The Asteroseismic Potential of the NASA TESS satellite

by

Mathew Schofield

A thesis submitted to the
University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Solar and Stellar Physics
School of Physics and Astronomy
University of Birmingham
Birmingham, B15 2TT

February 2019

Abstract

A new generation of space-based observations has revolutionised the field of asteroseismology. In this thesis, we present the asteroseismic preparations undertaken for a new mission; the NASA TESS satellite.

Like its predecessor *Kepler*, TESS is primarily an exoplanet-hunting satellite. It will observe planetary transits around bright stars. Here, a parametric algorithm was developed to select 'high priority' stars for TESS short cadence observation. These are main sequence, subgiant and red giant stars that will display detectable solar-like oscillations. By observing these stars at a short cadence, TESS can revolutionise both the fields of asteroseismology and exoplanetary science.

We also present predictions made about the overlap between these two fields; the numbers of exoplanet-host stars that display solar-like oscillations observed by TESS. In addition, the scaling relations used to construct the algorithm are thoroughly tested and found to be robust. As part of this robustness testing, analytical equations to calculate reliable uncertainties for $\Delta\nu$ and $\nu_{\rm max}$ were also developed.

Finally, a machine learning algorithm is developed to select solar-like oscillators using only stellar observables. This algorithm was developed for TESS, but can be easily adapted to automatically perform high priority target selection for any future photometric mission.

"We are still pioneers, we've barely begun. Our greatest accomplishments cannot be behind us, because our destiny lies above us."

Cooper

INTERSTELLAR

"Another day, another box of stolen pens."

Homer Simpson

The Simpsons

For Mum, Dad and Charlotte.

Acknowledgements

I would like to thank the Solar and Stellar Physics group at the University of Birmingham. My special thanks go to Bill Chaplin, who secured funding for my course through the School of Physics and Astronomy at the University of Birmingham. He supervised the PhD, and provided insight, advice and belief in me throughout. This was invaluable, and I am very grateful.

My thanks also go to Dan Huber, Guy Davies, Andrea Miglio and Mikkel Lund, whose knowledge and opinions have been a huge help throughout the PhD. They made time to help me when I needed it, and I am very grateful. I would also like to thank the other members of the Solar and Stellar Physics group, for providing great company to work with. My thanks go to: Lou, Yvonne, Rachel, Steve, Warrick, Andy, Diego, Hugo, James, Caitlin, Tom, Ben, Oli, Alex, Eddie, Vedad (Dad), Saniya, Matt and Walter.

Most of all, thank you to my parents Andrew and Linda Schofield, and my sister Charlotte. Without your support, none of this would have happened.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The study of stars has been revolutionised by the ability to observe them from space. Space enables continuous observations of stars to be made, without these observations being hampered by the Earth's atmosphere.

One area of stellar physics that has greatly benefited from space-based observations is asteroseismology - the science of studying stellar oscillations. *Astero* comes from the ancient Greek *astrum*, meaning star. *Seismos* also comes from ancient Greek, and means earthquake. The combination of these words gives us the name of the science of studying stellar oscillations.

This thesis is concerned with preparations for a new space-based mission - the Transiting Exoplanet Survey Satellite (TESS; Ricker et al. 2014). This satellite aims to detect exoplanets orbiting around bright stars. The aim of the work done in this thesis is to ensure that the satellite also observes the maximum number of solar-like oscillators. If successful, TESS will revolutionize the fields of both asteroseismology and exoplanet science.

To select stars for TESS to observe, an algorithm was developed to calculate the probability of detecting solar-like oscillations within a star. Chapter 2 describes how stars were chosen for TESS to observe using this algorithm. The list of solar-like oscillators selected for TESS was named the Asteroseismic Target List (ATL; Schofield et al. submitted).

The main focus of TESS is on exoplanets and their host stars. Some of the exoplanet-host stars that TESS will observe will also display solar-like oscillations. Predictions were made to quantify the number of exoplanet-host stars that will also display solar-like oscillations. The paper describing this work is bound in Appendix B.

The method used to select solar-like oscillators for TESS to observe relied upon scaling relations - relations that relate the properties of a star with those of the Sun. In Chapter 3, *Kepler* LEGACY data was used to test these relations and create an updated scaling relation to calculate the frequency width of the solar-like oscillation envelope.

The next chapter focusses on another aspect of the method used to select solar-like oscillators for TESS. The method used to select stars for TESS in this thesis relied upon the estimation of the uncertainty on the global asteroseismic parameters (we will meet these parameters later in the introduction, see Section 1.3.4). *Kepler* LEGACY timeseries data was used to produce analytical equations to calculate the uncertainties on the global asteroseismic parameters (Chapter 4).

Lastly, we look towards future space missions. In Chapter 5, a method of selecting solar-like oscillators for future missions to observe is developed. Here, stars are selected using Machine Learning. With this tool, solar-like oscillators can be automatically selected for observation by any future space mission.

It is important to start with an introduction to the areas being covered in the thesis. We begin with a review of the basic equations of stellar structure, and the evolutionary path that solar-like oscillators take. After this, we will look at solar-like oscillations themselves, before finishing with a review of some of the most influential satellites from the perspective of asteroseismology.

1.2 Stellar evolution

This thesis focusses on solar-like oscillators; cool main sequence, subgiant and red giant branch stars. Here we will review the evolution of these stars from the main sequence to

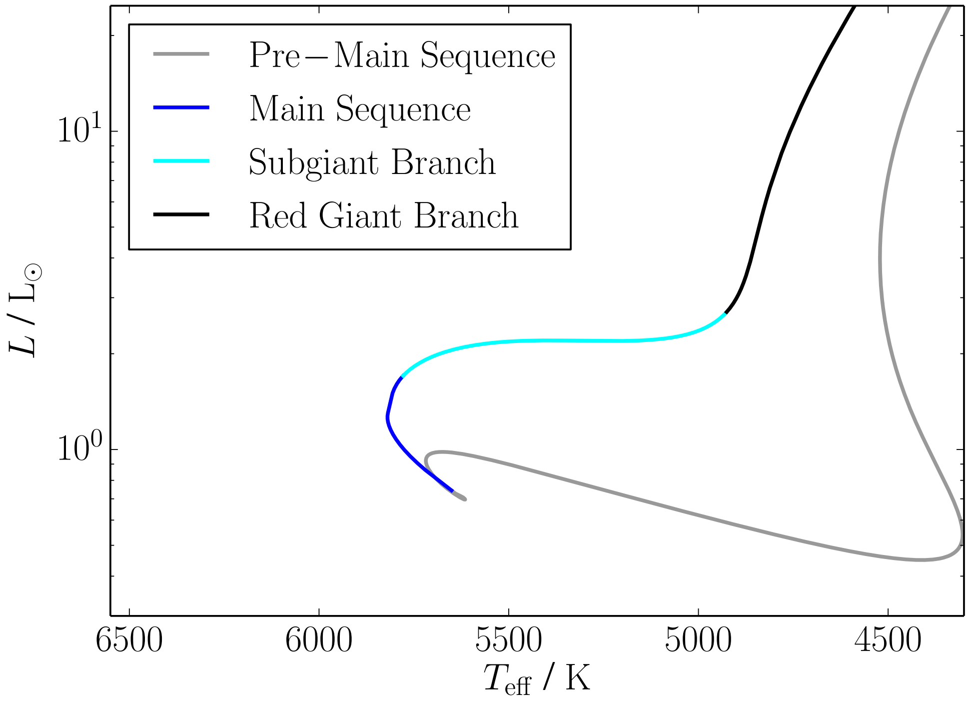


Figure 1.1: An HR diagram showing the stages of evolution of a $1\,M_\odot$ star from the main sequence to the red giant branch. The evolutionary track is from Rodrigues et al. (in prep), and was made using Modules for Experiments in Stellar Astrophysics (MESA)[1].

the red giant branch, in order to provide context for the work to come. Figure 1.1 shows these evolutionary stages on an HR diagram.

We will start off by reviewing some fundamental equations of stellar structure, before looking at different methods of heat transport inside stars. We will finish this section by reviewing the global composition of stars on the main sequence, the subgiant branch and the red giant branch, as well as how these stars undergo fusion.

1.2.1 Cool stars on the main sequence

Once a star has formed from the surrounding cloud of dust and gas, it will become opaque in the near-infrared (Larson, 2003). The star will exist at the base of the main sequence branch (the bottom of Figure 1.1), and core fusion will begin. To understand the physics that occurs here, we will review the equations of stellar structure.

---

[1] http://mesa.sourceforge.net/

1.2.2 Equations of stellar structure

The equations of stellar structure presented here will carry some implicit assumptions. We will assume that stars are perfectly spherical, that they are homogeneous, and that they do not change over time. These assumptions will allow us to broadly understand the physics of the stellar interior. We will start with an equation to describe the distribution of mass within the star.

Mass distribution

To describe the distribution of mass within a star, we can consider a spherical shell of mass $m(r,t)$. The mass of this shell will vary with radial distance from the star, and with time. This change is given by

$$dm = 4\,\pi\,r^2\rho\,dr \;-\; 4\,\pi\,r^2\rho\,v\,dt \tag{1.1}$$

(e.g. Kippenhahn et al. 2012). At a constant time, the mass contained within a spherical shell is given by the first term of equation 1.1;

$$dm = 4\,\pi\,r^2\rho\,dr\,. \tag{1.2}$$

The second term of equation 1.1 provides us with the mass flow out of a star of radius r (assuming spherically symmetric flow):

$$dm = -\,4\,\pi\,r^2\rho\,v\,dt\,, \tag{1.3}$$

where v is the velocity of the flowing mass. The minus sign of the equation refers to the direction of travel of this mass - outwards, away from the centre of the star.

Hydrostatic Equilibrium

During the majority of the stellar lifetime, stars are in a long-lasting state of equilibrium. The forces that act upon a spherical shell of mass cancel each other out. When this is the case, the star is said to be in hydrostatic equilibrium.

To obtain an equation for this hydrostatic equilibrium, we can equate the two forces acting upon our spherical shell of mass. These forces are the gradient of pressure (acting radially outwards) and the force due to gravity (acting radially inwards).

The mass of the spherical shell comes from equation 1.2. The mass per unit cross-sectional area of the shell is $\rho\, dr$. The force per unit cross-sectional area acting on this shell due to gravity is then $-g\rho\, dr$.

In order to maintain equilibrium, the net pressure force must act radially outwards and have the same magnitude as the force due to gravity. The net pressure acting on the spherical shell comes from the difference between the pressure acting at the lower boundary of the shell (P_{L}), and the pressure acting at the upper shell boundary (P_{U}). The net pressure force acting on the shell per unit area,

$$-\frac{\partial P}{\partial r}\, dr = P_{\mathrm{L}} - P_{\mathrm{U}}\,. \tag{1.4}$$

Since $P_{\mathrm{U}} < P_{\mathrm{L}}$, $-(\partial P/\partial r)$ is positive (the pressure on the shell acts radially outwards).

Equating the forces due to pressure and gravity gives the equation of hydrostatic equilibrium,

$$\frac{dP}{dr} = -\, g\, \rho\,. \tag{1.5}$$

Energy Conservation

To arrive at the equation for hydrostatic equilibrium, the pressure difference between the lower and upper boundaries of a spherical shell were considered (equation 1.4). To arrive at a similar equation for energy conservation, we instead consider the luminosity difference between the lower and upper boundaries of this shell.

We have already seen how the mass of a spherical shell changes with radius (assuming no flow of matter, equation 1.2):

$$dm = 4\,\pi\, r^2 \rho\, dr\,. \tag{1.6}$$

The change in luminosity between the boundaries of the shell, dL, is equal to $\epsilon\, dm$, where ϵ is the nuclear energy released per unit mass per unit time. This gives

$$dL = \epsilon\, dm = 4\,\pi\, r^2 \rho\, \epsilon\, dr\,. \tag{1.7}$$

Now we have seen how mass and luminosity can be described inside a star, as well as how the forces due to pressure and gravity can be balanced. Next, we will consider two methods of radial heat transfer: through radiation, and through convection.

Radiative Transport

Radiative transport inside stars can be thought of as a diffusive process. This is because the energy-transporting particles (the photons) have a very small mean free path l_p, compared to the characteristic length scale that radiative transport occurs over (the stellar radius R). To illustrate this, Kippenhahn et al. (2012) shows that $l_\mathrm{p}/R_\odot \sim 3 \times 10^{-11}$.

In order to obtain an equation for the radiative gradient of temperature, we start with the diffusive flux of radiative energy:

$$F = -D\, \nabla U\,, \tag{1.8}$$

where D is the coefficient of diffusion

$$D = \frac{1}{3}\, c\, l_\mathrm{p}\,, \tag{1.9}$$

and U is the energy density of radiation

$$U = a\, T^4\,. \tag{1.10}$$

a is the radiation-density constant, and c is the speed of light.

As the star is spherically symmetric, we will only consider the flux in the radial direction F_r. Once we differentiate equation 1.10, ∇U becomes

$$\frac{\partial U}{\partial r} = 4\,a\,T^3\,\frac{\partial T}{\partial r}\,. \tag{1.11}$$

Combining equations 1.8, 1.9 and 1.11 then gives

$$F_r = -\frac{4ac}{3}\,l_{\mathrm{p}}\,T^3\,\frac{\partial T}{\partial r}\,. \tag{1.12}$$

Finally, this can be solved for the gradient of temperature to give

$$\frac{\partial T}{\partial r} = -\frac{3}{16\pi ac}\,\frac{l_{\mathrm{loc}}}{r^2 T^3 l_{\mathrm{p}}}\,. \tag{1.13}$$

In equation 1.13, the diffusive flux of radiative energy F_r has been replaced by the local luminosity, $l_{\mathrm{loc}} = 4\pi r^2 F_r$. Equation 1.13 gives the temperature gradient associated with radiative heat transfer. Next, we will consider an alternative heat transport mechanism inside stars - convection.

Convective Transport

Granules are observed on the solar surface. These granules reveal that the Sun transports energy in the near-surface layers using convection. This is the process where hot parcels of gas rise slowly upwards (towards the stellar surface). This gas then cools, becomes denser and sinks around the edges of the granule (Iben, 2013).

If a parcel of gas were to have a higher temperature than the gas surrounding it (while maintaining a pressure equilibrium with the surroundings), this parcel would expand and become less dense. Buoyancy would then cause the parcel to rise. During this rise, the parcel of gas initially expands adiabatically, because the time scale of heat loss from the parcel to the surroundings is longer than the time scale for upward motion.

The parcel of gas will be lighter than the surroundings if

$$\left|\frac{\mathrm{dT}}{\mathrm{d}r}\right|_{\mathrm{rad}} > \left|\frac{\mathrm{dT}}{\mathrm{d}r}\right|_{\mathrm{ad}}\,, \tag{1.14}$$

where $|\mathrm{dT}/\mathrm{d}r|_{\mathrm{ad}}$ is the absolute adiabatic temperature gradient, and $|\mathrm{dT}/\mathrm{d}r|_{\mathrm{rad}}$ is the absolute temperature gradient of the surrounding matter. This is the Schwarzschild criterion for convective instability. When the Schwarzschild criterion is satisfied, the parcel of gas will be lighter than the surroundings and it will rise.

Eventually, heat transfer will cause the parcel to be in thermal equilibrium with its surroundings. Conversely, if a parcel of gas were to have a lower temperature than its surroundings, it would sink. Heat would be transferred to the sinking parcel from the surroundings, bringing the parcel into thermal equilibrium.

The main sequence

After the formation of a star, it will predominantly contain hydrogen and helium, with smaller amounts of heavier elements. The proportion of heavier elements within the star will depend on the cloud of material that the star formed from.

The core of the protostar will collapse, until the core density and temperature are high enough to commence nuclear fusion (Iben, 2013). Nuclear fusion provides the star with energy. This energy increases the temperature of the core, creating temperature and pressure gradients. These gradients provide a net outward force to combat the net inward force from gravity, maintaining hydrostatic equilibrium.

In lower mass main sequence stars, the energy to sustain an outward pressure gradient is provided by core hydrogen fusion through the proton-proton chain. In higher mass main sequence stars, the energy to sustain this outward pressure gradient is provided by hydrogen fusion through the CNO cycle. The CNO cycle is very temperature sensitive. This causes the radiative temperature gradient in the core to be much steeper than the adiabatic gradient, so the core will become convective.

The method of energy transportation throughout the star governs the internal structure of the star. The internal structure of a $1\,\mathrm{M}_{\odot}$ main sequence star is shown in Figure 1.2.

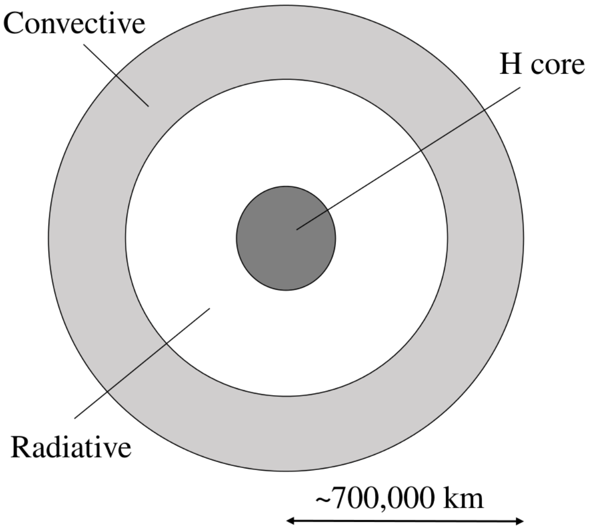


Figure 1.2: The internal structure of a 1 $\mathrm{M}_\odot$ main sequence star. The black circle shows the region where hydrogen fusion occurs (inner ∼20% by radius). A radiative region surrounds this core, and is shown in white (between ∼20-70% by radius). The light grey region surrounding this is the convective envelope, spanning the outer ∼30% of the star.

1.2.3 The subgiant branch

Hydrogen fusion on the main sequence will cause a build-up of helium in the core of the star. The cores of cool main sequence stars with masses $\leqslant 2.25\,\mathrm{M}_\odot$ will not be hot enough to undergo helium fusion. They will therefore contract, as there is a more gradual change in the outward pressure.

Eventually, the core of the star will stop contracting due to electron degeneracy pressure. A hydrogen-rich shell will surround this contracted core (Kippenhahn et al., 2012). The dense core will cause the shell surrounding it to compress due to gravity. The temperature of the shell will therefore increase, and hydrogen shell fusion will commence. Hydrogen shell fusion will heat the outer areas of the star, and the stellar luminosity will increase. The star will therefore expand and cool as it approaches the red giant branch.

1.2.4 Red giant branch stars

The star will continue to expand, while the surface temperature remains roughly constant. This will cause the luminosity of the star to increase. Throughout this phase, the helium core remains inert. Instead, hydrogen shell fusion will continue throughout the red giant

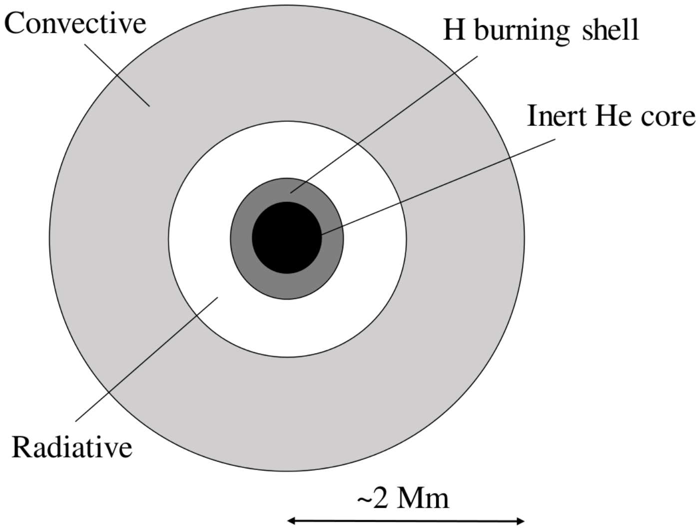


Figure 1.3: The internal structure of a $1\,\mathrm{M}_\odot$ low luminosity red giant branch star. The black circle represents the inert helium core, which occupies the inner ∼20% of the stellar radius. Hydrogen fusion occurs in a thin shell surrounding this (dark grey). A radiative region surrounds this shell, extending to ∼40% of the stellar radius (white). The expanded convective envelope is shown in light grey, and occupies the outer ∼60% of the stellar radius. Note that the radius of the star is heavily dependent on mass and evolutionary state, and is shown to give a rough sense of scale.

phase (Figure 1.3). A radiative zone will surround the hydrogen-fusing shell. Around the radiative zone there will be an expanded convective envelope.

The mass of the inert core increases during the red giant phase. This is because helium is accreted onto the core from the hydrogen-fusing shell. The core is compressed by the additional helium accreted onto it, and eventually becomes degenerate. As the core gains additional mass, the temperature will rise while maintaining a constant pressure. Once the core temperature is high enough ($\sim 10^8\,\mathrm{K}$), core helium fusion will commence. This marks the end of the red giant phase, and the beginning of the red clump phase.

We have now covered the evolution of solar-like oscillators, from the main sequence to the red giant branch. Next, we will discuss the modes inside of these stars, beginning with the theory of spherical harmonics.

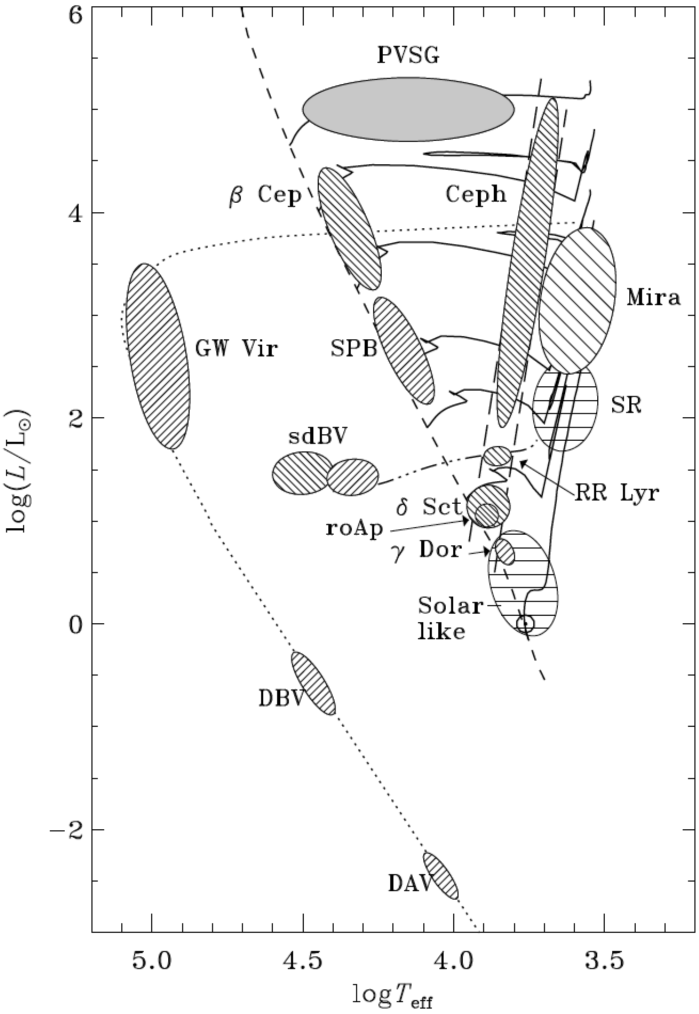


Figure 1.4: Different classes of stellar oscillator are shown here. This is commonly called a JCD diagram, after the author (Christensen-Dalsgaard, 2003). The Zero-Age Main-Sequence is shown by the dashed line. This thesis focusses on solar-like oscillators.

1.3 Asteroseismology

Asteroseismology is the science of studying stellar oscillations. There are many different classes of oscillating star. These are shown on Figure 1.4. We will focus on one of these classes - the solar-like oscillators. This class consists of the cool main sequence, subgiant and red giant branch stars that we discussed in the last section. Red Clump stars also show solar-like oscillations, but are not the focus of this thesis.

There are two types of mode associated with solar-like oscillators: gravity modes, and pressure modes. These modes have different restoring forces, and are sensitive to different regions of the star.

The restoring force of gravity modes is buoyancy (e.g. Chaplin & Miglio 2013). Gravity modes have high amplitudes in the stellar core, and low amplitudes at the stellar photosphere (Appourchaux et al., 2010). This is because for gravity modes, the parcel of gas driving the oscillation must be allowed to oscillate about its centre of gravity. In a convective region, parcels of hot, less dense gas rise towards the surface of the star. Simultaneously, cooler, more dense parcels of gas sink towards the interior. A gas parcel in a convective region is not able to oscillate around its centre of gravity.

The focus of this thesis is not on gravity modes, but on pressure modes (also referred to as p modes). In subsequent sections, we will go onto look at how pressure modes are intrinsically excited and damped. Before this, we will briefly review spherical harmonics.

1.3.1 Spherical harmonics

Providing that there are no large-scale structural changes to a sphere, any mode inside that sphere can be described using spherical harmonic functions. These functions have an angular dependence Y_l^m (Christensen-Dalsgaard, 2003; Campante et al., 2018). This dependence is described by an angular degree l and an azimuthal order m. Both l and m are integers, where $l \geqslant 0$ and $-l \leqslant m \leqslant l$.

Along with an angular degree l and an azimuthal order m, spherical modes are described with a third integer; their overtone number n. The overtone number of the mode corresponds to the number of nodes along the radius of the sphere.

As the ratio l/n tends to zero (i.e for high overtone modes), the asymptotic equation can be used to describe the cyclic frequencies of the modes ν_{nl}. To the second order, this equation is

$$\nu_{nl} = \Delta\nu\left(n + \frac{l}{2} + \epsilon\right) - l(l+1)D\,. \tag{1.15}$$

$\Delta\nu$ is the large frequency spacing between modes of the same angular degree l and consecutive overtone number n. ϵ is a surface correction to the mode frequencies (Ball et al., 2018). D is responsible for a frequency separation according to the azimuthal order m.

Christensen-Dalsgaard (2003) assigns the value 1.5μHz to this separation. In the Sun, the modes that are most excited have overtone numbers around $n = 20$.

The second term in equation 1.15 is known as the small frequency separation, and is a second order effect to the frequencies of the modes (e.g. Basu & Chaplin 2017). This is the difference in frequency between $\nu_{n+1,l-2}$ and ν_{nl} modes. For example, there is a small frequency separation between $l = 0$ and $l = 2$ modes, and between $l = 1$ and $l = 3$ modes. The large and small frequency separations will be visualised when we come to look at power spectra in Section 1.3.4. Before this, we will discuss the excitation of pressure modes.

1.3.2 Mode damping and excitation

The restoring force of pressure modes is the gradient of pressure. This pressure gradient is caused by turbulent convection in the near-surface layers of the star. A necessary condition for the existence of pressure modes within a star must therefore be the presence of near-surface convection.

This turbulent convection appears on the stellar surface as an irregular cellular pattern (Kallinger et al., 2014). Hot, less dense parcels of gas rise to the stellar surface, appearing as bright granules. The plasma then cools and sinks, once it has a higher density than its surroundings. The convective motion is turbulent, and stochastically excites an intrinsically-damped acoustic signal. These acoustic signals are called solar-like oscillations, and are the focus of this thesis (Christensen-Dalsgaard, 2002).

As we learnt in Section 1.2, cool main sequence, subgiant and red giant branch stars all have convective envelopes, so they exhibit pressure modes. This is why together they are referred to as the solar-like oscillators.

1.3.3 Modes at the stellar surface

In Section 1.3.1, we introduced the wavenumbers l and m. The angular degree l of a mode is the number of nodal lines across the surface of a star. The azimuthal order m of a mode is the number of these nodal lines that bisect the equator of the star.

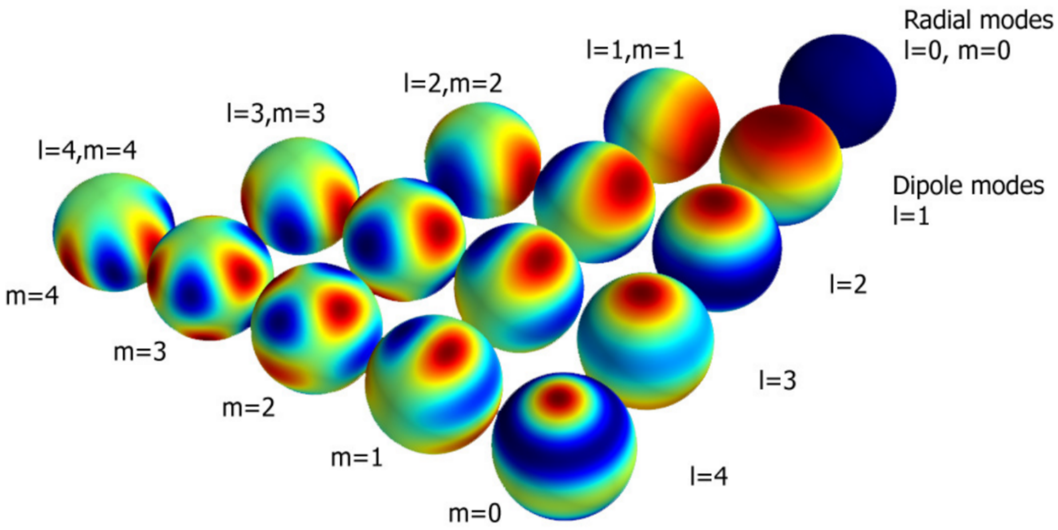


Figure 1.5: A visual representation of the displacement caused by pressure modes, with different combinations of the wavenumbers l and m. Blue areas represent displacement radially outwards, while red areas represent inward displacement of the gas. Figure from Beck (2013).

The pressure modes excited by near-surface turbulent convection distort solar-like oscillators. The wavenumbers l and m dictate the shape of this stellar distortion, as seen in Figure 1.5.

The simplest distortion comes from the mode with an angular degree l of 0; the radial mode. The radial mode causes the entire star to periodically expand and contract together, preserving the spherical shape of the star as it does so. This is not the case for the non-radial modes ($l > 0$). These modes cause some parts of the star to expand, while other parts are contracting (Figure 1.5).

In solar-like oscillators other than the Sun, it is only possible to detect pressure modes with angular degrees l of 0, 1 or 2 (although in some high signal-to-noise cases, $l = 3$ can also be detected). Higher angular degree modes cause smaller parts of the star to expand and contract (Figure 1.5); the amplitudes of higher angular degree modes average to zero. The luminosity variation from these high angular degree modes becomes indistinguishable from the background noise present in the observations.

We will now go on to see how pressure modes are visualised in the frequency domain.

This will tie together the asymptotic equation in Section 1.3.1 with the discussion here of low angular degree modes.

1.3.4 Modes in the power spectrum

Solar-like oscillations causes a displacement to the gas inside the star. They also cause a periodic change to the luminosity of the star. If the flux from the star is measured over time, the oscillations can be seen to cause the luminosity of the star to vary periodically (Figure 1.6 panel a).

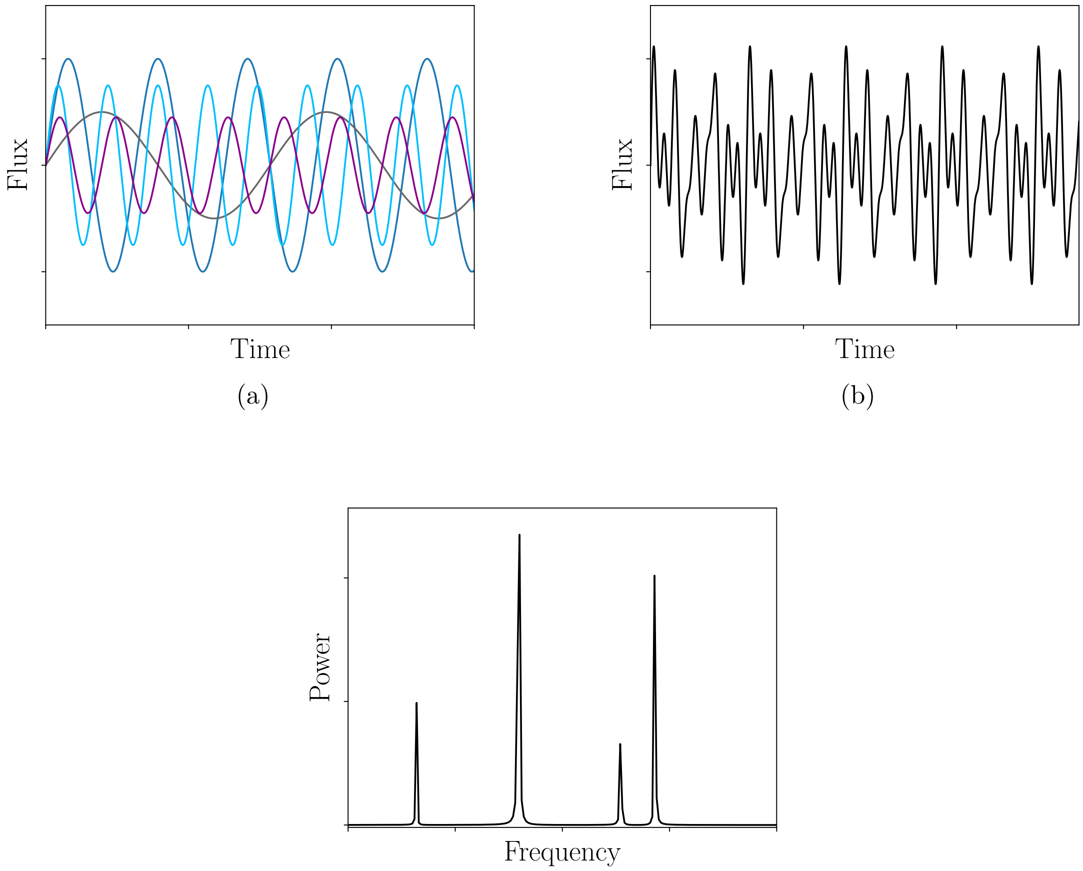
Each pressure mode will have a particular amplitude and frequency, which can become hard to distinguish from one another (Figure 1.6 panel b). If a Fourier Transform (FT) is performed on timeseries data, each periodic oscillation will be isolated in frequency space (Figure 1.6 panel c). By analysing the data in the frequency domain, the oscillations can be characterised more easily.

For simplicity, Figure 1.6 shows an ideal picture of four coherent modes. The pressure modes formed by stochastic turbulence inside solar-like oscillators are not coherent. These pressure modes are stochastically damped, and so have finite lifetimes.

There are also many other signals present in observational data. Some of these come from the star, and some signals are artefacts from the satellite. Other stars which appear close to the target (from the perspective of the observer) may also contribute some noise to the signal (known as blending). For the work done in this thesis, blending was not taken into account, as the stars in consideration are all much brighter than the surrounding stars. They therefore suffer very little from blending. Figure 1.7 shows the power spectrum of a cool main sequence star after observation with the *Kepler* satellite.

The power spectrum shown in Figure 1.7 consists of three components. The first component is due to stellar granulation (Section 1.3.2). In Figure 1.7, this was approximated using the Harvey law (Harvey, 1985), with an equation given in Kallinger et al. (2014). This equation will be described in more detail in Chapters 2 and Appendix B.

The second component of the power spectrum is shown in Figure 1.7 with a black



(a)

(b)

(c)

Figure 1.6: An example of a Fourier transform. Panel (a) shows four periodic sine waves, of different amplitudes and frequencies. The waves are superimposed in panel (b). The Fourier transform of this superposition is shown in panel (c).

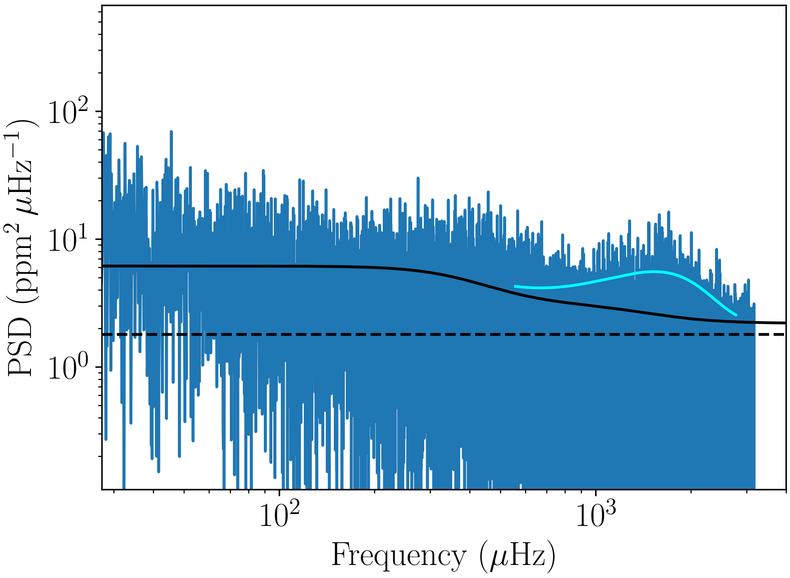


Figure 1.7: A power spectrum of KIC 2837475. The solid black line shows the signal from stellar granulation, modelled using a relation from Kallinger et al. (2014). The black dashed line shows the instrumental noise from the *Kepler* satellite, which made the original observation. The cyan line shows the power due to pressure modes.

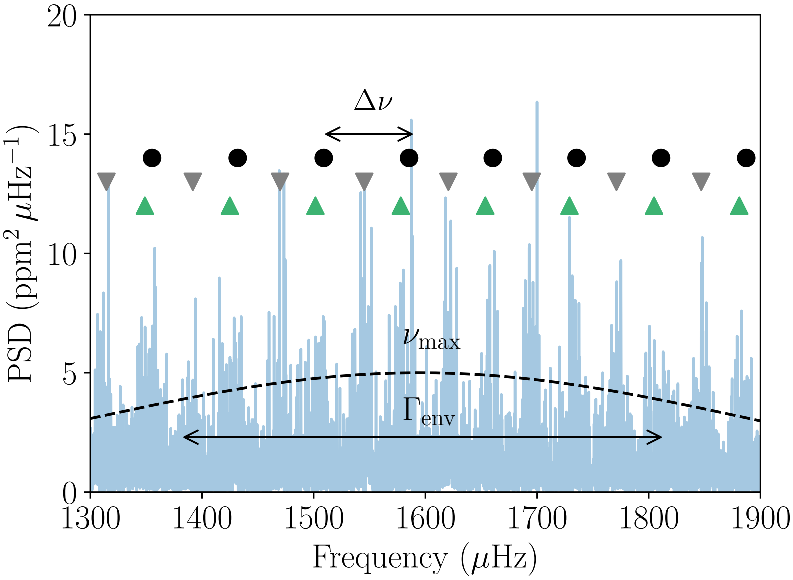


Figure 1.8: A power spectrum of KIC 2837475. The dashed line represents the solar-like oscillation envelope. The central frequency of this envelope, $\nu_{\rm max}$, and the FWHM envelope width, $\Gamma_{\rm env}$, have been annotated. The individual oscillations have also been marked: $l = 0$ with black circles, $l = 1$ with downward-facing triangles, and $l = 2$ with upward-facing triangles. The large frequency spacing between modes of the same angular degree $\Delta\nu$ is also shown.

dashed line. This comes from several background noise sources. It is frequency-independent, so is termed ‘white noise’. There are four main components to this noise. These are stellar shot noise, systematic noise, readout noise and blending. We will go into detail about what these noise sources are and how they can be modelled in Chapter 2.

The third component in Figure 1.7 comes from the solar-like oscillations themselves, which is shown with a Gaussian-shaped cyan line. If we take a closer look at the frequency range around the oscillations, regularly spaced peaks can be seen. Figure 1.8 again shows the power spectrum of KIC 2837475, after the background has been subtracted. The frequency range occupied by the pressure modes with highest amplitudes is shown here. The $l = 0, 1$ and 2 modes have been annotated, along with the Gaussian oscillation envelope.

In Section 1.3.3 we discussed how when observations of solar-like stars are made, higher angular degree l modes are not visible. Figure 1.8 is a visual example of this; modes with angular degrees $l \geqslant 3$ cannot be seen as their amplitudes are too low.

Figure 1.8 also shows that solar-like oscillations are contained within a Gaussian envelope (e.g. Chaplin et al. 2011). The central frequency of this envelope is known as ν_{max}. It scales to good approximation as $\nu_{\mathrm{max}} \propto g/\sqrt{T_{\mathrm{eff}}}$ (Kjeldsen & Bedding, 1995), where g is the surface gravity ($g = \mathrm{G}M/R^2$), and T_{eff} is the effective temperature of the star.

In this thesis, the probability of detecting pressure-mode oscillations was calculated. In order to make predictions about the detection probability for a star, ν_{max} first needed to be calculated. One way to calculate the ν_{max} value of a star is to compare it to the Sun. This can be done using a scaling relation:

$$\nu_{\mathrm{max}} = \nu_{\mathrm{max}\odot} \left(\frac{M}{M_\odot}\right) \left(\frac{R}{R_\odot}\right)^{-2} \left(\frac{T_{\mathrm{eff}}}{T_{\mathrm{eff}\odot}}\right)^{-0.5} . \tag{1.16}$$

where M is the stellar mass, and R is the stellar radius. $\nu_{\mathrm{max}\odot}$ is the solar value of ν_{max}, and has a value of $3100\,\mu$Hz (Huber et al., 2011; Chaplin et al., 2014a). Similarly, the solar effective temperature $T_{\mathrm{eff}\odot}$ is 5777 K.

In addition to $\nu_{\rm max}$, another useful parameter to be able to calculate is the frequency spacing between modes of the same angular degree, $\Delta\nu$. This was first introduced when we discussed the asymptotic equation in Section 1.3.1. $\Delta\nu$ is related to the sound speed $c_{\rm s}$ of the medium that the oscillations travel in,

$$\Delta\nu = \left(2\int_0^R \frac{dr}{c_{\rm s}}\right)^{-1}. \tag{1.17}$$

$\Delta\nu$ is an indicator of the mean density of the star; it scales to reasonable approximation as $\Delta\nu \propto \sqrt{\rho}$. It can be calculated using an equivalent scaling relation to equation 1.16:

$$\Delta\nu = \Delta\nu_{\odot}\left(\frac{M}{M_{\odot}}\right)^{0.5}\left(\frac{R}{R_{\odot}}\right)^{-1.5}, \tag{1.18}$$

where $\Delta\nu_{\odot} = 135.1\,\mu$Hz.

It follows that the large frequency separation $\Delta\nu$ can be directly related to the frequency of maximum oscillation amplitude $\nu_{\rm max}$. This relationship arises from the following: if equations 1.16 and 1.18 are combined, R can be eliminated. This gives

$$\Delta\nu = \nu_{\rm max}^{3/4}\, M^{-1/4}\, T_{\rm eff}^{-3/8}. \tag{1.19}$$

The mass M and effective temperature $T_{\rm eff}$ exponents are small - these parameters have a negligible effect on equation 1.19. Removing them reveals the relationship $\Delta\nu \propto \nu_{\rm max}^{0.75}$ (Basu & Chaplin, 2017).

In reality, the mass and effective temperature exponents of equation 1.19 are not negligible. By fitting to the $\Delta\nu$ and $\nu_{\rm max}$ values of bright stars with well constrained asteroseismic parameters, Stello et al. (2009) found the relationship to be $\Delta\nu \propto \nu_{\rm max}^{0.77}$.

Using the $\Delta\nu \propto \nu_{\rm max}^{0.77}$ proportionality, equations 1.16 and 1.18 can be simplified by removing the mass dependence from the equations. This leaves

$$\nu_{\rm max} = \nu_{\rm max\odot}\left(\frac{R}{R_{\odot}}\right)^{-1.85}\left(\frac{T_{\rm eff}}{T_{\rm eff\odot}}\right)^{0.92} \tag{1.20}$$

and

$$\Delta\nu = \Delta\nu_{\odot}\left(\frac{R}{R_{\odot}}\right)^{-1.42}\left(\frac{T_{\text{eff}}}{T_{\text{eff}\odot}}\right)^{0.71}. \tag{1.21}$$

Equations 1.16, 1.18, 1.20 and 1.21 are from Chaplin et al. (2011), with the coefficients in equations 1.20 and 1.21 from Campante et al. (2016).

ν_{max} and $\Delta\nu$ are together known as the global asteroseismic parameters. Equations 1.20 and 1.21 are both used throughout this thesis to calculate the asteroseismic properties of solar-like oscillators, so they have been introduced here.

Scaling relations such as equations 1.20 and 1.21 contain several implicit assumptions about the pressure modes within solar-like oscillators. Firstly, the scaling relations assume that all solar-like oscillators are similar to the Sun in all but their global properties. In the case of equations 1.20 and 1.21, only the stellar radius R and effective temperature T_{eff} are used to differentiate other stars from the Sun.

Equations 1.16 and 1.18 take into account the stellar mass M. These versions of the scaling relations were not used throughout this thesis, because in most cases this quantity was not available. Alongside the stellar mass, equations 1.20 and 1.21 ignore the effects of metallicity on the frequencies of the global asteroseismic parameters (e.g. Aguirre et al. 2015). This can reduce the precision of predictions made about ν_{max} and $\Delta\nu$.

Asteroseismic scaling relations assume that all solar-like oscillators are homologous to the Sun; they assume that you can get from any star to the Sun using a single number. In subsequent Chapters, we will look at these predictions in detail. We will also attempt to mitigate the effect of these unphysical predictions when stars are selected for the Asteroseismic Target List in Chapter 2.

In addition to ν_{max} and $\Delta\nu$, Figure 1.8 shows the FWHM of the oscillation envelope, Γ_{env}. The width of the Gaussian envelope is another useful descriptor of the solar-like oscillations within a star. In subsequent Chapters, we will calculate this parameter using pre-existing relations. The limitations with these relations will also be reviewed, and an updated relation to calculate the solar-like oscillation envelope width will be presented in Chapter 3.

Solar-like oscillations change as stars evolve. For main sequence stars, the solar-like oscillations at the highest amplitudes have relatively high frequencies. For example, in the Sun the Gaussian oscillation envelope is centred around 3100 μHz (Chaplin et al., 2011). As a star evolves, the envelope which contains solar-like oscillations moves towards lower frequencies. Figure 1.9 shows this process. It shows two main sequence stars (KIC 8006161 and KIC 12069424), two subgiants (KIC 6442183 and KIC 12508433) and a red giant branch star (KIC 6035199).

We have already seen that $\nu_{\rm max}$ is approximately proportional to $g/\sqrt{T_{\rm eff}}$. As solar-like oscillators evolve, their surface gravity g decreases while their effective temperature stays roughly the same. This implies that more evolved stars have Gaussian oscillation envelopes centred around lower $\nu_{\rm max}$ values.

In addition to the frequency of the Gaussian envelope decreasing as solar-like oscillators evolve, Figure 1.9 shows that the amplitude of oscillations increases during evolution (as shown by the vertical scale of the subplots in the Figure, c.f. Huber et al. 2010).

Kjeldsen & Bedding (1995) presents an empirical relation for the amplitude of the solar-like oscillation envelope. This equation relates the oscillation amplitude observed in intensity to the bolometric luminosity, mass and effective temperature of the star;

$$A_{\rm osc} \propto \frac{L}{M T_{\rm eff}^{0.5}}. \tag{1.22}$$

By making substitutions into equation 1.22, the oscillation amplitude can be shown to be proportional to $\log(g)$ and $T_{\rm eff}$ alone. Using $\log(g) \propto M/R^2$ and $L \propto R^2 T_{\rm eff}^4$, equation 1.22 becomes

$$A_{\rm osc} \propto \frac{T_{\rm eff}^{3.5}}{\log(g)}. \tag{1.23}$$

As a star evolves, its surface gravity decreases while its effective temperature stays roughly the same. This leads to larger oscillation amplitudes.

So far in this Chapter, we have neglected the observation of solar-like oscillations, and instead have focussed on the oscillations themselves. These observations can be made on

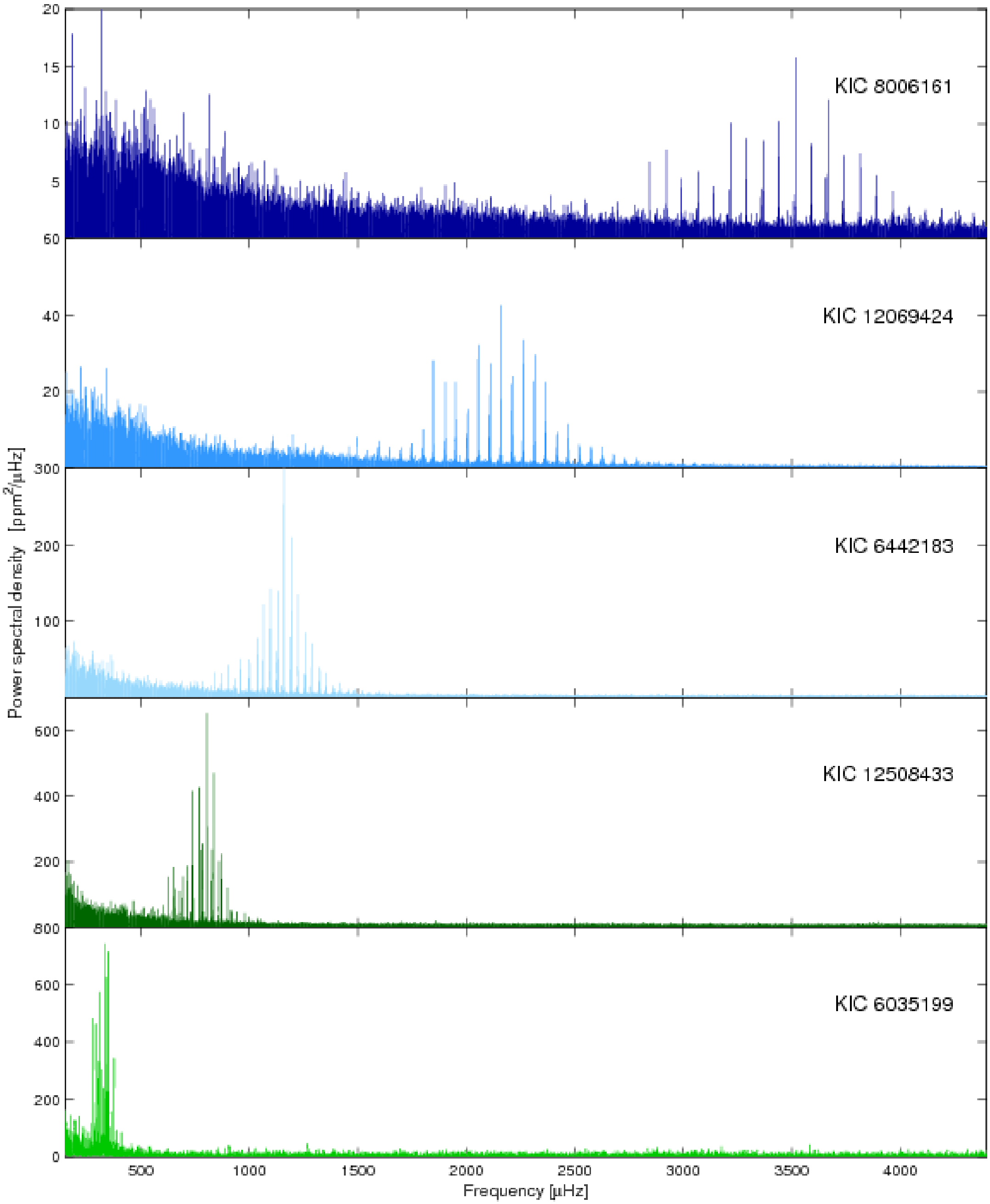


Figure 1.9: The oscillation envelope of 5 stars at different stages of evolution. KIC 8006161 (top panel) is a G8V dwarf star. The second panel shows KIC 12069424, a main sequence (G1.5V) star. KIC 6442183 (third panel) is a more evolved, G-type star. The fourth panel shows the K-type subgiant KIC 12508433. Finally the bottom panel shows the power spectrum of KIC 6035199, a G-type red giant. The Figure is from Chaplin & Miglio (2013). Spectral classifications were taken from Keenan & McNeil (1989) and Frasca et al. (2016).

the ground, or in space. This thesis used data from space-based missions, so we will now discuss some of the most influential satellites, from the perspective of asteroseismology.

1.4 Observations

The field of stellar astrophysics is currently undergoing a golden-age. This has been fueled by an enormous amount of data, from a number of satellites. Here we will review some of the most influential of these satellites.

1.4.1 *Gaia*

Gaia (Eyer et al., 2013) was launched in 2013 and is the follow-up mission to *Hipparcos* (Perryman et al., 1997). Like *Hipparcos*, *Gaia* is an interferometric satellite. It was commissioned with the aim of measuring the positions and distances of celestial bodies across the entire sky.

There are three main instruments on-board *Gaia*. Firstly, there is an astrometry instrument, which measures the 5 astrometric parameters: the positions (two angles), proper motions (from the two position angles) and the parallaxes (distances) of stars. The astrometric field is observed using 62 Charged Couple Device (CCD) cameras, out of a total of 106 CCDs for the satellite. The CCD configuration of the *Gaia* satellite is shown in Figure 1.10.

Gaia also contains a photometric instrument to measure stellar luminosities. It does this using one Blue and one Red Photometric band (BP and RP, respectively). The wavelength of the BP band is 330 - 680 nm, while the RP band has a wavelength of 640 - 1050 nm. The third instrument on-board *Gaia* is the Radial Velocity Spectrometer (RVS). This is used to measure stellar line-of-sight velocities through the Doppler shift of absorption lines. The RVS has a bandpass of 845 - 872 nm.

To date, there have been two data-releases from the *Gaia* satellite. The first of these is the Tycho-Gaia Asterometric Solution (TGAS; Michalik et al. 2015). TGAS combined positions from the *Tycho-2* catalogue (Høg et al., 2000) with proper motions, magnitudes

---

[2] https://www.cosmos.esa.int/web/gaia/focal-plane

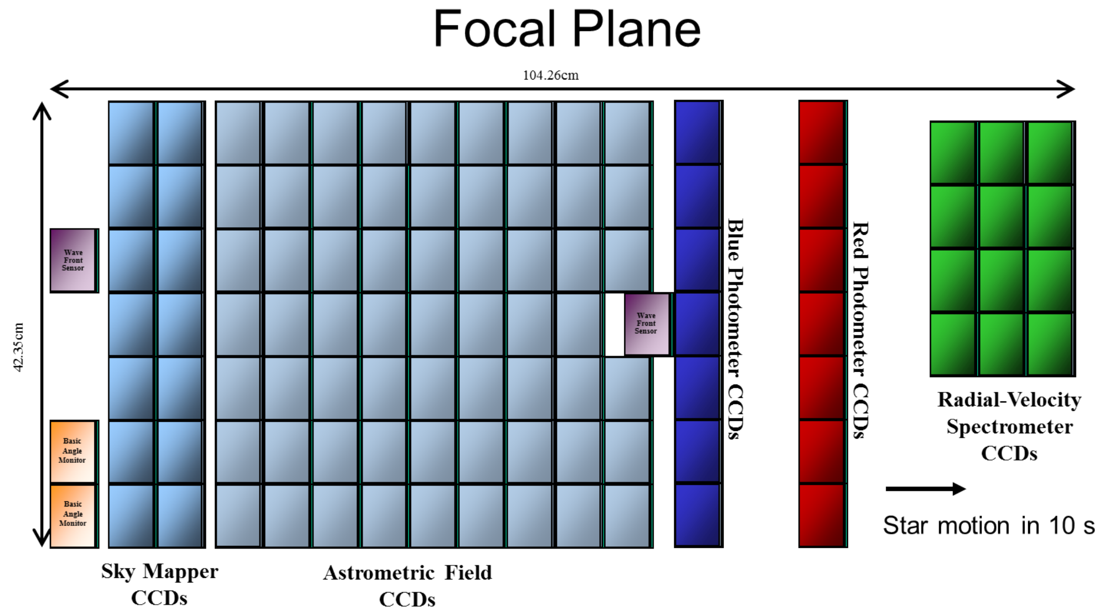


Figure 1.10: The focal plane of the *Gaia* satellite[2]. The layout of the CCDs for the astrometry, photometry, and radial velocity instruments are shown. Stars observed by *Gaia* will move across the image from left to right. The Wave Front Sensor (purple) and Basic Angle Monitor (orange) CCDs are used for satellite pointing corrections, the Sky Mapper CCDs (blue) are used for stellar detection; the Astrometric Field CCDs (light blue) measure the astrometric parameters, the Blue and Red photomultiplier CCDs provide photometric observations across different wavelength bands. Lastly, the CCDs used by the Radial Velocity Spectrometer (green) allow measurements of line-of-sight velocity to be made.

and parallaxes from the first year of *Gaia* observations. While this was an improvement on previous interferometric catalogues, asteroseismic parallaxes for eclipsing binary stars were shown to differ from those given in the TGAS catalogue (Davies et al., 2017).

On the 25th April 2018, there was another data release from the *Gaia* satellite; the *Gaia* Data Release 2 (DR2; Lindegren et al. 2018). The DR2 catalogue provides more precise parameters than TGAS, for a larger number of stars. The completeness and precision of DR2 makes it an ideal basis for the Asteroseismic Target List (c.f. Chapter 2).

1.4.2 *Kepler*

The *Kepler* satellite (Borucki et al., 2010) was designed with the aim of detecting Earth-sized planets orbiting around solar-type stars. It detected these planets by observing stellar transits; by detecting the decrease in flux caused by a planet passing in front of the star, with respect to the satellite.

Kepler has been in an Earth-trailing heliocentric orbit. To observe planetary transits, *Kepler* contains a primary mirror with a diameter of 1.4 meters, and is equipped with 42 CCD cameras. Rather than surveying the entire sky, the nominal *Kepler* mission observed a single 115 deg^2 area for 4 years, from 2009 to 2013 (Figure 1.11). During the nominal mission, *Kepler* achieved an impressive photometric precision of 25.4 ppm for 12th magnitude stars[3]. Photometric precision is a measure of the standard deviation of the observed light from a star. The lower the standard deviation for an observation, the higher the photometric precision.

This 4 years of observation produced extremely long-baseline timeseries, which yielded power spectra with excellent frequency resolution. Although *Kepler* was designed as an exoplanet-hunting satellite, it has revolutionised the field of asteroseismology by providing detailled power spectra of thousands of solar-like oscillators (e.g. Appourchaux et al. 2012, 2014; Davies & Miglio 2016). The data obtained by *Kepler* observations have been used

---

[3] https://keplergo.arc.nasa.gov/CalibrationSN.shtml

[4] https://www.sciencemag.org/news/2018/10/nasa-s-planet-hunting-space-telescope-nears-its-end

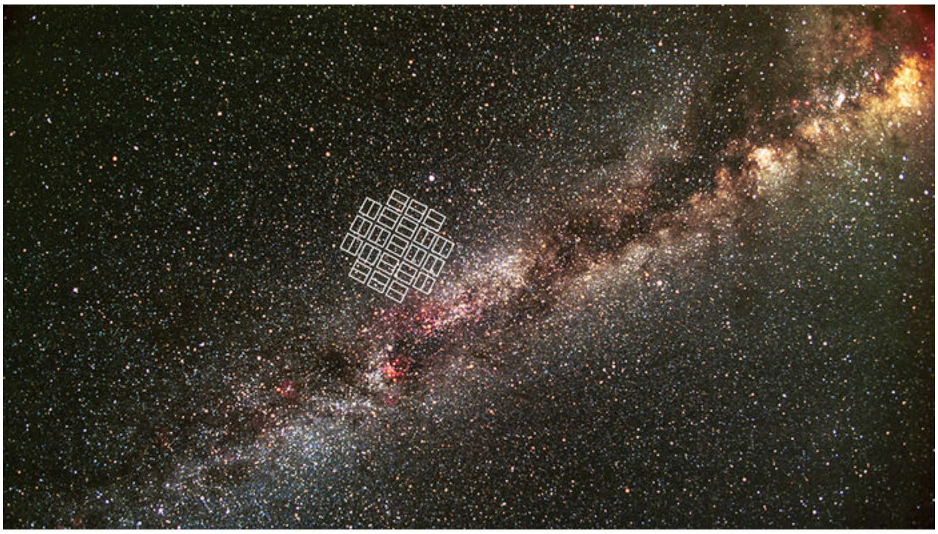


Figure 1.11: The field of view of the nominal *Kepler* mission[4].

to improve our understanding of the rotation profile (Lund et al., 2014), magnetic activity (Santos et al., 2018), age (Aguirre et al., 2015), global properties (Chaplin & Miglio, 2013) and evolution of stars (Miglio et al., 2012).

This thesis used *Kepler* data throughout. Firstly, *Kepler* data on main sequence stars was used to test the scaling relations (Chapter 3). These relations underpin the target-selection of solar-like oscillators. In addition, *Kepler* observations of the same main sequence stars were used to derive relations on the uncertainty on the global asteroseismic parameters ν_{max} and $\Delta\nu$ (Chapter 4). Thirdly, *Kepler* data on red giant stars was used to make an asteroseismic target-selection algorithm for solar-like oscillators. This algorithm can be used for any future photometric mission (Chapter 5).

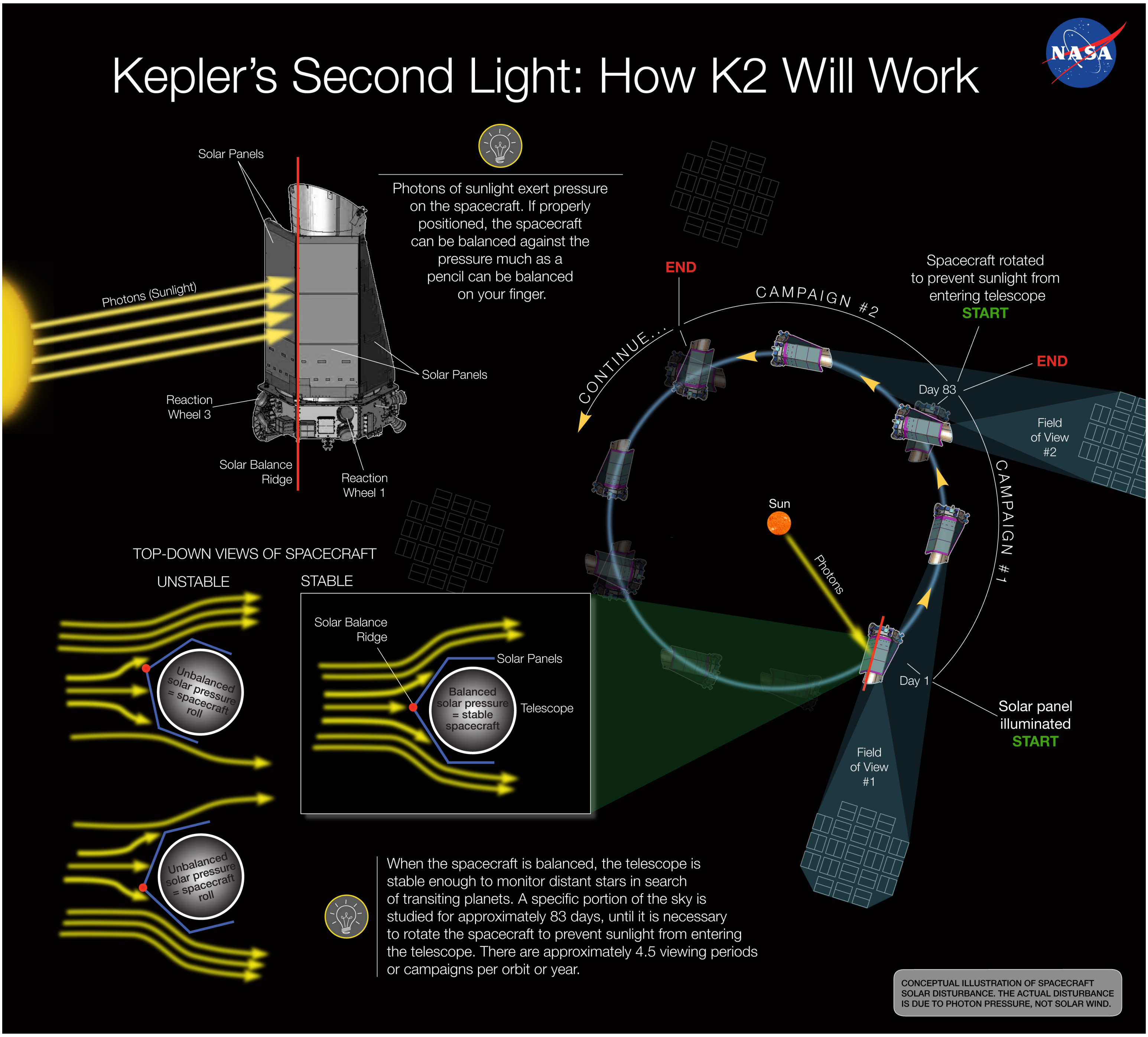
Kepler observed stars using two cadences. Approximately 150,000 targets were observed using a Long Cadence (LC) of 29.4 minutes. In addition, 512 targets could be observed at a Short Cadence (SC) of 58.9 seconds. The number of SC targets was limited to 512 because of the limited bandwidth on-board the satellite[6].

Kepler was equipped with four reaction wheels. Three of these are required to stabilise the spacecraft, and one is spare. In 2012, the first reaction wheel failed. In 2013, the second wheel failed, which meant that *Kepler* could not longer remain stable enough to

---

[5] https://keplerscience.arc.nasa.gov/k2-observing.html#campaigns

[6] https://keplerscience.arc.nasa.gov/data-products.html



Kepler's Second Light: How K2 Will Work

Photons of sunlight exert pressure on the spacecraft. If properly positioned, the spacecraft can be balanced against the pressure much as a pencil can be balanced on your finger.

When the spacecraft is balanced, the telescope is stable enough to monitor distant stars in search of transiting planets. A specific portion of the sky is studied for approximately 83 days, until it is necessary to rotate the spacecraft to prevent sunlight from entering the telescope. There are approximately 4.5 viewing periods or campaigns per orbit or year.

CONCEPTUAL ILLUSTRATION OF SPACECRAFT SOLAR DISTURBANCE. THE ACTUAL DISTURBANCE IS DUE TO PHOTON PRESSURE, NOT SOLAR WIND.

Figure 1.12: The field of view of the K2 mission, and the method used to maintain stability of the spacecraft[5]. The spacecraft is balanced against radiation pressure for up to 83 days as it orbits the ecliptic. During this time, solar panels at the side of the spacecraft will be illuminated. After this, the spacecraft will be rotated to prevent sunlight from entering the CCDs.

focus on a single area of sky. The *Kepler* mission was subsequently modified by pointing the satellite around the ecliptic plane, where radiation pressure from the Sun would act to stabilise the spacecraft's orbit (Figure 1.12). This modified mission was titled K2 (Howell et al. 2014).

K2 exceeded initial expectations, achieving a precision of 87 ppm for a 12th magnitude star[7]. Each campaign with K2 lasted approximately 83 days, before the satellite was readjusted to prevent contamination from sunlight. The K2 mission lasted for 19 full campaigns, during which hundreds of solar-like oscillators were detected (Lund et al., 2016; Huber et al., 2017). In October 2018, *Kepler* exhausted the last of its fuel and was decommissioned[8].

1.4.3 TESS

The focus of this thesis is the Transiting Exoplanet Survey Satellite (TESS). TESS will observe the majority of the sky over a period of 2 years. It's primary mission goal is to detect exoplanet transits around bright, nearby targets (Ricker et al., 2014).

To detect these transits, TESS will use four cameras, mounted vertically. Each camera contains an array of 2x2 CCDs, providing TESS with a total of 16 CCDs. The four cameras on-board TESS will observe one 'strip' of the sky at a time. The satellite will then rotate, until the majority of the southern hemisphere has been observed in these strips for at least 27 days. TESS will observe stars in these strips in a redder bandpass than *Kepler*; while *Kepler* observed stars in the range 400-850nm, TESS will observe stars at between 800-1000nm. This has the effect of reducing the intensity of the signal that TESS will observe, compared to *Kepler* (which will be discussed in the next Chapter).

TESS was launched on the 18th April 2018, and is in a 2:1 lunar resonant orbit[9]. This orbit causes TESS to reach perigee every 13.7 days, where data is downlinked to the Earth. TESS will observe each strip of sky for 27.4 days (for two orbits of the Earth). After one year, TESS will have finished observing the southern hemisphere using 13 observation

---

[7] https://keplergo.arc.nasa.gov/pages/photometric-performance.html

[8] https://www.nasaspaceflight.com/2018/10/kepler-retired-after-running-out-of-fuel/

[9] https://tess.gsfc.nasa.gov/documents/TESS_FactSheet_Oct2014.pdf

strips (Figure 1.13). The satellite will then rotate to observe the northern hemisphere in the same fashion. After two years, both the northern and southern hemispheres will have been observed, and thousands of exoplanets will have been discovered.

During these two years, TESS will observe the majority of targets at a Full-Frame Image (FFI) Cadence of 30 minutes. 200,000 high priority targets will be imaged at a Short Cadence (SC) of 2 minutes. 20,000 of these 2-minute cadence slots have been allocated to asteroseismology; for the observation of stellar modes of oscillation. The Asteroseismic Target List presented in Chapter 2 fills the majority of these 20,000 slots.

Each of TESS' four cameras will image a $24° \times 24°$ area in the sky, with a bandpass of 600 - 1000 nm. At any given time, the total collecting area of the satellite using it's four vertically-stacked cameras is $24°$ (longitude) $\times$ $96°$ (latitude) (see Figure 1.13).

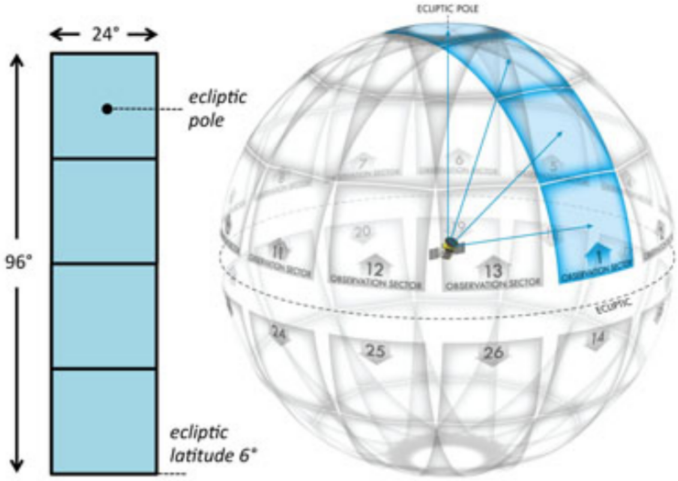


Figure 1.13: The 4 CCD cameras image a total area on the sky of $24°$ (longitude) $\times$ $96°$ (latitude). The CCD camera that will image the part of the sky at the highest latitude is centred around the ecliptic pole. TESS will observe 13 regions in the southern hemisphere, before observing 13 regions in the northern hemisphere. The image is from the NASA Transiting Exoplanet Survey Satellite website[10].

The cameras overlap at the ecliptic poles to increase the time coverage in these regions to up to 351 days (27 days × 13 regions in one hemisphere), Figures 1.14 and 1.15. A consequence of this overlap is that TESS will not observe $\pm 6°$ of the ecliptic latitude.

Using this all-sky coverage, TESS has huge potential to discover even more exoplan-

---

[10] http://tess.gsfc.nasa.gov/science.html

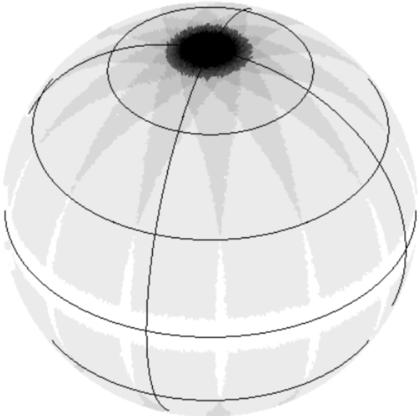


Figure 1.14: The TESS field of view on the celestial sphere. The black lines are lines of constant longitude or latitude. The colourbar represents how long a part of the sky will be observed by TESS. Different regions of observation will overlap with each other, so stars in certain parts of the sky will be imaged for longer. Stars close to the northern and southern ecliptic poles will be imaged for up to 351 days, as they will lie within 13 sectors of observation.

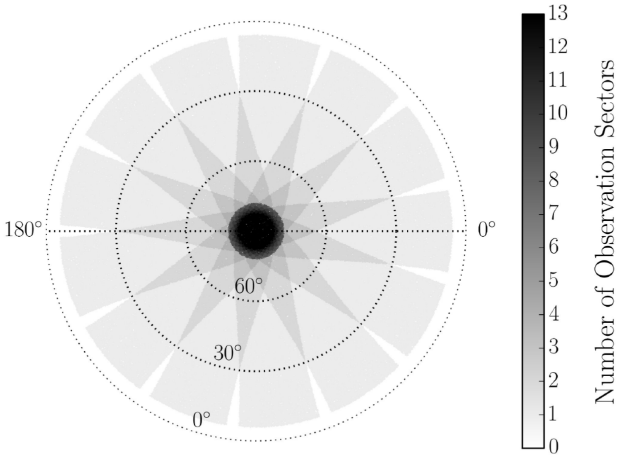


Figure 1.15: The field of view from one of the ecliptic poles. The colourbar represents how long a part of the sky will be observed by TESS. The dotted circles are lines of constant latitude (0°, 30° and 60°). The centre of the image has a latitude of 90°. The horizontal dotted line shows longitude values of 0° and 180$^{\odot}$.

ets that its predecessor, *Kepler*. The main aim of this thesis is to ensure that TESS is also successful at observing thousands of stars with potentially detectable solar-like oscillations, in order to advance the field of asteroseismology.

Chapter 2

The Asteroseismic Target List

This chapter is an extended, reformatted version of Schofield et al. (submitted). The paper has been submitted to the Astrophysical Supplement Series Volume 239[a]. It includes text and figures that are presented in that paper, as well as earlier work not presented there. The submitted paper is attached in Appendix A. A summary of the methodology and results of this Chapter were published beforehand in reports on the TESS Asteroseismic Science Operations Centre (TASOC) website[b]. Asteroseismic targets were selected for TESS in collaboration with Bill Chaplin, Dan Huber, Tiago Campante and the chairs of TASOC Working Groups 1 and 2.

