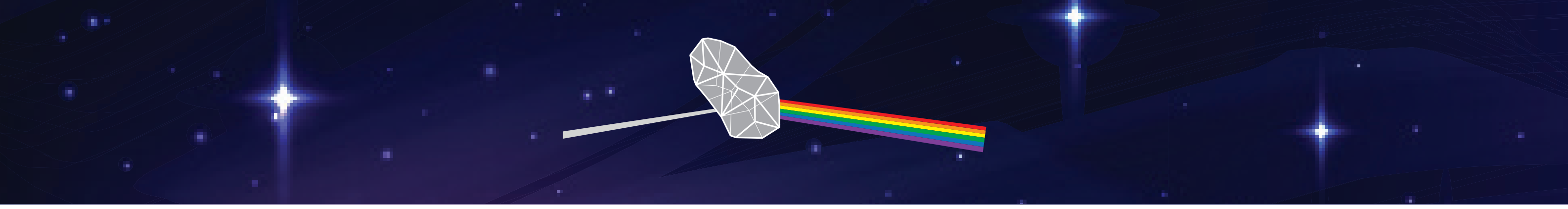


DATA MINING

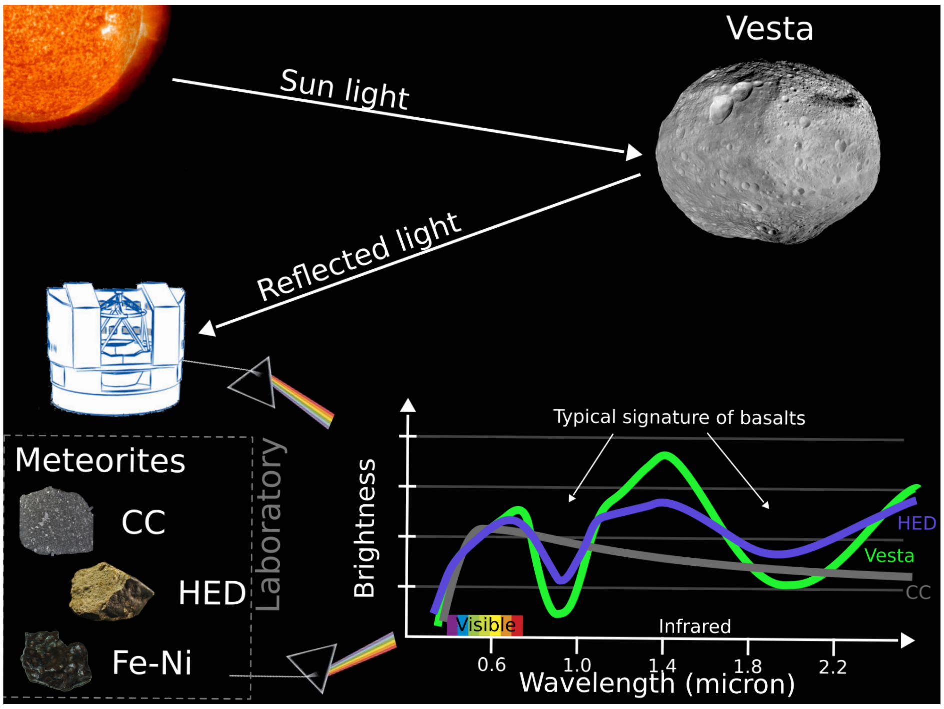
Benoit Carry
Observatoire de la Cote d'Azur

Near-Earth Asteroids (NEA) represent the vast majority of NEO: as of today, the astronomical community have discovered more than 28,000 NEOs. Owing to their orbits, they end up impacting planets or falling into the Sun. The potential threat represented by NEOs motivates discovery programmes (How many are they? Where are they?) and characterisation programmes (How big are they? What are they made of?), such as our NEOROCKS project.

