MSc thesis

Particle Physics and Astrophysical Sciences

Statistical analysis of lightcurves of active stars

András Kristóf Haris-Kiss

June 5, 2022

Supervisor: Mikko Tuomi

Examiners: Karri Muinonen
Mikko Tuomi

University of Helsinki
Faculty of Science

PL 64 (Gustaf Hällströmin katu 2a)
00014 Helsingin yliopisto

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty	Koulutusohjelma — Utbildningsprogram — Degree programme
Faculty of Science	Particle Physics and Astrophysical Sciences

Tekijä — Författare — Author

András Kristóf Haris-Kiss

Työn nimi — Arbetets titel — Title

Statistical analysis of lightcurves of active stars

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidantal — Number of pages
MSc thesis	June 5, 2022	72

Tiivistelmä — Referat — Abstract

Over the last thirty years more than five thousand exoplanets have been discovered around a wide variety of stellar objects. Most exoplanets have been discovered using the transit method, which relies on observing the periodic brightness changes of stars as their planet transits in front of them. The discovery efficiency of these planets has been strongly enhanced with the advent of space telescopes dedicated to the discovery of planets using the transit method.

Planetary signals in the photometric data of active stars can be challenging to find, as the surface features of the stars combined with their rotation might produce signals which are orders of magnitude stronger than those caused by the planetary transit.

The question of what statistical methods should be applied to account for the innate variability of stars in order to identify the transits of exoplanets in the lightcurves of active stars is being investigated in this thesis.

I test a number of statistical methods in order to combat stellar activity and to identify planetary transit signals. The rotation period of the star is investigated using the Lomb-Scargle and likelihood ratio periodograms. Starspot induced variability is approximated with a number of sinusoids, with periods based on the star's rotation period. Additional stellar activity is filtered out using autoregressive and moving average models.

Model fittings are performed with least squares fitting, and using samples generated by the Adaptive Metropolis algorithm. After the lightcurve has been detrended for stellar activity, the likelihoods of planetary transit signals are assessed with a box-fitting algorithm. Models are compared with the Bayesian and Akaike information criteria. Planetary characteristics are then estimated by modeling the shape of the transit lightcurve.

These methods are tested and performed on the lightcurve of HD 110082, a highly active young star with one confirmed planetary companion, based on the observations of the TESS space telescope. I find that stellar activity is sufficiently filtered out with a model containing four sinusoid signals. The signal corresponding to the planet is confirmed by the box fitting algorithm, agreeing with results available in scientific literature.

Avainsanat — Nyckelord — Keywords

Exoplanets, Stellar Activity, Observational Astronomy, Photometry

Säilytyspaikka — Förvaringsställe — Where deposited

Muita tietoja — Övriga uppgifter — Additional information

Acknowledgements

I owe an immense amount of gratitude to my supervisor, Mikko Tuomi, who guided me throughout my work and without whom this thesis would never have been possible to make.

I wish to thank Karri Muinonen for having accepted to be the examiner of this thesis.

I thank Thomas Hackman, Jyri Lehtinen and Teemu Willamo for welcoming me in their group, for their insightful comments and our regular Monday talks.

To all the people in the Opiskelijahuone, and Meridiaani, thank you for a wonderful year I got to spend with you. To Iida, who made everything more colorful, kiitos.

Finally, I want to thank my family who were always there for me and supported me in reaching for the stars. A special thanks for my Father, Mother, and Grandma. Köszönöm.

Contents

1. Introduction

1.1 Exoplanets

Planets outside our Solar System are called exoplanets. The first two exoplanets were discovered by Wolszczan and Frail (1992) around the pulsar PSR B1257+12. Three years later, the first exoplanet around a main-sequence star, 51 Pegasi has been found by Mayor and Queloz (1995). The latter discovery was rewarded in 2019 with a Nobel Prize in Physics.

Since the first discoveries, the number of confirmed exoplanets has been steadily increasing, reaching 5 059 by the first half of 2022[†]. These planets show a great diversity in their characteristics. Their orbital periods range from well below an hour (e.g., K2-137 b) to over 22 000 years (Oph 98 b). Their radii are between that of Mercury (e.g., Kepler-391 b) and more than twice that of Jupiter (e.g., ROXs 42B (AB) b). Exoplanets have also been found to orbit a wide variety of objects. They have been found around main-sequence stars, binary stars (the first planet found to orbit two main-sequence stars, Kepler-16b was discovered by Doyle et al. 2011), pulsars, and in some instances orbiting no star at all (e.g., PSO J318.522). This last type of planets, the so-called rogue planets are theorized to have been ejected from their planetary system through two body interactions at the time when their planetary system was forming.

The study of exoplanets has several applications. Observations of exoplanetary systems can be used to test theories about the formation and evolution of planetary systems. This also helps in developing a better understanding of the formation of our own Solar System and Earth.

Many exoplanets have been found in the habitable zones around their stars. These are regions where planets are able to sustain liquid water on their surfaces over geological timescales. The discovery of such planets sparked a renewed interest in astrobiology and the search of extraterrestrial life.

A variety of methods exists for the detection of exoplanets. Some of the most

[†] exoplanet.eu, accessed on May 29, 2022.

successful methods of exoplanet detection are:

- Transit method: The transit of a planet in front of its host star reduces the amount of light coming from the star to the observer. The periodic dimmings of the star can therefore be used to detect exoplanets.
- Radial velocity: A star's radial velocity can be measured from the Doppler shift of its spectral lines. If a star's radial velocity changes periodically, it is a sign that the star orbits a common center of mass with another object.
- Microlensing: As a foreground star moves in front of a background star, the background star becomes brighter due to the gravitational lensing effect of the foreground star. If a planet orbits the foreground star, an other brightening of the background star can be observed due to the gravitational lensing effect of the planet, close in time (typically within days) to the lensing event of the foreground star.
- Direct imaging: An image is taken of the host star and its environment, including possible planets.
- Timing measurements: If multiple planets are present in a system that has at least one transiting planet, the transit of a planet in front of its star does not occur strictly periodically, due to the perturbing effect of an other planet. Transit time variations can therefore be used to detect planets that would not necessarily transit their host star.

1.2 Transiting exoplanets

1.2.1 Transit method

As of the publication of this thesis, the most effective method for detecting exoplanets (with 3 552 discoveries) has been the transit method.

A schematic plot of the flux F coming from a star–planet system at different orbital phases of the planet is shown on Fig. 1.1. The brightness variations of an object as a function of time is an object's lightcurve. The amount of relative dimming when the planet is in front of its star $\delta = \Delta F/F$, the length of the transit event T, the mid-transit time (the time at which the planet is the closest to the center of the star) t_c, and the ingress time (the time in which the planet's disk is only partially over the stellar disk) τ can be measured directly from the shape of the system's lightcurve.

The amount of dimming in the star's lightcurve can be used to estimate the size of the planet, as a larger planet will cover a bigger fraction of the stellar disk. This

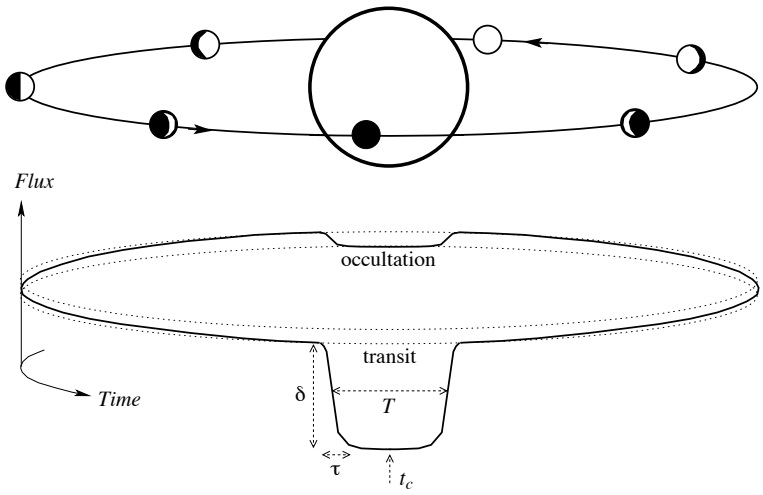

Figure 1.1: The lightcurve of a star with a transiting exoplanet. As the planet covers a part of the star's surface, the star's observed brightness decreases (primary transit). The incoming flux is not constant during the time the planet's disk is fully on the disk of the star, because the star shows some limb darkening. When the planet is closer to the edge of the stellar disk it covers parts from the star with lower surface brightness. This leads to a smaller decrease in brightness compared to when it is at the middle of the stellar disk, in front of high surface brightness parts. A secondary dimming can be observed at occultation, when the planet disappears behind the stellar disk (secondary transit). (Image credit: Winn, 2009)

leads to an increased level of dimming during the transit event. It follows from here that there is a bias in transit-method-based discoveries towards planets with larger radii.

If transit based radius measurements are combined with planetary mass data, which is calculated from radial velocity measurements, a density estimate can be made for the planet. Based on density, it is possible to approximate the composition, and put constraints on the habitability of the planet and the formation of the planetary system.

The orbit of a planet around its star on an elliptic orbit is described by six orbital elements. These elements are the 1) semi-major axis a, 2) eccentricity e, 3) inclination i, 4) longitude of the ascending node Ω, 5) argument of periapsis ω, and 6) true anomaly at epoch t_0, ν.

A requirement for the transit method to work is that we observe the orbital plane of the planet nearly from edge on. In other words, the planet's inclination (the angle between its orbital plane and the plane of the sky) must be close to 90°, otherwise its transit would not be observable from Earth. This is a considerable disadvantage of the transit method, since the probability of seeing a planetary system at the right inclination angle is relatively low. The transit probability of a planet on a circular orbit is given by the transit probability equation $p = R_*/a$, where R_* is the radius of

the star and a is the semi-major axis of the planet. This translates to a hot Jupiter (an exoplanet with a radius similar to Jupiter on an $a \approx 0.05$ au orbit) having a $\sim 10\%$ transit probability in front of a Sun-sized star.

When estimating the occurrence rate of different exoplanet populations in an observed sample of stars, geometrical considerations, such as the one proposed by the transit probability equation, have to be taken into account. This correction was, for example used by Howard et al. (2012), who report an occurrence rate of 0.005 ± 0.001 of hot Jupiters around G and K spectral class stars with Kepler magnitude $K_p < 16$ (Brown et al., 2011) in the sample of stars originally observed by the Kepler space telescope (the occurrence rate was also corrected for the instrumental capacities of the telescope).

A consequence of the transit probability equation is a bias towards smaller semi-major axes among the planets discovered with the transit method. This bias is further strengthened by the fact that planets with smaller semi-major axes have shorter orbital periods. Planets with shorter orbital periods transit their stars more frequently. This makes the transits more likely to be observed by a telescope at a random time. Fig. 1.2 shows the distribution of exoplanets with known semi-major axes and radii in terms of these two parameters. The planets are separated into two classes based on whether or not they were discovered with the transit method. The preference of the transit method towards smaller semi-major axes is demonstrated by the clustering of transiting planets at lower a values. Another consequence of the transit probabilities is that the presence

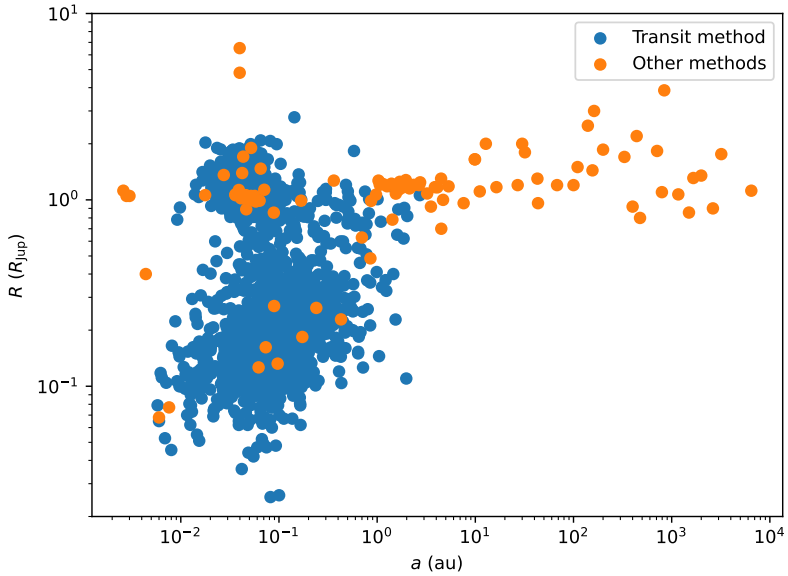

Figure 1.2: Semi-major axes and radii of planets where both parameters are known. Detection methods are marked with different colors. A clear bias towards smaller semi-major axes is visible among planets discovered with the transit method. Data from exoplanet.eu.

of planets around a star, especially on wider orbits cannot be ruled out by a lack of planetary transit observations.

As a planet transits the stellar disk, it might go in front of regions that have different surface brightnesses than what is expected from a featureless stellar disk (e.g., starspots). When this happens, the lightcurve changes its shape during the transit compared to the lightcurve model shown on Fig. 1.1, as the planet covers regions with a different brightness. The change in lightcurve shapes can be used to probe the star's surface for spots or other features. This technique is called transit mapping, and has been used for over a decade (e.g., Pont et al., 2007).

Exoplanet discoveries made with the transit method are particularly prone to false positives. Several phenomena are known that cause dimmings in the lightcurves of the stars similar to a planetary transit. One reason is, if the star falls close enough on the plane of the sky to an eclipsing binary, their lights are blended together. Due to the presence of the eclipsing binary, the star may show planetary transit-like variations. Another reason is, the star can itself also be part of a binary system. If the binary counterpart only grazes the stellar disk when it eclipses, it only covers a small fraction of the star's disk during its transit. This produces a lightcurve similar to a planetary transit. Finally, the binary partner can be a white or a brown dwarf. Both types of these celestial bodies have similar radii to planets. Even if these objects go in front of the disk of the primary star, the relative change in brightness will be similar to that caused by a planet.

Due to these many opportunities for a false detection, a planetary transit signal either has to be observed three times, or the planet has to be detected using an other detection method before the discovery is confirmed.

A considerable advantage of the transit method over radial velocity measurement lies in the number of stars that can be simultaneously observed. Surveys relying on the transit method can monitor several thousands of stars at the same time, the number is only restricted by the sensitivity of the telescope's sensor and field of view. Radial velocity surveys, in contrast are usually designed to observe one star at a time.

1.2.2 History

The first exoplanet to be observed using the transit method was HD 209458 b in 1999 at Fairborn Observatory in Arizona (Henry et al., 2000). This was a known object, discovered previously from radial velocity measurements.

The first planet to be discovered with the transit method was OGLE-TR-56b (Konacki et al., 2003) whose transit has first been observed at the Las Campanas Observatory in Chile. In the following years, several transiting exoplanets have been

discovered by ground-based surveys (e.g., SuperWasp, HATNet, Collier Cameron et al., 2007, Bakos et al., 2007).

The breakthrough in discovering transiting planets came with the advent of high-precision space-based photometry. Unlike Earth-based telescopes, space telescopes are not limited by daytime and nighttime variations. This allows for a long, continuous time series of observations. Furthermore, space based observations have much higher photometric precision than those based on the Earth. The most precise telescopes on Earth can detect a fractional brightness change of $\delta \sim 0.1\%$, whereas space telescopes can usually detect a fractional brightness change of $\delta \sim 10^{-4}$. The reason behind this difference in sensitivity is the lack of atmospheric effects for space based observations.

Using higher precision photometry, smaller brightness variations, caused by planets with smaller radii become visible and so space telescopes are more sensitive to smaller planets than Earth based telescopes. The discovery of Earth-size and smaller planets became feasible with space based programs. This shift towards smaller planets is illustrated on Fig. 1.3, which shows the radius distribution of planets discovered from the Earth and by the space telescopes CoRoT, Kepler, and TESS.

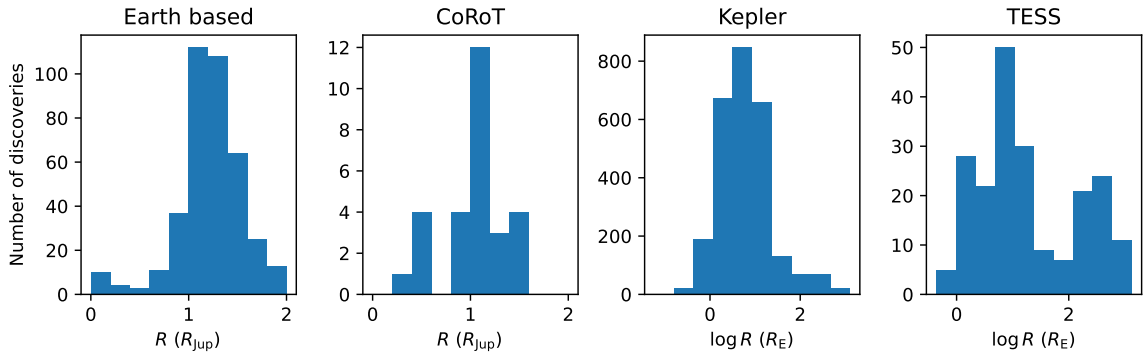

Figure 1.3: Radii of planets discovered by Earth based telescopes, and space telescopes CoRoT, Kepler, and TESS using the transit method. Observe the scaling and the units of measurements used for the individual histograms (The correspondence between Jupiter's and Earth's radius is $R_{\mathrm{Jup}} \approx 11.2\,R_{\mathrm{E}}$). Space based observatories have a higher photometric precision and so are able to discover planets with smaller radii in greater numbers.

An overview of these space telescopes and their surveys is presented in the following paragraphs.

The first space based observatory dedicated to the discovery of exoplanets using the transit method was CoRoT (Convection, Rotation and planetary Transits) space telescope. The mission's other aim was to measure solar-like oscillations in stars. It was launched in 2006 by the French Space Agency (CNES) and ESA, and was decommissioned in 2014. CoRoT was observing two areas on the sky, switching between them every 150 days. The telescope discovered a total of 34 exoplanets. One notable

discovery is CoRoT-7b, the first terrestrial exoplanet to be discovered (Léger et al., 2009).

CoRoT was followed by the Kepler space telescope, launched by NASA in 2009. The telescope was designed to discover Earth-size exoplanets. Between 2009 and 2013 it was aimed at a single area in the sky, observing about 150 000 main-sequence stars in the constellations Cygnus, Lyra, and Draco. In 2013, two of Kepler's four reaction wheels broke. These wheels were responsible for stabilizing the telescope, allowing it to point at the same area on the sky. This failure made it impossible to keep the telescope's field of view on its original target area. From 2014 until its deactivation in 2018, Kepler observed different parts of the sky along the ecliptic, for roughly 90 days at a time. The official designation of this second observing mission was K2. Over its nine year long operation, Kepler made 2 865 confirmed exoplanet discoveries – an overwhelming majority of all discovered exoplanets.

The latest, currently operational space telescope dedicated to the discovery of exoplanets with the transit method is called TESS (Transiting Exoplanet Survey Satellite). The space telescope was launched in 2018 and is operated by NASA. In the first two years of its operation, TESS observed $24° \times 96°$ regions on the sky in 27-day-long observing runs, covering 85 percent of the sky in total. These observations were mostly done with 30 minute cadences. Due to overlaps between different observed regions, areas closer to the ecliptic poles were observed for longer intervals. The observation pattern of the first two years of TESS is shown on Fig. 1.4. Regions missed in the first two years of TESS observations were partially covered in later observing sectors. As of May 2022, TESS has made 217 confirmed exoplanet discoveries.*

1.3 Stellar variability

Stars that show variations in their observed brightness are called variable stars. If the observations are done with high enough precision, all stars show brightness variations. When doing observations in visible light (the surveys described in the last section all operate in this range), we usually detect the brightness changes in the photosphere of stars. The photosphere is the innermost layer of stars that we can observe with optical measurements. One particularity of this layer is limb darkening, which is responsible for the stars having a lower apparent surface brightness closer to the edge of their disk. To quantify limb darkening, in this thesis, I adapt the linear limb darkening coefficient u (Milne, 1921). The stellar disk is u times brighter at its edge than at its center. It follows from this that when we describe the limb darkening of a star u has to be less

*https://exoplanetarchive.ipac.caltech.edu/docs/counts_detail.html accessed on May 29, 2022.

