

Mateus Cruz Targino

Análise do período de rotação de estrelas do tipo solar observadas pelo satélite Kepler

Natal – RN

Dezembro de 2019

Mateus Cruz Targino

Análise do período de rotação de estrelas do tipo solar observadas pelo satélite Kepler

Trabalho de Conclusão de Curso de Engenharia de Computação da Universidade Federal do Rio Grande do Norte, apresentado como requisito parcial para a obtenção do grau de **Bacharel em Engenharia de Computação**

Orientador: Jefferson Soares da Costa

Universidade Federal do Rio Grande do Norte – UFRN
Departamento de Engenharia de Computação e Automação – DCA
Curso de Engenharia de Computação

Natal – RN
Dezembro de 2019

Universidade Federal do Rio Grande do Norte - UFRN
Sistema de Bibliotecas - SISBI
Catalogação de Publicação na Fonte. UFRN - Biblioteca Central Zila Mamede

Targino, Mateus Cruz.

Análise do período de rotação de estrelas do tipo solar observadas pelo satélite Kepler / Mateus Cruz Targino. - BCZM, 2019.

52 f.: il.

Monografia (graduação) - Universidade Federal do Rio Grande do Norte, Centro de Tecnologia, Departamento de Engenharia de Computação e Automação, Curso de Engenharia de Computação e Automação. Natal, RN, 2019.

Orientador: Prof. Dr. Jefferson Soares da Costa.

Coorientador: Prof. Dr. Luiz Affonso Henderson Guedes de Oliveira.

1. Período Rotacional - Monografia. 2. Satélite Kepler - Observação - Monografia. 3. Rotação Estelar - Monografia. 4. Linguagem Python - Monografia. I. Costa, Jefferson Soares da. II. Oliveira, Luiz Affonso Henderson Guedes de. III. Título.

RN/UF/BCZM

CDU 52-327

Mateus Cruz Targino

Análise do período de rotação de estrelas do tipo solar observadas pelo satélite Kepler

Trabalho de Conclusão de Curso de Engenharia de Computação da Universidade Federal do Rio Grande do Norte, apresentado como requisito parcial para a obtenção do grau de **Bacharel em Engenharia de Computação**

Orientador: Jefferson Soares da Costa

Trabalho aprovado. Natal – RN, 06 de Dezembro de 2019:

Prof. Dr. Jefferson Soares da Costa - Orientador
UFRN

**Prof. Dr. Luiz Affonso Henderson Guedes de Oliveira -
Coorientador**
UFRN

Prof. Dr. Francys Anthony da Silva
UFRN

Natal – RN
Dezembro de 2019

Ao meu pai, que sempre me deu aquele empurrão necessário para seguir em frente.

AGRADECIMENTOS

Aos meus amigos Amanda, Felipe, Glauco, Igor e Malu. Sem vocês o curso não seria o mesmo.

À Patrícia, pelo amor e compreensão nesse período final.

Aos meus pais e minha irmã pelo apoio e suporte durante toda a minha vida acadêmica.

À equipe de desenvolvimento da Liga Contra o Câncer, pelas dicas, sugestões e companheirismo.

Ao Professor Jefferson Costa, pela excelente orientação e paciência, mesmos nos períodos difíceis.

Aos meus amigos Will, Jonas e Elson, pelos anos de amizade.

*“There is nothing noble in being superior to your fellow man;
true nobility is being superior to your former self.”*
— Ernest Hemingway

RESUMO

Os satélites caçadores de planetas, como Kepler e CoRoT, nos trouxeram uma vasta quantidade de informações sobre corpos celestes que orbitam outras estrelas pela nossa galáxia. Além disso, eles também coletaram dados sobre as estrelas hospedeiras desses planetas e muitas outras que estavam em seus campos de visão, sendo a grande parte delas parecida com o nosso Sol. Neste trabalho, foi desenvolvido uma aplicação na linguagem Python que utiliza as técnicas do Periodograma Lomb-Scargle e a Minimização da Dispersão de Fase (PDM) para obter os períodos de rotação de um conjunto de estrelas do tipo solar, essas estrelas foram observadas pelo satélite Kepler, sem a necessidade de análise visual das curvas de luz, como é feito em boa parte da literatura. O objetivo deste trabalho foi aquisição de forma automatizada dos dados provenientes do satélite e, em seguida, o cálculo do período de rotação. Foi observado que os resultados obtidos estão coerentes com os presentes na literatura, e sendo calculados de forma mais eficiente também.

Palavras-chaves: Estrelas, Kepler, Python, Período Rotacional.

ABSTRACT

The planet-hunting spacecrafts, like Kepler and CoRoT, brought us back a huge amount of information about celestial bodies that rotates around another stars through our galaxy. Beyond that, they gathered data about the stars that hosts these planets and many others that has been laying around in it's field of view, which many of those being. In this work, we developed an application in Python which makes use of Lomb-Scargle Periodogram and Phase Dispersion Minimization (PDM) techniques to obtain the rotation periods of a solar-type star in our sample, this sample is composed by solar-type stars observed by Kepler spacecraft. We make this without performing a visual inspection of the light curve as has been done in most of the literature. In this work, our main goal is to acquire the data of the satellite in an automatized manner and perform the rotation period determination. In the results, we have have that our determination presented a good agreement with the literature, and we also have obtained it in a more efficient way.

Keywords: Stars, Kepler, Python, Rotation Period.

LISTA DE ILUSTRAÇÕES

Figura 1.1 – Reprodução de um dos desenhos de manchas solares feitos por Galileu, em 23 de junho de 1612.	16
Figura 1.2 – Reprodução do esquema instrumental do Kepler	19
Figura 1.3 – Campo de visão do satélite Kepler. Os retângulos no centro da imagem representam os 42 CCDs utilizados na captação da luz.	20
Figura 3.1 – Exemplo de curva de luz e seu ajuste.	32
Figura 3.2 – Demonstração de trânsito planetário em uma curva de luz. Pode-se observar que a cada dois dias, aproximadamente, um corpo passa em frente à estrela.	33
Figura 3.3 – Curva de luz de uma estrela variável Delta Cephei, onde é possível observar a periodicidade das pulsações que ocorrem no astro.	34
Figura 3.4 – Representação da órbita de um sistema estelar binário.	35
Figura 3.5 – Estrela binária Albireo	35
Figura 3.6 – Representação da curva de luz de uma binária eclipsante associada a cada evento de sua órbita.	36
Figura 3.7 – Fluxo principal do algoritmo.	37
Figura 3.8 – Formulário de obtenção dos dados.	37
Figura 3.9 – Detalhes dos parâmetros do formulário.	38
Figura 3.10–Reprodução da página com o link de download	39
Figura 3.11–Trecho da curva de luz para KIC 5271947 antes (direta) e depois (esquerda) da interpolação linear. Os períodos resultantes foram 19.63 dias e 19.65 dias, antes e depois do preenchimento dos dados, respectivamente, mostrando que os métodos utilizados lidam bem com dados irregularmente espaçados.	40
Figura 4.1 – Curva de luz para KIC 5271947, com a intensidade luminosa normalizada em função do tempo de observação em dias.	44
Figura 4.2 – Gráficos dos periodogramas lomb-scargle de ambos os cálculos realizados para KIC 5271947. Os períodos escolhidos foram, respectivamente, 22.95 dias e 19.65 dias.	44
Figura 4.3 – Gráfico do periodograma para o método da PDM realizado para KIC 5271947. O período escolhido foi aquele com o menor valor de Theta, ou seja, 19.65 dias.	45
Figura 4.4 – Gráfico mostrando a relação dos períodos obtidos com os presentes na literatura. Foram utilizadas 191 estrelas que são em comum com a literatura.	47

Figura 4.5 – Histograma de todos períodos obtidos no trabalho. 48

LISTA DE TABELAS

Tabela 1 – Comparação dos resultados obtidos neste trabalho com os obtidos por (NASCIMENTO et al., 2014). A interrogação na coluna de Alias se refere a um erro desconhecido no valor obtido.	46
Tabela 2 – Comparação dos resultados obtidos neste trabalho com os obtidos por (MCQUILLAN; MAZEH; AIGRAIN, 2014)	46
Tabela 3 – Tabela mostrando os períodos obtidos utilizando ambos os métodos deste trabalho, assim como os resultados obtidos por (MCQUILLAN; MAZEH; AIGRAIN, 2014)	48

LISTA DE ABREVIATURAS E SIGLAS

HTML	<i>HyperText Markup Language</i>
PDM	<i>Phase Dispersion Minimization</i>
DFT	<i>Discrete Fourier Transform</i>
FFT	<i>Fast Fourier Transform</i>
CoRoT	<i>COvection ROTation and planetary Transits</i>
CCD	<i>Charged-coupled device</i>
JSON	<i>JavaScript Object Notation</i>
ASCII	<i>American Standard Code for Information Interchange</i>

SUMÁRIO

1	INTRODUÇÃO	15
1.1	<i>Observatórios Espaciais</i>	17
1.1.1	CoRoT	18
1.1.2	Kepler	18
1.2	<i>Automação do Processamento dos Dados</i>	20
1.3	<i>Plano de Trabalho</i>	21
2	ANÁLISE DE SÉRIES TEMPORAIS	22
2.1	<i>Transformada de Fourier</i>	22
2.1.1	Transformada Discreta de Fourier (DFT)	23
2.1.1.1	Função Pente de Dirac	24
2.1.1.2	Transformada Discreta de Fourier	24
2.1.1.3	<i>Fast Fourier Transform</i> (FFT)	24
2.2	<i>Periodograma Lomb-Scargle</i>	25
2.3	<i>Estatística Descritiva</i>	26
2.3.1	Média	26
2.3.2	Mediana	27
2.3.3	Moda	27
2.3.4	Desvio Padrão e Variância	27
2.4	<i>Minimização da Dispersão de Fase</i>	28
2.4.1	Definições	28
2.5	<i>Bibliotecas e Ferramentas</i>	29
2.5.1	Jupyter Notebook	29
2.5.2	NumPy	29
2.5.3	SciPy	30
2.5.4	Pandas	30
2.5.5	Matplotlib e Seaborn	30
2.5.6	AstroPy	30
2.5.7	pwkit	30
3	DADOS OBSERVACIONAIS	32
3.1	<i>Curvas de luz</i>	32
3.1.1	Manchas Estelares	33
3.1.2	Trânsito	33
3.1.3	Pulsação	33

3.1.4	Sistemas Binários	34
3.2	<i>Escolha dos dados a serem analisados</i>	36
3.3	<i>Processamento e análise dos dados</i>	36
3.3.1	Entrada e busca de dados	37
3.3.2	Download e execução do script	38
3.3.3	Processamento dos dados	39
3.3.4	Cálculo do período	40
4	RESULTADOS	43
4.1	<i>Download e Processamento</i>	43
4.2	<i>Validação dos Dados</i>	45
4.3	<i>Períodos</i>	46
4.4	<i>Discussão</i>	46
5	CONCLUSÃO	50
	REFERÊNCIAS	51

1 INTRODUÇÃO

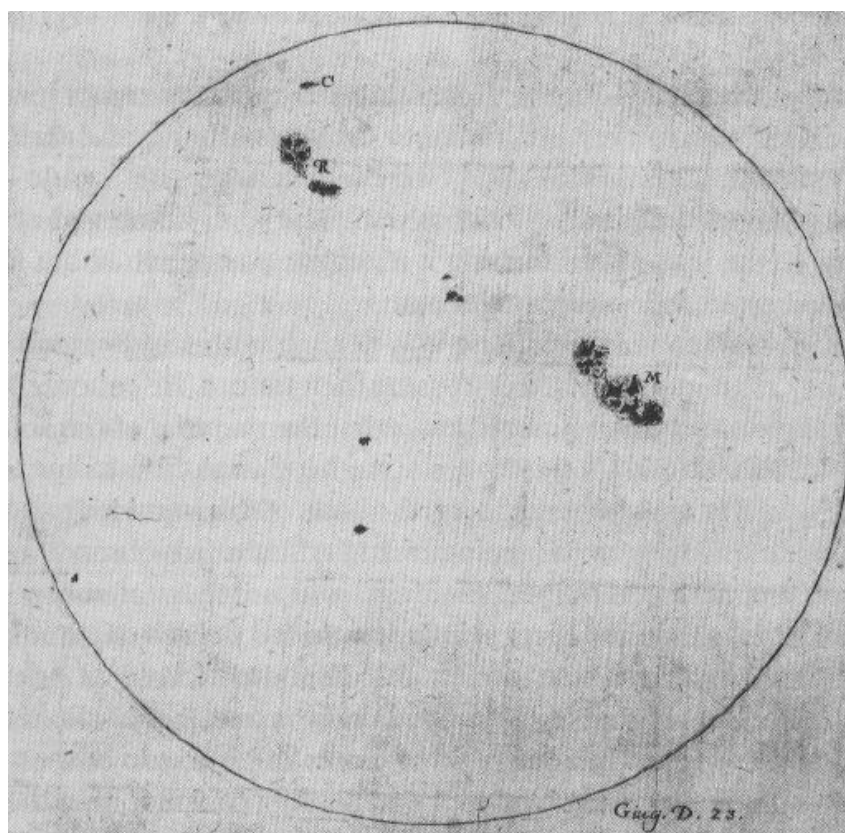
A humanidade sempre olhou para o céu em busca de respostas, seja para ter uma noção da passagem do tempo, para se orientar ou até mesmo para entender fenômenos naturais. Povos antigos associavam objetos celestiais a deuses e espíritos. A partir da posição e deslocamento desses objetos, tentavam definir períodos de chuva, seca, e até mesmo as estações e marés (KRUPP, 2003). Desta forma, a astronomia vem sendo importante para o desenvolvimento científico e tecnológico de várias civilizações, principalmente aquelas cuja economia era baseada em agricultura e pecuária. Estas se baseavam na posição dos astros para saber a época certa de plantio, ou qual seria o período de lua cheia, que era a única fonte de luz para viagens noturnas aos mercados das cidades (NILSSON, 1920). Desde o Antigo Egito já haviam tentativas de mapear os céus, onde o registro preciso mais antigo data de 1534 a.C. (SPAETH, 2000). Por volta da mesma época, astrônomos do Império Babilônico produziram os primeiros catálogos estelares (NORTH, 1994), onde estes possuíam listas de constelações, estrelas e planetas. Na civilização Ocidental, a catalogação realizada pelo astrônomo grego Hiparco (190 a.C. - 120 a.C.) foi uma das maiores e mais importantes, este catálogo possuía 1020 estrelas que foram mapeadas a olho nu e classificadas em seis categorias de acordo com o seu brilho aparente. O trabalho de Hiparco foi mais tarde utilizado no desenvolvimento do catálogo estelar de Ptolomeu, presente em seu célebre trabalho *Almagesto* (GRASSHOFF, 1990). Além do catálogo, o *Almagesto* também possuía uma definição para o movimento planetário, o modelo geocêntrico. Este modelo se baseava na hipótese de que a Terra era o centro do universo e os corpos celestes restantes giravam ao seu redor. Era defendido por vários filósofos da antiguidade, incluindo Aristóteles e foi desenvolvido de forma definitiva por Ptolomeu.