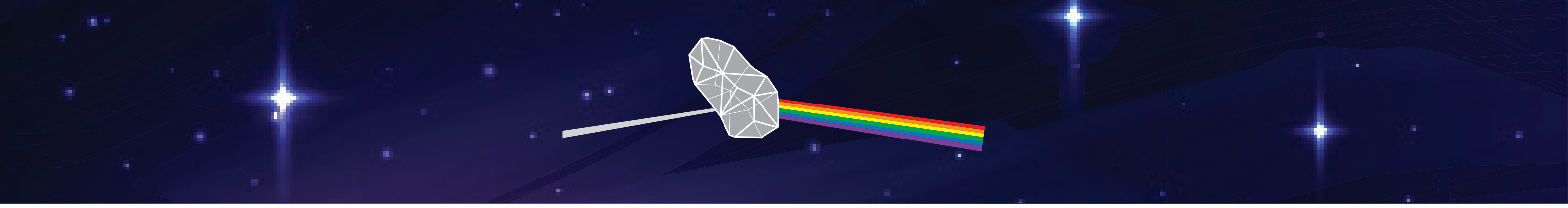
While the probability of impact with the Earth is very low, it is a natural hazard that we can predict. In the very unlikely case of an NEO heading towards us (Don't Look Up!), the mitigation strategy will depend on the characteristics of the asteroid, and in particular its diameter and its composition.

The level of threat is dictated by the mass of the asteroid, which is extremely complex to measure. We thus rely on measurement of the diameter and the density to estimate the mass. While density is also extremely complex to measure, it is closely linked with the composition of the asteroid, which can be determined from how it reflects the light of the Sun.

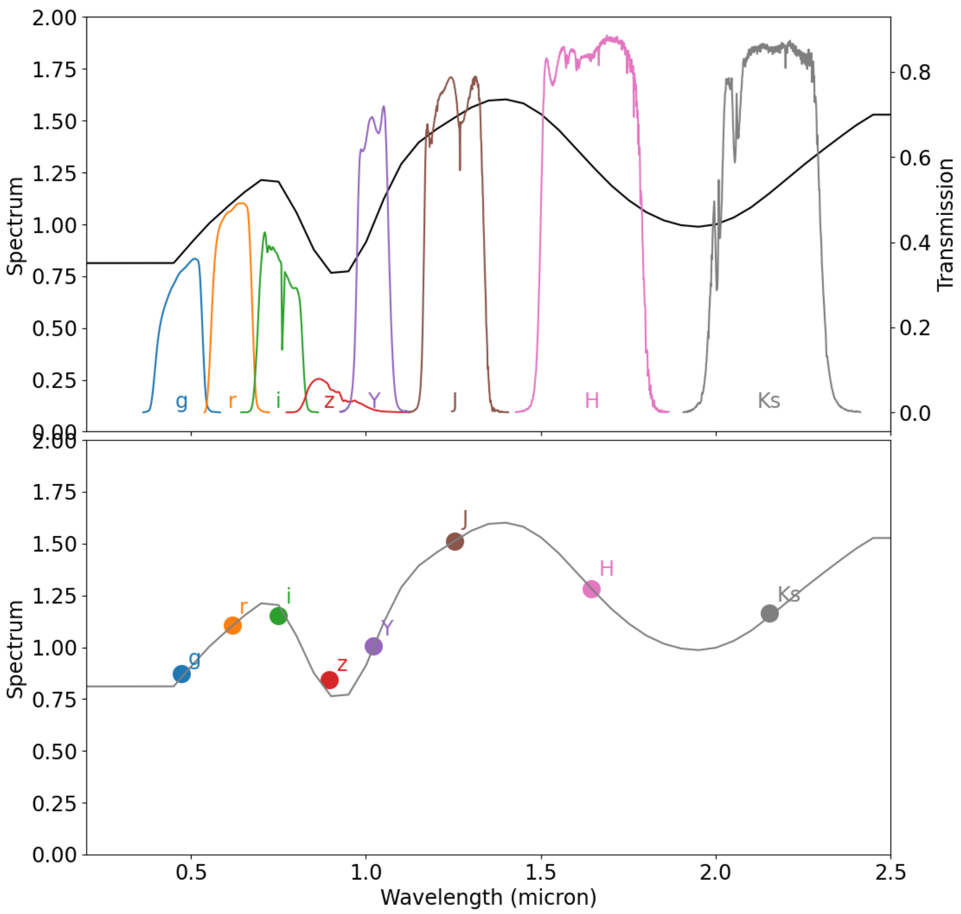
The composition of an asteroid is determined by comparing the light reflected by the surface with that of meteorites studied in the laboratory. In particular, we compare how the light spread across wavelengths (in plain English, we would say across colours). This technique is called spectroscopy. It is the most precise method to study the composition of an asteroid.



