PSL UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à Observatoire de Paris

Optimised modelling of pulsating stars for the extragalactic distance scale

Defended by
Garance Bras
on March 19th 2026

Doctoral school n°127
Astronomie et Astrophysique en Île-de-France

Speciality
Astronomy and Astrophysics

Prepared at
Laboratoire d'Instrumentation et de Recherche en Astrophysique

Jury members :

Yveline LEBRETON LIRA, Observatoire de Paris	*President* *Examinator*
Katrien KOLENBERG KU Leuven, University of Antwerp	*Reviewer*
Denis MOURARD Lagrange, Université de la Côte d'Azur	*Reviewer*
Louise BREUVAL ESA, STScI, Baltimore	*Examinator*
Grzegorz PIETRZYŃSKI CAMK, Warsaw	*Examinator*
Antoine MÉRAND ESO, Garching	*Invited*
Nicolas NARDETTO Lagrange, Observatoire de la Côte d'Azur	*Invited*
Pierre KERVELLA LIRA, Observatoire de Paris	*Supervisor*

LIRA Observatoire de Paris | PSL

“Ô nuit enchanteresse, divin ravissement”
Du zénith, à Nadir

Acknowledgments

Maintenant que ma thèse touche à sa fin, je voudrais sincèrement remercier Pierre pour son soutien, ses conseils, sa rigueur scientifique, mais surtout son incroyable gentillesse. Durant ces trois ans et demi, je me suis sentie entièrement accompagnée dans mes progrès. Chaque fois que je me suis sentie piétiner dans mon travail, nos discussions m'ont aidée à me relancer tout en me rassurant sur mes capacités. Merci en particulier pour toujours avoir eu un mot gentil qui a court-circuité chaque début de découragement. J'ai aussi beaucoup apprécié l'autonomie que j'ai pu avoir en travaillant avec toi, tout en n'étant jamais isolée.

My PhD began in a storm, at the RRL/Cep conference in La Palma. I retain a vivid memory of arriving at Madrid airport, followed by the sense of loneliness I felt upon learning that all flights to the Canary Islands had been cancelled due to a storm. Yet, even from afar, Pierre was able to give me the impetus to find a way to reach the island (which I eventually managed to do 48 hours later) and not to rush back to Paris. My heartfelt thanks go to Louise and Paulina for their tremendous support at that moment, even though we had not yet met. Since I am mentioning Louise, I would like to thank you in particular, not only for welcoming me at the conference and introducing me to various colleagues, but also for your friendship, your sorority, and your presence at every stage of my PhD. From the very beginning through to the writing phase (which I spent with a copy of your thesis on my desk throughout as a source of inspiration) you remained my light at the end of the tunnel, and having you as a member of my jury today is deeply moving.

I also thank the members of the jury for agreeing to serve on it and for their careful reading of my manuscript.

I had the opportunity to work with many pleasant colleagues. I thank Nicolas for your good humour (I will long remember our trips while in Marrakech), as well as for your scientific ideas, which greatly stimulated me, in particular your passion (or persistence) regarding the mysterious p-factor. I thank Antoine for your exceptional rigour and your willingness to respond to my SPIPS-related issues. I also thank Boris for teaching me how to become a real *spipser*. I thank Alexandre for guiding me through my first steps in interferometry. I must admit that I was apprehensive at first, but I ultimately greatly enjoyed the observations. Thank you to the

whole CHARA team on Mount Wilson for their warm welcome. To attend a quartet concert in the dome hosting Hubble's telescope was not part of my plan but I really enjoyed it.

Thanks to Grzegorz and Wolfgang for their interest in my work and for their encouragement. Thanks to all my Polish colleagues (Piotr and Bartek in particular), as well as those who are not quite Polish (hello Vincent), with whom I shared many very convivial moments, even without a particular fondness for vodka. I would also like to thank those with whom I had the pleasure of carrying out observations at Cerro Murphy, Cezary for whistling "An American in Paris", and Mikołaj for his excellent meals and for rescuing me from desert spiders.

I also think of the other PhD students in the team with whom I spent time. To Wilma, with whom I greatly enjoyed attending my second edition of RRL/Cep, and to Manon, for our walks on Porquerolles. I wish you, as well as the other new PhD students in the team, a journey at least as fulfilling as mine.

I would like to thank the various occupants of Building 5 for the positive atmosphere overall. In particular, I thank Miguel for showing me behind the scenes of the observatory and for being so attentive to the integration of newcomers. I also thank you all for the kindness and support you showed upon hearing the news of my pregnancy, and later at the birth of my son.

Pendant ces quelques années, il y a eu beaucoup de va-et-vient pour partager mon bureau. Parmi ceux qui sont restés le plus longtemps, je remercie Eitan pour sa bonne humeur et Lucie pour avoir supporté le ronron incessant du tire-lait, quand il finissait par m'agacer.

Je voudrais également remercier les professeurs que j'ai rencontrés qui m'ont encouragée à trouver ma voie. Merci à Jean-Michel Laclaverie de m'avoir permis de dénicher une formation qui me correspondait en me parlant du magistère d'Orsay.

Merci à Samuel pour nos rendez-vous tout au long de ma première année. Merci pour ton écoute et tes conseils qui m'ont permis de prendre un peu de recul. Également merci à Rhita pour nos discussions vis-à-vis de la maternité qui ont été très importantes pour moi.

Je remercie mes amis de longue date pour leur soutien immuable. Aurélie, Emma, Juliette P., Émilie (ça y est, la petite dernière a enfin terminé ses études), les copains de la raclette, Juliette V. (pour ne m'avoir presque pas retardée avec tes "podcasts"), Sara (pour les goûters crêpes et autres réjouissances), Arthur ~~et~~, Maurine, à ceux qui ont parcourus des kilomètres pour venir fêter mes 25 ans. Merci à Jack de m'avoir fait confiance pour me proposer d'investir le festival de Fleurance.

Pendant ces quelques années, j'ai eu le bonheur de me vider la tête avec diverses activités musicales. Je remercie toutes les personnes que j'ai côtoyées dans ce cadre pour les bons moments et le plaisir de faire de la musique ensemble. Les jubilants d'ACHOR, les camarades

violoncellistes de l'OSCO (oui Lucile il est enfin l'heure de sortir *legato* et le *Sartarello*), les copines de quatuor Blandine, Annie. Enfin un mot pour Catherine et Jean-Paul, même si je pourrais en faire un paragraphe, car c'est bien au-delà des moments musicaux que je me dois de vous remercier ! Merci pour tout le soutien que vous m'avez apportée pendant ces quelques années, et de m'avoir si gentiment libéré quelques heures sous un plafond étoilé pour finir dans les temps ma rédaction. J'y ai cherché une Céphéide mais n'en ai pas trouvé...

Je remercie toute ma famille ainsi que ma belle-famille pour avoir écouté mes histoires cosmiques, Rémi pour avoir, avec beaucoup de passion, rêvé de la respiration des étoiles. Merci à mon père de m'avoir accompagnée au bout de la nuit pour apprendre le ciel ensemble, même sous le froid, même menacés par un troupeau de vaches bien curieuses. Je remercie aussi ma mère de m'avoir permis de reconnaître mon goût pour les sciences. Merci à ma soeur, Perrine, de ne pas juger la qualité littéraire de ma thèse, qui est, je l'admets, bien plus modeste que tes productions.

Des Céphéides galactiques aux cookies galactogènes, merci à toi, Nadir, de m'avoir (parfois) permis de dormir et de m'avoir (occasionnellement) laissé rédiger lorsque je m'occupais de toi. Merci de m'avoir appris à relativiser, de tous les jours me rappeler que le bonheur ne se trouve pas forcément dans une thèse parfaite. À l'heure où je finis ce manuscrit, cela fait 8 mois et demi que tu es né. Pendant ce temps, tu as appris à manger, à rire, à babiller, à ramper, à te tenir assis, à jeter, tu as même dévoilé une petite dent. Il aura fallu autant de temps pour que tu mûrisses dans mon ventre et que tu te prépares à affronter le grand jour. Moi c'est le temps qu'il m'a fallu pour écrire ma thèse, comme quoi, j'ai bel et bien accouché une deuxième fois...

Remercier Nadir, c'est aussi remercier ma petite famille, et remercier celui qui partage ma vie. Alors, plus que tout, merci Elio, d'être là, toujours, en particulier depuis ma performance ratée sur cet exercice de mécanique... Merci de m'apporter la joie, le rire, de me soutenir dans mes projets les plus fous, d'être venu faire le voyage de mes rêves dans le désert du Chili et d'avoir fini par aimer l'espagnol avec moi. Merci de croire en moi et de me donner la confiance que je n'ai pas toujours. Merci pour bien plus de choses au-delà de la thèse sur lesquelles je ne suis pas forcément capable de mettre des mots...

Résumé

Les étoiles variables pulsantes, notamment les Céphéides et RR Lyrae, sont des outils puissants pour la détermination des distances. Les relations Période-Luminosité permettent de déduire leur luminosité intrinsèque, ensuite comparée à leur magnitude apparente. Par ailleurs, la méthode de la parallaxe de pulsation utilise la relation trigonométrique entre la variation du rayon et du diamètre angulaire d'une étoile pour en déterminer la distance. Cette dernière fait intervenir un paramètre, le facteur de projection, qui convertit les vitesses radiales observées à la vitesse de la photosphère. Ce paramètre, entièrement dégénéré avec la distance, montre d'importantes variations statistiques, nous empêchant d'utiliser la parallaxe de pulsation à grande échelle. Une compréhension plus fine de la pulsation est requise afin de déterminer l'origine de cette dispersion.

Dans cette thèse j'étudie tout d'abord le facteur de projection dans le cas des RR Lyrae, à l'aide du code SPIPS et de nouvelles mesures de vitesses radiales, afin de le comparer aux Céphéides. Je montre que la valeur et la dispersion du facteur de projection des RR Lyrae sont assez similaires à ce qui a été observé pour les Céphéides et que les différentes méthodes de mesure des vitesses radiales ont peu d'impact sur cette dispersion à condition de considérer des données homogènes. Je présente aussi mes travaux relatifs à la variation du coefficient d'assombrissement centre-bord lors de la pulsation, un ingrédient clé du facteur de projection, grâce à des observations interférométriques de la Céphéide η Aql.

Dans un second temps, je développe des courbes de lumière génériques pour les Céphéides, à l'aide des modèles multi-bandes SPIPS. Ces courbes génériques permettent pour une période de pulsation donnée de prédire la forme de la courbe de lumière dans une large sélection de bandes photométriques. Elles sont particulièrement efficaces pour modéliser des courbes de lumière sur des observations rares ou peu précises, comme c'est le cas pour les Céphéides extragalactiques.

Mots clés : Observations, Analyse de données, Apprentissage automatique, Physique Stellaire, Modélisation, Cosmologie, Etoiles Pulsantes, Echelles des distances

Abstract

Pulsating variable stars, particularly Cepheids and RR Lyrae stars, are powerful tools for determining distances. Period-luminosity relations allow us to deduce their intrinsic luminosity, which is then compared to their apparent magnitude. In addition, the parallax-of-pulsation method uses the trigonometric relationship between the radius variation and the angular diameter variation to determine the distance. This latter technique involves a parameter, the projection factor, which converts the observed radial velocity to the photospheric pulsation velocity. This parameter, which is entirely degenerate with distance, shows significant statistical variations, preventing us from using parallax-of-pulsation on a large scale. A more detailed understanding of pulsation is required in order to determine the origin of this dispersion.

In this thesis, I first study the projection factor of RR Lyrae stars, using the SPIPS code and new radial velocity measurements, in order to compare it with Cepheids. I show that the value and dispersion of the projection factor for RR Lyrae stars are quite similar to what has been observed for Cepheids and that the different methods of measuring radial velocities have little impact on this dispersion, provided that homogeneous data are considered. I also present my work on the variation of the centre-to-limb darkening coefficient during pulsation, a key ingredient of the projection factor, using interferometric observations of the Cepheid η Aql.

Secondly, I develop generic light curves for Cepheids using SPIPS multi-band models. These templates make it possible to predict the shape of the light curve for a given pulsation period in a large variety of photometric bands. They are particularly effective for modelling light curves based on rare or imprecise observations, as is the case with extragalactic Cepheids.

Keywords : Observations, Data Analysis, Machine Learning, Stellar Physics, Modelling, Cosmology, Pulsating Stars, Distance Scale

Contents

Acknowledgments i

Résumé iv

Abstract v

Table of Contents v

Introduction 1

1 Pulsating stars for the distance scale 3
1 Stellar pulsation . . . 3
1.1 The Colour-Magnitude Diagram . . . 3
1.2 Classes of pulsators . . . 4
1.3 The mechanism of pulsation . . . 6
2 Pulsation for distances . . . 7
2.1 Period-Luminosity relations . . . 7
2.2 Parallax-of-pulsation . . . 11
2.2.1 Interferometric measurements of pulsating stars . . . 12
2.2.2 Surface-Brightness-Colour Relations . . . 13
2.2.3 Radial velocities from spectroscopy . . . 13
2.2.4 The projection factor . . . 14
3 Other distance measurements . . . 15
3.1 Geometric distances . . . 15
3.2 Other empirical distance determination mehods . . . 16
4 The cosmic distance scale . . . 18
4.1 The classical distance scale . . . 18
4.2 A bit of cosmology . . . 19
4.3 The Hubble tension . . . 20

2 The projection factor of pulsating stars 22
1 The SPIPS parametrisation . . . 22
1.1 An implementation of the parallax-of-pulsation technique . . . 22
1.2 The full parametrisation . . . 23
1.3 The various observational datasets . . . 25
2 Data on Galactic RR Lyrae stars . . . 27
2.1 Radial Velocity data . . . 27
2.2 Photometric data . . . 29

2.3 Input parameters . . . 32
3 Results of the modelling . . . 33
3.1 Model convergence . . . 33
3.2 Results on the p-factor of RR Lyrae stars . . . 36
3.3 The lack of correlation with other parameters . . . 37
4 A more complex atmospheric motion? . . . 38
4.1 Discriminate layers in the atmosphere . . . 38
4.2 Role of the asymmetry of the spectral lines . . . 39
5 Paper I . . . 41
6 Possible decoupled motion between the photosphere and the atmosphere . . . 67
6.1 Interferometric observations of η Aql . . . 67
6.2 The limits of calibration . . . 68
6.2.1 Secondary calibration . . . 68
6.2.2 Systematic errors on the second lobe visibilities . . . 70
6.3 Theoretical and observed variation on limb-darkening parameter . . . 71

3 Modelling the light curves of Cepheids **75**
1 The standard approach . . . 75
1.1 Light curve . . . 76
1.2 Standard parametrisation . . . 78
1.2.1 Fourier decomposition . . . 78
1.2.2 Principal Component Analysis . . . 79
1.3 The link between light curves in different photometric bands . . . 80
2 Towards a more efficient modelling of light curves . . . 81
2.1 Our implementation of the principal component analysis . . . 82
2.1.1 The training sample . . . 82
2.1.2 Results of the Principal Component Analysis . . . 84
2.2 Predicting light-curve shapes . . . 84
2.3 Light-curve templates . . . 87
2.3.1 Test sample in observed filters . . . 87
2.3.2 Test sample in other filters . . . 88
3 Modelling extragalactic Cepheids . . . 89
3.1 Observation of extragalactic Cepheids . . . 90
3.1.1 HST photometry of Cepheids . . . 90
3.1.2 JWST photometry . . . 91
3.1.3 Period estimation . . . 91
3.2 Template fitting . . . 94
3.2.1 SH0ES procedure . . . 94
3.2.2 New fitting strategy . . . 96
3.2.3 Comparison and means of improvement . . . 97
4 Paper II . . . 99

4 Discussion and future research **121**
1 Observing the pulsation to understand the p-factor . . . 121
1.1 Tomographic analysis of the atmosphere of pulsating stars . . . 121
1.1.1 A look at Gaia Radial Velocity curves . . . 121
1.1.2 Tomographic spectral masks . . . 122
1.1.3 Complex cross-correlation profiles . . . 123
1.2 Variability of the limb-darkening: extension to other targets . . . 126
2 The limitation of light curve templates . . . 129

2.1 Average behaviour 129
2.2 Reproducing the secondary bump 131
2.3 Template fitting for precise light curve interpolation 133
2.3.1 Fitting K_i coefficients 133
2.3.2 Introducing mean magnitudes variations 134
2.3.3 Phase-coverage-dependent fitting approach 135

Conclusions and perspectives **137**

List of publications **140**

Acronyms **144**

Bibliography **147**

Introduction

The measurement of distances in astrophysics is a key to understand the structure and scale of the Universe. In particular, determining the distances to galaxies is crucial given the vast physical scales involved. Moreover, because of the finite speed of light, distance measurements also provide access to the cosmic time at which these structures are observed.

It is through the study of variable systems that many distance indicators have been discovered. In particular, pulsating variable stars constitute one of the main standard candles, thanks to the relation between their variability period and their intrinsic luminosity, discovered by Henrietta Leavitt at the beginning of the twentieth century. In addition, observations of Type Ia supernovae, resulting from the thermonuclear explosion of white dwarfs, have made it possible to estimate distances out to extremely distant galaxies. By combining these two indicators, it has been possible to quantify the cosmological recession of galaxies, characterised by the Hubble constant H_0.

Nevertheless, the observation of distant Cepheids requires state-of-the-art instrumentation, with high sensitivity and regular temporal sampling, in order to properly trace their variability curves and to extract the essential parameters: the pulsation period and the mean luminosity.

Furthermore, by comparing the variation of the physical size of pulsating stars with the variation of their angular diameter, it is possible to estimate their distance in a geometrical way, using the parallax-of-pulsation method. This method relies on the estimation of the projection factor, which is required to infer the variation of the physical radius from spectroscopic observations of the radial velocity of the line-forming regions in the stellar atmosphere. The determination of this factor requires both knowledge of the radial velocity and of the angular diameter variations, obtained through photometric and/or interferometric measurements.

My thesis focuses on the modelling of pulsating variable stars, in particular Cepheids and RR Lyrae stars, with the aim of improving distance measurements using these standard candles. I investigate the modelling of stellar pulsation for the application of the parallax-of-pulsation method, using multi-technique observations. I also present a new tool to model Cepheid light curves from sparse coverage, which are typical of extragalactic Cepheids.

In Chapter 1, I present a general introduction to the concepts necessary for understanding this thesis. I first explain the pulsation mechanisms of pulsating variable stars, focusing in particular on Cepheids and RR Lyrae stars. I then describe the different methods to measure distances based on these pulsating stars. Finally, I present several other distance determination methods which, when combined with Cepheid observations, make it possible to establish distances in the Universe through the extragalactic distance scale.

Chapter 2 presents my work on the projection factor of nearby RR Lyrae stars. I describe the modelling tool I used and the results of the modelling. The projection factor of both Cepheids and RR Lyrae stars, exhibits significant star-to-star variations, which remain unexplained. I propose two approaches to investigate this dispersion, based on spectroscopic observations on the one hand, and interferometric observations on the other hand.

Chapter 3 is dedicated to the modelling of Cepheid light curves through a multi-band photometric analysis. After recalling the classical methods employed to analyse light curves, I present the templates I developed using analysis methods applied to a training sample with well-characterised variability. I then propose the methodology to apply these templates to observations of extragalactic Cepheids.

In Chapter 4, I propose perspectives and discussion on the studies carried out in this thesis. Firstly, regarding the projection factor, I present several directions that would merit further attention in future research, extending the work already performed, both on radial velocity measurements and on the limb darkening of Cepheids. Finally, I present a critical discussion on the applicability of the photometric templates developed in this work.

Chapter 1

Pulsating stars for the distance scale

1 Stellar pulsation

This thesis is focused on pulsating stars. Before studying these objects, it is important to understand what are stellar pulsations and where they come from.

1.1 The Colour-Magnitude Diagram

To introduce the family of stars I have been studying, I will start from the Colour-Magnitude Diagram (CMD), as pulsating stars are often identified thanks to this graphical tool.

The CMD, the observational counterpart of the HR (Hertzsprung-Russell) diagram, is the main tool to visualise stellar evolution. It plots the absolute luminosity (or the absolute magnitude, as logarithmic measurements of flux) of the stars with respect to their temperature (or their colour, linked to the temperature through the Wien's displacement law). Several groups can be identified in this diagram. Firstly, the majority of stars appear in the so-called main sequence. Stars in this region follow a general relation with the temperature being correlated with the luminosity. The main sequence is characterised by fusion of hydrogen (H) in the core of the stars. The star is maintained in hydrostatic equilibrium, when thermal pressure from the H-burning exactly compensates the gravity forces.

After the exhaustion of H in the core, the star evolves to cooler temperatures. A H-burning shell is then established, leading to an expansion of the envelope and placing the star on the Red Giant Branch (RGB). This region in the CMD corresponds to very luminous and large stars, but cooler than main sequence stars. The high pressure in the core from gravity forces can then induce helium (He) burning in the inner regions, surrounded by a shell of burning H. Hence, the star moves to the hotter side of the HR diagram, before He is also exhausted. This causes what is called a blue loop in the evolutionary path of the star. For intermediate mass stars, the triggering of He-burning moves the star towards the so-called horizontal branch. Then, the rapid cooling starts again, corresponding to the Asymptotic Giant Branch (AGB).

The exact position of a star in the CMD tells a lot about its past and its future. From the

initial mass of the star, which is also related to its luminosity in the main sequence, we can predict at what point it will enter the RGB stage and how long all these steps will last. They are often represented in the CMD through evolutionary tracks.

The Instability Strip (IS) is a region identified in the CMD, where we observe a large fraction of pulsating variable stars. It is represented as a slightly inclined branch with a certain width in temperature, starting near the main sequence and crossing the RGB (see Fig. 1.1). In the region of the main sequence, the IS starts at temperatures around 7500K and extends towards the top of the CMD at cooler temperatures, around 5500K when crossing the RGB. The width of the band is around 1000K. In this region of the CMD, the stars exhibit a region of partial ionisation of He, which is a key ingredient for the pulsation mechanism, as I will detail afterwards. The IS is populated by stars in different stages of their evolution, but mainly stars that have already exhausted H in their core. During their evolution, stars can cross the IS several times. The first crossing occurs right after the end of core H burning, when stars are moving towards the RGB for the first time. This phase is extremely fast, hence only a few pulsating stars are observed in their first crossing of the IS. The subsequent crossings occur during the blue loop, which may, depending on the mass of the star, cross the IS one or two times. The IS visually includes a specific selection of pulsating variable stars. In general, the different classes of pulsators are identified in specific regions on the CMD. The IS itself is decomposed into several groups.

1.2 Classes of pulsators

There exists a large variety of pulsating stars. The location of these different classes are shown in the CMD in Fig. 1.1

We first focus on stars that are located in the IS of the CMD. Classification properties of pulsators can be found, for instance, in Clementini (2010). The most famous ones are the Cepheids, named after their prototype star δ Cephei. Classical Cepheids are more massive than the Sun, around 3 to 10 solar masses. Because of their relatively high mass, they evolve rapidly towards the RGB and are particularly young and luminous. They pulsate at the time scale of a day typically, with period ranging from about 1 to 100 days for Cepheids pulsating in their fundamental mode. When pulsating in the first overtone, their period may be shorter. The term Cepheids also refers to type II Cepheids, themselves divided in different subclasses. Type II Cepheids have smaller masses (about $0.8M_{\odot}$), but comparable periods. However, they are older populations than the Classical Cepheids, thanks to their lower masses. Classical Cepheids are mainly evolving on the blue loop of their evolutionary path and Type II Cepheids are evolving towards the AGB stage, after the horizontal branch. RR Lyrae stars have lower masses than the Sun (typically $0.8M_{\odot}$), and are less luminous and smaller than Cepheids. They pulsate more rapidly, with periods on the order of half a day. They are located at the intersection of the horizontal branch and the IS. Due to their lower initial mass, they are also older than Classical Cepheids. At the bottom of the IS, we find δ Scuti pulsators, which are distinct from the Cepheid and RR Lyrae classes. These pulsators have shorter cycles of a few hours and are main-sequence stars with masses up to $2M_{\odot}$. They exhibit smaller amplitudes, of the order of a tenth of magnitude

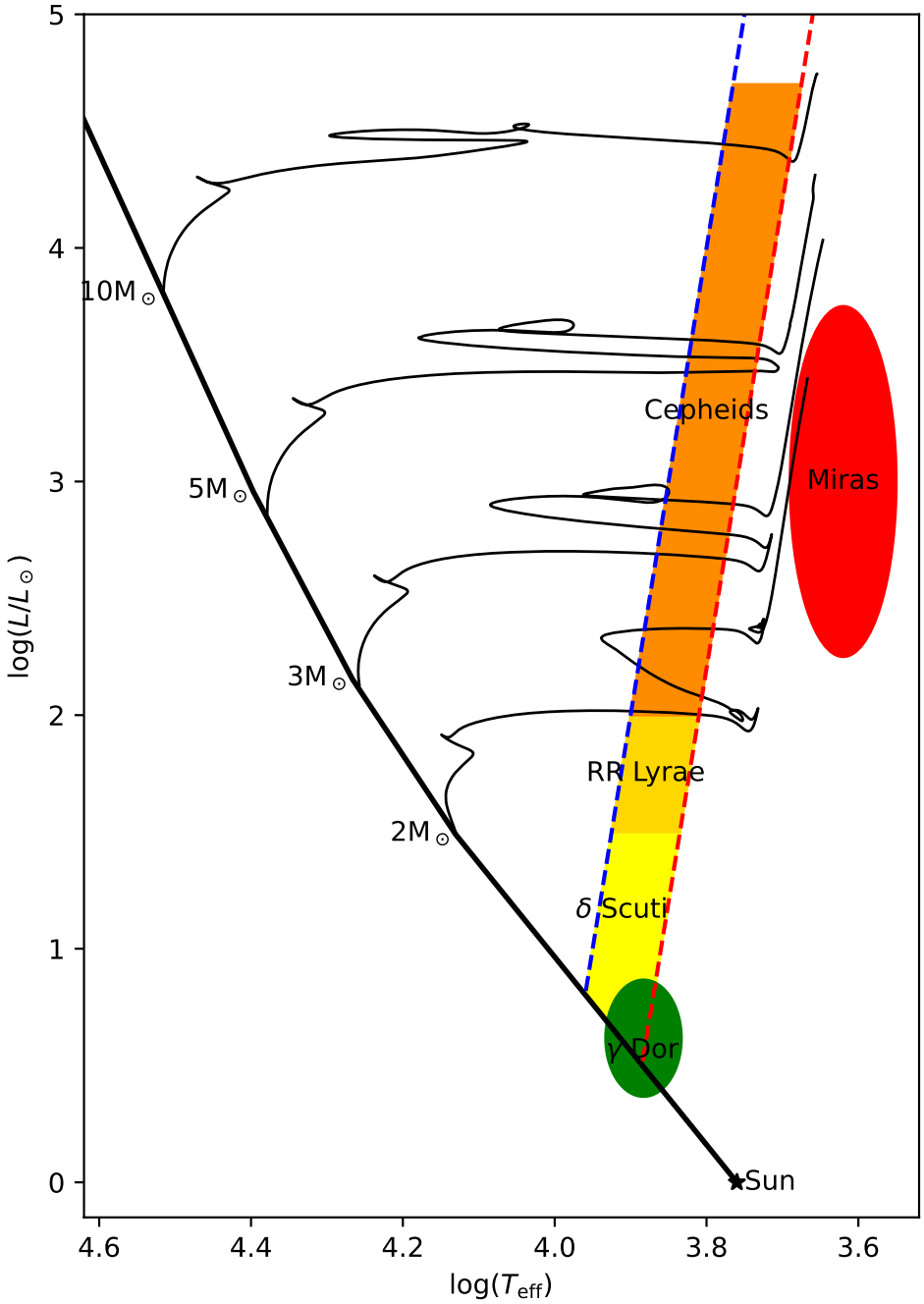


Figure 1.1: HR diagram with the most common classes of pulsators presented in the text. The approximate zero-age main sequence, theoretical evolutionary tracks for 2, 3, 5, $10M_{\odot}$ from Costa et al. (2025) and the classical IS are plotted.

in V compared to up to 2 magnitudes for Cepheids and RR Lyrae stars.

Other classes of pulsating variables can be found in other regions on the CMD. Mira pulsators refers to giant stars in their latest stage of evolution, just before evolving into white dwarfs. They present long periods (more than 100 days) and large amplitudes of variability. They can have oxygen or carbon-rich cores as they have undergone He fusion. As they have extended envelopes, evolving towards planetary nebulae, they often show dust clouds surrounding them. They typically show variable amplitudes, but their pulsation periods are precise enough to count them as regular pulsators. Various other variable sources are grouped in the class of semi-regular variables, which stand for stars indicating signs of periodic variations but with less regularity than the stars gathered in the previously detailed groups. It can refer either to very long-period and evolved stars like for instance Betelgeuse, or lower-amplitude and shorter-period variables. To finish this overview, we can mention the γ Doradus variables, which were first assigned as δ Scuti variables but are not pulsating in the same modes. Typical periods of γ Doradus variables are around one day, with small amplitudes of less than a tenth of magnitude. They are main sequence stars, outside the typical region for δ Scuti variables, towards the cooler side of the IS.

1.3 The mechanism of pulsation

Stellar oscillations may be explained mathematically from the various equations of stellar physics, i.e. the continuity equation, the momentum equation, the Poisson equation, the energy equation and the equation of state. Those equations simplify thanks to the adiabatic hypothesis, which is realistic in stellar interiors, as the typical time of energetic transfer is much larger than the typical time of oscillations, and using the linear perturbation form of the equations. The specific case of radial oscillations, which are observed in Cepheids and RR Lyrae, describes purely spherical oscillations. The resulting set of equations can be simplified by means of a detailed mathematical analysis to obtain a standard harmonic oscillator equation (Gautschy et al. 1995). Given the dispersion relation of the system, we can then identify two different modes of pulsation, p-modes and g-modes, with p-modes (resp. g-modes) describing oscillations for which the restoring force is pressure (resp. gravity). We also note that p-modes are the only modes that can produce radial pulsations.

Maintaining pulsations over a long duration requires a mechanism to excite the oscillation. In the case of classical pulsators (i.e. mainly Cepheids and RR Lyrae stars), this mechanism is called the κ-mechanism in reference to the opacity κ. The opacity generally decreases with the temperature, hence, tends to decrease towards the stellar interior. However, close to partial ionisation zones, the opacity is exceptionally high. Starting from the compression (Phase 1 in Fig. 1.2), the gas in the helium partial-ionisation zone is compressed and heated, which increases its opacity. As a result, radiation is trapped in this region and the internal pressure builds up (Phase 2), driving the expansion of the envelope (Phase 3). During expansion, the gas cools, leading to a decrease in opacity. Radiation can then escape more efficiently, reducing the pressure support and causing the region to contract again (Phase 4). This cycle can then repeat. It is

represented in Fig. 1.2.

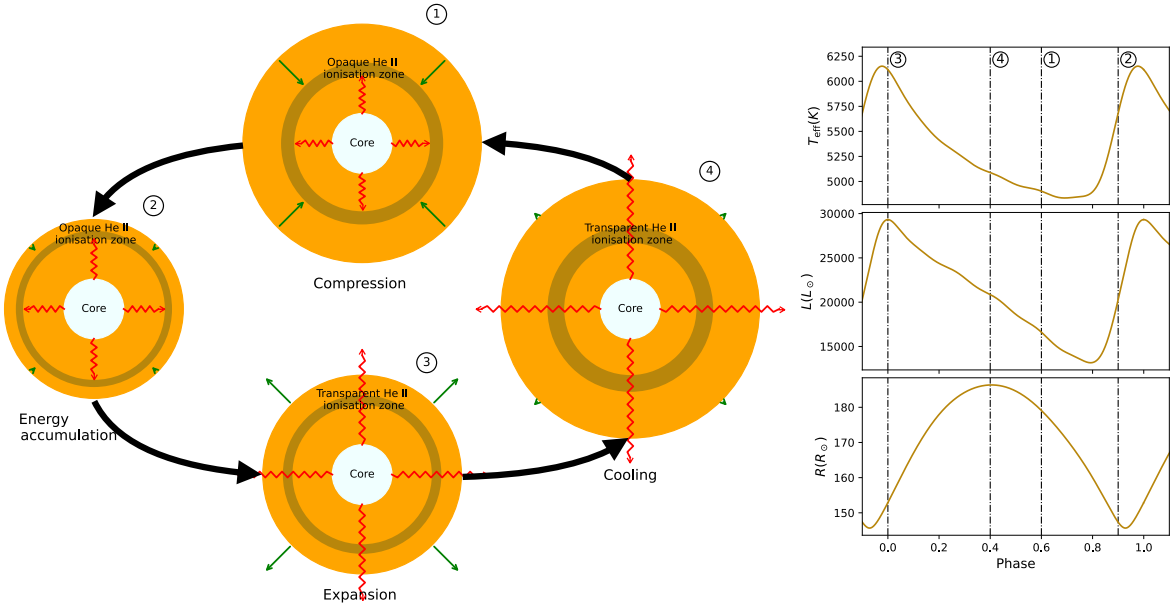


Figure 1.2: Schematic representation of the pulsation induced by the κ mechanism. Radiation is represented with red zigzag arrows and green arrows symbolise the photosphere velocity. On the left, the temperature, bolometric luminosity, and radius curves of a typical Cepheid (RS Pup) are shown.

In the case of Cepheids, the κ-mechanism occurs due to the second ionisation zone of He (De Somma et al. 2024), as well as for RR Lyrae stars and δ Scuti variables. For the κ-mechanism to be efficient, it requires that the pulsation period is of the same order as the typical thermal time scale. This is the case for He II in the classical IS, which explains the existence of this specific region in the CMD.

2 Pulsation for distances

2.1 Period-Luminosity relations

Period-Luminosity Relations (PLRs) have been discovered thanks to the work of the U.S. astronomer Henrietta Leavitt, at the beginning of the twentieth century. She is one of the most famous female figures of modern astronomy as her discovery revolutionised astronomy. Henrietta Leavitt used to work as a *human computer*, a job description that was used for groups of women scientists assigned to do the technical jobs. Her duty was to analyse photographic plates to tabulate the brightness and the positions of the stars. In particular, she did it for variable stars (mainly Cepheids) in the Magellanic Clouds (Leavitt 1907). Thanks to this work she identified the relation between the pulsation period of these variable stars and their observed brightness, which would later be known as the PL relation, or Leavitt law (Leavitt et al. 1912). At that time, she estimated that the Small Magellanic Cloud was roughly 100 times further than the

distance to the already known Cepheid prototype δ Cep, by measuring the apparent brightness of an analogue of δ Cep. The linear relation between the logarithm of the pulsation period and the apparent magnitude for Small Magellanic Cloud Cepheids is particularly interesting, as all Cepheids in this system are approximately at the same distance. This means that this relation reflects the intrinsic luminosity of the star. Recently, Breuval et al. (2025) investigated the work of Henrietta Leavitt with modern instrumentation and knowledge on PLRs and demonstrated the outstanding pioneering work realised at this epoch, as can be seen in Fig. 1.3.

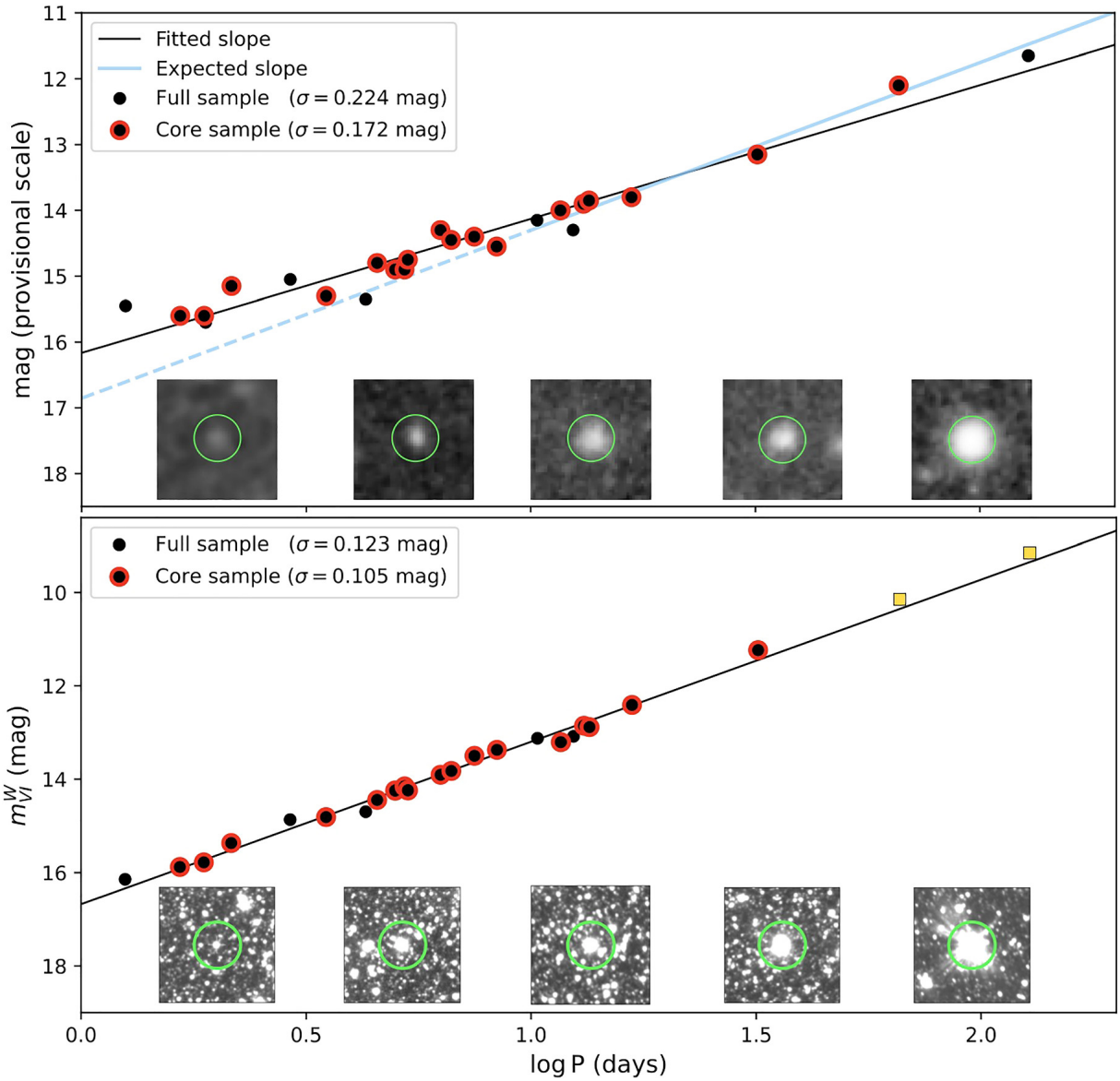


Figure 1.3: Leavitt law from Henrietta Leavitt's observations (top) and from more recent OGLE (Optical Gravitational Lensing Experiment) photometry (bottom). Figure reproduced from Breuval et al. (2025).

Period-Luminosity relations are a way to link the pulsation period to the intrinsic luminosity of a Cepheid. This elevates the Cepheids to the role of standard candles. This expression implies that observing a Cepheid anywhere in the sky, and after determining its pulsation period, we

can directly deduce its intrinsic luminosity, which is usually an unknown quantity. The comparison with the apparent luminosity makes it possible to estimate the distance of the star. In a simplified model, we can consider a pulsating star with radial adiabatic pulsations and with homogeneous density ρ. In such a physical object, the oscillation can be modelled as a sound wave. The period of the fundamental mode corresponds to the sound-crossing time of the star, which is linked to the sound velocity c_s, through the relation $P \propto \frac{R}{c_s}$. Moreover, in an ideal gas $c_s \propto \sqrt{\frac{p}{\rho}}$, where p is the pressure. At the level of orders of magnitude, the central pressure can be estimated as $p_c \sim \frac{GM^2}{R^4}$, thanks to the hydrostatic equilibrium equation. Combining these different orders finally gives that the pulsation period is inversely proportional to the square root of the density. This is the behaviour derived by August Ritter in 1879 from dimensional arguments, and then used by Eddington (1917). To go further, the density can be expressed in terms of the mass and the radius. Involving mass-luminosity relations and the Stefan-Boltzmann law, we end up with the period being a function of the luminosity and the temperature, which translates into $\log P = A \log L + B \log T + \text{const}$. This is formally a Period-Luminosity-Colour relation. The parameter in temperature reflects the width of the IS.

Most of the different pulsator populations can be identified from their pulsation period. In addition, among one class of pulsator, we can distinguish the stars that are pulsating in their fundamental modes or overtones. This translates in shorter periods for overtone pulsators for a star of given mass and structure. A star pulsating in the first overtone means that, instead of a global oscillation of the star, its internal structure is divided into two cavities, producing a radial node in the inner parts. Overtone pulsators do not follow the same PLRs than their fundamental analogues, making them important to identify. In particular, at the same period, first-overtone pulsators are intrinsically fainter than fundamental-mode pulsators, which is why they lie on a distinctly offset PL relation. In addition, around 10 days of period, PLRs can show peculiar behaviours. This is generally attributed to the resonance between the second overtone and the fundamental modes. It shows as a possible break in the PL relation, making some authors decide to fit two different slopes before and after the break (Bhardwaj et al. 2016; Ngeow et al. 2006).

Technically, PLRs are derived within a specific photometric system. Each PLR is a particular projection of a more global PL–colour dependence, as shown in Fig. 1.4. Specifically, the slope of the PL relation is steeper at visible wavelengths than in the infrared (IR). However, the dispersion of the PLR is not the same in all photometric bands. The most precise PL relations, i.e. the ones with minimal scatter, are found in the redder bands. This is due to the width of the IS and the dependence on temperature. At shorter wavelengths, the Spectral Energy Distribution (SED) is steeper, which makes temperature differences between the stars in the IS more apparent. Interestingly, it is possible to derive PLRs from minimum or maximum luminosities, as Henrietta Leavitt originally did, as well as from mean luminosities. The latter is the most commonly used, as the different amplitudes of Cepheid light curves bias the two other versions. It is also important to mention that the mean is taken in luminosity space,

and then the magnitude corresponding to this mean luminosity is computed. In this thesis, I will usually refer (abusively) to "mean magnitudes" when I actually mean the magnitude of the mean luminosity, for instance in Chapter 3.

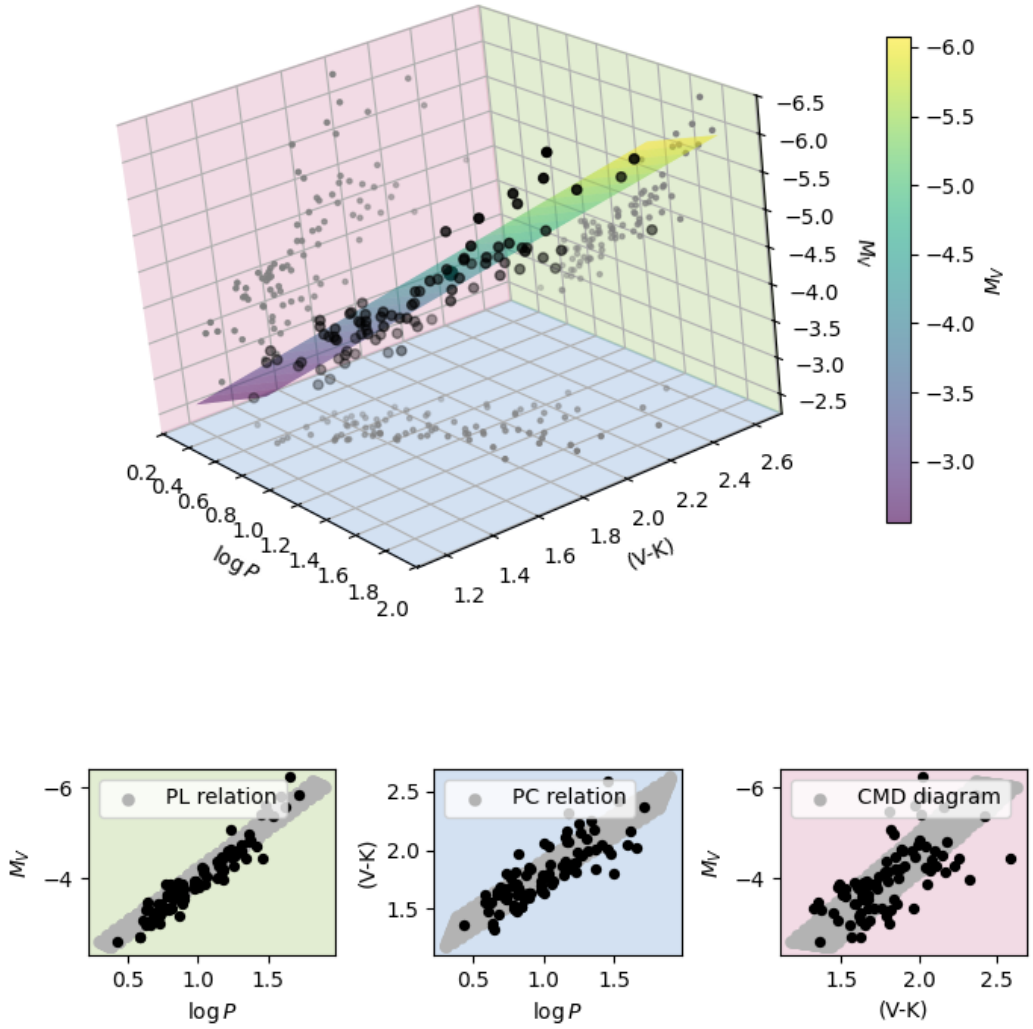


Figure 1.4: The Period-Luminosity-Colour plane of classical Cepheids and its 2D projections, leading to the CMD, PL and Period-Colour (PC) relations. Real data from Breuval et al. (2021) are added. Figure inspired by the representation from Madore et al. (1991).

When applying PL relations, it is important to take into account the effect of the interstellar dust. Indeed, apparent magnitudes are extinguished due to the dust along the line of sight, by a

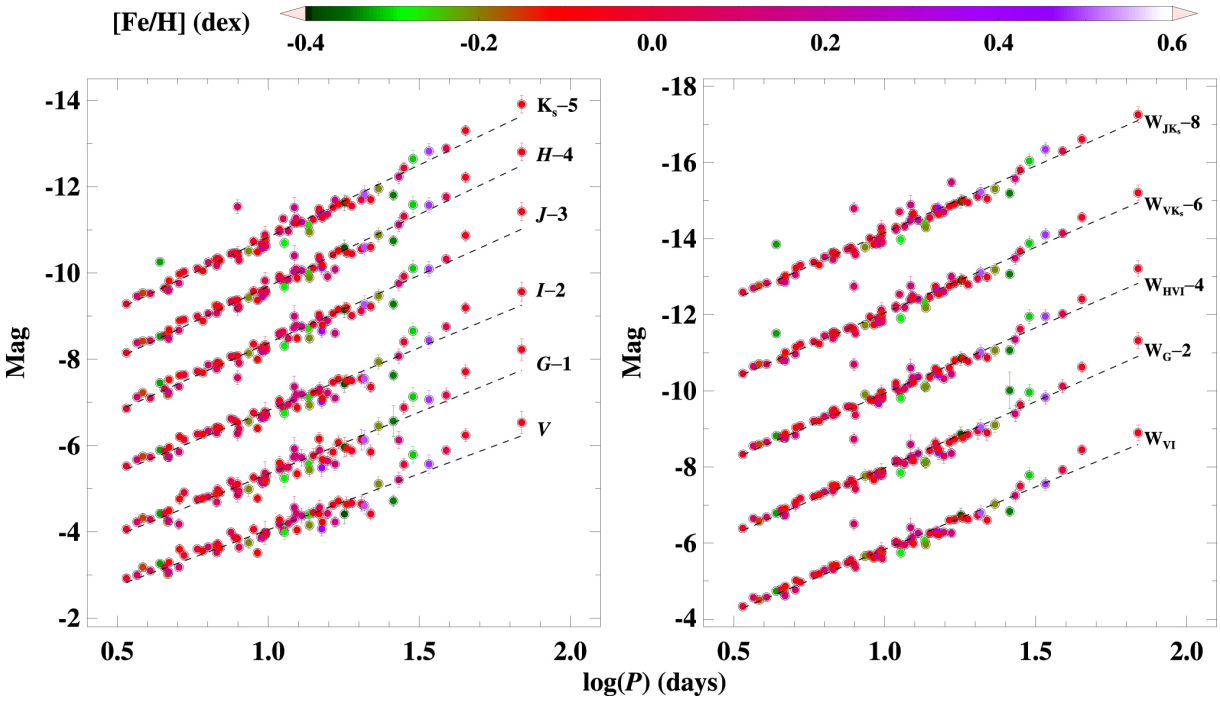


Figure 1.5: PLRs and PWRs for Milky Way Wepheids in various different photometric bands from Bhardwaj et al. (2023).

quantity that depends on the photometric band A_{band}. The extinction decreases with increasing wavelength. This makes astronomical objects appear redder than they really are, justifying the term of reddening from the interstellar dust. The different reddening values for the different stars are responsible for an additional dispersion in the observed PLRs. This makes reddening an essential parameter to determine in order to derive distances from PLRs. This effect is mitigated by the use of redder photometric bands. In addition, Madore (1982) introduced the Wesenheit function in PLRs, to express them in indexes independent of the reddening. It corresponds to a combination of magnitudes and colours such that $W = m_Z - R(m_X - m_Y)$ where R is the total to selective ratio $R = \frac{A_Z}{A_X - A_Y}$. In the case of the B and V band, this yields $W = V - R_V(B-V)$. Assuming the reddened magnitudes $V = V_0 + A_V$ and $B = B_0 + A_B$, one can show that W depends only on dereddened quantities, $W = V_0 - R_V(B_0 - V_0)$. This proves the potential of the Wesenheit function as a reddening-free magnitude. Period-Wesenheit relations (PWRs) then replace the standard PLRs, lowering the dispersion, thanks to a better treatment of the reddening. An example of calibrated PLR together with PWR is shown in Fig. 1.5. Wesenheit indexes however still rely on reddening laws through the R parameter.

2.2 Parallax-of-pulsation

Besides PLRs, distances to pulsating stars can also be determined using independent methods, in particular the parallax-of-pulsation technique

The general idea of the parallax-of-pulsation technique comes from the trigonometric relation between the distance d, the radius R and the angular diameter θ of an astrophysical object

$\tan \frac{\theta}{2} = \frac{R}{d}$. For a classical star, this method is not applied as it is rare to know the physical radius. For a pulsating star, however, we can retrieve the radius variation, by integrating the pulsation velocity with time, and compare it to the variation of angular diameter. Although already suggested by Lindemann (1918), the method is traditionally known as the Baade–Wesselink method (Baade 1926; Wesselink 1946). Throughout this thesis, I instead use the term parallax-of-pulsation, both because it describes the method more transparently and because it avoids the asymmetry that occurs when only some techniques are widely named after their (male) authors, whereas others, such as the Leavitt law, are often not. A scientifically driven terminology is, in my view, preferable.

A similar method is also used to determine the distance of Type II Supernovae. In this context, it is called the expanding photosphere method.

2.2.1 Interferometric measurements of pulsating stars

The variations of angular diameter can be observed directly by interferometry. By combining different telescopes, spaced by a baseline, an interferometer can reach the angular resolution of a telescope with a mirror of the same diameter as the baseline. This technique is essential to measure the angular sizes of distant stars. An interferometer measures the interference pattern between each pair of telescopes, by varying the optical delay between the two telescopes. These interference patterns are used to compute the visibility, a measure of the contrast of the fringes. The visibility reaches 1 where the interference shows high contrast and 0 when no interferences are observed. It is usually plotted against the spatial frequency, the ratio between the baseline and the wavelength of observation. The observation of a point-source, or a non-resolved object gives a constant visibility equal to 1. For resolved objects, the fringes are blurred at some spatial frequencies, depending on their angular size. Mathematical derivations give the visibility function for a Uniform Disk (UD) $V = \frac{2J_1(\pi\theta B/\lambda)}{\pi\theta B/\lambda}$, where J_1 is the Bessel function of the first kind of order one. This model is fitted to observations to measure a UD diameter. In the case of a limb-darkened disk, i.e. a disk with a decreasing intensity from centre to limb, the expression of the visibility differs. Different models of Limb-Darkening (LD) are used, the classical ones being the intensity decreasing with a power-law or a linear law. The first interferometric observation of a Cepheid was done by Mourard et al. (1997) on δ Cep, using the "Grand Interféromètre à 2 Telescopes". However, at that time, the uncertainty was too high on the visibility measurement to observe the variations of the angular diameter. The pulsation of ζ Gem was observed by Lane et al. (2000) using the Palomar Testbed Interferometer (PTI), which also gives the first application of the parallax-of-pulsation technique using interferometric measurement of a Cepheid. This method was also applied to measurements using the Navy Prototype Optical Interferometer (NPOI, Armstrong et al. 2001). Since then, a few other Cepheids were observed and their angular diameter curves were measured. The constant improvements of the instrumentation strengthen the different measures of the largest Cepheids (η Aql, ℓ Car, ζ Gem, δ Cep, β Dor, W Sgr, Y Sgr). The VINCI (Very Large Telescope INterferometer Commissioning Instrument) instrument at the VLTI (Very Large Telescope Interferometer) improved the precision of angular diameter measurements (Kervella et al. 2004a). The CHARA (Center for High Angular Reso-

lution Astronomy) array (Mérand et al. 2005) and the SUSI interferometer (Sydney University Stellar Interferometer, Davis et al. 2006) also provided various measurements.

2.2.2 Surface-Brightness-Colour Relations

The surface brightness F is the flux per surface area. It then relates to the total luminosity through the relation $L \propto FR^2$. Using the magnitude scale, the luminosity becomes magnitude and the surface brightness is noted S, which gives $S_\lambda = m_{\lambda_0} + 5\log(\theta_{LD})$. The LD diameter is used as it better corresponds to the photometric size of the star. In parallel, Stefan-Boltzmann law describes the dependence of the luminosity on $T_{\rm eff}^4$. Converting on the magnitude scale, we can use colour indices to constrain the temperature. This yields the Surface-Brightness-Colour Relations (SBCRs), of the form $F_{\lambda_1} = a(m_{\lambda_1} - m_{\lambda_2}) + b$ where a and b are the parameters calibrated on interferometric measurements of stellar diameters. The SBCRs together with the first equation derived, are then used to determine the diameter of any star. Those relations are calibrated in different pairs of photometric bands, as well as on different type of stars, to properly link the colour index to the temperature. Common calibrations used for pulsating stars are Fouqué et al. (1997), Graczyk et al. (2021), and Salsi et al. (2021), whereas Bailleul et al. (2025) and Kervella et al. (2004b) are exclusively calibrated on Cepheids. These relations have been extensively used in parallax-of-pulsation applications. For instance, Storm et al. (2011) applied them to Cepheids in the Magellanic Clouds to calibrate PLRs.

2.2.3 Radial velocities from spectroscopy

Radial velocities (RVs) are determined through spectroscopy, using the displacement of the spectral lines, induced by the Doppler effect. The velocity is given by the formula of the classical Doppler effect.

$$v_{\rm rad} = c\frac{\lambda_{\rm obs} - \lambda_{\rm ref}}{\lambda_{\rm ref}}$$

The position of each line in the spectrum is compared to the reference wavelength of the transition. Each line is associated to one measurement of RV. Consequently, different structures may be probed by different lines or set of lines. In order to improve the precision of these determinations, and to use spectra with a lower signal-to-noise ratio (S/N), it is common to use Cross-Correlation Functions (CCFs). The CCF method consists of cross-correlating a line mask with the spectra, in order to obtain a general CCF profile which is a mean profile of the lines included in the mask. The choice of the mask is important as it selects the region of the spectrum where the lines should appear. They should be adapted to the object that is observed. We will then be cautious to use masks corresponding to the spectral types of the considered stars. In addition, we distinguish between binary masks or templates for the cross-correlation. Binary masks select the wavelength windows where the lines should appear, whereas the templates add a weighting on the most visible lines. The determination of the velocity from the CCF profile is similar to what would be done on a single line profile. Due to the thermal broadening of the spectral lines, the common way is to fit a Gaussian profile to the lines.

Alternatively, it is possible to use the broadening function method (Rucinski 1999). This

approach consists of deconvolving the observed spectrum by a synthetic spectrum in order to retrieve the velocity distribution of the observed object. In that case, the global maximum of the broadening function provides an estimate of the object's systemic velocity, while additional peaks or asymmetries in the function reveal finer velocity substructures.

Velocities also need to be corrected from the motion of the observer. Velocities are usually considered in the barycentric reference frame.

2.2.4 The projection factor

From spectroscopic measurements, the velocity that is deduced is the RV, i.e. the velocity projected on the line of sight. However, we are interested in the pulsation velocity of the photosphere, which is directed from the centre to the exterior of the star. This translates into the RV being the integration of the projected pulsation velocity on the line of sight over the spherical hemisphere facing the observer. This computation yields a converting factor $v_{\rm puls} = 1.5 v_{\rm rad}$. The LD has also an effect on the intensity of the lines and hence on the weighting of the infinitesimal velocity vectors. The intensity profile should be added to the integration. This can be done knowing the LD law. For common computations using a linear emerging intensity profile parametrised by the coefficient u such that $I(\mu) = I_0 - u(1-\mu)$ where $\mu = cos(\theta)$ and θ is the angle between the line of sight and the normal to the surface, we obtain $v_{\rm puls} = (\frac{3}{2} - \frac{u}{6}) v_{\rm rad}$. Finally, the velocity determined by spectroscopy, even after the previous corrections, does not compare exactly to the velocity of the photosphere. It results that the radius curve is not totally comparable to the angular diameter curve either from SBCR or interferometry. Indeed, the line forming regions that we observe are in the upper atmosphere, above the photosphere. Neglecting this effect is equivalent to considering that the star is pulsating uniformly as an ensemble. The projection-factor (p-factor) is used to incorporate these differences between the radial and the pulsation velocity. It includes both measurable effects, such as the projection effect, and not measurable effects, from the limb-darkening to the velocity gradient in the atmosphere. By definition, it is such that

$$\Delta R(t) = p \int_0^t (v_{\rm rad}(\tilde{t}) - v_\gamma) d\tilde{t}$$

By comparing two integrated quantities of the star, the p-factor inherently incorporates the complex dynamics of pulsating stellar atmospheres, including the presence of shock waves (see Mihalas 2003, for a review), which induce velocity gradients and line profile asymmetries. These dynamical effects are now taken into account in theoretical models of pulsation (Nardetto et al. 2004), leading to a decomposition of the different contributions to the p-factor (Nardetto et al. 2007). The geometric p-factor includes only the projection effect and the effect of the LD. It is then multiplied given the velocity gradient and the conversion to the photospheric velocity. From the definition, and given the relation $d = \frac{2\Delta R}{\theta}$, it follows that the p-factor is fully degenerate with the distance. This means that it is mandatory to know the p-factor to determine the distance and reversely. For this reason, it is important to calibrate the p-factor to apply the parallax-of-pulsation technique as a distance determination technique. In the literature, $p = 1.36$ is a standard value taken for applications of the method in the visible domain (Burki

et al. 1982). Different authors performed independent calibrations of this factor, using different references for the distance, different datasets and different approaches. A general picture of the observationally calibrated period-p-factor relations is shown in Fig. 1.6. The relations are plotted with plain lines for the period range where they have been calibrated, and in dashed lines where they are extrapolated for other classes of pulsators. I added, as a dotted line, the relation corresponding to a purely geometric p-factor, $p = \frac{3}{2} - \frac{u(P)}{6}$, where the LD coefficient $u(P)$ is obtained using the stellar parameters of RR Lyrae and Cepheids respectively from Bras et al. (2024) and Trahin et al. (2021) and the tables of Claret et al. (2011).

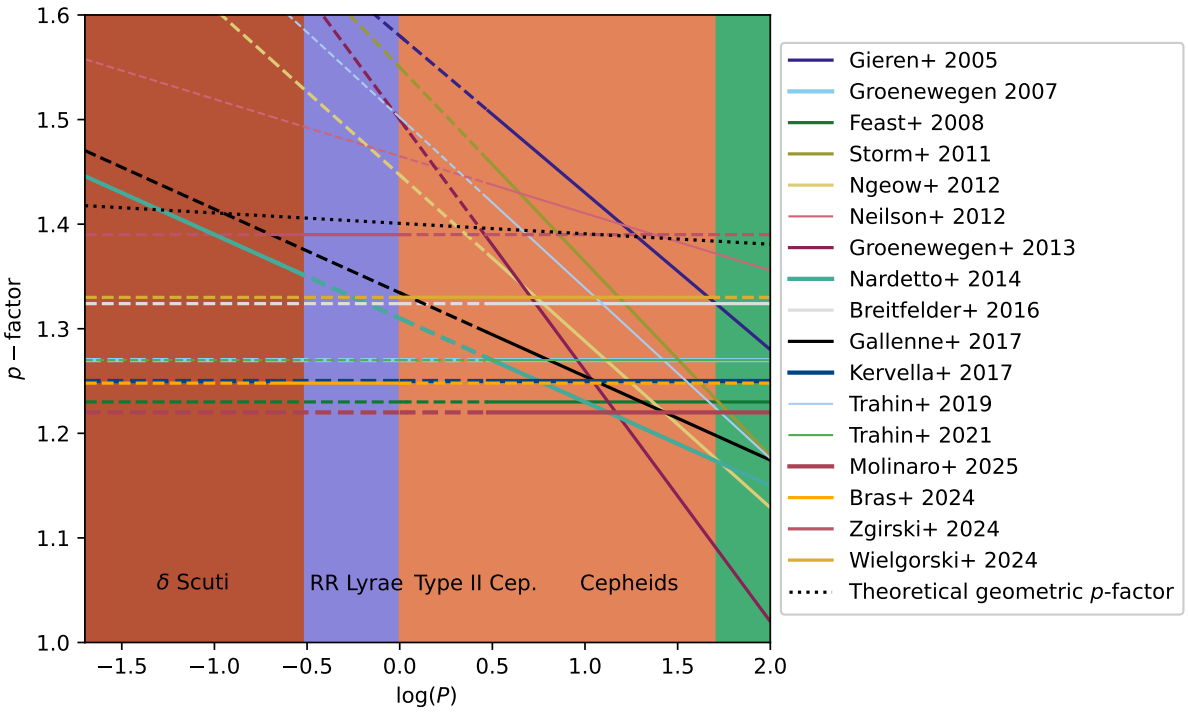


Figure 1.6: Period-p-factor relations from the literature from observational calibrations. (Updated figure from Trahin et al. (2021)).

This figure illustrates the current issue on the parallax-of-pulsation technique. The calibration of the underlying p-factor parameter is highly imprecise. This is however not only due to the different methodologies, as, even in single surveys, individual stars can present extremely scattered values. Interestingly, Trahin et al. (2021) found a p-factor value for a few Cepheids higher than 1.5, which would imply a limb-brightened stellar disk. These results suggest that we are missing some physical ingredient in our analysis of the p-factor.

3 Other distance measurements

3.1 Geometric distances

Parallax: The main method for geometric distance measurements is the annual parallax. This method is based on the apparent motion of a nearby object, compared to a fixed background,

when Earth is orbiting around the Sun. The apparent motion of the source on the sky is an ellipse whose semi-major axis is the parallax ϖ. The parallactic displacement is mathematically related to its distance d through the relation $\tan \varpi = \frac{1\text{au}}{d}$, which simplifies in $d = \frac{1}{\varpi}$ in astronomical units. When expressing ϖ in arcseconds, $\frac{1}{\varpi}$ gives the distance in parsecs. This method applies well for nearby objects, as the parallax ϖ decreases while the distance increases. For technical reasons, we are limited in the smallest parallax angle we can measure. However, it is the most robust distance measurements as it does not rely on any calibration of a relation such as PLRs. Today, parallax measurements are obtained thanks to the Gaia mission for a large number of objects. Gaia parallaxes are precise, but their accuracy is subject to small biases, which led the community to introduce zero-point offsets to correct the raw measurements (Lindegren et al. 2021). They are dependent on the magnitude, the colour and the position of the star, due to various technical and processing issues. Moreover, Gaia measures the parallax for the photocentre, which may bias the parallaxes for peculiar objects, in particular for binaries.

Eclipsing binaries: Eclipsing binaries are of great help to determine distances. They are systems of two stars for which the orbital plane is seen nearly edge-on, resulting in partial or total eclipses when they pass in front of each other. Observations of the photometric eclipses provide the relative radii of both components with respect to the separation of the system, as well as the flux ratio and the inclination. It is complemented with RV measurements which give information on the linear orbital scale. Combining both types of measurements makes it possible to determine the physical size of the two stars of the system which is compared to the angular size, obtained by using the same kind of SBCRs as for the parallax-of-pulsation. For nearby Galactic systems, direct interferometric measurements of angular diameters can also be used as an alternative. Detached eclipsing binaries (DEBs) are at the origin of the most precise measurements of the distances to the Magellanic Clouds to date (Graczyk et al. 2020; Pietrzyński et al. 2019). These measurements deeply rely on SBCRs, which have to be precisely calibrated. In particular, these relations must be calibrated to remain valid for the observed eclipsing binary systems, across the different temperature regimes and stellar properties involved.

3.2 Other empirical distance determination mehods

In addition, there exists other empirical methods than the PLRs, that are usually calibrated using geometric distances.

TRGB: The Tip of the Red Giant Branch (TRGB) method (Freedman et al. 2019, 2020; Lee et al. 1993) uses the statistics of stars in a particular regime, that are easily identified in the CMD. The tip is the region of the CMD at the end of the RGB, at the moment of the He flash, when the temperature of ignition of He is reached in the core. This causes the star to immediately move to the horizontal branch, causing a discontinuity in the CMD. From stellar physics, this point is reached at a given magnitude. The consequence is a visible edge on the upper part of the RGB in the CMD of a galaxy. In practice, once the photometry of all the stars of a galaxy is obtained, the RGB stars are identified using the CMD. Then, the luminosity function is studied

to pin the apparent magnitude where there is a brutal decrease of the number of objects, through the derivative of the luminosity function. This apparent magnitude is compared to the absolute magnitude of the TRGB to derive the distance. This method requires calibrating the absolute magnitude of the TRGB in the corresponding photometric band, typically in I band, where the TRGB is flat and less sensitive to metallicity.

JAGB: Following the same philosophy, a recent stellar distance indicator is the J-region Asymptotic Giant Branch (JAGB) method (Lee et al. 2025; Zgirski et al. 2021). It is based on the population of carbon-rich stars in the AGB stage. These stars are expected to have similar magnitudes, in particular in the NIR, as they are supposed to be the late stage of a mass-range restricted progenitor population, around 2 to 4 solar masses. The most common photometric band to observe them is in J, where they have an absolute magnitude of -6.2 mag. The JAGB population is easily identified in the CMD, using the combination of J and K magnitudes for instance (Madore et al. 2020), and the average magnitude is compared to the expected absolute magnitude to derive the distance. Such measurements only require one epoch of multi-band imaging on a galaxy to determine its distance, which explains the success of this method. In addition, AGB stars are particularly bright, and therefore visible up to large distances.

SNIa: Another distance indicator largely used for extragalactic distances is in the observation of Type Ia Supernovae (SNIa) (Brout et al. 2022; Riess et al. 1998). This phenomenon is triggered when a white dwarf in a multiple system is accreting material from its companion. This causes the white dwarf to exceed the Chandrasekhar mass, and to release energy in a very luminous explosion (around an absolute magnitude of -19.3 mag in V). Because of the similarity in mass of the SNIa progenitors, the explosions exhibit similar luminosities, leading to their use as standard candles. In practice, the light curve of a SNIa shows a singular pattern, with a bright luminosity peak followed by an exponential decay. In reality, SNIa are often considered only as "standardizable" candles as their actual peak brightness is related to the rate of decline of their light curve. Their brightnesses are also linked to the heavy element composition of the explosion. They are useful to determine distances, provided that the relation between the peak luminosity and the absolute luminosity is calibrated, thanks notably to Cepheids, TRGB, or JAGB. Because of their extreme luminosity, SNIa are observable in very distant galaxies, much farther than Cepheids, which is essential for the cosmic distance scale.

Among other distance measurements methods we can also cite masers (Reid et al. 2019), surface brightness fluctuations (SBF, Jensen et al. 2025), fundamental plane of elliptical galaxies (FP, Said et al. 2025), Tully-Fisher relation (TF, Boubel et al. 2024), type II supernovae (SNIIe, de Jaeger et al. 2020), and expanding photosphere method on SNIIe (EPM, Csörnyei et al. 2023).

4 The cosmic distance scale

As we have seen from the previous overview of distance indicators, many methods exist, but they can not be applied for the whole range of distances in the Universe. Some are more specific to close objects, other for large distances. In addition, several distance indicators need to be calibrated, based on other methods. The consequence is that one distance indicator is not enough to calibrate the full distance scale. This limitation introduced the need for a global methodology to reach the most distant galaxies, which is usually refereed to as the distance ladder. This approach is further motivated by its key role in measuring the expansion rate of the Universe.

4.1 The classical distance scale

The classical distance scale combines several distance indicators from the closest one to increasing distance, and includes the calibration of secondary distance indicators. It also offers some variants which I will not detail here. A global picture is depicted in H0DN Collaboration et al. (2025) and in Fig. 1.7.

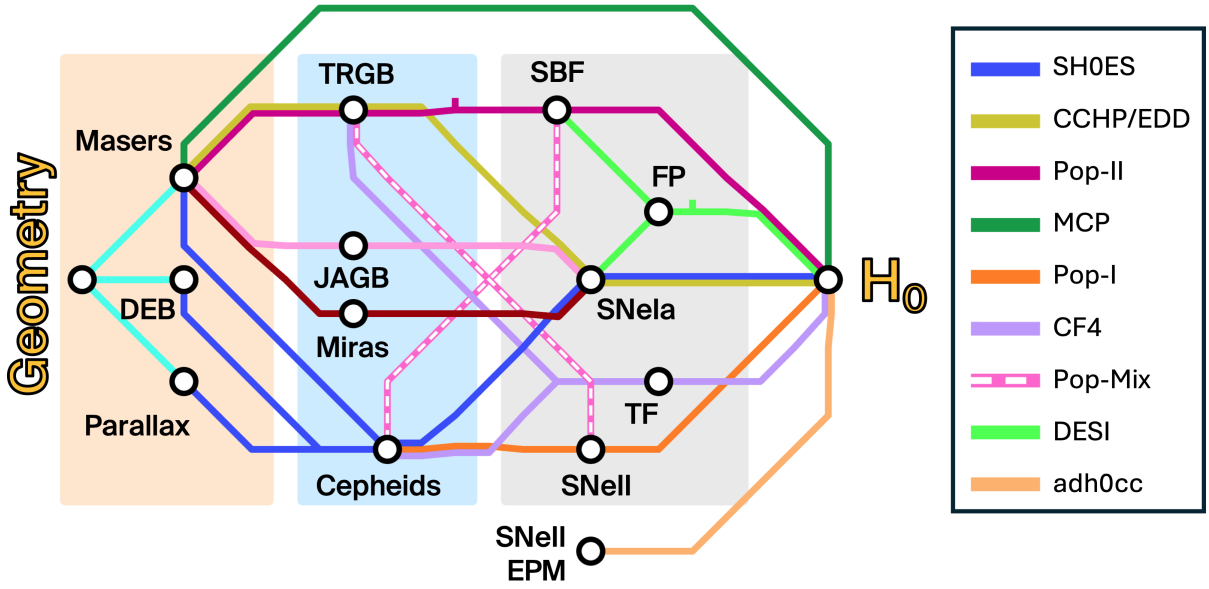


Figure 1.7: The classical distance scale as summarised by H0DN Collaboration et al. (2025).

As a first step, geometric distance indicators are used as they do not rely on other calibrations. Parallax measurements are applicable in the Milky Way, whereas eclipsing binaries are used up to the Magellanic Clouds. An additional geometric measurement through the water megamaser in NGC4258 is used. In these environments, Cepheids and Miras are observed. Geometric distances in the Milky Way and of these galaxies allow to calibrate PLRs of bright pulsating stars. Also, TRGB and JAGB distance indicators are calibrated in the Magellanic Clouds and NGC4258.

In a second step, PLRs, TRGB and JAGB methods are applied to nearby galaxies to determine their distances. This is possible thanks to photometric observations by the Hubble Space Telescope (HST) and the James Webb Space Telescope (JWST). Among the galaxies for which these methods can be applied, several hosted one or more SNIa. This enables to calibrate the absolute luminosity of SNIa, which is a secondary distance indicator.

Finally, distances to distant galaxies are measured for host of SNIa. This includes galaxies in the so-called Hubble flow, far enough so the recession velocity is dominated by the Universe expansion. For these galaxies, the redshift can be approximated to the cosmological redshift and $v = H_0 d$, as peculiar velocities become negligible compared to the Hubble flow. In practice, however, modern analyses still correct for these velocities to ensure a homogeneous treatment and to minimize residual systematic effects. The final product of the distance ladder is then an estimation of the Universe expansion rate, H_0. This determination is essential for measuring distances in a larger scale, as it intervenes in the relation between the redshift and the distance. Other paths include alternative distance measurements, such as other types of supernovae or the surface brightness fluctuations method. The SH0ES (Supernovae H_0 for the Equation of State) team provides the most recent complete measurement of H_0 using Cepheids and SNIa, in Breuval et al. (2024), using 42 SNIa for the calibration. In parallel, the CCHP (Chicago-Carnegie Hubble Program) team recently updated their determination using TRGB and JAGB with the JWST in Freedman et al. (2025).

4.2 A bit of cosmology

The expansion of the Universe was discovered in the early 20th century, both by observational and theoretical work. Its main consequence is that the distance between two galaxies is increasing due to cosmological expansion. This was observed using the Doppler shift of spectral lines in distant galaxies. It introduces the redshift measure, which quantifies how far a galaxy is. This observational view is currently modelled through the ΛCDM (Λ Cold Dark Matter) parametrisation of the Universe. The theoretical hypothesis supposes the origin of the Universe at the time of the Big Bang, which is the limit of validity of physical models. This phase is followed by the inflation, a rapid increase of the size of the Universe, followed by the current expansion. The mathematical behaviour of the Universe is described by the equations from general relativity, as follows, with the expansion being described by the scale factor $a(t)$.

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}G\rho - k\frac{c^2}{a^2} + \frac{\Lambda c^2}{3}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3P}{c^2}\right) + \frac{\Lambda c^2}{3}$$

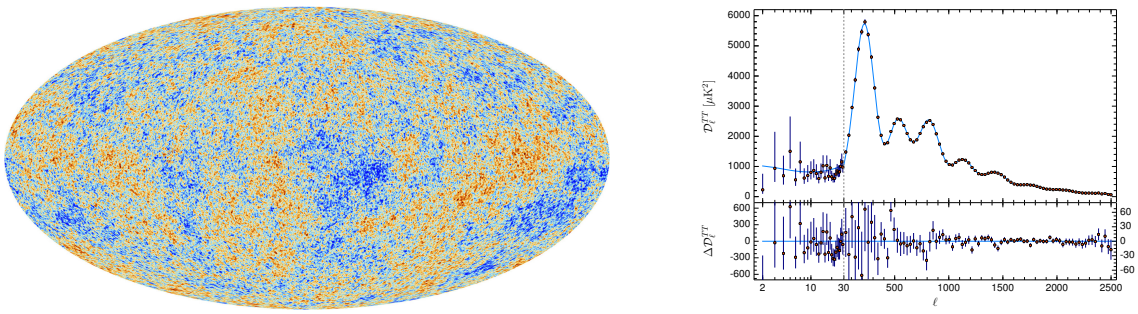
In these equations $\rho(t)$ is the density, $P(t)$ the pressure, k the curvature, G the gravitational constant, c the speed of light and Λ the cosmological constant which is used to describe the dark energy. This component was (re)-introduced after the discovery of the accelerated expansion of

the Universe using distant SNIa host galaxies. The different components are usually isolated in the equations using the density parameters $\Omega_i = \frac{\rho_i}{\rho_c}$, with $\rho_c = \frac{3H^2}{8\pi G}$ and $H(t) = \frac{\dot{a}}{a}$. Introducing a density parameter for matter, radiation, dark energy and curvature, respectively Ω_m, Ω_r, Ω_Λ and Ω_k, the first equation can be written as

$$\Omega_m + \Omega_r + \Omega_\Lambda + \Omega_k = 1$$

These density parameters are the one that are determined from the observation of the early Universe. The observation on which this determination is based is on the Cosmological Microwave Background (CMB). This black body emission at around 3K is caused by the cooling of the energy released when the Universe became transparent to radiation, approximately 380000 years after the Big Bang. This is the emission the closest to the ideal black body we observe in nature, as the temperature is particularly homogeneous in all directions. However, the observation of temperature inhomogeneities gives information on the structure of the Universe, and then on the cosmological parameters.