No século XVI ocorreu uma importante mudança de paradigma na astronomia, a revolução Copernicana. A revolução foi iniciada por Nicolau Copérnico, quando, em 1543, publicou o livro *De revolutionibus orbium coelestium* onde foi apresentado o modelo heliocêntrico, que estabelecia o Sol como centro do sistema solar. Apesar de um sistema heliocêntrico já ter sido sugerido anteriormente, a ampla aceitação do modelo desenvolvido por Copérnico ocorreu devido ao forte embasamento matemático presente no trabalho. Anos mais tarde, Tycho Brahe realizou observações mais acuradas que as de Copérnico e mesmo não sendo totalmente adepto ao heliocentrismo, suas contribuições fortaleceram ainda mais o modelo copernicano. Algum tempo depois, as observações de Brahe levaram Johannes Kepler, que era seu assistente, a formular as suas três leis do movimento planetário. A primeira lei, a lei das órbitas elípticas, estabelece que as órbitas dos planetas eram elípticas e tinham o Sol como um dos focos; já a segunda, a lei das áreas, diz que um segmento de linha que une o Sol a um planeta percorre áreas iguais em tempos iguais; por fim, a

lei dos períodos afirma que o quadrado do período orbital de um planeta é diretamente proporcional ao cubo do maior semi-eixo da sua órbita.

Várias observações importantes também foram realizadas por Galileu Galilei, através da sua versão aprimorada do telescópio, instrumento que tinha surgido naquele período e foi essencial para o desenvolvimento da astronomia observacional. Foi Galileu, aliás, um dos primeiros a observar as manchas solares, além de um dos protagonistas de um debate envolvendo esse fenômeno. O outro pivô dessa discussão foi Christoph Schneider, que acreditava que tais eventos eram satélites que orbitavam entre Mercúrio e o Sol. Galileu venceu a disputa, apresentando suas observações do fenômeno, não só afirmando que as manchas ocorriam de fato na superfície solar, como as usou para inferir que existe um movimento de rotação, assim como determinou o seu período como sendo próximo de um mês lunar. Na figura 1.1 pode ser vista a reprodução de uma das observações feitas por Galileu. Apesar da disputa entre Galilei e Schneider, a primeira observação registrada de uma mancha solar foi feita por Thomas Harriott, em 1610, e a primeira publicação de um estudo sobre o assunto foi por Johann Fabricius, no ano seguinte (MITCHELL, 1916).

Figura 1.1 – Reprodução de um dos desenhos de manchas solares feitos por Galileu, em 23 de junho de 1612.



Fonte – *Great Moments in the History of Solar Physics (1)* - Disponível em: https://web.archive.org/web/20060301083022/http://web.hao.ucar.edu/public/education/sp/great_moments.html

O estudo das estrelas evoluiu bastante no século XX, graças aos avanços realizados

nos equipamentos e técnicas de observação, como a fotometria. Tal progresso, permitiu o estudo das estrelas variáveis, aquelas que possuem uma variação na sua magnitude devido a vários fatores, como pulsação natural ou a presença de manchas estelares. Foi um grupo deste tipo de estrela, denominado Cefeidas, que foi utilizado em um estudo para a determinação da relação entre período de rotação e luminosidade (LEAVITT; PICKERING, 1912).

Nas décadas seguintes ocorreram várias descobertas, entre elas a do primeiro exoplaneta, 51 Pegasi b, através da técnica da velocidade radial¹ (MAYOR; QUELOZ, 1995). Não muito depois, foi detectado na estrela HD 209458 o primeiro trânsito planetário através de observações fotométricas em um telescópio terrestre (CHARBONNEAU et al., 2000). Esse resultado serviu de incentivo para a busca de trânsitos e também para o desenvolvimento de observatórios espaciais, com o objetivo de reduzir ruídos devido a fenômenos observacionais como a cintilação astronômica² e assim obter dados menos ruidosos e mais precisos.

1.1 *Observatórios Espaciais*

Realizar observações da terra é sempre muito complicado, pois a única informação que possuímos são os fótons e esses interagem com a atmosfera, interferindo nas medidas. Desta forma, é muito importante realizar observações do espaço. Assim, destacamos o lançamento das missões espaciais fotométricas, como MOST, CoRoT e Kepler, estas geraram avanços significativos para a astrofísica. A diferente abordagem na determinação de características estelares, assim como na sua classificação, iniciou uma nova era no estudo desses corpos celestes, pois até então eram informações obtidas através da observação terrestre. O grande diferencial foi que as estrelas passaram a ser observadas por longos intervalos de tempo na busca por planetas. Isso permitiu que estudos mais aprofundados de características como rotação e pulsação pudessem ser realizados de forma paralela. Essas longas observações eram necessárias, pois trânsitos podem levar entre meses e anos para serem detectados. Como na Terra existem fatores como clima e a própria rotação do planeta, esse feito é praticamente impossível de ser realizado por longos períodos. Nesta seção destacaremos os satélites CoRoT e Kepler no estudo de estrelas hospedeiras de planetas.

¹A presença de outro corpo causa um deslocamento nas linhas espectrais da estrela devido ao efeito doppler.

²Rápidas variações do brilho aparente de um objeto distante, visto através da atmosfera terrestre.

1.1.1 CoRoT

Desenvolvido pela agência espacial francesa (CNES) de forma cooperativa com a agência espacial europeia (ESA) e outros países, incluindo o Brasil, o CoRoT (COnvection, ROfation and planetary Transits) tinha como objetivo tanto a busca por planetas extrasolares, como estudar a asterosismologia das estrelas (BAGLIN et al., 2006). O satélite da missão foi lançado em dezembro de 2006 em um foguete Soyuz, da Rússia, no Cosmódromo de Baikonur, no Cazaquistão.

O telescópio possuía um espelho primário de 30cm de diâmetro, além de 4 sensores CCDs que coletavam a luz das estrelas observadas e alimentavam uma câmera que foi projetada em duas partes, enquanto uma apenas realizava a busca por exoplanetas, a outra, mais potente, observava variações sutis na luz emitida pelas estrelas com o intuito de realizar estudos asterosísmicos, sendo esses os dois programas da missão. Dessa forma, a missão foi dividida em períodos chamados de *runs*. Os períodos mais longos, denominados *long runs*, apresentavam observações com duração de até 150 dias, o satélite trabalhava em ambos os programas. Já nos períodos mais curtos, as *short runs*, que possuíam entre 20 e 40 dias (com algumas exceções que passavam dos 50 dias), a operação do satélite era exclusiva para o programa asterosísmico.

O satélite permaneceu em atividade por 7 anos, 5 meses e 20 dias, fornecendo mais de 160 mil curvas de luz e ajudando na descoberta de cerca de 40 novos planetas fora do Sistema Solar.

1.1.2 Kepler

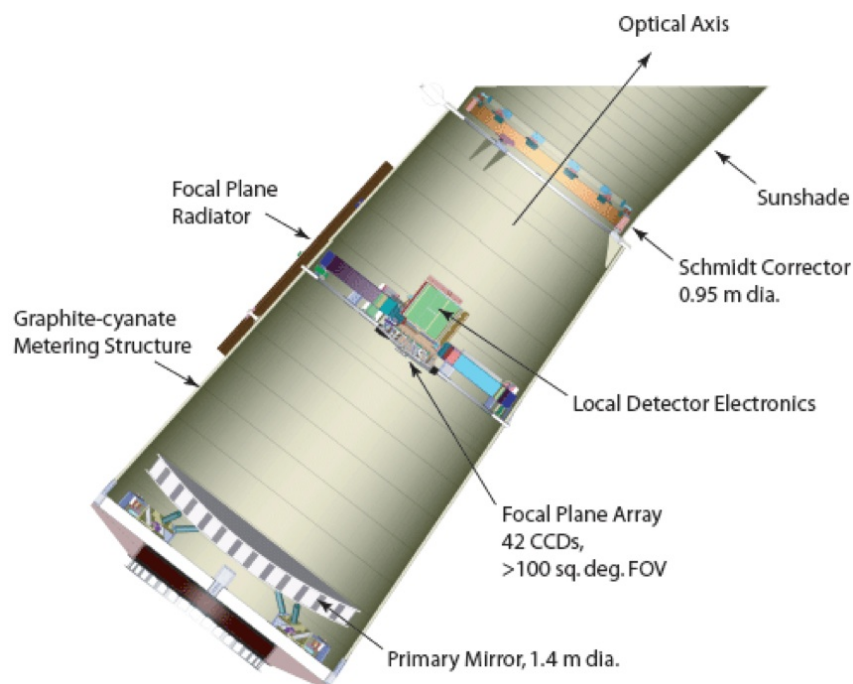
Em março de 2009, ainda com o CoRoT em atividade e como parte do Programa *Discovery*³, a agência espacial americana (NASA) lançou o telescópio espacial Kepler, cujo objetivo era detectar e observar a frequência com que planetas semelhantes à Terra apareciam na zona habitável, em especial em estrelas do tipo Solar. Nas primeiras 6 semanas de observação, a sonda monitorou mais de 150 mil estrelas, descobrindo cinco novos planetas de tamanhos variados (BORUCKI et al., 2010).

A grande quantidade de estrelas observadas foi possível devido ao fotômetro diferencial do Kepler, onde a luz absorvida passava por uma lente corretora de 0,95 metros para reduzir a aberração esférica⁴, depois era refletida por um espelho primário de 1,4 metro de diâmetro em direção aos 42 CCDs, o que permitia um campo de visão de 115 graus quadrados. O design do telescópio pode ser visto na figura 1.2, enquanto o campo de visão observado pelo Kepler pode ser visto na figura 1.3.

³Programa de pesquisa espacial que realiza experimentos científicos frequentes e de baixo custo.

⁴Fenômeno óptico em que os raios de luz incidentes próximos à borda das lentes são muito mais refratados do que os que estão mais próximos ao eixo óptico e os raios de luz incidentes próximo das bordas dos espelhos são refletidos além do foco.

Figura 1.2 – Reprodução do esquema instrumental do Kepler



Fonte – <https://keplerscience.arc.nasa.gov/the-kepler-space-telescope.html>

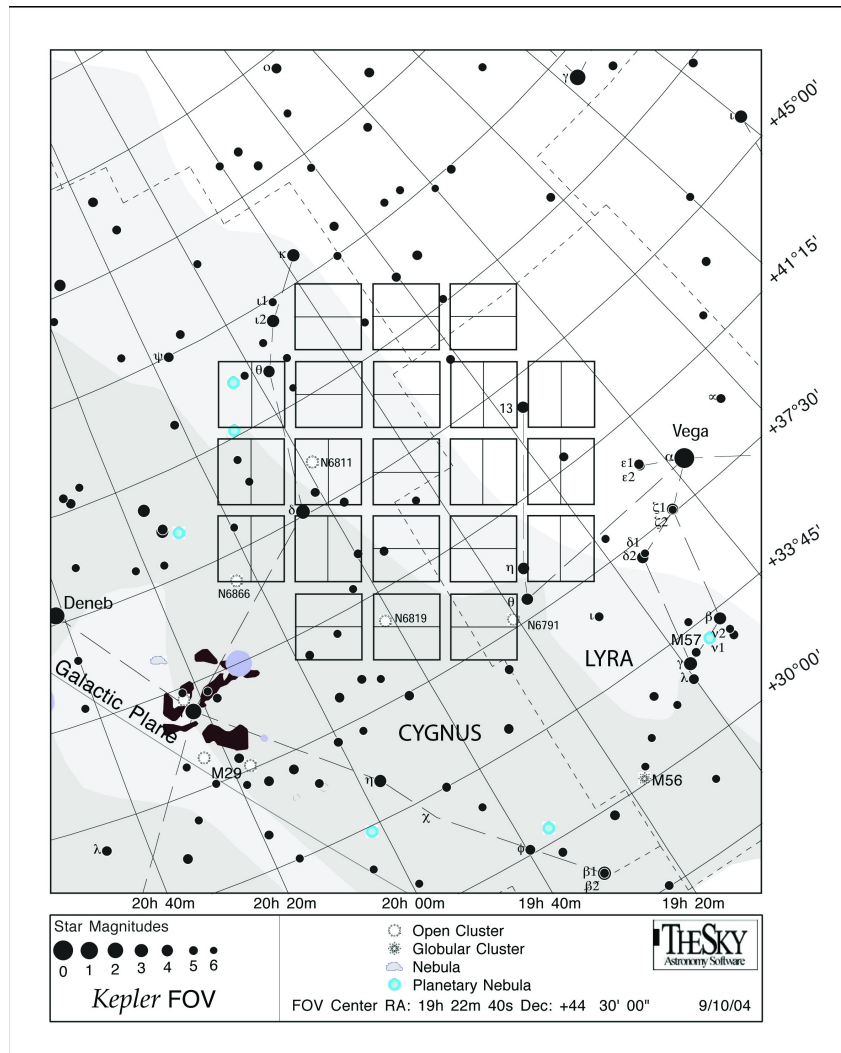
A missão inicialmente duraria três anos e meio. Nesse período, o telescópio observaria continuamente a região nas constelações de Lira e Cisne em busca de trânsitos planetários nas estrelas em seu campo de visão. Porém, em julho de 2012, uma parte do seu sistema giroscópico parou de funcionar, o que impediria o sucesso da mesma. Em maio do ano seguinte, a coleta de dados teve de ser interrompida devido à outra falha no mesmo sistema. Foi então que, com a ajuda da comunidade científica, a NASA anunciou em novembro daquele mesmo ano a missão K2. Sendo iniciada em maio do ano seguinte, a nova missão foi concebida de um modo que aproveitaria o satélite Kepler para realizar novas observações, levando em consideração as limitações e particularidades dos equipamentos, que cobririam uma área dez vezes maior que o campo de visão da missão original, onde ao invés de olhar para um ponto fixo no espaço, o satélite apontaria para vários pontos ao longo do plano eclíptico⁵ (HOWELL et al., 2014).

