UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE POSGRADO EN ASTROFÍSICA

INSTITUTO DE RADIOASTRONOMÍA Y ASTROFÍSICA

ANALYSIS OF MULTI-WAVELENGTH LIGHT CURVES
OF NEARBY EVOLVED STARS SURVEY TARGETS

T   E   S   I   S

QUE PARA OPTAR POR EL GRADO DE:

MAESTRO EN CIENCIAS (ASTROFÍSICA)

PRESENTA:

DIEGO ALEJANDRO VASQUEZ TORRES

DIRECTORES DE TESIS

DR. SUNDAR SRINIVASAN, IRyA

DR. LUIS HENRY QUIROGA-NUÑEZ, FLORIDA TECH

MIEMBROSDEL COMITÉ TUTOR

DR. BERNARDO CERVANTES SODI, IRyA

DR. STANLEY EUGENE KURTZ, IRyA

DRA. VERÓNICA LORA CASTELLANOS, ICN

MORELIA, MICHOACÁN, MÉXICO. NOVIEMBRE 2023

Las estrellas son hermosas, por una flor que no se ve...

The stars are beautiful, because of a flower that cannot be seen...

Les étoiles sont belles, à cause d'une fleur que l'on ne voit pas...

<div align="right">

The Little Prince

Antoine de Saint-Exupéry

</div>

Dedication

This thesis is dedicated to my parents, Carlos and Ofelia, and my brother, Andres, whom I love deeply. Despite the distance during my master's studies and the writing of this thesis, their support through calls has meant a lot. Thank you for always being there.

Esta tesis está dedicada a mis padres, Carlos y Ofelia, y a mi hermano, Andres, a quienes amo profundamente. A pesar de la distancia durante mi maestría y la escritura de esta tesis, su apoyo a través de llamadas ha significado mucho. Gracias por estar siempre ahí.

Personal Acknowledgments

I want to thank all the people who accompanied me during my master's in Mexico, the friends I made while being here, and who opened their homes to me, making me a part of their family. I am grateful to the researchers at IRyA who supported me in my educational journey and dedicated their time to answer my questions.

I am immensely grateful to Dr. Sundar Srinivasan for the support he provided during my master's, despite the language barrier. His patience made the writing of this thesis possible, and I also appreciate his willingness to assist me with personal matters to ensure my physical and emotional well-being.

Furthermore, I thank my psychologist, Cesar, who supported and listened to me during difficult moments of loneliness, sadness, or disappointment and always encouraged me with a smile at the end of each session.

Finally, I want to express my gratitude to the friends who now mean so much more in my life, in this case, Janis, Maritza, Karol, Natalia, Marco, and Sandino. Thank you for being there with me throughout this time.

Contents

List of Figures

Resumen

Las estrellas de la rama gigante asintótica (AGB, por sus siglas en inglés) son notables por sus temperaturas superficiales relativamente frías (aproximadamente 3000 K), luminosidades que van desde 10^3 hasta aproximadamente 5×10^4 L$_\odot$ y tasas significativas de pérdida de masa que oscilan entre 10^{-8} y 10^{-4} M$_\odot$ año^{-1}. Estas estrellas desempeñan un papel crucial en el enriquecimiento químico del medio interestelar, lo que hace esencial comprender los procesos subyacentes a su pérdida de masa para el estudio de la evolución química galáctica. Un fenómeno clave en las estrellas AGB es la pulsación, que se manifiesta como cambios en la luminosidad con el tiempo, con períodos que van desde 30 hasta 2000 días. Estas llamadas variables de período largo pueden clasificarse como semirregulares y Miras según la amplitud y el espectro de potencia de sus curvas de luz.

Los estudios ópticos rastrean la variabilidad de la fotosfera, mientras que los datos infrarrojos revelan los cambios correspondientes en la envoltura de polvo debido a la variación de la luminosidad estelar. Por lo tanto, se requiere un estudio sistemático de las curvas de luz de múltiples longitudes de onda para descubrir la conexión entre estos regímenes interrelacionados. Estudios previos han analizado la amplitud como función de la longitud de onda efectiva en filtros fotométricos para muestras pequeñas, como el trabajo de Smith et al., 2002, y algunos han examinado la variabilidad infrarroja en muestras grandes con muestreo de tiempo disperso, como Riebel et

al., 2015. Sin embargo, aún no se ha realizado una investigación sistemática de la dependencia de la longitud de onda de los períodos y amplitudes para muestras grandes de estrellas AGB. Esto es especialmente cierto para las estrellas AGB más cercanas, para las cuales existe un creciente archivo de curvas de luz bien muestreadas debido a instalaciones de encuesta en longitudes de onda ópticas e infrarrojas.

Esta tesis presenta la recopilación y análisis automatizados de curvas de luz en diferentes bandas fotométricas de múltiples encuestas para fuentes en la Encuesta de Estrellas Evolucionadas Cercanas (NESS; Scicluna et al. 2022), que incluye 852 estrellas evolucionadas en el Vecindario Solar dentro de un radio de 3 kpc. De la muestra original de 852 estrellas NESS, se encontró que 712 estrellas tenían al menos una curva de luz, con un total de 5114 curvas de luz. Basándose en criterios de calidad de las curvas de luz, 674 estrellas tenían al menos una curva de luz de alta calidad, lo que resultó en un total de 3634 curvas de luz para su procesamiento. Se extrajeron períodos fundamentales, magnitudes medias y amplitudes de las curvas de luz utilizando análisis de periodogramas y construcción de modelos utilizando el código de Procesos Gaussianos para Inferencia de Variabilidad Multilongitud de Onda en Python (PGMUVI; Scicluna et al. 2023).

Se llevaron a cabo comparaciones con relaciones existentes en la literatura científica con el objetivo de replicar estos resultados. Aunque el algoritmo PGMUVI aún está en proceso de implementar nuevas funciones, las tendencias generales vistas en la literatura en cuanto a las relaciones entre el período y el color, la tasa de producción de polvo o la tasa de pérdida de masa se reprodujeron con éxito para las estrellas Mira en NESS con determinaciones de período precisas (mejores que el 30%). Se exploró

el comportamiento de las amplitudes de las curvas de luz en relación con la longitud de onda y la banda fotométrica, y se replicaron con éxito resultados previamente informados en la literatura utilizando las mismas bases de datos, al menos para el infrarrojo cercano y medio. Se presentaron estrategias para mejorar el algoritmo PGMUVI y proporcionar información útil y de alta calidad en futuras investigaciones.

El trabajo descrito en esta tesis es un gran avance hacia la recopilación y el análisis automatizados de datos de series temporales de encuestas a gran escala y, por lo tanto, será relevante para encuestas de dominio temporal actuales y futuras, como Gaia y la Encuesta de Legado del Espacio y el Tiempo (LSST). Los datos, el código y los resultados de este trabajo se publicarán junto con un próximo artículo (Vasquez-Torres et al.) y estarán completamente disponibles para la comunidad científica. Complementarán las observaciones NESS existentes y futuras y los resultados de estas, proporcionando una comprensión más profunda de las estrellas AGB y su relevancia en el Vecindario Solar.

Abstract

Asymptotic giant branch (AGB) stars are notable for their relatively cool surface temperatures ($\approx$3000 K), luminosities ranging from 10^3 to $\sim 5 \times 10^4$ L$_\odot$, and significant mass-loss rates between 10^{-8} and 10^{-4} M$_\odot$ yr^{-1}. These stars play a crucial role in the chemical enrichment of the interstellar medium, making it essential to understand the processes underlying their mass loss for the study of galactic chemical evolution. A key phenomenon in AGB stars is pulsation, which manifests as changes in brightness over time, with periods ranging from 30 to 2000 days. These so-called long-period variables can be classified as semi-regulars-type and Miras-type based on the amplitude and power spectrum of their light curves.

Optical studies trace the variability of the photosphere, while infrared data reveals the corresponding changes in the dust shell due to the varying stellar luminosity. A systematic study of multiwavelength light curves is therefore required to uncover the connection between these interrelated regimes. Previous studies have analyzed the amplitude as a function of effective wavelength in photometric filters for small samples and some have examined infrared variability in large samples with sparse time sampling. However, a systematic investigation of the wavelength dependence of periods and amplitudes has not yet been conducted for large samples of AGB stars. This is especially true for the nearest AGB stars, for which a growing archive of well-sampled light curves exists due to survey facilities at optical and infrared wavelengths.

This thesis presents the automated collection and analysis of light curves in different photometric bands from multiple surveys for sources in the Nearby Evolved Stars Survey (NESS), which includes 852 evolved stars in the Solar Neighborhood within a radius of 3 kpc. From the original sample of 852 NESS stars, 712 stars were found to have at least one light curve, totaling 5114 light curves. Based on light curve quality criteria, 674 stars had at least one high-quality light curve, resulting in a total of 3634 light curves for processing. Fundamental periods, mean magnitudes, and amplitudes are extracted from light curves, using periodogram analysis and model construction using the Python Gaussian processes for MUltiwavelength Variability Inference (PGMUVI; Scicluna et al. 2023) code.

Comparisons to existing relationships in the scientific literature were carried out with the aim of replicating these results. Although the PGMUVI algorithm is still in the process of implementing new functions, the general trends seen in the literature for relationships between the period and color, dust-production rate, or mass-loss rate were successfully reproduced for Mira stars in NESS with precise period determinations (better than 30%). The behavior of light curve amplitudes in relation to wavelength and photometric band was explored, and previously reported results in the literature were successfully replicated using the same databases, at least for the near- and mid-infrared. Strategies were presented for enhancing the PGMUVI algorithm to provide useful and high-quality information in future research.

The work described in this thesis is a huge step towards the automated collection and analysis of time-series data from large-scale surveys, and will therefore be relevant to

current and upcoming time-domain surveys such as Gaia and the Legacy Survey of Space and Time (LSST). The data, code, and results of this work will be published along with an upcoming paper (Vasquez Torres et al.), and will be made fully available to the scientific community. They will complement existing and future NESS observations and results from these and provide a deeper understanding of AGB stars and their relevance in the Solar Neighbourhood.

1

Introduction

Stars in the asymptotic giant branch (AGB) phase are fascinating objects in the Universe due to their complex physical modeling and the numerous mysteries that remain to be solved for current astrophysics. Multiple physical processes occur both inside and outside the star, such as thermal pulses that result from nucleosynthesis in different layers of the star. In addition, the third dredge-up (3DU) occurs in the convective region, which takes material from the interior of stars and brings it to the surface, increasing the star's surface metallicity. Furthermore, the material outflows present in these stars' atmospheres transport the enriched matter. As a consequence, these stars have a history of mass loss in their atmospheres (see Figure 1.1), where it cools and allows the formation of regions with gas and dust. Another phenomenon linked to high luminosity is pulsations or changes in the star's brightness over time, which may result from the processing of stellar light by the circumstellar dust-forming region.

The main objective of this thesis is to study the variability of nearby AGB stars. This
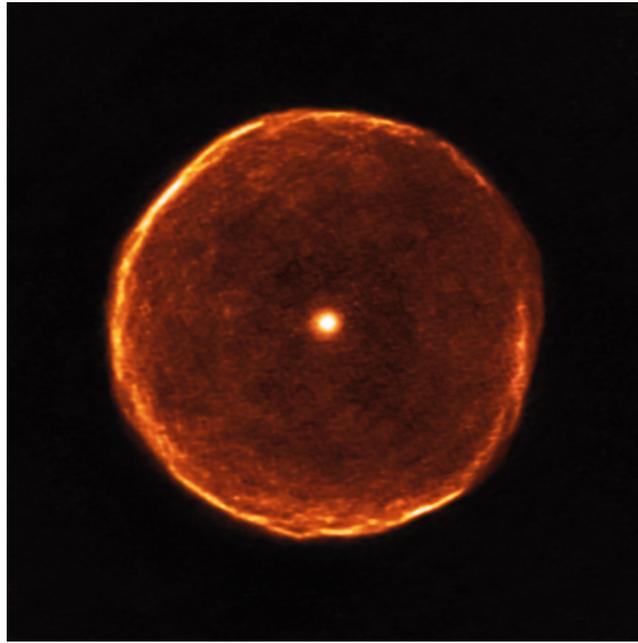
Figure 1.1: U Antliae is an AGB star that displays a history of mass loss. The detached shell, which signifies the end of the mass loss process as the ejection of matter by stellar winds has ceased, can be observed. This image was captured by the ALMA interferometer in Band 6 (230 GHz) and detects the vibrational emission line of the ^{13}CO j $= 2 - 1$ isotopologue. The shell has a radius of 0.05 pc [for more details Kerschbaum et al., 2017]. Image courtesy of: ALMA (ESO/NAOJ/NRAO)/F. Kerschbaum. European Southern Observatory [2017].

chapter presents relevant concepts about the nature of AGB sources, the role of dust, history of observations of AGB stars, the state of the art, a description of the main goals of the Nearby Evolved Stars Survey, and finally, the goals and scope of this thesis.

1.1 Evolution and structure of AGB stars

1.1.1 Evolution of low- and intermediate-mass stars from the main sequence to the pre-AGB phase.

Stars spend approximately 90% of their lifetime on the Main Sequence. This phase is character-

Evolution of low- and intermediate-mass stars from the main sequence to the pre-AGB phase.

ized by the fusion of hydrogen into helium in the star's core, a crucial process that averts the collapse of the star by providing a primary source of energy. This fusion process is primarily accomplished through two distinct mechanisms: for stars with lower masses, the proton-proton chain (pp) is the dominant fusion pathway, while in stars with masses exceeding $1.5 M_\odot$, the Carbon-Nitrogen-Oxygen cycle (CNO) becomes the prevailing fusion mechanism [Karttunen et al., 2017].

Figures 1.2 are Hertzsprung-Russell (HR) diagrams that show the theoretical stellar evolution for stars with masses of 1 and 5 $M_\odot$, respectively. The 1 $M_\odot$ star represents low-mass stars, while the 5 $M_\odot$ star represents intermediate-mass stars. It can be observed that both stars progress through the same evolutionary stages: the main sequence, the subgiant phase, the red giant phase, before climbing the AGB. The last stage gets its name from its location on the HR diagram, asymptotic to the red giant branch. It is noteworthy that only stars with masses in the range of 0.8 $M_\odot$ to 8.0 $M_\odot$ are capable of reaching the AGB phase.

As a star depletes the hydrogen in its core, it undergoes an expansion in its radius, resulting in an increase in its luminosity. However, there comes a point where the star can no longer generate enough energy through hydrogen fusion to maintain hydrostatic equilibrium in its core. At this point, the star departs from the main sequence. This event marks the beginning of the subgiant phase, during which the star expands and its temperature decreases, although its luminosity remains relatively constant compared to low-mass stars. For intermediate-mass stars, the decrease in luminosity relative to the end of the main sequence is relatively small. This transition occurs gradually over hundreds of years, which explains why stars are not observed in the region known as the "Hertzsprung gap" on an HR diagram of a globular cluster or the Galaxy.
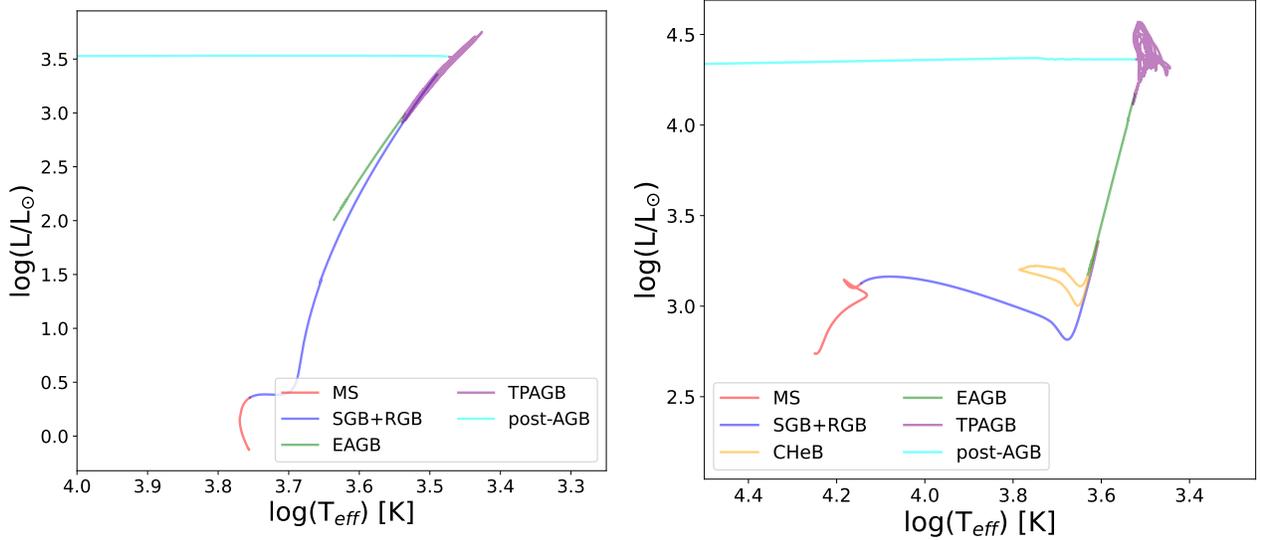
Figure 1.2: In these stellar evolution diagrams, we can see the different stages of a star's life cycle: Main Sequence (MS), Subgiant Branch (SGB), Red Giant Branch (RGB), Core Helium Burning (CHeB), Early AGB (EAGB), Thermally Pulsing AGB (TPAGB), and Post-AGB. On the left, the evolution of stars with 1 M⊙. On the right, the evolution of stars with 5 M⊙, both with the same solar metallicity. Each stage of evolution is differentiated by colors. These paths through the HR diagrams are defined by the star's metallicity and its initial mass on the main sequence. Figures generated using data from MIST[1] models.

For stars with a mass lower than 2 M⊙, the helium core remains inert while accumulating more helium from a surrounding region, where hydrogen burning still fuels it. However, once certain levels of density and temperature are reached, the fusion of helium into carbon is activated through the triple alpha process. This process generates a significant release of a large amount of energy in a short interval of time in the core of the star, known as a "helium flash". The helium flash occurs at the beginning of the red giant branch. After the flash, helium is being burned into carbon in the interior of the star, the structure of the star reorganizes, and energy transport is primarily carried out through convection. At this stage, the star evolved for millions of years as a red giant, expanding, becoming brighter, and decreasing its temperature. At the end of the red giant phase, the star has a degenerate core composed of oxygen and carbon. Following all these physical processes, the star has now entered the AGB phase. Stars with

masses > 2 M$_\odot$ do not experience the helium flash. Instead, the transition from hydrogen burning to helium burning occurs more smoothly.

Depending on the initial mass, one or two episodes of a phenomenon known as "dredge-up" can occur during pre-AGB stellar evolution. This phenomenon involves the transport of chemically enriched material, with an increase in metallicity, from the burning zones to the surface of the star through thermal convection. This dredge-up results in changes in the observed chemical abundances on the stellar surface. Initially, after the burning of hydrogen, the first dredge-up episode occurs. The second dredge-up episode takes place in stars with an initial mass between 4 and 8 solar masses [Habing & Olofsson, 2004]. Additionally, there is a third dredge-up episode during the AGB phase, which will be discussed later.

1.1.2 Internal structure of an AGB star

During the AGB stage the star experiences an increase in its radius and luminosity, while its effective temperature decreases (see Figure 1.2). As the star ascends the AGB branch, helium accumulates around its degenerate carbon-oxygen core. Under suitable conditions, such as high density, helium burning occurs in a shell through the triple-alpha process. This process manifests as a "helium flash" and initiates combustion in a helium layer, further fueling the core with more carbon-oxigen.

As the star depletes its helium, a hydrogen burning region forms, generating more helium that accumulates on the core. Then, when density and pressure are sufficiently high, helium burning initiates again. This cycle of helium and subsequent hydrogen burning is known as a thermal pulses (TP). These thermal pulses cause periodic expansions and contractions of the star, which are also reflected in periodic changes in its effective temperature and luminosity

as seen in Figure 1.3. However, the superficial variation is nowhere near as dramatic as those in the interior due to thermal pulses; moreover, the period of variations is markedly different; while thermal pulses happen over hundreds of thousands of years, surface pulsations vary over tens to hundreds of days.
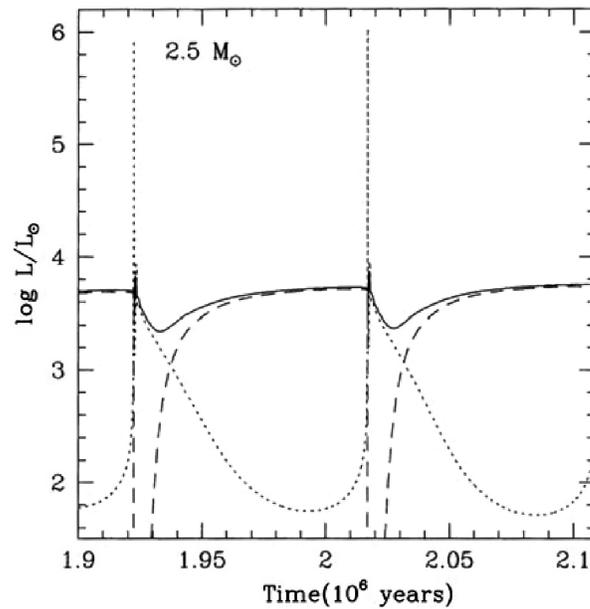
Figure 1.3: This is an example of what happens in the thermal pulse phase, showing the variation in the surface luminosity of the star (solid line). At the beginning of the pulse, there is a significant increase in the luminosity generated by helium (dotted line) in a short period of time, while in the case of hydrogen luminosity (dashed line), the opposite effect is observed. Adapted from Figure 2.9 of Lattanzio & Wood [2004].

In Figure 1.4, a diagram illustrating the different regions of the star in mass coordinates can be observed. These regions represent different states, such as the helium and hydrogen-burning zones. Due to the thermal pulses and the convective effect, these regions shift, and enriched material is transported out of the burning zones, eventually reaching the star's surface.

Internal structure of an AGB star
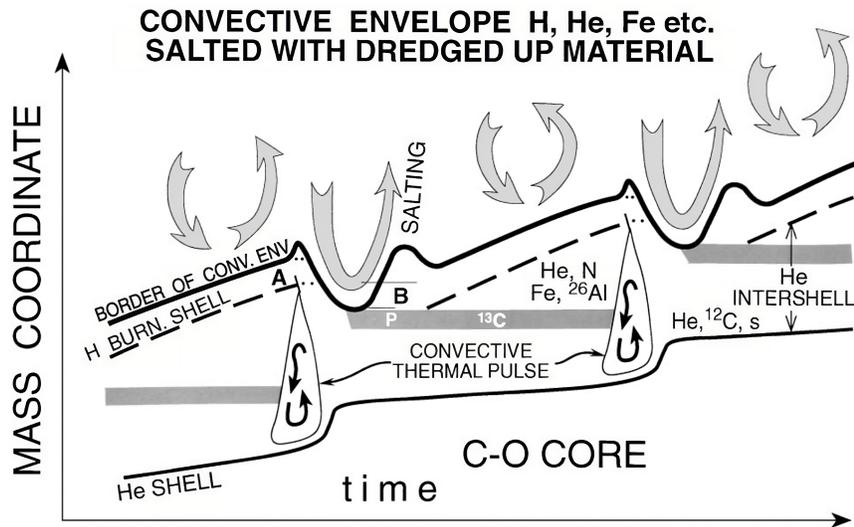
Figure 1.4: This image qualitatively shows the different mixing processes that occur between thermal pulses. As time progresses, the core increases in mass, as do the shells of helium and hydrogen burning. Figure 5 from Busso et al. [1999].

However, in the case of higher mass AGB stars, a nucleosynthesis process occurs through neutron capture (called the *s*-process), which takes place between the helium and hydrogen burning layers. An example of the generated elements is technetium, which serves as an observational indicator showing that the star is in the AGB phase. Additionally, there is a hot bottom burning that occurs in the more massive AGB stars. This is because the lower region of the convective layer reaches temperatures [40×10^6 K; García-Hernández et al., 2013] high enough for ^{12}C to capture neutrons and turn into ^{14}N before ^{12}C is transported to the surface of the star due to convection [Busso et al., 1999].

Figure 1.5 displays the internal and external structure of an AGB star, indicating the size of each region in stellar radii units. Additionally, it illustrates the physical and chemical processes occurring in each region. Convection around the core is particularly noteworthy as it is responsible for enriching the star's surface with carbon and *s*-process elements from the

13

Internal structure of an AGB star

interior. This enrichment is reflected in an increased relative abundance of carbon compared to other chemical elements present in the star, such as oxygen. Based on these relative abundance relationships, a chemical classification can be derived, which will be discussed later on.
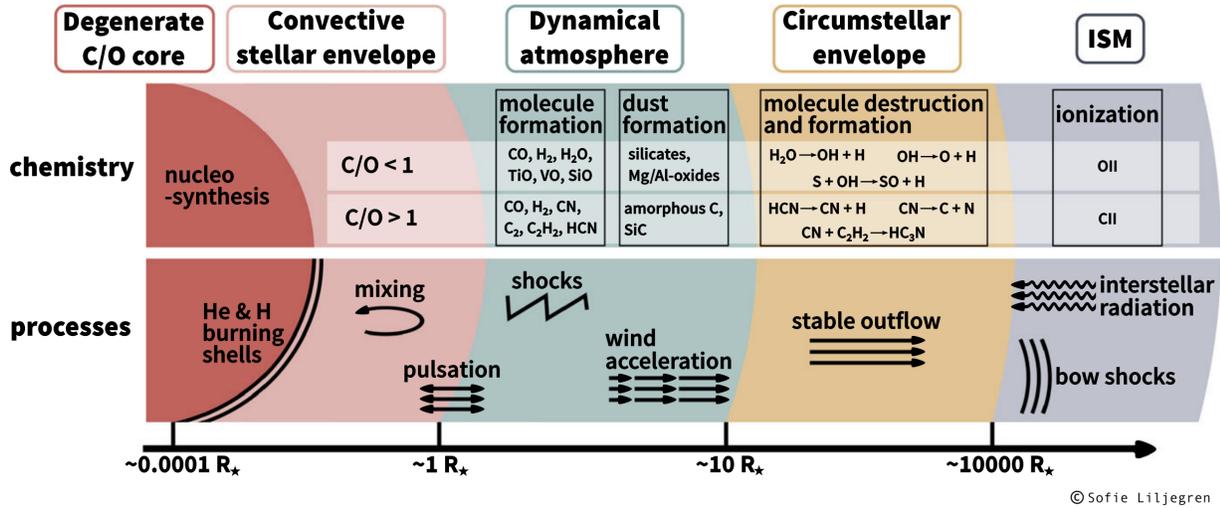
Figure 1.5: Structure of an AGB star, in the image various physical and chemical processes are observed to occur in different regions with respect to the star's radius. Figure 1 from Höfner & Olofsson [2018].

1.1.3   Superficial and circumstellar structure AGB star

The radial pulsations resulting from the release of energy from regions of partial hydrogen ionization below the optical photosphere of AGB stars (detailed in Section 1.2.2) cause the material in this region to levitate, leading to the formation of matter outflows that transport said material through the stellar atmosphere. This process enriches the atmosphere with metals that were previously extracted from the helium and hydrogen burning zones through stellar convection. These mechanisms generate temperature fluctuations that favor the formation of molecules ($T \lesssim 3000$ K) and dust [$T \lesssim 1800$ K; see, e.g., Gail et al., 2013] in that region. Figure 1.5 shows the regions where shocks, stable winds, and bow shocks with the interstellar medium occur.

AGB stars can be classified according to the relative abundance of carbon and oxygen present in their atmospheres. For example, if C/O < 1, they are classified as M-type stars (or in the context of AGB stars classified as O-type depending on the bibliographic source). Those with C/O ∼ 1 are classified as S-type stars, and those with C/O > 1 are classified as C-type stars, also known as carbon stars and [Höfner & Olofsson, 2018 and Iben & Renzini, 1983]. The chemical class can be determined on the basis of molecular absorption bands in the optical and near-infrared part of the spectrum. For example, M-type stars exhibit strong absorption from TiO and VO, while S-type stars show absorption from ZrO, and C-type stars have CN lines [Chen et al., 2022 and Palmer et al., 1982]. Figure 1.6 presents a color-magnitude diagram with the K_s and J bands from the 2 micron All-Sky Survey [2MASS; Skrutskie et al., 2006], where the different chemical populations of AGB stars in a sample from the Milky Way satellite, the Small Magellanic Cloud (SMC), can be distinguished. The figure shows the populations of different AGB stars in the SMC. Since they are statistically at the same distance, it is easy to separate the populations while ignoring the effects of extinction.

1.2   AGB variability

In this section, the most relevant concepts for this thesis are presented, as it discusses the variability of AGB stars and their classification, and how this is detected in their spectral energy distribution. En the study of light curves of variable stars, one of the most powerful tools available to astronomers is the representation of data in Fourier space through power spectra and periodograms. Using Fourier space provides a significant advantage due to its ability to transform variability in the time domain into the frequency domain. This transformation allows researchers to discern the underlying periodic patterns and identify the dominant frequencies,
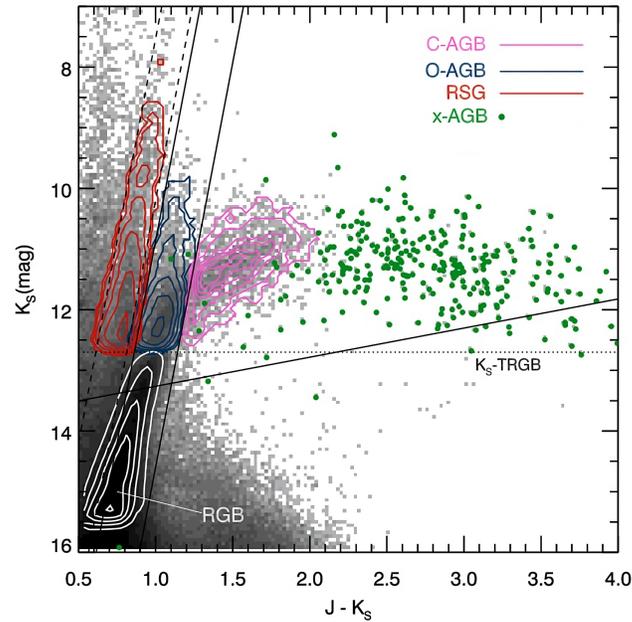
Figure 1.6: In the color-magnitude diagram, one can observe how the different chemical types and evolutionary states of evolved stars (red giant branch or RGB stars, red supergiants or RSGs, and AGB stars) in the Small Magellanic Cloud cluster together. AGB stars are further sub-classified into O–rich (M) and C–rich based on their chemical types. The dustiest AGB population (extreme AGB or x-AGB) contains both types. The effect of dust content in the stellar atmosphere causes these populations to become redder, thus locating them on the right side of the diagram. Figure 5 from Boyer et al. [2011].

uncovering hidden intrinsic variability within the data. Visualizing variability through power spectra and periodograms facilitates the extraction of essential frequency information, simplifies the identification of periodic components, and enables the comprehensive characterization of unique behaviors of variable stars, paving the way to a deeper understanding of their physical properties and evolutionary processes.

1.2.1 Luminosity in the optical and infrared and its variability

Stars in the AGB phase are characterized by having effective temperatures in the range $(2200 < T_{\text{eff}} < 4500$ K); using Wein 's Displacement Law, their photospheric emission therefore peaks at near-infrared wavelengths. This is reflected in the HR diagram, where they are classified

as K or M spectral type stars. Therefore, broad-band filters covering both the optical and infrared wavelengths enable a comprehensive study of the spectral energy distribution (SED) of AGB stars.