The CMB has been observed by several missions, which detected the microwave emission at 3K. The most recent one, the Planck satellite, which collected data between 2009 and 2013, is the heritage of the previous WMAP (Wilkinson Microwave Anisotropy Probe) and COBE (COsmic Background Explorer) missions. It mapped the temperature fluctuations around the sky, with very high sensitivity. These observations resulted in the map in Fig. 1.8. The power spectrum of temperature inhomogeneities is computed, and is fitted to the physical model which depends on the cosmological parameters including the density parameters and the Hubble constant $H(t = 0) = H_0$.



(a) (b)

Figure 1.8: (a) Temperature fluctuations in the CMB as observed by Planck. Credits: European Space Agency (ESA) and the Planck Collaboration. (b) Power spectrum and cosmological parameters fitting (Planck Collaboration et al. 2020)

4.3 The Hubble tension

The different studies evoked in the last paragraphs seem to show a very strong physical understanding of the Universe, with both a measurement of the expansion rate in the local Universe, on galaxies observations and of the fossil emission of the Universe at the epoch of

the recombination. However, the scientific community is currently divided because of significant differences in the results from the two methods. Most local measurements of H_0, whether based on Cepheids and SNIa or other distance indicators, agree on $H_0 = 73.50 \pm 0.81 \mathrm{km{\cdot}s^{-1}Mpc^{-1}}$ (H0DN Collaboration et al. 2025, for the baseline result). This very precise measurement is considering the standard distance scale from parallaxes, eclipsing binaries, maser distances, to SNIa, through Cepheids, Miras, TRGB and JAGB distances. Little variations are found when adding or removing peculiar methods, but they are all consistent with this standard result. In the opposite direction, recent results by the Planck collaboration (Planck Collaboration et al. 2020, Fig. 1.8) concluded on $H_0 = 67.4 \pm 0.5 \mathrm{km{\cdot}s^{-1}Mpc^{-1}}$, confirmed by independent surveys on the early-time Universe. These results are in a 7.5σ tension, which means that it is highly unlikely that both of them are compatible. A lot of efforts are put on both sides of the H_0 determination to find the origin of such a difference. Local measurements are generally limited by several step calibrations, which might introduce systematic uncertainties. They rely totally on empirical measurements. In addition, the distance indicators may be subject to specific systematics, including the reddening or the crowding effect for Cepheids, for example. Also, the sample of SNIa host galaxies is limited, and there might be a selection effect. For distant SNIa host galaxies, their systemic velocity is supposed to be negligible to attribute their velocity to cosmological expansion. On the other hand, predictions of H_0 from the early Universe are strongly model-dependent, as they all rely on the ΛCDM model.

The consequences of this tension are strong for cosmology. It indicates that the ΛCDM model might not be accurately describing our Universe. Several hypotheses to go beyond the standard model have been suggested and are currently investigated. Some of them include modifications in the physics describing the primordial Universe, other on the physics after the recombination. The nature of dark matter is also an object of debate, and even the general relativity theory is questioned. Measurements on the late Universe are becoming more precise and new methodologies appear (JAGB, Surface Brightness Fluctuations, gravitational wave standard sirens, masers...). They are used independently of the other methods to strengthen the value of H_0 from the local Universe.

Chapter 2

The projection factor of pulsating stars

1 The SPIPS parametrisation

In this chapter, I present my work on the modelling of RR Lyrae stars, carried out mostly during the first year of the PhD. For this, I used a computational tool called SPIPS for Spectro-Photo-Interferometry of Pulsating Stars, developed by Antoine Mérand (Mérand et al. 2015). This section describes the parametrisation used in this code and how it applies to pulsating star modelling.

1.1 An implementation of the parallax-of-pulsation technique

We have seen in Chapter 1 that the parallax-of-pulsation technique may be applied using Surface-Brightness-Colour Relations (SBCRs) to determine the angular diameter of stars from photometry in two different photometric bands. This requires to have a calibrated SBCR in two filters with good phase-coverage observations. Uncertainties on the predicted angular diameter are smaller when considering photometric bands at well separated wavelengths. Traditionally, the preferred SBCR is a combination of V and K photometric bands. However, it is possible to use various other combinations, which lead to different applications of the parallax-of-pulsation technique and affects the empirical calibration of the p-factor (Nardetto et al. 2023). One of the main motivations for the SPIPS tool is to use all photometric bands for which observations are available to improve the convergence to one consistent solution compatible with all the data that we have on a given star.

In addition, the angular diameter of nearby Cepheids is measurable by optical interferometry. As the number of points in the angular diameter curve may be limited, it is not always possible to perform a direct parallax-of-pulsation determination, but those points can improve or validate a solution. The SPIPS algorithm was designed to include interferometric measurements and compare with the diameter curve deduced from photometric and Radial Velocity (RV) measurements.

As it is restrictive to calibrate a SBCR for each pair of photometric bands, the solution adopted by SPIPS is to make use of atmospheric models to predict the surface brightness, from

the physical parameters of the Cepheid at each pulsation phase and convert it into magnitudes using the star's predicted spectrum and the transmission curves of the photometric passbands, retrieved from the Spanish Virtual Observatory.

The SPIPS code has been used mainly for Cepheids up to now, in the Milky Way and in the Magellanic Clouds, with or without interferometric measurements (Breitfelder et al. 2016; Gallenne et al. 2017, 2021; Kervella et al. 2017; Trahin et al. 2021).

1.2 The full parametrisation

The principle of the SPIPS algorithm is to optimise a set of stellar parameters, initialised in a dictionary. It is possible to select the fitted or fixed parameters.

Firstly, some parameters constrain the physical size of the star. It is done through the average angular diameter, the distance and the p-factor. Because of the degeneracy in the fit between the p-factor and the distance, either the p-factor should be fixed to apply the parallax-of-pulsation method or the distance should be fixed to calibrate the p-factor. Then, information on the period parameters contains at least the period and a reference time. Additional parameters are also possible to account for a changing period. The reference time for the zero phase is defined as the time of the maximum bolometric luminosity. The colour excess is given to correct magnitudes from the dust extinction. It should be noted that SPIPS adopts a specific definition of the reddening. The reddening from interstellar dust is taken into account as a corrective term in the apparent magnitude that depends on the wavelength, shorter wavelengths being more extinguished than longer wavelengths. This is parametrised by reddening laws $A(\lambda)/A_V$ where the extinction is normalised by the extinction in the V band, e.g. Fitzpatrick (1999). An object is then characterised by its colour excess $E(B-V) = A_B - A_V$. For the extinction over a photometric passband with a transmission curve $T(\lambda)$, the correction to the apparent magnitude will be the integral of the extinction parameter over the filter weighted by the intensity of the source $I(\lambda)$.

$$A = \int T(\lambda) I(\lambda) A(\lambda) d\lambda$$

It is usual to consider the standard $I(\lambda)$ spectrum of a Vega type star to determine the extinction parameter A_λ over a photometric passband. However, with SPIPS, we consider stars cooler than the 10000K Vega standard. This effect is taken into account in the SPIPS algorithm, with specific extinction parameters for the different photometric bands appropriated to the temperature of the pulsating star at each pulsation phase. This results in colour excesses $E(B-V)$ that are different from what is measured using the Vega standard in the literature. As a result, it is not possible to compare directly colour excesses from SPIPS with the literature, the main difference coming from the temperature of the used spectrum. Assuming the simplification that the SPIPS reddening is computed with a spectrum at 5500K, we can compute the ratio in reddening that this implies:

$$\frac{A_B(10000K) - A_V(10000K)}{A_B(5500K) - A_V(5500K)} = 1.16$$

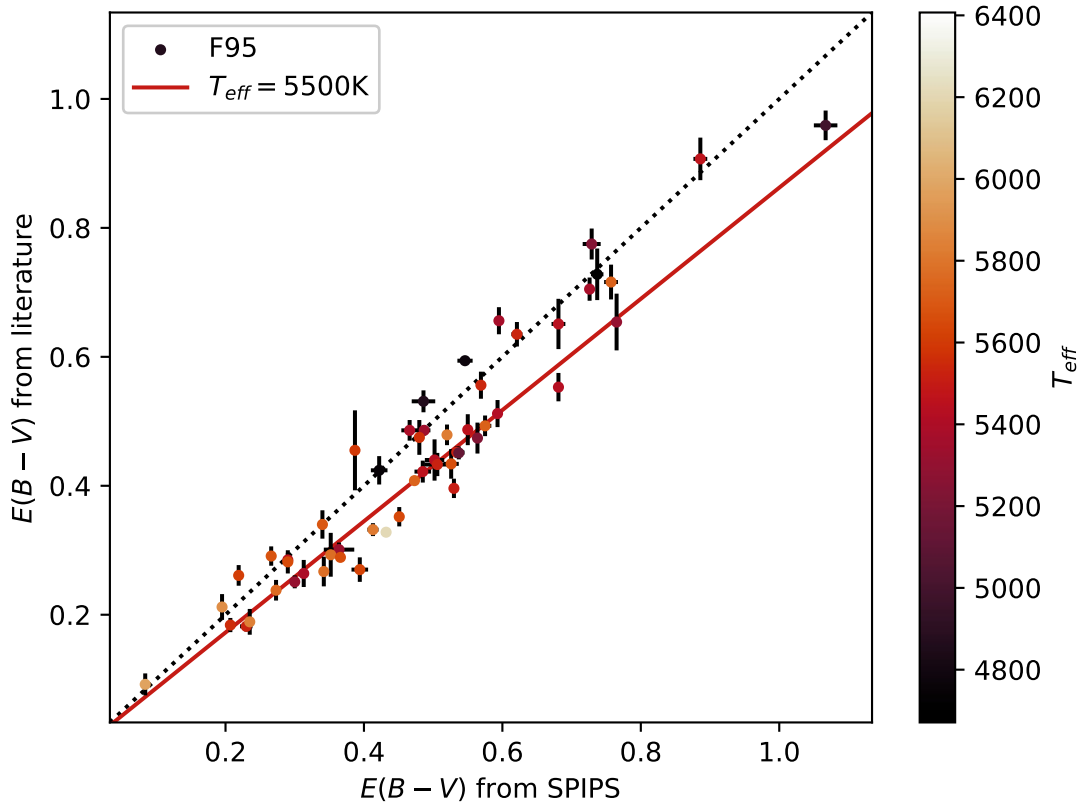


Figure 2.1: Comparison from colour excesses in the literature (Fernie et al. 1995) and SPIPS colour excess parameter, colour-coded by the mean temperature of the Cepheid.

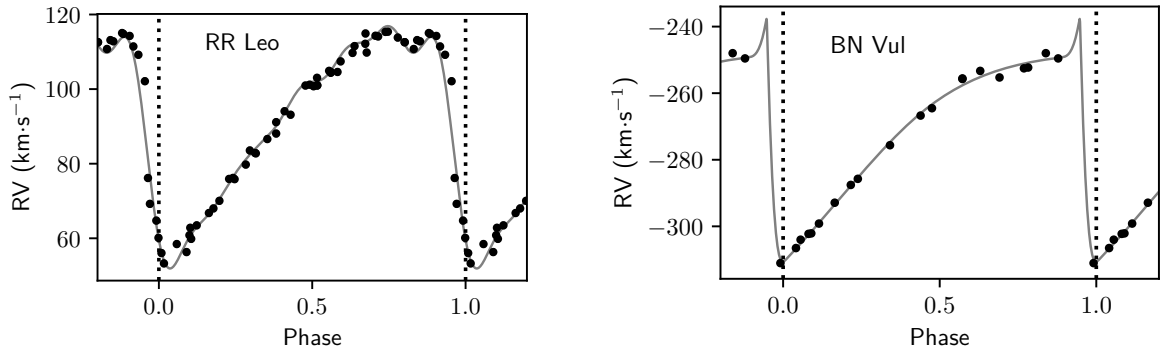
Then, in order to correct magnitudes, SPIPS will produce colour excess values that are 1.16 times higher than with the Vega based standard method. This is shown in Fig. 2.1.

In addition, SPIPS models the presence of a CircumStellar Envelope (CSE) around a Cepheid with a power law spectral contribution. As CSEs have not been observed yet for RR Lyrae stars, I do not detail this parametrisation here.

The fitted parameters also include a set of coefficients representing the RV curve (or alternatively the pulsation velocity curve). The interpolation can either use a Fourier decomposition or Splines functions. The advantage of the Fourier decomposition is that it has already proven good convergence for well sampled RV curves and it gives a smooth interpolation. However, it can introduce non-physical oscillations between data points. These oscillations can be avoided using Splines, but they might in turn introduce spikes for specific phases. These two behaviours are illustrated in Fig. 2.2.

In both cases, the number of parameters is set by the user, which offers the possibility to try different number of harmonics for instance for the Fourier interpolation.

In general, the Fourier interpolation is preferred as it is commonly used in the literature and



(a) (b)

Figure 2.2: (a) Example of unphysical oscillations in the interpolated RV curve from Fourier decomposition. Data from SOPHIE observations and Liu et al. (1989). (b) Example of an unphysical spike in the interpolated RV curve from Splines functions. Data from SOPHIE observations.

gives smooth variations of the curve. However, in case of limited phase-coverage, Spline functions can be considered, although they are less stable. The standard number of Fourier orders to include is typically between 6 and 8. In Fourier interpolation, when unphysical oscillations are observed, adding or removing one harmonic can help the solution to converge.

The interpolation of the temperature curve works the same way as for the RV but the fitted parameters are fully independent. It is possible to interpolate velocities with Splines and temperatures with Fourier series. As effective temperature curves present softer variations than RV curves, a Fourier interpolation is almost always a good choice.

1.3 The various observational datasets

The model fitting is conducted on the full set of data from various types of observations.

From spectroscopy, RV measurements are included, with their error bars. They can directly be compared with the model for the RV curve. They are critical for the model to converge and hence need good phase coverage.

Spectroscopic temperature measurements also constrain the model's temperature curve. Although they are not necessary for the model to converge, they are helpful to break the degeneracy between the temperature and the colour excess. Otherwise, it is recommended to fix the colour excess to independent determinations. Spectroscopic temperatures are in general obtained considering the depths of selected spectral lines in their spectra. On the other hand, these observations are sensitive to the metallicity and chemical composition of the stars. The method of line-depth ratios for temperature determination relies on the depth ratio of pairs of lines of the same chemical element to alleviate the dependence on chemical abundances. In particular, H lines are strongly sensitive to the temperature. However, in Cepheids, their profiles

are distorted due to the pulsation and the presence of H envelopes. Lists of temperatures from line depth ratios were published, notably by Kovtyukh et al. (2000) or more recently Lemasle et al. (2020). The uncertainties of this method are on the order of 100K.

Interferometric measurements are of particular interest as they link velocities with the photometry, but they are not available for most pulsating stars that are angularly too small, which is the case for RR Lyrae stars. When the star is well resolved, the angular diameter variations are compared with the integrated RV, corrected from the distance and the p-factor.

The bolometric luminosity curve is computed from the Rosseland radius and temperature curve using the relation

$$\frac{L}{L_\odot} = \left(\frac{R}{R_\odot}\right)^2 \left(\frac{T_{\rm eff}}{T_\odot}\right)^4$$

in which the temperature curve is directly obtained from the parametrisation and the radius curve is computed from the integrated RV curve and the distance, the p-factor, and the mean angular diameter. In addition, the Spectral Energy Distribution (SED) is predicted from ATLAS atmospheric models (Castelli et al. 2003), given the temperature, the size, and the surface gravity of the star at each pulsation phase. The mass to constrain the surface gravity is predicted from period-mass relations, for example from Bono et al. (2001).

For the model fitting, apparent magnitude curves in various photometric bands are produced by SPIPS. For each of them, the modelled light curve is computed using the SED and the transmission curve of the filter, and is extinguished from the distance and the interstellar medium using the reddening law from Fitzpatrick (1999) and $R_V = 3.1$.

It is interesting to note that the temperature variations dominate the photometry in the shorter wavelengths, while the contribution from the radius variation is dominant for infrared (IR) photometry.

An example of a panel from the SPIPS code is given in Fig. 2.3 in the case of the prototype Cepheid δ Cep. This figure shows the different curves from the SPIPS model. On the top left plot, the RV is shown with a Spline model in this case. Below is the effective temperature curve and on the bottom left panel interferometric measurements of the angular diameter are plotted together with the Rosseland diameter model. On the right part, each subplot corresponds to a light curve in one photometric band. The bands are sorted with the bluer ones on the top and the redder on the bottom. In addition, the photometric colour B-V is plotted at the end. In general, straight green lines represent the modelled curves that are directly fitted and the grey straight lines are computed from the other parameters. General information is given in the title.

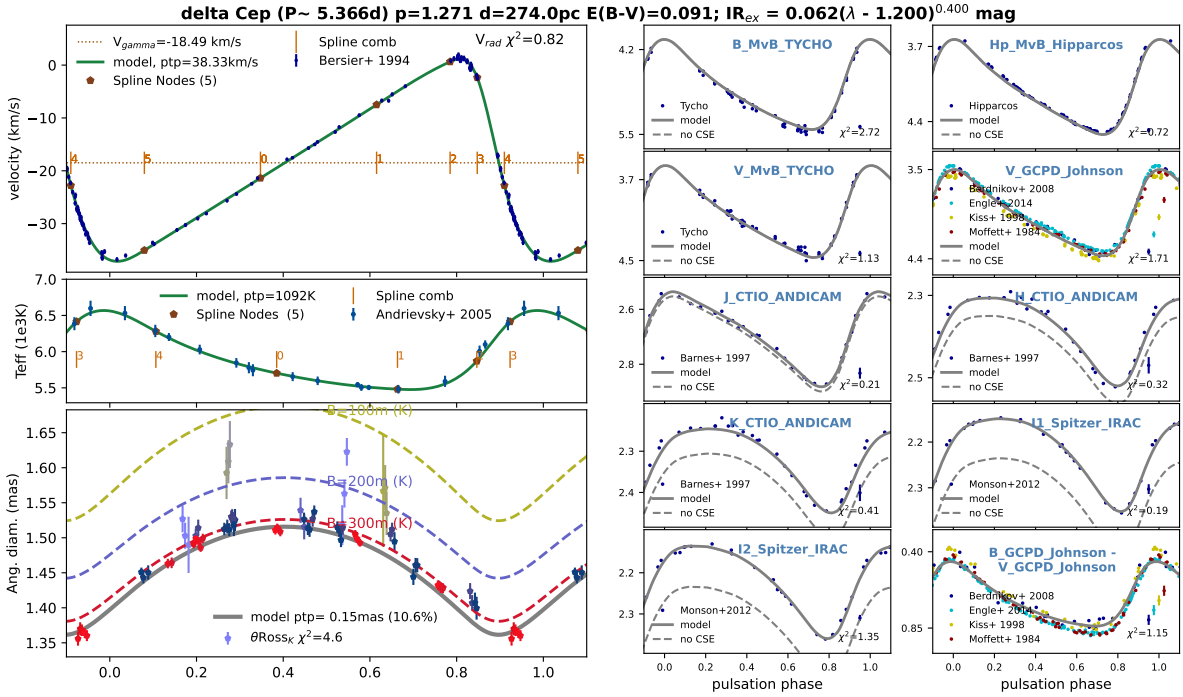


Figure 2.3: SPIPS model of δ Cep from Mérand et al. (2015).

2 Data on Galactic RR Lyrae stars

My project was focused on the SPIPS modelling of RR Lyrae stars. While the SPIPS code has already been extensively used to model Cepheids (Trahin et al. 2021), it can be applied to other pulsators like RR Lyrae. Boris Trahin, during his PhD, already tested the modelling of RR Lyrae stars, using archival data, in an unpublished work. I started from this preliminary work and extended it with new data. In this section, I describe the origin of the data used in the models I obtained.

2.1 Radial Velocity data

In order to conduct an analysis on the p-factor dispersions, it is important not to bias the determination by heterogeneous datasets and heterogeneous methods. This supports the need for homogeneous data for RVs for which the calibration is often not as precise as for photometric data.

We focused on the selection of 142 stars in the catalogue by Feast et al. (2008) and selected stars that are well covered by photometric studies. Stars showing evidence for cycle-to-cycle modulations, which is known as the Blazhko effect (Blažko 1907) were avoided. I acquired new spectroscopic data for a selection of 26 RR Lyrae in the Northern Hemisphere. The data were collected at Observatoire de Haute-Provence, using the SOPHIE spectrograph (Spectrographe pour l'Observation des PHénomènes des Intérieurs stellaires et des Exoplanètes, Perruchot et al. 2008) mounted on the 193 cm telescope. Typically used for exoplanet observations (Bouchy

et al. 2009; Santerne et al. 2016), the great stability of this instrument and its high spectral resolution enable us to determine very precise RV using short exposures, which is important as RR Lyrae stars are pulsating rapidly. The observations were divided in four runs to cover the range of right ascensions of the selected stars. Two runs were observed before the beginning of my PhD in preparation of the project, but I had the opportunity to observe on my own for the two last runs. To gain efficiency, I produced a tool to help me during the selection of the targets. We obtained series of 10 to 15 minutes exposures and alternated targets in order to properly cover the RV curve. In particular, we tried to do regular observations during the rising branch and increase the frequency while the star was at its rebound phases, as it is when the RV evolves very quickly. For example, according to Fig. 2.4, the night of the PhD defence at 1:30AM (UT time), it is interesting to observe the star VY Ser as it is high in the sky, and we lack some data at the corresponding phase, but we cannot observe AV Peg which is below the horizon. We would in any case not need additional RV for AV Peg at that particular phase as we already observed this phase.

The reduction and calibration were done using the reduction pipeline of the SOPHIE instrument. Initially, I took the values for RV obtained by the reduction as a first approximation. The determination was done from a Gaussian fit of the Cross-Correlation Function (CCF). However, as the instrument was tailored for exoplanet detection, the amplitude of RV around the systemic velocity was larger than what was expected by default by the reduction software. This led to the peak of the CCF to appear outside the velocity window where it was computed, and motivated me to produce my own program to derive RV. In addition to the SOPHIE reduction software, I hence familiarised with two different codes: iSpec (Blanco-Cuaresma et al. 2014; Blanco-Cuaresma 2019) and RaveSpan (Pilecki et al. 2017). I found iSpec easy to implement automatically for all my spectra and to adapt to my RR Lyrae project.

Unfortunately, quite a lot of the observing time was lost due to poor weather conditions leading us to revise the final sample for the modelling. I selected the stars from the phase-coverage of the RV curve, both from a visual inspection and using various numerical parameters. We initially expected around 15 measurements on each target, which is a reasonable number to consider for well phase-covered curves. In addition, we considered the number of measurements during the rebound of the stellar atmosphere around the minimum radius. Lacking those critical phases makes RV curve unusable, as they contain the information on the amplitude of the RV curve and hence on the p-factor. We looked for RV curves with more than 4 measurements in the 20% of the pulsation cycle around the rebound. Finally, we limited the phase gap between two measurements to 0.2 to avoid biasing the shape of the curve during the interpolation. We found 11 stars to satisfy and 6 others to be near these requirements. For some stars, we complemented the RV curve using archival data, offset to match our new measurements. We checked that the RV curve shape was similar to avoid introducing some bias on the RV measurements on our conclusions. Finally, the stars in the preliminary analysis by Boris Trahin were also considered, even without new RV measurements from SOPHIE to complement the statistical analysis.

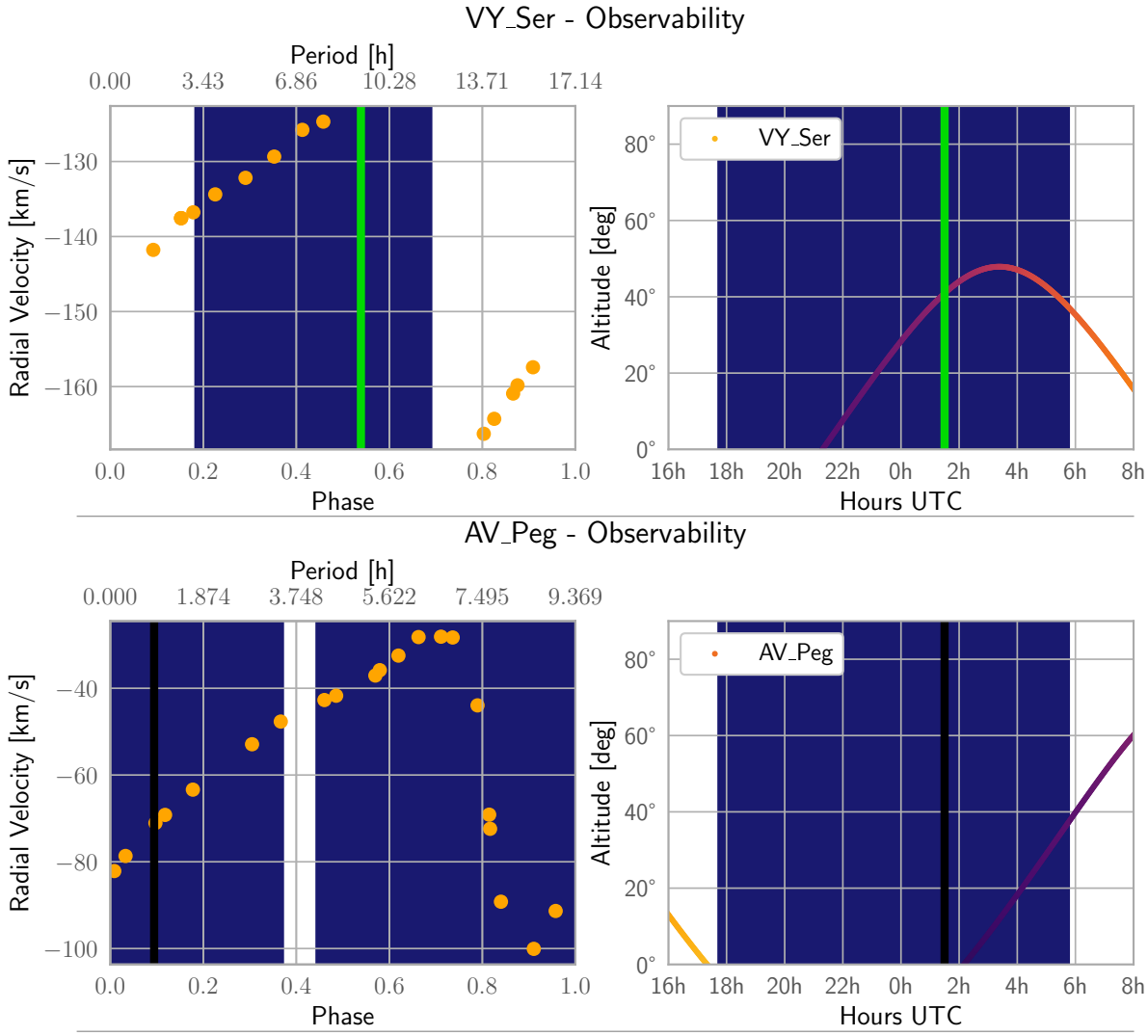


Figure 2.4: Observability of two RR Lyrae stars in our sample for the night of March 19th, 2026 at 1:30AM

2.2 Photometric data

Our final sample also depends on the available photometry, which is taken from various sources in the literature. In order to obtain reliable pulsation models, it is important to incorporate photometry in a large variety of photometric bands. However, we favour some filters that are well calibrated.

It has been observed that atmospheric models used in SPIPS failed to reproduce the photometry in the bluer parts of the spectrum. Indeed, in Casagrande et al. (2014), they show using MARCS models (Model Atmospheres with a Radiative and Convective Scheme) that the microturbulence has a role in the synthetic flux computed in blue and Ultra-Violet (UV) bands, especially for cool stars. This also applies for other atmosphere models such as Castelli et al. (2003), as the treatment of microturbulence and the values adopted are similar. These effects

propagate in SPIPS models when including shorter wavelengths; the shape is well reproduced, but there is an offset in the mean magnitude. This was observed for photometric bands bluer than the B Johnson band. For this reason, observations in those bands are not included in the fits.

Large photometric surveys are particularly valuable for such models as it is easy to include them for various sources, and we usually know the general caveats of their photometric data. For pulsating stars, the Gaia mission offers an opportunity to obtain very precise and well sampled light curves in three photometric bands. Observing the sky during a decade, this satellite launched by the European Space Agency (ESA) aimed at providing astrometric solutions for more than one billion of objects in the Milky Way and the Local Group, up to magnitude 20. Using spectrophotometry and a spectrometer on board, it also obtained photometry and spectroscopy for all these sources on top of their position and motion in the sky. Moreover, thanks to the multiple observations of each region, time series of all these observables were acquired. The publication of the data is still in process with various Data Releases (DRs). My PhD began just after the DR3, officially published on 13 June 2022. Specific to our study, this DR provided epoch photometry for 10.5 million sources, including many pulsating stars. For the SPIPS modelling, I retrieved the Gaia photometry in the three filters G, Bp and Rp for all the RR Lyrae stars in my sample. Unfortunately, the Bp band lies in the blue part of the electromagnetic spectrum and happen to fall in the domain where the atmospheric models are failing to predict the SED. In addition, the G band is in fact a combination of the two other filters, making it a broad band difficult to handle. The integration of the SED, leading to magnitudes, is tricky for this kind of broad filter as a small error will lead to big differences in the photometry. Hence, including this band sometimes yields some convergence problems of the fit. As Gaia photometry is an asset for our modelling, I still tried to include it as much as possible. However, when it was locking the fitting process in an invalid solution, I chose to remove it in order to help the model to converge. In the last DR, XP spectra have been published. These spectra are easy to integrate over photometric passbands to convert into photometry in the BP to RP domain. This could provide additional photometry in the future. However, only mean XP spectra are now available in the DR3. Moreover, the domain covered by these observations is generally well observed by other surveys, so the addition of this photometry should not change drastically the amount of available data.

The Gaia photometry was supplemented by various epoch photometry from the literature. In the following table, I list all the references I used, together with the photometric bands and the concerned stars. Rather than including multiple references for the same star in the same photometric region, it was more interesting to find one reference that covers the whole pulsation cycle with a low dispersion light curve. In general, I followed the criteria by order of priority that (i) they cover an interesting wavelength domain, (ii) dispersion and uncertainties are low, (iii) the phase coverage is satisfying, and (iv) data are relatively recent.

Reference	Filters visible	Filters IR	Stars
Barnes et al. (1988)	B, **V**, R, **I**		**DX Del**, TT Lyn, **TU UMa**, UU Vir
Barnes et al. (1992)	B, **V**, R, **I**	J, H, K	TT Lyn, **TU UMa**, UU Vir
Cacciari et al. (1987)	U, B, **V**, **R**, **I**		**SW Dra**
Carney et al. (1984)	u, v, b, y, V		VY Ser
Fernley et al. (1989)	u, v, b, B, V, R, I	**J**, H, **K**	**X Ari**
Fernley et al. (1990)	B, **V**, R, I	**J**, **H**, **K**	**SS Leo**, **VY Ser**
Fernley et al. (1993)		**J**, H, **K**	W CVn, **UY Cyg**, SV Eri, VX Her, WZ Hya, BN Vul
Jones et al. (1987a)	**V**, R	**J**, **H**, **K**	**X Ari**
Jones et al. (1987b)	B, **V**	**J**, **H**, **K**	**SW Dra**
Jones et al. (1988)	B, **V**	**K**	**RS Boo**, **VY Ser**, UU Vir
Layden et al. (2019)		**K**	W CVn, **UY Cyg**, VX Her, AB UMa
Liu et al. (1989)	U, B, **V**, **R**, **I**	**J**, **K**	**RR Cet**, **SU Dra**, **RX Eri**, RR Leo, **TT Lyn**, **AV Peg**, **TU UMa**, UU Vir
Meylan et al. (1986)	U, B_1, B, B_2, V_1, V, G		RR Cet, DX Del
Monson et al. (2017)	B, **V**, **I**	**I1**, **I2**	**X Ari**, **BB Pup**, **RR Cet**, **UY Cyg**, **DX Del**, **SU Dra**, **SW Dra**, **RX Eri**, SV Eri, VX Her, RR Leo, **TT Lyn**, **AV Peg**, **AB UMa**, **TU UMa**, UU Vir
Paczyński (1965)	U, B, V		AV Peg
Paczyński (1966)	U, B, V		DX Del
Preston et al. (1964)	U, B, V		X Ari, RR Cet, RX Eri, SV Eri, TU UMa
Skillen et al. (1989)	**V**	J, H, **K**	**DX Del**
Sturch (1966)	UBV		V0341 Aql, X Ari, W CVn, UY Cyg, RX Eri, SV Eri, VX Her, RR Leo, SS Leo, TT Lyn, AV Peg, VY Ser, TU UMa, UU Vir, BN Vul
Varsavsky (1960)	U, B, **V**		**VY Ser**

Table 2.1: Collection of photometric data from the literature. In bold the data used in the SPIPS modelling.

Finally, due to sensitivity problems including calibration issues and low signal-to-noise ratio

(S/N), I removed some filters from the fitting process. It is the case for WISE (Wide-field Infrared Survey Explorer) or R and I photometry.

The selected RR Lyrae stars are angularly too small to obtain interferometric measurements of their angular diameter, hence, all the models rely on the combination of photometry in the different photometric bands. In addition, temperature determinations during the pulsation cycle are hard to find, and many of the stars in my sample do not have any. Indeed, it requires high S/N spectra and the variability of the temperature prevents us from making longer exposures. From the SOPHIE spectra I acquired, it was not possible to derive precise temperature estimates as the S/N was around 20 to 30. Some few temperature determinations from Andrievsky et al. (2018) and Andrievsky et al. (2020, 2021) are added to the input data.

2.3 Input parameters

In addition to the observational data to fit the model, we initialise the fit with some specific parameters.

Period: The initial period is from the Gaia catalogue. Because some observations span over more than 50 years, and the pulsation period is only half-a-day, the phasing is particularly sensitive to the period and to period changes. We focused on a good phasing of the fitted data by implementing period variations when needed to help the fit to converge. The obtained period and period change may not be representative of the actual period of the star, as it may have smoothed some period variations, for example. An appropriate solution would have been to use phasing solutions from O-C (Observed-Computed) diagrams, but this would not have improved the fit significantly.

Reddening: The colour excess was fixed to the values determined using Stilism (Capitanio et al. 2017). It is appropriate in this case, as our stars are nearby, and the 3D maps are accurate in this range. As we do not have temperature measurements for the majority of our stars, it is impossible to determine a photometric reddening from the difference in magnitude in the visible and in the IR. This prevents us from fitting the colour excess parameter, and it is important to propagate the uncertainty. As it is not possible to have an uncertainty on an input parameter, I evaluated the impact of the error on the reddening in the outputs of the model by fitting the data with the highest (resp. lowest) possible colour excess, on the basis of the best fit model.

Distance: In order to calibrate the p-factor, the distance of the star has to be fixed. The Gaia DR3 provides parallaxes for all RR Lyrae in the sample. The distance is simply the inverted parallax but when it comes to measured quantities, it is more complicated as we need to take into account the error bars. The Bailer-Jones et al. (2021) catalogue provides distances from Gaia parallaxes for which the distribution of probability of the measured quantity of the parallax is properly inverted, and including the zero-point offset of Lindegren et al. (2021). This has an impact for relatively distant objects, which is not the case for our sample of RR Lyrae stars, as shown in Fig. 2.5.

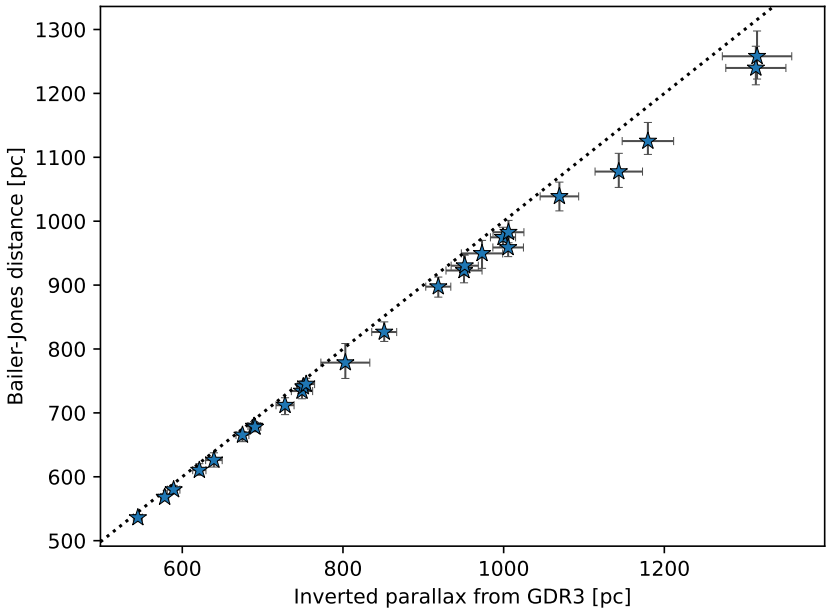


Figure 2.5: Distances of the RR Lyrae stars in my sample from inverted parallax and the Bailer-Jones catalogue.

Period-mass relations: The atmospheric models are selected from the temperature, the surface gravity and the metallicity. Surface gravity is known from the radius (direct product of the parametrisation) and the mass which is assumed. For RR Lyrae, we use the assumption of the mean mass in the sample from Marsakov et al. (2019) $M = 0.55 M_{\odot}$. For the low range of possible masses, the effect on the atmosphere models is negligible.

Metallicity: The metallicity is given to use the proper atmosphere model, but has a small impact on the fit. I used values tabulated in Monson et al. (2017) and evaluated the impact of a 0.15 dex error on the fit.

3 Results of the modelling

3.1 Model convergence

The convergence of the models was the most difficult part of that study. The SPIPS code uses a gradient descent to find the best fit model, and it requires some work to help the model converge towards the global minimum.

The gradient descent works by computing the χ^2 parameter, which quantifies the difference between the observational data and the model, normalised with the uncertainties. When using observational uncertainties, one point of RV has the same weight on the fit as one photometric point. However, as we combine different light curves in different bands, the photometric measurements are much more numerous than the RV ones and overall they have much more weight on the fit, leading to invalid solutions for RV curve to be found as the model will focus on

the agreement with the photometry. To tackle this issue, SPIPS uses normalised uncertainties on each group of data from the same technique. Observational uncertainties are then multiplied by a factor such that the groups of RV, temperature, angular diameter, and photometric measurements all have the same weight.

The new SOPHIE RV curves are generally composed of 20 to 30 measurements across the pulsation cycle. In the ideal case, it results in a mean spacing around 0.04 in phase. However, because of the difficulty to schedule the observations, gaps are present for some RV curves. A dense phase coverage of the RV curve is critical to avoid unphysical models. Even with the supplementation with data from the literature, it was not sufficient for example for the stars CG Peg or SV Eri.

Despite the normalisation by technique in the χ^2 function, a similar problem occurs in the set of photometric data. The light curves are gathered in various bands, with the shape being drastically different between visible and IR light curves. Moreover, IR photometry is generally sparser than visible photometry in the V band. It results that the IR photometry can be ignored if its weight happens to be too low compared to visible photometry due to fewer observations in that domain. It is thus important to balance the amount of observational data in the visible and in the IR, and not bias the fit with redundant photometry. The lack of IR photometry is the main reason that prevents the model from converging satisfactorily for the stars BN Vul, SW Aqr, V0341 Aql, VX Her and WZ Hya.

From the first tests I did on SW Dra, it turned out rapidly that a straight fit from scratch was particularly inefficient. Even when feeding the parameters with acceptable first guesses for the period and the diameter, the model would eventually converge to unacceptable solutions. In particular, the period determination is difficult for RR Lyrae stars. A very small change in the period could cause critical dephasing issues between data from the 80s and recent observations. It suggested that, when testing different parameters, variations on the period should be minimal compared to other parameters. This convergence issue was the subject of my visit to ESO (European Southern Observatory) in December 2022 to discuss with Antoine Mérand, developer of the SPIPS code. We ended up with an adapted incremental strategy to successively find the critical parameters and finish with fine-tuning of the parameters from a realistic set. This strategy follows the following steps.

- During the whole process, some parameters that cannot be determined independently are fixed: the distance (degenerate with the p-factor), the colour excess (degenerate with the temperature), and the metallicity. Finally, the reference time in the parameter input is fixed, modulo the period, to the mean time of the observations. The real reference time for the zero phase is computed by the model from the bolometric luminosity curve.

- First guess. The period is given from Gaia DR3, the p-factor is initialised at 1.27, close to the mean value for Cepheids. The first guess for RV and temperature is plotted in Fig. 2.6.

- Fit of the temperature curve. For pairs of photometric bands with simultaneous observa-

tions, we compute colour variations. Combined with available temperature observations, the temperature model is obtained. The only free parameters in this step are the parameters of the temperature curve. For the next steps, colours will not be used again.

- Fit of the RV curve. The RV parameters are fitted to the RV observations to establish a first guess of the RV model.
- Fit all data. All data are used to fit the other parameters. Models for RV and temperature are fixed. This step helps to refine the diameter and the p-factor.
- Fit everything except the period. All data are used to fit all the free parameters, except the period (and possible period changes). The values defined at the previous steps are adjusted considering the whole set of data.
- Finer fit with all data. This is the last step of the optimisation. Every free parameter is fitted, allowing smaller steps to avoid the period to vary too much. If needed, a change of period is fitted at this stage only.

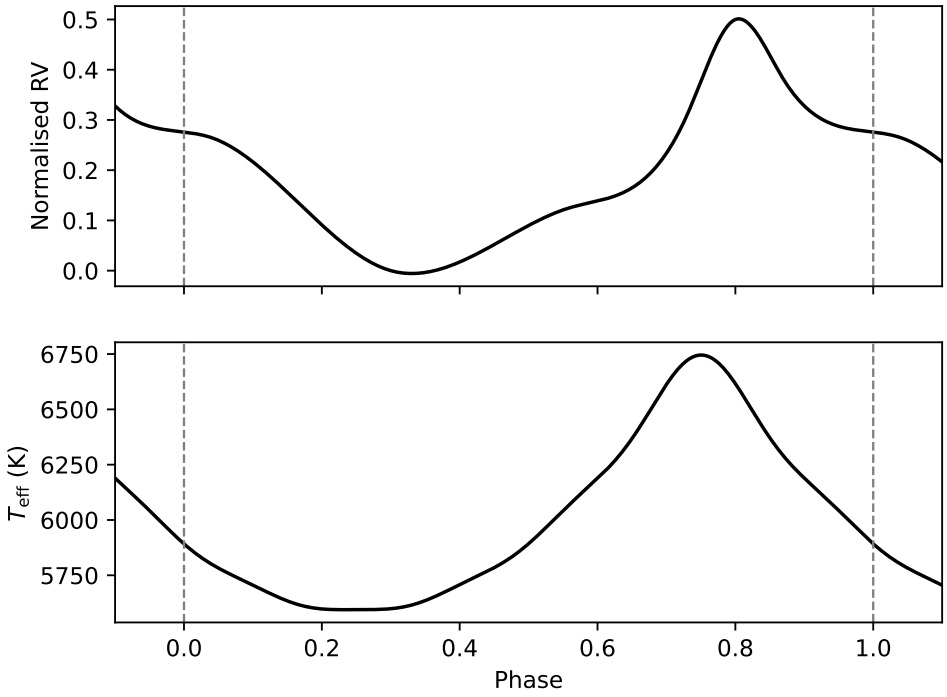


Figure 2.6: Starting solutions for RV (Splines) and temperature (Fourier). The RV solution is scaled according to the mean and amplitude of the observations.

Following the described strategy, I modelled the whole sample of RR Lyrae stars. I varied the number of Fourier or Splines parameters for both the temperature and the RV curve. Eventually, I obtained a model with a good visual agreement and a satisfying χ^2 value for 17 RR Lyrae stars, presented in the publication Bras et al. (2024), (see also Sect. 5). Unfortunately, I could not find a solution for the stars RR Leo, UU Vir, V0440 Sgr and W CVn, that do not fall in the

issues listed before. However, I am confident that these models would benefit from additional data.

3.2 Results on the p-factor of RR Lyrae stars

For the 17 RR Lyrae stars for which I obtained a satisfactory model, we have access to the p-factor, with its statistical error. In addition, it is important to take into account the uncertainties on the fixed parameters i.e. the distance, the metallicity, and the reddening. The uncertainty on the distance is simple to propagate as the model provides the value of d/p. The dependence on the metallicity and the colour excess is more difficult to predict. Therefore, when an appropriate best model was found, I changed the value of the metallicity (resp. the reddening) to the extrema of the error distribution, and determined the difference induced to the p-factor and to other parameters.

The detailed results of the models can be found in the publication linked to this chapter and I will only report here the essential conclusions.

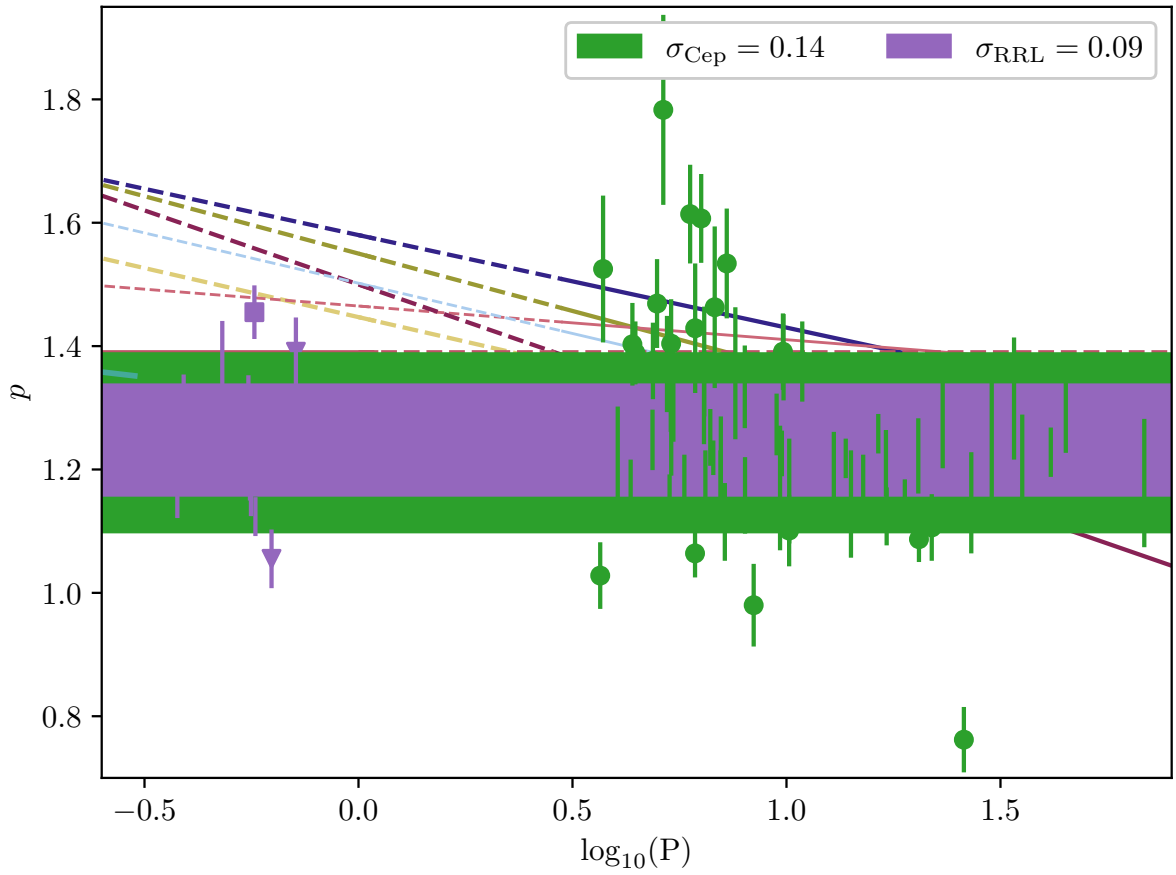


Figure 2.7: Period–p-factor relation for 17 RR Lyrae stars (in purple, Bras et al. 2024) and 63 Cepheids (in green, Trahin et al. 2021), with previous calibrations (see Fig. 1.6).

The values for the p-factor of the 17 RR Lyrae stars, plotted in Fig. 2.7, range from 1.06

for SS Leo to 1.46 for RV Oct with typical errors of about 0.04. I compared the results with the preliminary analysis of Boris Trahin for the stars BB Pup, RV Oct, SW Dra, WY Ant and X Ari, common to both samples. As the distance determination improved between our two analyses, the comparison should be made to the d/p quantity. In the meantime, I collaborated with Bartek Zgirski on his parallax-of-pulsation analysis using IR SBCRs for several southern RR Lyrae stars and independent data, including RX Eri, the only star I also analysed. In Zgirski et al. (2024), they report $p = 1.44 \pm 0.04$, compared to my value of $p = 1.25 \pm 0.02$. In order to quantify the effect of the methodology compared to the input dataset, I performed a SPIPS analysis based on the interpolation of the RV, V-band and K-band curves from his work for BB Eri, RX Eri and U Lep. Reversely, he tried his method from the modelled RV, V-band and K-band curves for AB UMa. From both sides, we ended up with a systematic difference of 0.05 using the SBCR from Kervella et al. (2004b), SPIPS p-factors being smaller.

Taking into account the sample of 17 RR Lyrae stars, the mean p-factor is 1.25. I investigated the dependence on the pulsation period and fitted the $p-P$ relation $p = -0.394_{\pm 0.298} \log_{10}(P) + 1.147_{\pm 0.078}$, which seems to indicate that there is not a clear correlation with the period. Adding the sample of 63 Cepheids from Trahin et al. (2021) since we used the same parallax-of-pulsation implementation, there is a good agreement between the p-factors of Cepheids and RR Lyrae stars, with a mean value of 1.25. This result is compatible with previous literature results, as gathered in Fig. 1.6.

For my sample of RR Lyrae stars, the Root Mean Square (RMS) scatter of the p-factor is 7%, compared to 12% for the 63 Cepheids. Even restricting the sample to the stars with homogeneous RV data, a large scatter is observed. Besides the systematic difference with the approach of Zgirski et al. (2024), a similar scatter was found on their sample of 6 RR Lyrae stars. In a similar study to Zgirski et al. (2024), Wielgórski et al. (2024) studied the p-factors of BL Her type II Cepheids and reported 4% scatter for a sample of 8 stars. On δ Scuti pulsators, Nardetto et al. (2014) reported a scatter of 2% between the four stars in their sample.

The various studies carried out to calibrate the p-factor of pulsating stars still lead to the same conclusion that the dispersion of the values is too high to use the parallax-of-pulsation technique for precise distance measurements.

3.3 The lack of correlation with other parameters

From the SPIPS models, several parameters are derived, allowing us to determine classical relations like PLRs and Period-Radius Relations (PRRs). The latter are particularly interesting in our case, due to the role of the radius curve in the parallax-of-pulsation technique. I derived the PR relation for my sample of RR Lyrae stars, which is particularly tight, and in agreement with the literature (see paper, figure 18). This suggests that the mean stellar radius is robustly determined in the modelling. The remaining dispersion observed in the p-factor therefore likely originates from uncertainties in the conversion of RVs into pulsation velocities, rather than from the radius determination itself.

In addition to the relation with the period, which did not show any clear tendency, we can try different correlation with other parameters. To investigate this aspect, I used the Spearman coefficient. This statistical tool indicates if two sets of data can be correlated monotonously. For both set of data X and Y, the rank R_X and R_Y are computed, describing the tendency of the variables. Then they are compared with the relation $r_S = \frac{Cov(R_x,R_y)}{\sigma_{R_x}\sigma_{R_y}}$. If the Spearman coefficient r_S is close to 1, it means that the two variables are correlated in a way that if X increases, Y increases too, including non-linear correlations. Respectively, if r_S is close to -1, an increase in X will induce a decrease in Y. When r_S is close to 0, the probability that the two variables are correlated in a monotonous way is low. From a theoretical viewpoint, the p-factor is expected to vary with the pulsation period, as longer-period stars have more extended atmospheres and are subject to stronger velocity gradients. Consequently, it may also depend on stellar parameters that correlate with the period, such as the effective temperature or surface gravity. In addition, a correlation with the pulsation amplitude, in particular the radial velocity amplitude, can be expected, as it directly impacts the atmospheric dynamics and the strength of velocity gradients. For this reason, I computed $r_S(p, T_{eff}) = -0.33$, $r_S(p, R) = -0.05$, $r_S(p, [Fe/H]) = 0.09$ and $r_S(p, A_{RV}) = 0.01$. The p-factor seems uncorrelated to the following quantities: the effective temperature, the radius, the metallicity and the amplitude of the RV curve.

A more complex dependence, for instance on a combination of two or more parameters, has not been explored.

4 A more complex atmospheric motion?

4.1 Discriminate layers in the atmosphere

As explained in the beginning of this chapter, I measured RV from SOPHIE spectra using the CCF method, detailed in the introduction. This technique is a way to stack the different spectral lines of the spectrum to have a mean profile of the lines and determine more easily the Doppler shift with this profile. It relies on a specific selection of lines, which I investigated.

The standard masks used for instance by the SOPHIE data reduction pipeline and for standard spectroscopic RV measurements depend on the spectral type of the star. Spectral lines are selected from the temperature of the star. An alternative approach uses synthetic spectrum at the temperature, surface gravity and metallicity of the star. I used F0 mask, corresponding to temperatures around 6000-7000K, and A0 mask to quantify the effect of this choice, and synthetic spectra with the same temperature, surface gravity of 3 to 3.5 and metallicity of -1 dex. This resulted in differences in RV of less than 2 km·s^{-1}. As a comparison, the typical error for RV measurements in Liu et al. (1989) is 1 km·s^{-1}.

In order to gain a better understanding on the role of the spectral lines on the RV measurements (van Hoof et al. 1953), I used a similar approach to the one by Borgniet et al. (2019).

Simon Borgniet was a former post-doc on the team, and I could have access to his codes. Adapting his code, I made my own specific masks tailored for RR Lyrae stars to disentangle the role of the different layers in the atmosphere. In this work, the spectral depth is used as a proxy for the optical depth of the lines, and we suppose that, generally, deep lines are forming in the higher atmosphere. From a synthetic PHOENIX spectrum (Husser et al. 2013), I extracted the unblended lines, and grouped them into weak, medium or deep lines according to their depth comparing to the limits I chose (respectively 0.2 to 0.45, 0.45 to 0.65 and 0.65 to 0.95).

The difference between RV measurements with the different masks is again of low amplitude, with variations of a few $\mathrm{km{\cdot}s^{-1}}$. The idea of that work was to investigate if different p-factors could be linked to the propagation of the pulsation towards the different layers of the atmosphere of the star. My intuition was that weak lines, that are forming near the photosphere should be more coupled to the motion of the photosphere and then to the radius variations coming from photometric measurements in the parallax-of-pulsation method (Nardetto et al. 2007). I therefore propagated the difference in RV measurements to the difference on the amplitude of the integrated RV curve, which is the main driver of the p-factor value, all else being equal. I concluded, based on two different RV curves for two different stars, that the effect is of the order of 1% on the amplitude of the RV curve and then on the p-factor, smaller than the typical error bars on the p-factor.

4.2 Role of the asymmetry of the spectral lines

A more important issue comes with the asymmetry of the spectral lines and hence of the CCF profile, which clearly impacts our measure of the RV. Asymmetry in line profiles of Cepheids have been observed first by Kraft (1956) on X Cyg and then on X Sgr and η Aql by Sasselov et al. (1989). The intrinsic asymmetry of spectral line profiles in radially pulsating stars arises naturally from the geometrical projection of the atmospheric velocity field onto the line of sight when integrating over the stellar disk.

The problem with the asymmetric CCF profiles is that it introduces a choice on where on this profile the Doppler shift should be computed with respect to the rest frame to derive the RV. Usually, the CCF profile is fitted with a Gaussian function, which, in the case of asymmetric profile will move the mean wavelength to the extended wing. To avoid this effect, we can use a bi-Gaussian profile defined as

$$\mathscr{B}(x,\mu,\sigma_1,\sigma_2) = \begin{cases} \mathscr{G}(x,\mu,\sigma_1) & \text{if } x < \mu, \\ \mathscr{G}(x,\mu,\sigma_2) & \text{if } x \geq \mu \end{cases}$$

where $\mathscr{G}$ is the Gaussian profile, σ_1 and σ_2 are the two standard deviations of the two Gaussian wings, and μ is the wavelength at which the Doppler shift is computed for the RV measurement. This profile has the advantage to fit well the majority of CCFs, but it traces mostly the peak of the CCF. Additionally, I computed the barycentre of the profile, also named the centroid method. The centroid velocity corresponds to the intensity-weighted mean velocity across the

spectral line profile, making it particularly suitable for pulsating stars with asymmetric line profiles. This definition is the least sensitive to the rotation of the star, which makes it theoretically the most robust option (Nardetto et al. 2006). However, it is extremely sensitive to the quality of the spectrum, and then is only possible for good S/N observations. An example of the application of these three methods are given in Fig. 2.8 for two spectral depths.

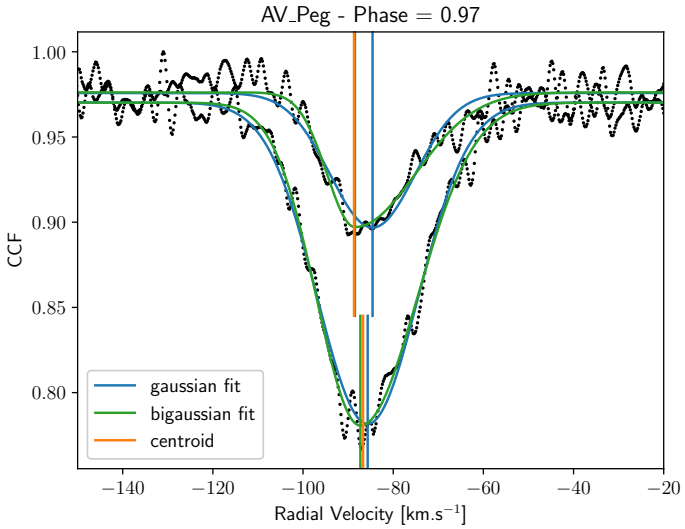


Figure 2.8: CCF and associated RV measurement from a bigaussian fit, a gaussian fit and the centroid method.

In a similar approach as to study the impact on the line selection, I could estimate the impact of the definition of the RV from the CCF profile on the p-factor, using the difference in the amplitude of the integrated RV curve. I concluded that this could have an impact of around 10% on the p-factor, which clearly exceeds the uncertainty on the p-factor. For the stars in my sample with only SOPHIE RV in their model, I computed an alternative model using the bi-Gaussian or the centroid method for RV measurements and retrieved the same 10% difference in the p-factor (see Table. 2.2).

Star	p_g	p_{2g}	p_c	ΔA_{RV}
AB UMa	1.25	1.14	1.26	11%
TT Lyn	1.20	1.09	1.13	8%

Table 2.2: Differences in p-factors from the different definitions of RV from the CCF profile.

If the method used on the CCF profile changes the individual p-factors, it has to be checked whether it has an impact on the dispersion of the p-factors. Indeed, the bi-Gaussian fit tends to increase the amplitude of the RV curve, yielding systematically smaller p-factors. Here, I only have 8 RR Lyrae stars for which I can determine the p-factor with each of the described methods, which are the ones with a good phase-coverage in SOPHIE observations. As shown in

Fig. 2.9, the scatter of the p-factor values is mildly affected by the RV measurement method. Quantitatively, the dispersion is of 0.048, 0.062, 0.055 respectively for the Gaussian fit, the centroid method and the bi-Gaussian fit.

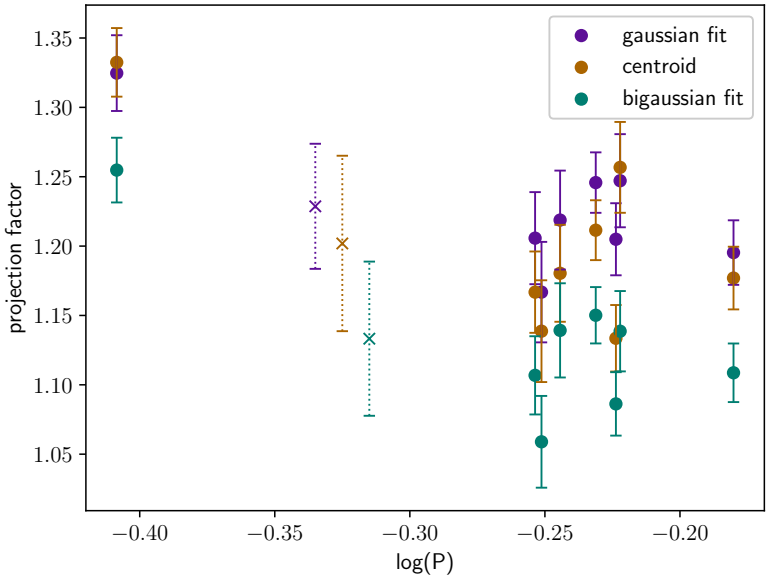


Figure 2.9: $p - P$ relation for p-factors of 8 RR Lyrae derived from different methods of RV measurement. The mean and standard deviation of the sample for each RV measurement method are shown as crosses with dotted error bars.

This is one of the major conclusions I drew from this first analysis. The takeaway I would like to highlight here is that it is extremely important, for p-factor comparison, to compare similar RV measurement methods. In that respect, it should be clearly stated in publications how RVs were obtained, to avoid introducing biases in their interpretation.

5 Paper I

This section presents the paper reporting the analysis I just described. In the previous pages, I presented the main results and the motivations of the publication, and further details can be found in the article.

A&A, 684, A126 (2024)
https://doi.org/10.1051/0004-6361/202348291
© The Authors 2024

Astronomy & Astrophysics

The Baade-Wesselink projection factor of RR Lyrae stars

Calibration from OHP/SOPHIE spectroscopy and *Gaia* DR3 parallaxes★,★★

Garance Bras[1], Pierre Kervella[1], Boris Trahin[2], Piotr Wielgórski[3], Bartłomiej Zgirski[4], Antoine Mérand[5], Nicolas Nardetto[6], Alexandre Gallenne[7,8], Vincent Hocdé[3], Louise Breuval[9], Anton Afanasiev[1], Grzegorz Pietrzyński[3], and Wolfgang Gieren[4]

[1] LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université Paris Cité, 5 place Jules Janssen, 92195 Meudon, France
e-mail: garance.bras@obspm.fr
[2] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[3] Nicolaus Copernicus Astronomical Center, Polish Academy of Sciences, Bartycka 18, 00-716 Warszawa, Poland
[4] Universidad de Concepción, Departamento de Astronomía, Casilla 160-C, Concepción, Chile
[5] European Southern Observatory, 85748 Garching, Munich, Germany
[6] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Nice, France
[7] Instituto de Astrofísica, Departamento de Ciencias Físicas, Facultad de Ciencias Exactas, Universidad Andrés Bello, Fernández Concha 700, Las Condes, Santiago, Chile
[8] French-Chilean Laboratory for Astronomy, IRL 3386, CNRS, Casilla 36-D, Santiago, Chile
[9] Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA

Received 16 October 2023 / Accepted 19 December 2023

ABSTRACT

Context. The application of the parallax-of-pulsation (PoP) technique to determine the distances of pulsating stars implies the use of a scaling parameter, namely the projection factor (p-factor), which is required to transform disc-integrated radial velocities (RVs) into photospheric expansion velocities. The value of this parameter is poorly known and is still debated. Most present applications of the PoP technique assume a constant p-factor. However, it may actually depend on the physical parameters of each star, as past studies designed to calibrate the p-factor (predominantly for Cepheids) led to a broad range of individual values.
Aims. We aim to calibrate the p-factors of a sample of RR Lyrae stars (RRLs) in order to compare them with classical Cepheids (CCs). Due to their higher surface gravity, RRLs have more compact atmospheres, and therefore provide a valuable comparison with their supergiant siblings.
Methods. We determined the p-factor of 17 RRLs by modelling their pulsation using the SPIPS code. The models are constrained using *Gaia* DR3 parallaxes, photometry, and new RVs that we collected with the OHP/SOPHIE spectrograph. We carefully examine the different steps of the PoP technique, in particular the method used to determine the RV from spectra using the classical approach based on the cross-correlation function (CCF).
Results. The method employed to extract the RV from the CCF has a strong impact on the p-factor, that is, of up to 10%. However, this choice of method results in a global scaling of the p-factor, and only marginally affects the scatter of p within the sample for a given method. Over our RRL sample, we find a mean value of $p = 1.248 \pm 0.022$ for RVs derived using a Gaussian fit of the CCF. There is no evidence for a different value of the p-factor of RRLs, although its distribution for RRLs appears significantly less scattered ($\sigma \approx 7\%$) than that for CCs ($\sigma \approx 12\%$).
Conclusions. The p-factor does not appear to depend in a simple way on fundamental stellar parameters (pulsation period, radius, metallicity, amplitude of the RV curve). We argue that large-amplitude dynamical phenomena occurring in the atmospheres of RRLs (and CCs) during their pulsation affect the relative velocity of the spectral line-forming regions compared to the velocity of the photosphere.

Key words. techniques: radial velocities – stars: distances – stars: oscillations – stars: variables: RR Lyrae

1. Introduction

Pulsating stars, and in particular Cepheids and RR Lyrae, are essential indicators for the determination of distances in the local Universe. Precise calibration of the relations between their apparent luminosity, period of pulsation (Leavitt & Pickering 1912; Breuval et al. 2022), and metallicity is therefore particularly important. These calibrations are based on geometrical determinations of the distance using, for instance, parallaxes (see Groenewegen 2018, for example) or detached eclipsing binaries (Riess et al. 2019; Pietrzyński et al. 2019). Such empirical relations play a central role in the extragalactic distance ladder. Period–luminosity relations of Cepheids in the Milky Way, the Magellanic Clouds, and NGC 4258 are used to calibrate the luminosity of type Ia supernovae, whose distances enable us to determine the Hubble constant H_0 (Riess et al. 2022). However, this empirical measurement of H_0 differs by 5σ from the prediction inferred from the *Planck* cosmological

★ Full Table 5 is available at the CDS via anonymous ftp to cdsarc.cds.unistra.fr (130.79.128.5) or via https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/684/A126
★★ Based in part on observations made at Observatoire de Haute-Provence (CNRS), France.

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Subscribe to A&A to support open access publication.

microwave background data assuming the ΛCDM model (Planck Collaboration VI 2020). In this context, there is growing interest in better understanding the physics of pulsating stars in order to explore possible biases that their period–luminosity relations may incur on the distance ladder.

The parallax-of-pulsation (PoP) technique to determine distances of pulsating stars, also known as the Baade-Wesselink technique (Baade 1926; Wesselink 1946), was first introduced by Lindemann (1918). It is based on a comparison of the variation of the linear radius of a pulsating star to the variation of its angular diameter, the ratio of these two quantities giving the distance of the star. The variation of photospheric radius is estimated from the integration of the pulsation velocity curve of the photosphere of the star over its pulsation cycle. The resulting radius curve is then compared to the angular diameter curve, which can be either measured directly by interferometry for nearby stars, or predicted using the surface-brightness-colour relation (SBCR; see Kervella et al. 2004; Nardetto et al. 2023, for instance). The SBCR is generally calibrated using interferometric measurements. Unfortunately, it is not possible to directly measure the pulsational velocity by spectroscopy, but only the radial velocity (RV), which is an averaged quantity over the stellar disc. The ratio $p = v_{\rm puls}/v_{\rm rad}$ between the pulsation velocity $v_{\rm puls}$ and the RV $v_{\rm rad}$ is called the projection factor, or p-factor. Considering a uniformly bright pulsating sphere, the projection on the line of sight of the spherical distribution of the pulsation velocity vector corresponds to $p = 1.5$. The p-factor also incorporates the effect of limb darkening, which results in a higher weight attributed to the RV in the centre of the star compared to the limb. This results in a lower value of p than the purely geometrical value, and typically around $p \sim 1.3$. In addition, p is also affected by velocity gradients that may exist in the atmosphere of the star between the line-forming regions and the continuum-forming photosphere. This aspect was studied theroretically in Nardetto et al. (2004) and with observations in Nardetto et al. (2007). Finally, the p-factor also depends on the method used to estimate the RV from the spectra. Such measurements need to be taken carefully, and well documented, as demonstrated by Anderson (2020) and Borgniet et al. (2019), but this is difficult because of the asymmetry of the lines in the spectrum of a pulsating star.

The intrinsic astrophysical complexity of the p-factor makes its calibration challenging. Using independent distance measurements (e.g. using *Gaia* parallaxes), several authors determined the p-factor of pulsating stars empirically (Storm et al. 2011; Ngeow et al. 2012; Breitfelder et al. 2016; Kervella et al. 2017; Nardetto et al. 2017, 2023; Trahin et al. 2021). The resulting large scatter of the determined p-factors led to a withdrawal of the PoP technique as a way to determine the distances of pulsating stars at the percent accuracy level. However, studying the p-factor is a powerful way to progress in our understanding of the dynamics of the atmosphere of pulsating stars. Among classical pulsators, Cepheids are known to exhibit relatively complex atmospheric dynamics (Fokin et al. 1996; Hocdé et al. 2020a), including mass loss (Neilson & Lester 2008) and circumstellar envelopes (Hocdé et al. 2020b; Gallenne et al. 2021). Conversely, RR Lyrae stars (RRLs), which present intrinsic differences in their atmospheres when compared to Cepheids (e.g. they are physically smaller than Cepheids, with a higher surface gravity, a more compact atmosphere, and a higher temperature), make for a valuable comparison in order to characterise their p-factor. As an example, Jurcsik et al. (2017) found a low dispersion of 4% of the distance of M3 using the Baade-Wesselink technique for a sample of 26 RRLs in this globular cluster. This is the motivation behind the present study, which is intended to expand the investigation into classical Cepheids presented by Trahin et al. (2021) towards the lower-mass RRLs.

In Sect. 2, we present the new RV data that we collected with the OHP/SOPHIE spectrograph, as well as the sources of ancillary literature data (photometry and effective temperature). In Sect. 3, we detail the methodology we used to extract RV measurements from our SOPHIE spectra. Section 4 is dedicated to the analysis of the pulsation of 17 RRLs using the SPIPS modelling tool. In Sect. 5, we summarise our results on the p-factor, as well as those on the physical properties of the modelled stars. Finally, we conclude this study in Sect. 6.

2. Observational data

2.1. Spectroscopic data

We focus our study on a selection of 23 fundamental RRLs observable from the Northern hemisphere; these are listed in Table 1. We obtained new spectra for these RRLs using the SOPHIE spectrograph installed at the 193 cm telescope of the Observatoire de Haute-Provence in southern France. SOPHIE is a cross-dispersed échelle spectrograph (Perruchot et al. 2011) with excellent stability. This instrument typically provides a signal-to-noise ratio (S/N) of 40 within 10 min on the brightest RRLs in our sample. This relatively short exposure time is important, as 10 min already corresponds to 1–2% of the pulsation cycle of a RRL with a period of 10 h. Longer exposure times would potentially result in smearing of the RV curve and underestimation of its amplitude. Our observations were divided into four runs of four or five nights to survey our RRL sample. We used the high resolution mode of SOPHIE, with the slow readout mode to limit the noise. During the night, we organised the sequence of stars according to the previous observations and the predicted phase of the star to obtain regular and complete coverage of the pulsation cycle. We were particularly careful to properly sample the phases between 0.8 and 1 when the stars rebound at their minimal radius, as the velocity changes quickly. For these phases, we obtained several consecutive exposures when possible. The automatic pipeline at OHP reduces the raw spectra and computes a cross-correlation function (CCF) with a specific mask. The RV is then determined as the mean value of a Gaussian profile adjusted to the CCF. For our RRLs, we initially chose an F0 type mask suitable for their typical effective temperature. In the present work, we recomputed the CCF using various methods that are detailed in Sect. 3. Within the list of 26 stars originally selected, we obtained a complete phase coverage for 11 of them (the number of observations is greater than 15, the maximum gap in phase between two points is lower than 0.2, and the number of observations at the rebound is more than four). Six additional stars have a satisfactory phase coverage, although some non-critical phases are incompletely covered. For four stars, part of the RV curve is missing, and two stars have too few RV measurements for an independent analysis. Three stars were not observed at all and are not presented in the table. Finally, the two stars RS Boo and VX Her showed a Blazhko effect during the observations (as confirmed by the literature) and were removed from our sample. We also note the binarity of TU UMa (Liška et al. 2016), which is remarkable when comparing different RV datasets. For the stars with insufficient phase coverage, we completed our RV measurements with data from the literature. A limitation of this approach is that these additional RV measurements are potentially not determined from the spectra using the same method.

Table 1. Selection of 23 RRLs in the northern hemisphere observed with the SOPHIE spectrograph at OHP, with their mean magnitude in band V (m_V), exposure time (Δt), and number of spectra acquired.

Name	m_V	Δt (s)	N	Maximum gap	$N_{\rm rebound}$	$< S/N >$	Refs RV	Refs $T_{\rm eff}$	Refs photometry
V0341 Aql$^{(\dagger)}$	10.89	900	13	0.17	2	22.0			35
X Ari$^{(\dagger)}$	8.97	600	25	0.27	3	26.0	6$^{(\star)}$	12,13	21$^{(\star)}$,24$^{(\star)}$,30$^{(\star)}$,33,35
RS Boo$^{(\dagger,B)}$	9.73	600	11	0.42	3	26.0	8$^{(\star)}$	12$^{(\star)}$,13,16	26$^{(\star)}$
RR Cet$^{(*)}$	9.15	600	10	0.22	3	25.0	9$^{(\star)}$,10	12,14,16	28$^{(\star)}$,29,30$^{(\star)}$,33
W CVn$^{(\ddagger)}$	11.18	900	19	0.08	4	25.0	5	12	23,27,35
UY Cyg$^{(\ddagger)}$	11.10	900	16	0.19	7	17.0			23$^{(\star)}$,27$^{(\star)}$,30$^{(\star)}$,35
DX Del$^{(*)}$	9.52	600	6	0.67	3	31.0	1,10$^{(\star)}$	12$^{(\star)}$,14$^{(\star)}$,15,16$^{(\star)}$	17$^{(\star)}$,29,30$^{(\star)}$,32,34$^{(\star)}$
SU Dra$^{(\ddagger)}$	9.27	600	19	0.18	8	29.0	9,11	13	28$^{(\star)}$,30$^{(\star)}$
SW Dra$^{(\ddagger)}$	10.55	600	19	0.12	7	22.0	2,7	12$^{(\star)}$	19$^{(\star)}$,25$^{(\star)}$,30$^{(\star)}$
RX Eri$^{(\ddagger)}$	9.22	600	18	0.12	7	25.0	9	12$^{(\star)}$,14$^{(\star)}$	28$^{(\star)}$,30$^{(\star)}$,33,35
SV Eri	9.96	900	4	0.62	0	24.0		12	23,30,33,35
VX Her$^{(*,B)}$	11.31	900	7	0.56	4	28.0		12,14,16	23,27,30,35
WZ Hya$^{(*)}$	10.27	900	9	0.43	3	16.0			23
RR Leo$^{(\ddagger)}$	9.94	900	25	0.11	9	22.0	9	12,14	28,30,35
SS Leo$^{(\dagger)}$	11.00	900	12	0.45	6	19.0	4$^{(\star)}$		22$^{(\star)}$,35
TT Lyn$^{(\ddagger)}$	9.51	600	24	0.09	7	33.0	1,9	12	17,18,28$^{(\star)}$,30$^{(\star)}$,35
AV Peg$^{(\ddagger)}$	9.87	600	22	0.13	7	21.0	9	12$^{(\star)}$,14	28$^{(\star)}$,30$^{(\star)}$,31,35
CG Peg	11.13	900	2	0.92	0	18.0			
VY Ser$^{(\dagger)}$	9.73	600	17	0.18	4	24.0	3$^{(\star)}$,11	13,15$^{(\star)}$,16	20,22$^{(\star)}$,26,35,36$^{(\star)}$
AB UMa$^{(\ddagger)}$	10.88	900	18	0.09	4	26.0			27,30$^{(\star)}$
TU UMa$^{(\ddagger)}$	9.26	600	24	0.09	7	32.0	1,9	12$^{(\star)}$	17$^{(\star)}$,18$^{(\star)}$,28$^{(\star)}$,30$^{(\star)}$,33,35
UU Vir$^{(\dagger)}$	9.90	900	10	0.24	3	22.0	1,8,9	12,13	17,18,26,28,30,35
BN Vul$^{(\ddagger)}$	11.13	900	23	0.10	6	24.0			23,35

Notes. The maximum gap and $N_{\rm rebound}$ are two criteria that we used to select RV curves that are sufficiently well covered. The three last columns list the references for the RV, effective temperature, and photometric data. Star symbols $^{(\star)}$ indicate the data sets that were effectively used in the model fitting. $^{(\ddagger)}$Refers to complete RV curves, $^{(\dagger)}$to satisfactorily covered RV curves, and $^{(*)}$to incomplete RV curves. $^{(B)}$ Indicates the stars that exhibit the Blazhko effect. The absence of any symbol indicates that there are too few observations.

