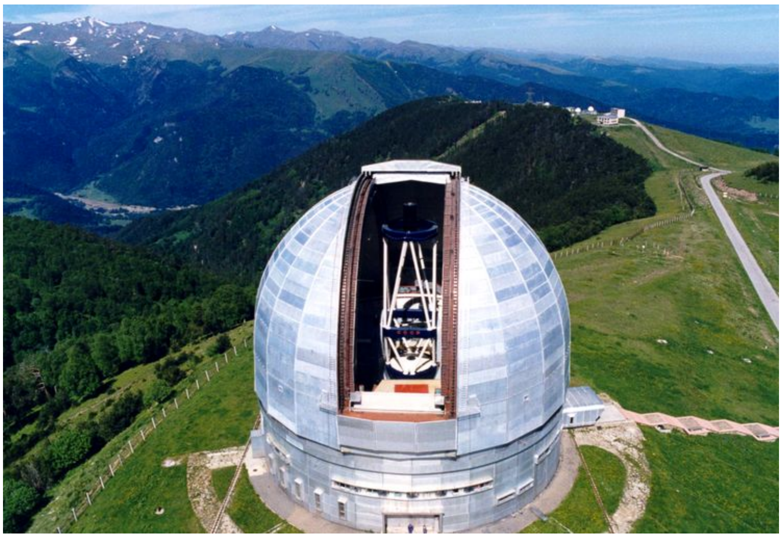
DEPARTMENT OF PHYSICS
INDIAN INSTITUTE OF TECHNOLOGY
MADRAS
CHENNAI - 600036

Search for variable stars in digitised archival photographic plates



A Project Report

Submitted by

YASHODHAN MANERIKAR

In the partial fulfilment of requirements

For the award of the

DUAL DEGREE (B.S. and M.S.)

June 2021

DEDICATION

In the beloved memory of George Reddy.

CERTIFICATE

This is to undertake that the project report titled **SEARCH FOR VARIABLE STARS IN DIGITISED ARCHIVAL PHOTOGRAPHIC PLATES**, submitted by me to the Indian Institute of Technology Madras, for the award of BS+MS, is a bona fide record of the research work done by me under the supervision of Prof. Nikolai Samus and Prof. Suresh Govindarajan. The contents of this project report in full or in parts, have not been submitted to any other institute or university for the award of any degree or diploma.

Place: Chennai 600 036

Date: 18th June 2021

Yashodhan Manerikar
Dual Degree student

Prof. Nikolai Samus
Research Guide
Lomonosov Moscow State University

Prof. Suresh Govindarajan
Research Co-Guide
Indian Institute of Technology, Madras.

ACKNOWLEDGEMENTS

I thank the director of the Institute of Astronomy of the Russian Academy of Sciences, prof. D.V. Bisikalo for inviting me to Moscow. I thank the director of Sternberg Astronomical institute for providing me with a workspace at SAI. I am incredibly grateful for Prof. Samus for everything he did for me. I thank Dr. Alexandra Zubareva for all her teaching and help for my work. I thank Dr. Kirill Sokolovsky for all his assistance with the technology. I thank prof. Suresh Govindarajan for his support and guidance. I wish to thank the DDP review seminar panel for their useful questions and guidance. I am grateful for my friends in Moscow who made my stay gleeful.

ABSTRACT

KEYWORDS: Variable stars, photographic plates, VaST, classification, machine learning.

Variable stars are those stars whose apparent observed brightness varies with time. Lightcurves of these stars can show a remarkable elegance and periodic behaviour. We can see stars expanding and contracting in precise ways, stars eclipsing smaller stars and gigantic explosions on the surface of white dwarfs through this study. They provide time series data for studying stellar dynamics and evolution. Variable stars are also used for measuring distances to stars, star clusters or galaxies. Their study is essential in understanding the evolution of our galaxy.

In this project we discovered new variable stars from scanned old photographs of a certain part of the sky. The photographic plates were borrowed from the Moscow plate collection, and they were taken between 1983 and 1996. The plates covered a large portion of the sky and the images were captured consistently.

The text is a master's project report. It prepares the background, exhibits the tools, illustrates the process of discovery, showcases the discoveries made in this project, explains the physics' of variability and talks about the scope for machine learning applications.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

This chapter is a short introduction to the organised research in variable stars. It will tell you what to expect from this project report. The aim of the chapter is to elucidate the state of research in variable stars and the significance and influence our research holds.

1.1 Introduction to variable stars research: past, present and future

1.1.1 The beginnings

Astronomy is a very old field of study, however the branch variable star astronomy is relatively recent. In the Aristotelian vision, stars were eternal, unchanging and pure, and that was the assumption with which we operated for centuries. The first variable stars which were detected varied in brightness with such an amplitude that their variability could be observed with naked eyes. The first variable star to be discovered and had its period determined was Omicron Ceti, a bright star in the Cetus constellation, near Pisces. It was discovered in 1595-1609 by David Fabricus. Its period was determined in 1638 by a Frisian astronomer named Johannes Holwarda. This star would be seen shining brightly and then in some months completely disappear from the view. At this point there was a discourse available on novae by Tycho Brahe and others. Nova is Latin for new; newly appeared stars were called novae. When Omicron Ceti, now renamed Mira appeared out of nowhere Holwarda believed it to be a nova at first. However when he saw it reappear, he knew he was looking at something else. He determined the period of this oscillation to be 11 months. This was the beginning of an era. The second variable star discovered was the famous devil star, Algol (al-ghul in Arabic). It is the second brightest star in the constellation of Perseus. Algol varies with a period of 2.85 days. The variability of Algol was noted by the Italian astronomer Geminiano Montanari in

1667. And a mechanism of its variability was proposed by an amateur astronomer John Goodricke, proposing that the dips in the brightness were caused by transit of a smaller darker body in front of the star.

There are ancient historical records that indicate that the knowledge of the variability of Algol and Mira was known to ancient civilisations like the Egyptians, Chinese, Babylonians or Greek; however we won't be getting into that.

Before photography, astronomers estimated brightness of the stars with their naked eyes pressed against the telescope's eyepiece and noted it down. Such an approach may sound prone to large errors, however with technique and scientific care they were able to construct lightcurves of variable stars and derive properties of the stars or star systems out of this data!

In 1844 German astronomer Friedrich Argelander published a list of 18 variable stars known at the time, and motivated a generation of amateur astronomers worldwide. Argelander was the first scientist to systematise the study of variable stars and contribute the step method of estimating brightness of the stars (which is called the Argelander method).

1.1.2 Photographic plate era

The advent of photographic plates revolutionised variable star research. By 1912 the number of known variable stars had gone up to over 4000! A great contribution came from Harvard College Observatory, Cambridge, Massachusetts. HCO had installed the biggest telescope in USA, and they had stations abroad as well. They conducted systematic sky surveys on photographic plates. Today over 500,000 photographic plates sit in the HCO photographic plate archives. The astronomers at HCO (one must note the women astronomers here!) discovered thousands of new variable stars, and formulated laws for their behaviour.

Next contributions to the field came from Germany and Russia; particularly the Sonneberg Observatory in Germany and Sternberg Astronomical Institute, Moscow. A mammoth of variable star research was a scientist named Cuno Hoffmeister. The founder of the Sonneberg Observatory, he discovered approximately 10,000 variable stars. He refined the technology and wrote the classic text on variable stars (Verändliche

Sterne). After WW2, the telescope at Sonneberg observatory was taken as a reparation by the Sovyet Union, and that is the mother telescope for our data. Sonneberg Observatory became a part of East German Academy of Sciences. The Sonneberg Observatory has over a 100,000 plates in their collection.

The Moscow Observatory was established in 1831 by the Moscow State University. In 1931 it became the Sternberg Astronomical Institute. The astronomers at SAI diligently maintained card catalogues of all the confirmed variable stars. A fact for which they were handed the responsibility of variable stars by the International Astronomical Union. In 1948 the Academy of Sciences of the SU published the first edition of 'General Catalogue of Variable Stars'.

We will discuss the photographic plates in another chapter. Photographic plates were superseded by the 'Charged Coupled Device' electronic receptor, and they fell out of use in the 90s. The use of spectroscopy in astronomy opened for astronomers many new dimensions of information. By splitting the star-light on a spectrum of wavelengths (a length scale for electromagnetic waves), we could observe the absorption lines and emission lines, indicating presence of certain elements and molecules in the atmospheres of the stars. The spectra of starlight also came in different *types*, what we call spectral classification of a star. Importantly, you could calculate a mean temperature of the stellar surface and the radial component of the relative velocity between the observatory and the star. This led to a deeper analysis and understanding of variable stars, especially of the intrinsic variables, which vary their brightness by complex astrophysical processes (as opposed to extrinsic or geometric variable stars like eclipsing binaries).

1.1.3 CCD and Space telescope era

Photometers have been around for over a century, however they remained primitive in the earlier phases. From the early seventies, photometers manufacturing technology improved exponentially and the new kinds of detectors known as Charged Coupled Devices swept away all the photographic plate or film technology both in commercial areas and personal.

CCD cameras capture pictures in form of data. With CCDs began a new era of astronomy. There was no need for developing a plate and preserve the plates physically. CCD images are 2D arrays that can be easily stored on electronic hard drive.
One could say that the history of space observatories begins with Sputnik 1, (the first artificial satellite launched in 1957) which sent back radio transmissions back to Earth. From the 60s, a variety of scientific space probes were launched into space. Notably, the Hubble Space Telescope launched into low Earth orbit in 1990 was a huge step for astronomy. Initially it was miscalibrated and it had to be fixed by astronauts, after which it provided with crisp long exposure pictures of the sky without any interference from atmosphere.
This is also the age of very large ground based observatories, which with the use of adaptive optics have started to produce high quality images comparable to HST. In the future, we are expecting the launch of James Webb Space Telescope, which would remain at the second Lagrange point of the Sun-Earth system. All of these are important for variable star astronomy.

In the future, the number of variable stars discovered by space telescopes will be so large and increase at such a rate that keeping a catalogue of variable stars is going to prove difficult. Currently the hot topic in variable stars is exoplanets. In the past decade we have discovered thousands of exoplanets from photometric data, especially the data taken by Kepler space telescope.

1.2 Goals of this project

The goal is to make new discoveries from old data using new techniques. Observatories worldwide have physical storage rooms of photographic plates, with the total number of large field photographic plates exceeding 2.5 million. The Moscow observatory (later turned into SAI) established in 1831 has over a 60,000 plates in its collection. These photographs were taken for various purposes and studies. A star or an area can be studied by pulling up photographs by coordinates. With digital scanning and the use of modern variability detection software we can use the archival data to make discoveries. The digitisation has been under way for around 15 years, and is still in process today.

The photographs can be accessed publicly through the internet.

There are some advantages in using the old data. Out of the 60,000 plates, 22,300 were taken on a German astrograph taken as a war reparation from East Germany. These pictures are all of the same format and hence it becomes possible to sew together pictures from different decades even on the same timeline without hassle. Archival data offers us a significantly longer time window than modern surveys. Though modern ground-based surveys offer incredible accuracy, typically one gets far fewer nights of observation per star as compared to photographic archives.

In this project we use 127 photographic plates with a general overlapping sky-field to discover new variable stars. We will study the discovered variable stars in detail. We will also make inferences about our findings and talk about any patterns observed. There is also the idea of an AI classifier.

1.3 Literature

The classic textbook for variable stars is "Verändliche Sterne" or "Variable Stars" by C. Hoffmeister, G. Richter, W. Wenzel. The English translation was written by S. Dunlop and is published by Springer-Verlag (Hoffmeister *et al.* (1985)). It is written by the astronomers working at the Sonneberg observatory, including the esteemed late prof. Hoffmeister. However, one should keep in mind that a lot of new discoveries have been made about variable stars since 1980s, when the book was published. Another good book is the "Practical guide to lightcurve photometry and analysis" by Brian D. Warner, it offers a practical guide to variable star astronomy, and the book is also accessible for amateur astronomers (Warner (2006)). John R. Percy's "Understanding Variable Stars" is a good textbook published in 2007 by Cambridge-Astrophysics (Percy (2007).

General Catalogue of Variable Stars is a catalogue published by a team from the Sternberg Astronomical Institute of Moscow State University and Institute of Astronomy, Russian academy of sciences, currently chaired by Prof. Samus. The latest version of the catalogue lists around 57,000 variable stars discovered and labeled until 2020 with their relevant information such as the variable star name per constellation, J2000 coordinates, variable star type, magnitude(s), period, new experimental classification, name

of the survey, proper motion etc (Samus *et al.* (2017)). The first edition was published in 1948, containing 10,840 stars. It was edited by B.V. Kukarkin and P.P. Parenago.
Variable star astronomy is found worldwide, and it was led by the astronomers from countries of Germany, USA and Russia. Harvard College Observatory, Sonneberg Observatory and Moscow Observatory played a crucial role in the development of the field. American Association of Variable Stars Observers (AAVSO) is a very valuable organisation today, receiving real discoveries and data from amateur astronomers of the USA. The International Astronomical Union's (old) commissions 27 and 42 were relevant to the variable stars research.
The AAVSO maintains an important online search tool for variable stars known as VSX: Variable Stars Index. It allows users to look up variable stars found in many different surveys and catalogues by coordinates, names and more. Details of VSX and other online tools is given in the "Tools of the trade" chapter.

CHAPTER 2

Basics and tools

The first section will catch you up on some vital astronomical concepts. It will be a quick course on basics of astronomy. The second section will discuss the mathematical tools we used, followed by introduction to the main software 'VaST' and astronomical databases.

2.1 Fundamentals of Astronomy

The goal of this chapter is to give you a ground-level understanding of astronomy. But we will only cover in depth the astronomical concepts that are relevant to our project; say colour, brightness and date and skip other basic concepts which are irrelevant such as polarisation, magnetic field and radial velocity. Radial velocity is indeed very important in the variable star business, however we will not be using any radial velocity values in this project; so it is skipped for brevity's sake.

