The Journal of Open Source Software

DBSP_DRP: A Python package for automated spectroscopic data reduction of DBSP data

Milan S. Roberson[*1,2], **Christoffer Fremling**[2], **and Mansi M. Kasliwal**[2]

1 Schmidt Academy of Software Engineering, California Institute of Technology **2** Division of Physics, Mathematics and Astronomy, California Institute of Technology

DOI: 10.21105/joss.03612

Software

- Review
- Repository
- Archive

Editor: Arfon Smith

Reviewers:

- @crhea93
- @arjunsavel

Submitted: 19 July 2021
Published: 28 February 2022

Summary

In astronomy, the spectrum of light emitted from astrophysical sources is of great use, allowing astronomers to classify objects and measure their properties. To measure the spectrum of a source, astronomers use spectrographs, in which dispersive elements spatially separate the incoming light by wavelength, and detectors, most commonly CCDs, image this dispersed light. But to do science with the spectrum, the 2D image in pixel coordinates taken by the CCD must be converted into a 1D spectrum of flux vs. wavelength. This process of converting 2D CCD images into 1D spectra is called data reduction.

To increase the signal-to-noise ratio, astronomers can take multiple exposures of the same object and coadd their 1D spectra to reveal faint absorption lines or increase the precision with which an important emission line can be measured. Many spectrographs have multiple paths that light can go through, and multiple detectors, each measuring a particular part of the spectrum, to increase the wavelength range that can be captured in a single exposure, or to allow the high resolution observation of distinct wavelength ranges. If two detectors cover an overlapping region, caused by partial reflectance of a dichroic (wavelength-dependent beam splitter), then the spectra from each detector need to be spliced together, combining the light collected by each detector.

DBSP_DRP is a python package that provides fully automated data reduction of data taken by the Double Spectrograph (DBSP) at the 200-inch Hale Telescope at Palomar Observatory (Oke & Gunn, 1982). The underlying data reduction functionality to extract 1D spectra, perform flux calibration and correction for atmospheric absorption, and coadd spectra together is provided by PypeIt (Prochaska et al., 2020). The new functionality that DBSP_DRP brings is in orchestrating the complex data reduction process by making smart decisions so that no user input is required after verifying the correctness of the metadata in the raw FITS files in a table-like GUI. Though the primary function of DBSP_DRP is to automatically reduce an entire night of data without user input, it has the flexibility for astronomers to fine-tune the data reduction with GUIs for manually identifying the faintest objects, as well as exposing the full set of PypeIt parameters to be tweaked for users with particular science needs. DBSP_DRP also handles some of the occasional quirks specific to DBSP, such as swapping FITS header cards, adding (an) extra null byte/s to FITS files making them not conform to the FITS specification, and not writing the coordinates of the observation to file. Additionally, DBSP_DRP contains a quicklook script for making real-time decisions during an observing run, and it can open a GUI displaying a minimally reduced exposure in under 15 seconds. Docker containers are available for ease of deploying DBSP_DRP in its quicklook configuration (without some large atmospheric model files) or in its full configuration.

*Present address: Department of Physics and Astronomy, University of California, Los Angeles.

Statement of Need

Palomar Observatory, located near San Diego, CA, is a multinational observatory with a broad user base. Users come from large and small institutions, and their observing experience ranges from novice to expert. One responsibility for serving such a diverse user base is to provide software data reduction pipelines for the most frequently used instruments, such as the Palomar Double Spectrograph (DBSP). Although DBSP was commissioned in 1982, it remains the workhorse instrument of the 200” Hale Telescope. It is used on 42% of the nights in a year, comprising nearly all of the valuable “dark” (moonless) time. In previous years, standard astronomical practice left the data reduction up to the user. However, attitudes in instrument-building have shifted since DBSP was built. A pipeline is now considered an indispensable component of an astronomical instrument. In fact, the difference between a good pipeline and a great pipeline means the difference between counting some of the photons vs. counting all of the photons.