The methodology used here was also used to make predictions for exoplanet-host stars. This was published in the Astrophysical Journal Volume 830 Article 138[c] (Campante et al., 2016), and can be found in Appendix B.

---

[a] http://iopscience.iop.org/journal/0067-0049

[b] https://tasoc.dk/

[c] http://iopscience.iop.org/volume/0004-637X/830

2.1 Introduction

In the previous chapter, we discussed solar-like oscillators; cool main sequence, subgiant and red giant branch stars. We also reviewed pressure mode oscillations inside these stars. Finally, we reviewed some of the satellites that observe these oscillations, finishing with a discussion of the TESS mission. This chapter combines these three areas. Here, an Asteroseismic Target List (ATL) is made to select stars for observation that have the highest probability of containing detectable oscillations.

Moreover, the focus of the ATL is on ‘solar-like oscillators’; stars which contain solar-like pressure-mode oscillations. These solar-like oscillators were selected for observation by the Transiting Exoplanet Survey Satellite (TESS; Ricker et al. 2014).

Section 2.2 will go on to explain why TESS - a predominantly exoplanet-finding satellite - has been allocated an asteroseismic cohort of stars to observe. This section will also explain the philosophy behind selecting stars in the ATL. After this, Section 2.3 describes the catalogues that the stars in the ATL were selected from. These sections will provide context to the Asteroseismic Target List, before the selection procedure is outlined.

The methodology used to select stars for the ATL is detailed in Section 2.4. This Section reviews cuts that were made to the data to only keep solar-like oscillators, the calculation of the total power due to solar-like oscillations, and an estimation of the background power that contributes to the observed signal. These oscillation and background powers were used together to calculate the probability of detecting solar-like oscillations within a star.

Calculating the probability that solar-like oscillations will be observed within a star is not novel. This was first done for the *Kepler* satellite by Chaplin et al. (2011). The method was subsequently developed for TESS in Campante et al. (2016) (Appendix B). As it is integral to this thesis, it will be reviewed here.

After the ATL was constructed in Section 2.4, the stars within the target list were ranked. Section 2.5 explains how this was done. Firstly, the HR diagram was separated into regions of $T_{\rm eff}\,/\,L$ space. After this was tried, stars were instead ranked using Kernel Density Estimation. After reviewing both of these ranking methods, they were disregarded. In their place, a mixture of detection probabilities was chosen to rank the stars in the ATL. This produced a target list which contains the stars that are most likely to contain detectable solar-like oscillations, alongside hotter stars that are on the transition towards showing no solar-like pressure mode oscillations. By also sampling this region, the transition between solar-like oscillators and hotter stars can be further explored, as this area had been poorly sampled by *Kepler*.

After the stars in the ATL were ranked, a limited number high-priority stars were selected from the ATL in Section 2.6. This limited number of high-priority stars will be observed at higher cadence, and so will be subject to less instrumental noise from signal attenuation. The stars in the ATL that were set aside as high-priority targets were the stars which are most similar to the Sun. These are the stars that will benefit most from higher cadence observation, and the lower instrumental noise level associated with it.

So far, detection probabilities have been calculated for the stars in the ATL. The ATL stars were then ranked, before high-priority stars were set aside. After all of this work, it is important to be sure that the detection probabilities calculated for the stars in the ATL are robust. Namely, that these probabilities can be relied upon as a precise measure of the detectability of solar-like oscillations within a star.

To check this, Monte Carlo trials were performed using the inputs into the target list. The inputs were perturbed according to their uncertainties, before a new detection probability was calculated. This was done iteratively for every star in the ATL, yielding 1000 detection probabilities for each star. The distribution of these 1000 detection probabilities is a measure of the robustness of the method used to select stars for the ATL: the smaller the spread of values, the more robust the method used to select stars. These Monte Carlo trials are described in Section 2.7.

The Monte Carlo trials confirmed the robustness of the detection probabilities calculated in the ATL. After this, the calculated properties of the stars in the ATL are compared to values from literature in Section 2.8. Between Sections 2.7 and 2.8, the method used to construct the ATL was shown to be reliable. Lastly, Section 2.9 summarises the Asteroseismic Target List and concludes the Chapter.

2.2 Philosophy behind construction of the ATL

In the introduction, TESS was described as a predominantly exoplanet-finding satellite. Like *Kepler* (Borucki et al., 2010), TESS observes these exoplanet-host stars using two cadences. The majority of stars will be observed with a Full-Frame Image Cadence of 30 minutes. TESS will also observe a 'high priority' cohort of 200,000 stars. These will be observed at a Short Cadence of 2 minutes.

20,000 of these Short Cadence slots were set aside for asteroseismology. These slots were allocated to the TESS Asteroseismic Science Operations Centre (TASOC). Through the study of asteroseismology, the field of exoplanetary science will also be advanced. Examples of this symbiosis between asteroseismology and exoplanetary science include Huber (2015) and Campante et al. (2016).

TASOC is a collaboration of the scientists involved in asteroseismology with TESS[1]. Within TASOC, different classes of stellar pulsator have been divided. They are separated into Working Groups 0-8. The Asteroseismic Target List (ATL) is a list of Short Cadence targets proposed by Working Groups 1 (Asteroseseismology of TESS exoplanet hosts) and 2 (Oscillations in solar-type stars). This ATL comprises the majority of the 20,000 targets allocated to TASOC.

Different stages of this work are presented in reports written for TASOC Working Groups 1 and 2 (Chaplin et al., 2016a,b,c). These were written to explain the method of high-priority asteroseismic target selection for TESS to the wider science community. The reports also provided regular updates to the work that was being done. The asteroseismic community provided feedback to the target-selection and ranking processes used to make the ATL. This feedback was reviewed during TESS Asteroseismic Science Consortium meetings.

When constructing the ATL for Working Groups 1 and 2, the aim was to select solar-like oscillators. These solar-like oscillators may be selected in one of two ways. One

---

[1] https://tasoc.dk/

method is to select stars that are cool enough to contain convective envelopes, and rank them from brightest to faintest, according to their apparent magnitudes. The work in this Chapter will show that ranking stars by apparent magnitude alone would drastically reduce the yield of stars with detectable solar-like oscillations.

Alternatively, the probability of detecting solar-like oscillations could be calculated for every solar-like star within a sample. This probability would take into account the power due to solar-like oscillations, the background granulation power, and the power due to instrumental noise. Once this detection probability is calculated for a sample of stars, it could be used to rank each target from most to least likely to exhibit detectable solar-like oscillations.

These detection probabilities were calculated for every star in the dataset, and used to rank the list. Before this method is explained in more detail, the datasets used to construct the ATL are reviewed.

2.3 Input Catalogues

The ATL was made by calculating detection probabilities for a sample of stars. Before calculating these probabilities, a suitable sample of stars needed to be chosen. Since TESS is an all-sky mission, targets were selected for the ATL using all-sky catalogues. Predominantly, the *Gaia* satellite's Data Release 2 (DR2; Lindegren et al. 2018) was used to select stars for the ATL. It will be shown that the vast coverage and astounding precision of the DR2 catalogue makes it an ideal choice for a dataset to use when making the ATL.

To supplement the DR2 catalogue at the brightest end, the Extended Hipparcos Compilation (XHIP) from the *Hipparcos* satellite was also used to construct the ATL (Anderson & Francis, 2012). Together, the DR2 and XHIP catalogues provide a comprehensive, all-sky dataset with which to select solar-like oscillators from.

Both DR2 and XHIP are interferometric catalogues; they provide a position, a stellar parallax, a (B-V) colour, and a V-band magnitude for every star. In order to calculate detection probabilities, these inputs were used to calculate ‘global’ stellar parameters for each target.

These parameters included effective temperature T_{eff}, luminosity L and radius R. After calculating the global parameters, cuts were made to the DR2 and XHIP catalogues to remove non solar-like oscillators. Finally, the probability of detecting oscillations was calculated for every star. Before this, the section below describes the process of calculating global parameters for every star in the DR2 and XHIP catalogues.

2.3.1 Global stellar properties

Before a detection probability could be calculated for every star, some stars in the DR2 and XHIP catalogues needed to be removed from the list, to only retain potential solar-like oscillators. To remove these stars, the global parameters of the targets in the DR2 and XHIP catalogues needed to be known. Calculating these global parameters is the topic of this section. The first parameter to be calculated was the apparent magnitude of each star.

The ATL was constructed at the same time as the TESS Input Catalogue (TIC) was being made. The TIC is a vast collection of stars that TESS will observe. Each TIC object is assigned a TESS-band magnitude T_{mag}. This is the apparent magnitude of the star in the photometric band that TESS will observe in.

As the TIC was not complete during the construction of the ATL, the I-band magnitude (I_{mag}) was used for each star in this work instead. I_{mag} has a very similar spectral response to T_{mag}, so it provides a good approximation to the TESS-band apparent magnitude. I_{mag} was also the apparent magnitude used to make predictions for TESS in Campante et al. (2016) (Appendix B).

While I-band magnitudes were available for the stars in the XHIP catalogue, they were not available in DR2. To calculate them, (B-V) colours and V-band magnitudes were used. Three sources of (B-V) and V_{mag} values were used for the stars in the DR2 catalogue.

Where available, colours and magnitudes from the revised Hipparcos catalogue were used (van Leeuwen, 2007). Where these were not available, Tycho-2 values from Høg et al. (2000) were substituted. In cases where there was no entry in the revised Hipparcos or Tycho-2 catalogues, values from the AAVSO All-Sky Photometric Survey (APASS) were used (Henden et al., 2009).

In order to calculate I_{mag} for the DR2 stars, (B-V) colour was first converted to (V-I). To do this, stars were first separated into 'giants' or 'dwarfs' using an empirical relation which relates absolute g-band magnitude with (B-V) colour:

$$g_{\mathrm{Mag}} = 6.5(B - V) - 1.8. \tag{2.1}$$

g_{Mag} is a proxy for the luminosity of a star, while (B-V) colour is a proxy for effective temperature (equation 2.3). Equation 2.1 therefore relates effective temperature with luminosity, and corresponds to a straight line on an HR diagram. Low luminosity red giant branch stars will lie above this straight line, while subgiant and main sequence stars lie below it. The position of a star with respect to this line determines whether a star is classified as a 'dwarf' or a 'giant'.

Once the stars were classified as dwarfs or giants, (V-I) colours were calculated for the DR2 stars using polynomials from Caldwell et al. (1993). These (V-I) colours can then be used to calculate I-band magnitudes for the stars:

$$I_{\mathrm{mag}} = V_{\mathrm{mag}} - (V - I). \tag{2.2}$$

Now, every star in the DR2 and XHIP catalogues has an I-band magnitude. After this, distances were obtained for every star in both catalogues. For the stars in DR2, this was straightforward.

Distances have been calculated for every star using inversions in Bailer-Jones et al. (2018). In that work, each stellar parallax is treated as an observed quantity, which has been drawn from a distribution. The width of this distribution is dictated by the parallax uncertainty. For the DR2 stars, these were the distances used in the ATL.

Sadly, this inversion process has not been performed for the stars in the XHIP catalogue. Instead, distances were calculated for each of these stars using the inverse of the parallax, $d = 1/\pi$.

Every star in the DR2 and XHIP catalogues now has a (B-V) colour, an I-band magnitude, and a distance. Because the stars lie at different distances across the Galaxy, the effect of reddening needed to be taken into account. This is an effect caused by the interaction of light with clouds of dust or gas in the galaxy. This interaction causes the wavelength of the light to be made 'redder'.

E(B-V) and A_V are reddening and extinction coefficients, respectively. They were used to de-redden (B-V) colours and V_{mag} magnitudes when calculating effective temperature and luminosity. These reddening and extinction values were calculated from Galactic coordinates and distances, using the 'Combined15' dust map from the MWDUST Python package (Marshall et al., 2006; Green et al., 2015; Drimmel et al., 2003; Bovy et al., 2016).

These de-reddening coefficients were subtracted from the colours and magnitudes in the DR2 and XHIP catalogues. Once de-reddened, the (B-V) colours were used to calculate effective temperatures T_{eff} from the polynomials and coefficients in Torres (2010);

$$\log(T_{\text{eff}}) = a + b(B - V) + c(B - V)^2 + ..., \tag{2.3}$$

with $-0.4 \leqslant (B - V) \leqslant 1.7$ limits imposed by Flower (1996).

Following this, stellar luminosities were calculated for every star. This was done using the equation given in Pijpers (2003),

$$\log(L/\mathrm{L}_{\odot}) = 4.0 + 0.4M_{\text{bol}\odot} - 2.0 \log \pi - 0.4 \left(V_{\text{mag}} - A_V + \mathrm{BC}_V\right). \tag{2.4}$$

This takes bolometric corrections BC_V from Flower (1996), as presented in Torres (2010). Along with these bolometric corrections, equation 2.4 uses stellar parallaxes, and de-reddened V-band magnitudes. Torres (2010) also gives the bolometric luminosity, $M_{\text{bol}\odot}$, as 4.73 ± 0.03 mag.

Once effective temperature and luminosity values were calculated from every DR2 and XHIP star, calculating equivalent stellar radii follows simply. This was done using the Stefan-Boltzmann law,

$$R/R_\odot = (L/L_\odot)^{0.5}\,(T_\mathrm{eff}/T_\odot)^{-2}, \tag{2.5}$$

where $T_\odot$ is the solar effective temperature, 5777 K.

After the stellar parameters were calculated for every star, some stars in the DR2 and XHIP sample were removed. Non solar-like oscillators were removed from the sample, alongside duplicate stars.

2.4 Constructing the ATL

In the previous section, global stellar parameters were calculated for every star in the DR2 and XHIP catalogues. Cuts were then made to both catalogues in order to retain cool stars with convective envelopes - the 'solar-like' stars.

2.4.1 Retaining solar-like oscillators

Several steps were taken to cut stars from the DR2 and XHIP catalogues. This was done to ensure that the ATL contained only potential solar-like oscillators. The first cut that was made was to parallax. A fractional parallax (σ_π/π) cut of 50% was imposed on the data. This removed the stars with the least precise stellar parameters.

Along with the parallax cut, stars were removed from the dataset if they lay on the hot side (the blue edge) of the δ-Scuti instability strip, i.e., if $T_\mathrm{eff} > T_\mathrm{red}$. This 'red-edge' effective temperature, T_red, was given in Chaplin et al. (2011) as

$$T_\mathrm{red} = 8907\,(L/L_\odot\,)^{-0.093}. \tag{2.6}$$

These cuts ensured that the ATL consisted solely of solar-like oscillators. Before the stars with detectable oscillations could be selected from this list, duplicate entries between the DR2 and XHIP catalogues were removed.

2.4.2 Combining the DR2 and XHIP catalogues

Once cuts were made to the DR2 and XHIP dataset, only solar-like stars remained. Even after these cuts, some of the stars listed in the DR2 catalogue were also listed in XHIP. These duplicated entries needed to be removed before a probability of detection could be calculated for every star in the dataset.

There were ∼ 17,000 stars which exist in both lists. These stars have two sets of input properties, which were calculated in Section 2.3.1. When combining the DR2 and XHIP catalogues, a decision had to be made about which set of properties to accept for these stars.

The precision across the ATL would be maximised if parameters were used from the catalogue with the smallest fractional parallax. σ_π/π from both DR2 and XHIP was calculated for the ∼ 17,000 'overlapping' stars. The set of stellar properties from the catalogue with the smallest fractional parallax σ_π/π value was used. Typical fractional parallax values quoted in the XHIP catalogue range from 10 - 30 %, while fractional parallaxes in DR2 rarely exceed an impressive 1 %. This resulted in XHIP properties being used for only a few stars, see Figure 2.1. The DR2 parallaxes are so precise that they were used for almost every star in the ATL.

After cuts were made to the dataset to retain solar-like stars, duplicate entries between the DR2 and XHIP catalogues were removed from the list. For the remaining stars, all that is left is to calculate P_{det}; a probability of detecting solar-like oscillations within a star.

2.4.3 Calculating a probability of detecting solar-like oscillations

Calculating the probability of detecting solar-like modes of oscillation within a star was first developed by Chaplin et al. (2011) for the nominal *Kepler* mission. This method was updated for TESS in Campante et al. (2016). This is how stars were selected for the

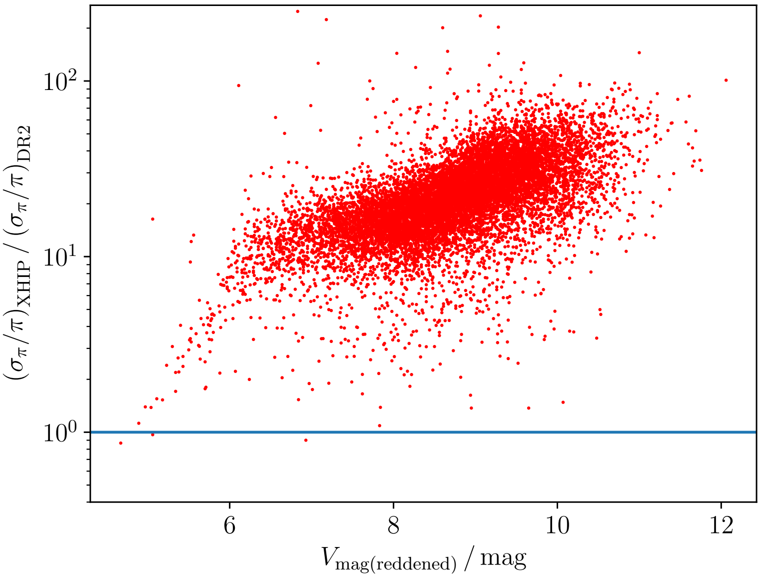


Figure 2.1: The stars with both DR2 and XHIP entries. The fractional parallax σ_π/π value was calculated from the DR2 and XHIP entries for each star separately. The ratio of these values is plotted on the vertical axis. The parameters from the catalogue with the lower σ_π/π value was chosen. XHIP properties were used for only the few stars that lie below the blue 1:1 ratio line.

ATL here.

The method is divided into four parts; the calculation of the observing time in the TESS field, the calculation of the power due to oscillations, the calculation of the total background power, and the calculation of the detection probabilities, P_{det}.

Calculating the Observation Time in the TESS field, T_{obs}

The motivation of the TESS mission and the observation strategy of the satellite were both explained in Section 1.4.3 of the introduction. Here we will explain how the observation time in the TESS field of view, T_{obs}, was calculated for every star in the ATL.

The field of view of TESS is defined using the ecliptic system; the plane covered by the Sun, as seen from Earth. The observation time in the TESS field of view was calculated for a given star using it's ecliptic coordinates (E_{long}, E_{lat}).

Before T_{obs} was calculated, some stars were removed from the ATL. In the introduction, we saw that TESS will not observe between 6° above and below the ecliptic plane. Stars within ± 6° of the ecliptic plane were therefore removed from the target list. At the time the ATL was constructed, the initial pointing of the satellite was not known. For this reason, all stars above +6° or below −6° were assumed to be observed for at least one month. When calculating the observation time in the TESS field of view, an initial pointing of $E_{\mathrm{long}} = 0°$ was used for the satellite.

After the stars that lie between ± 6° of the ecliptic were removed, T_{obs} was calculated for every star. T_{obs} may be calculated for a given star by defining the longitude values of the centre of each observation region ϕ_{CCD}, and the longitude range ϕ_{range} that the cameras cover at a given latitude (the latitude of the star). Figure 2.2 shows a pictorial representation of ϕ_{CCD} and ϕ_{range} at a constant latitude. ϕ_{range} is given by the equation

$$\phi_{\mathrm{range}} = \frac{24°}{\cos(E_{\mathrm{lat}})}. \tag{2.7}$$

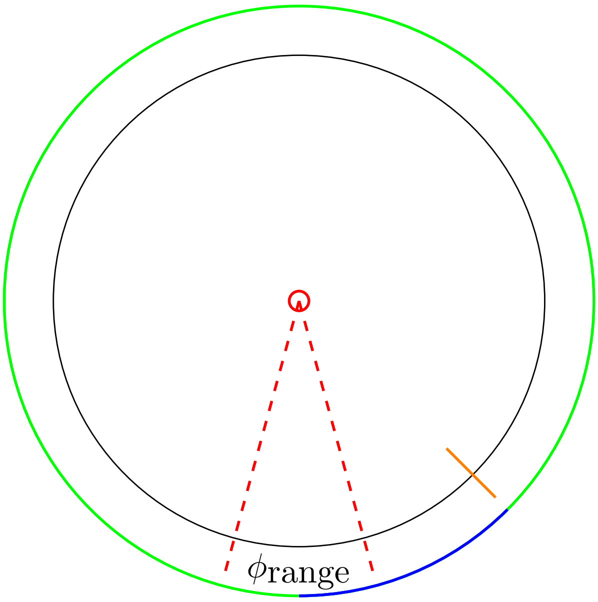


Figure 2.2: An example of what is calculated in order to determine whether a star lies inside ϕ_{range}. ϕ_{CCD} lies at the centre of ϕ_{range}. In this example, the star (represented by the orange line) lies outside of the satellite's field of view.

E_{lat} is the latitude of the star in question, and 24° is the width of the field covered by the CCD cameras at 0° latitude. If the longitude of a star lies within $\phi_{\mathrm{CCD}} \pm \phi_{\mathrm{range}}/2$, then the image of that star will be captured by a camera onboard TESS. In order to determine whether a star will be viewed by a camera, the longitude difference ϕ_{Diff} between E_{long} (the longitude of the star) and ϕ_{CCD} (the longitude of the centre of the camera) must be calculated. This is given by

$$\phi_{\mathrm{Diff}} = \begin{cases} |\phi_{\mathrm{CCD}} - E_{\mathrm{long}}|, \\ 360^\circ - |(\phi_{\mathrm{CCD}} - E_{\mathrm{long}})|. \end{cases} \tag{2.8}$$

ϕ_{CCD} lies at the centre of ϕ_{range}. When calculating the value of ϕ_{Diff} from Equation 2.8, it is important to calculate the value 'both ways around the circle' (shown by the blue and green lines in Figure 2.2). Only the smaller angle between E_{long} and ϕ_{CCD} (the blue line in Figure 2.2) should be used as the angle between a star and the centre of the CCD, ϕ_{Diff}.

The black circle in Figure 2.2 is a line of constant latitude. In this figure, the satellite is represented by the small red circle in the centre of the image. The red dashed lines show the width of the field of view of the satellite. These are the edges of ϕ_{range}. The longitudinal position of a star has been represented as the orange line. The green and blue arcs show the two values that equation 2.8 give. In Figure 2.2, the longitude difference show by the blue arc will be the accepted value for ϕ_{Diff}, as this is shorter than the green arc. As the star lies outside of ϕ_{range}, it will not be observed by TESS during this sector.

This calculation was performed for every sector, on every star in the target list. The largest contiguous number of sectors that a star is observed in was taken as the observation time of that star in the TESS field, T_{obs}. After this, the total power due to solar-like oscillations was calculated for each star.

The total power due to solar-like oscillations

Firstly, the total mode power due to oscillations across the solar-like mode envelope was found for every remaining star in the dataset. After this, the equivalent total background power across the solar-like mode envelope was determined. The ratio of these two power values could then be calculated. This ratio of power values was then used to calculate a detection probability for each star.

For the remainder of this chapter, the total power due to oscillations in the solar-like modes will be written as $P_{\rm tot}$. Conversely, the total background across the solar-like envelope will be $B_{\rm tot}$. The mode-to-background power ratio is known as the signal-to-noise ratio (SNR).

Estimating $P_{\rm tot}$ first requires the global asteroseismic parameters $\nu_{\rm max}$ and $\Delta\nu$; the frequency of maximum power due to solar-like oscillations and the large frequency separation between modes of the same spherical angular degree and successive overtone, respectively. The scaling relations to calculate $\nu_{\rm max}$ and $\Delta\nu$ were first introduced in Chapter 1; we remind the reader of the relations here. When constructing the ATL, mass estimates were not available for the target stars. The following scaling relations therefore do not rely on mass to provide values of the global stellar parameters $\nu_{\rm max}$ and $\Delta\nu$.

$$\nu_{\rm max} = \nu_{\rm max\odot}\left(\frac{R}{R_\odot}\right)^{-1.85}\left(\frac{T_{\rm eff}}{T_\odot}\right)^{0.92} \tag{2.9}$$

and

$$\Delta\nu = \Delta\nu_\odot\left(\frac{R}{R_\odot}\right)^{-1.42}\left(\frac{T_{\rm eff}}{T_\odot}\right)^{0.71} \tag{2.10}$$

give the global asteroseismic properties for the stars (Campante et al., 2016). The coefficients in equations 2.9 and 2.10 were derived from the $\Delta\nu \propto \nu_{\rm max}^{0.77}$ proportionality first given in Stello et al. (2009).

As well as $\nu_{\rm max}$ and $\Delta\nu$, the maximum amplitude for radial mode oscillations, $A_{\rm max}$, was required in order to estimate the total power contained in the Gaussian power-mode excess. The relation

$$A_{\rm max} = 0.85\times 2.5\,\beta\left(\frac{R}{R_\odot}\right)^{1.85}\left(\frac{T_{\rm eff}}{T_{\rm eff\odot}}\right)^{0.57}. \tag{2.11}$$

was adapted from Chaplin et al. (2011) with the inclusion of the 0.85 factor. This factor was included because of TESS' redder bandpass compared to *Kepler*'s (which the relation was originally developed for). 2.5 ppm corresponds to the maximum solar radial mode amplitude $A_{\mathrm{max}\,\odot}$ in the *Kepler* bandpass.

Lastly, β represents the suppression of modes close to the δ-Scuti instability strip (Houdek, 2006).

$$\beta = 1.0 - e^{-(T_{\mathrm{red}} - T_{\mathrm{eff}})/1550}\,, \tag{2.12}$$

where the 'red edge' temperature T_{red} is from equation 2.6. There is less energy associated with a mode inside a star that lies close to the instability strip. This is because these stars have thinner convective envelopes, and it is convective turbulence that drives the pressure mode oscillations (Chapter 1). The β factor describes this energy-reduction in terms of an amplitude suppression.

As well as the maximum radial mode amplitude A_{max}, the Full Width at Half Maximum (FWHM) of the Gaussian power excess (Γ_{env}) needed to be known for the stars in the dataset. Originally, this was estimated using ν_{max} as

$$\Gamma_{\mathrm{env}} = \begin{cases} 0.66\,\nu_{\mathrm{max}}^{0.88} & \text{if } \nu_{\mathrm{max}} \leqslant 100\,\mu\mathrm{Hz}\quad, \\ \nu_{\mathrm{max}}\,/\,2 & \text{if } \nu_{\mathrm{max}} > 100\,\mu\mathrm{Hz}\quad, \end{cases} \tag{2.13}$$

which comes from Stello et al. (2007) and Mosser et al. (2010, 2012).

Equation 2.13 was superseded by an updated relation to calculate the Gaussian envelope FWHM. This updated Γ_{env} relation was developed using LEGACY data (Aguirre et al. 2017; Lund et al. 2017), and is presented fully in Chapter 3. This updated relation was used to construct the ATL, and includes a temperature dependence:

$$\Gamma_{\mathrm{env}} = \begin{cases} 0.66\,\nu_{\mathrm{max}}^{0.88}\,(1 + (T_{\mathrm{eff}} - T_{\odot})\,6 \times 10^{-4}) & \text{if } T_{\mathrm{eff}} > 5600\,\mathrm{K}\quad, \\ 0.66\,\nu_{\mathrm{max}}^{0.88} & \text{if } T_{\mathrm{eff}} \leqslant 5600\,\mathrm{K}\quad, \end{cases} \tag{2.14}$$

where $T_{\odot}$ is the solar effective temperature.

Putting together the elements we have so far allows the total power contained within the oscillation envelope, $P_{\rm tot}$, to be calculated. This is given by

$$P_{\rm tot} = 0.5\, c\, \eta^2 A_{\rm max}^2 \, \frac{2\,\Gamma_{\rm env}}{\Delta\nu}\,, \tag{2.15}$$

from Chaplin et al. (2011). c is the effective number of modes per order; the sum of the relative visibilities of the $l = 0, 1, 2$ and 3 modes with respect to the radial ($l = 0$) mode within one overtone (Ballot et al., 2011). The value of c that was used for TESS in this work is 2.94. This was calculated by solving the matrix given in Bedding et al. (1996), with the coefficients u_2 and v_2 from Allen (1978). η is a damping factor attributed to apodization of the signal (equation 2.21). This is explained in more detail in the following section, when we discuss various sources of noise (c.f. Figure 2.5).

Equation 2.15 provides a measure of the total power due to solar-like oscillations within a star. After this, the equivalent total background power across the oscillation envelope was calculated for each star.

The total background power

Two separate noise sources were estimated in order to calculate the total background power across the frequency range occupied by solar-like modes, $B_{\rm tot}$. These noise sources are the instrumental noise from the TESS satellite, and stellar granulation.

Before the background noise level could be estimated, the number of pixels that are used in the aperture by each target needed to be known. The original equation used to calculate this was

$$N_{\rm aper} = 10 \times (n + 10)\,, \tag{2.16}$$

where n is

$$n = 10^{0.4(20 - I_{\rm mag}) - 5}\,. \tag{2.17}$$

This equation was updated to better describe the pixel mask size of targets observed

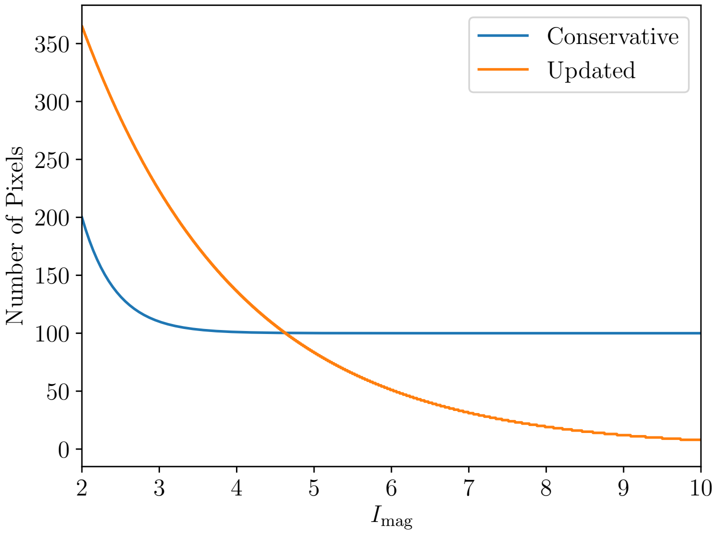


Figure 2.3: The TESS pixel cost per star as a function of I_{mag}. The function from equation 2.18 is shown in orange. The function from equation 2.16 is in blue. For very bright stars, the pixel cost very quickly becomes enormous.

by TESS. The updated mask size equation was used to produce the ATL, and is given by

$$N_{\text{aper}} = 10^{(0.8464 - 0.2144(I_{\text{mag}} - 10.0))} . \tag{2.18}$$

The mask size equations 2.16, 2.17 and 2.18 are from Dan Foreman-Mackey (private communication). The number of pixels calculated here is rounded up to the nearest whole number. This number is shown as a function of I_{mag} in Figure 2.3. In this Figure, equation 2.18 is shown in orange. The previous, conservative estimate of the number of pixels in the photometric aperture is from equation 2.16, and is shown in blue.

Once N_{aper} was calculated for the DR2 and XHIP stars, the instrumental noise sources could be estimated. These sources were estimated using the procedure from Sullivan et al. (2015).

There are four different sources of instrumental noise from TESS. The first of these is *sky noise* (also known as blending or dilution). This is the contribution from other stars which appear close to the target. The second source of noise is *readout noise*, which is an

---

[1] https://spacebook.mit.edu/

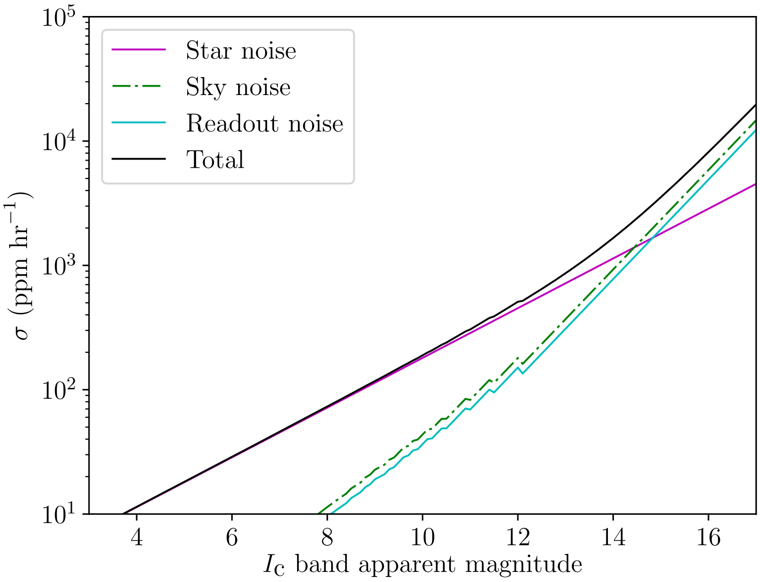


Figure 2.4: The instrumental sources of noise for TESS. The star noise, sky noise and pixel readout noise are all shown as functions of I_{mag}. The solid black line shows the sum of these, giving the total expected noise level for a star as a function of I_{mag}. This plot was first produced for the work I completed in Campante et al. (2016).

artefact of the photons from the signal being converted to electrons within a CCD camera. The camera does not have a shutter, so electrons are accumulated when the photons are being read out of the pixels. All of the electrons from a row of pixels are then summed over to process the image.

Alongside sky and readout noise is *systematic noise*. This is an intrinsic artefact of the satellite; it is an engineering requirement on the satellite's design and not attributed to a single noise source. It scales with the total observing time as $t^{-1/2}$. When constructing the ATL, this noise source was not known, so was assumed to be $0\,\mathrm{ppm\,hr}^{1/2}$. There are good reasons to assume this, given the excellent performance of *Kepler* (Caldwell et al., 2010).

The last source of noise is *stellar shot noise*. This is due to photon counting statistics. These four noise sources were estimated using the `calc_noise` IDL procedure from the TESS Wiki[2]. The noise function is shown in Figure 2.4.

---

[2] https://spacebook.mit.edu/

These four sources of instrumental noise were summed in quadrature. This sum, σ, was used to calculate the Power Spectral Density (PSD) value of the instrumental noise, b_{instr}. This is given by the equation

$$b_{\mathrm{instr}} = 2 \times 10^{-6}\,\sigma^2\,\Delta t, \tag{2.19}$$

from Chaplin et al. (2011). Δt is the 120-sec cadence for the target stars. The instrumental noise is measured in ppm$^2\,\mu$Hz^{-1}.

The second noise source is stellar granulation; the turbulent effect caused by hot plasma rising to the surface of the star, before cooling and sinking. This was estimated using an equation from Kallinger et al. (2014):

$$P(\nu)_{\mathrm{gran}} = \eta(\nu)^2 \left(\sum_{i=1}^{2} \frac{\zeta_i a(\nu)_i^2 / b(\nu)_i}{1 + \nu b(\nu)_i^c} \right). \tag{2.20}$$

The coefficients a and b are also from Kallinger et al. (2014). The value for c that is used is 4, and for ζ is $2\sqrt{2}/\pi$.

When calculating the granulation power observed in a star, the Nyquist frequency needs to be taken into account. The Nyquist frequency ν_{Nyq} of a target is half of the sampling frequency; $\nu_{\mathrm{Nyq}} = 1/(2\Delta t)$, where Δt is the cadence of observation.

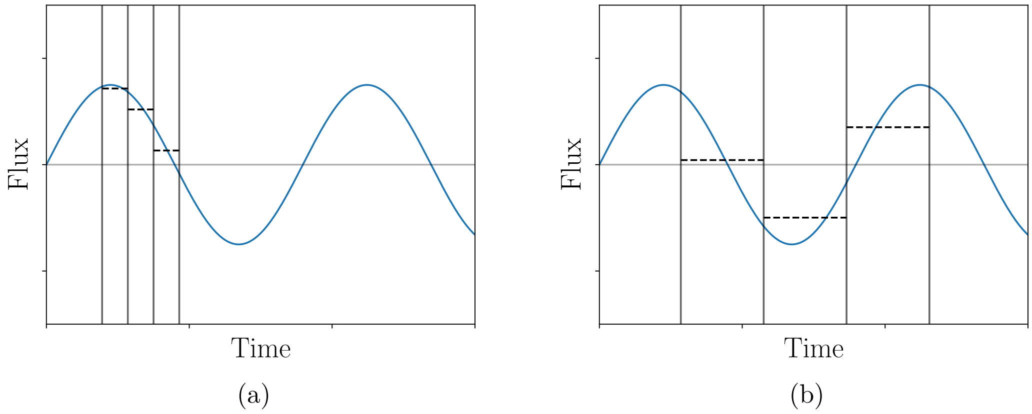
The Nyquist frequency causes two effects in the power spectrum; apodization (damping of the signal), and the appearance of aliases. The first effect, apodization, is shown in Figure 2.5. It was taken into account when calculating granulation using the damping factor $\eta(\nu)$. This is given as

$$\eta(\nu) = \mathrm{sinc}\left(\frac{\nu_{\mathrm{max}}}{2\nu_{\mathrm{Nyq}}}\right), \tag{2.21}$$

(e.g. Chaplin et al. 2014b).

As well as apodization, the Nyquist frequency causes a second effect: aliases. These aliases are reflections of the signal around the Nyquist frequency.

If a star has a ν_{max} value below the Nyquist frequency, it is known as a sub-Nyquist



(a) (b)

Figure 2.5: An example of apodization in the time domain. The cadence of observation is represented by the vertical black lines. The dashed lines represent the average of the signal between cadences. Panel (a) shows the effect of a fast cadence: the periodic signal can be well resolved. Conversely, panel (b) shows the effect of a much slower cadence: it becomes difficult to resolve the periodic sinusoid (the sinusoid will appear to have a lower amplitude).

$\nu_{\rm max}$ Region	$\nu_{\rm alias}$ value
$\nu_{\rm max} > \nu_{\rm Nyq}$	$\nu_{\rm Nyq} - (\nu_{\rm max} - \nu_{\rm Nyq})$
$\nu_{\rm max} < \nu_{\rm Nyq}$	$\nu_{\rm Nyq} + (\nu_{\rm Nyq} - \nu_{\rm max})$

Table 2.1: If $\nu_{\rm max}$ lies above $\nu_{\rm Nyq}$, the alias of the granulation will lie below $\nu_{\rm Nyq}$. Conversely if $\nu_{\rm max}$ lies below $\nu_{\rm Nyq}$, the alias will lie above $\nu_{\rm Nyq}$.

star. Conversely, if a star has a $\nu_{\rm max}$ value above the Nyquist frequency, it is a super-Nyquist star. Aliases mean that sub and super-Nyquist stars must be treated differently from each other. The power distribution around $\nu_{\rm max}$ differs depending on whether the modes of oscillation lie above or below the Nyquist frequency.

The total granulation power from equation 2.20 is the sum of the real and aliased granulation components (i.e the components above and below the Nyquist frequency). For each star, one component exists above the Nyquist frequency and one lies below (Table 2.1). The sum of these components is the total power due to granulation, $P_{\rm gran\ (tot)}$.

Equations 2.19 and 2.20 were summed together, after taking into account the effects of apodisation and aliases. This sum is then multiplied by the frequency range $2\,\Gamma_{\rm env}$. Although pressure modes are stochastically excited and damped at a much larger range of frequencies in the star, the vast majority of the mode power is contained in the range $\nu_{\rm max} \pm \Gamma_{\rm env}$, where $\Gamma_{\rm env}$ is the FWHM envelope width of the star. The sum is taken across $2\,\Gamma_{\rm env}$, giving the total background $B_{\rm tot}$;

$$B_{\rm tot} = 2\,\Gamma_{\rm env}\,(b_{\rm instr} + P_{\rm gran\ (tot)}). \tag{2.22}$$

The Detection Probability

So far in this section, the power due to solar-like oscillations $P_{\rm tot}$, and the power due to background noise $B_{\rm tot}$, have been calculated. The ratio of these powers can be used to calculate a detection probability, $P_{\rm det}$, for each star. This subsection explains how these detection probabilities were calculated.

Once the oscillation and background power values were calculated, the global signal-to-noise ratio could be obtained for each star. This is given by

$$\mathrm{SNR} = \frac{P_{\rm tot}}{B_{\rm tot}}. \tag{2.23}$$

This ratio applies over the number of frequency bins that span the Gaussian envelope, N. This number of bins was calculated using the envelope width and observing time of each star;

$$N = \frac{2\,\Gamma_{\rm env}\,T_{\rm obs}}{10^6}. \tag{2.24}$$

The factor of 10^6 ensures that the number of bins N is a dimensionless quantity - the units of $\Gamma_{\rm env}$ are μHz, and $T_{\rm obs}$ is given in seconds. The noise across these N bins is assumed to follow χ^2 $2N$ d.o.f statistics.

After calculating the signal-to-noise ratio for each star, a threshold signal-to-noise ratio, $\mathrm{SNR}_{\rm thresh}$, was also calculated. To calculate $\mathrm{SNR}_{\rm thresh}$, a false-alarm probability p of 5% was set. In other words, at $\mathrm{SNR}_{\rm thresh}$ there is a 95% chance that the signal is due to solar-like oscillations, rather than noise.

The probability P that the SNR lies above $\mathrm{SNR_{thresh}}$ is

$$P(\mathrm{SNR} \geqslant \mathrm{SNR_{thresh}}) = p. \tag{2.25}$$

Equation 2.25 is solved for $\mathrm{SNR_{thresh}}$ by substituting P with

$$P = \int_{x}^{\infty} \frac{e^{-x}}{\Gamma(N)} x^{N-1} dx, \tag{2.26}$$

where $\Gamma(N)$ is the Gamma function. To find $\mathrm{SNR_{thresh}}$, x is set to $1 + \mathrm{SNR}$.

Once $\mathrm{SNR_{thresh}}$ is found, Equation 2.26 is solved again. This time it is solved for P by setting $x = (1 + \mathrm{SNR_{thresh}})/(1 + \mathrm{SNR})$. This gives a detection probability, P_{det}, for each star.

This detection probability can be calculated in one of two ways. Firstly, it can be calculated by assuming that there is no mode damping from the δ-Scuti Instability strip. Mode damping can be suppressed by setting the β coefficient in equation 2.12 to 1. This is known as the 'fixed $\beta = 1$ case'. Hereafter is it denoted with $P_{\mathrm{det}}\ (\beta = 1)$.

Alternatively, P_{det} can be calculated by assuming that the mode damping is well described by equation 2.12. The detection probability calculated by assuming this is known as the 'variable β case', P_{det} (varied β). The difference between these two detection probabilities is discussed in Section 2.5.3.

So far in this Chapter, the stellar catalogues used to generate the ATL have been described. These catalogues were reduced to only keep solar-like stars. Afterwards, a detection probability was calculated for every star. The remainder of this Chapter is devoted to the methods used to rank the stars, and the tests that were performed to ensure that properties given in the ATL are robust. After the stars were ranked and the method used to generate the ATL was checked, the target list was ready to be sent to the TESS science team.

2.5 Ranking stars in the ATL

In the last section, a detection probability was calculated for every star in the ATL. After this, the stars in the ATL needed to be ranked. The first method used to rank stars involved separating them into regions of the Hertzsprung-Russell (HR) diagram, in order to ensure even coverage in effective temperature / luminosity space. This method was then improved using Kernel Density Estimation. Using a KDE ensured even more uniform coverage across the HR diagram.

Thirdly, a mixture of detection probabilities was used to rank the stars. This method also ensured even coverage across effective temperature / luminosity space, while maximising the number of stars that TESS would observe solar-like oscillations in. Before this is described, the 'regions' method of ranking stars is explained.

2.5.1 Regions of the Hertzsprung-Russell Diagram

While ranking stars in the the ATL, it was important to ensure that stars from across the HR diagram were chosen. This ensures that TESS will provide asteroseismic data for main sequence, subgiant and red giant stars, across a range of masses. Furthermore, selecting stars up to the δ-Scuti Instability strip will improve our understanding of the pressure-mode damping that occurs in this region of effective temperature / luminosity space.

One way of ensuring coverage across $T_{\text{eff}}\,/\,L$ space is to split the HR diagram up into different regions. The number of stars in each region of the HR diagram shows how well sampled the list is there. These numbers can be adjusted to ensure even coverage across the HR diagram.

To do this, the stars in the ATL were separated into 3 regions in $T_{\text{eff}}\,/\,L$ space, see Figure 2.6. These regions were for: main sequence and subgiant stars, low luminosity red giant branch stars, and stars which lie close to the δ-Scuti Instability strip. An example of what this looks like for the ATL is shown in Figure 2.7.

Figure 2.7 shows low luminosity red giant branch (RGB) stars in Region 2 (the blue

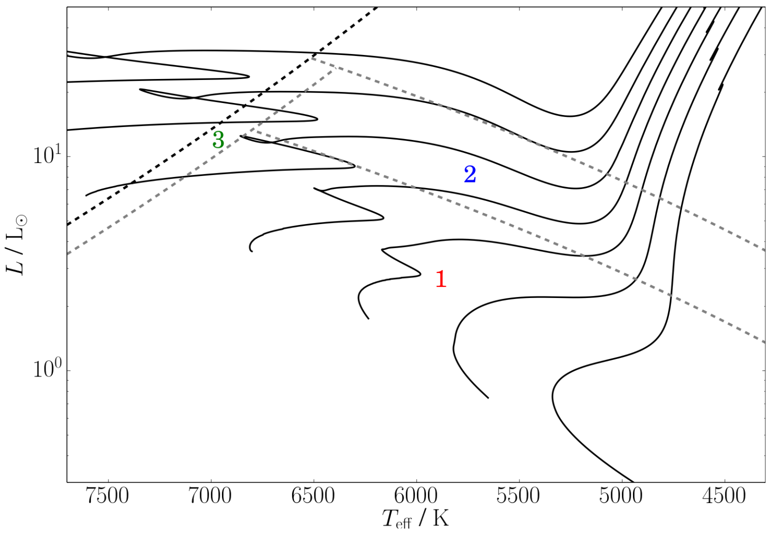


Figure 2.6: Different regions of the HR diagram. Evolutionary tracks with masses ranging from 0.8 - 2 $M_\odot$ have been plotted (solid black lines). The grey dashed lines separate different regions in $T_{\rm eff}$ /L space. Region 1 contains main-sequence and subgiant branch stars. Region 2 is for low luminosity red giant branch stars. Region 3 contains stars that lie close to the δ-Scuti instability strip (shown with a black dashed line).

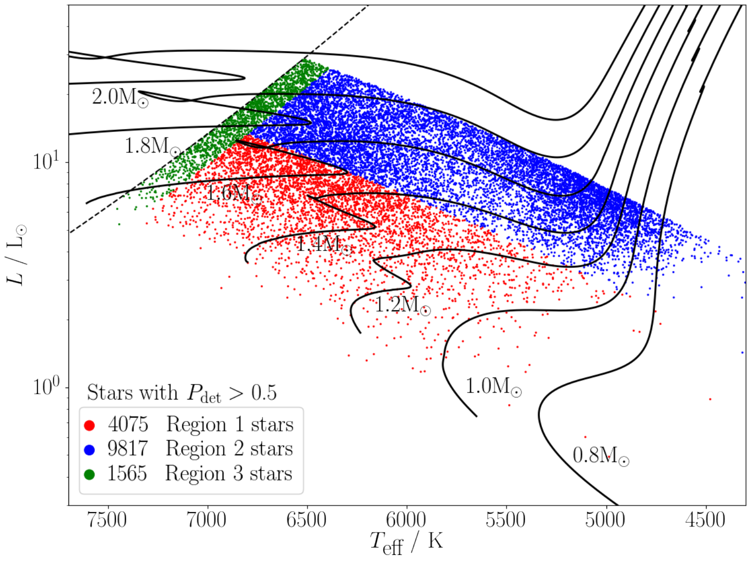


Figure 2.7: The ATL stars with a detection probability above 50 %, after being separated into 3 regions. Region 1 (red points) contains main-sequence and subgiant branch stars. Region 2 (blue points) is for low luminosity red giant branch stars. Region 3 (green points) contains stars that lie close to the δ-Scuti instability strip (black dashed line).

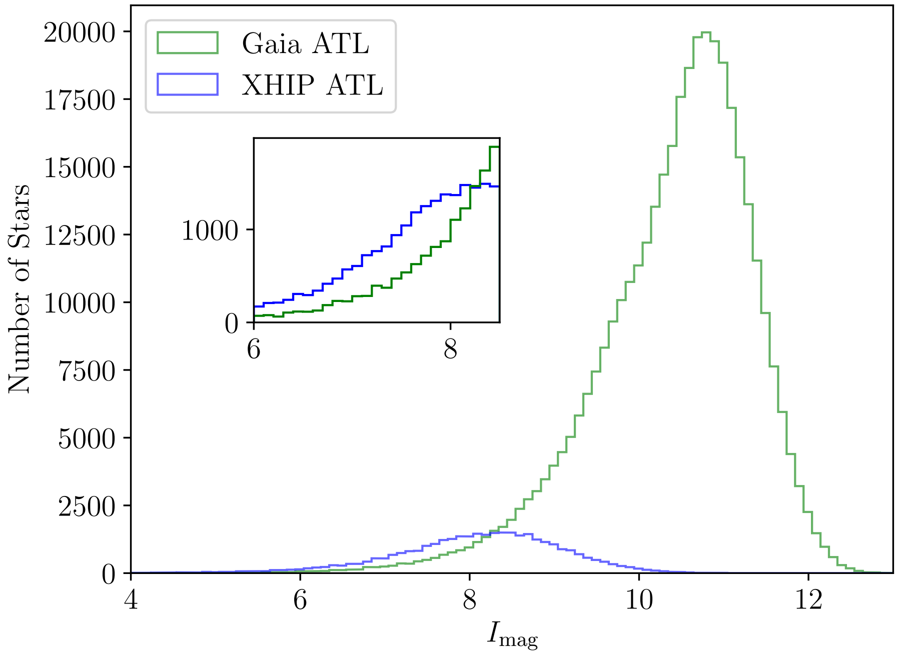


Figure 2.8: The $I_{\rm mag}$ distribution of the stars in the XHIP and DR2 catalogues. Shown before $P_{\rm det}$ was calculated for each star. The insert shows the $I_{\rm mag}$ region where the two catalogues overlap.

datapoints). Stars in this region had a tendency to be over-sampled when selecting targets for the ATL based on detection probability alone. The detection probabilities of RGB stars are (on average) higher than those of main sequence stars. This is because RGB stars have intrinsically larger solar-like mode amplitudes, compared to main-sequence stars.

As a main sequence star evolves into an RGB star, the radius and bolometric luminosity of the star will increase (Kjeldsen & Bedding, 1995). Since the amplitude of solar-like oscillations $A_{\rm osc}$ is proportional to $L \,/\, M$, if the bolometric luminosity of the star increases, so will the oscillation amplitude (assuming that the mass M of the star remains constant throughout evolution; Samadi et al. 2012). Larger oscillation amplitudes in RGB stars will lead to higher predicted detection probabilities.

As well as RGB stars having larger oscillation amplitudes, there is another reason why they could be over-sampled in the ATL. Using the DR2 catalogue as part of the ATL has introduced a very large number of fainter ($I_{\rm mag} \geqslant 8$) stars, see Figures 2.8 and 2.9.

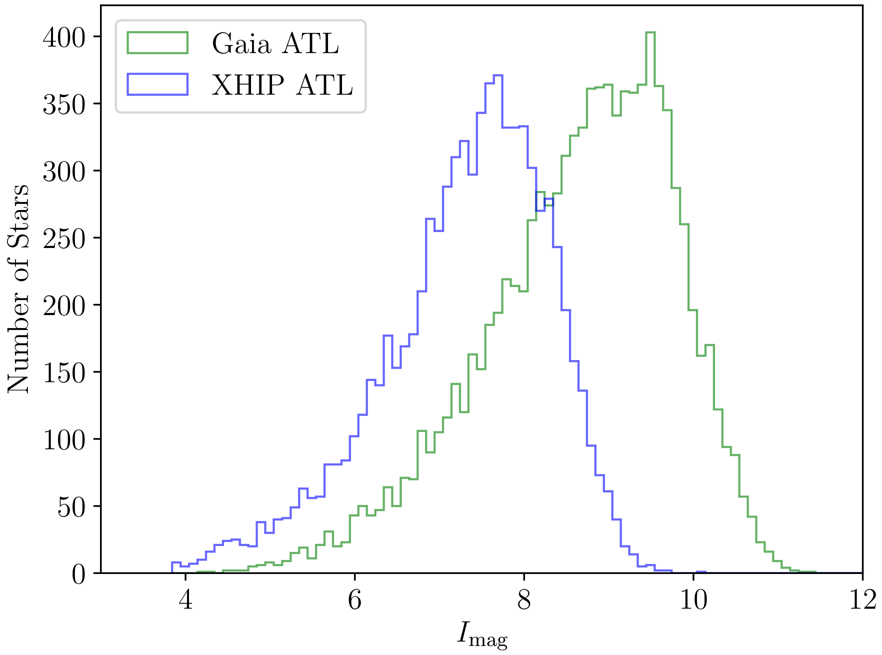


Figure 2.9: The I_{mag} distribution of the stars in the XHIP and DR2 catalogues that have a detection probability $P_{\mathrm{det}} \geqslant 0.5$.

Stars at magnitudes fainter than $I_{\mathrm{mag}} \approx 8$ will only have detected oscillations with TESS if they have evolved off of the main sequence (Campante et al., 2016). Since DR2 has introduced a lot of fainter stars into the ATL, there will be an extra bias towards RGB stars.

The boundaries of the three regions needed to be fixed in T_{eff} / L space with this bias in mind. Firstly, the upper ν_{max} limit of Region 2 was set at 240 μHz. This was selected as the point at which solar-like modes in Full Frame Image (FFI) stars would be too heavily apodized for pressure-mode oscillations to be detected. This is because the Full Frame cadence of observation with TESS is 30 minutes. This corresponds to a ν_{Nyq} frequency of 278 μHz, which would cause heavy apodization to stars with ν_{max} values above 240 μHz (equation 2.21).

After this, the boundary of the main sequence region was also defined. The lower ν_{max} cut of Region 2 at 600 μHz was used to separate low-luminosity red giant branch stars from subgiant stars. This limit was set ensure that main sequence stars would not be under-sampled when stars were selected using these three areas.

In addition to this, hotter stars could be biased against if a selection was made purely on detection probability. These 'hotter stars' lie close to the δ-Scuti instability strip (Region 3 in Figure 2.6). The β factor describes this mode damping of oscillations close the instability strip (equation 2.12). It was also important to keep this in mind when ranking the ATL, to ensure that a selection of hotter stars would be observed at a 2-minute cadence.

TESS short cadence imaging presents an excellent opportunity to gain understanding about the region close to the δ-Scuti instability strip. This is a typically under-sampled area of the HR diagram. Therefore, it was decided to separate the hotter part of T_{eff} / L space into a separate region; Region 3. This ensured that stars in this area were not under-sampled in the ATL. The 'cool edge' of Region 3 was chosen as the point at which solar-like oscillations were being damped too heavily for detections to be made. It lies parallel to the δ-Scuti instability strip.

In order to prevent a bias towards RGB stars or away from hotter stars, stars were first ranked separately in regions 1, 2 and 3. In this way, stars could be selected separately from the different regions of the HR diagram to ensure uniformity across T_{eff} / L space.

Stars could be ranked using observed parameters, such as parallax π or apparent magnitude I_{mag}, or inferred parameters such as detection probability P_{det} or signal to noise ratio. A ranking system based upon I_{mag} would cause the brightest stars to be ranked highest, while a P_{det}-based ranking system would bias towards stars where pressure-modes are most likely to be detected.

It is very likely that there is a correlation between an I_{mag}-based rank and a P_{det}-based rank. This is because brighter stars will have a smaller amount of instrumental noise, compared to their amplitude of oscillation (equation 2.18). If this correlation is strong, the I_{mag} Rank could be used to select targets for the ATL, while still retaining a large number of stars with detectable oscillations. I_{mag}-based ranking system would be more easily reproducible and requires fewer assumptions.

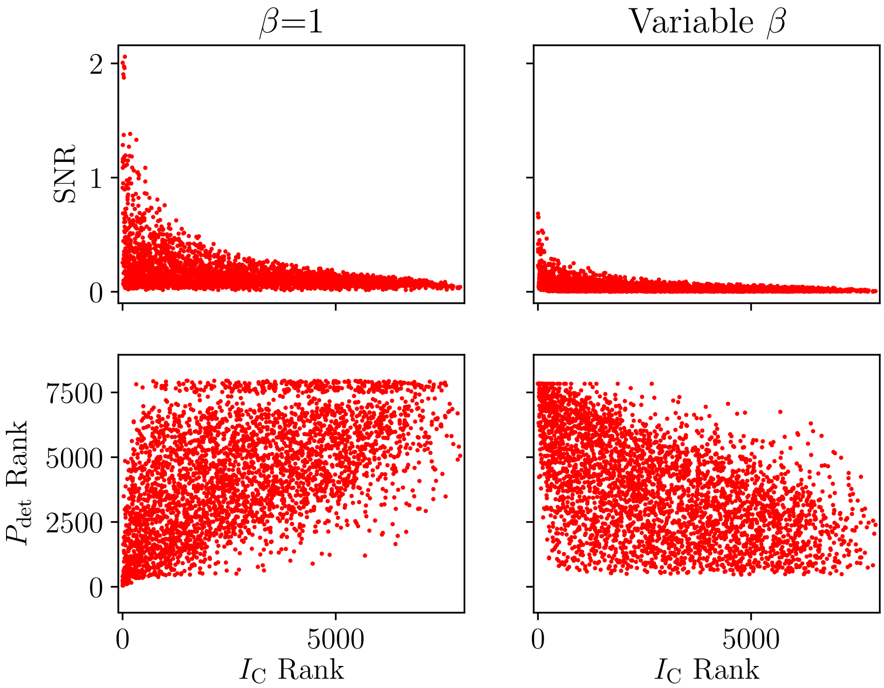


Figure 2.10: The relationship between ranking systems based upon $P_{\rm det}$, SNR and $I_{\rm mag}$, for the main-sequence stars in Region 1. The left panels show the comparison for the case were there is no damping due to modes near the $\delta-$Scuti instability strip. The panels on the right include this damping.

To determine if a $I_{\rm mag}$ - based rank could replicate a $P_{\rm det}$ - based rank, stars in the ATL were compared after they had been ranked using different properties. Figures 2.10 and 2.11 show $P_{\rm det}$ and SNR Ranks as a function of $I_{\rm mag}$ Rank, in Regions 1 and 2 respectively. This was done with the β damping factor turned off (left panels) and on (right panels).

The comparison shows that a ranking system based upon $P_{\rm det}$ cannot be replicated using $I_{\rm mag}$. There is more of a correlation for the stars in Region 2, but there is still a large spread between the two sets of ranks. This is due to the $P_{\rm det}$ Rank also taking into account the effects of effective temperature, luminosity, and observing time.

In a similar way to $I_{\rm mag}$ and signal to noise ratio, a rank based upon $T_{\rm eff}$ was also investigated. Figure 2.12 shows a comparison between an effective temperature-based rank and a $P_{\rm det}$-based rank, for the stars in Region 1 of the ATL. Although a correlation is seen between $P_{\rm det}$ and $T_{\rm eff}$, there is a very large amount of scatter. This shows that a ranking system based upon $P_{\rm det}$ could not be well replicated using effective temperature.

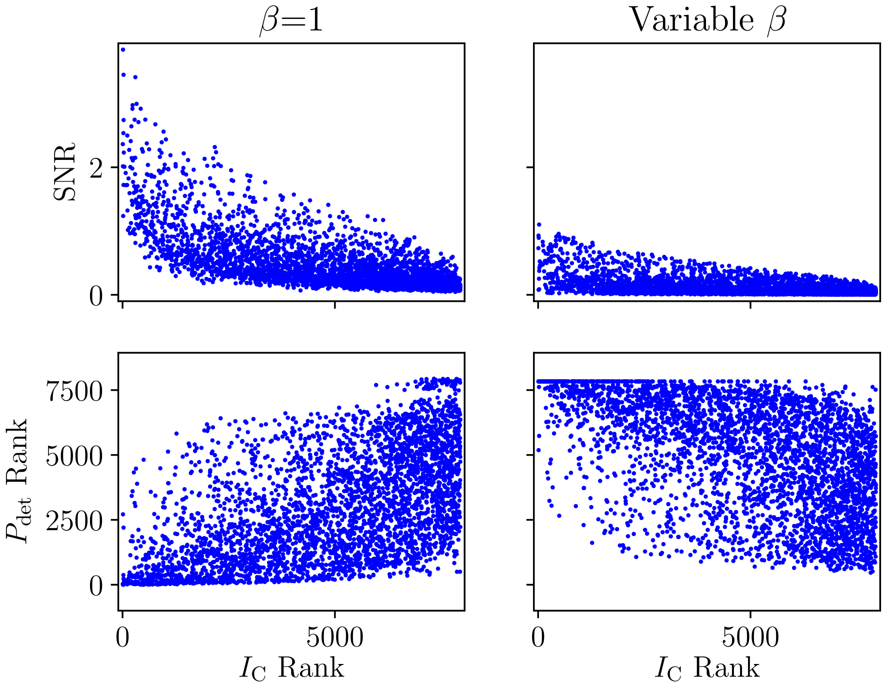


Figure 2.11: The same as Figure 2.10, but for the RGB stars in Region 2.

These results show that a rank based upon P_{det} cannot be reproduced using other parameters such as apparent magnitude, signal-to-noise ratio and effective temperature. Because of this, stars in the ATL will therefore be ranked according to P_{det} for the remainder of this Chapter.

Section 2.5.2 will go on to discuss another method of ranking targets using detection probabilities. This approach does not involve separating the HR diagram into different regions. Instead, Kernel Density Estimation was used to select stars from across T_{eff} / L space. KDEs can be finely-tuned to select stars more evenly across the HR diagram, or with a bias towards selecting stars with the highest detection probabilities.

2.5.2 Ranking targets using Kernel Density Estimation

Section 2.5.1 discussed separating the ATL into 3 regions in the HR diagram. This separation ensured coverage of solar-like asteroseismic targets across T_{eff} / L space.

There is an arguably more elegant way to ensure coverage across all T_{eff} / L space of

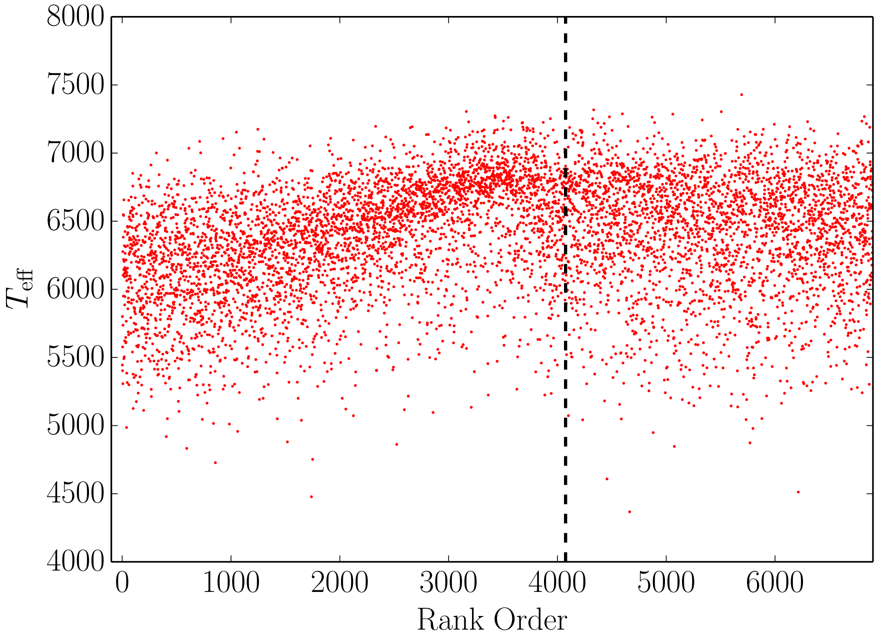


Figure 2.12: Effective temperature, as a function of $P_{\rm det}$-based rank order for the stars in region 1. Stars with lower effective temperatures tend to have higher ranks. However, this correlation is not strong enough to for a $T_{\rm eff}$-based ranking system to be relied upon when selecting asteroseismic targets for TESS.

interest in the ATL. This can be done by calculating the density of the ATL stars across the HR diagram. Stars in less dense regions of the HR diagram can then be chosen ahead of stars in very dense regions, guaranteeing even coverage in $T_{\rm eff}$ / L space.

Density estimates were made of the $T_{\rm eff}$ and L values for every star in the ATL. To estimate the density of the data, they first needed to be binned. Stars in the ATL could be binned discretely (with a histogram; Phillips & Tai 2017) or continuously (using Kernel Density Estimation; Scott 1979). This difference is illustrated in Figure 2.13.

Figure 2.13 shows that KDEs are better than histograms at describing the density of a continuous variable. Our data is continuous; stars can have any $T_{\rm eff}$ or L value. Therefore, the density of our data was estimated using a KDE.

Kernel Density Estimation was applied to the $T_{\rm eff}$ and L values of the ATL stars. A visual representation of this is shown in Figure 2.14. It shows that selecting stars for the ATL based upon $P_{\rm det}$ ($\beta = 1$) alone would result in a large number of sub-giant branch

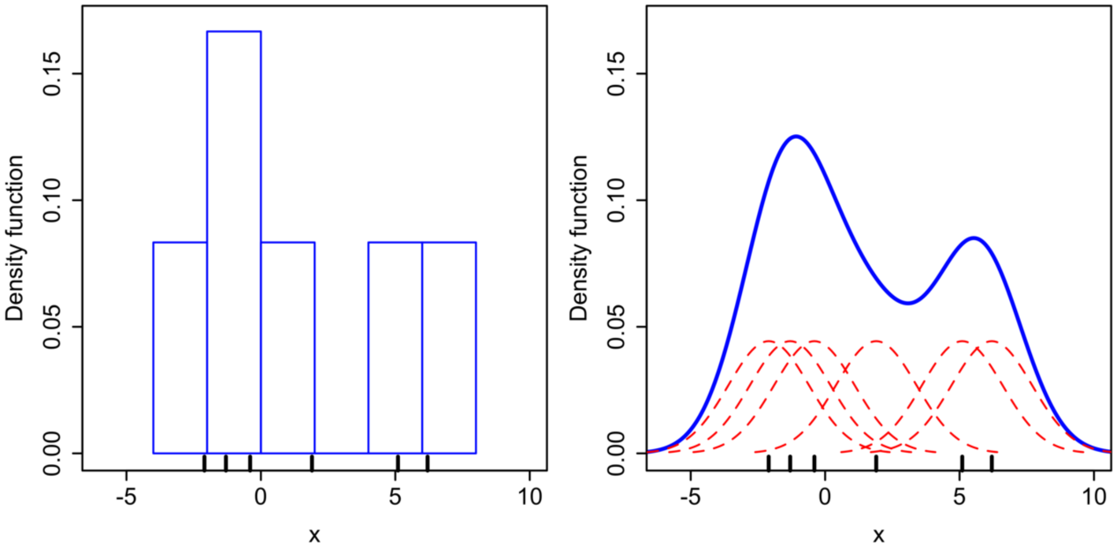


Figure 2.13: The difference between histograms and KDEs[3]. 6 data points (black lines) are shown plotted as a histogram (left). The posterior density function of a KDE is also shown for the same datapoints (right). In the histogram, each data point is treated as a discrete value within a bin. In the KDE, each data point is treated as a normal distribution (the red dashed lines). In both figures, the distributions are summed to give the Probability Density Functions shown in blue.

and low-luminosity red giant branch stars, and fewer main sequence targets.

As well as displaying a KDE on a single HR diagram (as in Figure 2.14), it is helpful to show how the distribution of stars varies throughout the ATL as a function of P_{det} value. A clear way of doing this is by plotting HR diagrams of the ATL in 2000-star chunks (Figures 2.15 and 2.16).

Kernel Density Estimation alone can be used to rank the stars in the ATL. A random uniform deviate x between 0 and 1 can be assigned to each star. If the (normalised) KDE value of the star is greater than x, the star is selected for observation. If the KDE $\leqslant x$, the star is not selected. Selecting stars this way would result in the smoothest possible distribution of stars in T_{eff} / L space.

This is not the only way that Kernel Densely Estimation could be used to make the ATL more uniformly populated in the HR diagram. The KDE-rank could be used alongside the P_{det}-rank. This could produce a target list where the stars with the highest ranks were more likely to have solar-like oscillations, and also existed in less dense areas

---

[3] This figure is from https://en.wikipedia.org/wiki/Kernel_density_estimation.

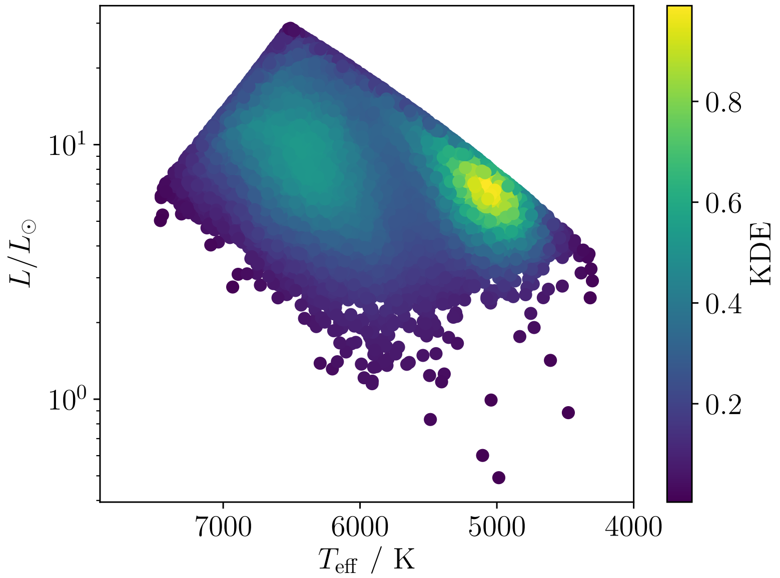


Figure 2.14: A visual representation of Kernel Density Estimation on a HR plot. It shows the highest ranked 25,000 stars, according to a $P_{\rm det}\,(\beta = 1)$ -based ranking system (where the stars with the highest $P_{\rm det}\,(\beta = 1)$ rank appear at the top of the ATL). The region of highest density is around the low-luminosity red giant branch.

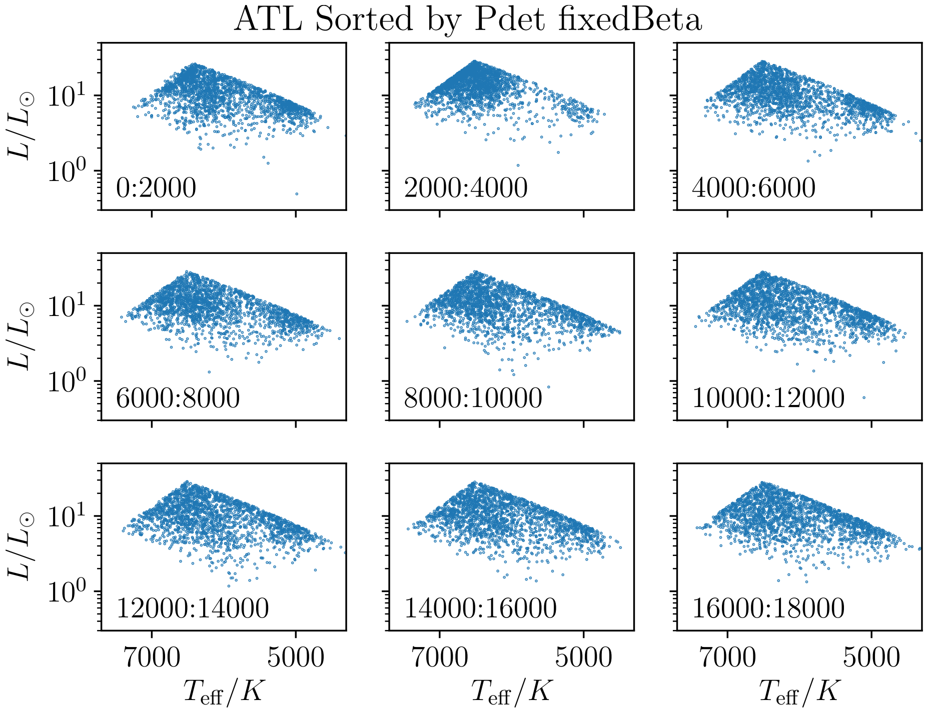


Figure 2.15: Subplots showing HR diagrams of the ATL, ranked by $P_{\rm det}\,(\beta = 1)$. Each subplot shows 2000 stars. The range of ranks is shown in the bottom left corner of each subplot. The Figure only contains stars with $P_{\rm det}\,(\beta = 1)\ > 0.5$.

