2892. We also consider a circular orbit (i.e e=0) and assumed a longitude of the periastron of ω=90°. We run the MCMC fit using a uniform distribution of the priors and run the sampler with 200 chains and 10,000 iterations with a burn-in phase of 2,000 iterations. As in the CoRoT exoplanet transits, we do this to ensure that the chains run for enough autocorrelation times and that the final mean acceptance fraction is between 0.25 and 0.5. In Table 13 we compare the transit parameters obtained by Crossfield et al. (2016) with the ones obtained for TFA and TFAW posterior probability distributions. The fitted parameter values are obtained from the 50% quantile and their upper and lower errors are computed from the 25% and 75% quantiles respectively.

For the best fit, the time of inferior conjunction, T0, for TFA and TFAW is compatible with the one reported by Crossfield et al. (2016), though for TFAW, the uncertainties are much lower than for the other two (~5× for TFA and ~20× compared to Crossfield et al. (2016)). For the semi-major axis of the orbit, a, TFA returns a slightly lower value and, TFAW an even lower one than the one by Mayo et al. (2018). However, both of them are compatible within the errors with the reported value of 8.856±1.49R_*. In addition, TFAW returns smaller uncertainties for this parameter. For the orbit inclination, i, both TFA and TFAW yield values compatible with the reported value of $87.3543^{+1.8561}_{-3.3003}$ (Mayo et al., 2018). Though, again, TFAW returns smaller uncertainties. Regarding the planetary to star radius ratio, p, both TFA and TFAW obtain slightly bigger values than the validated value of 0.0156±0.0012. Nonetheless, TFAW gives a compatible value within errors with the Crossfield et al. (2016) one and, again with a smaller uncertainty. Finally, for the period, P, the values found are compatible within the errors with the one in Crossfield et al. (2016). With the best fit parameters, TFA obtains a mean planetary radius of $3.014^{+0.292}_{-0.289}$ $R_\oplus$ and TFAW $2.867^{+0.272}_{-0.289}$ $R_\oplus$, slightly above the 2.72±0.32 $R_\oplus$ reported by Crossfield et al. (2016) but still, compatible within the errors.

Taking into account the results obtained for simulated light curves in Sect. 4.4.5, the TFAW fit for the semi-major axis of the orbit, the inclination and the planet to star radius ratio, should yield a closer representation to the real transit parameters. For K2-44 b, all transit parameters and the derived mean planetary radius are compatible with the ones given by Crossfield et al. (2016) and Mayo et al. (2018) but the results yielded by TFAW present lower uncertainties and, as discussed before, the parameters should be closer to the real ones. As in the case of CoRoT-21 b, we have confirmed the correct performance of TFAW's de-noising and signal reconstruction applied to a confirmed exoplanet of the K2 survey. Figure 59 shows the best fit obtained for TFA and TFAW over-plotted to their corresponding phase folded light curves.

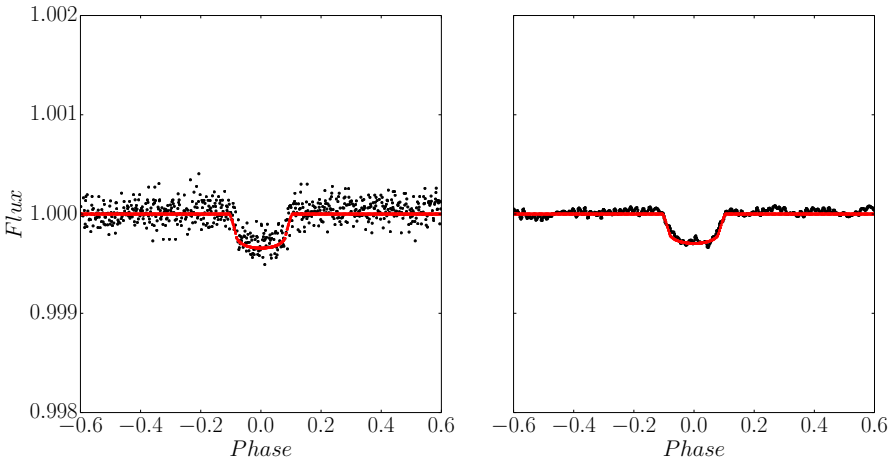


Figure 59: **Left**: Phase folded TFA light curve for planet K2-44 b. The solid red line corresponds to the best fit given by the parameters in Table 13. **Right**: Phase folded TFAW light curve for K2-44 b. As before, the solid red line corresponds to the best fit for TFAW following the parameters in Table 13.

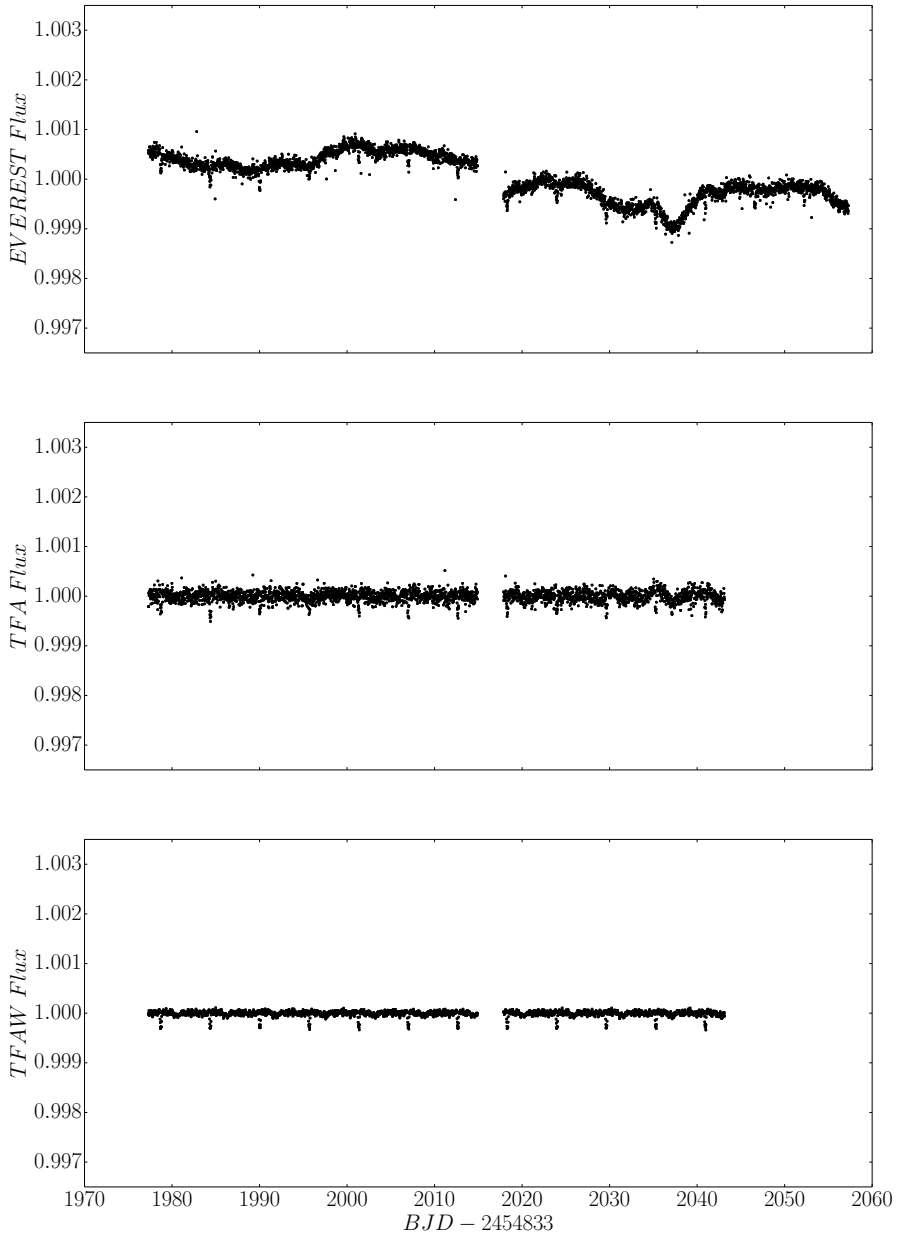


Figure 60: **Top**: EVEREST light curve for planet hosting star K2-44. **Middle**: TFA-reconstructed light curve for planet hosting star K2-44. **Bottom**: TFAW-reconstructed light curve for planet hosting star K2-44. Notice that the EVEREST light curve has a longer time base than the ones for TFA and TFAW due to 3,072 data points selected to run the SWT-based filter.

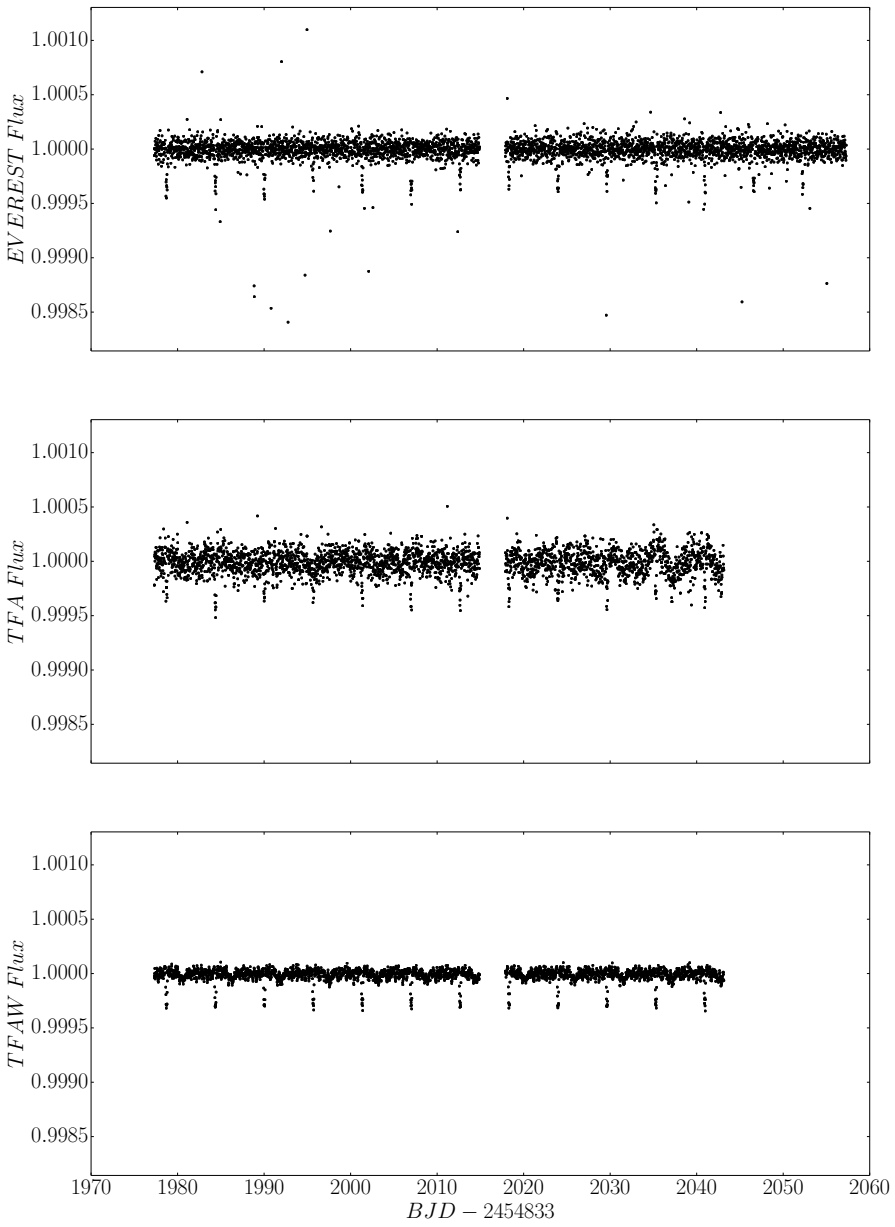


Figure 61: Same as Fig. 60 but, this time, showing the GP-corrected light curve EVEREST in top panel as per Luger et al. (2017).

Table 13: Top table: K2-44 b parameters from Crossfield et al. (2016). Posterior transit parameters values and their uncertainties (with the 25% and 75% quantile as the upper and lower errors) for TFA and TFAW's MCMC fits. Middle table: 95% confidence intervals of the highest probability density for K2-44 b transit parameters TFA and TFAW's MCMC fits. Bottom table: Derived planetary parameters for TFA and TFAW.

Parameters	To (BJD-2454833)	P (days)	a (R_*)	p	i (º)
Crossfield et al. (2016)	1,978.7176±0.0044	5.65688±0.00059	8.856±1.49*	0.0156±0.0012	$87.3543^{+1.8561}_{-3.3003}$ *
TFA MCMC	$1{,}978.7254^{+0.0009}_{-0.0011}$	5.65529±0.00014	$8.37313^{+0.57348}_{-0.71798}$	$0.01748^{+0.00035}_{-0.00028}$	$86.36954^{+0.92358}_{-1.11319}$
TFAW MCMC	1,978.7192±0.0002	5.65662±0.00003	$7.37125^{+0.14602}_{-0.15121}$	0.01663±0.00009	$84.84727^{+0.23430}_{-0.24664}$
95% confidence intervals of the highest posterior density					
TFA MCMC	1,978.7118 - 1978.7275	5.65498 - 5.65694	7.14021 - 9.19985	0.01700 - 0.01817	84.32285 - 87.79344
TFAW MCMC	1,978.7195 - 1978.7203	5.65656 - 5.65668	7.06853 - 7.66845	0.01647 - 0.01680	84.34842 - 85.32247

System parameters	R_p ($R_\oplus$)	b
Crossfield et al. (2016)	2.72±0.32	-
TFA	$3.014^{+0.292}_{-0.289}$	$0.53^{+7.71}_{-9.29}$
TFAW	$2.867^{+0.272}_{-0.289}$	$0.66^{+1.72}_{-1.81}$

6.2.2.3 *K2-35: a multiplanetary example*

K2-35 (EPIC201549860) is a V=14.35±0.06 mag (Henden et al., 2016) K4V star (Dressing et al., 2017a) located at $(\alpha,\ \delta)$ = (11:20:24.74009, +01:17:09.42617) (Gaia Collaboration, 2018). It was observed by the K2 mission during the K2-C1 monitoring campaign from May 30 to Aug 21, 2014. K2-35 hosts two confirmed close-in super-Earths. The outer planet, orbiting with a period of 5.6 days, was reported as a planet candidate by Foreman-Mackey et al. (2015) and Montet et al. (2015). This, and the inner one, orbiting with 2.4 days, were listed as planet candidates by Vanderburg et al. (2016). The system was validated by Sinukoff et al. (2016) as a ~K4, T_{eff}=4680±60 K, 0.72±0.04 $R_\odot$ star orbited by an inner planet of 1.40±0.17 $R_\oplus$ and an outer planet of $2.09^{+0.43}_{-0.31}$ $R_\oplus$. To characterize the host star they used Keck/HIRES optical spectra and applied the `SpecMatch` (Yee, Petigura, and von Braun, 2017) and `isochrones` (Morton, 2015) packages. Sinukoff et al. (2016) also confirmed, using Pan-STARRS1 3π survey (Kaiser et al., 2010), that no other sources fall within the 12" K2 photometry aperture to a limiting magnitude of r_{P1}=22 mag. Dressing et al. (2017a), using SpeX NIR spectra classify the host star as a K4V star with a $T_{eff}=4402^{+96}_{-93}$ K and a radius of 0.62±0.03 $R_\odot$. Dressing et al. (2017b) refitted the transit photometry to these new values and obtained a revised planetary radii of 1.32±0.08 $R_\oplus$ for the inner planet and $1.93^{+0.13}_{-0.11}$ $R_\oplus$ for the outer one.

As in the case of K2-44 b (see Section 6.2.2.2), we want to check the performance of TFA and TFAW applied to K2-35 b and K2-35 c light curves and compare them with the `EVEREST` ones. We also use 3,072 data points obtained from the K2-C1, Data Release 14, PDC-corrected flux to build K2-35's light curve. This K2 target was observed with CCD channel 63 using K2's long cadence (LC) mode. To build TFA and TFAW's sample of template stars we compute the Stetson's L variability index of all the stars in the CCD channel. We then select those with a value <10 to avoid the inclusion of bona-fide variable stars in the sample. During TFA and TFAW's *frequency analysis* steps, both periods are found in the BLS power spectrum; being the one for the outer planet the most significant one (see Figure 62). Once both periods have been detected, we run their *signal reconstruction* steps with TFA and TFAW. The *signal level* for both phase folded light curves have been set using the method explained in Section 3.3.1. The *noise level* has been set to the lowest decomposition level for both transits. Figure 63 shows K2-35 light curves for TFA and TFAW after the iterative *signal reconstruction* has been applied to the phase folded light curve of the outer

planet compared to the EVEREST one. As can be seen, though TFA efficiently detrends the light curve, the final photometric precision is worse (~1.6×) than the one achieved by EVEREST. On the other hand, TFAW uses TFA's capabilities to detrend the light curve while, at the same time, use the SWT-based filter to minimize the effects of the high frequency noise yielding an improved precision (~1.5×) with respect the EVEREST light curve. In addition, TFAW has been able to remove almost all the outliers thanks to the method explained in Section 3.4. Doing the same but applying the iterative *signal reconstruction* for the phase folded light curve of the inner transit yields similar results in terms of the photometric precision achieved both with TFA (~1.5× worse) and TFAW (~1.8× better) compared to EVEREST. Figures 65 and 66 show a comparison of the phase folded TFA and TFAW light curves for K2-35 b and K2-35 c, respectively, with the phase folded EVEREST ones. As for the K2-44 EVEREST light curve, the GP method described in Luger et al. (2017) has been applied to detrend it before being phase folded (see Fig. 64 for a comparison between the GP-corrected EVEREST light curve and TFA and TFAW ones).

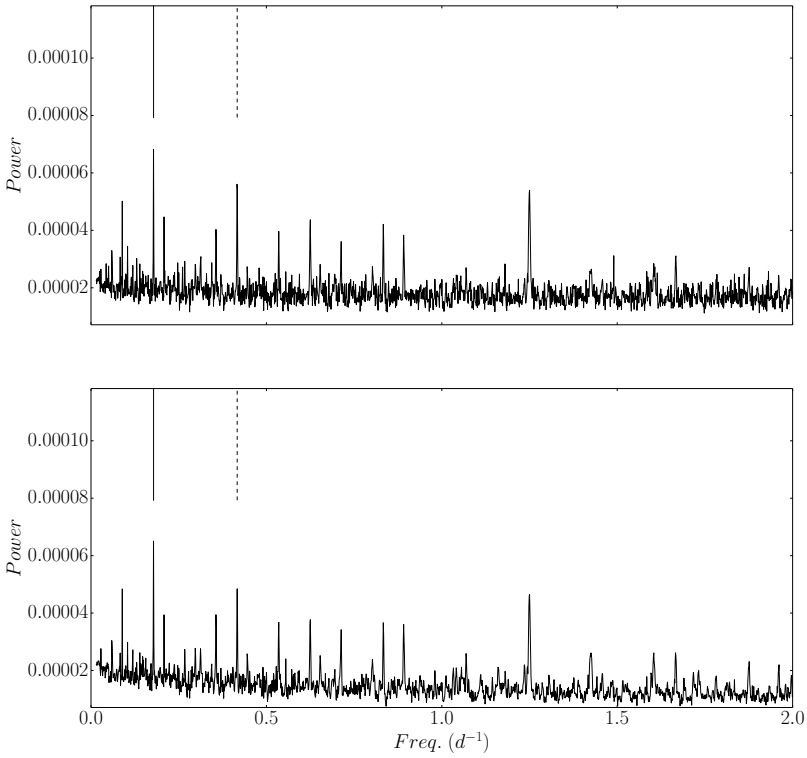


Figure 62: **Top**: TFA's BLS power spectrum for planet hosting star K2-35. **Bottom**: TFAW's BLS power spectrum for planet hosting star K2-35. The solid vertical line marks the outer planet period of 5.60835 days while the dashed vertical line marks the inner planet period of 2.39996 days as reported by Dressing et al. (2017b).

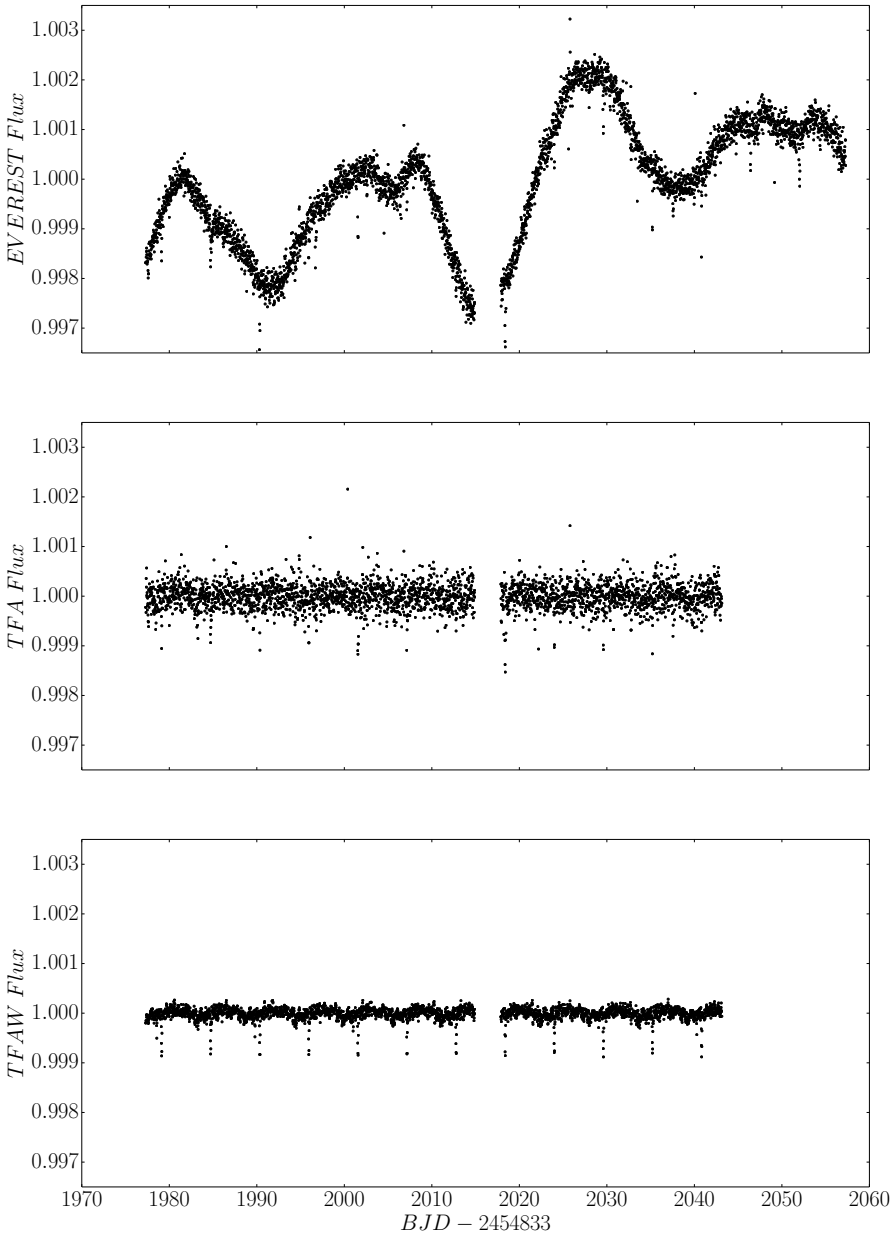


Figure 63: **Top**: EVEREST light curve for planet hosting star K2-35. **Middle**: TFA-reconstructed light curve for planet hosting star K2-35. **Bottom**: TFAW-reconstructed light curve for planet hosting star K2-35. Notice that the EVEREST light curve has a longer time base than the ones for TFA and TFAW due to 3,072 data points selected to run the SWT-based filter.

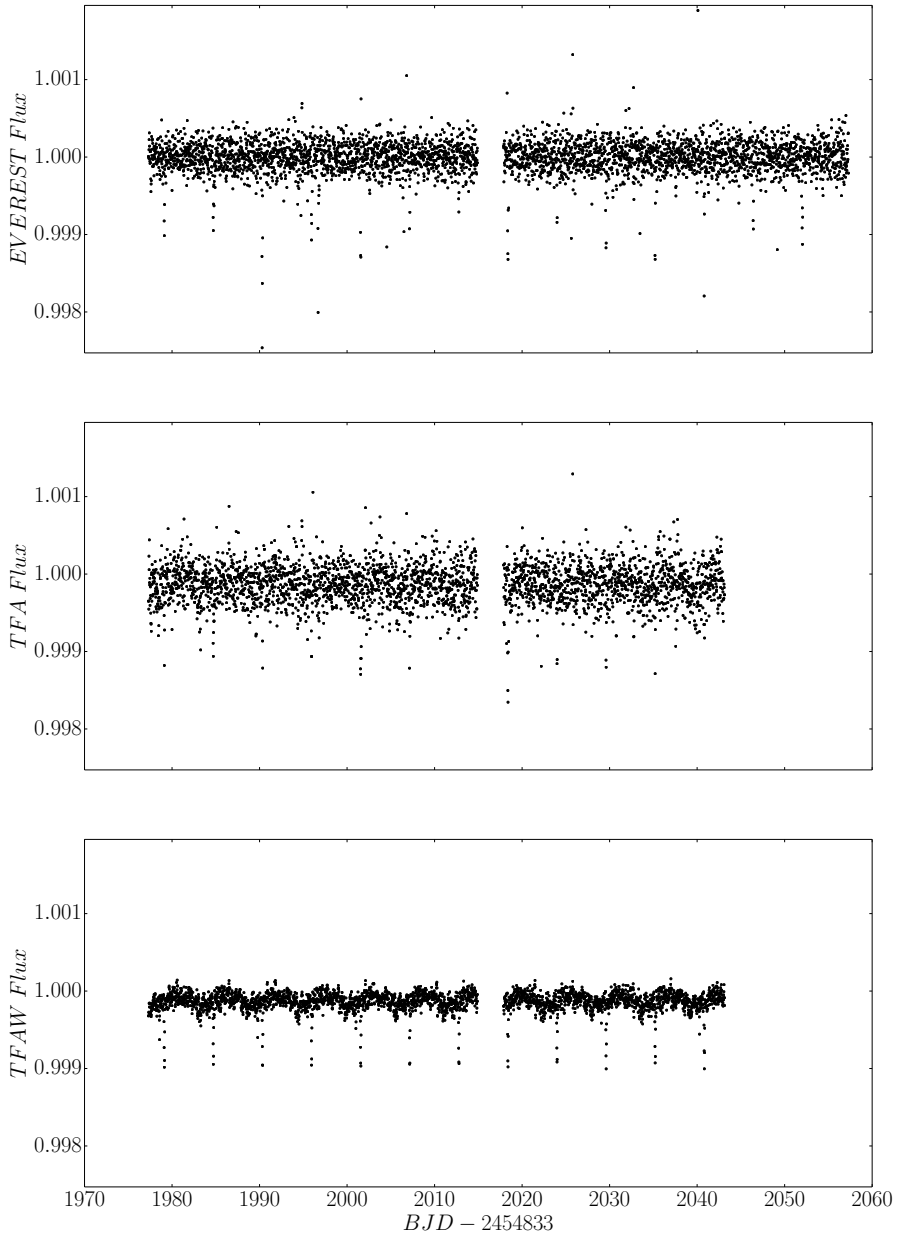


Figure 64: Same as Fig. 63 but, this time, showing the GP-corrected light curve EVEREST in top panel as per Luger et al. (2017).

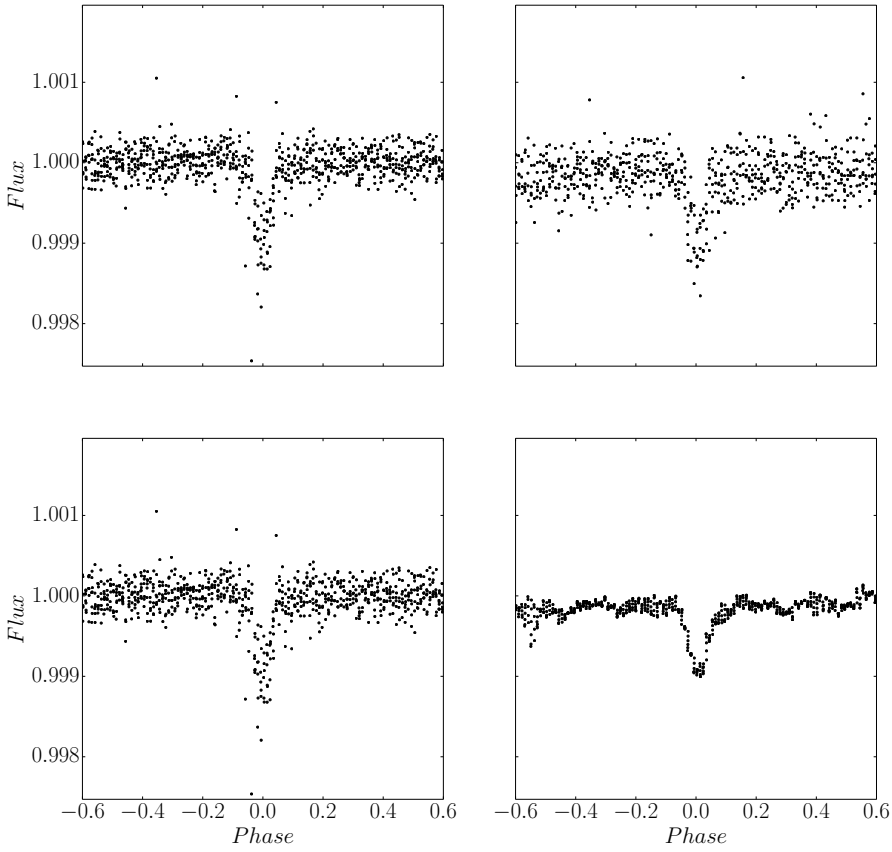


Figure 65: **Top left**: EVEREST light curve phase folded to the period of K2-35 b as reported by Dressing et al. (2017b). **Top right**: TFA-reconstructed and phase folded light curve for planet K2-35 b. **Bottom left**: EVEREST light curve phase folded for K2-35 b. **Bottom right** TFAW-reconstructed light curve for planet K2-35 b. The EVEREST light curve has been median filtered to detrend it prior to the phase folding.

