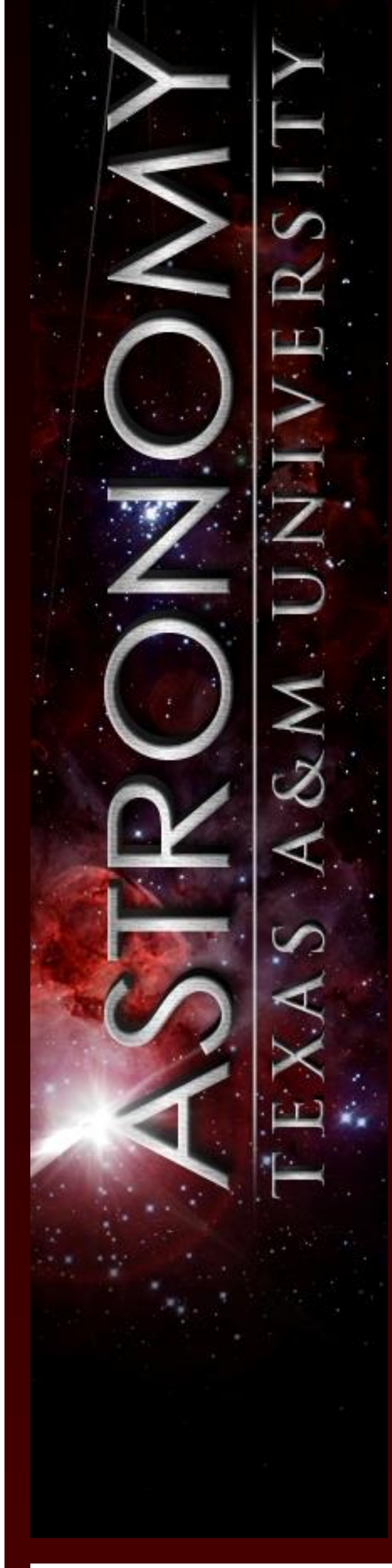
Classifying Variable Sources in SDSS Stripe 82

Christina W. Lindberg[1], James Long[2], Katelyn Stringer[3]

1. *Department of Astronomy & Department of Physics, University of Washington*
2. *Department of Statistics, Texas A&M*
3. *George P. and Cynthia W. Mitchell Institute for Fundamental Physics and Astronomy, Department of Physics and Astronomy, Texas A&M*



Abstract

As more astronomical data becomes available to the scientific community, less time can be equally dedicated to each object of scientific interest. SDSS (Sloan Digital Sky Survey) Stripe 82 is a well-documented and researched region of the sky that does not have all of its ~67,500 variable objects labeled yet. By collecting data from different catalogs, we are able to slowly fill in more classifications within the Stripe 82 catalog. The light curves of classified objects can then be used to determine features that most effectively separate the different types of variable objects. Upcoming large surveys such as LSST (the Large Synoptic Survey Telescope) could utilize these labeled variable sources as a training set for its own classifications.

Introduction

Stripe 82[1] is large central stripe covering the celestial equator that has been imaged several times by SDSS (Sloan Digital Sky Survey). The areas spans from a right ascension of 20 hours to 4 hours, with a declination of +/-1.25 degrees and contains just over a million detectable objects ranging from stars to quasars. By scanning the region repeatedly over multiple years in varying filters, it is possible to discern which objects in the region are variable in magnitude, an astronomical unit of brightness. Variable objects are useful because some are good distance estimators, since the period of their fluctuation and their luminosity can be correlated.

What is a variable object?

A variable object is any object in the night sky whose brightness fluctuates periodically. This means a variable object could be individual variable stars like RR Lyrae, quasars that spin rapidly while releasing energy, or multi-star systems like a binary star system where two stars orbit one another.

By taking observations of the same objects for a few days every half year, it is possible to see that the magnitudes varying over time as shown in Figure 1 (left). Once enough points have been gathered, we can estimate a period for the fluctuations. This period is very hard to see in the raw data, but once a period has been correctly estimated, all the data points can be plotted along the phase of the period, showing the light curve of the object, shown in Figure 1 (right). The object in Figure 1 is an RR Lyrae variable star, whose period is characteristically between 7 hours to 1 day. Figure 2 shows a binary star system. Notice how the folded light curves exhibit very different properties.