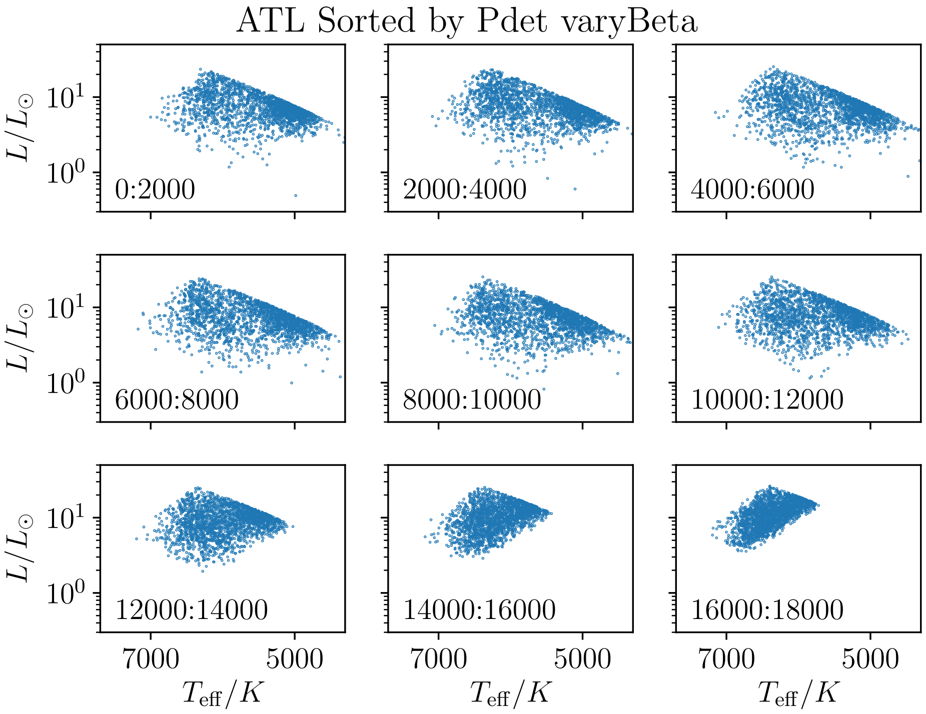


Figure 2.16: Subplots showing HR diagrams of the ATL, ranked by $P_{\rm det}$ (varied β) . Each subplot shows 2000 stars. The range of displayed ranks is shown in the bottom left corner of each subplot. The Figure only contains stars with $P_{\rm det}$ (varied β) > 0.5.

of the HR diagram.

To test this concept, Figure 2.17 was made. Figure 2.17 shows a comparison between several ranking systems. Firstly, it shows an unsmoothed ‘clumpy’ ranking system in black (when $P_{\rm det}$ $(\beta = 1)$ is used to rank stars without a KDE),

Figure 2.17 also shows the ‘smoothest’ possible ranking system, using only a KDE to rank stars. Thirdly, the Figure shows a mixture of ‘smooth’ and ‘clumpy’ ranking systems. The figure shows that the smoothest possible ranking system (the blue line) is far from the ideal way of selecting asteroseismic stars for the ATL; the $P_{\rm det}$ values of the highest-ranked stars decrease much faster using a ‘smooth’ ranking system, compared to a ‘clumpy’ ranking system based purely upon detection probability (the black line in Figure 2.17).

Using a mixture of KDE-rank and $P_{\rm det}$-rank is closer to ideal. This is shown by the green line in Figure 2.17. The plot shows that this is a better way of making the Astero-

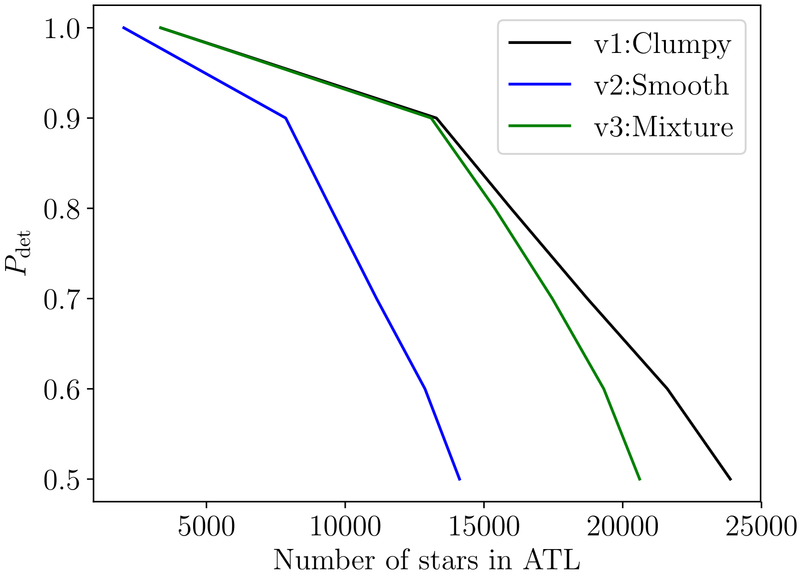


Figure 2.17: A comparison between an unsmoothed 'Clumpy' ranking system (such as P_{det} $(\beta = 1)$ in black), the smoothest possible ranking system (in blue), and a mixture of the two (in green). The lines have been plotted as a function of P_{det} $(\beta = 1)$. Every star in the list with $P_{\mathrm{det}} > 0.5$ is included.

seismic Target List selection function smoother in T_{eff} / L space, while ensuring that the number of stars with detectable solar-like oscillations is still high. Unfortunately, Figure 2.17 highlights the fact that there would be a sacrifice of about 3000 potentially detectable asteroseismic targets from the ATL if the 'mixture' ranking system was adopted, rather than a system based purely upon P_{det}.

The primary goal of the ATL is to select the most stars with solar-like oscillations for TESS to observe. Therefore, the cost of ∼3000 stars with potentially detectable solar-like oscillations was deemed too high to justify making the target list more uniform using a KDE (as seen by the difference between the green and black lines in Figure 2.17).

After Kernel Density Estimation was tested, it was found that increasing the coverage of the ATL across the HR diagram caused thousands of stars with detectable solar-like oscillations to be removed from the target list. The next section will go on to discuss how both P_{det} $(\beta = 1)$ and P_{det} (varied β) can be used to rank stars in the ATL together. This keeps the priority of the ATL on detecting solar-like modes of oscillation, while ensuring coverage across effective temperature / luminosity space.

2.5.3 Ranking targets using a mixture of detection probabilities

Section 2.5.1 reviewed how the HR diagram could be separated into three regions, in order to rank stars in the ATL evenly across $T_{\rm eff}$ / L space. Section 2.5.2 evolved this idea by using Kernel Density Estimation to rank stars. This section will present the final method used to rank targets in the ATL. This uses a mixture of detection probabilities, which include and exclude the mode damping factor β.

In Section 2.4, the probability of detecting solar-like oscillations within a star was calculated. This was done under the assumption that there is no mode damping as stars approach the δ-Scuti instability strip ($P_{\rm det}$ $(\beta = 1)$). $P_{\rm det}$ $(\beta = 1)$ therefore provides an overestimate of the detection probability for stars close to the instability strip. In addition to $P_{\rm det}$ $(\beta = 1)$, a detection probability was also calculated for every star assuming that modes are damped close to the instability strip (denoted by $P_{\rm det}$ (varied β)).

When ranking targets in the ATL, it is important to consider the types of physics we would like to test using these Short Cadence stars. One of the benefits to having up to 20,000 2-minute cadence observation slots with TESS is the ability to select a range of stars to test stellar physics across the HR diagram.

Stellar evolution is one area of stellar physics that can be tested. This is done by selecting a range of stars at different stages of evolution; main-sequence, subgiant branch and red giant branch stars can all be selected in the ATL. The change in the acoustic profile of a star can be seen by selecting stars at these different evolutionary stages.

There has already been a great deal of work on stellar evolution using asteroseismology (e.g Chaplin & Miglio 2013). An area of stellar physics which has not been tested as thoroughly is the difference between stars exhibiting solar-like modes of oscillation, and stars showing detectable κ-driven oscillations (in the δ-Scuti instability strip). This is the physical difference between stars that do have surface convection (solar-like oscillators), and those that do not.

Using $P_{\rm det}\,(\beta = 1)$ in the ranking metric will include stars in the ATL which lie up to the edge of the δ-Scuti instability strip. Including this range of stars will lead to a greater understanding of mode damping across $T_{\rm eff}$ / L space.

For this reason, the last ranking method to be tested involved both $P_{\rm det}\,(\beta = 1)$ and $P_{\rm det}$ (varied β). On its own, $P_{\rm det}\,(\beta = 1)$ selected too many hot stars close to the δ-Scuti instability strip (and not enough solar-like oscillators). One way of avoiding this was to include $P_{\rm det}$ (varied β) when ranking stars. Stars with a low β value will have heavily damped oscillations (equation 2.11), so will have lower detection probabilities.

The $P_{\rm det}\,(\beta = 1)$ and $P_{\rm det}$ (varied β) ranks were merged using equation 2.27. After merging, the combined detection probability was called $P_{\rm mix}$. Once $P_{\rm mix}$ was calculated, stars can simply be ranked from highest to lowest combined detection probability value.

$$P_{\rm mix} = (1 - \alpha) \times P_{\rm det}\,(\text{varied } \beta) + \alpha \times \mathrm{P}_{\rm det}\,(\beta = 1) \tag{2.27}$$

The α parameter allows $P_{\rm mix}$ to be tuned, depending on the number of 2-minute cadence target slots that should be allocated to test mode damping in stars close to the δ-Scuti instability strip. When ranking stars, it was decided to set $\alpha = 0.5$. This will allow the effect of β mode-damping to be tested thoroughly with TESS, while ensuring that thousands of observation slots are dedicated to solar-like oscillators.

The effect of $P_{\rm mix}$ on the ATL is shown in Figure 2.18. It shows HR diagram subplots of the ATL in 2000-star chunks. At the very top of the list (at ranks 0:2000), there are fewer hot stars with effective temperatures of ∼6500 - 7500 K. This is because these stars have lower $P_{\rm det}$ (varied β) detection probabilities; they experience heavy damping from the β effect.

Figure 2.19 also shows HR diagram subplots of the ATL after ranking by $P_{\rm mix}$. This plot only includes stars where $V_{\rm mag} < 6.5$. Figure 2.19 provides another clear picture of the benefit of using the $P_{\rm mix}$ ranking metric; bright stars with detectable solar-like modes of oscillation across the HR diagram will be included in the ATL.

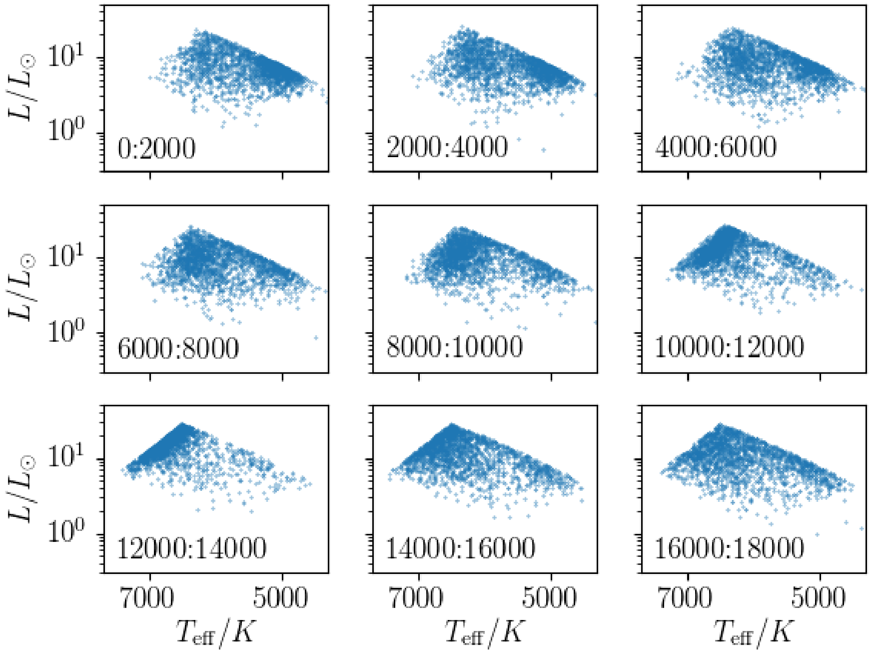


Figure 2.18: The highest ranked 18,000 stars using the $P_{\rm mix}$ metric. Each subplot shows 2000 stars. The range of $P_{\rm mix}$ ranks is shown in the bottom left corner of each subplot. There is a good distribution of stars across the HR diagram.

This Section has described three methods used to rank stars in the ATL. When ranking stars, the priority of selecting solar-like asteroseismic targets was tensioned against ensuring even coverage across the HR diagram. It was found that the $P_{\rm mix}$ ranking metric allowed the greatest number of stars with detectable oscillations to be ranked highest, whilst allowing for stars close to the δ-Scuti instability strip to be well observed by TESS.

After the stars in the 2-minute cadence ATL were ranked, a 'high priority' subset of the stars was chosen. This is a collection of stars to be observed at a cadence of 20-seconds. This cadence is reserved for the highest priority stars in the ATL.

2.6 Selecting high priority stars

As well as being allocated 20,000 2-minute cadence slots for asteroseismology, a number of 20-second cadence slots were set aside. 1500 20-second cadence slots had been assigned to asteroseismology. This section describes how these slots were allocated. During the creation of the ATL, these 20-second cadence slots were believed to be available, but it is

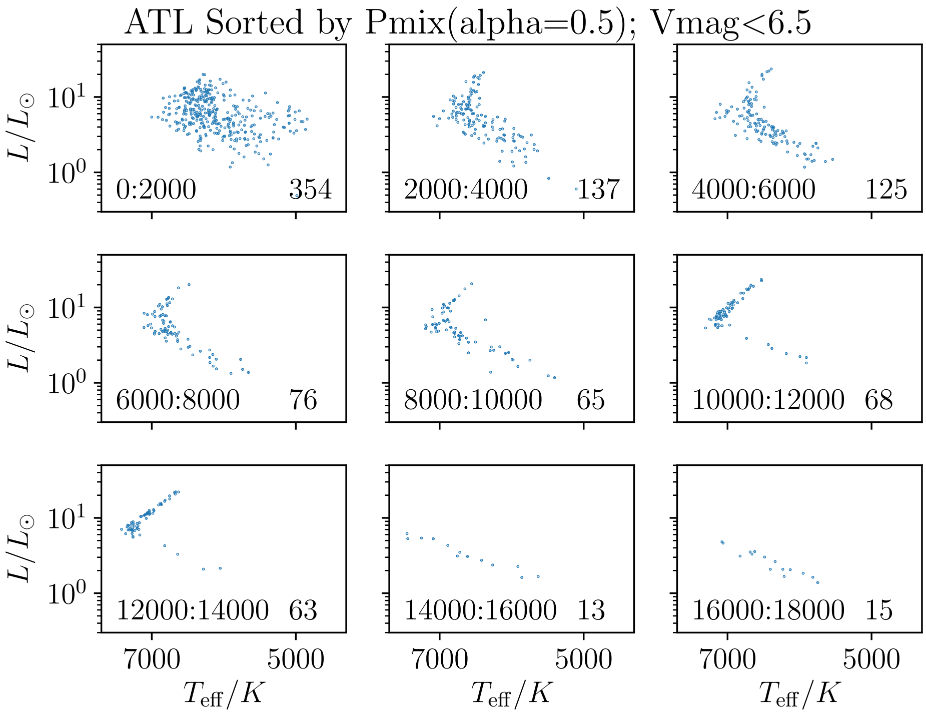


Figure 2.19: The highest ranked 18,000 stars using the $P_{\rm mix}$ metric with $V_{\rm mag} < 6.5$. The range of $P_{\rm mix}$ ranks is shown in the bottom left corner of each subplot. The number on the bottom right corner of each subplots shows the number of $V_{\rm mag} < 6.5$ stars in each range of ranks.

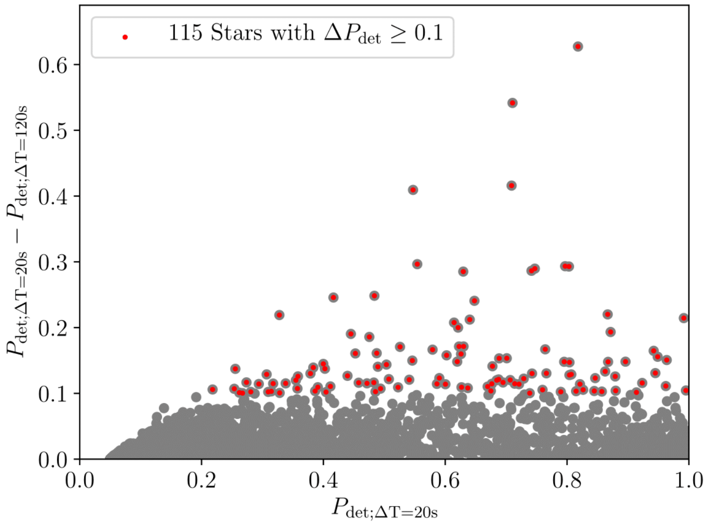


Figure 2.20: The difference between 20-second and 120-second cadence detection probabilities for the down-selected ATL. Stars with a detection probability difference greater than 0.1 are shown in red.

currently unclear that these slots will be available during the nominal TESS mission.

Short cadence observation slots are valuable tools for asteroseismologists. Decreasing the cadence of observation for solar-like stars will reduce the effect of apodization on the signal; the Nyquist frequency will not damp oscillations observed at a faster cadence as heavily (c.f. Section 2.4.3). Detections of solar-like oscillations will therefore become more likely.

With this effect in mind, there are several ways that stars could be selected for 20-second cadence observation. One option is to calculate the detection probabilities with 2-minute and 20-second cadences for every star in the ATL. The stars with the greatest difference in P_{det} could then be put onto the 20-second cadence list.

The detection probability difference, ΔP_{det}, between cadences for these stars is shown in Figure 2.20. The difference is due to the different levels of apodization between 20-second and 120-second cadence observations. Since there are few stars that this would make an appreciable difference for, 20-second cadence targets were not chosen using ΔP_{det}.

Another option is to select stars based upon both P_{det} and ν_{max} value. This way, the

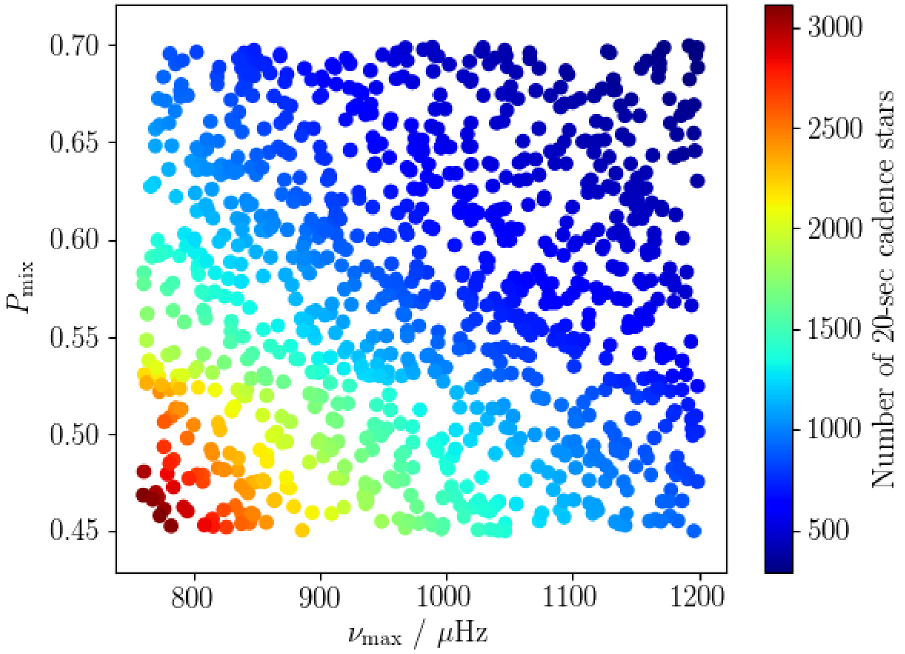


Figure 2.21: P_{mix} as a function of ν_{max}. The colourbar shows the number of stars that would be in the 20-second list, if this set of P_{mix} and ν_{max} cuts were used to select stars from the 120-second ATL.

main sequence stars with the highest detection probabilities can have photometry which is least affected by apodization. Figure 2.21 shows a set of 1000 ν_{max} and P_{mix} values drawn from uniform distributions. For each pair of ν_{max} and P_{mix} values, the number of stars in the ATL is shown by the colourbar.

Based on the results in Figure 2.21, ATL stars with a predicted $\nu_{\mathrm{max}} \geqslant 950\mu$Hz and $P_{\mathrm{det}} \geqslant 0.52$ were chosen as 20-second cadence targets. Figure 2.22 shows an HR diagram of these 1401 stars. This number of stars was chosen because in total, 1500 20-second cadence slots were allocated to Working Groups 0-8 for asteroseismology. The ATL comprises the target lists from Working Groups 1 and 2. A cut of $\nu_{\mathrm{max}} \geqslant 950\mu$Hz and $P_{\mathrm{det}} \geqslant 0.52$ was settled upon as the way to sample 20-second targets in the ATL.

So far in this Chapter, the method used to select and rank stars for the 120-second and 20-second asteroseismic target lists has been explained. It is now important to check that this method used to select stars was robust. Are the calculated stellar parameters

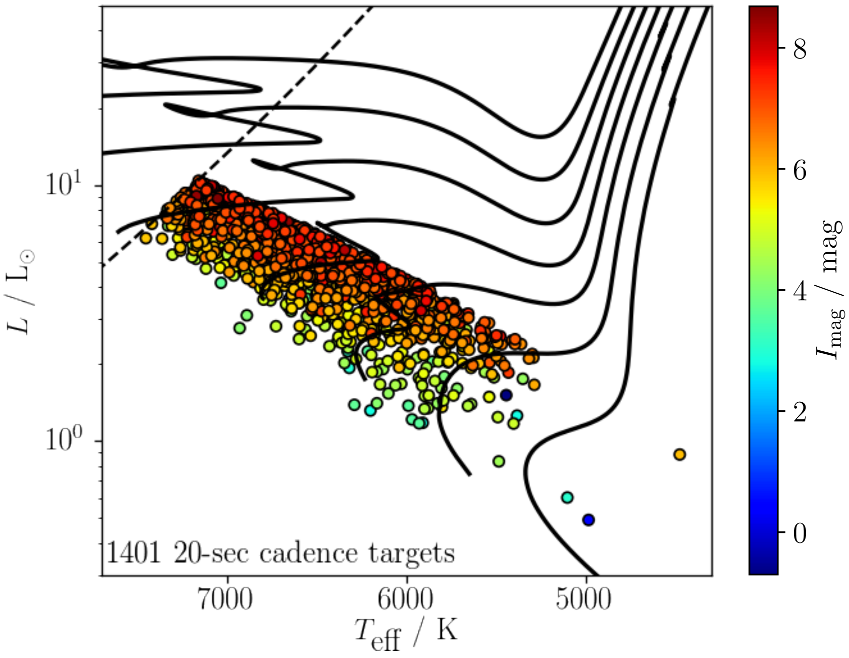


Figure 2.22: An HR diagram of the 20-second cadence targets from the ATL. These are the stars with a predicted $\nu_{\rm max} \geqslant 950\mu$Hz and $P_{\rm det} \geqslant 0.52$.

and detection probabilities reliable?

2.7 Monte Carlo trials

Stars were selected for the ATL based upon detection probability $P_{\rm det}$. A detection probability was calculated for every star in the ATL in Section 2.4. Monte Carlo trials were used to test the robustness of the method used to calculate these detection probabilities. These tests are explained here.

Before the Monte Carlo method is described, it is important to understand how errors propagate in the ATL. The global seismic parameters $\nu_{\rm max}$ and $\Delta\nu$ are calculated using scaling relations that depend on radius and temperature (equations 2.9 and 2.10). In these equations, radius is raised to a higher power than $T_{\rm eff}$. Radius therefore has a larger influence on the predicted $\Delta\nu$ and $\nu_{\rm max}$ values than effective temperature. Uncertainty introduced in radius will have a larger effect than $T_{\rm eff}$ on the seismic parameters.

Stellar radii were calculated with luminosity and effective temperature using the

Stefan-Boltzmann law. To do this, stellar luminosities were first calculated using equation 2.4. This relies on the Bolometric Correction, reddening and V_{mag}. In addition, T_{eff} is susceptible to changes in (B-V) colour.

The aim of this section is to estimate the standard deviation of the detection probabilities, $\sigma_{P_{\mathrm{det}}}$. As described above, there are a pyramid of equations used to calculate P_{det}. Calculating $\sigma_{P_{\mathrm{det}}}$ analytically would therefore be extremely difficult and prone to error. Because of this potential error, it is sensible to calculate $\sigma_{P_{\mathrm{det}}}$ with Monte Carlo trials.

If the detection probability algorithm is robust, small variations in the input stellar parameters (such as parallax and colour) will not drastically affect a star's P_{det} value. If the method is not robust, making small changes to the input parameters will cause a large scatter in the predicted P_{det} values. This Section describes how this scatter was calculated.

2.7.1 The Monte Carlo Method

The robustness of the method used to create the list of asteroseismic stars was tested. To do this, the input stellar parameters of the stars in the ATL were perturbed. The parameters (a) were perturbed using their factional or absolute uncertainties (σ_a). The uncertainties were then multiplied by a random number from a normal distribution (n) to scatter them. These 'scattered' uncertainties were added to the pristine parameters to perturb them:

$$a' = a + n\,\sigma_a\,. \tag{2.28}$$

The amount that the parameters were perturbed is given in Table 2.2.

The detection test was then re-run using these perturbed parameters a'. This process of perturbing the parameters and calculating P_{det} was repeated 1000 times for each star. This resulted in each star in the list having a distribution of 1000 P_{det} values. The scatter of these P_{det} values is a quantitative measure of how robust the detection test algorithm is.

Inputs	Outputs	Perturbation
$(B-V)$	T_{eff}	3% fractional T_{eff} uncertainty
$(B-V), V_{\mathrm{mag}}$	I_{C}	0.1 absolute I_{C} uncertainty
$V_{\mathrm{mag}}, A_{\mathrm{V}}, BC_{\mathrm{V}}, d$	L	$V_{\mathrm{mag}}: 0.02,\ A_{\mathrm{V}}: 0.03, BC_{\mathrm{V}}: 0.03.$
		Use actual distance (d) uncertainties.
T_{eff} , L	R	Exact equation; no perturbation
T_{eff} , R	$\Delta\nu$	1.8% fractional $\Delta\nu$ uncertainty
T_{eff} , R	ν_{max}	4% fractional ν_{max} uncertainty

Table 2.2: The amount the stellar parameters were perturbed in the Monte Carlo trials. ν_{max} and $\Delta\nu$ perturbations are from Campante et al. (2016).

2.7.2 Monte Carlo Results

For each of the 1000 Monte Carlo trials, a P_{det} value was obtained using perturbed stellar parameters. For each star, these 1000 detection probabilities were sorted from highest to lowest value.

The median detection probability from the trials is the 500th value. The standard deviation of the P_{det} values $\sigma_{P_{\mathrm{det}}}$ is half of the difference between the $1000 \times 0.67/2 = 335^{\mathrm{th}}$ value above and below the median.

To display the results, a histogram of $\sigma_{P_{\mathrm{det}}}$ values for all of the stars in the ATL is shown in Figure 2.23. This histogram shows the results of every star in the DR2 and XHIP catalogues, after non solar-like oscillators were removed (Section 2.4).

In Figure 2.24, ν_{max} is plotted against the difference between the upper 1-sigma P_{det} value from the MC trials, and the median P_{det} value from the trials. The Figure shows that the $\sigma_{P_{\mathrm{det}}}$ variation is larger for the stars with lower ν_{max} values. This is because these more evolved stars have larger intrinsic oscillation amplitudes. Altering the input parameters for these stars will cause larger variations in the oscillation amplitude, and so larger variations in the predicted P_{det} value.

By comparison, detections in brighter main sequence stars will be much more sensitive

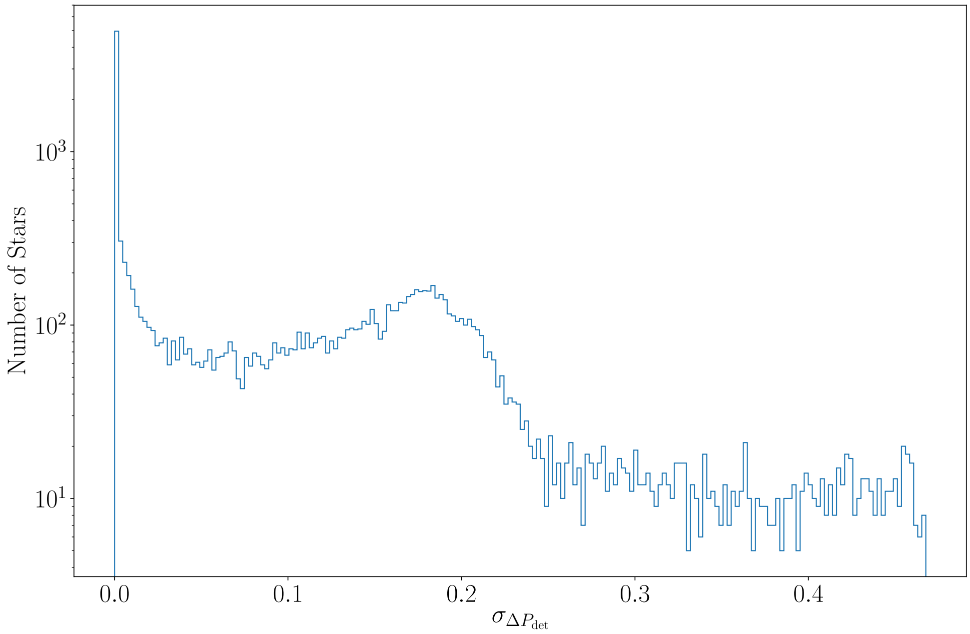


Figure 2.23: The scatter of the $P_{\rm det}$values after 1000 Monte Carlo trials for all the stars in the ATL, after parameters cuts (c.f Section 2.4).

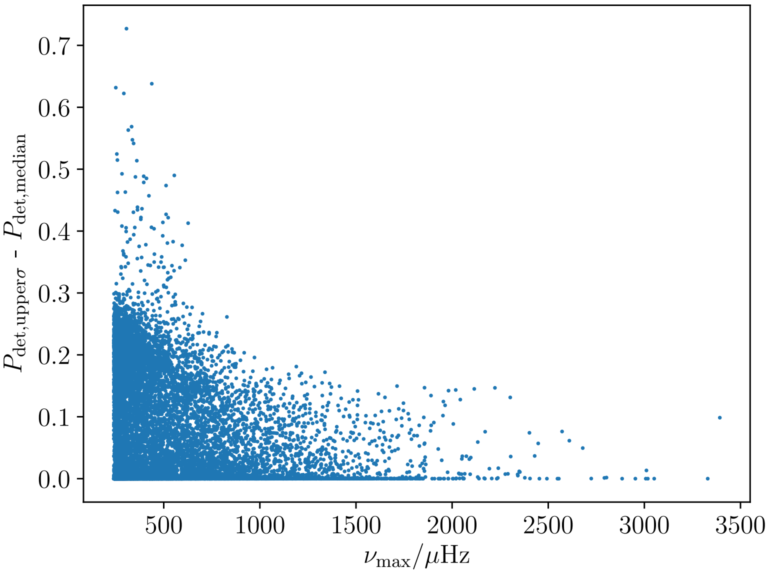


Figure 2.24: The difference between the upper 1-sigma $P_{\rm det}$ value from the trials, and the median $P_{\rm det}$ value from the trials, as a function of $\nu_{\rm max}$. There is more $\sigma_{P_{\rm det}}$ variation for stars with lower $\nu_{\rm max}$ values because more evolved stars have larger intrinsic oscillation amplitudes.

to a change in parameters. For example, if the apparent magnitude of an observed main sequence star is too faint, oscillations will not be detected at all with TESS.

The quantitative results of the robustness test are encouraging. $\sigma_{P_{\rm det}}$ was calculated for different subsets of the ATL. The results are:

- For the stars in the ATL with $P_{\rm det} > 0.5$: 9.8 %
- For the highest ranked 50,000 stars: 9.2 %
- For the highest ranked 10,000 stars: 8.3 %

To summarise this, for a given star in the ATL we can predict the $P_{\rm det}$ value to within 8 - 10 %. The method used to select stars for the ATL was therefore found to be reliable.

When calculating $\sigma_{P_{\rm det}}$, some of the calculated parameters were not scattered. Two of these include the total power due to oscillations $P_{\rm tot}$, and the power due to granulation $P_{\rm gran}$. No scatter was applied to these parameters because the uncertainty in the relations used to calculate them is not known. The results from the Monte Carlo trials performed here are therefore a conservative limit on the uncertainty of the detection test. Regardless of this, the standard deviation of the detection probabilities is reassuringly low. This leads to the conclusion that the method used to construct the ATL is robust.

Alongside confirming that the calculated detection probabilities were robust, the parameters that were calculated for each star were tested individually. These parameters included the stellar effective temperature and radius, alongside the global asteroseismic parameters $\Delta\nu$ and $\nu_{\rm max}$.

2.8 Checking individual stellar parameters

In the last section, Monte Carlo trials tested the robustness of the method used to calculate detection probabilities for every star in the ATL. In this section, these tests are expanded as individual stellar parameters are compared to literature values. In all of the comparison plots that follow, the black lines represent 1:1 lines; perfect agreement between the two datasets.

The first parameter to be tested is parallax. As this parameter is explicitly provided by the DR2 and XHIP catalogues, using precise values is fundamental to the accuracy of calculated detection probabilities.

2.8.1 Comparing parallax sources

Originally, the ATL was made with parallaxes from the TGAS catalogue. For a few bright stars with more precise values, parallaxes were instead taken from the XHIP catalogue. Where available, distances were taken from Astraatmadja & Bailer-Jones (2016). The TGAS catalogue was used because when the first version of the ATL was being constructed, this list provided the most precise parallaxes for stars across the sky.

In April 2018, a new *Gaia* data release was made available (DR2; Lindegren et al. 2018). DR2 promised to provide much more precise parallaxes than TGAS, for a much larger number of stars. Previously, TGAS provided information on $\sim$2,000,000 objects. By comparison, DR2 contains positions, proper motions and parallaxes for 1.6 *billion* stars. Using the same method as Astraatmadja & Bailer-Jones (2016) did for TGAS, distances for the DR2 stars have been inferred using Bayesian prior knowledge in Bailer-Jones et al. (2018).

The 'real' ATL presented throughout this Chapter uses the DR2 catalogue. For the tests in this section, an 'alternative' version of the ATL was made using the TGAS catalogue. The differences between the 'real' and 'alternative' target lists are compared. These differences are discussed here, and the effects are reviewed. In both versions of the ATL, parallaxes from the XHIP catalogue have been included to keep the brightest stars in the list.

Before the results from this DR2/TGAS comparison are shown, the potential effects are explained. Changing the sources of parallax and distance could affect the data in the following ways:

- Parallax is used to make rudimentary cuts to the stars in the ATL. The cuts remove stars with very uncertain parallaxes (with fractional values > 0.5), or negative parallaxes (as these are unphysical). If the parallaxes are incorrect, stars may be cut from the ATL unnecessarily.

- Distance is used to calculate the amount of reddening to apply to stellar magnitudes and colours. If distance is incorrectly inferred, the reddening coefficients may be wrong.
- An incorrect reddening value will lead to incorrect apparent magnitudes. If the apparent magnitude of the stars is incorrect, the wrong amount of instrumental and shot noise will be added to the signal of each star. This will alter the detection probabilities of the stars.
- The stellar luminosities depend on the distance to the star, which is determined from the parallax. An incorrect parallax will therefore lead to an incorrect luminosity.
- The effective temperatures of the stars use a polynomial with (B-V). If this colour is incorrect (because of incorrect reddening values), the effective temperatures will also be.
- Stellar radius is calculated from the Stefan-Boltzmann law. This uses both luminosity and effective temperature. As discussed, incorrect parallaxes and distances have the potential to perturb both the luminosity and effective temperature. This in turn could lead to poor estimates of the luminosities.
- The global seismic parameter $\nu_{\rm max}$ is calculated using a scaling relation involving both radius and effective temperature. If $\nu_{\rm max}$ is miscalculated, the predicted evolutionary states (and the amplitudes of the solar-like oscillations) for the ATL stars will be wrong.

These factors all contribute to the detection probabilities of the stars. Now that the effects have been discussed, the results of different parallax and distance sources on the stars in the ATL can be shown.

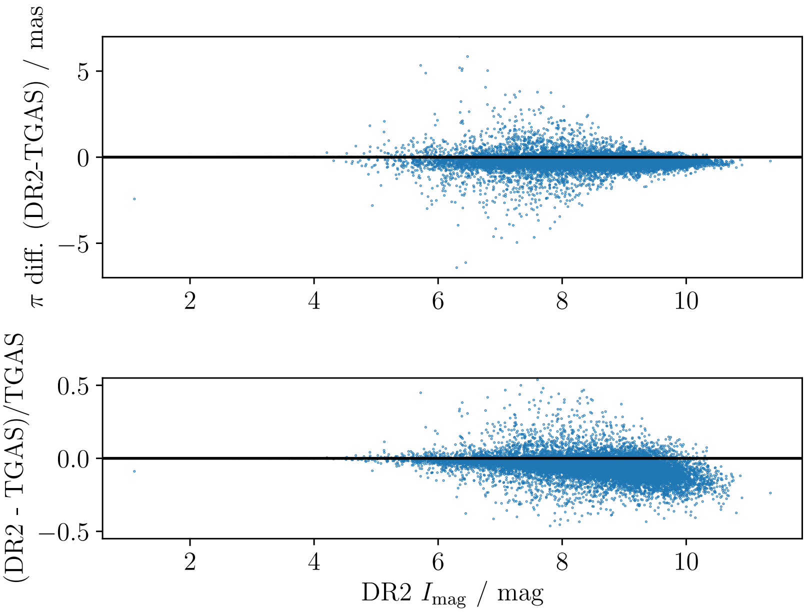


Figure 2.25: The parallax difference between the DR2 and TGAS catalogues, as a function of $I_{\rm mag}$.

In both the ‘real’ (using DR2) and ‘alternative’ (using TGAS) versions of the ATL, the brightest stars from the XHIP catalogue were added to the lists. The results of the comparison have been shown as a function of the apparent I-band magnitude of the stars for the highest ranked stars in the ‘real’ ATL.

Figure 2.25 shows the parallax difference between DR2 and TGAS, for the stars at the top of the ‘real’ ATL (with ranks $< 25,000$). There is a spread of $\sim 2\,$mas between the values in the astrometric catalogues. DR2 measures smaller parallaxes for the majority of the stars, particularly for the fainter targets in the ATL.

These stellar parallaxes are used to calculate luminosities for the stars in the ATL. Figure 2.26 shows differences between stellar luminosities calculated using parallaxes from the DR2 and TGAS catalogues. Calculating stellar luminosities incorrectly will cause stars to become misclassified; subgiant branch stars and low-luminosity red giant branch stars could be treated as main sequence stars (or visa versa) by mistake. The oscillation amplitudes - and detection probabilities - of the stars could then be grossly inaccurate. Figure 2.26 shows a large discrepancy between the luminosity values calculated using the two catalogues.

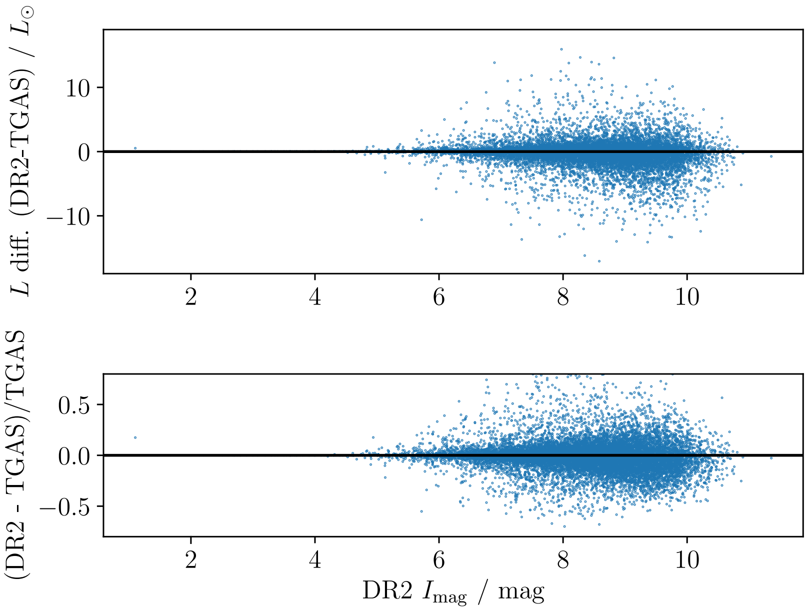


Figure 2.26: The luminosity difference in the ATL using parallaxes from the DR2 and TGAS catalogues. Results are shown as a function of I_{mag}.

This spread in parallax values also has consequences for the inferred distances of these stars (Figure 2.27). As expected, there is more of a discrepancy for fainter stars. For a few stars with apparent magnitudes of ~ 10 mag, there is a difference of up to 200 pc between the values inferred from the two catalogues.

These differences in distance values will manifest themselves in several ways. One of the consequences of these different differences is in the de-reddened apparent magnitudes (Figure 2.28). Distance affects I-band apparent magnitude through the extinction coefficient A_{I}. This extinction has a small effect on the apparent magnitudes (and therefore the calculated stellar shot noise levels of the stars). Figure 2.28 shows that incorrect distances (and hence reddening coefficients) can lead to both over and under-estimating the apparent magnitudes of the stars.

Like I_{mag}, effective temperature is another parameter that is indirectly affected by the different sources of parallax and distance. Using DR2 parallaxes and distances leads to a relatively small difference in the calculated effective temperatures of the stars. In the same way as I_{mag}, effective temperatures in the ATL are minimally affected between the different interferometric sources (Figure 2.29). This is due to different extinction

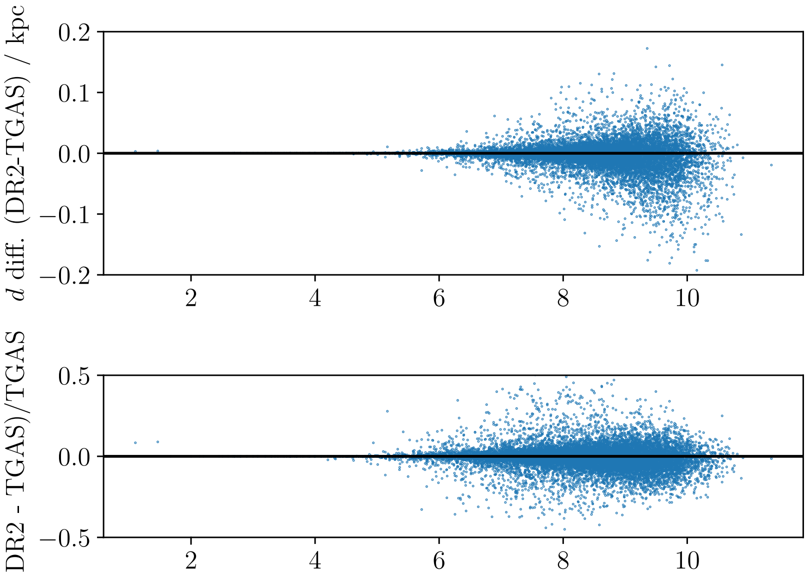


Figure 2.27: The inferred difference in distance between the DR2 and TGAS catalogues, as a function of I_{mag} (from Bailer-Jones et al. 2018 and Astraatmadja & Bailer-Jones 2016, respectively).

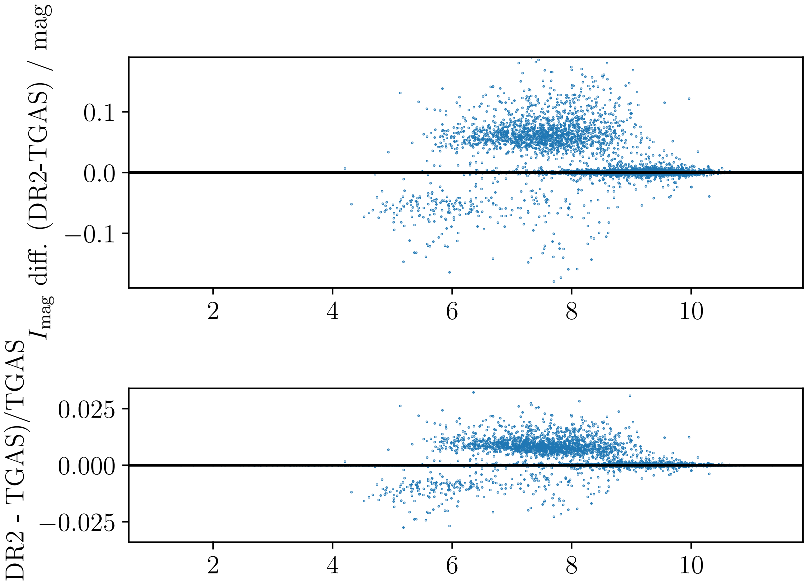


Figure 2.28: The I_{mag} difference in the ATL using the DR2 and TGAS catalogues, as a function of I_{mag} .

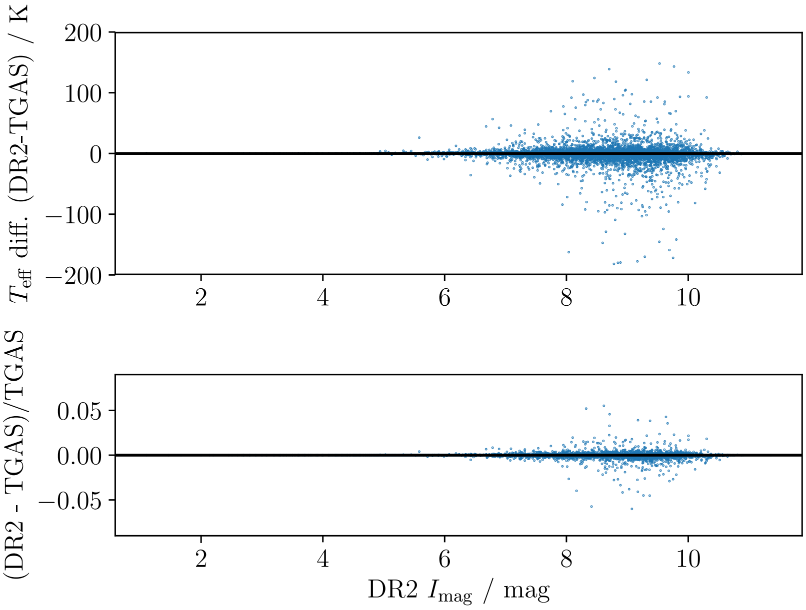


Figure 2.29: The T_{eff} difference in the ATL using the DR2 and TGAS catalogues, as a function of I_{mag} .

coefficients $E(B-V)$ being used, which are calculated using different distance sources.

Once both stellar luminosity and effective temperature are calculated for the stars in the ATL, radius follows from the Stefan-Boltzmann law. If the stellar luminosities and effective temperatures of the stars are different, the radii will also be. This difference is shown in Figure 2.30. The difference in radii is more prominent for fainter stars. This difference is due to more precise parallaxes (and inferred distances) using DR2, compared to TGAS.

In this section, differences between parallaxes (and inferred distances) from the DR2 and TGAS catalogues are shown. These differences were found by creating an ‘alternative’ version of the ATL, using the TGAS catalogue. Stellar parameters calculated using the TGAS catalogue were compared to the ‘real’ ATL, which uses parallaxes and distances from the DR2 catalogue. Results show that the stellar parameters differ significantly when the TGAS catalogue is used instead of DR2. Adopting the more precise DR2 values was therefore a sensible decision.

In the following sections, the stellar parameters in the ‘real’ ATL catalogue (based upon the DR2 and XHIP catalogues) are compared to values from literature. The first

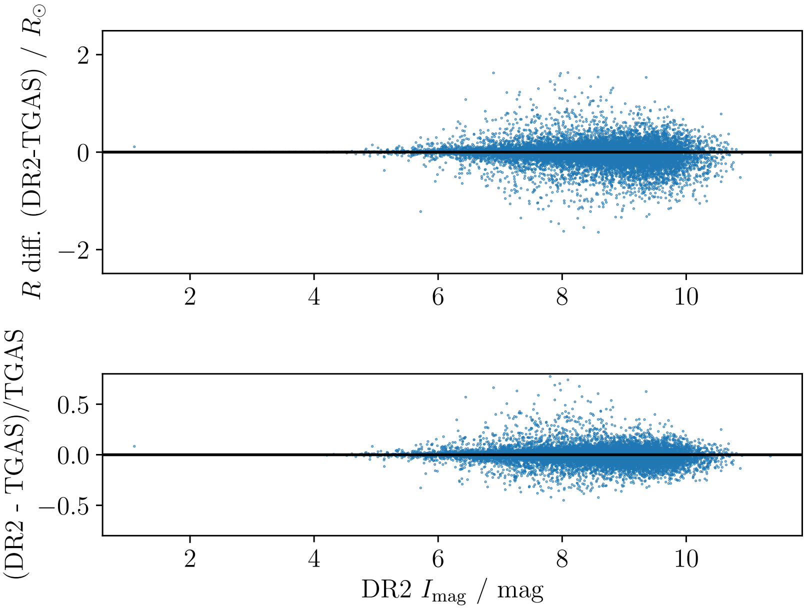


Figure 2.30: The radius difference in the ATL using the DR2 and TGAS catalogues, as a function of I_{mag} .

of these parameters is effective temperature. The results will show that not only can the calculated detection probabilities be relied upon (as shown in Section 2.7), but that the stellar parameters given in the ATL can also be used in further work.

2.8.2 Effective temperature

The P_{mix} method of ranking targets was chosen in Section 2.5. After selecting high priority stars (Section 2.6), the final work to be done on the ATL was to check the accuracy and reproducibility of the parameters used in the target selection. The overall issue of accuracy was tested in Section 2.7, when Monte Carlo trials were used to perturb input parameters in the detection algorithm.

Following this, this section looks at the accuracy of individual stellar parameters calculated in the ATL. This started with parallax, and now moves on to the global stellar parameters.

One of the most important parameters when calculating detection probability is effective temperature. If the T_{eff} values of the stars in the ATL cannot be trusted, the final detection probabilities will be inaccurate. Equation 2.3 shows the polynomial used

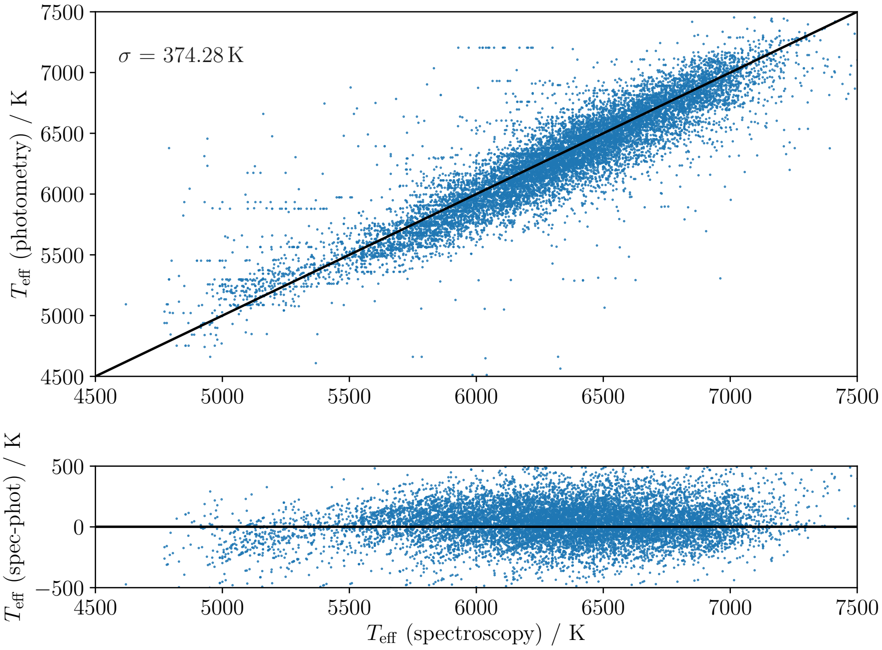


Figure 2.31: Comparison between the PASTEL catalogue and the ATL. There are horizontal lines in the data because the PASTEL catalogue gives several effective temperatures for some stars. No systematic offset is present between the catalogues.

to calculate $T_{\rm eff}$ from (B-V).

The temperatures in the ATL were compared to several catalogues to check a for systematic temperature offset. The first was the PASTEL catalogue (Soubiran et al., 2016). This catalogue contains spectroscopically-determined effective temperatures for over 60,000 stars. Figure 2.31 shows the comparison between temperatures in the ATL, and the PASTEL catalogue. Encouragingly, there is no systematic temperature offset between the spectroscopic PASTEL catalogue and the ATL.

Temperatures are also available in the catalogue presented in Huang et al. (2015). In this work, metallicity-dependent calibrations of effective temperature were made against colours for dwarf and giant stars, based on interferometric data. Huang et al. (2015) provides another useful catalogue to check against, in order to determine if effective temperatures derived in the ATL are reliable. If they are not reliable, metallicity may need to be taken into account when deriving $T_{\rm eff}$ from (B-V) colour.

Figure 2.32 shows that the ATL temperatures are consistent with measurements from

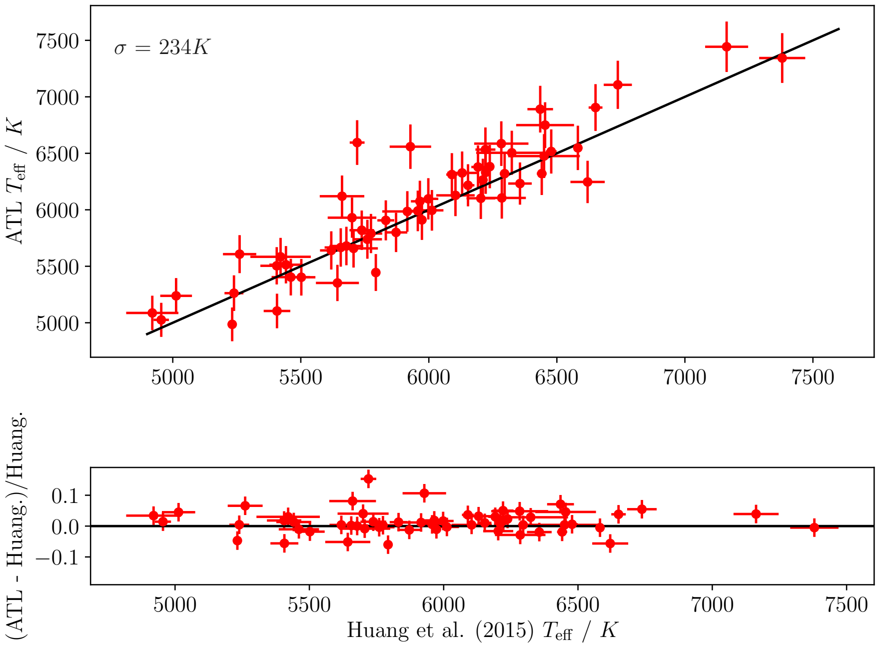


Figure 2.32: Comparison between the Huang et al. (2015) catalogue and the ATL. There is no systematic offset between the two sets of temperatures. The two catalogues are in good agreement.

interferometric data within error bars. This check, along with the comparison to the PASTEL catalogue, confirm that the effective temperatures in the ATL are not systematically offset and can be relied upon. After effective temperature, the next parameter in the ATL to be tested was stellar radius.

2.8.3 Stellar radius

Section 2.8.2 showed how the effective temperatures calculated for ATL stars agree well with literature values. Continuing this theme, this section compares radii values in the ATL with those in literature.

The first source of stellar radii comes from Silva Aguirre et al. (2012). In their work, the authors verified asteroseismically-determined parameters for *Kepler* stars. They did this using *Hipparcos* parallaxes, alongside photometry from the Tycho-2 (Høg et al., 2000) and 2 Micron All Sky Survey (2MASS; Skrutskie et al. 2006) catalogues.

Silva Aguirre et al. (2012) had already verified the radii of the stars they present. A comparison between these stars and the ATL is therefore a reliable way to highlight

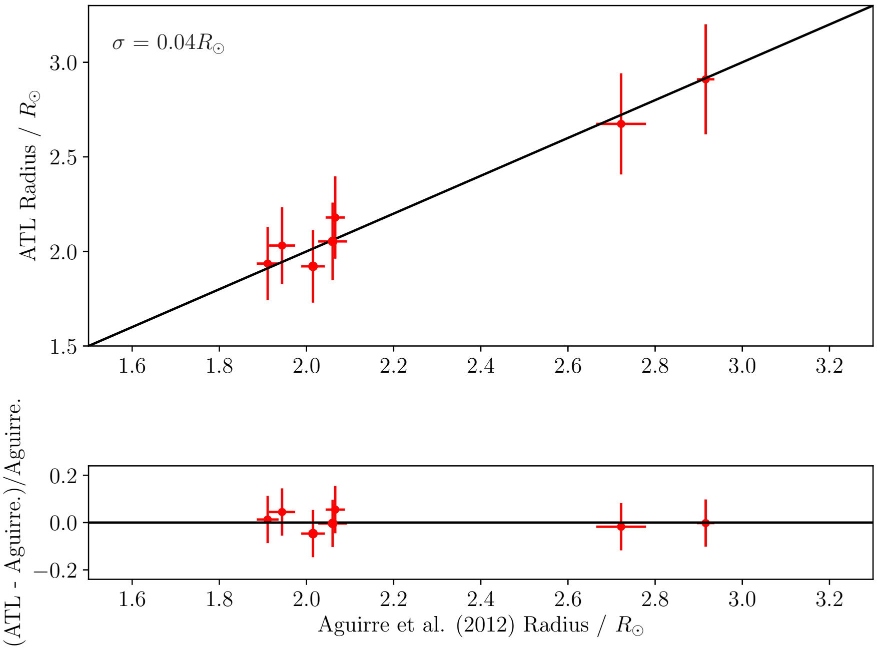


Figure 2.33: Comparison between the radii given in Silva Aguirre et al. (2012), and the ATL. The standard deviation is given in the top left corner of the plot. The ATL radii agree extremely well with the values given in Silva Aguirre et al. (2012).

systematic inaccuracies in the target list. Figure 2.33 shows this comparison. The results from Figure 2.33 are extremely encouraging; there is a standard deviation of only $0.04\,\mathrm{R}_\odot$ between radii in Silva Aguirre et al. (2012) and the ATL.

This scatter is consistent with the uncertainties on the stellar radii. When the formal error bars are larger, there is more expected scatter between the calculated radii, and radii values from literature. That is what is shown here.

Bruntt et al. (2010) is another catalogue where stellar radii are available to compare against. In their work, the authors used interferometry, asteroseismology and spectroscopy to determine stellar parameters. They were therefore able to calculate two sets of radii for their stars. The first set of radii in their paper are from interferometry, using the ‘limb darkening’ method:

$$R/R_\odot \; = 9.30 \times 10^{-3}\,\theta_{\mathrm{LD}}\,/\,\pi. \tag{2.29}$$

θ_{LD} is the limb-darkened angular diameter of the star, and π is the parallax. These radii are compared against ATL values in Figure 2.34. Encouragingly, the two sets of radii

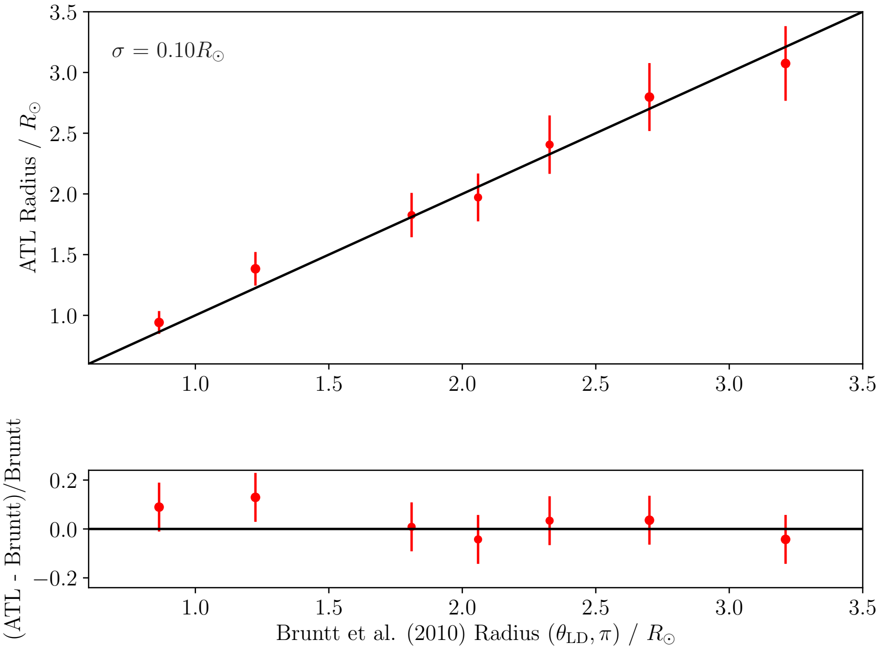


Figure 2.34: Comparison between the Bruntt et al. (2010) radii from the 'limb-darkening' method (using the Stefan-Boltzmann law) and the ATL. The standard deviation is given in the top left corner of the plot. The two sets of radii agree well.

agree with each other. There is a standard deviation of just 0.1 $R_\odot$ between the values.

The second set of radii in Bruntt et al. (2010) are from the 'direct' method, which used the Stefan-Boltzmann law:

$$R/R_\odot = (L/L_\odot)^{1/2} (T/T_\odot)^{-2}. \tag{2.30}$$

Once again, these radii were compared to the ATL values (Figure 2.35). For the third time, this radii comparison gives reassuring results. There is a standard deviation of only $0.21R_\odot$ between the 'direct' method from Bruntt et al. (2010) and the ATL.

The 3 sets of comparisons presented here have confirmed that the radii given in the ATL are precise. In all 3 cases, stellar radii from the ATL were compared to previously validated data, and found to agree. So far, the effective temperatures and stellar radii given in the ATL have both been checked and found to be precise. The next parameter to be checked in this way is the large frequency spacing $\Delta\nu$.

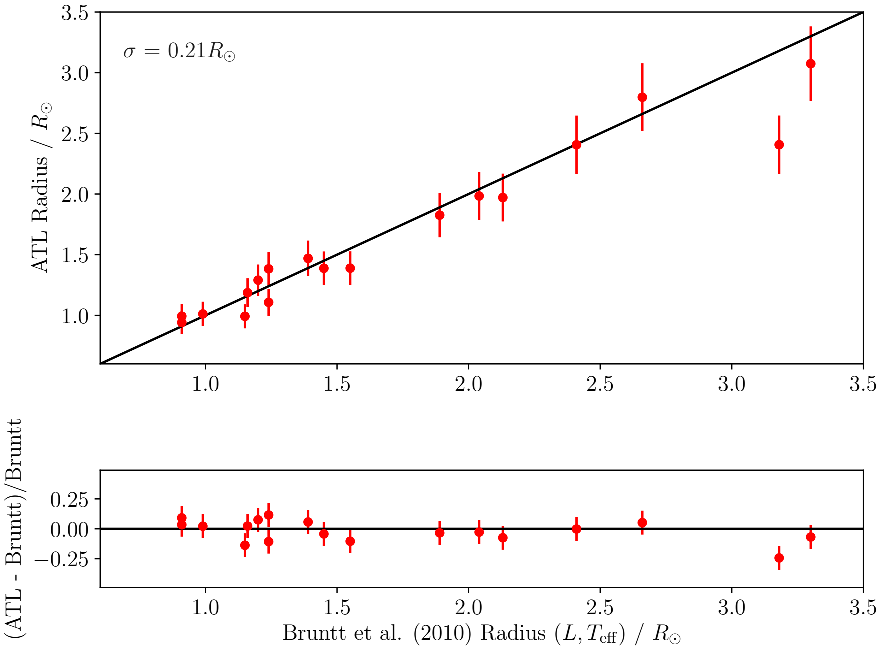


Figure 2.35: Comparison between Bruntt et al. (2010) radii from the 'direct' method (using the Stefan-Boltzmann law) and the ATL. The standard deviation is given in the top left corner of the plot. The two sets of radii agree well.

2.8.4 The large frequency separation

Section 2.8.3 compared radii from catalogues presented in Silva Aguirre et al. (2012) and Bruntt et al. (2010) to those given in the ATL. This section, and the section that follows it, will use the same catalogues from literature to check the precision of the global seismic parameters $\Delta\nu$ and $\nu_{\rm max}$ given in the ATL.

Silva Aguirre et al. (2012) presents the global seismic parameters from power spectra in Huber et al. (2009), using individual frequencies from Appourchaux et al. (2012). Large frequency separations in the ATL were also compared to values from Bruntt et al. (2010), which were taken from Carrier et al. (2005) and Mosser et al. (2008).

Checking the global seismic parameter values given in the ATL against those from literature is important: these parameters greatly influence the detection probability of a star, so determining them precisely is paramount when selecting stars for the ATL.

Moreover, the scaling relations used to calculate $\Delta\nu$ and $\nu_{\rm max}$ in this Chapter are heavily dependent on radius and effective temperature (equations 2.9 and 2.10). Confirming that

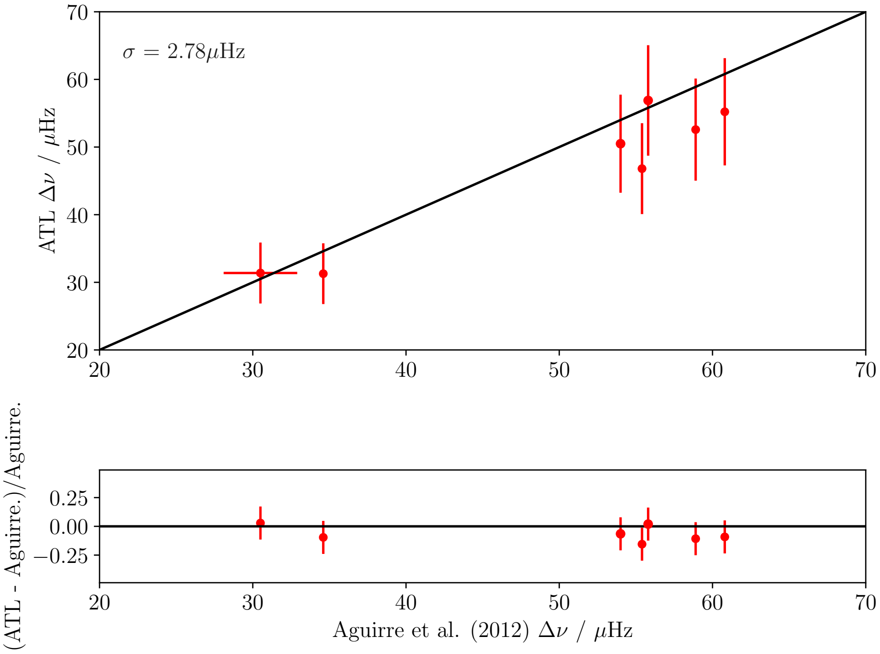


Figure 2.36: Comparison between Silva Aguirre et al. (2012) and the ATL. The standard deviation is given in the top left corner of the plot. There is less than a 3 μHz difference between the the catalogues. The uncertainties on individual $\Delta\nu$ values from Silva Aguirre et al. (2012) are large, so the standard deviation quoted in the figure may be an underestimate of the true scatter between datasets.

$\Delta\nu$ and $\nu_{\max}$ values in the ATL are accurate is therefore another test that the equations used to calculate radius and effective temperature are robust.

Firstly, $\Delta\nu$ values from Silva Aguirre et al. (2012) were compared to the ATL (Figure 2.36). The Figure shows that the standard deviation between the datasets is less than 3 μHz. This leads to the conclusion that the $\Delta\nu$ values in the ATL are remarkably precise.

Similarly, Figure 2.37 compares $\Delta\nu$ values in the ATL to those from Bruntt et al. (2010). Once again, the difference with the literature is small.

Through comparisons with Silva Aguirre et al. (2012) and Bruntt et al. (2010), the large frequency separation values given in the ATL are shown to be reliable. The final global seismic parameter to be compared to literature is $\nu_{\max}$.

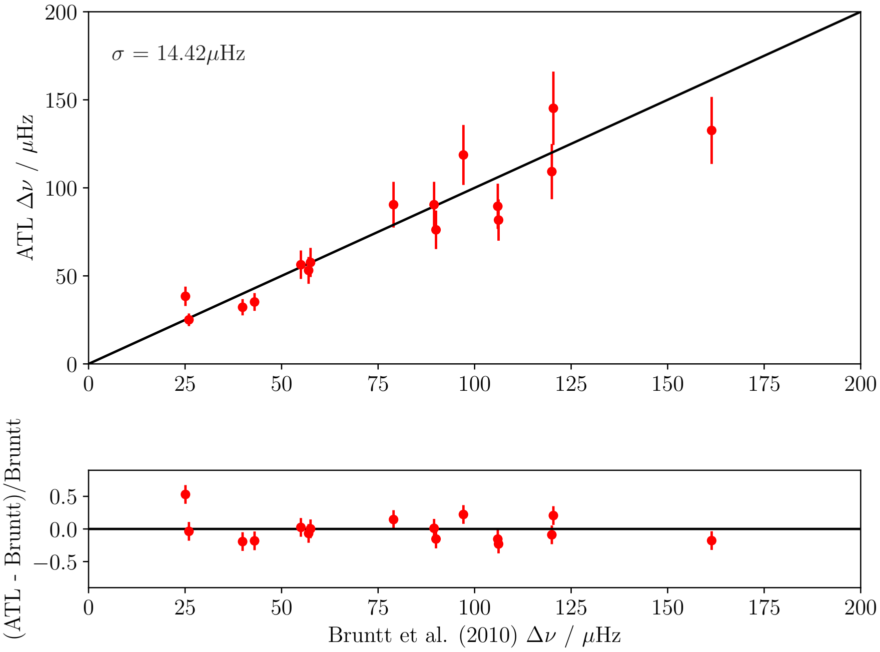


Figure 2.37: Comparison between Bruntt et al. (2010) and the ATL. The standard deviation is given in the top left corner of the plot. There is less than a 15 μHz difference between the the catalogues. The values are consistent with one another, given the uncertainties.

2.8.5 The frequency of maximum power

In the previous section, the values of the large frequency separation given in the ATL were compared to literature values. The $\Delta\nu$ values in the ATL were found to be reliable. This section continues these checks on the other global seismic parameter, this time on the frequency of maximum oscillation amplitude $\nu_{\max}$. $\nu_{\max}$ is an important parameter in the ATL because it is used to calculate the maximum oscillation amplitude height (equation 2.11). This parameters feeds directly into the detection probability of a star.

Firstly, the $\nu_{\max}$ values in the ATL were compared to those from Silva Aguirre et al. (2012) (Figure 2.38). The Figure shows that the discrepancy between the ATL and literature is $\leqslant 20\%$ of the $\nu_{\max}$ value, and lies within quoted uncertainties. Additionally, the ATL appears to slightly underestimate the $\nu_{\max}$ value. This result is expanded upon in Chapter 3, when the scaling relation to calculate $\nu_{\max}$ is tested using fitted oscillation frequencies.

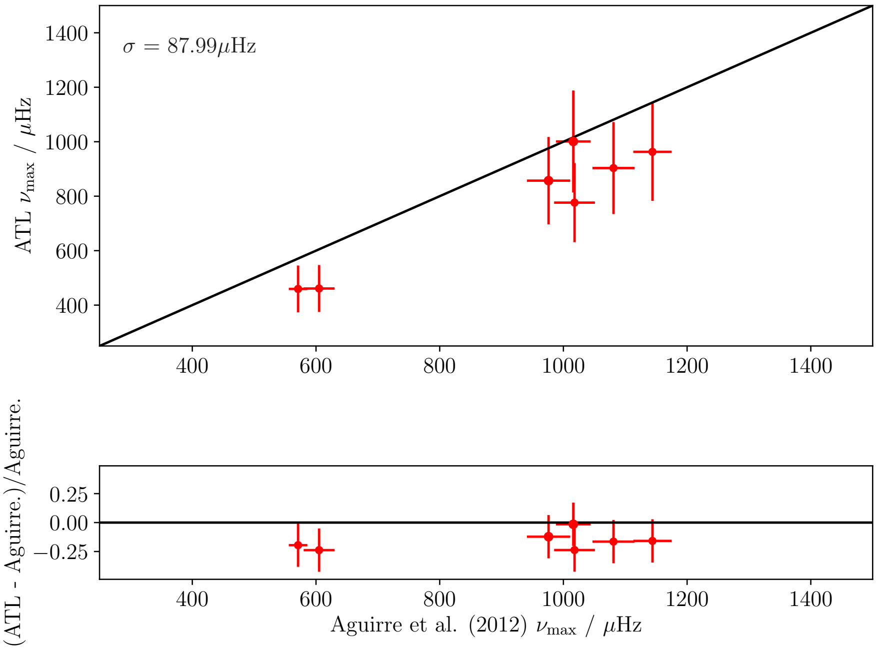


Figure 2.38: Comparison between Silva Aguirre et al. (2012) and the ATL. The standard deviation is given in the top left corner of the plot. The discrepancy between the ATL and literature is $\leqslant 10\%$ of the $\nu_{\rm max}$ value. 10% the expected accuracy of scaling relations.

It is important to check another source of literature values to verify if equation 2.9 underestimates the frequency of maximum oscillation power. To do this, Bruntt et al. (2010) was once again consulted.

The equivalent comparison plot against the ATL is shown in Figure 2.39. The discrepancy between the ATL and literature here seems to be taken into account by the uncertainties given in Bruntt et al. (2010). There is no systematic offset between the datasets. It is important to note that Bruntt et al. (2010) quoted $\nu_{\rm max}$ values in mHz. The precision of the literature values here is quite low, which could account for the standard deviation between the datasets being 466 μHz.

After the comparisons presented here, the $\nu_{\rm max}$ values in the ATL appear to be reliable within the quoted errorbars. This was the final set of checks to ensure that parameters given in the Asteroseismic Target List are robust.