2.1.1 Frame of Reference and the Coordinate system

For most of the history, astronomers used a geocentric frame of reference in which the Earth was a stationary sphere (though a daily rotating one) in the centre of the universe. From this point of view, the Sun and the Moon revolved around the Earth in circles while the planets traced complex orbits. We are now learning to see the universe from the perspective of the falling particle of dust in empty space, which is our real perspective. However we have not yet completely outgrown the geocentric tendencies in the coordinate systems for convenience's sake.
If you are familiar with the latitude and longitude system we use to map points on the surface of the Earth, the astronomical equatorial coordinates system will be very easy for you to understand. Both use the Earth's rotation as a reference in defining a pole and and an equator. Imagine the latitude longitude system on the surface of the Earth,

and then enlarge it about the Earth's center to exceed Earth, in a form of a grid that floats above the Earth in all directions. Change the angle of 'latitude' to the angle of 'declination', an angle that goes $-90°$to $+90°$ from the South pole to the North pole. Change longitude to Right Ascension, a circular scale divided in 24 hours, where each hour is divided into 60 minutes and each minute is divided in 60 seconds (and decimal digits after this). Now for calibration we rotate the sphere until the coordinate (0 RA,0 Dec) points in the (arbitrary) direction of vernal equinox (it is the direction of the radius vector from the Sun to Earth on the day of vernal equinox). Fix the grid here. Latitude and Longitude rotate with the Earth but RA and Dec do not.

A complication arises due to something known as the precession of the Earth. The Earth's rotational axis slowly rotates, tracing a circle about the normal to the plane of the Earth's orbit, completing one cycle every 26000 years. This means that each year vernal equinox is in a slightly different location. This is a small effect, hence we ignore it for 40 years or so. At the moment astronomers are using a J2000 coordinate system, meaning that we are using a grid whose (0 RA,0 Dec) points to the vernal equinox as it was in the year 2000. We will update it in the next couple of decades.

Equatorial Coordinate system is convenient because we only need two coordinates to map the sky. We would need 3 coordinates (x,y,z) in a rectangular system, but as we assume the stars to be absurdly far away, we can drop the radius coordinate in spherical coordinates and only use the angular coordinates. This however creates a false image of a 'celestial sphere', an idea (still useful in teaching astronomy) that stars are spherically distributed around the Earth.

Two other major systems in use are heliocentric coordinates and galactic coordinates.

2.1.2 Nomenclature

Prominent stars have folk names viz. Sirius, Dhruva or Altair, and they only form a minority in named stars. The IAU has made a list of 88 constellations (collected from a diaspora of cultures and folk-lore) which cover the entire celestial sphere without gaps, hence dividing the sky in 88 oddly shaped compartments. Out of the 88, 13 lie on the path of the Sun, and are known as the zodiacal constellations (the 13th constellation is Ophiuchus). All these constellations have a short-hand; for example 'Cyg' for Cygnus or 'Sgr' for Sagittarius.

Astronomers use the 'Bayer designation' for naming prominent stars. In this convention Greek letters alphabetically index brightness hierarchy within a constellation. For example, α Cyg and δ Cyg represent the first and fourth brightest stars in the Cygnus constellation. By this convention we can name up to 24 stars in each of the 88 constellations. Note that even if a star has a popular name like Betelgeuse, it is still assigned a Bayer designation (α Ori in Betelgeuse's case). Also note that it is a fixed convention, so Betelgeuse is α Ori despite it being the second brightest star in Orion. After the Bayer designated stars, if you wish to know the name of the star, you need context. Variable stars are named in their own convention, which we will come to in a moment. Once you search a star by its position online, the databases might provide you multiple names for the same star distinctly based on the catalogues or surveys it was named in.

In case of variable stars, if a star with Bayer designation is variable then the Bayer designation is allowed to stay. For example, δ Cep (Cepheus) is the prototype for the famous class of cepheid variables. Variables then are named like 'V0001', 'V0002' and so on in order of discovery in each constellation. However, an alphabetical indexing is used for the first 334 names due to tradition.

For example, 'V0001 Cep' is 'R Cep' and 'V0002 Cep' is 'S Cep'. The list follows from here alphabetically until it hits Z, then it goes into double letters. After Z we have RR, RS, RT and so on. Once we reach RZ, we start with SS, ST and so on. After ZZ comes AA. The 334th combination is the last one and it is QZ, the 335th variable star in Cepheus then is simply named 'V0335 Cep'.

2.1.3 Magnitude: Brightness

The old age astronomers came up with a system of estimating the star's brightness. The stars were divided in 6 classes, the brightest being class 1 stars and the dimmest being class 6 stars. In the nineteenth century this scale was redefined mathematically and the class turned into magnitude. It was defined so that a magnitude 1 star is exactly 100 times brighter than a magnitude 6 star. It is a logarithmic scale, in accordance with the Weber-Fechner law which claims that physiological sensations are logarithmically proportional to the physical stimuli. Saying the same thing mathematically,

$$M_1 - M_{ref} = -2.5 \log(\frac{I_1}{I_{ref}})$$

Where M's are the magnitudes and I's are the intensities. Intensities are related to the luminosities of the stars, the total power output. Note the negative sign. The scale is inverted, the brighter an object is, the smaller it is numerically.

This was done to adapt to the old system as closely as possible. For example, the magnitude/class of the star Antares was 1, and the modern magnitude is also close to 1. What is the Sun's brightness? As the Sun is a billion times brighter than all the other stars, the magnitude ought to be a lot lower than 1 and indeed the Sun's (apparent) magnitude comes out to be -26.75. Sun has the brightest apparent magnitude of all.

To give you a rough idea, the magnitudes of the top five planets are in the range -5 to 2. The brightest star Sirius is at a magnitude of -1.46.

But let us be very careful when we use the word 'brightness'. -1.46 is Sirius' "apparent magnitude". Brightness could refer to luminosity of a star, which these numbers are not indicative of. Luminosity of a star could be the total power output of a star. We only measure a fraction of it, known as intensity. Just by measuring the brightness of a star we cannot know the difference between a brighter star farther away from us or a fainter star closer to us. Without distance we only know the relative brightness, which we call as apparent magnitude. Once you know the distance to the star, you can calculate the 'absolute magnitude' which is the (apparent) magnitude observed at a fixed distance of 10 parsecs. The Sun's apparent magnitude is -26.75 while its absolute magnitude is 4.83. The Sun would be an average class 4 star in the night sky to Hipparchus if he lived 10 parsecs away from it.

2.1.4 Spectrum: Colour/Temperature

As we know, light constitutes of electromagnetic waves of different wavelengths. Light coming from a star can be made to split smoothly with use of a prism or a grating in the order of its constituent wavelengths. A prism spreads a beam of light across the wall in a spectrum of frequency or wavelength. A spectrum of a star gives a whole lot of additional information about the star, such as the surface temperature (colour) and the elemental constitution. We can also calculate the radial velocity of the star (the rate of change of radius drawn from the Earth to the star, negative/blue if decreasing and positive/red if increasing) from the spectrum.

Temperature is a statistical quantity and it gives you an idea about the kinetic energies of the particles which make up the object in question. Depending on the state of matter, the atoms and molecules in that object are vibrating, flying at great speeds in any directions, colliding with each other and so on. All objects with a temperature (all objects have a temperature as far as we know) emit electromagnetic radiation. As stars have fluids, their resulting electromagnetic radiation is distributed continuously across frequency instead of sharp emission lines as we see with gases.

Here two equations rule: Wien's law and Stefan-Boltzmann law. Wien's law relates the wavelength at the peak with the temperature of the body and Stefan-Boltzmann law relates the radiative power output of a star to its temperature and surface area.

$$\lambda_{max} = \frac{b}{T}$$

.... (Wien's Law)

$$P = \sigma \cdot A \cdot T^4$$

.... (Stefan Boltzmann Law)

Where λ_{max} is the peak wavelength, b is Wien's constant (equal to 0.0029 m·K), T is the temperature, σ is the Stefan Boltzmann constant (equal to $5.67 * 10^{-8} \frac{W}{m^2 K^4}$) and A is the surface area.

But for our matters, we will not be calculating temperatures of the stars in Kelvin, rather we will use the colour indices.

Colour index

When we observe light with a detector, that detector has a sensitivity. Our eyes are sensitive for visible light. Detectors can be made of different sensitivities, for bandwidths corresponding to different frequency zones near the visible light part of the electromagnetic spectrum. Traditionally detectors came in 3 bands: UBV: Ultraviolet, Blue and Visual. In the new age we use CCD detectors with a wide range and filters to only expose the detector in the selected colour band. The measured magnitude of a star depends on the filter you use to capture it. If you take pictures of the star through two different bands, you will get two different magnitudes. The difference between these two magnitudes is the colour index corresponding to those two bands.

For example in this project we will be using the J-K colour index, which gives you the difference in magnitude of a star when observed through a J filter and when observed through a K filter. Both J and K bands are in the near infrared regions with peak wavelengths at 1220nm and 2190nm respectively. A higher J-K value indicates a redder star. This gives us an idea about the colour of the star without having to obtain spectra.

HR Diagram

The Hertsprung-Russell diagram stands of great importance in astronomy. When stars were plotted with their luminosities vs spectral types or temperatures, they seemed to follow a pattern. That pattern is also the evolutionary path of stars. You can also see stellar populations lying outside the central trend: the white dwarfs in the bottom left and red giants in top right.

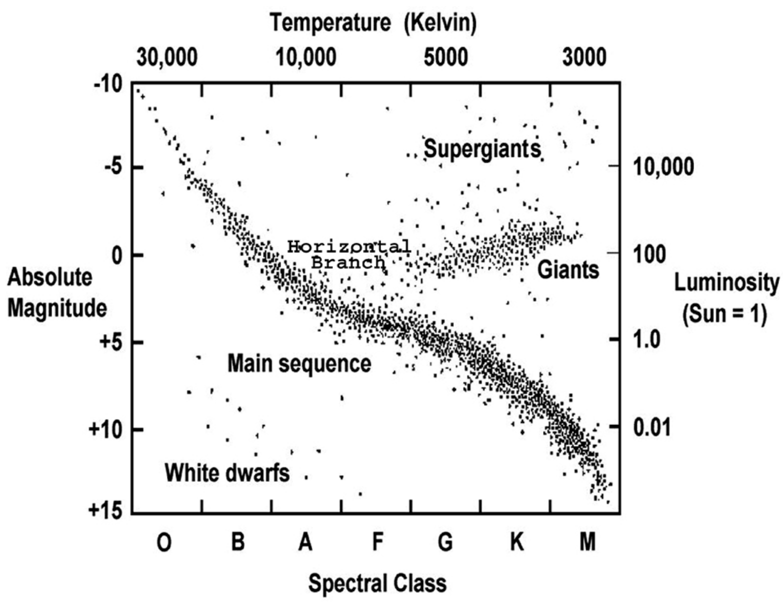


Fig. 2.1: The HR diagram. Courtesy: Chandra X-ray Observatory

2.1.5 Date

Tracking physical quantities with respect to time is very important for variable star research or time domain astronomy in general. We cannot use the Gregorian date in the numerical calculations. One cannot calculate the number of days between 2 March 2018 and 3 July 1967 with one subtraction. We need a simple time-line, hence we use a perfectly linear time-line called the Julian Date (JD). Today's JD is the number of Julian days passed since 0JD which was 2.4 million days ago. JD is a real number, meaning you can note down the Julian Date of an observation up to microseconds with the help of decimals.

The Julian Date for the second this is being written at is 2459365.85565, and just a second later now it is 2459365.85575. Greenwich Mean Time is considered 'Universal Time' and astronomers should remember to subtract the time difference between their longitude and GMT before calculating the Julian Date.

However this does not solve all our problems. Since speed of light is finite, it could take light coming from a star lying in the plane of the solar system 16 more minutes to reach the Earth in December as compared to June. While revolving around the Sun the Earth changes her distance to stars in a periodic manner, taking starlight different amounts of time to reach Earth. Julian Date being a geocentric system cannot account for it, so we have Heliocentric Julian Date (HJD), which uses a heliocentric frame of reference. HJD is calculated by adding a time correction (ranging from -8.1 min to +8.1 min roughly) to the JD by a function of heliocentric coordinates of the star and the heliocentric angular coordinate of the Earth. This calculation is not always needed, especially when dealing with large periods.

2.2 Tools of the trade

2.2.1 Mathematical tools

In this section we will learn to interpret visualisations that come out of variable star data analysis.

Physical quantities and time-series

A time series describes a physical variable with respect to time. In the study of variable stars one uses prominently 3 kinds of time series:

- **Lightcurve**
  The lightcurve is a plot of star's magnitude or luminosity values in y-axis and increasing time in x-axis. A lightcurve is the primary tool in the search for variable stars. It offers information about the size variation, geometric conditions of multiple star systems, rotation of stars etc. This is the only time-series we used in this project.
- **Radial Velocity**
  If you have spectral data available for your variable stars, you can calculate the radial velocity from the spectrum and create an RV time-series. This time-series offers complementary information to the lightcurve, such as rate of expansion of stars, stellar dynamics and rotation etc.
- **Temperature**
  From the spectrum one can also plot the surface temperature against time which gives us a good idea about the astrophysics of pulsating stars.

The three main physical quantities that go into identification or classification of variable stars are period, amplitude and colour index.

We tend to use periods more than frequencies (frequency=1/period) as it is numerically convenient. The periods for variable stars range from 0.1 day (about 2 and half hours) to 8000 days, however the list is short for long periods (periods greater than 300 days, let's say). Amplitude is the maximum observed variation in magnitude. The amplitudes range from fraction of 1 magnitude in small amplitude variables to the order of 7-10 magnitudes in the case of novae. In the classic book 'Variable Stars', prof. Cuno Hoffmeister writes that variable stars with amplitudes smaller than 0.1m cannot be called variables as their variation is just normal deviation in stellar brightness. All stars give out a slightly stochastic output; however the amplitudes (depth of the transit dips) for exoplanets is much smaller than 0.1m and they have been defined as variability class 'EP' in the GCVS. For understanding the colour indices see 2.1.4.

Periodogram: The Lafler Kinman method

Periodicities in data are not always conspicuous. If your sampling frequency is larger than the signal frequency, you will never see the curve hidden in the data by just looking at the lightcurve.

For example, if you observe a star's brightness once per day, and the star has a brightness variation cycle of 0.5 days; at best you will catch one data point per cycle. With one or two data points you certainly cannot visualise a trend or a curve. What you could do instead is taking observations for an extended period of time, arbitrarily catching data-points in different parts of the periodic curve, assuming that the periodicity is constant. When you look at the data you collected over a 100 days, you will not see 200 smooth curves, but rather something that looks like noise; but if you knew about the period being 0.5 days you can cut the entire dataset in sub-data-sets of width 0.5 days and stack them on top of each other. If you do this you will see the smooth variation that occurs in half a day.