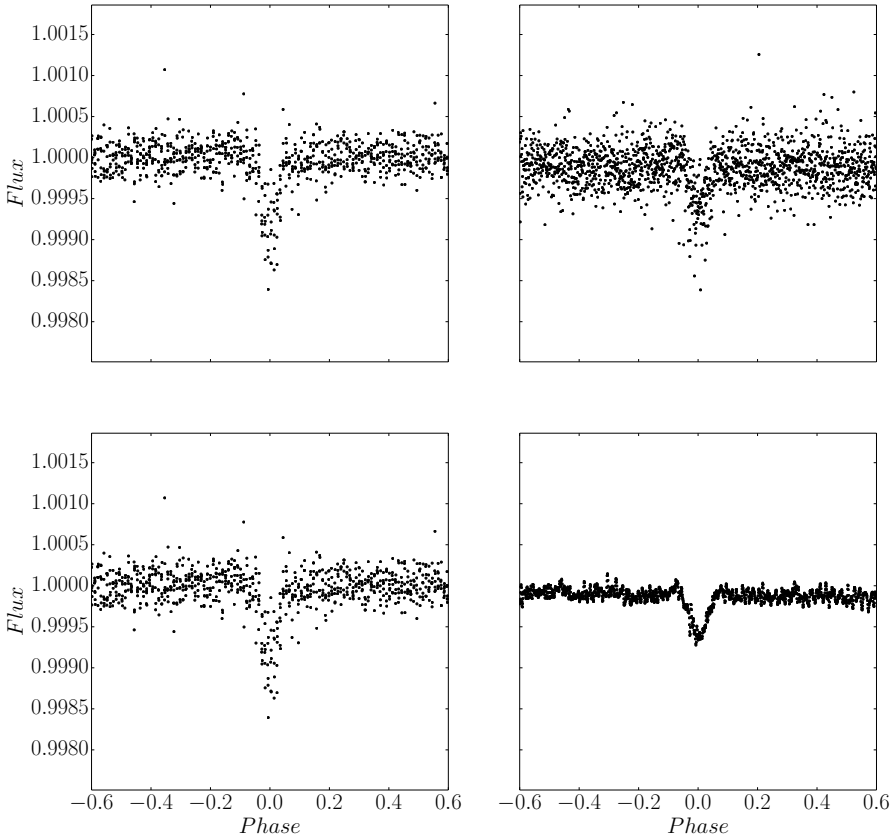


Figure 66: **Top left**: EVEREST light curve phase folded to the period of K2-35 c as reported by Dressing et al. (2017b). **Top right**: TFA-reconstructed and phase folded light curve for planet K2-35 c. **Bottom left**: EVEREST light curve phase folded for K2-35 c. **Bottom right** TFAW-reconstructed light curve for planet K2-35 c. The EVEREST light curve has been median filtered to detrend it prior to the phase folding.

Part IV

CONCLUSIONS

7 CONCLUSIONS

We began this PhD thesis with a brief review of the history of exoplanets and variable stars and their current discovery status. We also introduced the several limiting factors that affect most planetary transit and/or variability detection surveys. We presented wavelets as a perfect tool to analyze signals of non-stationary nature affected by singularities and noise thanks to their irregular shape and their compactly supported nature. With this we set up the framework in which to present the main contribution of this work, the Wavelet-based Trend Filtering Algorithm (TFAW), a Stationary Wavelet Transform (SWT)-based modification of the Trend Filtering Algorithm (TFA) (Kovács, Bakos, and Noyes, 2005). In this chapter we briefly summarize the main characteristics of the algorithm as well as list the most relevant results obtained from its application to simulated and real light curves.

7.0.1 *TFAW summary*

- TFAW is a totally generic, Python-based, parallelized algorithm useful to improve the performance of signal detection, reconstruction and characterization, provided that a set of comparison light curves sharing the same systematics and trends as the target time series is available.
- TFAW differs from other wavelet-based noise-filtering algorithms in that it does not require any parametric model fitting or any extra computational method. TFAW estimates the noise contribution of the signal from its SWT at each iteration step and the de-noising is done through the subtraction of this contribution from the signal. TFAW denoises the signal without modifying any of its intrinsic properties contrary to wavelet coefficient thresholding that can lead to distortions of the signal and introduce artificial oscillations or ripples around discontinuities.
- TFAW needs an even number of data points to work. The length of the target light curve will depend on the number of levels in which the signal wants to be decomposed (ideally, the length of the time series should be a power of two). The higher the available number of decomposition

levels, the better the different signal and noise contributions can be separated in the SWT decomposition of the signal. As future work we will study different ways of signal padding (like zero-, smooth-, or periodic-padding) and its effects to extend the original signal up to its closest power of two.

- TFAW modifies TFA's *frequency analysis* step to improve the detection of periodic variabilities in the target light curves. This improvement is achieved by the application of a SWT-based outlier-removal tool as well as the search of periods over a detrended and SWT-filtered version of the original light curve (i.e. the *estimated signal*).
- Once a significant period is found, the light curve is phase folded. A more accurate and precise estimation of the variable signal's shape is obtained through the SWT than with bin averaging.
- The algorithm decouples a large component of the noise contribution from the original signal, making use of a modified version of the original TFA's *iterative signal reconstruction* step. This de-noising is based on the SWT estimation of the noise at each iteration step, leading to an overall SNR improvement without alteration of the signal's time sampling or astrophysical characteristics.
- Compared to the original TFA, TFAW requires fewer reference stars to create the template used for the signal reconstruction and the noise filtering. Kovács, Bakos, and Noyes (2005) suggests the use of ~600 template stars, while TFAW operates stably with 200-300 template stars.

7.0.2 *Outlier detection*

- A custom distance-based outlier detection and removal method has been developed for this work.
- It uses the *estimated signal* computed from the SWT of the target light curve and the *signal level*.
- A threshold value can the be selected in such a way that any point in the time series whose distance to the *estimated signal* exceeds it is considered as an outlier and removed.
- We have proved that this method can be effectively used to remove any outlier without modifying the signal of interest.

7.0.3 *TFAW's performance summary: simulated data*

- Tests conducted on simulated TFAW-filtered light curves show an improvement of ~40% (although it can be higher) in their standard deviations with respect to the ones detrended with TFA, leading to a better characterization of the signal, without modifying its features (i.e. amplitude, phase, shape or depth).
- Results obtained for the *frequency analysis* step show that TFAW does not introduce any false periodicity, and that it can improve the overall power spectra of non-multiperiodic signals and the SDE of the peaks.
- In the case of simulated transits, TFAW's *frequency analysis* step improves the transit detection rate a factor ~2-5× for the low SNR signals with respect TFA and increases the SDEs up to a factor ~2.5×. In addition, TFAW is able to detect the transits for signals with ~2× higher standard deviation.
- We show that the SWT signal approximation provides a closer representation of the underlying signal with respect to bin averaging. In the case of planetary transits, this improvement is due to a better fit of the ingress and egress profiles.
- We demonstrate that the TFAW-filtered light curve yields better MCMC posterior distributions, diminishes the bias in the fitted transit parameters and their uncertainties and narrows the credibility intervals up to a factor ~10× for simulated transits.
- We show that TFAW is able to isolate the different underlying signals within a light curve with multiple periodic signals, such as multi-transit signals, transients, modulations or other kinds of stellar variabilities. As for single-period light curves, the SNR of each signal contribution improves due to the noise filtering capabilities of TFAW.

7.0.4 *TFAW's performance summary: real light curves*

- We have applied TFAW to two different ground-based survey data-sets (TFRM-PSES and The Evryscope) and two space-based surveys (CoRoT and Kepler). Each of them affected by different kinds of systematic effects and sources of noise.

For all these, TFAW is able to, first, detrend and, second, efficiently remove outliers and de-noise the light curves.

- As with the simulated light curves, TFAW-filtered light curves show an improvement of ~40% (although it can be higher) in their standard deviations with respect to the ones detrended with TFA.
- We have demonstrated TFAW's capability of de-noising different variabilities coming from diverse astrophysical sources ranging from δ-Scuti stars, fast rotating white dwarfs, eclipsing binaries or planetary transits. As in the case of simulated light curves, TFAW neither introduces any false signal or modifies the features of the variable signal.
- We have also shown TFAW's potential to improve the characterization and classification of variable signals. The example of CoRoT 102881832 shows how TFAW can help discriminate between a possible transit candidate and an eclipsing binary by increasing the SNR of shallow secondary transits. Also, the example of CoRoT 102588881, shows that TFAW-filtered light curves can yield more realistic planetary scenarios than TFA.
- We demonstrate that the TFAW-filtered real light curves yield better MCMC posterior distributions, diminish the bias in the fitted transit parameters and their uncertainties and narrow the credibility intervals.
- We have shown TFAW's capabilities of decoupling the different signal contributions in multiperiodic light curves both for CoRoT (CoRoT 102850921) and K2 (K2-35).
- For K2 light curves, we have demonstrated that TFAW can yield better photometric precision than `EVEREST` one of the most successful pipelines for this mission while at the same time, correctly detrending the light curve.

7.0.5 *Other results*

- We present 2MASS J10144313+5018191, a non-cataloged possible δ-Scuti star with a 0.1592-day period.
- We present 2MASS J13190996-8347115, a semi-detached contact binary of the W Ursae Majoris-type with a 0.658-day period. Detected during visual inspection of The Evryscope's light curves it was not cataloged in 2017 but was later confirmed by the ASAS-SN survey.

- We reinforce CoRoT 102881832 eclipsing binary nature by increasing the SNR of its shallow secondary transit.
- We obtain more realistic planetary parameters for CoRoT 102588881. A high orbital eccentricity, e=0.47384$^{+0.04982}_{-0.04997}$, planet of 1.20±0.28 R_J and 0.98±0.42 M_J in agreement with the RV observations.
- We study the multiperiodic nature of CoRoT 102850921 detecting three semi-sinusoidal components, the transit signal and two extra very high frequency signals that could hint towards the possible roAp nature of the host star. In addition, possible tidal distortions are observed during the transit egress.
- We study K2's confirmed planetary system K2-44 b and confirmed multiplanetary system K2-35, and demonstrate that TFAW yields better photometric precision than EVEREST.

Part V

APPENDIX

A

OTHER OBSERVATIONAL RESULTS

The observational results presented on this thesis have been focused on a few examples of each described exoplanet survey. However, as a by-product, many other objects have been analyzed by the TFAW. The main objective of this appendix is to provide an overview of some of the extra observational results obtained in the context of this thesis for objects showing some kind of variability in the TFRM-PSES and The Evryscope data sets. For a detailed description of the variability types presented in the following tables, we point out to the VSX description file available at `https://www.aavso.org/vsx/index.php?view=about.vartypes` and Samus' et al. (2017).

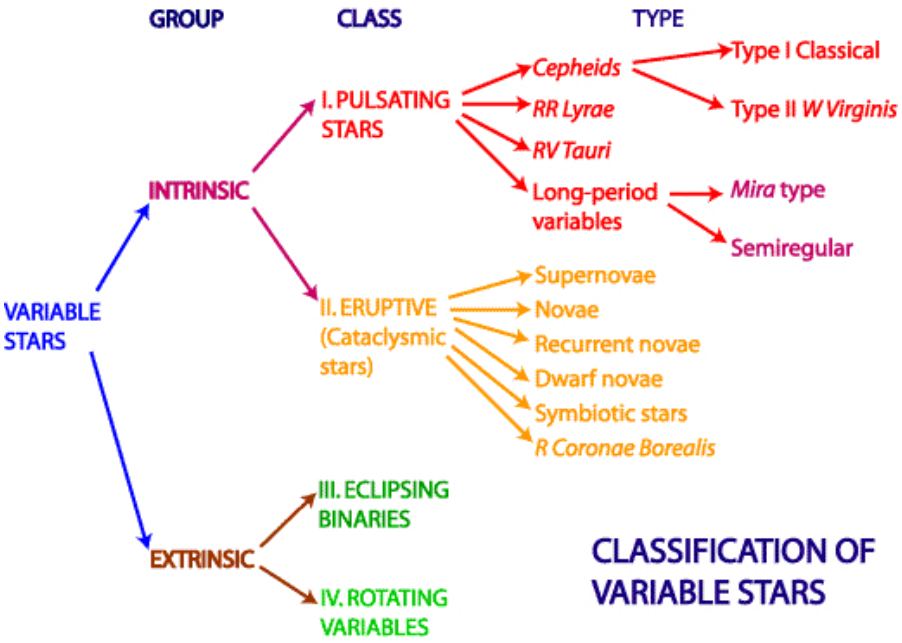


Figure 67: Classification diagram of stellar variabilities. Credit: ©Commonwealth Scientific and Industrial Research Organisation, 2015-2017

A.1 TFRM-PSES VARIABLES

Here we present the variables detected in the TFRM-PSES dataset. We have studied three of the TFRM-PSES fields centered at (α, δ) = (10:14:44, +48:30:00), (15:55:59, +35:04:00) and (10:18:51, +55:20:00), respectively. The light curves consist on 2,048 points coming from 30 nights observed during 2013, 2014 and 2015. As in the examples shown in Sect. 6.1.1.1, the light curves were generated with the APEX-2 reduction pipeline (Devyatkin et al., 2010) and had differential photometry applied prior to the application of TFA and TFAW. After visual inspection of those light curves with a SDE>6, 30 bona-fide variables were detected.

Of those, 24 (80%) were previously cataloged stars of different variability types. They are listed in Table 15 with more information regarding their coordinates, their TFAW periods, their cataloged periods, magnitude variations and variability types and the corresponding literature references. Their TFA and TFAW processed light curves are shown in Fig. 75 to 98.

Of the remaining 6 (20%) non-cataloged stars, 3 present sinusoidal like variabilities, 1 presents irregular sinusoidal features, 1 is an X-ray emitting star with irregular variability and the remaining one presents a transit like feature. They are listed in Table 14 alongside information about their TFAW periods, their positions in the sky and some notes regarding their variability nature. Their TFA and TFAW light curves are shown in Fig. 69 to 74.

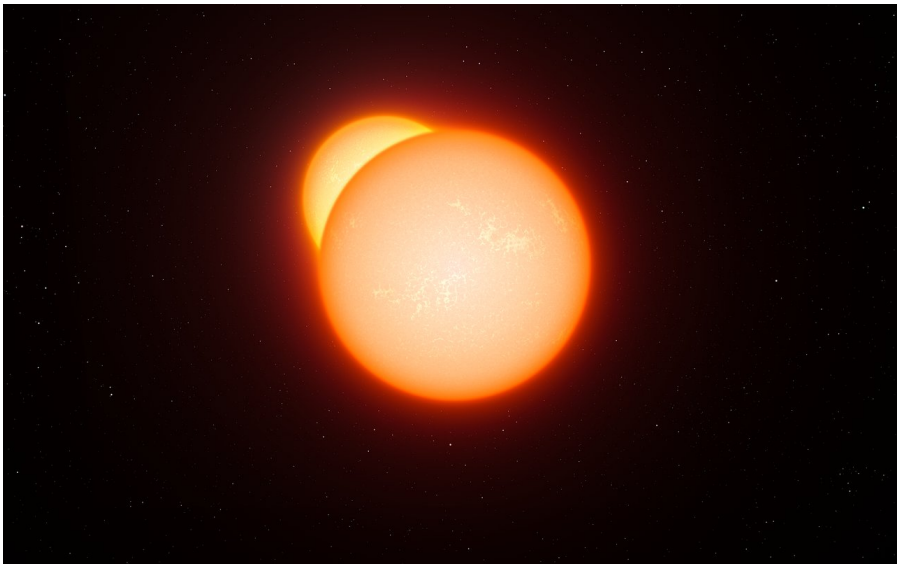


Figure 68: Artist's impression of an eclipsing binary system. Credit: ESO/L. Calçada

Table 14: TFRM-PSES non-cataloged variables and transiting candidates.

ID (Gaia DR2 or other)	α (J2000.0)	δ (J2000.0)	P_{TFAW} (d)	Notes
TYC 3437-773-1	151.7648	49.5481	0.28606	Sinusoidal like feature
847832561481232000	153.6796	50.3052	0.68564	Irregular sinusoidal like feature
822363194960911104	150.8177	47.2917	0.03533	Sinusoidal like feature
823735110594461824	151.7068	49.6116	0.63715	Transit like feature
SDSS J155651.09+351646.3	239.2130	35.2796	0.30751	Sinusoidal like feature
1RXS J101342.0+571451	153.4279	57.2458	1.30403	X-ray-emitting star

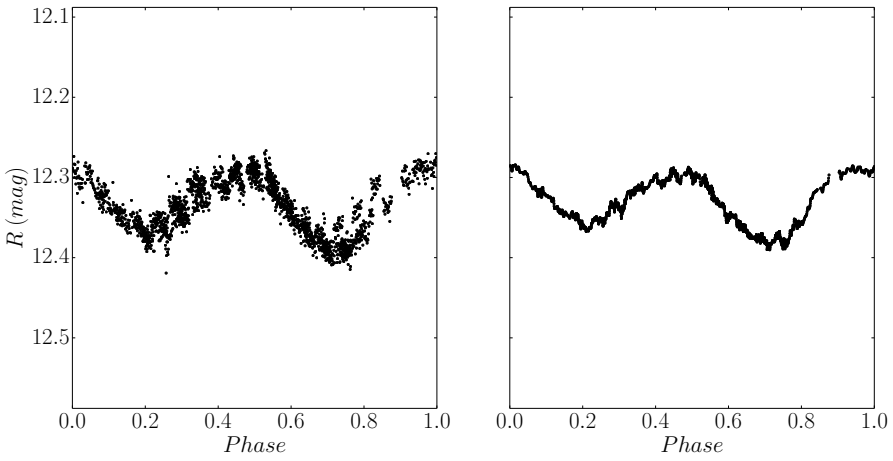


Figure 69: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable TYC 3437-773-1.

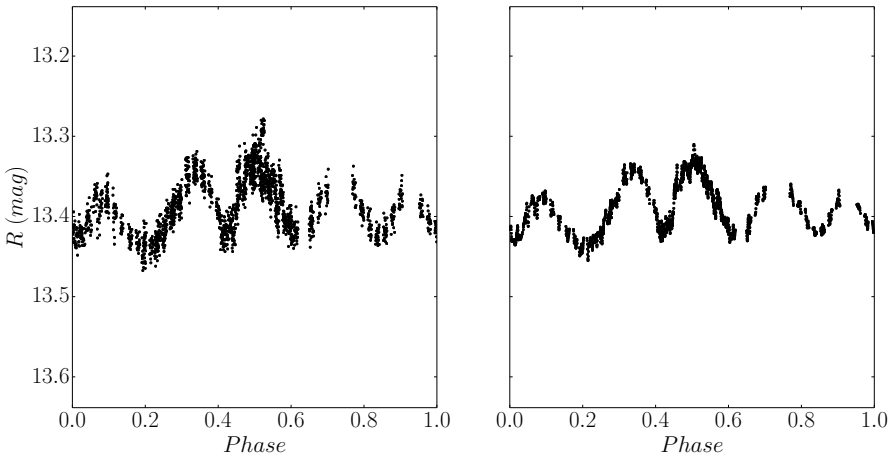


Figure 70: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable Gaia ID 847832561481232000.

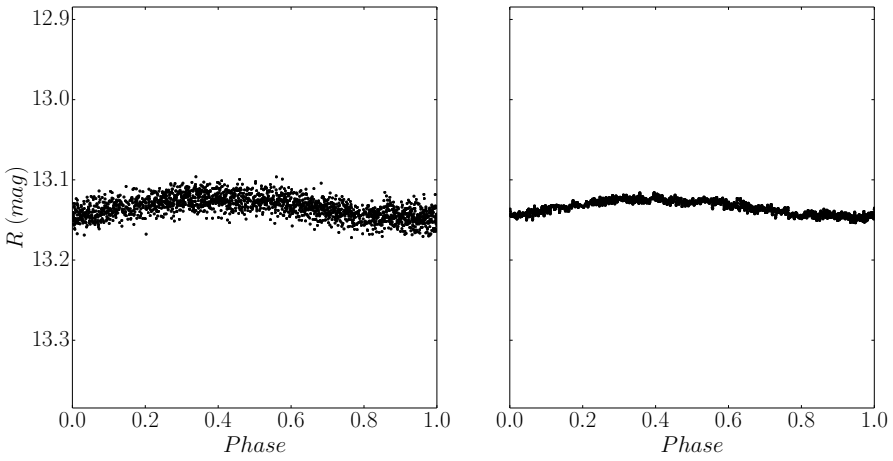


Figure 71: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable Gaia ID 822363194960911104.

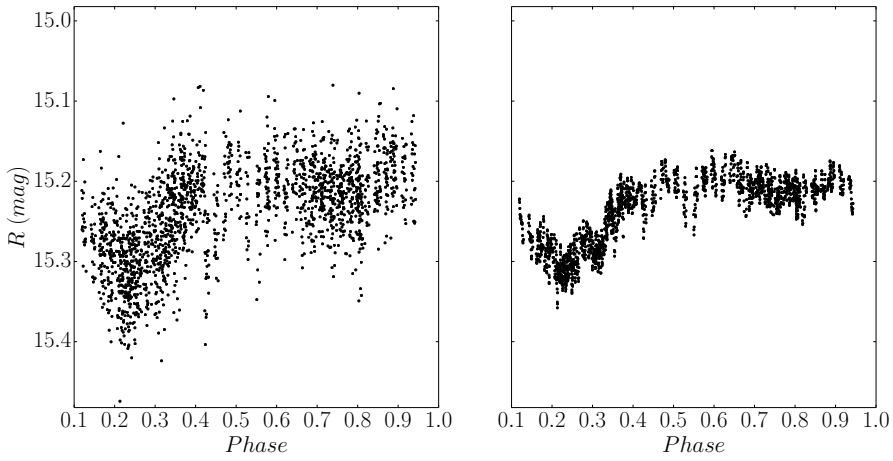


Figure 72: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable Gaia ID 8237351105944618 24.

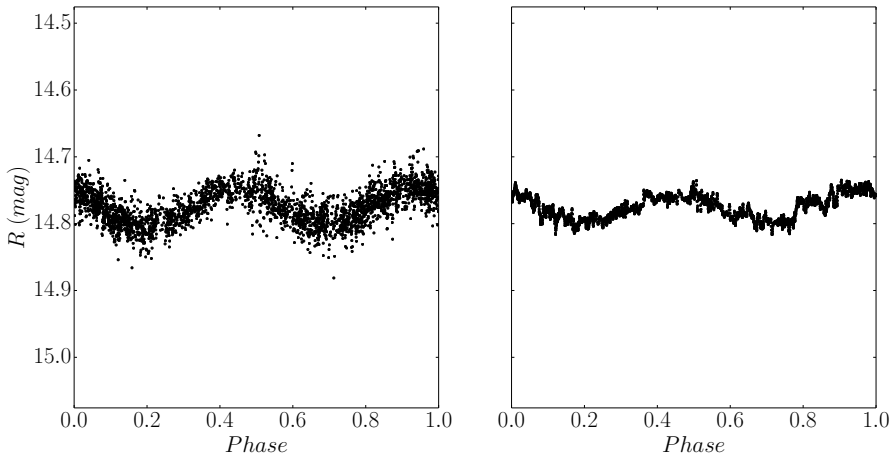


Figure 73: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable SDSS J155651.09+351646.3.

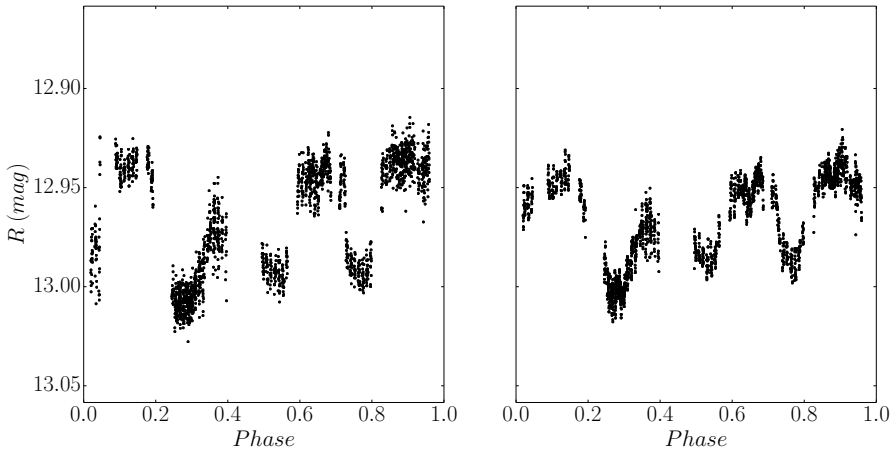


Figure 74: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable 1RXS J101342.0+571451.

Table 15: TFRM-PSES cataloged variables

Cataloged ID	α (J2000.0)	δ (J2000.0)	P_{TFAW} (d)	Cataloged var. type	Cataloged mag. var.	Cataloged P (d)	Cataloged ref.
NSVS 4922327	152.3089	48.1112	0.203129	EB|EW	12.047 - 12.397 (R1)	0.406257	2006AJ....131..621G
LINEAR 22079148	154.7320	50.4485	0.371819	EW	14.42 (V)	0.372173	2013AJ....146..101P
TSVSC1 TN-N232123120-15-67-2	153.3788	49.8128	0.250283	EW	13.72 (0.70) CV	0.2502057	2014ApJS..213....9D
NSVS 4924202	153.5261	46.4505	0.205827	EW	13.6 - 13.9 R1	0.411647	2017PZP....17....4K
V0337 UMa	152.7870	50.5892	0.191589	EW	13.70 - 14.25 R1	0.383360	2011IBVS.6008....1K
NSVS 4922773	152.6188	47.9386	0.241052	EW	14.03 - 14.5 R1	0.2741804	2013AJ....146..101P
NSVS 4921994	152.0622	47.0374	0.185512	EW	13.8 - 14.3 R1	0.371018	2004AJ....127.2436W
NSVS 4925836	154.7133	48.4649	0.164659	EW	14.15 - 14.8: R1	0.329342	2013AJ....146..101P
CSS_J102005.7+463141	155.0238	46.5282	0.317311	EW	14.61 (0.17) CV	0.316408	2014ApJS..213....9D
TSVSC1 TN-N232123120-15-67-2	153.3923	49.8148	0.250202	EW	13.72 (0.70) CV	0.2502057	2014ApJS..213....9D
CSS_J100854.4+494928	152.2268	49.8246	0.145371	EW	15.21 (0.17) CV	0.290751	2014ApJS..213....9D
CSS_J102002.9+472330	155.0122	47.3917	0.692106	RRAB	15.854 (0.63) CV	0.6921062	2013AJ....146..101P
AS CrB	240.0603	35.2090	0.1903003	EW	11.34 - 11.85 clear	0.380658	2008AJ....135..850D
USNO-A2.0 1200-07693323	238.8248	35.3051	0.522466	DSCT:	12.0 (0.03) V	0.175	2012NewA...17...65F
BQ CrB	238.0944	34.0242	1.724278	LB	13.2 - 13.9 R	-	2011IBVS.6008....1K
BP CrB	237.2507	35.2664	0.244414	RRAB	13.824 (1.05) CV	0.48004407	2017A&A...607A..11G
CSS_J154849.4+370609	237.2064	37.1025	0.444310	RRC	13.62 (0.34) CV	0.400477	2014ApJS..213....9D
CSS_J154612.8+343539	236.5537	34.5941	0.132265	EW	14.13 (0.35) CV	0.264498	2014ApJS..213....9D
T-CrBo-05430	238.0003	35.3284	0.559115	EA	14.12 (0.25) CV	1.15633	2014ApJS..213....9D
T-CrBo-07279	237.4596	34.8702	0.268033	EW	14.63 (0.19) CV	0.268024	2014ApJS..213....9D
CSS_J155029.2+360217	237.6221	36.0384	0.163318	EA	15.94 (0.29) CV	0.326596	2014ApJS..213....9D
CSS_J100616.9+531608	151.5639	53.2707	0.699863	RRAB	14.100 (0.79) CV	0.6997670	2013ApJ...763...32D
NSVS 2536063	153.8366	54.0205	0.35425	EB|EW	13.048 - 13.625 R1	0.354312	2006AJ....131..621G
LINEAR 22117001	154.0786	56.7029	0.24167	EW	14.41 (0.33) CV	0.275181	2013AJ....146..101P

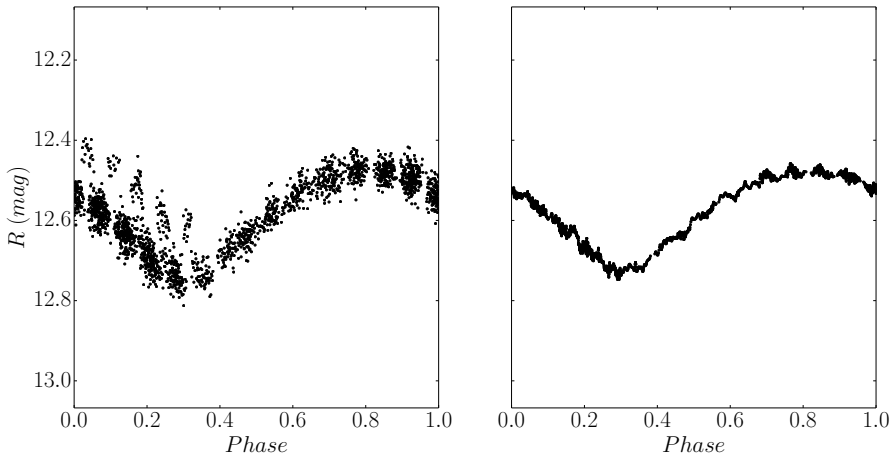


Figure 75: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable NSVS 4922327.