Using the comparisons between parameters in the ATL and those from literature, the ATL was shown to be a source of precise stellar parameters. To conclude the Chapter,

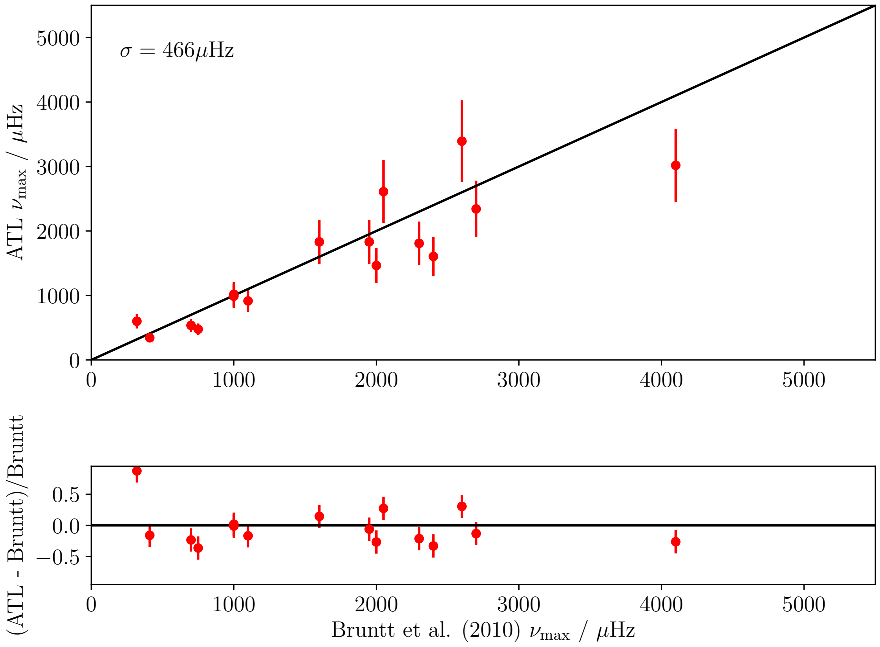


Figure 2.39: Comparison between Bruntt et al. (2010) and the ATL. The standard deviation is given in the top left corner of the plot.

the following section presents a summary of the final version of the ATL that was sent to the TESS Asteroseismic Science Operations Centre.

2.9 Summary of the ATL

This chapter has presented the Asteroseismic Target List; a list of stars to be observed with TESS at a Short Cadence of 2 minutes. The ATL consists of 20,000 stars that are likely to exhibit detectable solar-like oscillations. Figures 2.40, 2.41 and 2.42 show summary plots of this target list.

Figure 2.40 shows an HR diagram of the ATL. It includes a colourbar of the apparent magnitudes of the stars. In order to detect solar-like oscillations in a star, the star must be brighter than ~11th magnitude. Since oscillations in more evolved stars are larger, detections can be made to a fainter apparent magnitude. This result is confirmed in (Campante et al., 2016) (Appendix B).

Secondly, Figure 2.41 shows a histogram of the I-band magnitudes of the highest-ranked 20,000 stars in the ATL. This reinforces the HR diagram shown above: the majority

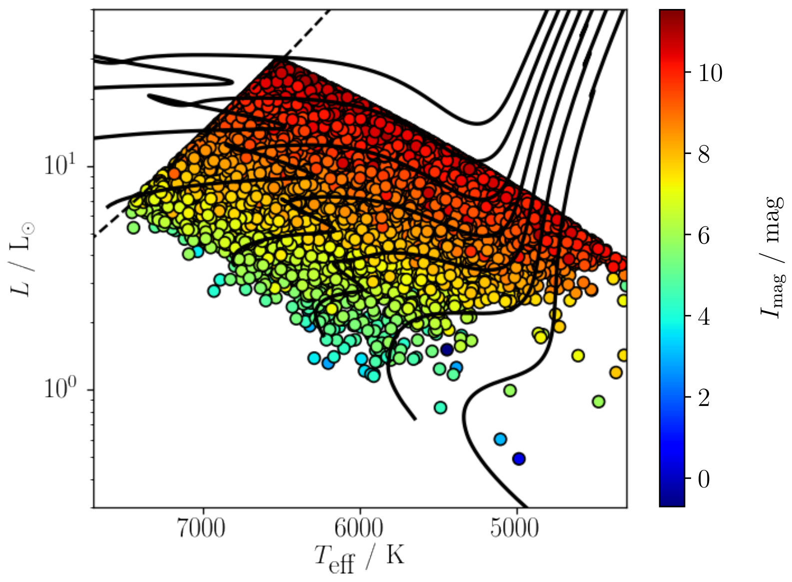


Figure 2.40: An HR diagram of the highest ranked 20,000 stars in the ATL. Black lines show 0.8 - 2.0 $\mathrm{M}_\odot$ evolutionary tracks. The datapoints are coloured with their apparent I-band magnitudes. This shows that only the brightest main sequence stars will have detectable oscillations with TESS. This results is confirmed by Campante et al. (2016).

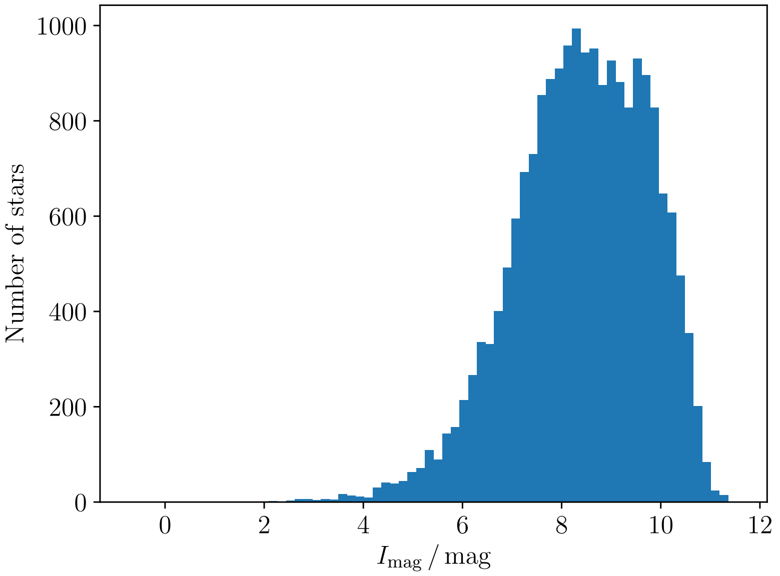


Figure 2.41: A histogram showing the distribution of the apparent I-band magnitudes of the highest ranked 20,000 stars in the ATL. This contains brighter main sequence and subgiant branch stars, and fainter low-luminosity red giant branch stars.

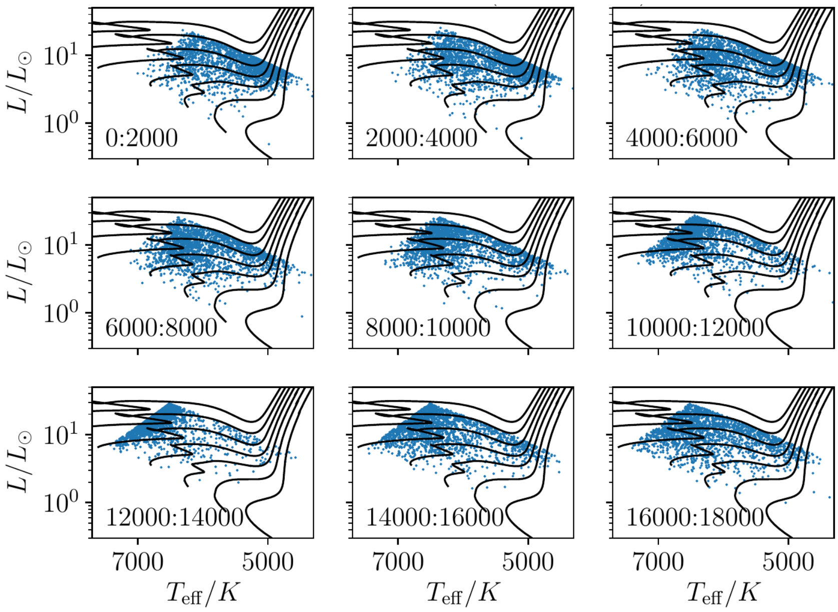


Figure 2.42: Subplots showing HR diagrams of the highest-ranked 18,000 stars in the ATL. Black lines show evolutionary tracks for stars with masses of 0.8 - 2.0 $\mathrm{M}_\odot$ evolutionary tracks. Numbers in the bottom left corner denote the range of ranks shown. There is a distribution of main sequence, red giant and hot stars across the subplots. This was the desire behind choosing the $P_{\rm mix}$ ranking metric.

of stars in the ATL have apparent magnitudes between 7 - 10 mags.

Thirdly, 2000-star subplots of the ATL are shown in Figure 2.42. The $P_{\rm mix}$-ranking metric allows the stars stars with the highest detection probabilities to be put at the top of the list, while selecting stars from across the HR diagram. The method used to select these stars, calculate their detection probabilities, rank them, and test the algorithm used to choose them are summarised below.

The chapter begins with a summary of the work (Section 2.1) and the motivation behind constructing the ATL (Section 2.2). The ATL was produced from the *Gaia* DR2 and *Hipparcos* XHIP interferometric catalogues (Lindegren et al., 2018; Anderson & Francis, 2012). These input catalogues are introduced in Section 2.3.

After the DR2 and XHIP catalogues were combined, cuts were made to the data to remove non solar-like stars. Stars with detectable oscillations then needed to be selected from this remaining list of targets. In order to select these solar-like oscillators from the remaining list, the probability of observing a power-excess due to solar-like modes of oscillation was calculated for each star (Section 2.4). This probability takes into account the power due to modes of oscillation, as well as the background power due to instrumental and granulation noise components.

Once a detection probability was calculated for every star in the target list, the stars needed to be ranked. The TESS Science Team could then select stars from the list, until the 2-minute Short Cadence allocation had been exceeded. Section 2.5 explains the three different ranking methods that were tested for the ATL.

Firstly, the Hertzsprung-Russell diagram was separated into three regions. This was done to ensure that stars from across effective temperature / luminosity space would be selected for Short Cadence observation, and that one area of the HR diagram would not be over-sampled. The HR diagram was separated into three areas: a region containing main sequence and subgiant branch stars, a region containing low-luminosity red giant branch stars, and a region containing stars that lie close to the δ-Scuti Instability strip.

These regions in the HR diagram were used to ensure even sampling from across effective temperature / luminosity space. This sampling method was then improved upon by using Kernel Density Estimation. KDEs can provide perfectly uniform sampling across a parameter space. This is achieved by treating every data point (in this case, every pair of T_{eff} / L values) as a normal distribution, rather than a discreet value.

Although Kernel Density Estimation provided perfectly uniform sampling, it resulted in thousands of solar-like oscillators being remove from the ATL. This sacrifice was not deemed worthwhile, as the purpose of the ATL is to select solar-like stars with potentially detectable oscillations.

The final method used to rank stars involved a mixture of detection probabilities, denoted by P_{mix}. P_{mix} provided the perfect balance between prioritising solar-like oscillators, while including stars that lie close to the δ-Scuti Instability strip. Until now, the oscillation and damping mechanisms in the region close to the instability strip have been poorly understood. The ATL provides the perfect opportunity to better understand these mechanisms.

After the list was ranked using a mixture of detection probabilities, a subset of the ATL was chosen in Section 2.6. This subset contained stars that would be observed by TESS at the shortest available cadence of 20 seconds. The subset was therefore reserved for the most important stars in the target list. Stars allocated to this subset were main sequence targets with the highest probability of containing detectable oscillations. Being observed at a shorter cadence will increase the probabilities that these stars display solar-like oscillations. This will improve the asteroseismic yield from TESS.

The final part of constructing the ATL involved checking the robustness of the method used to calculate detection probabilities. If the detection probabilities calculated in this Chapter cannot be trusted, the targets chosen in the ATL will not necessarily be the stars that are most likely to display solar-like oscillations.

The first robustness-check that was performed involved Monte Carlo trials (Section 2.7). Input parameters in the ATL were scattered according to their uncertainties. A new detection probability was then calculated for each star using the scattered parameters. This process was repeated 1000 times for each star. This produced a distribution of 1000 detection probabilities for every star.

The spread of these probabilities provided a measure of the robustness of the detection probability algorithm. If there was a large spread in the values, the calculated detection probabilities would have a low precision and could not be relied upon to select targets. Results from the trials showed that the precision of the detection probabilities ranged from 8 - 10 %. It was therefore concluded that the method used to construct the ATL was a reliable way of choosing solar-like asteroseismic targets.

After the Monte Carlo trials, individual stellar parameters were compared to values from literature in Section 2.8. The global stellar parameters R and T_{eff} were checked, alongside the asteroseismic variables $\Delta\nu$ and ν_{max}. All of the values given in the ATL for these parameters were found to agree with those from literature, to within reasonable uncertainties.

The effects of using a different parallax source (rather than DR2) were also shown to be significant. The DR2 catalogue was confirmed as the correct choice for the ATL. This concludes the work done to construct the Asteroseismic Target List for TESS.

The following Chapter presents an application of the same asteroseismic target selection method used here. The focus of that work is to predict the yield of asteroseismic exoplanet-host stars.

Chapter 3

Testing the scaling relations used to make the Asteroseismic Target List

> The work presented in this chapter used data from Chaplin et al. (2014a), Aguirre et al. (2017) and Lund et al. (2017). The updated envelope width relation presented in this chapter will appear in Lund, Schofield et al. (in prep). This paper has a wider focus, and covers more material.

3.1 Introduction

One way to better understand a star is to compare it to the Sun. This can be done using equations which relate the Sun's properties to those of other stars. Equations that do this are called scaling relations. Scaling relations were used to select stars for TESS to observe when constructing the Asteroseismic Target List (Chapter 2). They were also used to make predictions about the number of asteroseismic exoplanet-host stars observed by TESS (Appendix B).

The work done in both of those Chapters was centred around a common theme: comparing a star to the Sun using scaling relations, in order to constrain the properties of that star. The aim of this Chapter is to determine the following: to what extent scaling relations can be relied upon to provide accurate stellar parameters. This was investigated in several stages, starting with predictions that can be made about the solar-like oscillation envelope.

When viewed in a power spectrum, solar-like oscillations appear within a Gaussian envelope[1]. The properties of this envelope can be described using scaling relations. Namely, the width (Γ_{env}) and central height (H_{env}) of the envelope can both be determined.

Alongside these are the scaling relations used to calculate the global seismic properties ν_{max} and $\Delta\nu$ (e.g Huber et al. 2011). ν_{max} is the central frequency on the Gaussian envelope, while $\Delta\nu$ is the large frequency spacing between modes of the same angular degree l. When producing the ATL, scaling relations that require radius R and effective temperature T_{eff} were used. The relations used do not require a stellar mass M. These are the relations that are investigated in Section 3.2.

This work will show that the amount of information available when calculating the envelope properties of a star greatly informs the accuracy of the results. Having detailled knowledge of individual modes of oscillation leads to much more accurate values of the envelope properties. Conversely, when very little is known about a star (such as only a parallax and a colour), the properties of the solar-like oscillation envelope for that star may be poorly constrained with scaling relations.

Alongside providing an understanding of the oscillation envelope, scaling relations can also constrain the global parameters of stars. For example, (B-V) colour can be used to calculate the effective temperature of a star (Torres, 2010). Alongside effective temperature, these global parameters include the stellar luminosity and radius (e.g Huber et al. 2011; Epstein et al. 2014; Sharma et al. 2016). The scaling relations to calculate these global parameters are the subject of Section 3.3.

The work described in Section 3.3 follows a similar theme to Section 3.2. It aims to answer two questions: firstly, how accurately can scaling relations calculate stellar parameters? Furthermore, if different knowledge is available for a star (such as a different source of parallax), how much can predictions from the scaling relations be affected?

The results from Section 3.3 show that the source of parallax can affect predictions

---

[1] The exact form of this envelope is uncertain. Lund et al. (in prep) finds the oscillation envelope to be better described with a Lorentzian function.

from the scaling relations. It is therefore important to understand just how much the parallax values of a single star can vary between different sources. This was the motivation for the work in Section 3.4.

In Section 3.4, the parallax and distance values from different sources are compared to one another. These sources are the Extended Hipparcos catalogue (XHIP; Anderson & Francis 2012), the Tycho-Gaia asterometric solution (TGAS; Michalik et al. 2015), and the *Gaia* Data Release 2 (DR2; Lindegren et al. 2018). This comparison was done to show that parallax (and distance) values vary significantly between sources, and must be taken into account when inferring stellar properties from the scaling relations.

In Sections 3.2, 3.3 and 3.4, the focus is on improving our understanding of the scaling relations. In particular, work was aimed at understanding how different parameters given to the relations can lead to different results from them. Leading on from this, Section 3.5 presents the work that was done to improve the scaling relations. Specifically, the equation to calculate the width in frequency of the oscillation envelope was improved. A piecewise relation using both effective temperature and $\nu_{\max}$ was developed. This updated scaling relation is presented and justified in Section 3.5.

All the work in this Chapter was performed on the same set of stars. These are stars that were observed during the nominal *Kepler* mission. They all have short-cadence, long-baseline photometry for over one year. Chaplin et al. (2014a) (hereafter C14) published the global properties of these stars.

Further work was performed on a subset of the stars from C14. This subset was a group of 66 main sequence stars, titled the LEGACY sample (Aguirre et al. 2017; Lund et al. 2017, hereafter L17). In these papers, the solar-like modes of the stars were fitted - a process known as 'peak bagging'. L17 fitted for the frequencies, amplitudes, and linewidths of the individual modes within these stars.

3.2 Mode envelope properties with the scaling relations

The LEGACY stars were first used to compare predictions made about the Gaussian mode envelope properties. When constructing the asteroseismic target list for TESS, the parameters of the solar-like oscillation envelope were estimated. These envelope parameters were used to calculate the a probability of detecting oscillations, P_{det}, for each star. In order to make reliable predictions about P_{det}, the mode envelope needed to be estimated accurately.

The solar-like oscillation envelope can be modelled with a Gaussian function (e.g Chaplin et al. 2011). Gaussian functions are described using three parameters: the central abcissa value of the curve, the height at that centre, and the width of the curve. When describing the solar-like oscillation envelope, these parameters are denoted with ν_{max}, H_{env} and Γ_{env} respectively.

Alongside these three parameters, it is helpful to define a fourth global oscillation property when describing the modes within a star. This fourth property is the large frequency spacing between adjacent modes of the same angular degree. The scaling relations to calculate $\Delta\nu$ and ν_{max} are given in Chapter 2. Once again, the relation for $\Delta\nu$ is

$$\Delta\nu = \Delta\nu_{\odot} \left(\frac{R}{R_{\odot}}\right)^{-1.42} \left(\frac{T_{\mathrm{eff}}}{T_{\mathrm{eff}\odot}}\right)^{0.71} \mu\mathrm{Hz}\,. \tag{3.1}$$

The equivalent scaling relation for the central abscissa of the Gaussian (ν_{max}) is

$$\nu_{\mathrm{max}} = \nu_{\mathrm{max}\odot} \left(\frac{R}{R_{\odot}}\right)^{-1.85} \left(\frac{T_{\mathrm{eff}}}{T_{\mathrm{eff}\odot}}\right)^{0.92} \mu\mathrm{Hz} \tag{3.2}$$

(Campante et al., 2016). Both of these global asteroseismic properties can be calculated using the radius R and effective temperature T_{eff} of a star. They do not require the stellar mass M (as in Sharma et al. 2016). Removing mass from equations 3.1 and 3.2 reduced the number of parameters to be calculated when constructing the ATL, and hence the number of assumptions that needed to be made about those stellar parameters.

The properties of the solar-like oscillation envelope are described using the Full Width at Half Maximum of the envelope $\Gamma_{\rm env}$

$$\Gamma_{\rm env} \; = \nu_{\rm max} \, / \, 2 \tag{3.3}$$

(Stello et al., 2007) and the height of the centre of the mode envelope $H_{\rm env}$,

$$H_{\rm env} = \frac{A^2_{\rm max} \, \zeta}{\Delta\nu} \, {\rm ppm}^2 \, \mu{\rm Hz}^{-1} \, . \tag{3.4}$$

$\Gamma_{\rm env}$ is related to the Gaussian RMS width by

$$\sigma_{\rm env} = \frac{\Gamma_{\rm env}}{2\sqrt{2 \ln(2)}} \, . \tag{3.5}$$

Within a single $\Delta\nu$ -wide overtone, modes of different angular degrees are visible. In stars with high enough signal-to-noise ratios, $l = 0, 1, 2$ & 3 modes can be seen. These have different relative visibilities in the power spectrum. ζ is the sum in power of these $l = 0, 1, 2$ & 3 relative mode visibilities.

The $A_{\rm max}$ parameter in equation 3.4 is the RMS radial mode amplitude at the centre of the envelope (Chaplin et al., 2011). $A_{\rm max}$ was also calculated when producing the ATL in Chapter 2. Once again, it is given by

$$A_{\rm max} = 2.5 \, \beta \left(\frac{R}{R_\odot} \right)^{1.85} \left(\frac{T_{\rm eff}}{T_{\rm eff\odot}} \right)^{0.57} \; {\rm ppm}. \tag{3.6}$$

The solar radial mode amplitude $A_{\rm max\odot}$ depends on the bandpass of observation. For *Kepler*, $A_{\rm max\odot}$ was calculated as 2.5 ppm. As *Kepler* stars are used in this Chapter, the factor of 2.5 ppm is included in equation 3.6.

Finally, β represents a suppression factor of the solar-like oscillations within a star.

At hotter temperatures, convection ceases to become the prevalent heat transportation mechanism in the stellar envelope. Without this near-surface convection, solar-like oscillations cannot be excited. The modes within stars at hotter temperatures (i.e with thinner convective envelopes) appear to be suppressed in power spectra. Basu & Chaplin (2017) defines this suppression with

$$\beta = 1.0 - e^{-(T_{\mathrm{red}} - T_{\mathrm{eff}})\,/\,1550}, \tag{3.7}$$

with temperatures in Kelvin. This is the same β factor that was used in Chapter 2.

The area of the HR diagram where effective temperatures are too high for solar-like oscillations to exist is called the δ-Scuti Instability strip. It is helpful to estimate the temperature of the cooler, 'red edge' of this instability strip, T_{red}. It can be estimated using another relation from (Chaplin et al., 2011):

$$T_{\mathrm{red}} = T_{\mathrm{red}\odot}\, L^{-0.093}\,, \tag{3.8}$$

where the equivalent solar temperature $T_{\mathrm{red}\odot}$ is 8907 K. Together, the quantities $\Delta\nu$, ν_{max}, H_{env} and Γ_{env} describe the solar-like oscillation envelope of a star.

The properties of the solar-like envelope were calculated for the LEGACY stars. They were calculated in three ways. Firstly, they were calculated using all the information about individual fitted modes from L17. Secondly, the properties of the mode envelope were calculated using the scaling relations 3.1, 3.2, 3.3 and 3.4, with literature values of radius and effective temperature. Thirdly, the mode envelope properties were calculated using the same scaling relations, but with only literature values of (B-V) and V_{mag} (Cutri et al., 2003), alongside parallax (Anderson & Francis, 2012).

When only (B-V) values were available, effective temperatures were calculated from the equation given in Torres (2010),

$$\log(T_{\mathrm{eff}}) = a + b\,(B-V) + c\,(B-V)^2 + ... \tag{3.9}$$

These temperatures can be used alongside a bolometric correction (again from Torres 2010) to calculate stellar luminosity (Pijpers, 2003),

$$\log(L/\mathrm{L}_{\odot}) = 4.0 + 0.4M_{\mathrm{bol}\odot} - 2.0\log\pi - 0.4\left(V_{\mathrm{mag}} - A_{\mathrm{v}} + \mathrm{BC}_{v}\right) . \tag{3.10}$$

Torres (2010) gives the bolometric luminosity as $M_{\mathrm{bol},\odot} = 4.73 \pm 0.03\,\mathrm{mag}$. The extinction A_{v} values of these bright main sequence stars were assumed to be zero. T_{eff} and L from equations 3.9 and 3.10 were used together to calculate the stellar radius with the Stefan-Boltzmann law.

To calculate the 'true' values of $\Delta\nu$ for each LEGACY star, linear fits were performed using the 'peak bagged' radial mode frequencies and overtone numbers $\nu(n)$. The straight line to be fitted is $y = mx + c$. In the case of the fitted radial modes, this becomes $\nu = \Delta\nu n + c$. The gradients of these linear fits were used as the 'true' values of $\Delta\nu$ for the stars.

'True' values of ν_{max}, H_{env} and Γ_{env} were also calculated for each LEGACY star. To do this, the fitted radial mode heights H and widths Γ from L17 were used to calculate the height of the Gaussian envelope at each radial mode frequency (c.f. Chapter 2);

$$H_{\mathrm{gauss}} = \zeta\,\frac{\pi}{2}\,\frac{H\Gamma}{\Delta\nu}\,. \tag{3.11}$$

ζ is the same sum of the mode visibilities as in equation 3.4. $\Delta\nu$ is the 'true' large frequency spacing from the linear fits described above.

Once the height of the Gaussian H_{Gauss} was calculated at every radial mode frequency ν, a Gaussian curve was fitted to the $(\nu, H_{\mathrm{Gauss}})$ data of each LEGACY star. The central frequency, height and width of the Gaussian curve give the 'true' values of ν_{max}, H_{env} and Γ_{env}, respectively.

The 'true' values of $\Delta\nu$, ν_{max}, H_{env} and Γ_{env} from the linear and Gaussian fits were compared against values from scaling relations. The results of these comparisons are shown below.

3.2.1 Results from calculating the mode envelope properties

In this Section, the envelope properties of LEGACY stars were calculated in three ways. Firstly, linear and Gaussian functions were fitted to the 'peak bagged' radial mode frequencies, widths and heights from L17. The parameters from these fits were treated as the 'true' envelope values of these stars.

Secondly, scaling relations were used to calculate the Gaussian envelope properties. These scaling relations (equations 3.1, 3.2, 3.3 and 3.4) were given values of stellar radius R and effective temperature T_{eff} from C14.

Thirdly, the same scaling relations (equations 3.1, 3.2, 3.3 and 3.4) were used once more. This time, rather than using literature values of R and T_{eff}, only literature values of (B-V) and π were used. This third technique is analogous to the method used to construct the ATL in Chapter 2. It is therefore useful to compare results from this to 'true' values of the Gaussian envelope. The parallax values were taken from the XHIP catalogue. This Chapter will go on to show that the XHIP catalogue should not be used to infer stellar parameters from.

Figure 3.1 presents the results of comparisons between different values of $\Delta\nu$. The 'true' values of $\Delta\nu$ from linear fits to the 'peak bagged' radial modes are shown as blue points. Literature values of radius and effective temperature from C14 were also used to calculate $\Delta\nu$ from equation 3.1. These values are shown as orange data points. Thirdly, the green data points show large frequency spacing values if $\Delta\nu$ is instead calculated using estimated values of R and T_{eff}, rather than values from literature. These radii and effective temperatures come from the scaling relations presented in Section 3.2.

Figure 3.1 shows that the large frequency spacing is very well recovered with literature values of R and T_{eff}, which come from asteroseismic *Kepler* data (C14). If stellar radius and effective temperature are instead estimated using (B-V) and parallax, the scatter of $\Delta\nu$ about the true value increases but is still well estimated by equation 3.1.

Secondly, Figure 3.2 shows equivalent results for the envelope width Γ_{env}. The Figure shows that the envelope widths of the LEGACY stars are consistently overestimated by

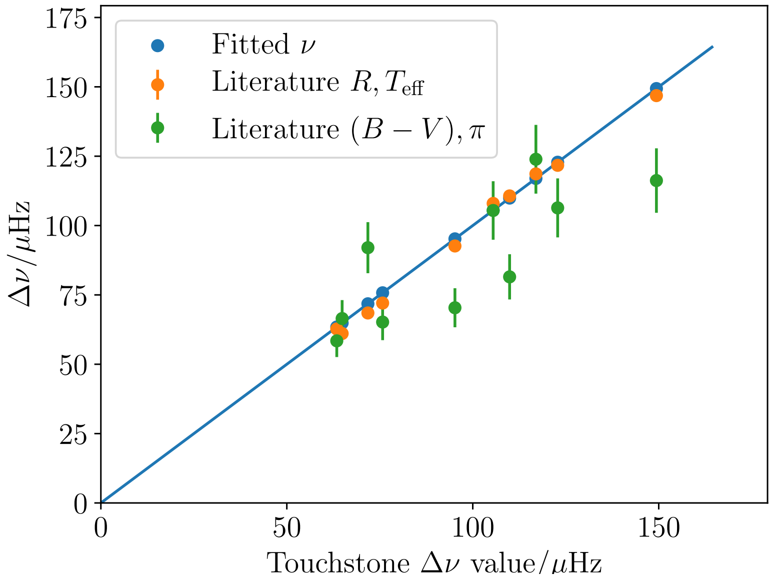


Figure 3.1: A plot showing the values of the large frequency spacing $\Delta\nu$. 'True' values of $\Delta\nu$ are from linear fits to L17 data (blue points). They are compared to values from the scaling relations. Orange points show the results of scaling relations using literature values of R and T_{eff}. Literature values of parallax and (B-V) were also used to calculate R and T_{eff} , before $\Delta\nu$ (green points). The spread in data is minimal, suggesting that the $\Delta\nu$ relation used provides a sensible value of $\Delta\nu$ for these stars.

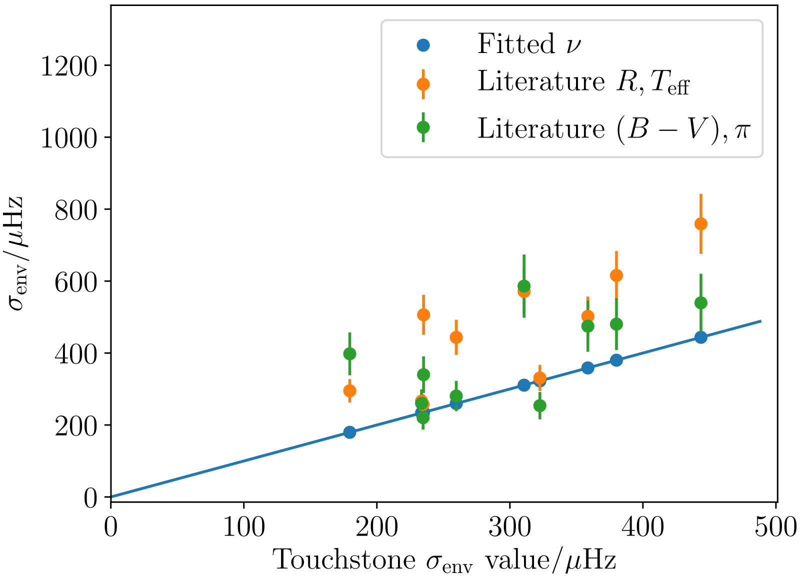


Figure 3.2: A plot showing values of the envelope width $\sigma_{\rm env}$. 'True' values of $\sigma_{\rm env}$ are from Gaussian fits to L17 data (blue points). They are compared to values from the scaling relations. Orange points show the results of scaling relations using literature values of R and $T_{\rm eff}$. Literature values of parallax and (B-V) were also used to calculate R and $T_{\rm eff}$, before $\sigma_{\rm env}$ (green points). The $\sigma_{\rm env}$ scaling relation consistently over-estimates the 'true' value of the Gaussian envelope width.

equation 3.3, even when literature values of R and $T_{\rm eff}$ are used. It was found that the envelope width of solar-like oscillations from Stello et al. (2007) has been over-estimated. A more accurate envelope width relation is needed. This result is developed later in this Chapter (Section 3.5), when an updated envelope width relation is presented.

Thirdly, Figure 3.3 shows the 'true' envelope heights of the LEGACY stars, alongside values from the scaling relations. Equation 3.4 predicts the envelope heights of the LEGACY stars reasonably well, especially if asteroseismically-inferred values of R and $T_{\rm eff}$ from literature are used.

There is a slight disagreement in $H_{\rm env}$ values between the 'true' envelope heights, and the heights from equation 3.4. This may be because equation 3.4 is not fully capturing how the envelope height varies in practice for different combinations of stellar properties. This leads to slightly more spread in the envelope heights than expected.

Lastly, $\nu_{\rm max}$ predictions for the LEGACY stars were compared to fitted values in

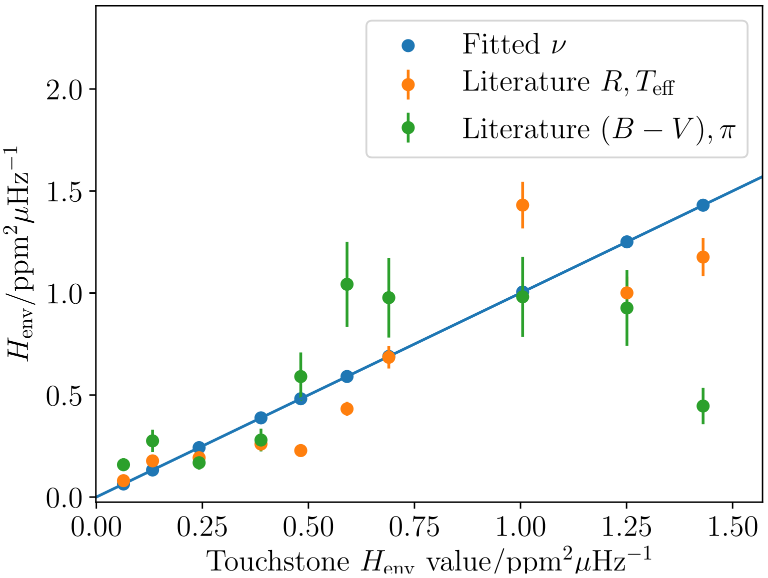


Figure 3.3: A plot where the observed values of the envelope height $H_{\rm env}$ from L17 (blue points) are compared to values from scaling relations. Orange points show the results using literature values of R and $T_{\rm eff}$. Literature values of parallax and (B-V) were also used to calculate R and $T_{\rm eff}$ beforehand (green points). For some LEGACY stars, the equation to estimate $H_{\rm env}$ appears to perform well when radius and $T_{\rm eff}$ are known. In other stars, the realisation noise seems to make a reliable $H_{\rm env}$ estimate impossible. When values of parallax and (B-V) are used, the envelope height is poorly estimated due to this noise.

Figure 3.4. $\nu_{\rm max}$ is consistently underestimated when equation 3.2 is used; the stars are assumed by equation 3.2 to be more evolved than in reality. When values of (B-V) and parallax are used (the green points), this can be partly explained by the stellar radii of these stars being over-estimated (c.f. Section 3.3).

A bias is also seen between the accepted $\nu_{\rm max}$ values and the values from equation 3.2 when literature values of R and $T_{\rm eff}$ are used (orange data points). This suggests that equation 3.2 gives underestimated $\nu_{\rm max}$ values, even when reliable values of R and $T_{\rm eff}$ are used.

In this Section, equations were used to calculate the envelope properties of a set of LEGACY stars. The results show that when only (B-V) and parallax values are used, the accuracy of predictions made about the solar-like oscillation envelope can be reduced. Moreover, biases can be introduced into the results.

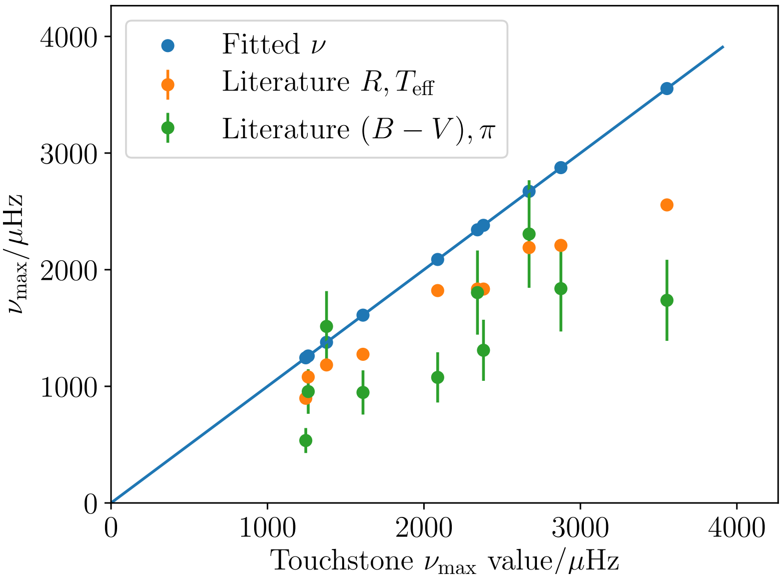


Figure 3.4: A plot where the observed values of ν_{max} from L17 (blue points) are compared to values from scaling relations. Orange points show the results using literature values of R and T_{eff}. Literature values of parallax and (B-V) were also used to calculate R and T_{eff} beforehand (green points). In almost all cases, the envelope is estimated at a lower value than in reality. This suggests that there is a problem with equation 3.2, or that the radii of these stars are systematically over-estimated.

Conclusions can also be drawn about some of the individual scaling relations used here. The results show that equation 3.2 under-estimates the $\nu_{\max}$ value of stars. Secondly, equation 3.3 was shown to systematically over-estimate the oscillation envelope width of a star. The following Sections will unpack these effects.

3.3 Global properties with the scaling relations

In the previous section, LEGACY data was used to test equations that describe the solar-like mode envelope. One of the conclusions from this work is that the equation to calculate $\nu_{\max}$ - equation 3.2 - systematically under-estimates the stellar frequency of maximum amplitude.

As well as exploring this $\nu_{\max}$ under-estimate in more detail, this section investigates whether the source of parallax used in the scaling relations affects the predictions from them. Are predictions from the scaling relations more or less accurate depending on the parallax source?

To determine if this is the case, scaling relations were used to calculate global properties of the LEGACY stars: their effective temperatures, luminosities, radii and $\nu_{\max}$ values. In the previous section, the 'true' envelope parameters of the LEGACY stars were obtained from Lorentzian fits to the solar-like oscillations of the stars (L17). Here, the 'true' global properties of the LEGACY stars were taken from Chaplin et al. (2014a) (C14). C14 performed grid-based modelling using the asteroseismic data of these stars.

In this section, different catalogues of parallax values were used to test the global scaling relations. These parallax sources are the Extended Hipparcos catalogue (XHIP; Anderson & Francis 2012), the Tycho-Gaia Astrometric Solution (TGAS; Gaia Collaboration et al. 2016) and the *Gaia* Data Release 2 (DR2; Lindegren et al. 2018). These different sources of parallax will produce different predictions about the global stellar properties. The more accurate the source of parallax, the closer the predicted stellar properties will be to the 'true' values from the literature.

The 'true' global stellar properties are from C14, and are shown as orange datapoints

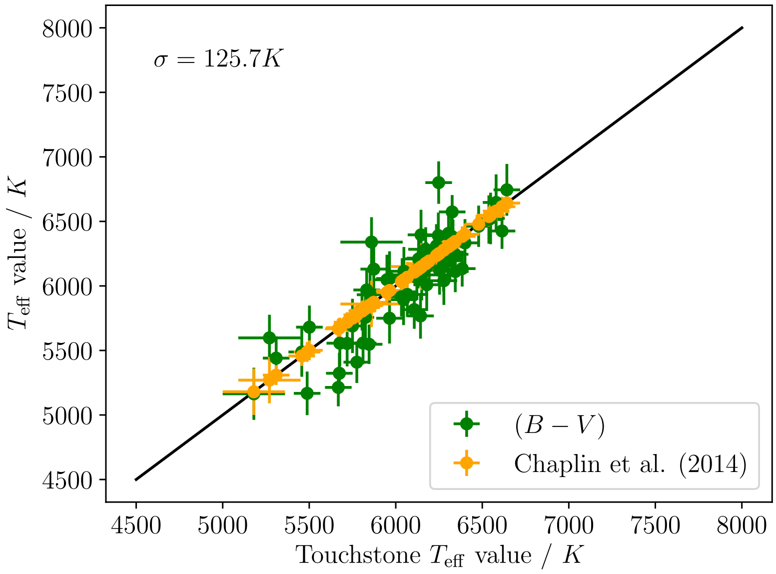


Figure 3.5: Literature 'touchstone' T_{eff} values, compared to scaling relation predictions using $(B-V)$ colour.

throughout the Section. Black lines have been added to the values from C14 as visual guides. These 'true' properties are compared to green datapoints, which show predictions using only (B-V) and parallax.

3.3.1 Effective temperature

Firstly, the effective temperatures of the LEGACY stars were calculated from their (B-V) colours. These temperatures were calculated using a polynomial given in Torres (2010). The effective temperatures are compared to literature values in Figure 3.5. The plot shows that there is minimal scatter from equation 3.9. Alongside this, no systematic temperature offset is introduced by the polynomial.

To confirm that equation 3.9 does not produce an effective temperature offset, the T_{eff} values in Figure 3.5 were compared to a different source of literature temperatures. In Figure 3.6, the effective temperatures from the polynomial are compared to values from Huang et al. (2015).

The blue line in Figure 3.6 shows the temperature scale produced by equation 3.9. The black points are spectroscopically-measured T_{eff} values from Huang et al. (2015). There is

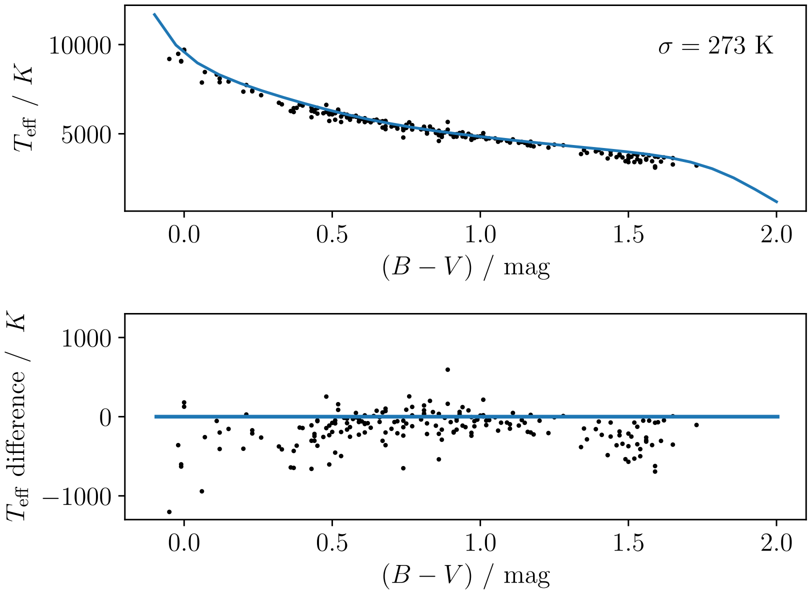


Figure 3.6: The $T_{\rm eff}$ values produced by equation 3.9 are subtracted from those given in Huang et al. (2015). The full effective temperature scale from Torres (2010) is overplotted with the blue line.

a small scatter between the from the polynomial and spectroscopic values temperatures.

The Huang et al. (2015) values appear to lead to an under-estimate of the effective temperatures. Hall et al. (in prep) found that the source of effective temperatures (in this case, spectroscopy or photometry) influences the temperature values themselves. Comparing a photometry-based $T_{\rm eff}$ scale (Torres, 2010) to effective temperatures from spectroscopy (Huang et al., 2015) may lead to systematic differences. Conclusions from Figure 3.6 should therefore be drawn with caution.

To summarise the results of Section 3.3.1, equation 3.9 does not produce an offset when compared to temperature values from grid-based models (Figure 3.5; C14). As a reminder, the model-derived $T_{\rm eff}$ values in C14 are from photometric observations. Results from this comparison should therefore be relied upon more than the comparison with spectroscopy (Figure 3.6; Huang et al. 2015).

The effective temperatures of the stars in the ATL were also compared to another source. They were compared to effective temperatures from the PASTEL catalogue (Soubiran et al., 2016) (Chapter 2 Section 2.8.2). The level of scatter between the ATL effective temperatures and the temperatures from the PASTEL catalogue is commensurate with the standard deviation seen in Figure 3.5.

In the next section, stellar luminosities from scaling relations are compared to those from literature. When calculating luminosity, the parallax source will influence the results (equation 3.10). Different sources of parallax are therefore used in these comparisons, and the results between them are compared.

3.3.2 Stellar luminosities

In the previous section, the equation to calculate effective temperature was compared to values from literature, and found to be reliable. In this section, the equation to calculate luminosity was tested.

Stellar luminosities were calculated using different sources of parallax. Differences between the calculated luminosity values will reveal the effect that the parallax source has on equation 3.10. Three different sources were used: the Extended Hipparcos catalogue (XHIP; Anderson & Francis 2012), the Tycho-Gaia asterometric solution (TGAS; Michalik et al. 2015) and the *Gaia* Data Release 2 (DR2; Lindegren et al. 2018).

Figures 3.7, 3.8 and 3.9 show the difference in luminosity between values from scaling relation 3.10 and literature. Scaling relation luminosities were calculated using parallaxes from the XHIP, TGAS and DR2 catalogues. The plots do not show the same number of LEGACY stars, because some catalogues do not provide parallaxes for every star in the sample.

Within each Figure, luminosities calculated from the scaling relations are shown as the green datapoints. The values from C14 come with uncertainties, which have been plotted. To calculate the error bars on the green datapoints, π, V_{mag} and T_{eff} values were perturbed about their uncertainties. Values of luminosity were calculated iteratively in 1000 Monte Carlo simulations using these scattered π, V_{mag} and T_{eff} values. The spread in the scattered luminosities provides the standard deviation of each datapoint. The annotation in the corner of each plot gives the standard deviation between the values from literature, and

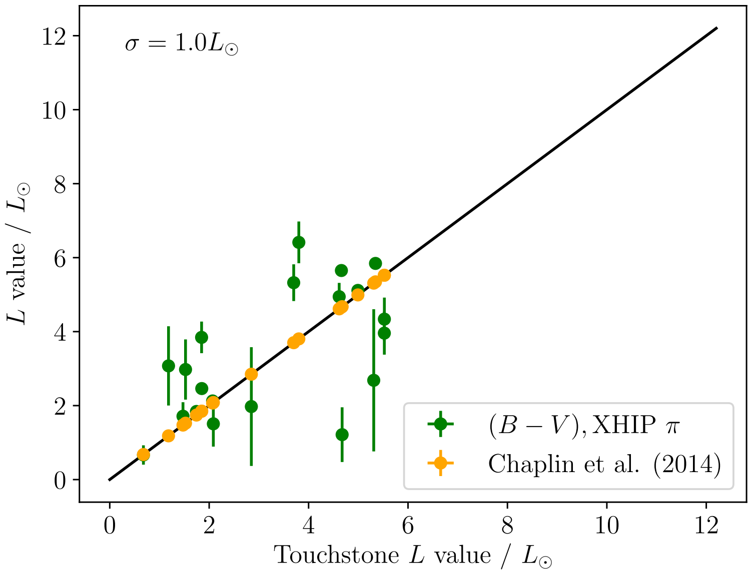


Figure 3.7: The literature 'touchstone' L values, compared to scaling relation values using parallaxes from XHIP.

those from the scaling relations.

Firstly, Figure 3.7 shows the difference between luminosities from literature, and those from equation 3.10 using XHIP parallaxes. A large scatter is seen between the values. The difference seen is not systematic, suggesting that the discrepancy is caused by imprecise parallax values from XHIP, rather than issues with the luminosity scaling relation.

Secondly, Figure 3.8 shows the difference between luminosities from literature, and those from equation 3.10 using TGAS parallaxes. For the majority of stars, less scatter is seen between values from equation 3.10 and C14. The standard deviation quoted in plot 3.8 is scattered by a few high, luminosity stars, whose luminosities are poorly estimated by equation 3.10. However across the whole distribution, TGAS parallaxes appear to provide more precise luminosities than when parallaxes from XHIP are used.

Thirdly, Figure 3.9 shows the difference between luminosities from literature, and those from equation 3.10 using DR2 parallaxes. The remarkably precise parallaxes from DR2 produce very little scatter when used to calculate luminosities with equation 3.10. They agree well with the literature values from C14. This is encouraging, as DR2 is the source of parallax used to calculate stellar parameters (including luminosity) in the ATL

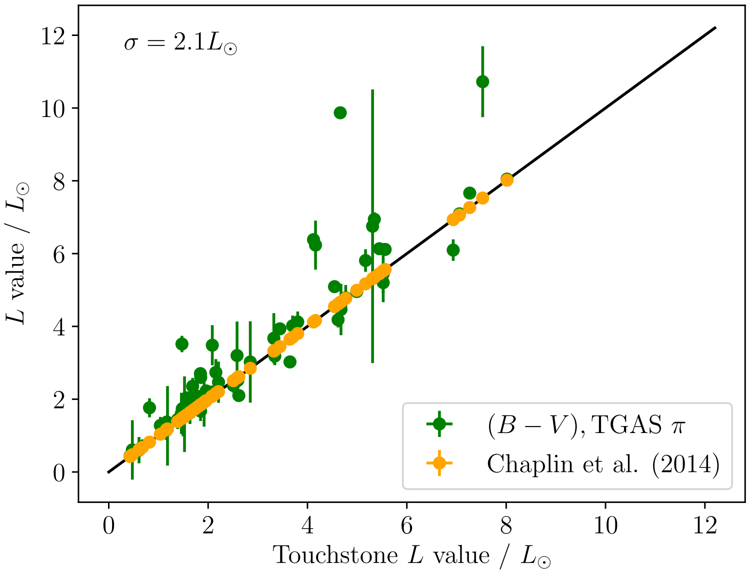


Figure 3.8: The literature 'touchstone' L values, compared to scaling relation values using parallaxes from TGAS.

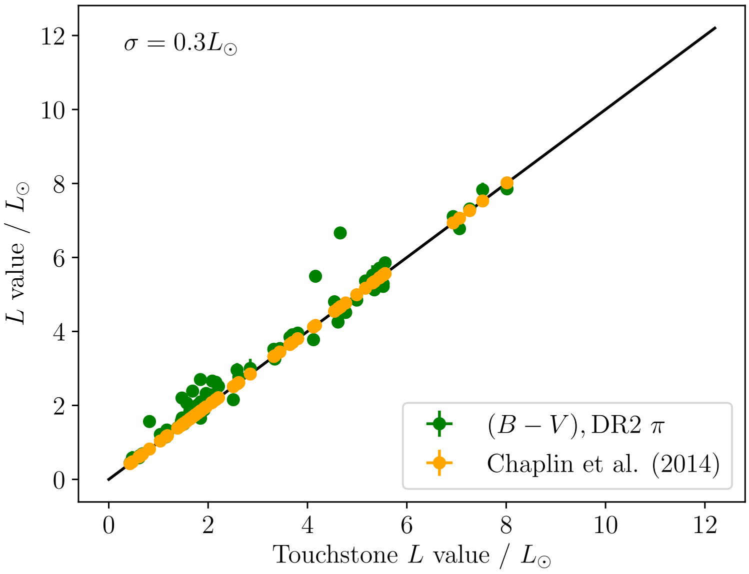


Figure 3.9: The literature 'touchstone' L values, compared to scaling relation values using parallaxes from DR2. The green dataspoints from scaling relations have errorbars on them from Monte Carlo simulations. However, the errorbars are so small that they do not appear on the plot.

(Chapter 2).

The inferior precision of the XHIP and TGAS parallaxes restricts their use for some stars. They will be tested in more detail in Section 3.4. If the XHIP and TGAS parallaxes are found to disagree with those from DR2, this can explain the deviation seen between the predicted and literature values of luminosity in Figures 3.7 and 3.8.

After luminosity, the radius values of the LEGACY stars were compared to those from Chaplin et al. (2014a). This was done by first calculating the luminosities and effective temperatures of the stars, before using the Stefan-Boltzmann law to infer stellar radii. This comparison of stellar radii with literature is a natural extension of Sections 3.3.1 and 3.3.2.

3.3.3 Stellar Radii

In the previous section, stellar luminosities of the LEGACY dataset were calculated from scaling relations. This parameter can be calculated using just (B-V) colour, V-band apparent magnitude, and a measure of the stellar parallax. Different parallaxes were used to calculate the luminosities for the main sequence stars. The first source of these parallaxes was the Hipparcos satellite (with the XHIP catalogue). The second parallax source was the *Gaia* satellite (which produced the TGAS and DR2 catalogues).

Section 3.3.2 showed that the source of parallax affects the luminosities from scaling relations. This section develops that idea: using effective temperatures from Section 3.3.1 and luminosities from Section 3.3.2, will predictions of stellar radii also be affected by the source of parallax?

First, the XHIP catalogue was used to generate luminosities and effective temperatures for the stars. Stellar radii can then be calculated with the Stefan-Boltzmann law and compared to literature values. Figure 3.10 shows that once again, the Hipparcos parallaxes underperformed when used in the scaling relations. A large scatter is seen in the results from scaling relations, and is due to the parallax precision of these stars in the XHIP catalogue.

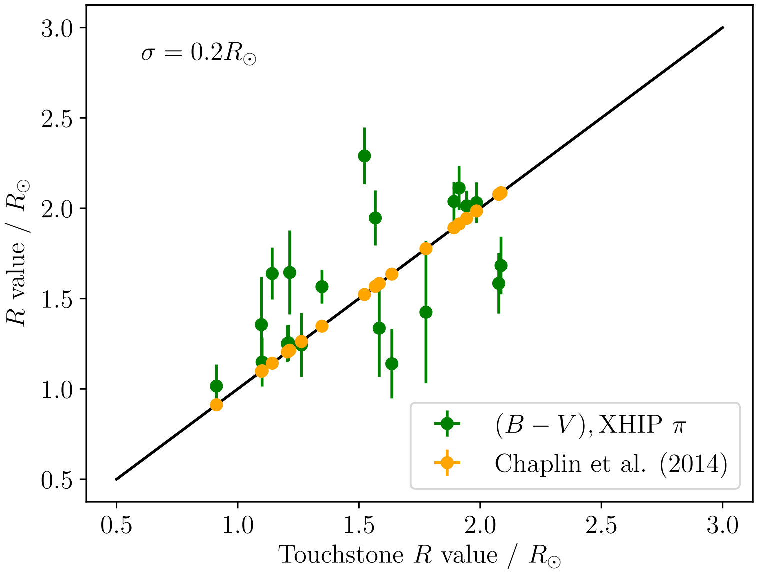


Figure 3.10: The literature 'touchstone' R values (orange), compared to scaling relation predictions using parallaxes from XHIP.

The same process of calculating radii from parallaxes and colours was repeated. This time, parallaxes were taken from the TGAS catalogue (Figure 3.11). The standard deviation between the scaling relations using TGAS parallaxes and those from literature is large. Rather than scattering radii above and below the 'true' values (as in Figure 3.10), TGAS parallaxes cause a slight systematic shift towards larger radii.

When reviewing the results from Figure 3.11, it is helpful to consult previous work done with the radius scaling relation. Huber et al. (2017) performed analysis on 2200 *Kepler* stars, and compared results to those from scaling relations using TGAS parallaxes. The spread in radius values in figure 3.11 is commensurate with results presented in that work.

Thirdly, (B-V) colours and parallaxes from DR2 were used to calculate the radii of the LEGACY stars. Figure 3.12 shows the results. Once more, the results show that using more precise parallax values leads to improvements in the calculated LEGACY parameters.

Even when precise parallaxes from DR2 are used, there is a slight bias towards larger stellar radii (Figure 3.12). This will likely cause a bias towards lower $\nu_{\max}$ values when using the scaling relations.

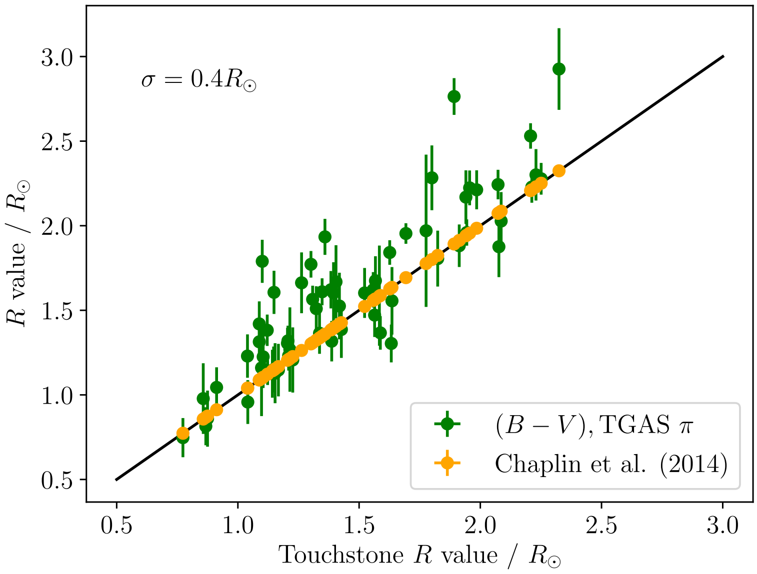


Figure 3.11: The literature 'touchstone' R values (orange), compared to scaling relation predictions using parallaxes from TGAS.

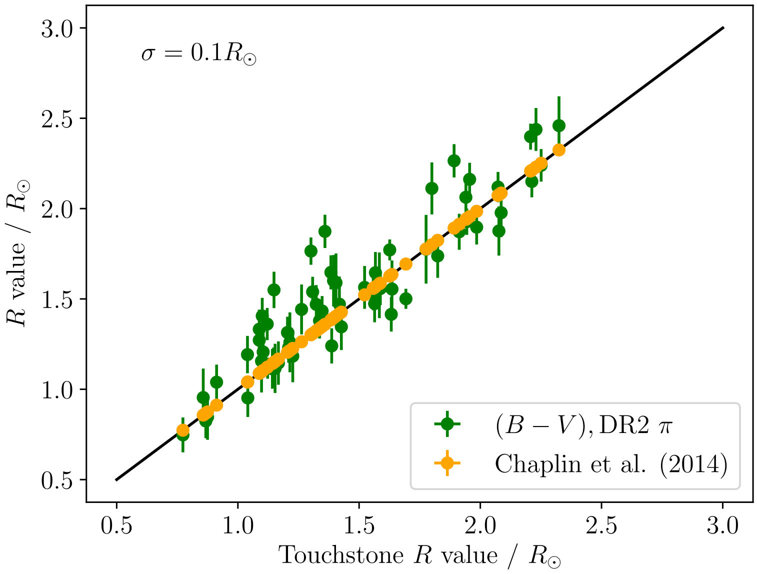


Figure 3.12: The literature 'touchstone' R values (orange), compared to scaling relation predictions using parallaxes from DR2.

3.3.4 The frequency of maximum oscillation power, $\nu_{\max}$

In Section 3.3, results from scaling relations are compared to those from scaling relations. These scaling relations use only (B-V) colour and parallax in order to calculate T_{eff}, L and R. This can also be done using the frequency of maximum oscillation power, $\nu_{\max}$. For consistency, all of the results from the scaling relations in Section 3.3 are compared to the same source of parameters (Chaplin et al. 2014a; C14).

Section 3.2 showed scaling relations to calculate the properties of the solar-like oscillation envelope. These use only the global stellar parameters. The equation to calculate $\nu_{\max}$ using radius and effective temperature was tested here, and the results are shown. As in all of Section 3.3, $\nu_{\max}$ values from scaling relations were calculated with parallaxes from different sources.

Firstly, Figure 3.13 shows the calculated values of $\nu_{\max}$ using XHIP parallaxes. Once again, values from the scaling relations were compared to those from literature (C14). Since $\nu_{\max}$ is an oscillation parameter, L17 provides the fitted $\nu_{\max}$ values for these LEGACY stars.

The values from L17 are also shown on the plot, however they agree so well with C14 that they cannot be seen! C14 was used as the source of 'true' stellar parameters throughout Section 3.3. The values from L17 agree very well with those from C14. This leads to the conclusion that C14 is a reliable source of stellar parameters, and can be used throughout this Chapter to compare results to.

Figure 3.14 also shows the $\nu_{\max}$ values of the LEGACY stars. Here, the $\nu_{\max}$ values were calculated using scaling relations with TGAS parallaxes. Once again, the scaling relation provides an under-prediction of the frequency of maximum power.

Thirdly, $\nu_{\max}$ was calculated for the LEGACY stars using DR2 parallaxes. Once again, scaling relation 3.2 under-predicts the $\nu_{\max}$ values of the stars. The offset from the 'true' $\nu_{\max}$ values is less severe than in Figures 3.13 and 3.14. Furthermore, the scatter about the 1:1 trend line is approximately half as much as seen in Figure 3.14.

From Figures 3.13, 3.14 and 3.15, all three sources of parallax under-predicted the

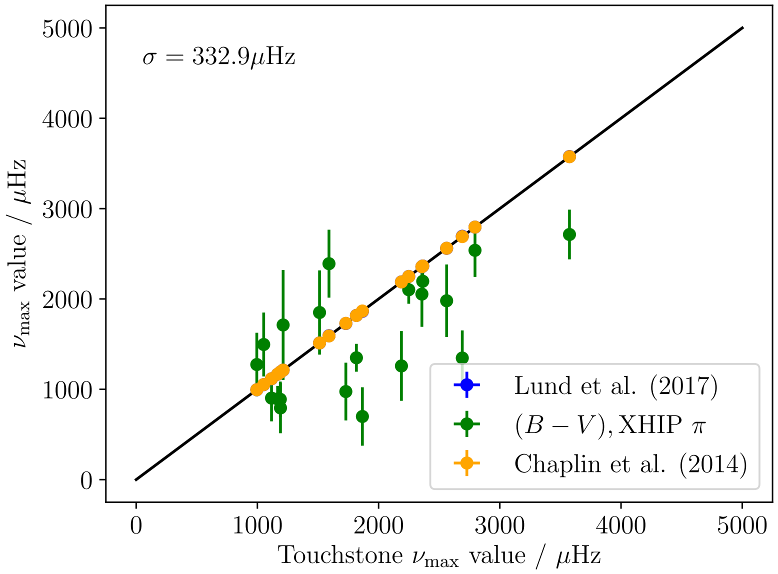


Figure 3.13: The literature 'touchstone' $\nu_{\rm max}$ values (orange), compared to scaling relation predictions using parallaxes from XHIP. The predictions from fits by Lund et al. (2017) are plotted in blue, and agree almost perfectly with those from Chaplin et al. (2014a).

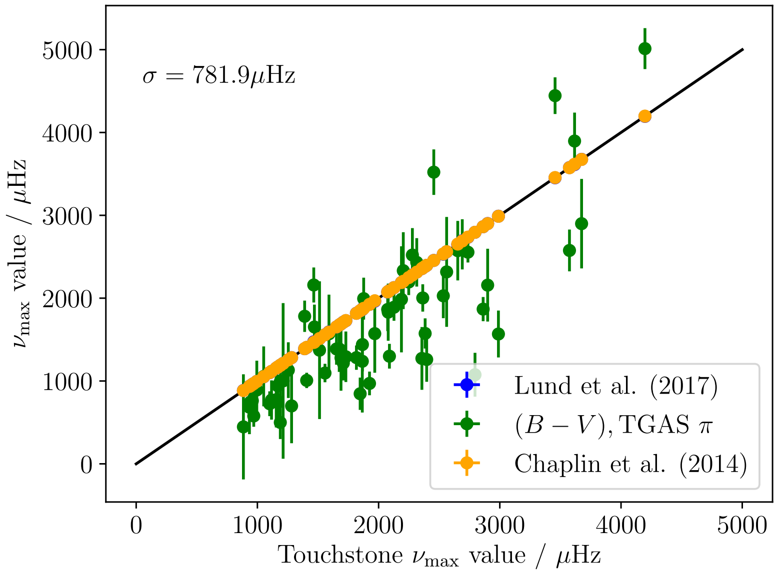


Figure 3.14: The literature 'touchstone' $\nu_{\rm max}$ values (orange), compared to scaling relation predictions using parallaxes from TGAS. The predictions from fits by Lund et al. (2017) are plotted in blue, and agree almost perfectly with those from Chaplin et al. (2014a).

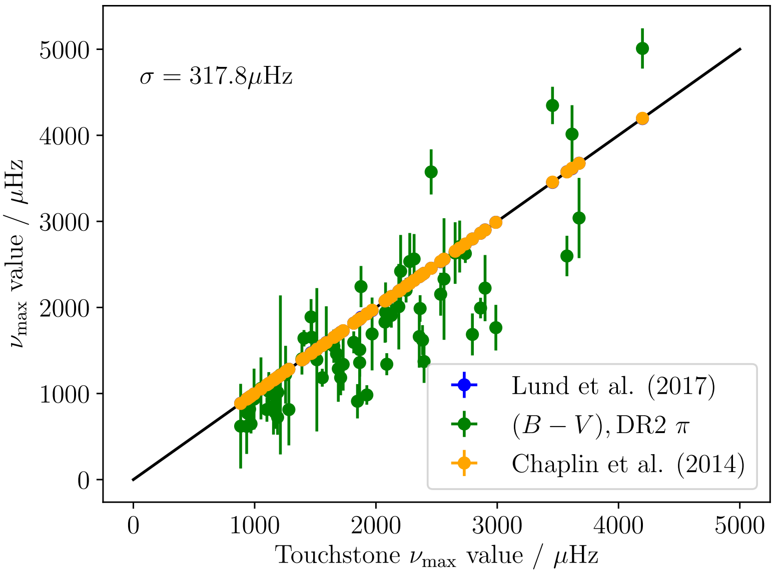


Figure 3.15: The literature 'touchstone' $\nu_{\rm max}$ values (orange), compared to scaling relation predictions using parallaxes from DR2. The predictions from fits by Lund et al. (2017) are plotted in blue, and agree almost perfectly with those from Chaplin et al. (2014a).

$\nu_{\rm max}$ values of the stars, compared to results from C14 and L17. As previously discussed, the stellar radii are systematically over-estimated (Section 3.3.3). This was expected to lead to the $\nu_{\rm max}$ values of the LEGACY stars being underestimated by equation 3.2. This is exactly what Figures 3.13, 3.14 and 3.15 show.

In this section, the stellar parameters $T_{\rm eff}$, luminosity, radius and $\nu_{\rm max}$ were calculated from scaling relations. These used only (B-V) colours (Cutri et al., 2003) and parallax. Parallax values from three catalogues were compared; from the XHIP (Anderson & Francis, 2012), TGAS (Michalik et al., 2015) and DR2 (Lindegren et al., 2018) catalogues.

This section has shown that Torres (2010) provides a reliable way to estimate effective temperature. Secondly, the luminosity equation given in Pijpers (2003) (equation 3.10) can be used with DR2 parallaxes to provide reliable stellar luminosities. Together, these temperatures and luminosities provide accurate stellar radii through the Stefan-Boltzmann law. However, it was shown that there is a slight tendency for the stellar radii to be over-estimated. Lastly, Section 3.3.4 showed that over-estimating the stellar radii leads to under-estimated $\nu_{\rm max}$ values.

The comparisons to literature also show that parallaxes from DR2 produce more accurate global stellar parameters than the values from the XHIP and TGAS catalogues. In the next section, the parallaxes from these three catalogues will be compared directly.

3.4 Comparing parallax and distances sources

So far, this Chapter has been focussed on testing the relations used to construct the ATL. Relations to calculate the stellar envelope (Section 3.2) and global stellar parameters (Section 3.3) have both been tested.

In both of these Sections, the LEGACY dataset was used to compare the results from equations to those from literature (C14). In this section, the LEGACY stars are used to analyse another crucial part of TESS asteroseismic target selection; the accuracy of parallax and distance values in three different interferometric catalogues.

As in the last Section, parallaxes from the XHIP, TGAS and DR2 catalogues were used here. Two distance catalogues were also compared to one another: Astraatmadja & Bailer-Jones (2016) gives distances for the stars in TGAS, while Bailer-Jones et al. (2018) provides distances for stars in the DR2 catalogue. In the plots that follow, the dashed line shows the 1:1 trend line; perfect agreement between the parallax or distance values from different sources.

Since LEGACY stars are used to compare parallax and distance values here, conclusions cannot be drawn about the quality of data in the complete XHIP, TGAS and DR2 catalogues. The work in this section was done to test the method used to construct the ATL. The ATL consists of bright, solar-like stars. The LEGACY dataset also consists of bright, solar-like stars, so it is an acceptable choice for a dataset to use for the comparisons in this section.

Firstly, parallaxes between the TGAS and XHIP catalogues are compared to one another (Figure 3.16). The Figure shows that the parallax values in the XHIP catalogue are slightly larger than the parallax values in the TGAS catalogue. To better understand which source is more accurate, the catalogues need to be compared to another

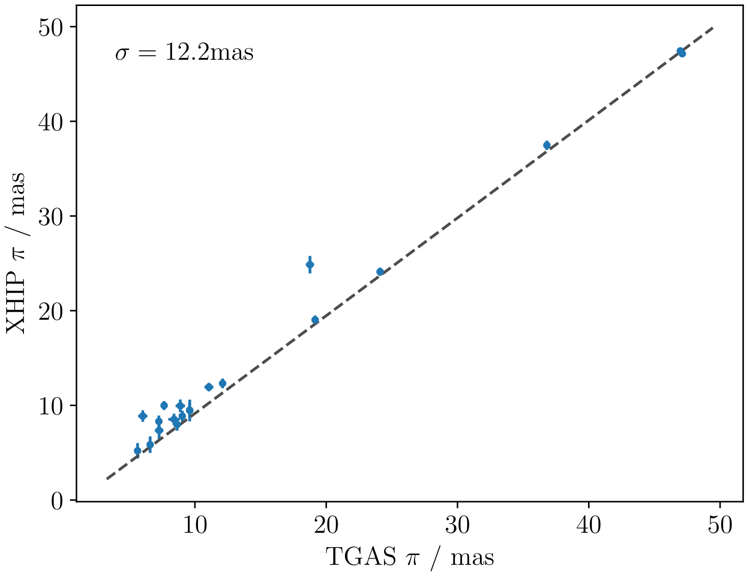


Figure 3.16: Parallaxes from the XHIP catalogue are compared to those from the TGAS catalogue. For these main sequence stars, parallax values in the XHIP catalogue are slightly larger than the corresponding parallax values in the TGAS catalogue.

interferometric dataset.

A similar comparison was then done between the XHIP and DR2 catalogues in Figure 3.17. The parallaxes in the XHIP catalogue are larger than those in DR2. The standard deviation quoted on the Figure shows that this difference is significant; the fractional difference in values lies between 50 - 100% for the LEGACY stars.

Thirdly, Figure 3.18 shows the parallax comparison between the TGAS and DR2 datasets. The standard deviation between parallax values is given in the corner of the plot. It shows that there is a significant difference between the two catalogues. The later DR2 catalogue release used longer timeline data, compared to TGAS. This longer timeline has led to more precise parallaxes in DR2, compared to those in TGAS.

Lastly, different sources of distance for these solar-like stars are compared directly in Figure 3.19. The distance catalogues that were compared both used the same method to infer values: with an exponentially-decreasing space density prior (Astraatmadja & Bailer-Jones, 2016). The distances used throughout this thesis are those calculated from the median of the Milky Way prior. This prior was used when distances were calculated

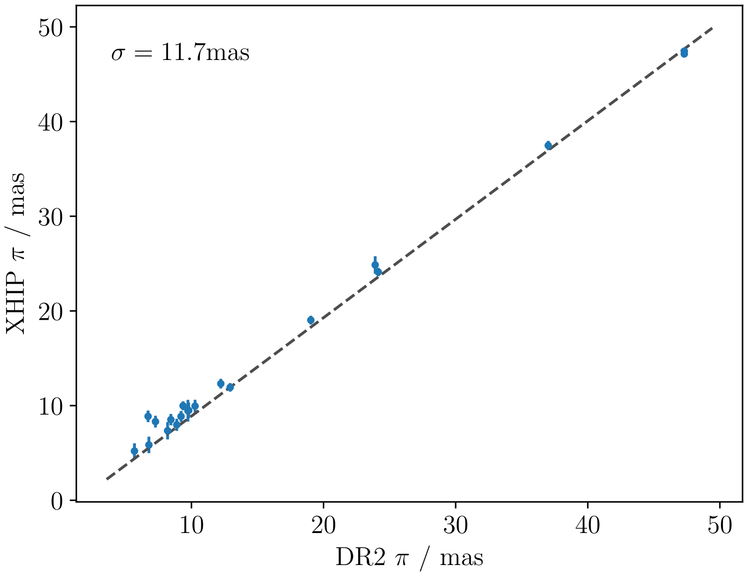


Figure 3.17: Parallaxes from the DR2 and XHIP catalogues are compared. The parallaxes in XHIP are larger than those in DR2.

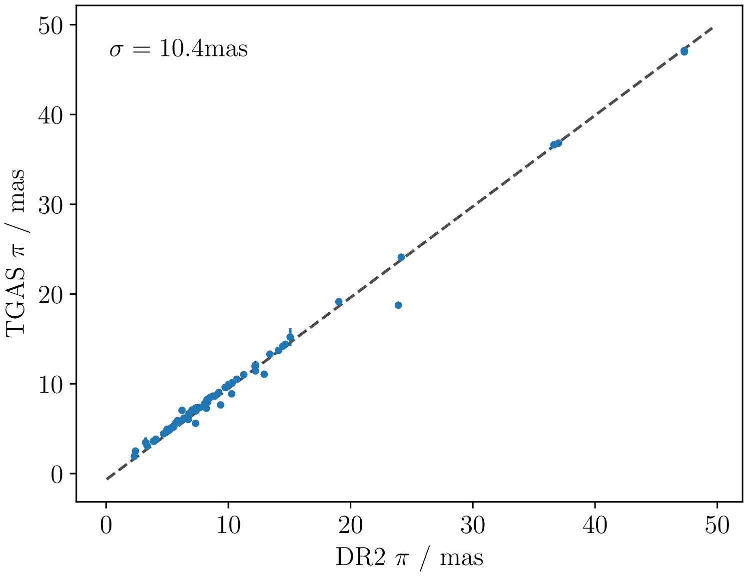


Figure 3.18: Parallaxes from the DR2 and TGAS catalogues are compared. At small parallaxes (i.e at larger distances), the TGAS catalogue appears to slightly over-predict the parallax values of the LEGACY stars, compared to DR2.