But when we look for new variable stars, we naturally do not know their periods. Here we use a computation technique known as the Lafler-Kinman technique. It is a computer program which once you feed in the magnitude vs time values, creates a periodogram and several phase diagrams. A periodogram is a plot which has periods or frequencies on the X-axis and a measure of how powerful those frequencies are in your data on the Y-axis. So for example, in our case, in the periodogram we will see a high value for period=0.5 days (or frequency=2/day), telling us immediately that the star has a period of 0.5 days.

The Lafler Kinman method works by iterating over a meaningful range of periods. For each period, the dataset is divided up into sub-data-sets, each with width equal to the period. These datasets are stacked on top of each other so that all the datapoints are in one period, corresponding to different phases. A measure of spread called θ is calculated. θ is the sum over phase of squared differences between magnitude values corresponding to each phase. Low θ value indicates a smoother curve.

What is great about this method is that you do not need to have sinusoidal curves for the periods to be picked up (like in the case of fourier transform) and you are allowed to have large gaps in your dataset. Sinusoidal wave-forms are rarely seen in variable star astrophysics so it makes the case for the Lafler-Kinman method which was first developed in 1964 to analyse stars of the RR Lyrae variability type (Lafler and Kinman (1965)).

Warning: One must beware of false periods arising out of sampling frequency and alias periods. For example let us go back to our consideration. If 0.5 days is a period then 1.0 days will also be reported as a period, so will 1.5 days and 2.0 days and so on. These are alias periods given by the formula:

$\frac{1}{P_{alias}} = \frac{m}{P_{true}} \pm \frac{1}{T}$

Where P(alias) is the alias period, P(true) is the true period, m is an integer and T is the regular spacing in the data.

The other case is of sampling frequency. Let us say you observe once per day at the same time. If you have a variable star with periodicity equal to exact one day, then you will never observe any changes in brightness as you happen to always catch the same phase everyday. In cases where there is a secular trend in the data, this 1-day period shows a significant value in the periodogram and one must remember that it is just because of sampling frequency. Obviously there exist genuine variables with period close to 1 day, but this caution takes the precedence.

Useful plots

- **The phase diagram** It is important that the reader understands how to read a phase diagram. In the x axis, the phase ranges from 0 to 1, or 0 to 2π in other words. Keep in mind that phase is a circular quantity and hence 0 and 1 are one and the same phase. For convenience we choose to view the phase from -0.5 to 1. Every point on the plot represents a brightness value, and remember that they are not arranged chronologically but as per the hypothetical phase they will have in case the data has this specific periodicity. If a variable is periodic with a meaningful smooth curve, we shall see that in the phase diagram. It is the most important tool of detection of variable stars.

- **The sigma-magnitude plot** This helpful plot allows us to instantly filter out the candidates for variable stars from a large pool. Here stars are positioned according to their mean magnitudes on the X axis, and the according to the standard deviation (called sigma) of their magnitudes on the Y axis.
  Most of the stars we observe come out to be non-variable. These stars have a natural standard deviation in brightness due to instrumental error and stellar surface activity. But in case a star has a large sigma as compared to the peers of the same

magnitude, it becomes a candidate for our survey, as it indicates a real change in amplitude. In general, the higher the amplitude the hotter the discovery.

2.2.2 Software

VaST: Variability Search Toolkit

VaST is a computer program developed by Dr. Kirill Sokolovsky (Sokolovsky and Lebedev (2018)). This software is the backbone of our project. The software takes a series of images of the sky (CCD images or digitised photographic plates), aligns them carefully, calculates their instrumental magnitudes, produces a sigma-mag plot and short-lists stars as candidates for variable stars. Once we provide information on the coordinates and catalogue magnitudes of a few stars in the field, it calculates coordinates and magnitudes of all the stars in the field. It can also launch periodogram calculation and search the catalogues for the star; telling us whether the variable star has already been discovered.
It is an open source software also hosted on GitHub. Visit the homepage (vast.sai.msu.ru) for more details.

Online databases

The astronomical community maintains many online databases from which you can pull a lot of information about the star(s) in your query. You can look them up using names or coordinates and the databases provide your information about them. We shall now see some of the prominent databases that we have used in this project.

- **VSX: The international Variable Star indeX**
  VSX is a database hosted by the AAVSO, it aims to bring variable stars discovered from different surveys all under one roof. On the VSX portal you can search for a variable star by name to obtain all the information available. In case of an unconfirmed candidate, you can search the sky in a small circle around the position of the star, and see if anybody else has already discovered a variable star in that area. We used this tool to know which of our discoveries were new and which were old news.

- **SIMBAD**
  SIMBAD stands for a Set of Identifications, Measurements and Bibliography for

Astronomical Data. It is hosted by the Strasbourg observatory. It is a general purpose star and non-stellar objects (asteroids or galaxies for example) database containing over 10 million objects. One can look up all the available information about a star here.

- **Aladin**
  Aladin is a desktop or web software which allows the user to view the universe through the eyes of multiple sky surveys, and fetch data from them. This app was used to find positions of the stars for which VaST failed to calculate positions for.

- **ASAS-SN sky patrol** ASAS-SN stands for All Sky Automated Survey for Super-Novae. You can use this database to fetch lightcurves of the candidate stars calculated from recent surveys.

CHAPTER 3

Data

3.1 Observation

Now we begin discussing the specifics of the project. This chapter will brief you about the observation process, the telescope and the camera used, the people involved and the procedures that lead up to the creation of data-sets.

Observatory

The data we use in this project in the form of photographs was taken by the old school ground based observatories. Here a telescope is kept inside a moving dome, pointing at the targeted sky field. Keeping the Earth's rotation in mind, the telescope needs to turn along the Earth's axis in the opposite direction to stay on the same stellar target. This is done by a counter-weight and motors. There was also a monitoring position through which a trained human could manually ensure that the telescope stays precisely at the target.

The observations we use here were taken for various purposes (though likely for variable stars hunting). To understand our data, we must understand basics of observation. The Sun lights up the whole atmosphere of the Earth of the Sun-facing side, making it impossible to study stars in the day (unless you are a radio astronomer, of course!). Hence we are limited to that first order of blind-spots. Secondly, as the Earth goes around the Sun, the stars behind the Sun are obscured, causing seasonal blind spots. The Earth itself blocks stars from our view. People living on the equator get to see stars from all the parts of the sky over the year, while people living towards the poles would never see certain amount of stars which are blocked by the Earth at all times. For example, having never been to the southern hemisphere, I have never seen the famous Magellanic Clouds. But since I have lived near the equator, I have seen the southern Cross often, which my European peers might not have had the pleasure of seeing.

These were the systematic blind-spots, they are important even in the latter analytical phases because they are periodic. Their periodicity can create a specious or spurious peak in the periodogram, and one must make a note to be suspicious of periodicity of exact 1 day in variable stars. The other problem with observations is the weather. Ground based observatories are vulnerable to weather, becoming useless when the clouds cover up the sky. To avoid this, we try to build observatories on hills and mountains so they can poke their head out of the clouds. India has optical observatories in high altitudes such as Nainital and Yelagiri.

Let us come back to the plate-archive. Out of the 60000 plates in Moscow, the largest subset is of 22300 plates which were taken by a German made astrograph taken as a war reparation by the Sovyet Union. This telescope took around 20800 photos at the Crimean station of SAI starting in 1958. These are the plates we are making of our interest.

The telescope is a refractor telescope using lenses (and not mirrors). The diameter is 40cm and the focal length is 160cm. It offers a large field of view of $10°$ x $10°$. For comparison, the full moon is only half a degree wide! The exposures are long, in our case the plates were exposed each for 45 minutes to the starlight.

Plates

These plates are with the Sternberg Astronomical Institute, stored in special cupboards vertical within casings.

Photographic plates were the first generation of detectors which stored the photographs. One could now take long exposure pictures. Due to the long exposures, stars which were faint to invisible came into the light. This gave us the technology to study a much greater number of stars than before. Photographs were also preserved and could be reassessed desirably. The photographic plates we work with in this project are squares of side 30cm. They are smooth glass plates layered with a photo-receptive emulsion. The telescope directs starlight onto the exposed surfaces of fresh plates. The emulsion reacts with the starlight and turns dark because of it.

An important thing to understand here is that a plate is coated with a specific emulsion, and that emulsion corresponds only to a certain bandwidth of the light's spectrum. The plates we use correspond to the 'B' region in the Johnson-Morgan scheme. It is a useful tool for astronomy. If one wishes to study a pulsating red giant, then it is a good idea to use the 'V' (for visual) emulsion. This is an instrumental and manufacturing limitation. The limiting magnitude of these plates is 17^m to 18^m, which is fairly good, meaning they can measure stars which are at least 50,000 times dimmer than the faintest star we could see with our naked eyes! 1960s to early 1990s was the peak period for the observatory, on good years the number of total plates taken that year would exceed 800. The plates offer a great survey of the Northern sky.

The following is a table of important subsets of the Moscow plate collection (Samus *et al.*).

D (cm)	f (cm)	Field size (deg)	Mag lim.	Years	N	Site
10	64	20x28	13 -14	1895 - 1933	1100	Moscow
16	82	16x22	14	1933 - 1956	2700	Moscow
23	230	6x6		1955 - 1991	10000	Moscow etc.
38	640	1.4x1.4	14	1902 - 1972	6400	Moscow
40	160	10x10	17 - 18	1948 - 1996	22300	Kuchino then Crimea
50	200	3.5x3.5	18 - 19	1958 - 2004	10000	Crimea
50	200	Spectra		1959 - 2004	2300	Crimea
70	1050	0.6x0.6	13 - 18	1961 - 1995	9500	Moscow

Table 3.1: Contents of Moscow Photographic Archive.

In addition to photographs, observatories also kept a logbook with Julian Date for each photograph (the mid-exposure date). The coordinates could be read from telescope calibration and they are usually written on the plates and also in a book. (check)

3.2 Data processing

In the nineteenth century astronomers estimated the magnitude of a target star with their eyes by comparing with two other stars, one brighter and one dimmer. This practice translated to specially crafted analysers for photographic plates. In the current era the CCD images from the telescope can be analysed directly on a computer, various adjustments and manipulations could be performed on software with a Graphical User Interface. The automated space surveys write their own data reduction and analysis software.

In this project we picked an array of photographs from the archive, scanned the photographic plates, converted them to the astronomical image file format, used the VaST software to perform lightcurve extraction, went through close to a thousand lightcurves, rediscovered many known variables and found some new variable stars hitherto unknown.

Let us now understand some key steps in this process.

(a) The first step is choosing our dataset, for this we chose a well photographed field from the archives. We will only use photographs from the same telescope taken at same field of view and focal length. The size of all the photographic plates must be the same. In our case we choose the field centered at the coordinates: 22:40:00 +54:00:00. There were 127 photographs available with those coordinates as their center. The field of view was $10°$ x $10°$. The plates are 30cm x 30cm in size.

(b) These frames are carefully scanned with the help of Epson Expression 11000XL scanner at 2400dpi resolution. The photographic plates on their own do not have a pixel resolution, and our scanning resolution is important. This resolution is great, however it costs a lot of storage. The scanned images are very large in size, so we will cut them into pieces and feed it to the software.

(c) The images originally in '.tiff' format are converted to the '.fits' file format (FITS: Flexible Image Transfer System). One can store many other information points about an observation in a *fits* file, the most important being the date and time of the photograph. This is done by manually writing the mid-exposure Julian Date into the file taken from the logbook kept by the observatory.

(d) The photograph is divided into sub-fields so handling is far easier and we can deal with the data one bite at a time. All photographs are supposed to have the same center, we draw a $1.2°$x$1.2°$ square around it and set up a grid of squares. We name the center field as having (00,00) discrete coordinates, the one to the left being (-1,00), the one on up side being (00,01) and so on.

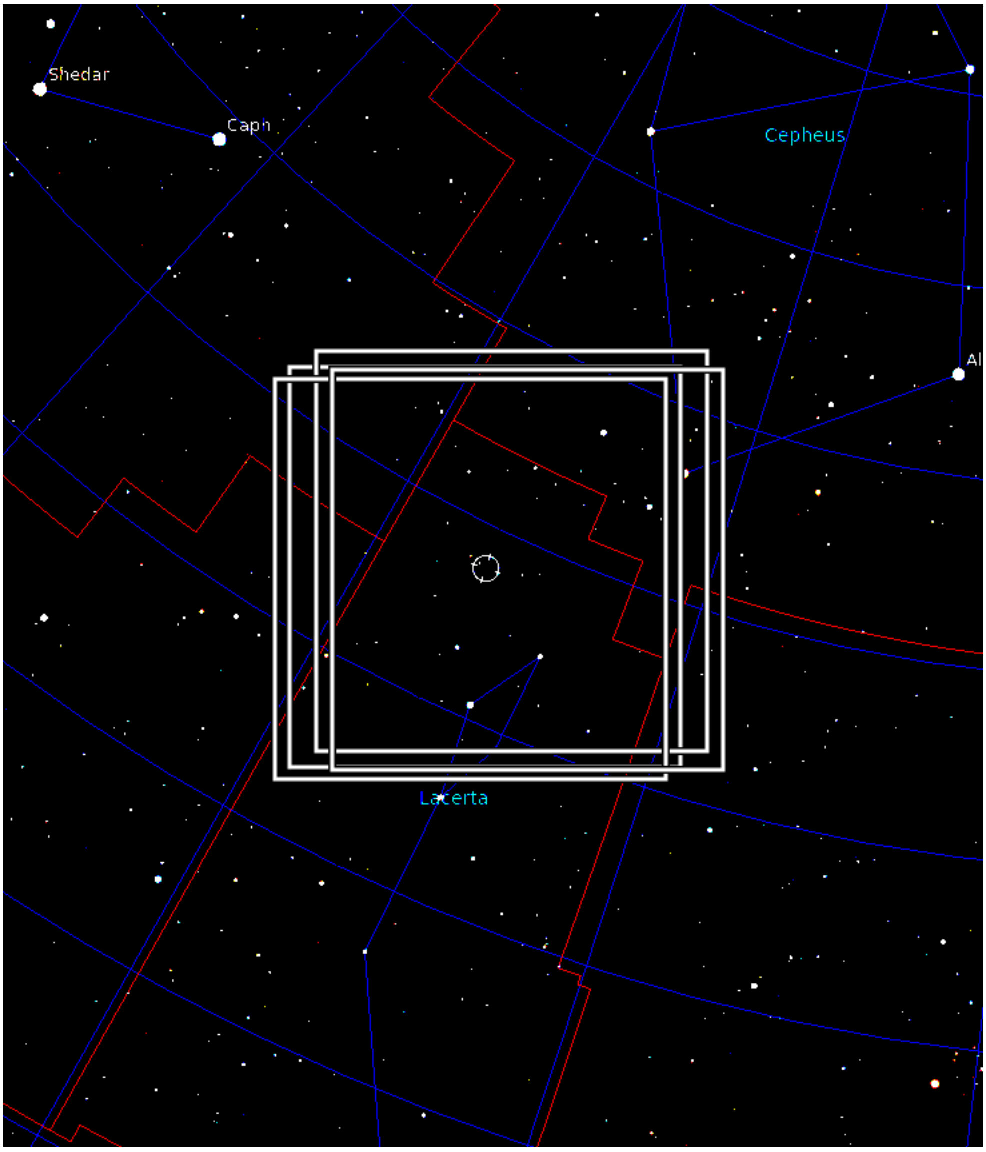


Fig. 3.1: A visualisation of the field chosen for our study. The squares symbolically represent various slightly different frames centered approximately at 22:40 +54:00. Software used for the image: stellarium.org