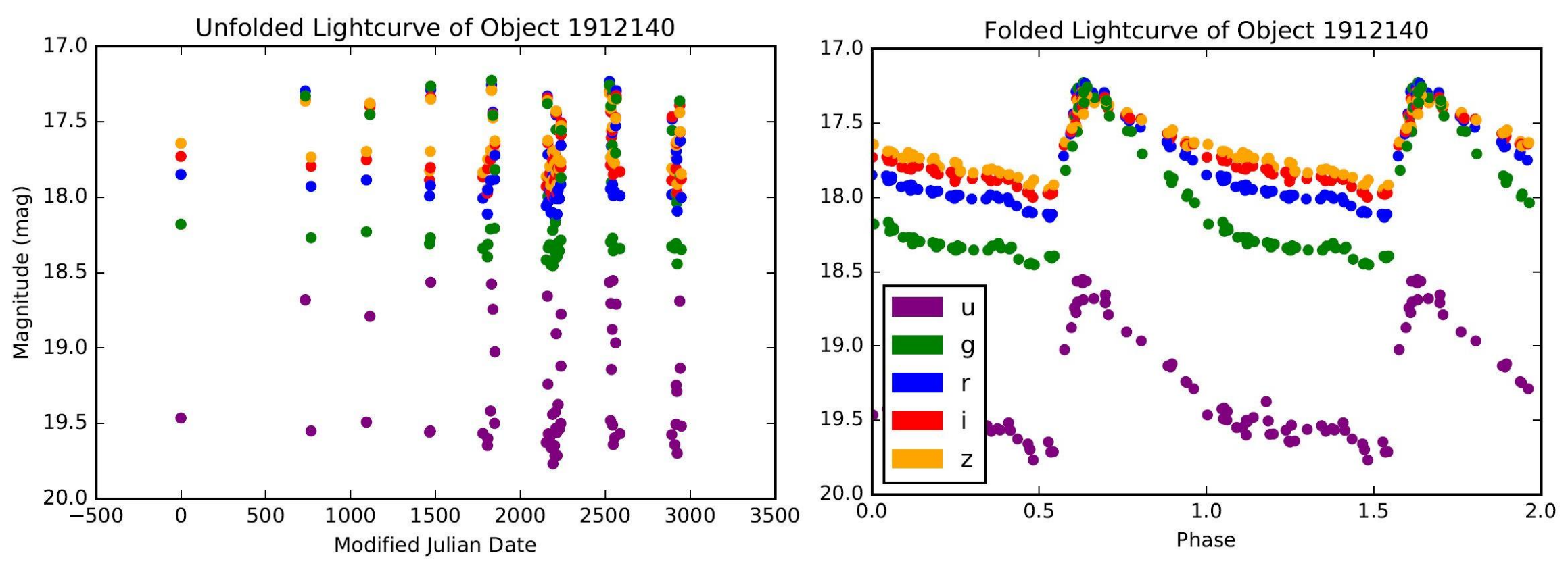


Figure 1. The raw, unfolded (left) light curve of an RR Lyrae star in different filters, and the folded RR Lyrae (right) plotted against double its phase.