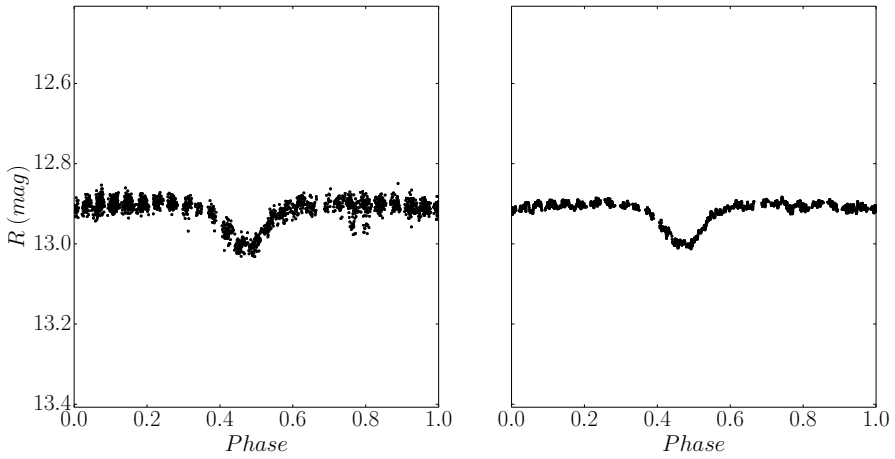


Figure 76: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable LINEAR 22079148.

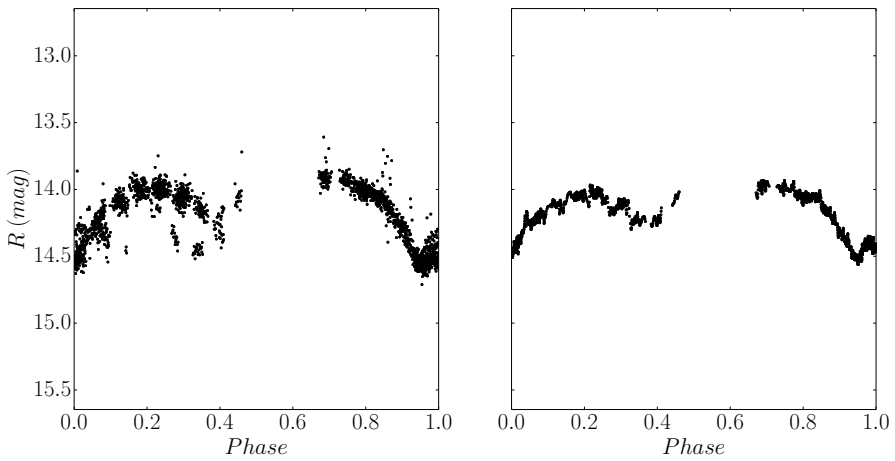


Figure 77: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable TSVSC1 TN-N232123120-15-67-2.

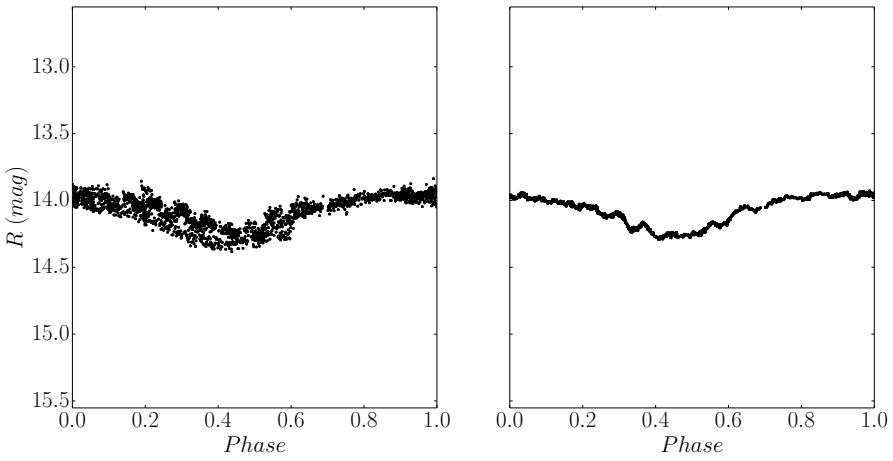


Figure 78: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable NSVS 4924202.

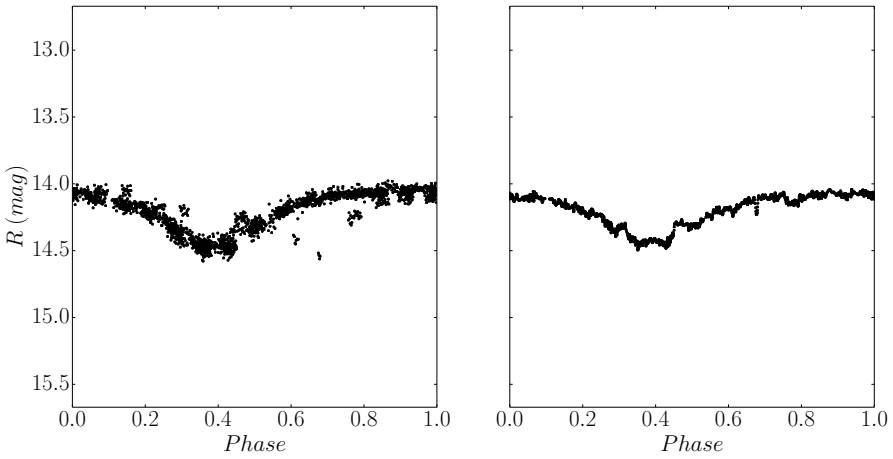


Figure 79: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable V0337 UMa.

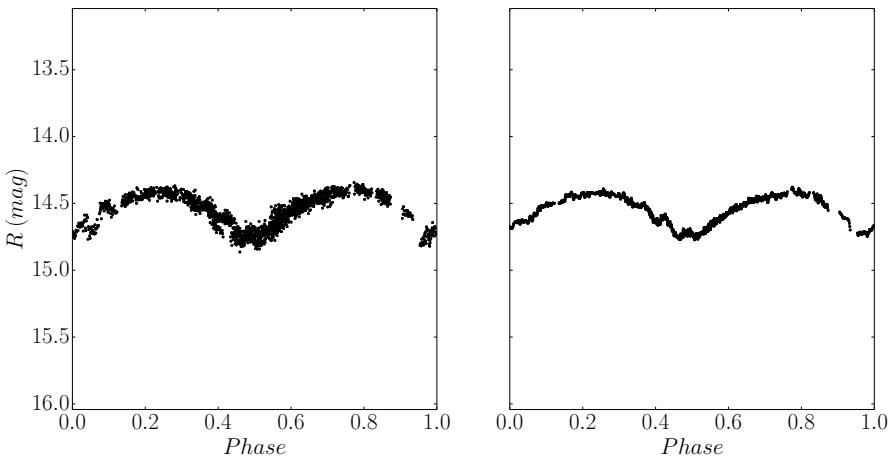


Figure 80: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable NSVS 4922773.

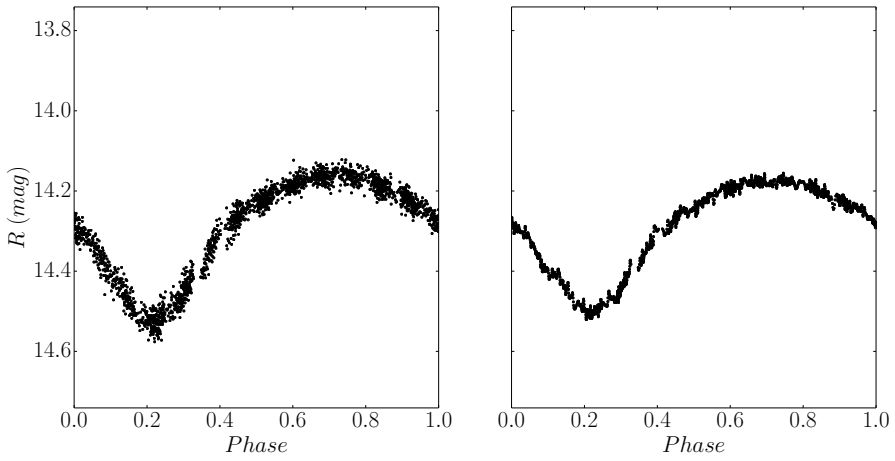


Figure 81: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable NSVS 4921994.

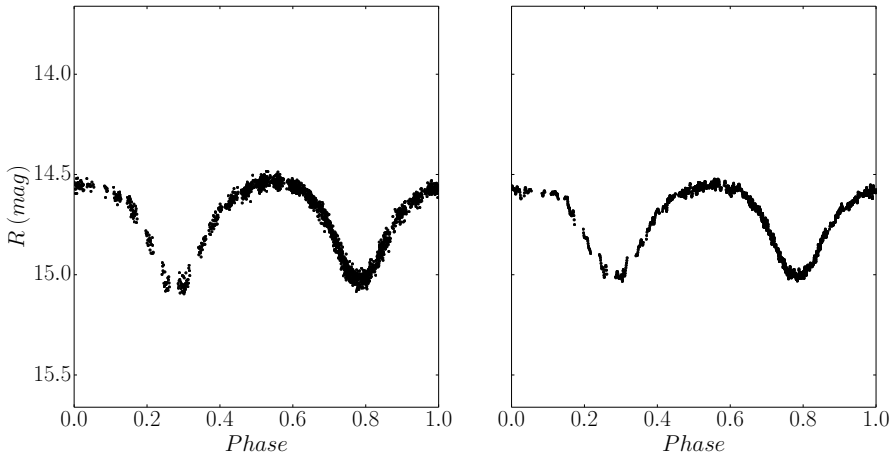


Figure 82: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable NSVS 4925836.

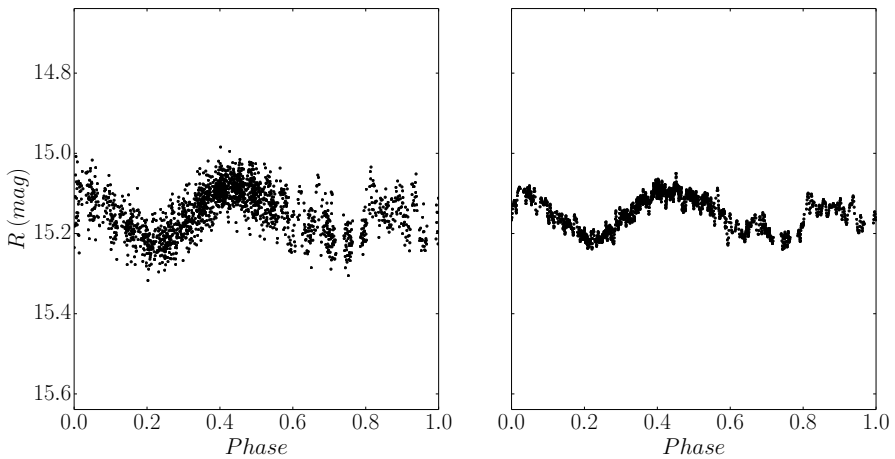


Figure 83: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable CSS_J102005.7+463141.

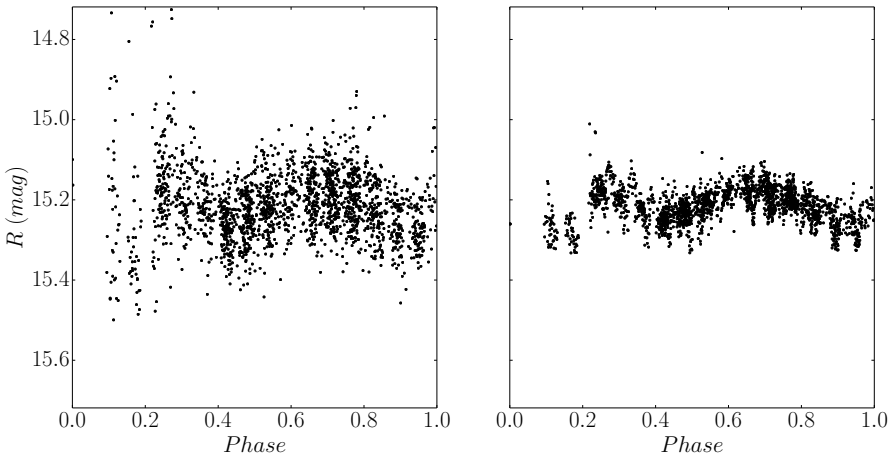


Figure 84: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable TSVSC1 TN-N232123120-15-67-2.

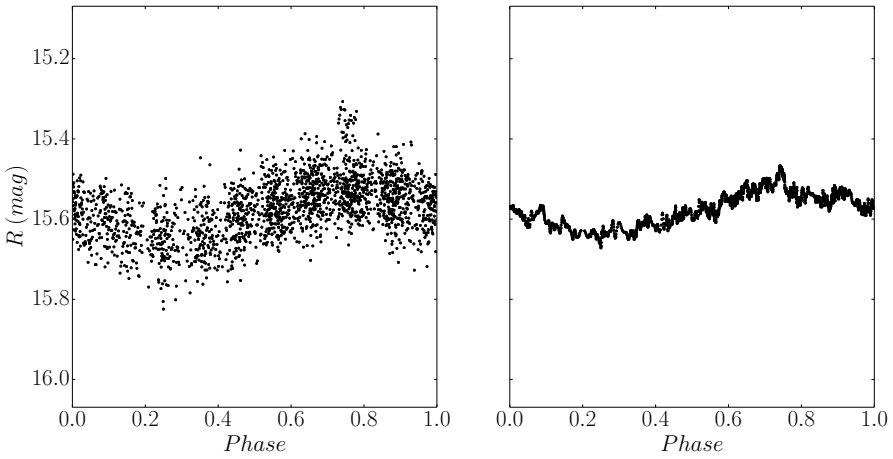


Figure 85: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable CSS_J100854.4+494928.

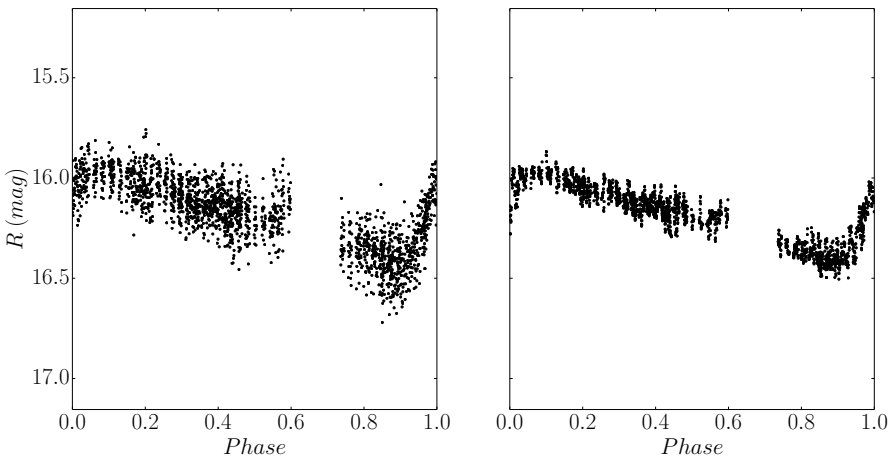


Figure 86: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable CSS_J102002.9+472330.

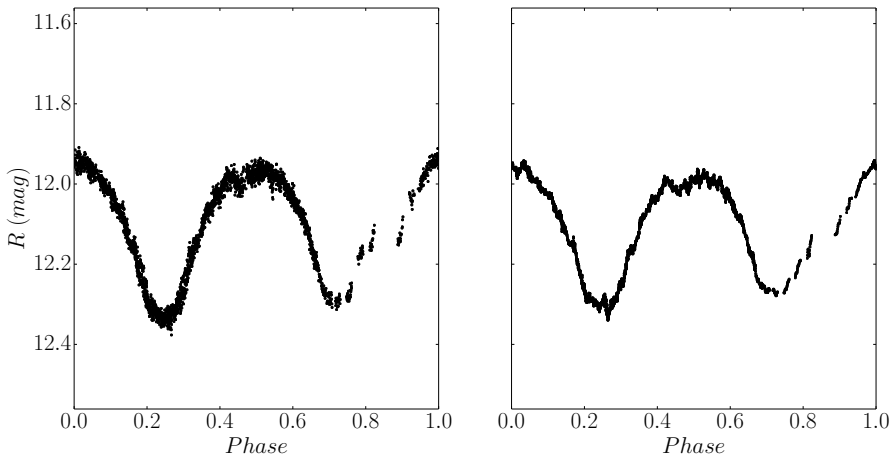


Figure 87: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable AC CrB.

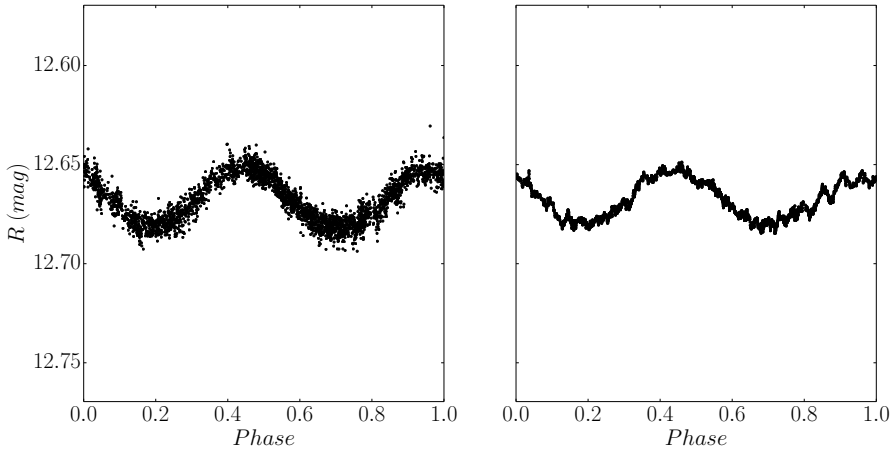


Figure 88: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable USNO-A2.0 1200-07693323.

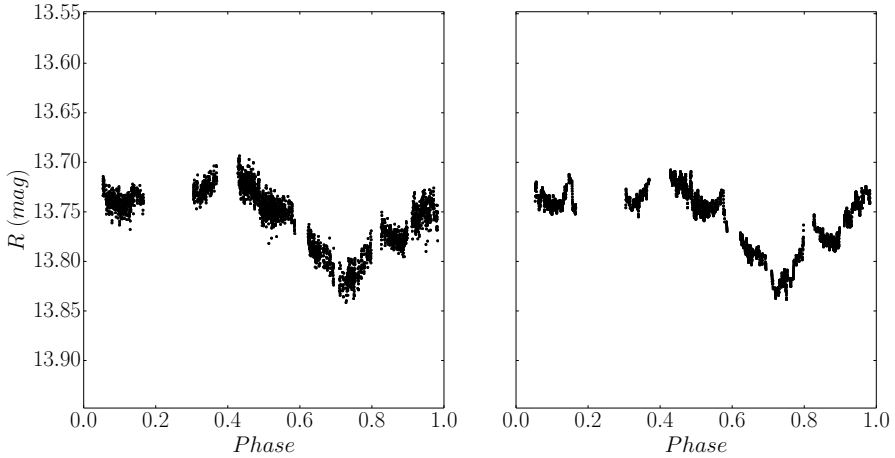


Figure 89: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable BQ CrB.

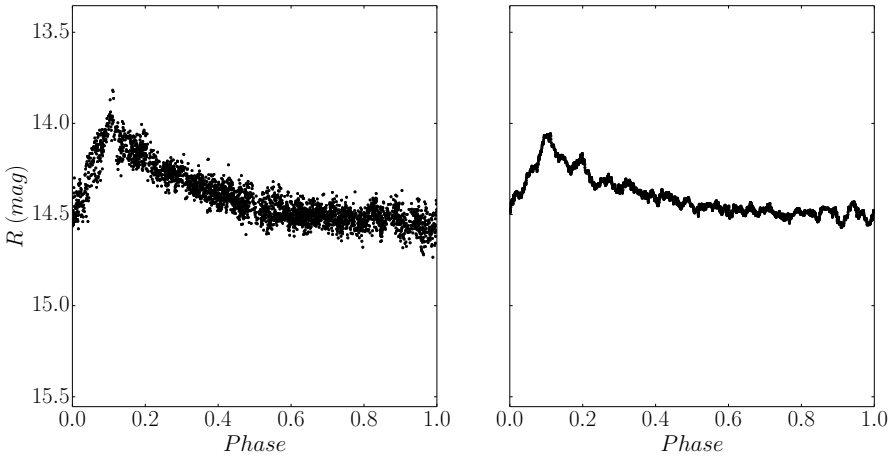


Figure 90: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable BP CrB.

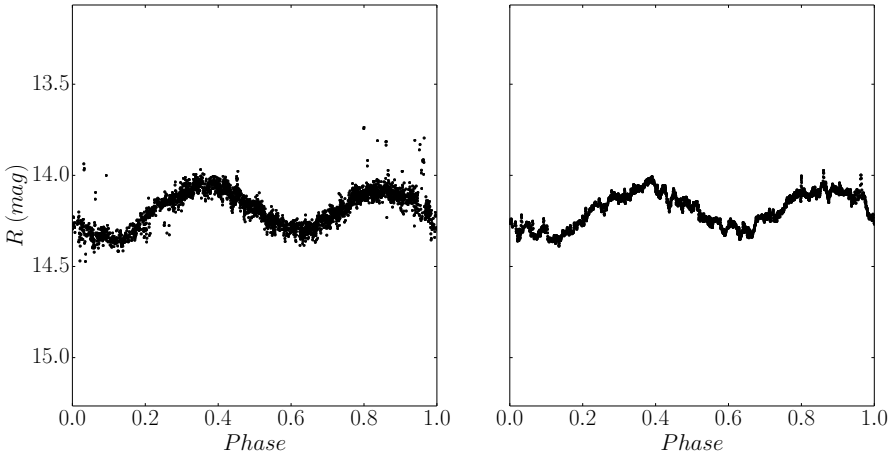


Figure 91: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable CSS_J154849+370609.

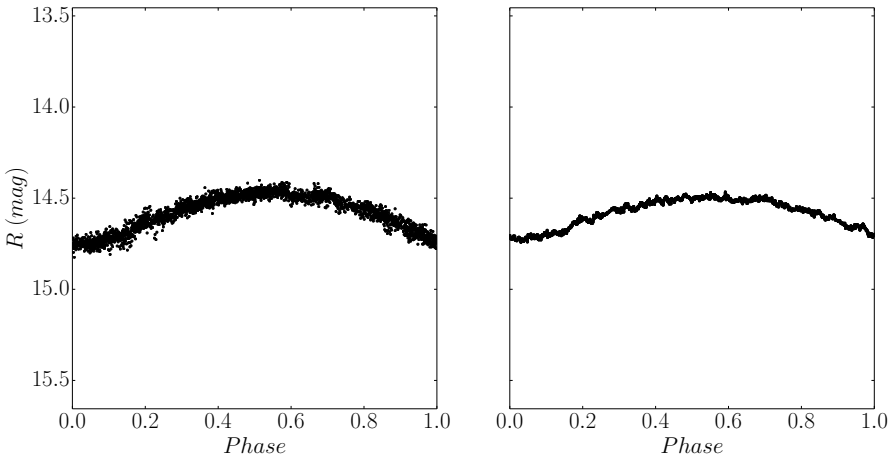


Figure 92: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable CSS_J154612.8+343539.

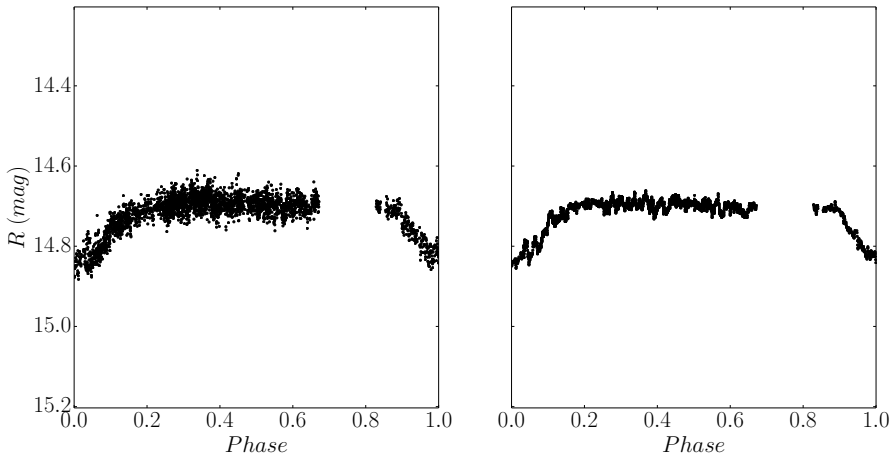


Figure 93: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable T-CrB0-05430.

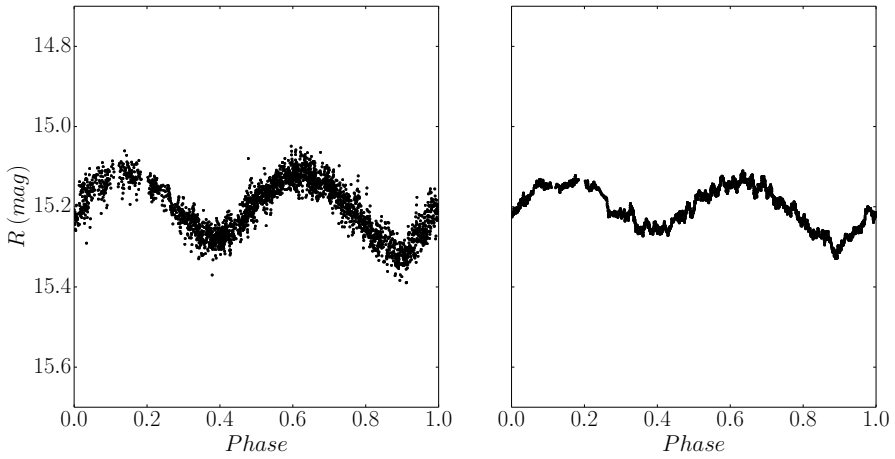


Figure 94: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable T-CrB0-07279.

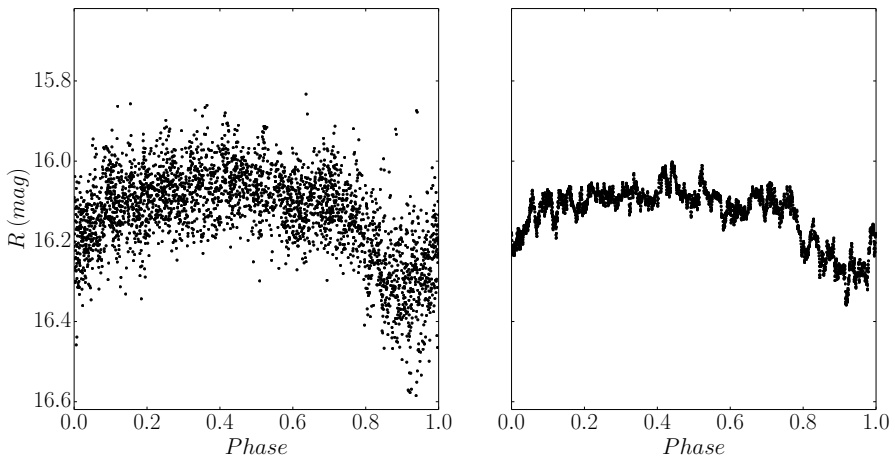


Figure 95: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable CSS_J155029.2+360217.

