Cross-matching Engine for Incremental Photometric Sky Survey

Jiří Nádvorník – Czech Technical University in Prague
Petr Škoda – Astronomical Institute, Czech Academy of Sciences

Contact: nadvornik.ji@gmail.com

Abstract

We present the Ondřejov Southern Photometry Survey, being conducted at the Danish 1.54m telescope in remote observing mode by several groups of Czech stellar astronomers.

The automatic astrometry and photometry pipelines run on every CCD frame combined with sophisticated parallelized cross-matching and clustering algorithms result in an on-the-fly generation of light curves of every single object in the field.

To allow powerful querying and visualization of current database of more than half billion of measurements, the technology of Virtual Observatory is used, combining IVOA protocols and powerful visualization tools as Aladin, TOPCAT and SPLAT-VO.

Introduction

Goals

- Access products during **all phases** of data processing (image management, preprocessing, data mining, …) with **VO protocols**:
  - Images (Simple image access - SIA)
  - Observations (Table access - TAP, Simple cone search - SCS)
  - Light curves (Simple spectral access - SSA)
- Extract **light curves** from **any photometry survey**.
- Use existing **VO clients** (Aladin, Topcat, Splat-VO)
- Centralized storage of all data in one database

Benefits

- **No need** for pre-planned observation fields
- Can change observation fields anytime
- **No need** to have regular grid for observation fields
- Dynamic light curve generation (after adding new images)

Motivation

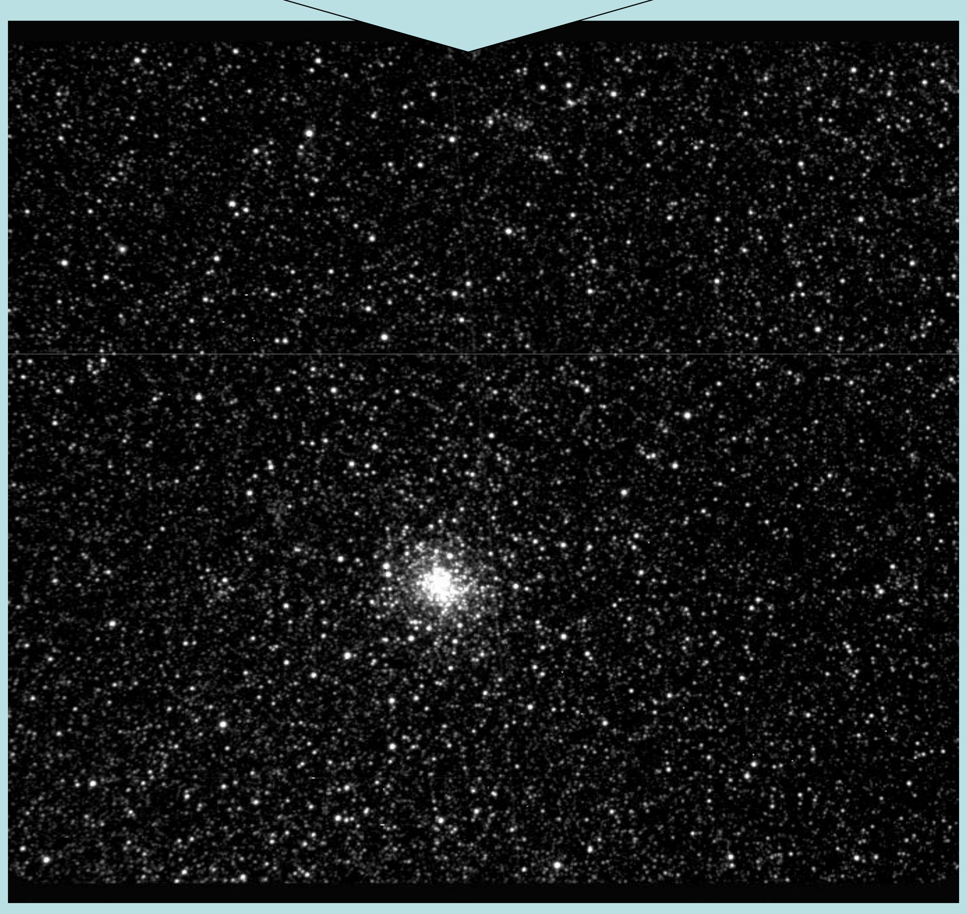
- Classical photometry analysis uses **only small fraction** of observed data (target variable + comparison stars)
- Can produce light curves from data of not primary interest