References. (1) Barnes et al. (1988); (2) Cacciari et al. (1987); (3) Carney & Latham (1984); (4) Fernley et al. (1990); (5) Fernley et al. (1993); (6) Jones et al. (1987a); (7) Jones et al. (1987b); (8) Jones et al. (1988); (9) Liu & Janes (1989); (10) Meylan et al. (1986); (11) Varsavsky (1960); (12) Andrievsky et al. (2018); (13) Andrievsky et al. (2020); (14) Andrievsky et al. (2021); (15) Takeda et al. (2006); (16) Takeda (2022); (17) Barnes et al. (1988); (18) Barnes et al. (1992); (19) Cacciari et al. (1987); (20) Carney & Latham (1984); (21) Fernley et al. (1989); (22) Fernley et al. (1990); (23) Fernley et al. (1993); (24) Jones et al. (1987a); (25) Jones et al. (1987b); (26) Jones et al. (1988); (27) Layden et al. (2019); (28) Liu & Janes (1989); (29) Meylan et al. (1986); (30) Monson et al. (2017); (31) Paczyński (1965); (32) Paczyński (1966); (33) Preston & Paczynski (1964); (34) Skillen et al. (1989); (35) Sturch (1966); (36) Varsavsky (1960).

2.2. Photometric data

Photometric data were taken exclusively from the literature. To perform the modelling as described below, a large number of observations is required, and in different bands and photometric filters, in order to constrain the pulsating star model. Photometric light curves from *Gaia* (Gaia Collaboration 2016, 2023) and HIPPARCOS (ESA 1997) were employed, as well as 2MASS magnitudes (Cutri et al. 2003). We also used the collection of photometric data assembled by Monson et al. (2017), and we added photometric measurements from the following sources: Barnes et al. (1988), Cacciari et al. (1987, 1989); Carney & Latham (1984); Fernley et al. (1993, 1989, 1990); Jones et al. (1987a,b, 1988); Liu & Janes (1989, 1990); Meylan et al. (1986); Skillen et al. (1989, 1993a,b). Photometry in the bluer filters (U, B, and *Gaia* B_P bands) is generally not included in the model fit, as the stellar atmosphere models that we are using are insufficiently accurate for comparison with observations in this wavelength range because of some effects of the microturbulence (Casagrande & VandenBerg 2014). The *Gaia* G band (G_GAIA_GAIA3 in the models) covers a very broad wavelength range, which leads to difficulties for atmosphere models and for reddening corrections. For this reason, the G band light curves are not systematically used to constrain the atmosphere models along the pulsation cycle. Generally speaking, light curves in visible bands are much more common than in infrared filters in the literature. However, as it provides a strong constraint on the effective temperature and is also mildly sensitive to interstellar reddening, infrared photometry provides a particularly valuable constraint. This encouraged us to only consider part of the visible photometric measurements available in the literature in order to give them comparable weight in the fit to infrared measurements. The choice of data sets was based on phase coverage, agreement with other data sets, and dispersion or uncertainties. The data that were effectively used for the modelling of each RRL of our sample are marked with a star symbol ★ in Table 1.

2.3. Effective temperature measurements

Spectroscopic effective temperature estimates offer a valuable constraint on the atmosphere models of the star. This is due to the fact that these estimates are independent of the interstellar

reddening of the star, and therefore help to mitigate the effects of the strong correlation that exists between the effective temperature and the colour excess $E(B-V)$ of the star. We took such measurements from Andrievsky et al. (2018, 2020, 2021), who used the Spectroscopy Made Easy tool, which compares a given spectrum with synthetic ones with different atmospheric parameters. Measurements from Takeda et al. (2006) and Takeda (2022) using the same approach were also used. APOGEE (Jönsson et al. 2020) and LAMOST (Cui et al. 2012) catalogues are of little use for this kind of work as they do not provide time-series of effective temperature. APOGEE temperatures are measured from a mean spectrum and LAMOST do not provide precise epochs of measurements. However, we were able to check that these measurements are compatible with the obtained models for effective temperature. The number of collected effective temperature estimates is small for a single star. In order to balance the different observables in the fit of the model, we set a similar weight to each kind of observable (RVs, effective temperatures, magnitudes in the visible, and magnitudes in the infrared) in the global χ^2. Even if photometric measurements are much more numerous and precise than RVs, the weight of the latter is properly balanced in the fit. A drawback of this choice is that the normalisation can potentially give disproportionate weight to the temperature measurements compared to the other observables. When temperature measurements were found to hamper the convergence of the model, leading to phase mismatching or unlikely light-curve shapes, we omitted them from the fitting process. However, models are still consistent with the measured temperatures.

3. Methods to determine RV and its impact on the *p*-factor

One key ingredient in the procedure to compute the p-factor of a pulsating star is the RV measurement. From the spectra acquired to the extraction of an RV value, some steps need to be carried out carefully. The most common way to do this is by fitting a Gaussian profile to the CCF. However, this implies a few choices. Firstly, the choice to compute the CCF with a specific mask or template. Secondly, there exists different definitions of the RV given the CCF; it can be defined as the centroid of the line profile as well as the mean of a Gaussian fit or as a parameter of another fitted function. Such differences in the measurements have a large impact on the p-factor as the angular diameter curve is compared to the integrated RV curve. One consequence is that RVs from different methods can lead to very different p-factors, which need to be quantified to identify the origin of the scatter of the p-factors in the pulsating-star population. In this section, we aim to quantitatively describe the impact of each of the aforementioned choices on the p-factor and to motivate one or the other choice, considering physical aspects but also the diversity observed in literature measurements.

To compare the various methods, we interpolate the RV curve using simple linear functions in order to avoid an impact of the choice of function used for the interpolation. This analysis is therefore only possible for really good phase coverage of the curve. We can then integrate $\Delta V = V - V_{\rm mean}$ with time on a pulsation cycle. This results in the integrated RV curve, of which we can measure the amplitude. This amplitude is directly correlated to the p-factor as it corresponds to the scaling of the integrated RV curve to the angular diameter curve. For each aspect studied in the following part, we present an example using a star with good phase coverage (especially at the rebound), along with its

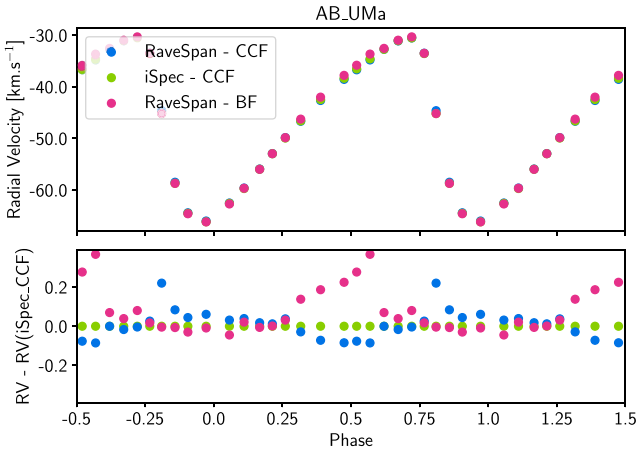


Fig. 1. Comparison between the iSpec and RaveSpan codes, and between the CCF and the BF approaches in analysing the RV curve for AB UMa.

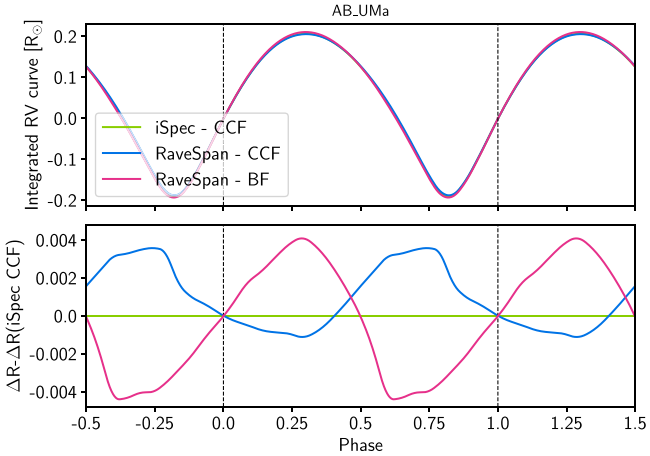


Fig. 2. Comparison between the iSpec and RaveSpan codes, and between the CCF and the BF approaches in analysing the integrated RV curves from Fig. 1 for AB UMa.

RV and integrated RV curves. A comparison with a reference is provided for the integrated RV curve, which is the method used afterwards. In this work, we used two different tools in addition to the SOPHIE pipeline. Firstly, iSpec (Blanco-Cuaresma et al. 2014; Blanco-Cuaresma 2019) was used to compute the CCF with different masks and templates, either standard or newly defined. In employing the CCF calculation method of this code, we tried several strategies, which are detailed in the following section. In parallel, the RaveSpan code (Pilecki et al. 2017) was also used because it allows a comparison of the CCF and broadening function (BF) approaches. Both functions show similar results but the BF has the advantage that it is easily computable from a poor S/N spectrum. Comparisons of those different tools and of the BF and CCF approaches are presented in Figs. 1 and 2 using a Gaussian profile fitting. The cross-correlation was done with the same template as that in Coelho et al. (2005), which is commonly used in RaveSpan, with $T_{\rm eff} = 6000$ K, $\log g = 3.0$, and [Fe/H] $= -1.0$. We compare the CCF and BF approaches using RaveSpan with the CCF computed by iSpec. The difference observed in the integrated RV curve is characterised by a difference of 1% and 2% in the amplitude for the CCF and BF, respectively.

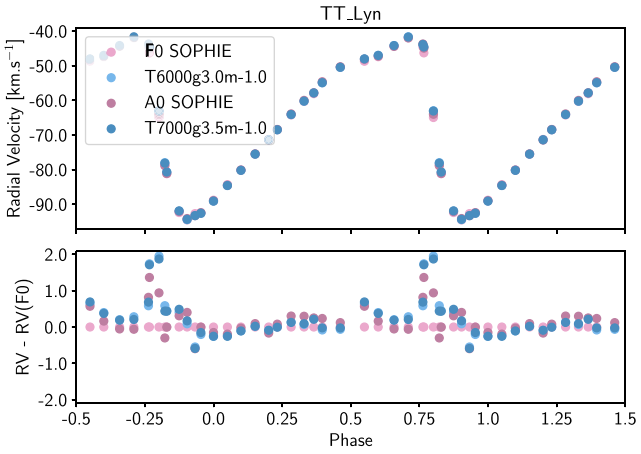


Fig. 3. Comparison between masks and templates in the RV curve for TT Lyn.

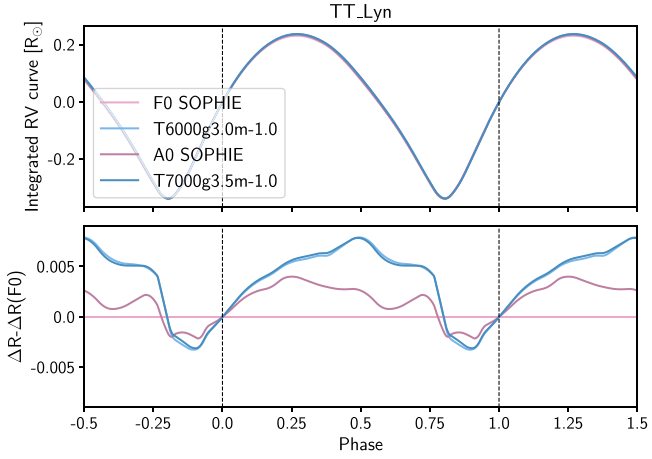


Fig. 4. Comparison between masks and templates in the integrated RV curves from Fig. 3 for TT Lyn. The bottom plot shows the difference between each integrated RV curve and that derived from the F0 mask.

3.1. Masks and templates

We studied the choice between binary masks and synthetic templates for the CCF computation. We compared the standard SOPHIE masks provided in the iSpec tools, and different templates derived from synthetic spectra. The comparison is not provided for spectral masks that are very far from the expected spectral type of RRLs, nor for templates with very different atmospheric properties. One can see the corresponding RV and integrated RV curves in Figs. 3 and 4, between two masks and two templates. Radial velocities were computed from a Gaussian fitting of the cross-correlation with an F0 and an A0 mask and with synthetic templates from Coelho et al. (2005) with properties similar to the masks. The corresponding error in the amplitude of the integrated RV curve is of 1%, and we conclude that it does not have a major impact on the p-factor.

Consequently, we have to keep in mind that the use of generic binary masks or templates, which contain a large number of spectral lines as they are usually used for SOPHIE observations, will probably average the gradient in RV in the atmosphere of the star, leading to a smoothed RV curve.

The choice between masks and templates leads us to the question of the line selection. The CCF is a profile computed from a large number of different spectral lines. However, it has been shown that not all lines are pertinent for the Baade-Wesselink method (Petterson et al. 2005). Indeed, different lines from different elements are forming at different levels on the atmosphere and do not offer a good comparison to angular diameter observations. We therefore studied the impact of the selection of the lines on the p-factor.

3.2. Line selection

The CCF is created by analysing a set of specific lines from a spectrum. These lines are determined by the mask or template used to do the cross-correlation. Nardetto et al. (2007) reported that the amplitude of the RV curve increases with the depth of the line used to measure the RV. Borgniet et al. (2019) observed the effect of cross-correlating lines that are forming at different levels near the photosphere of Cepheids. To do so, these latter authors built masks that select either weak, medium, or deep lines, considering the depths of lines to be a proxy for the optical depth. They built their masks by selecting lines in a synthetic spectrum at 5250K and with solar metallicity. We followed the same approach for our RRLs, with new masks using a synthetic spectrum with a metallicity of –1 dex, a surface gravity of $\log g = 3.0$, and a temperature of 6200 K in order to be closer to RRLs. From such a normalised synthetic spectrum computed with the PHOENIX code (Husser et al. 2013), we applied some masks to extract the position of lines with specific ranges of depths. We defined three ranges: the deep lines for which the line depth is between 0.65 and 0.95; medium lines, with the depth between 0.45 and 0.65; and weak lines with a depth of higher than 0.2 and lower than 0.45. We tried to extend the limits from the definition in Borgniet et al. (2019) for the weak lines mask to include more lines in order to improve the S/N of the CCF. In addition, a mask and a template including all these lines are also constructed. Contrary to the masks, the template is not binary, as it is weighted using the depth of each line in the synthetic spectrum. For this reason it is quite similar to the mask for deep lines as these lines have a major impact on the CCF for the general mask. All these masks are constructed in the wavelength range from 4500 Å to 6800 Å[1]. We note that the difference between the mask and the template for all lines is so small that it is not really pertinent to present both for each measurement. For each step in the analysis, we specify whether we use the binary mask or the template. Figure 5 shows part of the synthetic spectrum as an example, with the delimitation of the different ranges of depths used and the resulting masks. It is important to note that in the line selection, we only keep unblended lines in order to have a shape that is as accurate as possible for the CCF, and not smoothed by some blended lines. The properties of the masks are summarised in Table 2. Nardetto et al. (2007) proposed that the CCF of the weak lines is the most pertinent for the PoP technique as it measures the velocity of the lines that were formed near the photosphere (those probed by interferometric and photometric measurements). Unfortunately, the S/Ns of some of our observations were not sufficient to have a well-defined peak in the CCF, some of them being too noisy. The RV curves for the three depth-restricted masks are shown in Fig. 6. A satisfactory comparison for integrated RV curves for different depths is therefore only provided for deep and medium lines in Fig. 7, as for weak lines, the phase coverage was too poor to obtain a good interpolation on the RV curve. All RVs are measured by fitting a Gaussian profile to the CCF.

[1] Tables containing the list of wavelengths of the selected lines for all masks are available at the CDS.

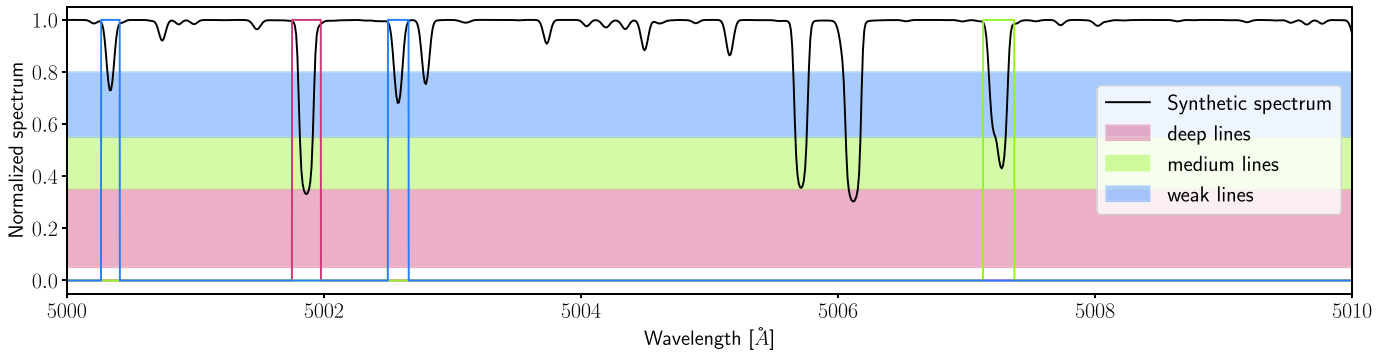


Fig. 5. Portion of the synthetic spectrum used for the template construction and the corresponding built binary masks. The colour encodes the three different depth ranges, both for the background colours that define the depth region of each mask and for the plot of the binary masks themselves.

Table 2. Characterisation of the masks.

	Deep lines	Medium lines	Weak lines
N	80	155	140
$\langle d \rangle$	0.72	0.55	0.37
$\langle \lambda \rangle$ (nm)	498.7	524.7	539.0
$\langle \sigma \rangle$ (nm)	0.029	0.020	0.017

Notes. We give the number of selected lines N, the mean depth $\langle d \rangle$, the mean wavelength $\langle \lambda \rangle$, and the mean width $\langle \sigma \rangle$.

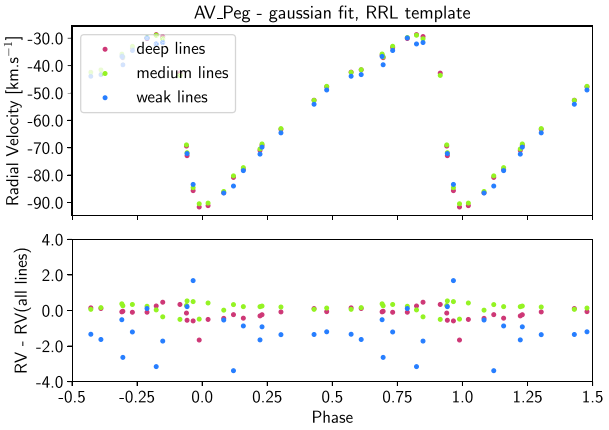


Fig. 6. Comparison between weak, medium, deep, and all lines. This figure shows the RV curve from the cross-correlation of the spectra with a binary mask built for different line selection of AV Peg. The bottom plot shows the difference with the mask keeping all lines with a weight associated to its depth.

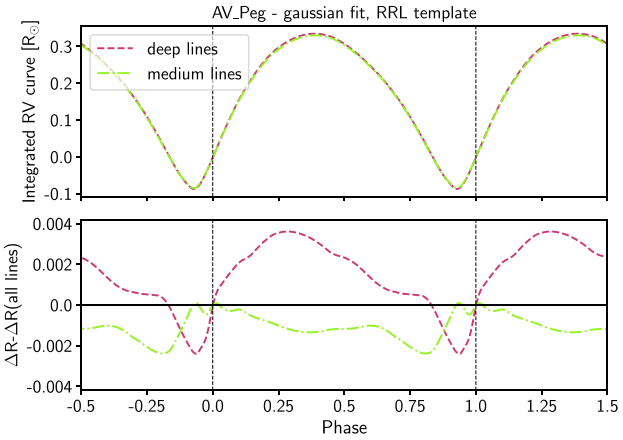


Fig. 7. Comparison between medium, deep, and all lines. This figure shows the integrated RV curve (showed in Fig. 6) for AV Peg. The bottom plot shows the difference with the mask keeping all lines with a weight associated to its depth.

We do not see a significant difference between RV measurements from masks on different line depths (1.3% and 0.3% in the amplitude of the integrated RV curve for deep lines and medium lines, respectively, with respect to all lines), which is similar to the findings of Borgniet et al. (2019; see Fig. 7). Also, Petterson et al. (2005) and Nardetto et al. (2007) showed that, for Cepheids, the difference in RV for different individual lines at different layers is stronger for long-period stars. We can extrapolate these observations and consider this effect to be negligible for RRLs. To further investigate the different CCFs using these different masks, we chose one star with high-quality spectra (S/N typically greater than 35 at 550 nm) and compared the CCF using the masks for the three ranges of depth. Among our 23 stars, DX Del provides good spectra for this analysis at different phases, and so we can monitor how the shape evolves during the pulsation. For this star, the asymmetry of the CCF, which reverses between the phases of higher and lower velocity, is visible to the naked eye (Fig. 8). This observation is compatible with the conclusions of Nardetto et al. (2006) based on single lines, who pointed out that the asymmetry of the line variation curve has, in the case of Cepheids and at first order, a shape similar to that of the RV curve itself. Some characterisation of the CCF is given in Table 3. We computed the bisector of the line, spanning over 3σ (with σ being the standard deviation of the Gaussian) on both sides of the centre of the line to avoid the noise from the wings. From this bisector, we obtain $\langle b \rangle$, the mean value of the bisector in this part of the line, and the amplitude in velocity Δb. The CCF is also fitted by a bi-Gaussian profile with widths of σ_1 and σ_2 on the bluer and redder side (positive RV) of the CCF, respectively. This table shows the correlation between the RV measurement from a Gaussian fit and the mean of the bisector b. The absolute value Δb and the signed value $\sigma_2^{2g} - \sigma_1^{2g}$ are markers of the asymmetry of the CCF. The sign of the latter indicates which of the CCF wings is more extended: if this parameter is positive or negative, the larger wing extension is on the red or blue part of the CCF, respectively. Both evolve in the same way with the phase or the depth of the mask. We note that the asymmetry is systematically higher for medium and weak lines than for deep lines, which is consistent with previous observations (Anderson 2016). However, it is either of the same order or higher for medium lines with respect to weak lines.

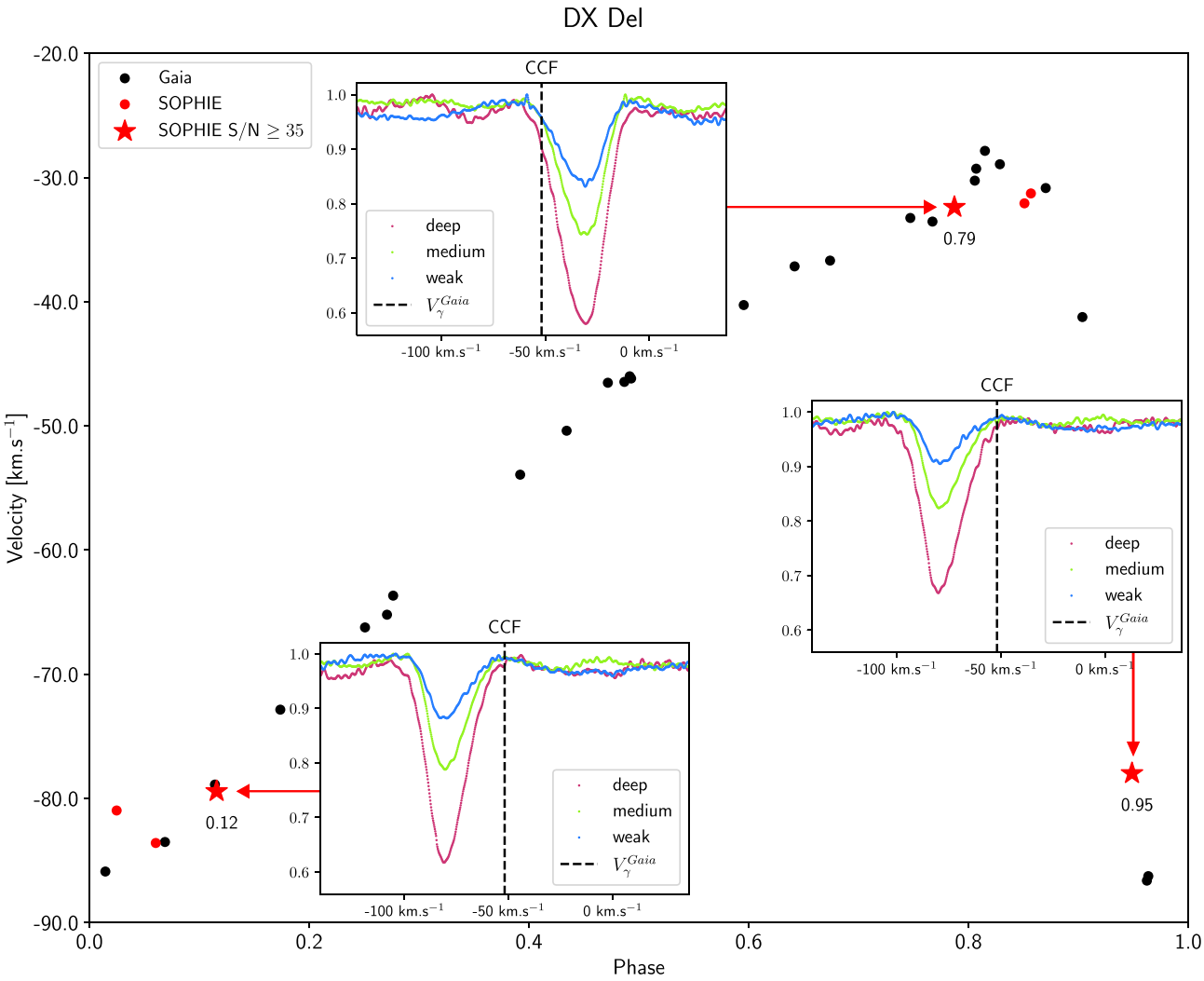


Fig. 8. Evolution of the asymmetry of the CCF during the pulsation cycle of DX Del. The *Gaia* RV curve for this star is shown in black for global visualisation. The red markers represent the measurements from SOPHIE and the red stars locate three spectra with a high S/N where we can compute the CCF using the three masks with different depths, shown in the inserts. The dashed line shows the mean RV extracted from the *Gaia* catalogue.

Borgniet et al. (2019) observed that weak lines are more sensitive to the asymmetry.

3.3. Measurement method on the CCF

The RV is then measured through the Doppler shift of the CCF. However, spectral lines in pulsating stars show a high level of asymmetry, which can bias the various possible estimators of this Doppler shift to a varying extent. For instance, a Gaussian fit of the CCF may be biased due to the asymmetry. An alternate approach is to fit a bi-Gaussian profile, defined as a Gaussian with different widths for the two sides of the profile. The bi-Gaussian profile reproduces the CCF of most of the data well (and therefore the line asymmetry derived from it will be properly determined), but does not produce a suitable RV for the projection factor (besides the additional information of the asymmetry that it provides, its mean value is almost consistent with the velocity associated to the minimum of the line, albeit less noisy) because it is highly sensitive to the line asymmetry generated, for example, by rotation. Instead, the centroid method, which is insensitive to rotation and width variation, is often preferred (Nardetto et al. 2006).

Here we present the effect of the choice of method on the RV curve (plotted in Fig. 9), and the resulting difference in the amplitude of the integrated RV curve (plotted in Fig. 10). Comparisons of RV curves using different techniques have been carried out for single lines (e.g. Nardetto et al. 2006), and on CCF for Cepheids (Anderson 2016) with the conclusion that there is a non-negligible difference in the RV curves when using one or the other method (i.e. Gaussian, centroid, bi-Gaussian etc.). We studied five definitions: the mean of a Gaussian fit, the mean of a bi-Gaussian fit, the mean of a Gaussian fit of the core of the CCF (defined as containing approximately 68% of the CCF absorption), the centroid, and the minimum using a fourth-order polynomial fit. For AB UMa, the difference in the amplitude of the curve ranges from 0.2% (centroid method) to 11% (bi-Gaussian fit) of the amplitude compared to the classic Gaussian fit method. Based on these differences, we estimate that the choice of the definition of the RV given a CCF can lead to a change in the p-factor of up to 10%.

As a summary of this analysis, Table 4 provides the major sources of variation in the amplitude of the integrated RV curve for each aspect studied, and for two different stars. We conclude that the main source of variation of the p-factor is caused by

Table 3. RV and asymmetry measurements on three spectra of DX Del.

		Deep	Medium	Weak
$\phi = 0.12$	RV_g	–79.71	–79.11	$-79.50_{\pm 0.01}$
	Q	6.6	6.5	5.8
	$\langle b \rangle$	–79.63	–78.85	–79.04
	Δb	2.6	4.2	3.5
	$\sigma_2^{2g} - \sigma_1^{2g}$	2.2	3.9	3.8
$\phi = 0.79$	RV_g	–32.35	–32.33	$-32.69_{\pm 0.01}$
	Q	6.2	6.1	5.7
	$\langle b \rangle$	–32.47	–32.54	–32.93
	Δb	3.5	4.5	3.9
	$\sigma_2^{2g} - \sigma_1^{2g}$	–2.8	–3.9	–3.7
$\phi = 0.95$	RV_g	–78.24	–77.67	$-78.11_{\pm 0.01}$
	Q	6.6	6.6	6.0
	$\langle b \rangle$	–78.11	–77.33	–77.68
	Δb	4.2	4.9	3.6
	$\sigma_2^{2g} - \sigma_1^{2g}$	3.4	4.8	4.0

Notes. All velocities are in km s^{-1}. Q is the quality factor as defined by Borgniet et al. (2019). Uncertainties smaller than 0.01 km s^{-1} in RV are not shown.

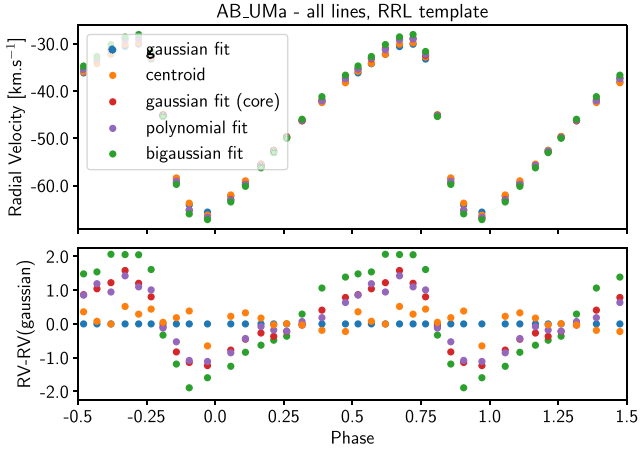


Fig. 9. Comparison between definitions of RV given a CCF. The cross-correlation is computed with all lines weighted by depth mask.

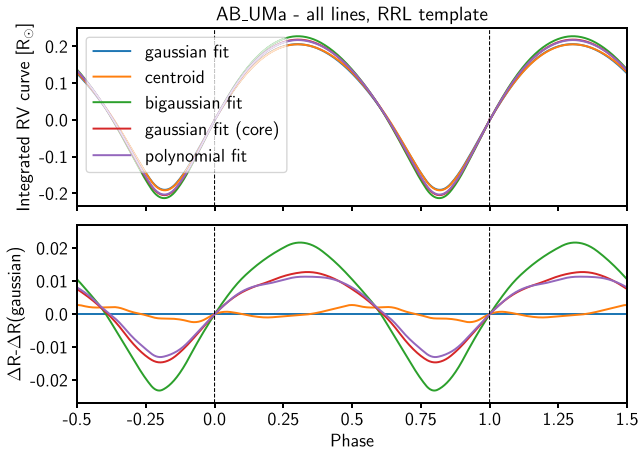


Fig. 10. Comparison of the impact of different CCF measurement methods on the radius curve of AB UMa (the RV curve is shown in Fig. 9). The bottom plot shows the difference from the more commonly used method, that is, the fit of a Gaussian profile to the CCF.

Table 4. Summary of the impact on the amplitude of the integrated RV curve when using different methods to determine RV for AB UMa and TT Lyn.

Source of error	AB UMa	TT Lyn
Code CCF	0.81%	0.67%
Masks/Templates	1.40%	1.11%
Line selection	1.42%	0.35%
CCF measurement method	11.17%	8.35%

the choice of the CCF measurement method – rather than the choice of mask or template, the code used to compute the CCF, or even the line selection. It therefore appears that the choice of CCF measurement method can significantly affect the RV, and therefore also the p-factor.

For the remainder of our analysis and the determination of the p-factor, we used the mean of the Gaussian fit of the CCF, as this is the most common method but is also the most precise with a poor S/N (Anderson 2016). This allows us to directly compare our values of the p-factor with previous determinations from the literature.

Velocity measurements are provided in Table 5 and are available at the CDS. All RV curves are plotted in Fig. 11, and are colour-coded according to pulsation period. SPIPS-modelled curves are shown in grey (see details in Sect. 4). For stars without a model, a dashed curve is plotted, which is the weighted mean of models of the most similar stars in terms of period (of those for which a model has been determined). This is interesting as it reveals the continuity of the shape of RV curves with changing period.

4. The SPIPS modelling

4.1. The SPIPS code

The Spectro-Photo-Interferometry for Pulsating Stars (SPIPS, Mérand et al. 2015) is a modelling code that simultaneously fits a broad range of observables of a pulsating star to determine its physical parameters (pulsation period, bolometric luminosity, distance or projection-factor, effective temperature, radius, reddening, etc.). SPIPS allows us to simultaneously fit different types of observations for a single star: RVs, spectroscopic effective temperatures, angular diameters from interferometry and photometric magnitudes in all bands and colours. Firstly, an interpolated curve is fitted to the RV and another for temperature data. Then, the other observables are predicted as a function of pulsation using the fitted parameters, and the photometry is modelled thanks to a grid of static ATLAS9 models (Castelli & Kurucz 2003) for the given temperature, metallicity, and surface gravity. The latter is estimated from the star's radius and the period–mass–radius relation from Bono et al. (2001). Such relations are calibrated for Cepheids but surface gravity does not have a large influence on the large-band photometry, with an impact on the p-factor of $\lesssim$1%. Filter transmissions and photometric zero points are used to compare with observations in a given set of filters.

4.2. The modelling strategy

To determine the p-factor for our RRLs, we adopted the distances from Bailer-Jones et al. (2021) adapted from the *Gaia* DR3 catalogue (Gaia Collaboration 2016, 2023) and considering the zero-point correction on parallax given by Lindegren et al. (2021). The distances of the stars in our sample

Table 5. Measurements of RV from spectra with the SOPHIE spectrograph.

Star	HJD	S/N	RV_g	σ_{RV_g}	RV_{2g}	$\sigma_{RV_{2g}}$	RV_{deep}	$\sigma_{RV_{deep}}$	RV_{medium}	$\sigma_{RV_{medium}}$
V0341 Aql	2459756.5572	25.2	–114.86	0.11	–119.67	0.52	–115.16	0.05	–	–
V0341 Aql	2459756.5951	15.2	–109.72	0.12	–110.11	0.56	–109.50	0.06	–108.99	0.17
V0341 Aql	2459756.4765	10.6	–51.49	0.03	–49.53	0.17	–50.92	0.02	–51.12	0.07

Notes. We provide values from a Gaussian (RV_g) and a bi-Gaussian (RV_{2g}) fit of the CCF of spectra with the computed binary mask containing all unblended lines and from a Gaussian fit of the CCF of spectra with the computed binary masks containing deep lines (RV_{deep}) or medium lines (RV_{medium}). The errors come from the fit of the Gaussian or bi-Gaussian fit and may be underestimated. This table is available in its entirety in the CDS.

range between 536 and 1258 pc. For such nearby stars, errors on the parallax are quite small, and therefore one could also choose the distance as the inverse of the parallax. We checked that this choice does not affect the result. The impact is around 2% on the p-factor, which is quite similar to the uncertainty on this value. As we do not have spectroscopic temperature measurements for all stars, we adopted the colour excess $E(B-V)$ predicted by the Stilism 3D maps using the same distances (Capitanio et al. 2017). Metallicities were set to the homogeneous values reported in Monson et al. (2017; which include original values from Fernley et al. 1998a,b; Fernley & Barnes 1997), and we considered a typical uncertainty of ±0.15 dex. Because of the large number of parameters in the global fit, we individually checked the convergence of the optimisation to convincing solutions. As stars in our sample are relatively close and bright, they have been well observed and realistic initial parameters are easy to find. We chose to initialise the period, the gamma velocity, and the amplitude of the RV curve to the *Gaia* values found by Clementini et al. (2023). The starting value of the star's radius is defined according to the initial period thanks to the period–radius relation of Marconi et al. (2015).

It is important to note that the various data used are separated by up to several decades, as many photometric measurements were collected in the 1980s. Such an interval corresponds to tens of thousands of pulsation cycles. Therefore, the fit is highly sensitive to the period. This is why we only fit the period in the last step, allowing only very small variations. We also set the reference time to be in the average of all the data in order to avoid adding the errors as many times as the number of cycles between the older and the more recent observation. Finally, we occasionally considered the temporal change in the pulsation period as an additional parameter, depending on the goodness of the fit.

Here is the strategy we adopted to model the pulsation cycle of a typical star in our sample. First, we only fit the effective temperature curve with temperature measurements (if any) and photometry. The curve is fitted using Fourier series with as many harmonics as necessary to properly reproduce the data (between 6 and 8). Photometric colours are not used for the general fit as we favoured single-band photometric measurements. In a second step, we only fit the RV model curve to RV measurements. Because the phase coverage is generally not very dense, Fourier series are not suitable for RV curves as they introduce small oscillations. We therefore used cubic splines, with between six and eight nodes. The choice of the number of Fourier series or nodes of splines to use was made through a visual inspection of the model and of the individual and global χ^2. Given these first guesses of both RV and temperature curves, we can then fit all the parameters but the period, before a finer and final tuning including the period determination. Metallicities, colour excesses, and distances are fixed during the whole procedure, but the impact of each of these parameters has been considered in the final p-factors and the associated uncertainties.

Because we were not able to obtain a good model with this procedure for some stars, we then adapted the code individually for each star to reach a good result in the fitting curves, that is, for RS Boo, AV Peg, and TU UMa. For these three stars, because of an observed phase shift between the model and some datasets that was not reduced by simply adding a period change, we performed a first model based on contemporaneous data and used the fitted parameters as initial parameters for a second run, including the other datasets. This helped the fitting process to converge. For RS Boo, the initial parameters were computed on data exclusively from Jones et al. (1988). For AV Peg we used the Monson et al. (2017) dataset together with the OHP/SOPHIE RV curve and Andrievsky et al. (2018) temperature for the first step. For TU UMa, the initial parameters were obtained using Monson et al. (2017) and *Gaia* photometry, the OHP/SOPHIE RV curve, and the temperature provided by Andrievsky et al. (2018) . In the particular case of RS Boo, its Blazhko effect prevented us from fitting one observable with various datasets, and so the RV curve from OHP/SOPHIE is not fitted. This leads to a higher uncertainty in the fit, especially for the period determination.

We found a good model for eight RRLs with new measurements taken at OHP using a Gaussian fit of the cross-correlation with a binary mask that includes all the identified unblended lines in a synthetic spectrum. We completed this sample with six additional stars from our list, using literature RVs that were offset to match our SOPHIE RV measurements (such an additive offset does not affect the p-factor). We also added three more southern stars using exclusively data from the literature.

Figures 12 and 13 present the result of the SPIPS modelling of two RRLs with new SOPHIE RV measurements. Others are provided in Appendix A[2]. The plot shows velocity, temperature, angular diameter, and photometric curves, both observed (points) and modelled (solid lines). The model curves shown in grey are computed based on the green curves whose parameters are optimised during the fitting process to fit the various observables simultaneously. For the photometric curves, the filter is indicated in blue.

5. Results

In total, we modelled 17 RRLs using the SPIPS code. The properties of the modelled RRLs are summarised in Table 6.

[2] Models are all available at the CDS.

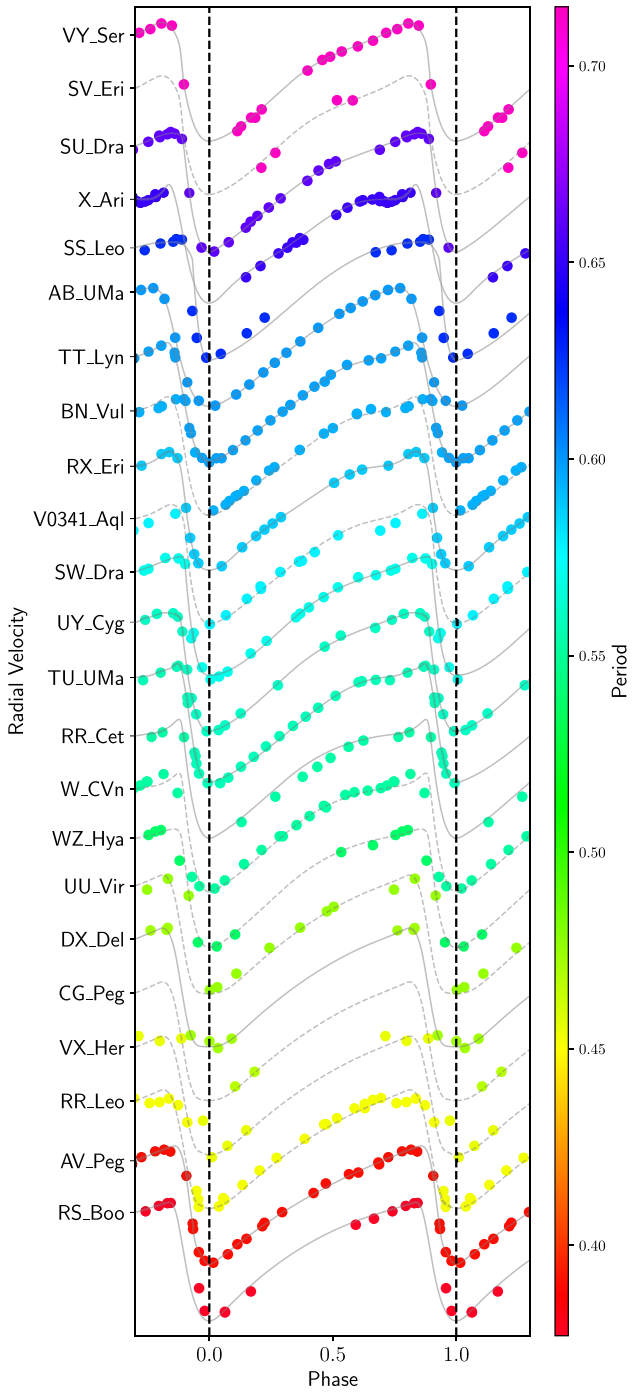


Fig. 11. All RV curves obtained with the SOPHIE spectrograph. The curves have been normalised using *Gaia* parameters if an insufficient number of measurements was obtained. An offset in velocity is added to see all curves in a single plot, which is related to the period.

5.1. p-factor

We compare our determinations of p-factors (listed in Table 6) with those of several previous Baade-Wesselink analysis – in terms of the ratio between the distance and the p-factor d/p. All these studies adopted a constant p-factor of between 1.3 and 1.36 to determine the distances of RRLs and their absolute magnitudes. A detailed comparison is provided in Table 7, and we summarise our main conclusions below.

– We note that our results for the ratio d/p are in good agreement with those of Fernley et al. (1989), Skillen et al. (1989), Fernley et al. (1990), Skillen et al. (1993a) for X Ari, DX Del, SS Leo, VY Ser, and RV Oct, but not for BB Pup and WY Ant, and are also in agreement with the results of Liu & Janes (1990) for RR Cet, RX Eri, AV Peg, and TU UMa, but not for TT Lyn and SW Dra.
– For the stars present in the studies by Jones et al. (1987a; 1987b; 1988; X Ari, SW Dra, VY Ser), our results do not agree in terms of the d/p ratio but the results for X Ari and VY Ser in the mentioned papers are also not consistent with the ones in Fernley et al. (1989, 1990), closer to the results of the present paper.
– Similarly, our results for RR Cet and DX Del are comparable with those in Skillen et al. (1989) and Liu & Janes (1990), but differ from the results of Burki & Meylan (1986b,a).
– For SW Dra, we find a different ratio d/p from all other authors (Jones et al. 1987b; Cacciari et al. 1989).
– Finally, it is worth noting that, even if they agree within the error bars, ratios d/p derived from the SPIPS algorithm are almost systematically higher than those in the literature, leading to smaller p-factors.

The resulting p-factors of the 17 RRLs (defined for a Gaussian fit of the CCF) are plotted against the pulsation period in Fig. 14, encoded according to the origin of the RV measurements, and together with a linear fit and a constant fit. We find a mean value of $p = 1.248 \pm 0.022$

In Fig. 15, we compare the dispersion of the RRL p-factors with that of the Cepheids studied by Trahin et al. (2021), who also used *Gaia* EDR3 parallaxes. The scatter for RRLs is smaller (7%) than for Cepheids (12%). It would be interesting to compare this to other classes of pulsating stars like δ Scuti (Nardetto et al. (2014) found higher p-factors with quite low dispersion of 2% for a sample of four stars) or type II Cepheids for instance (work by Wielgórski, in prep.). However, the scatter in p for RRLs remains too high to use the PoP technique to measure their distances with an accuracy of 1%.

Our p-factor distribution is compatible with a constant p-factor for all stars (RRLs and Cepheids) of $p = 1.245 \pm 0.015$, which is uncorrelated with period. This constant value is compatible with previous results obtained with the SPIPS algorithm for Cepheids in the Magellanic Clouds and in the Milky Way (Gallenne et al. 2017), and is relatively close to the result of Breitfelder et al. (2016).

Comparison of the p-factor of the RRLs in our sample with some of their parameters (see Fig. 16) did not reveal any correlation, that is, with the effective temperature, the radius, the metallicity, or the amplitude of the RV curve. A more complex dependence is not excluded but further analysis would be necessary to investigate this properly.

5.2. Period–luminosity relation

From the SPIPS models, we also have access to the main parameters of the star, which allows us to determine period–radius and period–luminosity relations. While period–luminosity relations are the principal tool for distance determinations, they are not always defined for RRLs (Bhardwaj et al. 2023). Figure 17 shows this latter relation for the dereddened absolute magnitude in the Wesenheit index $W_{JK} = K_S - 0.756(J - K_S)$ (using the reddening law from Fitzpatrick 1999 with $R_V = 3.1$), taking into account a metallicity effect. We find the relation $M_{W_{JK}} = -2.838 \log P + 0.126[\mathrm{Fe/H}] - 1.161$. The scatter around this relation is about 0.05 mag, which is smaller than for Cepheids (0.086 mag; see Breuval et al. 2022).

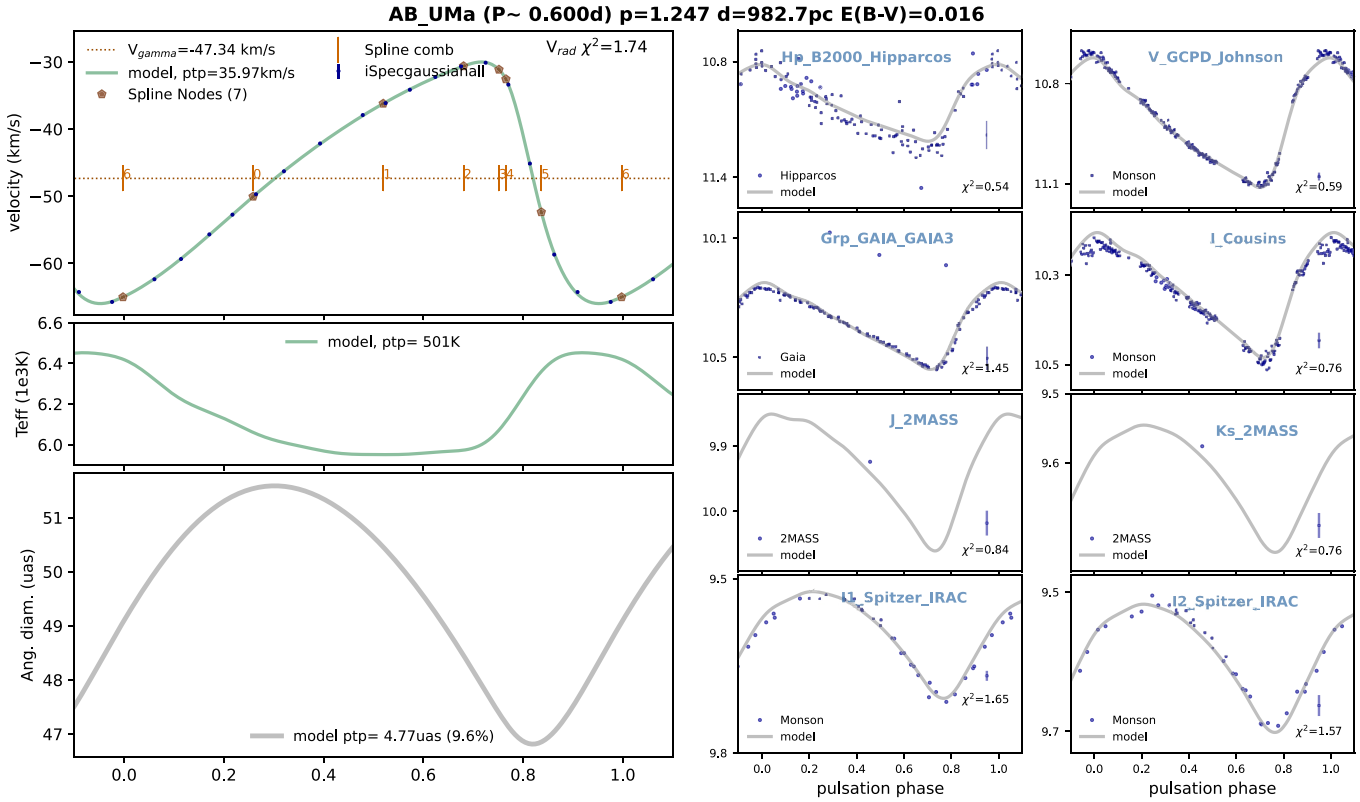


Fig. 12. SPIPS model of AB UMa with RV from SOPHIE spectra (Gaussian fit of the CCF using the binary mask built for RRL with all unblended lines). Photometry measurements are from Monson et al. (2017; 60% of available photometric measurements).

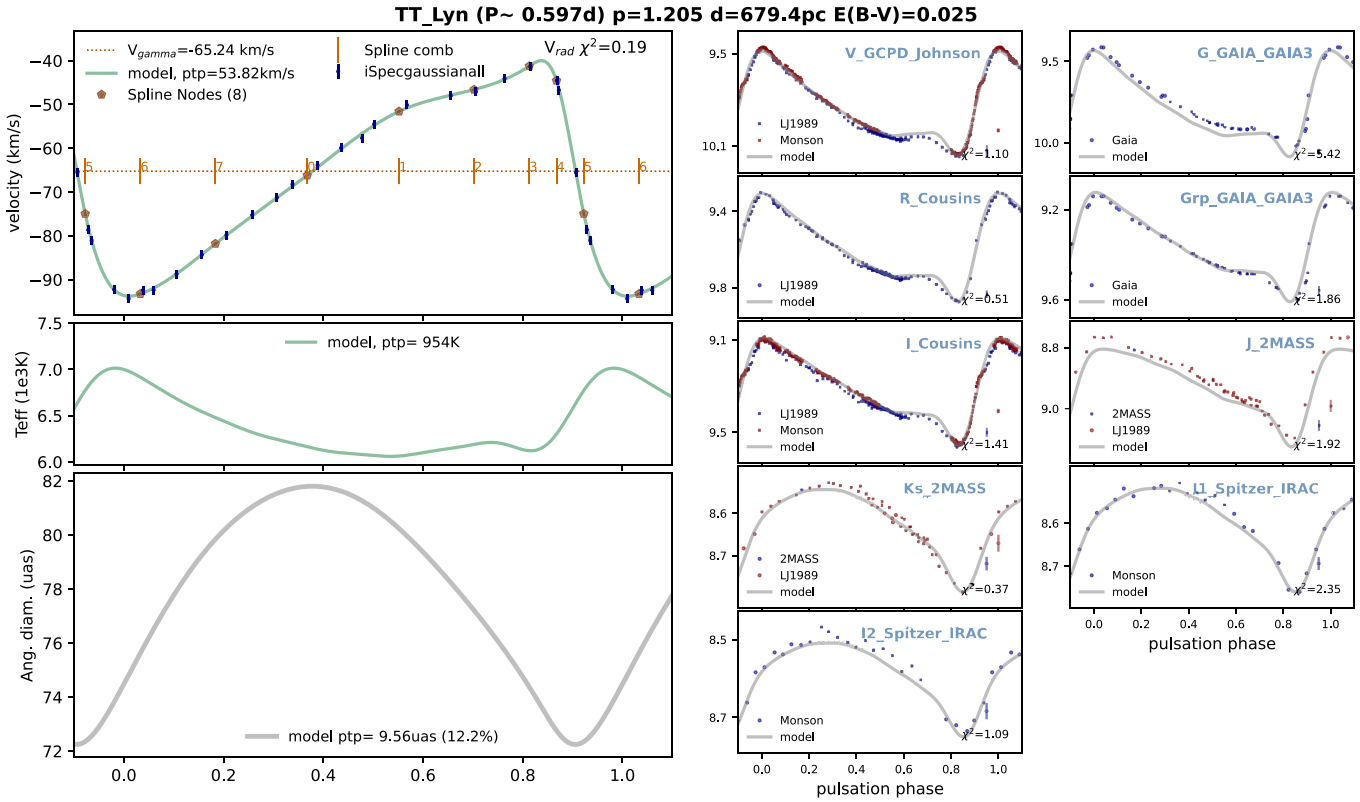


Fig. 13. SPIPS model of TT Lyn with RV from SOPHIE spectra (Gaussian fit of the CCF using the binary mask built for RRLs with all unblended lines). Photometry measurements are from Monson et al. (2017) and Liu & Janes (1989; 29% of available photometric measurements).

Table 6. Main properties of the best-fit models for 17 field RRLs.

Star	d (pc)	[Fe/H]	$E(B-V)$	P (h)	σ_P (h)	$\dot{P}$ (ms yr^{-1})	R ($R_\odot$)	$T_{\rm eff}$ (K)	V_γ (km s^{-1})	p	χ^2_r
UY Cyg	1039^{+22}_{-23}	–0.80	$0.149^{+0.026}_{-0.026}$	13.4569641	28×10^{-7}	–	$5.14_{\pm0.09}$	$6658_{\pm8}$	$7.62_{\pm3.01}$	$1.17_{\pm0.04}$	2.8
SU Dra	741^{+9}_{-7}	–1.80	$0.022^{+0.000}_{-0.014}$	15.8501639	12×10^{-7}	–	$6.15_{\pm0.15}$	$6458_{\pm4}$	$-164.74_{\pm0.95}$	$1.20_{\pm0.03}$	2.8
SW Dra	930^{+17}_{-16}	–1.12	$0.022^{+0.000}_{-0.010}$	13.6721226	8×10^{-7}	$0.88_{\pm0.64}$	$5.58_{\pm0.11}$	$6421_{\pm8}$	$-29.05_{\pm0.98}$	$1.22_{\pm0.04}$	2.0
RX Eri	580^{+7}_{-7}	–1.33	$0.031^{+0.013}_{-0.013}$	14.0939041	18×10^{-7}	–	$5.54_{\pm0.18}$	$6214_{\pm5}$	$67.90_{\pm1.03}$	$1.25_{\pm0.02}$	1.8
TT Lyn	679^{+7}_{-8}	–1.56	$0.025^{+0.000}_{-0.013}$	14.3383528	15×10^{-7}	$-3.76_{\pm0.77}$	$5.70_{\pm0.16}$	$6359_{\pm8}$	$-65.24_{\pm0.86}$	$1.20_{\pm0.03}$	1.6
AV Peg	665^{+5}_{-10}	–0.08	$0.062^{+0.003}_{-0.003}$	9.3691917	6×10^{-7}	$9.36_{\pm0.51}$	$3.87_{\pm0.11}$	$6603_{\pm6}$	$-56.07_{\pm0.53}$	$1.32_{\pm0.03}$	0.9
AB UMa	983^{+18}_{-17}	–0.49	$0.016^{+0.011}_{-0.015}$	14.3899565	28×10^{-7}	–	$5.25_{\pm0.10}$	$6137_{\pm5}$	$-47.34_{\pm0.10}$	$1.25_{\pm0.04}$	1.1
TU UMa	626^{+11}_{-11}	–1.51	$0.018^{+0.002}_{-0.022}$	13.3837597	5×10^{-7}	$1.71_{\pm0.02}$	$5.16_{\pm0.15}$	$6407_{\pm6}$	$96.69_{\pm0.69}$	$1.21_{\pm0.06}$	1.9
X Ari	536^{+5}_{-6}	–2.43	$0.254^{+0.003}_{-0.043}$	15.6279509	21×10^{-7}	$16.39_{\pm0.78}$	$6.20_{\pm0.21}$	$6789_{\pm16}$	$-36.00_{\pm1.01}$	$1.29_{\pm0.04}$	1.5
RS Boo	745^{+9}_{-10}	–0.36	$0.019^{+0.000}_{-0.019}$	9.0561197	32×10^{-7}	$47.36_{\pm0.20}$	$4.14_{\pm0.10}$	$6700_{\pm36}$	$-3.61_{\pm0.70}$	$1.19_{\pm0.07}$	4.8
RR Cet	610^{+10}_{-8}	–1.45	$0.028^{+0.003}_{-0.022}$	13.2726998	8×10^{-7}	–	$5.30_{\pm0.16}$	$6408_{\pm5}$	$-71.80_{\pm1.86}$	$1.32_{\pm0.03}$	1.6
DX Del	568^{+5}_{-5}	–0.39	$0.067^{+0.024}_{-0.024}$	11.3428460	3×10^{-7}	–	$4.59_{\pm0.15}$	$6429_{\pm4}$	$-56.43_{\pm0.72}$	$1.30_{\pm0.03}$	1.8
SS Leo	1258^{+40}_{-36}	–1.79	$0.023^{+0.005}_{-0.023}$	15.0321238	17×10^{-7}	–	$5.74_{\pm0.08}$	$6498_{\pm12}$	$163.40_{\pm0.60}$	$1.06_{\pm0.05}$	2.5
VY Ser	827^{+16}_{-15}	–1.79	$0.045^{+0.005}_{-0.037}$	17.1383252	38×10^{-7}	$-0.38_{\pm0.19}$	$6.55_{\pm0.15}$	$6241_{\pm11}$	$-144.48_{\pm0.81}$	$1.39_{\pm0.06}$	1.9
WY Ant	1019^{+22}_{-20}	–1.48	$0.042^{+0.012}_{-0.012}$	13.7842117	6×10^{-7}	–	$5.43_{\pm0.10}$	$6370_{\pm5}$	$203.53_{\pm0.27}$	$1.12_{\pm0.03}$	1.9
RV Oct	978^{+12}_{-11}	–1.71	$0.105^{+0.019}_{-0.019}$	13.7080606	4×10^{-7}	–	$5.57_{\pm0.11}$	$6306_{\pm6}$	$141.01_{\pm0.20}$	$1.46_{\pm0.04}$	1.9
BB Pup	1604^{+43}_{-39}	–0.64	$0.063^{+0.024}_{-0.064}$	11.5331719	5×10^{-7}	–	$4.70_{\pm0.05}$	$6390_{\pm7}$	$132.65_{\pm0.90}$	$1.35_{\pm0.09}$	1.5

Notes. The distance (d), metallicity ([Fe/H]), and colour excess $E(B-V)$ are fixed. We fit the period (P with uncertainty σ_P), if needed a change in period ($\dot{P}$), and the p-factor. From the RV and photometry curves, we determine the mean radius R, the mean effective temperature $T_{\rm eff}$, and the gamma velocity V_γ. The three groups correspond to stars with SOPHIE RV data exclusively, stars with some additional literature RV data, and stars with literature RV data only, respectively.

Table 7. Comparison of results for the ratio d/p from the present study (second column) and previous determinations (last column).

Name	$d_{\rm Bailer-Jones}/p_{\rm SPIPS}$	Reference BW	$d_{\rm BW}$ (pc)	$p_{\rm BW}$	$d_{\rm BW}/p_{\rm BW}$
WY Ant	910 ± 31	Skillen et al. (1993a)	1034 ± 10	1.33 (*)	777 ± 8
		Fernley et al. (1989)	505 ± 30	1.33 (*)	380 ± 23
X Ari	416 ± 14	Jones et al. (1987a)	440 ± 30	1.30 (*)	338 ± 23
		Jones et al. (1988)	466 ± 13	1.30 (*)	358 ± 10
RS Boo	626 ± 38	Jones et al. (1988)	741 ± 22	1.30 (*)	570 ± 17
RR Cet	462 ± 13	Burki & Meylan (1986b)	760 ± 40	1.36 (*)	559 ± 29
		Liu & Janes (1990)	587 ± 32	1.32 (*)	445 ± 24
DX Del	437 ± 11	Burki & Meylan (1986a)	750 ± 45	1.36 (*)	559 ± 29
		Skillen et al. (1989)	593 ± 50	1.33 (*)	446 ± 38
SU Dra	618 ± 17	Liu & Janes (1990)	640 ± 35	1.32 (*)	485 ± 27
		Cacciari et al. (1989)	840 ± 50	1.36 (*)	618 ± 37
SW Dra	762 ± 29	Jones et al. (1987b)	824 ± 40	1.30 (*)	634 ± 31
		Jones et al. (1988)	836 ± 23	1.30 (*)	643 ± 18
RX Eri	464 ± 9	Liu & Janes (1990)	568 ± 31	1.32 (*)	430 ± 23
SS Leo	1187 ± 68	Fernley et al. (1990)	1414 ± 80	1.33 (*)	1063 ± 60
TT Lyn	566 ± 16	Liu & Janes (1990)	654 ± 36	1.32 (*)	495 ± 27
RV Oct	670 ± 20	Skillen et al. (1993a)	904 ± 20	1.33 (*)	680 ± 15
AV Peg	504 ± 14	Liu & Janes (1990)	651 ± 36	1.32 (*)	493 ± 27
BB Pup	1188 ± 85	Skillen et al. (1993a)	1328 ± 20	1.33 (*)	998 ± 15
VY Ser	595 ± 28	Fernley et al. (1990)	750 ± 90	1.33 (*)	564 ± 68
		Jones et al. (1988)	682 ± 19	1.30 (*)	525 ± 15
TU UMa	517 ± 27	Liu & Janes (1990)	621 ± 34	1.32 (*)	470 ± 26

Notes. Asterisks (*) indicate the fixed parameter for determinations from other studies.

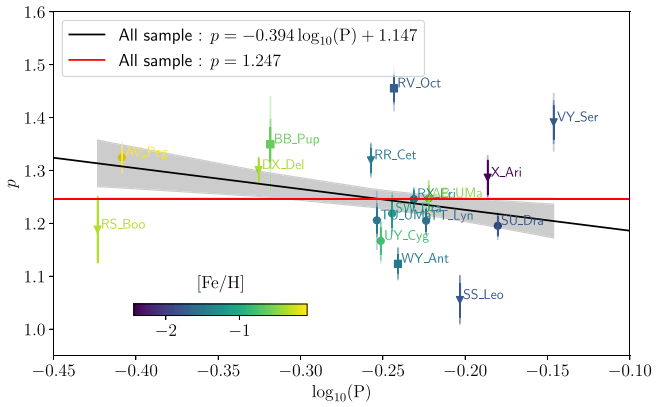


Fig. 14. p-factors for RRLs from the SPIPS model, fitted with a linear function in black and a constant in red. Points are for stars with an RV curve from the OHP exclusively. Inverted triangles represent stars for which the RV curve from the literature has been rescaled to OHP measurements. Finally, squares are the three additionnal stars for which all data are from the literature.

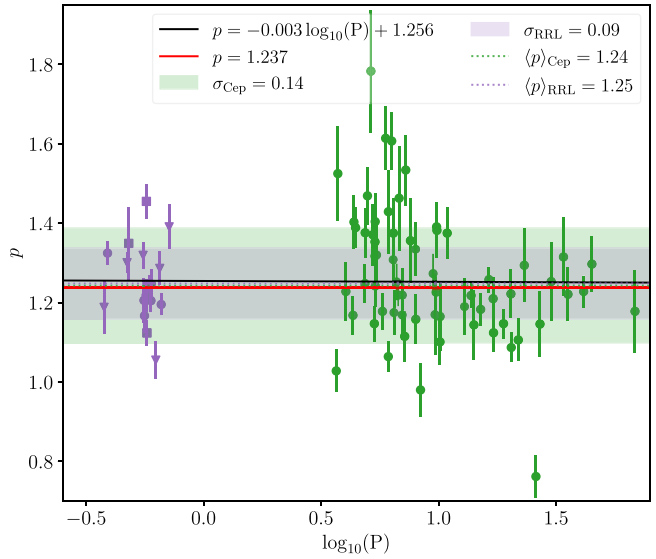


Fig. 15. Relation between the p-factor and the pulsation period for RRLs in this study and Cepheids from Trahin et al. (2021), who used the same tool for the modelling.

Another interesting relation is the period–radius relation, which is more complicated to obtain with the classical observables of a pulsating star, but is easy to reach thanks to SPIPS modelling. This relation is usually tighter than the period–luminosity relation. We show the relation in Fig. 18, the equation for which is $\log(R/R_\odot) = 0.770_{\pm 0.003} \log(P) + 0.9189_{\pm 0.0002}$. We compare with the relation found from non-linear convective models, where $\log(R/R_\odot) = 0.866_{\pm 0.003} + 0.55_{\pm 0.02} \log(P)$ (Marconi et al. 2015). This relation is within the confidence interval for our sample. These coefficients are in agreement with those that we find within the range of period containing the majority of our sample. It is truly remarkable that the period–radius relation is relatively tight. This is a clue that the conversion from photometry to a radius curve is realistic. Hence, the dispersion of the p-factors we obtain may come from a spectroscopic effect related to shocks or other specific behaviours of the atmosphere of pulsating stars.

5.3. HR diagram

Figure 19 presents the pulsation cycle of all RRLs in the Hertzsprung-Russell (HR) diagram, together with the predicted blue and red edges of the instability strip for a metallicity similar to our RRLs, as predicted by Marconi et al. (2015). We note that the RRLs do not usually cool down beyond the red edge. On the contrary, they tend to leave the strip on the blue edge during their pulsation cycle. Nevertheless, their average positions are well confined within the instability strip.

5.4. Applicability to a Gaia sample

In this paragraph, we aim to test the possibility of performing a precise calibration of the p-factor, taking advantage of the *Gaia* photometry and RV curves available in the DR3. As also reported by Clementini et al. (2023), we find some significant differences between the RV curves from ground-based spectrographs and from *Gaia*. For some stars, the *Gaia* Radial Velocity Spectrograph (RVS) curves exhibit a bump before the maximum velocity. We show an example in Fig. 20. This difference is likely due to the fact that this measurement is done by cross-correlating a template on a narrow spectrum centred on the calcium triplet. These spectra can be affected by shock waves in a specific way. Evidence of such shocks mainly comes from studies of Balmer lines (Gillet & Crowe 1988; Fokin & Gillet 1997; Gillet et al. 2017), but our knowledge of its effect on the entirety of the spectrum is still incomplete. Hocdé et al. (2020a) already showed that a desynchronisation between Hα and the calcium triplet lines may occur for long-period Cepheids, for which the surface gravity is lower. The presence of such a bump will increase the amplitude of the integrated RV curve and therefore poses a problem for the PoP method. For the star SW Dra, we estimate that the amplitude of the radius curve will differ by 15%, and the p-factor will decrease to $p = 1.04$.

All RV curves from the 23 RRLs in our sample with their *Gaia* RVS and other RV measurements from the literature are plotted in Fig. 21. To enhance readability, only one additional literature dataset is included in the plot. This dataset is either the one used in SPIPS (if available) or the one offering the best phase coverage. At least 12 (V0341 Aql, RR Cet, W CVn, UY Cyg, SW Dra, RX Eri, WZ Hya, RR Leo, SS Leo, TT Lyn, UU Vir and BN Vul) of these stars quite clearly exhibit the same feature as SW Dra. For 6 other stars, we are not able to carry out the comparison, as either *Gaia* or ground-based measurements are missing. Only DX Del appears to really show a similar shape between all curves.

It would be interesting to study the systematic uncertainties linked to the presence of this bump (if it is related to atmospheric parameters of the star or to some range of p-factors for instance), how it affects the curve, and why it appears for such spectral lines. This would require *Gaia* data release 4 (DR4), which will provide individual RVS spectra for each observing epoch.

6. Conclusions

In this paper, we present a calibration of the Baade-Wesselink projection factor for a selection of RRLs thanks to new measurements of RVs collected at OHP with the SOPHIE spectrograph. We computed the p-factor for 17 RRLs using *Gaia* DR3 parallaxes. We examined the influence of the RV measurement method on p by studying the cross-correlation of our spectra with various masks and templates. We find that the mask or template for CCF computation does not have a significant impact

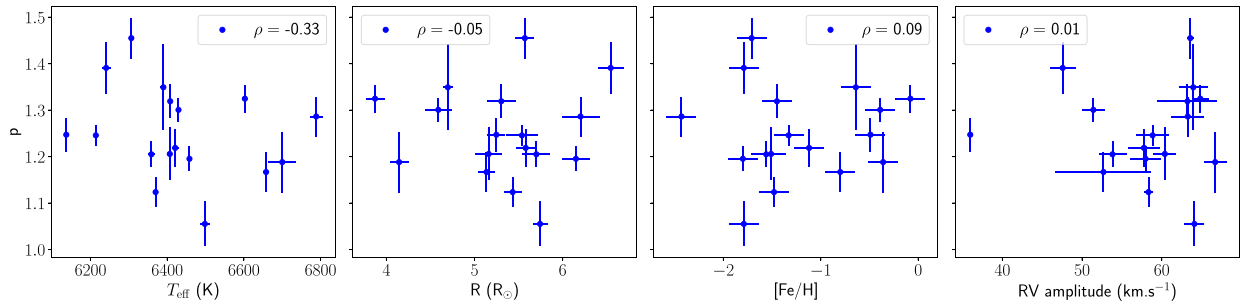


Fig. 16. Comparison between the p-factor of 17 RRLs and effective temperature, mean radius, metallicity, and amplitude of the RV curve. For each plot, the Spearman's coefficient is given to test the statistical probability of a monotonic relation between the p-factor and the other parameters.

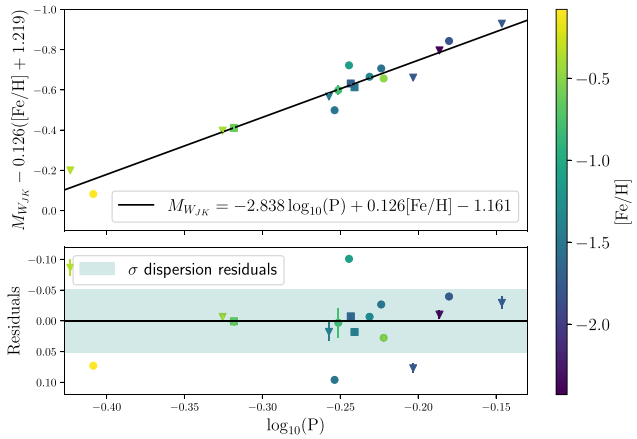


Fig. 17. Period–luminosity–metallicity relation in the Wesenheit index for the 17 RRLs of our sample. A plane has been fitted to the data and we show here the magnitude corrected for the metallicity effect versus the logarithm of the period.

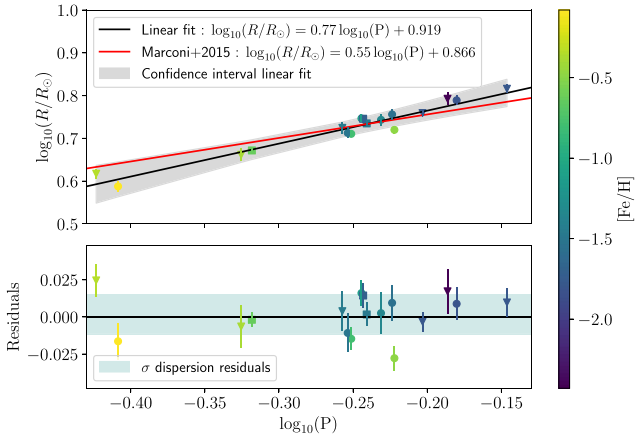


Fig. 18. Period–radius relation for the 17 RRLs of our sample. A comparison with theoretical predictions from Marconi et al. (2015) is given in red.

on p compared to the method employed to derive the RV from the CCF. Our study points out the importance of the precise documentation of how RVs are computed in the literature. However, for consistent RV series, the difference stemming from the methodology used to derive velocities should only shift the projection factor and not affect the scatter. We calibrated the p-factor using the SPIPS modelling tool. As for Cepheids (Trahin et al. 2021), we see an important scatter in the results, which is greater than the uncertainties in *Gaia* parallaxes. This result has been observed by various authors who calibrated the

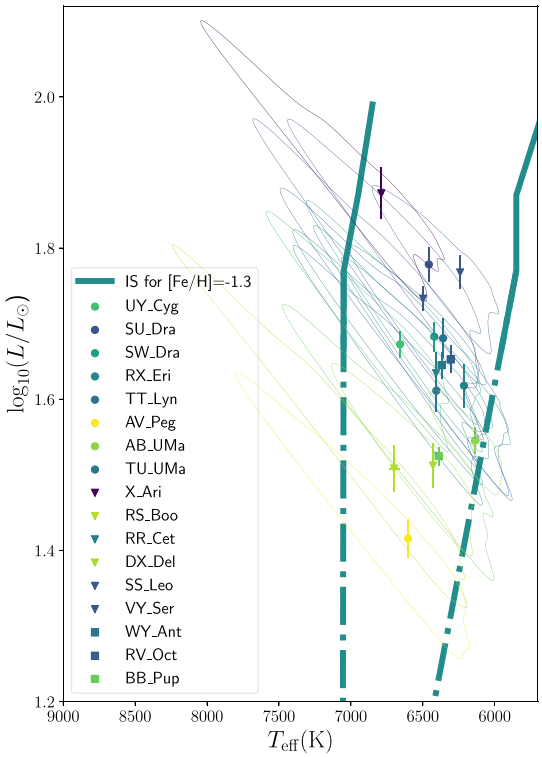


Fig. 19. Hertzsprung-Russel diagram showing the pulsation cycle of the 17 modelled RRLs of our sample. Markers indicate the type of RV measurements used in the model (see legend of Fig. 14). Thick lines are the edges of the instability strip from models (Marconi et al. 2015) for a metallicity globally corresponding to RRLs. These predictions have been extended linearly to lower luminosities in semi-dotted lines.

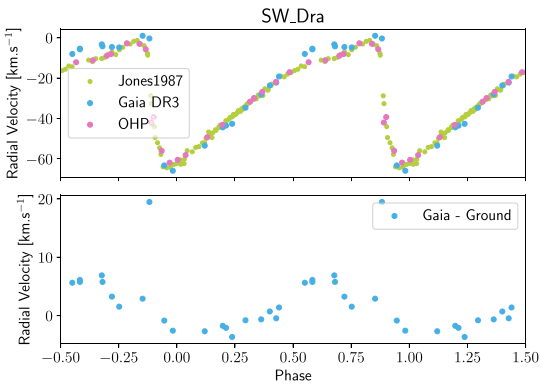


Fig. 20. Comparison between the RV curves from the ground (Jones et al. 1987b, SOPHIE) and the *Gaia* series of SW Dra. The bottom plot is the difference between the *Gaia* RV series and the total ground RV curve.