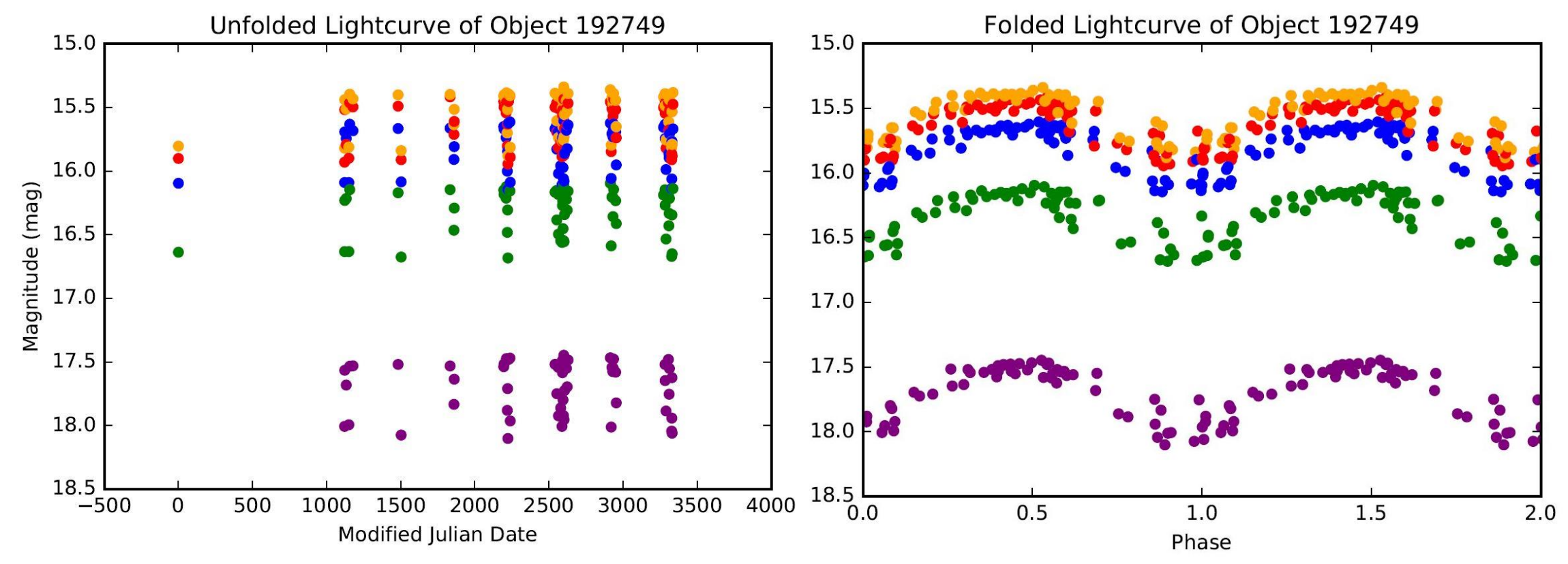


Figure 2. The raw, unfolded (left) light curve of an eclipsing binary star system in different filters, and the folded eclipsing binary star system (right) plotted against double its phase.

Why variable stars?

The pulsation period of a few of these types of variable stars correlate with their luminosity, meaning the longer the period of the star is, the brighter it gets. We can then calculate the distance to them using the distance modulus.

It is often difficult to measure the distances to astronomical structures accurately, so the more options there are for measuring (such as variable stars, supernovae, parallax, and redshift) the more we can confirm the distances we measure and the smaller our uncertainly becomes.

Thus, it is imperative to identify as many of these distance indicators as we can, so that we can map a better three dimensional structure of our surrounding galaxies. At the same time, it is also important to figure out what class each variable object is, so that we do not assume an object is a distance indicator, when it is not.

Merging catalogs

By referencing other research papers, we have been able to identify ~7,300 objects within the Stripe 82 catalog consisting of 67,466 variable objects, leaving around 60,200 objects unlabeled, or around 89%. If more objects can be classified, it will be possible to determine which features are most representative of each class of objects, and from that we will be able to classify other unlabeled objects that exhibit the same features.

We obtained a catalog of variable objects from the Catalina Sky Survey (CSS). While the survey's main objective is to discover and track near-earth objects, they have also constructed a catalog containing 47,000 periodic variable objects that have had their periods determined [2]. Out of the 47,000 objects in the Catalina catalog, there were about 820 objects that were contained within the region of Stripe 82.

While this seems like a small fraction, it is justified by the small area overlap between the two catalogs, as shown in Figure 3.

