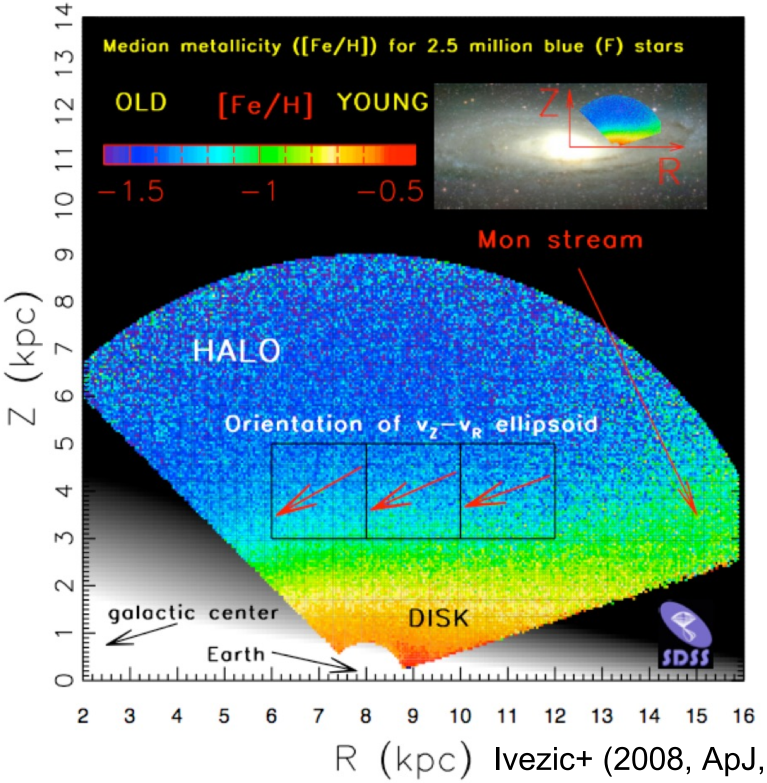
Rubin/LSST photo-D: Brief Introduction

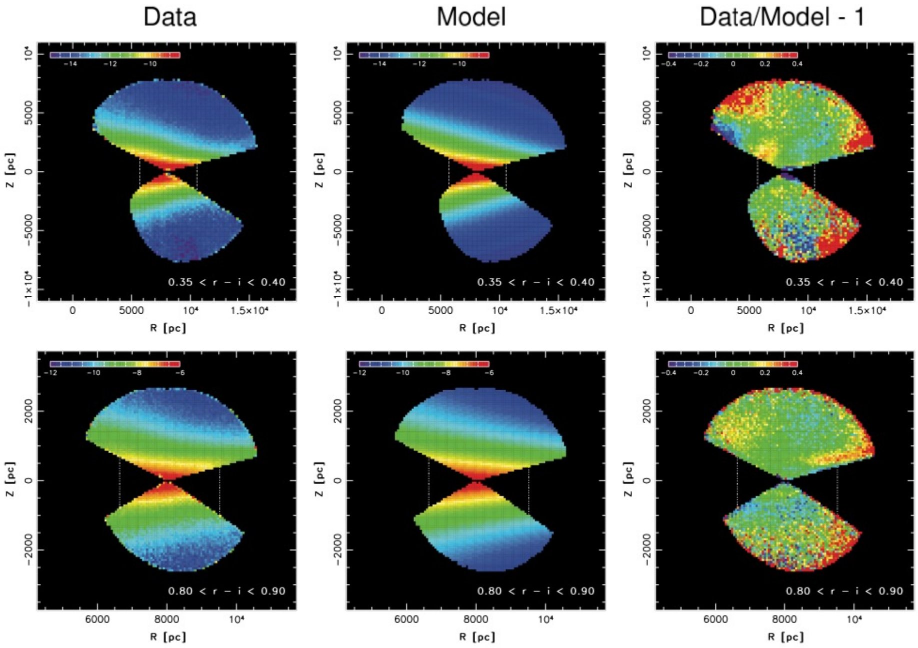
- Photo-D: estimate distances to stars using LSST photometry
  - akin to photo-z in the context of galaxies (redshift and galaxy type)
- Method: given T_{eff}, [Fe/H], and log(g), Luminosity follows and thus D
  - shown to work well with SDSS and Gaia data (and others)
- T_{eff}, [Fe/H], log(g) and dust A_V can be constrained using LSST photometry
  - need full Bayesian treatment ("if MS then best D is…; if WD then best D is…")
  - not compute intensive per star, but 10 billion stars
  - a good code base (chi2 minimizer with model SEDs) exists but too slow: BEAST
  - investigating grid search with empirically calibrated isochrones
- The methods and implementation must be thoroughly tested
  - tests based on simulations and SDSS and Gaia datasets
- The products must be made available on RSP in timely fashion and user friendly
  - Bura at UniRi, perhaps UW cluster, cloud, elsewhere?

