The Journal of Open Source Software

lightcurver: A Python Pipeline for Precise Photometry of Multiple-Epoch Wide-Field Images

Frédéric Dux [1,2]

1 European Southern Observatory, Alonso de Córdova 3107, Vitacura, Santiago, Chile **2** Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

DOI: 10.21105/joss.06775

Software

- Review
- Repository
- Archive

Editor: Ivelina Momcheva

Reviewers:

- @Onoddil
- @benjaminrose

Submitted: 24 March 2024
Published: 15 October 2024

Summary

lightcurver is a photometry pipeline for time series astronomical imaging data, designed for the semi-automatic extraction of precise light curves from small, blended targets. Such targets include, but are not limited to, lensed quasars, supernovae, or Cepheids in crowded fields. lightcurver is not a general purpose photometry, astrometry, and classification pipeline like legacypipe (Legacy Surveys Collaborations, 2024). Instead, it is a framework tailored to the precise study of a small region of interest (ROI) in wide-field images, utilizing stars surrounding the ROI to calibrate the frames.

At its core, lightcurver leverages STARRED (Michalewicz et al., 2023; Millon et al., 2024) to generate state-of-the-art empirical point spread function (PSF) models for each image. It then determines the relative zeropoints between images by combining the PSF-photometry fluxes of several stars in the field of view. Subsequently, STARRED is used again to simultaneously model the calibrated pixels of the ROI across all epochs. This process yields light curves of the point sources and a high-resolution image model of the ROI, cumulating the signal from all epochs.

lightcurver aims to be maintainable, fast, and incremental in its processing approach. As such, it can enable the daily photometric analysis of a large number of blended targets in the context of the upcoming Rubin Observatory Legacy Survey of Space and Time (LSST; Vera C. Rubin Observatory LSST Solar System Science Collaboration et al., 2021).

Statement of need

The LSST survey will generate an unprecedented amount of imaging data, revisiting the same regions of the sky every four days, with irregular pointings due to its observing strategy. Processing data at this cadence will require robust pipelines capable of ingesting new observations and providing immediate photometric calibration and analysis. This is particularly important for time-sensitive targets of opportunity, where rapid reaction to changes is essential for timely follow-up. An existing pipeline that performs this precise deblending and photometric measurement task, COSMOULINE (Magain et al., 1998; The COSMOGRAIL collaboration, 2010), requires too much manual intervention to be run on a daily basis.

On the other hand, STARRED is a powerful PSF modelling and deconvolution package, ideal for this task. However, by its nature, it cannot include an infrastructure that makes it convenient to apply to large datasets without manual intervention (e.g., visually identifying appropriate stars, extracting cutouts, and all subsequent processing steps leading to a light curve). Particularly, STARRED modelling requires a very stable zeropoint across modelled epochs, as it emulates the constant components of the ROI as one grid of pixels common to all epochs, which it simultaneously optimizes together with the fluxes of the variables. Achieving such precise relative zeropoint calibration (typically one millimag), especially in an automated manner, comes with challenges.

`lightcurver` addresses this challenge by automatically selecting calibration stars, modelling them, and robustly combining their fluxes to calibrate the zeropoints, making it suitable as a daily running pipeline on a large number of ROIs.