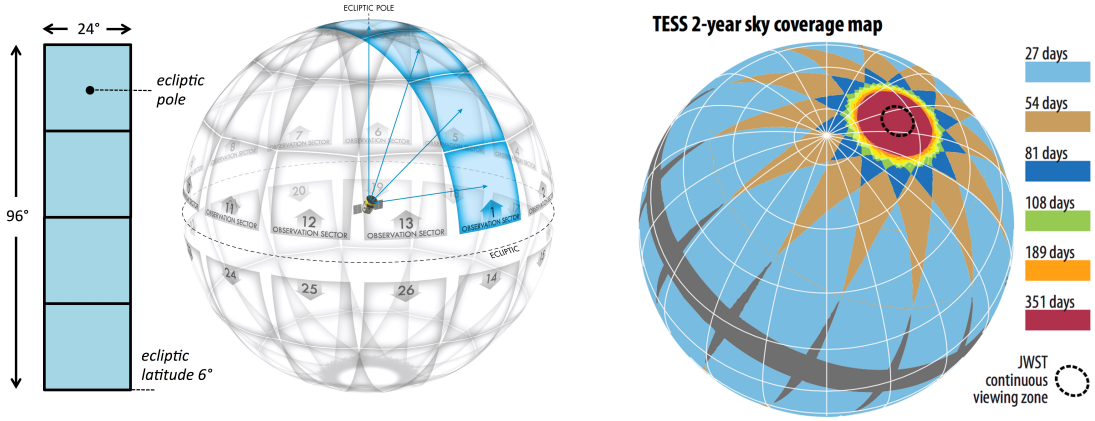

Figure 1.4: Observing pattern of the TESS space telescope in the first two years of its operation. *Left:* The telescope observed 24° × 96° "Sectors" for 27 days, covering most of the sky save for the areas falling close to the ecliptic. *Right:* Observation arcs of different regions on the sky. The color coding corresponds to how long certain regions were monitored. Areas closer to the ecliptic poles have been observed for more than 27 days, due to the overlap between individual sectors. In the regions around the ecliptic poles, TESS obtained a quasi-continuous time series of 351 days (Image credit: NASA, MIT).

than one.

Variable stars are divided into intrinsic or extrinsic variable stars, based on the source of brightness variations.

When the observed brightness of an intrinsic variable star changes, the luminosity (the amount of energy radiated by star the star over time) of the star itself changes.

Extrinsic variable stars change their brightness without significant variation in their luminosity, as a result of some external or geometric effect. This thesis deals with the brightness variations of this second class of variable stars.

Extrinsic variable stars change their brightness for two reasons. Firstly a star might have a non-uniform surface flux from the presence of some features in the photosphere. As the star rotates, these features appear at different angles for the observer, or get behind the stellar disc completely. Secondly, if the star is in a binary or planetary system, its brightness periodically fluctuates as other bodies within those systems orbit in front of it.

The following few sections contain short explanations for the causes behind the uneven surface brightnesses of stars. Combined with stellar rotation, these phenomena are responsible for a significant amount of stellar variability in the case of extrinsic stars.

Starspots

Similarly to our Sun (Fig. 1.5), other stars have been observed to show starspots. These spots are created at places where the magnetic field lines enter the surface of the stars. Regions with strong, complex magnetic fields are referred to as active regions. As the magnetic field lines enter the stellar surface, they hinder convection and essentially block hotter material from within the star from reaching the surface. As a consequence, these regions have lower temperatures (temperature differences typically range between 500 K and 2 000 K) and surface fluxes and appear darker than the rest of the stellar surface.

The number of spots to be observed on the Sun follows the solar activity cycle. Around the maximum of the solar cycle, the number of spots is more, around the minimum of the solar cycle, the number of spots is less.

When observed over an extended period of time, the sunspots closer to the equator seem to rotate faster around the axis of the Sun than those closer to its poles. The phenomenon whereby different latitudes rotate with different angular velocities is known as the differential rotation and has been observed in several stars (e.g., Lanza et al., 2014).

Starspots observed on other stars can cover up to 40% of the stellar disk at the same time (Strassmeier, 1999). The lifetimes of starspots can vary greatly. Based on the example of the Sun, the lifetime of smaller starspots, such as those on the sun is on the order of magnitude of days to weeks. Larger spots can have lifetimes up to several years (Berdyugina, 2005).

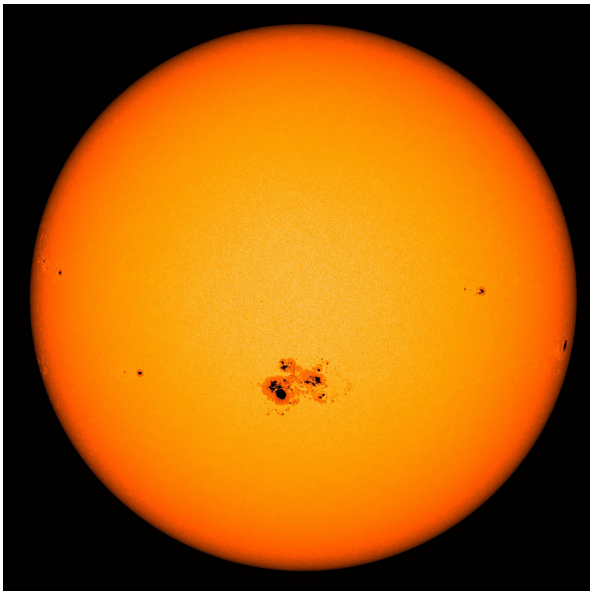

Figure 1.5: Starspots on the surface of the Sun around the peak of its last activity cycle in 2014. Limb darkening is observable close to the edge of the solar disk (Image credit: NASA/SDO).

Flares

Flares are sudden, localized bursts of electromagnetic radiation on the stellar surface. They occur due to the interaction of highly accelerated charged particles with the plasma in the stellar atmosphere. The source of acceleration is the recombination of that star's magnetic field lines. For this reason, flares are inherently linked to active regions. The lowest energy flares can typically be observed in X-ray bands. The highest energy flares are also observable in optical bands and so affect the measurements of exoplanet-hunting space telescopes. These are the so-called white flares. The first observed example of a white flare is referred to as the Carrington Event, after its observer (Carrington, 1859). Flares have typical lifetimes on the scale of minutes. Owing to their short lifetimes, they can be safely ignored when the lightcurve fitting is performed with the aim of looking for or fitting exoplanetary transits, which have timescales on the order of hours.

Stellar flares are especially important with regards to the question of the habitability of planets (Yamashiki et al., 2019, and references therein). Stars release vast amounts of energy in the form of flares, which might affect the atmosphere and habitability of stars that show frequent flaring activity. The subject is further complicated by the fact that the most common types of stars, namely M dwarfs, show particularly strong stellar activity with frequent flares (e.g., West et al., 2008). These stars have effective temperatures between 2 400 K and 3 700 K. This places their habitable zones to 10^{-2} au $\lesssim a \lesssim -10^{-1}$ au, where the effects of flares may be more severe than at greater orbital distances.

Granulation

Granulation is the result of the convective heat transfer in the outer layers of solar-type stars. In the photosphere, it shows as a network of brighter regions. These bright regions are called granules and they correspond to the convection cells in the photosphere of stars, through which hot plasma from the inner regions of the stars reach their surface. There are approximately two million granules on the surface of the Sun with diameters around 2 000 km at any given time (Fig. 1.6). The size of granules is predicted to be several orders of magnitude ($\sim 10^4$) larger on the surfaces of giant and supergiant stars, due to their low surface gravity (Schwarzschild, 1975). The lifetime of one granule is around ten minutes on the Sun, but is less known for stars with higher radii (Paladini et al., 2018).

The effect of granulation per se does not prevent a planetary transit from being observed. However, if one aims to derive planetary parameters from the shape of the lightcurve, granulation has a non-negligible effect in the case of G and K spectral class

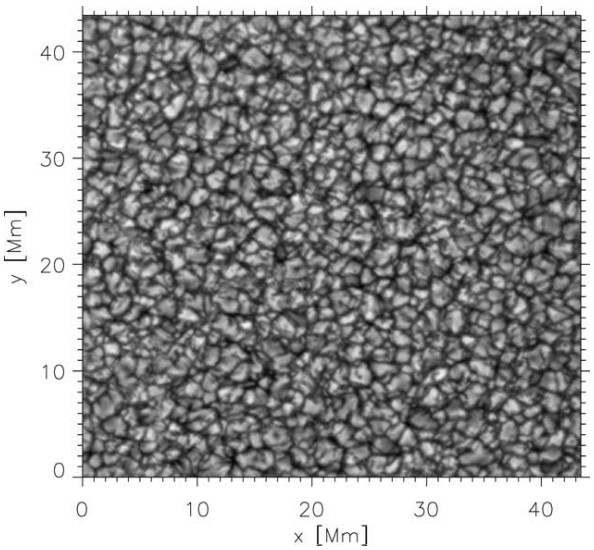

Figure 1.6: Granulation on the surface of the Sun captured with the Swedish Vacuum Solar Telescope. (Image credit Müller et al., 2001)

stars (Chiavassa et al., 2017).

Faculae

Faculae are bright formations in the photosphere that follow a string-like shape (Fig. 1.7). They are produced by magnetic field lines entering the photosphere, yet these field lines have lower magnetic fluxes than those causing starspots. The plasma within faculae is hotter than the temperature of the photosphere. Faculae make a significant contribution to the total solar irradiance per unit area (a quantity also referred to as the solar constant). Their number density, similarly to that of sunspots, follows the solar cycle. The number of faculae peaks around solar maximum. Even though the amount of sunspots is the highest around the solar maximum, the increased amount of bright faculae is able to compensate for the dimming effect of sunspots.

Faculae have typical lifetimes on the order of minutes.

Stellar activity cycles

Similarly to the 11 year long cycle of the Sun, other stars also have been observed to show activity cycles (e.g., Baliunas and Vaughan, 1985). These variations, however, happen on a timescale so long (from years to decades) that for the purposes of this thesis they can safely be ignored.

The typical lifetimes of different phenomena are summarized in Table 1.1.

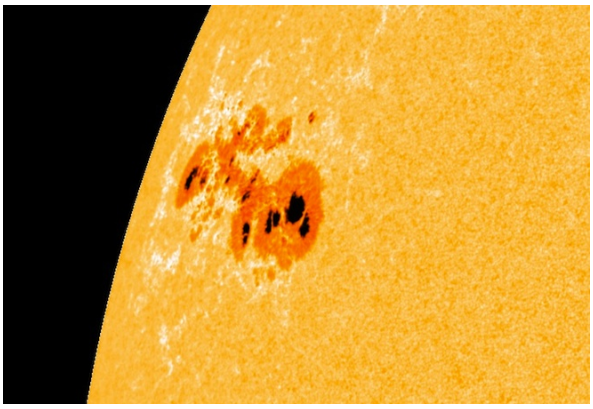

Figure 1.7: Active region at the edge of the Sun's disk. Faculae are the brighter string-like formations around the sunspots (Image credit: NASA/SDO).

Table 1.1: Lifetimes of the most common phenomena responsible for variations in the lightcurves of extrinsic variable stars.

Phenomenon	Lifetime
Starspots	days – years
Flares	minutes
Granulation	$\sim$ 10 minutes
Faculae	minutes
Stellar activity cycles	years – decades

1.4 Aims of this thesis

In this thesis, I examine different models that serve to filter out variations in a star's lightcurve that happen on different timescales than planetary transits.

I do this with the aim of finding an optimal model that removes the effects of stellar variability whose source is not the presence of a transiting exoplanet. These models are then assessed based on how strong the planetary transit's signal is after they are applied to the original observations, and how closely they follow the observations.

In Chapter 2, I investigate what models are generally used to describe stellar lightcurves. Since observations contain some variations that change from one measurement to the next, I include autoregression models that have the possibility to account for these random fluctuations. Subsequently, I take a look at how the lightcurve changes when a planet transits its star.

Once these models are introduced, Chapter 3 explains the different methods to fit these models to observed data (least squares method, Monte Carlo Markov chains). The chapter then shows the equipment to be used to assess the efficiency of the fitted

models in describing the observed data (periodograms and information criteria).

All models and methods described in previous chapters are put into practice in Chapter 4. In this chapter, I use the variability models on the lightcurve of the star HD 110082, observed by TESS space telescope. The star has one confirmed transiting planet, roughly the size of Neptune, orbiting it every 10.18 days. The transit signal from the planet is one order of magnitude weaker than the brightness variations of the star itself. After finding the stellar lightcurve model that leaves us the strongest planetary signal, I characterize the planet based on the lightcurves of its transit.

The thesis ends with conclusions in Chapter 5.

2. Modeling stellar lightcurves

2.1 Sinusoid model

One of the goals of this thesis is to find models to fit the large scale variations of stars that happen on the timescales of days. As flares have typically much shorter, and starspot evolution has typically longer timescales than that, the primary phenomena to be modeled are those arising from the rotation of the stars. As a consequence of the star's rotation, the observer looks at the surface features of the star at continuously changing angles. The most prominent features that have lifetimes that typically exceeding the rotation period of stars and thus introduce a stable, periodic signal into the lightcurves of stars are starspots. These spots disappear and reappear as they rotate around the axis of the star.

It is for these reasons that the lightcurves of the stars with prominent surface features can be approximated with periodic functions. A common form of this is through modeling the time dependence of the brightness as the sum of some freely chosen, n number of sine waves:

$$f(t) = C + \sum_{i=1}^{n} A_i \cos\left(2\pi f_i t + \varphi_i\right) \ , \tag{2.1}$$

where C is some constant, A_i is the amplitude, f_i is the frequency and φ_i is the phase of a sine wave. Throughout this thesis, the function $f(t) = C + A\cos(2\pi f t + \varphi)$ is referred to as a sinusoid.

In most cases, the lightcurve will not follow the form presented in (2.1), as the observations inevitably contain some excess noise. In this case, an observation m_i has an additional noise component ϵ_i. This ϵ_i is different for every observation and is most usually modeled as having a Gaussian distribution, with a mean $\mu = 0$ and variance σ^2. These Gaussian random numbers are independent and identically distributed (i.i.d.). The common notation for random variables ϵ_i following a Gaussian distribution with given μ and σ^2 values is $\epsilon_i \sim \mathcal{N}(\mu, \sigma^2)$.

A commonly used statistic to describe the goodness of fit of a model (including the variance of the noise) to the observed data is the likelihood l. For a single observation,

it is calculated as

$$l_i = [2\pi(\sigma_i^2 + \sigma^2)]^{-1/2} \exp\left[-\frac{1}{2}\frac{(m_i - f(t))^2}{\sigma_i^2 + \sigma^2}\right] \tag{2.2}$$

where m_i denotes the observed value, $f(t)$ is the value predicted by the model, σ_i^2 is the known instrumental noise of the observations and σ^2 is the variance of the excess noise of the model. The likelihood of the whole fit is given as the product of the likelihoods of individual observations

$$l = \prod_i l_i \,. \tag{2.3}$$

The calculation of likelihood l can lead to values so high that they impose computational problems. It is a common approach to use the logarithm of the likelihood, or log likelihood, $\log l$ instead. The log likelihood of individual observations is

$$\log l_i = -\frac{1}{2}\left[\log(2\pi) + \log(\sigma_i^2 + \sigma^2) + \frac{(m_i - f(t))^2}{\sigma_i^2 + \sigma^2}\right] . \tag{2.4}$$

The log likelihood of the whole fit is the sum of the log likelihoods of the individual observations

$$\log l = \sum_i \log l_i \,. \tag{2.5}$$

The lightcurve of the star may be divided into intervals on the scale of couple of days, to account for spot evolution and the differential rotation of the star, which changes the positions of starspots relative to each other. Differential rotation is also responsible for the need for sinusoids with different frequencies, describing the lightcurve of a single rotating body.

2.2 Autoregressive and moving average models

A sinusoidal fit might not be able to follow the lightcurve of the star so closely, that when the brightness variations are removed by the fit, the transits become visible. A way to compensate for the noise above the sinusoidal variations is presented by ARMA models. The word ARMA is an acronym, based on the two components of the model, the autoregressive (AR) and moving average (MA) models. These models describe an observation as a function of previous observations (autoregressive models) or their random deviations from the mean of the observations (moving average models).

2.2.1 Autoregressive model

Autoregressive process models operate with autoregression, i.e., the assumption that there is a correlation between an observation and the ones preceding it. Using this

assumption, one can give a prediction for a data point based on previous data points and account for correlated noise in which the value of the measurement depends on the previous ones.

The most simple autoregressive process model, the first-order autoregression, denoted as AR(1) assumes a linear correlation between an observation, m_i and the one preceding it, m_{i-1}. It fulfills the equation

$$m_i = c + \varphi\, m_{i-1} + \epsilon_i\,. \tag{2.6}$$

where c is some constant, φ is the parameter of the model, which quantifies the dependence of m_i on the previous observed value m_{i-1}, and ϵ_i are i.i.d. Gaussian random numbers with standard deviation σ ($\epsilon_i \sim \mathcal{N}(0, \sigma^2)$).

The general form of autoregressive process models is a pth order autoregression, also noted as AR(p). In this model, the value of a measurement depends on the values of the preceding p measurements. An AR(p) process satisfies the equation

$$m_i = c + \sum_{j=1}^{p} \varphi_j\, m_{i-j} + \epsilon_i\,. \tag{2.7}$$

The observations might not be evenly spaced in time, and so it might be useful to introduce an exponential smoothing to the model, where the dependence of m_i on the previous measurements decays exponentially over time:

$$m_i = c + \sum_{j=1}^{p} \varphi_j \exp\left(\frac{t_i - t_{i-j}}{\tau}\right) m_{i-j} + \epsilon_i\,. \tag{2.8}$$

Here, t_i is the time at which observation m_i is taken and τ is the parameter that determines the timescale of exponential smoothing. This guarantees that the dependence of an observation on a previous one becomes negligible if they are separated by a long time interval $t_i - t_{i-j}$.

The parameters $c, \varphi_1, ..., \varphi_p, \sigma$ can be estimated with Monte Carlo methods (see Section 3.2). For this, one needs to take the likelihood function or alternatively, the log likelihood function of the autoregressive model. Since in autoregressive models, the individual measurements are dependent on each other, the joint likelihood of the observations is conditional, such that

$$l(m_1, ..., m_N) = l(m_1) \times l(m_2|m_1) \times ... \times l(m_N|m_{N-1}, ..., m_{N-p})\,, \tag{2.9}$$

where N is the total number of observations. This means that the likelihood function of the AR(p) model can be given recursively. The log likelihood of the first observation in the AR(p) model is

$$\log l(m_1) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{(m_i - c)^2}{2\sigma^2}\,, \tag{2.10}$$

the log likelihood of the second observation in the AR(1) model is

$$\log l(m_2) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{(m_i - c - \varphi m_1)^2}{2\sigma^2}\,, \tag{2.11}$$

similarly, the likelihood of the ith observation in the AR(1) model is

$$\log l(m_i) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{(m_i - c - \varphi m_{i-1})^2}{2\sigma^2}\,. \tag{2.12}$$

The log likelihood of the ith observation of an AR(p) model is

$$\log l(m_i) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{(m_i - c - \sum_{j=1}^{p}\varphi_j m_{i-j})^2}{2\sigma^2}\,, \tag{2.13}$$

given that $p \leq i$. If $p > i$, the summing in the right hand term is over indices $j = 1, ..., i$.

The log likelihood of the model fit is

$$\log l = \sum_{i=1}^{N} \log l(m_i)\,. \tag{2.14}$$

An artificially generated AR(1) process, fitted using the mean of parameters generated by the Adaptive Metropolis (AM) algorithm (see Section 3.2) is shown on Fig. 2.1.

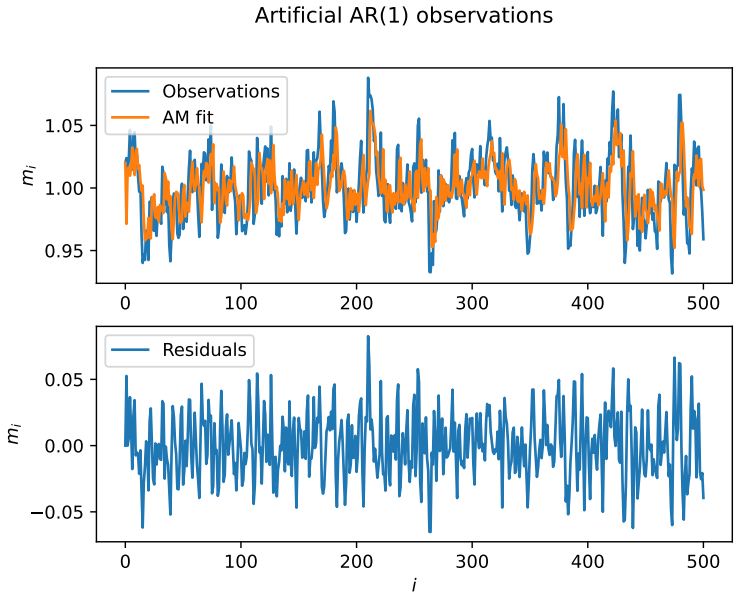

Figure 2.1: An artificially generated AR(1) dataset, and its fit, calculated by generating the parameters describing the process with the Adaptive Metropolis (AM) algorithm (see Section 3.2). The AR(1) process has the parameters $c = 0.3, \varphi = 0.7, \sigma = 0.02$.