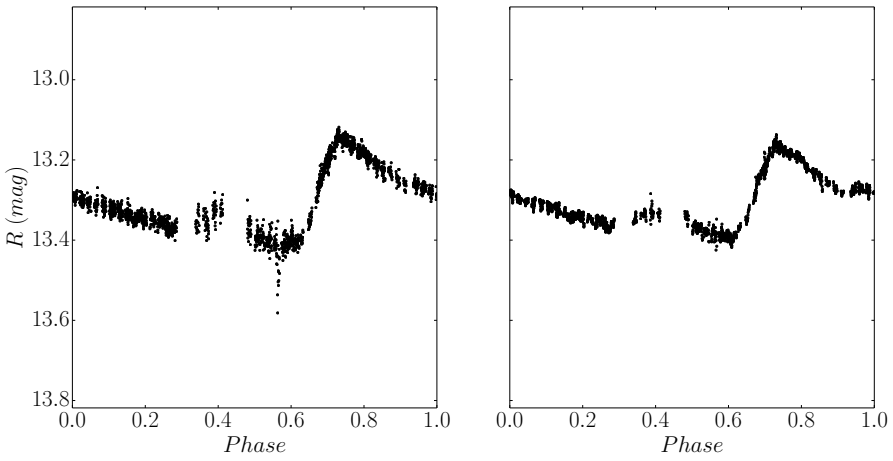


Figure 96: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable CSS_J100616.9+531608.

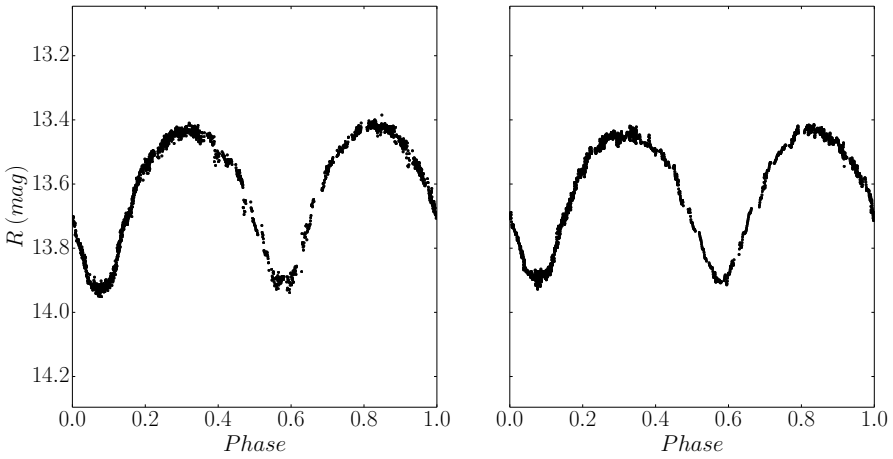


Figure 97: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable NSVS 2536063.

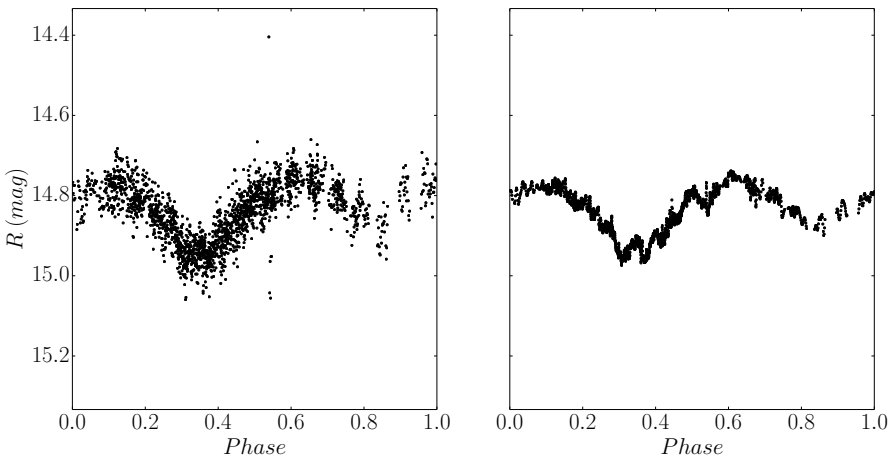


Figure 98: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable LINEAR 22117001.

A.2 THE EVRYSCOPE VARIABLES

Here we present the variables detected in The Evryscope data sample for five, ~3° × 3° regions (i.e ~0.6% of the total Evryscope's FoV) of the camera pointed towards the South Celestial Pole. As the examples shown in Sect. 6.1.2, each light curve comprises 10,240 data points observed from January 10, 2016 to February 26, 2016. After visual inspection of those light curves with a SDE>15, 43 bona-fide variables were detected.

Of those, 24 (55.81%) were previously cataloged variables of different variability types. They are presented in Table 17 along with more information regarding their positions in the sky, their periods, their cataloged magnitude variations and variability types and their bibliographic references. Their TFA and TFAW light curves are shown in Fig. 119 to 142. Light curves without TFAW period are variables with periods longer than the analyzed time-base.

Of the remaining 19 (44.19%) non-cataloged stars, 6 present transit-like features, 1 is most probably an eclipsing binary and the others are either sinusoidal-like variables or long period variables. They are presented in Table 16 with information about their TFAW periods, their coordinates and some notes regarding their possible variability origin. Their TFA and TFAW light curves are shown in Fig. 100 to 118.

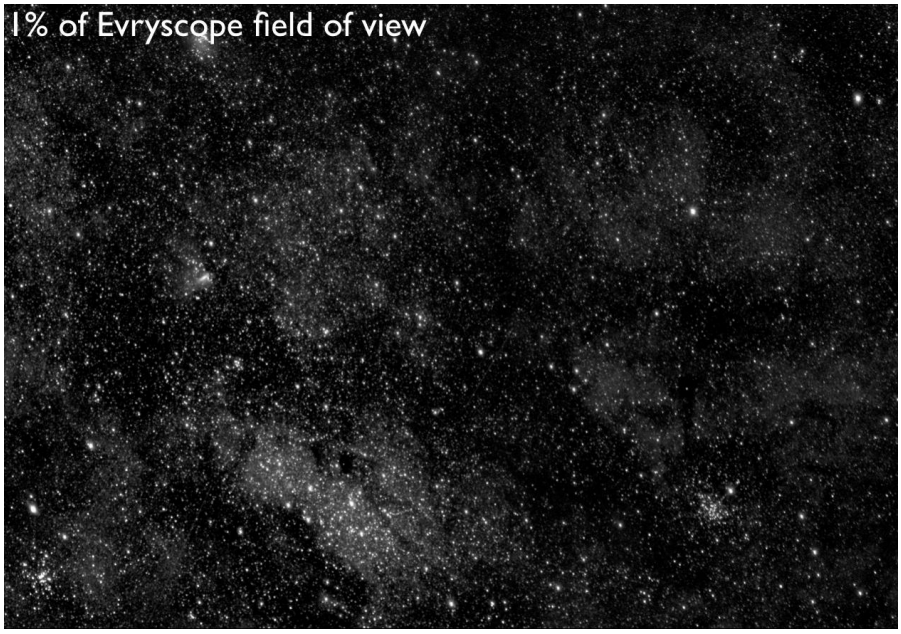


Figure 99: 1% of The Evryscope's FoV.

Table 16: The Evryscope non-cataloged variables and transiting candidates.

Gaia DR2 ID	Evryscope ID	α (J2000.0)	δ (J2000.0)	P_{TFAW} (d)	Notes
5770239054090955904	18974388	181.262	-85.4629	4.324238	Transit like feature
4613765126197369216	18964795	39.0111	-85.409	2.473688	Transit like feature
5767245874203061120	18978076	242.464	-85.3169	1.8647858	Transit like feature
5767144989713190912	19021806	251.104	-86.1164	2.650729	Eclipsing Binary candidate
5766463292503813376	19018247	230.618	-85.8456	2.31345	Eclipsing Binary candidate
5769603978748830976	19010070	193.775	-85.2212	1.27347	Transit like feature
5770320555390624768	63578985	186.254	-84.7953	21.161955	Long period variable candidate
5770505204625238144	63729238	198.275	-83.7165	0.03151	Sinusoidal like feature
4613744823887608704	18964242	31.6983	-85.7202	0.177328	Sinusoidal like feature
4613771383965141888	18964916	40.2699	-85.3484	15.093289	Long period variable candidate
5767220585435415168	18978271	246.468	-85.7283	0.423683	Sinusoidal like feature
5767113520490189056	18978962	259.925	-85.3383	0.1203477	Sinusoidal like feature. Blended with other sources.
5767118193414597632	18979330	261.432	-85.2144	15.324587	Sinusoidal long period variable
5767049336498764160	18979819	267.270	-85.6117	0.1792033	Sinusoidal like feature
5767045556927517952	18979915	268.413	-85.6935	0.1861069	Sinusoidal like feature
5767421692981997696	63583325	247.385	-84.2174	17.4658	Long period variable candidate
5769505503738599680	18975059	190.175	-85.3402	2.131039	Sinusoidal like feature
5769508944011260288	18975147	191.54	-85.2888	0.306067	Sinusoidal like feature
5769303507135132544	18975965	209.155	-85.4846	0.23072	Sinusoidal like feature

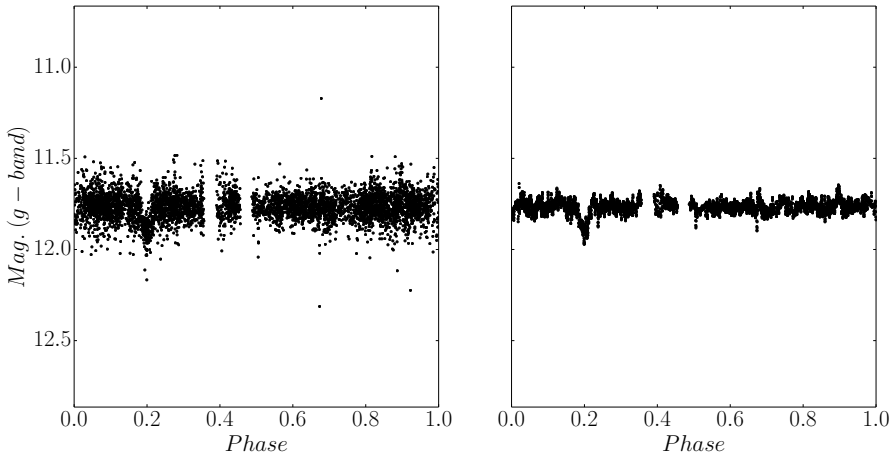


Figure 100: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 18974388.

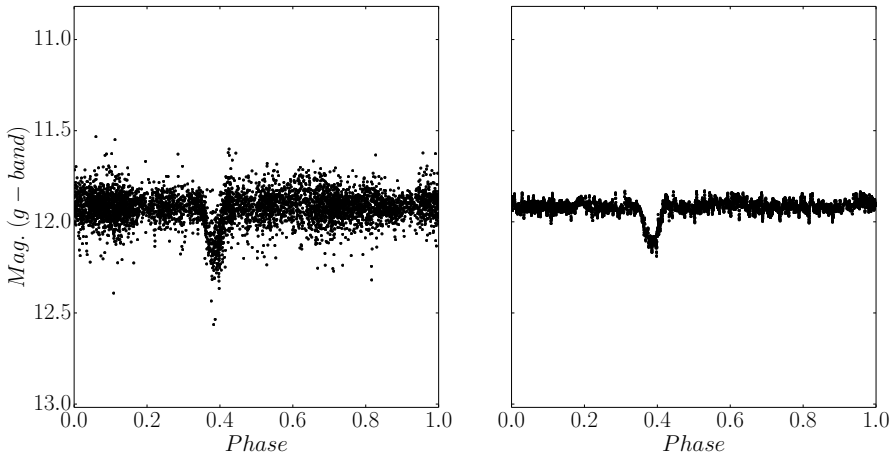


Figure 101: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 18964795.

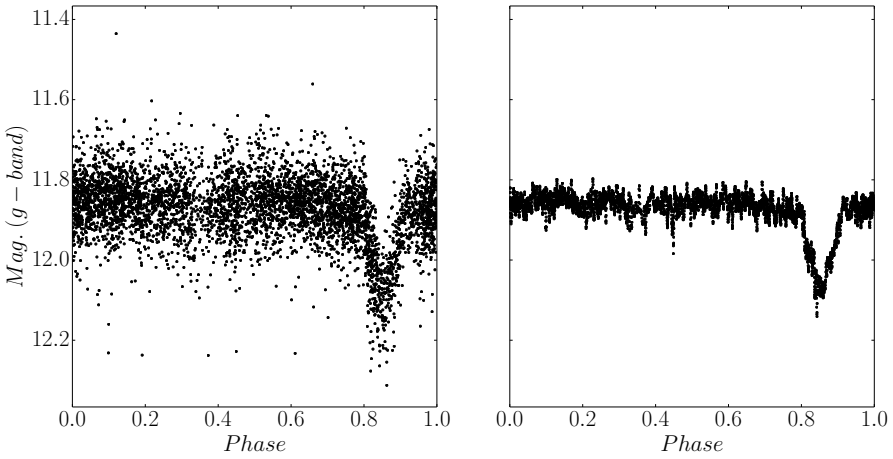


Figure 102: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 18978076.

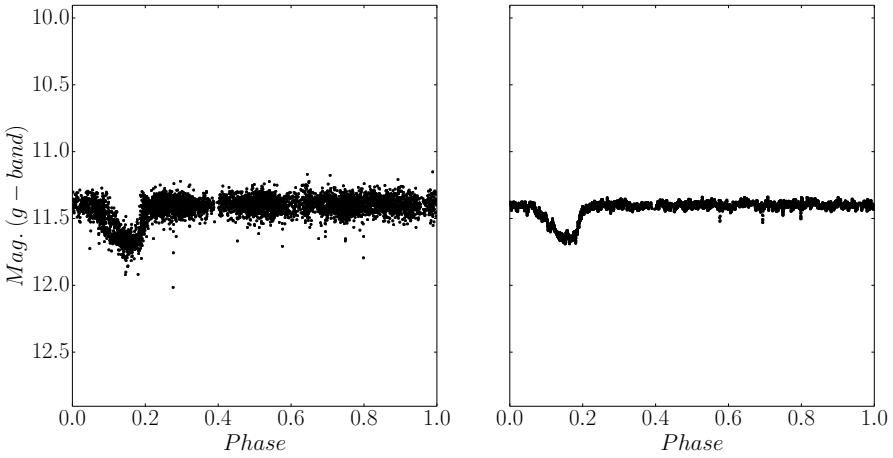


Figure 103: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 19021806.

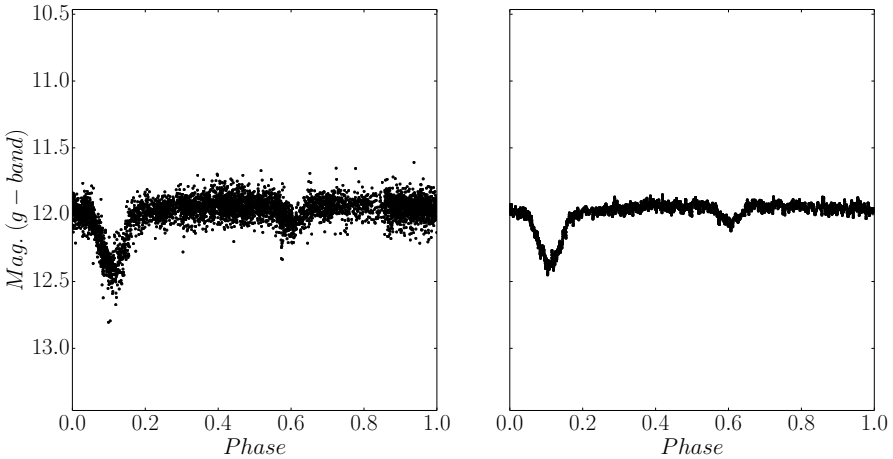


Figure 104: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 19018247.

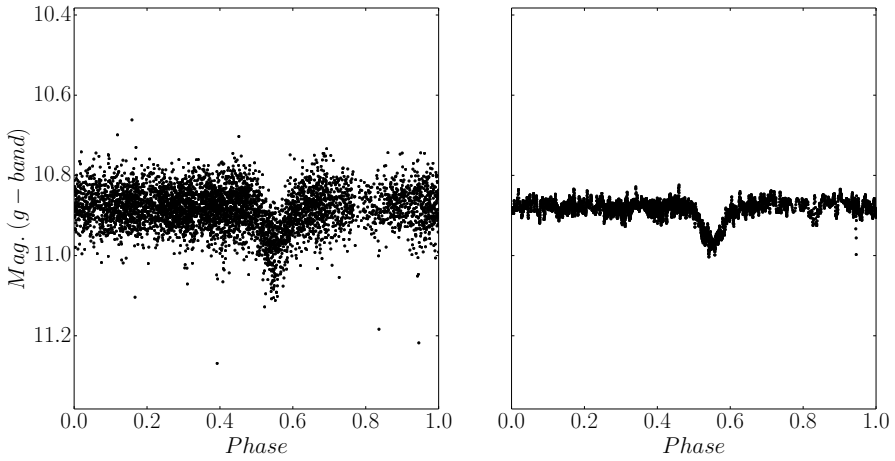


Figure 105: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 19010070.

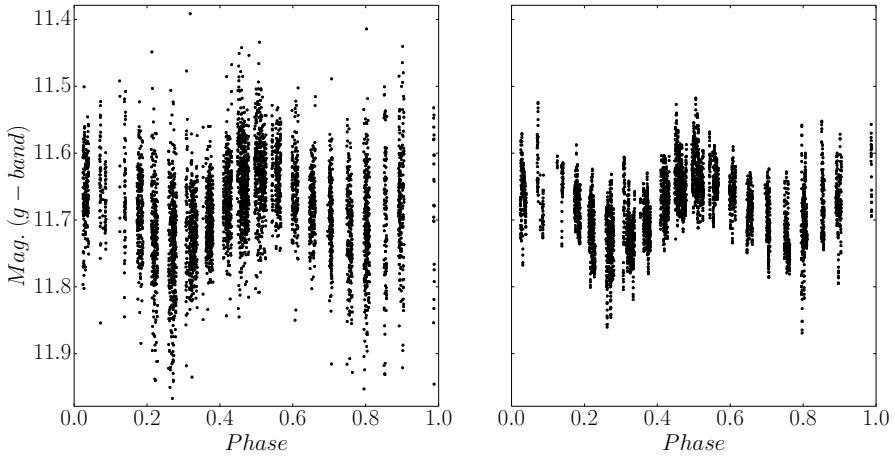


Figure 106: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 63578985.

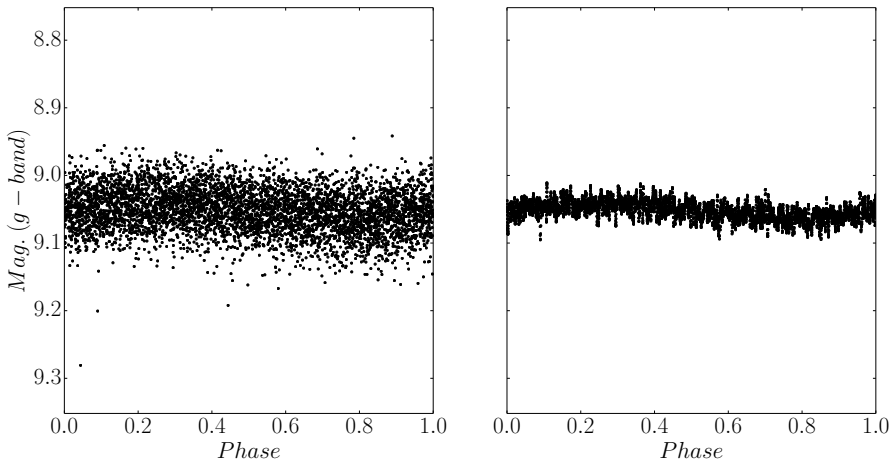


Figure 107: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 63729238.

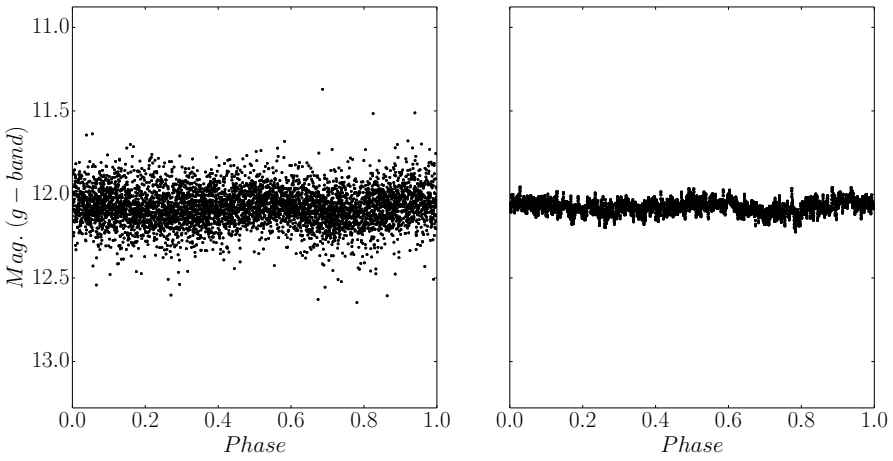


Figure 108: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 18964242.

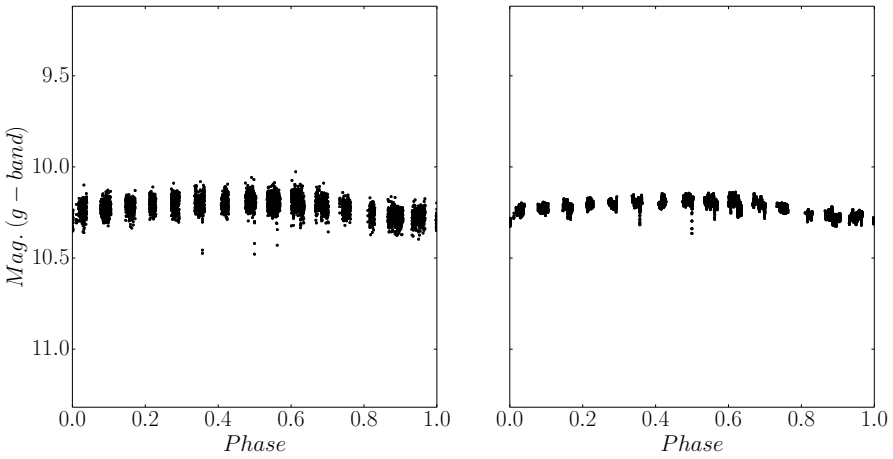


Figure 109: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 18964916.

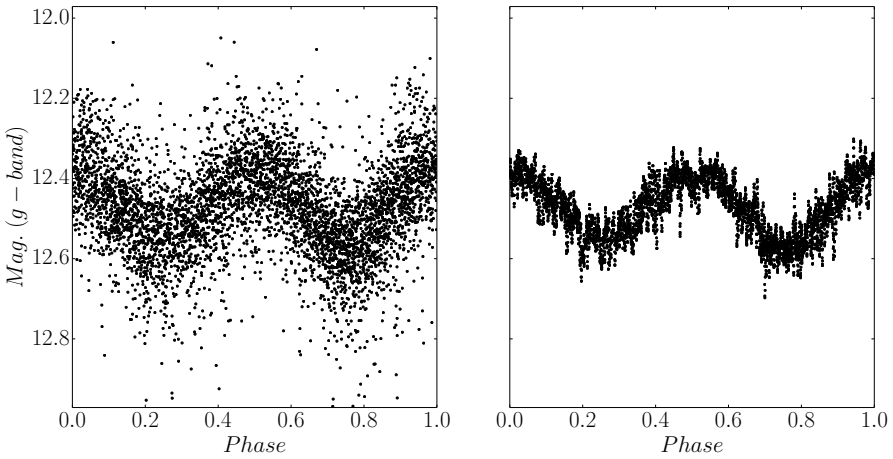


Figure 110: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 18978271.

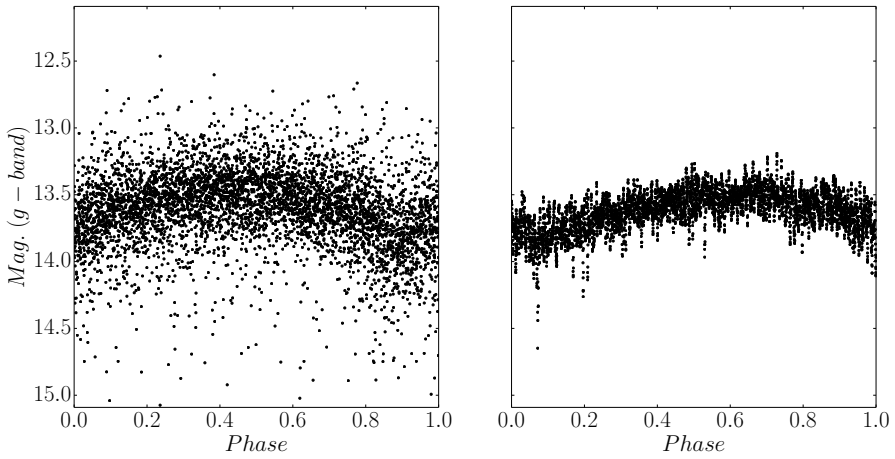


Figure 111: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 18978962.

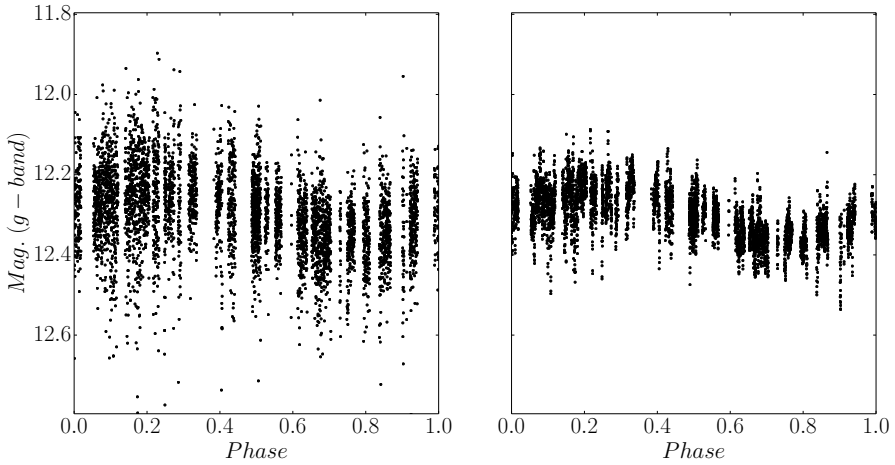


Figure 112: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 18979330.

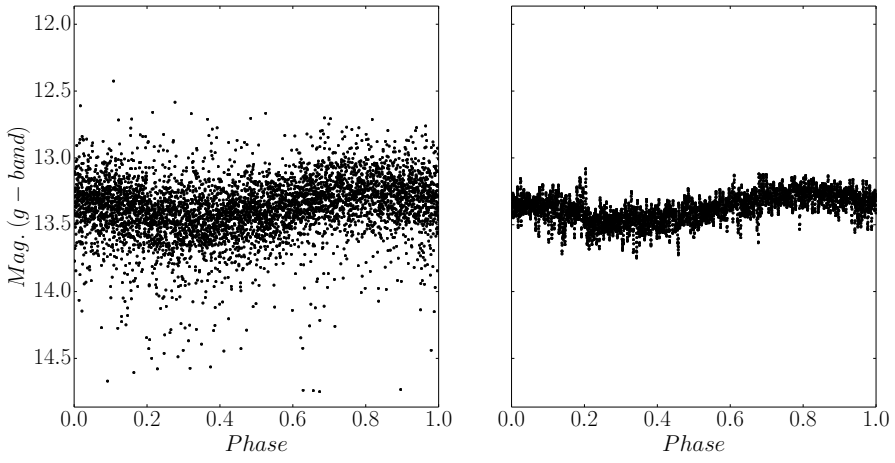


Figure 113: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 18979819.