The flux density of these stars varies over time, which means that by taking measurements on different days or months, significant changes in the apparent magnitude of the star can be detected. By collecting magnitude information chronologically, it is possible to construct a time series, commonly known as a light curve. The qualitative properties of these time series are of great interest in the study of variable stars, as they can provide insights into parameters such as pulsation periods, pulse amplitudes, or average apparent magnitude depending on the photometric filter used for detection. It is known that the pulsation periods of AGB stars correlate with their luminosities [e.g., Lebzelter et al., 2019]. Therefore, it becomes conceivable that by measuring the average flux of the star and determining its pulsation period, it may be possible to estimate the distance to the star through the period-luminosity relationship and distance modeling.

Without delving into the pulsation mechanism of AGB stars that causes changes in magnitude, it has been observed that the amplitude of brightness variation in the light curve is influenced by the effective wavelength of the photometric filter. Specifically, it is higher in the optical range than in the infrared. This phenomenon is illustrated in Figure 1.7, which depicts how the ratio of the flux density at its maximum to that at its minimum decreases as the wavelength increases.

This relationship suggests the existence of physical phenomena in the atmospheres of stars that modulate the amplitude of light curves depending on the observed wavelength. It is known

Luminosity in the optical and infrared and its variability
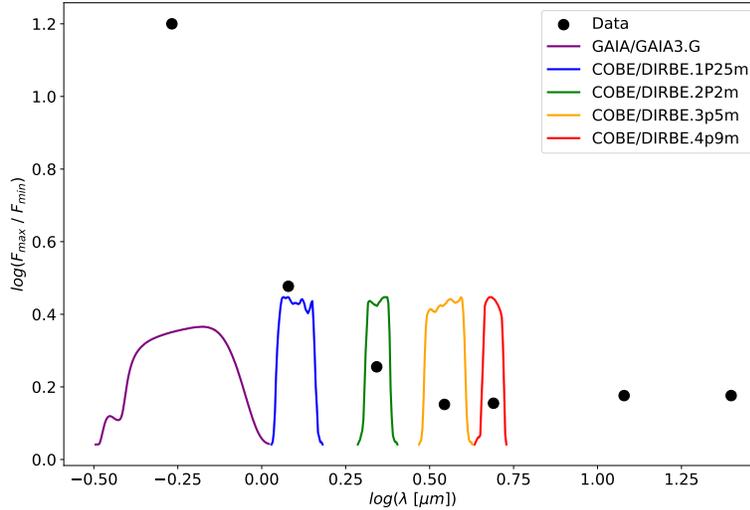
Figure 1.7: This graph illustrates the decreasing trend in the ratio of maximum-to-minimum fluxes of a light curve as the wavelength increases. Filter response curves are presented to demonstrate which photometries bands enable the identification of magnitude changes. The data used to construct this graph are derived from the study conducted by Smith et al. [2002] on the variable star S Cep.

that the dynamic atmosphere of AGB stars contains regions of molecular and dust formation, which can cause extinction in different intervals of their SEDs and to varying degrees. The interaction of optical radiation with gas is distinct from its interaction with dust, and the same applies to infrared radiation with gas and dust.We will briefly describe how radiation flux is modulated based on the opacity of the stellar atmosphere. This process is referred to as the κ mechanism.

1.2.2 Theoretical approach to the variability of AGB stars

As of today, in order to try to understand the variability in magnitude over time of AGB stars, models are being developed that focus on quantifying changes in opacity in stellar atmospheres [Freytag et al., 2017]. For this purpose, there are codes that solve the non-linear equations resulting from coupling the equations of radiative transport with hydrodynamic equations

Theoretical approach to the variability of AGB stars

[Freytag & Höfner, 2008]. The goal is to create 3D models that realistically represent the dynamics of the atmosphere and the circumstellar envelope of an AGB star. These models are an improvement over the older ones, such as those by Bowen et al. [1988] and Wood [1979], which were based on assumptions of spherical symmetry, null magnetic fields, and were solved in only one dimension.

All of these models primarily focus on understanding the effect of radial oscillations in the optical photosphere of AGB stars, assuming that these oscillations exhibit periodic behavior. This approach is commonly known as the piston model [Bowen, 1992]. The energy to sustain this oscillatory piston motion is derived from the energy released in the form of radiation from the partially ionized hydrogen regions located beneath the optical photosphere that expands the star and the gravitational potential of the star that contracts it [Keeley, 1970]. The radial pulsations periodically transfer momentum to the matter present in the stellar atmosphere, causing acoustic waves to propagate through it. As the material reaches specific temperature and density regions, it facilitates the formation of molecular gas, followed by nucleation and the subsequent creation and growth of dust particles. This process enhances the opacity of the region, leading to the absorption of the star's radiation in the optical by the molecular gas and in the infrared and optical by the dust. Consequently, the dust creation and growth region undergoes heating. As a result, thermal pressure increases, and acoustic waves merge to create shock fronts. These shock fronts, combined with radiation pressure, propel the dust particles away from the formation region[Dreyer et al., 2009]. Friction causes the dust to drag the gas as it moves, resulting in the loss of mass of both components. This mass ejection, combined with the suppresion of dust formation due to the shocks, decreases the opacity, increasing the observed optical flux of the star. The reduction in density leads to a drop in the temperature and thermal pressure, which allows for a new cycle of dust and molecular formation. As this is a

cyclical process, it leads to changes in magnitude over time, modulated by the opacity (κ) of the stellar atmosphere [Dreyer et al., 2009]. This is the κ mechanism. It is important to note that the cycle of dust formation is periodic, and this timescale is correlated with the fundamental modes present in the light curves [Höfner & Freytag, 2022].

In Figure 1.8, a schematic representation of the physical processes and how different parameters of the stellar atmosphere behave within the pulsation cycle of AGB stars in luminosity is shown.
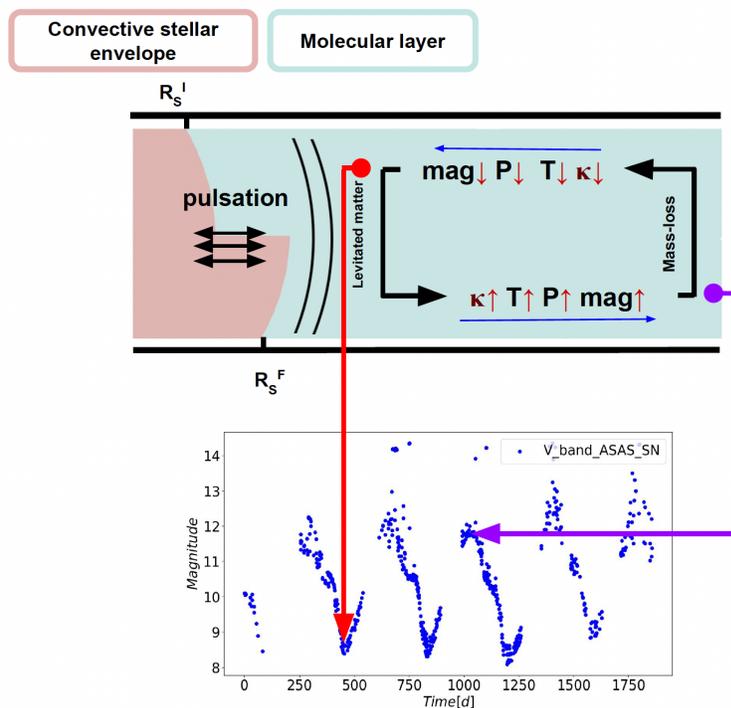
Figure 1.8: In this diagram, the relationship between the different variables characterizing the stellar atmosphere after each pulsation cycle is illustrated. Here, P represents pressure, T represents atmospheric temperature, κ represents the opacity of molecular gas or dust, and mag represents the observed magnitude resulting from the radiation that manages to escape the atmosphere after each pulsation cycle. The blue arrows indicate the direction in which one parameter affects the other.

1.2.3 Long period variables

Stars with pulsation periods in the range 30 to 2000 days are called long-period variables (LPV). Historically, the first LPV star to be discovered was the star Mira, located in the constellation of Cetus. Its name comes from the Latin "mīra", which means "The Wonderful", and was named by the astronomer Johannes Hevelius. The discovery of this star was made by the astronomer Johannes Holwarda in the year 1638, although in 1596 the astronomer Fabricius had already noticed its disappearance from the sky and related it to the phenomenon of a stellar nova. Holwarda noticed that the star decreased and increased its apparent magnitude with a periodicity of months, which caused it to appear and disappear in the sky [Hoffleit, 1997].

Thanks to technological advances in astronomy, such as the photographic recording of the sky on plates, by the year 1896 a total of 251 stars with characteristics similar to Mira had been recorded [Habing & Olofsson, 2004]. These stars showed changes in their magnitude and a large periodicity, which are now known as LPVs. In addition, stars with large periodicity (periods greater than a month) of the regular, semiregular, and irregular types were identified.

In the 1960s and early 1970s, a correlation was discovered between the period of variation in the intensity of the OH 1612 MHz maser line and the period of variation in magnitude in different infrared photometric bands for stars such as NML Cyg, a red supergiant. This finding allowed for the connection between LPVs and AGB stars [Harvey et al., 1974].

LPV classification

LPV stars are classified according to the characteristics observed in their light curves. Mira-type stars are expected to have only the fundamental mode (P_0). This implies that their light curve

can be characterized by a single period. On the other hand, semi-regular (SR) stars are a combination of several pulsation modes (P_n), which can follow the rule of harmonics. Regarding the periods, this can be expressed using Equation 1.1, where $n = 0, 1, 2, ...$:

$$P_n = \frac{P_0}{n+1}.$$ (1.1)

Observationally, Miras and SR stars are differentiated based on their periods and amplitudes in addition to the periodicity of their light curves. Stars in the General Catalog of Variable Stars [GCVS; Samus' et al., 2017], are divided into three groups: Mira, semi-regular, and irregular. Mira-type stars exhibit periods in the range 100—2000 d [e.g., Höfner & Olofsson, 2018 and Riebel et al., 2010] and amplitudes greater than 2.5 mag, based on measurements in the V-band of the Johnson system. On the other hand, SR stars show amplitudes less than 2.5 mag and typically have periods shorter than Miras [30–300 d; Riebel et al., 2010]. For irregular stars, it can only be stated that they exhibit variability without defined periodicity.

Figure 1.9 displays two periodograms presenting power spectra in the period space for two source. In the left panel, it can be observed that there is only one fundamental mode characterizing the light curve, which is typical of a Mira-type star. On the other hand, in the right hand panel, the power spectrum of a semi-regular star is depicted, showing at least four distinct peaks.

Stars of different variability types are found to arrange themselves along different sequences on a period-luminosity (or PL) diagram. These sequences reveal PL relationships that allow for distance estimation based on the recovered pulsation mode from the star's light curve. Figure 1.10 shows one such example of a PL diagram, based on data for stars from the Large Magellanic Cloud (LMC). Periods determined from optical light curves of the MAssive Compact Halo
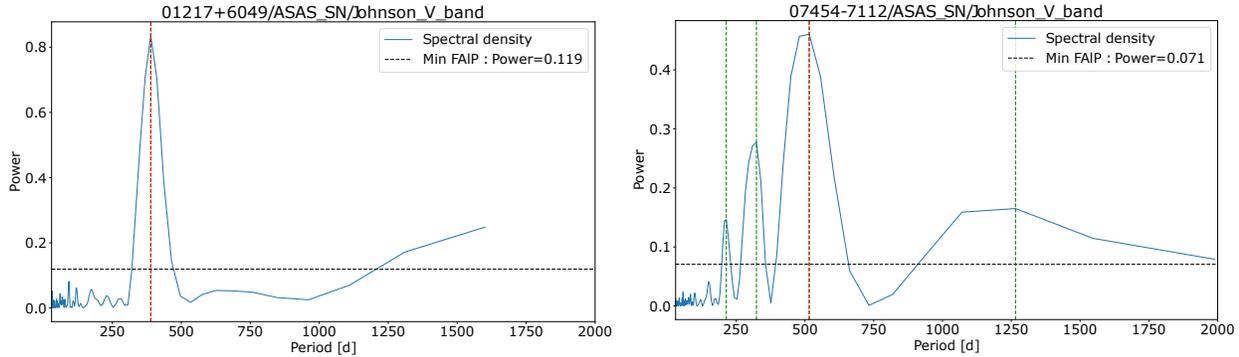
Figure 1.9: Left: Periodogram of a Mira-type star. The highest peak in amplitude corresponds to the "fundamental mode", highlighted by the dashed red line. There are no additional peaks surpassing this line, which represents the minimum power threshold to consider a peak as a significant signal [False alarm probability; Scargle, 1982]. Right: Periodogram of an SR-type star. Multiple harmonic peaks are visible, including the fundamental mode and three additional peaks. One of these is located to the right of the fundamental mode and is known as the "long secondary period." [Pawlak, 2021]

Objects [MACHO; Fraser et al., 2008 and The MACHO collaboration et al., 2000] project are compared with brightnesses in the Spitzer Space Telescope 3.6 μm band [Riebel et al., 2010]. It is important to note that since these stars belong to the LMC, they are statistically located at the same distance, meaning that the studies are independent of distance and extinction.

1.2.4   Dust and AGB stars

Enriched by AGB stars, the interstellar medium receives gas and dust, which have a significant impact on radiative processes and the formation of molecules within it. For instance, dust interacts with light from various astronomical sources and changes the shape of the SED by absorbing optical and ultraviolet spectra and re-emitting them in the infrared range. Additionally, dust acts as a catalyst in converting hydrogen gas to molecular hydrogen, which is crucial for the formation of new stars and planets in molecular gas clouds [van de Sande et al., 2020]. The enrichment of the interstellar medium with gas and dust generated in the stellar atmospheres of AGB stars regulates galactic chemical evolution [Renzini, 2015]. It is worth
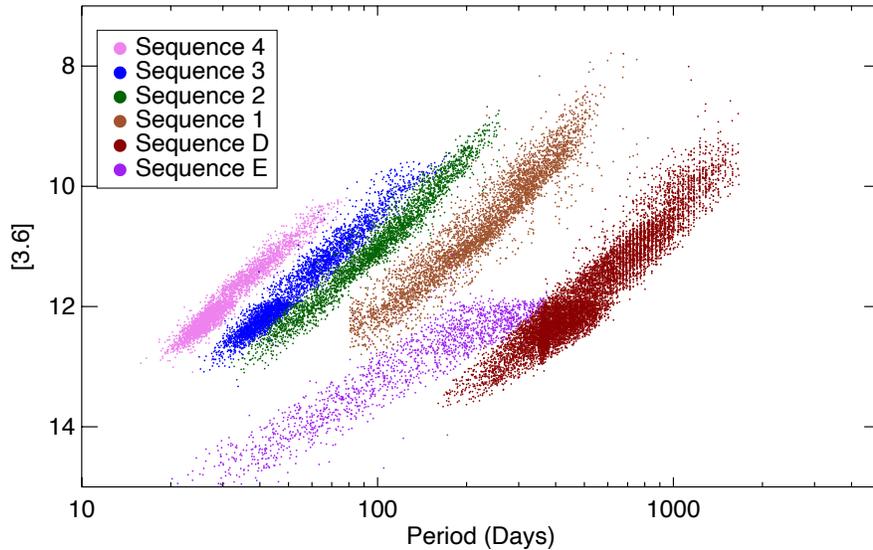
Figure 1.10: Relationship of PL with detections in Spitzer's 3.6 μm photometric band on the ordinate axis and the periods in days on the abscissa. Different sequences of the PL relationship related to the various pulsation modes of LPV stars are observed. Sequence 1 is linked to the fundamental mode and consists of Mira-type LPVs and SRs. On the other hand, sequences 2 and 3 correspond to SR stars displaying the first and second pulsation modes, respectively. Sequence 4, in turn, is associated with SRs exhibiting the third pulsation mode. Additionally, the figure also includes sequences E and D, which represent another type of stellar objects such as binaries. Figure 1, adapted from Riebel et al. [2010].

noting that low- and intermediate-mass stars evolve into the AGB phase and, according to the initial mass function, are more abundant and constitute a larger fraction of the galactic mass [Karakas, 2018]. Therefore, AGB stars are among the dominant dust contributors and they play a fundamental role in the chemistry and evolution of present-day galaxies.

In the realm of optical wavelengths, several significant surveys have been conducted, each progressively contributing to the identification of a substantial number of LPV stars. These surveys, in chronological order of inception, include the American Association of Variable Star Observers (AAVSO), the Sloan Digital Sky Survey (SDSS), the Variable Star Index (VSX), the All-Sky Automated Survey for Supernovae (ASAS-SN), the Zwicky Transient Facility (ZTF),

the Palomar Transient Factory (PTF), the Gaia mission and the Panoramic Survey Telescope and Rapid Response System (PAN-STARRS)[1]. These efforts have collectively facilitated the discovery and categorization of a multitude of variable stars, enriching our understanding of stellar evolution.

More recently, the focus has expanded to infrared surveys, enabling the investigation of variability across longer wavelengths. Projects such as the Diffuse Infrared Background Experiment [DIRBE; Boggess et al., 1992] and the Near-Earth Object Wide-Field Infrared Survey Explorer [NEOWISE; Mainzer et al., 2014] have made notable contributions in this domain. Notably, these infrared surveys possess the capability to observe highly evolved AGB stars, a category often underrepresented in optical surveys because of the significant dust obscuration they experience.

Currently, there are many research efforts to fit data from the surveys such as the ones described above in order to derive physical parameters for LPVs. Iwanek et al. [2022] used data from the Optical Gravitational Lensing Experiment to analyze these stars. In his research, he employed the Lomb-Scargle technique to identify their periods and developed subroutines to detect power peaks in the periodograms. However, it's important to mention that this methodology was primarily applied to sources in the Galactic Bulge and a limited sample of the Galactic Disk. Unfortunately, the data from the latter set has not yet been made available for analysis as part of a broader study. Lebzelter et al. [2023] used data from Gaia Data Release 3 and relied on least-squares fitting to determine the periods. Jayasinghe et al. [2018] analyzed LPV light curves using Lomb-Scargle with data from ASAS-SN. While the above examples do not employ Gaussian process-based methods (see Chapter 3), they have been used to model

---

[1]For more information, consult Table 2.2, which contains the reference of some of the surveys presented here

light curves and estimate periods by others. In the study by Soraisam et al. [2018], a Gaussian processes technique was employed to fit the light curves of 255 red supergiants in the M31 galaxy, comparing them with the Lomb-Scargle method and other approaches. Furthermore, Sanders et al. [2022] applied Gaussian processes fitting to multiple sources, analyzing over 1700 light curves of pulsating variables in the near-infrared, within the nuclear disk of our Galaxy. In this context, the present thesis aims to contribute to the field of modeling light curves of stars in the Solar Neighborhood, using Gaussian processes. Section 3.2.1 provides details of the method and describes the specific kernel that will be used in the analysis.

The estimation of critical parameters, such as luminosity and dust-production rate (DPR), for these stars necessitates an accurate understanding of their distances. While this poses fewer challenges for extragalactic AGB populations given the known distances to nearby galaxies, the issue is more complex for Galactic stars due to the presence of foreground Milky Way dust extinction. Recent advancements, exemplified by the Gaia mission [Jordi et al., 2010], have provided high-precision parallaxes for less evolved stars. However, complications arise in the case of AGB stars due to the influence of pulsation and dust obscuration, leading to compromised optical parallax estimates.

To address this uncertainty in distance, a feasible approach involves the selection of a volume-complete sample of nearby AGB stars. Their proximity not only enhances the likelihood of acquiring both optical and infrared light curves, but also enables comprehensive statistical analysis of their variability characteristics. The primary objective of this thesis is the modeling of multiwavelength light curves of the sample of long-period variables obtained from the Nearby Evolved Stars Survey [NESS; Scicluna et al., 2022]. To achieve this objective, we first collect light curves for these sources from a large set of databases. We will then fit these curves using

the PGMUVI code [Scicluna et al., 2023], which will provide us with periods, amplitudes, and mean magnitudes for each light curve, as well as models for their time variation. These data and results will unveil the variability properties of a large, complete sample of nearby AGB stars over optical as well as infrared wavelengths, thus connecting the variations of the photospheric and dust properites of these sources. The study will therefore enable a more profound understanding of their variability properties as a function of wavelength and contribute to the broader comprehension of stellar evolution.

1.3   The Nearby Evolved Stars Survey

The survey design, sample selection, and main scientific objectives are detailed in Scicluna et al. [2022]. The program is briefly outlined here: the NESS project aims to investigate a comprehensive set of evolved stars within a 3 kpc radius from the Solar System. This initiative employs radio observations conducted by the James Clerk Maxwell Telescope (JCMT; for stars in the Northern Hemisphere) and the Atacama Pathfinder Experiment (APEX) radio telescope (for stars in the Southern Hemisphere). The primary aim of NESS is to characterize the attributes of AGB stars in order to enhance comprehension of the Galaxy's chemical evolution. To attain this objective, NESS examines the characteristics of gas and dust in stellar atmospheres, utilizing indicators that facilitate the calculation of stellar wind velocities and the rate of mass loss for each individual star. This is made possible by APEX and JCMT detecting the molecular transitions $J = 3 \rightarrow 2$ and $J = 2 \rightarrow 1$ of the CO molecule, as well as the emission of the ^{13}CO isotope in the same transitions, which allows for the estimation of the relative abundances between ^{13}C/^{12}C.

The main goals of the NESS are Scicluna et al. [2022]:

- Estimate the current mass return rate to the interstellar medium.

- Estimate the gas-to-dust ratio and the properties of the dust.

- Determine the physics of mass loss in evolved stars.

- Understand wind properties and reconstruct the mass loss histories of AGB stars.

NESS targets were selected from the Infrared Astronomical Satellite [IRAS; Neugebauer et al., 1984] Point Source Catalogue [PSC; Helou & Walker, 1988]. The IRAS 12 and 25 μm features are sensitive to warm dust. The NESS targets were selected based on IRAS colours that indicated the presence of dust. We therefore expect that the majority of NESS sources are evolved stars and LPVs. The work in this thesis will aid the NESS objectives by constraining the variability of the sample.

The NESS team has combined the IRAS photometry for its targets with matching photometry from all-sky surveys spanning optical (Gaia, Pan-STARRS) and infrared (DIRBE, 2MASS, WISE, Akari) wavelengths in order to construct SEDs. Of the 852 NESS targets, this thesis uses 786 objects for which reliable SEDs are currently available. The compilation of NESS SEDs is described in an upcoming paper (McDonald et al., in prep). Figure 1.11 shows the locations of the NESS targets overlaid on a Gaia map of the Milky Way, with chemical classifications based on mid-infrared spectroscopy.

1.4 Justification of this thesis and objectives

As mentioned earlier, systematic studies aimed at analyzing the variability of magnitude in extensive samples of LPVs primarily focus on extragalactic AGB stars. In the case of our Galaxy, the emphasis has been on wavelengths within the optical range, and the sampled stars are not
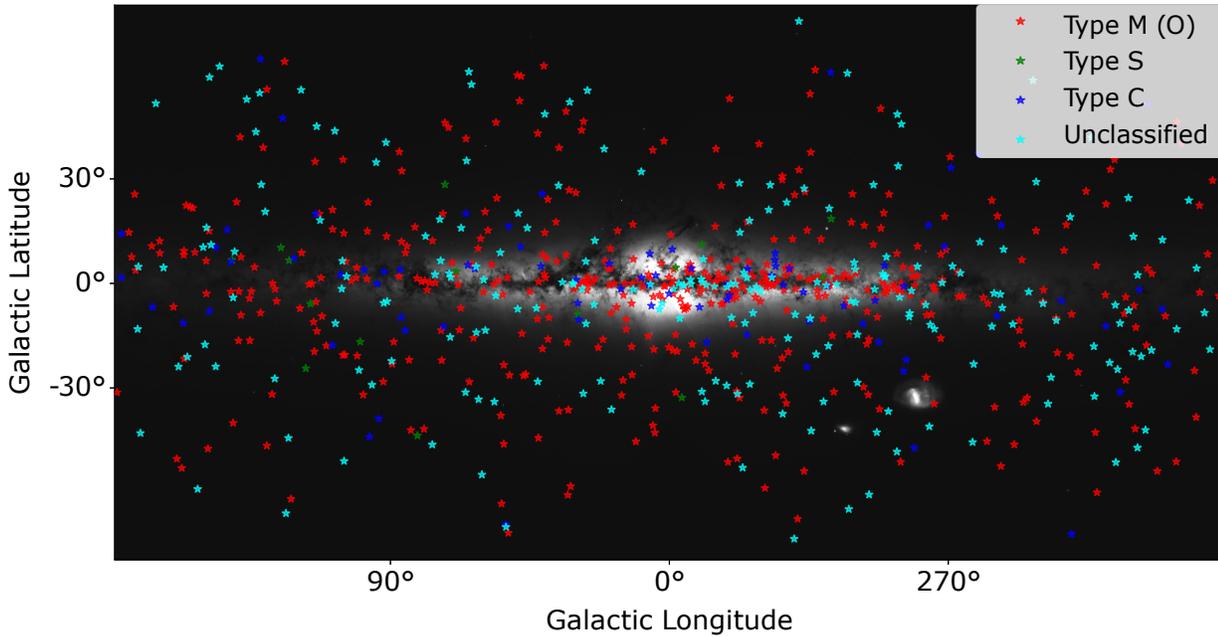
Figure 1.11: The distribution of NESS targets in the sky. The background image of the Galaxy is a star map generated with data from Gaia DR3 [Prusti et al., 2016].

well-defined. While there are studies of AGB stars in our Galaxy that encompass a large number of LPVs, they lack an approach that examines variability through multiwavelength light curves spanning the optical, near-infrared, and mid-infrared.

Therefore, this thesis strives to provide information that contributes to characterizing the pulsations of AGB stars using data collected from multiple wavelengths, including a broad set of stars with light curves obtained at different epochs. The work in this thesis employs Gaussian process regression to model the light curves of nearby evolved stars. Gaussian-process regression has previously demonstrated success in this application, although its conventional usage has primarily been confined to objects like transients, AGN, and exoplanets, as evidenced by prior studies [e.g., Aigrain & Foreman-Mackey, 2022 and Soraisam et al., 2023]. For a comprehensive

understanding of Gaussian processes and the methodology utilized, readers are referred to Chapter 3. The work described in this thesis derives physical parameters (period, amplitude, mean magnitude) for each light curve for each source, and also allows us to probe statistical trends in and among the parameters as well as investigate correlations with other parameters (e.g., dust-production rate, mass-loss rate) related to the star. This statistical approach allows us to extend the numerical models computed for this sample towards the investigation of LPVs in other environments.

This work is significant because it provides valuable information for the NESS Collaboration in terms of understanding and characterizing the properties of stellar atmospheres of evolved stars in the catalog. Additionally, as mentioned earlier, the pulsation modes are related to the timescale at which dust forms, so this thesis provides parameters to understand the characteristics of dust creation.

In order to achieve these objectives, the following goals have been established:

- Develop Python codes to automate the download of light curves from multiple surveys.

- Use the PGMUVI [Python Gaussian processes for MUltiwavelength Variability Inference; Scicluna et al., 2023] code to obtain the fundamental pulsation mode present in each light curve using a single component. This is the first application of Gaussian-process based fitting of AGB-star light curves.

- Build a catalog of light curve models associated with NESS stars.

- Investigate the statistical and phenomenological relationship between stellar variability and the characteristics of the stars and their circumstellar environment.

- Establish correlations between the obtained parameters such as period (fundamental mode) and amplitude of the light curve models with other properties of AGB stars, such as mass-loss rate (MLR), dust content (using the broadband color as a proxy), as well as chemical type.

The rest of the thesis is organized as follows. In Chapter 2, the acquisition of light curves is addressed, where the consulted catalogs and their characteristics are described in detail. It explains how the information was homogenized and the storage method used. Furthermore, the amounts of information obtained in this initial stage of the work are provided. In Chapter 3, data processing and parameter extraction are explored further using algorithms dedicated to periodic signal analyzes. The results obtained through this process are presented in detail. Chapter 4 is dedicated to the analyzes of the estimated parameters and examines the existing correlation with the associated sources. Finally, in Chapter 5, a comprehensive discussion about the obtained results takes place, and the reached conclusions are presented.

Collecting light curves for NESS survey objects

This chapter details the methodology used to collect light curves for NESS targets, providing a thorough description of the photometric characteristics of the databases or telescopes consulted. It explains the automated methods used for downloading and storing the light curves, and shows how parameters were computed from the information obtained in each survey. The chapter also documents the structure of the final data file, and includes tables summarizing the amount of data collected.

2.1 Methodology for collecting light curves

As mentioned in Section 1, the NESS sample is derived from the IRAS point-Source catalog (PSC). This means that each NESS source is associated with a unique IRAS PSC identifier (e.g., IRAS 00042+4248). Each IRAS identifier is cross-matched with a unique identifier from every other survey. The advantage of associating unique identifiers (strings or long integers) with a match is that we can leverage existing matches that other surveys have already computed, rather

than having to search for neighbors within a given radius around a pair of coordinates ("cone search"). In case there are multiple matches within this radius, we also need to decide which one is the true counterpart of our source. Matching by unique identifier avoids this duplication of effort, which is an important step towards reproducibility of the procedure. Whenever possible, we find light curves in a survey by finding the matching identifier. For example, for the IRAS 00042+4248 source, it is possible to find the following identifiers, as shown in Table 2.1:

Table 2.1: Example of identifiers for a NESS source with an IRAS ID in other surveys. An additional example is that knowing the Gaia DR2 ID makes it possible to obtain the Gaia DR3 ID and thus find information on the most recent light curves.

Source in IRAS	Survey as explorer	ID in the Survey.
00042+4248	DIRBE	D00065274P4305021
	AllWISE	J000652.77+430502.2
	Gaia DR2	384627244915398144

During this study, two main forms were identified, each with two variants, for downloading light curves. If the surveyed database has a Table Access Protocol (TAP)[1], the query can be performed using a related identifier or a cone search. If this service is not available, Uniform Resource Locators (URLs) can be constructed to download files in HTML (Hypertext Markup Language) or JSON (JavaScript Object Notation) format and process them as plain text files. Constructing these URLs involves identifying the pattern of character line writing, which generally take object IDs or coordinates as arguments within their data server. In most cases, this simply involves replacing these parameters in the URLs provided on their web pages. As an example, the light curve for the source IRAS 00042+4248 in the Zwicky Transient Facility survey can be downloaded from the following URL: ZTF URL. Figure 2.1 shows a flowchart summarizing the logic.