Pipeline

- Starts at La Silla robotically controlled telescope
- Changing priorities and targets
- Survey coverage is irregular, without any grid usable for differential photometry
- Individual photometry and astrometry needs to be calculated for each image, continues within Light curve identification

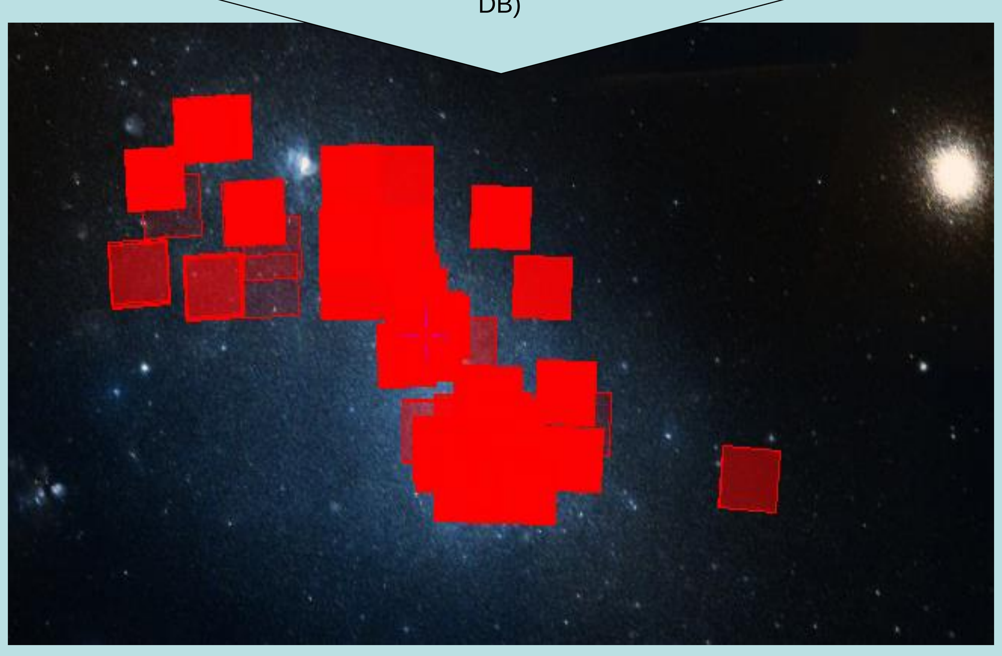
Danish DK1.54 telescope in Chile

Calibrating images
(produces FITS images)



Reduced image
(displayed with SIA in Aladin or web browser)

Storing images
(product is image metadata stored in DB)



Survey coverage of Small Magellanic Cloud
(displayed in Aladin with SIA)

Light curve identification

- Starts with astrometry and photometry for individual observations
- Light curve identifiers need to be generated
- For speedup, HEALPix tessellation is used
- All of the intermediate data products are published via VO protocols and can be visualized via VO tools