Spectroscopy is a severe bottleneck in time-domain astronomy; currently less than 10% of discoveries are spectroscopically classified. Without a pipeline, data reduction is a difficult process and the standard method without a pipeline is to use IRAF (Tody, 1986, 1993), a 35-year-old program on which development and maintenance was discontinued in 2013 and whose use is discouraged by many in the field (e.g. Ogaz & Tollerud (2018)). Needless to say, data reduction sans existing pipeline is extremely time-consuming. There is a clear need for a modern and stable automated data reduction pipeline for DBSP.

During observing runs, one would like to be able to quickly inspect data as it is taken, in order to ensure that it is of sufficient quality to do the desired science with. For objects whose brightness may have changed between a previous observation and the current observing run, or for nights with highly variable cloud cover, the observer may be unsure how long of an exposure is needed to produce quality data. For very faint objects, objects in crowded fields, or objects with uncertain positions (e.g. due to high or uncertain motion across the sky), the observer may not even be sure that the telescope is pointed at the right object! A quicklook functionality, which can do a rudimentary reduction to correct for instrumental signatures and subtract light from the sky, revealing the spectra of the objects observed, can answer questions of exposure time and whether the object observed is the right one.

DBSP_DRP is currently being used by the ZTF Bright Transient Survey (Fremling et al., 2020; Perley et al., 2020), the ZTF Census of the Local Universe (De et al., 2020), and a program investigating ZTF Superluminous Supernovae (Lunnan et al., 2020; Chen et al., in preparation). Ravi et al. (2021) is the first (known) publication that used DBSP_DRP for data reduction. The development of DBSP_DRP also lays the groundwork towards a fully automated pipeline for the Next Generation Palomar Spectrograph that is planned to be deployed on the Palomar 200-inch Hale Telescope in 2023.

Acknowledgements

MSR acknowledges funding from the Schmidt Academy of Software Engineering, which is supported by the generosity of Eric and Wendy Schmidt by recommendation of the Schmidt Futures program.

We thank the following members of the time domain astronomy group at Caltech for beta-testing and providing valuable feedback during the development of this pipeline: Andy Tzanidakis, Lin Yan, Aishwarya Dahiwale, Yuhan Yao, Yashvi Sharma, and Igor Andreoni.

MSR is extremely grateful to the welcoming, friendly, and helpful team of developers on the PypeIt team, without whom this package would not exist.

This research made use of Astropy,[1] a community-developed core Python package for Astronomy (Astropy Collaboration et al., 2018, 2013).

References

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., Lim, P. L., Crawford, S. M., Conseil, S., Shupe, D. L., Craig, M. W., Dencheva, N., Ginsburg, A., Vand erPlas, J. T., Bradley, L. D., Pérez-Suárez, D., de Val-Borro, M., Aldcroft, T. L., Cruz, K. L., Robitaille, T. P., Tollerud, E. J., … Astropy Contributors. (2018). The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package. *The Astronomical Journal*, *156*(3), 123. https://doi.org/10.3847/1538-3881/aabc4f

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., Greenfield, P., Droettboom, M., Bray, E., Aldcroft, T., Davis, M., Ginsburg, A., Price-Whelan, A. M., Kerzendorf, W. E., Conley, A., Crighton, N., Barbary, K., Muna, D., Ferguson, H., Grollier, F., Parikh, M. M., Nair, P. H., … Streicher, O. (2013). Astropy: A community Python package for astronomy. *Astronomy and Astrophysics*, *558*, A33. https://doi.org/10.1051/0004-6361/201322068