---

[1]An example of a TAP service is Vizier: https://tapvizier.u-strasbg.fr/adql/.
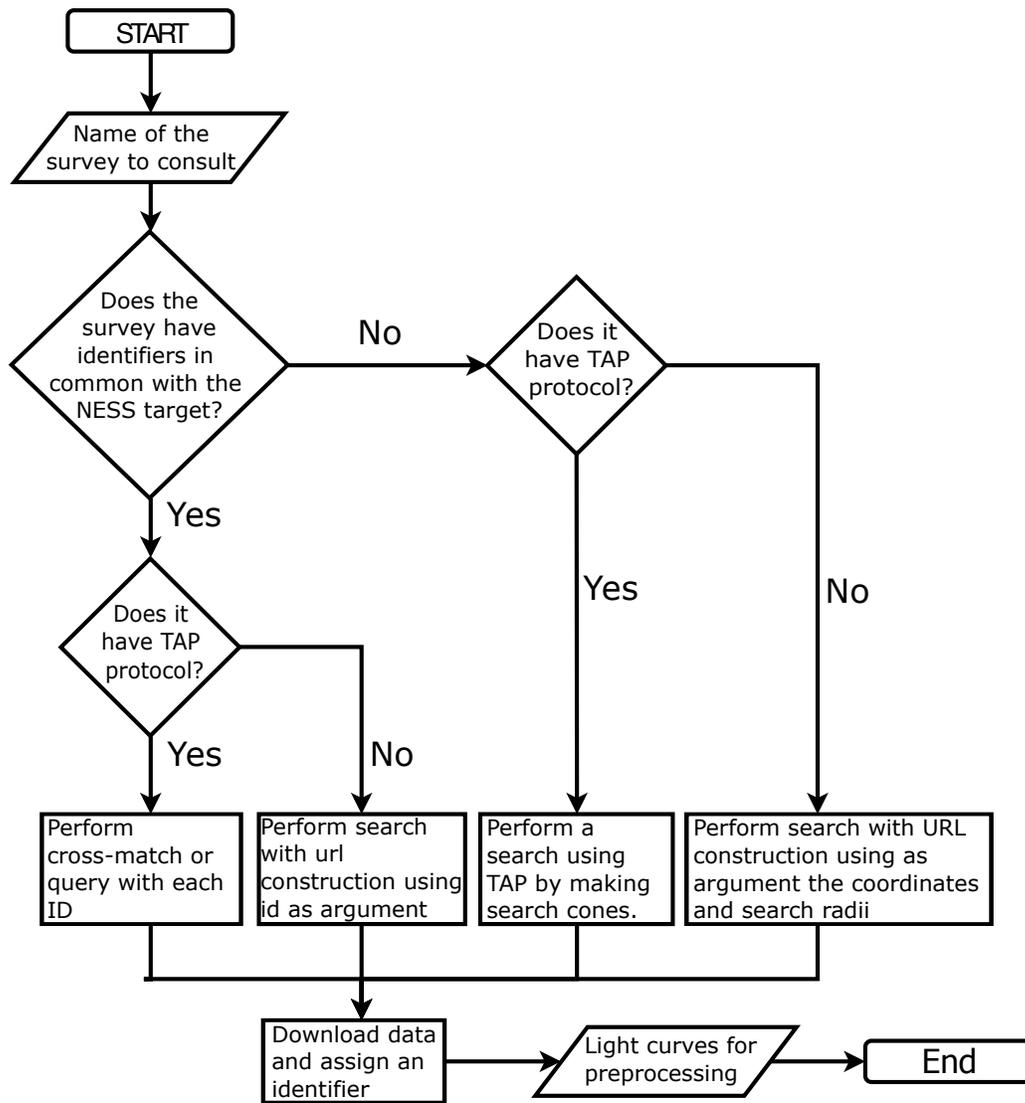
Figure 2.1: Flowchart showing the logic for downloading light curves with possible options for creating a Python code. While the detailed process of each code is not shown, it provides a broad overview of the search decisions that need to be made.

A Python code was developed to construct a file containing the J2000 coordinates of objects in the NESS database for searches based on coordinates. The code uses the .SkyCoord.from_name() function from the astropy library, which queries in *Centre de Données astronomiques de Strasbourg* (CDS) with the identifier and coordinate system of the object and returns a SkyCoord

object that can be decomposed into its right ascension and declination.

As not all NESS targets have identifiers in each database, queries are prioritized based on surveys with greater accuracy in position detection. Specifically, queries are first made with the Gaia ID, then the 2MASS ID, and finally the AllWise ID in order of precision. The resulting file can be used as a plugin to perform specific surveys, which are listed below.

2.2 Photometric characteristics of the different databases consulted

In Table 2.2 we list the different surveys consulted that comply with the previous approach and contains: the name of the survey, the key name in this thesis, and the main article(s) in which the data or technical information of the survey was published, and finally, the year in which the information was published.

Table 2.2: The names, key identifiers in this thesis, and main articles of the nine surveys, collaborations, or catalogs consulted are presented.

Telescope, observatory or survey	ID in this thesis	Main papers	Release year
The Serendipitous Survey (ASAS SN)	ASAS_SN	Kochanek et al. [2017] and Jayasinghe et al. [2020]	2020
The AAVSO Photometric All-Sky Survey dr10	APASS_DR10	Henden [2019]	2018
DIRBE Near-Ifrared Stellar Lihgt Curves	DIRBE	Price et al. [2010]	2010
Gaia Data Release 3	GAIA_DR3	Gaia Collaboration et al. [2022]	2022
The Kilodegree Extremely Little Telescope	KELT	Pepper et al. [2007] and Arnold et al. [2020]	2022
The Near-Earth Object Wide-field Infrared Survey Explorer	NEOWISE	Mainzer et al. [2014]	2022
The PAN-STARRS1 Surveys dr2	PANSTARRS	Chambers et al. [2016]	2019
Near-infrared photometry of carbon stars	WHITELOCK	Whitelock et al. [2006]	2006
The Zwicky Transient Facility	ZTF	Bellm et al. [2019] and Masci et al. [2019]	2022

A summary of the photometric characteristics of the different surveys can be found in Table 2.3, such as: key identifier of the survey, ranges of the electromagnetic spectrum (optical and

infrared) for its photometric filters, which each contain the technical name of the filter, for example: the name it has in Spanish Virtual Observatory database[SVO; Rodrigo & Solano, 2020 and Rodrigo et al., 2012] or the name of the photometric system to which the filter belongs according to the survey and their respective effective wavelengths λ_{eff}.

2.3 Brief description of the catalogs and their respective repositories

The physical characteristics of the telescopes, cameras and photometric systems are discussed, as well as their location (space or terrestrial) among other characteristics of each survey.

ASAS_SN

The "All-Sky Automated Survey for Supernovae" is primarily dedicated to the study of supernova events. It collects photometric information over time using various stations around the world. The published catalog contains data from the Las Cumbres and Cerro Tololo sites, which 14cm aperture telephoto lenses and can record a sky field equivalent to 4.5 deg^2 with an exposure of 90 s. Currently, light curves can be obtained in photometric filters V and g [Johnson's system; Kochanek et al., 2017]. From this instrumentation, the "Serendipitous Survey" was created, consisting of 27,479 periodic variable sources and 38,700 irregular ones identified only in the V band [Jayasinghe et al., 2018]. Since ASAS SN calibrates its zero point in flux using information from APASS [Kochanek et al., 2017], the photometry is adopted with the Misc/APASS.V filter.

To obtain the light curves in the V band, we cross match the Serendipitous Survey table "II/366/catalog", which is available at VizieR, with the NESS table. We used the AllWISE IDs as the key, since "II/366/catalog" had more NESS objects with that identifier than with

Table 2.3: This table shows the basic photometric information of the surveys consulted. In the Filter ID column is the name of the filter supplied for each survey and also the name with which the band can be searched in our database.

ID in this thesis	Spectral Range	Photometric Bands	
		Filter ID[†]	λ_{eff} [nm]
ASAS_SN	Optical	V: Misc_APASS.V	539.3
APASS_DR10	Optical	B: TYCHO_TYCHO.B	428.0
		V: TYCHO_TYCHO.V	534.0
		g': SLOAN_SDSS.gprime_filter	482.0
		r': SLOAN_SDSS.rprime_filter	628.7
		i': SLOAN_SDSS.iprime_filter	745.7
	Infrared	z: SLOAN_SDSS.z	899.2
DIRBE	Infrared	1.25 μm: COBE_DIRBE.1P25m	1255.4
		2.2 μm: COBE_DIRBE.2P2m	2210.1
		3.5 μm: COBE_DIRBE.3P5m	3491.1
		4.9 μm: COBE_DIRBE.4P9m	4867.4
GAIA_DR3	Optical	BP: GAIA_GAIA3.Gbp	531.9
		G: GAIA_GAIA3.G	671.9
		RP: GAIA_GAIA3.Grp	793.9
KELT	Optical	R_k: OSN_Johnson.Cousins_R3 (_0, _1)	645.5
NEOWISE	Infrared	$W1$: WISE_WISE.$W1$	3352.6
		$W2$: WISE_WISE.$W2$	4602.8
PANSTARRS	Optical	g: PAN-STARRS_PS1.g	481.0
		r: PAN-STARRS_PS1.r	615.5
		i: PAN-STARRS_PS1.i	750.3
		z: PAN-STARRS_PS1.z	866.8
		y: PAN-STARRS_PS1.y	961.3
WHITELOCK	Infrared	J: CTIO_OSIRIS.J	1258.1
		H: CTIO_OSIRIS.H	1616.2
		K: CTIO_OSIRIS.K	2163.0
		L: IRTF_NSFCAM.L	3472.0
ZTF	Optical	g: Palomar_ZTF.g	472.2
		r: Palomar_ZTF.r	633.8
		i: Palomar_ZTF.i	788.7

[†] Filter characteristics for all surveys available from the Spanish Virtual Observatory's (SVO) Filter Profile Service.

other IDs like Gaia or 2MASS. The light curves can be displayed on the VizieR website using URLs provided by the Serendipitous Survey. By appending the ".csv" string to these URLs,

we were able to download the raw data using the `urllib` library[1] in Python, specifically the `request.urlopen` method from this library. We downloaded these tables and stored them in CSV format using the `.Table` module and the `.write()` function from the Python Astropy library, with the name ASASSN_«IRAS_ID».CSV. An example of the raw data can be found in Table 2.4, which includes the time, the instrument used to record the data, the magnitude, and the flux returned by VizieR after the query.

Table 2.4: Main parameters obtained by consulting the light curve of IRAS 00042+4248 reported by ASAS SN in VizieR

Return parameter by ASAS SN	Raw values	Parameter description
hjd	2457212.02074	Time measured in the Julian Date
camera	bb	Instrument identifier
mag	16.839	Magnitude in Vega system in band V
mag_err	0.13814	Uncertainty in magnitude
flux	0.705	V-band flux in mJy
flux_err	0.09	V-band flux uncertainty in mJy

APASS_DR10

The AAVSO[2] Photometric All-Sky Survey (APASS) is a survey that covers the entire celestial sphere by capturing photographic records from two stations: one located in Mayhill, New Mexico in the Northern Hemisphere, and the other in Cerro Tololo, Chile [American Association of Variable Star Observers, 2019]. Both stations employ identical instrumentation− **20 cm aperture telescopes −that record an** apparent area of the sky of 2.9 x 2.9 degrees square [Smith et al., 2010].

---

[1] https://docs.Python.org/3/library/urllib.request.html
[2] American Association of Variable Star Observers https://www.aavso.org/

OTA1 captures magnitudes in bands B, g', z, and z' with a recording time of 180 s for each observation, while OTA2 captures magnitudes in bands V, r', and i' with an exposure time of 90 s for each observation. The APASS Data Release 10, published in late 2018, contains the photometric record of 128 million celestial objects. Notably, the records of bands B and V are reported in the Vega magnitude system, while the other bands (g', r', i', z, and z', or Sloan photometry) report magnitudes in the AB system [American Association of Variable Star Observers, 2019].

To download individual light curves in CSV format, URLs were constructed that allow the connection to the APASS server. These URLs perform queries through search cones and are passed as the main argument of the `urlopen()` method. The result of this function is then passed as a parameter of the `.read()` function, which allows the requested information to be downloaded and temporarily stored in memory.

The coordinates used for this process are those obtained in Section 2.1. Using the `BeautifulSoup`[1] Python library, the information is converted to text to identify the metadata and separate it from the columns containing the light curves. After this separation is achieved, the text containing the columns is transformed into an Astropy Table object, which is stored under the name APASS_«IRAS_ID».csv. An example of the raw data can be found in the Table 2.5.

When searching for an object in APASS that belongs to NESS, an indicator that it was not found is when the metadata displays the text "No rows were returned by query." In such cases, the associated identifier is discarded and the search continues with the next star.

---

[1]https://pypi.org/project/beautifulsoup4/

Table 2.5: Raw data for one of the sources of NESS present is the APASS catalog

Return parameter by APPAS	Raw values	Parameter description
hjd−24e5	56199.69356	Time measured in the Julian Date where the day zero is $24 \cdot 10^5$
mag	7.1291	Magnitude in Vega system for B and V filters and in AB system for g',r',i',z and z' filters
errmag	0.001	Uncertainty in magnitude
filter	'B'	filter identifier
dra_arcsec	−0.22	Uncertainty in right ascension
ddec_arcsec	−0.41	declining uncertainty

DIRBE

The Cosmic Background Explorer (COBE) was launched into space in 1989 with the objective of studying the spectrum of the Cosmic Microwave Background Radiation (CMB). The mission comprised three experiments that were specifically designed to detect spectra in the microwave and infrared range [Boggess et al., 1992]. Among these experiments was the Diffuse Infrared Background Experiment (DIRBE), which collected data in 10 photometric bands ranging from 1 to 224 μm. It consisted of a 19 cm telescope with a field of view of 0.7×0.7 square degrees.[Silverberg et al., 1993]

Thanks to the catalog created for the paper "3.6 YEARS OF DIRBE NEAR-INFRARED STELLAR LIGHT CURVES" [Price et al., 2010], light curves for 2652 stars are available at 1.25, 2.2, 3.5, and 4.9 μm from DIRBE. This work demonstrated that thanks to the 3.6 years of photometric data collection of the entire sky by COBE, it is possible to study LPV stars from the infrared with a large enough sample of epochs to observe pulsations. This is not possible with only the 10 months of the cold mission of COBE. The technical details for obtaining fluxes from the complete maps created in DIRBE are also explained in this work.

The VizieR database provides access to a table J/ApJS/190/203/var that lists the stars with DIRBE identifiers and NESS stars, enabling the cross match of tables using the DIRBE ID as a key. The essential parameters for making queries are the DIRBE ID and the photometric band IDs, which indicate the bands where DIRBE has information on mean flux and pulsation period. By using the .get function of the requests module and passing the URLs as the main parameter, it is possible to download the data and store each light curve in a data table in ASCII format. This format is adopted because the fluxes for different bands are grouped in individual tables and separated by an empty row, which flags the beginning of each band. Therefore, it is crucial to know the bands where DIRBE has measured fluxes as this information is essential in the post-processing stage, where the curves must be separated into individual photometric bands. An example of raw data can be seen in Table 2.6.

Finally, it is important to mention that of the 2652 stars identified by Price et al. [2010], in table J/ApJS/190/203/var only 597 objects identified as variables or candidates for variable sources are reported.

Table 2.6: Information obtained from the DIRBE sources in the table catalog J/ApJS/190/203/var

Return parameter by DIRBE	Raw values	Parameter description
#JD-2440000	7875.00	Time measured in the Julian Date with day zero at 2440000
Flux (arbitrary units)	11.92	Flux of each band in units of Jy

GAIA (3rd Data Release)

The Gaia mission was launched to study the Milky Way in three dimensions using three photometric bands: G, GBP, and GRP [Table 2.3; Jordi et al., 2010]. Since 2014, Gaia has

conducted surveys of the sky and has published data on three occasions. Gaia collects an average of 70 data points per star over a 5-year interval [Prusti et al., 2016], providing useful light curves for variability studies. In the most recent publication, 11,754,237 time series were reported, and 10,509,536 variable sources were identified and classified into 24 types of variable objects. Of these, 1,720,588 are long-period variables [LPVs; Gaia Collaboration et al., 2022].

To download the light curves of Gaia NESS objects, their Gaia DR2 IDs were first used to find their Gaia DR3 counterparts using the "gaiaedr3.dr2_neighbourhood" table available at https://gea.esac.esa.int/tap-server/tap. This was done using code that connects to Gaia and VizieR's TAP service through the pyvo library and the TAPService functions.

With the Gaia DR3 IDs, each light curve is downloaded using the .search() function, which takes a stream with ADQL language commands that return time data, flux, and magnitude in the three photometric bands for the identified DR3 source. The light curves are stored in the "I/355/epphot " [Gaia collaboration, 2022] table available from VizieR. Each raw light curve is stored in CSV format under the name "GAIADR3«IRAS_ID».csv". An example of the raw data for a particular source can be seen in Table 2.7.

KELT

The Kilodegree Extremely Little Telescope (KELT) is a robotic telescope designed primarily for detecting extrasolar planets by studying the light curves that vary due to planetary transits. To carry out this task, it consists of two stations located in different places: one in Arizona (USA), known as the Winer Observatory, and another in the South African Astronomical Observatory (SAAO) near Sutherland, South Africa. Each telescope has a 42 mm aperture lens, and it also features an automated system that captures images of the sky every 30 minutes.. These

Table 2.7: Raw data of times, magnitudes and fluxes in the three Gaia photometric bands for the source with IRAS ID: 00193-4033

Return parameter by Gaia DR3	Raw values	Parameter description
Source	4997070998651580416	Time measured in the Julian calendar
TimeG	1753.303559	Time in Julian Date with day zero at 2455197.5
FG	29758.3623979809	Flux in band G in units of e^-/s
e_FG	174.115270137742	Uncertainty flux in band G in units of e^-/s
Gmag	14.503344	Magnitude in Vega system in band G
TimeBP	1753.3039	Time in Julian Date with day zero at 2455197.5
FBP	123.120248482789	Flux in band BP in units of e^-/s
e_FBP	16.4142657566337	Uncertainty flux in band BP in units of e^-/s
BPmag	20.112719	Magnitude in Vega system in band BP
TimeRP	1753.304	Time in Julian Date with day zero at 2455197.5
FRP	65140.5509004643	Flux in band RP in units of e^-/s
e_FRP	148.254668130072	Uncertainty flux in band RP in units of e^-/s
RPmag	12.713267	Magnitude in Vega system in band RP

images cover portions of the sky equivalent to 26 x 26 square degrees with a resolution of 23 arcseconds per pixel. The KELT has been collecting light curves since 2005. The photometric filter used is the Kodak Wratten No. 8, which is equivalent to the Johnson $R-$band filter [Pepper et al., 2007]. The photometry used for the KELT project in this thesis is based on the OSN/Johnson.Cousins_R filter provided by the "Observatorio Sierra Nevada.", **available through the SVO.**

The light curves are the result of the study titled "Long-period High-amplitude Red Variables in the KELT Survey" conducted by Arnold et al. [2020]. This study focuses on AGB stars in the Galaxy observed with the KELT telescope, aiming to determine if they are Mira-type variable stars. In this work, 4132 stars were reported as Miras, representing approximately 70% of the total, while the rest were classified as semi-regular stars. Additionally, a study was carried out to determine the chemical composition of these stars.

To download the light curves, an automated query was made using the "J/ApJS/247/44/table2" table stored in VizieR. Common identifiers were cross-matched using the 2MASS IDs, which are also available in NESS, utilizing the TAP service. This table contains a column that stores the names assigned by KELT to the light curves, enabling the construction of URLs that, along with the `.get` module of the `requests` library, facilitate downloading in text format. Headers were removed, and each source was downloaded using the functions provided by the Table library in the Astropy module. Each raw light curve is stored in CSV format under the name "KELT«IRAS_ID».csv". An example of the raw data for a particular source can be seen in Table 2.8.

Table 2.8: Main parameters obtained by consulting the light curve of IRAS 00042+4248 reported by KELT in VizieR

Return parameter by KELT	Raw values	Parameter description
JD	4036.718103	Time measured in the Julian Date from day 2450000
Rmag	10.26074	Magnitude in vega system in band R
e_R	0.1	Uncertainty in magnitude
E/W	1.0	Observation from the east (SAAO) if it is 0.0 or from the west (Winer Observatory) if it is 1.0

NEOWISE

The Wide-field Infrared Survey Explorer [WISE; Wright et al., 2010] was a NASA mission launched in 2009 to map the entire celestial sphere in the mid-infrared (bands $W1$, $W2$, $W3$, and $W4$, centred at 3.3, 4.6, 12, and 24 μm respectively) during its "cold mission", which lasted for 10 months [Infrared Processing and Analysis Center, 2013]. This ambitious task was important because it would allow scientists to identify previously unknown objects, such as brown dwarfs, galaxies, and asteroids, and study the evolution of these objects over time. WISE accomplished

this task using a 40 cm primary mirror which provided a field of view of 47 x 47 arcmin^2 [Jet Propulsion Laboratory, 2009].

Following the successful completion of its initial "cold mission", WISE continued to operate in a new phase called NEOWISE, which lasted for an additional 4 months. During this phase, the explorer focused primarily on using the $W1$ and $W2$ bands to detect near-Earth objects (NEOs) and further expand our understanding of the solar system. After completing this phase, WISE was put into hibernation around 2011 [Wright, 2009]. However, in 2013, the explorer was reactivated and underwent an operation and calibration phase to continue the search for NEOs. Since its reactivation, the explorer has been referred to as NEOWISE and has been releasing new data every year for the study of objects in the solar system [Infrared Processing and Analysis Center, 2021], as well as objects with variable brightness in the near-infrared.

To download the NEOWISE light curves, the `.query_region()` method from the Infrared Science Archive (IRSA) module of the `astroquery` library was used. The `IRSA` module is a collection of tools for accessing and analyzing data from infrared and submillimeter observatories, and the astroquery library is a Python module for querying astronomical databases. The `.query_region()` method takes several parameters, including an identifier to provide coordinates for a search cone (in this case, Gaia DR2 identifiers), the name of the table where the light curves are stored ("neowiser_p1bs_ps", available in https://irsa.ipac.caltech.edu/Missions/wise.html), a search radius of 3 arcsec, and a character string with the desired columns to obtain, which are presented in a table of raw data.

After the query, the obtained table is checked to see if it is empty. If the table is empty, the query is discarded and the next object is processed. If the query contains information,

it is saved using the Table module and stored in CSV format. The name of each table is NEOWISE«IRSA_ID».csv.

NEOWISE conducts periodic observations of a specific section of the sky approximately every 180 days, for about a day. This generates multiple records (on average 8) of a particular object in that region, separated by only a few hours. Visual inspection of the light curves (e.g., Figure 2.2. NEOWISE light curve panel) shows a significant short-term variation in the magnitude (of approximately one magnitude within an interval of approximately one day). This observational artifact leads to contamination in the obtained parameters from these curves and affects the results of the different surveys conducted.

It is known (NEOWISE mission's website) that the $W1$ and $W2$ bands suffer from saturation issues for nearby bright IR objects. These saturation issues affect time-series observations in NEOWISE, and it is suggested that objects with $W1 < 2$ mag and $W2 < 0$ mag should be discarded, as there is no way to perform corrections for all sources [Groenewegen, 2022 and Suh, 2021]. While light curves for this dataset of NESS targets were collected, the decision was made not to use them as part of this thesis work due to concerns about saturation. However, the process of obtaining the light curves will still be explained, leaving open the possibility of utilizing this data in the future.

PANSTARRS

The Panoramic Survey Telescope and Rapid Response System (Pan-STARRS) is a survey project that uses the PS1 telescope located at Haleakala, Hawaii to continuously map the sky for the purpose of identifying Near-Earth Objects (NEOs) and studying the structure of

Table 2.9: Raw data for a NEOWISE source where the information obtained after the consultation in IRSA/IPAC is shown

Return parameter by NEOWISE	Raw values	Parameter description
ra	6.9820737	Right ascension (J2000)
dec	-33.0071868	Declination (J2000)
clon	00h27m55.70s	Right ascention in Sexagesimal format.
clat	-33d00m25.87s	Declination in Sexagesimal format
allwise_cntr	71033401351042267	Closest associated ALLWISE Catalog source cntr principal neowiser
tmass_key	98906296	Closest associated 2MASS All-Sky Release PSC key
mjd	57551.47656175	Julian date of the mid-point of the observation of the frame principal
$W1$mpro	-1.196	Magnitude in band $W1$ in vega system
$W1$sigmpro	-1.322	Uncertainty in magnitude in band $W1$
$W2$mpro	0.061472	Magnitude in band $W2$ in Vega system
$W2$sigmpro	169.535837	Uncertainty in magnitude in band $W2$

the Milky Way [Institute for Astronomy at the University of Hawaii, 2023]. The telescope is equipped with 1.8 m diameter primary mirror that covers approximately 7 square degrees of the sky in the photometric bands g, r, i, z, and y, with exposure time varying for each band but not exceeding one minute. Additionally, the telescope can scan up to 1000 square degrees of the sky in a single night of observation. The data collected by the telescope is processed and made available on the Mikulski Archive for Space Telescopes [MAST; Chambers et al., 2016].

To download light curves, the PANSTARRS Data Release 2 (DR2) from 2019 was utilized [Flewelling et al., 2020], and Python functions provided by MAST were used to access their API and make queries (available on MAST's Github page). As NESS has PANSTARRS IDs, they were used as an argument for the search with MAST. However, it is worth noting that not all objects have light curves, so queries were made to confirm the existence of the object within DR2. Raw data was then downloaded by requesting specific parameters such as time, flux, and

Table 2.10: Table with the information retrieved from MAST for a PANSTARRS source

Return parameter by PANSTARRS	Raw values	Parameter description
objID	144760178163592618	ID in PANSTARRS
detectID	136648175450000002	Unique detection identifier
filterID	3	Filter ID encoded in PANSTARRS
obsTime	55467.4820299	Modified Julian Date
ra	17.81636912	Right ascension
dec	30.63506952	Declination
psfFlux	0.0670306980609893	Flux from Point Spread Function (PSF) fit
psfFluxErr	4.4578398956218734e-05	Error on flux from PSF fit.
filter	i	Filter name

position. An example of the raw data and its parameters can be seen in Table 2.10.

WHITELOCK

The Whitelock et al. survey is based on the study of 239 carbon-rich evolved variable stars, during which their characteristics such as K-band period, distance, bolometric magnitude, rate of mass loss, and other properties were determined. The observations were made at the South African Astronomical Observatory (SAAO) using the 0.75 m aperture telescope and supplementing some observation nights with the 1.9 m telescope, in the J, H, K and L filters [Whitelock et al., 2006] from the Glass photometric system [Glass, 1973]. Since the photometry in the J, H, and K bands of Glass is an old system without digital support, it was decided to utilize the photometry from The Ohio State InfraRed Imager/Spectrometer (OSIRIS) instrument on the CTIO telescope, as this dataset is endorsed by the SVO. For the Glass L band, the L band information from the NSFCAM instrument at the NASA Infrared Telescope Facility was adopted.

Light curves can be obtained with VizieR's TAP service by querying the `.dal.TAPService` module and the .search() function. The table containing all the curves is called "J/MN-

RAS/369/751/table2". To find out which of the NESS stars are in this survey, we consulted the table "J/MNRAS/369/751/stars" . Both tables are available on VizieR. The table "J/MN-RAS/369/751/stars" is downloaded and saved in CSV format so that it can later be divided by each star and each observed band. For this survey, the apparent magnitude of the bands (H, J, K, and L) are obtained from the table. Table 2.11 shows the raw data.

Table 2.11: First rows of the table "J/MNRAS/369/751/table2" , with time and magnitude records in bands H, J, K and L . The table contains 4615 records for 258 stars.

Return parameter by WHITELOCK	Raw values	Parameter description
recno	1	Reference for metadata in VizieR
Name	R Scl	SIMBAD identifier
HJD	2443123.5	Time in Julian Date
Jmag		$J-$band magnitude in the Vega system
u_Jmag		Uncertainty in magnitude in $J-$band in the Vega system
Hmag	0.6	$H-$band magnitude in the Vega system
u_Hmag		Uncertainty in magnitude in $H-$band in the Vega system
Kmag	-0.03	$K-$band magnitude in the Vega system
u_Kmag		Uncertainty in magnitude in $K-$band in the Vega system
Lmag	-0.62	$L-$band magnitude in the Vega system
u_Lmag		Uncertainty in magnitude in $L-$band in the Vega system

ZTF

The ZTF survey is designed to study the temporary changes of astronomical sources with variable luminosity, and it succeeds The Palomar Transient Factory [PTF; Bellm et al., 2019]. To construct this survey, a robotic system is used to map the sky with the 48-inch telescope of the Palomar observatory, which offers a field of view of 47.7 deg^2 [Masci et al., 2019] in three photometric bands (g, r, and i; see Table 2.3) and provides detections on the ZTF website.

To download each ZTF light curve, URLs were built with parameters that request the object's coordinates, a search radius, the bands to be consulted, and the data download format. The coordinates were obtained from the file consulted with `SkyCoord`, and the search radius was determined based on the identifier used for the coordinate consultation (0.5 arcsec for Gaia DR2 queries and 0.8 arcsec for 2MASS or AllWISE queries). The photometric bands consulted were: g, i, and z . The format selected as VOTable [a standard XML format designed for the use of astronomical data; Demleitner et al., 2015]. The data was downloaded from the ZTF servers using the `urlopen()` method.

Since the data in the cache was in VOTable format, the length of the columns was used to detect if there was data or not. If the table extension was null, it was discarded. This is the way to identify if there are objects in common between ZTF and NESS. Finally, the light curves were stored as files in CSV format and saved under the name "ZTF_«IRAS_ID».csv" in the Table 2.12 as an example of the raw data.

2.4 Comments on light curve downloads

In Figure 2.2, various light curves corresponding to different NESS targets are presented, each taken in different photometric bands. This collection has been designed with the purpose of showcasing a variety of behavior patterns. When examining these light curves, certain cases stand out that display regular and easily predictable patterns in terms of the source's period. These patterns are visible through visual inspection, as illustrated in Figures 2.2.b, 2.2.c, 2.2.d, 2.2.e, 2.2.h and 2.2.i. In contrast, light curves with missing data or irregular sampling are also observed, exemplified by figures 2.2.a and 2.2.g.

Table 2.12: Data table for the raw light curve of IRAS 00247+6922, showing the most important parameters that ZTF returns when consulting with the URL.

Return parameter by ZTF	Raw values	Parameter description
oid	833211400019365	Object ID
expid	49048552	Exposure ID
hjd	2458244.982574490	Time in Heliocentric Julian Date
mjd	58244.4855208	Time in Modified Julian Date
mag	19.396907806396500	Magnitude in respective photometric band
magerr	0.0687364861369133	Uncertainty in magnitude
catflags	0	Catalog flags for source from PSF-fitting catalog
filtercode	zr	Photometric band identifier in which the magnitude was recorded
ra	6.9219081	Right Ascension of source
dec	69.6476697	Declination of source

Additionally, distinguishable light curves can be identified that exhibit regular sampling, where the data sampling pattern visually prevails over the inherent variability of the star, as seen in Figure 2.2.f. Simultaneously, highly informative light curves exist, as shown in figures 2.2.b, 2.2.c, 2.2.e and 2.2.i. Another noteworthy aspect is that the pulsation behavior is consistent in light curves with appropriate sampling, allowing the identification of pulsation patterns in different photometric bands of the same survey for a given source. This is clearly illustrated in figures 2.2.c, 2.2.d, 2.2.h and 2.2.i, where the behavior is similar in all bands, with differences solely in the curve amplitudes.