2.2.2 Moving average model

Moving average models describe the measurement m_i based on the the random deviations ϵ_i of previous measurements from the mean value of the time series. Based on this, it is possible to give a prediction for an observation by taking into account the deviations of the previous observations. These models can be used to account for the correlated noise from the deviations of previous measurements.

The most simple moving average model is the first order moving average model, denoted as MA(1). This model only uses the random deviation of the previous measurement and satisfies

$$m_i = c + \epsilon_i + \theta\epsilon_{i-1} \tag{2.15}$$

where c is some constant, $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$ that are i.i.d., and θ quantifies the dependence of m_i on the random deviation of the previous measurement.

The general form of moving average models, where the observation m_i depends on the random deviations of the previous q measurements is a qth order moving average model and is denoted as MA(q). The MA(q) model satisfies

$$m_i = c + \epsilon_i + \sum_{j=1}^{q} \theta_j \epsilon_{i-j} \,. \tag{2.16}$$

The joint log likelihood function of the MA models are conditional, similarly to that of the AR models. The log likelihood of an MA(1) model can be given using the following steps. If ϵ_{i-1} is known, then

$$m_i | \epsilon_{i-1} \sim \mathcal{N}(c + \theta\epsilon_{i-1}, \sigma^2) \,. \tag{2.17}$$

Let us assume that $\epsilon_0 = 0$. If m_1 is known, the value of ϵ_1 can be given as $\epsilon_1 = m_1 - c$. If ϵ_1 is known, then ϵ_2 can be calculated as $\epsilon_2 = m_2 - c - \theta\epsilon_1$. Similarly, all ϵ_i values can be calculated recursively, as

$$\epsilon_i = m_i - c - \theta\epsilon_{i-1} \,. \tag{2.18}$$

The log likelihood of the first observation is the same as in (2.10). The log likelihood of the rest of the observations is

$$\log l(m_i) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{\epsilon_i^2}{2\sigma^2} \,. \tag{2.19}$$

To get the log likelihood of the whole model, the log likelihoods of the individual observations are summed up similarly as in (2.14). Exponential decay between the observations can be similarly introduced as for the AR models.

The log likelihood function of a qth order moving average model, MA(q) is

$$\log l = -\frac{N}{2}\log(2\pi\sigma^2) - \sum_{i=1}^{N} \frac{\epsilon_i^2}{2\sigma^2} \,, \tag{2.20}$$

where if $i = 1$, $\epsilon_i = m_1 - c$, and if $i > 1$,

$$\epsilon_i = m_i - c - \sum_{j=1}^{q} \theta_j \epsilon_{i-j} \tag{2.21}$$

given that $q \leq i$. If $q > i$, the summing in the right hand term is over indices $j = 1, ..., i$.

2.2.3 ARMA models

The values and deviations from the mean of previous measurements can be used together to predict the value of future measurements. The models that take into account both types of information are called ARMA models. These models account for the correlated noise in the observation introduced by the dependence on both the previous observation and its random deviation from the data mean. The combination of a pth order autoregressive AR(p), and a qth order moving average model is denoted as ARMA(p, q). The most simple form of an ARMA model is the ARMA(1, 1) model, for an observation m_i which satisfies

$$m_i = c + \epsilon_i + \varphi m_{i-1} + \theta \epsilon_{i-1} \,. \tag{2.22}$$

For an observation m_i, the ARMA(p, q) model fulfills

$$m_i = c + \epsilon_i + \sum_{j=1}^{p} \varphi_j m_{i-j} + \sum_{j=1}^{q} \theta_j \epsilon_{i-j} \,. \tag{2.23}$$

The log likelihood of observation of the first observation in the ARMA models is the same as in (2.10). The likelihood of the rest of the observations in an ARMA(1,1) model is

$$\log l(m_i) = -\frac{1}{2} \log(2\pi\sigma^2) - \frac{(m_i - c - \varphi m_{i-1} - \theta \epsilon_{i-1})^2}{2\sigma^2} \,, \tag{2.24}$$

where ϵ can be given recursively, assuming that $\epsilon_0 = 0$ and $\epsilon_1 = m_1 - c$:

$$\epsilon_i = m_i - c - \varphi m_{i-1} - \theta \epsilon_{i-1} \,. \tag{2.25}$$

The log likelihood of the rest of the observations in an ARMA(p, q) is

$$\log l(m_i) = -\frac{1}{2} \log(2\pi\sigma^2) - \frac{\epsilon_i^2}{2\sigma^2} \,, \tag{2.26}$$

where

$$\epsilon_i = m_i - c - \sum_{j=1}^{p} \varphi_j m_{i-j} - \sum_{j=1}^{q} \theta_j \epsilon_{i-j} \,. \tag{2.27}$$

A practical set of initial values for the iteration in (2.27) for $i = 1, 2, ..., N$ are $y_s = c/(1 - \sum_{i=1}^{p} \varphi_i)$ for $s = 0, -1, ..., -p + 1$ and $\epsilon_s = 0$ for $s = 0, -1, ..., -q + 1$ (e.g.,

Hamilton, 1994, Chapter 5). The log likelihood of the whole model is given similarly as in (2.14).

Similarly to the AR(p) and MA(q) models, exponential smoothing can also be applied here, in order to account for the uneven sampling intervals.

In all the above equations, the constant c can be replaced by the values of a function $f(t)$. This is particularly useful when dealing with signals that can be fit with a simple function (e.g., a sinusoid, when modeling rotating starspots or stellar radial velocity data in order to detect exoplanets).

A considerable disadvantage of the models described in this section in general, is that they can adapt rather easily to the variations they are supposed to describe. In other words, they can be almost as effective at filtering out the signal as filtering out the noise, which they are supposed to target.

2.3 Transit models

Transit models make predictions about the shape of the lightcurve when the exoplanet transits the star. They make these predictions based on the parameters of the planet and the star.

If the stellar, planetary, and orbital parameters are chosen correctly, a transit model is able to provide a fit for the lightcurve when the transit happens. From this point on, this part of the lightcurve is referred to as a transit lightcurve.

Such models have been proposed by Mandel and Agol (2002) and Giménez (2007) who used analytical formulae to model the amount of incoming stellar flux during the transit. These models give a highly accurate lightcurve, accounting for limb darkening with highly accurate quadratic or non-linear formulae. However, when dealing with highly noisy data, the nuances in the lightcurve caused by the uneven surface brightness of the star are negligible. In this case, it is often sufficient to model the transit lightcurve using a linear approximation.

2.3.1 Linear approximation

One model using a linear approximation for the shape of the transit lightcurve was introduced by Carter et al. (2008). The model assumes a star of radius R_* with a uniform surface flux f_0, and a planet with radius R_p, orbital period P, semi-major axis a, eccentricity e, argument of periapsis ω, and inclination i. The ratio of the radii of the two bodies is denoted as $r = R_p/R_*$ The amount of incoming flux from the star

$F^l(t)$ changes as

$$F^l(t) = \begin{cases} f_0 - \delta, & |t - t_c| \leq T/2 - \tau/2, \\ f_0 - \delta + (\delta/\tau) \times (|t - t_c| - T/2 + \tau/2), & T/2 - \tau/2 < |t - t_c| < T/2 + \tau/2, \\ f_0, & |t - t_c| \geq T/2 + \tau/2\,. \end{cases} \tag{2.28}$$

where t_c is the mid-transit time of the planet. The amount of relative dimming caused by the planet, i.e., the transit depth is given as

$$\delta = f_0 r^2 = f_0 \left(\frac{R_p}{R_*}\right)\,. \tag{2.29}$$

The difference between t and t_c is compared to parameters T and τ, which determine if the planet is fully or partially eclipsing the star (ingressing or egressing the stellar disk) or does not eclipse it at all. These parameters are defined as

$$T = 2\tau_0\sqrt{1 - b^2}\,, \tag{2.30}$$

$$\tau = 2\tau_0 \frac{r}{\sqrt{1 - b^2}}\,. \tag{2.31}$$

Here b is the impact parameter, which is proportional to the angular distance between the center of the planet and the center of the star in the middle of the transit:

$$b = \frac{a\cos i}{R_*}\left(\frac{1 - e^2}{1 + e\sin\omega}\right)\,. \tag{2.32}$$

The impact parameter can change between 0 (corresponding to a transit in which the planet's center passes in front of the star's center) and 1 (corresponding to a transit where the edge of the planet touches the edge of the star).

The parameter τ_0 is defined as

$$\tau_0 = \frac{R_*}{an}\left(\frac{\sqrt{1 - e^2}}{1 + e\sin\omega}\right)\,, \tag{2.33}$$

where $n = 2\pi/P$ is the mean motion of the planet.

2.3.2 Stellar limb darkening

The uneven surface brightness distribution of stars has to be taken into account when fitting highly accurate transit data. The surface brightness distribution of a star can be approximated by a linear limb-darkening law (Milne, 1921), according to which the intensity of the stellar surface changes as

$$\frac{I(\mu)}{I(0)} = 1 - u\sum_{i=1}^{N} A_i(1 - \mu)^i \tag{2.34}$$

where $\mu = \sqrt{1 - z^2}$, A is the flux at the center of the stellar disk, u is the linear limb darkening coefficient, N is the order of linear limb darkening* and $0 \leq z \leq 1$ is the normalized radial coordinate on the disk of the star (Mandel and Agol, 2002). For a circular orbit, z can be given as

$$z(t) = aR_*^{-1}\sqrt{[\sin n(t - t_c)]^2 + [\cos i \cos n(t - t_c)]^2} \tag{2.35}$$

(e.g., Carter et al., 2008). A special case, the first order linear limb darkening arises by choosing $N = 1$.

In order to take into account the changes in the linear approximation of the transit lightcurve, introduced by the limb darkening, the transit depth δ needs to be multiplied by $I(\mu)$ at every point where the model lightcurve $F^l(t)$ is evaluated.

*An alternative, commonly used description is $\mu = \cos\gamma$, where γ is the angle between the line of sight and the direction of the emerging flux from the star.

3. Computational methods

3.1 Least squares method

The least squares method aims at finding the parameters of a model that give the best fit for some observed data.

Let us denote the individual observations of the independent variables of a dataset with $x_i,\ i = 1, ..., n$, and the individual observations of the dependent variable with $y_i,\ i = 1, ..., n$. Let us assume that the function $f(x_i, \boldsymbol{\beta})$, where $\boldsymbol{\beta} = (\beta_1, ..., \beta_k)$ is a vector of parameters, describes the dependence between the independent and dependent variables. The residual, i.e., the difference between the observation of the dependent variable and the value predicted by the function $f(x_i, \boldsymbol{\beta})$ at x_i is denoted as $r_i = y_i - f(x_i, \boldsymbol{\beta})$.

The goal of least squares problems is finding the parameters $\boldsymbol{\beta}$ by minimizing the sum of squared residuals

$$S = \sum_{i=1}^{n} r_i^2 \,. \tag{3.1}$$

To minimize S, the parameters $\boldsymbol{\beta}$ need to satisfy

$$\frac{\partial S}{\partial \beta_i} = 0,\ i = 1, ..., k \,. \tag{3.2}$$

In other words, the least squares problem is solved, when the gradient of the sum of squared residuals is equal to zero.

Least squares problems are divided into linear and nonlinear least squares problems, depending on whether the correlation between the independent and dependent variables is linear or not. For the purposes of this thesis, we only need to familiarize ourselves with linear least squares problems.

3.1.1 Linear model estimation

Linear least squares estimation can be used when the dependent variables can be described as some linear combination of the independent variables.

Let us denote the vector containing the ith observation of the independent variables as $\boldsymbol{x}_i = (x_{i1}, ..., x_{ik})$. These vectors can be collected to a $n \times k$ data matrix

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & \cdots & x_{1k} \\ & \ddots & \\ x_{n1} & \cdots & x_{nk} \end{bmatrix} . \tag{3.3}$$

Let us further denote the vector containing the observed dependent variables as $\boldsymbol{y} = (y_1, ..., y_n)$. In linear models, the dependent variables can be given as

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} , \tag{3.4}$$

and a single observation of the dependent variable as

$$y_i = \boldsymbol{x_i} \cdot \boldsymbol{\beta} = \beta_1 x_{i1} + ... + \beta_k x_{ik} . \tag{3.5}$$

The sum of squared residuals in linear models is

$$S = \sum_{i=1}^{n} (y_i - \beta_1 x_{i1} - ... - \beta_k x_{ik})^2 = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^{\mathrm{T}} (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) = \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 . \tag{3.6}$$

The most common way to give an estimate $\hat{\boldsymbol{\beta}}$ for the parameters $\boldsymbol{\beta}$ is the ordinary least squares method. This minimizes S through the solution

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y} . \tag{3.7}$$

For the purposes of this thesis, the most important least squares fitting problem to be performed is one involving the sum of sinusoidal functions. A sinusoid $f(t) = C + A\cos(\omega t + \varphi)$ can be fitted using linear least squares fitting, by writing it up as the sum of two trigonometric functions:

$$C + A\cos(\omega t + \varphi) = C + A_1 \cos(2\pi f t) + A_2 \sin(2\pi f t) . \tag{3.8}$$

Here, C is some constant and $A = \sqrt{A_1^2 + A_2^2}$. The vector containing the ith observation of independent variables becomes then $\boldsymbol{x}_i = (1, \cos(2\pi f, t_i), \sin(2\pi f, t_i))$. The parameter vector is $\boldsymbol{\beta} = (C, A_1, A_2)$. The parameter estimate can then be given as in (3.7).

If the frequency f is unknown, then a round of iteration is needed on top of the fitting routine described above. This iteration goes over different test frequencies. The best estimate for f is given by the f value used fitting round with the lowest sum of squared residuals S.

When the signal is a sum of sinusoidal functions, it can be described in the form

$$f(t) = C + \sum_{i=1}^{n} A_{i,1} \cos(2\pi f_i t) + A_{i,2} \sin(2\pi f_i t) . \tag{3.9}$$

The parameters $\boldsymbol{\beta} = (C, A_{1,1}, A_{1,2}, ..., A_{n,1}, A_{n,2})$ can be determined with linear least squares fitting. In this case, the vector of the ith observation of independent variables is $\boldsymbol{x}_i = (1, \cos(2\pi f_1 t_i), \sin(2\pi f_1 t_i), ..., \cos(2\pi f_n t_i), \sin(2\pi f_n t_i))$. For this fitting it is necessary that all frequencies be unequal, $f_1 \neq f_2 \neq ... \neq f_n$.

3.2 Markov chain Monte Carlo methods

Markov chain Monte Carlo (MCMC) methods are algorithms that sample probability distributions using computer generated random numbers.

In such algorithms, the sampling is done by constructing so-called Markov chains. Markov chains are discrete, random processes that satisfy the Markov property. The Markov property, without going into any details, means that a system's future state can be determined based solely on its current state. In other words, knowing the previous states of the system, besides its current one, does not improve the predictions about its future state.

By taking samples from these Markov chains, one can access information about the probability distribution, expected value, or variance of a variable.

Several instances of MCMC methods have been formulated. One of the most commonly used one of these is the Metropolis-Hastings algorithm (Hastings, 1970). This algorithm generates Markov chains by comparing the proposal distribution's value at the randomly generated points and by rejecting some of the newly proposed points.

For the purposes of this thesis, I am using an updated form of the Metropolis-Hastings algorithm, the Adaptive Metropolis (AM) algorithm, introduced by Haario et al. (2001).

3.2.1 Adaptive Metropolis algorithm

Let the target distribution to be sampled by the AM algorithm be denoted by π.

For every iteration t, a random point $Y \in \mathbb{R}^d$ is generated from a Gaussian proposal distribution $q_t(\cdot|X_0, ..., X_{t-1})$, where $X_0, ..., X_{t-1} \in \mathbb{R}^d$ are the previous values of the chain, with d being the number of dimensions of the generated points. The mean of this distribution is the last generated point X_{t-1}. The way the covariance matrix of the distribution, C_t is calculated depends on whether or not the number of iterations exceeds a certain index, $t_0 > 0$ (the choice of which is free, but a higher t_0 leads to the adaptation taking effect later):

$$C_t = \begin{cases} C_0, & t \leq t_0, \\ s_d \operatorname{cov}(X_0, ..., X_{t-1}) + s_d \varepsilon I_d, & t > t_0. \end{cases} \tag{3.10}$$

Here C_0 indicates a strictly positive definite, initial covariance matrix, which may be chosen arbitrarily, according to our best prior knowledge of the target distribution,* s_d is a scaling parameter, which is chosen as $s_d = (2.4)^d/f$ (from Gelman et al., 1996), ε is a constant that we may choose to be a very small positive number, I_d is a d dimensional identity matrix, and $\mathrm{cov}(X_0, ..., X_{t-1})$ is the covariance matrix of points $X_0, ..., X_{t-1}$.

The empirical covariance matrix of points $x_0, ..., x_k \in \mathbb{R}^d$ is

$$\mathrm{cov}(x_0, ..., x_k) = \frac{1}{k}\left(\sum_{i=0}^{k} x_i x_i^{\mathrm{T}} - (k+1)\bar{x}_k \bar{x}_k^{\mathrm{T}}\right) \tag{3.11}$$

where

$$\bar{x}_k = \frac{1}{k+1}\sum_{i=0}^{k} x_i \tag{3.12}$$

and the elements $x_i \in \mathbb{R}^d$ are column vectors.

The key difference between the AM algorithm, and previously proposed MCMC algorithms is in the treatment of previously generated points. From (3.10), it is visible that the proposal distribution is updated by the knowledge of the target distribution accumulated in previous iterations. If the covariance matrix would not be updated, we would deal with a regular Metropolis-Hastings algorithm.

The recursive formula for the covariance

$$C_{t+1} = \frac{t-1}{t} C_t + \frac{s_d}{t}\left[t\,\bar{X}_{t-1}\,\bar{X}_{t-1}^{\mathrm{T}} - (t+1)\,\bar{X}_t\,\bar{X}_t^{\mathrm{T}} + X_t\,X_t^{\mathrm{T}} + \varepsilon\, I_d\right] \tag{3.13}$$

can be used for iterations $t \geq t_0 + 1$, which alleviates the computational cost of the algorithm.

The freshly generated point Y is accepted with probability

$$\alpha(X_{t-1}, Y) = \min\left(1, \frac{\pi(Y)}{\pi(X_{t-1})}\right) \tag{3.14}$$

in which case the point returned in the current iteration is $X_t = Y$, otherwise it remains that of the value of the previous iteration $X_t = X_{t-1}$.

The probability distribution of points X_t generated by the algorithm follows the target distribution π. This is demonstrated on Fig. 3.1, which shows AM algorithm generated points whose target distribution is the reciprocal Rosenbrock function

$$f(x, y) = [(a - x)^2 + b(y - x^2)^2]^{-1}\,, \tag{3.15}$$

with parameters $a = 1$ and $b = 100$.

*I chose the initial covariance matrix to be the identity matrix, multiplied by some small number n, i.e., $C_0 = n\,I_d$, with n typically ranging between 10^{-6} and 10^{-9}. Other initial covariance matrices may be freely chosen.

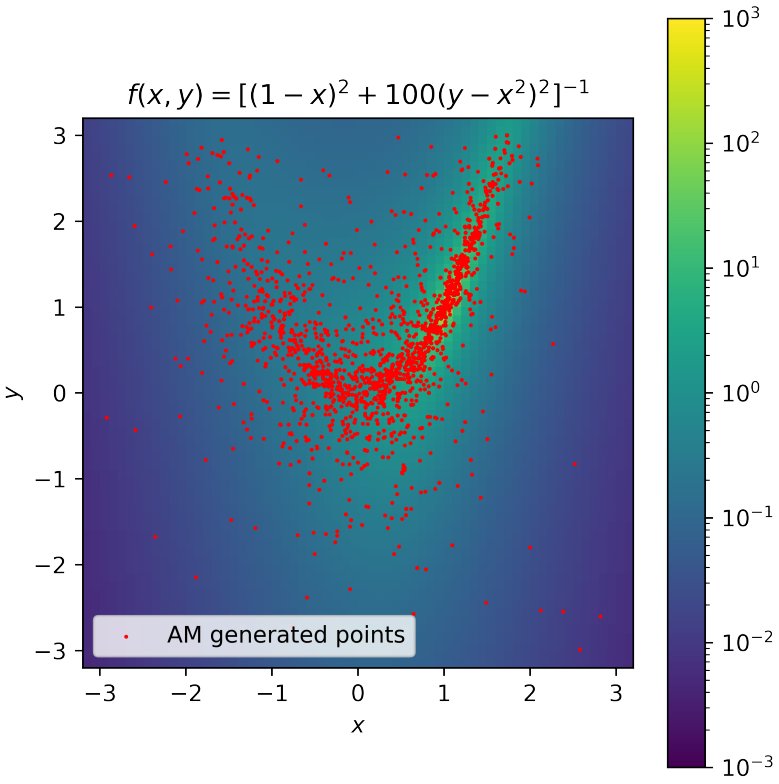

Figure 3.1: AM algorithm generated random points sampling the reciprocal Rosenbrock function $f(x, y)$. The algorithm is more likely to generate a point at such (x, y) coordinates, where the function takes a higher value.