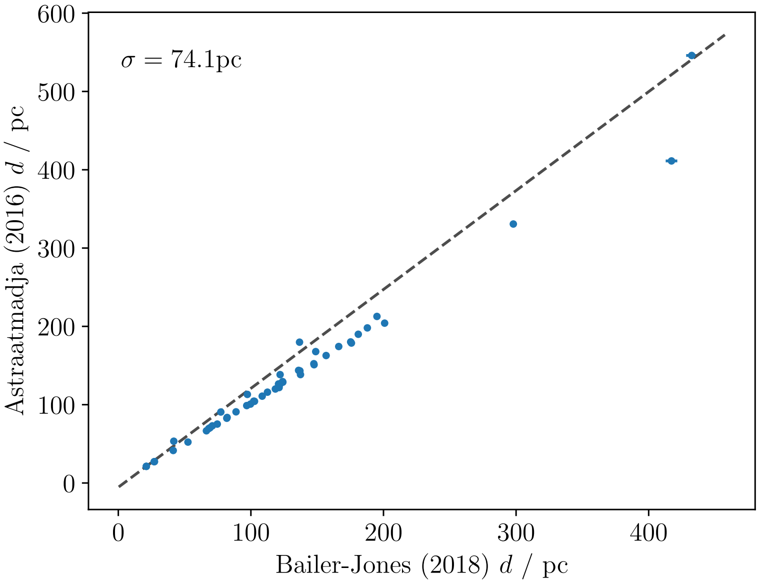


Figure 3.19: Distances are compared between those inferred from the TGAS and DR2 catalogues (Astraatmadja & Bailer-Jones 2016 and Bailer-Jones et al. 2018, respectively). Distance values in Astraatmadja & Bailer-Jones (2016) are systematically smaller than those in Bailer-Jones et al. (2018).

from parallax by inversion.

The catalogues being compared in Figure 3.19 are from Astraatmadja & Bailer-Jones (2016) (created using TGAS) and Bailer-Jones et al. (2018) (created using DR2). The improved quality of the data in DR2 has led to a significant difference in the inferred distances, compared to those inferred from TGAS.

To summarise, Section 3.3 found that when calculating stellar parameters from scaling relations, DR2 parallaxes and distances provide significantly better results for the solar-like stars used here. Because of this, DR2 was used as ‘true’ parallax and distance source in this section. This section has found that parallax and distance values vary significantly between catalogues.

When producing the ATL, DR2 parallaxes were used. Alongside these, distances from Bailer-Jones et al. (2018) were adopted. The results shown in this section confirm these decisions.

3.5 An updated envelope width relation

The LEGACY sample was used to test relations for the envelope properties (Section 3.2) and the global stellar properties (Section 3.3). The LEGACY stars were then used to determine the differences between parallaxes in the XHIP, TGAS and DR2 catalogues, as well as differences in the distances that were inferred from them (Section 3.4).

In Section 3.2, the equation to calculate the envelope width from Stello et al. (2007) (equation 3.3) was shown. It was found to over-estimate the width of the solar-like mode power excess. In this section, an updated envelope width relation is presented. This relation was tested using the LEGACY dataset.

Envelopes of different shapes were fitted to power spectra. The envelopes were fitted to the power excess due to solar-like modes of oscillation in those power spectra. Fits were made for the 66 LEGACY stars in Lund et al. (2017). The envelopes were fitted with Gaussian, Lorentzian and Voigt functions (a Voigt function is a convolution of Gaussian and Lorentzian profiles; Olivero & Longbothum 1977).

Envelopes were fitted to the LEGACY stars in order to better characterise the properties of this oscillation power excess. One of these fitted parameters is the Full Width at Half Maximum (FWHM) of the solar-like oscillation envelope, Γ_{env}.

Lund et al. (in prep) calculated values of Γ_{env} for the LEGACY stars using several functions, fitted to individual mode frequencies and power values. Of the functions fitted to the power spectra, the Lorentzian offered the best overall fits to the data. As a result, the work in this section used the Γ_{env} values from the Lorentzian fits from Lund et al. (in prep).

Until now, scaling relations to calculate the envelope width have been given as a function of ν_{max} alone (e.g Stello et al. 2007 and Mosser et al. 2010, 2012). It was thought that envelope width may scale as a function of effective temperature T_{eff}, surface gravity $\log(g)$ and the frequency of maximum power ν_{max}. This was suspected because these properties are all related to each other. Namely, $\nu_{\text{max}} \propto \log(g) / \sqrt{T_{\text{eff}}}$ (Chapter 1).

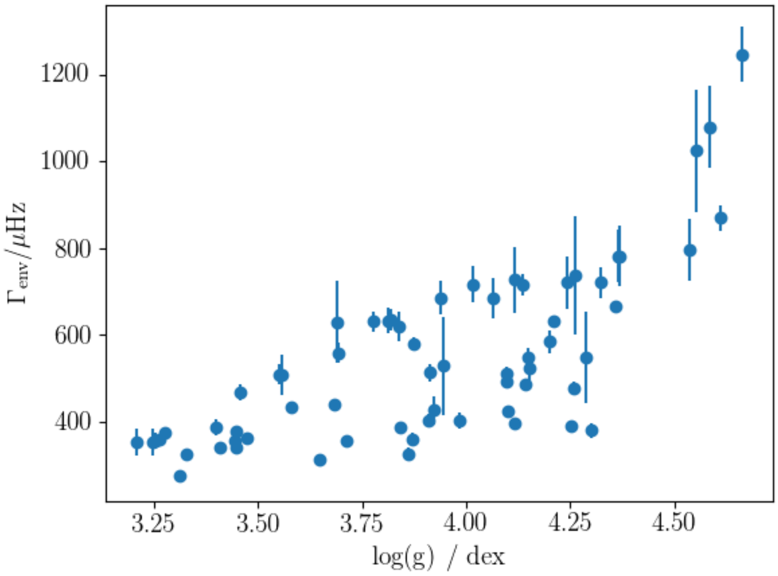


Figure 3.20: The envelope width as a function of $\log(g)$. A larger surface gravity leads to a wider mode-envelope. Although this trend is present, the correlation is not tight.

The aim of this work was to determine the relationship between T_{eff}, $\log(g)$, $\nu_{\max}$ and Γ_{env}. An improved scaling relation to calculate envelope width can then be produced.

To determine the relationship between $\log(g)$ and Γ_{env}, the fitted envelope widths of the LEGACY stars were plotted against $\log(g)$ (Figure 3.20). Although envelope width does appear to correlate with $\log(g)$, Γ_{env} appears to become degenerate to $\log(g)$ at values below 4.25 dex.

The corresponding relationship between Γ_{env} and $\nu_{\max}$ was then tested. The literature already contains several Γ_{env} equations using this parameter. Firstly, Stello et al. (2007) published an envelope width relation, which was given in equation 3.3. Alongside this, Mosser et al. (2012) also published a Γ_{env} relation as a function of $\nu_{\max}$:

$$\Gamma_{\mathrm{env}} = 0.66\,\nu_{\max}^{0.88}\,. \tag{3.12}$$

Equation 3.12 was calibrated for red giants, so it is not immediately obvious that it would apply to main sequence stars.

Figure 3.21 shows equations 3.3 and 3.12 with this LEGACY data. Equation 3.12

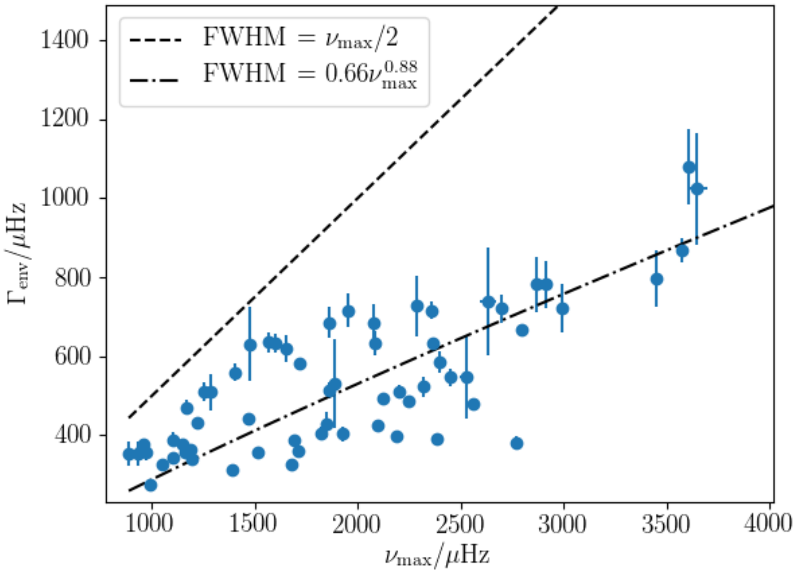


Figure 3.21: The LEGACY stars plotted as a function of $\nu_{\rm max}$. Relations from Stello et al. (2007) and Mosser et al. (2012) have been over-plotted. For these Main-Sequence stars, the relation from Mosser et al. (2012) matches the data more closely, although there is still significant scatter as stars approach lower $\nu_{\rm max}$ values.

replicates the envelope widths of these stars more closely than equation 3.3, although the data is still scattered around this line. Equation 3.12 was therefore used to investigate an updated $\Gamma_{\rm env}$ relation further.

Thirdly, a relationship between $\Gamma_{\rm env}$ and $T_{\rm eff}$ was investigated (Figure 3.22). As temperature increases, envelope width is seen to decrease. Figure 3.22 shows that envelope width is a function of $T_{\rm eff}$, but not $T_{\rm eff}$ alone.

This led to a relation which uses both $T_{\rm eff}$ and $\nu_{\rm max}$, i.e $\Gamma_{\rm env}(T_{\rm eff}, \nu_{\rm max})$. This relation is shown in Figures 3.23 and 3.24. A piecewise relation was fitted to Figure 3.24. Below a temperature of 5600 K, equation 3.3 was fitted to the data. Above this, a line of best fit to the data was used, which involved $T_{\rm eff}$;

$$\Gamma_{\rm env} = \begin{cases} 0.66\,\nu_{\rm max}^{0.88}\,(1 + (T_{\rm eff} - T_{\odot}) \times 6 \times 10^{-4}) & \text{if } T_{\rm eff} > 5600\,\text{K} \;, \\ 0.66\,\nu_{\rm max}^{0.88} & \text{if } T_{\rm eff} \leqslant 5600\,\text{K} \;. \end{cases} \tag{3.13}$$

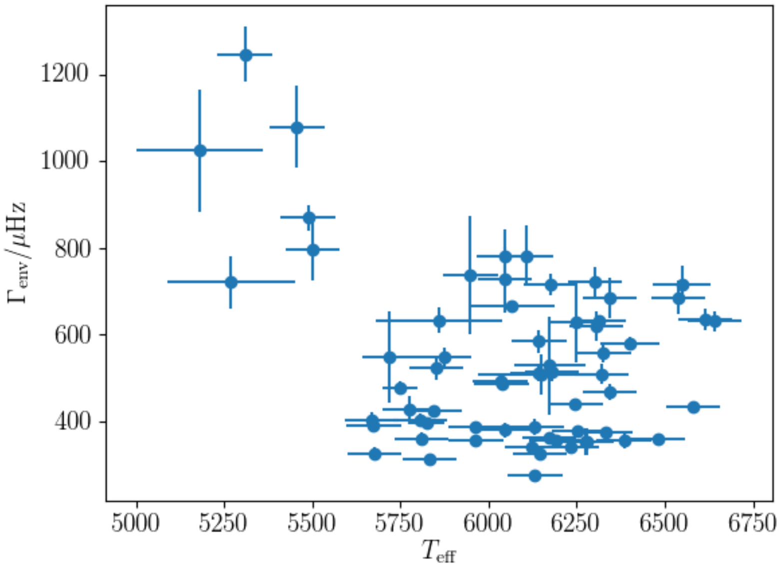


Figure 3.22: A plot of envelope width, as a function of $T_{\rm eff}$. Effective temperature does appear to correlate with envelope width.

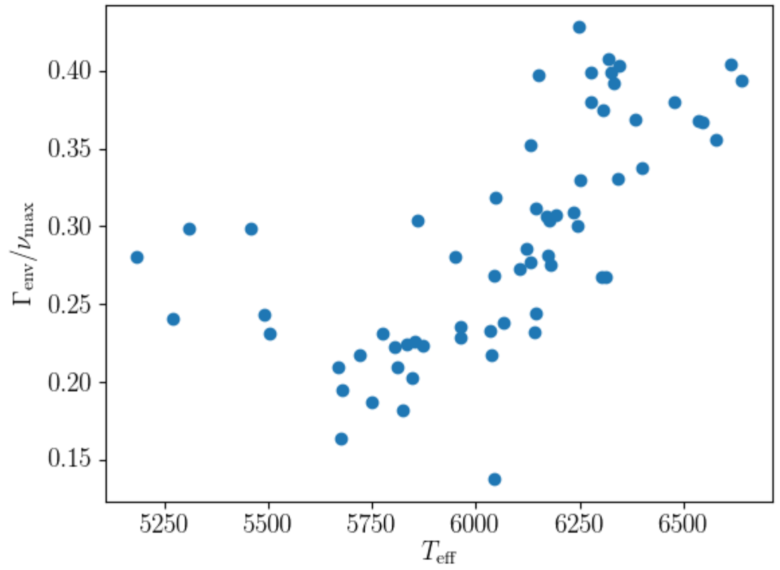


Figure 3.23: A plot of envelope width divided by $\nu_{\rm max}$, as a function of $T_{\rm eff}$. For the majority of the LEGACY stars, $\Gamma_{\rm env}$ appears to scale as a function of $T_{\rm eff}$.

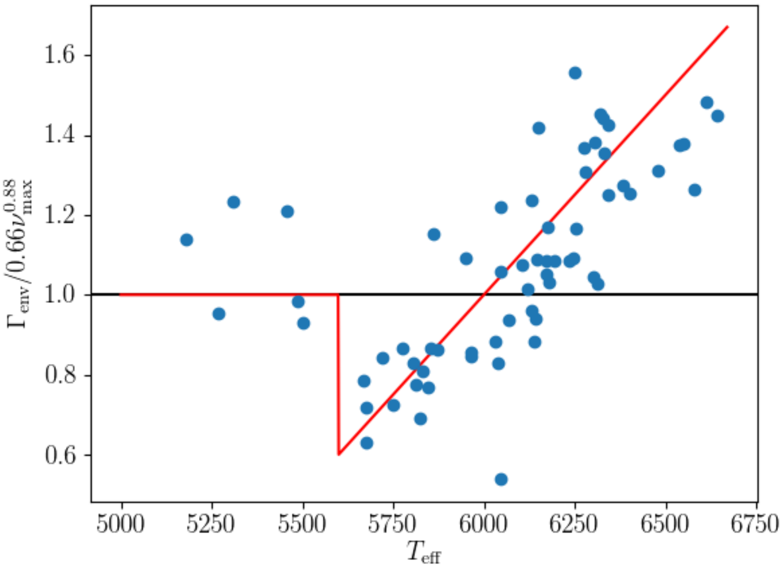


Figure 3.24: A plot of envelope width divided by equation 3.12, as a function of T_{eff}. A piecewise relation (red) was fitted to the LEGACY data. Below a temperature of 5600 K, equation 3.12 was fitted to the data. Above this, a temperature-dependent line of best fit was used.

Figure 3.24 shows that equation 3.13 describes the envelopes of these main sequence stars better than equation 3.12 alone. The piecewise relation is far from perfect. Although it is an improvement, scatter is still seen between the relation and the ‘true’ LEGACY envelope widths. Because of this scatter, equation 3.13 was not used when constructing the ATL itself in Chapter 2.

Additionally, there is a sparse amount of data at lower effective temperatures in the LEGACY sample. In order to improve the relation further, the behaviour of envelope width would need to be better understood at these lower temperatures. Equation 3.13 may also slightly over-estimate the envelope width around the solar effective temperatures (5777 K). More fitted data would confirm this.

3.6 Conclusion

In this Chapter, the method used to generate the Asteroseismic Target List was tested. It was tested using the *Kepler* LEGACY sample from Chaplin et al. (2014a) (C14). This

LEGACY sample is a collection of 66 main sequence stars with photometry spanning several years. C14 presented the stellar parameters of these 66 solar-like oscillators.

Further work was done on the LEGACY stars by Lund et al. (2017) and Aguirre et al. (2017) (L17). They fitted the individual solar-like oscillations of each main sequence star. They present the following individual fitted mode parameters: the mode frequencies ν, heights H and linewidths Γ.

In this Chapter, Gaussian curves were fitted to the radial mode frequencies and heights from L17. The height, width and central abscissa of these Gaussian fits were used as the 'true' values of the mode envelope; H_{env}, Γ_{env} and ν_{max}, respectively. In order to obtain 'true' values of the large frequency spacing $\Delta\nu$, linear fits were performed on the radial mode frequencies and overtone numbers from L17. The gradients of these straight lines were used as the 'true' values of $\Delta\nu$ in this Chapter.

Firstly, the relations which describe the mode power-excess were tested in Section 3.2. These 'envelope relations' were tested on the *Kepler* LEGACY stars. Results were compared to 'true' values of $\Delta\nu$, H_{env}, Γ_{env} and ν_{max} from fits to the 'peak bagged' modes from L17.

In Section 3.2, the envelope parameters of the LEGACY stars were calculated with scaling relations, using literature values of radius and effective temperature from C14. These envelope parameters were compared to the 'true' values from fits performed to L17 data. Additionally, literature values of (B-V) (Cutri et al., 2003) and parallax (Anderson & Francis, 2012) were used to calculate effective temperatures and radii of the stars, before calculating the envelope parameters of the LEGACY dataset with the scaling relations.

These comparisons showed that using (B-V) and parallax alone could introduce significant uncertainty when calculating the properties of the solar-like envelope. Conclusions were also drawn about individual scaling relations used to calculate these envelope parameters. It was found that equation 3.2 under-estimates the ν_{max} value of stars. Additionally, equation 3.3 was shown to systematically over-estimate the oscillation envelope width of a star. These results were expanded upon in Section 3.3.

In Section 3.3, the global scaling relations were tested using LEGACY stars. Namely, the scaling relations used to calculate radius R, effective temperature T_{eff}, stellar luminosity L and the frequency of maximum power ν_{max} were all tested.

These tests were performed by comparing literature values of the global parameters from C14, to results from scaling relations using only (B-V) and colour. To determine if the source of parallax affected the results, three different parallax sources were used. Parallaxes were taken from the XHIP (Anderson & Francis, 2012) TGAS (Michalik et al., 2015) and DR2 (Lindegren et al., 2018) catalogues. It was found that when parallaxes from DR2 were used, there was significant improvement to the accuracy of the results.

The work in Section 3.3 assessed the accuracy of the scaling relations that were used to construct the ATL (Chapter 2). Section 3.3 concluded that the source of parallax used significantly influences predictions from the scaling relations.

After this, Section 3.4 presented the results of direct comparisons between different parallax and distance sources using the LEGACY stars. Once more, results from the XHIP, TGAS and DR2 interferometric catalogues were compared. Additionally, distances inferred from the TGAS and DR2 catalogues were also compared (Astraatmadja & Bailer-Jones 2016 and Bailer-Jones et al. 2018, respectively).

It was found that there are significant differences between the parallax and distance values of the LEGACY stars in these catalogues. When constructing the ATL, DR2 parallaxes were used, along with distances from Bailer-Jones et al. (2018). The results in Sections 3.3 and 3.4 support this decision.

Lastly, an updated envelope width relation was presented in Section 3.5. This work was motived by the results from Section 3.2, where it was found that the envelope relation from literature (equation 3.3) over-estimates the width of the power excess due to solar-like oscillations.

To produce this updated relation, tests were performed using the fitted envelope widths of the LEGACY stars from L17. It was found that the solar-like mode envelope width could be better described using both ν_{max} and effective temperature, in a piecewise relation.

The following Chapters will evolve from making predictions of stellar parameters, to using detailled fitted solar-like oscillations to advance the field of asteroseismology.

In Chapter 4, fitted solar-like oscillations will be used to develop analytical equations to calculate the uncertainty of the global seismic parameters $\Delta\nu$ and ν_{max}. In Chapter 5, an automated pipeline to select solar-like oscillators for future missions is developed. Rather than using a series of scaling relations (as in Chapter 2), this automated pipeline learns connections between parameters using machine learning. It can be used to select stars for any future mission, without the need to construct an Asteroseismic Target List.

Chapter 4

Analytical $\Delta\nu$ and $\nu_{\rm max}$ Uncertainty Equations

> The work presented in this Chapter is entirely my own. The analytical equations are tested using data from Santos et al. (2018).

4.1 Introduction

In Chapter 2, a list of stars that contain detectable solar-like oscillations with TESS was presented. The method used to select these stars was tested using Monte Carlo simulations. To perform the Monte Carlo simulations, the uncertainty on the global asteroseismic parameters $\Delta\nu$ and $\nu_{\rm max}$ needed to be known.

Here, equations to calculate $\sigma_{\nu_{\rm max}}$ and $\sigma_{\Delta\nu}$ are presented. Alongside being able to produce high-fidelity simulations from artificial populations, being able to predict $\sigma_{\Delta\nu}$ and $\nu_{\rm max}$ using analytical equations accurately is of great use in the field of asteroseismology, including in stellar modelling and galactic archaeology (e.g Yu et al. 2016).

The work done in this Chapter used *Kepler* LEGACY stars (Aguirre et al. 2017, Lund et al. 2017). The time series' of these stars were separated into sub series and subsequently 'peak-bagged' in Santos et al. (2018) (hereafter S18). S18 fitted the frequencies, uncertainties, heights and linewidths of the solar-like modes within every 90-day sub series, across the stars.

The time series' of these 10 main sequence stars were divided into several 90-day long segments by S18. Rather than having one long time series for each star, each star now had 20-23 shorter 90-day long time series'. Each 90-day long time series is known as a sub series. S18 had fitted 20 sub series' for some of the LEGACY stars, while other stars had up to 23 fitted sub series'. The number of sub series' for a star varied depending on the quality of the original time series data, and the total duration of observation.

Having several different sub series' for a single star is useful because each of these contains a different realisation of the underlying χ^2 2 DOF noise. Different measurements of a single parameter (for example, the linewidth of a single mode) can be made for each star. Measuring one parameter in a single star across several different sub series' will give different results. This is because the χ^2 2 DOF noise level in each sub series is different. If enough different measurements of the same quantity are made (in this case, the linewidth of a mode), the distribution of results will tend towards a Gaussian centred about the true value of that quantity (the 'true' mode linewidth). This is an example of the Central Limit Theorem.

Separating the timeseries of a star is especially useful here because each sub series can be treated as a separate observation. The analytical $\sigma_{\nu_{\max}}$ and $\sigma_{\Delta\nu}$ equations presented here were tested using several 90-day sub series for a single star, rather than only once per star on the full-length timeseries. Assuming these parameters do not actually change with time, the scatter in $\sigma_{\nu_{\max}}$ and $\sigma_{\Delta\nu}$ robustly tests the analytical formulae.

Section 4.2 describes the LEGACY data: where the data came from, why the data were originally 'peak bagged', and how the data was used here. Steps were also taken in this work to prepare the data before calculating $\sigma_{\Delta\nu}$ and $\sigma_{\nu_{\max}}$. Section 4.2 details these steps.

Section 4.3 then presents analytical equations to calculate the uncertainty on $\Delta\nu$ and $\nu_{\max}$. One equation to determine $\sigma_{\nu_{\max}}$ is derived and presented here. Two equations to calculate $\sigma_{\Delta\nu}$ are also given. The differences between the two $\sigma_{\Delta\nu}$ equations are explained.

The analytical equations in Section 4.3 must be tested against 'true' measures of

$\sigma_{\Delta\nu}$ and $\sigma_{\nu_{\max}}$ to determine if they work well on real data. The $\sigma_{\Delta\nu}$ and $\sigma_{\nu_{\max}}$ equations were tested by performing least-squares fits to the data from S18. These fits gave measures of $\sigma_{\Delta\nu}$ and $\sigma_{\nu_{\max}}$ that were treated as the 'true' values of these parameters. This fits are described in Section 4.4.

The 'true' uncertainties from Section 4.4 were compared against the uncertainties from the analytical equations from Section 4.3. The results of this comparison are given in Section 4.5. The results from individual stars across multiple sub series' are presented. Summary results are also given across all 10 LEGACY stars. Lastly, these results are summarised in Section 4.6.

4.2 The Dataset

The stars used in this work were observed during the nominal *Kepler* mission (Borucki et al., 2010). They are all stars which were first studied as part of the LEGACY dataset (Aguirre et al. 2017, Lund et al. 2017). This LEGACY dataset is a collection of 66 main sequence stars with photometry spanning several years. Lund et al. (2017) presented oscillation mode parameters for the 66 stars using the full-length photometry.

Santos et al. (2018) (S18) used the full-length lightcurves of the 66 LEGACY stars to look for signs of magnetic activity. To do this, S18 did not take the mode parameters given in Lund et al. (2017). Instead, they separated the full-length lightcurves of the 66 stars into 90-day sub series'. These 90-day sub series' were overlapped by 45 days. S18 fitted oscillation parameters for each sub series of the 66 stars. Fitting mode parameters is known as 'peak bagging'.

In order to perform the analysis done in this Chapter, S18 provided the 90-day sub series' of 10 LEGACY stars. The 'peak bagged' oscillation parameters for each sub series of the 10 stars were also given. Namely, S18 gave: individual radial mode frequencies ν, uncertainties σ_ν, heights H and linewidths Γ.

S18 provided several sub series' of each star. The number of sub series' of the 10 stars varied between 20 - 23 (due to issues with the quality of the oscillation spectra). In total, 224 90-day sub series' of the 10 stars were made available. The oscillation parameters of the 224 sub series' were used to calculate $\Delta\nu$ and ν_{max} uncertainties. This dataset provides ample mode parameters to test predictions of $\sigma_{\Delta\nu}$ and $\sigma_{\nu_{\mathrm{max}}}$.

Before the oscillation parameters from S18 could be used to calculate $\sigma_{\Delta\nu}$ and $\sigma_{\nu_{\mathrm{max}}}$, the data was prepared;

1. Firstly, the height of the Gaussian mode envelope at each oscillation frequency was calculated. This height comes from the mode power of a Lorentzian oscillation, $P = \pi H\Gamma/2$ (e.g. Houdek 2006). The power is scaled using the relative contributions of the four modes closest to an oscillation frequency, ζ. This power is contained within a frequency window of length $\Delta\nu$, which is divided out to give the gaussian height at a frequency:
$$H_{\mathrm{gauss}} = \zeta\,\frac{\pi}{2}\,\frac{H\Gamma}{\Delta\nu}\,. \tag{4.1}$$
H and Γ are the 'peak bagged' radial mode heights and linewidths, respectively. ζ is the sum of the mode visibilities $V_l^2 = [1, 1.5, 0.5, 0.03]$ for the angular degrees $l = [0, 1, 2, 3]$ as observed by *Kepler* (Ballot et al., 2011). $\Delta\nu$ is the large frequency spacing between modes of the same angular degree l and successive overtone number n.

   The mean $\Delta\nu$ value for each sub series was calculated using the individual radial mode frequencies. This process is also described in Chapter 3, where LEGACY data from L17 was used.

2. Secondly, modes that have a 'peak bagged' radial mode frequency uncertainty σ_ν greater than $2.5\,\mu$Hz were removed from the sample. This ensured that only well constrained oscillations were used to calculate $\sigma_{\Delta\nu}$ and $\sigma_{\nu_{\mathrm{max}}}$.

3. Thirdly, a straight line was fitted to the 'peak bagged' radial mode frequencies ν and overtone numbers n. The straight line fit provided 'true' values of the radial mode frequency spacings. The fit gave 'true' values of $\Delta\nu$ (the gradient of the line) and

$\sigma_{\Delta\nu}$ (the uncertainty of the gradient). The straight line fit is described in Section 4.4.1.

4. Lastly, a Gaussian was fitted to the H_{Gauss} values as a function of frequency (equation 4.1). This fit returned the 'true' parameters of the oscillation envelope: H_{env} (the height at the centre of the envelope), ν_{max} (the frequency of the centre of the envelope), Γ_{env} (the FWHM width of the envelope) and $\sigma_{\nu_{\text{max}}}$ (the uncertainty on the frequency at the centre of the envelope). These are the fitted envelope parameters of the LEGACY stars. This fit is described in detail in Section 4.4.2.

Each sub series of the 10 LEGACY stars was treated with these steps. The fitted $\sigma_{\Delta\nu}$ values from step 3 are the 'true' $\Delta\nu$ uncertainties. Similarly, the fitted $\sigma_{\nu_{\text{max}}}$ values from step 4 are the 'true' ν_{max} uncertainty.

These 'truth' values will be used to test predictions made by analytical equations. If the predictions from the equations match the 'true' values, then the equations are reliable and can be used elsewhere. These analytical equations are presented in Section 4.3.

4.3 Calculating $\Delta\nu$ and ν_{max} uncertainties using analytical equations

This Section presents derivations of the analytical equations used to calculate $\sigma_{\Delta\nu}$ and $\sigma_{\nu_{\text{max}}}$. Section 4.4 will then go on to explain how the results of these analytical equations were compared to least-squares analysis results from numerical simulations.

4.3.1 The first equation to calculate $\sigma_{\Delta\nu}$

The equation to calculate $\sigma_{\Delta\nu}$ comes from the uncertainty on the gradient of a linear fit of the radial-mode frequencies $\nu(n)$ to the radial order n. For a generic straight line

$$y = mx + c\,, \tag{4.2}$$

the uncertainty on the gradient m is

$$\sigma_m = \left(\frac{(N-2)^{-1} \sum_{i=1}^{N} (y_i - \hat{y})^2}{\sum_{i=1}^{N} (x_i - \bar{x})^2} \right)^{1/2} . \tag{4.3}$$

For example, see Miller (1991). The numerator of equation 4.3 is the reduced χ^2 value of the straight line, with $\hat{y}$ representing the weighted mean value of the dependent variable y. The weighted mean $\hat{y}$ can be used to calculate the uncertainty on the gradient σ_m if the uncertainties on each y value are known. For clarity, the y-residuals can be written in terms of the standard deviation of each data point about the weighted mean;

$$\sum_{i=1}^{N} (y_i - \hat{y})^2 = \sum_{i=1}^{N} \sigma_i^2 . \tag{4.4}$$

Equation 4.4 assumes that the y uncertainties σ_y are all drawn from a normal distribution with standard deviation σ. For our purposes, the y uncertainties are the σ_ν values of the strongest modes in the spectrum. These modes typically do have very similar uncertainties, so this is a valid assumption.

In our case, x_i are the radial mode overtone numbers n. y_i are the frequency values of the radial modes ν_i (each with frequency uncertainties σ_{ν_i}). The gradient m of this straight line fit is the large frequency separation $\Delta\nu$, with uncertainty $\sigma_{\Delta\nu}$. The equation for the uncertainty on this linear fit comes from substituting the variables into equation 4.3;

$$\sigma_{\Delta\nu} = \left(\frac{(N-2)^{-1} \sum_{i=1}^{N} \sigma_{\nu_i}^2}{\sum_{i=1}^{N} (n_i - \bar{n})^2} \right)^{1/2} . \tag{4.5}$$

4.3.2 The second equation to calculate $\sigma_{\Delta\nu}$

Using equation 4.5, the uncertainty of the gradient ($\sigma_{\Delta\nu}$) can be calculated straightforwardly. When deriving the equation, an assumption was made. Equation 4.5 assumes that the radial mode frequencies ν are all drawn from a Gaussian distribution with some standard deviation σ. In practice, the σ_ν values of the strongest radial modes are all similar, so this assumption is reasonable when considering radial modes in the envelope.

Ideally, we would not need to make any assumptions about the distribution of ν when calculating $\sigma_{\Delta\nu}$. This was the motivation behind deriving a second equation to calculate $\sigma_{\Delta\nu}$ analytically.

This second analytical equation was derived using Barlow (1989) and Miller (1991). As in Section 4.3.1, the second equation also comes from a straight line fit. Unlike in Section 4.3.1, here the σ_ν errors are not assumed to be normally distributed.

In general, the gradient estimate $\hat{m}$ can be found by rewriting the straight line equation $y(x)$ (equation 4.2) as a sum. Differentiating this sum with respect to m gives

$$\hat{m} = \sum_i \frac{x_i - \overline{x}}{N(\overline{x^2} - \overline{x}^2)} y_i \,, \tag{4.6}$$

where N is the number of x values. This leads to an expression of the variance on the gradient of the straight line,

$$V(\bar{m}) = \frac{\sigma^2}{N(\overline{x^2} - \overline{x}^2)} \,, \tag{4.7}$$

where σ^2 is the variance of the y values.

As a reminder, we are not assuming that the y uncertainties (σ_i) are uniform. In this general case, the uncertainty on the gradient m is obtained from a weighted straight line fit. Rather than being uniform, the uncertainties σ_i in equation 4.7 are described with

$$\overline{\sigma^2} = \frac{N}{\sum 1/\sigma_i^2} \,. \tag{4.8}$$

Equation 4.8 is the alternative to assuming uniform σ_i values, as in equation 4.4. Here, the distribution of uncertainties $\overline{\sigma^2}$ is explicitly given using each individual uncertainty σ_i.

Combining equations 4.7 and 4.8 gives an analytical expression for the uncertainty on the gradient of a weighted straight line. This gradient uncertainty takes into account the individual σ_i values, regardless of their distribution. The number of data points N cancels out from equations 4.7 and 4.8, leaving

$$\sigma_m = \sqrt{\frac{1}{(\sum 1/\sigma_i^2)(\overline{x^2} - \overline{x}^2)}}\,. \tag{4.9}$$

For our purposes, it is helpful to substitute notation into equation 4.9. In order to calculate $\sigma_{\Delta\nu}$, the straight line $\nu(n)$ must be fitted. The y values here are the radial mode frequencies ν. These frequencies have uncertainties σ_ν. x represents the radial mode overtone numbers n. m is the gradient of the fit $\Delta\nu$, with a corresponding uncertainty $\sigma_{\Delta\nu}$. Finally, we have

$$\sigma_{\Delta\nu} = \sqrt{\frac{1}{(\sum 1/\sigma_{\nu_i}^2)(\overline{n^2} - \overline{n}^2)}}\,. \tag{4.10}$$

4.3.3 The equation to calculate $\sigma_{\nu_{\mathbf{max}}}$

Equations to calculate the uncertainty on the global asteroseismic parameter $\Delta\nu$ were given in Sections 4.3.1 and 4.3.2. Here, an equivalent equation to calculate the uncertainty on the frequency of maximum oscillations power, $\nu_{\max}$, is presented.

This derivation has been adapted from work generalised by Libbrecht (1992), and expanded upon by Toutain & Appourchaux (1994). The same argument has been used here to obtain an equation for $\sigma_{\nu_{\max}}$.

In those works, an equation to determine the frequency uncertainty of a p-mode is obtained by using Maximum Likelihood Estimators (MLEs). MLEs are used to fit Lorentzian profiles to individual p-modes. MLEs work by determining the maximum likelihood of the parameters that are fitted. This maximum value in ‘normal’ likelihood space is equivalent to the minimum value in log-likelihood space (Toutain & Appourchaux, 1994).

The parameters of these Lorentzian fits are described using a multi-normal probability distribution. With this multi-normal probability distribution, the uncertainties on the parameters can be calculated from the inverse Hessian matrix. A Hessian matrix is a matrix of partial derivatives. The diagonal elements of this matrix give the uncertainties on the parameters of the Lorentzian fits.

In their work, Toutain & Appourchaux (1994) computed the expectation value for a Lorentzian mode in the power spectrum. This expectation value is the mode profile without χ^2 2 degrees-of-freedom noise.

They calculated the Hessian matrix of this mode expectation value, to give the uncertainties of the parameters in the Lorentzian fit. The Hessian matrix was calculated from the partial derivatives of the parameters. One of the elements of this Hessian matrix is $h_{\nu_0 \nu_0}$. This element of the matrix is equal to $\sigma_{\nu_0}^{-2}$; the inverse of the squared frequency uncertainty.

Equivalent steps to Toutain & Appourchaux (1994) were taken here. In this work, rather than isolating an individual p-mode, the entire Gaussian oscillation envelope in the power spectrum was treated. The expectation value of this power spectrum M is given by

$$M(\nu) = H_{\mathrm{env}} \exp\left(-\frac{(\nu - \nu_{\mathrm{max}})^2}{2\sigma_{\mathrm{env}}^2}\right) + B(\nu)\,. \tag{4.11}$$

The first term on the right hand side represents the Gaussian envelope due to solar-like modes of oscillation. For a single realisation, the mode envelope is approximated by a Gaussian. The second term on the right hand side represents the background due to shot noise, instrumental effects and granulation. $B(\nu)$ is assumed to be slowly varying across the frequency range occupied by the Gaussian mode envelope. H_{env} is the height of the envelope at its central frequency, ν_{max}. Lastly, σ_{env} represents the width of the Gaussian envelope.

$\sigma_{\nu_{\mathrm{max}}}$ can be calculated from equation 4.11 using a Hessian matrix. Specifically, the element $h_{\nu_{\mathrm{max}}\nu_{\mathrm{max}}}$ of the matrix is what we need: this is equal to $\sigma_{\nu_{\mathrm{max}}}^{-2}$. Computing this element of the Hessian matrix gives the precision of ν_{max};

$$\sigma_{\nu_{\mathrm{max}}}^{-2} = T \int \frac{1}{M(\nu)^2} \left[\frac{\partial M(\nu)}{\partial \nu_{\mathrm{max}}}\right]^2 d\nu\,. \tag{4.12}$$

T is the observation time. In our case, T is 90 days; the length of each sub series.

The integral in equation 4.12 can be solved analytically (see Brandt 1970 Section 7, with the solution quoted in Libbrecht 1992). Using the notation consistent with this chapter, the solution of the integral is

$$\sigma_{\nu_{\max}} = \sqrt{\left(\frac{\Gamma_{\mathrm{env}}}{4T\sqrt{\pi^3 \ln(2)}}\right) f(\beta)}\,. \tag{4.13}$$

Γ_{env} is the Full Width at Half Maximum height of the Gaussian mode envelope, and is related to σ_{env} by

$$\Gamma_{\mathrm{env}} = \sigma_{\mathrm{env}}\,(2\sqrt{2\ln(2)})\,. \tag{4.14}$$

β is the inverse signal-to-noise ratio; the background at $\nu_{\max}$, $B_{\max}$, divided by the mode height at $\nu_{\max}$, H_{env}. $B_{\max}$ was fitted for each sub series by S18. The function of the inverse signal-to-noise ratio β in equation 4.13 is given by

$$f(\beta) = \sqrt{1+\beta}\,(\sqrt{1+\beta} + \sqrt{\beta})^3 \tag{4.15}$$

The functional form of equation 4.15 is given in Libbrecht (1992) and Toutain & Appourchaux (1994). $f(\beta)$ is the result of assuming that $B(\nu) \approx B_{\max}$. $B_{\max}$ can then be divided through equation 4.11 to leave a factor of β. After this, equation 4.11 is used to solve for $\sigma_{\nu_{\max}}$ (equation 4.12). The process of differentiating, squaring and integrating $M(\nu)$ results in the factors of β in equation 4.15.

H_{env} and Γ_{env} come from a Gaussian fit to the radial mode frequencies and H_{Gauss} values of each sub series from S18. This Gaussian fit is the method used to obtain 'true' values of $\sigma_{\nu_{\max}}$ to compare against predictions from equation 4.13. The method is described in Section 4.4.2.

Section 4.3 described the derivations for the analytical $\sigma_{\Delta\nu}$ equations 4.5 and 4.10, and the derivation for the analytical $\sigma_{\nu_{\max}}$ equation 4.13. In order to test these analytical equations, we will need 'ground truth' values of $\sigma_{\Delta\nu}$ and $\sigma_{\nu_{\max}}$. These 'true' $\sigma_{\Delta\nu}$ and $\sigma_{\nu_{\max}}$ values were obtained from linear and Gaussian fits to the data, respectively. These fits are explained in the next Section.

4.4 Calculating 'true' $\Delta\nu$ and ν_{max} uncertainties

In Section 4.2, the sub series' of the 10 LEGACY stars were prepared. This was done in four steps. In step 3, a straight line was fitted to the 'peak bagged' radial mode frequencies ν and overtone numbers n. The linear fit gave 'true' measures of $\sigma_{\Delta\nu}$ for every sub series. This linear fit is explained in more detail in Section 4.4.1.

Similarly in Section 4.2 step 4, a Gaussian was fitted to the frequencies ν and heights H_{Gauss} of the 'peak bagged' radial modes. This fit gave 'true' measures of $\sigma_{\nu_{\text{max}}}$ for every sub series. The Gaussian fit is described in Section 4.4.2.

4.4.1 Linear fit to estimate $\sigma_{\Delta\nu}$

The $\sigma_{\Delta\nu}$ values from Sections 4.3.1 and 4.3.2 needed to be compared to the formal 'true' uncertainty of the large frequency spacing. This 'true' uncertainty was calculated with a least-squares weighted straight line fit to the data from S18.

The 'true' uncertainties were calculated for every sub series of every LEGACY star. This was done because each sub series is a different realisation of the oscillation spectra. While performing the linear fits, it was assumed that the underlying oscillation parameters do not change significantly over the full duration of the observations (i.e that the star does not evolve during this time).

To calculate the 'true' $\Delta\nu$ uncertainties, a straight line was fitted to the radial mode overtone numbers n, and the frequency values of the radial modes $\nu(n)$. The radial mode frequency uncertainties σ_ν were used to weight the straight line fit.

The weighted straight line $y = mx + c$ needed to be fit. After substituting the values from S18, the line to be fitted became $\nu = \Delta\nu n + c$, with corresponding values of σ_ν. The line was fitted using the `Scipy` package `curve_fit`, written in `Python` (Jones et al., 2001). `curve_fit` returned the 'true' gradient of the line $\Delta\nu$, along with the 'true' uncertainty on that gradient, $\sigma_{\Delta\nu}$.

4.4.2 Gaussian fit to estimate $\sigma_{\nu_{\max}}$

$\sigma_{\nu_{\max}}$ was estimated for each sub series by fitting a Gaussian to the $l = 0$ radial mode frequencies and H_{Gauss} values (equation 4.1). A Gaussian was fitted to the H_{Gauss} values using the same `curve_fit` package as in Section 4.4.1. The general form of the Gaussian to be fitted is

$$f = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} \,. \tag{4.16}$$

The mean value of the Gaussian is $\bar{x}$, and the standard deviation of the distribution is σ. When fitting the solar-like oscillation envelope, $x = \nu$, $\bar{x} = \nu_{\max}$ and $\sigma = \sigma_{\mathrm{env}}$. The function f outputs values of $H_{\mathrm{Gauss}}\,(\nu)$:

$$H_{\mathrm{Gauss}}\,(\nu) = \frac{1}{\sqrt{2\pi\sigma_{\mathrm{env}}^2}}\, e^{-\frac{(\nu-\nu_{\max})^2}{2\sigma_{\mathrm{env}}^2}} \,. \tag{4.17}$$

It is important to provide accurate starting values for the parameters to be fitted, as these will give robust results. These starting values were calculated using real, fitted data from S18.

Before starting values for the fit could be calculated, the heights of the Gaussian at each radial mode frequency (H_{Gauss}) needed to be known. H_{Gauss} was calculated using the 'peak bagged' data from S18: the radial mode heights H, linewidths Γ and the large frequency spacing between radial modes $\Delta\nu$. These 'true' $\Delta\nu$ values for each sub series were obtained from the straight line fits of $\nu(n)$ in Section 4.4.1.

After H_{Gauss} values were obtained from the 'peak bagged' data, starting values for the Gaussian fit could be calculated. The starting values that were given to the fit were the centroid and standard deviation values of the Gaussian. The centroid value of the Gaussian (in this case, $\nu_{\max}$) was given by a weighted mean of the radial mode frequencies ν and Gaussian heights;

$$\nu_{\max} = \frac{\sum H_{\mathrm{Gauss}}\,\nu}{\sum H_{\mathrm{Gauss}}} \,. \tag{4.18}$$

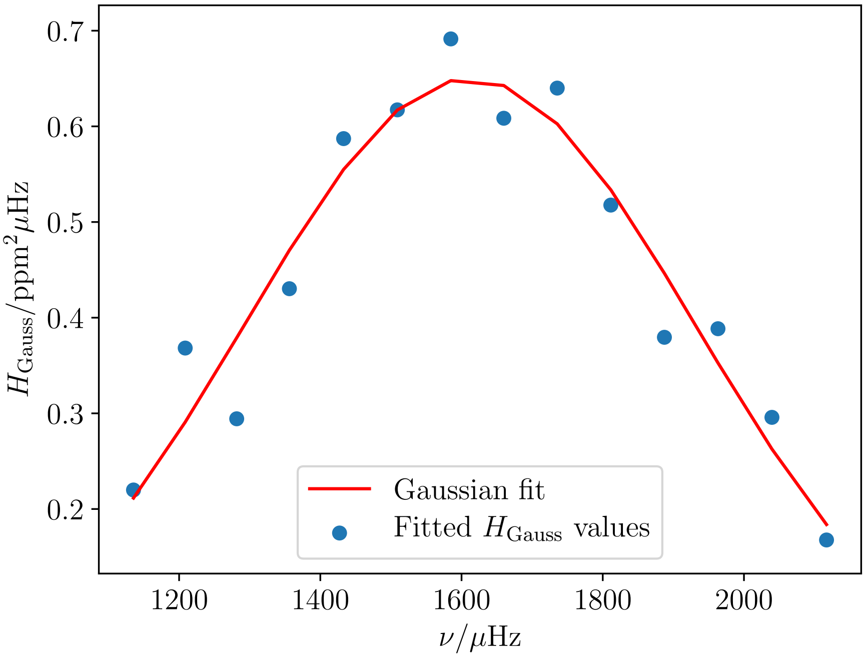


Figure 4.1: The $l = 0$ frequencies and $H_{\rm Gauss}$ values for one sub series of KIC 2837475. The fitted Gaussian from `curve_fit` is shown in red, and has been evaluated at each radial mode frequency and height (the blue points).

The standard deviation of this Gaussian was defined by

$$\sigma = \sqrt{\frac{\sum H_{\rm Gauss}\,(\nu - \nu_{\rm max}\,)^2}{\sum H_{\rm Gauss}}}\,. \tag{4.19}$$

The weighted mean and standard deviation values (equations 4.18 and 4.19 respectively) were the starting values given to the Gaussian fits for each sub series of the 10 LEGACY stars. One of the Gaussian fits for KIC 2837475 is shown in Figure 4.1 to give an example of the results.

The fits provided 'true' values of the centroid uncertainty $\sigma_{\nu_{\rm max}}$. These were compared against the results from equation 4.13. Along with the linear fits (Section 4.4.1), the results from the analytical equations (Section 4.3) could now be compared to 'ground truth' measures of the uncertainty. The results of this comparison are given in the next Section.

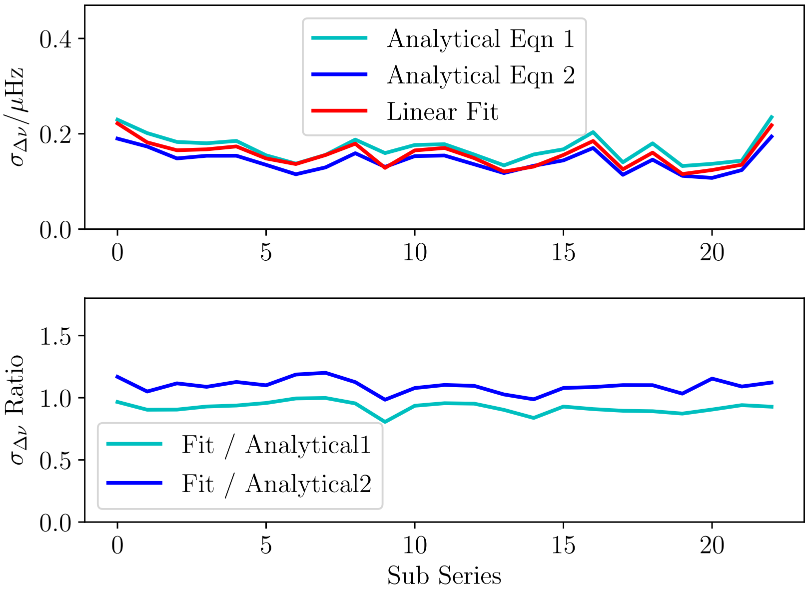


Figure 4.2: The analytical and 'true' linear fit $\sigma_{\Delta\nu}$ values of 23 different sub series' for KIC 2837475.

4.5 Results

This Section presents the results from the analytical equations, and compares them to the 'true' fitted $\sigma_{\Delta\nu}$ and $\sigma_{\nu_{\max}}$ values. Results are given for every sub series of the 10 stars. Firstly, the results of some individual LEGACY stars are highlighted to show the variation of results over time (across the different sub series'). Summary results are also presented across all 224 sub series' of the 10 stars. These results show how well the analytical equations perform.

4.5.1 $\sigma_{\Delta\nu}$ results

Firstly, we compare the results of the analytical $\sigma_{\Delta\nu}$ equations to those from the linear fit (the 'true' $\Delta\nu$ uncertainties). For the $\sigma_{\Delta\nu}$ comparisons, Analytical Eqn 1 refers to equation 4.5, and Analytical Eqn 2 refers to equation 4.10.

Figure 4.2 shows the $\Delta\nu$ uncertainty predictions for KIC 2837475. The upper subplot shows the absolute uncertainties across the 23 sub series'. By comparison, the lower subplot shows the ratio of uncertainty values (the 'true' linear fit values divided by the analytical equations).

The 'true' uncertainties from the linear fits show a small amount of variation across the 23 sub series' ($\sim 0.08\,\mu$Hz). Regardless of this variation, the results from the analytical equations remain acceptably close to the 'true' values. As a reminder, Analytical Eqn 1 (equation 4.5) assumes that the radial mode frequency uncertainties are uniform. This assumption leads to a very slight overestimate of the $\Delta\nu$ uncertainty compared to the 'true' values. By comparison, Analytical Eqn 2 (equation 4.10) does not assume uniform uncertainties. This equation predicts $\sigma_{\Delta\nu}$ correctly, but may lead to a *very* small underestimate of the uncertainty.

It is important to stress just how small the difference is between the analytical $\sigma_{\Delta\nu}$ values and the 'true' uncertainties. Typically, the analytical equations predicted correct $\Delta\nu$ uncertainties to within just $\pm\,0.05\,\mu$Hz. This is more than sufficient for any practical purpose. Moreover, if Analytical Eqns 1 and 2 are used together, they provide accurate upper and lower bounds to $\sigma_{\Delta\nu}$.

Figures 4.3 and 4.4 show similar results for two other LEGACY stars in the dataset. Both analytical equations accurately predict values of $\sigma_{\Delta\nu}$ across the 23 sub series' in both stars. As in Figure 4.2, Analytical Eqn 1 either correctly predicts the uncertainty, or is a very slight overestimate. Conversely, Analytical Eqn 2 either correctly predicts the uncertainty, or is a very slight underestimate. In both cases, the differences are negligible.

To summarise these results, Figure 4.5 shows uncertainty comparisons of all 10 stars, across every sub series. For each star, the 'true' value from the linear fit was divided by the value from the analytical equations. This gives the ratio of the uncertainties. In total, there were 20 - 23 uncertainty ratios for each star (depending on the number of sub series' available). The median value of these 20 - 23 ratios was then calculated. The median values of all 10 LEGACY stars are plotted in Figure 4.5.

Figure 4.5 shows that Analytical Eqn 1 will always overestimate the uncertainty on $\Delta\nu$, while Analytical Eqn 2 will always underestimate it. In addition, Figure 4.5 shows

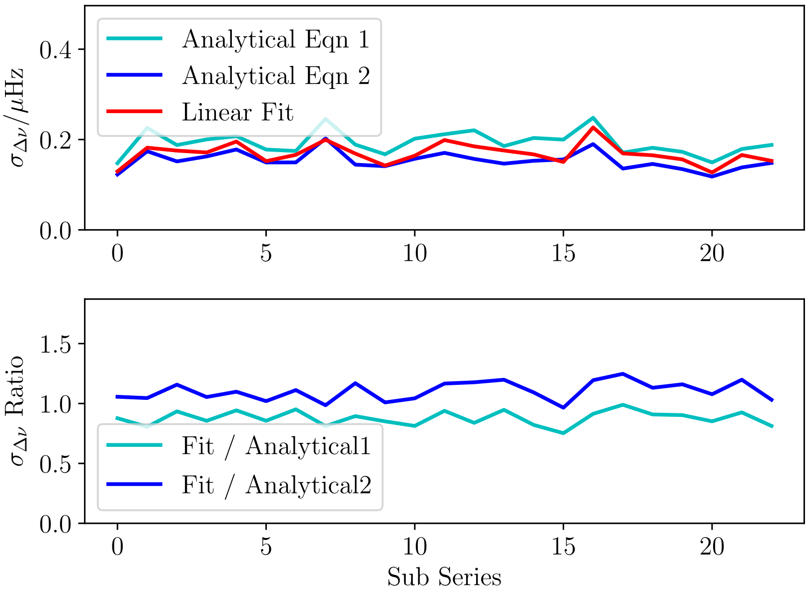


Figure 4.3: The analytical and 'true' linear fit $\sigma_{\Delta\nu}$ values of 23 different sub series' for KIC 9812850.

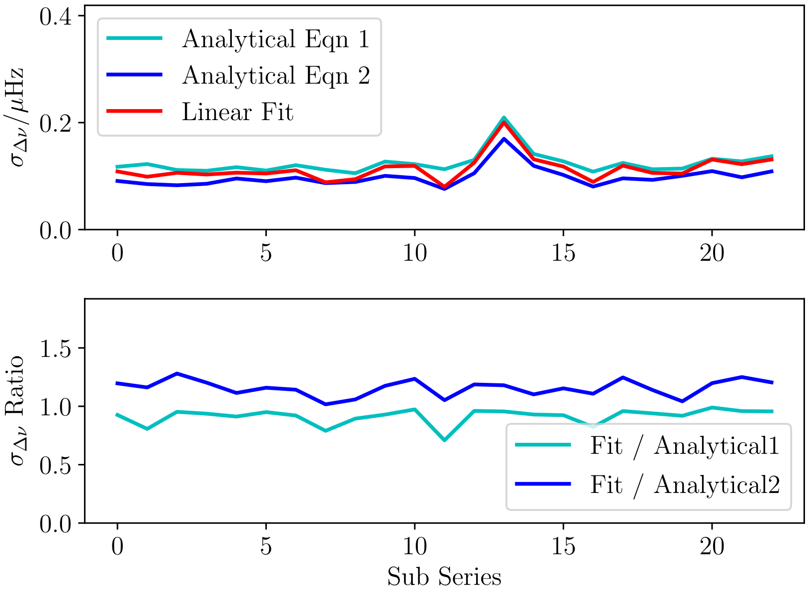


Figure 4.4: The analytical and 'true' linear fit $\sigma_{\Delta\nu}$ values of 23 different sub series' for KIC 12317678.

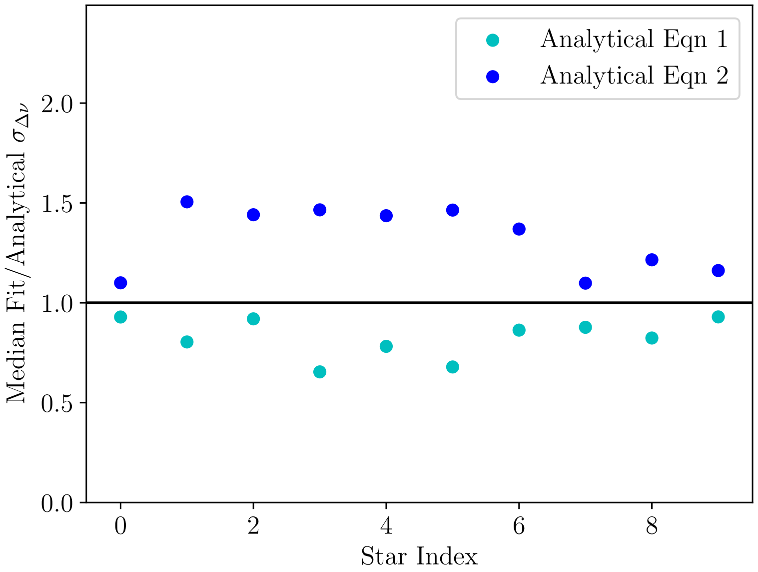


Figure 4.5: The $\Delta\nu$ uncertainty values from the linear fits (the 'true' uncertainties), divided by the uncertainties from equations 4.5 and 4.10, giving two uncertainty ratios for every sub series (one ratio for each equation). The median of these uncertainty ratios is plotted for all 10 LEGACY stars.

that both analytical equations make better $\sigma_{\Delta\nu}$ predictions for some stars than others. For all 10 stars across every sub series, the *absolute* values of both analytical equations are very close to the 'true' values from linear fits.

4.5.2 $\sigma_{\nu_{\max}}$ results

In Section 4.5.1, we compared the results from analytical equations 4.5 and 4.10 to the 'true' results from linear fits. The results showed that both analytical equations predict $\sigma_{\Delta\nu}$ well. Equation 4.5 always provides an upper bound on this uncertainty, while equation 4.10 always provides a lower bound.

In this Section, we make equivalent comparisons for $\sigma_{\nu_{\max}}$. How well do the predictions of analytical equation 4.13 match the 'true' uncertainties from Gaussian fits? To assess that, we compared the results of individual LEGACY stars across different sub series'. These are shown in Figures 4.6, 4.7 and 4.8. For the $\sigma_{\nu_{\max}}$ comparisons, Analytical Eqn 1 refers to equation 4.13.

Firstly, we compare the results from the analytical equation to the 'true' numax un-

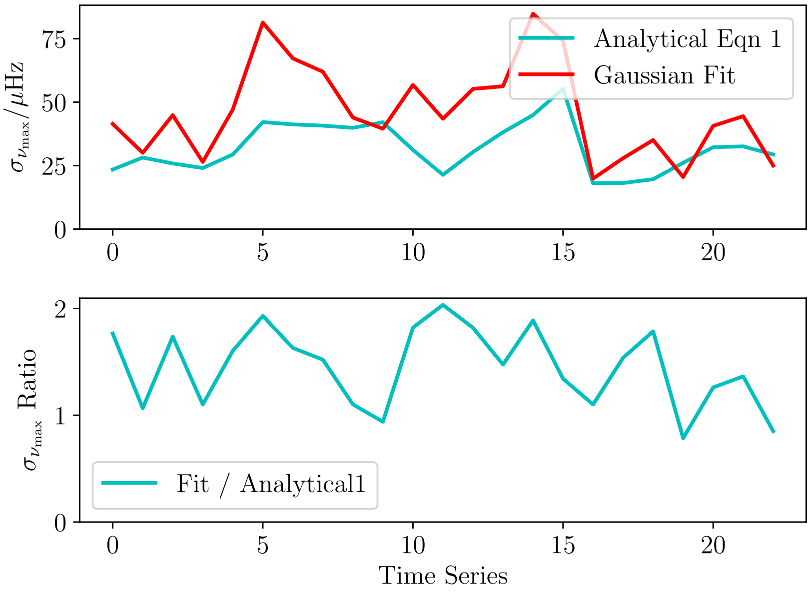


Figure 4.6: The analytical and 'true' Gaussian fit $\sigma_{\nu_{\max}}$ values of 23 different sub series' for KIC 8006161.

certainties for a single star. We start by reviewing the results for KIC 8006161 across 23 sub series' (Figure 4.6). Figure 4.6 shows that the analytical equation either predicts the correct $\nu_{\max}$ uncertainty, or is an underestimate. The upper subplot shows the frequency uncertainty values themselves. For KIC 8006161, the difference between $\sigma_{\nu_{\max}}$ values ranges from $0\,\mu$Hz to approximately $40\,\mu$Hz.

The same results are found for other stars, including KIC 9139151. Figure 4.7 shows that the frequency differences for this star also ranges from $0\,\mu$Hz to approximately $40\,\mu$Hz. For both KIC 8006161 and KIC 9139151, predictions from equation 4.13 range from 50 % - 100 % of the 'true' values from Gaussian fits (see in the lower subplots of Figures 4.6 and 4.7 for the fractional differences between the results).

Thirdly, Figure 4.8 shows results for KIC 10644253. The 'true' uncertainties from the Gaussian fits vary much more across the different sub series' of this star. Figure 4.8 is shown to illustrate just how much the $\nu_{\max}$ uncertainty can vary, depending on the realisation noise in the power spectra.

A common theme is starting to emerge - equation 4.13 either predicts $\sigma_{\nu_{\max}}$ correctly, or it is an underestimate. While the analytical equation provides a lower bound to $\sigma_{\nu_{\max}}$,

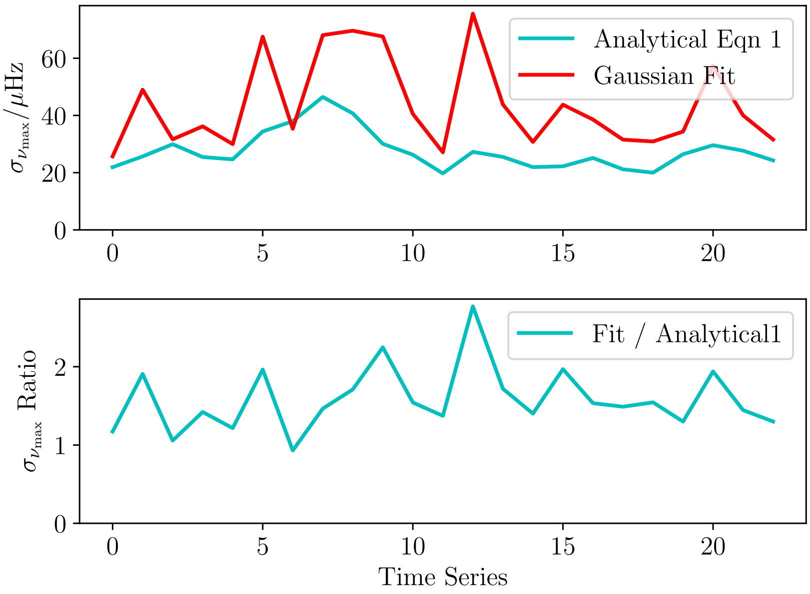


Figure 4.7: The analytical and 'true' Gaussian fit $\sigma_{\nu_{\max}}$ values of 23 different sub series' for KIC 9139151.

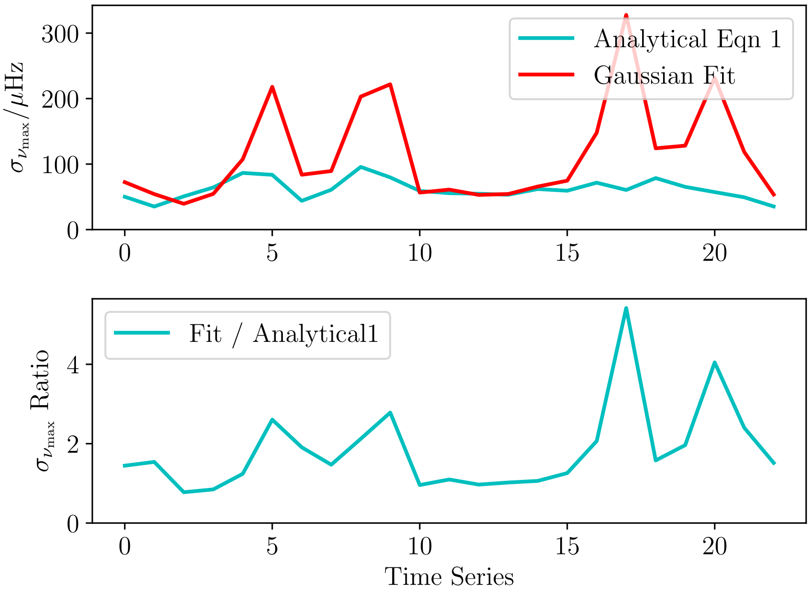


Figure 4.8: The analytical and 'true' Gaussian fit $\sigma_{\nu_{\max}}$ values of 23 different sub series' for KIC 10644253.

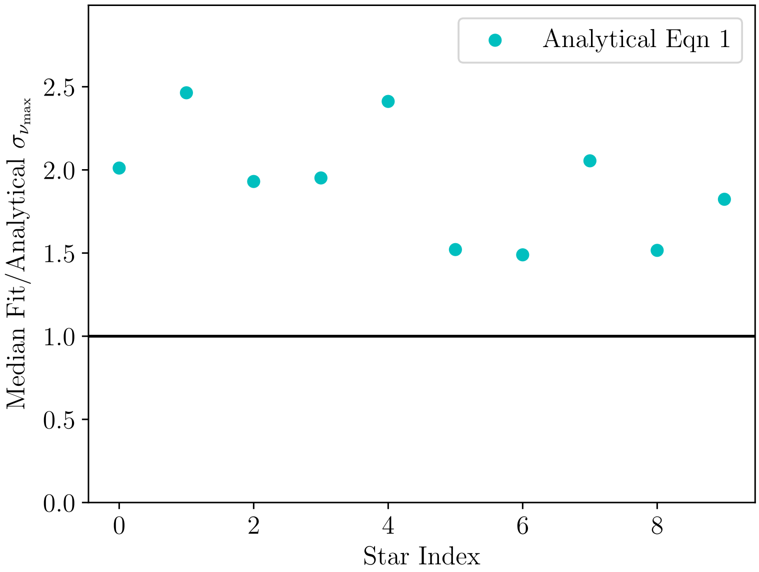


Figure 4.9: The median $\nu_{\max}$ uncertainty values from the Gaussian fits (the 'true' uncertainties), divided by the uncertainties from equation 4.13 for all 10 LEGACY stars.

it cannot reproduce the added scatter from realisation noise.

We have now reviewed the $\sigma_{\nu_{\max}}$ results from three LEGACY stars across every sub series. Figure 4.9 gives a summary of the results from these three stars, alongside the other seven LEGACY stars that have not be shown individually. It shows that across every sub series, on average equation 4.13 under-predicts the uncertainty on $\nu_{\max}$. This is because for a single sub series, equation 4.13 either correctly predicts $\sigma_{\nu_{\max}}$, or under-predicts it. Averaging the results from each sub series will therefore always show an underestimate.

To summarise this, the analytical equation to calculate $\sigma_{\nu_{\max}}$ provides a lower bound to the uncertainty of $\nu_{\max}$. While it successfully predicts $\sigma_{\nu_{\max}}$ in some sub series, it can give far too small a prediction in others. Equation 4.13 struggles to reproduce the added uncertainty from realisation noise. Let us look at why this is.

The analytical equation to calculate $\sigma_{\nu_{\max}}$ depends on the envelope width Γ_{env}, observation time T and the inverse signal-to-noise ratio β (equation 4.13). These parameters are all weakly sensitive to the realisation noise in the observation; they will not vary drastically from one sub series to another.

By comparison, the Gaussian fits were made using the individual radial mode frequencies ν and Gaussian heights H_{Gauss}. H_{Gauss} depends on the individual radial mode heights H and linewidths Γ, as well as the large frequency spacing between radial modes. These parameters are all much more sensitive to realisation noise. This means that the 'true' $\sigma_{\nu_{\max}}$ values from the Gaussian fits will vary much more from one sub series to another (e.g. Figure 4.8).

4.6 Conclusion

Correctly predicting uncertainties in astronomy is immensely valuable. If the uncertainties of some data are too small or too large, incorrect conclusions may be drawn from that data. The aim of this work was to prevent this happening when working with the global asteroseismic parameters $\Delta\nu$ and $\nu_{\max}$.

This work uses the LEGACY data from Aguirre et al. (2017) and Lund et al. (2017). In S18, the time series' of these stars were separated into 90-day sub series'. These sub series' were subsequently 'peak-bagged'. S18 fitted the frequencies, uncertainties, heights and linewidths of the radial solar-like modes within every 90-day sub series, across all 10 stars. Here, each sub series was treated independently to assess the variation of $\sigma_{\Delta\nu}$ and $\sigma_{\nu_{\max}}$ for different observations of a single star.

Firstly, we present two analytical equations to calculate the uncertainty of the large frequency spacing $\Delta\nu$ (equations 4.5 and 4.10). Both equations can be evaluated using only the mode frequency uncertainties and overtone numbers. In addition, we give an analytical equation to calculate the uncertainty of $\nu_{\max}$, the frequency of maximum oscillation power. This uses no information about individual modes of oscillation. Instead it calculates $\sigma_{\nu_{\max}}$ using the envelope width, observation time and the inverse signal-to-noise ratio at $\nu_{\max}$.

To assess the accuracy of the predictions from the analytical equations, the results for the LEGACY stars were compared to ground-truth values. These 'true' $\sigma_{\Delta\nu}$ and $\sigma_{\nu_{\max}}$ values were obtained by performing non-linear least-squares fits to the LEGACY data. To calculate 'true' values of $\sigma_{\Delta\nu}$, a straight line was fitted to the radial mode frequencies ν and overtone numbers n; $\nu(n)$. The gradient of this line is the frequency spacing between radial modes, $\Delta\nu$. The least squares fit provides an uncertainty on that gradient. This was used as the 'true' value of $\sigma_{\Delta\nu}$ to compare results against.

A similar approach was taken for $\sigma_{\nu_{\rm max}}$. Least squares fitting was once again applied to the data. This time, a Gaussian was fitted to the radial mode frequencies ν and their corresponding heights $H_{\rm Gauss}$. This least-squares Gaussian fit provided a value of the central frequency $\nu_{\rm max}$, as well as an uncertainty on that central frequency; the 'true' value of $\sigma_{\nu_{\rm max}}$.

After performing the least-squares fits, we find that the analytical equations to calculate $\sigma_{\Delta\nu}$ work extremely well. They provide accurate values of the uncertainty on $\Delta\nu$ to within just $\pm\,0.05\,\mu$Hz. This is extremely encouraging, and shows that both (or either) equation can be used with confidence elsewhere in the field of asteroseismology. Equation 4.5 consistently provides a *slightly* larger $\sigma_{\Delta\nu}$ value than the 'true' uncertainty from linear fits. This is because it assumes that the radial mode frequencies are all drawn from the same Gaussian distribution. Equation 4.10 does not make the assumption of uniform uncertainties. and so does not over-estimate $\sigma_{\Delta\nu}$.

An analytical equation to calculate the uncertainty on $\nu_{\rm max}$ was also presented here (equation 4.13). In a given power spectrum, it consistently provided a lower bound to the 'true' value of $\sigma_{\nu_{\rm max}}$. While it correctly predicted $\sigma_{\nu_{\rm max}}$ for some sub series, it underestimated the uncertainty on $\nu_{\rm max}$ in others.

Equation 4.13 cannot reproduce the added uncertainty from realisation noise. This is because it is only given information of the mode envelope width, the observation time of the star and the signal-to-noise ratio at $\nu_{\rm max}$. By comparison, the 'true' uncertainty on $\nu_{\rm max}$ was calculated using the individual radial mode heights, linewidths and large frequency spacings. These parameters are all much more sensitive to realisation noise. As a result, the 'true' $\sigma_{\nu_{\rm max}}$ values vary much more from one observation to another than the values of $\sigma_{\nu_{\rm max}}$ from the analytical equation.

The analytical equations presented here can all be used in the field of asteroseismology with confidence. Both analytical $\sigma_{\Delta\nu}$ equations are extremely accurate. For a given observation, the equations differ from the 'true' $\Delta\nu$ uncertainty with maximum uncertainties of just $\pm\,0.05\,\mu$Hz. If used together, they provide accurate upper and lower bounds to the uncertainty on $\Delta\nu$.

The analytical $\sigma_{\nu_{\max}}$ equation gives a lower bound on the uncertainty on $\nu_{\max}$. This is still a useful parameter to calculate, providing that it is used with a caveat: the values from the equation are a lower bound to the 'true' uncertainty on $\nu_{\max}$. In the case of low realisation noise, the analytical equation provides accurate values of $\sigma_{\nu_{\max}}$. When larger realisation noise is present, the analytical equation should be used to give a lower bound to the uncertainty on $\nu_{\max}$. When observing solar-like oscillators with TESS, the realisation noise will typically be relatively high (Chapter 2; Ricker et al. 2014), so equation 4.13 should be used bearing this in mind.

In the next Chapter, we will go on to describe how asteroseismic target selection can be done without using scaling relations. Instead, stellar observables (such as parallax and $\log(g)$), along with individual mode detection probabilities are used to train a machine. This machine learns the relationship between the observables and the detectability of modes, in order to select stars for observation by any future mission.

Chapter 5

TESS Asteroseismic Predictions for red giants using Machine Learning

Guy Davies provided the radial mode frequencies, widths and amplitudes for the 1000 *Kepler* stars used in this chapter. Bill Chaplin provided insight into how various parameters affect mode detectability. Andrea Miglio provided insight into stellar evolution.

The work done here is being prepared for publication.

5.1 Introduction

In this work, machine learning was used to investigate the potential of predicting the detectability of oscillations in stars given basic observables as input. If this is possible, this would mean an alternative to the construction of detection algorithms (e.g see Chaplin et al. 2011 for *Kepler*, or Campante et al. 2016 for TESS).

A classifier was used here to make predictions about the detectability of solar-like modes of oscillation. Specifically, predictions of mode detectability for *Kepler* red giant branch stars were made. After this, these *Kepler* red giants were degraded to produce 1-year and 27-day TESS-like datasets. Mode detectability predictions were then made using the same classifier on these TESS-like observations.

In the past, supervised classifiers have been used to determine the evolutionary state of variable stars (see Debosscher et al. 2007, Sarro et al. 2009, Nun et al. 2014, Elorrieta et al. 2016). Similarly, unsupervised classifiers (Valenzuela & Pichara, 2018) and regression (Ness et al., 2015) have also been used to determine evolutionary stages. As well as classifiers, a neural network has been used to identify the evolutionary stage of red giant stars (Hon et al. 2017, Hon et al. 2018).