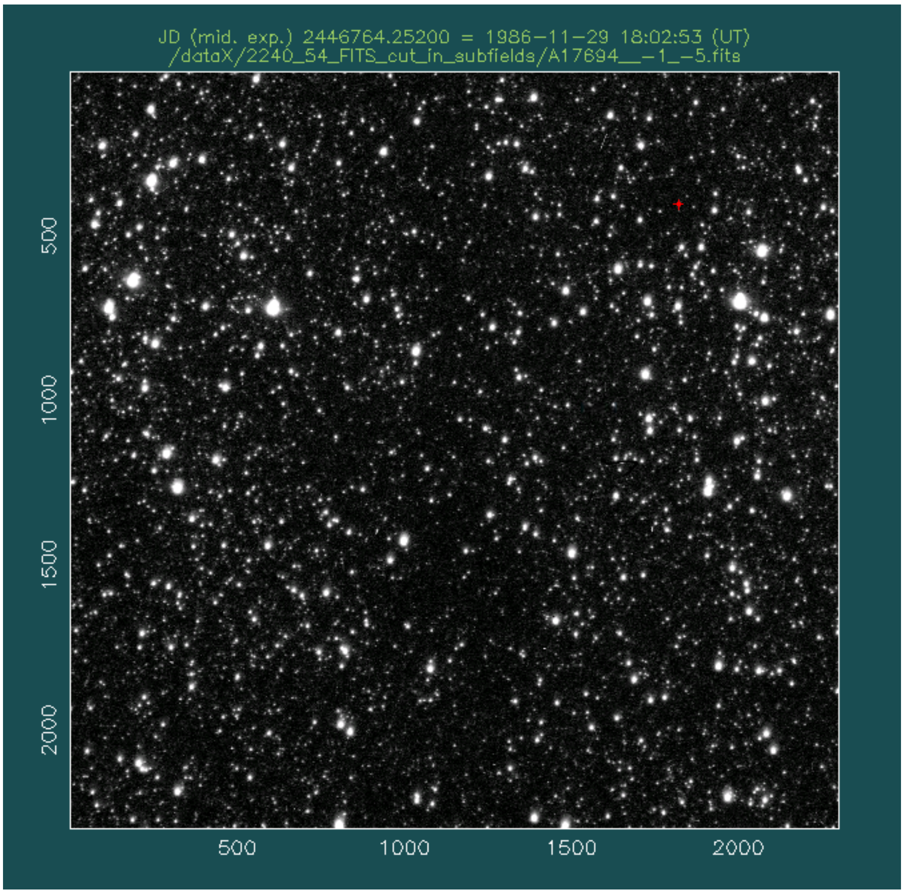


Fig. 3.2: One of the sub-fields of our plates

Please note that the frames might be a little bit rotated, or shifted in axis so a particular sub-field in all the frames will not be identically the same. The software takes care of this issue; but what it practically means is that we have different number of datapoints for lightcurves in the same sub-field. Stars that are present in all the frames will have 127 data points (this is very rare), the edge stars might only get 40 data points.

(e) Now we have split our set of 127 images into over 100 sets of 127 smaller images. We apply the software VaST on each sub-field. It is done by writing a script that automatically runs VaST on all the sub-fields and saves the lightcurve information for each star in '.dat' files in special save folders. The software identifies the same stars in all the images, uses a tool named SExtractor to calculate the instrumental magnitude of the stars.

(f) Many important stars are identified in the field, their apparent magnitudes are found in catalogues or online searches. With these catalogues we then calibrate the whole field, translating the instrumental magnitudes to actual magnitudes.

Our processing is now complete and we are ready to look at the lightcurves.

CHAPTER 4

The search and results

In this important chapter we will discuss our search for variable stars and the results of our search. First we will learn how to look at lightcurves, phase diagrams, periodogram, colour index, photographs etc in order to detect variable stars. Then we will study the procedure of classifying the stars and confirming variable star discoveries. In the second section we will look at the results of the search so far. The technology and terminology has already been explained.

4.1 How to search for and classify variable stars

(a) The first step in shortlisting candidates for variability is the sigma-magnitude plot. It plots all the stars in a field with respect to their magnitude and the standard deviation in their magnitude. Ideally one would expect a flat line which grows thicker towards the end as the accuracy of magnitude calculation decreases for low magnitude stars; but we observe that many stars have unusually high standard deviations. The other normal stars form a band, whose standard deviation comes from various errors between observations. The expected standard deviation is a result of noise, limited resolution, differing weather conditions etc.
It gets interesting when a star has a standard deviation much larger than the expected value, because it is the first indicator that the star is varying its magnitude. We manually open the lightcurves of these stars (see 4.1) to proceed. It is not worth opening lightcurves of stars whose standard deviation is not much larger than the average; because any patterns you see could just be noise and you cannot prove otherwise.

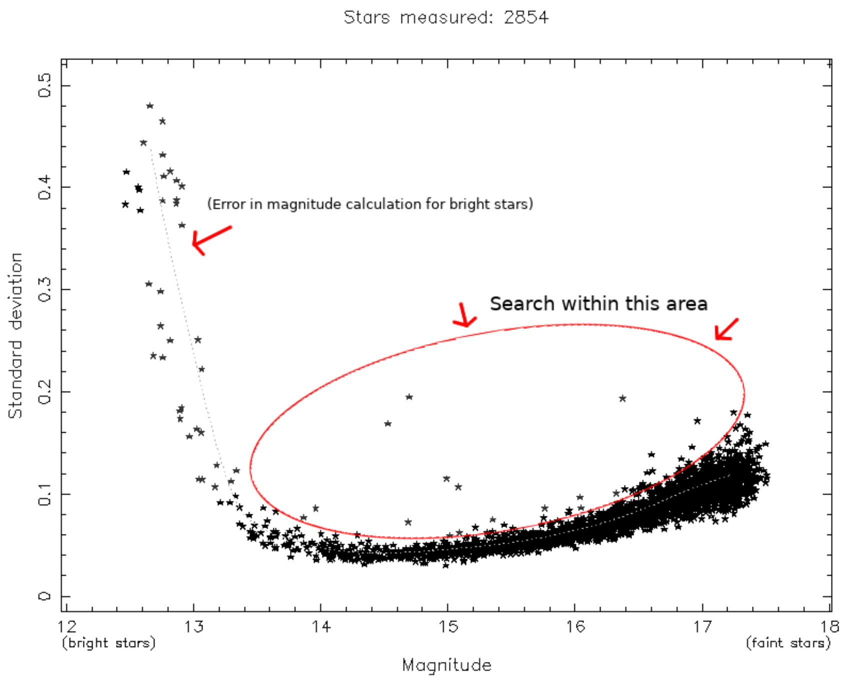
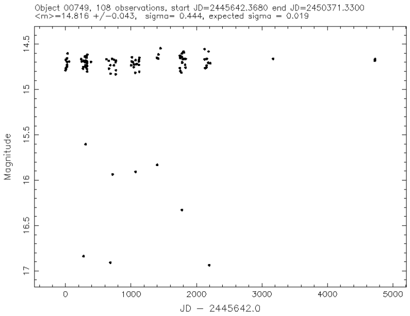
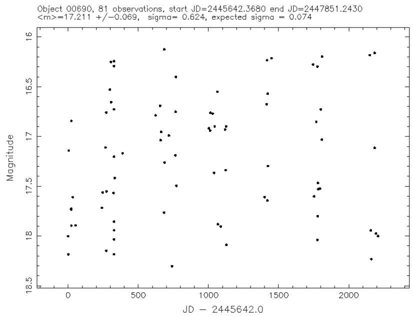


Fig. 4.1: A sigma - magnitude plot.

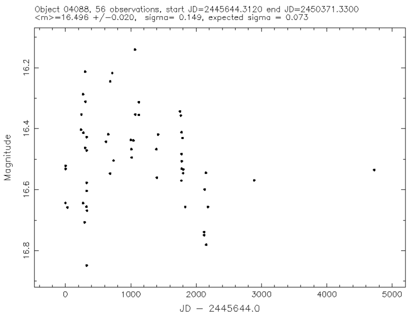
(b) The second step is taking a look at the lightcurve. Lightcurve can give you a quick idea of what the star could be. Keep in mind that the data we have spans roughly 100 observations spread across thousands of days so the smooth lightcurves as seen in textbooks will seldom appear. However, for the long period variables or certain types of irregular variables one immediately sees a long term behaviour in the lightcurve.
Apart from that, in obvious cases you can still recognise what kind of variable star you are looking at just by looking at the lightcurve. For example, in cases of eclipsing stars of the Algol kind, you can distinctly see the eclipses apart from the normal magnitude.
In datasets where you have multiple observations per night, you can use press 'Z' in VaST to zoom into the lightcurve of only one night, and you can often see smooth variations in magnitude (see 4.3d).
If a nova outburst is present in your photos, you could observe it very easily in the lightcurve.



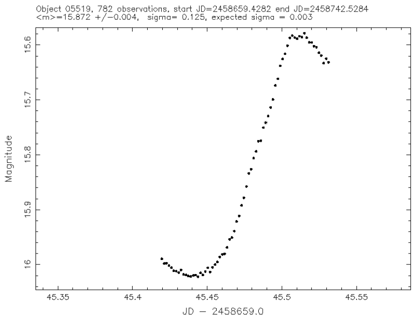
(a) Eclipsing type, like Algol.



(b) Pulsator, RR Lyr



(c) Long period variable



(d) RR Lyr CCD close-up

Fig. 4.3: Some recognisable lightcurves

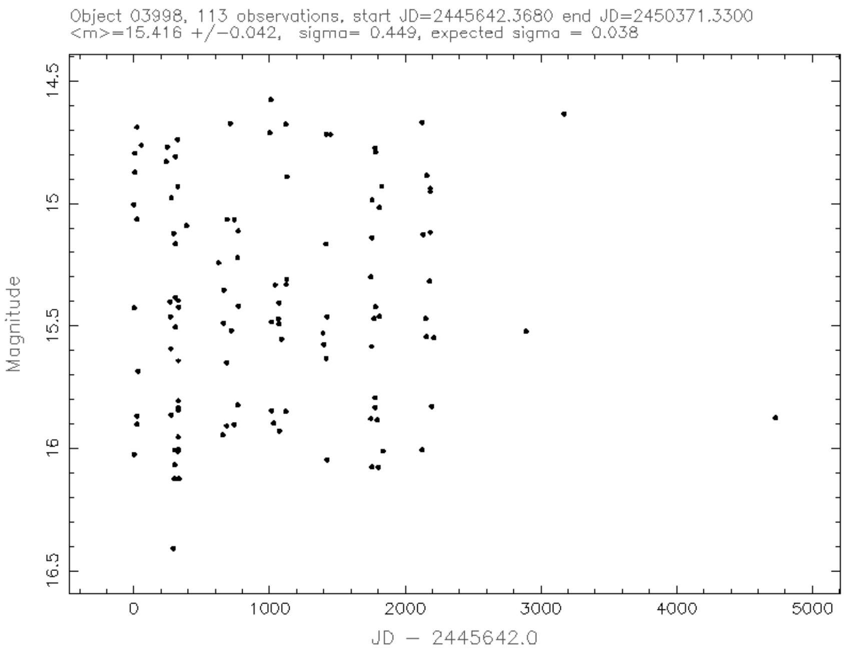


Fig. 4.2: A typical lightcurve in our project

(c) At this point it is important to take a look at the photographs. Especially in cases of novae, one should make sure that the supposed explosion is not reported

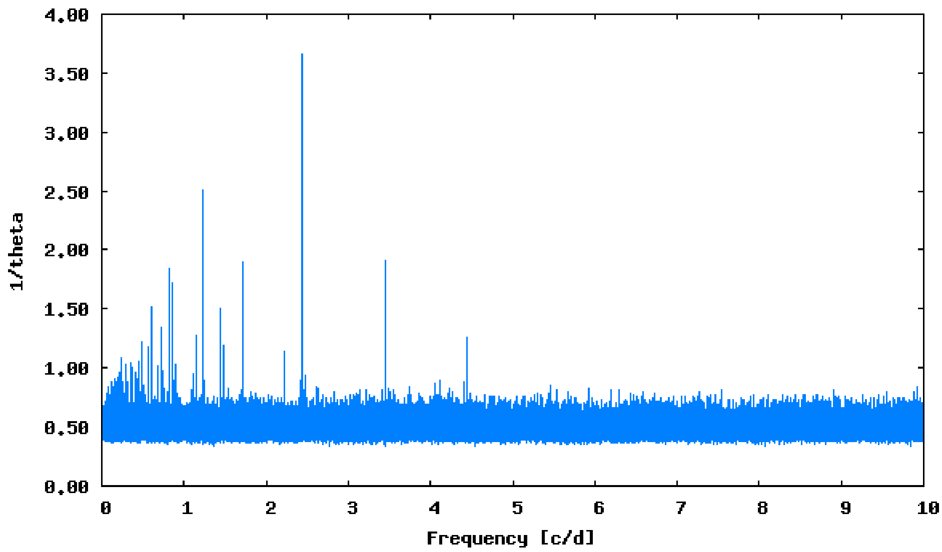


Fig. 4.4: What a good periodogram looks like for periodic variables.

because of a piece of dust on the photographic plate. So quickly inspect photographs for photographic artifacts and errors. You should also check if your star is too close to other bright stars, if they are not clearly resolved, then the fluctuations may cause variation. Many times low magnitude values are caused by incorrect calibration owing to bad weather or plate edge. So it is always a good idea to look at the pictures.