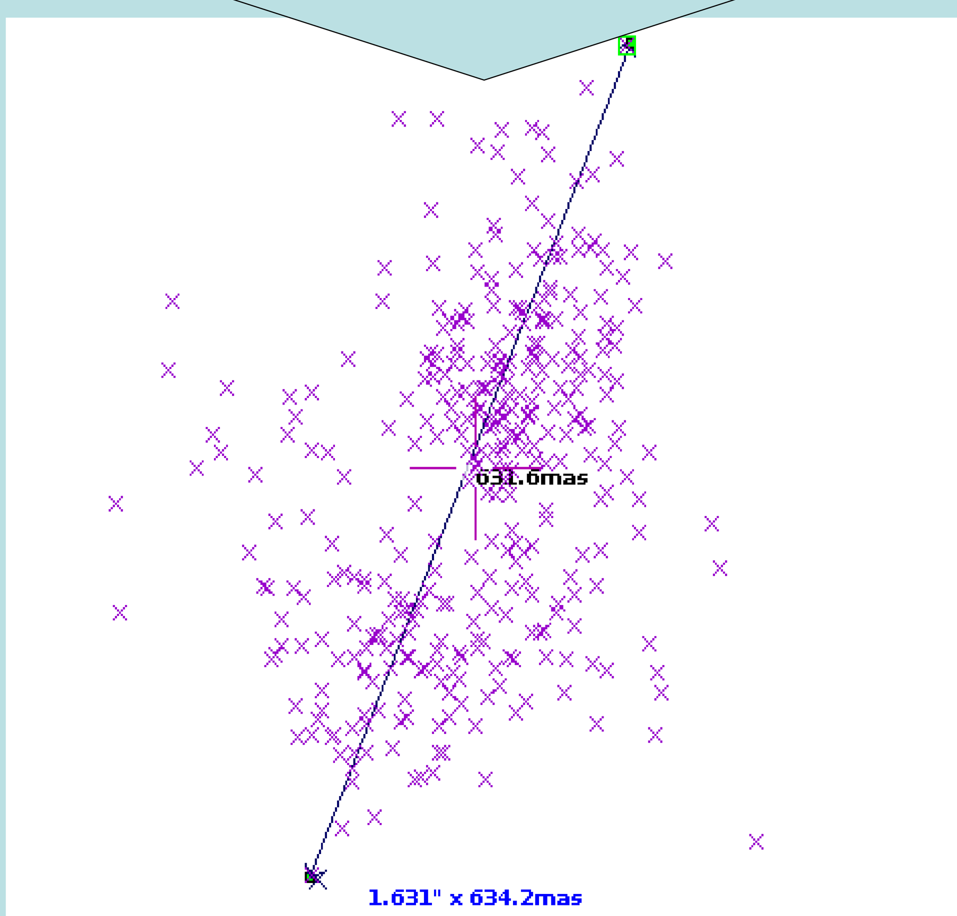
Individual sources
(Displayed via SCS or TAP in Topcat)

Table Browser for 1: smc-Cone-6s

	raj2000	dej2000	MAG	MAGERR	band
1	13.15438	-72.80147	19.1522	0.011065	R
2	13.15437	-72.80146	19.2179	0.014162	R
3	13.15429	-72.80144	19.108	0.011758	R
4	13.15435	-72.80145	19.3138	0.013023	R
5	13.15435	-72.80145	19.1899	0.009383	R
6	13.15437	-72.80145	19.1304	0.012005	R
7	13.15434	-72.80144	19.7292	0.014355	V
8	13.15432	-72.80144	19.2347	0.013462	R
9	13.15432	-72.80144	19.6423	0.014513	V
10	13.15436	-72.80144	19.0584	0.011104	R
11	13.15425	-72.80143	19.7798	0.013159	V
12	13.15425	-72.80143	19.2128	0.016681	R

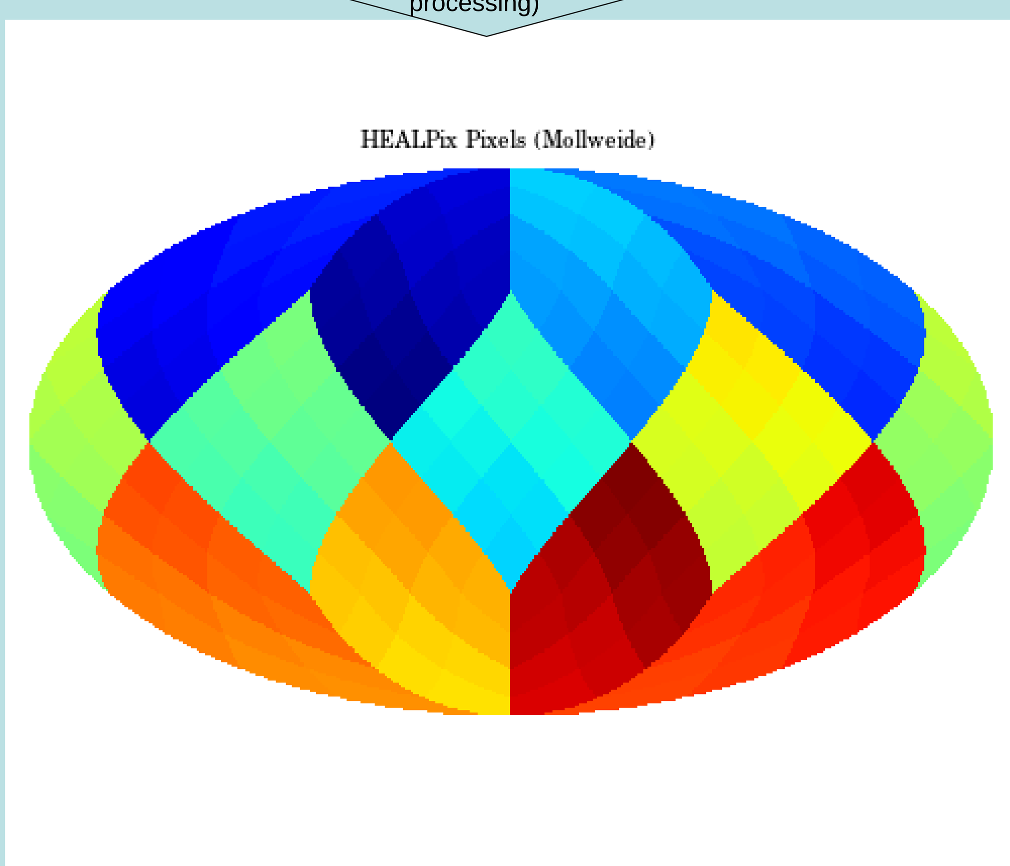
Observations of one star in multiple bands (displayed in Topcat via SCS)