Após nove anos de atividade, em outubro de 2018 o telescópio precisou ser desativado devido ao fim do seu combustível. Como legado, o Kepler deixou um total de 530.506 estrelas observadas e ajudou a descobrir 2.662 exoplanetas durante o seu tempo de serviço⁶. Nosso trabalho se concentrará em estudar as estrelas observadas pelo Kepler.

⁵Plano da órbita terrestre ao redor do Sol.

⁶<https://www.nytimes.com/2018/10/30/science/nasa-kepler-exoplanet.html>

Figura 1.3 – Campo de visão do satélite Kepler. Os retângulos horizontais representam os 42 CCDs utilizados na captação da luz.



Fonte – <https://keplerscience.arc.nasa.gov/the-kepler-space-telescope.html>

1.2 Automação do Processamento dos Dados

As estrelas do tipo solar são aquelas que apresentam características físicas semelhantes ao nosso Sol e possuem a mesma classificação estelar, a classe espectral G. O satélite Kepler tinha como objetivo buscar planetas nas zonas habitáveis dessas estrelas, logo, uma quantidade massiva de dados sobre este tipo de estrelas foi coletada (BORUCKI et al., 2010). O estudo desse tipo de estrela é importante, pois nos ajuda a entender melhor o nosso Sol, a sua relação com outras estrelas e a capacidade de outros planetas sustentarem o tipo de vida que conhecemos (SODERBLOM; KING, 1998).

Atualmente, para obter as curvas de luz geradas pelas observações do Kepler, é preciso realizar um processo envolvendo vários passos, o que acaba se tornando algo contraproducente. Portanto, para realizar uma análise de uma quantidade massiva de dados,

muito tempo seria perdido apenas obtendo os dados necessários ao invés de realizando a análise em si. Após o *download* das informações, ainda é feito um tratamento para filtrar os dados necessários, isso além de calcular o período de rotação das estrelas.

Tem-se, como objetivo central deste trabalho, a automação do processo descrito anteriormente, de forma que este seja relativamente fácil de ser realizado por qualquer pessoa, assim como feito de uma maneira que aumente a produtividade de quem o realize.

1.3 *Plano de Trabalho*

Este trabalho apresenta uma introdução sobre o tema, mostrando os fatores que motivam a implantação da ideia, além da justificativa e dos objetivos. Em sequência, o Capítulo 2 aborda a metodologia, explicando os conceitos utilizados no trabalho, o Capítulo 3 trata dos dados observados, onde discutiremos suas características e a abordagem utilizada no seu tratamento, o Capítulo 4 apresenta os resultados obtidos, e por fim, o Capítulo 5 traz as principais conclusões e perspectivas associadas a esse trabalho.

2 ANÁLISE DE SÉRIES TEMPORAIS

Para entender a análise de séries temporais, é necessário ter duas definições em mente: a de séries temporais e a de sinais. O conceito de séries temporais é bastante comum entre estaticistas e analistas de dados. Pode-se definir que uma série temporal é um conjunto de observações de uma variável ao longo do tempo, em que estas são feitas de maneira regular e organizadas sequencialmente. Já o sinal, é uma função que transporta informações sobre propriedades ou atributos de um fenômeno. Sinais podem ocorrer naturalmente ou serem sintetizados (PRIEMER, 1991). Assim, podemos dizer que uma série temporal pode ser definida também como um sinal, permitindo o uso de técnicas de processamento de sinais para a análise da série. Há também o caso das séries que possuem espaçamento de dados irregular, onde as observações não são feitas de maneira constante, e isso ocorre continuamente na astronomia observacional. Para esses casos, é necessário o uso de técnicas mais sofisticadas para realizar uma análise mais coerente, conforme veremos no decorrer deste capítulo.

2.1 Transformada de Fourier

Em séries temporais, geralmente temos sinais periódicos simples, como senoides e cossenoides. Esses sinais, através da Série de Fourier, podem ser utilizados para representar sinais mais complexos, facilitando a análise. Para a decomposição dessas funções em termos de base sinusoidal, foi desenvolvida a Transformada de Fourier, que reduz uma função temporal em frequências. Este é considerado o método mais antigo para realizar a análise de sinais periódicos, sendo o precursor de vários outros. Diversos métodos vêm sendo desenvolvidos para o estudo da periodicidade de funções, porém a compreensão desta técnica, devido à sua simplicidade, ajuda no entendimento das outras que vieram posteriormente, inclusive daquelas que realizam uma abordagem diferente para o mesmo fim. Porém, antes de partirmos para a discussão de métodos mais complexos e elaborados, estudaremos em detalhes a Transformada de Fourier.

A definição da transformada de Fourier para uma função $g(t)$ é dada pela seguinte expressão:

$$\mathcal{F}\{g\} = \hat{g}(f) \equiv \int_{-\infty}^{\infty} g(t)e^{-2\pi ift} dt \quad (2.1)$$

e para recuperar a função original, temos a Transformada Inversa de Fourier, que é dada por:

$$\mathcal{F}^{-1}\{\hat{g}\} = g(t) \equiv \int_{-\infty}^{\infty} \hat{g}(f)e^{2\pi ift} df \quad (2.2)$$

Para exemplificar, aplicaremos a Transformada em um sinal periódico simples como $g(t) = \text{sen}(2\pi f_0 t)$, onde f_0 é uma frequência qualquer. Substituindo $g(t)$ na Eq. 2.1, obtemos a equação

$$\mathcal{F}\{g\} = \mathcal{F}\{\sin(2\pi f_0 t)\} = \int_{-\infty}^{\infty} \sin(2\pi f_0 t) e^{-2\pi i f t} dt \quad (2.3)$$

que apesar de sua solução possuir um grau de dificuldade elevado, pode ser simplificada, usaremos a fórmula de Euler e obteremos uma expressão para o $\sin(x)$, tal que $\sin(x) = \frac{e^{ix} - e^{-ix}}{2i}$. Ao aplicarmos o resultado na Eq. 2.2, temos que

$$\mathcal{F}\{\sin(2\pi f_0 t)\} = \frac{1}{2i} \int_{-\infty}^{\infty} (e^{2\pi i f_0 t} - e^{-2\pi i f_0 t}) e^{-2\pi i f t} dt \quad (2.4)$$

Sabendo que, por definição, a função Delta de Dirac, $\delta(t - t_0)$, pode ser expressa da seguinte forma $\int_{-\infty}^{\infty} e^{2\pi i f(t-t_0)} df$, podemos concluir que

$$\delta(f - f_0) = \mathcal{F}\{e^{2\pi i f_0 t}\} \quad (2.5)$$

assim, a Eq. 2.4 tem seu resultado dado por

$$\mathcal{F}\{\sin(2\pi f_0 t)\} = \frac{1}{2i} [\delta(f - f_0) - \delta(f + f_0)] \quad (2.6)$$

portanto, a transformada de Fourier de um sinal senoidal com período $T = 1/f_0$ possui um resultado que pode ser descrito como um conjunto de impulsos em $\pm f_0$. Propriedades como essas, permitem que a transformada de qualquer sinal que seja composto por senoides consista de uma soma de impulsos, o que torna a transformada de Fourier uma poderosíssima ferramenta na análise de sinais, pois permite identificar as frequências que constituem aquele sinal, assim como as suas magnitudes e fases.

A partir da Transformada de Fourier, podemos obter a densidade espectral de um sinal $x(t)$, essa densidade é definida como sendo o quadrado da magnitude da transformada do sinal:

$$P_x(f) = |\hat{x}(f)|^2 \quad (2.7)$$

onde a contribuição de cada frequência f_i presente no sinal é quantificada a partir de f .

2.1.1 Transformada Discreta de Fourier (DFT)

Na maioria das aplicações do mundo real, não é feita a análise de um sinal contínuo, pois geralmente lidamos com dados discretos. Esses dados são finitos e obtidos através de algum tipo de sensor que os obtém de forma igualmente espaçada temporalmente. Para realizar a análise de Fourier desses sinais, foi desenvolvida, a partir da função Pente de Dirac, a Transformada Discreta de Fourier (DFT).

2.1.1.1 Função Pente de Dirac

A função Pente de Dirac, também conhecida como trem de impulsos ou função de amostragem, é uma distribuição periódica construída a partir de sinais de impulso dado um período T e pode ser definida como $III_T(t) \equiv \sum_{k=-\infty}^{\infty} \delta(t - kT)$. Por ser uma função periódica, também pode ser representada como uma série de Fourier:

$$III_T(t) = \frac{1}{T} \sum_{n=-\infty}^{\infty} e^{i2\pi n \frac{t}{T}} \quad (2.8)$$

cuja transformada é

$$\mathcal{F}\{III_T(t)\} = \sum_{n=-\infty}^{\infty} e^{i2\pi f n T} \quad (2.9)$$

2.1.1.2 Transformada Discreta de Fourier

Quando uma função contínua possui uma amostragem feita em intervalos regulares, os impulsos no pente de Dirac servem como forma de compactar a integral de Fourier em uma soma de Fourier, o que nos ajuda a chegar na forma geral da DFT (VANDERPLAS, 2018). Portanto, a partir de um sinal real $g(t)$ contínuo e infinito, podemos observá-lo e amostrá-lo por um período $\Delta(t)$. Logo, série temporal resultante $g_{obs}(t)$ é dada por:

$$g_{obs}(t) = g(t)III_{\Delta t}(t) \quad (2.10)$$

Utilizando a Eq. 2.8, obtemos a transformada de Fourier desta série temporal:

$$\hat{g}_{obs}(f) = \sum_{n=-\infty}^{\infty} g(n\Delta t)e^{i2\pi f n \Delta t} \quad (2.11)$$

onde, de acordo com os limites de Nyquist (SHANNON, 1949), a frequência f deve pertencer ao intervalo $0 \leq f \leq \frac{1}{\Delta t}$. Entretanto, quando tratamos de casos reais, não temos uma quantidade infinita de dados, mas um número finito N de observações. Sendo assim, definindo $g_n \equiv g(n\Delta t)$ e $\hat{g}_k \equiv \hat{g}_{obs}(k\Delta f)$, onde $\Delta f = \frac{1}{N\Delta t}$, podemos reescrever a Eq. 2.10 como

$$\hat{g}_k = \sum_{n=0}^{N-1} g_n e^{i2\pi k n / N} \quad (2.12)$$

que é a forma padrão da DFT.

2.1.1.3 Fast Fourier Transform (FFT)

Apesar da sua grande utilidade em uma enorme variedade de campos, o custo computacional para o cálculo da transformada discreta de Fourier a partir da sua definição é muito alto. Conforme visto na Eq. 2.11, cada transformada \hat{g}_k realiza N cálculos e cada \hat{g}_k é calculado N vezes. Temos assim, uma complexidade de $O(n^2)$, o que não é uma boa notícia. Para resolver esse problema, foi desenvolvido um algoritmo capaz de

realizar este cálculo com uma complexidade de apenas $O(n \log n)$. Este método foi chamado Transformada Rápida de Fourier, também conhecido como FFT.

A FFT não possui uma definição específica, podendo ser implementada de diversas maneiras. Apesar disso, um dos primeiros algoritmos desenvolvidos ainda é o mais utilizado. Popularizado por Cooley e Tukey em 1965 (COOLEY; TUKEY, 1965) e consequentemente chamado de *Cooley-Tukey*, o algoritmo decompõe recursivamente a DFT em várias DFTs menores, juntamente com o uso dos chamados *twiddle factors*, que servem para evitar cálculos de operações repetidas ao longo da DFT, assim como a realização de operações simples e.g. multiplicações por 0 e 1 (GENTLEMAN; SANDE, 1966).

Uma das implementações mais comuns do *Cooley-Tukey* é o caso *radix-2*. Possuindo a limitação de ser aplicado apenas em transformadas de tamanho 2^n , o algoritmo divide a DFT em duas DFTs de tamanho $N/2$ a cada passo da recursão.