3.2.2 Efficiency of Markov chains

The convergence of Markov chains cannot ever be completely proven. Therefore, one needs to perform tests proving the non-convergence of Markov Chains. If the tests of non-convergence fail, they suggest (but not entirely prove) that a Markov Chain is convergent (Chen et al., 2002).

One such test for non-convergence is the German-Rubin statistic $\hat{R}$ (Gelman and Rubin, 1992). This method relies on the generation of several Markov chains. A short description of this statistics, after Ford (2006) is presented below.

The Gelman-Rubin statistics compares the variance of any quantity $z(\theta)$, generated by individual Markov chains with the variance of $z(\theta)$ values from all the chains combined. Here θ marks the parameter to be estimated by the Markov chain.

Let the z_{ic} stand for the parameter z generated in the ith iteration of the cth Markov chain. Let the number of Markov chains be N_c, all with length L_c. The mean value of z_{ic} of an individual Markov chain with index c can be estimated as

$$\bar{z}_{\cdot c} = \frac{1}{L_c} \sum_{i=1}^{L_c} z_{ic} \,. \tag{3.16}$$

The average of the variances of z_{ic} within each chain can be estimated

$$W(z) = \frac{1}{N_c} \sum_{c=1}^{N_c} \frac{1}{L_c - 1} \sum_{i=1}^{L_c} (z_{ic} - \bar{z}_{\cdot c})^2 \,. \tag{3.17}$$

The mean value of z_{ic} from all the Markov chains can be estimated as

$$\bar{z}_{\cdot\cdot} = \frac{1}{N_c} \sum_{c=1}^{N_c} \bar{z}_c = \frac{1}{L_c N_c} \sum_{c=1}^{N_c} \sum_{i=1}^{L_c} z_{ic} \tag{3.18}$$

and the variance of the estimates of a single chain as

$$B(z) = \frac{L_c}{N_c - 1} \sum_{c=1}^{N_c} (\bar{z}_{\cdot c} - \bar{z}_{\cdot\cdot})^2 \tag{3.19}$$

The variance of all z_{ic} values can then estimated as a weighted average of $W(z)$ and $B(z)$

$$\widehat{\mathrm{var}^+}(z) = \frac{L_c - 1}{L_c} W(z) + \frac{1}{L_c} B(z) \,. \tag{3.20}$$

This is an unbiased estimator of the variance of z_{ic} if the initial states of the Markov chains were selected from the target distribution or in the limit $L \to \infty$. The Gelman-Rubin statistics $\hat{R}(z)$ can then be given as

$$\hat{R}(z) = \sqrt{\frac{\widehat{\mathrm{var}^+}(z)}{W(z)}} \tag{3.21}$$

As the Markov chains continue to show a lack of non-convergence, the value of $\hat{R}(z)$ approaches 1. If $\hat{R}(z)$ gets sufficiently close to 1, one can conclude that there is no evidence for the non-convergence of the Markov chains. The value that $\hat{R}(z)$ needs to reach for all parameters θ before the chain can be used to infer these parameters is up for choice. Ford (2005) required that $\hat{R}(z)$ for all parameters, before using the chains for inference.

The Gelman-Rubin statistics of five different example Markov chains estimating parameter z is shown on Fig. 3.2.

3.3 Periodograms

Periodograms are used in order to find and characterize the relative strengths of periodic signals in a time series dataset.

In this section, three periodograms are presented. The first two, the Lomb-Scargle and the likelihood ratio periodograms are most capable at identifying signals that show a sinusoidal time variation.

The third type of periodograms, the box fitting periodogram was designed specifically to identify the periodic variations caused by the transits of planets in front of the stellar disk.

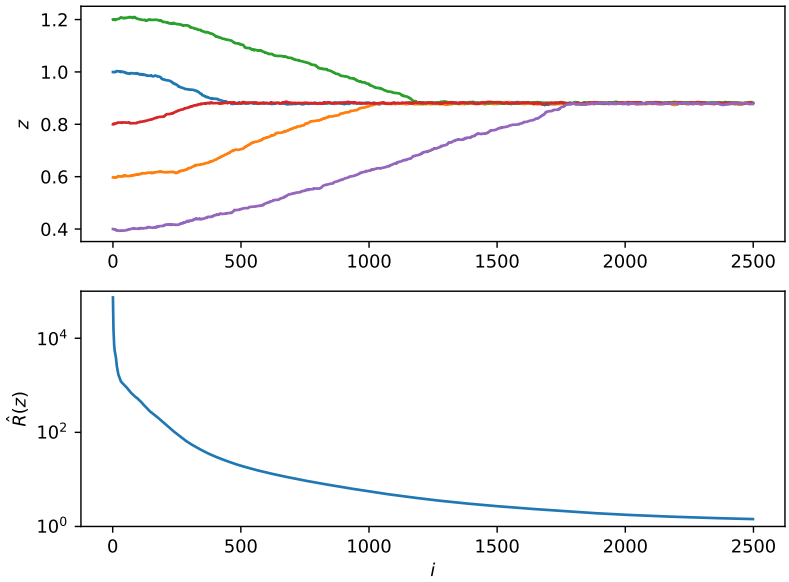

Figure 3.2: Five different Markov chains estimating parameter z (above) and their corresponding Gelman-Rubin statistics $\hat{R}(z)$ for every iteration i (below). As the variance of the generated z values of all chains approaches the mean of variances from individual chains, the Gelman-Rubin statistics converges towards 1.

3.3.1 Lomb-Scargle periodogram

One variation of periodograms, widely used within astronomy, the Lomb-Scargle periodogram, $P_{\mathrm{LS}}(f)$ has the strength of being able to be used for observations with uneven sampling (Lomb, 1976, Scargle, 1982). For observations $y_n,\ n = 1, 2, ..., N$ taken at times $t_n,\ n = 1, 2, ..., N$, the periodogram for frequencies f can be given as

$$\begin{aligned} P_{\mathrm{LS}}(f) =& \frac{1}{2}\left\{\left(\sum_n y_n \cos(2\pi f[t_n-\tau])\right)^2 / \sum_n \cos^2(2\pi f[t_n-\tau])\right. \\ &\left. + \left(\sum_n y_n \sin(2\pi f[t_n-\tau])\right)^2 / \sum_n \sin^2(2\pi f[t_n-\tau])\right\} \end{aligned} \tag{3.22}$$

where τ is calculated for each f to keep timeshift invariance:

$$\tau = \frac{1}{4\pi f}\tan^{-1}\left(\frac{\sum_n \sin(4\pi f t_n)}{\sum_n \cos(4\pi f t_n)}\right). \tag{3.23}$$

To obtain the periodogram $P_{\mathrm{LS}}(f)$, a signal (in our case the stellar flux at different times) must be sampled at different frequencies f. The minimum difference between the frequencies at which the signal should be sampled (also known as the Nyquist sampling limit) is $f_{\mathrm{Ny}} = 2B$. Here B is the bandwidth of the signal that is to say, the difference between the maximum and the minimum frequencies.

Frequencies can be converted using the $P = 1/f$ relation to find the periods of the signals. This is useful, as the period with the highest P_{LS} value might correspond to the rotation period of the star.

An example of the Lomb-Scargle periodogram is shown on Fig. 3.3. The figure shows an artificially generated time series, along with its Lomb-Scargle periodogram.

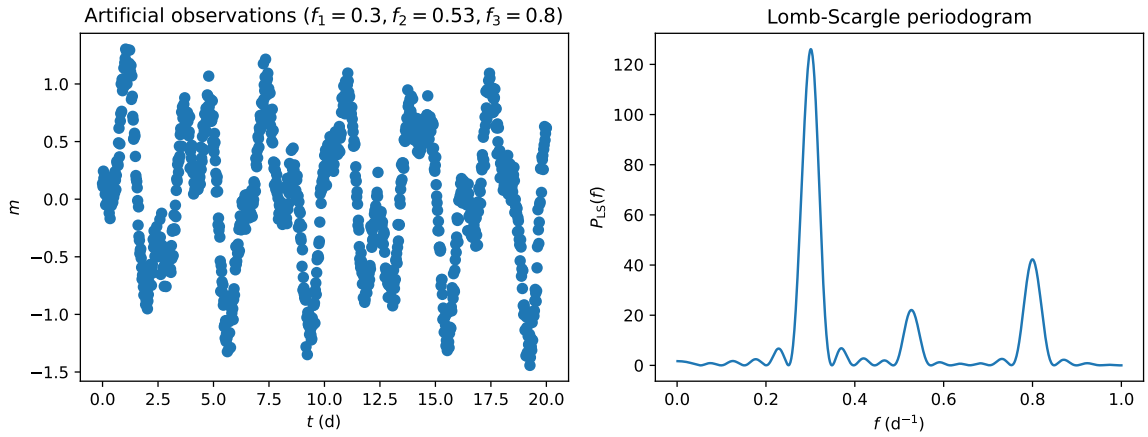

Figure 3.3: An example of the Lomb-Scargle periodogram on artificial data. *Left:* Artificially generated observations, obtained as the sum of three sinusoidal signals, with frequencies $f_1 = 0.3$ d^{-1}, $f_2 = 0.53$ d^{-1}, $f_3 = 0.8$ d^{-1}, and amplitudes $A_1 = 0.7, A_2 = 0.25, A_3 = 0.4$, and an additional Gaussian noise with $\mu = 0, \sigma = 0.1$. *Right:* The Lomb-Scargle periodogram of the artificial observations. The peaks correspond to the frequencies of signals in the observations. The heights of the peaks are proportional to the signal strengths at given frequencies.

3.3.2 Likelihood ratio periodogram

An alternative method of assessing the importance of different periodic signals in the lightcurve is the likelihood ratio periodogram (Anglada-Escudé and Tuomi, 2012, Tuomi et al., 2018).

The basis of this algorithm is the comparison of least-squares-fit sinusoid models with different frequencies to a reference model. The reference model is a constant signal with the value of the mean of the observed signals $\bar{m}$.

For every frequency f, a least squares fitting of a sinusoid

$$f(t) = A\cos(2\pi ft + \varphi) \tag{3.24}$$

is performed with the amplitude A and phase φ being free parameters. The relative power of the signal with period f is then equal to the likelihood ratio r of the best fit sinusoid versus the reference model.

Wilks' theorem (Wilks, 1938) states that the statistics resulting from the logarithm of the likelihood ratio $\log r$ multiplied by -2 are distributed asymptotically

according to the χ^2 distribution. It is therefore possible to take the logarithm of likelihood ratios, as

$$-2\log r = \chi^2_{\rm mod} - \chi^2_{\rm ref} \tag{3.25}$$

where $\chi^2_{\rm mod}$ is the χ^2 statistics of the best fit model with frequency f and $\chi^2_{\rm ref}$ is the χ^2 statistics of the reference model. The value of χ^2 is calculated as

$$\chi^2 = \sum_{i=1}^{N} \frac{(m_i - f(t))^2}{\sigma_i^2}\,, \tag{3.26}$$

where m_i is the ith observation, $f(t)$ is the value predicted by the model at the time of the observation, and σ_i^2 is the error associated with the ith observation.

The advantage of this periodogram over the Lomb-Scargle periodogram is in the ease of assessing the likelihood of a signal at a given frequency. For given degrees of freedom (DF), the model is accepted if the likelihood ratio exceeds the certain thresholds, listed in Table 3.1. The table contains the likelihood ratio statistics to be exceeded, so that it can be stated that the reference model cannot explain the observation with probability P. In case of the periodogram, the number of free parameters is two (amplitude and phase), which results in DF = 2.

Table 3.1: A model with DF degrees of freedom is accepted if its likelihood ratio value r exceeds the values below. The different columns marked with P signify that the probability of the signal being produced by constant stellar activity (the reference model) is $P = 0.1$, $P = 0.05$, $P = 0.01$, $P = 0.001$, respectively.

DF				
1	2.71	3.84	6.63	10.83
2	4.61	5.99	9.21	13.82
3	6.25	7.81	11.34	16.27
4	7.78	9.49	13.28	18.47
5	9.24	11.07	15.09	20.52
6	10.64	12.59	16.81	22.46
7	12.02	14.07	18.48	24.32
8	13.36	15.51	20.09	26.12
9	14.68	16.92	21.67	27.88
10	15.99	18.31	23.21	29.59
P	0.10	0.05	0.01	0.001

A value similar to P is the false alarm possibility (FAP) that indicates how likely is it that the observed signal can be explained by something else than the proposed model.

3.3.3 Box-fitting algorithm

The box-fitting algorithm in its current form has been proposed by Kovács et al. (2002) with the aim of finding the periodic signals of transiting exoplanets in stellar lightcurves. They employ a model which assumes a strictly periodic signal with period P_0 that only takes two discrete values: a higher H value that corresponds to the received flux from the star when the planet is not transiting it, and a lower L value that corresponds to the received flux when the planet is passing in front of the star. The length of the transit is quantified with the fractional transit length q, which gives the ratio between the time spent transiting the star and the orbital period of the planet and is typically assumed to be a small number ($\approx 0.01 - 0.05$). This means that the amount of time that the planet spends in front of the star is qP_0. The mid-transit time is denoted as t_c.

The algorithm aims to find the five free parameters P_0, q, L, H, t_c that produce the best fit model to the observations. The number of free parameters can be reduced to four, by assuming that the average observed flux is zero. This way, one can describe the higher flux value as $H = -Lq/(1-q)$.

Let the observations be denoted as $x_i,\ i = 1, 2, ..., n$. Every observation x_i has an additive zero-mean Gaussian noise with standard deviation σ_i. Each observation is given a weight

$$w_i = \sigma_i^{-2} \left[\sum_{j=1}^{n} \sigma_j^{-2} \right]^{-1} . \tag{3.27}$$

The term $w_i x_i$ is assumed to have a zero arithmetic average.

The algorithm iterates through different trial periods, P_{trial}. For each trial period, the algorithm takes a folded lightcurve. To create a folded lightcurve the timestamps of the observations have to be reset. A modulo operation is performed on the timestamps of the observations with modulus P_{trial}, so that the times of all observations fall between 0 and P_{trial}. The values of the observations stay intact during this conversion. It is advised to divide the data into bins, and perform the following steps on the binned values, to alleviate the computational costs of the algorithm.

For each folded lightcurve of different trial periods P_{trial} different fractional transit lengths are considered (usually $q = 0.01, ..., 0.05$, the number of intermediate transit lengths is freely chosen).

Let the folded observations be denoted as $\tilde{x}_i$ and their corresponding weights as $\tilde{w}_i$. The planet transits the star between times i_1 and i_2.

The algorithm fits a step function to the folded lightcurve, which takes the value $\hat{L}$ between times $[i_1, i_2]$ and $\hat{H}$ otherwise. For each (i_1, i_2), the average square deviation

$$\mathcal{D} = \sum_{i=1}^{i_1-1} \tilde{w}_i \left(\tilde{x}_i - \hat{H}^2\right) + \sum_{i=i_2+1}^{n} \tilde{w}_i \left(\tilde{x}_i - \hat{H}^2\right) + \sum_{i=i_1}^{i_2} \tilde{w}_i \left(\tilde{x} - \hat{L}\right)^2 \tag{3.28}$$

is minimized. $\mathcal{D}$ is minimized, if

$$\hat{L} = \frac{s}{r},\ \hat{H} = -\frac{s}{1-r} \tag{3.29}$$

where

$$\begin{aligned} s &= \sum_{i=i_1}^{i_2} \tilde{w}_i \tilde{x}_i \,, \\ r &= \sum_{i=i_1}^{i_2} \tilde{w}_i \,. \end{aligned} \tag{3.30}$$

Taking advantage of the notation above, the average square deviation can be expressed as

$$\mathcal{D} = \sum_{i=1}^{n} \tilde{w}_i \tilde{x}_i^2 - \frac{s^2}{r(1-r)} \,. \tag{3.31}$$

$\mathcal{D}$ needs to be calculated for all possible (i_1, i_2) pairs separately, to find the absolute minimum of $\mathcal{D}$ for any trial period. The first term in (3.31) is independent of the choice of (i_1, i_2), and so it is sufficient to only take into account the second term of the equation to characterize the quality of the fit. The best fit is provided by the parameters that minimize $\mathcal{D}$ over the whole parameter space.

For any trial period, the confidence of the transit signal for any (i_1, i_2) can be expressed with the log likelihood of the fit

$$\log l = -\frac{1}{2} \left(\sum_{i=1}^{i_1-1} \frac{(\tilde{x}_i - \hat{H})^2}{\sigma_i^2} + \sum_{i=i_2+1}^{n} \frac{(\tilde{x}_i - \hat{H})^2}{\sigma_i^2} + \sum_{i=i_1}^{i_2} \frac{(\tilde{x}_i - \hat{L})^2}{\sigma_i^2} \right) + c \tag{3.32}$$

where

$$c = -\frac{1}{2} \sum_{i=1}^{n} \log(2\pi\sigma_i^2) \,. \tag{3.33}$$

The maximum of $\log l$ values from the iterations over (i_1, i_2) gets accepted as the log likelihood of the trial period. The indices (i_1, i_2) resulting in the maximum $\log l$ correspond to the indices of the minimum of $\mathcal{D}$.

A box fitting periodogram is created by calculating the likelihood ratio r, or log likelihood ratio $\log r$ for all trial periods at a fixed fractional transit length. This results in the use of only three free parameters $(\hat{H}, \hat{L}, i_1)$.

The log likelihood of the box fitting model for different periods and a fixed fractional transit length of HD 110082 is shown on Fig. 3.4.

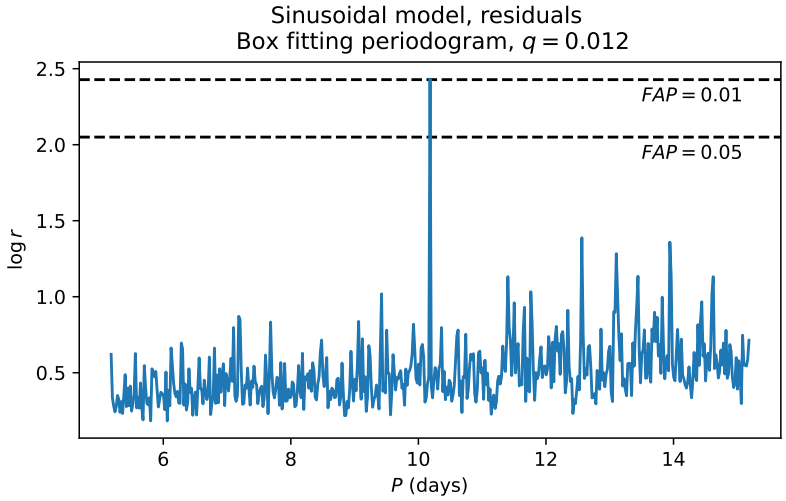

Figure 3.4: Box fitting periodogram of HD 110082 with fixed fractional transit length $q = 0.012$. The dashed lines correspond to a 5% (below) and a 1% (above) probability of the transit signal being falsely identified (false alarm probability). The peak at $P \approx 10.18$ days corresponds closely to the orbital period of HD 110082 b.

3.4 Information criteria and model selection

In order to select the model that gives the best description of the observed data, it is necessary to apply a tool that quantifies the goodness of the individual models. A practical overview of these statistical tools, the model selection criteria have been given by Liddle (2007). They note that the two set of tools that statistics literature contains are either based on information theory (a prime example of this is the Akaike Information Criterion) or Bayesian inference (where a commonly used tool is the Bayesian Information Criterion).