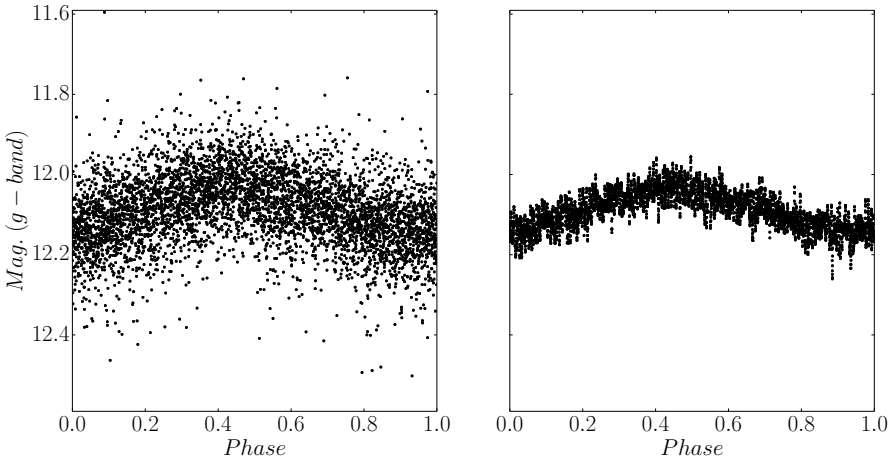


Figure 114: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 18979915.

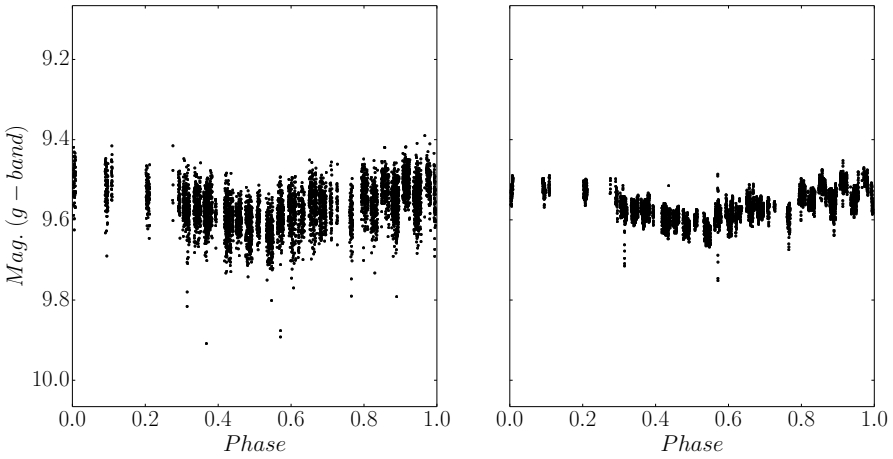


Figure 115: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 63583325.

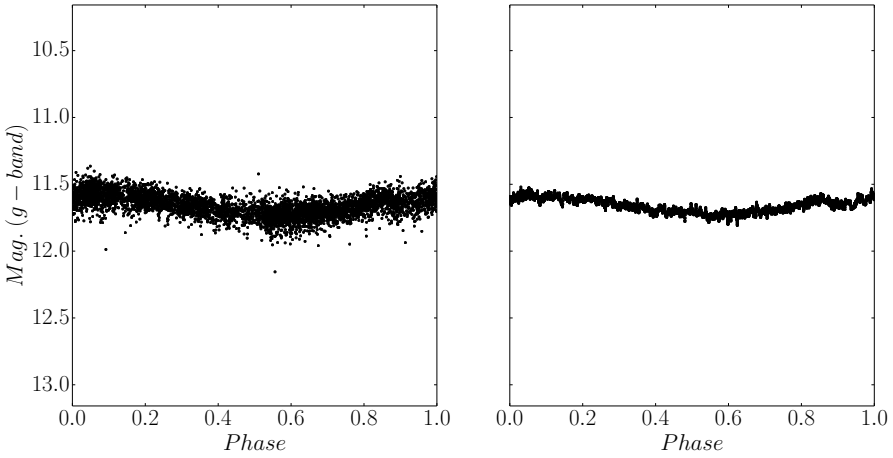


Figure 116: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 18975059.

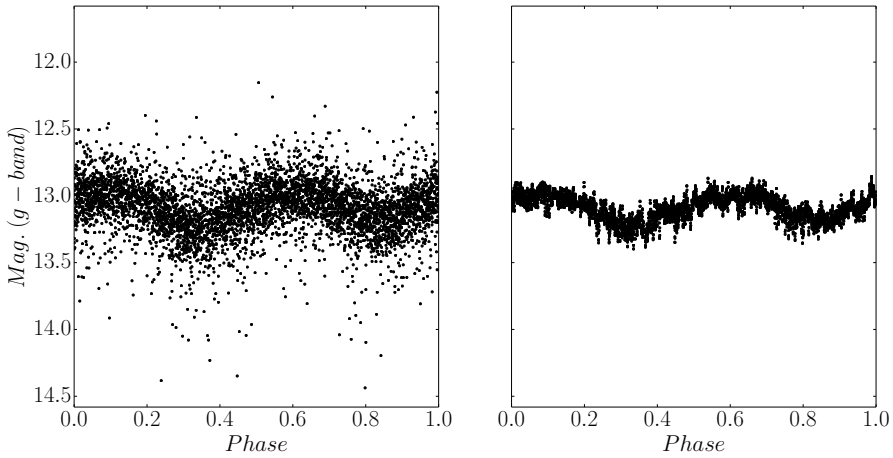


Figure 117: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 18975147.

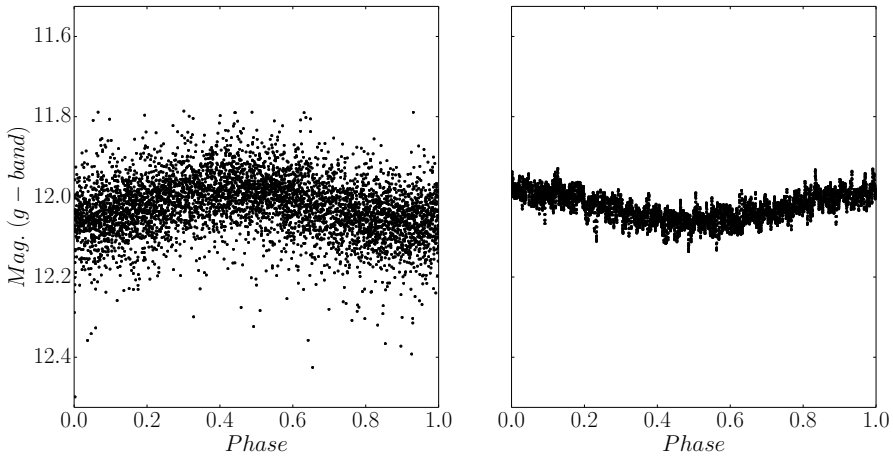


Figure 118: Phase folded TFA (left) and TFAW (right) light curves for non-cataloged variable Evryscope ID 18975965.

Table 17: The Evryscope cataloged variables

Cataloged ID	α (J2000.0)	δ (J2000.0)	P_{TFAW} (d)	Cataloged var. type	Cataloged mag. var.	Cataloged P (d)	Cataloged ref.
NSV 6086	13 08 08.57	-83 47 59.8	0.554553	RRAB/BL	13.45 - 14.20 V	0.554855	2017PASP..129j4502K
U Oct	13 24 32.62	-84 13 30.9	-	M	7.0 - 14.1 V	304	2002AcA....52..397P
ASAS J132837-8405.6	13 28 37.99	-84 05 38.0	43.002	MISC	12.52 (0.55) V	52.720501	2002AcA....52..397P
TX Oct	03 05 43.34	-85 44 59.4	0.9072314	EA/SD	12.6 - 13.9 p	0.907137	2002AcA....52..397P
ASAS J033718-8516.7	03 37 19.19	-85 16 44.5	47.04861	MISC	10.08 (0.23) V	51.9	2002AcA....52..397P
ASAS J164514-8545.8	16 45 12.67	-85 45 44.4	0.183343	EC|ESD	13.41 (0.65) V	0.366716	2002AcA....52..397P
ASAS J171002-8524.5	17 10 03.38	-85 24 32.1	24.84187	MISC	10.57 (0.13) V	82.827164	2002AcA....52..397P
ASAS J162726-8602.7	16 27 26.44	-86 02 37.2	-	MISC	11.21 (0.64) V	63.9	2002AcA....52..397P
ASASSN-V J164326.80-861144.8	16 43 26.80	-86 11 44.8	0.202501	EA|EB	14.02 - 14.61 V	0.404988	2018MNRAS.477.3145J
EQ Oct	17 10 29.01	-86 23 00.1	-	SRA	9.9 - 12.3 V	267	2008IBVS.5863....1K
FO Oct	17 19 28.83	-86 38 26.9	-	SRB	11.5 - 14.4 V	86.2	2015IBVS.6151....1K
ASASSN-V J173520.26-865008.5	17 35 20.26	-86 50 08.5	-	L	12.61 - 13.12 V	-	2018MNRAS.477.3145J
Z Oct	17 44 36.63	-86 29 10.6	-	M	10.15 - 13.7 V	331	2002AcA....52..397P
ASAS J174554-8654.5	17 45 52.48	-86 54 28.0	0.76672	ESD	12.9 (0.46) V	0.766716	2002AcA....52..397P
ASASSN-V J163423.01-841701.7	16 34 23.01	-84 17 01.7	0.39678	EW	13.10 - 13.58 V	0.396747	2018MNRAS.477.3145J
ASASSN-V J165050.25-843635.3	16 50 50.27	-84 36 35.4	1.07729	EA	13.74 - 14.13 V	2.15545	2018MNRAS.477.3145J
ASAS J165834-8458.6	16 58 36.18	-84 58 35.0	0.54937	RRAB	13.051 - 13.824 V	0.5493159	2009arXiv0906.2199S
ASAS J130015-8520.2	13 00 12.96	-85 20 13.7	-	MISC	10.65 (0.39) V	53.8	2002AcA....52..397P
ASAS J130710-8509.0	13 07 10.43	-85 08 58.1	-	MISC	9.79 (0.07) V	20.761246	2002AcA....52..397P
ASAS J130807-8503.5	13 08 03.40	-85 03 28.6	1.29596	RS	11.85 - 12.15 V	0.51448	2002AcA....52..397P
ASAS J121332-8646.7	12 13 32.70	-86 46 43.1	-	MISC	12.68 (0.5) V	283.01886	2002AcA....52..397P
ASASSN-V J121952.08-861641.8	12 19 52.08	-86 16 41.8	-	L	12.71 - 13.13 V	-	2018MNRAS.477.3145J
ASAS J131308-8528.6	13 13 09.52	-85 28 33.8	0.276453	EC|RRC	11.51 (0.27) V	0.552796	2002AcA....52..397P
ASASSN-V J134322.05-845650.8	13 43 22.05	-84 56 50.8	0.405308	EA|EB	12.74 - 13.17 V	0.405349	2018MNRAS.477.3145J

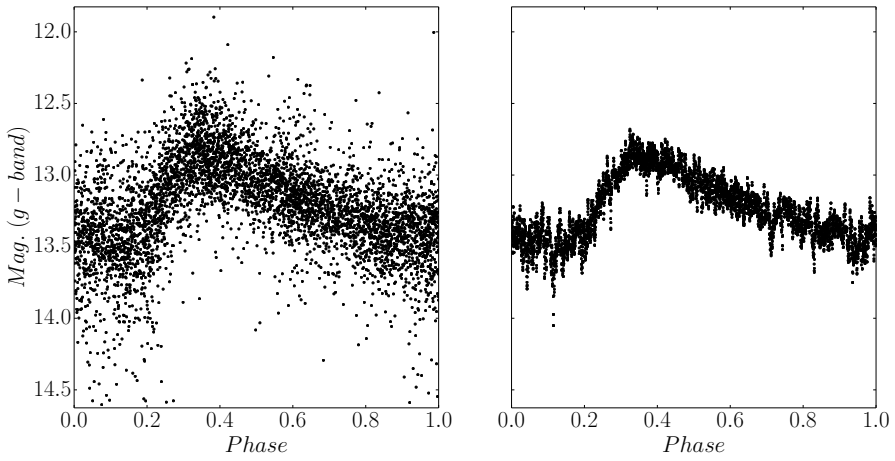


Figure 119: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable NSV6086.

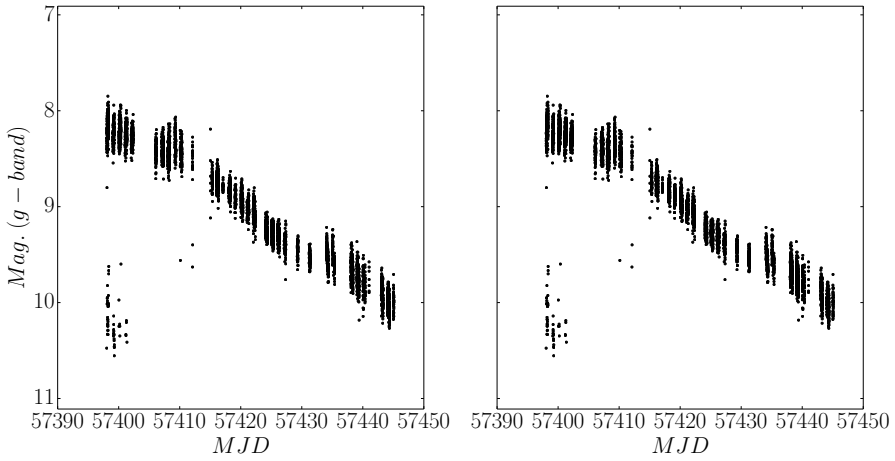


Figure 120: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable U Oct.

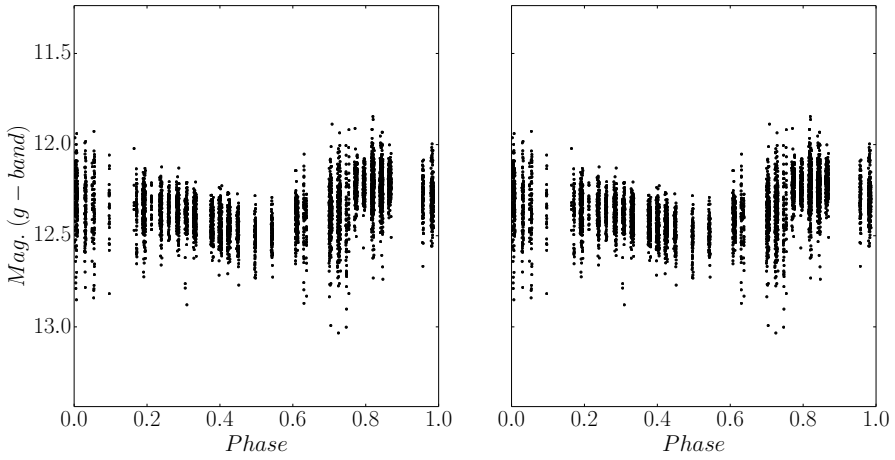


Figure 121: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASAS J132837-8405.6.

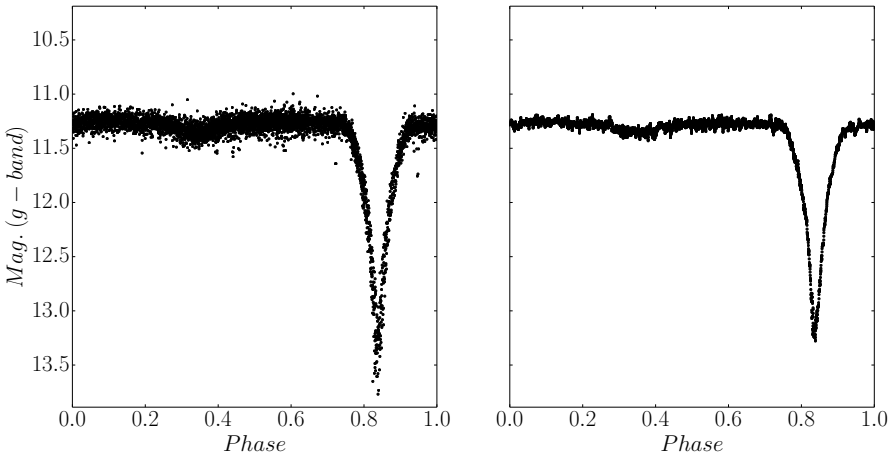


Figure 122: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable TX Oct.

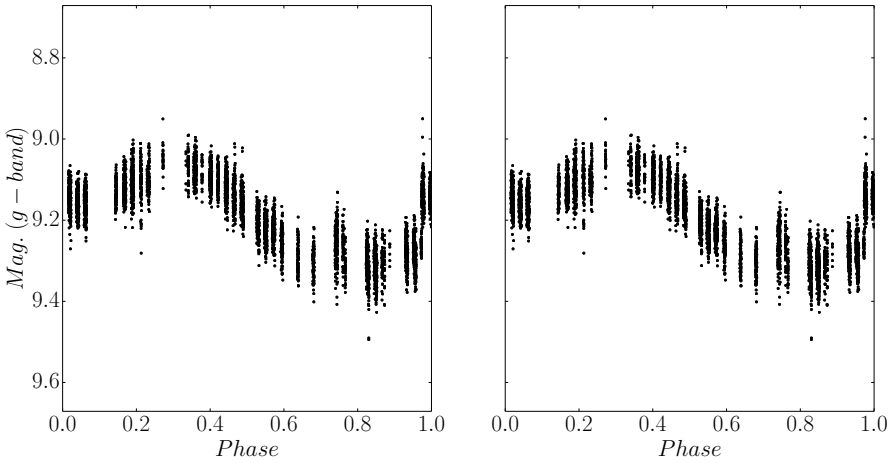


Figure 123: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASAS J033718-8516.7.

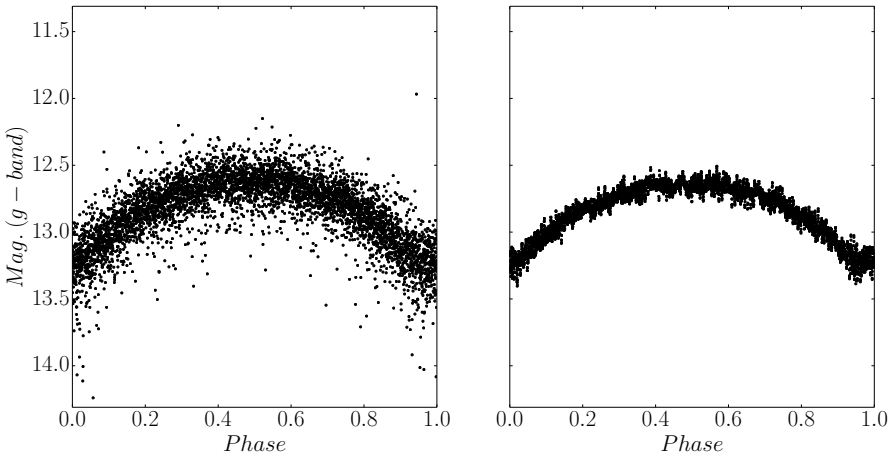


Figure 124: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASAS J164514-8545.8.

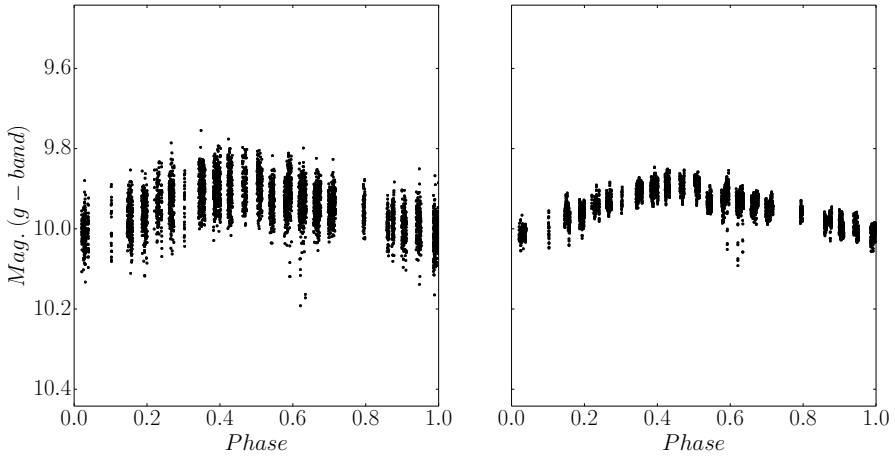


Figure 125: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASAS J171002-8524.5.

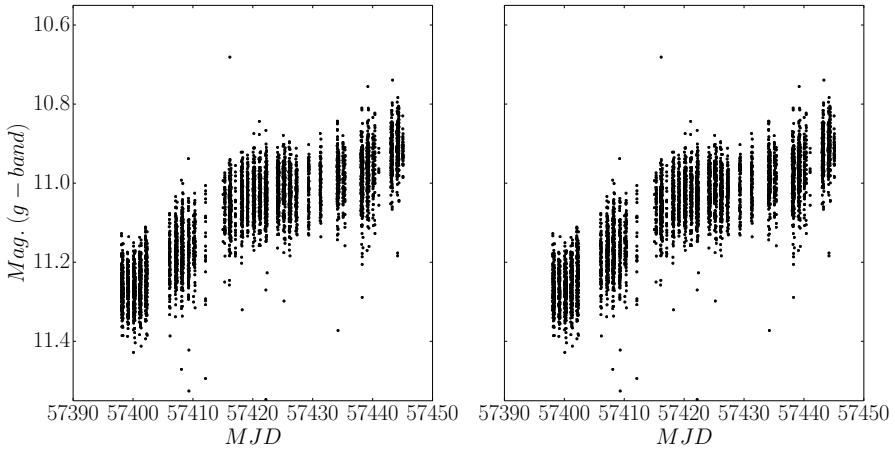


Figure 126: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASAS J162726-8602.7.

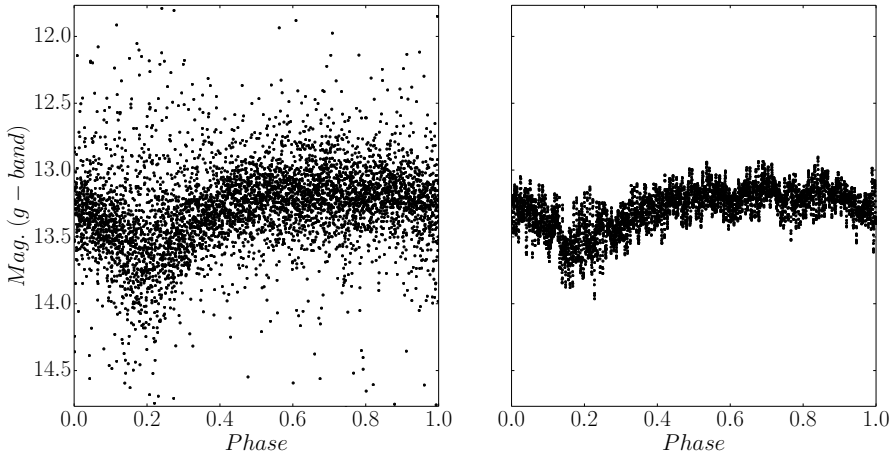


Figure 127: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASASSN-V J164326.8-861144.8.

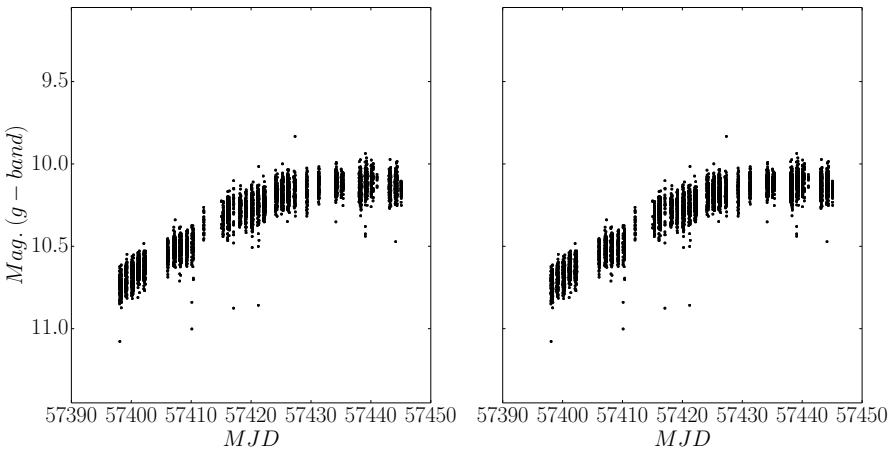


Figure 128: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable EQ Oct.

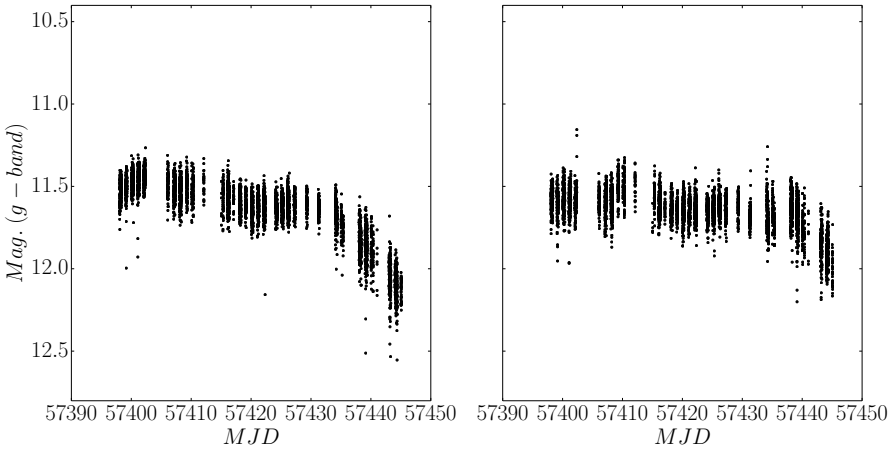


Figure 129: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable FO Oct.

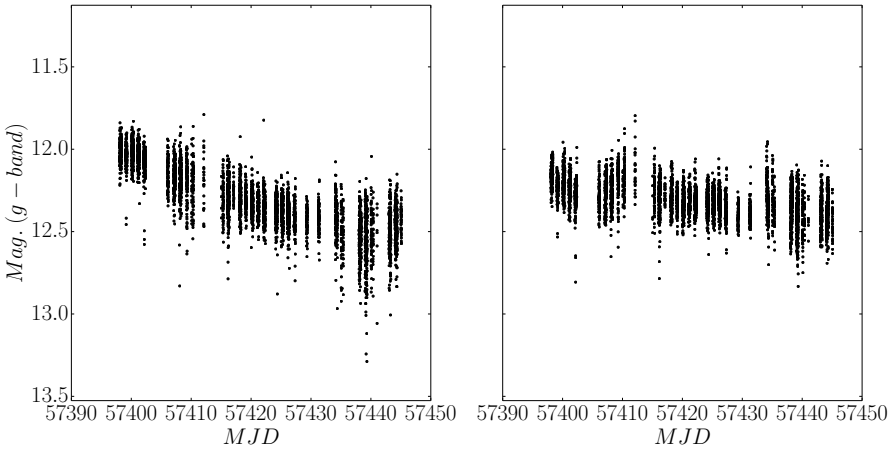


Figure 130: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASASSN-V J173520.26-865008.5.

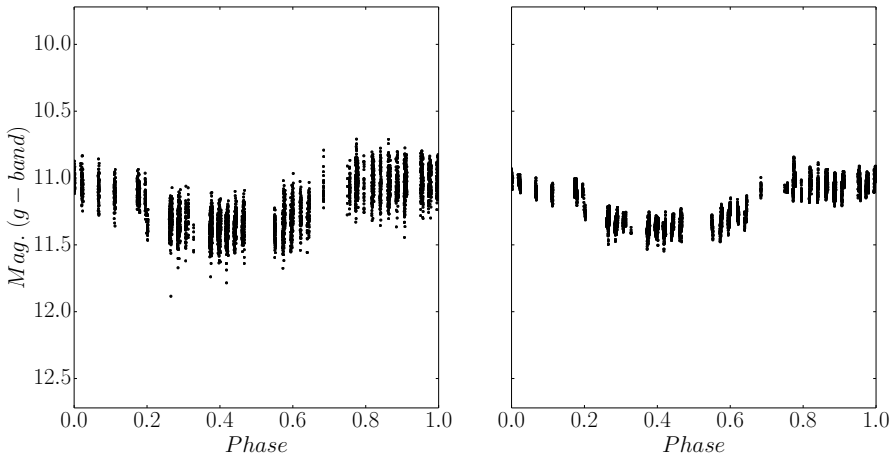


Figure 131: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable Z Oct.

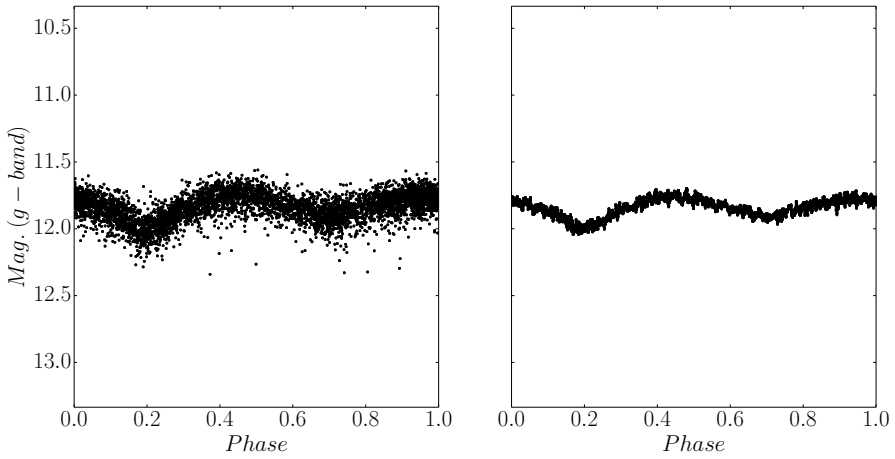


Figure 132: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASAS J174554-8654.5.

