Advanced Astroinformatics - Student Project

Introduction (topic, TESS light curves)

Dr. Nina Hernitschek
June 3, 2022

Motivation

Nowadays large-scale astronomical surveys enable us to see the universe in much more detail:

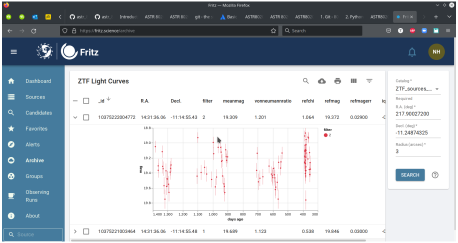
deeper (fainter objects), **wider** (larger on-sky footprint), **faster** (higher temporal resolution, cadence)

⇒ methodological approaches to analyze the data changes

what you will learn in this project

this project will prepare you for “doing science” with current and upcoming large astronomical surveys:

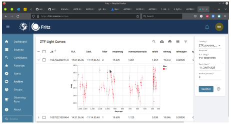
accessing astronomical survey data



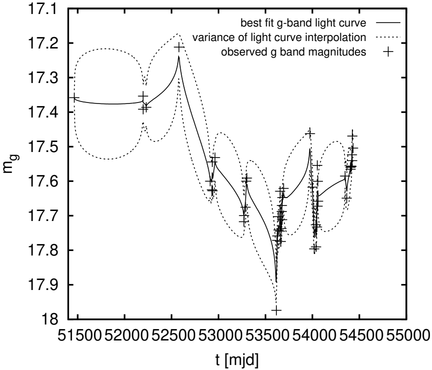
what you will learn in this project

this project will prepare you for "doing science" with current and upcoming large astronomical surveys:

accessing astronomical survey data



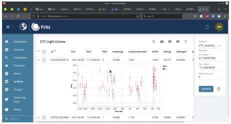
time series analysis



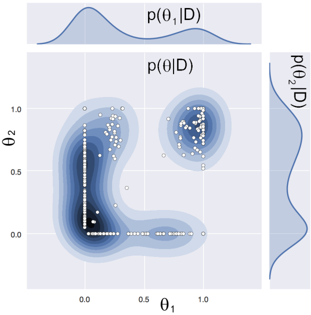
what you will learn in this project

this project will prepare you for “doing science” with current and upcoming large astronomical surveys:

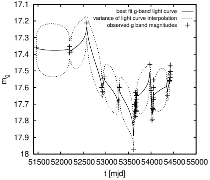
accessing astronomical survey data



statistical methods



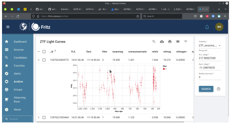
time series analysis



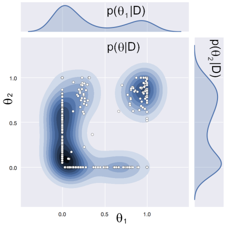
what you will learn in this project

this project will prepare you for “doing science” with current and upcoming large astronomical surveys:

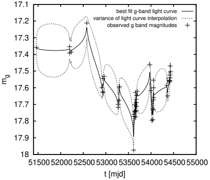
accessing astronomical survey data



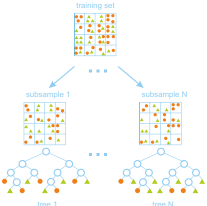
statistical methods



time series analysis



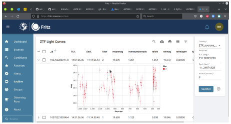
machine learning



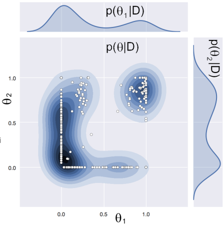
what you will learn in this project

this project will prepare you for "doing science" with current and upcoming large astronomical surveys:

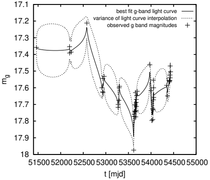
accessing astronomical survey data



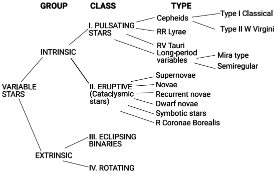
statistical methods



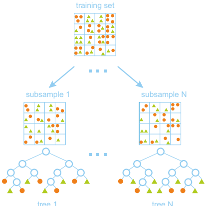
time series analysis



classification



machine learning



Project Overview

The project is structured as follows:

1) an overview of machine-learning techniques applicable for astronomical data

2) preparing astronomical data for processing them with machine learning techniques

3) applying machine-learning techniques to astronomical data

4) understanding the benefits, relevance and limitations of machine learning classification on the example of the light-curve data provided for this project

The **outcome** of this project will be a machine-learning program to classify variable star light curves.

Plan for Today

- github
- `Python3.x` environment
- TESS data

git

`git` is a **version control system**. Specifically, `git` is a distributed version control system: each user actually clones the entire **repository** locally.

Think of a repository as a collaborative directory to which multiple people have access, and that it tracks how changes are made (who, when, what).

`git` vs. `github`: `git` is a version control system. `github` is a cloud-based hosting service that lets you manage Git repositories.

git

In this class, all work will be submitted through `git`.

Please create your own github repository for your own code.

The files for this student project are available on github:

`https://github.com/ninahernitschek/advanced_astroinformatics`

Python

Python is an open-source, object-orientated high-level scripting language that is useful for manipulating data. It is widely used in science, especially physics and astronomy.

As with any programming language, Python has some undesirable features, such as some relatively slow processes. But Python can be used to wrap faster code such as C. (I use this often.)

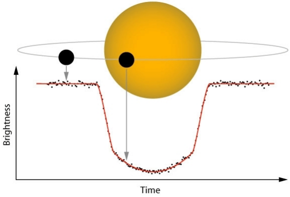
advantages: Python it has a large number of existing packages for manipulating large datasets including data access, machine learning and plotting.

Python

we continue with `notebook_1a.ipynb` from the `github` repository

Transiting Exoplanet Survey Satellite (TESS)

TESS is a space telescope for NASA's Explorer program, designed to search for exoplanets using the **transit method**.

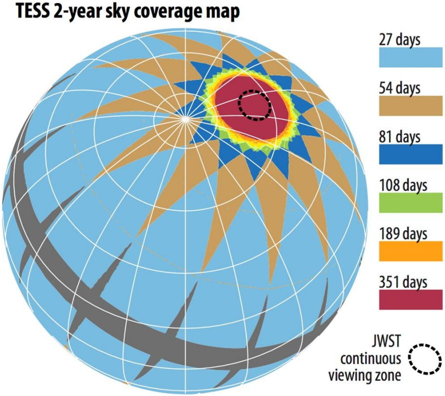


TESS is designed to survey over 85% of the sky (400 times larger than covered by Kepler) to search for planets around nearby stars (within ∼200 light years). TESS stars are typically 30-100 times brighter than those observed with Kepler.

TESS websites:
`https://www.nasa.gov/tess-transiting-exoplanet-survey-satellite`
`https://tess.mit.edu/`

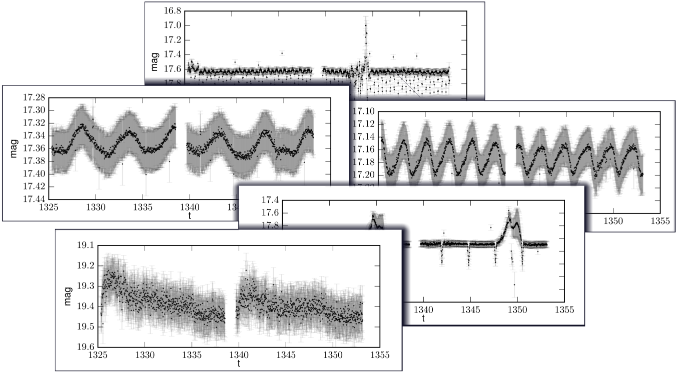
Transiting Exoplanet Survey Satellite (TESS)

Using an **input catalog**, TESS observes the sky in **sectors**. Most sectors are observed for about 27 days.