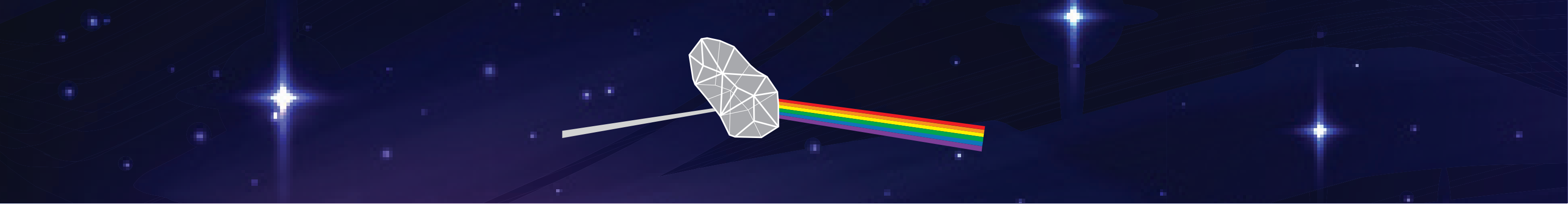
Concept of the analysis of asteroid composition



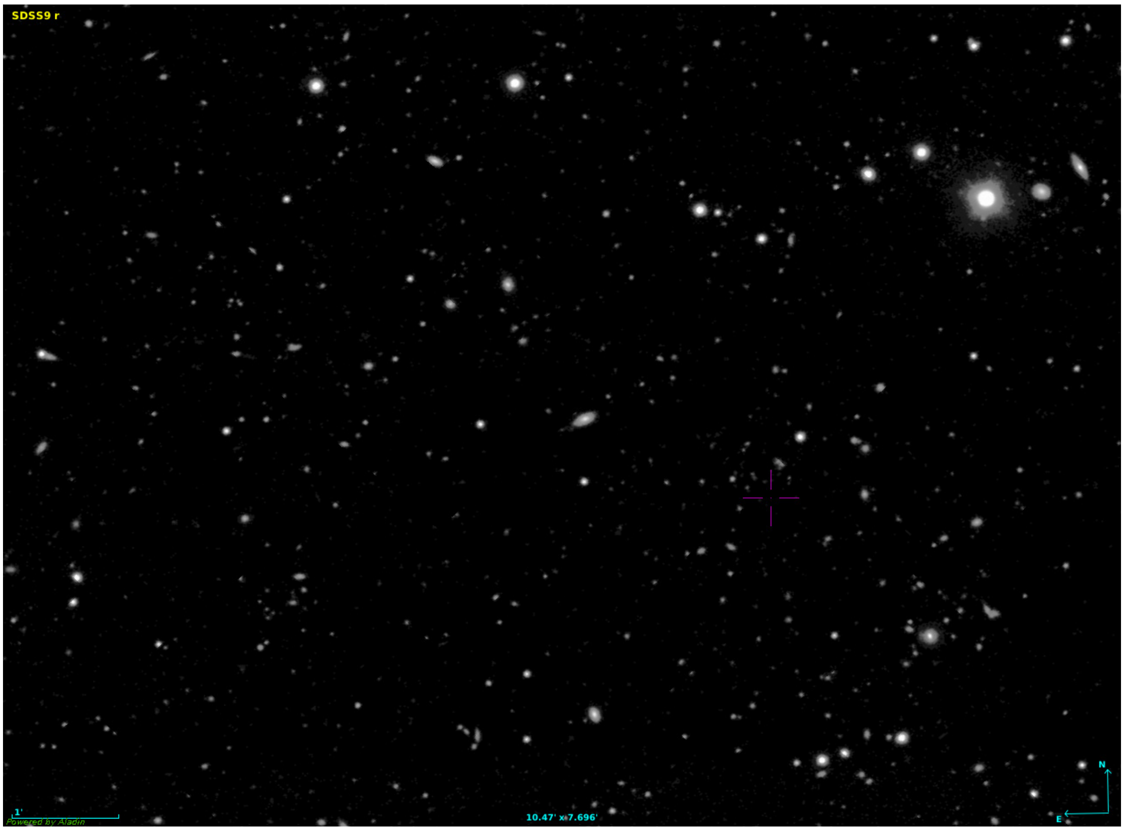
However, acquiring these spectra is time consuming. It requires nights, and nights, and nights of observations to collect a few tens or hundreds of spectra. These spectra are extremely important as they provide detailed information on the composition of an object. So, how can we increase the number of characterization of compositions?
In NEOROCKS, we decided to complement these reference spectroscopic observations with another approach: colours.
We call colours the comparison of the amount of light between two wavelengths. It is similar to the usual, daily, conception: if an asteroid reflects more light in the wavelength corresponding to the "red" than in the wavelength corresponding to the "blue", we will say that the asteroid is "red". Beyond a simple red/green/blue perception, the colours really describe the spectrum of the NEOs. They are simply a low-resolution version of the spectrum (an analogy could be a pixelized version of a portrait that still allows to recognise the person).
In astronomy, we use many different filters to take images and these select only some wavelengths. In the figure here, the g and r filters correspond to what the eye recognize as blue and red. However, we also use filter in the near-infrared, represented by the z, Y, J, H, Ks in this example. Thus, we can obtain the colour by measuring the amount of light in two images, taken with two different filters. In NEOROCKs, we also measure colours of NEOs.