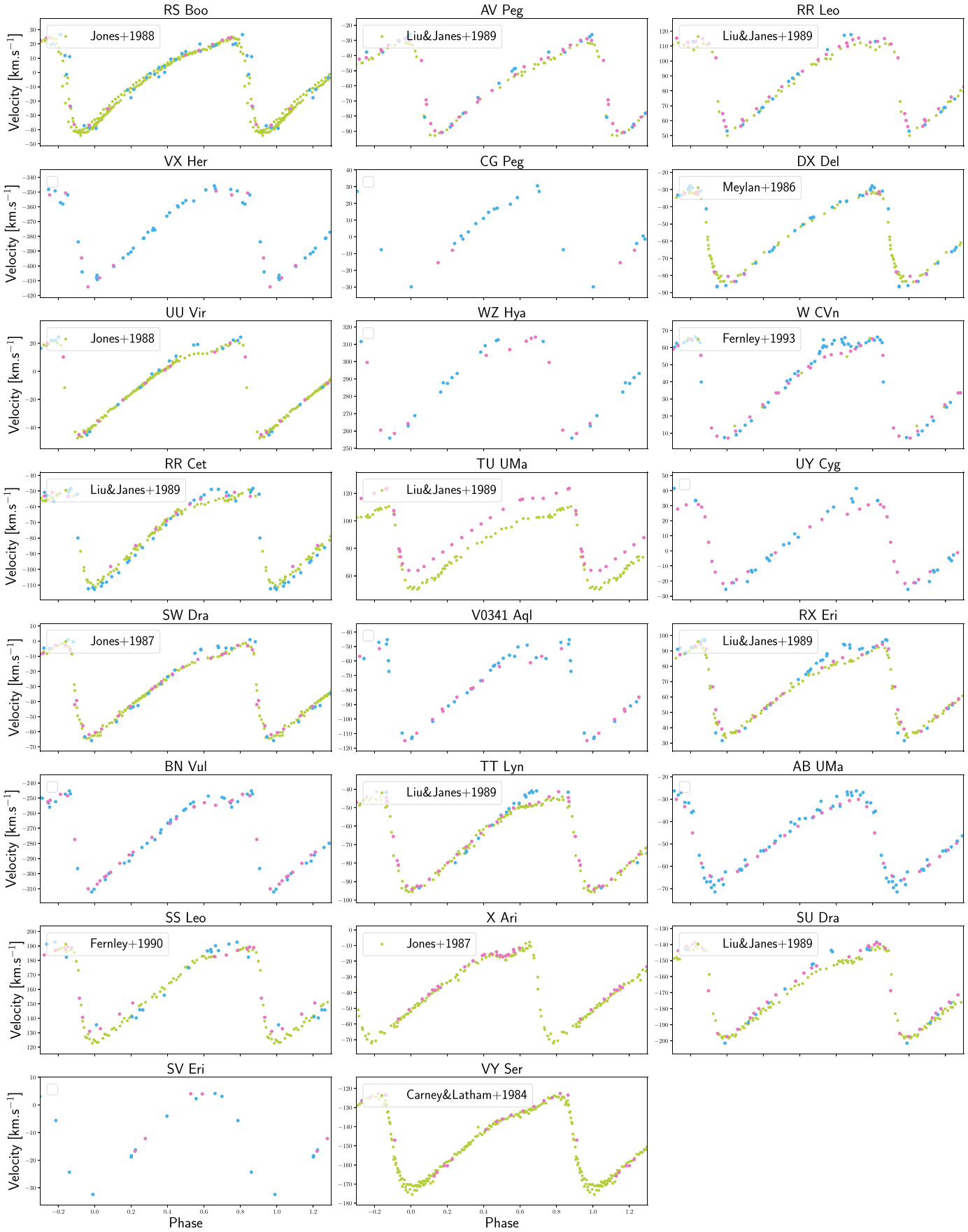


Fig. 21. The RV curves for the 23 RRLs from measurements with the SOPHIE spectrograph (this work, in pink), data from the references listed in Table 1 (in green), and *Gaia* RVS (in blue). We can see the effect of the companion of TU UMa in its RV curve.

projection factor, and does not seem to be uniquely correlated with any of the main physical parameters of the stars (period, temperature, metallicity, etc.). Interestingly, as shown in Fig. 15, the scatter of p seems to be the highest for Cepheids with short periods (below 10 days). However, whether it is a selection effect due to the size of the sample or a real phenomenon is uncertain. We note that we do not find any infrared excess for these RRLs. Although it can have an effect on the p-factors of Cepheids (Nardetto et al. 2023), the possible presence of a circumstellar envelope cannot explain the great diversity of the p-factors of pulsating stars.

Another remaining uncertainty in the SPIPS analysis is the estimation of the colour excess $E(B-V)$, which we determined from 3D maps of the interstellar medium. Although this value does not have a drastic importance for the p-factor, it is still poorly known, because of its degeneracy with the star's effective temperature. Meanwhile, RRLs in this study are generally nearby, and so the uncertainty is low for this parameter and does not represent a limitation to the present study.

The source of the scatter in the resulting p-factors is still poorly understood, but we argue that it is linked to a misunderstanding of the pulsation of the atmosphere of the star. There appears to be a missing ingredient between the pulsational velocity and the motion of line-forming regions, which may be linked to the observations of peculiar features in RVS from *Gaia* DR3.

Acknowledgements. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (http://www.cosmos.esa.int/gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, http://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. The research leading to these results has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (projects CepBin, grant agreement 695099, and UniverScale, grant agreement 951549). This research has been supported by the Polish-French Marie Skłodowska-Curie and Pierre Curie Science Prize awarded by the Foundation for Polish Science. A.G. aknowledges support from the CONICYT/FONDECYT grant No. 3130361 and ANID-ALMA fund No. ASTRO20-0059. We also acknowledge support from the DIR/WK/2018/09 grant of the Polish Ministry of education and science. This research has made use of Astropy (http://www.astropy.org/), a community-developed core Python package for Astronomy (Astropy Collaboration 2013, 2018), the Numpy library (Harris et al. 2020), the Astroquery library (Ginsburg et al. 2019) and the Matplotlib graphics environment (Hunter 2007). We used the SIMBAD and VizieR databases and catalogue access tool at the CDS, Strasbourg (France), and NASA's Astrophysics Data System Bibliographic Services. The original description of the VizieR service was published in Ochsenbein et al. (2000). This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

References

Anderson, R. I. 2016, MNRAS, 463, 1707
Anderson, R. I. 2020, in Stars and their Variability Observed from Space, eds. C. Neiner, W. W. Weiss, D. Baade, et al., 61
Andrievsky, S., Wallerstein, G., Korotin, S., et al. 2018, PASP, 130, 024201
Andrievsky, S. M., Wallerstein, G., Korotin, S. A., et al. 2020, Astron. Nachr., 341, 899
Andrievsky, S. M., Korotin, S. A., Kovtyukh, V. V., Khrapaty, S. V., & Rudyak, Y. 2021, Astron. Nachr., 342, 887
Astropy Collaboration (Robitaille, T. P., et al.) 2013, A&A, 558, A33
Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, AJ, 156, 123
Baade, W. 1926, Astron. Nachr., 228, 359
Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Demleitner, M., & Andrae, R. 2021, AJ, 161, 147
Barnes, T. G., III, Moffett, T. J., Hawley, S. L., Slovak, M. H., & Frueh, M. L. 1988, ApJS, 67, 403
Barnes, T. G., III, Moffett, T. J., & Frueh, M. L. 1992, PASP, 104, 514
Bhardwaj, A., Marconi, M., Rejkuba, M., et al. 2023, ApJ, 944, L51
Blanco-Cuaresma, S. 2019, MNRAS, 486, 2075
Blanco-Cuaresma, S., Soubiran, C., Heiter, U., & Jofré, P. 2014, A&A, 569, A111
Bono, G., Gieren, W. P., Marconi, M., Fouqué, P., & Caputo, F. 2001, ApJ, 563, 319
Borgniet, S., Kervella, P., Nardetto, N., et al. 2019, A&A, 631, A37
Breitfelder, J., Mérand, A., Kervella, P., et al. 2016, A&A, 587, A117
Breuval, L., Riess, A. G., Kervella, P., Anderson, R. I., & Romaniello, M. 2022, ApJ, 939, 89
Burki, G., & Meylan, G. 1986a, A&A, 159, 255
Burki, G., & Meylan, G. 1986b, A&A, 156, 131
Cacciari, C., Clementini, G., Prevot, L., et al. 1987, A&AS, 69, 135
Cacciari, C., Clementini, G., Prevot, L., & Buser, R. 1989, A&A, 209, 141
Capitanio, L., Lallement, R., Vergely, J. L., Elyajouri, M., & Monreal-Ibero, A. 2017, A&A, 606, A65
Carney, B. W., & Latham, D. W. 1984, ApJ, 278, 241
Casagrande, L., & VandenBerg, D. A. 2014, MNRAS, 444, 392
Castelli, F., & Kurucz, R. L. 2003, in Modelling of Stellar Atmospheres, 210, eds. N. Piskunov, W. W. Weiss, & D. F. Gray, A20
Clementini, G., Ripepi, V., Garofalo, A., et al. 2023, A&A, 674, A18
Coelho, P., Barbuy, B., Meléndez, J., Schiavon, R. P., & Castilho, B. V. 2005, A&A, 443, 735
Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., et al. 2012, Res. Astron. Astrophys., 12, 1197
Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, The IRSA 2MASS All-Sky Point Source Catalog, NASA/IPAC Infrared Science Archive
ESA 1997 in The Hipparcos and Tycho catalogues. Astrometric and photometric star catalogues derived from the ESA Hipparcos Space Astrometry Mission, ESA Special Publication, 1200
Fernley, J., & Barnes, T. G. 1997, A&AS, 125, 313
Fernley, J. A., Lynas-Gray, A. E., Skillen, I., et al. 1989, MNRAS, 236, 447
Fernley, J. A., Skillen, I., Jameson, R. F., et al. 1990, MNRAS, 247, 287
Fernley, J. A., Skillen, I., & Burki, G. 1993, A&AS, 97, 815
Fernley, J., Barnes, T. G., Skillen, I., et al. 1998a, A&A, 330, 515
Fernley, J., Skillen, I., Carney, B. W., Cacciari, C., & Janes, K. 1998b, MNRAS, 293, L61
Fitzpatrick, E. L. 1999, PASP, 111, 63
Fokin, A. B., & Gillet, D. 1997, A&A, 325, 1013
Fokin, A. B., Gillet, D., & Breitfellner, M. G. 1996, A&A, 307, 503
For, B.-Q., Preston, G. W., & Sneden, C. 2011, ApJS, 194, 38
Gaia Collaboration (Prusti, T., et al.) 2016, A&A, 595, A1
Gaia Collaboration (Vallenari, A., et al.) 2023, A&A, 674, A1
Gallenne, A., Kervella, P., Mérand, A., et al. 2017, A&A, 608, A18
Gallenne, A., Mérand, A., Kervella, P., et al. 2021, A&A, 651, A113
Gillet, D., & Crowe, R. A. 1988, A&A, 199, 242
Gillet, D., Mauclaire, B., Garrel, T., et al. 2017, A&A, 607, A51
Ginsburg, A., Sipőcz, B. M., Brasseur, C. E., et al. 2019, AJ, 157, 98
Groenewegen, M. A. T. 2018, A&A, 619, A8
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357
Hocdé, V., Nardetto, N., Borgniet, S., et al. 2020a, A&A, 641, A74
Hocdé, V., Nardetto, N., Lagadec, E., et al. 2020b, A&A, 633, A47
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Husser, T. O., Wende-von Berg, S., Dreizler, S., et al. 2013, A&A, 553, A6
Jones, R. V., Carney, B. W., Latham, D. W., & Kurucz, R. L. 1987a, ApJ, 312, 254
Jones, R. V., Carney, B. W., Latham, D. W., & Kurucz, R. L. 1987b, ApJ, 314, 605
Jones, R. V., Carney, B. W., & Latham, D. W. 1988, ApJ, 332, 206
Jönsson, H., Holtzman, J. A., Allende Prieto, C., et al. 2020, AJ, 160, 120
Jurcsik, J., Smitola, P., Hajdu, G., et al. 2017, MNRAS, 468, 1317
Kervella, P., Bersier, D., Mourard, D., et al. 2004, A&A, 428, 587
Kervella, P., Trahin, B., Bond, H. E., et al. 2017, A&A, 600, A127
Layden, A. C., Tiede, G. P., Chaboyer, B., Bunner, C., & Smitka, M. T. 2019, AJ, 158, 105
Leavitt, H. S., & Pickering, E. C. 1912, Harvard Coll. Observatory Circ., 173, 1
Lindegren, L., Bastian, U., Biermann, M., et al. 2021, A&A, 649, A4
Lindemann, F. A. 1918, MNRAS, 78, 639
Liška, J., Skarka, M., Mikulášek, Z., Zejda, M., & Chrastina, M. 2016, A&A, 589, A94
Liu, T., & Janes, K. A. 1989, ApJS, 69, 593
Liu, T., & Janes, K. A. 1990, ApJ, 354, 273
Marconi, M., Coppola, G., Bono, G., et al. 2015, ApJ, 808, 50
Mérand, A., Kervella, P., Breitfelder, J., et al. 2015, A&A, 584, A80
Meylan, G., Burki, G., Rufener, F., et al. 1986, A&AS, 64, 25
Monson, A. J., Beaton, R. L., Scowcroft, V., et al. 2017, AJ, 153, 96
Nardetto, N., Fokin, A., Mourard, D., et al. 2004, A&A, 428, 131
Nardetto, N., Mourard, D., Kervella, P., et al. 2006, A&A, 453, 309

Nardetto, N., Mourard, D., Mathias, P., Fokin, A., & Gillet, D. 2007, A&A, 471, 661
Nardetto, N., Poretti, E., Rainer, M., et al. 2014, A&A, 561, A151
Nardetto, N., Poretti, E., Rainer, M., et al. 2017, A&A, 597, A73
Nardetto, N., Gieren, W., Storm, J., et al. 2023, A&A, 671, A14
Neilson, H. R., & Lester, J. B. 2008, ApJ, 684, 569
Ngeow, C. C., Neilson, H. R., Nardetto, N., & Marengo, M. 2012, A&A, 543, A55
Ochsenbein, F., Bauer, P., & Marcout, J. 2000, A&AS, 143, 23
Paczyński, B. 1965, Acta Astron., 15, 103
Paczyński, B. 1966, Acta Astron., 16, 97
Perruchot, S., Bouchy, F., Chazelas, B., et al. 2011, SPIE Conf. Ser., 8151, 815115
Petterson, O. K. L., Cottrell, P. L., Albrow, M. D., & Fokin, A. 2005, MNRAS, 362, 1167
Pietrzyński, G., Graczyk, D., Gallenne, A., et al. 2019, Nature, 567, 200
Pilecki, B., Gieren, W., Smolec, R., et al. 2017, ApJ, 842, 110
Planck Collaboration VI. 2020, A&A, 641, A6
Preston, G. W., & Paczynski, B. 1964, ApJ, 140, 181
Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. 2019, ApJ, 876, 85
Riess, A. G., Yuan, W., Macri, L. M., et al. 2022, ApJ, 934, L7
Skillen, I., Fernley, J. A., Jameson, R. F., Lynas-Gray, A. E., & Longmore, A. J. 1989, MNRAS, 241, 281
Skillen, I., Fernley, J. A., Stobie, R. S., Marang, F., & Jameson, R. F. 1993a, South African Astron. Obs. Circ., 15, 90
Skillen, I., Fernley, J. A., Stobie, R. S., & Jameson, R. F. 1993b, MNRAS, 265, 301
Storm, J., Gieren, W., Fouqué, P., et al. 2011, A&A, 534, A94
Sturch, C. 1966, ApJ, 143, 774
Takeda, Y. 2022, MNRAS, 514, 2450
Takeda, Y., Honda, S., Aoki, W., et al. 2006, PASJ, 58, 389
Trahin, B., Breuval, L., Kervella, P., et al. 2021, A&A, 656, A102
Varsavsky, C. M. 1960, ApJ, 131, 623
Wesselink, A. J. 1946, Bull. Astron. Inst. Netherlands, 10, 88

Appendix A: SPIPS models

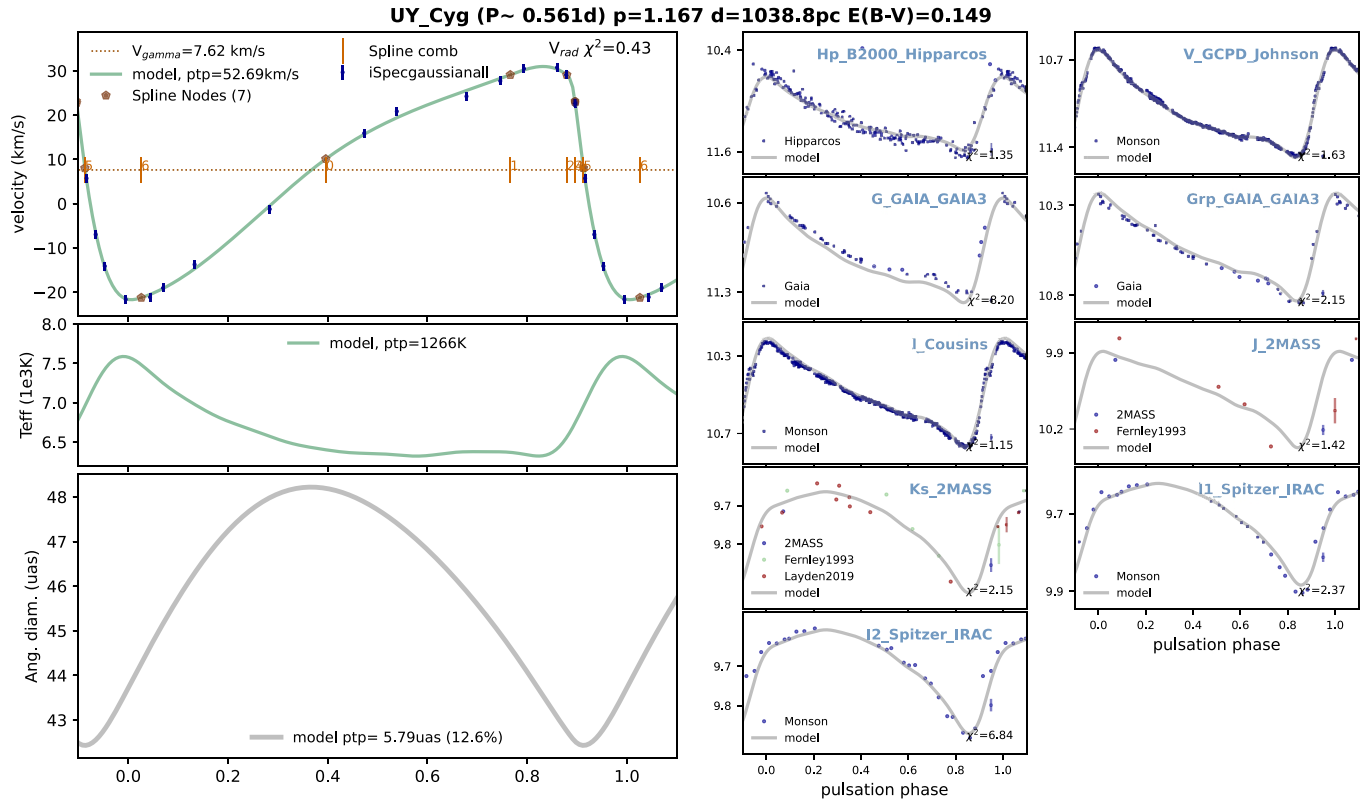


Fig. A.1. SPIPS model of UY Cyg with RV from SOPHIE spectra (Gaussian fit of the CCF using the binary mask built for RRLs with all unblended lines). Photometry measurements are from Monson et al. (2017), Fernley et al. (1993), and Layden et al. (2019) (70% of available photometric measurements).

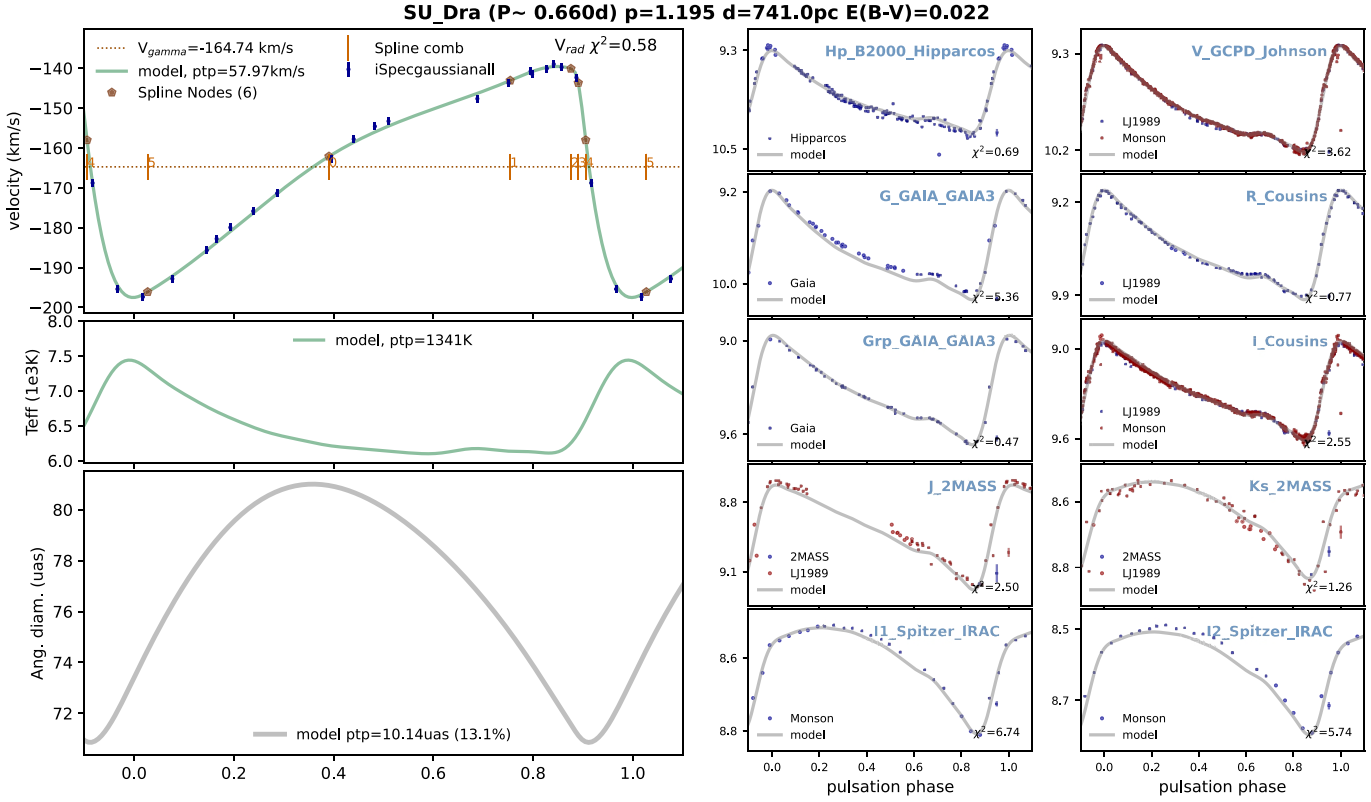


Fig. A.2. SPIPS model of SU Dra with RV from SOPHIE spectra (Gaussian fit of the CCF using the binary mask built for RRLs with all unblended lines). Photometry measurements are from Monson et al. (2017) and Liu & Janes (1989) (70% of available photometric measurements).

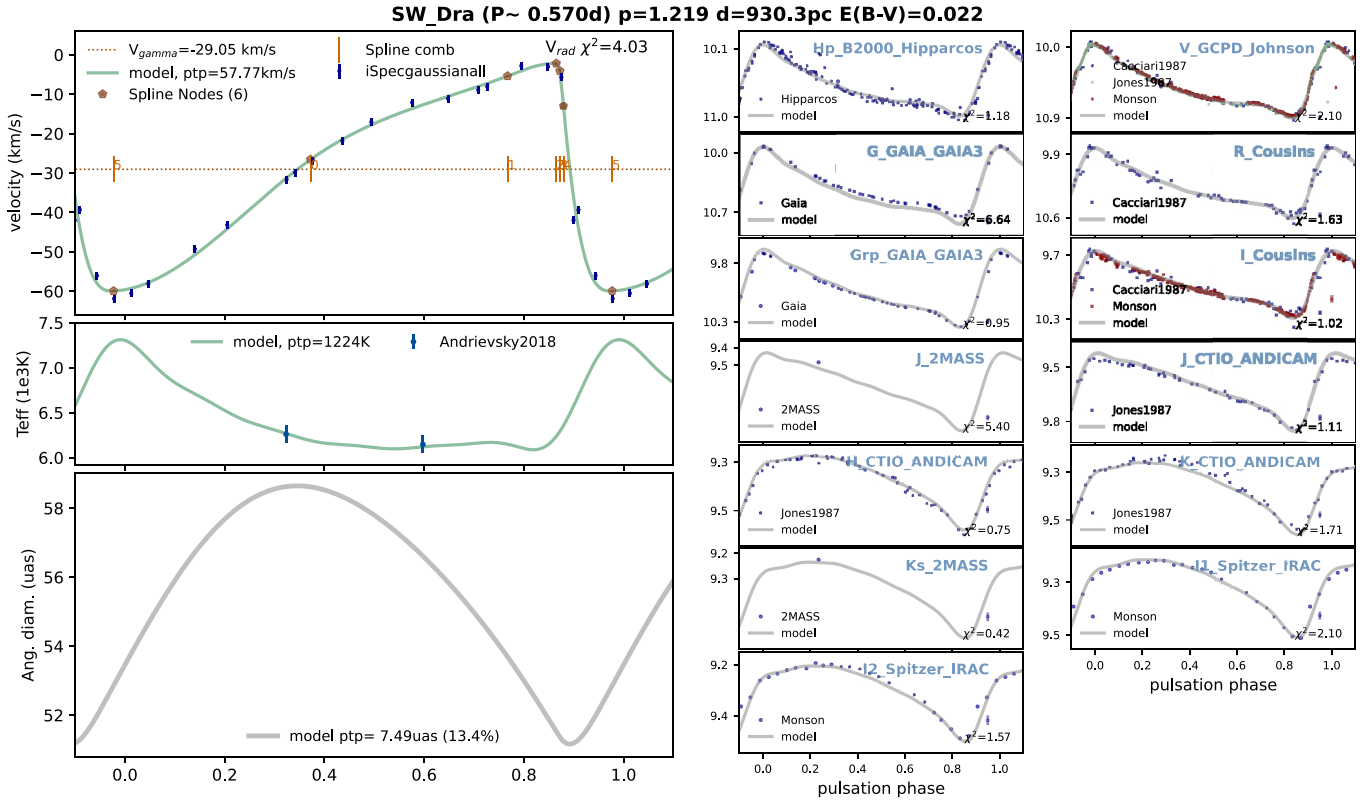


Fig. A.3. SPIPS model of SW Dra with RV from SOPHIE spectra (Gaussian fit of the CCF using the binary mask built for RRLs with all unblended lines). Photometry measurements are from Monson et al. (2017), Cacciari et al. (1987), and Jones et al. (1987b) (68% of available photometric measurements).

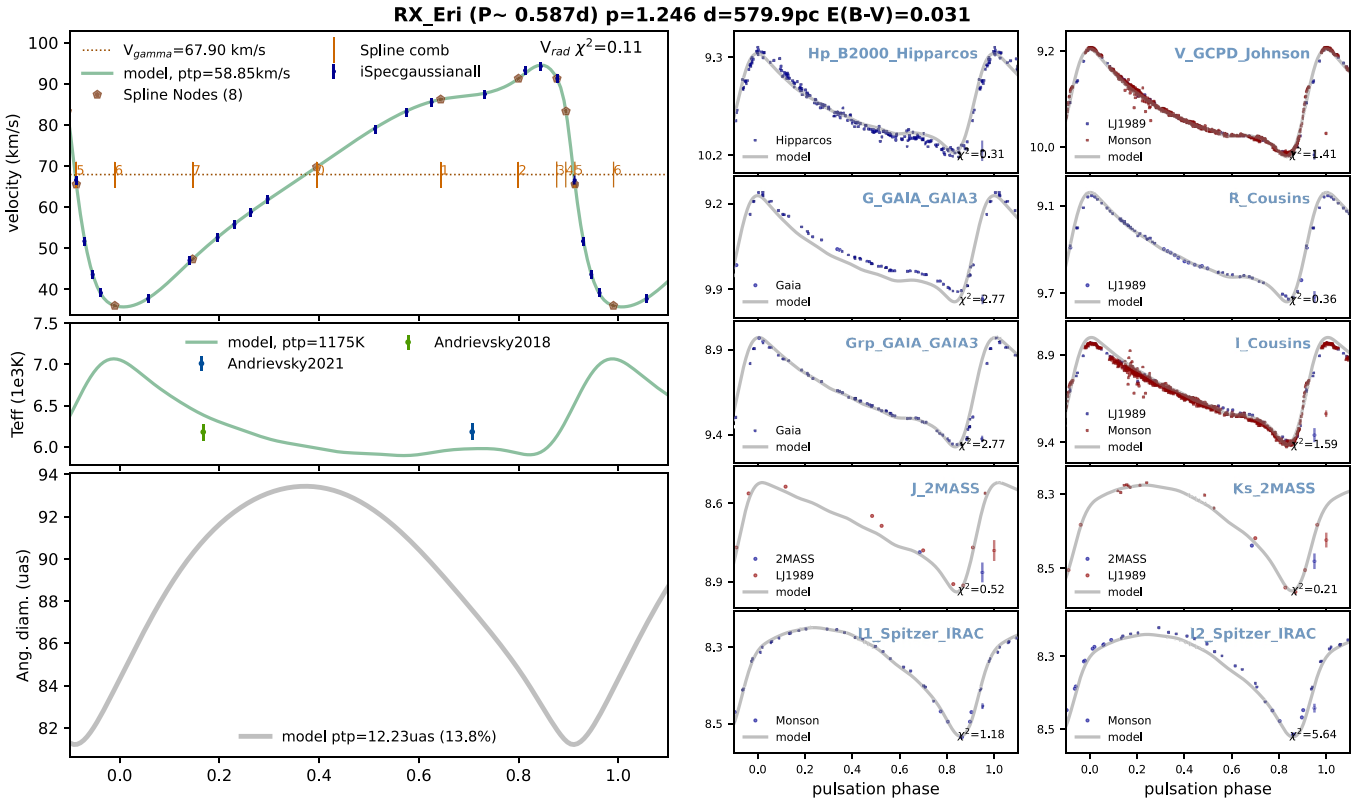


Fig. A.4. SPIPS model of RX Eri with RV from SOPHIE spectra (Gaussian fit of the CCF using the binary mask built for RRLs with all unblended lines). Photometry measurements are from Monson et al. (2017) and Liu & Janes (1989) (64% of available photometric measurements).

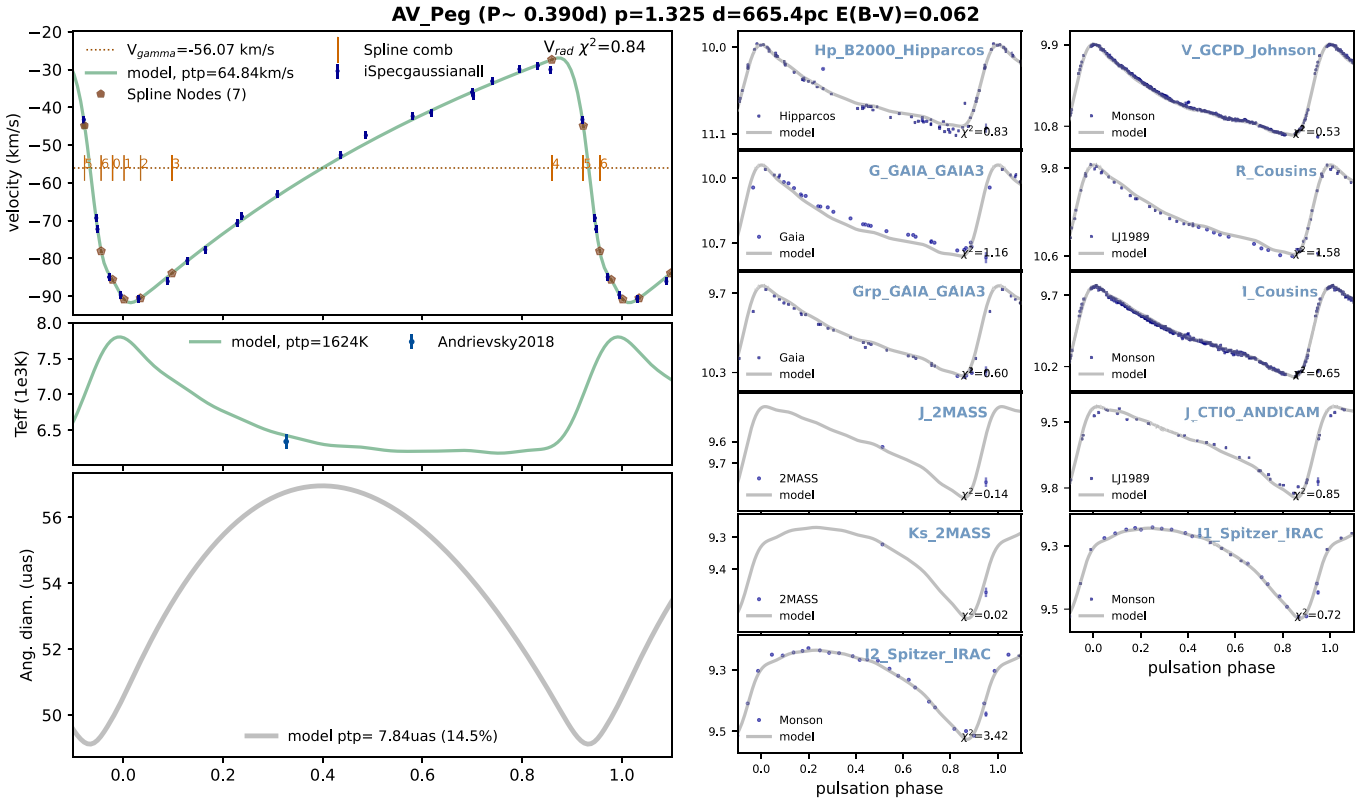


Fig. A.5. SPIPS model of AV Peg with RV from SOPHIE spectra (Gaussian fit of the CCF using the binary mask built for RRLs with all unblended lines). Photometry measurements are from Monson et al. (2017) and Liu & Janes (1989) (56% of available photometric measurements).

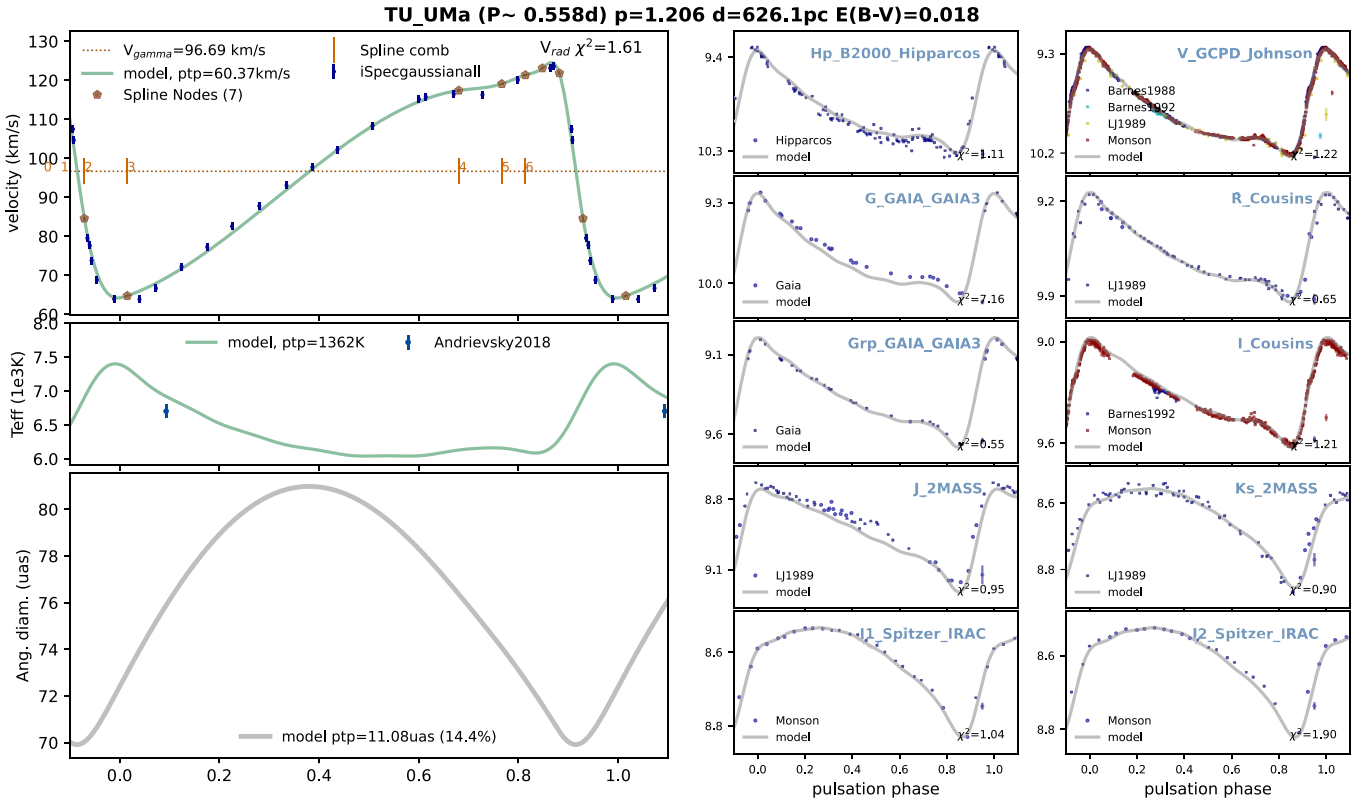


Fig. A.6. SPIPS model of TU UMa with RV from SOPHIE spectra (Gaussian fit of the CCF using the binary mask built for RRLs with all unblended lines). Photometry measurements are from Monson et al. (2017), Liu & Janes (1989), Barnes et al. (1988), and Barnes et al. (1992) (42% of available photometric measurements).

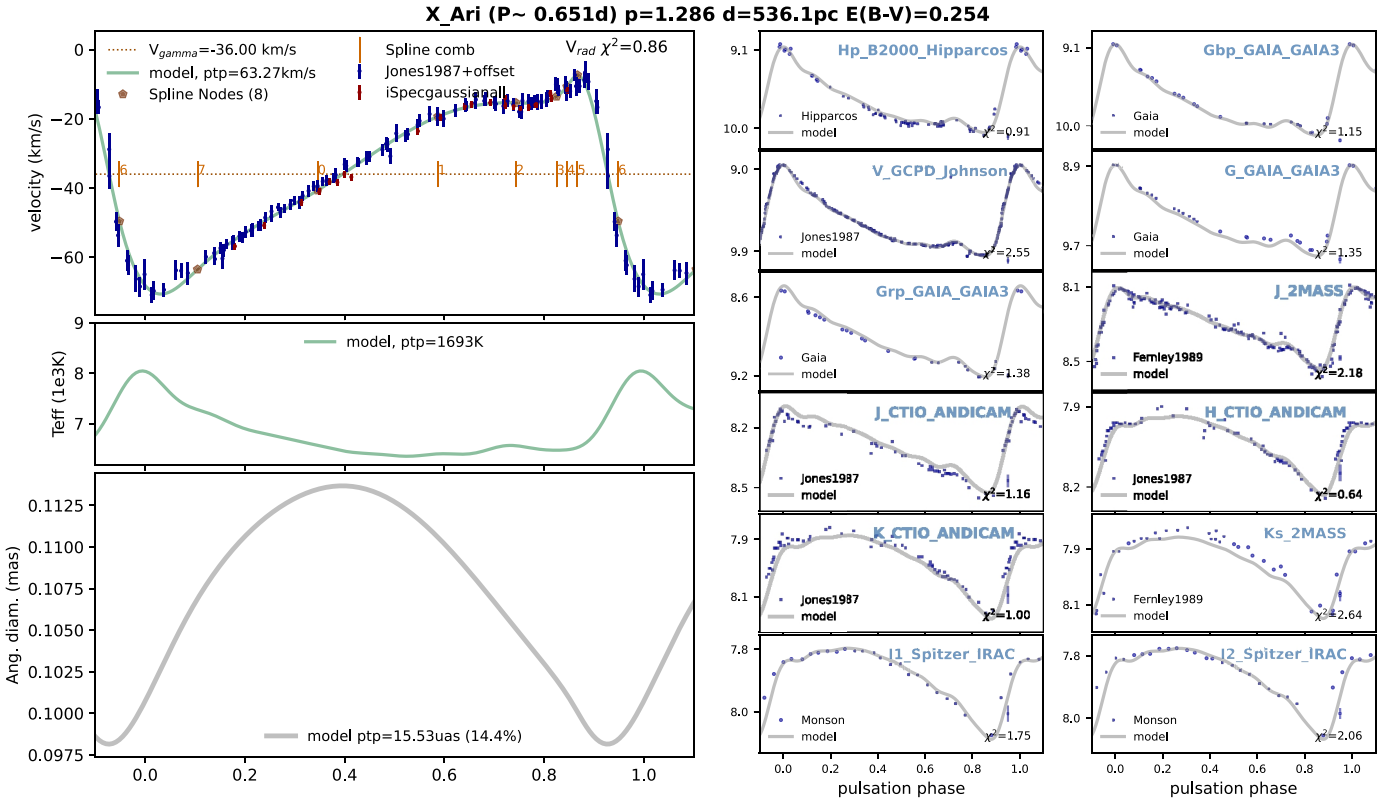


Fig. A.7. SPIPS model of X Ari with RV from Jones et al. (1987a) scaled to RV measurements (Gaussian fit of the CCF using the binary mask built for RRLs with all unblended lines), offset=1.285013 km.s^{-1}. Photometry measurements are from Monson et al. (2017), Jones et al. (1987a), and Fernley et al. (1989) (23% of available photometric measurements).

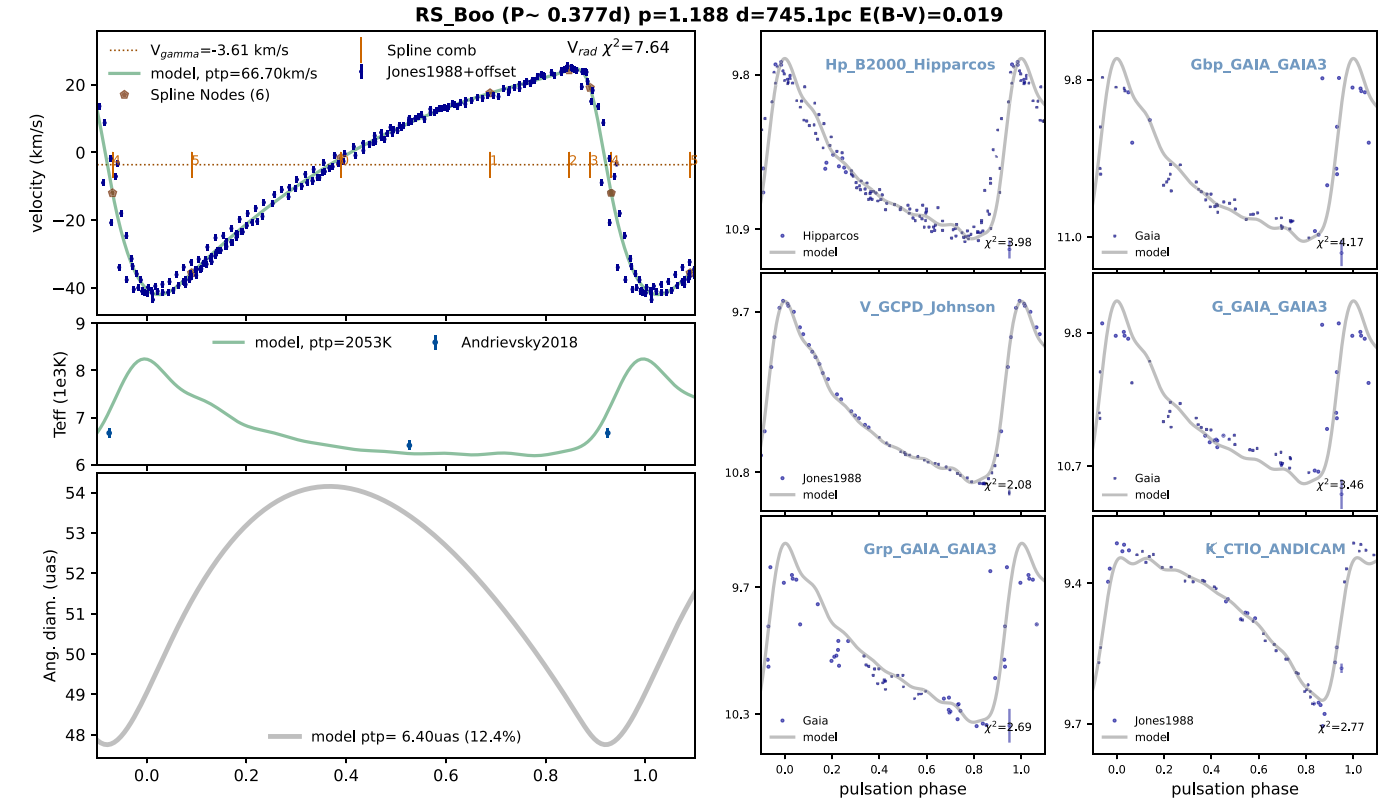


Fig. A.8. SPIPS model of RS Boo with RV from Jones et al. (1988) scaled to RV measurements (Gaussian fit of the CCF using the binary mask built for RRLs with all unblended lines), offset=0.64241 km.s^{-1}. Photometry measurements are from Jones et al. (1988) (87% of available photometric measurements).

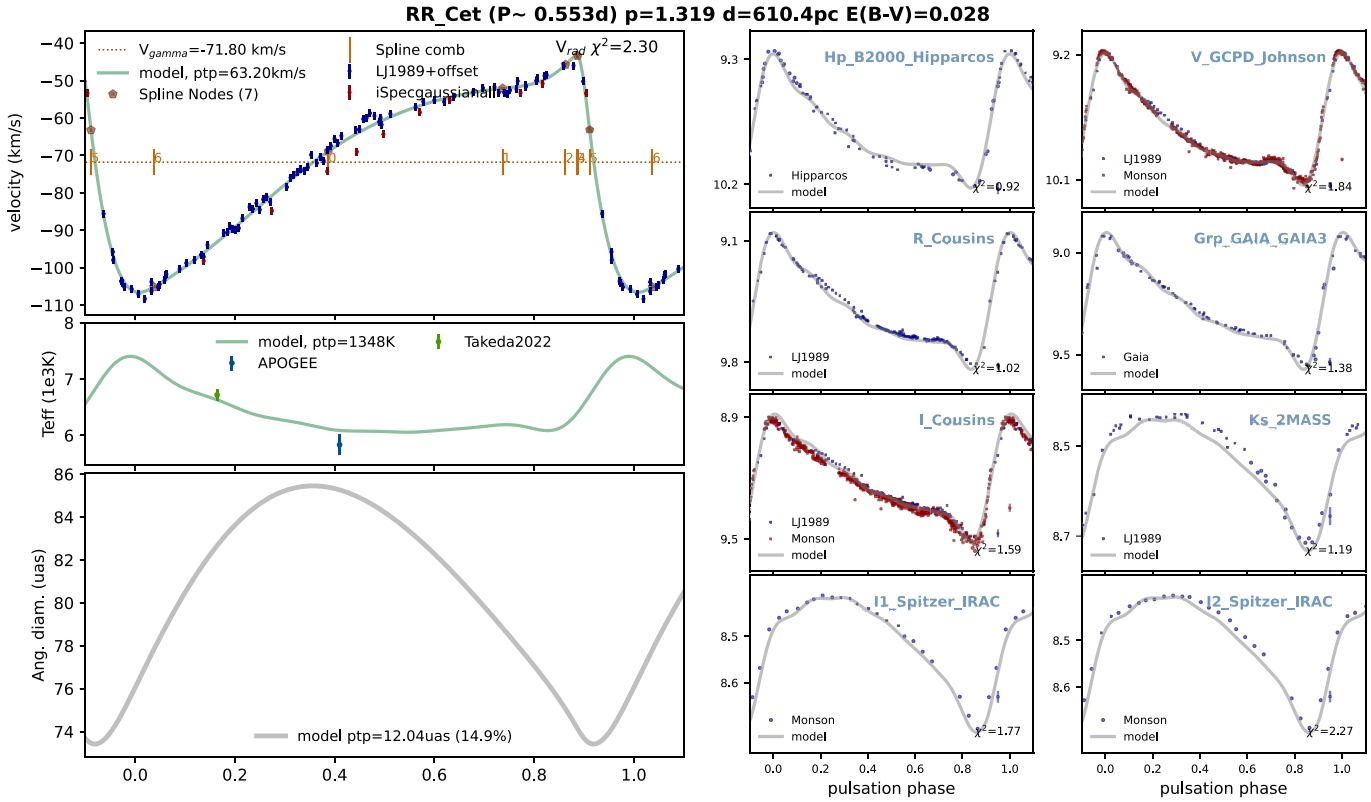


Fig. A.9. SPIPS model of RR Cet with RV from Liu & Janes (1989) scaled to RV measurements (Gaussian fit of the CCF using the binary mask built for RRLs with all unblended lines), offset=2.952771 km.s^{-1}. Photometry measurements are from Monson et al. (2017) and Liu & Janes (1989) (21% of available photometric measurements).

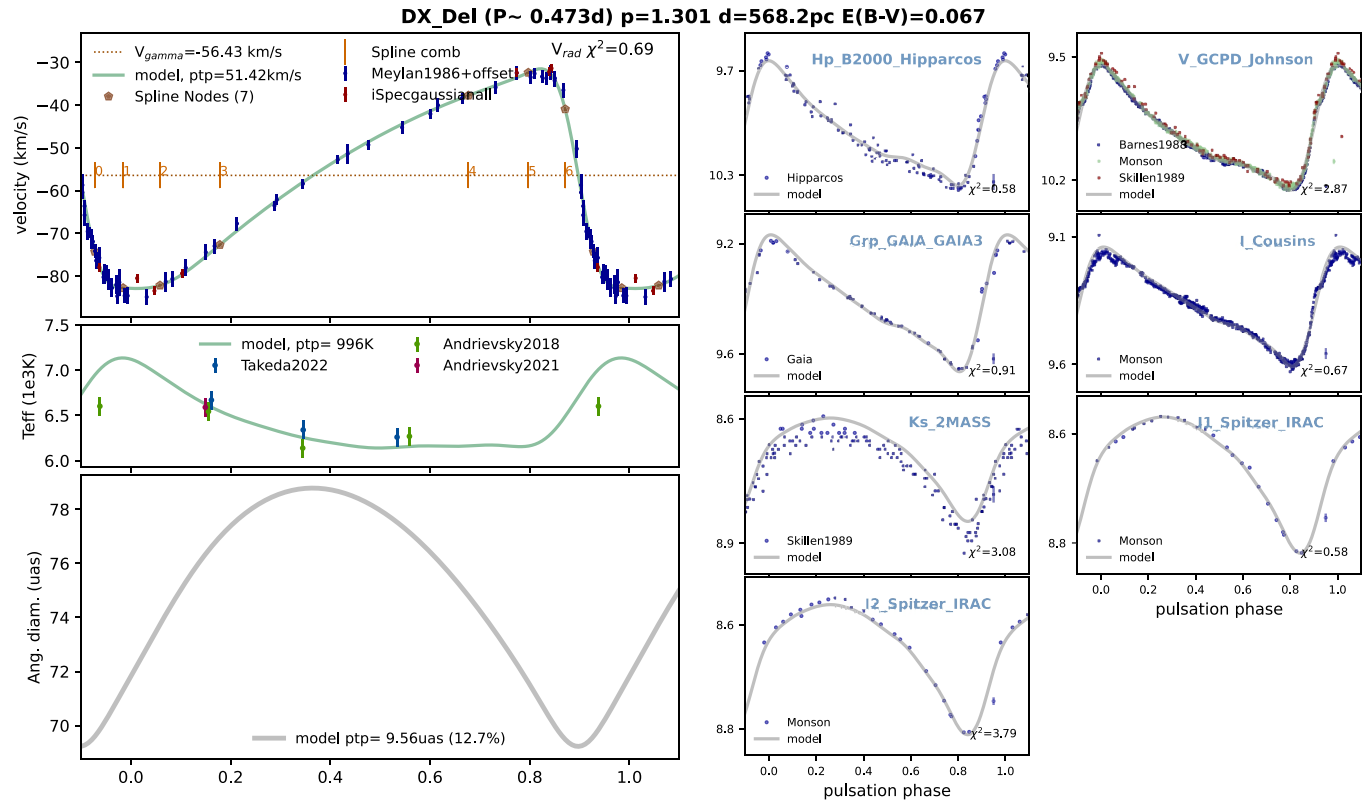


Fig. A.10. SPIPS model of DX Del with RV from Meylan et al. (1986) scaled to RV measurements (Gaussian fit of the CCF using the binary mask built for RRLs with all unblended lines), offset=-1.173494 km.s^{-1}. Photometry measurements are from Monson et al. (2017), Barnes et al. (1988), and Skillen et al. (1989) (27% of available photometric measurements).

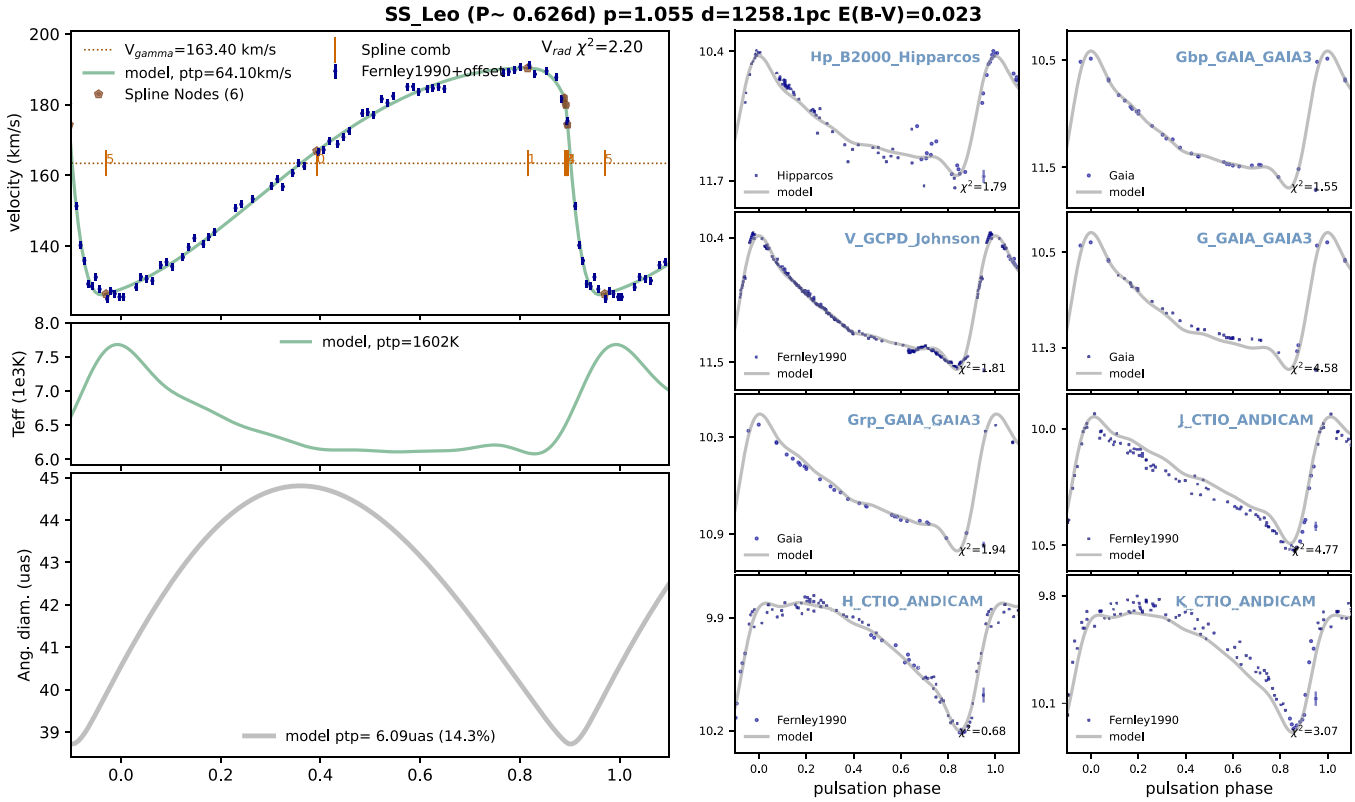


Fig. A.11. SPIPS model of SS Leo with RV from Fernley et al. (1990) scaled to RV measurements (Gaussian fit of the CCF using the binary mask built for RRLs with all unblended lines), offset=2.219878 km.s^{-1}. Photometry measurements are from Fernley et al. (1990) (44% of available photometric measurements).

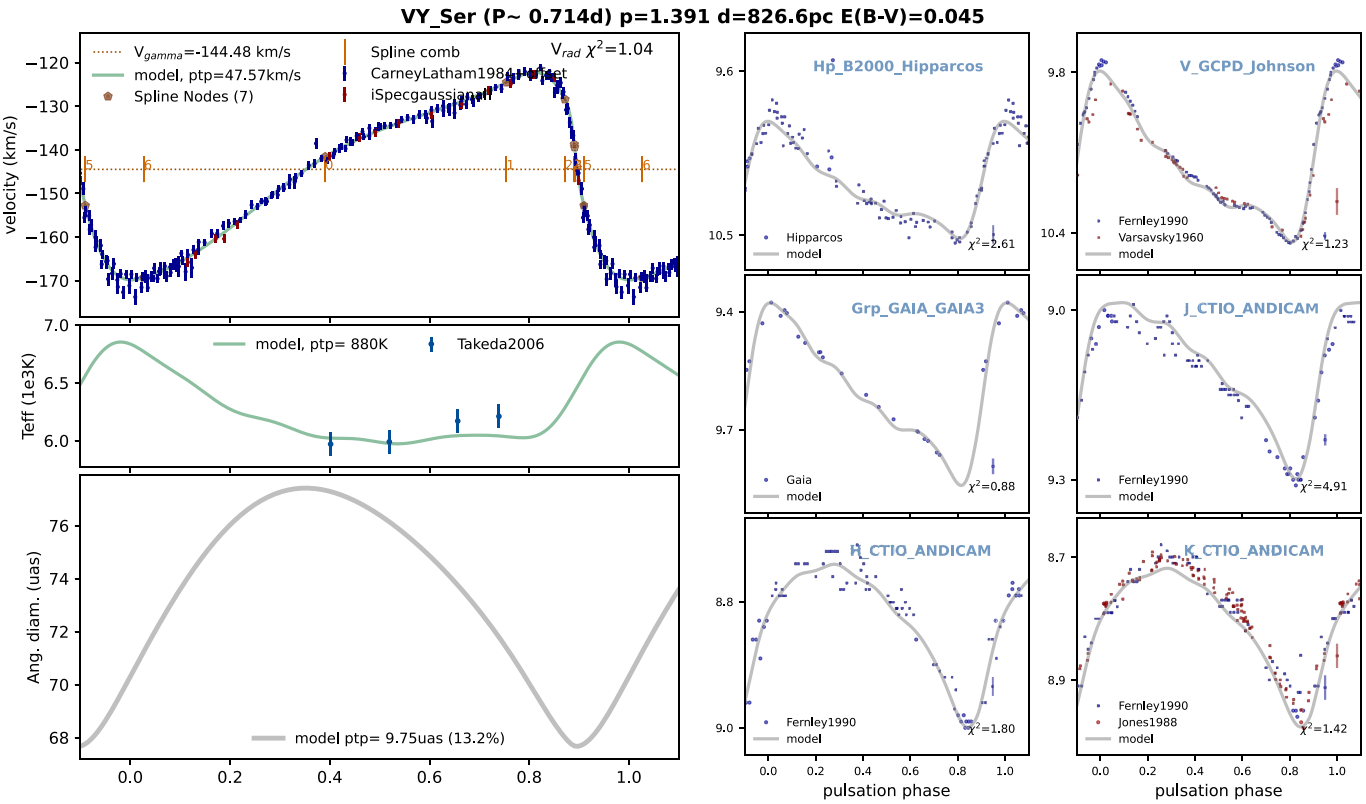


Fig. A.12. SPIPS model of VY Ser with RV from Carney & Latham (1984) scaled to RV measurements (Gaussian fit of the CCF using the binary mask built for RRLs with all unblended lines), offset=1.658486 km.s^{-1}. Photometry measurements are from Fernley et al. (1990) and Varsavsky (1960) (41% of available photometric measurements).

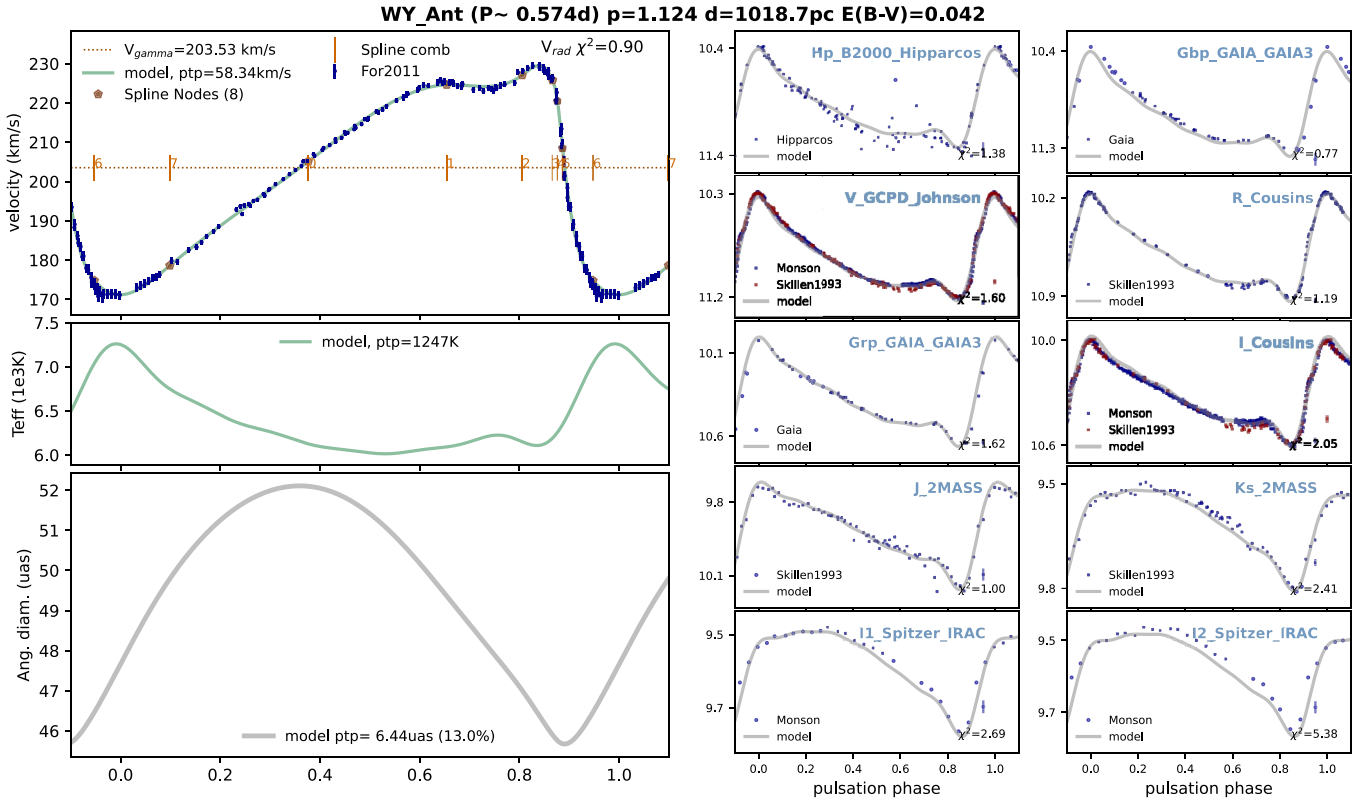


Fig. A.13. SPIPS model of WY Ant with RV from For et al. (2011) (not observed at OHP). Photometry measurements are from Monson et al. (2017) and Skillen et al. (1993a,b) (66% of available photometric measurements).

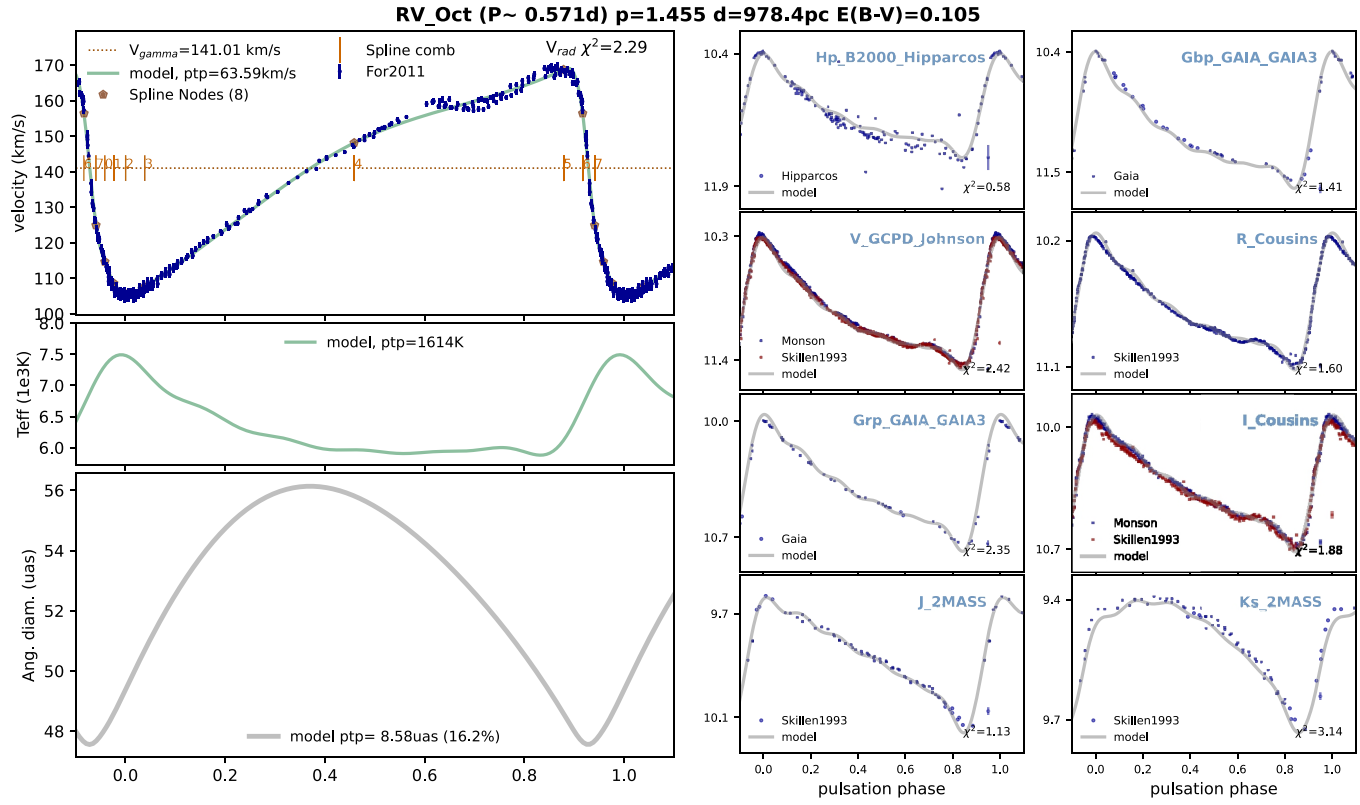


Fig. A.14. SPIPS model of RV Oct with RV from For et al. (2011) (not observed at OHP). Photometry measurements are from Monson et al. (2017) and Skillen et al. (1993a,b) (80% of available photometric measurements).

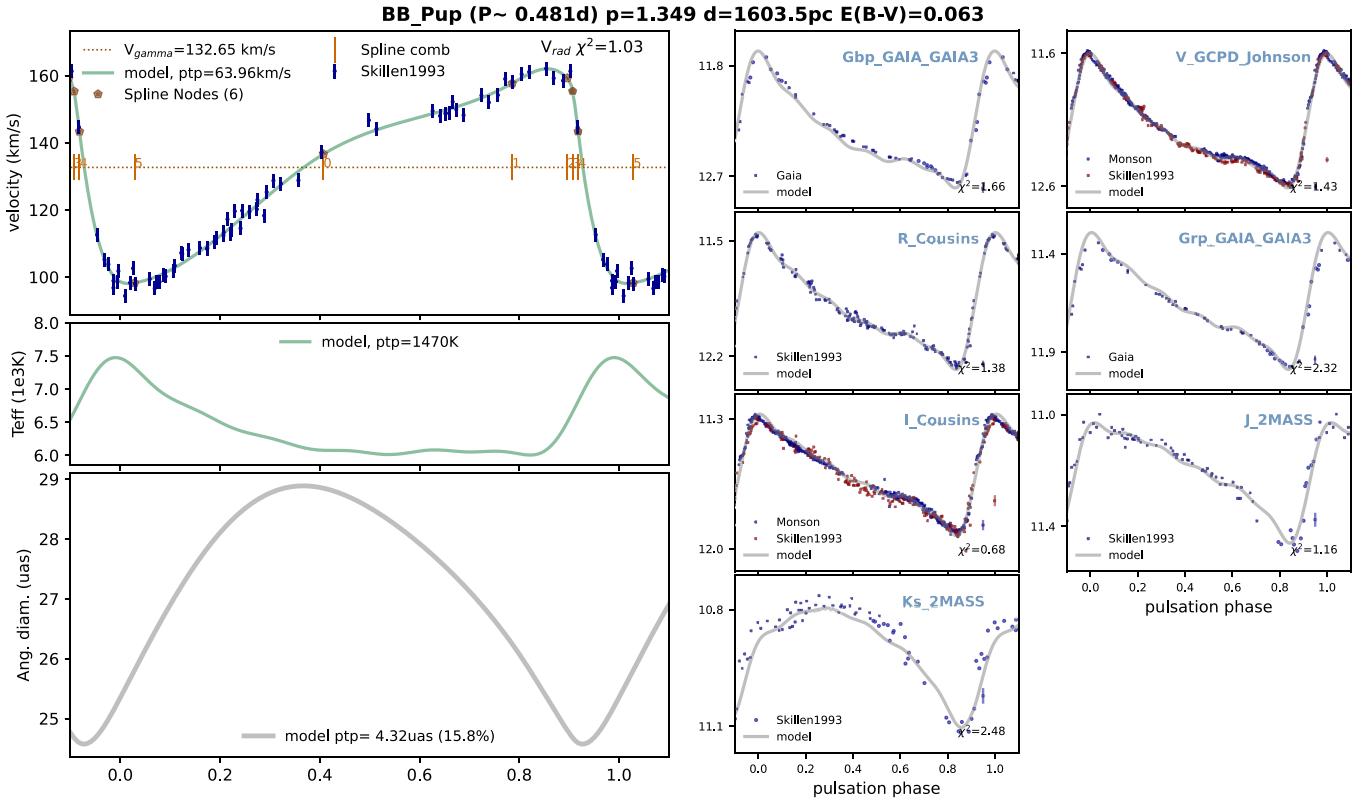


Fig. A.15. SPIPS model of BB Pup with RV from Skillen et al. (1993a,b) (not observed at OHP). Photometry measurements are from Monson et al. (2017) and Skillen et al. (1993a,b) (64% of available photometric measurements).

6 Possible decoupled motion between the photosphere and the atmosphere

Many studies (see references in Fig. 1.6) produced calibrations of the p-factor of pulsating stars, with the objective to find a single value applicable to all stars, but without success. From the most recent studies, it seems that the dispersion of the p-factor values does not correlate to the sample size or the heterogeneity of observations. It is now probable that we miss a physical ingredient in the description of this factor, that traces the comparison of line forming regions motion with the photospheric pulsation. An insight in how the photosphere reacts to the pulsation may be a path to understand the cause of the discrepancy.

6.1 Interferometric observations of η Aql

Observationally, it is possible to directly monitor the photospheric motion thanks to interferometry. However, these observations are quite demanding. It requires angularly large targets and cutting-edge instruments. I conducted an observing program in this direction focused on the Cepheid η Aql with the CHARA (Center for High Angular Resolution Astronomy) array.

The CHARA (ten Brummelaar et al. 2005) interferometric array is located in the Mount Wilson Observatory, near Los Angeles, in California. It is composed of six 1-metre telescopes separated by various distances ranging from 34 to 331m. Light from the six telescopes is collected simultaneously and gathered in the beam combining lab. Several beam combiners are available for different science cases. We used the MIRC-X (Michigan InfraRed Combiner-eXeter, Anugu et al. 2020; Monnier et al. 2004) and MYSTIC (Michigan Young STar Imager at CHARA, Monnier et al. 2018; Setterholm et al. 2023) beam combiners that can be operated simultaneously, observing in the J and K-band respectively. The light from each telescope is conducted along adaptable delay lines. Consisting of moving carts, the goal is to find the right position for each cart, so the optical path difference between each pair of telescopes is compensated by the delay lines. At this exact position, the combination of light will produce a constructive interference pattern, traditionally called interference fringes. When six telescope are used, the fringes have to be found for each pair of telescopes, for a total of 15 interference patterns. The six telescopes are distributed in a Y-shape configuration, with two telescopes in the direction of the south (S1, S2), spaced by 34m (shortest baseline), two telescopes in the direction of the east (E1, E2), spaced by 66m and two telescopes in the direction of the west (W1, W2), spaced by 108m. The different lengths between each pair of telescope are important to improve the coverage of spatial frequencies and hence the visibility curve. For this work, we need to sufficiently resolve the star to measure the fringe visibility in the second lobe, to have access to the information on the Limb-Darkening (LD). This is possible for the Cepheid η Aql, with its angular size around 1.7 mas, considering the baseline lengths of CHARA.

The observing run was scheduled in July 2024, with 5 half-nights allocated on 17, 19, 21, 23 and 25, in order to cover the pulsation cycle of η Aql. Besides the science observations, different

calibrators were observed. They correspond to smaller stars, near the science target, for which the angular size is well-known. In our case, the choice of calibrators was particularly difficult as η Aql is extremely bright and science and calibrators need to be observed with approximately the same exposure parameters. Our selection includes rather large calibrators, which is affecting the precision of the calibration. I had the opportunity to go to the observing site to acquire myself the data. Alexandre Gallenne came with me for the first half of the run and taught me how interferometric observations are done. Over the five nights of observations, we benefited from good weather conditions. During science observations, the shutter was half-closed to keep the same gain as for calibrators. The observations consisted of 2×5 minutes of observation, in cycles of 30 minutes including the fringe search and necessary additional data (shutters, sky, darks). The key step of interferometric observations is to find where the interference occurs. After the target is centred in the field, the search for the fringes can start. The way to do is to move the carts until fringes appear for a pair of telescopes. Then, the difference between the two corresponding carts is locked to keep the same delay, but moves compared to the other telescopes to find the other interferences.

I performed the reduction using the reduction pipeline, on the remote machine in Atlanta. As our target is so bright, it was possible to set a very high standard for the selection of data to remove poor observations, when small clouds pass above one of the telescope for instance. The first quick reduction showed encouraging visibility curves, even if we could identify some bad data. I reduced a second time with more appropriate parameters in order to avoid these bad data to be kept. The calibration was done using the IDL code written by John Monnier, as he advised to do. A cleaning for outliers is included in the procedure, and is supplemented with visual inspection for each individual baseline and triangle.

6.2 The limits of calibration

We removed one of our calibrators, which is the largest one we observed, as its visibility curve is reaching zero at the higher spatial frequencies observed, leading to high discrepancy on the data calibrated with this calibrator compared to the other ones. At first, the calibrator diameters were taken from the JMMC (Jean-Marie Mariotti Centre) SearchCal tool (Chelli et al. 2016), using photometry-based methods. The consecutive analysis showed beautiful results including unexpected limb-brightening. As we were working on a publication of these results, John Monnier warned us about the uncertainties on the calibrator diameters. After several exchanges, it turned out that the error bars on the calibrators sizes were not properly propagated to the visibility curves. I present here the refined analysis of the data from the CHARA interferometer, with an appropriate treatment of systematic uncertainties.

6.2.1 Secondary calibration

To exploit our data, it is then necessary to have good control on the calibrators sizes. The diameters in the SearchCal catalogue have uncertainties of around 10% of their diameter, and

on top of that, are quite large. This induces large uncertainties on the higher spatial frequencies, which is the region of interest for the LD. To tackle this issue, we took advantage of the snapshot imaging mode of CHARA to perform secondary calibrations of our calibrators, and improve the precision of their angular diameter. The visibilities were fitted using a simple Uniform Disk (UD) model. These new data improved the precision of calibrator diameters up to 2%. This is a nice improvement but is still large for our purpose. To refine again the uncertainty, we asked for GRAVITY time (GRAVITY Collaboration et al. 2017), taking advantage of the Director's Discretionary Time (DDT). Indeed, when it comes to angular diameter measurements, GRAVITY is much more precise than MIRC-X/MYSTIC instruments. Also, with the largest configuration, we can reach high-precision visibilities reaching 50%, allowing us to have good constraint on the angular diameters. The data obtained could not be used, however, because the visibility curves are unusual. This may result from a binary secondary calibrator. In parallel, Wilma Kiviaho, currently PhD student in the team, is an expert on SED fitting. Using photometry on the different calibrators, she was able to find independent determinations of the angular diameters of the calibrators. She used photometry from Gaia DR3, 2MASS and AllWISE. The SED is fitted using atmosheric models from Castelli et al. (2003) with the appropriate parameters. The surface gravity is fixed from the JMMC determinations, as it has a minor effect in the fit, and the temperature to Gaia determinations to avoid the degeneracy with the reddening. The reddening law from Gordon et al. (2023) is adopted. The obtained diameters are limb-darkened and can be converted to UD diameters in H and K bands using Claret et al. (2011) tables.