2.5 Pre-processing

In order to properly collect and store light curves associated with NESS targets, we have to
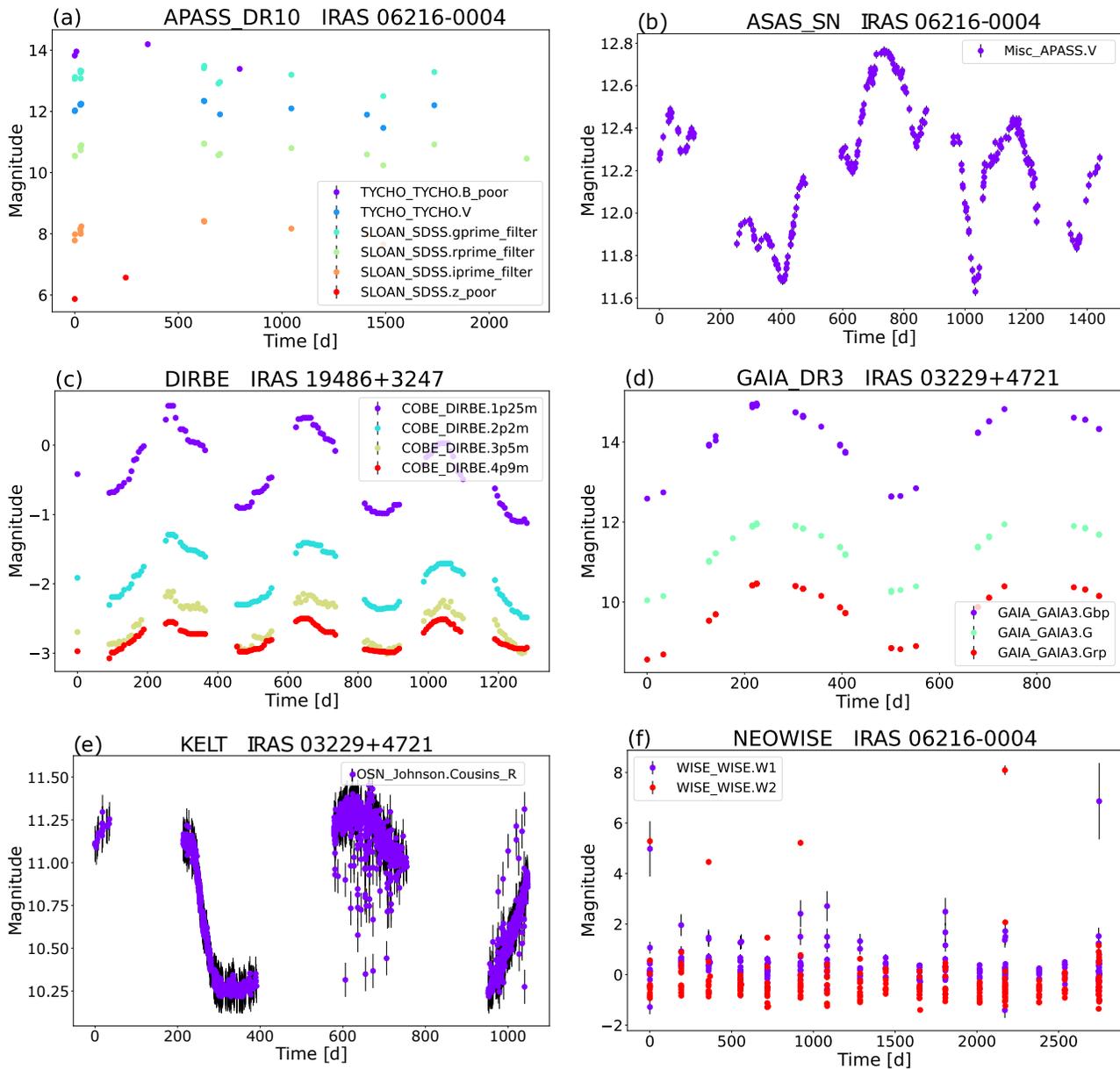
Figure 2.2: Collection of six surveys where the light curves of a NESS target can be seen. Although these light curves exhibit a periodic pattern, this is not always the case, as there are multiple observational interferences that can affect the sampling of the signal

account for heterogeneity in the data downloaded from surveys due to each one having its own standards for storing information. In order to standardize the downloaded data, it has been decided that all light curves must include at least one column with the times (epochs) and
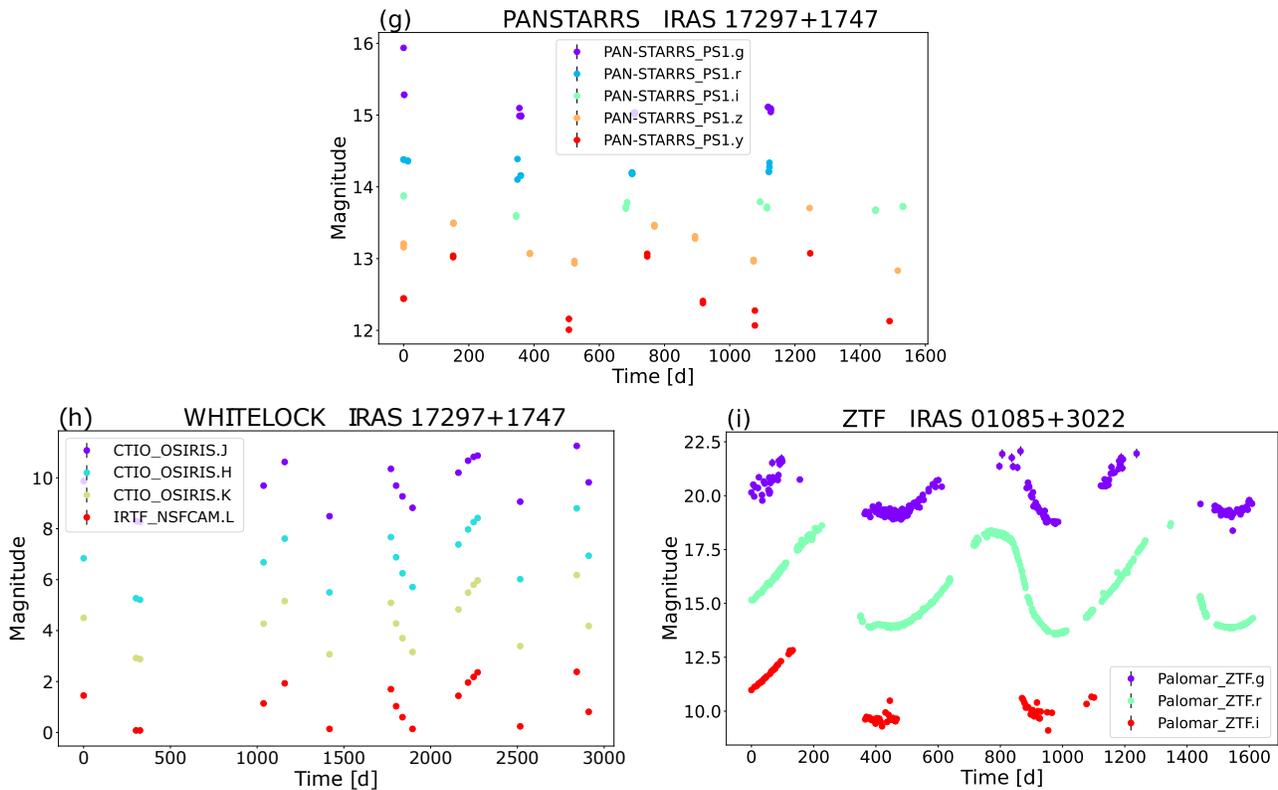
Figure 2.2: Three additional surveys with their respective light curves in all photometric bands for NESS sources.

another column with the corresponding magnitudes, and these magnitudes, regardless of the survey they come from, they must be in the Vega magnitude system.

This implies that transformations between magnitude systems or conversions from flux to magnitude must be performed, depending on the availability of the data. If the data includes information about fluxes, they are stored alongside the time and magnitude columns. Additionally, if uncertainties in the flux are available, they are added to the data. Furthermore, based on these data, the uncertainty in magnitude is estimated and included as another data column.

It is important to note that the data may contain empty cells, such as "Not a Number"

(NaN) values, or characters that do not correspond to numbers. To address this situation, the removal of NaNs and the necessary conversions between magnitude systems or fluxes will be carried out.

2.5.1   Removing NaN values and empty cells

Before being stored, all light curves from each of the surveys undergo analysis using lines of code (implemented in Python) that remove data rows containing empty cells or NaN values, both in the time columns and in the magnitude or flux columns. This process is repeated for light curves that require a conversion between magnitude systems or when magnitudes are calculated from flux values. In some cases, the uncertainties are unavailable or are set to NaN. We replace these with median values estimated from the points where uncertainties are available.

2.5.2   Missing parameter calculation

Since some surveys only have the flux reported for their detectors and the uncertainty associated with the measurement instrument, the apparent magnitude m and magnitude uncertainty σ_m parameters can be estimated using as parameters: detected flux F_sur, flux uncertainty σ_F and flux at zero point F_0^band as follows:

$$m = -2.5 \cdot \log\left(\frac{F_\mathrm{sur}}{F_0^\mathrm{band}}\right).$$

The uncertainty σ_m is then calculated according to

$$\sigma_\mathrm{m} = -2.5 \cdot \log\left(1 - \frac{\sigma_\mathrm{F}}{F_\mathrm{sur}}\right).$$

The reference flux can correspond to the Vega flux determined for each photometric band

and thus, the estimated magnitude will be in the Vega magnitude system. The Vega magnitude system is chosen to have a catalog of homogeneous light curves in this aspect. To get the fluxes at the zero point, another Pyphot function is used, which is called `get_pyphot_filter`. This function uses as a parameter the technical name of the filter reported in the (see Table 2.3) to return the flow in the desired units. For example: Jansky.

Magnitudes are placed on the Vega system as follows:

$$m_{\text{Vega}}^{\text{band}} = m_{\text{sur}}^{\text{band}} - 2.5\log_{10}\left(\frac{F_{0,\text{Vega}}^{\text{band}}}{F_{0,\text{sur}}^{\text{band}}}\right).$$

In the above equation, $m_{\text{Vega}}^{\text{band}}$ represents the magnitude being estimated in the Vega magnitude system, while $m_{\text{sur}}^{\text{band}}$ represents the magnitude obtained from the survey during the download. $F_{0,\text{Vega}}^{\text{band}}$ flux is the zero-point flux in the Vega system, which depends on the photometric band, and $F_{0,\text{sur}}^{\text{band}}$ is the flux at the zero point in the photometric band according to the magnitude system of the survey. The `get_pyphot_filter` function from the pyphot library allows obtaining zero point fluxes from the SVO website.

2.5.3 Segmentation of light curves into different photometric bands

To facilitate the processing of light curves, they are divided for each source according to their survey and photometric band. This means that in each data table, there is only one column of magnitudes corresponding to the same photometric filter. Therefore, in the best case scenario, a single target can have up to 29 light curves. The algorithms for separating the light curves into different photometric bands vary, as they depend on how they were stored. In some cases, they were already separated on the server from which the download was performed, while in others, it required a segmentation process based on the fact that magnitude-time pairs could be

differentiated with an identifier that refers to the filter in which this information was captured. At this point, the light curves are ready for processing.

2.5.4 Documentation of the file containing the resulting light curves

All light curves were stored in a single file with the extension .h5, which is created using the read and write functions of the Hierarchical Data Format 5 (HDF5) library available in Python. A light curve can be accessed by specifying its file address in the format: "«IRAS ID»/«Survey key»/«photometric band name»". In this context, the IRAS ID is the identifier from IRASPSC available in NESS. The survey key pertains to the survey's name in this study, provided in Table 2.3. The table also includes the names of the photometric bands. With this information, it is possible to access the light curve, provided it exists. Now, to query a specific data column, you only need to use the column's key name. Below are the key names for each column and their characteristics.With this information, it is possible to access the light curve, provided it exists. Now, to query a specific data column, you only need to use the column's key name. Below are the key names for each column and their characteristics.

- JD: Time in which each flux or magnitude was measured

- Magnitude: Apparent magnitude for its respective photometric band

- Magnitude_err: uncertainty associated with the apparent magnitude.

- Flux: Reported Flux Survey

- Flux_err: Uncertainty associated with Flux

From now on this database will be called NESS_LC.

2.5.5 NESS_LC repository

The codes for automatic downloads, data segmentation according to the photometric band, and pre-processing, as well as the NESS_LC database containing light curves, are hosted in a private repository on the GitHub platform. The repository will be made public once the article derived from this thesis is published. It is clarified that each code is duly documented and has examples so that an external user can search in the future using the strategies discussed in the previous subsection.

2.6 Summary of data obtained

After executing all the codes for downloading, preprocessing, and storing the light curves, 712 sources were found that have at least one light curve in at least one of the 29 photometric bands in the 9 consulted catalogs. For the other 74 sources that were not found in any catalog, there is a possibility that their optical emission is very weak due to the dust absorption effect. Another possibility is that some of these sources are outside the sky mapping range for groundbased telescopes. Additionally, it is possible that the used search radii are too small and do not allow the detection of these sources.

Table 2.13 lists, for each survey, the number of sources found with at least one light curve from the survey. As can be seen, the most common sources between NESS and the different surveys are NEOWISE and ASAS_SN. However, as mentioned before, the light curves of NEOWISE will not be taken into account for processing due to their high dispersion of information. The study with the fewest common objects is the Whitelock et al. study, which is evident given that this study was based solely on 239 carbon stars in the Galaxy [Whitelock et al., 2006].

Table 2.13: Number of NESS sources identified in each catalog consulted

ID in this thesis	Number of sources found
ASAS_SN	614
APASS_DR10	499
DIRBE	222
GAIA_DR3	400
KELT	77
NEOWISE	644
PANSTARRS	95
WHITELOCK	38
ZTF	117

In the Table 2.14, the number of light curves obtained for each photometric band (column: Number of LCs obtained) is shown. A total of 6774 light curves were successfully obtained. After visually reviewing the results of the light curve processing for parameter estimation, it has been decided not to consider those with fewer than 5 data epochs, as they are considered of low quality and provide little information. However, these light curves are still available in the NESS_LC database. To access them, in addition to the regular path, the string "_poor" should be added to the photometric band name. Therefore, the path will be: «IRAS ID»/«survey key»/«photometric band name +_poor».

When labeling light curves with limited information as unusable, a total of 5114 light curves containing at least 5 observation epochs are obtained. This information is also found in Table 2.14 (column: Number of usable LCs). Finally, by excluding the NEOWISE light curves, a total of 3838 usable light curves are obtained across the 27 photometric bands of the 8 surveys. These curves correspond to 700 sources from NESS, as 12 of them only have information in the two NEOWISE bands, and since these bands are not used, they don't have light curves to process.

Table 2.14: The table below presents the number of light curves obtained from the NESS sources in each catalog, as well as the number of light curves obtained for each photometric band in which a data record was obtained. Additionally, the table indicates the number of light curves that have at least five observed epochs and are therefore suitable for use in this study.

ID in this thesis	FIlter ID	Number of LCs obtained	Number of usable LC
ASAS_SN	V: Misc_APASS.V	614	614
APASS_DR10	B: TYCHO_TYCHO.B	323	66
	V: TYCHO_TYCHO.V	392	128
	g': SLOAN_SDSS.gprime_filter	373	108
	r': SLOAN_SDSS.rprime_filter	386	96
	i': SLOAN_SDSS.iprime_filter	386	101
	z: SLOAN_SDSS.zprime_filter	153	6
DIRBE	1.25 μm: COBE_DIRBE.1P25m	194	194
	2.2 μm: COBE_DIRBE.2P2m	209	209
	3.5 μm: COBE_DIRBE.3P5m	220	220
	4.9 μm: COBE_DIRBE.4P9m	221	221
GAIA_DR3	BP: GAIA_GAIA3.Gbp	400	400
	G: GAIA_GAIA3.G	400	400
	RP: GAIA_GAIA3.Grp	400	400
KELT	R_k: OSN_Johnson.Cousins_R3 (_0, _1)	77	77
NEOWISE	$W$1: WISE_WISE.W1	644	638
	$W$2: WISE_WISE.W2	644	638
PANSTARRS	g: PAN-STARRS_PS1.g	83	67
	r: PAN-STARRS_PS1.r	87	67
	i: PAN-STARRS_PS1.i	89	67
	z: PAN-STARRS_PS1.z	78	44
	y: PAN-STARRS_PS1.y	72	49
WHITELOCK	J: CTIO_OSIRIS.J	38	29
	H: CTIO_OSIRIS.H	38	33
	K: CTIO_OSIRIS.K	38	34
	L: IRTF_NSFCAM.L	38	34
ZTF	g: Palomar_ZTF.g	96	95
	r: Palomar_ZTF.r	60	59
	i: Palomar_ZTF.i	21	20

Finally, Figure 2.3 presents a histogram showing the quantity of light curves as a function of the number of observed epochs. From this graph, it can be concluded that there is a tendency to have light curves with a low number of epochs (less than a thousand), as well as light curves that reach over 9000 observation epochs, specifically those from the KELT survey, followed

by ZTF and ASAS_SN. Meanwhile, light curves with a smaller number of data points are APASS_DR10 and PANSTARRS.
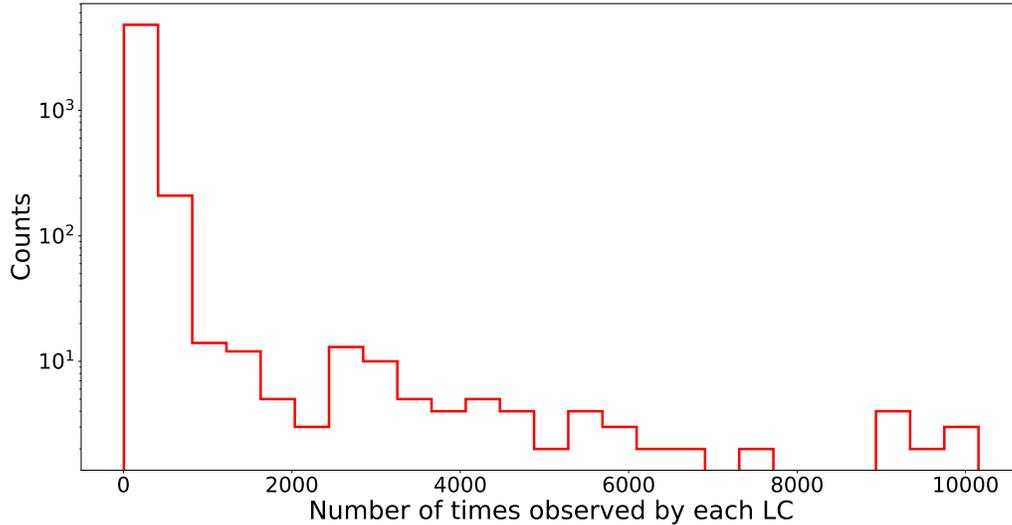
Figure 2.3: Histogram showing the number of light curves as a function of the number of observed epochs. A trend is observed towards having fewer epochs in the entire sample of light curves obtained

2.7   Quality of the light curves

Up to this point, the data collection and preprocessing phase has been completed, allowing for the creation of a consistent catalog of light curves and their proper storage. In this thesis, data refinement has not been carried out in order to improve the desired results, such as the application of smoothing filters to reduce the noise present in the light curves. However, we analyzed the quality of the light curves as follows. For each curve, we computed the number of unique time stamps (as time values are sometimes repeated for some light curves), N_uniq_pts_LC, and compared this to the maximum time separation between consecutive epochs, max_timesep_LC (this quantity accounts for differences in sampling), divided by the total time baseline baseline_LC of the light curve (baseline_LC is the difference between the

latest and earliest times at which observations were made). This relationship is presented in Figure 2.4. Among the optical surveys, the ASAS-SN, Gaia, and ZTF data have the best sampling. Similarly, in the infrared, the DIRBE light curves are well-sampled. Most of the APASS and PAN-STARRS light curves have fewer than 10 points and are also the most irregularly-sampled. In the future, diagnostic figures like this one will be used to provide valuable constraints on the required observation cadences for AGB variability studies.
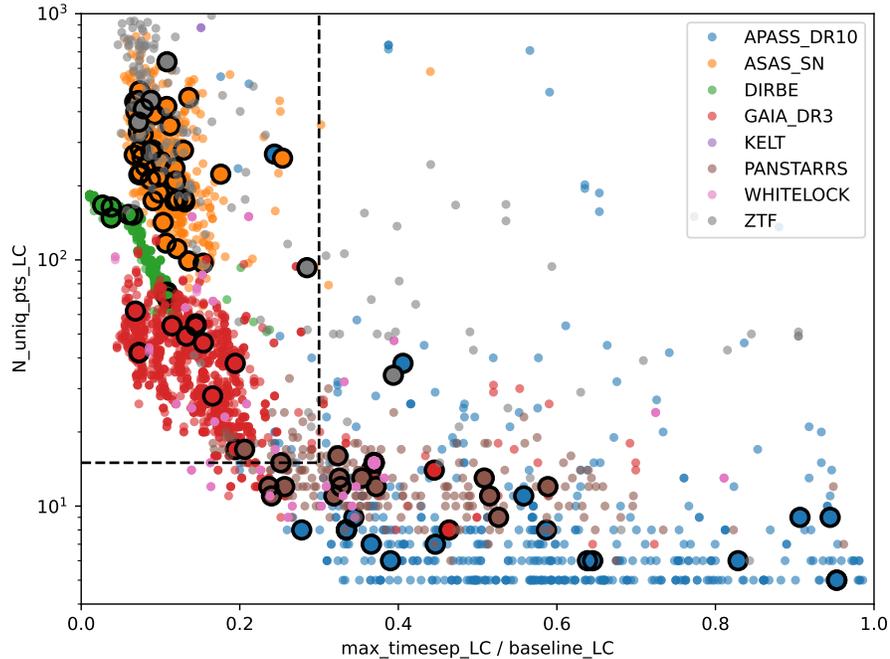
Figure 2.4: The number of unique time stamps (epochs) in a given light curve as a function of the maximum separation between consecutive epochs, relative to the total time baseline. The highest-quality light curves are those that fall in the rectangular region at the top left of this figure. Symbol colors represent various surveys for which light curves were collected. Larger symbols represent sources whose predicted amplitudes differ from the values estimated based on their light curves by more than 5 mag. See text for details.

From the PGMUVI models (as explained in section 3.2.4), peak-to-peak amplitudes are subsequently calculated. Significant disparities between the predicted and observed amplitudes

are anticipated due to various factors. For instance, the ZTF, APASS, and ASAS-SN surveys are performed in similar bands, but have very different saturation, completeness, and confusion limits. Saturation at the light-curve maximum and/or non-detection of the light-curve minimum could result in an underestimate of the amplitude, which would in turn result in a larger discrepancy with the predicted amplitude. In future work, a systematic visual inspection of the light curves will be carried out to determine if sources that are saturated or lost in some surveys have poorer model fits in those surveys. For now, the sources with a discrepancy of more than 5 magnitudes between their predicted and estimated amplitudes in Figure 2.4 were identified using larger symbols.

It is important to note that we expect light curves from a given survey to be located close to each other on this diagram because of their comparable time baselines as a result of the study's design. However, the maximum separations between consecutive epochs will vary depending on the characteristics of the telescope used for monitoring and the geographic location of the telescope (such as APASS, ASAS-SN, PAN-STARRS, KELT, Whitelock et al., and ZTF), which may result in intervals without data records for some sources. Additionally, in the case of space telescopes (COBE-DIRBE and Gaia), the maximum observation time will be determined by the telescope's orbit properties, meaning the telescope may repeatedly point to the same region of the sky at certain time intervals.

For this analysis, only light curves containing five or more data points have been considered. However, as illustrated in the figure, a more conservative criterion to select the highest-quality light curves may be to restrict their maximum time separation to under 30% of the total baseline, while requiring at least 15 unique epochs in their light curves (corresponding to the region delineated by the dashed lines in Figure 2.4). This is because, even though some of these curves

may have a substantial number of data points, they may also exhibit significant gaps in their sampling time. Although all light curves in the study will be processed, this preliminary work will help identify those light curves with higher quality, which will be crucial for subsequent analysis.

Period estimation

In this chapter, key concepts and implementation details of two methods used to determine the parameters that characterize light curves are discussed. These parameters include the fundamental pulsation period, amplitude and mean magnitude. The first method discussed is Lomb-Scargle periodograms, while the second method is Gaussian process modeling with the Gaussian Processes for Multiband Variable Inference (PGMUVI) algorithm. We use the periods obtained from the Lomb-Scargle run as initial guesses for our PGMUVI run. Both the Lomb-Scargle algorithms discussed below accept uncertainties in the magnitudes as input; however, this makes the code more computationally intensive. As we are only interested in obtaining initial guesses for the period, we do not pass uncertainties to the Lomb-Scargle algorithms.

This study then presents the statistical analysis derived for 3838 light curves of the 700 NESS targets that contain sufficient information for processing with the above methods.

3.1   Lomb-Scargle algorithm

Periodograms are a valuable tool in the analysis of light curves from astronomical sources, as they allow the determination of source variability and the identification of different pulsation modes of the objects, recovering their characteristic periods. A classical approach to this is the Schuster periodogram [Schuster et al., 1898]. The power $P(f)$ as a function of the frequency f is given by the square of the magnitude of the Fourier transform:

$$P(f) = \frac{1}{N} \left| \sum_{n=1}^{N} g_n e^{-2\pi i f t_n} \right|^2.$$

(3.1)

In this equation, g_n represents the measured variable at time t_n, where $t_n = n\Delta t$. Here, Δt represents the regular time sampling interval and n takes values from 1, 2, 3,..., N. N, in turn, corresponds to the number of observations used to construct the light curve. Two important conditions here are that the sampling is constant and that the mean value of the light curve is zero.

An important feature of light curves of astronomical objects is that it is difficult to guarantee a constant sampling interval. Various phenomena can affect observations, such as weather conditions interfering with ground-based observations, periods when astronomical objects are not visible due to solar occultation and the possibility that the implemented algorithm may not be able to identify these occultations during observation detection. Additionally, it may happen that, for some epochs, no data is found within the specified search radius, which would cause an irregularly-sampled light curve. This implies that the tool used must be capable of constructing periodograms using light curves with observed data at irregular time intervals. The LS method excels in this task as it presents a generalization of the classical Fourier transform. This method allows working with irregular samplings, is insensitive to global time shifts in the

data, allows the user to incorporate uncertainties and, in its most generalized version, does not require pre-processing of the light curve to obtain a mean value equal to zero. We refer the reader to Lomb [1976] and Scargle [1982] for details. Nowadays, LS algorithms found in libraries like Astropy in Python use a generalized version of LS that can perform fitting for light curves whose mean value is different from zero.

3.1.1 False alarm probability

Noise embedded in the data of a light curve can generate peaks in the periodogram that appear to be part of the signal due to statistical coincidences. To address this problem, there is a tool called False Alarm Probability [FAlP; Scargle, 1982]. Its aim is to discriminate peaks in the periodogram that are false positives, meaning those peaks that are misclassified as part of the signal when they are actually the result of noise. The FAlP determines the minimum power required, given a probability value, for a peak in the periodogram to be considered statistically significant and classified as part of the real signal, rather than being considered noise [Scargle, 1982]. In this work, we choose a significance level (probability threshold) of 0.05. The `LombScargle` function is built into the `astropy.timeseries` module and it takes a power value as an argument and returns a probability value.

3.1.2 Multiband Lomb-Scargle periodogram

Thanks to the work carried out by VanderPlas & Ivezić [2015], a method has been developed that allows the combination of light curves from multiple photometric bands corresponding to the same object. This multiband LS (MLS) method represents an even greater generalization of the well-known LS periodogram.

The algorithm that implements MLS is found as the library function `LombScargleMultiband` in the module `gatspy.periodic` of the Gatspy package. It has functions similar to LS in Astropy, but they have not yet developed the FAlP function due to the complexity of the models. The `LombScargleMultiband` function requires three one-dimensional arrays as input: the array of epochs, the fluxes corresponding to each epoch, and a unique string identifier denoting the wavelength band in which the flux was obtained. This allows us to easily concatenate the light curves of multiple bands together and feed it to the function.

3.1.3   Implementation of LS and MLS

Using LS and MLS, it is possible to obtain the fundamental mode present in the various light curves. In this study, a Python code was developed to generate periodograms and analyze them to determine the fundamental period of NESS stars. Additionally, the ocation and value of the full width at half maximum (FWHM) around the fundamental period was obtained. These three parameters are relevant for PGMUVI training and also serve as reference values for comparison with the results obtained by PGMUVI.

Figure 3.1 illustrates the flowchart of the algorithm implemented. Essentially, the algorithm focuses on determining the number of available light curves for each star and subsequently decides whether to construct the periodogram using LS or MLS. If a star has only one light curve, its data will be processed using LS. In this case, FAlP is employed to determine if the peak with the highest power meets the minimum threshold required to be classified as a signal rather than noise. On the other hand, if a star has more than one light curve, they are concatenated as explained in the previous section before being used as input for the MLS procedure.

The algorithm utilizes the function `findpeaks` from the findpeaks library to identify all the
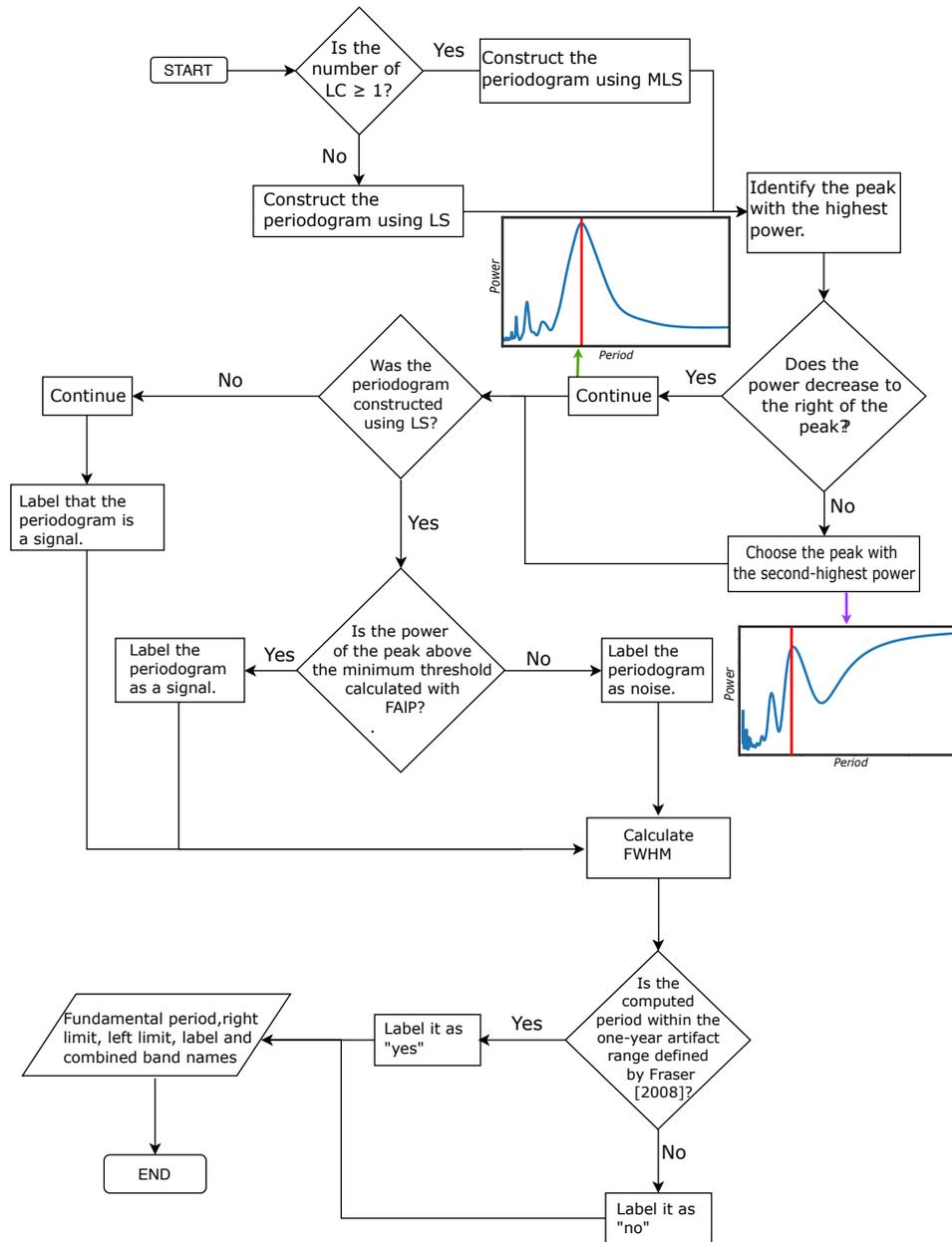
Figure 3.1: This flowchart presents the algorithm developed to identify fundamental modes in periodograms generated through LS-MLS, along with key steps for distinguishing power peaks. The algorithm's implementation was carried out using various functions and libraries in Python. The periodograms included in the diagram are provided to aid the reader in understanding peak selection. For instance, the periodogram marked with a purple arrow displays a peak without an immediate decrease to its right, unlike the one marked with a green arrow, which does exhibit such a falloff.

local maxima and minima [1] of the periodogram. Then, the peak with the highest power is selected. However, it is important to consider certain special cases. Sometimes, the peak with the highest power does not show a well-defined minimum to its right (i.e., its low-frequency end). This typically occurs for peaks at the highest periods. An example these cases can be seen in the periodogram of Figure 3.2.a. In such cases, the initially selected peak with the highest power is discarded and the second peak with the highest power is chosen as the result. This is done to avoid selecting peaks which could indicate the presence of noise rather than a significant signal.