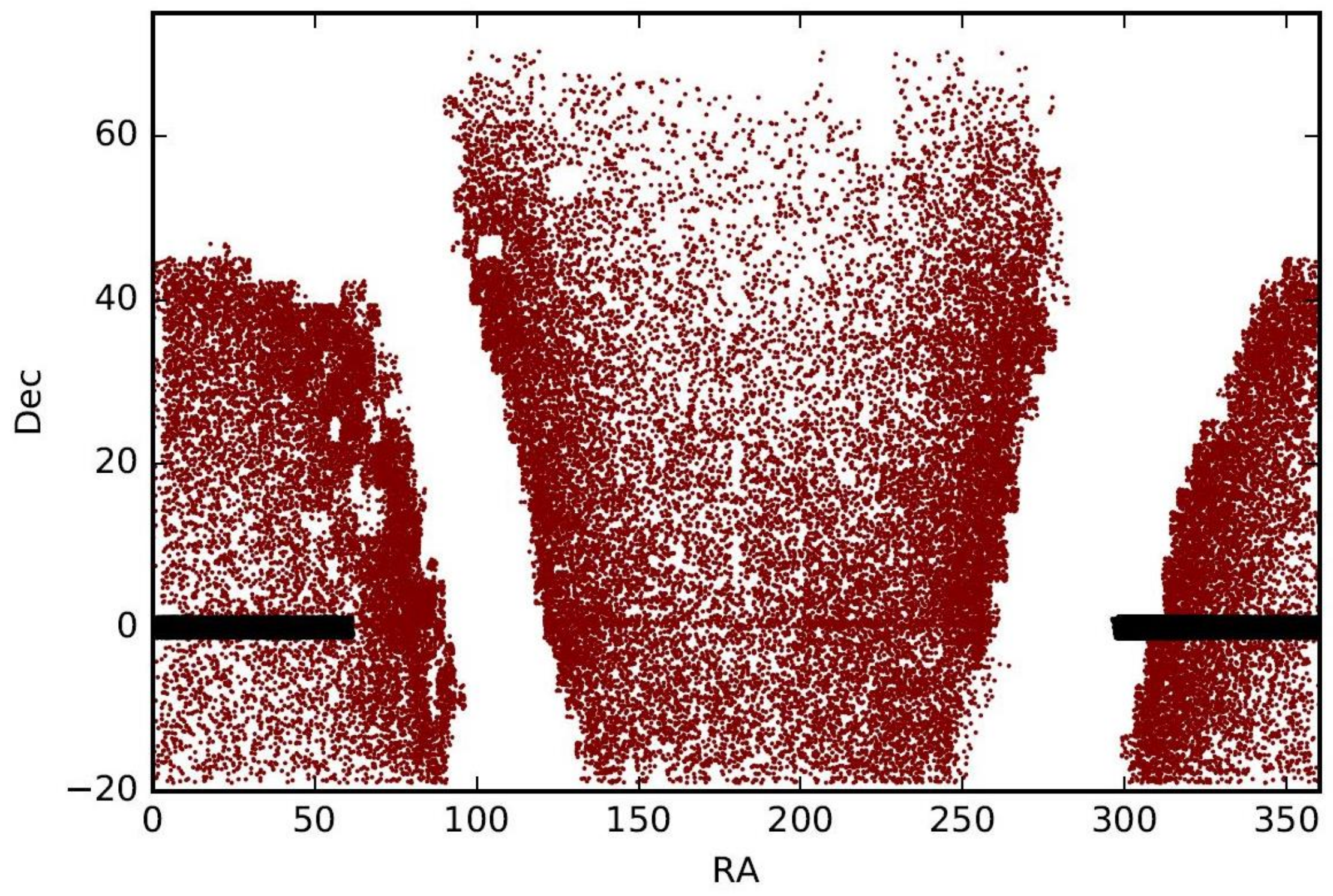


Figure 3. The overlap region between SDSS Stripe 82 (blue) and the Catalina catalog (red).

By converting the Catalina catalog right ascension units from hours-minutes-seconds (hms) to degrees and using the python package SkyCoord[3], we were able to match 300 objects from the Catalina catalog to objects in the Stripe 82 catalog, using a maximum distance separation of 3 arcseconds. The number of matches was reconfirmed with another matching algorithm that accounts for the difference in magnitude between the objects being matched. In order to compare the magnitudes, since Catalina was only observed in V, it was necessary to convert the Stripe 82 magnitudes from g and r to V.