The Akaike Information Criterion (AIC) is defined as

$$\mathrm{AIC} \equiv -2 \log l_{\max} + 2k \tag{3.34}$$

where $l_{\max}$ is the maximum likelihood achievable by the model and k is the number of parameters of the model (Akaike, 1974). Note that the definition uses the maximum of all possible likelihood values and not the likelihood value of an individual fit. This is useful as the mean of the generated parameters of Monte Carlo Markov chains does not necessarily correspond to a model with the highest likelihood, only one close to it.

The best model is the one that minimizes the AIC value. For models that have been sampled with Monte Carlo Markov chains, the maximum likelihood is readily available.

The Bayesian Information Criterion (BIC) is defined as

$$\mathrm{BIC} \equiv -2 \log l_{\max} + k \log N \tag{3.35}$$

where N is the number of datapoints used for the fit (Schwarz, 1978). Similarly to AIC, a lower BIC value means a better model.

4. Results

The methods described in the previous chapters were tested on observations of HD 110082, made with the TESS space telescope. HD 110082 is a star showing periodic variations on a timescale of days. It has one confirmed planet, HD 110082 b on a 10.18 day orbit (Tofflemire et al., 2021).

After the period of the large scale brightness variations was established with the help of periodograms, the variability of HD 110082 was modeled using three different methods:

1. A sum of four sinusoids (Eq. 2.1 with $n = 4$) is fitted on the lightcurve with the least squares method (also referred to as the sinusoidal fit in this Chapter).

2. An ARMA(1,1) model fit is generated by taking the mean of the model parameters sampled with the AM algorithm.

3. A combined version of the ARMA(1,1) and the sinusoidal fit (also referred to as the sinusoids + ARMA(1,1) model).

This chapter contains the description and the assessment of these model fits. It closes with a report on the fitting of the transit lightcurve of the planet.

4.1 TESS lightcurve

HD 110082 is a young, active F8V spectral type star, with one confirmed planet, HD 110082 b.

The planet, HD 110082 b is a sub-Neptune on a 10.1827 day orbit. A list of select stellar and planetary parameters is given in Table 4.1 (Tofflemire et al., 2021).

HD 110082 shows variations on the scale of $\delta \sim 4 \times 10^{-3}$, mostly as a combined result of stellar rotation and starspots. The planet's signal depth is $\delta \sim 6 \times 10^{-4}$.

The star has been observed by the TESS in Sectors 12 (May 21–June 19, 2019), 13 (June 19–July 18, 2019) and 27 (July 4 – July 30, 2020). These observations are shown on Fig. 4.1. The star was a so-called TESS Object of Interest. Such objects are selected before the observations, with the purpose of detecting or confirming possible

Table 4.1: Parameters of HD 110082 and HD 110082 b.

	Parameter	Value
HD 110082	RA (J2000)*	12:50:22.020
	Dec. (J2000)*	−88:07:15.72
	M_* ($M_\odot$)	1.21 ± 0.06
	R_* ($R_\odot$)	1.19 ± 0.06
	$P_{\rm rot}$ (d)	2.34 ± 0.07
	Age (Myr)	250^{+50}_{-70}
HD 110082 b	t_0 (BJD)	$2\,458\,629.909 \pm 0.001$
	$P_{\rm orb}$ (d)	10.18271 ± 0.00004
	a (au)	$0.113^{+0.009}_{-0.013}$
	R_p ($R_\oplus$)	3.2 ± 0.1

planets around them. TESS Objects of Interest get a unique identifier within the TESS mission (in this case, TOI-1098) and are observed with two-minute cadence.

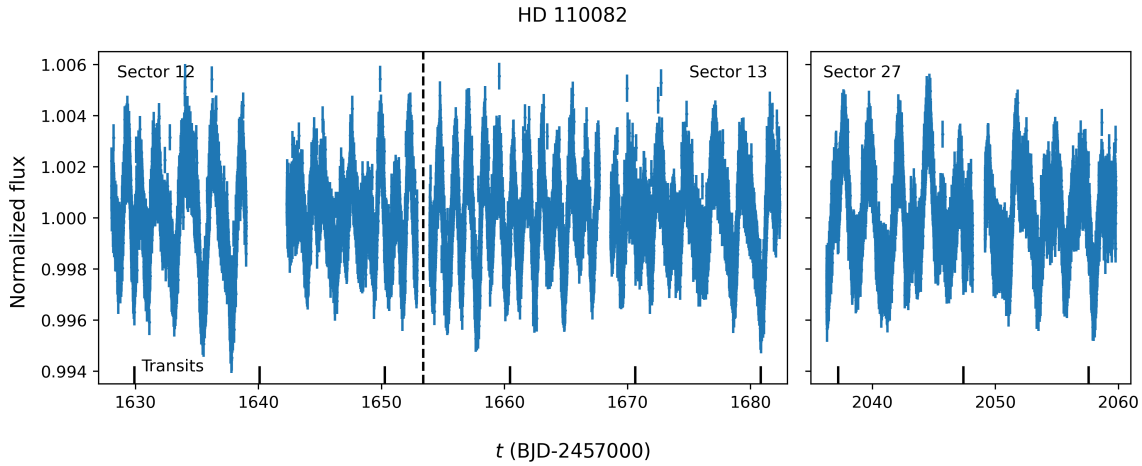

Figure 4.1: TESS lightcurve of HD 110082. The star was observed by the space telescope in three of its observing sectors (Sectors 12 and 13 are shown left, separated with a dashed line, Sector 27 is shown right). The times of transit events are marked with black lines.

A total of eight transits are to be seen in the recorded lightcurve of HD 110082. These transits are not necessarily consecutive ones. The transit with center time t_0 = BJD-2 458 629.909 is henceforward referred to as Transit 1. Transits with center times $t_c = t_0 + n \times P_{\rm orb}$, where $P_{\rm orb}$ stands of the orbital period of the planet are refereed to as Transit n.

*Based on Gaia DR2 astrometry (Gaia Collaboration, 2018).

4.2 Period search

Time series data from HD 110082 was first analyzed by the Lomb-Scargle and likelihood ratio periodograms. The three TESS observation sectors were analyzed separately. The likelihood ratio periodograms for the three sectors are shown on Fig. 4.2. The Lomb-Scargle periodograms follow a similar curve shape, only with different scaling. Two common features on the periodograms for all sectors are a spike around 1.15 and 2.3 days. The higher one of these periods corresponds to the rotation period of the star P_{rot}.

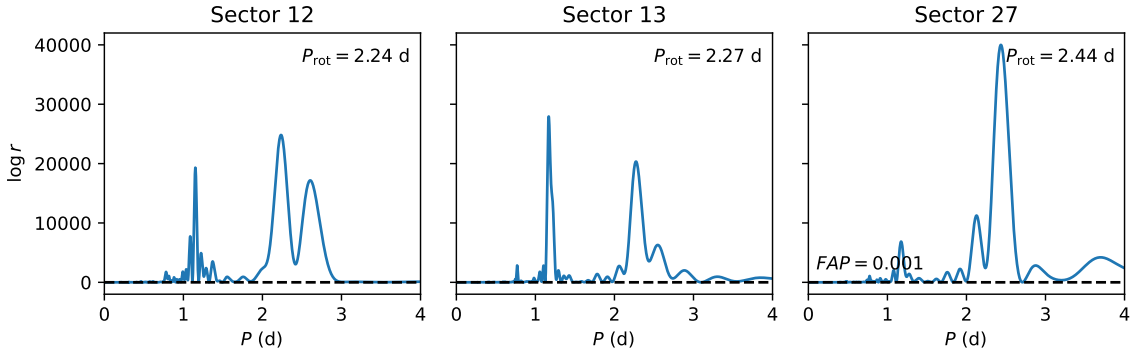

Figure 4.2: Likelihood ratio periodograms of HD 110082 for different TESS observing sectors. The changing likelihood ratios of different periods is an indicator of spot evolution on the star's surface. The black dashed line corresponds to a false alarm possibility of 0.001.

The explanation for choosing the spike with the longer period to be the rotation period of the star is as follows. The large scale variations in the lightcurve are most likely to be caused by starspots, viewed at different angles as the star rotates. First, let us imagine that there are two active regions on the star, at exactly 180-degree latitude difference. As the star rotates, at every half-rotation, a group of starspots will face the observer. In this scenario, the highest spike of the periodogram is at $P \approx P_{\mathrm{rot}}/2$.

If we assume the presence of only a single active region, periodic dimmings caused by the starspots facing us will only happen once every rotation. If this happens, the highest spike of the periodogram is at $P \approx P_{\mathrm{rot}}$.

The presence of multiple spikes on the periodograms thus is an indicator of multiple spot groups and their relative strength reflects on their relative sizes. The interpretation of change in spike positions and strengths is therefore a change in the spot pattern of the star. The presence of multiple active regions produces two spikes in the periodogram, one at $P \approx P_{\mathrm{rot}}/2$ and one at $P \approx P_{\mathrm{rot}}$.

The rotation periods for the different sectors based on the periodograms are $P_{\mathrm{rot}} = 2.24 \pm 0.11$ days for Sector 12, $P_{\mathrm{rot}} = 2.27 \pm 0.09$ days for Sector 13, and

$P_{\rm rot} = 2.44 \pm 0.09$ days for Sector 27. The error weighted average of these values is $P_{\rm rot} = 2.31 \pm 0.09$.

There is a clear shift towards a longer rotation period from Sectors 12 and 13 to Sector 27. This can be explained with a combination of starspot evolution and the differential rotation of the star. If a spot appears at a latitude that has a long orbital period, the periodogram shows a peak at a long period, corresponding to the rotation period of that latitude. The strength of that peak is proportional to the signal produced by the spot. In a similar way, spots appearing at latitudes with faster rotation produce peaks at shorter periods in the periodogram. The periodograms can therefore be interpreted as the sum of signals produced by starspots at different latitudes of the star. The highest period, clearly distinguishable individual peak from all periodograms of HD 110082 appears at $P \approx 2.6$ days (Sector 12), while the lowest period individual peak around the main calculated rotation period appears at $P \approx 2.1$ days (Sector 27).

4.3 Sinusoidal model

In order to account for the variations in flux, caused by the starspots in the lightcurve of HD 110082, the lightcurve was first fit with a sum of four sinusoids (the functional form is presented in Eq. 2.1, with $n = 4$) using a least squares fit.

The observations were divided into 1.5 day long sections, over which the fitting was performed. Since the sections were fitted independently from each other, when a combined lightcurve is created of the different sections, it might show some sudden "jumps" at the edges of the individual sections. To avoid these jumps, nine additional least squares fits were performed, all starting 0.15 days later than the previous one. The final least squares fit is the mean of these individual fits. The least squares fit along with the original lightcurve is shown on Fig. 4.3. Parts of this lightcurve along with the fit showing the planetary transit events are available in Section B of the Appendix.

The frequencies f of the individual sinusoids were searched for in accordance with the rotational period of the star for each TESS observing sector separately. Iterations on f were performed around $f \sim 1/P_{\rm rot}$ for one pair of sinusoids and $f \sim 2/P_{\rm rot}$ for the other pair of sinusoids. Using this choice of frequencies, was possible to cover the most significant peaks introduced by starspots and rotation in the periodograms of all sectors.

The lightcurve of two individual transit events, along with the least squares fit sinusoidal model, and a model for the signal the planet is expected to produce (see Section 4.7) is shown on Fig. 4.4. The fitted sinusoidal model follows the long timescale brightness variations of the star, but seems to spare the few hours long signal produced by the planetary transit. The lightcurve of all transiting events along with the model

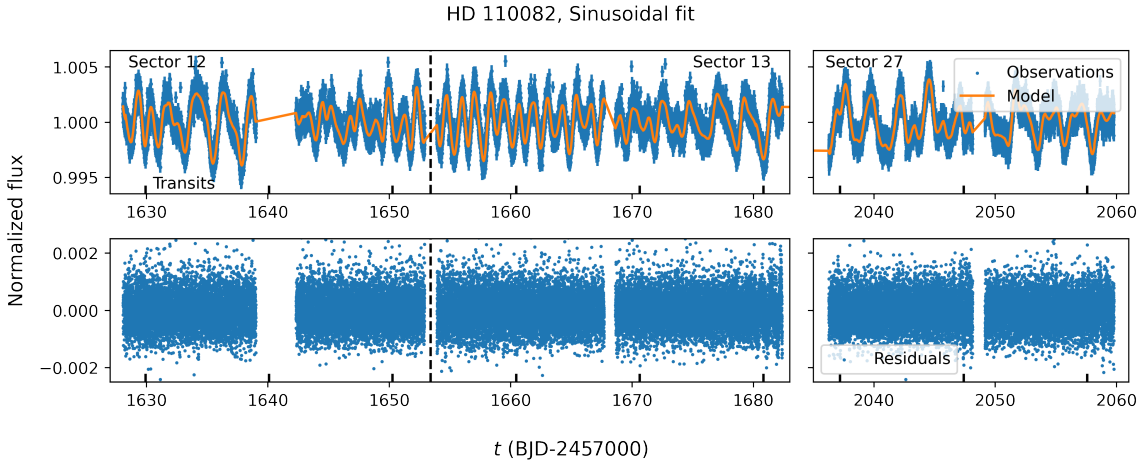

Figure 4.3: TESS lightcurve of HD 110082 fitted with a four component sinusoidal curve. The standard deviation of the residuals is $\sigma = 5.8 \times 10^{-4}$.

is shown on Fig. B.1 in the Appendix.

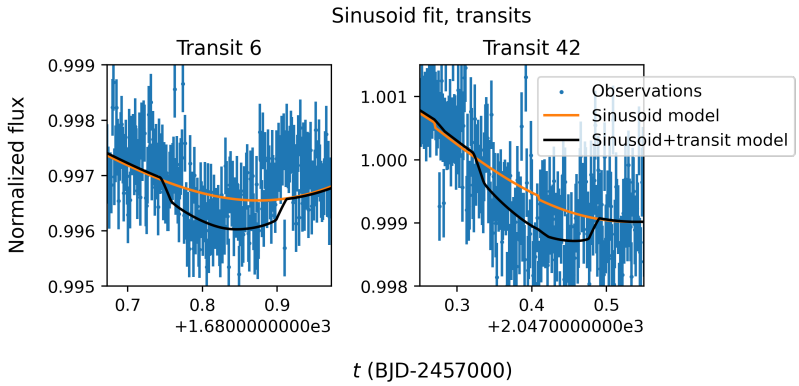

Figure 4.4: Transits 6 and 42 from the lightcurve of HD 110082. The least squares fit is marked with orange. A planetary transit model above the least squares fit (see Section 4.7) is marked with a black line.

An efficient way to visualize the efficiency of the fit in detrending the large scale variations of the lightcurve is supplied by the periodogram of the residuals. A likelihood ratio periodogram of the residuals is shown on Fig. 4.5. The effects of the two most significant peaks at $P_{\rm rot}$ and $P_{\rm rot}/2$ have been strongly filtered out by the sinusoidal model. The log likelihood of both $P_{\rm rot}/2$ and $P_{\rm rot}$ has weakened by a factor of 10^5. The periodogram of the residuals is dominated by a forest of spikes for periods shorter than one day. These variations are, however, insignificant compared to the prominent effects of starspots in the original TESS lightcurve. The variations appear to be on timescales that do not characteristically correspond to any photospheric phenomenon.

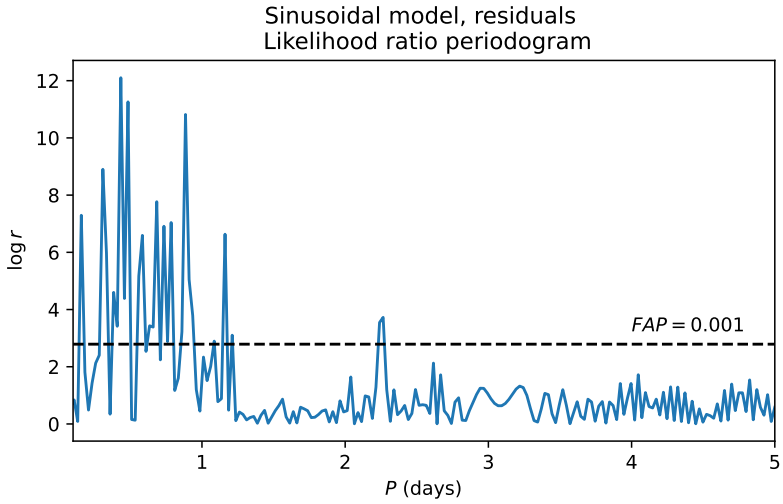

Figure 4.5: Likelihood ratio periodogram of the residuals of the sinusoid fit. A false alarm probability of 0.001 is marked with a black, dashed line.

4.4 MCMC sampling

The Adaptive Metropolis (AM) algorithm was used for the fitting of two models on the lightcurve of HD 110082. These models were a simple ARMA(1,1) model, and an ARMA(1,1) model where the constant c has been replaced with a sum of four sinusoidal functions (sinusoids+ARMA(1,1) model).

The appropriate length of the Markov chains for parameter estimation was determined for both models by assessing the Gelman-Rubin statistic $\hat{R}(z)$ of the Markov chains used for parameter estimation.

To create different Markov chains for the calculation of $\hat{R}(z)$, all initial parameters of a given model $z_1, z_2, ..., z_N$ were shifted by some random number ϵ_i. This random number was generated from a Gaussian distribution with $\epsilon_i \sim \mathcal{N}(0, 0.2)$ for all initial values individually, so that the new initial values of the Markov chains were $z_1 + \epsilon_1, z_2 + \epsilon_2, ..., z_N + \epsilon_N$. The threshold for convergence was chosen to be $\hat{R}(z) < 1.1$ for all parameters z.

For the combined sinusoid and ARMA(1,1) fitting, the TESS lightcurve was divided into 1.5-day-long sections. The calculation of $\hat{R}(z)$ for all individual sections using that model would have imposed a disproportionately large computational cost. For that reason, $\hat{R}(z)$ was only calculated for five randomly selected, 1.5-day-long sections. The Markov chain length of the section whose $\hat{R}(z)$ values have shown the slowest convergence towards one was chosen to be the required Markov chain length for all sections.

4.4.1 ARMA(1,1) model

An ARMA(1,1) model was fit to the whole TESS lightcurve with the AM algorithm using 140 000 iteration rounds.

As the ARMA(1,1) model does not take into account the frequency changes and other effects resulting from the evolution of the stellar surface, it was not necessary to divide the lightcurve into smaller sections, as it was with the least squares fitting. The whole MCMC fitting was performed on the whole lightcurve as a single piece.

The lightcurve along with the ARMA(1,1) model with the above parameters is shown on Fig. 4.6. A consequence of the ARMA(1,1) fit being very sensitive to the

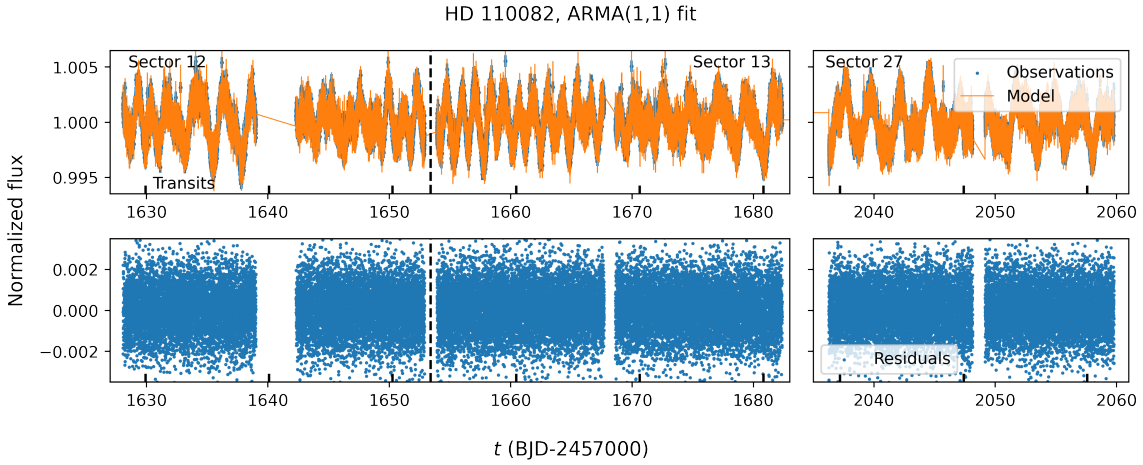

Figure 4.6: TESS lightcurve of HD 110082 along with the best fit ARMA(1,1) model. The standard deviation of the residuals is $\sigma = 1.6 \times 10^{-3}$.

minute changes of the lightcurve is the almost complete disappearance of the transit signal. This is demonstrated on Fig. 4.7. The figure shows two parts of the lightcurve around individual transiting events (the lightcurves of all transit events along with the ARMA(1,1) fit is shown on Fig. B.2 in the Appendix). The ARMA(1,1) model seems to closely follow the measurements. This is further shown when plotting the box fitting periodogram with the expected fractional transit length $q = 0.012$ for the residuals of the model (see Section 4.5).