O seu pseudocódigo é dado por (JOHNSON; FRIGO, 2009):

```

1:  $\mathbf{Y}[0, \dots, n-1] \leftarrow \text{fft2}(n, \mathbf{X}, s)$ :
2: if  $n = 1$  then
3:    $Y[0] \leftarrow X[0]$ 
4: else
5:    $\mathbf{Y}[0, \dots, n/2 - 1] \leftarrow \text{fft2}(n/2, \mathbf{X}, 2s)$ 
6:    $\mathbf{Y}[n/2, \dots, n - 1] \leftarrow \text{fft2}(n/2, \mathbf{X} + s, 2s)$ 
7:   for  $k_1 = 0$  to  $(n/2) - 1$  do
8:      $t \leftarrow \mathbf{Y}[k_1]$ 
9:      $\mathbf{Y}[k_1] \leftarrow t + \omega_n^{k_1} \mathbf{Y}[k_1 + n/2]$ 
10:     $\mathbf{Y}[k_1 + n/2] \leftarrow t - \omega_n^{k_1} \mathbf{Y}[k_1 + n/2]$ 
11:   end for
12: end if

```

Apesar da limitação de tamanho dita anteriormente, existem variações deste caso chamadas de *mixed-radix*, onde diferentes tipos de fatorização podem ser aplicados.

2.2 Periodograma Lomb-Scargle

A partir da DFT, definida na Eq. 2.12, podemos aplicar a definição da densidade de potência espectral da Eq. 2.7 para obter o que é chamado tanto de *periodograma clássico* como de *periodograma de Schuster* (SCHUSTER, 1898):

$$P_S(f) = \frac{1}{N} \left| \sum_{n=1}^N g_n e^{-2\pi i f t_n} \right|^2 \quad (2.13)$$

onde N é o número de medidas realizadas. A partir da fórmula de Euler, podemos reescrever a Eq. 2.13 como:

$$P(f) = \frac{1}{N} \left[\left(\sum_n g_n \cos(2\pi f t_n) \right)^2 + \left(\sum_n g_n \sin(2\pi f t_n) \right)^2 \right] \quad (2.14)$$

Inicialmente, foi desenvolvido para estimar através de várias observações, o espectro de potência de um sinal. Infelizmente, na natureza nem sempre os dados são igualmente espaçados e nesses casos, os resultados obtidos pelo periodograma não podem ser expressados de maneira geral e analítica. Foi então que o astrofísico Jeffrey D. Scargle percebeu o problema e propôs uma correção para o mesmo (SCARGLE, 1982), tornando o periodograma invariante ao tempo e considerando intervalos irregulares de medida. A partir de uma forma mais generalizada do periodograma clássico,

$$P(f) = \frac{A^2}{N} \left(\sum_n g_n \cos(2\pi f[t_n - \tau]) \right)^2 + \frac{B^2}{N} \left(\sum_n g_n \sin(2\pi f[t_n - \tau]) \right)^2, \quad (2.15)$$

onde A, B e τ são funções arbitrárias em f e nos tempos de observação t_n . Scargle mostrou que ao escolher valores únicos para essas funções, poderíamos obter um periodograma que seria insensitivo a variações temporais nos dados e que em caso de observações feitas de forma constante. A partir dessas propriedades, foi obtido o que ficou conhecido como *Periodograma de Lomb-Scargle*:

$$P_{LS}(f) = \frac{1}{2} \left\{ \frac{\left(\sum_n g_n \cos(2\pi f[t_n - \tau]) \right)^2}{\sum_n \cos^2(2\pi f[t_n - \tau])} + \frac{\left(\sum_n g_n \sin(2\pi f[t_n - \tau]) \right)^2}{\sum_n \sin^2(2\pi f[t_n - \tau])} \right\} \quad (2.16)$$

onde, para garantir a invarância temporal, τ é especificado como:

$$\tau = \frac{1}{4\pi f} \tan^{-1} \left(\frac{\sum_n \sin(4\pi f t_n)}{\sum_n \cos(4\pi f t_n)} \right). \quad (2.17)$$

Uma das razões da importância do uso do Periodograma de Lomb-Scargle é que, para cada f , possui um resultado idêntico ao da distribuição χ^2 de uma interpolação com uma curva senoidal. O trabalho de Scargle foi feito como uma continuação do trabalho de Lomb (1976) e por isso o método leva o nome de ambos.

2.3 Estatística Descritiva

Ao analisar os dados de uma série temporal, muitas vezes é necessário organizar os dados para melhor descrevê-los e até mesmo encontrar alguma tendência entre eles. Para tal fim, existe a estatística descritiva, que é um ramo da estatística cujo objetivo é sintetizar um conjunto de observações para transmitir a maior quantidade de informações da maneira mais simples possível. As medidas mais comumente utilizadas são as de tendência central¹ e as de variabilidade ou dispersão².

2.3.1 Média

A média pode ser definida como a aritmética simples de um conjunto de observações, sendo ela uma medida sensível aos valores amostrados, o seu uso é mais adequado quando

¹Média, mediana e moda

²Desvio padrão e variância

os dados estão distribuídos de forma uniforme e sem grandes discrepâncias. O seu cálculo é dado por:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.18)$$

onde x_i são os dados observados e n é a quantidade destes.

2.3.2 Mediana

Um dos conceitos mais utilizados em síntese estatística, a mediana representa o valor central de um conjunto de dados. Para obter a mediana, basta ordenar os valores amostrados de maneira crescente ou decrescente e obter aquele está no centro. Caso o número de elementos seja par, é então calculada a média dos dois valores do meio. Uma das principais vantagens da mediana é que, ao contrário da média, o seu valor é mais resistente à poluições na amostra e.g. *outliers*³.

2.3.3 Moda

Há casos em que tanto a média como a mediana não servem para descrever qual é a tendência dos dados, principalmente quando eles não são numéricos. Para situações como essas, temos o conceito de moda. A moda representa o valor que aparece mais frequentemente em um conjunto de observações e pode ser utilizada tanto para valores numéricos quanto, e principalmente, para valores nominais. É uma medida extremamente robusta devido à sua insensibilidade aos *outliers*. Um conjunto de dados pode possuir duas modas, nesse caso diz-se que ele é um conjunto bimodal.

2.3.4 Desvio Padrão e Variância

O desvio padrão é uma medida de dispersão da estatística descritiva. É utilizada para descobrir se os valores estão ou não próximo da média. A forma mais comum de realizar o seu cálculo é através da seguinte fórmula:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2} \quad (2.19)$$

onde N é a quantidade de dados observados e μ é a média aritmética do conjunto.

Já a variância, representa o quão longe os valores estão da média. Sendo o seu valor o quadrado do desvio padrão.

³Observação que apresenta um grande afastamento das demais da série (MADDALA, 1992).

2.4 Minimização da Dispersão de Fase

Dentre as técnicas de análise de periodicidade em séries temporais, a minimização da dispersão de fase, ou PDM (*Phase Dispersion Minimization*), atua onde métodos baseados em Fourier não podem ser utilizados, como, por exemplo, séries com espaçamento irregular entre observações. Essa técnica foi desenvolvida para ser uma versão automatizada de outra técnica já utilizada anteriormente conhecida como *data-folding*, que consiste em estimar vários períodos de teste para a série e dividir os dados em várias subséries com uma duração igual ao período sendo testado. Os dados são então traçados em relação à fase, ou em uma escala de 0 a 1 relativa ao período de testes. O período escolhido será aquele que produzir menos dispersamento em relação à curva de luz média (STELLINGWERF, 1978). A principal vantagem desta técnica em relação à outras como FFT ou o Periodograma Lomb-Scargle, é que a PDM não utiliza uma função específica, como seno ou cosseno, para realizar o ajuste da curva.

2.4.1 Definições

Um conjunto discreto de observações, ou uma série temporal, pode ser representado por dois vetores: as magnitudes \mathbf{x} e os tempos de observação \mathbf{t} , onde a i -ésima observação é dada por (x_i, t_i) , havendo um total N pontos onde $i = (1, N)$. Sendo σ^2 a variância de \mathbf{x} , dada por

$$\sigma^2 = \frac{1}{N-1} \sum (x_i - \mu)^2 \quad (2.20)$$

onde μ é a média, tal que $\mu = \frac{1}{N} \sum x_i$. Como consequência da Equação 2.20, temos a variância amostral total s^2 definida por

$$s^2 = \frac{\sum (n_j - 1) s_j^2}{\sum n_j - M} \quad (2.21)$$

sendo M a quantidade de amostras distintas, que possuem variâncias s_j^2 , com $j = (1, M)$ e n_j pontos.

É desejado reduzir a variância dos dados em respeito à curva de luz média. Então, sendo Π um período de teste, é computado o vetor de fases $\Phi : \phi_i = t_i/\Pi - [t_i/\Pi]$, onde os colchetes indicam a parte inteira. O vetor de fases também pode ser escrito como $\Phi = \mathbf{t} \bmod (\Pi)$. Então, são retiradas M amostras de \mathbf{x} utilizando o critério de que todos os elementos da amostra j devem possuir similar fase ϕ_i . A variância das amostras nos dá uma medida da dispersão ao redor da curva de luz média definida pelas médias de x_i em cada amostra, considerada uma função de ϕ . Nós então definimos a estatística

$$\Theta = \frac{s^2}{\sigma^2} \quad (2.22)$$

onde σ^2 é dado pela Equação 2.20 e s^2 é dado pela Equação 2.21. Caso Π não seja um período válido, então $s^2 \approx \sigma^2$ e $\Theta \approx 1$, caso contrário, Θ vai cair em um mínimo local em relação aos períodos vizinhos, provavelmente com um valor próximo de zero.

Esta técnica foi nomeada de "minimização da dispersão de fase"(PDM) pois visa reduzir a dispersão dos dados em uma fase constante. Matematicamente, nada mais é que uma aproximação por mínimos quadrados, mas ao invés de realizar uma aproximação para uma curva dada, a aproximação é relativa à curva média que foi definida pelas médias de cada intervalo (ou *bin*). A PDM, portanto, é um "Fouriergrama" de ordem infinita, já que todos os harmônicos estão inclusos na função ajustada. A série de Fourier, que também é um método de mínimos quadrados, geralmente requer restrições adicionais ao invés de ordens elevadas para variações não-senoidais.

2.5 Bibliotecas e Ferramentas

Este trabalho foi implementado em Python, uma linguagem de programação interpretada e orientada a objetos. Os principais motivos de ter sido escolhida para o projeto foram a grande quantidade de bibliotecas que simplificam o desenvolvimento de rotinas complexas, o fato de possuir uma programação mais interativa, que ajuda bastante na hora de encontrar erros, e a sua vasta comunidade de desenvolvedores, que ajuda bastante com dúvidas pontuais. As principais ferramentas utilizadas serão brevemente descritas abaixo.

2.5.1 Jupyter Notebook

Aplicação que permite a execução dos códigos de forma interativa, facilitando a visualização do processo em ação⁴. É rodada diretamente no navegador e normalmente, de forma local, ou seja, com o computador do usuário agindo como servidor.

2.5.2 NumPy

Principal biblioteca da linguagem para trabalhar com operações numéricas e científicas⁵. Possui estruturas de dados que facilitam o trabalho com conceitos de álgebra linear, como vetores e matrizes, de forma que o processamento dos dados nessas estruturas utilize menos recursos computacionais e seja mais eficiente (WALT; COLBERT; VAROQUAUX, 2011). Entre as utilidades da biblioteca pode-se incluir a geração de números aleatórios, a realização cálculos envolvendo a transformada de Fourier e também a integração de códigos em C/C++ e Fortran.

⁴<https://jupyter.org/>

⁵<https://www.numpy.org/>

2.5.3 SciPy

A biblioteca SciPy⁶ foi desenvolvida em cima da NumPy para fazer uso das suas ferramentas básicas, de forma que operações mais complexas possam ser realizadas, como a manipulação de sinais, interpolação de funções e o uso de matrizes esparsas.

2.5.4 Pandas

Biblioteca *open source* que fornece estruturas e ferramentas de fácil uso para análise de dados⁷, como o DataFrame, que permite uma visualização dos dados em formato de tabela, além de simplificar a sua manipulação. Várias operações são realizadas em conjunto com as bibliotecas NumPy e SciPy.

2.5.5 Matplotlib e Seaborn

Matplotlib é uma biblioteca utilizada para a visualização de gráficos 2D⁸, permitindo, de forma simples, a geração de gráficos de linha, histogramas, gráficos de dispersão, etc. Já a biblioteca Seaborn⁹, pode ser usada como complemento à Matplotlib. Ela fornece mais formatos de gráficos, assim como uma grande variedade de formas de estilizar a visualização.

2.5.6 AstroPy

Biblioteca cujo objetivo é reunir ferramentas úteis para realizar operações relacionadas à astronomia, assim como para serem utilizadas como base para outras implementações¹⁰. Possui utilidades como a leitura de arquivos da extensão .fits e de tabelas ASCII. Também possui implementações de melhorias de ferramentas de outras bibliotecas, como a cálculo de convoluções do SciPy, que aqui possui um tratamento diferenciado de valores nulos¹¹. Neste trabalho, a biblioteca foi utilizada para o cálculo do período através do Periodograma Lomb-Scargle e para a leitura de arquivos com extensão .tbl, que são tabelas ASCII.

2.5.7 pwkit

Biblioteca que reúne várias ferramentas desenvolvidas por Williams et al. (2017) para realizar cálculos científicos e astronômicos. Ela possui uma variedade de ferramentas,

⁶<https://www.scipy.org/>

⁷<https://pandas.pydata.org/>

⁸<https://matplotlib.org/>

⁹<https://seaborn.pydata.org/>

¹⁰<https://www.astropy.org/about.html>

¹¹<http://docs.astropy.org/en/stable/convolution/index.html>

que varia desde a conversão para notação sexagesimal¹², até cálculos de propriedades de anãs ultra-frias¹³. Esta biblioteca foi utilizada neste trabalho para calcular períodos através da minimização da dispersão de fase.