photo-D methodology: science motivation



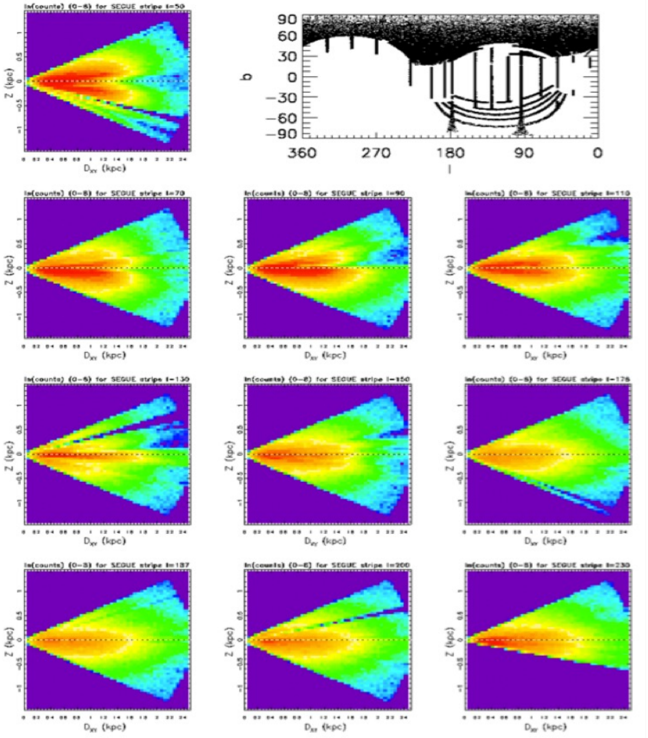
Ivezic+ (2008, ApJ, 684, 287)

- if we know stellar distances, we can study the Milky Way structure



Juric+ (2008, ApJ, 673, 864)

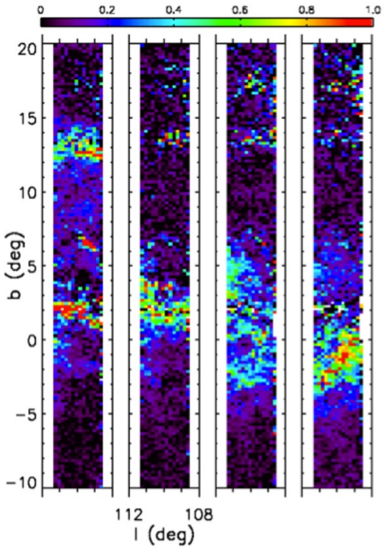
photo-D methodology: science motivation



Stellar counts

Berry+ (2012, ApJ, 757, 166)

- if we know stellar distances, we can study the Milky Way structure
- furthermore, at low galactic latitudes, we can map dust (and its properties), too



- left: differences in median A_r for D~ 1, 1.5, 2, 2.5 kpc
- dust at b ~ 2^o and b~13^o is confined to D~1-1.5 kpc
- dust at -3^o < b < 0^o is at D~2 kpc