The Table. 2.3 summarises the angular diameters of the calibrators from the various setups, compared in Fig. 2.10.

Calibrator	JMMC, H	JMMC, K	Aug25, H	Aug25, K	SED fitting
HD189322	$1.043_{\pm 0.270}$	$1.050_{\pm 0.272}$	$0.989_{\pm 0.014}$	$0.993_{\pm 0.034}$	$1.069_{\pm 0.028}$
HD191067	$0.969_{\pm 0.065}$	$0.976_{\pm 0.066}$	$0.889_{\pm 0.033}$	$0.869_{\pm 0.041}$	$0.958_{\pm 0.022}$
HD181907	$0.984_{\pm 0.087}$	$0.991_{\pm 0.088}$	$1.039_{\pm 0.024}$	$1.048_{\pm 0.057}$	$1.025_{\pm 0.038}$

Table 2.3: Calibrator sizes from the JMMC SearchCal tool, the snapshot program at CHARA in August 2025, and SED fitting.

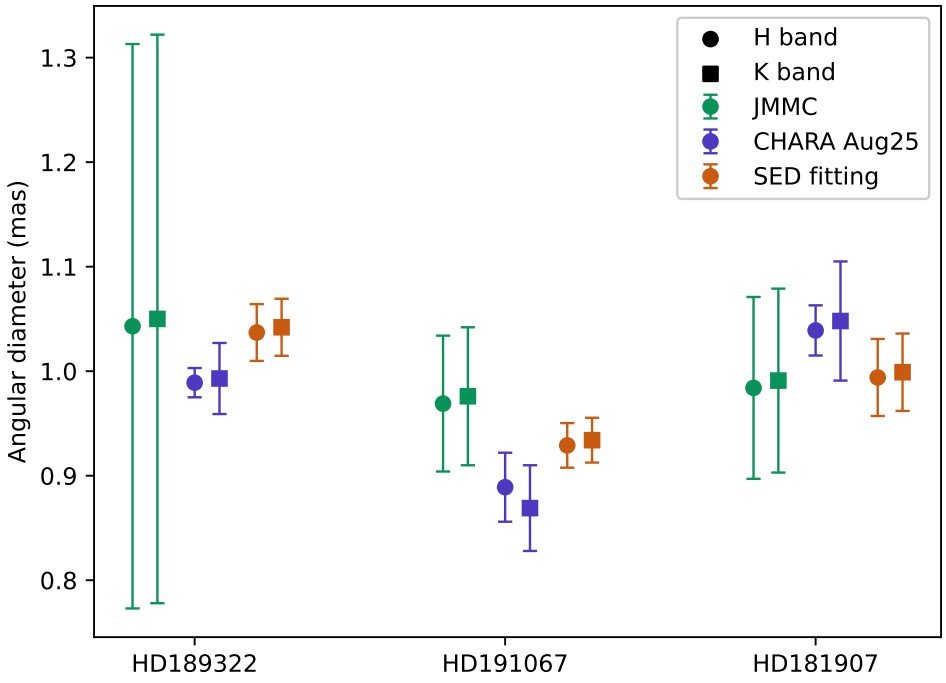


Figure 2.10: Calibrator diameters from the different methods.

6.2.2 Systematic errors on the second lobe visibilities

Having well-constrained angular diameters, with appropriate uncertainties, it is important to propagate the systematic errors in the analysis. This was done carefully in the final analysis. As the calibrators are partially resolved, their visibility curves decrease at higher spatial frequencies, depending on their size. An error in the calibrator size will then induce an error in the transfer function used to calibrate the science visibilities. We can quantify the impact of the error on the calibrator diameter as $\Delta_{\rm CAL} = \frac{\mathcal{V}_{\rm UD}(\theta_{\rm CAL}-e_\theta)-\mathcal{V}_{\rm UD}(\theta_{\rm CAL}+e_\theta)}{2\mathcal{V}_{\rm UD}(\theta_{\rm CAL})}$. It follows that the systematic error on the calibrated visibility is given by multiplying the calibrated visibility with the $\Delta_{\rm CAL}$ function. For multiple calibrators during the night, the transfer function is interpolated between the calibrators observations. We will then take the mean between $\Delta_{\rm CAL1}$, $\Delta_{\rm CAL2}$.... The systematic uncertainty due to calibrator size is not affecting much the angular diameter measurements, as it is mostly determined from the spatial frequency where the visibility reaches its minimum. However, it is of particular importance when quantifying the amplitude of the second lobe of visibility, especially when the uncertainty in the visibility curve is higher at high spatial frequencies. Therefore, the uncertainty on the LD parameter is dominated by this systematic uncertainty. In order to avoid biased calibration and to be able to reject a bad calibrator, we started the run with four calibrators. Moreover, not all of them were observable during the whole night. Our observations were scheduled either in the first half or in the second half of the night. It led us to observe different calibrators from one night to another. This strategy, while justified, presents some risks for our analysis. Indeed, if the size of one calibrator is overestimated, the transfer function will be underestimated and so the calibrated visibility. But when the inaccurate calibrator will not be observed, the calibrated visibility will not be biased. Comparing the two

observations may lead us to interpret apparent visibility variations as physical changes in the LD, since the second visibility lobe is the region most affected by this bias. However, these variations are unphysical. This is exactly what happened with the first analysis with estimates of angular diameter from SearchCal.

6.3 Theoretical and observed variation on limb-darkening parameter

The goal of the η Aql observations is to monitor the LD over its pulsation cycle. This could constrain the geometrical part of the p-factor and its variation with the pulsation.

When it comes to pulsating stars, a standard approach for predicting the LD is to use a sequence of static models reproducing the parameter at each pulsation phase. This is done in SPIPS, using SATLAS models (Neilson et al. 2013). In parallel, I used predictions from the Limb-Darkening ToolKit (Parviainen et al. 2015), which uses the PHOENIX library (Husser et al. 2013). For both cases, it is possible to compute the variations of the LD with the pulsation phase, which are low. This approach is limiting, since it hides the possible dynamical effects due to the pulsation and its impact on the LD. In order to understand these effects, hydrodynamical models have been studied in the past (Marengo et al. 2003; Nardetto et al. 2006). The conclusion from these predictions is that only small variations of the LD coefficient are expected, around 1%.

Even after the secondary calibration, the uncertainties on the calibrator diameters are still large, leading to large confidence intervals in the second lobe of the visibility function, as shown in Fig. 2.11.

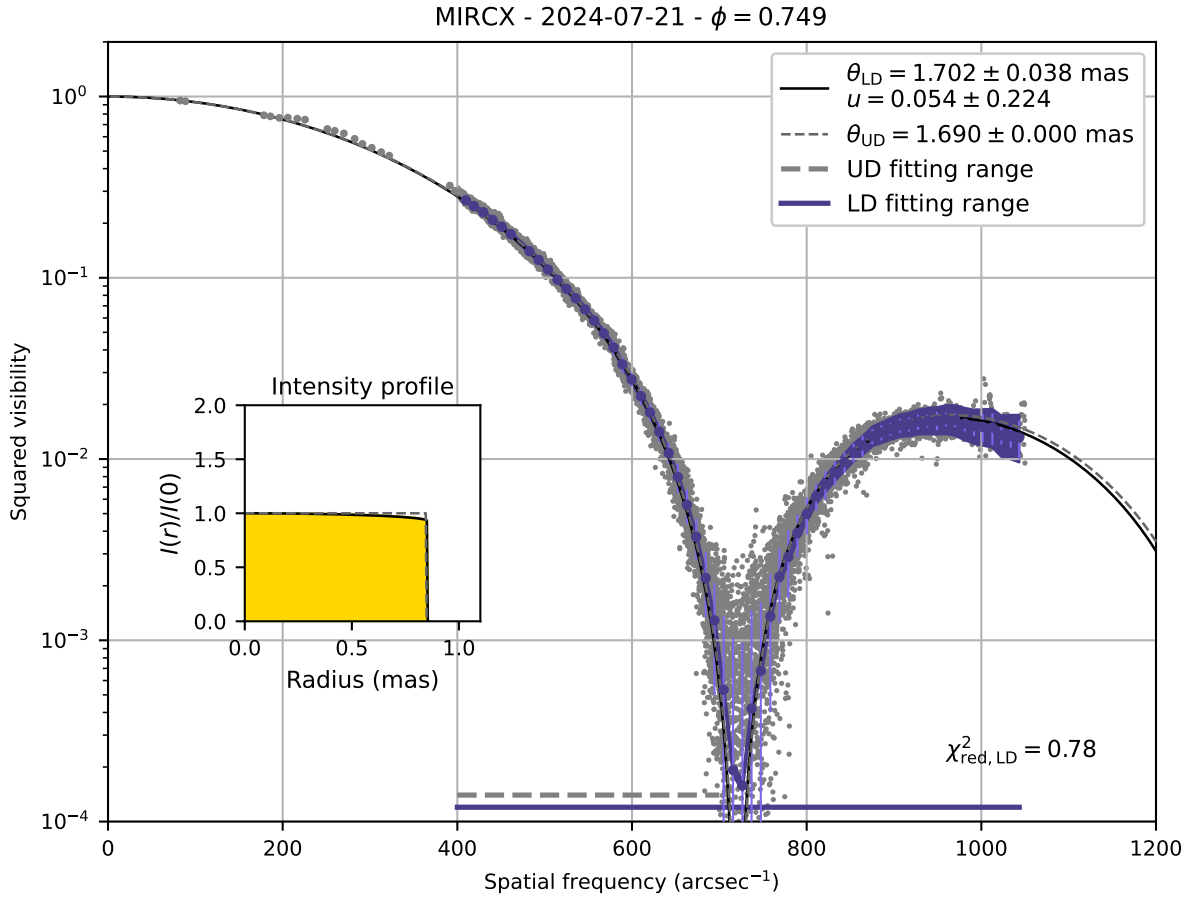


Figure 2.11: Visibility function of η Aql taking into account all calibrators. The systematic uncertainty is shown with a shaded area, increasing with spatial frequency, whereas the error bars represent only statistical uncertainties. See also the caption of Fig. 2.12.

In order to tackle this issue, I propose here an alternative differential approach, in which I only evaluate the variation of the LD on the different epochs, without relying on an absolute value. This is possible by neglecting the systematic error on the visibilities induced by the uncertainties on the calibrator diameters. However, this is unbiased only if the same calibrators are used for the different epochs. I therefore calibrated the data using the unique calibrator HD189322 that was observed in the first four nights of the program. In order to simplify the plot and the fitting process, the data are firstly binned in spatial frequency, taking into account the associated uncertainties. The resulting uncertainty in the binned data shows the dispersion of individual data points. This preliminary step also saves computational time, as the density of points was initially large. The binned data are then fitted with the visibility curve of a UD and a LD disk. A linear law is adopted for LD due to its simplicity. A power law could have been another option, but in any case, the precision of the data does not enable us to identify differences between these two laws. To estimate the errors on the LD diameter and coefficient, I performed a Monte-Carlo sampling of the binned data, given their uncertainties. The results of the fit are shown in Fig. 2.12.

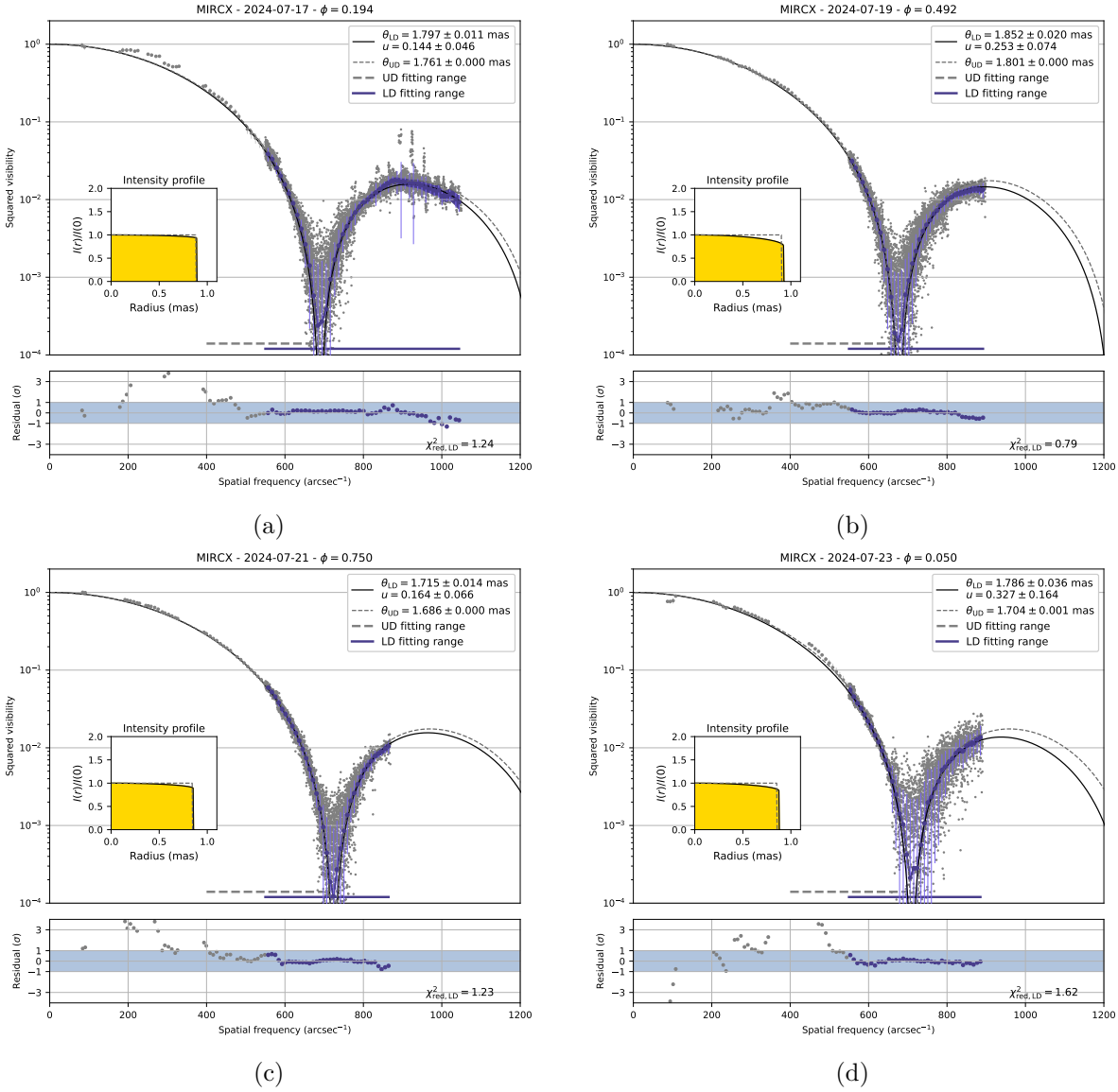


Figure 2.12: Squared visibility functions calibrated using HD189322 on four different nights. Original data (grey) are binned (indigo) and then fitted by a UD (dashed) and an LD (plain) model. The corresponding intensity profile of the LD model is shown in boxes.

For the observations on July 23, the errors on the binned date are particularly large, coming from a large dispersion of the individual points. This gives large uncertainties on the parameters inferred.

The variation curves of LD diameter and coefficient are plotted in Sect. 6.3, together with the predictions from the quasi-static approach, using SPIPS models and LDTk. As a reminder, these results do not account for the systematic uncertainty coming from the calibrator diameter error, so an offset can be expected. The variation is comparable to the model predictions, apart from the July 23 data, which are highly uncertain.

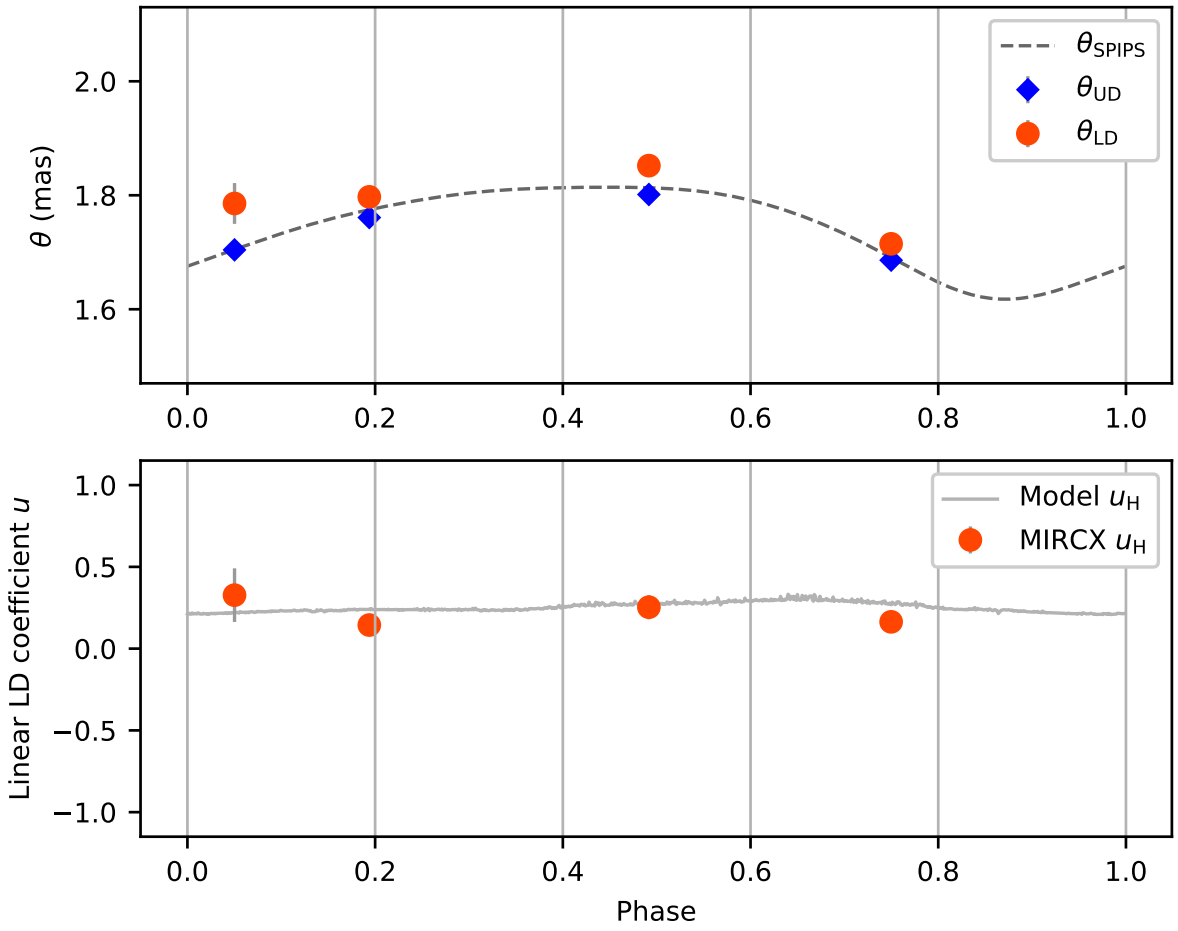


Figure 2.13: Phase variation of LD parameters of η Aql.

Numerically, the observed variation of the LD coefficient between July 17 and July 21 is of the order of 0.1, compared to 0.05 from the atmosphere models.

The results presented here are preliminary, as they only include part of the observed data. Restricting to one unique calibrator not only reduces the number of observing dates, missing the pulsation phase 0.3, but it also reduces the resolution, as the mentioned calibrator was not observable on the same telescopes as our science target for the four dates. This is seen in particular by comparing the data of July 21 between Fig. 2.11 and Fig. 2.12.

Although the present study remains inconclusive due to the high precision required and the current uncertainties associated with the observations, it offers a promising route to observationally constrain the geometric p-factor. Higher angular resolution and sensitivity improvements are required to further develop this approach. The extension to other targets, together with new observations at shorter wavelengths and with next-generation instrumentation, should significantly enhance limb-darkening measurements in Cepheids, as discussed in Chapter 4.

Chapter 3

Modelling the light curves of Cepheids

1 The standard approach

Contrary to the difficulties encountered by the parallax-of-pulsation technique, the road of distance determinations from pulsating stars using Period-Luminosity Relations (PLRs) is less chaotic. However, to reach the high precision necessary to address the Hubble tension, it requires accurate determinations of periods and mean magnitudes. In parallel, observations of extragalactic Cepheids can be extremely time-demanding when reaching distant galaxies. In addition, the current facilities that have the required sensitivity to perform epoch photometry of Cepheids beyond the Local Group is limited to the Hubble Space Telescope (HST) and the James Webb Space Telescope (JWST). In order to mitigate the telescope time, the number of epochs of observation is limited, and in a minimum number of photometric bands. In the next decades, new instruments (Vera Rubin Observatory, Extremely Large Telescope, Roman Space Telescope) will complement the existing data. From HST and JWST observations, we reach up to 41 photometry epochs unequally distributed in 7 photometric bands for the most observed galaxy. As an example, Fig. 3.1 shows a synthetic dataset for a 40-day period Cepheid in the most observed galaxy in the SH0ES (Supernova H_0 for the Equation of State) sample. The mock data are created from the templates I built using typical uncertainties for this kind of data (Hoffmann et al. 2016; Riess et al. 2023). The theoretical light curve from the template is shown in dashed line in each photometric band. The exact number of observing epochs in this galaxy is respected. In particular, this shows that only two epochs are available for the reddest wavelengths, which motivates the need for a multi-wavelength fitting approach.

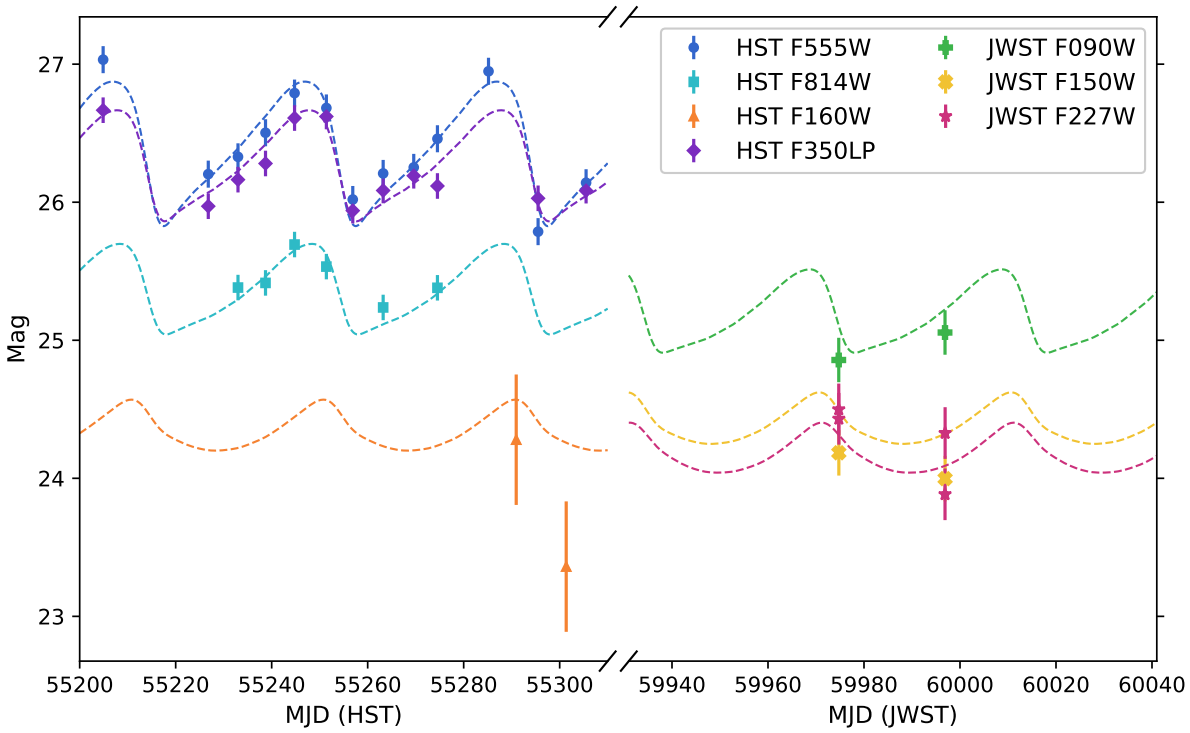


Figure 3.1: Synthetic epoch photometry for a 40-day period Cepheid in NGC5584 observed with HST and JWST at realistic observing times, and corresponding template light curves.

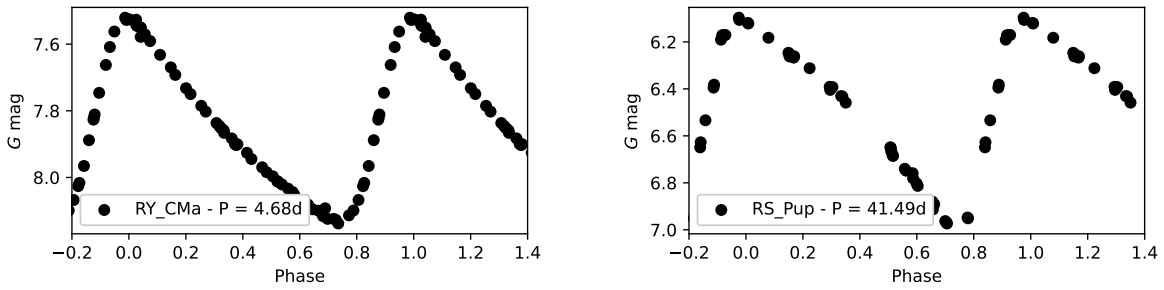
In order to compare Cepheids from photometric data, it is important to compare their light curves. In this section, I describe the usual way to parameterise light curves.

1.1 Light curve

The term light curve refers to the variation of apparent or absolute luminosity, bolometric or in a specific photometric band, during the variation cycle of a pulsating star. We usually refer to folded light curves, i.e, the variation of light with respect to the phase of the observations. The phase of an observation taken at time t is computed as $\phi = \frac{t-t_0}{P} \mod 1$, where P is the period and t_0 a reference time, set conventionally, usually to the time of maximum flux in a specific band. In this part of the analysis of pulsating stars, the period is supposed to be already known.

The luminosity of pulsating stars varies during the pulsation cycle as a result of the combined and phase-shifted variations of the stellar radius and effective temperature. Knowing precisely the shape of the light curve is needed to extract the mean magnitude, that is used in PL relations. Fortunately, Cepheids are easy to observe in our galaxy, and numerous campaigns were devoted to the monitoring of their light curves. From these observations, a large variety of light curve shapes was identified. The first characteristic is that the shape depends strongly on the period of the star, for the same instrument in the same filter and the same class of pulsators. The linear pulsation theory predicts a sinusoidal shape for the light curve. However, non-linear effects, notably related to the κ-mechanism, produce asymmetric light curves, looking more like

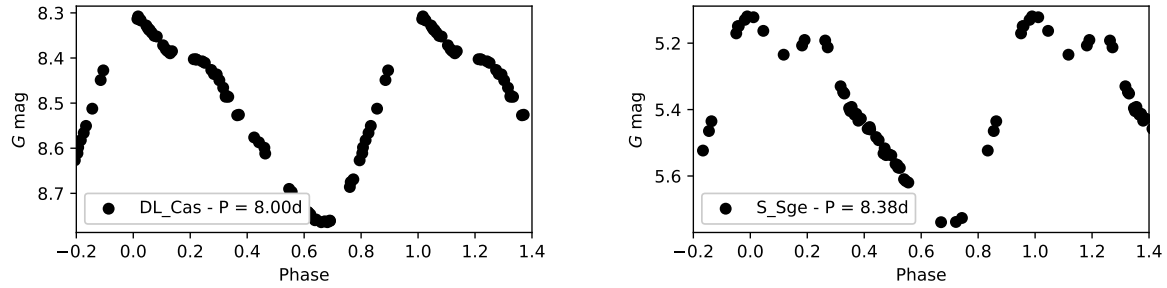
sawtooth, with a long decreasing branch and a sudden increase in luminosity towards the end of the cycle. An example is given for two Cepheids of different periods in Fig. 3.2 from the Gaia G band light curves. Between those two periods, the shape evolves continuously. A similar behaviour is observed for fundamental-modes RR Lyrae stars.



(a) (b)

Figure 3.2: Comparison between two typical light curves for a short-period Cepheid (a) and a long-period Cepheid (b).

Besides the general evolution of the light curve shape with the period, it is possible to observe the presence of a secondary bump, near the maximum light. It appears for a lot of stars around 6 days of period in the descending branch and moves towards the maximum of luminosity around 10 days of period. Finally it disappears smoothly on the uprising branch around 14 days of period. This behaviour was first discovered by Hertzsprung (Hertzsprung 1926) following the study of the best covered light curves at that time. Now known as the Hertzsprung progression, it is characterised by the central period, around which we can see this specific secondary bump. The centre appears at the transition where the secondary bump reaches the maximum light and hence flattens the curve. The origin of the bump is still discussed, but two main hypotheses are favoured. It may come from the resonance between the fundamental and the first overtone mode of the pulsation or to an echo produced by the pressure excess in the Helium ionisation zone, propagating first towards the stellar core to rebound up to the surface during the next pulsation cycle. From the systematic observation of light-curve shapes, with increasing amount of well-covered Cepheids, it is now possible to predict it for every Cepheid in the Hertzsprung period range, however with various amplitudes. Two different behaviours for similar periods are shown in Fig. 3.3 from the Gaia light curves in G band.



(a) (b)

Figure 3.3: Characteristic secondary bumps of the Hertzsprung progression for two Cepheids with a period of around 8 days (descending branch), one with a moderate bump (a), and one with a very prominent bump (b).

Interestingly, this dependency of the light-curve shape on the period was also observed in extragalactic Cepheids. It is shown by the stacking of observations for Cepheid in period bins in Figure 6 from Riess et al. (2022). They show as well the general evolution of the shape and the presence of the secondary maximum. The obtained shapes compare remarkably well with similar Galactic Cepheids.

1.2 Standard parametrisation

To precisely know the light curve shape, we need a method to convert discrete data into a continuous light curve. For this, two main methods exist.

1.2.1 Fourier decomposition

As we already know, light curves of Cepheids are periodic signals. To that extent they can be modelled by standard Fourier series. The Fourier decomposition is indeed casually used for light curve interpolation. The light curve is then approximated by the following equation.

$$m(t) = A_0 + \sum_{k=1}^{N} A_k \cos\left(2\pi k \frac{t - t_0}{P} + \phi_k\right) \tag{3.1}$$

In this equation, P is the period of the star, A_i and ϕ_i are the Fourier coefficients and t_0 is the reference time. It can be chosen conventionally to set the zero phase at the maximum of luminosity, or sometimes to the mean luminosity in the uprising branch. It is also common to use the derived dimensionless Fourier parameters $R_{i1} = \frac{A_i}{A_1}$ and $\phi_{i1} = \phi_i - i\phi_1$. Finally, N is the order of the Fourier decomposition. It has to be set to fit the complexity of the light curve. Commonly it is between 4 and 6. Adding too many modes may overfit the noise but a too simple parametrisation can bias the curve shapes. The parametrisation using Fourier parameters shows well the dependence of the light curve shape on the period, as illustrated by Fig. 3.4.

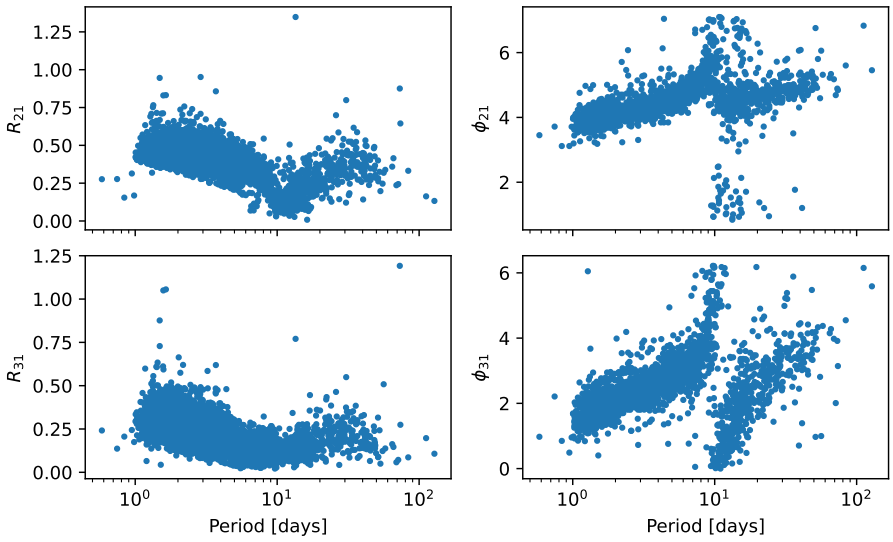


Figure 3.4: Fourier parameters in G band from Gaia database (Gaia Collaboration et al. 2023) for the 4415 fundamental Cepheids with at least 40 observations in G.

In addition, they have shown specific dependence to other characteristic quantities like the temperature (as linked to the period) and the metallicity. (see for instance Hocdé et al. 2023).

1.2.2 Principal Component Analysis

The Principal Component Analysis (PCA) is an analytical tool that aims at decorrelating variables to efficiently extract the information and help dimension reduction. It allows representing a multi-dimensional dataset in smaller dimension, preserving the total variance. Its mathematical framework can be explained using linear algebra. Let us describe the data by the matrix X. This matrix is filled with n realisations of an experiment, in which p quantities are evaluated. The dimension of X is $n \times p$ The covariance of the data is explained by $C_X = \frac{1}{n-1} XX^T$. In the basis vectors given as an input in the X matrix, correlations may appear. The goal of PCA is to find an alternative basis with the transformation matrix, in which the data are uncorrelated. Mathematically, we look for the matrix Y of the data, and the transformation P such that $Y = PX$ and C_Y, the covariance matrix of the data in the second basis, is diagonal. Let us remark that $C_Y = \frac{1}{n-1} YY^T = \frac{1}{n-1} PXX^T P^T = PC_X P^T$. From the spectral theorem in finite dimension, we know that a real symmetric matrix, like the covariance matrix C_X, is diagonalisable, and the transformation matrix is orthogonal. This means that there exists an orthogonal matrix Q such that $Q^T C_X Q$ is diagonal. Taking $P = Q^T = Q^{-1}$ is bringing us back to C_Y. From that derivation, we understand that the new basis can be easily constructed by looking for the eigenvectors and eigenvalues of the covariance matrix of the data. In addition, eigenvalues of the covariance matrix give information on the weight of this direction in the dispersion of the data. Hence, sorting the new basis by the eigenvalues allows us to reduce the dimension to the most important directions. The PCA gives new variables, represented by the matrix Y, that are uncorrelated. The first component gives the most significant variations while the last ones encode smaller amplitude noise. The variance fraction, represented by the

quantity $f_k = \frac{\lambda_k}{\sum_{i=1}^{n} \lambda_i}$, where λ_i are the eigenvalues of the covariance matrix C_X, gives how much the k component contributes to the total variance. In order to reduce the noise from the dataset, we can keep only the components that have a significant impact on the variance, by setting a limit number of component m such that $\sum_{k=1}^{m} f_k \geq 0.95$ for instance. Here, keeping the m-th first components will reproduce 95% of the variance of the dataset.

The Fig. 3.5 illustrates the simplest case of PCA with $n = 2$, $p = 200$.

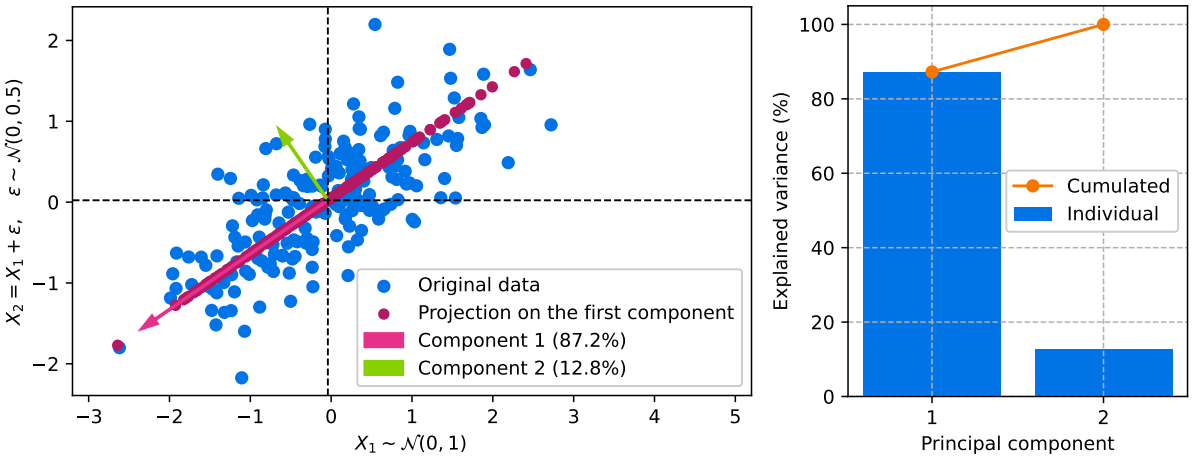


Figure 3.5: Simple illustration of the PCA principle, with a representation of the dataset (left) and the variance plot (right)

PCA is a powerful tool for linearly correlated datasets. Non-linear cases would need other tools. However, its application requires to prepare the data in the appropriate format, which means that they need to be standardised and with a zero-mean. Also, it is not easy to understand the physical meaning of each principal component. In the case of Cepheid light curves, there are two main ways to standardise the datasets. Either we work on the Fourier coefficients to reduce the number of coefficients, but it relies on a good Fourier decomposition, or can work on the interpolated light curves, which also requires a good quality interpolation of the data to sample the light curves the same way.

Once the PCA is performed, it offers a way to approximate a variety of light curves. The principal components form a set of basis functions that replace the usual sine and cosine functions in the Fourier decomposition, and we can determine the set of coefficients in this new basis of functions. The main advantage of PCA is that the resulting basis functions are better suited to reproducing light-curve shapes than simple sinusoidal functions.

1.3 The link between light curves in different photometric bands

The description of light-curve shape is more difficult due to the diversity of photometric bands in which pulsating stars are observed. The shape of the curves may differ according to the wavelength domain of the observations. This leads us to develop strategies to combine homogeneous data; otherwise, it prevents us from using data from different instruments to interpolate the same light curve. In the beginning of this chapter, I presented various light

curves in the visible domain, close to the V-band. In contrast, in the IR, light curves are smoother, and more sinusoidal, even for short-period Cepheids. The amplitude of variability is also lower. This is represented by the light curves of RS Pup in Fig. 3.6.

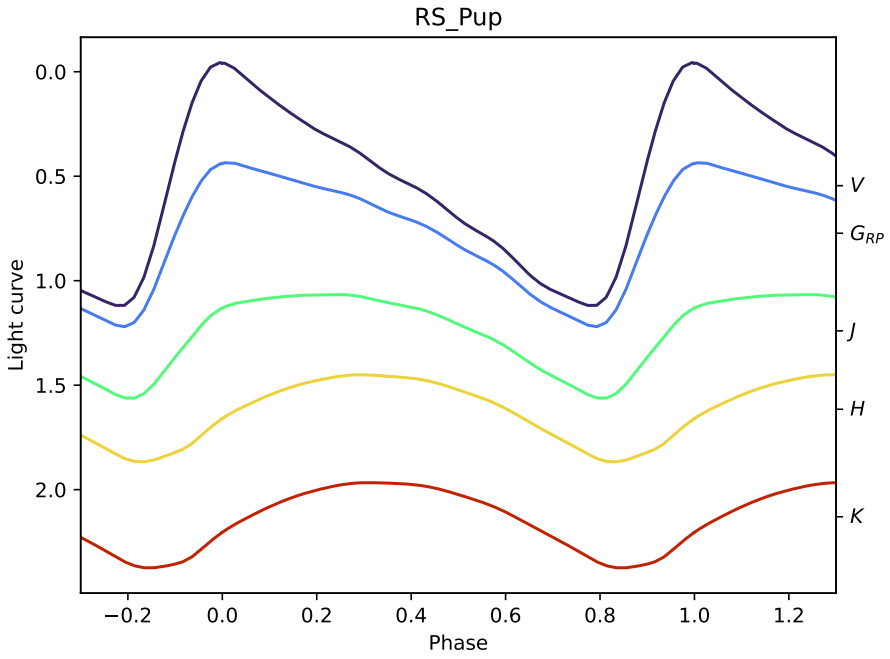


Figure 3.6: Light curve of RS Pup in five different photometric bands. Light curves are normalised to zero-mean and offset vertically by the wavelength in microns to show the evolution of the shape.

We expect that the link between the different photometric bands can be predicted. In the SPIPS code (Spectro-Photo-Interferometry of Pulsating Stars), it is intrinsically done by computing the Spectral Energy Distribution (SED) at each phase. In the standard approach, it is common to study the different filters independently. For instance, when doing Fourier decomposition, each light curve is fitted by a Fourier series with its own parameters A_i^f and ϕ_i^f. They can then be linked by calibrating interrelations between the Fourier coefficients. When comparing different light curves in different bands, another important quantity is the phase delay between the maximum light in the various filters. If this quantity does not change the shape of one light curve in one band, it can change the global modelling of the pulsation in all the observable bands.

2 Towards a more efficient modelling of light curves

The second part of my PhD project aimed at producing templates of Cepheid light curves, with the objective to apply them to HST images of distant galaxies from the SH0ES team. For this project, I started from the principle explained in Javanmardi et al. (2021), and I developed a specific methodology to reach my objectives. This section intentionally does not detail the full analysis, as the details are provided in the attached A&A paper (Bras et al. 2025). Here,

I aim at clarifying the general method and describe the difficulties I encountered during the implementation.

2.1 Our implementation of the principal component analysis

The motivation behind our new implementation of PCA for light curve approximation is that Cepheids are observed with various facilities, with specific photometric filters and every light curve from all these telescopes are very similar but combining them blindly is subject to systematic biases. It is then important to find a way to use them simultaneously in a global modelling of light curve shapes.

2.1.1 The training sample

For our analysis, the standardisation of the input data is particularly interesting. The idea is to have one PCA for all filters. This can be done by concatenating the light curves in the various photometric bands, sampled the same way. We sampled light curves at 0.001 in phase, in 43 different filters (including several similar filters like $K_{\rm CTIO}$ and $K_{\rm 2MASS}$ for instance). One example of these concatenated light curves can be represented by the following list

$$[\ \underbrace{m_1^{f_1}, m_2^{f_1}, \ldots, m_{1000}^{f_1}}_{\text{Light curve in the first filter } f_1}, m_1^{f_2}, m_2^{f_2}, \ldots, m_{1000}^{f_2}, \underbrace{\ldots}_{\text{Light curve in the } i\text{-th filter } f_i}, m_1^{f_{43}}, m_2^{f_{43}}, \ldots, m_{1000}^{f_{43}}]$$

The 43 filters contain a selection of the commonly used filters, including also HST and JWST filters. The list can easily be extended upon usage. It is particularly interesting for new facilities, for instance for the upcoming Vera Rubin Observatory.

The concatenated light curve is then normalised to have a zero mean, which is necessary for the PCA. I call the normalising parameter $m_{\rm ref}$ as it has the same dimension as a magnitude.

Reversely, we can extract the original light curve in a filter f_i by truncating the concatenated light curve F:

$$m^{f_i}(\phi_i) = m_{\rm ref} + F_{1000i:1000(i+1)}$$

In order to have a standardised training dataset, it is necessary to have the F vector for a selection of different Cepheids. This means that we need to have an interpolated light curve in each of the 43 filters for each Cepheid. Hence, we need the SED and its variation along the pulsation cycle. This is one of the products of the SPIPS algorithm. Therefore, I used the SPIPS models of Cepheids obtained by Boris Trahin and Antoine Mérand (developer of SPIPS). I gathered all the SPIPS models with a satisfying solution obtained for Cepheids, for a total of 75 fundamental-mode Cepheids in the training sample. The SPIPS output files of these 75 stars gives the modelled light curves only in the bands in which observations are available. Hence, the first necessary step is to compute the light curves in the 43 filters, from the stellar parameters ($\theta(\phi)$, $T_{\rm eff}(\phi)$, $\log g(\phi)$, and [Fe/H]) and the atmospheric models. A particular aspect in the SPIPS code is its treatment of the reddening, as explained in Sect. 1.2. It uses a variable

correction, depending on the instantaneous effective temperature of the star. This is not possible in general, outside SPIPS models, as we need to know the temperature variations. Hence, the template can only be applied with a fixed correction, not depending on the temperature. In order to be consistent in my analysis, and to recover the light curves of my training set when applying the templates, it is then required to remove from the photometry computed by SPIPS the effect of the varying reddening. For this reason, I first applied the reddening effect in the SPIPS way on the light curves, and then dereddened the resulting photometry using a constant extinction parameter. I however used extinction parameters computed at 5500K, as explained in Sect. 1.2, as this temperature is better suited to Cepheids and the use of $E(B-V)$ in the SPIPS products. The process is illustrated in Fig. 3.7. The synthetic photometry from atmosphere models (referred to as "Initial") is corrected for the interstellar extinction using a varying parameter (giving the "Reddened" curve), and then the reddening effect is removed with a fixed extinction parameter. The resulting "Corrected" curve, where the effet of the reddening has been corrected, is subsequently used as training data for the PCA. In both the "Initial" and "Corrected" curve, an appropriate reddening correction gives the "Reddened" curve that corresponds to the observations. In the first case, the reddening correction should take into account the temperature of the star at each phase, whereas the second yields the "Reddened" curve provided a fixed colour excess computed for a star at 5500K. As the temperature at each phase is not accessible in general, the "Corrected" curve is the only one that can be used for large scale applications.

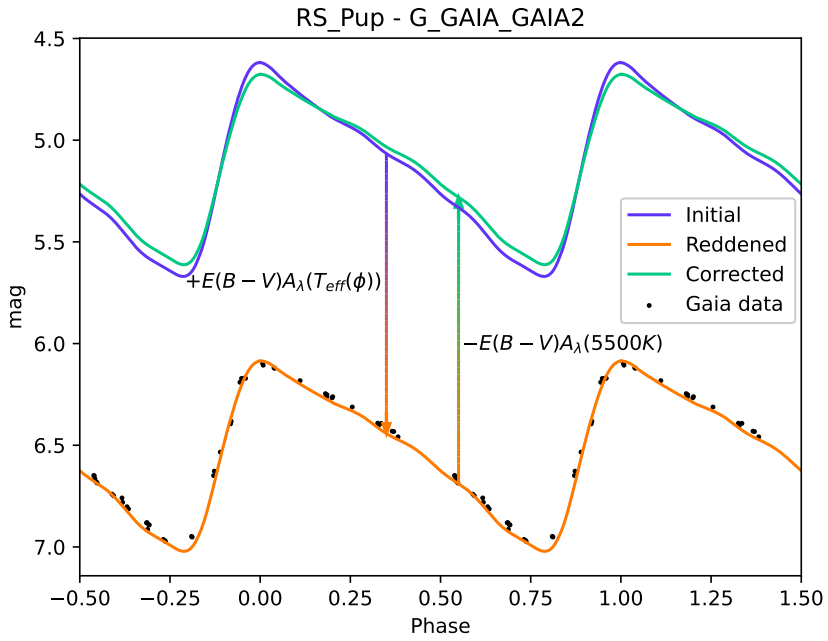


Figure 3.7: Light curve in the Gaia G band of the Cepheid RS Puppis. The G band gives a good comparison as this band is broad and hence very sensitive to the integration of temperature dependent reddening.

The light curves are also corrected for the distance, by subtracting the distance modulus

$\mu = 5\log(\frac{d}{0.01})$ to the magnitude, with d the distance in kpc. Finally, the infrared (IR) excess due to CircumStellar Envelopes (CSEs) is added to the luminosity, so the resulting light curve is comparable to the observed light curves.

2.1.2 Results of the Principal Component Analysis

The PCA is performed using the scikit-learn python library (Pedregosa et al. 2011) and I analysed the different products. To evaluate the performance of the PCA I investigated the variance fraction. The quick increase of the cumulated variance ratio indicates that the PCA is well-suited, and a small number of parameters will be sufficient to recover typical light curves. The final threshold of the dimension reduction is set using this estimation, and assuming a limit on the percentage of variance we want to reproduce. In my case, I fixed the limit number of components to four to keep mainly the parameters that depend on the period. As the corresponding cumulated variance ratio is 96.5%, this choice is satisfactory. The PCA first extracts a mean vector C_0. The other products are the principal components C_i and the transformation matrix, which gives the decomposition of the training sample in the new basis, i.e, the coefficients associated to each component for each one of the stars in the training sample. The concatenated light curve for a specific star can hence be recovered by the formula $F(star) = C_0 + \sum_{k=0}^{75} K_k(star)C_k$. As the PCA is fed with concatenated light curves, the principal components are also a concatenation of curves. I had a closer look on the shape of the principal components in the various filters, using the truncation of the principal components $C_k^{f_i} = C_{k_{1000i:1000(i+1)}}$.

In Fig. 3.8 I show the C_0 function as obtained by the PCA. It consists of a concatenation of the mean light curve in each photometric band. Bands are sorted to go from the bluest band to the reddest. This function contains the information of the global shape of the light curves as well as the general differences in absolute magnitudes of the different photometric bands. The magnitude step between each LC is directly related to the temperature and the colour excess. In general, a redder or more extinguished star will present a similar concatenated LC in terms of individual shapes but the steps will be higher, with the reddest wavelengths more luminous, compared to the bluer ones. As a remark, we see in this figure one LC that does not follow the general trend. It is associated to the Walraven V band. Magnitudes in this system are not expressed using the same law as the other systems which causes the scale difference.

2.2 Predicting light-curve shapes

At this stage, the PCA is completed, and we have obtained a small set of functions that are enough to reproduce the large variety of light curves. The next step is to predict the light curve shape from the minimal knowledge of a Cepheid, which is its pulsation period.

As the light curve shapes depend mostly on the pulsation period, as discussed in Sect. 1.1, it is possible to relate it only to the period of the Cepheid. To do this, I investigated the distribution of the coefficients associated to the four first components I kept with respect to the pulsation

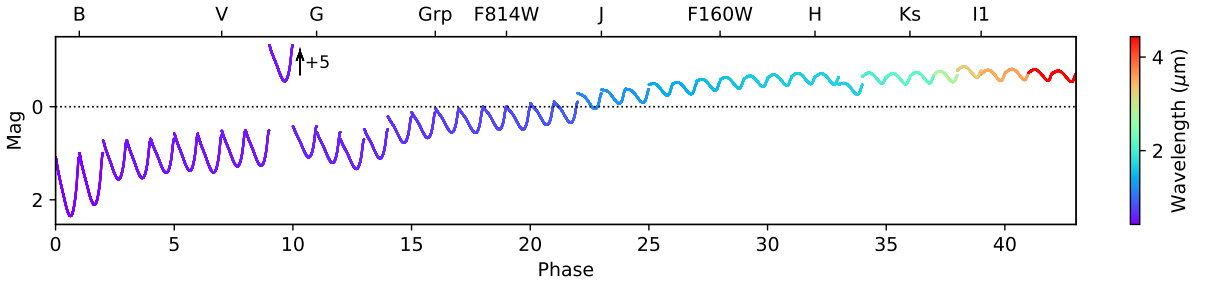


Figure 3.8: The concatenated mean light curve from the PCA. The Walraven V-band LC is offset for readability.

period for the 75 stars in the training sample. A close look at these distributions shows complex dependences on the period. We can distinguish easily by eye the population of short period Cepheids ($P < 10$d) and long-period Cepheids ($P > 10$d). Around 10 days, Cepheids tends to gather in a specific region on these diagrams, which is clearly visible for components 2 and 4. Simple linear relations can not be applied given these distributions to connect the K_i to the period. The relations I derived are computed to recover the average trend of the distributions. I used a moving average with a Gaussian weight. This means that for each period, I take the mean coefficient from the stars in the training sample, weighted by the difference in period, so the similar stars give the dominant constraint. Mathematically, it can be summarised as follows,

$$K_i(P) = \frac{\sum_j K_i(j) e^{-\frac{(P(j)-P)^2}{2w^2}}}{\sum_j e^{-\frac{(P(j)-P)^2}{2w^2}}}$$

where j is the index of each of the 75 stars in the training sample and w is the width of the moving box. I fixed its value from the visual inspection of the relations and using the validation test described afterwards. Small variations of the smoothing parameter w do not have a significant impact on the final results.

The caveats of this methodology, which relies on a set of 75 stars, is that we need to be sure that those Cepheids constitute a representative sample. In particular, as they are supported by SPIPS pulsation models, the sample includes mainly nearby and well-observed Cepheids. Moreover, this type of statistical analysis is always limited by the size of the training sample. Some efforts to improve the statistics are needed. In our case, it is crucial to be careful with the predictions for long-period Cepheids as few of them have a SPIPS model and are included in the training sample. Unfortunately, while Radial Velocity (RV) curves are not used directly in the PCA, they are necessary to achieve good models and predict the stellar parameters, essentials for obtaining the synthetic photometry in the 43 filters. Hence, the number of stars in the training sample for the PCA is too limited. Fortunately, the sample of stars for which we have access to well-sampled light curves in various filters is more extended. This motivated me to include in the analysis several additional Cepheids with good photometry but no SPIPS models, with a particular interest for long-period Cepheids. For this selection of stars, I fitted the light curves

in the available photometric bands with the principal components in the results of the PCA (C_i). The coefficients obtained by the fit (K_i) are then compared with the previous sample. The selection of Cepheids was done based on the quality of the photometry essentially, and I mostly considered long-period Cepheids as the number of short-period Cepheids was already acceptable. It is more interesting to reach a good balance between the different period bins. This selection was then refined with respect to the convergence of the fit and the accurate reproduction of the light curves. It should be noted that when fitting the photometry of these additional stars, the colour excess has to be fixed as it can be degenerate with some of the K_i coefficients (and in particular with K_1). Colour excess values are therefore gathered from the literature and rescaled using the factor determined in Sect. 1.2. The coefficients associated with the principal components, i.e. the K_i values for this set of stars, were compared with the ones from the training sample. Eventually, the relations $K_i(P)$ were derived from the two samples together, using the moving average previously described. They are illustrated in Fig. 3.9. In this plot, the amplitude of the y-axis is the same between the four main plots to highlight the decreasing importance of the components of higher order.

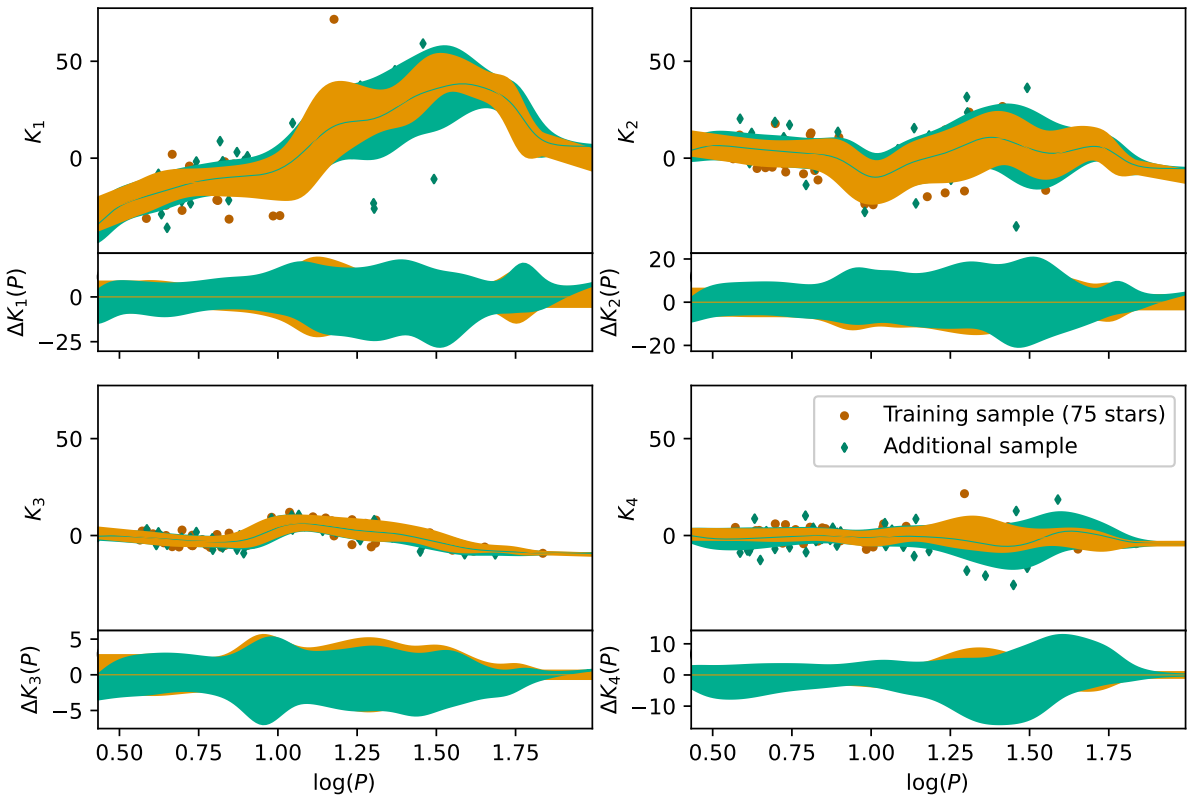


Figure 3.9: The relations $K_i(P)$ for the first four components using only the training sample (brown points, and brown line) and both samples (brown and green points, green line). The bottom plot for each panel represents the variation of the relations when adding the additional stars.

2.3 Light-curve templates

The combination of the principal components (C_i) with the relations $K_i(P)$ allows us to predict the light curve of a Cepheid of a given period in the 43 different filters. In addition, they give information on the magnitude differences in these 43 filters. Practically, this means that knowing the period P_* of the Cepheid only, we have access to the relative shape of the light curve in any of these filters (f) by using the formula $m_f(\phi) = [C_0(\phi)]_f + \sum_{i=1}^{4} K_i(P_*)[C_i(\phi)]_f$ where the notation $[\ldots]_f$ represents the truncation of the function corresponding to the filter f. The correction from the interstellar colour excess can be achieved by recovering the right differences in the photometric bands including the effect of interstellar dust, and the reddened magnitudes thus become $m_f(\phi) + E(B-V)A_f(5500K)$. Finally, when compared to real data, one last parameter has to be used to rescale the global photometry. This $m_{\rm ref}$ parameter is characteristic of a Cepheid and applies similarly to all bands. It basically takes into account the distance of the Cepheid. Its specific usage summarises the whole originality of this method as it is standard to scale template light curves independently for every band. This section describes the application of the developed templates to the validation samples. I distinguish between two validation samples: the first one is focused on the modelling of light curves in observed filters, whereas the second tests the prediction of the light curve in other filters for which observations are not available.

2.3.1 Test sample in observed filters

The first test sample is composed of a selection of Cepheids with various photometry (including several data points, in the visible and in the IR) that are not included in the development of the templates. It contains stars with poor phase-coverage in some photometric band or for which the fit of the coefficient K_i did not converge. For the fit, all the photometry is used simultaneously to find the three free parameters: the scaling magnitude $m_{\rm ref}$, the colour excess $E(B-V)$ and the phase offset. The latter parameter is necessary to align the folded observed light curves to the templates, for which the zero phase is set to the maximum flux in bolometric luminosity, as defined in the SPIPS analysis. The fits can be discussed independently for each Cepheid, but it is more interesting to evaluate the possible biases of the templates for this test sample. In particular, as the templates are particularly meant for the determination of mean magnitudes, I studied the possible differences in the mean magnitudes derived by the template fitting and using classical Fourier interpolation. For the large majority of the filters for which I could do the comparison, the agreement is satisfying. Details of the comparison can be found in the attached publication.

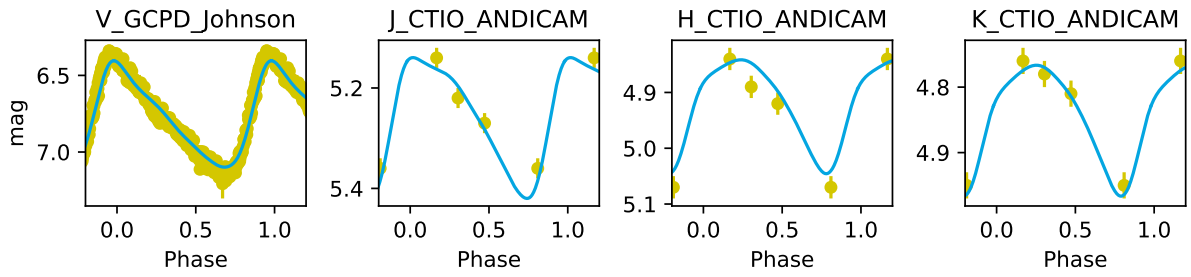
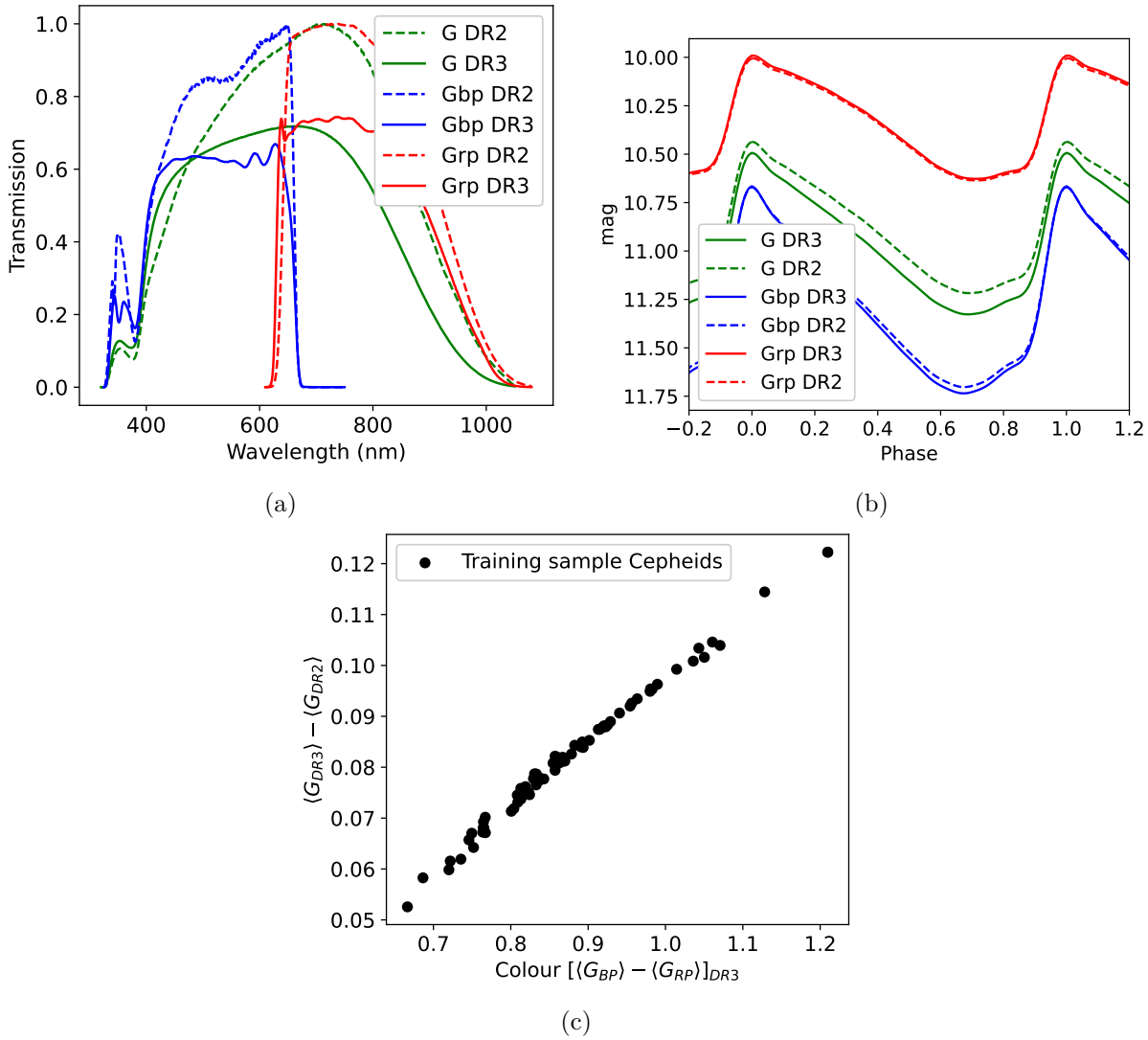


Figure 3.10: Example of template fitting for the Cepheid R Cru in the test sample.

However, for Gaia filters, some difficulties appear. Firstly, between the second (DR2) and third data release (DR3), the transmission curves of the Gaia filters have been recalibrated and are significantly different. Consequently, the synthetic photometry has changed for the same stellar parameters. This is characterised by a systematic offset between filters using the DR2 and DR3 calibration for the templates, particularly in the G band. However, DR2 data are also included in the DR3. The result is that the template will never be able to fit Gaia data in both cases. In practice, we observe that using the transmission curves from the DR2 gives photometry compatible with the template fitted from the other filters, but not for the DR3. The simple solution is then to use the definition of the filters in the DR2 when using Gaia data, but this is not ideal, and needs to be investigated in the context of the upcoming DR4. The plausible reason for which the DR3 calibration does not work is that the weight of the bluer part of the spectrum in the G filter increased with the most recent calibration. However, we already stated that stellar atmosphere models fail to reproduce the photometry at shorter wavelengths than V band. The error introduced by the bluer part of the spectrum is then increased and makes the G band impossible to fit. This is also supported by the fact that the offset between the templates in the Rp filters from the previous to the new calibration is lower. The difference between the DR2 and DR3 transmission curves and an illustration of the effect on the templates are provided in Fig. 3.11.

2.3.2 Test sample in other filters

The second test sample is focused on the validation of mean magnitudes on HST filters. What makes this approach interesting is that the test is done for photometry fitted by the present templates without any data in the mentioned filters. The comparison is made using the study from the SH0ES team in Riess et al. (2021). In this paper, the authors publish the mean photometry in HST filters from HST time series observations. However, the epoch photometry is still unpublished. I decided to take advantage of this work as it is accessible, in order to demonstrate the originality of my templates. In comparison, the usual approach to derive mean magnitudes in unobserved filters is to use empirical conversion relations. The sample includes all Cepheids in Riess et al. (2021). The large majority of these stars are included in one of the previous samples but the template fitting is done independently of previous results. The fit follows the same strategy as the first test sample, but I extract in addition the predicted light curves in the HST filters. The results are detailed in the attached publication. Globally,



(a) (b)

(c)

Figure 3.11: (a) Transmission curves of Gaia filters calibrated for the DR2 and the DR3. (b) Difference in the template for a 20-day period Cepheid. (c) Dependance of the offset in the template in G band on the colour.

the templates are efficient in retrieving mean magnitudes in these filters. In addition, statistics on the differences in mean magnitudes are used to evaluate the potential biases of the method. Such a fitting strategy also provides estimates of the colour excess for each Cepheid, which, when compared to literature measurements, are particularly accurate: the standard deviation of the difference with previous determinations is on the order of the classical error on colour excess measurements.

3 Modelling extragalactic Cepheids

The templates presented in this chapter are motivated by the analysis of HST images of nearby galaxies for the calibration of the distance ladder. They are part of a larger project to calibrate independently the distance ladder and the Hubble constant. The main work was

done by Anton Afanasiev, post-doctoral researcher in the team during the time of my PhD. This section is dedicated to the application of my templates in the context of Anton's work on extragalactic Cepheids.

3.1 Observation of extragalactic Cepheids

In order to present the data used for the template fitting, I hereby describe the main steps for extracting the photometry of extragalactic Cepheids, as done by Anton.

3.1.1 HST photometry of Cepheids

The HST is one of the most important facilities for observational astronomy. Launched in 1990 and still in operation, it is composed of a 2.4-m telescope and up to five instruments, including mainly imaging instruments. Useful in various fields of astrophysics including planet atmospheres, galaxy formation and dark matter detection, it is not surprising that one of its main mission was to study Cepheids in the context of the distance scale, as it is named after the astronomer Edwin Hubble, the first to estimate the age of the Universe using its expansion.

An important heritage of Hubble is the SH0ES program, devoted to the observation of several galaxies hosting a type Ia Supernova (SNIa). The HST images of SNIa host galaxies are now public in the Barbara A. Mikulski Archive for Space Telescopes (MAST).

The data for extragalactic Cepheids in Riess et al. (2022) include HST observations for 37 galaxies between 2009 and 2020 from various observing programs led by the SH0ES team, and supplemented by some other observations. From this date, the data obtained by the SH0ES team use only four selected filters on the same instrument of the HST, in order to have homogeneous data and avoid calibration of photometric zero-points and the systematic uncertainty it brings. They use the large filter F350LP, the two optical filters F555W and F814W and the IR filter F160W of the WFC3 (Wide-Field Camera). The standard data are composed of around 12 epochs spreading around 60 to 90 days to retrieve both short and long-period Cepheids. The F350LP filter is mostly used for the Cepheid detection, identification and phasing in the other filters. The epoch photometry on the other filters allows deriving a Wesenheit magnitude and calibrate Period-Wesenheit Relations (PWRs).

In a previous analysis of the team (Javanmardi et al. 2021), the photometry of Cepheids in one galaxy was obtained thanks to the HST images of this galaxy and the Cepheid identification from the SH0ES team. The work led by Anton Afanasiev does not take into account the identification from the SH0ES analysis. Epoch photometry is obtained by Point-Spread Function (PSF) fitting on each source, for each epoch. The identification of Cepheids is done using several indicators. First, we rely on the position in the Colour-Magnitude Diagram and select the Instability Strip (IS). Secondly, stars are selected from the variability of their epoch photometry, as the amplitude of Cepheids is quite stable. Finally, the probability of a star being a Cepheid can be evaluated by the quality of Cepheid-template fitting. Indeed, fitting a star with the templates I produced should fail and give a large χ^2 value if the variability is not

characteristic of a Cepheid. Anton provided me with time-series photometry for a selection of Cepheid candidates in the galaxy NGC5584 to test the template-fitting strategy and evaluate the feasibility of the Cepheid identification using template fitting. NGC5584 is a favorable case study, as it is the galaxy that was most observed by HST in the SH0ES programs, providing the largest number of photometric epochs.

3.1.2 JWST photometry

Sometimes considered as the successor of the HST, the JWST is actually its IR twin. While having similar resolution in the red to mid-IR domain, thanks to its 6.5-m aperture, it can observe much fainter and distant objects. Launched in December 2021, with a first image released in July 2022 it is now fully operational and has today the highest oversubscription rate of all astronomical telescopes, which means that the time asked during call for proposals is largely exceeding the allocated time.

The JWST telescope is game-changing in the field of the distance scale, thanks to its sensitivity in the IR, less depending on the reddening, and to its high spatial resolution, most useful in crowded fields. The photometry of individual stars in galaxy images requires estimating the background luminosity, which may be composed of dust, clouds, or other stars. The estimate of the background level is usually done by incorporating artificial stars in the field, with known brightness. The difference with the recovered brightness using the measurement method (that can be PSF or aperture photometry) gives the quantity that has to be substracted to the measured luminosity for stars in the same field. To optimise the JWST observing time, observations with HST are used to identify the Cepheids, also taking advantage of the larger amplitude of variability in the visible. Two JWST epochs are then sufficient to complement the HST observations. Observations were collected by both the SH0ES and Chicago-Carnegie Hubble Program (CCHP) teams (Freedman et al. 2025; Riess et al. 2023). As JWST data are progressively moving to the public domain, it is possible to analyse the images to extract epoch JWST photometry of the Cepheids identified in the HST images thanks to Anton's work. As I already included JWST filters in my templates, it is simple to include JWST photometry to refine the global model of each individual Cepheid, including both visible and NIR precise data. Adding JWST photometry is moreover a good opportunity to improve the template fitting, as NIR F160W data from HST have a lower resolution but are essential for the model to converge.

3.1.3 Period estimation

For the template fitting, the period has two different roles. First, it is used to phase the data and then to constrain the shape of the light curve and the differences in the several photometric bands used. Its estimation is then extremely important but also quite difficult. We can integrate it in the template fitting, which means that when the test period is changed, the template is changed and the other parameters also. This is not a suitable approach for the gradient descent and will make the fit difficult to converge. On the other hand, it is possible to do a preliminary determination of the period to phase the data. In this case, we will not take into account the

shape sketched by the available data to constrain the period, which would have been valuable. The strategy I employed uses a combination of the two ideas, which I will detail in the following part. We rely on other methods to obtain a first estimation of the period for sparse data. I relied on the different methods exposed in Giertych et al. (2024) and investigated two different approaches.

The string-length method is one of the most popular phase-based approaches. It is based on the idea that when phasing data with inaccurate period, it will introduce strong variations in the phase-folded curves, whereas the right period should give a smooth curve. The smoothness is characterised by the sum of the distance between each point in the 2D-plane intensity (the periodic variable) versus the phase. The right period minimises this distance. In the case of Cepheids in NGC5584, we have phase observations in four different filters. In order to keep all the information from these observations, I decided to sum the string-length functions for all four filters in a global string-length function. The string-length function for one filter f with N observations can be expressed with the following formula,

$$L_f = \sqrt{(\tilde{m}_0 - \tilde{m}_N)^2 + (\phi_0 - \phi_N)^2} + \sum_{t=1}^{N} \sqrt{(\tilde{m}_i - \tilde{m}_{i-1})^2 + (\phi_i - \phi_{i-1})^2}$$

where $\tilde{m} = \frac{m_i - \min(m)}{2(\max(m) - \min(m))}$ is a normalised quantity to have the same dimension as the phase. In practice, the string-length function is exploitable only for F350LP and F555W for which the observations are more numerous but adding the two other filters will add flat functions which is not a problem. An example of the string-length function for the four filters is given for the test Cepheid 473829, whose fit is shown in the associated paper, in Fig. 3.12. The period found is sensitive to the spacing of trial periods, which I chose logarithmically spaced.

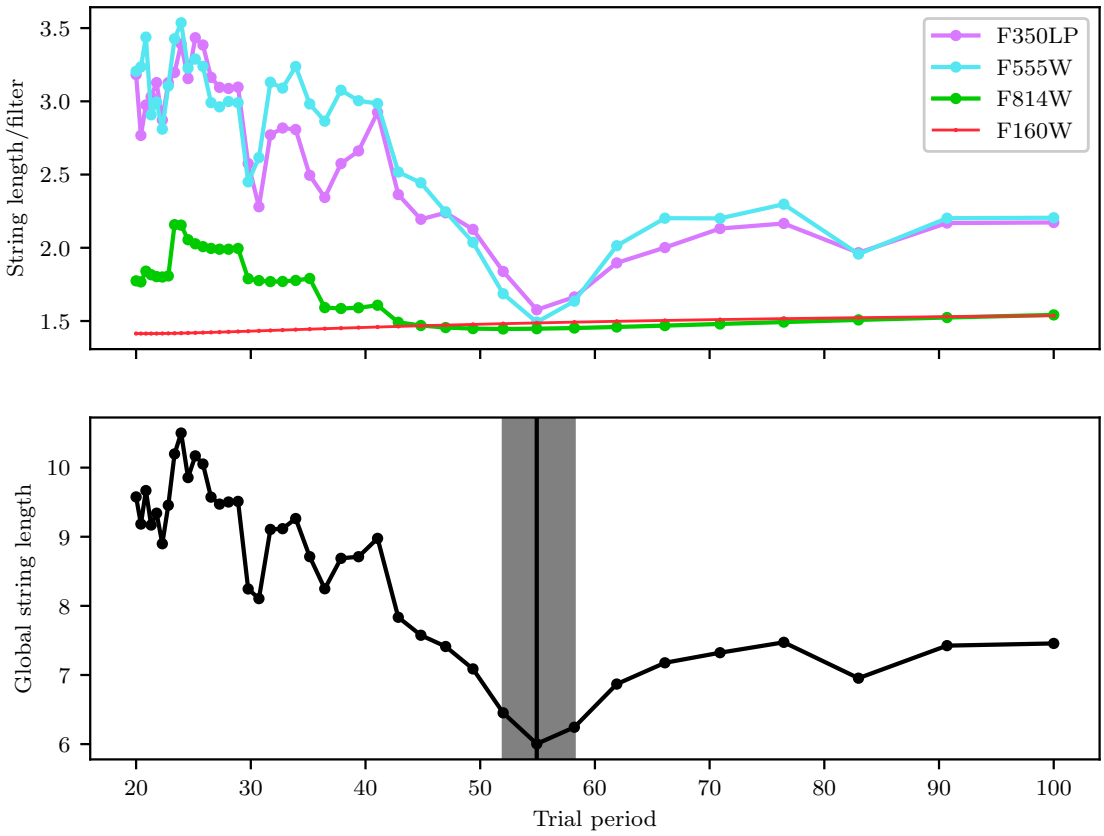


Figure 3.12: String length for the Cepheid 473829 in NGC5584. The optimal period from this method is 54.93 days.