¹²<https://pwkit.readthedocs.io/en/latest/science/basic-astro/>

¹³https://pwkit.readthedocs.io/en/latest/science/pwkit-ucd_physics/

3 DADOS OBSERVACIONAIS

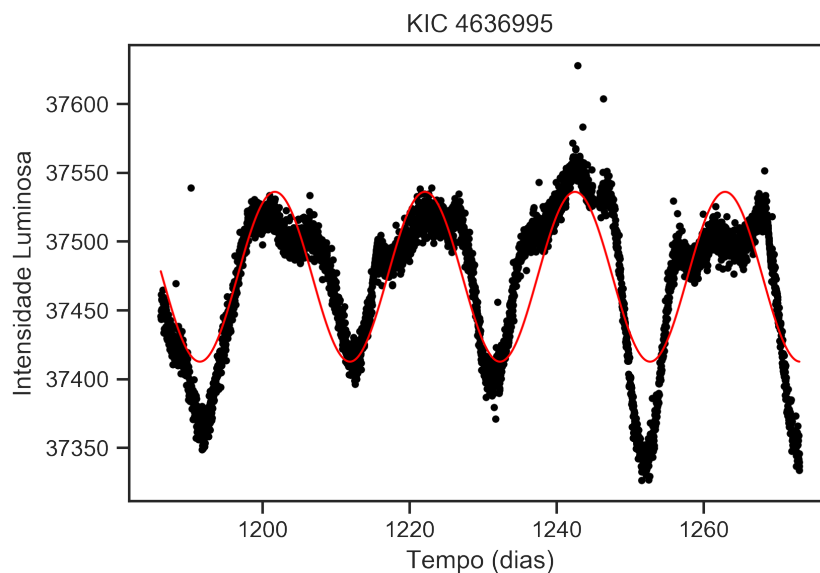
Este trabalho tem como objetivo o desenvolvimento de uma ferramenta que facilite o cálculo do período de rotação estelar ao realizá-lo de forma automatizada. Logo, é necessário um entendimento sobre a periodicidade das curvas de luz, assim como os seus tipos. Em seguida serão explicados os tipos de estrelas, suas características, além de justificar a aplicação dos métodos apresentados no Capítulo 2.

3.1 *Curvas de luz*

Curvas de luz são séries temporais que indicam a intensidade luminosa emitida por um objeto ou corpo celeste, seja ele uma estrela ou até mesmo uma galáxia. Normalmente, o sinal da curva de luz possui um formato senoidal, como pode ser observado na figura 3.1. Ao realizar o ajuste da curva, é obtida uma senoide pura, permitindo o cálculo do período ao aplicar a FFT sobre ela.

A análise da periodicidade de tais curvas possui uma complexidade variável de acordo com o conhecimento que se tem sobre o que está sendo observado. Portanto, é necessário entender as causas mais comuns dessa periodicidade.

Figura 3.1 – Exemplo de curva de luz e seu ajuste.



Fonte – *Elaborada pelo autor.*

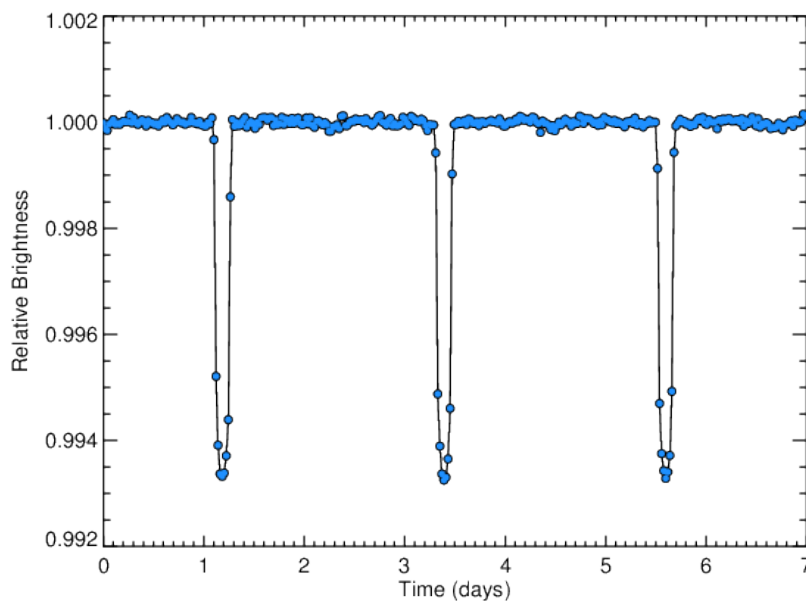
3.1.1 Manchas Estelares

A principal razão para a oscilação no brilho de uma estrela, é a presença de manchas em suas superfícies. Conforme observado por Galileu e outros cientistas de sua época, o Sol, nossa principal estrela, possuía manchas em sua superfície externa e através delas era possível não apenas detectar que havia um movimento de rotação no astro, como também obter a periodicidade com que tal movimento era realizado em sua plenitude (MITCHELL, 1916).

3.1.2 Trânsito

Outra forma de periodicidade detectada em curvas de luz é o trânsito planetário, só que nesse caso o período observado é o tempo que o provável planeta leva para dar uma volta completa ao redor da estrela. Na figura 3.2 é possível ver como o trânsito de um objeto afeta a curva de luz obtida.

Figura 3.2 – Demonstração de trânsito planetário em uma curva de luz. Pode-se observar que a cada dois dias, aproximadamente, um corpo passa em frente à estrela.



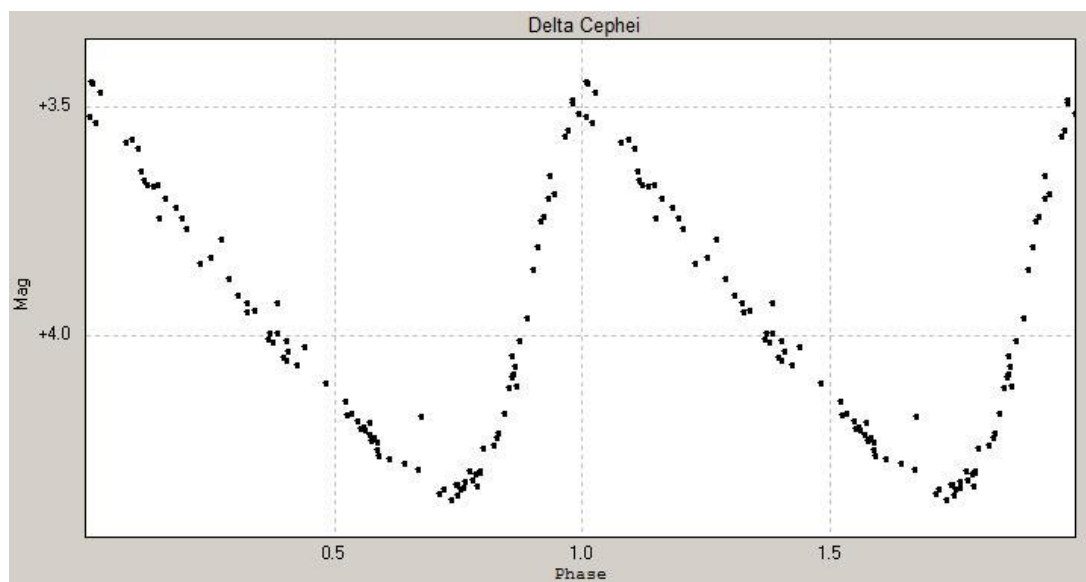
Fonte – Cfa/Harvard - <https://www.cfa.harvard.edu/avanderb/tutorial/tutorial2.html>

3.1.3 Pulsação

Normalmente, as estrelas possuem uma luminosidade praticamente constante, com pouca ou quase nenhuma variação ao longo do tempo. Porém, existem certos tipos de estrelas que seu brilho varia de forma bem perceptível e em um relativamente curto intervalo de tempo. Essas estrelas, chamadas de *variáveis*, são categorizadas e agrupadas

de acordo com a razão da sua variação de luminosidade. Um desses grupos é o das *variáveis pulsantes*, cuja variação na luminosidade decorre da expansão e contração radial da estrela. Na maioria dos casos, ao analisar uma curva de luz, pode ser observada uma periodicidade na pulsação da estrela, como pode ser observado na figura 3.3.

Figura 3.3 – Curva de luz de uma estrela variável Delta Cephei, onde é possível observar a periodicidade das pulsações que ocorrem no astro.



Fonte – https://en.wikipedia.org/wiki/Stellar_pulsation#/media/File:Delta_Cephei_lightcurve.jpg

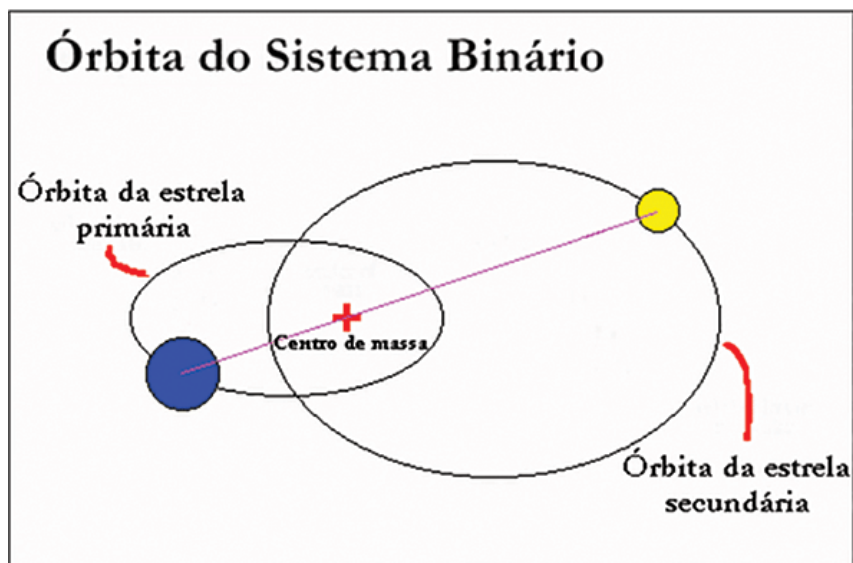
3.1.4 Sistemas Binários

Sistemas estelares são chamados de binários, ou estrelas binárias, quando são formados por duas estrelas que orbitam o mesmo centro de massa. Um esquema que representa um sistema desse tipo pode ser visto na figura 3.4.

As estrelas binárias podem ser classificadas em:

- *Binárias Visuais* são aquelas cuja binariedade pode ser observada diretamente a partir de telescópios, sejam eles terrestres ou espaciais. Tal observação é possível quando as estrelas estão distantes o suficiente entre si, possuem brilho suficiente e/ou estão próximas o suficiente da Terra para serem observadas separadamente. Albireo (figura 3.5), Mira e Sirius são exemplos de binárias visuais.
- *Binárias Espectroscópicas* são sistemas onde a distância entre as estrelas é muito pequena e a velocidade orbital delas é muito alta. Como não é possível observá-las visualmente, essas binárias são detectadas a partir da observação da variação da sua velocidade radial, que é feita a partir de uma análise espectroscópica da luz emitida para examinar o efeito Doppler nas linhas espectrais da estrela.

Figura 3.4 – Representação da órbita de um sistema estelar binário.



Fonte – R.C. Mello, Daniel. (2014). *Aprendendo física com as estrelas binárias*. *Revista Brasileira de Ensino de Física*. 36. 1-9. 10.1590/S1806-11172014000300007.

Figura 3.5 – Estrela binária Albireo.

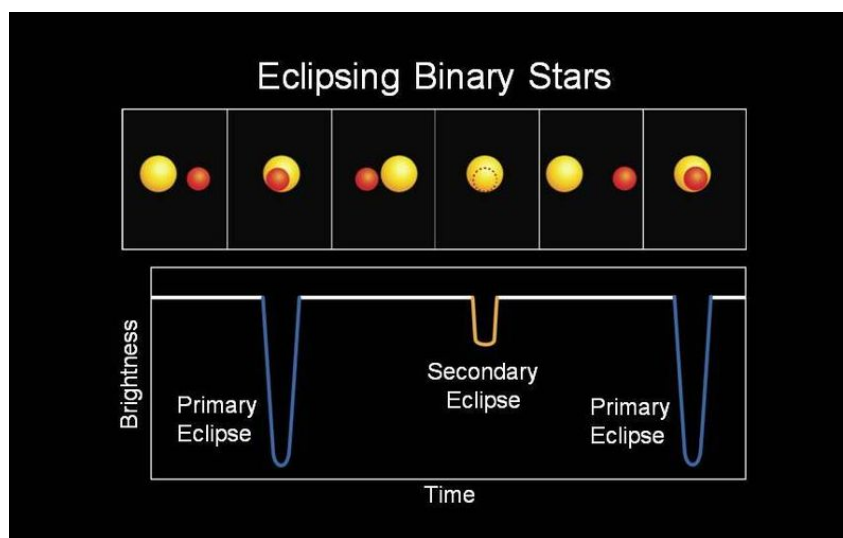


Fonte – APOD: 30 de agosto de 2005 - Albireo: A Bright and Beautiful Double. NASA.

- *Binárias Astrométricas* são aquelas onde apenas uma das estrelas pode ser observada através de telescópios, devido à baixa luminosidade da segunda estrela. A presença dessa companheira se faz notável devido ao deslocamento na órbita da estrela mais brilhante, que ocorre graças à atração gravitacional que ocorre entre elas.
- *Binárias Eclipsantes* são estrelas que estão em um plano orbital próximo da nossa linha observacional, assim, para um observador terrestre, essas estrelas realizam

eclipses mútuas durante as suas órbitas. Tais eclipses resultam em variações no brilho dessas estrelas, como pode ser observada na figura 3.6. A análise da curva de luz dessas estrelas permite obter o período orbital e o tamanho relativo de cada estrela do sistema.

Figura 3.6 – Representação da curva de luz de uma binária eclipsante associada a cada evento de sua órbita.