Another task of this algorithm is to determine if the estimated fundamental period falls within the interval of 338.84 to 398.10 days, referred to in this thesis as the "Fraser interval," as identified by Fraser et al. [2008]. Periods within this interval have the potential to be considered as noise generated by the data's sampling frequency rather than the nature of the star. An example of a power peak falling within the Fraser interval is shown in Figure 3.2.d. (highlighted in the red region in the figure). This is due to the presence of the one-year artifact [2] in the data that causes peaks in the periodogram that are actually related to the sampling periodicity, as mentioned in the work of Fraser et al. [2008].

Finally, the function `peak_widths`, from module `scipy.signal` from the SciPy library was used to calculate the relative full width at half maximum. It is important to note that in this calculation, the mean power is not used to determine the half maximum height. Instead, one of the surrounding minima with the highest power is used. An example of this is shown in panels b. and c. of Figure 3.2.

---

[1]The `findpeaks` function selects the last value of the ordinate as a minimum, as long as it detects a defined peak beforehand, meaning those peaks whose amplitudes are significant

[2]The one-year artifact is a noise signal present in the power spectrum of a light curve, originating from the sampling cadence of the survey. In some cases, the power of this noise signal can even be greater than that of the fundamental mode signal, or it can combine with the fundamental mode and shift the signal peak.
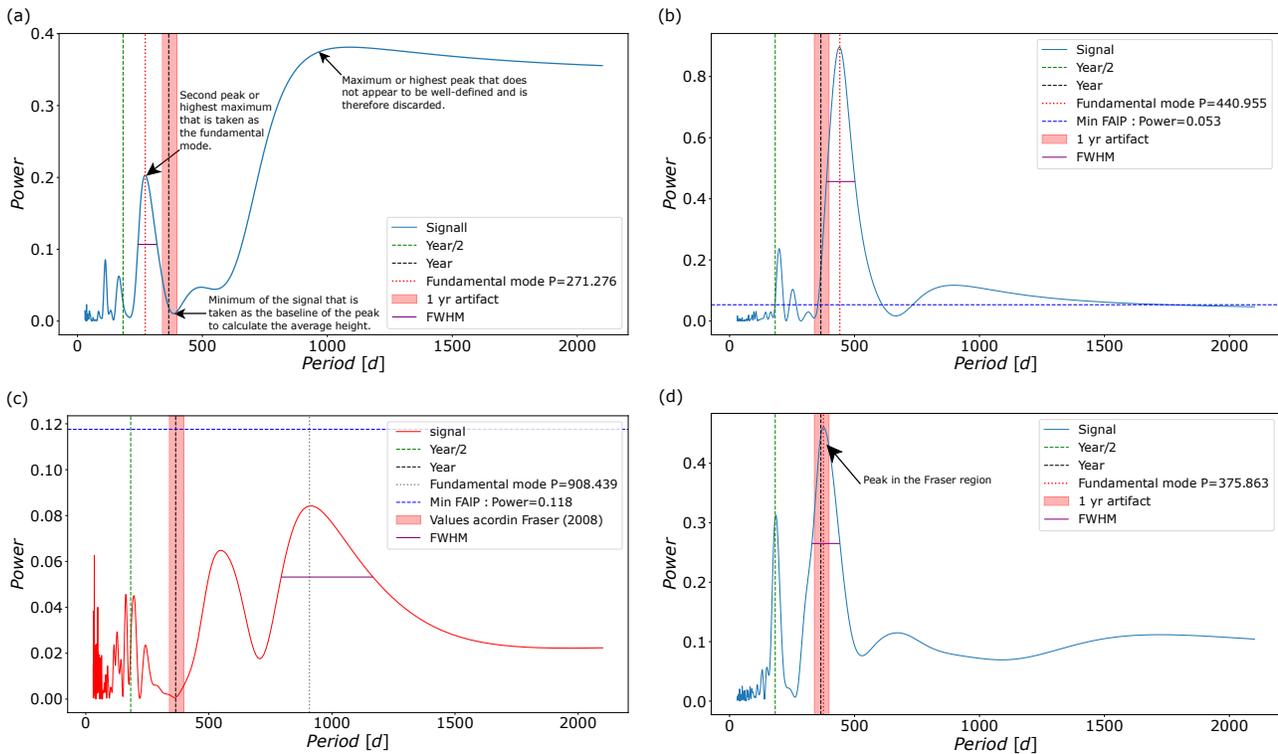
Figure 3.2: a): periodogram for IRAS 09147-5719, flagged with 30, built using MLS and combining 5 light curves (LC). b): periodogram for IRAS 10403-7612, flagged with 10, constructed with LS and containing only 1 LC. c: periodogram for IRAS 10193+4145, flagged with 20, built using LS and having 1 LC, but it is identified as noise. d): periodogram for IRAS 14390+3147, flagged with 31, built using MLS and combining 4 LC. The main peak falls within the Fraser region.

We encode a flag column to summarise the results of the application of LS and MLS to each source (see Table 3.1). The flag is set to 0 if no fit was possible for a given source. There are two such sources in our data. The remaining flag values have two digits. The least significant digit is set to 0 if the fundamental peak falls within the one-year artifact, and to 1 if it is outside this region. If the periodogram was calculated using the LS method, the most significant digit is set to either 1 or 2 depending on whether the fundamental peak was below the FAIP or above it. If the periodogram was calculated using the MLS technique, the most significant digit is set to 3. In the MLS case, the FAIP check has not yet been implemented.

Table 3.1: Flag value assigned to period determined from LS and MLS. Flag values of 10 and 30 represent the ideal cases, where the determined period is statistically significant and is not part of the 1-year artifact.

Flag	Description	Condition	Periodograms with this flag
0	Error in data	-	2
10	Periodogram by LS	Peak not in Fraser interval	84
11	Periodogram by LS	Peak in Fraser interval	8
20	Periodogram by LS	Peak not in interval, considered noise by FAlP	32
21	Periodogram by LS	Peak in Fraser interval considered noise by FAlP	3
30	Periodogram by MLS	Peak not in Fraser interval	495
31	Periodogram by MLS	Peak in Fraser interval	50

3.1.4 Results of the implementation of LS and MLS

In summary, this algorithm returns the period with the highest power and the location and value of the FWHM, identifies the photometric bands used to generate the periodograms and provides a flag indicator that distinguishes between different cases. The different cases are shown in Table 3.1.

Without distinguishing by flags, the data from Table 3.1 comprises a total of 672 periods obtained from LS and MLS of the initial 700 objects in NESS that possessed at least one light curve, which are assumed to be fundamental pulsation modes extracted from the periodograms. Of these, 579 have well-determined fundamental periods from either the LS or MLS method that are distinct from the 1-year artifact. Thirty two periods are discarded because their peaks do not meet the minimum power criteria to be considered as a signal. Nonetheless, this information is preserved for comparison with the results from PGMUVI.

Figure 3.3 shows a probability density function (PDF) representing the distribution of

obtained periods, a broad distribution with two apparent components, the first centered at approximately 180 days and the second at 420 days. The first peak could be attributed to the presence of a second artifact with a periodicity equivalent to half of an Earth year, as some periodograms present dominant peaks around the midpoint of the year. The second peak could represent a general trend in the pulsation mode of the NESS stars, which could be contaminated by the annual artifact. This PDF was generated using the kernel density estimation technique [KDE; Parzen, 1962], a statistical method employed to estimate and represent the shape of a data distribution without the need to predefine the number of intervals (bins) in a histogram. KDE produces smoothed non-parametric[1] distributions, enabling a more accurate visualization of the obtained distributions.
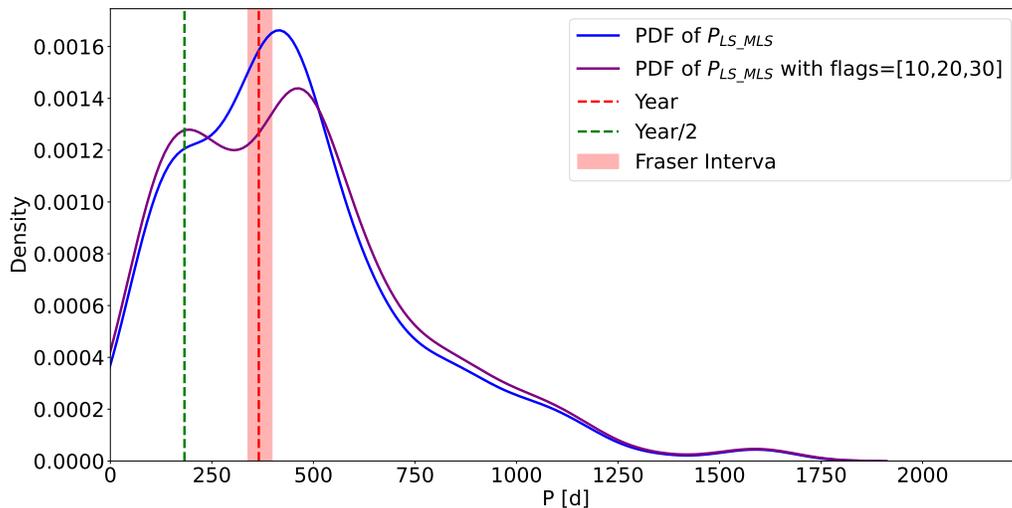
Figure 3.3: The PDF shows the distribution of the periods obtained using the LS and MLS methods. In the distribution that includes all LS and MLS period values (blue line), two peaks are observed. The first one is located at the value corresponding to the semi-annual artifact (green line), and the second peak is at 417 d. In the second distribution (purple line), sources with flags 11, 12, and 31 are excluded, and only the distribution of period values outside the Fraser interval is presented. In this distribution, it is noticeable that the second period peak shifts to 465 d.

---

[1]See the next section for a discussion of the difference between parametric and non-parametric models.

Examples of the results obtained after implementing LS and MLS are shown in Table 3.2. For each source, the table shows the fundamental period determined by either the LS or MLS method, the corresponding locations that define the FWHM, the flags and the number of combined bands along with the bands that were combined.

Table 3.2: Fundamental periods and locations that define the FWHM around these periods after the implementation of LS and MLS. The flag associated with each determination, as defined in Table 3.1, is also shown, as are the total number of LCs that were combined to compute the period and the bands in which these LCs were measured. Only 10 rows of the total 672 are shown here.

IRASPSC	$P_{LS,MLS}$ [d]	P_{right} [d]	P_{left} [d]	Flag	Base line time [d]	No. of combined LC	List of combined LC
00042+4248	747.61	604.07	948.8	30	12013.82	7	Johnson_V_band|COBE_DIRBE.1p25m|COBE_DIRBE.2p2m|COBE_DIRBE.3p5m|COBE_DIRBE.4p9m|Palomar_ZTF.g|Palomar_ZTF.r
00084-1851	387.25	340.01	439.96	31	1834.99	4	Johnson_V_band|GAIA_GAIA3.Gbp|GAIA_GAIA3.G|GAIA_GAIA3.Grp
00121-1912	380.31	349.28	435.89	11	1834.99	1	Johnson_V_band
00192-2020	164.62	153.87	183.72	30	10501.56	5	Johnson_V_band|COBE_DIRBE.1p25m|COBE_DIRBE.2p2m|COBE_DIRBE.3p5m|COBE_DIRBE.4p9m
00193-4033	691.7	573.77	1014.67	30	910.83	3	GAIA_GAIA3.Gbp|00193-4033/GAIA_DR3/GAIA_GAIA3.G|GAIA_GAIA3.Grp
00205+5530	445.24	387.97	532.22	30	10576.78	5	Johnson_V_band|COBE_DIRBE.1p25m|COBE_DIRBE.2p2m|COBE_DIRBE.3p5m|COBE_DIRBE.4p9m
00213+3817	404.86	378.59	441.07	30	4415.14	2	Johnson_V_band|Johnson_R_band_0
00245-0652	153.23	146.51	159.2	30	2216.85	4	Johnson_V_band|GAIA_GAIA3.Gbp|GAIA_GAIA3.G|GAIA_GAIA3.Grp
00247+6922	658.56	521.44	896.29	30	985.62	3	GAIA_GAIA3.Gbp|GAIA_GAIA3.G|GAIA_GAIA3.Grp
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
R09061-4313	301.23	258.37	344.11	10	965.34	1	Johnson_V_band

3.2 PGMUVI

The Lomb-Scargle periodogram is a valuable tool for the analysis of periodicity in light curves and has been used since the 1970s to determine the modes and pulsation periods of variable sources. However, it has some drawbacks – it is sensitive to outliers in the data, which prevent it from fitting light curves that are poorly sampled or have saturation/non-detection issues. The periodogram is better suited for fitting sinusoidal light curves. Moreoever, while it estimates the parameters of the light curve, it cannot produce a model of the light curve itself. Gaussian-process based techniques are built to handle exactly these issues. For this reason, PGMUVI is

better suited for the aims of this thesis. Using the non-parametric models provided by PGMUVI, it is possible to obtain information in the time intervals where the light curve was not sampled. Additionally, it estimates magnitude values that were not recorded or that fall outside the operational limits of the sensors in each survey, as it can make predictions based on the data. The virtue of the kernel it implements allows for modeling the number of pulsation modes of the star and returns parameters that characterize the light curves (explained in the next section), such as periods, weights of each mode, pulsation period, and mean magnitude.

As mentioned above, LS and MLS methods assume a sinusoidal model for the light curve, and they solve for the associated parameters (period, mean magnitude, amplitude). An alternative method is to fit the light curve using a family (or distribution) of functions, such that the combination of these functions is able to reproduce the observed shape. With no prior knowledge of the parameters, this family of functions is infinite-dimensional (for each value of time, there are infinitely many functions that are combined to reproduce the observed brightness) and is referred to as a *non-parametric* model [Roberts et al., 2013]. If we had a strong prior for the parameter values, then we could reduce the dimensionality of the family down to a finite value, thus turning it into a *parametric* model. In other words, while a parametric model might quantify the relationship between the time and the brightness in the form of an analytical function, a non-parametric model instead draws functions from an infinite-dimensional set and combines them in such a way that they reproduce the observed brightness at each available time.

Gaussian-process techniques are non-parametric methods. PGMUVI thus fits non-parametric models to the data. In the following section, basic information about GP will be discussed, the kernel used by PGMUVI will be described, the computational implementation will be explained and finally, the achieved values up to this point will be discussed.

3.2.1 Basics of Gaussian processes

Gaussian processes are an infinite collection of random variables, in which a subset of them has a joint probability distribution that is Gaussian. A Gaussian process can also be understood as a realization over a distribution of random functions [Rasmussen & Williams, 2006]. Mathematically, this is expressed as:

$$f(\mathbf{t}) \sim \mathcal{GP}(m(\mathbf{t}), k(\mathbf{t}, \mathbf{t}')), \tag{3.2}$$

where the function $m(\mathbf{t})$ represents the mean value at each of an array of times $\mathbf{t}$, which is generally modeled as zero or with a constant value for simplicity. Additionally, a covariance matrix $k(\mathbf{t}, \mathbf{t}')$ is used. In Figure 3.4, the most important steps to go from observational data to a non-parametric model are shown, explaining the algorithm involved. To achieve this, the Gaussian process requires several steps and ingredients:

(a) The random variables that are desired to be modeled are defined, which in this case are the data from the light curves (magnitude vs. time). The magnitude may contain uncertainties that are assumed to behave like Gaussian noise.

(b) The Gaussian process is modeled by selecting a mean function and a correlation function (these choices will be explained later).

(c) The hyperparameter[1] inference is performed using Bayesian inference and optimization. Bayes' Theorem is used to combine the prior distributions for the hyperparameters with the likelihood computed from the data and the model including the GP kernel to derive

---

[1]The hyperparameters in a Gaussian process can be physical quantities of interest depending on what is being modeled. These control both the mean function and the kernel function and exist in a space where they take on different values. After optimization and Bayesian inference, they return as the best parameters that describe the Gaussian process. These are the values of interest for this thesis, as in the kernel, as will be explained later, they can be the pulsation periods.
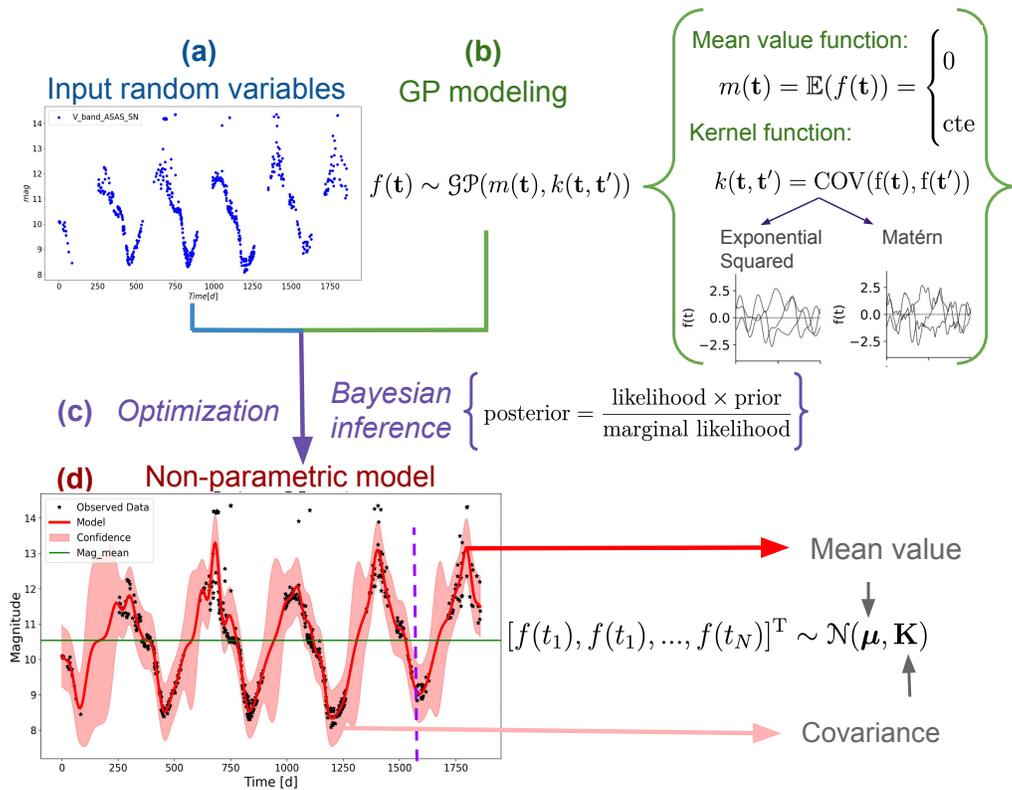
Figure 3.4: The diagram represents the structure of a pseudo-algorithm for the selection of non-parametric models using Gaussian processes. In step (c), a high level of programming expertise is required to identify the most effective numerical methods for minimizing computational expenses. This step involves numerical integration and multiple iterations, which consume substantial resources when compared to traditional methods for studying light curves, such as LS. Additionally, in the figure, an example of kernels typically used in GP, the square exponential kernel and the Matern kernel, are shown [Rasmussen & Williams, 2006].

the posterior probability distribution for each hyperparameter. The latter distribution is then used to find the most probable value as well as the credible interval for each hyperparameter [see, e.g., Rasmussen & Williams, 2006, for more details].

(d) Finally, a probability distribution is obtained from a multivariate Gaussian in the time domain.

PGMUVI allows the user to select a mean function according to their convenience, either by selecting a function equal to zero or a constant function [Roberts et al., 2013]. On the

Basics of Gaussian processes

other hand, the choice of the correlation function is crucial as it determines how the data are correlated with each other. It can also incorporate known physical information of the problem without inferring its functional behavior beforehand. This allows for the incorporation of prior knowledge and can improve the model's performance in capturing the underlying patterns.

For example, if it is known that the underlying process is linear, the exponential kernel can be used, which strongly correlates nearby points and decays exponentially for more distant points [Aigrain & Foreman-Mackey, 2022 and Rasmussen & Williams, 2006]. However, sampling can affect how these correlations are constructed, so kernels that vary more rapidly can also be chosen to improve correlation construction. Some additional examples of kernels are: the exponential quadratic and Matérn. These correlation functions depend on hyperparameters that are not known beforehand and must be obtained through Bayesian inference.

PGMUVI specifically implements a static kernel called the spectral mixture kernel [SM kernell; Wilson & Adams, 2013]. This kernel consists of an exponential part and a sinusoidal part, where the exponential part is responsible for correlating nearby points and decaying exponentially for more distant points, while the sinusoidal part models the correlation for distant neighbors due to the oscillatory behavior of light curves. Figure 3.5 illustrates an example of how the SM kernel works and show the SM kernel function.

In PGMUVI, Bayesian inference is performed using the spectral density of the SM kernel. This spectral density is obtained by taking the Fourier transform of the kernel, allowing the researchers to employ a Gaussian mixture model to model the power spectral density. Each Gaussian in this mixture has its hyperparameters, which consist of the weight w_q, the scale factor v_q and, most importantly for the study, $1/\mu_q$, corresponding to the Q periods present
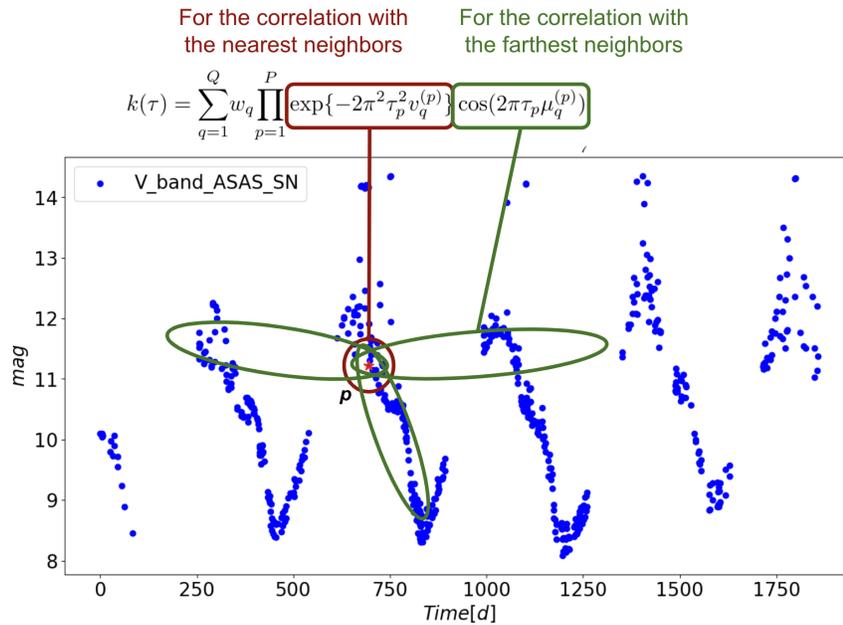
Figure 3.5: The image illustrates how the SM kernel correlates the evaluation point within the light curve with its nearby and distant neighbors. In the case of the equation, the variable τ causes the kernel to be static (making the covariance time-invariant) in such a way that $\tau = t - t'$.

in the light curves [Wilson & Adams, 2013]. These hyperparameters are obtained through the inference process.

In this point, following the topic of Bayesian inference, it is allowed to obtain the hyperparameters through the implementation of Bayes' theorem, which provides the posterior probability distribution based on the prior probability distribution, the likelihood and the marginal likelihood. Once the model is set up and the method for obtaining hyperparameters is known, an optimization algorithm is required to maximize the likelihood. By achieving this maximization, fits are made to the spectral density to obtain the parameters of the intrinsic set of hyperparameters that lead to this maximization and provide the best fit [Aigrain & Foreman-Mackey, 2022]. This optimization stage is one of the innermost layers of PGMUVI

and the details of the numerical processes implemented are not the focus of this thesis, so, for now, it is treated as a black box. Some details on this optimization can be found in [Scicluna et al., 2023].

3.2.2 Implementation of PGMUVI

We specify initial guesses for the period and mean magnitude parameters for the PGMUVI fitting procedure. We use the value for the fundamental period estimated from the implementation of LS and MLS (values from Table 3.2). As PGMUVI fits two components to each light curve, we provide half the estimate for the fundamental period as an initial guess for the second component. These initial guesses are specified for fits to all the light curves of the same star regardless of wavelength band. Meanwhile, the median of the magnitudes from the data in each band is employed.

The implementation of this algorithm is simple, as it features the PGMUVI class that defines the `lightcurve` object. Thanks to the various functions available in this class, it is possible to efficiently obtain and provide parameters derived from this object. Figure 3.6 displays one-component fits computed for four different bands, showing a periodic pulsation pattern. Upon visual inspection, it is evident that the obtained mean magnitude consistently aligns with the light curves. Something relevant about these models is that upon visual inspection of Figure 3.6 for the G band of Gaia, for example, it is observed that the peak-to-peak period with good sampling is consistent with the period values for the other three models. Thus, it can be noted here that in the portion of the light curve (times greater than 90 days), those sparse data points impact the final estimation. It can be concluded that sparse and poor sampling affects the parameters returned by PGMUVI. This is discussed later in the discussion of the results.

79
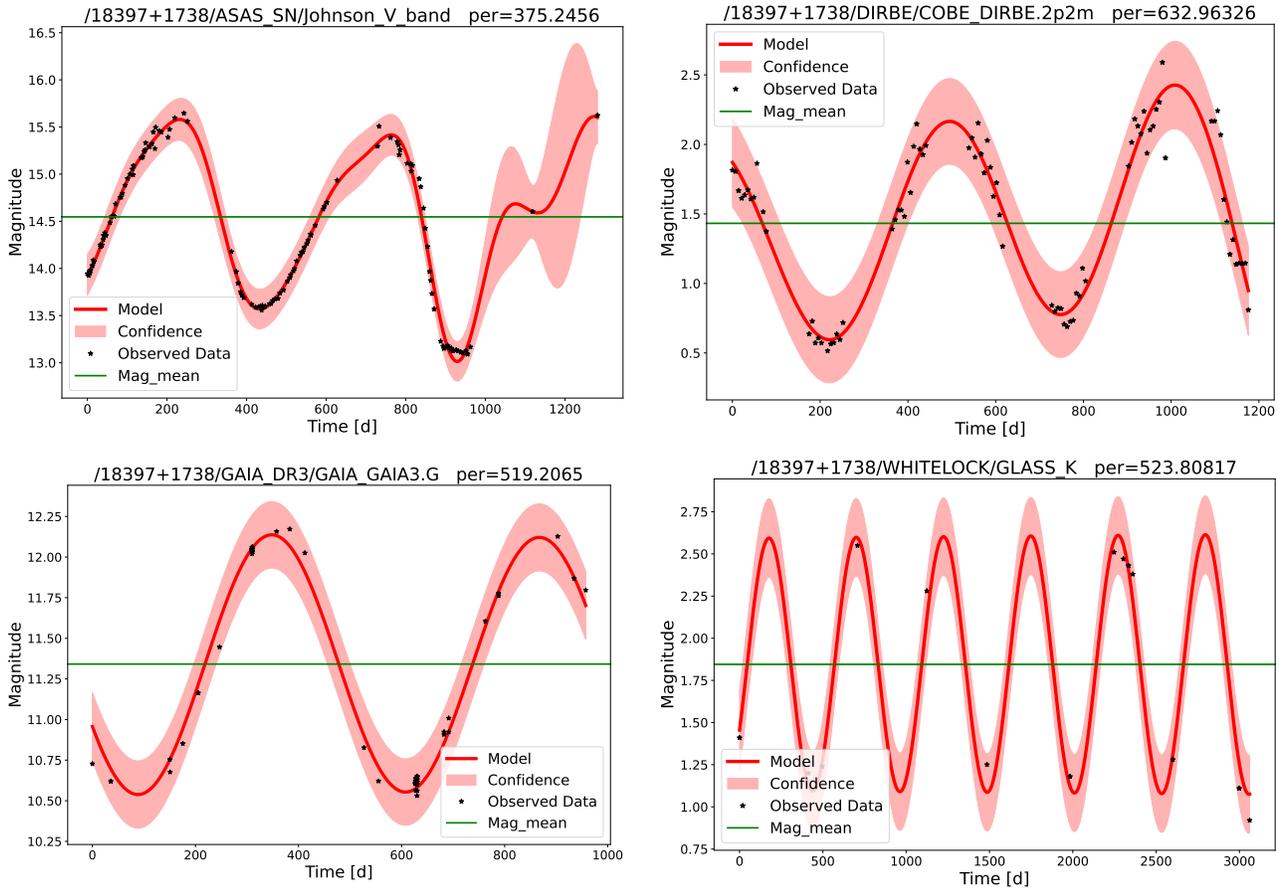
Figure 3.6: PGMUVI was applied to each of the 16 light curves of the source IRAS 18397+1738 in different optical and infrared bands. Four models are shown, demonstrating a good fit and the period values (labeled as "per" in the upper right corner of each subplot) are close to the training value of 508.41 days (except for the results of figure a, which is consistent with this period at earlier times, but which struggle to fit the final, poorly sampled, ∼200 days). This is an example of the multiple fits that were performed for each light curve in the database of this study.