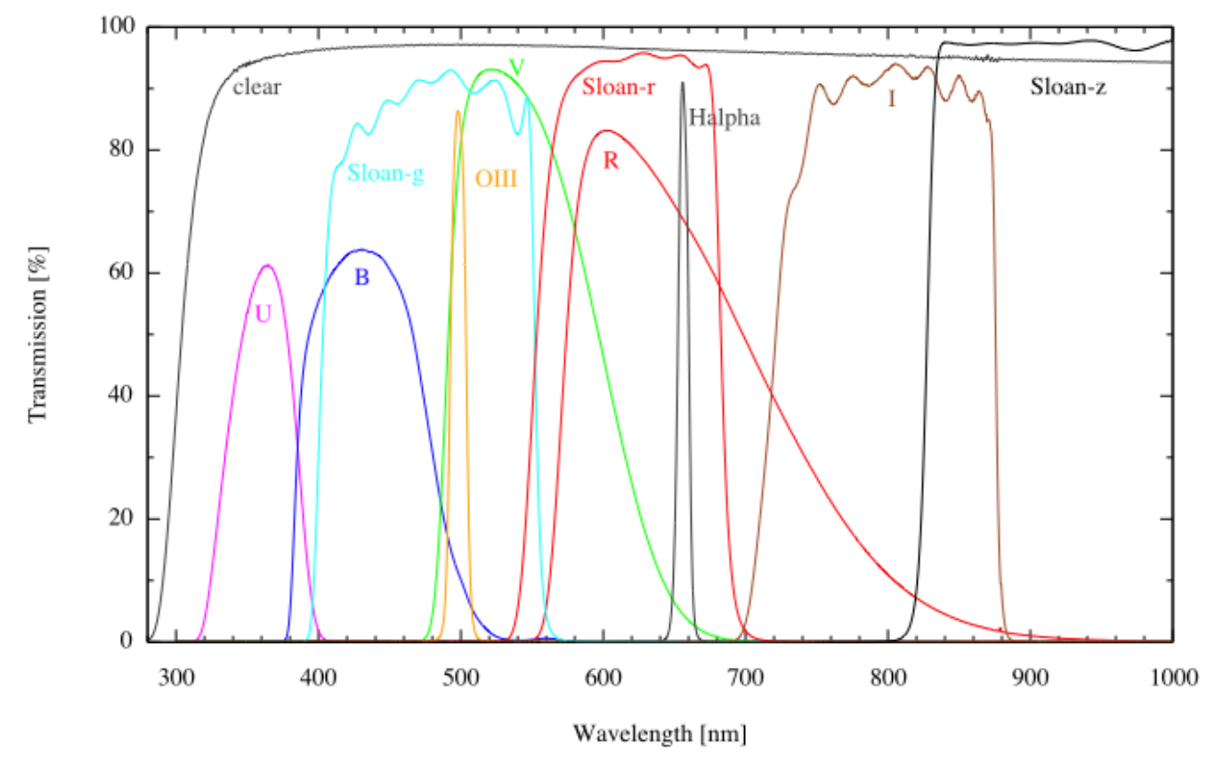


Figure 4. Some of the more common telescope filters. Credit: MONET Telescope

Feature extraction

Once enough of one class has been labeled, we can start analyzing the different features each class exhibits as a whole. If we can identify enough features, then it will be possible to start classifying unlabeled objects, and testing how our predictions match up.

With the Catalina catalog, we were able to add around 175 new eclipsing binary systems, labeled 'ew' and 'ea'. By plotting some of the various features, color, period, and variability index, of each class, one can start to outline rules that would divide each class into separate bins. In Figure 5, 'ew' and 'ea' objects are characterized well by their small period, while any other feature does not distinguish it much from any other class. RR Lyrae and quasars have been plotted alongside to show contrast.