(d) The next step is calculating the Lafler-Kinman periodogram and phase diagrams corresponding to important periods. You can do that by pressing 'L' in VaST. In case of periodic variables with constant periods such as eclipsing variables or δ Cepheids you can expect to see proper curve shapes. In the periodogram, good periods will have tall peaks and alias periods on both the sides.
In case of long period variables, you will probably see a spurious period of exact 1 day (and multiples of it). Such a period is detected when there is a overall trend in the larger lightcurve (for example if the mean magnitude is increasing linearly). The semi-regulars will be harder to catch from periodograms, one would need to take a refined look at the lightcurves and photographs to make sure if they are seeing irregular or semi-regular variability.

(e) At this stage it is possible to assign a preliminary classification. One should check the colour index, as the spectral class of the star could heavily contribute to the classification. Periods also give an idea about what kind of variable it must be. For example, RR Lyr and Mira are both pulsating stars, but they lie in different parts of the HR diagram and RR Lyr stars vary on the scale of one day while Mira stars have periods of many months.

(f) Once we create a list of candidates, we then check catalogues and online databases for stars at those coordinates. A large majority of the stars that we detect as variable will already be discovered (especially in our case as our data has lower accuracy). Out of the stars that remain, one must ensure as much as humanly possible that the stars are not already discovered and described somewhere.

4.2 Discoveries from the 1983-1996 data

This section will describe the achievements of the project up till this point. The project is still in process at the time of submission of this report and we will only describe here the variable stars found so far in this field.

The work done on this project by mewas done partially remotely and partially in person in Moscow. So far I have covered 31 sub-fields, each of size $1.2°$ x $1.2°$. In each field I looked at between 20 to 30 lightcurves of stars with unusually high sigma. The vision of identifying variable stars comes slowly and your speed and efficiency increases with time. The 'Found variables' list has 168 entries, out of which 140 were known already, 17 are newly discovered variable stars worth publishing and 11 are suspected variables which need more data to be discerned. **??** is the table of all the discoveries made so far in the project.

Field	Coordinates (J2000)	Variability	Mag range	P (d)	J-K	Exp. Class.
-1_-4	22 34 48.43 +57 07 10.2	Periodic	15.8-16.6	0.55	0.40	EB
00_03	22 42 37.33 +50 37 37.7	Periodic	16.0-17.2	0.46	0.47	EW
00_04	22 42 16.89 +50 04 04.6	Irregular	16.1-17.5	0.99(!)	1.06	TBD
00_-3	22 45 24.41 +56 16 51.5	Long period	16.1-16.8	?	0.75	L
00_04	22 41 47.89 +50 17 10.8	Periodic	15.2-15.8	1.00	0.30	EA
01_-5	22 51 05.66 +57 33 21.1	Long period	14.0-14.4	?	0.39	L
01_-5	22 48 55.09 +57 15 19.8	Periodic	15.8-16.4	2.59	0.38	EA
02_-3	22 56 35.78 +55 26 53.9	Periodic	16.0-17.0	1.33	0.30	EA
02_-4	22 52 36.28 +56 54 54.1	Periodic	15.6-16.2	1.12	0.30	EA
01_-5	22 38 28.08 +57 19 18.3	Irregular	15.3-15.8	0.99(!)	0.58	L
00_-5	*22 42 20.77 +57 38 26.6	Outburst	17.5-16.0	NA	?	Nova like
-1_-4	23 36 17.66 +56 48 04.7	Periodic	16.3-17.6	0.59	0.39	EW
02_05	22 52 53.13 +48 57 16.8	Periodic	14.8-15.6	0.125	0.37	EA
02_05	22 52 14.33 +49 34 14.9	Periodic	12.8-13.0	0.99(!)	0.70	SR?
02_-4	22 57 45.30 +56 24 24.6	Slow irregular	15.2-15.7	NA	?	L
02_05	22 53 01.41 +48 55 32.6	Periodic	13.0-14.5	1.99	0.65	EA
00_05	22 43 32.35 +48 58 20.4	Periodic	14.8-15.1	?	0.14	L

Table 4.1: Table of discoveries

'?' - Unknown

'(!)' - Warning for spurious periods

'*' - Coordinates belonging to a star nearby

'TBD' - To be determined

Generally in any photometric survey, about half of the discoveries are of contact binary stars (of the classification 'EW' or 'EB') hence it is unsurprising that it also the case with our data. Owing to the long term nature of our data, we have found more long period variables than expected. We have not so far discovered new stars which are strong periodic pulsators (such as δ Cep or RR Lyr), however we did detect a large number of known RR Lyr. The candidates for novae often turned out to be dust on the plate, or in one case it was an asteroid which had moved in front of the star that night.

In cases where there was a reasonable suspicion of outbursts, we could only find the outburst recorded on one plate. Discoveries of new rotating variables was not going to be possible as their amplitudes are too small.

The lightcurves of these discoveries are shown in the next chapter. It is interesting to note that despite the data being old the discoveries are new. These photos sat on a shelf for decades with useful data inside them. Automation allows us calculate lightcurves of tens of thousands of stars without human analytical effort in a matter of hours. The long span of our study also gives us an edge.

4.3 Scope for Machine Learning application

This section requires the reader to have basic knowledge of machine learning algorithms and terminology.

The authors of the book 'Variable Stars' wrote about the developments in computer technology, when they were writing the book in 1980s. Back then they did the analysis and computations manually, as computers were still in infancy. From their writings it is clear that they could not have predicted how the computer age would come about and swipe away the manual methods. They predicted that the technology would use punch cards and wrote that the process has many errors and the computers are expensive. Today we stand at a similar point with machine learning. It is hard to predict where this technology will take us in the future and what it would mean for variable star astronomy.

Currently a human analyst (viz me) is still needed to look at lightcurves, phase diagrams, periodograms and spectra to classify a star. This process can be completely automated soon enough, and the algorithms are presently in development. See Antipin *et al.* (2018) to read more about the classification attempts using random forest algorithm. I wanted to write an automatic classifier programme for this report, unfortunately time did not permit me to do so. I shall now describe what I have learned so far.

There are two potential ideas, one project is classification or clustering based machine learning while the second project is training neural networks to recognise phase diagrams like we train them to recognise hand-written digits and letters. The mathematics of the first project is simpler, however the problem might be the more challenging one. The second project could give powerful results if configured properly.

Datasets are crucial in machine learning, and for this we have the GCVS catalogue containing 60,000 stars. The catalogue lists as the attributes the coordinates, period, magnitudes, spectral type and the classification as the label. Spatial coordinates are only important in some classifications, as certain type of variables are majorly found in certain parts of our host galaxy. Such information is perfect for a random forest classifier, for example. However when you consider that not all the stars have all the attributes, the list shrinks down to only 10,000 stars who have all their attributes written. The magnitudes are not provided to GCVS in a standard format, so sometimes a range is mentioned, while sometimes minimum and amplitude is provided. That can be solved by simple list manipulation. The next challenge is the variety of spectral types. The total number of unique spectral types in the GCVS is easily over 50. You could possibly solve this by only taking the first two letters of the spectral type and deleting the rare cases. However, the biggest challenge in classification is the classification types. Not counting multiple designations, the number of unique classification types in GCVS is roughly 80. 80 labels is too many labels. One could attempt to fix this by clubbing similar types together into one parent type (RR, RRC, RRAB as just RR for example). That would still leave one with 20 or so types and that is also too many. The last resort could be simply training the algorithm to classify the variable star as one of the broad categories (eclipsing, pulsating, cataclysmic, eruptive, rotating). But it might be a waste of time because the distributions of period, spectral type and magnitude between these categories are similar anyway.

I believe that instead of supervised classification, unsupervised clustering would work wonders for variable star research. All our current variability types have come about because scientists did the clustering process. Clustering is creating clusters of similar data points. If we do this, we can expect the RR stars or Mira stars to all come together in one cluster. And we might even find new relations and clusters between variable stars. Being unsupervised, it might lead to some arbitrary results, ones we can't make sense out of.

Once it is known that a variable star is of eclipsing type, further classification is simple;

but that is not the case with pulsating variables where there are so many times with so many conditions. That is why my initial idea was only to create a tool to give you possibilities of classification, to aid manual classification.

A pulsating star and and an eclipsing system might have the same period, magnitude and spectral type, hence the algorithm cannot know what it really is. The most important of our tools, the lightcurve remains unused in these attempts. This is where the second project comes in. The 'hello world' of deep learning is writing a programme that recognises hand-written digits after training on thousands of images with given answers. Phase diagrams are not far from symbols of digits and hence I reckon it might be possible to do this. Data could be obtained from Kepler databases. Such training will probably require a big computer, unless transfer learning is used and we train an already trained algorithm.

Machine learning applications is a great field to work in right now, the possibilities are limitless!

CHAPTER 5

Variable star physics

In this chapter we will understand the interesting phenomena and dynamics that give rise to different kinds of variability in stellar magnitudes in the light of our discoveries and findings.
In the second section we will study the use knowledge of variable stars serves and how this project contributes to the research. Please visit the appendix A to get a list of all variability types and their descriptions.

5.1 Variable star physics

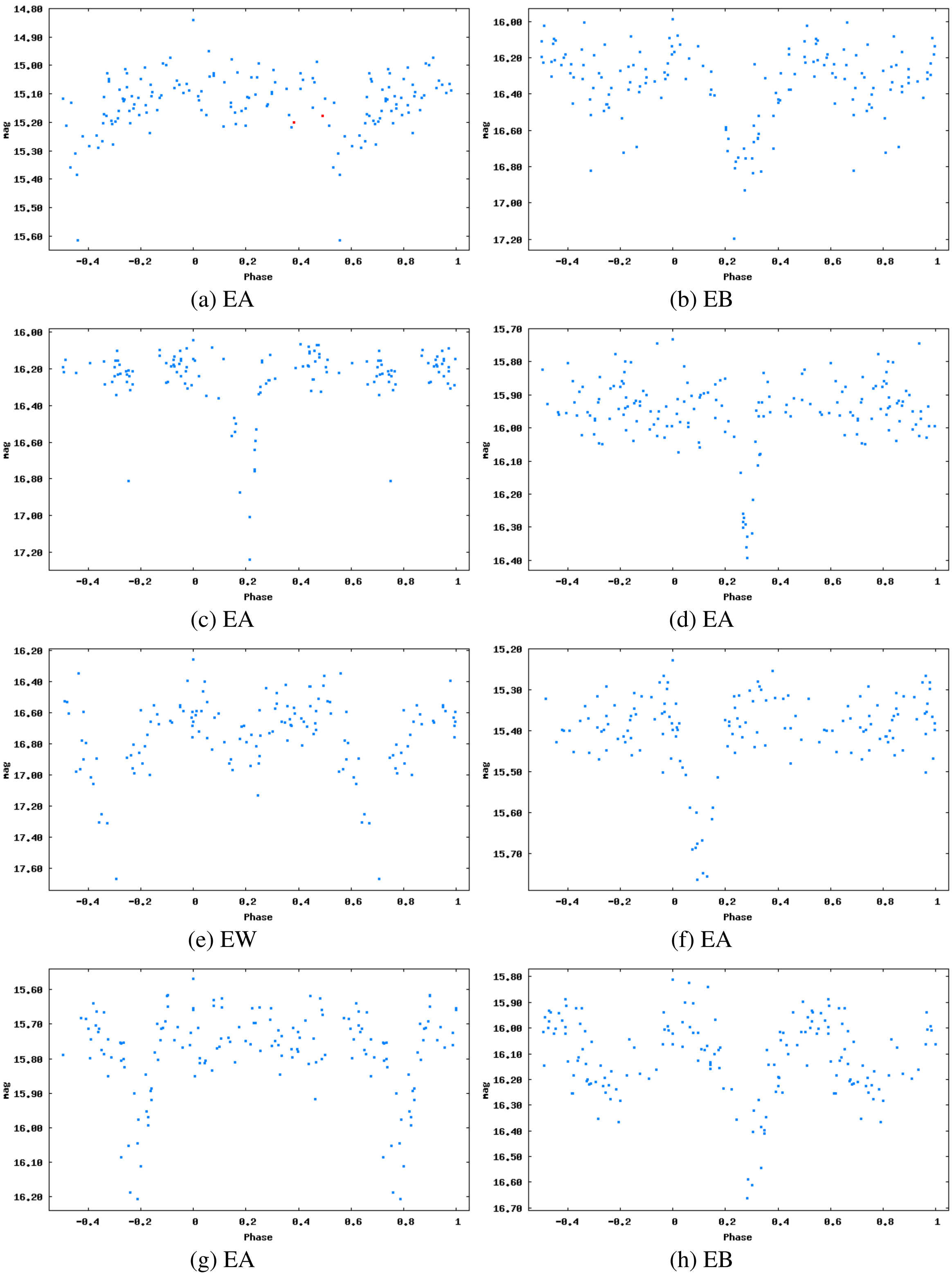
We will be discussing the broad categorisations of the variable stars arranged according to their frequency in our findings. It is interesting to see the kind of information about the stars' physical properties could be deduced from their lightcurves and spectral classes.

5.1.1 Eclipsing systems

This is what we call an extrinsic variable, the changes in the apparent magnitude of this 'star' are caused by external effects of eclipses and transits rather than internal astrophysical changes. As a result one observes a really well defined periodicity and curve shape. They are found really commonly, especially the ones with small orbital periods; in fact the majority of the variable stars found from general variable star search surveys are of the GCVS classifications EW, E, EA or EB; which are different types of eclipsing variables.

Lightcurves

Below are the phase diagrams of the **newly discovered** eclipsing variables found in our data.



(a) EA (b) EB

(c) EA (d) EA

(e) EW (f) EA

(g) EA (h) EB

Fig. 5.1: Phase diagrams of newly discovered eclipsing binaries

History

The second non-nova variable star ever discovered was Algol (al-ghul in arabic, ie. *the demon*). It is named 'β Per' under the Bayer designation. This star dimmed twenty times (however according to Fechner's principle it only feels like 1.3 times) every three days. British amateur astronomer John Goodricke proposed an eclipsing mechanism as the explanation for this variability, a revolutionary idea at the time.
Since then thousands of eclipsing variables have been found, prompting a great study into binary stars.

Occurrence and Probability

We terrans are used to looking at a singular star as our primary Sun. The Sun is spherical in shape and is single. But our study shows that it might even be a minority. We discover binary stars very commonly. Some of them are so close that they share an atmosphere, resembling the shape of a peanut. Others have really large distances between them.

