ASTR 598: Astro-statistics and Machine Learning

Andy Connolly & Željko Ivezić

University of Washington, Winter Quarter 2018

Class Project: Deep Proper Motion Catalog for SDSS Stripe 82 Region

Scientific Goals: The main scientific aim of this quarter-long class project is to produce a catalog with improved proper motion measurements for faint stars ($14 < r < 22$) in the SDSS Stripe 82 region (250 sq. deg.). The SDSS-based catalog[1] for 3.7 million objects constructed by Bramich et al. will be augmented with newer additional data obtained with the Dark Energy Camera[2], and possibly from other surveys, and the proper motions will be refit. If all goes well, the results should be publishable in a top astronomical journal.

Learning Goals: While working on this project, students will develop a working knowledge of the NOAO Datalab interface, astropy, pandas, astroML and other astronomical python tools, and of selected methods from astro-statistics (e.g., robust regression, Bayesian statistics, clustering, visualization).

Prerequisites: The students taking this class are required to open an account at the NOAO Data Lab site[3]. The Data Lab allows users to

- access, search, and filter databases containing large catalogs;
- create custom databases and analyses from large catalogs using familiar tools;
- combine catalog databases with data from NOAO telescopes, analysis results, and data from external archives in one place;
- share custom results easily with collaborators and create and publish catalogs derived from large data sets through a central workspace;
- experiment with tools being developed for LSST using existing large data sets.

We will use GitHub and Jupyter notebooks for progress tracking.

[1] Bramich et al. 2008, "Light and motion in SDSS Stripe 82: the catalogues", MNRAS 386, 887.
[2] See, for example, http://legacysurvey.org/decamls/
[3] http://datalab.noao.edu

Brief Project Outline

We will discuss detailed work plan in class (the 3rd week) and here we list only an overview of the main steps:

1. Download Bramich et al. catalog (HLC files) from the SDSS website[4] and perform preliminary analysis (e.g., the variation of proper motion errors with position and magnitude; also, see examples in the last bullet below and, in particular, reproduce Figs. 22 and 23 from Sesar et al.); Using Data Lab, perform similar analysis for DECaLS and other available data from the Stripe 82 region.

2. Cross-match to DECaLS data using Data Lab (technical details TBD).

3. Preliminary analysis and quality assurance of the assembled data set (e.g., coordinates vs. time plots for stars with large SDSS proper motions); decision whether astrometric recalibration is required (perhaps using galaxies).

4. Fit proper motions (see eqs. 3-6 in Bramich et al.) using the Bramich et al. SDSS values as priors. For coordinate sytstems, please see Section 2.6 in Bond et al. (2010, ApJ 716, 1).

5. Analyze updated proper motions (e.g., Fig. 2 in Vidrih et al. 2007, MNRAS 382, 515; Figs. 22 and 23 from Sesar et al. 2010, AJ 708, 717, color-coded with the mean proper motions), including using quasars for the assessment of systematic errors.

6. Write paper(s) and become rich and famous!

[4] http://das.sdss.org/va/stripe_82_variability/SDSS_82_public/