Studying stellar populations with machine learning

G. Baume[1,2], M.J. Rodriguez[2], C. Feinstein[1,2] & E. Gularte[1,3]

1. Facultad de Ciencias Astronómicas y Geofísicas de La Plata (UNLP)
2. Instituto de Astrofísica de La Plata (Conicet - UNLP)
3. Geodesia Espacial y Aeronomía (FCAG - UNLP)

Contact: gbaume@fcaglp.unlp.edu.ar

Abstract

Stellar populations in NGC 1313 galaxy were studied using multi-band photometric data obtained with the Hubble Space Telescope. Machine learning techniques were applied to recognize both the stellar populations and the groups of stars in the youngest population. In both cases, different clustering algorithms were used and their efficiency were evaluated. Additionally, we characterized the spatial distribution of each population. It was possible to identify the youngest populations with a hierarchical structure and the most evolved ones with a homogeneous distribution, except for very large scale fluctuations.

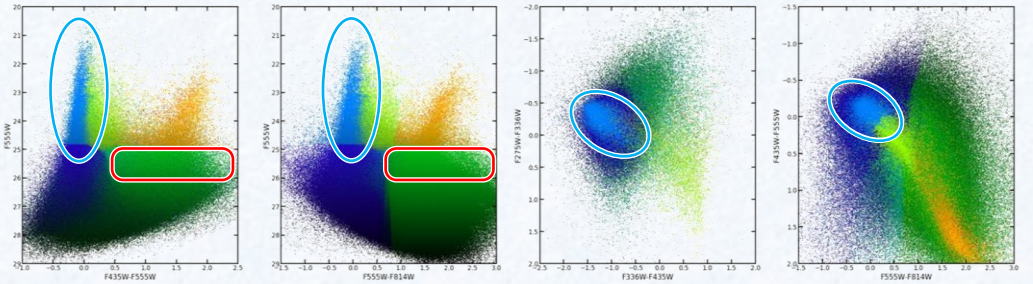

Figure 2: Photometric diagrams (CMDs and TCDs) of NGC 1313. Colors indicate different stellar groups. Bluest-bright population and red-faint one are indicated inside blue and red ellipses, respectivelly.

Introduction

- Galaxies are made up of diverse stellar populations
- These populations allow to know the star formation history (SFH) in the galaxy
- The Hubble Space Telescope (HST) has generated an large database including multiband photometry (Dalcanton et al. 2009; Lee et al. 2014)
- HST high spatial resolution (see Fig. 1) allows to detect stellar components in nearby galaxies and their photometric diagrams are an important tool to identify different stellar populations, but their separation is a complex task.
- There are several methods from unsupervised machine learning that allow to group data based on common characteristics ("clustering")

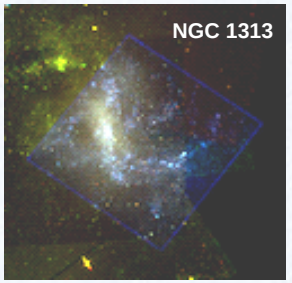

Figure 1: Color image of NGC 1313 galaxy. Based on data from the ***"HST Legacy Extragalactic UV Survey"***

Identification of stellar populations

- Color-magnitude diagrams (CMDs) and color-color diagrams (TCDs) of NGC 1313 galaxy were obtained from multiband HST observations
- We selected data with successive ranges in F555W (0.25 mag wide) and different stellar components were identified in each case using GMM over the corresponding TCDs (see Fig. 2)
- Different gaussian components were grouped. In this procedure, “AgC" method was applied over the corresponding gaussian centers in a CMDs-TCDs space of features. They also were tell apart using F555W = 25 as a split value. Therefore, seven different stellar populations were identified.

Clustering methods

- **GMM:** *Gaussian mixture method*
- **AgC:** *Agglomerative clustering*
- **HDBSCAN:** *Hierarchical Density based spatial clustering of applications with noise*
- **PLC:** *Path linkage criterion*

See details in Battinelli et al. 2000, Pedregosa et al. 2011 and McInnes et al. 2017

Bands: *F275W*, *F336W*, *F435W*, *F555W*, *F814W*

Studying stellar populations with machine learning

G. Baume[1,2], M.J. Rodriguez[2], C. Feinstein[1,2] & E. Gularte[1,3]

1. Facultad de Ciencias Astronómicas y Geofísicas de La Plata (UNLP)
2. Instituto de Astrofísica de La Plata (Conicet - UNLP)
3. Geodesia Espacial y Aeronomía (FCAG - UNLP)