Lightcurves and classification

Initially the eclipsing lightcurves were classified according to the shape of thei lightcurve. Initial classification was simply 'E' for eclipsing variables. Once the diversity became apparent, 'EA' was introduced for Algol type stars, EB for β Lyrae type stars, EW for W Ursae Majoris type stars and EP for exoplanets. EA is ascribed to those lightcurves which have a deep primary minima, very distinct from the secondary minima. The secondary minima is often difficult to see. EB is for lightcurves with rounded curves between minima. EW is for contact binaries with really short periods, typically around half a day. The primary and secondary minima for these stars are consecutive.

Today a physical classification is used to describe the stars, based on 3 dimensional gravitational equipotential surfaces around the two stars. Langrangian mechanics plays a great role here. Between two stars, there exists a point where the total gravitional force experienced by an observer equals 0 as the stars cancel each other's effect at that point. The equipotential surfaces around a binary system resemble a peanut shell or an hour-glass. The tear-drop shaped equipotential surface around a star in a binary system

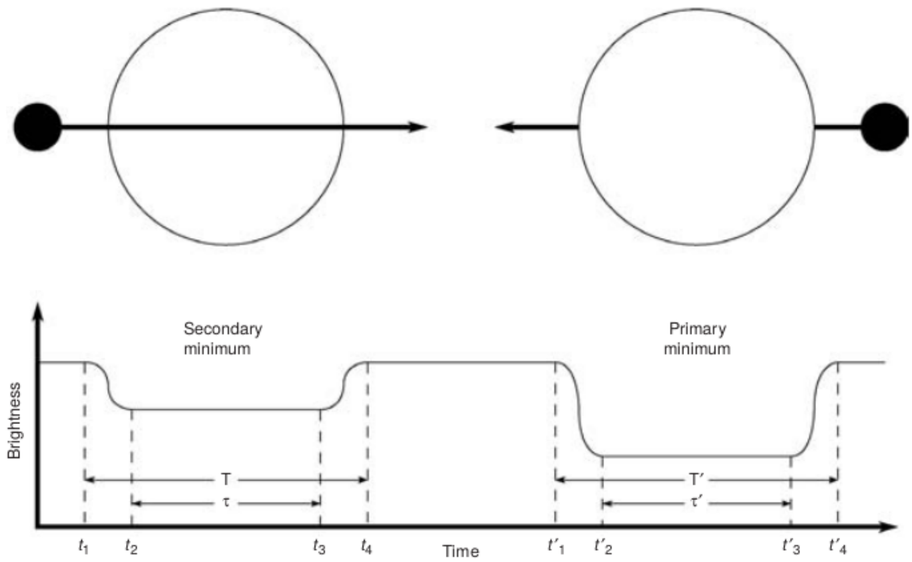


Fig. 5.2: The mechanism of brightness dips. Courtesy: 'Understanding Variable Stars'- J.R.Percy.

is known as a 'Roche lobe'. The stars may or may not fill up their Roche lobe. It even happens so that a giant star grows in size due to its own nuclear processes, and fills up the Roche lobe in the shape of the Roche lobe. Once it starts to overflow, it leaks from the Roche center into the Roche lobe of another star. Stars might even be so close that they share the same atmosphere!

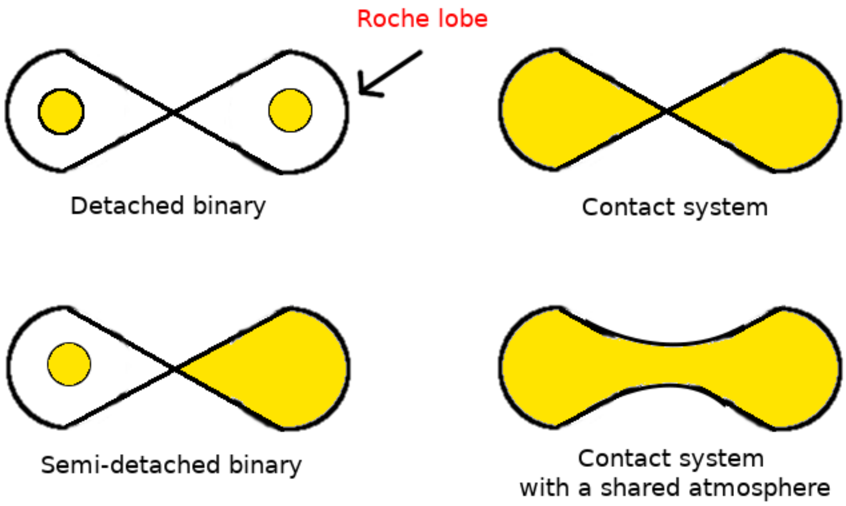


Fig. 5.3: Types of binary systems

Physics

One can calculate many many properties of the stars constituting a binary system with just the lightcurve, spectrum and radial velocity data.
From the lightcurve, we can ascertain two important things: period and luminosity ratio. At the primary minima we only observe the dimmer star as the brighter star has gone behind that one, making it easy to calculate how much dimmer it is from the first one. Once you know the temperatures of these stars from the spectra, you can calculate the individual apparent magnitudes of these stars. From the shape of the curve leading to the minima we can calculate the shape of the smaller star. That is the time it crosses the boundary of the bigger star. Radial velocity data combined with period and Kepler's third law gives us the relative masses of the stars, and that information is very important.

5.1.2 Pulsating stars

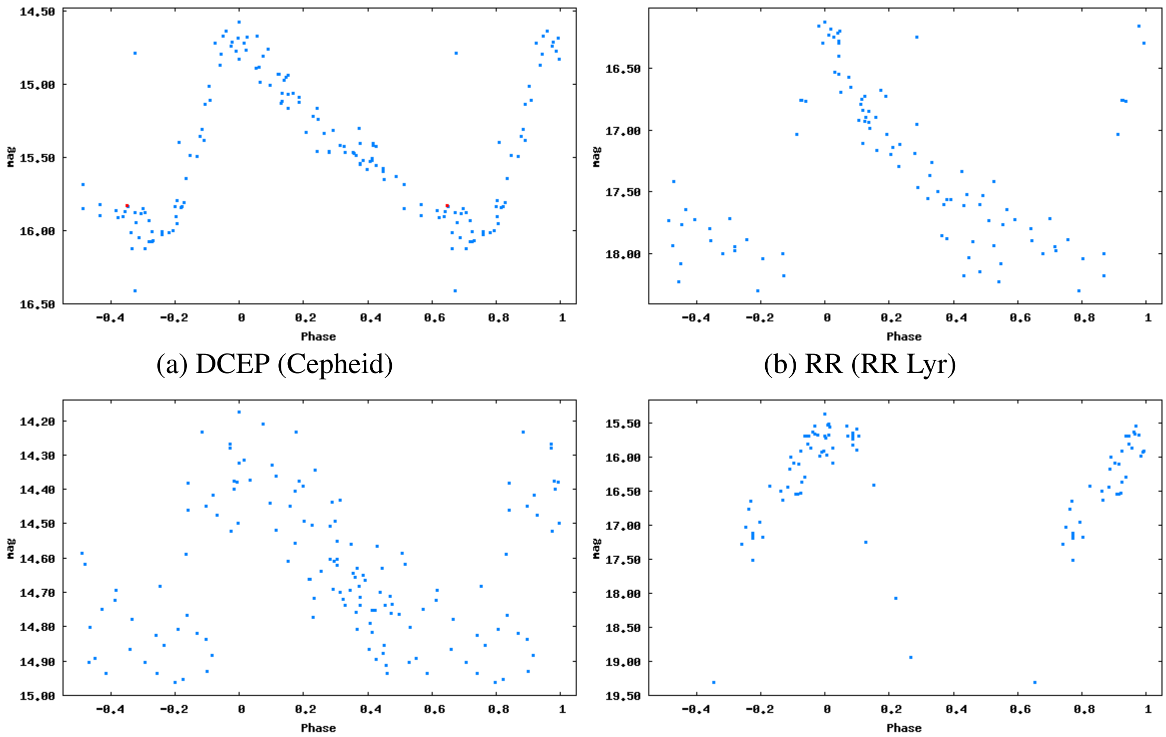
The first periodic variable star ever discovered was a slow pulsating variable with a high amplitude. These were the variables with a neatly defined periodic lightcurve. These curves did not resemble the sine function, and they came in many different shapes such as Mira, Delta Cepheids or RR Lyrae.

History and basics

o Cet or Mira and δ Cep or Cepheid variables had been discovered very early on. However a satisfying explanation was not offered until the second half of 1910's. Then the theory of pulsation was formulated and applied. This theory was a success. It explained the lightcurve from a pulsation mechanism, suggesting that stars expanded and contracted in size like a vibrating string. When the HR diagram came into picture, an interesting fact was seen. The stars with the same lightcurve shapes tended to be clustered in specific regions of the HR diagram; this led to the idea that the vibrations are not characteristic of the particular star, instead vibrations are a state caused by regular evolutionary changes in the star's life. Today many kinds of pulsators have been studied, some pulsating in multiple modes, some pulsating non radially and some pulsating semi-regularly.

Lightcurves

The below are the lightcurves of some of the pulsating variables we found in our data. All of these were already known. We have not discovered a good new pulsating variable yet.



(a) DCEP (Cepheid)

(b) RR (RR Lyr)

(c) RRAB (Asymmetric RR Lyr)

(d) M (Mira)

Fig. 5.4: Examples of phase diagrams of pulsating variables in our data.

Analysis and modelling

Pulsating stars are located in the so called "instability strip" within the Hertzsprung Russell diagram. When stars leave their main sequence timeline to become giants they redden and pass through this portion of the HR diagram. Different kinds of pulsators are located in their individual distinct zones.

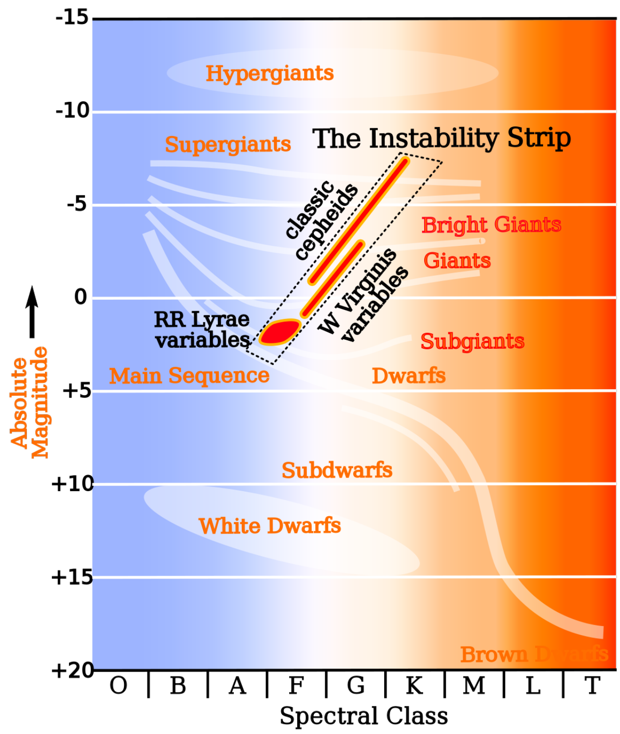


Fig. 5.5: The instability strip. Courtesy: user Rursus@Wikipedia

In 1908 Henrietta Swan Leavitt of the Harvard College Observatory first discovered the important 'period-luminosity relation' in cepheid variables from the Magellanic Clouds. She noticed that when plotted against the log of period, the magnitudes of these stars followed a linear trend, implying brighter cepheids pulsated faster. This simple empirical relation proved to be paramount importance to astronomy as it lets us estimate the absolute magnitude of a cepheid variable by finding the period of pulsation. Later it was found that many other pulsators follow a period-luminosity-colour relation. The general period-luminosity-colour relation is as follows:

$$\log P = a + b \cdot M_v + c \cdot (B - V)_0$$

Where P is the period of pulsations, M_v is the magnitude in the V band and $(B - V)_0$ is the subtraction of magnitudes in B and V bands, the colour index.

5.1.3 Irregular or long period variables

In the GCVS classification scheme, these have been included under pulsating variables. Despite knowing this, we create a different section to discuss them as they are not clearly understood, and since the nature of our data has led to many discoveries of this type in our project.

?? shows the lightcurves of long period variables we have **newly** discovered.

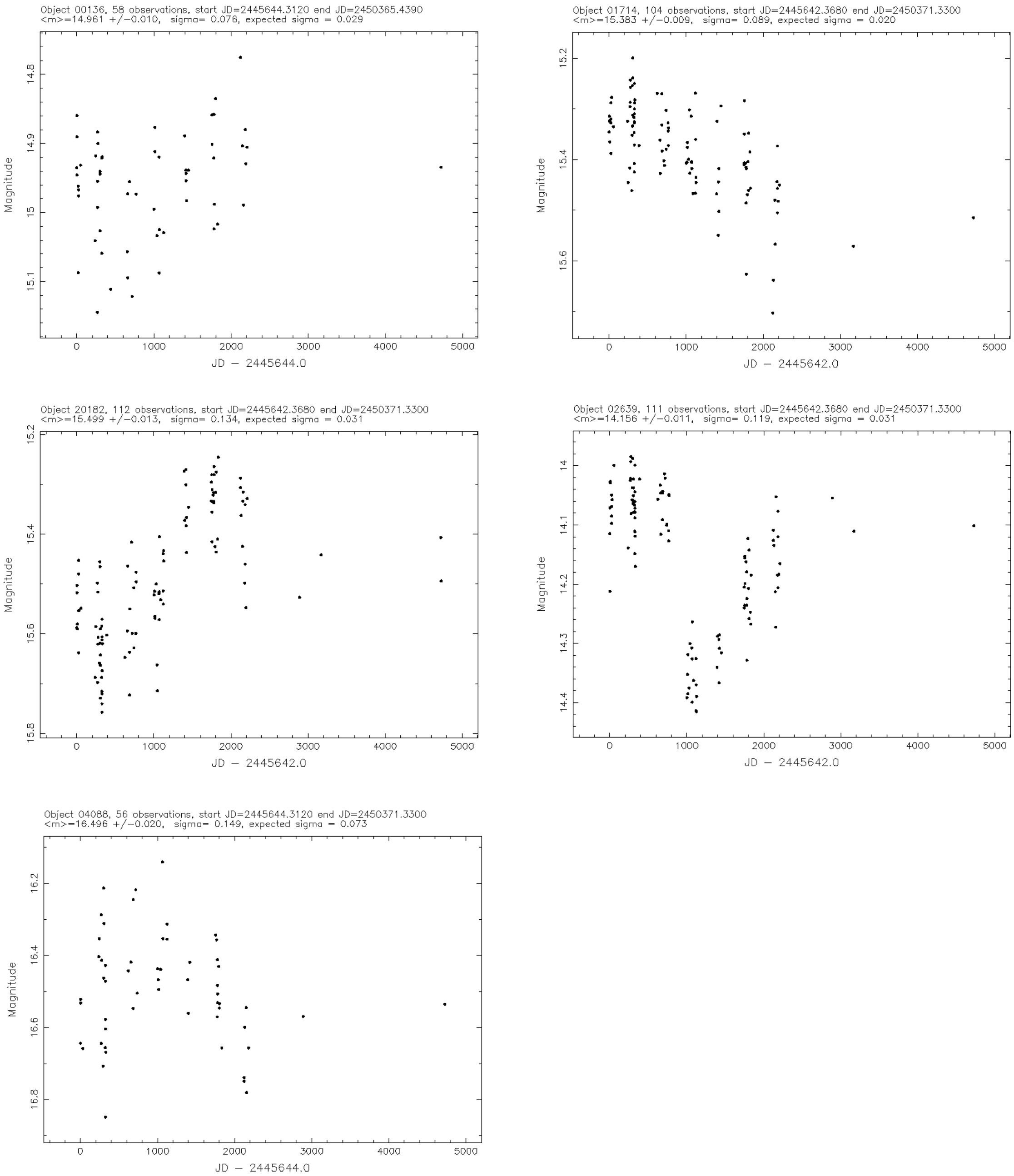


Fig. 5.6: Lightcurves of newly discovered long period variable stars

Irregular variable stars show variability in their magnitudes without any discernible periodicity. If periodicity is seen, it is only temporary. Some stars show such a slow variation nothing can be said about them, apart from the observation of changing brightness. These stars typically tend to be very red, with J-K colour indices over 1.0.
GCVS lists 3 classifications for this category: L, LB and LC.