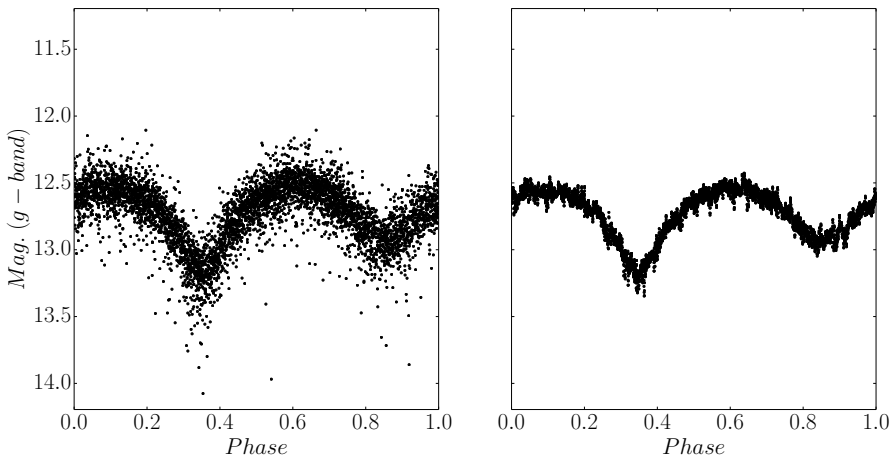


Figure 133: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASASSN-V J163423.01-841701.7.

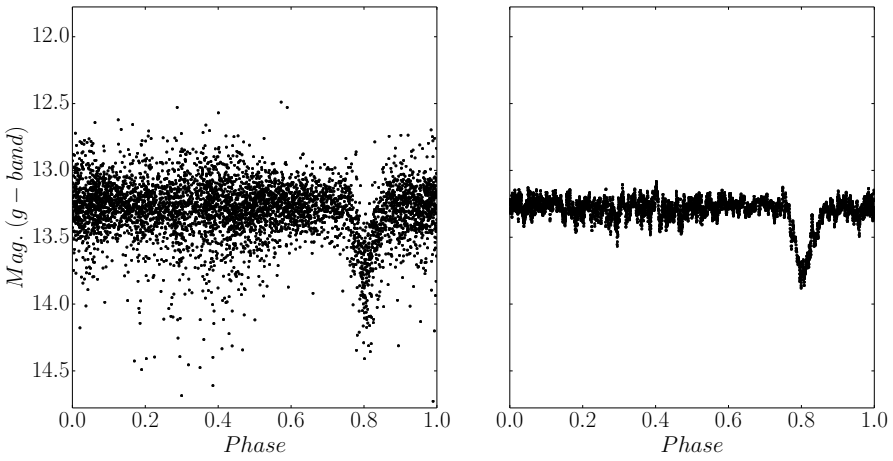


Figure 134: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASASSN-V J165050.25-843635.3.

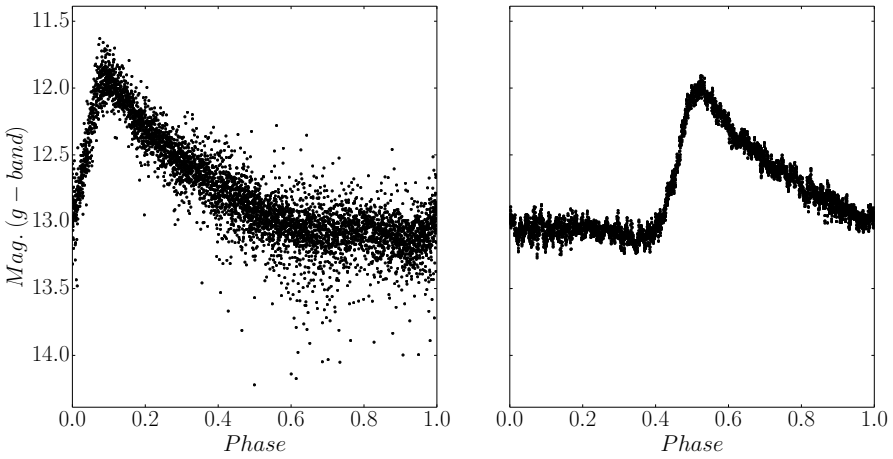


Figure 135: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASAS J165834-8458.6.

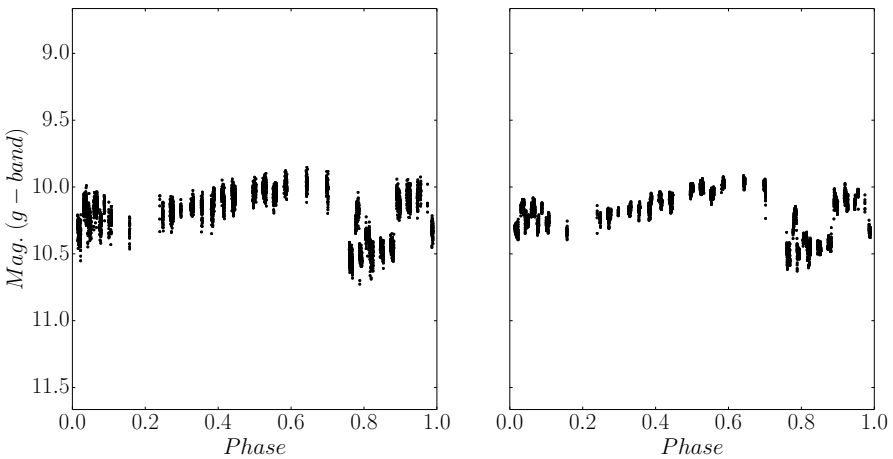


Figure 136: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASAS J130015-8520.2.

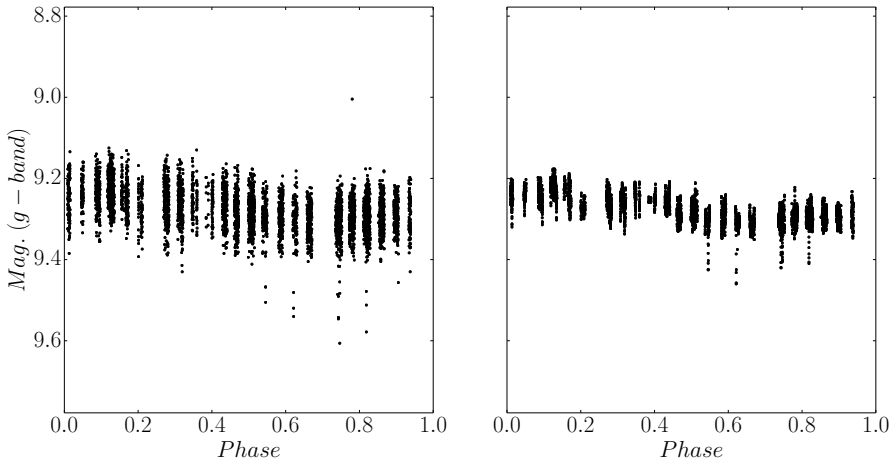


Figure 137: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASAS J130710-8509.0.

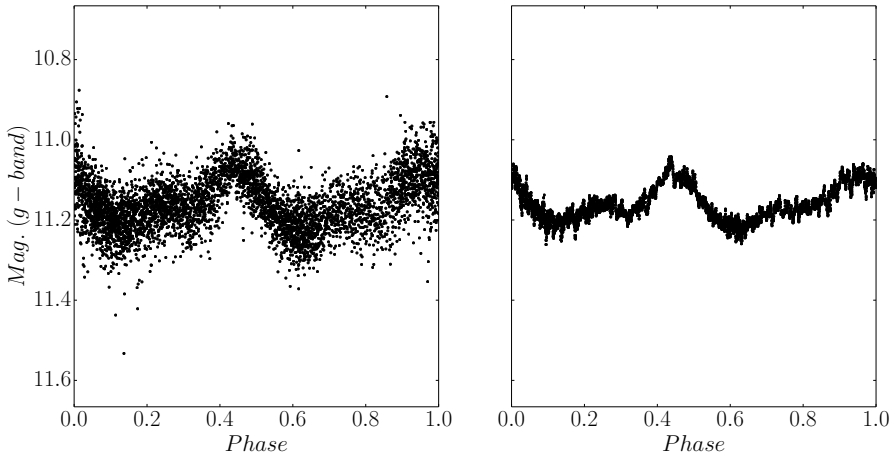


Figure 138: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASAS J130807-8503.5.

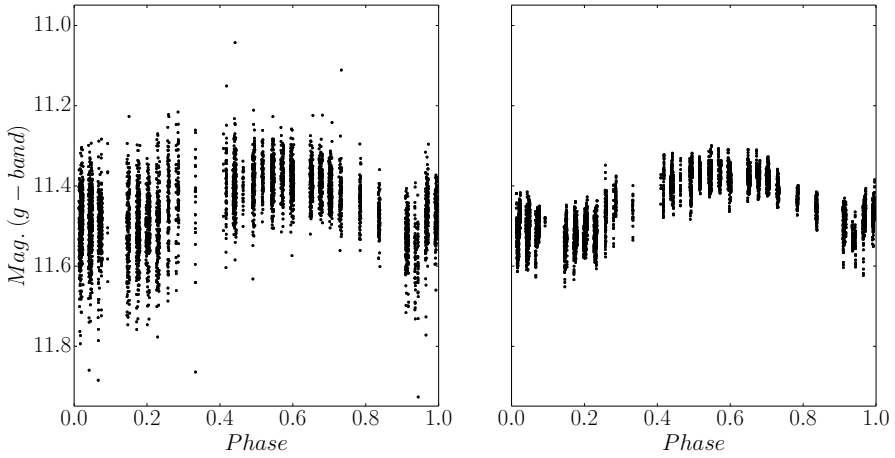


Figure 139: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASAS J121332-8646.7.

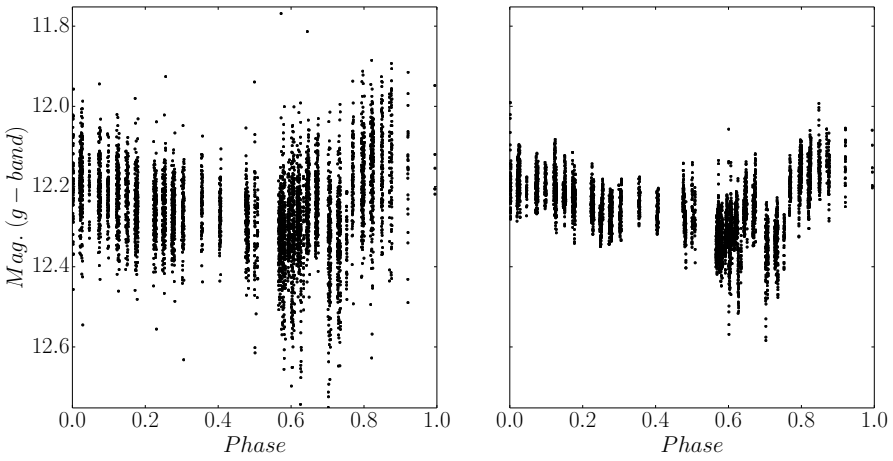


Figure 140: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASASSN-V J121952.08-861641.8.

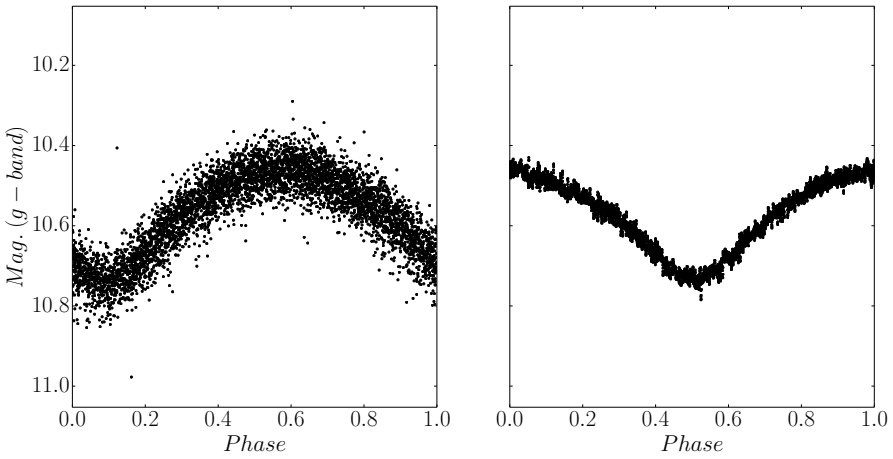


Figure 141: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASAS J121332-8646.7.

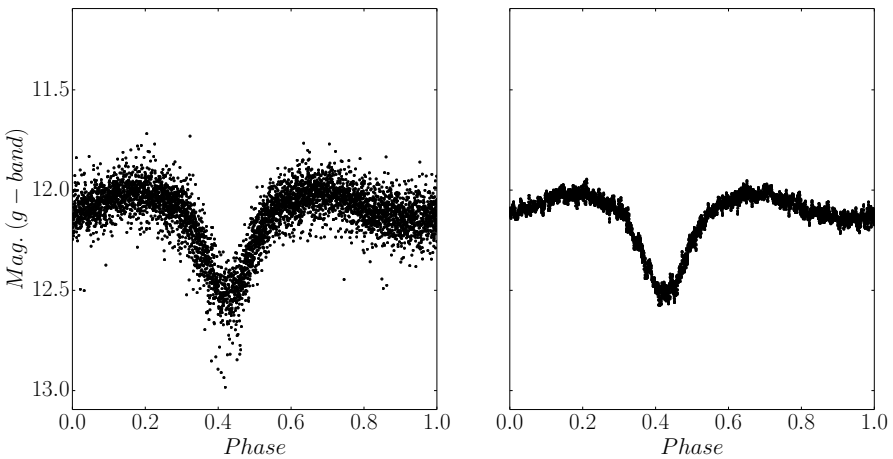


Figure 142: Phase folded TFA (left) and TFAW (right) light curves for cataloged variable ASASSN-V J134322.05-845650.8.

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

ADU Analogue-to-Digital Units

BEB Background Eclipsing Binary

BLS Box Least Square

BNC Baker-Nunn Camera

CCD Charged Couple Device

CVZ CoRoT continuous viewing zone

CNES Centre National d'études Spatiales

CWT Continuous Wavelet Transform

CoRoT COnvection ROtation and planetary Transits

CTIO Cerro Tololo Inter-American Observatory

DFT Discrete Fourier Transform

DQD Data Quality Daemon

DWT Discrete Wavelet Transform

ESA European Space Agency

FFT Fast Fourier Transform

FoV Field-of-view

FT Fourier Transform

GCVS General Catalogue of Variable Stars

GP Gaussian Process

HZ Habitable Zone

ICWT Inverse Continuous Wavelet Transform

IDWT Inverse Discrete Wavelet Transform

ISWT Inverse Stationary Wavelet Transform

LS Lomb-Scargle

LSST Large Synoptic Survey Telescope

MAD Median Absolute Deviation

MCMC Markov chain Monte Carlo

MRA Multi-Resolution Analysis

PSF Point Spread Function

RACAB Reial Acadèmia de Ciències i Arts de Barcelona

ROA Real Instituto y Observatorio de la Armada

RV Radial Velocity

SDE Signal Detection Efficiency

SNR Signal-to-noise Ratio

SWT Stationary Wavelet Transform

TEP Transits of Extrasolar Planets

TESS Transiting Exoplanet Survey Satellite

TFA Trend Filtering Algorithm

TFAW Wavelet-based Trend Filtering Algorithm

TFRM Telescope Fabra-ROA at Montsec

TFRM-PSES TFRM-Preselected Super-Earth Survey

UTC Coordinated Universal Time

VSX Variable Star Index

WPS Wavelet Power Spectrum

BIBLIOGRAPHY

Abramovich, Felix, Trevor C Bailey, and Theofanis Sapatinas 2000, Wavelet analysis and its statistical applications. Journal of the Royal Statistical Society: Series D (The Statistician) 49.1, pp. 1–29 (cit. on pp. 35, 37).

Affer, L., G. Micela, F. Favata, and E. Flaccomio 2012, The rotation of field stars from CoRoT* data. Monthly Notices of the Royal Astronomical Society 424.1, pp. 11–22. DOI: 10.1111/j.1365-2966.2012.20802.x. eprint: /oup/backfile/content_public/journal/mnras/424/1/10.1111_j.1365-2966.2012.20802.x/3/mnras0424-0011.pdf (cit. on p. 129).

Aigrain, S., H. Parviainen, and B. J. S. Pope 2016, K2SC: flexible systematics correction and detrending of K2 light curves using Gaussian process regression. MNRAS 459, pp. 2408–2419. DOI: 10.1093/mnras/stw706. arXiv: 1603.09167 [astro-ph.SR] (cit. on p. 93).

Alcock, C. et al. 2000, The MACHO Project Large Magellanic Cloud Variable-Star Inventory. IX. Frequency Analysis of the First-Overtone RR Lyrae Stars and the Indication for Nonradial Pulsations. APJ 542, pp. 257–280. DOI: 10.1086/309530. eprint: astro-ph/0005361 (cit. on p. 55).

Almenara, J. M. et al. 2009, Rate and nature of false positives in the CoRoT exoplanet search*. A&A 506.1, pp. 337–341. DOI: 10.1051/0004-6361/200911926 (cit. on p. 105).

Anglada-Escudé, G. et al. 2016, A terrestrial planet candidate in a temperate orbit around Proxima Centauri. Nature 536, pp. 437–440. DOI: 10.1038/nature19106. arXiv: 1609.03449 [astro-ph.EP] (cit. on p. 3).

Antoja, T., F. Figueras, D. Fernández, and J. Torra 2008, Origin and evolution of moving groups. I. Characterization in the observational kinematic-age-metallicity space. A&A 490, pp. 135–150. DOI: 10.1051/0004-6361:200809519. arXiv: 0809.0511 (cit. on p. 10).

Antoja, T. et al. 2012, Kinematic groups beyond the solar neighbourhood with RAVE. MNRAS 426, pp. L1–L5. DOI: 10.1111/j.1745-3933.2012.01310.x. arXiv: 1205.0546 (cit. on p. 10).

Armstrong, D. J., J. Kirk, K. W. F. Lam, J. McCormac, S. R. Walker, D. J. A. Brown, H. P. Osborn, D. L. Pollacco, and J. Spake 2015, K2 Variable Catalogue: Variable stars and eclipsing binaries in K2 campaigns 1 and 0. A&A 579, A19, A19. DOI: 10.1051/0004-6361/201525889. arXiv: 1502.04004 [astro-ph.SR] (cit. on p. 44).

Arnalte-Mur, P., A. Labatie, N. Clerc, V. J. Martínez, J.-L. Starck, M. Lachièze-Rey, E. Saar, and S. Paredes 2012, Wavelet analysis of

baryon acoustic structures in the galaxy distribution. A&A 542, A34, A34. DOI: 10.1051/0004-6361/201118017. arXiv: 1101.1911 (cit. on p. 10).

Aschwanden, M. J., B. Kliem, U. Schwarz, J. Kurths, B. R. Dennis, and R. A. Schwartz 1998, Wavelet Analysis of Solar Flare Hard X-Rays. APJ 505, pp. 941–956. DOI: 10.1086/306200 (cit. on p. 10).

Bakos, G. et al. 2008, HAT-South: A global network of southern hemisphere automated telescopes to detect transiting exoplanets. Proceedings of the International Astronomical Union 4.S253, pp. 354–357. DOI: 10.1017/S174392130802663X (cit. on pp. xi, 3, 47).

Balona, L. A. 2011, Rotational light variations in Kepler observations of A-type stars. MNRAS 415, pp. 1691–1702. DOI: 10.1111/j.1365-2966.2011.18813.x (cit. on p. 129).

Balona, L et al. 2012, Unusual high-frequency oscillations in the KeplerScuti star KIC4840675. MNRAS 424, pp. 1187–1196 (cit. on p. 129).

Barclay, T. et al. 2013, A sub-Mercury-sized exoplanet. Nature 494, pp. 452–454. DOI: 10.1038/nature11914. arXiv: 1305.5587 [astro-ph.EP] (cit. on p. 89).

Barros, S. C. C. et al. 2014, Revisiting the transits of CoRoT-7b at a lower activity level. A&A 569, A74. DOI: 10.1051/0004-6361/201423939 (cit. on p. 87).

Barstow, M. A., S. Jordan, D. O'Donoghue, M. R. Burleigh, R. Napiwotzki, and M. K. Harrop-Allin 1995, RE J0317 – 853: the hottest known highly magnetic DA white dwarf. Monthly Notices of the Royal Astronomical Society 277.3, pp. 971–985. DOI: 10.1093/mnras/277.3.971. eprint: /oup/backfile/content_public/journal/mnras/277/3/10.1093/mnras/277.3.971/2/mnras277-0971.pdf (cit. on pp. 100–103).

Batalha, N. M. et al. 2011, Kepler's First Rocky Planet: Kepler-10b. ApJ 729, 27, p. 27. DOI: 10.1088/0004-637X/729/1/27. arXiv: 1102.0605 [astro-ph.EP] (cit. on p. 89).

Bernardo, J.M., J. O. Berger, A. P. Dawid, and A. F. M. Smith 1996, *Bayesian statistics 5: proceedings of the fifth Valencia international meeting, June 5-9, 1994*. Oxford science publications v. 5. Clarendon Press (cit. on p. 117).

Bertin, E. and S. Arnouts 1996, SExtractor: Software for source extraction. A&AS 117, pp. 393–404. DOI: 10.1051/aas:1996164 (cit. on p. 80).

Bilen, Canan and S. Huzurbazar 2002, Wavelet-Based Detection of Outliers in Time Series. Journal of Computational and Graphical Statistics 11.2, pp. 311–327 (cit. on p. 43).

Bolt, M., T. Hockey, J.A. Palmeri, V. Trimble, T.R. Williams, K. Bracher, R. Jarrell, J.D. Marché, and F.J. Ragep 2007, *Biographical Encyclopedia of Astronomers*. Biographical Encyclopedia of Astronomers. Springer New York (cit. on p. 5).

Bonomo, A. S. et al. 2012, Detection of Neptune-size planetary candidates with CoRoT data. Comparison with the planet occurrence rate derived from Kepler. A&A 547, A110, A110. DOI: 10.1051/0004-6361/201219780. arXiv: 1209.4815 [astro-ph.EP] (cit. on p. 87).

Bonomo, A. S. et al. 2017, A deeper view of the CoRoT-9 planetary system. A small non-zero eccentricity for CoRoT-9b likely generated by planet-planet scattering. A&A 603, A43, A43. DOI: 10.1051/0004-6361/201730624. arXiv: 1703.06477 [astro-ph.EP] (cit. on p. 87).

Borucki, William J. et al. 2010, Kepler Planet-Detection Mission: Introduction and First Results. Science 327.5968, pp. 977–980. DOI: 10.1126/science.1185402. eprint: http://science.sciencemag.org/content/327/5968/977.full.pdf (cit. on pp. xi, 2, 4, 89).

Borucki, William J. et al. 2012, Kepler-22b: A 2.4 Earth-radius Planet in the Habitable Zone of a Sun-like Star. The Astrophysical Journal 745.2, p. 120 (cit. on p. 3).

Bravo, J. P., S. Roque, R. Estrela, I. C. Leão, and J. R. De Medeiros 2014, Wavelets: a powerful tool for studying rotation, activity, and pulsation in Kepler and CoRoT stellar light curves. A&A 568, A34, A34. DOI: 10.1051/0004-6361/201323032. arXiv: 1407.1084 [astro-ph.IM] (cit. on pp. 10, 28).

Breiman, Leo 1995, Better subset regression using the nonnegative garrote. Technometrics 37.4, pp. 373–384 (cit. on p. 35).

Bruno, G. 1584, *De l'infinito, universo e mondi* (cit. on p. 1).

Bryson, S. T. et al. 2010, The Kepler Pixel Response Function. ApJ 713, pp. L97–L102. DOI: 10.1088/2041-8205/713/2/L97. arXiv: 1001.0331 [astro-ph.EP] (cit. on p. 92).

Bryson, Stephen T. et al. 2013, Identification of Background False Positives from Kepler Data. Publications of the Astronomical Society of the Pacific 125, p. 889. DOI: 10.1086/671767. arXiv: 1303.0052 [astro-ph.IM] (cit. on p. 107).

Burrus, C.S., R.A. Gopinath, and H. Guo 1998, *Introduction to Wavelets and Wavelet Transforms: A Primer*. Prentice Hall Series in Advanced. Prentice Hall (cit. on p. 17).

Calderón, A. 1964, Intermediate spaces and interpolation, the complex method. eng. Studia Mathematica 24.2, pp. 113–190 (cit. on p. 9).

Campbell, B., G. A. H. Walker, and S. Yang 1988, A search for substellar companions to solar-type stars. ApJ 331, pp. 902–921. DOI: 10.1086/166608 (cit. on p. 1).

Carone, L. et al. 2012, Planetary transit candidates in the CoRoT LRa01 field. A&A 538, A112. DOI: 10.1051/0004-6361/201116968 (cit. on pp. 107, 113, 120–122, 124, 125, 127).

Carpano, S. et al. 2009, Planetary transit candidates in Corot-IRa01 field *. A&A 506.1, pp. 491–500. DOI: 10.1051/0004-6361/200911882 (cit. on pp. 107–110, 128).

Carter, J. A. and J. N. Winn 2009, Parameter Estimation from Time-series Data with Correlated Errors: A Wavelet-based Method and its Application to Transit Light Curves. ApJ 704, pp. 51–67. DOI: 10.1088/0004-637X/704/1/51. arXiv: 0909.0747 [astro-ph.EP] (cit. on pp. 11, 52).

Charbonneau, D. and D. Deming 2007, The Dynamics-Based Approach to Studying Terrestrial Exoplanets. Submitted to the Exoplanet Task Force (AAAC. arXiv: 0706.1047 (cit. on p. 77).

Charbonneau, D., T. M. Brown, D. W. Latham, and M. Mayor 2000, Detection of Planetary Transits Across a Sun-like Star. ApJ 529, pp. L45–L48. DOI: 10.1086/312457. eprint: astro-ph/9911436 (cit. on pp. 2, 3).

Charbonneau, D. et al. 2005, Detection of Thermal Emission from an Extrasolar Planet. ApJ 626, pp. 523–529. DOI: 10.1086/429991. eprint: astro-ph/0503457 (cit. on pp. xi, 4).

Charbonneau, D. et al. 2009, A super-Earth transiting a nearby low-mass star. Nature 462, pp. 891–894. DOI: 10.1038/nature08679. arXiv: 0912.3229 [astro-ph.EP] (cit. on p. 3).

CoRot Team 2016, *The CoRoT Legacy Book: The adventure of the ultra high precision photometry from space, by the CoRot Team*. DOI: 10.1051/978-2-7598-1876-1 (cit. on p. 87).

Cohen, A., Ingrid Daubechies, and J.-C. Feauveau 1992, Biorthogonal bases of compactly supported wavelets. Communications on Pure and Applied Mathematics 45.5, pp. 485–560. DOI: 10.1002/cpa.3160450502 (cit. on pp. 10, 24, 26, 52).

Cooley, James and John Tukey 1965, An Algorithm for the Machine Calculation of Complex Fourier Series. Mathematics of Computation 19.90, pp. 297–301 (cit. on p. 12).

Copernicus, N. 1543, *D revolutionibus orbium coelestium* (cit. on p. 1).

Corbett, H., O. Fors, J. Ratzloff, D. del Ser, and N. M. Law 2016, *The Evryscope: Data Analysis Pipeline for the First Gigapixel-Scale Telescope*. Presentation at the Research Corporation for Science Advancement (cit. on p. 83).

Corbett, H., N. Law, E. Goeke, J. Ratzloff, W. Howard, O. Fors, D. del Ser, and R. M. Quimby 2018, Pre-Discovery Detection of ASASSN-18fv by Evryscope. The Astronomer's Telegram 11467 (cit. on p. 235).

Croisier, A., D. Esteban, and C. Galand 1976, Perfect channel splitting by use of interpolation/decimation/tree decomposition techniques. Proceedings of International Conference on Information Sciences and System, pp. 443–446 (cit. on p. 18).

Crossfield, Ian J. M. et al. 2016, 197 Candidates and 104 Validated Planets in K2's First Five Fields. ApJS 226.1, p. 7 (cit. on pp. 133–138, 141).