Fonte – NASA (2013).

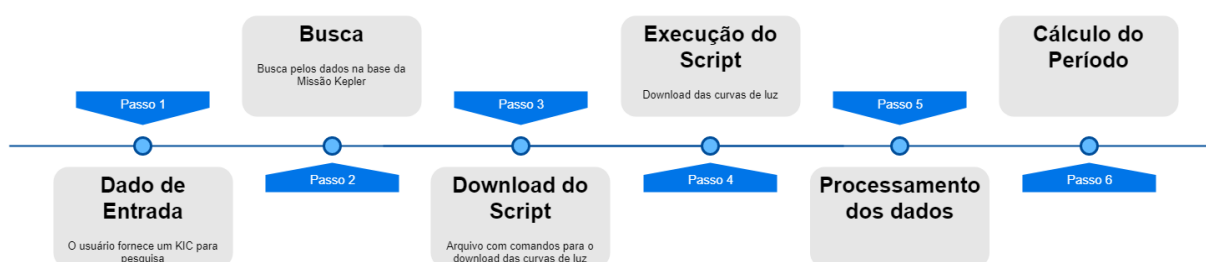
3.2 Escolha dos dados a serem analisados

No desenvolvimento deste trabalho, foram utilizados os dados obtidos pelo Kepler e que estão disponíveis. Desses dados, foram filtradas aleatoriamente mil estrelas do tipo Solar, que são aquelas que, de acordo com Chaplin et al. (2014), possuem uma temperatura efetiva (T_{eff}) entre $5200K$ e $6060K$, e uma gravidade superficial ($\log g$) entre $3.63 dex$ e $4.40 dex$. Foram eliminadas as estrelas que possuem algum tipo de trânsito, seja planetário ou estrelas binárias. É importante frisar que os dados da missão estão divididos em dezessete "quartos", que são os períodos de observação. Cada estrela, possui várias curvas de luz, onde cada uma é referente à cada um desses "quartos". Desses dezessete, foram removidos da análise os primeiros três quartos devido à sua curta duração, assim como defeitos instrumentais (MCQUILLAN; MAZEH; AIGRAIN, 2014).

3.3 Processamento e análise dos dados

O fluxo principal do algoritmo do processo implementado neste trabalho é mostrado na figura 3.7 e será detalhado a seguir.

Figura 3.7 – Fluxo principal do algoritmo.



Fonte – *Elaborada pelo autor.*

3.3.1 Entrada e busca de dados

A aquisição manual dos dados é feita a partir do formulário mostrado na figura 3.8, que se encontra no site da missão¹. Nele, é passado apenas o ID (código de identificação) da estrela. Para poder implementar essa etapa do processo, foi necessário descobrir as ações executadas ao clicar no botão de *Download*.

Figura 3.8 – Formulário de obtenção dos dados.

Search Kepler Stellar Data

The Kepler Stellar data set is too large to display effectively in an interactive table, so use the search form to specify a data subset. Results are returned in a new window. Clicking Submit Search without any constraints will return the complete data set, but note that performance will be compromised. By default, queries search the complete stellar table, including Q1-12, Q1-Q16, Q1-17 DR24, Q1-17 DR 25, and Q1-Q17 DR 25 Supplemental data. Download the complete Q1-Q17 DR 25 table [here](#), or see the [Kepler Stellar documentation page](#) for more download options.

Time Series Lookup

Enter a Kepler ID (integer only) to view its time series, or to generate a download script for all related time series files.

Kepler ID

Fonte – *Elaborada pelo autor.*

Inspecionando o formulário, foi possível descobrir qual requisição é feita e quais os parâmetros utilizados, como pode ser visto na figura 3.9.

Utilizando esses parâmetros, foi montado o código abaixo, que realiza uma requisição HTTP do tipo POST². A variável 'kic' é o código da estrela no *Kepler Input Catalog*, que é um banco de dados público que foi utilizado para armazenar os dados obtidos pelo satélite. 'BASE_URL' é o endereço do site da missão.

¹<https://exoplanetarchive.ipac.caltech.edu/>

²Protocolo que envia dados a um servidor.

Figura 3.9 – Detalhes dos parâmetros do formulário.

```

▼<form name="dwnlld" onsubmit="return onSubmitDWNLD()
;" method="post" action="/cgi-bin/IERDownload/nph-
IERDownload" target="_blank" enctype="multipart/
form-data"> == $0
  <input type="hidden" name="id" value="8279274">
  <input type="hidden" name="inventory_mode" value=
  "id_single">
  <input type="hidden" name="idtype" value=
  "source">
  <input type="hidden" name="dataset" value=
  "Kepler">
  <input type="hidden" name="resultmode" value=
  "webpage">
  <input class="dwnlldButton" type="submit" name=
  "submit" value="Download">
</form>

```

Fonte – Desenvolvido pelo autor.

```

r = requests.post(BASE_URL+"/cgi-bin/IERDownload/nph-IERDownload",
                  data={'id': kic,
                        'inventory_mode': 'id_single',
                        'idtype': 'source',
                        'dataset': 'kepler',
                        'resultmode': 'webpage'})

```

3.3.2 Download e execução do script

No fluxo normal da página, após a consulta o usuário é redirecionado para a página reproduzida na figura 3.10. Já no código, a requisição retorna uma *string* com o conteúdo HTML da página. Para obter o endereço de download, foi necessário realizar outra inspeção da página para descobrir a localização do mesmo. Sabendo que o *link* se encontra dentro de uma *tag* `<big>`, bastou apenas filtrar o texto de resposta para obter apenas o que estava dentro da *tag* e retirar o endereço. Para realizar esse procedimento, foi escrita a função abaixo, que recebe o texto completo da página e retorna o *link*.

```

def get_download_url(page_text):
    big = page_text[page_text.find('<big>'):page_text.find('</big>')]
    link = big[big.find('/cgi'):big.find('t')+1]
    return link

```

Figura 3.10 – Reprodução da página com o link de download



Fonte – Elaborada pelo autor.

Após a obtenção do endereço, é realizada uma requisição HTTP do tipo GET³. O retorno da requisição é um arquivo `.bat`, que possui um conjunto de comandos sequenciais. Nesse caso específico, os comandos presentes no arquivo executam o programa Wget⁴ com um endereço como parâmetro para, finalmente, realizar o *download* dos arquivos contendo as curvas de luz.

3.3.3 Processamento dos dados

Ao lidar com uma grande quantidade de dados é necessário filtrá-los para manter apenas as informações essenciais, pois assim reduz o espaço ocupado, além de deixar a visualização mais limpa. Por mais que a diferença individual de tamanho seja mínima, ao observar o conjunto inteiro, se torna bem significativa.

Os arquivos obtidos através dos scripts vêm em dois formatos: `.fits` e `.tbl`. Os arquivos `.fits`, apesar de possuírem tamanho menor em relação ao outro formato, estão codificados de uma forma que dificulta a conversão para estrutura de dados do tipo `DataFrame`⁵, portanto, foram utilizados os arquivos no formato `.tbl`.

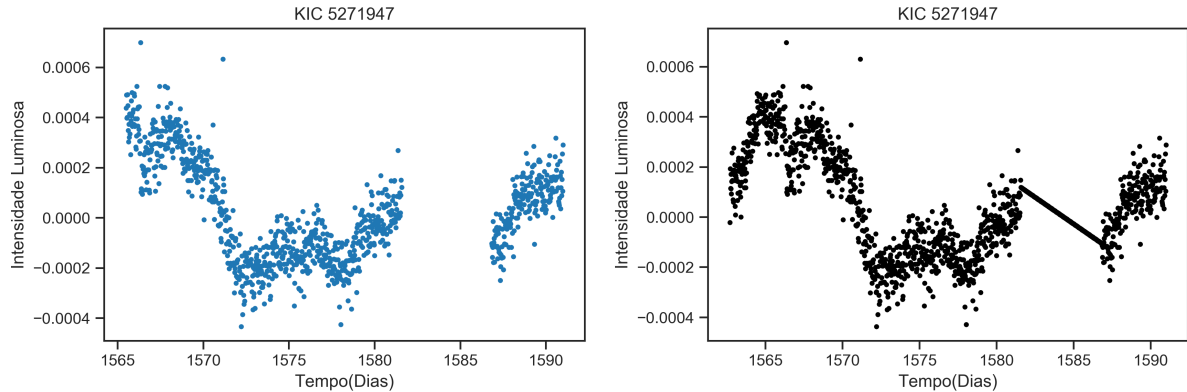
O processamento é iniciado ao ler os dados do arquivo `.tbl` através da estrutura `DataFrame`, que pode ser interpretado como uma tabela. Logo após é realizada uma interpolação linear como uma forma de preencher os dados que faltam na curva de luz. Apesar de os métodos utilizados não serem afetados por esses espaçamentos, existem métodos que são sensíveis à essas irregularidades e se fossem utilizados, mesmo que de forma auxiliar, o resultado poderia divergir. Assim, esses espaços são preenchidos de forma linear com pontos baseados nas observações anteriores; a interpolação foi feita de forma linear pois é a que melhor se ajusta à curva sem alterar o seu formato, como pode ser vista na figura 3.11. Então a tabela é reescrita apenas com as colunas que fazem referência ao

³Protocolo que recupera dados.

⁴Programa livre que realiza download de arquivos através de um endereço fornecido pelo usuário.

⁵Da biblioteca Pandas. O seu uso facilita a visualização e manipulação de dados.

Figura 3.11 – Trecho da curva de luz para KIC 5271947 antes (direita) e depois (esquerda) da interpolação linear. Os períodos resultantes foram 19.63 dias e 19.65 dias, antes e depois do preenchimento dos dados, respectivamente, mostrando que os métodos utilizados lidam bem com dados irregularmente espaçados.



Fonte – *Elaborada pelo autor.*

tempo ou à intensidade luminosa. Para uma melhor visualização, as colunas referentes à luminosidade têm seus valores normalizados entre 0 e 1 e armazenados em novas colunas, de forma que os dados originais sejam mantidos. Por fim, a nova tabela é salva em um arquivo .csv, devido à sua facilidade de leitura e por necessitar de menos armazenamento que os arquivos .tbl. Todos os arquivos originais são deletados durante o processo.

Abaixo está o código que realiza o processo descrito anteriormente.

```
def process_data(folder_path):
    columns=['TIME', 'SAP_FLUX', 'PDCSAP_FLUX', 'SAP_FLUX_ERR', 'PDCSAP_FLUX_ERR', 'CADENCE']
    filenames_tbl = get_filenames(folder_path, 'tbl')
    filenames_fits = get_filenames(folder_path, 'fits')
    for filename in filenames_fits:
        os.remove(folder_path+filename)
    for idx, filename in enumerate(filenames_tbl):
        data = ascii.read(folder_path + filename).to_pandas()
        data = data.interpolate(method='linear', columns=columns)
        data = data[columns].dropna()
        data = normalize_data(data)

        data = data.dropna()
        data.to_csv(folder_path+filename.replace('.tbl', '.csv'), index=False)
        os.remove(folder_path+filename)
```

3.3.4 Cálculo do período

O cálculo do período de rotação é dividido em etapas. Primeiramente são obtidas as frequências de cada quarto. Esse cálculo é feito utilizando o periodograma Lomb-Scargle.

Nesse algoritmo, um dos parâmetros necessários é um *range*, ou conjunto, de frequências que será utilizado para obter a que melhor se encaixa na curva de luz. Para obter esse conjunto, foi usado um método semelhante ao que é utilizado pela ferramenta online da Caltech⁶. São necessários três parâmetros: frequência mínima, frequência máxima e o intervalo entre cada elemento do conjunto, sendo as frequências calculadas a partir de períodos. O período mínimo é definido como o dobro do intervalo tempo entre as observações, sendo algo próximo da frequência de amostragem de *Nyquist* (SHANNON, 1949). O período máximo é todo o intervalo de tempo observado. Já o passo entre cada frequência é definido pela fórmula $(f_{max} - f_{min}) / (N * 10)$, onde N é o número de elementos da série temporal, f_{max} é o inverso do período mínimo e f_{min} é o inverso do período máximo. Abaixo está o código que realiza o cálculo desse conjunto.

```
def get_range_freqs(t):
    min_period = 2*(t[1] - t[0])
    max_period = t.max() - t.min()

    max_freq = 1/min_period
    min_freq = 1/max_period
    step_freq = (max_freq - min_freq)/(t.size*10)

    frequencies = np.arange(min_freq, max_freq, step_freq)
    return frequencies
```

Então, é obtida a frequência do quarto sendo analisado no momento utilizando o periodograma Lomb-Scargle.

```
def get_freq_LS(t, y, dy):
    frequencies = get_range_freqs(t)
    ls = LombScargle(t, y, dy)
    power = ls.power(frequencies)
    best_freq = frequencies[np.argmax(power)]
    return best_freq
```

Após obter as frequências de todos os quartos, elas são utilizadas para o cálculo do período de rotação utilizando a curva de luz inteira da estrela. Onde o período, que é o inverso da frequência, com a maior potência é o escolhido como o que melhor representa a curva de luz.

```
def get_period(t, y, dy, frequencies=None):
    if frequencies == None:
        frequencies = get_range_freqs(t)

    ls = LombScargle(t, y, dy)
    power = ls.power(frequencies)
```

⁶https://exoplanetarchive.ipac.caltech.edu/docs/pgram/pgram_parameters.html

```
best_freq = frequencies [np.argmax(power)]
period = round(1/best_freq,3)
return period
```