L This classification is used for general slow irregular variables. This classification may be ascribed to poorly studied stars who show slow irregular variation. No condition on spectral class is mentioned.

LB This classification is used for red slow irregular variables. To count as LB, stars have to be of late spectral type (K,M,C or S) and they need to be giants.

LC These are red super giants with amplitude of variation being about 1^m.

5.1.4 Cataclysmic variables

Cataclysmic variables are those where a sudden brightness surplus appears in a star's lightcurve. Novae were the first variable stars ever discovered independently by various cultures throughout the history. Nova is Latin for new (same as the word 'Nava' in Marathi!), and it was a designation used for stars which were absent in the sky until suddenly they became so bright that it felt like a new star had popped up.
There are four main reasons why we see a sudden spark in a star's brightness. The most cataclysmic one of them all is a supernova explosion wherein a large star implodes, causing intense nuclear reactions and expulsions of gas, dust and metals at near-c velocities. The second case is a nova, now understood as an explosion on the surface of a white dwarf caused by collision of matter leaked from a thicker giant star's atmosphere into the Roche lobe of the white dwarf, where both these stars are in a close binary system. The third case is a dwarf-nova, in which the outburst happens due to the accretion disk around a white dwarf. The fourth is the most benign, it is caused by flares (solar surface activity) erupting on the said star, however they are short lived and their light gets integrated into one frame in long exposures such as ours.

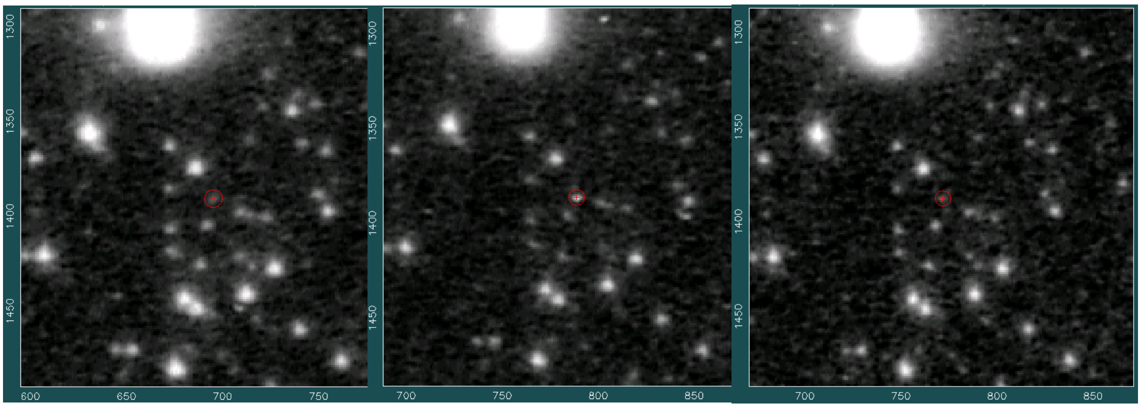


Fig. 5.7: The star that only appeared on one frame

There were many candidates for eruptive variables in our data but they often turned out to be errors on closer inspection. However, one case stood out. As you can see below, a star is only visible on one of the frames. It can't be confirmed as a nova from these observations only, so the next step will be to find more data near these coordinates.

5.1.5 Rotating stars

Stars rotate about an internal axis just like the planets (however since they are not solids, they rotate at different rates at different latitudes). When a star has dark spots or flares on its surface, the rotation leads to (almost) periodic variations in lightcurve. These are generally low in amplitude and hence difficult to detect with photographic plates. Despite that, we did find some variables of this kind in our data. These were already known, I just happened to check and find them.

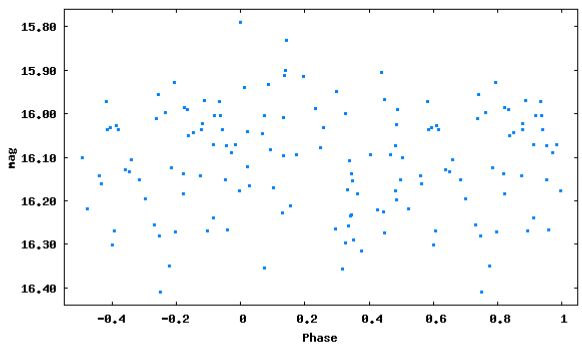


Fig. 5.8: Phase diagram of a BY Draconis variable found in our data (already known)

5.2 The applications and significance

Variable stars is now a very broad central topic of study encompassing a wide variety of astrophysics.

Pulsating stars are the sole reason we have been able to calculate distances to far away stars and galaxies due to the period-luminosity relation which certain types of pulsators follow combined with their apparent brightness. We know the dimensions of our galaxy because of Cepheid variables.

Eclipsing stars provide rich information about the stellar populations and dynamics. Radial velocity data of binary stars allows us to calculate masses of these stars from orbital parameters. It is also possible now with data spanning over decades to witness the binary system evolve in real time as the nuclear fuels of constituents run dry. We are now also detecting and studying exoplanets with photometry. The study of nearby habitable exoplanets will be very useful if we were to ever spread to other solar systems.

Variable stars also teach us about stellar evolution, as many kinds of pulsating variabilities are only observed in specific parts of the HR diagram.

In general time domain data allows us to see what's going on, even if we are only observing a white point in the sky. The discoveries we made in this project will have a direct (albeit minor) impact on astronomy. It is not only the discoveries that count, stars which were already known to be variables still need to be studied. Through our study we came across many cases where the classification was incorrect or the period was off, something we can now help with our data.

As was always the case, there remain variables we don't yet understand. Interesting physics may come out of these stars in the future.

5.3 Conclusion

We successfully discovered variable stars out of old archived data. These results are promising, they prove that old data is still useful. Such projects should help fuel the ongoing digitisation process. These discoveries are worth publishing, and we will be publishing these later this year. Detection of new variable stars is just one of the possibilities with the archival data. Once digitised completely, researchers will be able to try

new ideas.
We also reviewed the possibilites of machine learning projects. There is a good scope for programmers to write interesting software.
This is an old field of study, now evolving into many different individual branches of study. I had the opprtunity to study a vast area in astronomy and astrophysics.

APPENDIX A

GCVS variability types

(The following is a shortened list of all variability types as written in the General Catalogue of Variable Stars. The list shortly summarises all the variability types and is given for the reader's reference. The list is available online at (hyperlink): Samus and Durlevich (2016))

An improved system of variability classification is used in the fourth edition of the GCVS, based on recent developments in classification principles and taking into account the suggestions of a number of specialists. Variability types are grouped according to the major astrophysical reasons for variability, viz.,

1. **eruptive (BE, FU, GCAS, I, IA, IB, IN, INA, INB, INT, IT, IN(YY), IS, ISA, ISB, RCB, RS, SDOR, UV, UVN, WR),**
2. **pulsating (ACYG, BCEP, BCEPS, BLBOO, CEP, CEP(B), CW, CWA, CWB, DCEP, DCEPS, DSCT, DSCTC, GDOR, L, LB, LC, LPB, M, PVTEL, RPHS, RR, RR(B), RRAB, RRC, RV, RVA, RVB, SR, SRA, SRB, SRC, SRD, SRS, SXPHE, ZZ, ZZA, ZZB, ZZO),**
3. **rotating (ACV, ACVO, BY, ELL, FKCOM, PSR, R, SXARI),**
4. **cataclysmic (explosive and novalike) variables (N, NA, NB, NC, NL, NR, SN, SNI, SNII, UG, UGSS, UGSU, UGZ, ZAND),**
5. **eclipsing binary systems (E, EA, EB, EP, EW, GS, PN, RS, WD, WR, AR, D, DM, DS, DW, K, KE, KW, SD),**
6. **intense variable X-ray sources (AM, X, XB, XF, XI, XJ, XND, XNG, XP, XPR, XPRM, XM),**
7. **other symbols (BLLAC, CST, GAL, L:, QSO, S, *, +, :).**

All of these classes include objects of a dissimilar nature that belong to different types of light variability. On the other hand, an object may be variable because of almost all of the possible reasons or because of any combination of them. If a variable belongs to several types of variability, the types are joined in the data field by a "+" sign, e.g., E+UG, UV+BY.

Despite considerable success in understanding stellar variability pro- cesses, the classification adopted in the Catalogue is far from perfect. This is especially the case for explosive, symbiotic and novalike variables; X-ray sources; and peculiar objects.

The new variability types (ZZO, AM, R, BE, BLBOO, EP, SRS, LPB) have been added in the Name-Lists 67- 77 and in the GCVS vol.V.

1. **Eruptive Variable Stars**

BE It becomes more and more clear that, although the majority of Be stars are photometrically variable, not all of them could be properly called GCAS variables. Quite a number of them show small-scale variations not necessarily related to shell events; in some cases the variations are quasi-periodic. By now we are not able to present an elaborated system of classification for Be variables, but we adopt a decision that in the cases when a Be variable cannot be readily described as a GCAS star we give simply BE for the type of variability.

FU Orion variables of the FU Orionis type.

GCAS Eruptive irregular variables of the Gamma Cas type.

I Poorly studied irregular variables with unknown features of light variations and spectral types.

IA Poorly studied irregular variables of early (O-A) spectral type.

IB Poorly studied irregular variables of intermediate (F-G) to late (K-M) spectral type.

IN Orion variables. Irregular, eruptive variables connected with bright or dark diffuse nebulae or observed in the regions of these nebulae. This type may be divided into the following subtypes:

INA Orion variables of early spectral types (B-A or Ae). They are often characterized by occasional abrupt Algol-like fadings (T Ori);

INB Orion variables of intermediate and late spectral types, F-M or Fe-Me (BH Cep, AH Ori);

INT,IT Orion variables of the T Tauri type;

IN(YY) Some Orion variables (YY Ori) show the presence of absorption components on the redward sides of emission lines, indicating the infall of matter toward the stars' surfaces. In such cases, the symbol for the type may be accompanied by the symbol "YY".

IS Rapid irregular variables having no apparent connection with diffuse nebulae and showing light changes of about 0.5 - 1.0 mag within several hours or days.

ISA Rapid irregular variables of the early spectral types, B-A or Ae;

ISB Rapid irregular variables of the intermediate and late spectral types, F-M and Fe-Me.

RCB Variables of the R Coronae Borealis type.

RS Eruptive variables of the RS Canum Venaticorum type.

SDOR Variables of the S Doradus type. These are eruptive, high-luminosity Bpec-

Fpec stars showing irregular (sometimes cyclic) light changes with amplitudes in the range 1-7 mag in V.

UV Eruptive variables of the UV Ceti type, these are K Ve-M Ve stars sometimes displaying flare activity with amplitudes from several tenths of a magnitude up to 6 mag in V.

UVN Flaring Orion variables of spectral types Ke-Me.

WR Eruptive Wolf-Rayet variables.

2. **Pulsating Variable Stars**

ACYG Variables of the Alpha Cygni type, which are nonradially pulsating supergiants of Bep-AepIa spectral types.

BCEP Variables of the Beta Cephei type (Beta Cep, Beta CMa), which are pulsating O8-B6 I-V stars with periods of light and radial-velocity variations in the range of 0.1 - 0.6 days and light amplitudes from 0.01 to 0.3 mag in V.

BCEPS A short-period group of Beta Cep variables.

BLBOO The so-called "anomalous Cepheids", i.e. stars with periods characteristic of comparatively long-period RRAB variables, but considerably brighter by luminosity (BL Boo = NGC 5466 V19).

CEP Cepheids. Radially pulsating, high luminosity (classes Ib-II) vari- ables with periods in the range of 1-135 days and amplitudes from several hundredths to 2 mag in V (in the B band, the amplitudes are greater).

CEP(B) Cepheids (TU Cas, V 367 Sct) displaying the presence of two or more simultaneously operating pulsation modes (usually the fundamental tone with the period P0 and the first overtone P1).

CW Variables of the W Virginis type. They may be separated into the following subtypes:

CWA W Vir variables with periods longer than 8 days (W Vir);

CWB W Vir variables with periods shorter than 8 days (BL Her).

DCEP These are the classical cepheids, or Delta Cep-type variables.

DCEPS These are Delta Cep variables having light amplitudes <0.5 mag in V (<0.7 mag in B) and almost symmetrical light curves (M-m approx. 0.4 - 0.5 periods); as a rule, their periods do not exceed 7 days.

DSCT Variables of the Delta Scuti type.

DSCTC Low amplitude group of Delta Sct variables (light amplitude <0.1 mag in V).

GDOR Gamma Doradus stars. Early type F dwarfs showing (multiple) periods from several tenths of a day to slightly in excess of one day.