An alternative method is the Lomb-Scargle periodogram. This method relies on the Fourier series basics, and the fact that any periodic signal can be written as a sum of sinusoidal functions. The frequencies of these sinusoidal functions gives the power spectral density. The largest peak in the power spectral density indicates the lower order periodic signals and in our case, will be associated to the pulsation period of the Cepheid. The Lomb-Scargle analysis aims at recovering the power spectral density from sparse data. Similarly to the string-length method, Fig. 3.13 represents the power spectral density of the same Cepheid for the three visible filters and the summed one, mostly constrained by the F350LP and F555W filters. The optimal period is the one that maximises the power spectral density and is the same as for the string-length method for this star, as the initial spacing of trial periods is the same.

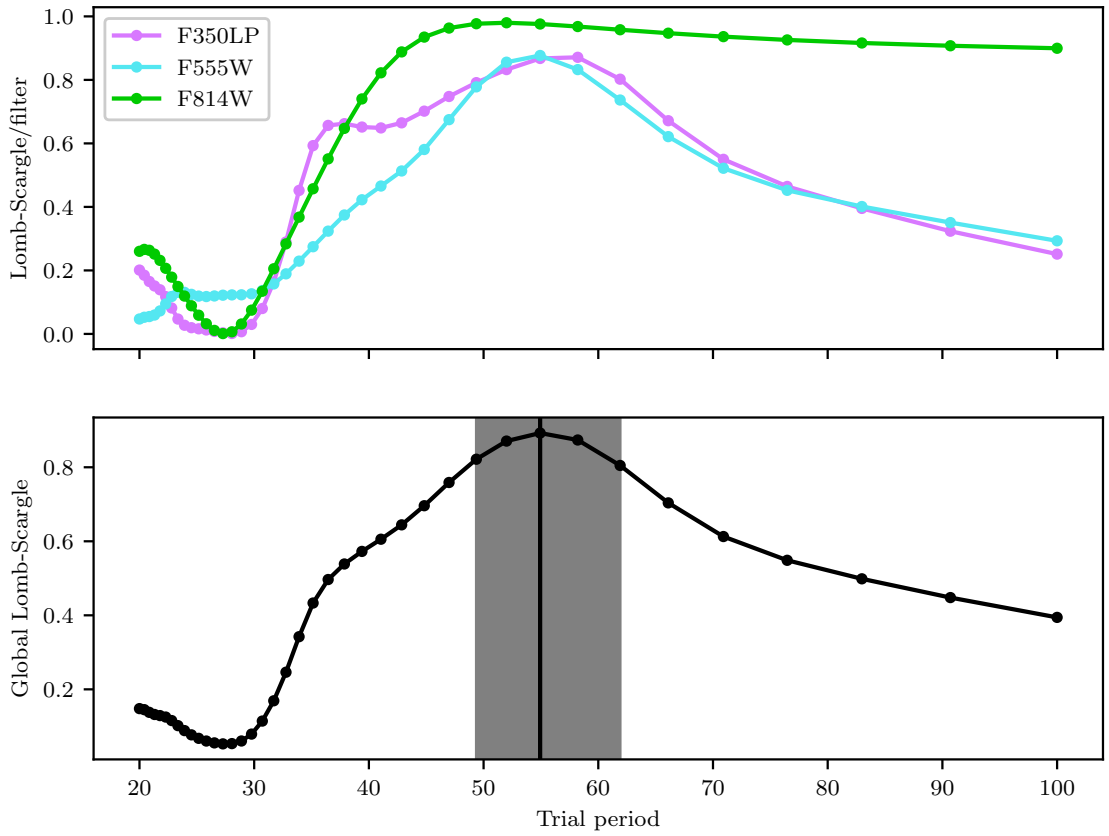


Figure 3.13: Lomb-Scargle periodogram for the Cepheid 473829 in NGC5584. The optimal period from this method is also 54.93 days.

3.2 Template fitting

Once the photometry is obtained for extragalactic Cepheids, it is possible to obtain approximations of mean magnitudes from the mean of the observations, based on the random-phase observations. This means that we assume that overall, the different observations of one Cepheid are taken at totally random phases and that it is unlikely that all measurements are in the maximum branch of the light curve, for instance. The observation cadence in the SH0ES observing programs was chosen specifically to maximise the probability of random-phase observations for the range of typical periods expected for Cepheids at this distance. However, this strategy is introducing uncertainty on the PWRs, and it is more satisfying to obtain a full modelling of extragalactic Cepheids, for optimising our determination of the distance ladder.

3.2.1 SH0ES procedure

I first describe here the procedure for template fitting that is used in recent SH0ES papers, presented as in Hoffmann et al. (2016). The reference work used by the SH0ES team for the templates is published in Yoachim et al. (2009). Their approach starts with a Fourier fitting of the V and I light curves of a large quantity of Cepheids in the Milky Way and the Magellanic Clouds, up to the 8th order. Each interpolated light curve is subtracted from the mean light

curve and the PCA is done on the residuals. They use polynomial fit of the relations between the eigenvalues of the PCA and the period and end up with template light curve depending on the period. In their case, a Cepheid can then be fitted with their template with two phasing parameters (period and phase) and as many mean magnitudes as the studied light curve, which is limited to two. With these four parameters, they are able to reproduce light curves of any Cepheid in the V and I bands exclusively. The V and I band templates are used to fit F555W and F814W observations respectively, thanks to the proximity of their effective wavelengths. For our two different set of templates, we have access to the general shape of the light curve in at least V and I bands for a Cepheid at a given period. I therefore compared the resulting shape, normalizing the templates and removing the phase offset from the maximum luminosity in the given band. I noticed remarkable differences in the shape of the light curve around the resonance period of 10 days which may be due to the different training set, as the behaviour of Cepheids around this period are particular. For a 20-day period Cepheid, the agreement is satisfactory. In addition, I compared the mean magnitudes from the template fitting with my templates and the ones by Yoachim et al. (2009) in the V band. The fitting strategy for the application of the latter templates is the one suggested by the authors. As we see in Fig. 3.14, the mean magnitude is similar with both methods. In this figure, the statistics are computed for the stars for which the fit with the templates from Yoachim et al. (2009) are satisfying. The indicator I used for the quality of the fit was that the period was retrieved, with less than 5% difference. Some fits are not converging to a good solution, which generally results in wrong period measurements and then bad estimates of the mean magnitudes (represented with light green points).

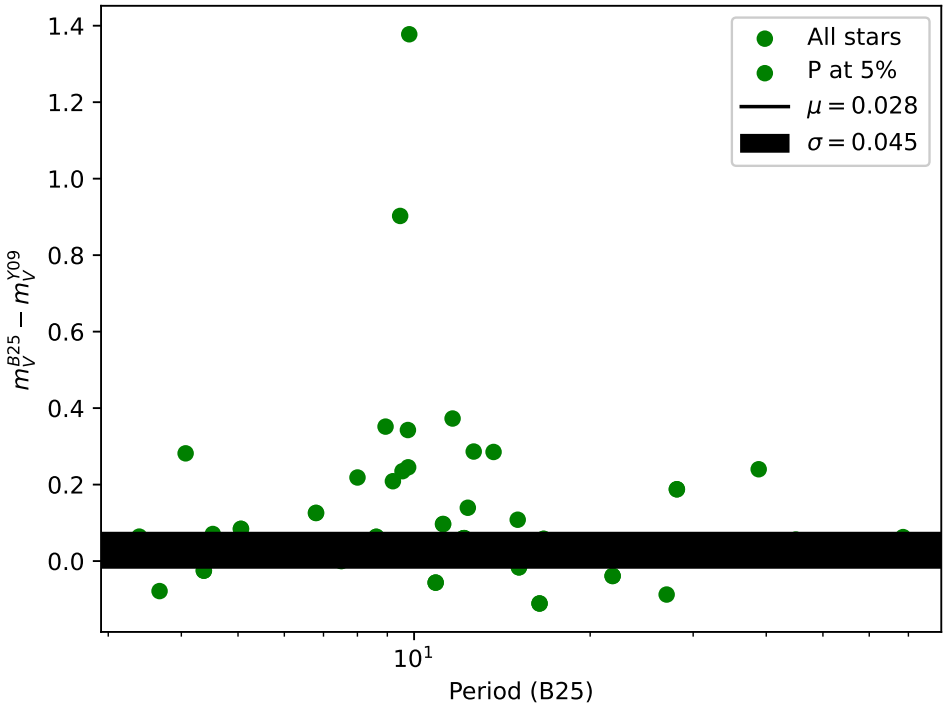


Figure 3.14: Difference in mean magnitude in the V-band in the template fitting from Bras et al. (2025) (B25), compared to Yoachim et al. (2009) (Y09).

In order to avoid bad period determinations when fitting extragalactic Cepheids, Hoffmann et al. (2016) decided to test 150 trial periods between 10 and 100 days. For each of the starting period in this comb, four parameters are then fitted; the period, a phase offset and the amplitude in the two different bands, F555W and F814W. The six bests solutions are then visually inspected to identify potential evidence of bad periods which led to the final determination of the four variable parameters. This implementation is done for every variable source with epoch photometry and the parameters are then used to reject non-Cepheid variables.

3.2.2 New fitting strategy

In our new implementation, we wanted to avoid any kind of visual inspection as it is time-demanding and not reproducible. However, we addressed the difficulties related to period determination, explained in Sect. 3.1.3. I eventually adopted a hybrid strategy between an independent determination of the period, a fit, and a multiplication of the trial periods. In order not to test 150 periods on a very broad range, which is computationaly expensive, I first use the standard period estimators described previously (string-length and Lomb-Scargle) to determine a range of plausible periods. The range of selected periods is illustrated in Figs. 3.12 and 3.13 with shaded areas in the bottom panel. It is broader for longer periods, as long-period variables are less sensitive to the same absolute period difference. In this range, I selected 10 different trial periods for the fit. The next step is to fit the four parameters (period, phase offset, colour excess and scaling magnitude) to all data in the four filters F350LP, F555W, F814W, F160W. The period is locked between the previous and the next trial period. From the evolution of the

χ^2 parameter, I check that we did not miss the real period. If the minimum χ^2 is for the first or the last period tested, I move the trial period window accordingly and test 10 new periods. This avoids relying strongly on the first period estimate. The results of this step are mainly the phasing parameters of the fit (period and phase offset). The observations in the F350LP and F555W filters are more numerous and they dominate the fit. The consequence is that the colour excess and scaling magnitude parameters are poorly fitted, as they are not constrained sufficiently by the F814W and F160W bands, which however contain critical information. The next step is therefore to fit properly the two magnitude parameters, taking into account the longer wavelength data. Inspired by the SPIPS normalisation of the error bars, I compute a reduced χ^2 for each band and minimise the mean of these reduced χ^2 in the fit. Finally, I obtain errors on the magnitude parameters and in the magnitudes in the different bands from random draws of the data within their error distributions. The phasing parameters are fixed during that last step.

3.2.3 Comparison and means of improvement

I presented the comparison of the template fitting from my new templates in the attached paper. In particular, I show the result for one of the Cepheids in NGC5584, using the epoch photometry from Javanmardi et al. (2021). Two additional examples are shown in Fig. 3.15 for Cepheids 549082 and 143986 (identification numbers from Hoffmann et al. 2016). While the fit for the second Cepheid is relatively poor, the observed photometric variability has a particularly low amplitude, which may call into question its classification as a Cepheid. This naturally highlights the potential of template fitting for Cepheid identification. Although this has not yet been formally established, the use of an appropriate numerical indicator of the goodness of the fit could help in diagnosing misclassified Cepheids.

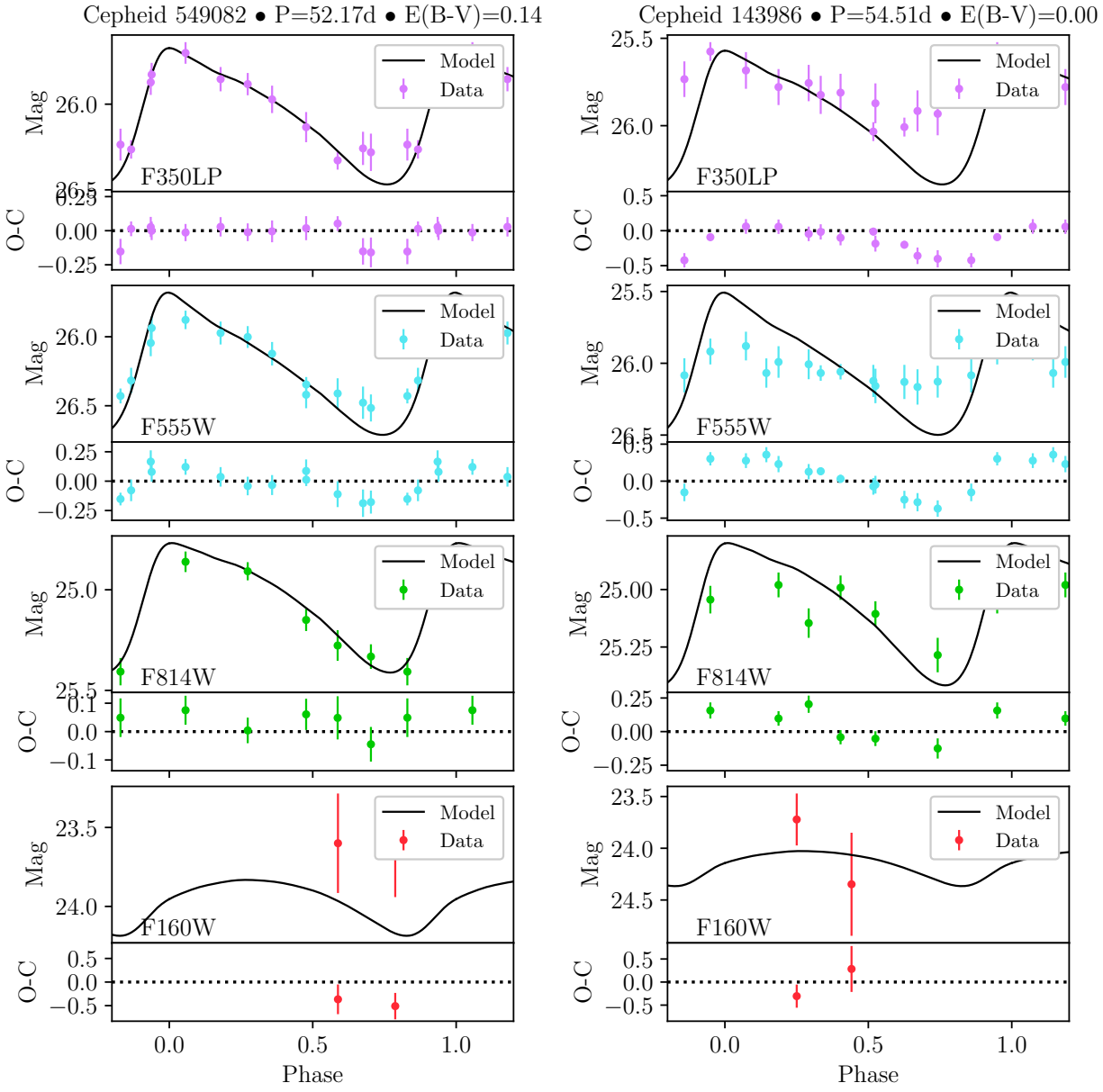


Figure 3.15: Multi-band light-curve fit of two stars identified as Cepheids by Hoffmann et al. (2016) in NGC5584.

The future steps for this work will be to apply it to the new photometry and for all Cepheids in the different galaxies observed by HST. However, our analysis is particularly sensitive to the quality of the F160W photometry. If data are not available in this filter, the results can be strongly biased. For this reason, its application will be greatly improved by adding the JWST photometry. This will be easily done thanks to the inclusion of JWST filters in the template derivation, but will require adjustments in the fitting procedure. Phasing the JWST data to HST data requires an accurate determination of the period and can make the fit difficult to converge. Alternative strategies are presented for instance in Riess et al. (2023).

4 Paper II

My second A&A paper as first author on the development of light-curve templates for Cepheids is attached hereafter. Essential details on the efficiency of the templates can be found in the following pages and is not included in the previous presentation. In this Chapter, I mainly expose the context that motivated the development of such templates.

A&A, 698, A97 (2025)
https://doi.org/10.1051/0004-6361/202553783
© The Authors 2025

Astronomy & Astrophysics

Versatile light curve templates of Cepheids

Multi-band fitting of photometric time series

Garance Bras[1,⋆], Pierre Kervella[1,2], Antoine Mérand[3], Anton Afanasiev[1], Vincent Hocdé[4], Nicolas Nardetto[5], Alexandre Gallenne[6], Grzegorz Pietrzyński[4], Manon Bailleul[5], and Wilma Kiviaho[1,2]

[1] LIRA, Observatoire de Paris, Université PSL, Sorbonne Université, Université Paris Cité, CY Cergy Paris Université, CNRS, 5 place Jules Janssen, 92195 Meudon, France
[2] French-Chilean Laboratory for Astronomy, IRL 3386, CNRS and U. de Chile, Casilla 36-D, Santiago, Chile
[3] European Southern Observatory, Karl-Schwarzschild-Straße 2, 85748 Garching, Germany
[4] Nicolaus Copernicus Astronomical Center, Polish Academy of Sciences, Bartycka 18, 00-716 Warszawa, Poland
[5] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Nice, France
[6] Instituto de Alta Investigación, Universidad de Tarapacá, Casilla 7D, Arica, Chile

Received 16 January 2025 / Accepted 27 March 2025

ABSTRACT

Context. To use the period–luminosity relations of Cepheids to derive distances of extragalactic Cepheids, it is crucial to have reliable estimations of the luminosity of such stars. However, the light curves (LCs) of extragalactic Cepheids are sparsely sampled, which makes the interpolation difficult.
Aims. Our purpose is to create general templates of Cepheid LCs that can be efficiently applied with only a few parameters, while ensuring they are reliable in terms of resulting mean magnitudes. We have also aimed to provide adaptable templates for the large variety of photometric bands to enlarge their possible applications.
Methods. We performed a principal component analysis (PCA) on LCs in different bands from the optical to the mid-infrared (MIR) fitted by atmosphere models along the pulsation cycle for 75 galactic Cepheids. All the photometric bands were treated together to incorporate the results of the effective temperature variation. To link the different shapes of LCs, we also derived the relations between the period and the coefficients associated to the PCA.
Results. We present our templates and a further analysis to validate their reliability. We also offer an example of their application to one Cepheid in the galaxy NGC 5584. We find a good agreement with previous estimations of mean magnitudes in the HST filters. We show examples of a template fitting for Cepheids in various environments to demonstrate the variety of applications of these templates.
Conclusions. The templates we propose here can be used in a number of different situations (good phase coverage in many filters, poor phase coverage in one specific filter, but good phase coverage in some others, or noisy and sparse data), thereby mobilising the adapted fitting strategies. Moreover, the broad adaptability of this LC prediction tool is provided by the multiple photometric bands in which they can be inferred.

Key words. methods: data analysis – techniques: photometric – stars: distances – stars: variables: Cepheids – distance scale

1. Introduction

Determining distances is one of the key interrogations across the field of astrophysics. For this purpose, a large variety of approaches and indicators are available, which depend on the order of magnitude of the distance. However, a lot of distance measurements rely on the so-called global distance ladder. This ladder consists of achieving distances to observable galaxies by stepping through different calibrations of various indicators (see e.g. Gaia Collaboration 2023; Pietrzyński et al. 2019; Reid et al. 2019; Leavitt & Pickering 1912; Phillips 1993, for trigonometric parallaxes, eclipsing binaries, mega-maser, Cepheids, and Type Ia Supernovae, SNIa, respectively).

This paper is focussed on multi-band light curve (LC) templates for the determinations of periods and mean magnitudes of Cepheids. When they are in other galaxies, such stars are observed e.g., using image series from the *Hubble* Space Telescope (HST) and the *James Webb* Space Telescope (JWST). However, with such facilities, the number of measurements per star is very restricted. It is important to have a robust procedure to analyse the sparse photometry of extragalactic Cepheids and extract the main properties without any loss of information. Fortunately, a range of Cepheids in the Milky Way (MW) has been observed extensively and are well-characterized. Cepheids are useful objects as their LC properties, shape, amplitude, and phase variation of the different features depend mostly on their pulsation period. We here take advantage of these properties.

This idea has already led to the use of LC templates, that is to say, generic LCs created on the basis of a known period of a Cepheid, which have been derived by several teams in the past. For instance, Riess et al. (2022) made use of the templates derived by Yoachim et al. (2009). These templates are specific to two photometric bands: the V and I bands. Their work relies on a principal component analysis (PCA) of Fourier interpolations of the LCs of MW, Large Magellanic Cloud (LMC), and Small

⋆ Corresponding author: garance.bras@obspm.fr

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Subscribe to A&A to support open access publication.

Magellanic Cloud (SMC) Cepheids in these two photometric bands. For a given pulsation period, specific coefficients were then derived to estimate the LC shape in both bands. These templates were used to derive the main properties of extragalactic Cepheids in Hoffmann et al. (2016).

Other works have been performed to derive LC templates. Stetson (1996) introduced continuous relations between Fourier parameters and pulsation period. Ngeow et al. (2003) kept this idea together with interrelations between Fourier coefficients in different photometric bands. It allowed them to use fewer parameters and, specifically, to apply them to extremely sparse data in the I-band. In a related approach, Inno et al. (2015) used the mean LCs in specific period bins with Fourier amplitudes and phase relations between well-sampled visible LCs and single-epoch measurements in near-infrared (NIR) LCs to predict the LCs in the latter filters. In Chown et al. (2021), the authors derive scaling and phasing relations between LCs in different photometric bands and a mean template LC, but removed the impact of the period in their templates. Recently, Madore et al. (2024) studied the specific case of dealing with single-epoch observation in one band, using good phase coverage in at least another band, following an approach first implemented by Freedman (1988). The PCA approach was first used by Hendry et al. (1999) using the Fourier interpolation parameters in both V and I band as a training sample. They also made use of relations between Fourier parameters to be able to find mean magnitudes in other filters. The method was used by Kanbur et al. (2002) to understand the structure of the LC and its changes with the pulsation period. The PCA was then used by Tanvir et al. (2005) and Yoachim et al. (2009), based on the Fourier parameters in the standard V and I band. In these works, the correlations between the LC shapes in both bands are encoded in the PCA. However, mean magnitudes in those bands are still fully independent. In order to decompose the LC shape of Cepheids, Freedman & Madore (2010) and Pejcha & Kochanek (2012) introduced a physical model that extracts photometric curves from variation of the radius and the temperature. These models have the advantage that they are applicable to other filters than the ones used for the training in PCA, while making it possible to reduce the dimensionality of the fitting procedure.

Our work encodes the link in the LC shapes as well as the magnitude differences in the different filters, while being trained on parallax-of-pulsation models. It is an extension of a work that was presented in Javanmardi et al. (2021). What we looked for eventually is the ability to find the main properties of a Cepheid with the smallest possible number of parameters to fit and we kept only the uncorrelated parameters. We were also motivated to produce templates that could be applied to any new facility with new calibrated filters easily, without any preliminary observations in those new filters. In addition, the encoding of relations between the different filters, which is correlated to the temperature of the star, enables us to determine the reddening. It is an interesting choice to bypass the use of specific colour combinations to empirically compensate for the effect of the reddening. In our work, the relations between the LCs in the different photometric bands are constrained by the underlying atmosphere models, at each pulsation phase. This approach guarantees some versatility in the produced templates, as they can be produced for any photometric band, knowing its transmission curve.

This paper is organised as follows. We present our PCA implementation in Sect. 2, followed by the procedure to generate templates in Sect. 3. Then, a validation is provided in Sect. 4 and an application in Sect. 5. Finally, we discuss this approach and its results in Sect. 6. We provide our conclusive remarks in Sect. 7. The essential data to reproduce the analysis are provided on Zenodo, with parallax-of-pulsation models of the stars in the training sample and the output of the PCA. An example of how to use it to generate a template LC is presented in Appendix A. We present in Fig. 1 a schematic view of the different steps detailed in this paper to present a general overview. Details of each step are provided in the text.

2. Performing a PCA

To build LC templates for Cepheids, we used the PCA approach. This method is particularly suited to such a study as the shape of the LC is mainly driven by the mass of the Cepheid (Balázs & Kovács 2025). The PCA allows us to extract the main shape of the curve and characterise it using a small number of components, with as many parameters. A mathematical description of the principle of PCA can be found for instance in Deeming (1964). In short (and in the context of the application to LCs), the PCA can be used to find a proper base of function in which a given training set can be decomposed. These functions are orthogonal to avoid redundant information and by ordering them by their order of magnitude in the training set, we can choose to reduce the dimension of the base of functions to keep the most relevant parts. In this work, we refer to the different base functions (or base vectors) as C_i, with i being an integer from 1 to the dimension of the PCA, which depends on the training sample. We set the C_0 component to be the mean function of the training sample, which will always be weighted similarly. Each vector, f, can then be represented as a linear combination of the base vectors, $(C_i)_{i\in\{0,\dots,k\}}$, as follows

$$f(x) = C_0(x) + \sum_{i=1}^{k} K_i C_i(x). \tag{1}$$

In the following, we explain how we determined the set of vectors that best describes the range of LCs.

In this analysis, we base our template generation on various datasets. We defined three particular samples, named 'gold' for the training sample of the PCA, 'silver' for the sample used to derive the relations between the period and the shape of the LC, and 'bronze' for the main validation sample. Each sample is composed of a specific selection of Cepheids depending on the available observational data. The use of these samples in our PCA implementation is visually described using colour-coded boxes in Fig. 1.

2.1. Training datasets

We trained our PCA in all filters of interest using previous pulsation parametrisation of Cepheids. We made use of the Spectro-Photo-Interferometry for Pulsating Stars (SPIPS) results of a set of 75 Cepheids. The details of the SPIPS analysis method can be found in Mérand et al. (2015). The general idea was to combine observations from different techniques: spectroscopy for radial velocity (RV) and effective temperature, interferometry for angular diameter, and photometry in different passbands. We fit them using a phase dependent pulsation model. The parametrisation is composed of an RV curve, an effective temperature curve and a set of stellar parameters. Atmospheric models (from Castelli et al. 2003) were used to compute the synthetic LC in the different filters. Any filter can be used, providing its wavelength passband is known. This allows us to obtain photometric LC predictions for the modelled stars in every filter, including filters for

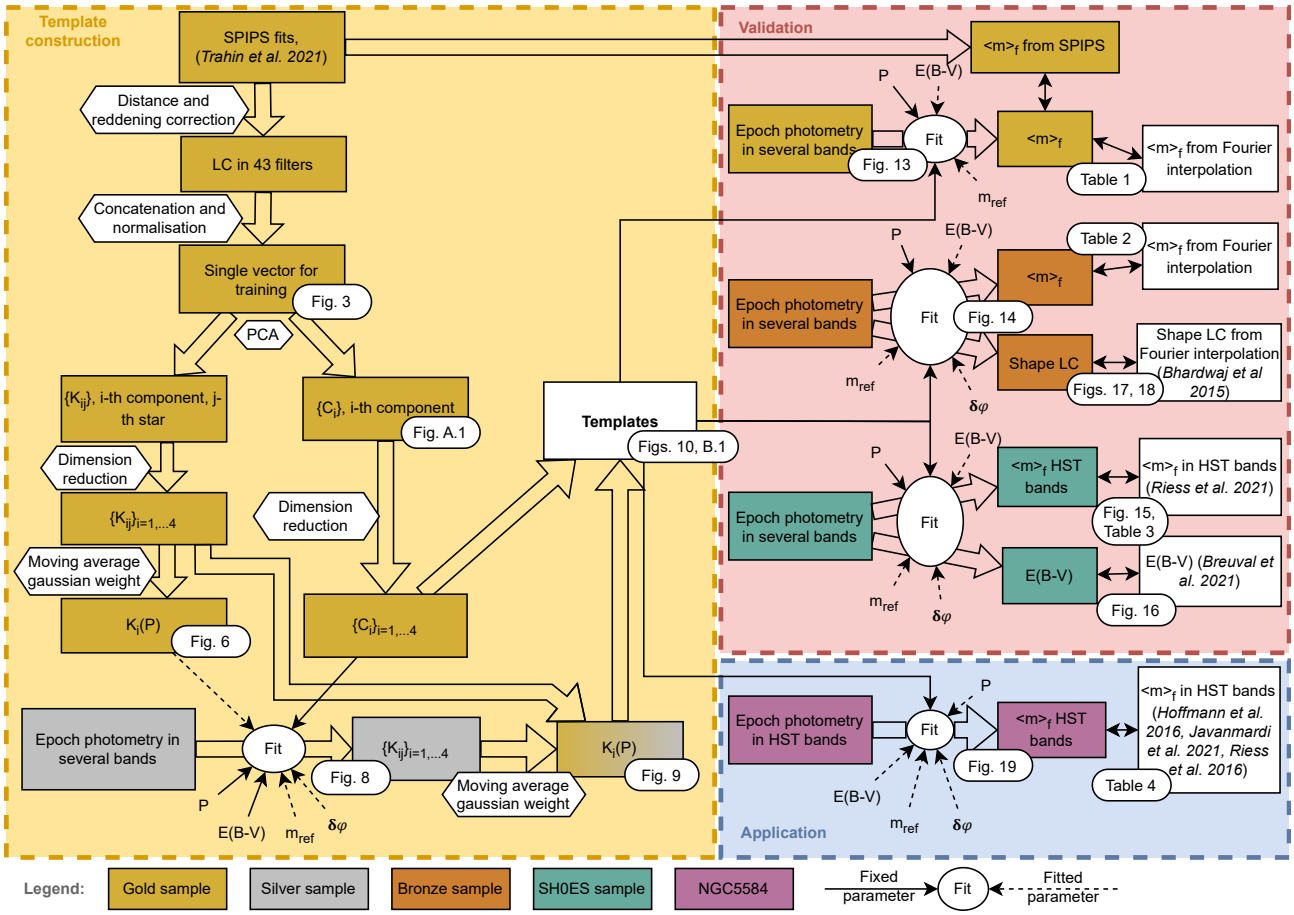


Fig. 1. Graphical representation of the entire process described in this paper.

which no observations is available. This is a strong advantage of the SPIPS tool, as Cepheids are observed in various photometric systems, but the heterogeneity of the photometry can be a problem when calibrating from one system to another. Instead, here, we are considering these various observations simultaneously, allowing us to cross-check their consistency. The use of the SPIPS parametrisation is particularly powerful as the predicted LC depends on interpolations using Fourier series of the radial velocity and effective temperature curves, which provides fits to all the observables. The predicted LCs show less unphysical behaviour, even in bands where the phase coverage is sparse. In this paper, we sometimes refer to these results as pulsation models, but it has to be noted that they are distinct from hydrodynamical pulsation models. This is only a mathematical parametrisation, which allows us to reproduce the different observables in the pulsation of a Cepheid.

To train the PCA, we accumulated the gold sample comprising of 63 Cepheids whose χ^2 values for the fit are low and their observed curves are well-reproduced from the large sample presented in Trahin et al. (2021). This sample was supplemented by 12 other fits that were considered to be of sufficient quality. We were particularly careful to check the satisfactory reproduction of the LCs in the different passbands that does not introduce spurious features. The list of stars used for the training sample is given in Table E.1, with the references of the data used for the SPIPS parametrisation. An example of these additional fit results is given in Fig. B.1. We tried to find a balance between the periods of the stars in the gold sample; however, a reliable SPIPS fit relies on precise and well-sampled photometry and spectroscopy

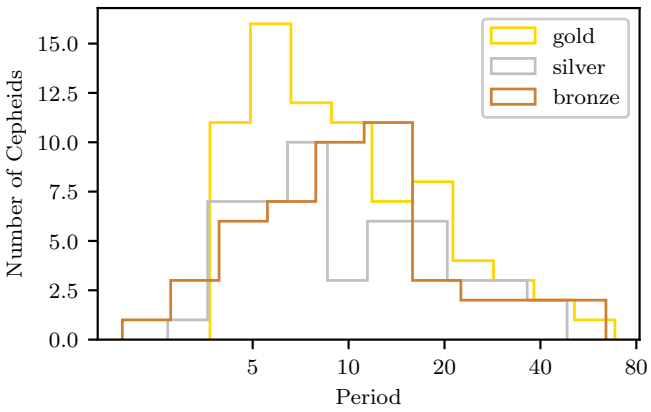


Fig. 2. Histogram of periods of Cepheids in the samples described in this paper.

(at least) and preferably with interferometric measurements. For this reason, our sample includes a lot of nearby Cepheids, with quite short periods, which we chose to keep to avoid reducing the size of the training sample. The distribution in terms of the period of the training sample is shown in gold in Fig. 2. In Fig. 1, we synthesise the steps of the processing for the Cepheid LCs in the gold sample (gold boxes).

To gain efficiency in the analysis of extragalactic Cepheids and versatility for the LC fitting of Cepheids in general, we aimed to generate templates that combine information from all

photometric observations at one time. This can be done by performing only one PCA on all filters at the same time, which is possible by generating photometry in any given list of filters for any star in the gold sample. In a technical aspect, we concatenated all LCs in the different filters in one single vector for the PCA (see Fig. 3). In the process, the mean temperature of the Cepheid is intrinsically incorporated when we perform the PCA, taking into account the differences in magnitude in the different bands. This method also allows us to drastically limit the number of parameters in the final templates. For this work, we arbitrarily chose a set of filters commonly used in ground-based observations for galactic Cepheids (*BVRI* Johnson, *JHK*) to validate our templates, but also HST and JWST filters to apply the templates to extragalactic Cepheids, which can be extended. We computed the synthetic LC for every star in our gold sample in the same selection of 43 filters. The list of the filters with the adopted mean wavelengths and labelled with the name that should be used when applying the templates is provided in the attached file. We include the *V* filter in the Walraven system, which is not in the same magnitude scale as the other filters. However, thanks to the numerical approach of the PCA, we have been able to use the direct magnitudes observed in this system, without recalibration. The change of scale will be encoded in the difference in magnitudes between this filter and the other filters, especially other filters in the *V* band.

All synthetic LCs were corrected for their apparent distance moduli, using distances from *Gaia* DR3 (Gaia Collaboration 2021; Lindegren et al. 2021) and from the reddening. The SPIPS parametrisation corrects for the reddening taking into account the temperature of the star at each pulsation phase. Generally, the correction is applied by integrating the A_λ parameter over the photometric passband ponderated by a standard Vega stellar spectrum. Because Cepheids are much cooler than 10 000 K, this can introduce an inaccurate correction from the interstellar dust, particularly in the broad filters. In practice, as observed by Trahin et al. (2021), the SPIPS approach of the reddening correction will convert into larger values for the colour excess $E(B-V)$ than reported in the literature. For our templates, we simplified the SPIPS treatment by considering a constant temperature of 5500 K for all Cepheids during their whole pulsation phase. This is an approximation, but the variation induced in the reddening correction is very small compared to the variation with the classical approach of considering that all Cepheids have a constant temperature of 10 000 K. For this reason, all colour excesses that will be presented here are not in the same scale as the usual values for $E(B-V)$ in the literature (for instance those collected by Breuval et al. 2021). We can estimate, however, how they relate to each other to compare and apply the literature estimations, from the following formula and using the Fitzpatrick (1999) reddening law,

$$\frac{E(B-V)_{5500\,\mathrm{K}}}{E(B-V)_{10\,000\,\mathrm{K}}} = \frac{A_B(5500\,\mathrm{K}) - A_V(5500\,\mathrm{K})}{A_B(10\,000\,\mathrm{K}) - A_V(10\,000\,\mathrm{K})} \simeq 1.16. \quad (2)$$

A photometric correction for circumstellar environments (CSEs) is also possible in SPIPS, using a power law parametrisation. The photometric contribution from such an envelope has been preserved in the derived LCs. Indeed, we used the LCs that offer a good match to the observations, as extragalactic Cepheids can also have CSEs (Gallenne et al. 2017). Finally, the vectors used in the PCA are normalised individually for each star, so the concatenated LC has a zero-mean value. This results in one unique additional parameter to reproducing the observations from such decomposition. We plotted the vector that corresponds to the long-period Cepheid RS Pup in Fig. 3 as an example. In this figure, the magnitude level of each sawtooth is important and is directly related to the temperature of the Cepheid only and specific to an unreddened star. Only the concatenated curve has been normalised, which explains why when looking at specific filters, the mean value will not be zero.

Our selection of stars is only composed of Galactic Cepheids, with solar-like metallicities. As it is possible that the metallicity of the star changes the shape of its LC, identified through the distribution of Fourier interpolation coefficient by, for instance, Hocdé et al. (2023), the present templates are only suitable for solar-like metallicities. An extension for metal-poor stars is discussed in Sect. 6.

2.2. Analysis

The PCA was performed using the implementation in the scikit-learn Python library (Pedregosa et al. 2011). The synthetic LCs are sampled at 0.001 in phase for all filters, while 75 components are kept in this computation to account for the variations in the 75 stars.

To reduce the number of dimensions, we extracted the components that account for the major differences in terms of shape. We assumed that the presence of peculiar small bumps that appear randomly in the LCs of certain stars will be represented in minor components; this is not crucial in our templates as these features are not important for the determination of mean magnitudes.

We retained the first four components, which explain 96.5% of the variance in the training sample, as shown in Fig. 4. This is similar to the findings of Javanmardi et al. (2021), where four components explained 97.1% of the variance, with a smaller sample and long-period Cepheids only. The large significance of a small number of components is expected from the SPIPS parametrisation, which is mainly encoded in the variations of the star's radius and temperature. Beyond the fifth component, variations in the LC shape become particularly small, while encoding smaller features such as bumps. Also, the coefficients associated to the fifth component does not seem related to the period. We have aimed to retain only the variance related to the period to obtain meaningful templates. Including more components would be useless, as the interpolation of a coefficient to the period compensates for the added information. When adding the fourth component, the induced variation in mean magnitudes is smaller than 0.1 dex for every filter used, which explains why we stopped at four components.

Following the vectors in the training sample, the components are a concatenation of the LCs in the different filters. To obtain the components in a specific filter, we only need to take a truncation that corresponds to the given filter. This was done in Fig. C.1, where we plot the mean curves and the four first components for the *V* and *K* filters.

2.3. Coefficient distribution

It is interesting to find links between the empirical coefficients of the PCA and physical properties using the stellar parameters derived by SPIPS. We note that the effective temperature of the star constrains the coefficient of the first PCA component, as shown in Fig. 5. The temperature is the most important parameter as it drives the difference in magnitude among the different filters, which means that the first component generally defines the mean magnitude for each filter. Its close correlation with the luminosity and, therefore, the period of the star is crucial to building templates characterised by the period only.

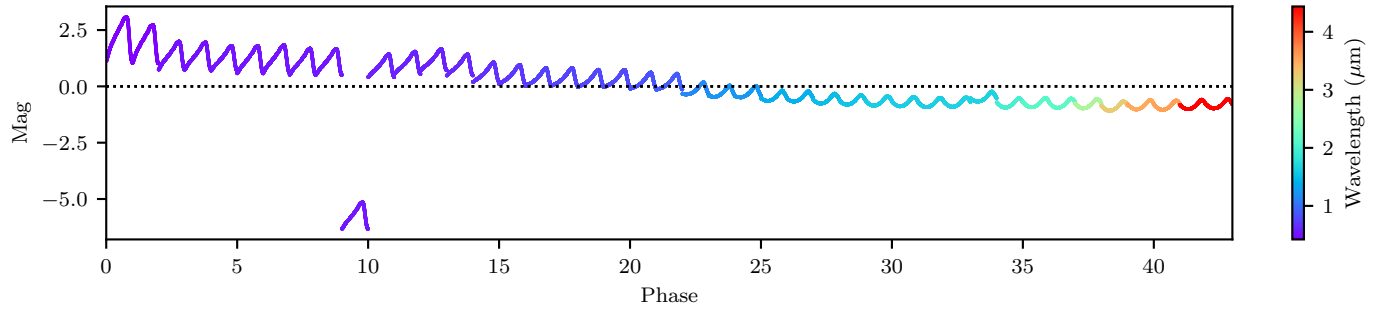


Fig. 3. Concatenated LC of RS Pup as one vector of the training sample for the PCA. Each sawtooth corresponds to one filter, colour-coded by its central wavelength. The LC around −6 in magnitude corresponds to the *V* Walraven system, which explains why it scales so differently from the other bands.

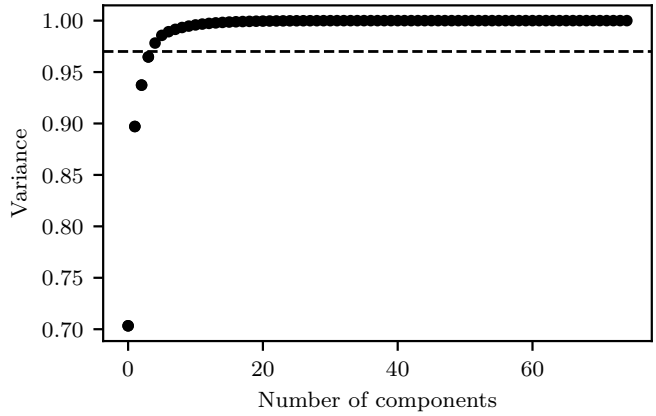


Fig. 4. Explained variance compared to the number of components considered. The dashed line corresponds to 97% of the variance of the training sample reproduced, acting as our threshold for the templates.

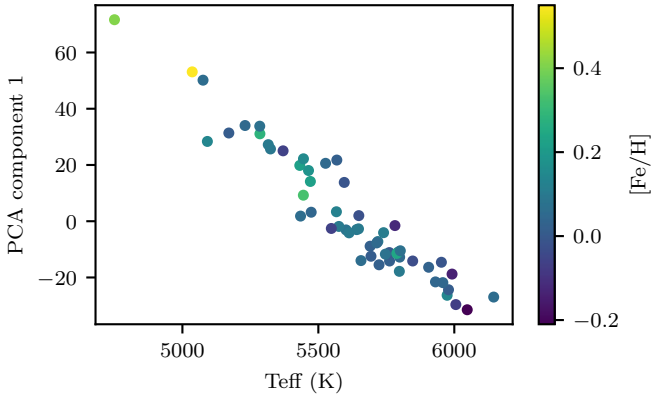


Fig. 5. Distribution of the coefficient associated to the first component with respect to the effective temperature of the star, derived by SPIPS. The metallicity from Breuval et al. (2021) is represented in colours.

3. Generation of templates

3.1. Relation between the period and the coefficients

To define our set of templates, we need to predict the coefficients associated with the PCA from the period only, as it is the other parameter required to produce the PL relation. Deriving linear relations does not seem appropriate for each of the four components. However, for each period, we can estimate the probable values for the coefficients of the PCA from the mean value of the distribution of the coefficients of stars with similar periods. This leads us to define relations by using a moving average with a Gaussian weight. The standard deviation of the Gaussian weight was tuned manually and set to 0.07 in logarithmic scale to account for the variations of the coefficients without introducing unrealistic oscillations in terms of periods. From this method, we also estimated a confidence interval for the final templates from the dispersion of the coefficients for a certain period range. The distribution and the corresponding relations are shown for the four main components in Fig. 6. This averaging method will lead to smooth the variety of LCs in a same period range, which means averaging over the width of the instability strip and, in particular, they will average over the range of amplitudes of a given period bins as the amplitude of the LC of a Cepheid is not purely correlated to the pulsation period. As it could then be biased by the distribution of the instability strip, we plot in Fig. 7 the positions in the Hertzsprung–Russel diagram for the 75 stars in the gold sample.

Our relations are currently defined over a period ranging from 3.7 to 68.6 days, limited by the stars in our training sample. In nearby galaxies, it is common however to observe Cepheids with longer periods. Without any longer period Cepheid to increase this range, we used an extrapolation for relations between the period and the coefficients, assuming that the trend followed by these relations is also followed for longer periods and that the small number of long-period Cepheids considered here are representative.

3.2. Addition of a supplementary dataset

We were motivated by the lack of SPIPS results for very-long-period Cepheids to expand our dataset by including stars that are not in the gold sample, but for which good photometry was available from the literature. For this additional dataset, the silver sample, we made use of good photometric observations of Cepheids for which the SPIPS tool was not able to converge (due to a lack of spectroscopic data, e.g. either temperature or RV). With no model for the radius and the temperature of the star, we were not able to build the LC in any filter that we liked. Hence, we could not use this dataset as a proper training sample for the PCA, but we can use the fully phase-covered photometric data, as shown in the silver boxes in Fig. 1.

The silver sample is composed of a selection of 48 stars from the phase coverage of their LCs in various different filters, including IR filters. The main properties and references of datasets for the stars in this sample are tabulated in Table E.1. Photometric observations of stars in the silver sample were

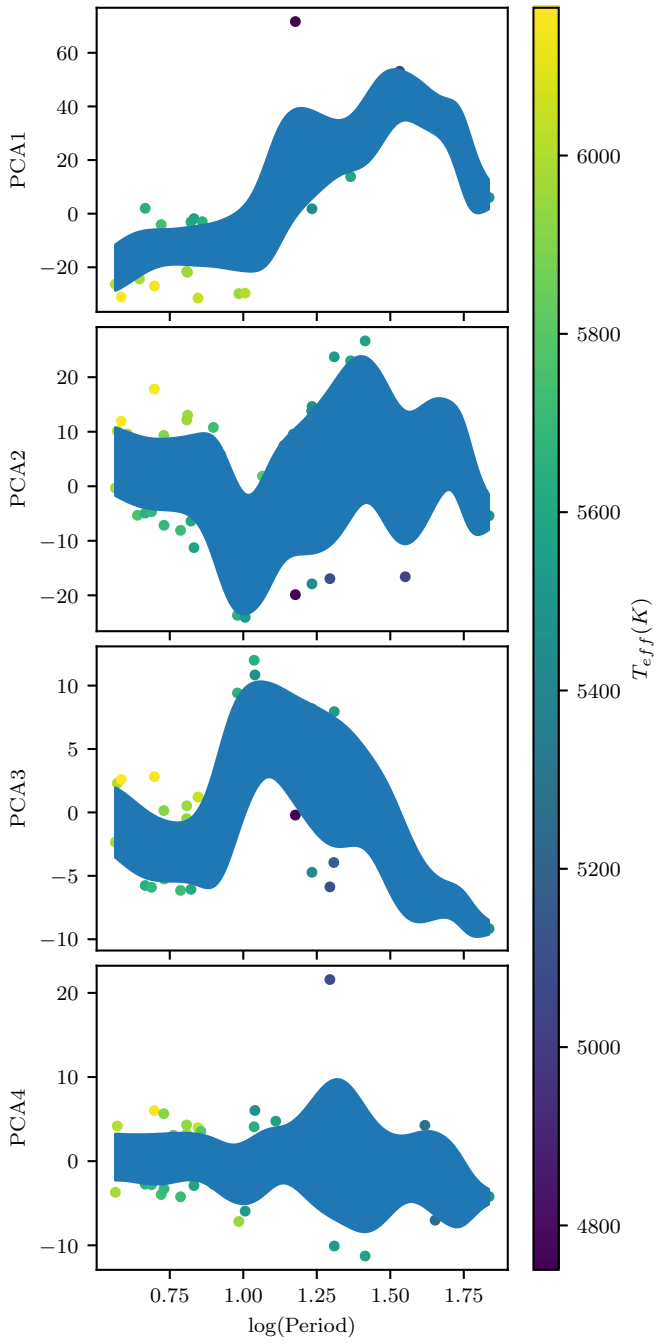


Fig. 6. Distribution of the coefficients associated to the four main components of the PCA as a function of the period, with the relation from a moving average with a Gaussian weight. The shaded area is the confidence interval derived from the standard deviation of the coefficients.

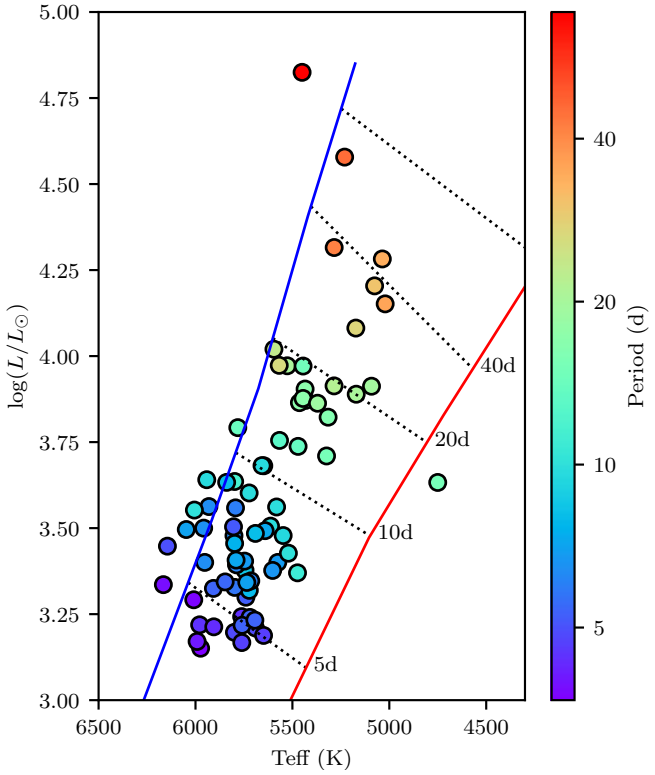


Fig. 7. Distribution in the instability strip of the 75 stars in the gold sample. The edges of the instability strip and the predicted periods are plotted according to Anderson (2016), for the first crossing, with a rotation of $\omega_{\rm ini} = 0.5$ and a solar metallicity of $Z = 0.014$ for comparison.

fitted using the principal components, C_i, obtained from the gold sample, leaving the associated coefficients, K_i, free. The colour excess is fixed to literature determinations, rescaled with our factor of 1.16 as defined in Eq. (2), to break the degeneracy with the temperature and, hence, the coefficient K_1. In SPIPS, data are phased taking the maximum bolometric luminosity as a reference. This quantity is generally unknown for the stars in the silver sample as it relies on the global pulsation model, which we do not have. We thus included an offset parameter to take into account the difference in the reference date in fraction of period in the processing of the photometry of the stars in the silver sample. One unique parameter for all filters is used to scale the luminosity of the star, which we call the reference magnitude, $m_{\rm ref}$. It acts as a global additive offset in magnitude and takes into account the normalisation of the vectors in the PCA and the distance modulus of the star.

The detailed parametrisation is the application of the PCA, as described by Eq. (1), where the f function is the magnitude, offseted by the reference magnitude, $m_{\rm ref}$, is given in Eq. (3), with $C_0^\lambda(\phi)$ being the mean of the PCA and $C_i^\lambda(\phi)$ the i-th components of the PCA in the filter at wavelength λ, coming from the PCA on the gold sample described previously. The components are associated with the coefficients K_i (independent of the filter). Here, A_λ is the correction from the interstellar extinction given the value of $E(B-V)$ in the filter λ and $d\phi$ is the phase offset to correct the date for the zero-phase. The m_λ curve is only a truncated part of the whole vector containing all the filters,

$$m_\lambda(\phi) = m_{\rm ref} + C_0^\lambda(\phi + d\phi) + \sum_{i=1}^{4} K_i C_i^\lambda(\phi + d\phi) + A_\lambda. \quad (3)$$

From this parametrisation, we set $m_{\rm ref}$, the four K_i, and $d\phi$ as free parameters, fitted to the data in all photometric bands simultaneously. We give an example of such a fit for the Cepheid SW Vel in Fig. 8. Each and every fit has been visually checked in order to keep the ones for which the LCs are well reproduced. The rejected ones were moved to the bronze sample (see Sect. 4). This sample allows us to increase the statistics to compute relations between the period and the coefficients associated to the PCA (hereafter, the P–K_i relations). We used the same

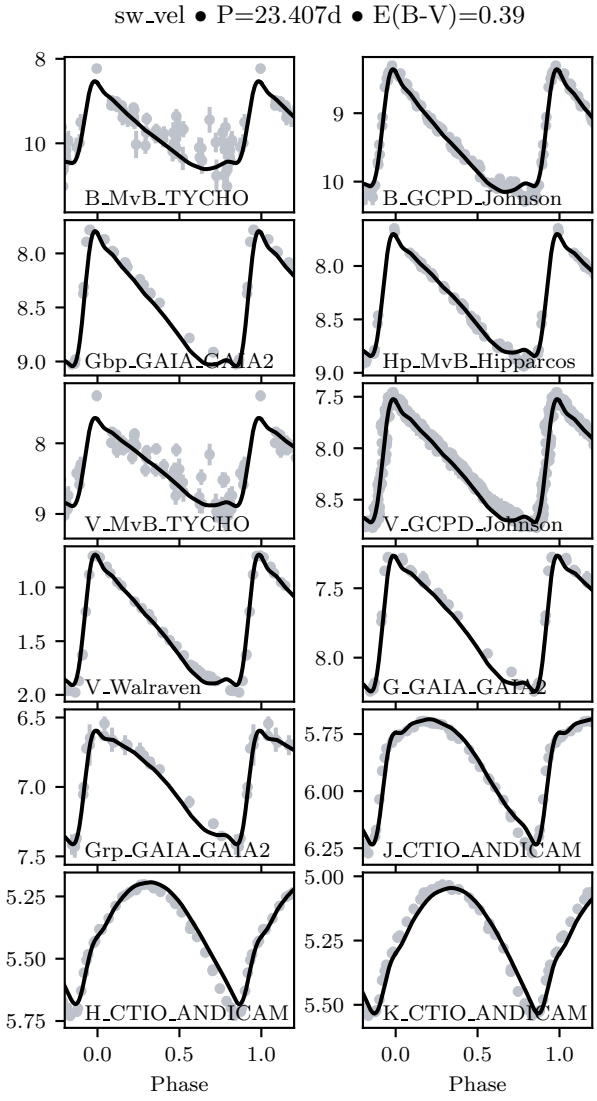


Fig. 8. Fit of SW Vel in the silver sample. The period and the $E(B-V)$ are fixed. The predicted curves for several filters are shown in grey lines, the observations in silver-coloured points.

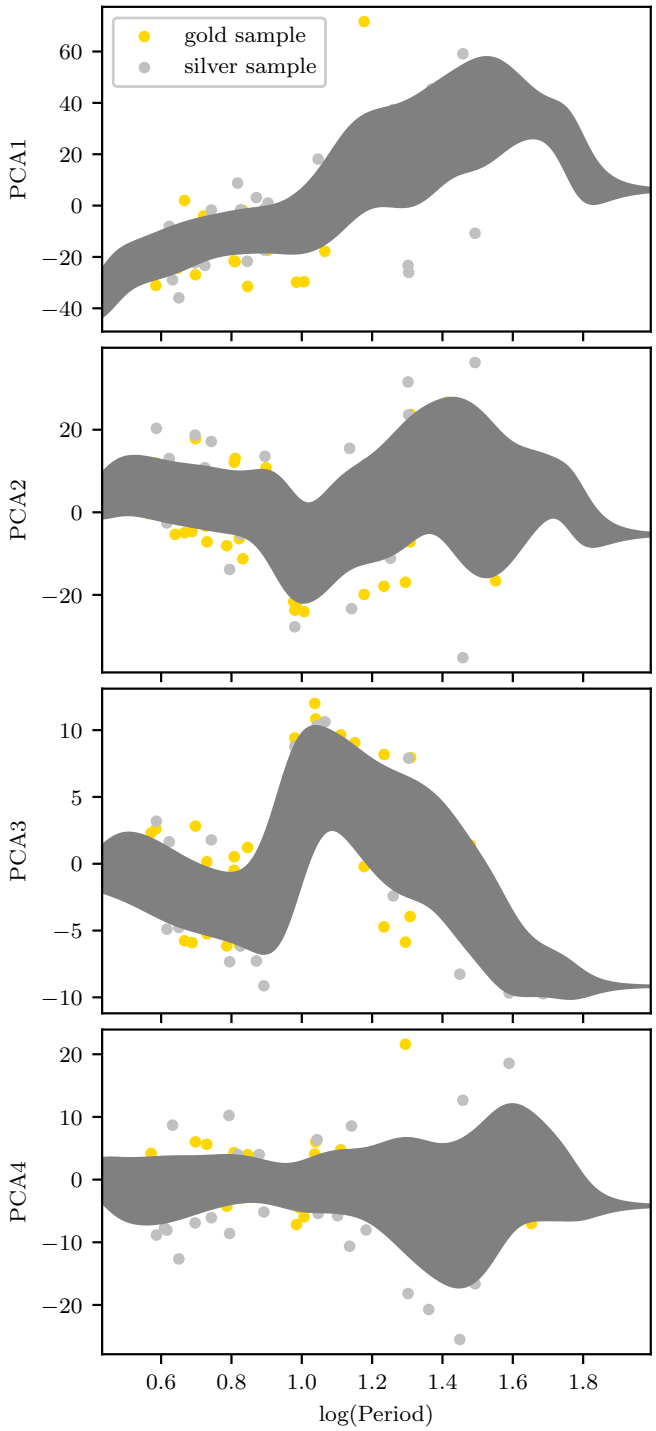


Fig. 9. Distribution of the K_i for the gold and silver samples. The moving average with a Gaussian weight and the confidence interval from the total distribution (gold and silver samples) are also plotted similarly as in Fig. 6.

moving average method as previously described, taking into account both gold and silver stars. We show them in Fig. 9.

3.3. Templates and Hertzsprung progression

The derived $P-K_i$ relations, together with the principal components C_i, allow us to predict the shape of the LC for a Cepheid of any period in any filter. An example of such a template is plotted in Fig. 10. This template is provided for the common V Johnson filter for a 20-day-period Cepheid, and J, H, and K CTIO filters are shown in Fig. D.1. The confidence interval shows the possible range for the particular features that we could observe for a Cepheid of this period. It does not represent a general offset of the curve.

These templates naturally allow us to study how the LC shape evolves with the Cepheid's period in any filter. We will focus on the common V Johnson filter, since the shape of the LC changes with the wavelength of the filter, as reported by Bhardwaj et al. (2015) in the context of the Hertzsprung progression. We plot different templates for periods between 6 and 40 days in Fig. 11 to show the evolution in the skewness of the LC. Around 10 days, we show particular change in the observed LC shape and the secondary bump of the Hertzsprung progression. This bump was first highlighted by Hertzsprung (1926) in the LCs of Cepheids between 6 and 14 days. In Fig. 11, even if this feature is smoothed during the template generation, we can still identify it (highlighted by the red arrows for LCs for 6 and 13 days).

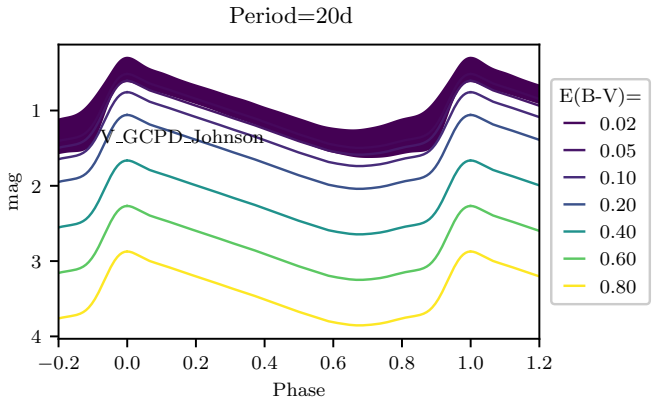


Fig. 10. Example of a template for a Cepheid of a period of 20 days. The different lines correspond to different values of $E(B-V)$. The shaded area is the confidence interval and is given only for no interstellar reddening for readability. For other filters, see Fig. D.1.

The Hertzsprung progression can also be characterised by the centre period of the progression, around which bump Cepheids are found. Marconi et al. (2024) showed, using theoretical models compared to observations in the *Gaia* database, that the centre of the Hertzsprung progression corresponds to a local minimum in the variation of the peak-to-peak amplitude with respect to the pulsation period, in the *Gaia G* band. The evolution of the amplitude of the LCs in our templates is shown in Fig. 12. The central period is located at 8.67 days, knowing that the mean metallicity of our sample is around 0.07 (individual metallicities are from the compilation by Breuval et al. 2021, and references therein). This is compatible with the results in Marconi et al. (2024), who reported the central period in this band is between 7.5 and 9.5 days, with longer period associated to lower metallicities, with the trend being confirmed by the theoretical models.

4. Validation of the template LCs

In this section, we validate the derived templates with respect to how well they reproduce the multi-band LCs and their shapes. We describe its performance both internally and externally.

4.1. Search for systematic errors

As our templates are suited for the determinations of mean magnitudes, we need to estimate the error in this parameter. This can be done by comparing the predicted mean magnitudes from templates to the one retrieved by a fit of an LC using standard methods, such as the Fourier series. In both cases, the mean magnitudes were computed by converting into magnitude the average flux in the different filters.

We performed this analysis in two steps, using similar methods. First, we propose an internal validation using the training sample (gold sample) and an external validation using an independent validation sample (bronze sample). They are represented in the top part of the left panel of Fig. 1

For the internal validation, the 75 stars of the gold sample were fitted with the templates and the P–K_i relations without introducing the prior knowledge on the coefficients K_i. Only two parameters were fitted in this process; the reference magnitude, $m_{\rm ref}$, and the colour excess, $E(B-V)$. We show, as an example, the LC fitting for the Cepheid η Aql in Fig. 13. Interestingly, for this star, the bump present in the V band is not reproduced by the template, which smooths the curve at these phases. This feature is supposed to be too small to have an impact on the mean observed magnitude. We also plot the LC in the V band from the templates described in Yoachim et al. (2009). For this template, the secondary bump is also not fitted and the amplitude seems to be underestimated.

In Table 1, we compare the mean magnitudes from the fitted template to the mean magnitudes in the SPIPS results and from a Fourier interpolation of the LC when the

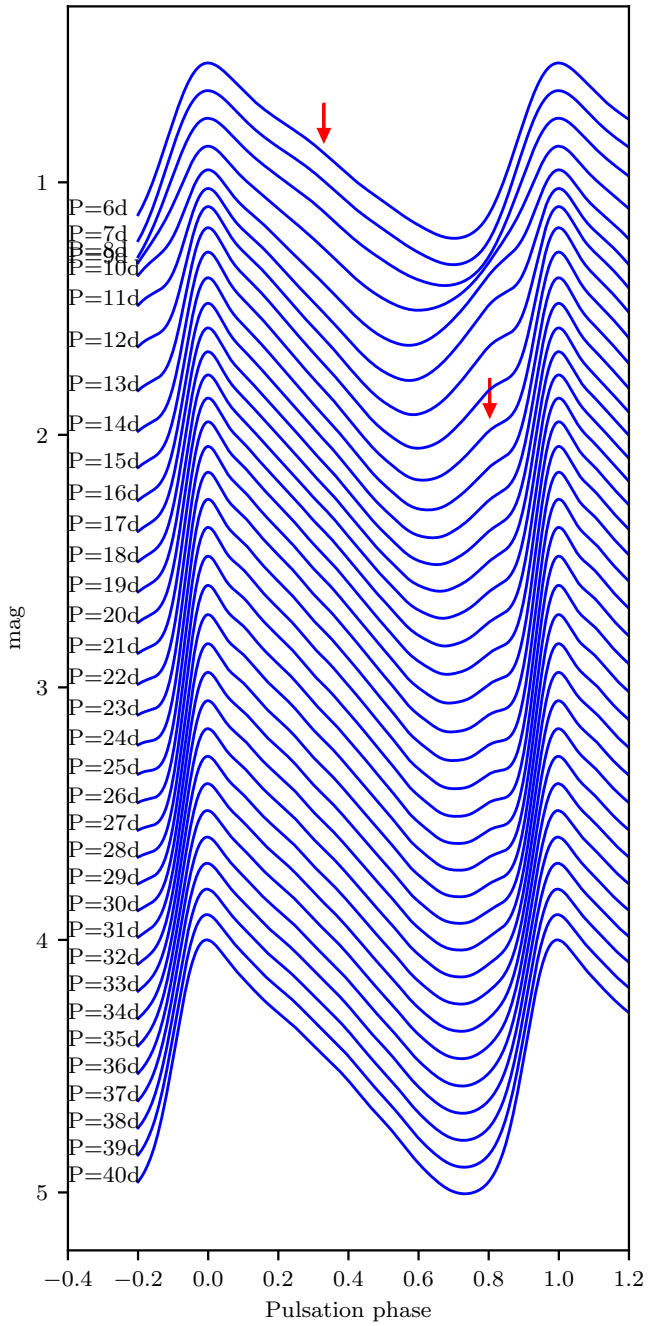


Fig. 11. Evolution of the shape of the LC with the period in the Johnson V band for periods from 6 to 40 days as predicted by our templates. The red arrow shows the secondary bump from the Hertzsprung progression.

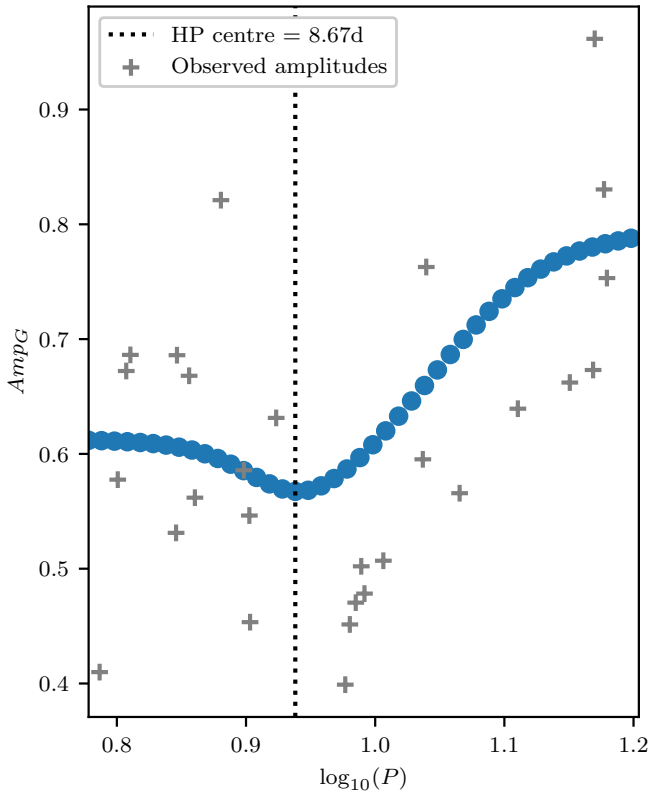


Fig. 12. Amplitude of the template LC in the *Gaia G* band and its evolution with the period from 6 to 16 days. The amplitude from the *Gaia* LC for the stars in the gold sample are also plotted for comparison with real data.

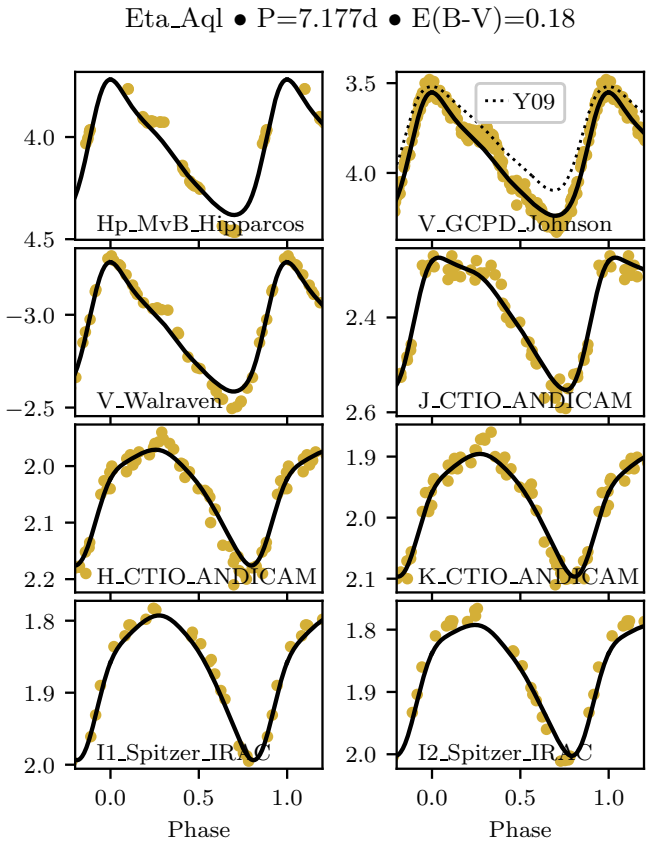


Fig. 13. Fitting procedure of η Aql in the gold sample, as a validation of the templates. Only two parameters are fitted to all the observations, $m_{\rm ref}$ and $E(B-V)$.

phase-coverage is sufficient. For the reference comparison, we used determinations from several works (Monson & Pierce 2011; Gaia Collaboration 2018; Breuval et al. 2021; Gaia Collaboration 2023). The comparison with the SPIPS results provides us proof of self-consistence of the process and is used to validate the approach of P–K_i relations. The small offsets and dispersion found here indicate that while our templates smooth the LCs by omitting high-frequency terms, they remain effective in obtaining accurate mean magnitudes across all filters.

For most photometric bands, we can see from the mean error that the mean magnitude is well retrieved, except for *Gaia G* and *BP* bands, especially for the DR3. Between the two latest data releases, the photometric passbands of *Gaia* have been recalibrated and the shape of the distribution in wavelengths have changed quite drastically. This explains why so much difference is obtained in the modelled magnitudes from atmospheric models. The problem for the newly calibrated passbands in the DR3 may lie in the difficulty for atmospheric models used in SPIPS to get a good photometry in the bluer bands, as stated by Casagrande & VandenBerg (2014). Although this problem is concerning in the context of using *Gaia* data more extensively, it does not impact the prediction in the other bands. We however note that the calibration of *Gaia* passbands from the DR2 seems to provide good predictions for the LC, even when compared to *Gaia* DR3 data. A quite large offset is also found for mean magnitudes in *JHK* filters in the 2MASS system (0.07 mag in *H*). This is surprising as the reference mean magnitudes that are quoted are computed from the same data sample as used in the template fitting, which were in the CTIO system, transformed into the 2MASS system, using transformation relations from Monson & Pierce (2011). A similar offset is found by comparing the LCs in *JHK* in the 2MASS system from our templates directly with the transformation using the same relations of the predicted LCs in *JHK* in the CTIO system. Given this comparison, and the fact that the photometry from the 2MASS survey seems to be well-fitted (even though one data point is clearly not sufficient to validate the fit), we conclude that this offset comes from the transformation relations between the CTIO and 2MASS systems. A comparison with the literature in HST filters will be discussed in details in Sect. 4.2.

The bronze sample is composed of 34 stars with available photometry both in the visible and the NIR, but not used for the template generation (neither in the gold nor in the silver sample). In particular, we selected stars that were already studied for the calibration of the Period-Wesenheit relations in HST filters (Riess et al. 2021). Photometric data for these stars come from various references, listed in Table E.1. We fit a template at the pulsation period of the star taken from the International Variable Star Index (VSX) database[1] or, alternatively, from Berdnikov (2008). We left the reference magnitude, the colour excess, and an offset in phase as free parameters. Due to the heterogeneity in observation epochs (spanning many periods), the phasing of the data can be highly sensitive to the given period. Before the template fitting, we found a variation in the period (as a fraction of the initial period), by minimizing the dispersion in the phase-folded LCs to obtain cleaner phase-folded LCs in all filters. This was done independently of any predicted LC, avoiding a degeneracy with

[1] Operated at AAVSO, Cambridge, Massachusetts, USA.

Table 1. Distribution (average and standard deviation) of errors between mean magnitudes from a template fitting process and reference values from the literature.

Filter	SPIPS		Reference		
	$\overline{\Delta\langle m_\lambda\rangle}$	$\sigma_{\Delta\langle m_\lambda\rangle}$	$\overline{\Delta\langle m_\lambda\rangle}$	$\sigma_{\Delta\langle m_\lambda\rangle}$	Literature
B MvB TYCHO	0.012	0.054	–	–	
B GCPD Johnson	−0.010	0.030	–	–	
Gbp GAIA GAIA3	–	–	−0.090	0.049	Gaia Collaboration (2023)
Gbp GAIA GAIA2	−0.004	0.016	−0.026	0.045	Gaia Collaboration (2018)
Hp MvB Hipparcos	−0.005	0.016	–	–	
V MvB TYCHO	0.001	0.018	–	–	
F555W HST WFC3 UVIS2	−0.008	0.022	−0.070	0.054	Riess et al. (2021)
V GCPD Johnson	−0.004	0.017	0.003	0.018	Breuval et al. (2021)
V Walraven	−0.005	0.019	–	–	
G GAIA GAIA3	–	–	−0.179	0.082	Gaia Collaboration (2023)
G GAIA GAIA2	−0.000	0.018	−0.015	0.048	Gaia Collaboration (2018)
Grp GAIA GAIA3	–	–	−0.015	0.031	Gaia Collaboration (2023)
Grp GAIA GAIA2	−0.002	0.013	0.007	0.046	Gaia Collaboration (2018)
F814W HST WFC3 UVIS2	−0.003	0.016	−0.014	0.043	Riess et al. (2021)
J 2MASS	−0.001	0.014	−0.024	0.014	Breuval et al. (2021)
J CTIO ANDICAM	−0.002	0.016	−0.010	0.016	Monson & Pierce (2011)
F160W HST WFC3 IR	−0.002	0.008	0.031	0.043	Riess et al. (2021)
H CTIO ANDICAM	−0.002	0.010	−0.007	0.010	Monson & Pierce (2011)
H 2MASS	−0.001	0.008	0.069	0.011	Breuval et al. (2021)
K CTIO ANDICAM	−0.002	0.010	−0.004	0.014	Monson & Pierce (2011)
Ks 2MASS	−0.002	0.010	0.003	0.015	Breuval et al. (2021)
I1 Spitzer IRAC	0.005	0.010	–	–	
I2 Spitzer IRAC	0.009	0.021	–	–	

Notes. The different filters are ordered by central wavelength. The errors are computed with respect to the SPIPS results and to a value from the literature. The statistical parameters are provided on the 75 stars of the gold sample. Not all the stars have literature values.

the template fit. The resulting period is used only for phasing and not to define the template as variations are supposed to be minimal. While this step has little impact on the retrieved mean magnitudes, it helps visualise the agreement between the template and observations. We do not claim to have a high level of precision in terms of the derived period, as we only aim to phase our data. In particular, period changes were not accounted for to maintain a low number of degrees of freedom. Finally, for extragalactic Cepheids, this step becomes less critical due to shorter time spans between epochs and noisier photometry, except when combining HST and JWST data. An example of fit of the LC for the Cepheid YZ Car using our templates is given in Fig. 14. We also show in the V and I bands (as done in Fig. 13), the LCs from the templates described in Yoachim et al. (2009). Both templates are especially similar in shape for this star.

In the fitting procedure, we ignored a few specific filters that are not reproduced well by the template, but also generally in the SPIPS analysis. This accounts for blue filters (Johnson U for instance), but also other filters that are usually not included in the SPIPS analysis (R, I_c Johnson). For more information regarding the filters discarded in the SPIPS analysis, we refer to the SPIPS papers (e.g. Mérand et al. 2015; Breitfelder et al. 2016; Trahin et al. 2021; Bras et al. 2024).

Table 2 gives the statistical difference between the mean magnitudes predicted by the templates and a mean magnitude from a reference paper (Monson & Pierce 2011; Breuval et al. 2021; Gaia Collaboration 2023). These results were used to define the best choice for the width of the box in the moving average procedure to determine the P–K_i relations.

4.2. Validation with the SH0ES sample of Galactic Cepheids

In this section, we focus on stars in the sample described in Riess et al. (2021). Among these stars, 28 are in our gold sample, 12 in our silver sample, and 23 in our bronze sample; for the remaining 10 stars, we were not able to find any NIR photometry (except for 2MASS single measurements) to have confidence in the fit and add them to the bronze sample. We included these stars in the 'bronze +' sample to allow for some comparison. In Fig. 1, this step corresponds to the green boxes in the bottom of the validation panel.

We fit the photometry of all stars with the described templates, with the same free parameters: a reference magnitude, the colour excess, and a phase offset, except for the gold sample for which the data are already phased taking the maximum in bolometric luminosity as a reference. The LCs in the HST bands are predicted from the data in the other bands and are not based on any actual measurement. We also express the mean magnitudes in the Wesenheit index, defined as W_H = F160W − 0.386(F555W − F814W) (Riess et al. 2021). In our case, we can produce an LC in the Wesenheit index and compute its mean magnitude or we can (as done in the literature) express the mean magnitudes in W_H from the mean magnitudes in the F160W, F555W, and F814W filters. As the comparison from the second option is intrinsically given through the comparison of mean magnitudes in the other filters, we here compare the Wesenheit magnitude from Riess et al. (2021) with the mean magnitude from our templates constructed in the Wesenheit index. This enables us to validate more deeply the relation between the shape of the different LCs.