Cubillos, P., J. Harrington, T. J. Loredo, N. B. Lust, J. Blecic, and M. Stemm 2017, On Correlated-noise Analyses Applied to Exoplanet Light Curves. AJ 153, 3, p. 3. DOI: 10.3847/1538-3881/153/1/3. arXiv: 1610.01336 [astro-ph.EP] (cit. on p. 11).

Cutri, R. M. et al. 2003, VizieR Online Data Catalog: 2MASS All-Sky Catalog of Point Sources (Cutri+ 2003). VizieR Online Data Catalog, II/246, pp. II/246 (cit. on p. 120).

Daubechies, I. 1992, *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics. DOI: 10.1137/1.9781611970104. eprint: http://epubs.siam.org/doi/pdf/10.1137/1.9781611970104 (cit. on p. 15).

Daubechies, Ingrid 1988, Orthonormal bases of compactly supported wavelets. Communications on Pure and Applied Mathematics 41.7, pp. 909–996. DOI: 10.1002/cpa.3160410705 (cit. on p. 10).

Debosscher, J., Sarro, L. M., Aerts, C., Cuypers, J., Vandenbussche, B., Garrido, R., and Solano, E. 2007, Automated supervised classification of variable stars* - I. Methodology. A&A 475.3, pp. 1159–1183. DOI: 10.1051/0004-6361:20077638 (cit. on p. 110).

Debosscher, J. et al. 2009, Automated supervised classification of variable stars in the CoRoT programme *** - Method and application to the first four exoplanet fields. A&A 506.1, pp. 519–534. DOI: 10.1051/0004-6361/200911618 (cit. on p. 110).

Deeg, H. J. et al. 2009, Ground-based photometry of space-based transit detections: photometric follow-up of the CoRoTssion *. A&A 506.1, pp. 343–352. DOI: 10.1051/0004-6361/200912011 (cit. on p. 107).

Deeg, H. J. et al. 1998, Near-term detectability of terrestrial extrasolar planets: TEP network observations of CM Draconis. A&A 338, pp. 479–490. eprint: astro-ph/9806371 (cit. on p. 2).

Deleuil, M. et al. 2009, Exo-Dat: An Information System in Support of the CoRoT/Exoplanet Science. AJ 138, pp. 649–663. DOI: 10.1088/0004-6256/138/2/649 (cit. on p. 86).

Deleuil, M. et al. 2018, Planets, candidates, and binaries from the CoRoT/Exoplanet programme: the CoRoT transit catalogue. ArXiv e-prints. arXiv: 1805.07164 [astro-ph.EP] (cit. on pp. 86, 87).

Deming, D., S. Seager, L. J. Richardson, and J. Harrington 2005, Infrared radiation from an extrasolar planet. Nature 434, pp. 740–743. DOI: 10.1038/nature03507. eprint: astro-ph/0503554 (cit. on pp. xi, 4).

Devor, J. 2005, Solutions for 10,000 Eclipsing Binaries in the Bulge Fields of OGLE II Using DEBiL. ApJ 628, pp. 411–425. DOI: 10.1086/431170. eprint: astro-ph/0504399 (cit. on p. 113).

Devyatkin, A. V., D. L. Gorshanov, V. V. Kouprianov, and I. A. Verestchagina 2010, Apex I and Apex II software packages for the reduction of astronomical CCD observations. Solar System Research 44.1, pp. 68–80. DOI: 10.1134/S0038094610010090 (cit. on pp. 78, 80, 95, 160).

Djorgovski, S. G. et al. 2011, The Catalina Real-Time Transient Survey (CRTS). ArXiv e-prints. arXiv: 1102.5004 [astro-ph.IM] (cit. on p. 47).

Donoho, D. L. 1995, De-noising by Soft-thresholding. IEEE Trans. Inf. Theor. 41.3, pp. 613–627. DOI: 10.1109/18.382009 (cit. on p. 10).

Donoho, David L. and Iain M. Johnstone 1994a, Ideal denoising in an orthonormal basis chosen from a library of bases. Comptes Rendus de l'Académie des Sciences, Série I: Mathématique 319.12, pp. 1317–1322 (cit. on pp. 10, 34, 38).

— 1994 b, Ideal spatial adaptation by wavelet shrinkage. Biometrika 81.3, p. 425. DOI: 10.1093/biomet/81.3.425. eprint: /oup/backfile/Content_public/Journal/biomet/81/3/10.1093/biomet/81.3.425/2/81-3-425.pdf (cit. on pp. 10, 33–37).

Donoho, David L and Iain M Johnstone 1995a, Adapting to unknown smoothness via wavelet shrinkage. Journal of the american statistical association 90.432, pp. 1200–1224 (cit. on p. 36).

— 1995 b, Adapting to unknown smoothness via wavelet shrinkage. Journal of the american statistical association 90.432, pp. 1200–1224 (cit. on p. 36).

Doyle, L. R. et al. 2011, Kepler-16: A Transiting Circumbinary Planet. Science 333, p. 1602. DOI: 10.1126/science.1210923. arXiv: 1109.3432 [astro-ph.EP] (cit. on p. 89).

Dressing, Courtney D., Elisabeth R. Newton, Joshua E. Schlieder, David Charbonneau, Heather A. Knutson, Andrew Vanderburg, and Evan Sinukoff 2017a, Characterizing K2 Candidate Planetary Systems Orbiting Low-mass Stars. I. Classifying Low-mass Host Stars Observed during Campaigns 1–7. ApJ 836.2, p. 167 (cit. on p. 142).

Dressing, Courtney D. et al. 2017b, Characterizing K2 Candidate Planetary Systems Orbiting Low-mass Stars. II. Planetary Systems Observed During Campaigns 1–7. ApJ 154.5, p. 207 (cit. on pp. 142, 143, 146, 147).

E., Van Cleve J. and Caldwell D. A 2016, *Kepler Instrument Handbook*. KSCI-19033-001. Moffett Field: NASA Ames Research Center (cit. on pp. 88, 89).

Efromovich, Sam 1999, Quasi-linear wavelet estimation. Journal of the American Statistical Association 94.445, pp. 189–204 (cit. on p. 36).

Erikson, A. et al. 2012, Planetary transit candidates in the CoRoT-SRc01 field. A&A 539, A14. DOI: 10.1051/0004-6361/201116934 (cit. on p. 87).

Evans, D. F. et al. 2018, High-resolution Imaging of Transiting Extrasolar Planetary systems (HITEP). II. Lucky Imaging results from 2015 and 2016. A&A 610, A20, A20. DOI: 10.1051/0004-6361/201731855 (cit. on p. 134).

Eylen, Vincent Van et al. 2016, The K2-ESPRINT Project. II. Spectroscopic Follow-up of Three Exoplanet Systems from Campaign 1 of K2. ApJ 820.1, p. 56 (cit. on p. 133).

Ferrario, Lilia, S. Vennes, D. T. Wickramasinghe, J. A. Bailey, and D. J. Christian 1997, EUVE J0317 — 855: a rapidly rotating, high-field magnetic white dwarf. MNRAS 292.2, pp. 205–217. DOI: 10.1093/mnras/292.2.205. eprint: /oup/backfile/content_public/journal/mnras/292/2/10.1093_mnras_292.2.205/1/292-2-205.pdf (cit. on p. 101).

Foreman-Mackey, D., D. W. Hogg, D. Lang, and J. Goodman 2013, emcee: The MCMC Hammer. PASP 125, p. 306. DOI: 10.1086/670067. arXiv: 1202.3665 [astro-ph.IM] (cit. on pp. 68, 116, 117).

Foreman-Mackey, Daniel, Benjamin T. Montet, David W. Hogg, Timothy D. Morton, Dun Wang, and Bernhard Schölkopf 2015, A Systematic Search for Transiting Planets in the K2 Data. ApJ 806.2, p. 215 (cit. on p. 142).

Fors, O., A. Richichi, X. Otazu, and J. Núñez 2008, A new wavelet-based approach for the automated treatment of large sets of lunar occultation data. A&A 480, pp. 297–304. DOI: 10.1051/0004-6361:20078987. arXiv: 0711.0537 (cit. on p. 10).

Fors, O., J. Núñez, J. L. Muiños, F. J. Montojo, R. Baena-Gallé, J. Boloix, R. Morcillo, M. T. Merino, E. C. Downey, and M. J. Mazur 2013, Telescope Fabra ROA Montsec: A New Robotic Wide Field Baker-Nunn Facility. PASP 125, pp. 522–538. DOI: 10.1086/670941. arXiv: 1211.5581 [astro-ph.IM] (cit. on pp. xi, 3, 47, 75).

Fors, O., N. M. Law, J. Ratzloff, D. del Ser, P. J. Wulfken, and D. Kavanaugh 2015a, Evrystats for Evryplanets: planets from the first all-sky gigapixel scale telescope. IAU General Assembly 22, 2253010, p. 2253010 (cit. on p. 236).

— 2015 b, The Evryscope and extrasolar planets. IAU General Assembly 22, 2258237, p. 2258237 (cit. on p. 236).

Fors, O., N. M. Law, J. Ratzloff, H. Corbett, D. del Ser, W. Howard, and S. Cox 2017, Engaging AAVSO members in Stellar Astrophysics Follow-up from The Evryscope Data (Abstract). Journal of the American Association of Variable Star Observers (JAAVSO) 45, p. 129 (cit. on p. 235).

Fridlund, M., A. Baglin, J. Lochard, and L. Conroy, eds. 2006, *The CoRoT Mission Pre-Launch Status - Stellar Seismology and Planet Finding*. Vol. 1306. ESA Special Publication (cit. on pp. xi, 4, 84, 87).

Gabor, D. 1946, Theory of communication. Part 1: The analysis of information. Electrical Engineers - Part III: Radio and Communi-

cation Engineering, Journal of the Institution of 93.26, pp. 429–441. DOI: 10.1049/ji-3-2.1946.0074 (cit. on p. 14).

Gaia Collaboration 2018, VizieR Online Data Catalog: Gaia DR2 (Gaia Collaboration, 2018). VizieR Online Data Catalog, I/345, pp. I/345 (cit. on pp. 99, 101, 109, 112, 120, 128, 133, 142).

García-Berro, Enrique, Pablo Lorén-Aguilar, Gabriela Aznar-Siguán, Santiago Torres, Judit Camacho, Leandro G. Althaus, Alejandro H. Córsico, Baybars Külebi, and Jordi Isern 2012, Double Degenerate Mergers as Progenitors of High-field Magnetic White Dwarfs. ApJ 749.1, p. 25 (cit. on p. 101).

Gillon, M. et al. 2017, Seven temperate terrestrial planets around the nearby ultracool dwarf star TRAPPIST-1. Nature 542, pp. 456–460. DOI: 10.1038/nature21360. arXiv: 1703.01424 [astro-ph.EP] (cit. on p. 3).

Giménez de Castro, C. G., J.-P. Raulin, C. H. Mandrini, P. Kaufmann, and A. Magun 2001, Multi-resolution wavelet analysis of high time resolution millimeter wavelength observations of solar bursts. A&A 366, pp. 317–325. DOI: 10.1051/0004-6361:20000088 (cit. on p. 10).

Gonzalez, G., M. K. Carlson, and R. W. Tobin 2010, Parent stars of extrasolar planets - X. Lithium abundances and v sini revisited. MNRAS 403, pp. 1368–1380. DOI: 10.1111/j.1365-2966.2009.16195.x. arXiv: 0912.1621 [astro-ph.SR] (cit. on p. 123).

Górski, K. M., E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann 2005, HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. ApJ 622, pp. 759–771. DOI: 10.1086/427976. eprint: astro-ph/0409513 (cit. on p. 84).

Grané, Aurea and Helena Veiga 2010, Wavelet-based detection of outliers in financial time series. Computational Statistics & Data Analysis 54.11. The Fifth Special Issue on Computational Econometrics, pp. 2580 –2593. DOI: http://dx.doi.org/10.1016/j.csda.2009.12.010 (cit. on p. 43).

Grossmann, A. and J. Morlet 1984, Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape. SIAM Journal of Mathematical Analysis 15.4, pp. 723–736. DOI: 10.1137/0515056 (cit. on pp. 9, 14).

Grubbs, Frank E. 1950, Sample Criteria for Testing Outlying Observations. Ann. Math. Statist. 21.1, pp. 27–58. DOI: 10.1214/aoms/1177729885 (cit. on p. 43).

Grziwa, S., J. Korth, and M. Pätzold 2014, VARLET and PHALET two wavelet based filter methods to separate stellar variation, orbital disturbances and instrumental effects from transit events in stellar light curves. European Planetary Science Congress 2014, EPSC Abstracts, Vol. 9, id. EPSC2014-156 9, EPSC2014-156, EPSC2014–156 (cit. on pp. 11, 52, 87).

Grziwa, S., M. Pätzold, and L. Carone 2012, The needle in the haystack: searching for transiting extrasolar planets in CoRoT stellar light curves. MNRAS 420, pp. 1045–1052. DOI: 10.1111/j.1365-2966.2011.19970.x (cit. on pp. 87, 113, 118).

Grziwa, S., J. Korth, M. Paetzold, and KEST 2016, Wavelet-based filter methods for the detection of small transiting planets: Application to Kepler and K2 light curves. AAS/Division for Planetary Sciences Meeting Abstracts. Vol. 48. AAS/Division for Planetary Sciences Meeting Abstracts, p. 122.02 (cit. on pp. 11, 52, 60).

Guenther, E. W. and Tal-Or, L. 2010, High-resolution infrared spectroscopy as a tool to detect false positives of transit search programs*. A&A 521, A83. DOI: 10.1051/0004-6361/200913530 (cit. on p. 107).

Guenther, E. W. et al. 2013, High angular resolution imaging and infrared spectroscopy of CoRoT candidates. A&A 556, A75. DOI: 10.1051/0004-6361/201220902 (cit. on pp. 120, 124).

Haar, A. 1910, Zur Theorie der orthogonalen Funktionensysteme. (Erste Mitteilung). Mathematische Annalen 69, pp. 331–371 (cit. on pp. 12, 20, 24).

Haas, M. R. et al. 2010, Kepler Science Operations. ApJ 713, pp. L115–L119. DOI: 10.1088/2041-8205/713/2/L115. arXiv: 1001.0437 [astro-ph.EP] (cit. on p. 88).

Hale, Alan and Laurance R. Doyle 1994, The Photometric Method of Extrasolar Planet Detection Revisited. Ap&SS 212, pp. 335–348. DOI: 10.1007/BF00984537 (cit. on p. 2).

Hall, Peter and Prakash Patil 1996, Effect of threshold rules on performance of wavelet-based curve estimators. Statistica Sinica, pp. 331–345 (cit. on p. 36).

Hambleton, K. et al. 2018, KIC 8164262: a heartbeat star showing tidally induced pulsations with resonant locking. MNRAS 473, pp. 5165–5176. DOI: 10.1093/mnras/stx2673. arXiv: 1706.05051 [astro-ph.SR] (cit. on p. 129).

Hatzes, A. P., W. D. Cochran, M. Endl, B. McArthur, D. B. Paulson, G. A. H. Walker, B. Campbell, and S. Yang 2003, A Planetary Companion to γ Cephei A. ApJ 599, pp. 1383–1394. DOI: 10.1086/379281. eprint: astro-ph/0305110 (cit. on p. 1).

Hayward, T. L., B. Brandl, B. Pirger, C. Blacken, G. E. Gull, J. Schoenwald, and J. R. Houck 2001, PHARO: A Near-Infrared Camera for the Palomar Adaptive Optics System. PASP 113, pp. 105–118. DOI: 10.1086/317969 (cit. on p. 133).

Hebb, L. et al. 2009, WASP-12b: The Hottest Transiting Extrasolar Planet Yet Discovered. ApJ 693, pp. 1920–1928. DOI: 10.1088/0004-637X/693/2/1920. arXiv: 0812.3240 (cit. on p. 2).

Henden, A. A., D. L. Welch, D. Terrell, and S. E. Levine 2009, The AAVSO Photometric All-Sky Survey (APASS). American Astro-

nomical Society Meeting Abstracts #214. Vol. 214. American Astronomical Society Meeting Abstracts, p. 669 (cit. on p. 84).

Henden, A. A., D. Terrell, D. Welch, and T. C. Smith 2010, New Results From The AAVSO Photometric All-sky Survey (APASS). American Astronomical Society Meeting Abstracts #215. Vol. 42. Bulletin of the American Astronomical Society, p. 515 (cit. on p. 84).

Henden, A. A., M. Templeton, D. Terrell, T. C. Smith, S. Levine, and D. Welch 2016, VizieR Online Data Catalog: AAVSO Photometric All Sky Survey (APASS) DR9 (Henden+, 2016). VizieR Online Data Catalog 2336 (cit. on pp. 84, 142).

Henize, K. G. 1957, The Baker-Nunn Satellite-Tracking Camera. S&T 16 (cit. on p. 75).

Henry, G. W., G. Marcy, R. P. Butler, and S. S. Vogt 1999, HD 209458. IAU Circ. 7307 (cit. on p. 2).

Hermes, J. J., B. T. Gänsicke, Nicola Pietro Gentile Fusillo, R. Raddi, M. A. Hollands, E. Dennihy, J. T. Fuchs, and S. Redfield 2017, When flux standards go wild: white dwarfs in the age of Kepler. MNRAS 468.2, pp. 1946–1952. DOI: 10.1093/mnras/stx567. eprint: /oup/backfile/content_public/journal/mnras/468/2/10.1093_mnras_stx567/1/stx567.pdf (cit. on pp. 101, 102).

Holschneider, M., R. Kronland-Martinet, J. Morlet, and Ph. Tchamitchian 1989, *A Real-Time Algorithm for Signal Analysis with the Help of the Wavelet Transform*. Ed. by Jean-Michel Combes, Alexander Grossmann, and Philippe Tchamitchian. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 286–297. DOI: 10.1007/978-3-642-97177-8_28 (cit. on pp. 10, 22, 52).

Holwarda, Johannes Phocylides 1640, *Panselenos ekleiptike diaugazou [romanized form], id est, Dissertatio astronomica quae occasione ultimi lunaris anni 1638 deliquii [microform] : manuductio sit ad cognoscendum, I. Statum astronomiae, praaesertim Lansbergianae, II. Novorum phenomenon exortu & interitum / autore Ioanne Phocylide Holwarda*. Latin. Typis Idzardi Alberti, ejusdem[que] & Ioannis Fabiani Theuring, impensis Franekerae, [24], 288 p. : (cit. on pp. xii, 4).

Howard, W. S. et al. 2018a, Evryscope Detection of the First Proxima Superflare: Impacts on the Atmosphere and Habitability of Proxima b. LPI Contributions 2065, 2039, p. 2039 (cit. on p. 235).

Howard, W. S., N. Law, O. Fors, H. T. Corbett, J. Ratzloff, and D. del Ser 2018b, Stellar activity for every TESS star in the Southern sky. American Astronomical Society Meeting Abstracts #231. Vol. 231. American Astronomical Society Meeting Abstracts, p. 310.03 (cit. on p. 235).

Howard, W. S. et al. 2018c, The First Naked-eye Superflare Detected from Proxima Centauri. ApJ 860, L30, p. L30. DOI: 10.3847/2041-8213/aacaf3. arXiv: 1804.02001 [astro-ph.EP] (cit. on p. 235).

Howard, W., O. Fors, J. Ratzloff, H. Corbett, D. del Ser, and N. Law 2017, EvryFlare: Flare rates and intensities for every 10 < g < 15 solar-type and red dwarf star in the Southern sky. Radio Exploration of Planetary Habitability (AASTCS5). Vol. 49, p. 202.05 (cit. on p. 235).

Howell, S. B. et al. 2014, The K2 Mission: Characterization and Early Results. PASP 126, p. 398. DOI: 10.1086/676406. arXiv: 1402.5163 [astro-ph.IM] (cit. on pp. 90, 91).

Huygens Christiaan, 1629-1695 1698, *Cosmotheoros, sive De Terris Coelestibus*. Ed. by Apud Adrianum Moetjens. Hagae-Comitum : Apud Adrianum Moetjens (cit. on p. 1).

Irwin, Jonathan, David Charbonneau, Philip Nutzman, and Emilio Falco 2009, The MEarth project: searching for transiting habitable super-Earth planets around nearby M-dwarfs. AIP Conference Proceedings 1094.1, pp. 445–448. DOI: 10.1063/1.3099143. eprint: http://aip.scitation.org/doi/pdf/10.1063/1.3099143 (cit. on pp. xi, 3, 47, 77).

J. E. Van Cleve J. L. Christiansen, J. M. Jenkins D. A. Caldwell T. Barclay S. T. Bryson C. J. Burke J. Campbell J. Catanzarite B. D. Clarke J. L. Coughlin F. Girouard M. R. Haas T. C. Klaus J. J. Kolodziejczak J. Li S. D. McCauliff R. L. Morris F. Mullally E. V. Quintana J. Rowe A. Sabale S. Seader J. C. Smith M. D. Still P. G. Tenenbaum S. E. Thompson J. D. Twicken A. K. Uddin and K. Zamudio 2016, *Kepler Data Characteristics Hand book*. KSCI-19040-005. Moffett Field: NASA Ames Research Center (cit. on p. 90).

Jacob, W. S. 1855, On certain Anomalies presented by the Binary Star 70 Ophiuchi. MNRAS 15, p. 228. DOI: 10.1093/mnras/15.9.228 (cit. on p. 1).

Jain, Anil, Sharath Pankanti, and The Pennsylvania State University CiteSeer Archives 2001, Automated Fingerprint Identification and Imaging Systems. English (cit. on p. 10).

Jayasinghe, T et al. 2018, The ASAS-SN catalogue of variable stars I: The Serendipitous Survey. Monthly Notices of the Royal Astronomical Society 477.3, pp. 3145–3163. DOI: 10.1093/mnras/sty838. eprint: /oup/backfile/content_public/journal/mnras/477/3/10.1093_mnras_sty838/2/sty838.pdf (cit. on p. 99).

Jenkins, J. M. et al. 2010, Overview of the Kepler Science Processing Pipeline. ApJ 713, pp. L87–L91. DOI: 10.1088/2041-8205/713/2/L87. arXiv: 1001.0258 [astro-ph.EP] (cit. on pp. 92, 93).

Johnstone, Iain M and Bernard W Silverman 1997, Wavelet threshold estimators for data with correlated noise. Journal of the royal statistical society: series B (statistical methodology) 59.2, pp. 319–351 (cit. on p. 35).

Kaiser, Nick, William Burgett, Ken Chambers, Larry Denneau, Jim Heasley, Robert Jedicke, Eugene Magnier, Jeff Morgan, Peter On-

aka, and John Tonry 2010, *The Pan-STARRS wide-field optical/NIR imaging survey*. DOI: 10.1117/12.859188 (cit. on pp. 47, 142).

Kalas, P., J. R. Graham, E. Chiang, M. P. Fitzgerald, M. Clampin, E. S. Kite, K. Stapelfeldt, C. Marois, and J. Krist 2008, Optical Images of an Exosolar Planet 25 Light-Years from Earth. Science 322, p. 1345. DOI: 10.1126/science.1166609. arXiv: 0811.1994 (cit. on p. 3).

Keller, S. C. et al. 2007, The SkyMapper Telescope and The Southern Sky Survey. PASA 24, pp. 1–12. DOI: 10.1071/AS07001. eprint: astro-ph/0702511 (cit. on p. 47).

Kilic, Mukremin, Alexandros Gianninas, Keaton J. Bell, Brandon Curd, Warren R. Brown, J. J. Hermes, Patrick Dufour, John P. Wisniewski, D. E. Winget, and K. I. Winget 2015, A Dark Spot on a Massive White Dwarf. ApJ 814.2, p. L31 (cit. on p. 101).

Kim, Joong Hoon, Zong Woo Geem, and Eung Seok Kim 2001, PARAMETER ESTIMATION OF THE NONLINEAR MUSKINGUM MODEL USING HARMONY SEARCH1. JAWRA Journal of the American Water Resources Association 37.5, pp. 1131–1138. DOI: 10.1111/j.1752-1688.2001.tb03627.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1752-1688.2001.tb03627.x (cit. on pp. 113, 119).

Kipping, D. M. and D. S. Spiegel 2011, Detection of visible light from the darkest world. MNRAS 417, pp. L88–L92. DOI: 10.1111/j.1745-3933.2011.01127.x. arXiv: 1108.2297 [astro-ph.EP] (cit. on p. 3).

Knorr, Edwin M., Raymond T. Ng, and Vladimir Tucakov 2000, Distance-based outliers: algorithms and applications. The VLDB Journal 8.3, pp. 237–253. DOI: 10.1007/s007780050006 (cit. on p. 43).

Konacki, M., G. Torres, S. Jha, and D. D. Sasselov 2003, An extrasolar planet that transits the disk of its parent star. Nature 421, pp. 507–509. DOI: 10.1038/nature01379 (cit. on p. 2).

Koposov, S. and O. Bartunov 2006, Q3C, Quad Tree Cube – The new Sky-indexing Concept for Huge Astronomical Catalogues and its Realization for Main Astronomical Queries (Cone Search and Xmatch) in Open Source Database PostgreSQL. Astronomical Data Analysis Software and Systems XV. Ed. by C. Gabriel, C. Arviset, D. Ponz, and S. Enrique. Vol. 351. Astronomical Society of the Pacific Conference Series, p. 735 (cit. on p. 84).

Kovacs, G. and G. A. Bakos 2008, Application of the Trend Filtering Algorithm in the search for multiperiodic signals. Communications in Asteroseismology 157, pp. 82–86. arXiv: 0812.2824 (cit. on pp. 50, 64).

Kovács, G., G. Bakos, and R. W. Noyes 2005, A trend filtering algorithm for wide-field variability surveys. MNRAS 356, pp. 557–567. DOI: 10.1111/j.1365-2966.2004.08479.x. eprint: astro-ph/0411724 (cit. on pp. xiii, 6, 39, 47–51, 54, 56, 96, 118, 151, 152).

Kovács, G., S. Zucker, and T. Mazeh 2002, A box-fitting algorithm in the search for periodic transits. A&A 391, pp. 369–377. DOI: 10.1051/0004-6361:20020802. eprint: astro-ph/0206099 (cit. on pp. 49, 53, 55, 56, 58).

Kovtyukh, V. V., C. Soubiran, S. I. Belik, and N. I. Gorlova 2003, High precision effective temperatures for 181 F-K dwarfs from line-depth ratios. A&A 411, pp. 559–564. DOI: 10.1051/0004-6361:20031378. eprint: astro-ph/0308429 (cit. on p. 123).

Kreidberg, L. 2015, batman: BAsic Transit Model cAlculatioN in Python. PASP 127, p. 1161. DOI: 10.1086/683602. arXiv: 1507.08285 [astro-ph.EP] (cit. on pp. 38, 55).

Kreidberg, L., J. L. Bean, J.-M. Désert, B. Benneke, D. Deming, K. B. Stevenson, S. Seager, Z. Berta-Thompson, A. Seifahrt, and D. Homeier 2014, Clouds in the atmosphere of the super-Earth exoplanet GJ1214b. Nature 505, pp. 69–72. DOI: 10.1038/nature12888. arXiv: 1401.0022 [astro-ph.EP] (cit. on pp. xi, 4).

Kumar, Pawan, Chi On Ao, and Eliot J. Quataert 1995, Tidal Excitation of Modes in Binary Systems with Applications to Binary Pulsars. ApJ 449, p. 294. DOI: 10.1086/176055. arXiv: astro-ph/9503053 [astro-ph] (cit. on p. 129).

Kurtz, D. W. 1982, Rapidly oscillating Ap stars. MNRAS 200.3, pp. 807–859. DOI: 10.1093/mnras/200.3.807. eprint: /oup/backfile/content_public/journal/mnras/200/3/10.1093/mnras/200.3.807/2/mnras200-0807.pdf (cit. on p. 129).

Kushniruk, I., T. Schirmer, and T. Bensby 2017, Kinematic structures of the solar neighbourhood revealed by Gaia DR1/TGAS and RAVE. A&A 608, A73, A73. DOI: 10.1051/0004-6361/201731147. arXiv: 1709.06905 (cit. on p. 10).

Latham, D. W., R. P. Stefanik, T. Mazeh, M. Mayor, and G. Burki 1989, The unseen companion of HD114762 - A probable brown dwarf. Nature 339, pp. 38–40. DOI: 10.1038/339038a0 (cit. on p. 1).

Law, N. M. et al. 2009, The Palomar Transient Factory: System Overview, Performance, and First Results. PASP 121, p. 1395. DOI: 10.1086/648598. arXiv: 0906.5350 [astro-ph.IM] (cit. on p. 47).

Law, N. M., O. Fors, P. Wulfken, J. Ratzloff, and D. Kavanaugh 2014, The Evryscope: the first full-sky gigapixel-scale telescope. Ground-based and Airborne Telescopes V. Vol. 9145. Proc. SPIE, 91450Z. DOI: 10.1117/12.2057031. arXiv: 1407.0026 [astro-ph.IM] (cit. on pp. xi, 4, 47, 81).

Law, N. M. et al. 2015, Evryscope Science: Exploring the Potential of All-Sky Gigapixel-Scale Telescopes. PASP 127, pp. 234–249. DOI: 10.1086/680521. arXiv: 1501.03162 [astro-ph.IM] (cit. on pp. xi, 47, 81, 236).