As well as identifying the evolutionary stage of stars, Machine Learning has been used to calculate stellar parameters (Bellinger et al., 2016) and oscillation frequencies (Davies et al., 2016).

For the first time, this work makes predictions about the detectability of solar-like modes inside *Kepler* and TESS targets using a classifier. This same technique can be applied to any future space mission, such as PLATO (Rauer et al., 2014). As well as being able to apply the same technique to any future mission, Machine Learning can also reverse-engineer target selection bias, for example in K2 (Lund et al., 2015) and CoRoT (Baglin et al., 2006). This can provide insights into the formation history of the Galaxy (Thomas et al., 2017).

Firstly, Section 5.2 describes the *Kepler* red giant dataset from Davies et al. (in prep). These 1000 stars have measured values for the global properties [`log`(g)`,` π`,` T_{eff}`,` `[M/H],` I_{mag}`]`, as well as fitted radial mode frequencies, heights and widths. Most (but not all) of these red giant stars have been observed in the I-band. Section 5.2.1 describes how regression was used to infer the I_{mag} values of the missing stars, using the rest of the red giant dataset.

After the *Kepler* dataset was prepared, the stars were modified to mimic observations by TESS. Section 5.4 explains how changes were made to the data in the time-domain to simulate 1 year (stars in the Continuous Viewing Zone) and 27 days (stars observed for 1 sector) of TESS-like observations.

3 datasets were available to classify: the original *Kepler* data, 1 year of simulated TESS observations, and 27 days of simulated TESS observations. Each dataset was

made up of the same 1000 stars, after their apparent magnitudes had been perturbed 100 times. Section 5.5 describes the detection test that was run on the individual modes of the stars in these datasets. This determined which radial solar-like modes of oscillation were detectable by *Kepler* and TESS.

Once detection probabilities were calculated for every mode, Section 5.6 explains how a classifier was used on the 3 datasets separately. The detection probabilities of each mode were put into three bins. Qualitatively, there were: not detected (0 - 50 %), potentially detected (51 - 90 %), very like to be detected (91 - 100 %)). The global properties of the red giants (`[log(`g`), `π`, `T_{eff}`, [M/H], `I_{mag}`]`) were given to the classifier alongside the binned detection probabilities to train the classifier.

Once trained, the classifier was given only `[log(`g`), `π`, `T_{eff}`, [M/H], `I_{mag}`]` and required to predict the binned detection probabilities. It did this with with a 98 % precision for the original *Kepler* data, 90 % precision for 1 year of TESS data, and 81% precision for 27 days of TESS data.

5.2 The dataset

The data to perform Machine Learning on are the 1000 *Kepler* red giants from Davies et al. (in prep). In that work, the radial modes of the *Kepler* stars were fitted. Specifically, Davies et al. (in prep) provided the frequency, height, width and background of each radial mode in the 1000 stars. The spectroscopic parameters of these stars (T_{eff}, $\log(g)$, [M/H]) were available from APOKASC (Pinsonneault et al., 2014). The apparent magnitudes are from the *Tycho*-2 catalogue (Høg et al., 2000). Lastly, the parallaxes are from *Gaia* DR2 (Lindegren et al., 2018). These stars are shown in Figure 5.1.

5.2.1 Obtaining missing Imag values

741 of the 1000 *Kepler* stars have measured I_{mag} values from $Tycho-2$ (Høg et al., 2000). I_{mag} is needed for every star to calculate the TESS shot noise level. In order to keep the remaining stars in the dataset, the I_{mag} values that are missing from the dataset were imputed (values were assigned to the missing magnitudes by inference).

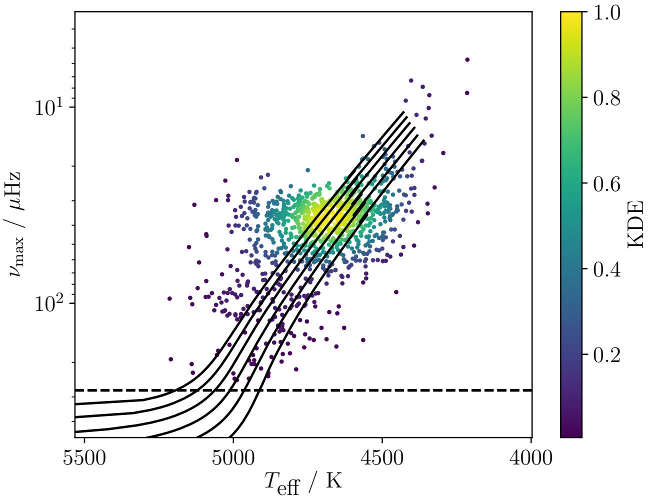


Figure 5.1: The 1000 *Kepler* red giants from Davies et al. (in prep). The colourbar shows the relative density of points with Kernel Density Estimation. The evolutionary tracks range from 0.9 - 1.5 $M_{\odot}$. The dashed line shows the *Kepler* Nyquist frequency for Full Frame Images (278 μHz). The evolutionary tracks range from 0.9 - 1.5 $M_{\odot}$. These tracks were generated using CLÉS (Scuflaire et al., 2008).

I_{mag} values for the missing stars were inferred using random forest regression. Like a random forest classifier, random forest regression is another example of supervised learning: in both cases there is a known label to predict. In this case, the label to predict is I_{mag}.

Supervised learning involves separating the data into training and testing datasets. In classification, the data are grouped by similarity. In regression, the difference between the regression model and 'true' values is evaluated, and iteratively reduced. This difference is evaluated using the Sum of Squared Errors (SSE);

$$\text{SSE} = \sum_{i=1}^{N} (x_i - \bar{x})^2 \,. \tag{5.1}$$

The random forest regression was done in 3 stages: the algorithm was trained, tested, and used to calculate I_{mag} values where they are missing in the 1000-star dataset. To train and test the algorithm, the 741 *Kepler* stars with [K_p, [M/H], T_{eff}] and known I_{mag} values were used. These 741 stars were split into a training dataset, and a testing dataset.

70 % of the 741-star dataset were used to train the algorithm; 30 % were used to test its accuracy. The regression achieves an explained variance score[1] of 0.54 (the best possible score is 1). The explained variance is given by

$$\texttt{explained\_variance}(I_{\mathrm{T}}, I_{\mathrm{P}}) = 1 - \frac{Var\{I_{\mathrm{T}} - I_{\mathrm{P}}\}}{Var\{I_{\mathrm{T}}\}} \,. \tag{5.2}$$

I_{T} are the 'true' I_{mag} values of the stars used to train the algorithm (70 % of the 741-star dataset), and I_{P} are the predicted I_{mag} values from the regression for these stars. The explained variance quantifies how well the predicted dataset accounts for variance in the 'true' I_{mag} values.

The results of this random forest regression are shown in Figures 5.2 and 5.3. Figure 5.2 shows that the distribution of predicted value matches the 'true' I_{mag} values closely. Most importantly, Figure 5.3 shows that there is no offset between the predicted and 'true' values; the mean difference between the two is 0.06 mag. Furthermore, the standard deviation of the difference is only 0.40 mag. To summarise this, random forest regression predicts the I_{mag} values in the test dataset without introducing bias or adding a large uncertainty.

Figure 5.3 shows that the regressor does not predict 'extreme' I_{mag} values (less than 9.5 or greater than 11.5) as accurately as values that lie between 9.5 - 11.5. This is because some 'extreme' I_{mag} values in the testing set lie at the edge or outside of the training distribution (Figure 5.2). The regressor cannot predict values outside of the training set that it is given, so will always tend to predict extreme magnitudes (less than 9.5 or greater than 11.5 in our case) closer to the centre of the training distribution.

After the algorithm was trained and tested, it was used to calculate the I_{mag} values of the 259 stars where they are unavailable. The *Kepler* apparent magnitudes, K_p, of the stars with known and unknown I_{mag} values are shown in Figure 5.4. Similarly, the distribution of the known, and predicted (previously unknown) I_{mag} values is shown in Figure 5.5. In both Figures 5.4 and 5.5, the stars with unknown I_{mag} values lie toward the

---

[1] http://scikit-learn.org/stable/modules/model_evaluation.html#explained-variance-score

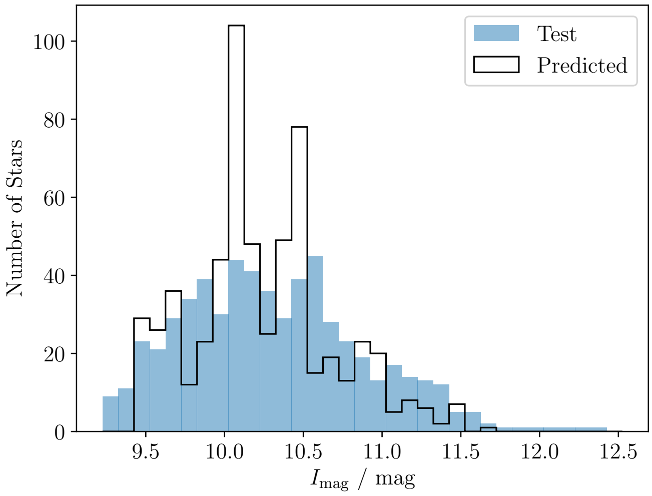


Figure 5.2: The $I_{\rm mag}$ distribution of the *Kepler* stars used to test the random forest regression. The 'true' values used to test the algorithm are shown in blue. The black histogram shows the distribution of values that the random forest regression predicted for those stars. A scatter plot of the differences between these distributions is shown in Figure 5.3.

fainter end of the distribution. This implies that the random forest regression is making correct inferences about the $I_{\rm mag}$ values of the stars.

The majority of *Kepler* stars in the Davies et al. (in prep) dataset had measured $I_{\rm mag}$ values. Using regression, the $I_{\rm mag}$ values of the remaining stars were evaluated. After this, the size of the dataset was increased from 1000 to 81,100.

5.3 Increasing the sizes of the *Kepler* and TESS datasets

Machine Learning performs best when a large dataset is available to train the algorithm on. In order to increase the size of the dataset above 1000 (Table 5.1), the magnitude of each *Kepler* star from Davies et al. (in prep) was perturbed. Each star had its apparent magnitude perturbed 100 times (Table 5.2).

The instrumental (shot) noise model of *Kepler* was used as the PDF to draw apparent magnitudes from. This was used to increase the size of the dataset because it provides a realistic distribution of the number of stars at different magnitudes that the satellite

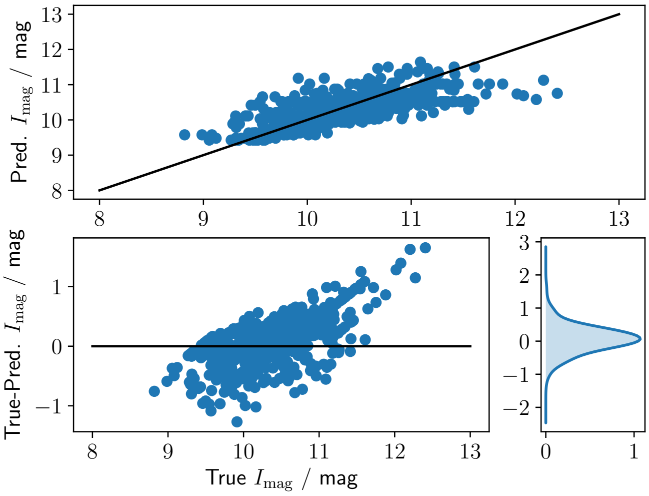


Figure 5.3: The true I_{mag} values of the *Kepler* stars use to test the algorithm, compared to their predicted values. The mean difference between the two sets of values is 0.06 mag, with a standard deviation of just 0.40 mag.

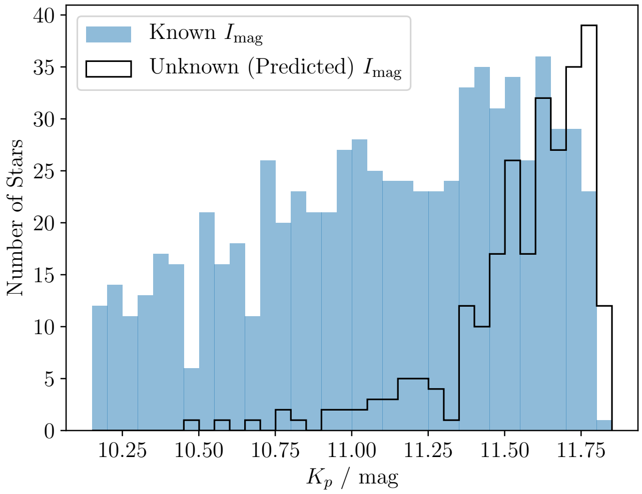


Figure 5.4: The K_p distribution of the *Kepler* stars with known I_{mag} values (741 stars) is shown in blue. The K_p values of the stars without I-band magnitudes (259 stars) are shown in black. The majority of these stars are fainter.

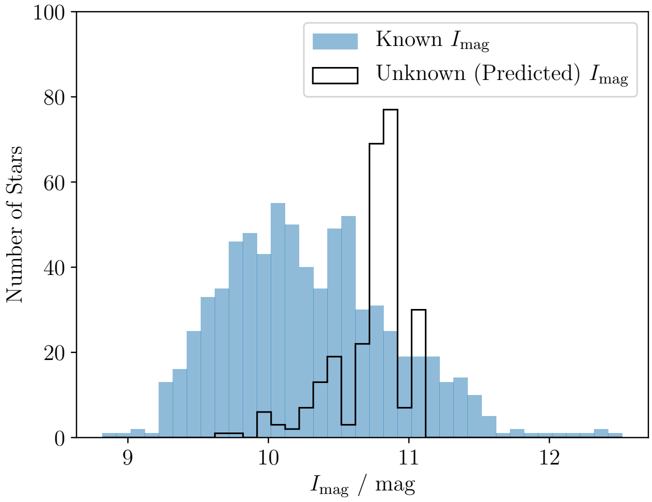


Figure 5.5: The I_{mag} distribution of the *Kepler* stars with known values (741 stars) is shown in blue. The I_{mag} values predicted by the regression (259 stars) are shown in black. Like in Figure 5.4, the predicted magnitudes are fainter than the majority of known values.

observed. Many more fainter stars are observed than brighter stars because the volume of space that contains stars increases as the distance of observation increases.

The shot noise model of *Kepler* depends on the *Kepler* magnitude of the star, K_p. This noise model is from Gilliland et al. (2010). The `RMS` noise, σ_{Kepler}, is given by

$$\sigma_{Kepler} = \frac{10^6}{c}\sqrt{c + 9.5\times 10^5\left(\frac{14}{K_p}\right)^5}\ \mathrm{ppm}\,, \tag{5.3}$$

where c is the number of detected electrons per cadence. It is given by

$$c = 1.28\times 10^{0.4\,(12\,-\,K_p)\,+7}\,. \tag{5.4}$$

The *Kepler* noise model was used as the PDF to draw stellar magnitudes from. 100 magnitudes were drawn for every *Kepler* red giant. For every star in the original 1000-star dataset, there were 100 copies with identical $\log(g)$, π, T_{eff} and [M/H] values (Table 5.2). Each of these identical copies had a different apparent magnitude value.

After this, stars that do not have available measurements of $\log(g)$, π, T_{eff} and [M/H] were removed. This left 81,100 datasets (out of 100,000). For comparison to Table 5.2,

KIC	Star Index	$\log(g)$	π	T_{eff}	[M/H]	I_{mag}
9205705	1	2.758	0.688	4685	-0.39	9.89
2554924	2	2.799	0.969	4594	0.27	8.46
6508328	3	2.467	0.655	4712	0.29	11.27
3748691	4	2.482	0.606	4632	0.07	8.26
...						
9704774	1000	2.442	0.569	4679	0.03	10.56

Table 5.1: The original 1000 *Kepler* red giants from Davies et al. (in prep). The spectroscopic parameters of these stars (T_{eff} , $\log(g)$, [M/H]) were available from APOKASC. The apparent magnitudes are from the *Tycho*-2 catalogue. Lastly, the parallaxes are from *Gaia* DR2.

an exert of the original 1000-star dataset is shown in Table 5.1.

After the *Kepler* magnitudes of every star were perturbed 100 times and gaps in the timeseries data were removed, this left 81,100 out of 100,000 datasets. 3 copies of these 81,100 datasets were then made. In Section 5.4, the timeseries of one copy of 81,100 datasets will be treated like 4 years of *Kepler* observations. The timeseries of the second copy will be degraded to mimic 1 year of TESS observations. The timeseries of the third copy of 81,100 datasets will be degraded to look like 27 days of TESS observation.

5.4 Modifying Kepler data to look like TESS

Before a classifier could be used on the red giant sample, the timeseries data from *Kepler* needed to be modified for TESS. These adjustments were made in the time domain before the signal was converted to the frequency domain.

Several different adjustments needed to be made to the *Kepler* data. One difference between the missions is the length of observation. The *Kepler* mission observed stars for up to 4 years. The nominal 2 year long TESS mission will observe stars for between 27 days to 1 year, according to the ecliptic latitude of the stars (Ricker et al., 2014).

As well as reducing the dataset length, the bandpass of observation needed to be adjusted. TESS will observe in a much redder bandpass than that of *Kepler*. This has the effect of reducing the amplitude of stellar signals (i.e. the signals due to stellar granulation and solar-like oscillations) (Ballot et al., 2011). Campante et al. (2016) found this correction to the oscillation intensity amplitude to be 0.85 for TESS.

Thirdly, the noise level for a given star in *Kepler* is lower than the noise level in TESS for a star of the same brightness. The TESS shot noise model was taken from Sullivan et al. (2015). This `RMS` noise model takes into account photon counting noise, the noise from background stars, the readout noise and the systematic noise. These four noise sources were then summed in quadrature to give the total TESS noise, σ_{TESS}. σ_{TESS} was then added to the *Kepler* timeseries.

These three adjustments - to the length of observation, to the bandpass, and to the noise level - were performed in the time domain. From the original *Kepler* timeseries, the following adjustments were made:

1. Apply a 4-σ clip to the dataset to remove spurious points.
2. Shorten the 4 years of timeseries data down to the reduced dataset length (either 27 days or 1 year). When reducing the length of the timeseries, take the section (27 days or 1 year) with the fewest gaps in observation out of the 4 year *Kepler* data.
3. Adjust the bandpass by multiplying the flux by 0.85.
4. Add TESS instrumental noise to the timeseries. For each flux value in the *Kepler* timeseries (F_{Kepler}) draw a random number from the normal distribution (N). Multiply the TESS `RMS` noise level (σ_{TESS}, Section 5.3) by the random numbers N. Add these to the original flux values in the *Kepler* timeseries to get the TESS flux:

$$F_{\mathrm{TESS}} = F_{\mathrm{Kepler}} + N\,\sigma_{\mathrm{TESS}}\,. \tag{5.5}$$

5. Use the Lomb-Scargle periodogram to convert the signal from the time to the frequency domain.

The results from the original *Kepler* observation are compared to TESS-like power spectra of the same stars in Figures 5.6 and 5.7. Power spectra were generated for the original 81,100 4-year *Kepler* datasets, for the 81,100 datasets made to mimic 1 year of

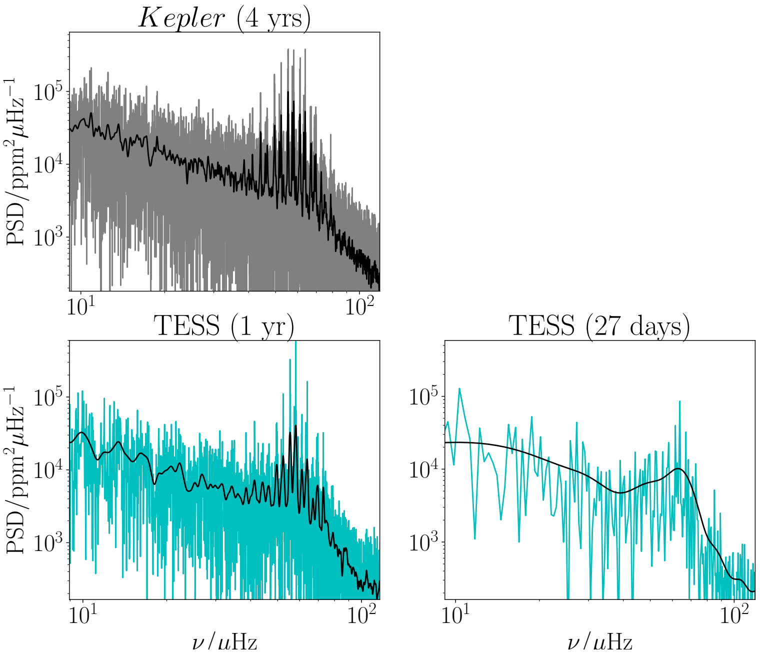


Figure 5.6: The Power spectra of KIC 9535399 is plotted three times with moving medians in black. The original Power Spectra is plotted in grey. The power spectra after making the data look like TESS are plotted in blue. The subplot on the bottom left shows 1 year of TESS observation (the maximum). The subplot on the bottom right column shows 27-days (the minimum).

TESS observations and for the 81,100 datasets made to mimic 27 days of TESS observations. After this, a detection test was run on the radial modes of the three 81,100 datasets.

5.5 Detection Test

Section 5.4 described the method to transform the *Kepler* lightcurves into TESS-like power spectra. A detection test was then run on the stars to determine which modes were detectable by TESS, and which were not.

First, a moving median was used to estimate the underlying background spectrum. The solar-like mode envelope width was used as the frequency range of this moving median. This envelope width was calculated as

$$\Gamma_{\text{env}} = \begin{cases} 0.66\,\nu_{\text{max}}^{0.88}\,(1 + (T_{\text{eff}} - T_{\odot}) \times 6 \times 10^{-4}) & \text{if } T_{\text{eff}} > 5600\,\text{K} \ , \\ 0.66\,\nu_{\text{max}}^{0.88} & \text{if } T_{\text{eff}} \leqslant 5600\,\text{K} \ . \end{cases} \tag{5.6}$$

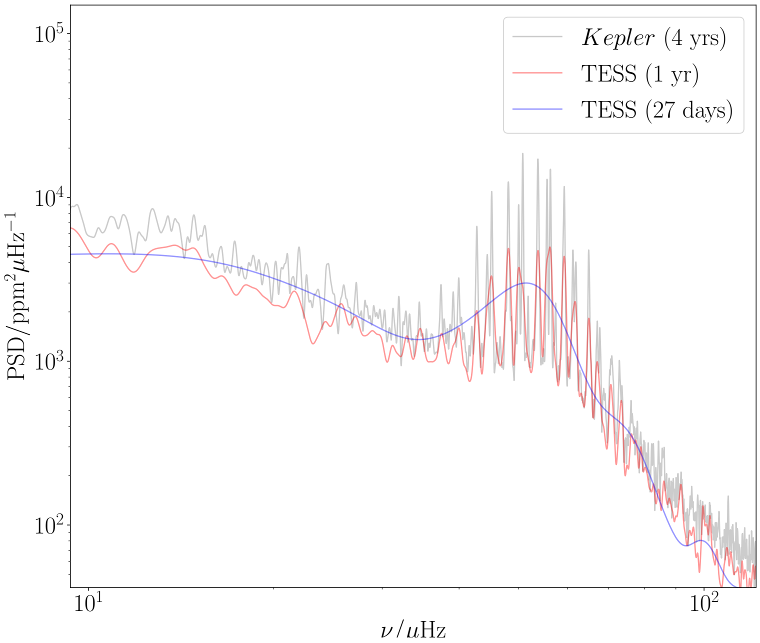


Figure 5.7: The Power Spectrum of KIC 6768319. The original power spectrum is in grey. The data were transformed into 1 year and 27 days of TESS observation. The moving medians of these transformations are overplotted.

from Chapter 3 equation 3.13. The moving median provided an estimate of the background B in the power spectrum (ppm$^2\,\mu$Hz^{-1}). This background was divided out of the power P (ppm$^2\,\mu$Hz^{-1}) in the power spectrum to the get signal-to-noise ratio of the spectrum,

$$\mathrm{SNR} = P\,/\,B\,. \tag{5.7}$$

Once the SNR spectrum for the star was recovered, the SNR values at the mode frequencies were extracted. To ensure the correct SNR values of every mode were used, a window was fitted around each mode frequency from Davies et al. (in prep). The size of the window was given as the fitted linewidth of the mode. The highest value in the window was taken as the SNR of the mode. An example of this for KIC 10587122 is shown in Figure 5.8.

Once all mode SNR values for the star were calculated, a detection test from Chaplin et al. (2011) was run on each mode. The detection test as described in Chaplin et al. (2011) was used across the entire oscillation envelope. Here, the same detection test was instead applied to individual modes. This method is described below.

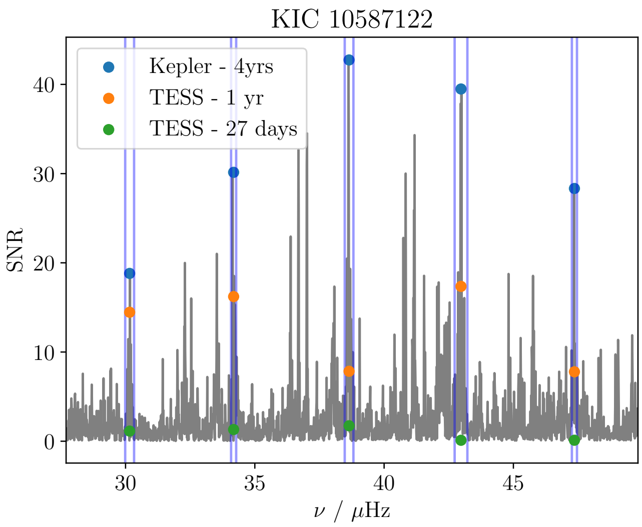


Figure 5.8: The SNR spectrum of KIC 10587122 after background subtraction. The SNR values of the radial modes in the star were extracted from this spectrum. The highest SNR value within the linewidth of each mode is taken to be the SNR value of that mode. The mode linewidths are shown as blue lines. The values of every mode in the original Kepler spectrum are plotted as blue points. The overplotted orange points are the SNR values after degrading the signal to 1 year of TESS observation. Similarly, the green points are the SNR values of 27 days of TESS observations. The white noise level and reduced observation time severely reduce the SNR of TESS observations compared to *Kepler*.

The probability P that the SNR value of a solar-like mode of oscillation lies above some threshold $\mathrm{SNR}_{\mathrm{thresh}}$ is

$$P(\mathrm{SNR} \geqslant \mathrm{SNR}_{\mathrm{thresh}}) = p\,. \tag{5.8}$$

A false-alarm probability p of 5% was set; there is a 95% chance that the signal is not due to noise. Equation 5.8 is solved for $\mathrm{SNR}_{\mathrm{thresh}}$ by substituting P with

$$P = \int_{x}^{\infty} \frac{e^{-x}}{\Gamma(N)}\, x^{N-1}\, dx\,. \tag{5.9}$$

N is the number of frequency bins that the mode occupies. The linewidth of each mode was used as the value of N in equation 5.9.

$\Gamma(N)$ is the Gamma function. The lower bound of Equation 5.9 is set to $x = 1 + \mathrm{SNR}_{\mathrm{thresh}}$. The noise in the N bins is assumed to follow χ^2 $2N$ d.o.f statistics.

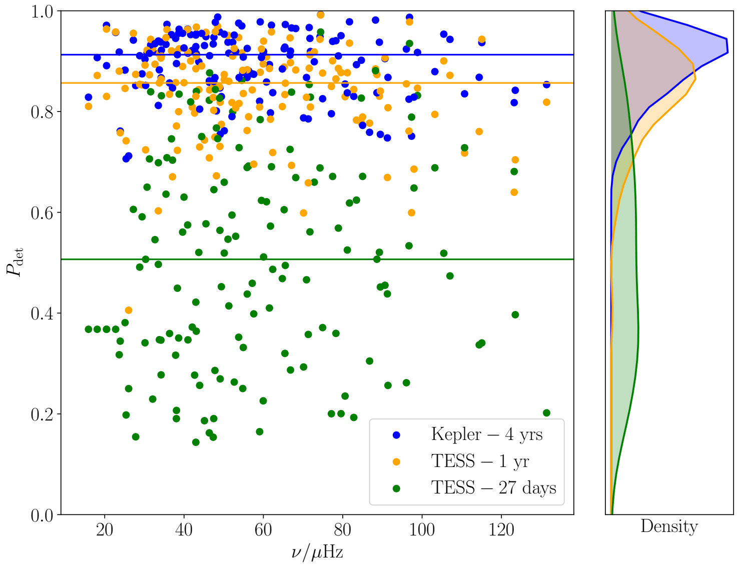


Figure 5.9: A plot showing the result of the detection test, after running on every mode in 20 stars. The results of the original power spectra are plotted in blue. The results of 1 year of TESS observation are in orange. 27 days of TESS observation is in green. At this short an observation, detecting individual modes will be extremely difficult.

Once $\mathrm{SNR}_{\mathrm{thresh}}$ is found, Equation 5.9 is solved again. This time it is solved for P by setting $x = (1 + \mathrm{SNR}_{\mathrm{thresh}}) \,/\, (1 + \mathrm{SNR})$. Here, SNR is the observed signal-to-noise ratio calculated from Equation 5.7. This calculates the probability that the measured SNR value was due to stochastic excitation in the convective envelope of the star.

This detection test was applied to every mode of every star in the sample. Figure 5.9 shows the mode detection probabilities from this test for the original *Kepler* dataset, for 1-year of TESS observation and for 27 days of TESS observation. After the P_{det} values for the 3 datasets were calculated, a classifier was used to see if these results could be predicted rather than calculated. This is described in Section 5.6.

5.6 Classification

Section 5.4 describes how the lightcurves of every star were treated before a detection test was run on the oscillations in Section 5.5. After this, a random forest classifier was used to predict the detection probability of the modes in the red giants.

Classification algorithms work by assessing similarity[2]. In classification, the training set is separated into groups based on the similarity of the data. The more information that was gained by splitting the data, the better. The data continues to be split until a prediction can be made (in this case about the detection probabilities of 3 radial modes). In decision tree classification, this is done once.

Here, a random forest classifier was used. This is made up of many independent Decision Trees which have been weighted to predict the detection probabilities of the 3 radial modes centred around $\nu_{\rm max}$. Section 5.6.1 will explain how the data were prepared before the random forest classifier was used.

5.6.1 Preparing the data

After removing stars without APOKASC information, 811 stars were left from the original 1000-star dataset. In Section 5.3, each star had it's $I_{\rm mag}$ value perturbed 100 times, giving 81,100 datasets. This was done because different $I_{\rm mag}$ values will have different shot noise levels (Section 5.3). After perturbing the apparent magnitudes, 3 copies of the 81,100 datasets were made. In Section 5.4, these 3 copies of the 81,100 datasets were treated in 3 ways: as *Kepler* targets (observed for up to 4 years), as TESS targets in the Continuous Viewing Zone (1 year of observation), and as TESS targets observed in 1 sector (for 27 days).

In Section 5.5, a mode-detectability test was performed on all 3 sets of 81,100 datasets. After this, the same classification method was performed once on each set of 81,100 datasets. Before a classifier could be used, the data needed to be prepared. That preparation is described in this Section.

Each calculated detection probability from Section 5.5 was put into a discrete bin (or *class*) depending on how likely the mode is to be detected. These discrete classes are

---

[2] http://www.simafore.com

given in equation 5.10.

$$P_{\mathrm{det}} = \begin{cases} 2 & \text{if } 1.0 \geq P_{\mathrm{det}} > 0.9 \quad , \\ 1 & \text{if } 0.9 \geq P_{\mathrm{det}} > 0.5 \quad , \\ 0 & \text{if } 0.5 \geq P_{\mathrm{det}} > 0.0 \quad . \end{cases} \tag{5.10}$$

Using equation 5.10, every mode was assigned a discrete class [0, 1 or 2], depending on how high the probability of detection was for that mode. The same three radial modes (3 *features*) were used for every star: the mode closest to the centre of the power-excess due to solar-like oscillations $\nu_{\mathrm{max;n}}$, the radial mode one overtone below that $\nu_{\mathrm{n-1}}$, and one overtone above that, $\nu_{\mathrm{n+1}}$ [$P_{\mathrm{det}}(1)$, $P_{\mathrm{det}}(2)$, $P_{\mathrm{det}}(3)$]. It was important to use the same modes for every star so that the algorithm could be trained on the patterns between the variables.

A classifier is an algorithm that can learn a relationship between variables. The classifier will map from some initial information about the star (the X data), to some unknown information (the Y data). In this work, the X data features were magnitude (K_p or I_{mag}), $\log(g)$, π, T_{eff} and [M/H]. $\log(g)$, T_{eff} and [M/H] values are from Pinsonneault et al. (2014), π is from *Gaia* DR2 (Lindegren et al., 2018) and I_{mag} is from Høg et al. (2000), with some values imputed using regression (Section 5.2.1).

The Y data features were the P_{det} values of 3 radial modes centred around ν_{max}. An example of the final dataset for 1 year of TESS-like observations are shown in Tables 5.2 and 5.3.

5.6.2 Target selection using a classifier

The 3 sets of 81,100 datasets were prepared in Section 5.6.1 by assigning bins to every mode detection probability in each dataset. The first set consisted of *Kepler* targets (observed for up to 4 years). The second set comprised TESS targets that were observed in the Continuous Viewing Zone (1 year of observation). The third set contained the TESS targets that were observed in 1 sector (for 27 days). After this, the same random forest classification was performed once on each of the 3 sets of 81,100 datasets. The

KIC	Star Index	Iteration	$\log(g)$	π	T_{eff}	[M/H]	I_{mag}
9205705	1	1	2.758	0.688	4685	-0.39	9.89
9205705	1	2	2.758	0.688	4685	-0.39	9.19
9205705	1	3	2.758	0.688	4685	-0.39	9.79
9205705	1	4	2.758	0.688	4685	-0.39	11.30
...							
9205705	1	100	2.758	0.688	4685	-0.39	7.81
2554924	2	1	2.799	0.969	4594	0.27	8.46
2554924	2	2	2.799	0.969	4594	0.27	9.26
...							

Table 5.2: An example of the X-dataset for 1 year of TESS-like observations. Every star has it's magnitude perturbed 100 times (Section 5.3). See Table 5.3 for the equivalent Y-dataset. There are 81,100 rows in the X-dataset.

KIC	Star Index	Iteration	$P_{\text{det}}(1)$	$P_{\text{det}}(2)$	$P_{\text{det}}(3)$
9205705	1	1	1	2	2
9205705	1	2	1	2	2
9205705	1	3	1	2	2
9205705	1	4	1	2	1
...					
9205705	1	100	1	2	2
2554924	2	1	2	2	2
2554924	2	2	2	2	2
...					

Table 5.3: An example of the Y-dataset for 1 year of TESS-like observations. Every star has it's magnitude perturbed 100 times. White noise is then added to the timeseries and mode detection probabilities are calculated for 3 radial modes centred around ν_{max} . Lastly, these probabilities are put into discrete classes [0, 1 or 2]. The radial mode closest to ν_{max} is $P_{\text{det}}(2)$. There are 81,100 rows in the Y-dataset. See Table 5.2 for the equivalent X-dataset.

classifier was given 1 set of data at a time.

For each set of 81,100 datasets, the random forest classification was done in 3 stages: the data were separated, the classifier was trained, and it was tested. Firstly, the 81,100 datasets (from 811 stars) were separated into a training dataset and a testing set. 70 % of the data were used to train the classifier (56,770 datasets); 30 % of the stars were used to test the algorithm (24,330 datasets).

Secondly, the classifier was trained using the 56,770 datasets. The X and Y data in the training set were given to the algorithm ($\mathrm{X_{train}}$ and $\mathrm{Y_{train}}$). $\mathrm{X_{train}}$ comprises 56,770 sets of `log(`g`)`, π, T_{eff}, `[M/H]` and I_{mag} values. This is accompanied by $\mathrm{Y_{train}}$ with the corresponding 56,770 detection probabilities (Tables 5.2 and 5.3). The classifier trained itself by using $\mathrm{X_{train}}$ and $\mathrm{Y_{train}}$ to define distinct groups of `log(`g`)`, π, T_{eff} and `[M/H]` in the data[3].

Thirdly, the classifier was tested with $\mathrm{X_{test}}$ and $\mathrm{Y_{test}}$. $\mathrm{X_{test}}$ comprises 24,330 sets of `log(`g`)`, π, T_{eff}, `[M/H]` and I_{mag} values which the classifier is given. Without having access to the corresponding $P_{\mathrm{det}}(1)$, $P_{\mathrm{det}}(2)$ and $P_{\mathrm{det}}(3)$ values in $\mathrm{Y_{test}}$, the classifier predicted a set of Y data ($\mathrm{Y_{pred}}$). This was compared to the actual Y data for the testing set ($\mathrm{Y_{test}}$). The more similar $\mathrm{Y_{pred}}$ is to $\mathrm{Y_{test}}$, the better the classifier has replicated the data.

Two metrics were used to measure the performance of the algorithm. The first was the precision of the classifier. To calculate the precision, the differences between the 'true' detection probabilities ($\mathrm{Y_{test}}$) and the values predicted by the classifier ($\mathrm{Y_{pred}}$) were calculated. This was done for each feature ($P_{\mathrm{det}}(1)$, $P_{\mathrm{det}}(2)$ and $P_{\mathrm{det}}(3)$) separately. The mean of these differences was calculated, and weighted by the number of true-positive values in each feature.

This weighted precision $\mathcal{P}$ is given as

$$\mathcal{P} = t_{\mathrm{p}} \,/\, (t_{\mathrm{p}} + f_{\mathrm{p}})\,, \tag{5.11}$$

---

[3] https://docs.marklogic.com/

Satellite	T_{obs}	Precision	Hamming loss
Kepler	4 years	0.98	0.02
TESS	1 year	0.90	0.09
TESS	27 days	0.81	0.19

Table 5.4: Results of the classifier on the original *Kepler* dataset, and the 1-year and 27-day TESS datasets. The 'Precision' column gives the average weighted precision of the classifier across the 3 classes [0, 1, 2] and 3 features [$P_{\mathrm{det}}(1)$, $P_{\mathrm{det}}(2)$, $P_{\mathrm{det}}(3)$].

where t_{p} are true-positives and f_{p} are false-positives. The classifier's precision is its ability to not label a negative sample as positive[4].

The second was the Hamming loss[4] (Wegner, 1960) of the algorithm. This was used to give another measure of similarity between the predicted P_{det} values $\mathrm{Y}_{\mathrm{pred}}$, and the testing P_{det} values $\mathrm{Y}_{\mathrm{test}}$;

$$H_{\mathrm{loss}}(Y_{\mathrm{test}}, Y_{\mathrm{pred}}) = \frac{1}{n_{\mathrm{classes}}} \sum_{j=0}^{n_{\mathrm{classes}}-1} 1\left(Y_{\mathrm{pred}} \neq Y_{\mathrm{test}}\right). \tag{5.12}$$

A Hamming loss score of 0.0 means that $\mathrm{Y}_{\mathrm{pred}}$ is identical to $\mathrm{Y}_{\mathrm{test}}$. A score of 1.0 means that there are no similar values between $\mathrm{Y}_{\mathrm{pred}}$ and $\mathrm{Y}_{\mathrm{test}}$. Precision and Hamming loss are similar (but not identical) measures of the success of a classifier. By using both metrics, the usefulness of the classifier can be better gauged. The precision and Hamming loss of the classifier on the *Kepler* and TESS datasets are shown in Table 5.4.

We also tested the impact of increasing the number of classes in equation 5.10, and the range of P_{det}values for each class. The number of classes was varied from 2 (i.e. the mode was detected (1) or it was not (0)) to 6. The width of each bin was also varied to ensure that bins were not underpopulated. The 3 classes and P_{det} ranges given in equation 5.10 gave the best predictions for the 3 sets of 81,100 datasets (4 years of *Kepler* observation, 1 year of TESS observation and 27 days of TESS observation).

The results for the original *Kepler* dataset and for 1-year of TESS data are very good; the classifier was able to replicate the mode detection predictions of the stars robustly. This means that a classifier can be used as a tool for target selection for future missions.

---

[4] http://scikit-learn.org

Once the classifier predicted P_{det} values, the stars could be ranked from those with many detected modes, to those with the fewest. In this way, the classifier could be used as the target selection function of solar-like oscillators for TESS.

For the 27-day TESS dataset, the classifier was able to correctly predict the P_{det} values of 81 % of the modes. The classifier's precision is lower than for the 1-year TESS dataset because with these briefly-observed TESS targets, the data is very susceptible to the realization noise of the observation. While an individual mode may be detectable in one 27-day observation, the realisation noise may render it undetectable in another.

5.6.3 Feature Importance

Here, a classifier was used to predict the detection probabilities of individual modes in *Kepler* red giant stars. As well as being much faster than a conventional mode detection test, an added bonus of the classifier is that it returns the 'feature importance' of each label in the X-data.

As Table 5.2 shows, the X-data labels (or features) that are given to the classifier are `[log(`g`),` π`,` T_{eff}`, [M/H],` I_{mag}`]`. Some of these X-data labels are more important than others when predicting the detection probability of solar-like modes `[`P_{det}`(1),` P_{det}`(2),` P_{det}`(3)]`. The feature importance of an X-data label is a measure of how informative that label is when predicting `[`P_{det}`(1),` P_{det}`(2),` P_{det}`(3)]`.

The feature importance of an X-data label is calculated as the number of splits in the decision trees that include that label, divided by the total number of splits in all of the decision trees. For example, if parallax is used in 10 % of all the splits made across all of the decision trees, then parallax will have a feature importance of 0.1. The feature importance of all 5 X-data labels sums to 1.

The feature importances of the X-data labels are given in Figure 5.10. The feature with the highest influence on mode detectability is apparent stellar magnitude I_{mag}. This is not surprising: fainter stars have higher shot-noise levels. This will reduce the signal-to-noise ratio (equation 5.7) of the modes in those stars, making them less likely to be

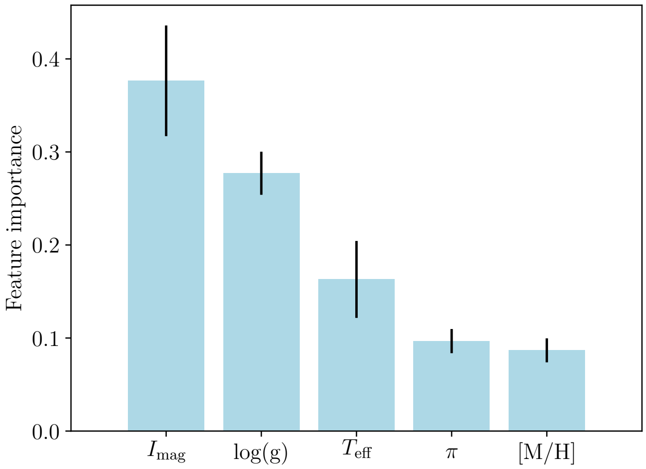


Figure 5.10: The feature importance of the 5 X-data labels.

detected.

After the apparent stellar magnitude, surface gravity $\log(g)$ also has a heavy influence on $P_{\rm det}$ value. This can be explained in terms of the oscillation amplitude of stars. Kjeldsen & Bedding (1995) showed that the oscillation amplitude observed in intensity is proportional to the bolometric luminosity, mass and effective temperature of the star;

$$A_{\rm osc} \propto \frac{L}{M\,T_{\rm eff}^{0.5}} \,. \tag{5.13}$$

By making substitutions into equation 5.13, the oscillation amplitude can be shown to be proportional to $\log(g)$ and $T_{\rm eff}$ alone. Using $\log(g) \propto M/R^2$ and $L \propto R^2\,T_{\rm eff}^4$, equation 5.13 becomes

$$A_{\rm osc} \propto \frac{{T_{\rm eff}}^{3.5}}{\log(g)} \,. \tag{5.14}$$

As a star evolves ‘up’ the Hertzsprung-Russell diagram, its surface gravity decreases while its effective temperature stays roughly the same (see the evolutionary tracks in Figure 5.1). This leads to larger oscillation amplitudes which are more likely to be detected. Mathur et al. (2011) gives a good explanation of how the granulation properties (and hence the oscillation profile) change as a star evolves.

Equation 5.13 shows that oscillation amplitude has a larger dependence on effective temperature than on surface gravity. T_{eff} might therefore be expected to have a higher feature importance than $\log(g)$ when predicting `[`P_{det}`(1),` P_{det}`(2),` P_{det}`(3)]`. Despite this, Figure 5.10 shows that T_{eff} is less informative than $\log(g)$ when predicting detection probability. This is because T_{eff} varies very little for red giant stars, see Figure 5.1.

If an X-data label has a low feature importance, this leads to two conclusions; either the label is *irrelevant*, or it is *redundant*. An irrelevant label is completely unrelated to the features that are being predicted. A redundant label is relevant to the feature being predicted, but is very similar to one or more other X-data labels in the training dataset. For example, including both I_{mag} and V_{mag} X-data labels would result in one magnitude being disregarded by the classifier, and having a low feature importance score. Although Figure 5.10 shows that apparent magnitude is an important feature when predicting detection probability, including both I_{mag} and V_{mag} would not give the classifier more information to predict detection probability, compared to only including I_{mag}. V_{mag} would be a redundant label.

Parallax π has a relatively low feature importance when predicting detection probability. Parallax is proportional to the inverse distance to the star, which is related to the apparent magnitude of the star. As well as apparent magnitude, parallax is also related to luminosity (for example, see Torres et al. 2010). Luminosity is proportional to oscillation amplitude, so including parallax does provide a small amount of extra information to the classifier.

Metallicity [M/H] is the least informative label when predicting P_{det}. Oscillation amplitude does not have a large dependence on metallicity; metallicity is less relevant than the other X-data labels when predicting detection probability.

5.6.4 Comparing results between different evolutionary states

So far in this paper, the evolutionary state of the red giant population from Davies et al. (in prep) has been ignored when predicting mode detection probability P_{det}. In reality,

Evolution	Satellite	T_{obs}	Precision	Hamming loss
RGB	*Kepler*	4 years	0.99	0.01
RGB	TESS	1 year	0.88	0.11
RGB	TESS	27 days	0.82	0.18
RC	*Kepler*	4 years	0.98	0.02
RC	TESS	1 year	0.92	0.07
RC	TESS	27 days	0.82	0.19
2CL	*Kepler*	4 years	0.98	0.02
2CL	TESS	1 year	0.85	0.13
2CL	TESS	27 days	0.75	0.26

Table 5.5: Results of the classifier when RGB, RC and 2CL stars are separated. Results are shown when the data are treated like *Kepler* stars, and when they are degraded to look like 1-year and 27-day TESS observation. The 'Precision' column gives the average weighted precision of the classifier across the 3 classes [0, 1, 2] and 3 features [$P_{\text{det}}(1)$, $P_{\text{det}}(2)$, $P_{\text{det}}(3)$].

the 1000 *Kepler* red giants used in this work are a mixture of red giant branch (RGB), red clump (RC) and secondary clump (2CL) stars. This Section investigates whether there is any difference in the predictions if the stars are first separated into RGB, RC and 2CL groups. The evolutionary states of the stars in the sample were taken from Elsworth et al. (2017).

If a red giant branch star is massive enough, it will undergo the Helium flash and become a red clump star. Red clump stars are Helium-core burning stars, and all have very similar core masses to each other. These stars have very different $g-$mode period spacings to RGB stars, but are otherwise difficult to differentiate (Chaplin & Miglio 2013, Bedding 2011, Beck et al. 2011). If a star is more massive than $\gtrapprox 1.8\,\mathrm{M}_\odot$, it will instead become a secondary clump star. This means that it will undergo Helium-core burning without the Helium flash.

The red giant dataset was separated into 3 subsets; the RGB, RC and 2CL stars. Each subset was treated in the same way as in Table 5.4. Table 5.5 gives the full list of results when the RGB, RC and 2CL stars are separated.

Table 5.5 shows that there is a negligible difference in predictions between stars undergoing Helium-core burning (RC and 2CL stars) and those that are not (RGB stars). This leads to the conclusion that a classifier can predict the mode detectability of red giants at different evolutionary states equally well. It is therefore not necessary to separate out these stars into different evolutionary states before predicting $P_{\rm det}$.

5.7 Conclusion

The solar-like oscillations of 1000 *Kepler* red giant stars from Davies et al. (in prep) were used to show that asteroseismic target selection can be done using a classifier. A classifier was used here to predict which individual solar-like oscillations inside the red giants could be detected with a future space mission. The mission in question here was TESS, although the same technique can be easily applied to other missions such as PLATO. This classifier can also be used to understand target selection bias in previous missions, such as K2 and CoRoT.

Firstly, the number of datasets was increased by perturbing the stellar magnitudes 100 times for each star. These perturbed magnitudes were drawn from a PDF of the *Kepler* shot noise model (Section 5.3). After removing stars where complete timeseries data was unavailable, this left 81,100 out of 100,000 *Kepler* datasets.

Three copies of these 81,100 datasets were made. In Section 5.4, the timeseries of one copy of 81,100 datasets was treated like 4 years of *Kepler* observations. The timeseries of the second copy was degraded to mimic 1 year of TESS observations. The timeseries of the third copy was degraded to look like 27 days of TESS observation. This was done by reducing the length of the dataset, adding white noise to the timeseries and reddening the bandbass of observation. A moving median was then calculated for the power spectra to estimate the total background in the signal. This was divided out of the spectra, leaving a signal-to-noise ratio at every frequency bin.

A detection test was then run on the SNR values at every mode frequency (Section 5.5). This gave a detection probability $P_{\rm det}$ between 0.0 and 1.0 for every mode. In order to prepare the detection probabilities before classification, each continuous $P_{\rm det}$ value was assigned a discrete class ([0,1 or 2]; equation 5.10).

In Section 5.6, a classifier was given the global photometric and spectroscopic properties of the red giant sample, along with mode detection probabilities for each star. The parameters `[log(`g`), `π`, `T_{eff}`, [M/H], `I_{mag}`]` from APOKASC (Pinsonneault et al., 2014), TGAS (Gaia Collaboration et al., 2016) and *Tycho*-2 (Høg et al., 2000) were the 5 X-data features (Table 5.2). The P_{det} values of 3 radial modes centred around ν_{max} were used as the Y-data features (Table 5.3). The classifier used the global stellar parameters (the X-data) to make predictions about mode detectability (the Y-data). The stars with the largest number of detected modes could then be selected as the red giants for observation by TESS.

The classifier successfully made predictions about the original 4 years of *Kepler* data; the algorithm had a weighted precision of 0.98 across the 3 P_{det} features. This confirms the proof of concept that classifiers can be used as a way to select solar-like asteroseismic targets before future missions. As well as this, this shows that it can be used to investigate any possible target selection bias, for example in K2 or CoRoT. Classification vastly reduces the computation time required to produce a target selection function, especially when large datasets are involved ($\geq$50,000 stars).

Degrading the red giant data to make predictions for 1 year of TESS observations was also successful. When the different evolutionary states were kept together, the predicted mode detections scored a weighted precision of 0.90 across the 3 P_{det} features (Table 5.4). This illustrates that classification is a valid target selection method for TESS targets in the Continuous Viewing Zone (CVZ, Ricker et al. 2014).

Using the classifier on 27 days of TESS data returned detection predictions with a precision of 0.81. This is too low for the classifier to be used to select solar-like oscillators for 27 days of TESS observation. The precision is lower when stars are observed by TESS for 27 days because the white noise level is too high and the length of observation is too short to make robust predictions of individual solar-like oscillations. While individual modes may be detectable in one observation, a different 27-day observation (and noise realisation) may render the modes undetectable. It may be that individual solar-like oscillations cannot be detected in 27 days of TESS data. If this is the case, then the classifier should not be expected to make robust detection predictions for these stars.

Chapter 6

Conclusion

This thesis has been focussed on the NASA Transiting Exoplanet Survey Satellite - TESS. The nominal TESS mission will span two years, during which the satellite will observe the entire sky (excluding the ecliptic plane). TESS will discover thousands of exoplanets orbitting around bright, nearby stars.

This thesis has been focussed on making asteroseismic preparations for TESS. These preparations will ensure that the satellite is successful at detecting solar-like oscillations in stars across the galaxy, in addition to discovering previously unknown exoplanets.

There were five aspects of these asteroseismic preparations. The majority of work done for this thesis concerned the selection of solar-type stars for TESS to observe at a 'high priority' cadence of 2 minutes. The solar-type stars that were selected for TESS will contain *p*-mode oscillations that the satellite will be able to detect.

To select these stars for TESS to observe, an algorithm was developed to calculate the probability of detecting solar-like oscillations within a star (Chapter 2). The list of solar-like oscillators selected for TESS was named the Asteroseismic Target List (ATL; Schofield et al. submitted).

The algorithm used two all-sky interforemetric catalogues. The majority of targets were taken from *Gaia* Data Release 2 (DR2). To supplement DR2 with additional bright stars, the Extended Hipparcos compliation (XHIP) was also used. These catalogues proved a parallax, a colour and a magnitude for each star. These were used to calculate global stellar parameters using scaling relations. The calculated parameters included the effective temperature, stellar luminosity and the frequency of maximum oscillation power, ν_{max}.

In addition, the total power due to solar-like oscillations across the Gaussian envelope was calculated. The equivalent background power in this frequency range was also calculated. The ratio of these powers - the signal-to-noise ratio - was then used to calculate the probability of detecting solar-like oscillations within the star above the stellar background. Stars were selected for TESS to observe based on the probability of detection solar-like oscillations within them.

The paper describing the method used to generate the Asteroseismic Target List was submitted to the Astrophysical Supplement Series Volume 239[1]. It is attached in Appendix A.

When constructing the ATL in Chapter 2, scaling relations were relied on to select solar-like oscillators. These are relations that relate the properties of a star with those of the Sun. Chapter 3 reviews the relations used in this thesis. It was found that these scaling relations can be relied upon to make predictions about the properties of solar-type stars.

The next chapter focusses on another aspect of the method used to construct the ATL. The method used to select stars for TESS in this thesis relied upon the estimation of the uncertainty on the global asteroseismic parameters. In Chapter 4, analytical equations to calculate the uncertainties on these global asteroseismic parameters. The equations agreed well with uncertainties from fits to *Kepler* LEGACY data.

Finally, work was done towards target election for future space missions (Chapter 5). In Chapter 5, a method of selecting solar-like oscillators for future missions to observe was developed. Here, stars are selected using Machine Learning. Using this tool, solar-like oscillators can be automatically selected for observation by any future space mission.

These will be a number of stars that TESS will observe that will have exoplanets

---

[1] http://iopscience.iop.org/journal/0067-0049

orbitting them, and will also contain solar-like oscillations. In Appendix B, predictions were made to quantify the number of exoplanet-host stars that will also display solar-like oscillations. These results were produced before the target list described in Chapter 2 was created. They have been published in the Astrophysical Journal Volume 830 Article 138[2] (Campante et al., 2016).

TESS was predicted to detect a few dozen solar-type stars which are both solar-like oscillators and exoplanet hosts. In addition, TESS will detect a few hundred equivalent low luminosity red giant hosts with oscillations.

The work done in this thesis will ensure that TESS will observe thousands of stars that contain solar-like oscillations. These observations can be used to better understand the oscillations themselves. This knowledge can then inform stellar models, and our understanding of the formation and evolution of the Galaxy.

---

[2] http://iopscience.iop.org/volume/0004-637X/830

Appendix A
The Asteroseismic Target List

The paper describing the method used to generate the Asteroseismic Target List was submitted to the Astrophysical Supplement Series Volume 239[a]. The paper contains a condensed version of the work described in Chapter 2. The work was done in collaboration with the listed coauthors.

---

[a] http://iopscience.iop.org/journal/0067-0049

Draft version December 5, 2018
Typeset using LATEX **twocolumn** style in AASTeX62

THE ASTEROSEISMIC TARGET LIST (ATL) FOR SOLAR-LIKE OSCILLATORS OBSERVED IN 2-MINUTE CADENCE WITH THE TRANSITING EXOPLANET SURVEY SATELLITE (TESS)

Mathew Schofield,[1, 2] William J. Chaplin,[1, 2] Daniel Huber,[3] Tiago L. Campante,[4, 5, 1] Guy R. Davies,[1, 2] Andrea Miglio,[1, 2] Warrick H. Ball,[1, 2] Thierry Appourchaux,[6] Sarbani Basu,[7] Timothy R. Bedding,[8, 2] Jørgen Christensen-Dalsgaard,[2] Orlagh Creevey,[9] Rafael A. García,[10] Rasmus Handberg,[2] Steven D. Kawaler,[11] Hans Kjeldsen,[2] David W. Latham,[12] Mikkel N. Lund,[2] Travis S. Metcalfe,[13, 14] George R. Ricker,[15] Aldo Serenelli,[16, 17] Victor Silva Aguirre,[2] Dennis Stello,[8, 2] and Roland Vanderspek[15]

[1] *School of Physics and Astronomy, University of Birmingham, Birmingham B15 2TT, UK*
[2] *Stellar Astrophysics Centre (SAC), Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark*
[3] *Institute for Astronomy, University of Hawai'i, 2680 Woodlawn Drive, Honolulu, HI 96822, USA*
[4] *Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, Rua das Estrelas, PT4150-762 Porto, Portugal*
[5] *Departamento de Física e Astronomia, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre, s/n, PT4169-007 Porto, Portugal*
[6] *Université Paris-Sud, Institut dAstrophysique Spatiale, UMR 8617, CNRS, Bâtiment 121, F-91405 Orsay Cedex, France*
[7] *Department of Astronomy, Yale University, PO Box 208101, New Haven, CT 06520-8101, USA*
[8] *Sydney Institute for Astronomy (SIfA), School of Physics, 2006 University of Sydney, Australia*
[9] *Université Côe d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304, Nice Cedex 4, France*
[10] *Laboratoire AIM, CEA/DSM CNRS - Univ. Paris*
Diderot IRFU/SAp, Centre de Saclay
91191 Gif-sur-Yvette Cedex, France
[11] *Department of Physics and Astronomy, Iowa State University, Ames, IA 50011, USA*
[12] *Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*
[13] *Space Science Institute, 4750 Walnut Street, Suite 205, Boulder, CO 80301, USA*
[14] *Max-Planck-Institut für Sonnensystemforschung, Justus-von-Liebig-Weg 3, 37077, Göttingen, Germany*
[15] *MIT Kavli Institute for Astrophysics and Space Research, 70 Vassar St., Cambridge, MA 02139, USA*
[16] *Institute of Space Sciences (ICE, CSIC) Campus UAB, Carrer de Can Magrans s/n, 08193 Barcelona, Spain*
[17] *Institut dEstudis Espacials de Catalunya (IEEC), C/ Gran Capità, 2-4, 08034 Barcelona, Spain*

ABSTRACT

We present the target list of solar-type stars to be observed in short-cadence (2-min) for asteroseismology by the NASA *Transiting Exoplanet Survey Satellite* (TESS) during its 2-year nominal survey mission. The solar-like Asteroseismic Target List (ATL) is comprised of bright, cool main-sequence and subgiant stars and forms part of the larger target list of the TESS *Asteroseismic Science Consortium* (TASC). The ATL uses Gaia DR2 and the Extended *Hipparcos* Compilation (XHIP) to derive fundamental stellar properties, calculate detection probabilities and produce a rank-ordered target list. We provide a detailed description of how the ATL was produced and calculate expected yields for solar-like oscillators based on the nominal photometric performance by TESS. We also provide publicly available source code which can be used to reproduce the ATL, thereby enabling comparisons of asteroseismic results from TESS with predictions from synthetic stellar populations.

Keywords: space vehicles: instruments — catalogs — surveys — stars: oscillations — stars: fundamental parameters

1. INTRODUCTION

NASA's *Transiting Exoplanet Survey Satellite* (TESS) was launched on 2018 April 18 with the main goal to detect small planets orbiting nearby stars using the transit method (Ricker et al. 2014). Its photometric data

w.j.chaplin@bham.ac.uk

will also enable high-fidelity studies of stars, and other astrophysical objects and phenomena (e.g. transients, galaxies, solar-system objects etc.). TESS is observing bright stars, including those visible to the naked eye, opening up a new discovery space to characterize stars several magnitudes brighter than those observed by the NASA *Kepler* Mission. While *Kepler* (Borucki et al. 2010) and its re-purposed follow-on Mission known as K2 (Howell et al. 2014) observed stars in only dedicated fields, TESS will survey over 90 % of the sky during its 2-year nominal mission, covering first the southern and then the northern equatorial hemispheres. TESS thus promises to provide a unique census of bright stars in the solar neighborhood.

TESS will produce Full-Frame Image (FFI) data every 30 min for the entire field of view, and 2-min (short-cadence) data on a total of approximately 200,000 targets. The short-cadence target list is comprised of several cohorts: high-priority targets for exoplanet transit searches, which form the Candidate Target List (CTL) (Stassun et al. 2018); targets from the TESS Guest Investigator (GI) program[1], the Directors Discretionary Target (DDT) and out-of-cycle Target of Opportunity (ToO) programs[2]; and targets for asteroseismic studies of stars (e.g. Chaplin & Miglio 2013).

The high-precision, high-cadence, near continuous photometric data that TESS will provide are well suited to asteroseismology. As with *Kepler* (Gilliland et al. 2010), the international asteroseismology community is coordinating efforts through the TESS *Asteroseismic Science Consortium* (TASC)[3]. Owing to their short oscillation periods, there are several classes of stars that require short-cadence data for asteroseismology. The most prominent examples are solar-type stars, here defined as cool main-sequence and sub-giant stars which show solar-like oscillations that are stochastically excited and intrinsically damped by near-surface convection. *Kepler* and K2 have provided asteroseismic detections in approximately 700 solar-type stars (Chaplin et al. 2011, 2014; Lund et al. 2016a), including about 100 *Kepler* planet hosts (Huber et al. 2013; Lundkvist et al. 2016). The main limitation for the asteroseismic yield of *Kepler*/K2 was the limited number of short-cadence target slots; there were around 500 available at any one time to the mission. That constraint will be eased dramatically for TESS, giving the potential to provide detections in thousands of solar-type stars. In addition to asteroseismic characterizations of already known planet hosts (Campante et al. 2016), TASC will also provide the TESS Science Team with such data on the bright solar-type hosts around which TESS will discover planets.

TESS will dedicate around 20,000 short-cadence targets to asteroseismology, and it is the responsibility of TASC to provide the target list. In this paper we describe the construction of the prioritized Asteroseismic Target List (ATL) of solar-like oscillators, which forms part of the overall TASC list. The breakdown of the rest of the paper is as follows. We begin in Section 2 by describing the basic philosophy underlying the construction of the ATL. Section 3 summarizes the input data. In Section 4 we discuss in detail the steps followed to produce a prioritized target list. Then in Section 5 we provide an overview of the rank-ordered list, including a prediction of the overall asteroseismic yield. We finish in Section 6 with a summary overview of the list, including information on how to access both the ATL in electronic form[4] and the `Python` codes used to construct it (the latter a `Github` repository[5][6]).

2. PHILOSOPHY FOR THE CONSTRUCTION OF THE ATL

Our goal was to produce an all-sky rank-ordered target list based on basic observables from all-sky catalogs and derived quantities which can be easily be duplicated for simulated populations (to facilitate stellar populations studies). The most obvious approach would be to select stars that are expected to show solar-like oscillations (i.e., stars cool enough to have convective envelopes), and then rank by apparent magnitude (either in the TESS bandpass, T, or Johnson I-band which is a good proxy of the TESS magnitude). However, we must also consider whether solar-like oscillations are likely to be detected in a potential target. This requires a prediction of expected photometric amplitudes of the solar-like oscillations, stellar granulation, and the expected shot and instrumental noise. A simple rank-order approach based on apparent magnitude would significantly compromise the potential yield of asteroseismic detections, and omit targets for which we expect to make asteroseismic detections.

We therefore base the ranking in our list on predictions of asteroseismic detectability, which were made using the basic methodology developed for and applied successfully to *Kepler* target selection (Chaplin et al. 2011).

---

[1] https://heasarc.gsfc.nasa.gov/docs/tess/proposing-investigations.html

[2] https://tess.mit.edu/science/ddt/

[3] http://tasoc.dk

[4] https://figshare.com/s/e62b08021fba321175d6

[5] https://github.com/MathewSchofield/ATL_public

[6] https://figshare.com/s/aef960a15cbe6961aead

While this approach is more complicated it is worth stressing that the asteroseismic predictions use simple analytical formulae, which may be applied straightforwardly to synthetic populations. All codes and data used to produce the target list are publicly available to facilitate reproducibility and the comparison with synthetic stellar populations.

3. INPUT DATA

3.1. *Input catalogs*

The ATL is mainly based on targets in Gaia Data Release 2 (DR2)[7] (Lindegren et al. 2018), supplemented at bright magnitudes by the eXtended *Hipparcos* Compilation (XHIP) (Anderson & Francis 2012). The basic set of data used to construct the ATL comprises the astrometric distances, magnitudes in the I and V bands, $(B - V)$ color, and the sky positions. From these input data we may estimate the photometric variability in the TESS bandpass caused by solar-like oscillations, granulation and shot/instrumental noise, as well as the expected duration of the TESS observations. Using these derived quantities, we then calculate the probability of detecting solar-like oscillations.

3.2. *Fundamental Stellar Properties*

Distances for *Gaia* DR2 stars (Lindegren et al. 2018) were taken from Bailer-Jones et al. (2018) using the median of the posterior calculated using their Milky Way prior. This set of distances was chosen because the Milky Way prior performs better for stars closer than 2 kpc, where the vast majority of the ATL targets are located. Distances for XHIP stars were derived by inverting the parallax. We added a zeropoint offset of 0.029 mas to all Gaia DR2 parallaxes (Luri et al. 2018; Zinn et al. 2018). After this, we discarded all targets in both catalogs which have a fractional parallax uncertainty $\sigma_\pi/\pi > 0.5$. Reddening and extinction in the V and I bands were calculated from the derived distances and sky positions (Galactic coordinates) using the Combined15 dust map from the MWDUST Python package (Marshall et al. 2006; Green et al. 2015; Drimmel et al. 2003; Bovy et al. 2016).

While I-band magnitudes are available for XHIP targets, this is not the case for most of the *Gaia* DR2 targets. This is important because the I magnitudes are needed to estimate the shot noise in the TESS bandpass. We therefore used $(B - V)$ colors and apparent V magnitudes to derive the required values. The preferred source for both inputs was the revised *Hipparcos* catalog (van Leeuwen 2007). If those data were unavailable, we used the *Tycho-2* catalog (Hog et al. 2000); and failing that, we took values from the AAVSO All-Sky Photometric Survey (APASS; Henden et al. 2009).

The input $(B-V)$ colors were first de-reddened, using the previously calculated $E(B-V)$, and then converted to $(V-I)$ using the polynomials in Caldwell et al. (1993). The coefficients of the polynomial depend upon whether the target is classified as a "giant" or "dwarf". Here, we separated targets using an empirically derived relation in M_g, the absolute magnitude in the *Gaia* bandpass, and $(B-V)$, classifying stars with $M_g > 6.5\times(B-V)-1.8$ as dwarfs, and the rest as giants. Once $(V - I)$ had been calculated for all of the stars, the I magnitudes were estimated from V and $(V-I)$. The derived I magnitudes were then reddened using the previously estimated A_I to calculate the TESS noise (see Section 4).

Dereddened $(B-V)$ colors were used to estimate stellar effective temperatures $T_{\rm eff}$, using color-temperature relations of the form:

$$\log(T_{\rm eff}) = a + b(B-V) + c(B-V)^2 + ..., \quad (1)$$

where the best-fitting coefficients were taken from Torres (2010). Luminosities, L, were calculated from

$$\log(L/{\rm L}_\odot) = 4.0 + 0.4M_{\rm bol\odot} - 2.0\log\pi - 0.4\,(V_{\rm mag} - A_V + {\rm BC}_V)\,. \quad (2)$$

Note that V magnitudes were first de-reddened using the previously calculated A_V, while the bolometric corrections, $BC_{\rm V}$, were taken from Flower (1996), as presented in Torres (2010), with $M_{{\rm bol},\odot} = 4.73 \pm 0.03$ mag. Finally, we estimated radii using the Stefan-Boltzmann law $L \propto R^2 T_{\rm eff}^4$, using $T_{{\rm eff},\odot} = 5777\,{\rm K}$.

3.3. *Comparison to Literature Values*

We compared our estimated stellar properties with several literature sources. The PASTEL catalog (Soubiran et al. 2016) includes spectroscopically-determined effective temperatures for over 60,000 stars. Figure 1 compares our derived temperatures with PASTEL for stars that are common to both lists. We observe a good agreement, with a residual median and scatter of 102 K and 146 K, respectively. We furthermore compared our temperatures with values listed in Huang et al. (2015), which compiled empirical temperatures derived from optical long-baseline interferometry (e.g. Mozurkewich et al. 2003; Boyajian et al. 2012a,b, 2013). Figure 2 again shows good agreement, with a residual median and scatter of 109 K and 173 K, respectively. Both comparisons show that our temperatures are on average ~ 100 K hotter, which is comparable to previously found offsets

---

[7] http://www.cosmos.esa.int/gaia

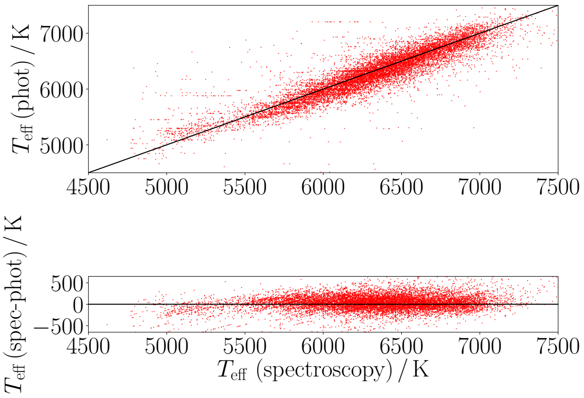


Figure 1. Comparison between effective temperatures from high-resolution spectroscopy (as listed in the PASTEL catalogue) and the ATL. The solid line shows the 1:1 relation. The horizontal lines of datapoints exist because the PASTEL catalogue gives several effective temperatures for some stars. The residual median and scatter is 102 K and 146 K, respectively.

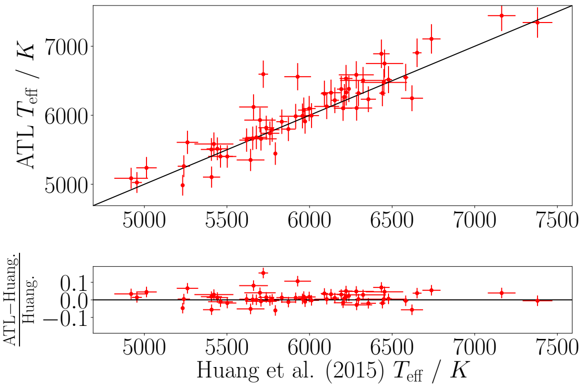


Figure 2. Comparison between effective temperatures from long-baseline interferometry (as compiled by Huang et al. 2015) and the ATL. The solid line shows the 1:1 relation. The residual median and scatter are 109 K and 173 K, respectively.

between temperature scales (Pinsonneault et al. 2012) and well within the systematic uncertainty of the fundamental interferometric temperature scale itself (e.g. White et al. 2018). Based on these comparisons we have adopted a conservative uncertainty of 3% on the temperatures in the ATL, which encompasses both random and systematic uncertainties from the literature comparisons.

Next, we compared radii in the ATL to a selection of bright stars in Silva Aguirre et al. (2012) and Bruntt et al. (2010). Silva Aguirre et al. (2012) derived radii for a small number of *Kepler* solar-type stars that have detections of solar-like oscillations as well as precise *Hipparcos* parallaxes. Bruntt et al. (2010) estimated the radii of even brighter stars using two approaches: first, using measurements of limb-darkened stellar angular diameters and stellar parallaxes; and second, using the Stefan-Boltzmann law with luminosities derived from V-band magnitudes, bolometric corrections and parallaxes, and spectroscopic temperatures, i.e., the basic approach we have used but with some different observables. Figure 3 shows the comparison with between ATL and those literature values. We observe excellent agreement, with a residual median and scatter of 0.04 % and 0.07 %, respectively. Overall, these comparisons confirm that the stellar properties derived in the ATL do not suffer from large systematic errors when compared with literature values.

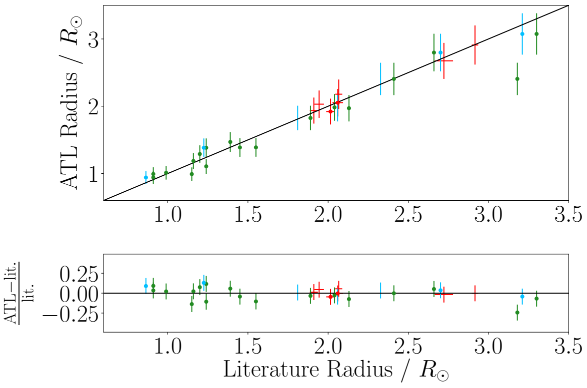


Figure 3. Comparison between the literature radii derived from parallaxes (red and blue symbols) and interferometry (green symbols). We observe good agreement, with a residual median and scatter of 0.04 % and 0.07 %.

