Faro: A framework for measuring the scientific performance of petascale Rubin Observatory data products

Leanne P. Guy[1], Keith Bechtol[1,2], Jeffrey L. Carlin[1], Erik Dennihy[1], Peter S. Ferguson[2], K. Simon Krughoff[1], Robert H. Lupton[3], Colin T. Slater[4], Krzysztof Findeisen[4], Arun Kannawadi[3], Lauren A. MacArthur[3], Sophie L. Reed[3], Dan S. Taranu[3], and W. Michael Wood-Vasey[5]

[1]Rubin Observatory Project Office, 950 N. Cherry Ave., Tucson, AZ 85719, USA
[2]Department of Physics, University of Wisconsin-Madison, Madison, WI 53706, USA
[3]Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
[4]University of Washington, Dept. of Astronomy, Box 351580, Seattle, WA 98195, USA
[5]Department of Physics and Astronomy, University of Pittsburgh, 3941 O'Hara Street, Pittsburgh, PA 15260, USA

2022-01-04

ABSTRACT

The Vera C. Rubin Observatory will advance nearly every area of astronomy over the next decade with its unique wide-fast-deep multi-color imaging survey, the Legacy Survey of Space and Time (LSST). The LSST will produce approximately 20TB of raw data per night, which will be automatically processed by the LSST Science Pipelines to generate science-ready data products – processed images, catalogs and alerts. To ensure that these data products enable transformative science with LSST, stringent requirements have been placed on their quality and scientific fidelity, for example on image quality and depth, astrometric and photometric performance, and object recovery completeness. In this paper we introduce `faro`, a framework for automatically and efficiently computing scientific performance metrics on the LSST data products for units of data of varying granularity, ranging from single-detector to full-survey summary statistics. By measuring and monitoring metrics, we are able to evaluate trends in algorithmic performance and conduct regression testing during development, compare the performance of one algorithm against another, and verify that the LSST data products will meet performance requirements by comparing to specifications. We present initial results using `faro` to characterize the performance of the data products produced on simulated and precursor data sets, and discuss plans to use `faro` to verify the performance of the LSST commissioning data products.

100 word summary

Once operational, up to 20TB of raw imaging data will be collected by Rubin Observatory's Legacy Survey of Space and Time (LSST) per night and processed by the LSST Science Pipelines to produce science-ready data products – processed images, catalogs and alerts. To ensure that these data products enable transformative science with LSST, stringent requirements have been placed on their quality and scientific fidelity. In this paper we introduce `faro`, a package for automatically and efficiently computing scientific performance metrics on the LSST data products for units of data of varying granularity, ranging from single-detector to full-survey summary statistics.

Keywords: Rubin Observatory, LSST, performance characterization, verification

1. REFERENCES AND PRIOR ART

The LSST high-level science requirements are outlined in the SRD.[1]

Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST) Science Pipelines,[?,2]

Introducing validate_drp: Calculate SRD Key Performance Metrics for an output repository:[?]

The Rubin Observatory LSST Science Pipelines:[2] and[?]

LSST Science Pipeline Characterization Metric Report[3]

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Cooperative Agreement 1258333 managed by the Association of Universities for Research in Astronomy (AURA), and the Department of Energy under Contract No. DE-AC02-76SF00515 with the SLAC National Accelerator Laboratory. Additional funding for Rubin Observatory comes from private donations, grants to universities, and in-kind support from LSSTC Institutional Members.

APPENDIX A. REFERENCES

REFERENCES

[1] Ivezić, Ž. and The LSST Science Collaboration, "LSST Science Requirements Document," (Jan. 2018).

[2] Bosch, J., AlSayyad, Y., Armstrong, R., Bellm, E., Chiang, H.-F., Eggl, S., Findeisen, K., Fisher-Levine, M., Guy, L. P., Guyonnet, A., Ivezić, Ž., Jenness, T., Kovács, G., Krughoff, K. S., Lupton, R. H., Lust, N. B., MacArthur, L. A., Meyers, J., Moolekamp, F., Morrison, C. B., Morton, T. D., O'Mullane, W., Parejko, J. K., Plazas, A. A., Price, P. A., Rawls, M. L., Reed, S. L., Schellart, P., Slater, C. T., Sullivan, I., Swinbank, J. D., Taranu, D., Waters, C. Z., and Wood-Vasey, W. M., [*An Overview of the LSST Image Processing Pipelines*], vol. 523 of *Astronomical Society of the Pacific Conference Series*, 521 (2019).

[3] Carlin, J., "Characterization metric report: Science pipelines version 22.0.0," (July 2021).

APPENDIX B. ACRONYMS

Acronym	Description
AURA	Association of Universities for Research in Astronomy
DE	dark energy
DMTN	DM Technical Note
DMTR	DM Test Report
LPM	LSST Project Management (Document Handle)
LSST	Legacy Survey of Space and Time (formerly Large Synoptic Survey Telescope)
LSSTC	LSST Corporation
SLAC	SLAC National Accelerator Laboratory
SRD	LSST Science Requirements; LPM-17