Contact: gbaume@fcaglp.unlp.edu.ar

Fractal analysis

- Stellar density map of each population was built using KDE method. Therefore, different nested structures were identified using dendrograms (see Fig. 3a).
- Perimeter-area method (Fig 3b; Mandelbrot 1982; Rodirguez et al. 2019), was applied at stellar density maps of each stellar population.
- A dimension value compatible with a fractal structure was obtained for the youngest population (blue ellipses in Fig. 2). However, it was not possible to obtain a reliable value for the most evolved population (red ellipses in Fig. 2)

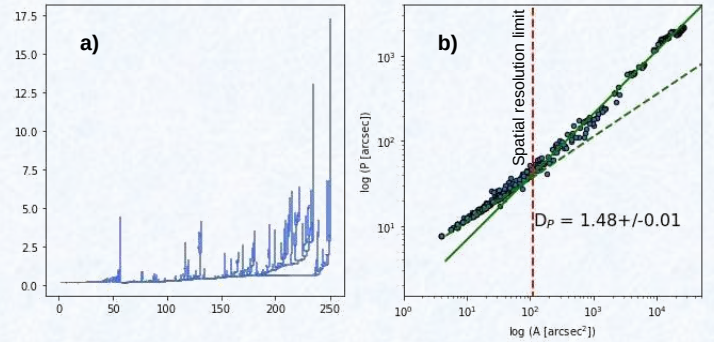

Figure 3: Youngest stellar population. a) Dendrogram of stellar density structures. b) Perimeter-area method to estimate fractal dimension.

Preliminary conclusions

- Stellar populations of NGC 1313 galaxy were discriminated based on the photometric diagrams of high spatial resolution of the HST
- Using HDBSCAN and PLC, the stellar groups associated with the youngest population were identified and parametrized.
- The youngest population appears to have a structure compatible with a fractal behavior, while the red-faint population is homogeneous except for very large-scale fluctuations.

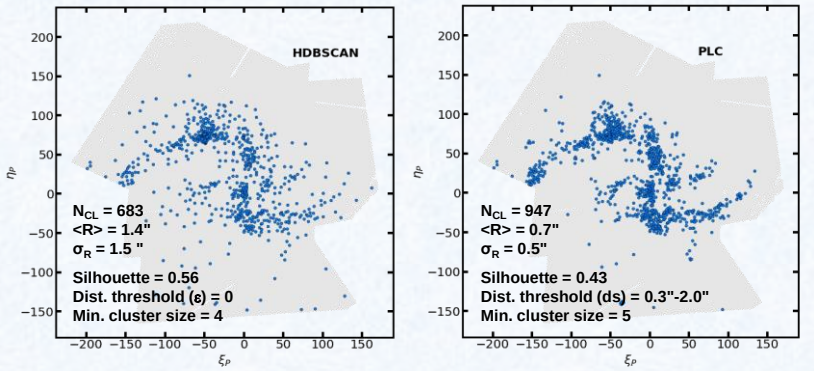

Figure 4: Stellar clusters identified in the youngest stellar population using HDSCAN (left) and PLC (right) and their main parameters.

Young stellar clusters

HDBSCAN and PLC clustering methods were applied over the spatial distribution of the youngest population (blue ellipses in Fig. 2). The obtained cluster centers and their most relevant parameters are presented in Fig 4. In general, we found that:

- Both methods identified stellar clusters linked to stellar over-densities
- HDBSCAN provided stellar groups with a wider range of sizes and stellar densities than PLC
- Both methods revealed Silhouette indices (Rousseeuw 1987) with acceptable values (> 0)

References

- Battinelli, P., Capuzzo-Dolcetta, R., Hodge, P. W., Vicari, A., & Wyder, T. K. 2000, A&A, 357, 437
- Dalcanton J.J., Williams B.F., Seth A.C., Dolphin, A., Holtzman J. et al. 2009 ApJS 183, 67
- Lee J.C., Calzetti D., Adamo A. Aloisi A. Andrews J.~E et al. 2014 AAS 22321701
- Mandelbrot B.B., 1982, The fractal geometry of nature, vol. 1, WH freeman New York
- McInnes L., Healy J. & Astels S. 2017 Journal of Open Source Software, The Open Journal, 2, 11
- Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B. et al 2011 JMLR 12, 2825
- Rodríguez M.J., Baume G., Feinstein C., 2019, A&A, 626,A35
- Rousseeuw P.. 1987 Journal of Computational and Applied Mathematics. 20. 53.