Dust tomography

photo-D methodology: stellar astrophysics

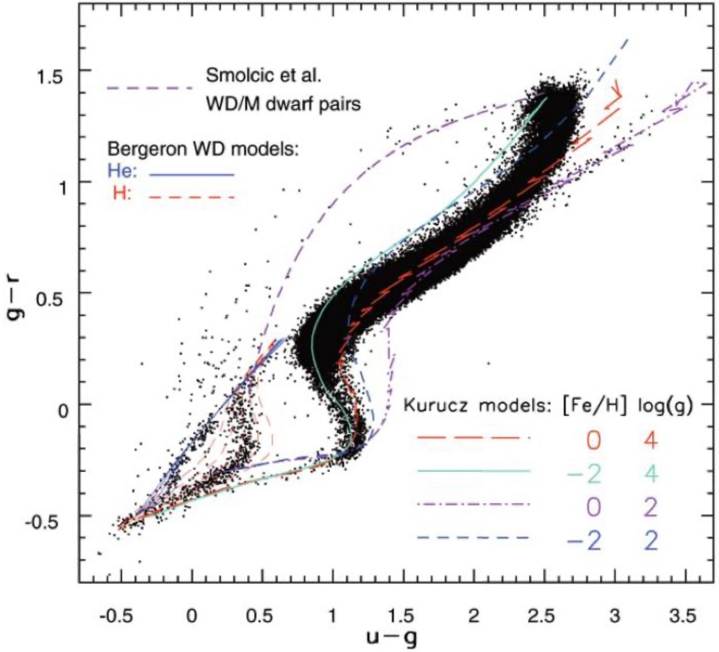


FIG. 23.—The $g - r$ vs. $u - g$ color-color diagrams for all nonvariable point sources constructed with the improved averaged photometry (*dots*). Various stellar models (Kurucz 1979; Bergeron et al. 1995; Smolčić et al. 2004) are shown by lines, as indicated in the figure. Berry+ (2012, ApJ, 757, 166)

- SDSS color-color diagram (corrected for dust):
  - stellar SEDs determined by: T_{eff}, [Fe/H], log(g)
  - different populations: MS, giants, WDs, binaries
  - **assuming** pop: from colors get best-fit SED
  - best-fit SED gives **L**uminosity constraint
  - pop **probability** can be gauged from goodness of fit, variability, priors, and other info (WISE?)
  - but there is dust:

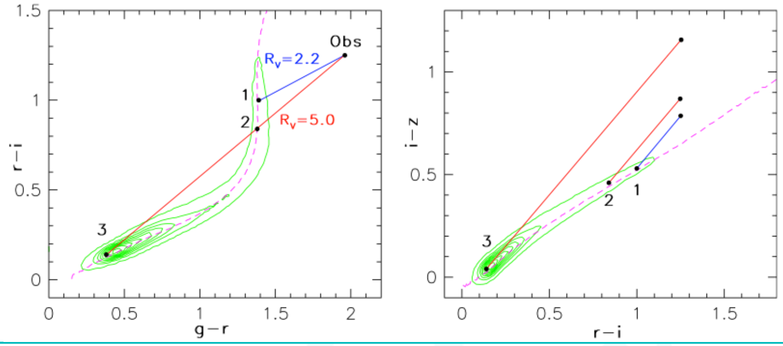


photo-D methodology: impact of dust

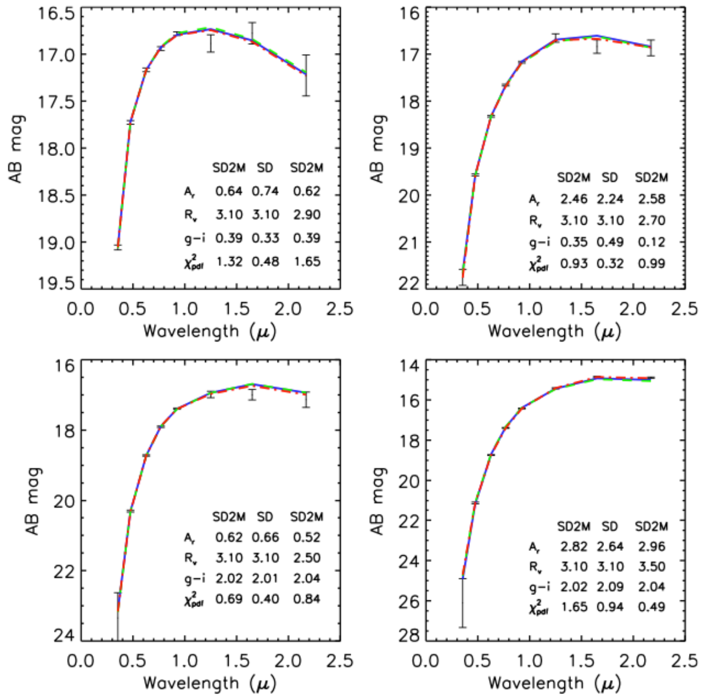


Figure 31. Comparison of three different types of best-fit SEDs: using only-SDSS data with fixed $R_V = 3.1$ (blue line) and using joint SDSS–2MASS data set with fixed R_V (green line) and with free R_V (red line). As demonstrated by the similarity of best-fit lines, the differences in best-fit parameters, listed in each panel, are due to degeneracies between intrinsic stellar color, amount of dust, and R_V. The shown cases correspond to blue and red stars (top row vs. bottom row), and small and large A_r (left column vs. right column).