4. ATL CONSTRUCTION

4.1. *Consolidation of DR2 and XHIP entries*

Having removed stars with large fractional parallax uncertainties (see Section 3.2), we combined the retained stars from DR2 and XHIP into a single list to be treated homogeneously. This combined list contained over 300,000 stars. Most had entries in the DR2 catalog, with only a small number in XHIP. However, there were $\sim$17,000 stars which existed in both lists. We broke this degeneracy by using data and derived parameters from the catalog whose target entry had the smaller fractional parallax uncertainty of the two. Not surprisingly, in the vast majority of cases the DR2 entries were selected, with only a handful of bright XHIP targets being re-

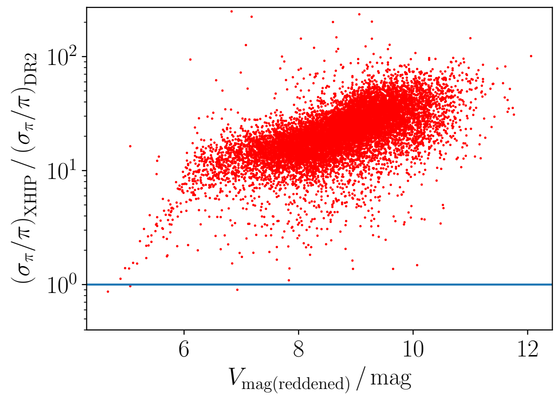


Figure 4. The stars with both DR2 and XHIP entries. The fractional parallax σ_π/π value was calculated from the DR2 and XHIP entries for each star separately. The parameters from the catalogue with the lower σ_π/π value was chosen. XHIP properties were only used for the few stars below the blue line.

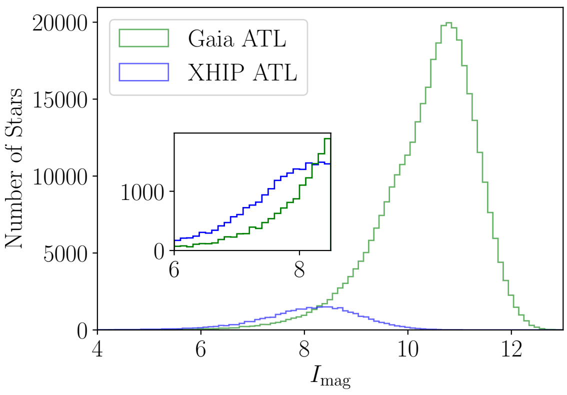


Figure 5. The I_{mag} distribution of the XHIP and DR2 catalogues. Shown here after T_{eff} /L cuts, but before P_{det} was calculated. The inset shows the I_{mag} region where the two catalogues overlap.

tained where *Hipparcos* outperforms *Gaia* (see Figure 4).

4.2. *Down-selection to solar-like oscillating short-cadence targets*

From the above combined list, we selected targets that are potential solar-like oscillators. To do this, we retained all stars that lie on the cool side (redwards) of the δ Scuti instability strip, i.e., those having $T_{\mathrm{eff}} < T_{\mathrm{red}}$, with the red-edge temperature defined as Chaplin et al. (2011):

$$T_{\mathrm{red}} = 8907\,\mathrm{K} \times (L/L_\odot)^{-0.093}. \tag{3}$$

We further restricted to targets that have predicted dominant oscillation frequencies requiring the TESS short-cadence (2-min) data. Solar-like oscillators present a rich spectrum of detectable overtones, with oscillation power following a Gaussian-like envelope centered on the so-called frequency of maximum oscillations power, ν_{max}. We retained all targets having $\nu_{\mathrm{max}} \geqslant 240\,\mu$Hz. This represents, to reasonable approximation, an upper-limit cut in luminosity that discards low-luminosity red-giants at or just above the base of the red-giant branch, i.e., giants whose solar-like oscillations can be very readily resolved in the TESS 30-min long-cadence FFI data. The 240 μHz limit was set deliberately to lie below the FFI Nyquist frequency of 278 μHz to account for uncertainties in the ATL-based predictions and also to provide a reasonable sample of targets in short-cadence whose oscillation spectra are reasonably close to the Nyquist limit. Experience from *Kepler* has shown that such spectra can be difficult to analyze using long-cadence data only, due to aliasing about the Nyquist frequency (e.g. Yu et al. 2016).

The boundary in the L-T_{eff} plane for the ν_{max} cut follows from the approximate relation (see Campante et al. 2016):

$$\nu_{\mathrm{max}} = \nu_{\mathrm{max},\odot} \left(\frac{R}{R_\odot}\right)^{-1.85} \left(\frac{T_{\mathrm{eff}}}{T_{\mathrm{eff},\odot}}\right)^{0.92}, \tag{4}$$

which, combined with $L \propto R^2 T_{\mathrm{eff}}^4$ and setting $\nu_{\mathrm{max}} = 240\,\mu$Hz, defines the boundary

$$L/L_\odot \leqslant 16.7 \times \left(\frac{T_{\mathrm{eff}}}{T_{\mathrm{eff},\odot}}\right)^5 \tag{5}$$

for retaining targets. Figure 5 shows the I magnitude distribution of the Hipparcos and Gaia subsamples of the ATL after these H-R diagram cuts have been applied. As expected, Gaia dominates the faint end of the ATL, and the drop-off at $I \sim 11$ is caused by the fractional parallax precision cut described in Section 3.2.

4.3. *Estimation of asteroseismic detection probabilities*

To calculate asteroseismic detection probabilities we used the approach developed by Chaplin et al. (2011), which has been applied successfully to short-cadence target selection for *Kepler* Objects of Interest in the *Kepler* nominal mission and, more recently, to short-cadence target selection for solar-type stars observed with K2 (e.g. see Chaplin et al. 2015; Lund et al. 2016a,b). The approach is based on predicting the global signal-to-noise ratio in the oscillation spectrum, i.e., the predicted total power in the observed solar-like oscillations divided by the total power from granulation, shot and instrumental noise, summed across the range in frequency occupied by the modes. The total oscillation and granulation power across the frequency range of interest centered on the predicted $\nu_{\max}$ may be calculated from the previously derived L, $T_{\rm eff}$ and R. The shot and instrumental noise depend on the instrumental performance and the apparent magnitude of targets in the instrumental bandpass. The duration of the observations is also an important factor: at a given global signal-to-noise ratio, the detection probability will rise as the length of observations is increased.

The formulation by Chaplin et al. (2011) was updated for the TESS instrumental specifications in Campante et al. (2016). We followed that revised recipe in detail here, and refer the reader to Section 3 of Campante et al. (2016) for the relevant steps and relations. We have made some changes to the estimation of the TESS noise, to reflect updates to information that is available on the instrumental performance. We describe those small changes next in Section 4.3.1.

4.3.1. *Updates to noise predictions*

The predicted instrumental noise is dominated by the shot noise, but also includes contributions to represent contamination from nearby stars, and readout noise. Since the ATL targets are bright, contamination is expected to be modest, in spite of the large point-spread function of TESS. Note that we assumed that the systematic noise floor of 60 ppm per hour (Sullivan et al. 2015) is negligible, since this is a design threshold requirement for meeting core exoplanet science deliverables and will not reflect the actual performance.

As in Campante et al. (2016), we used the `calc_noise` IDL procedure on the TESS Science Team Wiki[8] to calculate the instrumental noise, which takes the I-band magnitude as its main input. There are two updates: (i) the absolute calibration of the expected noise levels is now slightly higher, due to a reduced estimated effective aperture size for the instrument; and (ii) the expected number of pixels, $N_{\rm mask}$, in each stellar pixel mask is now smaller, which has the effect of reducing noise levels. Updated mask sizes were calculated using the simple parametric model[9]:

$$N_{\rm mask} = 10^{0.8464-0.2144\times(I_{\rm mag}-10.0)}, \qquad (6)$$

and the number of pixels was rounded up to the nearest whole number. Once calculated, the individual instrumental noise contributions (see Figure 6) were summed in quadrature to give the total instrumental noise per 2-min cadence.

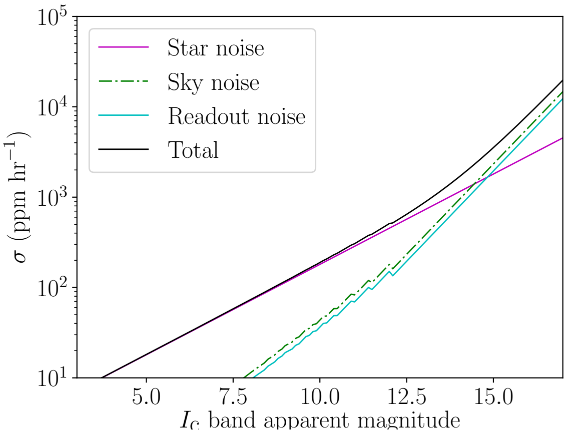


Figure 6. Individual noise contributions (coloured lines) and total noise budget (black line) as a function of apparent I magnitude used to calculate asteroseismic detection probabilities.

4.3.2. *The Observation time in the TESS field of view*

TESS comprises four CCD cameras. Each CCD images a 24° × 24° area on the sky, with the total collecting area of the four cameras at any given time being a strip of dimensions 24° (ecliptic longitude) × 96° (ecliptic latitude). TESS will survey the sky south of the ecliptic in its first year of science operations, with the hemisphere divided into 13 strips. Each resulting *sector* pointing will last, on average, about 27.4 days. The durations of each sector pointing differ by up to 1.5 days due to variations in the length of the spacecraft's orbit. The sky north of the ecliptic will be observed in the second year of nominal science operations.

The majority of TESS targets will be observed over only one 27-day sector. The duration increases for latitudes significantly above or below the ecliptic plane, because targets may then be observed in more than one sector pointing, reaching a maximum of 13 sectors, i.e., about 351 days, at the ecliptic poles (the Continuous Viewing Zone, see Figure 7). TESS will not observe targets within ±6° of the ecliptic during its nominal mission, and those stars were removed from our list.

Figure 7 shows that there will be small gaps between sectors. At the time the ATL was delivered, the initial pointing at the commencement of science operations was not known. As can be seen from the figures, that will influence not only which stars are missed by TESS (i.e., those falling in the sector-to-sector gaps at low ecliptic latitudes) but also the numbers of sectors for which targets at higher latitudes will be observed. Here, we ig-

---

[8] https://spacebook.mit.edu/

[9] https://dfm.io/

nore the low-latitude gaps and assume that all stars with ecliptic latitudes beyond ±6° are potentially observable. Targets that fall in the gaps will be discarded when the TESS team compiles actual target lists for each known pointing.

For higher-latitude targets, there are several options open to us. We could adopt a particular pointing and then compute the resulting number of observation sectors for each target for input to the asteroseismic detection recipe. We could instead estimate the minimum and maximum potential number of observation sectors for each target, which depend on the ecliptic latitude but not the exact pointing, and use one or the other as input to the detection recipe. While this choice will affect the rank ordering of targets based on the detection probability, it turns out that the resulting changes in ranking are typically a few hundred places or less, a change that is very unlikely to influence whether targets with potentially detectable oscillations are observed by TESS. As such we adopted the simpler first option and assumed an initial pointing of $E_{\rm long} = 0°$.

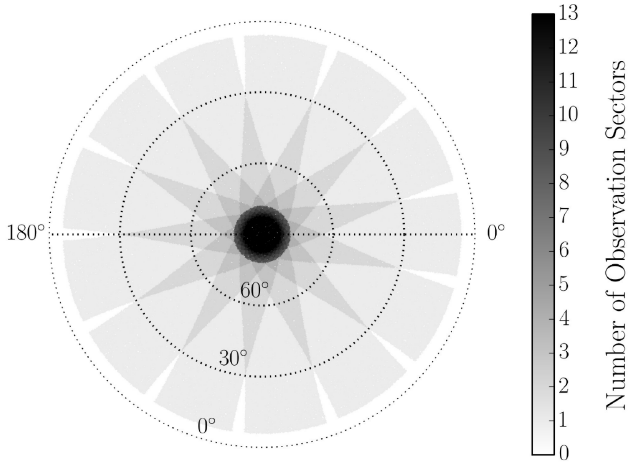


Figure 7. The field of view from one of the ecliptic poles. The colourbar represents how long a part of the sky will be observed by TESS. The dotted circles are lines of constant latitude (0°, 30° and 60°). The centre of the image has a latitude of 90°. The outer dotted circle at 0° latitude has labels for 0° and 180° longitude. The horizontal dotted line represents longitude values of 0° and 180°.

The ecliptic position $(E_{\rm long}, E_{\rm lat})$ determines how long a star can be observed. To determine whether a star is observable in any given sector pointing we must define the longitudes of the center of each observing sector, $E_{\rm CCD}$, and the longitude range, $\phi_{\rm range}$, that the cameras cover at a given latitude (i.e., the latitude of the star). Figure 8 gives a pictorial representation of $\phi_{\rm range}$. The black circle in Figure 8 is a line of constant latitude. In Figure 8, the satellite is represented by the small red circle in the center of the image. The red dashed lines show the width of the field-of-view of TESS. These are the edges of $\phi_{\rm range}$.

$\phi_{\rm range}$ is given by

$$\phi_{\rm range} = \frac{24°}{\cos(E_{\rm lat})}, \tag{7}$$

where $E_{\rm lat}$ is the latitude of the star in question and 24° is the width of the field covered by the CCD cameras at 0° latitude. If the longitude of the star lies within $\pm\phi_{\rm range}/2$ then the image of the star will be captured by a camera. In order to check this, the difference between the center of the CCD ($E_{\rm CCD}$) and the longitude of the star ($E_{\rm long}$) must be calculated. This difference is given by

$$\phi_{\rm Diff} = \begin{cases} |E_{\rm CCD} - E_{\rm long}|, \\ 360° - |E_{\rm CCD} - E_{\rm long}|. \end{cases} \tag{8}$$

Equation 8 will produce two values of $\phi_{\rm Diff}$, as shown by the blue and green lines in Figure 8. Only the smaller distance between $E_{\rm long}$ and $E_{\rm CCD}$ should be taken as the distance between the star and the center of the field-of-view. The longitudinal position of a star is marked in Figure 8 by the orange line.

Now, if $\phi_{\rm range} \geqslant \phi_{\rm Diff}$, the star will be observed in that region. In Figure 8, the length of the blue arc will be the accepted value for $\phi_{\rm Diff}$, since it is shorter than the green arc. However, although the blue arc is the shorter of the two, $\phi_{\rm range} < \phi_{\rm Diff}$ and so the star will not be observed in this sector.

Using Equations 7 and 8, we determined which stars would be observable in the first sector pointing of each ecliptic hemisphere, again taking each to be centered on $E_{\rm CCD} = 0°$. The same calculations were then repeated for each subsequent pointing, with every adjacent pointing shifted by $E_{\rm CCD} = 360°/13 = 27.7°$. The observing time $T_{\rm obs}$ was then obtained from the maximum contiguous number of sectors that each star is observed.

4.3.3. *Rank-ordering the ATL using the detection probabilities*

The analysis of the *Kepler* sample demonstrated that amplitudes of solar-like oscillations are progressively reduced relative to predictions from scaling relations when moving from late to early F-type stars. Chaplin et al. (2011) attempted to explicitly capture this effect by introducing an attenuation factor β, which was also adopted by Campante et al. (2016) to describe the maximum amplitude for radial mode oscillations in the TESS

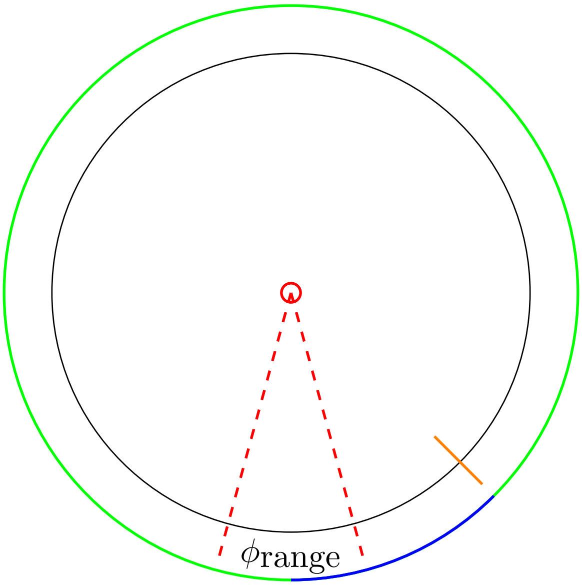


Figure 8. An example of what is calculated in order to determine whether a star lies inside $\phi_{\rm range}$. In this example, the star's longitude (represented by the orange line) lies outside of the satellite's field of view marked by the red dashed lines (Equation 7).

bandpass:

$$A_{\rm max} = 0.85 \times 2.5\,{\rm ppm} \times \beta \times \left(\frac{R}{R_\odot}\right)^{1.85} \times \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{0.57}. \quad (9)$$

Here, β is given by:

$$\beta = 1.0 - \exp\left[(T_{\rm red} - T_{\rm eff})/1550\,{\rm K}\right], \quad (10)$$

where $T_{\rm red}$ is the previously defined temperature on the red-edge of the δ Scuti instability strip at the luminosity of the target (see Equation 3). The attenuation given by β reduces predicted mode amplitudes in hotter stars, and hence lowers detection probabilities and the associated rank-ordering of those targets. Figure 9 illustrates the effect by showing all ATL stars with detection probabilities greater than 50 % with and without including the β factor. As expected, the β factor strongly reduces the number of stars with significant detection probabilities, especially towards the instability strip.

Rank-ordering the ATL using the detection probabilities including the β factor would optimize the yield of asteroseismic detections with TESS. However, using the β factor would also strongly bias against making new discoveries in stars that do not fit the trend in asteroseismic amplitudes shown by the *Kepler* sample on which the detection recipe is based (which is, by definition, an already biased sample).

The group most affected by this comprises hot F-type stars, which lie at the boundary where solar-like oscillations diminish to undetectable amplitudes and classical pulsations driven by the κ mechanism start to become excited. Determining the details of this transition is of considerable interest for understanding the driving and damping of oscillations, and intriguing examples of "hybrid stars" showing signatures of solar-type oscillations and classical pulsators have already been detected (Kallinger & Matthews 2010; Antoci et al. 2011), leading to suggestions of new pulsation driving mechanisms (Antoci et al. 2014). The sampling of targets in this region was sparse for *Kepler*, and limited by the small number of short-cadence target slots available at any one time. There is now the potential to address those issues with TESS.

To mitigate the strong bias against hot stars in the ATL we define a new probability, $p_{\rm mix}$, as follows:

$$p_{\rm mix} = (1-\alpha)\,p_{\rm vary} + \alpha\,p_{\rm fix}. \quad (11)$$

Here, $p_{\rm vary}$ is the detection probability calculated using the β factor, $p_{\rm fix}$ is the detection probability calculated by fixing $\beta = 1$ for all stars (i.e. ignoring amplitude attenuation), and α regulates the relative weighting between $p_{\rm vary}$ and $p_{\rm fix}$. After investigating the rank ordered lists using a range of values of α, we found that $\alpha = 0.5$ (i.e. equal weighting between $p_{\rm vary}$ and $p_{\rm fix}$) provides the best overall compromise between obtaining a significant yield and including enough hot stars at high ranks. For the remainder of the paper, all ranked lists in the ATL were calculated with detection probabilities using $\alpha = 0.5$.

5. OVERVIEW OF THE ASTEROSEISMIC TARGET LIST

5.1. *Distribution across H-R Diagram*

Figure 10 shows the distribution of the 18,000 top-ranked stars in the ATL in an H-R diagram, split into bins of 2000 stars each. Similar to Figure 9, the sharp edges are caused by the down-selection of solar-type dwarfs and sub-giants using Equations 3 and 4. As expected, the top-ranked stars are dominated by cool, high-luminosity sub-giants with intrinsically high detection probabilities. Progressing towards lower ranks, a larger number of hot stars appear, a direct consequence of relaxing the β amplitude dilution factor described in Section 4.3.3.

The distribution of targets in Figure 10 demonstrates the well known bias of asteroseismic detections against cool, low-mass stars due to their intrinsically low oscillation amplitudes (see, e.g., Chaplin et al. 2011). In total, only six stars ranked among the top 25,000 in the

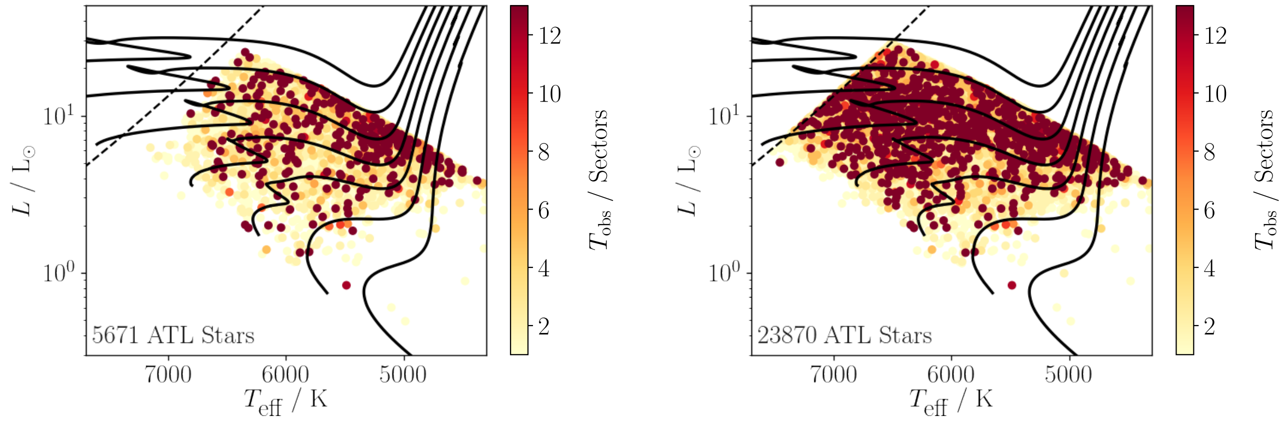


Figure 9. H-R diagram of all ATL stars with detection probabilities greater than 50 % with (left panel) and without (right panel) including the β factor, which accounts for the attenuation of oscillation amplitudes towards the red edge of the instability strip (dashed line). Solid lines show solar metallicity evolutionary tracks with masses from $0.8\,M_\odot$ to $2.0\,M_\odot$ in steps of $0.2\,M_\odot$. Note that the sharp edges are due to cuts at the red edge of the instability strip (Equation 3) and stars oscillating with frequencies accessible with TESS FFI data (Equation 4).

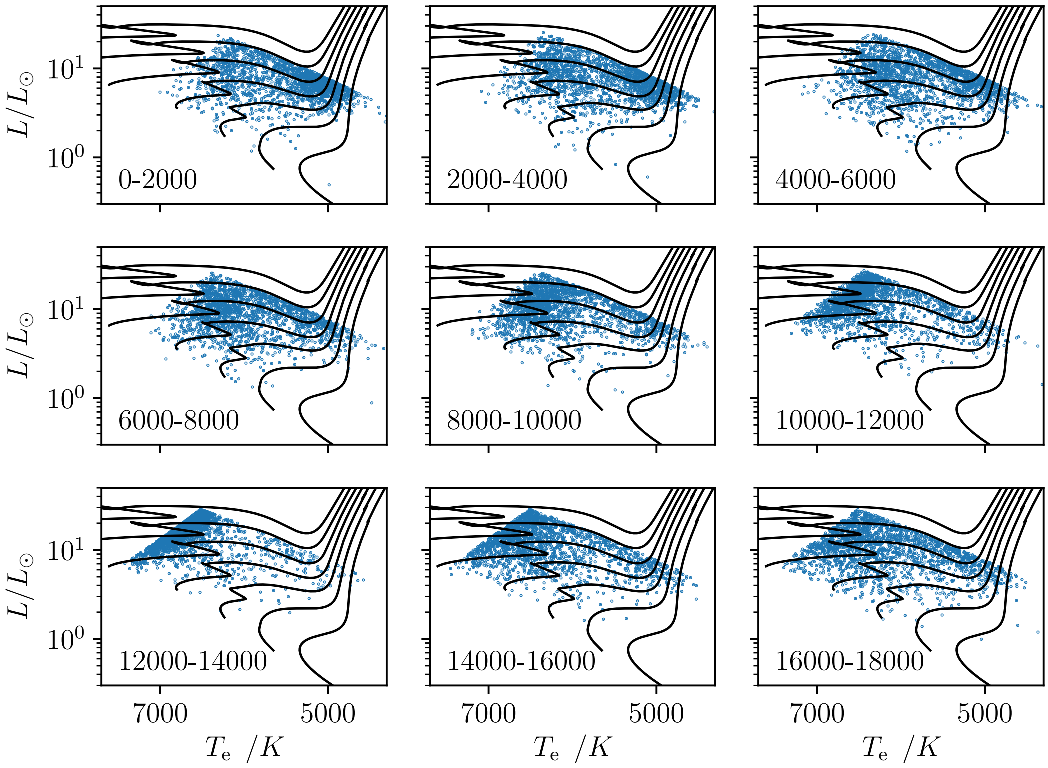


Figure 10. H-R diagram of stars in the ATL. Each panel shows two thousand stars according to their ranking, with highest to lowest ranked stars running from the top left to the bottom right panel. Black lines show evolutionary tracks with masses from $0.8\,M_\odot$ to $2\,M_\odot$ in steps of $0.2\,M_\odot$.

ATL have luminosities less than solar. This makes the ATL highly complementary to the exoplanet target list, which prioritizes cool dwarfs due to the improved probability of finding small transiting exoplanets. We note that all solar-type stars having a magnitude $T < 6$ in the TESS bandpass are automatically included in the TESS 2-minute cadence target list (Stassun et al. 2018), irrespective of their position on the ATL.

5.2. *Expected Yield*

To estimate the expected yield of asteroseismic detections, we performed a Monte-Carlo simulation as follows. For each star in the ATL, we drew a uniform random number n between zero and unity and counted the target as a potential seismic detection if $n < p_{\rm vary}$. We recall here that $p_{\rm vary}$ provides a conservative yield, since the amplitude dilution factor β may be overestimated. To determine whether the target would be observed in 2-minute cadence, we adopted a starting ecliptic longitude of zero degrees for the first observing sector and picked the top 450 targets (the per-sector allocation of TASC) in the ATL that fall on silicon in that sector. We then repeated this for each of the 13 sectors in the southern ecliptic hemisphere, adding new detections to the list each time.

The predicted TESS yield for the first full year of science operations, corresponding to Guest Investigator Cycle 1, is ~ 2500 oscillating targets, already a five-fold increase over the yield from the *Kepler* mission. Of these detections, the majority are observed for a single sector (~ 1500), while ~ 200 targets are expected to be observed for 10 sectors or more. The second year of nominal science operations (Cycle 2) is expected to produce a similar yield, bringing the total expected number of detections to 5000 stars. We emphasize that these estimates only take into account stars on the ATL and ignore potential overlaps with other target lists (such as the CTL and Guest Investigator Program), which would result in a slightly higher yield. They also assume that our adopted noise model provides a good description of the actual, in-flight photometric precision.

Figure 11a compares the predicted asteroseismic yield of TESS to detections for dwarfs and sub-giants from the *Kepler* mission (Chaplin et al. 2014). As expected, the TESS yield is skewed towards evolved sub-giants with intrinsically larger amplitudes, and contains a smaller number of cool dwarfs (for which higher photometric precision is required for a detection). Importantly, Figure 11b demonstrates that the TESS detections will be on average 4-5 magnitudes brighter than *Kepler*, which follows from the difference in aperture size. Similar to the characterization of transiting exoplanets, this will enable significantly more powerful complementary follow-up observations, including measurements of angular diameters using optical long-baseline interferometry (which was only possible for a handful of *Kepler* dwarfs and subgiants Huber et al. 2012; White et al. 2013). Overall this demonstrates that TESS will excel in a significantly different parameter space than *Kepler*, in particular for evolved subgiants which exhibit mixed modes that allow powerful constraints on the interior structure (e.g. Chaplin & Miglio 2013; Hekker & Christensen-Dalsgaard 2017).

6. SUMMARY AND CONCLUSIONS

We have presented the construction of the Asteroseismic Target List (ATL) for solar-like oscillators to be observed in 2-minute cadence by the TESS Mission. The main characteristics of the ATL can be summarized as follows:

- The ATL includes 25,000 bright main-sequence and subgiant stars that have at least a 5% probability of detecting solar-like oscillations with TESS. Detection probabilities were calculated from stellar properties estimated from colors, parallaxes and apparent TESS magnitudes. The ranking of targets is based on a mixture of detection probability and the prioritization of hot stars, for which the oscillation amplitudes are poorly understood.

- We have validated our derived stellar properties against spectroscopy, asteroseismology and interferometry, finding good agreement. In addition to the asteroseismic detection probabilities, the ATL provides a homogeneous catalog of stellar properties for bright solar-type stars observed by TESS.

- Based on the nominal TESS photometric performance and the number of target slots assigned to the ATL, we expect that TESS will increase the number of solar-type stars with detected oscillations by an order of magnitude over *Kepler*. Most of the detections will be in evolved subgiants, with only a small number of detections in unevolved main-sequence stars.

- The `Python` code used to produce the ATL is publicly available on `Github`[10][11], allowing full reproducibility of the asteroseismic target selection for comparison with population synthesis models.

---

[10] https://github.com/MathewSchofield/ATL_public

[11] https://figshare.com/s/aef960a15cbe6961aead

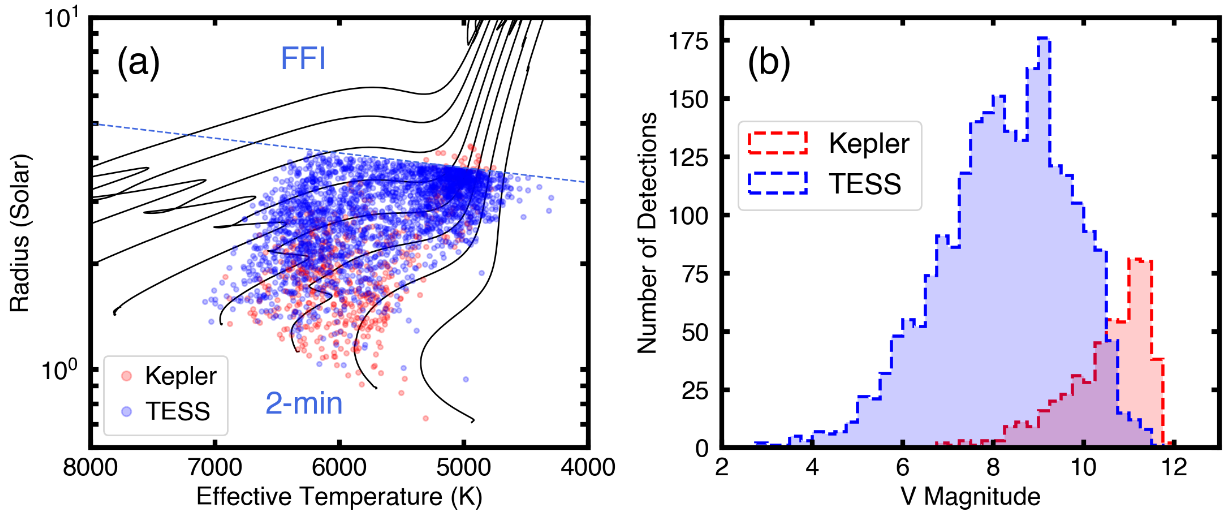


Figure 11. The predicted asteroseismic yield for the first year of TESS science operations (Cycle 1). Panel (a): Radius versus effective temperature for all expected TESS detections (blue) and the detections for dwarfs and sub-giants by *Kepler* (red). The blue dashed line marks the approximate radius limit above which oscillations can be confidently detected using FFI light curves. Black lines show evolutionary tracks. Panel (b): Approximate V magnitude distribution of the expected TESS yield (blue) and the *Kepler* yield (red).

The ATL itself is available in electronic form[12]. The columns of the ATL are shown in Table 1.

The yield of solar-like oscillators with TESS is expected to continue the asteroseismic revolution initiated by CoRoT and *Kepler*. In particular, TESS is expected to deliver detections in the nearest solar-type stars for which strong complementary constraints (e.g. from Hipparcos/Gaia parallaxes and interferometry) are available, allowing powerful inferences on the interior structure of stars and stellar ages, including exoplanet host stars. Our improved understanding of the excitation mechanism of solar-like oscillations probed by the large sample of TESS stars observed in 2-minute cadence will also be helpful to optimize target selection for future missions such as PLATO (Rauer et al. 2014).

M.S. acknowledges support from the University of Birmingham. W.J.C., G.R.D., A.M. and W.H.B. acknowledge support from the UK Science and Technology Facilities Council (STFC). D.H. acknowledges support by the National Aeronautics and Space Administration (NASA) under grant NNX14AB92G issued through the *Kepler* Participating Scientist Program. T.L.C. acknowledges support from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 792848 and from grant CIAAUP-12/2018-BPD. T.S.M. acknowledges support from NASA grant NNX16AB97G. A.S. acknowledges partial support from grants ESP2017-82674-R (Spanish Ministry of Economy) and SGR2017-1131 (Generalitat de Catalunya). Funding for the Stellar Astrophysics Center is provided by The Danish National Research Foundation (Grant agreement no.: DNRF106).

REFERENCES

Anderson, E., & Francis, C. 2012, Astronomy Letters, 38, 331, doi: 10.1134/S1063773712050015

Antoci, V., Handler, G., Campante, T. L., et al. 2011, Nature, 477, 570, doi: 10.1038/nature10389

Antoci, V., Cunha, M., Houdek, G., et al. 2014, ApJ, 796, 118, doi: 10.1088/0004-637X/796/2/118

Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, AJ, 156, 58, doi: 10.3847/1538-3881/aacb21

Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977, doi: 10.1126/science.1185402

Bovy, J., Rix, H.-W., Green, G. M., Schlafly, E. F., & Finkbeiner, D. P. 2016, The Astrophysical Journal, 818, 130, doi: 10.3847/0004-637X/818/2/130

---

[12] https://figshare.com/s/e62b08021fba321175d6

Table 1. Column headers of the ATL

01:	TESS *Input Catalog* (TIC) ID
02:	*Tycho* 2 ID
03:	*Hipparcos* ID
04:	*Gaia* Data Release 1 ID
05:	*Gaia* Data Release 2 ID
06:	Maximum number of contiguous observing sectors (1-13)
07:	Rank based on $p_{\rm mix}$
08:	(Flag) 1: Rank manually adjusted or star added to list afterwards
09:	(Flag) 1: High priority star (for 20-sec cadence); 0: 120-sec cadence star
10:	Ecliptic latitude (deg)
11:	Ecliptic Longitude (deg)
12:	Galactic latitude (deg)
13:	Galactic longitude(deg)
14:	Equatorial declination (deg)
15:	Equatorial right ascension (deg)
16:	TESS-band apparent magnitude (mag)
17:	V-band apparent magnitude (mag)
18:	I-band apparent magnitude (mag)
19:	Extinction in I-band (mag)
20:	Extinction in V-band (mag)
21:	$(B-V)$ color (mag)
22:	$(B-V)$ color uncertainty (mag)
23:	Reddening of $(B-V)$ color (mag)
24:	Parallax (mas)
25:	Parallax uncertainty (mas)
26:	(Flag) 1: *Hipparcos* parallaxes used; NaN: DR2 parallaxes used
27:	Distance (Kpc)
28:	Distance uncertainty (Kpc)
29:	(Flag) 1: Bailer-Jones et al. (2018) distances are provided
30:	Bailer-Jones et al. (2018) distance (upper limit) (pc)
31:	Bailer-Jones et al. (2018) distance (median value) (pc)
32:	Bailer-Jones et al. (2018) distance (lower limit) (pc)
33:	Luminosity L (in $L_\odot$)
34:	$\nu_{\rm max}$ (μHz)
35:	Radius R (in $R_\odot$)
36:	$T_{\rm eff}$ (K)
37:	Global asteroseismic SNR ($\beta = 1$)
38:	Global asteroseismic SNR ($0 \leq \beta \leq 1$)
39:	$p_{\rm mix}$ composite probability
40:	$p_{\rm vary}$ probability ($0 \leq \beta \leq 1$)
41:	$p_{\rm fix}$ probability ($\beta = 1$)

Boyajian, T. S., McAlister, H. A., van Belle, G., et al. 2012a, ApJ, 746, 101, doi: 10.1088/0004-637X/746/1/101

Boyajian, T. S., von Braun, K., van Belle, G., et al. 2012b, ApJ, 757, 112, doi: 10.1088/0004-637X/757/2/112

—. 2013, ApJ, 771, 40, doi: 10.1088/0004-637X/771/1/40

Bruntt, H., Bedding, T. R., Quirion, P.-O., et al. 2010, Monthly Notices of the Royal Astronomical Society, 405, 1907, doi: 10.1111/j.1365-2966.2010.16575.x

Caldwell, J. A. R., Cousins, A. W. J., Ahlers, C. C., van Wamelen, P., & Maritz, E. J. 1993, South African Astronomical Observatory Circular, 15, 1

Campante, T. L., Schofield, M., Kuszlewicz, J. S., et al. 2016, ArXiv e-prints, 1608, arXiv:1608.01138

Chaplin, W. J., & Miglio, A. 2013, Annual Review of Astronomy and Astrophysics, 51, 353, doi: 10.1146/annurev-astro-082812-140938

Chaplin, W. J., Kjeldsen, H., Christensen-Dalsgaard, J., et al. 2011, Science, 332, 213, doi: 10.1126/science.1201827

Chaplin, W. J., Kjeldsen, H., Bedding, T. R., et al. 2011, The Astrophysical Journal, 732, 54, doi: 10.1088/0004-637X/732/1/54

Chaplin, W. J., Basu, S., Huber, D., et al. 2014, ApJS, 210, 1, doi: 10.1088/0067-0049/210/1/1

Chaplin, W. J., Lund, M. N., Handberg, R., et al. 2015, PASP, 127, 1038, doi: 10.1086/683103

Drimmel, R., Cabrera-Lavers, A., & Lpez-Corredoira, M. 2003, Astronomy and Astrophysics, 409, 205, doi: 10.1051/0004-6361:20031070

Flower, P. J. 1996, The Astrophysical Journal, 469, 355, doi: 10.1086/177785

Gilliland, R. L., Brown, T. M., Christensen-Dalsgaard, J., et al. 2010, Publications of the Astronomical Society of the Pacific, 122, 131, doi: 10.1086/650399

Green, G. M., Schlafly, E. F., Finkbeiner, D. P., et al. 2015, The Astrophysical Journal, 810, 25, doi: 10.1088/0004-637X/810/1/25

Hekker, S., & Christensen-Dalsgaard, J. 2017, A&A Rv, 25, 1, doi: 10.1007/s00159-017-0101-x

Henden, A. A., Welch, D. L., Terrell, D., & Levine, S. E. 2009, in , 407.02. http://adsabs.harvard.edu/abs/2009AAS...21440702H

Hog, E., Fabricius, C., Makarov, V. V., et al. 2000, Astronomy and Astrophysics, 355, L27

Howell, S. B., Sobeck, C., Haas, M., et al. 2014, Publications of the Astronomical Society of the Pacific, 126, 398, doi: 10.1086/676406

Huang, Y., Liu, X.-W., Yuan, H.-B., et al. 2015, Monthly Notices of the Royal Astronomical Society, 454, 2863, doi: 10.1093/mnras/stv1991

Huber, D., Ireland, M. J., Bedding, T. R., et al. 2012, ApJ, 760, 32, doi: 10.1088/0004-637X/760/1/32

Huber, D., Chaplin, W. J., Christensen-Dalsgaard, J., et al. 2013, ApJ, 767, 127, doi: 10.1088/0004-637X/767/2/127

Kallinger, T., & Matthews, J. M. 2010, ApJL, 711, L35, doi: 10.1088/2041-8205/711/1/L35

Lindegren, L., Hernandez, J., Bombrun, A., et al. 2018, ArXiv e-prints, 1804, arXiv:1804.09366

Lund, M. N., Chaplin, W. J., Casagrande, L., et al. 2016a, PASP, 128, 124204, doi: 10.1088/1538-3873/128/970/124204

Lund, M. N., Basu, S., Silva Aguirre, V., et al. 2016b, MNRAS, 463, 2600, doi: 10.1093/mnras/stw2160

Lundkvist, M. S., Kjeldsen, H., Albrecht, S., et al. 2016, Nature Communications, 7, 11201, doi: 10.1038/ncomms11201

Luri, X., Brown, A. G. A., Sarro, L. M., et al. 2018, A&A, 616, A9, doi: 10.1051/0004-6361/201832964

Marshall, D. J., Robin, A. C., Reyl, C., Schultheis, M., & Picaud, S. 2006, Astronomy and Astrophysics, 453, 635, doi: 10.1051/0004-6361:20053842

Mozurkewich, D., Armstrong, J. T., Hindsley, R. B., et al. 2003, AJ, 126, 2502, doi: 10.1086/378596

Pinsonneault, M. H., An, D., Molenda-Żakowicz, J., et al. 2012, ApJS, 199, 30, doi: 10.1088/0067-0049/199/2/30

Rauer, H., Catala, C., Aerts, C., et al. 2014, Experimental Astronomy, 38, 249, doi: 10.1007/s10686-014-9383-4

Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003, doi: 10.1117/1.JATIS.1.1.014003

Silva Aguirre, V., Casagrande, L., Basu, S., et al. 2012, The Astrophysical Journal, 757, 99, doi: 10.1088/0004-637X/757/1/99

Soubiran, C., Campion, J.-F. L., Brouillet, N., & Chemin, L. 2016, arXiv:1605.07384 [astro-ph]

Stassun, K. G., Oelkers, R. J., Pepper, J., et al. 2018, AJ, 156, 102, doi: 10.3847/1538-3881/aad050

Sullivan, P. W., Winn, J. N., Berta-Thompson, Z. K., et al. 2015, The Astrophysical Journal, 809, 77, doi: 10.1088/0004-637X/809/1/77

Torres, G. 2010, The Astronomical Journal, 140, 1158, doi: 10.1088/0004-6256/140/5/1158

van Leeuwen, F. 2007, in . http://adsabs.harvard.edu/abs/2007ASSL..350.....V

White, T. R., Huber, D., Maestro, V., et al. 2013, MNRAS, 433, 1262, doi: 10.1093/mnras/stt802

White, T. R., Huber, D., Mann, A. W., et al. 2018, MNRAS, 477, 4403, doi: 10.1093/mnras/sty898

Yu, J., Huber, D., Bedding, T. R., et al. 2016, MNRAS, 463, 1297, doi: 10.1093/mnras/stw2074

Zinn, J. C., Pinsonneault, M. H., Huber, D., & Stello, D. 2018, ArXiv e-prints. https://arxiv.org/abs/1805.02650

Appendix B
The Asteroseismic Potential of TESS: Exoplanet-host stars

The work done in this Appendix was published in the Astrophysical Journal Volume 830 Article 138[a] (Campante et al., 2016). The work presented here is my own, but was written by Tiago Campante. As a result, I am the second author.

I collaborated with Bill Chaplin, Daniel Huber and Tiago Campante to produce these predictions. Figure 11 (the Mass-period diagram of known exoplanets) was produced by Daniel Huber.

---

[a] http://iopscience.iop.org/volume/0004-637X/830

The Astrophysical Journal, 830:138 (15pp), 2016 October 20 doi:10.3847/0004-637X/830/2/138

THE ASTEROSEISMIC POTENTIAL OF *TESS*: EXOPLANET-HOST STARS

T. L. Campante[1,2], M. Schofield[1,2], J. S. Kuszlewicz[1,2], L. Bouma[3], W. J. Chaplin[1,2], D. Huber[2,4,5],
J. Christensen-Dalsgaard[2], H. Kjeldsen[2], D. Bossini[1,2], T. S. H. North[1,2], T. Appourchaux[6], D. W. Latham[7],
J. Pepper[8], G. R. Ricker[3], K. G. Stassun[9,10], R. Vanderspek[3], and J. N. Winn[3,11]
[1] School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK; campante@bison.ph.bham.ac.uk
[2] Stellar Astrophysics Centre (SAC), Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark
[3] MIT Kavli Institute for Astrophysics and Space Research, 70 Vassar St., Cambridge, MA 02139, USA
[4] Sydney Institute for Astronomy, School of Physics, University of Sydney, Sydney, Australia
[5] SETI Institute, 189 Bernardo Avenue #100, Mountain View, CA 94043, USA
[6] Université Paris-Sud, Institut d'Astrophysique Spatiale, UMR 8617, CNRS, Bâtiment 121, F-91405 Orsay Cedex, France
[7] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[8] Department of Physics, Lehigh University, Bethlehem, PA 18015, USA
[9] Vanderbilt University, Physics & Astronomy Department, 1807 Station B, Nashville, TN 37235, USA
[10] Fisk University, Department of Physics, 1000 17th Ave. N, Nashville, TN 37208, USA
[11] Department of Physics, 77 Massachusetts Ave., Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 2016 April 15; revised 2016 August 2; accepted 2016 August 2; published 2016 October 18

ABSTRACT

New insights on stellar evolution and stellar interior physics are being made possible by asteroseismology. Throughout the course of the *Kepler* mission, asteroseismology has also played an important role in the characterization of exoplanet-host stars and their planetary systems. The upcoming NASA *Transiting Exoplanet Survey Satellite* (*TESS*) will be performing a near all-sky survey for planets that transit bright nearby stars. In addition, its excellent photometric precision, combined with its fine time sampling and long intervals of uninterrupted observations, will enable asteroseismology of solar-type and red-giant stars. Here we develop a simple test to estimate the detectability of solar-like oscillations in *TESS* photometry of any given star. Based on an all-sky stellar and planetary synthetic population, we go on to predict the asteroseismic yield of the *TESS* mission, placing emphasis on the yield of exoplanet-host stars for which we expect to detect solar-like oscillations. This is done for both the target stars (observed at a 2-minute cadence) and the full-frame-image stars (observed at a 30-minute cadence). A similar exercise is also conducted based on a compilation of known host stars. We predict that *TESS* will detect solar-like oscillations in a few dozen target hosts (mainly subgiant stars but also in a smaller number of F dwarfs), in up to 200 low-luminosity red-giant hosts, and in over 100 solar-type and red-giant known hosts, thereby leading to a threefold improvement in the asteroseismic yield of exoplanet-host stars when compared to *Kepler*'s.

Key words: asteroseismology – planets and satellites: detection – space vehicles: instruments – surveys – techniques: photometric

1. INTRODUCTION

Asteroseismology is proving to be particularly relevant for the study of solar-type and red-giant stars (for a review, see Chaplin & Miglio 2013, and references therein), in great part due to the exquisite photometric data made available by the French-led *COnvection ROtation and planetary Transits* satellite (*CoRoT*; Michel et al. 2008), NASA's *Kepler* space telescope (Borucki et al. 2010), and, more recently, by the repurposed *K2* mission (Howell et al. 2014). These stars exhibit solar-like oscillations, which are excited and intrinsically damped by turbulence in the outermost layers of a star's convective envelope. The information contained in solar-like oscillations allows fundamental stellar properties (e.g., mass, radius, and age) to be precisely determined, while also allowing the internal stellar structure to be constrained to unprecedented levels, provided that individual oscillation mode parameters are measured. As a result, asteroseismology of solar-like oscillations is quickly maturing into a powerful tool whose impact is being felt more widely across different domains of astrophysics.

A noticeable example is the synergy between asteroseismology and exoplanetary science. Asteroseismology has been playing an important role in the characterization of exoplanet-host stars and their planetary systems, in particular over the course of the *Kepler* mission (Huber et al. 2013b; Silva Aguirre et al. 2015; Davies et al. 2016). Transit observations—as carried out by *Kepler*—are an indirect detection method, and are consequently only capable of providing planetary properties relative to the properties of the host star. The precise characterization of the host star through asteroseismology thus allows for inferences on the absolute properties of its planetary companions (e.g., Carter et al. 2012; Howell et al. 2012; Barclay et al. 2013; Campante et al. 2015; Gettel et al. 2016). Moreover, information on the stellar inclination angle as provided by asteroseismology can lead to a better understanding of the planetary system dynamics and evolution (e.g., Chaplin et al. 2013; Huber et al. 2013a; Campante et al. 2016). Another domain of application is that of orbital eccentricity determination based on the observed transit timescales (Sliski & Kipping 2014; Van Eylen & Albrecht 2015). Finally, the potential use of asteroseismology in measuring the levels of near-surface magnetic activity and in probing stellar activity cycles may help constrain the location of habitable zones around Sun-like stars.

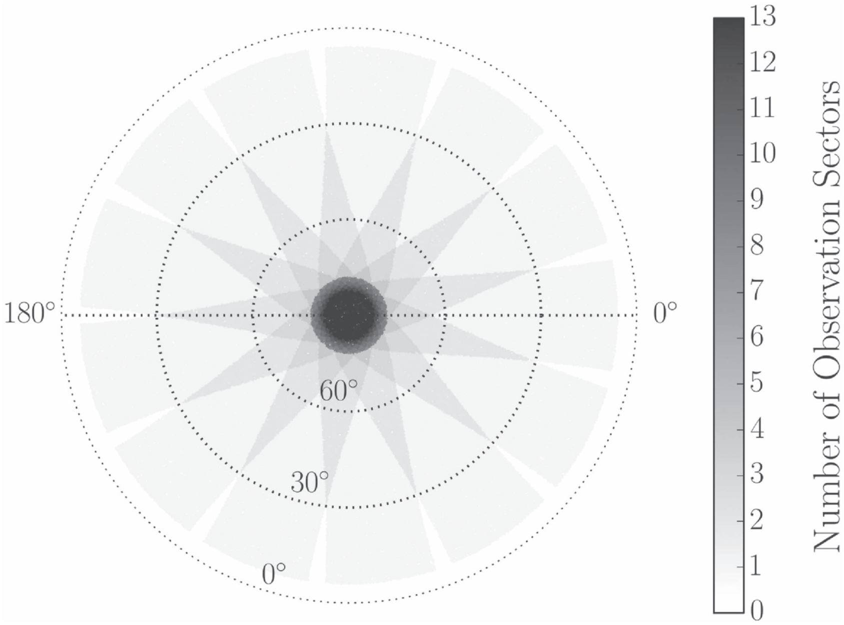


Figure 1. Polar projection illustrating *TESS*'s coverage of a single ecliptic hemisphere.

The *Transiting Exoplanet Survey Satellite*[12] (*TESS*; Ricker et al. 2015) is a NASA-sponsored Astrophysics Explorer mission that will perform a near all-sky survey for planets that transit bright nearby stars. Its launch is currently scheduled for 2017 December. During the primary mission duration of two years, *TESS* will monitor the brightness of several hundred thousand main-sequence, low-mass stars over intervals ranging from one month to one year, depending mainly on a star's ecliptic latitude. Monitoring of these pre-selected target stars will be made at a cadence of 2 minutes, while full-frame images (FFIs) will also be recorded every 30 minutes. Being 10–100 times brighter than *Kepler* targets and distributed over a solid angle that is nearly 300 times larger, *TESS* host stars will be well suited for follow-up spectroscopy. Sullivan et al. (2015; hereafter S15) predicted the properties of the transiting planets detectable by *TESS* and of their host stars. *TESS* is expected to detect approximately 1700 transiting planets from 2×10^5 pre-selected target stars. The majority of the detected planets will have their radii in the sub-Neptune regime (i.e., 2–4 $R_\oplus$). Analysis of the FFIs will lead to the additional detection of several thousand planets larger than 1.25 $R_\oplus$ orbiting stars that are not among the pre-selected targets.

Furthermore, *TESS*'s excellent photometric precision, combined with its fine time sampling and long intervals of uninterrupted observations, will enable asteroseismology of solar-type and red-giant stars, whose dominant oscillation periods range from several minutes to several hours. In this paper we aim at investigating the asteroseismic yield of the mission, placing emphasis on the yield of exoplanet-host stars for which we expect to detect solar-like oscillations. A broader study of the asteroseismic detections for stars that are not necessarily exoplanet hosts will be presented in a subsequent paper. The rest of the paper is organized as follows. A brief overview of *TESS* covering the mission design and survey operations is given in Section 2. Chaplin et al. (2011b) provides a simple recipe for estimating the detectability of solar-like oscillations in *Kepler* observations. In Section 3 we revisit that work and perform the necessary changes (plus a series of important updates) to make the recipe applicable to *TESS* photometry. Based on an existing all-sky stellar and planetary synthetic population, we then go on in Section 4 to predict the yield of *TESS* exoplanet-host stars with detectable solar-like oscillations. A similar exercise is conducted in Section 5, although now based on a compilation of known (i.e., confirmed) host stars. We summarize and discuss our results in Section 6.

2. OVERVIEW OF *TESS*

Four identical cameras will be employed by *TESS*, each consisting of a lens assembly and a detector assembly with four 2048 × 2048 charge-coupled devices (CCDs). Each of the four lenses has an entrance pupil diameter of 10.5 cm and forms a 24° × 24° image on the four-CCD mosaic in its focal plane, hence leading to a pixel scale of $21''.1$. The effective collecting area of each camera is 69 cm^2. The four camera fields are stacked vertically to create a combined field of view of 24° × 96° (or 2304 deg^2).

TESS will observe from a thermally stable, low-radiation High Earth Orbit. *TESS*'s elliptical orbit will have a nominal perigee of 17 $R_\oplus$ and a 13.7 day period in 2:1 resonance with the Moon's orbit. Over the course of the two-year duration of the primary mission, *TESS* will observe nearly the whole sky by dividing it into 26 observation sectors, 13 per ecliptic hemisphere. Each sector will be observed for 27.4 days (or two spacecraft orbits). Science operations will be interrupted at perigee for no more than 16 hr to allow for the downlink of the data, thus resulting in a high duty cycle of the observations. Figure 1 shows a polar projection illustrating the coverage of a single ecliptic hemisphere. The partially overlapping observation sectors are equally spaced in ecliptic longitude, extending from an ecliptic latitude of 6° to the ecliptic pole and beyond (the top camera is centered on the ecliptic pole). Successive sectors are positioned in order of increasing longitude (i.e.,

---

[12] http://tess.gsfc.nasa.gov/

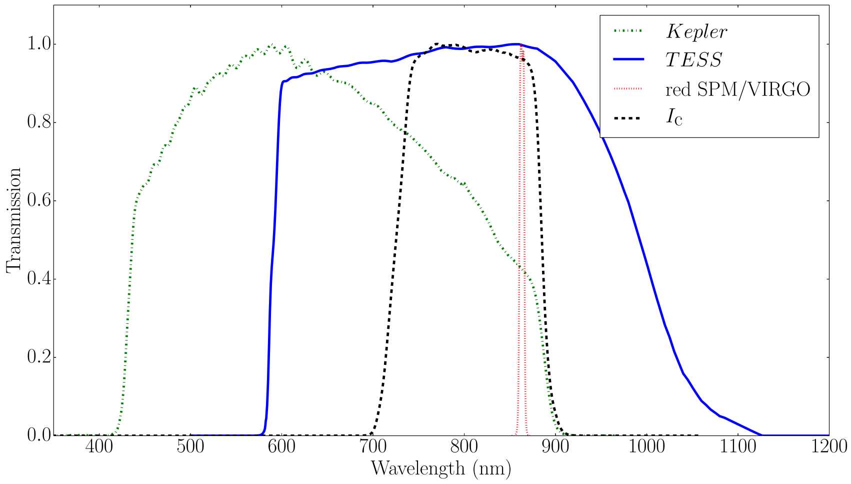


Figure 2. *TESS* spectral response function. Also shown for comparison are the spectral response functions of *Kepler* and of the red channel of the SPM/VIRGO instrument on board *SOHO*, as well as the Johnson–Cousins I_C filter curve. Each curve has been normalized to have a maximum value of unity.

eastwardly), with the first pointing[13] centered at 0° of longitude. Approximately 30,000 $\deg^2$ will be observed for at least 27.4 days. Moreover, observation sectors overlap near the ecliptic poles for increased sensitivity to smaller and longer-period planets in the *James Webb Space Telescope*'s (Beichman et al. 2014) continuous viewing zone.

The *TESS* spectral response function is shown in Figure 2. It is defined as the product of the long-pass filter transmission curve and the detector quantum efficiency curve. An enhanced sensitivity to red wavelengths is desirable, since cool red dwarfs will be preferentially targeted by *TESS* in the search for small transiting planets. The bandpass thus covers the range 600–1000 nm, being approximately centered on the Johnson–Cousins I_C band. The spectral response functions of *Kepler* and that of the red channel of the SPM/VIRGO instrument[14] (Fröhlich et al. 1995) on board the *Solar and Heliospheric Observatory* spacecraft are also shown in Figure 2.

New images will be acquired by each camera every 2 s. However, due to limitations in onboard data storage and telemetry, these 2 s images will be stacked (before being downlinked to Earth) to produce two primary data products with longer effective exposure times: (i) subarrays of pixels centered on several hundred thousand pre-selected target stars will be stacked at a 2 minute cadence, while (ii) FFIs will be stacked every 30 minutes. Up to 20,000 2 minute cadence slots (or the equivalent to ~10% of the pre-selected target stars) will be allocated to the *TESS* Asteroseismic Science Consortium (TASC) over the course of the mission. In addition, a number of slots (notionally 1500) with faster-than-standard sampling, i.e., 20 s, will be reserved for the investigation of asteroseismic targets of special interest (mainly compact pulsators and main-sequence, low-mass stars).

A catalog of pre-selected target stars ($\gtrsim 2 \times 10^5$) will be monitored by *TESS* at a cadence of 2 minutes. This catalog will ideally include main-sequence stars that are sufficiently bright to maximize the prospects for detecting the transits of small planets (i.e., $R_p < 4\,R_\oplus$). This leads to a limiting magnitude that will depend on spectral type, with $I_C \lesssim 12$ for FGK dwarfs and $I_C \lesssim 13$ for the smaller M dwarfs. In addition to the pre-selected targets, *TESS* will return FFIs with a cadence of 30 minutes, which will expand the search for transits to any sufficiently bright stars in the field of view that may have not been pre-selected. The longer integration time of the FFIs will, however, reduce the sensitivity to transits with a short duration. Over the course of the mission, the FFIs will be the source of precise photometry for approximately 20 million bright objects ($I_C < 14$–15).

3. PREDICTING THE DETECTABILITY OF SOLAR-LIKE OSCILLATIONS

Solar-like oscillations are predominantly acoustic standing waves (or p modes). The oscillation modes are characterized by the radial order n (related to the number of radial nodes), the spherical degree l (specifying the number of nodal surface lines), and the azimuthal order m (with $|m|$ specifying how many of the nodal surface lines cross the equator). Radial modes have $l = 0$, whereas non-radial modes have $l > 0$. Values of m range from $-l$ to l, meaning that there are $2l+1$ azimuthal components for a given multiplet of degree l. Observed oscillation modes are typically high-order modes of low spherical degree, with the associated power spectrum showing a pattern of peaks with near-regular frequency separations. The most prominent separation is the large frequency separation, $\Delta\nu$, between neighboring overtones with the same spherical degree. The large frequency separation essentially scales as $\langle\rho\rangle^{1/2}$, where $\langle\rho\rangle \propto M/R^3$ is the mean density of a star with mass M and radius R. Moreover, oscillation mode power is modulated by an envelope that generally assumes a bell-shaped appearance. The frequency at the peak of the power envelope is referred to as the frequency of maximum oscillation amplitude, $\nu_{\max}$. This frequency scales to very good approximation as $g\,T_{\rm eff}^{-1/2}$, where g is the surface gravity and $T_{\rm eff}$ is the effective temperature. The fact that $\nu_{\max}$ mainly depends on g makes it an indicator of the evolutionary state of a star.

---

[13] This is the convention used in this work and in S15. The actual pointing coordinates will depend on the spacecraft's launch date.

[14] The three-channel Sun photometer (SPM) enables Sun-as-a-star helioseismology.

3.1. Detection Test

In this work we adopt the test developed by Chaplin et al. (2011b) to estimate the detectability of solar-like oscillations in any given *Kepler* target, which looked for signatures of the bell-shaped power excess due to the oscillations (see also Campante et al. 2014). Below we revisit that work and detail the necessary changes (plus a series of important updates) to make the detection test applicable to *TESS* photometry.

Estimation of the detection probability, $p_{\rm detect}$. The detection test is based upon the ratio of total mean mode power due to p-mode oscillations, $P_{\rm tot}$, to the total background power across the frequency range occupied by the oscillations, $B_{\rm tot}$. This quantity provides a global measure of the signal-to-noise ratio (S/N), in the oscillation spectrum, i.e.,

$$({\rm S/N})_{\rm tot} = P_{\rm tot}/B_{\rm tot}. \tag{1}$$

A total of N independent frequency bins in the power spectrum enter the estimation of $P_{\rm tot}$ and $B_{\rm tot}$, and hence $({\rm S/N})_{\rm tot}$:

$$N = W\,T, \tag{2}$$

where

$$W = \begin{cases} 1.32\,\nu_{\rm max}^{0.88} & \text{if } \nu_{\rm max} \leqslant 100\ \mu{\rm Hz}, \\ \nu_{\rm max} & \text{if } \nu_{\rm max} > 100\ \mu{\rm Hz}. \end{cases} \tag{3}$$

Here T represents the length of the observations and is based on the maximum number of contiguous observation sectors for a given star. Moreover, we have assumed that the mode power is contained either within a range $\pm 0.66\,\nu_{\rm max}^{0.88}$ (Mosser et al. 2012) or $\pm\nu_{\rm max}/2$ (Stello et al. 2007; Mosser et al. 2010) around $\nu_{\rm max}$, with frequencies expressed in μHz. The width, W, of this range corresponds to twice the FWHM of the power envelope (where a Gaussian-shaped envelope in frequency has been assumed). Note that any asymmetries of the power envelope have been disregarded.

When binning over N bins, the statistics of the power spectrum of a pure noise signal is taken to be χ^2 with 2N degrees of freedom (Appourchaux 2004). We begin by testing the null (or H_0) hypothesis that we observe pure noise. After specifying a false-alarm probability (or p-value) of 5%, we numerically compute the detection threshold $({\rm S/N})_{\rm thresh}$:

$$p = \int_x^{\infty} \frac{\exp(-x')}{\Gamma(N)} x'^{(N-1)}\,dx', \tag{4}$$

where $x = 1 + ({\rm S/N})_{\rm thresh}$ and Γ is the gamma function. Finally, the probability, $p_{\rm detect}$, that $({\rm S/N})_{\rm tot}$ exceeds $({\rm S/N})_{\rm thresh}$ is once more given by Equation (4), but now setting $x = (1 + ({\rm S/N})_{\rm thresh})/(1 + ({\rm S/N})_{\rm tot})$. This last step can be thought of as testing the alternative (or H_1) hypothesis that we observe a signal embedded in noise. Throughout this work, we assume to be able to detect solar-like oscillations only in stars for which $p_{\rm detect} > 0.5$. Next, we in turn detail how $P_{\rm tot}$ and $B_{\rm tot}$ are predicted.

Estimation of the total mean mode power, $P_{\rm tot}$. The total mean mode power may be approximately predicted following

$$P_{\rm tot} \approx 0.5\,c\,A_{\rm max}^2\,\eta^2(\nu_{\rm max}) D^{-2}\,\frac{W}{\Delta\nu}\quad {\rm ppm}^2, \tag{5}$$

where $A_{\rm max}$ corresponds to the maximum oscillation amplitude of the radial ($l = 0$) modes. The factor c measures the effective number of p modes per order ($c = 2.94$) and was computed following Bedding et al. (1996) for a weighted wavelength of 797 nm representative of the *TESS* bandpass. We disregard the dependence of c on $T_{\rm eff}$, $\log g$, and the metallicity, which could amount to relative variations of a few percent (Ballot et al. 2011). The fraction in the above equation takes into account the contribution from all segments of width $\Delta\nu$ that fall in the range where mode power is present. On average, the power of the contributing segments will be ~0.5 times that of the central segment, thus explaining the extra 0.5 factor in Equation (5). The attenuation factor $\eta^2(\nu)$ takes into account the apodization of the oscillation signal due to the finite integration time. It is given by ${\rm sinc}^2\left[\pi/2\left(\frac{\nu}{\nu_{\rm Nyq}}\right)\right]$ for an integration duty cycle of 100%, where $\nu_{\rm Nyq}$ is the Nyquist frequency. Finally, a dilution (or wash-out) factor D is introduced, which is defined as the ratio of the total flux in the photometric aperture from neighboring stars and the target star to the flux from the target star. This factor will be available for the simulated host stars introduced in Section 4, being otherwise set to $D = 1$ (i.e., an isolated system).

The maximum oscillation amplitude, $A_{\rm max}$, is predicted based on

$$A_{\rm max} = (0.85)(2.5)\beta\left(\frac{R}{R_\odot}\right)^2\left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{0.5}\quad {\rm ppm}, \tag{6}$$

where

$$\beta = 1 - \exp\left(-\frac{T_{\rm red} - T_{\rm eff}}{1550}\right) \tag{7}$$

and

$$T_{\rm red} = (8907)(L/L_\odot)^{-0.093}\quad {\rm K}. \tag{8}$$

Here and throughout we use $T_{\rm eff,\odot} = 5777$ K. Equation (6) is based on the prediction that the rms oscillation amplitude, $A_{\rm rms}$, observed in photometry at a wavelength λ, scales as $A_{\rm rms} \propto (L/M)^s/(\lambda T_{\rm eff}^r)$ (Kjeldsen & Bedding 1995), with M subsequently eliminated using the scaling relation $M \propto T_{\rm eff}^{1.5}$ (cf. Chaplin et al. 2011b). Accordingly, amplitudes are predicted to increase with increasing luminosity along the main sequence and relatively large amplitudes are expected for red giants. The exponent s has been examined both theoretically and observationally, and found to lie in the range $0.7 < s < 1.5$ (e.g., Corsaro et al. 2013, and references therein). Here we adopt $s = 1$ (Chaplin et al. 2011b). The value of r is chosen to be $r = 2$ following a fit to observational data in Kjeldsen & Bedding (1995). The factor β is introduced to correct for the overestimation of oscillation amplitudes in the hottest solar-type stars, with the luminosity-dependent quantity $T_{\rm red}$ representing the temperature on the red edge of the radial-mode δ Scuti instability strip. The solar rms value $A_{\rm max,\odot}$, as it would be measured by *Kepler*, is $A_{\rm max,\odot} \sim 2.5$ ppm. However, the absolute calibration of the predicted oscillation and granulation amplitudes depends on the spectral response of the instrument. *TESS* has a redder response than *Kepler* (cf. Figure 2), meaning observed amplitudes will be lower in the *TESS* data. Starting from the estimated *TESS* response, we

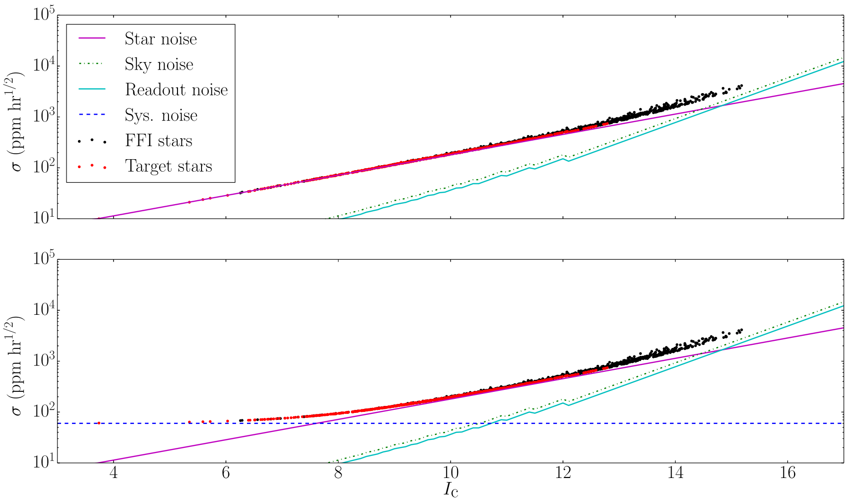


Figure 3. Photometric noise model. Predicted rms noise, σ, per a 1 hr integration as a function of the apparent magnitude I_C. The several noise components are represented by different line styles/colors. In the bottom panel a systematic noise level of $\sigma_{\rm sys} = 60$ ppm hr$^{1/2}$ is assumed, while this systematic error term is absent from the top panel. The overall rms noise levels for a synthetic population (see Section 4) of host stars are also displayed (target stars in red and FFI stars in black).

followed the procedures outlined in Ballot et al. (2011) to calculate a fractional multiplicative correction. We find that *TESS* oscillation (and granulation) amplitudes will be ~0.85 times those observed with *Kepler*.

Even though Equation (6) has been calibrated based on solar-type stars alone (Chaplin et al. 2011b), it is also used here to predict the maximum oscillation amplitudes of red-giant stars. As a sanity check, we compared the red-giant oscillation amplitudes as predicted by Equation (6) with those obtained using the similar models $\mathcal{M}_1$ and $\mathcal{M}_{1,\beta}$ of Corsaro et al. (2013), whose calibration was based on over 1000 *Kepler* long-cadence targets. Having run such a test for a sequence of red-giant-branch (solar-calibrated) stellar models along a 1 $M_\odot$ track, we obtained an rms relative difference of either 12% (model $\mathcal{M}_1$) or 7% (model $\mathcal{M}_{1,\beta}$).

When predicting $A_{\rm max}$, the effect of stellar activity should be considered. Evidence has been found that high levels of stellar activity, tied to the magnetic field and rotation period of the star, tend to suppress the amplitudes of oscillation modes (García et al. 2010; Chaplin et al. 2011a). In order to incorporate an appropriate correction to the predicted mode amplitudes, the stellar activity levels must first be predicted from the fundamental stellar properties. This has, however, proven to be difficult, for a variety of reasons. The initial difficulty lies in describing how stellar activity can be measured from photometric time series. Throughout the *Kepler* mission, several activity proxies have been used (e.g., Basri et al. 2011; Campante et al. 2014; Mathur et al. 2014; Gilliland et al. 2015) that show a high degree of correlation among them. However, predicting the absolute level of stellar activity remains a challenge. For instance, Gilliland et al. (2011) attempted to predict stellar activity levels in *Kepler* stars by first predicting the chromospheric emission activity index $R'_{\rm HK}$, before converting this to a photometric measure. The prediction of $R'_{\rm HK}$ requires knowledge of the rotation period of the star, which can in principle be predicted from gyrochronology for low-mass stars ($M < 1.3\,M_\odot$) if the age of the star is also known (Skumanich 1972; Aigrain et al. 2004). This is only applicable to main-sequence stars, since for more evolved stars the rotation period is no longer coupled to the stellar age in the same fashion. An additional problem with this procedure is that it in no way accounts for an activity cycle like the one observed in the Sun. Several challenges thus remain unsurmounted before stellar activity levels can be accounted for in the detection test and we ignore such a correction for the time being.

Estimation of the total background power, B_{tot}. The total background power is approximately given by

$$B_{\rm tot} \approx b_{\rm max}\, W \quad {\rm ppm}^2, \tag{9}$$

where $b_{\rm max}$ is the background power spectral density from instrumental/shot noise and granulation at $\nu_{\rm max}$:

$$b_{\rm max} = b_{\rm instr} + P_{\rm gran} \quad {\rm ppm}^2\,\mu{\rm Hz}^{-1}. \tag{10}$$

The power spectral density due to instrumental/shot noise is given by (e.g., Chaplin et al. 2008)

$$b_{\rm instr} = 2 \times 10^{-6}\,\sigma^2\,\Delta t \quad {\rm ppm}^2\,\mu{\rm Hz}^{-1}, \tag{11}$$

where Δt is the observational cadence. We use the photometric noise model for *TESS* presented in S15 to predict the rms noise, σ, per a given exposure time. This photometric noise model includes the photon-counting noise from the star (star noise), that from zodiacal light and background stars (sky noise), as well as the readout and systematic noise (instrumental noise). Figure 3 shows the contributions from the several noise components to the overall rms noise. The jagged appearance of the sky and readout noise components is due to the discretization of the number of pixels in the optimal photometric aperture. A systematic error term of $\sigma_{\rm sys} = 60$ ppm hr$^{1/2}$ is included in the bottom panel of Figure 3. This is an engineering requirement that is imposed on the design of the *TESS* photometer and not an estimate of the anticipated systematic noise level on 1 hr timescales. The systematic error term is assumed to scale with the total observing length as $T^{-1/2}$. It is perhaps unrealistic to assume that the systematic error will surpass 60 ppm for timescales shorter than one hour. Throughout this paper we will thus explore the implications of having $\sigma_{\rm sys} = 0$ ppm hr$^{1/2}$ (ideal

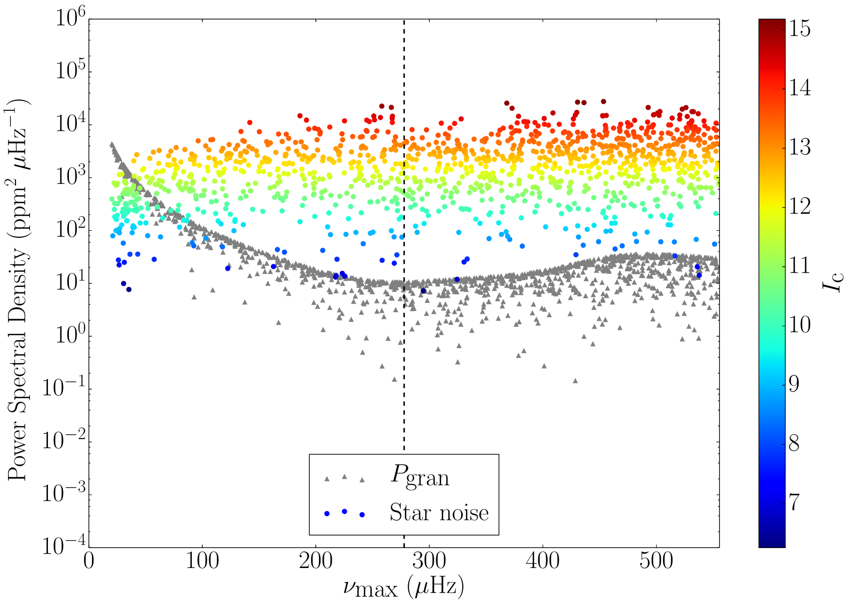


Figure 4. Contributions from granulation ($P_{\rm gran}$) and stellar shot noise to the background power spectral density of a synthetic population (see Section 4.2) of FFI host stars. The contribution from stellar shot noise is color-coded according to $I_{\rm C}$. The vertical dashed line represents $\nu_{\rm Nyq,FFI} \sim 278$ μHz.

case) and $\sigma_{\rm sys} = 60$ ppm hr$^{1/2}$ (regarded as a worst-case scenario).