De, K., Kasliwal, M. M., Tzanidakis, A., Fremling, U. C., Adams, S., Aloisi, R., Andreoni, I., Bagdasaryan, A., Bellm, E. C., Bildsten, L., Cannella, C., Cook, D. O., Delacroix, A., Drake, A., Duev, D., Dugas, A., Frederick, S., Gal-Yam, A., Goldstein, D., … Yao, Y. (2020). The Zwicky Transient Facility Census of the Local Universe. I. Systematic Search for Calcium-rich Gap Transients Reveals Three Related Spectroscopic Subclasses. *The Astrophysical Journal*, *905*(1), 58. https://doi.org/10.3847/1538-4357/abb45c

Fremling, C., Miller, A. A., Sharma, Y., Dugas, A., Perley, D. A., Taggart, K., Sollerman, J., Goobar, A., Graham, M. L., Neill, J. D., Nordin, J., Rigault, M., Walters, R., Andreoni, I., Bagdasaryan, A., Belicki, J., Cannella, C., Bellm, E. C., Cenko, S. B., … Kulkarni, S. R. (2020). The Zwicky Transient Facility Bright Transient Survey. I. Spectroscopic Classification and the Redshift Completeness of Local Galaxy Catalogs. *The Astrophysical Journal*, *895*(1), 32. https://doi.org/10.3847/1538-4357/ab8943

Lunnan, R., Yan, L., Perley, D. A., Schulze, S., Taggart, K., Gal-Yam, A., Fremling, C., Soumagnac, M. T., Ofek, E., Adams, S. M., Barbarino, C., Bellm, E. C., De, K., Fransson, C., Frederick, S., Golkhou, V. Z., Graham, M. J., Hallakoun, N., Ho, A. Y. Q., … Yao, Y. (2020). Four (Super)luminous Supernovae from the First Months of the ZTF Survey. *The Astrophysical Journal*, *901*(1), 61. https://doi.org/10.3847/1538-4357/abaeec

Ogaz, S., & Tollerud, E. (2018). Removing the Institute's Dependence on IRAF (You can do it too!). *STScI Newsletter*, *35*(03).

Oke, J. B., & Gunn, J. E. (1982). An Efficient Low Resolution and Moderate Resolution Spectrograph for the Hale Telescope. *Publications of the Astronomical Society of the Pacific*, *94*, 586. https://doi.org/10.1086/131027

Perley, D. A., Fremling, C., Sollerman, J., Miller, A. A., Dahiwale, A. S., Sharma, Y., Bellm, E. C., Biswas, R., Brink, T. G., Bruch, R. J., De, K., Dekany, R., Drake, A. J., Duev, D. A., Filippenko, A. V., Gal-Yam, A., Goobar, A., Graham, M. J., Graham, M. L., … Yan, L. (2020). The Zwicky Transient Facility Bright Transient Survey. II. A Public Statistical Sample for Exploring Supernova Demographics. *The Astrophysical Journal*, *904*(1), 35. https://doi.org/10.3847/1538-4357/abbd98

Prochaska, J. X., Hennawi, J. F., Westfall, K. B., Cooke, R. J., Wang, F., Hsyu, T., Davies, F. B., Farina, E. P., & Pelliccia, D. (2020). PypeIt: The python spectroscopic data reduction pipeline. *Journal of Open Source Software*, *5*(56), 2308. https://doi.org/10.21105/joss.02308

[1] http://www.astropy.org

Ravi, V., Law, C. J., Li, D., Aggarwal, K., Burke-Spolaor, S., Connor, L., Lazio, T. J. W., Simard, D., Somalwar, J., & Tendulkar, S. P. (2021). The host galaxy and persistent radio counterpart of FRB 20201124A. *arXiv e-Prints*, arXiv:2106.09710. https://arxiv.org/abs/2106.09710

Tody, D. (1993). IRAF in the Nineties. In R. J. Hanisch, R. J. V. Brissenden, & J. Barnes (Eds.), *Astronomical data analysis software and systems II* (Vol. 52, p. 173).

Tody, D. (1986). The IRAF Data Reduction and Analysis System. In D. L. Crawford (Ed.), *Instrumentation in astronomy VI* (Vol. 627, p. 733). https://doi.org/10.1117/12.968154