- There are degeneracies with dust:
  - need to adopt an extinction curve
  - RV fixed or not?
- Two fitting philosophies:
  1) use stellar models to fit SEDs, or
  2) use high-latitude observations to fit dust-extincted low-latitude data
- The role of priors...
- Hierarchical Bayes (megaBEAST)?
- **We need a lot of testing!**

photo-D methodology: statistical treatment

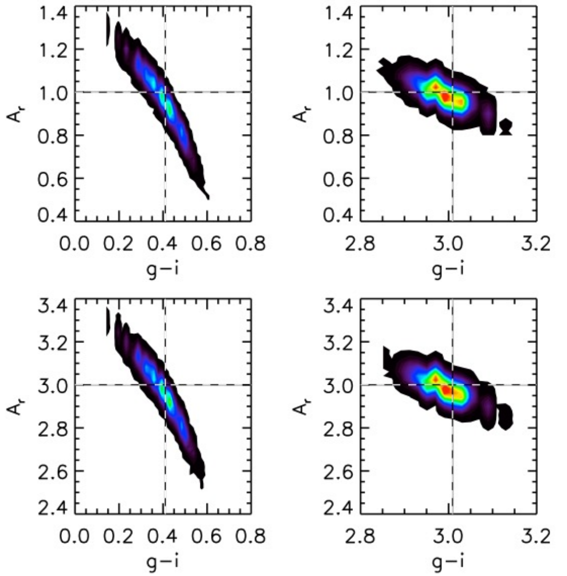


Figure 12. Analysis of the covariance in the best-fit values for A_r and $g-i$ using a simulated data set. The panels show the distributions of the best-fit values for A_r and $g-i$ for two different fiducial stars (left column: a blue star with true $g-i=0.4$; right column: a red star with true $g-i=3.0$), and two different extinction values (top panels: $A_r=1$; bottom panels: $A_r=3$). Photometric errors in the *ugriz* bands are generated using Gaussian distributions with $\sigma=0.02$ mag (uncorrelated between different bands). Note that the A_r vs. $g-i$ covariance is larger for the blue star, and does not strongly depend on assumed A_r.

- Berry+ (2012): Fit SEDs constructed from high-b observations and a dust model parametrized by R_V (shape vs. wavelength) and Av (how much dust)

- Compute best-fit via a brute force χ^2-minimization process:

$$\chi^2_{\mathrm{pdf}} = \frac{1}{N-k}\sum_{i=1}^{N}\left(\frac{c_i^{\mathrm{obs}} - c_i^{\mathrm{mod}}}{\sigma_i}\right)^2$$

- c_i and σ_i are N adjacent colors and errors (e.g., u - g, g - r, etc
- the number of fitting parameters is k = 2 (the position along the locus and A_r) for fixed-R_V (R_V=3.1), and k = 3 for free-R_V
- model colors are constructed by:

$$c^{\mathrm{mod}} = c^{\mathrm{lib}}(t) + [C_{\lambda 2}(R_V) - C_{\lambda 1}(R_V)]\ A_r$$

photo-D methodology: needed improvements

- Use of robust SED fitter? e.g. BEAST (https://beast.readthedocs.io/)
  - many pros and cons here... or **empirical isochrones?**
- Use of priors for stellar populations, as a function of (l,b) and magnitude
  - can rely on TRILEGAL models that are available from NOIRLab's DataLab, see https://arxiv.org/abs/2208.00829v1
- For a fixed (assumed population), use of priors for fitted parameters
  - distances must be consistent with Gaia results
- Automated quality assurance using Gaia data products (Bailer-Jones+ distances)
  - if we design plans well, ample opportunities to engage students
- Better code and documentation management

What we want to do is nicely described in **Green et al. 2014 (ApJ, 783:114)**

photo-D in Green et al. (2014)

THE ASTROPHYSICAL JOURNAL, 783:114 (16pp), 2014 March 10

- Left: **empirical isochrones**
- Middle: [Fe/H] prior as a function of distance from the Galactic plane (Z)
- Right: distance prior