Identifying **astrometry error**
(crucial parameter for physical object identification)



Observations of one star (displayed in Aladin via SCS)
Note the highlighted diameter of astrometry accuracy, cca 0.6 arcsec.

HEALPix tesselation
(slicing sky for parallel processing)



HEALPix tessellation

Assigning observations to object IDs
(**parallelized K-means** based clustering)

Clusters with their centers and underlying HEALPix pixels
(displayed in Aladin via SCS)
Note the highlighted distance of cca 1.1 arcsec between observed objects which is comparable with the astrometry accuracy.

Light curve generation (on the fly)

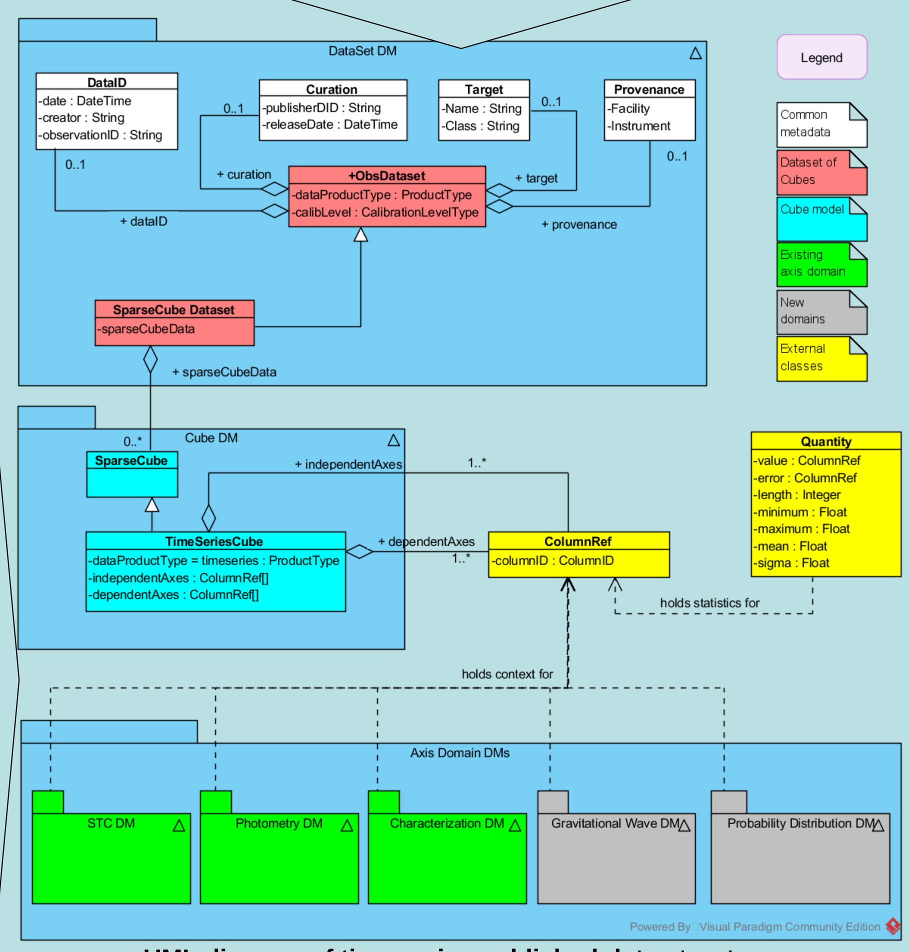
Results

- We maintain data with **GAVO DaCHS** package as wrapper around:
  - Data ingestion
  - Data processing
  - Data publishing
- All products are accessible in VO-compatible tools
- Publishing these products:
  - Raw + Reduced images (SIAP)
  - Observations identified on images (SCS, TAP)
  - Light curves (SSA)
- Light curve generation:
  - Light curves for **400 mil.** observations generated **under 2h** (10 cores, 32GB RAM)
  - Consuming ~8GB RAM per 100 mil. observations
  - Performance not dependent on number of objects observed

Publication

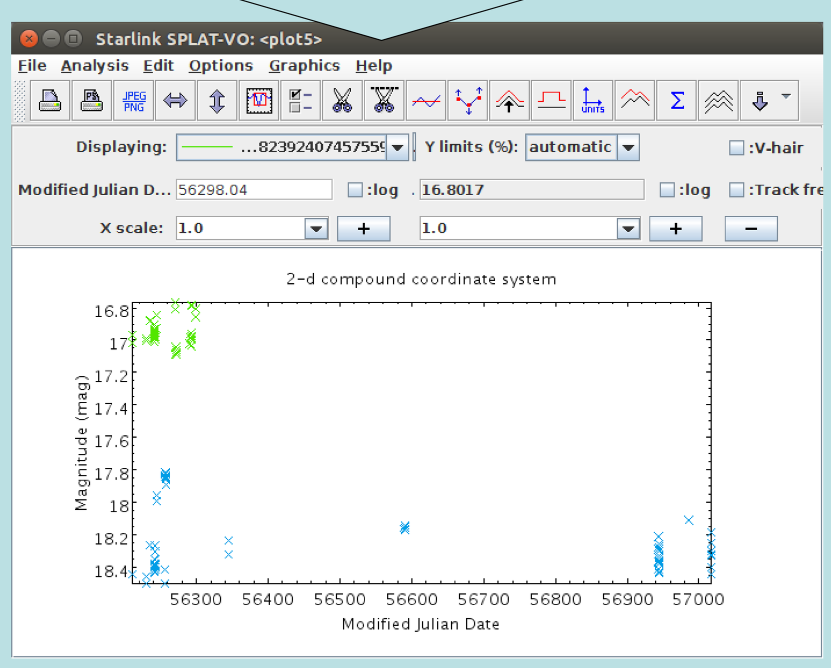
- We use Time Series Data Cube DM for publication
- Latest effort driven by needs of all time series data, not only light curves