In contrast, Figure 3.7 shows an example of a source with bad quality fits to some of its light curves, while others are fit well. The bad fits can be attributed to the limitation of available data, the need to increase the number of iterations in the fitting process, or the possibility that certain parameters were initialized incorrectly. Considering these factors is of utmost importance to enhance the quality of the fits in future analyses.
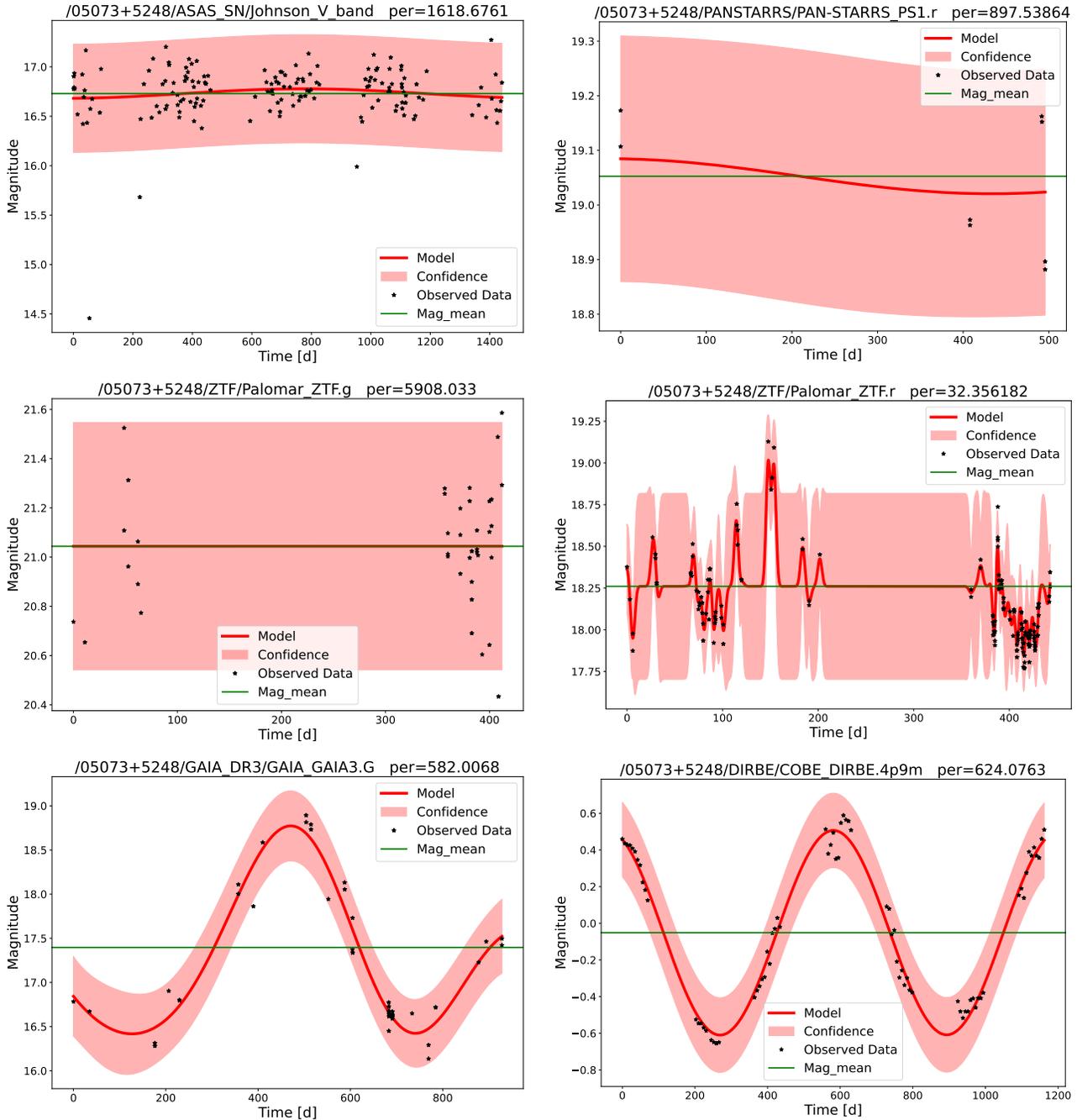
Figure 3.7: For the source IRAS 05073+5248, four light curves (4 top panels) out of the 15 available exhibit poor fits. Although the average magnitude (green line) appears to be consistent, no clear patterns indicating periodicity are observed in the model. This is likely due to the limited amount of data, inadequate sampling, or incorrect parameter initialization. Additionally, two light curves from Gaia or DIRBE (two bottom panels) are shown, where better results are observed for the same source. It is worth noting that the periods of each model are very different from the initial value of 638.8 days.

Implementation of PGMUVI

It is anticipated that the evolved objects from the NESS sample will manifest as fundamental mode pulsators. This implies that a successful prediction of their variation and accurate estimation of their periods can be achieved through a single-component fit. Nonetheless, the sample also encompasses semi-regular variables and is expected to include a notable count of overtone pulsators and/or sources exhibiting multiperiodicity. In such instances, a one-component model fit might not yield realistic period estimations.

Since the PGMUVI algorithm involves an optimization process, it is important to highlight that multiple executions under the same initial conditions can produce different combinations of optimal parameter settings. This phenomenon is subsequently illustrated in the example that presents fits of 1 and 2 components. Therefore, it is crucial to conduct a thorough sampling of the parameter space to identify the combination corresponding to the global minimum.

The Markov Chain Monte Carlo [MCMC; Speagle, 2019], is the appropriate tool to address this task. However, it is important to note that MCMC entails a high computational cost and its feasibility within the PGMUVI algorithm, and this tool will be implemented in future versions of the code. In the context of this thesis, the approach focuses on calculating a fit for each light curve in the sample and analyzing the statistical properties of the optimal parameter settings. In future work, the intention is to perform fits using MCMC for each light curve. Currently, we employ a simpler approach on a single light curve to emphasize the need for a 2-component fit, as explained below.

A comparison was conducted between one-component and two-component fits to the light curve of the source IRAS 04307+6210 within the survey ASAS_SN in band Misc_APASS.V (as illustrated in Figure 3.8). Evidently, the light curve displays multiperiodic behavior, with

a longer period of approximately 500 days. One hundred single-component adjustments to the light curve were calculated and the median period was determined. The same procedure was repeated with the 2-component model, yielding median values for both best-fit periods. It should be noted that the fitting procedure does not distinguish between the two components in terms of designating one as primary and the other as secondary. Consequently, there can be degenerate parameter combinations which are permutations of each other. For instance, two different attempts to fit a light curve might yield periods of (300 days, 100 days) and (100 days, 300 days), respectively. To address this degeneracy, post-processing is applied to the results, identifying the longest and shortest periods per fit based on the outcomes of the two components. When computing the median periods of the 2-component model, we ensured that we always selected the longer period as the "primary".
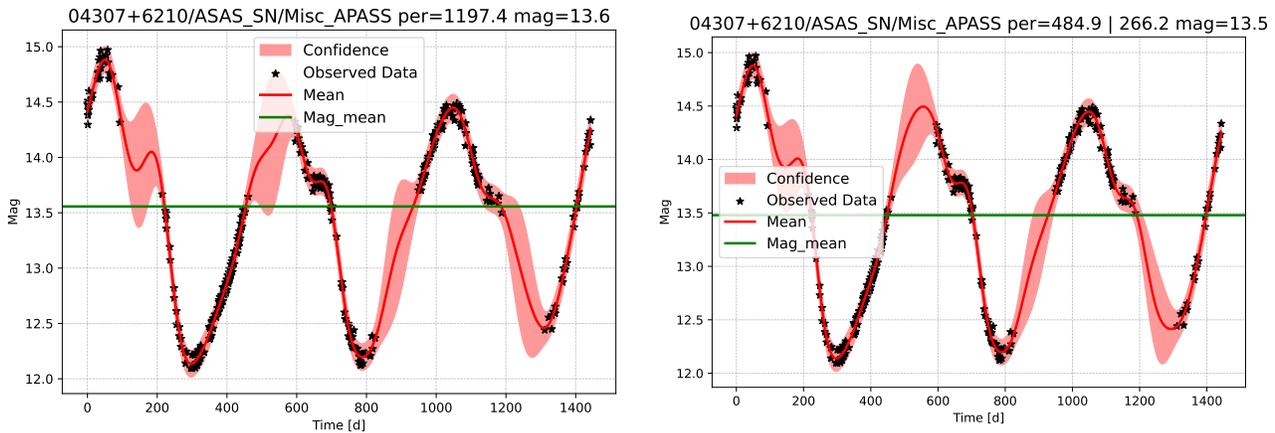
Figure 3.8: Comparison between a single-component fit (left) and a two-component fit (right). It can be observed that in the single-component fit, the estimated period is not consistent with the peak-to-peak time interval of the light curve or the best-fit model, whereas the two-component model has at least one period that matches.

The choice of the fundamental period in this context poses an additional problem. If a second long period in the power spectrum has a higher power than the first overtone, then, due

Implementation of PGMUVI

to the choice of the longer period, the second long period will be interpreted as the fundamental, while the fundamental period will be considered as the first overtone or the second component detected by PGMUVI. However, it is important to note that these cases are uncommon, as it is observed in power spectra that the second long period rarely exhibits higher power than the first overtone.

Figure 3.8 compares the results of the 1- and 2-component fits. The 1-component fit results in a median period of 1443 d. This value is much larger than the peak-to-peak interval; in fact, it is comparable to the total baseline. In contrast, the median of the longer period for the 2-component fits is 485 d, which aligns with the time interval between the peaks of the model. For the second component, a period estimate of 270 d is obtained. This can be interpreted as the first harmonic. In summary, adopting a multi-component model leads to more precise estimations of the periods.

One final detail to discuss is the case of the KELT light curves. Although these curves were separated for analysis using the two telescopes that have the same photometric system, it did not reduce the amount of data. Because the PGMUVI algorithm consumes significant computational resources, processing these light curves involves a long computation time, making their analysis challenging. Therefore, only the light curves with fewer than 1000 observed epochs were processed, leaving the rest for future work, where optimizations can be performed or a more advanced computing setup can be utilized for their analysis. The above reduces the total number of light curves from 3838 to 3634 light curves processed in the end.

3.2.3   PGMUVI results

A total of 3634 light curve models were processed and analyzed after implementing PGMUVI. A representative sample of these data is presented in Table 3.3. This table displays the periods, mean magnitudes, amplitudes and uncertainties corresponding to one of the 28 observed bands, encompassing some of the 674 stars subjected to the PGMUVI methodology.

Table 3.3: Example of the values obtained after the extensive application of PGMUVI to the entire NESS_LS database. The result is a table with data for the fundamental period and second period (P_0 and P_1), mean magnitude (m_{mean}), amplitude (Δm), time baseline (ΔT), among other values. Not all available values are shown, such as the mixture weight or loss, among others. This example shows only 10 light curves from ASAS-SN. The table will be available in electronic format once the respective article is published.

IRASID	P_0 [d]	P_1 [d]	m_{mean} [mag]	Δm [mag]	Δm_{err} [mag]	ΔT [d]	$\ldots$
03377+6303	498.33	494.64	5.86	0.47	1.39	1443.04	$\ldots$
09076+3110	1225.99	263.66	7.17	4.97	1.08	2468.18	$\ldots$
09273-5157	459.55	188.56	11.05	4.14	1.50	887.68	$\ldots$
11287-6918	613.73	472.17	12.47	2.48	2.08	894.80	$\ldots$
12226+5703	1923.80	191.65	6.68	0.36	0.30	1973.37	$\ldots$
14003-7633	365.79	123.16	6.21	1.29	0.85	1586.82	$\ldots$
14248-5927	276.86	100.64	8.85	0.56	0.33	740.12	$\ldots$
15361+2441	162.49	59.89	7.84	6.24	1.11	2004.64	$\ldots$
16450-4448	409.52	222.77	9.79	0.24	2.60	926.79	$\ldots$
20000+4954	1359.91	219.92	10.81	4.58	0.86	1699.61	$\ldots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Since determining a single fundamental period through PGMUVI is currently unfeasible, the approach was taken to consider the set of estimated periods with the highest significance for different light curves of the same star. From this set, the median was calculated and assumed as the characteristic period of the fundamental mode. The associated uncertainty was obtained using the median absolute deviation (MAD) of the periods in that set. For a more detailed analysis, a discussion will be provided in the upcoming chapter. Examples of the period medians

and uncertainties can be found in Table 3.4.

Table 3.4: The mean, median and standard deviation of the periods for each of the stars in our sample. Only ten rows are shown here. Additionally, the number of periods for each star combined for the calculation of the median and mean are shown

IRASPSC	P_0 [d]	P_{MAD}^0[d]	P_1 [d]	P_{MAD}^1 [d]
00042+4248	732.5	165.1	733.4	474.0
00084-1851	360.9	8.2	98.8	5.4
00121-1912	1042.9	–	230.4	–
00192-2020	158.2	1.7	86.4	5.6
00193-4033	662.1	63.5	196.6	40.9
00205+5530	445.2	124.3	361.7	91.0
00213+3817	214.9	196.9	108.5	90.5
00245-0652	174.6	21.2	102.5	31.9
00247+6922	654.7	93.5	279.1	136.6
⋮	⋮	⋮	⋮	⋮
R09061-4313	328.2	–	28.3	–

Before comparing PGMUVI and LS and MLS results, it is necessary to clarify the differences in period determinations between these methods. In the LS and MLS methods, the period uncertainty is determined in terms of the FWHM of the peak in the periodogram corresponding to the fundamental period (see Section 3.1.3). Sources with light curves at multiple wavelengths are processed using the MLS method, which returns a single period per source. In this thesis, we use PGMUVI to fit the light curve at a single wavelength (fitting multiple wavelengths will be investigated in the future); for a given source, this results in as many period estimates as there are light curves. Then, the median of these periods is calculated for a source, and we use the range of periods determined from the various light curves to estimate the dispersion around this median. It should be noted that PGMUVI currently does not provide uncertainties for the estimated parameters; therefore, sources with a single light curve do not have a range of periods that can be assigned as uncertainty. Consequently, a direct comparison of the uncertainties

obtained from the two methods should be performed with caution. Multi-wavelength light curve analysis with PGMUVI will allow us to properly constrain the period uncertainties in the future.

Additionally, in Figure 3.9, the Probability Density Functions (PDFs) of the medians of the periods corresponding to the fundamental mode (P_0) and the second period (P_1) are presented, as well as the fundamental period obtained through LS and MLS ($P_{\text{LS\_MLS}}$). This PDF was generated using the kernel density estimation [KDE; Parzen, 1962] technique, a statistical method employed to estimate and represent the shape of a data distribution without the need to predefine the number of intervals (bins) in a histogram. KDE produces smoothed non-parametric distributions, enabling a more accurate visualization of the obtained distributions. It can be observed that the PDF of P_0 exhibits its maximum value near the artifact of half-year. This could be due to PGMUVI effectively modeling this artifact, or it could be the first overtone corresponding to the fundamental mode P_1. A detailed study is necessary to discern the factors contributing to this distribution centered around this value.

When comparing the PDF of P_0 with the PDFs of $P_{\text{LS\_MLS}}$, it is noticeable that they are centered around very similar values: the PDF of $P_0 = 422.2$ days and the PDF of $P_{\text{LS\_MLS}} = 417.8$ days. Furthermore, these two peaks of the distributions are outside the artifact of a year previously identified by Fraser et al. [2008]. It is also noticeable that the PDF of $P_{\text{LS\_MLS}}$ appears to be a combination of two distributions, one of which is within the half-year artifact, while the latter disappears in the PDF of P_0. Therefore, a preliminary conclusion is that PGMUVI does not seem to be influenced by the half-year artifact, at least in the case of the PDF of P_0 and provides a better estimation of periods associated with the fundamental pulsation mode.

In Figure 3.10, we compare the predictions of PGMUVI for P_0 with the periods of LS and MLS. We find that 54.3% of the 672 values fall within the $\pm 25\%$ interval (highlighted in red),
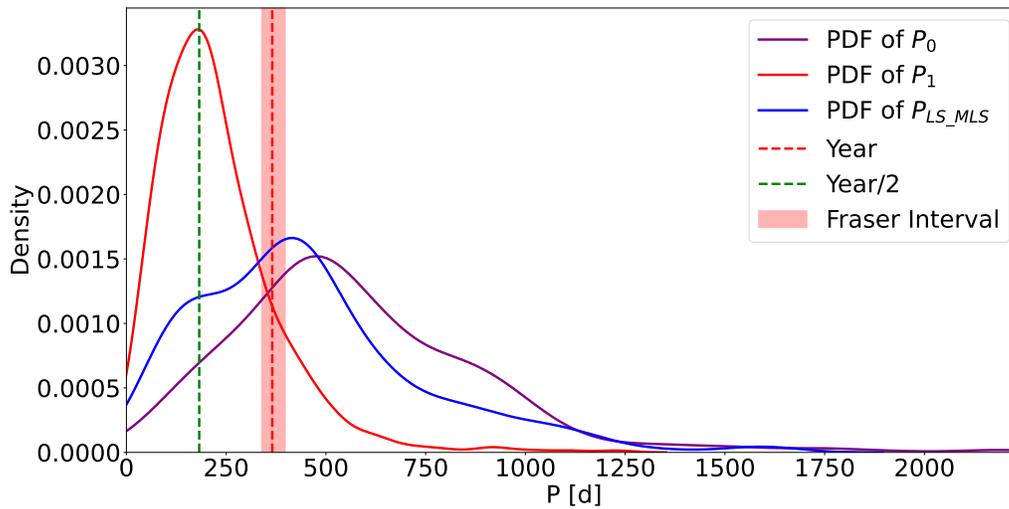
Figure 3.9: Different PDFs constructed with KDE are used to visualize how the periods are distributed and whether there are peaks in the distribution related to artifacts equivalent to a year or half a year, or that fall within the Fraser interval.

while 70.5% are within the ±50% interval (shaded areas in gray and red). In addition, due to the typical uncertainties associated with periods along both axes (purple cross in figure), this region might include more points. These results are encouraging, particularly considering the inherent uncertainties in period determinations, as detailed below. It can also be inferred from this figure that PGMUVI adopts period values from LS and MLS to calculate larger period values. This explains the dispersion observed above the gray region, which encompasses differences greater than 50%. On the other hand, cases in which PGMUVI estimates periods shorter than those estimated by LS and MLS are rare.

It is important to note, upon observing Figure 3.10, that the inclusion of the Fraser interval in the proximity of the shaded regions leads to an apparent homogeneity in the distribution of points. This observation has a significant impact on the correlations presented in the following chapter.
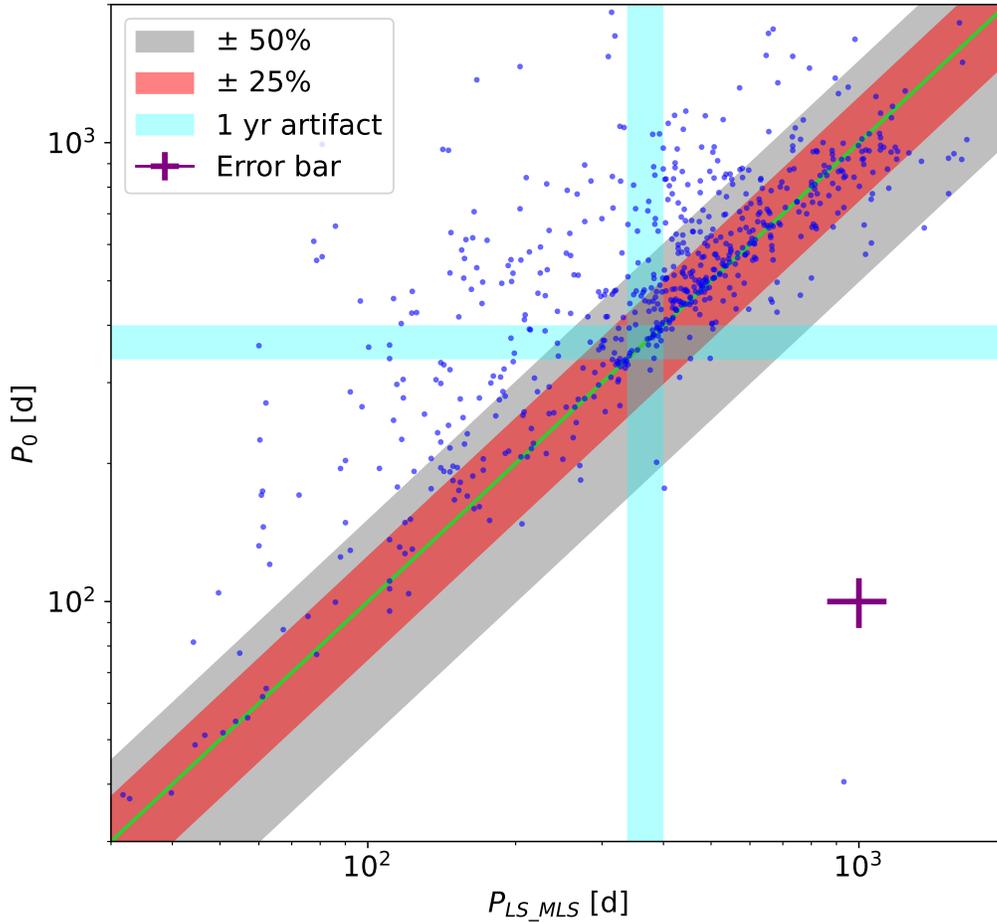
PGMUVI results

Figure 3.10: The periods obtained by PGMUVI results are compared to the LS and MLS values. The median is chosen because it is a robust statistical measure. Most of the PGMUVI periods are within 50% (grey band) of the 1:1 line (green curve), which indicates the cases where the PGMUVI fit equals the initial guess from the LS and MLS methods. The typical error bar (crosshair) demonstrates that a 50% agreement is, in fact, consistent with the inherent uncertainties.

Finally, we examine the relative uncertainties of the periods. To estimate this quantity, we divide the median absolute deviation from the median (MADM) by the median. This is a robust estimator of the relative uncertainty, as the mean and standard deviations are sensitive to outliers. It is observed that more than 50% of the sources exhibit relative uncertainties below

About the amplitude of the light curves

25%. It is important to highlight that this result is valid only for those sources that possess multiple light curves, which allow for the estimation of the MADM.

3.2.4 About the amplitude of the light curves

As part of this thesis, we implemented a method to estimate the amplitude using the non-parameteric model returned by PGMUVI for each light curve. We set the amplitude to be the absolute difference between the maximum and minimum magnitude values from the non-parametric model. This approximation holds as long as the period estimated by PGMUVI is shorter than the time baseline of the light curve.

The determination of amplitudes in other studies relies on the identification of maximum and minimum magnitude values within observational data. To exemplify this, consider the work conducted by Silverberg et al. [1993], which focused on observations from COBE-DIRBE. In this study, to estimate the maximum (or minimum) magnitude, the authors computed the median of three points with the highest (or lowest) values. This procedure produces more robust estimates of the extreme values. [Jayasinghe et al., 2018] employ a more robust method to estimates V band amplitudes for the ASAS-SN data: they define the minimum and maximum magnitudes as the 2.5^{th} and 97.5^{th} percentiles of the observations. Despite being robust against outliers, even the latter methodology can pose challenges. For example, if the star is at the detection limit in terms of magnitude, there may be parts of the light curve with magnitude peaks (corresponding to luminosity minima) that are not recorded. In such cases, the resulting estimation will only reflect a fraction of the true amplitude. On the other hand, if the star's magnitude is at its minimum value, there is a possibility of sensor saturation, which would also affect the accuracy of the estimation.

As PGMVUI is able to predict magnitudes in regions with missing data, we expect that PGMUVI fits would perform better in cases of low sensitivity or saturation. Figure 3.11 presents the distributions of PGMUVI amplitudes for the V band of ASAS SN (orange line) compared to the distribution of amplitudes estimated by Jayasinghe et al. [2018] (blue line). In both cases, a bimodal distribution is observed. The smaller peak is located in the region above 2.5 mag, which, by definition, is classified as Mira-type stars. Stars with magnitudes below 2.5 mag are classified as non-Mira stars (Here,"Non-Miras" refers to semi-regular variable stars or NESS objects that are not necessarily LPV), suggesting that these two populations are present in both our estimates and the previous work. Finally, it can be noticed that the peak of Mira stars in our estimates is distributed at higher amplitudes, which is consistent with our strategy for estimating amplitudes based on the model. In contrast, the amplitudes from ASAS-SN are affected by sensitivity limits, resulting in Mira stars with smaller amplitudes compared to those calculated in this thesis.
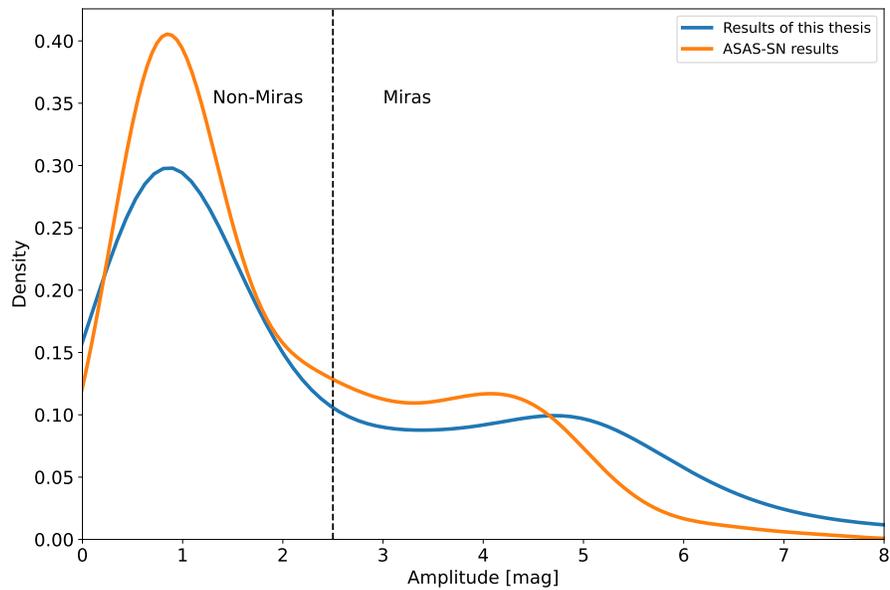
Figure 3.11: Comparison of the amplitudes of NESS sources in the V band of ASAS-SN between this thesis and the work of Jayasinghe et al. [2018]. A similar bimodal behavior is observed in both distributions, one for Mira-type stars and another for non-Mira stars. The amplitude PDF calculated in this thesis shows a wider range for Mira-type stars, as our models allow us to infer the maximum and minimum magnitude values that are lost due to the combined nature of the sensor and the observed star.

Analysis and exploration of correlations

In this chapter, previously established correlations in the literature will be examined in relation to aspects such as period, amplitude, and mass-loss rate (MLR), as well as color, among others. These correlations will be compared with the relationships constructed using the data obtained through the light curves for NESS targets processed with PGMUVI. In this section, we study trends between parameters such as period and amplitude calculated in this study and other quantities derived from observations. This is done as a first step to demonstrate that our results can reproduce trends noted in the literature. It is crucial for the reader to grasp that some of the relationships established in the literature were built upon limited and meticulously characterized samples, whereas the replications carried out in this study involve a set of 674 sources with a diverse range of physical properties, which could lead to significant discrepancies in these relationships, or dispersions around the relationships confirmed with smaller, better-characterized samples. Before discussing these trends, we remind the reader that, as noted in Section 3.2.3, we must be careful when comparing the uncertainties of periods

derived using the LS and MLS methods with those of PGMUVI.

4.1 Period-color relationship

Period-color (PC) relations are highly useful when investigating sources for which determining distance or the interstellar medium extinction affecting their radiation is challenging. Moreover, color enables the identification of infrared excess, which greatly aids in identifying stars with a significant dust content in their atmospheres if interstellar extinction is already accounted for.

In Figure 4.1, we present the PC diagram using the $K_{\rm s}$ band (effective wavelength: $\lambda_{\rm eff} \sim 2.1\,\mu$ m) of 2MASS and the W4 band ($\lambda_{\rm eff} \sim 22\,\mu$m) from WISE. This is done to analyze dust emission as a function of period. AGB dust does not significantly contribute to emission or absorption in the $K_{\rm s}$ band (it is dominated by AGB photospheric emission), allowing for the study of the star's SED without the effects of circumstellar dust. On the other hand, the $W4$ band traces the emission of warm circumstellar dust [approximately 300 K; McDonald & Zijlstra, 2016]. In general terms, the variation is similar to Figure 2 from McDonald et al. [2018]: a nearly constant color in the range of 60 to 300 days, followed by a rapid increase in reddening as the period increases beyond 300 days. In this figure, we distinguish sources with periods more precise than 10% from lower-precision sources. The choice of the cutoff value in precision does not have a significant effect on the resulting distributions, we pick 10% just to be able to point out the trend in the data. Sources with relative uncertainties < 0.1 are shown in red. As mentioned in Section 3.2.3, the relative uncertainties were calculated using the MADM and the median of the periods for each source. Throughout the remainder of the chapter, these sources will be highlighted in additional relationships, while those with $\sigma_{\rm Relative} \geq 0.1$ will be represented with a level of transparency.
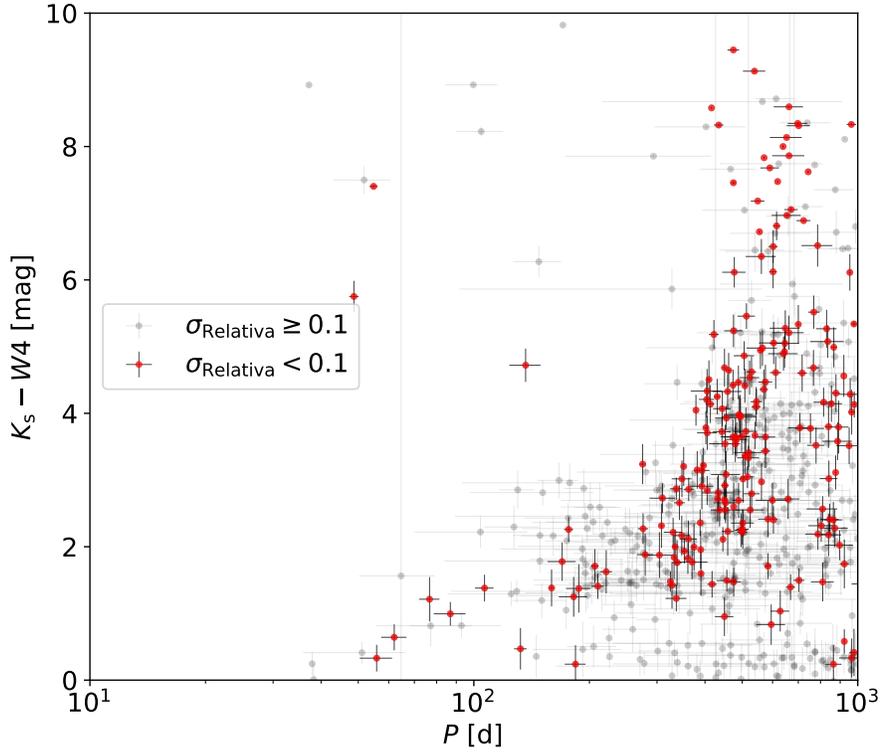
Figure 4.1: The color-period relationship constructed using periods estimated through PGMUVI. Periods with precision higher than 10% are shown in red. The color increases as the period increases beyond a period of approximately 300 d.

