D3.11.2: Demonstration Software for One Example Catalogue

WP3.11: Cross-Matching and Astrometry at LSST Depths

Project Acronym	LUSC-B
Project Title	UK Involvement in the Legacy Survey of Space and Time
Document Number	LUSC-B-11

Submission date	7/Jan/2021
Version	1.0
Status	Review
Author(s) inc. institutional affiliation	Tom J. Wilson (Exeter) Tim Naylor (Exeter)
Reviewer(s)	Bob Mann (UEDIN), Raphael Shirley (SOTON)

Dissemination level
Restricted to project

Version History

Version	Date	Comments, Changes, Status	Authors, Contributors, Reviewers
1.0	07/01/21	First draft for review	Tom J. Wilson

Table of Contents

List of Figures

List of Tables

1 Introduction

WP3.11 is investigating the astrometry of LSST. An unprecented parameter space in optical photometric surveys, the depths to which the Rubin Observatory will probe during its 10-year initial survey will challenge data reduction and analysis tools. We are focusing on the problems surrounding the creation of value-added catalogues: those combining the LSST dataset with external, ancillary catalogues, to improve the science that can be done. To achieve this, we must identify those sources in common between the two catalogues which are truly one physical object on the sky, for which we have two detections. These must be, in turn, differentiated from sources which simply happen to appear very close in the sky to one another, but are two independent sources in the sky. This process of creating a single, composite dataset is the "cross-matching" of the two catalogues.

The challenge for LSST is primarily one of sheer source counts. Never before has an optical survey reached the expected depths of LSST – 27th magnitude at its full completeness limit – and thus never had to cope with the increase in source detection rate that goes along with such sensitivity. And yet the survey will still be limited to ground-based resolutions, with the atmospheric distortions that "spread" out otherwise point source objects into a larger pool of light on the CCD of the telescope. When crowding of sources, as we expect for LSST, becomes significant, these "point spread functions" (PSFs) can and will start to overlap one another. Thus a very bright source can "hide" within its PSF a faint object, which the detection algorithms will fail to include in the final catalogue of objects produced by the "LSST stack". Unfortunately, the faint source can influence the measured position of the bright source, even if it was not detected, by subtly tugging on the centroid of the bright object during its position determination.

Usually a photometric catalogue records the measured position of a source, and a corresponding uncertainty: the precision with which the algorithms and pipelines were able to pinpoint the determined position. For LSST, these precisions will be very good, and thus the positions will likely be quoted to high confidence. Crucially, however, these "tugs" from faint sources, buried within the light of the brighter source, can be larger than the precisions with which sources are pinpointed, and thus the simple questions asked by traditional cross-match algorithms "given the positions and corresponding precisions of these two sources, how likely are they to have the same sky position?" will break down. It is overcoming this limiting factor, expected to be significant for LSST, that WP3.11 is most interested in.

2 Demonstration Software

To enable more robust cross-matches of LSST and other catalogues, including the effects of position perturbation from blended objects, WP3.11 is mainly tasked with the creation of new software to allow for such cross-matches to occur. These matches will then be hosted on the UKDAC, accessible to users. Thus, after we had established it was possible to overcome or sidestep a few implementation challenges (see the D3.11.1 report for more details), we began the establishment of the codebase core, on to which the more advanced, and WP-focused, algorithms could be added.

2.1 Deliverable Name

The formal deliverable for D3.11.2 is titled "Demonstration Software for One Example Catalogue". This is because, during the proposal phase for this WP, it was envisaged that demonstration code would be written to specifically match an example catalogue pair, the code then

moved to run on the DAC and subsequently generalised to allow for any catalogue combinations in D3.11.3. In practice we thought it would save time to go directly to writing the software for the DAC, skipping the demonstration software. Thus D3.11.2 should really be "Demonstration of Preliminary Software for DAC Integration". The software that forms D3.11.2 is an end-to-end complete cross-match code, installable by the end user, featuring documentation, but does not yet have the sophisticated algorithms which are WP3.11's main task and will be the focus of D3.11.3.

2.2 Preliminary Software

The software that delivers D3.11.2 is a full, end-to-end, "many-to-many" cross-match code, capable of accepting two catalogues and producing posterior probabilities of likely matches and non-matches between "islands" – sources potentially astrometrically correlated with one another and definitely not correlated with any other source – of sources across the two catalogues. This limit to two catalogues, but allowing for inclusion of multiple matches simultaneously, is what makes it a "many-to-many" match software. It accepts the astrometric uncertainty – the precision with which sources were centroided – and uses the normalised on-sky separation of two sources to calculate their relative match likelihood. This simple match assumption – the natural assumption of most match algorithms, ignoring any other systematic effects which might cause separation between sources, such as proper motions or blending – and the non-inclusion of any other discriminating information that might be used to distinguish between two competing counterparts for a source mean that this implementation is currently a "naive Bayes" cross-match. However, these assumptions hold for a large number of previous generation photometric catalogues, and even in its half-finished state the cross-match codebase provides useful functionality.

3 Deliverables

The main deliverables for this work, being software development, are available online. As per the Project Management Plan, software development should be maintained in a version control repository. The main codebase is therefore located at https://github.com/Onoddil/macauff. It features a full test suite for validation, as well as functionality to generate test data to input into the cross-match algorithm for end-to-end verification.

In addition, to aid in the review, preliminary documentation – guides to installation and getting started with the codebase – is available at https://onoddil.github.io/macauff/.

4 Future Work

The main focus of D3.11.3, now that WP3.11 has an established codebase for its software development aims, is to implement the extended algorithmic aspects of its remit. First, we must include the algorithmic components that extend the Astrometric Uncertainty Function from the simple assumption that centroiding uncertainties are the only component, and include a prescription for the uncertainty in position that derives from the perturbation of a bright source by a fainter contaminant source within its PSF. Second, we plan to include an additional improvement to better provide users confident cross-matches: the use of the photometry of the sources to distinguish between true coevality in sources and happenstance astrometry; this can be achieved using the idea that as an ensemble sources of 15th magnitude in one optical bandpass

are going to have a roughly 15th magnitude source near to them in a similar optical wavelength, but unlikely to experience a 22nd magnitude source being coeval with them.

In addition, we plan to extend and improve the user documentation, both for the current features and for the additional features to be implemented. This is crucial to ensure that the codebase is useable by the LSST:UK project after the end of this current work package.