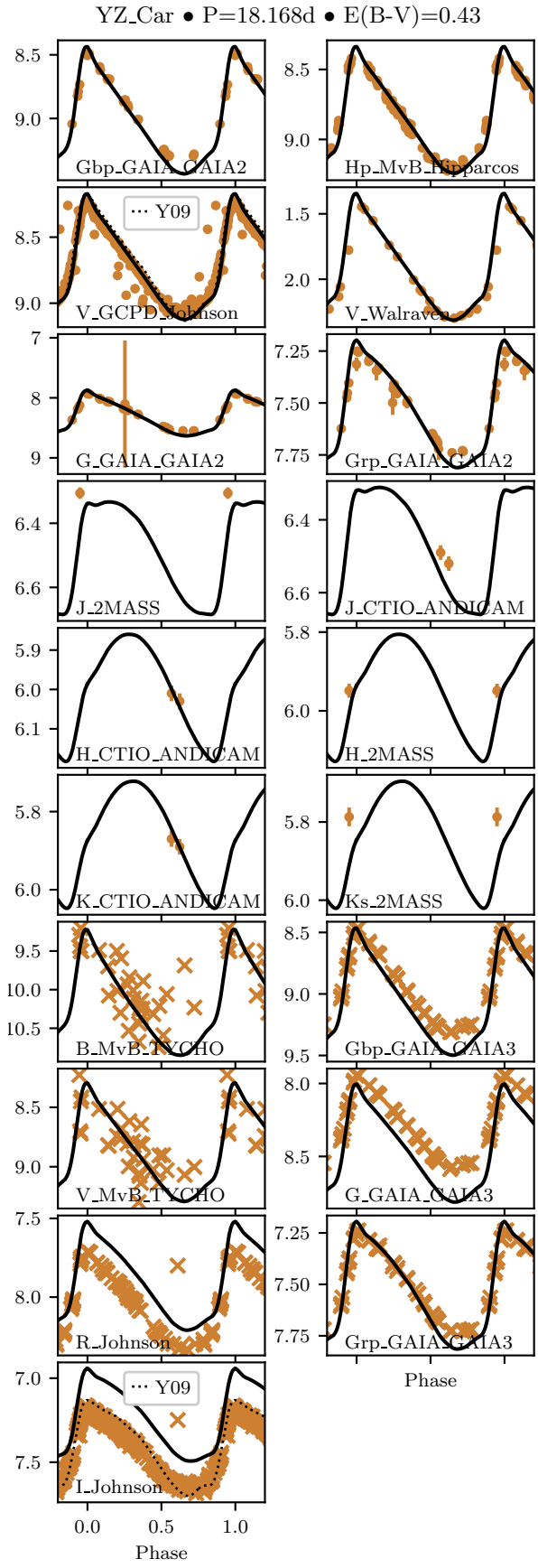


Fig. 14. Example of template fitting (grey lines) on YZ Car in the bronze sample. Observations are shown in bronze-coloured points and crosses. Crosses show data that are excluded from the fitting procedure.

Table 2. Same as Table 1 but for the stars in the bronze sample.

Filter	N	$\overline{\Delta\langle m_\lambda\rangle}$	$\sigma_{\Delta\langle m_\lambda\rangle}$	Reference
V Johnson	21	0.003	0.041	1
J 2MASS	8	−0.030	0.040	1
H 2MASS	8	0.051	0.037	1
Ks 2MASS	8	0.000	0.040	1
J CTIO	12	−0.008	0.054	2
H CTIO	12	−0.023	0.037	2
K CTIO	12	−0.011	0.038	2
G GAIA3	32	−0.224	0.161	3
Gbp GAIA3	32	−0.131	0.100	3
Grp GAIA3	32	−0.006	0.052	3

Notes. The errors are computed with respect to a value from the literature for the N stars for which a reference value could be found. References: (1) Breuval et al. (2021); (2) Monson & Pierce (2011); (3) Gaia Collaboration (2023).

We compare the mean magnitudes in the HST bands with the mean magnitudes in Table 1 of Riess et al. (2021) in Fig. 15 and Table 3. We removed the correction from the count-rate non-linearity of the IR detector in Table 1 in Wesenheit magnitudes in Riess et al. (2021), to compare the actual measurements to the results of our templates. The standard deviation in the difference in mean magnitudes retrieved gives a hint on the precision we can achieve from such a method, which still depends on the number of filters for which photometry is available and is of order of the dispersion observed by Riess et al. (2021) when comparing their photometry to literature photometry (see their Fig. 1). It is not comparable to the precision from direct measurements of the LCs in those filters, but should be representative of the errors using the template fitting approach. However, without individual measurements in the filters considered here, it is difficult to decide whether the offset and the dispersion we observe come from bad fits of the LCs or from systematics in our method. However, this offset does not have clear consequences for the distance ladder as we still need to compare Cepheids in anchor galaxies and in distant galaxies with the same tool and this offset would presumably be the same in both cases.

Our approach also enables us to determine the colour excess given the magnitudes in the different filters. We provide in Fig. 16 a comparison with previous determinations in the literature, from Breuval et al. (2021), multiplied by our 1.16 factor (see Eq. 2).

4.3. Validation of the LC shape

Besides the mean magnitudes, this section aims to investigate how well the LC shape is reproduced. We made use of our previous template fitting of our bronze sample to compare zero-mean LCs in each filter, as synthetised in the middle part of the validation panel in Fig. 1.

The comparison of the shape can be made by comparing the produced template with a parametrisation from a Fourier decomposition, as commonly done in the literature. A possible reference for such a decomposition could be the *Gaia* parameters, but as the *Gaia* bands are quite badly reproduced with by templates, it is not well suited for that purpose. Bhardwaj et al. (2015) provided the first Fourier parameters for a large number of stars in the V band and for a smaller number of stars in JHK bands; thus, it is fully adapted for the purposes of our investigation as well.

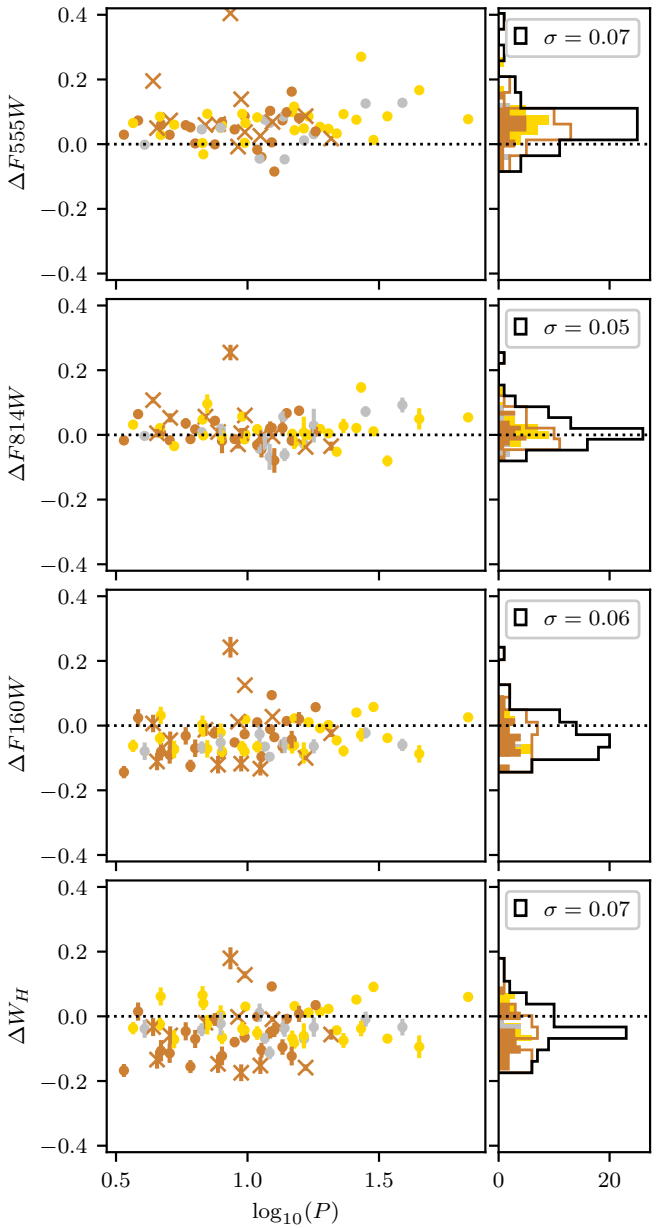


Fig. 15. Difference in mean magnitudes retrieved from the template with the mean magnitude in Riess et al. (2021) ($m_{\text{This work}} - m_{\text{Riess2021}}$). The colours correspond to the gold, silver, and bronze samples accordingly. The right panel shows the distribution. The black unfilled step is for the whole sample. Bronze crosses are for stars without NIR photometry. They are added to the bronze sample to give the unfilled bronze step distribution in the right panel.

Both computed curves were phased to minimise the total difference between them. We differentiated only among the differences in the shape of the curve, thus, we did not consider a possible offset in magnitude and we simply retrieved the intensity-averaged magnitude for the Fourier interpolation and the template independently. The observations were plotted to visually check whether the variations in shape come from real observations or from an oscillation in the Fourier interpolation. An example is given in Fig. 17. The template curve shown here is only dependent on the period of the star (due to the normalisation in terms of magnitude) and not on any observation or fitting process.

We plot in Fig. 18 the normalised difference between the template and the Fourier interpolation for the V-band LCs of all stars in our bronze sample, except AP Pup (as it has no Fourier decomposition in Bhardwaj et al. 2015). The curves are shown colour-coded with the pulsation period of the star, but there do not seem to be any clear tendencies that ought to be highlighted. The standard deviation of the distribution ($\leq$1) could suggest an overestimation of the confidence interval for the template; however, this may also account for the difference in the mean magnitude, which is not considered here.

5. Fitting the templates to extragalactic Cepheids

As a proof of concept, here we present an application of our templates to extragalactic Cepheids, observed with the HST in the context of the SH0ES project (see the bottom of the left panel in Fig. 1, shown as purple boxes). We made use of the epoch photometry in four filters (F350LP, F555W, F814W, and F160W) derived by Javanmardi et al. (2021) from HST observations of the SNIa host galaxy NGC5584. However, it is interesting to note that this galaxy has been quite intensively observed and the number of data points for the Cepheids in this galaxy is not fully representative of the observations of SNIa host galaxies in general. Still, testing the template fitting on this galaxy is valuable in terms of demonstrating the potential of our method.

Our templates have allowed us to simultaneously use information from all filters to derive consistent parameters. However, the determination of the period is still a delicate matter. Using the templates in V and I bands of Yoachim et al. (2009), Hoffmann et al. (2016) tested several periods that were equally spaced in a logarithmic scale, involving a visual inspection. As this process can become computationally expensive, we suggest narrowing the probable period range using standard methods such as the Lomb-Scargle periodogram or the string-length method (see Giertych et al. 2024). For this example, we used a string-length method on both the F350LP and F555W filters, summing the string-length functions to determine a plausible period. We then explored ten starting periods, logarithmically spaced around this plausible period, with a range of ~1 day around 20 days for the period and ~20 days around 100 days for the period. For each trial period, we fit all the photometry with our templates and four free parameters:

- the period (varying within the limit of the other trial periods);
- a phase offset to align the reference in time;
- the reference magnitude that offsets all LCs homothetically;
- the colour excess $E(B-V)$.

The fit was performed in two steps. First, we found the phasing parameters (period and phase offset) from a standard χ^2 minimisation in the four filters simultaneously and selected the trial period with the lowest χ^2. Because F160W has fewer observations, it has little impact on the phasing, which is mostly determined by LCs in the visible. However, these observations are critical for constraining the colour excess; hence, in the second step, we refined the reference magnitude and $E(B-V)$, weighting the data by the number of observations per filter, using fixed phasing parameters. Uncertainties were derived by fitting random realisations of observations within their error bars. A result of such a fit (Fig. 19 in this work) is given for one of the representative Cepheids given in Hoffmann et al. (2016, specifically, Fig. 4) and Javanmardi et al. (2021, specifically, Fig. 7).

We compare our parameters for the presented Cepheid with previous estimations in Table 4. From a visual inspection of the curves, we would like to lower the magnitude in the F160W filter and increase the magnitude in the F814W filter. This is not

Table 3. Distribution of mean magnitudes errors in HST filters for the different samples (see also Fig. 15).

	F555W		F814W		F160W		W_H	
	$\overline{\Delta\langle m_\lambda\rangle}$	$\sigma_{\Delta\langle m_\lambda\rangle}$	$\overline{\Delta\langle m_\lambda\rangle}$	$\sigma_{\Delta\langle m_\lambda\rangle}$	$\overline{\Delta\langle m_\lambda\rangle}$	$\sigma_{\Delta\langle m_\lambda\rangle}$	$\overline{\Delta\langle m_\lambda\rangle}$	$\sigma_{\Delta\langle m_\lambda\rangle}$
Gold	0.070	0.054	0.014	0.043	−0.031	0.043	−0.014	0.052
Silver	0.044	0.055	0.004	0.049	−0.062	0.021	−0.041	0.030
Bronze	0.044	0.053	0.006	0.037	−0.036	0.058	−0.058	0.065
Bronze+	0.063	0.080	0.017	0.057	−0.030	0.079	−0.054	0.083
Total	0.056	0.056	0.010	0.043	−0.039	0.047	−0.034	0.057
Total+	0.063	0.068	0.014	0.051	−0.035	0.062	−0.037	0.068

Notes. The '+' indicates selections where the stars with no NIR photometry were included (excluded otherwise).

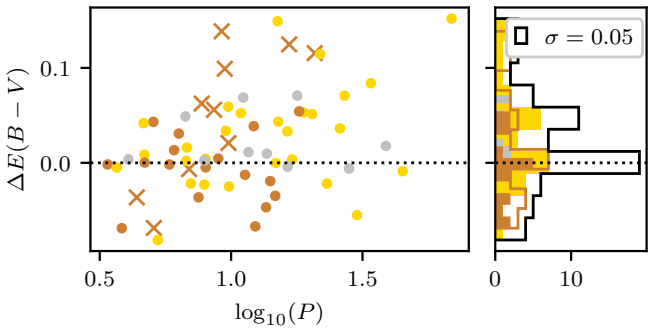


Fig. 16. Difference in $E(B-V)$ retrieved from the template fitting with determinations from Breuval et al. (2021). The legend is the same as for Fig. 15.

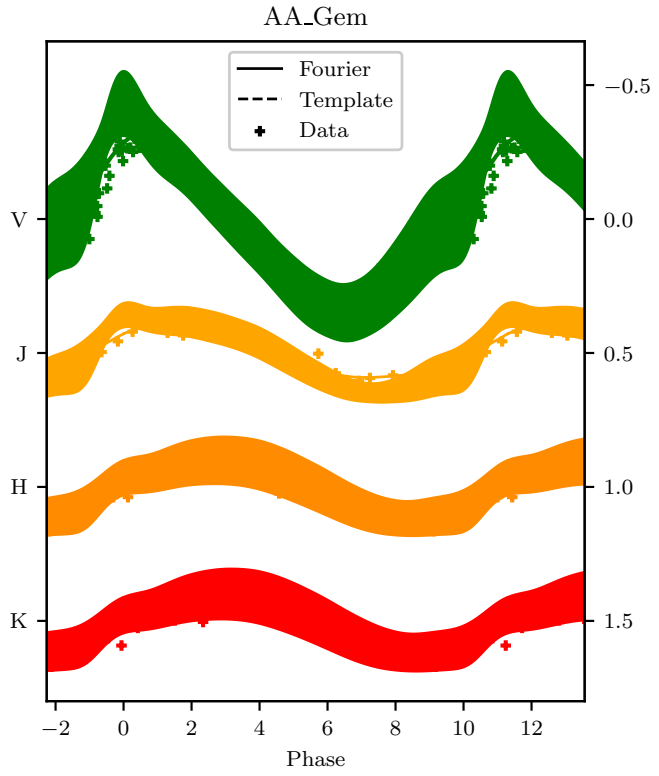


Fig. 17. Shape validation of AA Gem (bronze sample). We compare the template, associated to the period of the star, and the Fourier interpolation of observations from Bhardwaj et al. (2015).

possible in our multi-band approach as differences of mean magnitudes are fully determined by the period through the template and the colour excess. A better constraint on IR filters would be very valuable and we are making progress in adding JWST data to improve the template fitting process. Adding in these data will certainly improve our results even in the HST bands, as we will determine the colour excess more accurately. This means that we have access to dereddened magnitudes to obtain the period-luminosity relation in the visible range, without the scatter introduced by reddening effects. From the visualisation of the fit, it seems that we are underestimating the magnitude in F814W, probably because of an error in the determination of the colour excess. In the F160W filter, we obtained a mean magnitude that is lower than both previous determinations. However, increasing this magnitude does not seem to fit well with the data showed in Fig. 19.

6. Discussion

In this section, we point out some improvements to the present work or future paths that might be followed to extract the complete potential of these templates.

First of all, we emphasise that our generated templates are mainly aimed to the estimation of mean magnitudes in different filters. They are not properly suited for an interpolation of a well-observed Cepheid with good phase coverage, unless we relax the constraint on the coefficients associated to the PCA. In particular, they will not reproduce the amplitude of any Cepheid exactly or specific features like the bump of the Hertzsprung progression, because the goal in this work has been to link the general shape of the LC to the pulsation period only.

It is important to treat the reddening issue meticulously, as it has an important role in the mean magnitudes in the different filters. For the generation of our templates, we used appropriate corrections from the reddening due to interstellar dust, taking into account the fact that Cepheids have temperatures around 5500K, which has a significant impact on the photometry in broad filters such as F350LP. The present work enables us to study the suitable R_V parameter for reddening laws in various other galaxies and the reddening law itself (see also Mörtsell et al. 2022, for a discussion on the characterisation of reddening laws between anchor and SNIa host galaxies), as we have been able to study dereddened PL relations in the different filters. These questions are not raised by the standard approach by the SH0ES team as they used the Wesenheit index, rather than via the R_W parameter, in defining the Wesenheit magnitudes.

On the other hand, metallicity has a limited impact on the distance scale as it is quite similar in SNIa host galaxies and

Table 4. Properties of the Cepheid 473829 in NGC5584.

	This work	Hoffmann+2016	Javanmardi+2021	Riess+2016
Period	$56.394_{\pm 0.432}$	55.142	56.557	55.140
F350LP	$25.783_{\pm 0.023}$	–	$25.629_{\pm 0.012}$	–
F555W	$25.983_{\pm 0.034}$	$26.048_{\pm 0.020}$	$25.906_{\pm 0.015}$	–
F814W	$24.668_{\pm 0.048}$	$24.950_{\pm 0.017}$	$24.716_{\pm 0.020}$	–
F160W	$23.444_{\pm 0.098}$	–	$23.527_{\pm 0.045}$	$23.530_{\pm 0.250}$

Notes. Comparison is given between this work and results from Hoffmann et al. (2016), Riess et al. (2016), Javanmardi et al. (2021).

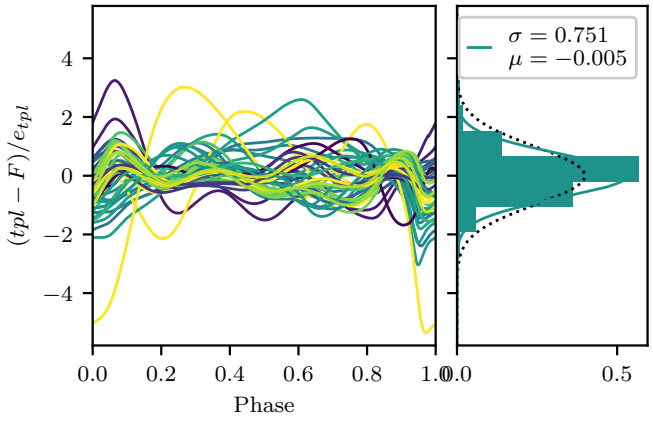


Fig. 18. Normalised residuals between the templates and a Fourier interpolation in the V band of stars in the bronze sample and its distribution (right panel). The Gaussian distribution is given for comparison in dashed black line.

in the MW (Riess et al. 2022), which explains why we considered the galactic metallicity in this work. However, it seems to have an impact on the shape of Cepheids LCs (Hocdé et al. 2023), and was proven to be particularly striking for short-period Cepheids (below 6 days). For longer periods, it is possible that the variation in the shape would occur at higher frequencies. This raises the question of whether this effect would impact the template fitting to metal-poor Cepheids in the Magellanic Clouds, for instance. It would be interesting to quantify this effect based on our previous analysis. Using our principal components, we could study the LCs of low-metallicity Cepheids and analyse the coefficient distribution. This could reveal analogous Cepheids (in terms of shape and luminosities) when going to lower metallicity, but a longer period.

Moreover, in this approach, we include in our templates the potential contribution from CSE, by making the hypothesis that their impact is homogeneous in galactic and extragalactic Cepheids or its presence will add dispersion to the $P - K_i$ relations. It is then important to understand the CSE properties across the instability strip to reduce the uncertainty of our templates and (if it does indeed exist) the link between the CSE and the metallicity. So far, no particular correlation with the location on the instability strip could be identified by Gallenne et al. (2021). They also highlighted that the impact of the CSE mainly results in a variation in the ratio of magnitude between the different filters, which will likely be absorbed in the fitted parameters m_{ref} and the colour excess, $E(B-V)$.

Finally, we introduced the algebraic parameter $m_{\rm ref}$ to move all LCs in all filters homothetically, which depends on the filters used at the beginning of our analysis and does not have

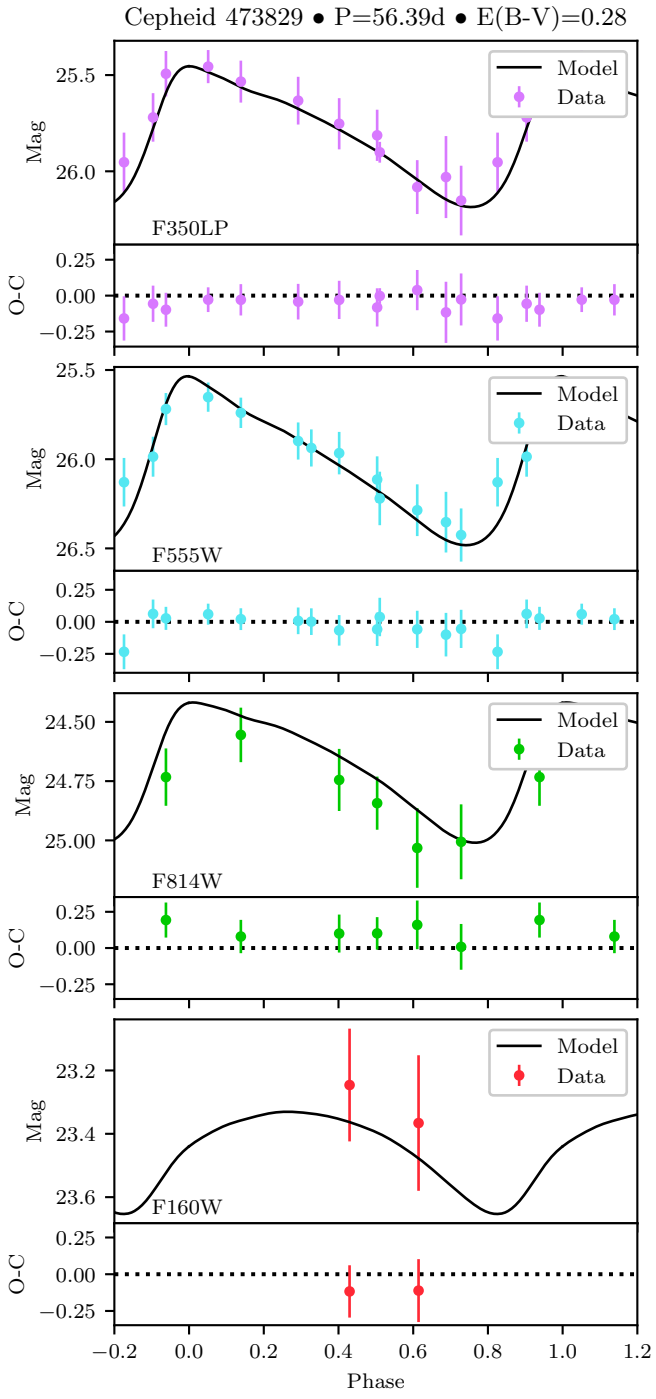


Fig. 19. Example of the Cepheid 473829 in NGC 5584. For each filter the observations, the fitted template, and the residuals are plotted. Data come from the work of Javanmardi et al. (2021).

a proper physical meaning, but we note that it is similar to a distance modulus. The result is that for more distant Cepheids, the template fitting procedure will give higher values for $m_{\rm ref}$, as all magnitudes in all filters will be higher. We could think of another parametrisation of the PCA where this parameter would be directly linked to the distance of the Cepheid. The only free parameters will then be reduced to the distance, the reddening, the period, and a time reference. By doing so, the calibration of the PL relation would be incorporated into the $P-K_i$ relations directly and will no longer be needed, as the fit with such derived templates would also estimate the distance. This could help in setting better constraints on the uncertainties when fitting extragalactic Cepheids in future studies.

7. Conclusion

We present a new approach to constructing LC templates for Cepheids, which intrinsically includes the information of the temperature of the star and a proper correction for the reddening from interstellar dust, with the limitation that they are not able to reproduce the variety of Cepheid LCs in the same period bin. This approach allows us to apply these templates to a broad variety of science cases, thanks to their versatility in terms of the photometric filters used. They incorporated a physical understanding of the variation of the photosphere, through the use of SPIPS parametrisation for the training sample. In our study, we introduced different validation samples in order to test the possibilities of the use of such templates. We recommend the user to adapt the template fitting procedure to their scientific question. In particular, if the aim is to fit the shape of the LC with a good phase coverage for the data, but a large number of filters, it would be necessary to allow some variation in the shape of LCs by relaxing the constraints on the values of the coefficients associated to the PCA, to retrieve fine features. We have also demonstrated their application in the case of extragalactic Cepheids for the distance ladder.

These templates will be used for an independent analysis of HST images of the sample of SNIa host galaxies studied in Riess et al. (2022) to validate the second rung of the distance ladder performed by this team, together with new photometry of the stars in these galaxies. The result will be published in a forthcoming paper. The templates will also be pertinent to interpret JWST data. We are already able to produce LCs of Cepheids in JWST filters and including them for extragalactic Cepheids will be feasible without the need for any new analysis.

Data availability

The necessary data to reproduce the analysis is accessible on Zenodo (DOI:10.5281/zenodo.15102196). The record contains the fits output files of the SPIPS analysis of the 75 stars of our gold sample and a fits file including all the information to generate LC templates (extinction coefficients for the selection of filters, principal components and $P-K_i$ relations).

Acknowledgements. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (http://www.cosmos.esa.int/gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, http://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. The research leading to these results has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (projects CepBin, grant agreement 695099, and UniverScale, grant agreement 951549). This research has been supported by the Polish-French Marie Skłodowska-Curie and Pierre Curie Science Prize awarded by the Foundation for Polish Science. The authors acknowledge the support of the French Agence Nationale de la Recherche (ANR), under grant ANR-23-CE31-0009-01 (Unlockpfactor). A. G. acknowledges the support of the Agencia Nacional de Investigación Científica y Desarrollo (ANID) through the FONDECYT Regular grant 1241073. We also acknowledge support from the Polish Ministry of Science and Higher Education grant DIR-WSIB.92.2.2024. This research has made use of Astropy (Available at http://www.astropy.org/), a community-developed core Python package for Astronomy (Astropy Collaboration 2013, 2018), the Numpy library (Harris et al. 2020), the Astroquery library (Ginsburg et al. 2019), the scikit-learn package (Pedregosa et al. 2011) and the Matplotlib graphics environment (Hunter 2007). We acknowledge with thanks the variable star observations from the AAVSO International Database contributed by observers worldwide and used in this research. We used the SIMBAD and VizieR databases and catalogue access tool at the CDS, Strasbourg (France), and NASA's Astrophysics Data System Bibliographic Services. The original description of the VizieR service was published in Ochsenbein et al. (2000). This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

References

Anderson, R. I. 2016, MNRAS, 463, 1707
Astropy Collaboration (Robitaille, T. P., et al.) 2013, A&A, 558, A33
Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, AJ, 156, 123
Balázs, L. G., & Kovács, G. B. 2025, New Astron., 116, 102317P
Barnes, T. G., Fernley, J. A., Frueh, M. L., et al. 1997, PASP, 109, 645
Berdnikov, L. N. 2008, VizieR Online Data Catalog: II/285
Bhardwaj, A., Kanbur, S. M., Singh, H. P., Macri, L. M., & Ngeow, C.-C. 2015, MNRAS, 447, 3342
Bras, G., Kervella, P., Trahin, B., et al. 2024, A&A, 684, A126
Breitfelder, J., Mérand, A., Kervella, P., et al. 2016, A&A, 587, A117
Breuval, L., Kervella, P., Wielgórski, P., et al. 2021, ApJ, 913, 38
Casagrande, L., & VandenBerg, D. A. 2014, MNRAS, 444, 392
Castelli, F., & Kurucz, R. L. 2003, in Modelling of Stellar Atmospheres, eds. N. Piskunov, W. W. Weiss, & D. F. Gray, 210, A20
Chown, A. H., Scowcroft, V., & Wuyts, S. 2021, MNRAS, 500, 817
Coulson, I. M., & Caldwell, J. A. R. 1985, South African Astronomical Observatory Circular, 9, 5
Coulson, I. M., Caldwell, J. A. R., & Gieren, W. P. 1985, ApJS, 57, 595
Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, 2MASS All Sky Catalog of Point Sources
Deeming, T. J. 1964, MNRAS, 127, 493
Engle, S. G., Guinan, E. F., Harper, G. M., Neilson, H. R., & Remage Evans, N. 2014, ApJ, 794, 80
ESA 1997, in The HIPPARCOS and TYCHO Catalogues. Astrometric and Photometric Star Catalogues Derived from the ESA HIPPARCOS Space Astrometry Mission, ESA Special Publication, 1200
Feast, M. W., Laney, C. D., Kinman, T. D., van Leeuwen, F., & Whitelock, P. A. 2008, MNRAS, 386, 2115
Fitzpatrick, E. L. 1999, PASP, 111, 63
Freedman, W. L. 1988, ApJ, 326, 691
Freedman, W. L., & Madore, B. F. 2010, ApJ, 719, 335
Gaia Collaboration (Brown, A. G. A., et al.) 2018, A&A, 616, A1
Gaia Collaboration (Brown, A. G. A., et al.) 2021, A&A, 649, A1
Gaia Collaboration (Vallenari, A., et al.) 2023, A&A, 674, A1
Gallenne, A., Kervella, P., Mérand, A., et al. 2017, A&A, 608, A18
Gallenne, A., Mérand, A., Kervella, P., et al. 2021, A&A, 651, A113
Gieren, W. 1981, ApJS, 47, 315
Gieren, W. P. 1985, ApJ, 295, 507
Giertych, N., Shaban, A., Haravu, P., & Williams, P., Jr. 2024, Reports on Progress in Physics, 87, 078401
Ginsburg, A., Sipőcz, B. M., Brasseur, C. E., et al. 2019, AJ, 157, 98
Harris, H. C. 1981, AJ, 86, 707
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357
Hendry, M. A., Tanvir, N. R., & Kanbur, S. M. 1999, in Harmonizing Cosmic Distance Scales in a Post-HIPPARCOS Era, eds. D. Egret, & A. Heck, ASP Conf. Ser., 167, 192
Hertzsprung, E. 1926, Bull. Astron. Inst. Neth., 3, 115
Hocdé, V., Smolec, R., Moskalik, P., Ziółkowska, O., & Singh Rathour, R. 2023, A&A, 671, A157
Hoffmann, S. L., Macri, L. M., Riess, A. G., et al. 2016, ApJ, 830, 10
Hunter, J. D. 2007, Computing in Science and Engineering, 9, 90
Inno, L., Matsunaga, N., Romaniello, M., et al. 2015, A&A, 576, A30

Javanmardi, B., Mérand, A., Kervella, P., et al. 2021, ApJ, 911, 12
Kanbur, S. M., Iono, D., Tanvir, N. R., & Hendry, M. A. 2002, MNRAS, 329, 126
Kervella, P., Trahin, B., Bond, H. E., et al. 2017, A&A, 600, A127
Kiss, L. L. 1998, MNRAS, 297, 825
Laney, C. D., & Stobie, R. S. 1992, A&AS, 93, 93
Leavitt, H. S., & Pickering, E. C. 1912, Harvard College Observatory Circular, 173, 1
Lindegren, L., Bastian, U., Biermann, M., et al. 2021, A&A, 649, A4
Lloyd Evans, T. 1980, South African Astronomical Observatory Circular, 1, 163
Madore, B. F. 1975, ApJS, 29, 219
Madore, B. F., Freedman, W. L., & Owens, K. 2024, AJ, 167, 201
Marconi, M., De Somma, G., Molinaro, R., et al. 2024, MNRAS, 529, 4210
Mérand, A., Kervella, P., Breitfelder, J., et al. 2015, A&A, 584, A80
Moffett, T. J., & Barnes, T. G., I 1984, ApJS, 55, 389
Monson, A. J., & Pierce, M. J. 2011, ApJS, 193, 12
Monson, A. J., Freedman, W. L., Madore, B. F., et al. 2012, ApJ, 759, 146
Mörtsell, E., Goobar, A., Johansson, J., & Dhawan, S. 2022, ApJ, 933, 212
Ngeow, C.-C., Kanbur, S. M., Nikolaev, S., Tanvir, N. R., & Hendry, M. A. 2003, ApJ, 586, 959
Ochsenbein, F., Bauer, P., & Marcout, J. 2000, A&AS, 143, 23
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825
Pejcha, O., & Kochanek, C. S. 2012, ApJ, 748, 107
Pel, J. W. 1976, A&AS, 24, 413
Phillips, M. M. 1993, ApJ, 413, L105
Pietrzyński, G., Graczyk, D., Gallenne, A., et al. 2019, Nature, 567, 200
Reid, M. J., Pesce, D. W., & Riess, A. G. 2019, ApJ, 886, L27
Riess, A. G., Macri, L. M., Hoffmann, S. L., et al. 2016, ApJ, 826, 56
Riess, A. G., Casertano, S., Yuan, W., et al. 2021, ApJ, 908, L6
Riess, A. G., Yuan, W., Macri, L. M., et al. 2022, ApJ, 934, L7
Schechter, P. L., Avruch, I. M., Caldwell, J. A. R., & Keane, M. J. 1992, AJ, 104, 1930
Stetson, P. B. 1996, PASP, 108, 851
Szabados, L. 1977, Commmunications of the Konkoly Observatory Hungary, 70, 1
Szabados, L. 1980, Commmunications of the Konkoly Observatory Hungary, 76, 1
Szabados, L. 1981, Commmunications of the Konkoly Observatory Hungary, 77, 1
Tanvir, N. R., Hendry, M. A., Watkins, A., et al. 2005, MNRAS, 363, 749
Trahin, B., Breuval, L., Kervella, P., et al. 2021, A&A, 656, A102
Walraven, J. H., Tinbergen, J., & Walraven, T. 1964, Bull. Astron. Inst. Neth., 17, 520
Welch, D. L., Wieland, F., McAlary, C. W., et al. 1984, ApJS, 54, 547
Yoachim, P., McCommas, L. P., Dalcanton, J. J., & Williams, B. F. 2009, AJ, 137, 4697

Appendix A: Python programme for using complementary data to generate templates of LCs

We present here a python script easy to run to produce template LCs for a Cepheid given a set of parameters that may be changed. The fits file needed is joined to this paper on Zenodo.

```
import numpy as np
from astropy.io import fits
from scipy.interpolate import interp1d
import astropy.units as u

# Read supplementary data
hdu = fits.open('PCA_data.fits')

# Parameters
photometric_band = 'V_GCPD_Johnson'
R_v = 3.1
Period_in_days = 20
mag_ref = 10
EBV = 0.5

# Compute template LC and its error
mask = np.where(hdu[1].data['Filter']==photometric_band)
LC = hdu[1].data['C0'][mask]
error_LC = np.zeros(LC.shape)
for i in range (1,5):
   Ki = interp1d(hdu[2].data['log_P'], hdu[2].data['K'+str(i)])(np.log10(Period_in_days))
   e_Ki = interp1d(hdu[2].data['log_P'], hdu[2].data['e_K'+str(i)])(np.log10(Period_in_days))
   LC += Ki*hdu[1].data['C'+str(i)][mask]
   error_LC += (e_Ki*hdu[1].data['C'+str(i)][mask])**2
error_LC = np.sqrt(error_LC)
LC += mag_ref + EBV*hdu[0].header['A_LAMBDA_5500K␣'+photometric_band]
```

Appendix B: Example of a SPIS fit for Cepheids in the gold sample

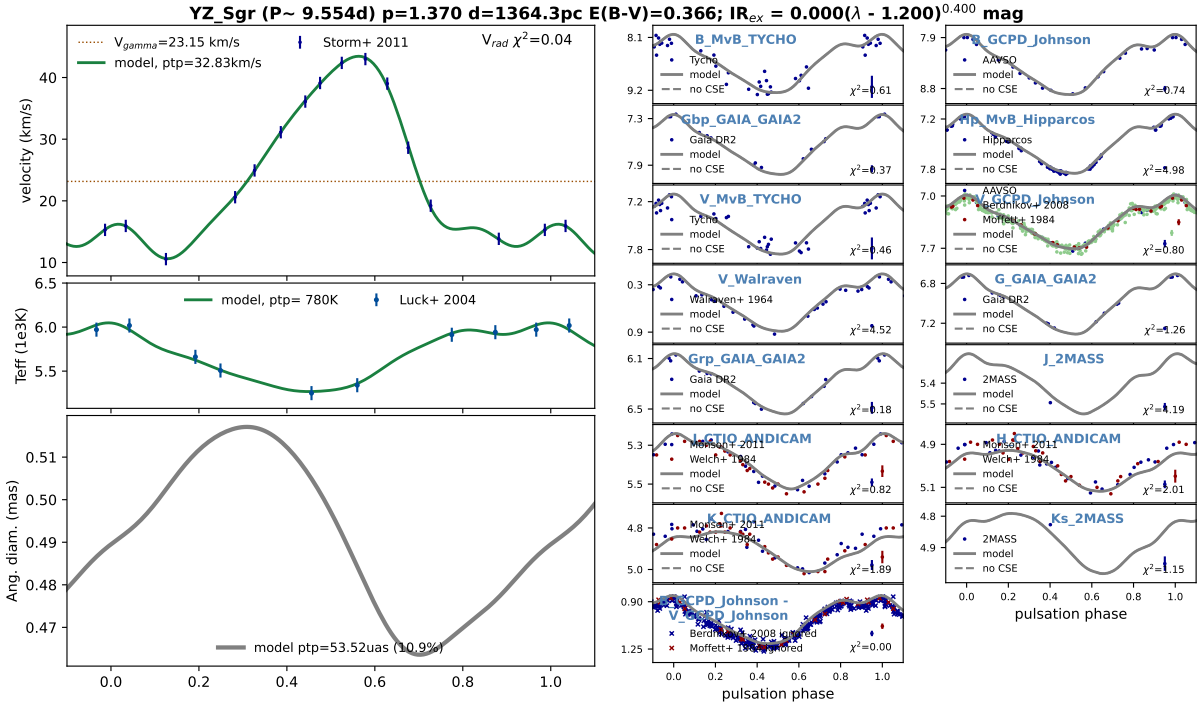


Fig. B.1. Pulsation parametrisation from the SPIPS fitting tool for YZ Sgr in the gold sample, but not included in Trahin et al. (2021). The χ^2 for this fit is 1.92.

Appendix C: PCA components for two filters

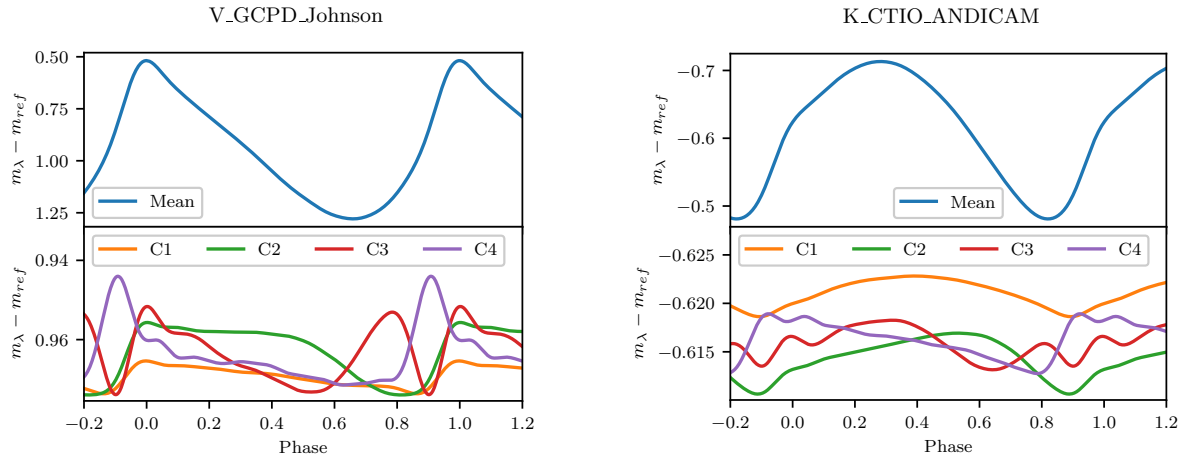


Fig. C.1. Mean LC (top) and the four first components (bottom) associated to the Johnson V (a) and CTIO *K* (b) filters.

Appendix D: Example of templates for some standard filters

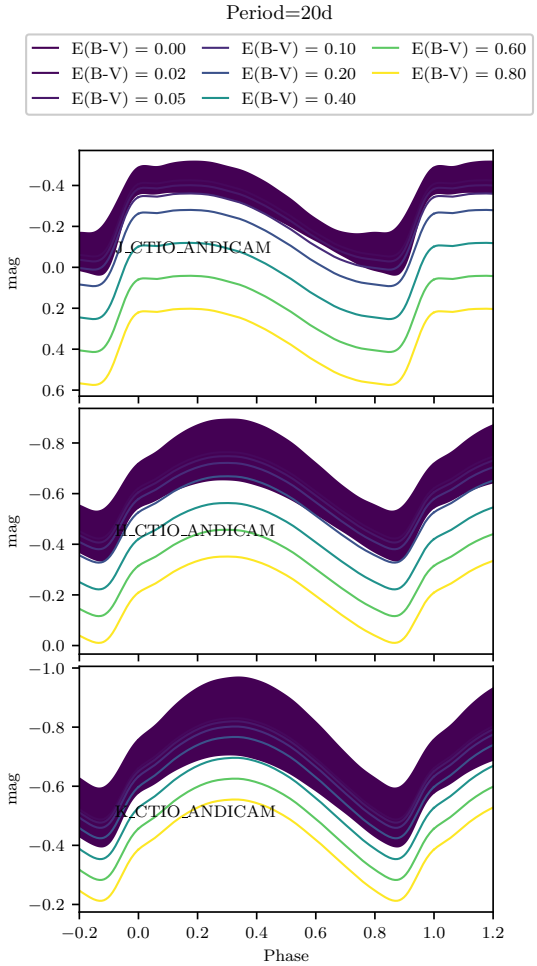


Fig. D.1. Example of a template for a Cepheid of a period of 20 days. The legend is similar as in Fig. 10

Appendix E: Properties of the stars used in this paper

Table E.1. Galactic Cepheids studied for the templates generation and validation.

Star	Period (days)	E(B-V) (mag)	Sample	References
AA GEM	11.3024, 11.3037[a]	0.387[b]	Bronze, SH0ES	10, 11, 18, 21, 22, 24, 28
AC MON	8.0143	0.588	Silver	1, 4, 18, 21, 22, 24, 28
AD PUP	13.5963, 13.5975[a]	0.374[b]	Bronze, SH0ES	3, 4, 17, 18, 21, 22, 28
AG CRU	3.8373, 3.8372[a]	0.212[b]	Bronze, SH0ES	4, 9, 12, 21, 22, 28
AP PUP	5.0843, 5.0844[a]	0.221[b]	Bronze +, SH0ES	1, 4, 18, 21, 22
AP SGR	5.0579, 5.0580[a]	0.256[b]	Bronze, SH0ES	1, 4, 9, 11, 12, 18, 21, 22, 28
AQ CAR	9.7690, 9.7691[a]	0.216[b]	Bronze +, SH0ES	2, 14, 18, 21, 22, 28
AQ PUP	30.1667	0.528, 0.561[b]	Gold, SH0ES	1, 3, 6, 11, 16, 17, 18, 21, 22, 28
AS PER	4.9725	0.793	Silver	5, 11, 18, 21, 22, 24, 28
AW PER	6.4635	0.553	Gold	8, 11, 18, 20, 21, 22, 24, 28
AY SGR	6.5696	0.974	Silver	1, 4, 18, 21, 22, 24, 28
BB SGR	6.6370	0.323	Gold	1, 4, 11, 12, 16, 18, 21, 22, 28
BE MON	2.7055	0.637	Silver	1, 5, 18, 21, 22, 24, 28
BETA DOR	9.8426	0.063	Gold	4, 16, 18, 21, 22, 25
BF OPH	4.0677	0.303, 0.306[b]	Silver, SH0ES	1, 4, 7, 9, 11, 12, 16, 18, 21, 22, 28
BG LAC	5.3322	0.324	Gold	11, 18, 19, 21, 22, 24, 28
BG VEL	6.9237, 6.9239[a]	0.497[b]	Bronze +, SH0ES	1, 4, 18, 21, 22, 28
BK AUR	8.0024, 8.0034[a]	0.451[b]	Bronze, SH0ES	8, 18, 21, 22, 24, 28
BM PER	22.9568	1.066	Silver	1, 21, 22, 24
BN PUP	13.6727	0.489, 0.499[b]	Silver, SH0ES	3, 4, 13, 16, 17, 18, 21, 22, 28
BR VUL	5.1972, 5.1978[a]	0.803[b]	Bronze	18, 21, 24, 28
CD CAS	7.8009	0.864	Silver	18, 21, 22, 24, 28
CD CYG	17.0745	0.598, 0.597[b]	Gold, SH0ES	6, 10, 11, 12, 18, 21, 22, 24, 25, 28
CF CAS	4.8754	0.623	Gold	6, 11, 18, 21, 22, 24, 25, 28
CK SCT	7.4152, 7.4152[a]	0.905[b]	Bronze	1, 4, 18, 21, 24, 28
CP CEP	17.8646	0.790, 0.860[b]	Silver, SH0ES	18, 21, 22, 24, 28
CR CAR	9.7596, 9.7596[a]	0.733[b]	Bronze, SH0ES	4, 17, 21, 22, 28
CR CEP	6.2329	0.816	Silver	1, 8, 11, 18, 21, 22, 24, 28
CR SER	5.3013	1.129	Silver	18, 21, 22, 24, 28
CS ORI	3.8894, 3.8880[a]	0.430[b]	Bronze	18, 21, 22, 24, 28
CS VEL	5.9046	0.830	Silver	12, 16, 17, 21, 22, 28
CV MON	5.3786	0.826	Gold	4, 8, 11, 16, 18, 21, 22, 24, 25, 28
CY AUR	13.8476	0.912, 0.866[b]	Silver, SH0ES	21, 22, 24, 28
CY CAS	14.3769, 14.3793[a]	1.113[b]	Bronze	18, 21, 22, 24, 28
DD CAS	9.8113	0.514, 0.539[b]	Gold, SH0ES	8, 11, 18, 21, 22, 24, 28
DELTA CEP	5.3663	0.092	Gold	8, 11, 18, 19, 21, 22, 25, 26
DL CAS	8.0007	0.568, 0.569[b]	Gold, SH0ES	6, 8, 11, 18, 21, 22, 24, 25, 28
DR VEL	11.2078, 11.2078[a]	0.796[b]	Bronze +, SH0ES	3, 4, 13, 18, 21, 28
ER CAR	7.7185, 7.7188[a]	0.191[b]	Bronze +, SH0ES	1, 2, 18, 21, 22, 28
ETA AQL	7.1768	0.162	Gold	4, 8, 11, 12, 18, 19, 20, 22, 25
EZ VEL	34.5446, 34.5446[a]	1.012[b]	Bronze	17, 21, 22
FF AQL	4.4709	0.227	Silver	4, 5, 11, 12, 18, 20, 22, 25, 28
FK CAR	23.2445, 23.2445[a]	1.139[b]	Bronze	17, 21, 22
FM AQL	6.1143	0.669	Gold	4, 8, 11, 12, 18, 19, 21, 22, 24, 28
FN AQL	9.4832	0.531	Gold	11, 12, 18, 19, 21, 22, 24, 28
GH CYG	7.8179	0.705	Silver	8, 18, 21, 22, 24, 28
GH LUP	9.2767, 9.2763[a]	0.503[b]	Bronze	1, 4, 13, 16, 18, 21, 22
GQ ORI	8.6161, 8.6161[a]	0.316[b]	Bronze +, SH0ES	4, 8, 11, 21, 22, 28
GY SGE	51.6865, 51.6870[a]	1.379[b]	Bronze	12, 16, 21, 22, 24, 25
HW CAR	9.1994, 9.1995[a]	0.348[b]	Bronze +, SH0ES	18, 21, 22, 28
II CAR	64.2400, 64.2368[a]	1.742[b]	Bronze	17, 21, 22
KK CEN	12.1803, 12.1816[a]	0.682[b]	Bronze, SH0ES	1, 3, 4, 17, 18, 21, 22, 28
KN CEN	34.0191	0.874, 0.928[b]	Gold, SH0ES	2, 3, 4, 6, 16, 18, 21, 22, 28
KQ SCO	28.7050	0.988	Silver	1, 12, 13, 16, 18, 21, 22, 28
KX CYG	20.0467	1.923	Silver	21, 22, 24
L CAR	35.5523	0.223	Gold	3, 4, 16, 18, 21, 22, 25
LS PUP	14.1473	0.606	Gold	16, 17, 18, 21, 22, 28

Table E.1. Continued.

Star	Period (days)	E(B-V) (mag)	Sample	References
MW CYG	5.9547	0.754	Gold	8, 11, 18, 21, 22, 24, 28
OT PER	26.0910, 26.0889[a]	1.537[b]	Bronze	21, 22, 24
QY CEN	17.7524	1.407	Silver	3, 17, 21, 22, 28
R CRU	5.8258, 5.8257[a]	0.179[b]	Bronze, SH0ES	1, 4, 12, 18, 21, 22, 28
R MUS	7.5102, 7.5104[a]	0.136[b]	Bronze, SH0ES	1, 2, 7, 12, 18, 21, 22, 28
R TRA	3.3893, 3.3892[a]	0.192[b]	Bronze, SH0ES	4, 9, 12, 18, 21, 22, 28
RR LAC	6.4165	0.388	Gold	1, 6, 8, 11, 18, 21, 22, 24, 28
RS ORI	7.5666	0.385	Silver	1, 4, 8, 11, 18, 21, 22, 24, 28
RS PUP	41.4544	0.590	Gold	3, 4, 6, 11, 12, 16, 18, 21, 22, 27, 28
RT AUR	3.7283	0.096	Gold	1, 5, 11, 18, 19, 20, 21, 22, 24, 25, 28
RU SCT	19.7051	0.953	Gold	11, 12, 16, 18, 24, 25
RV SCO	6.0613, 6.0611[a]	0.411[b]	Bronze, SH0ES	4, 9, 11, 12, 18, 21, 22, 28
RW CAM	16.4149	0.686, 0.682[b]	Silver, SH0ES	1, 10, 11, 18, 21, 22, 24, 28
RW CAS	14.7937	0.502, 0.510[b]	Gold, SH0ES	10, 11, 18, 21, 22, 24, 28
RX AUR	11.6234	0.393	Gold	10, 11, 18, 22, 24
RX CAM	7.9120	0.630, 0.610[b]	Gold, SH0ES	8, 11, 18, 22, 24, 28
RY CAS	12.1373	0.737, 0.722[b]	Silver, SH0ES	21, 22, 24, 28
RY CMA	4.6786	0.273, 0.285[b]	Gold, SH0ES	1, 4, 11, 18, 21, 22, 24, 28
RY SCO	20.3228	0.747, 0.810[b]	Gold, SH0ES	3, 4, 6, 11, 12, 16, 21, 22, 28
RY VEL	28.1359	0.625, 0.619[b]	Silver, SH0ES	1, 3, 4, 12, 13, 16, 18, 21, 22, 28
RZ GEM	5.5290	0.574	Silver	8, 11, 18, 21, 22, 24, 28
RZ VEL	20.3953	0.437	Gold	1, 3, 4, 16, 18, 21, 22, 28
S CRU	4.6900, 4.6897[a]	0.200[b]	Bronze, SH0ES	1, 4, 9, 12, 18, 21, 22, 28
S MUS	9.6598	0.287	Gold	1, 2, 12, 16, 18, 21, 22, 28
S NOR	9.7542	0.290, 0.270[b]	Gold, SH0ES	2, 6, 12, 16, 18, 21, 22, 25, 28
S SGE	8.3821	0.148	Gold	2, 8, 11, 12, 18, 19, 20, 21, 22, 25, 28
S TRA	6.3235, 6.3236[a]	0.131[b]	Bronze, SH0ES	4, 9, 12, 18, 21, 22, 28
S VUL	68.5517	1.023, 1.004[b]	Gold, SH0ES	6, 12, 16, 21, 22, 24, 25
SS CMA	12.3553, 12.3553[a]	0.572[b]	Bronze, SH0ES	1, 3, 4, 13, 17, 18, 21, 22, 28
SS SCT	3.6713	0.398, 0.389[b]	Gold, SH0ES	4, 11, 12, 18, 21, 24, 28
ST TAU	4.0342	0.401	Gold	1, 11, 18, 19, 21, 22, 24, 28
SU CAS	1.9493, 1.9493[a]	0.282[b]	Bronze	1, 5, 11, 18, 19, 20, 21, 22, 28
SU CYG	3.8456	0.118	Gold	5, 11, 12, 18, 20, 21, 22, 24, 28
SV MON	15.2338	0.306	Silver	4, 10, 11, 18, 21, 22, 24, 28
SV PER	11.1293	0.376, 0.444[b]	Silver, SH0ES	10, 11, 18, 21, 22, 24, 28
SV VEL	14.0971, 14.0999[a]	0.417[b]	Bronze, SH0ES	1, 2, 3, 12, 18, 21, 22, 28
SV VUL	44.9385	0.564, 0.541[b]	Gold, SH0ES	1, 10, 11, 12, 16, 18, 19, 20, 21, 22, 24, 25, 28
SW CAS	5.4411	0.536	Gold	1, 8, 11, 18, 21, 22, 24, 28
SW VEL	23.4003	0.392	Silver	1, 3, 4, 13, 16, 18, 21, 22, 28
SX PER	4.2900	0.623	Silver	5, 18, 21, 22, 24, 28
SX VEL	9.5499	0.275	Silver	1, 2, 16, 18, 21, 22, 28
SY AUR	10.1440	0.510	Gold	6, 10, 11, 18, 21, 22, 24, 28
SY NOR	12.6452, 12.6491[a]	0.732[b]	Bronze, SH0ES	3, 4, 18, 21, 22, 28
SZ AQL	17.1415	0.692	Gold	4, 10, 11, 12, 16, 18, 19, 21, 22, 24, 25, 28
SZ CYG	15.1094	0.691, 0.732[b]	Gold, SH0ES	6, 10, 11, 18, 21, 22, 24, 28
T MON	27.0256	0.271, 0.312[b]	Gold, SH0ES	4, 6, 10, 11, 12, 13, 16, 18, 21, 22, 24, 25, 28
T VEL	4.6398	0.290, 0.369[b]	Gold, SH0ES	1, 4, 15, 16, 18, 21, 22, 28
T VUL	4.4355	0.079	Gold	1, 5, 11, 12, 18, 19, 20, 21, 22, 25, 28
TT AQL	13.7550	0.551	Gold	4, 10, 11, 12, 18, 19, 20, 21, 22, 24, 25, 28
TV CMA	4.6700	0.666	Silver	4, 18, 21, 22, 24, 28
TW CMA	6.9955	0.434	Silver	4, 11, 18, 21, 22, 24, 28
TX CYG	14.7098, 14.7116[a]	1.267[b]	Bronze, SH0ES	1, 10, 11, 18, 21, 22, 24, 28
TX MON	8.7017, 8.7017[a]	0.641[b]	Bronze	1, 4, 11, 18, 21, 22, 24, 28
TY SCT	11.0520	1.078	Silver	1, 3, 4, 18, 21, 22, 24, 28
TZ MON	7.4282	0.503	Silver	1, 4, 18, 21, 22, 24, 28
U AQL	7.0239	0.417, 0.415[b]	Gold, SH0ES	4, 11, 12, 18, 20, 21, 22, 24, 25
U CAR	38.8484	0.320, 0.338[b]	Silver, SH0ES	1, 3, 4, 12, 13, 16, 18, 22, 25, 28
U NOR	12.6437	1.006	Silver	3, 4, 13, 16, 18, 21, 28
U SGR	6.7453	0.486, 0.475[b]	Gold, SH0ES	1, 4, 6, 11, 12, 16, 18, 21, 22, 24, 25, 28
U VUL	7.9907	0.666	Gold	8, 11, 18, 19, 20, 21, 22, 24, 25, 28

Table E.1. Continued.

Star	Period (days)	E(B-V) (mag)	Sample	References
UU MUS	11.6361	0.500, 0.511[b]	Silver, SH0ES	1, 2, 3, 13, 16, 18, 21, 28
UZ SCT	14.7486	1.174	Gold	4, 21, 22, 24
V0339 CEN	9.4660, 9.4667[a]	0.593[b]	Bronze +, SH0ES	1, 4, 13, 18, 21, 22, 28
V0340 ARA	20.8140, 20.8123[a]	0.751[b]	Bronze +, SH0ES	1, 3, 18, 21, 22, 28
V0350 SGR	5.1545	0.337	Gold	11, 12, 18, 21, 22, 28
V0386 CYG	5.2580	0.929, 0.970[b]	Gold, SH0ES	8, 11, 18, 21, 22, 24, 28
V0396 CYG	33.2470, 33.2450[a]	1.820[b]	Bronze	21, 22, 24
V0402 CYG	4.3651	0.416	Gold	5, 11, 18, 21, 22, 24, 28
V0459 CYG	7.2517	0.801	Gold	11, 18, 21, 22, 24, 28
V0482 SCO	4.5278, 4.5278[a]	0.420[b]	Bronze +, SH0ES	21, 22, 28
V0508 MON	4.1337	0.359	Silver	18, 21, 22, 24, 28
V0538 CYG	6.1191	0.668	Gold	8, 18, 21, 22, 24, 28
V0609 CYG	31.0880	1.441	Silver	21, 22, 24, 28
V0636 SCO	6.7969	0.234, 0.279[b]	Gold, SH0ES	1, 2, 12, 18, 21, 22, 28
V1162 AQL	5.3760	0.225	Gold	18, 21, 22, 24, 28
V1467 CYG	48.5117	1.777	Silver	21, 22, 24
V CAR	6.6968	0.190, 0.239[b]	Silver, SH0ES	1, 4, 16, 18, 21, 22, 28
V CEN	5.4940	0.307	Silver	4, 9, 12, 16, 18, 21, 22, 25, 28
V LAC	4.9836	0.430	Gold	5, 11, 18, 21, 22, 24, 28
V VEL	4.3710, 4.3711[a]	0.224[b]	Bronze +, SH0ES	1, 18, 21, 22, 28
VV CAS	6.2071	0.567	Silver	8, 18, 21, 22, 24, 28
VW CEN	15.0371	0.421, 0.641[b]	Gold, SH0ES	4, 13, 16, 18, 21, 22, 28
VX CYG	20.1334	1.101	Silver	21, 22, 24, 28
VX PER	10.8863, 10.8863[a]	0.522[b]	Bronze, SH0ES	21, 22, 24, 28
VY CAR	18.9019	0.347, 0.366[b]	Gold, SH0ES	1, 3, 4, 6, 12, 16, 21, 22, 28
VY CYG	7.8571	0.691	Silver	1, 8, 11, 18, 21, 22, 24, 28
VZ CYG	4.8645	0.287	Gold	5, 11, 12, 18, 19, 21, 22, 24, 28
VZ PUP	23.1741	0.536, 0.480[b]	Gold, SH0ES	1, 3, 4, 6, 13, 16, 17, 18, 21, 22, 28
W GEM	7.9136	0.306, 0.309[b]	Silver, SH0ES	1, 4, 8, 11, 18, 20, 21, 22, 24, 28
W SGR	7.5950	0.149	Gold	1, 2, 11, 12, 18, 21, 22, 25, 28
WX PUP	8.9370, 8.9366[a]	0.359[b]	Bronze, SH0ES	4, 11, 17, 18, 21, 22, 28
WZ CAR	23.0132	0.413	Silver	16, 17, 21, 22
WZ SGR	21.8476	0.631, 0.644[b]	Gold, SH0ES	2, 3, 6, 11, 12, 13, 16, 18, 21, 22, 24, 25, 28
X CYG	16.3859	0.297, 0.324[b]	Gold, SH0ES	6, 10, 11, 12, 18, 19, 20, 21, 22, 25, 28
X PUP	25.9710	0.547, 0.495[b]	Gold, SH0ES	3, 4, 6, 11, 16, 18, 21, 22, 28
X SCT	4.1980	0.674	Silver	4, 18, 21, 22, 24, 28
X SGR	7.0128	0.277	Gold	1, 4, 11, 12, 18, 21, 22, 23, 25, 28
X VUL	6.3203	0.806	Gold	8, 11, 18, 21, 22, 24, 28
XX CAR	15.7089, 15.7089[a]	0.366[b]	Bronze, SH0ES	2, 3, 14, 18, 21, 22, 28
XX CEN	10.9540	0.298	Gold	2, 12, 14, 16, 18, 22
XX SGR	6.4243	0.633	Gold	11, 12, 18, 21, 22, 28
XY CAR	12.4464, 12.4464[a]	0.538[b]	Bronze +, SH0ES	2, 3, 14, 18, 21, 28
XZ CAR	16.6499, 16.6519[a]	0.556[b]	Bronze +, SH0ES	1, 2, 3, 14, 18, 21, 22, 28
Y AUR	3.8595	0.445	Silver	5, 18, 21, 22, 24, 28
Y LAC	4.3238	0.206	Gold	5, 11, 18, 19, 21, 22, 24, 28
Y OPH	17.1250	0.688	Gold	4, 11, 12, 13, 16, 18, 22, 25, 28
Y SGR	5.7734	0.271	Gold	1, 4, 11, 12, 18, 21, 22, 25, 28
YZ AUR	18.1957	0.635	Silver	10, 18, 21, 22, 24, 28
YZ CAR	18.1680, 18.1669[a]	0.430[b]	Bronze, SH0ES	2, 3, 12, 13, 18, 21, 22, 28
YZ SGR	9.5537	0.366, 0.369[b]	Gold, SH0ES	1, 2, 11, 12, 18, 21, 22, 24, 28
Z LAC	10.8857	0.460, 0.460[b]	Gold, SH0ES	1, 6, 10, 11, 18, 19, 21, 22, 24, 28
Z SCT	12.9014	0.642	Gold	3, 4, 6, 11, 18, 21, 24, 28
ZETA GEM	10.1498	0.040	Gold	10, 11, 18, 20, 21, 22, 23, 25

References. (1) AAVSO; (2) Walraven et al. (1964); (3) Madore (1975); (4) Pel (1976); (5) Szabados (1977); (6) Harris (1981); (7) Lloyd Evans (1980); (8) Szabados (1980); (9) Gieren (1981); (10) Szabados (1981); (11) Moffett & Barnes (1984); (12) Welch et al. (1984); (13) Coulson & Caldwell (1985); (14) Coulson et al. (1985); (15) Gieren (1985); (16) Laney & Stobie (1992); (17) Schechter et al. (1992); (18) ESA (1997); (19) Barnes et al. (1997); (20) Kiss (1998); (21) Cutri et al. (2003); (22) Berdnikov (2008); (23) Feast et al. (2008); (24) Monson & Pierce (2011); (25) Monson et al. (2012); (26) Engle et al. (2014); (27) Kervella et al. (2017); (28) Gaia Collaboration (2018).

Notes. [a]Phasing period (bronze sample) [b]Fitted value during template fitting (bronze and SH0ES samples)

Chapter 4

Discussion and future research

1 Observing the pulsation to understand the p-factor

The most recent results on the calibration of the p-factor of pulsating stars, whether for Cepheids or RR Lyrae confirmed the individual dispersion of the p-factor values. The conclusions I drew in Bras et al. (2024) suggest that we are lacking an ingredient in our understanding of the pulsation, which is revealed when comparing spectroscopic integrated Radial Velocities (RVs) and photometric diameters in the parallax-of-pulsation technique. In particular, we suspect a possible decoupling in the different layers of the star during the pulsation. This section presents two different directions to find possible evidence for such a decoupling, either by observing the photosphere or the line-forming regions.

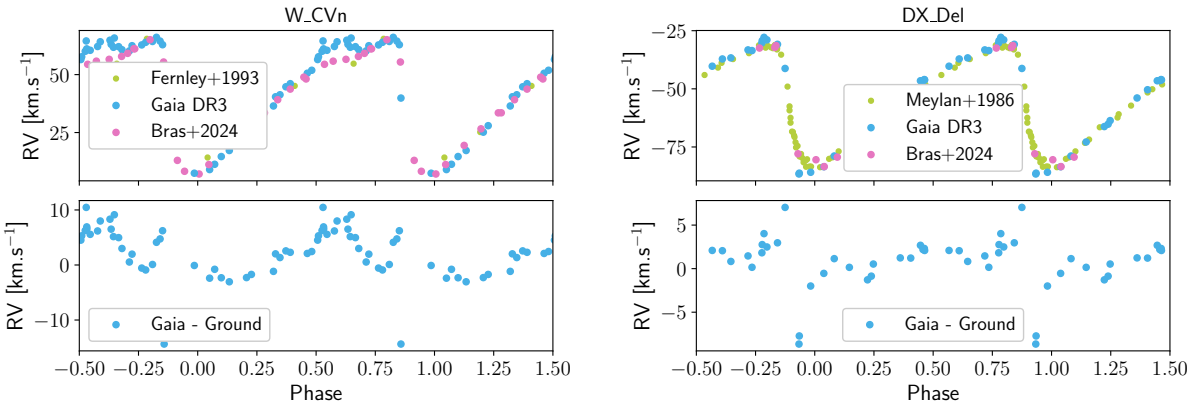
1.1 Tomographic analysis of the atmosphere of pulsating stars

The first idea was to observe the pulsation on the different layers of the atmosphere of the star by differentiating different line-forming regions, isolating specific set of lines.

1.1.1 A look at Gaia Radial Velocity curves

With the recent Gaia Data Realeas 3 (DR3), published just before the beginning of my PhD, epoch RVs have been published for 1100 RR Lyrae stars and 798 Cepheids in the Milky Way. More will become public in the DR4 and the final DR5 releases. This RV survey provides a powerful tool for pulsating star characterisation. For technical reasons, the spectral region of the Gaia Radial Velocity Spectrometer (RVS) had to be as narrow as possible. This was achieved by choosing the region around the calcium II triplet (847 to 874 nm), as it contains major information for the RV and the stellar parameter determination. It follows that the calcium element drives the RV determination, done using cross-correlation with a template spectrum. However, these lines are supposed to form high in the chromosphere of the pulsating star, and may be less tied to the photosphere pulsation. Having a large homogeneous set of RVs and photometry of pulsating stars is particularly valuable for the parallax-of-pulsation technique. However, the specific region observed by the RVS provides different RV measurements as the

ones obtained from the ground using the Cross-Correlation Function (CCF) approach. Ground-based measurements were compared with RVS determinations for several RR Lyrae stars in Clementini et al. (2023). They showed generally a good agreement, however, for some stars, they pointed out a difference near the maximum of RV. In order to add Gaia RV measurements to my study of the p-factor of RR Lyrae stars, I also investigated this peculiar feature. We observe that the calcium II lines are reaching a first maximum velocity a little in advance compared to CCF velocities, followed by the global maximum, common to all lines. The difference observed is quite important, reaching around 10 km.s^{-1} for SW Dra for instance. In Clementini et al. (2023), this effect was observed only for 5 out of the 19 RR Lyrae stars analysed, however, I could identify a similar feature in at least 12 out of the 20 stars in my sample with Gaia RV data. From the visual inspection of the data in my sample, it seems that this effect may be of different amplitude for the different targets. I present in Fig. 4.1 two comparisons between ground and Gaia RV curves. For the star W CVn, a strong bump before the maximum is observable in the Gaia RV data, however, there is no strong evidence for such a feature in DX Del. For W CVn, it is only visible thanks to the new data I obtained on this star.



(a) (b)

Figure 4.1: RV curve from ground-based observations compared to Gaia RVS data for (a) W CVn and (b) DX Del.

I estimated in Bras et al. (2024) the impact of taking into account this bump in the amplitude of the integrated RV to 15% on SW Dra, leading to a potential 15% difference on the p-factor. However, this difference is reduced for some other stars, for instance for AB UMa, where this feature is less pronounced. A refined estimation of the effect on the p-factor for the whole sample, including on the dispersion of the p-factors, would need a good interpolation, and therefore a good phase sampling. The Gaia DR4 will include more RV measurements as well as epoch spectroscopy, which will enable us to explore the origin of the bump and constrain its statistical effect on the p-factor.

1.1.2 Tomographic spectral masks

One idea investigated in Bras et al. (2024) was to derive RV curves for different set of spectral lines forming in different regions of the star's atmosphere. We used as an indicator of

the optical depth of the lines their spectral depths. From this approach, we could not identify a strong effect on the RV curves and hence on the p-factor because of the limited signal-to-noise (S/N) ratio of our observations. To complement this study, we discussed with Kateryna Kravchenko, who developed a tomographic method to recover the velocity distribution of the atmospheres of cool giants and supergiants. The idea of Kravchenko et al. (2018) was to compute different tomographic masks for lines forming at different optical depths, from synthetic spectra. Applying such masks to real spectra requires high-quality data, and the SOPHIE (Spectrographe pour l'Observation des Phénomènes des Intérieurs stellaires et des Exoplanètes) spectra have too low S/N for such a study, as the exposure time was limited. I used HARPS (High Accuracy Radial velocity Planet Searcher) spectra from Borgniet et al. (2019), and selected the Cepheid RZ Vel for that study. From its stellar parameters, K. Kravchenko computed the CCF masks adapted to this Cepheid, using the corresponding MARCS (Model Atmospheres with a Radiative and Convective Scheme) atmosphere model, with $T_{\rm eff} = 5500$K, $\log g = 1.0$, solar mass and solar metallicity. This resulted in 6 masks for 6 different layers of optical depths. The mask numbered 1 has the highest optical depth meaning that it probes the innermost layer of the atmosphere and reversely the mask numbered 6 the outermost layers. The masks correspond to a list of central wavelength of the lines forming at the corresponding regions. I converted them into cross-correlating masks adding a width of 0.02 nm. Each mask was cross-correlated with the different spectra of RZ Vel at the different pulsation phases.

1.1.3 Complex cross-correlation profiles

A typical set of CCF for the different masks is shown in Fig. 4.2.

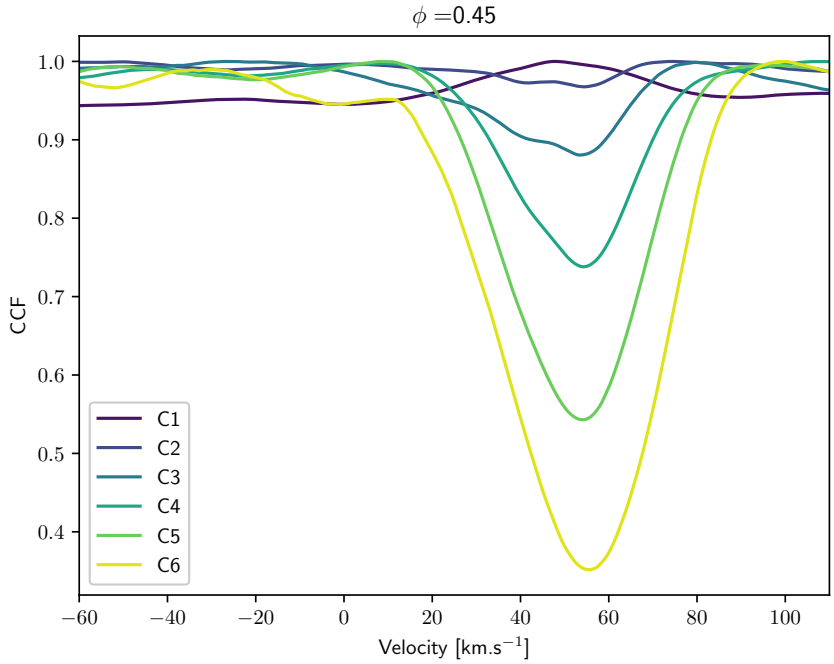


Figure 4.2: CCF of HARPS spectrum of RZ Vel at $\phi = 0.45$ with the 6 tomographic masks. C1 selects the innermost lines and C6 the outermost.

First, we observe that the signal for the CCF with the deepest mask (in terms of optical depth, i.e. C1) is too low to be exploited. A second remark is that this figure argues is favour of the assumption made for the study of RR Lyrae, as the spectral depth decreased for the masks from 6 to 1. Another interesting observation is the shape of the CCF profiles, which is quite far from a Gaussian. A two-sided CCF seems to appear for C2, C3 and C4. Also, for C6, a wing appears in the red outside the main CCF profile. The overall shape makes it difficult to associate a single RV measurements with standard approaches. The RV curve from a fourth order polynomial fit of the CCF is given in Fig. 4.3 for the four masks with the highest signal. Additionally, the shape of the CCF differs significantly at the different phases of pulsation, including the depth and the width of the CCF as well as their asymmetry.

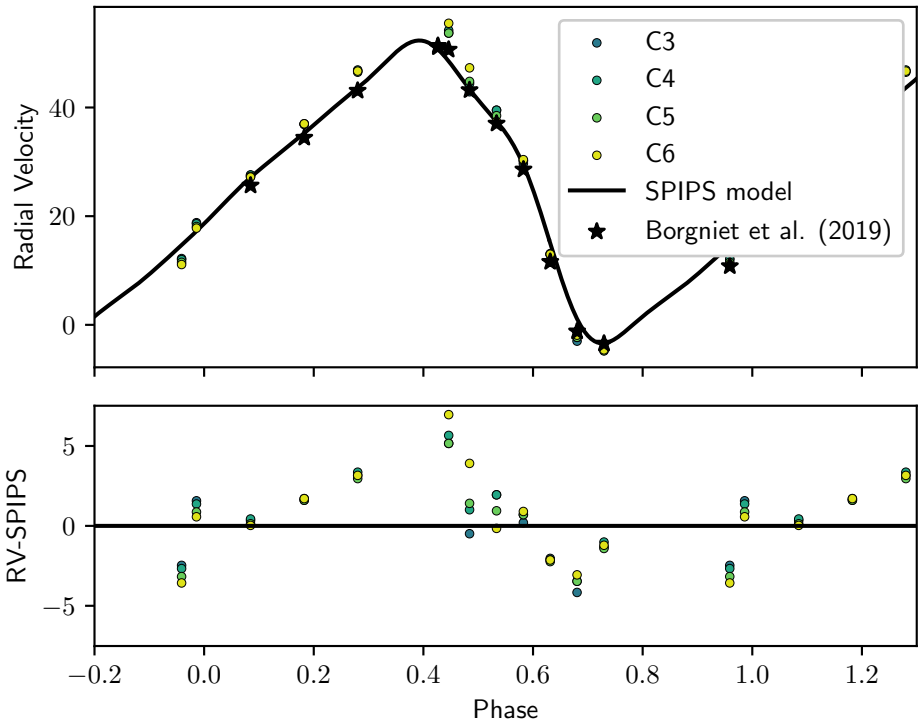


Figure 4.3: The RV curve of the Cepheid RZ Vel obtained with four tomographic masks, and its difference compared to standard measurements (Gaussian fit, Borgniet et al. 2019)

The complete picture of the motion of the different layers of spectral lines is given in Fig. 4.4.