Também é calculado outro período utilizando o método descrito anteriormente, com um *range* de frequências, pois assim reduz as chances e *alias*. Caso o segundo período seja de um valor muito alto ou muito baixo, algo que pode acontecer ao analisar uma série temporal muito grande, é escolhido o primeiro, caso contrário é utilizado o método da minimização da dispersão de fase (PDM) para decidir qual dos períodos é o mais apropriado.

```
def get_period_pdm(t,y,dy,periods_list,nbins):
    periods = round_elements(periods_list,3)
    _pdm = pdm.pdm(t,y,dy,periods,nbins)
    period = _pdm.pmin
    return [period, _pdm.thetas[_pdm.imin]]
```

4 RESULTADOS

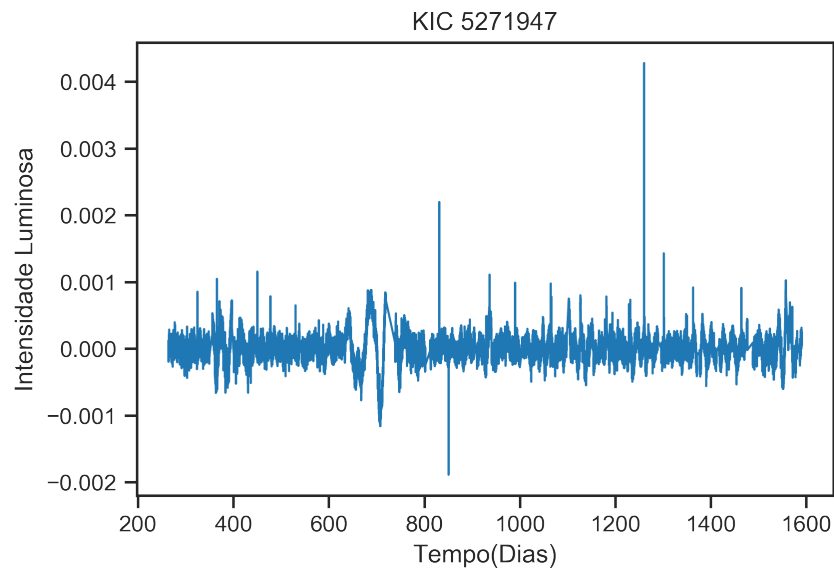
Neste capítulo serão apresentados os resultados referentes à aplicação desenvolvida, cujo objetivo é o de obter as curvas de luz e calcular períodos rotacionais de estrelas do tipo solar, de modo que o processo possa ser feito de forma automática e em grande escala. Inicialmente escolhemos de forma aleatória uma amostra de 1000 estrelas que se encaixam nos requisitos descritos na seção 3.2, durante o processamento dos dados, foram mantidas apenas as estrelas que possuíam curvas de luz a partir do quarto período de observação, pois os três primeiros possuem curta duração, além de defeitos instrumentais que afetam diretamente o resultado (MCQUILLAN; MAZEH; AIGRAIN, 2014). No total, foram 837 períodos calculados. Desta forma, analisamos o período de rotação para essas estrelas aplicando as seguintes técnicas em conjunto, como descrito na seção 3.3: o Periodograma Lomb-Scargle e a Minimização da Dispersão de Fase (PDM).

Na figura 4.1 apresentamos a curva de luz da estrela KIC 5271947, com o seu fluxo normalizado em relação ao tempo de observação. Nela, podemos ver algumas valores bem extremos que são causados por ruído instrumental. Realizamos uma tentativa de filtrar esses valores através de métodos como a média móvel, entretanto esse processo afetou significativamente os resultados obtidos quando comparados com os da literatura, portanto não houve um tratamento dos dados nesse sentido. Já na figura 4.2 apresentamos os periodogramas Lomb-Scargle obtidos ao realizar o cálculo tanto utilizando os períodos obtidos por quarto (figura da esquerda), como a curva de luz em sua totalidade (figura da direita). Nesse método, os períodos escolhidos são aqueles com maior energia. Portanto, foram escolhidos os valores 22.95 dias no cálculo por quarto e 19.65 dias utilizando a curva inteira. Por fim, na figura 4.3, temos o periodograma que foi produzido ao calcular a PDM para decidir entre os dois períodos obtidos. Nesse procedimento, o período selecionado é o que possui o menor valor de θ , logo 19.65 é o valor final para a curva de luz selecionada.

4.1 *Download e Processamento*

O download dos dados acabou sendo uma das partes do processo que levou mais tempo para ser finalizada, pois, apesar dos dados serem pequenos (cerca de 800 kB por quarto), há outro fator que influencia a velocidade da sua obtenção, que é a taxa de upload do servidor que está alocando esses dados. Apesar desse problema ser algo fora de controle, o fato de não precisar ir à página e obter um script de cada vez, já é uma otimização significativa, já que era algo feito 100% manualmente. Manualmente, a obtenção dos dados de uma estrela leva em torno de 220 segundos, já a automatizada foi de 200 segundos, o que consiste em um processo mais rápido. Apesar da diferença ser baixa, quando escala

Figura 4.1 – Curva de luz para KIC 5271947, com a intensidade luminosa normalizada em função do tempo de observação em dias.



Fonte – *Elaborada pelo autor.*

Figura 4.2 – Gráficos dos periodogramas lomb-scargle de ambos os cálculos realizados para KIC 5271947. Os períodos escolhidos foram, respectivamente, 22.95 dias e 19.65 dias.

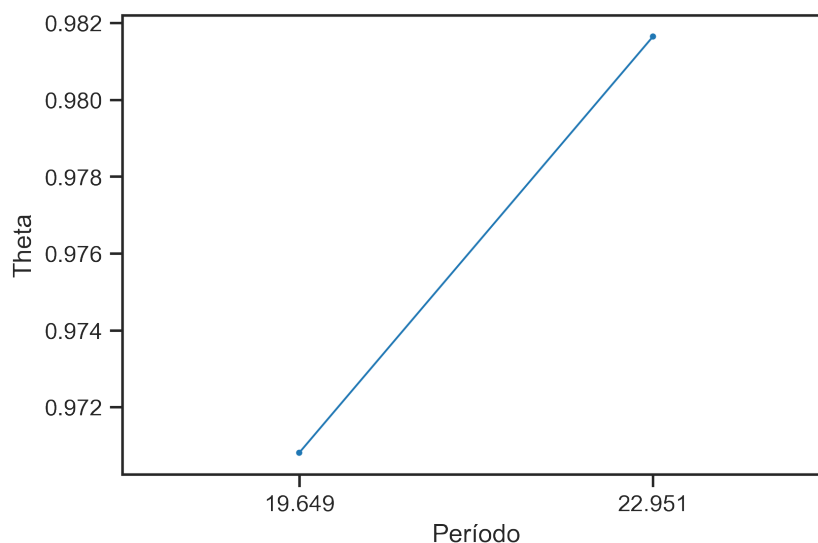


Fonte – *Elaborada pelo autor.*

esse processo para 1000 estrelas, como foi no caso desse trabalho, há uma diferença de tempo de aproximadamente 5.5 horas.

O processamento, que envolve várias etapas de manipulação de arquivos e dados, levou cerca de 30 segundos para cada elemento. No total, foram 8 horas de processamento para o conjunto completo de dados. Já o cálculo dos períodos levou em torno de 4 horas, levando em conta que os métodos utilizados são bastante complexos e utilizam uma grande quantidade de recursos computacionais. As configurações do computador utilizado durante a realização desse trabalho também são importantes para o tempo de processamento; a

Figura 4.3 – Gráfico do periodograma para o método da PDM realizado para KIC 5271947. O período escolhido foi aquele com o menor valor de Theta, ou seja, 19.65 dias.



Fonte – *Elaborada pelo autor.*

CPU consiste de um processador Intel Core i5 da 4^a geração, com uma velocidade de 3.00 GHz e a memória RAM é de 16 GB DDR3, com uma velocidade 1600 MHz. Em computadores mais potentes o resultado será bem mais satisfatório.

4.2 Validação dos Dados

Para validar os resultados obtidos, foram utilizados 191 resultados que também estão presentes em outros artigos, como base de comparação, sendo 51 deles provenientes de (NASCIMENTO et al., 2014) e uma parte deles podem ser vistos na Tabela 1. Os 140 restantes são originários de (MCQUILLAN; MAZEH; AIGRAIN, 2014) e uma amostra está na Tabela 2. Desse total, 17.3% foram *alias* e 4.2% possuíam algum valor que não foi aproximado ou *alias*. Isso pode ser bem observado na figura 4.4, onde apresentamos um gráfico $x=y$ em que os períodos obtidos estão plotados em relação aos períodos presentes na literatura. Os pontos acima da linha representam os períodos maiores que os determinados anteriormente, assim como os que estão abaixo representam os menores. Pode ser observado a maioria dos períodos de rotação obtidos neste trabalho estão muito próximos, ou até mesmo junto à linha, o que nos mostra que a maioria dos resultados estão em acordo com a literatura, e alguns poucos diferem.

Tabela 1 – Comparação dos resultados obtidos neste trabalho com os obtidos por (NASCIMENTO et al., 2014). A interrogação na coluna de Alias se refere a um erro desconhecido no valor obtido.

KIC	Alias	P_{Trab} (dias)	P_{Rot} (dias)	Erro rel.
3118654		15.9	15.6	1.92%
2718678		25.3	24.5	3.27%
5472881		6.6	6.5	1.54%
6605816	?	87.3	24.0	263.75%
8017116	x	10.3	20.8	50.5%
8046864		10.5	10.5	0.0%

Tabela 2 – Comparação dos resultados obtidos neste trabalho com os obtidos por (MCQUILLAN; MAZEH; AIGRAIN, 2014)

KIC	Alias	P_{Trab} (dias)	P_{Rot} (dias)	Erro rel.
757099		0.4	0.4	0.0%
3338421	x	2.3	4.8	52.1%
3437548		1.8	1.8	0.0%
3529147		14.4	14.2	1.41%
3869891	x	23.0	11.2	105.4%
8029079		25.5	28.0	8.93%

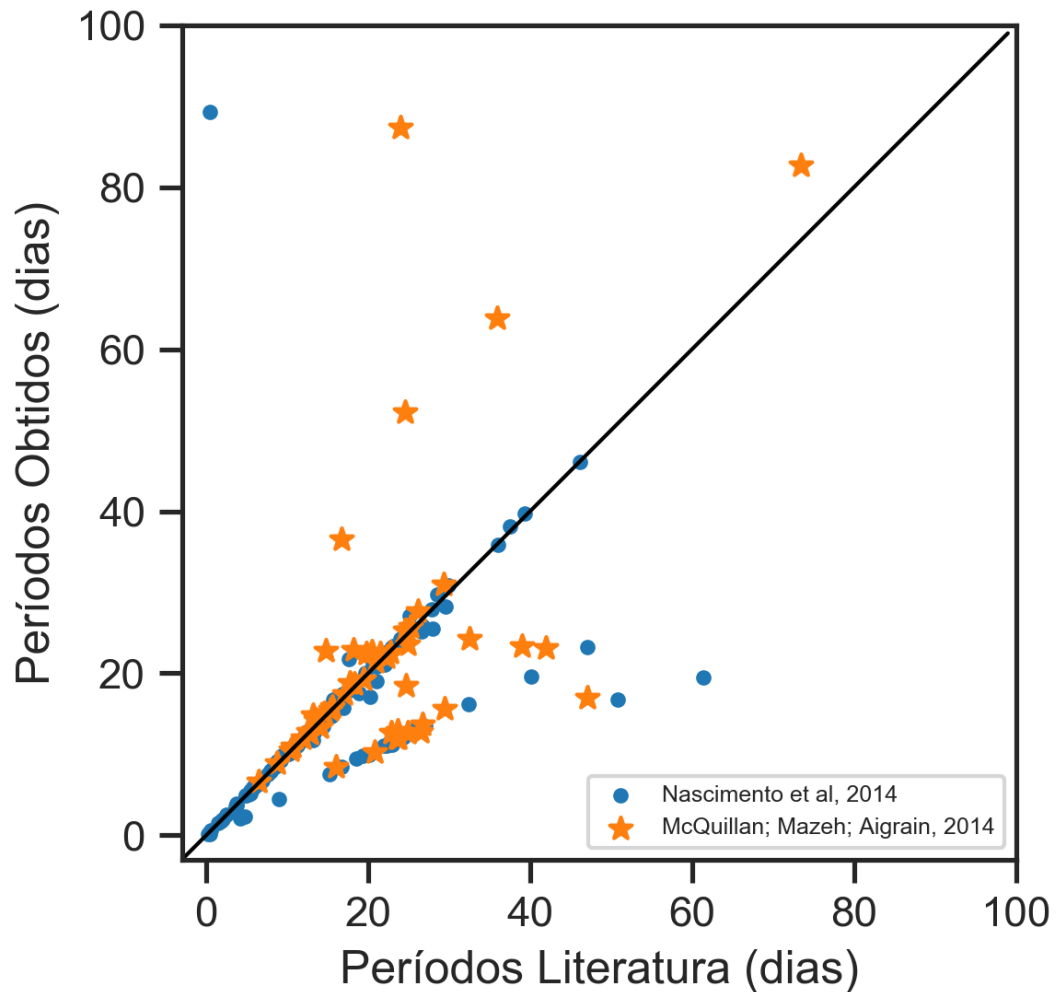
4.3 Períodos

Na Tabela 3 temos um resumo dos períodos obtidos neste trabalho, juntamente com os períodos apresentados por (MCQUILLAN; MAZEH; AIGRAIN, 2014) e na figura 4.5 temos um histograma de todos períodos obtidos neste trabalho.

4.4 Discussão

A automatização do processo de cálculo do período de rotação estelar se mostrou bem eficiente, produzindo resultados com um bom nível de proximidade com os presentes na literatura e em um tempo relativamente bom, visto que os métodos utilizados anteriormente tinham a verificação visual como parte do processo (NASCIMENTO et al., 2014; NIELSEN et al., 2013; MCQUILLAN; MAZEH; AIGRAIN, 2014), sendo isso algo que aumenta de forma considerável o tempo levado na obtenção dos resultados. Neste trabalho, boa parte do tempo de execução da ferramenta foi utilizado no processamento da PDM, que, aliás, foi utilizada em uma tarefa simples (decidir entre dois períodos previamente calculados)

Figura 4.4 – Gráfico mostrando a relação dos períodos obtidos com os presentes na literatura. Foram utilizadas 191 estrelas que são em comum com a literatura.