Also shown in Figure 3 are the predicted rms noise levels for the simulated host stars of Section 4 (target stars in red and FFI stars in black). The observed scatter is a result of the minute dependence of the overall noise on $T_{\rm eff}$ and a star's celestial coordinates. It can be seen that, for the brightest stars, the photometric precision is limited by the systematic noise floor (when present). We note that the central pixels of a stellar image will saturate for stars with $I_{\rm C} \lesssim 7.5$ during the 2 s exposures, although high photometric precision is still expected down to $I_{\rm C} \approx 4$ or brighter. For most of the stars in Figure 3, whose magnitudes lie in the range $I_{\rm C} \approx 7$–15, the photometric precision is instead dominated by stellar shot noise.

To model the granulation power spectral density, we adopt model F (with no mass dependence) of Kallinger et al. (2014) and evaluate it at $\nu_{\max}$:

$$P_{\rm gran,real}(\nu_{\max}) = \eta^2(\nu_{\max}) D^{-2} \sum_{i=1}^{2} \frac{(2\sqrt{2}/\pi) a_i^2/b_i}{1 + (\nu_{\max}/b_i)^4} \times \text{ppm}^2\ \mu\text{Hz}^{-1}, \quad (12)$$

where the rms amplitude, $a_{1,2}$, and the characteristic frequencies, b1 and b_2, are given by

$$a_{1,2} = (0.85)(3382)\nu_{\max}^{-0.609} \quad \text{ppm}, \quad (13a)$$

$$b_1 = 0.317\ \nu_{\max}^{0.970} \quad \mu\text{Hz}, \quad (13b)$$

$$b_2 = 0.948\ \nu_{\max}^{0.992} \quad \mu\text{Hz}. \quad (13c)$$

This model was found by Kallinger et al. (2014) to be statistically preferred after a Bayesian model comparison that considered different approaches to quantifying the signature of stellar granulation. The model consists of two super-Lorentzian functions representing separate classes of physical processes such as stellar activity and/or different granulation scales. Model parameters have been calibrated via fits to the power spectra of a large set of *Kepler* targets, hence explaining the 0.85 multiplicative correction in Equation (13*a*) to convert to *TESS* granulation amplitudes.

When a continuous signal is being sampled that contains frequency components above the Nyquist frequency, $\nu_{\rm Nyq} \equiv 1/(2\Delta t)$, these will give rise to an effect known as aliasing and the signal is then said to be undersampled. The aliased granulation power at $\nu_{\max}$, $P_{\rm gran,aliased}(\nu_{\max})$, is given by[15]

$$P_{\rm gran,aliased}(\nu_{\max}) \equiv P_{\rm gran,real}(\nu'_{\max}), \quad (14)$$

with the folded frequency $\nu'_{\max}$ defined as

$$\nu'_{\max} = \begin{cases} \nu_{\rm Nyq} + (\nu_{\rm Nyq} - \nu_{\max}) & \text{if } \nu_{\max} \leqslant \nu_{\rm Nyq}, \\ \nu_{\rm Nyq} - (\nu_{\max} - \nu_{\rm Nyq}) & \text{if } \nu_{\rm Nyq} < \nu_{\max} \leqslant 2\nu_{\rm Nyq}, \end{cases} \quad (15)$$

where we restrict ourselves to the range [0, 2 $\nu_{\rm Nyq}$]. The total granulation power spectral density (at $\nu_{\max}$) is then given by

$$P_{\rm gran} = P_{\rm gran,real}(\nu_{\max}) + P_{\rm gran,aliased}(\nu_{\max}). \quad (16)$$

The formalism above allows us to correctly predict the detectability of solar-like oscillations both in stars with $\nu_{\max}$ in the sub- ($\nu_{\max} \leqslant \nu_{\rm Nyq}$) and super-Nyquist ($\nu_{\rm Nyq} < \nu_{\max} \leqslant 2\nu_{\rm Nyq}$) regimes. The latter regime is particularly relevant for stars in FFIs (cf. Chaplin et al. 2014b), for which $\nu_{\rm Nyq,FFI} \sim 278$ μHz, although not as much for target stars, since we do not expect to detect solar-like oscillations with $\nu_{\max}$ above $\nu_{\rm Nyq,target} \sim 4167$ μHz.

Figure 4 shows the contributions from granulation ($P_{\rm gran}$) and stellar shot noise to the background power spectral density (Equation (10)) of the simulated FFI host stars in Section 4.2. The observed scatter for $P_{\rm gran}$ is entirely due to the varying dilution factor, D. Stellar shot noise is seen to dominate over granulation across most of the plotted frequency range. This is

---

[15] Note that although $P_{\rm gran,real}(\nu'_{\max})$ is computed at $\nu'_{\max}$, the coefficients $a_{1,2}$ and $b_{1,2}$ are evaluated at $\nu_{\max}$.

in stark contrast to what was observed with *Kepler* photometry (e.g., Mathur et al. 2011; Karoff et al. 2013; Kallinger et al. 2014) and is mostly due to the smaller (by a factor of $\sim 10^2$) effective collecting area of the individual *TESS* cameras. While this will likely make robust modeling of the granulation profile a challenge, it does not necessarily mean that oscillations cannot be detected, as shown below.

Estimation of $\nu_{\max}$ *and* $\Delta\nu$. The values of $\nu_{\max}$ and $\Delta\nu$ used as input in the detection test are predicted from the stellar mass (when available; cf. Section 3.2), stellar radius, and effective temperature according to the scaling relations (e.g., Kallinger et al. 2010, and references therein):

$$\nu_{\max} = \nu_{\max,\odot}\left(\frac{M}{M_\odot}\right)\left(\frac{R}{R_\odot}\right)^{-2}\left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{-0.5} \quad (17)$$

and

$$\Delta\nu = \Delta\nu_\odot\left(\frac{M}{M_\odot}\right)^{0.5}\left(\frac{R}{R_\odot}\right)^{-1.5}, \quad (18)$$

with $\nu_{\max,\odot} = 3090\ \mu$Hz and $\Delta\nu_\odot = 135.1\ \mu$Hz. If no stellar mass is available (cf. Sections 4 and 5), we then eliminate M from Equations (17) and (18) using the relation (Stello et al. 2009a)

$$\Delta\nu \propto \nu_{\max}^{0.77}, \quad (19)$$

whose calibration was based on a cohort of stars with $\nu_{\max}$ in the range $15 \lesssim \nu_{\max} \lesssim 4500\ \mu$Hz. We note that the exponent in the previous equation varies slightly depending on the range in $\nu_{\max}$ being considered (Huber et al. 2011). However, for the purpose of this work, the use of a "unified" relation such as Equation (19) seems justified. The resulting scaling relations for $\nu_{\max}$ and $\Delta\nu$ in terms of the stellar radius and effective temperature are

$$\nu_{\max} = \nu_{\max,\odot}\left(\frac{R}{R_\odot}\right)^{-1.85}\left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{0.92} \quad (20)$$

and

$$\Delta\nu = \Delta\nu_\odot\left(\frac{R}{R_\odot}\right)^{-1.42}\left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{0.71}. \quad (21)$$

As a sanity check, we compared the output values from Equations (20) and (21) with those from Equations (17) and (18) across the full $\nu_{\max}$ and $\Delta\nu$ ranges. Based on a sequence of (solar-calibrated) stellar models along a 1 $M_\odot$ track, we obtained an rms relative difference of 3.9% for $\nu_{\max}$ and 1.8% for $\Delta\nu$, commensurate with typical fractional uncertainties measured by *Kepler* for these global parameters (e.g., Kallinger et al. 2010; Chaplin et al. 2014a).

3.2. Detectability of Solar-like Oscillations across the H-R Diagram

Figures 5–7 depict the detectability of solar-like oscillations with *TESS* across the Hertzsprung–Russell (H-R) diagram. We focus on that portion of the H-R diagram populated by solar-type and low-luminosity red-giant stars (i.e., up to the red-giant branch bump), bound at high effective temperatures by the red edge of the δ Scuti instability strip. The detection code was applied along several solar-calibrated stellar-model tracks spanning the mass range 0.8–2.0 $M_\odot$ (in steps of 0.2 $M_\odot$). These stellar models were computed using the Modules for Experiments in Stellar Astrophysics (MESA; Paxton et al. 2011, 2013) evolution code.

In Figure 5 we consider two different observing lengths (corresponding to 1 and 13 observation sectors) and a cadence of $\Delta t = 2$ minute. Further assuming a systematic noise level of $\sigma_{\rm sys} = 60$ ppm hr$^{1/2}$, detection of solar-like oscillations in main-sequence stars will not be possible for $T = 27$ day. Increasing the observing length to $T = 351$ day (relevant for stars near the ecliptic poles) may lead to the marginal detection of oscillations in (very bright) main-sequence stars more massive than the Sun. In both cases, detection of oscillations in subgiant and red-giant stars is nonetheless made possible, owing to their higher intrinsic amplitudes. As one would expect, this situation is significantly improved as the systematic noise level is brought down to $\sigma_{\rm sys} = 0$ ppm hr$^{1/2}$, with detections now being made possible for the brightest main-sequence stars over a range of masses. The longer 30-minute cadence is considered in Figures 6 and 7, where we have assumed a systematic noise level of $\sigma_{\rm sys} = 60$ ppm hr$^{1/2}$ only. FFIs will allow detecting oscillations in red-giant stars down to relatively faint magnitudes. Furthermore, it becomes apparent from Figure 7 that it should be possible to detect oscillations in the super-Nyquist regime for the brightest red giants.

4. ASTEROSEISMIC YIELD BASED ON SIMULATED DATA

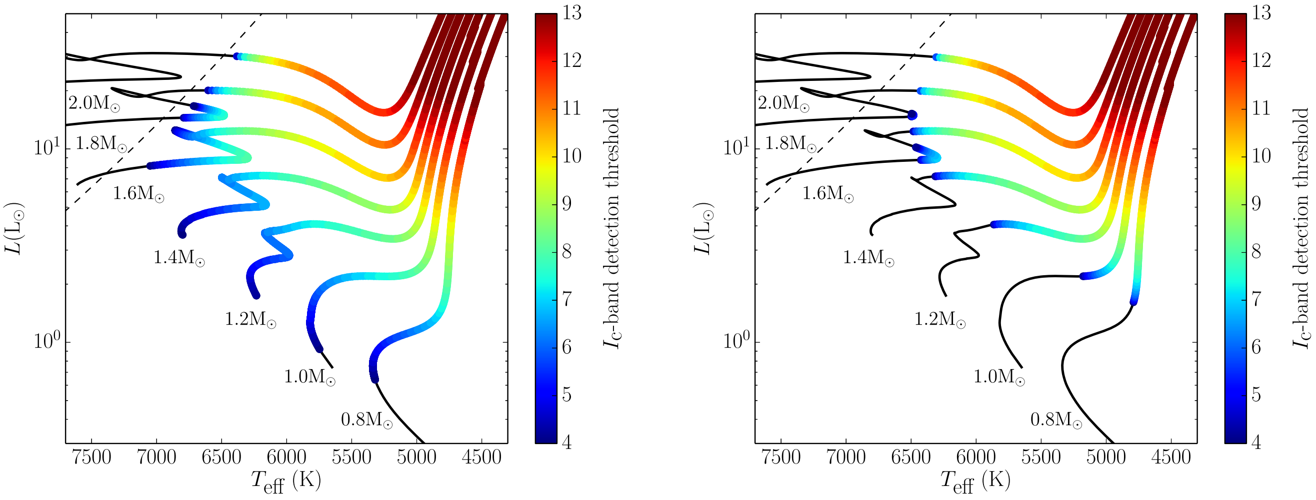
In S15 the authors predicted the properties of the transiting planets detectable by *TESS* and of their host stars, having done so for both the cohorts of target and FFI systems. Predictions were also made of the population of eclipsing binary stars that produce false-positive photometric signals. These predictions are based on a Monte Carlo simulation of a population of nearby stars generated using the TRIdimensional modeL of thE GALaxy (TRILEGAL; Girardi et al. 2005) population synthesis code. Any star in the above simulation that could in principle be searched for transiting planets is included in a so-called "bright catalog" (with 2MASS K_S magnitude $K_S < 15$) containing 1.58×10^8 stars. The 2×10^5 target stars are then selected from this catalog. The simulation employs planet occurrence rates derived from *Kepler* (Fressin et al. 2013; Dressing & Charbonneau 2015) whose completeness is high for the planetary periods and radii relevant to *TESS*, and a model for the photometric performance of the *TESS* cameras. In the present section, we apply the detection test to the synthetic population of host stars obtained in this way in order to predict the yield of *TESS* hosts with detectable solar-like oscillations.

4.1. TESS Target Hosts

The procedure by which target stars are selected in the simulation aims at maximizing the prospects for detecting the transits of small planets, and hence is mainly driven by stellar radius and apparent magnitude. In practice, this is done[16] by determining whether a fiducial planet with an orbital period of 20 days could be detected by *TESS* transiting a given star. This results in a target star catalog that is approximately complete

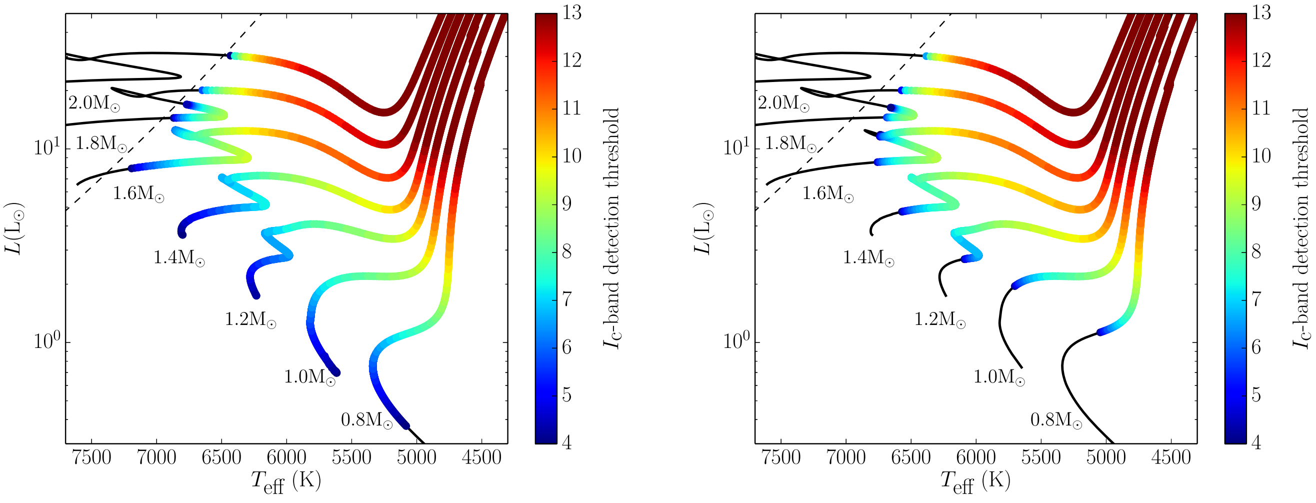
---

[16] The actual target selection procedure differs slightly from the one adopted in the simulation: stars will be selected for which a 2.25-$R_\oplus$ planet can be detected in a single 4 hr transit at the 5σ level.



(a) $T = 27$ d, $\sigma_{\rm sys} = 0$ ppm hr$^{1/2}$.

(b) $T = 27$ d, $\sigma_{\rm sys} = 60$ ppm hr$^{1/2}$.



(c) $T = 351$ d, $\sigma_{\rm sys} = 0$ ppm hr$^{1/2}$.

(d) $T = 351$ d, $\sigma_{\rm sys} = 60$ ppm hr$^{1/2}$.

Figure 5. Detectability of solar-like oscillations with *TESS* across the H-R diagram for a cadence of $\Delta t = 2$ minutes. Solar-calibrated evolutionary tracks spanning the mass range 0.8–2.0 $M_\odot$ (in steps of 0.2 $M_\odot$) are displayed. $I_{\rm C}$-band detection thresholds are color-coded (no detection is possible along those portions of the tracks shown as a thin black line). Modeled stars were assumed to be isolated (i.e., $D = 1$). The slanted dashed line represents the red edge of the δ Scuti instability strip. The several panels consider different combinations of the length of the observations (T) and systematic noise level ($\sigma_{\rm sys}$), as indicated.

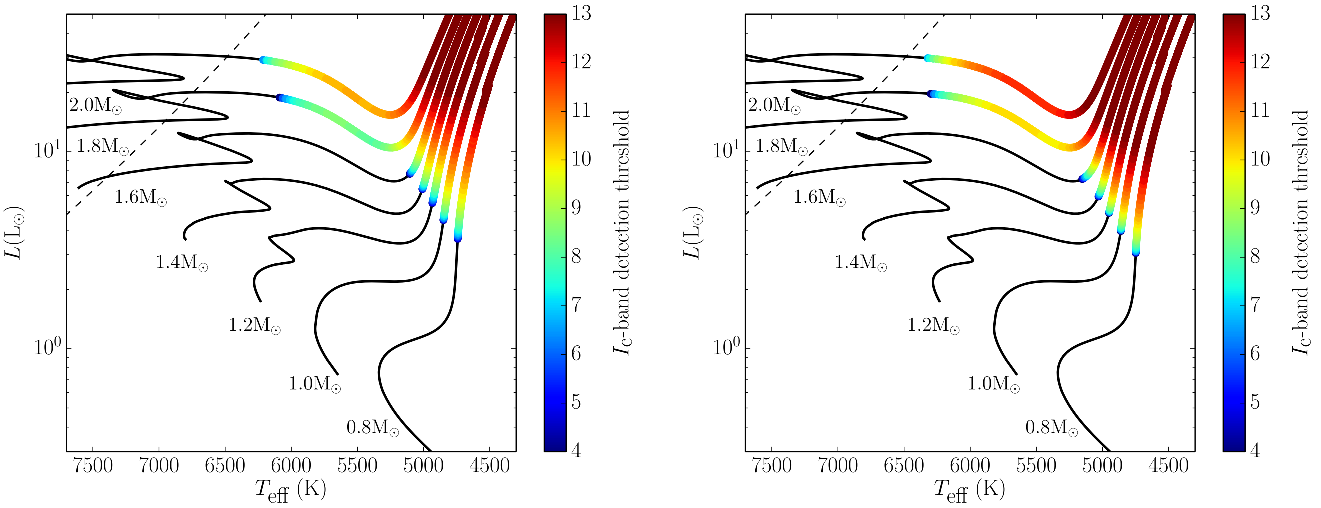
for short-period planets smaller than 2.25 $R_\oplus$. From a stellar perspective, this also means that nearly all bright main-sequence stars with $T_{\rm eff} < 6000$ K are selected, while a decreasing fraction of hotter stars make it into the target star catalog. In effect, a limiting apparent magnitude $I_{\rm C} \lesssim 12$ is imposed for FGK dwarfs (cf. Figure 17 of S15). Given this limiting apparent magnitude, virtually all main-sequence stars for which the detection of solar-like oscillations will be possible should already be included in the target star catalog (see Figure 5).

Furthermore, according to Figure 16 of S15, a non-negligible[17] number of subgiants end up being selected as target stars, even though they are far from optimal for transiting planet detection. Once *Gaia* (Perryman et al. 2001) parallaxes become available, we expect to have excellent knowledge of target stellar radii and that information could then be used to screen out, or else to deliberately target, subgiants. Here we advocate for the latter. As can be seen from Figure 5, bright subgiants are attractive targets for the 2-minute cadence slots reserved for asteroseismology. In what follows, we assess the overall asteroseismic potential of subgiant stars and the resulting impact on the asteroseismic yield of target hosts.

Having access to the all-sky bright catalog from where target stars have been selected, we made use of the known stellar properties to isolate all subgiant stars that fall into *TESS*'s field

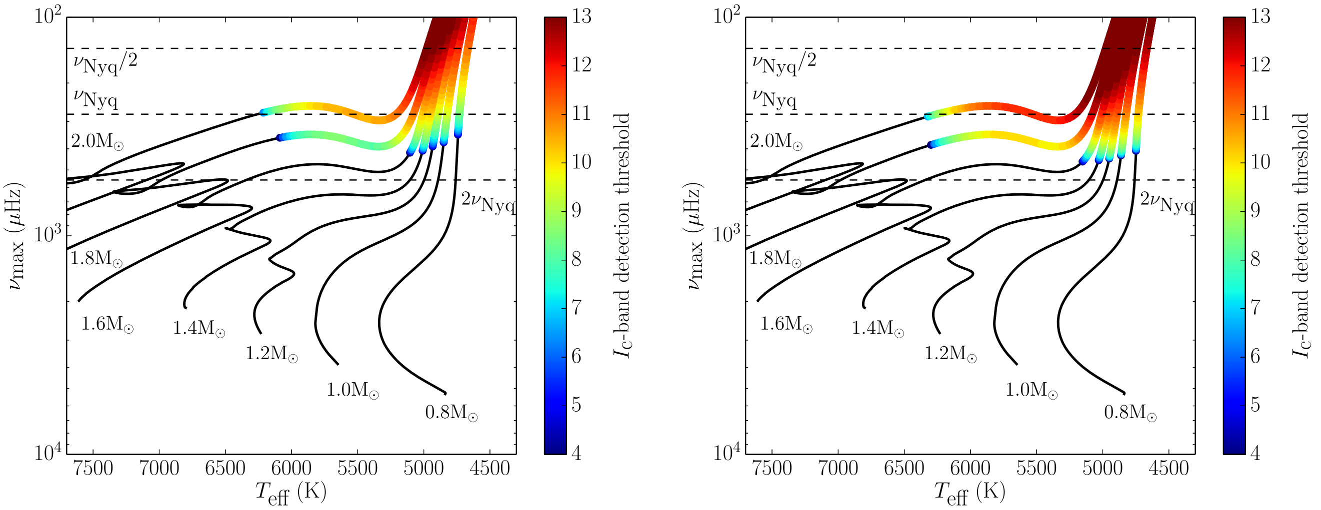
---

[17] Using flicker measurements of 289 bright *Kepler* candidate exoplanet-host stars with 4500 K < $T_{\rm eff}$ < 6650 K, Bastien et al. (2014) found that a Malmquist bias is responsible for a contamination of the sample by evolved stars, being that nearly 50% of those stars are in fact subgiants.



(a) $T = 27$ d, $\sigma_{\mathrm{sys}} = 60$ ppm hr$^{1/2}$. (b) $T = 351$ d, $\sigma_{\mathrm{sys}} = 60$ ppm hr$^{1/2}$.

Figure 6. Detectability of solar-like oscillations with *TESS* across the H-R diagram for a cadence of $\Delta t = 30$ minute. Solar-calibrated evolutionary tracks spanning the mass range 0.8–2.0 $M_\odot$ (in steps of 0.2 $M_\odot$) are displayed. I_{C}-band detection thresholds are color-coded (no detection is possible along those portions of the tracks shown as a thin black line). Modeled stars were assumed to be isolated (i.e., $D = 1$). The slanted dashed line represents the red edge of the δ Scuti instability strip. The two panels consider different lengths of the observations (T) and a systematic noise level of $\sigma_{\mathrm{sys}} = 60$ ppm hr$^{1/2}$, as indicated.



(a) $T = 27$ d, $\sigma_{\mathrm{sys}} = 60$ ppm hr$^{1/2}$. (b) $T = 351$ d, $\sigma_{\mathrm{sys}} = 60$ ppm hr$^{1/2}$.

Figure 7. Detectability of solar-like oscillations with *TESS* across an asteroseismic H-R diagram for a cadence of $\Delta t = 30$ minute. Note that $\nu_{\max}$ is now plotted along the vertical axis and not luminosity. Horizontal dashed lines indicate $\nu_{\mathrm{Nyq}}/2$, ν_{Nyq}, and $2\nu_{\mathrm{Nyq}}$. Solar-calibrated evolutionary tracks spanning the mass range 0.8–2.0 $M_\odot$ (in steps of 0.2 $M_\odot$) are displayed. I_{C}-band detection thresholds are color-coded (no detection is possible along those portions of the tracks shown as a thin black line). Modeled stars were assumed to be isolated (i.e., $D = 1$). The two panels consider different lengths of the observations (T) and a systematic noise level of $\sigma_{\mathrm{sys}} = 60$ ppm hr$^{1/2}$, as indicated.

of view. These stars were then ranked in order of decreasing brightness and the detection test was run assuming they would be observed at the 2-minute cadence. Simply ranking stars by brightness does not necessarily constitute the optimal procedure for selecting potential asteroseismic targets, as there is also a dependence of the detectability of solar-like oscillations on stellar mass and effective temperature along the subgiant branch (cf. Figure 5), not to mention the effect of the length of the observations. This simple approach is nonetheless suitable for arguing our point and also allows setting an upper bound on the number of pixels required to accommodate these potential asteroseismic subgiants.

Figure 8 summarizes the outcome of this exercise. The horizontal axes in the top panels of Figure 8 represent the total number of selected subgiants (after being ranked in order of decreasing brightness), with the vertical axes representing the relative (top left) and absolute (top right) yield of asteroseismic subgiants. The bottom left panel provides an alternative perspective, by plotting the cumulative yield of asteroseismic subgiants as a function of limiting apparent magnitude. The cumulative number of pixels in the target masks is shown in the bottom right panel. If, for instance, we were to select the brightest 1×10^4 (5×10^3) subgiants in *TESS*'s field of view, one would be able to detect solar-like oscillations in ~43%

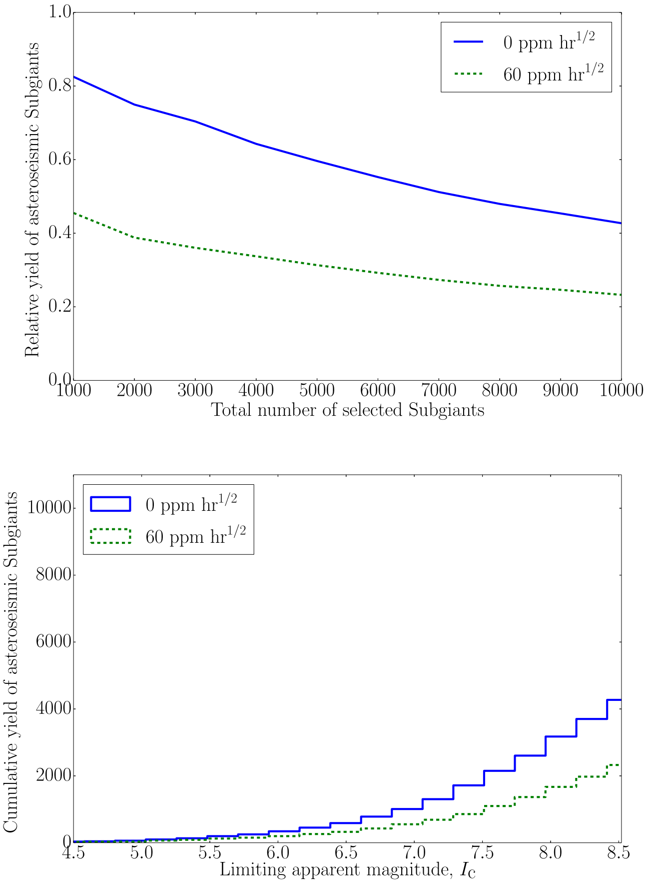


Figure 8. Asteroseismic potential of subgiant stars. Top panels: relative (left) and absolute (right) asteroseismic yield as a function of the total number of subgiants selected as target stars (ranked in order of decreasing brightness). Bottom left panel: cumulative yield of asteroseismic subgiants as a function of limiting apparent magnitude. Bottom right panel: cumulative pixel cost as a function of the total number of subgiants selected as target stars (ranked in order of decreasing brightness). Systems were assumed to be isolated (i.e., $D = 1$). A systematic noise level of either $\sigma_{\rm sys} = 0$ ppm hr$^{1/2}$ or $\sigma_{\rm sys} = 60$ ppm hr$^{1/2}$ was considered, as indicated.

(∼60%) of those stars assuming a systematic noise level of $\sigma_{\rm sys} = 0$ ppm hr$^{1/2}$. Furthermore, this would equate to a cumulative pixel cost of $\sim 1.2 \times 10^6$ ($\sim 6.1 \times 10^5$) pixels over the course of the mission or $\sim 1.1 \times 10^4$ ($\sim 5.8 \times 10^3$) pixels on average per camera for any given observation sector. For reference, due to onboard storage and bandwidth limitations, an allocation of ∼1.4 megapixels per camera for all types of 2 minute cadence targets has been set as the design goal.

Let us then assume that we select the brightest 1×10^4 subgiants in *TESS*'s field of view and observe them at the 2 minute cadence. What impact could this potentially have on the asteroseismic yield of target hosts? Doing this corresponds to setting a limiting apparent magnitude $I_C \sim 8.5$ (cf. bottom left panel of Figure 8). We now apply this magnitude cut to the synthetic population of subgiant hosts in FFIs and run the detection test.[18] Figure 9 shows the asteroseismic yield of exoplanet-host target stars for a single representative trial. This is dominated by subgiant stars. We assume Poisson statistics in estimating the statistical uncertainties and obtain 24 ± 5 or 14 ± 4 host stars depending on whether $\sigma_{\rm sys} = 0$ ppm hr$^{1/2}$ or $\sigma_{\rm sys} = 60$ ppm hr$^{1/2}$ (to be compared to 16 ± 4 or 8 ± 3 before inclusion of the brightest subgiants). For intermediate values of $\sigma_{\rm sys}$, the yield can be simply estimated by linear interpolation.

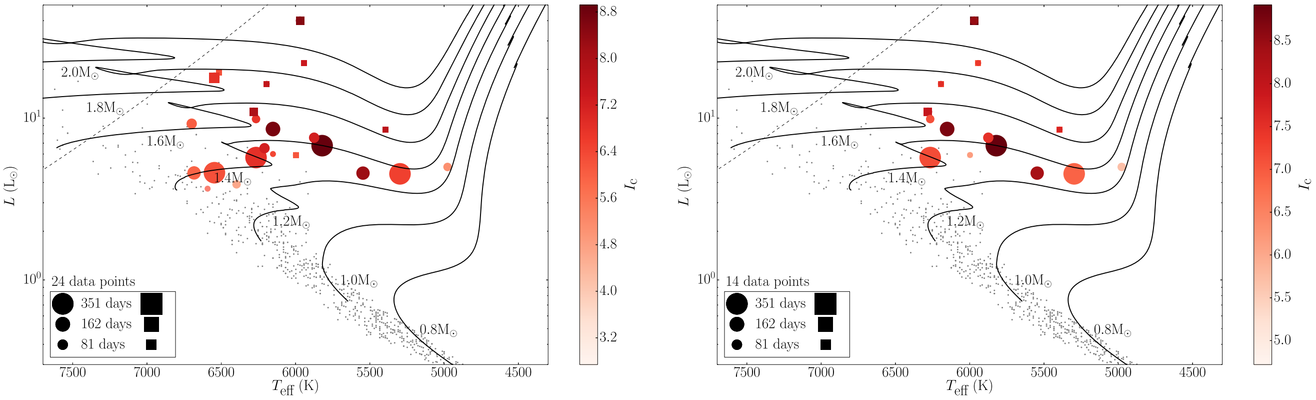
We note that this yield may be affected by biases in the planet occurrence rates upon which the simulation is based. S15 point out that such biases may be as high as ∼40% across all planetary sizes and periods. We further note that the adopted occurrence rates do not account for the expected effects of post-main-sequence evolution on the occurrence of planets migrating into close-in orbits (e.g., Frewen & Hansen 2016).

4.2. *TESS* FFI Hosts

Figure 10 shows the asteroseismic yield of exoplanet-host FFI stars for a single representative trial. The vast majority of host stars depicted are low-luminosity red giants. Assuming Poisson statistics, we obtain 191 ± 14 or 188 ± 14 host stars depending on whether $\sigma_{\rm sys} = 0$ ppm hr$^{1/2}$ or $\sigma_{\rm sys} = 60$ ppm hr$^{1/2}$. We note that the adopted occurrence

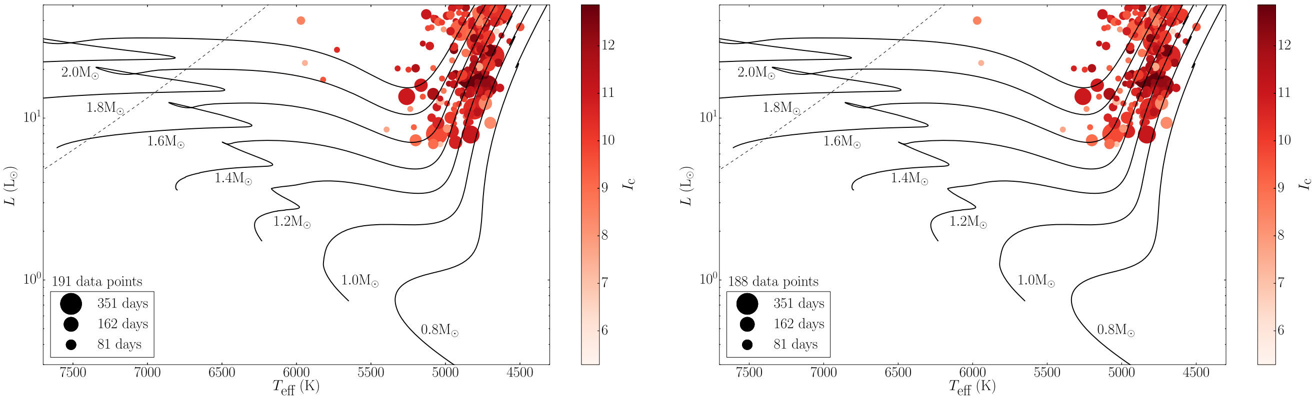
---

[18] The procedure described will in principle only provide a lower bound on the asteroseismic yield of subgiant hosts. The planet yield for FFI stars is estimated based on a 30-minute cadence, which can smear out short-duration and/or high-impact-parameter transits. Were we to observe the brightest 1×10^4 subgiants in *TESS*'s field of view at a 2-minute cadence, planets that would otherwise remain undetectable using the 30-minute cadence could now in principle be detected. We tested this by seeding these bright subgiants with planets, after which we simulated *TESS* observations at the 2- and 30-minute cadences. The resulting lack of difference between the two planet yields (i.e., obtained for either cadence) can be understood in terms of the long transit durations of planets about large stars (with a mean duration of 18 hr for the detected planets in this exercise), so that switching from a 30- to a 2-minute cadence does not lead to a significant improvement in the planet yield.



(a) $\sigma_{\rm sys} = 0\ {\rm ppm\,hr}^{1/2}$. (b) $\sigma_{\rm sys} = 60\ {\rm ppm\,hr}^{1/2}$.

Figure 9. Asteroseismic yield of exoplanet-host stars (target stars). The yield is computed for a single trial. Data points are color-coded according to apparent magnitude and their size is proportional to the observing length. Squares correspond to those extra stars with asteroseismic detections once the brightest subgiants have been included during target selection. Gray dots represent the underlying synthetic population of host stars from S15. Solar-calibrated evolutionary tracks spanning the mass range 0.8–2.0 $M_\odot$ (in steps of 0.2 $M_\odot$) are shown as continuous lines. The slanted dashed line represents the red edge of the δ Scuti instability strip. A systematic noise level of either $\sigma_{\rm sys} = 0$ ppm hr$^{1/2}$ or $\sigma_{\rm sys} = 60$ ppm hr$^{1/2}$ was considered, as indicated.



(a) $\sigma_{\rm sys} = 0\ {\rm ppm\,hr}^{1/2}$. (b) $\sigma_{\rm sys} = 60\ {\rm ppm\,hr}^{1/2}$.

Figure 10. Asteroseismic yield of exoplanet-host stars (FFI stars). The yield is computed for a single trial. Data points are color-coded according to apparent magnitude and their size is proportional to the observing length. Solar-calibrated evolutionary tracks spanning the mass range 0.8–2.0 $M_\odot$ (in steps of 0.2 $M_\odot$) are shown as continuous lines. The slanted dashed line represents the red edge of the δ Scuti instability strip. A systematic noise level of either $\sigma_{\rm sys} = 0$ ppm hr$^{1/2}$ or $\sigma_{\rm sys} = 60$ ppm hr$^{1/2}$ was considered, as indicated.

rates (from Fressin et al. 2013, for $T_{\rm eff} > 4000$ K) do not account for physical and orbital changes of planets as their parent stars evolve off the main sequence. Such evolutionary effects might be substantial, as there seem to be fewer close-in giant planets around evolved stars than main-sequence stars (e.g., Bowler et al. 2010; Johnson et al. 2010). An investigation of evolutionary effects on planet occurrence rates, and hence on *TESS* planet yields, is beyond the scope of this work. Since in S15 at least two transits need to be observed for a planet to be flagged as detectable, the yield shown in Figure 10 does not take into account single-transit events associated with long-period planets, which can be followed up with radial-velocity (RV) observations in order to characterize the planet (e.g., Yee & Gaudi 2008). Given the large expected yield of red-giant stars with detectable solar-like oscillations, it is likely that there will be a significant number of such single-transit events around asteroseismic hosts.

Figure 11 shows a mass–period diagram of known exoplanets orbiting red-giant-branch stars (adapted from Huber 2015). Despite the dearth of close-in giant planets (with $M_{\rm p} \gtrsim 0.5\ M_{\rm J}$) unveiled by RV surveys (e.g., Johnson et al. 2007), data from *Kepler* have led to the discovery of several giant planets with short orbital periods ($P_{\rm orb} \lesssim 50$ day) orbiting asteroseismic red-giant branch stars (four planets in three systems, to be precise; Huber et al. 2013a; Lillo-Box et al. 2014; Ciceri et al. 2015; Quinn et al. 2015). The latter may be hinting at the existence of a population of warm sub-Jovian planets around evolved stars that has remained elusive to RV surveys. The shaded area in Figure 11 approximately corresponds to the parameter space that will be probed by *TESS*, which will be mainly sensitive to planets with orbital periods[19] $P_{\rm orb} \lesssim 20$ day. Such parameter space is inaccessible to RV surveys at the low planetary-mass range.

---

[19] A fiducial planet with an orbital period of 13 days in a circular orbit around a low-luminosity red giant will have $a/R \sim 5$, where a is the semimajor axis, hence well above the Roche limit.

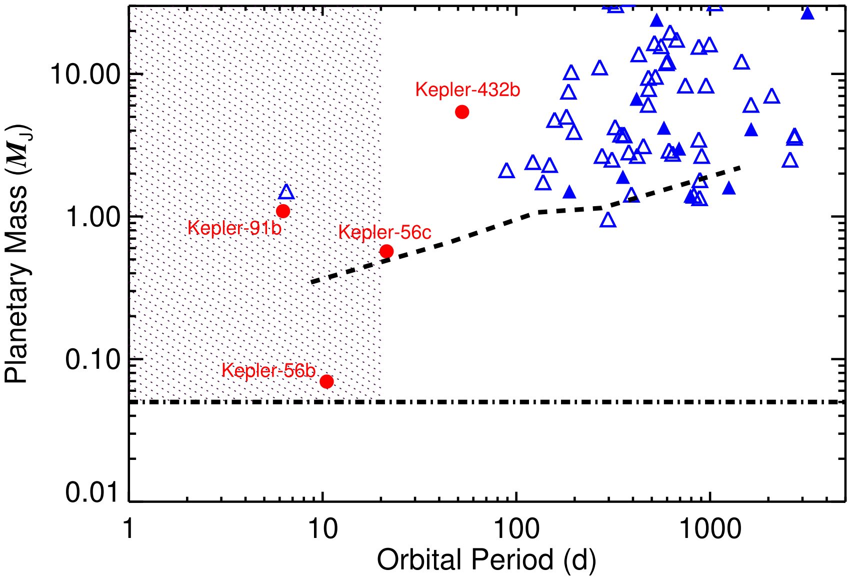


Figure 11. Mass–period diagram of known exoplanets orbiting red-giant-branch stars. Planets detected by the transit method are depicted as red circles and those detected in RV surveys as blue triangles (open triangles correspond to mean planetary masses assuming random orbital orientations). The dashed line represents the median RV detection threshold for mean masses from Bowler et al. (2010). The dash-dotted line marks the mass of Neptune and represents an approximate *TESS* detection limit. The shaded area approximately corresponds to the parameter space that will be probed by *TESS*.

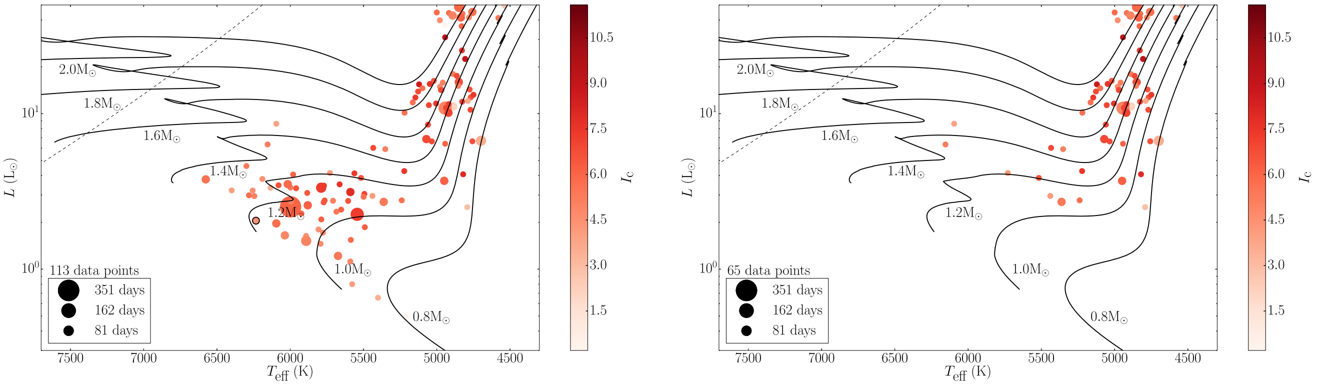
5. ASTEROSEISMIC YIELD OF CONFIRMED EXOPLANET-HOST STARS

We are now interested in assessing *TESS*'s asteroseismic yield of known (i.e., confirmed) exoplanet-host stars, assuming these will all be selected as target stars. We used the NASA Exoplanet Archive[20] (Akeson et al. 2013) to identify all known host stars (1182 at the time of writing after discarding the few known circumbinary planetary systems). The minimum amount of information on a given star that must be available in order to compute the probability of detecting solar-like oscillations comprises its celestial coordinates, I_C-band magnitude, $T_{\rm eff}$, and R (we henceforth enforce the simplifying assumption that stars are isolated, i.e., $D = 1$). While celestial coordinates are readily available for all known hosts, the same is not true for the remaining three quantities, and we will often need to derive them based on ancillary stellar properties. We started by grouping the known host stars according to the set of available properties, as follows:

1. Stars with an entry in the *Hipparcos* catalog. For the known hosts with an entry in the Extended *Hipparcos* Compilation (XHIP; Anderson & Francis 2012), I_C-band magnitudes are readily available. Whenever available in the Exoplanet Archive, $T_{\rm eff}$ and/or R values were used. When not available, these then had to be derived. The effective temperature was calculated using the $(B - V)$–$T_{\rm eff}$ relation from Torres (2010), which uses the $B - V$ color index as input. In order to compute the stellar radius, the stellar luminosity was first calculated via the *Hipparcos* parallax, π, using (Pijpers 2003)
$$\log(L/L_\odot) = 4.0 + 0.4M_{\rm bol,\odot} - 2.0\log\pi\,[\rm mas] - 0.4(V - A_V + {\rm BC}_V), \qquad (22)$$
where we have adopted $M_{\rm bol,\odot} = 4.73$ mag (Torres 2010) for the bolometric magnitude of the Sun, V is the apparent visual magnitude (available in XHIP), A_V is the extinction (assumed negligible), and ${\rm BC}_V$ are the bolometric corrections from the Flower (1996) polynomials presented in Torres (2010), which use $T_{\rm eff}$ as input. Stellar radii were then computed by rearranging the Stefan–Boltzmann law. Only stars with fractional parallax errors smaller than 25% were retained. A total of 385 stars fell under this group.
2. I_C-band magnitude, $T_{\rm eff}$, and R directly available from the Exoplanet Archive. These were used in the case of 33 host stars.
3. No available I_C-band magnitude. For the numerous *Kepler* and *K2* hosts, estimates of $T_{\rm eff}$ and R are generally available in the Exoplanet Archive, but an estimate of I_C is usually not. In such cases, we start by computing the Johnson–Cousins $R - I_C$ color index from 2MASS JHK_S colors on the main sequence (Bilir et al. 2008):
$$R - I_C = 0.954(J - H) + 0.593(H - K_S) + 0.025. \qquad (23)$$
The previous equation is then used in combination with the Johnson–Cousins $UBVRI_C$ to SDSS *ugriz* transformations from Jordi et al. (2006), to give I_C in terms of 2MASS JHK_S and SDSS r photometry, i.e.,
$$I_C = r - 1.239(R - I_C) - 0.104. \qquad (24)$$
This enabled us to gather all input quantities needed to run the detection test for 362 *Kepler* and *K2* hosts. Alternatively, for other families of hosts the I_C-band magnitude could be estimated based on the statistical color–color relation of Caldwell et al. (1993) provided $B - V$ and V are available (with separate sets of coefficients tabulated according to luminosity class). Further requiring that R is available (since $T_{\rm eff}$ could always be estimated from the $B - V$ color index), this ended up providing all input quantities for an additional 182 hosts. We note that for 133 of these stars we had to

---

[20] http://exoplanetarchive.ipac.caltech.edu/



(a) $\sigma_{\rm sys} = 0$ ppm hr$^{1/2}$.

(b) $\sigma_{\rm sys} = 60$ ppm hr$^{1/2}$.

Figure 12. Asteroseismic yield of known exoplanet-host stars for a cadence of $\Delta t = 2$ minutes. Data points are color-coded according to apparent magnitude and their size is proportional to the observing length. All but one system had their first planet (i.e., with suffix "b") detected using RV measurements. The only non-RV host had its first planet detected through direct imaging (represented by a circle with a black rim in the left panel). Solar-calibrated evolutionary tracks spanning the mass range 0.8–2.0 $M_\odot$ (in steps of 0.2 $M_\odot$) are shown as continuous lines. The slanted dashed line represents the red edge of the δ Scuti instability strip. A systematic noise level of either $\sigma_{\rm sys} = 0$ ppm hr$^{1/2}$ or $\sigma_{\rm sys} = 60$ ppm hr$^{1/2}$ was considered, as indicated.

rely on the properties available through the Exoplanet Orbit Database[21] (Han et al. 2014).

Stars that do not fall into one of the groups above were discarded. There are 962 known hosts for which all the relevant input quantities are available. Of these, 832 occupy that portion of the H-R diagram populated by solar-type and (low-luminosity) red-giant stars, and for which we ran the detection test. Figure 12 shows the asteroseismic yield of known exoplanet-host stars assuming either $\sigma_{\rm sys} = 0$ ppm hr$^{1/2}$ or $\sigma_{\rm sys} = 60$ ppm hr$^{1/2}$. For intermediate values of $\sigma_{\rm sys}$, the yield can again be estimated by linear interpolation. By considering the faster-than-standard 20 s cadence, we may still expect to detect solar-like oscillations in a few extra high-$\nu_{\rm max}$ hosts. Allocation of these slots will only be relevant for stars with $\nu_{\rm max}$ larger than $\nu_{\rm Nyq,target}/2 \sim 2084\ \mu$Hz, for which the attenuation factor, $\eta^2(\nu_{\rm max})$, exceeds ~20%.

We remind the reader that the actual pointing coordinates will depend on the spacecraft's launch date. The yield, however, remains virtually unchanged if we were to adopt different pointing coordinates. Furthermore, we notice how the asteroseismic yield of known exoplanet-host stars is an order of magnitude greater than that of target hosts (cf. left panels of Figures 9 and 12). This is simply the result of a selection effect. First, *TESS* target stars are preferentially bright main-sequence stars with spectral types F5 and later, thus maximizing the prospects for detecting the transits of small planets. second, *TESS* target hosts are restricted to transiting systems with short orbital periods, whereas known hosts are in their vast majority RV systems (hence allowing for a range of orbital inclinations) whose planets span a wider range in terms of orbital period (the median orbital period of planet "b" around main-sequence hosts in the left panel of Figure 12 is 480.3 days).

With over 100 solar-type and red-giant known hosts with detectable solar-like oscillations, this represents an invaluable stellar sample. The impact of having additional constraints from *TESS* asteroseismology on the characterization of known exoplanet-host stars, and consequently of their planetary systems, remains to be fully assessed. Also, we note that all but one system in Figure 12 were discovered using RV measurements and hence will be potential prime targets for the upcoming ESA *CHaracterising ExOPlanet Satellite* (*CHEOPS*; Fortier et al. 2014). *CHEOPS* will be monitoring bright ($V < 12$) known hosts anywhere in the sky for transiting planets. Consequently, *TESS* could be providing asteroseismic measurements for a significant number of potential *CHEOPS* targets, a link that is yet to be explored.

6. SUMMARY AND DISCUSSION

We have developed a simple test to estimate the detectability of solar-like oscillations in *TESS* photometry of any given star (Section 3.1). The detection test looks for signatures of the bell-shaped power excess due to the oscillations. We applied the detection test along stellar-model tracks spanning a range of masses in order to predict the detectability of solar-like oscillations across the H-R diagram (Section 3.2).

Detection of the power excess due to the oscillations as considered here, and hence the ability to measure $\nu_{\rm max}$, will generally mean that the large frequency separation $\Delta\nu$ can be readily extracted. Fundamental stellar properties can be estimated by comparing these two global asteroseismic parameters and complementary spectroscopic observables to the outputs of stellar evolutionary models. This so-called grid-based approach to the determination of stellar properties is currently well established (e.g., Stello et al. 2009b; Basu et al. 2010, 2012; Creevey et al. 2012). A systematic study of *Kepler* planet-candidate hosts using asteroseismology was performed by Huber et al. (2013b), in which fundamental properties were determined for 66 host stars (with typical uncertainties of 3% and 7% in radius and mass, respectively) based on their average asteroseismic parameters. A similar approach was followed by Chaplin et al. (2014a) in estimating the fundamental properties of more than 500 main-sequence and subgiant field stars that had been observed for one month each with *Kepler*. For a subset of 87 of those stars, for which spectroscopic estimates of $T_{\rm eff}$ and metallicity were available, the median uncertainties obtained were 2.2% in radius and 5.4% in mass, with 57% of the stars having age uncertainties smaller than 1 Gyr. An outlook on the precision achievable by

---

[21] http://www.exoplanets.org/

TESS on the estimation of stellar properties for a fiducial low-luminosity red giant is given in Davies & Miglio (2016).

Furthermore, novel strategies have been developed that allow determining the stellar surface gravity for large samples of stars by directly measuring the amplitude of the brightness variations due to granulation and acoustic oscillations in the light curves (Bastien et al. 2013; Kallinger et al. 2016). However, owing to the shorter duration of *TESS* time series compared to *Kepler*'s and the fact that the instrumental/shot noise is now expected to dominate over granulation (cf. Figure 4), the robustness of such techniques when applied to *TESS* photometry remains to be tested. We have not addressed this issue here.

Based on an existing all-sky stellar and planetary synthetic population, we predicted the asteroseismic yield of the *TESS* mission, placing emphasis on the yield of exoplanet-host stars for which we expect to detect solar-like oscillations. This was done for both the target hosts (Section 4.1) and the full-frame-image or FFI hosts (Section 4.2). We predict that asteroseismology will become possible for a few dozen target hosts (mainly subgiant stars but also for a smaller number of F dwarfs) and for up to 200 FFI hosts (at the low-luminosity end of the red-giant branch). We also conducted a similar exercise based on a compilation of known host stars (Section 5), with the prediction being that over 100 solar-type and red-giant known hosts will have detectable solar-like oscillations. Altogether, this equates to a threefold improvement in the asteroseismic yield of exoplanet-host stars when compared to *Kepler*'s.

In Section 4.1 we further advocate for the inclusion of as many bright subgiants as possible in the 2-minute cadence slots reserved for asteroseismology, where we assess the overall asteroseismic potential of subgiant stars and the resulting impact on the asteroseismic yield of target hosts. We should be able to use parallaxes from the ongoing *Gaia* mission to deliberately target these bright subgiants. More generally, *Gaia*-derived luminosities could be used as strong constraints on the asteroseismic modeling, which should help improve the accuracy of the inferred stellar properties, in particular the stellar age.

The authors acknowledge the support of the UK Science and Technology Facilities Council (STFC). Funding for the Stellar Astrophysics Centre is provided by The Danish National Research Foundation (Grant DNRF106). D.H. acknowledges support by the Australian Research Council's Discovery Projects funding scheme (project number DE140101364) and support by the National Aeronautics and Space Administration under Grant NNX14AB92G issued through the *Kepler* Participating Scientist Program. This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program. This research has made use of the Exoplanet Orbit Database and the Exoplanet Data Explorer at exoplanets.org.

Facility: TESS.

REFERENCES

Aigrain, S., Favata, F., & Gilmore, G. 2004, A&A, 414, 1139
Akeson, R. L., Chen, X., Ciardi, D., et al. 2013, PASP, 125, 989
Anderson, E., & Francis, C. 2012, AstL, 38, 331
Appourchaux, T. 2004, A&A, 428, 1039
Ballot, J., Barban, C., & van't Veer-Menneret, C. 2011, A&A, 531, A124
Barclay, T., Rowe, J. F., Lissauer, J. J., et al. 2013, Natur, 494, 452
Basri, G., Walkowicz, L. M., Batalha, N., et al. 2011, AJ, 141, 20
Bastien, F. A., Stassun, K. G., Basri, G., & Pepper, J. 2013, Natur, 500, 427
Bastien, F. A., Stassun, K. G., & Pepper, J. 2014, ApJL, 788, L9
Basu, S., Chaplin, W. J., & Elsworth, Y. 2010, ApJ, 710, 1596
Basu, S., Verner, G. A., Chaplin, W. J., & Elsworth, Y. 2012, ApJ, 746, 76
Bedding, T. R., Kjeldsen, H., Reetz, J., & Barbuy, B. 1996, MNRAS, 280, 1155
Beichman, C., Benneke, B., Knutson, H., et al. 2014, PASP, 126, 1134
Bilir, S., Ak, S., Karaali, S., et al. 2008, MNRAS, 384, 1178
Borucki, W. J., Koch, D., Basri, G., et al. 2010, Sci, 327, 977
Bowler, B. P., Johnson, J. A., Marcy, G. W., et al. 2010, ApJ, 709, 396
Caldwell, J. A. R., Cousins, A. W. J., Ahlers, C. C., van Wamelen, P., & Maritz, E. J. 1993, SAAOC, 15, 1
Campante, T. L., Barclay, T., Swift, J. J., et al. 2015, ApJ, 799, 170
Campante, T. L., Chaplin, W. J., Lund, M. N., et al. 2014, ApJ, 783, 123
Campante, T. L., Lund, M. N., Kuszlewicz, J. S., et al. 2016, ApJ, 819, 85
Carter, J. A., Agol, E., Chaplin, W. J., et al. 2012, Sci, 337, 556
Chaplin, W. J., Basu, S., Huber, D., et al. 2014a, ApJS, 210, 1
Chaplin, W. J., Bedding, T. R., Bonanno, A., et al. 2011a, ApJL, 732, L5
Chaplin, W. J., Elsworth, Y., Davies, G. R., et al. 2014b, MNRAS, 445, 946
Chaplin, W. J., Houdek, G., Appourchaux, T., et al. 2008, A&A, 485, 813
Chaplin, W. J., Kjeldsen, H., Bedding, T. R., et al. 2011b, ApJ, 732, 54
Chaplin, W. J., & Miglio, A. 2013, ARA&A, 51, 353
Chaplin, W. J., Sanchis-Ojeda, R., Campante, T. L., et al. 2013, ApJ, 766, 101
Ciceri, S., Lillo-Box, J., Southworth, J., et al. 2015, A&A, 573, L5
Corsaro, E., Fröhlich, H.-E., Bonanno, A., et al. 2013, MNRAS, 430, 2313
Creevey, O. L., Doğan, G., Frasca, A., et al. 2012, A&A, 537, A111
Davies, G. R., & Miglio, A. 2016, arXiv:1601.02802
Davies, G. R., Silva Aguirre, V., Bedding, T. R., et al. 2016, MNRAS, 456, 2183
Dressing, C. D., & Charbonneau, D. 2015, ApJ, 807, 45
Flower, P. J. 1996, ApJ, 469, 355
Fortier, A., Beck, T., Benz, W., et al. 2014, Proc. SPIE, 9143, 91432J
Fressin, F., Torres, G., Charbonneau, D., et al. 2013, ApJ, 766, 81
Frewen, S. F. N., & Hansen, B. M. S. 2016, MNRAS, 455, 1538
Fröhlich, C., Romero, J., Roth, H., et al. 1995, SoPh, 162, 101
García, R. A., Mathur, S., Salabert, D., et al. 2010, Sci, 329, 1032
Gettel, S., Charbonneau, D., Dressing, C. D., et al. 2016, ApJ, 816, 95
Gilliland, R. L., Chaplin, W. J., Dunham, E. W., et al. 2011, ApJS, 197, 6
Gilliland, R. L., Chaplin, W. J., Jenkins, J. M., Ramsey, L. W., & Smith, J. C. 2015, AJ, 150, 133
Girardi, L., Groenewegen, M. A. T., Hatziminaoglou, E., & da Costa, L. 2005, A&A, 436, 895
Han, E., Wang, S. X., Wright, J. T., et al. 2014, PASP, 126, 827
Howell, S. B., Rowe, J. F., Bryson, S. T., et al. 2012, ApJ, 746, 123
Howell, S. B., Sobeck, C., Haas, M., et al. 2014, PASP, 126, 398
Huber, D. 2015, arXiv:1511.07441
Huber, D., Bedding, T. R., Stello, D., et al. 2011, ApJ, 743, 143
Huber, D., Carter, J. A., Barbieri, M., et al. 2013a, Sci, 342, 331
Huber, D., Chaplin, W. J., Christensen-Dalsgaard, J., et al. 2013b, ApJ, 767, 127
Johnson, J. A., Fischer, D. A., Marcy, G. W., et al. 2007, ApJ, 665, 785
Johnson, J. A., Howard, A. W., Bowler, B. P., et al. 2010, PASP, 122, 701
Jordi, K., Grebel, E. K., & Ammon, K. 2006, A&A, 460, 339
Kallinger, T., De Ridder, J., Hekker, S., et al. 2014, A&A, 570, A41
Kallinger, T., Hekker, S., Garcia, R. A., Huber, D., & Matthews, J. M. 2016, SciA, 2, 1500654
Kallinger, T., Mosser, B., Hekker, S., et al. 2010, A&A, 522, A1
Karoff, C., Campante, T. L., Ballot, J., et al. 2013, ApJ, 767, 34
Kjeldsen, H., & Bedding, T. R. 1995, A&A, 293, 87
Lillo-Box, J., Barrado, D., Moya, A., et al. 2014, A&A, 562, A109
Mathur, S., García, R. A., Ballot, J., et al. 2014, A&A, 562, A124
Mathur, S., Hekker, S., Trampedach, R., et al. 2011, ApJ, 741, 119
Michel, E., Baglin, A., Auvergne, M., et al. 2008, Sci, 322, 558
Mosser, B., Belkacem, K., Goupil, M.-J., et al. 2010, A&A, 517, A22
Mosser, B., Elsworth, Y., Hekker, S., et al. 2012, A&A, 537, A30
Paxton, B., Bildsten, L., Dotter, A., et al. 2011, ApJS, 192, 3
Paxton, B., Cantiello, M., Arras, P., et al. 2013, ApJS, 208, 4
Perryman, M. A. C., de Boer, K. S., Gilmore, G., et al. 2001, A&A, 369, 339
Pijpers, F. P. 2003, A&A, 400, 241
Quinn, S. N., White, T. R., Latham, D. W., et al. 2015, ApJ, 803, 49
Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, JATIS, 1, 014003
Silva Aguirre, V., Davies, G. R., Basu, S., et al. 2015, MNRAS, 452, 2127

Skumanich, A. 1972, ApJ, 171, 565
Sliski, D. H., & Kipping, D. M. 2014, ApJ, 788, 148
Stello, D., Bruntt, H., Kjeldsen, H., et al. 2007, MNRAS, 377, 584
Stello, D., Chaplin, W. J., Basu, S., Elsworth, Y., & Bedding, T. R. 2009a, MNRAS, 400, L80
Stello, D., Chaplin, W. J., Bruntt, H., et al. 2009b, ApJ, 700, 1589
Sullivan, P. W., Winn, J. N., Berta-Thompson, Z. K., et al. 2015, ApJ, 809, 77
Torres, G. 2010, AJ, 140, 1158
Van Eylen, V., & Albrecht, S. 2015, ApJ, 808, 126
Yee, J. C., & Gaudi, B. S. 2008, ApJ, 688, 616

Bibliography

Aguirre V. S., et al., 2015, Monthly Notices of the Royal Astronomical Society, 452, 2127

Aguirre V. S., et al., 2017, The Astrophysical Journal, 835, 173

Allen C. W., 1978, Astrophysical Quantities. The Athlone Press

Anderson E., Francis C., 2012, Astronomy Letters, 38, 331

Appourchaux T., et al., 2010, The Astronomy and Astrophysics Review, 18, 197

Appourchaux T., et al., 2012, Astronomy and Astrophysics, 543, A54

Appourchaux T., et al., 2014, Astronomy and Astrophysics, 566, A20

Astraatmadja T. L., Bailer-Jones C. A. L., 2016, The Astrophysical Journal, 833, 119

Baglin A., et al., 2006. p. 3749, `http://adsabs.harvard.edu/abs/2006cosp...36.3749B`

Bailer-Jones C. A. L., Rybizki J., Fouesneau M., Mantelet G., Andrae R., 2018, preprint, 1804, arXiv:1804.10121

Ball W. H., Themeßl N., Hekker S., 2018, Monthly Notices of the Royal Astronomical Society, 478, 4697

Ballot J., Barban C., van't Veer-Menneret C., 2011, Astronomy and Astrophysics, 531, A124

Barlow R. J., 1989, Statistics. John Wiley & Sons

Basu S., Chaplin W. J., 2017, Asteroseismic Data Analysis, `https://press.princeton.edu/titles/11170.html`

Beck P. G., 2013, Ph.D. Thesis

Beck P. G., et al., 2011, Science, 332, 205

Bedding T. R., 2011, arXiv:1107.1723 [astro-ph]

Bedding T. R., Kjeldsen H., Reetz J., Barbuy B., 1996, Monthly Notices of the Royal Astronomical Society, 280, 1155

Bellinger E. P., Angelou G. C., Hekker S., Basu S., Ball W. H., Guggenberger E., 2016, The Astrophysical Journal, 830, 31

Borucki W. J., et al., 2010, Science, 327, 977

Bovy J., Rix H.-W., Green G. M., Schlafly E. F., Finkbeiner D. P., 2016, The Astrophysical Journal, 818, 130

Brandt S., 1970, Statistical and computational methods in data analysis. `http://adsabs.harvard.edu/abs/1970scmd.book.....B`

Bruntt H., et al., 2010, Monthly Notices of the Royal Astronomical Society, 405, 1907

Caldwell J. A. R., Cousins A. W. J., Ahlers C. C., van Wamelen P., Maritz E. J., 1993, South African Astronomical Observatory Circular, 15, 1

Caldwell D. A., et al., 2010. p. 59.01, `http://adsabs.harvard.edu/abs/2010DPS....42.5901C`

Campante T. L., et al., 2016, preprint, 1608, arXiv:1608.01138

Campante T., Santos N., Monteiro M., eds, 2018, Asteroseismology and Exoplanets: Listening to the Stars and Searching for New Worlds: IVth Azores International Advanced School in Space Sciences. Astrophysics and Space Science Proceedings, Springer International Publishing, `//www.springer.com/gb/book/9783319593142`

Carrier F., Eggenberger P., D'Alessandro A., Weber L., 2005, New Astronomy, 10, 315

Chaplin W. J., Miglio A., 2013, Annual Review of Astronomy and Astrophysics, 51, 353

Chaplin W. J., et al., 2011, The Astrophysical Journal, 732, 54

Chaplin W. J., et al., 2014a, The Astrophysical Journal Supplement Series, 210, 1

Chaplin W. J., Elsworth Y., Davies G. R., Campante T. L., Handberg R., Miglio A., Basu S., 2014b, Monthly Notices of the Royal Astronomical Society, 445, 946

Chaplin W. J., Schofield M., Campante T., Huber D., WG1/2 Chairs 2016a, TESS Asteroseismic Target List for Solar-Like Oscillators, TASC WG1/2 target selection report (see http://tasoc.dk)

Chaplin W. J., Schofield M., Campante T., Huber D., WG1/2 Chairs 2016c, TESS Asteroseismic Target List (ATL) for Solar-Like Oscillators II. ATL made with TGAS and Hipparcos parallaxes, TASC WG1/2 target selection report (see http://tasoc.dk)

Chaplin W. J., Schofield M., Campante T., Huber D., WG1/2 Chairs 2016b, TESS Asteroseismic Target List for Solar-Like Oscillators, TASC WG1/2 target selection report (see http://tasoc.dk)

Christensen-Dalsgaard J., 2002, Reviews of Modern Physics, 74, 1073

Christensen-Dalsgaard J., 2003, Lecture notes on Stellar Oscillations, `http://w.astro.berkeley.edu/~eliot/Astro202/2009_Dalsgaard.pdf`

Cutri R. M., et al., 2003, VizieR Online Data Catalog, p. II/246

Davies G. R., Miglio A., 2016, Astronomische Nachrichten, 337, 774

Davies G. R., et al., 2016, Monthly Notices of the Royal Astronomical Society, 456, 2183

Davies G. R., et al., 2017, preprint, 1701, arXiv:1701.02506

Debosscher J., Sarro L. M., Aerts C., Cuypers J., Vandenbussche B., Garrido R., Solano E., 2007, Astronomy and Astrophysics, 475, 1159

Drimmel R., Cabrera-Lavers A., López-Corredoira M., 2003, Astronomy and Astrophysics, 409, 205

Elorrieta F., et al., 2016, Astronomy and Astrophysics, 595, A82

Elsworth Y., Hekker S., Basu S., R. Davies G., 2017, Monthly Notices of the Royal Astronomical Society, 466, 3344

Epstein C. R., et al., 2014, The Astrophysical Journal, 785, L28

Eyer L., Holl B., Pourbaix D., Mowlavi N., Siopis C., Barblan F., Evans D. W., North P., 2013, Central European Astrophysical Bulletin, 37, 115

Flower P. J., 1996, The Astrophysical Journal, 469, 355

Frasca A., et al., 2016, Astronomy and Astrophysics, 594, A39

Gaia Collaboration et al., 2016, Astronomy and Astrophysics, 595, A2

Gilliland R. L., et al., 2010, The Astrophysical Journal Letters, 713, L160

Green G. M., et al., 2015, The Astrophysical Journal, 810, 25

Harvey J., 1985. `http://adsabs.harvard.edu/abs/1985ESASP.235..199H`

Henden A. A., Welch D. L., Terrell D., Levine S. E., 2009. p. 407.02, `http://adsabs.harvard.edu/abs/2009AAS...21440702H`

Høg E., et al., 2000, Astronomy and Astrophysics, 355, L27

Hon M., Stello D., Yu J., 2017, Monthly Notices of the Royal Astronomical Society, 469, 4578

Hon M., Stello D., Yu J., 2018, Monthly Notices of the Royal Astronomical Society

Houdek G., 2006, ArXiv Astrophysics e-prints, pp arXiv:astro–ph/0612024

Howell S. B., et al., 2014, Publications of the Astronomical Society of the Pacific, 126, 398

Huang Y., Liu X.-W., Yuan H.-B., Xiang M.-S., Chen B.-Q., Zhang H.-W., 2015, Monthly Notices of the Royal Astronomical Society, 454, 2863

Huber D., 2015, arXiv:1511.07441 [astro-ph]

Huber D., Stello D., Bedding T. R., Chaplin W. J., Arentoft T., Quirion P.-O., Kjeldsen H., 2009, Communications in Asteroseismology, 160, 74

Huber D., et al., 2010, The Astrophysical Journal, 723, 1607

Huber D., et al., 2011, The Astrophysical Journal, 743, 143

Huber D., et al., 2017, The Astrophysical Journal, 844, 102

Iben Jr. I., 2013, Stellar Evolution Physics, Volume 1: Physical Processes in Stellar Interiors. `http://adsabs.harvard.edu/abs/2013sepp.book.....I`

Jones E., Oliphant T., Peterson P., 2001

Kallinger T., et al., 2014, Astronomy and Astrophysics, 570, A41

Keenan P. C., McNeil R. C., 1989, The Astrophysical Journal Supplement Series, 71, 245

Kippenhahn R., Weigert A., Weiss A., 2012, Stellar Structure and Evolution, doi:10.1007/978-3-642-30304-3. , `http://adsabs.harvard.edu/abs/2012sse..book.....K`

Kjeldsen H., Bedding T. R., 1995, Astronomy and Astrophysics, 293, 87

Larson R. B., 2003, Reports on Progress in Physics, 66, 1651

Libbrecht K. G., 1992, The Astrophysical Journal, 387, 712

Lindegren L., et al., 2018, preprint, 1804, arXiv:1804.09366

Lund M. N., et al., 2014, Astronomy and Astrophysics, 570, A54

Lund M. N., Handberg R., Davies G. R., Chaplin W. J., Jones C. D., 2015, The Astrophysical Journal, 806, 30

Lund M. N., et al., 2016, Monthly Notices of the Royal Astronomical Society, 463, 2600

Lund M. N., et al., 2017, The Astrophysical Journal, 835, 172

Marshall D. J., Robin A. C., Reylé C., Schultheis M., Picaud S., 2006, Astronomy and Astrophysics, 453, 635

Mathur S., et al., 2011, The Astrophysical Journal, 741, 119

Michalik D., Lindegren L., Hobbs D., 2015, Astronomy and Astrophysics, 574, A115

Miglio A., Montalbán J., Noels A., eds, 2012, Red Giants as Probes of the Structure and Evolution of the Milky Way. Astrophysics and Space Science Proceedings, Springer-Verlag, Berlin Heidelberg, `//www.springer.com/gb/book/9783642184178`

Miller J. N., 1991, Analyst, 116, 3

Mosser B., Deheuvels S., Michel E., ThÃľvenin F., Dupret M. A., Samadi R., Barban C., Goupil M. J., 2008, Astronomy and Astrophysics, 488, 635

Mosser B., et al., 2010, Astronomy and Astrophysics, 517, A22

Mosser B., et al., 2012, Astronomy & Astrophysics, 537, A30

Ness M., Hogg D. W., Rix H.-W., Ho A. Y. Q., Zasowski G., 2015, The Astrophysical Journal, 808, 16

Nun I., Pichara K., Protopapas P., Kim D.-W., 2014, The Astrophysical Journal, 793, 23

Olivero J. J., Longbothum R. L., 1977, Journal of Quantitative Spectroscopy and Radiative Transfer, 17, 233

Perryman M. A. C., et al., 1997, Astronomy and Astrophysics, 323

Phillips J. M., Tai W. M., 2017, preprint, 1710, arXiv:1710.04325

Pijpers F. P., 2003, Astronomy and Astrophysics, 400, 241

Pinsonneault M. H., et al., 2014, The Astrophysical Journal Supplement Series, 215, 19

Rauer H., et al., 2014, Experimental Astronomy, 38, 249

Ricker G. R., et al., 2014, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Samadi R., Belkacem K., Dupret M.-A., Ludwig H.-G., Baudin F., Caffau E., Goupil M.-J., Barban C., 2012, Astronomy and Astrophysics, 543, A120

Santos A. R. G., et al., 2018, The Astrophysical Journal Supplement Series, 237, 17

Sarro L. M., Debosscher J., López M., Aerts C., 2009, Astronomy and Astrophysics, 494, 739

Scott D. W., 1979, Biometrika, 66, 605

Scuflaire R., Théado S., Montalbán J., Miglio A., Bourge P.-O., Godart M., Thoul A., Noels A., 2008, Astrophysics and Space Science, 316, 83

Sharma S., Stello D., Bland-Hawthorn J., Huber D., Bedding T. R., 2016, arXiv:1603.05661 [astro-ph]

Silva Aguirre V., et al., 2012, The Astrophysical Journal, 757, 99

Skrutskie M. F., et al., 2006, The Astronomical Journal, 131, 1163

Soubiran C., Campion J.-F. L., Brouillet N., Chemin L., 2016, arXiv:1605.07384 [astro-ph]

Stello D., et al., 2007, Monthly Notices of the Royal Astronomical Society, 377, 584

Stello D., Chaplin W. J., Basu S., Elsworth Y., Bedding T. R., 2009, Monthly Notices of the Royal Astronomical Society, 400, L80

Sullivan P. W., et al., 2015, The Astrophysical Journal, 809, 77

Thomas A., Stevenson E., Gittins F. W. R., Miglio A., Davies G., Girardi L., Campante T. L., Schofield M., 2017. eprint: arXiv:1610.08862, p. 05006, doi:10.1051/epjconf/201716005006, `http://adsabs.harvard.edu/abs/2017EPJWC.16005006T`

Torres G., 2010, The Astronomical Journal, 140, 1158

Torres G., Andersen J., GimÃĺnez A., 2010, Astronomy and Astrophysics Review, 18, 67

Toutain T., Appourchaux T., 1994, Astronomy and Astrophysics, 289, 649

Valenzuela L., Pichara K., 2018, Monthly Notices of the Royal Astronomical Society, 474, 3259

Wegner P., 1960, Commun. ACM, 3, 322

Yu J., Huber D., Bedding T. R., Stello D., Murphy S. J., Xiang M., Bi S., Li T., 2016, Monthly Notices of the Royal Astronomical Society, p. stw2074

van Leeuwen F., 2007. , doi:10.1007/978-1-4020-6342-8, `http://adsabs.harvard.edu/abs/2007ASSL..350.....V`