L Slow irregular variables.

LB Slow irregular variables of late spectral types (K, M, C, S); as a rule, they are giants (CO Cyg).

LC Irregular variable supergiants of late spectral types having amplitudes of about 1 mag in V (TZ Cas).

LPB The comparatively long-period pulsating B stars (periods exceeding one day).

M Mira (Omicron) Ceti-type variables.

PVTEL Variables of the PV Telescopii type.

RPHS Very rapidly pulsating hot (subdwarf B) stars.

RR Variables of the RR Lyrae type, which are radially-pulsating giant A-F stars having amplitudes from 0.2 to 2 mag in V.

RR(B) RR Lyrae variables showing two simultaneously operating pulsation modes, the fundamental tone with the period P0 and the first overtone, P1 (AQ Leo). The ratio P1/P0 is approximately 0.745;

RRAB RR Lyrae variables with asymmetric light curves (steep ascending branches), periods from 0.3 to 1.2 days, and amplitudes from 0.5 to 2 mag in V;

RRC RR Lyrae variables with nearly symmetric, sometimes sinusoidal, light curves, periods from 0.2 to 0.5 days, and amplitudes not greater than 0.8 mag in V (SX UMa).

RV Variables of the RV Tauri type. These are radially pulsating supergiants having spectral types F-G at maximum light and K-M at minimum. Two subtypes, RVA and RVB, are recognized:

RVA RV Tauri variables that do not vary in mean magnitude (AC Her);

RVB RV Tauri variables that periodically (with periods from 600 to 1500 days and amplitudes up to 2 mag in V) vary in mean magnitude (DF Cyg, RV Tau).

SR Semiregular variables, which are giants or supergiants of intermediate and late spectral types showing noticeable periodicity in their light changes, accompanied or sometimes interrupted by various irregularities.

SRA Semiregular late-type (M, C, S or Me, Ce, Se) giants displaying persistent periodicity and usually small (<2.5 mag in V) light amplitudes (Z Aqr).

SRB Semiregular late-type (M, C, S or Me, Ce, Se) giants with poorly defined periodicity (mean cycles in the range of 20 to 2300 days) or with alternating intervals of periodic and slow irregular changes, and even with light constancy intervals (RR CrB, AF Cyg).

SRC Semiregular late-type (M, C, S or Me, Ce, Se) supergiants (Mu Cep) with amplitudes of about 1 mag and periods of light variation from 30 days to several thousand days;

SRD Semiregular variable giants and supergiants of F, G, or K spectral types, sometimes with emission lines in their spectra.

SRS Semiregular pulsating red giants with short period (several days to a month), probably high-overtone pulsators. Prototype: AU Ari.

SXPHE Phenomenologically, these resemble DSCT (Delta Sct) variables and are pulsating subdwarfs of the spherical component, or old disk galactic population, with spectral types in the range A2-F5.

ZZ ZZ Ceti variables. These are nonradially pulsating white dwarfs that change their brightnesses with periods from 30 s to 25 min and amplitudes from 0.001 to 0.2 mag in V.

ZZA ZZ Cet-type variables of DA spectral type (ZZ Cet) having only hydrogen absorption lines in their spectra;

ZZB ZZ Cet-type variables of DB spectral type having only helium absorption lines in their spectra.

ZZO ZZ Cet type variables of the DO spectral type showing HeII and and CIV absorpion lines in their spectra.

3. Rotating Variable Stars

ACV Alpha2 Canum Venaticorum variables. These are main-sequence stars with spectral types B8p-A7p and displaying strong magnetic fields.

ACVO Rapidly oscillating Alpha2 CVn variables. These are nonradially pulsating, rotating magnetic variables of Ap spectral type (DO Eri).

BY BY Draconis-type variables, which are emission-line dwarfs of dKe-dMe spectral type showing quasiperiodic light changes with periods from a fraction of a day to 120 days and amplitudes from several hundredths to 0.5 mag in V.

ELL Rotating ellipsoidal variables (b Per, Alpha Vir). These are close binary systems with ellipsoidal components, which change combined brightnesses with periods equal to those of orbital motion because of changes in emitting areas toward an observer, but showing no eclipses.

FKCOM FK Comae Berenices-type variables. These are rapidly rotating giants with nonuniform surface brightnesses, which have G-K spectral types with broad H and K Ca II emission and sometimes Halpha.

PSR Optically variable pulsars (CM Tau), which are rapidly rotating neutron stars with strong magnetic fields, radiating in the radio, optical, and X-ray regions.

R Close binary systems characterized by the presence of strong reflection (re-radiation) of the light of the hot star illuminating the surface of the cooler companion.

SXARI SX Arietis-type variables. These are main-sequence B0p-B9p stars with variable-intensity He I and Si III lines and magnetic fields.

4. **Cataclysmic (Explosive and Novalike) Variables**

N Novae. Close binary systems with orbital periods from 0.05 to 230 days. One of the components of these systems is a hot dwarf star that suddenly, during a time interval from one to several dozen or several hundred days, increases its brightness by 7-19 mag in V, then returns gradually to its former brightness over several months, years, or decades. According to the features of their light variations, novae are subdivided into fast (NA), slow (NB), very slow (NC), and recurrent (NR) categories.

NA Fast novae displaying rapid light increases and then, having achieved maximum light, fading by 3 mag in 100 or fewer days (GK Per);

NB Slow novae that fade after maximum light by 3 mag in >= 150 days (RR Pic);

NC Novae with a very slow development and remaining at maximum light for more than a decade, then fading very slowly;

NL Novalike variables, which are insufficiently studied objects resembling novae by the characteristics of their light changes or by spectral features.

NR Recurrent novae, which differ from typical novae by the fact that two or more outbursts (instead of a single one) separated by 10-80 years have been observed (T CrB).

SN Supernovae (B Cas, CM Tau). Stars that increase, as a result of an outburst, their brightnesses by 20 mag and more, then fade slowly;

SNI Type I supernovae;

SNII Type II supernovae.

UG U Geminorum-type variables, quite often called dwarf novae. They are close binary systems consisting of a dwarf or subgiant K-M star that fills the volume of its inner Roche lobe and a white dwarf surrounded by an accretion disk.

UGSS SS Cygni-type variables (SS Cyg, U Gem). They increase in brightness by 2-6 mag in V in 1-2 days and in several subsequent days return to their original brightnesses;

UGSU SU Ursae Majoris-type variables. These are characterized by the presence of two types of outbursts called "normal" and "supermaxima";

UGZ Z Camelopardalis-type stars. These also show cyclic outbursts, differing from UGSS variables by the fact that sometimes after an outburst they do not return to the original brightness, but during several cycles retain a magnitude between maximum and minimum.

ZAND Symbiotic variables of the Z Andromedae type. They are close binaries consisting of a hot star, a star of late type, and an extended envelope excited by the hot star's radiation.

5. **Close Binary Eclipsing Systems**

We adopt a triple system of classifying eclipsing binary systems: according to the shape of the combined light curve, as well as to physical and evolutionary characteristics of their components. The classification based on light curves is simple, traditional, and suits the observers; the second and third classification methods take into account positions of the binary-system components in the (MV ,B-V) diagram and the degree of inner Roche lobe filling. Estimates are made by applying the simple criteria proposed by Svechnikov and Istomin (1979). The symbols for the types of eclipsing binary systems that we use are given below.

a) Classification based on the shape of the light curve

E Eclipsing binary systems. These are binary systems with orbital planes so close to the observer's line of sight (the inclination i of the orbital plane to the plane orthogonal to the line of sight is close to 90 deg) that the components periodically eclipse each other.

EA Algol (Beta Persei)-type eclipsing systems. Binaries with spherical or slightly ellipsoidal components.

EB Beta Lyrae-type eclipsing systems. These are eclipsing systems having ellipsoidal components and light curves for which it is impossible to specify the exact times of onset and end of eclipses because of a continuous change of a system's apparent combined brightness between eclipses; secondary minimum is observed in all cases, its depth usually being considerably smaller than that of the primary minimum; periods are mainly longer than 1 day.

EP Stars showing eclipses by their planets. Prototype: V0376 Peg.

EW W Ursae Majoris-type eclipsing variables. These are eclipsers with periods shorter than 1 days, consisting of ellipsoidal components almost in contact and having light curves for which it is impossible to specify the exact times of onset and end of eclipses.

b) Classification according to the components' physical characteristics

GS Systems with one or both giant and supergiant components; one of the components may be a main sequence star.

PN Systems having, among their components, nuclei of planetary nebulae (UU Sge).

RS RS Canum Venaticorum-type systems. A significant property of these systems is the presence in their spectra of strong Ca II H and K emission lines of variable intensity, indicating increased chromospheric activity of the solar type.

WD Systems with white-dwarf components.

WR Systems having Wolf-Rayet stars among their components (V 444 Cyg).

c) Classification based on the degree of filling of inner Roche lobes

AR Detached systems of the AR Lacertae type. Both components are subgiants not filling their inner equipotential surfaces.

D Detached systems, with components not filling their inner Roche lobes.

DM Detached main-sequence systems. Both components are main-sequence stars and do not fill their inner Roche lobes.

DS Detached systems with a subgiant. The subgiant also does not fill its inner critical surface.

DW Systems similar to W UMa systems in physical properties (KW, see below), but not in contact.

K Contact systems, both components filling their inner critical surfaces.

KE Contact systems of early (O-A) spectral type, both components being close in size to their inner critical surfaces.

KW Contact systems of the W UMa type, with ellipsoidal components of F0-K spectral type.

SD Semidetached systems in which the surface of the less massive com- ponent is close to its inner Roche lobe.

6. **Optically Variable Close Binary Sources of Strong, Variable X-ray Radiation (X-ray Sources)**

AM AM Her type variables; close binary systems consisting of a dK-dM type dwarf and of a compact object with strong magnetic field, characterized by variable linear and circular polarization of light.

X Close binary systems that are sources of strong, variable X-ray emission and which do not belong to or are not yet attributed to any of the above types of variable stars. These objects may be subdivided into the following types:

XB X-ray bursters. Close binary systems showing X-ray and optical bursts, their duration being from several seconds to ten minutes, with amplitudes of about 0.1 mag in V (V 801 Ara, V 926 Sco);

XF Fluctuating X-ray systems showing rapid variations of X-ray (Cygnus X-1 = V1357 Cyg) and optical (V821 Ara) radiation on time scales of dozens of milliseconds;

XI X-ray irregulars. Close binary systems consisting of a hot compact object surrounded by an accretion disk and a dA - dM-type dwarf;

XJ X-ray binaries characterized by the presence of relativistic jets evident at X-ray and radio wavelengths, as well as in the optical spectrum in the form of emission components showing periodic displacements with relativistic velocities (V1343 Aql);

XND X-ray, novalike (transient) systems containing, along with a hot compact object, a dwarf or subgiant of G-M spectral type;

XNG X-ray, novalike (transient) systems with an early-type supergiant or giant primary component and a hot compact object as a companion. Following the main component's outburst, the material ejected by it falls onto the compact object and causes, with a significant delay, the appearance of X rays;

XP X-ray pulsar systems. The primary component is usually an ellipsoidal early-type supergiant;

XPR X-ray pulsar systems featuring the presence of the reflection effect. They consist of a dB-dF-type primary and an X-ray pulsar, which may also be an optical pulsar;

XPRM, X-ray systems consisting of a late-type dwarf (dK-dM) and a pulsar;

XM with a strong magnetic field. Matter accretion on the compact object's magnetic poles is accompanied by the appearance of variable linear and circular polarization; hence, these systems are sometimes known as "polars".

7. **Other Symbols**

In addition to the variable-star types described above, certain other symbols that need to be explained will be found in the Type data field:

BLLAC Extragalactic BL Lacertae-type objects. These are compact quasistellar objects showing almost continuous spectra with weak emission and absorption lines and relatively rapid irregular light changes with amplitudes up to 3 mag in V or more.

CST Nonvariable stars, formerly suspected to be variable and hastily designated. Further observations have not confirmed their variability.

GAL Optically variable quasistellar extragalactic objects (active galactic nuclei [AGNs]) considered to be variable stars by mistake.

L: Unstudied variable stars with slow light changes.

QSO Optically variable quasistellar extragalactic sources (quasars) that earlier were erroneously considered to be variable stars.

S Unstudied variable stars with rapid light changes.

* Unique variable stars outside the range of the classifications de- scribed above. These probably represent either short stages of transition from one variability type to another or the earliest and latest evolutionary stages of these types, or they are insufficiently studied members of future new types of variables.

+ If a variable star belongs to several types of light variability simultaneously, the types are joined in the Type field by a "+" sign (e.g., E+UG, UV+BY).

: Uncertainty flag on Type of Variability

REFERENCES

1. **Antipin, S.**, **I. Becker**, **A. A. Belinski**, **D. M. Kolesnikova**, **K. Pichara**, **N. N.Samus**, **K. V. Sokolovsky**, **A. V. Zharova**, and **A. M. Zubareva** (2018). New variable stars from the photographic archive:semi-automated discoveries, attempts of automatic classification, and the new field 104 her. *Research in Astronomy and Astrophysics*. URL `https://arxiv.org/pdf/1802.02575.pdf`.

2. **Hoffmeister, C.**, **G. Richter**, and **W. Wenzel** (1985). *Variable Stars*. Springer-Verlag.

3. **Lafler, J.** and **T. Kinman** (1965). The calculation of rr lyr periods by electronic computer. *Astrophysical Journal Supplement*, **11**, `doi:10.1086/190116`.

4. **Percy, J. R.** (2007). *Understanding Variable Stars*. Cambridge University Press.

5. **Samus, N.** and **O. V. Durlevich** (2016). Gcvs variability types. URL `http://www.sai.msu.su/gcvs/gcvs/iii/vartype.txt`.

6. **Samus, N.**, **E. Kazarovets**, **O. Durlevich**, **N. Kireeva**, and **E. Pastukhova** (2017). The general catalogue of variable stars. URL `www.sai.msu.su/gcvs/gcvs`.

7. **Samus, N. N.**, **L. Sat**, **S. Vereshchagin**, and **A. Zharova** (). Moscow astronomical plate archives: Contents, digitisation, current and possible applications. URL `http://www.sai.msu.su/gcvs/digit/preprint.htm`.

8. **Sokolovsky, K. V.** and **A. A. Lebedev** (2018). Vast: A variability search toolkit. `doi:10.1016/j.ascom.2017.12.001`.

9. **Warner, B. D.** (2006). *A practical guide to lightcurve photometry and analysis*. Springer.