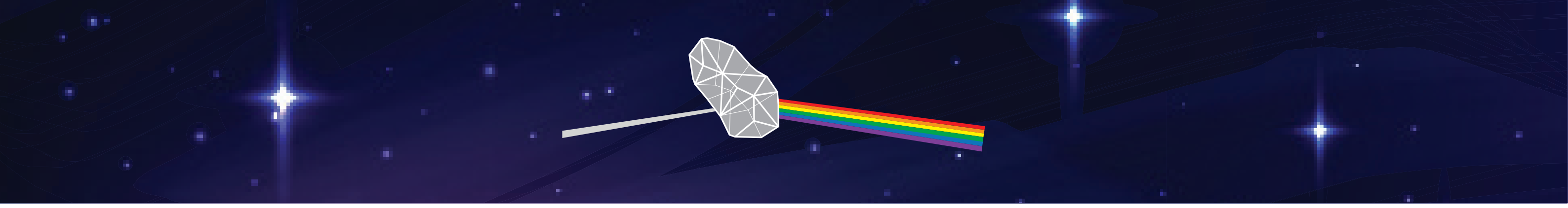
Filters select the light over some specific wavelength ranges. Combining several filters reproduce the spectrum, with less resolution



Why should we use a cruder version of the information? Why do we not acquire only spectra? The answer is that it is much faster to acquire two images in two filters than to build a complete spectrum. The filters collect the light over a broad range of wavelengths, while by essence spectroscopy spreads it. It is thus more efficient to acquire colours than spectra, but we lose spectral details. These colours can, therefore, be used to classify asteroids broadly into similar groups, the composition of which can be understood thanks to spectroscopy. By doing both spectroscopy and colours in NEOROCKS, we win on both sides: more characterised NEOs and more understanding!
We also decided to complement our own observations with another approach: mining of colours from astronomical archives. There are many different telescopes in the world. For many years, they have been observing every night and collecting data. Most of the time, the images acquired by these telescopes have other purposes than NEOs: some researchers may be studying galaxies, stars, nebulae, exoplanets… There is nevertheless a tremendous amount of images taken over many different regions of the sky every night, in many different filters.
In NEOROCKS, we are hunting for NEOs present in these images by mere chance, These could be NEOs orbiting the Sun, which are constantly moving, and usually appear starlike in a single image taken for other purposes. The complexity of this task is to be sure that it is definitely an NEO in the image. Moreover, in the vast majority of images, there will only be stars, or galaxies, but no NEOs. So, it becomes a bit like searching for a needle in a haystack.



Example of an image from the SDSS. From a single image, there is a priori no way to identify the source: stars, galaxies, NEOs?



This approach, thus, requires both an analysis of many images to find NEOs and the development of tools to ascertain their identification. In NEOROCKS, we have searched NEOs in images obtained by two telescopes that have imaged very large portions of the sky: the Sloan Digital Sky Survey (SDSS) from 1998 to 2008, and the SkyMapper Southern Survey (SMSS) since 2014. We have computed the predicted positions of all asteroids for each of the 1,400,000 images of the SDSS and the 200,000 images of the CMSS. If there was a source detected on the image at the predicted position of the asteroid, we considered it was a good candidate for being a NEO.

We then checked that these candidates were genuine NEOs. Indeed, the simple match on position can lead to many false association with stars. These checks are multiple, from the motion of the candidate over several frames, to its colours, to its proximity with known stars, among others.

As a result, we have managed to measure the colours of 1652 NEOs from the SDSS (five filters each), and of 669 NEOs (two to four filters) from the SkyMapper, without actually undertaking the observation ourselves! This is a clear advantage of this mining / hunting approach: we can extract many, many observations. We benefit from hundreds of nights of observations, which were initially taken for other purposes. The drawback is that have not yet been able to choose the NEO for which we would report the colours. This is the next step, to be carried out with our own, targeted observations. This is why we combine both in NEOROCKS, so stay in touch to find out more.