To compare the results in Figure 4.1, we consider the findings of the study conducted by McDonald et al. [2018]. In their study, the authors aimed to determine if the infrared excess follows the same behavior as the mass-loss rate determined by the rotational emission of CO. Their Figure 2 (see top panel in Figure 4.2) presents the color as $K_s-[22]$. In this case, the notation [22] refers to the $W4$ band of WISE. In the figure, the PC distributions of LPV stars from the General Catalog of Variable Stars (GCVS) are shown in the background, with the stars considered by McDonald et al. in their study superimposed (stars from M13, ω Centauri,
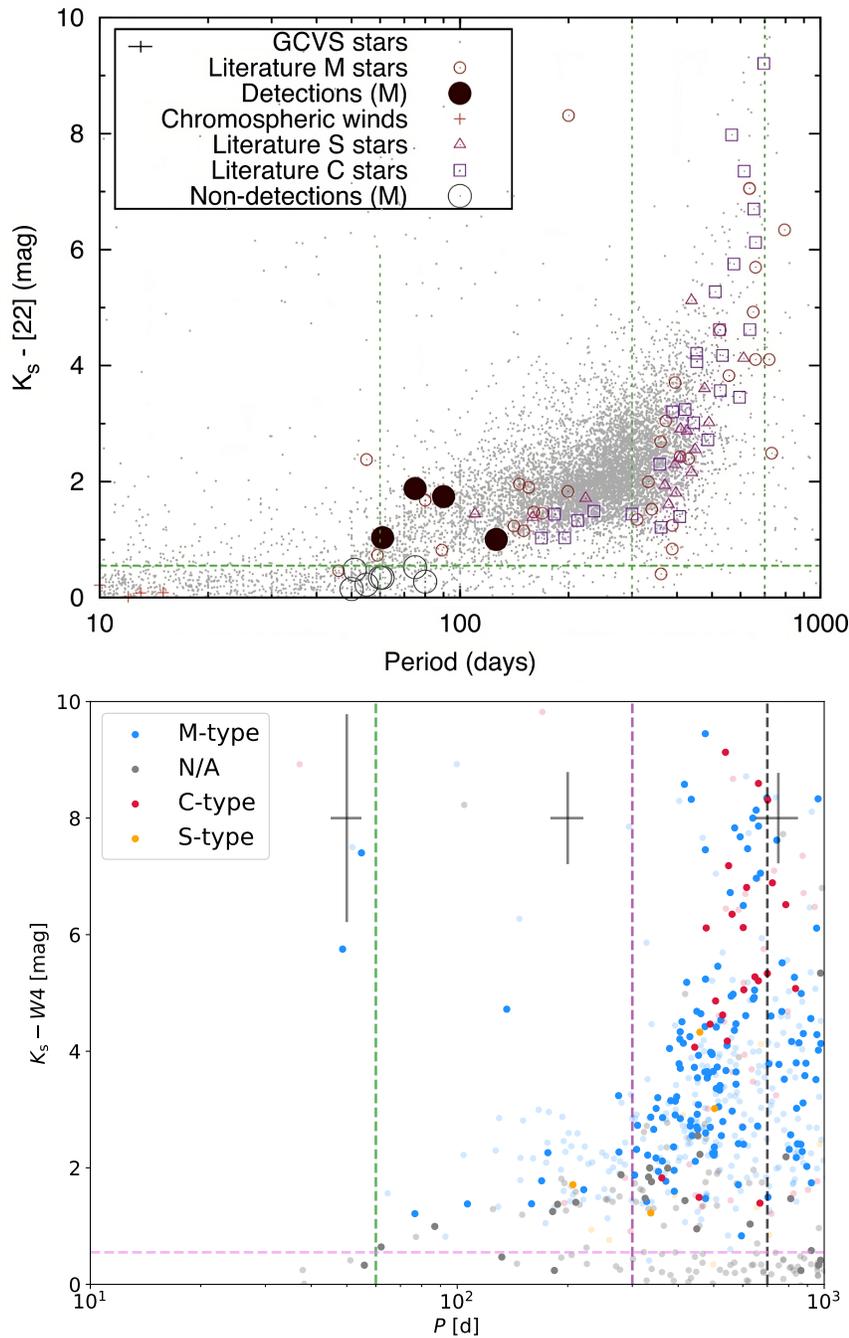
Figure 4.2: Top: Figure 2 of the [McDonald et al., 2018] study presents a distribution of LPV (Long-Period Variables) sourced from the GCVS, with data points categorized by chemical classifications and clearly identified sources. Bottom: Results of these thesis for periods and colors for NESS targets. Furthermore, the gray crosses show representative uncertainties for the periods in three intervals. The dashed violet line represents the limit of stars with low dust production from McDonald et al. [2018]. The high-quality PGMUVI sample of periods faithfully reproduces the distribution seen in their paper.

NGC 2808, etc.). Several details can be pointed out about this figure. First, sources with periods of less than 60 days have very low color values, $K_s - W4 < 0.55$, indicating limited dust production and no detected CO emission from these sources. Moving into the range of 60 to 300 days with $K_s - W4 > 0.55$, there is dust production and detected or reported CO emission in the literature, causing stars to become redder in this region. Beyond 300 days, the dust content increases dramatically, and stars tend to become redder as their periods increase.

The PC diagram from this thesis is depicted in the bottom panel of Figure 4.2. It is distinguished by the chemical types of NESS stars, and it is compared with Figure 2 from McDonald et al. [2018]. Our results reproduce the trends seen for each chemical type from their figure. In our figure, we highlight the sources with high-precision periods to be able to compare their distribution with that of Figure 2 of McDonald et al. [2018]. There is a consistent color distribution within the 60 to 300-day interval (between the green and purple dashed vertical lines), followed by a tendency for colors to become redder as the period extends beyond 300 days. The McDonald et al. sample does not have a large population of sources with confirmed chemistries above $P = 700$ d; however, there are many GCVS sources in this region of their figure. Our data contains 188 sources in this region, most of which have chemical classifications.

Comparing the distributions of different chemical types, it is observed that in the PC diagram of this thesis, carbon-rich stars are grouped in intervals of 300 to 1000 days and are dispersed in a range of colors from 1 to 10 mag. On the other hand, oxygen-rich stars dominate the NESS sample, spanning periods from 60 to 1000 days. These stars tend to cluster in the color range of 2 to 6 magnitudes, where the point density decreases for colors bluer than 1 mag and colors redder than 6 mag. This can be explained by the fact that the bluest stars are not dusty enough and are not included in the NESS sample or do not have significant MLRs, and the reddest

stars correspond to the dustiest and are therefore the most massive; there are very few stars in this mass range. Overall, the behavior of these oxygen-rich stars becomes redder as the period increases.

Regarding S-type stars, they are less common in the NESS sample. They exhibit colors bluer than 5 mag and seem to follow a distribution similar to that of oxygen-rich stars. Finally, unclassified stars show a wide range of periods and are located in the bluer colors. It is worth noting that the results obtained for the PC diagram in this thesis align with the distribution of different chemical types presented in Figure 2 of McDonald et al. [2018] work.

4.2   Dust-production rate/mass-loss rate relationships

One of the main characteristics of stars in the AGB phase is mass loss. As mentioned in Section 1.2.2, this mass loss is linked to pulsations that lead to the condensation of gas into dust and acoustic waves that create shocks. Therefore, it is feasible to establish a relationship between the mass-loss rate and the pulsation period.

Dust-production rate (DPR) estimates for the NESS targets are provided in Scicluna et al. [2022, also Trejo et al., in prep.]. These DPRs were computed by fitting their SEDs with models from the Grid of RSG and AGB ModelS[GRAMS; Sargent et al., 2011 and Srinivasan et al., 2011]. Figure 4.3 illustrates the relationship between the DPR or MLR and the period for our sample. In each case, we highlight sources with high-quality periods (relative uncertainty $< 10\%$). In the upper panel, DPR estimates for the NESS targets are presented. These estimates have been scaled by a factor of 200 in order to have a direct comparison with Groenewegen et al. [1998],
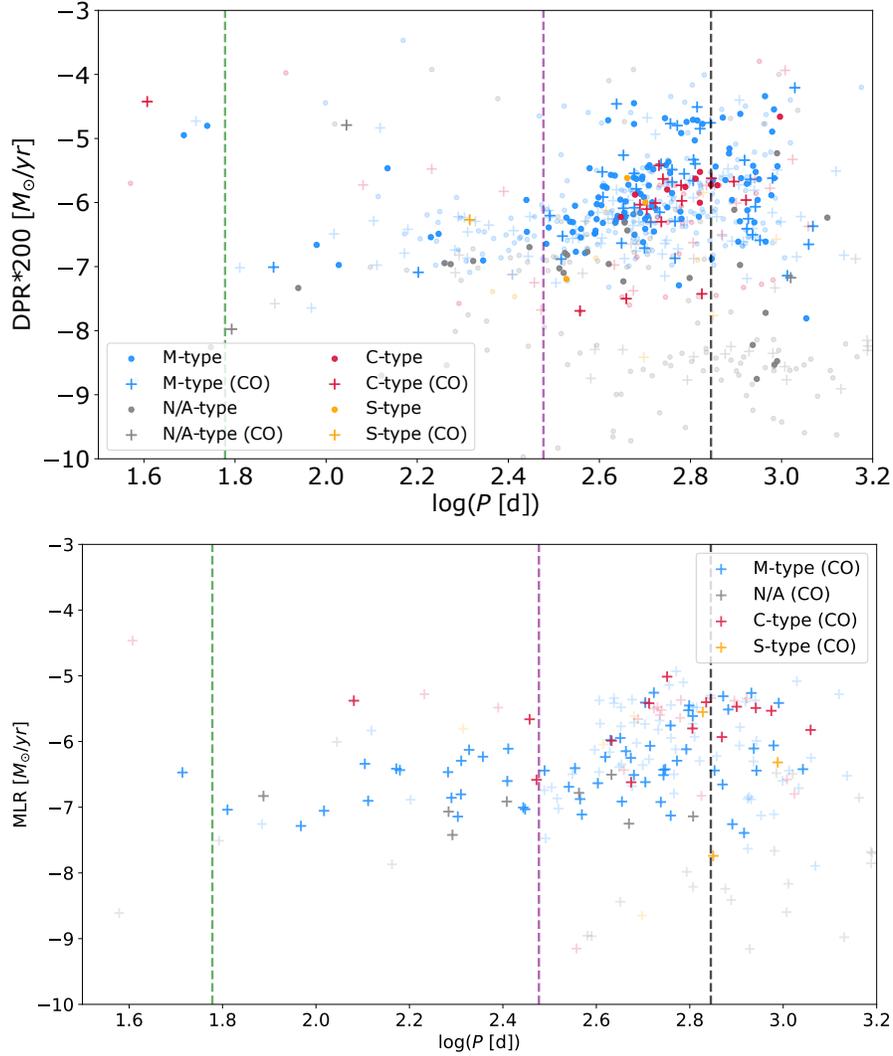
Figure 4.3: Relationships period-MLR, symbols marked with "+" indicate sources with CO detection. Top: Relationships constructed with data from Trejo-Cruz et al. in prep. for the DPR modeled with GRAMS and the median of estimated periods with PGMUVI The relative uncertainty for this panel is 21% for the DPR. Bottom: Relationships constructed with data from Wallström et al. in prep. for the MLR derived from the detection of CO $J = 2 - 1$ and the same periods. The relative uncertainty for this panel is 87% for the MLR. In both panels, the relative uncertainty in period is 12%.

who use the same value to convert their DPRs to MLRs. In the lower panel, gas MLR estimates from Wallström et al. in prep are used, derived from CO $J = 2 - 1$ emission. NESS has 421 targets with CO measurements, but only 195 of them have $MLR_{CO(2-1)} > 0$. For these

sources, the "+" marker was used to distinguish them from those where $MLR_{CO(2-1)}$ has not been detected or is indeterminate. This was done to allow easy comparison between the DPR and MLR plots (the latter only contain the sources with CO detections and valid MLRs). This convention will be maintained in the remaining graphs in this section.

In each plot, different colors are used to represent various chemical types (M-rich, C-rich, and S-type), which are determined from mid-infrared spectra. The comparison between both panels in Figure 4.3 reveals a similar trend for oxygen-rich stars. Most of these stars are concentrated in the range of 10^{-7} to 10^{-5} $M_\odot$ yr^{-1}, with a higher density of data points when the period exceeds 300 days. Stars without chemical classifications tend to exhibit mass-loss rates lower than 10^{-5} $M_\odot$ yr^{-1} in both figures. For carbon stars, a wider range is observed in the upper panel, ranging from 10^{-8} to 10^{-4} $M_\odot$ yr^{-1}. In contrast, in the lower panel, these stars show rates ranging from 10^{-7} to 10^{-5} $M_\odot$ yr^{-1}, following the same pattern as oxygen-rich stars. Finally, S-type stars exhibit behavior similar to that of carbon-rich or oxygen-rich stars.

The work of Vassiliadis & Wood [1993] is utilized to draw a comparison between the data they reported and the data presented in this thesis, as it is a widely cited article to this day and holds significant importance in the study of mass loss in the AGB phase. Nevertheless, it's crucial to emphasize that the primary objective of Vassiliadis & Wood [1993]'s work was to explore the evolution of AGB stars during the final thermal pulse cycles. In their evolutionary models, Vassiliadis & Wood [1993]. incorporated empirical relationships that describe the mass loss of these stars. Additionally, they explored the period-MLR relationship using Mira stars of different chemical types (C, M, and S) in the Galaxy as well as OH/IR stars and AGB stars from the LMC.

To determine the MLR in these different stellar populations, Vassiliadis & Wood [1993] derived

values based on CO emissions in the microwave range or through models using flux measurements with the 60 μm IRAS filter. They concluded that for stars with periods shorter than 500 days, the MLR increases exponentially with the period. On the other hand, for stars with periods longer than 500 days, AGB stars enter the superwind phase, and the MLR remains constant (around $10^{-4}M_\odot$ yr^{-1}) as a function of the period. As their work is based on a dataset that combines small samples of Milky Way and Magellanic Cloud stars, their results might not be representative of AGB stars in the Solar Neighborhood. This fact should be taken into account when comparing our results to theirs.

These findings can be observed in Figure 4.4. The top panel reproduces Figure 1 from Vassiliadis & Wood [1993], while the central and bottom panels compare mass-loss rates derived from gas and dust data with PGMUVI periods from this work. For the central panel, we convert DPRs estimated from GRAMS fits (Trejo et al., in prep) to MLRs by multiplying them by 200, the canonical value for the gas-to-dust ratio assumed by Groenewegen et al. [1998], for easy comparison with their paper. The bottom panel shows gas MLRs estimated from NESS CO data (Wallstöm et al., in prep). For a direct comparison of our results with those of Vassiliadis & Wood [1993] and Groenewegen et al. [2009], we only display sources in our sample that have been classified as Miras by various authors. We also highlight the high-quality period determinations for this sub-sample in the figure.

We use the same chemical types (C, M, and S) and stars classified as Miras in the works of Kharchenko et al. [2002], Whitelock et al. [2008], Arnold et al. [2020], Merchan-Benitez et al. [2023] and Jayasinghe et al. [2018]. Out of the 674 NESS stars with periods estimated
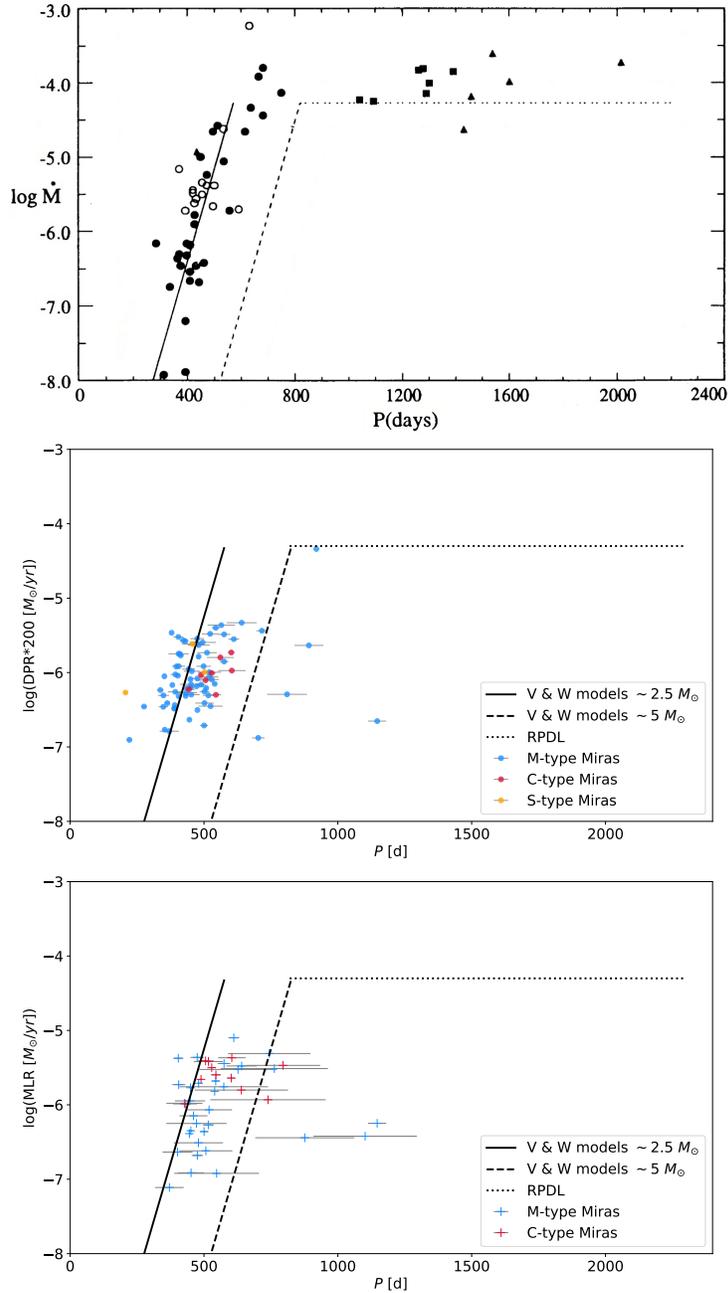
Figure 4.4: Top: The relationship MLR–period, as extracted from Figure 1 of Vassiliadis & Wood [1993] (filled symbols: O–rich and S stars, open symbols: carbon stars). Center: The relationship constructed using DPR with a factor of 200. Bottom: Constructed MLR data derived from CO emissions. It can be observed that the Vassiliadis & Wood model for 2.5 M$_\odot$ stars (represented by the black line) closely aligns with the distributions in the central and lower panels. The Vassiliadis & Wood model passes through a data region with small uncertainties relative to the data itself. However, carbon stars deviate from this line and tend to be disrupted in this model, as well as in the Vassiliadis & Wood model for 5 M$_\odot$ stars (black dashed line).

by PGMUVI, the classification of 204 Mira-type LPVs is derived from estimated amplitudes in the optical bands.We also show the Vassiliadis & Wood [1993] empirical relationship between mass-loss rate and period in the figure for two different values of stellar mass (solid line: 2 $M_\odot$, dashed line: 5 $M_\odot$).

It is noteworthy that the distribution obtained in this thesis appears to follow the trend of the 2.5 $M_\odot$ model. Even the distribution of chemical types shows similarities with Figure 2 of Vassiliadis & Wood [1993], where carbon stars are found in the range of 10^{-6} $M_\odot$ $yr^{-1} <$ MLR $< 10^{-5}$ $M_\odot$ yr^{-1}. In the central panel, this population exhibits values close to 10^{-6} $M_\odot$ yr^{-1}, and in the lower panel, a range similar to that of figure 2 of Vassiliadis & Wood [1993] is observed (see top panel in Figure 4.4). Therefore, for this selected sample of sources with relative uncertainties of 0.1 in period for the central panel and 0.3 in period for the lower panel (we lowered the precision threshold in this case to increase the number of stars with CO MLRs), it is possible to reproduce the Vassiliadis & Wood model, both in terms of period and the chemical distribution along the MLR axis.

Regarding the horizontal dashed line (MLR $\sim 10^{-4}$ $M_\odot$ yr^{-1}) resulting from the maximum estimated MLR value in the superwind phase, it can be observed that this line provides an upper limit for the MLR estimated by CO. On the other hand, in the central panel of the figure, a source can be seen right at the superwind phase boundary (see Figure 4.4 with period ~ 900 d and $\log(DPR^*200) \sim -4 \log(M_\odot/yr)$), which is identified as IRAS 20015+3019. This source is a confirmed dusty OH/IR star [Chengalur et al., 1993] and has been classified as an LPV [Gaia collaboration, 2022]. This classification is consistent with the source being in the superwind phase, based on its location in the central panel of Figure 4.4.

Also compared is the MLR–period relationship for our sample with the one from Groenewegen et al. [2009], who investigated the relationship between MLR and period, as well as luminosity, using radiative transfer models to simulate the SEDs of evolved stars in the LMC and SMG. In their study, the model results are compared with those of the fit determined by Groenewegen et al. [1998] (represented by the black straight line) and the Vassiliadis & Wood model for 2.5 $M_\odot$ (represented by the black curved line), taking into account observational data and separating by chemical types of stars. This can be observed in Figure 22 of Groenewegen et al. [2009] (top panel of Figure 4.5, with red color for type C stars and blue color for type M stars). While the locations and distributions of these stars would, in principle, be affected by host galaxy metallicity [e.g., Cioni, 2009] for this simple statistical comparison, this detail is not relevant.

Before we compare our results, we need to account for the fact that the Groenewegen et al. [2009] "MLR" are simply DPRs converted to MLRs assuming a gas: dust ratio of 200. This is the reason we have also converted our DPRs in the same fashion. Another effect must be accounted for: due to the differing optical constants used for amorphous carbon in Groenewegen et al. [2009] and Srinivasan et al. [2011], the Groenewegen et al. [1998] opacities are higher than the GRAMS opacities. As a results, GRAMS DPRs for carbon stars are smaller by a factor of approximately 4 compared to their values. To account for this, we scale down the Groenewegen et al. [1998] relation by the same factor for a direct comparison with our data. There is still the possibility that the actual gas:dust ratio for each star is different from the assumed value of 200, and this may also introduce a scatter around the fit of Groenewegen et al. [1998].

The two bottom rows in Figure 4.5 compare the fit of Groenewegen et al. (solid line) and the
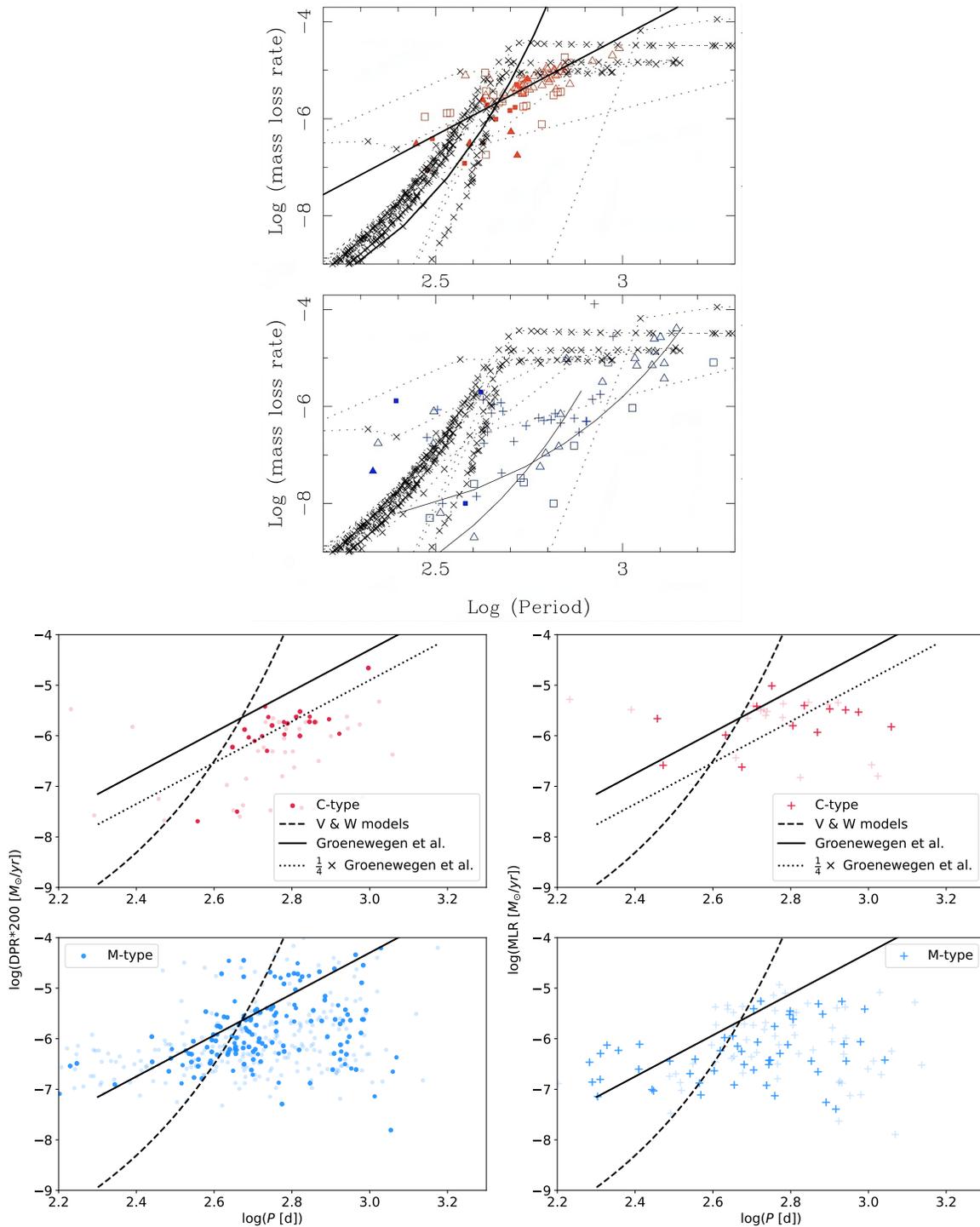
Figure 4.5: Period-MLR relationships separated by the chemical type of the star, according to Groenewegen et al. [2009] (top two rows) and according to this study (bottom two rows): in each case, carbon-rich stars are shown in the top row, while oxygen-rich stars are shown in the bottom row. As before, open circles represent sources with high-quality period determinations. An additional model, a black dashed line, adjusts the Groenewegen et al. [1998] carbon-star mass-loss rate relation for the 4× systematic offset with GRAMS DPRs.

2.5 $M_\odot$ model of Vassiliadis & Wood [1993] (dashed curve) with our results. As before, we highlight the sources with high-quality periods. Once the line is shifted to account for the difference in opacities (dashed line), we see that the DPRs for our carbon stars are consistent with those of Groenewegen et al. The CO-based MLR has a larger dispersion and it is not straightforward to determine whether they follow one of the two fits. As the fits are based on DPR, it is possible that the dispersion and the difference are due to the gas:dust ratio being different from the assumed value. The carbon stars in our sample with high-precision periods all seem to lie to the right of the Vassiliadis & Wood model; while this indicates that their masses are higher than 2.5 $M_\odot$, more work is required to properly interpret this observation.

The bottom row in the figure shows the case for O–rich AGB and S stars. In this case, there is a larger spread in the data due to the fact that these stars span a larger range of masses (as evidenced by the fact that our sources lie on either side of the Vassiliadis & Wood model for 2.5 $M_\odot$) and consist of low- and high-mass populations. Based on this sample and our current results, the period–MLR relation for oxygen-rich stars is more complicated and requires further analysis.

4.3   Relationship of amplitude as a function of effective wavelength

We can now investigate the variation of the amplitude $\Delta\mathrm{mag}_\lambda$, calculated as described in Section 3.2.4, with wavelengths ranging from the optical to the mid-infrared.

In Figure 4.6, we show the variation of amplitude with effective wavelength of the filter[1]

---

[1]These λ_{eff} are recorded in Table 3.3.

for three sources from Smith et al. [2002] which are also NESS targets. We find that we reproduce their amplitude estimates, and also obtain an additional amplitude value for one of the sources, IRAS 09185−4918 (RW Vel), this is consistent with the functional behavior of the amplitude-wavelength relationship for that source. Figure 4.7 illustrates seven more examples of the relationship for stars in our sample. Sources with the highest number of recorded light curves were selected, spanning the optical and mid-infrared ranges. In both figures, the amplitude shows an overall decrease with effective wavelength, consistent with the findings of Smith et al. [2002] in their study using V-band data and COBE-DIRBE bands (left panel in Figure 4.7). According to him, the variation in amplitude is due to fluctuations in the temperature of the molecular layers and in the temperature affecting the condensation of dust, and both components affect the radiation escaping from the star and passing through each of the regions.
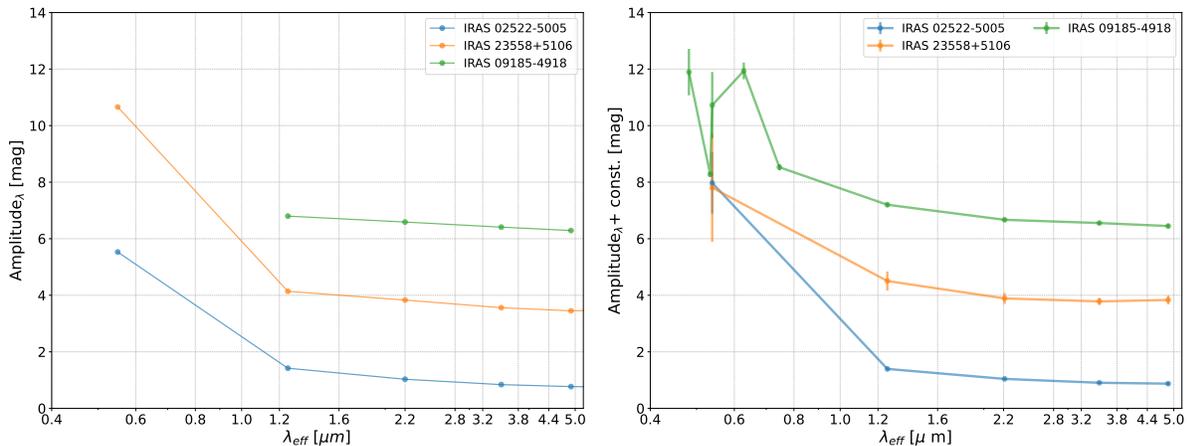
Figure 4.6: Reproduction of the relationship between amplitude and wavelength for three NESS sources, right panel: results plotted using data from Smith et al. [2002]. Left panel: results from this thesis for the same sources. It is noteworthy that for bands with wavelengths greater than 1 μm, the same values are obtained, and in general, the same decrease in amplitude with wavelength is observed. In this work, more data points were obtained in the optical range compared to what Smith showed for IRAS 09185–4918 (shown in green). For clarity, the curves are shifted by a constant factor along the amplitude axis.

We can probe the amplitude dependence over a total of 29 bands using the work in this thesis. In the future, our results will allow us to study the amplitude distribution as a function

of wavelength. A plot showing such a variation over six bands in the near- and mid-infrared can be found in Sanders et al. [2022, see their Figure 7].

An initial conclusion is that the nonparametric models generated by PGMUVI allow one to obtain amplitude estimates that agree with what is reported in the literature. It is expected that in the future, PGMUVI will enable obtaining functions that provide better estimates, reduce uncertainties, and focus on the overall behavior of the model rather than just its maximum and minimum values.
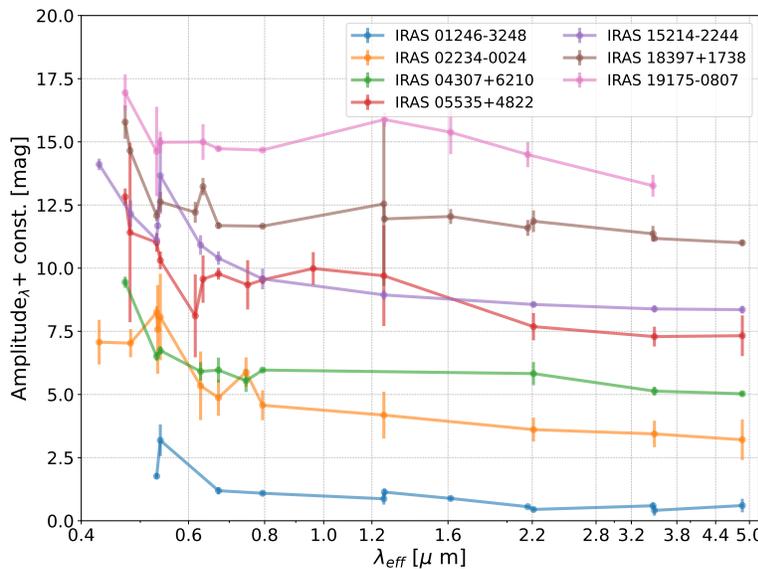
Figure 4.7: An example is presented of 7 sources of NESS that have more than 10 light curves, obtained through different photometric filters. This allows us to examine the relationship between the amplitude change based on the effective wavelength. The scale used on the wavelength axis is semi-logarithmic, and labels have been chosen to allow the reader to easily identify the corresponding wavelength of each point. The light curves are displaced along the amplitude axis by a constant factor for clarity.

In conclusion, the results obtained in this thesis exhibit qualities and behaviors that are established in the literature, as long as we consider data with high-precision period determinations ($<30\%$). Further improvements in the implementation of PGMUVI are still needed to ensure

that most sources meet this criterion.