The likelihood ratio periodogram of residuals of the model fit is shown on Fig. 4.8. The most significant peaks in the periodogram correspond to the rotation and half rotation period of the star. Although the likelihood ratio values of these peaks have decreased significantly compared to the original TESS lightcurve, they still dominate the periodogram. The reason behind this is, despite the fact that ARMA(1,1) fit follows the original TESS lightcurve rather closely, it tends to overestimate the minima of the lightcurve, and underestimate the maxima of the lightcurve. This causes the residuals to keep the period of the original observations.

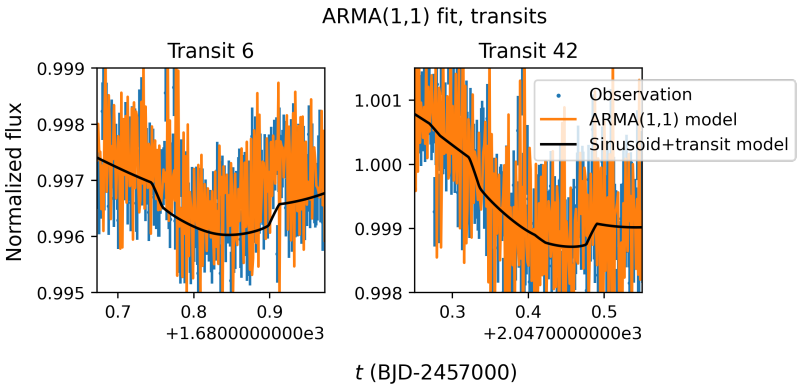

Figure 4.7: Transits 6 and 42 from the ARMA(1,1) fitting of the lightcurve of HD 110082. The least squares fit combined with a planetary transit model (see Section 4.7) is marked with a black line.

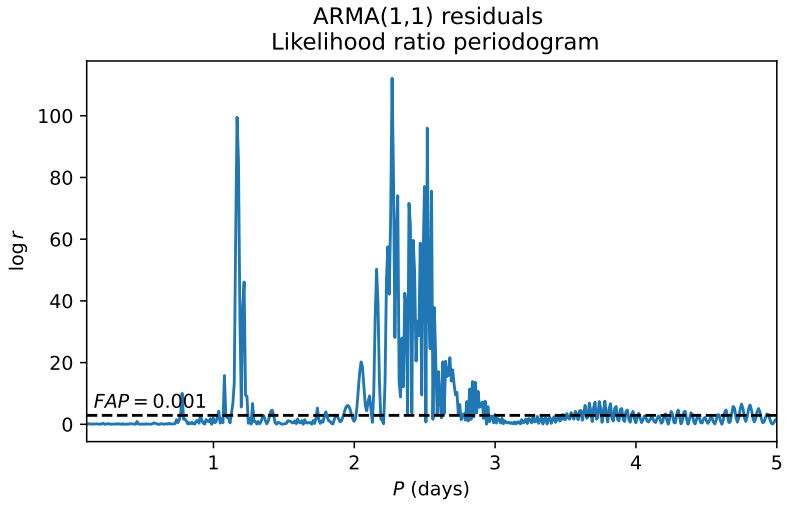

Figure 4.8: Likelihood ratio periodogram of the ARMA(1,1) model fit residuals.

4.4.2 Sinusoids + ARMA(1,1) model

The model employed for this section was an ARMA(1,1) model where the constant c has been replaced by a function, described as the sum of four sinusoids (Eq. 2.1, with $n = 4$). The lightcurve was divided into shorter sections similarly to what is described in Section 4.3. The fitting of the model was performed with the AM algorithm. The initial values describing the sinusoids (A_i, f_i, φ_i) were supplied by a least squares fit. The initial parameters describing the ARMA(1,1) part of the fit were set as $(\theta, \varphi, \sigma) = (0, 0, \sigma_X)$, where σ_X is the standard deviation of observations within the given section of the lightcurve. The chain length was set to 100 000, corresponding to the criterion $\hat{R}(z) < 1.1$.

The generated fit, along with the observed lightcurve is shown on Fig. 4.9. The shape of the fitted lightcurve bears a close resemblance to the simple sinusoidal fit.

However, as this model contains an ARMA(1,1) component, this fit is more sensitive to the smaller individual deviations of each observation from the sinusoidal curve.

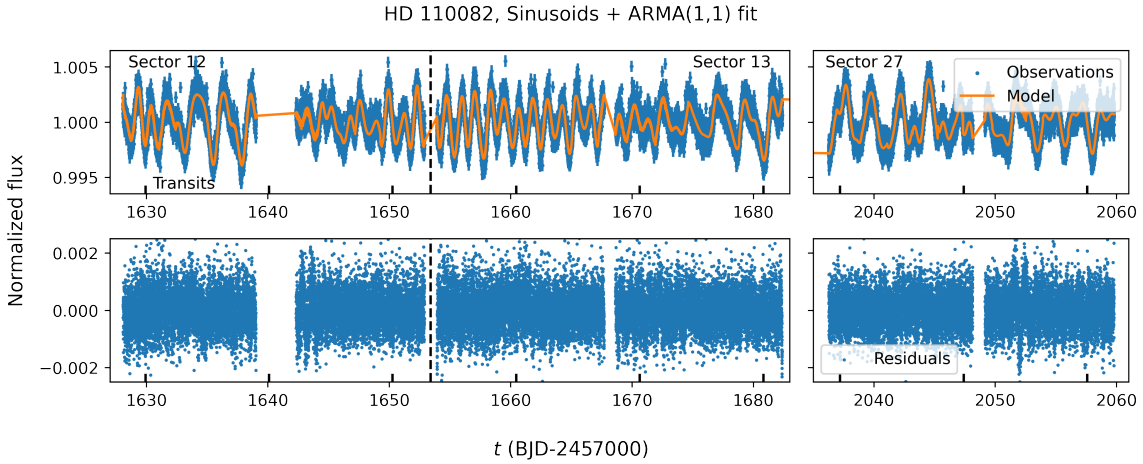

Figure 4.9: TESS lightcurve of HD 110082, fitted with the combined sinusoid + ARMA(1,1) model. The standard deviation of the residuals is $\sigma = 6.4 \times 10^{-4}$.

The likelihood ratio periodogram of the residuals is shown on Fig. 4.10. The effects of stellar rotation have been damped significantly. The peaks at $P_{\rm rot}$ and $P_{\rm rot}/2$ have log likelihood ratio values that are three orders of magnitude lower than those on the periodograms of the TESS observations. Indeed, the rotation induced variations appear to have been filtered out so completely that they disappear in the forest of signals of other periods. The periodogram is dominated by variations with periods less than a day, similarly to the simple sinusoid fit (see Fig. 4.5). Parts of the lightcurve showing individual transit events along with the fit are shown on Fig. B.3 in the Appendix.

4.5 Box fitting periodograms

The TESS lightcurves detrended with the sinusoid, ARMA(1,1), and combined sinusoids+ARMA(1,1) variability models were all analyzed with the box fitting periodogram with a fractional transit length of $q = 0.012$, predicted for HD 110082 b. The box fitting periodograms around the orbital period $P = 10.18271$ days for the three models are shown on Fig. 4.11. The box fitting periodograms over an extended range are given on Figs. C.4-C.6 in Section C.2 of the Appendix.

The log likelihood ratios, likelihood rations and false alarm probabilities at the orbital period $P_{\rm orb}$ are listed for each model in Table 4.2.

Of the three tested models, two, namely the sinusoidal and the sinusoid+ARMA(1,1) show a peak at the orbital period of the planet, with a false alarm probability $FAP < 0.05$.

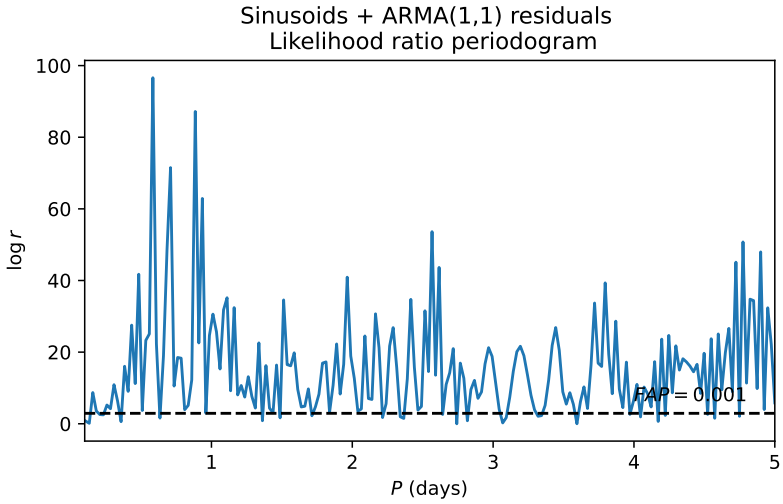

Figure 4.10: Likelihood ratio periodogram of the residuals from the Sinusoidal+ARMA fit. The standard deviation of the residuals is $\sigma = 6.4 \times 10^{-4}$.

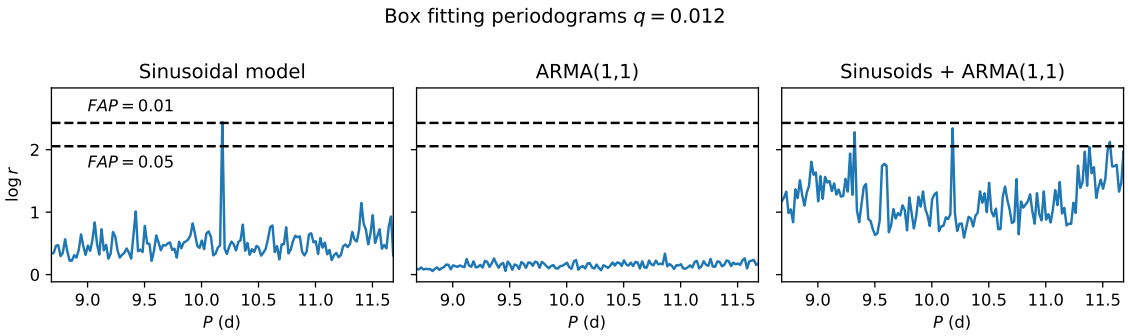

Figure 4.11: Box fitting periodogram for the residuals of the three models analyzed in this thesis. The fractional transit length $q = 0.012$ was chosen in accordance with the predictions for the planet. Black dashed lines represent the different false alarm possibilities.

The peak at $P_{\rm orb}$ on the box-fitting periodogram of the residuals of the sinusoid model stands out as a single solution. The sinusoids+ARMA(1,1) model was able to detrend the variations of the star efficiently enough so that the transit signal is returned with a $FAP < 0.05$. This signal, however, does not appear to be a unique peak as in the case of the simple sinusoid model. On the contrary, several peaks surpass the $FAP = 0.05$ threshold, and a signal with a stronger log likelihood is found at $P = 14.44$ days (see Fig. C.5 in the Appendix). These changes in the periodogram compared to that of the simple sinusoidal model is an artifact caused by the ARMA(1,1) component. While the simple sinusoidal model only fits the long timescale variations, the autoregression present in the model introduces a sensitivity to the individual observations. In this case, this results in the fitted lightcurve becoming more prone to producing false positive signals.

Table 4.2: Likelihood ratios and false alarm probabilities at the orbital period P of HD 110082 b per each tested model used for fitting the whole TESS lightcurve.

Model	Method of fitting	$\log r$ at P_{orb}	r at P_{orb}	FAP
Sinusoidal model	least squares	2.443	11.508	0.01
ARMA(1,1)	AM	0.179	1.196	0.77
Sinusoids + ARMA(1,1)	AM	2.343	10.412	0.016

The simple ARMA(1,1) model followed the TESS observations so closely that no significant peak is shown through the whole range in which the periodogram was evaluated.

4.6 Information criteria

The Akaike and Bayesian Information Criteria (AIC and BIC respectively) have been calculated for four different variability models. These models are 1) a reference model containing a single sinusoidal signal, 2) a sum of four sinusoids, 3) the ARMA(1,1) model, and 4) a sum of four sinusoids combined with the ARMA(1,1) model. All these models contain an excess noise component with standard deviation σ as a free parameter. The model uses the excess noise as a means to compensate for the deviation between the observed and expected values.

All models have been fitted for the first 6.5 days ($N = 4\,500$ observations) of the TESS lightcurve of HD 110082. This part of the lightcurve includes one transit of HD 110082 b (Transit 1). All fittings were performed for the whole of the 6.5 day long period as a single piece. This was necessary to avoid any possible discontinuity in the fitted model lightcurves.

The range of 6.5 days was chosen as an approximate maximum possible length for the reference model. Due to spot evolution and differential rotation, the lightcurve of HD 110082 is not likely to show a pattern that can be approximated with a single sinusoid. All models were fit with the AM algorithm to explore the possible range of log likelihood values $\log l$. The maximum of these $\log l$ values were used for the calculation of the AIC and BIC criteria.

The AIC and BIC values of different models, as well as the number of their free parameters k and average excess noise $\bar{\sigma}$ is presented in Table 4.3.

The stellar variability model with both the lowest AIC and BIC values is the ARMA(1,1) model (the best model is said to be the one that minimizes these values). This model, however, is too sensitive to the variations happening from one observation to the next. Indeed, this capacity is in the definition of ARMA models. As the model

Table 4.3: Information criteria and excess noise of different models to describe a section of the lightcurve of HD 110082.

Model	k	AIC	BIC	$\bar{\sigma}$
Single sinusoid	4	−519	−380	1.19×10^{-3}
Four sinusoids	13	−6 032	−5 949	5.29×10^{-4}
ARMA(1,1)	4	−10 497	−10 494	3.98×10^{-4}
Four sinusoids + ARMA(1,1)	16	−6 092	−5 989	4.05×10^{-4}

follows all the minute variations in the lightcurve exceedingly closely, even though it is the best model found in terms of AIC and BIC criteria, it is utterly useless for finding planetary signals.

The two models with the second and third lowest AIC and BIC values are the sum of four sinusoids combined with the ARMA(1,1) model and without the ARMA(1,1) model, respectively. Of these two, the one containing an ARMA(1,1) component proves to be the better model, as it is more sensitive to the minute changes taking place between observations. Nevertheless, the information criteria values of both of these models fall remarkably close to each other. The differences of the two models show more clearly in the average standard deviation of the model fit. The sinusoidal model with the ARMA(1,1) component showed a $\bar{\sigma}$ value that falls much closer to the $\bar{\sigma}$ value of the ARMA(1,1) fit compared to the simple four sinusoidal fit.

The model with the highest AIC and BIC values proved to be the one containing a single sinusoid. This model was the least able to follow the variations of the lightcurve. Its inability to produce a fit similar in quality to the other models is based on two reasons. The model did not contain a component that would help it adapt to the values of single observations and was not able to follow the possible differential rotation of the star, as it contained only a single frequency.

4.7 Transit modeling

Planetary parameters were estimated based on the transit model proposed by Carter et al. (2008) combined with a first-order linear limb darkening of the stellar disk (see Section 2.3.2).

Individual transiting events have been selected from the original TESS lightcurve of HD 110082. Observations in a 0.5-day-long interval both before and after the transit event have been fit with a four-component sinusoidal function, using the least squares method. The individual transit events themselves were left out from the fitted data-points. The least squares fitting of a simple sum of sinusoids was preferred over an AM

based, combined sinusoid and ARMA method as it gives results much faster.

These individual lightcurve sections, containing the transits were then combined into a phase-folded and binned lightcurve, with 400 bins in total. This phase-folded and binned lightcurve was then fitted with the transit model described in Section 2.3, for parameters $r, i, t_c - t_0$, and u – i.e., the radius ratio of the planet and its star, the planetary orbit's inclination, the difference between the center of the transit event and the value found in literature, and the linear limb darkening coefficient. The fitting was performed using the following assumptions:

1. The limb darkening of the star can be described as a first-order linear limb darkening.

2. The eccentricity of the planet is $e = 0$. This is a reasonable expectation, as the planet orbits its star on a very close orbit ($a = 0.113$ au), where the tidal interactions are predicted to circularize the orbit of a planet.

3. The maximum possible inclination achievable by the orbit of the planet is $i = 90°$. Solutions with higher inclinations are possible, but it is impossible to distinguish them from the solutions with $i < 90°$, as they are geometrically identical.

The fitting was performed using the AM algorithm. This makes it possible to get a full image on the probability distributions of the fitted parameters. A cornerplot, showing the histograms of the parameters generated for the AM fits, and their correlations is plotted on Fig. 4.12. The parameter estimate for each value is shown above the histogram of the parameter.

The phase folded and binned lightcurve around the transit event, along with 50 randomly sampled AM fits is shown on Fig. 4.13.

4.8 Transit mapping

Additional information on the stellar surface might be available from the shape of the individual transit lightcurves. If the planet passes in front of a darker region on the surface of the star, the received flux from the star increases. This can be observed as a small, temporary increase in the lightcurve of the transit event. In order for this increase to be credibly explained as a spot, the observed brightness should not exceed the value observed when the planet is not in front of the stellar disk.

The transit lightcurves of HD 110082 b contain a number of possible instances in which the planet transits starspots. Fig. 4.14 shows two individual transit events, Transit 1 ($t_c = 2\,458\,650.3$ BJD) and Transit 42 ($t_c = 2\,459\,047.4$ BJD). Long timescale variations have been removed in a similar way to what is reported in Section 4.7.

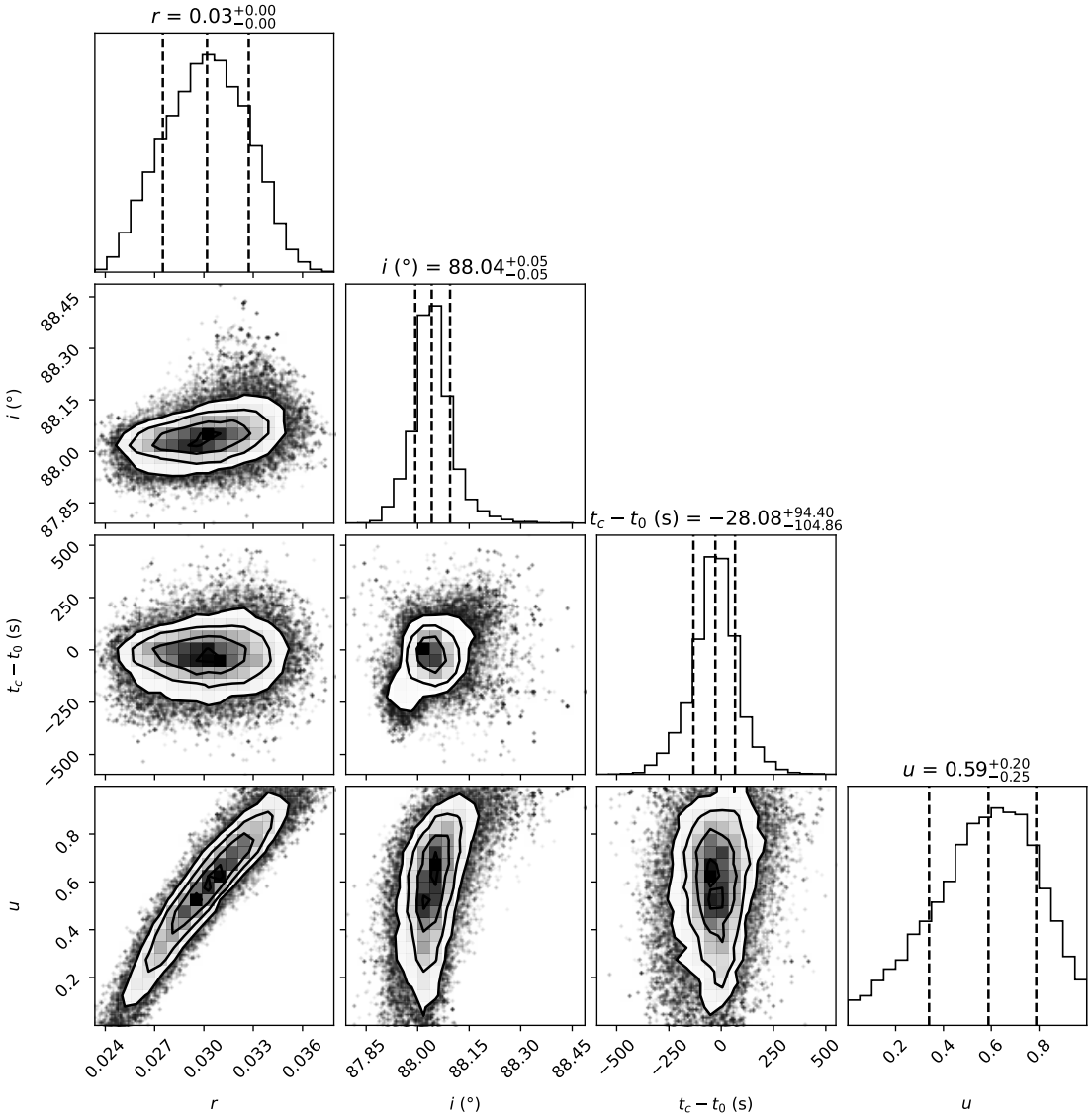

Figure 4.12: Cornerplot of parameters estimated by fitting the transit lightcurve with a linear approximation transit lightcurve, assuming a circular orbit, a known orbital period, stellar mass and radius. The parameter values were generated with the AM algorithm.