Law, N. M., O. Fors, J. Ratzloff, H. Corbett, D. del Ser, and P. Wulfken 2016, The Evryscope: design and performance of the first full-sky gigapixel-scale telescope. Society of Photo-Optical Instrumenta-

tion Engineers (SPIE) Conference Series. Vol. 9906. Proc. SPIE, p. 99061M. DOI: 10.1117/12.2233349 (cit. on pp. 47, 81, 83, 235).

Law, Nicholas M. et al. 2014, Robotic Laser Adaptive Optics Imaging of 715 Kepler Exoplanet Candidates Using Robo-AO. The Astrophysical Journal 791.1, p. 35 (cit. on p. 107).

Léger, A. et al. 2009, Transiting exoplanets from the CoRoT space mission * - VIII. CoRoT-7b: the first super-Earth with measured radius. A&A 506.1, pp. 287–302. DOI: 10.1051/0004-6361/200911933 (cit. on pp. 105, 107).

Lépine, S. and E. Gaidos 2011, An All-sky Catalog of Bright M Dwarfs. AJ 142, 138, p. 138. DOI: 10.1088/0004-6256/142/4/138. arXiv: 1108.2719 [astro-ph.SR] (cit. on p. 77).

Lépine, S. and M. M. Shara 2005, A Catalog of Northern Stars with Annual Proper Motions Larger than 0.15" (LSPM-NORTH Catalog). AJ 129, pp. 1483–1522. DOI: 10.1086/427854. eprint: astro-ph/0412070 (cit. on p. 77).

Liu, Yonggang, X San Liang, and Robert H Weisberg 2007, Rectification of the bias in the wavelet power spectrum. Journal of Atmospheric and Oceanic Technology 24.12, pp. 2093–2102 (cit. on p. 28).

Llebaria, A. and P. Guterman 2006, Building up Photometric Apertures for the Exoplanet Channel. The CoRoT Mission Pre-Launch Status - Stellar Seismology and Planet Finding. Ed. by M. Fridlund, A. Baglin, J. Lochard, and L. Conroy. Vol. 1306. ESA Special Publication, p. 293 (cit. on p. 85).

Luger, R., E. Agol, E. Kruse, R. Barnes, A. Becker, D. Foreman-Mackey, and D. Deming 2016, EVEREST: Pixel Level Decorrelation of K2 Light Curves. AJ 152, 100, p. 100. DOI: 10.3847/0004-6256/152/4/100. arXiv: 1607.00524 [astro-ph.EP] (cit. on pp. 93, 133).

Luger, R., E. Kruse, D. Foreman-Mackey, E. Agol, and N. Saunders 2017, An update to the EVEREST K2 pipeline: Short cadence, saturated stars, and Kepler-like photometry down to Kp = 15. ArXiv e-prints. arXiv: 1702.05488 [astro-ph.IM] (cit. on pp. 93, 133, 136, 140, 143, 145).

Machado, D. P., A. Leonard, J.-L. Starck, F. B. Abdalla, and S. Jouvel 2013, Darth Fader: Using wavelets to obtain accurate redshifts of spectra at very low signal-to-noise. A&A 560, A83, A83. DOI: 10.1051/0004-6361/201219857. arXiv: 1309.3579 (cit. on p. 10).

Malavolta, Luca et al. 2018, An Ultra-short Period Rocky Super-Earth with a Secondary Eclipse and a Neptune-like Companion around K2-141. The Astronomical Journal 155.3, p. 107 (cit. on p. 3).

Mallat, S. G. 1989, A theory for multiresolution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 11.7, pp. 674–693. DOI: 10.1109/34.192463 (cit. on pp. 10, 16, 17).

Mallat, S. 2008, *A Wavelet Tour of Signal Processing: The Sparse Way*. Elsevier Science (cit. on pp. 26, 27, 34, 35, 38, 52).

Mallat, Stephane and Wen Liang Hwang 1992, Singularity detection and processing with wavelets. IEEE transactions on information theory 38.2, pp. 617–643 (cit. on p. 43).

Mandel, K. and E. Agol 2002, Analytic Light Curves for Planetary Transit Searches. ApJ 580, pp. L171–L175. DOI: 10.1086/345520. eprint: astro-ph/0210099 (cit. on pp. 38, 51, 55, 60, 67, 68, 113, 116, 119, 123, 127, 137).

Marois, C., B. Macintosh, T. Barman, B. Zuckerman, I. Song, J. Patience, D. Lafrenière, and R. Doyon 2008, Direct Imaging of Multiple Planets Orbiting the Star HR 8799. Science 322, p. 1348. DOI: 10.1126/science.1166585. arXiv: 0811.2606 (cit. on p. 3).

Martinez, V. J., S. Paredes, and E. Saar 1993, Wavelet analysis of the multifractal character of the galaxy distribution. MNRAS 260, pp. 365–375. DOI: 10.1093/mnras/260.2.365 (cit. on p. 10).

Martínez, Vicent J., Jean-Luc Starck, Enn Saar, David L. Donoho, Simon C. Reynolds, Pablo de la Cruz, and Silvestre Paredes 2005, Morphology of the Galaxy Distribution from Wavelet Denoising. The Astrophysical Journal 634.2, p. 744 (cit. on p. 10).

Mayo, Andrew W. et al. 2018, 275 Candidates and 149 Validated Planets Orbiting Bright Stars in K 2 Campaigns 0–10. ApJ 155.3, p. 136 (cit. on pp. 92, 134, 137, 138).

Mayor, M. and D. Queloz 1995, A Jupiter-mass companion to a solar-type star. Nature 378, pp. 355–359. DOI: 10.1038/378355a0 (cit. on pp. xi, 1, 2).

Meyer, Y. 1992, *Ondelettes et algorithmes concurrents*. Actualités scientifiques et industrielles. Hermann (cit. on p. 13).

Meyers, S. D., S. D. Kelly, and J. J. O'Brien 1993, An Introduction to Wavelet Analysis in Oceanography and Meteorology: With Application to the Dispersion of Yanai Waves. Monthly Weather Review 121.10, pp. 2858–2866. DOI: 10.1175/1520-0493(1993)121<2858:AITWAI>2.0.CO;2. eprint: http://dx.doi.org/10.1175/1520-0493(1993)121<2858:AITWAI>2.0.CO;2 (cit. on p. 28).

Montet, Benjamin T., Timothy D. Morton, Daniel Foreman-Mackey, John Asher Johnson, David W. Hogg, Brendan P. Bowler, David W. Latham, Allyson Bieryla, and Andrew W. Mann 2015, Stellar and Planetary Properties of K2 Campaign 1 Candidates and Validation of 17 Planets, Including a Planet Receiving Earth-like Insolation. ApJ 809.1, p. 25 (cit. on pp. 133, 142).

Morton, T. D. 2015, *isochrones: Stellar model grid package*. Astrophysics Source Code Library. ascl: 1503.010 (cit. on p. 142).

Moudden, Y., J.-F. Cardoso, J.-L. Starck, and J. Delabrouille 2005, Blind Component Separation in Wavelet Space: Application to CMB Analysis. EURASIP Journal on Applied Signal Processing 2005,

484606, p. 484606. DOI: 10.1155/ASP.2005.2437. eprint: astro-ph/0407053 (cit. on p. 10).

Moutou, C. et al. 2009, Planetary transit candidates in the CoRoT initial run: resolving their nature ***. A&A 506.1, pp. 321–336. DOI: 10.1051/0004-6361/200911911 (cit. on pp. 109, 128).

Nason, G. P. and B. W. Silverman 1995, «The Stationary Wavelet Transform and some Statistical Applications.» Wavelets and Statistics. Ed. by Anestis Antoniadis and Georges Oppenheim. New York, NY: Springer New York, pp. 281–299. DOI: 10.1007/978-1-4612-2544-7_17 (cit. on pp. 18, 22–24).

Nason, Guy P 1996, Wavelet shrinkage using cross-validation. Journal of the Royal Statistical Society. Series B (Methodological), pp. 463–479 (cit. on p. 37).

Nefs, S. V., Snellen, I. A. G., and de Mooij, E. J. W. 2012, Minimizing follow-up for space-based transit surveys using full lightcurve analysis. A&A 543, A63. DOI: 10.1051/0004-6361/201118140 (cit. on pp. 110, 128).

Nikolov, N. et al. 2018, An absolute sodium abundance for a cloud-free 'hot Saturn' exoplanet. Nature 557, pp. 526–529. DOI: 10.1038/s41586-018-0101-7. arXiv: 1806.06089 [astro-ph.EP] (cit. on pp. xi, 4).

Núñez, Jorge and Xavier Otazu 1996, Multiresolution image reconstruction using wavelets. Vistas in Astronomy 40.4. Vision modeling and information coding, pp. 555 –562. DOI: https://doi.org/10.1016/S0083-6656(96)00041-4 (cit. on p. 10).

O'Donovan, F. T. et al. 2006, TrES-2: The First Transiting Planet in the Kepler Field. ApJ 651, pp. L61–L64. DOI: 10.1086/509123. eprint: astro-ph/0609335 (cit. on p. 3).

O'Toole, S. J., C. G. Tinney, H. R. A. Jones, R. P. Butler, G. W. Marcy, B. Carter, and J. Bailey 2009, Selection functions in doppler planet searches. MNRAS 392, pp. 641–654. DOI: 10.1111/j.1365-2966.2008.14051.x. arXiv: 0810.1589 (cit. on p. 127).

Ofir, A. 2014, Optimizing the search for transiting planets in long time series. A&A 561, A138, A138. DOI: 10.1051/0004-6361/201220860. arXiv: 1307.7330 [astro-ph.EP] (cit. on p. 58).

Ofir, A. et al. 2010, The SARS algorithm: detrending CoRoT light curves with Sysrem using simultaneous external parameters. MNRAS 404, pp. L99–L103. DOI: 10.1111/j.1745-3933.2010.00843.x. arXiv: 1003.0427 [astro-ph.EP] (cit. on pp. 47, 87).

Otazu, X., M. Ribó, M. Peracaula, J. M. Paredes, and J. Núñez 2002, Detection of superimposed periodic signals using wavelets. MNRAS 333, pp. 365–372. DOI: 10.1046/j.1365-8711.2002.05396.x. eprint: astro-ph/0202107 (cit. on p. 10).

Pätzold, M. et al. 2012, Transiting exoplanets from the CoRoT space mission - XXIII. CoRoT-21b: a doomed large Jupiter around a

faint subgiant star. A&A 545, A6. DOI: 10.1051/0004-6361/201118425 (cit. on pp. 107, 113–119).

Peirce, B. 1852, Criterion for the rejection of doubtful observations. AJ 2, pp. 161–163. DOI: 10.1086/100259 (cit. on p. 43).

Pepper, J. et al. 2007, The Kilodegree Extremely Little Telescope (KELT): A Small Robotic Telescope for Large-Area Synoptic Surveys. PASP 119, pp. 923–935. DOI: 10.1086/521836. arXiv: 0704.0460 (cit. on pp. xi, 3).

Pollacco, D. L. et al. 2006, The WASP Project and the SuperWASP Cameras. PASP 118, pp. 1407–1418. DOI: 10.1086/508556. eprint: astro-ph/0608454 (cit. on pp. xi, 3, 47).

Pont, F., S. Zucker, and D. Queloz 2006, The effect of red noise on planetary transit detection. MNRAS 373, pp. 231–242. DOI: 10.1111/j.1365-2966.2006.11012.x. eprint: astro-ph/0608597 (cit. on p. 33).

Quintana, E. V. et al. 2010, Pixel-level calibration in the Kepler Science Operations Center pipeline. Software and Cyberinfrastructure for Astronomy. Vol. 7740. Proc. SPIE, p. 77401X. DOI: 10.1117/12.857678 (cit. on p. 92).

Ratzloff, J. K., N. M. Law, O. Fors, D. d. Ser, and H. T. Corbett 2016, Evryscope Robotilter automated camera / ccd alignment system. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. Vol. 9908. Proc. SPIE, 99080W. DOI: 10.1117/12.2233384 (cit. on pp. 81, 235).

Rauer, H. 2013, PLATO: PLAnetary Transits and Oscillations of stars. EGU General Assembly Conference Abstracts. Vol. 15. EGU General Assembly Conference Abstracts, EGU2013–2581 (cit. on pp. xi, 4).

Reid, I. N., S. L. Hawley, and J. E. Gizis 1995, The Palomar/MSU Nearby-Star Spectroscopic Survey. I. The Northern M Dwarfs - Bandstrengths and Kinematics. AJ 110, p. 1838. DOI: 10.1086/117655 (cit. on p. 77).

Ricker, G. R. 2014, The Transiting Exoplanet Survey Satellite Mission. Journal of the American Association of Variable Star Observers (JAAVSO) 42, p. 234 (cit. on pp. xi, 4).

Rosenblatt, F. 1971, A Two-Color Photometric Method for Detection of Extra solar Planetary Systems. Icarus 14, pp. 71–93. DOI: 10.1016/0019-1035(71)90103-5 (cit. on p. 2).

Samus', N. N., E. V. Kazarovets, O. V. Durlevich, N. N. Kireeva, and E. N. Pastukhova 2017, General catalogue of variable stars: Version GCVS 5.1. Astronomy Reports 61, pp. 80–88. DOI: 10.1134/S1063772917010085 (cit. on pp. xii, 5, 159).

Santerne, A., R. F. Díaz, J. M. Almenara, A. Lethuillier, M. Deleuil, and C. Moutou 2013, Astrophysical false positives in exoplanet transit surveys: why do we need bright stars ? SF2A-2013: Proceedings of the Annual meeting of the French Society of Astronomy and

Astrophysics. Ed. by L. Cambresy, F. Martins, E. Nuss, and A. Palacios, pp. 555–560. arXiv: 1310.2133 [astro-ph.EP] (cit. on pp. 105, 107).

Scargle, J. D. 1982, Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data. ApJ 263, pp. 835–853. DOI: 10.1086/160554 (cit. on pp. 49, 53, 101).

Sebastian, D., E. W. Guenther, V. Schaffenroth, D. Gandolfi, S. Geier, U. Heber, M. Deleuil, and C. Moutou 2012, Multi-object spectroscopy of stars in the CoRoT fields. I. Early-type stars in the CoRoT-fields IRa01, LRa01, LRa02. A&A 541, A34, A34. DOI: 10.1051/0004-6361/201118032. arXiv: 1203.1143 [astro-ph.EP] (cit. on p. 128).

Shensa, M. J. 1992, The discrete wavelet transform: wedding the a trous and Mallat algorithms. IEEE Transactions on Signal Processing 40.10, pp. 2464–2482. DOI: 10.1109/78.157290 (cit. on pp. 10, 22).

Shupe, David L., Russ R. Laher, Lisa Storrie-Lombardi, Jason Surace, Carl Grillmair, David Levitan, and Branimir Sesar 2012, More flexibility in representing geometric distortion in astronomical images. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. Vol. 8451. Proc. SPIE, pp. 84511M–84511M–13. DOI: 10.1117/12.925460 (cit. on p. 84).

Sigurdsson, S., H. B. Richer, B. M. Hansen, I. H. Stairs, and S. E. Thorsett 2003, A Young White Dwarf Companion to Pulsar B1620-26: Evidence for Early Planet Formation. Science 301, pp. 193–196. DOI: 10.1126/science.1086326. eprint: astro-ph/0307339 (cit. on p. 3).

Sing, D. K. 2010, Stellar limb-darkening coefficients for CoRot and Kepler*. A&A 510, A21. DOI: 10.1051/0004-6361/200913675 (cit. on pp. 116, 123, 134, 137).

Sinukoff, Evan et al. 2016, Eleven Multiplanet Systems from K2 Campaigns 1 and 2 and the Masses of Two Hot Super-Earths. ApJ 827.1, p. 78 (cit. on p. 142).

Skottfelt, J. et al. 2015, The two-colour EMCCD instrument for the Danish 1.54 m telescope and SONG. A&A 574, A54. DOI: 10.1051/0004-6361/201425260 (cit. on p. 134).

Smith, T. C., A. Henden, and D. Terrell 2010, AAVSO Photometric All-Sky Survey Implementation at the Dark Ridge Observatory. Society for Astronomical Sciences Annual Symposium 29, pp. 45–53 (cit. on p. 84).

Starck, J.-L. and F. Murtagh 1994, Image restoration with noise suppression using the wavelet transform. A&A 288, pp. 342–348 (cit. on p. 10).

Starck, J.L., F.D. Murtagh, and A. Bijaoui 1998, *Image Processing and Data Analysis: The Multiscale Approach*. Cambridge University Press (cit. on pp. 10, 38).

Stassun, K. G., K. A. Collins, and B. S. Gaudi 2017, Accurate Empirical Radii and Masses of Planets and Their Host Stars with Gaia Parallaxes. AJ 153, 136, p. 136. DOI: 10.3847/1538-3881/aa5df3. arXiv: 1609.04389 [astro-ph.EP] (cit. on p. 127).

Stein, Charles M. 1981, Estimation of the Mean of a Multivariate Normal Distribution. The Annals of Statistics 9.6, pp. 1135–1151 (cit. on p. 36).

Stetson, P. B. 1996, On the Automatic Determination of Light-Curve Parameters for Cepheid Variables. PASP 108, p. 851. DOI: 10.1086/133808 (cit. on p. 99).

Struve, O. 1952, Proposal for a project of high-precision stellar radial velocity work. The Observatory 72, pp. 199–200 (cit. on p. 2).

Szatmary, K., J. Vinko, and J. Gal 1994, Application of wavelet analysis in variable star research. I. Properties of the wavelet map of simulated variable star light curves. A&AS 108 (cit. on p. 10).

Tamuz, O., T. Mazeh, and S. Zucker 2005, Correcting systematic effects in a large set of photometric light curves. MNRAS 356, pp. 1466–1470. DOI: 10.1111/j.1365-2966.2004.08585.x. eprint: astro-ph/0502056 (cit. on pp. 6, 47, 87).

Taubman, D. S. and M. W. Marcellin 2002, JPEG2000: standard for interactive imaging. Proceedings of the IEEE 90.8, pp. 1336–1357. DOI: 10.1109/JPROC.2002.800725 (cit. on p. 10).

Thompson S. E. Fraquelli D., Van Cleve J. E. and Caldwell D. A. 2016, *Kepler Archive Manual*. KDMC-10008-006. Moffett Field: NASA Ames Research Center (cit. on p. 90).

Thompson, S. E. et al. 2012, A Class of Eccentric Binaries with Dynamic Tidal Distortions Discovered with Kepler. ApJ 753, 86, p. 86. DOI: 10.1088/0004-637X/753/1/86. arXiv: 1203.6115 [astro-ph.SR] (cit. on p. 129).

Tonry, J. L. 2011, An Early Warning System for Asteroid Impact. PASP 123, p. 58. DOI: 10.1086/657997. arXiv: 1011.1028 [astro-ph.IM] (cit. on p. 47).

Torrence, Christopher and Gilbert P Compo 1998, A practical guide to wavelet analysis. Bulletin of the American Meteorological society 79.1, pp. 61–78 (cit. on p. 28).

Tukey, John W. 1949, Comparing Individual Means in the Analysis of Variance. Biometrics 5.2, pp. 99–114 (cit. on p. 43).

Twicken, J. D., H. Chandrasekaran, J. M. Jenkins, J. P. Gunter, F. Girouard, and T. C. Klaus 2010, Presearch data conditioning in the Kepler Science Operations Center pipeline. Software and Cyberinfrastructure for Astronomy. Vol. 7740. Proc. SPIE, 77401U. DOI: 10.1117/12.856798 (cit. on p. 93).

Udalski, A., K. Zebrun, M. Szymanski, M. Kubiak, I. Soszynski, O. Szewczyk, L. Wyrzykowski, and G. Pietrzynski 2002, The Optical Gravitational Lensing Experiment. Search for Planetary and Low-Luminosity Object Transits in the Galactic Disk. Results of 2001 Campaign – Supplement. Acta Astron. 52, pp. 115–128. eprint: astro-ph/0207133 (cit. on p. 2).

Vanderburg, A. 2014, Reduced Light Curves from Campaign 0 of the K2 Mission. ArXiv e-prints. arXiv: 1412.1827 [astro-ph.IM] (cit. on p. 91).

Vanderburg, A. and J. A. Johnson 2014, A Technique for Extracting Highly Precise Photometry for the Two-Wheeled Kepler Mission. PASP 126, p. 948. DOI: 10.1086/678764. arXiv: 1408.3853 [astro-ph.IM] (cit. on pp. 91, 93).

Vanderburg, A. et al. 2015, Characterizing K2 Planet Discoveries: A Super-Earth Transiting the Bright K Dwarf HIP 116454. ApJ 800, 59, p. 59. DOI: 10.1088/0004-637X/800/1/59. arXiv: 1412.5674 [astro-ph.EP] (cit. on p. 92).

Vanderburg, Andrew, David W. Latham, Lars A. Buchhave, Allyson Bieryla, Perry Berlind, Michael L. Calkins, Gilbert A. Esquerdo, Sophie Welsh, and John Asher Johnson 2016, Planetary Candidates from the First Year of the K2 Mission. ApJS 222.1, p. 14 (cit. on p. 142).

Vennes, Stéphane, G. D. Schmidt, L. Ferrario, D. J. Christian, D. T. Wickramasinghe, and A. Kawka 2003, A Multiwavelength Study of the High-Field Magnetic White Dwarf EUVE J0317–85.5 (=RE J0317–853). ApJ 593.2, p. 1040 (cit. on p. 101).

Voss, H. 2006, «Developing a ground-based search system for transits of extrasolar planets.» PhD thesis. Technischen Universität Berlin (cit. on pp. 81, 84, 95).

Waldmann, I. P. 2014, On Signals Faint and Sparse: The ACICA Algorithm for Blind De-trending of Exoplanetary Transits with Low Signal-to-noise. The Astrophysical Journal 780.1, p. 23 (cit. on pp. 11, 52).

Weyrich, Norman and Gregory T Warhola 1995, «De-noising using wavelets and cross validation.» Approximation Theory, Wavelets and Applications. Springer, pp. 523–532 (cit. on p. 37).

Wheatley, P. J. et al. 2013, The Next Generation Transit Survey (NGTS). European Physical Journal Web of Conferences. Vol. 47. European Physical Journal Web of Conferences, p. 13002. DOI: 10.1051/epjconf/20134713002. arXiv: 1302.6592 [astro-ph.EP] (cit. on pp. xi, 4, 47).

Wolszczan, A. 1994, Confirmation of Earth-Mass Planets Orbiting the Millisecond Pulsar PSR B1257+12. Science 264, pp. 538–542. DOI: 10.1126/science.264.5158.538 (cit. on pp. xi, 1).

Wolszczan, A. and D. A. Frail 1992, A planetary system around the millisecond pulsar PSR1257 + 12. Nature 355, pp. 145–147. DOI: 10.1038/355145a0 (cit. on pp. xi, 1).

Woźniak, P. R. et al. 2004, Northern Sky Variability Survey: Public Data Release. AJ 127, pp. 2436–2449. DOI: 10.1086/382719. eprint: astro-ph/0401217 (cit. on p. 95).

Yee, S. W., E. A. Petigura, and K. von Braun 2017, Precision Stellar Characterization of FGKM Stars using an Empirical Spectral Library. ApJ 836, 77, p. 77. DOI: 10.3847/1538-4357/836/1/77. arXiv: 1701.00922 [astro-ph.SR] (cit. on pp. 133, 142).

Zacharias, N., C. T. Finch, T. M. Girard, A. Henden, J. L. Bartlett, D. G. Monet, and M. I. Zacharias 2012, VizieR Online Data Catalog: UCAC4 Catalogue (Zacharias+, 2012). VizieR Online Data Catalog, I/322A, I/322A (cit. on pp. 109, 112, 128, 133).

Ziegler, C., N. M. Law, C. Baranec, T. Morton, R. Riddle, D. Atkinson, and L. Nofi 2016, The Robo-AO KOI survey: laser adaptive optics imaging of every Kepler exoplanet candidate. Adaptive Optics Systems V. Vol. 9909. Proc. SPIE, 99095U. DOI: 10.1117/12.2231185. arXiv: 1608.00575 [astro-ph.EP] (cit. on p. 107).

de Freitas, D. B., I. d. C. Leão, B. L. Canto Martins, and J. R. De Medeiros 2010, Wavelet analysis of stellar rotation and other periodicities. ArXiv e-prints. arXiv: 1009.5090 [astro-ph.SR] (cit. on pp. 10, 28).

del Ser, D., O. Fors, and J. Núñez 2018, TFAW: wavelet-based signal reconstruction to reduce photometric noise in time-domain surveys. Accepted in A&A. arXiv: 1702.06547v2 [astro-ph.IM] (cit. on pp. 11, 51, 235).

del Ser, D., O. Fors, J. Núñez, H. Voss, A. Rosich, and V. Kouprianov 2015, Detrending Algorithms in Large Time Series: Application to TFRM-PSES Data. Living Together: Planets, Host Stars and Binaries. Ed. by S. M. Rucinski, G. Torres, and M. Zejda. Vol. 496. Astronomical Society of the Pacific Conference Series, p. 301. arXiv: 1411.5320 [astro-ph.IM] (cit. on p. 236).

van de Kamp, P. 1963, Astrometric study of Barnard's star from plates taken with the 24-inch Sproul refractor. AJ 68, pp. 515–521. DOI: 10.1086/109001 (cit. on p. 1).

PUBLICATIONS

Directly related to this thesis

1. D. del Ser, O. Fors, and J. Núñez 2018, TFAW: wavelet-based signal reconstruction to reduce photometric noise in time-domain surveys. Accepted in A&A. arXiv: 1702.06547v2 [astro-ph.IM]

Other contributions

1. W. S. Howard et al. 2018c, The First Naked-eye Superflare Detected from Proxima Centauri. ApJ 860, L30, p. L30. DOI: 10.3847/2041-8213/aacaf3. arXiv: 1804.02001 [astro-ph.EP]

2. W. S. Howard et al. 2018b, Stellar activity for every TESS star in the Southern sky. American Astronomical Society Meeting Abstracts #231. Vol. 231. American Astronomical Society Meeting Abstracts, p. 310.03

3. W. S. Howard et al. 2018a, Evryscope Detection of the First Proxima Superflare: Impacts on the Atmosphere and Habitability of Proxima b. LPI Contributions 2065, 2039, p. 2039

4. H. Corbett et al. 2018, Pre-Discovery Detection of ASASSN-18fv by Evryscope. The Astronomer's Telegram 11467

5. O. Fors et al. 2017, Engaging AAVSO members in Stellar Astrophysics Follow-up from The Evryscope Data (Abstract). Journal of the American Association of Variable Star Observers (JAAVSO) 45, p. 129

6. W. Howard et al. 2017, EvryFlare: Flare rates and intensities for every 10 < g < 15 solar-type and red dwarf star in the Southern sky. Radio Exploration of Planetary Habitability (AASTCS5). Vol. 49, p. 202.05

7. N. M. Law et al. 2016, The Evryscope: design and performance of the first full-sky gigapixel-scale telescope. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. Vol. 9906. Proc. SPIE, p. 99061M. DOI: 10.1117/12.2233349

8. J. K. Ratzloff et al. 2016, Evryscope Robotilter automated camera / ccd alignment system. Society of Photo-Optical

Instrumentation Engineers (SPIE) Conference Series. Vol. 9908. Proc. SPIE, 99080W. DOI: 10.1117/12.2233384

9. D. del Ser et al. 2015, Detrending Algorithms in Large Time Series: Application to TFRM-PSES Data. Living Together: Planets, Host Stars and Binaries. Ed. by S. M. Rucinski, G. Torres, and M. Zejda. Vol. 496. Astronomical Society of the Pacific Conference Series, p. 301. arXiv: 1411.5320 [astro-ph.IM]

10. O. Fors et al. 2015b, The Evryscope and extrasolar planets. IAU General Assembly 22, 2258237, p. 2258237

11. O. Fors et al. 2015a, Evrystats for Evryplanets: planets from the first all-sky gigapixel scale telescope. IAU General Assembly 22, 2253010, p. 2253010

12. N. M. Law et al. 2015, Evryscope Science: Exploring the Potential of All-Sky Gigapixel-Scale Telescopes. PASP 127, pp. 234–249. DOI: 10.1086/680521. arXiv: 1501.03162 [astro-ph.IM]

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*".

FUNDING ACKNOWLEDGMENTS

The author acknowledges financial support from the RACAB and the Universitat de Barcelona. The TFRM project is supported, in part, by the Departament d'Empresa i Coneixement de la Generalitat de Catalunya. We also acknowledge support by the Spanish Ministerio de Economía y Competitividad (MINECO/ FEDER, UE) under grants AYA2013-47447-C3-1-P, AYA2016-76012-C3-1-P, MDM-2014-0369 of ICCUB (Unidad de Excelencia 'María de Maeztu'), and the Catalan DEC grant 2014 SGR 86. The Evryscope project is supported by the National Science Foundation CAREER grant 1555175, and the Research Corporation Scialog grants 23782 and 23822.