4.4 Relationship of amplitude as a function of period

In Figure 4.8, the relationship between period and amplitude is presented for NESS sources, separated by photometric band and classified into two categories: Miras (left) and Non-Miras (right). The dashed black lines mark the division between Mira stars and Non-Mira stars. For the V band, the separation value is around 2.5 mag according to GCVS Variability Types. This separation is at approximately 0.8 mag in the near-infrared [Wood et al., 1996], as illustrated, for example, at 1.25 μm from DIRBE. Most of our amplitudes are consistent with the literature-based classifications (see Section 4.2 for reference). Disagreements could arise from the fact that, as explained above, PGMUVI is able to account for saturations/non-detections and is able to produce a more robust light curve.

When analyzing the sources located in the top left panel of the Figure 4.8 corresponding to the V band of ASAS-SN, we observe that the Non-Mira sources in this sample exhibit periods ranging from 30 to 1000 days. This is consistent with the periods for non-Miras seen in the ASAS-SN sample (the maximum period for this sub-sample was found to be approximately 950 d). There is a trend toward longer periods with amplitudes ranging up to 2.5 magnitudes. Regarding the distribution of different chemical types, it can be observed that carbon stars tend
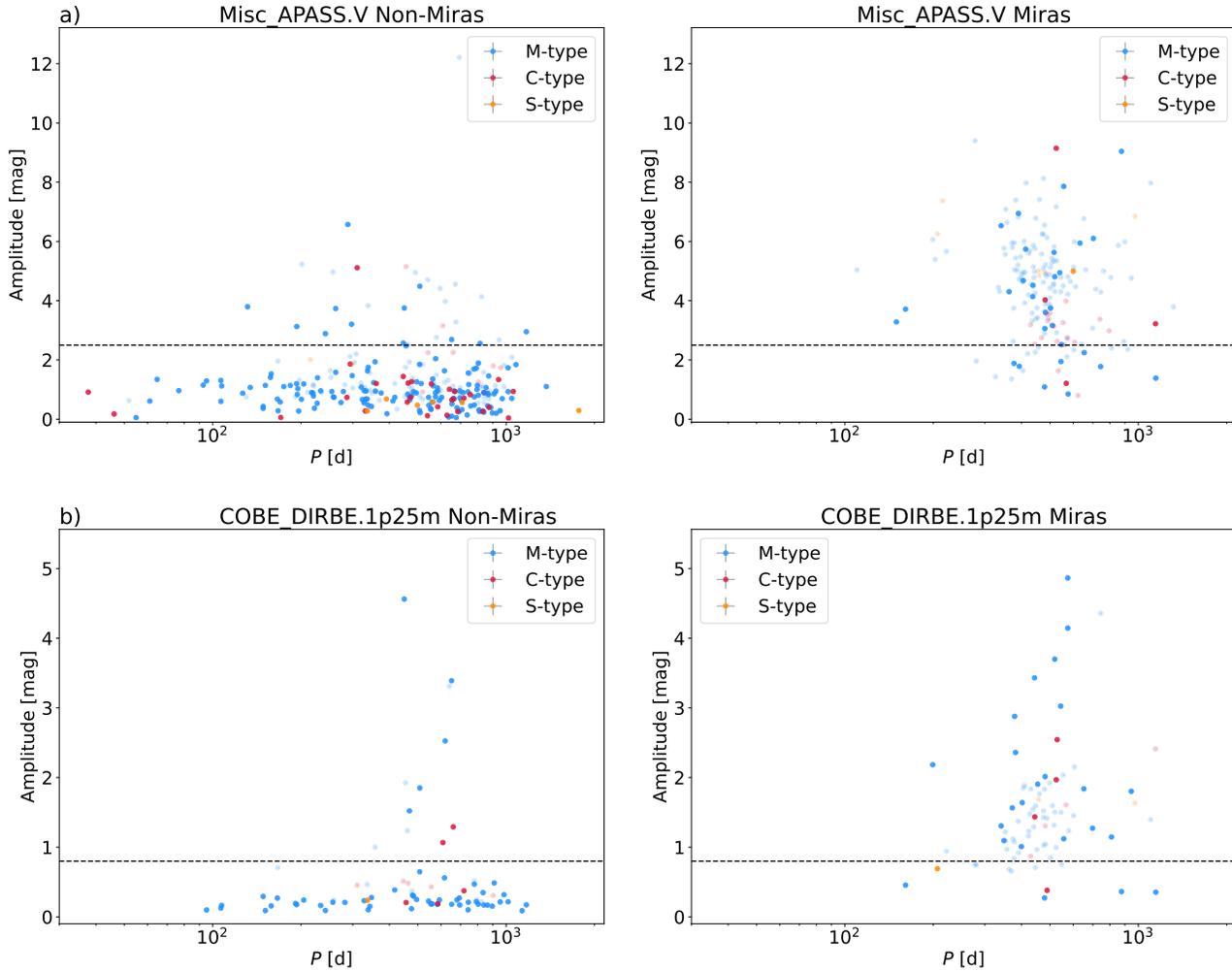
Figure 4.8: Period-Luminosity Relationship for NESS Sources. In the top left panel, the V band of ASAS-SN is depicted, where non-Mira stars are observed to be normally distributed with periods ranging from 30 to 2000 days and amplitudes of less than 2.5 mag. In contrast, in the right panel, Mira stars can be observed, showing a significant dispersion in their amplitudes and having periods ranging from 100 to 1000 days. The bottom panels show the same parameters but for the DIRBE 1.25 μm band. On the bottom left, non-Mira stars concentrate with amplitudes less than 0.8 mag and are distributed in periods between 100 and 1000 days. This could be due to the lack of light curves available for stars with periods between 30 and 2000 days in this band. There is no clear trend observed in the distributions based on the chemical type of the star.

to concentrate in periods ranging from 300 to 1000 days. Furthermore, in the right panel of the same figure, it is evident that Mira-type stars exhibit a dispersion in period values between 300 and 1000 days, as well as amplitudes ranging from 2 to 8 magnitudes. This same pattern of

behavior is repeated in the relationship between amplitude and period in the 1.25 μm DIRBE band, as shown in the lower panels of the corresponding figure.

4.5   Discussion

This thesis demonstrated that PGMUVI is a promising algorithm for the retrieval of multi-component light curve fits, and that it has a promising future of potential applications. This thesis has helped to put it into operation with real data and has allowed the authors of the code to develop improvement strategies. For example, a function will be incorporated that allows the user to initialize parameters using the LS method, which is an idea originating from this thesis. Additionally, the need to employ MCMC to search for optimal parameters has been emphasized, since the results obtained so far correspond to a minimization in a local minimum rather than a global one, which is reflected in a high dispersion of the obtained periods.

Discussion and Conclusions

In this work, fundamental concepts regarding variability in evolved stars, particularly those transitioning to the asymptotic giant branch phase, were synthesized. The physical principles associated with variations in magnitude, stemming from oscillations in the optical photosphere of a star, were discussed in accordance with existing literature. In summary, the observed radiation flux from the star, as well as the temperature and pressure of the stellar atmosphere, are modulated by the opacity of this region, which, in turn, is influenced by the material inflow and outflow. Inflow is attributed to matter levitating above the photosphere due to radial pulsations, whereas outflow is a result of the formation of molecules and dust particles accelerated by radiation pressure and friction between gas and dust.

These two factors, in turn, impact absorption in both the optical (molecular and dust) and infrared (dust) spectra. As a result of these variations, there is observed variability in the apparent magnitude of the star over time, allowing them to be classified as long-period

variables (ranging from 30 to 2000 days). However, the challenge in studying these stars lies in the difficulty of accurately determining their intrinsic luminosity, because of the combined effect of the atmosphere and interstellar medium. This, in turn, complicates the precise determination of their distances.

Extragalactic AGB stars have the advantage of statistically experiencing the same extinction, which allows for the establishment of period-luminosity relationships that facilitate distance estimation and determination of intrinsic luminosities. The uncertain distances to Galactic stars impedes their study. By collecting multiwavelength photometry for these stars, and by using AGB luminosities (which are well constrained from studies in the Magellanic Clouds), NESS has produced more reliable estimates of these distances.

We have complemented our catalog of light curves with physical parameters that are embedded in the light curves of the LPV stars from the NESS sample and we can now determine how well the parameters obtained with PGMUVI perform. Periods characterizing the fundamental mode of pulsation were provided using a total of 18 codes (9 for retrieval and 9 for homogenization and storage). These codes consulted nine surveys that offered photometric bands in both optical and near-infrared, resulting in a total of 29 photometric bands consulted[1]. The number of collected light curves is shown in Table 2.14, while the relationship between the number of data points, the maximum sampling interval, and the time baseline is illustrated in Figure 2.4. This is the first time such a multiwavelength compilation has been attempted for nearby AGB stars. The entire database will be made publicly available with the submitted paper, along with the code required to compile it.

---

[1]Although the NEOWISE survey was consulted for the W1 and W2 bands, which provided approximately 1200 light curves, nearby AGB stars are saturated in these WISE bands. Therefore, this survey was excluded from the analysis and left for future research.

With the available light curves, we proceeded to obtain derived parameters through their analysis. To achieve this, two algorithms were applied: the first one, a combination of Lomb-Scargle (LS) and Multiband Lomb-Scargle (MLS: Technique that combines light curves in different photometric bands to obtain a single periodogram.), allowed us to obtain the fundamental model through the analysis of periodograms. The Multiwavelength Lomg-Scargle algorithm is a recent development [Seilmayer et al., 2020 and VanderPlas & Ivezić, 2015]; as far as we are aware, this work marks the first time that MLS is used to study the properties of LPV light curves on a large scale.

The obtained periods were used to feed the PGMUVI algorithm, which uses Gaussian processes to fit the light curve of each source and computes parameters such as the fundamental period, the second period, the mean magnitude, and amplitudes, along with a non-parametric numerical model. Fits for individual light curves are shown in Figures 3.6, 3.7, and 3.8, demonstrating that the results are similar to those from other applications of Gaussian processes [e.g., Figure 12 in Sanders et al., 2022].

This thesis represents the first application of PGMUVI, and it has served as a way to test the code. The full code for and results of our fitting procedure will also be made available in an upcoming paper. GP-based algorithms have been applied to LPV light curves in recent studies of individual sources [e.g., Soraisam et al., 2023] as well as large samples [Sanders et al., 2022]. As far as we are aware, this thesis is the first application of such a technique towards LPVs in the Solar Neighborhood. More work is required to determine how well PGMUVI compares to other methods of fitting light curves.

The distributions of periods obtained with LS, MLS, and PGMUVI are shown in Figure 3.9. One of the challenges when studying sources with long periods is that the sampling can appear as additional signals in the periodogram, causing artifacts, especially at annual or semi-annual intervals. It is observed that the LS-MLS technique prevents the maximum value of the distribution from falling into the annual artifact, although it persists in the semi-annual one.

On the other hand, the distribution of fundamental periods from PGMUVI shifts the distribution away from the artifacts, and there is no evidence of a density in the semi-annual artifact. It is important to note that PGMUVI still does not combine information from multiple photometric bands to obtain a single fundamental period. In this thesis, the median value of the available periods for each source was taken as representative of the fundamental pulsation mode and was used to construct the previously mentioned distribution.

Finally, with the obtained period distributions, analyses of color-period relationships and mass-loss rates as a function of period were carried out. For the color-period relationships, it was found that the stars follow the trend observed by McDonald et al. [2018] in their work: a longer period corresponds to a more evolved star with increased dust content, resulting in increased extinction of starlight and therefore in redder colors.

As for the MLR-period relationships, data involving dust-production rates and CO ($J=0 \hookrightarrow 1$) emission were used. These relationships showed agreement with the models of Vassiliadis & Wood [1993] and Groenewegen et al. [2009], despite the significant dispersion in the estimated periods that makes it challenging to identify trends in the diagrams. It is important to note that these studies involve AGB stars from the Galaxy as well as the Magellanic Clouds and they are models for stars of specific masses. Determining the mass of the NESS stars was not

the focus of this thesis, so this aspect was not deeply explored.

Another relationship explored during this study was the behavior of light curve amplitudes as a function of the effective wavelength of the photometric filter used to observe the stars. In accordance with previous studies such Smith et al. [2002], the amplitude was found to decrease with wavelength, with flux variations being more pronounced in the optical spectrum than in the infrared. This can be explained by the greater sensitivity of molecular absorption to temperature changes, which in turn causes more significant variation in the optical spectrum. It is the first time that the amplitude variation with wavelength has been determined for such a large sample of Galactic AGB stars.

In summary, as a result of this thesis, a database composed of 712 NESS stars with at least one light curve available in one of the 30 photometric bands used by nine different surveys has been compiled. In total, 6774 light curves were obtained across all bands, of which 5114 are considered to contain useful information. Of this initial dataset, light curves that suffered from saturation/irregular sampling issues (NEOWISE), or curves that had too many data points, leading to high computational costs (KELT), were not fit. PGMUVI fits were obtained for the remaining 3634 light curves. Consequently, it is possible to retrieve periods and amplitudes for 674 NESS stars from these processed light curves. Additionally, there are 3634 non-parametric models of light curves available, which can be visualized graphically or used numerically. Given the need to initialize PGMUVI with LS and MLS, 672 periodograms are available to validate these PGMUVI models.

In the future, we will conduct detailed analyses to assess the quality of the analyzed light curves, applying preprocessing techniques to smooth the curves in the hope of eliminating noise

from their power spectra and extracting valuable information from the NEOWISE light curves. We will implement PGMUVI in the version that combines multiple wavelengths to obtain a unique fundamental period.

On several previous occasions, it has been mentioned that the PGMUVI is still being developed and is considered an essential tool for optimization using Bayesian inference processes. At the date of this thesis, the MCMC method, which plays a crucial role in enabling the search for local minima to optimize the model and obtain the best parameters in the hyperparameter space, has not been implemented. The inclusion of MCMC would have provided a solution to problems related to parameter degeneracy, especially in two-component models, where it is possible to adjust the relative weights of the components, making the solutions practically indistinguishable from each other. In this context, MCMC would have simplified the selection of the fundamental mode, which does not necessarily coincide with the value of the longest period obtained. In the near future, this tool is anticipated to be implemented and fine-tuned, enabling the acquisition of more precise results using the data set gathered in this study. All of these outcomes will be featured in an article highlighting innovation in the study of AGB stars in the Milky Way through the utilization of Gaussian processes, with a special emphasis on the PGMUVI algorithm.

With this work, we now have a framework for analysing large samples of multiwavelength light curves that can also be extended to AGB populations in the Galactic Bulge or in nearby galaxies.

Acknowledgements to institutions

Referencias

European Southern Observatory. 2017, Delicate bubble of expelled material around the cool red star U Antliae | ESO España. https://www.eso.org/public/spain/images/eso1730a/?lang 8

Aigrain, S., & Foreman-Mackey, D. 2022, arXiv e-prints, arXiv:2209.08940, doi: 10.48550/arXiv.2209.08940 29

Aigrain, S., & Foreman-Mackey, D. 2022, arXiv e-prints. http://arxiv.org/abs/2209.08940 77, 78

American Association of Variable Star Observers. 2019, APASS – AAVSO Photometric All-Sky Survey, https://www.aavso.org/apass 38, 39

Arnold, R. A., McSwain, M. V., Pepper, J., et al. 2020, The Astrophysical Journal Supplement Series, 247, 44, doi: 10.3847/1538-4365/ab6bdb 35, 43, 101

Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019, Publications of the Astronomical Society of the Pacific, 131, 018002, doi: 10.1088/1538-3873/aaecbe 35, 49

Boggess, N. W., Mather, J. C., Weiss, R., et al. 1992, The Astrophysical Journal, 397, 420, doi: 10.1086/171797 25, 40

Bowen, G. H. 1992, iesh, 104. https://ui.adsabs.harvard.edu/abs/1992iesh.conf..104B/abstract 19

Bowen, G. H., Bowen, & H., G. 1988, ApJ, 329, 299, doi: 10.1086/166378 19

Boyer, M. L., Srinivasan, S., Loon, J. T. V., et al. 2011, The Astronomical Journal, 142, 103, doi: 10.1088/0004-6256/142/4/103 16

Busso, M., Gallino, R., & Wasserburg, G. J. 1999, Annual Review of Astronomy and Astrophysics. www.annualreviews.org vii, 13

Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv e-prints, arXiv:1612.05560. https://arxiv.org/abs/1612.05560 35, 47

Chen, J., Luo, A.-L., Li, Y.-B., et al. 2022, ApJ, 931, 133, doi: 10.3847/1538-4357/ac66de 15

Chengalur, J. N., Lewis, B. M., Eder, J., & Terzian, Y. 1993, ApJS, 89, 189, doi: 10.1086/191843 103

Cioni, M. R. L. 2009, A&A, 506, 1137, doi: 10.1051/0004-6361/200912138 104

Demleitner, M., Derriere, S., Gray, N., Louys, M., & Ochsenbein, F. 2015, IVOA recommendation: Units in the VO, doi: 10.5479/ADS/bib/2014ivoa.spec.0523D 50

Dreyer, C., Hegmann, M., & Sedlmayr, E. 2009, Astronomy & Astrophysics, 499, 765, doi: 10.1051/0004-6361/200811603 19, 20

Flewelling, H. A., Magnier, E. A., Chambers, K. C., et al. 2020, ApJS, 251, 7, doi: 10.3847/1538-4365/abb82d 47

Fraser, O. J., Hawley, S. L., & Cook, K. H. 2008, The Astronomical Journal, 136, 1242, doi: 10.1088/0004-6256/136/3/1242 23, 69, 87

Freytag, B., & Höfner, S. 2008, A&A, 483, 571, doi: `10.1051/0004-6361:20078096` 19

Freytag, B., Liljegren, S., & Höfner, S. 2017, A&A, 600, A137, doi: `10.1051/0004-6361/201629594` 18

Gaia collaboration. 2022, Gaia DR3 Part 1. Main source, Version 05-Dec-2022 (last modified), Centre de Donnees astronomique de Strasbourg (CDS), doi: `https://doi.org/10.26093/cds/vizier.1355` 42, 103

Gaia Collaboration, Vallenari, A., Brown, A., & Prusti, T. 2022, Astronomy & Astrophysics 35, 42

Gail, H. P., Wetzel, S., Pucci, A., & Tamanai, A. 2013, A&A, 555, A119, doi: `10.1051/0004-6361/201321807` 14

García-Hernández, D. A., Zamora, O., Yagüe, A., et al. 2013, Astronomy & Astrophysics, 555, L3, doi: `10.1051/0004-6361/201321818` 13

Glass, I. S. 1973, Monthly Notices of the Royal Astronomical Society, 164, 155, doi: `10.1093/mnras/164.2.155` 48

Groenewegen, M. A., Sloan, G. C., Soszyński, I., & Petersen, E. A. 2009, Astronomy & Astrophysics, 506, 1277, doi: `10.1051/0004-6361/200912678` viii, 101, 104, 105, 115

Groenewegen, M. A., Whitelock, P. A., Smith, C. H., & Kerschbaum, F. 1998, Monthly Notices of the Royal Astronomical Society, 293, 18, doi: `10.1046/J.1365-8711.1998.01113.X` 98, 101, 104, 105

Groenewegen, M. A. T. 2022, A&A, 659, 145, doi: `10.1051/0004-6361/202142648` 46

Habing, H. J., & Olofsson, H. 2004, AGB Stars: History, Structure, and Characteristics (Springer New York), 1–21, doi: 10.1007/978-1-4757-3876-6_1 11, 21

Harvey, P. M., Bkcinsj, K. P., Wilson, W. J., & Ball, J. A. 1974, The Astrophysical Journal Supplement Series, 27, 331 21

Helou, G., & Walker, D. W. 1988, iras, 7, 1. https://ui.adsabs.harvard.edu/abs/1988iras....7.....H/abstract 28

Henden, A. 2019, AAVSO – APASS DR10 Has Arrived! (Abstract). https://app.aavso.org/jaavso/article/3493/ 35

Hoffleit, D. 1997, The Journal of the American Association of Variable Star Observers, vol. 25, no. 2, p. 115-136, 25, 115. https://ui.adsabs.harvard.edu/abs/1997JAVSO..25..115H/abstract 21

Höfner, S., & Freytag, B. 2022, in The Origin of Outflows in Evolved Stars, ed. L. Decin, A. Zijlstra, & C. Gielen, Vol. 366, 165–172, doi: 10.1017/S1743921322000199 20

Höfner, S., & Olofsson, H. 2018, The Astronomy and Astrophysics Review, 26, 1, doi: 10.1007/s00159-017-0106-5 14, 15, 22

Iben, I., & Renzini, A. 1983, Annual Review of Astronomy and Astrophysics, 21, 271, doi: 10.1146/annurev.aa.21.090183.001415 15

Infrared Processing and Analysis Center. 2013, Explanatory Supplement to the AllWISE Data Release Products, https://wise2.ipac.caltech.edu/docs/release/allsky/expsup/index.html 44

—. 2021, NEOWISE Data Release, https://wise2.ipac.caltech.edu/docs/release/neowise/ 45

Institute for Astronomy at the University of Hawaii. 2023, PANSTARRS - Panoramic Survey Telescope and Rapid Response System, https://outerspace.stsci.edu/display/PANSTARRS/ 47

Iwanek, P., Soszyński, I., Kozłowski, S., et al. 2022, The Astrophysical Journal Supplement Series, 260, 46, doi: 10.3847/1538-4365/ac6676 25

Jayasinghe, T., Kochanek, C. S., Stanek, K. Z., et al. 2018, Monthly Notices of the Royal Astronomical Society, 477, 3145, doi: 10.1093/mnras/sty838 25, 36, 90, 91, 92, 101

Jayasinghe, T., Stanek, K. Z., Kochanek, C. S., et al. 2020, MNRAS, 491, 13, doi: 10.1093/mnras/stz2711 35

Jet Propulsion Laboratory. 2009, NASA's Wide-field Infrared Survey Explorer (WISE) Mission Press Kit, https://www.jpl.nasa.gov/news/press_kits/wise-launch.pdf 45

Jordi, C., Gebran, M., Carrasco, J. M., et al. 2010, Astronomy & Astrophysics, 523, A48, doi: 10.1051/0004-6361/201015441 26, 41

Karakas, A. I. 2018, Proceedings of the International Astronomical Union, 14, 79, doi: 10.1017/S1743921318006166 24

Karttunen, H., Kröger, P., Oja, H., Poutanen, M., & Donner, K. J. 2017, Fundamental Astronomy, 6th edn. (Springer), 255–226 9

Keeley, D. A. 1970, The Astrophysical Journal, 161, 643, doi: 10.1086/150566 19

Kerschbaum, F., Maercker, M., Brunner, M., et al. 2017, Astronomy and Astrophysics, 605, A116, doi: 10.1051/0004-6361/201730665 8

Kharchenko, N., Kilpio, E., Malkov, O., et al. 2002, A&A, 384, 925, doi: `10.1051/0004-6361:20020084` 101

Kochanek, C. S., Shappee, B. J., Stanek, K. Z., et al. 2017, PASP, 129, 104502, doi: `10.1088/1538-3873/aa80d9` 35, 36

Kochanek, C. S., Shappee, B. J., Stanek, K. Z., et al. 2017, Publications of the Astronomical Society of the Pacific, 129, 104502, doi: `10.1088/1538-3873/aa80d9` 36

Lattanzio, J. C., & Wood, P. R. 2004, in Asymptotic giant branch stars (Springer), 23–104 12

Lebzelter, T., Trabucchi, M., Mowlavi, N., et al. 2019, A&A, 631, A24, doi: `10.1051/0004-6361/201936395` 17

Lebzelter, T., Mowlavi, N., Lecoeur-Taibi, I., et al. 2023, A&A, 674, A15, doi: `10.1051/0004-6361/202244241` 25

Lomb, N. R. 1976, Astrophysics and Space Science, 39, 447, doi: `10.1007/BF00648343` 66

Mainzer, A., Bauer, J., Cutri, R. M., et al. 2014, ApJ, 792, 30, doi: `10.1088/0004-637X/792/1/30` 25, 35

Masci, F. J., Laher, R. R., Rusholme, B., et al. 2019, Publications of the Astronomical Society of the Pacific, 131, 018003, doi: `10.1088/1538-3873/aae8ac` 35, 49

McDonald, I., Beck, E. D., Zijlstra, A. A., & Lagadec, E. 2018, Monthly Notices of the Royal Astronomical Society, 481, 4984, doi: `10.1093/mnras/sty2607` 94, 95, 96, 97, 98, 115

McDonald, I., & Zijlstra, A. A. 2016, The Astrophysical Journal Letters, 823, L38, doi: `10.3847/2041-8205/823/2/L38` 94

Merchan-Benitez, P., Uttenthaler, S., & Jurado-Vargas, M. 2023, A&A, 672, 165, doi: `10.1051/0004-6361/202245593` 101

Neugebauer, G., Habing, H. J., van Duinen, R., et al. 1984, ApJL, 278, L1, doi: `10.1086/184209` 28

Palmer, L. G., Wing, R. F., Palmer, L. G., & Wing, R. F. 1982, AJ, 87, 1739, doi: `10.1086/113264` 15

Parzen, E. 1962, https://doi.org/10.1214/aoms/1177704472, 33, 1065, doi: `10.1214/AOMS/1177704472` 72, 87

Pawlak, M. 2021, Astronomy & Astrophysics, 649, A110, doi: `10.1051/0004-6361/202038642` 23

Pepper, J., Pogge, R. W., DePoy, D. L., et al. 2007, Publications of the Astronomical Society of the Pacific, 119, 923, doi: `10.1086/521836` 35, 43

Price, S. D., Smith, B. J., Kuchar, T. A., Mizuno, D. R., & Kraemer, K. E. 2010, Astrophysical Journal, Supplement Series, 190, 203, doi: `10.1088/0067-0049/190/2/203` 35, 40, 41

Prusti, T., de Bruijne, J. H. J., Brown, A. G. A., et al. 2016, Astronomy & Astrophysics, 595, A1, doi: `10.1051/0004-6361/201629272` 29, 42

Rasmussen, C. E., & Williams, C. K. I. 2006, Gaussian Processes for Machine Learning. `www.GaussianProcess.org/gpml` 75, 76, 77

Renzini, A. 2015, in Astronomical Society of the Pacific Conference Series, Vol. 497, Why Galaxies Care about AGB Stars III: A Closer Look in Space and Time, ed. F. Kerschbaum, R. F. Wing, & J. Hron, 1, doi: `10.48550/arXiv.1412.6379` 23

Riebel, D., Meixner, M., Fraser, O., et al. 2010, The Astrophysical Journal, doi: `10.1088/0004-637X/723/2/1195` 22, 23, 24

Roberts, S., Osborne, M., Ebden, M., et al. 2013, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 371, 20110550, doi: `10.1098/rsta.2011.0550` 74, 76

Rodrigo, C., & Solano, E. 2020, in XIV.0 Scientific Meeting (virtual) of the Spanish Astronomical Society, 182 36

Rodrigo, C., Solano, E., & Bayo, A. 2012, SVO Filter Profile Service Version 1.0, doi: `10.5479/ADS/bib/2012ivoa.rept.1015R` 36

Samus', N. N., Kazarovets, E. V., Durlevich, O. V., et al. 2017, ARep, 61, 80, doi: `10.1134/S1063772917010085` 22

Sanders, J. L., Matsunaga, N., Kawata, D., et al. 2022, MNRAS, 517, 257, doi: `10.1093/mnras/stac2274` 26, 108, 114

Sargent, B. A., Srinivasan, S., & Meixner, M. 2011, ApJ, 728, 93, doi: `10.1088/0004-637X/728/2/93` 98

Scargle, J. D. 1982, The Astrophysical Journal, 263, 835, doi: `10.1086/160554` 23, 66

Schuster, A., Schuster, & Arthur. 1898, TeMag, 3, 13, doi: `10.1029/TM003I001P00013` 65

Scicluna, P., Waterval, S., Vasquez-Torres, D. A., Srinivasan, S., & Jamal, S. 2023, arXiv e-prints, arXiv:2308.00132, doi: `10.48550/arXiv.2308.00132` 27, 30, 79

Scicluna, P., Kemper, F., McDonald, I., et al. 2022, Monthly Notices of the Royal Astronomical Society, 512, 1091, doi: `10.1093/mnras/stab2860` 26, 27, 98

Seilmayer, M., Garcia Gonzalez, F., & Wondrak, T. 2020, arXiv e-prints, arXiv:2001.10200, doi: 10.48550/arXiv.2001.10200 114

Silverberg, R. F., Hauser, M. G., Boggess, N. W., et al. 1993, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 2019, Infrared Spaceborne Remote Sensing, ed. M. S. Scholl, 180–189, doi: 10.1117/12.157825 40, 90

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163, doi: 10.1086/498708 15

Smith, B. J., Leisawitz, D., Castelaz, M. W., & Luttermoser, D. 2002, The Astronomical Journal, 123, 948, doi: 10.1086/338647 ix, 18, 107, 116

Smith, T. C., Henden, A., & Terrell, D. 2010, in Society for Astronomical Sciences Annual Symposium, Vol. 29, 45–53 38

Soraisam, M. D., Bildsten, L., Drout, M. R., et al. 2018, The Astrophysical Journal, 859, 73, doi: 10.3847/1538-4357/aabc59 26

Soraisam, M. D., Szalai, T., Van Dyk, S. D., et al. 2023, arXiv e-prints, arXiv:2306.10783, doi: 10.48550/arXiv.2306.10783 29, 114

Speagle, J. S. 2019, A Conceptual Introduction to Markov Chain Monte Carlo Methods. https://arxiv.org/abs/1909.12313v2 82

Srinivasan, S., Sargent, B. A., & Meixner, M. 2011, A&A, 532, A54, doi: 10.1051/0004-6361/201117033 98, 104

Suh, K.-W. 2021, The Astrophysical Journal Supplement Series, 256, 43, doi: 10.3847/1538-4365/ac1274 46

The MACHO collaboration, Alcock, C., Allsman, R. A., et al. 2000, ApJ, 542, 281, doi: `10.1086/309512` 23

van de Sande, M., Walsh, C., & Danilovich, T. 2020, Monthly Notices of the Royal Astronomical Society, 495, 1650, doi: `10.1093/mnras/staa1270` 23

VanderPlas, J. T., & Ivezić, 2015, The Astrophysical Journal, 812, 18, doi: `10.1088/0004-637X/812/1/18` 66, 114

Vassiliadis, E., & Wood, P. R. 1993, ApJ, 413, 641, doi: `10.1086/173033` viii, 100, 101, 102, 103, 106, 115

Whitelock, P. A., Feast, M. W., & Leeuwen, F. V. 2008, Monthly Notices of the Royal Astronomical Society, 386, 313, doi: `10.1111/j.1365-2966.2008.13032.x` 101

Whitelock, P. A., Feast, M. W., Marang, F., & Groenewegen, M. A. T. 2006, Monthly Notices of the Royal Astronomical Society, 369, 751, doi: `10.1111/j.1365-2966.2006.10322.x` 35, 48, 57

Wilson, A. G., & Adams, R. P. 2013, Gaussian Process Kernels for Pattern Discovery and Extrapolation 77, 78

Wood, P. R. 1979, ApJ, 227, 220, doi: `10.1086/156721` 19

Wood, P. R., Sebo, K. M., Wood, P. R., & Sebo, K. M. 1996, MNRAS, 282, 958, doi: `10.1093/MNRAS/282.3.958` 109

Wright, E. L. 2009, WISE - Wide-field Infrared Survey Explorer, `https://www.astro.ucla.edu/~wright/WISE/index.html` 45

Wright, E. L., Eisenhardt, P. R., Mainzer, A. K., et al. 2010, Astronomical Journal, 140, 1868, doi: `10.1088/0004-6256/140/6/1868` 44