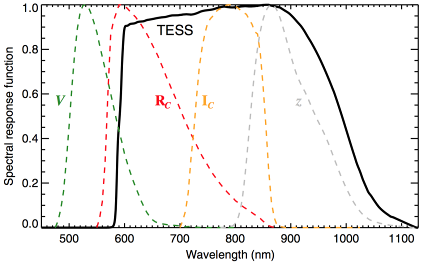
Transiting Exoplanet Survey Satellite (TESS)

typical TESS light curves:



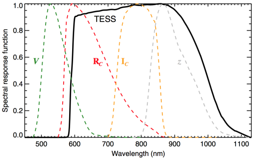
characteristics (sector01):
baseline: 27.4 days
cadence: 30 minutes (in addition: 2-minute cadence for some stars)
wavelength: single-band, 600 - 1000 nm (same for all TESS observations)

Transiting Exoplanet Survey Satellite (TESS)



The TESS spectral response function (black line). For comparison, the Johnson-Cousins V, R_C, and I_C and the SDSS z filter curve are shown. Credit: Ricker et al. (2015)

Transiting Exoplanet Survey Satellite (TESS)



The TESS spectral response function (black line). For comparison, the Johnson-Cousins V, R_C, and I_C and the SDSS z filter curve are shown. Credit: Ricker et al. (2015)

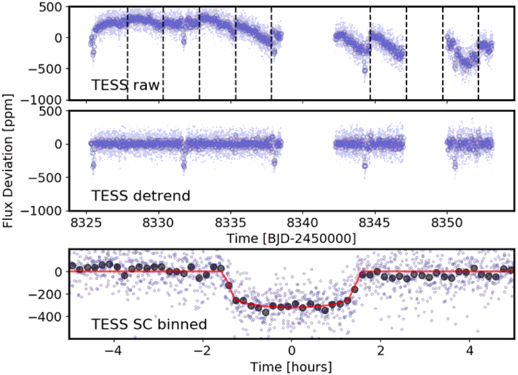
This wide, red-optical bandpass reduces photon-counting noise, thus increasing sensitivity.

TESS observed a large number of M dwarfs:

- Planets are easier to detect around these small stars (the planets induce larger transit signals).
- Most nearby stars are M dwarfs. Because M dwarfs are cool and red, the TESS bandpass will be more sensitive to red wavelengths.

Transiting Exoplanet Survey Satellite (TESS)

With a baseline of 27.4 days in most sectors, the satellite is most sensitive to exoplanets with a periods of less than 13 days (so that at least two transits are used for discovery).

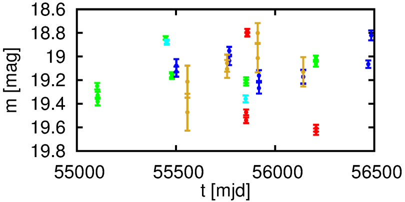


Huang et al. (2018) *TESS Discovery of a Transiting Super-Earth in the* Π *Mensae System*

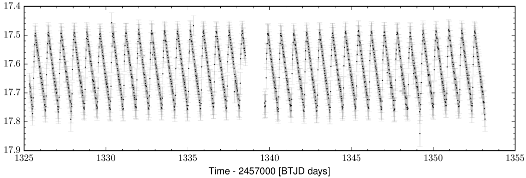
Transiting Exoplanet Survey Satellite (TESS)

We can also use the light curves to analyze stellar variability.

A RR Lyrae star seen in the PS1 3π survey



... and with TESS:



Variable Stars

From supernovas to stars with exoplanets, variable stars are of high interest to astronomical research.

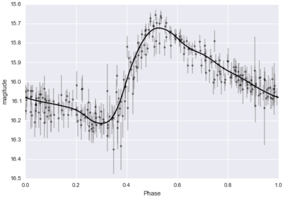
Research on variable stars is important because it **provides information about stellar properties**, such as mass, radius, luminosity, temperature, internal and external structure, composition, and evolution. Some of this information would be difficult or impossible to obtain any other way.