Besides the Gaussian jitter, observable through the whole lightcurve, two individual periods with a higher observed flux can be observed. These are marked as Feature 1 and 2. The features fulfill the criteria of a scenario in which the planet transits a starspot. A collection of all transit events is shown on Fig. B.4 in the Appendix.

The individual times of these spot eclipses are 0.12 hours and 0.28 hours. Assuming that the planet orbits on a circular orbit and that the features on the stellar disk remain motionless for the time of the transit, one can give an estimate on the minimal possible size $d_{\min}$ of these spots. These values are $d_{\min} \approx 122\,000$ km and $d_{\min} \approx 52\,000$ km for Features 1 and 2 respectively.

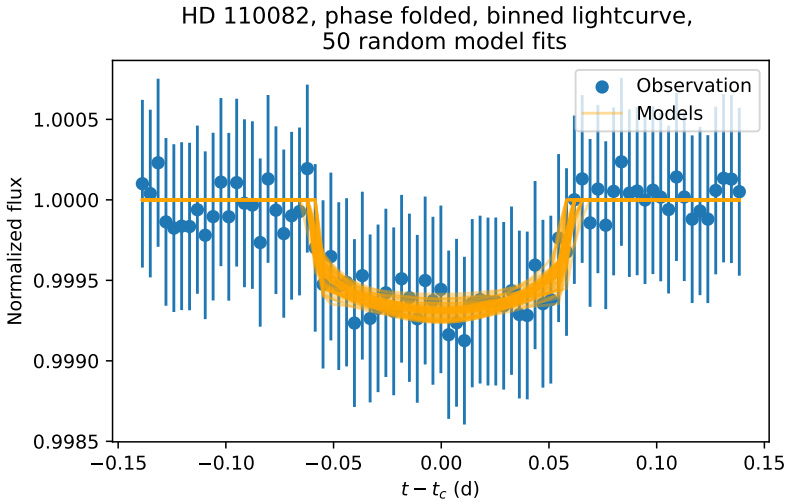

Figure 4.13: Phase-folded and binned lightcurve of HD 110082 around the transit event. The lightcurve was fit with a first-order linear limb darkening transit model, using the AM algorithm. The orange curves represent fifty different randomly drawn AM generated fits.

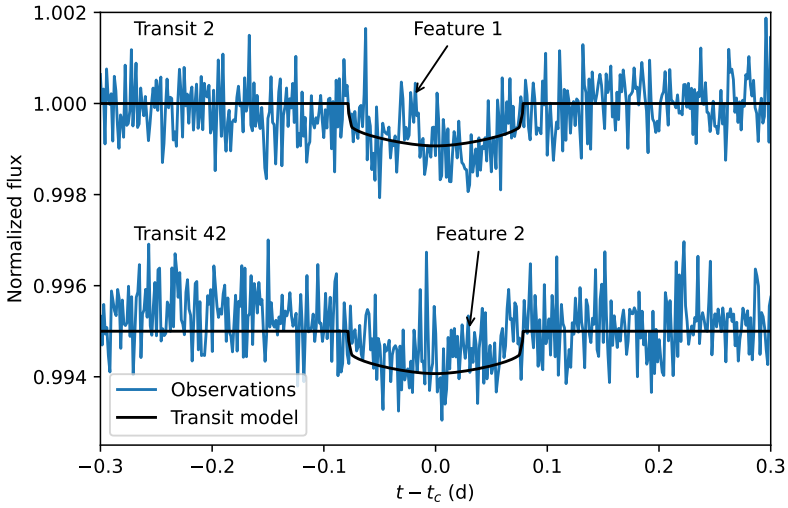

Figure 4.14: Individual transit events in the lightcurve of HD 110082. The two possible surface features on the star have been identified as Feature 1 and 2. The mean generated transit model is shown with a black line.

The scope of this thesis does not extend to a deeper analysis of the signals that might arise from the presence of starspots in the transit lightcurves. Nevertheless, such models would by all means provide valuable information about the spots on the stellar surface and produce a more accurate fit for the transit lightcurve.

5. Discussion

In this thesis, I applied a number of different approaches to describe the activity of the young star HD 110082. The lightcurve of HD 110082 is a prime example of an extrinsic variable star with a highly fluctuating brightness.

Using the Lomb-Scargle and likelihood ratio periodograms, I have found signs of spot evolution and differential rotation on the surface of HD 110082. The derived values for the rotation period of the star in this work are compared to those derived by Tofflemire et al. (2021) in Table 5.1. The values from the two sources are within the margin of error.

Table 5.1: Rotation periods P_{rot} of HD 110082 for different TESS observing sectors, and an average derived in this work and by Tofflemire et al. (2021).

	This thesis (d)	Tofflemire et al. (2021) (d)
Sector 12	2.24 ± 0.11	2.28 ± 0.05
Sector 13	2.27 ± 0.09	2.29 ± 0.03
Sector 27	2.44 ± 0.09	2.43 ± 0.03
Average	2.31 ± 0.09	2.34 ± 0.07

The minimum and maximum rotational period found based on the periodograms of different TESS observing sectors are $P \approx 2.1$ days and $P \approx 2.6$ days.

In the paper reporting on the discovery of HD 110082 b, Tofflemire et al. (2021) also explain the inferred rotational period change of HD 110082 with differential rotation. They state that the shift between rotation periods is well within the range observed for active stars in the lightcurves collected by the Kepler space telescope (Reinhold et al., 2013, Lanza et al., 2014). I found this statement also to be true for the individual periods $P \approx 2.1$ days and $P \approx 2.6$ days.

The whole TESS lightcurve of the star was fitted with three different models. The first model was composed as a sum of four sinusoidal signals with frequencies close to either the rotation period, or half rotation period of the star. The second model was an ARMA(1,1) model. The aim of using this model was to reduce the noise in the observations, from dependence on the last observation and its deviation

from the signal average. The model was fit with the AM algorithm. The third model was a combination of the four component sinusoid model and the ARMA(1,1) model. The frequencies of the sinusoids were set similarly as in the first model. Fitting was performed with the AM algorithm. The TESS lightcurve of HD 110082 along with each model fit is shown in Chapter A of the Appendix.

Both models containing sinusoids were able to weaken the effects of large timescale ($P > 1$ day) stellar variations so much that the transit signal of the planet became detectable. The likelihood ratio periodograms of the residuals from every model fit are shown in Section C.1. The sinusoidal model without the ARMA(1,1) component proved to produce a higher likelihood ratio signal at the expected orbital period of the planet.

The combined sinusoid and ARMA(1,1) model did manage to find the planetary transit signal, yet with a smaller likelihood ratio. An explicit disadvantage of this model is the fact that it was much more prone to identifying false positives than any other model examined. The simple ARMA(1,1) model followed all the variations of the star so closely that it filtered out the planetary transit signal to a level that it became undetectable.

All stellar variability models, along with a model containing a single sinusoidal curve have been assessed in terms of the Bayesian and Akaike Information Criteria. Based on these criteria, the best model is the simple ARMA(1,1) model. However, this model filtered out the planetary transit signals and therefore is not to be used in the survey for exoplanets.

The four component sinusoid combined with an ARMA(1,1) model, and a simple four component sinusoid had the second lowest and third lowest AIC and BIC values. The information criteria of these models fell rather close to each other.

Due to the relative similarities in information criteria of the aforementioned two models and the higher probability of false positives in the model containing an ARMA(1,1) component, the most effective model in order to find planetary signals proves to be one that can be expressed as a simple sum of sinusoids.

After the stellar activity has been subtracted from the observations, the transit signal of the planet was analyzed separately. The shape of the transit lightcurve was fit with a model assuming a linear limb darkening of the stellar disk. The fitting was performed with the AM algorithm. The parameters derived from the AM fitting for this thesis, as well as those calculated by Tofflemire et al. (2021), are listed in Table 5.2. The most striking difference between the two sets of values is the discrepancy between the radius ratios. This can be attributed to the fact that Tofflemire et al. (2021) used a different method with a quadratic limb darkening law to describe the stellar limb darkening. Assuming a uniform intensity stellar disk, the same AM model fit results

Table 5.2: Parameters estimated from the transit fitting.

	This thesis	Tofflemire et al. (2021)
r	0.030 ± 0.003	0.025 ± 0.001
i (°)	88.0 ± 0.05	$88.2^{+1.1}_{-0.7}$
$t_c - t_0$ (s)	$-28.1^{+94.4}_{-104.9}$	0 ± 86.4
u	$0.59^{+0.20}_{-0.25}$	–

in $r = 0.025 \pm 0.001$.

Finally, the individual transits were examined separately in order to find possible active regions on the star. The inclusion of more information about starspots has a potential at further improving the analysis of lightcurves.

Appendix A. TESS lightcurves

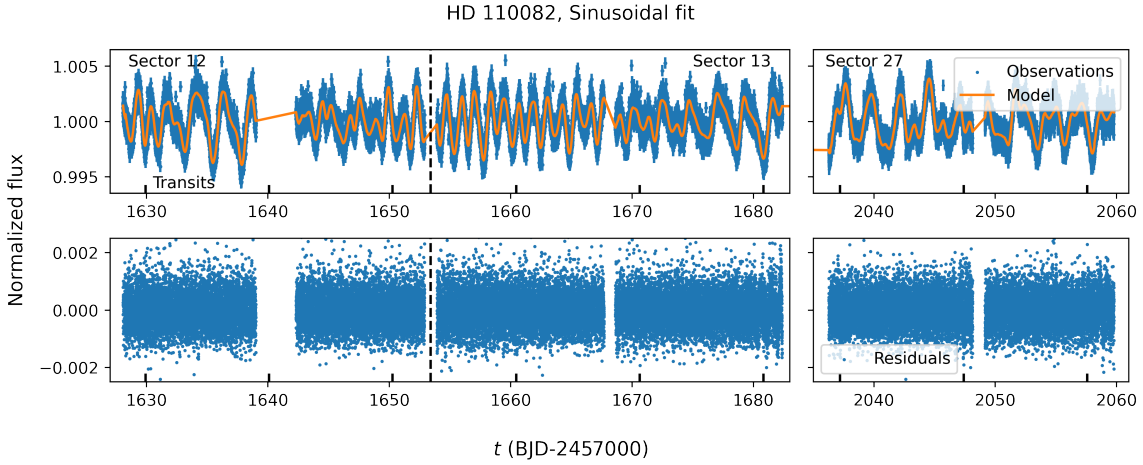

Figure A.1: TESS lightcurve of HD 110082 fitted with a four component sinusoidal curve. The standard deviation of the residuals is $\sigma = 5.8 \times 10^{-4}$.

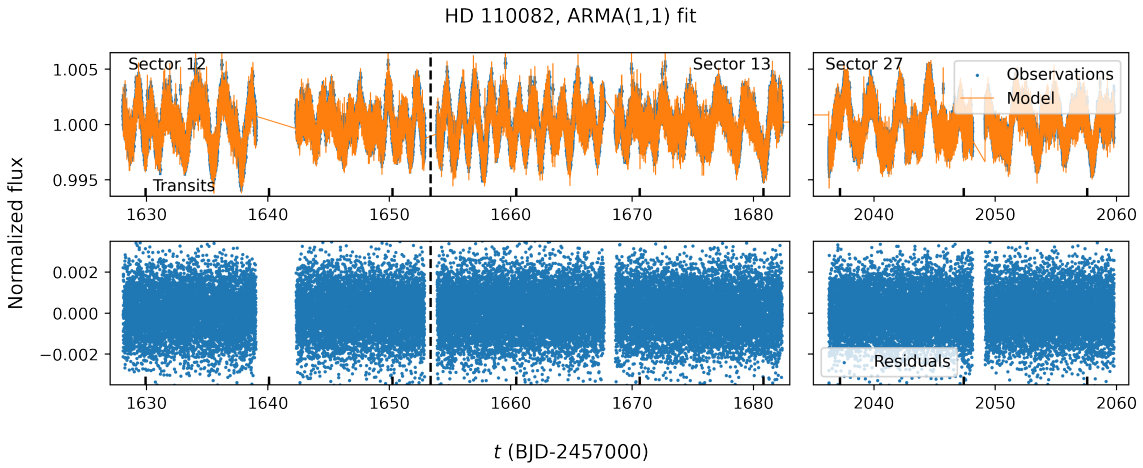

Figure A.2: TESS lightcurve of HD 110082 along with the best fit ARMA(1,1) model. The standard deviation of the residuals is $\sigma = 1.6 \times 10^{-3}$.

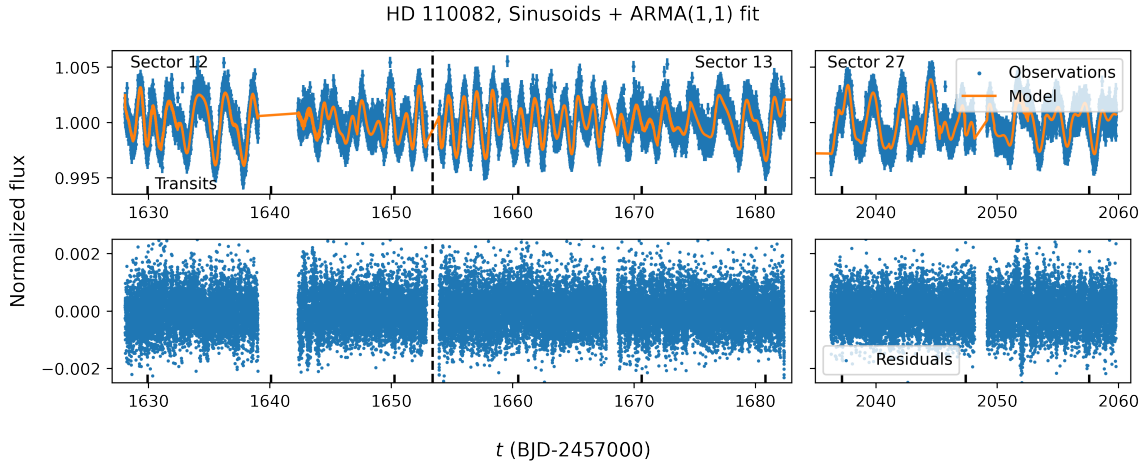

Figure A.3: TESS lightcurve of HD 110082, fitted with the combined sinusoid + ARMA(1,1) model. The standard deviation of the residuals is $\sigma = 6.4 \times 10^{-4}$.

Appendix B. Transit lightcurves

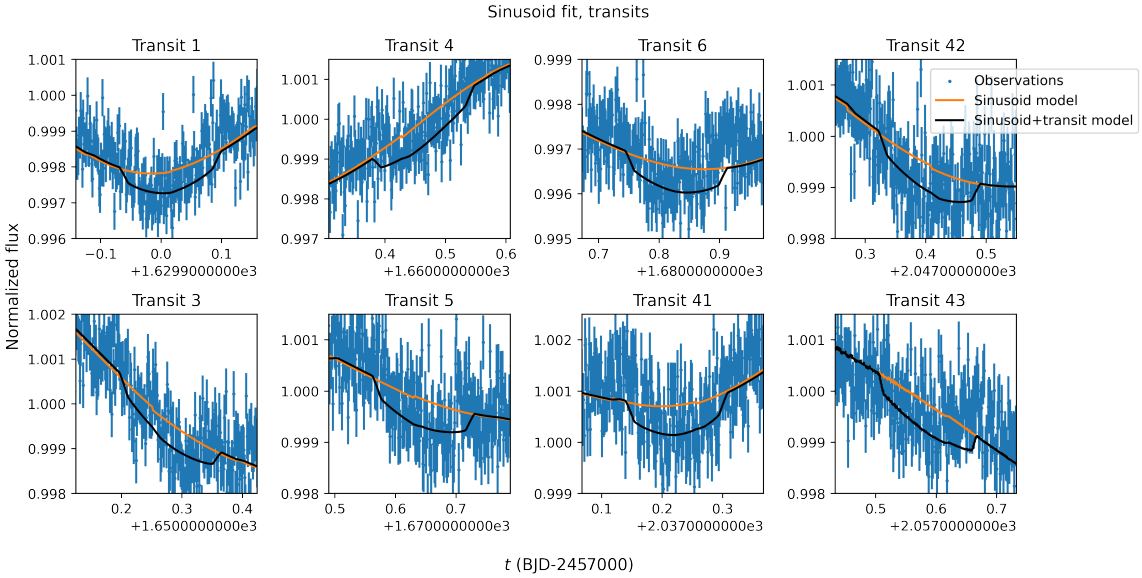

Figure B.1: TESS lightcurve of HD 110082 around individual transit events. The least squares fit of a four-component sinusoid model is shown in orange. Transit signals are shown with a black line for comparison.

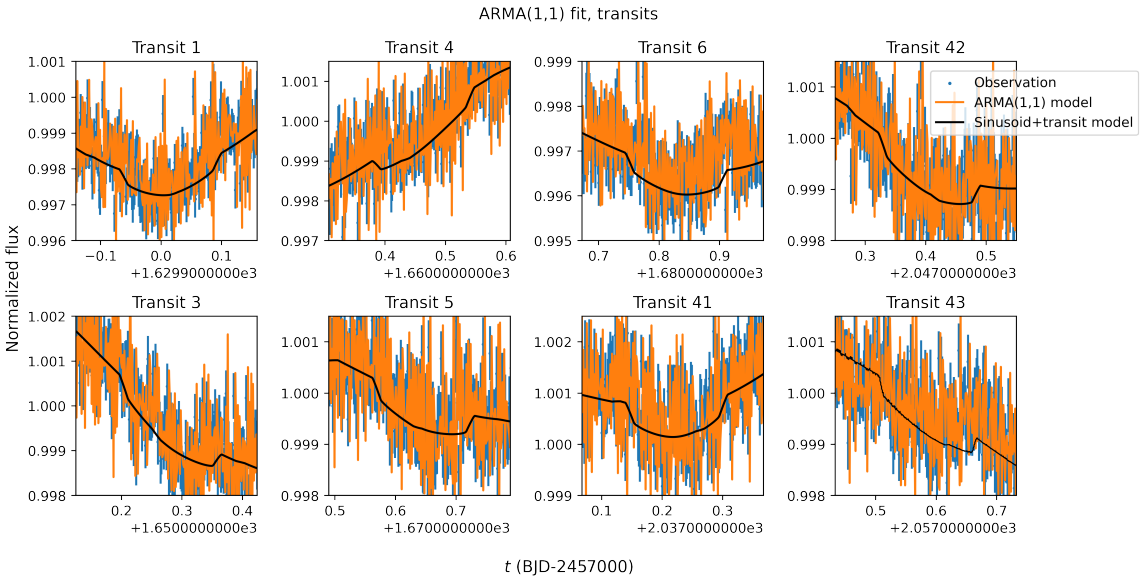

Figure B.2: TESS lightcurve of HD 110082 around individual transit events. The AM generated ARMA(1,1) is shown in orange. Transit signals are shown with a black line for comparison.

