

PDAC Tests: March 2017

Krzysztof Suberlak,¹★ Željko Ivezić,¹ and the PDAC team

¹*Department of Astronomy, University of Washington, Seattle, WA, United States*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

A report of progress made in testing the Preliminary Data Access Center user interface, infrastructure, and database ingestion. We list conducted tests commenting on the outcomes. This will help direct the focus of the PDAC development to make it user-friendly and efficient.

INTRODUCTION

Data overview

We continued testing the structure of the S82 database ingested in PDAC, performed tests of individual lightcurves, and Box queries to test source density. We started investigating the periodogram tool and ways in which the database can be made more user-friendly through improved metadata. We also tried ways of going beyond the built - in functions to analyse the data with user-defined functions.

Performed tests

- Tested whether it is possible to run code parallel to the database server, or locally to the data (at the NCSA), using a jupyter notebook. This would be helpful in the analysis of forced photometry, such as using the Bayesian faint flux pipeline developed at the University of Washington. In a discussion we found that the database level (qserv, scisql) User-Defined Functions (UDF) are not yet implemented. UDFs, and ability to run code next to shard server is FY 2019-2020 specification. The ability to run jupyter notebooks locally to the data (within the same computer cluster at eg. NCSA where the data is stored for current implementation of PDAC) is not yet implemented, but is projected to be available towards the end of FY 2017.
- Tested source density using Box queries. We found that to properly compare the PDAC data to the Deep Source catalog, locally stored at the UW, we need to account for the magnitude cutoff made to create the UW catalog ($i < 23.5$ mag). We found that a straightforward SQL query was illegal due to limitations of the server in understanding the alias. The solution was to use HAVING rather than WHERE clause. This issue also uncovered a degradation of quality of error messages as they are transferred from server to server (mariaDB to qserv, etc.)
- Tested the raw lightcurve ingestion to the database. We queried against a given position, and compared the raw data stored at PDAC to the UW-stored data. We found that direct comparison is not possible due to lack of documentation about the units of `flux_psf` column in RunDeepForcedSource Catalog of S82 data. Converting the PDAC data with the scisql functions to calibrated magnitudes we find that all the negative forced photometry measurements (where `flux_psf < 0`) are by default discarded by the scisql function.

This paper has been typeset from a T_EX/L^AT_EX file prepared by the author.