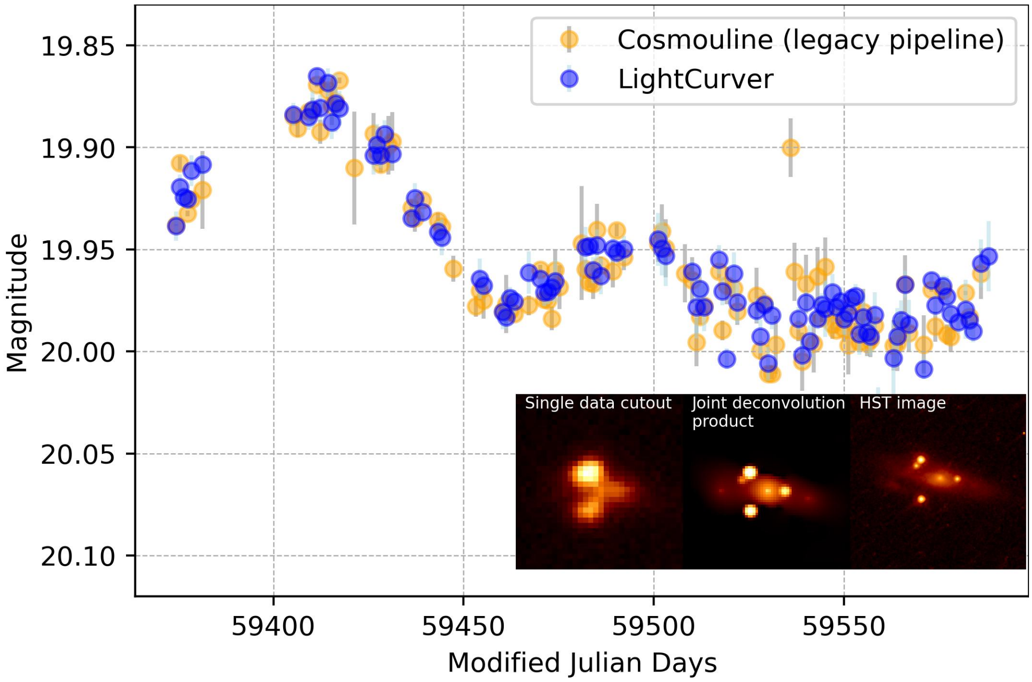

Figure 1: Light curve of a lensed image of a quasar (J0030-1525), extracted once with the existing code base (`COSMOULINE`), requiring a week of investigor's time, and another time with `lightcurver`, requiring about an hour of investigator's time. HST image: PI Tommaso Treu, proposal GO 15652 (Treu, 2018).

Functionality

`lightcurver` utilizes an SQLite3 database to track data processing stages and relies on SQL queries to manage its workflow, identifying the processing required at each step. First, the frames undergo background subtraction, and the sources are extracted using `sep` (Barbary, 2016; Bertin & Arnouts, 1996). The positions of the extracted sources are then used to plate-solve each frame, primarily with `Astrometry.net` (Lang et al., 2010). This permits sanity checks with `pyephem` (Rhodes, 2011), but also allows for an automatic selection of calibration stars around the ROI by querying Gaia (Gaia Collaboration, 2016) with `astroquery` (Ginsburg et al., 2019) for suitable stars. The pointings and field rotations do not need to be stable across epochs, as each frame is assigned its own calibration stars with the help of `shapely` (Gillies et al., 2024).

Subsequently, cutouts of the ROI and stars are extracted using `astropy` (Astropy Collaboration, 2013; Astropy Collaboration et al., 2018, 2022), masked, cleaned from cosmic rays with the help of `astroscrappy` (McCully et al., 2018; van Dokkum, 2001), and stored in an HDF5 file (Fortner, 1998). The PSF model is then calculated for each frame with `STARRED` before being stored in the same HDF5 file. Next, the fluxes of all the calibration stars in all frames are measured using PSF photometry, and the resulting fluxes of the stars are scaled and combined to obtain precise relative zeropoints. To avoid clearly failed fits from corrupting the reduction, each fitting procedure (PSF or photometry) stores its reduced chi-squared statistic in the database, allowing downstream steps to filter which frame, PSF, or flux to proceed with.

The software is divided into three main sub-packages: `processes`, containing individual data processing tasks; `pipeline`, defining the sequence of these tasks to ensure orderly data analysis;

and `structure`, containing the database schema and handles user configuration. Users can customize the processing of datasets through a YAML configuration file, allowing flexibility in handling various data characteristics. Typically, the YAML configuration file needs to be configured once when executing the pipeline on the first few frames, but the subsequent addition of new frames as they are observed requires no further manual intervention.

`lightcurver`, in comparison to `COSMOULINE`, achieves equal or better photometric precision in a much more automated fashion. Figure 1 presents the light curve of the southernmost lensed image of the quasar J0030-1525 (Lemon et al., 2018), extracted from the same dataset (ESO program 0106.A-9005(A), PI Courbin) using both `COSMOULINE` and `lightcurver`. The stable zeropoint across frames enables STARRED to reliably fit the constant components, in this case, two galaxies visible in the image. This reliable deblending of the different flux components yields both light curves and a high resolution image, the morphology of which is confirmed by comparison with Hubble Space Telescope imaging.