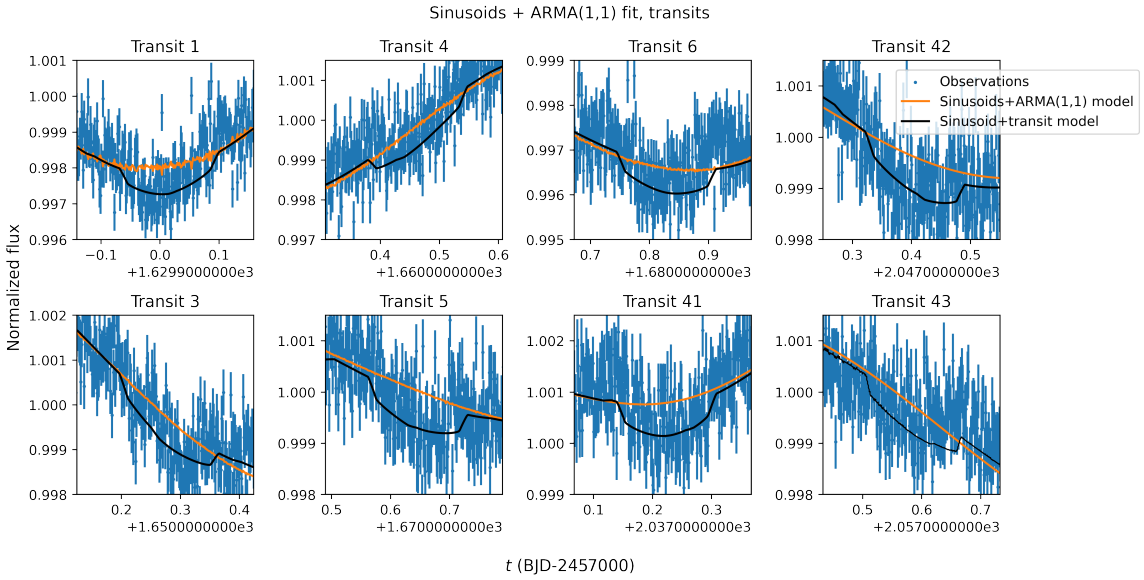

Figure B.3: TESS lightcurve of HD 110082 around individual transit events. The mean of modeled composite sinusoid+ARMA(1,1) lightcurves, generated by the AM algorithm are shown in orange. The least squares fit sinusoidal model with transits is shown as a black line for comparison.

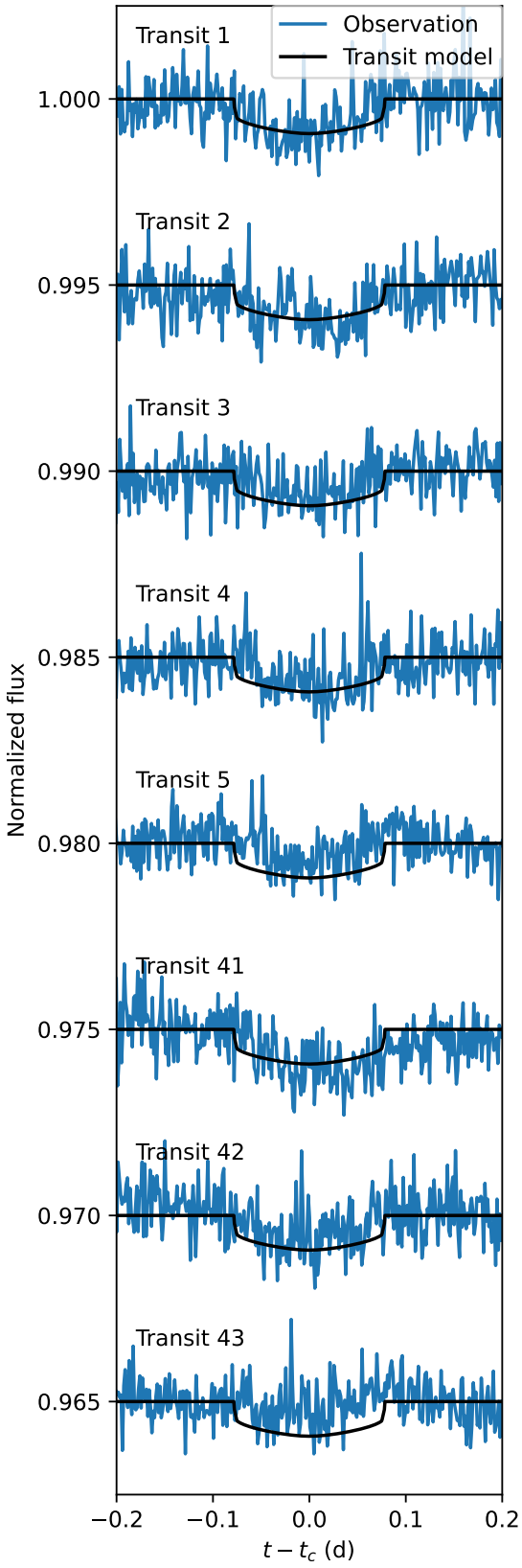

Figure B.4: Individual transit events of HD 110082 b. Long scale stellar brightness variations have been removed with a least-squares-fitted, four-component sinusoidal model (see Section 4.7). Brightness increases below the out-of-transit flux level, during transit events might be explained by the planet transiting a spot on the star's surface.

Appendix C. Periodograms of residuals

C.1 Likelihood ratio periodograms

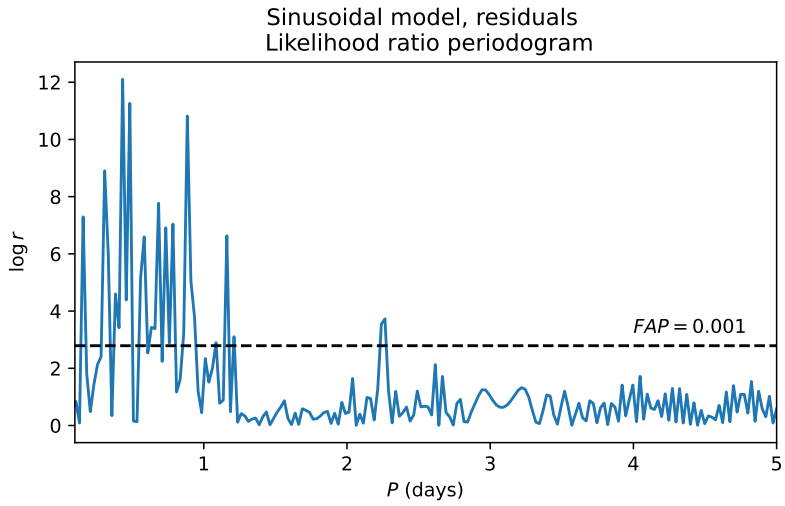

Figure C.1: Likelihood ratio periodogram of the sinusoidal model fit residuals.

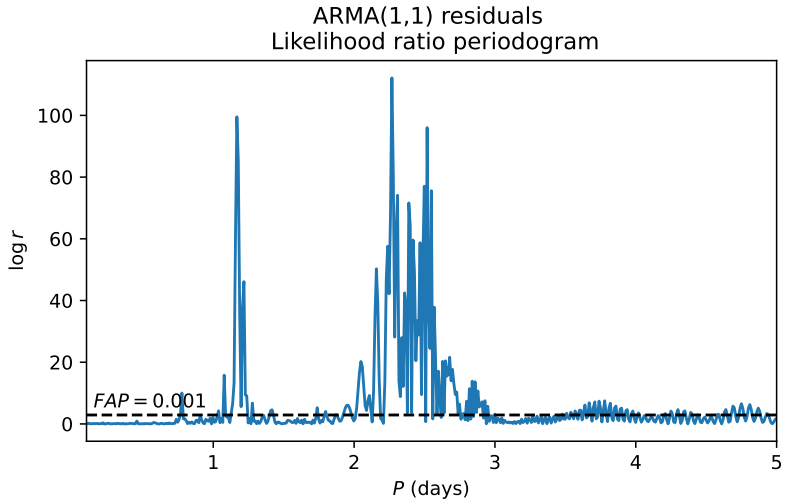

Figure C.2: Likelihood ratio periodogram of the ARMA(1,1) model fit residuals.

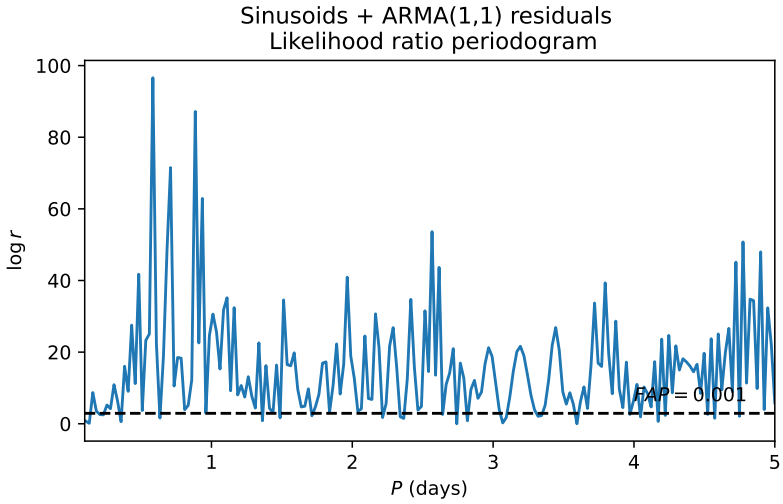

Figure C.3: Likelihood ratio periodogram of the residuals from the Sinusoidal+ARMA fit.

C.2 Box fitting periodograms

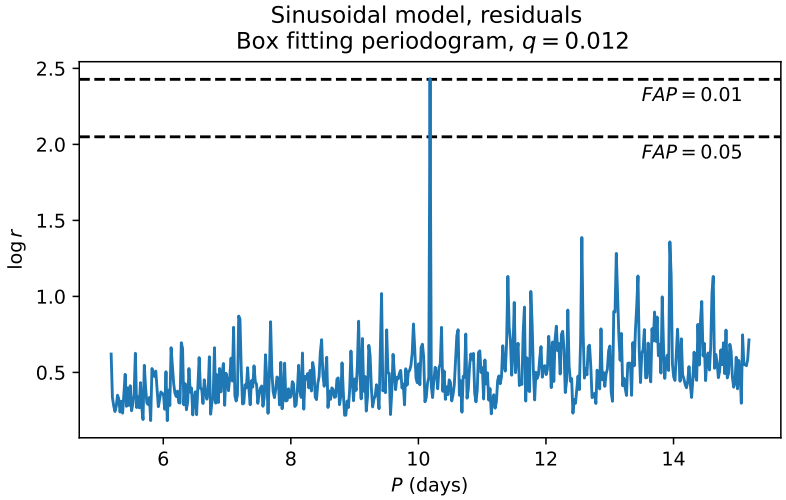

Figure C.4: Box fitting periodogram of the residuals of the sinusoid model fitting. The log likelihood of a transit at the observed orbital period of HD 110082 b is $\log r = 2.443$.

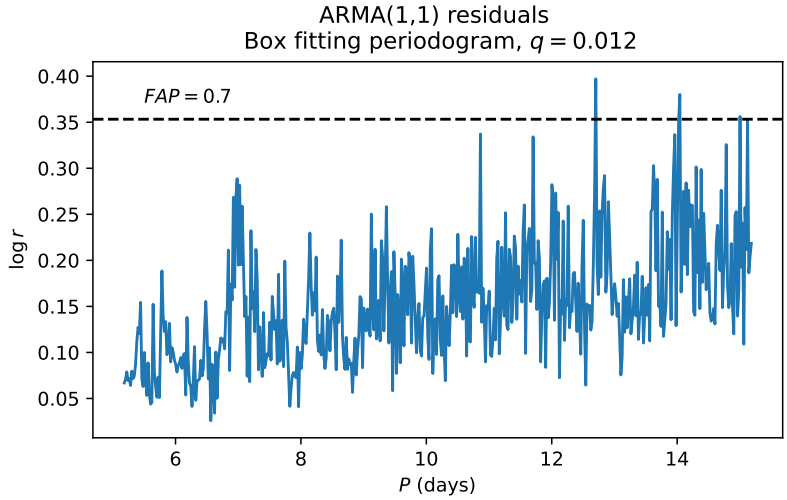

Figure C.5: Box fitting periodogram of the residuals of the ARMA(1,1) fitting. The log likelihood of a transit at the observed orbital period of HD 110082 b is $\log r = 0.179$.

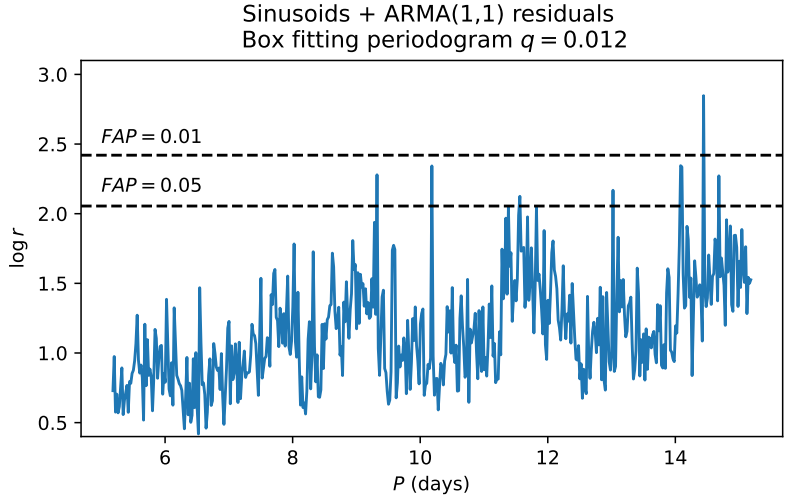

Figure C.6: Box fitting periodogram of the residuals of the combined sinusoid + ARMA(1,1) fitting. The log likelihood of a transit at the observed orbital period of HD 110082 b is $\log r = 2.343$.

Bibliography

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 19:716–723.

Anglada-Escudé, G. and Tuomi, M. (2012). A planetary system with gas giants and super-Earths around the nearby M dwarf GJ 676A. Optimizing data analysis techniques for the detection of multi-planetary systems. *Astronomy & Astrophysics*, 548:A58.

Bakos, G. Á. et al. (2007). HAT-P-1b: A Large-Radius, Low-Density Exoplanet Transiting One Member of a Stellar Binary. *The Astrophysical Journal*, 656(1):552–559.

Baliunas, S. L. and Vaughan, A. H. (1985). Stellar activity cycles. *Annual Review of Astronomy and Astrophysics*, 23:379–412.

Berdyugina, S. V. (2005). Starspots: A Key to the Stellar Dynamo. *Living Reviews in Solar Physics*, 2(1):8.

Brown, T. M., Latham, D. W., Everett, M. E., and Esquerdo, G. A. (2011). Kepler Input Catalog: Photometric Calibration and Stellar Classification. *The Astronomical Journal*, 142(4):112.

Carrington, R. C. (1859). Description of a Singular Appearance seen in the Sun on September 1, 1859. *Monthly Notices of the Royal Astronomical Society*, 20:13–15.

Carter, J. A., Yee, J. C., Eastman, J., Gaudi, B. S., and Winn, J. N. (2008). Analytic Approximations for Transit Light-Curve Observables, Uncertainties, and Covariances. *The Astrophysical Journal*, 689(1):499–512.

Chen, R., Liu, J. S., and Wang, X. (2002). Convergence analyses and comparisons of Markov chain Monte Carlo algorithms in digital communications. *IEEE Transactions on Signal Processing*, 50(2):255–270.

Chiavassa, A. et al. (2017). Measuring stellar granulation during planet transits. *Astronomy & Astrophysics*, 597:A94.

Collier Cameron, A. et al. (2007). WASP-1b and WASP-2b: two new transiting exoplanets detected with SuperWASP and SOPHIE. *Monthly Notices of the Royal Astronomical Society*, 375(3):951–957.

Doyle, L. R. et al. (2011). Kepler-16: A Transiting Circumbinary Planet. *Science*, 333(6049):1602.

Ford, E. B. (2005). Quantifying the Uncertainty in the Orbits of Extrasolar Planets. *The Astronomical Journal*, 129(3):1706–1717.

Ford, E. B. (2006). Improving the Efficiency of Markov Chain Monte Carlo for Analyzing the Orbits of Extrasolar Planets. *The Astrophysical Journal*, 642(1):505–522.

Gaia Collaboration (2018). Gaia Data Release 2. Summary of the contents and survey properties. *Astronomy & Astrophysics*, 616:A1.

Gelman, A., Roberts, G. O., and Gilks, W. R. (1996). Efficient metropolis jumping rules. In Bernardo, J. M., Berger, J. O., Dawid, A. P., and Smith, A. F. M., editors, *Bayesian Statistics*, pages 599–608. Oxford University Press, Oxford.

Gelman, A. and Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7:457–472.

Giménez, A. (2007). Equations for the analysis of the light curves of extra-solar planetary transits. *Astronomy & Astrophysics*, 474(3):1049–1049.

Haario, H., Saksman, E., and Tamminen, J. (2001). An adaptive metropolis algorithm. *Bernoulli*, 7(2):223–242.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

Hastings, W. K. (1970). Monte Carlo Sampling Methods using Markov Chains and their Applications. *Biometrika*, 57(1):97–109.

Henry, G. W., Marcy, G. W., Butler, R. P., and Vogt, S. S. (2000). A Transiting "51 Peg-like" Planet. *The Astrophysical Journal Letters*, 529(1):L41–L44.

Howard, A. W. et al. (2012). Planet Occurrence within 0.25 AU of Solar-type Stars from Kepler. *The Astrophysical Journal Supplement Series*, 201(2):15.

Konacki, M., Torres, G., Jha, S., and Sasselov, D. D. (2003). An extrasolar planet that transits the disk of its parent star. *Nature*, 421(6922):507–509.

Kovács, G., Zucker, S., and Mazeh, T. (2002). A box-fitting algorithm in the search for periodic transits. *Astronomy & Astrophysics*, 391:369–377.

Lanza, A. F., Das Chagas, M. L., and De Medeiros, J. R. (2014). Measuring stellar differential rotation with high-precision space-borne photometry. *Astronomy & Astrophysics*, 564:A50.

Léger, A. et al. (2009). Transiting exoplanets from the CoRoT space mission. VIII. CoRoT-7b: the first super-Earth with measured radius. *Astronomy & Astrophysics*, 506(1):287–302.

Liddle, A. R. (2007). Information criteria for astrophysical model selection. *Monthly Notices of the Royal Astronomical Society*, 377(1):L74–L78.

Lomb, N. R. (1976). Least-Squares Frequency Analysis of Unequally Spaced Data. *Astrophysics and Space Science*, 39(2):447–462.

Mandel, K. and Agol, E. (2002). Analytic Light Curves for Planetary Transit Searches. *The Astrophysical Journal Letters*, 580(2):L171–L175.

Mayor, M. and Queloz, D. (1995). A Jupiter-mass companion to a solar-type star. *Nature*, 378(6555):355–359.

Milne, E. A. (1921). Radiative equilibrium in the outer layers of a star. *Monthly Notices of the Royal Astronomical Society*, 81:361–375.

Müller, D. A. N., Steiner, O., Schlichenmaier, R., and Brandt, P. N. (2001). Time-slice diagrams of solar granulation. *Solar Physics*, 203(2):211–232.

Paladini, C. et al. (2018). Large granulation cells on the surface of the giant star π^1 Gruis. *Nature*, 553(7688):310–312.

Pont, F. et al. (2007). Hubble Space Telescope time-series photometry of the planetary transit of HD 189733: no moon, no rings, starspots. , 476(3):1347–1355.

Reinhold, T., Reiners, A., and Basri, G. (2013). Rotation and differential rotation of active Kepler stars. *Astronomy & Astrophysics*, 560:A4.

Scargle, J. D. (1982). Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal*, 263:835–853.

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461 – 464.

Schwarzschild, M. (1975). On the scale of photospheric convection in red giants and supergiants. *The Astrophysical Journal*, 195:137–144.

Strassmeier, K. G. (1999). Doppler imaging of stellar surface structure. XI. The super starspots on the K0 giant HD 12545: larger than the entire Sun. *Astronomy & Astrophysics*, 347:225–234.

Tofflemire, B. M. et al. (2021). TESS Hunt for Young and Maturing Exoplanets (THYME). V. A Sub-Neptune Transiting a Young Star in a Newly Discovered 250 Myr Association. *The Astrophysical Journal*, 161(4):171.

Tuomi, M. et al. (2018). AD Leonis: Radial Velocity Signal of Stellar Rotation or Spin-Orbit Resonance? *The Astronomical Journal*, 155(5):192.

West, A. A. et al. (2008). Constraining the Age-Activity Relation for Cool Stars: The Sloan Digital Sky Survey Data Release 5 Low-Mass Star Spectroscopic Sample. *The Astronomical Journal*, 135(3):785–795.

Wilks, S. S. (1938). The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics*, 9(1):60 – 62.

Winn, J. N. (2009). Measuring accurate transit parameters. In Pont, F., Sasselov, D., and Holman, M. J., editors, *Transiting Planets*, volume 253, pages 99–109.

Wolszczan, A. and Frail, D. A. (1992). A planetary system around the millisecond pulsar PSR1257 + 12. *Nature*, 355(6356):145–147.

Yamashiki, Y. A. et al. (2019). Impact of Stellar Superflares on Planetary Habitability. *The Astrophysical Journal*, 881(2):114.