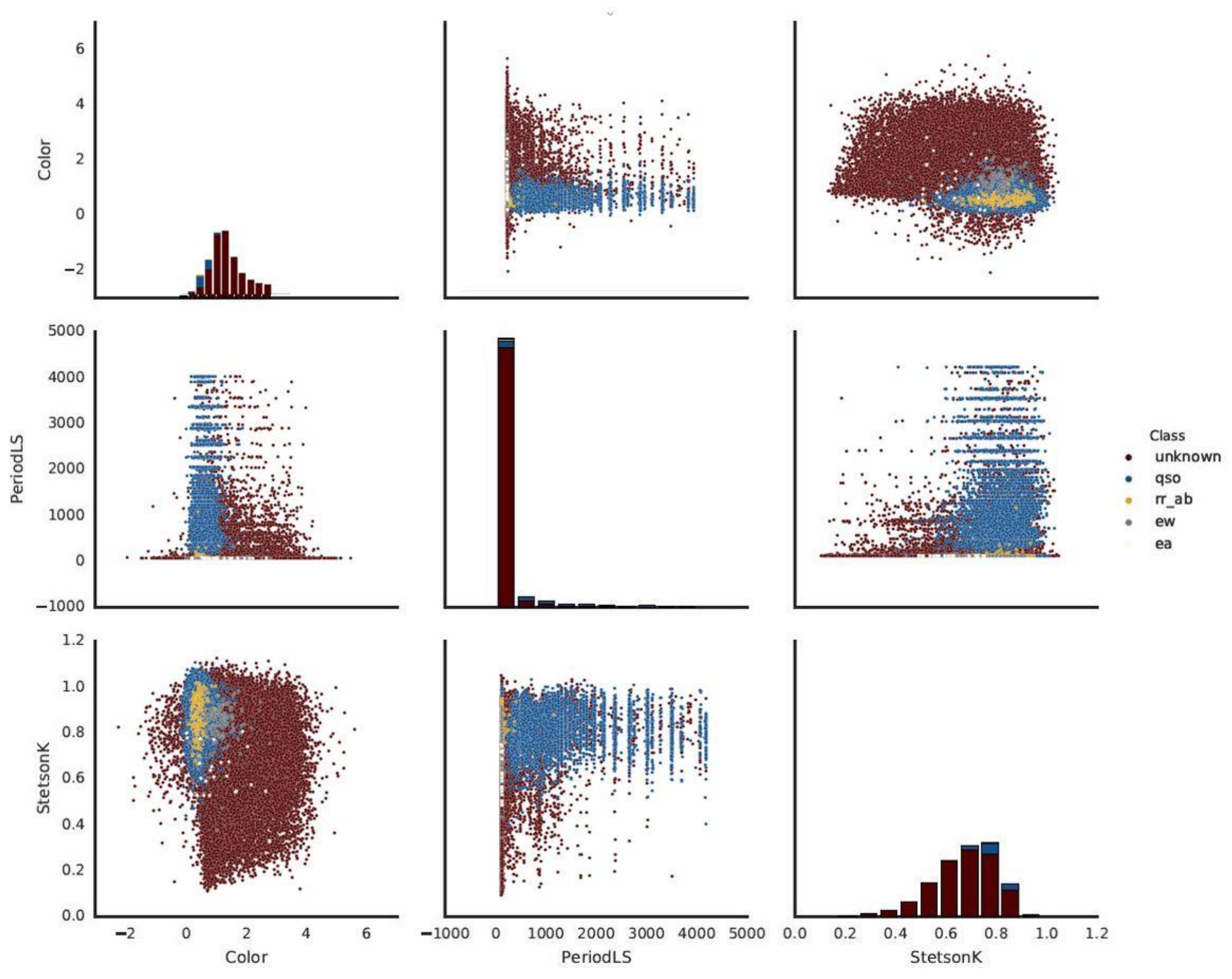


Figure 5. Different feature distribution for each class: eclipsing binary star system (grey and white), RR Lyrae (yellow), quasars (blue), and unknown (maroon)

Once enough rules have been written to successfully identify the different types of variable objects, the rest of the unlabeled objects in the dataset can be classified. Upcoming large surveys such as LSST (the Large Synoptic Survey Telescope) could utilize these labeled variable sources as a training set for its own classifications.

Citations

1. Ivezic, Z., Smith, J.A., Miknaitis, G., et al. 2007, Astronomical Journal, 134, 973
2. Drake, A. J., Djorgovski, S. G., Catelan, M., et al. 2017, MNRAS, 469, 3688
3. Astropy Collaboration, Robitaille, T. P., Tremonti, C., et al. 2003, MNRAS, 346, 1055

Acknowledgments

Texas A&M University thanks Charles R. '62 and Judith G. Munnerlyn, George P. '40 and Cynthia Woods Mitchell, and their families for support of astronomical instrumentation activities in the Department of Physics and Astronomy.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is http://www.sdss.org/.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

Texas A&M University Department of Physics and Astronomy is an institutional member of:

Munnerlyn Astronomical Instrumentation Lab
Texas A&M University

GMT

HETDEX Illuminating the Darkness
Hobby-Eberly Telescope Dark Energy Experiment

LSST Large Synoptic Survey Telescope

DARK ENERGY SURVEY