In summary, `lightcurver` is a robust and efficient photometry pipeline designed for the semi-automatic extraction of precise light curves from small, blended targets in cadenced astronomical imaging data. By leveraging the power of STARRED for state-of-the-art PSF modelling and deconvolution, and employing an automated flux calibration process, `lightcurver` achieves equal or better photometric precision compared to existing pipelines, while requiring significantly less manual intervention.

Acknowledgments

Special thanks to Martin Millon, Malte Tewes, Vivien Bonvin and Frederic Courbin for conceptualizing and creating the `COSMOULINE` pipeline. While this code base it not a re-write of `COSMOULINE`, it certainly shares the philosophy of some of the processing steps.

I am grateful to the two reviewers, Tom Wilson and Benjamin Rose, for sharing their expertise in identifying and reporting potential issues in `lightcurver`, and to the JOSS team for making this publication platform possible.

This project also made use of some of the backbone packages of scientific Python computing: NumPy (Harris et al., 2020), SciPy (Virtanen et al., 2020), Matplotlib (Hunter, 2007) and Pandas (McKinney, 2010; pandas, 2020). For first guess photometry, this software also benefited from the functions of `photutils` (Larry Bradley et al., 2024), and for finding transformations between frames, `astroalign` (Beroiz et al., 2020).

Astropy Collaboration. (2013). Astropy: A community Python package for astronomy. *Astronomy and Astrophysics*, *558*. https://doi.org/10.1051/0004-6361/201322068

Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., Earl, N., Starkman, N., Bradley, L., Shupe, D. L., Patil, A. A., Corrales, L., Brasseur, C. E., N"othe, M., Donath, A., Tollerud, E., Morris, B. M., Ginsburg, A., Vaher, E., Weaver, B. A., Tocknell, J., Jamieson, W., … Astropy Project Contributors. (2022). The Astropy Project: Sustaining and growing a community-oriented open-source project and the latest major release (v5.0) of the core package. *The Astrophysical Journal*, *935*(2), 167. https://doi.org/10.3847/1538-4357/ac7c74

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

Barbary, K. (2016). SEP: Source Extractor as a library. *Journal of Open Source Software*, *1*(6), 58. https://doi.org/10.21105/joss.00058

Beroiz, M., Cabral, J. B., & Sanchez, B. (2020). Astroalign: A Python module for astronomical image registration. *Astronomy and Computing*, *32*, 100384. https://doi.org/10.1016/j.ascom.2020.100384

Bertin, E., & Arnouts, S. (1996). SExtractor: Software for source extraction. *Astronomy and Astrophysics Supplement*, *117*, 393–404. https://doi.org/10.1051/aas:1996164

Fortner, B. (1998). HDF: The hierarchical data format. *Dr Dobb's J Software Tools Prof Program*, *23*(5), 42.

Gaia Collaboration. (2016). The Gaia mission. *Astronomy and Astrophysics*, *595*. https://doi.org/10.1051/0004-6361/201629272

Gillies, S., Wel, C. van der, Van den Bossche, J., Taves, M. W., Arnott, J., Ward, B. C., & others. (2024). *Shapely*. Zenodo. https://doi.org/10.5281/zenodo.5597138

Ginsburg, A., Sipőcz, B. M., Brasseur, C. E., Cowperthwaite, P. S., Craig, M. W., Deil, C., Guillochon, J., Guzman, G., Liedtke, S., Lian Lim, P., Lockhart, K. E., Mommert, M., Morris, B. M., Norman, H., Parikh, M., Persson, M. V., Robitaille, T. P., Segovia, J.-C., Singer, L. P., … a subset of the astropy collaboration. (2019). Astroquery: An astronomical web-querying package in Python. *The Astronomical Journal*, *157*, 98. https://doi.org/10.3847/1538-3881/aafc33

Harris, C. R., Millman, K. J., Walt, S. J. van der, Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M. H. van, Brett, M., Haldane, A., Río, J. F. del, Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Lang, D., Hogg, D. W., Mierle, K., Blanton, M., & Roweis, S. (2010). Astrometry.net: Blind astrometric calibration of arbitrary astronomical images. *The Astronomical Journal*, *139*(5), 1782–1800. https://doi.org/10.1088/0004-6256/139/5/1782