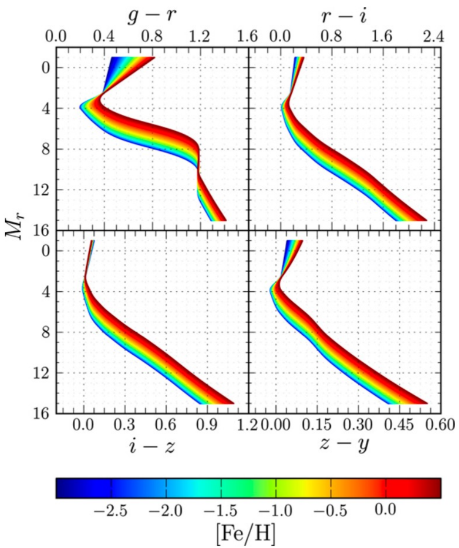


Figure 1. Model stellar colors as a function of absolute r magnitude and metallicity in Pan-STARRS 1 passbands. The stellar templates are based on PS1 color–color relations, and color is related to absolute magnitude and metallicity by SDSS observations of globular clusters (Ivezić et al. 2008a). Our empirical templates therefore assume an old stellar population. While the main sequence below the turnoff is nearly invariant with age, the giant branch and the location of the turnoff do, in reality, vary considerably with age. For this reason, we expect our inferences for main-sequence stars to be more accurate than those for giants. The narrowness of the kink at $M_r \simeq 2.4$ is an artifact of our models (see Section 4.1).

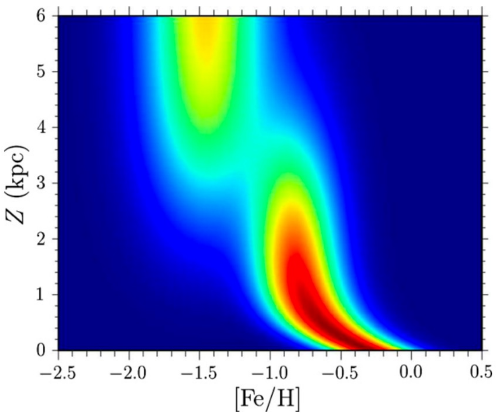


Figure 3. Metallicity prior, $p([\mathrm{Fe/H}] | Z)$, in the solar neighborhood ($R = 8$ kpc). High above the plane of the Galaxy, where the halo dominates, the metallicity distribution has a constant mean and variance. In the plane, where the disk dominates, the mean decreases with scale height. Adapted from Figure 9 of Ivezić et al. (2008a).

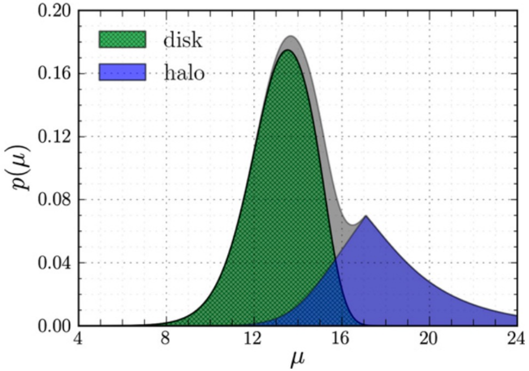


Figure 2. Distance prior for $(\ell, b) = (90°, 10°)$. The contributions of the disk and halo are shown individually in green and purple, respectively, while the total prior is given by the gray contour. The break in the contribution from the halo is due to the use of a broken power law for the number density of stars in this component.

photo-D in Green et al. (2014)

- Left: test of isochrones in PS1 color-color diagrams
- Below: test using PS1 data for globular clusters

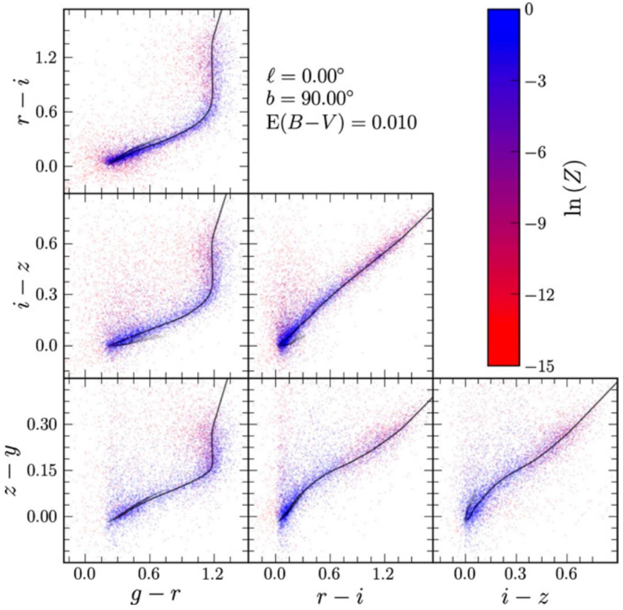


Figure 10. Comparison of PS1 stellar colors in the vicinity of the North Galactic Pole with our model colors. Each object is colored according to the evidence Z we compute. Objects represented by red dots have a low probability of being drawn from our stellar model and are rejected for the line-of-sight reddening determination. The solid black line traces our model stellar colors. Our main-sequence model colors do not depend on metallicity, while the model colors for the giant branch have a slight metallicity dependence.