- Unified publishing standard for all time series data
- Provided by VO technology

Provides protocols and standards



UML diagram of time series published data structure

Publication via VO protocols



Final SMC object light curve for bands U and I
(displayed in Splat-VO via SSA protocol)

Usages

- Usage of common **VO clients** instead of implementing new ones
- **Publishing** any time series related data products **with VO protocols** (images, tabular data, light curves, …)
- Convenient for heterogeneous surveys produced by different groups still can provide light curves
- Data mining from image sets where light curve extraction was **not possible** (need individual astrometry and photometry)
- HEALPix based parallelization can be used for **localized processing of any spherical** data (astronomy, geodesy, …)

Acknowledgement

This work was supported by grants 13-08195S of Czech Science Foundation and LD-15113 of Ministry of Education Youth and Sports of the Czech Republic.

References

- Demleitner, M. et al., (2014) *Virtual Observatory publishing with DaCHS,* Astronomy and Computing 7-8, 27
- Škoda, P. et al., (2014) *Spectroscopic analysis in the Virtual Observatory environment with SPLAT-VO*, Astronomy and Computing 7-8, 108
- Nádvorník, J. et al., (2017), *Time Series Cube Data Model*, arXiv:1702.01393
- Nádvorník, J., (2015), *Cross-matching Engine for Incremental Photometric Sky Survey*, arXiv:1506.07208