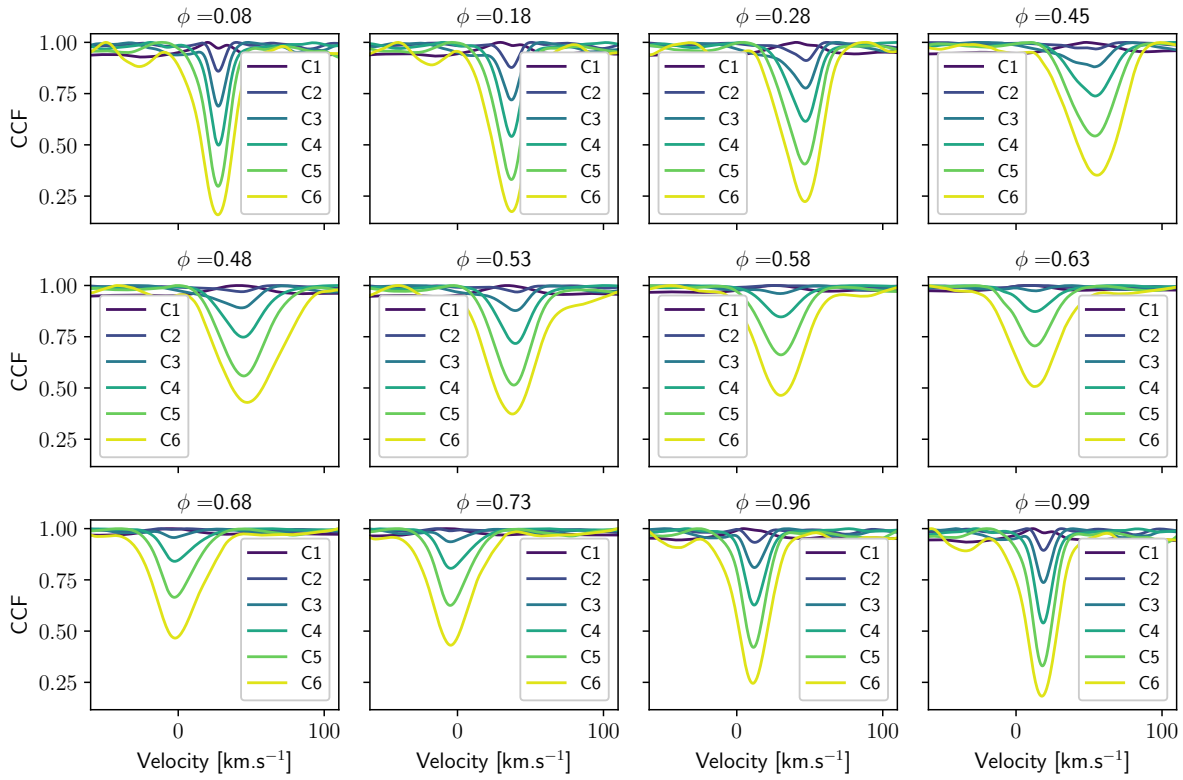


Figure 4.4: The varying CCF profiles at the different pulsation phases for the Cepheid RZ Vel with the 6 tomographic masks.

Although it gave promising results, I could not go further on this study as it would need more extensive work to provide conclusions on the RV curves and eventually on the p-factor. Nevertheless these CCF profiles suggest complex underlying stellar atmosphere physics that deserves further investigation.

1.2 Variability of the limb-darkening: extension to other targets

The other direction we followed for inspecting the source of the dispersion of p-factors is the variability of the Limb-Darkening (LD). In Chapter 2, I presented the work done on the Northern Cepheid η Aql. However, it would be interesting to extend the analysis to other Cepheids in order to see if we can draw different conclusions.

In the Northern Hemisphere, the only optical interferometric instrument available is the CHARA (Center for High Angular Resolution Astronomy) interferometer. With the MIRC-X (Michigan InfreRed Combiner-eXeter) beam combiner, operating in the H band (1662nm) and a maximum baseline of 331m, we know that the first zero of visibility occurs at $\theta = \frac{1.22\lambda}{B}$. This means that the second lobe is observable for stars slightly larger than $\frac{1.22\lambda}{B} = 1.26$mas. It reduces significantly the number of Cepheids for which the second lobe is observable with CHARA. A possible target is Polaris (α UMi). However Polaris is not the most standard Cepheid. It pulsates in the first overtone, with a particularly low amplitude. It has also shown rapid period change

and even a sudden period change around 1960. Besides these peculiarities, its large angular diameter is around 3.1 mas, which makes it well suitable for LD measurement. In practice, it is observable with four out of the six telescopes, and we are able to resolve up to the third lobe of visibility. Additionally, Polaris has been extensively observed in the past, to study its triple stellar system. Recent observations by Evans et al. (2024) with CHARA were used to reconstruct images of the surface of the star, showing spots and surface inhomogeneities. Our observing proposal for 2024A period was focused on observations of both η Aql and Polaris, separated in two different runs to cover their respective pulsation cycles. However, the weather during the observing run for Polaris made the observations impossible. During our July 2024 run on η Aql, we added a few observations of Polaris, to have a preliminary investigation on its visibility curves. Preliminary reduced data are shown in Fig. 4.5. This figure shows well the high resolving power of the instrument for this star. The second lobe is fully resolved by MIRC-X, allowing us to well constrain the LD parameter.

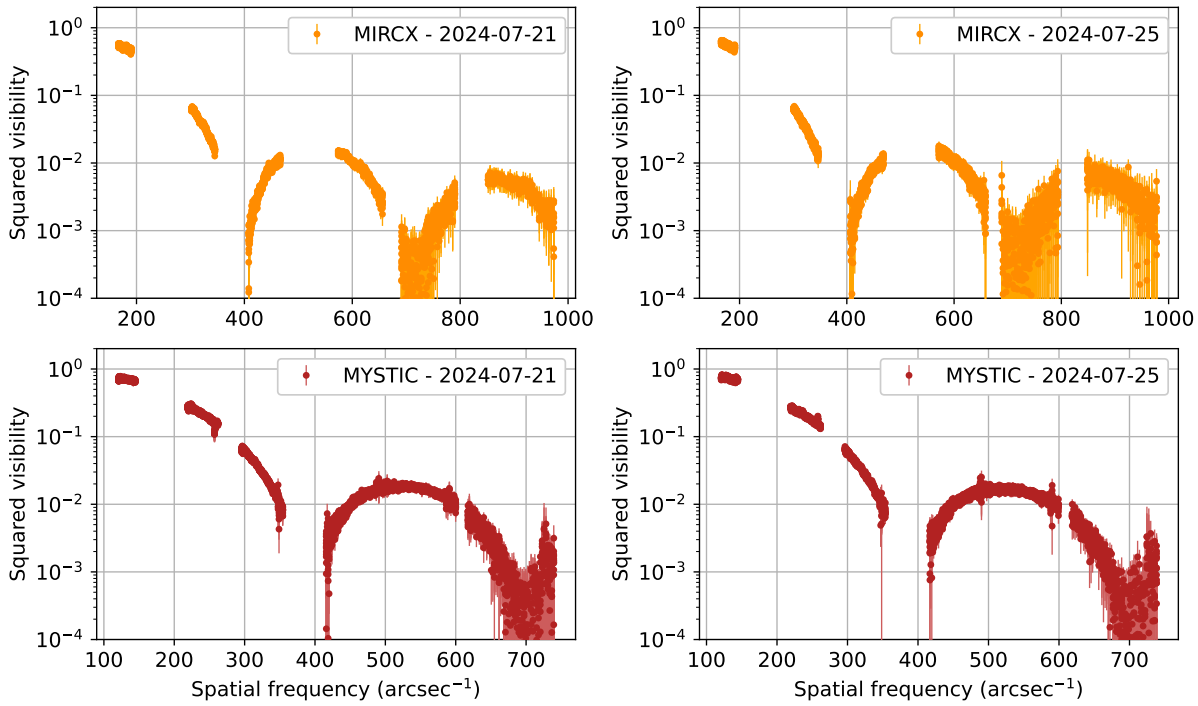


Figure 4.5: Squared visibilities for preliminary data of Polaris during the 2024A observing program.

We asked for more observations in the next cycle, which occurred in July 2025. The data are now acquired and are being reduced and analysed. The aim is the same as for η Aql, in addition with a look to more refined LD profiles and image reconstruction to probe the photosphere of the Cepheid during its pulsation. The presence of photospheric spots can be studied thanks to closure phase signals, as shown in Fig. 4.6 for July 2024 data. A purely spherical object, with no signs of asymmetry would give a closure phase signal always equal to 0 or $\pm 180°$.

In parallel, it is possible to increase the number of targets for our study by two different means. Either we can increase the baselines, which is part of the projects for CHARA, with the

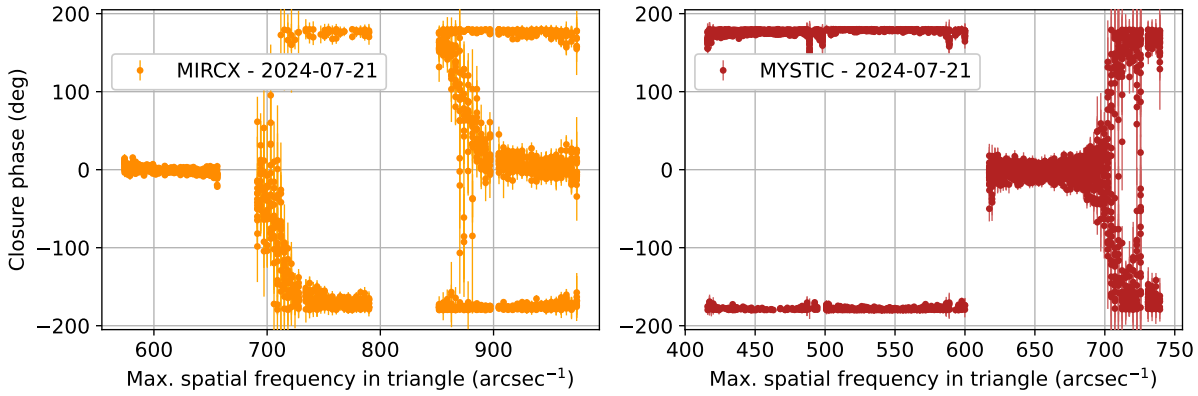


Figure 4.6: Closure phases for preliminary data of Polaris during the 2024A observing program.

objectives to go towards the kilometre baseline. Or we can observe at shorter wavelengths, which is made possible by the new SPICA (Stellar Parameters and Images with a Cophased Array) instrument. SPICA works in the V-band and can be used simultaneously with the MIRC-X and MYSTIC (Michigan Young STar Imager at CHARA) instruments. The instrument is currently available in shared risk. For the moment, it is not reaching its full capabilities because of technical issues on site. In the future, it will be a major instrument to study the LD of Cepheids in the visible.

In the Southern Hemisphere, interferometry is done from the Very Large Telescope Interferometer (VLTI) in Paranal, Chile. The VLTI can combine observations from the 4 Unit Telescopes of 8.2m diameter and the 4 Auxiliary Telescopes of 1.8m diameter. The Auxiliary Telescopes can be moved to cover the (u,v) plane, with the maximum baseline reaching 202m. The current instrumentation includes three beam combiners. MATISSE (Multi AperTure mid-Infrared SpectroScopic Experiment) operates in the infrared (IR) bands L, M and N, GRAVITY in the K band and PIONIER (Precision Integrated-Optics Near-infrared Imaging ExpeRiment) in the H band. Following the basic equation for the limit of resolved stars up to the zero of visibility, the limit for these three instruments is of 4mas, 2.5mas and 1.8mas respectively. This basically limits the potential targets to the Cepheid ℓ Car for the study of the LD. The new Asgard Suite is currently being mounted in order to complement the instrumentation already available. It will be composed of four modules. Very recently, in summer 2025, the BIFROST combiner have been integrated on site. It is now in its commissioning phase with time allocated to the community on the P117 ESO call. The advantage of the BIFROST instrument is that it will be observing in the Y and J bands, typically at 1.35 μm. The second lobe of β Dor and the third lobe of ℓ Car should be accessible with this capability. We hope that the commissioning will provide the expected specifications, and we already submitted a proposal to observe the two largest Cepheids in the Southern Hemisphere to increase the number of Cepheids in our study. In particular, ℓ Car is interesting to complement the observations of η Aql and Polaris as its pulsation period is 30 days, placing it in the long-period Cepheid population of the Milky Way.

2 The limitation of light curve templates

The other part of my work was the development of generic templates for light-curve fitting of Cepheids. In this section I want to highlight the intrinsic limitations of this approach and how my templates can be used, depending on the science case, to benefit from their flexibility.

2.1 Average behaviour

The light-curve templates are constructed in order to minimise the number of free parameters in the characterisation of the shape. The drawback of this approach is that it will only incorporate the average behaviour in the whole sample of Cepheids.

From pulsation theory, we know that the pulsation period is related to the location in the Instability Strip (IS) of the Hertzsprung-Russell (HR) diagram. Therefore, the shape of the light curve is expected to be related to its position in the IS as well. In particular, we observe a high correlation between the first Principal Component Analysis (PCA) coefficient K_1 and the temperature. From that observation, we understand that we average the variety of Cepheids in the width of the IS (perpendicular to the period variation). In Fig. 4.7, I fitted the linear relation between the temperature and the first PCA coefficient. I also checked that the dispersion around this fitted relation does not correspond to the dispersion inside the IS. This means that we still have individual specificities in the light-curve shapes that does not simply correlate with the stellar parameters.

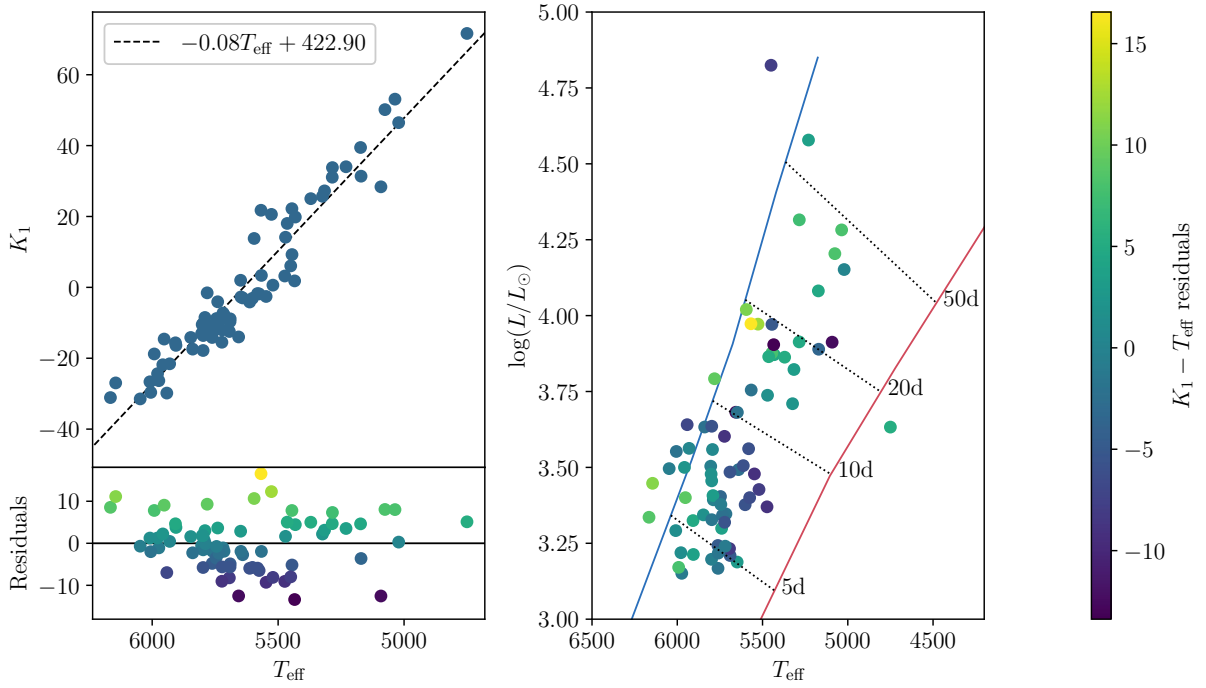


Figure 4.7: Relations between the first PCA coefficient and the stellar parameters involved in the HR diagram.

In particular, a light curve can be characterised by its amplitude. From the past observations, the amplitude is not fully correlated with the period of the star, and can differ between Cepheids with similar periods. If no additional parameter that allows for such variations is applied in the template fitting process, the amplitude will be forced to the average amplitude for Cepheids with the same period. As a test, I compared the amplitude given by my templates in the V band with measured amplitudes as collected in the catalogue of Berdnikov (2008). In Fig. 4.8, I show the measured amplitude with the sliding median of observed amplitudes, with period bins of 0.08 in $\log(P)$. The Period-Amplitude relation of my templates is shown as a solid line. We estimate that approximately 70% of the Cepheids have their amplitude underestimated when using the templates, and the other 30% have their amplitude overestimated.

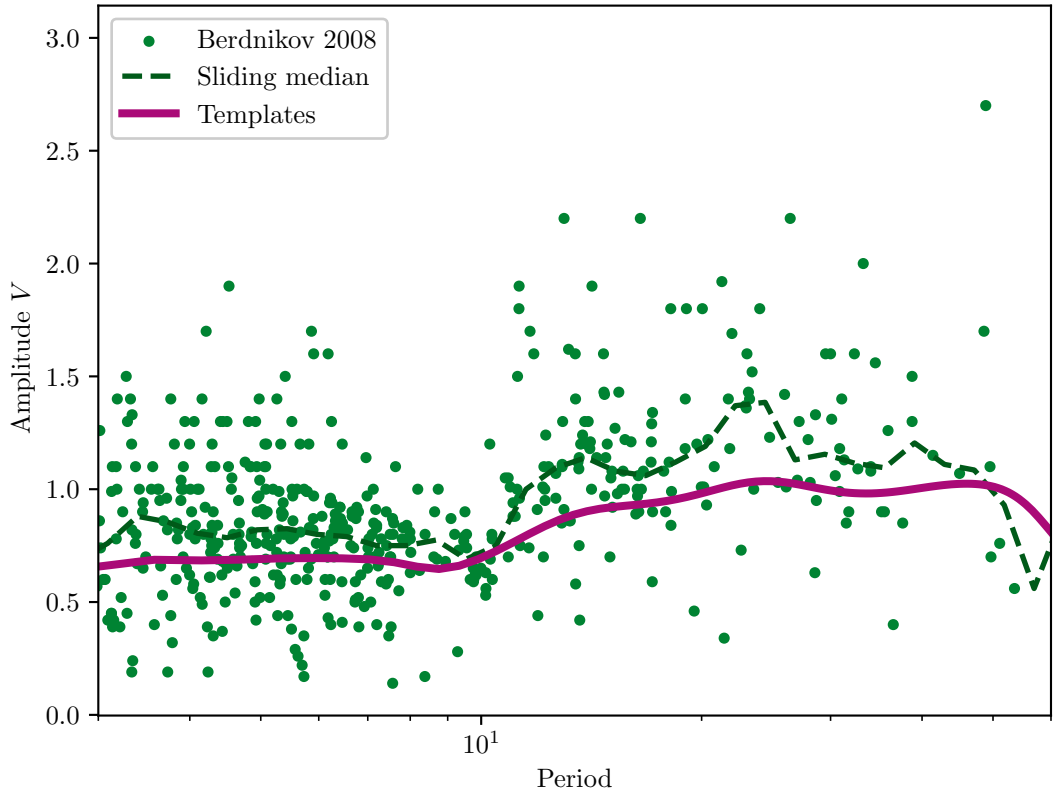
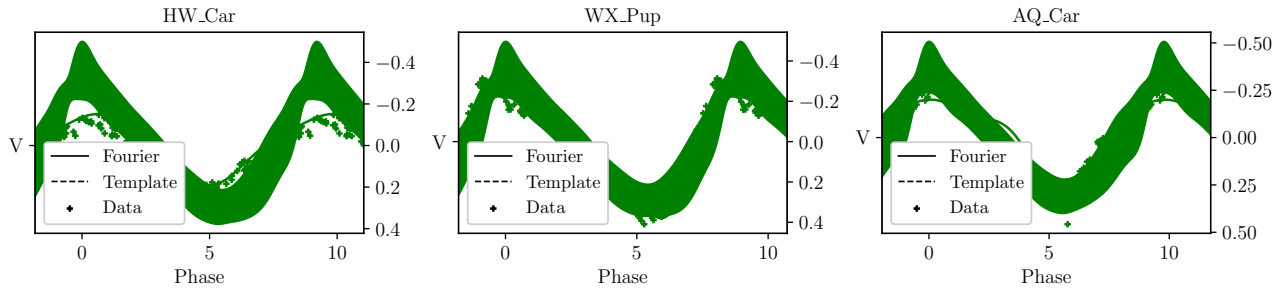


Figure 4.8: Period-amplitude relation of my templates in the V band compared to observed amplitudes.

The templates are generalising the light-curve shapes of Cepheids. They depend on the input training sample, which should include the diversity of shapes observables. In Bras et al. (2025), I compared the light-curve shape obtained from my templates for a set of Cepheids with the Fourier interpolation performed by Bhardwaj et al. (2015). The comparison is in favour of the templates for the shape agreement. However, I show in Fig. 4.9 other examples that are not as convincing as the one presented in Bras et al. (2025). The case of HW Car is typical of a star for which the observed amplitude is considerably different from the average amplitude for this period range. For WX Pup, the global shape differs, mainly from the presence of a

sharp bump in the observed magnitudes (the period of this star is around 9 days, close to the Hertzsprung progression centre). The ascending branch is also shorter in reality than predicted by the templates. The case of AQ Car is different as the shape does not agree between the Fourier interpolation and the template at the corresponding period but the Fourier interpolation is not fitting well the observations plotted for visual inspection. In the opposite, the templates seem particularly adequate for this star.



(a) (b) (c)

Figure 4.9: Comparison between the Fourier interpolation and the predicted curve from my templates for three galactic Cepheids; (a) for HW Car (P=9.20 d), (b) for WX Pup (P=8.94 d) and (c) for AQ Car (P=9.77 d).

In addition, a comparison with the template shapes of Yoachim et al. (2009) can be found in Sect. 3.2.

2.2 Reproducing the secondary bump

One characteristic high-frequency feature of light curves of Cepheids is the presence of the secondary maximum, highlighted by the Hertzsprung progression. This effect concerns specifically the period bin around 10 days, which is not characteristic of the main extragalactic Cepheid population we are interested in. Nonetheless, it is important to keep in mind the limitations of the templates in reproducing such features.

In agreement with the training sample provided for the template generation, the Hertzsprung progression can be observed in the templates. For a 6-day period Cepheid, the secondary bump is identifiable in the descending branch of the light curve, although with low amplitude. At 10 days, the shape of the ascending branch changes significantly as the bump shows up. It then disappears near the maximum for longer wavelengths. By determining the amplitude of the global curve it is possible to determine the central period of the progression. Indeed, the centre period corresponds to the period where the secondary maximum hits the global maximum, which tends to flatten the curve. The amplitude is minimal around this central period. This method has been used in Marconi et al. (2015). From the amplitude of the templates, we can clearly see the decrease around 10 days of period, for instance in Fig. 4.8. The centre period obtained with this method is in agreement with what is observed on real light curves. However, the amplitude of the secondary bump is quite small compared to what is observed on several stars. We argue

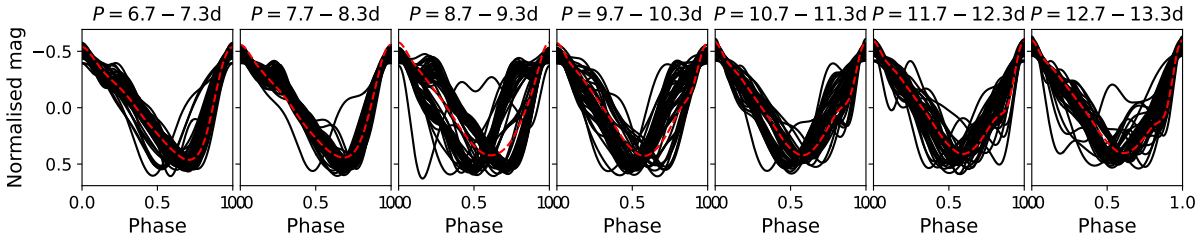


Figure 4.10: Fourier decomposition of fundamental Cepheids in 7 period ranges around the Hertzsprung progression centre (black) and the corresponding templates (red, dashed line).

that this comes from averaging the statistics on this secondary bump. Different stars may show a more or less strong bump in their light curves, and it can also be a bit dephased for Cepheids with similar periods, leading to a general smoothing of the feature. This approach has the drawback that strong bumps would not be recovered by the template fitting, which is particular to this type of analysis. As a comparison, the shape of the templates from Yoachim et al. (2009) also shows mitigated bumps, especially on the descending branch, for the shorter periods. In addition, as the purpose of these templates is primarly to optimise light curves for extragalactic Cepheids, which generally have periods significantly longer than 10 days to be detected. The Fig. 4.10, represents the Hertzsprung progression as observed in Gaia light curves together with the evolution of my templates. The bump for shorter periods is present on observed light curves, but is too small to appear on the templates. Around P=9 days, we can see the drastic change in the light curve shape. The ascending branch becomes suddenly longer, with the presence of the secondary maximum, separating the curves in two populations. For longer periods, we can see that the templates tend to average the secondary maximum in the rising branch.

The fact that the generic templates do not perfectly reproduce the secondary maximum is of secondary importance compared to their ability to retrieve accurate mean magnitudes. Indeed, the main idea of the templates is not to reproduce high-frequency structures, but to improve the determination of Period-Luminosity Relations (PLRs). For this, they aim at capturing the general pulsation curve shape. Hence, the effect of the amplitude of the secondary maximum and its difference with how it appears in the templates on the mean magnitudes remains small. In Bras et al. (2025), I compared the mean magnitudes in Hubble Space Telescope (HST) filters for 73 Cepheids in Riess et al. (2021). Around the centre period of the Hertzsprung progression, we do not see an increase of this difference, which rejects the possibility of a strong bias on mean magnitudes due to the bump of the Hertzsprung progression. I also show in Fig. 4.11 the distribution of the differences observed in the mean magnitudes in the F555W HST filter between the template fitting procedure and the real modelled curves from the training sample (in green). I added the comparison with the Riess et al. (2021) as in the Figure 15 of Bras et al. (2025). The sliding average lines are able to trace a potential bias due to the secondary bump of the Hertzsprung progression. However, for both comparison, the sliding average remains flat around the centre period, which is in favour of the low impact of the secondary maximum on the global mean magnitude.

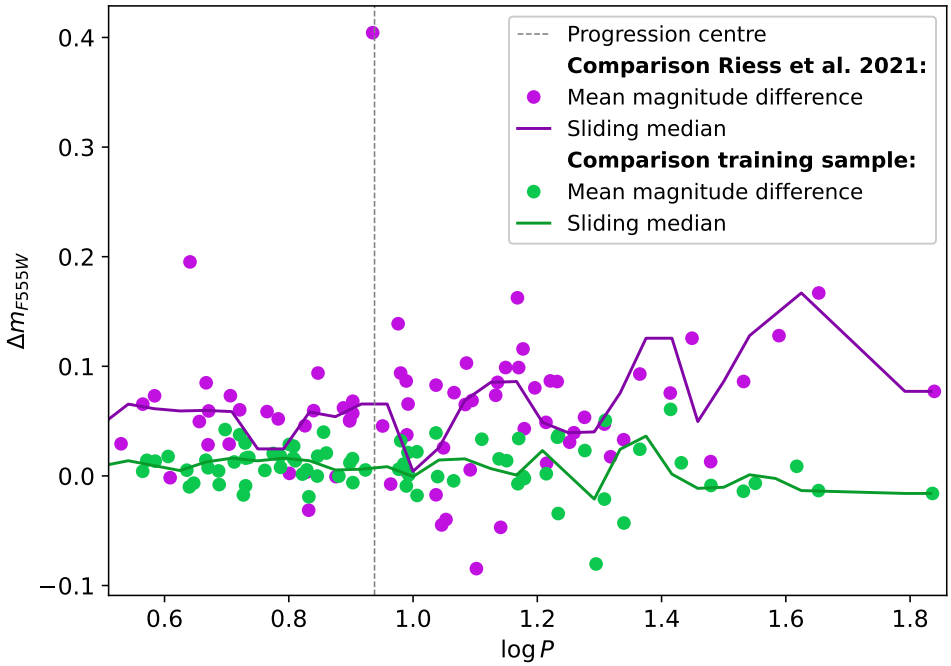


Figure 4.11: Difference in the mean magnitudes with the template fitting for the stars in Riess et al. (2021) and the modelled curve of the training sample plotted against the period.

2.3 Template fitting for precise light curve interpolation

Knowing the limits of the templates enables us to use them wisely, depending on the science objectives and the quality of the data. To conclude this chapter, I suggest different approaches to use my templates and avoid their potential biases.

2.3.1 Fitting K_i coefficients

The first alternative fitting strategy requires a good phase coverage of the pulsation, typically one point every 0.05 in phase. If this condition is respected, then it is possible to keep more than two parameters for one light curve (period and mean magnitude for the template fitting). Moreover, more parameters are needed to recover precisely the shape of the light curve. This can be achieved by relaxing the constraints on the PCA coefficients imposed by the $P-K_i$ relations. In practice, these relations can be used as initial guesses, while allowing the coefficients to vary around the predicted values in order to better reproduce the light-curve shape. This approach also combines different degrees of adaptability. The PCA coefficients can evolve freely, and the fit is similar to decomposition by Fourier series although benefitting from the appropriate basis of functions. Alternatively, we can limit the potential solutions by setting the space of the K_i parameters to the confidence intervals computed together with the $P-K_i$ relations. By performing such kind of fit, we are taking full advantage of the available observations. In particular, this can avoid the smoothing of high-frequency structures such as the secondary maximum of the Hertzsprung progression, or poor estimations of the amplitude of the light curve. This is illustrated in Fig. 4.12 with the Cepheid SX Vel ($P = 9.54993$ d). The template with fixed

PCA coefficient poorly reproduces the secondary maximum and the amplitude, particularly in the IR bands. By fitting the PCA coefficient, the light curve shape is more accurately recovered. The template fitting method requires only 4 parameters in total, where the PCA decomposition needs 8 parameters, including in addition the K_i coefficients up to the fourth PCA component.

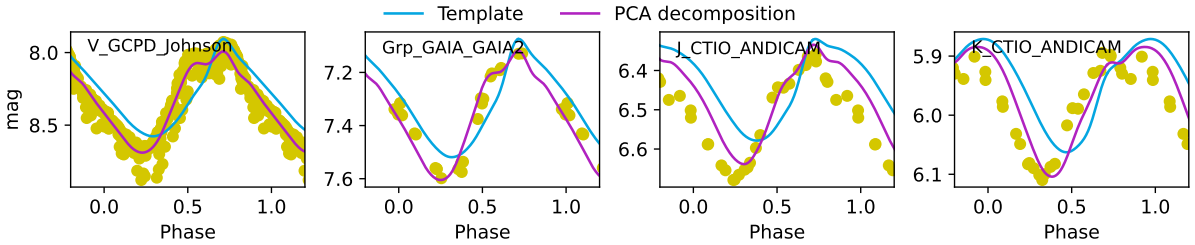


Figure 4.12: Observations and light curve fitting of SX Vel by two methods.

In addition, it could be a powerful road to estimate the systematic differences in the shape of the light curves in different environments. To follow up on my previous work, I would like to study the dependence of the light-curve shape on metallicity, thanks to my templates and the OGLE observations of Magellanic Clouds Cepheids. The OGLE project (Optical Gravitational Lensing Experiment), is a long-time survey focused on the search for microlensing events in the Galactic Centre and the Magellanic Clouds. The project started in 1992, and observed for many years variable sources with a 1-metre telescope (1992-1995) and a 1.3-metre telescope (1996-today) in Las Campanas (Chile). Thanks to its frequent observations of variable sources, it provided photometry for a large quantity of variable stars, including Cepheids and RR Lyrae in the galactic bulge and the Magellanic Clouds. In particular, the most recent OGLE IV catalogue (2015) (Soszyński et al. 2015; Udalski et al. 2015) published observed light curves in V and I bands for 9535 identified Cepheids in the Magellanic Clouds. Fitting these light curves with the PCA decomposition I already obtained would give information on the similarity of Magellanic Cloud, i.e. more metal-poor Cepheids with Galactic ones, through the distribution of K_i coefficients.

2.3.2 Introducing mean magnitudes variations

Another specificity of this approach is that the mean magnitudes in all the different bands are fully parametrised by only two parameters, the scaling magnitude and the colour excess. This is different from the approach by Yoachim et al. (2009) for instance who add one parameter per photometric band studied to cover the scaling of the template light curve on the observed magnitudes. This approach is interesting to minimise the number of free parameters and if some photometric bands have only a few observations. The drawback is that it requires to have a bit of photometry in the largest possible number of photometric bands, including IR. However, even if it is not what my templates were designed for, it is also possible to modify the fitting strategy accordingly and apply a similar approach. If for instance only two photometric bands are available, it would be more efficient to relax the constraint on mean magnitudes from the templates, and use one magnitude parameter per band. In this case, it does not change the

number of parameters as the parametrisation only changes from one global scaling magnitude and the colour excess to predict the difference between the two bands, to two independent magnitude parameters. Using the colour excess parametrisation with only two bands may lead to biased results for this parameter if the bands have too close effective wavelengths.

2.3.3 Phase-coverage-dependent fitting approach

As a conclusion, I wanted to highlight that the templates I built are primarily meant for deriving the mean parameters of extragalactic Cepheids with a few observations, in HST and JWST filters. Their application to other science cases are however possible, with an adaptation of the fitting strategy. In the meantime, as the templates are intrinsically generated for all photometric bands, they are usable to interpret the photometry from many datasets. I present in Fig. 4.13 a schematic summary in view of potential further applications.

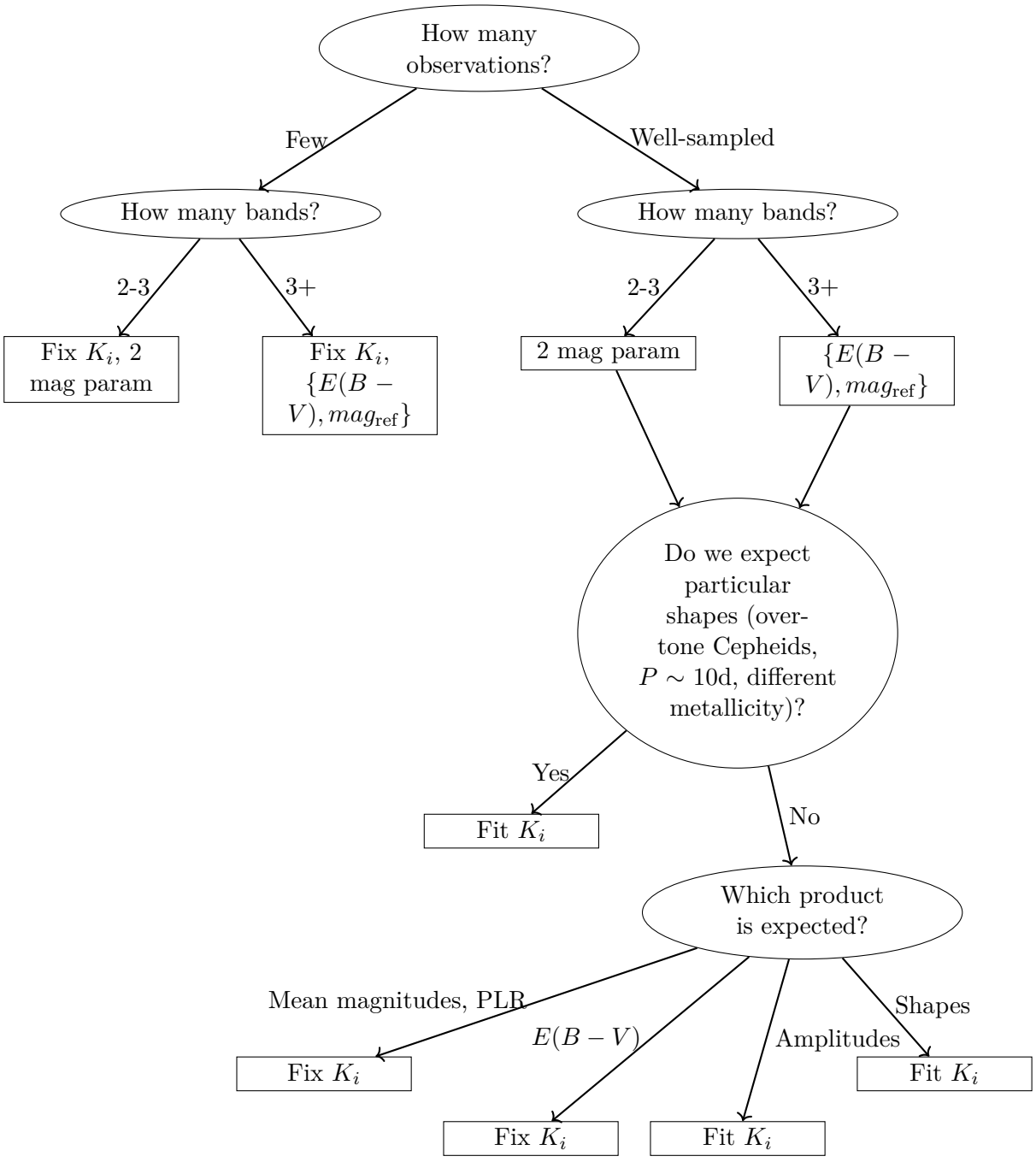


Figure 4.13: Schematic summary of the application of the multi-band light curve templates.

Conclusions and perspectives

This thesis is set within the framework of distance determinations using pulsating variable stars, in particular Cepheids and RR Lyrae stars. It is motivated by the need to improve the robustness of pulsating star modelling and the applicability of such modelling to distance determination methods.

In a first publication, I confirm the dispersion of the p-factor for pulsating stars and extend this result to RR Lyrae stars. In addition, I demonstrate the impact of combining data obtained using different observational methodologies, in particular with respect to RV measurements. This highlights the importance of carefully documenting the methods used when measuring radial velocities.

This work addresses the question of the origin of the p-factor dispersion and proposes the possibility of a desynchronisation of the pulsation motion between the stellar photosphere and the atmosphere. I also exclude the possibility of a simple dependence of the p-factor on the main physical properties of the star, in particular on the pulsation period. In this respect, RR Lyrae stars appear to follow the extension of the trend observed for Cepheids, exhibiting a dispersion similar to that of short-period Cepheids.

In a second publication, I propose an innovative methodology for modelling Cepheid light curves based on a limited number of observations. This approach relies on a statistical analysis of pulsation models derived from long-term observations of Galactic Cepheids. The proposed method provides the pulsation period and the mean magnitudes in all photometric bands, and by extension in combinations of photometric bands, which are key quantities for period–luminosity relations, through an optimisation involving a minimal number of free parameters.

This work is directly applicable to the observational data currently available for Cepheids in Type Ia supernova host galaxies, in particular from the HST and the JWST.

Overall, my thesis fits more broadly within the field through its applicability to future telescopes, surveys and data releases. New photometric surveys of Cepheids in Local Group galaxies will soon become available thanks to the Legacy Survey of Space and Time (LSST) at the Vera C. Rubin Observatory. In addition, the forthcoming launch of the Roman Space Telescope is expected to provide new observations. These data will enable the application of

photometric templates to derive increasingly precise distances to nearby galaxies. On the other hand, the next Gaia DR will make epoch spectroscopy publicly available for a large number of Cepheids and RR Lyrae stars, allowing a deeper investigation of the differences observed in RV measurements. This will also make it possible to apply the parallax-of-pulsation method on a large scale, and to achieve a better characterisation of the p-factor. The Extremely Large Telescope (ELT) will begin its operations by 2030. The multi-object spectrograph MOSAIC (Multi Object Spectrograph with Adaptive Image Correction) will then enable the measurement of RV curves of Cepheids in the Local Group, allowing an even broader application of the parallax-of-pulsation method. This further underlines the importance of the work carried out to understand the origin of the p-factor dispersion.

These studies aim to improve the calibration of the distance scale. This is achieved first through the introduction of the geometrical parallax-of-pulsation method, provided that an accurate calibration of the p-factor is available, thereby increasing the number of anchor galaxies with precisely determined distances. In addition, improved determinations of pulsation periods and mean magnitudes are obtained through the use of photometric templates.

I have also identified a number of limitations. Firstly, the modelling of pulsating stars remains constrained by the requirement of an appropriate phase coverage of their variability. Modelling with the SPIPS tool is particularly demanding in terms of observational data, as it requires at minimum a well-sampled RV curve and multi-band photometry in both the optical and IR domains. In addition, my analysis relies entirely on the use of stellar atmosphere models to convert photometric measurements between different bands. Although these models are strongly constrained by various observational data and empirical relations, this assumption still has limitations, particularly in the modelling of short-wavelength photometry. The treatment of interstellar reddening also introduces additional modelling assumptions, notably through the adopted reddening laws. Due to observational limitations, the studied samples - whether for the p-factor of RR Lyrae stars (17 objects) or for the training sample used to construct the photometric templates (75 stars) - remain relatively small. The impact of this selection is still difficult to fully assess.

Finally, regarding the calibration of the p-factor, several assumptions are made that could be challenged. The parallax-of-pulsation method assumes a phase-independent p-factor, neglecting variations in LD throughout the pulsation cycle. At present, these variations are constrained only theoretically, through quasi-static and hydrodynamical models. The application of photometric templates, on the other hand, assumes the universality of Cepheids and a negligible effect of metallicity, in particular between the light-curve shapes of Galactic Cepheids and those in Type Ia supernova host galaxies.

New observational perspectives naturally emerge from the questions raised in this work. The development of high-resolution spectroscopy opens the way to more detailed analyses of pulsating stellar atmospheres, in particular through tomographic studies. The next generation of interferometers (Asgard at the VLTI, SPICA at CHARA) will soon make it possible to

observationally constrain the geometrical component of the projection factor. Large photometric surveys will also play an important role in multi-wavelength analyses of light curves.

Improvements in pulsation models and in the convergence of the SPIPS algorithm, the extension of photometric templates to other classes of pulsating stars or to different environments, and the incorporation of LD variations into the p-factor represent methodological perspectives that could naturally extend the work presented in this thesis.

Finally, the large-scale application of photometric templates in the context of Hubble constant measurements, and thus an indirect contribution to precision cosmology, should be pursued in order to fully exploit the originality and potential of the work presented here.

List of publications

Refereed publications

As first author

Bras G., Kervella P., Mérand A. et al. "*Versatile light curve templates of Cepheids: Multi-band fitting of photometric time series*". A&A 698, A97.
DOI: 10.1051/0004-6361/202553783

Bras G., Kervella P., Trahin B. et al. "*The Baade-Wesselink projection factor of RR Lyrae stars. Calibration from OHP/SOPHIE spectroscopy and Gaia DR3 parallaxes*". A&A 684, A126.
DOI: 10.1051/0004-6361/202348291

As co-author

Bailleul M., Nardetto N., ..., **Bras G.** et al. "*Surface brightness–colour relations of Cepheids calibrated by optical interferometry*". A&A 698, A46 (2025).
DOI: 10.1051/0004-6361/202553740

Hocdé V., Matter A., ..., **Bras G.** et al. "*Circumstellar emission of Cepheids across the instability strip: Mid-infrared observations with VLTI/MATISSE*". A&A 694, A101 (2025).
DOI: 10.1051/0004-6361/202452827

Zgirski B., Gieren W., ..., **Bras G.** et al. "*Infrared surface brightness technique applied to RR Lyrae stars from the solar neighborhood*". A&A 690, A295 (2024).
DOI: 10.1051/0004-6361/202449850

Wielgórski P., Pietrzyński G., ..., **Bras G.** et al. "*Projection factor and radii of Type II Cepheids: BL Her stars*". A&A 689, A241 (2024).
DOI: 10.1051/0004-6361/202450182

Evans N., Schaefer G., ..., **Bras G.** et al. "*The Orbit and Dynamical Mass of Polaris: Observations with the CHARA Array*". ApJ 971, 190 (2024).
DOI: 10.3847/1538-4357/ad5e7a

Nardetto N., Hocdé V., ..., **Bras G.** et al. "*The orbital parameters of the δ Cep inner binary system determined using 2019 HARPS-N spectroscopic data*". A&A 684, L9 (2024).
DOI: 10.1051/0004-6361/202449600

Zgirski B., Gieren W., ..., **Bras G.** et al. "*Distance to the globular cluster M3 from the infrared surface brightness technique applied to RR Lyrae stars*". accepted in A&A.

Submitted manuscripts

Bailleul M, Nardetto N, ..., **Bras G.** et al. "*Surface brightness colour relations of Milky Way and Magellanic clouds Classical Cepheids based on Gaia magnitudes*". submitted to A&A.

Gallenne A., Kervella P., ..., **Bras G.** et al. "*High-contrast imaging of Galactic Cepheids with VLT/SPHERE*". submitted to A&A.

Manuscripts in preparation

Bras G., Kervella P., Gallenne A. et al. "*Centre-to-limb darkening in the classical Cepheid η Aql*". in prep.

International conferences

IAU Symposium 376 Budapest (April 2023): Poster - "*Observational calibration of the projection factor of RR Lyrae stars using the SPIPS pulsation modelling*".

MIAPbP Program "Extragalactic Distance Scale" (July 2023): Discussion session - "*The parallax-of-pulsation method*".

Araucaria collaboration workshop (September 2024): Oral presentation - "*Light-curve templates of Cepheids*".

RRL/Cep Marrakech (November 2024): Oral presentation - "*Observational calibration of the p-factor of RR Lyrae stars using the SPIPS modelling tool*".

RRL/Cep Marrakech (November 2024): Poster - "*Versatile templates of Cepheid light curves*".

Other presentations

Meeting of the High Angular Resolution group at LESIA (January 2023): Oral presentation - "*Observational calibration of the projection factor of RR Lyrae stars for distance determinations*".

Elbereth conference of the PhD students in astrophysics in Paris (March 2023): Oral presentation - "*Observational calibration of the projection factor of pulsating stars to determine distances*".

PhD day of the LESIA laboratory (June 2023): Oral presentation - "*Observational calibration of the projection factor of pulsating stars to determine distances*".

Action Initiative Physique Stellaire seminar (December 2023): Oral presentation - "*Observational calibration of the projection factor of RR Lyrae stars using OHP/SOPHIE radial velocity measurements and Gaia DR3 parallaxes*".

Elbereth conference of the PhD students in astrophysics in Paris (February 2024): Oral presentation - "*The calibration of the projection factor of pulsating stars: an unanswered question?*".

QuercY Physics Junior Workshop on Quantum Systems, Elementary Fields, Radiation, Cosmos and Interactions (April 2024): Oral presentation - "*A journey from time to space: unravelling distances with pulsating stars*".

PhD day of the LESIA laboratory (May 2024): Oral presentation - "*Predictive modelling of Cepheid light curves for enhanced distance determination*".

Meeting of the High Angular Resolution group at LESIA (January 2026): Oral presentation - "*Monitoring the limb-darkening variations in Cepheids*".

Acronyms

ΛCDM Lambda Cold Dark Matter

(N)IR (Near) InfraRed

AGB Asymptotic Giant Branch

CCF Cross-Correlation Function

CCHP Chicago-Carnegie Hubble Program

CHARA Center for High Angular Resolution Astronomy

CMB Cosmic Microwave Background

CMD Colour-Magnitude Diagram

CSE CircumStellar Envelope

DR Data Release

ELT Extremely Large Telescope

ESA European Space Agency

ESO European Southern Observatory

H Hydrogen

HARPS High Accuracy Radial velocity Planet Searcher

He Helium

HR Hertzsprung-Russell

HST Hubble Space Telescope

IS Instability Strip

JAGB J-region Asymptotic Giant Branch

JMMC Jean-Marie Mariotti Centre

JWST James Webb Space Telescope

LD Limb-Darkening

LSST Legacy Survey of Space and Time

MARCS Model Atmospheres with a Radiative and Convective Scheme

MIRC-X Michigan InfraRed Combiner-eXeter

MOSAIC Multi Object Spectrograph with Adaptive Image Correction

MYSTIC Michigan Young STar Imager at CHARA

O-C Observations-Computed

OGLE Optical Gravitational Lensing Experiment

PC(R) Period-Colour (Relation)

PCA Principal Component Analysis

PL(R) Period-Luminosity (Relation)

PR(R) Period-Radius (Relation)

PSF Point-Spread Function

PW(R) Period-Wesenheit (Relation)

RGB Red Giant Branch

RV Radial Velocity

RVS Radial Velocity Spectrometer

S/N Signal-to-Noise Ratio

SBCR Surface-Brightness-Colour Relation

SED Spectral Energy Distribution

SH0ES Supernova H_0 for the Equation of State

SNIa Type Ia Supernova

SOPHIE Spectrographe pour l'Observation des PHénomènes des Intérieurs stellaires et des Exoplanètes

SPICA Stellar Parameters and Images with a Cophased Array

SPIPS Spectro-Photo-Interferometry of Pulsating Stars

TRGB Tip of the Red Giant Branch

UD Uniform Disk

UV Ultra-Violet

VLTI Very Large Telescope Interferometer

WISE Wide-field Infrared Survey Explorer

Bibliography

Andrievsky, S. et al. (2018). "The Relationship of Sodium and Oxygen in Galactic Field RR Lyrae Stars". PASP 130.984, p. 024201.

Andrievsky, Sergei M. et al. (2020). "NLTE C and O abundances in RR lyrae stars". *Astronomische Nachrichten* 341.9, pp. 899–907.

Andrievsky, Sergei M. et al. (2021). "NLTE CNO abundances in a sample of nine field RR Lyr type stars". *Astronomische Nachrichten* 342.6, pp. 887–897.

Anugu, Narsireddy et al. (2020). "MIRC-X: A Highly Sensitive Six-telescope Interferometric Imager at the CHARA Array". AJ 160.4, 158, p. 158.

Armstrong, J. T. et al. (2001). "Diameters of δ Cephei and η Aquilae Measured with the Navy Prototype Optical Interferometer". AJ 121.1, pp. 476–481.

Baade, W. (1926). "Über eine Möglichkeit, die Pulsationstheorie der δ Cephei-Veränderlichen zu prüfen". *Astronomische Nachrichten* 228.20, p. 359.

Bailer-Jones, C. A. L. et al. (2021). "Estimating Distances from Parallaxes. V. Geometric and Photogeometric Distances to 1.47 Billion Stars in Gaia Early Data Release 3". AJ 161.3, 147, p. 147.

Bailleul, M. C. et al. (2025). "Surface brightness–colour relations of Cepheids calibrated by optical interferometry". A&A 698, A46, A46.

Barnes Thomas G., III et al. (1988). "Contemporaneous Photometry and Radial Velocities of RR Lyrae Stars". ApJS 67, p. 403.

Barnes Thomas G., III et al. (1992). "BVRIJHK Photometry of RR Lyrae Variables". PASP 104, p. 514.

Berdnikov, L. N. (2008). "Photoelectric observations of Cepheids in UBV(RI)c". *VizieR Online Data Catalog: II/285*.

Bhardwaj, Anupam et al. (2015). "On the variation of Fourier parameters for Galactic and LMC Cepheids at optical, near-infrared and mid-infrared wavelengths". MNRAS 447.4, pp. 3342–3360.

Bhardwaj, Anupam et al. (2016). "Large Magellanic Cloud Near-Infrared Synoptic Survey - III. A statistical study of non-linearity in the Leavitt Laws". MNRAS 457.2, pp. 1644–1665.

Bhardwaj, Anupam et al. (2023). "High-resolution Spectroscopic Metallicities of Milky Way Cepheid Standards and Their Impact on the Leavitt Law and the Hubble Constant". ApJ 955.1, L13, p. L13.

Blanco-Cuaresma, S. et al. (2014). "Determining stellar atmospheric parameters and chemical abundances of FGK stars with iSpec". A&A 569, A111, A111.

Blanco-Cuaresma, Sergi (2019). "Modern stellar spectroscopy caveats". MNRAS 486.2, pp. 2075–2101.

Blažko, S. (1907). "Mitteilung über veränderliche Sterne". *Astronomische Nachrichten* 175, p. 325.

Bono, G. et al. (2001). “Improving the Mass Determination of Galactic Cepheids”. ApJ 563.1, pp. 319–324.

Borgniet, S. et al. (2019). “Consistent radial velocities of classical Cepheids from the cross-correlation technique”. A&A 631, A37, A37.

Boubel, Paula et al. (2024). “Large-scale motions and growth rate from forward-modelling Tully-Fisher peculiar velocities”. MNRAS 531.1, pp. 84–109.

Bouchy, F. et al. (2009). “The SOPHIE search for northern extrasolar planets . I. A companion around HD 16760 with mass close to the planet/brown-dwarf transition”. A&A 505.2, pp. 853–858.

Bras, Garance et al. (2024). “The Baade-Wesselink projection factor of RR Lyrae stars. Calibration from OHP/SOPHIE spectroscopy and Gaia DR3 parallaxes”. A&A 684, A126, A126.

Bras, Garance et al. (2025). “Versatile light curve templates of Cepheids: Multi-band fitting of photometric time series”. A&A 698, A97, A97.

Breitfelder, J. et al. (2016). “Observational calibration of the projection factor of Cepheids. II. Application to nine Cepheids with HST/FGS parallax measurements”. A&A 587, A117, A117.

Breuval, Louise et al. (2021). “The Influence of Metallicity on the Leavitt Law from Geometrical Distances of Milky Way and Magellanic Cloud Cepheids”. ApJ 913.1, 38, p. 38.

Breuval, Louise et al. (2024). “Small Magellanic Cloud Cepheids Observed with the Hubble Space Telescope Provide a New Anchor for the SH0ES Distance Ladder”. ApJ 973.1, 30, p. 30.

Breuval, Louise et al. (2025). “The Legacy of Henrietta Leavitt: A Re-analysis of the First Cepheid Period–Luminosity Relation”. PASP 137.4, 044001, p. 044001.

Brout, Dillon et al. (2022). “The Pantheon+ Analysis: Cosmological Constraints”. ApJ 938.2, 110, p. 110.

Burki, G. et al. (1982). “The peculiar classical cepheid HD 7308.” A&A 109, pp. 258–270.

Cacciari, C. et al. (1987). “The Baade-Wesselink method applied to field RR Lyrae stars. I. UBVRI photoelectric and radial velocity data.” A&AS 69, pp. 135–156.

Capitanio, L. et al. (2017). “Three-dimensional mapping of the local interstellar medium with composite data”. A&A 606, A65, A65.

Carney, B. W. et al. (1984). “The Baade-Wesselink method and the distances to RR Lyrae stars. I. The field star VY Serpentis.” ApJ 278, pp. 241–254.

Casagrande, L. et al. (2014). “Synthetic stellar photometry - I. General considerations and new transformations for broad-band systems”. MNRAS 444.1, pp. 392–419.

Castelli, F. et al. (2003). “New Grids of ATLAS9 Model Atmospheres”. *Modelling of Stellar Atmospheres*. Ed. by N. Piskunov et al. Vol. 210, A20.

Chelli, Alain et al. (2016). “Pseudomagnitudes and differential surface brightness: Application to the apparent diameter of stars”. A&A 589, A112, A112.

Claret, A. et al. (2011). “Gravity and limb-darkening coefficients for the Kepler, CoRoT, Spitzer, uvby, UBVRIJHK, and Sloan photometric systems”. A&A 529, A75, A75.

Clementini, G. (2010). “Pulsating variable stars, powerful tools for galactic structure and evolution”. *EAS Publications Series*. Ed. by C. Turon et al. Vol. 45. EAS Publications Series. EDP, pp. 267–272.

Clementini, G. et al. (2023). “Gaia Data Release 3. Specific processing and validation of all-sky RR Lyrae and Cepheid stars: The RR Lyrae sample”. A&A 674, A18, A18.

Costa, G. et al. (2025). “Evolutionary tracks, ejecta, and ionizing photons from intermediate-mass to very massive stars with PARSEC”. A&A 694, A193, A193.

Csörnyei, G. et al. (2023). “Consistency of Type IIP supernova sibling distances”. A&A 672, A129, A129.

Davis, John et al. (2006). "SUSI: an update on instrumental developments and science". *Advances in Stellar Interferometry*. Ed. by John D. Monnier et al. Vol. 6268. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 626804, p. 626804.

de Jaeger, T. et al. (2020). "Studying Type II supernovae as cosmological standard candles using the Dark Energy Survey". MNRAS 495.4, pp. 4860–4892.

De Somma, Giulia et al. (2024). "Stellar Pulsation and Evolution: A Combined Theoretical Renewal and Updated Models (SPECTRUM). I. Updating Radiative Opacities for Pulsation Models of Classical Cepheid and RR-Lyrae". ApJ 977.1, 1, p. 1.

Eddington, A. S. (1917). "The pulsation theory of Cepheid variables". *The Observatory* 40, pp. 290–293.

Evans, Nancy Remage et al. (2024). "The Orbit and Dynamical Mass of Polaris: Observations with the CHARA Array". ApJ 971.2, 190, p. 190.

Feast, Michael W. et al. (2008). "The luminosities and distance scales of type II Cepheid and RR Lyrae variables". MNRAS 386.4, pp. 2115–2134.

Fernie, J. D. et al. (1995). "A Database of Galactic Classical Cepheids". *Information Bulletin on Variable Stars* 4148, p. 1.

Fernley, J. A. et al. (1989). "The absolute magnitudes of RR Lyrae stars - I. X Arietis." MNRAS 236, pp. 447–473.

Fernley, J. A. et al. (1990). "The absolute magnitudes of RR Lyrae stars - IV. V 445 , SS Leo and VY Serpentis." MNRAS 247, p. 287.

Fernley, J. A. et al. (1993). "Infrared photometry and radial velocities of field RR Lyraes." A&AS 97, pp. 815–826.

Fitzpatrick, Edward L. (1999). "Correcting for the Effects of Interstellar Extinction". PASP 111.755, pp. 63–75.

Fouqué, P. et al. (1997). "An improved calibration of Cepheid visual and infrared surface brightness relations from accurate angular diameter measurements of cool giants and supergiants." A&A 320, pp. 799–810.

Freedman, Wendy L. et al. (2019). "The Carnegie-Chicago Hubble Program. VIII. An Independent Determination of the Hubble Constant Based on the Tip of the Red Giant Branch". ApJ 882.1, 34, p. 34.

Freedman, Wendy L. et al. (2020). "Calibration of the Tip of the Red Giant Branch". ApJ 891.1, 57, p. 57.

Freedman, Wendy L. et al. (2025). "Status Report on the Chicago-Carnegie Hubble Program (CCHP): Measurement of the Hubble Constant Using the Hubble and James Webb Space Telescopes". ApJ 985.2, 203, p. 203.

Gaia Collaboration et al. (2023). "Gaia Data Release 3. Summary of the content and survey properties". A&A 674, A1, A1.

Gallenne, A. et al. (2017). "Observational calibration of the projection factor of Cepheids. IV. Period-projection factor relation of Galactic and Magellanic Cloud Cepheids". A&A 608, A18, A18.

Gallenne, A. et al. (2021). "Extended envelopes around Galactic Cepheids. V. Multi-wavelength and time-dependent analysis of IR excess". A&A 651, A113, A113.

Gautschy, Alfred et al. (1995). "Stellar Pulsations Across The HR Diagram: Part 1". ARA&A 33, pp. 75–114.

Giertych, Naomi et al. (2024). "A statistical primer on classical period-finding techniques in astronomy". *Reports on Progress in Physics* 87.7, 078401, p. 078401.

Gordon, Karl D. et al. (2023). "One Relation for All Wavelengths: The Far-ultraviolet to Mid-infrared Milky Way Spectroscopic R(V)-dependent Dust Extinction Relationship". ApJ 950.2, 86, p. 86.

Graczyk, D. et al. (2021). "The surface brightness-colour relations based on eclipsing binary stars and calibrated with Gaia EDR3". A&A 649, A109, A109.

Graczyk, Dariusz et al. (2020). "A Distance Determination to the Small Magellanic Cloud with an Accuracy of Better than Two Percent Based on Late-type Eclipsing Binary Stars". ApJ 904.1, 13, p. 13.

GRAVITY Collaboration et al. (2017). "First light for GRAVITY: Phase referencing optical interferometry for the Very Large Telescope Interferometer". A&A 602, A94, A94.

H0DN Collaboration et al. (2025). *The Local Distance Network: a community consensus report on the measurement of the Hubble constant at 1% precision.*

Hertzsprung, E. (1926). "On the relation between period and form of the light-curve of variable stars of the δ Cephei type". Bull. Astron. Inst. Netherlands 3, p. 115.

Hocdé, V. et al. (2023). "Metallicity estimations of MW, SMC, and LMC classical Cepheids from the shape of the V- and I-band light curves". A&A 671, A157, A157.

Hoffmann, Samantha L. et al. (2016). "Optical Identification of Cepheids in 19 Host Galaxies of Type Ia Supernovae and NGC 4258 with the Hubble Space Telescope". ApJ 830.1, 10, p. 10.

Husser, T.-O. et al. (2013). "Astrophysics A new extensive library of PHOENIX stellar atmospheres". *A&A* 553, A6.

Javanmardi, Behnam et al. (2021). "Inspecting the Cepheid Distance Ladder: the Hubble Space Telescope Distance to the SN Ia Host Galaxy NGC 5584". ApJ 911.1, 12, p. 12.

Jensen, Joseph B. et al. (2025). "The TRGB−SBF Project. III. Refining the HST Surface Brightness Fluctuation Distance Scale Calibration with JWST". ApJ 987.1, 87, p. 87.

Jones, Rodney V. et al. (1987a). "The Baade-Wesselink Method and the Distances to RR Lyrae Stars. II. The Field Star X ARIETIS". ApJ 312, p. 254.

— (1987b). "The Baade-Wesselink Method and the Distances to RR Lyrae Stars. III. The Field Star SW Draconis". ApJ 314, p. 605.

Jones, Rodney V. et al. (1988). "The Baade-Wesselink Method and the Distances to RR Lyrae Stars. VI. The Field Stars RS Bootis, TW Herculis, VY Serpentis, and UU Virginis, and the Absolute Magnitudes of RR Lyrae Stars". ApJ 332, p. 206.

Kervella, P. et al. (2004a). "Cepheid distances from infrared long-baseline interferometry. I. VINCI/VLTI observations of seven Galactic Cepheids". A&A 416, pp. 941–953.

Kervella, P. et al. (2004b). "Cepheid distances from infrared long-baseline interferometry. III. Calibration of the surface brightness-color relations". A&A 428, pp. 587–593.

Kervella, Pierre et al. (2017). "Observational calibration of the projection factor of Cepheids. III. The long-period Galactic Cepheid RS Puppis". A&A 600, A127, A127.

Kovtyukh, V. V. et al. (2000). "Precise temperatures of classical Cepheids and yellow supergiants from line-depth ratios". A&A 358, pp. 587–592.

Kraft, Robert P. (1956). "Double Lines in the Spectrum of the Classical Cepheid X Cygni". PASP 68.401, p. 137.

Kravchenko, K. et al. (2018). "Tomography of cool giant and supergiant star atmospheres. I. Validation of the method". A&A 610, A29, A29.

Lane, B. F. et al. (2000). "Direct detection of pulsations of the Cepheid star ζ Gem and an independent calibration of the period-luminosity relation". Nature 407.6803, pp. 485–487.

Layden, Andrew C. et al. (2019). "Infrared K-band Photometry of Field RR Lyrae Variable Stars". AJ 158.3, 105, p. 105.

Leavitt, Henrietta S. (1907). "1777 variables in the Magellanic Clouds". *Annals of Harvard College Observatory* 60, pp. 87–108.3.

Leavitt, Henrietta S. et al. (1912). "Periods of 25 Variable Stars in the Small Magellanic Cloud." *Harvard College Observatory Circular* 173, pp. 1–3.

Lee, Abigail J. et al. (2025). "The Chicago–Carnegie Hubble Program: The JWST J-region Asymptotic Giant Branch Extragalactic Distance Scale". ApJ 985.2, 182, p. 182.

Lee, Myung Gyoon et al. (1993). "The Tip of the Red Giant Branch as a Distance Indicator for Resolved Galaxies". ApJ 417, p. 553.

Lemasle, Bertrand et al. (2020). "Atmospheric parameters of Cepheids from flux ratios with ATHOS. I. The temperature scale". A&A 641, A71, A71.

Lindegren, L. et al. (2021). "Gaia Early Data Release 3. Parallax bias versus magnitude, colour, and position". A&A 649, A4, A4.

Lindemann, F. A. (1918). "Note on the pulsation theory of Cepheid variables". MNRAS 78, p. 639.

Liu, Tianxing et al. (1989). "The Luminosity Scale of RR Lyrae Stars with the Baade-Wesselink Method. I. Photometry and Radial Velocities". ApJS 69, p. 593.

Madore, B. F. (1982). "The period-luminosity relation. IV. Intrinsic relations and reddenings for the Large Magellanic Cloud Cepheids." ApJ 253, pp. 575–579.

Madore, Barry F. et al. (1991). "The Cepheid Distance Scale". PASP 103, p. 933.

— (2020). "Astrophysical Distance Scale: The AGB J-band Method. I. Calibration and a First Application". ApJ 899.1, 66, p. 66.

Marconi, M. et al. (2015). "On a New Theoretical Framework for RR Lyrae Stars. I. The Metallicity Dependence". ApJ 808.1, 50, p. 50.

Marengo, Massimo et al. (2003). "Theoretical Limb Darkening for Classical Cepheids. II. Corrections for the Geometric Baade-Wesselink Method". ApJ 589.2, pp. 968–975.

Marsakov, V. A. et al. (2019). "Masses of RR Lyrae Stars with Different Chemical Abundances in the Galactic Field". *Astronomy Reports* 63.3, pp. 203–211.

Mérand, A. et al. (2005). "The projection factor of δ Cephei. A calibration of the Baade-Wesselink method using the CHARA Array". A&A 438.1, pp. L9–L12.

Mérand, A. et al. (2015). "Cepheid distances from the SpectroPhoto-Interferometry of Pulsating Stars (SPIPS). Application to the prototypes δ Cephei and η Aquilae". A&A 584, A80, A80.

Meylan, G. et al. (1986). "RR Lyrae, delta Scuti, SX Phoenicis stars and Baade-Wesselink method.I. Photometric and radial velocity measurements of four field stars :RR Cet, DX Del, BS AQR and DY Peg." A&AS 64, pp. 25–51.

Mihalas, D. (2003). "Cepheid Atmospheres". *Stellar Atmosphere Modeling*. Ed. by Ivan Hubeny et al. Vol. 288. Astronomical Society of the Pacific Conference Series, p. 471.

Monnier, John D. et al. (2004). "The Michigan Infrared Combiner (MIRC): IR imaging with the CHARA Array". *New Frontiers in Stellar Interferometry*. Ed. by Wesley A. Traub. Vol. 5491. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 1370.

Monnier, John D. et al. (2018). "MYSTIC: Michigan Young STar Imager at CHARA". *Optical and Infrared Interferometry and Imaging VI*. Ed. by Michelle J. Creech-Eakman et al. Vol. 10701. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1070122, p. 1070122.

Monson, Andrew J. et al. (2017). "Standard Galactic Field RR Lyrae. I. Optical to Mid-infrared Phased Photometry". AJ 153.3, 96, p. 96.

Mourard, D. et al. (1997). "The mean angular diameter of δ Cephei measured by optical long-baseline interferometry." A&A 317, pp. 789–792.

Nardetto, N. et al. (2004). "Self consistent modelling of the projection factor for interferometric distance determination". A&A 428, pp. 131–137.

Nardetto, N. et al. (2006). "High resolution spectroscopy for Cepheids distance determination. I. Line asymmetry". A&A 453.1, pp. 309–319.

Nardetto, N. et al. (2007). "High-resolution spectroscopy for Cepheids distance determination. II. A period-projection factor relation". A&A 471.2, pp. 661–669.

Nardetto, N. et al. (2014). "Understanding the dynamical structure of pulsating stars. HARPS spectroscopy of the δ Scuti stars ρ Puppis and DX Ceti". A&A 561, A151, A151.

Nardetto, N. et al. (2023). "HARPS-N high spectral resolution observations of Cepheids. II. The impact of the surface-brightness color relation on the Baade-Wesselink projection factor of η Aql". A&A 671, A14, A14.

Neilson, H. R. et al. (2013). "Spherically-symmetric model stellar atmospheres and limb darkening. I. Limb-darkening laws, gravity-darkening coefficients and angular diameter corrections for red giant stars". A&A 554, A98, A98.

Ngeow, C. et al. (2006). "Nonlinear Period-Luminosity Relation for the Large Magellanic Cloud Cepheids: Myths and Truths". ApJ 650.1, pp. 180–188.

Paczyński, B. (1965). "Three-Colour Photometry of RR Lyrae Stars. I." ActaA 15, p. 103.

— (1966). "Three-Colour Photometry of RR Lyrae Stars. III." ActaA 16, p. 97.

Parviainen, Hannu et al. (2015). "ldtk: Limb Darkening Toolkit". *MNRAS* 453.4, pp. 3821–3826.

Pedregosa, F. et al. (2011). "Scikit-learn: Machine Learning in Python". *Journal of Machine Learning Research* 12, pp. 2825–2830.

Perruchot, S. et al. (2008). "The SOPHIE spectrograph: design and technical key-points for high throughput and high stability". *Ground-based and Airborne Instrumentation for Astronomy II*. Ed. by Ian S. McLean et al. Vol. 7014. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 70140J, 70140J.

Pietrzyński, G. et al. (2019). "A distance to the Large Magellanic Cloud that is precise to one per cent". Nature 567.7747, pp. 200–203.

Pilecki, Bogumił et al. (2017). "Mass and p-factor of the Type II Cepheid OGLE-LMC-T2CEP-098 in a Binary System". ApJ 842.2, 110, p. 110.

Planck Collaboration et al. (2020). "Planck 2018 results. VI. Cosmological parameters". A&A 641, A6, A6.

Preston, George W. et al. (1964). "Atmospheric Phenomena in the RR Lyrae Stars. I. The Singly Periodic Variables." ApJ 140, p. 181.

Reid, M. J. et al. (2019). "An Improved Distance to NGC 4258 and Its Implications for the Hubble Constant". ApJ 886.2, L27, p. L27.

Riess, Adam G. et al. (1998). "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant". AJ 116.3, pp. 1009–1038.

Riess, Adam G. et al. (2021). "Cosmic Distances Calibrated to 1% Precision with Gaia EDR3 Parallaxes and Hubble Space Telescope Photometry of 75 Milky Way Cepheids Confirm Tension with ΛCDM". ApJ 908.1, L6, p. L6.

Riess, Adam G. et al. (2022). "A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km s^{-1} Mpc^{-1} Uncertainty from the Hubble Space Telescope and the SH0ES Team". ApJ 934.1, L7, p. L7.

Riess, Adam G. et al. (2023). "Crowded No More: The Accuracy of the Hubble Constant Tested with High-resolution Observations of Cepheids by JWST". ApJ 956.1, L18, p. L18.

Rucinski, S. (1999). "Determination of Broadening Functions Using the Singular-Value Decomposition (SVD) Technique". *IAU Colloquium 170: Precise Stellar Radial Velocities*. Ed. by J. B. Hearnshaw et al. Vol. 185. Astronomical Society of the Pacific Conference Series, p. 82.

Said, Khaled et al. (2025). "DESI peculiar velocity survey - Fundamental Plane". MNRAS 539.4, pp. 3627–3644.

Salsi, A. et al. (2021). "Progress on the calibration of surface brightness-color relations for early- and late-type stars". A&A 652, A26, A26.

Santerne, A. et al. (2016). "SOPHIE velocimetry of Kepler transit candidates. XVII. The physical properties of giant exoplanets within 400 days of period". A&A 587, A64, A64.

Sasselov, Dimitar D. et al. (1989). “Infrared Spectroscopy of Cepheids: Peculiar Velocity Structure and Its Effect on Radii and Distances”. ApJ 337, p. L29.

Setterholm, Benjamin R. et al. (2023). “MYSTIC: a high angular resolution K-band imager at CHARA”. *Journal of Astronomical Telescopes, Instruments, and Systems* 9, 025006, p. 025006.

Skillen, I. et al. (1989). “The absolute magnitudes of RR Lyrae stars - II. DX Delphini.” MNRAS 241, pp. 281–294.

Soszyński, I. et al. (2015). “The OGLE Collection of Variable Stars. Classical Cepheids in the Magellanic System”. ActaA 65.4, pp. 297–312.

Storm, J. et al. (2011). “Calibrating the Cepheid period-luminosity relation from the infrared surface brightness technique. I. The p-factor, the Milky Way relations, and a universal K-band relation”. A&A 534, A94, A94.

Sturch, C. (1966). “Intrinsic UBV colors of RR LYR stars.” ApJ 143, p. 774.

ten Brummelaar, T. A. et al. (2005). “First Results from the CHARA Array. II. A Description of the Instrument”. ApJ 628.1, pp. 453–465.

Trahin, B. et al. (2021). “Inspecting the Cepheid parallax of pulsation using Gaia EDR3 parallaxes. Projection factor and period-luminosity and period-radius relations”. A&A 656, A102, A102.

Udalski, A. et al. (2015). “OGLE-IV: Fourth Phase of the Optical Gravitational Lensing Experiment”. ActaA 65.1, pp. 1–38.

van Hoof, A. et al. (1953). “Differential Radial Velocities in Beta Canis Majoris”. PASP 65.384, p. 158.

Varsavsky, Carlos M. (1960). “Radial Velocities and Theree-Color Light-Curves of Some RR Lyrae Variables.” ApJ 131, p. 623.

Wesselink, A. J. (1946). “Discussion of radial velocities and photometric observations of δ Cephei”. Bull. Astron. Inst. Netherlands 10, p. 88.

Wielgórski, P. et al. (2024). “Projection factor and radii of Type II Cepheids: BL Her stars”. A&A 689, A241, A241.

Yoachim, Peter et al. (2009). “A Panoply of Cepheid Light Curve Templates”. AJ 137.6, pp. 4697–4706.

Zgirski, Bartłomiej et al. (2021). “The Araucaria Project. Distances to Nine Galaxies Based on a Statistical Analysis of their Carbon Stars (JAGB Method)”. ApJ 916.1, 19, p. 19.

Zgirski, Bartłomiej et al. (2024). “Infrared surface brightness technique applied to RR Lyrae stars from the solar neighborhood”. A&A 690, A295, A295.

RÉSUMÉ

Les étoiles variables pulsantes, notamment les Céphéides et RR Lyrae, sont des outils puissants pour la détermination des distances. Les relations Période-Luminosité permettent de déduire leur luminosité intrinsèque, ensuite comparée à leur magnitude apparente. Par ailleurs, la méthode de la parallaxe de pulsation utilise la relation trigonométrique entre la variation du rayon et du diamètre angulaire d'une étoile pour en déterminer la distance. Cette dernière fait intervenir un paramètre, le facteur de projection, qui convertit les vitesses radiales observées à la vitesse de la photosphère. Ce paramètre, entièrement dégénéré avec la distance, montre d'importantes variations statistiques, nous empêchant d'utiliser la parallaxe de pulsation à grande échelle. Une compréhension plus fine de la pulsation est requise afin de déterminer l'origine de cette dispersion.

Dans cette thèse j'étudie tout d'abord le facteur de projection dans le cas des RR Lyrae, à l'aide du code SPIPS et de nouvelles mesures de vitesses radiales, afin de le comparer aux Céphéides. Je montre que la valeur et la dispersion du facteur de projection des RR Lyrae sont assez similaires à ce qui a été observé pour les Céphéides et que les différentes méthodes de mesure des vitesses radiales ont peu d'impact sur cette dispersion à condition de considérer des données homogènes. Je présente aussi mes travaux relatifs à la variation du coefficient d'assombrissement centre-bord lors de la pulsation, un ingrédient clé du facteur de projection, grâce à des observations interférométriques de la Céphéide η Aql.

Dans un second temps, je développe des courbes de lumière génériques pour les Céphéides, à l'aide des modèles multi-bandes SPIPS. Ces courbes génériques permettent pour une période de pulsation donnée de prédire la forme de la courbe de lumière dans une large sélection de bandes photométriques. Elles sont particulièrement efficaces pour modéliser des courbes de lumière sur des observations rares ou peu précises, comme c'est le cas pour les Céphéides extragalactiques.

MOTS CLÉS

Observations, Analyse de données, Apprentissage automatique, Physique Stellaire, Modélisation, Cosmologie, Etoiles Pulsantes, Echelles des distances

ABSTRACT

Pulsating variable stars, particularly Cepheids and RR Lyrae stars, are powerful tools for determining distances. Period-luminosity relations allow us to deduce their intrinsic luminosity, which is then compared to their apparent magnitude. In addition, the parallax-of-pulsation method uses the trigonometric relationship between the radius variation and the angular diameter variation to determine the distance. This latter technique involves a parameter, the projection factor, which converts the observed radial velocity to the photospheric pulsation velocity. This parameter, which is entirely degenerate with distance, shows significant statistical variations, preventing us from using parallax-of-pulsation on a large scale. A more detailed understanding of pulsation is required in order to determine the origin of this dispersion.

In this thesis, I first study the projection factor of RR Lyrae stars, using the SPIPS code and new radial velocity measurements, in order to compare it with Cepheids. I show that the value and dispersion of the projection factor for RR Lyrae stars are quite similar to what has been observed for Cepheids and that the different methods of measuring radial velocities have little impact on this dispersion, provided that homogeneous data are considered. I also present my work on the variation of the centre-to-limb darkening coefficient during pulsation, a key ingredient of the projection factor, using interferometric observations of the Cepheid η Aql.

Secondly, I develop generic light curves for Cepheids using SPIPS multi-band models. These templates make it possible to predict the shape of the light curve for a given pulsation period in a large variety of photometric bands. They are particularly effective for modelling light curves based on rare or imprecise observations, as is the case with extragalactic Cepheids.

KEYWORDS

Observations, Data Analysis, Machine Learning, Stellar Physics, Modelling, Cosmology, Pulsating Stars, Distance Scale