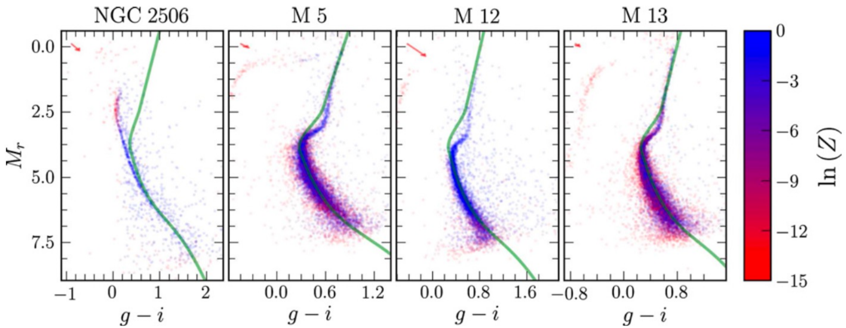
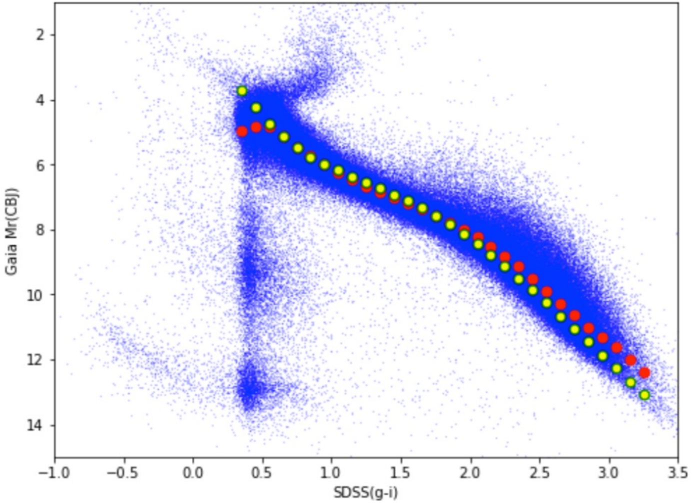


Figure 11. PS1 color–magnitude diagrams of three globular and one open cluster. For each cluster, the model isochrone with the catalog metallicity of the cluster is overplotted. The stellar photometry has been de-reddened and shifted by the catalog distance modulus to produce absolute magnitudes. The reddening vector is plotted in the top left corner of each panel in red for reference. Each star is colored by its evidence, with red stars unlikely to be drawn from our stellar model. In particular, stars which are blueward of the main-sequence turnoff, which are bluer than any star in our template library, have low evidence.

Comparison: Bailer-Jones+ distances from Gaia

- Method paper and also the latest results: https://arxiv.org/pdf/2012.05220.pdf
- Combines parallax measurement at the bright end and color-based distances
- Somewhat similar to what we want to do with LSST data (1 colors vs. 5 colors...)
- Their code is R, unclear if they plan to do anything with LSST data...



- left: Gaia-SDSS sample from SDSS Stripe 82
- blue: individual stars
- X axis: g-i color measured by SDSS
- Y axis: Mr from Bailer-Jones distances
- red symbols: median Mr in g-i bins
- yellow symbols: photo-D relation from SDSS Tomography papers (some mean [Fe/H])

Long-term plans, short-term plans...

- plans for the next 12 months (approximately to ComCam first light)
  - team responsibilities and (self)management
  - overall timelines for project “WBS”
  - collaboration & reporting tools
  - coarse plans for the following 12+12 months
    - Phase 1: now to ComCam first light (Sep ‘22 to Aug ‘23)
    - Phase 2: commissioning + early ops (Sep ‘23 to Aug ‘24)
    - Phase 3: early ops and towards DR1 (Sep ‘24 to Aug ‘25)

- first papers to write?
  - reproducing Gaia distances with SDSS Stripe 82 photometry and BEAST etc.
  - reproducing distances in TRILEGAL mock catalogs with chosen method