In addition, stellar variability can provide **distance information** (keyword: distance ladder).

Variable Stars

periodic variable stars allow for distance calculation:

Cepheid and RR Lyrae stars are variable stars with the period being directly related to their true (absolute) brightness.

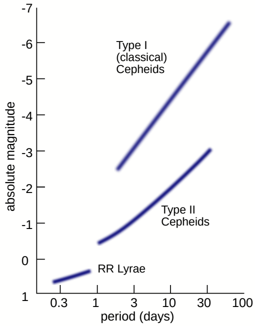


- measure apparent mean brightness m
- measure period P

Variable Stars

periodic variable stars allow for distance calculation:

Cepheid and RR Lyrae stars are variable stars with the period being directly related to their true (absolute) brightness.



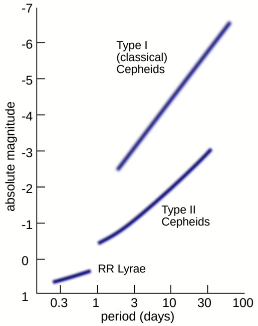
- measure apparent mean brightness m
- measure period P
- using period-luminosity relation, get absolute brightness M
- solve for distance using equation $d = 10^{(m-M+5)}/5$ parsec

  where 1 parsec $= 3.086^{16}$ m $= 3.26156$ light years

Variable Stars

periodic variable stars allow for distance calculation:

Cepheid and RR Lyrae stars are variable stars with the period being directly related to their true (absolute) brightness.



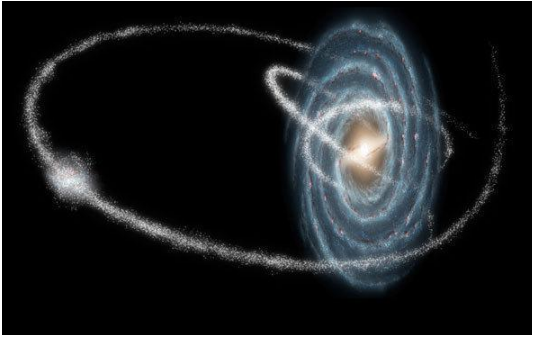
- measure apparent mean brightness m
- measure period P
- using period-luminosity relation, get absolute brightness M
- solve for distance using equation $d = 10^{(m-M+5)}/5$ parsec
  where 1 parsec $= 3.086^{16}$ m $= 3.26156$ light years

$\Rightarrow$ allow us to create *3D maps* of structures in our Milky Way

Variable Stars

RR Lyrae stars as tracers for old Milky Way substructure:

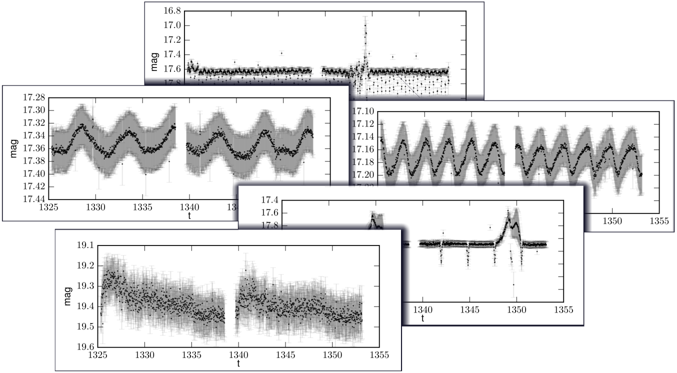
- old: $\sim 10^9$ years
- high-precision 3D mapping of the (old) Milky Way



artistic image, www.spitzer.caltech.edu

Transiting Exoplanet Survey Satellite (TESS)

typical TESS light curves:

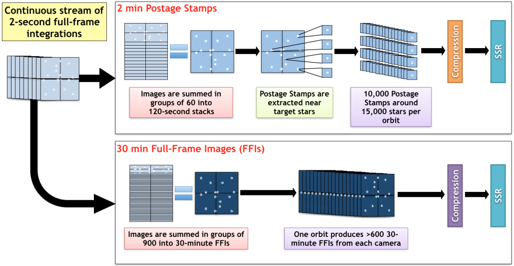


obvious systematics: momentum dumps

a problem when analyzing variability

back to: Transiting Exoplanet Survey Satellite (TESS)

where do these light curves come from?