Larry Bradley, Brigitta Sipőcz, Thomas Robitaille, Erik Tollerud, Zé Vinícius, Christoph Deil, Kyle Barbary, Tom J Wilson, Ivo Busko, Axel Donath, Hans Moritz Günther, Mihai Cara, P. L. Lim, Sebastian Meßlinger, Zach Burnett, Simon Conseil, Michael Droettboom, Azalee Bostroem, E. M. Bray, … Gabriel Perren. (2024). *astropy/photutils: 1.12.0*. Zenodo. https://doi.org/10.5281/zenodo.596036

Legacy Surveys Collaborations. (2024). Legacypipe: Image reduction pipeline for the DESI legacy imaging surveys. In *GitHub repository*. GitHub. https://github.com/legacysurvey/legacypipe

Lemon, C. A., Auger, M. W., McMahon, R. G., & Ostrovski, F. (2018). Gravitationally lensed quasars in Gaia – II. Discovery of 24 lensed quasars. *Monthly Notices of the Royal Astronomical Society*, *479*(4), 5060–5074. https://doi.org/10.1093/mnras/sty911

Magain, P., Courbin, F., & Sohy, S. (1998). Deconvolution with correct sampling. *The Astrophysical Journal*, *494*(1), 472–477. https://doi.org/10.1086/305187

McCully, C., Crawford, S., Gabor Kovacs, Tollerud, E., Betts, E., Bradley, L., Craig, M., Turner, J., Streicher, O., Sipocz, B., Robitaille, T., & Deil, C. (2018). *astropy/astroscrappy: v1.0.5 Zenodo Release*. Zenodo. https://doi.org/10.5281/zenodo.1482018

McKinney, Wes. (2010). Data structures for statistical computing in Python. In Stéfan van der Walt & Jarrod Millman (Eds.), *Proceedings of the 9th Python in Science Conference* (pp. 56–61). https://doi.org/10.25080/Majora-92bf1922-00a

Michalewicz, K., Millon, M., Dux, F., & Courbin, F. (2023). STARRED: a two-channel deconvolution method with Starlet regularization. *The Journal of Open Source Software*, *8*(85), 5340. https://doi.org/10.21105/joss.05340

Millon, M., Michalewicz, K., Dux, F., Courbin, F., & Marshall, P. J. (2024). Image deconvolution and point-spread function reconstruction with STARRED: A wavelet-based two-channel method optimized for light-curve extraction. *The Astronomical Journal*, *168*(2), 55. https://doi.org/10.3847/1538-3881/ad4da7

pandas, development team. (2020). *pandas-dev/pandas: Pandas* (latest). Zenodo. https://doi.org/10.5281/zenodo.3509134

Rhodes, B. C. (2011). *PyEphem: Astronomical ephemeris for Python*. Astrophysics Source Code Library, record ascl:1112.014.

The COSMOGRAIL collaboration. (2010). COSMOULINE. In *GitHub repository*. GitHub. https://github.com/cosmograil/cosmouline

Treu, T. L. (2018). *H0, the stellar initial mass function, and other dark matters from a large sample of quadruply imaged quasars* (p. 15652). HST Proposal. Cycle 26, ID. #15652.

van Dokkum, P. G. (2001). Cosmic-ray rejection by Laplacian edge detection. *Publications of the Astronomical Society of the Pacific*, *113*(789), 1420–1427. https://doi.org/10.1086/323894

Vera C. Rubin Observatory LSST Solar System Science Collaboration, Jones, R. L., Bannister, M. T., Bolin, B. T., Chandler, C. O., Chesley, S. R., Eggl, S., Greenstreet, S., Holt, T. R., Hsieh, H. H., Ivezic, Z., Juric, M., Kelley, M. S. P., Knight, M. M., Malhotra, R., Oldroyd, W. J., Sarid, G., Schwamb, M. E., Snodgrass, C., … Trilling, D. E. (2021). The scientific impact of the Vera C. Rubin Observatory's Legacy Survey of Space and Time (LSST) for solar system science. *Bulletin of the AAS*, *53*(4). https://doi.org/10.3847/25c2cfeb.d8909f28

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2