Fonte – Elaborada pelo autor.

justamente por causa do seu alto custo computacional. De acordo com a documentação¹ da biblioteca utilizada para realizar este cálculo, o tempo de execução é diretamente proporcional ao número de períodos amostrados passados para o algoritmo. Como o objetivo é realizar o processamento em um tempo consideravelmente rápido, não é viável utilizar esse método com uma quantidade maior de dados.

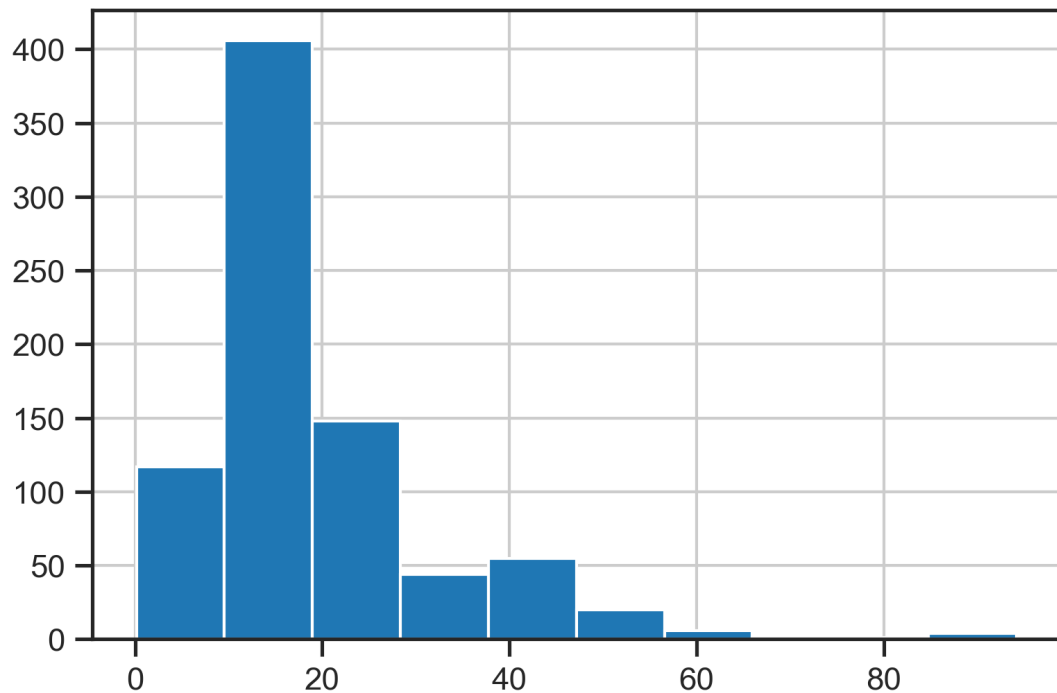
Como pode ser observado nas Tabelas 1, 2 e na figura 4.4, nem todos os períodos utilizados para validação foram exatamente iguais ou aproximados, porém grande parte desses resultados diferentes foram *alias* ou próximos aos *alias* presentes na literatura. A minoria que possui uma variação de resultado desconhecida, possivelmente foi devido à ruídos na curva de luz ou pela diferença na forma de obtenção dos períodos, visto que o método da Autocorrelação (MCQUILLAN; MAZEH; AIGRAIN, 2013) foi utilizado nos

¹<https://pwkit.readthedocs.io/en/latest/science/pwkit-pdm/>

Tabela 3 – Tabela mostrando os períodos obtidos utilizando ambos os métodos deste trabalho, assim como os resultados obtidos por (MCQUILLAN; MAZEH; AIGRAIN, 2014)

KIC	$P_{Trab} - LS1$ (dias)	$P_{Trab} - LS2$ (dias)	$P_{Trab} - PDM$ (dias)	P_{Rot} (dias)	Erro rel.
757099	0.363	0.363	0.363	0.4	0.0%
3338421	2.353	2.338	2.338	4.8	52.1%
3437548	1.829	1.829	1.829	1.8	0.0%
3529147	14.425	14.404	14.404	14.2	1.41%
4636995	10.611	19.009	19.009	21.0	9.48%
4648919	0.626	0.626	0.626	0.6	0.0%
5221138	20.098	20.013	20.013	19.7	1.5%
5271947	22.951	19.649	19.649	40.1	50.87%
5524851	6.897	6.896	6.896	—	—
8054838	14.6	14.023	14.023	14.2	1.27%
8456181	19.565	19.626	19.626	19.5	0.5%
10280371	20.048	16.275	16.275	—	—

Figura 4.5 – Histograma de todos períodos obtidos no trabalho.



Fonte – Elaborada pelo autor.

estudos cujos resultados foram comparados. Assim, podemos dizer que alguns dos períodos de rotação obtidos neste trabalho estão bem satisfatórios quando comparados com os obtidos anteriormente. No histograma da figura 4.5 é possível ver como estão distribuídos os períodos

5 CONCLUSÃO

Neste trabalho, foi desenvolvida uma aplicação que possibilitou a obtenção de dados estelares obtidos pelo satélite Kepler e a partir desses dados, calcular o período rotacional de uma estrela, ou várias, através de dois métodos distintos que foram utilizados em conjunto. Essa abordagem foi bastante eficaz, pois apesar de ter aumentado um pouco o custo computacional devido à complexidade do método da PDM, os resultados foram em boa parte precisos, reduzindo a probabilidade de ocorrer *alias*. Como foi observado no capítulo anterior, isso ocorreu em apenas 17.3% do total dos valores utilizados para validação.

A automatização completa do processo se mostrou bastante benéfica mesmo durante o desenvolvimento do código, pois agilizava a realização dos testes necessários. Houve uma economia de 5,5 horas apenas na obtenção dos dados, de forma que houve um aumento na produtividade. No total, o processo inteiro, desde a aquisição dos dados até a determinação dos períodos de rotação, levou aproximadamente 3 dias. Algo que levaria um tempo bem maior caso fosse feito manualmente. Como todo o processo é feito de forma automática, o usuário pode realizar outras atividades em paralelo sem atrapalhar o fluxo da operação.

Com base nos resultados apresentados, pode-se dizer que o objetivo do trabalho foi alcançado, porém ainda há possibilidades de melhorias, como o desenvolvimento de interfaces gráficas para facilitar o uso da ferramenta, integração com banco de dados e a utilização de conceitos como orientação à objetos que podem melhorar a eficiência do código. Ajustes no código podem ser feitos para processar os dados da missão K2, assim como do TESS¹. Também é interessante encontrar métodos eficientes para retirar ruídos das curvas de luz de forma que não afete de forma significamente a determinação do período de rotação.

¹ *Transiting Exoplanet Survey Satellite*, satélite sucessor do Kepler lançado em 2018 pela agência espacial americana.

REFERÊNCIAS

- BAGLIN, A. et al. Scientific Objectives for a Minisat: CoRoT. In: Fridlund, M. et al. (Ed.). *The CoRoT Mission Pre-Launch Status - Stellar Seismology and Planet Finding*. [S.l.: s.n.], 2006. (ESA Special Publication, v. 1306), p. 33. 18
- BORUCKI, W. J. et al. Kepler planet-detection mission: Introduction and first results. *Science*, American Association for the Advancement of Science, v. 327, n. 5968, p. 977–980, 2010. ISSN 0036-8075. Disponível em: <<https://science.sciencemag.org/content/327/5968/977>>. 18, 20
- CHAPLIN, W. J. et al. Asteroseismic Fundamental Properties of Solar-type Stars Observed by the NASA Kepler Mission. , v. 210, p. 1, jan. 2014. 36
- CHARBONNEAU, D. et al. Detection of Planetary Transits Across a Sun-like Star. , v. 529, p. L45–L48, jan. 2000. 17
- COOLEY, J.; TUKEY, J. An algorithm for the machine calculation of complex fourier series. *Mathematics of Computation*, v. 19, n. 90, p. 297–301, 1965. 25
- GENTLEMAN, W. M.; SANDE, G. Fast fourier transforms: For fun and profit. In: *Proceedings of the November 7-10, 1966, Fall Joint Computer Conference*. New York, NY, USA: ACM, 1966. (AFIPS '66 (Fall)), p. 563–578. Disponível em: <<http://doi.acm.org/10.1145/1464291.1464352>>. 25
- GRASSHOFF, G. *The history of Ptolemy's star catalogue*. New York: Springer-Verlag, 1990. ISBN 9780387971810. 15
- HOWELL, S. B. et al. The K2 Mission: Characterization and Early Results. , v. 126, p. 398, abr. 2014. 19
- JOHNSON, S. G.; FRIGO, M. *Implementing FFTs in Practice*. 2009. 25
- KRUPP, E. C. *Echoes of the ancient skies : the astronomy of lost civilizations*. Mineola, N.Y: Dover Publications, 2003. ISBN 0486428826. 15
- LEAVITT, H. S.; PICKERING, E. C. Periods of 25 Variable Stars in the Small Magellanic Cloud. *Harvard College Observatory Circular*, v. 173, p. 1–3, mar. 1912. 17
- LOMB, N. R. Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science*, v. 39, n. 2, p. 447–462, Feb 1976. ISSN 1572-946X. Disponível em: <<https://doi.org/10.1007/BF00648343>>. 26
- MADDALA, G. S. *Introduction to econometrics*. New York Toronto New York: Macmillan Pub. Co. Maxwell Macmillan Canada Maxwell Macmillan International, 1992. ISBN 9780023745454. 27
- MAYOR, M.; QUELOZ, D. A Jupiter-mass companion to a solar-type star. , v. 378, p. 355–359, nov. 1995. 17

MCQUILLAN, A.; MAZEH, T.; AIGRAIN, S. Stellar rotation periods of the kepler objects of interest: A dearth of close-in planets around fast rotators. *The Astrophysical Journal*, IOP Publishing, v. 775, n. 1, p. L11, Sep 2013. ISSN 2041-8213. Disponível em: <<http://dx.doi.org/10.1088/2041-8205/775/1/L11>>. 47

MCQUILLAN, A.; MAZEH, T.; AIGRAIN, S. ROTATION PERIODS OF 34,030keplermain-SEQUENCE STARS: THE FULL AUTOCORRELATION SAMPLE. *The Astrophysical Journal Supplement Series*, IOP Publishing, v. 211, n. 2, p. 24, mar 2014. Disponível em: <<https://doi.org/10.1088%2F0067-0049%2F211%2F2%2F24>>. 11, 36, 43, 45, 46, 48

MITCHELL, W. M. The history of the discovery of the solar spots. *Popular Astronomy*, n. 24, p. 22–ff, 1916. 16, 33

NASCIMENTO, J. J.-D. do et al. ROTATION PERIODS AND AGES OF SOLAR ANALOGS AND SOLAR TWINS REVEALED BY THEKEPLERMISSION. *The Astrophysical Journal*, IOP Publishing, v. 790, n. 2, p. L23, jul 2014. Disponível em: <<https://doi.org/10.1088%2F2041-8205%2F790%2F2%2F123>>. 11, 45, 46

NIELSEN, M. B. et al. Rotation periods of 12 main-sequence kepler stars: Dependence on stellar spectral type and comparison with $v \sin i$ observations. *A&A*, v. 557, p. L10, 2013. Disponível em: <<https://doi.org/10.1051/0004-6361/201321912>>. 46

NILSSON, M. P. *Primitive time-reckoning, a study in the origins and first development of the art of counting time among the primitive and early culture peoples*, by Martin P. Nilsson, ... Lund: C.W.K. Gleerup, 1920. 15

NORTH, J. *The Norton history of astronomy and cosmology*. New York: Norton, 1994. ISBN 9780393036565. 15

PRIEMER, R. *Introductory Signal Processing*. [S.l.]: World Scientific Publishing Co. Pte. Ltd, 1991. 22

SCARGLE, J. D. Studies in astronomical time series analysis. 2. Statistical aspects of spectral analysis of unevenly spaced data. *Astrophys. J.*, v. 263, p. 835–853, 1982. 26

SCHUSTER, A. On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena. *Terrestrial Magnetism (Journal of Geophysical Research)*, v. 3, p. 13, 1898. 25

SHANNON, C. E. Communication in the presence of noise. *Proceedings of the IRE*, v. 37, n. 1, p. 10–21, 1949. 24, 41

SODERBLOM, D. R.; KING, J. R. Solar-Type Stars: Basic Information on Their Classification and Characterization. In: Hall, J. C. (Ed.). *Solar Analogs: Characteristics and Optimum Candidates*. [S.l.: s.n.], 1998. p. 41. 20

SPAETH, O. V. Dating the oldest egyptian star map. *Centaurus*, v. 42, n. 3, p. 159–179, 2000. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1034/j.1600-0498.2000-.420301.x>>. 15

STELLINGWERF, R. F. Period determination using phase dispersion minimization. , v. 224, p. 953–960, set. 1978. 28

VANDERPLAS, J. T. Understanding the Lomb-Scargle Periodogram. , v. 236, n. 1, p. 16, May 2018. 24

WALT, S. van der; COLBERT, S. C.; VAROQUAUX, G. The numpy array: a structure for efficient numerical computation. *CoRR*, abs/1102.1523, 2011. Disponível em: <<http://arxiv.org/abs/1102.1523>>. 29

WILLIAMS, P. K. G. et al. *pwkit: Astronomical utilities in Python*. 2017. ascl:1704.001 p. 30