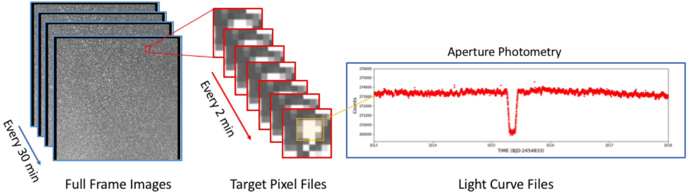


Representation of how the postage stamps and FFIs are created. source: `https://heasarc.gsfc.nasa.gov/docs/tess/data-products.html`

A Full Frame Image (FFI) is a collection of pixels observed simultaneously across all CCDs of a given camera. FFIs were taken every 30 minutes during science operations in the primary mission.

back to: Transiting Exoplanet Survey Satellite (TESS)

where do these light curves come from?



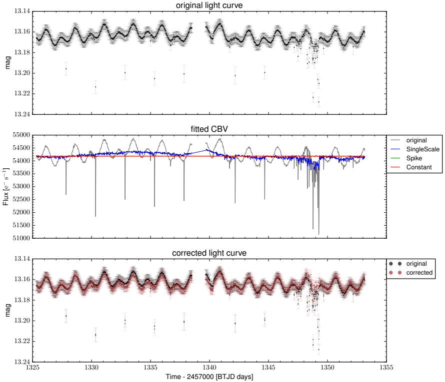
Photometric data products. source:

`https://heasarc.gsfc.nasa.gov/docs/tess/data-products.html`

back to: Transiting Exoplanet Survey Satellite (TESS)

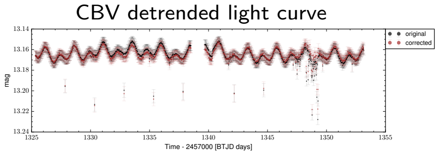
In order to do variable star science, systematics and outliers must be removed while preserving stellar variability.

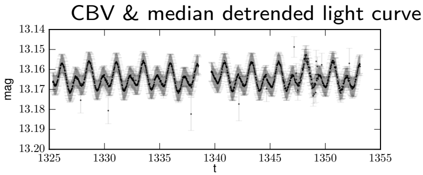
1. Removing trends using CBV (cotrending basis vectors)



Co-trending basis vectors (CBVs) represent the set of systematic trends present in the ensemble flux data. CBVs are provided for each TESS sector.

2. median detrending



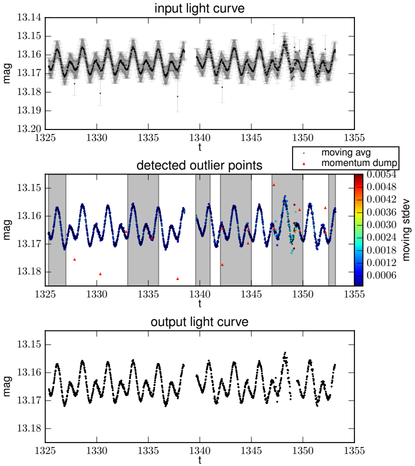


back to: Transiting Exoplanet Survey Satellite (TESS)

3. outlier cleaning
The previous steps remove systematics but not outliers. They also don't remove very distinct systematics.

approach:

- removing outliers at known times of momentum dumps
- moving standard deviation to detect outliers
- detection of conspicious Gaussians in the light curve



Break & Questions

afterwards we continue with **`notebook_1b.ipynb`** from the `github` repository